报文格式大全(V1.0)

目录

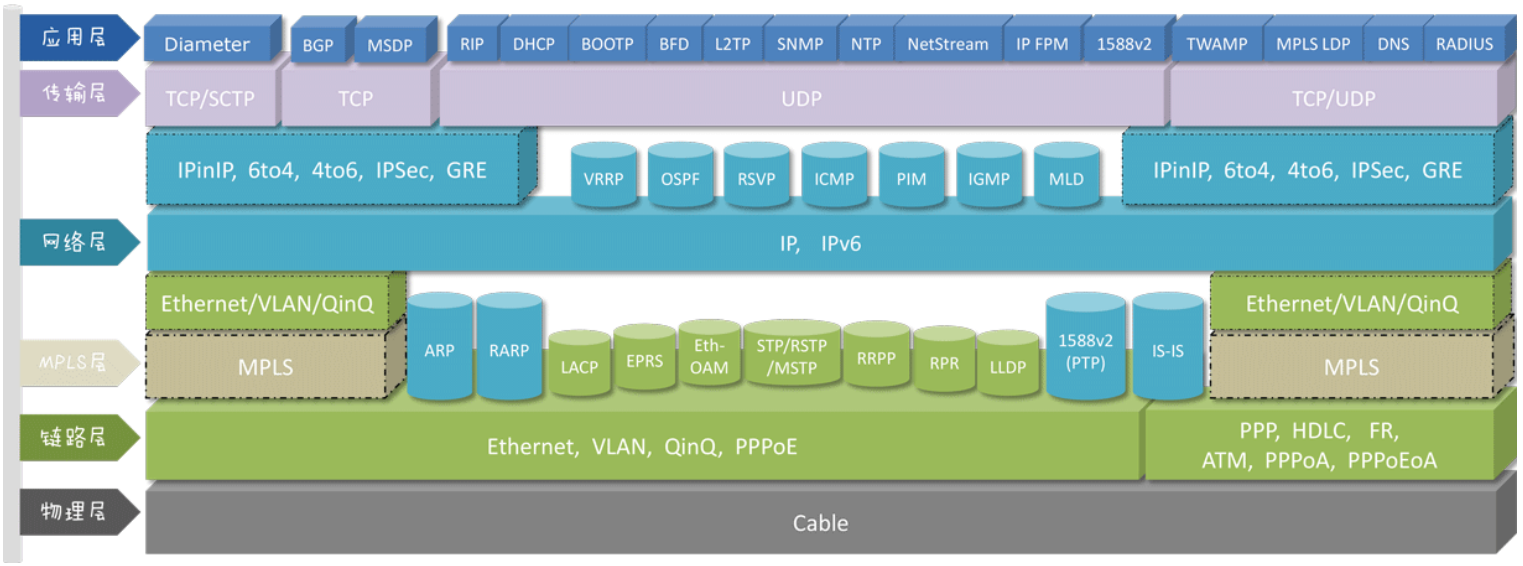
1. 协议地图 5
2. 链路层 5
2.1 以太帧格式 6
2.1.1 Ethernet Ⅱ 以太帧 6
2.1.2 Netware 以太帧格式 9
2.1.3 802.3 SAP 以太帧 10
2.1.4 802.3 LLC SNAP 以太帧格式 12
2.2 VLAN 帧格式 14
2.3 QinQ 帧格式 15
2.4 PPP 帧格式 17
2.5 PPPoE 报文格式 21
2.6 HDLC 帧格式 27
2.7 ATM 信元格式 29
2.7.1 ATM 信元格式 29
2.7.2 AAL5 多协议封装通用格式 31
2.7.3 IPoEoA/PPPoEoA 格式 33
2.7.4 PPPoA 格式 33
2.7.5 ATM PWE3 封装格式 34
2.8 STP/RSTP/MSTP 帧格式 34
2.9 RPR 帧格式 41
2.10 RRPP 帧封装格式 44
2.11 LACP 报文格式 47
2.12 以太 OAM 报文格式 53
2.13 ERPS 帧格式 59
2.14 LLDP 报文格式 62
2.15 IS-IS 报文格式 67
2.15.1 IS-IS 报文通用格式 67
2.15.2 IS-IS Hello 消息格式 71
2.15.3 IS-IS LSP 消息格式 73
2.15.4 IS-IS SNP 消息格式 75
3. MPLS 层 77
3.1 MPLS 报文格式 77
3.2 MPLS L2/L3VPN 报文结构 79
3.3 MPLS Ping/Tracert (MPLS Echo)报文格式 84
4. 网络层 92
4.1 ARP/RARP 报文格式 92
4.2 GRE 报文格式 96
4.3 ICMP 报文格式 98
4.3.1 ICMP 报文通用格式 98
4.3.2 ICMP Echo Request/Reply 消息格式 101
4.3.3 ICMP 目的不可达消息格式 103
4.3.4 ICMP 重定向消息格式 105
4.3.5 ICMP 超时消息格式 106
4.3.6 ICMP 参数问题消息格式 107
4.3.7 ICMP 源端被关闭消息格式 108
4.4 ICMPv6 报文格式 109
4.4.1 ICMPv6 报文通用格式 109

4.4.2 ICMPv6 回显请求/应答消息 112
4.4.3 ICMPv6 目的不可达消息 113
4.4.4 ICMPv6 重定向消息 114
4.4.5 ICMPv6 超时消息 116
4.4.6 ICMPv6 参数错误消息 117
4.4.7 ICMPv6 路由器请求（Router Solicitation）消息 118
4.4.8 ICMPv6 路由器通告消息 119
4.4.9 ICMPv6 邻居请求（Neighbor Solicitation）消息 123
4.4.10 ICMPv6 邻居通告消息 125
4.5 IGMP 报文格式 127
4.5.1 IGMPv1 报文格式 127
4.5.2 IGMPv2 报文格式 128
4.5.3 IGMPv3 报文格式 130
4.6 IP in IP 报文格式 136
4.7 IP 报文格式 138
4.8 IPv6 报文格式 144
4.9 IPv6 in IP (6to4)报文格式 146
4.10 MLD 报文格式 147
4.11 OSPF 报文格式 154
4.11.1 OSPF 报文头格式 154
4.11.2 OSPF Hello 报文格式 155
4.11.3 OSPF DD 报文格式 157
4.11.4 OSPF LSR 报文格式 159
4.11.5 OSPF LSU 报文格式 161
4.11.6 OSPF LSAck 报文格式 168
4.12 OSPFv3 报文格式 169
4.12.1 OSPFv3 报文头格式 170
4.12.2 OSPFv3 Hello 报文格式 171
4.12.3 OSPFv3 DD 报文格式 172
4.12.4 OSPFv3 LSR 报文格式 174
4.12.5 OSPFv3 LSU 报文格式 174
4.12.6 OSPFv3 LSAck 报文格式 187
4.13 PIM 报文格式 187
4.13.1 PIM 报文通用格式 188
4.13.2 PIM Hello 消息格式 189
4.13.3 PIM Register 消息格式 191
4.13.4 PIM Register-Stop 消息格式 193
4.13.5 PIM Join/Prune 消息格式 195
4.13.6 PIM Graft/Graft-Ack 消息格式 198
4.13.7 PIM Bootstrap 消息格式 201
4.13.8 PIM Assert 消息格式 203
4.13.9 PIM C-RP Advertisement 消息格式 205
4.14 RSVP 报文格式 208
4.15 VRRP 报文格式 217
5. 传输层 220
5.1 TCP 报文格式 221
5.2 UDP 报文格式 229
5.3 SCTP 报文格式 230
5.3.1 SCTP 通用报文格式 231
5.3.2 SCTP ABORT 报文格式 234
5.3.3 SCTP COOKIE ACK 格式 235
5.3.4 SCTP COOKIE ECHO 数据块格式 236

5.3.5 SCTP DATA 数据块格式 237
5.3.6 SCTP ERROR 数据块格式 239
5.3.7 SCTP HEARTBEAT 数据块格式 243
5.3.8 SCTP HEARTBEAT ACK 数据块格式 244
5.3.9 SCTP INIT 数据块格式 245
5.3.10 SCTP INIT ACK 数据块格式 248
5.3.11 SCTP SACK 数据块格式 250
5.3.12 SCTP SHUTDOWN 消息格式 253
5.3.13 SCTP SHUTDOWN ACK 数据块格式 253
5.3.14 SCTP SHUTDOWN COMPLETE 数据块格式 253
6. 应用层 253
6.1 1588v2（PTP）报文格式 254
6.1.1 1588v2（PTP）报文通用格式 255
6.1.2 1588v2 Sync 消息和 Delay_Req 消息 258
6.1.3 1588v2 Follow_Up 消息 262
6.1.4 1588v2 Delay_Resp 消息 265
6.1.5 1588v2 Pdelay_Req 消息 269
6.1.6 1588v2 Pdelay_Resp 消息 273
6.1.7 1588v2 Pdelay_Resp_Follow_Up 消息 276
6.1.8 1588v2 Signaling 消息 280
6.1.9 1588v2 Management 消息 283
6.2 BFD 控制报文格式 286
6.3 BGP 报文格式 292
6.3.1 BGP 报文头基本格式（RFC4271） 292
6.3.2 BGP OPEN 报文格式 293
6.3.3 BGP UPDATE 报文格式 297
6.3.4 BGP 的 NOTIFICATION 报文格式 302
6.3.5 BGP KEEPALIVE 报文格式 307
6.3.6 BGP 的 REFRESH 报文格式 308
6.4 BOOTP 报文格式 309
6.5 DHCP 报文格式 312
6.6 DHCPv6 报文格式 318
6.7 Diameter 协议报文格式 324
6.8 DNS 报文格式 328
6.9 IP FPM 报文格式 333
6.10 IPSec 报文格式 338
6.11 L2TP 报文格式 344
6.12 MPLS LDP 报文格式 347
6.13 MSDP 报文格式 361
6.14 NetStream 报文格式 364
6.15 RIP 报文格式 372
6.16 RIPng 的报文格式 376
6.17 NTP 报文格式 378
6.18 RADIUS 报文格式 382
6.19 SNMP 报文格式 390
6.19.1 SNMPv1 Packet FormatSNMPv1 报文格式 390
6.19.2 SNMPv2c Packet FormatSNMPv2c 报文格式 393
6.19.3 SNMPv3 报文格式 396
6.20 TWAMP 报文格式 399

1. 协议地图

单击下图的方块或圆柱，可进入对应协议的报文格式介绍及抓包图。



2. 链路层

- 以太帧格式
- VLAN 帧格式
- QinQ 帧格式
- PPP 帧格式
- PPPoE 报文格式
- HDLC 帧格式
- ATM 信元格式
- STP/RSTP/MSTP 帧格式
- RPR 帧格式
- RRPP 帧封装格式
- LACP 报文格式
- 以太 OAM 报文格式
- ERPS 帧格式
- LLDP 报文格式
- IS-IS 报文格式

2.1 以太帧格式

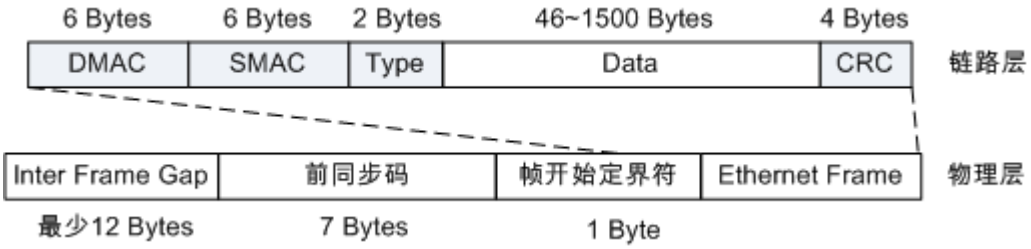
- Ethernet Ⅱ 以太帧
- Netware 以太帧格式
- 802.3 SAP 以太帧
- 802.3 LLC SNAP 以太帧格式

父主题： 链路层

2.1.1 Ethernet Ⅱ 以太帧

帧格式

图 1 Ethernet Ⅱ帧格式



字段	长度	含义
DMAC	6 字节	目的 MAC 地址，IPV4 为 6 字节，该字段确定帧的接收者。
SMAC	6 字节	源 MAC 地址，IPV4 为 6 字节，该字段标识发送帧的工作站。
Type	2 字节	协议类型。下表列出了链路直接封装的协议。
Data	变长	数据字段的最小长度必须为 46 字节以保证帧长至少为 64 字节，这意味着传输一字节信息也必须使用 46 字节的数据字段。 如果填入该字段的信息少于 46 字节，该字段的其余部分也必须进行填充。数据字段的最大长度为 1500 字节。
CRC	4 字节	用于帧内后续字节差错的循环冗余检验（也称为 FCS 或帧检验序列）。

表 1 各 Type 值对应的协议

值	协议

字段	长度	含义
	0x0800	Internet Protocol (IP) [RFC894]
	0x0801	X.75 Internet
	0x0805	X.25 Level 3
	0x0806	Address Resolution Protocol (ARP)[RFC7042]
	0x0808	Frame Relay ARP [RFC1701]
	0x8000	IS-IS
	0x8035	Reverse Address Resolution Protocol (RARP) [RFC903]
	0x8137	Novell NetWare IPX/SPX (old)
	0x8138	Novell, Inc.
	0x8100	IEEE Std 802.1Q - Customer VLAN Tag Type
	0x814C	SNMP over Ethernet [RFC1089]
	0x86DD	IP Protocol version 6 (IPv6) [RFC7042]
	0x8808	IEEE Std 802.3 - Ethernet Passive Optical Network (EPON) [RFC7042]
	0x880B	Point-to-Point Protocol (PPP) [RFC7042]
	0x880C	General Switch Management Protocol (GSMP)
	0x8847	MPLS (multiprotocol label switching) label stack - unicast [RFC 3032]
	0x8848	MPLS (multiprotocol label switching) label stack - multicast [RFC 3032]
	0x8863	PPP over Ethernet (PPPoE) Discovery Stage [RFC2516]

字段	长度	含义
	0x8864	PPP over Ethernet (PPPoE) Session Stage [RFC2516]
	0x888E	IEEE Std 802.1X - Port-based network access control
	0x88A8	IEEE Std 802.1Q - Service VLAN tag identifier (S-Tag)
	0x88B7	IEEE Std 802 - OUI Extended Ethertype
	0x88C7	IEEE Std 802.11 - Pre-Authentication (802.11i)
	0x88CC	IEEE Std 802.1AB - Link Layer Discovery Protocol (LLDP)
	0x88E5	IEEE Std 802.1AE - Media Access Control Security
	0x88F5	IEEE Std 802.1Q - Multiple VLAN Registration Protocol (MVRP)
	0x88F6	IEEE Std 802.1Q - Multiple Multicast Registration Protocol (MMRP)

帧示例

```
⊞ Frame 1: 122 bytes on wire (976 bits), 122 bytes captured (976 bits)
⊟ Ethernet II, Src: HuaweiTe_ca:dd:a3 (00:25:9e:ca:dd:a3), Dst: HuaweiTe_03:33:00 (00:e
  ⊟ Destination: HuaweiTe_03:33:00 (00:e0:fc:03:33:00)
      Address: HuaweiTe_03:33:00 (00:e0:fc:03:33:00)
      .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
      .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
  ⊟ Source: HuaweiTe_ca:dd:a3 (00:25:9e:ca:dd:a3)
      Address: HuaweiTe_ca:dd:a3 (00:25:9e:ca:dd:a3)
      .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
      .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
    Type: IP (0x0800)
    Trailer: 2930bbaf
⊞ Internet Protocol Version 4, Src: 129.1.47.124 (129.1.47.124), Dst: 129.1.46.15 (129.
⊞ Transmission Control Protocol, Src Port: diameter (3868), Dst Port: 65382 (65382), Se
⊞ Diameter Protocol
```

```
0000  00 e0 fc 03 33 00 00 25  9e ca dd a3 08 00 45 00   ....3..% ......E.
0010  00 68 80 ee 00 00 40 06  9a 14 81 01 2f 7c 81 01   .h....@. ..../|..
0020  2e 0f 0f 1c ff 66 aa f0  71 66 c2 28 09 97 50 18   .....f.. qf.(..P.
0030  7f ff 2e 6f 00 00 01 00  00 40 80 00 01 18 00 00   ...o.... .@......
0040  00 00 00 0c 50 b7 00 0c  50 b7 00 00 01 08 40 00   ....P... P.....@.
0050  00 17 70 63 72 66 2e 68  75 61 77 65 69 2e 63 6f   ..pcrf.h uawei.co
0060  6d 00 00 00 01 28 40 00  00 12 68 75 61 77 65 69   m....(@. ..huawei
0070  2e 63 6f 6d 00 00 29 30  bb af                     .com..)0 ..
```

参考标准

标准	描述
IEEE 802.3	Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications

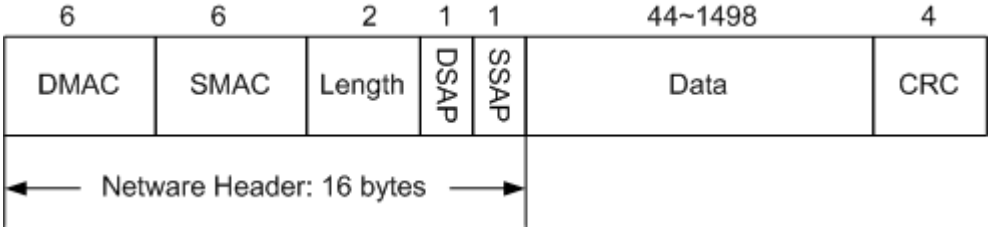
标准	描述
IEEE 802.3ae	Media Access Control (MAC) Parameters, Physical Layers, and Management parameters for 10Gb/s Operation
RFC 894	A Standard for the Transmission of IP Datagrams over Ethernet Networks
RFC 1042	A Standard for the Transmission of IP Datagrams over IEEE 802 Networks

2.1.2 Netware 以太帧格式

帧格式

Netware-以太网帧对 IEEE802.3 的数据字段进行了专门分隔以便传输 NetWare 类型的数据。

图 1 Netware 帧格式



字段	长度（字节）	含义
DMAC	6	目的 MAC 地址
SMAC	6	源 MAC 地址
Length	2	指后续数据的字节长度，但不包括 CRC 检验码
DSAP	1	目的服务访问点
SSAP	1	源服务访问点
Data	44~1498	负载

字段	长度（字节）	含义
CRC	4	用于帧内后续字节差错的循环冗余检验（也称为 FCS 或帧检验序列）

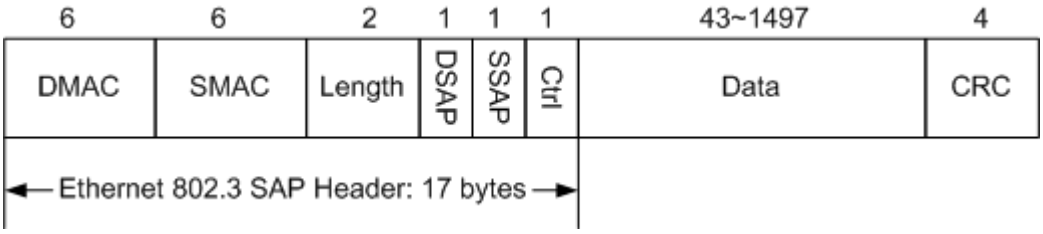
参考标准

标准	描述
IEEE 802.3	Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications
IEEE 802.3ae	Media Access Control (MAC) Parameters, Physical Layers, and Management parameters for 10Gb/s Operation
RFC 894	A Standard for the Transmission of IP Datagrams over Ethernet Networks
RFC 1042	A Standard for the Transmission of IP Datagrams over IEEE 802 Networks

2.1.3 802.3 SAP 以太帧

帧格式

图 1 802.3 SAP 以太帧格式

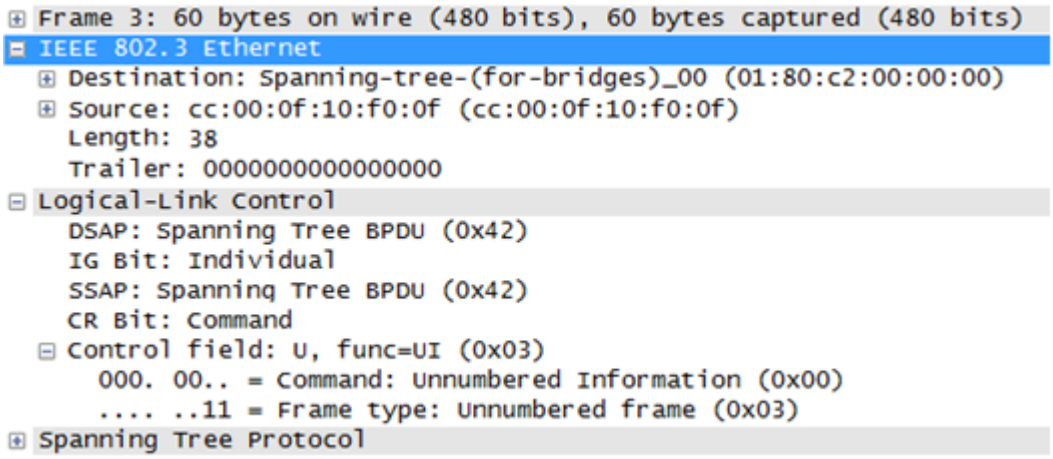


字段	长度（字节）	含义
DMAC	6	目的 MAC 地址。
SMAC	6	源 MAC 地址。
Length	2	指后续数据的字节长度，但不包括 CRC 检验码。

字段	长度（字节）	含义
DSAP	1	目的服务访问点，若后面类型为 IP 帧值设为 0x06。
SSAP	1	源服务访问点，若后面类型为 IP 帧值设为 0x06。
Ctrl	1	该字段值通常设为 0x03，表示无连接服务的 IEEE 802.2 无编号数据格式。
Data	44~1498	负载。
CRC	4	用于帧内后续字节差错的循环冗余检验（也称为 FCS 或帧检验序列）。

帧示例

图 2 802.3 LLC NAP 以太帧



参考标准

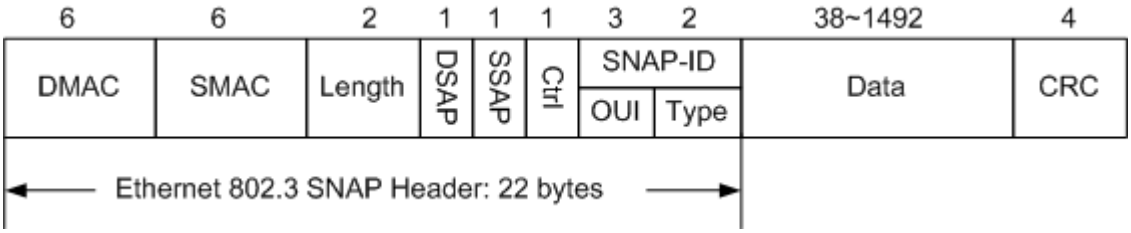
标准	描述
IEEE 802.3	Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications
IEEE 802.3ae	Media Access Control (MAC) Parameters, Physical Layers, and Management parameters for 10Gb/s Operation
RFC 894	A Standard for the Transmission of IP Datagrams over Ethernet Networks

标准	描述
RFC 1042	A Standard for the Transmission of IP Datagrams over IEEE 802 Networks

2.1.4 802.3 LLC SNAP 以太帧格式

帧格式

图 1 SNAP 以太帧格式

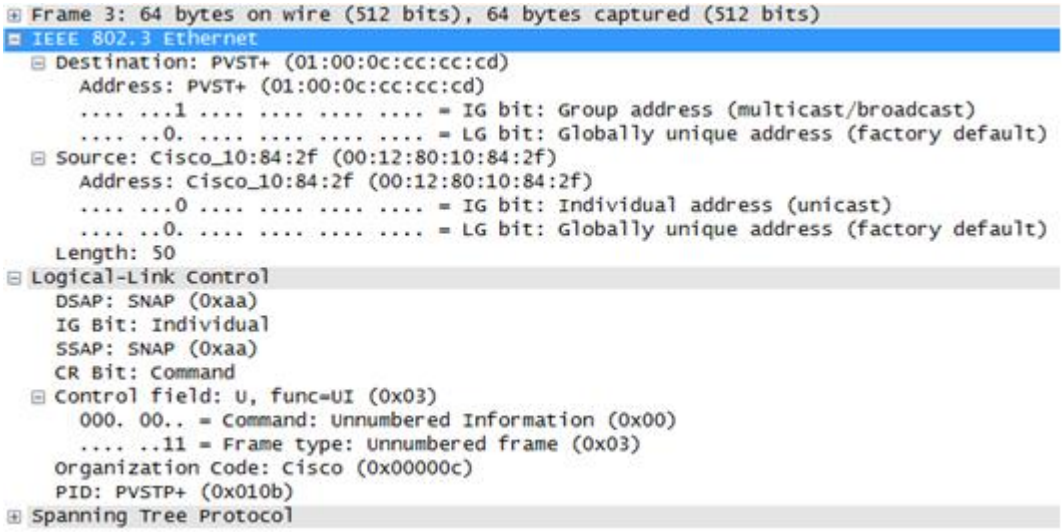


字段	长度（字节）	含义
DMAC	6	目的 MAC 地址。
SMAC	6	源 MAC 地址。
Length	2	指后续数据的字节长度，但不包括 CRC 检验码。
DSAP	1	目的服务访问点，若后面类型为 IP 帧值设为 0x06。
SSAP	1	源服务访问点，若后面类型为 IP 帧值设为 0x06。
Ctrl	1	该字段值通常设为 0x03，表示无连接服务的 IEEE 802.2 无编号数据格式。
SNAP-ID	5	由 OUI 和 Type 两部分组成。
OUI	3	3 字节的组织唯一标识符（Organizationally Unique Identifier），其值通常等于 MAC 地址的前 3 字节，即网络适配器厂商代码。

字段	长度（字节）	含义
Type	2	标识以太网帧所携带的上层数据类型。
Data	44~1498	负载。
CRC	4	用于帧内后续字节差错的循环冗余检验（也称为 FCS 或帧检验序列）。

帧示例

图 2 802.3 LLC SNAP 以太帧



参考标准

标准	描述
IEEE 802.3	Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications
IEEE 802.3ae	Media Access Control (MAC) Parameters, Physical Layers, and Management parameters for 10Gb/s Operation
RFC 894	A Standard for the Transmission of IP Datagrams over Ethernet Networks
RFC 1042	A Standard for the Transmission of IP Datagrams over IEEE 802 Networks

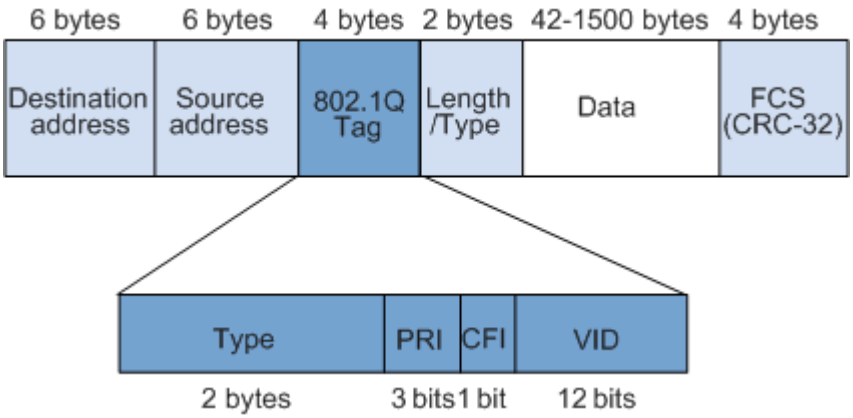
2.2 VLAN 帧格式

帧格式

IEEE 802.1Q 标准对 Ethernet 帧格式进行了修改，在源 MAC 地址字段和协议类型字段之间加入 4 字节的 802.1Q Tag。

VLAN 帧最小帧长为 64 字节。

图 1 VLAN 帧格式

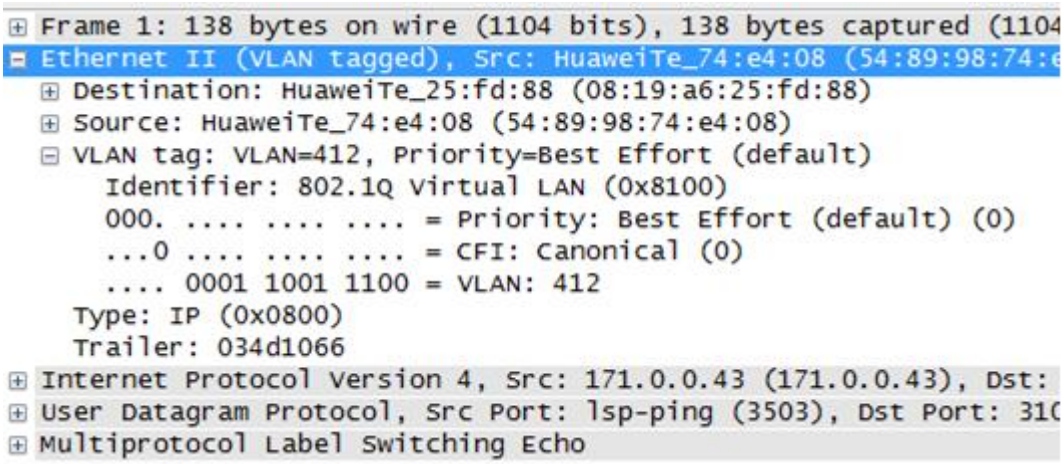


字段	长度	含义
Destination address	6 字节	目的 MAC 地址。
Source address	6 字节	源 MAC 地址。
Type	2 字节	长度为 2 字节，表示帧类型。取值为 0x8100 时表示 802.1Q Tag 帧。如果不支持 802.1Q 的设备收到这样的帧，会将其丢弃。
PRI	3 比特	Priority，长度为 3 比特，表示帧的优先级，取值范围为 0～7，值越大优先级越高。用于当阻塞时，优先发送优先级高的数据包。 如果设置用户优先级，但是没有 VLANID，则 VLANID 必须设置为 0x000。
CFI	1 比特	CFI (Canonical Format Indicator)，长度为 1 比特，表示 MAC 地址是否是经典格式。CFI 为 0 说明是标准格式，CFI 为 1 表示为非标准格式。用于区分以太网帧、FDDI（Fiber Distributed Digital Interface）帧和令牌环网帧。在以太网中，CFI 的值为 0。
VID	12 比特	LAN ID，长度为 12 比特，表示该帧所属的 VLAN。在 VRP 中，可配置的 VLAN ID 取值范围为 1～4094。0 和 4095 协议中规定为保留的 VLAN ID。 三种类型：

字段	长度	含义
		• Untagged 帧：VID 不计 • Priority-tagged 帧：VID 为 0x000 • VLAN-tagged 帧：VID 范围 0～4095 三个特殊的 VID： • 0x000：设置优先级但无 VID • 0x001：缺省 VID • 0xFFF：预留 VID
Length/Type	2 字节	指后续数据的字节长度，但不包括 CRC 检验码。
Data	42~1500 字节	负载（可能包含填充位）。
CRC	4 字节	用于帧内后续字节差错的循环冗余检验（也称为 FCS 或帧检验序列）。

帧示例

图 2 VLAN 帧



参考标准

标准	描述
RFC 3069	VLAN Aggregation for Efficient IP Address Allocation
IEEE 802.1Q	IEEE Standards for Local and Metropolitan Area Networks : Virtual Bridged Local Area Networks

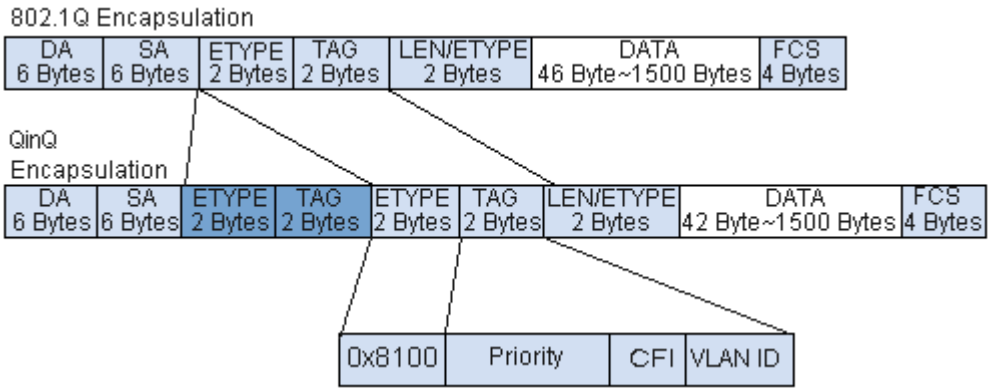
2.3 QinQ 帧格式

QinQ 报文有固定的格式，就是在 802.1Q 的标签之上再打一层 802.1Q 标签，QinQ 报文比 802.1Q 报文多四个字节。

VLAN 帧最小帧长为 68 字节。

帧格式

图 1 QinQ 帧格式

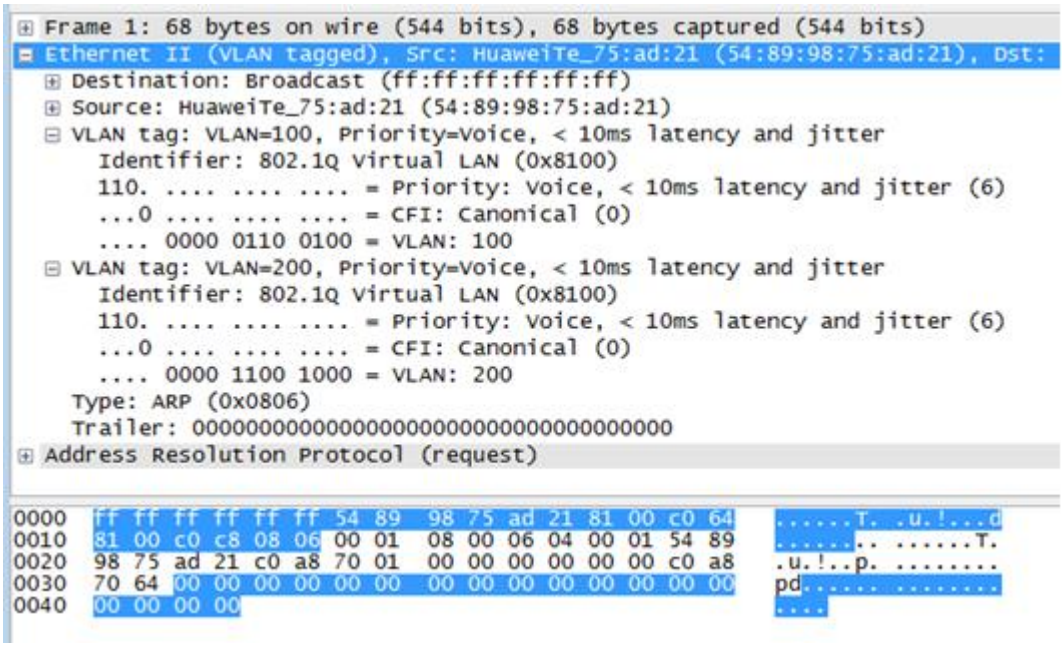


字段	长度	含义
Destination address	6 字节	目的 MAC 地址。
Source address	6 字节	源 MAC 地址。
Type	2 字节	长度为 2 字节，表示帧类型。取值为 0x8100 时表示 802.1Q Tag 帧。如果不支持 802.1Q 的设备收到这样的帧，会将其丢弃。 对于内层 VLAN tag，该值设置为 0x8100；对于外层 VLAN tag，有下列几种类型 • 0x8100：思科路由器使用 • 0x88A8：Extreme Networks switches 使用 • 0x9100：Juniper 路由器使用 • 0x9200：Several 路由器使用
PRI	3 比特	Priority，长度为 3 比特，表示帧的优先级，取值范围为 0～7，值越大优先级越高。用于当交换机阻塞时，优先发送优先级高的数据包。
CFI	1 比特	CFI (Canonical Format Indicator)，长度为 1 比特，表示 MAC 地址是否是经典格式。CFI 为 0 说明是经典格式，CFI 为 1 表示为非经典格式。用于区分以太网帧、FDDI（Fiber Distributed Digital Interface）帧和令牌环网帧。在以太网中，CFI 的值为 0。
VID	12 比特	LAN ID，长度为 12 比特，表示该帧所属的 VLAN。在 VRP 中，可配置的 VLAN ID 取值范围为 1～4094。
Length/Type	2 字节	指后续数据的字节长度，但不包括 CRC 检验码。

字段	长度	含义
Data	42~1500 字节	负载（可能包含填充位）。
CRC	4 字节	用于帧内后续字节差错的循环冗余检验（也称为 FCS 或帧检验序列）。

帧示例

图 2 QinQ 帧



参考标准

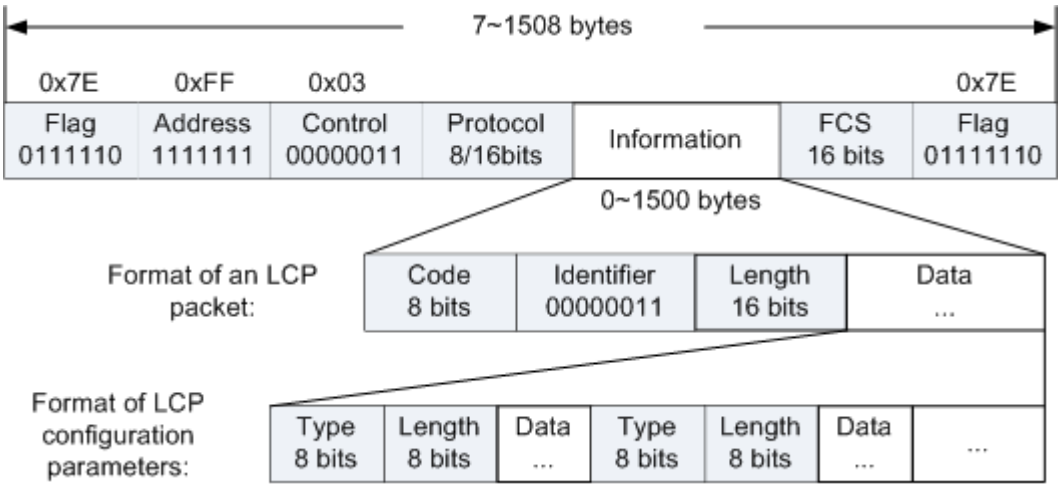
标准	描述
RFC 3069	VLAN Aggregation for Efficient IP Address Allocation
IEEE 802.1q	IEEE Standards for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks
IEEE 802.1ad	IEEE Standards for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks-Amendment 4

2.4 PPP 帧格式

帧格式

PPP 帧的内容是指 Address、Control、Protocol 和 Information 四个域的内容。各字段的含义如下。

图 1 PPP 帧格式



字段	长度	含义
Flag	1 字节	Flag 域标识了一个物理帧的起始和结束，该字节为 0x7E。
Address	1 字节	PPP 协议室被运用在点对点的链路上，它可唯一标识对方，因此无须知道对方数据链路层地址。所以该字节无任何意义，按协议规定填充为全 1 广播地址。
Control	1 字节	同 Address 域一样，PPP 数据帧的 Control 域也没实际意义，规定值为 0x03，该域与 Address 域一起标识了 PPP 报文，即 PPP 报文头为 FF03。
Protocol	1 字节或 2 字节	协议域，可用来区分 PPP 数据帧中信息域所承载的数据报文的内容。协议域的内容必须依据 ISO 3309 的地址扩展机制所给出的规定。该机制规定协议域所填充的内容必须为奇数，也就是要求低字节的最低位为“1”，高字节的最低位为“0”。如果当发送端发送的 PPP 数据帧的协议域字段不符合上述规定，接收端则会认为此数据帧是不可识别的。接收端向发送端发送一个 Protocol-Reject 报文，在该报文尾部将填充被拒绝报文的协议号。 • 0021: IP 报文 • 002b: Novell IPX • 002d: Van Jacobson Compressed TCP/IP • 002f: Van Jacobson Uncompressed TCP/IP • 0057: IPV6 报文 • 8021: IPCP 报文 • 802b: Novell IPX Control Protocol • 8031: Bridging NC • 8057: IPv6 CP 报文 • C021: LCP 报文

字段	长度	含义
		• C023: Password Authentication Protocol • C223: Challenge Handshake Authentication Protocol
Information	0~1500 字节	信息域最大长度是 1500 字节,其中包括填充域的内容。信息域的最大长度等于 PPP 协议中 MRU(Maximum Receive Unit）的缺省值。在实际应用当中可根据实际需要进行信息域最大封装长度选项的协商。 如果信息域长度不足 1500 字节，可被填充，但不是必须的。如果填充则需通信双方的两端能辨认出有用与无用的信息方可正常通信。
FCS	0/1/2 字节	FCS 域计算范围是除了 flag 域的其他域。 校验域的功能主要对 PPP 数据帧传输的正确性进行检测。 在数据帧中引入了一些传输的保证机制，会引入更多的开销，这样可能会增加应用层交互的延迟。
Code	1 字节	代码域，主要是用来标识 LCP 数据报文的类型。在链路建立阶段，接收方接收到 LCP 数据报文。当其代码域的值无效时，就会向对端发送一个 LCP 的代码拒绝报文（Code-Reject 报文）。 如果是 IP 报文，则不存在此域，取而代之的是 IP 报文内容。 常见 Code 值如下： 0x01: Configure-Request 0x02: Configure-Ack 0x03: Configure-Nak 0x04: Configure-Reject 0x05: Terminate-Request 0x06: Terminate-Ack 0x07: Code-Reject 0x08: Protocol-Reject 0x09: Echo-Request 0x0a: Echo-Replyt 0x0b: Discard-Request 0x0c: Reserved
Identifier	1 字节	标识域的值表示进行协商报文的匹配关系。 标识域目的是用来匹配请求和响应报文。 一般而言，在进入链路建立阶段时，通信双方任何一端都会连续发送几个配置请求报文（Configure-Request 报文)。这几个请求报文的数据域的值可能是完全一样的，只是它们的标志域不同。 通常一个配置请求报文的 ID 是从 0x01 开始逐步加 1 的。

字段	长度	含义
		当对端接收到该配置请求报文后，无论使用何种报文回应对方，但必须要求回应报文中的 ID 要与接收报文中的 ID 一致。当通信设备收到回应后就可以将该回应与发送时的进行比较来决定下一步的操作。
Length	2 字节	长度域表示此协商报文长度，它包含 Code 域及 Identifier 域的长度。长度域的值就是该 LCP 报文的总字节数据。它是代码域、标志域、长度域和数据域四个域长度的总和。 长度域所指示字节数之外的字节将被当作填充字节而忽略掉，而且该域的内容不能超过 MRU 的值。
Data	变长	数据域所包含的是协商报文的内容。 • Type 为协商选项类型。 常见 Type 中的协商类型值： 0x01: Maximum-Receive-Unit 0x02: Async-Control-Character-Map 0x03: Authentication-Protocol 0x04: Quality-Protocol 0x05: Magic-Number 0x06: RESERVED 0x07: Protocol-Field-Compression 0x08: Address-and-Control-Field-Compression • Length 为协商选项长度，它是指 Data 域的总长度，也就是包含 Type、Length 和 Data。 • Data 为协商的选项具体内容。

参考标准

标准	描述
RFC 1661	The Point-to-Point Protocol （PPP）
RFC 1055	A NONSTANDARD FOR TRANSMISSION OF IP DATAGRAMS OVER SERIAL LINES: SLIP
RFC 1144	Compressing TCP/IP headers for low-speed serial links
RFC1717	The PPP Multilink Protocol (MP)
RFC1332	The PPP Internet Protocol Control Protocol （IPCP）

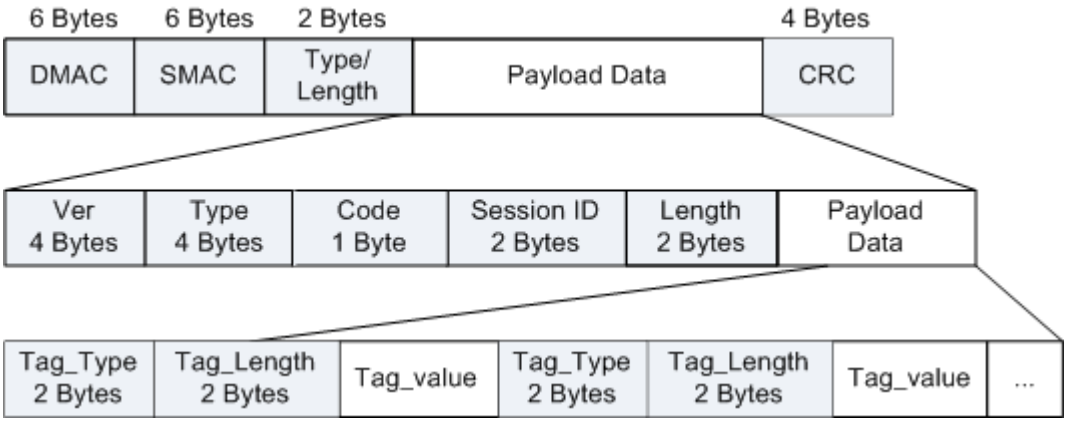
标准	描述
RFC1962	The PPP Compression Control Protocol （CCP）

2.5 PPPoE 报文格式

PPPoE 是对 PPP 的扩展，它可以使 PPP 协议应用于以太网。

帧格式

PPPoE 帧格式如下：



PPPoE 帧中字段解释：

字段	长度	含义
DMAC	6 字节	以太网单播目的地址或者以太网广播地址（0xFFFFFFFF）。在 Discovery 数据包中，该域的值是以太网广播地址；在 PPPoE 会话流量中，该域必须是 Discovery 阶段已经确定的通信对方的单播地址。ry stage.
SMAC	6 字节	源设备的以太网 MAC 地址。
Type/Length	2 字节	链路直接封装的协议，当值为 0x8863 时表示 Discovery 阶段；当值为 0x8864 时表示 PPPoE 会话阶段。
Payload Data	可变	以太帧的数据字段。The Ethernet payload.

字段	长度	含义
CRC	4 字节	用于帧内后续字节差错的循环冗余检验（也称为 FCS 或帧检验序列）。
Ver	4 字节	PPPoE 版本，必须设置为 0x01。
Type	4 字节	PPPoE 类型，必须设置为 0x01。
Code	1 字节	其定义在后面的 Discovery 和 PPPoE 会话中分别指定。
Session ID	2 字节	是一个网络字节序的无符号值。其值在后面 Discovery 数据包中定义。对一个给定的 PPPoE 会话来说该值是一个固定值，并且与以太网 Source_address 和 Destination_address 一起实际地定义了一个 PPPoE 会话。值 0xFFFF 为将来的使用保留，不允许使用。
Length	2 字节	该值是 PPPoE 的 Payload 长度。它不包括以太网头部和 PPPoE 头部的长度。
Payload Data	可变	PPPoE 的 Payload，包含 0 个或多个 Tag。一个 Tag 是一个 TLV（Type-Length-Value）结构。
Tag_type	2 字节	网络字节序。下表列出了各种 Tag_Type 和 Tag_Value 的对应关系和含义。
Tag_Length	2 字节	是一个网络字节序的无符号值，表明 Tag_Value 的字节数。 如果收到的 Discovery 数据包中包含未知的 Tag_Type,则必须忽略掉该 Tag。It is an unsigned number in network byte order, indicating the length in octets of the TAG_VALUE.
Tag_value	可变	Tag 的数据字段。Value of a Tag.

表 1 TLV

Tag_Value	Tag_Type	含义
0x0000	End-Of-List	该 Tag 值表明是最后一个 Tag。该 Tag 的 Tag_Length 必须总是 0。 不要求使用该标签，它是为了向后兼容。

字段	长度	含义
0x0101	Service-Name	该 Tag 表明后面紧跟的是服务的名称。 • Tag_Value 是不以 NULL 结束的字符串。 • 当 Tag_Length 为 0 时，该 TAG 用于表明接受任何服务。 使用 Service-Name 标签的例子是表明 Internet 服务提供商 ISP 或者一类服务或者服务的质量。
0x0102	AC-Name	该 Tag 表明后面紧跟的字符串唯一地表示了某个特定的接入服务器。 它可以是商标、型号以及序列号等信息的集合，或者该接入服务器 MAC 地址的一个简单表示。它不以 NULL 来结束。
0x0103	Host-Uniq	该 Tag 由主机用于把接入服务器的响应报文（PADO 或者 PADS）与主机的某个唯一特定的请求联系起来。Tag_Value 是主机选择的长度和值，可以是任意的二进制数据。它不能由接入服务器解释。 主机可以在 PADI 或者 PADR 中包含一个 Host-Uniq 标签。如果接入服务器收到了该标签，它必须在对应的 PADO 或者 PADS 中不加改变的包含该标签。
0x0104	AC-Cookie	该 Tag 由接入服务器用于防止服务攻击。接入服务器可以在 PADO 数据包中包含该 Tag。如果主机收到了该标签，它必须在接下来的 PADR 中不加改变的包含该标签。 Tag_Value 的长度和值都是任意的二进制数据。
0x0105	Vendor-Specific	该 Tag 用来传送厂商自定义的信息。Tag_Value 的前 4 个字节包含了厂商的识别码，其余字节尚未定义。 厂商识别码的高字节为 0，低 3 个字节为网络字节序的厂商的 SMI 网络管理专用企业码。 不推荐使用该 Tag。为了确保互操作性，在实现过程中，可以忽略 Vendor-Specific Tag。
0x0110	Relay-Session-Id	该 Tag 可由中继流量的中间代理加入到 Discovery 数据包中。 Tag_Value 对主机和接入服务器都是不透明。如果主机或接入服务器收到该 Tag，则它们必须在所有的 Discovery 数据包中包含该 Tag 以作为响应。 所有的 PADI 数据包必须保证足够空间来加入 Tag_Value 长度为 12 字节的 Relay-Session-Id 标签。 如果 Discovery 数据包中已经包含一个 Relay-Session-Id 标签，则不允许再加入该标签。这种情况下，中间代理应该使用该 Relay-Session-Id 标签。 如果它不能使用现有的标签，或者没有足够空间来增加一个 Relay-Session-Id 标签，那么它应该向发送者返回一个 Generic-Error 标签。

字段	长度	含义
0x0201	Service-Name-Error	该 Tag 典型的有一个长度为零的数据部分。 它表明了由于某种原因，没有理睬所请求的 Service-Name。如果有数据部分，并且数据部分的头一个字节非 0，那么它必须是一个可打印字符串，解释请求被拒绝的原因。 该字符串可以不以 NULL 结束。
0x0202	AC-System-Error	该 Tag 表明了接入服务器在处理主机请求时出现了某个错误。例如没有足够资源来创建一个虚拟电路。PADS 数据包中可以包含该标签。 如果有数据，并且数据的第一个字节不为 0，那么数据必须是一个可打印字符串，该字符串解释了错误的性质。 该字符串可以不以 NULL 结束。
0x0203	Generic-Error	该 Tag 表明发生了一个错误。 当发生一个不可恢复的错误并且没有其它合适的 Tag 时，它可被加到 PADO、PADR 或 PADS 数据包中。 如果出现数据部分，那么数据必须是一个解释错误性质的字符串。 该字符串不允许以 NULL 结束。

帧示例

图 1 PPPoED (PADI)

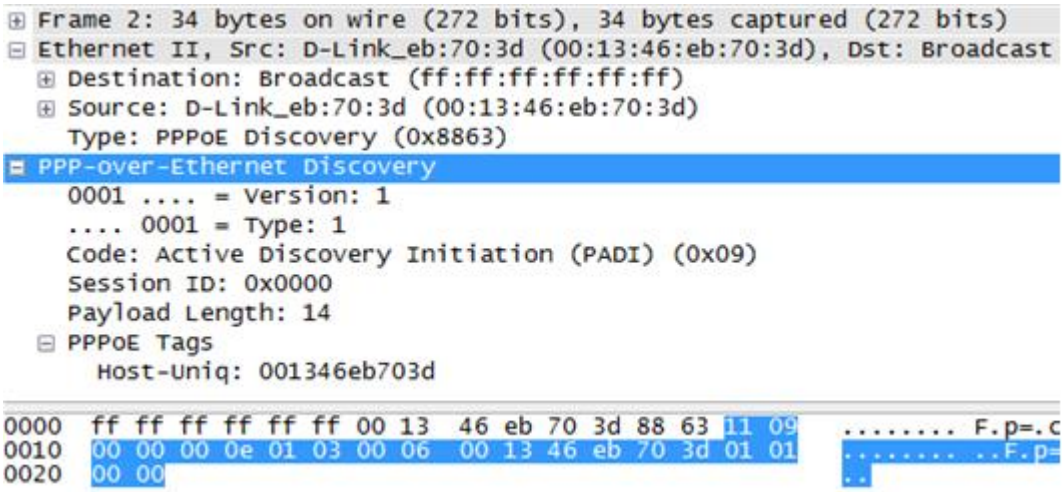


图 2 PPPoED (PADO)

```
⊞ Frame 3: 60 bytes on wire (480 bits), 60 bytes captured (480 b
⊟ Ethernet II, Src: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf), Dst:
  ⊞ Destination: D-Link_eb:70:3d (00:13:46:eb:70:3d)
  ⊞ Source: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf)
    Type: PPPoE Discovery (0x8863)
⊟ PPP-over-Ethernet Discovery
    0001 .... = Version: 1
    .... 0001 = Type: 1
    Code: Active Discovery Offer (PADO) (0x07)
    Session ID: 0x0000
    Payload Length: 25
  ⊟ PPPoE Tags
      Host-Uniq: 001346eb703d
      AC-Name: Quidway

0000  00 13 46 eb 70 3d 00 e0  fc 59 b1 cf 88 63 11 07   ..F.p=.
0010  00 00 00 19 01 03 00 06  00 13 46 eb 70 3d 01 01   .......
0020  00 00 01 02 00 07 51 75  69 64 77 61 79 00 00 00   ......C
0030  00 00 00 00 aa aa aa aa  00 00 00 00               .......
```

图 3 PPPoED (PADR)

```
⊞ Frame 4: 34 bytes on wire (272 bits), 34 bytes captured (27
⊟ Ethernet II, Src: D-Link_eb:70:3d (00:13:46:eb:70:3d), Dst:
  ⊞ Destination: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf)
  ⊞ Source: D-Link_eb:70:3d (00:13:46:eb:70:3d)
    Type: PPPoE Discovery (0x8863)
⊟ PPP-over-Ethernet Discovery
    0001 .... = Version: 1
    .... 0001 = Type: 1
    Code: Active Discovery Request (PADR) (0x19)
    Session ID: 0x0000
    Payload Length: 14
  ⊟ PPPoE Tags
      Host-Uniq: 001346eb703d

0000  00 e0 fc 59 b1 cf 00 13  46 eb 70 3d 88 63 11 19   ...Y.
0010  00 00 00 0e 01 03 00 06  00 13 46 eb 70 3d 01 01   .....
0020  00 00                                             ..
```

图 4 PPPoED (PADS)

```
⊞ Frame 5: 60 bytes on wire (480 bits), 60 bytes captured (480 b
⊟ Ethernet II, Src: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf), Dst:
  ⊞ Destination: D-Link_eb:70:3d (00:13:46:eb:70:3d)
  ⊞ Source: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf)
    Type: PPPoE Discovery (0x8863)
⊟ PPP-over-Ethernet Discovery
    0001 .... = Version: 1
    .... 0001 = Type: 1
    Code: Active Discovery Session-confirmation (PADS) (0x65)
    Session ID: 0x004f
    Payload Length: 14
  ⊟ PPPoE Tags
      Host-Uniq: 001346eb703d

0000  00 13 46 eb 70 3d 00 e0  fc 59 b1 cf 88 63 11 65   ..F.p=..
0010  00 4f 00 0e 01 03 00 06  00 13 46 eb 70 3d 01 01   .O......
0020  00 00 00 00 00 00 00 00  00 00 00 00 00 00 00 00   ........
0030  00 00 00 00 aa aa aa aa  00 00 00 00               ........
```

图 5 PPP LCP

```
⊞ Frame 6: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
⊟ Ethernet II, Src: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf), Dst: D-Link_eb:70:3d (00:1
  ⊞ Destination: D-Link_eb:70:3d (00:13:46:eb:70:3d)
  ⊞ Source: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf)
    Type: PPPoE Session (0x8864)
⊟ PPP-over-Ethernet Session
    0001 .... = Version: 1
    .... 0001 = Type: 1
    Code: Session Data (0x00)
    Session ID: 0x004f
    Payload Length: 21
⊟ Point-to-Point Protocol
    Protocol: Link Control Protocol (0xc021)
⊟ PPP Link Control Protocol
    Code: Configuration Request (0x01)
    Identifier: 0x00
    Length: 19
  ⊟ Options: (15 bytes)
      Maximum Receive Unit: 1492
    ⊟ Authentication protocol: 5 bytes
        Authentication protocol: Challenge Handshake Authentication Protocol (0xc223)
        Algorithm: CHAP with MD5 (0x05)
      Magic number: 0x0200106f
```

```
0000  00 13 46 eb 70 3d 00 e0  fc 59 b1 cf 88 64 11 00   ..F.p=.. .Y...d..
0010  00 4f 00 15 c0 21 01 00  00 13 01 04 05 d4 03 05   .O...!.. ........
0020  c2 23 05 05 06 02 00 10  6f 00 00 00 00 00 00 00   .#...... o.......
0030  00 00 00 00 aa aa aa aa  00 00 00 00               ........ ....
```

图 6 PPP CHAP

```
⊞ Frame 12: 60 bytes on wire (480 bits), 60 bytes captured (480 bit
⊟ Ethernet II, Src: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf), Dst: D-L
  ⊞ Destination: D-Link_eb:70:3d (00:13:46:eb:70:3d)
  ⊞ Source: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf)
    Type: PPPoE Session (0x8864)
⊟ PPP-over-Ethernet Session
    0001 .... = Version: 1
    .... 0001 = Type: 1
    Code: Session Data (0x00)
    Session ID: 0x004f
    Payload Length: 23
⊟ Point-to-Point Protocol
    Protocol: Challenge Handshake Authentication Protocol (0xc223)
⊟ PPP Challenge Handshake Authentication Protocol
    Code: Challenge (1)
    Identifier: 1
    Length: 21
  ⊟ Data
      Value Size: 16
      Value: 2dd3ccaaab2ab3eb14bcffc56c393433
```

```
0000  00 13 46 eb 70 3d 00 e0  fc 59 b1 cf 88 64 11 00   ..F.p=.. .Y...d..
0010  00 4f 00 17 c2 23 01 01  00 15 10 2d d3 cc aa ab   .O...#.. ...-....
0020  2a b3 eb 14 bc ff c5 6c  39 34 33 00 00 00 00 00   *......l 943.....
0030  00 00 00 00 aa aa aa aa  00 00 00 00               ........ ....
```

图 7 PPP IPCP

```
⊞ Frame 19: 56 bytes on wire (448 bits), 56 bytes captured (44
⊟ Ethernet II, Src: D-Link_eb:70:3d (00:13:46:eb:70:3d), Dst:
  ⊞ Destination: HuaweiTe_59:b1:cf (00:e0:fc:59:b1:cf)
  ⊞ Source: D-Link_eb:70:3d (00:13:46:eb:70:3d)
    Type: PPPoE Session (0x8864)
⊟ PPP-over-Ethernet Session
    0001 .... = Version: 1
    .... 0001 = Type: 1
    Code: Session Data (0x00)
    Session ID: 0x004f
    Payload Length: 36
⊟ Point-to-Point Protocol
    Protocol: IP Control Protocol (0x8021)
⊟ PPP IP Control Protocol
    Code: Configuration Request (0x01)
    Identifier: 0x02
    Length: 34
  ⊟ Options: (30 bytes)
      IP address: 10.164.19.144
      Primary DNS server IP address: 10.72.255.100
      Secondary DNS server IP address: 10.72.55.101
      Primary WINS server IP address: 10.164.33.252
      Secondary WINS server IP address: 10.164.33.250
```

```
0000  00 e0 fc 59 b1 cf 00 13  46 eb 70 3d 88 64 11 00   ...Y.
0010  00 4f 00 24 80 21 01 02  00 22 03 06 0a a4 13 90   .O.$.
0020  81 06 0a 48 ff 64 83 06  0a 48 37 65 82 06 0a a4   ...H.
0030  21 fc 84 06 0a a4 21 fa                           !....
```

参考标准

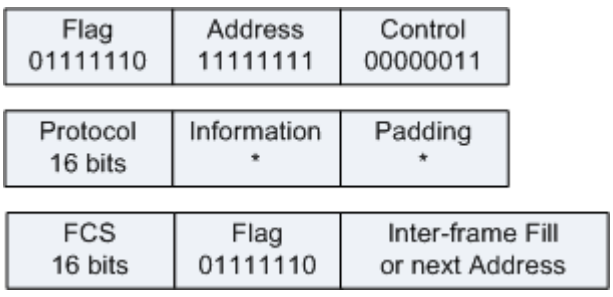
标准	描述
RFC 2516	A Method for Transmitting PPP Over Ethernet (PPPoE)

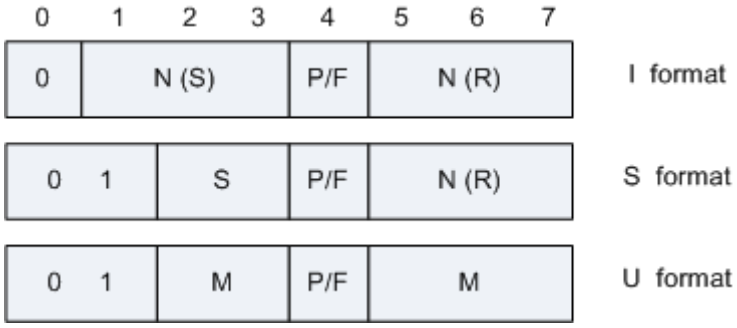
2.6 HDLC 帧格式

帧格式

在 HDLC 中，数据和控制报文均以帧的标准格式传送。HDLC 中的帧类似于 BSC 的字符块，但不是独立传输的。HDLC 的完整的帧由标志字段（F）、地址字段（A）、控制字段（C）、信息字段（I）、帧校验序列字段（FCS）等组成：

图 1 HDLC 帧格式



字段	长度	含义
Flag	1 字节	标志字段，为 01111110(0x7e)的比特模式，用以标志帧的开始与结束，也可以作为帧与帧之间的填充字符。通常，在不进行帧传送的时刻，信道仍处于激活状态，在这种状态下，发送方不断地发送标志字段，而接收方则检测每一个收到的标志字段，一旦发现某个标志字段后面不再是一个标志字段，便可认为新的帧传动已经开始。采用“0 比特插入法”可以实现数据的透明传输。
Address	1 字节	地址字段，内容取决于所采用的操作方式，有主节点、从节点、组合节点之分。每个从节点与组合节点都被分配一个唯一的地址，命令帧中的地址字段携带的是对方节点的地址，而响应帧中的地址字段所携带的地址是本节点的地址。某一地址也可分配给不止一个节点，这种地址称为组地址，利用一个组地址传输的帧能被组内所有拥有该地址的节点接收。但当一个节点或组合节点发送响应时，它仍应当用它唯一的地址。还可以用全“1”地址来表示包含所有节点的地址，称为广播地址，含有广播地址的帧传送给链路上所有的节点。另外还规定全 0 的地址为无节点地址，不分配给任何节点，仅作为测试用。
Control	1 字节	控制字段，用于构成各种命令及响应，以便对链路进行监视与控制。发送方主节点或组合节点利用控制字段来通知被寻址的从节点或组合节点执行约定的操作；相反，从节点用该字段作为对命令的响应，报告已经完成的操作或状态的变化。该字段是 HDLC 的关键。 由于 Control 字段的构成不同，可以把 HDLC 帧分为三种类型：信息帧、监控帧、无编号帧，分别简称 I 帧(Information)、S 帧(Supervisory)、U 帧(Unnumbered)。在控制字段中，第 1 位是“0”为 I 帧，第 1、2 位是“1 ”为 S 帧，第 1、2 位是“11”为 U 帧。 **图 2** HDLC Control 字段格式  Control 字段帧中的各字段含义如下： • N(S): Send Sequence Number • N(R): Receive Sequence Number • P/F: Poll Bit command frame/Final Bit response frame • M: Modifier Function • S: Supervisory Function 控制字段的第五位是 P/F 位，即轮询/终止位（POLL/Final）位。

字段	长度	含义
Protocol	2 字节	协议字段。表示 Information 域中的数据封装的协议类型。
Information	0~N 字节	信息字段。可以是任意的二进制比特串，长度未作限定。其上限由 FCS 字段或通信节点的缓冲容量来决定，目前国际上用得较多的是 1000～2000 比特，而下限可以是 0，即无信息字段。但是监控帧中不可有信息字段。
FCS	2 字节	FCS(Frame Check Sequence)：帧检验序列字段，可以使用 16 位 CRC，对两个标志字段之间的整个帧的内容进行校验。FCS 的生成多项式是 CCITT V.41 建议的 X'16+X'12+X'5+1。

2.7 ATM 信元格式

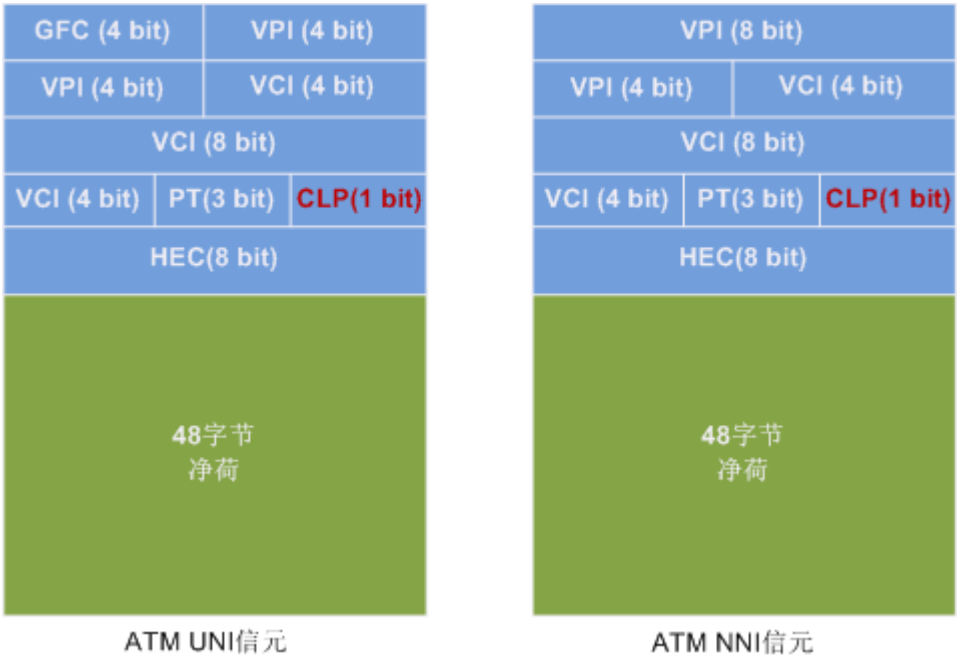
- ATM 信元格式
- AAL5 多协议封装通用格式
- IPoEoA/PPPoEoA 格式
- PPPoA 格式
- ATM PWE3 封装格式

父主题： 链路层

2.7.1 ATM 信元格式

ATM 信元头根据网络接口的不同分为 UNI 和 NNI 两种信元头类型。

图 1 ATM 信元头格式



- GFC：一般流量控制，长度是 4 比特，只用于 UNI 接口，进行流量控制或在共享媒体的网络中标识不同的接入访问。
- VPI：虚通道标识符，在 UNI 中长度是 8 比特，可标识 256 个 VP；在 NNI 中长度是 12 比特，可标识 4096 个 VP。
- VCI：虚通路标识符，长度是 16 比特，可标识 65536 个 VC。
- CLP：信元丢弃优先等级（Cell Loss Priority），长度是 1 比特，用于拥塞控制。发生拥塞时优先丢弃 CLP=
- PTI：净荷类型标识（Payload Type Indicator），长度是 3 比特，用于标识净荷的类型。
- HEC：信头差错控制，长度是 8 比特，用于信元头中的差错控制和信元定界。可纠正 1 位错码，发现多位错码。在物理层进行 HEC 处理。

对于一些特定的 VPI/VCI 值已经保留作为特殊信元使用，下面对它们进行简单介绍。

- 空闲信元：VPI=
- 未赋值信元：VPI=
- OAM 信元：
  - 对于 VP，VCI=
  - 对于 VC，PTI=
- 信令信元：它分为以下三种类型：
  - 元信令信元：VPI 为任意值，VCI=
  - 一般广播信令信元：VPI 为任意值，VCI=
  - 点对点信令信元：VPI 为任意值，VCI=5。
- 净荷类型 PT（Payload Type）：该域长度是 3 比特。用于标识信息域，也就是净荷的类型。下面列出的是 ITU-T I.361 已定义的 PT 值及其含义。
  - PT=000：用户数据信元，未经历拥塞，ATM 层用户到 ATM 层用户指示 AUU（ATM User to User）为 0。
  - PT=
  - PT=
  - PT=
  - PT=100：OAM F5 段相关信元。
  - PT=101：OAM F5 端到端相关信元。
  - PT=
  - PT=

由此可见，当信元用于承载用户数据时：

- PT 第一位为 0。
- 第二位标识信元是否经历拥塞，这一位可通过处于拥塞的网络节点设置。
- 第三位是 AUU 指示，其中，AUU=0 表明对应的 SAR-PDU 是起始段或中间段，AUU=1 表明为结束段。

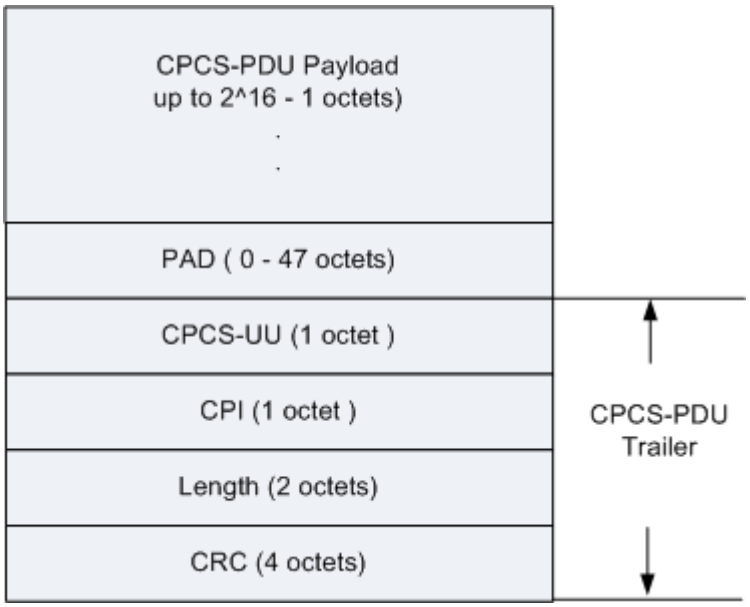
2.7.2 AAL5 多协议封装通用格式

在 ATM 适配层，针对不同的业务，有不同的 AAL 协议，比如，AAL2 针对实时性要求高、数据量较小的话音业务，AAL5 针对数据量较大但没有实时性要求的数据传输。

AAL5 PDU 格式

AAL5 CPCS 子层 CPCS-PDU 格式如下：

图 1 AAL5 CPCS-PDU 格式



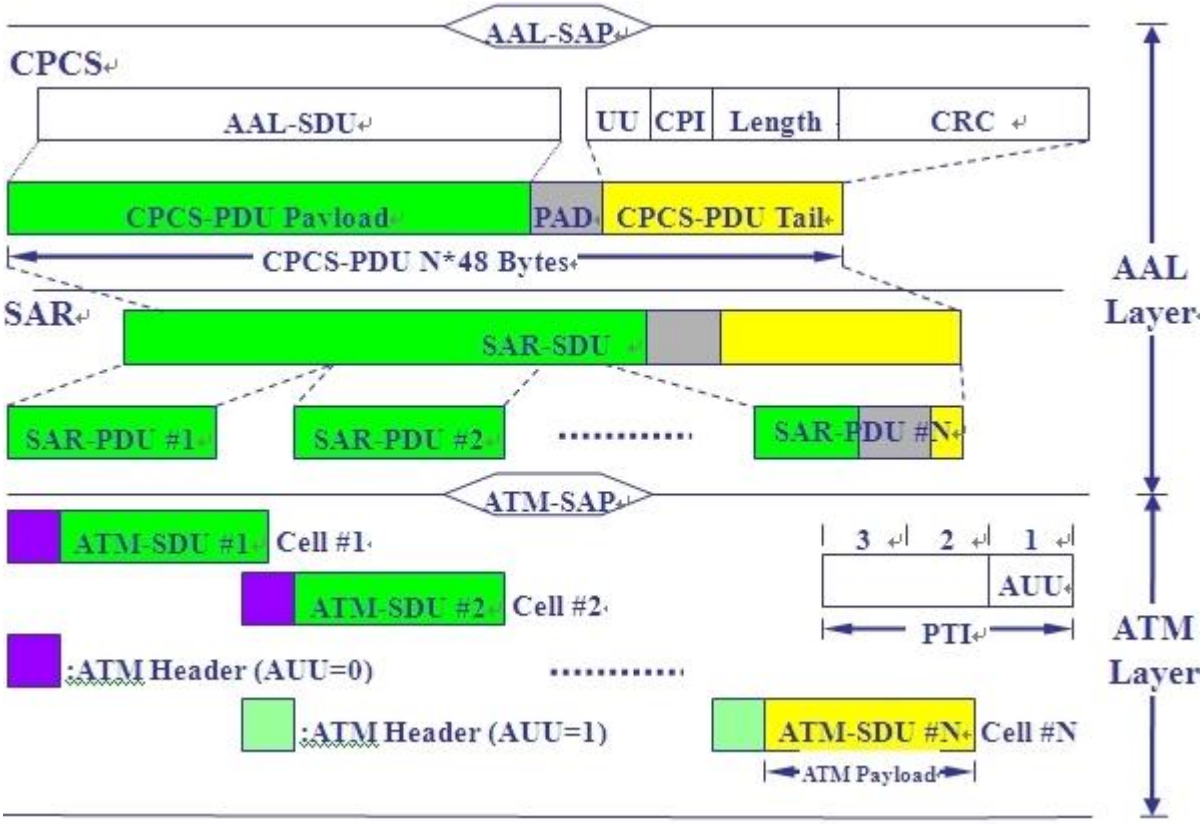
字段	含义
CPCS-PDU Payload	CPCS－PDU 净荷：用于承载 CPCS 用户信息，长度是可变的，范围为 1～65535 字节。
PAD	PAD：填充位，使整个 CPCS－PDU 长度为 48 字节的整数倍，长度范围是 0～47 字节。
CPCS-UU	CPI：用作使 CPCS－PDU 尾部长为 8 个字节，占 8 位。
CPI	CPI：用作使 CPCS－PDU 尾部长为 8 个字节，占 8 位。
Length	指示 CPCS－PDU 净荷长度，占 16 位。
CRC	CRC：循环冗余校验，占 32 位。

AAL5 SAR

AAL5 SAR 将 CPCS-PDU 分成 48 字节的 SAR-PDU。

AAL5 适配过程

图 2 AAL5 适配过程



在 AAL 的 CPCS 子层，业务数据单元由 AAL5 在 CPCS-PDU 净荷的尾部加了 CPCS-PDU 尾，然后由 PAD 把整个 CPCS-PDU 填充成为 48 字节的整数倍（图中表示成 N*48）的数据单元。这样 CPCS 完成了它的任务，它就把这个 N*48 的 CPCS-PDU 数据单元发送给 SAR 子层，AAL5-SAR 将这个 CPCS-PDU 分成 N 个 48 字节的 SAR-PDU 单元。这样 AAL5 就完成了它的任务，它把这 N 个 48 字节的 SAR-PDU 单元传送给 ATM 层。

在 ATM 层，SAR-PDU 被表示成为 ATM-PDU，每一个 ATM-PDU 被加上一个信元头。

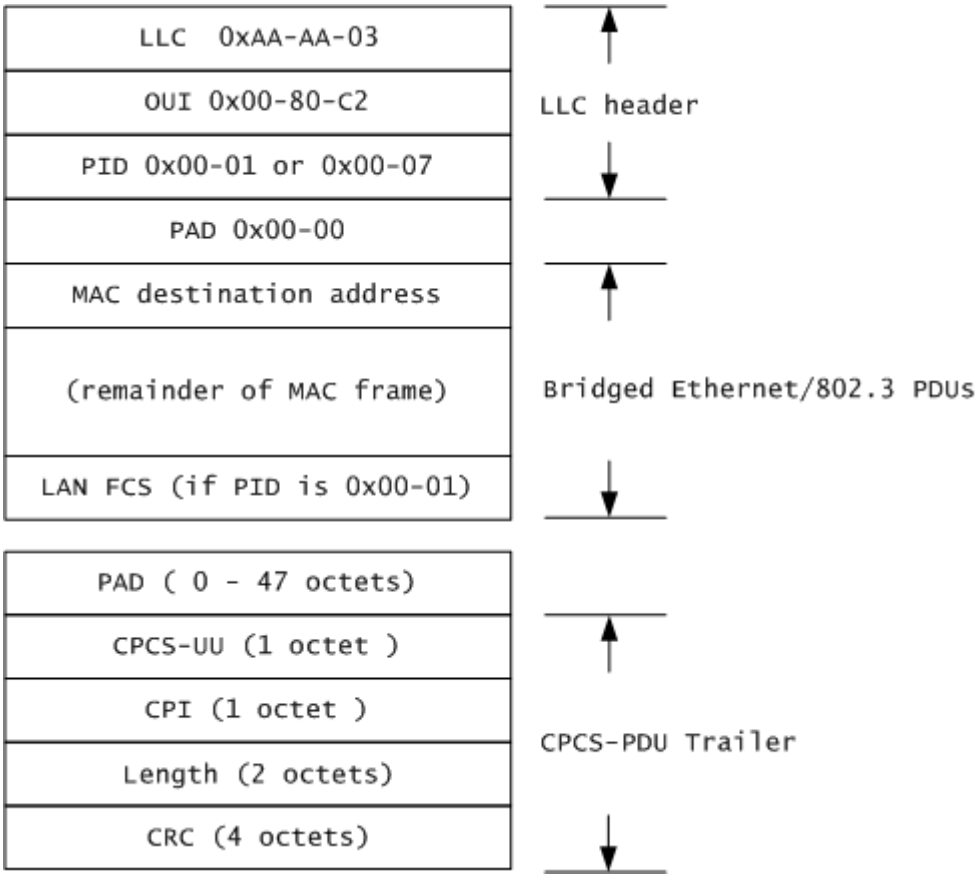
ATM 传送信息的基本单位是 53 字节长的信元，由 5 个字节的信元头和 48 字节的净荷组成。由 AAL 传到 ATM 层的就是一个一个的 48 字节净荷，这些 48 字节的净荷在 ATM 层就被加上了 5 个字节的信元头，这样就组成了 53 字节的信元。

Reference

标准	描述
RFC 1483	Multiprotocol Encapsulation over ATM Adaptation Layer 5
RFC 2364	PPP Over AAL5
RFC 2684	Multiprotocol Encapsulation over ATM Adaptation Layer 5

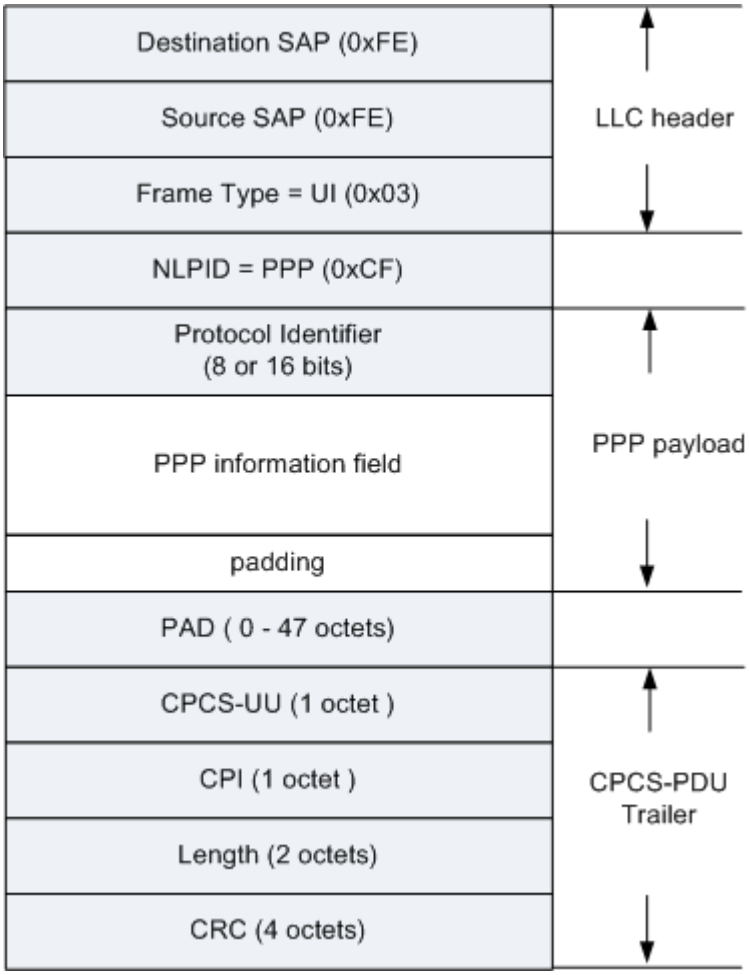
2.7.3 IPoEoA/PPPoEoA 格式

IPoEoA/PPPoEoA AAL5 封装格式



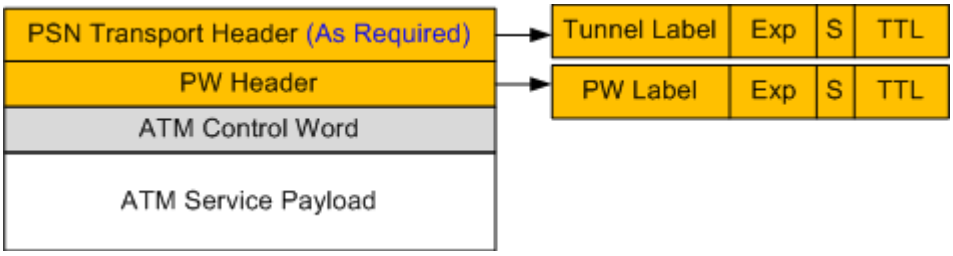
2.7.4 PPPoA 格式

图 1 PPPoA 格式



2.7.5 ATM PWE3 封装格式

图 1 ATM PWE3 封装格式



2.8 STP/RSTP/MSTP 帧格式

STP 帧格式

图 1 STP 帧格式

	Octet
Protocol Identifier	1-2
Protocol Version Identifier	3
BPDU Type	4
Flags	5
Root Identifier	6-13
Root Path Cost	14-17
Bridge Identifier	18-25
Port Identifier	26-27
Message Age	28-29
Max Age	30-31
Hello Time	32-33
Forward Delay	34-35

字段内容	说明
Protocol Identifier	协议 ID＝“0”
Protocol Version Identifier	协议版本标识符，STP 为 0，RSTP 为 2，MSTP 为 3。
BPDU Type	BPDU 类型，MSTP 为 0x02。 • 0x00：STP 的 Configuration BPDU • 0x80：STP 的 TCN BPDU（Topology Change Notification BPDU） • 0x02：RST BPDU（Rapid Spanning-Tree BPDU）或者 MST BPDU（Multiple Spanning-Tree BPDU）
Flags	对于“标记域”（Flags），第一个 bit（左边、高位 bit）表示“TCA（拓扑改变响应）”，最后一个 bit（右边、低位 bit）表示“TC（拓扑改变）”。
Root Identifier	网桥 ID 都是 8 个字节——前两个字节是网桥优先级，后 6 个字节是网桥 MAC 地址。
Root Path Cost	根路径开销，本端口累计到根桥的开销。
Bridge Identifier	发送者 BID，本交换机的 BID。
Port Identifier	发送端口 PID，发送该 BPDU 的端口 ID。

字段内容	说明
Message Age	该 BPDU 的消息年龄。
Max Age	消息老化年龄。
Hello Time	发送两个相邻 BPDU 间的时间间隔。
Forward Delay	控制 Listening 和 Learning 状态的持续时间。

RSTP 帧格式

在 BPDU 的格式上，除了保证和 STP 格式基本一致之外，RSTP 作了一些小的变化。一个是在 Type 字段，配置 BPDU 类型不再是 0 而是 2，版本号也变成了 2。所以运行 STP 的交换机收到该类 BPDU 时会丢弃。

另一个变化是在 Flag 字段，把原来保留的中间 6 位使用起来。这样改变了的配置 BPDU 叫做 RST BPDU。

RSTP Flag 字段格式：

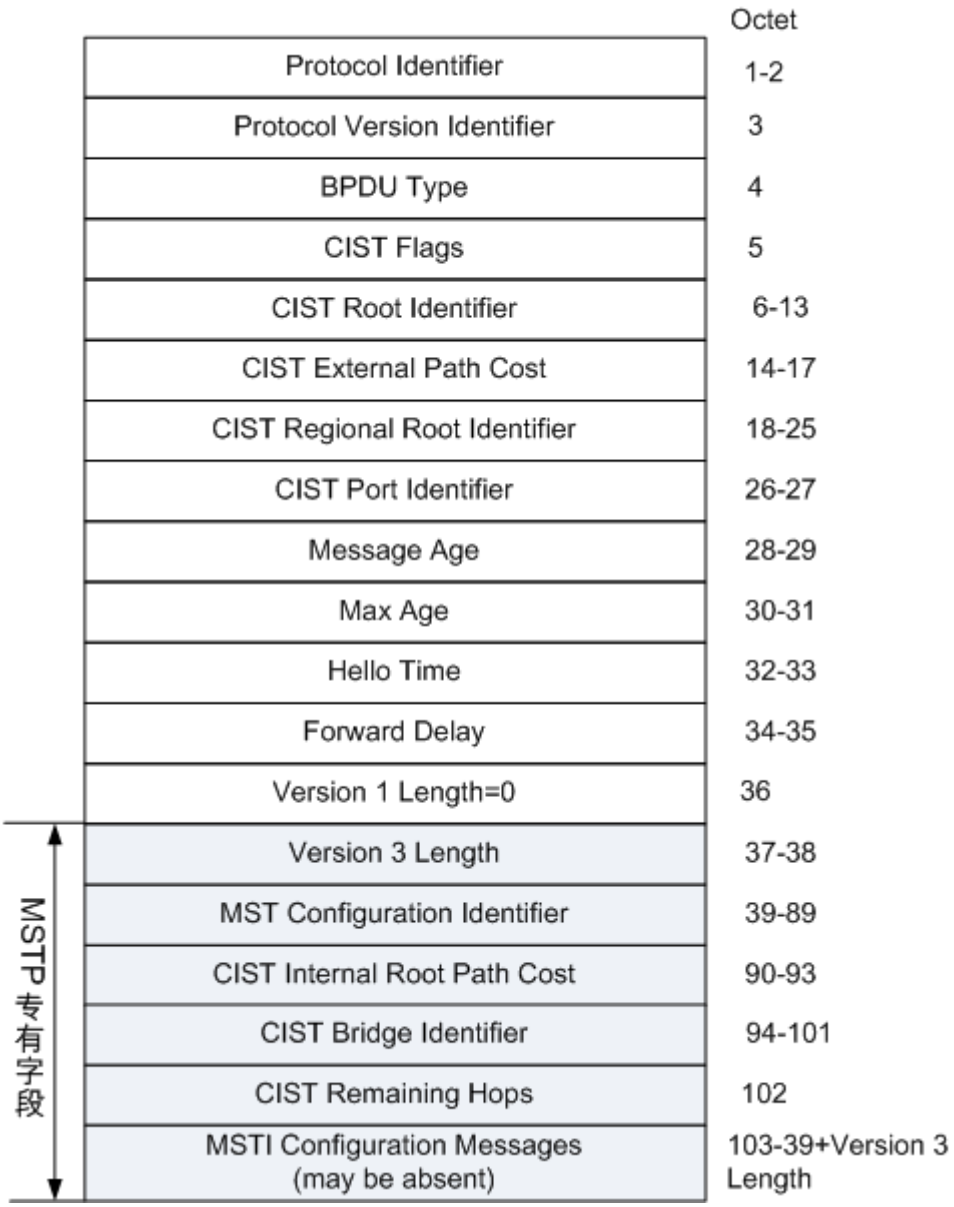
- Bit7：TCA
- Bit6：Agreement
- Bit5：Forwarding
- Bit4：Learning
- Bit3 和 Bit2：端口角色
    - 00：未知
    - 01：根端口
    - 10：Alternate / Backup
    - 11：指定端口
- Bit1：Proposal
- Bit0：TC

MSTP 帧格式

多生成树协议 MSTP 是生成树协议的一种，用于消除网络环路，它兼容生成树协议 STP 和快速生成树 RSTP 协议，并且弥补了两者的缺陷。

MSTP 使用多生成树桥协议数据单元 MST BPDU（Multiple Spanning Tree Bridge Protocol Data Unit）作为生成树计算的依据。 MST BPDU 报文用来计算生成树的拓扑、维护网络拓扑以及传达拓扑变化记录。MST BPDU 报文结构如下图所示：

图 2 MSTP 帧格式

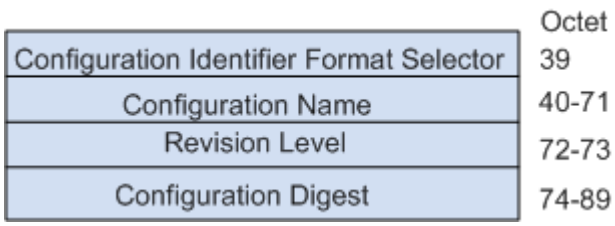


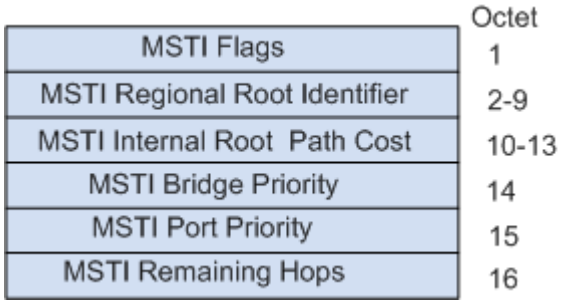
无论是域内的 MST BPDU 还是域间的，前 35 个字节和 RST BPDU 相同。

从第 36 个字节开始是 MSTP 专有字段。最后的 MSTI 配置信息字段由若干 MSTI 配置信息组连缀而成。

MST BPDU 中的主要信息如下表所示。

字段	说明
Protocol Identifier	协议标识符。
Protocol Version Identifier	协议版本标识符，STP 为 0，RSTP 为 2，MSTP 为 3。
BPDU Type	BPDU 类型，MSTP 为 0x02。 • 0x00：STP 的 Configuration BPDU

字段	说明
	• 0x80：STP 的 TCN BPDU（Topology Change Notification BPDU） • 0x02：RST BPDU（Rapid Spanning-Tree BPDU）或者 MST BPDU（Multiple Spanning-Tree BPDU）
CIST Flags	CIST 标志字段。
CIST Root Identifier	CIST 的总根交换机 ID。
CIST External Path Cost	CIST 外部路径开销指从本交换机所属的 MST 域到 CIST 根交换机的累计路径开销。CIST 外部路径开销根据链路带宽计算。
CIST Regional Root Identifier	CIST 的域根交换机 ID，即 IST Master 的 ID。 如果总根在这个域内，那么域根交换机 ID 就是总根交换机 ID。
CIST Port Identifier	本端口在 IST 中的指定端口 ID。
Message Age	BPDU 报文的生存期。
Max Age	BPDU 报文的最大生存期，超时则认为到根交换机的链路故障。
Hello Time	Hello 定时器，缺省为 2 秒。
Forward Delay	Forward Delay 定时器，缺省为 15 秒。
Version 1 Length	Version1 BPDU 的长度，值固定为 0。
Version 3 Length	Version3 BPDU 的长度。
MST Configuration Identifier	MST 配置标识，表示 MST 域的标签信息，包含 4 个字段：  • Configuration Identifier Format Selector：固定为 0。 • Configuration Name：“域名”，32 字节长字符串。 • Revision Level：2 字节非负整数。 • Configuration Digest：利用 HMAC-MD5 算法将域中 VLAN 和实例的映射关系加密成 16 字节的摘要。 只有 MST Configuration Identifier 中的四个字段完全相同的，并且互联的交换机，才属于同一个域。
CIST Internal	CIST 内部路径开销指从本端口到 IST Master 交换机的累计路径开销。CIST 内部路径开销根据链路带宽计算。

字段	说明
Root Path Cost	
CIST Bridge Identifier	CIST 的指定交换机 ID。
CIST Remaining Hops	BPDU 报文在 CIST 中的剩余跳数。
MSTI Configuration Messages (may be absent)	MSTI 配置信息。每个 MSTI 的配置信息占 16 bytes，如果有 n 个 MSTI 就占用 n×16bytes。单个 MSTI Configuration Messages 的字段说明如下：  • MSTI Flags：MSTI 标志。 • MSTI Regional Root Identifier：MSTI 域根交换机 ID。 • MSTI Internal Root Path Cost：MSTI 内部路径开销指从本端口到 MSTI 域根交换机的累计路径开销。MSTI 内部路径开销根据链路带宽计算。 • MSTI Bridge Priority：本交换机在 MSTI 中的指定交换机的优先级。 • MSTI Port Priority：本交换机在 MSTI 中的指定端口的优先级。 • MSTI Remaining Hops：BPDU 报文在 MSTI 中的剩余跳数。

帧示例

图 3 STP 帧格式

```
⊞ Frame 1: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
⊞ IEEE 802.3 Ethernet
⊞ Logical-Link Control
⊟ Spanning Tree Protocol
    Protocol Identifier: Spanning Tree Protocol (0x0000)
    Protocol Version Identifier: Spanning Tree (0)
    BPDU Type: Configuration (0x00)
  ⊟ BPDU flags: 0x00
      0... .... = Topology Change Acknowledgment: No
      .... ...0 = Topology Change: No
  ⊟ Root Identifier: 32768 / 100 / 00:1c:0e:87:78:00
      Root Bridge Priority: 32768
      Root Bridge System ID Extension: 100
      Root Bridge System ID: 00:1c:0e:87:78:00
    Root Path Cost: 4
  ⊟ Bridge Identifier: 32768 / 100 / 00:1c:0e:87:85:00
      Bridge Priority: 32768
      Bridge System ID Extension: 100
      Bridge System ID: 00:1c:0e:87:85:00
    Port identifier: 0x8004
    Message Age: 1
    Max Age: 20
    Hello Time: 2
    Forward Delay: 15
```

图 4 RSTP 帧格式

```
⊞ Frame 1320: 64 bytes on wire (512 bits), 64 bytes captured (512 bits)
⊞ IEEE 802.3 Ethernet
⊞ Logical-Link Control
⊟ Spanning Tree Protocol
    Protocol Identifier: Spanning Tree Protocol (0x0000)
    Protocol Version Identifier: Rapid Spanning Tree (2)
    BPDU Type: Rapid/Multiple Spanning Tree (0x02)
  ⊟ BPDU flags: 0x2c (forwarding, Port Role: Designated)
      0... .... = Topology Change Acknowlegement: No
      .0.. .... = Agreement: No
      ..1. .... = Forwarding: Yes
      ...0 .... = Learning: No
      .... 11.. = Port Role: Designated (3)
      .... ..0. = Proposal: No
      .... ...0 = Topology Change: No
  ⊟ Root Identifier: 32768 / 0 / 00:16:c8:97:0e:40
      Root Bridge Priority: 32768
      Root Bridge System ID Extension: 0
      Root Bridge System ID: 00:16:c8:97:0e:40
    Root Path Cost: 0
  ⊟ Bridge Identifier: 32768 / 0 / 00:16:c8:97:0e:40
      Bridge Priority: 32768
      Bridge System ID Extension: 327
      Bridge System ID: 00:16:c8:97:0e:40
    Port Identifier: 0x8001
    Message Age: 0
    Max Age: 20
    Hello Time: 2
    Forward Delay: 15
    Version 1 Length: 0
```

参考标准

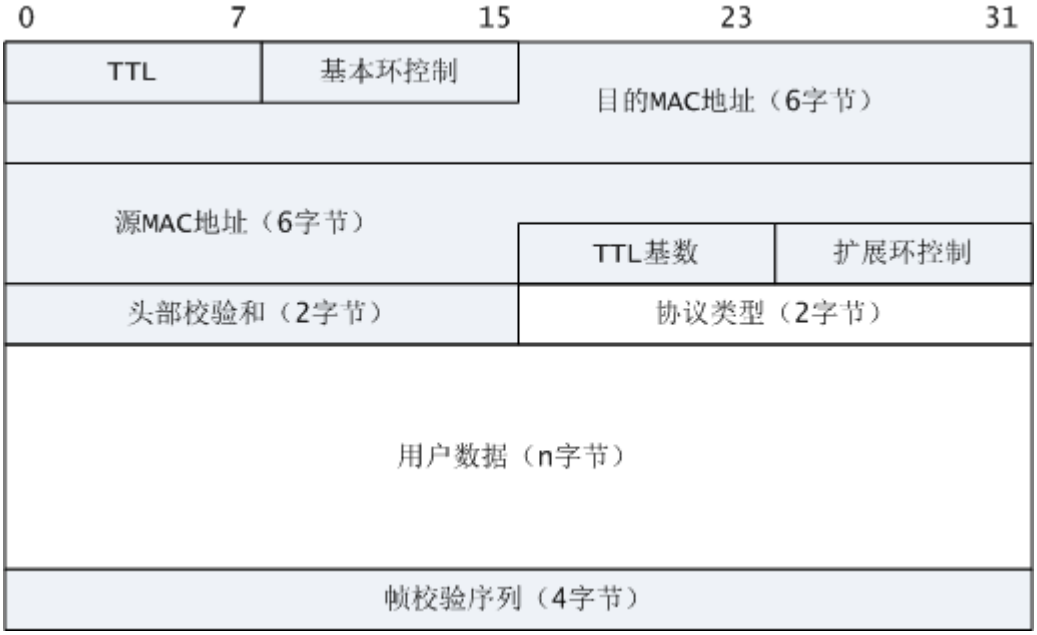
标准	描述
IEEE Std 802.1S	Virtual Bridged Local Area Networks—Amendment 3: Multiple Spanning Trees

标准	描述
IEEE Std 802.1D	Information technology — Telecommunicationsand information exchange between systems — Local and metropolitan area networks—Common specifications—Part 3:Media Access Control (MAC) Bridges.]
IEEE Std 802.1W	IEEE Standard for Information technology — Telecommunications and information exchange between systems—Local and metropolitan area networks—Common specifications—Part 3: Media Access Control (MAC) Bridges —Amendment 2: Rapid Reconfiguration. [Amendment to IEEE Std 802.1D, 1998 Edition (ISO/IEC 15802-3:1998) and IEEE Std 802.1t-2001].

2.9 RPR 帧格式

RPR 有 4 种帧格式，包括：（1）基本数据帧（2）控制帧（3）公平算法帧（4）Idle 帧。

基本数据帧

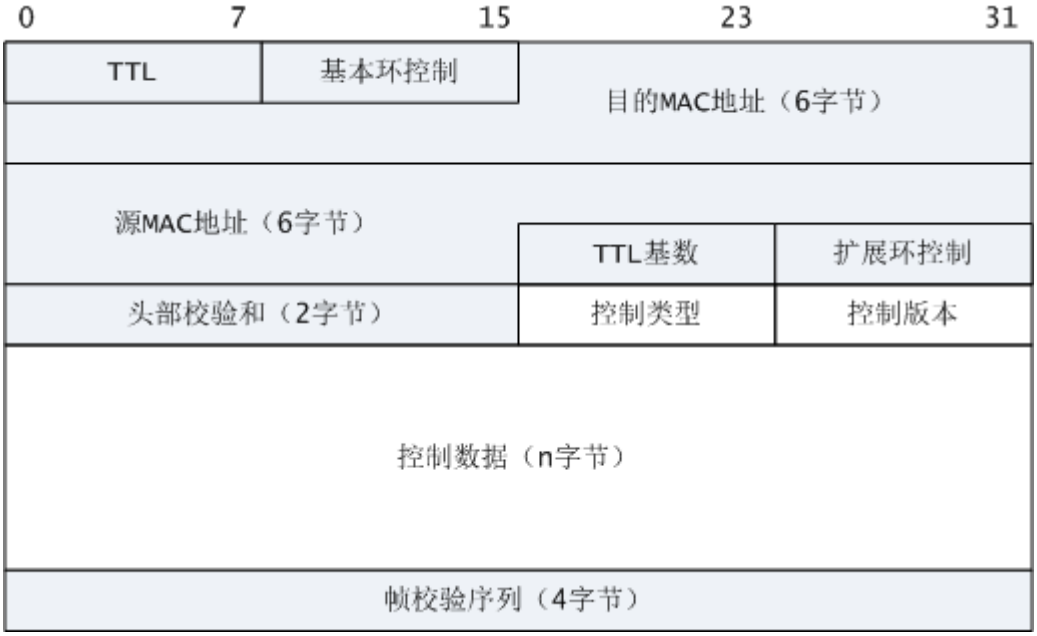


RPR 基本数据帧中几个主要字段的含义：

字段	长度（字节）	描述
TTL	1	TTL（Time to Live），TTL 的值决定数据帧在 RPR 环网上被转发的最大跳数。每经过一个节点，TTL 值被减 1，当 TTL 等于 0 时，数据帧被丢弃。这种机制可以防止数据帧在环网上被无限制的循环转发。
基本环控制	1	包含选环信息、公平算法控制信息、帧类型、流类型（服务级别）、Wrap 控制信息和奇偶校验位。
目的 MAC 地址	6	48 位目的 MAC 地址。对于单播报文，该字段在目的节点被删除；对于多播报文，该字段在源节点被删除。
源 MAC 地址	6	48 位源 MAC 地址。发送数据的源站点的 MAC 地址，它在传送过程中一直保持不变，用于目的节点回应消息的地址。
TTL 基数	1	TTL 的初始值。在数据帧转发过程中，TTL 基数保持不变。用 TTL 基数减去 TTL 值，可以得到数据帧转发到当前节点经过的跳数。
扩展环控制	1	RPR 报文扩展头，扩展帧标志、泛洪标志、泛洪类型、通过源节点标志、严格顺序帧标志和保留位。
HEC	2	HEC（Header Error Check），头部校验和，16 位 CRC（Cyclic Redundancy Check）校验。头部校验和是对 TTL、基本控制信息、目的 MAC 地址、源 MAC 地址、TTL 基数和扩展控制信息 6 个字段的数据进行计算而得到的。
协议类型	2	当此值小于 1536（十进制）时表示帧的长度。当此值大于等于 1536（十进制）时表示数据载荷的协议类型。“数据”字段承载的协议和“协议类型”字段的取值的对应关系如下： • 0x0800: IPv4 • 0x86dd: IPv6 • 0x8847: Tag Unicast • 0x8848: Tag Multicast • 0x0806: ARP (Address Resolution Protocol)

字段	长度（字节）	描述
用户数据	n	有效数据帧，此字段的长度是可变的。
FCS	4	FCS（Frame Check Sequence），帧校验序列，32 位 CRC 校验。是对协议类型和数据两个字段进行 CRC 校验得到的。

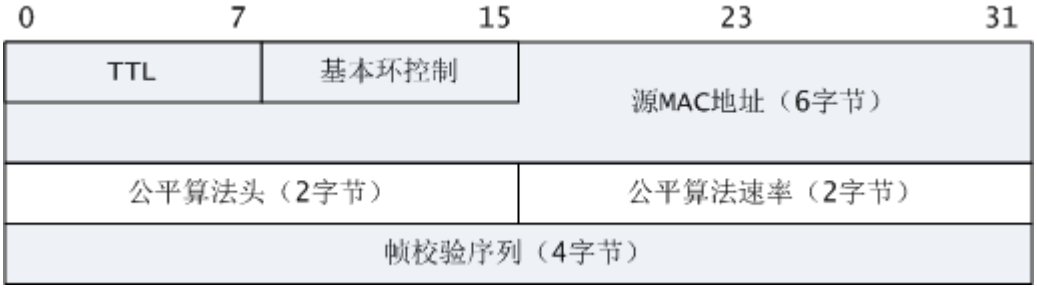
控制帧



控制类型”字段的取值决定控制帧的类型。控制帧有 10 种类型，控制帧的类型与“控制类型”字段的取值的对应关系：

- 0x01: Attribute Discovery（ATD），节点属性发现帧。
- 0x02: Topology and Protection packet（TP），拓扑与保护协议帧。
- 0x03: Topology Checksum（TC），拓扑校验帧。
- 0x04: Link round trip time measurement（LRTT） request，链路往返时间测量请求。
- 0x05: Link Round Trip Time measurement （LRTT） response，链路往返时间测量响应。
- 0x06: Fairness Differential Delay（FDD）
- 0x07: OAM echo request
- 0x08: OAM echo response
- 0x09: OAM flush
- 0x0A: OAM organization specific
- All others: Reserved

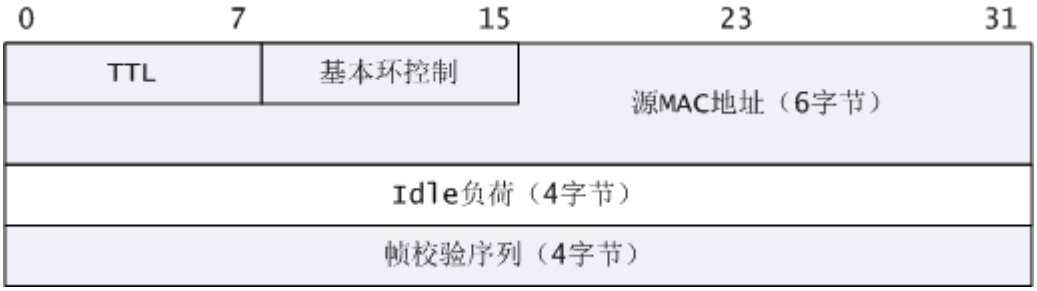
公平算法帧



公平算法帧格式中主要字段的解释如下：

- 源 MAC 地址：当前节点或者拥塞节点的 MAC 地址。
- 公平算法头：
  - Version＝0：Single Choke Fairness，每个公平算法周期发送一个，点对点传输。
  - Version＝1：Muti Choke Fairness，每 10 个公平算法周期发送一个，广播传输。
- FCS：32 位 CRC 校验。校验时不包含 TTL 和基本环控制字段。

Idle 帧



Idle 帧用于调整节点间的速率同步，在相邻的两个节点之间进行点对点传输。

Idle 负荷域默认为全 0。

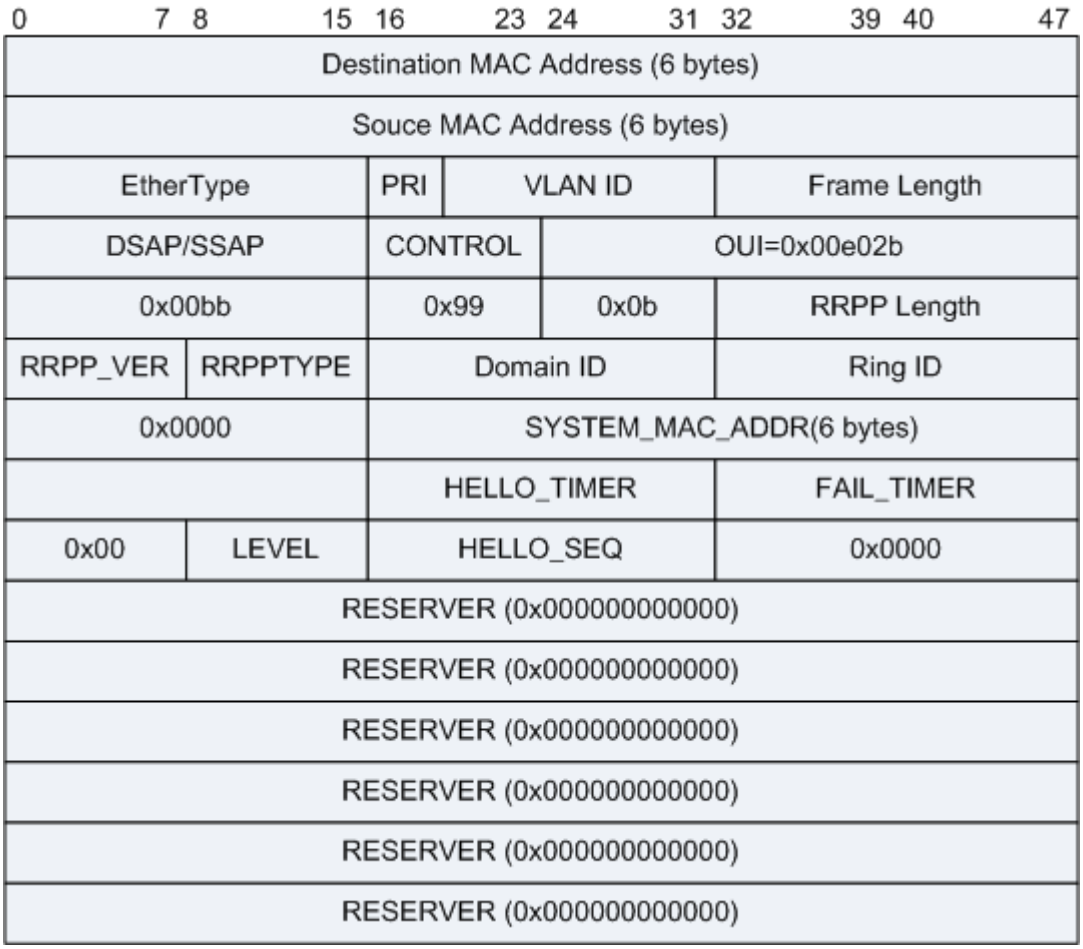
参考标准

标准	描述
IEEE P802.17-2004	IEEE Standard for information technology—Telecommunications and information exchange between systems— Local and metropolitan area networks—Specific requirements Part 17: Resilient packet ring (RPR) access method and physical layer specifications

2.10 RRPP 帧封装格式

帧格式

图 1 RRPP 帧格式



各域的说明如下：

字段	长度	说明
Destination MAC Address	48 比特	协议帧的目的 MAC。
Source Mac Address	48 比特	协议帧的源 MAC，固定值为 0x000fe203fd75。
EtherType	8 比特	帧封装类型域，固定值为 0x8100，表示 Tagged 封装。
PRI	4 比特	COS(Class of Service)优先级，固定值为 0xe0。
VLAN ID	12 比特	帧所属 VLAN 的 ID。
Frame Length	16 比特	以太网帧长度，固定值为 0x48。

字段	长度	说明
DSAP/SSAP	16 比特	目的服务访问点/源服务访问点，固定值为 0xaaaa。
CONTROL	8 比特	该字段无实际意义，固定值为 0x03。
OUI	24 比特	该字段无实际意义，固定值为 0x00e02b。
RRPP_LENGTH	16 比特	RRPP 协议数据单元长度，固定值为 0x40。
RRPP_VERS	8 比特	RRPP 版本信息，当前是 0x0001。
RRPP TYPE	8 比特	RRPP 帧类型： • 0x05 (HEALTH)：健康检测帧，由主节点发起，对网络进行环路检测。 • 0x06 (COMPLETE-FLUSH-FDB)：链路 DOWN 帧，由传输节点、边缘节点或者辅助边缘节点发起，通知主节点有端口 DOWN，环路消失。 • 0x07 (COMMON-FLUSH-FDB)：刷新 FDB（Forwarding Database）帧，由主节点发起，通知传输节点更新各自 MAC 地址转发表和 ARP 表。 • 0x08 (LINK-DOWN)： 由主节点发起，通知传输节点、边缘节点或者辅助边缘节点更新各自 MAC 地址转发表和 ARP 表。 同时通知传输节点解除临时阻塞数据 VLAN 的端口的阻塞状态。 • 0x0a (EDGE-HELLO)：主环完整性检查帧，由子环的边缘节点发起，同子环的辅助边缘节点接收，子环通过此帧检查其所在域主环的环路完整性。 • 0x0b (MAJOR-FAULT)：主环故障通知帧，当子环的辅助边缘节点在规定时间内收不到边缘节点发送的 EDGE-HELLO 帧时发起，向边缘节点报告其所在域主环发生故障。
DOMAIN_ID	16 比特	帧所属 RRPP 域的 ID。
RING_ID	16 比特	帧所属 RRPP 环的 ID。
SYSTEM_MAC_ADDR	48 比特	发送帧节点的桥 MAC。
HELLO_TIMER	16 比特	发送帧节点使用的 Hello 定时器的超时时间，单位为秒。
FAIL_TIMER	16 比特	发送帧节点使用的 Fail 定时器的超时时间，单位为秒。

字段	长度	说明
LEVEL	8 比特	帧所属 RRPP 环的级别。
HELLO_SEQ	16 比特	Hello 帧的序列号。

2.11 LACP 报文格式

报文格式

基于 IEEE802.3ad 标准的 LACP，链路汇聚控制协议是一种实现链路动态聚合与解聚合的协议。LACP 协议通过 LACPDU（Link Aggregation Control Protocol Data Unit）与对端交互信息。

LACPDU 报文为慢协议(平均每秒发送的协议报文不超过 5 个),如果接口板收到报文的 DMAC 是特殊的组播地址 0x01-80-c2-00-00-02，二层协议类型字段为 0x8809，协议子类型为 0x01，则说明此数据报文为 LACPDU 报文。

图 1 LACPDU 格式

Destination Address
Source Address
Length/Type
Subtype
Version Number
TLV_type
Actor_Information_Length
Actor_Port
Actor_State
Actor_System_Priority
Actor_System
Actor_key
Actor_Port_Priority
Reserved
Partner_Information_Length
Partner_Port
Partner_State
Partner_System_Priority
Partner_System
Partner_key
Partner_Port_Priority
Reserved
Collector_Information_Length
CollectorMaxDelay
Reserved
Terminator_Length
Reserved
FCS

报文中各域的说明如下：

字段	长度	说明
Destination Address	6字节	目的 MAC 地址，是一个组播地址（01-80-C2-00-00-02）

字段	长度	说明
Source Address	6 字节	源 MAC 地址，发送端口的 MAC 地址
Length/Type	2 字节	协议类型：0x8809
Subtype	1 字节	报文子类型：0x01，说明是 LACP 报文
Version Number	1 字节	协议版本号：0x01
TLV_type	1 字节	• 0x00 代表 Terminator 字段 • 0x01 代表 Actor 字段 • 0x02 代表 Partner 字段 • 0x03 代表 Collector 字段
Actor_Information_Length	1 字节	actor 信息字段长度，为 20 字节
Actor_Port	2 字节	端口号，根据算法生成，由接口所在的槽位号、子卡号和端口号决定
Actor_State	1 字节	本端状态信息： • LACP_Activity：代表链路所在的聚合组参与 LACP 协商的方式。主动的 LACP 被编码为 1，主动方式下会主动发送 LACPDU 报文给对方，被动方式不会主动发送协商报文，除非收到协商报文才会参与。

字段	长度	说明
		• LACP_Timeout：代表链路接收 LACPDU 报文的周期，有两种，快周期 1s 和慢周期 30s，超时时间为周期的 3 倍。短超时被编码为 1，长超时被编码为 0。 • Aggregation：标识该链路能否被聚合组聚合。如果编码为 0，该链路被认为是独立的，不能被聚合，即，这个链路只能作为一个个体链路运行。 • Synchronization：代表该链路是否已被分配到一个正确的链路聚合组，如果该链路已经关联了一个兼容的聚合器，那么该链路聚合组的识别与系统 ID 和被发送的运行 Key 信息是一致的。编码为 0，代表链路当前不在正确的聚合里。 • Collecting：帧的收集使能位，假如编码为 1，表示在这个链路上进来的帧的收集是明确使能的；即收集当前被使能，并且不期望在没有管理变化或接收协议信息变化的情况下被禁止。其它情况下这个值编码为 0。 • Distributing：帧的分配使能位，假如编码为 0，意味着在这个链路上的外出帧的分配被明确禁止，并且不期望在没有管理变化或接收协议信息变化的情况下被使能。其它情况下这个值编码为 1。 • Default：诊断调试时使用，编码为 1，代表接收到的对端的信息是管理配置的。假如编码为 0，正在使用的运行伙伴信息在接收到的 LACPDU 里。该值不被正常 LACP 协议使用，仅用于诊断协议问题。 • Expired：诊断调试时使用，编码为 1，代表本端的接收机是处于 EXPIRED 超时状态；假如编码为 0，本端接收状态机处于正常状态。该值不被正常 LACP 协议使用，仅用于诊断协议问题。
Actor_System_Priority	2 字节	本端系统优先级，可以设置，默认情况下为 32768
Actor_System	6 字节	系统 ID，本端系统的 MAC 地址
Actor_key	2 字节	端口 KEY 值，系统根据端口的配置生成，是端口能否成为聚合组中的一员的关 键因素，影响 Key 值得因素有 trunk ID、接口的速率和双工模式
Actor_Port_Priority	2 字节	接口优先级，可以配置，默认为 0x8000

字段	长度	说明
Reserved	3 字节	保留字段，可用于功能调试以及扩展
Partner_Information_Length	1 字节	Partner 信息字段长度。 Partner 字段代表了链路接口接收到对端的系统信息、接口信息和状态信息，与 actor 字段含义一致。在协商最开始未收到对端信息时，partner 字段填充 0，接收到对端信息后会把收到的对端信息补充到 partner 字段当中。
Partner_Port	2 字节	对端端口号
Partner_State	2 字节	对端状态信息
Partner_System_Priority	2 字节	对端系统优先级
Partner_System	6 字节	对端系统 ID，对端系统的 MAC 地址
Partner_key	2 字节	对端端口 KEY 值
Partner_Port_Priority	2 字节	对端接口优先级
Reserved	2 字节	保留字段
Collector_Information_Length	1 字	Collector 信息字段长度：0x10

字段	长度	说明
	节	
CollectorMaxDelay	2 字节	最大延时：默认情况下为 0xffff
Reserved	12 字节	保留字段
Terminator_Length	1 字节	Terminator 信息字段长度：0x00
Reserved	50 字节	保留字段，全置 0
FCS	4 字节	用于帧内后续字节差错的循环冗余检验（也称为 FCS 或帧检验序列）。

报文示例

```
⊞ Frame 1: 128 bytes on wire (1024 bits), 128 bytes captured (1024 bits)
⊞ Ethernet II, Src: HuaweiTe_93:e1:98 (28:6e:d4:93:e1:98), Dst: Slow-Protocol
⊟ Link Aggregation Control Protocol
    Slow Protocols subtype: LACP (0x01)
    LACP Version Number: 0x01
    Actor Information: 0x01
    Actor Information Length: 0x14
    Actor System Priority: 1
    Actor System: HuaweiTe_93:e1:98 (28:6e:d4:93:e1:98)
    Actor Key: 6449
    Actor Port Priority: 100
    Actor Port: 260
  ⊟ Actor State: 0xc7 (Activity, Timeout, Aggregation, Defaulted, Expired)
      .... ...1 = LACP Activity: Yes
      .... ..1. = LACP Timeout: Yes
      .... .1.. = Aggregation: Yes
      .... 0... = Synchronization: No
      ...0 .... = Collecting: No
      ..0. .... = Distributing: No
      .1.. .... = Defaulted: Yes
      1... .... = Expired: Yes
    Reserved: 000000
    Partner Information: 0x02
    Partner Information Length: 0x14
    Partner System Priority: 0
    Partner System: 00:00:00_00:00:00 (00:00:00:00:00:00)
    Partner Key: 0
    Partner Port Priority: 0
    Partner Port: 0
  ⊞ Partner State: 0xc7 (Activity, Timeout, Aggregation, Defaulted, Expired)
    Reserved: 000000
    Collector Information: 0x03
    Collector Information Length: 0x10
    Collector Max Delay: 0
    Reserved: 00000000000000000000000000
    Terminator Information: 0x00
    Terminator Length: 0x00
    Reserved: 000000000000000000000000000000000000000000000000000...
```

参考标准

标准	描述
IEEE 802.3ad	Link Aggregation Control Protocol

2.12 以太 OAM 报文格式

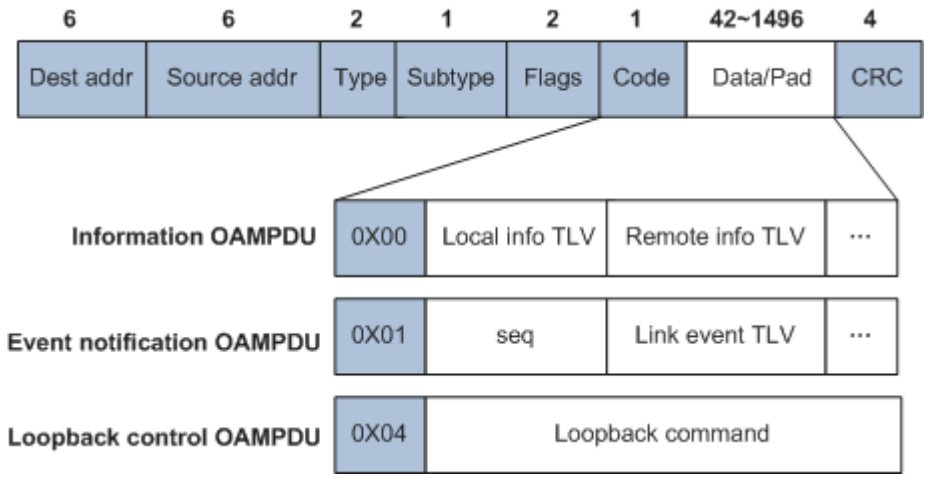
以太网 OAM 技术是分级实现，分为以下两个级别：

- 1）链路级以太网 OAM 技术：多应用于网络的 PE 设备-CE 设备-用户设备之间（也叫最后一公里）的以太网物理链路，用于监测用户网络与运营商网络之间的链路状态，典型协议为 EFM OAM 协议，参考标准 802.3ah。
- 2）网络级以太网 OAM 技术：多应用于网络的接入汇聚层，用于监测整个网络的连通性、定位网络的连通性故障，典型协议为 CFD（Connectivity Fault Detection）协议，参考标准为 802.1ag 和 Y.1731。

EFM（802.3ah）报文格式

EFM 工作在数据链路层，其协议报文被称为 OAM PDU。EFM 通过设备之间定时交互 OAM PDU 来报告链路状态，使网络管理员能够对网络进行有效的管理。下图为 OAMPDU 的格式及其几种常见的 OAMPDU。

图 1 EFM OAM PDU 格式



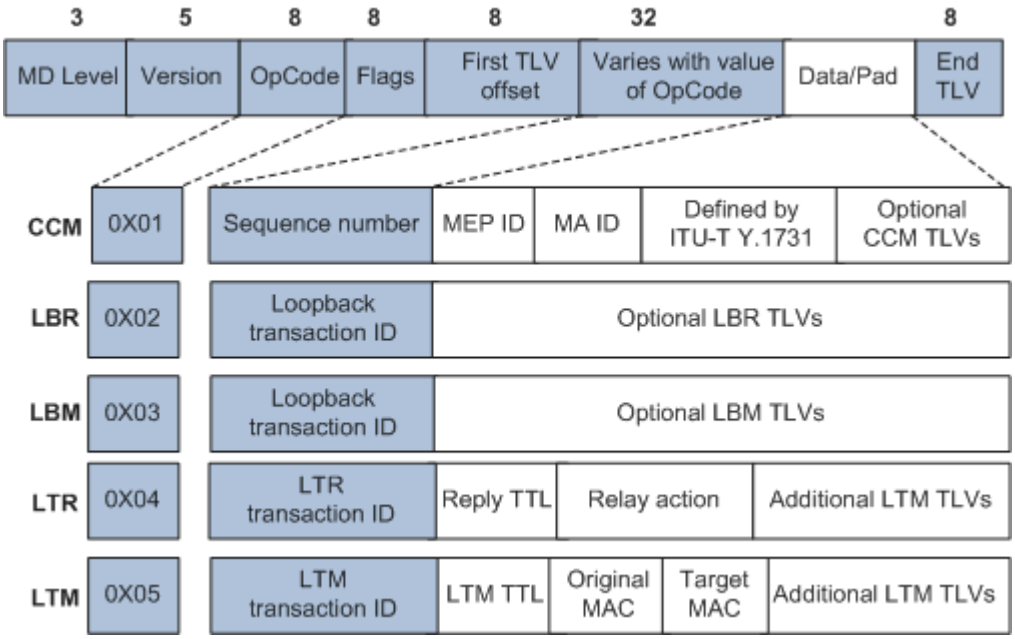
字段	长度	含义
Destination Address (DA)	6 字节	The DA in OAMPDUs is the Slow_Protocols_Multicast address. 0x0180-C200-0002.目的 MAC 地址，为慢速协议组播地址：0x0180-C200-0002。慢速协议报文的特点就是不能被网桥转发，因此无论是否具备 OAM 功能或 OAM 功能是否激活，EFM OAMPDU 都不能跨多跳转发。
Source Address (SA)	6 字节	源 MAC 地址，为发送端的端口 MAC 地址（若没有则采用该设备的桥 MAC 地址），是一个单播 MAC 地址。
Length/Type	2 字节	取值为 0x8809，用于标识慢协议类型。
Subtype	1 字节	用于标识慢协议的子协议类型，此处取值为 0x03，用于标识 EFM 协议。
Flags	2 字节	OAMPDU. EFM 实体的状态信息： • Remote Stable • Remote Evaluating • Local Stable • Local Evaluating • Critical Event • Dying Gasp • Link Fault
Code	1 字节	消息编码，不同取值表示不同类型的 OAMPDU： • 0x00: Information OAMPDU. 信息 OAMPDU，也称为心跳报文。用于在本端与远端的 OAM 实体之间交互各种状态信息（包括本地信息 TLV、远端信息 TLV 和组织自定义信息 TLV）。

字段	长度	含义
		• 0x01: Event Notification OAMPDU 事件通知 OAMPDU，用于对连接本端与远端 OAM 实体的链路上所发生的故障进行告警。 • 0x04: Loopback Control OAMPDU 环回控制 OAMPDU，用于检测链路质量和定位链路故障，该报文中带有使能/去使能信息，用来开启/关闭远端环回功能。
Local info TLV	变长	本地信息 TLV。
Remote infor TLV	变长	对端信息 TLV。
seq	变长	序列号。
Link Event TLV	变长	• 0x01 错误信号事件(Errored Symbol Period Event):单位时间内的错误信号数量超过定义的阈值 • 0x02 错误帧事件(Errored Frame Event):单位时间内的错误帧数量超过定义的阈值 • 0x03 错误帧周期事件(Errored Frame Period Event):指定帧数 N 为周期，在收到 N 个帧的周期内错误帧数超过定义的阈值 • 0x04 错误帧秒数事件(Errored Frame Seconds Event):指定 M 秒数下有错误帧的秒数超过了定义的阈值
Command Description	变长	通过非以太网 OAM 协议报文的环回来检测链路故障。主动模式下的 OAM 实体向对端(远端)发送除 OAMPDU 以外的所有其它报文时，对端收到报文后不按其目的地址进行转发，而是将其按原路返回给本端。远端环回只有在以太网 OAM 连接建立之后才能实现。 • 0x01: Enable OAM Remote Loopback • 0x02: Disable OAM Remote Loopback • 0x00, 0x03-0xFF: 预留，OAM 客户端忽略此字段。
Data/Pad	变长	OAMPDU 数据和填充字段。
CRC	4 字节	校验字段。

CFM OAM (802.1ag)协议报文格式

CFM 是通过携带不同标记的 CFM 协议报文实现链路的故障检测和定位的。CFM 帧工作在数据链路层，Type = 0x8902。

图 2 CFM PDU 格式



字段	长度	含义
MD level	3 比特	维护域的级别，取值范围为 0～7，取值越大表示级别越高
Version	5 比特	协议版本号，为 0。
OpCode	8 比特	消息编码，不同取值表示不同类型的 CFM PDU，常见的 CFM PDU 如表 2 所示。
Flags	8 比特	Flag 域，该字段在不同类型的 CFM PDU 中表示不同的含义。
Sequence number	8 比特	序列号，初始值为一个随机值，以后维护端点每发送一个 CCM PDU，该字段的取值就会加 1。
Loopback transaction ID	32 比特	处理编号，初始值为 0，以后维护端点每发送一个 LBR/LBM PDU，该字段的取值就会加 1。
LTR transaction ID	32 比特	处理编号，初始值为 0，以后维护端点每发送一个 LTR PDU，该字段的取值就会加 1。
LTM	32 比特	处理编号，初始值为 0，以后维护端点每发送一个 LTM PDU，该字段的取值就会加 1。

字段	长度	含义
transaction ID		
TLV(Type, Length, Value)	变长	• 0: End TLV 终了 TLV，长度和数值字段都不用。 • 1: Sender ID TLV • 1: Port Status TLV • 1: Data TLV • 1: Interface Status TLV • 1: Reply Ingress TLV • 1: Reply Egress TLV • 1: LTM Egress Identifier TLV • 1: LTR Egress Identifier TLV • 9-30: Reserved for IEEE 802.1 • 31: Organization-Specific TLV • 32: Test TLV, Defined by ITU-T Y.1731 • 33-63: Reserved for ITU-T Y.1731 • 64-255: Reserved for IEEE 802.1

表 1 消息编码与 PDU 类型

OpCode	PDU 类型	目的 MAC 地址	说明
0x01	CCM（Continuity Check Message）连续性检测报文	01-80-C2-00-00-3x（组播地址）	用于连续性检测，各维护端点均可发出。x 的取值： MD level x 的取值 y 的取值 7 7 F 6 6 E 5 5 D 4 4 C 3 3 B 2 2 A 1 1 9 0 0 8

字段	长度	含义	
0x02	LBR (Loopback Reply) 环回应答	环回发起端的 MAC（单播地址）	用于环回，由环回对端回应。
0x03	LBM (Loopback Message) 环回消息	环回目的端的 MAC（单播地址）	用于环回，由环回发起端发出。
0x04	LTR (Linktrace Reply) 链路跟踪应答	链路跟踪发起端的 MAC（单播地址）	用于链路跟踪，由链路跟踪对端回应。
0x05	LTM (Linktrace Message) 链路跟踪消息	01-80-C2-00-00-3y（组播地址）	用于链路跟踪，由链路跟踪发起端发出。y 的取值： MD level x 的取值 y 的取值 7 7 F 6 6 E 5 5 D 4 4 C 3 3 B 2 2 A 1 1 9 0 0 8
0, 6-31	预留给 IEEE 802.1	-	-
32-63	由 ITU-T Y.1731 定义： • 33: AIS • 35: LCK • 37: TST • 39: APS • 41: MCC • 43: LMM • 42: LMR • 45: 1DM • 47: DMM • 46: DMR	-	-
64-255	预留给 IEEE 802.1	-	-

参考标准

文档	描述
IEEE Std 802.3ah-2004	Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment: Media Access Control Parameters, Physical Layers, and Management Parameters for Subscriber Access Networks
IEEE Std 802.1ag-2007	IEEE Standard for Local and metropolitan area networks-Virtual Bridged Local Area Networks Amendment 5: Connectivity Fault Management
IEEE 802.1ag/Draft7.0	Virtual Bridged Local Area Networks— Amendment 5: Connectivity Fault Management
ITU-T Y.1731 (02/2008)	Y.1731 is an OAM protocol organized by the ITU-T. It covers not only the contents defined by IEEE802.1ag but also combinations of OAM messages, including the Alarm Indication Signal (AIS), Remote Defect Indication (RDI), Locked Signal (LCK), Test Signal, Automatic Protection Switching (APS), Maintenance Communication Channel (MCC), Experimental (EXP), and Vendor Specific (VSP) for fault management and performance monitoring, such as frame loss measurement (LM) and delay measurement (DM).
IITU-T G.8013/Y.1731 Amendment 1 (05/2012)	

2.13 ERPS 帧格式

ERPS 协议的报文只有一种，即 RAPS（Ring Auto Protection Switching） PDU（Protocol Data Unit）报文，RAPS PDU 报文包含 ERPS 环信息，在 ERPS 环上传递以实现各设备端口信息的互通。

ERPS 工作在数据链路层，Ethernet Type = 0x8902。

图 1 RAPS PDU 基本格式

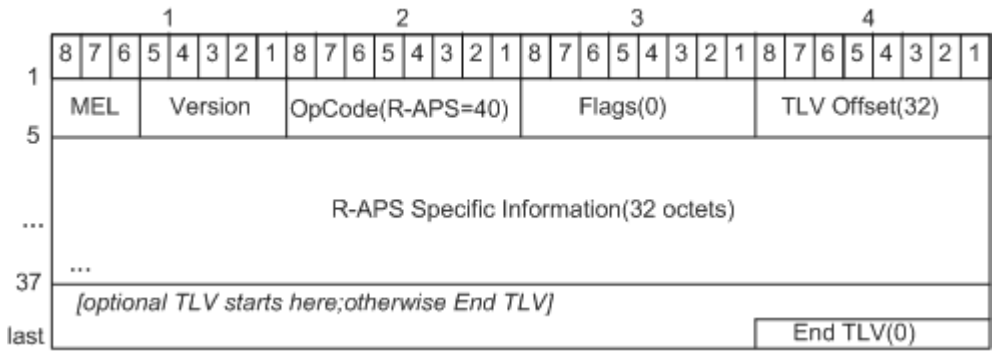


表 1 RAPS PDU 字段含义

字段名称	长度	说明

表 1 RAPS PDU 字段含义

字段名称	长度	说明
MEL（Maintenance Entity Group Level）	3 位	标识维护实例等级。
Version	5 位	• 0x00：v1 版本 • 0x01：v2 版本
OpCode	8 位	固定值 0x28，标识该 PDU 是 RAPS PDU。
Flags	8 位	固定值 0x00，该字段在接收的过程中会被忽略。
TLV Offset	8 位	固定值 0x20。
R-APS Specific Information	32x8 位	该字段携带 RAPS 环信息，是 RAPS PDU 的核心字段。对于该字段，v1 版本和 v2 版本在某些子字段的定义上存在一定的差异。图 2 描述 ERPSv1 版本该字段具体包含的各子字段，图 3 描述 ERPSv2 版本该字段具体包含的各子字段。
TLV（type-length-value）	无限制	描述报文中需要加载的信息，其中 End TLV 是固定值 0x00。

图 2 ERPSv1 版本 RAPS Specific Information 格式

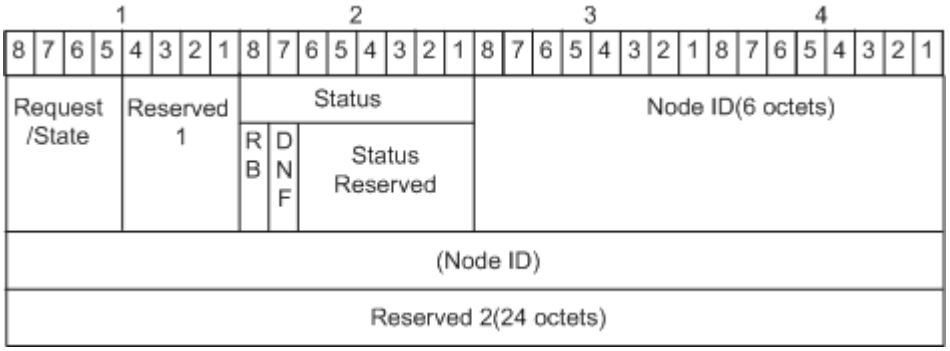


图 3 ERPSv2 版本 RAPS Specific Information 格式

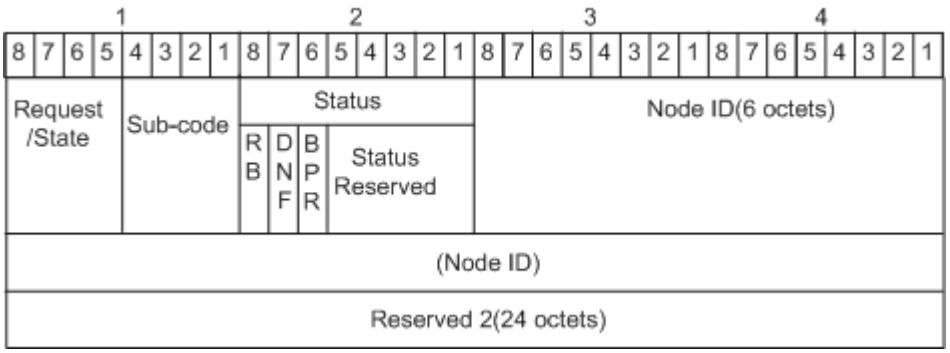


表 2 R-APS Specific Information 各子字段含义

字段名称	长度	说明
Request/State	4 位	标识该信息是请求信息或当前状态信息： • 1101：FS（Forced Switch） RAPS • 1110：Event 报文 • 1011：SF（Signal Fail） RAPS • 0111：MS（Manual Switch） RAPS • 0000：NR（No Request） RAPS • 其他：保留字段
Reserved 1 Sub-code	4 位	对于 ERPSv1，该字段是“Reserved 1”，表示保留字段，留作以后报文应答或是保护类型标识。 对于 ERPSv2，该字段是“Sub-code”： • 当“Request/State”字段的取值为 1110 时，该字段为 0000 表示 FDB 表项刷新请求。 • 当“Request/State”字段取其他值时，该字段的取值为全 0，为保留字段，且在接收过程中会被忽略。
Status	8 位	标识状态信息： • RB（RPL Blocked，1 位）：RB=1 标识 RPL 链路被阻塞；RB=0 标识 RPL 链路解除阻塞。非 RPL Owner 设备在发送 RPL PDU 时将该字段置为 0。 • DNF（Do Not Flush，1 位）：DNF= • BPR（Blocked port reference，1 位）：阻塞端口标志位，该字段为 0 表示阻塞第一个端口，该字段为 1 表示阻塞第二个端口。 只有 ERPSv2 版本支持该字段。 • Status Reserved：保留字段。在发送过程中，此字段全置为 0，且在接收的过程中会被忽略。该字段在 ERPSv1 版本有 6 位，在 ERPSv2 版本有 5 位。
Node ID	6x8 位	标识 RAPS 环节点的 MAC 地址，该字段属于提示信息，不影响 RAPS 环的保护切换操作。
Reserved 2	24x8 位	保留字段，在发送过程中，此字段全置为 0，且在接收的过程中会被忽略。

参考标准

标准	描述
ITU-T G.8032/Y1344	Recommendation ITU-T G.8032/Y.1344 defines the automatic protection switching (APS) protocol and protection switching mechanisms for ETH layer Ethernet ring topologies. Included are details

标准	描述
	pertaining to Ethernet ring protection characteristics, architectures and the ring APS protocol.

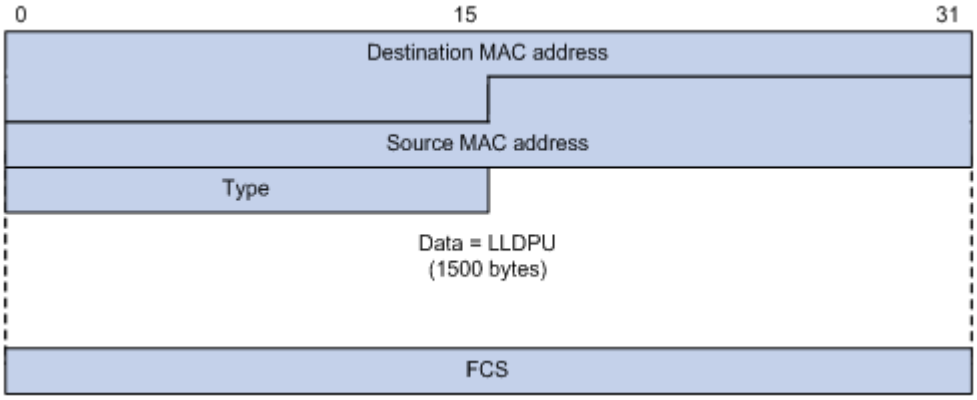
2.14 LLDP 报文格式

LLDP（Link Layer Discovery Protocol，链路层发现协议）提供了一种标准的链路层发现方式，可以将本端设备的的主要能力、管理地址、设备标识、接口标识等信息组织成不同的 TLV（Type/Length/Value），并封装在 LLDPDU（Link Layer Discovery Protocol Data Unit，链路层发现协议数据单元）中发布给与自己直连的邻居，邻居收到这些信息后将其以标准 MIB（Management Information Base，管理信息库）的形式保存起来，以供网络管理系统查询及判断链路的通信状况。

报文格式

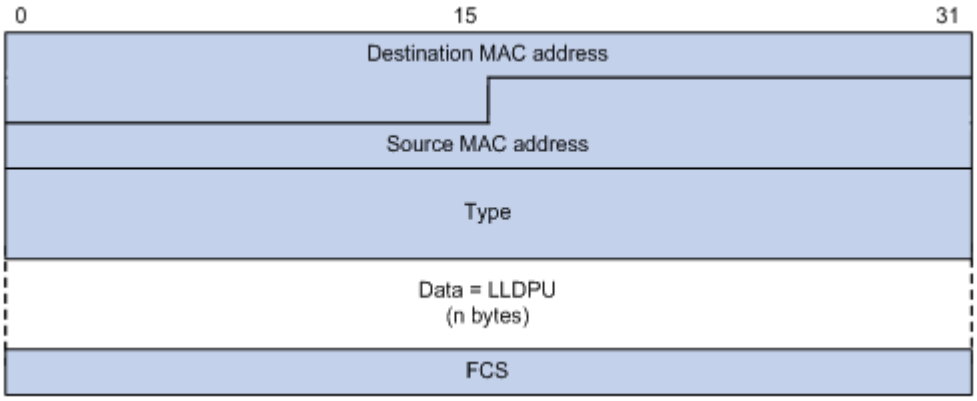
封装有 LLDPDU 的报文称为 LLDP 报文，其封装格式有两种：Ethernet II 和 SNAP（Subnetwork Access Protocol，子网访问协议）。

图 1 Ethernet II 格式封装的 LLDP 报文



- Destination MAC address：目的 MAC 地址，为固定的组播 MAC 地址 0x0180-C200-000E。
- Source MAC address：源 MAC 地址，为端口 MAC 地址或设备桥 MAC 地址（如果有端口地址则使用端口 MAC 地址，否则使用设备桥 MAC 地址）。
- Type：报文类型，为 0x88CC。
- Data：数据，为 LLDPDU。
- FCS：帧检验序列。

图 2 SNAP 格式封装的 LLDP 报文



- Destination MAC address：目的 MAC 地址，为固定的组播 MAC 地址 0x0180-C200-000E。
- Source MAC address：源 MAC 地址，为端口 MAC 地址或设备桥 MAC 地址（如果有端口地址则使用端口 MAC 地址，否则使用设备桥 MAC 地址）。
- Type：报文类型，为 0xAAAA-0300-0000-88CC。
- Data：数据，为 LLDPDU。
- FCS：帧检验序列。

LLDPDU 就是封装在 LLDP 报文数据部分的数据单元。在组成 LLDPDU 之前，设备先将本地信息封装成 TLV 格式，再由若干个 TLV 组合成一个 LLDPDU 封装在 LLDP 报文的数据部分进行传送。

图 3 LLDPDU 格式

每个 LLDPDU 最多可携带 28 种 TLV，其中深蓝色的 Chasis ID TLV、Port ID TLV、TTL TLV 和 End TLV 这四种是必须携带的，其余的 TLV 则为可选携带。

表 1 基本 TLV

TLV 名称	说明	是否必须发布
End of LLDPDU	标识 LLDPDU 结束	是
Chassis ID	发送设备的桥 MAC 地址	是
Port ID	标识 LLDPDU 发送端的端口。当设备不发送 MED TLV 时，内容为端口名称；当设备发送 MED TLV 时，内容为端口的 MAC 地址，没有端口 MAC 时使用桥 MAC	是
Time To Live	本设备信息在邻居设备上的存活时间	是

表 1 基本 TLV

TLV 名称	说明	是否必须发布
Port Description	以太网端口的描述字符串	否
System Name	设备的名称	否
System Description	系统描述	否
System Capabilities	系统的主要功能以及已使能的功能项	否
Management Address	管理地址，以及对应的接口号和 OID（Object Identifier，对象标识）	否

表 2 IEEE 802.1 组织定义的 TLV

TLV 名称	说明
Port VLAN ID	端口的 VLAN ID
Port And Protocol VLAN ID	端口的协议 VLAN ID
VLAN Name	端口 VLAN 的名称
Protocol Identity	端口支持的协议类型

表 3 IEEE 802.3 组织定义的 TLV

TLV 名称	说明
MAC/PHY	端口的速率和双工状态、是否支持端口速率自动协商、是否已使能自动协商功能以及当前的速率和双工

表 3 IEEE 802.3 组织定义的 TLV

TLV 名称	说明
Configuration/Status	状态
Power Via MDI	端口的供电能力
Link Aggregation	端口是否支持链路聚合以及是否已使能链路聚合
Maximum Frame Size	端口支持的最大帧长度，取端口配置的 MTU（Max Transmission Unit，最大传输单元）

表 4 LLDP-MED TLV

TLV 名称	说明
LLDP-MED Capabilities	当前设备的 MED 设备类型以及在 LLDPDU 中可封装的 LLDP-MED TLV 类型
Network Policy	端口的 VLAN ID、支持的应用（如语音和视频）、应用优先级以及使用策略等
Extended Power-via-MDI	当前设备的供电能力
Hardware Revision	MED 设备的硬件版本
Firmware Revision	MED 设备的固件版本
Software Revision	MED 设备的软件版本
Serial Number	MED 设备的序列号
Manufacturer Name	MED 设备的制造厂商
Model Name	MED 设备的模块名
Asset ID	MED 设备的资产标识符，以便目录管理和资产跟踪
Location	位置标识信息，供其它设备在基于位置的应用中使用

表 3 IEEE 802.3 组织定义的 TLV

TLV 名称	说明
Identification	

LLDP-MED TLV 为 VoIP（Voice over IP，在 IP 上传送语音）提供了许多高级的应用，包括基本配置、网络策略配置、地址信息以及目录管理等，满足了语音设备的不同生产厂商在成本有效、易部署、易管理等方面的要求，并解决了在以太网中部署语音设备的问题，为语音设备的生产者、销售者以及使用者提供了便利。

报文示例

```
⊞ Frame 27: 341 bytes on wire (2728 bits), 341 bytes captured (2728 bits)
⊞ Ethernet II, Src: e0:24:7f:b6:ea:4d (e0:24:7f:b6:ea:4d), Dst: LLDP_Multicast
⊟ Link Layer Discovery Protocol
  ⊟ Chassis Subtype = MAC address, Id: e0:24:7f:b6:ea:4d
      0000 001. .... .... = TLV Type: Chassis Id (1)
      .... ...0 0000 0111 = TLV Length: 7
      Chassis Id Subtype: MAC address (4)
      Chassis Id: e0:24:7f:b6:ea:4d (e0:24:7f:b6:ea:4d)
  ⊟ Port Subtype = Interface name, Id: GigabitEthernet1/0/7
      0000 010. .... .... = TLV Type: Port Id (2)
      .... ...0 0001 0101 = TLV Length: 21
      Port Id Subtype: Interface name (5)
      Port Id: GigabitEthernet1/0/7
  ⊟ Time To Live = 120 sec
      0000 011. .... .... = TLV Type: Time to Live (3)
      .... ...0 0000 0010 = TLV Length: 2
      Seconds: 120
  ⊟ System Name = PE_MS_TB3
      0000 101. .... .... = TLV Type: System Name (5)
      .... ...0 0000 1001 = TLV Length: 9
      System Name = PE_MS_TB3
  ⊟ System Description = Huawei Versatile Routing Platform Software\r\nVRP (R)
      0000 110. .... .... = TLV Type: System Description (6)
      .... ...0 1011 0011 = TLV Length: 179
      System Description = Huawei Versatile Routing Platform Software\r\nVRP (
  ⊟ Port Description = HUAWEI, Quidway Series, GigabitEthernet1/0/7 Interface
      0000 100. .... .... = TLV Type: Port Description (4)
      .... ...0 0011 0110 = TLV Length: 54
      Port Description: HUAWEI, Quidway Series, GigabitEthernet1/0/7 Interface
  ⊟ Capabilities
      0000 111. .... .... = TLV Type: System Capabilities (7)
      .... ...0 0000 0100 = TLV Length: 4
    ⊟ Capabilities: 0x0014
        .... .... .... .1.. = Bridge
        .... .... ...1 .... = Router
    ⊟ Enabled Capabilities: 0x0014
        .... .... .... .1.. = Bridge
        .... .... ...1 .... = Router
  ⊟ Management Address
      0001 000. .... .... = TLV Type: Management Address (8)
      .... ...0 0001 1101 = TLV Length: 29
      Address String Length: 5
      Address Subtype: IPv4 (1)
      Management Address: 10.133.113.1 (10.133.113.1)
      Interface Subtype: ifIndex (2)
      Interface Number: 34
      OID String Length: 18
      Object Identifier: 060f2b060104018f5b051929010201010100
  ⊟ End of LLDPDU
      0000 000. .... .... = TLV Type: End of LLDPDU (0)
      .... ...0 0000 0000 = TLV Length: 0
```

参考标准

标准	描述
IEEE 802.1AB	IEEE Standard for Local and Metropolitan Area Networks: Station and Media Access Control Connectivity Discovery

2.15 IS-IS 报文格式

- IS-IS 报文通用格式
- IS-IS Hello 消息格式
- IS-IS LSP 消息格式
- IS-IS SNP 消息格式

父主题： 链路层

2.15.1 IS-IS 报文通用格式

报文格式

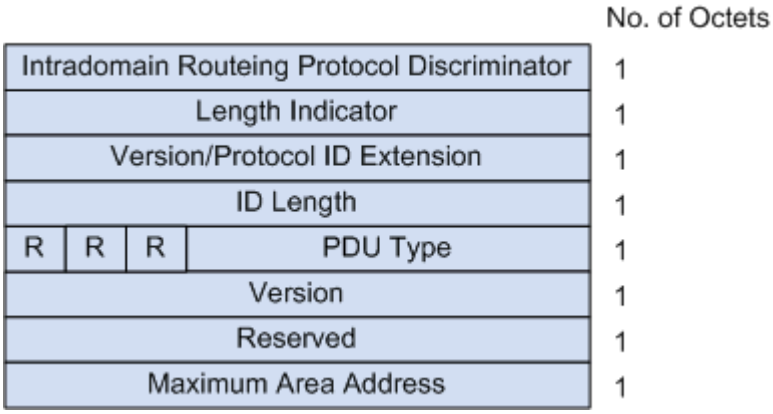
IS-IS 报文是直接封装在数据链路层的帧结构中的。PDU 可以分为两个部分，报文头和变长字段部分。其中头部又可分为通用头部和专用头部。对于所有 PDU 来说，通用报头都是相同的，但专用报头根据 PDU 类型不同而有所差别。

IS-IS 的 PDU 有 4 种类型：Hello 报文，LSP，CSNP，PSNP。

图 1 IS-IS 的 PDU 格式

所有的 PDU 都有相同的通用头格式：

图 2 IS-IS 通用报头格式



- Intradomain Routing Protocol Discriminator：域内路由选择协议鉴别符，设置为 0x83。
- Length Indicator：PDU 头部的长度（包括通用头部和专用头部），以字节为单位。

- Version/Protocol ID Extension：版本/协议标识扩展，设置为 1（0x01）。
- ID Length：NSAP 地址或 NET 中 System ID 区域的长度。值为 0 时，表示 System ID 区域的长度为 6 字节。值为 255 时，表示 System ID 区域为空（即长度为 0）。
- R（Reserved）：保留，设置为 0。
- PDU Type：PDU 的类型。IS-IS PDU 共有 9 种类型，详细信息请参考下表。
- Version：设置为 1（0x01）。
- Maximum Area Address：支持的最大区域个数。设置为 1～254 的整数，表示该 IS-IS 进程实际所允许的最大区域地址数；设置为 0，表示该 IS-IS 进程最大只支持 3 个区域地址数。

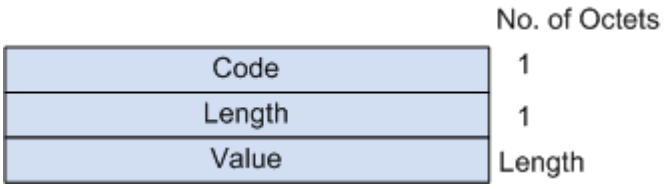
表 1 PDU 类型对应关系表

类型值	PDU 类型	简称
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

CLV 报文格式

PDU 中的变长字段部分是多个 CLV（Code-Length-Value）三元组。CLV 也称为 TLV（Type-Length-Value）。其格式如下图所示。

图 3 CLV 格式



不同 PDU 类型所包含的 CLV 是不同的。

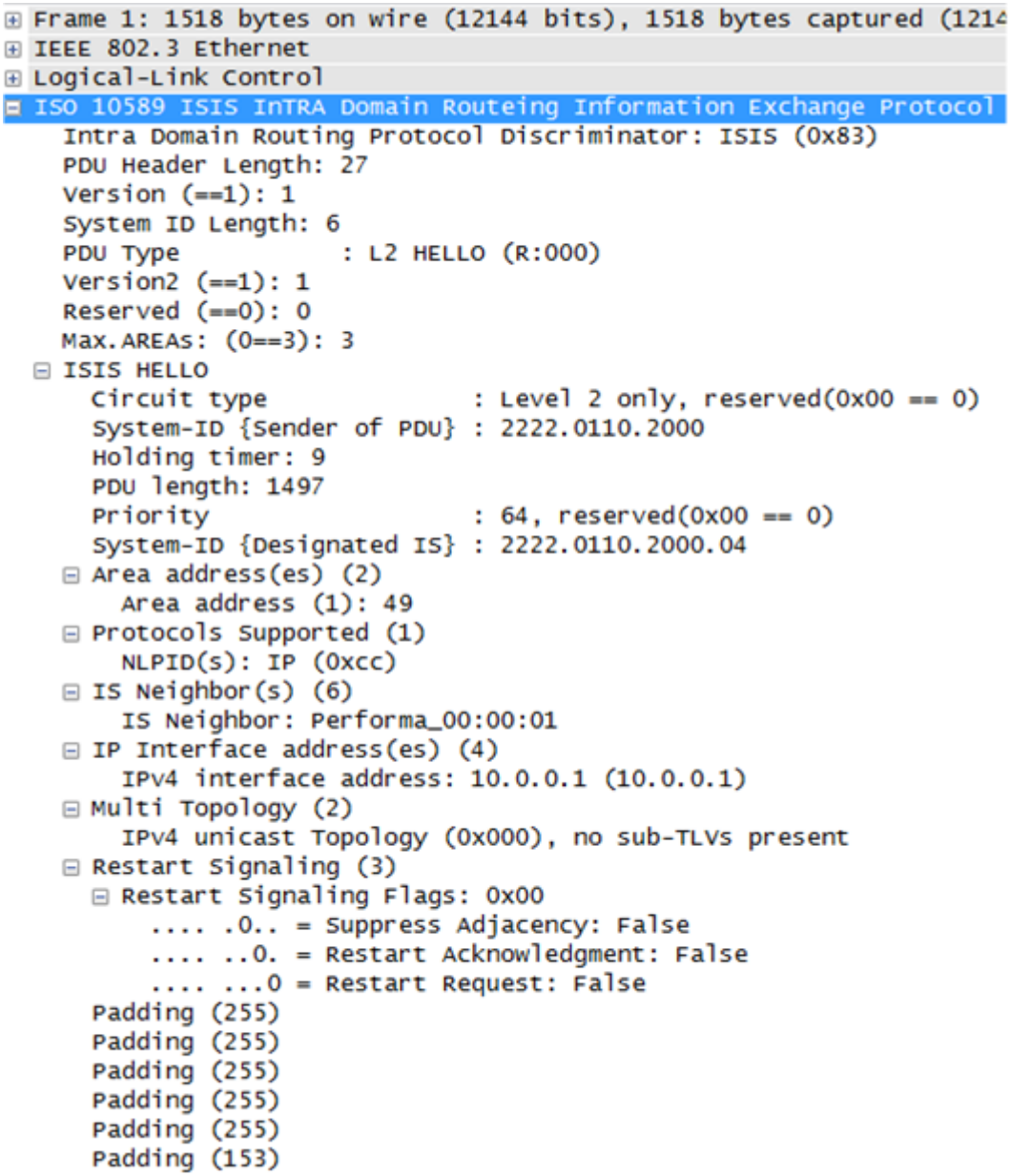
CLV Code	名称	所应用的 PDU 类型
1	Area Addresses	IIH、LSP
2	IS Neighbors (LSP)	LSP
4	Partition Designated Level2 IS	L2 LSP
6	IS Neighbors (MAC Address)	LAN IIH
7	IS Neighbors (SNPA Address)	LAN IIH
8	Padding	IIH
9	LSP Entries	SNP
10	Authentication Information	IIH、LSP、SNP
128	IP Internal Reachability Information	LSP
129	Protocols Supported	IIH、LSP
130	IP External Reachability	L2 LSP

CLV Code	名称	所应用的 PDU 类型
	Information	
131	Inter-Domain Routing Protocol Information	L2 LSP
132	IP Interface Address	IIH、LSP

其中，Code 值从 1 到 10 的 CLV 在 ISO10589 中定义（有 2 类未在上表中列出），其他几种 CLV 在 RFC1195 中定义。

报文示例

图 4 IS-IS 报文格式



参考标准

标准	描述
ISO 10589	ISO IS-IS Routing Protocol
RFC 1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments

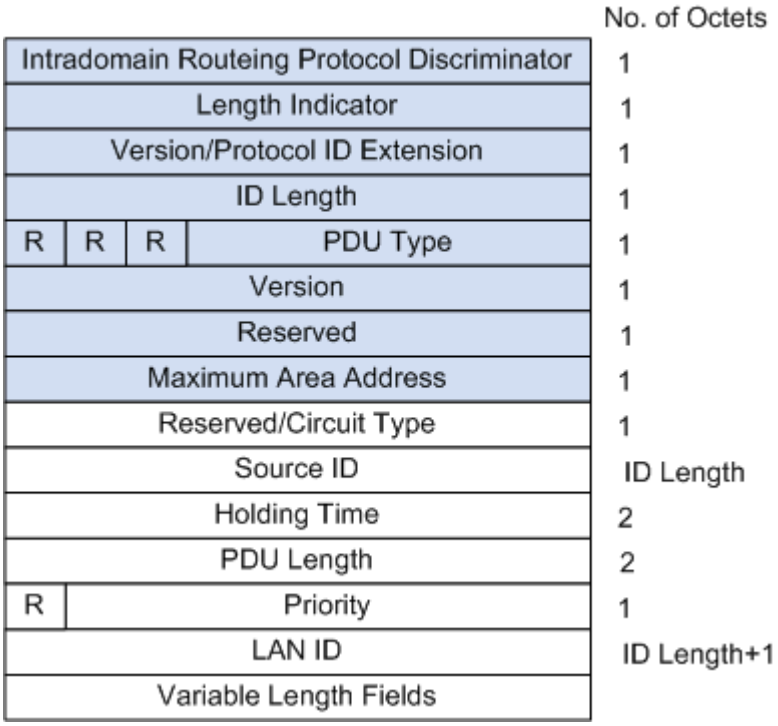
2.15.2 IS-IS Hello 消息格式

报文格式

Hello 报文用于建立和维持邻居关系，也称为 IIH（IS-to-IS Hello PDUs）。其中，广播网中的 Level-1 路由器使用 Level-1 LAN IIH；广播网中的 Level-2 路由器使用 Level-2 LAN IIH；非广播网络中则使用 P2P IIH。它们的报文格式有所不同。

广播网中的 Hello 报文格式如下图所示（蓝色部分是通用报文头）。

图 1 L1/L2 LAN IIH 格式



主要字段的解释如下：

- Reserved/Circuit Type：高位的 6 比特保留，值为 0。低位的 2 比特表示路由器的类型（01 表示 L1，10 表示 L2，11 表示 L1/L2）。
- Source ID：发出 Hello 报文的路由器的 System ID。
- Holding Time：保持时间。在此时间内如果没有收到邻居发来的 Hello 报文，则中止已建立的邻居关系。
- PDU Length：PDU 的总长度，单位是字节。

- Priority：选举 DIS 的优先级，取值范围为 0～127。数值越大，优先级越高。
- LAN ID：包括 DIS 的 System ID 和一字节的伪节点 ID。

点到点网络中的 Hello 报文格式如下图所示。

图 2 P2P IIH 格式

	No. of Octets
Intradomain Routeing Protocol Discriminator	1
Length Indicator	1
Version/Protocol ID Extension	1
ID Length	1
R R R PDU Type	1
Version	1
Reserved	1
Maximum Area Address	1
Reserved/Circuit Type	1
Source ID	ID Length
Holding Time	2
PDU Length	2
Local Circuit ID	1
Variable Length Fields	

从图中可以看出，P2P IIH 中的多数字段与 LAN IIH 相同。不同的是没有 Priority 和 LAN ID 字段，而多了一个 Local Circuit ID 字段，表示本地链路 ID。

报文示例

图 3 IS-IS Hello

```
⊞ Frame 1: 1518 bytes on wire (12144 bits), 1518 bytes captured (1214
⊞ IEEE 802.3 Ethernet
⊞ Logical-Link Control
⊟ ISO 10589 ISIS InTRA Domain Routeing Information Exchange Protocol
    Intra Domain Routing Protocol Discriminator: ISIS (0x83)
    PDU Header Length: 27
    Version (==1): 1
    System ID Length: 6
    PDU Type           : L2 HELLO (R:000)
    Version2 (==1): 1
    Reserved (==0): 0
    Max.AREAs: (0==3): 3
  ⊟ ISIS HELLO
      Circuit type              : Level 2 only, reserved(0x00 == 0)
      System-ID {Sender of PDU} : 2222.0110.2000
      Holding timer: 9
      PDU length: 1497
      Priority                  : 64, reserved(0x00 == 0)
      System-ID {Designated IS} : 2222.0110.2000.04
    ⊟ Area address(es) (2)
        Area address (1): 49
    ⊟ Protocols Supported (1)
        NLPID(s): IP (0xcc)
    ⊟ IS Neighbor(s) (6)
        IS Neighbor: Performa_00:00:01
    ⊟ IP Interface address(es) (4)
        IPv4 interface address: 10.0.0.1 (10.0.0.1)
    ⊟ Multi Topology (2)
        IPv4 unicast Topology (0x000), no sub-TLVs present
    ⊟ Restart Signaling (3)
      ⊟ Restart Signaling Flags: 0x00
          .... .0.. = Suppress Adjacency: False
          .... ..0. = Restart Acknowledgment: False
          .... ...0 = Restart Request: False
      Padding (255)
      Padding (255)
      Padding (255)
      Padding (255)
      Padding (255)
      Padding (153)
```

参考标准

标准	描述
ISO 10589	ISO IS-IS Routing Protocol
RFC 1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments

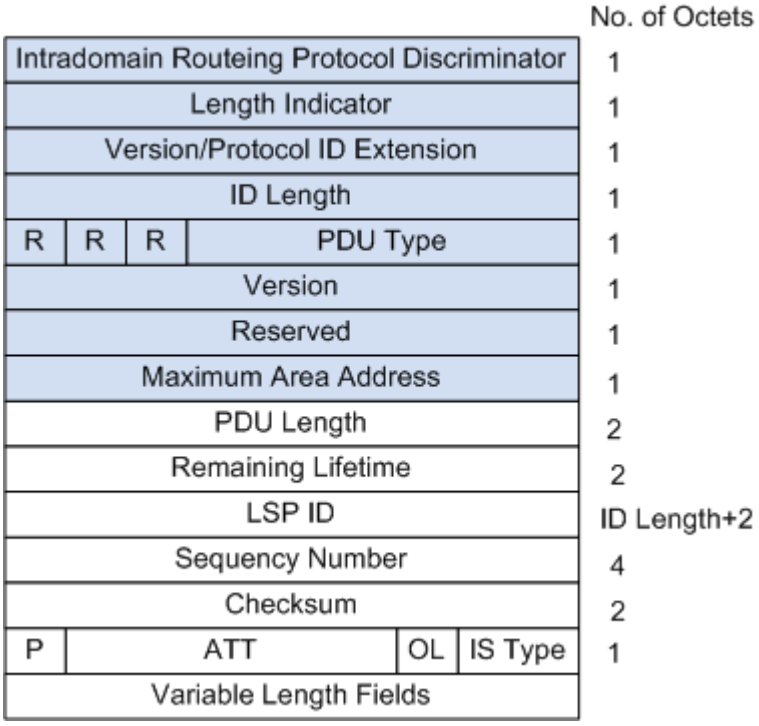
2.15.3 IS-IS LSP 消息格式

报文格式

链路状态报文 LSP（Link State PDUs）用于交换链路状态信息。LSP 分为两种：Level－1 LSP 和 Level－2 LSP。Level－1 LSP 由 Level-1 路由器传送，Level－2 LSP 由 Level-2 路由器传送，Level-1-2 路由器则可传送以上两种 LSP。

两类 LSP 有相同的报文格式。

图 1 L1/L2 LSP 格式



主要字段的解释如下：

- PDU Length：PDU 的总长度，以字节为单位。
- Remaining Lifetime：LSP 的生存时间，以秒为单位。
- LSP ID：由三部分组成，System ID、伪节点 ID（一字节）和 LSP 分片后的编号（一字节）。
- Sequency Number：LSP 的序列号。
- Checksum：LSP 的校验和。
- P（Partition Repair）：仅与 L2 LSP 有关，表示路由器是否支持自动修复区域分割。
- ATT（Attachment）：由 Level-1-2 路由器产生，用来指明始发路由器是否与其它区域相连。虽然此标志位也存在于 Level-1 和 Level-2 的 LSP 中，但实际上此字段只和 Level-1-2 路由器始发的 L1 LSP 有关。此字段有 4bit，用来表明相连的区域所使用的度量方式。

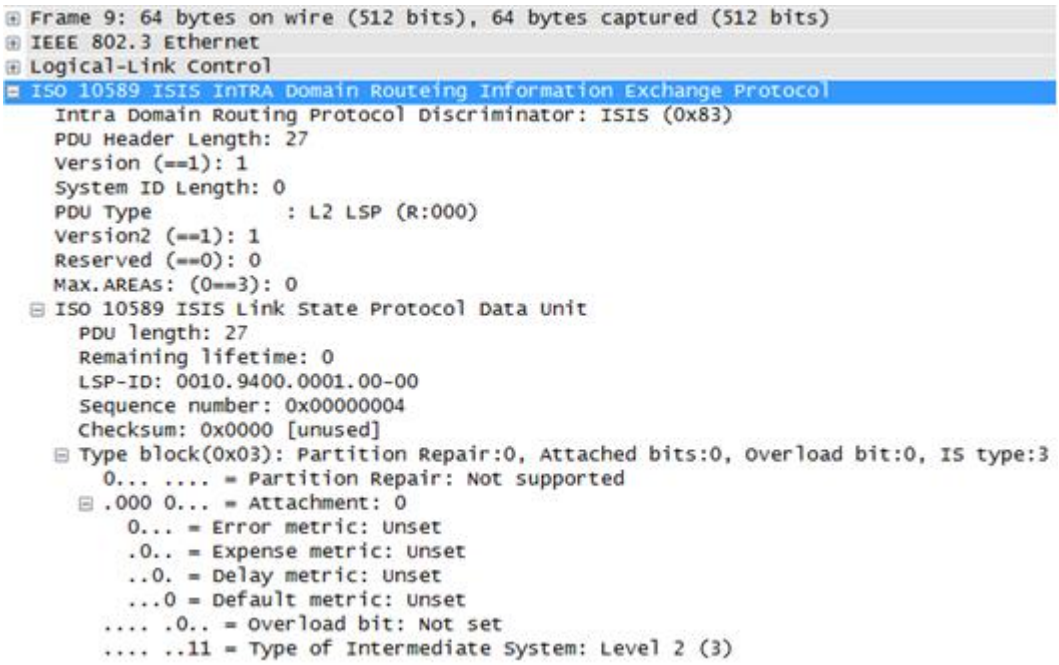
  从右至左这 4 位依次表示如下所示：

  - 第 4 位：缺省度量；
  - 第 5 位：时延度量；
  - 第 6 位：代价度量；
  - 第 7 位：差错度量。
- OL（LSDB Overload）：过载标志位。设置了过载标志位的 LSP 虽然还会在网络中扩散，但是在计算通过超载路由器的路由时不会被采用。即，对路由器设置过载位后，其它路由器在进行 SPF 计算时不会考虑这台路由器。当路由器内存不足时，系统自动在发送的 LSP 报文中设置过载标志位。

- IS Type：生成 LSP 的路由器的类型。用来指明是 Level-1 还是 Level-2 路由器（01 表示 Level-1，11 表示 Level-2）。

报文示例

图 2 IS-IS LSP 消息



参考标准

标准	描述
ISO 10589	ISO IS-IS Routing Protocol
RFC1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments

2.15.4 IS-IS SNP 消息格式

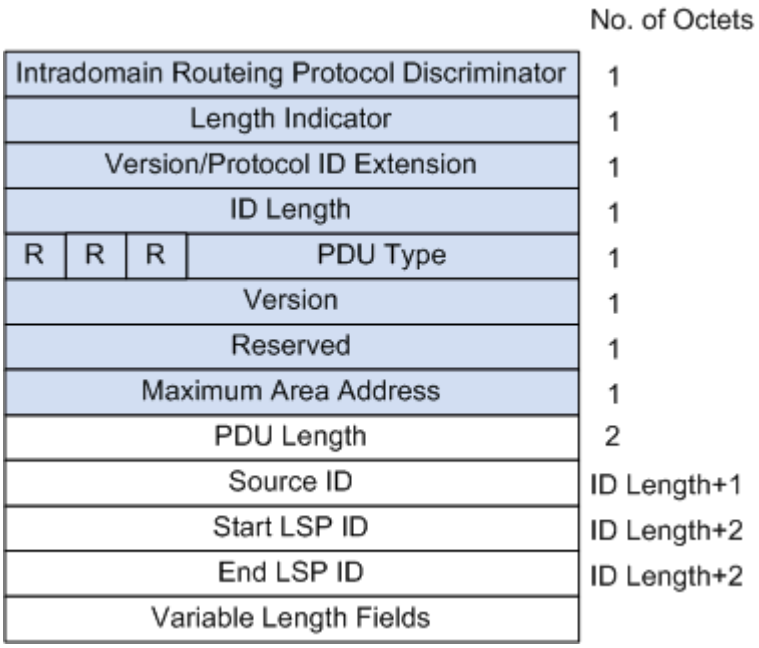
报文格式

时序报文 SNP（Sequence Number PDUs）通过描述全部或部分数据库中的 LSP 来同步各 LSDB（Link-State DataBase），从而维护 LSDB。

SNP 包括 CSNP（Complete SNP，全时序报文）和 PSNP（Partial SNP，部分时序报文），进一步又可分为 L1 CSNP、L2 CSNP、L1 PSNP 和 L2 PSNP。

CSNP 包括 LSDB 中所有 LSP 的摘要信息，从而可以在相邻路由器间保持 LSDB 的同步。在广播网络上，CSNP 由 DIS 定期发送（缺省的发送周期为 10 秒）；在点到点链路上，CSNP 只在第一次建立邻接关系时发送。

图 1 L1/L2 CSNP 消息格式

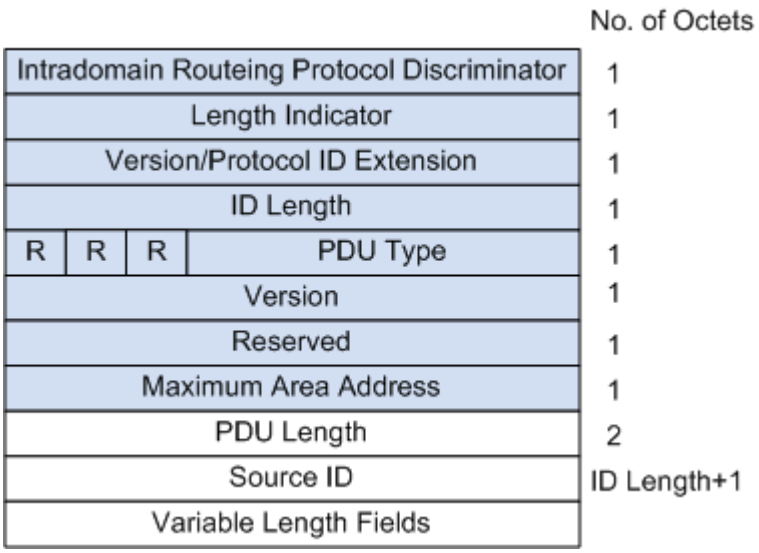


主要字段的解释如下：

- Source ID：发出 SNP 报文的路由器的 System ID。
- Start LSP ID：CSNP 报文中第一个 LSP 的 ID 值。
- End LSP ID：CSNP 报文中最后一个 LSP 的 ID 值。

PSNP 只列举最近收到的一个或多个 LSP 的序号，它能够一次对多个 LSP 进行确认，当发现 LSDB 不同步时，也用 PSNP 来请求邻居发送新的 LSP。

图 2 L1/L2 PSNP 格式



报文示例

图 3 IS-IS CSNP 消息

```
⊞ Frame 2: 168 bytes on wire (1344 bits), 168 bytes captured (1344 bits)
⊞ IEEE 802.3 Ethernet
⊞ Logical-Link Control
⊟ ISO 10589 ISIS InTRA Domain Routeing Information Exchange Protocol
    Intra Domain Routing Protocol Discriminator: ISIS (0x83)
    PDU Header Length: 33
    Version (==1): 1
    System ID Length: 6
    PDU Type           : L2 CSNP (R:000)
    Version2 (==1): 1
    Reserved (==0): 0
    Max.AREAs: (0==3): 3
  ⊟ ISO 10589 ISIS Complete Sequence Numbers Protocol Data Unit
      PDU length: 147
      Source-ID:    2222.0110.2000.00
      Start LSP-ID: 0000.0000.0000.00-00
      End LSP-ID: ffff.ffff.ffff.ff-ff
    ⊟ LSP entries (112)
      ⊟ LSP-ID: 0010.9400.0001.00-00, Sequence: 0x00000003, Lifetime:  1077s, Checksum: 0x388e
          LSP-ID:              : 0010.9400.0001.00-00
          LSP Sequence Number  : 0x00000003
          Remaining Lifetime   : 1077s
          LSP checksum         : 0x388e
      ⊟ LSP-ID: 0010.9400.0001.00-01, Sequence: 0x00000003, Lifetime:  1077s, Checksum: 0xfa2b
          LSP-ID:              : 0010.9400.0001.00-01
          LSP Sequence Number  : 0x00000003
          Remaining Lifetime   : 1077s
          LSP checksum         : 0xfa2b
      ⊞ LSP-ID: 0010.9400.0002.00-00, Sequence: 0x00000003, Lifetime:  1075s, Checksum: 0xd735
      ⊞ LSP-ID: 0010.9400.0002.00-01, Sequence: 0x00000003, Lifetime:  1076s, Checksum: 0xd928
      ⊞ LSP-ID: 2222.0110.2000.00-00, Sequence: 0x00000001, Lifetime:  1068s, Checksum: 0xed6f
      ⊞ LSP-ID: 2222.0110.2000.03-00, Sequence: 0x00000001, Lifetime:  1068s, Checksum: 0xfbef
      ⊞ LSP-ID: 2222.0110.2000.04-00, Sequence: 0x00000001, Lifetime:  1068s, Checksum: 0xdd0e
```

参考标准

标准	描述
ISO 10589	ISO IS-IS Routing Protocol
RFC 1195	Use of OSI IS-IS for Routing in TCP/IP and Dual Environments

3. MPLS 层

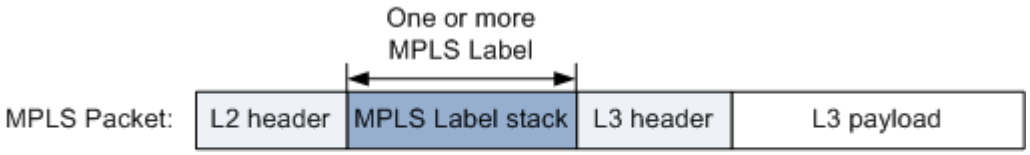
- MPLS 报文格式
- MPLS L2/L3VPN 报文结构
- MPLS Ping/Tracert (MPLS Echo)报文格式

3.1 MPLS 报文格式

报文格式

MPLS 标签（Label）是一个短而定长的、只具有本地意义的标识符，用于唯一标识一个分组所属的 FEC。在某些情况下，例如要进行负载分担，对应一个 FEC 可能会有多个入标签，但是一台路由器上，一个标签只能代表一个 FEC。标签与 ATM 的 VPI/VCI 以及 Frame Relay 的 DLCI 类似，是一种连接标识符。标签长度为 4 个字节，封装在链路层和网络层之间。这样，标签能够被任意的链路层所支持。标签在分组中的封装位置如下图所示。

图 1 MPLS 报文格式



字段	长度	含义
Label	20 比特	标签值字段，用来标识一个 FEC。
EXP	3 比特	用于扩展。现在通常用做 CoS（Class of Service），其作用与 Ethernet802.1p 的作用类似。
S	1 比特	MPLS 支持多重标签。值为 1 时表示为最底层标签。
TTL	8 比特	和 IP 分组中的 TTL 意义相同，可以用来防止环路。

报文示例

图 2 MPLS 报文（含 2 层标签）

```
⊞ Frame 1: 162 bytes on wire (1296 bits), 162 bytes captured (1296 bits)
⊟ Ethernet II (VLAN tagged), Src: HuaweiTe_25:fd:88 (08:19:a6:25:fd:88), Dst:
  ⊞ Destination: HuaweiTe_74:e4:08 (54:89:98:74:e4:08)
  ⊞ Source: HuaweiTe_25:fd:88 (08:19:a6:25:fd:88)
  ⊞ VLAN tag: VLAN=412, Priority=Voice, < 10ms latency and jitter
    Type: MPLS label switched packet (0x8847)
⊟ MultiProtocol Label Switching Header, Label: 1127, Exp: 0, S: 0, TTL: 255
    MPLS Label: 1127
    MPLS Experimental Bits: 0
    MPLS Bottom Of Label Stack: 0
    MPLS TTL: 255
⊟ MultiProtocol Label Switching Header, Label: 1098, Exp: 0, S: 1, TTL: 255
    MPLS Label: 1098
    MPLS Experimental Bits: 0
    MPLS Bottom Of Label Stack: 1
    MPLS TTL: 255
⊞ Internet Protocol Version 4, Src: 171.0.0.41 (171.0.0.41), Dst: 127.0.0.1 (
⊞ User Datagram Protocol, Src Port: 31015 (31015), Dst Port: lsp-ping (3503)
⊞ Multiprotocol Label Switching Echo
```

参考标准

标准	描述
RFC 3031	Multiprotocol Label Switching Architecture
RFC 3032	MPLS Label Stack Encoding

标准	描述
RFC 3034	Use of Label Switching on Frame Relay Networks Specification
RFC 3035	MPLS using LDP and ATM VC Switching
RFC 2547	BGP/MPLS VPNs

3.2 MPLS L2/L3VPN 报文结构

MPLS L2/L3VPN 报文结构

图 1 MPLS L3VPN 报文结构

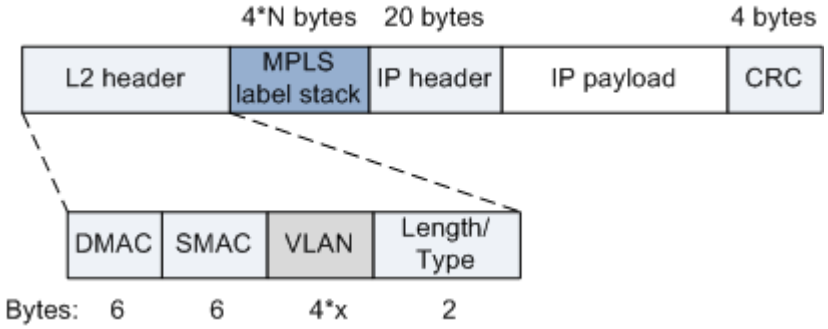
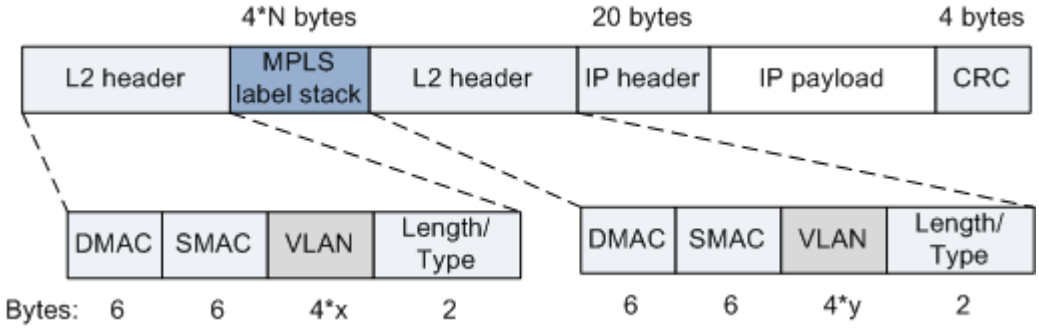


图 2 MPLS L2VPN 报文结构



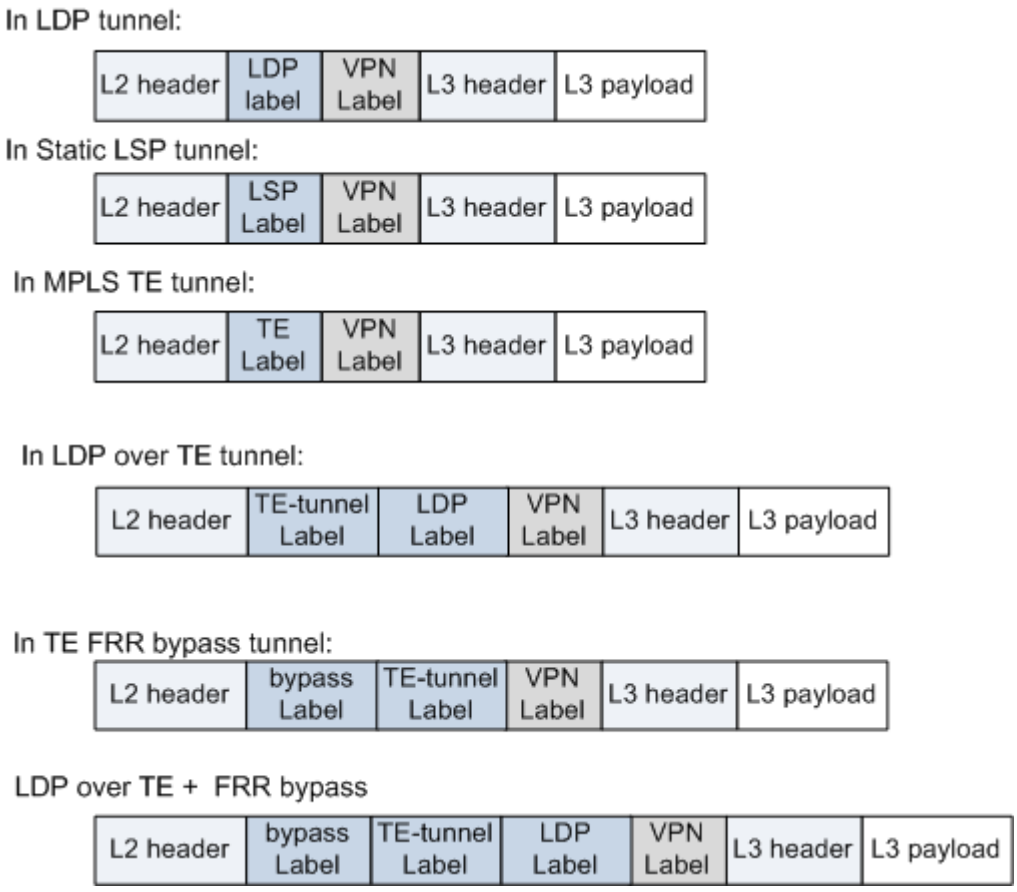
各种公网隧道场景下的 MPLS 报文携带的标签

在 VPN 场景下，MPLS 报文携带 M 层私网标签+N 层公网隧道标签，M 取决于 VPN 的组网场景（请参见后文），N 的取值取决于公网隧道类型。

场景	公网隧道标签层数 N
隧道为 LDP LSP	1 层公网标签
隧道为 static LSP	1 层公网标签

场景	公网隧道标签层数 N
隧道为 TE 隧道	1 层公网标签
隧道为 LDP over TE 隧道，报文在 TE 隧道上传输时	2 层公网标签
TE FRR 场景下，报文在 bypass 隧道上传输时	2 层公网标签
隧道为 LDP over TE 且部署 TE FRR 场景下，被保护接口出现故障，业务切换到 bypass 隧道，报文在 bypass 隧道上传输时	3 层公网标签

图 3 各种公网隧道场景下的 MPLS 报文格式

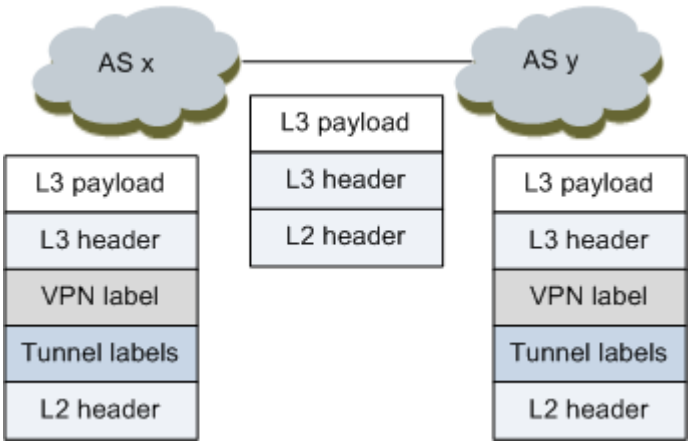


跨域 VPN OptionA 场景

在各 AS 域内传递时，报文携带 1 层私网 MPLS 标签 + N 层公网隧道标签（公网隧道标签参见上文）。

在 AS 域间传递时，报文不携带 MPLS 标签。

图 4 跨域 VPN OptionA 场景下的报文结构

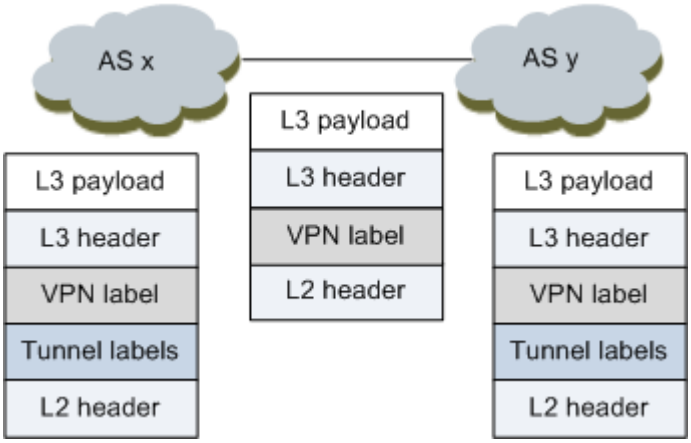


跨域 VPN OptionB 场景

在各 AS 域内传递时，报文携带 1 层私网 MPLS 标签 + N 层公网隧道标签（公网隧道标签参见上文）。

在 AS 域间传递时，报文携带 1 层私网 MPLS 标签。

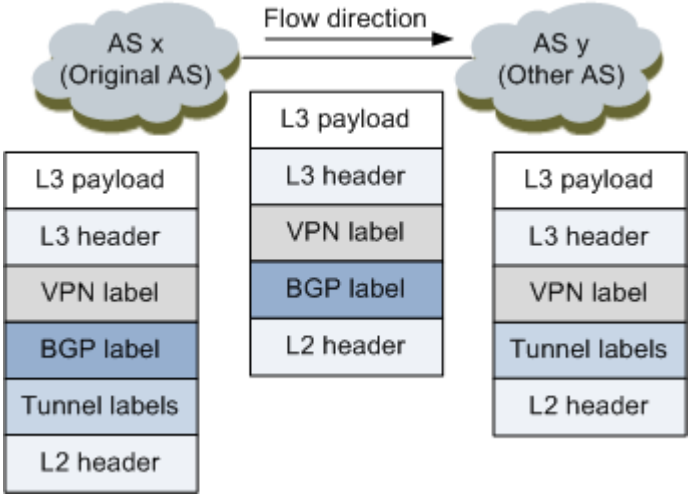
图 5 跨域 VPN OptionB 场景下的报文结构



跨域 VPN OptionC 场景

- 在首发的 AS 域内传递时，报文携带 1 层私网标签 + 1 层 BGP 标签 + N 层公网隧道标签
- 在 AS 域间传递时，报文携带 1 层私网标签+1 层 BGP 标签
- 在第 2 个 AS 域间传递时，报文携带 1 层私网标签+ N 层公网隧道标签

图 6 跨域 VPN OptionC 场景下的报文结构

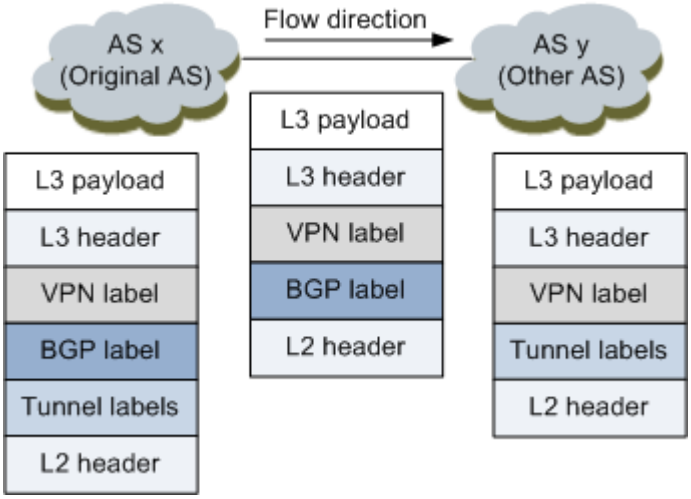


HoVPN/HVPLS 场景

- 核心层传递时，报文携带 1 层内层标签+N 层公网隧道标签

- UPE 和 SPE 之间传递时，报文携带 1 层内层标签

图 7 跨域 VPN OptionC 场景下的报文结构



运营商的运营商 CSC（Carriers’ Carrier）场景

与普通 BGP/MPLS IP VPN 相比，运营商的运营商的实现关键在于一级运营商 CE 接入到一级运营商 PE 这一部分。而二级运营商可能只是普通 SP，也可能是 BGP/MPLS IP VPN 服务提供商。

- 二级运营商是普通 SP 时，需要在一级运营商 CE 之间建立 BGP 会话交换外部路由。二级运营商内部通过 IGP 或 BGP 扩散二级运营商的外部路由。二级运营商用户侧 PE 可以不部署 MPLS。
- 二级运营商是 BGP/MPLS IP VPN 服务提供商时，其 PE 也需要运行 MPLS，与一级运营商 CE 之间运行 IGP 和 LDP。二级运营商 PE 之间通过 MP-BGP 会话交换外部路由。

图 8 运营商的运营商场景(二级运营商为普通 SP)

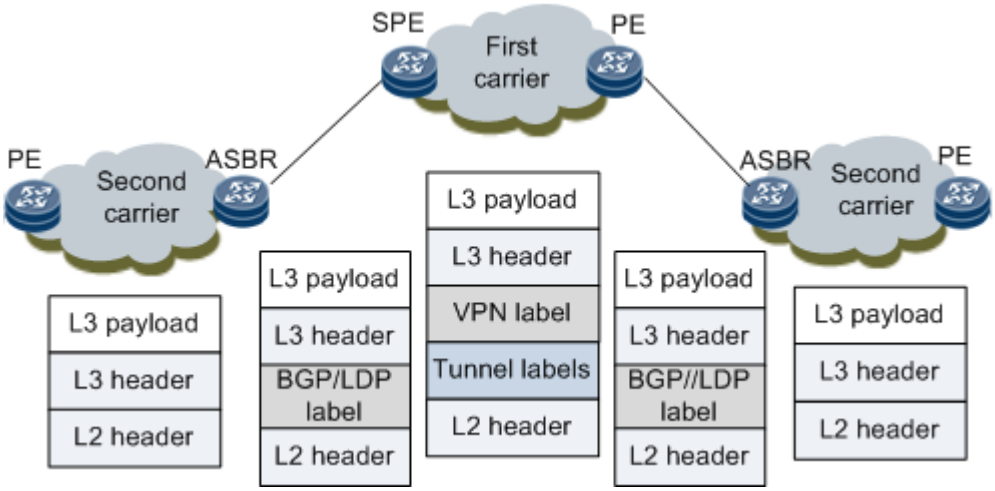
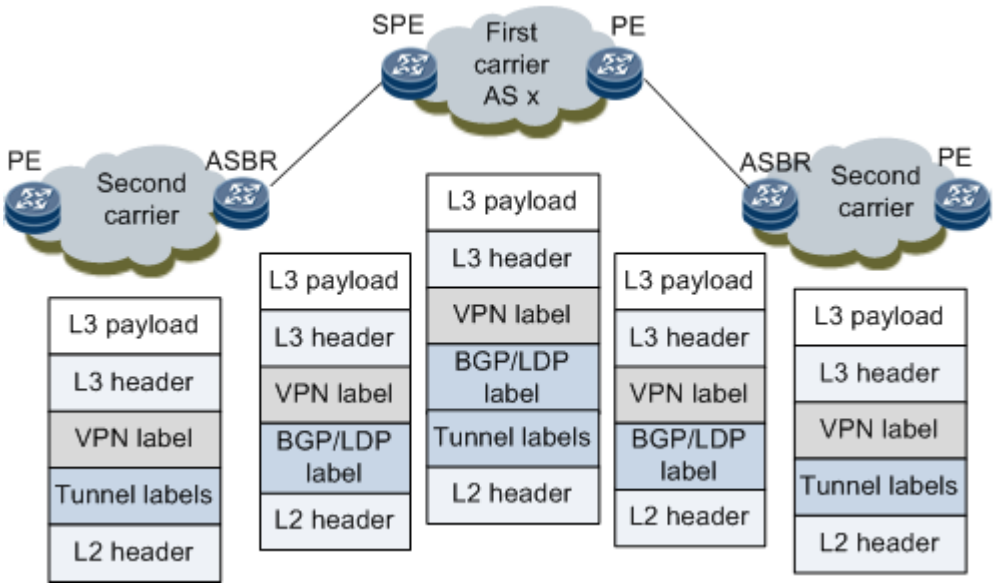


图 9 运营商的运营商场景(二级运营商也为 MPLS VPN SP)



报文示例

图 10 MPLS L2VPN 报文

```
Frame 1: 96 bytes on wire (768 bits), 96 bytes captured (768 bits)
Ethernet II, Src: Ibm_22:00:23 (00:22:00:22:00:23), Dst: HuaweiTe_b2:29:93 (0
  Destination: HuaweiTe_b2:29:93 (00:25:9e:b2:29:93)
  Source: Ibm_22:00:23 (00:22:00:22:00:23)
  Type: MPLS label switched packet (0x8847)
MultiProtocol Label Switching Header, Label: 1033, Exp: 3, S: 0, TTL: 255
  MPLS Label: 1033
  MPLS Experimental Bits: 3
  MPLS Bottom Of Label Stack: 0
  MPLS TTL: 255
MultiProtocol Label Switching Header, Label: 1032, Exp: 3, S: 1, TTL: 255
  MPLS Label: 1032
  MPLS Experimental Bits: 3
  MPLS Bottom Of Label Stack: 1
  MPLS TTL: 255
Ethernet II (VLAN tagged), Src: Ibm_22:00:23 (00:22:00:22:00:23), Dst: IPv4mc
  Destination: IPv4mcast_00:00:b8 (01:00:5e:00:00:b8)
  Source: Ibm_22:00:23 (00:22:00:22:00:23)
  VLAN tag: VLAN=4095, Priority=Network Control
  Type: IP (0x0800)
  Trailer: 692d2756
Internet Protocol Version 4, Src: 7.8.10.239 (7.8.10.239), Dst: 224.0.0.184 (
User Datagram Protocol, Src Port: 49352 (49352), Dst Port: bfd-control (3784)
BFD Control message
```

参考标准

标准	描述
RFC 3031	Multiprotocol Label Switching Architecture
RFC 3032	MPLS Label Stack Encoding
RFC 3034	Use of Label Switching on Frame Relay Networks Specification
RFC 3035	MPLS using LDP and ATM VC Switching

标准	描述
RFC 2547	BGP/MPLS VPNs
RFC 4090	Fast Reroute Extensions to RSVP-TE for LSP Tunnels
RFC 3107	Carrying Label Information in BGP-4

3.3 MPLS Ping/Tracert (MPLS Echo)报文格式

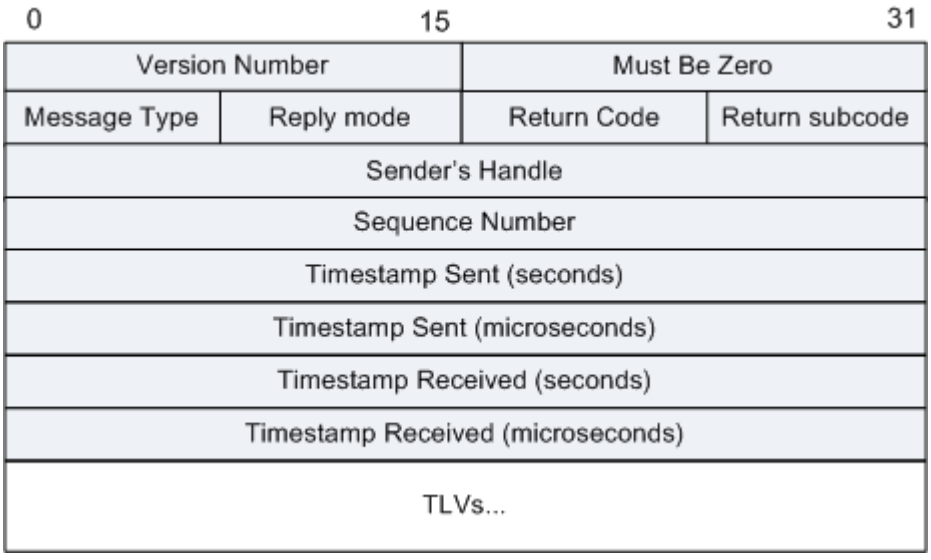
MPLS LSP Ping/Tracert 通过发送 MPLS Echo 消息实现。MPLS Echo 消息使用 IPv4/IPv6 的 UDP 协议封装，UDP 端口为 3503，只有使能 MPLS 的路由器才能够识别该端口号。

MPLS LSP Traceroute 和传统的 Traceroute 类似，通过连续发送一个 TTL 步进为 1 的 Echo Request 报文，让 LSP 沿途的每一个 LSR 都会收到 TTL 超时的 Echo Request 报文，同时回送一个携带下游信息（可选）以及相应返回码的 Echo Reply 给发送者。这样发送者就会得到该 LSP 沿途每一个节点的信息。

报文格式

MPLS Echo 消息封装在 IPv4/IPv6 的 UDP 报文中，可能还带有 MPLS 标签。MPLS Echo 消息中，UDP 负载的格式如下：

图 1 MPLS LSP Echo 报文格式



字段	长度	含义
Version Number	2 字节	标识 MPLS Echo 的版本号，目前为 1。

字段	长度	含义
Must Be Zero	2 字节	必须填全 0，接收时忽略。.
Message Type	1 字节	标识该 MPLS Echo 消息的类型： • 1: MPLS Echo 请求消息 • 2: MPLS Echo 响应消息
Reply mode	1 字节	指示 Reply Router 采用什么方式回应这个消息： • 1: Do Not Reply • 2: Reply via an IPv4 UDP Packet • 3: Reply via an IPv4 UDP packet with Router Alert
Return Code	1 字节	发送端设置为 0，接收端可以设置为如下值之一： • 0: No return code • 1: Malformed Echo Request Received • 2: One Or More of the TLVs Was Not Understood • 3: Replying Router Is an Egress for the FEC at stack-depth <RSC> • 4: Replying Router Has No Mapping for the FEC at stack-depth <RSC> • 5: Downstream Mapping Mismatch • 6: Upstream Interface Index Unknown • 7: Reserved
Return subcode	1 字节	Return Subcode 字段包含了标签栈的处理结束的指针。如果 Return subcode 值为 0，标识报文没有携带标签，不需要处理标签。否则，报文携带了标签。
Sender's Handle	4 字节	发送者句柄，是用来标识一个 MPLS Echo 的，其值是在应用程序发送一个 MPLS Echo Request 时随机生成的。单次的 LSP Ping 操作可以产生多个 Echo Request，但是这些 Echo Request 所包含的 Sender's Handle 的值是相同，即单次 LSP Ping 操作仅能产生一个 Sender's Handle 的 Echo Request。
Sequence Number	4 字节	序列号，Sequence Number 同样是用来标识 MPLS Echo 的，它是一个进程的概念，进程内有效，可以用来检测丢失的 Reply 的个数，从而可以对网络进行延时和抖动统计。单次 LSP Ping 操作可以产生多个 Sequence Number，其值一般从零开始逐一递增。
Timestamp	4 字节	时间戳，采用 NTP 协议的时间格式，包含两部分：收到的时间戳和发送时间戳；可以用来计算报文从一个节点到另一个节点所需要花费的时间。

字段	长度	含义
TLVs	可变	TLV (Type, Length, Value): • Type = 1: Target FEC Stack TLV • Type = 2: Downstream mapping TLV • Type = 3: PAD TLV • Type = 4: Error Code TLV • Type = 5: Vendor Enterprise Code TLV Length: Value 字段的长度，字节为计数单位。 Value: 取决于 Type 的取值，如果 TLV 不足 4 字节的整数倍，需要填充。

报文示例

图 2 MPLS LSP ping request

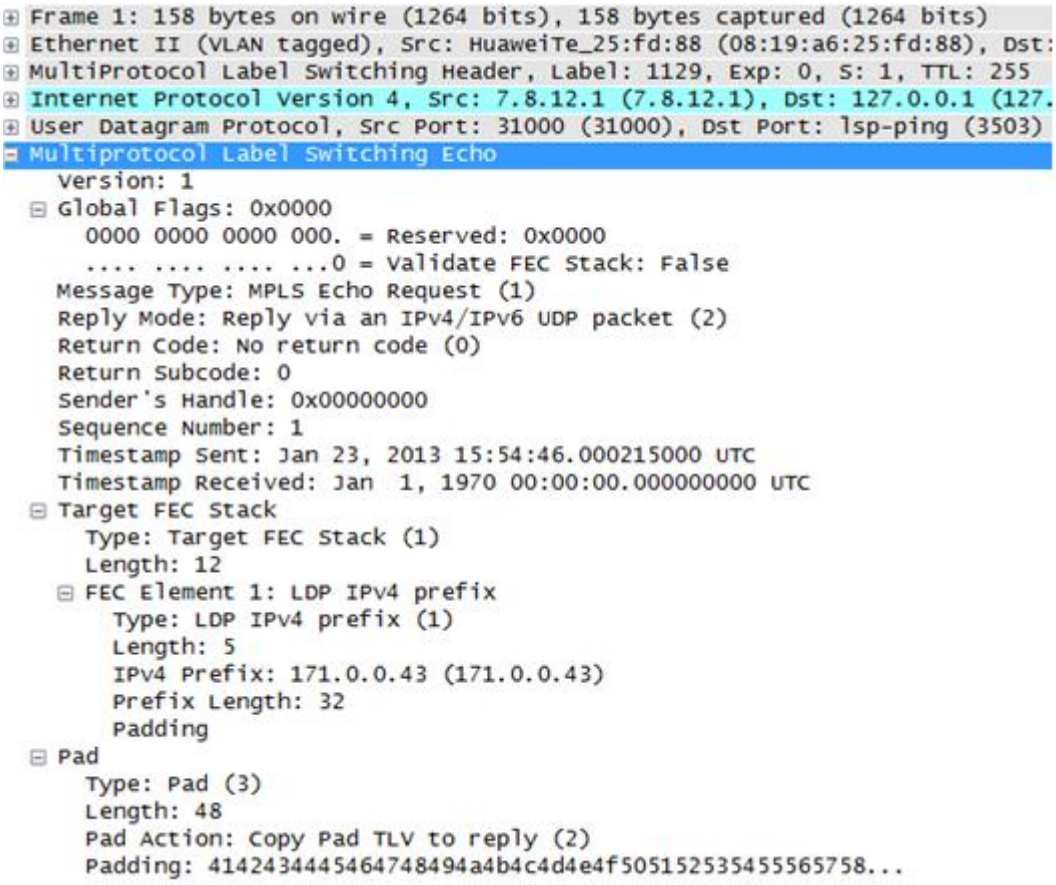


图 3 MPLS LSP ping reply

```
⊞ Frame 1: 150 bytes on wire (1200 bits), 150 bytes captured (1200 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:08 (54:89:98:74:e4:08), Dst: Hu
⊞ Internet Protocol Version 4, Src: 171.0.0.43 (171.0.0.43), Dst: 7.8.12.1 (7.8.
⊞ User Datagram Protocol, Src Port: lsp-ping (3503), Dst Port: 31000 (31000)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊟ Global Flags: 0x0000
      0000 0000 0000 000. = Reserved: 0x0000
      .... .... .... ...0 = Validate FEC Stack: False
    Message Type: MPLS Echo Reply (2)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: Replying router is an egress for the FEC at stack depth RSC (3)
    Return Subcode: 1
    Sender's Handle: 0x00000000
    Sequence Number: 1
    Timestamp Sent: Jan 23, 2013 15:54:46.000215000 UTC
    Timestamp Received: Jan 24, 2013 15:53:19.000099000 UTC
  ⊟ Target FEC Stack
      Type: Target FEC Stack (1)
      Length: 12
    ⊟ FEC Element 1: LDP IPv4 prefix
        Type: LDP IPv4 prefix (1)
        Length: 5
        IPv4 Prefix: 171.0.0.43 (171.0.0.43)
        Prefix Length: 32
        Padding
  ⊟ Pad
      Type: Pad (3)
      Length: 48
      Pad Action: Copy Pad TLV to reply (2)
      Padding: 414243444546474849 4a4b4c4d4e4f505152535455565758...
```

图 4 MPLS TE ping request

```
⊞ Frame 1: 158 bytes on wire (1264 bits), 158 bytes captured (1264 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_25:fd:88 (08:19:a6:25:fd:88), Ds
⊞ MultiProtocol Label Switching Header, Label: 1139, Exp: 0, S: 1, TTL: 255
⊞ Internet Protocol Version 4, Src: 7.8.12.1 (7.8.12.1), Dst: 127.0.0.1 (12
⊞ User Datagram Protocol, Src Port: 31005 (31005), Dst Port: lsp-ping (3503
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊟ Global Flags: 0x0000
      0000 0000 0000 000. = Reserved: 0x0000
      .... .... .... ...0 = Validate FEC Stack: False
    Message Type: MPLS Echo Request (1)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: No return code (0)
    Return Subcode: 0
    Sender's Handle: 0x00000005
    Sequence Number: 2
    Timestamp Sent: Jan 23, 2013 20:42:58.000133000 UTC
    Timestamp Received: Jan  1, 1970 00:00:00.000000000 UTC
  ⊟ Target FEC Stack
      Type: Target FEC Stack (1)
      Length: 24
    ⊟ FEC Element 1: RSVP IPv4 Session Query
        Type: RSVP IPv4 Session Query (3)
        Length: 20
        IPv4 Tunnel endpoint address: 171.0.0.43 (171.0.0.43)
        Must Be Zero: 0
        Tunnel ID: 1
        Extended Tunnel ID: 0xAB000029 (171.0.0.41)
        IPv4 Tunnel sender address: 171.0.0.41 (171.0.0.41)
        Must Be Zero: 0
        LSP ID: 3
  ⊟ Pad
      Type: Pad (3)
      Length: 36
      Pad Action: Copy Pad TLV to reply (2)
      Padding: 414243444546474849 4a4b4c4d4e4f505152535455565758...
```

图 5 MPLS TE ping reply

```
⊞ Frame 1: 150 bytes on wire (1200 bits), 150 bytes captured (1200 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:08 (54:89:98:74:e4:08), Dst: Hu
⊞ Internet Protocol Version 4, Src: 171.0.0.43 (171.0.0.43), Dst: 7.8.12.1 (7.8.
⊞ User Datagram Protocol, Src Port: lsp-ping (3503), Dst Port: 31005 (31005)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊟ Global Flags: 0x0000
      0000 0000 0000 000. = Reserved: 0x0000
      .... .... .... ...0 = Validate FEC Stack: False
    Message Type: MPLS Echo Reply (2)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: Replying router is an egress for the FEC at stack depth RSC (3)
    Return Subcode: 1
    Sender's Handle: 0x00000005
    Sequence Number: 2
    Timestamp Sent: Jan 23, 2013 20:42:58.000133000 UTC
    Timestamp Received: Jan 24, 2013 20:41:31.000039000 UTC
  ⊟ Target FEC Stack
      Type: Target FEC Stack (1)
      Length: 24
    ⊟ FEC Element 1: RSVP IPv4 Session Query
        Type: RSVP IPv4 Session Query (3)
        Length: 20
        IPv4 Tunnel endpoint address: 171.0.0.43 (171.0.0.43)
        Must Be Zero: 0
        Tunnel ID: 1
        Extended Tunnel ID: 0xAB000029 (171.0.0.41)
        IPv4 Tunnel sender address: 171.0.0.41 (171.0.0.41)
        Must Be Zero: 0
        LSP ID: 3
  ⊟ Pad
      Type: Pad (3)
      Length: 36
      Pad Action: Copy Pad TLV to reply (2)
      Padding: 414243444546474849 4a4b4c4d4e4f505152535455565758...
```

图 6 VCCV ping request

```
⊞ Frame 1: 166 bytes on wire (1328 bits), 166 bytes captured (1328 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_25:fd:88 (08:19:a6:25:fd:88), Dst
⊞ MultiProtocol Label Switching Header, Label: 1127, Exp: 0, S: 0, TTL: 255
⊞ MultiProtocol Label Switching Header, Label: 1097, Exp: 0, S: 1, TTL: 255
⊞ PW Associated Channel Header
⊞ Internet Protocol Version 4, Src: 171.0.0.41 (171.0.0.41), Dst: 127.0.0.1
⊞ User Datagram Protocol, Src Port: 31010 (31010), Dst Port: lsp-ping (3503)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊞ Global Flags: 0x0000
    Message Type: MPLS Echo Request (1)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: No return code (0)
    Return Subcode: 0
    Sender's Handle: 0x0000000a
    Sequence Number: 1
    Timestamp Sent: Jan 24, 2013 19:04:41.000044000 UTC
    Timestamp Received: Jan  1, 1970 00:00:00.000000000 UTC
  ⊟ Target FEC Stack
      Type: Target FEC Stack (1)
      Length: 20
    ⊟ FEC Element 1: FEC 128 Pseudowire (new)
        Type: FEC 128 Pseudowire (new) (10)
        Length: 14
        Sender's PE Address: 171.0.0.41 (171.0.0.41)
        Remote PE Address: 171.0.0.43 (171.0.0.43)
        VC ID: 2001
        Encapsulation: Ethernet VLAN (4)
        MBZ: 0x0000
        Padding
  ⊟ Pad
      Type: Pad (3)
      Length: 40
      Pad Action: Copy Pad TLV to reply (2)
      Padding: 0000000000000000000000000000000000000000000000000000...
```

图 7 VCCV ping request alert

```
⊞ Frame 1: 150 bytes on wire (1200 bits), 150 bytes captured (1200 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:08 (54:89:98:74:e4:08), Dst: Hua
⊞ Internet Protocol Version 4, Src: 171.0.0.43 (171.0.0.43), Dst: 171.0.0.41 (171
⊞ User Datagram Protocol, Src Port: lsp-ping (3503), Dst Port: 31010 (31010)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊞ Global Flags: 0x0000
    Message Type: MPLS Echo Reply (2)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: Replying router is an egress for the FEC at stack depth RSC (3)
    Return Subcode: 1
    Sender's Handle: 0x0000000a
    Sequence Number: 1
    Timestamp Sent: Jan 24, 2013 19:04:41.000044000 UTC
    Timestamp Received: Jan 25, 2013 19:03:14.000052000 UTC
  ⊟ Target FEC Stack
      Type: Target FEC Stack (1)
      Length: 20
    ⊟ FEC Element 1: FEC 128 Pseudowire (new)
        Type: FEC 128 Pseudowire (new) (10)
        Length: 14
        Sender's PE Address: 171.0.0.41 (171.0.0.41)
        Remote PE Address: 171.0.0.43 (171.0.0.43)
        VC ID: 2001
        Encapsulation: Ethernet VLAN (4)
        MBZ: 0x0000
        Padding
  ⊟ Pad
      Type: Pad (3)
      Length: 40
      Pad Action: Copy Pad TLV to reply (2)
      Padding: 00000000000000000000000000000000000000000000000000...
```

图 8 VPLS MAC Ping Request

```
⊞ Frame 1: 222 bytes on wire (1776 bits), 222 bytes captured (1776 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_25:fd:88 (08:19:a6:25:fd:88), Dst: HuaweiTe_7
⊞ MultiProtocol Label Switching Header, Label: 1127, Exp: 0, S: 0, TTL: 255
⊞ MultiProtocol Label Switching Header, Label: 1 (Router Alert), Exp: 0, S: 0, TTL: 255
⊞ MultiProtocol Label Switching Header, Label: 1099, Exp: 0, S: 1, TTL: 255
⊞ Ethernet II, Src: HuaweiTe_25:fd:80 (08:19:a6:25:fd:80), Dst: 00:00:00_00:00:01 (00:00
⊞ Internet Protocol Version 4, Src: 171.0.0.41 (171.0.0.41), Dst: 127.0.0.1 (127.0.0.1)
⊞ User Datagram Protocol, Src Port: 31020 (31020), Dst Port: lsp-ping (3503)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊞ Global Flags: 0x0000
    Message Type: MPLS Echo Request (1)
    Reply Mode: Unknown (5)
    Return Code: No return code (0)
    Return Subcode: 0
    Sender's Handle: 0x00000014
    Sequence Number: 1
    Timestamp Sent: Jan 24, 2013 19:35:25.000033000 UTC
    Timestamp Received: Not representable
  ⊞ Target FEC Stack
  ⊟ Vendor Private
      Type: Vendor Private (64512)
      Length: 56
      Vendor Id: HUAWEI Technology Co.,Ltd (2011)
      Value: 000100300004000000000000100000819a625fd8000000000...
  ⊟ Pad
      Type: Pad (3)
      Length: 22
      Pad Action: Copy Pad TLV to reply (2)
      Padding: 414243444546474849 4a4b4c4d4e4f505152535455
```

图 9 VPLS MAC Ping Reply

```
⊞ Frame 1: 286 bytes on wire (2288 bits), 286 bytes captured (2288 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:08 (54:89:98:74:e4:08), Dst: Hua
⊞ MultiProtocol Label Switching Header, Label: 1 (Router Alert), Exp: 0, S: 0, T
⊞ MultiProtocol Label Switching Header, Label: 1092, Exp: 0, S: 1, TTL: 255
⊞ Ethernet II, Src: HuaweiTe_25:fd:80 (08:19:a6:25:fd:80), Dst: 00:00:00_00:00:0
⊞ Internet Protocol Version 4, Src: 7.8.23.3 (7.8.23.3), Dst: 127.0.0.1 (127.0.0.
⊞ User Datagram Protocol, Src Port: lsp-ping (3503), Dst Port: 31020 (31020)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊞ Global Flags: 0x0000
    Message Type: MPLS Echo Reply (2)
    Reply Mode: Unknown (5)
    Return Code: Replying router is an egress for the FEC at stack depth RSC (3)
    Return Subcode: 1
    Sender's Handle: 0x00000014
    Sequence Number: 1
    Timestamp Sent: Jan 24, 2013 19:35:25.000033000 UTC
    Timestamp Received: Jan 25, 2013 19:33:58.000043000 UTC
  ⊞ Target FEC Stack
  ⊞ Vendor Private
  ⊞ Pad
```

图 10 VPLS PW Ping Request

```
⊞ Frame 1: 166 bytes on wire (1328 bits), 166 bytes captured (1328 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_25:fd:88 (08:19:a6:25:fd:88), Dst: HuaweiTe_7
⊞ MultiProtocol Label Switching Header, Label: 1127, Exp: 0, S: 0, TTL: 255
⊞ MultiProtocol Label Switching Header, Label: 1 (Router Alert), Exp: 0, S: 0, TTL: 255
⊞ MultiProtocol Label Switching Header, Label: 1099, Exp: 0, S: 1, TTL: 255
⊞ Internet Protocol Version 4, Src: 171.0.0.41 (171.0.0.41), Dst: 127.0.0.1 (127.0.0.1)
⊞ User Datagram Protocol, Src Port: 31018 (31018), Dst Port: lsp-ping (3503)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊞ Global Flags: 0x0000
    Message Type: MPLS Echo Request (1)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: No return code (0)
    Return Subcode: 0
    Sender's Handle: 0x00000012
    Sequence Number: 2
    Timestamp Sent: Jan 24, 2013 19:24:00.000048000 UTC
    Timestamp Received: Jan  1, 1970 00:00:00.000000000 UTC
  ⊟ Target FEC Stack
      Type: Target FEC Stack (1)
      Length: 20
    ⊟ FEC Element 1: FEC 128 Pseudowire (new)
        Type: FEC 128 Pseudowire (new) (10)
        Length: 14
        Sender's PE Address: 171.0.0.41 (171.0.0.41)
        Remote PE Address: 171.0.0.43 (171.0.0.43)
        VC ID: 1
        Encapsulation: Ethernet VLAN (4)
        MBZ: 0x0000
        Padding
  ⊟ Pad
      Type: Pad (3)
      Length: 40
      Pad Action: Copy Pad TLV to reply (2)
      Padding: 414243444546474849 4a4b4c4d4e4f505152535455565758...
```

图 11 VPLS Multi-Hop Ping Request

```
⊞ Frame 1: 185 bytes on wire (1480 bits), 185 bytes captured (1480 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_25:fd:88 (08:19:a6:25:fd:88), Dst:
⊞ MultiProtocol Label Switching Header, Label: 1127, Exp: 0, S: 0, TTL: 255
⊞ MultiProtocol Label Switching Header, Label: 1 (Router Alert), Exp: 0, S: 0
⊞ MultiProtocol Label Switching Header, Label: 1099, Exp: 0, S: 1, TTL: 255
⊞ Ethernet II, Src: HuaweiTe_25:fd:80 (08:19:a6:25:fd:80), Dst: IPv4mcast_01:
⊞ Internet Protocol Version 4, Src: 171.0.0.41 (171.0.0.41), Dst: 225.1.1.1 (
⊞ User Datagram Protocol, Src Port: 31022 (31022), Dst Port: lsp-ping (3503)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊞ Global Flags: 0x0000
    Message Type: MPLS Echo Request (1)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: No return code (0)
    Return Subcode: 0
    Sender's Handle: 0x00000016
    Sequence Number: 4
    Timestamp Sent: Jan 24, 2013 19:39:37.000213000 UTC
    Timestamp Received: Jan  1, 1970 00:00:00.000000000 UTC
  ⊟ Target FEC Stack
      Type: Target FEC Stack (1)
      Length: 20
    ⊞ FEC Element 1: FEC 128 Pseudowire (new)
  ⊟ Vendor Private
      Type: Vendor Private (64512)
      Length: 28
      Vendor Id: HUAWEI Technology Co.,Ltd (2011)
      Value: 00050014e1010101ab0000290819a625fd80000000000000
  ⊟ Pad
      Type: Pad (3)
      Length: 13
      Pad Action: Copy Pad TLV to reply (2)
      Padding: 41424344454647484 94a4b4c
```

图 12 VPLS Multi-Hop Ping Reply

```
⊞ Frame 1: 138 bytes on wire (1104 bits), 138 bytes captured (1104 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:08 (54:89:98:74:e4:08), Dst: Hu
⊞ Internet Protocol Version 4, Src: 171.0.0.43 (171.0.0.43), Dst: 171.0.0.41 (17
⊞ User Datagram Protocol, Src Port: lsp-ping (3503), Dst Port: 31022 (31022)
⊟ Multiprotocol Label Switching Echo
    Version: 1
  ⊞ Global Flags: 0x0000
    Message Type: MPLS Echo Reply (2)
    Reply Mode: Reply via an IPv4/IPv6 UDP packet (2)
    Return Code: Replying router is an egress for the FEC at stack depth RSC (3)
    Return Subcode: 1
    Sender's Handle: 0x00000016
    Sequence Number: 4
    Timestamp Sent: Jan 24, 2013 19:39:37.000213000 UTC
    Timestamp Received: Jan 25, 2013 19:38:10.000223000 UTC
  ⊟ Target FEC Stack
      Type: Target FEC Stack (1)
      Length: 20
    ⊞ FEC Element 1: FEC 128 Pseudowire (new)
  ⊟ Vendor Private
      Type: Vendor Private (64512)
      Length: 28
      Vendor Id: HUAWEI Technology Co.,Ltd (2011)
      Value: 00050014e1010101ab0000290819a625fd80000000000000
```

参考标准

标准	描述
RFC 4379	Detecting Multi-Protocol Label Switched (MPLS) Data Plane Failures
RFC 4377	Operations and Management (OAM) Requirements for Multi-Protocol Label Switched (MPLS) Networks

标准	描述
RFC 5085	Pseudowire Virtual Circuit Connectivity Verification (VCCV): A Control Channel for Pseudowires
RFC 4447	Pseudowire Setup and Maintenance Using the Label Distribution Protocol (LDP)

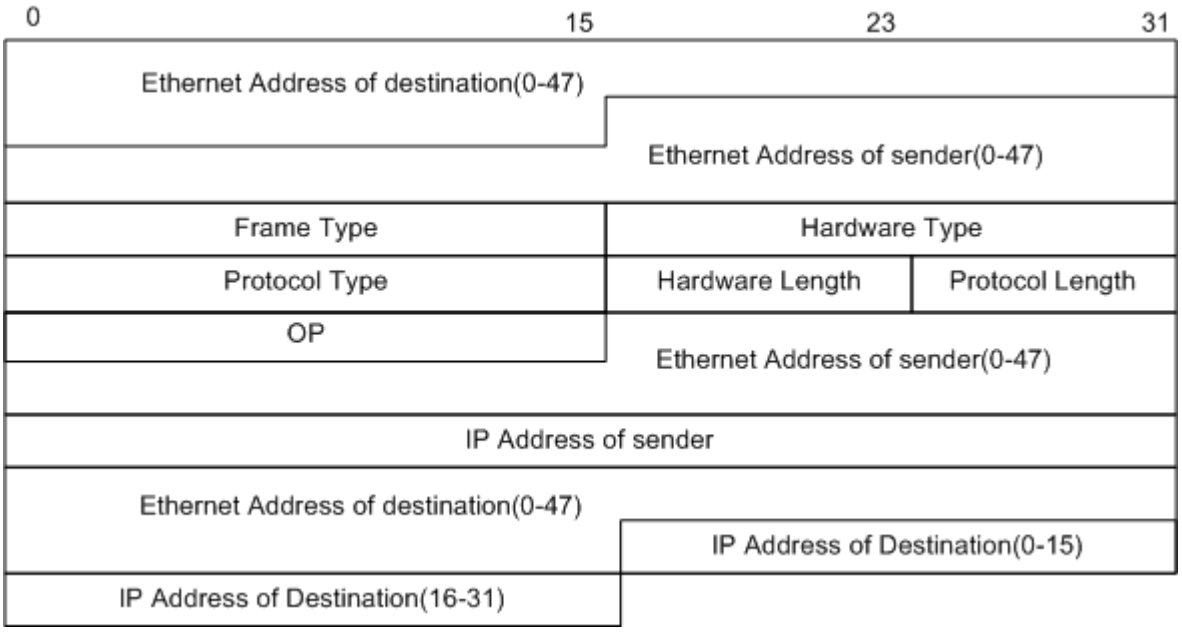
4. 网络层

- ARP/RARP 报文格式
- GRE 报文格式
- ICMP 报文格式
- ICMPv6 报文格式
- IGMP 报文格式
- IP in IP 报文格式
- IP 报文格式
- IPv6 报文格式
- IPv6 in IP (6to4)报文格式
- MLD 报文格式
- OSPF 报文格式
- OSPFv3 报文格式
- PIM 报文格式
- RSVP 报文格式
- VRRP 报文格式

4.1 ARP/RARP 报文格式

地址解析协议 ARP（Address Resolution Protocol）是用来将 IP 地址解析为 MAC 地址的协议。

报文格式



字段	长度（bit）	含义
Ethernet Address of destination	48 比特	目的以太网地址。发送 ARP 请求时，为广播的 MAC 地址，0xFF.FF.FF.FF.FF.FF。
Ethernet Address of sender	48 比特	源以太网地址。
Frame Type	16 比特	表示后面数据的类型。对于 ARP 请求或应答来说，该字段的值为 0x0806。
Hardware Type	16 比特	表示硬件地址的类型。对于以太网，该类型的值为“1”。
Protocol Type	16 比特	表示发送方要映射的协议地址类型。对于 IP 地址，该值为 0x0800。
Hardware Length	8 比特	表示硬件地址的长度，单位是字节。对于 ARP 请求或应答来说，该值为 6。
Protocol Length	8 比特	表示协议地址的长度，单位是字节。对于 ARP 请求或应答来说，该值为 4。
OP	16 比特	操作类型: • 1 ARP 请求 • 2 ARP 应答 • 3 RARP 请求

字段	长度（bit）	含义
		• 4 RARP 应答
Ethernet Address of sender	48 比特	发送方以太网地址。这个字段和 ARP 报文首部的源以太网地址字段是重复信息。
IP Address of sender	32 比特	发送方的 IP 地址。
Ethernet Address of destination	48 比特	接收方的以太网地址。发送 ARP 请求时，该处填充值为 0x00.00.00.00.00.00。
IP Address of destination	32 比特	接收方的 IP 地址。

报文示例

图 1 免费 ARP 报文格式

图 2 ARP 请求报文格式

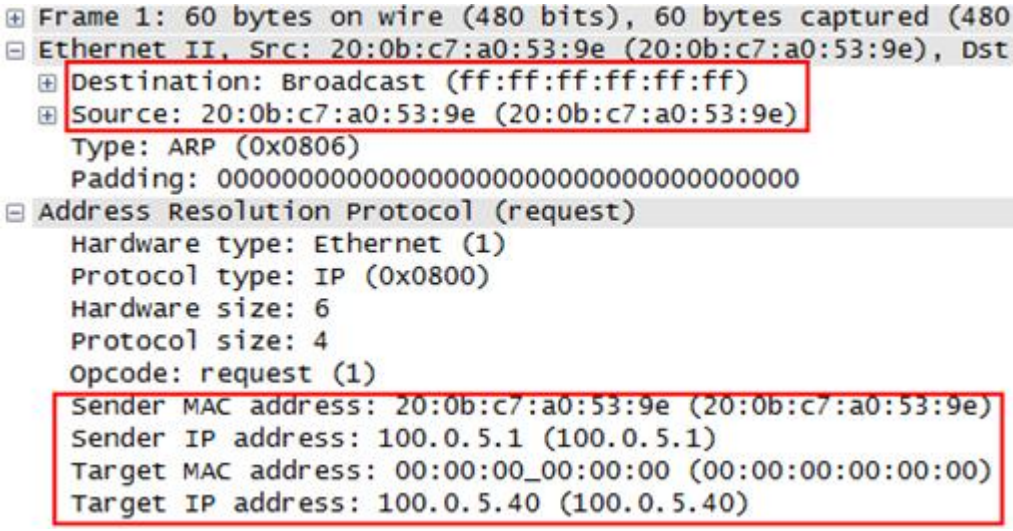


图 3 ARP 应答报文格式

```
⊞ Frame 7: 60 bytes on wire (480 bits), 60 bytes captured (480 b
⊟ Ethernet II, Src: HuaweiTe_9f:24:cb (4c:1f:cc:9f:24:cb), Dst:
  ⊞ Destination: 20:0b:c7:a0:53:9e (20:0b:c7:a0:53:9e)
  ⊞ Source: HuaweiTe_9f:24:cb (4c:1f:cc:9f:24:cb)
    Type: ARP (0x0806)
    Padding: 000000000000000000000000000000000000
⊟ Address Resolution Protocol (reply)
    Hardware type: Ethernet (1)
    Protocol type: IP (0x0800)
    Hardware size: 6
    Protocol size: 4
    Opcode: reply (2)
    Sender MAC address: HuaweiTe_9f:24:cb (4c:1f:cc:9f:24:cb)
    Sender IP address: 100.0.5.40 (100.0.5.40)
    Target MAC address: 20:0b:c7:a0:53:9e (20:0b:c7:a0:53:9e)
    Target IP address: 100.0.5.1 (100.0.5.1)
```

图 4 RARP 请求报文格式

```
⊞ Frame 1: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
⊞ Ethernet II, Src: Marquett_12:dd:88 (00:00:a1:12:dd:88), Dst: Broadcast
⊟ Address Resolution Protocol (reverse request)
    Hardware type: Ethernet (1)
    Protocol type: IP (0x0800)
    Hardware size: 6
    Protocol size: 4
    Opcode: reverse request (3)
    [Is gratuitous: False]
    Sender MAC address: Marquett_12:dd:88 (00:00:a1:12:dd:88)
    Sender IP address: 0.0.0.0 (0.0.0.0)
    Target MAC address: Marquett_12:dd:88 (00:00:a1:12:dd:88)
    Target IP address: 0.0.0.0 (0.0.0.0)
```

```
0000  ff ff ff ff ff ff 00 00  a1 12 dd 88 08 06 00 01   ........ ........
0010  08 00 06 04 00 03 00 00  a1 12 dd 88 00 00 00 00   ........ ........
0020  00 00 a1 12 dd 88 00 00  00 00 00 00 00 00 00 00   ........ ........
0030  00 00 00 00 00 00 00 00  00 00 00 00              ........ ....
```

参考标准

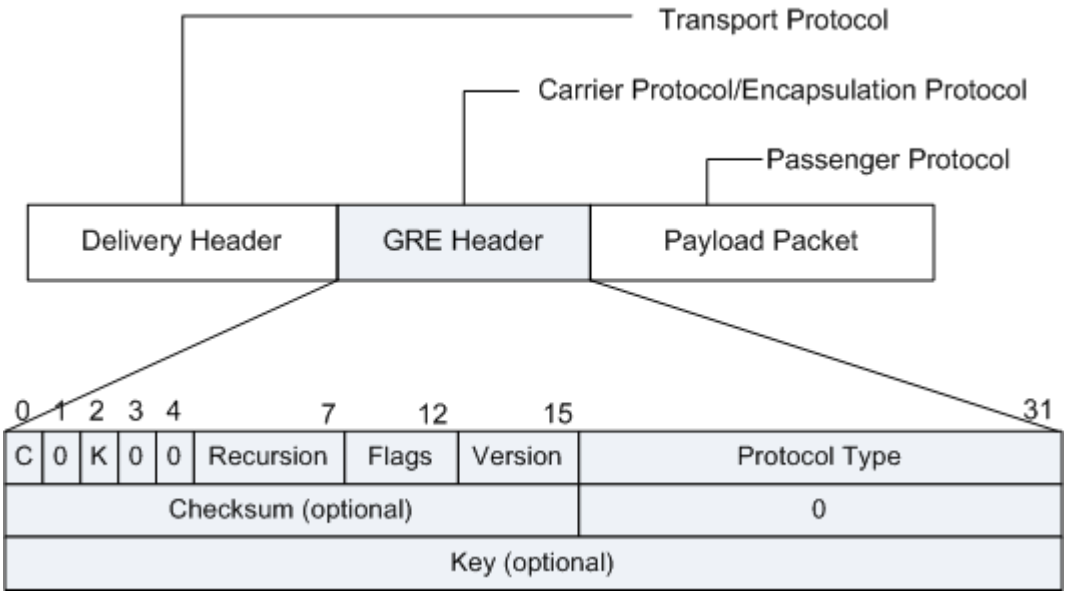
标准	描述
RFC 826	Ethernet Address Resolution Protocol
RFC 903	Reverse Address Resolution Protocol
RFC 1027	Using ARP to Implement Transparent Subnet Gateways
RFC 1042	Standard for the Transmission of IP Datagrams over IEEE 802 Networks
RFC 5227	IPv4 Address Conflict Detection

4.2 GRE 报文格式

报文格式

系统收到需要进行封装和路由的某网络层协议（如 IPX）数据时，将首先对其加上 GRE 报文头，使之成为 GRE 报文，再将其封装在另一协议（如 IP）中。这样，此报文的转发就可以完全由 IP 协议负责。封装后的报文的格式如下图所示：

图 1 GRE 报文格式



其中：

- 净荷（Payload）：系统收到的需要封装和路由的数据报称为净荷。
- 乘客协议（Passenger Protocol）：封装前的报文协议称为乘客协议。
- 封装协议（Encapsulation Protocol）：上述的 GRE 协议称为封装协议，也称为运载协议（Carrier Protocol）。
- 传输协议（Transport Protocol 或者 Delivery Protocol）：负责对封装后的报文进行转发的协议称为传输协议。

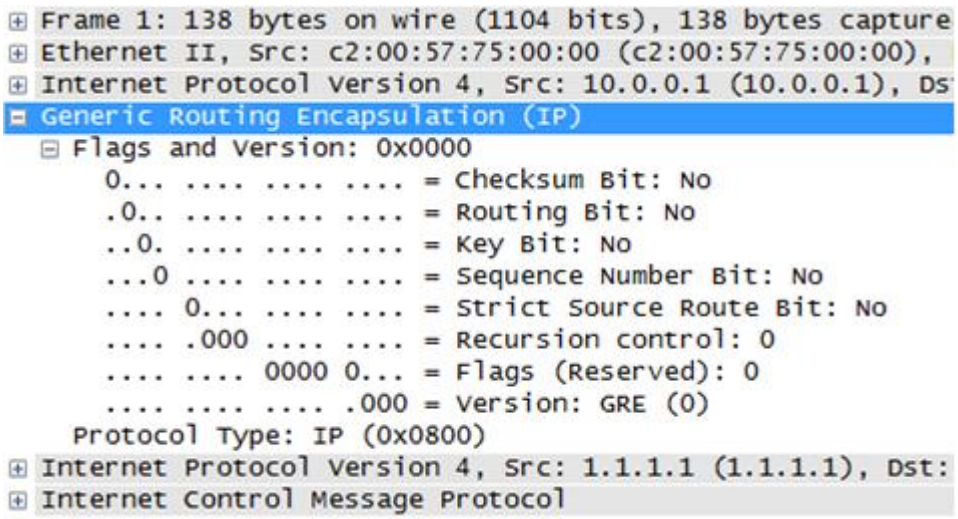
GRE 首部各字段解释如下：

字段	长度	描述
C	1 bit	校验和验证位。如果该位置 1，表示 GRE 头插入了校验和（Checksum）字段；该位为 0 表示 GRE 头不包含校验和字段。
K	1 bit	关键字位。如果该位置 1，表示 GRE 头插入了关键字（Key）字段；该位为 0 表示 GRE 头不包含关键字字段。
Recursion	3 bits	用来表示 GRE 报文被封装的层数。完成一次 GRE 封装后将该字段加 1。如果封装层数大于 3，则丢弃该报文。该字段的作用是防止报文被无限次的封装。

字段	长度	描述
Flags	5 bits	预留字段。当前必须设为 0。
Version	3 bits	版本字段，必须置为 0。Version 为 1 是使用在 RFC2637 的 PPTP 中。
Protocol Type	16 bits	乘客协议的协议类型。
Checksum	16 bits	对 GRE 头及其负载的校验和字段。
Key	31 bits	关键字字段，隧道接收端用于对收到的报文进行验证。

因为 VRP 中的 GRE 头不包含源路由字段，因此 Bit 1、Bit 3 和 Bit 4 都置为 0。

报文示例



参考标准

标准	描述
RFC 1701	Generic Routing Encapsulation (GRE)
RFC 2784	Generic Routing Encapsulation (GRE)

4.3 ICMP 报文格式

- ICMP 报文通用格式
- ICMP Echo Request/Reply 消息格式
- ICMP 目的不可达消息格式
- ICMP 重定向消息格式
- ICMP 超时消息格式
- ICMP 参数问题消息格式
- ICMP 源端被关闭消息格式

父主题： 网络层

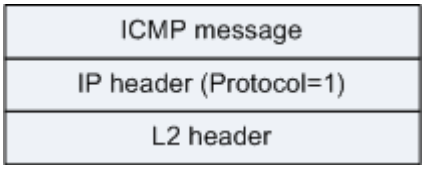
4.3.1 ICMP 报文通用格式

报文格式

有很多情况都会发送 ICMP 消息，例如，报文无法发送到目的地址，再如，网关设备没有足够的缓存来存储转发报文。

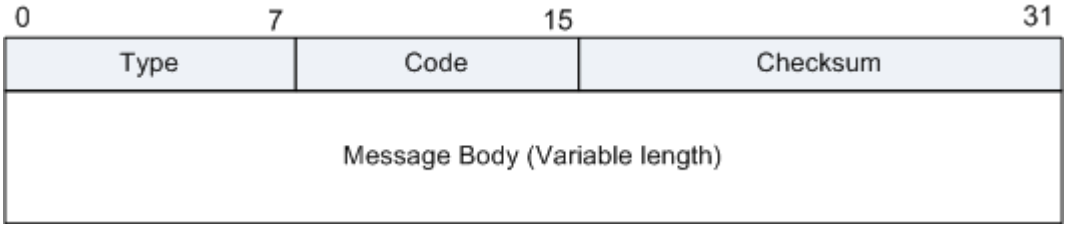
ICMP 消息封装在 IP 报文中，格式如下：

图 1 ICMP 消息封装格式



ICMP 消息头部格式如下：

图 2 ICMP 消息头部格式



其中，最后一个字段的长度和内容，取决于消息的类型和代码。对应的列表如下：

表 1 ICMP 消息类型代码对应表

类型 Type	代码 Code	描述

表 1 ICMP 消息类型代码对应表

类型 Type	代码 Code	描述
0	0	回显应答（ping 应答）
3	0	网络不可达
3	1	主机不可达
3	2	协议不可达
3	3	端口不可达
3	4	需要进行分片但设置不分片比特
3	5	源站选路失败
3	6	目的网络不认识
3	7	目的主机不认识
3	8	源主机被隔离（作废不用）
3	9	目的网络被强制禁止
3	10	目的主机被强制禁止
3	11	由于 TOS，网络不可达
3	12	由于 TOS，主机不可达
3	13	由于过滤，通信被强制禁止
3	14	主机越权

表 1 ICMP 消息类型代码对应表

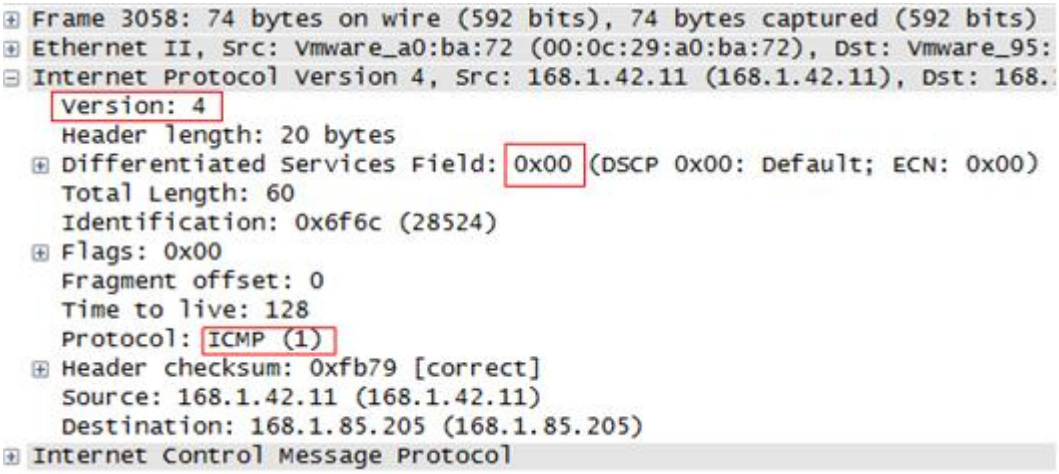
类型 Type	代码 Code	描述
3	15	优先权中止生效
4	0	源端被关闭
5	0	对网络重定向
5	1	对主机重定向
5	2	对服务类型和网络重定向
5	3	对服务类型和主机重定向
8	0	请求回显（ping 请求）
9	0	路由器通告
10	0	路由器请求告
11	0	传输期间生存时间为 0
11	1	在数据报组装期间生存时间为 0
12	0	坏的 IP 首部
12	1	缺少必须的选项
13	0	时间戳请求（作废不用）
14	0	时间戳应答（作废不用）
15	0	信息请求（作废不用）

表 1 ICMP 消息类型代码对应表

类型 Type	代码 Code	描述
16	0	信息应答（作废不用）
17	0	地址掩码请求
18	0	地址掩码应答

报文示例

图 3 封装了 ICMP 消息的 IP 头部格式示例



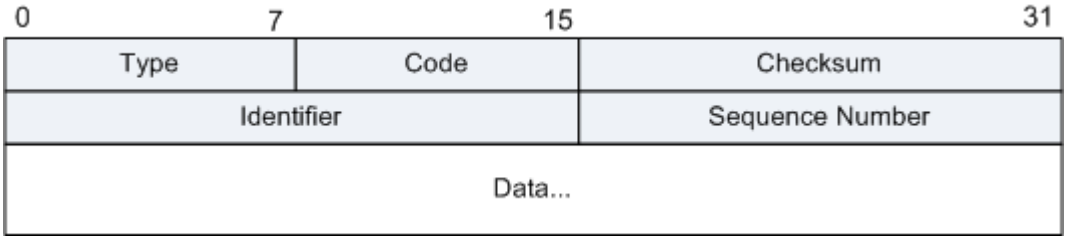
参考标准

标准	描述
RFC 792	Internet Control Message Protocol

4.3.2 ICMP Echo Request/Reply 消息格式

报文格式

图 1 ICMP Echo Request/Reply 消息格式



字段	长度	含义
Type	1 字节	消息类型： • 0：回显应答报文 • 8：请求回显报文
Code	1 字节	消息代码，此处值为 0。
Checksum	2 字节	检验和。
Identifier	2 字节	标识符，发送端标示此发送的报文
Sequence Number	2 字节	序列号，发送端发送的报文的顺序号。每发送一次顺序号就加 1。
Data	可变	选项数据，是一个可变长的字段，其中包含要返回给发送者的数据。回显应答通常返回与所收到的数据完全相同的数据。

报文示例

图 2 ICMP 请求消息

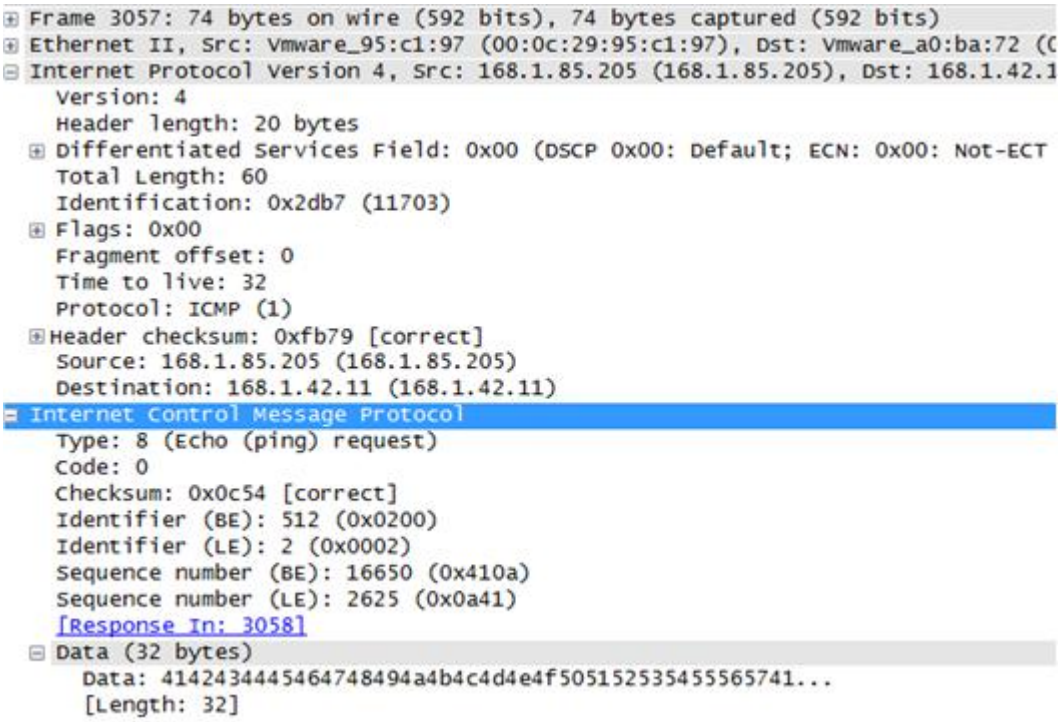


图 3 ICMP 应答消息

```
⊞ Frame 3058: 74 bytes on wire (592 bits), 74 bytes captured (592 bits)
⊞ Ethernet II, Src: Vmware_a0:ba:72 (00:0c:29:a0:ba:72), Dst: Vmware_95:c1:97 (0(
⊟ Internet Protocol Version 4, Src: 168.1.42.11 (168.1.42.11), Dst: 168.1.85.205
    Version: 4
    Header length: 20 bytes
  ⊞ Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00: Not-ECT
    Total Length: 60
    Identification: 0x6f6c (28524)
  ⊞ Flags: 0x00
    Fragment offset: 0
    Time to live: 128
    Protocol: ICMP (1)
  ⊞ Header checksum: 0xfb79 [correct]
    Source: 168.1.42.11 (168.1.42.11)
    Destination: 168.1.85.205 (168.1.85.205)
⊟ Internet Control Message Protocol
    Type: 0 (Echo (ping) reply)
    Code: 0
    Checksum: 0x1454 [correct]
    Identifier (BE): 512 (0x0200)
    Identifier (LE): 2 (0x0002)
    Sequence number (BE): 16650 (0x410a)
    Sequence number (LE): 2625 (0x0a41)
    [Response To: 3057]
    [Response Time: 0.443 ms]
  ⊟ Data (32 bytes)
      Data: 4142434445464748494a4b4c4d4e4f505152535455565741...
      [Length: 32]
```

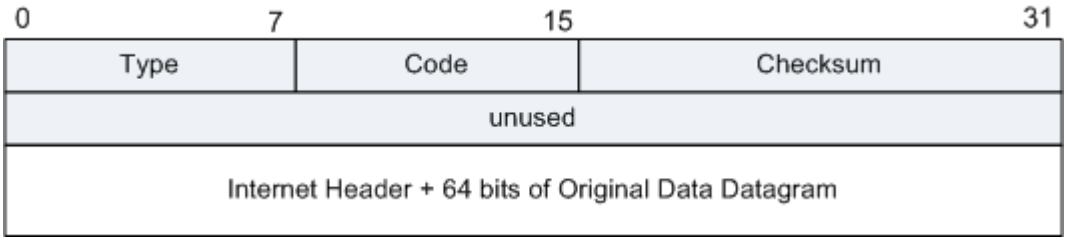
参考标准

标准	描述
RFC 792	Internet Control Message Protocol

4.3.3 ICMP 目的不可达消息格式

报文格式

图 1 ICMP 目的不可达消息格式

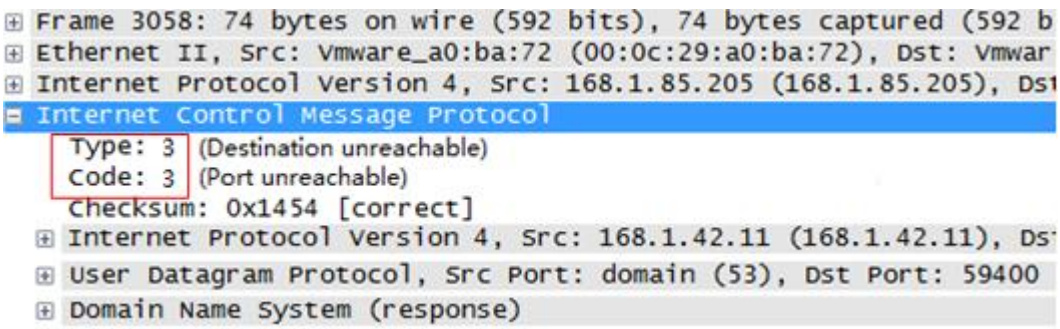


字段	长度	含义
Type	1 字节	消息类型，此处值为 3。
Code	1 字节	消息代码： • 0 = net unreachable;网络不可达 • 1 = host unreachable;主机不可达

字段	长度	含义
		• 2 = protocol unreachable;协议不可达 • 3 = port unreachable; 端口不可达，Tracert 时发送的 ICMP 报文即为此类。 • 4 = fragmentation needed and DF set;需要进行分片但设置不分片比特 • 5 = source route failed.源站选路失败 • 6 = Destination network unknown 目的网络不认识 • 7 = Destination host unknown 目的主机不认识 • 8 = Source host isolated (obsolete)源主机被隔离（作废不用） • 9 = Destination network administratively prohibited 目的网络被强制禁止 • 10 = Destination host administratively prohibited 目的主机被强制禁止 • 11 = Network unreachable for TOS 由于 TOS，网络不可达 • 12 = Host unreachable for TOS 由于 TOS，主机不可达 • 13 = Communication administratively prohibited by filtering 由于过滤，通信被强制禁止 • 14 = Host precedence violation 主机越权 • 15 = Precedence cutoff in effect 优先权中止生效
Checksum	2 字节	检验和。
unused	4 字节	未使用，必须填 0。
Internet Header + 64 bits of Original Data Datagram	可变	IP 首部+原始数据包的前 8 字节： • IP 首部：如果 IP 首部没有选项字段时为 20 字节 • 原始数据包的前 8 字节：UDP 首部的 8 字节或者 TCP 首部的 8 字节。 该数据是主机用来匹配消息。对于更高层协议的用户端口号，原始数据包的前 64 比特的这些数据会被重组。

报文示例

图 2 ICMP 目的不可达消息



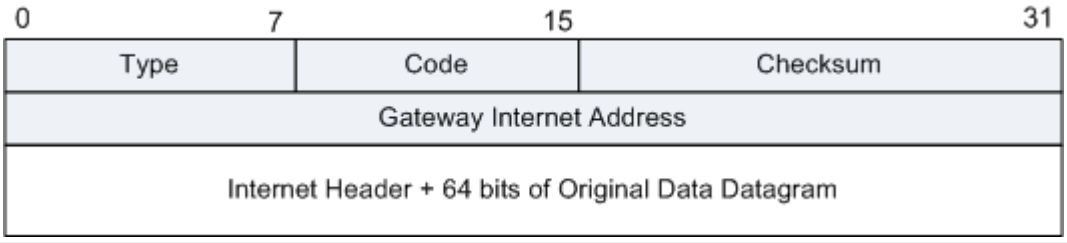
参考标准

标准	描述
RFC 792	Internet Control Message Protocol

4.3.4 ICMP 重定向消息格式

报文格式

图 1 ICMP 重定向消息格式



字段	长度	含义
Type	1 字节	消息类型，此处值为 5。
Code	1 字节	消息代码： • 0 = Redirect datagrams for the Network • 1 = Redirect datagrams for the Host. • 2 = Redirect datagrams for the Type of Service and Network. • 3 = Redirect datagrams for the Type of Service and Host.
Checksum	2 字节	检验和。
Gateway Internet Address	4 字节	即原始数据包里的 IP 目的地址域。
Internet Header + 64 bits of Original Data Datagram	可变	IP 头和原始数据包的前 64 比特数据。该数据是主机用来匹配消息。对于更高层协议的用户端口号，原始数据包的前 64 比特的这些数据会被重组。

报文示例

图 2 ICMP 重定向消息

```
⊞ Frame 3058: 74 bytes on wire (592 bits), 74 bytes captured (592 bits)
⊞ Ethernet II, Src: Vmware_a0:ba:72 (00:0c:29:a0:ba:72), Dst: Vmware_95:c1:
⊞ Internet Protocol Version 4, Src: 123.1.1.3(123.1.1.3), Dst: 168.1.85.205
⊟ Internet Control Message Protocol
    Type: 5 (Redirect)
    Code: 1 (Redirect for host)
    Checksum: 0x01f6 [correct]
    Gateway address:123.1.1.2 (123.1.1.2)
  ⊟ Internet Protocol, Src: 123.1.1.3(123.1.1.3), Dst: 24.1.1.4 (24.1.1.4)
      Version:4
      Header length: 20 bytes
    ⊞ Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00)
      Total Length: 100
      Identification: 0x0041 (65)
    ⊞ Flags: 0x00
      Fragment offset: 0
      Time to live: 254
      Protocol: ICMP (1)
    ⊞ Header checksum: 0x274f [correct]
      Source: 123.1.1.3 (123.1.1.3)
      Destination: 168.1.85.205 (168.1.85.205)
    ⊟ Internet Control Message Protocol
        Type: 8 (Echo (ping) request)
        Code: 0
        Checksum: 0x74fa [incorrect, should be 0xf7f2]
        Identifier: 0x000d
        Sequence number: 0 (0x0000)
        Sequence number (LE):  0  (0x0000)
```

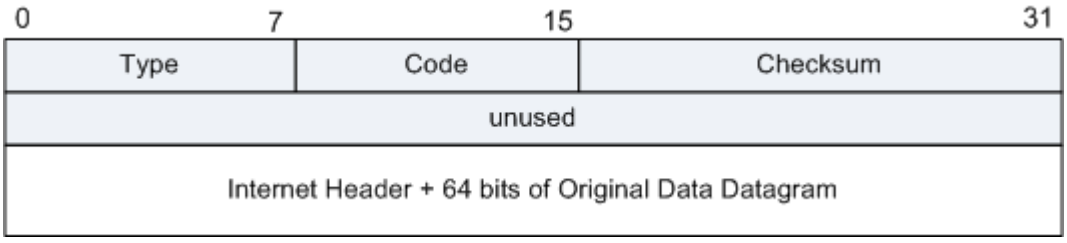
参考标准

标准	描述
RFC 792	Internet Control Message Protocol

4.3.5 ICMP 超时消息格式

报文格式

图 1 ICMP 超时消息格式



字段	长度	含义
Type	1 字节	消息类型，此处值为 3。
Code	1 字节	消息代码：

字段	长度	含义
		• 0 = time to live exceeded in transit • 1 = fragment reassembly time exceeded
Checksum	2 字节	检验和。
Internet Header + 64 bits of Original Data Datagram	可变	IP 头和原始数据包的前 64 比特数据。该数据是主机用来匹配消息。对于更高层协议的用户端口号，原始数据包的前 64 比特的这些数据会被重组。

报文示例

图 2 ICMP 超时消息

```
⊞ Frame 1232: 70 bytes on wire (560 bits), 70 bytes captured (560 bits)
⊞ Ethernet II, Src: HuaweiTe_9f:24:cb (4c:1f:cc:9f:24:cb), Dst: 20:0b:c7
⊞ Internet Protocol, Src:172.1.1.2 (172.1.1.2), Dst: 192.168.1.212(192.1
⊟ Internet Control Message Protocol
    Type: 11 (Time-to-live exceeded)
    Code: 1 (Fragment reassembly time exceeded)
    Checksum: 0x0c54 [correct]
  ⊞ Internet Protocol, Src:192.168.1.212(192.168.1.212), Dst: 172.1.1.2
  ⊟ Internet Control Message Protocol
      Type: 0 (Echo (ping) reply)
      Code: 0
      Checksum: 0x1454 [incorrect, should be 0xffff]
      Identifier (BE): 512 (0x0200)
      Sequence number (BE): 16650 (0x410a)
```

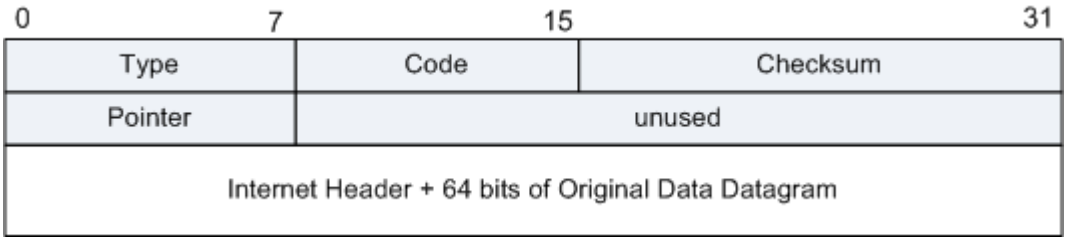
参考标准

标准	描述
RFC 792	Internet Control Message Protocol

4.3.6 ICMP 参数问题消息格式

报文格式

图 1 ICMP 参数问题消息的格式

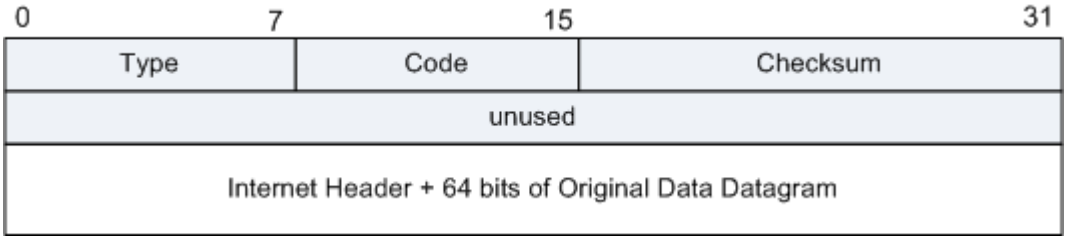


字段	长度	含义
Type	1 字节	消息类型，此处值为 12。
Code	1 字节	消息代码： • 0 = pointer indicates the error.
Checksum	2 字节	检验和。
Pointer	1 字节	标识出报文中出现错误地方的 8 位片偏移量。
Internet Header + 64 bits of Original Data Datagram	可变	IP 头和原始数据包的前 64 比特数据。该数据是主机用来匹配消息。对于更高层协议的用户端口号，原始数据包的前 64 比特的这些数据会被重组。

4.3.7 ICMP 源端被关闭消息格式

报文格式

图 1 ICMP 源端被关闭消息格式



字段	长度	含义
Type	1 字节	消息类型，此处值为 4。
Code	1 字节	消息代码，此处值为 0。
Checksum	2 字节	检验和。

字段	长度	含义
Internet Header + 64 bits of Original Data Datagram	可变	IP 头和原始数据包的前 64 比特数据。该数据是主机用来匹配消息。对于更高层协议的用户端口号，原始数据包的前 64 比特的这些数据会被重组。

4.4 ICMPv6 报文格式

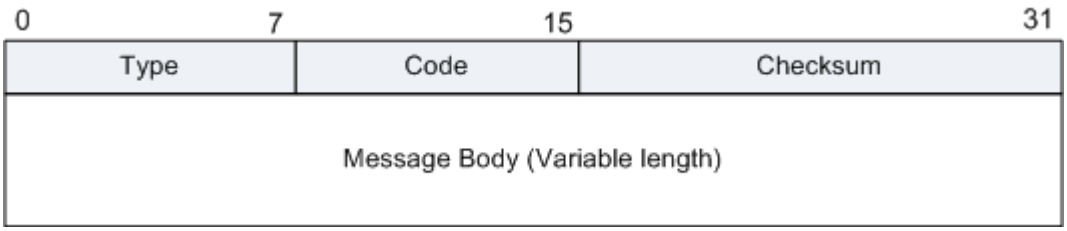
- ICMPv6 报文通用格式
- ICMPv6 回显请求/应答消息
- ICMPv6 目的不可达消息
- ICMPv6 重定向消息
- ICMPv6 超时消息
- ICMPv6 参数错误消息
- ICMPv6 路由器请求（Router Solicitation）消息
- ICMPv6 路由器通告消息
- ICMPv6 邻居请求（Neighbor Solicitation）消息
- ICMPv6 邻居通告消息

4.4.1 ICMPv6 报文通用格式

报文格式

ICMPv6 报文有如下的通用格式：

图 1 ICMPv6 报文通用格式



字段	长度	描述
Type	1 字节	定义了报文的类型，该字段决定了其它部分的报文格式。
Code	1 字节	该字段依赖 TYPE 字段，在 TYPE 字段的基础上，它被用来在基本类型上创建更详细的报文等级。
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
Message Body	可变	大体上说，ICMPv6 报文可以被分为 2 大类：差错报文与消息报文。这 2 类报文是依靠报文中的 TYPE 字段来标识的，当 TYPE 字段的最高位置 0，即在 0～127 范围时（TYPE 字段长度为 1 字节），被标识为差错报文，TYPE 字段值为 128～255 范围时，则标识为消息报文。

表 1 ICMPv6 报文类型对应表

Type	Code	消息名
1	0	没有路由到达目的地
1	1	与目的地的通信由于管理被禁止
1	2	超过了源地址的范围
1	3	地址不可达
1	4	端口不可达
1	5	源地址的入口/出口策略失败
1	6	拒绝路由到达目的地
2	0	包太大
3	0	传输过程中“hop-limit”超时
3	1	分片重组超时

字段	长度	描述
4	0	参数错误
4	1	错误的首部字段
4	2	不可识别的 Next Header 类型
4	3	不可识别的 IPv6 选项
100	x	私有实验用
101	x	私有实验用
127	x	ICMPv6 差错报文扩展保留
128	0	回显请求
129	0	回显应答
133	x	路由请求
134	x	路由通告
135	x	邻居请求
136	x	邻居通告
137	x	重定向
143	x	MLDv2 (Multicast Listener Report Message v2)
200	x	私有实验用
201	x	私有实验用
255	x	ICMPv6 消息报文扩展保留

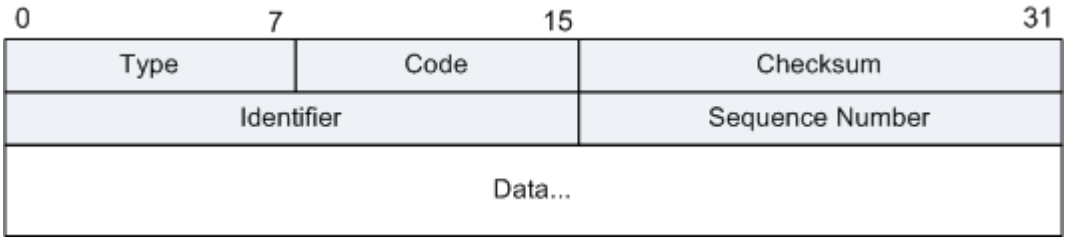
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

4.4.2 ICMPv6 回显请求/应答消息

报文格式

图 1 ICMPv6 回显请求/应答消息的格式



字段	长度	含义
Type	1 字节	消息类型： • 128: Echo Request • 128: Echo Reply
Code	1 字节	消息代码，此处值为 0。
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
Identifier	4 字节	请求与应答报文能够彼此匹配的一个标识，可能是全 0；
Sequence Number	4 字节	请求与应答报文能够彼此匹配的一个标识，可能是全 0；
Data	变长	0 或任意数据的 8 位组。

报文示例

图 2 ICMPv6 Echo Request Message

```
⊞ Frame 108: 118 bytes on wire (944 bits), 118 bytes captured (944
⊞ Ethernet II, Src: 00:46:4b:d8:28:c7 (00:46:4b:d8:28:c7), Dst: Rea
⊞ Internet Protocol Version 6, Src: 2008::246:4bff:fed8:28c7 (2008:
⊟ Internet Control Message Protocol v6
    Type: Echo (ping) request (128)
    Code: 0
    Checksum: 0x97ca [correct]
    Identifier: 0xabd0
    Sequence: 1
    [Response In: 109]
  ⊟ Data (56 bytes)
      Data: 043aa29273e44c78504e4701043a037f0001020304050607...
      [Length: 56]
```

图 3 ICMPv6 Echo Reply Message

```
⊞ Frame 109: 118 bytes on wire (944 bits), 118 bytes captured (944 bits
⊞ Ethernet II, Src: RealtekS_88:5a:81 (00:e0:4c:88:5a:81), Dst: 00:46:4
⊞ Internet Protocol Version 6, Src: 2008::230 (2008::230), Dst: 2008::2
⊟ Internet Control Message Protocol v6
    Type: Echo (ping) reply (129)
    Code: 0
    Checksum: 0x96ca [correct]
    Identifier: 0xabd0
    Sequence: 1
    [Response To: 108]
    [Response Time: 0.025 ms]
  ⊟ Data (56 bytes)
      Data: 043aa29273e44c78504e4701043a037f0001020304050607...
      [Length: 56]
```

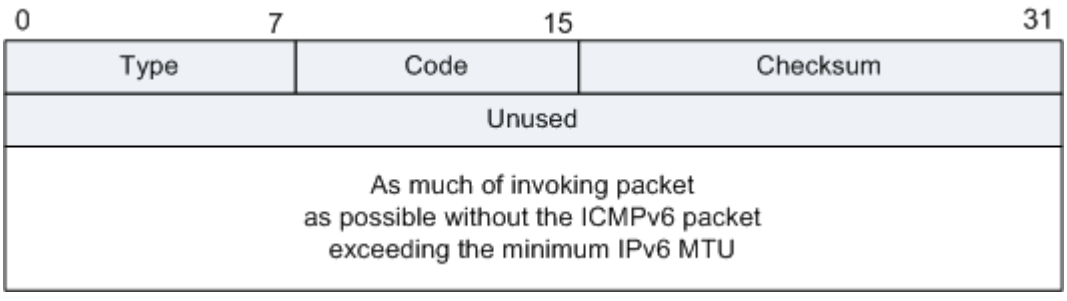
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

4.4.3 ICMPv6 目的不可达消息

报文格式

图 1 ICMPv6 目的不可达消息的格式



字段	长度	含义

字段	长度	含义
Type	1 字节	消息类型，此处值为 1。
Code	1 字节	消息代码： • 0 - 没有到目的地址的路由 • 1 - 禁止与目的地址通讯 • 2 - 超出源地址的范围 • 3 - 地址不可达 • 4 - 端口不可达 • 5 - 源地址入口/出口策略失败 • 6 - 拒绝到目的地址的路由
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
unused	4 字节	该字段对所有的 code 值均未使用。在报文的发送端，该字段必须被初始化为 0，且在接收端，该字段应该被忽略。

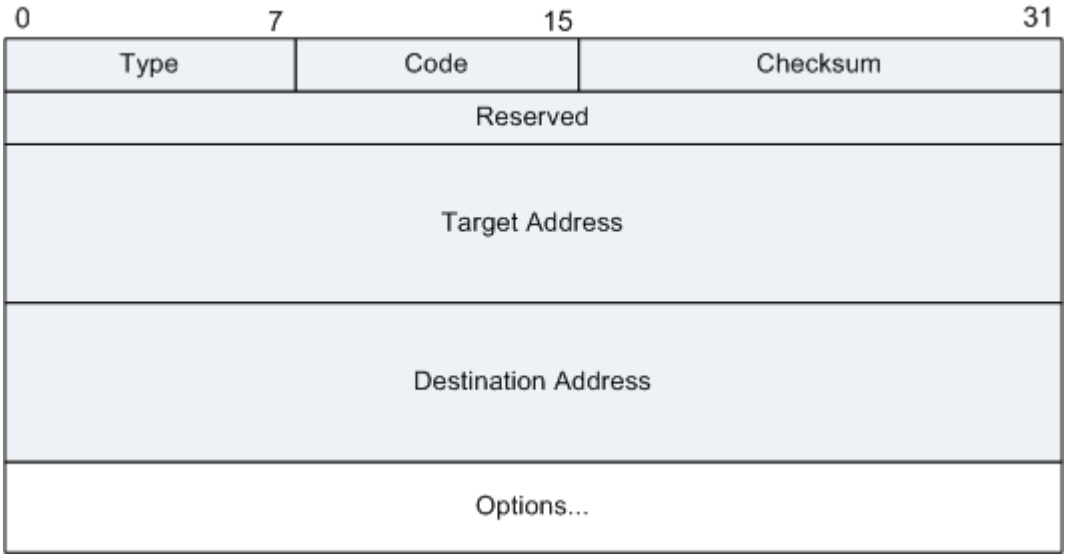
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

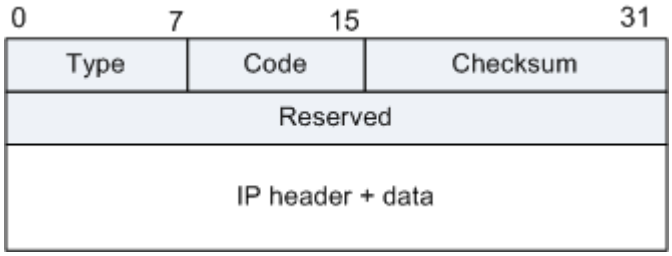
4.4.4 ICMPv6 重定向消息

报文格式

图 1 ICMPv6 重定向消息的格式



字段	长度	含义
Type	1 字节	消息类型，此处值为 137。
Code	1 字节	该 ICMPv6 差错报文的始发者必须将该字段置为 0，且接收端忽略该字段。
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
Reserved	4 字节	此字段未使用。它必须由发送者初始化为 0，接收者必须忽略它。
Target Address	16 字节	更好的下一跳地址。当目标是实际通信端点时，即，目的地是邻居，Target Address 字段必须包括与 ICMP Destination Address 字段相同的值。其他情况，目标是更好的第一跳路由器并且 Target Address 必须是该路由器的链路本地地址，以便主机能够唯一地识别路由器。
Destination Address	16 字节	重定向到目标的目的地的 IP 地址。
Options	可变	选项，TLV 格式。 • Source link-layer address: 目标链路层地址。该目标的链路层地址。应当包括它(如果知道)在内。注意，在 NBMA 链路上，主机们或许根据 Redirect 消息中 Target Link-Layer Address 选项的存在，作为确定邻居们的链路层地址的方法。在此情况，此选项必须包括在 Redirect 消息中。 TLV 格式字段含义： ▪ Type：1 字节，Source Link-layer Address 的取值为 1。

字段	长度	含义
		▪ Length:1 字节,选项的长度(包括类型字段和长度字段)以 8 字节为单位计算。例如,IEEE802 地址的长度是 1。 ▪ Link-Layer Address：可变长度的链路层地址。此字段的内容和形式(包括字节和比特顺序)一般由描述 IPv6 在不同链路层上如何运行的特定文件中规定。 • Redirected Header: 不造成重定向分组超过在 IPv6 协议规定的最小 MTU 情况下，尽可能多地触发发送 Redirect 的 IP 分组。其他 Neighbor Discovery 消息必须忽略此选项。 格式如下： **图 2** Redirected Header 字段格式  ▪ Type: = 4. ▪ Length: 此选项的长度，以 8 字节为单位。 ▪ Reserved: 这些字段未使用。它们必须被发送者初始化为 0，接收者必须忽略它们。 ▪ IP header + data: 原始分组被截短,以便确保重定向消息大小不超过 IPv6 要求的最小 MTU。

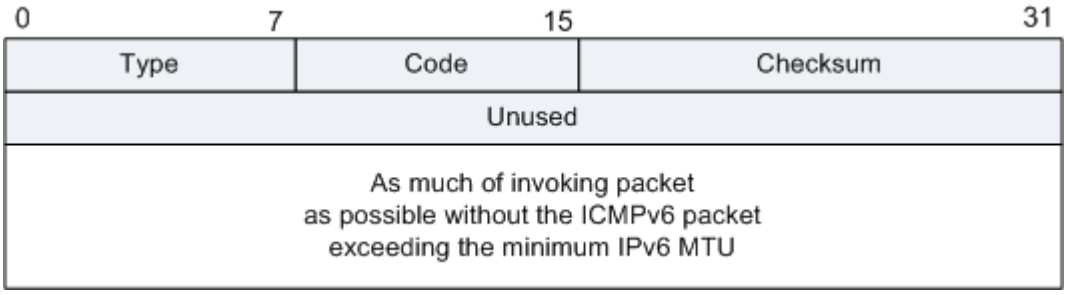
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

4.4.5 ICMPv6 超时消息

报文格式

图 1 ICMPv6 超时消息的格式



字段	长度	含义
Type	1 字节	消息类型，此处值为 2。
Code	1 字节	• 0 - Hop limit exceeded in transit 传输过程中“hop-limit”超时； • 1 - Fragment reassembly time exceeded 分片重组超时；
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
unused	4 字节	该字段对所有的 code 值均未使用。在报文的发送端，该字段必须被初始化为 0，且在接收端，该字段应该被忽略。

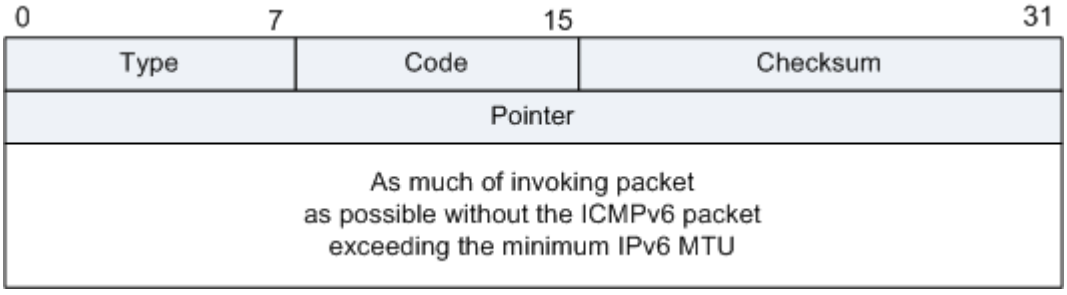
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

4.4.6 ICMPv6 参数错误消息

报文格式

图 1 ICMPv6 参数错误消息的格式



字段	长度	含义

字段	长度	含义
Type	1 字节	消息类型，此处值为 4。
Code	1 字节	• 0 - 错误的首部字段； • 1 - 不可识别的 Next Header 类型； • 2 - 不可识别的 IPv6 选项。
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
Pointer	4 字节	标识出报文中出现错误地方的 8 位片偏移量；若原报文中出现错误的地方在 ICMPv6 差错报文达到最大长度时也不能被包括在内，指针的值将超过 ICMPv6 差错报文的长度。

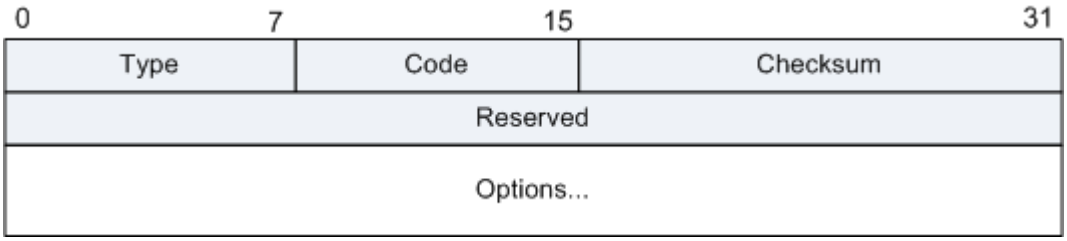
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

4.4.7 ICMPv6 路由器请求（Router Solicitation）消息

报文格式

图 1 ICMPv6 路由器请求消息的格式

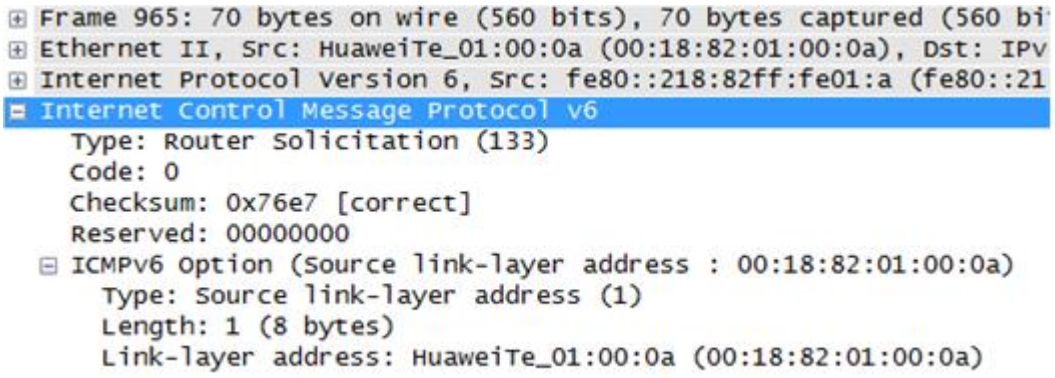


字段	长度	含义
Type	1 字节	消息类型，此处值为 133。
Code	1 字节	该 ICMPv6 差错报文的始发者必须将该字段置为 0，且接收端忽略该字段。
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。

字段	长度	含义
Reserved	4 字节	此字段不使用。它必须由发送者初始化为 0，接收者必须忽略它。
Options	可变	选项 • 源链路层地址：发送者的链路层地址，如果知道。如果 Source Address 是未指定地址，必须不 包括在内。否则，有地址的链路层上应当包括源链路层地址。 为 TLV 格式，各字段含义如下： ▪ Type: = 1，长度是 1 字节。 ▪ Length: 1 字节，标识选项的长度(包括类型字段和长度字段)以 8 字节为单位计算。例如，IEEE802 地址的长度是 1。Length 值最少为 1，0 为非法值，须丢弃。 ▪ Link-Layer Address: 可变长度的链路层地址。此字段的内容和形式(包括字节和比特顺序)一般由描述 IPv6 在不同链路层上如何运行的特定文件中规定。

报文示例

图 2 ICMPv6 路由器请求消息



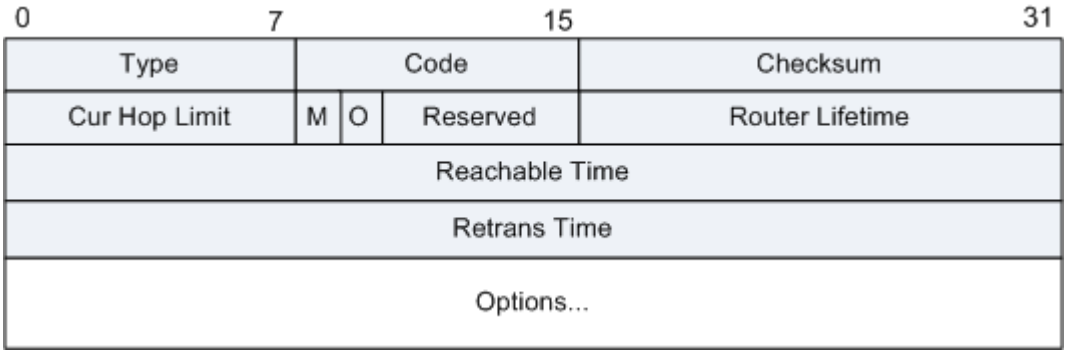
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

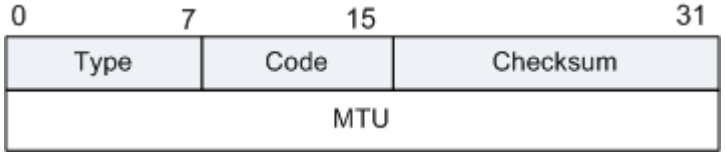
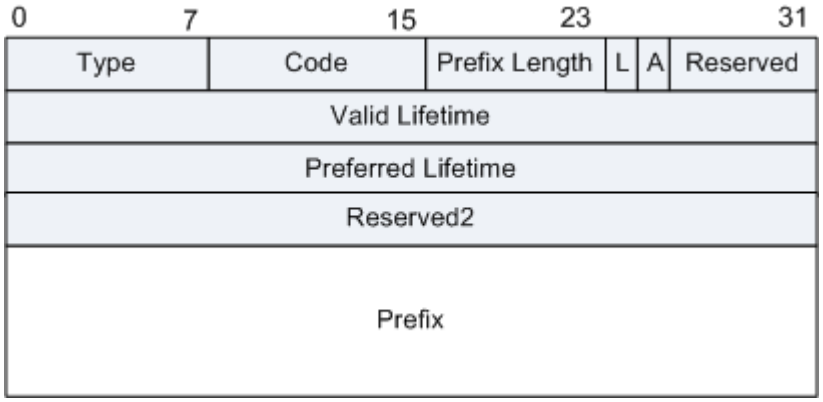
4.4.8 ICMPv6 路由器通告消息

报文格式

图 1 ICMPv6 路由器通告消息的格式



字段	长度	含义
Type	1 字节	消息类型，此处值为 134。
Code	1 字节	该 ICMPv6 差错报文的始发者必须将该字段置为 0，且接收端忽略该字段。
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
Cur Hop Limit	1 字节	8 位无符号整数。默认值应当放置在发出 IP 分组的 IP 首部的 Hop Count 字段中。 取 0 值意味着未(由该路由器)规定。
M	1 比特	1 位“管理地址配置”标记。当置 1 时，它指出地址可通过 Dynamic Host Configuration 协议获得。 如果 M 标记置 1，则 O 标记为冗余，可以忽略，因为 DHCPv6 将返回所有可用配置信息。
O	1 比特	1 位“其他配置”标记。 当 M= 如果 M=
Reserved	6 比特	6 位未使用字段。它必须由发送者初始化为 0，接收者必须忽略它。
Router Lifetime	2 字节	16 位无符号整数。与默认路由器关联的生存期，以秒为单位。最大值 18.2 小时。取 0 值的 Lifetime 指出路由器不是默认路由器并且不应当出现在默认路由器列表中。Router Lifetime 仅适用于作为默认路由器的路由器应用；对包括在其他消息字段或选项中的信息不适用。需要对它们的信息规定时间限制的选项有它们自己的生存期字段。
Reachable Time	4 字节	32 位无符号整数。此时间以毫秒计，在收到可达性确认后节点假定该邻居是可到达的。它由 Neighbor Unreachability Detection 算法使用(参阅第 7-3 节)。此值为 0 意味着没有(由此路由器)作出规定。
Retrans Timer	4 字节	32 位无符号整数。重发的 Neighbor Solicitation 消息间隔时间，以毫秒计。由地址解析和 Neighbor Unreachability Detection 算法使用。此值为 0 意味着没有(由此路由器)作出规定。
Options	可变	选项

字段	长度	含义
		• Source link-layer address：源链路层地址，发出 Router Advertisement 的接口的链路层地址。仅在有地址的链路层上使用。路由器可以忽略此选项，以便能够使入境负载跨多个链路层地址共享。 为 TLV 格式，各字段含义如下： ▪ Type: = 1，长度是 1 字节。 ▪ Length: 1 字节，选项的长度(包括类型字段和长度字段)，以 8 字节为单位计算。例如，IEEE802 地址的长度是 1。 ▪ Link-Layer Address: 可变长度的链路层地址。此字段的内容和形式(包括字节和比特顺序)一般由描述 IPv6 在不同链路层上如何运行的特定文件中规定。 • MTU：在有可变 MTU 的链路上应当按此发送流量(正如在描述特定链路类型上如何 运行 IP 的文件中规定的)。可以按此在其他链路上发送流量。 MTU 格式： **图 2** MTU 选项的格式  ▪ Type = 5 ▪ Length = 1 ▪ Reserved: 此字段未使用。它必须被发送者初始化为 0，接收者必须忽略它。 ▪ MTU: 32 位无符号整数。是为此链路推荐的 MTU。 • Prefix Information：这些选项规定了前缀，这些前缀是 on-link 的，和/或被用于地址自动配置。路由器应当包括所有它的 on-link 前缀(链路本地前缀除外)，所以多归属第主机有 完整的前缀信息，这些前缀是关于主机们附着的链路的 on-link 目的地的。如果缺乏完整信息，当发送流量到它的邻居们时，多归属地主机或许不能够选择正确的出接口。 格式如下： **图 3** Prefix 选项格式 

字段	长度	含义
		▪ Type: = 3 ▪ Length: = 4 ▪ Prefix Length: 8 位无符号整数。在合法前缀中领先比特的数目。其值范围是 0 到 128。前缀长度字段为 on-link 确定提供必须的信息(当与前缀信息选项中 L 标记相结合时)。它也帮助实现地址自动配置，对此存在更多关于前缀长度的限制。 ▪ L: 1 位 on-link 标记。当置 1 时，指出此前缀可用于 on-link 确定。当没有置 1 时，通告对此前缀的 on-link 或 off-link 性质没有说明。换句话讲，如果 L 标记没有置 1，主机不能推断出从该前缀引申出的地址是 off-link。即，主机不能更新先前关于地址是 on-link 的指示。 ▪ A: 1 位自动地址配置标记。当置 1 时，指出此前缀可用于无状态地址自动配置。 ▪ Reserved1: 6 位未使用字段。必须被发送者初始化为 0，接收者必须忽略它。 ▪ Valid Lifetime: 32 位无符号整数。时间长度以秒为单位(相对于分组被发送的时间)，在此时间内此前缀对于 on-link 确定来说是合法的。全 1 比特值(0xffffffff)表示无限。 ▪ Preferred Lifetime: 32 位无符号整数。时间长度以秒为单位(相对于分组被发送的时间)。在此时间 内经无状态地址自动配置，根据此前缀生成的地址保有优先权[ADDRCONF]。全 1 比特值(0xffffffff)表示无限。注意，此字段的值不能超过 Valid Lifetime 字段的值，以避免优先的地址不再合法。 ▪ Reserved2: 此字段未使用。它必须被发送者初始化为 0，接收者必须忽略它。 ▪ Prefix: IP 地址或 IP 地址的前缀。Prefix Length 字段包含此前缀中有效领先比特的数目。在前缀中，在前缀长度之后的这些位被保留，并且必须被发送者初始化为 0，接收者必须忽略它们。路由器不应当发送链路本地前缀的前缀选项，主机应当忽略这种前缀选项。

报文示例

图 4 ICMPv6 路由器通告报文

```
⊞ Frame 169: 166 bytes on wire (1328 bits), 166 bytes captured (1328 b
⊞ Ethernet II, Src: RealtekS_88:5a:81 (00:e0:4c:88:5a:81), Dst: IPv6mc
⊞ Internet Protocol Version 6, Src: fe80::2e0:4cff:fe88:5a81 (fe80::2e
⊟ Internet Control Message Protocol v6
    Type: Router Advertisement (134)
    Code: 0
    Checksum: 0xd11a [correct]
    Cur hop limit: 0
  ⊟ Flags: 0x18
      0... .... = Managed address configuration: Not set
      .0.. .... = Other configuration: Not set
      ..0. .... = Home Agent: Not set
      ...1 1... = Prf (Default Router Preference): Low (3)
      .... .0.. = Proxy: Not set
      .... ..0. = Reserved: 0
    Router lifetime (s): 7200
    Reachable time (ms): 0
    Retrans timer (ms): 0
  ⊟ ICMPv6 Option (Source link-layer address : 00:e0:4c:88:5a:81)
      Type: Source link-layer address (1)
      Length: 1 (8 bytes)
      Link-layer address: RealtekS_88:5a:81 (00:e0:4c:88:5a:81)
  ⊟ ICMPv6 Option (MTU : 1500)
      Type: MTU (5)
      Length: 1 (8 bytes)
      Reserved
      MTU: 1500
  ⊟ ICMPv6 Option (Prefix information : fec0:0:0:4::/64)
      Type: Prefix information (3)
      Length: 4 (32 bytes)
      Prefix Length: 64
    ⊟ Flag: 0xc0
        1... .... = On-link flag(L): Set
        .1.. .... = Autonomous address-configuration flag(A): Set
        ..00 0000 = Reserved: 0
      Valid Lifetime: 172800
      Preferred Lifetime: 1800
      Reserved
      Prefix: fec0:0:0:4:: (fec0:0:0:4::)
  ⊞ ICMPv6 Option (Prefix information : 2002:ac00:26e6:4::/64)
  ⊟ ICMPv6 Option (Route Information : Low 2002::/16)
      Type: Route Information (24)
      Length: 2 (16 bytes)
      Prefix Length: 16
    ⊟ Flag: 0x18
        ...1 1... = Route Preference: Low (3)
        000. .000 = Reserved: 0
      Route Lifetime: 7200
      Prefix: 2002:: (2002::)
```

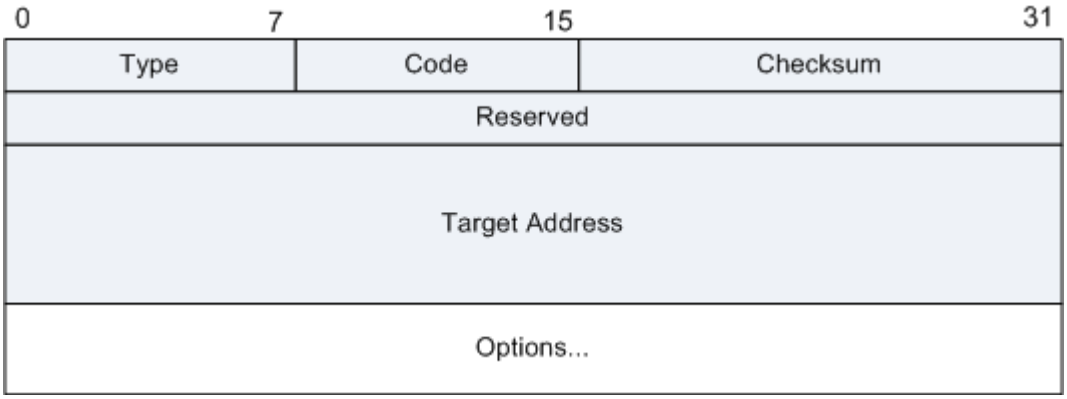
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

4.4.9 ICMPv6 邻居请求（Neighbor Solicitation）消息

报文格式

图 1 ICMPv6 邻居请求消息



字段	长度	含义
Type	1 字节	消息类型，此处值为 135。
Code	1 字节	该 ICMPv6 差错报文的始发者必须将该字段置为 0，且接收端忽略该字段。
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
Reserved	4 字节	此字段不使用。它必须由发送者初始化为 0，接收者必须忽略它。
Target Address	16 字节	请求的目标的 IP 地址。它必须不是多播地址。
Options	可变	选项： • Source link-layer address 源链路层地址：发送者的链路层地址，如果知道。如果 Source Address 是未指定地址，必须不包括在内。否则，有地址的链路层上应当包括源链路层地址。 为 TLV 格式： ▪ Type：1 字节，取值为 1。 ▪ Length：1 字节，表示选项的长度(包括类型字段和长度字段)，以 8 字节为单位计算。例如，IEEE802 地址的长度是 1。 ▪ Link-Layer Address：可变长度的链路层地址。此字段的内容和形式(包括字节和比特顺序)一般由描述 IPv6 在不同链路层上如何运行的特定文件中规定。

报文示例

图 2 ICMPv6 邻居请求报文示例

```
⊞ Frame 164: 86 bytes on wire (688 bits), 86 bytes captured (688 bit
⊞ Ethernet II, Src: 00:46:4b:d8:28:c7 (00:46:4b:d8:28:c7), Dst: Real
⊞ Internet Protocol Version 6, Src: fe80::246:4bff:fed8:28c7 (fe80::
⊟ Internet Control Message Protocol v6
    Type: Neighbor Solicitation (135)
    Code: 0
    Checksum: 0x49e8 [correct]
    Reserved: 00000000
    Target Address: 2008::230 (2008::230)
  ⊟ ICMPv6 Option (Source link-layer address : 00:46:4b:d8:28:c7)
      Type: Source link-layer address (1)
      Length: 1 (8 bytes)
      Link-layer address: 00:46:4b:d8:28:c7 (00:46:4b:d8:28:c7)
```

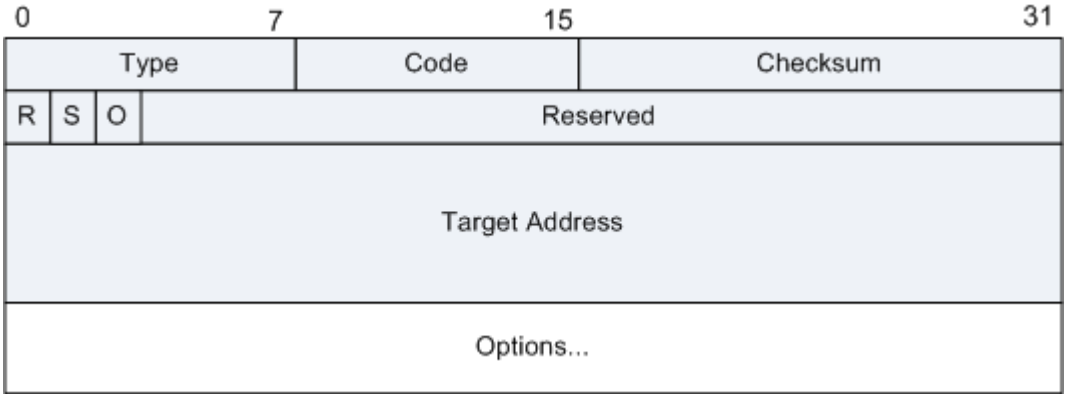
参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

4.4.10 ICMPv6 邻居通告消息

报文格式

图 1 ICMPv6 邻居通告消息的格式



字段	长度	含义
Type	1 字节	消息类型，此处值为 136。
Code	1 字节	该 ICMPv6 差错报文的始发者必须将该字段置为 0，且接收端忽略该字段。
Checksum	2 字节	用来在 ICMPv6 报文中检验数据和部分 IPv6 首部的完整性。
R	1 比特	路由器标记。当置 1 时，R 位指出发送者是路由器。R 位由 Neighbor Unreachability Detection 使用，用

字段	长度	含义
		于检测改变为主机的路由器。
S	1 比特	请求标记。当置 1 时，S 位指出通告被发送以响应来自目的地地址的 Neighbor Solicitation。S 位用作 Neighbor Unreachability Detection 的可达性确认。在多播通告和非请求单播通告中置 0。
0	1 比特	替代标记。替代标志，1 表示通告中的信息替代缓存，如更新链路层地址时，对于任播的回应则不应置位。在针对任播地址的请求通告中，以及在请求的前缀通告中它不能被置 1。在其他请求通告中和在非请求通告中它应当被置 1。
Reserved	29 比特	29 位未使用字段。它必须由发送者初始化为 0，接收者必须忽略它。
Target Address	16 字节	对于请求的通告，是在 Neighbor Solicitation 消息(该消息催促这个通告)中的 Target Address 字段。对于非请求通告，是其链路层地址已经改变的地址。Target Address 必须不是多播地址。
Options	可变	选项： • Target link-layer address: 目标的链路层地址，即，通告发送者。当响应多播请求时，在有地址的链路层上必须包括此选项。当响应单播 Neighbor Solicitation 时应当包括此选项。 当对端节点由于没有缓存条目从而不能返回一个 Neighbor Advertisements 消息时，为了避免无休止的 Neighbor Solicitation“递归”，对于多播请求必须包括 此选项。当响应单播请求时，可忽略此选项，因为请求的发送者有正确的链路层地址；其他情况，此选项不能在第一位置发送单播请求。然而，在此情况，包括链路层地址仅增加了少许开销，却消除了潜在的竞争条件，那里在收到对先前的请求的响应之前，发送者删除缓存的链路层地址。 为 TLV 格式： ▪ Type: = ▪ Length: 1 字节，表示选项的长度(包括类型字段和长度字段)，以 8 字节为单位计算。例如，IEEE802 地址的长度是 1。 ▪ Link-Layer Address: 可变长度的链路层地址。此字段的内容和形式(包括字节和比特顺序)一般由描述 IPv6 在不同链路层上如何运行的特定文件中规定。

报文示例

图 2 Example for ICMPv6 Neighbor Advertisement MessageICMPv6 邻居通告报文示例

```
⊞ Frame 165: 86 bytes on wire (688 bits), 86 bytes captured (688 bits
⊞ Ethernet II, Src: RealtekS_88:5a:81 (00:e0:4c:88:5a:81), Dst: 00:46
⊞ Internet Protocol Version 6, Src: 2008::230 (2008::230), Dst: fe80:
⊟ Internet Control Message Protocol v6
    Type: Neighbor Advertisement (136)
    Code: 0
    Checksum: 0x34e3 [correct]
  ⊟ Flags: 0xe0000000
      1... .... .... .... .... .... .... .... = Router: Set
      .1.. .... .... .... .... .... .... .... = Solicited: Set
      ..1. .... .... .... .... .... .... .... = Override: Set
      ...0 0000 0000 0000 0000 0000 0000 0000 = Reserved: 0
    Target Address: 2008::230 (2008::230)
  ⊟ ICMPv6 Option (Target link-layer address : 00:e0:4c:88:5a:81)
      Type: Target link-layer address (2)
      Length: 1 (8 bytes)
      Link-layer address: RealtekS_88:5a:81 (00:e0:4c:88:5a:81)
```

参考标准

标准	描述
RFC 4443	Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification

4.5 IGMP 报文格式

IGMP 消息封装在 IP 报文中。

IP 报文头的协议类型字段值为 2，用来标识数据部分封装了 IGMP 消息。

IP 报文头的目的地址字段用来标识该 IGMP 消息的目的接收端。

IP 报文头的 TTL 字段值为 1，表示 IGMP 消息只在本地网段传播。

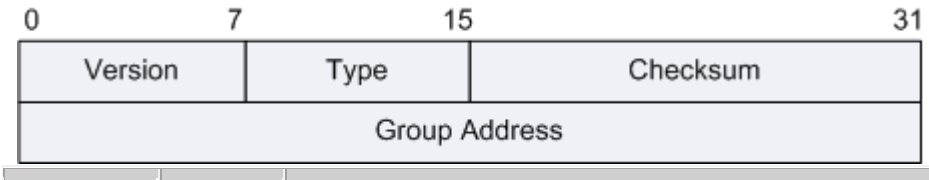
IGMP 有三个版本，不同版本报文格式不一样。

不同版本的 IGMP 协议，支持的 IGMP 消息也不同。

4.5.1 IGMPv1 报文格式

报文格式

图 1 IGMPv1 报文格式

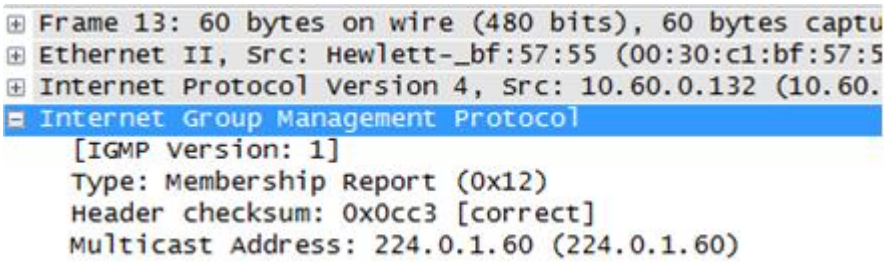


字段	长度	描述

字段	长度	描述
Version	4 比特	IGMP 版本号，在 IGMPv1 中应为 0x1。
Type	4 比特	即 IGMP 报文类型： • 1 = Host Membership Query 主机成员查询 • 2 = Host Membership Report 主机成员报告
Unused	8 比特	未使用的字段，发送时必须填 0，接收时忽略。
Checksum	16 比特	IGMP 消息的校验和。该字段在进行校验计算时设为 0。当传送报文的时候，必须计算该校验字并插入到该字段中去。当接收包的时候，该校验字必须在处理该包之前进行检验。
Group Address	32 比特	组播地址。

报文示例

图 2 IGMPv1 Membership Report



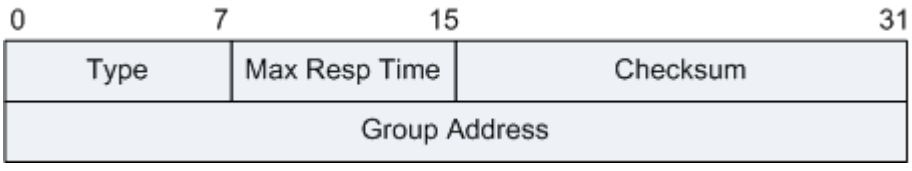
参考标准

标准	描述
RFC 1112	Host extensions for IP multicasting

4.5.2 IGMPv2 报文格式

报文格式

图 1 IGMPv2 报文格式



字段	长度	描述
Type	8 比特	报文类型，有以下几种类型： • 0x11 = Membership Query IGMP 查询消息。 • 0x12 = Version 1 Membership Report IGMPv1 成员报告消息。 • 0x16 = Version 2 Membership Report IGMPv2 成员报告消息。 • 0x17 = Leave Group 离开消息。 在 IGMP 版本 2 中，旧的 4 位版本字段和旧的 4 位类型字段拼成了一个新的 8 位类型字段，通过分别将成员查询（版本 1 和版本 2 的）及版本 1 的成员报告报文的 IGMP 版本 2 的类型代码置为 0x11 和 0x12，保持了 IGMP 版本 1 和版本 2 包格式的向后兼容。
Max Resp Time	8 比特	在发出响应报告前的以 1/10 秒为单位的最长时间，缺省值为 10 秒。 新的最大响应时间（以 1/10 秒为单位）字段允许查询用路由器为它的查询报文指定准确的查询间隔响应时间。IGMP 版本 2 主机在随机选择它们的响应时间值时以此作为上限。 这样在查询响应间隔时有助于控制响应的爆发。
Checksum	16 比特	IGMP 消息的校验和。传送报文时，必须计算校验和并填入该字段中；接收报文时，必须在处理报文之前检验校验和，以判断 IGMP 消息在传输过程中是否发生了错误。
Group Address	32 比特	组播组地址（如果是通用查询则为 0.0.0.0）。 除了在通用查询时这一字段置为 0.0.0.0 外，这一字段和 IGMP 版本 1 中的这一字段意义相同。

报文示例

图 2 IGMPv2 成员查询消息

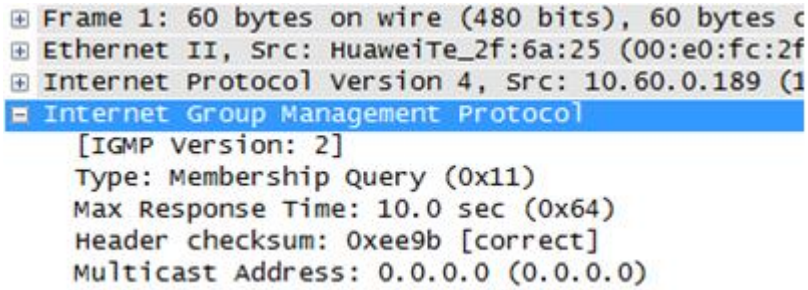


图 3 IGMPv2 成员报告消息

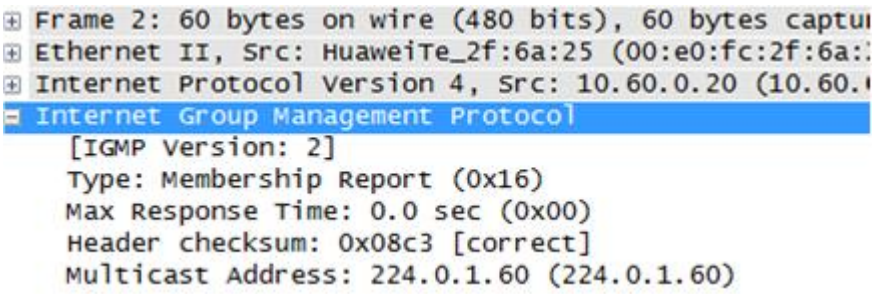


图 4 IGMPv2 离开组消息

```
⊞ Frame 12: 64 bytes on wire (512 bits), 64 bytes captured (
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_9c:68:e4 (08:19:a
⊞ Internet Protocol Version 4, Src: 0.0.0.0 (0.0.0.0), Dst:
⊟ Internet Group Management Protocol
    [IGMP Version: 2]
    Type: Leave Group (0x17)
    Max Response Time: 0.0 sec (0x00)
    Header checksum: 0xf897 [correct]
    Multicast Address: 239.255.0.104 (239.255.0.104)
```

参考标准

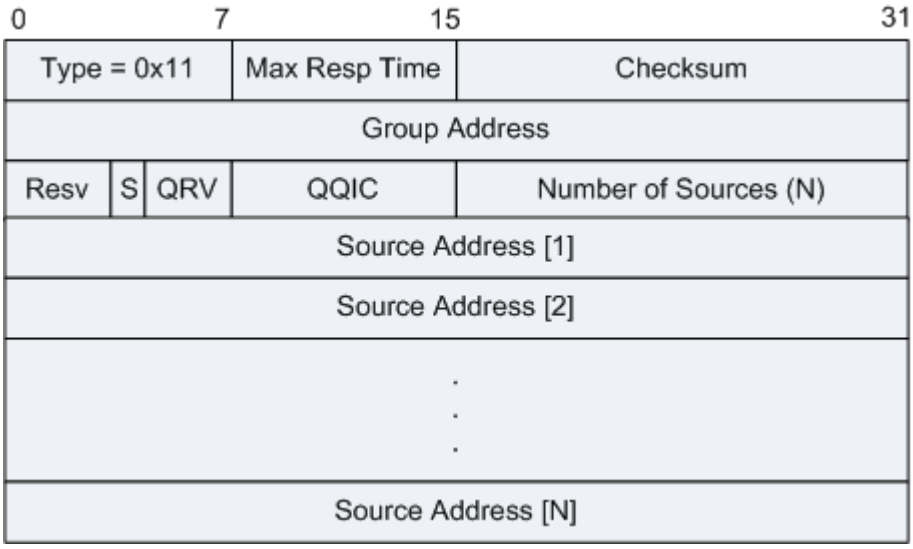
标准	描述
RFC 2236	Internet Group Management Protocol, Version 2

4.5.3 IGMPv3 报文格式

报文格式

IGMPv3 包含查询报文和报告报文两种不同格式的报文。

图 1 IGMPv3 查询报文格式



字段	长度	描述
Type	8 比特	成员关系查询 Type = 0x11。
Max Resp Code	8 比特	设备接收到查询消息后发出响应报文的最大延迟时间，超过该时间没有发出响应报文，则查询设备认为此次查询超时，单位是 1/10 秒。

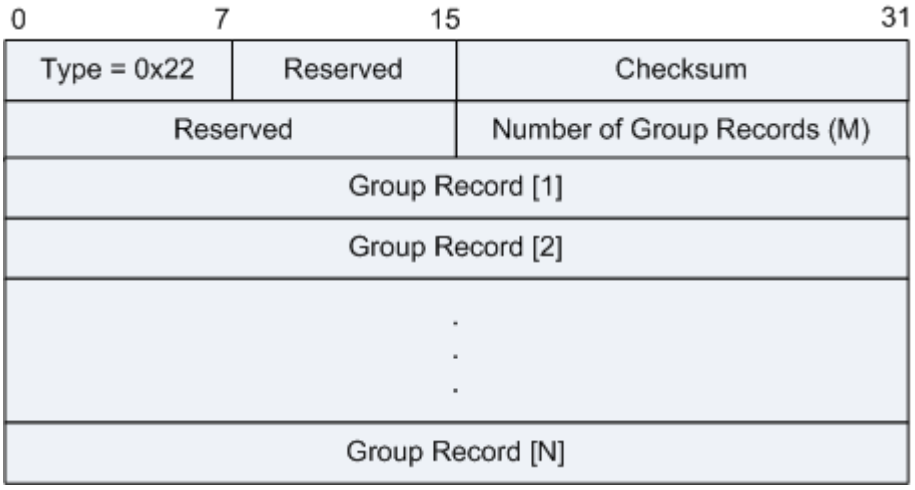
字段	长度	描述
Checksum	16 比特	IGMP 消息的校验和。传送报文时，必须计算校验和并填入该字段中；接收报文时，必须在处理报文之前检验校验和，以判断 IGMP 消息在传输过程中是否发生了错误。
Group Address	32 比特	对于普遍组查询消息，该字段置 0。对于特定组查询消息、特定组/源查询消息，该字段为设置为欲查询的组播组的地址。
Resv	4 比特	保留字段，发送报文时置 0；接收到报文时，对该字段不做任何处理。
S	1 比特	该比特位置 1 时，所有收到此查询消息的其他路由器不启动定时器刷新过程，但是此查询消息并不抑制查询器选举过程和路由器的主机侧处理过程。
QRV	3 比特	查询者的健壮变量，如果不为 0，QRV 中包含中一个被查询者使用的[健壮变量]的值，如果查询者的健壮变量的值超过 7，即 QRV 字段的最大值，那么 QRV 被设成 0。路由器取最近收到的查询中的 QRV 值作为它们自己的健壮性变量的值，除非最近收到的 QRV 是 0，在这种情况下，接收者使用缺省的健壮性变量值，或者是一个静态配置的值。
QQIC	8 比特	查询器的查询间隔，单位是秒。非查询器收到查询报文时，如果发现该字段非 0，则将自己的查询间隔参数调整为该字段的值。
Number of Sources (N)	16 比特	消息中包含的组播源的数量。对于普遍组查询报文和特定组查询报文，该字段为 0；对于特定组/源地址查询报文，该字段非 0。此参数的大小受到所在网络 MTU 大小的限制。
Source Address [i]	32 比特	组播源地址，其数量受到 Number of Sources 字段值大小的限制。

查询消息有三种类型的变体：

- 1、“普通查询”由多播路由器发出，用于获知邻接接口(即查询所传输的网络中所相连的接口)的完整的多播接收状态。在一个普通查询中，组地址字段和源数量(N)字段都为 0。
- 2、“指定组查询”由一台多播路由器发出，用于获知邻接接口中跟某一个 IP 地址相关的多播接收状态。在指定组查询中，“组地址”字段含有需要查询的那个组地址，源数量(N)字段为 0。

- 3、“指定组和源查询”由一台多播路由器发出，用于获知邻接接口是否需要接收来自指定的这些源的，发往指定组的多播数据报。在一个指定组和源的查询中，组地址字段含有要查询的多播地址,源地址[i]字段含有相关的源地址。

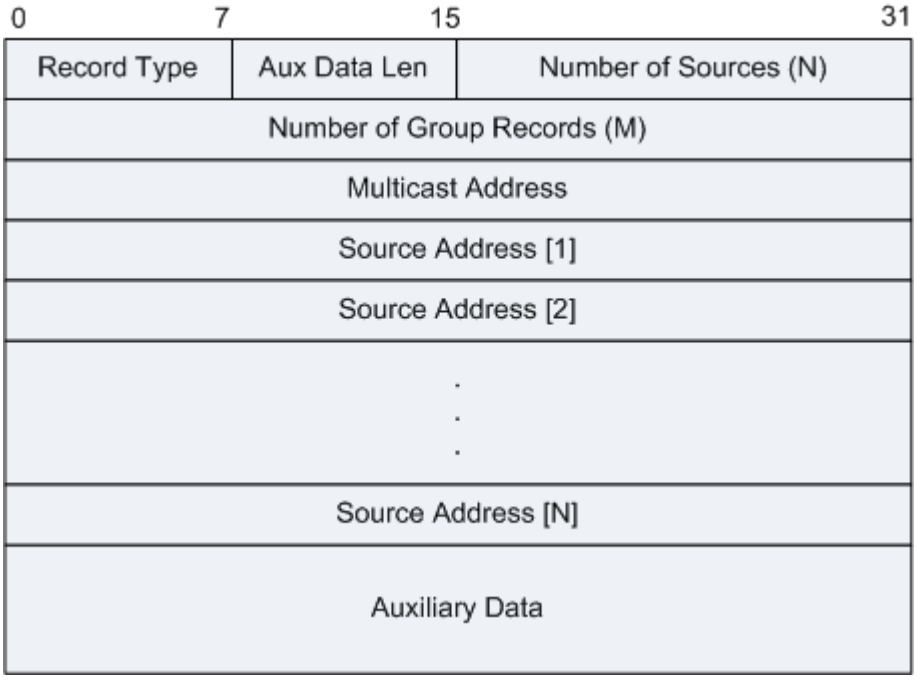
图 2 成员报告消息的格式



成员报告消息是主机向组播路由器发送的报告消息，用报告加入某组播组并只接收由指定组播源发往该组的数据。

封装该消息的 IP 报文头的目的地址字段为 224.0.0.22，本地网段上的所有 IGMPv3 路由器都能识别并接收。

图 3 Group Record 字段格式



字段	长度	描述
Type	8 比特	Type = 0x22 成员关系报告
Reserved	8 比特	保留字段，在发送的时候是以 0 填充，在接收的时候是不作任何处理的。
Checksum	16 比特	校验和是对整个 IGMP 消息以 16 位为一段进行取反求和。为了计算校验和，校验和字段首先必须被置 0。当收到一个数据，在处理之前，必须先对校验和进行验证。

字段	长度	描述
Reserved	16 比特	T 保留字段，在发送的时候是以 0 填充，在接收的时候是不作任何处理的。
Number of Group Records (M)	16 比特	该字段表示该报告报文中包含有几个组记录。
Group Record	变长	一个主机可能需要点播多个组播地址的组播业务，每个记录包含了对应于其中一个组播地址的源地址列表等信息，它受到 Number_of_Group_Records 的大小的影响。 每一个组记录字段是一整块数据，其含有的信息是关于发送者在报告发送接口上的某一个多播组的成员关系。
Record Type	8 比特	Group Record 消息的类型。 • MODE_IS_INCLUDE：接收源地址列表包含的源发往该组的组播数据。如果指定源地址列表为空，该消息为无效消息。 • MODE_IS_EXCLUDE：不接收源地址列表包含的源发往该组的组播数据。 • CHANGE_TO_INCLUDE_MODE：过滤模式由 EXCLUDE 转换到 INCLUDE，接收源地址列表包含的新组播源发往该组播组的数据。如果指定源地址列表为空，主机离开组播组。 • CHANGE_TO_EXCLUDE_MODE：过滤模式由 INCLUDE 转换到 EXCLUDE，拒绝源地址列表中新组播源发往该组的组播数据。 • ALLOW_NEW_SOURCES：表示在现有的基础上，需要接收源地址列表包含的源发往该组播组的组播数据。如果当前对应关系为 INCLUDE，则向现有源列表中添加这些组播源；如果当前对应关系为 EXCLUDE，则从现有阻塞源列表中删除这些组播源。 • BLOCK_OLD_SOURCES：表示在现有的基础上，不再接收从源地址列表包含的源组播源发往该组播组的组播数据。如果当前对应关系为 INCLUDE，则从现有源列表中删除这些组播源；如果当前对应关系为 EXCLUDE，则向现有源列表中添加这些组播源。
Aux Data Len	8 比特	辅助数据长度含有在组记录中的辅助数据的实际长度，其单位是 32bit 字。它有可能是 0，这就表示辅助数据不存在。
Number of Sources (N)	16 比特	源数量(N)字段标明在组记录中存在多少源地址。
Multicast Address	32 比特	多播地址字段标明该组记录从属的多播 IP 地址。
Source Address [i]	32 比特	源地址[i]字段是一个数组，含有 n 个单播地址。n 就是该记录的源数量(N)字段的值。

字段	长度	描述
Additional Data	变长	附加数据。如果收到的报告中的 IP 首部的数据报长度字段标明在最后一个组记录后面有附加的数据存在。IGMPv3 的实现必须在计算和验证校验和的时候包含这些附加数据，但是同时必须忽略这些附加数据。当发送一个报告时，一个 IGMPv3 的实现在最后一个组记录后面不能包含附加数据。

报文示例

图 4 IGMPv3 Membership Report (CHANGE_TO_INCLUDE_MODE)

```
⊞ Frame 1: 64 bytes on wire (512 bits), 64 bytes captured
⊞ Ethernet II (VLAN tagged), Src: Performa_00:00:01 (00:1
⊟ Internet Protocol Version 4, Src: 192.85.1.3 (192.85.1.
    Version: 4
    Header length: 24 bytes
  ⊞ Differentiated Services Field: 0xc0 (DSCP 0x30: Class
    Total Length: 40
    Identification: 0xf001 (61441)
  ⊞ Flags: 0x00
    Fragment offset: 0
    Time to live: 1
    Protocol: IGMP (2)
  ⊞ Header checksum: 0x929f [correct]
    Source: 192.85.1.3 (192.85.1.3)
    Destination: 224.0.0.22 (224.0.0.22)
  ⊞ Options: (4 bytes)
⊟ Internet Group Management Protocol
    [IGMP Version: 3]
    Type: Membership Report (0x22)
    Header checksum: 0xf9fc [correct]
    Num Group Records: 1
  ⊟ Group Record : 225.0.0.1  Change To Include Mode
      Record Type: Change To Include Mode (3)
      Aux Data Len: 0
      Num Src: 0
      Multicast Address: 225.0.0.1 (225.0.0.1)
```

图 5 IGMPv3 Membership Report (CHANGE_TO_EXCLUDE_MODE)

```
⊞ Frame 8: 68 bytes on wire (544 bits), 68 bytes capturec
⊞ Ethernet II (VLAN tagged), Src: Performa_00:00:01 (00:1
⊞ Internet Protocol Version 4, Src: 192.85.1.3 (192.85.1.
⊟ Internet Group Management Protocol
    [IGMP Version: 3]
    Type: Membership Report (0x22)
    Header checksum: 0xf8fc [correct]
    Num Group Records: 1
  ⊟ Group Record : 225.0.0.1  Change To Exclude Mode
      Record Type: Change To Exclude Mode (4)
      Aux Data Len: 0
      Num Src: 0
      Multicast Address: 225.0.0.1 (225.0.0.1)
```

图 6 IGMPv3 Membership Report (BLOCK_OLD_SOURCES)

```
⊞ Frame 1: 68 bytes on wire (544 bits), 68 bytes cap
⊞ Ethernet II (VLAN tagged), Src: Performa_00:00:01
⊞ Internet Protocol Version 4, Src: 192.85.1.3 (192.
⊟ Internet Group Management Protocol
    [IGMP Version: 3]
    Type: Membership Report (0x22)
    Header checksum: 0x2f52 [correct]
    Num Group Records: 1
  ⊟ Group Record : 232.0.0.1  Block Old Sources
      Record Type: Block Old Sources (6)
      Aux Data Len: 0
      Num Src: 1
      Multicast Address: 232.0.0.1 (232.0.0.1)
      Source Address: 192.168.0.1 (192.168.0.1)
```

图 7 IGMPv3 Membership Report (ALLOW_NEW_SOURCES)

```
⊞ Frame 1: 68 bytes on wire (544 bits), 68 bytes captu
⊞ Ethernet II (VLAN tagged), Src: Performa_00:00:01 (0
⊞ Internet Protocol Version 4, Src: 192.85.1.3 (192.8
⊟ Internet Group Management Protocol
    [IGMP Version: 3]
    Type: Membership Report (0x22)
    Header checksum: 0x3052 [correct]
    Num Group Records: 1
  ⊟ Group Record : 232.0.0.1  Allow New Sources
      Record Type: Allow New Sources (5)
      Aux Data Len: 0
      Num Src: 1
      Multicast Address: 232.0.0.1 (232.0.0.1)
      Source Address: 192.168.0.1 (192.168.0.1)
```

图 8 IGMPv3 Membership Query (General)

```
⊞ Frame 1: 68 bytes on wire (544 bits), 68 bytes captured (54
⊞ Ethernet II (VLAN tagged), Src: 00:e1:fc:00:29:d7 (00:e1:fc
⊞ Internet Protocol Version 4, Src: 192.168.0.1 (192.168.0.1)
⊟ Internet Group Management Protocol
    [IGMP Version: 3]
    Type: Membership Query (0x11)
    Max Response Time: 10.0 sec (0x64)
    Header checksum: 0xec5f [correct]
    Multicast Address: 0.0.0.0 (0.0.0.0)
  ⊟ QRV=2 S=Do not suppress router side processing
      .... 0... = S: Do not suppress router side processing
      .... .010 = QRV: 2
    QQIC: 60
    Num Src: 0
```

图 9 IGMPv3 Membership Query (Special)

```
⊞ Frame 1: 68 bytes on wire (544 bits), 68 bytes capture
⊞ Ethernet II (VLAN tagged), Src: 00:e1:fc:00:29:d7 (00:
⊞ Internet Protocol Version 4, Src: 192.168.0.1 (192.168
⊟ Internet Group Management Protocol
    [IGMP Version: 3]
    Type: Membership Query (0x11)
    Max Response Time: 1.0 sec (0x0a)
    Header checksum: 0x3df2 [correct]
    Multicast Address: 232.0.0.1 (232.0.0.1)
  ⊟ QRV=0 S=SUPPRESS router side processing
      .... 1... = S: SUPPRESS router side processing
      .... .000 = QRV: 0
    QQIC: 0
    Num Src: 1
    Source Address: 192.0.1.0 (192.0.1.0)
```

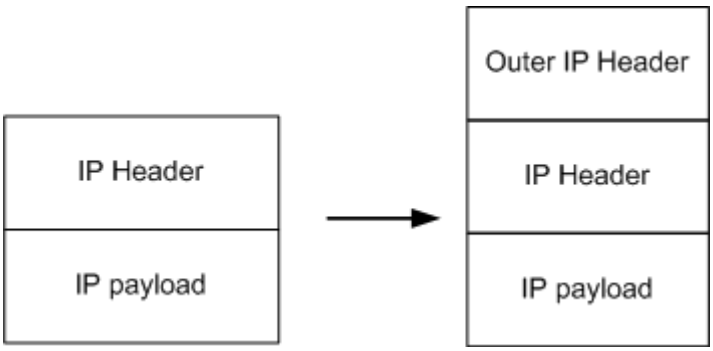
参考标准

标准	描述
RFC 3376	Internet Group Management Protocol, Version 3

4.6 IP in IP 报文格式

报文格式

图 1 IPinIP 报文格式



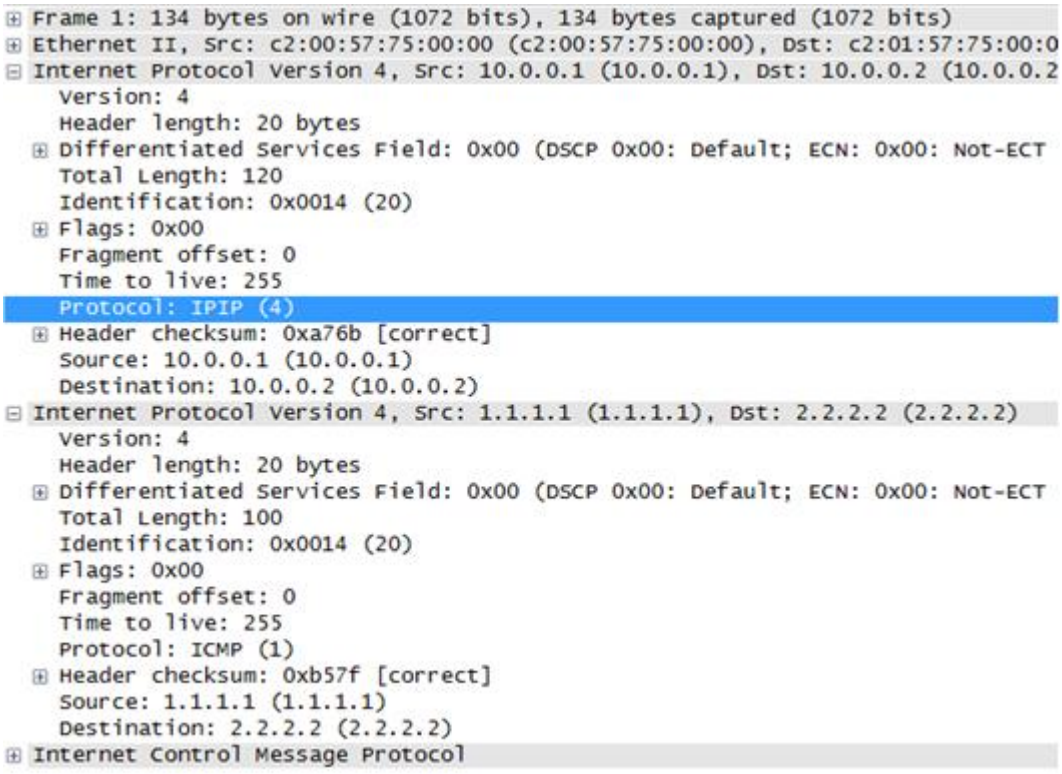
其中，内层 IPv4 头部和普通 IPv4 报文头部相同，IPv4 报文头详细解释请参见 IPv4 报文格式。外层 IPv4 头部处理如下：

字段	含义
Version	=4
IHL	指外层 IP 头部长度，以 32 比特为计算单位。
TOS	从内层 IP 头部复制。
Total Length	指整个 IP 负载的长度，包括外层 IP 头，内层 IP 头和 IP 负载。
Identification, Flags, Fragment Offset	这三个字段的含义与 RFC791 的定义相同。注意，如果内层 IP 头部的 DF 位置位，外层 IP 头部的 DF 位也必须置位。如果内层 IP 头的 DF 未置位，外层 IP 头部的 DF 位可以置位也可以不置位。
Time to Live	外层 IP 头部的 TTL 域设置为发送该数据包到隧道目的端的合适的值。

字段	含义
Protocol	=4
Header Checksum	外层 IP 头部的校验字段。
Source Address	执行该 IPinIP 隧道头封装的隧道入口设备的 IP 地址。
Destination Address	执行该 IPinIP 隧道头解封装的隧道出口设备的 IP 地址。
Options	内层 IP 头部的任何选项字段通常不被复制到外层 IP 头部。隧道路径上的设备可以添加新的选项字字段。 内层 IP 头部的安全选项字段的类型可能影响外层 IP 头部的安全选项字段的选择。

报文示例

图 2 IP in IP 报文



参考标准

标准	描述
RFC 2003	IP Encapsulation within IP

4.7 IP 报文格式

报文格式

图 1 IP 头格式

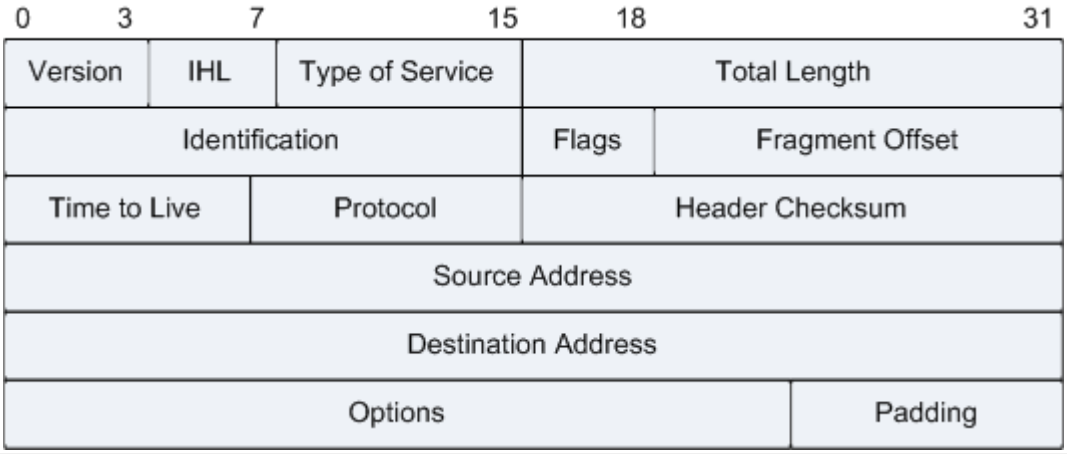


表 1 IP 头字段解释

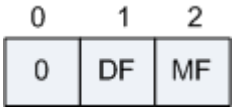
字段	长度	含义
Version	4 比特	• 4：表示为 IPV4； • 6：表示为 IPV6。
IHL	4 比特	首部长度，如果不带 Option 字段，则为 20，最长为 60，该值限制了记录路由选项。以 4 字节为一个单位。
Type of Service	8 比特	服务类型。只有在有 QoS 差分服务要求时这个字段才起作用。
Total Length	16 比特	总长度，整个 IP 数据报的长度，包括首部和数据之和，单位为字节，最长 65535，总长度必须不超过最大传输单元 MTU。
Identification	16 比特	标识，主机每发一个报文，加 1，分片重组时会用到该字段。
Flags	3 比特	标志位： **图 2** IP Flag 字段格式  • Bit 0: 保留位，必须为 0。 • Bit 1: DF（Don't Fragment），能否分片位，0 表示可以分片，1 表示不能分片。 • Bit 2: MF（More Fragment），表示是否该报文为最后一片，0 表示最后一片，1 代表后面还有。
Fragment Offset	12 比特	片偏移：分片重组时会用到该字段。表示较长的分组在分片后，某片在原分组中的相对位置。以 8 个字节为偏移单位。

表 1 IP 头字段解释

字段	长度	含义
Time to Live	8 比特	生存时间：可经过的最多路由数，即数据包在网络中可通过的路由器数的最大值。
Protocol	8 比特	协议：下一层协议。指出此数据包携带的数据使用何种协议，以便目的主机的 IP 层将数据部分上交给哪个进程处理。 常见值： • 0: 保留 Reserved • 1: ICMP, Internet Control Message [RFC792] • 2: IGMP, Internet Group Management [RFC1112] • 3: GGP, Gateway-to-Gateway [RFC823] • 4: IP in IP (encapsulation) [RFC2003] • 6: TCP Transmission Control Protocol [RFC793] • 17: UDP User Datagram Protocol [RFC768] • 20: HMP Host Monitoring Protocol [RFC 869] • 27: RDP Reliable Data Protocol [RFC908] • 46: RSVP (Reservation Protocol) • 47: GRE (General Routing Encapsulation) • 50: ESP Encap Security Payload [RFC2406] • 51: AH (Authentication Header) [RFC2402] • 54: NARP (NBMA Address Resolution Protocol) [RFC1735] • 58: IPv6-ICMP (ICMP for IPv6) [RFC1883] • 59: IPv6-NoNxt (No Next Header for IPv6) [RFC1883] • 60: IPv6-Opts (Destination Options for IPv6) [RFC1883] • 89: OSPF (OSPF Version 2) [RFC 1583] • 112: VRRP (Virtual Router Redundancy Protocol) [RFC3768] • 115: L2TP (Layer Two Tunneling Protocol) • 124: ISIS over IPv4 • 126: CRTP (Combat Radio Transport Protocol) • 127: CRUDP (Combat Radio User Protocol) • 132: SCTP (Stream Control Transmission Protocol) • 136: UDPLite [RFC 3828]

表 1 IP 头字段解释

字段	长度	含义
		• 137: MPLS-in-IP [RFC 4023]
Header Checksum	16 比特	首部检验和，只检验数据包的首部，不检验数据部分。这里不采用 CRC 检验码，而采用简单的计算方法。
Source Address	32 比特	源 IP 地址。
Destination Address	32 比特	目的 IP 地址。
Options	可变	选项字段，用来支持排错，测量以及安全等措施，内容丰富（请参见下表）。选项字段长度可变，从 1 字节到 40 字节不等，取决于所选项的功能。
Padding	可变	填充字段，全填 0。

表 2 IP Header Options

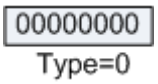
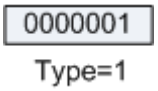
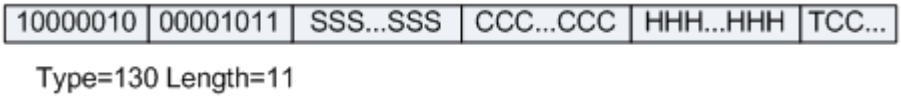
CLASS	NUMBER	长度	含义
0	0	-	Code 为 0 代表了选项列表的结束，放在所有选项链表的后面，用来补字节对齐。 该选项无长度字段，占一个字节。 选项格式如下： 
0	1	-	表示无操作的选项。用在各种选项之间，占一个字节。用于填充 4 字节对齐。 选项格式如下： 
0	2	11 字节	表示安全和处理限制的选项。 该选项提供一种主机可以发送安全、分隔、处理限制及 TCC（关闭使用组）的参数功能。 选项格式如下：  • **Type=**占 1 字节，code 的值此处设为 130

表 1 IP 头字段解释

字段		长度	含义
			• **length=**占 1 字节，长度选项固定为 11，表示该选项的长度为 11 字节 • **SSS：**占 2 字节，表示安全域，下面列出了 16 种不同的安全标准，其中 8 个至今还没使用，预留将来使用，列表如下： 00000000 00000000 - Unclassified 11110001 00110101 - Confidential 01111000 10011010 - EFTO 10111100 01001101 - MMMM 01011110 00100110 - PROG 10101111 00010011 - Restricted 11010111 10001000 - Secret 01101011 11000101 - Top Secret 00110101 11100010 - (Reserved for future use) 10011010 11110001 - (Reserved for future use) 01001101 01111000 - (Reserved for future use) 00100100 10111101 - (Reserved for future use) 00010011 01011110 - (Reserved for future use) 10001001 10101111 - (Reserved for future use) 11000100 11010110 - (Reserved for future use) 11100010 01101011 - (Reserved for future use) • **CCC：**占 2 个字节，表示分隔域，当传输的数据没被分隔的时候，此值设为 0，其他的具体值可以从国防情报局获取。 • **HHH：**占 2 个字节，操作限制域，该值由国防情报局 DIAM 手册 65－19，《标准安全记号》描述。 • **TCC：**占 3 个字节，传输控制码。提供一种传输隔离的手段，该值为 3 字母词，可用值从 HQ DCA Code 530。
0	3	可变	松散的源站选路（为数据报指定一系列必须经过的 IP 地址） 选项格式如下： 

表 1 IP 头字段解释

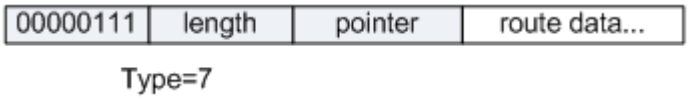
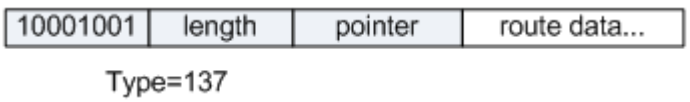
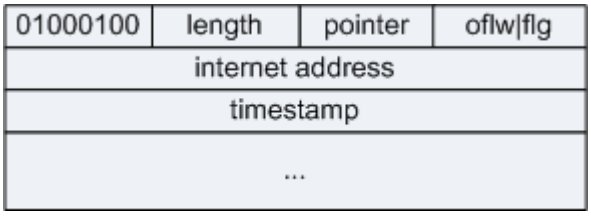
字段		长度	含义
			• Type：类型。占 1 字节，此处设为 131。 • length：占 1 字节，记录整个选项的长度。 • pointer：指针项，占 1 个字节，指向下一个被处理的源站地址，最小值为 4。 • route data：路由数据。
0	7	可变	记录路径（让每个路由器都记下它的 IP 地址）。 选项格式如下：  Type=7 • Type：指明 IP 选项的类型。此处值为 7。 • length：选项的总字节长度。不包含填充的长度，IP 填满时最大为 39;。 • pointer：它是一个基于 1 的指针，指向存放下一个 IP 地址的位置。它的最小值为 4，指向存放第一个 IP 地址的位置。随着每个 IP 地址存入清单，ptr 的值分别为 8，12，16，最大到 36，当记录下 9 个 IP 地址后，ptr 的值为 40，表示清单已满。 • route data：路由数据。
0	8	4 字节	流标识选项。该选项长度固定为 4 字节，code 值为 136，后面的字段固定为 0x02，流 ID 为 2 字节。该选项提供了一种携带 SATNET 流标识符通过不支持流方式的网络。 选项格式如下：  Type=136 Length=4
0	9	可变	严格的源站选路选项。与宽松的源站选路类似，但是要求只能经过指定的这些地址，不能经过其他的地址。 选项格式如下：  Type=137 • **Type=**占 1 字节，code 的值此处设为 137。 • **length：**占 1 字节，记录整个选项的长度。 • **pointer：**指针项，占 1 个字节，指向下一个被处理的源站地址，最小值为 4。
2	4	可变	时间戳选项。

表 1 IP 头字段解释

字段	长度	含义
		选项格式如下：  • **Type (01000100)**：时间戳选项，代码为 68； • **length**：选项的总长度（一般为 36 或 40）； • **ponter**：指向下一个可用空间的指针（5，9，13 等）； • **oflw**：表示溢出字段； • **flg**：表示标志字段： ▪ 0：只记录时间戳。 ▪ 1：每台路由器都记录它的 IP 地址和时间戳。在选项列表中只有存放 4 对地址和时间戳的空间。 ▪ 3：发送端对选项列表进行初始化，存放了 4 个 IP 地址和 4 个取值为 0 的时间戳值。只有当列表中的下一个 IP 地址与当前路由器地址相匹配时，才记录它的时间戳。

报文示例

```
⊞ Frame 91: 60 bytes on wire (480 bits), 60 bytes captured (480 bits)
⊞ Ethernet II, Src: HuaweiTe_2f:6a:25 (00:e0:fc:2f:6a:25), Dst: Dell_b5:ca:35 (00:11:43:
⊟ Internet Protocol Version 4, Src: 83.111.62.34 (83.111.62.34), Dst: 172.21.2.4 (172.21
    Version: 4
    Header length: 20 bytes
  ⊟ Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00: Not-ECT (Not ECN
      0000 00.. = Differentiated Services Codepoint: Default (0x00)
      .... ..00 = Explicit Congestion Notification: Not-ECT (Not ECN-Capable Transport)
    Total Length: 40
    Identification: 0x0911 (2321)
  ⊟ Flags: 0x02 (Don't Fragment)
      0... .... = Reserved bit: Not set
      .1.. .... = Don't fragment: Set
      ..0. .... = More fragments: Not set
    Fragment offset: 0
    Time to live: 127
    Protocol: TCP (6)
  ⊟ Header checksum: 0xb314 [correct]
      [Good: True]
      [Bad: False]
    Source: 83.111.62.34 (83.111.62.34)
    Destination: 172.21.2.4 (172.21.2.4)
⊞ Transmission Control Protocol, Src Port: ratio-adp (1108), Dst Port: ftp (21), Seq: 1,
```

参考标准

标准 Standard	Description
RFC 791	INTERNET PROTOCOL

4.8 IPv6 报文格式

报文格式

图 1 IPv6 报文头格式

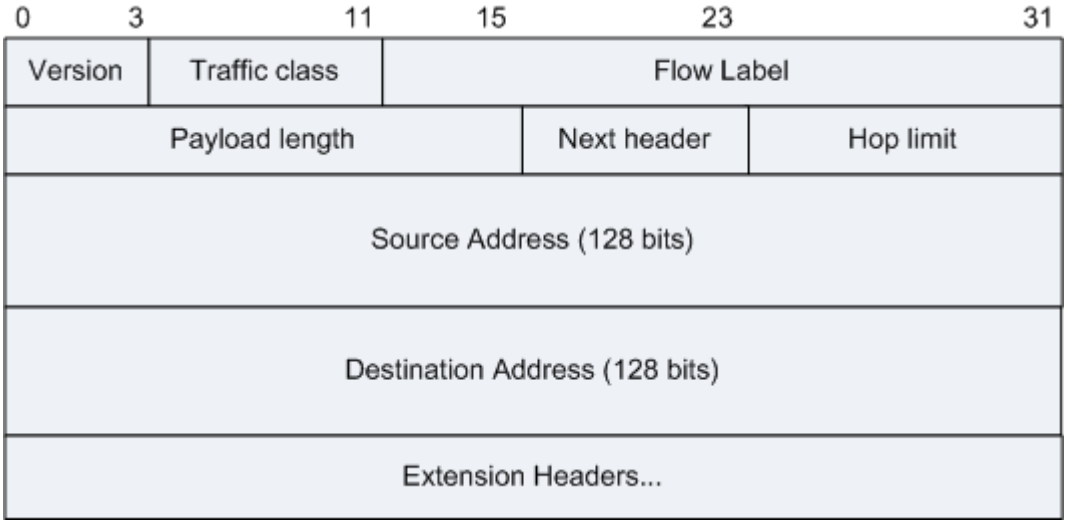


表 1 IP 头字段解释

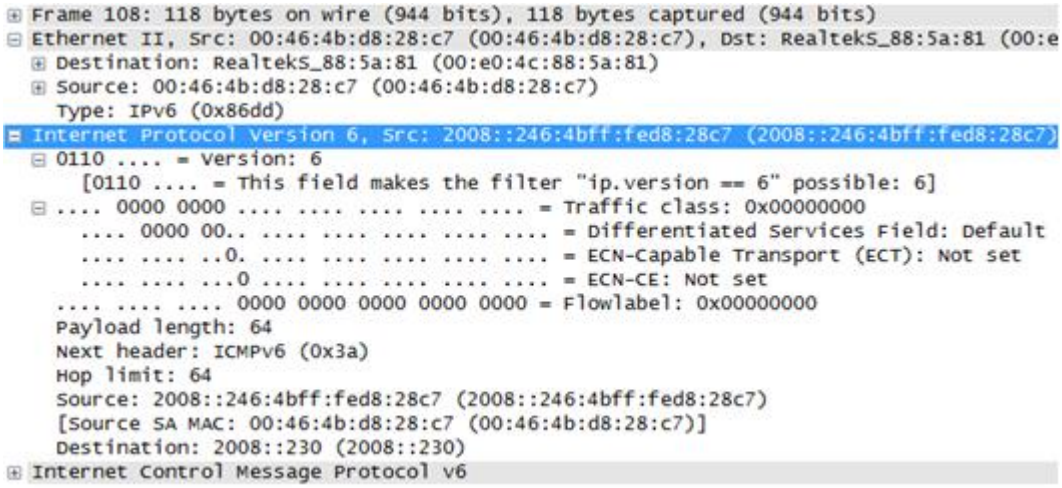
字段	长度	含义
Version	4 比特	• 4：表示为 IPV4； • 6：表示为 IPV6。
Traffic class	8 比特	流量类别。该字段及其功能类似于 IPv4 的业务类型字段。该字段以区分业务编码点（DSCP）标记一个 IPv6 数据包，以此指明数据包应当如何处理。
Flow Label	20 比特	流标签。该字段用来标记 IP 数据包的一个流，当前的标准中没有定义如何管理和处理流标签的细节。
Payload length	16 比特	该字段表示有效载荷的长度，有效载荷是指紧跟 IPv6 基本报头的数据包，包含 IPv6 扩展报头。
Next header	8 比特	下一报头，该字段指明了跟随在 IPv6 基本报头后的扩展报头的信息类型。
Hop limit	8 比特	跳数限制，该字段定义了 IPv6 数据包所能经过的最大跳数，这个字段和 IPv4 中的 TTL 字段非常相似。
Source Address	128 比特	该字段表示该报文的源地址。
Destination Address	128 比特	该字段表示该报文的目的地址。

表 1 IP 头字段解释

字段	长度	含义
Extension Headers	可变	扩展报头。IPv6 取消了 IPv4 报头中的选项字段，并引入了多种扩展报文头，在提高处理效率的同时还增强了 IPv6 的灵活性，为 IP 协议提供了良好的扩展能力。当超过一种扩展报头被用在同一个分组里时，报头必须按照下列顺序出现： • IPv6 基本报头 • 逐跳选项扩展报头 • 目的选项扩展报头 • 路由扩展报头 • 分片扩展报头 • 授权扩展报头 • 封装安全有效载荷扩展报头 • 目的选项扩展报头（指那些将被分组报文的最终目的地处理的选项。） • 上层扩展报头 不是所有的扩展报头都需要被转发路由设备查看和处理的。路由设备转发时根据基本报头中 Next Header 值来决定是否要处理扩展头。 除了目的选项扩展报头出现两次（一次在路由扩展报头之前，另一次在上层扩展报头之前），其余扩展报头只出现一次。

报文示例

图 2 IPv6 报文



参考标准

标准 Standard	Description

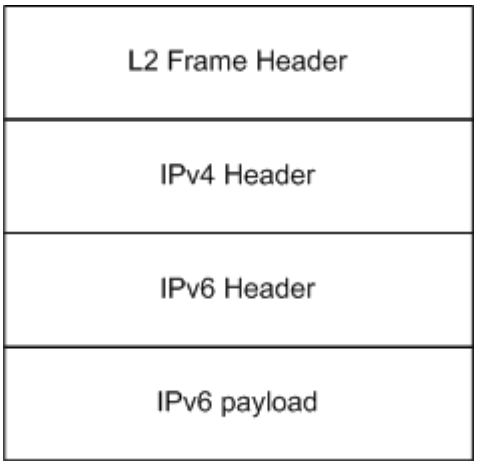
标准 Standard	Description
RFC 2460	Internet Protocol, Version 6 (IPv6) Specification (Updated by RFC 7045, RFC 5871, RFC 5722, RFC 6935, RFC 7112, RFC 6437, RFC 6946, RFC 6564, RFC 5095)

4.9 IPv6 in IP (6to4)报文格式

报文格式

RFC3056 定义了 6to4 隧道。6to4 隧道是 IETF 较为重视并得到深入研究的有广阔应用前景的一种网络过渡机制，可以使连接到纯 IPv4 网络上的孤立 IPv6 子网或 IPv6 站点与其他同类站点在尚未获得纯 IPv6 连接时彼此间进行通信。在 IPv4 网络内可以采用多种路由协议（OSPF、BGP、RIP、IS-IS 等），在两个 6to4 域之间可以通过 MP-BGP 路由方式实现路由可达。

图 1 IPv6 in IP (6to4)报文格式



IPv6 in IP (6to4)报文格式中，IPv4 头部和 IPv6 头部格式与普通 IPv4 和 IPv6 报文头部相同，详细请参见 IPv4 报文格式和 IPv6 报文格式。

报文示例

```
⊞ Frame 1: 134 bytes on wire (1072 bits), 134 bytes captured (1072 bits)
⊟ Ethernet II, Src: c2:00:42:02:00:00 (c2:00:42:02:00:00), Dst: c2:01:42:02:00:00 (c2:01:42:02:00:00)
  ⊞ Destination: c2:01:42:02:00:00 (c2:01:42:02:00:00)
  ⊞ Source: c2:00:42:02:00:00 (c2:00:42:02:00:00)
    Type: IP (0x0800)
⊟ Internet Protocol Version 4, Src: 10.0.0.1 (10.0.0.1), Dst: 10.0.0.2 (10.0.0.2)
    Version: 4
    Header length: 20 bytes
  ⊞ Differentiated Services Field: 0x00 (DSCP 0x00: Default; ECN: 0x00: Not-ECT (Not ECN-Capable Transport))
    Total Length: 120
    Identification: 0x0009 (9)
  ⊞ Flags: 0x00
    Fragment offset: 0
    Time to live: 255
    Protocol: IPv6 (41)
  ⊞ Header checksum: 0xa751 [correct]
    Source: 10.0.0.1 (10.0.0.1)
    Destination: 10.0.0.2 (10.0.0.2)
⊟ Internet Protocol Version 6, Src: 2001:db8:0:1::1 (2001:db8:0:1::1), Dst: 2001:db8:0:1::2 (2001:db8:0:1::2)
  ⊟ 0110 .... = Version: 6
      [0110 .... = This field makes the filter "ip.version == 6" possible: 6]
  ⊟ .... 0000 0000 .... .... .... .... .... = Traffic class: 0x00000000
      .... 0000 00.. .... .... .... .... .... = Differentiated Services Field: Default (0x00000000)
      .... .... ..0. .... .... .... .... .... = ECN-Capable Transport (ECT): Not set
      .... .... ...0 .... .... .... .... .... = ECN-CE: Not set
    .... .... .... 0000 0000 0000 0000 0000 = Flowlabel: 0x00000000
    Payload length: 60
    Next header: ICMPv6 (0x3a)
    Hop limit: 64
    Source: 2001:db8:0:1::1 (2001:db8:0:1::1)
    Destination: 2001:db8:0:1::2 (2001:db8:0:1::2)
⊞ Internet Control Message Protocol v6
```

参考标准

标准	描述
RFC 4213	Basic Transition Mechanisms for IPv6 Hosts and Routers
RFC 2529	Transmission of IPv6 over IPv4 Domains without Explicit Tunnels

4.10 MLD 报文格式

MLD（Multicast Listener Discovery）组播监听者发现协议。MLD 和 IPv4 网络中的 IGMP 功能类似，用于 IPv6 路由器发现其直连网段上组播监听者（Multicast Listener）、建立、维护组成员关系。

现有 MLDv1 和 MLDv2 两个版本，MLDv2 在 MLDv1 的基础上增加了对 SSM 模型的支持。

报文格式

MLD 消息是 ICMPv6 消息的一个子集，封装在 IPv6 报文中。

图 1 MLD 报文封装格式

- IPv6 报文头的源地址字段为 MLD 消息发送者的 IPv6 本地链路地址。
- IPv6 报文头的目的地址字段用来标识 MLD 消息所属的组播组。
- IPv6 报文头的 Hop Limit 字段值为 1，表示 MLD 消息只在本地网段传播。
- 逐跳选项头（Hop-by-Hop Option Header）。其中下一报头（Next Header）字段值为 58，表示该报文是 ICMPv6 消息；路由器告警选项（RTR-ALERT）的值是 0x05020000，表明该报文是 MLD 消息。
- 不同版本的 MLD 协议，使用不同的消息格式，支持不同的消息类型。

图 2 MLDv1 消息格式

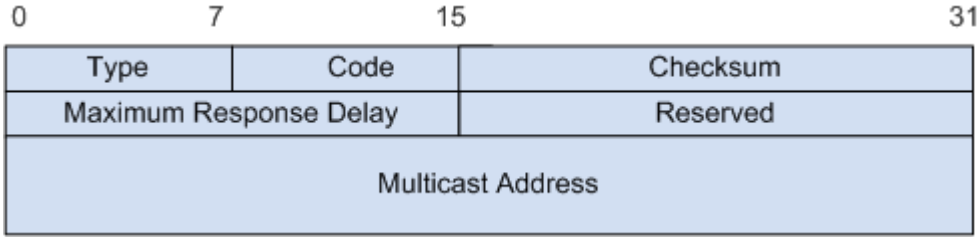


表 1 MLDv1 消息格式字段解释

字段	长度	描述

表 1 MLDv1 消息格式字段解释

字段	长度	描述
Type	1 字节	MLDv1 消息类型，有 3 种取值，如表 2 所示。
Code	1 字节	代码。发送时此字段设置为 0，接收时不处理此字段。
Checksum	2 字节	标准的 ICMPv6 校验和，覆盖所有 MLD 消息以及 IPv6 首部区域中的伪首部。在计算校验和时，此字段以零计算。发送报文时必须计算校验和并将结果写入此字段。接收报文时首先验证校验和，然后才处理报文。
Maximum Response Delay	2 字节	最大响应时间。这个字段只有在查询消息中才有意义。在其他类型的消息中，发送时这个字段被清零，接收时不处理这个字段。
Reserved	2 字节	保留位。发送时此字段被清零，接收时不处理此字段。
Multicast Address	16 字节	组地址。

表 2 MLDv1 消息类型

取值	消息类型	备注
130	MLDv1 普遍组查询消息	普遍组查询消息是查询器定期向共享网段内所有主机以组播方式发送的查询消息，用于查询哪些组播组存在成员。封装该消息的 IPv6 报文头的目的地址字段为 FF02::1。组地址字段为全 0，表示不指定组播组。
	MLDv1 特定组查询消息	特定组查询消息是查询器向共享网段内特定组播组成员发送的消息，用于查询该组播组是否存在成员。封装该消息的 IPv6 报文头的目的地址字段为被查询的组播组的 IP 地址，网络中属于该组播组的成员才能识别并响应。组地址字段为被查询的 IPv6 组播组地址。
131	MLDv1 的成员报告消息	成员报告消息是主机向组播路由器发送的报告消息，用于申请加入某个组播组或者应答查询消息。封装该消息的 IPv6 报文头的目的地址字段为主机要加入的 IPv6 组播组地址，只有网络中的组播路由器和该组成员才能识别并接收。组地址字段为主机要加入的 IPv6 组播组地址。
132	MLDv1 离开消息	离开消息是主机主动离开组播组时向组播路由器发送的消息，用于宣告自己离开了某个组播组。封装该消息的 IPv6 报文头的目的地址字段为 FF02::2。组地址字段为主机要离开的 IPv6 组播组地址。

MLDv2 有以下几种消息：

- MLDv2 查询消息（Type=
- MLDv1 成员报告消息（Type=
- MLDv1 离开消息（Type=
- MLDv2 成员报告消息（Type=

图 3 MLDv2 查询消息

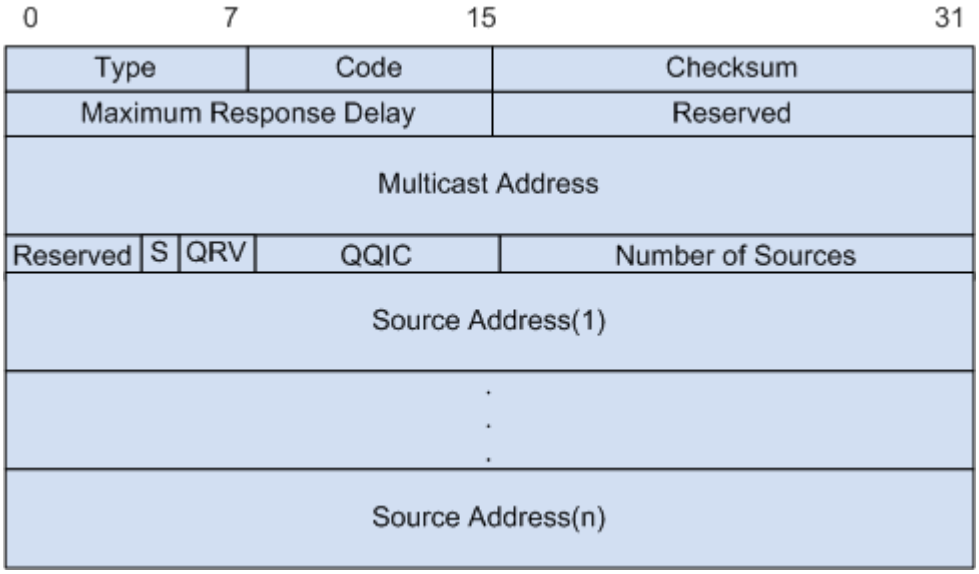


表 3 MLDv2 查询消息字段解释

字段	长度	描述
Type	1 字节	消息类型，该字段取值为 130。
Code	1 字节	发送时此字段设置为 0，接收时不处理此字段。
Checksum	2 字节	标准的 ICMPv6 校验和，覆盖所有 MLD 消息以及 IPv6 首部区域中的伪首部。在计算校验和时，此字段以零计算。发送报文时必须计算校验和并将结果写入此字段。接收报文时首先验证校验和，然后才处理报文。
Maximum Response Delay	2 字节	主机发送报告消息前允许的最长响应延迟。
Reserved	2 字节	保留字段。发送时此字段设置为 0，接收时不处理此字段。
Multicast Address	16 字节	组地址。 • 普遍组查询消息中，此字段设置为 0。

表 3 MLDv2 查询消息字段解释

字段	长度	描述
		• 特定组查询消息中，此字段设置为待查询的组播地址。 • 特定源-组查询消息中，此字段设置为待查询的组播地址。
Reserved	4 比特	保留字段。发送时此字段设置为 0，接收时不处理此字段。
S(Suppress Router-Side Processing)	1 比特	标识位，表示路由器接收到查询消息后是否对定时器更新进行抑制。
QRV(Querier ’ s Robustness Variable)	3 比特	查询器健壮系数。
QQIC(Querier ’ s Query Interval Code)	1 字节	查询器查询间隔。
Number of Sources	2 字节	组播源个数。 • 普遍组查询消息中，此字段设置为 0。 • 特定组查询消息中，此字段设置为 0。 • 特定源-组查询消息中，此字段表示查询消息中包含的源地址个数。
Source Address(i)	16 字节	组播源地址列表。 • 普遍组查询消息中，此字段设置为 0。 • 特定组查询消息中，此字段设置为 0。 • 特定源-组查询消息中，此字段表示指定查询的组播源地址（i＝1,2,…,n, n 表示源地址个数）。

图 4 MLDv2 成员报告消息

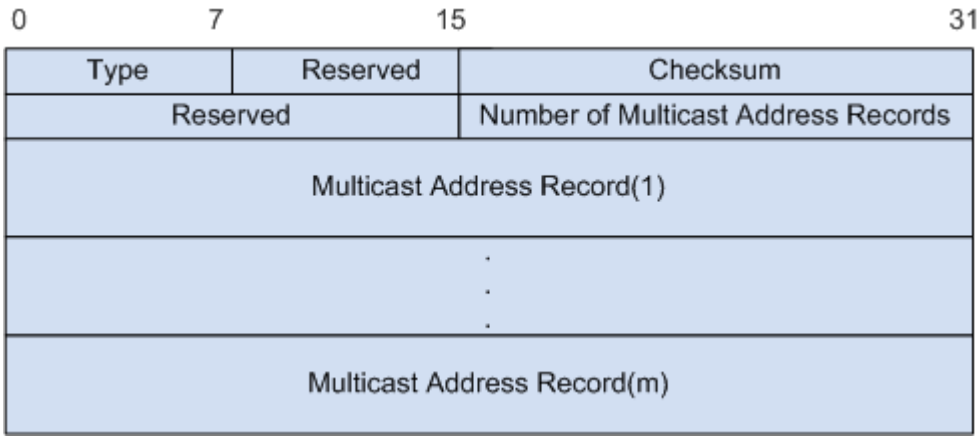
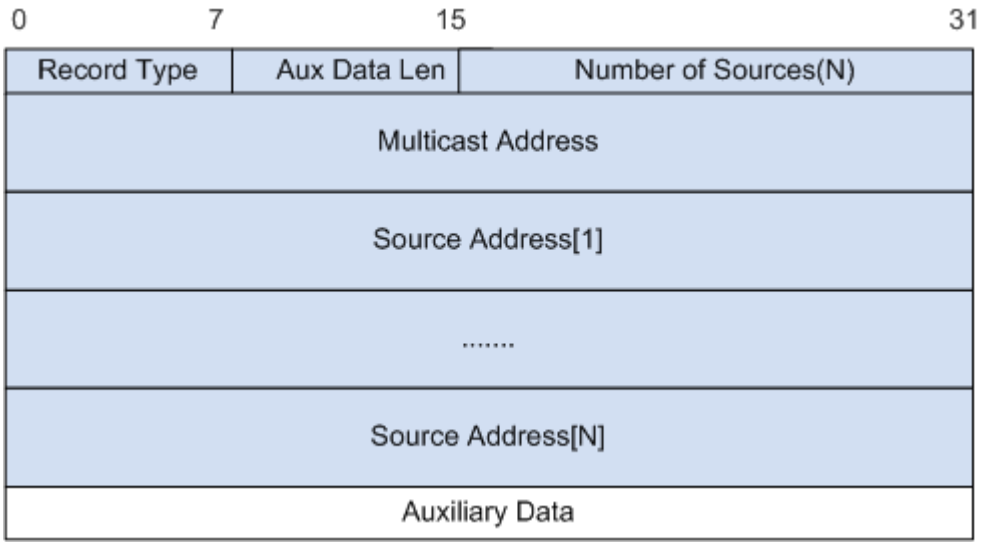


表 4 MLDv2 成员报告消息字段解释

字段	长度	描述
Type	1 字节	消息类型，该字段取值为 143。
Reserved	1 字节	发送时此字段设置为 0，接收时不处理此字段。
Checksum	2 字节	标准的 ICMPv6 校验和，覆盖所有 MLD 消息以及 IPv6 首部区域中的伪首部。在计算校验和时，此字段以零计算。发送报文时必须计算校验和并将结果写入此字段。接收报文时首先验证校验和，然后才处理报文。
Reserved	2 字节	保留字段。发送时此字段设置为 0，接收时不处理此字段。
Number of Multicast Address Record	2 字节	组播地址记录的个数。
Multicast Address Record (i)	变长	组播地址记录，表示主机在接口上侦听到的每个组播地址信息，包括记录类型、组播地址、源地址等。（i＝1,2,…,m, m 表示组播地址记录的个数）

图 5 Multicast Address Record 格式



字段	长度	说明
Record Type	1 字节	记录类型： • MODE_IS_INCLUDE，表示接口和指定组之间的关系是 INCLUDE，源地址列表中会包含的源。该类型的 Record 不会包含空的源列表。 • MODE_IS_EXCLUDE，表示组播组与源列表之间的对应方式为 EXCLUDE，即接收从指定源列表以外的组播源发往该组播组的数据。 • CHANGE_TO_INCLUDE_MODE，表示主机的组播组与源列表之间的对应方式由 EXCLUDE 转换到 INCLUDE。 • CHANGE_TO_EXCLUDE_MODE，表示主机的组播组与源列表之间的对应方式由 INCLUDE 转换到 EXCLUDE。 • ALLOW_NEW_SOURCES，表示在现有的基础上，还希望从某些组播源接收组播数据。如果当前对应关系为 INCLUDE，则向现有源列表中添加某些组播源；如果当前对应关系为 EXCLUDE，则从现有源列表中删除某些组播源。 • BLOCK_OLD_SOURCES，表示在现有的基础上，不再希望从某些组播源接收组播数据。如果当前对应关系为 INCLUDE，则从现有源列表中删除某些组播源；如果当前对应关系为 EXCLUDE，则向现有源列表中添加某些组播源。
Aux Data Len	1 字节	在组播地址 Record 中附加 Auxiliary Data 的长度。
Number of Sources	2 字节	本记录中包含的源地址数量。
Multicast Address	16 字节	组地址。

字段	长度	说明
Source Address	16 字节	组播源地址。
Auxiliary Data	变长	组播地址 Record 的附加信息。

报文示例

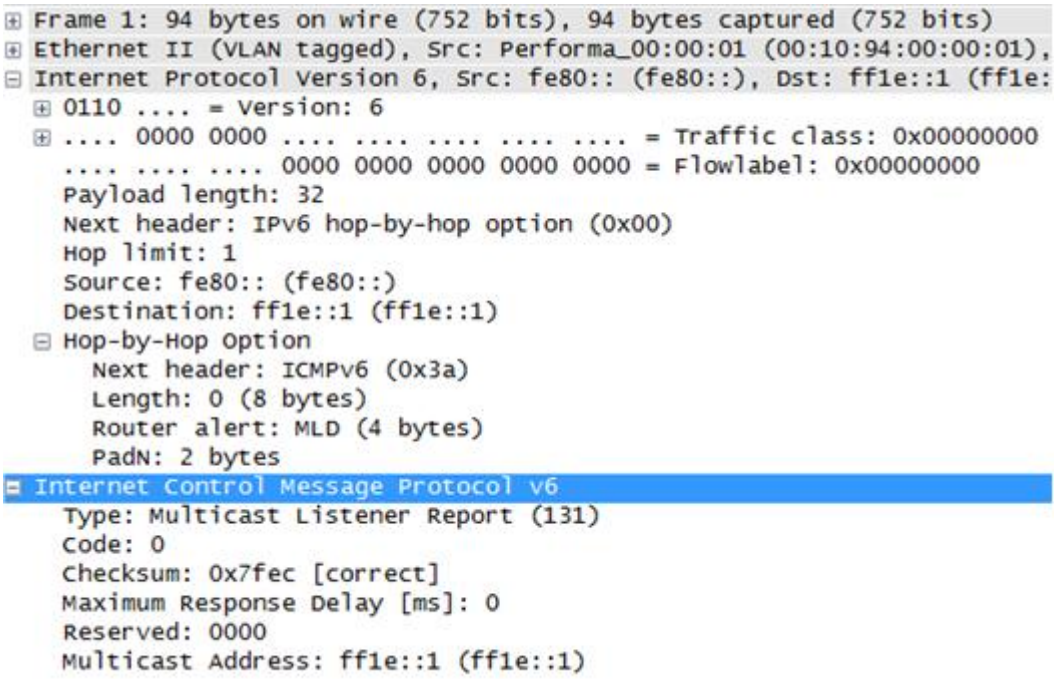
图 6 MLD Membership Query (General) message

```
⊞ Frame 1: 94 bytes on wire (752 bits), 94 bytes captured (752 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_00:1d:af (00:e0:fc:00:1d:af),
⊟ Internet Protocol Version 6, Src: fe80:: (fe80::), Dst: ff02::1 (ff02:
  ⊞ 0110 .... = Version: 6
  ⊞ .... 0000 0000 .... .... .... .... .... = Traffic class: 0x00000000
    .... .... .... 0000 0000 0000 0000 0000 = Flowlabel: 0x00000000
    Payload length: 32
    Next header: IPv6 hop-by-hop option (0x00)
    Hop limit: 1
    Source: fe80:: (fe80::)
    Destination: ff02::1 (ff02::1)
  ⊟ Hop-by-Hop Option
      Next header: ICMPv6 (0x3a)
      Length: 0 (8 bytes)
      Router alert: MLD (4 bytes)
      PadN: 2 bytes
⊟ Internet Control Message Protocol v6
    Type: Multicast Listener Query (130)
    Code: 0
    Checksum: 0x5918 [correct]
    Maximum Response Delay [ms]: 10000
    Reserved: 0000
    Multicast Address: :: (::)
```

图 7 MLD Membership Query (Special) message

```
⊞ Frame 1: 94 bytes on wire (752 bits), 94 bytes captured (752 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_f9:62:64 (00:18:82:f9:62:64),
⊟ Internet Protocol Version 6, Src: fe80:: (fe80::), Dst: ff1e::1 (ff1e:
  ⊞ 0110 .... = Version: 6
  ⊞ .... 0000 0000 .... .... .... .... .... = Traffic class: 0x00000000
    .... .... .... 0000 0000 0000 0000 0000 = Flowlabel: 0x00000000
    Payload length: 32
    Next header: IPv6 hop-by-hop option (0x00)
    Hop limit: 1
    Source: fe80:: (fe80::)
    Destination: ff1e::1 (ff1e::1)
  ⊟ Hop-by-Hop Option
      Next header: ICMPv6 (0x3a)
      Length: 0 (8 bytes)
      Router alert: MLD (4 bytes)
      PadN: 2 bytes
⊟ Internet Control Message Protocol v6
    Type: Multicast Listener Query (130)
    Code: 10
    Checksum: 0x59d2 [correct]
    Maximum Response Delay [ms]: 10000
    Reserved: 0000
    Multicast Address: ff1e::1 (ff1e::1)
```

图 8 MLD Membership Report message



参考标准

标准	描述
RFC 2710	Multicast Listener Discovery (MLD) for IPv6
RFC 3810	Multicast Listener Discovery Version2(MLDv2) for IPv6

4.11 OSPF 报文格式

- OSPF 报文头格式
- OSPF Hello 报文格式
- OSPF DD 报文格式
- OSPF LSR 报文格式
- OSPF LSU 报文格式
- OSPF LSAck 报文格式

父主题： 网络层

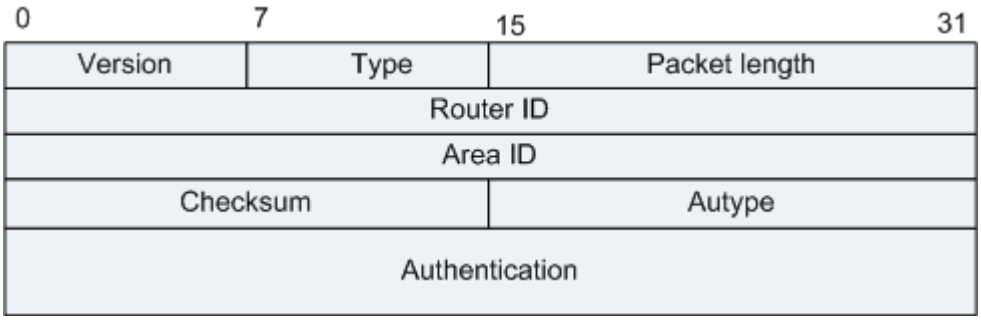
4.11.1 OSPF 报文头格式

报文格式

OSPF 用 IP 报文直接封装协议报文，协议号为 89。OSPF 分为 5 种报文：Hello 报文、DD 报文、LSR 报文、LSU 报文和 LSAck 报文。

OSPF 这五种报文具有相同的报文头格式，长度为 24 字节。

图 1 OSPF 报文头格式



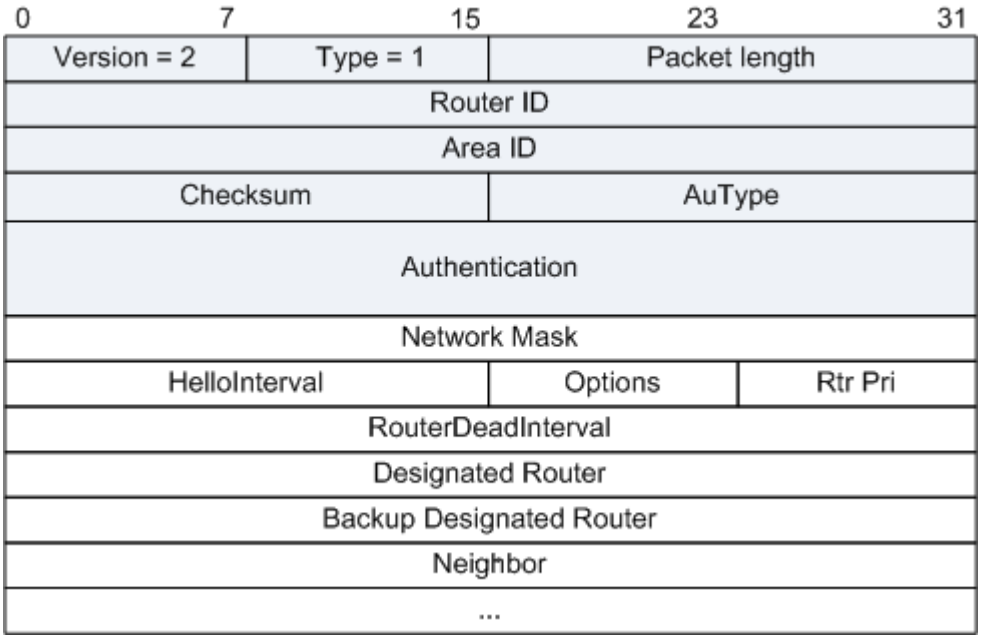
字段	长度	含义
Version	1 字节	版本，OSPF 的版本号。对于 OSPFv2 来说，其值为 2。
Type	1 字节	类型，OSPF 报文的类型，有下面几种类型： • 1：Hello 报文； • 2：DD 报文； • 3：LSR 报文； • 4：LSU 报文； • 5：LSAck 报文。
Packet length	2 字节	OSPF 报文的总长度，包括报文头在内，单位为字节。
Router ID	4 字节	发送该报文的路由器标识。
Area ID	4 字节	发送该报文的所属区域。
Checksum	2 字节	校验和，包含除了认证字段的整个报文的校验和。
AuType	2 字节	验证类型，值有如下几种表示， 0：不验证；1：简单认证；2：MD5 认证。
Authentication	8 字节	鉴定字段，其数值根据验证类型而定。当验证类型为 0 时未作定义；类型为 1 时此字段为密码信息；类型为 2 时此字段包括 Key ID、MD5 验证数据长度和序列号的信息。 MD5 验证数据添加在 OSPF 报文后面，不包含在 Authenticaiton 字段中。

4.11.2 OSPF Hello 报文格式

报文格式

Hello 报文是最常用的一种报文，其作用为建立和维护邻接关系，周期性的在使能了 OSPF 的接口上发送。报文内容包括一些定时器的数值、DR、BDR 以及自己已知的邻居。

图 1 OSPF Hello 报文格式



字段	长度	含义
Network Mask	32 比特	发送 Hello 报文的接口所在网络的掩码。
HelloInterval	16 比特	发送 Hello 报文的时间间隔。
Options	8 比特	可选项： • E：允许 Flood AS-External-LSAs • MC：转发 IP 组播报文 • N/P：处理 Type-7 LSAs • DC：处理按需链路
Rtr Pri	8 比特	DR 优先级。默认为 1。如果设置为 0，则路由器不能参与 DR 或 BDR 的选举。
RouterDeadInterval	32 比特	失效时间。如果在此时间内未收到邻居发来的 Hello 报文，则认为邻居失效。
Designated Router	32 比特	DR 的接口地址。
Backup Designated Router	32 比特	BDR 的接口地址。
Neighbor	32 比特	邻居，以 Router ID 标识。

报文示例

图 2 OSPF Hello

```
⊞ Frame 1: 82 bytes on wire (656 bits), 82 bytes captured (656 bits)
⊞ Ethernet II, Src: 1b:00:00:40:64:00 (1b:00:00:40:64:00), Dst: Air2u_00:ff:
⊞ Internet Protocol Version 4, Src: 161.81.0.1 (161.81.0.1), Dst: 224.0.0.5
⊟ Open Shortest Path First
  ⊟ OSPF Header
      OSPF Version: 2
      Message Type: Hello Packet (1)
      Packet Length: 48
      Source OSPF Router: 220.0.10.2 (220.0.10.2)
      Area ID: 0.0.0.0 (Backbone)
      Packet Checksum: 0xcec3 [correct]
      Auth Type: Null
      Auth Data (none)
  ⊟ OSPF Hello Packet
      Network Mask: 255.255.255.0
      Hello Interval: 1 seconds
    ⊟ Options: 0x02 (E)
        0... .... = DN: DN-bit is NOT set
        .0.. .... = O: O-bit is NOT set
        ..0. .... = DC: Demand Circuits are NOT supported
        ...0 .... = L: The packet does NOT contain LLS data block
        .... 0... = NP: NSSA is NOT supported
        .... .0.. = MC: NOT Multicast Capable
        .... ..1. = E: External Routing Capability
        .... ...0 = MT: NO Multi-Topology Routing
      Router Priority: 0
      Router Dead Interval: 4 seconds
      Designated Router: 161.81.0.40
      Backup Designated Router: 0.0.0.0
      Active Neighbor: 166.96.0.40
```

参考标准

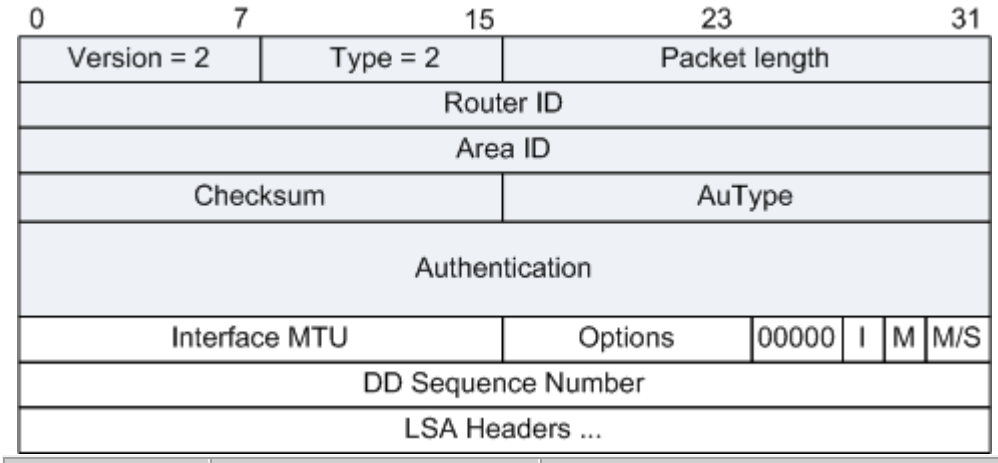
标准	描述
RFC2328	OSPF Version 2

4.11.3 OSPF DD 报文格式

报文格式

两台路由器在邻接关系初始化时，用 DD 报文（Database Description Packet）来描述自己的 LSDB，进行数据库的同步。报文内容包括 LSDB 中每一条 LSA 的 Header（LSA 的 Header 可以唯一标识一条 LSA）。LSA Header 只占一条 LSA 的整个数据量的一小部分，这样可以减少路由器之间的协议报文流量，对端路由器根据 LSA Header 就可以判断出是否已有这条 LSA。在两台路由器交换 DD 报文的过程中，一台为 Master，另一台为 Slave。由 Master 规定起始序列号，每发送一个 DD 报文序列号加 1，Slave 方使用 Master 的序列号作为确认。

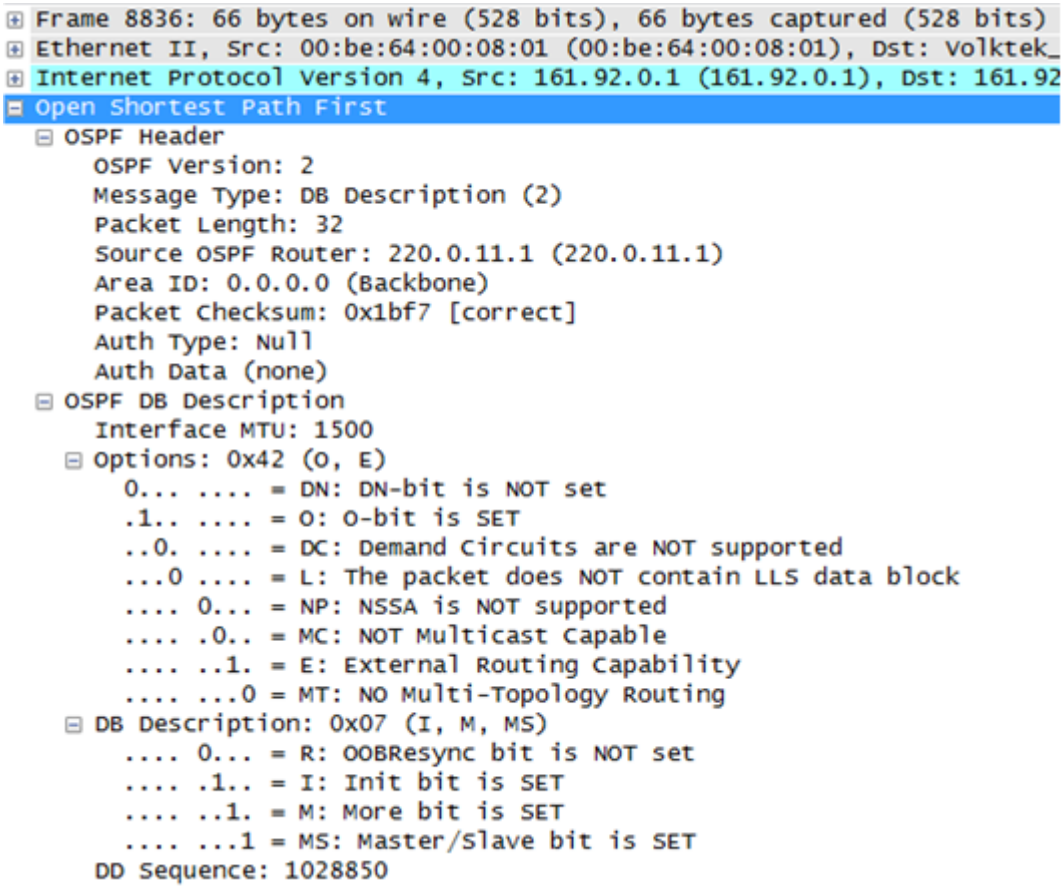
图 1 DD 报文格式



字段	长度	含义
Interface MTU	16 比特	在不分片的情况下，此接口最大可发出的 IP 报文长度。
Options	8 比特	可选项： • E：允许 Flood AS-External-LSAs； • MC：转发 IP 组播报文； • N/P：处理 Type-7 LSAs； • DC：处理按需链路。
I	1 比特	当发送连续多个 DD 报文时，如果这是第一个 DD 报文，则置为 1，否则置为 0。
M (More)	1 比特	当发送连续多个 DD 报文时，如果这是最后一个 DD 报文，则置为 0。否则置为 1，表示后面还有其他的 DD 报文。
M/S (Master/Slave)	1 比特	当两台 OSPF 路由器交换 DD 报文时，首先需要确定双方的主从关系，Router ID 大的一方会成为 Master。当值为 1 时表示发送方为 Master。
DD sequence number	32 比特	DD 报文序列号。主从双方利用序列号来保证 DD 报文传输的可靠性和完整性。
LSA Headers	可变	该 DD 报文中所包含的 LSA 的头部信息。

报文示例

图 2 OSPF DD 报文

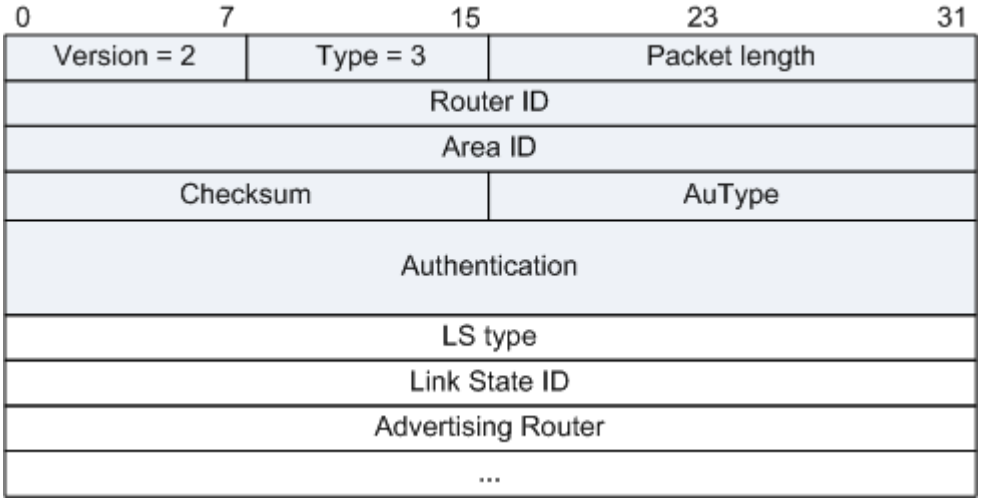


参考标准

标准	描述
RFC2328	OSPF Version 2

4.11.4 OSPF LSR 报文格式

两台路由器互相交换过 DD 报文之后，知道对端的路由器有哪些 LSA 是本地的 LSDB 所缺少的和哪些 LSA 是已经失效的，这时需要发送 LSR 报文（Link State Request Packet）向对方请求所需的 LSA。内容包括所需要的 LSA 的摘要。LSR 报文格式如下图所示，其中 LS type、Link State ID 和 Advertising Router 可以唯一标识出一个 LSA，当两个 LSA 一样时，需要根据 LSA 中的 LS sequence number、LS checksum 和 LS age 来判断出所需要 LSA 的新旧。



字段	长度	含义
LS type	32 比特	LSA 的类型号。
Link State ID	32 比特	根据 LSA 中的 LS Type 和 LSA description 在路由域中描述一个 LSA。
Advertising Router	32 比特	产生此 LSA 的路由器的 Router ID。

报文示例

图 1 OSPF LSR

```
Frame 8853: 70 bytes on wire (560 bits), 70 bytes captured (560 b
Ethernet II, Src: 1b:00:00:40:64:00 (1b:00:00:40:64:00), Dst: Voll
Internet Protocol Version 4, Src: 161.92.0.40 (161.92.0.40), Dst:
Open Shortest Path First
  OSPF Header
      OSPF Version: 2
      Message Type: LS Request (3)
      Packet Length: 36
      Source OSPF Router: 166.96.0.40 (166.96.0.40)
      Area ID: 0.0.0.0 (Backbone)
      Packet Checksum: 0x894b [correct]
      Auth Type: Null
      Auth Data (none)
  Link State Request
      Link-State Advertisement Type: Router-LSA (1)
      Link State ID: 220.0.11.1
      Advertising Router: 220.0.11.1 (220.0.11.1)
```

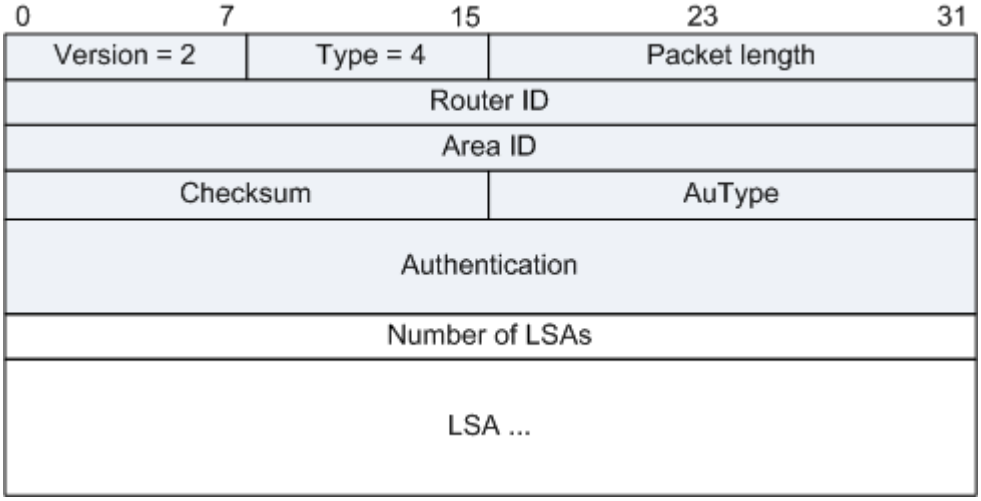
参考标准

标准	描述

标准	描述
RFC2328	OSPF Version 2

4.11.5 OSPF LSU 报文格式

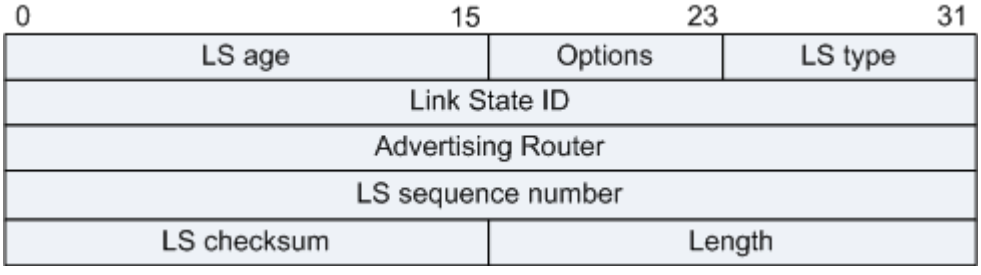
用来向对端 Router 发送其所需要的 LSA 或者泛洪自己更新的 LSA，内容是多条 LSA（全部内容）的集合。LSU 报文（Link State Update Packet）在支持组播和广播的链路上是以组播形式将 LSA 泛洪出去。为了实现 Flooding 的可靠性传输，需要 LSAck 报文对其进行确认。对没有收到确认报文的 LSA 进行重传，重传的 LSA 是直接发送到邻居的。



字段	长度	含义
Number of LSAs	32 比特	LSA 的数量。

常用的 LSA 共有 5 种，分别为：Router-LSA、Network-LSA、Network-summary-LSA、ASBR-summary-LSA 和 AS-External-LSA。

所有的 LSA 都有相同的报文头：



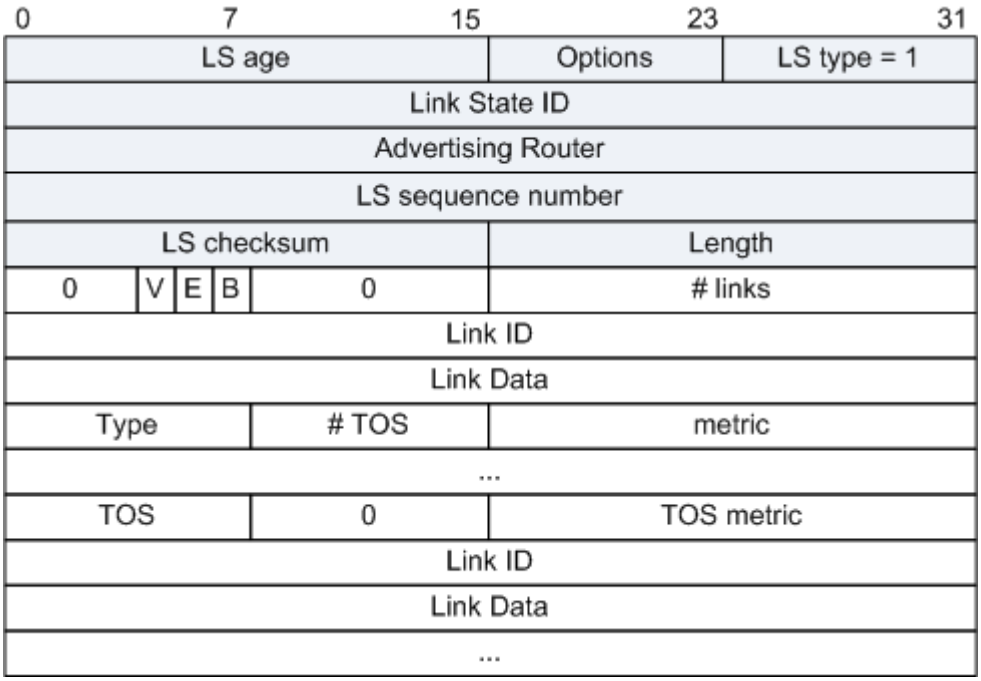
字段	长度	含义

字段	长度	含义
LS age	16 比特	LSA 产生后所经过的时间，以秒为单位。无论 LSA 是在链路上传送，还是保存在 LSDB 中，其值都会在不停的增长。
Options	8 比特	可选项： • E：允许泛洪 AS-External-LSA； • MC：转发 IP 组播报文； • N/P：处理 Type-7 LSA； • DC：处理按需链路。
LS type	8 比特	LSA 的类型： • Type1：Router-LSA • Type2：Network-LSA • Type3：Network-summary-LSA • Type4：ASBR-summary-LSA • Type5：AS-External-LSA • Type7：NSSA-LSA
Link State ID	32 比特	与 LSA 中的 LS Type 和 LSA description 一起在路由域中描述一个 LSA。
Advertising Router	32 比特	产生此 LSA 的路由器的 Router ID。
LS sequence number	32 比特	LSA 的序列号。其他路由器根据这个值可以判断哪个 LSA 是最新的。
LS checksum	16 比特	除了 LS age 外其它各域的校验和。
length	16 比特	LSA 的总长度，包括 LSA Header，以字节为单位。

Router-LSA

Router-LSA（Type1）：每个路由器都会产生，描述了路由器的链路状态和花费，在所属的区域内传播。

图 1 Router-LSA 格式



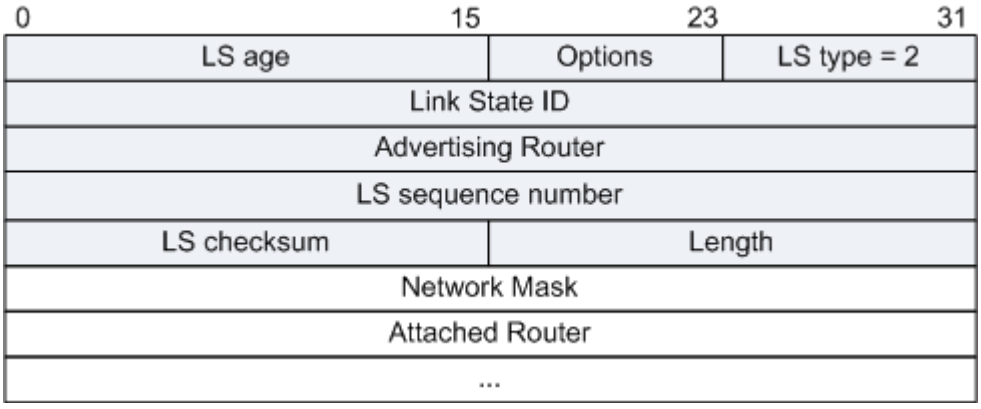
字段	长度	含义
Link State ID	32 比特	生成 LSA 的 Router ID。
V (Virtual Link)	1 比特	如果产生此 LSA 的路由器是虚连接的端点，则置为 1。
E (External)	1 比特	如果产生此 LSA 的路由器是 ASBR，则置为 1。
B (Border)	1 比特	如果产生此 LSA 的路由器是 ABR，则置为 1。
# links	16 比特	LSA 中所描述的链路信息的数量，包括路由器上处于某区域中的所有链路和接口。
Link ID	32 比特	路由器所接入的目标，其值取决于连接的类型： • 1：Router ID； • 2：DR 的接口 IP 地址； • 3：网段 / 子网号； • 4：虚连接中对端的 Router ID。
Link Data	32 比特	连接数据，其值取决于连接的类型： • unnumbered P2P：接口的索引值； • stub 网络：子网掩码；

字段	长度	含义
		• 其它连接：路由器接口的 IP 地址。
Type	8 比特	路由器连接的基本描述： • 1：点到点连接到另一台路由器； • 2：连接到传输网络； • 3：连接到 stub 网络； • 4：虚拟链路。
# TOS	8 比特	连接不同的 TOS 数量。
metric	16 比特	链路的开销值。
TOS	8 比特	服务类型。
TOS metric	16 比特	和指定 TOS 值相关联的度量。

Network-LSA

Network-LSA（Type2）：由广播网或 NBMA 网络中的 DR 产生，Network-LSA 中记录了这一网络上所有路由器的 Router ID，描述本网段的链路状态，在所属的区域内传播。

图 2 Network-LSA 格式



字段	长度	含义
Link State ID	32 比特	DR 的接口 IP 地址。
Network Mask	32 比特	该广播网或 NBMA 网络地址的掩码。
Attached Router	32 比特	连接在同一个网络上的所有路由器的 Router ID，也包括 DR 的 Router ID。

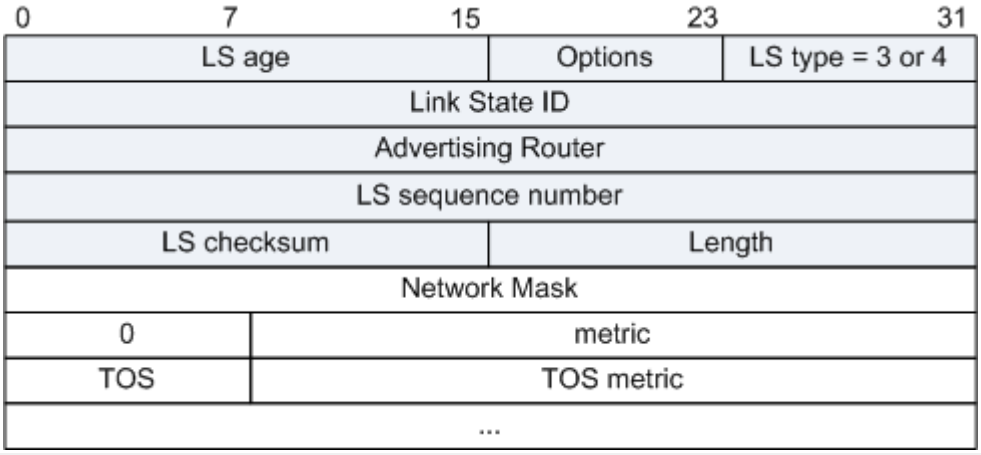
Summary-LSA 格式

Network-summary-LSA（Type3）：描述区域内所有网段的路由，并通告给其他相关区域。

ASBR-summary-LSA（Type4）：描述到 ASBR 的路由，通告给除 ASBR 所在区域的其他相关区域。

Type3 和 Type4 的 LSA 有相同的格式，它们都是由 ABR 产生。

图 3 Summary-LSA 格式



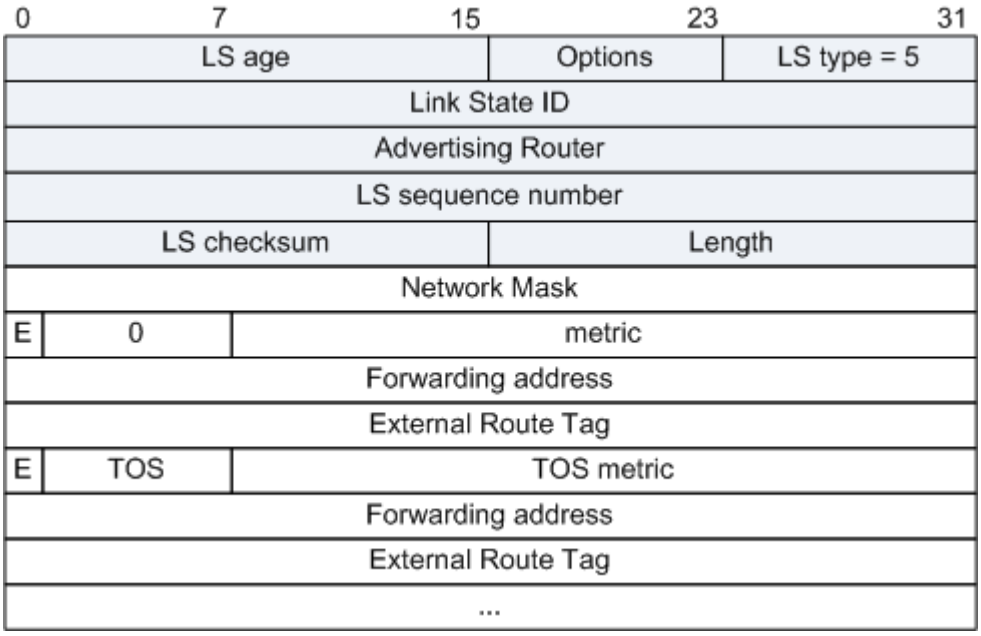
字段	长度	含义
Link State ID	32 比特	通告的网络地址。如果是 ASBR Summary LSA，此字段表示 ASBR 的 Router ID。
Network Mask	32 比特	该广播网或 NBMA 网络地址的掩码。如果是 ASBR Summary LSA，此字段无意义，设置为 0.0.0.0。
metric	24 比特	到目的地址的路由开销。
TOS	8 比特	服务类型。
TOS metric	24 比特	和指定 TOS 值相关联的度量。

通告缺省路由时，Link State ID 和 Network Mask 都设置为 0.0.0.0。

AS-External-LSA

AS-External-LSA（Type5）：由 ASBR 产生，描述到 AS 外部的路由，这是五种 LSA 中，唯一一种通告到所有区域（除了 Stub 区域和 NSSA 区域）的 LSA。

图 4 AS-External-LSA 格式



字段	长度	含义
Link State ID	32 比特	通告的网络地址。
Network Mask	32 比特	通告的目的地址的掩码。
E	1 比特	外部度量值类型： • 0：第一类外部路由； • 1：第二类外部路由。
metric	24 比特	到目的地址的路由开销。
Forwarding Address	32 比特	到所通告的目的地址的报文将被转发到这个地址。
External Route Tag	32 比特	添加到外部路由上的标记。OSPF 本身并不使用这个字段，它可以用来对外部路由进行管理。
TOS	8 比特	服务类型。
TOS metric	24 比特	TOS 附加距离信息。

Type5 的 LSA 可以用来通告缺省路由，此时 Link State ID 和 Network Mask 都设置为 0.0.0.0。

报文示例

图 5 Network-LSA

```
⊞ Frame 38246: 130 bytes on wire (1040 bits), 130 bytes captured (1040 bits)
⊞ Ethernet II, Src: e4:00:08:01:00:be (e4:00:08:01:00:be), Dst: MdiSecur_01:0
⊞ Internet Protocol Version 4, Src: 161.80.0.40 (161.80.0.40), Dst: 224.0.0.5
⊟ Open Shortest Path First
  ⊞ OSPF Header
  ⊟ LS Update Packet
      Number of LSAs: 2
    ⊞ LS Type: Router-LSA
    ⊟ LS Type: Network-LSA
        LS Age: 1 seconds
        Do Not Age: False
      ⊟ Options: 0x02 (E)
          0... .... = DN: DN-bit is NOT set
          .0.. .... = O: O-bit is NOT set
          ..0. .... = DC: Demand Circuits are NOT supported
          ...0 .... = L: The packet does NOT contain LLS data block
          .... 0... = NP: NSSA is NOT supported
          .... .0.. = MC: NOT Multicast Capable
          .... ..1. = E: External Routing Capability
          .... ...0 = MT: NO Multi-Topology Routing
        Link-State Advertisement Type: Network-LSA (2)
        Link State ID: 161.80.0.40
        Advertising Router: 166.96.0.40 (166.96.0.40)
        LS Sequence Number: 0x80000003
        LS Checksum: 0x6f86
        Length: 32
        Netmask: 255.255.255.0
        Attached Router: 166.96.0.40
        Attached Router: 220.0.11.2
```

图 6 Router-LSA

```
⊞ Frame 8811: 166 bytes on wire (1328 bits), 166 bytes captured (1328 bits)
⊞ Ethernet II, Src: 00:00:00_00:00:00 (00:00:00:00:00:00), Dst: Air2u_00:00:00
⊞ Internet Protocol Version 4, Src: 223.225.0.40 (223.225.0.40), Dst: 224.0.0.5
⊟ Open Shortest Path First
  ⊟ OSPF Header
      OSPF Version: 2
      Message Type: LS Update (4)
      Packet Length: 132
      Source OSPF Router: 166.96.0.40 (166.96.0.40)
      Area ID: 1.1.1.1
      Packet Checksum: 0xd255 [correct]
      Auth Type: Null
      Auth Data (none)
  ⊟ LS Update Packet
      Number of LSAs: 3
    ⊟ LS Type: Router-LSA
        LS Age: 21 seconds
        Do Not Age: False
      ⊟ Options: 0x02 (E)
          0... .... = DN: DN-bit is NOT set
          .0.. .... = O: O-bit is NOT set
          ..0. .... = DC: Demand Circuits are NOT supported
          ...0 .... = L: The packet does NOT contain LLS data block
          .... 0... = NP: NSSA is NOT supported
          .... .0.. = MC: NOT Multicast Capable
          .... ..1. = E: External Routing Capability
          .... ...0 = MT: NO Multi-Topology Routing
        Link-State Advertisement Type: Router-LSA (1)
        Link State ID: 223.225.0.1
        Advertising Router: 223.225.0.1 (223.225.0.1)
        LS Sequence Number: 0x80000001
        LS Checksum: 0x18f2
        Length: 36
      ⊟ Flags: 0x00
          .... .0.. = V: NO Virtual link endpoint
          .... ..0. = E: NO AS boundary router
          .... ...0 = B: NO Area border router
        Number of Links: 1
      ⊟ Type: Stub     ID: 223.225.0.0     Data: 255.255.255.0   Metric: 1
          IP network/subnet number: 223.225.0.0
          Link Data: 255.255.255.0
          Link Type: 3 - Connection to a stub network
          Number of TOS metrics: 0
          TOS 0 metric: 1
    ⊞ LS Type: Router-LSA
    ⊞ LS Type: Network-LSA
```

图 7 Summary-LSA

```
⊞ Frame 47214: 90 bytes on wire (720 bits), 90 bytes captured (720 bits)
⊞ Ethernet II, Src: 00:be:64:00:08:01 (00:be:64:00:08:01), Dst: Volktek_00:
⊞ Internet Protocol Version 4, Src: 10.41.22.1 (10.41.22.1), Dst: 224.0.0.5
⊟ Open Shortest Path First
  ⊞ OSPF Header
  ⊟ LS Update Packet
      Number of LSAs: 1
    ⊟ LS Type: Summary-LSA (IP network)
        LS Age: 1 seconds
        Do Not Age: False
      ⊟ Options: 0x02 (E)
          0... .... = DN: DN-bit is NOT set
          .0.. .... = O: O-bit is NOT set
          ..0. .... = DC: Demand Circuits are NOT supported
          ...0 .... = L: The packet does NOT contain LLS data block
          .... 0... = NP: NSSA is NOT supported
          .... .0.. = MC: NOT Multicast Capable
          .... ..1. = E: External Routing Capability
          .... ...0 = MT: NO Multi-Topology Routing
        Link-State Advertisement Type: Summary-LSA (IP network) (3)
        Link State ID: 20.20.20.1
        Advertising Router: 10.41.22.1 (10.41.22.1)
        LS Sequence Number: 0x80000001
        LS Checksum: 0x5e75
        Length: 28
        Netmask: 255.255.255.255
        Metric: 2
```

参考标准

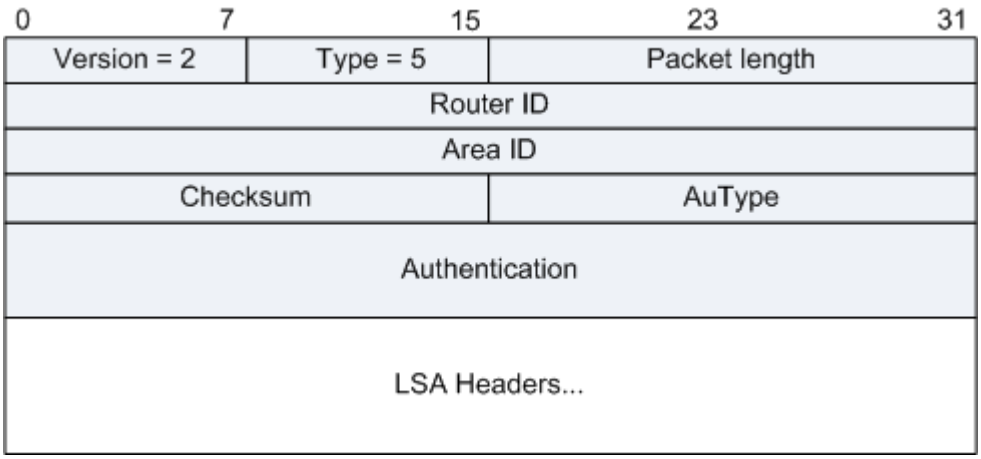
标准	描述
RFC2328	OSPF Version 2

4.11.6 OSPF LSAck 报文格式

报文格式

用来对接收到的 LSU 报文进行确认。内容是需要确认的 LSA 的 Header（一个 LSAck 报文可对多个 LSA 进行确认）。LSAck（Link State Acknowledgment Packet）报文根据不同的链路以单播或组播的形式发送。

图 1 OSPF LSAck 报文格式



字段	长度	含义
LSAs Headers	可变	通过 LSA 的头部信息确认收到该 LSA。

报文示例

图 2 OSPF LSAck

```
⊞ Frame 8815: 118 bytes on wire (944 bits), 118 bytes captured (944 bits)
⊞ Ethernet II, Src: 00:00:00_00:00:00 (00:00:00:00:00:00), Dst: Air2u_00:00:00
⊞ Internet Protocol Version 4, Src: 223.225.0.1 (223.225.0.1), Dst: 224.0.0.6
⊟ Open Shortest Path First
  ⊟ OSPF Header
      OSPF Version: 2
      Message Type: LS Acknowledge (5)
      Packet Length: 84
      Source OSPF Router: 223.225.0.1 (223.225.0.1)
      Area ID: 1.1.1.1
      Packet Checksum: 0xa1c4 [correct]
      Auth Type: Null
      Auth Data (none)
  ⊟ LSA Header
      LS Age: 6 seconds
      Do Not Age: False
    ⊟ Options: 0x02 (E)
        0... .... = DN: DN-bit is NOT set
        .0.. .... = O: O-bit is NOT set
        ..0. .... = DC: Demand Circuits are NOT supported
        ...0 .... = L: The packet does NOT contain LLS data block
        .... 0... = NP: NSSA is NOT supported
        .... .0.. = MC: NOT Multicast Capable
        .... ..1. = E: External Routing Capability
        .... ...0 = MT: NO Multi-Topology Routing
      Link-State Advertisement Type: Router-LSA (1)
      Link State ID: 166.96.0.40
      Advertising Router: 166.96.0.40 (166.96.0.40)
      LS Sequence Number: 0x80000002
      LS Checksum: 0xac84
      Length: 36
  ⊞ LSA Header
  ⊞ LSA Header
```

参考标准

标准	描述
RFC2328	OSPF Version 2

4.12 OSPFv3 报文格式

- OSPFv3 报文头格式
- OSPFv3 Hello 报文格式
- OSPFv3 DD 报文格式
- OSPFv3 LSR 报文格式
- OSPFv3 LSU 报文格式

- OSPFv3 LSAck 报文格式

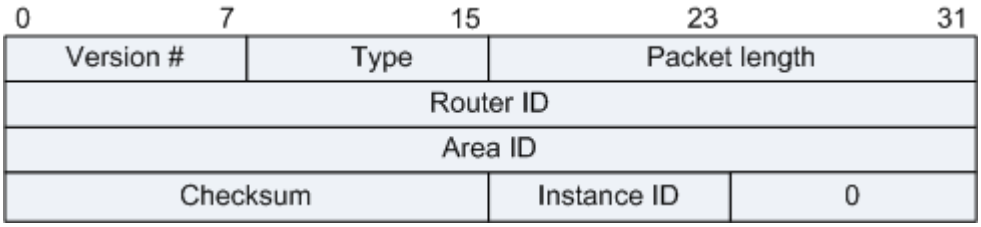
父主题： 网络层

4.12.1 OSPFv3 报文头格式

OSPFv3 用 IPv6 报文直接封装协议报文，协议号为 89，在 IPv6 Next Header 里标识。OSPFv3 分为 5 种报文：Hello 报文、DD 报文、LSR 报文、LSU 报文和 LSAck 报文。

OSPFv3 报文头格式

OSPFv3 这五种报文具有相同的报文头格式，长度为 24 字节。



字段	长度	含义
Version	1 字节	版本，OSPF 的版本号。对于 OSPFv3 来说，其值为 3。
Type	1 字节	类型，OSPFv3 报文的类型，有下面几种类型： • 1：Hello 报文； • 2：DD 报文； • 3：LSR 报文； • 4：LSU 报文； • 5：LSAck 报文。
Packet length	2 字节	OSPFv3 报文的总长度，包括报文头在内，单位为字节。
Router ID	4 字节	始发此包的路由器的 Router ID。
Area ID	4 字节	发送该报文的所属区域。
Checksum	2 字节	使用 IPv6 标准 16 位校验和。校验内容包括前导的 IPv6 伪头和 OSPF 协议包头。伪头中的 Upper-Layer Packet Length 字段值等于 OSPF 包头中的 Packet length 字段值。如果包长度不是 16 位的整数倍，则用 0 填充后进行计算。计算校验和时校验和字段本身设置为 0。
Instance ID	1 字节	缺省值为 0。允许在一个链路上运行多个 OSPFv3 的实例。每个实例应该具有唯一的

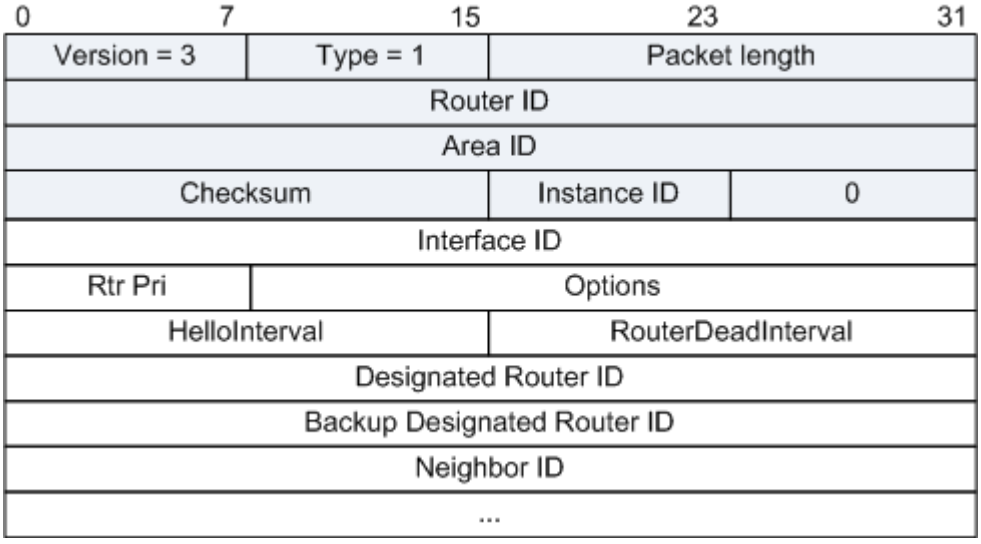
字段	长度	含义
		Instance ID。Instance ID 只在本地链路上有意义。如果接收到的 OSPF 包的 Instance ID 和本接口的 Instance ID 不同，则丢弃这个包。
0	1 字节	保留字段，必须填 0。

参考标准

标准	描述
RFC 2740	OSPF for IPv6

4.12.2 OSPFv3 Hello 报文格式

Hello 报文是最常用的一种报文，其作用为建立和维护邻接关系，周期性的在使能了 OSPF 的接口上发送。报文内容包括一些定时器的数值、DR、BDR 以及自己已知的邻居。



字段	长度	含义
Interface ID	32 比特	唯一标识了建立连接的（发送 Hello 报文的）接口。
Rtr Pri	8 比特	DR 优先级。默认为 1。如果设置为 0，则路由器不能参与 DR 或 BDR 的选举。
Options	24 比特	可选项： • E：允许 Flood AS-External-LSAs

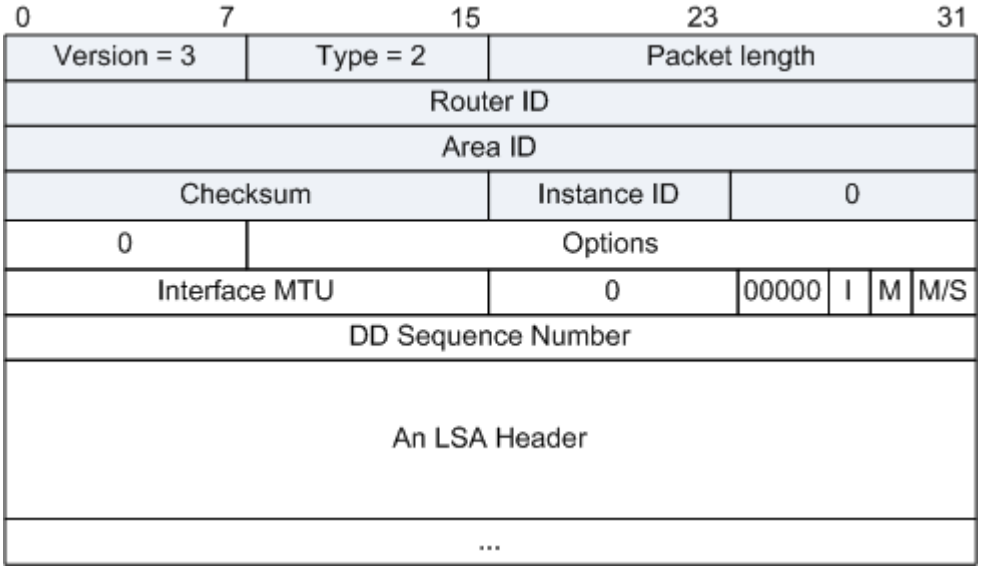
字段	长度	含义
		• MC：转发 IP 组播报文 • N/P：处理 Type-7 LSAs • DC：处理按需链路
HelloInterval	16 比特	发送 Hello 报文的时间间隔。
RouterDeadInterval	16 比特	失效时间。如果在此时间内未收到邻居发来的 Hello 报文，则认为邻居失效。
Designated Router ID	32 比特	DR 的接口地址。
Backup Designated Router ID	32 比特	BDR 的接口地址。
Neighbor ID	32 比特	邻居，以 Router ID 标识。

参考标准

标准	描述
RFC 2740	OSPF for IPv6

4.12.3 OSPFv3 DD 报文格式

两台路由器在邻接关系初始化时，用 DD 报文（Database Description Packet）来描述自己的 LSDB，进行数据库的同步。报文内容包括 LSDB 中每一条 LSA 的 Header（LSA 的 Header 可以唯一标识一条 LSA）。LSA Header 只占一条 LSA 的整个数据量的一小部分，这样可以减少路由器之间的协议报文流量，对端路由器根据 LSA Header 就可以判断出是否已有这条 LSA。在两台路由器交换 DD 报文的过程中，一台为 Master，另一台为 Slave。由 Master 规定起始序列号，每发送一个 DD 报文序列号加 1，Slave 方使用 Master 的序列号作为确认。



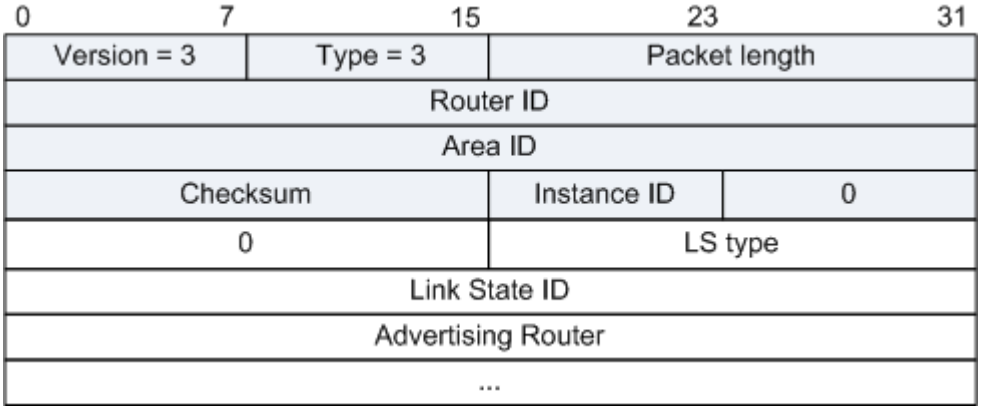
字段	长度	含义
Options	24 比特	可选项： • E：允许 Flood AS-External-LSAs； • MC：转发 IP 组播报文； • N/P：处理 Type-7 LSAs； • DC：处理按需链路。
Interface MTU	16 比特	在不分片的情况下，此接口最大可发出的 IP 报文长度。
I	1 比特	当发送连续多个 DD 报文时，如果这是第一个 DD 报文，则置为 1，否则置为 0。
M (More)	1 比特	当发送连续多个 DD 报文时，如果这是最后一个 DD 报文，则置为 0。否则置为 1，表示后面还有其他的 DD 报文。
M/S (Master/Slave)	1 比特	当两台 OSPF 路由器交换 DD 报文时，首先需要确定双方的主从关系，Router ID 大的一方会成为 Master。当值为 1 时表示发送方为 Master。
DD sequence number	32 比特	DD 报文序列号。主从双方利用序列号来保证 DD 报文传输的可靠性和完整性。
LSA Header	可变	该 DD 报文中所包含的 LSA 的头部信息。

参考标准

标准	描述
RFC 2740	OSPF for IPv6

4.12.4 OSPFv3 LSR 报文格式

两台路由器互相交换过 DD 报文之后，知道对端的路由器有哪些 LSA 是本地的 LSDB 所缺少的和哪些 LSA 是已经失效的，这时需要发送 LSR 报文（Link State Request Packet）向对方请求所需的 LSA。内容包括所需要的 LSA 的摘要。LSR 报文格式如下图所示，其中 LS type、Link State ID 和 Advertising Router 可以唯一标识出一个 LSA，当两个 LSA 一样时，需要根据 LSA 中的 LS sequence number、LS checksum 和 LS age 来判断出所需要 LSA 的新旧。



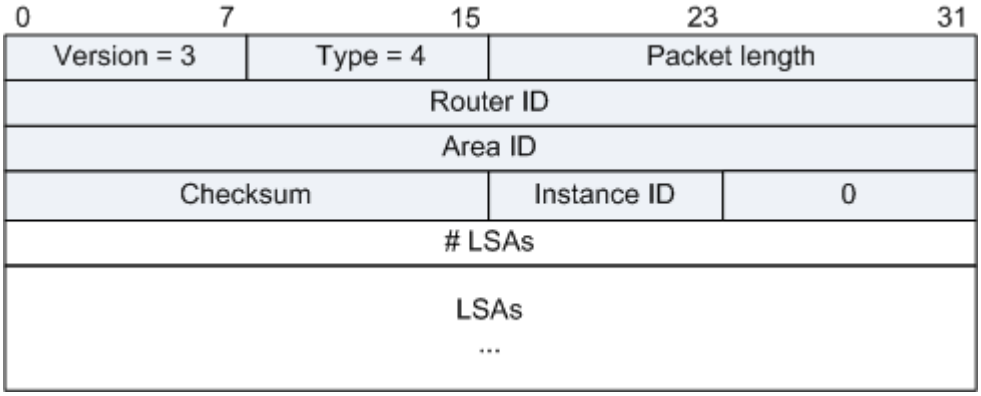
字段	长度	含义
LS type	16 比特	LSA 的类型号。
Link State ID	32 比特	根据 LSA 中的 LS Type 和 LSA description 在路由域中描述一个 LSA。
Advertising Router	32 比特	产生此 LSA 的路由器的 Router ID。

参考标准

标准	描述
RFC 2740	OSPF for IPv6

4.12.5 OSPFv3 LSU 报文格式

用来向对端 Router 发送其所需要的 LSA 或者泛洪自己更新的 LSA，内容是多条 LSA（全部内容）的集合。LSU 报文（Link State Update Packet）在支持组播和广播的链路上是以组播形式将 LSA 泛洪出去。为了实现 Flooding 的可靠性传输，需要 LSAck 报文对其进行确认。对没有收到确认报文的 LSA 进行重传，重传的 LSA 是直接发送到邻居的。

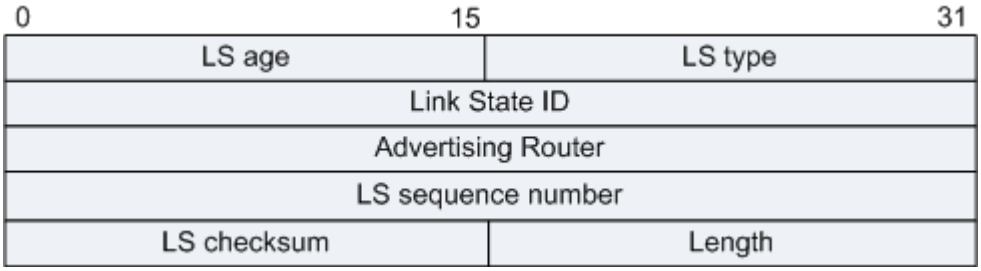


字段	长度	含义
# LSAs	32 比特	LSA 的数量。

OSPFv3 常用的 LSA 共有 7 种，分别为：

- Router-LSA（Type1）
- Network-LSA（Type2）
- Inter-Area-Prefix-LSA（Type3）
- Inter-Area-Router-LSA（Type4）
- AS-external-LSA（Type5）
- Link-LSA（Type8）
- Intra-Area-Prefix-LSA（Type9）

图 1 LSA 头部结构



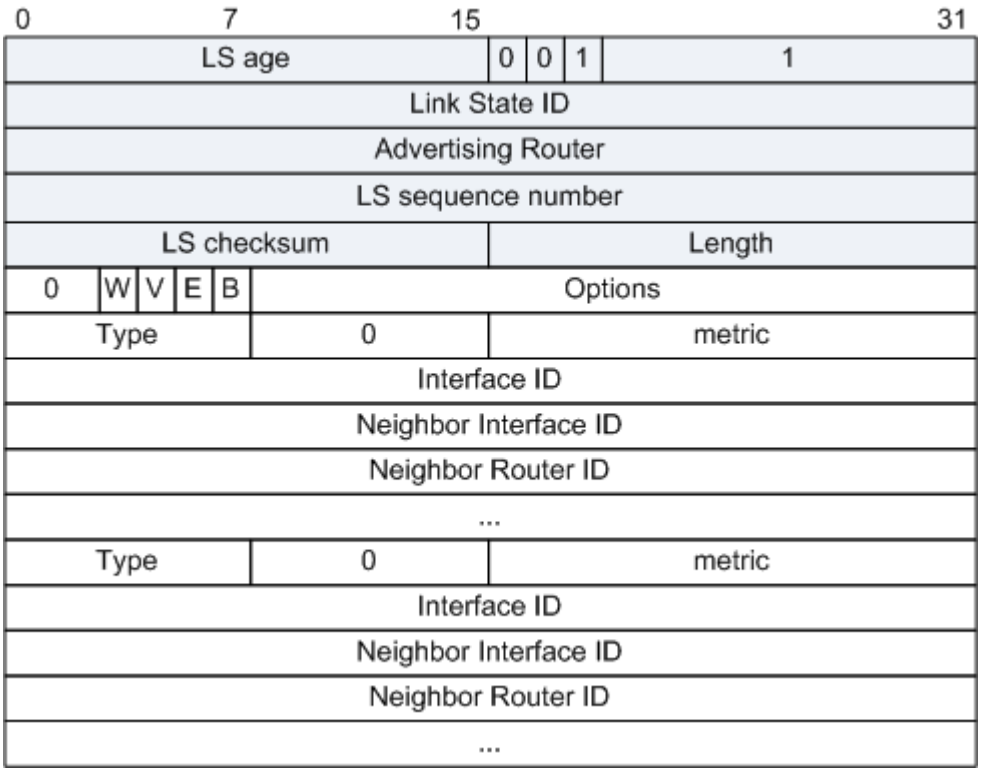
字段	长度	含义
LS age	16 比特	LSA 产生后所经过的时间，以秒为单位。无论 LSA 是在链路上传送，还是保存在 LSDB 中，其值都会在不停的增长。
LS type	16 比特	LSA 的类型，标识了 LSA 的功能。该字段的高 3 位标识 LSA 的通用属性，剩下的比特位标识 LSA 的特定功能。LS Type 字段的格式如下：

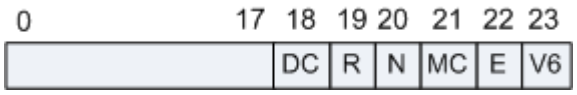
字段	长度	含义
		图 2 LS Type 字段格式  U 比特位标识标识对未知 LSA 的处理方法： • 0：把此 LSA 当作具有链路本地泛洪范围来对待，从而只能泛洪到本地链路上。 • 1：把此 LSA 当作类型已知的 LSA 来处理，也就是存储下来并泛洪出去。 S1 和 S2 比特位标识了 LSA 的泛洪范围： • S2 S1 = 0 0: 链路本地范围内，即只在始发链路上泛洪。 • S2 S1 = 0 1: 区域范围内，即泛洪到始发区域内的所有路由器。 • S2 S1 = 1 0: AS 范围内，即泛洪到本 AS 的所有路由器。 • S2 S1 = 1 1: 预留 LSA 的功能代码定义如下： • Type1: Router-LSA (LS Type = 0x2001) • Type2: Network-LSA (LS Type = 0x2002) • Type3: Inter-Area-Prefix-LSA (LS Type = 0x2003) • Type4: Inter-Area-Router-LSA (LS Type = 0x2004) • Type5: AS-external-LSA (LS Type = 0x2005) • Type8: Link-LSA (LS Type = 0x2008) • Type9: Intra-Area-Prefix-LSA (LS Type = 0x2009)
Link State ID	32 比特	与 LSA 中的 LS Type 和 LSA description 一起在路由域中描述一个 LSA。
Advertising Router	32 比特	产生此 LSA 的路由器的 Router ID。
LS sequence number	32 比特	LSA 的序列号。其他路由器根据这个值可以判断哪个 LSA 是最新的。
LS checksum	16 比特	除了 LS age 外其它各域的校验和。
length	16 比特	LSA 的总长度，包括 LSA Header，以字节为单位。

Router-LSA

Router-LSA（Type1）：每个路由器都会产生，描述了路由器的链路状态和花费，在所属的区域内传播。

图 3 Router-LSA 格式



字段	长度	含义
Link State ID	32 比特	生成 LSA 的 Router ID。
W	1 比特	如果置 1，标识该路由器是个组播“通吃者”(wild-card receiver)。当运行 MOSPF 时，无论目的地址是什么，这些路由器接收所有的组播数据。
V (Virtual Link)	1 比特	如果产生此 LSA 的路由器是虚连接的端点，则置为 1。
E (External)	1 比特	如果产生此 LSA 的路由器是 ASBR，则置为 1。
B (Border)	1 比特	如果产生此 LSA 的路由器是 ABR，则置为 1。
Options	24 比特	Options 字段使 OSPF 路由器能支持可选的能力，并且与其它路由器互相通告其能力。通过这种机制，具有不同能力的路由器可以在一个 OSPF 路由域中混合工作。其格式为： **图 4** Options 字段格式  • DC：当且仅当路由器可以正确处理 LSA 的 LS age 字段中出现的 DoNotAge 位的时候设置为

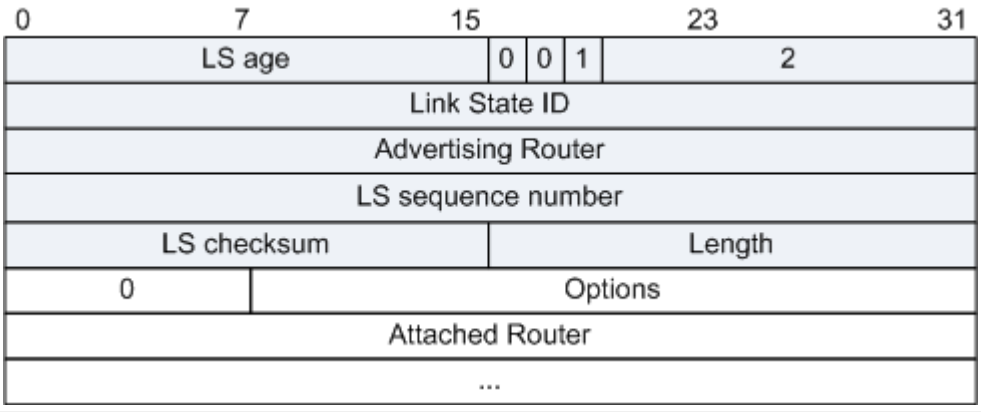
字段	长度	含义
		1。当且仅当路由器想把相连的点到点电路当作按需电路时，它把 Hello 和 DD 包的 DC 位设置为 1。在点到点按需电路上，两个端点设备都必须支持对 Hello 包的抑止，以便节省带宽资源。为了确保这一点，路由器在按需电路的接口发送的 Hello 包和 Database Description 包上把 DC 位设置为 1。如果对方发送的 DD 包或返回的 Hello 包的 DC 位为 0，说明对方不支持对 Hello 的抑止功能，所以路由器必须继续周期性地发送 Hello 包。 按需电路特性的支持只需要在两个端点之一配置即可。如果一个路由器实现了按需电路特性，但是没有配置，那么当它接收到设置了 DC 位的 Hello 包时，应该把这个点到点电路按照按需电路来对待，从而改变其相应进程。但是 LSA 中的 DC 位含义却有微妙的差别。设置 LSA 的 DC 位并不意味着路由器能够作为按需电路的一个端点，而是说明它能否正确处理设置了 DoNotAge 位的 LSA。当且仅当路由器可以正确处理 LSA 的 LS age 字段中出现的 DoNotAge 位的时候，它才把自己建立的 LSA 中的 DC 位设置为 1。 处于不同位置的 Options 字段会对协议运作产生不同的影响。某些选项的不匹配，会阻止邻接关系的形成，例如如果过 Hello 包的发送和接收，两个路由器发现 E 位不匹配，就不能形成邻居关系。某些选项的不匹配，会阻止特定类型的 LSA 的泛洪，例如如果通过 DD 包的交换，两个路由器发现彼此的 MC 位不匹配，那么组播路由使用的 Group-membership-LSA 就不能泛洪。某些选项如果不匹配，会使路由器不能被纳入某一种或多种路由的计算，例如如果路由器从 LSA 中发现某个路由器的 MC 位没有设置，那么在组播路由计算时它就不会考虑这个路由器。 在发送 Hello 包、发送 DD 包和建立 LSA 时，路由器应该对 Options 字段中不认可的位进行清零。在接收 Hello 包、接收 DD 包和接收 LSA 时，路由器应该忽略 OSPF Options 字段中不认可的位，并且正常处理这个包或 LSA。 • R：设置为 1。指出该公告者是否一个路由器。如果清零，则说明该公告者并不能路由数据。所以经过该公告者的路由不能纳入路由计算。如果多宿主机希望分享 OSPF 路由信息，但又不希望转发数据时，可以使用之。 • N：描述了路由器对 Type-7 LSA 的处理。当且仅当一个接口的所属区域为 NSSA 区域时设置为 1。 • MC：描述路由器是否运行了 MOSPF。当且仅当路由器运行 MOSPF 时设置为 1。 • E:当且仅当所属区域为 stub 区域时设置为 0。描述 AS-external-LSA 的泛洪方式。在 Hello 包中，当且仅当这个区域能够处理 AS-external-LSA 的时候，E 位设置为 1（例如在非 stub 区域中），否则为 0。如果 E 位设置不正确，邻接关系就不能形成。在 DD 包中，当一个链路所属的区域是非 stub 区域时，E 位设置为 1，否则为 0。而在 LSA 头中，E 位表现了相应的 LSA 的特性——骨干区域的 LSA、非 stub 区域的 LSA 和 AS-external-LSA 的 E 位都应该设置为 1；而在 stub 区域的 LSA 中应该设置为 0。LSA 中的 E 位设置仅仅是为了发布信息，而不影响路由表计算。 • V6：设置为 1。表示这个路由器或链路是不是在路由 IPv6。如果清零，这个路由器或链路

字段	长度	含义
		不应该纳入 IPv6 路由计算。 所有未定义的位都应该清零。
Type	8 比特	链路的类型: • 1: 点到点连接到另一台路由器 • 2: 连接到穿越(Transit)网 • 3: 保留 • 4: 虚连接
metric	16 比特	流量出接口的开销值。
Interface ID	32 比特	接口 ID。
Neighbor Interface ID	32 比特	邻居的接口 ID。
Neighbor Router ID	32 比特	邻居的路由器 ID。

Network-LSA

Network-LSA（Type2）：由广播网或 NBMA 网络中的 DR 产生，Network-LSA 中记录了这一网络上所有路由器的 Router ID，描述本网段的链路状态，在所属的区域内传播。

图 5 Network-LSA 格式



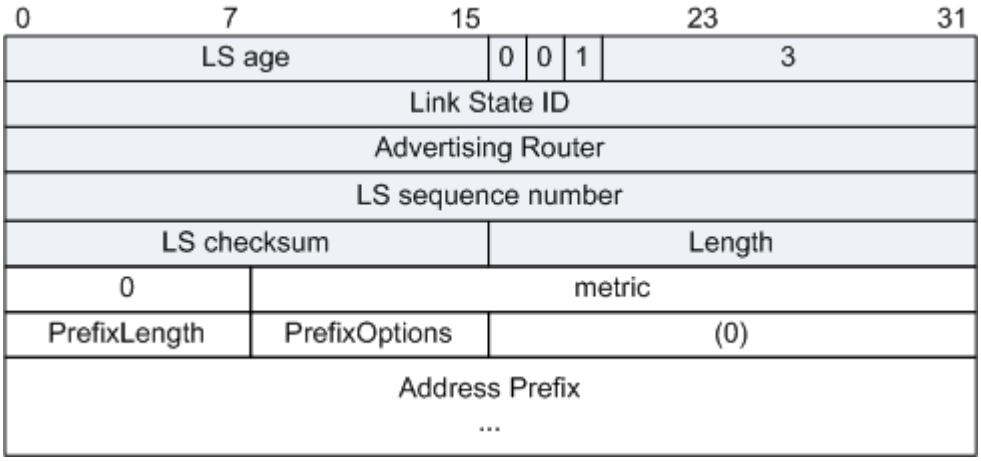
字段	长度	含义
Options	24 比特	参见 Router-LSA 的字段解释。
Attached	32 比特	连接在同一个网络上的所有路由器的 Router ID，也包括 DR 的 Router ID。

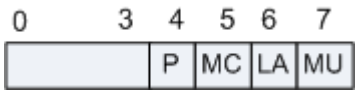
字段	长度	含义
Router		

Inter-Area-Prefix-LSA 格式

IPv6 的这些 LSA 与 IPv4 的 OSPFv2 中的 type 3 summary-LSAs 等同。由区域边界路由器始发，这些 LSA 描述了到其他区域的 IPv6 地址前缀。每个 IPv6 地址前缀单独发一个 Inter-Area-Prefix-LSA。

图 6 Inter-Area-Prefix-LSA 格式



字段	长度	描述
metric	24 bits	到目的地址的开销值。
PrefixLength	8 bits	前缀的比特数。
PrefixOptions	8 bits	用来表达这个前缀的一些特性，以便在各种不同的路由计算时做相应的判断和处理。例如希望在特定情况下忽略一个前缀的计算。 由 LSA 公告的每个前缀都拥有一个自己的 PrefixOptions 字段。PrefixOptions 字段格式如下： **图 7** PrefixOptions 字段格式  • P 位：传播位。如果一个 NSSA 区域的前缀需要被 ABR 传播出去，就需要设置这一位。 • MC 位：组播位。如果设置为 1，则这个前缀应该纳入组播计算中，否则不纳入组播计算。 • LA 位：本地地址位。如果设置为 1，则这个前缀是路由器的一个接口地址。

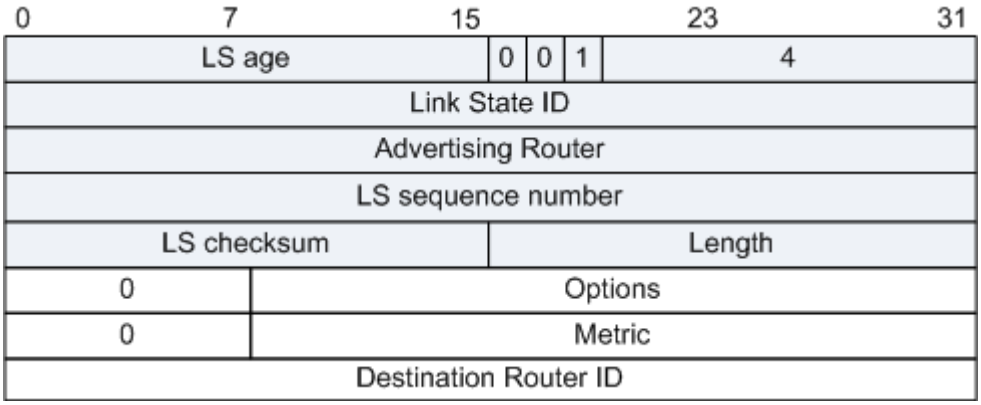
字段	长度	描述
		• NU 位：非单播位。如果设置为 1，则这个前缀不会纳入 IPv6 单播路由计算中。
Address Prefix	变长	IPv6 地址前缀。

缺省路由的前缀长度为 0。

Inter-Area-Router-LSA 格式

IPv6 的这些 LSA 与 IPv4 的 OSPFv2 中的 Type 4 summary-LSAs 等同。由区域边界路由器始发，这些 LSA 描述了到其他区域的 IPv6 地址前缀。每个 LSA 描述了到某台路由器的一条路由。

图 8 Inter-Area-Prefix-LSA 格式

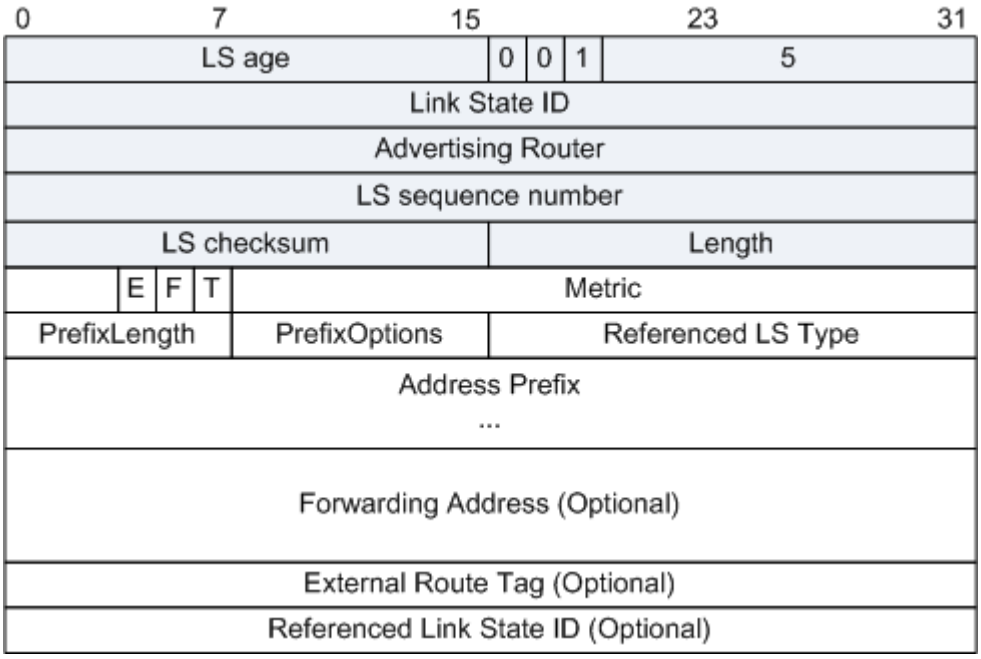


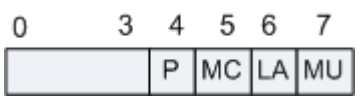
Field	Length	Description
Options	24 比特	Options 字段描述的不是源路由器的能力，而是目的路由器所支持的能力，所以此字段值应该等于目的路由器的 router-LSA 的 Options 字段值。
metric	24 比特	到目的地址的开销值。
Destination Router ID	32 比特	LSA 中描述的目的路由器的 Router ID。

AS-External-LSA

每个 AS-external-LSA 描述到达自治系统外部的一个前缀的路径。

图 9 AS-External-LSA 格式



字段	长度	含义
E	1 比特	外部路由的 Metric 类型。如果设置为 1，表示此为 2 类外部路由，其 Metric 不随着路由的传递而增长。如果设置为 0，表示此为 1 类外部路由，其 Metric 随着路由的传递而增长。
F	1 比特	如果设置为 1，则表示后面的 Forwarding Address 可选字段存在。
T	1 比特	如果设置为 1，则表示后面的 External Route Tag 可选字段存在。
metric	24 比特	到目的地址的路由开销。
PrefixLength	8 比特	前缀的比特数。
PrefixOptions	8 比特	用来表达这个前缀的一些特性，以便在各种不同的路由计算时做相应的判断和处理。例如希望在特定情况下忽略一个前缀的计算。 由 LSA 公告的每个前缀都拥有一个自己的 PrefixOptions 字段。PrefixOptions 字段格式如下： **图 10** PrefixOptions 字段格式  • P 位：传播位。如果一个 NSSA 区域的前缀需要被 ABR 传播出去，就需要设置这一位。

字段	长度	含义
		• MC 位：组播位。如果设置为 1，则这个前缀应该纳入组播计算中，否则不纳入组播计算。 • LA 位：本地地址位。如果设置为 1，则这个前缀是路由器的一个接口地址。 • NU 位：非单播位。如果设置为 1，则这个前缀不会纳入 IPv6 单播路由计算中。
Referenced LS type	16 比特	表明这个 LSA 是参考一个 Router-LSA，还是一个 Network-LSA。1 表示参考一个 router-LSA，2 表示参考一个 Network-LSA。
Address Prefix	变长	IPv6 地址前缀。
Forwarding Address	32 比特	可选的 128 位 Pv6 地址。当前面的 F 位为 1 时存在。表示到达目的的数据应该转发到这个地址。在公告路由器不是最优的下一跳的时候可以使用。
External Route Tag	32 比特	可选的标记位。可以用于 ASBR 之间的通信。一个比较常见的例子是，在 OSPF 自治系统的两个边界路由器上进行路由分发时，通过对引入的路由进行标记，可以很方便地进行路由过滤。
Referenced Link State ID	32 比特	当设置了前面的 Reference LS Type 时存在。如果存在，说明此条外部路由有一些相关信息需要参考另外一个 LSA。

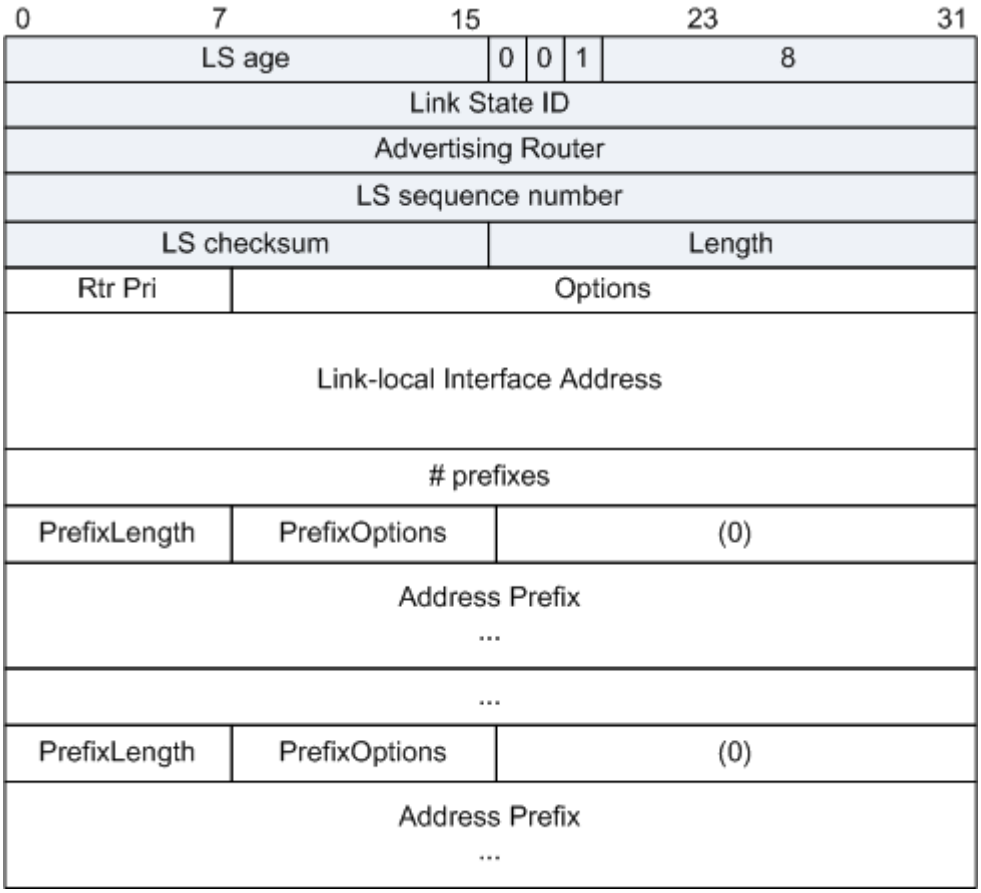
Link-LSA

每个连接的链路产生一个 Link-LSA。

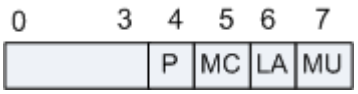
作用：

- 向该链路上其他路由器通知本地的 Link-Local 地址，即到本地的下一跳地址；
- 收集本路由器在该链路上配置的所有的 IPv6 前缀，并通知该链路上其他路由器；
- 向 Network-LSA 提供选项信息。收集该链路上所有的 Link-LSA，与操作得到的该链路上 Network-LSA 中选项信息。

图 11 Link-LSA format



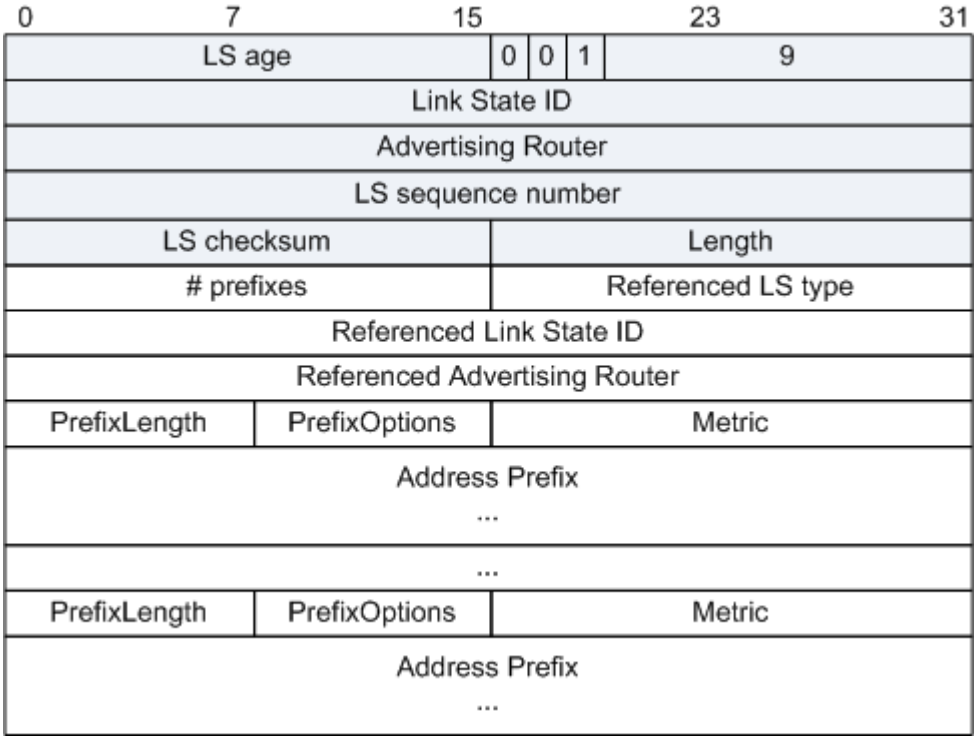
Field	Length	Description
Rtr Pri	8 bits	该路由器在该链路上的优先级(Router Priority)。
Options	24bits	提供给 Network LSA 的 Options。
Link-local Interface Address	128 bits	路由器与该链路相连的接口上配置的 Link Local 地址(Link Local 地址只出现在 Link LSA 中)。
# prefixes	32 bits	该 LSA 中携带了多少 IPv6 地址 Prefix。
PrefixLength	8 bits	前缀的比特数。
PrefixOptions	8 bits	用来表达这个前缀的一些特性，以便在各种不同的路由计算时做相应的判断和处理。例如希望在特定情况下忽略一个前缀的计算。 由 LSA 公告的每个前缀都拥有一个自己的 PrefixOptions 字段。 PrefixOptions 字段格式如下： **图 12** PrefixOptions 字段格式

Field	Length	Description
		 • P 位：传播位。如果一个 NSSA 区域的前缀需要被 ABR 传播出去，就需要设置这一位。 • MC 位：组播位。如果设置为 1，则这个前缀应该纳入组播计算中，否则不纳入组播计算。 • LA 位：本地地址位。如果设置为 1，则这个前缀是路由器的一个接口地址。 • NU 位：非单播位。如果设置为 1，则这个前缀不会纳入 IPv6 单播路由计算中。
Address Prefix	Variable	IPv6 地址前缀。

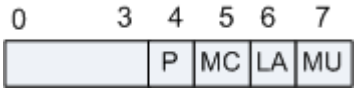
Intra-Area-Prefix-LSA

Intra-Area-Prefix-LSA 携带区域内 IPv6 Prefix 信息。

图 13 Intra-Area-Prefix-LSA 格式



字段	长度	描述
# prefixes	16 bits	在 LSA 中包含的 IPv6 前缀数量。必要的话，可以通过多个 intra-area-prefix-LSA 来携带前缀，这样可以控制 LSA 的长度。
Referenced LS type	16 bits	表明这个 LSA 是参考一个 Router-LSA，还是一个 Network-LSA。1 表示参考一个 router-LSA，2 表示参考一个 Network-LSA。

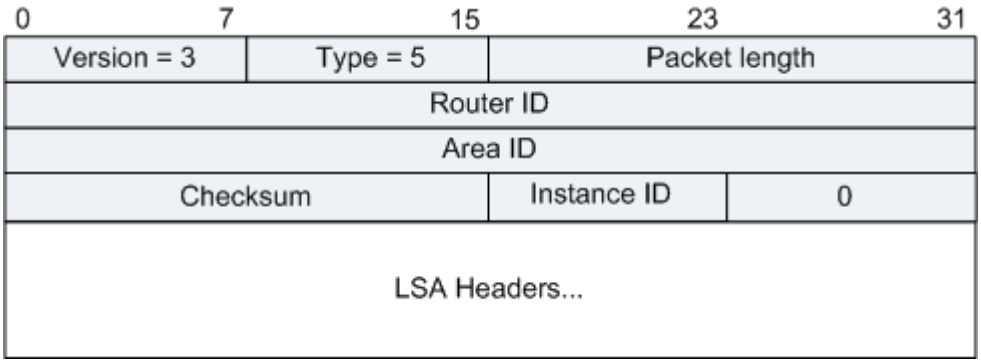
字段	长度	描述
Referenced Link State ID	32 bits	当这个 LSA 是参考一个 Router-LSA 时，设置为 0。当这个 LSA 是参考一个 Network-LSA 时，设置为该链路的 DR 的 Interface ID。
Referenced Advertising Router	32 bits	当这个 LSA 是参考一个 Router-LSA 时，设置为这个路由器的 Router ID。当这个 LSA 是参考一个 Network-LSA 时，设置为该链路的 DR 的 Router ID。
PrefixLength	8 bits	前缀的比特数。
PrefixOptions	8 bits	用来表达这个前缀的一些特性，以便在各种不同的路由计算时做相应的判断和处理。例如希望在特定情况下忽略一个前缀的计算。 由 LSA 公告的每个前缀都拥有一个自己的 PrefixOptions 字段。PrefixOptions 字段格式如下： **图 14** PrefixOptions 字段格式  • P 位：传播位。如果一个 NSSA 区域的前缀需要被 ABR 传播出去，就需要设置这一位。 • MC 位：组播位。如果设置为 1，则这个前缀应该纳入组播计算中，否则不纳入组播计算。 • LA 位：本地地址位。如果设置为 1，则这个前缀是路由器的一个接口地址。 • NU 位：非单播位。如果设置为 1，则这个前缀不会纳入 IPv6 单播路由计算中。
Metric	16 bits	前缀开销值。与 Router-LSA 的接口开销值相同单位。
Address Prefix	Variable	IPv6 地址前缀。

参考标准

标准	描述
RFC 2740	OSPF for IPv6

4.12.6 OSPFv3 LSAck 报文格式

用来对接收到的 LSU 报文进行确认。内容是需要确认的 LSA 的 Header（一个 LSAck 报文可对多个 LSA 进行确认）。LSAck（Link State Acknowledgment Packet）报文根据不同的链路以单播或组播的形式发送。



字段	长度	含义
LSAs Headers	可变	通过 LSA 的头部信息确认收到该 LSA。

参考标准

标准	描述
RFC 2740	OSPF for IPv6

4.13 PIM 报文格式

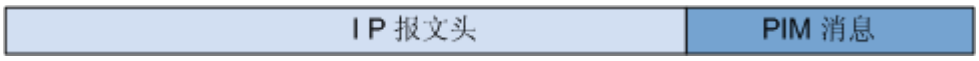
- PIM 报文通用格式
- PIM Hello 消息格式
- PIM Register 消息格式
- PIM Register-Stop 消息格式
- PIM Join/Prune 消息格式
- PIM Graft/Graft-Ack 消息格式
- PIM Bootstrap 消息格式
- PIM Assert 消息格式
- PIM C-RP Advertisement 消息格式

4.13.1 PIM 报文通用格式

PIM（Protocol Independent Multicast）称为协议无关组播，作为一种组播路由解决方案，也可以支持 IPv4 和 IPv6 网络，在实践中得到广泛的应用。

PIM 通过路由器之间交互 PIM 控制消息实现组播路由功能。PIM 控制消息使用 IP 报文封装。

图 1 PIM 消息的封装格式



- IP 报文头的协议类型字段值为 103，用来标识数据部分封装了 PIM 消息。
- IP 报文头的目的地址字段用来标识该 PIM 消息的目的接收者。可以是单播地址，也可以是组播地址。
- PIM-DM 协议与 PIM-SM 协议，支持不同的控制消息。

PIM 消息通用头部格式

所有的 PIM 控制消息头部有相同的格式，如下图：

图 2 PIM 消息头部格式

字段	长度	说明
Version	4 比特	PIM 版本，值为 2。
Type	4 比特	消息类型，取值如下： • 0：Hello（PIM-DM 与 PIM-SM 都适用） • 1：Register（只适用于 PIM-SM） • 2：Register-Stop（只适用于 PIM-SM） • 3：Join/Prune（PIM-DM 与 PIM-SM 都适用） • 4：Bootstrap（只适用于 PIM-SM） • 5：Assert（PIM-DM 与 PIM-SM 都适用） • 6：Graft（只适用于 PIM-DM） • 7：Graft-Ack（只适用于 PIM-DM） • 8：Candidate-RP-Advertisement（只适用于 PIM-SM） • 9：State Refresh（只适用于 PIM-DM）

字段	长度	说明
Reserved	8 比特	保留。
Checksum	16 比特	校验和。

参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)
RFC 3973	Protocol Independent Multicast - Dense Mode （PIM-DM）:Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

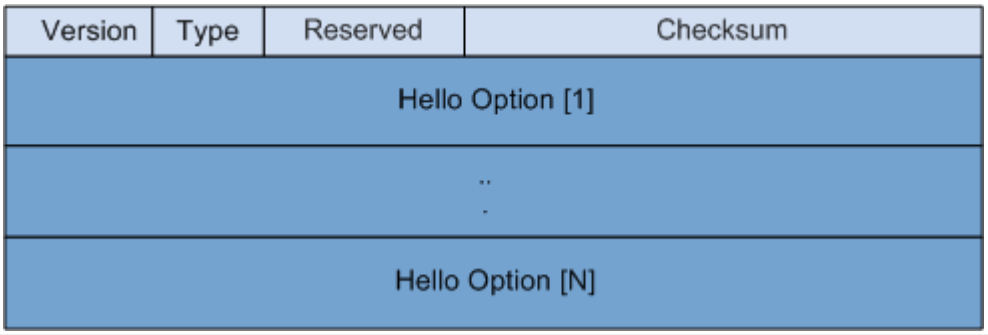
4.13.2 PIM Hello 消息格式

Hello 消息

PIM 路由器之间通过交互 Hello 消息，发现 PIM 邻居并维护邻居关系。Hello 消息同时在 PIM-DM 与 PIM-SM 中使用。Hello 消息中无法区分是 PIM-DM 还是 PIM-SM。

封装 Hello 消息的 IP 报文源地址为本地接口地址，目的地址为 224.0.0.13，TTL 值为 1。使用组播方式发送。

图 1 Hello 消息格式



字段	长度	说明
Version	4 比特	PIM 版本，值为 2。

字段	长度	说明
Type	4 比特	消息类型，值为 0。
Reserved	8 比特	保留字段，发送时设置为 0，接收时忽略此值。
Checksum	16 比特	校验和。
Hello Option [1]...[N]	16 比特	采用 Type-Length-Value（TLV）格式，其中： • Type：2 字节，Option 参数类型。 • Length：Value 字段的长度，字节为单位。 • Value：Option 参数值。 Type 值对应的参数名： • 1: Holdtime，表示保持邻居为可达状态的超时时间，若超时仍没有收到 Hello 消息则认为邻居不可达。 • 2: 该字段由三部分组成： LAN Prune Delay：在共享网段上传递 Prune 消息的延迟时间。 Override Interval：在共享网段上执行剪枝前的否决时间。 • T：Join 消息抑制能力位。 • 19: DR Priority，表示各路由器接口竞选 DR 的优先级，优先级越高越容易获胜。 • 20: Generation ID，Hello 消息中携带的随机数，表示当前邻居状态。如果状态发生更新则随机数也会更新。当路由器发现接收到的来自上游的 Hello 消息中包含不同 Generation ID 值，则认为上游邻居已经丢失或上游邻居状态已经改变。 • 21: State Refresh Capable，表示邻居状态刷新时间间隔。 • 24: Address List，PIM 接口的从地址列表。

报文示例

图 2 PIM Hello 消息（IPv4）

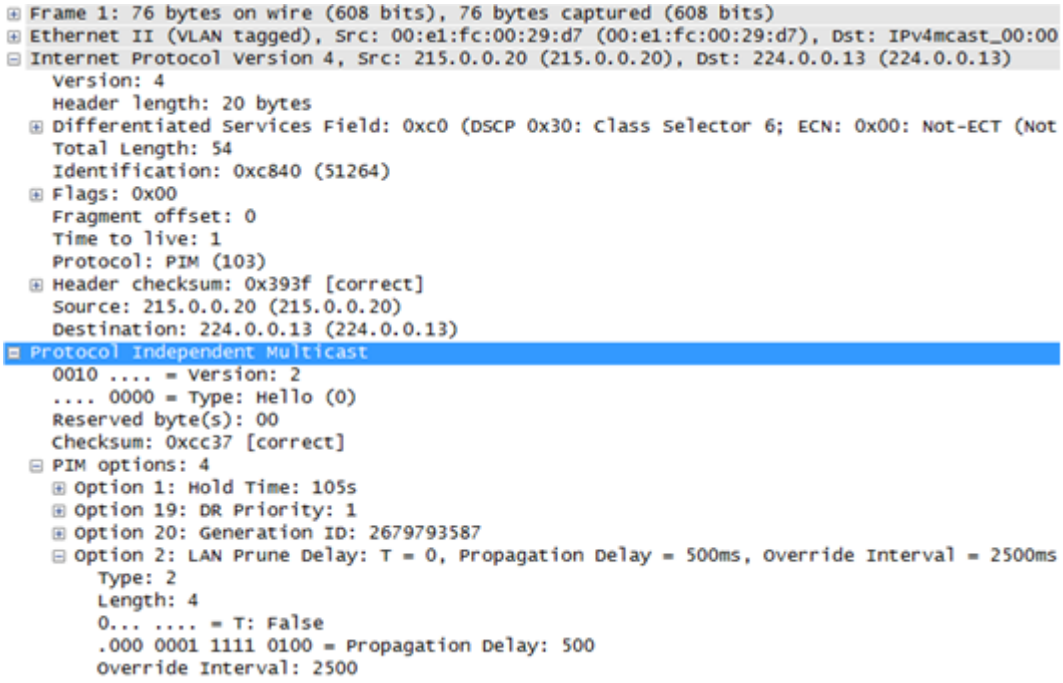
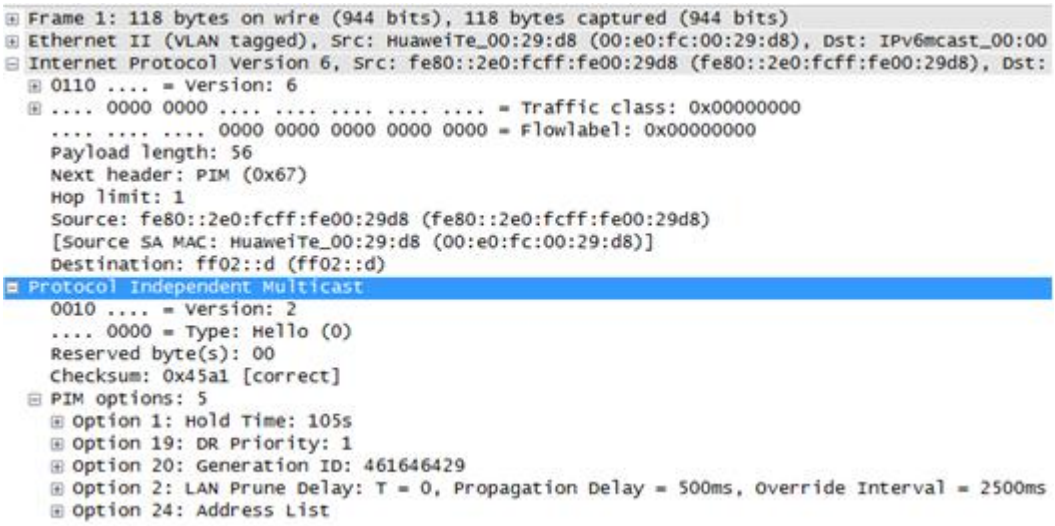


图 3 PIM Hello 消息（IPv6）



参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)
RFC 3973	Protocol Independent Multicast - Dense Mode (PIM-DM) :Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

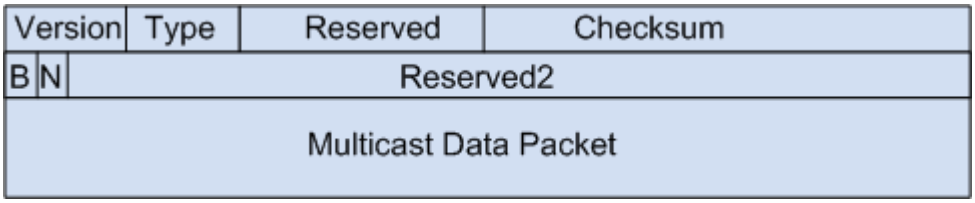
4.13.3 PIM Register 消息格式

Register 消息

当 PIM-SM 网络中出现活跃组播源时，源端 DR 向 RP 发送 Register 消息，进行源注册。Register 消息只在 PIM-SM 中使用。

封装 Register 消息的 IP 报文源地址为源端 DR，目的地址为 RP。使用单播方式发送。

图 1 Register 消息格式



字段	长度	说明
Version	4 比特	PIM 版本，值为 2。
Type	4 比特	消息类型，值为 1。
Reserved	8 比特	保留位。发送时此字段被清零，接收时不处理此字段。
Checksum	16 比特	校验和。
B	1 比特	边界位。
N	1 比特	空注册位。
Reserved2	30 比特	保留位。发送时此字段被清零，接收时不处理此字段。
Multicast data packet	变长	组播数据报文。源端 DR 将接收到的组播数据报文封装在 Register 消息中发往 RP。RP 解封装后，学习到该组播数据报文的（S，G）信息。

报文示例

图 2 PIM Register 消息

```
⊞ Frame 1: 78 bytes on wire (624 bits), 78 bytes captured (624 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_00:29:d6 (00:e0:fc:00:29:d6), Dst: HuaweiTe
⊟ Internet Protocol Version 4, Src: 2.2.2.3 (2.2.2.3), Dst: 1.2.3.4 (1.2.3.4)
    Version: 4
    Header length: 20 bytes
  ⊞ Differentiated Services Field: 0xc0 (DSCP 0x30: Class Selector 6; ECN: 0x00: Not-E
    Total Length: 56
    Identification: 0x4d90 (19856)
  ⊞ Flags: 0x00
    Fragment offset: 0
    Time to live: 255
    Protocol: PIM (103)
  ⊞ Header checksum: 0x6504 [correct]
    Source: 2.2.2.3 (2.2.2.3)
    Destination: 1.2.3.4 (1.2.3.4)
⊟ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0001 = Type: Register (1)
    Reserved byte(s): 00
    Checksum: 0xdeff [correct]
  ⊟ PIM options
    ⊟ Flags: 0x00000000
        0... .... .... .... .... .... .... .... = Not border
        .0.. .... .... .... .... .... .... .... = Not Null-Register
    ⊟ Internet Protocol Version 4, Src: 2.2.2.2 (2.2.2.2), Dst: 228.1.1.1 (228.1.1.1)
        Version: 4
        Header length: 20 bytes
      ⊞ Differentiated Services Field: 0xc0 (DSCP 0x30: Class Selector 6; ECN: 0x00: N
        Total Length: 28
        Identification: 0x0000 (0)
      ⊞ Flags: 0x00
        Fragment offset: 0
        Time to live: 254
        Protocol: UDP (17)
      ⊞ Header checksum: 0xd30a [correct]
        Source: 2.2.2.2 (2.2.2.2)
        Destination: 228.1.1.1 (228.1.1.1)
    ⊟ User Datagram Protocol, Src Port: 0 (0), Dst Port: 0 (0)
        Source port: 0 (0)
        Destination port: 0 (0)
        Length: 8
      ⊟ Checksum: 0xfff7 [validation disabled]
          [Good Checksum: False]
          [Bad Checksum: False]
```

参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)
RFC 3973	Protocol Independent Multicast - Dense Mode (PIM-DM) :Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

4.13.4 PIM Register-Stop 消息格式

Register-Stop 消息

在 PIM-SM 网络中，在以下三种情况下，RP 将会向组播源端 DR 发送 Register-Stop 消息。

- 接收者不再通过 RP 接收发往某组播组的数据
- RP 不再为某组播组服务

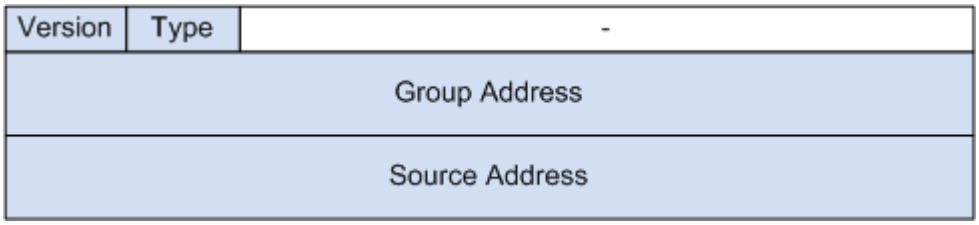
- 组播数据转发路径已经由 RPT 切换到 SPT

组播源端 DR 收到 Register-Stop 消息后，停止使用 Register 注册消息封装组播数据报文，并进入注册抑制状态。

Register-Stop 消息只在 PIM-SM 中使用。

封装 Register-Stop 消息的 IP 报文源地址为 RP，目的地址为源端 DR。使用单播方式发送。

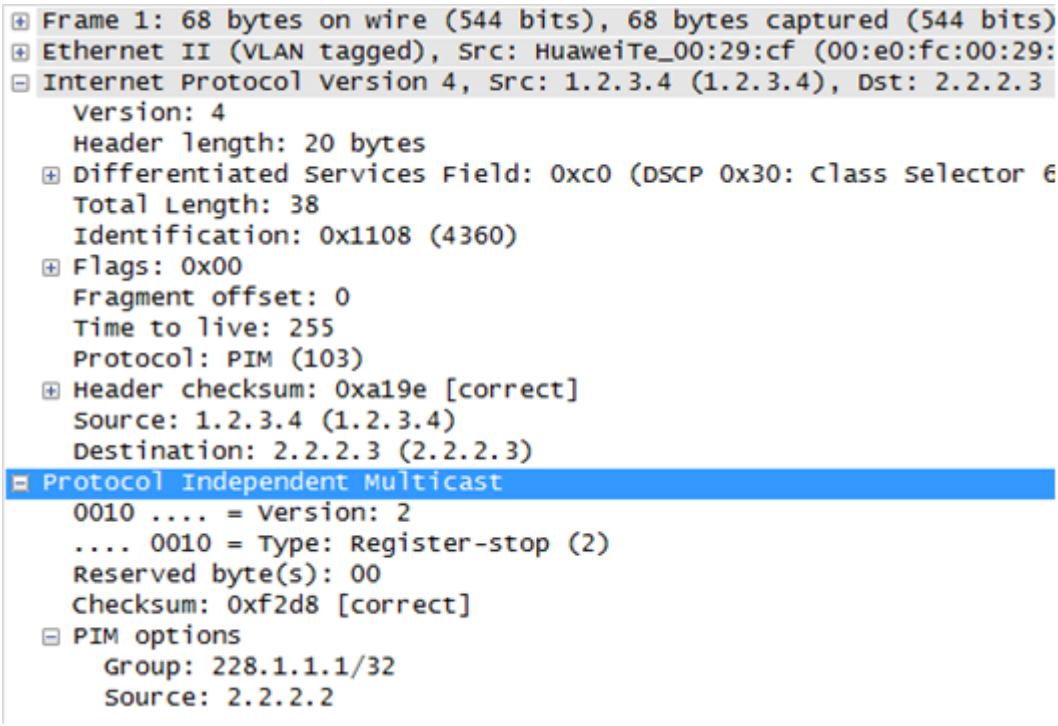
图 1 Register-Stop 消息格式



字段	长度	说明
Version	4 比特	PIM 版本，值为 2。
Type	4 比特	消息类型，值为 2。
Reserved	24 比特	保留位。发送时此字段被清零，接收时不处理此字段。
Group Address	32 比特	组播组地址 G。
Source Address	32 比特	组播源地址 S。

报文示例

图 2 PIM Register-Stop 消息



参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)
RFC 3973	Protocol Independent Multicast - Dense Mode （PIM-DM）:Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

4.13.5 PIM Join/Prune 消息格式

Join/Prune 消息

一条 Join/Prune 消息中可以同时包含 Join 信息和 Prune 信息。只包含 Join 信息的 Join/Prune 消息称为 Join 消息。只包含 Prune 信息的 Join/Prune 消息称为 Prune 消息。

Join/Prune 消息同时在 PIM-DM 和 PIM-SM 中使用。

封装 Join/Prune 消息的 IP 报文源地址为本地接口地址，目的地址为 224.0.0.13，TTL 值为 1。使用组播方式发送。

图 1 Join/Prune 消息格式

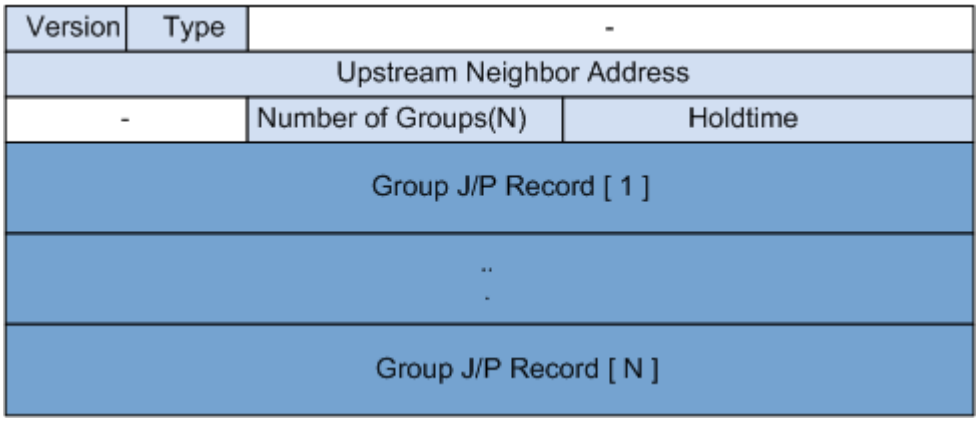
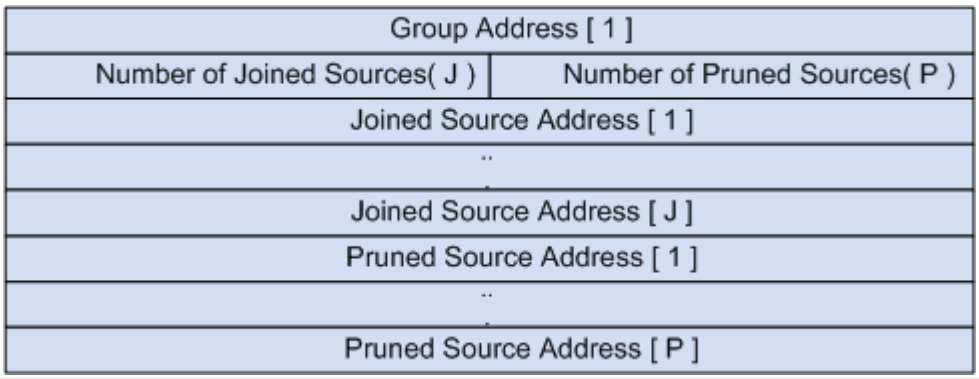


图 2 Group J/P Record 字段格式



字段	长度	说明
Version	4 比特	PIM 版本，值为 2。
Type	4 比特	消息类型，值为 3。
Upstream Neighbor Address	32 比特	上游邻居地址。也就是收到 Join/Prune 消息的路由器上，进行 Join 或 Prune 操作的下游接口地址。
Number of Groups	8 比特	消息中包含的组播组数目。
Holdtime	16 比特	接收 Join/Prune 消息的路由器保持相应接口加入/剪枝状态的时间。
Group Address	32 比特	组播组地址。
Number of Joined Sources	16 比特	针对该组播组，请求加入的组播源总数。

字段	长度	说明
Number of Pruned Sources	16 比特	针对该组播组，请求剪枝的组播源总数。
Joined Source Address	32 比特	请求加入的组播源地址。
Pruned Source Address	32 比特	请求剪枝的组播源地址。

报文示例

图 3 PIM Join 消息

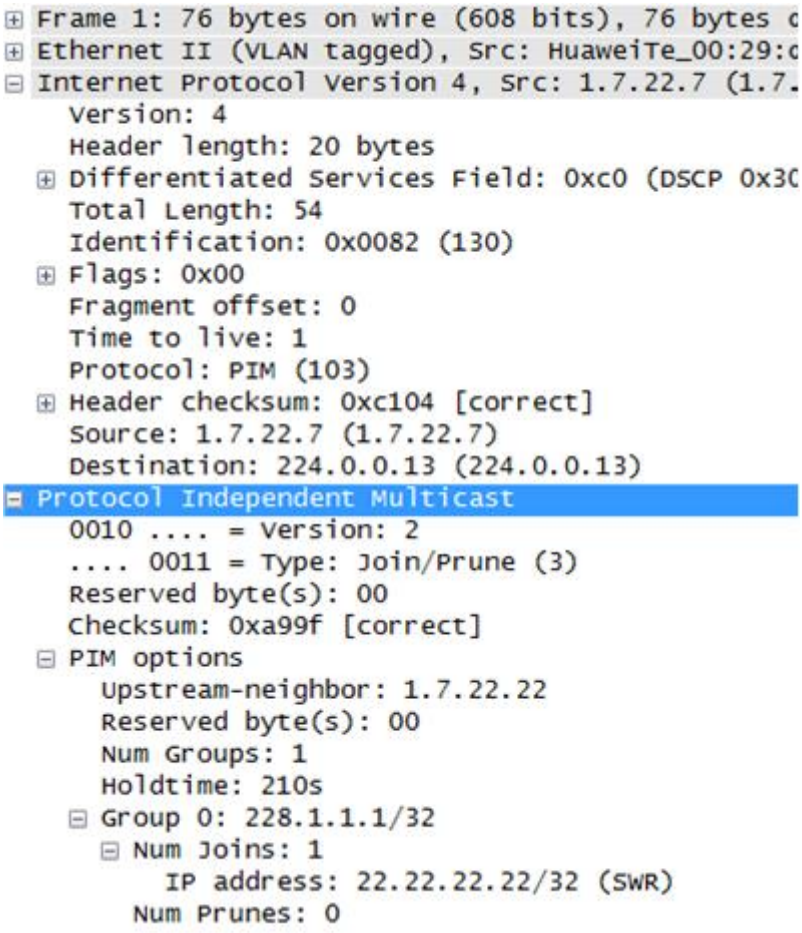
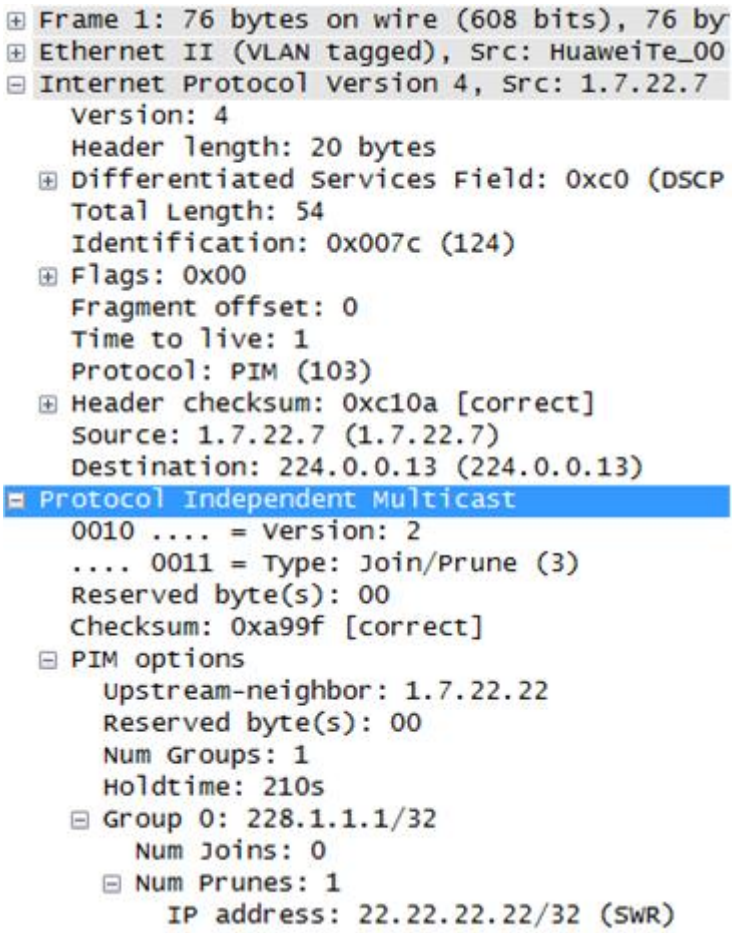


图 4 PIM Prume 消息



参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)
RFC 3973	Protocol Independent Multicast - Dense Mode (PIM-DM) :Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

4.13.6 PIM Graft/Graft-Ack 消息格式

Graft/Graft-Ack 消息

在 PIM-DM 网络中，路由器上出现组成员时，如果本身不在 SPT 上，则从对应的（S，G）表项的上游接口发送 Graft 消息。上游邻居立即恢复下游接口的转发，同时从该下游接口发出 Graft-Ack 消息，表示已经接受嫁接请求。如果上游邻居不在 SPT 上，则继续向上游发送 Graft 消息。

封装 Graft-Ack 消息的 IP 报文源地址为下游接口地址，目的地址为 Graft 消息的发出者。使用单播方式发送。

Graft 消息格式与 Join/Prune 消息相同，仅部分字段取值存在差异。

Graft-Ack 消息与 Graft 消息格式相同，并复制了 Graft 消息的内容。其中不同的是，Upstream Neighbor Address 字段，在 Graft-Ack 消息中填为 Graft 消息的发出者地址。

图 1 Join/Prune 消息格式

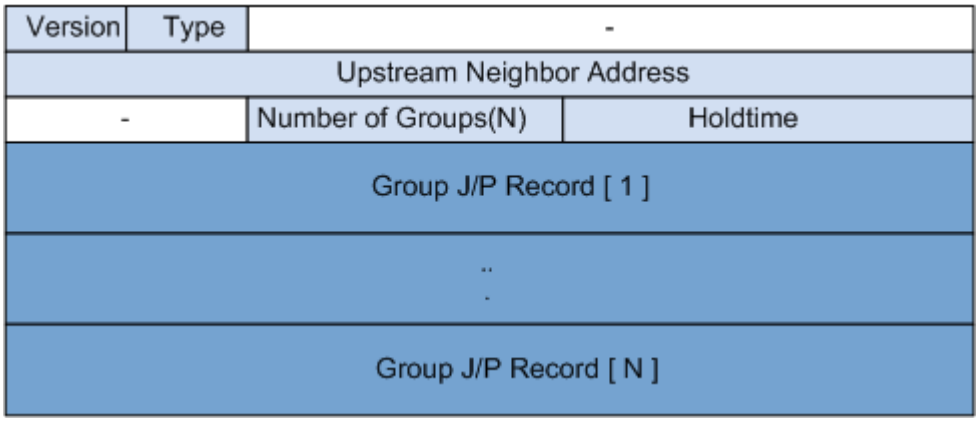
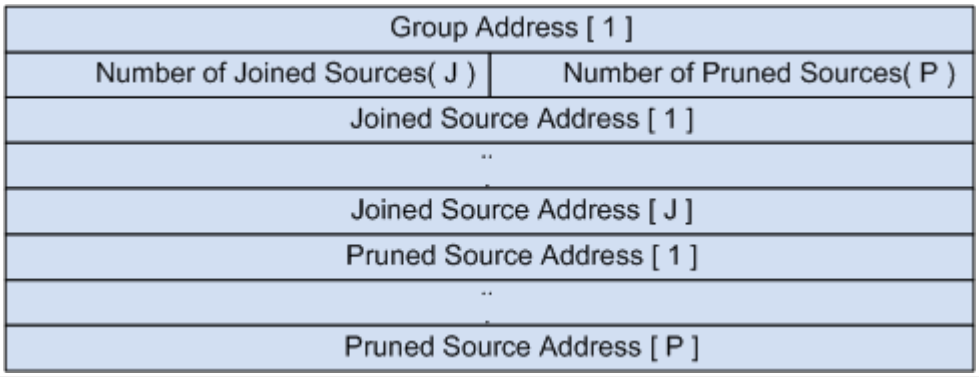


图 2 Group J/P Record 字段格式

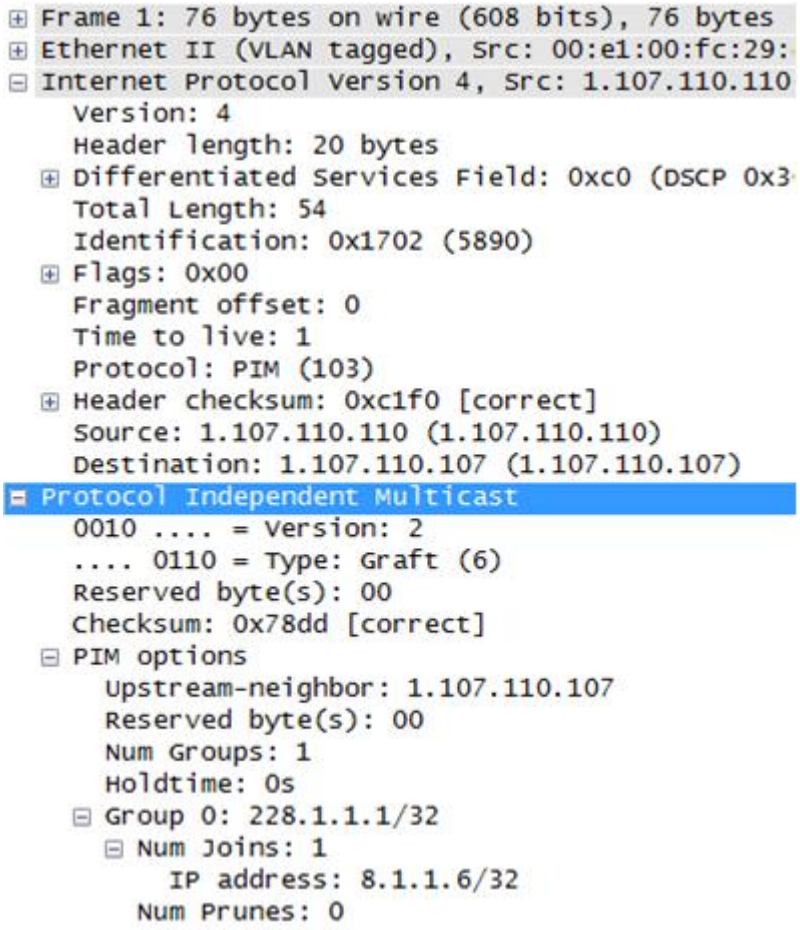


字段	长度	说明
Version	4 比特	PIM 版本，值为 2。
Type	4 比特	消息类型，Graft 值为 6，Graft-Ack 值为 7。
Upstream Neighbor Address	32 比特	在 Graft 消息中填上游邻居地址。也就是收到嫁接消息的路由器上，进行嫁接操作的下游接口地址。 在 Graft-Ack 消息中填为 Graft 消息的发出者地址。
Number of Groups	8 比特	消息中包含的组播组数目。
Holdtime	16 比特	该字段为 0。
Group Address	32 比特	组播组地址。
Number of Joined Sources	16 比特	针对该组播组，请求加入的组播源总数。

字段	长度	说明
Number of Pruned Sources	16 比特	该字段为 0。
Joined Source Address	32 比特	待嫁接的（S，G）的源地址。

报文示例

图 3 PIM Graft 消息



参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)
RFC 3973	Protocol Independent Multicast - Dense Mode （PIM-DM）:Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

4.13.7 PIM Bootstrap 消息格式

Bootstrap 消息

当 PIM-SM 网络中使用动态 RP 时，配置了 C-BSR 的路由器从所有 PIM 接口周期性的发送 Bootstrap 消息，参与 BSR 竞选。竞选获胜者，继续发送 Bootstrap 消息，向域内所有 PIM 路由器发布 RP-Set 信息。

Bootstrap 消息只在 PIM-SM 中使用。

封装 Bootstrap 消息的 IP 报文源地址为 C-BSR 地址，目的地址为 224.0.0.13，使用组播方式发送。TTL 为 1，在 PIM-SM 网络中逐跳转发，最终达到全网泛滥。

图 1 Bootstrap 消息格式

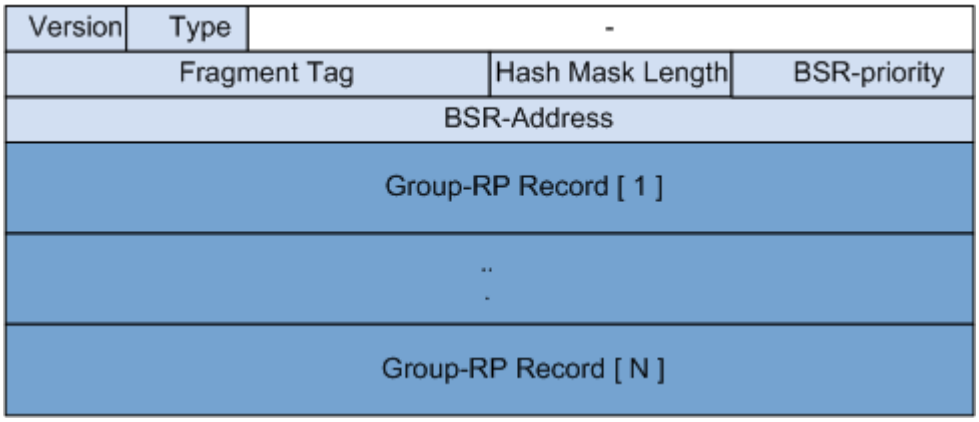
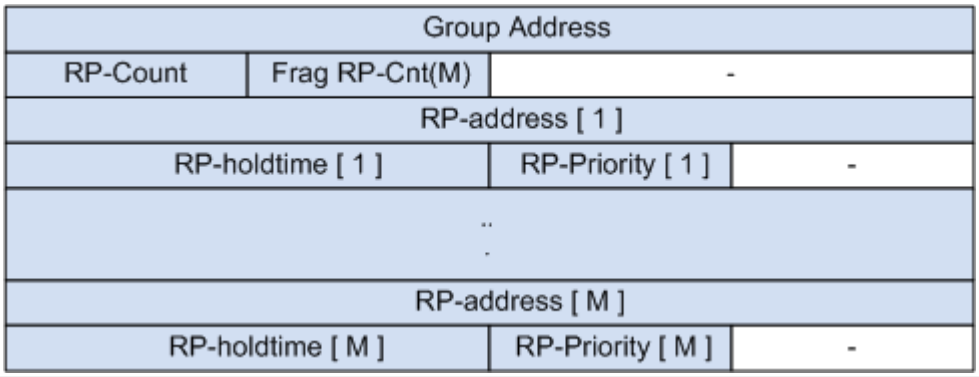


图 2 Group-RP Record 字段格式

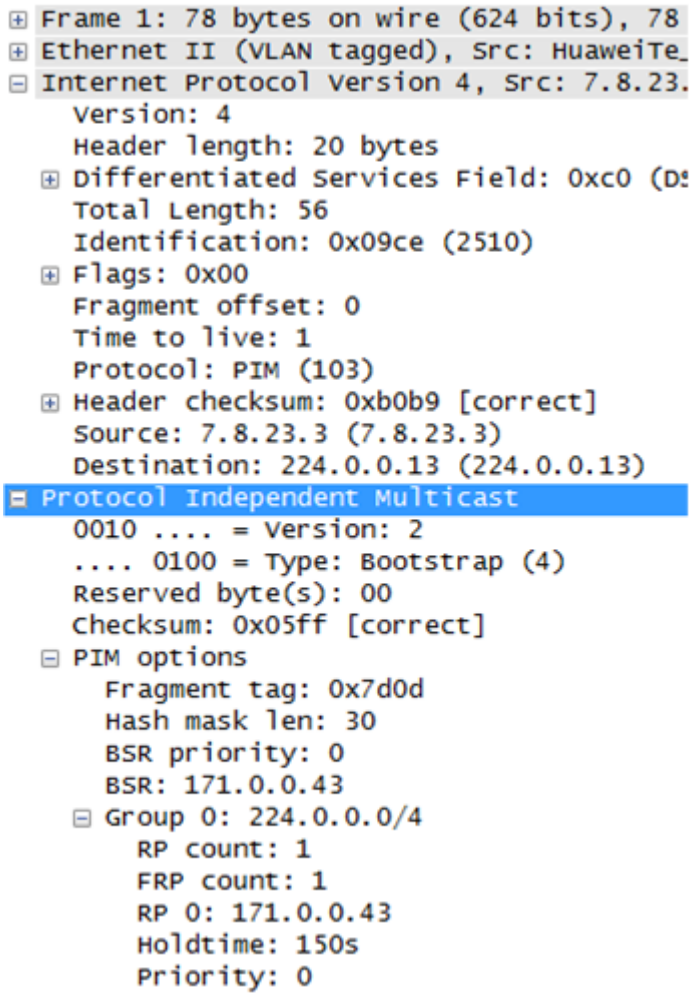


字段	长度	说明
Version	4 比特	PIM 版本，值为 2。
Type	4 比特	消息类型，值为 4。
Fragment Tag	16 比特	随机数，用来区分 Bootstrap 消息。
Hash Mask length	8 比特	C-BSR 的 Hash 掩码长度。

字段	长度	说明
BSR-priority	8 比特	C-BSR 优先级。
BSR-Address	32 比特	C-BSR 地址。
Group Address	32 比特	组播组地址。
RP-Count	8 比特	希望为该组服务的 C-RP 的总数。
Frag RP-Cnt	8 比特	在本段内包含的 C-RP 地址的个数。对于一个给定的组来说，如果 Bootstrap 消息分片，Frag RP-Cnt 字段便于将 RP-Set 分片。
RP-address	32 比特	C-RP 的地址。
RP-holdtime	16 比特	C-RP 发出的 advertisement 消息的老化时间，表示 C-RP 的有效时间。
RP-Priority	8 比特	C-RP 的优先级。

报文示例

图 3 PIM Bootstrap 消息



参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)
RFC 3973	Protocol Independent Multicast - Dense Mode (PIM-DM) :Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

4.13.8 PIM Assert 消息格式

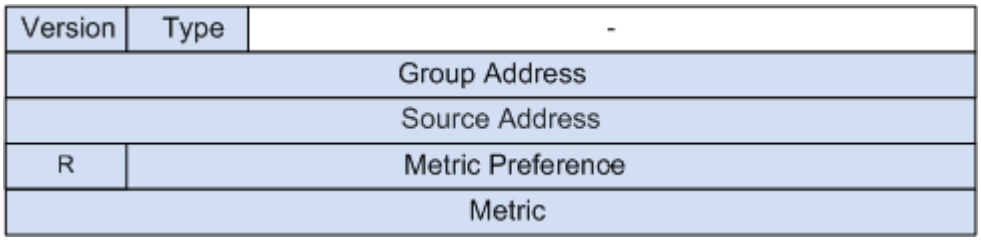
Assert 消息

在共享网段上，如果 PIM 路由器从（S，G）或（*，G）表项的下游接口收到（S，G）报文，则表示该网段存在其他的转发者。路由器从该下游接口发出 Assert 消息，参与竞选。竞选落败者停止下游接口的转发。

Assert 消息同时在 PIM-DM 和 PIM-SM 中使用。

封装 Assert 消息的 IP 报文源地址为本地接口地址，目的地址为 224.0.0.13，TTL 值为 1。使用组播方式发送。

图 1 Assert 消息格式



字段	长度	说明
Version	4 比特	PIM 版本，值为 2。
Type	4 比特	消息类型，值为 5。
Group Address	32 比特	组播组地址。
Source address	32 比特	如果竞选（S，G）表项的唯一转发者，则为组播源地址。如果竞选（*，G）表项的唯一转发者，则为 RP 地址。
R	4 比特	RPT 位。如果竞选（S，G）表项的唯一转发者，该位为 0；如果竞选（*，G）表项的唯一转发者，该位为 1。
Metric Preference	28 比特	到 Source address 的单播路径的优先级。
Metric	32 比特	到 Source address 的单播路由的开销。

报文示例

图 2 PIM Assert 消息

```
⊞ Frame 2: 68 bytes on wire (544 bits), 68 bytes
⊞ Ethernet II (VLAN tagged), Src: 00:ee:44:44:44:
⊟ Internet Protocol Version 4, Src: 130.130.130.2
    Version: 4
    Header length: 20 bytes
  ⊞ Differentiated Services Field: 0xc0 (DSCP 0x3
    Total Length: 46
    Identification: 0x6b85 (27525)
  ⊞ Flags: 0x00
    Fragment offset: 0
    Time to live: 1
    Protocol: PIM (103)
  ⊞ Header checksum: 0x6892 [correct]
    Source: 130.130.130.2 (130.130.130.2)
    Destination: 224.0.0.13 (224.0.0.13)
⊟ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 0101 = Type: Assert (5)
    Reserved byte(s): 00
    Checksum: 0xe10d [correct]
  ⊟ PIM options
      Group: 227.0.7.8/32
      Source: 13.1.0.200
      0... .... = RP Tree: False
      Metric Preference: 0
      Metric: 0
```

参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)
RFC 3973	Protocol Independent Multicast - Dense Mode (PIM-DM) :Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

4.13.9 PIM C-RP Advertisement 消息格式

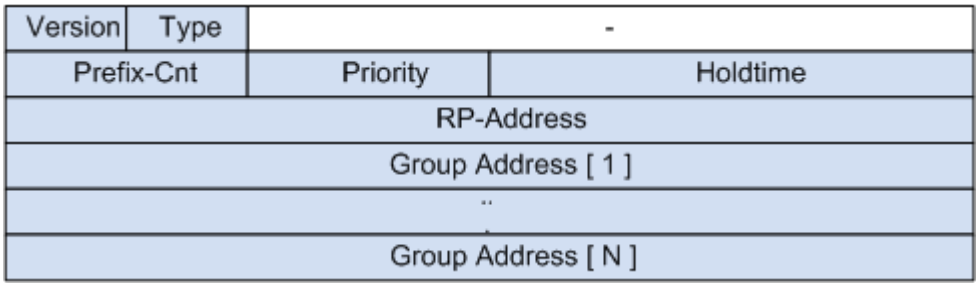
C-RP Advertisement 消息

当 PIM-SM 网络中使用动态 RP 时，配置了 C-RP 的路由器周期性的向 BSR 发送 Advertisement 消息，通告希望服务的组范围。

C-RP Advertisement 消息只在 PIM-SM 中使用。

封装 Advertisement 消息的 IP 报文源地址为源端 C-RP，目的地址为 BSR。使用单播方式发送。

图 1 Advertisement 消息格式



字段	长度	说明
Version	4 比特	PIM 版本，值为 2。
Type	4 比特	消息类型，值为 8。
Prefix-Cnt	8 比特	组播地址前缀值。
Priority	8 比特	C-RP 优先级。
Holdtime	16 比特	Advertisement 消息的老化时间。
RP-Address	32 比特	C-RP 地址。
Group Address	32 比特	组播组地址。

报文示例

图 2 PIM C-RP Advertisement 消息

```
⊞ Frame 1: 68 bytes on wire (544 bits), 68 bytes capt
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:04 (
⊟ Internet Protocol Version 4, Src: 171.0.0.42 (171.0
    Version: 4
    Header length: 20 bytes
  ⊞ Differentiated Services Field: 0xc0 (DSCP 0x30: C
    Total Length: 42
    Identification: 0x0acb (2763)
  ⊞ Flags: 0x00
    Fragment offset: 0
    Time to live: 255
    Protocol: PIM (103)
  ⊞ Header checksum: 0x598c [correct]
    Source: 171.0.0.42 (171.0.0.42)
    Destination: 171.0.0.43 (171.0.0.43)
⊟ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 1000 = Type: Candidate-RP-Advertisement (8)
    Reserved byte(s): 00
    Checksum: 0x493a [correct]
  ⊟ PIM options
      Prefix-count: 1
      Priority: 0
      Holdtime: 150s
      RP: 171.0.0.42
      Group 0: 224.0.0.0/4
```

图 3 PIM C-RP Advertisement 消息（注销）

```
⊞ Frame 1: 68 bytes on wire (544 bits), 68 bytes capt
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:04 (
⊟ Internet Protocol Version 4, Src: 171.0.0.42 (171.0
    Version: 4
    Header length: 20 bytes
  ⊞ Differentiated Services Field: 0xc0 (DSCP 0x30: C
    Total Length: 42
    Identification: 0x0c48 (3144)
  ⊞ Flags: 0x00
    Fragment offset: 0
    Time to live: 255
    Protocol: PIM (103)
  ⊞ Header checksum: 0x580f [correct]
    Source: 171.0.0.42 (171.0.0.42)
    Destination: 171.0.0.43 (171.0.0.43)
⊟ Protocol Independent Multicast
    0010 .... = Version: 2
    .... 1000 = Type: Candidate-RP-Advertisement (8)
    Reserved byte(s): 00
    Checksum: 0x49d0 [correct]
  ⊟ PIM options
      Prefix-count: 1
      Priority: 0
      Holdtime: 0s
      RP: 171.0.0.42
      Group 0: 224.0.0.0/4
```

参考标准

标准	描述
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM):Protocol Specification (Revised)

标准	描述
RFC 3973	Protocol Independent Multicast - Dense Mode （PIM-DM）:Protocol Specification (Revised)
RFC 4607	Source-Specific Multicast for IP

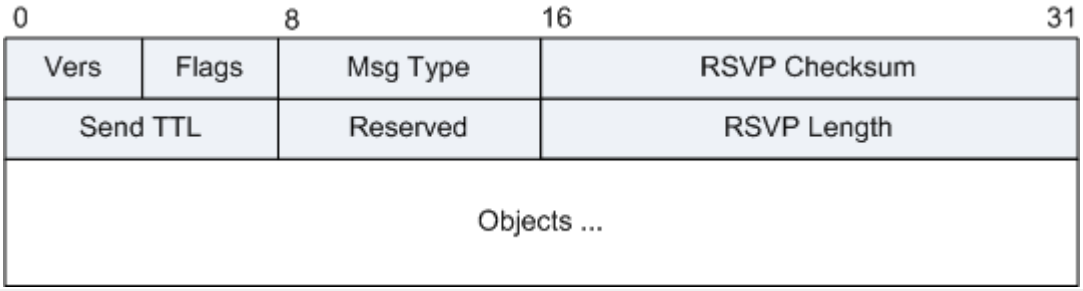
4.14 RSVP 报文格式

RSVP（Resource Reservation Protocol），资源预留协议，工作在传输层，但不参与应用数据的传送，是一种网络上的控制协议，类似于 ICMP。

报文格式

RSVP 各类消息都包含一个通用头部，随后是多个可变长度、类型的消息对象。

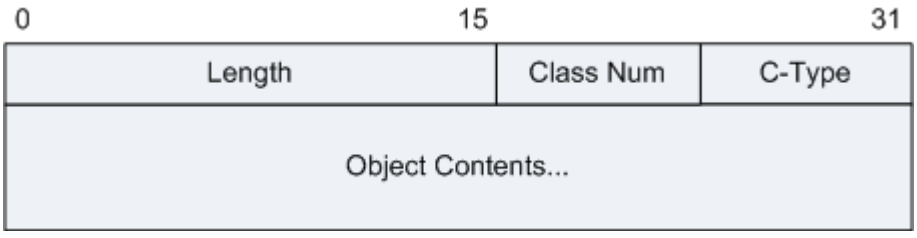
图 1 RSVP 消息格式



字段	长度	描述
Vers	4 比特	RSVP 版本号，当前版本为 1。
Flags	4 比特	标识位，一般值为 0。RFC2961 扩展其用来标识是否支持摘要刷新（Srefresh）。如果支持 Srefresh，则 Flags 置为 0x01。
Message Type	8 比特	表示消息的类别，下面列出一些类别： • 1－Path • 2－Resv • 3－PathErr • 4－ResvErr • 5－PathTear

字段	长度	描述
		• 6－ResvTear • 7－ResvConf • 8－DREQ • 9－DREP • 10－ResvTearConfirm • 11－Unassigned • 12－Bundle • 13－ACK • 14－Reserved • 15－Srefresh • 20－Hello • 21－Notify Message • 25－Integrity Challenge • 26－Integrity Response • 66－DSBM_willing • 67－I_AM_DSBM
RSVP Checksum	16 比特	表示 RSVP 的校验和，值为消息的补码的反码。如果值为 0，表示消息传输过程中不进行检验和检查。
Send TTL	8 比特	IPv4 的 TTL 的值，随报文一起发送。当节点接收到 RSVP 消息时，通过比较 Send_TTL 和 IP 首部的 TTL 值可以计算出该报文在非 RSVP 域中经过的跳数。
Reserved	8 比特	保留。
RSVP Length	16 比特	报文总长度，包括公共头及后面的 TLV objects，以字节为单位。
Object	变长	消息对象。每个 RSVP 消息都包含多个对象。不同类型的消息，包含的对象不同。

图 2 消息对象格式



字段	长度	描述

字段	长度	描述
Length	16 比特	示对象的总长度，以字节为单位。Length 必须是 4 的倍数，最小值为 4。
Class Num	8 比特	对象的类别： • 0：NULL，它的长度至少为 4，但是可以是任意 4 的倍数。NULL 对象可以出现在对象序列中的任意位置，它的内容会被接收者忽略掉。 • 1：SESSION，RSVP 会话相关信息，包括：Destination Address、Tunnel ID、Extend Tunnel ID。 • 2：unassigned • 3：RSVP_HOP，发送 Path 消息的上一跳的出接口地址和接口句柄。 • 4：INTEGRITY，携带了验证源节点的加密数据，它用于验证这个 RSVP 消息的内容。 • 5：TIME_VALUES，包含了消息创建者使用的刷新周期，在每个 Path 消息和 Resv 消息中必须存在。 • 6：ERROR_SPEC，指定了在 PathErr、ResvErr 消息中的 Error，或者在 ResvConf 消息中的确认。 ▪ Error Code = 00：确认 ▪ Error Code = 01：准入控制失败 ▪ Error Code = 02：策略控制失败 ▪ Error Code = 03：Resv 消息没有路径信息。 ▪ Error Code = 04：Resv 消息没有发送者信息 ▪ Error Code = 05：冲突的预留风格 ▪ Error Code = 06：不可识别的预留风格 ▪ Error Code = 07：冲突的目的端口 ▪ Error Code = 08:冲突的发送者端口 ▪ Error Code = 09，10，11：（保留） ▪ Error Code = 12:服务抢占 ▪ Error Code = 13：不可识别的对象 Class ▪ Error Code = 14：不可识别的对象 C-Type ▪ Error Code = 15-19：（保留） ▪ Error Code = 20：为 API 保留 ▪ Error Code = 21：流量控制错误

字段	长度	描述
		Sub-code = 01：服务冲突 Sub-code = 02：不支持的服务 Sub-code = 03：不正确的流规范值 Sub-code = 04：不正确的 Tspec 值 Sub-code = 05：不正确的 Adspec 值 ▪ Error Code = 22：流量控制系统错误 ▪ Error Code = 23：RSVP 系统错误 ▪ Error Code = 24：路由错误 Sub-code = 01：错误的 ERO Sub-code = 02：错误的严格显式路径节点 Sub-code = 03：错误的松散显式路径节点 Sub-code = 04：错误的初始子对象 Sub-code = 05：没有到目的地址的可用路由 Sub-code = 06：不可接受的标签值 Sub-code = 07：RRO 包含了路由环路 Sub-code = 08：MPLS 正在协商，但是路径上存在不支持 MPLS 的路由器。 Sub-code = 09：MPLS 标签分配错误 Sub-code = 10：不支持的 L3PID ▪ Error Code = 25：错误通告 Sub-code = 01：超过 MTU 值的 RRO Sub-code = 02：RRO 通告 Sub-code = 03：隧道本地修复 • 7：SCOPE，携带了一个朝向信息转发方向的发送者主机的显式列表。可能出现在 Resv、ResvErr 或者 ResvTear 消息中。 • 8：STYLE，定义了预留风格和在 FLOWSPEC 或者 FILTER_SPEC 对象中没有提及的指定风格的信息。在每个 Resv 消息中必须存在。 • 9：FLOWSPEC，指明了数据流的 QoS 特征，存在于 Resv 消息中。 • 10：FILTER_SPEC，发送节点的 IP 地址和 LSP ID，存在于 Resv 消息中。 • 11：SENDER_TEMPLATE，指定了发送节点的 IP 地址和 LSP ID，存在于 Path 消息中。(将会成为 Resv 消息中的 FILTERSPEC 对象的内容。) • 12：SENDER_TSPEC，指明了数据流的流量特征。存在于 Path 消息中。(将会成为 Resv 消息中

字段	长度	描述
		的 FLOWSPEC 对象的内容。） • 13：ADSPEC，用于收集路径上的实际 QoS 相关参数，例如，路径带宽估计、最小路径时延、Path MTU。存在于 Path 消息中。 • 14：POLICY_DATA，携带了允许一个本地策略模块来决定一个相关的预留是否是可管理的许可的信息。可能出现在 Path、Resv、PathErr 或者 ResvErr 消息中。 • 15：RESV_CONFIRM，预留确认请求，携带了请求预留确认的节点的 IP 地址。 • 16：RSVP_LABEL 表示分配的标签。 • 19：LABEL_REQUEST 标识 LABEL_REQUEST 对象，只在 Path 消息中携带。 • 20：EXPLICIT_ROUTE ERO（Explicit Route Object）描述 LSP 经过的路径信息，可以为严格显式路径也可以是松散显式路径。Path 消息沿 ERO 指定的路径转发，不受 IGP 最短路径约束。 ERO 包含了包含了 C-Type= 1 IPv4 前缀 2 IPv6 前缀 32 AS 号 • 21：ROUTE_RECORD RRO（Record Route Object）记录了 Path 消息实际途经的 LSR 的列表。RRO 可用于收集实际的路径信息，发现路由环路，还可以被复制到下一条 Path 消息中以实现路由锁定。 RRO 包含了 C-Type= 1 IPv4 地址 2 IPv6 地址 3 标签 • 22：HELLO C_Type = 1：HELLO REQUEST 对象 C_Type = 2 ：HELLO ACK 对象 • 207：SESSION_ATTRIBUTE 指定了建立优先级、保持优先级、预留风格、亲和属性等属性。
C-Type	8 比特	对象类型，表示同一类对象中不同的类型。Class_Number 与 C-Type 唯一标识了一个对象。
Object	变长	对象内容，可变长度。

字段	长度	描述
Content		

报文示例

图 3 RSVP Hello 消息

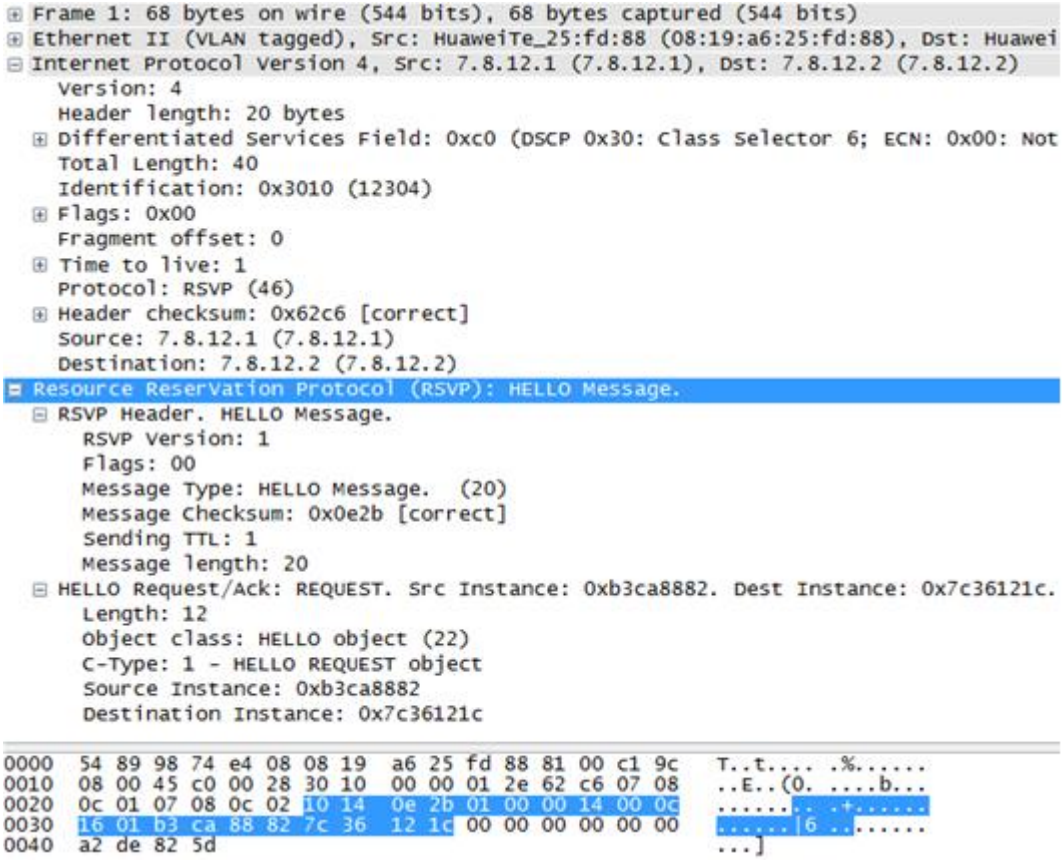


图 4 RSVP Path 消息

```
⊞ Frame 2: 242 bytes on wire (1936 bits), 242 bytes captured (1936 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_25:fd:88 (08:19:a6:25:fd:88), Dst:
⊞ Internet Protocol Version 4, Src: 171.0.0.41 (171.0.0.41), Dst: 171.0.0.43 (
⊟ Resource ReserVation Protocol (RSVP): PATH Message. SESSION: IPv4-LSP, Desti
  ⊟ RSVP Header. PATH Message.
      RSVP Version: 1
      Flags: 00
      Message Type: PATH Message.  (1)
      Message Checksum: 0xf210 [correct]
      Sending TTL: 254
      Message length: 196
  ⊟ SESSION: IPv4-LSP, Destination 171.0.0.43, Tunnel ID 1, Ext ID ab000029.
      Length: 16
      Object class: SESSION object (1)
      C-type: 7 - IPv4 LSP
      Destination address: 171.0.0.43 (171.0.0.43)
      Tunnel ID: 1
      Extended Tunnel ID: 2868903977 (171.0.0.41)
  ⊟ TIME VALUES: 30000 ms
      Length: 8
      Object class: TIME VALUES object (5)
      C-type: 1
      Refresh interval: 30000 ms (30 seconds)
  ⊟ SENDER TSPEC: IntServ, Token Bucket, 0 bytes/sec.
      Length: 36
      Object class: SENDER TSPEC object (12)
      C-type: 2 - Integrated Services
      Message format version: 0
      Data length: 7 words, not including header
      Service header: 1 - Traffic specification
      Length of service 1 data: 6 words, not including header
    ⊟ Token Bucket TSpec: Rate=0 Burst=1000 Peak=0 m=0 M=1500
        Parameter 127 - Token bucket
        Parameter 127 flags: 0x00
        Parameter 127 data length: 5 words, not including header
        Token bucket rate: 0
        Token bucket size: 1000
        Peak data rate: 0
        Minimum policed unit [m]: 0
        Maximum packet size [M]: 1500
  ⊟ SESSION ATTRIBUTE: SetupPrio 7, HoldPrio 7, SE Style,  [Tunnel0/0/1]
      Length: 20
      Object class: SESSION ATTRIBUTE object (207)
      C-type: 7 - IPv4 LSP (No Resource Affinities)
      Setup priority: 7
      Hold priority: 7
    ⊟ Flags: 0x04
        .... ...0 = Local protection not desired
        .... ..0. = Label recording not desired
        .... .1.. = SE style desired
        .... 0... = Bandwidth protection not desired
        ...0 .... = Node protection not desired
      Name length: 11
      Name: Tunnel0/0/1
  ⊟ ADSPEC
      Length: 48
      Object class: ADSPEC object (13)
      C-type: 2
      Message format version: 0
      Data length: 10 words, not including header
    ⊟ Default General Parameters
        Service header 1 - Default General Parameters
        Break bit not set
        Data length: 8 words, not including header
        IS Hop Count - 1 (type 4, length 1)
        Path b/w estimate - 1250000 (type 6, length 1)
        Minimum path latency - 0 (type 8, length 1)
        Composed MTU - 1500 (type 10, length 1)
    ⊟ Controlled Load
        Service header 5 - Controlled Load
        Break bit not set
        Data length: 0 words, not including header
  ⊟ LABEL REQUEST: Basic: L3PID: IP (0x0800)
      Length: 8
      Object class: LABEL REQUEST object (19)
      C-type: 1
      L3PID: IP (0x0800)
  ⊟ HOP: IPv4, 7.8.12.1
      Length: 12
      Object class: HOP object (3)
      C-type: 1 - IPv4
      Neighbor address: 7.8.12.1
```

图 5 RSVP Path Tear 消息

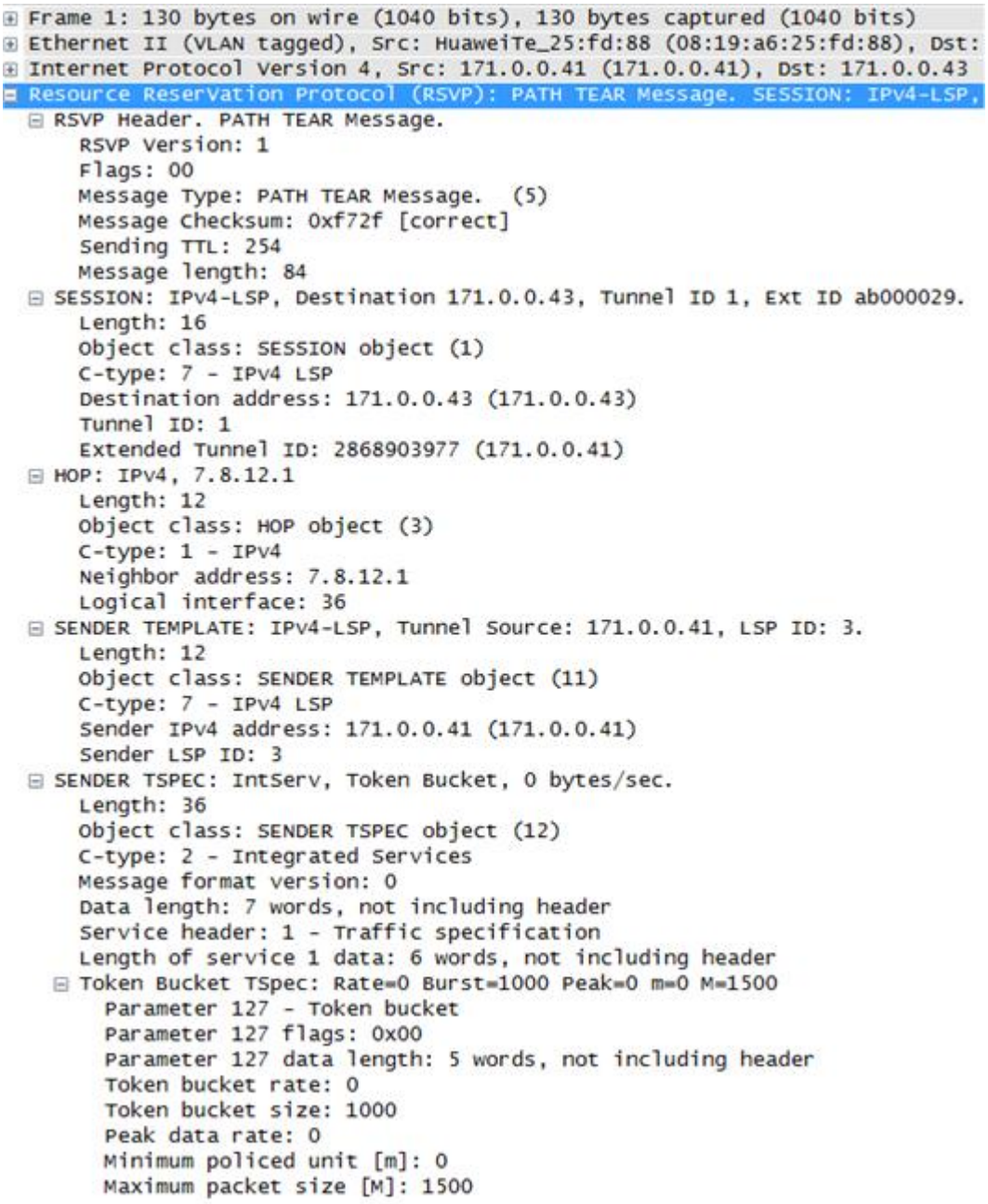


图 6 RSVP Resv 消息

```
⊞ Frame 2: 150 bytes on wire (1200 bits), 150 bytes captured (1200 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:08 (54:89:98:74:e4:08), Dst
⊞ Internet Protocol Version 4, Src: 7.8.12.2 (7.8.12.2), Dst: 7.8.12.1 (7.8.
⊟ Resource Reservation Protocol (RSVP): RESV Message. SESSION: IPv4-LSP, Des
  ⊟ RSVP Header. RESV Message.
      RSVP Version: 1
      Flags: 00
      Message Type: RESV Message.  (2)
      Message Checksum: 0x5f47 [correct]
      Sending TTL: 255
      Message length: 108
  ⊟ SESSION: IPv4-LSP, Destination 171.0.0.43, Tunnel ID 1, Ext ID ab000029.
      Length: 16
      Object class: SESSION object (1)
      C-type: 7 - IPv4 LSP
      Destination address: 171.0.0.43 (171.0.0.43)
      Tunnel ID: 1
      Extended Tunnel ID: 2868903977 (171.0.0.41)
  ⊟ HOP: IPv4, 7.8.12.2
      Length: 12
      Object class: HOP object (3)
      C-type: 1 - IPv4
      Neighbor address: 7.8.12.2
      Logical interface: 36
  ⊟ TIME VALUES: 30000 ms
      Length: 8
      Object class: TIME VALUES object (5)
      C-type: 1
      Refresh interval: 30000 ms (30 seconds)
  ⊟ STYLE: Shared-Explicit (18)
      Length: 8
      Object class: STYLE object (8)
      C-type: 1
      Flags: 0x00
      Style: 0x000012 - Shared-Explicit
  ⊟ FLOWSPEC: Controlled Load: Token Bucket, 0 bytes/sec.
      Length: 36
      Object class: FLOWSPEC object (9)
      C-type: 2
      Message format version: 0
      Data length: 7 words, not including header
      Service header: 5 - Controlled Load
      Length of service 5 data: 6 words, not including header
    ⊟ Token Bucket: Rate=0 Burst=1000 Peak=0 m=0 M=1500
        Parameter 127 - Token bucket
        Parameter 127 flags: 0x00
        Parameter 127 data length: 5 words, not including header
        Token bucket rate: 0
        Token bucket size: 1000
        Peak data rate: 0
        Minimum policed unit [m]: 0
        Maximum packet size [M]: 1500
  ⊟ FILTERSPEC: IPv4-LSP, Tunnel Source: 171.0.0.41, LSP ID: 4.
      Length: 12
      Object class: FILTER SPEC object (10)
      C-type: 7 - IPv4 LSP
      Sender IPv4 address: 171.0.0.41 (171.0.0.41)
      Sender LSP ID: 4
  ⊟ LABEL: 1140
      Length: 8
      Object class: LABEL object (16)
      C-type: 1 (Packet Label)
      Label: 1140
```

图 7 RSVP Resv-Tear 消息

```
⊞ Frame 1: 134 bytes on wire (1072 bits), 134 bytes captured (1072 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_74:e4:08 (54:89:98:74:e4:08), Dst
⊞ Internet Protocol Version 4, Src: 7.8.12.2 (7.8.12.2), Dst: 7.8.12.1 (7.8.
⊟ Resource ReserVation Protocol (RSVP): RESV TEAR Message. SESSION: IPv4-LSP
  ⊟ RSVP Header. RESV TEAR Message.
      RSVP Version: 1
      Flags: 00
      Message Type: RESV TEAR Message.  (6)
      Message Checksum: 0xee0a [correct]
      Sending TTL: 255
      Message length: 92
  ⊟ SESSION: IPv4-LSP, Destination 171.0.0.43, Tunnel ID 1, Ext ID ab000029.
      Length: 16
      Object class: SESSION object (1)
      C-type: 7 - IPv4 LSP
      Destination address: 171.0.0.43 (171.0.0.43)
      Tunnel ID: 1
      Extended Tunnel ID: 2868903977 (171.0.0.41)
  ⊟ HOP: IPv4, 7.8.12.2
      Length: 12
      Object class: HOP object (3)
      C-type: 1 - IPv4
      Neighbor address: 7.8.12.2
      Logical interface: 36
  ⊟ STYLE: Shared-Explicit (18)
      Length: 8
      Object class: STYLE object (8)
      C-type: 1
      Flags: 0x00
      Style: 0x000012 - Shared-Explicit
  ⊟ FLOWSPEC: Controlled Load: Token Bucket, 0 bytes/sec.
      Length: 36
      Object class: FLOWSPEC object (9)
      C-type: 2
      Message format version: 0
      Data length: 7 words, not including header
      Service header: 5 - Controlled Load
      Length of service 5 data: 6 words, not including header
    ⊞ Token Bucket: Rate=0 Burst=1000 Peak=0 m=0 M=1500
  ⊟ FILTERSPEC: IPv4-LSP, Tunnel Source: 171.0.0.41, LSP ID: 3.
      Length: 12
      Object class: FILTER SPEC object (10)
      C-type: 7 - IPv4 LSP
      Sender IPv4 address: 171.0.0.41 (171.0.0.41)
      Sender LSP ID: 3
```

参考标准

标准	描述
RFC 2205	Resource ReSerVation Protocol
RFC 2209	Resource ReSerVation Protocol (RSVP) -- Version 1 Message Processing Rules
RFC 3209	RSVP-TE: Extensions to RSVP for LSP Tunnels

4.15 VRRP 报文格式

报文格式

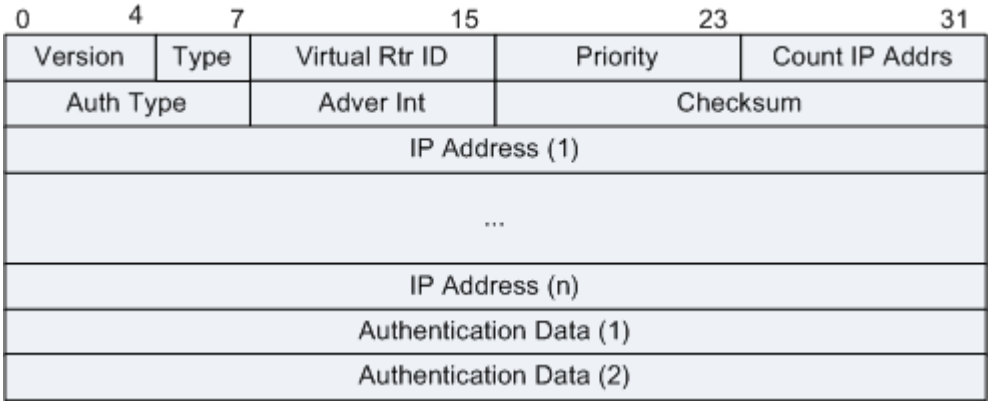
VRRP 报文被封装在 IP 包中。使用专门的 VRRP IPv4 组播地址。（协议号 112，组播地址 224.0.0.18）

IANA 分配给 VRRP 的 IP 协议号为 112（十进制）。

IANA 给 VRRP 分配的 IP 组播地址为 224.0.0.18。这是一个本地范围的多播地址。不论 TTL 的值是多少，路由器都被禁止转发以此地址为目标地址的报文。

VRRP 报文的 IP 头中，TTL 必须为 255。当 VRRP 路由器收到 TTL 不等于 255 的 VRRP 协议报文后，必须丢弃。

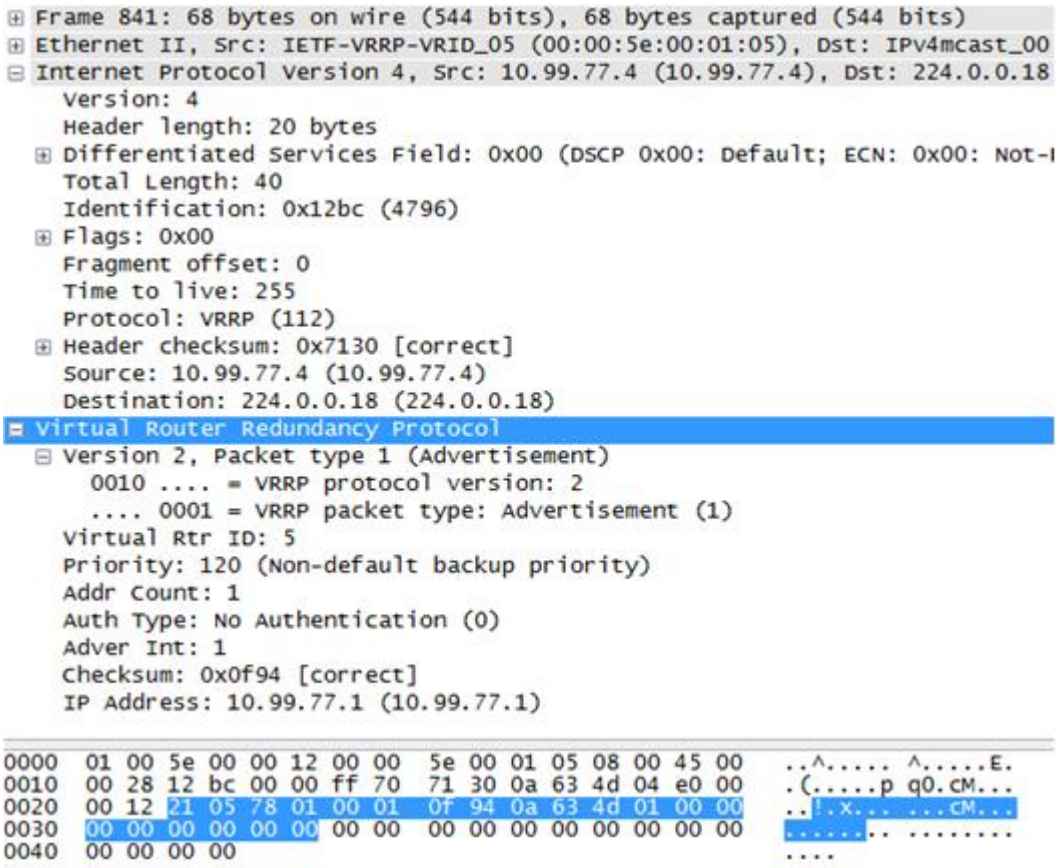
图 1 VRRP 报文格式



字段	长度	描述
Version	4 比特	指 VRRP 协议版本，本文档定义版本号 2。
Type	4 比特	定义了 VRRP 报文的类型。本版本的协议仅定义了一个报文类型： 1：ADVERTISEMENT 带有未知类型的报文必须被丢弃。
Virtual Rtr ID8	8 比特	虚拟路由器标识(VRID)字段标识了此报文所报告状态的虚拟路由器。可配置的范围是 1--255。没有缺省值。
Priority	8 比特	Priority 字段申明了发送此报文的 VRRP 路由器的优先级。值越高优先级越高。该字段为 8 位无符号整型。 如果 VRRP 路由器是虚拟路由器地址的 IP 地址所有者，那么其优先级必须为 255。起备用作用的 VRRP 路由器的优先级必须在 1--254 之间。缺省的 VRRP 路由器优先级为 100。 优先级值 0 用于指示当前虚拟路由器的主路由器停止参与 VRRP 组。主要用于触发备用路由器快速地迁移到主路由器，而不用等待当前主路由器超时。
Count IP Addrs	8 比特	在此 VRRP 通告中包含的 IP 地址的数量。
Auth Type	8 比特	认证类型字段用于标识要用到的认证方法。在一个虚拟路由器组内认证类型是唯一的。认证类型字段是一个 8 位无符号整型。如果报文携带未知的认证类型或者该认证类型和本地配置的认证方法不匹配，那么该报文必须被丢弃。 目前定义的认证方法有：

字段	长度	描述
		• 0 - No Authentication 不认证 该认证类型表明 VRRP 协议报文的交换不需要认证。在发送 VRRP 协议报文时，Authentication Data 字段将被置为 0；而在接收协议报文时，Authentication Data 字段被忽略。 • 1 - Reserved 保留 • 2 - Reserved 保留 **说明：** VRRP 的早期版本 定义了一些认证类型[RFC2338]。这些认证类型的定义已经在本文档中被删除，因为根据实际经验表明，这些认证方法并不能提供任何真正的安全保障，并且仅会导致在一个 VRRP 组内出现多个 Master 的情况。
Adver Int	8 比特	VRRP 通告间隔时间，单位为秒。缺省为 1 秒。这个字段主要用于错误配置路由器时的故障定位和解决。
Checksum	16 比特	校验和字段用于检测 VRRP 消息的数据是否出错。 校验和是从 version 字段开始的整个 VRRP 消息的 1 的 16 位补码和。(RFC1071 描述了校验和的计算细节)。
IP Address	32 比特	IP 地址字段为虚拟路由器的一个或者多个 IP 地址。IP 地址的数量在"Count IP Addrs"字段中说明。IP 地址字段用于错误配置路由器时的故障定位和解决。
Authentication Data	32 比特	认证字符串仅仅用于对 RFC2338 的向后兼容。在发送 VRRP 报文时该字段应该被置为 0，而在接收 VRRP 报文时该字段应该被忽略。

报文示例



参考标准

标准	描述
RFC 2338	Virtual Router Redundancy Protocol
RFC 2787	Definitions of Managed Objects for the Virtual Router Redundancy Protocol
RFC 3768	Virtual Router Redundancy Protocol (version number Two 2004)
RFC 5798	Virtual Router Redundancy Protocol Version 3 for IPv4 and IPv6

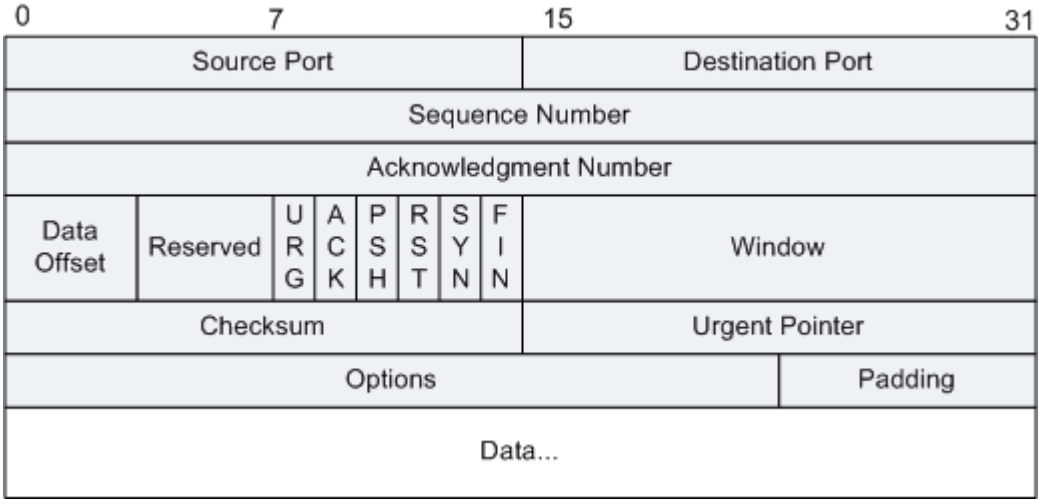
5. 传输层

- TCP 报文格式
- UDP 报文格式
- SCTP 报文格式

5.1 TCP 报文格式

报文格式

图 1 TCP 首部格式



字段	长度	含义
Source Port	16 比特	源端口，标识哪个应用程序发送。
Destination Port	16 比特	目的端口，标识哪个应用程序接收。
Sequence Number	32 比特	序号字段。TCP 链接中传输的数据流中每个字节都编上一个序号。序号字段的值指的是本报文段所发送的数据的第一个字节的序号。
Acknowledgment Number	32 比特	确认号，是期望收到对方的下一个报文段的数据的第 1 个字节的序号，即上次已成功接收到的数据字节序号加 1。只有 ACK 标识为 1，此字段有效。
Data Offset	4 比特	数据偏移,即首部长度,指出 TCP 报文段的数据起始处距离 TCP 报文段的起始处有多远，以 32 比特（4 字节）为计算单位。最多有 60 字节的首部，若无选项字段，正常为 20 字节。
Reserved	6 比特	保留，必须填 0。
URG	1 比特	紧急指针有效标识。它告诉系统此报文段中有紧急数据，应尽快传送（相当于高优先级的数据）。

字段	长度	含义
ACK	1 比特	确认序号有效标识。只有当 ACK=
PSH	1 比特	标识接收方应该尽快将这个报文段交给应用层。接收到 PSH = 1 的 TCP 报文段，应尽快的交付接收应用进程，而不再等待整个缓存都填满了后再向上交付。
RST	1 比特	重建连接标识。当 RST=
SYN	1 比特	同步序号标识，用来发起一个连接。SYN=
FIN	1 比特	发端完成发送任务标识。用来释放一个连接。FIN=
Window	16 比特	窗口：TCP 的流量控制，窗口起始于确认序号字段指明的值，这个值是接收端正期望接收的字节数。窗口最大为 65535 字节。
Checksum	16 比特	校验字段，包括 TCP 首部和 TCP 数据，是一个强制性的字段，一定是由发端计算和存储，并由收端进行验证。在计算检验和时，要在 TCP 报文段的前面加上 12 字节的伪首部。
Urgent Pointer	16 比特	紧急指针，只有当 URG 标志置 1 时紧急指针才有效。TCP 的紧急方式是发送端向另一端发送紧急数据的一种方式。紧急指针指出在本报文段中紧急数据共有多少个字节（紧急数据放在本报文段数据的最前面）。
Options	可变	选项字段。TCP 协议最初只规定了一种选项，即最长报文段长度（数据字段加上 TCP 首部），又称为 MSS。MSS 告诉对方 TCP“我的缓存所能接收的报文段的数据字段的最大长度是 MSS 个字节”。 新的 RFC 规定有以下几种选型：选项表结束，无操作，最大报文段长度，窗口扩大因子，时间戳。 • 窗口扩大因子：3 字节，其中一个字节表示偏移值 S。新的窗口值等于 TCP 首部中的窗口位数增大到（16+S），相当于把窗口值向左移动 S 位后获得实际的窗口大小。 • 时间戳：10 字节，其中最主要的字段是时间戳值（4 字节）和时间戳回送应答字段（4 字节）。

字段	长度	含义
		• 选项确认选项：
Padding	可变	填充字段，用来补位，使整个首部长度是 4 字节的整数倍。
data	可变	TCP 负载。

报文示例

图 2 TCP 报文(正常报文)

```
⊞ Frame 2: 1514 bytes on wire (12112 bits), 1514 bytes captured (
⊞ Ethernet II, Src: HughesNe_a9:4f:87 (00:80:ae:a9:4f:87), Dst: C
⊞ Internet Protocol Version 4, Src: 172.25.4.24 (172.25.4.24), Ds
⊟ Transmission Control Protocol, Src Port: icon-discover (2799),
    Source port: icon-discover (2799)
    Destination port: tht-treasure (1832)
    [Stream index: 0]
    Sequence number: 1    (relative sequence number)
    [Next sequence number: 1461    (relative sequence number)]
    Acknowledgement number: 1    (relative ack number)
    Header length: 20 bytes
  ⊟ Flags: 0x10 (ACK)
      000. .... .... = Reserved: Not set
      ...0 .... .... = Nonce: Not set
      .... 0... .... = Congestion Window Reduced (CWR): Not set
      .... .0.. .... = ECN-Echo: Not set
      .... ..0. .... = Urgent: Not set
      .... ...1 .... = Acknowledgement: Set
      .... .... 0... = Push: Not set
      .... .... .0.. = Reset: Not set
      .... .... ..0. = Syn: Not set
      .... .... ...0 = Fin: Not set
    Window size value: 65535
    [Calculated window size: 65535]
    [Window size scaling factor: -1 (unknown)]
  ⊟ Checksum: 0x2e11 [validation disabled]
      [Good Checksum: False]
      [Bad Checksum: False]
  ⊟ [SEQ/ACK analysis]
      [Bytes in flight: 1460]
⊟ Data (1460 bytes)
    Data: 5559db44bdee803e1f216f515e347d5e67095ec71aca0c30...
    [Length: 1460]
```

图 3 TCP 报文（Keepalive）

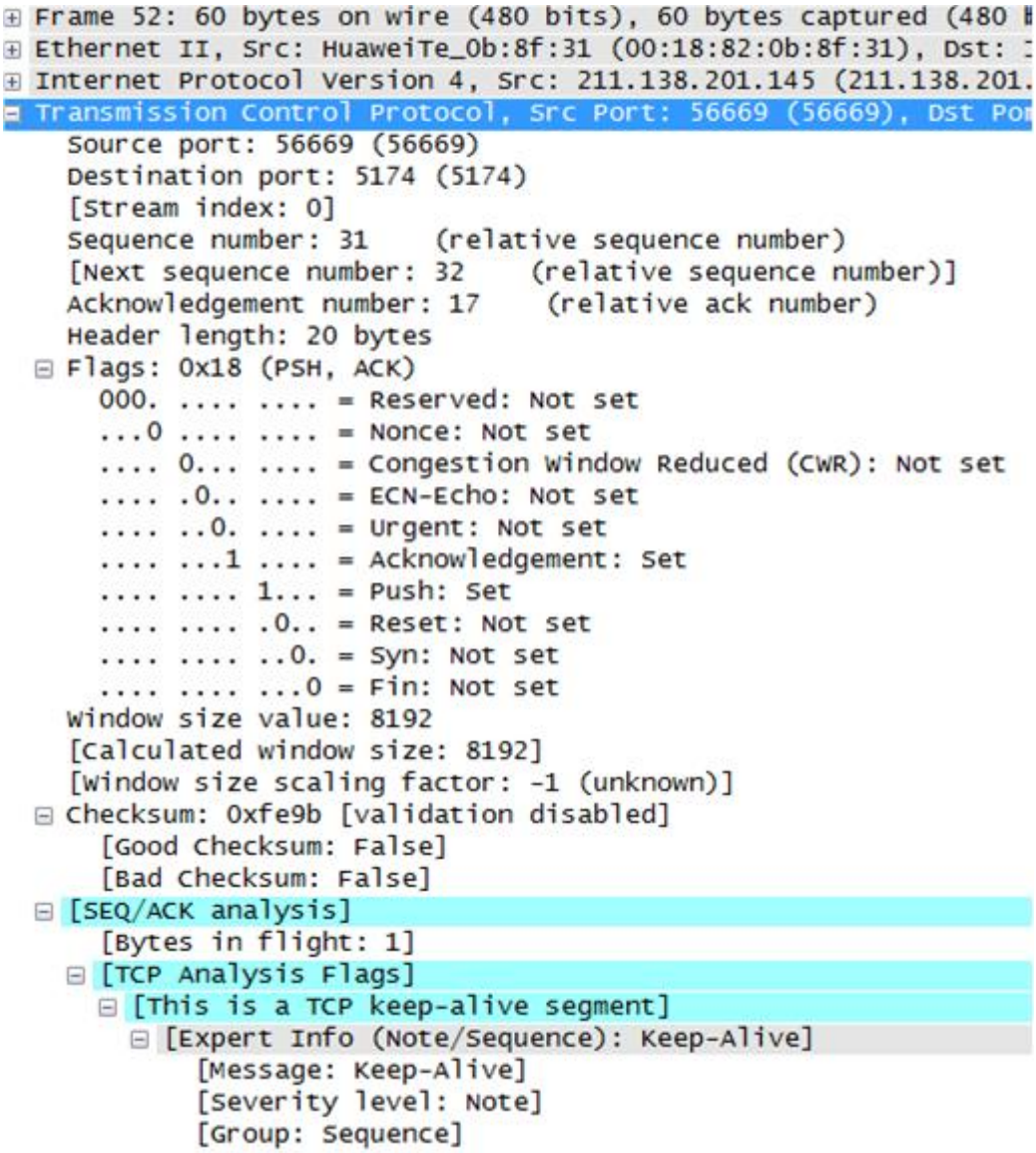


图 4 TCP 报文（Keepalive ACK）

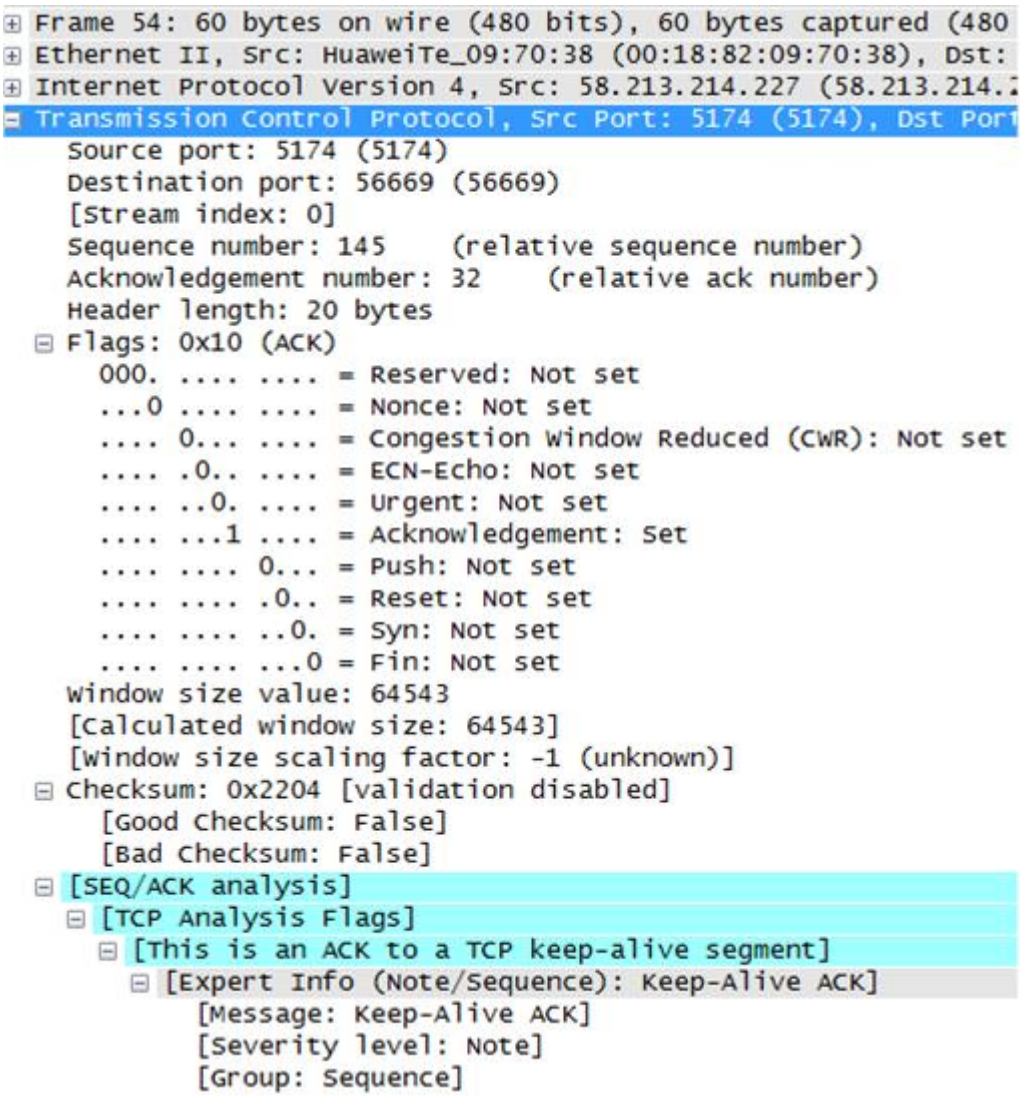


图 5 TCP 报文（Duplicate ACK）

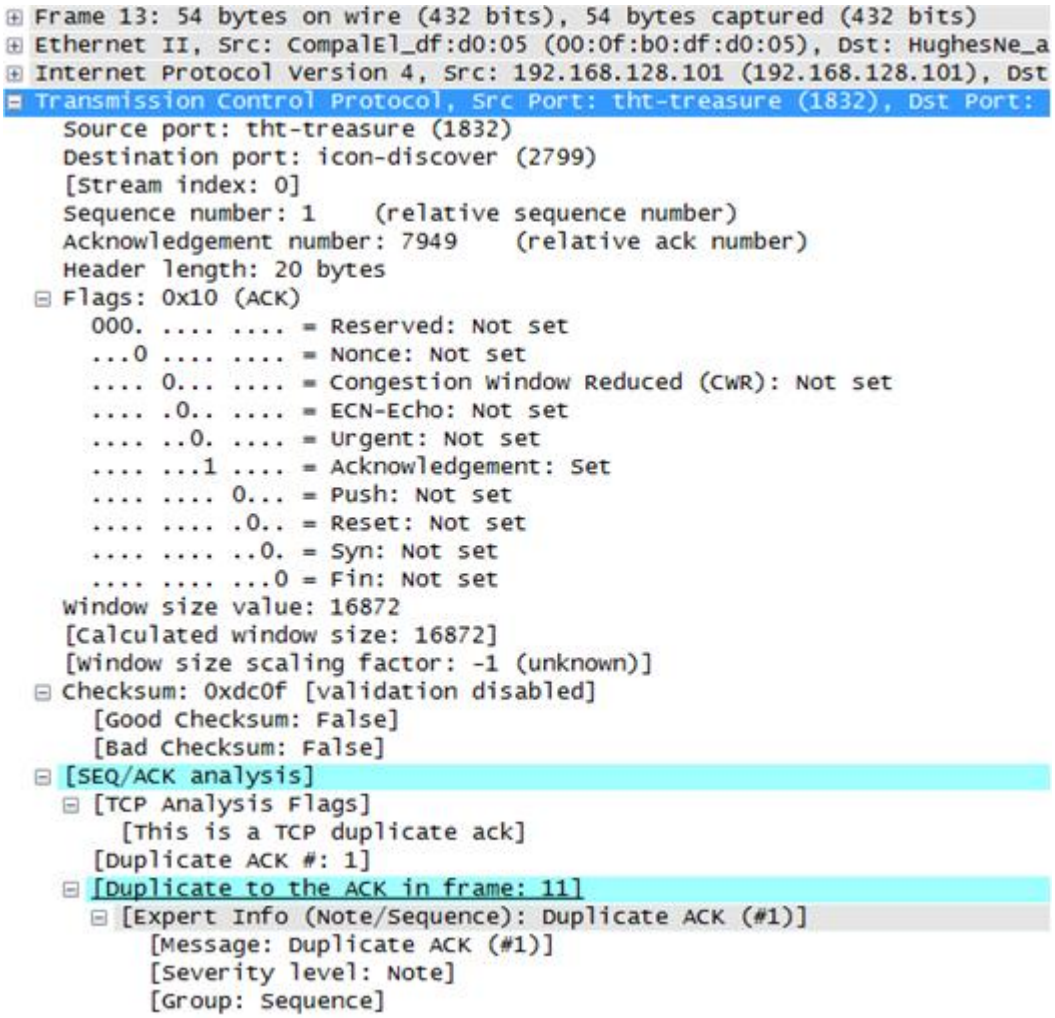

Frame 13: 54 bytes on wire (432 bits), 54 bytes captured (432 bits)
Ethernet II, Src: CompalEl_df:d0:05 (00:0f:b0:df:d0:05), Dst: HughesNe_a
Internet Protocol Version 4, Src: 192.168.128.101 (192.168.128.101), Dst
Transmission Control Protocol, Src Port: tht-treasure (1832), Dst Port:
Source port: tht-treasure (1832)
Destination port: icon-discover (2799)
[Stream index: 0]
Sequence number: 1 (relative sequence number)
Acknowledgement number: 7949 (relative ack number)
Header length: 20 bytes
Flags: 0x10 (ACK)
000. .... = Reserved: Not set
...0 .... = Nonce: Not set
.... 0... .... = Congestion Window Reduced (CWR): Not set
.... .0.. .... = ECN-Echo: Not set
.... ..0. .... = Urgent: Not set
....1 .... = Acknowledgement: Set
.... 0... = Push: Not set
.... .0.. = Reset: Not set
....0. = Syn: Not set
....0 = Fin: Not set
Window size value: 16872
[Calculated window size: 16872]
[Window size scaling factor: -1 (unknown)]
Checksum: 0xdc0f [validation disabled]
[Good Checksum: False]
[Bad Checksum: False]
[SEQ/ACK analysis]
[TCP Analysis Flags]
[This is a TCP duplicate ack]
[Duplicate ACK #: 1]
[Duplicate to the ACK in frame: 11]
[Expert Info (Note/Sequence): Duplicate ACK (#1)]
[Message: Duplicate ACK (#1)]
[Severity level: Note]
[Group: Sequence]

图 6 TCP 报文（重传）

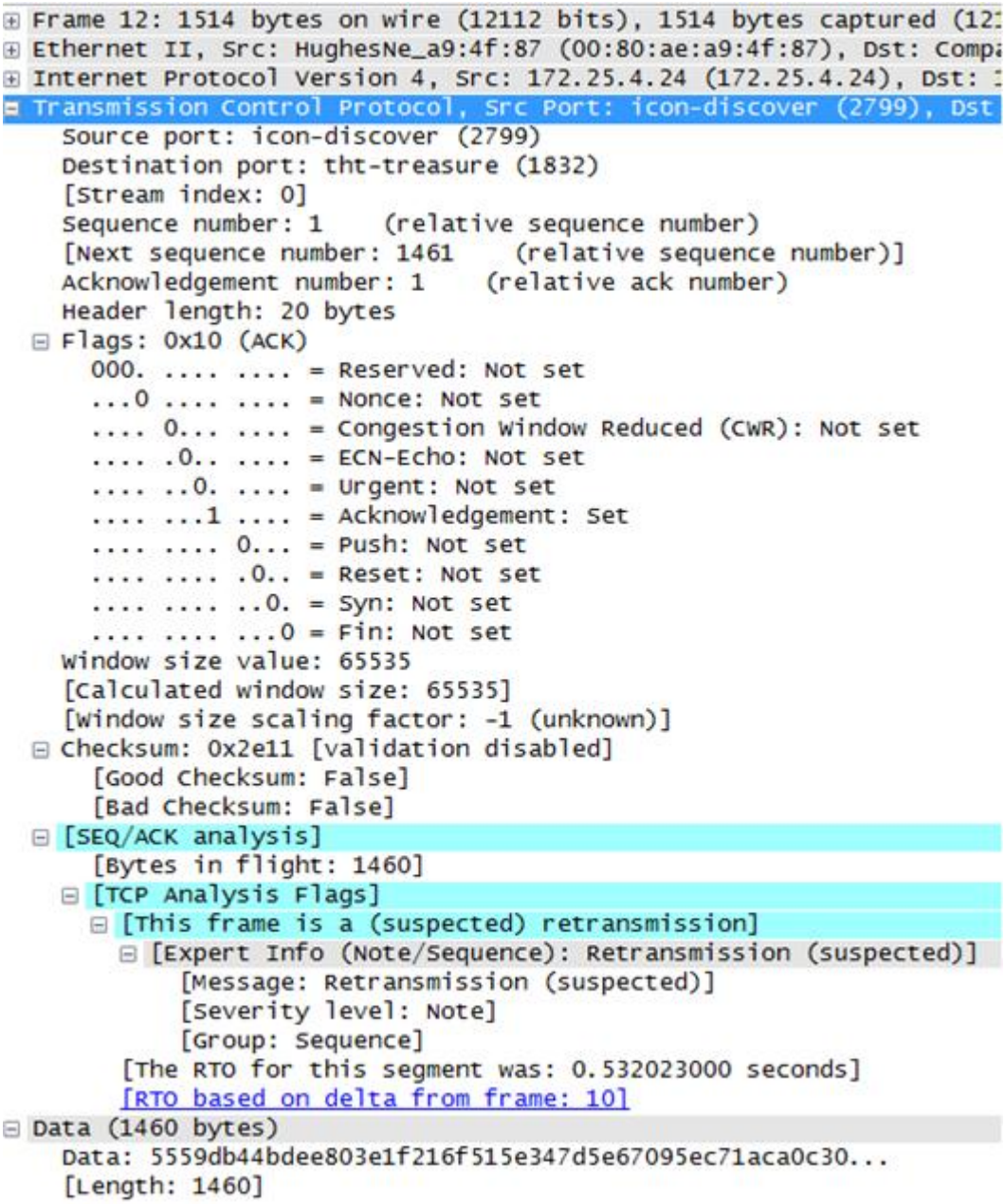


图 7 TCP 报文（Out-Of-Order 乱序）

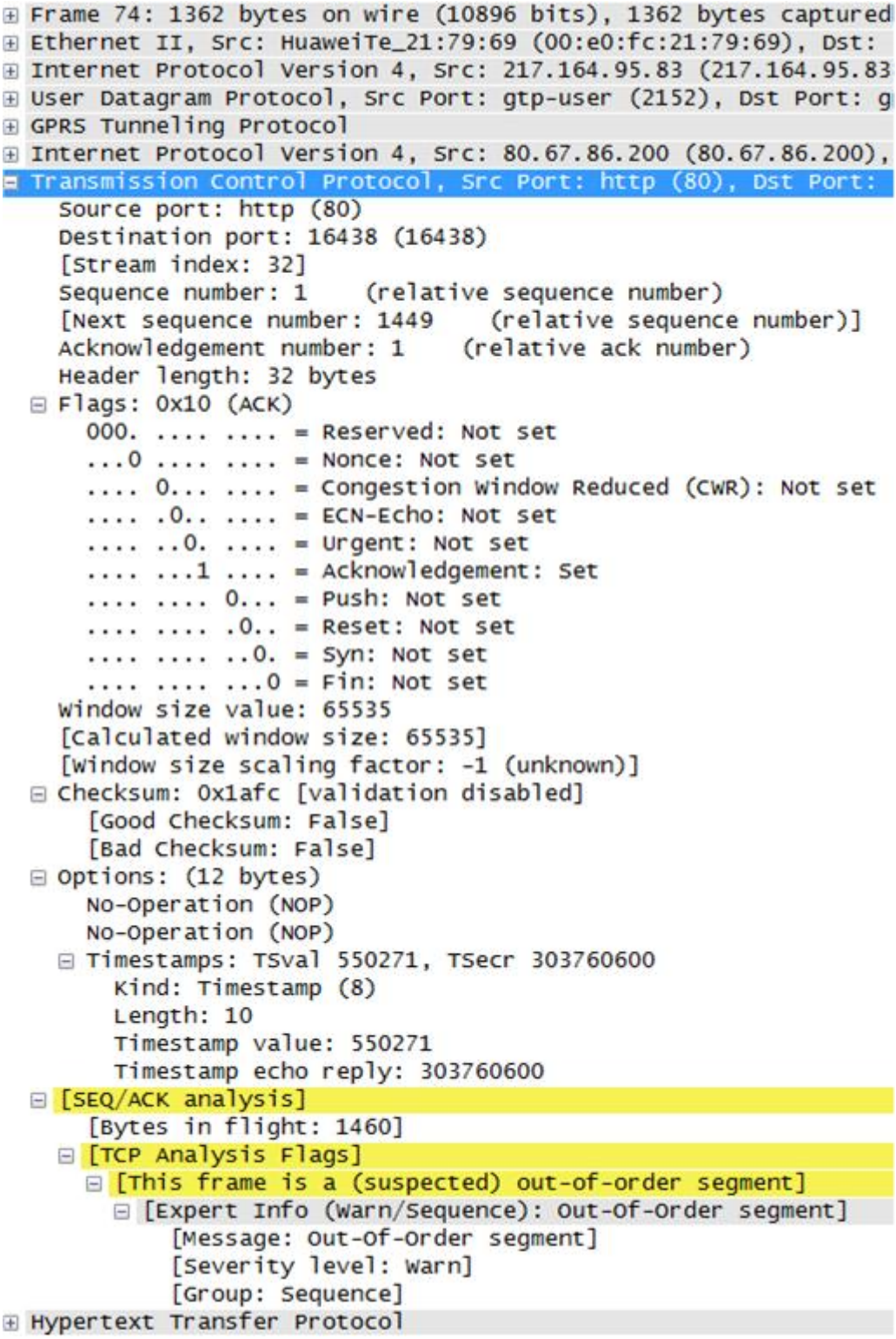


图 8 TCP 报文（Window Update）

```
⊞ Frame 57: 60 bytes on wire (480 bits), 60 bytes captured (480
⊞ Ethernet II, Src: HuaweiTe_0b:8f:31 (00:18:82:0b:8f:31), Dst:
⊞ Internet Protocol Version 4, Src: 211.138.201.145 (211.138.20
⊟ Transmission Control Protocol, Src Port: 56669 (56669), Dst P
    Source port: 56669 (56669)
    Destination port: 5174 (5174)
    [Stream index: 0]
    Sequence number: 32    (relative sequence number)
    Acknowledgement number: 145    (relative ack number)
    Header length: 20 bytes
  ⊟ Flags: 0x10 (ACK)
      000. .... .... = Reserved: Not set
      ...0 .... .... = Nonce: Not set
      .... 0... .... = Congestion Window Reduced (CWR): Not set
      .... .0.. .... = ECN-Echo: Not set
      .... ..0. .... = Urgent: Not set
      .... ...1 .... = Acknowledgement: Set
      .... .... 0... = Push: Not set
      .... .... .0.. = Reset: Not set
      .... .... ..0. = Syn: Not set
      .... .... ...0 = Fin: Not set
    Window size value: 8192
    [Calculated window size: 8192]
    [Window size scaling factor: -1 (unknown)]
  ⊟ Checksum: 0xfe23 [validation disabled]
      [Good Checksum: False]
      [Bad Checksum: False]
  ⊟ [SEQ/ACK analysis]
      [This is an ACK to the segment in frame: 56]
      [The RTT to ACK the segment was: 0.590543000 seconds]
    ⊟ [TCP Analysis Flags]
      ⊟ [This is a tcp window update]
        ⊟ [Expert Info (Chat/Sequence): Window update]
            [Message: Window update]
            [Severity level: Chat]
            [Group: Sequence]
```

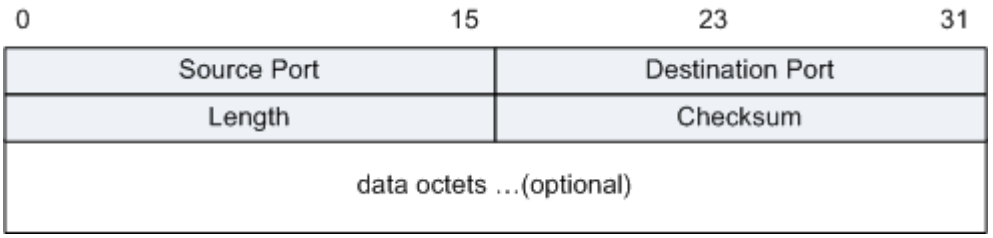
参考标准

标准	描述
RFC 793	Transmission Control Protocol

5.2 UDP 报文格式

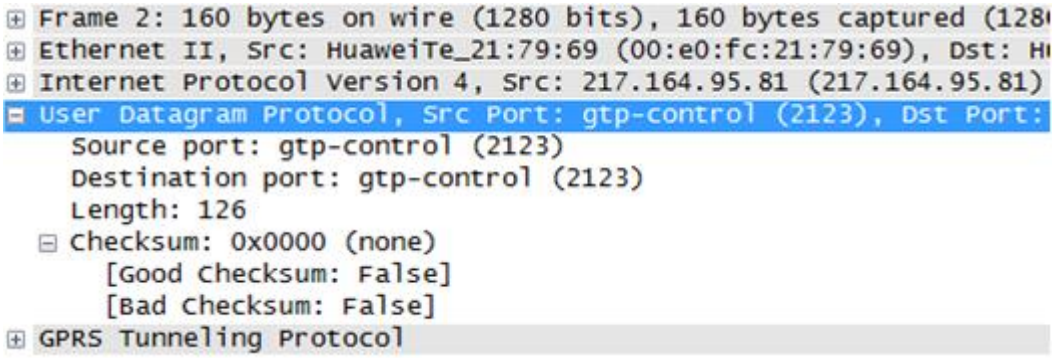
报文格式

图 1 UDP 报文格式



字段	长度	描述
Source Port	2 字节	标识哪个应用程序发送（发送进程）。
Destination Port	2 字节	标识哪个应用程序接收（接收进程）。
Length	2 字节	UDP 首部加上 UDP 数据的字节数，最小为 8。
Checksum	2 字节	覆盖 UDP 首部和 UDP 数据，是可选的。
data octets	变长	UDP 负载，可选的。

报文示例



参考标准

标准	描述
RFC 768	User Datagram Protocol

5.3 SCTP 报文格式

- SCTP 通用报文格式
- SCTP ABORT 报文格式
- SCTP COOKIE ACK 格式
- SCTP COOKIE ECHO 数据块格式
- SCTP DATA 数据块格式
- SCTP ERROR 数据块格式

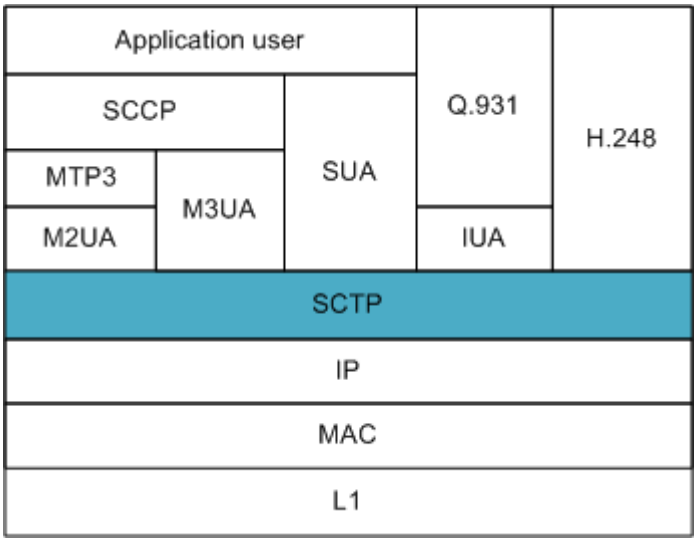
- SCTP HEARTBEAT 数据块格式
- SCTP HEARTBEAT ACK 数据块格式
- SCTP INIT 数据块格式
- SCTP INIT ACK 数据块格式
- SCTP SACK 数据块格式
- SCTP SHUTDOWN 消息格式
- SCTP SHUTDOWN ACK 数据块格式
- SCTP SHUTDOWN COMPLETE 数据块格式

5.3.1 SCTP 通用报文格式

SCTP（ Stream Control Transmission Protocol ），即流媒体控制传输协议，是一种可靠的基于无连接数据包网络如 IP 网络之上传输协议。他被设计用来在 IP 网络上传输 PSTN 在窄带信令消息，同时也能支持宽带信令消息的传输。

SCTP 可以看作 OSI 层次结构中的传输层，它的上层作为 SCTP 用户应用，下层为分组网络 IP 层。

图 1 SCTP 所处的协议栈结构

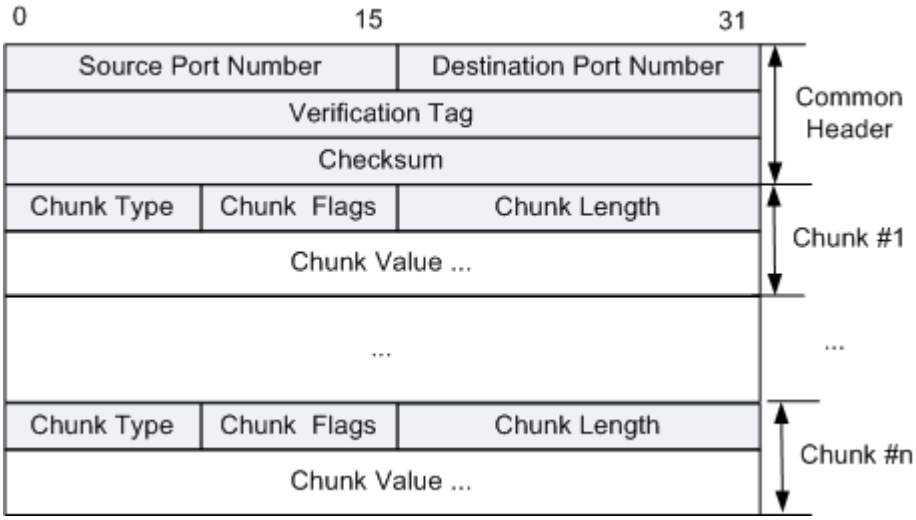


SCTP 报文格式

一个 SCTP 分组含了一个公共的分组头（Common Header）和若干数据块（Chunk），每个数据块中既可以包含控制信息，也可以包含用户数据。

除了 INIT、INIT ACK 和 SHUTDOWN COMPLETE 数据块外，其他类型的多个数据块可以捆绑在一个 SCTP 分组中，以满足对 MTU 大小的要求。当然，这些数据块也可以不与其他数据块捆绑在一个分组中。如果一个用户消息不能放在一个 SCTP 分组中，这个消息可以被分成若干个数据块。

图 2 SCTP 报文格式



字段	长度	描述
Source Port Number	16 比特的无符号整数	源端口号，识别 SCTP 发送端点的 SCTP 端口号。接收方可以使用源端口号、源 IP 地址、目的端口号和目的 IP 地址标识该 SCTP 分组所属的偶联。
Destination Port Number	16 比特的无符号整数	目的端口号，为目的端点的 SCTP 端口号。接收主机可以使用目的端口号将 SCTP 分组复用到正确的端点或应用中。
Verification Tag	32 比特的无符号整数	验证标签是偶联建立时，本端端点为这个偶联生成一个随机标识。偶联建立过程中，双方会交换这个 TAG，到了数据传递时，发送端必须在公共分组头中带上对端的这个 TAG，以备校验。 • 包含 INIT 数据块的分组中验证标签必须为 0。 • 在包含 SHUTDOWN-COMPLETE 数据块且设置了 T 比特的分组中，验证标签必须要从包含 SHUTDOWN-ACK 数据块的分组中复制。 • 在包含 ABORT 数据块的分组中，验证标签必须要从触发这个 ABORT 发送的分组中复制。
Checksum	32 比特的无符号整数	SCTP 通过对用户数据使用 ADLER-32 算法，计算出一个 32 位的校验码，带在数据报中，在接收端进行同样的运算，通过检查校验码是否相等来验证用户数据是否遭到破坏。
Chunk Type	8 比特的无符号整数	块类型定义在块值（Chunk Value）中消息所属的类型。包括：INIT、INIT ACK、SACK、ABORT、ERROR、SHUTDOWN、COOKIE ACK 等 13 种数据块类型。 该参数的取值范围为 0～254，255 留作今后的扩展。 数据块类型字段的编码分配如下： • 0：净荷数据(DATA) • 1：启动(INIT) • 2：启动证实 (INIT ACK) • 3：选择证实 (SACK)

字段	长度	描述
		• 4：Heartbeat 请求(HEARTBEAT) • 5：Heartbeat 证实(HEARTBEAT ACK) • 6：中止 (ABORT) • 7：关闭(SHUTDOWN) • 8：关闭证实(SHUTDOWN ACK) • 9：操作差错(ERROR) • 10：状态 Cookie(COOKIE ECHO) • 11：Cookie 证实(COOKIE ACK) • 12：为明确拥塞通知响应(ECNE)预留 • 13：为降低拥塞窗口(CWR)预留 • 14：关闭完成(SHUTDOWN COMPLETE) • 15～62：IETF 预留 • 63：IETF 定义的数据块扩展 • 64～126：IETF 预留 • 127：IETF 定义的数据块扩展 • 128～190：IETF 预留 • 191：IETF 定义的数据块扩展 • 192～254：IETF 预留 • 255：IETF 定义的数据块扩展 Chunk type 的高两位 bit 指示了收端不认识对应的 chunk type 的处理原则: • 00：停止处理数据报并丢弃，不再处理报中的其他 Chunk。 • 01：与 00 相同处理外，还要在 ERROR 或 INIT ACK 中上报，原因为不认识的参数类型。 • 10：忽略该 Chunk ，继续处理数据报中的其他 Chunk。 • 11：同 10 相同处理外，还要在 ERROR 中上报，原因为不认识的 Chunk 类型。
Chunk Flags	8 比特的无符号整数	块标志位用法由块类型决定。除非被置为其他值，块标记在传送过程中会被置 0 而且接收端点会忽视块标记。
Chunk Length	16 比特的无符号整数	块长度用来表示包括块类型、块标记、块长度和块值在内的字节数，长度使用二进制表示。
Chunk Value	变长	块值字段是在该数据块中真正传送的信息，内容由数据块类型决定。块值的长度为不定长。

SCTP 报文交互流程

图 3 Association 建立流程

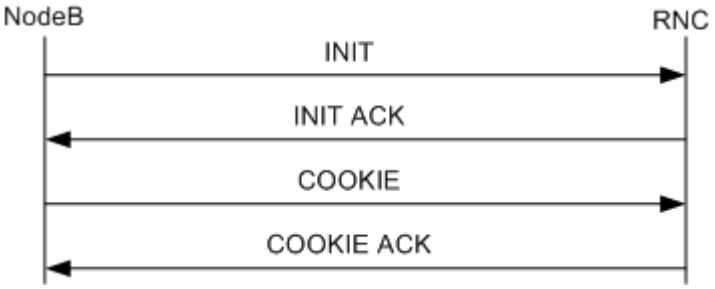


图 4 Association 关闭流程（Ungraceful）

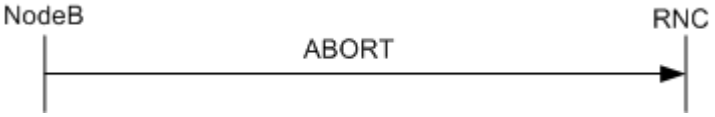


图 5 Association 关闭流程（Graceful）

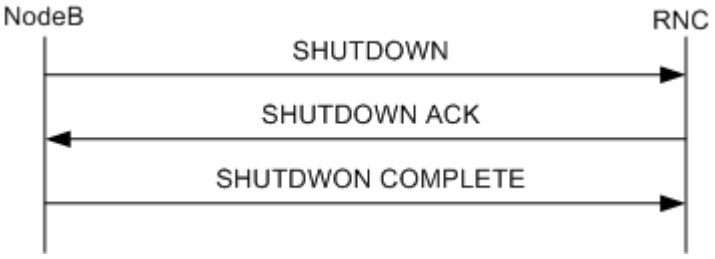
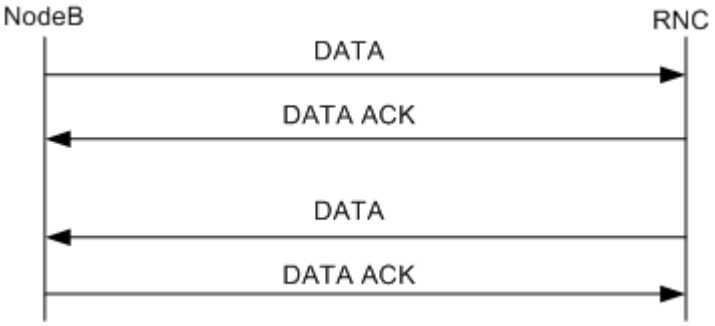


图 6 数据包的发送和确认流程



参考标准

标准	描述
RFC 2960	Stream Control Transmission Protocol

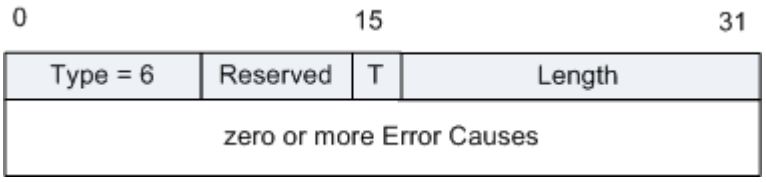
5.3.2 SCTP ABORT 报文格式

中止(ABORT)数据块的格式

SCTP 端点发送 ABORT 数据块来中止到对等端的偶联，ABORT 数据块中可以包含原因参数用来通知接收 ABORT 数据块的一方中止该偶联的原因。DATA 数据块不能同 ABORT 数据块捆绑在一个 SCTP 分组中。控制数据块(除 INIT、INIT ACK 和 SHUTDOWN COMPLETE)均可以同 ABORT 进行捆绑在一个 SCTP 分组中，但这些控制块都应放在 SCTP 分组中的 ABORT 数据块之前，否则这些控制数据块将被接收方忽略。

如果一个端点收到了格式错误或与不存在的偶联相关的 ABORT 消息，则应当舍弃该消息。此外，在任何情况下，端点收到一个 ABORT 消息后，都不能通过发送 ABORT 消息作为响应。

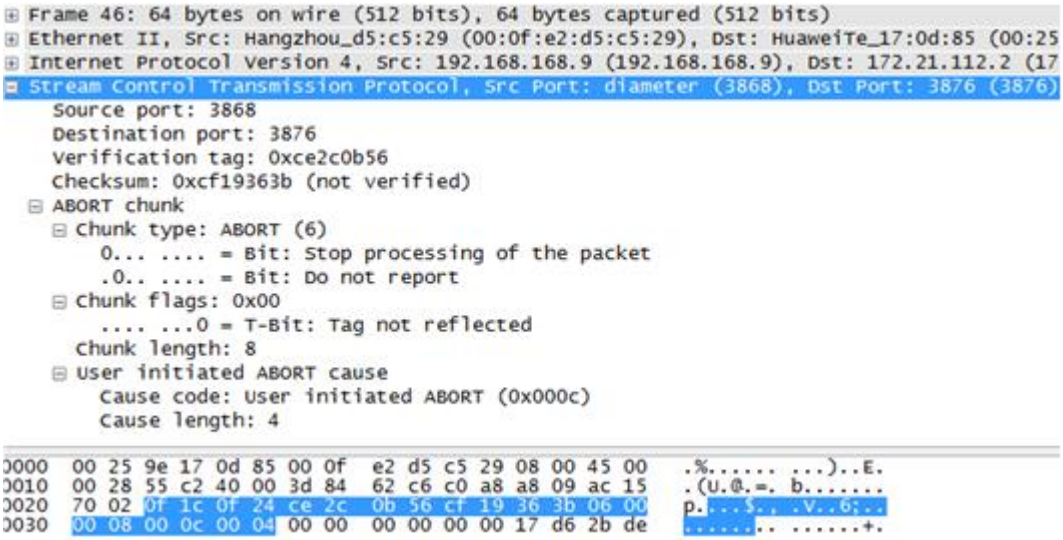
图 1 SCTP ABORT 报文格式



- Reserved：7 比特，在发送方设置为全 0，并在接收方忽略。
- T (1bit)：当发送方有一个 TCB 被破坏时则该 T 比特设置为 0，如果发送方没有 TCB，则把该比特设置为 1。
- Length (16bit 无符号整数)：设置为该数据块的长度，包括数据块头和所有包含的差错原因字段。

ABORT 数据块还可以包含 0 个或多个差错原因参数

报文示例



参考标准

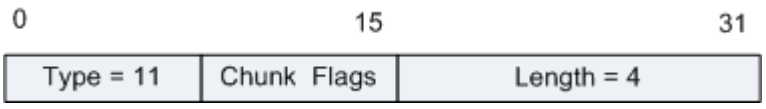
标准	描述
RFC 2960	Stream Control Transmission Protocol

5.3.3 SCTP COOKIE ACK 格式

COOKIE 证实(COOKIE ACK)数据块的格式

这个数据块只在启动偶联时使用，它用来证实收到 COOKIE EHCO 数据块。这个数据块必须在该偶联上发送任何 DATA 或 SACK 数据块前发送，但这个数据块可以与一个或多个 DATA 或 SACK 数据块捆绑在一个 SCTP 分组中发送。COOKIE ACK 数据块中没有其他参数。

图 1 SCTP COOKIE ACK 格式



COOKIE ACK 数据块中只包含数据块标志(8bit)，在发送方设置为全 0，并在接收方忽略。

报文示例

```
⊞ Frame 44: 64 bytes on wire (512 bits), 64 bytes captured (512 bits)
⊞ Ethernet II, Src: Hangzhou_d5:c5:29 (00:0f:e2:d5:c5:29), Dst: HuaweiTe_17:0d:85 (00:25
⊞ Internet Protocol Version 4, Src: 192.168.168.9 (192.168.168.9), Dst: 172.21.112.2 (17
⊟ Stream Control Transmission Protocol, Src Port: diameter (3868), Dst Port: 3876 (3876)
    Source port: 3868
    Destination port: 3876
    Verification tag: 0xce2c0b56
    Checksum: 0xc3520f0b (not verified)
  ⊟ COOKIE_ACK chunk
    ⊟ Chunk type: COOKIE_ACK (11)
        0... .... = Bit: Stop processing of the packet
        .0.. .... = Bit: Do not report
      Chunk flags: 0x00
      Chunk length: 4
```

```
0000  00 25 9e 17 0d 85 00 0f  e2 d5 c5 29 08 00 45 00   .%...... ...)..E.
0010  00 24 55 c0 40 00 3d 84  62 cc c0 a8 a8 09 ac 15   .$U.@.=. b.......
0020  70 02 0f 1c 0f 24 ce 2c  0b 56 c3 52 0f 0b 0b 00   p....$., .V.R....
0030  00 04 00 00 00 00 00 00  00 00 00 00 e5 ce 25 72   ........ ......%r
```

参考标准

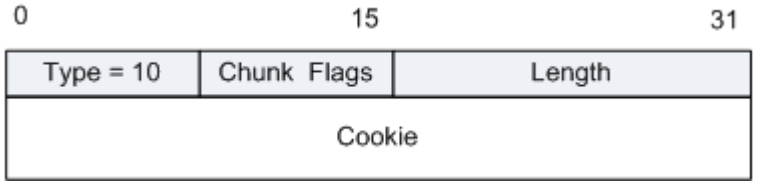
标准	描述
RFC 2960	Stream Control Transmission Protocol

5.3.4 SCTP COOKIE ECHO 数据块格式

状态 COOKIE(COOKIE ECHO)数据块的格式

该数据块只在启动偶联时使用，它由偶联的发起者发送到对端点，用来完成启动过程。这个数据块必须在该偶联上发送的 DATA 数据块前发送，但可以同其他的 DATA 数据块捆绑到同一个 SCTP 分组中。

图 1 SCTP COOKIE ECHO 数据块格式



- Chunk Flags(8bit)：在发送方设置为全 0，在接收方忽略。
- Length：(16bit 的无符号整数)：设置为该数据块长度的字节数，包括 4 字节的数据块头和 COOKIE 的长度。
- Cookie(可变长度)：该字段必须包含从前一个 INIT ACK 数据块的状态 COOKIE 参数中收到的准确的 COOKIE，使用 COOKIE 时应尽可能的小从而保证互操作性。

报文示例

```
⊞ Frame 43: 226 bytes on wire (1808 bits), 226 bytes captured (1808 bits)
⊞ Ethernet II, Src: Hangzhou_d5:c5:29 (00:0f:e2:d5:c5:29), Dst: Hangzhou_2f:d7:18 (00:0f
⊞ Internet Protocol Version 4, Src: 172.21.112.2 (172.21.112.2), Dst: 192.168.168.9 (192
⊟ Stream Control Transmission Protocol, Src Port: 3876 (3876), Dst Port: diameter (3868)
    Source port: 3876
    Destination port: 3868
    Verification tag: 0xa0f73ac7
    Checksum: 0x67052617 (not verified)
  ⊟ COOKIE_ECHO chunk (Cookie length: 172 bytes)
    ⊟ Chunk type: COOKIE_ECHO (10)
        0... .... = Bit: Stop processing of the packet
        .0.. .... = Bit: Do not report
      Chunk flags: 0x00
      Chunk length: 176
      Cookie: 4e4f4b4941205343545008003811000003c00000000000000...
```

```
0000  00 0f e2 2f d7 18 00 0f  e2 d5 c5 29 08 00 45 e0   .../.... ...)..E.
0010  00 d0 af 46 00 00 3f 84  45 ba ac 15 70 02 c0 a8   ...F..?. E...p...
0020  a8 09 0f 24 0f 1c a0 f7  3a c7 67 05 26 17 0a 00   ...$.... :.g.&...
0030  00 b0 4e 4f 4b 49 41 20  53 43 54 50 08 00 38 11   ..NOKIA  SCTP..8.
0040  00 00 3c 00 00 00 00 00  00 00 00 00 00 00 ce 2c   ..<..... .......,
0050  0b 56 a0 f7 3a c7 01 2c  4a 4e 20 83 0c 00 ac 15   .V..:.., JN .....
0060  70 02 00 00 00 00 00 00  00 00 00 00 00 00 05 00   p....... ........
0070  00 00 c0 a8 a8 09 00 00  00 00 00 00 00 00 00 00   ........ ........
0080  00 00 05 00 00 00 00 00  00 00 0f 24 0f 1c 01 00   ........ ...$....
0090  00 00 01 00 00 00 01 00  00 1c ce 2c 0b 56 00 07   ........ ...,.V..
00a0  6c 64 00 10 00 10 00 00  00 00 00 05 00 08 ac 15   ld...... ........
00b0  70 02 02 00 00 c8 a0 f7  3a c7 00 02 ee 00 00 02   p....... :.......
00c0  08 00 ce 24 35 1f 80 00  00 04 76 00 57 2d a2 f1   ...$5... ..v.W-..
00d0  e8 0f 1d 4b b3 d8 35 71  33 3c f3 5c 7c f7 db dd   ...K..5q 3<.\|...
00e0  6f 35                                             o5
```

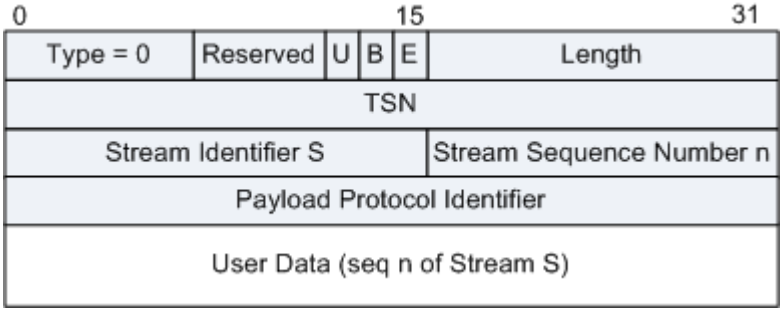
参考标准

标准	描述
RFC 2960	Stream Control Transmission Protocol

5.3.5 SCTP DATA 数据块格式

净荷数据(DATA)数据块的格式

图 1 SCTP DATA 数据块格式



字段	长度	描述
Reserved	5 比特	预留，应当设置为全 0，在接收方忽略。
U	1 比特	U 比特称为非顺序比特。如果该比特设置为 1，则指示这是一个非顺序的 DATA 数据块，不需要给该数据块分配流顺序号码，所有接收方必须忽略流顺序号码。在重新组装完成后(如果需要)，非顺序的数据块不需要尝试任何重新排序的过程，可以由接收方直接递交到高层；如果一个非顺序的用户消息被分段，

字段	长度	描述
		则消息的每个分段中的 U 比特必须被设置为 1。
B	1 比特	B 比特称为分段开始比特。如果该比特被设置，则指示这是用户消息的第一个分段。
E	1 比特	E 比特称为分段结束比特。如果该比特被设置，则指示这是用户消息的最后一个分段。一个未分段的用户消息应当把所有的 B 和 E 比特设置为 1。如果 B 和 E 比特都设置为 0，则表明这是一个分段的用户消息的一个中间分段。当用户消息被分段到多个数据块中，接收方需要使用 TSN 对消息进行重组，这意味着给分段的用户消息的每个分段都必须要使用连续的 TSN。 BE 比特的取值含义如下： • B=1, E= • B=0, E= • B=0, E= • B=1, E=
Length	16 比特无符号整数	该字段用来指示 DATA 数据块从类型字段开始到用户数据字段结束之间的字节数，但不包含任何填充字节，如果 DATA 数据块的用户数据字段为 0，则长度字段设为 16。
TSN	32 比特无符号整数	该值表示该数据块的 TSN，TSN 的有效值从 0～2 -1。TSN 的值达到 4294967295 后将回转到 0。
Stream Identifier S	16 比特无符号整数	该字段用来识别用户数据属于的流。
Stream Sequence Number n	16 比特无符号整数	该值用来表示所在流中的用户数据的顺序号码。该字段的有效值为 0～65535。当一个用户消息被 SCTP 分段后，则必须在消息的每个分段中都带有相同的流顺序号码。
Payload Protocol Identifier	32 比特无符号整数	该值表示一个应用(或上层协议)特定的协议标识符。这个值由高层协议传递到 SCTP 并发送到对等层。这个标识符不由 SCTP 使用，但却可以由特定的网络实体或对等的应用来识别在 DATA 数据块中携带的信息类型。甚至在每个分段的 DATA 数据块中也应包含该字段(以确保对网络中间的代理可用)。0 表示高层未对该协议净荷规定应用标识符。其中“M2UA”协议净荷使用编码 2；“M3UA”协议净荷使用编码 3；“SUA”协议净荷使用编码 4；“M2PA”协议净荷使用的编码待定。
User Data	变长	用来携带用户数据净荷。该字段必须被填充为 4 字节的整数情，发送方填充的字节数应不超过 3 个字节，接收方忽略所有的填充字节。

报文示例

```
Frame 272: 202 bytes on wire (1616 bits), 202 bytes captured (1616 bits)
Ethernet II, Src: HuaweiTe_88:76:66 (00:18:82:88:76:66), Dst: HuaweiTe_ce:5a:e9 (78:1d:ba:ce:5a:e9)
Internet Protocol Version 4, Src: 204.11.0.65 (204.11.0.65), Dst: 181.4.3.4 (181.4.3.4)
Stream Control Transmission Protocol, Src Port: s1-control (36412), Dst Port: s1-control (36412)
    Source port: 36412
    Destination port: 36412
    Verification tag: 0x1753b478
    Checksum: 0x013c95f0 (not verified)
  DATA chunk(ordered, complete segment, TSN: 3367083215, SID: 2, SSN: 5, PPID: 18, payload length: 137 bytes)
    Chunk type: DATA (0)
        0... .... = Bit: Stop processing of the packet
        .0.. .... = Bit: Do not report
    Chunk flags: 0x03
        .... ...1 = E-Bit: Last segment
        .... ..1. = B-Bit: First segment
        .... .0.. = U-Bit: Ordered delivery
        .... 0... = I-Bit: Possibly delay SACK
      Chunk length: 153
      TSN: 3367083215
      Stream Identifier: 0x0002
      Stream sequence number: 5
      Payload protocol identifier: S1 Application Protocol (S1AP) (18)
      Chunk padding: 000000
S1 Application Protocol

0000  78 1d ba ce 5a e9 00 18  82 88 76 66 08 00 45 e0   x...Z... ..vf..E.
0010  00 bc 30 89 00 00 40 84  c4 00 cc 0b 00 41 b5 04   ..0...@. .....A..
0020  03 04 8e 3c 8e 3c 17 53  b4 78 01 3c 95 f0 00 03   ...<.<.S .x.<....
0030  00 99 c8 b1 9c cf 00 02  00 05 00 00 00 12 00 09   ........ ........
0040  00 80 84 00 00 08 00 00  00 02 00 5e 00 08 00 03   ........ ...^....
```

参考标准

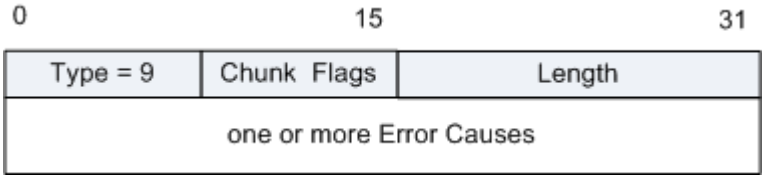
标准	描述
RFC 2960	Stream Control Transmission Protocol

5.3.6 SCTP ERROR 数据块格式

操作差错(ERROR)数据块的格式

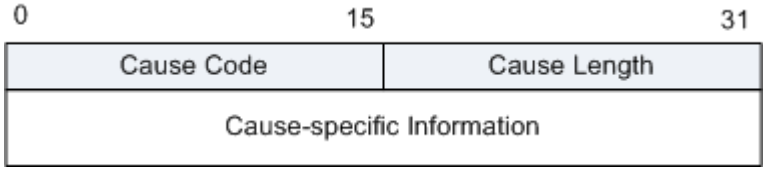
SCTP 端点发送该数据块向其对端点通知一些特定的差错情况。该数据块中可以包含一个或多个差错原因。一般操作差错不一定被看作是致命的，致命差错情况的报告一般使用 ABORT 数据块。

图 1 SCTP ERROR 数据块的格式



- 数据块标志位(8bit)：在发送方设置为全 0，并在接收方忽略。
- 可变长度(16bit 无符号整数)：设置为该数据块的字节数，包括数据块头和所有包含的差错原因字段的长度。
- 差错原因：可变长度的参数。

  图 2 SCTP ERROR 差错参数格式

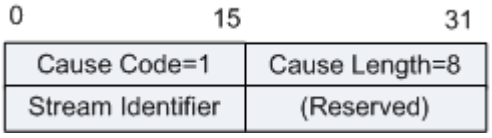
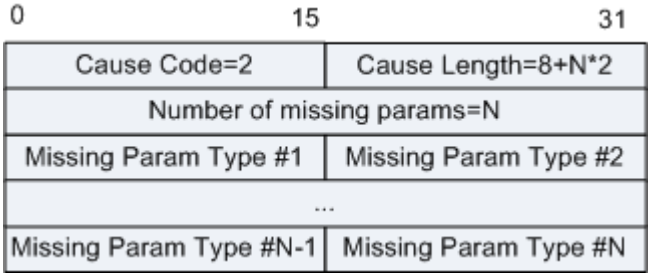
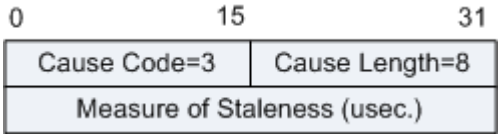
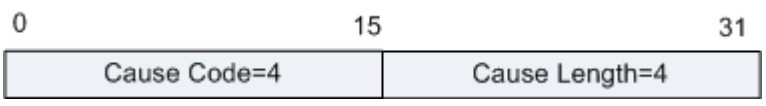
  - 原因编码 (16bit 无符号整数)：

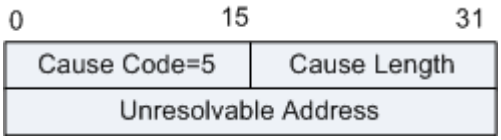
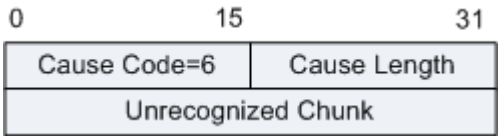
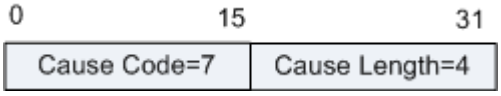
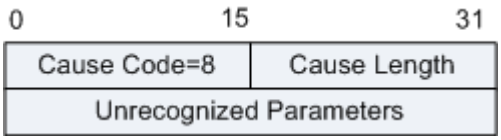
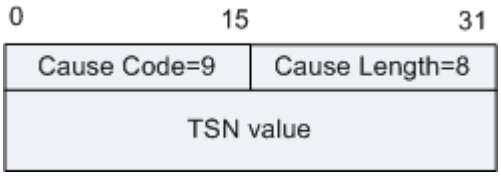
定义了被报告的差错情况的类型。

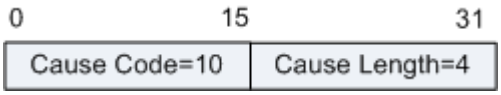
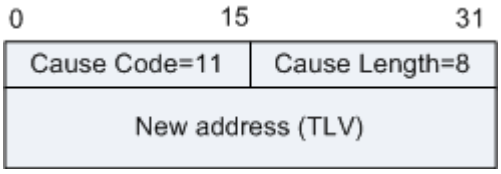
原因编码	描述
1	无效的流标识符
2	丢失必备参数
3	过期的 Cookie 差错
4	资源耗尽
5	无法解析的地址
6	不识别的数据块类型
7	无效的必备参数
8	不识别的参数
9	无用户数据
10	关闭阶段收到 COOKIE
11	使用新的地址重新启动偶联

- 原因长度(16bit 无符号整数)：设置为该参数的字节数，包括原因编码、原因长度和原因特定的信息字段。
- 原因特定的信息：可变长度，该字段用来携带差错的详细情况。

原因编码	含义	字段格式

原因编码	含义	字段格式
1	无效的流标识符	差错原因无效的流识别符用来指示端点收到了一个关于不存在的流的 DATA 数据块。  ▪ 流识别符：(16bit 无符号整数)：包含了接收购差错的 DATA 数据块的流标识符。 ▪ 备用字段(16bit)：由发送方设为全 0，在接收方忽略。
2	丢失必备参数	丢失必备参数差错原因用来指示一个或多个必备的参数在收到的 INIT 或 INIT ACK 数据  ▪ 丢失的参数个数(32bit 无符号整数)：该字段用来指示丢失的参数个数。 ▪ 丢失的参数类型(16bit 无符号整数)：每个字段都应包含丢失的必备参数号。
3	过期的 Cookie 差错	过期的 COOKIE 差错原因参数用来指示收到的有效的 State Cookie 已经过期了。  ▪ 过期测量(32bit 无符号整数)：该字段包含了当前时间和 State Cookie 过期时的 毫秒表示)。该差错原因的发送方可以通过在该字段中包含一个非 0 的值来报告 过期了多长时间。如果发送方不希望提供这个信息，则该字段设置为 0。
4	资源耗尽	资源耗尽差错原因用来指示发送方的资源已经耗尽,通常情况下该查错原因与 ABORT 数据 
5	不可解析的地址	U 不可解析的地址用来指示发送方不能解析特定的地址参数(即发送方不支持该类地址类 况下该查错原因与 ABORT 数据块一起发送。

原因编码	含义	字段格式
		 ▪ 不可解析的地址：可变长度，不可解析的地址字段中包括不能解析的完整的地址 名参数(类型、长度和地址值)。
6	不识别的数据块类型	如果接收方不理解数据块且数据块类型比特中的高位比特设为 1，则把不识别的数据块类 给数据块的产生者  ▪ 不识别的数据块(可变长度)：该字段包含 SCTP 分组中不识别数据块的数据块类 志和数据块长度。
7	无效的必备参数	当一个必备参数被设置成无效值时，则向 INIT 或 INITACK 的生成者返回无效的必备参数 
8	不识别的参数	如果接收方不能识别 INIT ACK 数据块中一个或多个任选参数，则向 INIT ACK 数据块的 识别的参数的差错原因。  ▪ 不识别的参数：可变长度，该参数字段包含了从 INIT ACK 数据块中复制的完整 数。当 COOKIE ECHO 数据块的发送者希望报告不识别的参数时，这个参数通常是 数据块中与 COOKIE ECHO 数据块捆绑在一起发送作为对 INIT ACK 的响应。
9	无用户数据	如果收到的 DATA 数据块中未包含用户数据，则把这个差错原因返回给 DATA 数据块的产  ▪ TSN 的值(32bit 无符号整数)：该字段包含接收到的这个没有用户数据的 DATA 数

原因编码	含义	字段格式
		这个原因值通常是在 ABORT 数据块中返回的。
10	关闭期间收到 Cookie	当端点处于 SHUTDOWN-ACK-SENT 状态时，又收到 COOKIE ECHO 时则发送该差错原因。返回 因的 ERROR 数据块通常与重发的 SHUTDOWN ACK 数据块捆绑在一起发送。 
11	使用新地址重新启动偶联	当在现存的偶联上收到了 COOKIE ECHO 数据块，而 COOKIE ECHO 数据块又向该偶联中增加 的地址，此时使用该错误原因，并把新增加的地址作为差错信息在该参数中传送，这个差 都在 ABORT 中发送，用来拒绝 COOKIE ECHO 数据块。 

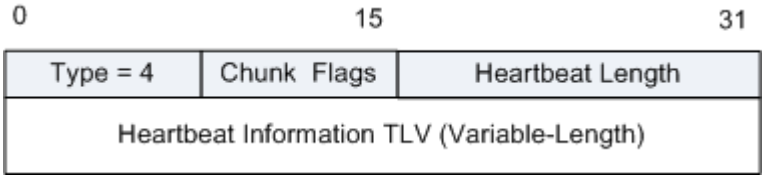
5.3.7 SCTP HEARTBEAT 数据块格式

HeartBeat 请求(HEARTBEAT)数据块的格式

SCTP 端点通过向对端点发送这个数据块用来检测定义在该偶联上到特定目的地传送地址的可达性。

参数字段包含 HEARTBEAT 信息，它是一个可变长度的非透明数据结构，其信息通常只需要发送方明白即可。

图 1 SCTP HEARTBEAT 数据块格式



- Chunk Flags (8bit)：在发送方设置为全 0，并在接收方忽略。
- HEARTBEAT Length(16bit)：设置为数据块长度的字节数，包括数据块头和 HEARTBEAT 信息字段。
- HEABTBEAT Information TLV：当该 HEARTBEAT 数据块发送到目的地传送地址时，发送方特定的 HEARTBEAT 信息字段通常包括关于发送方当前的时间信息。

报文示例

```
Frame 2: 74 bytes on wire (592 bits), 74 bytes captured (592 bits)
Ethernet II, Src: HuaweiTe_43:9c:e2 (78:1d:ba:43:9c:e2), Dst: HuaweiTe_88:76:68 (0
Internet Protocol Version 4, Src: 181.6.9.2 (181.6.9.2), Dst: 206.11.0.65 (206.11.
Stream Control Transmission Protocol, Src Port: m3ua (2905), Dst Port: m3ua (2905)
    Source port: 2905
    Destination port: 2905
    Verification tag: 0xe6f68e28
    Checksum: 0x5282fb7b (not verified)
  HEARTBEAT chunk (Information: 24 bytes)
    Chunk type: HEARTBEAT (4)
        0... .... = Bit: Stop processing of the packet
        .0.. .... = Bit: Do not report
      Chunk flags: 0x00
      Chunk length: 28
    Heartbeat info parameter (Information: 20 bytes)
      Parameter type: Heartbeat info (0x0001)
          0... .... .... .... = Bit: Stop processing of chunk
          .0.. .... .... .... = Bit: Do not report
        Parameter length: 24
        Heartbeat information: 0000000c000000022be18fc000010008ce0b0041
```

```
0000  00 18 82 88 76 68 78 1d  ba 43 9c e2 08 00 45 02   ....vhx. .C....E.
0010  00 3c f4 8a 00 00 fe 84  3b 5c b5 06 09 02 ce 0b   .<...... ;\......
0020  00 41 0b 59 0b 59 e6 f6  8e 28 52 82 fb 7b 04 00   .A.Y.Y.. .(R..{..
0030  00 1c 00 01 00 18 00 00  00 0c 00 00 00 02 2b e1   ........ ......+.
0040  8f c0 00 01 00 08 ce 0b  00 41                     ........ .A
```

参考标准

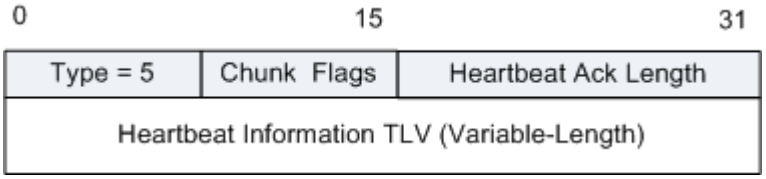
标准	描述
RFC 2960	Stream Control Transmission Protocol

5.3.8 SCTP HEARTBEAT ACK 数据块格式

HeartBeat 证实(HEARTBEAT ACK)数据块的格式

SCTP 端点在收到对端点发来的 HEARTBEAT 数据块后，则发送该数据块作为响应。HeartBeat 证实总是向包含 HEARTBEAT 数据块的 IP 数据报中的起源 IP 地址发送，来作为对该 HEARTBEAT 数据块的响应。

图 1 SCTP HEARTBEAT ACK 数据块格式



- 数据块标志(8bit)：在发送方设置为全 0，并在接收方忽略。
- HEARTBEAT 长度(16bit)：设置为数据块长度的字节数，包括数据块头和 HEARTBEAT 信息字段。
- HEABTBEAT 信息：可变长度，该字段的内容应当把 HEARTBEAT 请求数据块中的 HEARTBEAT 信息参数作为回送的响应，该参数字段包含一个可变长度的非透明的数据结构。

报文示例

```
⊞ Frame 3: 74 bytes on wire (592 bits), 74 bytes captured (592 bits)
⊞ Ethernet II, Src: HuaweiTe_88:76:68 (00:18:82:88:76:68), Dst: HuaweiTe_bb:b8:5c (2
⊞ Internet Protocol Version 4, Src: 206.11.0.65 (206.11.0.65), Dst: 181.6.9.2 (181.6
⊟ Stream Control Transmission Protocol, Src Port: m3ua (2905), Dst Port: m3ua (2905)
    Source port: 2905
    Destination port: 2905
    Verification tag: 0x2aaf9bf7
    Checksum: 0xd9bc8e79 (not verified)
  ⊟ HEARTBEAT_ACK chunk (Information: 24 bytes)
    ⊟ Chunk type: HEARTBEAT_ACK (5)
        0... .... = Bit: Stop processing of the packet
        .0.. .... = Bit: Do not report
      Chunk flags: 0x00
      Chunk length: 28
    ⊟ Heartbeat info parameter (Information: 20 bytes)
      ⊟ Parameter type: Heartbeat info (0x0001)
          0... .... .... .... = Bit: Stop processing of chunk
          .0.. .... .... .... = Bit: Do not report
        Parameter length: 24
        Heartbeat information: 0000000c000000022be18fc000010008ce0b0041
```

```
0000  28 6e d4 bb b8 5c 00 18  82 88 76 68 08 00 45 e0   (n...\.. ..vh..E.
0010  00 3c 30 76 00 00 40 84  bc 93 ce 0b 00 41 b5 06   .<0v..@. .....A..
0020  09 02 0b 59 0b 59 2a af  9b f7 d9 bc 8e 79 05 00   ...Y.Y*. .....y..
0030  00 1c 00 01 00 18 00 00  00 0c 00 00 00 02 2b e1   ........ ......+.
0040  8f c0 00 01 00 08 ce 0b  00 41                     ........ .A
```

参考标准

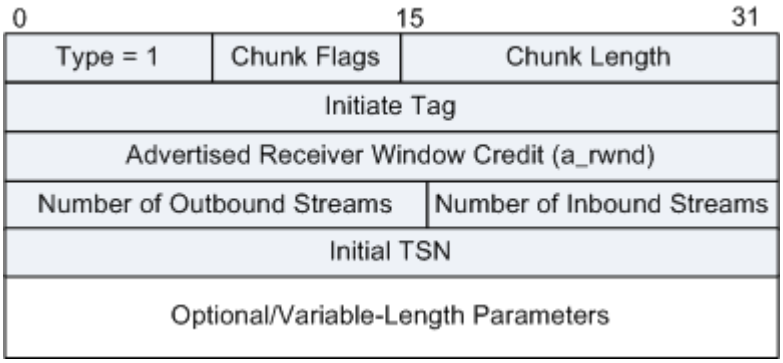
标准	描述
RFC 2960	Stream Control Transmission Protocol

5.3.9 SCTP INIT 数据块格式

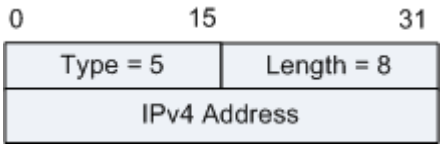
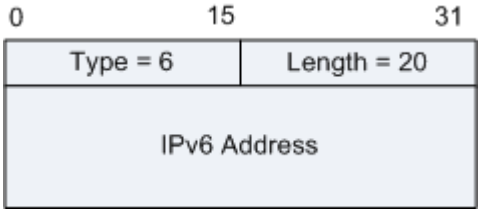
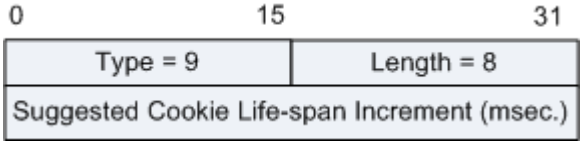
Initiation (INIT)数据块格式

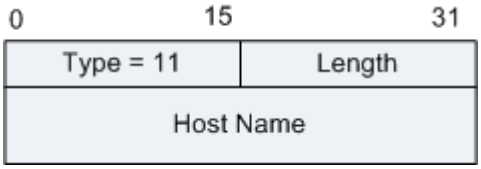
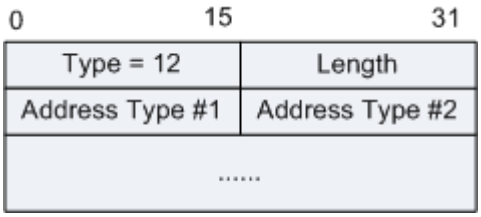
该数据块用来启动两个 SCTP 端点间的一个偶联，其格式如下：

图 1 SCTP INIT 数据块格式



字段	长度	描述
启动标签	32 比特无符号整数	INIT 的接收方(响应端)记录启动标签参数的值。这个值必须被放置到 INIT 的接收方发送的与该偶联相关的每个 SCTP 分组中的验证标签字段中。启动标签允许除 0 以外的的任何值。如果在收到的 INIT 数据块中的启动标签为 0，则接收方必须作为错误处理，并且发送 ABORT 数据块中止该偶联。

字段	长度	描述
A 通告的接收方窗口信用(a_rwnd)	32 比特无符号整数	表示指定的缓冲区的容量，用字节数表示，为 INIT 发送方为偶联预留的窗口大小。在偶联存活期间，这个缓冲区的容量不应减少(即不应把该偶联的专用缓冲区取走)，但端点可以在发送的 SACK 数据块中修改 a_rwnd 的值。
输出流数量	16 比特无符号整数	用来定义发送 INIT 数据块的一方希望在该偶联中创建的输出流的数量。该值不允许为 0，接收方收到该参数为 0 的 INIT 数据块后应中止该偶联。
输入流数量	16 比特无符号整数	定义了发送这个 INIT 块的一方允许对端在该偶联中所创建的流的数量。该值不允许为 0，接收方收到该参数为 0 的 INIT 数据块后应中止该偶联。
初始的 TSN	32 比特无符号整数	定义发送方将使用的初始的 TSN，该值可以设置为启动标签字段的值。
任选 / 可变长参数	可变长度	• IPv4 地址参数(5)  IPv4 地址（32bit 无符号整数）：包含发送方端点的 IPv4 地址，采用二进制编码。 • IPv6 地址参数(6)  IPv6 地址(128bit 的无符号整数)：包含发送方端点的 IPv6 地址，采用二进制编码。发送方不必把 IPv4 地址映射到 IPv6 地址中，可以直接在 IPv4 地址参数中使用 IPv4 地址。 • 防止 Cookie 过期参数(9) INIT 的发送方应使用这个参数来建议 INIT 的接收方提供较长存活跨度的状态 CooKIE。  建议的 COOKIE 存活跨度增量(32bit 的无符号整数)，该参数用来向接收方指示发送方希望接收方为其缺省的 COOKIE 的存活跨度增加的毫秒数。

字段	长度	描述
		由于失效的 cookie 操作差错原因，前一次尝试与对等端建立偶联失败后，又重新尝试偶联建立时，这个任选参数应能由发送方添加到 INIT 数据块中。接收方出于安全的考虑可以选择忽略建议的 COOKIE 存活跨度增量。 • 主机名地址(11) INIT 发送方使用这个参数把其主机名(在其 IP 地址的位置中)传递到对等层。这个对等层负责解析这个主机名，用这个参数可以使偶联工作通过 NAT box 进行工作。  主机名：可变长度，该字段包含了按照 RFC1123 规定的“主机名句法”定义的主机名，主机名地址的解析不在本标准中规定，该参数中至少有一个非空的中止符包含在主机名字符串中，并且应包含长度。 • 支持的地址类型(12) INIT 的发送方使用该参数列出其所支持的全部地址类型。  地址类型(16bit 无符号整数)：该参数使用对应的地址类型的类型值(例如：IPv4=

报文示例

```
⊞ Frame 41: 78 bytes on wire (624 bits), 78 bytes captured (624 bits)
⊞ Ethernet II, Src: Hangzhou_d5:c5:29 (00:0f:e2:d5:c5:29), Dst: Hangzhou_2f:d7:18 (00:0f
⊞ Internet Protocol Version 4, Src: 172.21.112.2 (172.21.112.2), Dst: 192.168.168.9 (192
⊟ Stream Control Transmission Protocol, Src Port: 3876 (3876), Dst Port: diameter (3868)
    Source port: 3876
    Destination port: 3868
    Verification tag: 0x00000000
    Checksum: 0x8bbee7b7 (not verified)
  ⊟ INIT chunk (Outbound streams: 16, inbound streams: 16)
    ⊟ Chunk type: INIT (1)
        0... .... = Bit: Stop processing of the packet
        .0.. .... = Bit: Do not report
      Chunk flags: 0x00
      Chunk length: 28
      Initiate tag: 0xce2c0b56
      Advertised receiver window credit (a_rwnd): 486500
      Number of outbound streams: 16
      Number of inbound streams: 16
      Initial TSN: 0
    ⊟ IPv4 address parameter (Address: 172.21.112.2)
      ⊟ Parameter type: IPv4 address (0x0005)
          0... .... .... .... = Bit: Stop processing of chunk
          .0.. .... .... .... = Bit: Do not report
        Parameter length: 8
        IP Version 4 address: 172.21.112.2 (172.21.112.2)
```

```
0000  00 0f e2 2f d7 18 00 0f  e2 d5 c5 29 08 00 45 e0   .../.... ...)..E.
0010  00 3c af 45 00 00 3f 84  46 4f ac 15 70 02 c0 a8   .<.E..?. FO..p...
0020  a8 09 0f 24 0f 1c 00 00  00 00 8b be e7 b7 01 00   ...$.... ........
0030  00 1c ce 2c 0b 56 00 07  6c 64 00 10 00 10 00 00   ...,.V.. ld......
0040  00 00 00 05 00 08 ac 15  70 02 1f 8f 2d 33         ........ p...-3
```

参考标准

标准	描述
RFC 2960	Stream Control Transmission Protocol

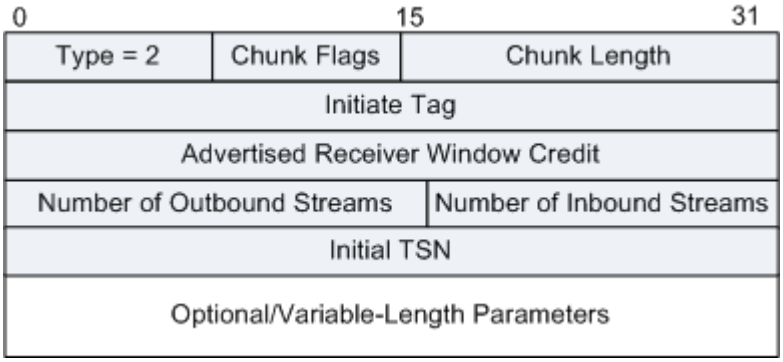
5.3.10 SCTP INIT ACK 数据块格式

Format of Initiation Acknowledgement (INIT ACK)

INIT ACK 数据块用来确认 SCTP 偶联的启动。

INIT ACK 的参数部分与 INIT 数据块的参数部分相同，它额外还使用两个的可变长度的参数即：状态 COOKIE(STATE COOKIE)和未识别的参数。

图 1 SCTP INIT ACK 数据块格式



INIT ACK 数据块应包含以下参数：

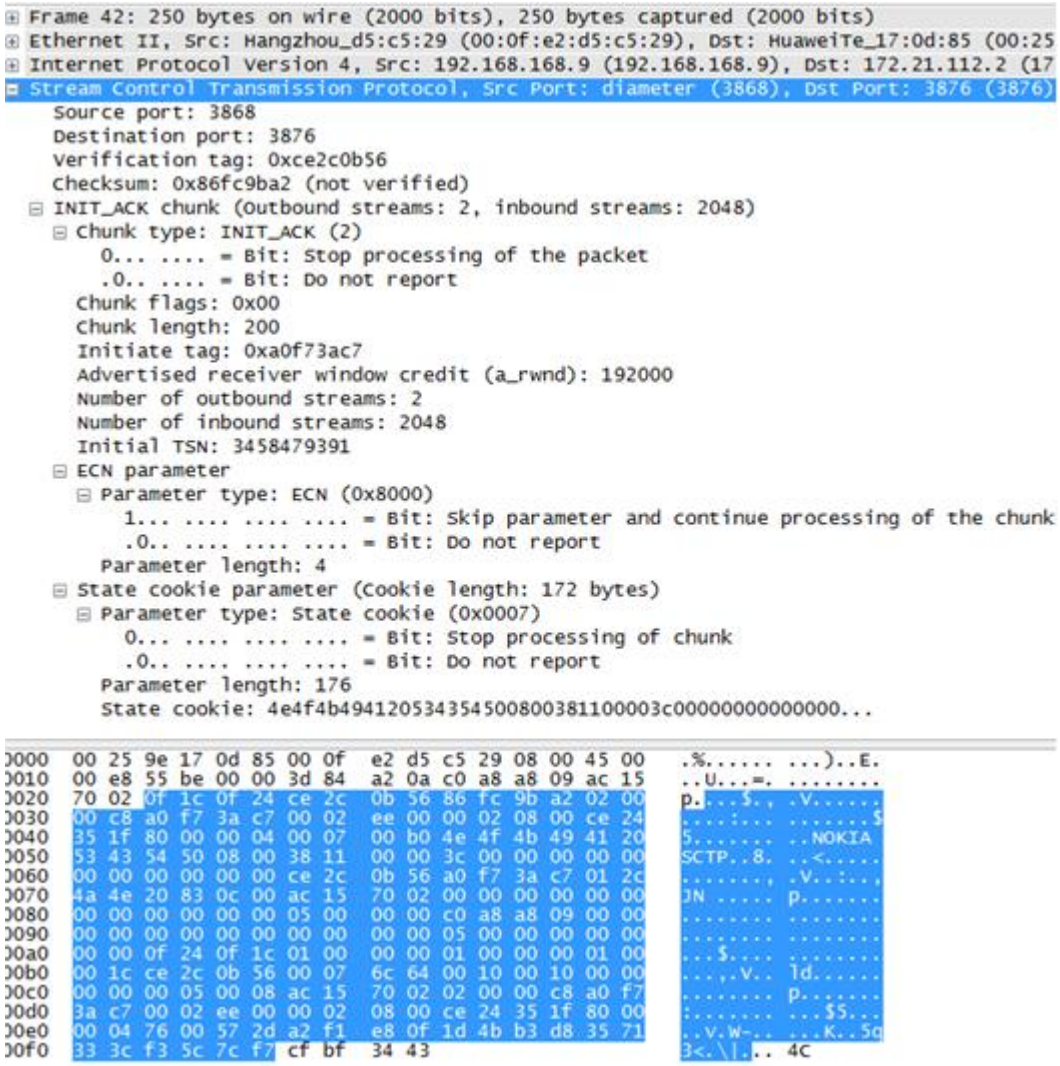
- 必备参数：
  - 启动标签；
  - 通告的接收方窗口信用；
  - 输入流数量；
  - 输出流数量；
  - 初始 TSN。
- 可变长参数：
  - STATECOOKIE 类型值=7 必备
  - IPv4 地址 类型值=5 任选
  - IPv6 地址 类型值=6 任选
  - 未识别的参数 类型值=9 任选
  - ECN 能力预留 类型值=32768(0x8000) 任选

- 主机名地址 类型值=11 任选

字段	长度	描述
启动标签（Initiate Tag）	32 比特无符号整数	INIT ACK 的接收方记录启动标签参数的值，并把该值放到 INITACK 接收方需要在相应的偶联上发送的每个 SCTP 分组中的验证标签中。启动标签不允许为 0。如果收到的 INIT ACK 数据块中的启动标签为 0，则接收方当作错误来处理并通过发送 ABORT 来关闭偶联。
通告的接收方窗口信用值 Advertised Receiver Window Credit (a_rwnd)	32 比特无符号整数	这个值表示指定的缓冲区的容量，用字节数表示，是 INIT ACK 发送方为偶联预留的窗口，在偶联存活期间，这个缓冲区的容量不应减少(即不应把该偶联的专用缓冲区取走)。
输出流数量 Number of Outbound Streams (OS)	16 比特无符号整数	定义发送 INIT ACK 数据块的一方希望在该偶联中创建的输出流的数量。该值不允许为 0，接收方收到该参数为 0 的 INIT ACK 数据块后应中止该偶联并舍弃 TCB。
输入流数量 Number of Inbound Streams (MIS)	16 比特无符号整数	定义发送 INIT ACK 数据块的一方允许对端点在该偶联中所创建的流的最大数量。该值不允许为 0，接收方收到该参数为 0 的 INIT ACK 数据块后应中止该偶联并舍弃该 TCB。
Initial TSN(I-TSN)	32 比特无符号整数	定义发送方将使用的初始的 TSN，该值可以设置为启动标签字段的值。
Optional/Variable-length Parameters	变长	• State Cookie：该参数类型为 7，为可变长度参数，该参数长度取决于 COOKIE 的长度，该参数值的取值必须包含由 INIT ACK 发送方创建该偶联所需的所有状态和参数信息，连同消息授权码。 • 不识别的参数：该参数类型为 8，可变长度参数。该参数内容是 INIT 数据块中包含的一个不识别的参数，该参数用来返回给 INIT 数据块的产生者一个指示，这个参数值字段包含了从 INIT 数据块中复制过来的不识别参数的完整的参数类型、长度和参数

字段	长度	描述
		值。

报文示例



参考标准

标准	描述
RFC 2960	Stream Control Transmission Protocol

5.3.11 SCTP SACK 数据块格式

选择证实(SACK)数据块的格式

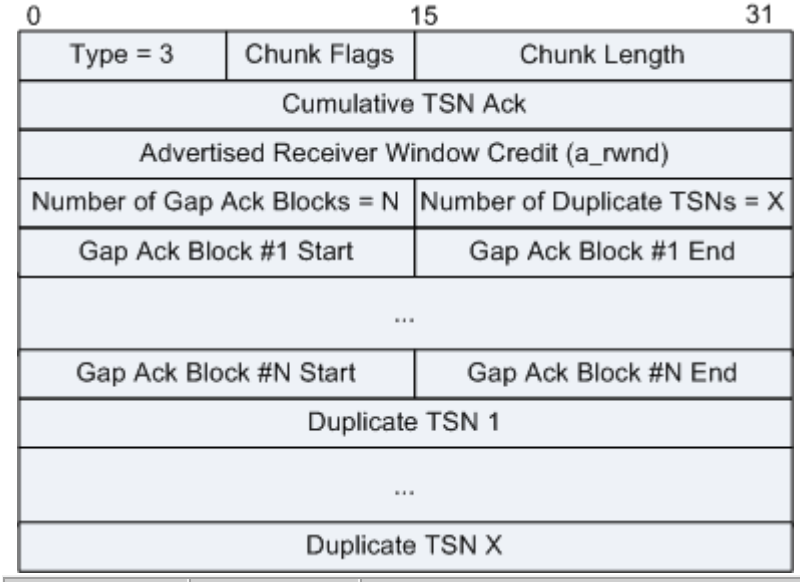
这个数据块通过使用 DATA 数据块中的 TSN 用来向对等的端点确认接收到的 DATA 数据块，并通知对等的端点在收到的 DATA 数据块中的间隔。所谓间隔就是指收到的 DATA 数据块的 TSN 不连续的情况。

SACK 必须包含累积的 TSN 证实和通告的接收方窗口信用(a_rwnd)参数。

累积的 TSN 证实参数的值是指收到的 TSN 顺序断开前的最后一个 TSN 号码，下一个 TSN 则是在发送 SACK 的端点尚未收到的 TSN 值。所以这个参数确认已经收到了小于或等于该值的所有 TSN。

SACK 中可以包含 0 个或多个间隔证实块，每个间隔证实块确认了在一个不连续 TSN 后所收到的 TSN 序列，所有通过间隔证实块确认的 TSN 值都应比累积 TSN 证实的值大。

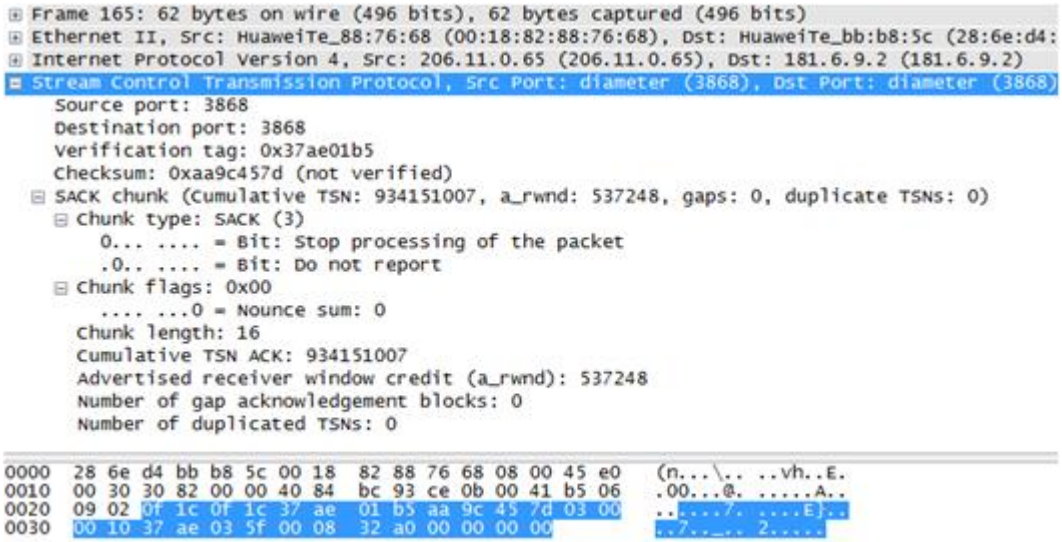
图 1 SCTP SACK 数据块格式



字段	长度	描述
数据块标志位 Chunk Flags	8 比特	设为全 0 并由接收方忽略。
Cumulative TSN Ack	32 比特无符号整数	该参数包含了在收到 TSN 序列的间隔前的最后一个 TSN 值。
Advertised Receiver Window Credit (a_rwnd)	32 比特无符号整数	该字段指示修改了 SACK 的发送方的接收缓冲容量的字节数。
Number of Gap Ack Blocks	16 比特无符号整数	用来指示 SACK 数据块中包含的间隔证实块的数目。
Number of Duplicate	16 比特	该字段包含了该端点收到的重复的 TSN 的数目。每个重复的 TSN 都列在间隔证实块列表后。

字段	长度	描述
TSNs		
Gap Ack Block	变长	这个字段中包含了间隔证实块，根据间隔证实块数量字段给出的值，间隔证实块重复若干次。所有 TSN 大于或等于累积 TSN 证实+间隔证实块开始的 DATA 数据块，或者是小于或等于每个间隔证实块的累积 TSN 证实+间隔证实块结束的 DATA 数据块都被看作是被正确地接收了。
Gap Ack Block Start	16 比特无符号整数	该字段用来指示这个间隔整数块的起始 TSN 偏移，为了计算实际的 TSN 号码必须要用累积 TSN 证实加上偏移号码。计算出的 TSN 标识用来识别第一个在这个间隔证实块中被收到的 TSN。
Gap Ack Block End	16 比特无符号整数	用来指示这个间隔证实块的结束 TSN 偏移，为了计算实际的 TSN 需要把累积 TSN 证实加上这个偏移号码。这个计算出的 TSN 用来识别在这个间隔证实块中最后收到的 DATA 数据块。
Duplicate TSN	32 比特无符号整数	用来指示一个在上一个 SACK 发送后收到的 TSN 重复的个数。每次一个接收者收到一个重复的 TSN(在发送 SACK 前)，则把这个 TSN 加到重复的 TSN 列表中。每发送一次 SACK 后则把统计重复 TSN 的计数器重新清 0。

报文示例



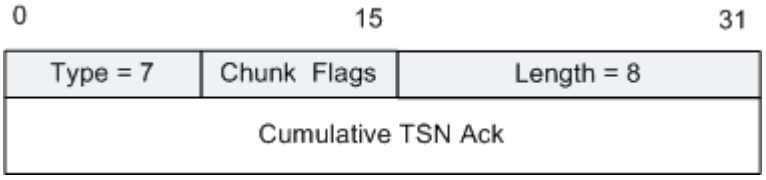
参考标准

标准	描述
RFC 2960	Stream Control Transmission Protocol

5.3.12 SCTP SHUTDOWN 消息格式

报文格式

图 1 SCTP SHUTDOWN 消息格式



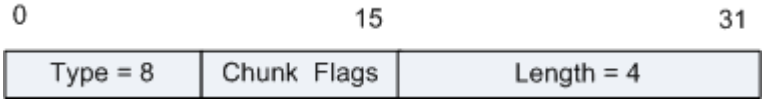
- Chunk Flags: 块标识，8 比特。
- Length: 16 比特（无符号整数），标识参数的长度，值为 8。
- Cumulative TSN Ack: 32 比特（无符号整数），包含在任何间隔之前接收的最后一个 TNS。

5.3.13 SCTP SHUTDOWN ACK 数据块格式

关闭证实(SHUTDOWN ACK)数据块的格式

在完成了关闭程序后必须使用该数据块来确认收到的 SHUTDOWN 数据块。

图 1 SCTP SHUTDOWN ACK 数据块格式



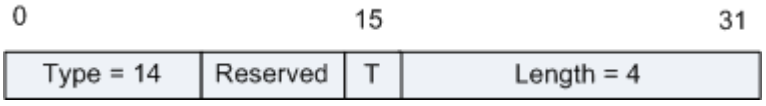
数据块标志位(8bit)：在发送方设置为全 0，并在接收方忽略。SHUTDOWN ACK 中不再包含其他参数。

5.3.14 SCTP SHUTDOWN COMPLETE 数据块格式

关闭完成(SHUTDOWN COMPLETE)数据块的格式

该数据块在完成关闭程序后用来确认收到 SHUTOOWN ACK 数据块。SHUTDOWN COMPLETE 数据块中不包含其他参数。

图 1 SCTP SHUTDOWN COMPLETE 数据块格式



- Reserved：保留位。7 比特，在发送方设为全 0，在接收方忽略。
- T：1 比特。当发送方有一个 TCB 被破坏时则该 T 比特设置为 0；如果发送方没有 TCB，则把该比特设置为 1。

6. 应用层

- 1588v2（PTP）报文格式
- BFD 控制报文格式

- BGP 报文格式
- BOOTP 报文格式
- DHCP 报文格式
- DHCPv6 报文格式
- Diameter 协议报文格式
- DNS 报文格式
- IP FPM 报文格式
- IPSec 报文格式
- L2TP 报文格式
- MPLS LDP 报文格式
- MSDP 报文格式
- NetStream 报文格式
- RIP 报文格式
- RIPng 的报文格式
- NTP 报文格式
- RADIUS 报文格式
- SNMP 报文格式
- TWAMP 报文格式

6.1 1588v2（PTP）报文格式

- 1588v2（PTP）报文通用格式
- 1588v2 Sync 消息和 Delay_Req 消息
- 1588v2 Follow_Up 消息
- 1588v2 Delay_Resp 消息
- 1588v2 Pdelay_Req 消息
- 1588v2 Pdelay_Resp 消息
- 1588v2 Pdelay_Resp_Follow_Up 消息
- 1588v2 Signaling 消息

- 1588v2 Management 消息

6.1.1 1588v2（PTP）报文通用格式

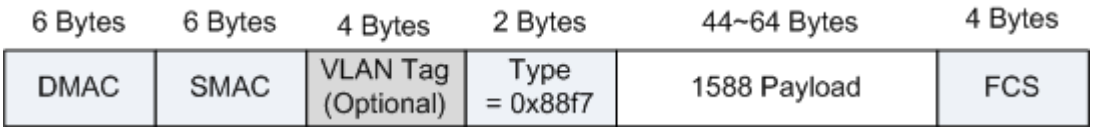
IEEE 1588v2 协议附录 D 定义了 1588V2 over IPv4 的报文封装，附录 E 定义了 1588V2 over IPv6 的报文封装，附录 F 定义了 1588v2 over IEEE 802.3 /Ethernet 的报文封装。其他如 1588v2 over MPLS 封装，业界还没有成熟的标准。除此之外，在实际应用中还可能携带 VLAN。

- PTP over Ethernet
- PTP over UDP over IPv4
- PTP over UDP over IPv6
- PTP over MPLS

PTP 通用消息格式 PTP General Message Format

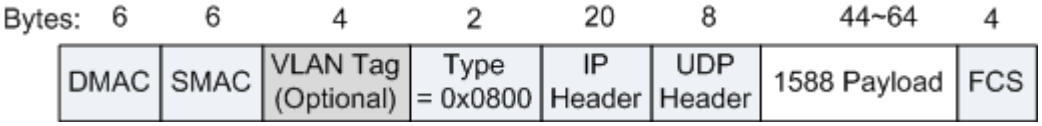
以太封装的 PTP (Precision Timing Protocol)报文，其帧头中以太类型值=

图 1 以太封装 PTP 报文 PTP over Ethernet



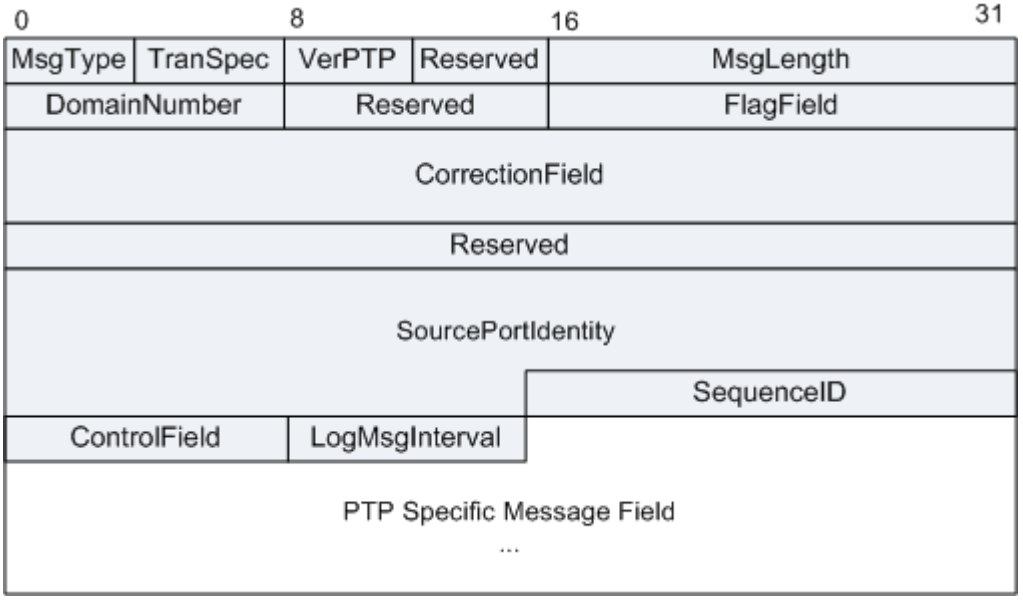
IPv4 封装 PTP 报文，EVENT 消息头的 UDP 目的端口号是 319，General 消息的 UDP 目的端口号是 320。

图 2 IPv4 封装 PTP 报文 Format of PTP packet over UDP over IPv4



PTP 消息头格式 Format of a PTP header

1588v2 消息必须包含消息头、消息体和消息扩展字节，扩展字节长度可能为 0。



字段	长度	含义
TranSpec	4 比特	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 比特	表示消息类型。1588V2 消息分为两类：事件消息（EVENT Message）和通用消息（General Message）。事件报文是时间概念报文，进出设备端口时需要打上精确的时间戳，而通用报文则是非时间概念报文，进出设备不会产生时戳。类型值 0~3 的为事件消息，8~D 为通用消息。 • 0x00: Sync • 0x01: Delay_Req • 0x02: Pdelay_Req • 0x03: Pdelay_Resp • 0x04-7: Reserved • 0x08: Follow_Up • 0x09: Delay_Resp • 0x0A: Pdelay_Resp_Follow_Up • 0x0B: Announce • 0x0C: Signaling • 0x0D: Management • 0x0E-0x0F: Reserved
Reserved	4 比特	保留字段。
VerPTP	4 比特	表示 1588 协议的版本。
MsgLength	2 字节	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 字节	域编号，表示发送该消息时钟所属的域。
Reserved	1 字节	保留字段。
FlagField	2 字节	标志域。取值请参见表 2。
CorrectionField	64 比特	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。

字段	长度	含义
Reserved	32 比特	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 字节	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 字节	控制域，由消息类型决定： • 0x00：Sync • 0x01：Delay_Req • 0x02：Follow_Up • 0x03：Delay_Resp • 0x04：Management • 0x05：All others • 0x06-0xFF：reserved
LogMsgInterval	1 字节	录入消息周期，PTP 消息的发送时间间隔，由消息类型决定。
PTP Specified Message Field	变长	PTP 消息体和消息扩展字节。

表 1 flagField 的取值

字节	比特位	消息类型	消息名	说明
0	0	Announce, Sync, Follow_Up, Delay_Resp	alternateMasterFlag	如果发送侧端口处于 MASTER 状态，则为 FALSE。
0	1	Sync, Pdelay_Resp	twoStepFlag	对于一步时钟，twoStepFlag 取值要求为 FALSE。对于双步时钟，twoStepFlag 取值要求为 TRUE。
0	2	ALL	unicastFlag	如果此消息发送到的传送层协议地址是一个单播地址，则置为 TRUE。如果此消息发送到的传送层协议地址是一个多播地址，则置为 FALSE。

字段			长度	含义
0	5	ALL	PTP profile Specific 1	由一个备选 PTP 模板定义；否则置为 FALSE。
0	6	ALL	PTP profile Specific 2	由一个备选 PTP 模板定义；否则置为 FALSE。
0	7	ALL	Reserved	此比特保留用于实验性安全机制.
1	0	Announce	leap61	-
1	1	Announce	leap59	-
1	2	Announce	currentUtcOffsetValid	-
1	3	Announce	ptpTimescale	-
1	4	Announce	timeTraceable	-
1	5	Announce	frequencyTraceable	-

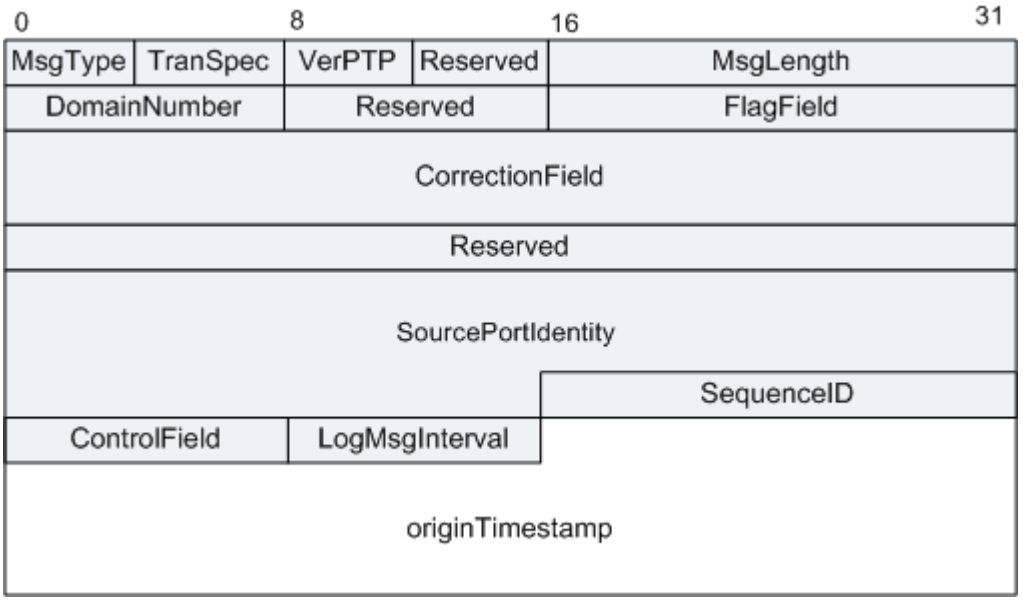
参考标准

标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.1.2 1588v2 Sync 消息和 Delay_Req 消息

报文格式

图 1 Sync 消息和 Delay_Req 消息的格式



字段	长度	含义
TranSpec	4 bits	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 bits	• 0x00: Sync • 0x01: Delay_Req
Reserved	4 bits	保留字段。
VerPTP	4 bits	表示 1588 协议的版本。
MsgLength	2 bytes	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 byte	域编号，表示发送该消息时钟所属的域。
Reserved	1 byte	保留字段。
FlagField	2 bytes	标志域。
CorrectionField	64 bits	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。

字段	长度	含义
Reserved	32 bits	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 bytes	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 byte	控制域，由消息类型决定。
LogMsgInterval	1 byte	录入消息周期，PTP 消息的发送时间间隔。
originTimestamp	10 bytes	源时间标签。

报文示例

图 2 基于 UDP 的 Sync 消息

```
⊞ Frame 1: 90 bytes on wire (720 bits), 90 bytes captured (720 bits)
⊞ Ethernet II, Src: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13), Dst: HuaweiTe_98:f2
⊞ Internet Protocol Version 4, Src: 10.1.1.1 (10.1.1.1), Dst: 224.0.1.129 (224
⊞ User Datagram Protocol, Src Port: ptp-event (319), Dst Port: ptp-event (319)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = V1 Compatibility: False
    .... 0000 = messageId: Sync Message (0x00)
    .... 0010 = versionPTP: 2
    messageLength: 44
    subdomainNumber: 0
  ⊟ flags: 0x022c
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .0.. .... .... = PTP_UNICAST: False
      .... ..1. .... .... = PTP_TWO_STEP: True
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .1.. = PTP_UTC_REASONABLE: True
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 0.000000 nanoseconds
      correction: Ns: 0 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00259e1000000001
    SourcePortID: 1280
    sequenceId: 1208
    control: Sync Message (0)
    logMessagePeriod: -10
    originTimestamp (seconds): 0
    originTimestamp (nanoseconds): 0
```

图 3 基于以太的 Sync 消息

```
⊞ Frame 2: 64 bytes on wire (512 bits), 64 bytes captured (512 b
⊟ Ethernet II, Src: HuaweiTe_83:7b:0d (00:25:9e:83:7b:0d), Dst:
  ⊞ Destination: HuaweiTe_46:8f:7e (28:6e:d4:46:8f:7e)
  ⊞ Source: HuaweiTe_83:7b:0d (00:25:9e:83:7b:0d)
    Type: PTPv2 over Ethernet (IEEE1588) (0x88f7)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = 802.1as conform: False
    .... 0000 = messageId: Sync Message (0x00)
    .... 0010 = versionPTP: 2
    messageLength: 44
    subdomainNumber: 0
  ⊟ flags: 0x0628
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .1.. .... .... = PTP_UNICAST: True
      .... ..1. .... .... = PTP_TWO_STEP: True
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .0.. = PTP_UTC_REASONABLE: False
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: -137816927759630.000000 nanoseconds
      correction: Ns: -137816927759630 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00aa12fffe431112
    SourcePortID: 1024
    sequenceId: 12243
    control: Sync Message (0)
    logMessagePeriod: 127
    originTimestamp (seconds): 0
    originTimestamp (nanoseconds): 0
```

图 4 基于 UDP 的 Delay_Req 消息

```
⊞ Frame 5: 94 bytes on wire (752 bits), 94 bytes captured (752 bits)
⊞ Ethernet II (VLAN tagged), Src: HuaweiTe_00:00:11 (00:18:82:00:00:11), Dst:
⊞ Internet Protocol Version 4, Src: 82.0.1.2 (82.0.1.2), Dst: 82.0.1.50 (82.0.
⊞ User Datagram Protocol, Src Port: ptp-event (319), Dst Port: ptp-event (319)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = V1 Compatibility: False
    .... 0001 = messageId: Delay_Req Message (0x01)
    .... 0010 = versionPTP: 2
    messageLength: 44
    subdomainNumber: 0
  ⊟ flags: 0x050a
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .1.. .... .... = PTP_UNICAST: True
      .... ..0. .... .... = PTP_TWO_STEP: False
      .... ...1 .... .... = PTP_ALTERNATE_MASTER: True
      .... .... ..0. .... = FREQUENCY_TRACEABLE: False
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .0.. = PTP_UTC_REASONABLE: False
      .... .... .... ..1. = PTP_LI_59: True
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 0.000000 nanoseconds
      correction: Ns: 0 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x704433fffe297564
    SourcePortID: 4363
    sequenceId: 48672
    control: Delay_Req Message (1)
    logMessagePeriod: 127
    originTimestamp (seconds): 0
    originTimestamp (nanoseconds): 0
```

图 5 基于以太的 Delay_Req 消息

```
⊞ Frame 74: 64 bytes on wire (512 bits), 64 bytes captured (512
⊟ Ethernet II, Src: 0e:e0:f0:00:00:02 (0e:e0:f0:00:00:02), Dst:
  ⊞ Destination: IeeeI&MS_00:00:00 (01:1b:19:00:00:00)
  ⊞ Source: 0e:e0:f0:00:00:02 (0e:e0:f0:00:00:02)
    Type: PTPv2 over Ethernet (IEEE1588) (0x88f7)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = 802.1as conform: False
    .... 0001 = messageId: Delay_Req Message (0x01)
    .... 0010 = versionPTP: 2
    messageLength: 44
    subdomainNumber: 0
  ⊟ flags: 0x0128
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .0.. .... .... = PTP_UNICAST: False
      .... ..0. .... .... = PTP_TWO_STEP: False
      .... ...1 .... .... = PTP_ALTERNATE_MASTER: True
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .0.. = PTP_UTC_REASONABLE: False
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: -22895715404183.000000 nanoseconds
      correction: Ns: -22895715404183 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00259e1000000003
    SourcePortID: 3330
    sequenceId: 19267
    control: Delay_Req Message (1)
    logMessagePeriod: 127
    originTimestamp (seconds): 0
    originTimestamp (nanoseconds): 0
```

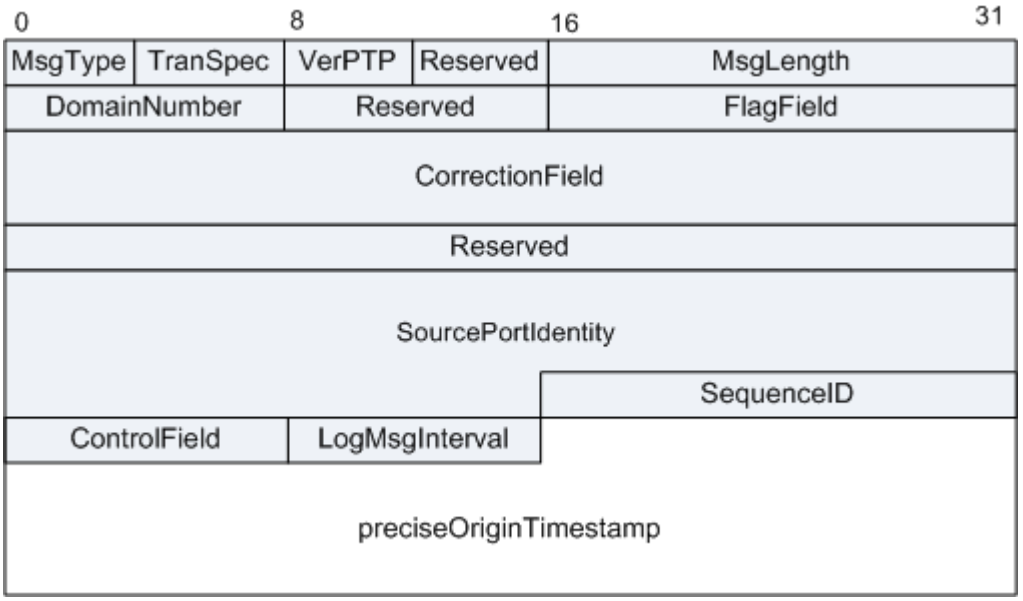
参考标准

标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.1.3 1588v2 Follow_Up 消息

报文格式

图 1 Follow_Up 消息格式



字段	长度	含义
TranSpec	4 bits	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 bits	消息类型。
Reserved	4 bits	保留字段。
VerPTP	4 bits	表示 1588 协议的版本。
MsgLength	2 bytes	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 byte	域编号，表示发送该消息时钟所属的域。
Reserved	1 byte	保留字段。
FlagField	2 bytes	标志域。
CorrectionField	64 bits	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。

字段	长度	含义
Reserved	32 bits	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 bytes	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 byte	控制域，由消息类型决定。
LogMsgInterval	1 byte	录入消息周期，PTP 消息的发送时间间隔。
preciseOriginTimestamp	10 bytes	精确源时间标签。PTP 提供传输时间戳的机制，这个时间戳包括事件消息产生的时刻和相应的修正域，通过这个机制保证接收方接收到的是最精确的时间戳。在实际应用中时间戳分布在：originTimestamp 或者 preciseOriginstamp 和 correctionField 中，由具体的执行决定。

报文示例

图 2 基于 UDP 的 Follow_Up 消息

```
⊞ Frame 2: 90 bytes on wire (720 bits), 90 bytes captured (720 bits)
⊞ Ethernet II, Src: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13), Dst: HuaweiTe_98:f2:6
⊞ Internet Protocol Version 4, Src: 10.1.1.1 (10.1.1.1), Dst: 224.0.1.129 (224.0
⊞ User Datagram Protocol, Src Port: ptp-event (319), Dst Port: ptp-general (320)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = V1 Compatibility: False
    .... 1000 = messageId: Follow_Up Message (0x08)
    .... 0010 = versionPTP: 2
    messageLength: 44
    subdomainNumber: 0
  ⊟ flags: 0x022c
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .0.. .... .... = PTP_UNICAST: False
      .... ..1. .... .... = PTP_TWO_STEP: True
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .1.. = PTP_UTC_REASONABLE: True
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 0.000000 nanoseconds
      correction: Ns: 0 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00259e1000000001
    SourcePortID: 1280
    sequenceId: 1208
    control: Follow_Up Message (2)
    logMessagePeriod: -10
    preciseOriginTimestamp (seconds): 947131157
    preciseOriginTimestamp (nanoseconds): 758319957
```

图 3 基于以太的 Follow_Up 消息

```
⊞ Frame 3: 64 bytes on wire (512 bits), 64 bytes captured (
⊟ Ethernet II, Src: HuaweiTe_83:7b:0d (00:25:9e:83:7b:0d),
  ⊞ Destination: HuaweiTe_46:8f:7e (28:6e:d4:46:8f:7e)
  ⊞ Source: HuaweiTe_83:7b:0d (00:25:9e:83:7b:0d)
    Type: PTPv2 over Ethernet (IEEE1588) (0x88f7)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = 802.1as conform: False
    .... 1000 = messageId: Follow_Up Message (0x08)
    .... 0010 = versionPTP: 2
    messageLength: 44
    subdomainNumber: 0
  ⊟ flags: 0x0628
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .1.. .... .... = PTP_UNICAST: True
      .... ..1. .... .... = PTP_TWO_STEP: True
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .0.. = PTP_UTC_REASONABLE: False
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 0.000000 nanoseconds
      correction: Ns: 0 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00aa12fffe431112
    SourcePortID: 1024
    sequenceId: 12243
    control: Follow_Up Message (2)
    logMessagePeriod: 127
    preciseOriginTimestamp (seconds): 946878901
    preciseOriginTimestamp (nanoseconds): 214732091
```

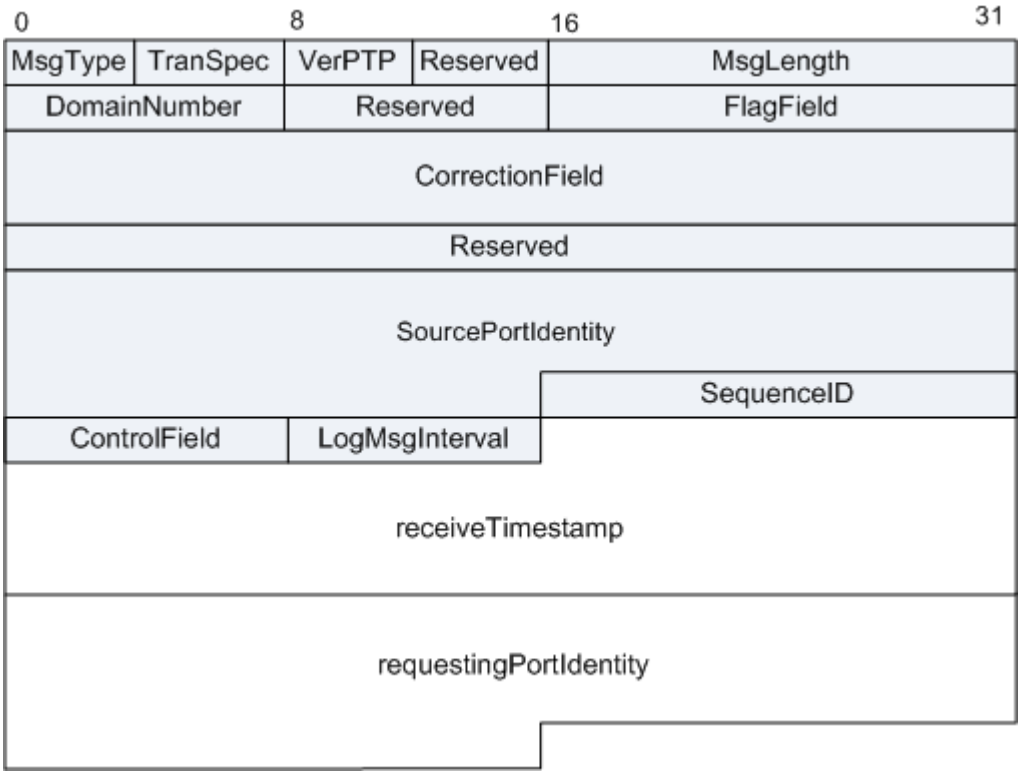
参考标准

标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.1.4 1588v2 Delay_Resp 消息

报文格式

图 1 Delay_Resp 消息格式



字段	长度	含义
TranSpec	4 bits	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 bits	表示消息类型。
Reserved	4 bits	保留字段。
VerPTP	4 bits	表示 1588 协议的版本。
MsgLength	2 bytes	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 byte	域编号，表示发送该消息时钟所属的域。
Reserved	1 byte	保留字段。
FlagField	2 bytes	标志域。

字段	长度	含义
CorrectionField	64 bits	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。
Reserved	32 bits	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 bytes	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 byte	控制域，由消息类型决定。
LogMsgInterval	1 byte	录入消息周期，PTP 消息的发送时间间隔。
receiveTimestamp	10 bytes	接收时间戳。
requestingPortIdentity	10 bytes	请求端口标识。

报文示例

图 2 基于 UDP 的 Delay_Resp 消息

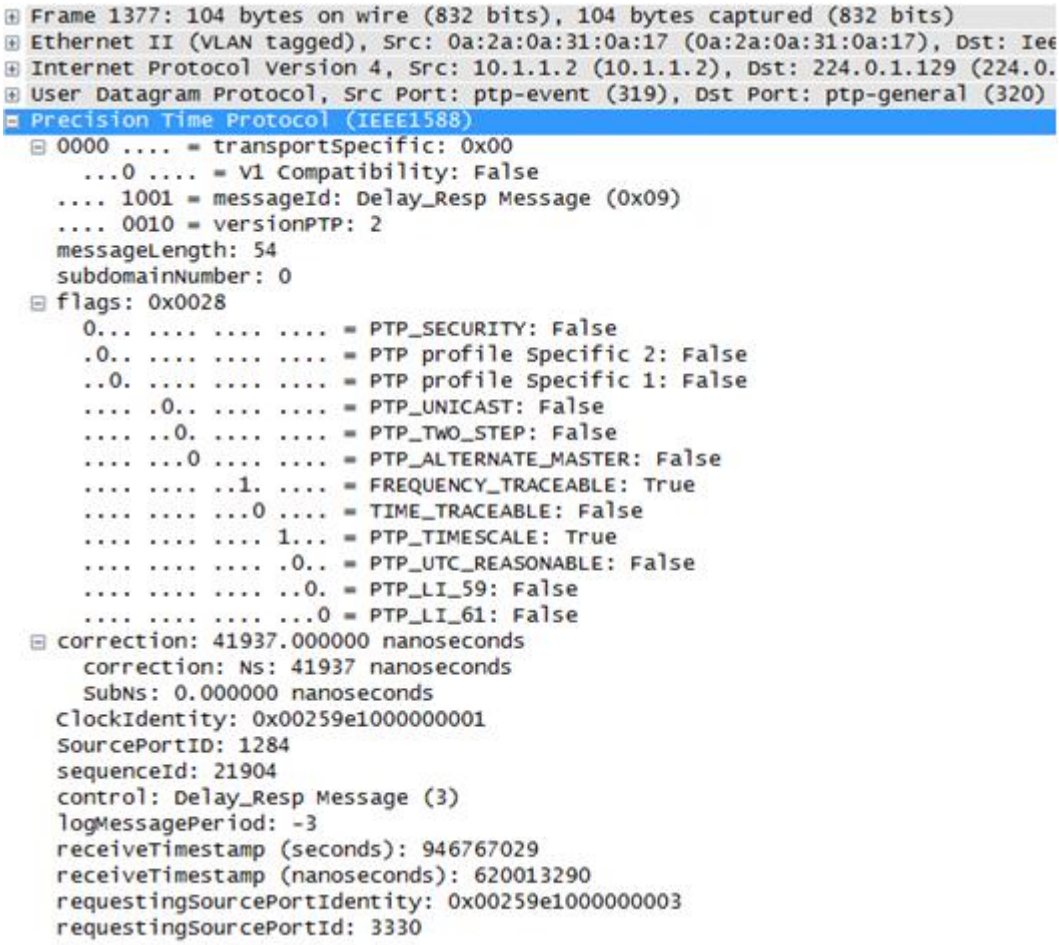


图 3 基于以太的 Delay_Resp 消息

```
⊞ Frame 75: 76 bytes on wire (608 bits), 76 bytes captured (60
⊟ Ethernet II (VLAN tagged), Src: 0a:2a:0a:31:0a:17 (0a:2a:0a:
  ⊞ Destination: 0e:e0:f0:00:00:02 (0e:e0:f0:00:00:02)
  ⊞ Source: 0a:2a:0a:31:0a:17 (0a:2a:0a:31:0a:17)
  ⊞ VLAN tag: VLAN=10, Priority=Best Effort (default)
    Type: PTPv2 over Ethernet (IEEE1588) (0x88f7)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = 802.1as conform: False
    .... 1001 = messageId: Delay_Resp Message (0x09)
    .... 0010 = versionPTP: 2
    messageLength: 54
    subdomainNumber: 0
  ⊟ flags: 0x0428
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .1.. .... .... = PTP_UNICAST: True
      .... ..0. .... .... = PTP_TWO_STEP: False
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .0.. = PTP_UTC_REASONABLE: False
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 37288.000000 nanoseconds
      correction: Ns: 37288 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00259e1000000001
    SourcePortID: 1284
    sequenceId: 19267
    control: Delay_Resp Message (3)
    logMessagePeriod: 127
    receiveTimestamp (seconds): 946766674
    receiveTimestamp (nanoseconds): 120913862
    requestingSourcePortIdentity: 0x00259e1000000003
    requestingSourcePortId: 3330
```

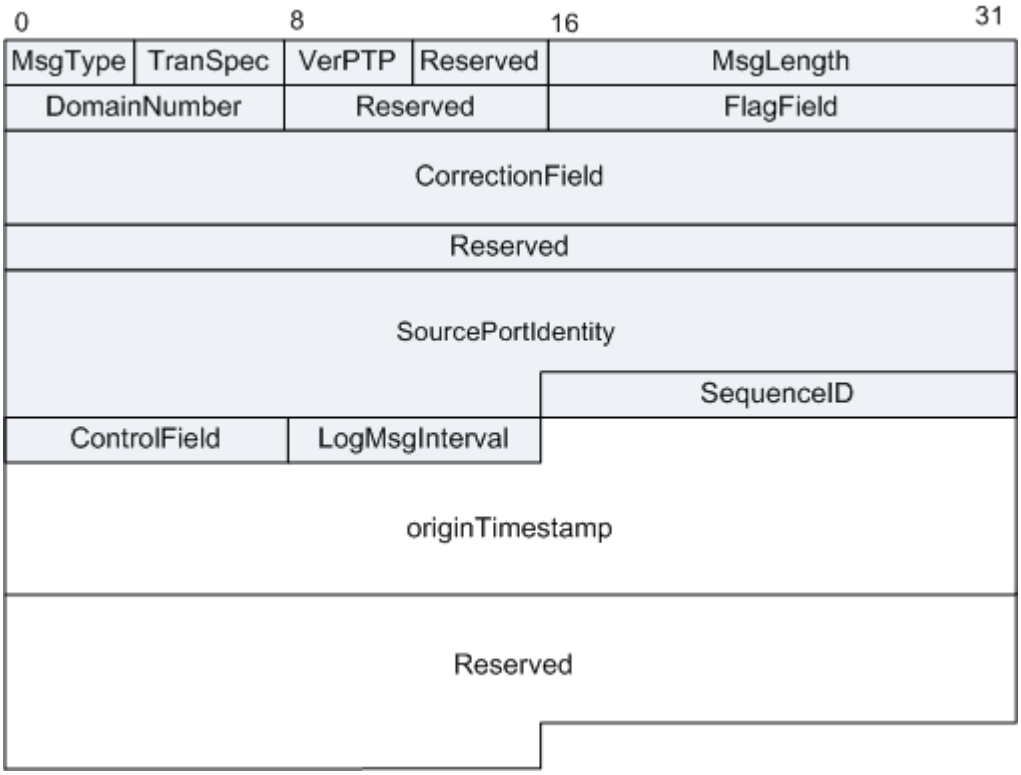
参考标准

标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.1.5 1588v2 Pdelay_Req 消息

报文格式

图 1 Pdelay_Req 消息格式



字段	长度	含义
TranSpec	4 bits	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 bits	表示消息类型。
Reserved	4 bits	保留字段。
VerPTP	4 bits	表示 1588 协议的版本。
MsgLength	2 bytes	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 byte	域编号，表示发送该消息时钟所属的域。
Reserved	1 byte	保留字段。
FlagField	2 bytes	标志域。

字段	长度	含义
CorrectionField	64 bits	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。
Reserved	32 bits	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 bytes	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 byte	控制域，由消息类型决定。
LogMsgInterval	1 byte	录入消息周期，PTP 消息的发送时间间隔。
originTimestamp	10 bytes	源时间戳
Reserved	10 bytes	保留

报文示例

图 2 基于 UDP 的 Pdelay_Req 消息

```
⊞ Frame 65: 100 bytes on wire (800 bits), 100 bytes captured (800 bits)
⊞ Ethernet II, Src: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13), Dst: HuaweiTe_98:f2
⊞ Internet Protocol Version 4, Src: 10.1.1.1 (10.1.1.1), Dst: 224.0.0.107 (224.
⊞ User Datagram Protocol, Src Port: ptp-event (319), Dst Port: ptp-event (319)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = V1 Compatibility: False
    .... 0010 = messageId: Path_Delay_Req Message (0x02)
    .... 0010 = versionPTP: 2
    messageLength: 54
    subdomainNumber: 0
  ⊟ flags: 0x022c
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .0.. .... .... = PTP_UNICAST: False
      .... ..1. .... .... = PTP_TWO_STEP: True
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .1.. = PTP_UTC_REASONABLE: True
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 0.000000 nanoseconds
      correction: Ns: 0 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00259e1000000001
    SourcePortID: 1280
    sequenceId: 18877
    control: Other Message (5)
    logMessagePeriod: 127
    originTimestamp (seconds): 0
    originTimestamp (nanoseconds): 0
```

图 3 基于以太的 Pdelay_Req 消息

```
⊞ Frame 4: 72 bytes on wire (576 bits), 72 bytes captured (576 b
⊟ Ethernet II, Src: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13), Dst:
  ⊞ Destination: HuaweiTe_98:f2:6e (00:25:9e:98:f2:6e)
  ⊞ Source: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13)
    Type: PTPv2 over Ethernet (IEEE1588) (0x88f7)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = 802.1as conform: False
    .... 0010 = messageId: Path_Delay_Req Message (0x02)
    .... 0010 = versionPTP: 2
    messageLength: 54
    subdomainNumber: 0
  ⊟ flags: 0x062c
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .1.. .... .... = PTP_UNICAST: True
      .... ..1. .... .... = PTP_TWO_STEP: True
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .1.. = PTP_UTC_REASONABLE: True
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 0.000000 nanoseconds
      correction: Ns: 0 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00259e1000000001
    SourcePortID: 1280
    sequenceId: 20333
    control: Other Message (5)
    logMessagePeriod: 127
    originTimestamp (seconds): 0
    originTimestamp (nanoseconds): 0
```

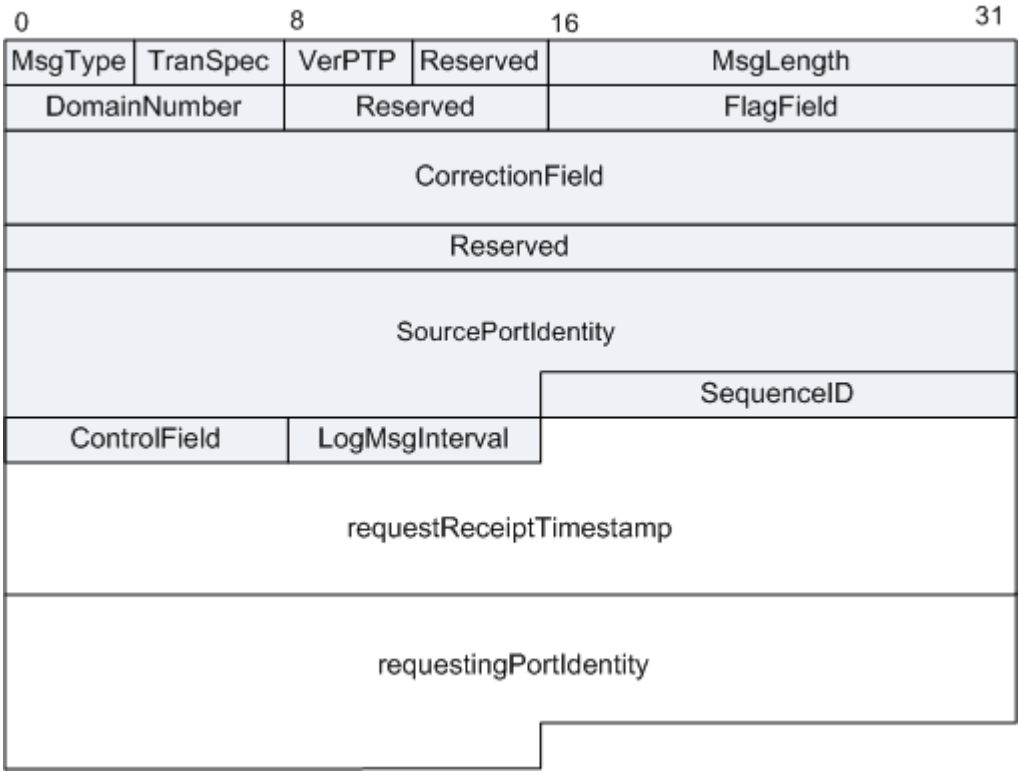
参考标准

标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.1.6 1588v2 Pdelay_Resp 消息

报文格式

图 1 Pdelay_Resp 消息格式



字段	长度	含义
TranSpec	4 bits	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 bits	表示消息类型。

字段	长度	含义
Reserved	4 bits	保留字段。
VerPTP	4 bits	表示 1588 协议的版本。
MsgLength	2 bytes	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 byte	域编号，表示发送该消息时钟所属的域。
Reserved	1 byte	保留字段。
FlagField	2 bytes	标志域。
CorrectionField	64 bits	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。
Reserved	32 bits	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 bytes	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 byte	控制域，由消息类型决定。
LogMsgInterval	1 byte	录入消息周期，PTP 消息的发送时间间隔。
requestReceiptTimestamp	10 bytes	请求接收时间戳。
requestingPortIdentity	10 bytes	请求端口标识。

报文示例

图 2 基于 UDP 的 Pdelay_Resp 消息

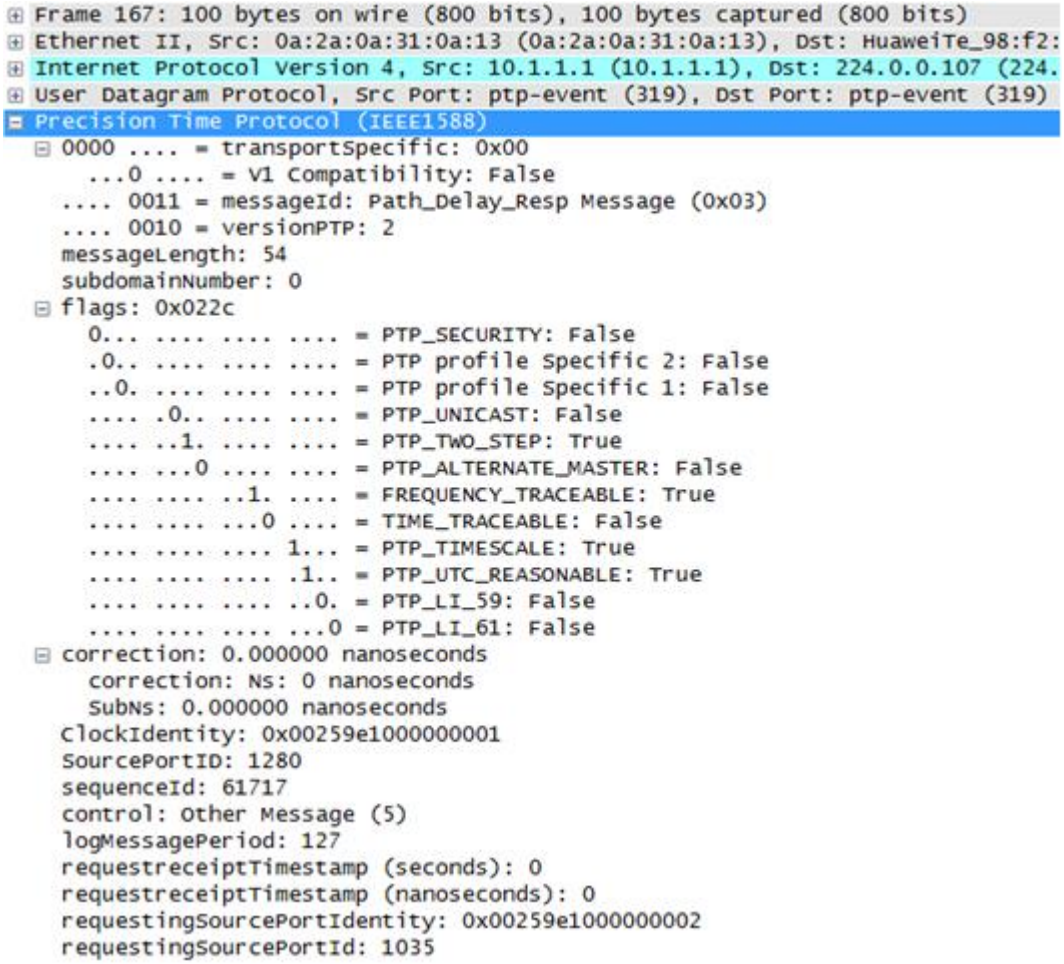


图 3 基于以太层的 Pdelay_Resp 消息

```
⊞ Frame 5: 72 bytes on wire (576 bits), 72 bytes captured (576 bi
⊟ Ethernet II, Src: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13), Dst: H
  ⊞ Destination: HuaweiTe_98:f2:6e (00:25:9e:98:f2:6e)
  ⊞ Source: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13)
    Type: PTPv2 over Ethernet (IEEE1588) (0x88f7)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = 802.1as conform: False
    .... 0011 = messageId: Path_Delay_Resp Message (0x03)
    .... 0010 = versionPTP: 2
    messageLength: 54
    subdomainNumber: 0
  ⊟ flags: 0x062c
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .1.. .... .... = PTP_UNICAST: True
      .... ..1. .... .... = PTP_TWO_STEP: True
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .1.. = PTP_UTC_REASONABLE: True
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 0.000000 nanoseconds
      correction: Ns: 0 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00259e1000000001
    SourcePortID: 1280
    sequenceId: 63173
    control: Other Message (5)
    logMessagePeriod: 127
    requestreceiptTimestamp (seconds): 0
    requestreceiptTimestamp (nanoseconds): 0
    requestingSourcePortIdentity: 0x00259e1000000002
    requestingSourcePortId: 1035
```

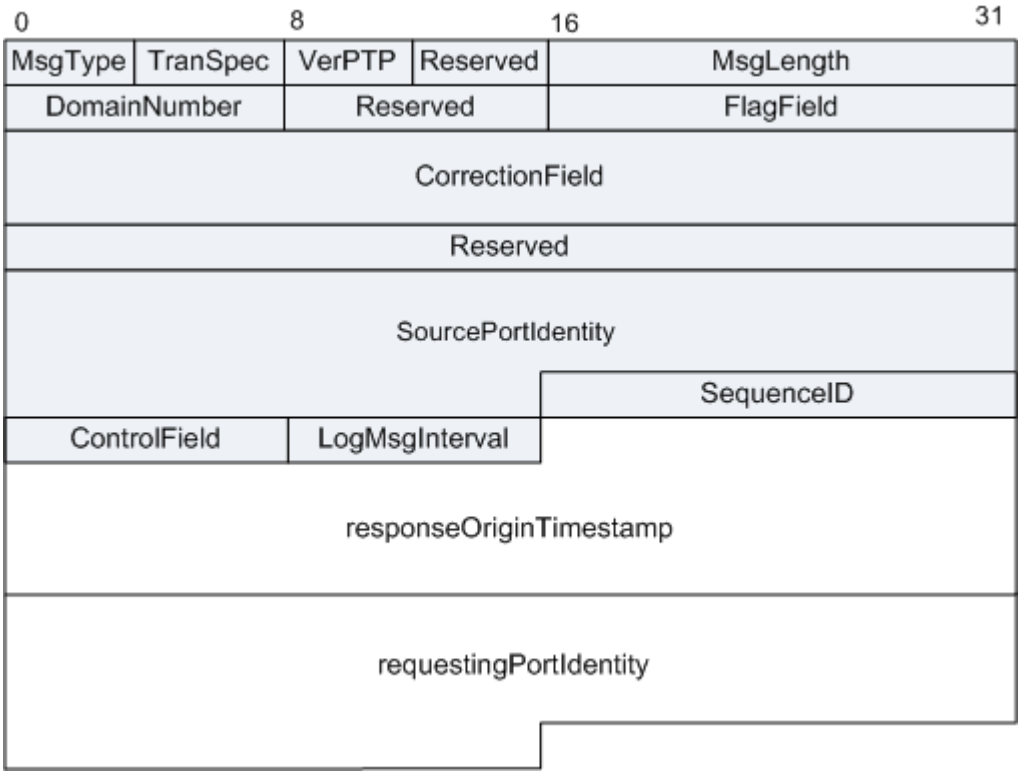
参考标准

标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.1.7 1588v2 Pdelay_Resp_Follow_Up 消息

报文格式

图 1 Pdelay_Resp_Follow_Up 消息格式



字段	长度	含义
TranSpec	4 bits	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 bits	表示消息类型。
Reserved	4 bits	保留字段。
VerPTP	4 bits	表示 1588 协议的版本。
MsgLength	2 bytes	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 byte	域编号，表示发送该消息时钟所属的域。
Reserved	1 byte	保留字段。
FlagField	2 bytes	标志域。

字段	长度	含义
CorrectionField	64 bits	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。
Reserved	32 bits	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 bytes	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 byte	控制域，由消息类型决定。
LogMsgInterval	1 byte	录入消息周期，PTP 消息的发送时间间隔。
responseOriginTimestamp	10 bytes	响应源时间戳
requestingPortIdentity	10 bytes	请求端口标识

报文示例

图 2 基于 UDP 的 Pdelay_Resp_Follow_Up 消息

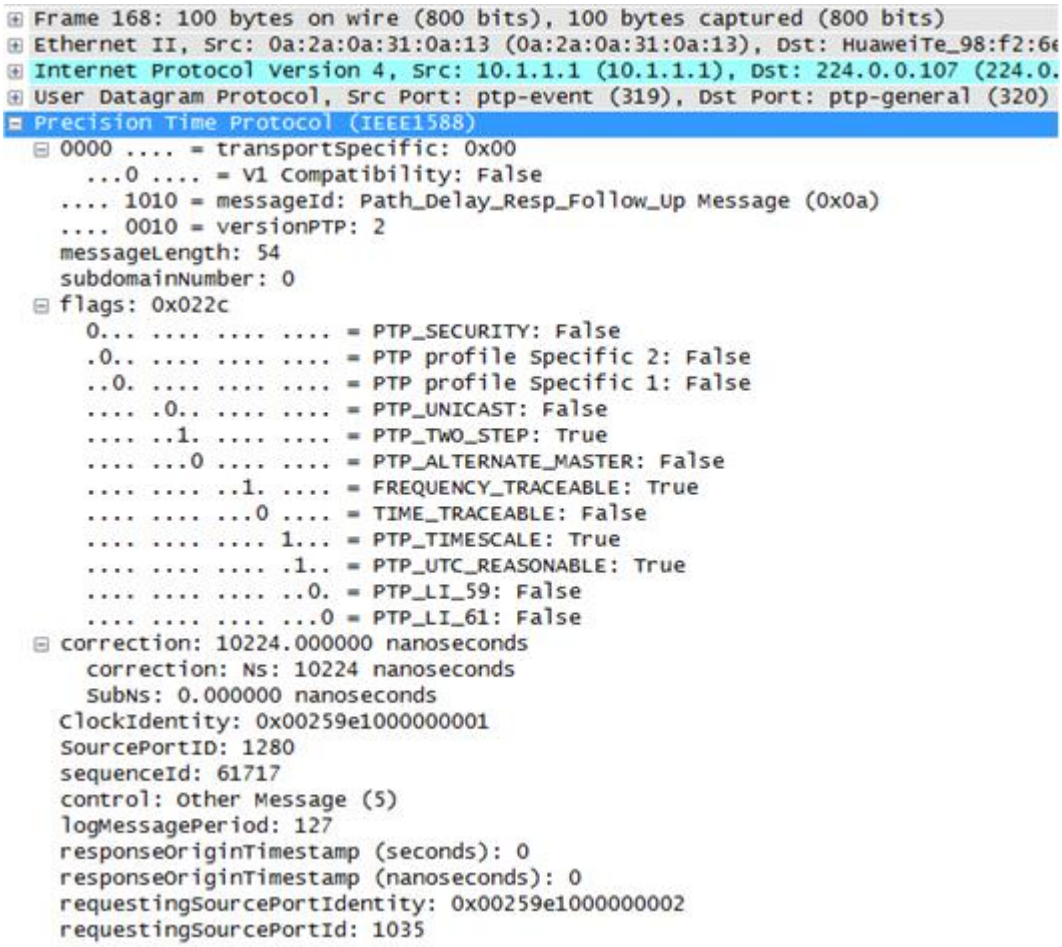


图 3 基于以太的 Pdelay_Resp_Follow_Up 消息

```
⊞ Frame 6: 72 bytes on wire (576 bits), 72 bytes captured (576 bits)
⊟ Ethernet II, Src: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13), Dst: Huawe
  ⊞ Destination: HuaweiTe_98:f2:6e (00:25:9e:98:f2:6e)
  ⊞ Source: 0a:2a:0a:31:0a:13 (0a:2a:0a:31:0a:13)
    Type: PTPv2 over Ethernet (IEEE1588) (0x88f7)
⊟ Precision Time Protocol (IEEE1588)
  ⊟ 0000 .... = transportSpecific: 0x00
      ...0 .... = 802.1as conform: False
    .... 1010 = messageId: Path_Delay_Resp_Follow_Up Message (0x0a)
    .... 0010 = versionPTP: 2
    messageLength: 54
    subdomainNumber: 0
  ⊟ flags: 0x062c
      0... .... .... .... = PTP_SECURITY: False
      .0.. .... .... .... = PTP profile Specific 2: False
      ..0. .... .... .... = PTP profile Specific 1: False
      .... .1.. .... .... = PTP_UNICAST: True
      .... ..1. .... .... = PTP_TWO_STEP: True
      .... ...0 .... .... = PTP_ALTERNATE_MASTER: False
      .... .... ..1. .... = FREQUENCY_TRACEABLE: True
      .... .... ...0 .... = TIME_TRACEABLE: False
      .... .... .... 1... = PTP_TIMESCALE: True
      .... .... .... .1.. = PTP_UTC_REASONABLE: True
      .... .... .... ..0. = PTP_LI_59: False
      .... .... .... ...0 = PTP_LI_61: False
  ⊟ correction: 10430.000000 nanoseconds
      correction: Ns: 10430 nanoseconds
      SubNs: 0.000000 nanoseconds
    ClockIdentity: 0x00259e1000000001
    SourcePortID: 1280
    sequenceId: 63173
    control: Other Message (5)
    logMessagePeriod: 127
    responseOriginTimestamp (seconds): 0
    responseOriginTimestamp (nanoseconds): 0
    requestingSourcePortIdentity: 0x00259e1000000002
    requestingSourcePortId: 1035
```

参考标准

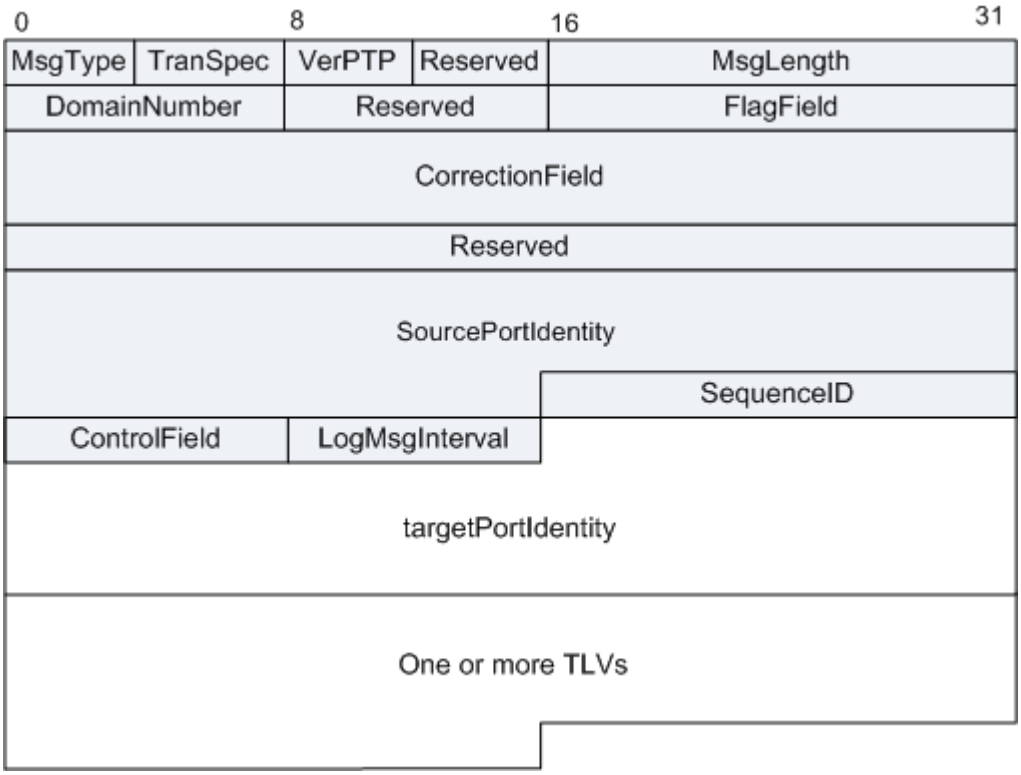
标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.1.8 1588v2 Signaling 消息

Signaling 消息用于传送一个或多个 TLV 实体序列。Signaling 消息从一个时钟传送到一个或多个其它时钟。

报文格式

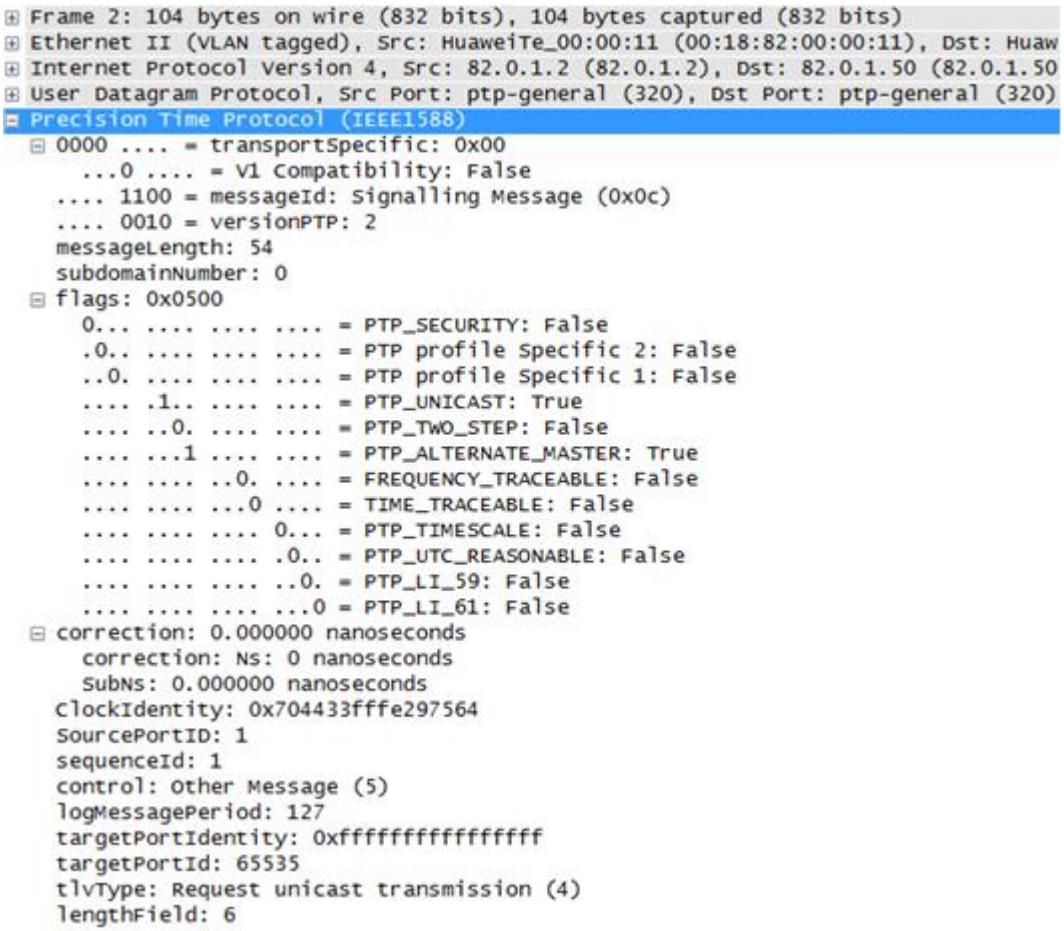
图 1 Signaling 消息格式



字段	长度	含义
TranSpec	4 bits	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 bits	表示消息类型。
Reserved	4 bits	保留字段。
VerPTP	4 bits	表示 1588 协议的版本。
MsgLength	2 bytes	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 byte	域编号，表示发送该消息时钟所属的域。
Reserved	1 byte	保留字段。
FlagField	2 bytes	标志域。

字段	长度	含义
CorrectionField	64 bits	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。
Reserved	32 bits	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 bytes	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 byte	控制域，由消息类型决定。
LogMsgInterval	1 byte	录入消息周期，PTP 消息的发送时间间隔。
targetPortIdentity	10 bytes	目的端口标识。targetPortIdentity 的取值要求为本消息目的地址对应端口的 portIdentity。

报文示例



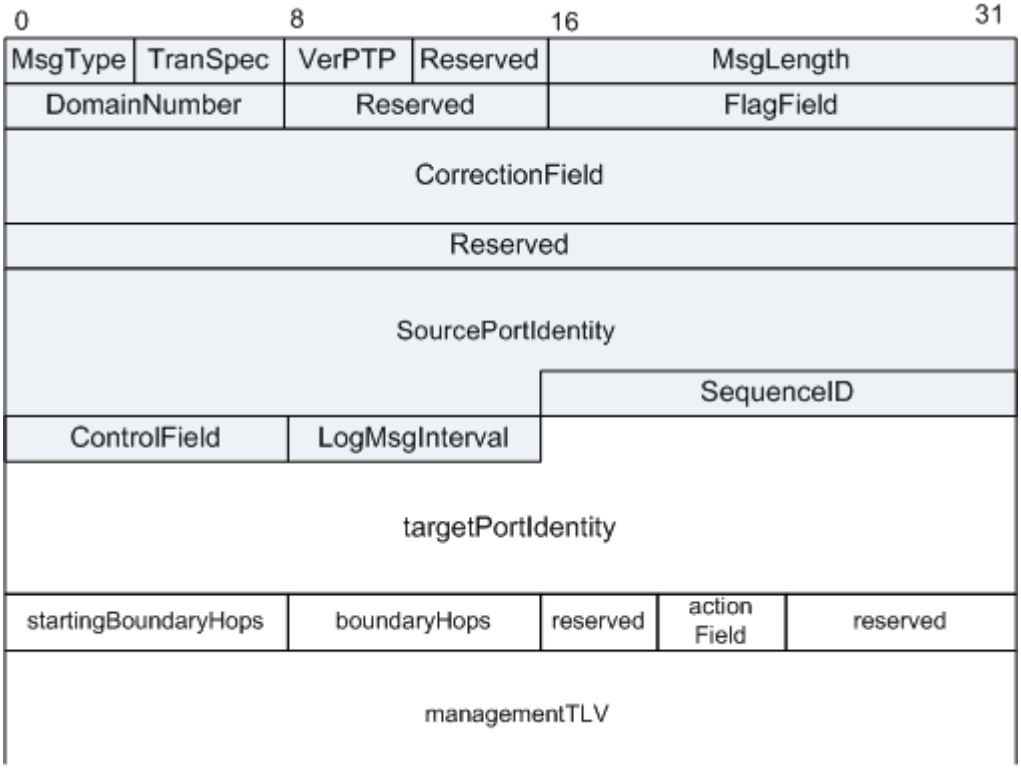
参考标准

标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.1.9 1588v2 Management 消息

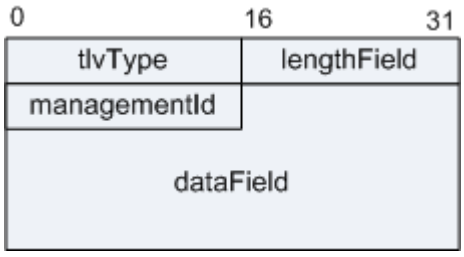
报文格式

图 1 Management 消息格式



字段	长度	含义
TranSpec	4 bits	传送相关。 • 0 表示 PTP 消息由 1588 协议使用 • 1 表示 PTP 消息由 802.1as 协议使用
MsgType	4 bits	消息类型值。
Reserved	4	保留字段。

字段	长度	含义
	bits	
VerPTP	4 bits	表示 1588 协议的版本。
MsgLength	2 bytes	PTP 消息的长度，即 PTP 消息的全部字节数目。计入字节始于报头的第一个字节，同时包含并收尾于任何尾标的最后一个字节，或是无尾标成员时收尾于消息的最后一个字节。
DomainNumber	1 byte	域编号，表示发送该消息时钟所属的域。
Reserved	1 byte	保留字段。
FlagField	2 bytes	标志域。
CorrectionField	64 bits	修正域，各报文都有，主要用在 Sync 报文中，用于补偿网络中的传输时延，E2E 的频率同步。
Reserved	32 bits	保留字段。
SourcePortIdentity		源端口标识符，发送该消息时钟的 ID 和端口号。
SequenceID	2 bytes	序列号 ID，表示消息的序列号，以及关联消息的对应关系。
ControlField	1 byte	控制域，由消息类型决定。
LogMsgInterval	1 byte	录入消息周期，PTP 消息的发送时间间隔。
targetPortIdentity	10 bytes	管理消息产生动作节点或端口的 portIdentity。通过 targetPortIdentity 标识的端口不一定是接收到该管理消息的端口。在时钟发送管理消息到管理者的情况下，targetPortIdentity 字段应该被设置为它所响应的管理消息的 sourcePortIdentity。
startingBoundaryHops	1 byte	用于那些不是用来响应从另外的管理消息请求而发起的消息，startingBoundaryHops 的值应该是从请求消息中的 startingBoundaryHops 和 boundaryHops 字段计算出来的，值为 startingBoundaryHops 减去 boundaryHops。 当接收到管理消息时，这个差异的绝对值指示该消息经过边界时钟重新发送的次数。

字段	长度	含义
boundaryHops	1 byte	指示边界时钟接收消息，该管理剩余的连续重新转发次数。当第一次由初始的时钟发送时，boundaryHops 的值应该和 startingBoundaryHops 字段中的值相同。
reserved	1 byte	预留。
actionField	1 byte	指示对接收到消息将采取的动作。
reserved	1 byte	预留。
managementTLV	M bytes	管理消息要求以 0 或一个 TLV 结尾。 管理 TLV 的格式：  • tlvType (Enumeration16): tlvType 的值要求为 MANAGEMENT。 • lengthField (UInteger16): lengthField 的值是 2+N，这里的 N 是一个偶数。 • managementId (Enumeration16): managementId 的实现特定范围由制造商来分配，定义他们自有设备独有的管理功能。不期望互联互通能力，并且用户必须保证这些 TLV 送往到合适的设备。 ▪ 0000 - 1FFF: 对所有节点适用。 ▪ 2000 - 3FFF: 适用普通时钟和边界时钟。 ▪ 4000 to 5FFF: 适用于透明时钟。 ▪ 6000 - 7FFF: 适用普通，边界和透明时钟。 ▪ C000 - DFFF: 本范围的值用来实现特定的标识。 ▪ E000 - FFFE: 本范围的值由备选的 PTP 模板来分配。 ▪ FFFF: 保留。

参考标准

标准	描述

标准	描述
IEEE 1588 V2	Precision Clock Synchronization Protocol for Networked Measurement and Control Systems
IEEE P1588™ D2.2	Draft Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems

6.2 BFD 控制报文格式

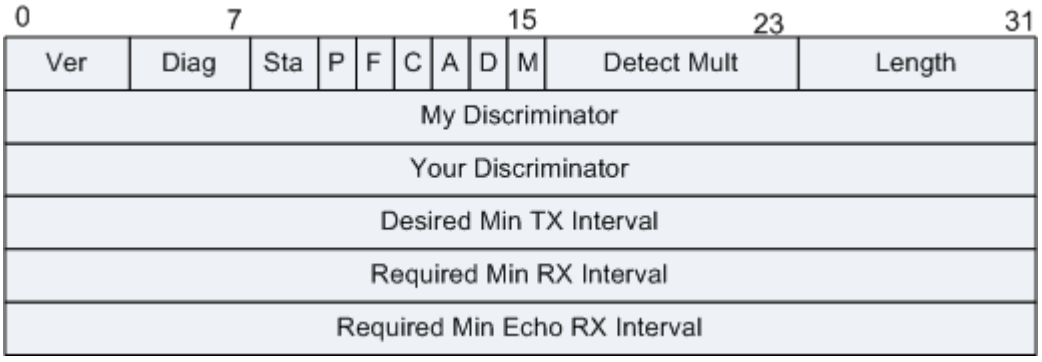
BFD 控制报文封装在 UDP 报文中传送，对于单跳检测其 UDP 目的端口号为 3784，对于多跳检测其 UDP 目的端口号为 4784 或 3784。

BFD 控制报文根据场景不同封装不同。BFD 控制报文包括两部分：强制部分和可选的认证字段。不同的认证类型，认证字段的格式不同。

报文格式

BFD 控制报文强制部分的格式如下（RFC5880）：

图 1 强制部分的格式

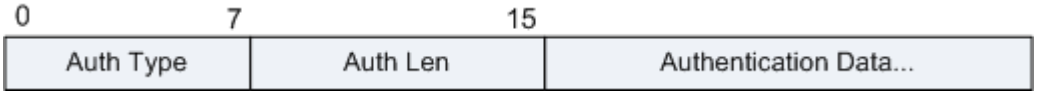


字段	长度	含义
Version (Vers)	3 bits	BFD 协议版本号，目前为 1。
Diagnostic (Diag)	5 bits	诊断字，标明本地 BFD 系统最近一次会话状态发生变化的原因，取值及含义： 0 -- No Diagnostic 1 -- Control Detection Time Expired 2 -- Echo Function Failed 3 -- Neighbor Signaled Session Down 4 -- Forwarding Plane Reset 5 -- Path Down

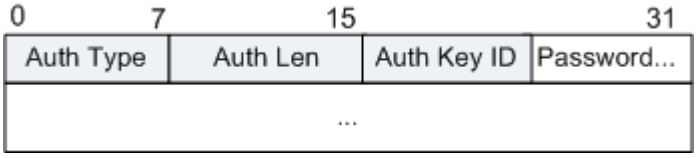
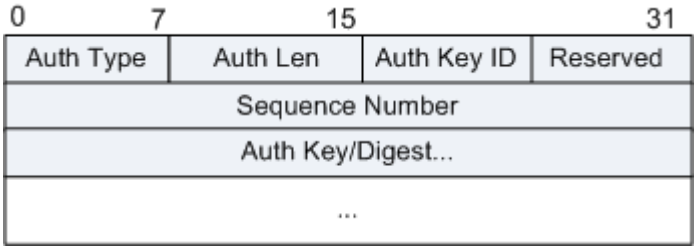
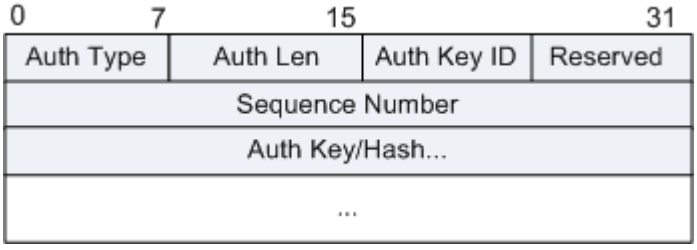
字段	长度	含义
		6 -- Concatenated Path Down 7 -- Administratively Down 8 -- Reverse Concatenated Path Down 9-31 -- Reserved for future use •
State (Sta)	2 bits	BFD 本地状态。 0 -- AdminDown 1 -- Down 2 -- Init 3 -- Up
Poll (P)	1 bit	参数发生改变时，发送方在 BFD 报文中置该标志，接收方必须立即响应该报文。 • 1：表示发送系统请求进行连接确认，或者发送请求参数改变的确认。 • 0：表示发送系统不请求确认。
Final (F)	1 bit	响应 P 标志置位的回应报文中必须将 F 标志置位。 • 1：表示发送系统响应一个接收到 P 比特为 1 的 BFD 包。 • 0：表示发送系统不响应一个 P 比特为 1 的包。
Control Plane Independent (C)	1 bit	转发/控制分离标志，一旦置位，控制平面的变化不影响 BFD 检测，如：控制平面为 IS-IS，当 IS-IS 重启/GR 时，BFD 可以继续监测链路状态。 • 1：表示发送系统的 BFD 实现不依赖于它的控制平面。即，BFD 报文在转发平面传输，即使控制平面失效，BFD 仍然能够起作用。 • 0：表示 BFD 报文在控制平面传输。
Authentication Present (A)	1 bit	认证标识，置 1 代表会话需要进行验证。
Demand (D)	1 bit	查询请求，置位代表发送方期望采用查询模式对链路进行监测。 • 1：表示发送系统希望工作在查询模式。 • 0：表示发送系统不希望、或不能工作在查询模式。
Multipoint (M)	1 bit	为 BFD 将来支持点对多点扩展而设的预留位。
Detect Mult	8 bits	检测超时倍数，用于检测方计算检测超时时间。

字段	长度	含义
		• 查询模式：采用本地检测倍数。 • 异步模式：采用对端检测倍数。
Length	8 bits	报文长度，单位为字节。
My Discriminator	32 bits	BFD 会话连接本地标识符。发送系统产生的一个唯一的、非 0 鉴别值，用来区分一个系统的多个 BFD 会话。
Your Discriminator	32 bits	BFD 会话连接远端标识符。从远端系统接收到的鉴别值，这个域直接返回接收到的“My Discriminator”，如果不知道这个值就返回 0。
Desired Min TX Interval	32 bits	本地支持的最小 BFD 报文发送间隔，单位为微秒。
Required Min RX Interval	32 bits	本地支持的最小 BFD 报文接收间隔，单位为微秒。
Required Min Echo RX Interval	32 bits	本地支持的最小 Echo 报文接收间隔，单位为微秒（如果本地不支持 Echo 功能，则设置 0）。

图 2 BFD 控制报文可选部分的格式



字段	长度	含义
Auth Type	8 bits	BFD 控制报文使用的认证类型。不同值表示的认证类型如下： 0 - Reserved 1 - Simple Password 2 - Keyed MD5 3 - Meticulous Keyed MD5 4 - Keyed SHA1 5 - Meticulous Keyed SHA1 6-255 - Reserved for future use
Auth Len	8 bits	认证字段的长度，包括认证类型与认证长度字段，单位为字节。

字段	长度	含义
Authentication Data	Variable	认证字段净荷。 如果 A 比特位置 1 且认证类型值为 1（Simple Password），则认证字段格式如下：  • **Auth Type:**认证类型，对于简单密码认证，值为 1。 • **Auth Len:**认证字段长度，对于简单密码认证，长度等于密码长度+3。 • **Auth Key ID:**该报文使用的认证秘钥 ID，该值允许同时激活多个秘钥。 • **Password:** 密码值，是个二进制字符串，长度为 1~16 字节。 如果 A 比特位置 1 切认证类型值为 2（Keyed MD5）或 3（Meticulous Keyed MD5），则认证字段格式如下：  • **Auth Type:**认证类型，值为 2 (Keyed MD5)或 3 (Meticulous Keyed MD5)。 • **Auth Len:**认证字段长度，对于 Keyed MD5 和 Meticulous Keyed MD5 认证，长度值为 24。 • **Auth Key ID:**该报文使用的认证秘钥 ID，该值允许同时激活多个秘钥。 • **Reserved:**在传输过程中，该值必须为 0，接收端接收时忽略此字段。 • **Sequence Number:** 报文顺序号。对于 Keyed MD5 认证，该值随机增加，对于 Meticulous Keyed MD5 认证，同一个会话中的报文按顺序逐渐增加。该值用于预防重放攻击。 • **Auth Key/Digest:** 该字段携带 16 字节 MD5 摘要信息。当 MD5 摘要被计算后，该字段填的是 MD5 共享秘钥，并按需尾填充 16 字节的 0。 如果 A 比特位置 1 切认证类型值为 4（Keyed SHA1）或 5（Meticulous Keyed SHA1），则认证字段格式如下：  • **Auth Type:**认证类型，值为 4 (Keyed SHA1)或 5 (Meticulous Keyed SHA1)。 • **Auth Len:**认证字段长度，对于 Keyed SHA1 和 Meticulous Keyed SHA1 认证，长度为 28。

字段	长度	含义
		• **Auth Key ID:**该报文使用的认证秘钥 ID，该值允许同时激活多个秘钥。 • **Reserved:**在传输过程中，该值必须为 0，接收端接收时忽略此字段。 • **Sequence Number:** 报文顺序号。对于 Keyed SHA1 认证，该值随机增加，对于 Meticulous Keyed SHA1 认证，同一个会话中的报文按顺序逐渐增加。该值用于预防重放攻击。 • **Auth Key/Digest:** 该字段携带 20 字节 SHA1 哈希值。当 hash 被计算后，，该字段填的是 SHA1 共享秘钥，并按需尾填充 20 字节的 0。

报文示例

图 3 BFD 报文格式（不认证）

```
BFD Control message
    001. .... = Protocol Version: 1
    ...0 0001 = Diagnostic Code: Control Detection Time Expired (0x01)
    11.. .... = Session State: Up (0x03)
  Message Flags: 0x20 (P)
    1... .. = Poll: Set
    .0.. .. = Final: Not set
    ..0. .. = Control Plane Independent: Not set
    ...0 .. = Authentication Present: Not set
    .... 0. = Demand: Not set
    .... .0 = Multipoint: Not set
  Detect Time Multiplier: 3 (= 3000 ms Detection time)
  Message Length: 24 bytes
  My Discriminator: 0x000001f5
  Your Discriminator: 0x00002093
  Desired Min TX Interval: 1000 ms (1000000 us)
  Required Min RX Interval: 1000 ms (1000000 us)
  Required Min Echo Interval:    0 ms (0 us)
```

图 4 BFD 报文格式（简单认证）

```
Frame 1: 79 bytes on wire (632 bits), 79 bytes captured (632 bits)
Ethernet II, Src: Performa_00:00:02 (00:10:94:00:00:02), Dst: Xerox_00:00:0
Internet Protocol Version 4, Src: 192.85.1.2 (192.85.1.2), Dst: 192.0.0.1 (
User Datagram Protocol, Src Port: 1024 (1024), Dst Port: bfd-control (3784)
BFD Control message
    001. .... = Protocol Version: 1
    ...0 0000 = Diagnostic Code: No Diagnostic (0x00)
    01.. .... = Session State: Down (0x01)
  Message Flags: 0x04 (A)
    0... .. = Poll: Not set
    .0.. .. = Final: Not set
    ..0. .. = Control Plane Independent: Not set
    ...1 .. = Authentication Present: Set
    .... 0. = Demand: Not set
    .... .0 = Multipoint: Not set
  Detect Time Multiplier: 5 (= 5000 ms Detection time)
  Message Length: 33 bytes
  My Discriminator: 0x00000001
  Your Discriminator: 0x00000000
  Desired Min TX Interval: 1000 ms (1000000 us)
  Required Min RX Interval: 1000 ms (1000000 us)
  Required Min Echo Interval:    0 ms (0 us)
  Authentication: Simple Password: secret
    Authentication Type: Simple Password (1)
    Authentication Length: 9 bytes
    Authentication Key ID: 2
    Password: secret

0000  00 00 01 00 00 01 00 10  94 00 00 02 08 00 45 00   ........ ......E.
0010  00 3d 00 00 00 00 0a 11  2f 58 c0 55 01 02 c0 00   .=...... /X.U....
0020  00 01 04 00 0e c8 00 29  72 31 20 44 05 21 00 00   .......) r1 D.!..
0030  00 01 00 00 00 00 00 0f  42 40 00 0f 42 40 00 00   ........ B@..B@..
0040  00 00 01 09 02 73 65 63  72 65 74 4e 0a 90 40      .....sec retN..@
```

图 5 BFD 报文格式（MD5 认证）

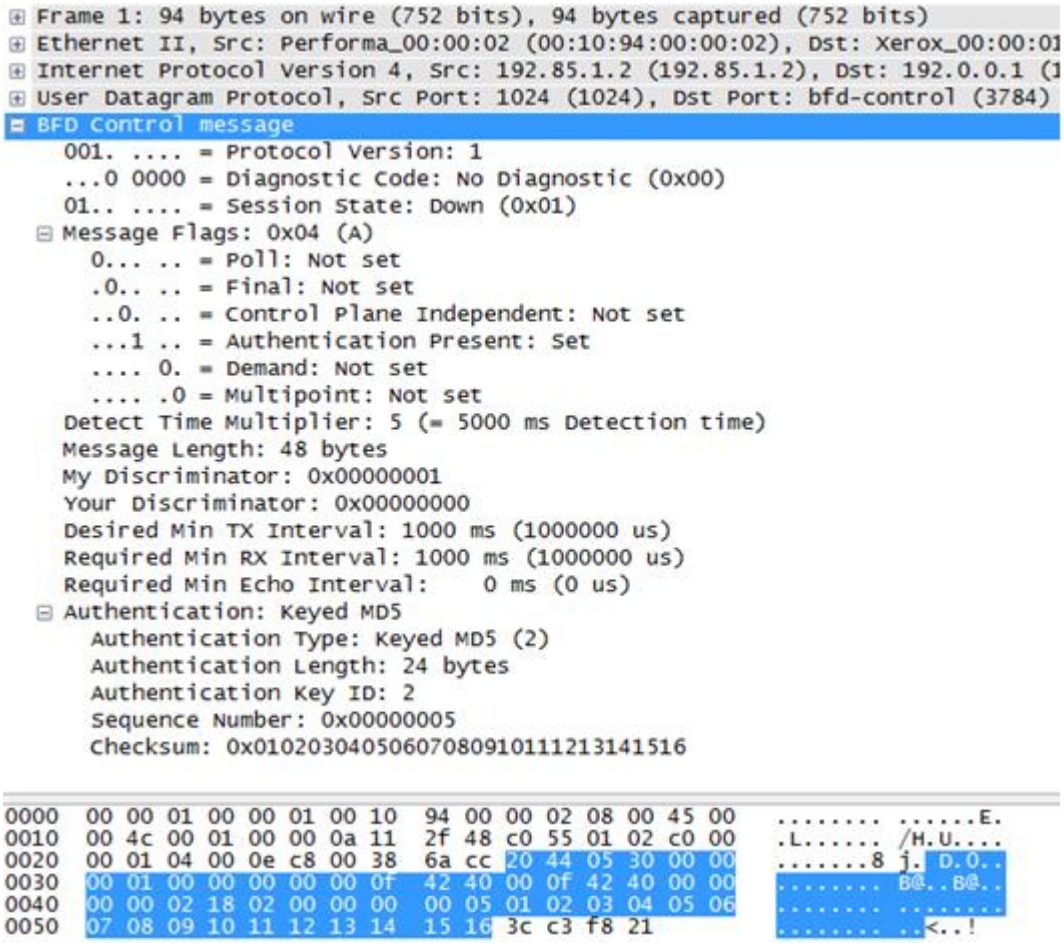
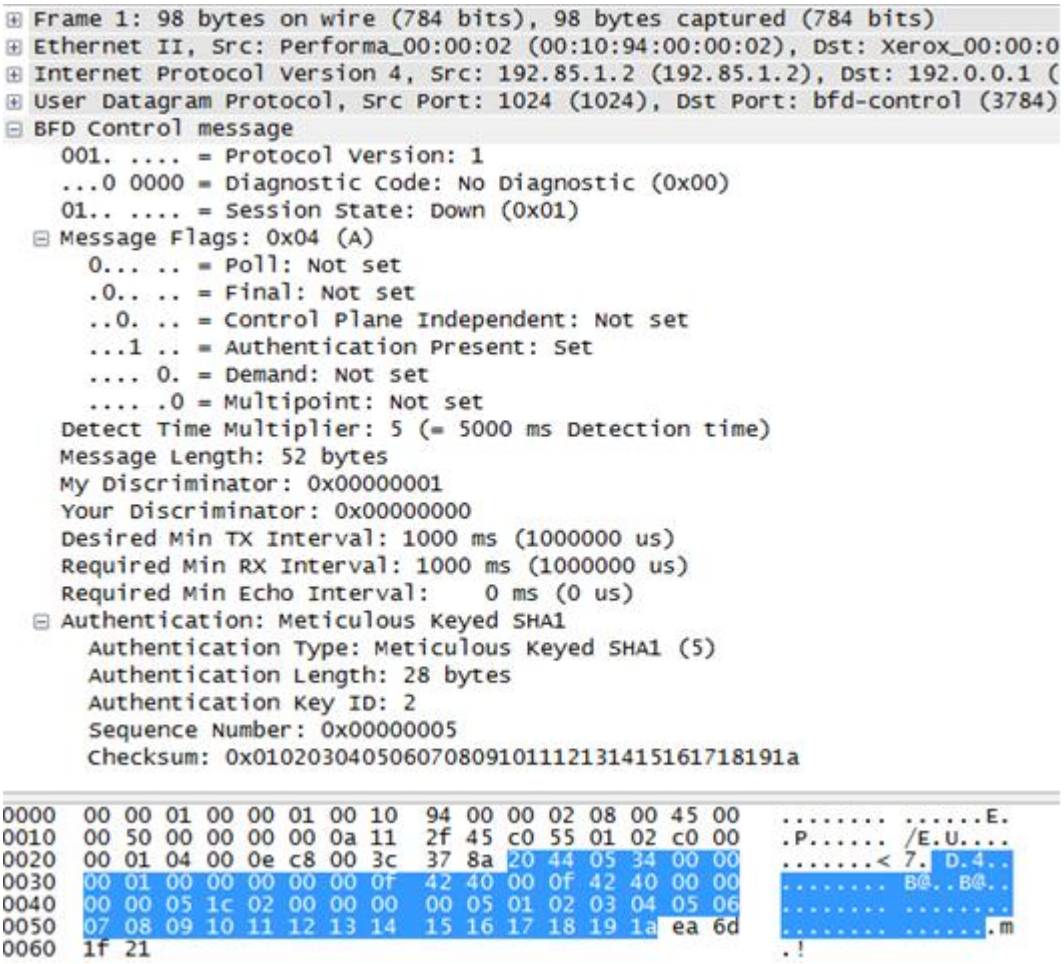


图 6 BFD 报文格式（SHA1 认证）



参考标准

标准	描述
RFC 5880	Bidirectional Forwarding Detection

6.3 BGP 报文格式

- BGP 报文头基本格式（RFC4271）
- BGP OPEN 报文格式
- BGP UPDATE 报文格式
- BGP 的 NOTIFICATION 报文格式
- BGP KEEPALIVE 报文格式
- BGP 的 REFRESH 报文格式

父主题： 应用层

6.3.1 BGP 报文头基本格式（RFC4271）

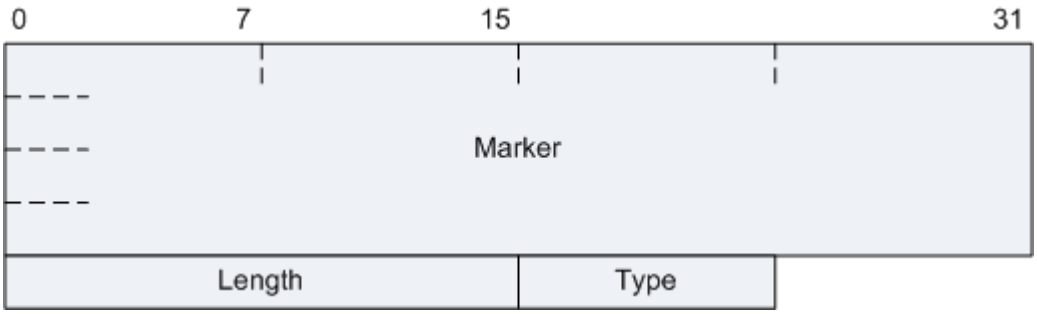
报文格式

BGP 报文由 BGP 报文头和具体报文内容两部分组成。（RFC4271）

BGP 的运行是通过消息驱动的，共有 5 种消息类型，这些消息有相同的报文头。这些消息通过 TCP 协议进行传播（端口号是 179）。消息最长为 4096 字节，最短为 19 字节（只包含报文头）。

BGP 报文头包括三的部分，总长 19 字节。各个部分的格式和功能如下：（RFC4271）

图 1 BGP 报文头格式



- **Marker：** 占 16 字节，用于检查 BGP 对等体的同步信息是否完整，以及用于 BGP 验证的计算。不使用验证时所有比特均为 1（十六进制则全“FF”）。
- **Length：** 占 2 个字节（无符号位），BGP 消息总长度（包括报文头在内），以字节为单位。长度范围是 19～4096。

- **Type**：占 1 个字节（无符号位），BGP 消息的类型。Type 有 5 个可选值，表示 BGP 报文头后面所接的 5 类报文（其中，前四种消息是在 RFC4271 中定义的，而 Type5 的消息则是在 RFC2918 中定义的）：

TYPE 值	报文类型
1	OPEN
2	UPDATE
3	NOTIFICATION
4	KEEPALIVE
5	REFRESH（RFC2918）

参考标准

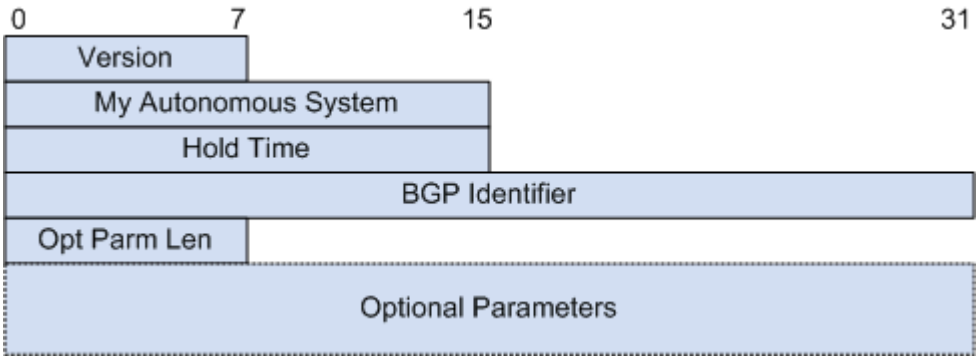
标准	描述
RFC 827	Exterior Gateway Protocol (EGP)
RFC 2918	Route Refresh Capability for BGP-4
RFC 4271	A Border Gateway Protocol 4 (BGP-4)

6.3.2 BGP OPEN 报文格式

报文格式

如果 BGP 报文头中的 TYPE 为 1，则该报文为 OPEN 报文。报文头后面所接的报文内容如下，OPEN 报文用于建立 BGP 连接：

图 1 OPEN 报文格式



字段	长度	含义
Version	1 个字节（无符号位）	表示协议的版本号，现在 BGP 的版本号为 4。
My Autonomous System	2 个字节（无符号位）	发送者自己的 AS 域号
Hold Time	2 个字节（无符号位）	发送者自己设定的 hold time 值（单位：秒），用于协商 BGP 对等体间保持建立连接关系，发送 KEEPALIVE 或 UPDATE 等报文的时间间隔。BGP 的状态机必须在收到对等体的 OPEN 报文后，对发出的 OPEN 报文和收到的 OPEN 报文两者的 hold time 时间作比较，选择较小的时间作为协商结果。Hold Time 的值可为零（不发 KEEPALIVE 报文）或大于等于 3，我们系统的默认为 180。
BGP Identifier	4 个字节（无符号位）	发送者的 router id。
Opt Parm Len	1 个字节（无符号位）	表示 Optional Parameters（可选参数）的长度。如果此值为 0，表示没有可选参数。
Optional Parameters		此值为 BGP 可选参数列表，每一个可选参数是一个 TLV 格式的单元(RFC3392)。 0 1 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-... | Parm. Type | Parm. Length | Parameter Value (variable) +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-... • **Parm. Type**：占 1 个字节（无符号位），为可选参数类型。我们现在的实现中，只在 type 值为 2 时有意义，表示携带的参数为协商能力。

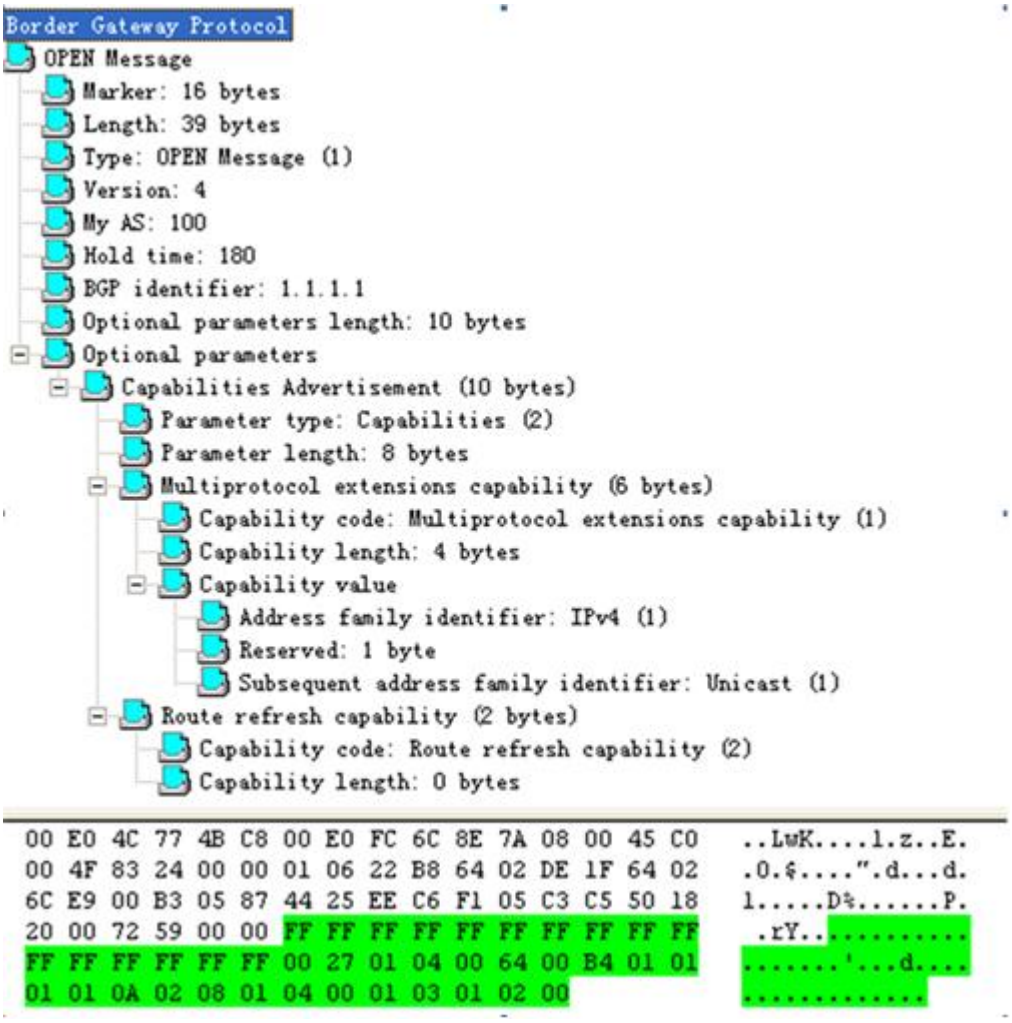
字段	长度	含义
		• **Parm. Length**：占 1 个字节（无符号位），为 Parameter Value 的长度。 • **Parameter. Value**：根据 Parm.Type 的不同值填写不同的参数内容，在 Parm.Type 为 2 表示协商能力时，Parameter.Value 是表示所支持的各种协商能力的列表，列表中的每一个单元是如下的一个 TLV 三元组： +------------------------------+ | Capability Code (1 octet) | +------------------------------+ | Capability Length (1 octet) | +------------------------------+ | Capability Value (variable) | +------------------------------+ ▪ Capability Code：所支持的能力编号，占 1 个字节。Code 为 1 时，表示支持的地址族能力；Code 为 2 时，表示支持 REFRESH 能力。 ▪ Capability Length：表示 Capability Value 的长度，占 1 个字节。 ▪ Capability Value：根据 Code 值的不同其内容与长度也不同。 Capability Code 为 1： Capability Value 值是一个 TLV 三元组，共占 4 个字节： 0 7 15 23 31 +-------+-------+-------+-------+ | AFI | Res. | SAFI | +-------+-------+-------+-------+ **AFI**：地址族标识(Address Family Identifier)，占 2 个字节，能力所支持地址族标识信息，用以和 SAFI 一同确定网络层协议和 IP 地址间的关系，编码方式与多协议扩展中的规定相同。其值按照 RFC1700 中 ADDRESS FAMILY NUMBERS 的定义； **Res**：保留位，占 1 个字节，发送者应将其设置为零，在接受的时候忽略； **SAFI**：子地址族标识(Address Family Identifier)，占 1 个字节，能力所支持的子地址族标识信息，用以和 AFI 一同确定网络层协议和 IP 地址间的关系，编码方式与多协议扩展中的规定相同。其值按照 RFC1700 中 ADDRESS FAMILY NUMBERS 的定义。 Capability Code 为 2（RFC2918） 表示支持路由刷新能力，即 Route Refresh Capability。此能力的 code 为 2，length 为零，无 value 部分。 需要说明的是，只有在能力协商中使能了支持 Route Refresh Capability，路由器才

字段	长度	含义
		能处理 REFRESH 报文。我们的实现是默认情况下，支持 IPv4 单播能力与路由刷新能力，其他能力需要另外设定。

表 1 AFI 及 SAFI 编码说明

AFI 编码	AFI 说明	SAFI 编码	SAFI 说明	说明
1	IPv4 地址族	1	单播	IPv4 单播
		2	组播	IPv4 组播
		128	VPN	IPv4 的 L3VPN
2	IPv6 地址族	1	单播	IPv6 单播
		2	组播	IPv6 组播
		128	VPN	IPv6 的 L3VPN
196	二层	128	VPN	L2VPN 的 Kompella 方式

报文实例



参考标准

标准	描述
RFC 827	Exterior Gateway Protocol (EGP)
RFC 2918	Route Refresh Capability for BGP-4
RFC 4271	A Border Gateway Protocol 4 (BGP-4)

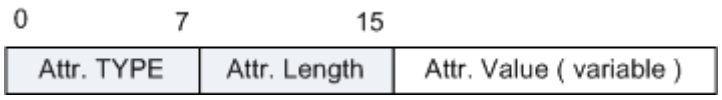
6.3.3 BGP UPDATE 报文格式

如果 BGP 报文头中的 TYPE 为 2，则该报文为 UPDATE 报文。报文头后面所接的报文内容如下（RFC 4271），UPDATE 报文用于通告路由。

报文格式

图 1 UPDATE 报文格式

Unfeasible Routes Length (2 octets)
Withdrawn Routes (variable)
Total Path Attribute Length (2 octets)
Path Attributes (variable)
Network Layer Reachability Information (variable)

字段	长度	含义
Withdrawn Routes Length	2 个字节（无符号位）	标明 Withdrawn Routes 部分的长度。其值为零时，表示没有撤销的路由。
Withdrawn Routes	变长	包含要撤销的路由列表，列表中的每个单元包含 1 字节的 Length 域和可变长度的 Prefix 域。 • **Length**：待撤销路由的掩码。其值为零时，表示匹配所有的路由。 • **Prefix**：传送的 IP 地址前缀必须用整字节表示。例如：假定待撤销的路由为 200.200.200.200，其编码用 16 进制表示可如下： Mask 掩码(十进制) Length Prefix 32 20 C8 C8 C8 C8 25 19 C8 C8 C8 80 20 14 C8 C8 C0 15 0F C8 C8
Total Path Attribute Length	2 个字节（无符号位）	标明 Path Attributes 部分和 Network Layer Reachability Information 两部分的长度。其值为零时，表示没有路由及其路由属性要通告。
Path Attributes	变长	包含要更新的路由属性列表，按其类型号从小到大的顺序排序，填写更新的路由的所有属性。每一个属性单元包括属性类型，属性长度，属性值三部分。其编码采用 TLV 格式。如下所示。 **图 2** BGP 路径属性 TLV 格式  其中，Attr.TYPE 占 2 个字节（无符号位），包括 1 字节的 Flags（无符号位）和 1 字节的 Type Code（无符号位）。 **图 3** TLV 结构－Type

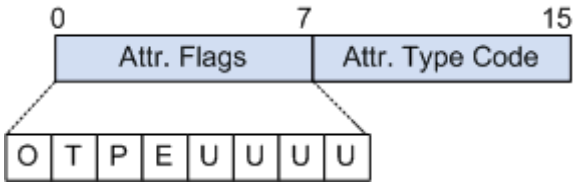
字段	长度	含义
		 **Attr.Flags**：占 1 个字节（8 个 bit），表示属性的标记，其每个 bit 位的意义如下显示： O: Optional bit, 属性的可选性。决定属性是否为必携带属性。带可选属性（optional）设为 1，公认属性（well-known）设为零。 T: Transitive bit 属性的可传递性。对于可选属性，是可传递的设为 1，非可传递的设为 0。对于公认属性必须设为 1。 P: Partial bit 属性的局部性。对于可传递的可选属性是局部的设为 1，是完全的设为零。对于非可传递的的可选属性和公认属性，必须设为零。 E: Extended Length bit 决定该属性的长度的字段（即 Attr. Length）是否需要扩展。不需要扩展则设为零，Attr. Length 占 1 个字节；需要扩展则设为 1，Attr. Length 占 2 个字节。 U: Unused bits 低 4 位没有使用，发送时必须全部设为零，并且在接收时被忽略。 **Attr.Type Code**：占 1 个字节（无符号位），表示属性的类型号。设置如下表 2。 **Attr.Value**：根据不同属性的类型填写不同内容。
Network Layer Reachability Information （NLRI）	变长	包含要更新的地址前缀列表,每一个地址前缀单元由一个LV二元组(prefix length, the prefix of the reachable route）组成，其编码填写方法与 Withdrawn Routes 的填写方法相同。

表 1 路由属性的类型号列表

属性类型	属性值
1：Origin	IGP
	EGP
	Incomplete
2：As_Path	AS_SET

字段	长度	含义
		AS_SEQUENCE
		AS_CONFED_SET
		AS_CONFED_SEQUENCE
	3：Next_Hop	下一跳的 IP 地址
	4：Multi_Exit_Disc	MED 用于判断流量进入 AS 时的最佳路由
	5：Local_Pref	Local_Pref 用于判断流量离开 AS 时的最佳路由
	6：Atomic_Aggregate	BGP Speaker 选择聚合后的路由，而非具体的路由
	7：Aggregator	发起聚合的路由器 ID 和 AS 号
	8：Community	团体属性
	9：Originator_ID	反射路由发起者的 Router ID
	10：Cluster_List	反射路由经过的反射器列表
	14：MP_REACH_NLRI	多协议可达 NLRI
	15：MP_UNREACH_NLRI	多协议不可达 NLRI
	16：Extended Communtities	扩展团体属性

报文示例

图 4 withdrawn 路由的更新报文

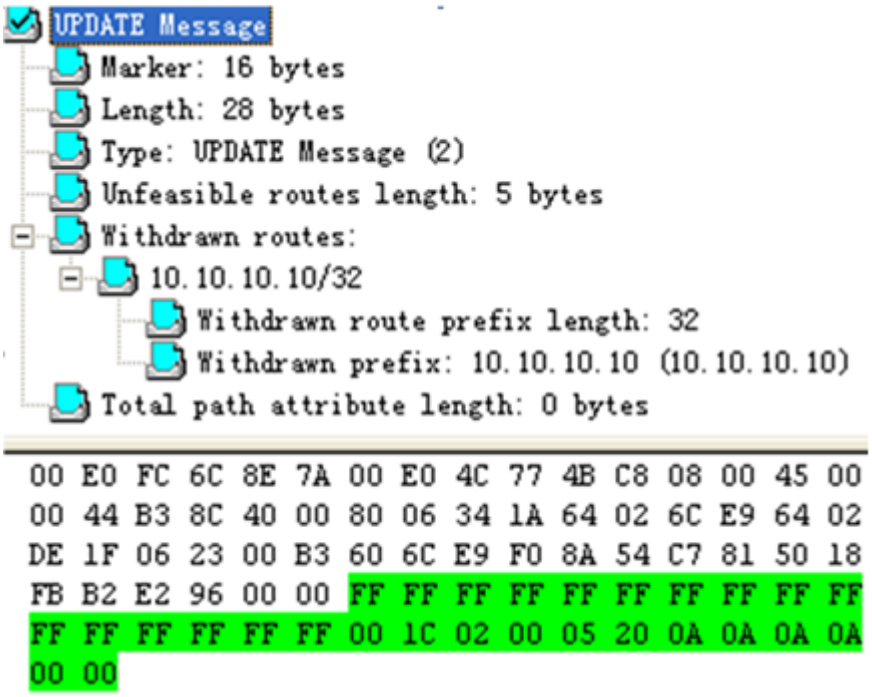
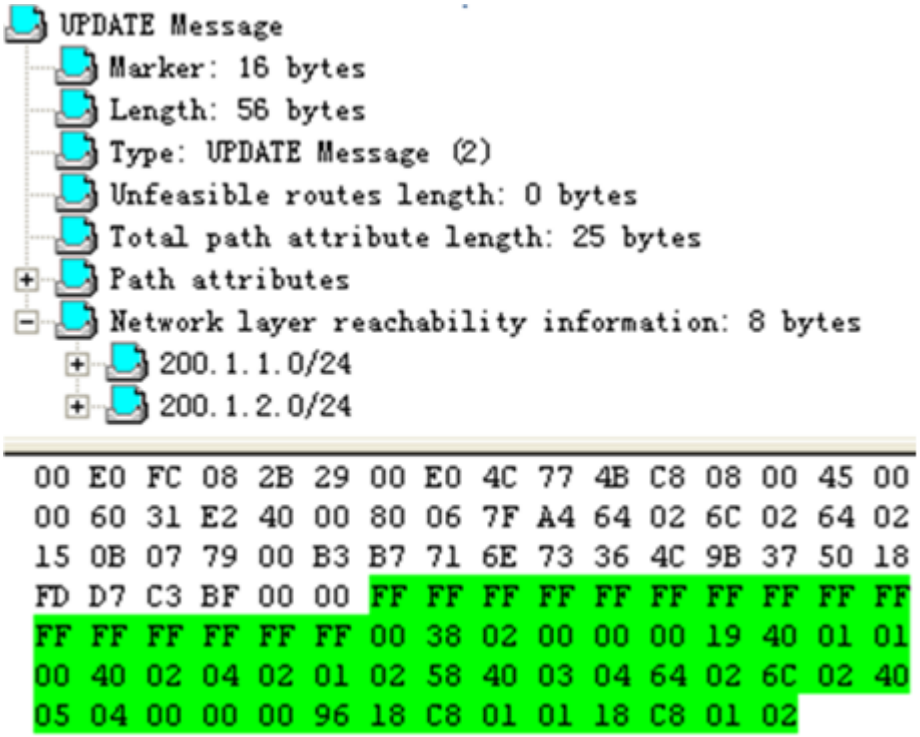


图 5 添加路由的更新报文



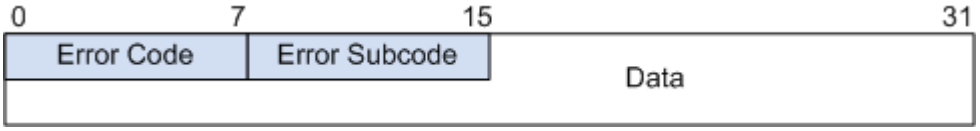
参考标准

标准	描述
RFC 827	Exterior Gateway Protocol (EGP)
RFC 2918	Route Refresh Capability for BGP-4
RFC 4271	A Border Gateway Protocol 4 (BGP-4)

6.3.4 BGP 的 NOTIFICATION 报文格式

如果 BGP 报文头中的 TYPE 为 3，则该报文为 NOTIFICATION 报文。报文头后面所接的报文内容如下（RFC 4271），NOTIFICATION 报文用于处理 BGP 进程中的各种错误。

图 1 NOTIFICATION 报文格式



各字段解释如下：

- **Error code**：占 1 个字节（无符号位），定义错误的类型，非特定的错误类型用零表示。
- **Error subcode**：占 1 个字节（无符号位），指定错误细节编号，非特定的错误细节编号用零表示。
- **Data**：指定错误数据内容。

错误码	错误子码
1：消息头错误	1：连接未同步
	2：错误的消息长度
	3：错误的消息类型
2：Open 消息错误	1：不支持的版本号
	2：错误的对等 AS
	3：错误的 BGP 标识符
	4：不支持的可选参数
	5：认证失败
	6：不可接受的保持时间

错误码	错误子码
	7：不支持的能力
3：Update 消息错误	1：畸形属性列表
	2：不可识别的公认属性
	3：缺少公认属性
	4：属性标志错误
	5：属性长度错误
	6：无效 Origin 属性
	7：AS 路由环路
	8：无效 Next_Hop 属性
	9：可选属性错误
	10：无效网络字段
	11：畸形 AS_Path
4：Hold Timer 溢出	0：没有特别的错误子码定义。
5：有限状态机错误	0：没有特别的错误子码定义。
6：终止	1：前缀超过最大值。

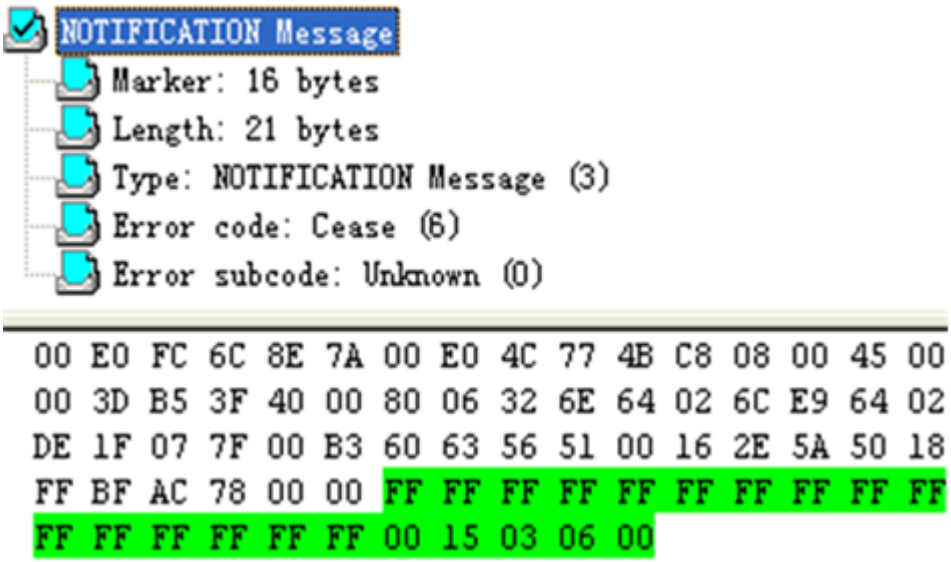
错误码	错误子码
	2：管理关闭
	3：删除邻居
	4：管理重置
	5：连接失败
	6：其他配置改变
	7：连接冲突
	8：资源短缺
	9：BFD 断开连接

Error Code	Error Subcode
1: Message header error	1: connection not synchronized
	2: error message length
	3: error message type
2: Open message error	1: unsupported version number
	2: error peer AS
	3: error BGP identifier

错误码	错误子码
	4: unsupported optional parameter
	5: authentication failed
	6: unacceptable Holdtime
	7: unsupported capability
3: Update message error	1: malformed attribute list
	2: unrecognized well-known attribute
	3: well-known attribute is missing
	4: attribute flags error
	5: attribute length error
	6: invalid origin attribute
	7: AS routing loop
	8: invalid Next-Hop attribute
	9: error optional attribute
	10: invalid network field
	11: abnormal AS-Path

错误码	错误子码
4: Hold timer expired	0: no special definition of the error subcode
5: Finite state machine error	0: no special definition of the error subcode
6: Cease	1: maximum number of prefixes reached
	2: administrative shutdown
	3: peer de-configured
	4: administrative reset
	5: connection rejected
	6: other configuration change
	7: connection collision resolution
	8: out of resources
	9: BFD session Down

报文实例



参考标准

标准	描述
RFC 827	Exterior Gateway Protocol (EGP)
RFC 2918	Route Refresh Capability for BGP-4
RFC 4271	A Border Gateway Protocol 4 (BGP-4)

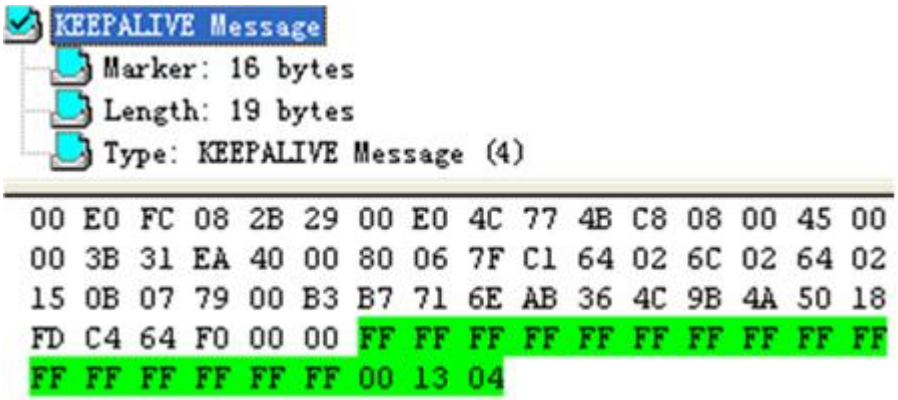
6.3.5 BGP KEEPALIVE 报文格式

报文格式

如果 BGP 报文头中的 TYPE 为 4，则该报文为 KEEPALIVE 报文。KEEPALIVE 报文用于保持 BGP 连接。

KEEPALIVE 报文只有 BGP 报文头，没有具体内容，故其报文长度应固定为 19 个字节。

报文实例



参考标准

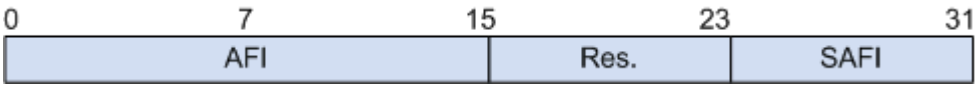
标准	描述
RFC 827	Exterior Gateway Protocol (EGP)
RFC 2918	Route Refresh Capability for BGP-4
RFC 4271	A Border Gateway Protocol 4 (BGP-4)

6.3.6 BGP 的 REFRESH 报文格式

如果 BGP 报文头中的 TYPE 为 5，则该报文为 REFRESH 报文。报文头后面所接的报文内容如下（RFC 2918），REFRESH 报文用于动态的请求 BGP 路由发布者重新发布 UPDATE 报文，进行路由更新。

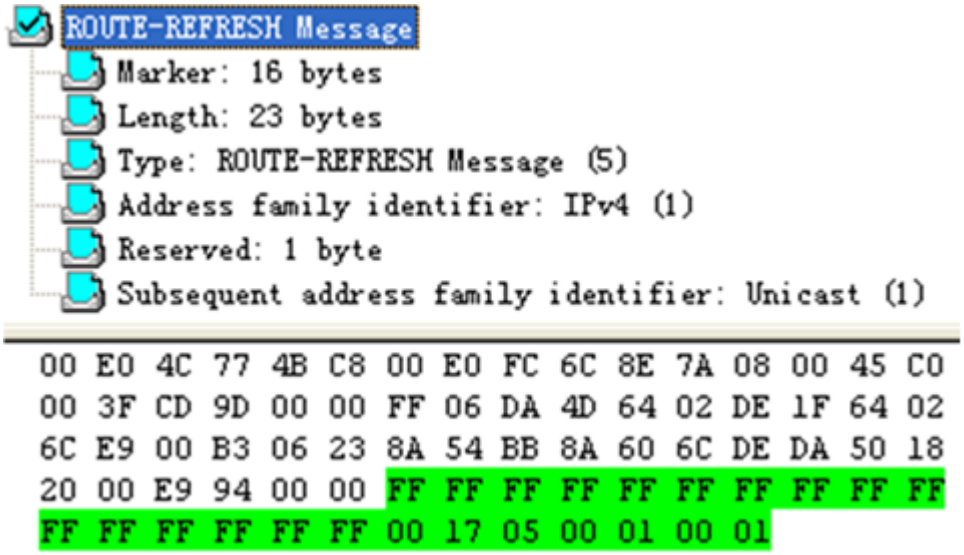
报文格式

图 1 REFRESH 报文格式



Field 字段	Length 长度	Description 描述
AFI	2 字节（无符号位）	表示地址族 id，与 UPDATE 报文中的定义相同。
Res.	1 字节（无符号位）	所有为应全为零，在接收报文时，此位被忽略。
SAFI	1 字节（无符号位）	与 UPDATE 报文中的定义相同。

报文实例



参考标准

标准	描述
RFC 827	Exterior Gateway Protocol (EGP)
RFC 2918	Route Refresh Capability for BGP-4
RFC 4271	A Border Gateway Protocol 4 (BGP-4)

6.4 BOOTP 报文格式

BOOTP (Boot Protocol)是一种 IP/UDP 引导协议，可以使一个无盘客户端获取自己的 IP 地址、服务器的主机地址和一个需要放在内存中运行的指定名称的引导文件。

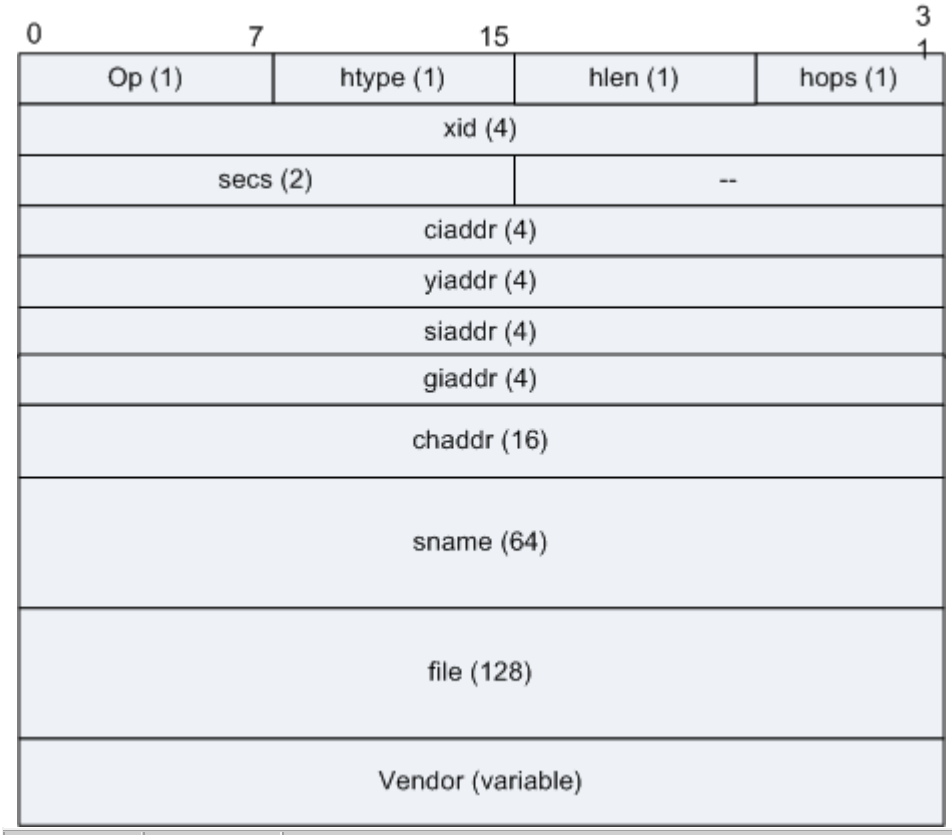
我们希望客户端在启动时可以有一种无须用户参与的完整的引导方式，即一种无人值守的上电启动方式，幸运的是，BOOTP 帮我们做到了这一点。通过 BOOTP 协议，客户端可以自动获得 IP 地址、服务器地址和引导文件等，从而完成地址绑定和引导启动。

文件传送一般使用 TFTP 协议，但 BOOTP 也能够与其它协议如 SFTP 或 FTP 一起工作。

报文格式

BOOTP 协议报文被封装在 UDP 中。

图 1 BOOTP 报文格式



字段	长度（字节）	含义
op	1	操作码/消息类型，取值为 1 或 2： • 1 = BOOTREQUEST（引导请求） • 2 = BOOTREPLY（引导应答）
htype	1	Hardware address type，硬件地址类型，如为 1 时表示客户端的网络硬件是 10M 以太网类型。
hlen	1	Hardware address length，硬件地址长度，如为 6 时表示客户端的网络硬件地址长度为 6byte。
hops	1	客户端设置成 0，在跨越网关引导时网关可选择使用（加 1）。
xid	4	事务 ID，一个随机数，用来匹配引用请求和应答。
secs	2	客户端引导开始后的过去的秒数，由客户端填写。
--	2	未使用
ciaddr	4	Client IP address，客户端 IP 地址，如果客户端知道就在引导请求中填入。
yiaddr	4	Your (client) IP address，你的（客户端）IP 地址，如果客户端不知道它的地址（ciaddr 是 0），服务器填入。
siaddr	4	Server IP address，服务器 IP 地址由服务器在引导应答返回。
giaddr	4	Gateway IP address，网关 IP 地址，在跨越网关引导中可以选择使用。

字段	长度（字节）	含义
chaddr	16	Client hardware address，客户端硬件地址，由客户端填写。
sname	64	Server host name，服务器主机名。可选的，如果填写，必须为一个以 0 结尾的字符串。
file	128	Boot file name，引导文件名，可选的，如果填写，必须为一个以 0 结尾的字符串。
vendor	64	Vendor-specific area，可选的商家指定的区域。可以是请求硬件类型/序列、或应答的性能/远端文件系统配置。这些信息留给第三方分析引导或程序使用。 注：vendor 字段，对于 DHCP，又称为“option”字段。此字段采用“CLV”方式构成，即 code：标识号，唯一标识后面的信息内容，占 1bytes；length：长度，表示后面信息内容的长度，占 1bytes；value：信息内容，其长度为 length 所指定，以 bytes 为单位。

BOOTP 报文示例

图 2 BOOTP 请求报文

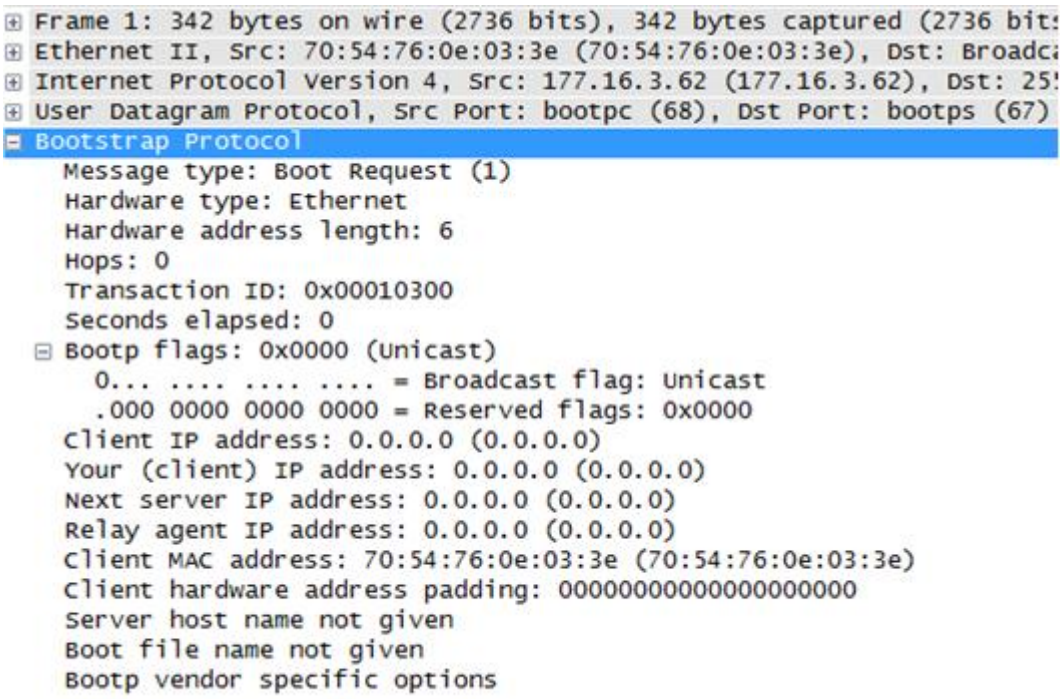


图 3 BOOTP 响应报文

```
⊞ Frame 1: 342 bytes on wire (2736 bits), 342 bytes captured (2736 bit
⊞ Ethernet II, Src: 70:54:76:0e:03:40 (70:54:76:0e:03:40), Dst: 70:54:
⊞ Internet Protocol Version 4, Src: 177.16.3.64 (177.16.3.64), Dst: 17
⊞ User Datagram Protocol, Src Port: bootps (67), Dst Port: bootpc (68)
⊟ Bootstrap Protocol
    Message type: Boot Reply (2)
    Hardware type: Unknown (0xef)
    Hardware address length: 176
    Hops: 56
    Transaction ID: 0x00010300
    Seconds elapsed: 27802
  ⊟ Bootp flags: 0xae38 (Broadcast)
      1... .... .... .... = Broadcast flag: Broadcast
      .010 1110 0011 1000 = Reserved flags: 0x2e38
    Client IP address: 181.127.0.0 (181.127.0.0)
    Your (client) IP address: 10.16.0.1 (10.16.0.1)
    Next server IP address: 10.16.0.2 (10.16.0.2)
    Relay agent IP address: 1.0.0.0 (1.0.0.0)
    Client address not given
    Server host name not given
    Boot file name not given
    Bootp vendor specific options
```

参考标准

标准	描述
RFC 951	Bootstrap Protocol (BOOTP)
RFC 1542	Clarifications and Extensions for the Bootstrap Protocol

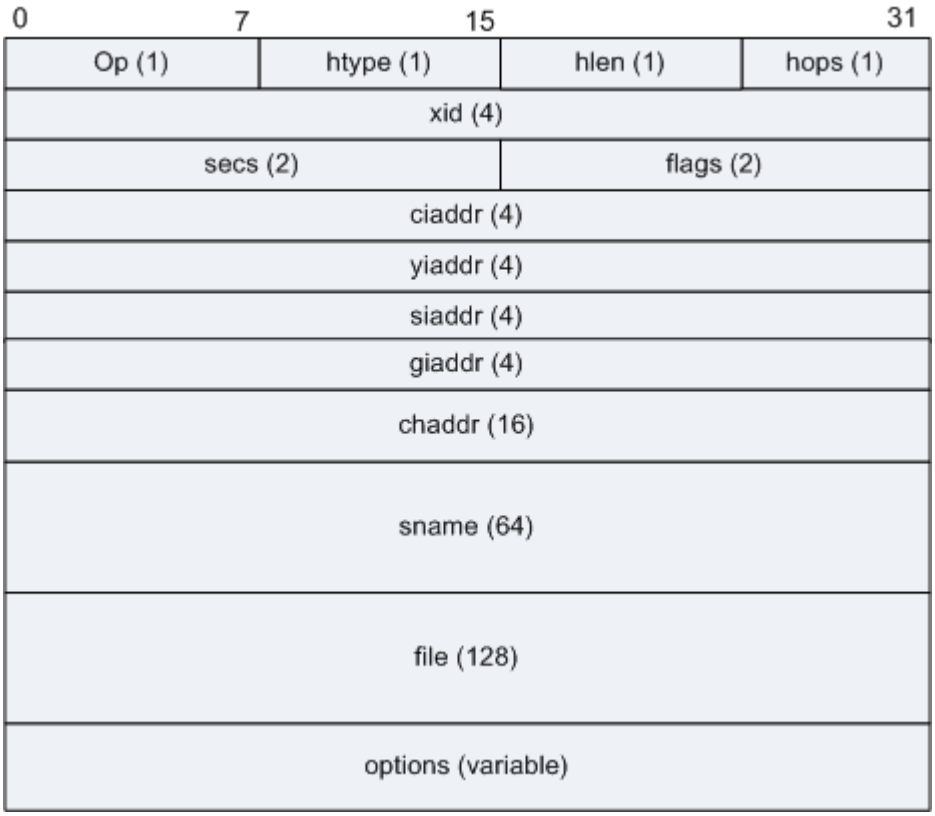
6.5 DHCP 报文格式

报文格式

DHCP 报文是承载于 UDP 上的高层协议报文，采用 67（DHCP 服务器）和 68（DHCP 客户端）两个端口号。

DHCP 的报文格式如下图所示。

图 1 DHCP 报文格式



DHCP 报文中各字段的含义：

字段	长度	含义
OP	1 字节	表示报文的类型： • 1：客户端请求报文 • 2：服务器响应报文
htype	1 字节	表示硬件地址的类型。对于以太网，该类型的值为“1”。
hlen	1 字节	表示硬件地址的长度，单位是字节。对于以太网，该值为 6。
Hops	1 字节	跳数。客户端设置为 0，也能被一个代理服务器设置。
xid	4 字节	事务 ID，由客户端选择的一个随机数，被服务器和客户端用来在它们之间交流请求和响应，客户端用它对请求和应答进行匹配。该 ID 由客户端设置并由服务器返回，为 32 位整数。
secs	2 字节	由客户端填充，表示从客户端开始获得 IP 地址或 IP 地址续借后所使用了的秒数。
flags	2 字节	此字段在 BOOTP 中保留未用，在 DHCP 中表示标志字段。 **图 2** Flags 字段格式  只有标志字段的最高位才有意义，其余的位均被置为 0。 最左边的字段被解释为广播响应标志位，内容如下所示：

字段	长度	含义
		• 0：客户端请求服务器以单播形式发送响应报文 • 1：客户端请求服务器以广播形式发送响应报文
ciaddr	4 字节	客户端的 IP 地址。只有客户端是 Bound、Renew、Rebinding 状态，并且能响应 ARP 请求时，才能被填充。
yiaddr	4 字节	"你自己的"或客户端的 IP 地址。
siaddr	4 字节	表明 DHCP 协议流程的下一个阶段要使用的服务器的 IP 地址。
giaddr	4 字节	该字段表示第一个 DHCP 中继的 IP 地址（注意：不是地址池中定义的网关）。当客户端发出 DHCP 请求时，如果服务器和客户端不在同一个网络中，那么第一个 DHCP 中继在转发这个 DHCP 请求报文时会把自己的 IP 地址填入此字段。服务器会根据此字段来判断出网段地址，从而选择为用户分配地址的地址池。服务器还会根据此地址将响应报文发送给此 DHCP 中继，再由 DHCP 中继将此报文转发给客户端。 若在到达 DHCP 服务器前经过了不止一个 DHCP 中继，那么第一个 DHCP 中继后的中继不会改变此字段，只是把 Hops 的数目加 1。
chaddr	16 字节	该字段表示客户端的 MAC 地址，此字段与前面的“Hardware Type”和“Hardware Length”保持一致。当客户端发出 DHCP 请求时，将自己的硬件地址填入此字段。对于以太网，当“Hardware Type”和“Hardware Length”分别为“1”和“6”时，此字段必须填入 6 字节的以太网 MAC 地址。
sname	64 字节	该字段表示客户端获取配置信息的服务器名字。此字段由 DHCP Server 填写，是可选的。如果填写，必须是一个以 0 结尾的字符串。
file	128 字节	该字段表示客户端的启动配置文件名。此字段由 DHCP Server 填写，是可选的，如果填写，必须是一个以 0 结尾的字符串。
options	可变	该字段表示 DHCP 的选项字段，至少为 312 字节，格式为"代码+长度+数据"。DHCP 通过此字段包含了服务器分配给终端的配置信息，如网关 IP 地址，DNS 服务器的 IP 地址，客户端可以使用 IP 地址的有效租期等信息。

DHCP Options

Option id	长度(字节)	描述
1	4	Subnet Mask
3	n*4	Router(网关)
6	n*4	DNS Server

Option id	长度(字节)	描述
7	n*4	Log Server
26	2	Interface MTU
33	n*8	Static route
35	4	ARP cache timeout
42	n*4	NTP servers
51	4	IP address lease time
53	1	Message type: • 1-DHCPDISCOVER • 2-DHCPOFFER • 3-DHCPREQUEST • 4-DHCPDECLINE • 5-DHCPACK • 6-DHCPNAK • 7-DHCPRELEASE • 8-DHCPINFORM
54	4	DHCP Server Identifier
60	n	华为自定义：可配置该终端设备在发起 DHCP 请求时，通过 Option 60 携带域信息。收到 DHCP 报文时，可根据 Option 60 中携带的域信息来分配 IP 地址。
82	n	华为自定义：作为 DHCP Relay，在中继用户 DHCP 报文时，可在 Option 82 中填写用户的物理位置信息，通知 DHCP 服务器按物理位置信息对为用户分配 IP 地址。

报文示例

图 3 DHCP 报文格式（discover 阶段）

```
⊞ Frame 1: 314 bytes on wire (2512 bits), 314 bytes captured (2512 bit
⊞ Ethernet II, Src: Grandstr_01:fc:42 (00:0b:82:01:fc:42), Dst: Broadc
⊞ Internet Protocol Version 4, Src: 0.0.0.0 (0.0.0.0), Dst: 255.255.25
⊞ User Datagram Protocol, Src Port: bootpc (68), Dst Port: bootps (67)
⊟ Bootstrap Protocol
    Message type: Boot Request (1)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0x00003d1d
    Seconds elapsed: 0
  ⊟ Bootp flags: 0x0000 (Unicast)
      0... .... .... .... = Broadcast flag: Unicast
      .000 0000 0000 0000 = Reserved flags: 0x0000
    Client IP address: 0.0.0.0 (0.0.0.0)
    Your (client) IP address: 0.0.0.0 (0.0.0.0)
    Next server IP address: 0.0.0.0 (0.0.0.0)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: Grandstr_01:fc:42 (00:0b:82:01:fc:42)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: DHCP
  ⊟ Option: (t=53,l=1) DHCP Message Type = DHCP Discover
      Option: (53) DHCP Message Type
      Length: 1
      Value: 01
  ⊞ Option: (t=61,l=7) Client identifier
  ⊞ Option: (t=50,l=4) Requested IP Address = 0.0.0.0
  ⊞ Option: (t=55,l=4) Parameter Request List
    End Option
    Padding
```

图 4 DHCP 报文格式（offer 阶段）

```
⊞ Frame 2: 342 bytes on wire (2736 bits), 342 bytes captured (2736 bit
⊞ Ethernet II, Src: DellComp_ad:f1:9b (00:08:74:ad:f1:9b), Dst: Grands
⊞ Internet Protocol Version 4, Src: 192.168.0.1 (192.168.0.1), Dst: 19
⊞ User Datagram Protocol, Src Port: bootps (67), Dst Port: bootpc (68)
⊟ Bootstrap Protocol
    Message type: Boot Reply (2)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0x00003d1d
    Seconds elapsed: 0
  ⊟ Bootp flags: 0x0000 (Unicast)
      0... .... .... .... = Broadcast flag: Unicast
      .000 0000 0000 0000 = Reserved flags: 0x0000
    Client IP address: 0.0.0.0 (0.0.0.0)
    Your (client) IP address: 192.168.0.10 (192.168.0.10)
    Next server IP address: 192.168.0.1 (192.168.0.1)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: Grandstr_01:fc:42 (00:0b:82:01:fc:42)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: DHCP
  ⊟ Option: (t=53,l=1) DHCP Message Type = DHCP Offer
      Option: (53) DHCP Message Type
      Length: 1
      Value: 02
  ⊞ Option: (t=1,l=4) Subnet Mask = 255.255.255.0
  ⊞ Option: (t=58,l=4) Renewal Time Value = 30 minutes
  ⊞ Option: (t=59,l=4) Rebinding Time Value = 52 minutes, 30 seconds
  ⊞ Option: (t=51,l=4) IP Address Lease Time = 1 hour
  ⊞ Option: (t=54,l=4) DHCP Server Identifier = 192.168.0.1
    End Option
    Padding
```

图 5 DHCP 报文格式（request 阶段）

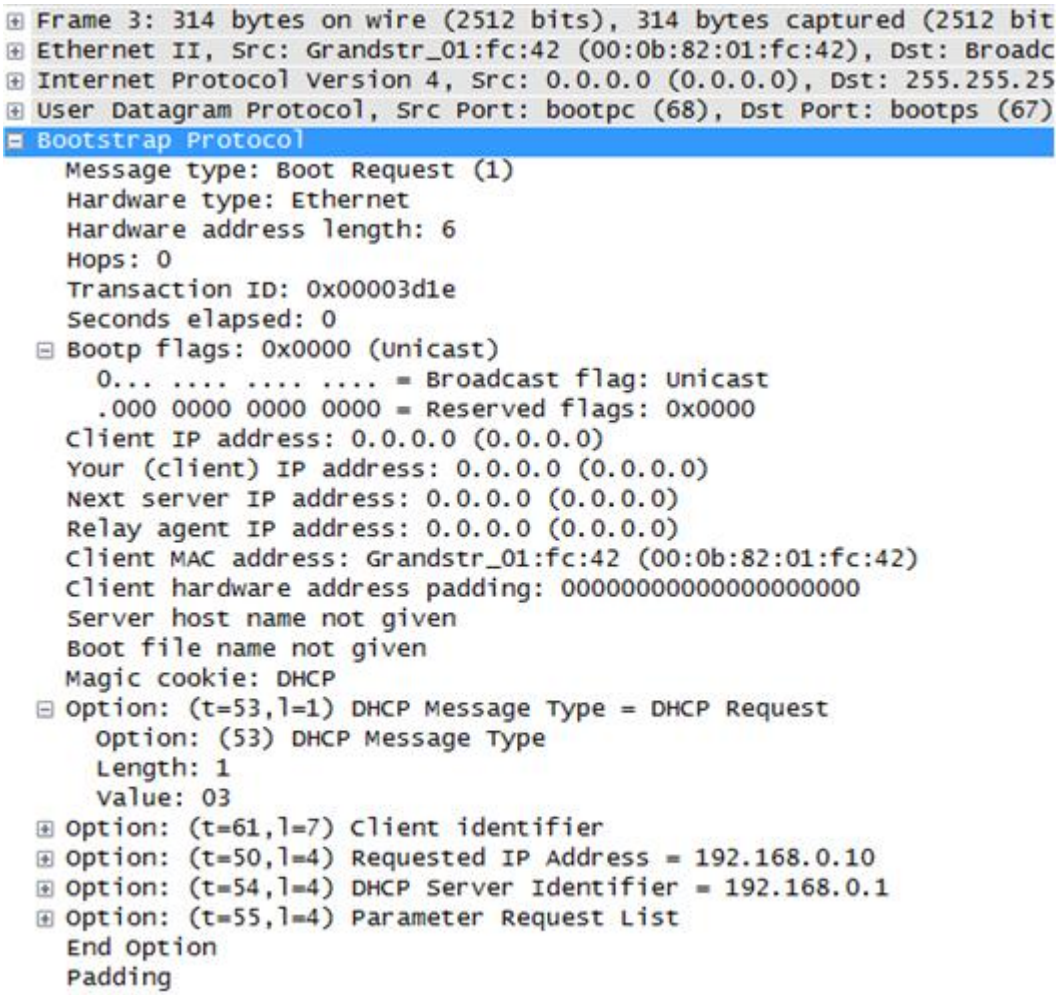


图 6 DHCP 报文格式（ACK 阶段）

```
⊞ Frame 4: 342 bytes on wire (2736 bits), 342 bytes captured (2736 bit:
⊞ Ethernet II, Src: DellComp_ad:f1:9b (00:08:74:ad:f1:9b), Dst: Grands
⊞ Internet Protocol Version 4, Src: 192.168.0.1 (192.168.0.1), Dst: 19
⊞ User Datagram Protocol, Src Port: bootps (67), Dst Port: bootpc (68)
⊟ Bootstrap Protocol
    Message type: Boot Reply (2)
    Hardware type: Ethernet
    Hardware address length: 6
    Hops: 0
    Transaction ID: 0x00003d1e
    Seconds elapsed: 0
  ⊟ Bootp flags: 0x0000 (Unicast)
      0... .... .... .... = Broadcast flag: Unicast
      .000 0000 0000 0000 = Reserved flags: 0x0000
    Client IP address: 0.0.0.0 (0.0.0.0)
    Your (client) IP address: 192.168.0.10 (192.168.0.10)
    Next server IP address: 0.0.0.0 (0.0.0.0)
    Relay agent IP address: 0.0.0.0 (0.0.0.0)
    Client MAC address: Grandstr_01:fc:42 (00:0b:82:01:fc:42)
    Client hardware address padding: 00000000000000000000
    Server host name not given
    Boot file name not given
    Magic cookie: DHCP
  ⊟ Option: (t=53,l=1) DHCP Message Type = DHCP ACK
      Option: (53) DHCP Message Type
      Length: 1
      Value: 05
  ⊞ Option: (t=58,l=4) Renewal Time Value = 30 minutes
  ⊞ Option: (t=59,l=4) Rebinding Time Value = 52 minutes, 30 seconds
  ⊞ Option: (t=51,l=4) IP Address Lease Time = 1 hour
  ⊞ Option: (t=54,l=4) DHCP Server Identifier = 192.168.0.1
  ⊞ Option: (t=1,l=4) Subnet Mask = 255.255.255.0
    End Option
    Padding
```

参考标准

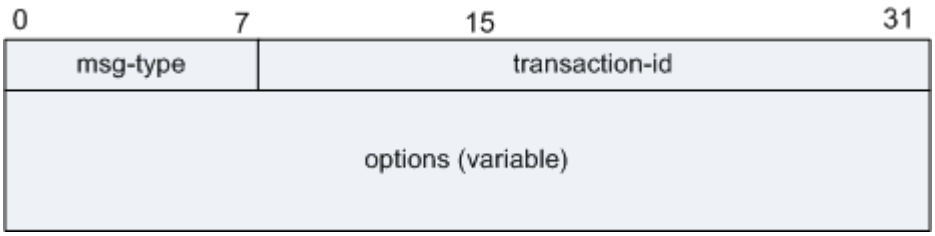
标准	描述
RFC 951	Bootstrap Protocol (BOOTP)
RFC 1542	Clarifications and Extensions for the Bootstrap Protocol
RFC 2131	Dynamic Host Configuration Protocol
RFC 2132	DHCP Options and BOOTP Vendor Extensions

6.6 DHCPv6 报文格式

DHCPv6 报文是承载于 UDP 上的高层协议报文，RFC 推荐采用 547（DHCPv6 服务器/Relay）和 546（DHCPv6 客户端）两个端口号。

报文格式

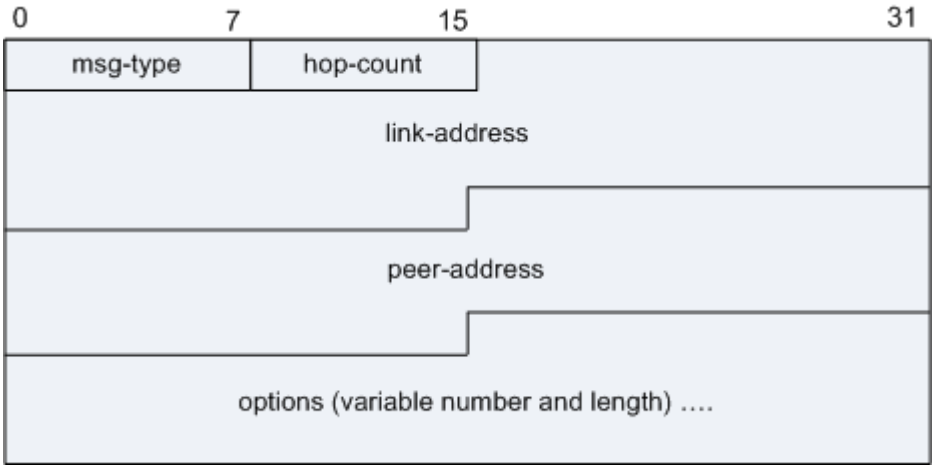
图 1 客户端/服务器端的 DHCPv6 消息格式



各字段的含义：

字段	长度	含义
msg-type	1 字节	报文类型： • SOLICIT (1) • ADVERTISE (2) • REQUEST (3) • CONFIRM (4) • RENEW (5) • REBIND (6) • REPLY (7) • RELEASE (8) • DECLINE (9) • RECONFIGURE (10) • INFORMATION-REQUEST (11) • RELAY-FORW (12) • RELAY-REPL (13)
transaction-id	2 字节	消息 ID
options	可变	选项字段

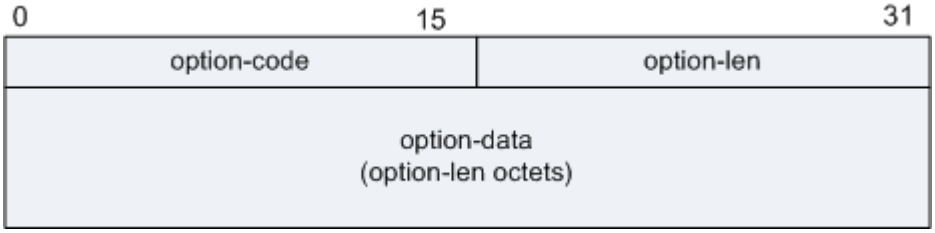
图 2 中继的 DHCPv6 消息格式



各字段的含义：

字段	长度	含义
msg-type	1 字节	• Relay-forward Message: RELAY-FORW • Relay-reply Message: RELAY-REPL
hop-count	1 字节	• Relay-forward Message: Number of relay agents that have relayed this message. • Relay-reply Message: Copied from the Relay-forward message.
link-address	12 字节	• Relay-forward Message: A global or site-local address that will be used by the server to identify the link on which the client is located. • Relay-reply Message: Copied from the Relay-forward message.
peer-address	12 字节	• Relay-forward Message: The address of the client or relay agent from which the message to be relayed was received. • Relay-reply Message: Copied from the Relay-forward message.
options	可变	必须包含名为"Relay Message option"的选项，当然，中继可以添加其他选项。

图 3 DHCPv6 Options 字段的格式



字段	长度	含义
option-code	2 字节	无符号整数，标识选项的类型： • OPTION_CLIENTID (1)：标识客户端身份，用于识别客户。

字段	长度	含义
		• OPTION_SERVERID (2): 用于识别服务器。 • OPTION_IA_NA (3): 非临时地址集合选项。 • OPTION_IA_TA (4): 临时地址集合选项。 • OPTION_IAADDR (5): 用于携带地址选项。 • OPTION_ORO (6): 选项请求选项，用来在客户端和服务器之间标识一系列选项。 • OPTION_PREFERENCE (7) • OPTION_ELAPSED_TIME (8): Elapsed Time Option • OPTION_RELAY_MSG (9): 在 Relay-forward 消息或 Relay-reply 消息中传递 DHCP 消息。 • OPTION_AUTH (11): Authentication Option • OPTION_UNICAST (12): Server Unicast Option • OPTION_STATUS_CODE (13): Status Code Option • OPTION_RAPID_COMMIT (14): Rapid Commit Option • OPTION_USER_CLASS (15): User Class Option • OPTION_VENDOR_CLASS (16): Vendor Class Option • OPTION_VENDOR_OPTS (17): Vendor-specific Information Option • OPTION_INTERFACE_ID (18): 用于标识用户接入接口 • OPTION_RECONF_MSG (19): Reconfigure Message Option • OPTION_RECONF_ACCEPT (20): Reconfigure Accept Option • DNS Recursive Name Server (23) • DNS Domain Search List (24) • IA_PD (25): (Identity association for prefix delegation)，授权的前缀集合选项。 • IA_PD Prefix (26): 用于携带前缀选项。 • Relay Agent Remote-ID (37) • Relay Agent Subscriber-ID (38) • AFTR Name (64): 用于 DS-Lite 方案中携带 AFTR 的域名。
option-len	2 字节	无符号整数，标示 option-data 字段的字节数。
option-data	可变	选项的数据部分。

报文示例

图 4 DHCPv6 Solicit 报文(IAPD & IANA)

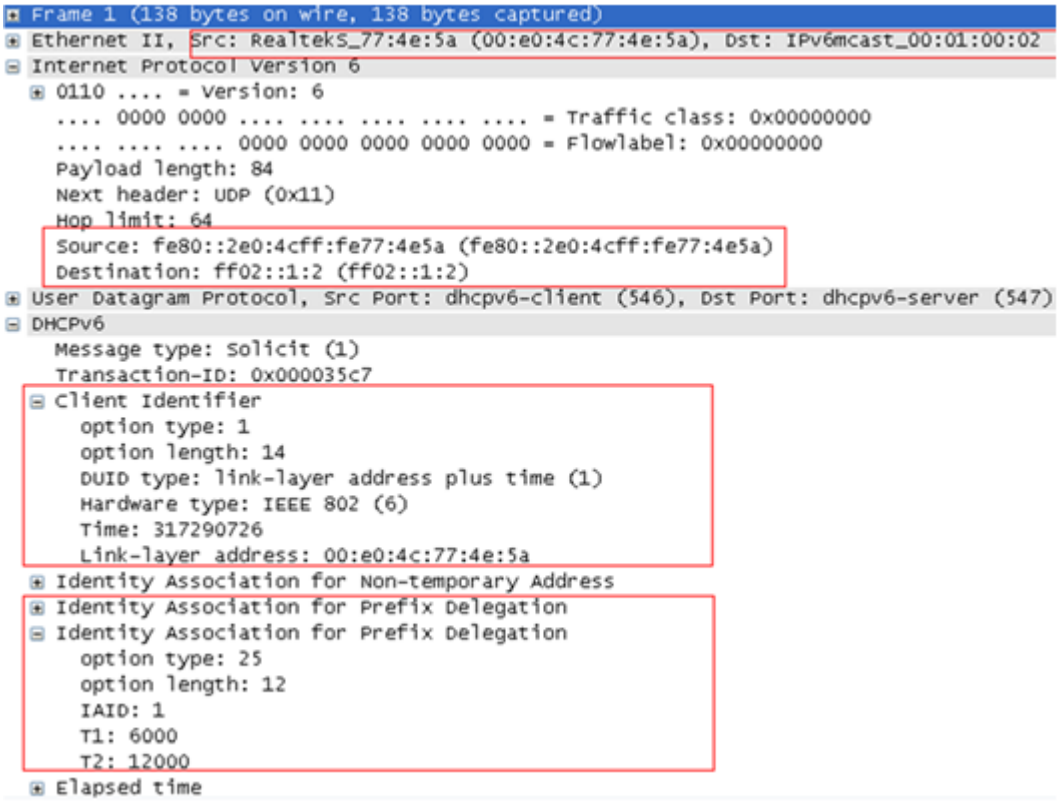


图 5 DHCPv6 Request 报文

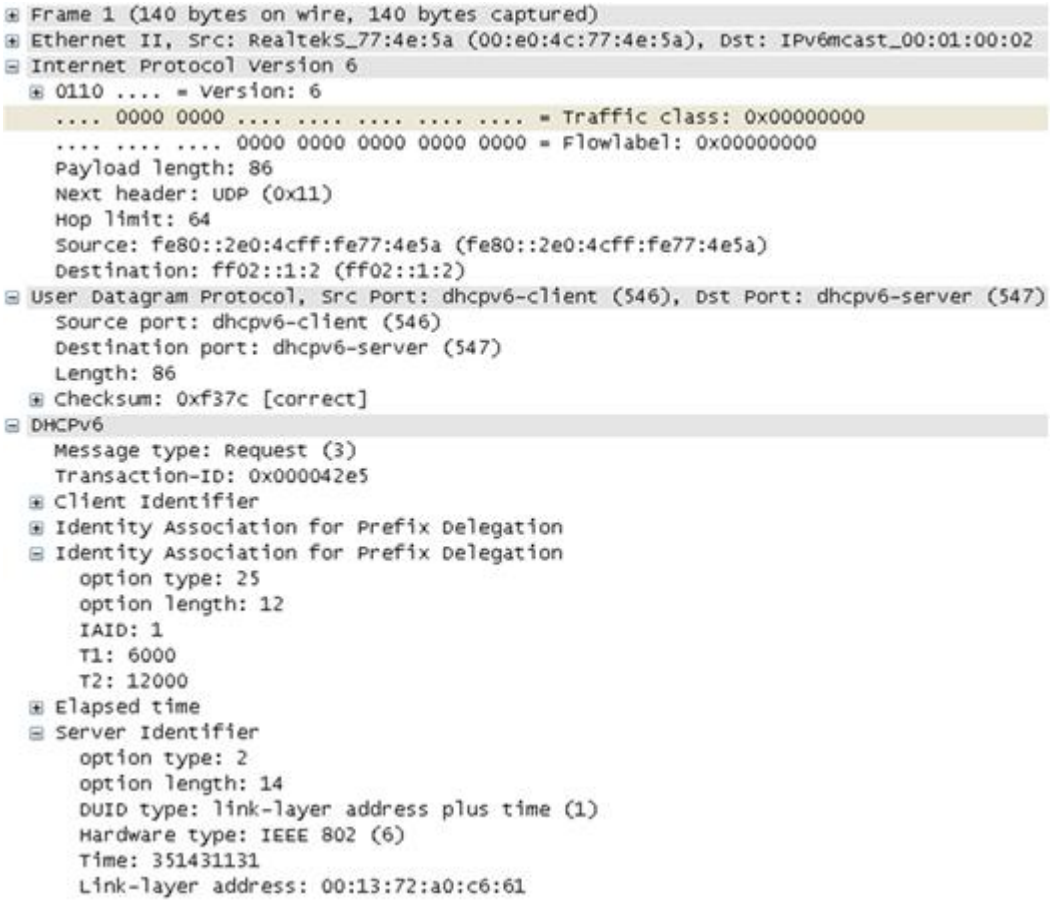


图 6 DHCPv6 Relay-Forw 报文

```
⊞ Frame 1 (228 bytes on wire, 228 bytes captured)
⊞ Ethernet II, Src: 02:00:e0:00:01:33 (02:00:e0:00:01:33), Dst: Dell_a0:c6:61 (00:13:7
⊞ Internet Protocol Version 6
⊞ User Datagram Protocol, Src Port: dhcpv6-server (547), Dst Port: dhcpv6-server (547)
⊟ DHCPv6
    Message type: Relay-forw (12)
    Hop count: 0
    Link-address: 4001::1
    Peer-address: fe80::2e0:4cff:fe77:4e5a
  ⊟ Interface-Id
      option type: 18
      option length: 44
      Interface-ID
  ⊟ Relay Message
      option type: 9
      option length: 76
    ⊟ DHCPv6
        Message type: Solicit (1)
        Transaction-ID: 0x000035c7
      ⊞ Client Identifier
      ⊞ Identity Association for Non-temporary Address
      ⊞ Identity Association for Prefix Delegation
      ⊞ Identity Association for Prefix Delegation
      ⊞ Elapsed time
```

```
0000  00 13 72 a0 c6 61 02 00  e0 00 01 33 86 dd 60 13   ..r..a.. ...3..`.
0010  33 33 00 aa 11 ff 40 01  00 00 00 00 00 00 00 00   33....@. ........
0020  00 00 00 00 00 01 30 01  00 00 00 00 00 00 00 00   ......0. ........
0030  00 00 00 00 00 01 02 23  02 23 00 aa 19 7d 0c 00   .......# .#...}..
0040  40 01 00 00 00 00 00 00  00 00 00 00 00 00 00 01   @....... ........
0050  fe 80 00 00 00 00 00 00  02 e0 4c ff fe 77 4e 5a   ........ ..L..wNZ
0060  00 12 00 2c 30 37 31 36  2e 30 30 30 30 2e 30 30   ...,0716 .0000.00
0070  65 30 34 63 37 37 34 65  35 61 3a 47 69 67 61 62   e04c774e 5a:Gigab
0080  69 74 45 74 68 65 72 6e  65 74 30 2f 30 2f 34 35   itEthern et0/0/45
0090  00 09 00 4c 01 00 35 c7  00 01 00 0e 00 01 00 06   ...L..5. ........
```

图 7 DHCPv6 Relay-Reply 报文

```
⊞ Frame 1 (385 bytes on wire, 385 bytes captured)
⊞ Ethernet II, Src: Dell_a0:c6:61 (00:13:72:a0:c6:61), Dst: 02:00:e0:00:01:33 (02:00:e
⊟ Internet Protocol Version 6
  ⊞ 0110 .... = Version: 6
    .... 0000 0000 .... .... .... .... .... = Traffic class: 0x00000000
    .... .... .... 0000 0000 0000 0000 0000 = Flowlabel: 0x00000000
    Payload length: 331
    Next header: UDP (0x11)
    Hop limit: 128
    Source: 3001::1 (3001::1)
    Destination: 4001::1 (4001::1)
⊞ User Datagram Protocol, Src Port: dhcpv6-server (547), Dst Port: dhcpv6-server (547)
⊟ DHCPv6
    Message type: Relay-reply (13)
    Hop count: 0
    Link-address: 4001::1
    Peer-address: fe80::2e0:4cff:fe77:4e5a
  ⊞ Interface-Id
  ⊟ Relay Message
      option type: 9
      option length: 237
    ⊟ DHCPv6
        Message type: Reply (7)
        Transaction-ID: 0x00005066
      ⊞ Client Identifier
      ⊞ Server Identifier
      ⊞ Identity Association for Non-temporary Address
      ⊞ Identity Association for Prefix Delegation
```

参考标准

标准	描述
RFC3315	Dynamic Host Configuration Protocol for IPv6 (DHCPv6)
RFC3633	IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version6

标准	描述
RFC3646	DNS Configuration options for Dynamic Host Configuration Protocol for IPv6 (DHCPv6)
RFC3736	Stateless Dynamic Host Configuration Protocol (DHCP) Service for IPv6
RFC6221	Lightweight DHCPv6 Relay Agent
RFC2131	Dynamic Host Configuration Protocol
RFC2131	DHCP Options and BOOTP Vendor Extensions
RFC3406	DHCP Relay Agent Information Option

6.7 Diameter 协议报文格式

Diameter 协议是 IETF 的 AAA 工作组作为下一代的 AAA 协议标准，由 RADIUS 协议演进而来。

Diameter 目前主要应用于移动通信系统，固网接入主要使用 RFC2865 的 Radius 协议。随着固网与无线网络融合，统一认证计费授权服务器的趋势越来越迫切，固网接入支持 Diameter 协议栈已成为普遍需求。

报文格式

图 1 Diameter 协议栈结构

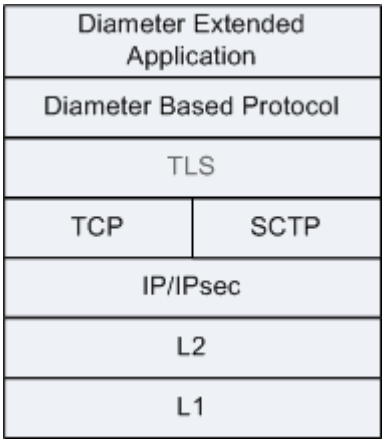
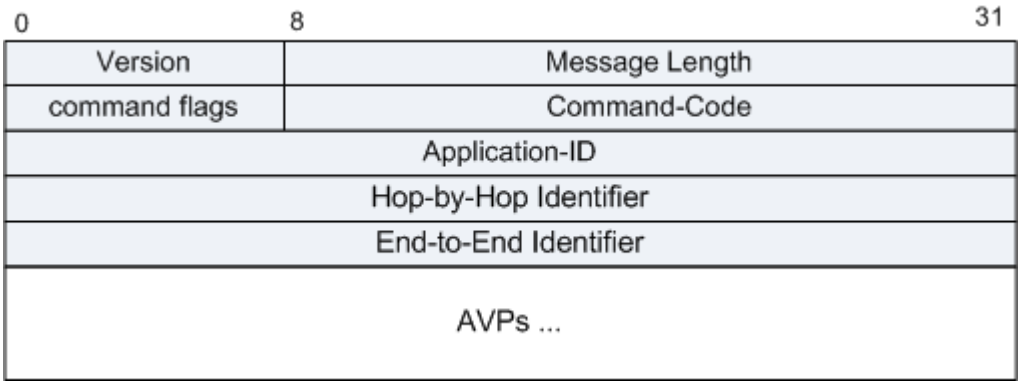


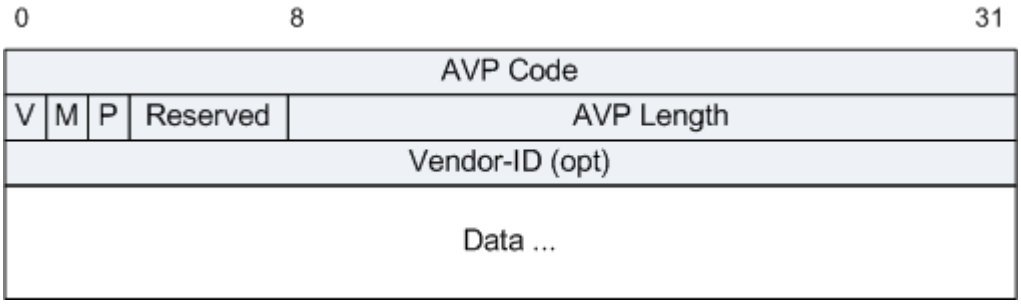
图 2 Diameter 报文格式



字段	长度	含义
Version	1 字节	必须置为 1，表示 Diameter 协议版本号为 1。
Message Length	3 字节	表示 Ditameter 消息长度，包括 Diameter 的头部域。
Command Flags	1 字节	此字段格式如下： 0 1 2 3 4 5 6 7 +-+-+-+-+-+-+-+-+ |R P E T r r r r| +-+-+-+-+-+-+-+-+ • R(equest) - 如果置 1，表示消息为请求消息，如果置 0，表示消息为应答消息。 • P(roxiable) - 如果置 1，表示消息可能被代理、中继或重定向。如果置 0，表示消息必须本地处理。 • E(rror) - 如果置 1，表示消息包含协议错误，不符合 ABNF 的定义。带有 E 比特置位的通常表示错误消息。在请求消息中该比特不应该置位。 • T(Potentially re-transmitted message) - 当发送的请求还没得到确认时，此标记置位，表示可能因为链路故障导致消息的重复。 第 1 次发送的请求消息中此标记必须置 0。应答消息中此标记也应该置 0。 • r(eserved) - 预留将来使用，必须设置为 0，接收时忽略。
Command-Code	3 字节	命令代码，代码值由 IANA 分配，其中 0xFFFFFE -0xFFFFFF 预留给实验用。 • Abort-Session-Request (ASR): 274 • Abort-Session-Answer (ASA): 274 • Accounting-Request (ACR): 271 • Accounting-Answer (ACA): 271 • Capabilities-Exchange-Request (CER): 257

字段	长度	含义
		• Capabilities-Exchange-Answer (CEA): 257 • Device-Watchdog-Request (DWR): 280 • Device-Watchdog-Answer (DWA): 280 • Disconnect-Peer-Request (DPR): 282 • Disconnect-Peer-Answer (DPA): 282 • Re-Auth-Request (RAR): 258 • Re-Auth-Answer (RAA): 258 • Session-Termination-Request (STR): 275 • Session-Termination-Answer (STA): 275
Application-ID	4 字节	用来标记消息的应用，该应用可能是认证、计费或者厂家特殊应用。
Hop-by-Hop Identifier	4 字节	逐跳标记，用来匹配请求和应答，通常是一个自动增加的编号，从一个随机的数开始增加。应答消息里的此字段如果无法识别，消息将被丢弃。
End-to-End Identifier	4 字节	用于检测重复消息。
AVPs	变长	Diameter 消息使用 AVP 来封装信息。

图 3 AVP 消息格式



字段	长度	含义
AVP Code	4 bytes	AVP Code 字段和 Vendor-Id 字段一起唯一标识了一个属性。1 - 255 预留用于和 RADIUS 后向兼容，不需要携带 Vendor-Id 字段。256 及以上的值用于 Diameter 协议，由 IANA 分配。
V	1 bit	V（Vendor-Specific bit）位用来标识 AVP 头部是否必须携带 Vendor-ID 字段。
M	1 bit	M（Mandatory）位用来标识此 AVP 是否必须携带。

字段	长度	含义
P	1 bit	P 位用来标识是否需要加密。
Reserved	5 bits	保留位。
AVP Length	3 bytes	表示 AVP 的字节数，包括 AVP Code、AVP Length、AVP Flags、Vendor-ID、AVP data 字段。
Vendor-ID	4 bytes	IANA 分配的厂家标识。如果 V 比特置位，AVP 必须携带 Vendor-ID 字段。
Data	Variable	包括 0 个或多个属性。

报文示例

图 4 Diameter 报文（使用 TCP 封装）

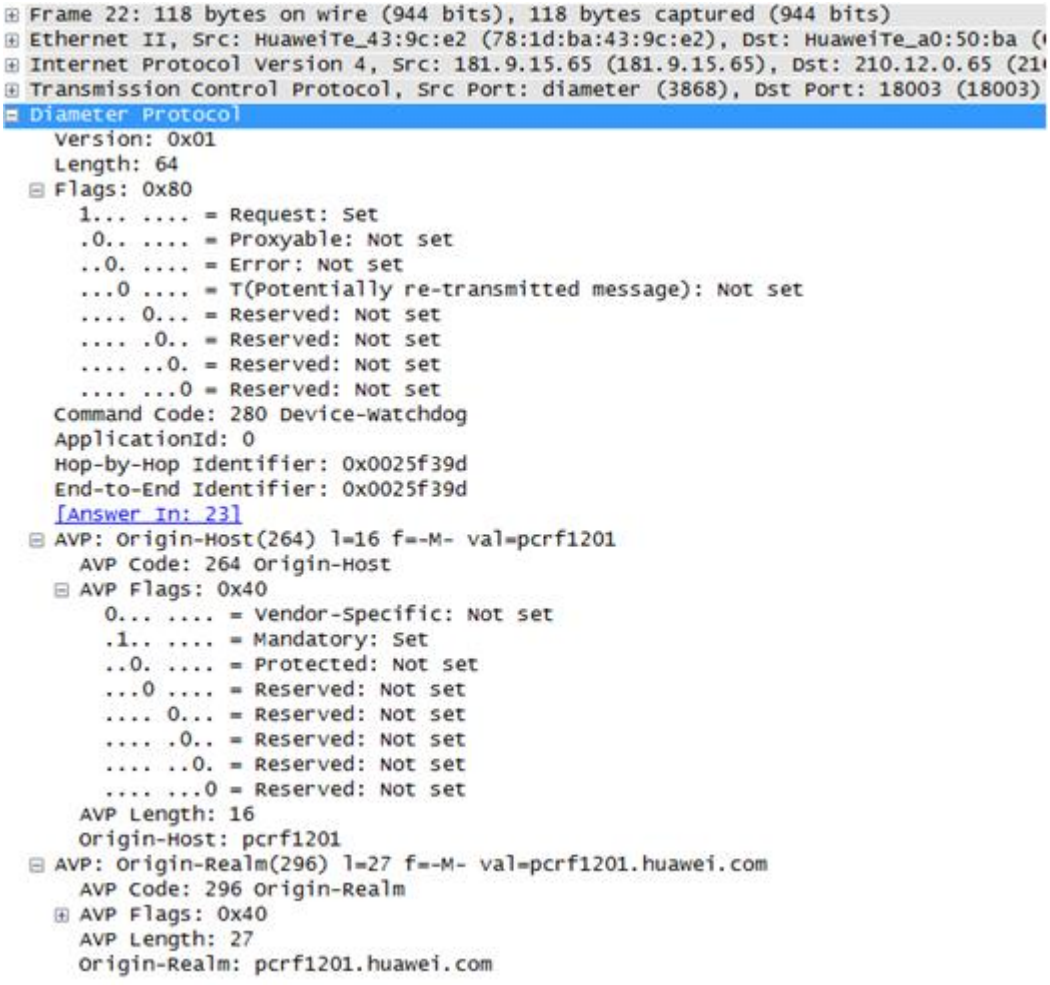


图 5 Diameter 报文（使用 SCTP 封装）

```
⊞ Frame 7: 302 bytes on wire (2416 bits), 302 bytes captured (2416 bits)
⊞ Ethernet II, Src: Hangzhou_d5:c5:29 (00:0f:e2:d5:c5:29), Dst: Hangzhou_2f:d7:18 (00:0f
⊞ Internet Protocol Version 4, Src: 172.21.112.2 (172.21.112.2), Dst: 192.168.168.9 (192
⊞ Stream Control Transmission Protocol, Src Port: 3876 (3876), Dst Port: diameter (3868)
⊟ Diameter Protocol
    Version: 0x01
    Length: 236
  ⊟ Flags: 0x80
      1... .... = Request: Set
      .0.. .... = Proxyable: Not set
      ..0. .... = Error: Not set
      ...0 .... = T(Potentially re-transmitted message): Not set
      .... 0... = Reserved: Not set
      .... .0.. = Reserved: Not set
      .... ..0. = Reserved: Not set
      .... ...0 = Reserved: Not set
    Command Code: 257 Capabilities-Exchange
    ApplicationId: 0
    Hop-by-Hop Identifier: 0x00001ff4
    End-to-End Identifier: 0x00001ff4
  ⊞ AVP: Origin-Host(264) l=53 f=-M- val=huawei2.mme.epc.mnc001.mcc460.3gppnetwork.org
  ⊞ AVP: Origin-Realm(296) l=41 f=-M- val=epc.mnc001.mcc460.3gppnetwork.org
  ⊞ AVP: Host-IP-Address(257) l=14 f=-M- val=172.21.112.2 (172.21.112.2)
  ⊞ AVP: Vendor-Id(266) l=12 f=-M- val=2011
  ⊞ AVP: Product-Name(269) l=11 f=--- val=MME
  ⊞ AVP: Inband-Security-Id(299) l=12 f=-M- val=NO_INBAND_SECURITY (0)
  ⊞ AVP: Vendor-Specific-Application-Id(260) l=32 f=-M-
  ⊞ AVP: Vendor-Specific-Application-Id(260) l=32 f=-M-
```

参考标准

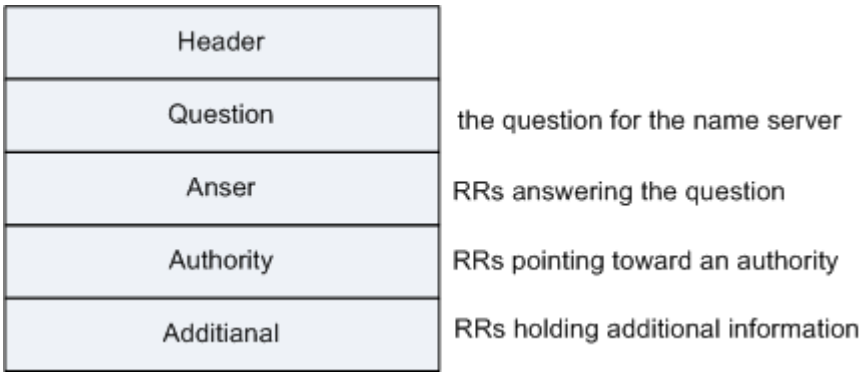
标准	描述
RFC 3588	Diameter Base Protocol

6.8 DNS 报文格式

报文格式

DNS 报文格式分为 DNS 查询和响应的报文格式。这个报文由 12 字节长的首部和 4 个长度可变的字段组成。报文中问题字段是由客户填入的，由服务器返回问题的回答、授权和附加信息字段。报文格式如下(RFC 1035)：

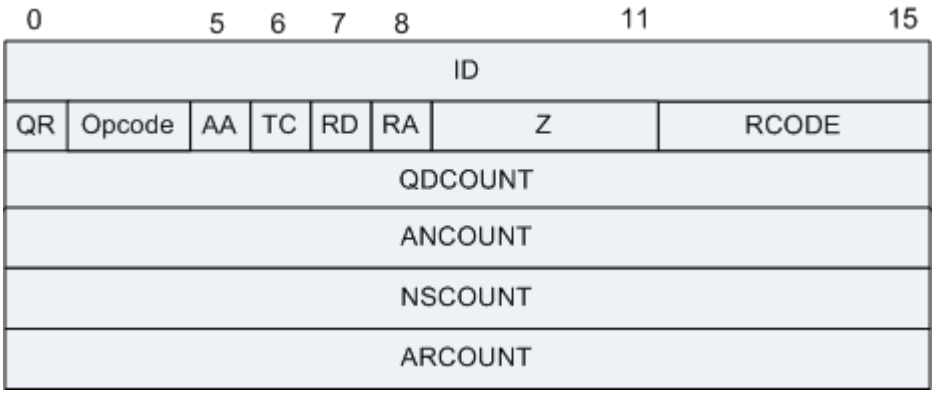
图 1 DNS 报文封装格式



Header 段是必须存在的，它定义了报文是请求还是应答，也定义了其他段是否需要存在，以及是标准查询还是其他。

头部字段格式如下：

图 2 头部字段格式

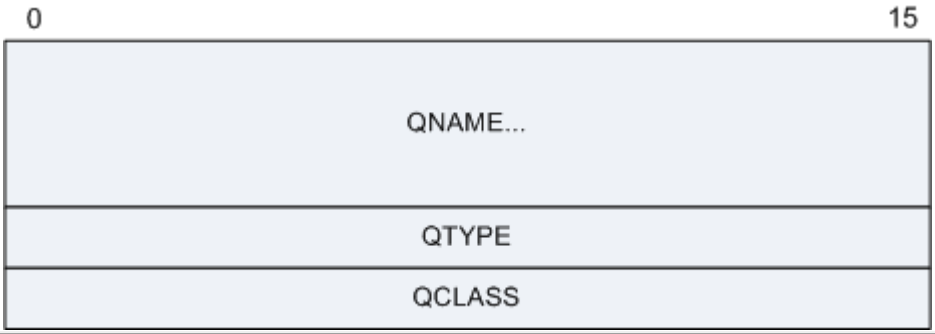


字段	长度	描述
ID	16 bit	标识字段，客户通过标识字段来确定 DNS 响应是否与查询请求匹配。
QR	1 bit	操作类型： • 0：查询报文 • 1：响应报文
OPCODE	4 bit	查询类型： • 0：标准查询 • 1：反向查询 • 2：服务器状态查询 • 3～15：保留未用 反向查询是客户端请求服务器根据回答生成导致此回答的问题，这个查询类型的使用并不多。
AA	1 bit	若置位，则表示该域名解析服务器是授权回答该域的。
TC	1 bit	若置位，则表示报文被截断。 使用 UDP 传输时，应答的总长度超过 512 字节时，只返回报文的前 512 个字节内容。
RD	1 bit	客户端希望域名解析服务器采取的解析方式： • 0：表示希望域名解析服务器采取迭代解析 • 1：表示希望域名解析服务器采取递归解析
RA	1 bit	域名解析服务器采取的解析方式： • 0：表示域名解析服务器采取迭代解析 • 1：表示域名解析服务器采取递归解析
Z	3 bit	全部置 0，保留未用。

字段	长度	描述
RCODE	4 bit	响应类型： • 0：无差错 • 1：查询格式错 • 2：服务器失效 • 3：域名不存在 • 4：查询没有被执行 • 5：查询被拒绝 • 6-15：保留未用
QDCOUNT	16 bit	无符号 16 位整数表示报文请求段中的问题记录数。
ANCOUNT	16 bit	无符号 16 位整数表示报文回答段中的回答记录数。
NSCOUNT	16 bit	无符号 16 位整数表示报文授权段中的授权记录数。
ARCOUNT	16 bit	无符号 16 位整数表示报文附加段中的附加记录数。

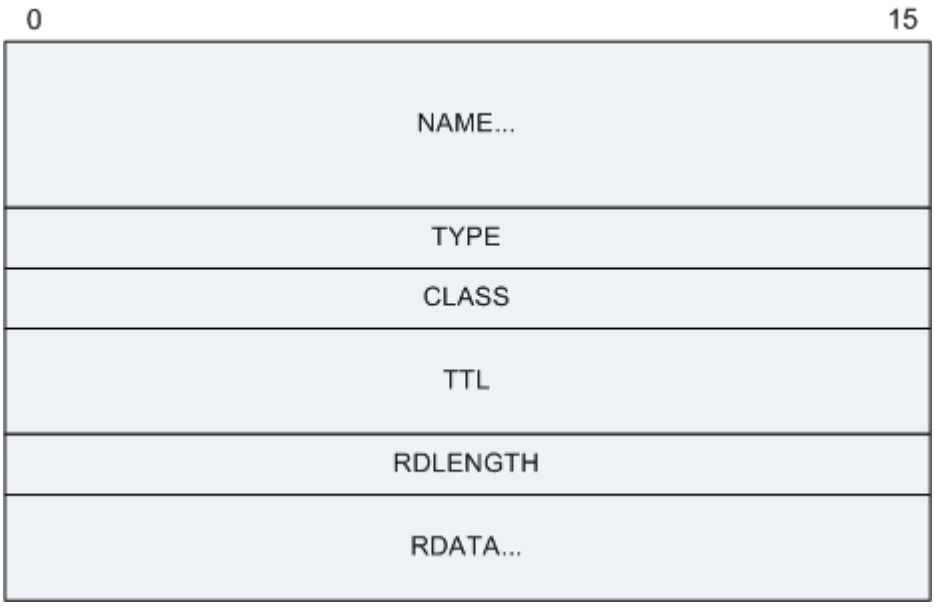
大多数查询中，Question 段包含着问题(question)，比如，指定问什么。这个段包含 QDCOUNT(usually 1)个问题，每个问题为下面的格式：

图 3 Question 字段的格式



字段	长度	描述
QNAME	变长	域名被编码为一些 labels 序列，每个 labels 包含一个字节表示后续字符串长度，以及这个字符串，以 0 长度和空字符串来表示域名结束。注意这个字段可能为奇数字节，不需要进行边界填充对齐。
QTYPE	2 个字节	表示查询类型，.取值可以为任何可用的类型值，以及通配码来表示所有的资源记录。
QCLASS	2 个字节	表示查询的协议类，比如，IN 代表 Internet。

图 4 资源记录字段的格式



应答，授权，附加段都共用相同的格式：多个资源记录，资源记录的个数由报文头段中对应的几个数值确定，每个资源记录格式如下：

字段	长度	描述
NAME	不定长	资源记录包含的域名。
TYPE	2 个字节	表示资源记录的类型，指出 RDATA 数据的含义。
CLASS	2 个字节	表示 RDATA 的类。
TTL	4 字节	无符号整数，表示资源记录可以缓存的时间。0 代表只能被传输，但是不能被缓存。
RDLENGTH	2 个字节	无符号整数，表示 RDATA 的长度。
RDATA	不定长	字符串，表示记录，格式跟 TYPE 和 CLASS 有关。比如，TYPE 是 A，CLASS 是 IN，那么 RDATA 就是一个 4 个字节的 ARPA 网络地址。

报文示例

图 5 DNS query 消息

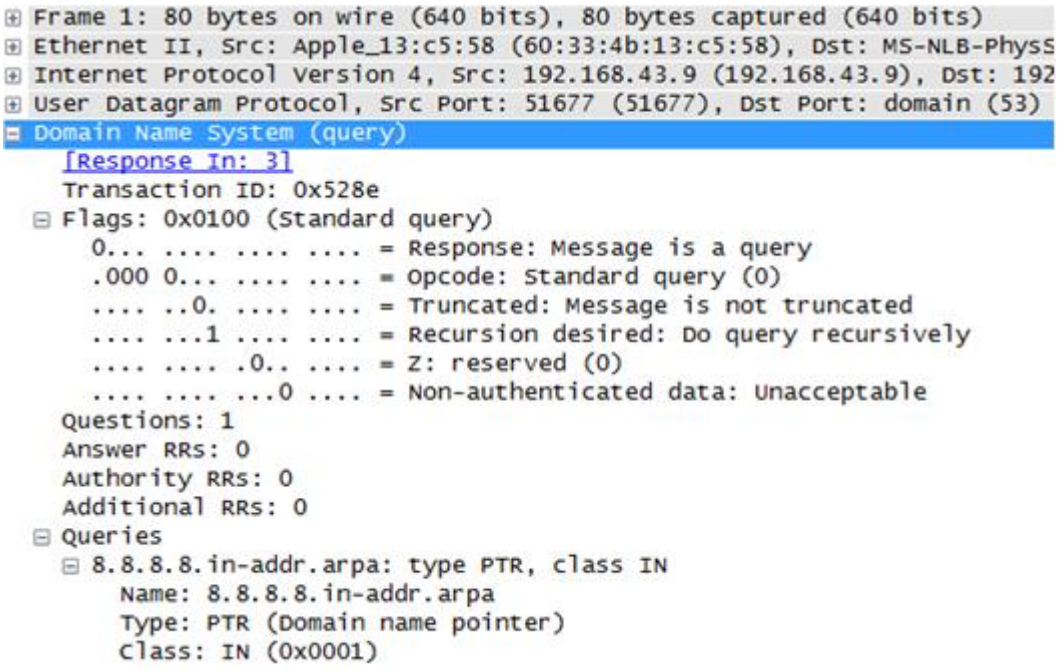
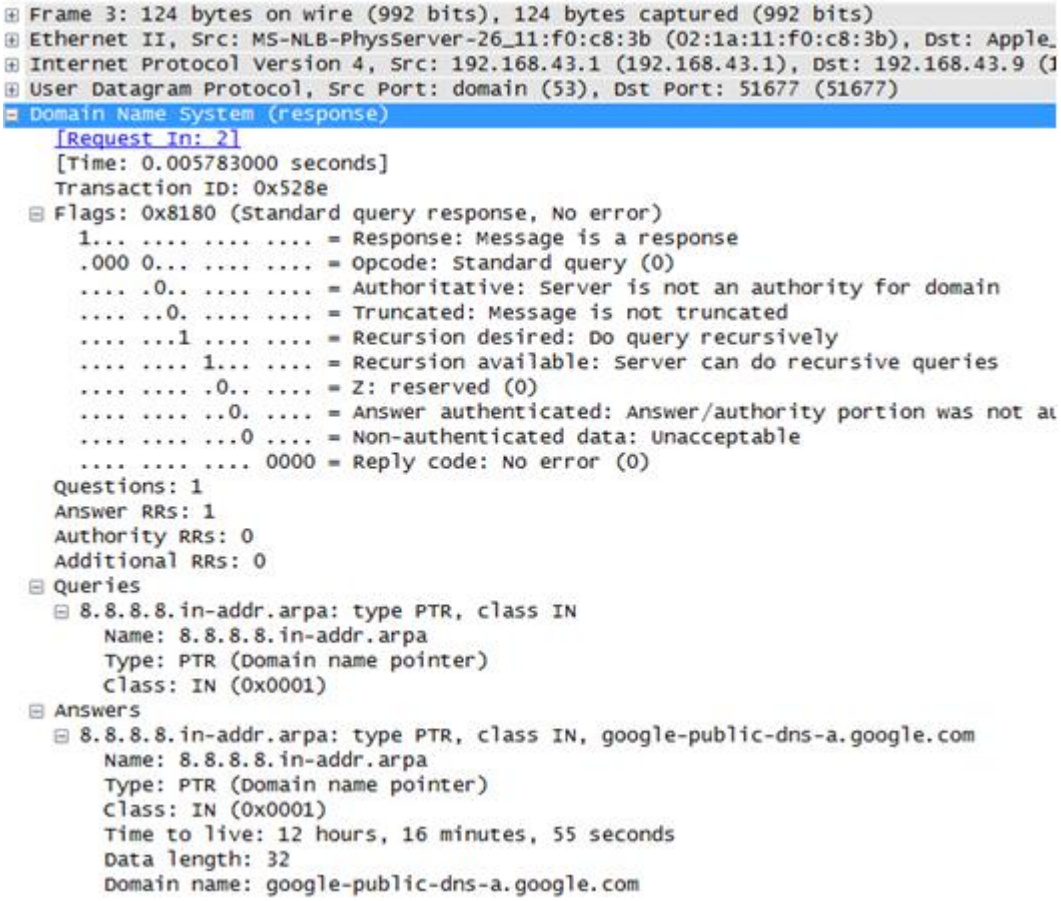


图 6 DNS Response 消息



参考标准

标准	描述
RFC 1034	DOMAIN NAMES - CONCEPTS AND FACILITIES
RFC 1035	DOMAIN NAMES - IMPLEMENTATION AND SPECIFICATION

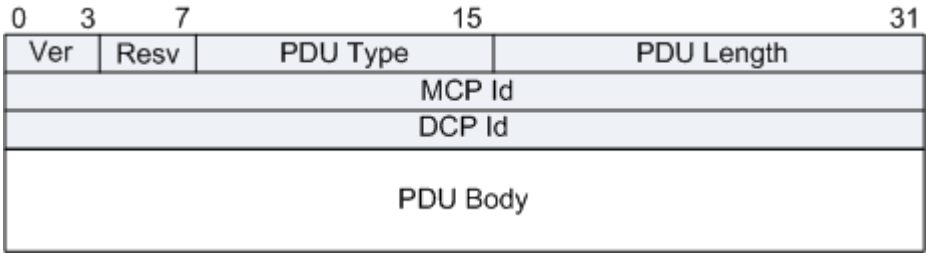
6.9 IP FPM 报文格式

IP FPM（IP Flow Performance Measurement）是一种基于端到端，直接对业务报文进行测量，从而得到 IP 网络的真实丢包率、时延等性能指标的检测方式。

报文格式

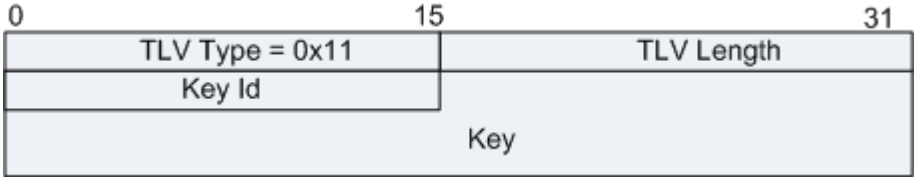
IP FPM 报文格式如下：

图 1 IP FPM 报文格式



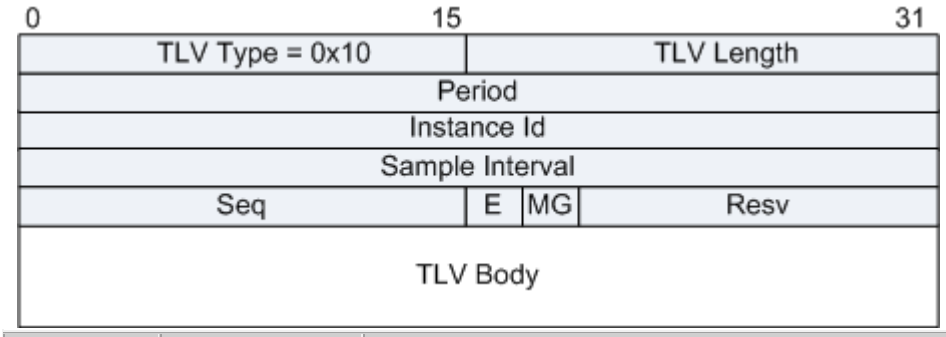
字段	长度	含义
Ver	4 bits	协议版本号，=
Resv	4 bits	保留字段，填 0。
PDU Type	8 bits	PDU 类型，目前只支持一种类型：值为 0x06。
PDU Length	16 bits	PDU 长度，包括头部和内容的整个长度，字节为单位。
MCP Id	32 bits	IP FPM 统计系统的测量控制点 MCP（Measurement Control Point）的 IP 地址。
DCP Id	32 bits	IP FPM 统计系统的数据收集点 DCP（Data Collecting Point）的 IP 地址。
PDU Body	变长	包含 1~n 个 Instance TLV。 包含 0 或 1 个 Auth TLV。

图 2 认证 TLV 格式



字段	长度	含义
TLV Type	16 bits	TLV 类型，Instance TLV 的类型值为 0x10。
TLV Length	16 bits	TLV body 的长度，以字节为计数单位。
key id	16 bits	认证类型 ID。
key	变长	Key 的长度取决于用户配置的认证类型，目前只能支持 hmac-sha256，目前为 32 字节。

图 3 实例 TLV 格式

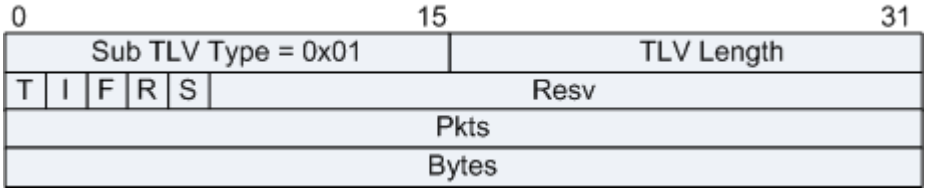


字段	长度	含义
TLV Type	16bits	TLV 类型，Instance TLV 的类型值为 0x10。
TLV Length	16 bits	TLV body 的长度，以字节为计数单位。
Period	8 bits	采样周期号。
Instance Id	16 bits	采样实例 ID。
Sample Interval	32 bits	采样周期，单位是秒。
Seq	32 bits	描述当前实例数据分片号； 当 E bit 为 1 时，当前为最后一个分片，seq 为分片号，同时描述分片数。

字段	长度	含义
E	1 bit	E 比特位为 1 时，标识当前为最后一个分片；E 比特位为 0 时，表示当前为非最后一个分片。
MG	1 bit	只在首分片有意义。 • 0：单播数据 • 1：组播数据
Resv	14 bits	保留
TLV body	变长	实例 TLV 的内容： • 包含 0..n 个 Loss 子 TLV • 包含 0..n 个 Delay 子 TLV • 包含 0..1 个 Error 子 TLV

丢包数据 TLV (Loss TLV)的格式

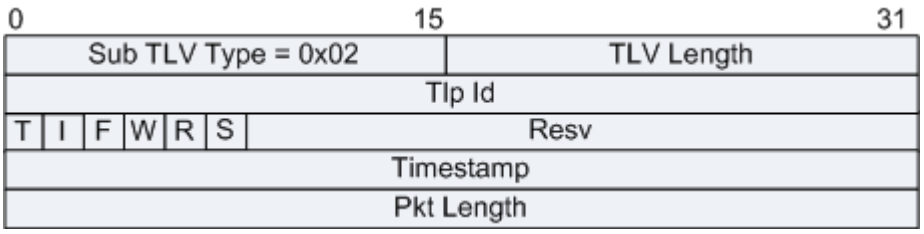
图 4 Loss TLV 格式



字段	长度	含义
TLV Type	16 bits	子 TLV 的类型，Lost TLV 的类型值为 0x01。
TLV Length	16 bits	TLV body 的长度，以字节为计数单位。
T	1 bit	观测点（TLP）的类型： • 0: 端到端 TLP • 1: 逐点 TLP
I	1 bit	观测点位置标记： • 0: ingress • 1: egress
F	1 bit	方向标记： • 0: forward • 1: backward

字段	长度	含义
R	1 bit	组播源的主备标记： • 0：main 主组播源 • 1：backup 备组播源
S	1 bit	源切换状态标记： • 0：未发生源切换 • 1：发生源切换
Resv	27 bits	保留位。
Pkts	64 bits	周期内监测到的流的报文总数。
Bytes	64 bits	周期内监测到的流的字节总数。记录的字节数是 IP 头中的 LENGTH 字段，即不包括报文的 ETH 头、VLAN 头和 4 字节的校验码。

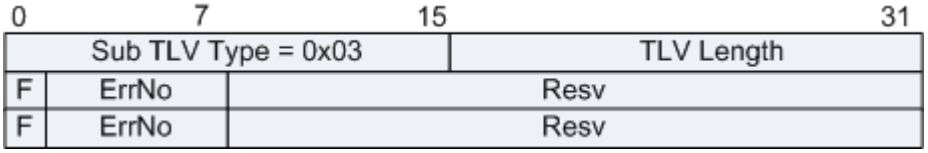
图 5 Delay TLV 格式



字段	长度	含义
TLV Type	16 bits	TLV 类型，Delay TLV 的类型值为 0x02。
TLV Length	16 bits	TLV body 的长度，以字节为计数单位。
Tlp Id	32 bits	TLP 编号
T	1 bit	观测点（TLP）的类型： • 0：端到端 TLP • 1：逐点 TLP
I	1 bit	节点标记： • 0：ingress • 1：egress

字段	长度	含义
F	1 bit	方向标记： • 0：forward • 1：backward
W	1 bit	方向标记： • 0：one-way • 1：two-way
R	1 bit	组播源的主备标记： • 0：main 主组播源 • 1：backup 备组播源
S	1 bit	源切换状态标记： • 0：未发生源切换 • 1：发生源切换
Resv	26 bit	保留位。
Timestamp	64 bit	时间戳。
Pkt Length	32 bit	周期内检测到的报文字节数。记录的字节数是 IP 头中的 LENGTH 字段，即不包括报文的 ETH 头、VLAN 头和 4 字节的校验码。

图 6 Error 子 TLV 格式



字段	长度	含义
TLV Type	16 bits	TLV 类型，Error 子 TLV 的类型值为 0x03。
TLV Length	16 bits	TLV body 的长度，以字节为计数单位。
F		方向标记： • 0：forward

字段	长度	含义
		• 1: backward
ErrNo	7 bits	错误码。
Resv	24 bits	保留字段。

参考标准

IP FPM 为华为私有协议。

6.10 IPSec 报文格式

IPSec 协议族是 IETF（Internet Engineering Task Force）制定的一系列协议，它为 IP 数据报提供了高质量的、可互操作的、基于密码学的安全性。

IPSec 通过认证头 AH（Authentication Header，协议号 51）和封装安全载荷 ESP（Encapsulating Security Payload）这两个安全协议来实现。AH 可提供数据源验证和数据完整性校验功能；ESP 除可提供数据验证和完整性校验功能外，还提供对 IP 报文的加密功能。

IPSec 协议有两种封装模式：

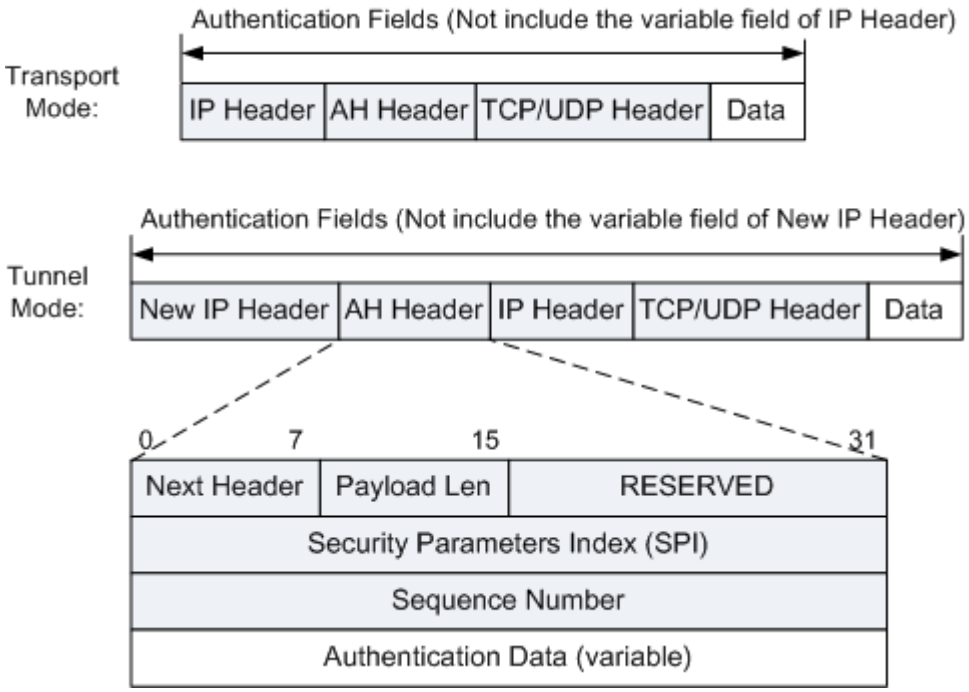
- 传输模式。在传输模式下，AH 或 ESP 被插入到 IP 头之后但在所有传输层协议之前，或所有其他 IPSec 协议之前。
- 隧道模式。在隧道模式下，AH 或 ESP 插在原始 IP 头之前，另外生成一个新 IP 头放到 AH 或 ESP 之前。

传输模式用于两台主机之间的通讯，或者是一台主机和一个安全网关之间的通讯。在传输模式下，对报文进行加密和解密的两台设备本身必须是报文的原始发送者和最终接收者。

通常，在两个安全网关（路由器）之间的数据流量，绝大部分都不是安全网关本身的通讯量，因此在安全网关之间一般不使用传输模式，而总是使用隧道模式。在一个安全网关被加密的报文，只有另一个安全网关能够解密。因此必须对 IP 报文进行隧道封装，即增加一个新的 IP 头，进行隧道封装后的 IP 报文被发送到另一个安全网关，才能够被解密。

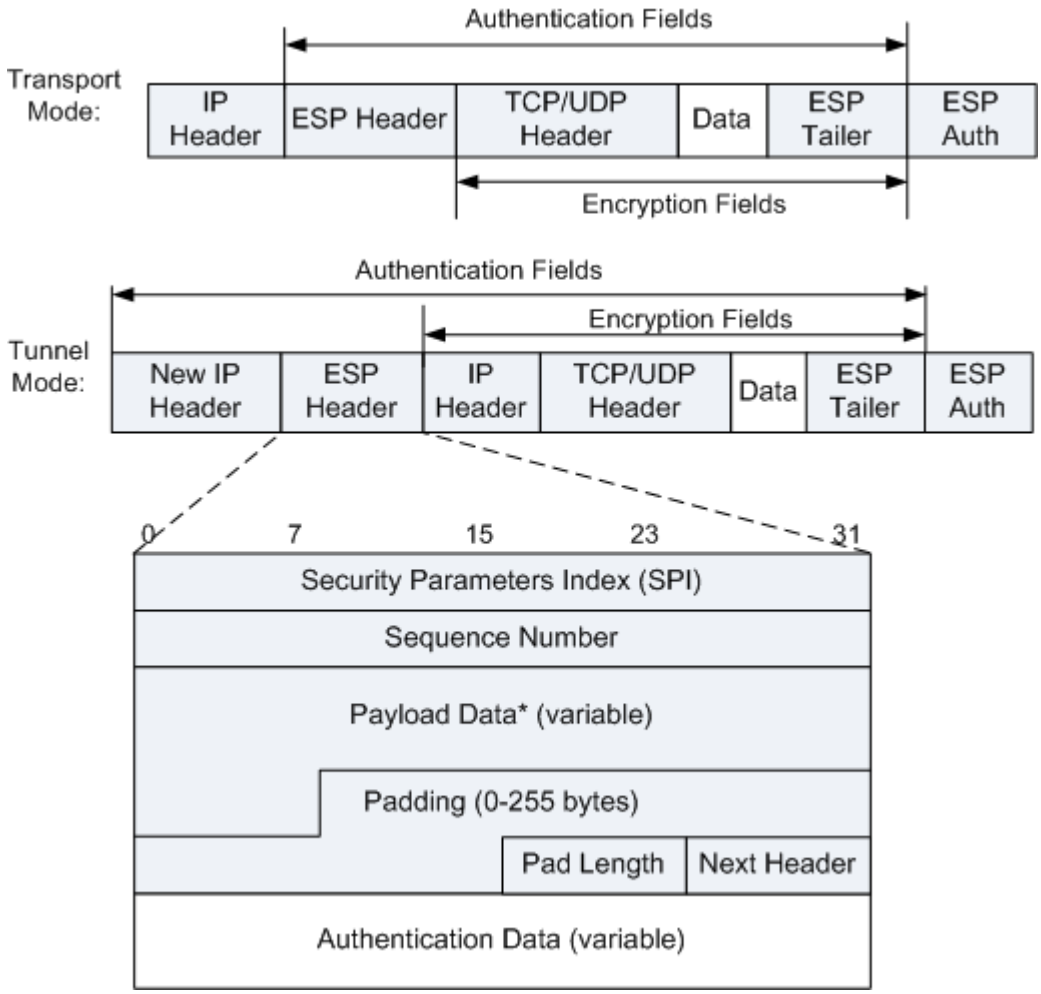
报文格式

图 1 AH 封装及头部格式



字段	长度	描述
Next Header	8 比特	表示认证头部之后的下一个负载。
Payload Len	8 比特	AH 的长度减 2，4 字节为计数单位。例如，有个 96 比特的认证值，长度将是"4"（即头部固定的 3 个 4 字节 + ICV 的 3 个 4 字节 - 2）。对于 IPv6，头部总长度必须为 8 字节的倍数。
RESERVED	16 比特	预留将来使用。必须置 0，接收时忽略。
Security Parameters Index	32 比特	用于给报文接收端识别 SA
Sequence Number Field	32 比特	序列号，每发送一个报文，计数加 1，例如每发一个 SA 报文序列号增加 1。
Integrity Check Value-ICV	变长	报文的 ICV 字段，可变长度，长度必须为 32 比特的整数倍。

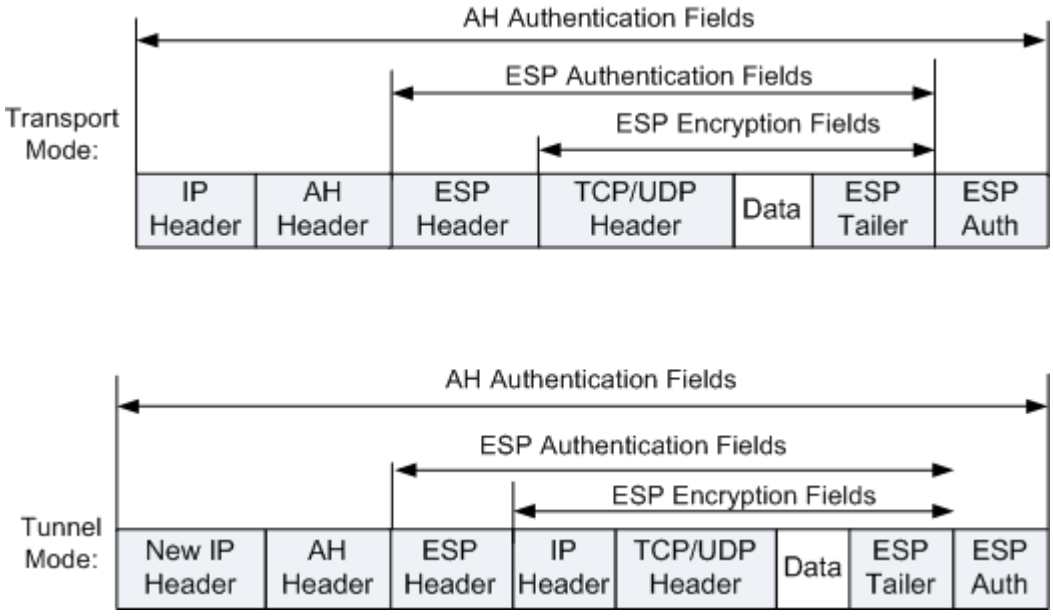
图 2 ESP 封装及头部格式



字段	长度	描述
Security Parameters Index	32 比特	安全参数索引。
Sequence Number	32 比特	序列号。
Payload Data*	变长	有效载荷数据（可变）。
Padding	0 - 255 字节	填充字段。
Pad Length	8 比特	填充字段长度。
Next Header	8 比特	下一个头。

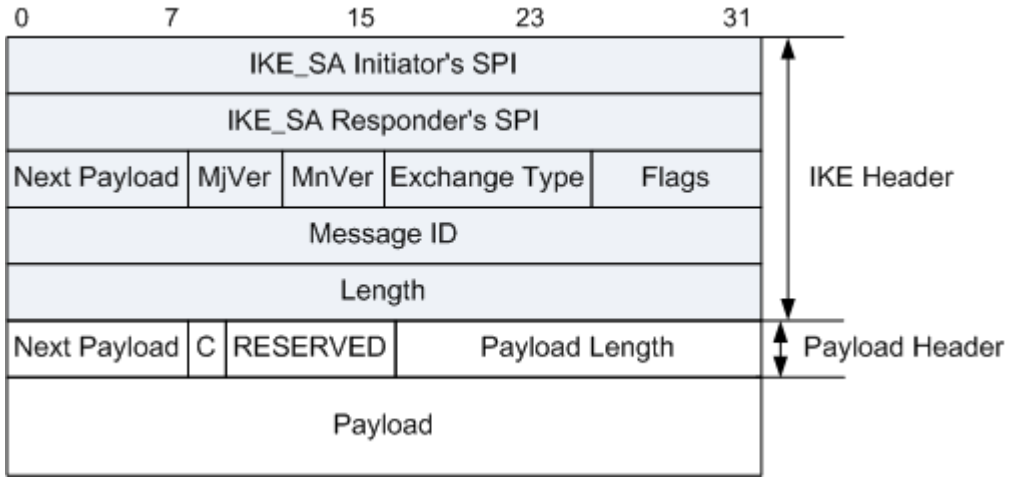
字段	长度	描述
Integrity Check Value-ICV	变长	验证数据。

图 3 AH 和 ESP 协议组合使用



因特网密钥交换协议 IKE（Internet Key Exchange）是 IPSEC 的信令协议。

图 4 IKE Header Format



字段	长度	描述
IKE_AS Initiator's SPI	8 bytes	发送者用来唯一标识一个 IKE 安全联盟，该值不能设置为 0。

字段	长度	描述
IKE_AS Responder's SPI	8 bytes	应答者用来唯一标识一个 IKE 安全联盟，对于 IKE 初始交互的消息该值必须为 0，其他消息不能为 0。
Next Payload	1 byte	仅随头部之后的负载的类型。
MjVer	4 bits	标识所使用的 IKE 协议的最大版本。
MnVer	4 bits	标识所使用的 IKE 协议的最小版本。
Exchange Type	1 byte	• 0-33: RESERVED • 34: IKE_SA_INIT • 35: IKE_AUTH • 36: CREATE_CHILD_SA • 37: INFORMATIONAL • 38-239: RESERVED TO IANA • 240-255: Reserved for private use.
Flags	1 byte	消息中设置的特定选项。如果 Flag 域置位表示带有选项。 • X(reserved) (bits 0-2) - 发送时必须清 0，接收时忽略。 • I(nitiator) (bit 3 of Flags) - IKE_SA 原始发送者在发送消息是必须将此位置 1，源回应者发送的消息必须清零。 • V(ersion) (bit 4 of Flags) - 标识转发者支持的版本比 Major 字段标识的版本更高 IKEv2 版本的实现中，此比特必须置 0，接收时忽略。 • R(esponse) (bit 5 of Flags) - 标识此消息是对相同 Message-ID 的消息的一个回应消息。所有请求消息中此位需置 0，所有回应消息置 1。 • X(reserved) (bits 6-7 of Flags) - 发送时需置 0，接收时忽略。
Message ID	4 bytes	消息标识符，用来对请求消息和呼应消息的匹配，以便控制丢弃消息的重复发送。这在抑制重放攻击时对保障协议的安全性很关键。
Length	4 bytes	整个消息的长度（报文头+负荷），以字节为单位。
Next Payload	1 byte	标识消息中的下一个负载的类型。如果当前的负载是消息的最后一个，则此字段置 0。

字段	长度	描述
		• 0: No Next Payload • 1-32: RESERVED • 33: Security Association (SA) • 34: Key Exchange (KE) • 35: Identification - Initiator (IDi) • 36: Identification - Responder (IDr) • 37: Certificate (CERT) • 38: Certificate Request (CERTREQ) • 39: Authentication (AUTH) • 40: Nonce (Ni, Nr) • 41: Notify (N) • 42: Delete (D) • 43: Vendor ID (V) • 44: Traffic Selector - Initiator (TSi) • 45: Traffic Selector - Responder (TSr) • 46: Encrypted (E) • 47: Configuration (CP) • 48: Extensible Authentication (EAP) • 49-127: RESERVED TO IANA • 128-255: PRIVATE USE

C (Critical)	1 bit	• 如果发送者想让接收者在无法识别当前一个负载的 Next Payload 域是能够跳过此域，可将此位置 0。
		• 如果接收者能够识别负载的类型代码，则忽略此位。
		• 负载类型为以下情况时，此位必须设置为 0。
		  ▪ Security Association (SA)
		  ▪ Key Exchange (KE)
		  ▪ Identification - Initiator (IDi)
		  ▪ Identification - Responder (IDr)
		  ▪ Certificate (CERT)
		  ▪ Certificate Request (CERTREQ)

字段	长度	描述
		▪ Authentication (AUTH) ▪ Nonce (Ni, Nr) ▪ Notify (N) ▪ Delete (D) ▪ Vendor ID (V) ▪ Traffic Selector - Initiator (TSi) ▪ Traffic Selector - Responder (TSr) ▪ Encrypted (E) ▪ Configuration (CP) ▪ Extensible Authentication (EAP) 注意，C 比特应用于当前负载，而不是下一个负载。
RESERVED	7 bits	发送时必须置 0，接收时忽略。
Payload Length	2 bytes	当前负载的长度，包括通用负载的头部，以字节为单位。

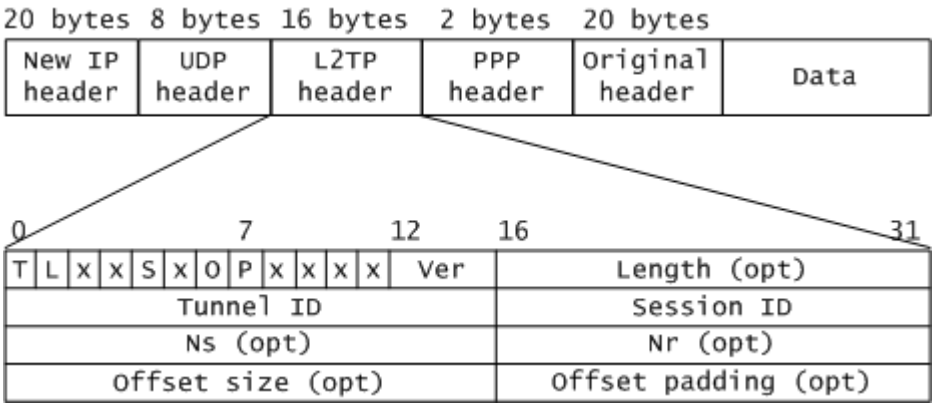
参考标准

标准	描述
RFC 4302	IP Authentication Header
RFC 4303	IP Encapsulating Security Payload (ESP)
RFC 4306	Internet Key Exchange (IKEv2) Protocol

6.11 L2TP 报文格式

报文格式

L2TP 的控制消息和数据消息使用相同的报文头。



L2TP 报文头中标记为可选（opt）的字段，是指在数据消息中可选，在控制消息中则是必选的。

字段	长度	描述
T	1 比特	类型（Type），取值为“0”时表示数据消息，取值为“1”时表示控制消息。
L	1 比特	长度在位标志，取值为“1”时表示报文头中存在长度字段 Length。控制消息中必须为“1”。
x	1 比特	保留位
S	1 比特	顺序字段在位标志，取值为“1”时表示报文头中存在 Ns 和 Nr 字段。控制消息中必须为“1”。
O	1 比特	取值为“1”时表示报文头中存在 offset size 字段。控制消息中必须为“0”。
P	1 比特	优先级（Priority），只用于数据消息。控制消息中必须为“0”。
Ver	4 比特	版本号，对于 L2TPv2 协议取值为“2”。
Length	16 比特	消息的总长度，单位为字节。
Tunnel ID	16 比特	隧道标识符，只具有本地意义。Hello 控制消息具有全局性，其 Tunnel ID 必须为 0。
Session ID	16 比特	会话标识符，只具有本地意义。

字段	长度	描述
Ns	16 比特	当前消息的顺序号。
Nr	16 比特	希望接收的下一条控制消息的顺序号。数据消息中是保留字段。
Offset size	16 比特	偏移值，指示载荷数据开始的位置。
Offset padding	16 比特	填充位。

Packet Example

图 1 L2TP Control Message

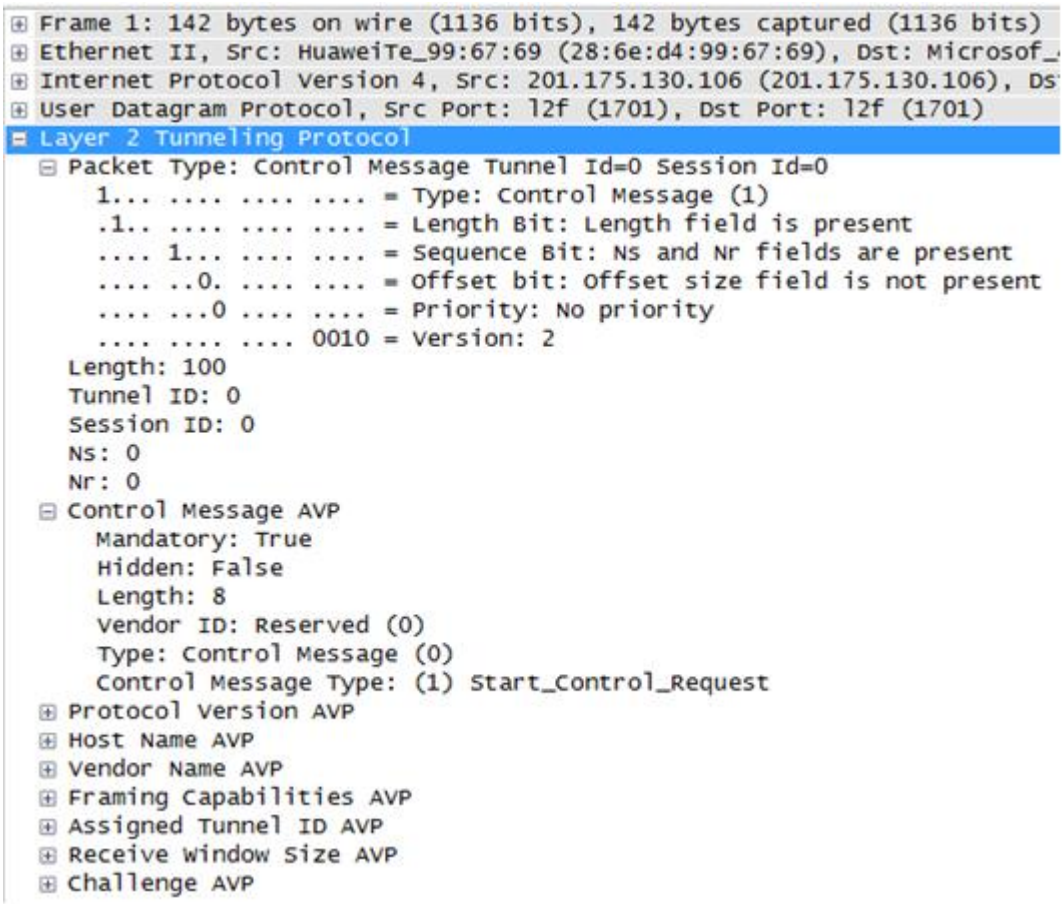


图 2 L2TP Data Message

```
⊞ Frame 148: 1080 bytes on wire (8640 bits), 1080 bytes captured (8640 bits)
⊞ Ethernet II, Src: HuaweiTe_99:67:69 (28:6e:d4:99:67:69), Dst: Microsof_40
⊞ Internet Protocol Version 4, Src: 201.175.130.106 (201.175.130.106), Dst:
⊞ User Datagram Protocol, Src Port: l2f (1701), Dst Port: l2f (1701)
⊟ Layer 2 Tunneling Protocol
  ⊟ Packet Type: Data    Message Tunnel Id=3 Session Id=1
      0... .... .... .... = Type: Data Message (0)
      .0.. .... .... .... = Length Bit: Length field is not present
      .... 0... .... .... = Sequence Bit: Ns and Nr fields are not present
      .... ..0. .... .... = Offset bit: Offset size field is not present
      .... ...0 .... .... = Priority: No priority
      .... .... .... 0010 = Version: 2
    Tunnel ID: 3
    Session ID: 1
⊞ Point-to-Point Protocol
⊞ Internet Protocol, Src: 192.168.0.3 (192.168.0.3), Dst: 192.168.0.1 (192.
⊞ Internet Control Message Protocol
```

参考标准

标准	描述
RFC 2661	Layer Two Tunneling Protocol "L2TP"

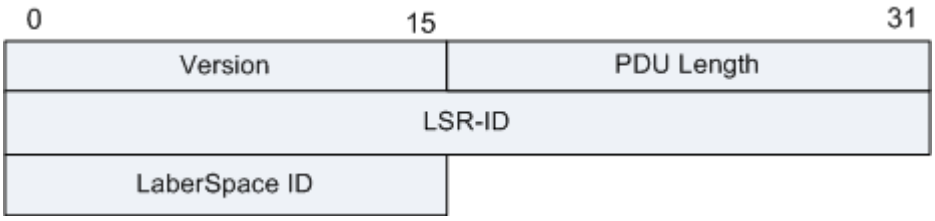
6.12 MPLS LDP 报文格式

LDP 消息头部格式

为保证 LDP 消息的可靠发送，除了 Discovery（Hello）消息使用 UDP（端口 646）外，LDP 的 Session 消息、Advertisement 消息和 Notification 消息都使用 TCP（端口 646）传输。

LDP 协议消息头部格式如下（PDU，协议数据单元，每个 LDP PDU 有个 LDP 消息头，后面跟着一个或多个 LDP 消息内容）。

图 1 LDP 协议消息头部格式



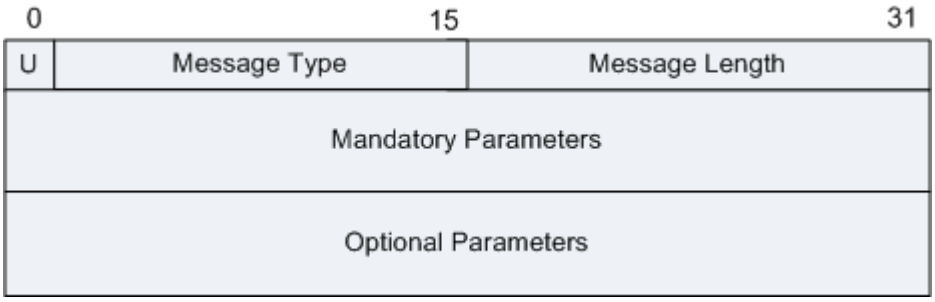
字段	长度	说明
Version	2 字节	表示版本号。目前 LDP 的版本号始终为 1。
PDU Length	2 字节	表示 PDU 的总长度，包括 LDP ID 和整组 LDP 消息，不包括 Version 和 PDU Length 字段。 例如某个 LDP 报文中包含 3 个 Hello 消息，则该报文的 PDU length = 3 * Message length。

字段	长度	说明
LSR-ID	4 字节	LDR-ID 标识一台 LSR，必须全局唯一。
LaberSpace ID	2 字节	标识了 LSR 内的标签空间。对于平台范围标签空间，这些数值都应当为 0。
Bunch of messages	变长	是一组 LDP 消息的集合，可以是一个或者多个 LDP 消息。 • 当 LDP 报文以 UDP 方式传输时，“Bunch of messages”只能是 Hello 消息的集合。 • 当 LDP 报文以 TCP 方式传输时，“Bunch of messages”可以是除 Hello 消息外任意类型的 LDP 消息的集合。

LDP 消息格式

所有 LDP 消息的格式如下：

图 2 LDP 消息格式



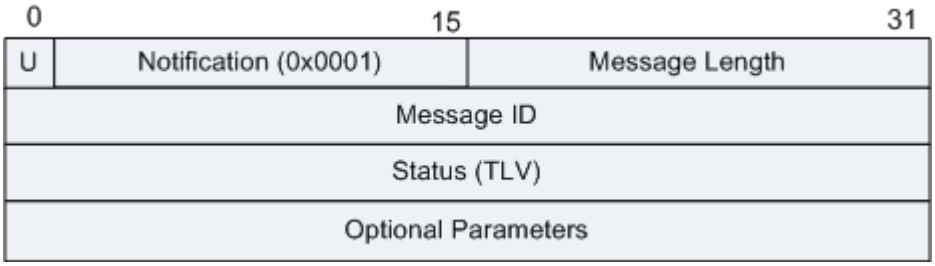
字段	长度	说明
U	1 比特	未知的消息。 • 如果对端收到的 LDP 消息中“Message type”字段为未知的 LDP 消息类型，且“U”字段的值为“0”，则向源端发送通知（Notification）消息。 • 如果对端收到的 LDP 消息中“Message type”字段为未知的 LDP 消息类型，且“U”字段的值为“1”，则忽略该未知消息。
Message Type	15 比特	LDP 消息的类型。
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Mandatory Parameters	变长	LDP 消息的强制参数。

字段	长度	说明
Optional Parameters	变长	LDP 消息的可选参数，包含 0~n 个 TLV。

通告（Notification）消息

LSR 发送通告消息来通知重要事件到 LDP 对等体。通告消息通知致命错误或提供咨询信息，如处理 LDP 消息的结果或 LDP 会话的状态。

图 3 Notification 消息格式

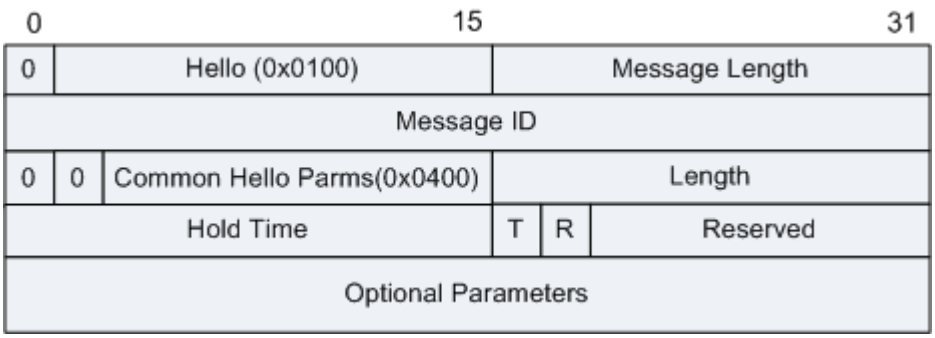


字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Status TLV	变长	标识一个事件。
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

Hello 消息

用于通告和维护网络中 LSR 的存在。

图 4 Hello 消息格式

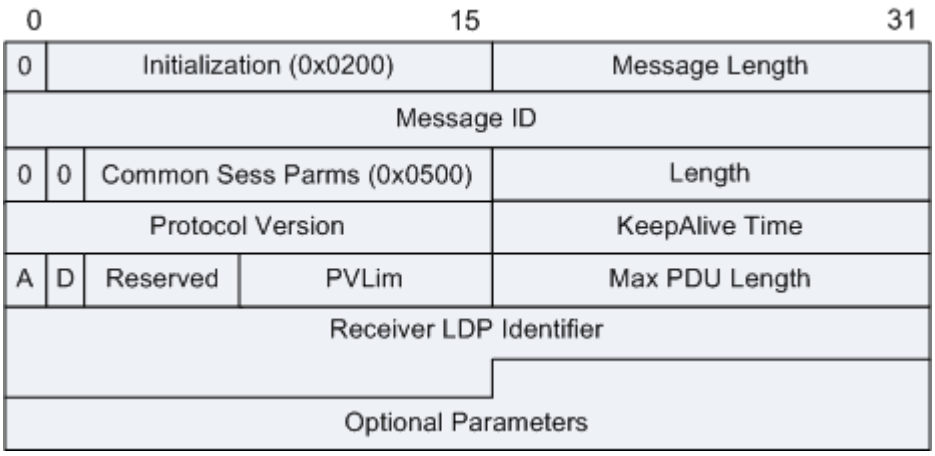


字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。
Hold Time	16 比特	保持时间，以秒为单位的 Hello 保持时间。LSR 维护来自潜在同伴的 Hello 的记录。为 0 的数值意味着使用缺省值。0xffff 的数值意味着无穷大。
T	1 比特	T（Targeted Hello）值为 1 表示为远端 Hello 消息，值为 0 表示本地 Hello 消息。
R	1 比特	R（Request Send Targeted Hellos）值为 1 表示请求接收者周期性发送远端 Hello 消息给该 Hello 的发送源端，值为 0 表示没有此需求。
Reserved	14 比特	保留字段，必须置 0，接收端忽略此字段。
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

Initialization 消息

LDP 的 Initialization 消息在 LDP 回家建立阶段发送，格式如下：

图 5 Initialization 消息格式



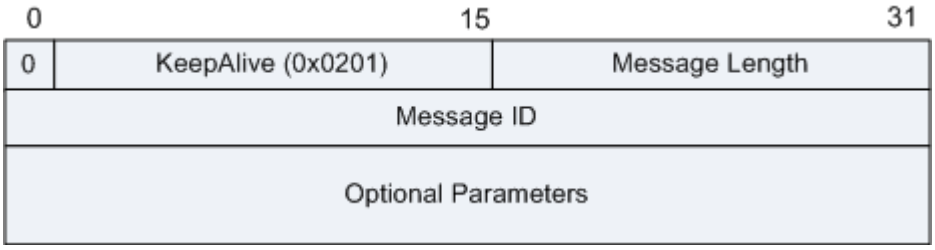
字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。
Protocol Version	16 比特	表示协议版本号 = 1。
KeepAlive Time	16 比特	TCP 连接的保持时间，这个定时器的刷新并不是收到 KeepAlive 消息才会刷新，而是通过 TCP 连接收到的 LDP PDU 时都会刷新
A	1 比特	表明标签分配方式 (0 = DU ; 1 = DoD)
D	1 比特	表明是否使能了环路检测功能(0 = Disable ; 1 = Enable)。
Reserved	6 比特	保留字段，必须置 0，接收端忽略此字段。
PVLim	8 比特	PVLim (Path Vector Limit)，LSP 支持的最大跳数(只有在使能了 Loop detection 功能时有效，默认值为 32)。
Max PDU Length	16 比特	LDP PDU 的最大长度，默认值为 4096 字节。
Receiver LDP Identifier	6 字节	Initialization 消息的接收者的 LDP 标识符(LDP ID)。

字段	长度	说明
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

KeepAlive 消息

Keepalive 消息无 Mandatory Parameters 字段及后面的字段，用于维护 SESSION 的状态，所以这里不需要什么特别的内容，只要对方知道自己还存在就好。

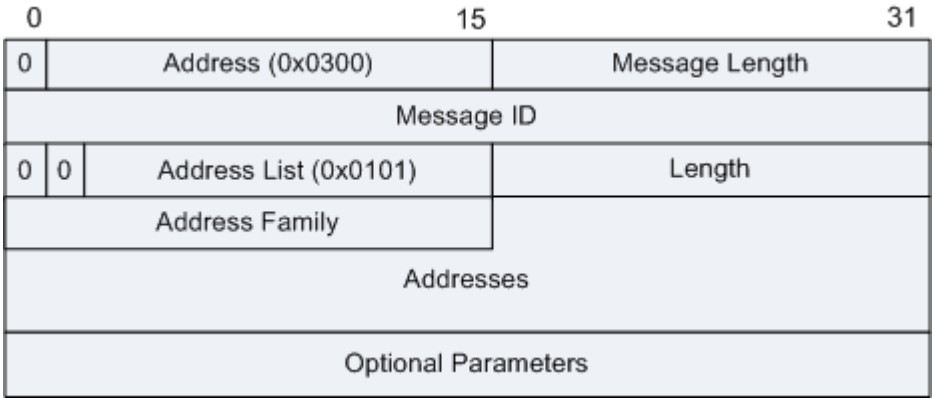
图 6 KeepAlive 消息格式



地址（Address）消息

Address 消息用于 LSR 发送地址消息到 LDP 邻居，以公告其接口地址。

图 7 Address 消息格式



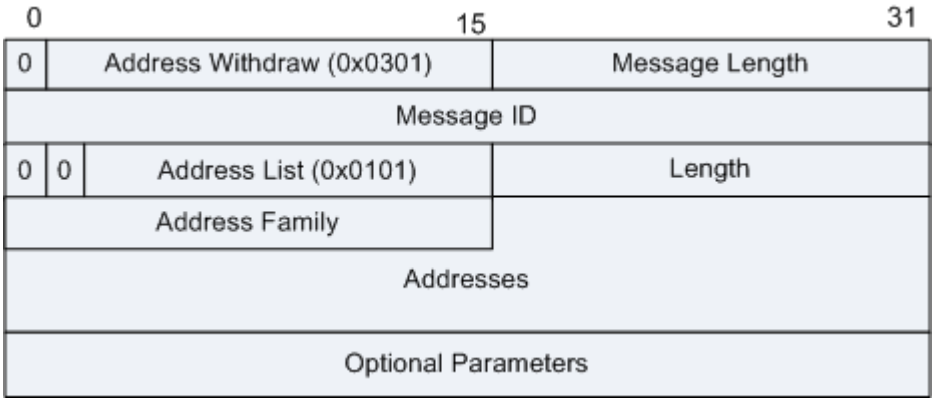
字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。

字段	长度	说明
Address Family	16 比特	地址族编号。
Addresses	变长	指定地址族的地址列表，格式取决于地址族类型： • IPv4 地址为 4 字节 • IPv6 地址为 16 字节
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

地址撤销（Address Withdraw）消息

LSR 发送 Address Withdraw 消息到 LDP 对等体，以撤销之前公告的接口地址。当接口地址被删除或接口 down 后，就会发送 Address Withdraw 消息。

图 8 Address Withdraw 消息格式



字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。
Address	16 比特	地址族编号。

字段	长度	说明
Family		
Addresses	变长	指定地址族的地址列表，格式取决于地址族类型： • IPv4 地址为 4 字节 • IPv6 地址为 16 字节
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

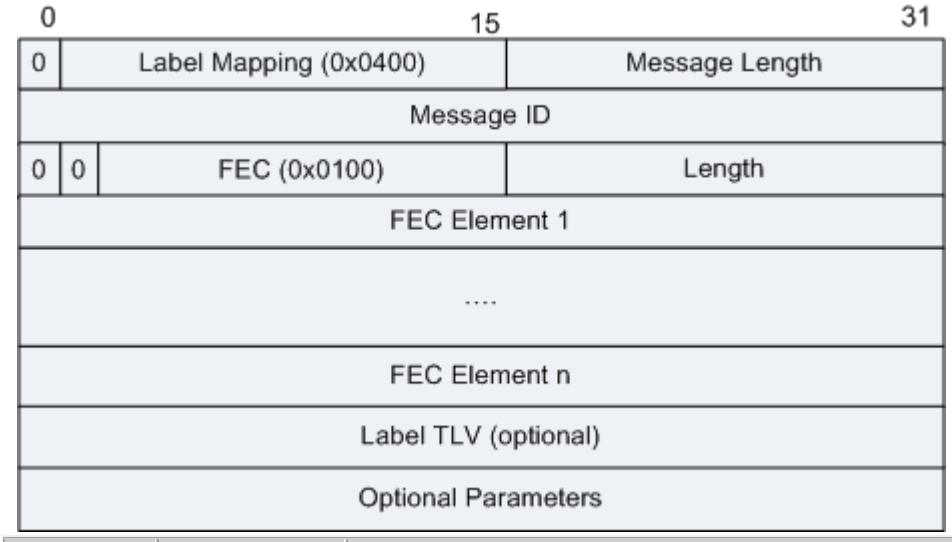
Label Mapping 消息

对于特定的 FEC，下游使用 Label Mapping 消息为上游分配标签。

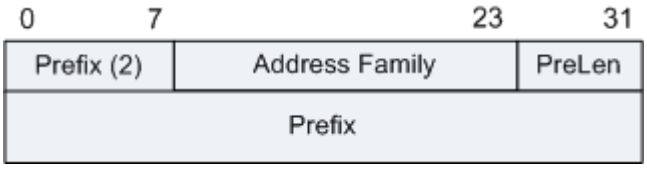
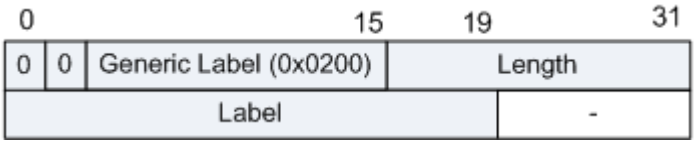
Label Mapping 消息均由下游发往上游节点，Label Mapping 消息的发送方式因标签分配方式或者标签控制方式的不同而不同：

- DU 模式下：下游无需等待上游的 Label Request 消息可以直接为某 FEC 向上游 LDP 邻居发送 Label Mapping 消息；
- DoD 模式下：下游必须等待上游的 Label Request 消息才能为指定的 FEC 向上游 LDP 邻居发送 Label Mapping 消息；
- Independent 模式下：中间节点无需等待收到下游为指定 FEC 发送的 Label Mapping 消息后才向它的上游 LDP 邻居发送 Label Mapping 消息；
- Order 模式下：中间节点必须等待收到下游为指定 FEC 发送的 Label Mapping 消息后才能向它的上游 LDP 邻居发送 Label Mapping 消息。

图 9 Label Mapping 消息格式



字段	长度	说明
Message	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。

字段	长度	说明
Length		
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。
FEC Element 1 to FEC Element n	每个 Element 1 字节	表明该标签是为哪个 FEC 而分配的。 FEC Element 的格式定义： • Type = 0x01: 反掩码。只在 Label Withdraw 和 Label Release 消息中使用。 • Type = 0x02: 前缀。 格式如下： **图 10** 前缀 FEC Element 的格式 
Label TLV	52 比特	表明下游为该 FEC 分配了什么标签。 **图 11** Label TLV 的格式  Label 字段是个 20 比特的标签值。
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

Label Request 消息

LSR 发送 Label Request 消息给 LDP 对等体请求 FEC 对应的标签。

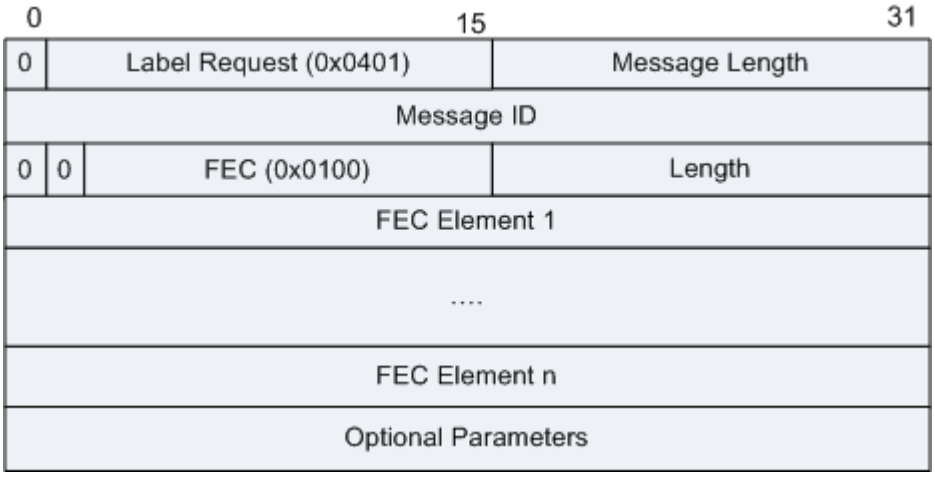
Label Request 消息只能由作为 Ingress 的 LSR 发起，不能由作为 Transit 的 LSR 发起（Order＋dod）。下游收到 Label Request 消息后需要向上游返回 Label Mapping 消息，Label Mapping 消息中携带 Label Request 消息中的 Message ID TLV。

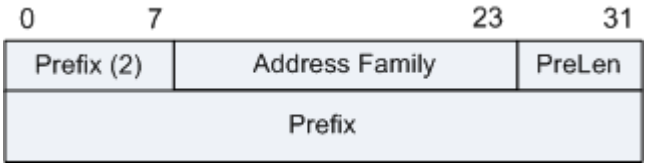
下列情况下，下游收到 Label Request 消息后不会向上游返回 Label Mapping 消息：

- 下游找不到与 Label Request 中 FEC TLV 中对应的路由，此时返回 No Route 的 Notification 消息；
- 下游没有足够的标签来分配，此时返回 No Label Resource 的 Notification 消息；

- 下游检测到有环路发生，此时返回 Loop Detected 的 Notification 消息。

图 12 Label Request 消息格式

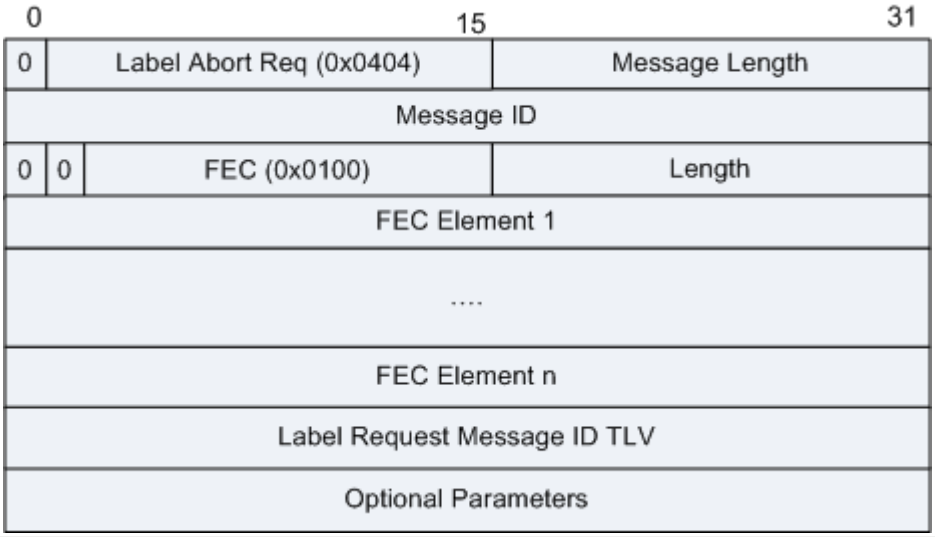


字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。
FEC Element 1 to FEC Element n	每个 Element 1 字节	表明该标签是为哪个 FEC 而分配的。 FEC Element 的格式定义： • Type = 0x01: 反掩码。只在 Label Withdraw 和 Label Release 消息中使用。 • Type = 0x02: 前缀。 格式如下： **图 13** 前缀 FEC Element 的格式 
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

Label Abort Request 消息

上游 LSR 发送了 Label Request 消息后但还没有收到 Label Mapping 消息前，发现 FEC 对应的下一跳变化了或者其他可能的原因需要发送新的 Label Request 消息时，上游会向下游发送 Label Abort Request 消息。

图 14 Label Abort Request 消息格式



字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。
FEC Element 1 to FEC Element n	每个 Element 1 字节	表明该标签是为哪个 FEC 而废弃的。 FEC Element 的格式定义： • Type = 0x01: 反掩码。只在 Label Withdraw 和 Label Release 消息中使用。 • Type = 0x02: 前缀。 **图 15** 前缀 FEC Element 的格式 
Label Request Message ID TLV	-	要被废弃的 Label Request 消息的消息 ID。
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

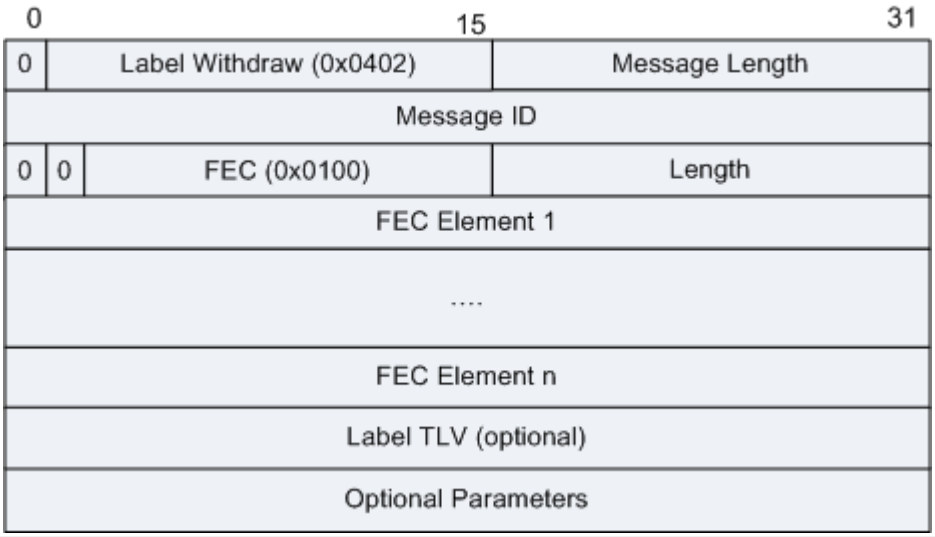
Label Withdraw 消息

Label Withdraw 消息一般由下游 LSR 发往上游 LSR，通知上游 LSR 之前通告的与某 FEC 对应的 Label 不再使用，上游 LSR 需要解除 Label 和 FEC 的映射关系。

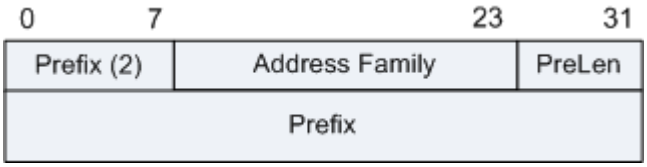
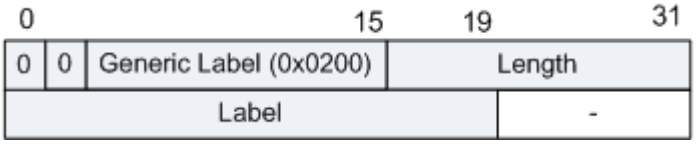
下列情况下会发送 Label Withdraw 消息：

- 下游节点不再有某条 FEC，如果已经为该 FEC 发送了 Label Mapping 消息，则发送 Label Withdraw 消息；
- 下游单方面的决定不再使用标签转发时也会发送 Label Withdraw 消息。

图 16 Label Withdraw 消息格式



字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。
FEC Element 1 to FEC Element n	每个 Element 1 字节	表明该标签是为哪个 FEC 而撤销的。 FEC Element 的格式定义： • Type = 0x01: 反掩码。只在 Label Withdraw 和 Label Release 消息中使用。 • Type = 0x02: 前缀。 **图 17** 前缀 FEC Element 的格式

字段	长度	说明
		
Label TLV	52 比特	表明该 FEC 对应的标签。 **图 18** Label TLV 的格式 
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

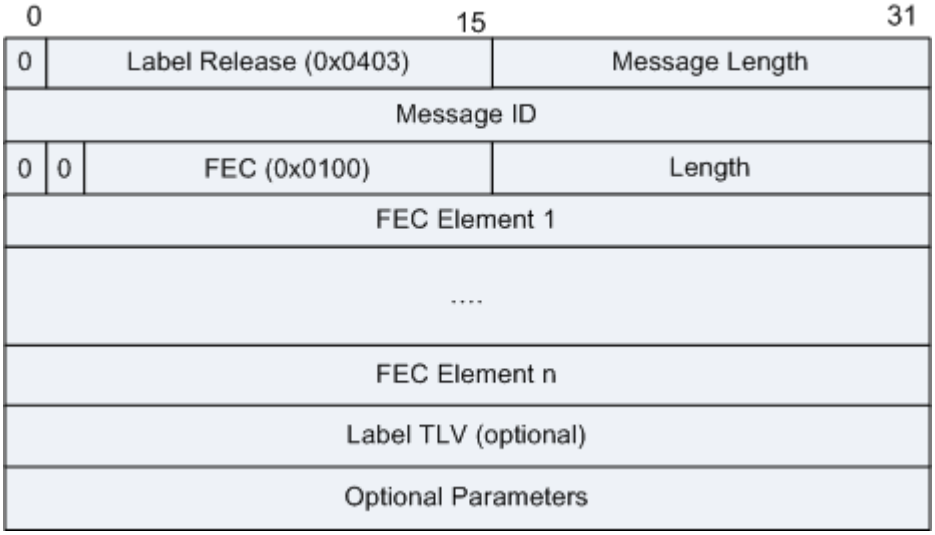
Label Release 消息

Label Release 消息一般由上游发往下游，通知撤销 Label 和 FEC 的邦定，该消息相当于 Label Request 消息的逆过程。

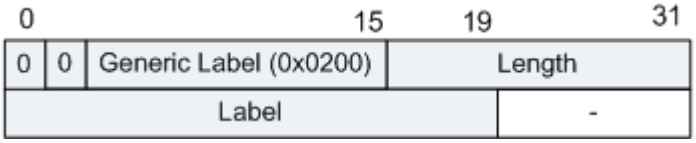
在下列情况下会发送 Label Release 消息：

- 上游 LSR 的标签保持方式是保守方式，发送 Label Mapping 消息的 LSR 不再是 FEC 的下一跳时，上游 LSR 需要发送 Label Release 消息来撤销 Label 和 FEC 的映射关系；
- 上游 LSR 的标签保持方式是保守方式，从不是 FEC 的下一跳收到 Label Mapping 消息后，上游 LSR 需要发送 Label Release 消息；
- LSR 收到 Label Withdraw 消息后需要发送 Label Release 消息。

图 19 Label Release 消息格式



字段	长度	说明

字段	长度	说明
Message Length	16 比特	LDP 消息的长度，是 Message ID、强制参数和可选参数的长度的总和。
Message ID	32 比特	LDP 消息的编号，用于唯一地标识一个 LDP 消息。
Length	16 比特	LDP 消息是以 LDP PDU 中的 TLV 形式定义的. 每个 LDP TLV 有一个 2 字节的 Type 域，2 字节的 Length 域和变长的 Value 域。这里的 Length 就表示 TLV 的 Value 域的字节数。
FEC Element 1 to FEC Element n	每个 Element 1 字节	表明该标签对应的 FEC。 FEC Element 的格式定义： • Type = 0x01: 反掩码。只在 Label Withdraw 和 Label Release 消息中使用。 • Type = 0x02: 前缀。 **图 20** 前缀 FEC Element 的格式 
Label TLV	52 比特	**图 21** Label TLV 的格式 
Optional Parameters	变长	可选参数，包含 0~n 个 TLV。

LDP 消息示例

图 22 LDP Hello 消息

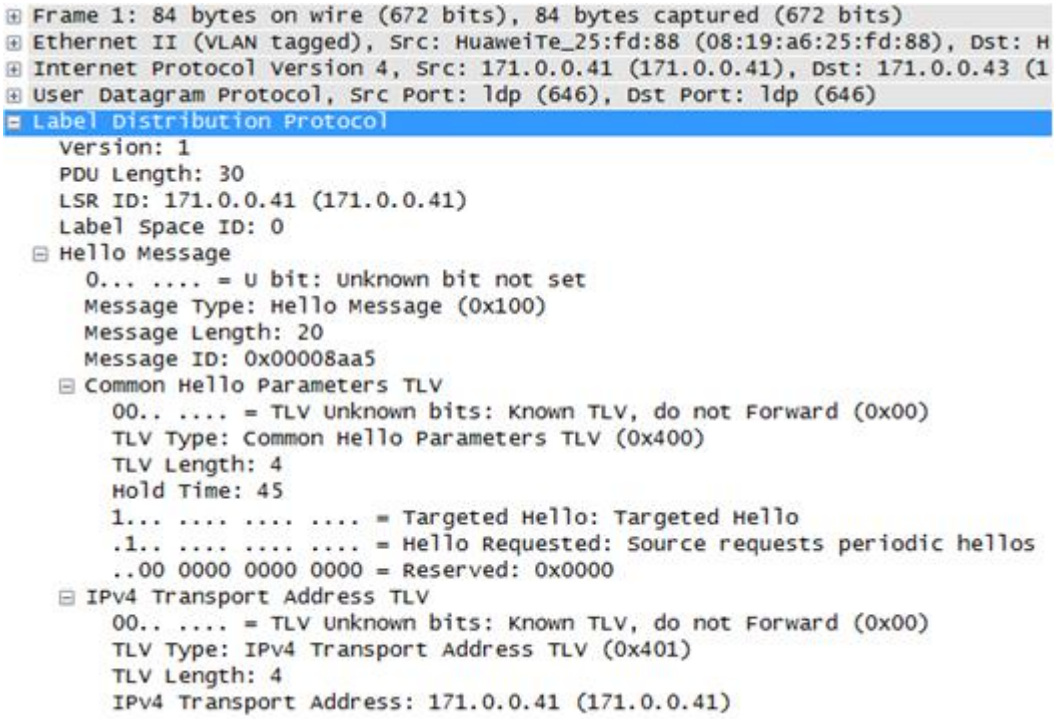
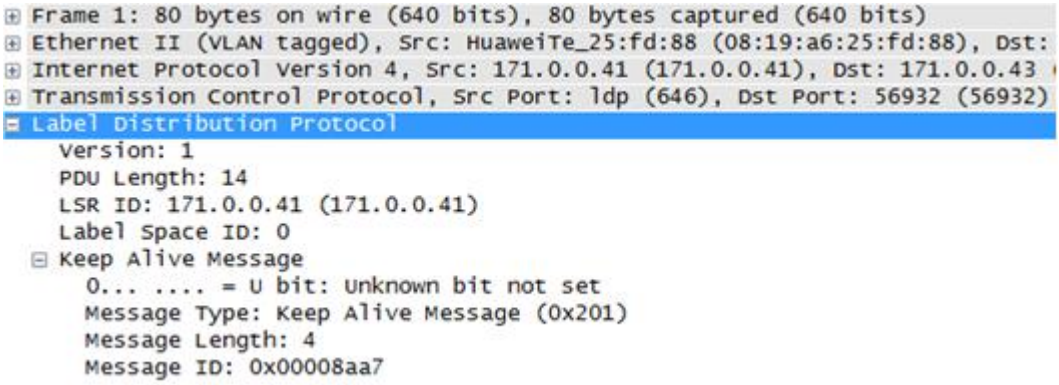


图 23 LDP KeepAlive 消息



参考标准

标准	描述
RFC 5036	LDP Specification

6.13 MSDP 报文格式

MSDP（Multicast Source Discovery Protocol）称为组播源发现协议，是基于多个 PIM-SM（Protocol Independent Multicast Sparse Mode）域互连而开发的一种域间组播解决方案。

适用条件：域内组播路由协议必须是 PIM-SM。MSDP 仅对 ASM（Any-Source Multicast）模型有意义。

报文格式

MSDP 支持四种消息，都符合标准的 TLV（Type-Length-Value）消息格式，通过 TCP 连接交互信息。

字段	长度	说明
Type	8 比特	消息类型 • 1: Source-Active，携带多组（S，G）信息，在多个 RP 之间传递，或者封装 PIM-SM 组播数据。 • 2: Source-Active Request，请求指定组 G 的（S，G）列表，减少源加入延迟 • 3: Source-Active Response，对 Source-Active Request 消息的响应 • 4: KeepAlive，保持 MSDP 对等体的连接关系 • 5: Reserved • 6: MSDP traceroute in progress • 7: MSDP traceroute reply
Length	16 比特	消息长度，包含 Type、Length 和 Value 字段的长度，字节为计数单位。除了 Keepalive 消息外，其他消息要求最小长度为 4 字节。最大长度为 9192 字节。
Value	Variable length	消息内容，因消息类型而异。

图 1 MSDP SA 控制消息的格式

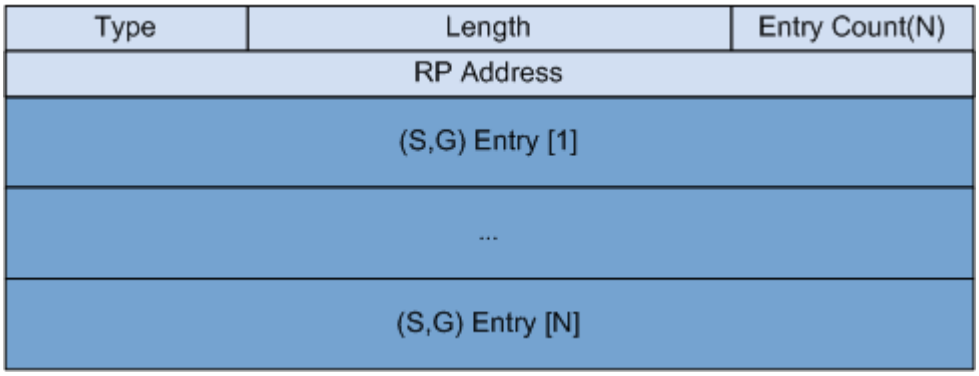
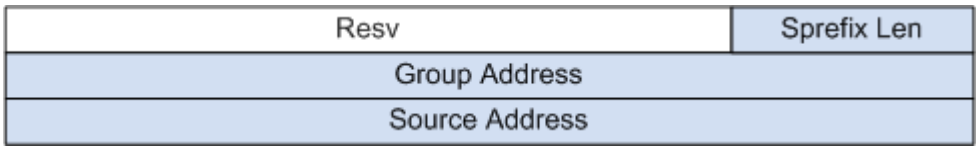


图 2 （S，G）Entry 字段格式

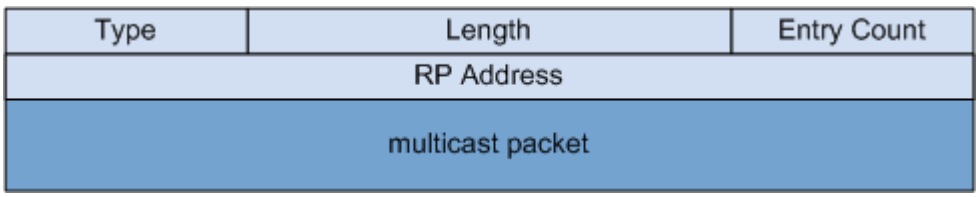


The maximum size SA message that can be sent is 9192 octets. The 9192 octet size does not include the TCP, IP, layer-2 headers.

字段	长度	说明
Type	1 字节	消息类型，值为 1。

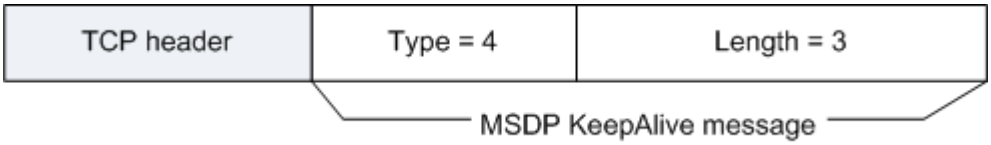
字段	长度	说明
Length	2 字节	整个 TLV 的长度。
Entry Count	1 字节	消息中包含的（S，G）项总数。
RP Address	4 字节	源 RP 地址。
（S，G）Entry [1]...[N]	变长	（源，组）信息。
Resv	3 字节	保留字段，发送报文时置 0；接收到报文时，对该字段不做任何处理。
Sprefix Len	1 字节	源地址掩码长度，以 32 位传输。
Group Address	4 字节	组地址。
Source Address	4 字节	组播源地址。

图 3 MSDP SA 数据消息的格式



MSDP SA 数据消息的 Entry Count 值为 1。

图 4 MSDP KeepAlive 消息的格式



参考标准

标准	描述

标准	描述
RFC 3618	Multicast Source Discovery Protocol (MSDP)

6.14 NetStream 报文格式

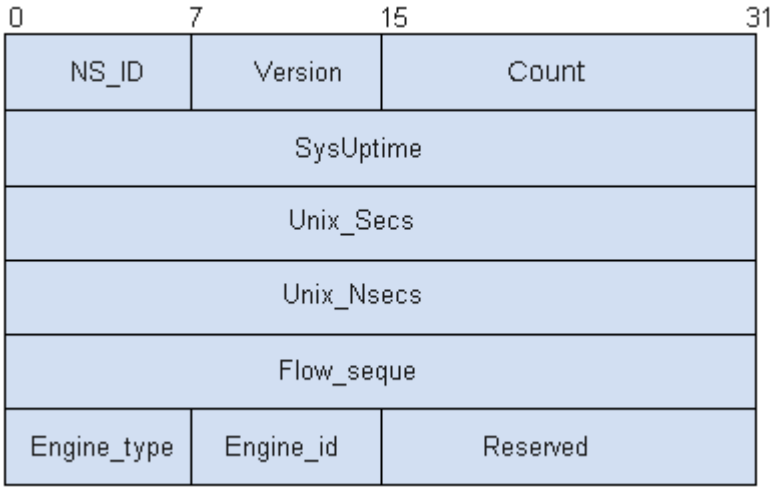
NetStream 是华为公司的专利技术，是一种基于网络流信息的统计与发布技术。NetStream 可以对网络中的通信量和资源使用情况进行分类和统计，实现对各种业务和不同的 QoS 进行管理和计费。

目前 Netstream 输出的报文主要有 5、8、9 三个版本，其他的版本处于实验阶段，没有商用。所有的版本都是通过 UDP 协议传递统计信息的。每个数据包都包括一个 Packet Header 再加上一条或者几条流的记录信息。

NetStream 原始流输出报文支持版本 5 和版本 9 两种报文格式，聚合流输出支持版本 8 和版本 9 两种报文格式。

NetStream 版本 5 报文头格式

图 1 版本 5 报文头格式



字段	长度	描述
NS_ID	1 字节	NetStream 的标识位，第 7 个比特位 0，表示入接口统计报文，第 7 个比特位 1，表示出接口统计报文
version	1 字节	NetStream 输出报文格式版本编号，对于 V5，为 0x05。

字段	长度	描述
count	2 字节	当前报文中的流记录数（1-30）
SysUptime	4 字节	报文产生的时间，是系统启动以来的毫秒数
unix_secs	4 字节	从 1970 年 1 月 1 日 0 时起，到报文产生时间的整秒数
unix_nsecs	4 字节	报文产生时间的纳秒数，也即不足一秒的余下的纳秒数
flow_sequence	4 字节	输出的流记录的顺序号 在第一个 NetStream 报文中，此值为 0，count = c1， 在第二个 NetStream 报文中，此值为 c1，count = c2， 在第三个 NetStream 报文中，此值为 c2 + c1， … 在第 n - 1 个 NetStream 报文中，此值为 fs(n - 1)，count = c(n - 1) 在第 n 个 NetStream 报文中，此值为 fs(n - 1) + c(n - 1)。 利用此值可以判断报文是否丢失。 当流序列号溢出时，按自然溢出继续进行。
engine_type	1 字节	流交换引擎类型
engine_id	1 字节	交换引擎槽号
reserved	2 字节	保留字段，全零

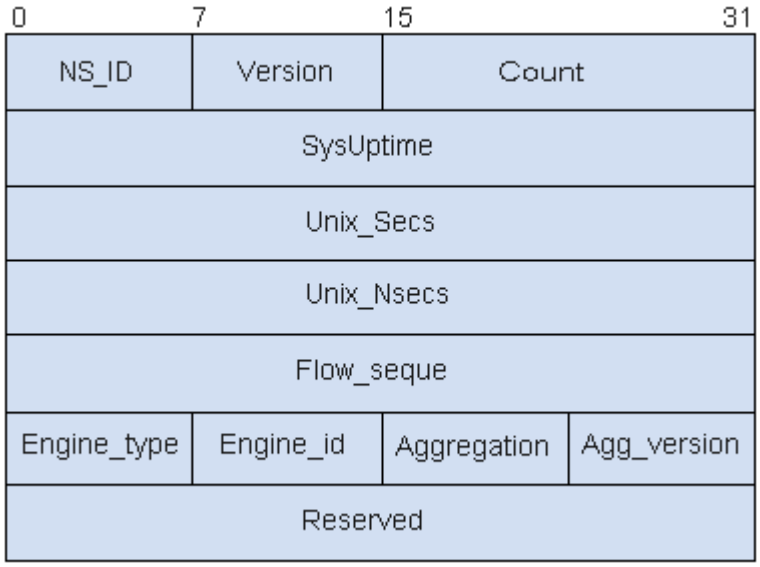
版本 5 包括以下信息：

- 时间信息：流建立的时间、流结束的时间

- 统计信息：包计数、字节计数
- 协议信息：协议类型
- 路由信息：目的 IP、源 IP、下一跳 IP、目的 IP 掩码长度、源 IP 掩码长度、源 AS 域编号、目的 AS 域编号
- 接口信息：入接口、出接口
- 传输层信息：源端口号、目的端口号、TCP Flags
- 服务等级信息：ToS

NetStream 版本 8 报文头

图 2 版本 8 报文头格式



字段	长度	描述
NS_ID	1 字节	NetStream 的标识位
version	1 字节	NetStream 输出报文格式版本编号
count	2 字节	当前报文中的流记录数，而不是流的总数
SysUptime	4 字节	报文产生的时间，是系统启动以来的毫秒数
unix_secs	4 字节	从 1970 年 1 月 1 日 0 时起，到报文产生时间的整秒数
unix_nsecs	4 字节	报文产生时间的纳秒数，也即不足一秒的余下的纳秒数

字段	长度	描述
flow_sequence	4 字节	输出的流记录的顺序号，
engine_type	1 字节	流交换引擎类型
engine_id	1 字节	交换引擎槽号
aggregation	1 字节	聚合策略，分别如下： • AS：0x01 • Protocol-Port：0x02 • rc-Prefix：0x03 • Dst-Prefix：0x04 • SPrefix：0x05
Agg_version	1 字节	聚合版本
Reserved	4 字节	保留字段，全零

版本 8 通过 AS 域聚合，这种聚合方式主要统计在一个路由器上从一个 AS 域到另一个 AS 域的包和字节信息。可以用以运营商之间结算。每个记录包括以下信息：

- 时间信息：流建立时间、流结束时间
- 统计信息：包计数、字节计数
- 接口信息：入接口、出接口
- AS 自治域信息：源 AS 域编号、目的 AS 域编号
- 构成聚合流的流总数：聚合流的总数

通过协议类型聚合，这种方式针对四层的协议类型（对于 TCP 和 UDP，还包括源和目的端口号）进行聚合。每个记录包括：

- 时间信息：流建立时间、流结束时间
- 统计信息：包计数、字节计数
- 协议信息：协议类型、源端口号、目的端口号
- 构成聚合流的流总数：聚合流的总数

根据源 IP 和目的 IP 的前缀进行聚合，这种方式下，源 IP 地址的前缀部分和目的 IP 地址的前缀部分都参与聚合。每个记录包括：

- 时间信息：流建立时间、流结束时间
- 统计信息：包计数、字节计数
- 接口信息：入接口、出接口
- IP 地址信息：源 IP 前缀、源 AS 域编号、目的 IP 前缀、目的 AS 域编号
- 构成聚合流的流总数：聚合流的总数

TOS＋AS 域的方式进行流聚合，每个记录包括：

- 时间信息：流建立时间、流结束时间
- 统计信息：包计数、字节计数
- 接口信息：入接口、出接口
- ToS + AS 信息：IP 头的 Tos、源 AS 域编号、目的 AS 域编号
- 构成聚合流的流总数：聚合流的总数

按照 TOS 加协议类型的聚合，每个记录包括：

- 时间信息：流建立时间、流结束时间
- 统计信息：包计数、字节计数
- 接口信息：入接口、出接口
- ToS + 协议信息：IP 头的 Tos、协议类型、源端口号、目的端口号
- 构成聚合流的流总数：聚合流的总数

IP 地址前缀＋TOS＋协议类型的聚合，每个记录包括：

- 时间信息：流建立时间、流结束时间
- 统计信息：包计数、字节计数
- 接口信息：入接口、出接口
- IP 信息：IP 头的 Tos、源前缀、目的前缀
- 协议信息：协议类型、源端口号、目的端口号
- 构成聚合流的流总数：聚合流的总数

TOS＋源 IP 前缀的聚合，每个记录包括：

- 时间信息：流建立时间、流结束时间
- 统计信息：包计数、字节计数
- ToS + 源 IP 信息：IP 头的 Tos、源前缀、源 AS 域编号、入接口

- 构成聚合流的流总数：聚合流的总数

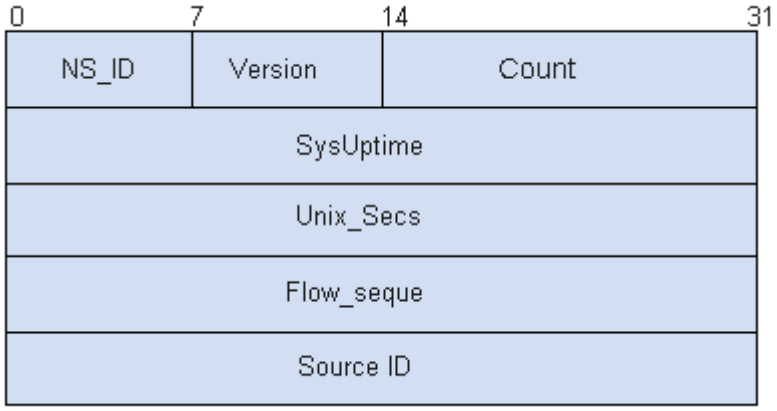
TOS＋目的 IP 前缀的聚合（和 TOS＋源 IP 前缀的聚合方式的方式类似），每个记录包括：

- 时间信息：流建立时间、流结束时间
- 统计信息：包计数、字节计数
- 接口信息：入接口、出接口
- ToS + 目的 IP 信息：IP 头的 Tos、目的前缀、源前缀
- 构成聚合流的流总数：聚合流的总数

版本 9 报头格式

版本 9 最显著的特点是基于模板的方式，使统计信息的输出更为灵活，而且更容易扩展新的定义流的元素以及生成新的记录。使用版本 9 可以实现 NAT、组播、MPLS、BGP 下一跳的统计。

图 3 版本 9 报头格式



字段	长度	描述
NS_ID	1 字节	NetStream 的标识位
version	1 字节	NetStream 输出报文格式版本编号
count	2 字节	该报文包含的 FlowSet records(包括 Template 和 Data)数目
SysUptime	4 字节	报文产生的时间，是系统启动以来的毫秒数
UNIX Seconds	4 字节	从 1970 年 1 月 1 日 0 时起，到报文产生时间的整秒数

字段	长度	描述
Sequence Number	4 字节	所有输出报文的顺序号
Source ID	4 字节	用来保证从一台路由器中输出的所有流的唯一性(Source ID 等同于 V5、V8 报文头中的 engine type 和 engine ID)。该值可以由用户定义。

版本 9 报文有两种输出情况：

- Export Packet 中仅有 Template FlowSet。在用户使能 NetStream 功能，对流量进行采集时，为了使网管能够正确的解析流量，系统会向 NSC 发送模板。为了保证网管可靠的对接收到流量统计信息进行解析，设备会定时重新发送模板给 NSC。另外，模板具有有效时间，超过有效时间 NSC 会删除超时的模板，因此，需要定时的发送 Template FlowSet 到 NSC，如果需要发送的时候没有 Data FlowSet 生成，则此时只发送 Template FlowSet。报文格式如下：

- Export Packet 中仅有 Data FlowSet。如果 Template ID 都已经定义好了，使能 NetStream 的路由器传递给 NSC 的 Export Packet 一般属于这种情况。报文格式如下：

Template FlowSet 和 Data FlowSet 是独立的。Data FlowSet 中的 Data Record 由 collector 已知的模板解释（换句话说，NSC 已经知道了 Data Record 中的 Template ID 对应的模板了）。而 Template FlowSet 是告诉 NSC 一个即将被使用的模板，NSC 使用这个模板的时候只能是针对后续的 Export Packet。

Template FlowSet 是版本 9 的灵魂。使用模板后，NSC 的程序无需预先设置好按照什么样的格式解析 Export Packet，只需做成通用的方式，然后通过路由器发过来的模板来解释流记录的信息。模板极大的增强了 NetStream 流记录的灵活性和可扩展性，方便了第三方软件的开发，和后续 NetStream 功能的增强。

图 4 Template FlowSet 的格式

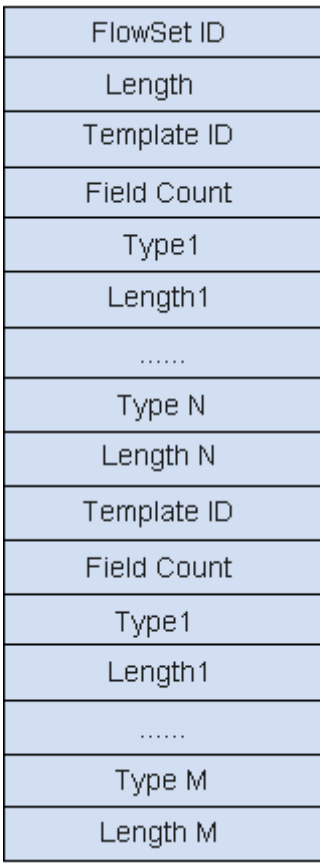


图 5 Data FlowSet 的格式

字段	长度	含义
FlowSet ID	2 字节	用于在 Export Packet 中对 Template FlowSet 进行编号，同时区分出 Template FlowSet 和 Data FlowSet。对于 Template FlowSet，FlowSet ID 的取值是 0～255，对于 Data FlowSet，取值从 256 开始，这样 collector 就可以在 Export Packet 中识别出 Template FlowSet。
Length	2 字节	用于决定下一个 FlowSet 的其实位置，取值是上面图中全部的字节数（包括 FlowSet ID 和 Length 自身）。

字段	长度	含义
Template ID	2 字节	为模板定义一个编号，取值从 256 开始，原因是 0～255 被 FlowSet ID 所用。注意一个 Data FlowSet 对应一个 Template ID Field count 表示模板中 type 的总个数。
Field type	2 字节	表示类型的名称，这个字段的解释 router 和 collector 必须约定好。比如如果支持按照目的 IP 地址、协议类型、TOS 和 MPLS 标签进行统计，则这四种信息都有一个 type 的定义。
Field length	2 字节	对应的 type 的长度，对于目的 IP 地址，取值是 4，表示 4 个字节。

参考标准

标准	描述
RFC 3954	Cisco Systems NetFlow Services Export Version

6.15 RIP 报文格式

RIP 是 Routing Information Protocol(路由信息协议)的简称。它是一种较为简单的内部网关协议 IGP(Interior Gateway Protocol)，主要用于规模较小的网络中，比如校园网以及结构较简单的地区性网络。对于更为复杂的环境和大型网络，一般不使用 RIP。

RIP 是一种基于距离矢量（Distance-Vector）算法的协议，它通过 UDP 报文进行路由信息的交换，使用的端口号为 520。

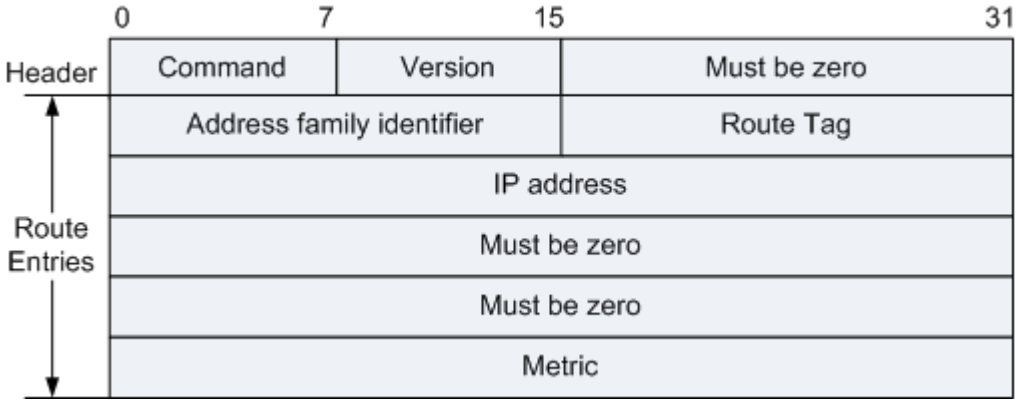
RIP 有两个版本：RIP-1 和 RIP-2。

- RIP-1 是有类别路由协议（Classful Routing Protocol），它只支持以广播方式发布协议报文。RIP-1 的协议报文中没有携带掩码信息，它只能识别 A、B、C 类这样的自然网段的路由，因此 RIP-1 无法支持路由聚合，也不支持不连续子网（Discontiguous Subnet）。
- RIP-2 是一种无分类路由协议（Classless Routing Protocol），有两种报文传送方式：广播方式和组播方式，缺省将采用组播方式发送报文，使用的组播地址为 224.0.0.9。当接口运行 RIP-2 广播方式时，也可接收 RIP-1 的报文。

RIP-1 的报文格式

RIP-1 报文由头部（Header）和多个路由表项（Route Entries）部分组成。在一个 RIP 报文中，最多可以有 25 个路由表项。RIP 是一个基于 UDP 协议的，并且 RIP-1 的数据包不能超过 512 字节。

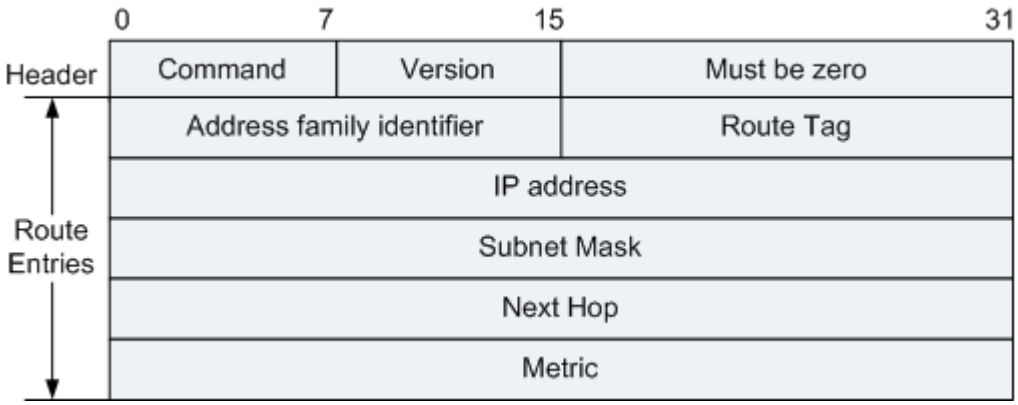
图 1 RIP-1 的报文格式



字段名	长度	含义
Command	8 比特	标识报文的类型： • 1：Request 报文，向邻居请求全部或部分路由信息； • 2：Reponse 报文，发送自己全部或部分路由信息，一个 Response 报文中最多包含 25 个路由表项。
Version	8 比特	RIP 的版本号： • 1：RIP-1 • 2：RIP-2
Must be zero	16/32 比特	必须为零字段。
AFI（Address family identifier）	16 比特	地址族标识，其值为 2 时表示 IP 协议。对于 Request 报文，此字段值为 0。
IP Address	32 比特	该路由的目的 IP 地址，可以是自然网段的地址，也可以是子网地址或主机地址。
Metric	32 比特	路由的开销值。对于 Request 报文，此字段值为 16。

RIP-2 的报文格式

图 2 RIP-2 的报文格式

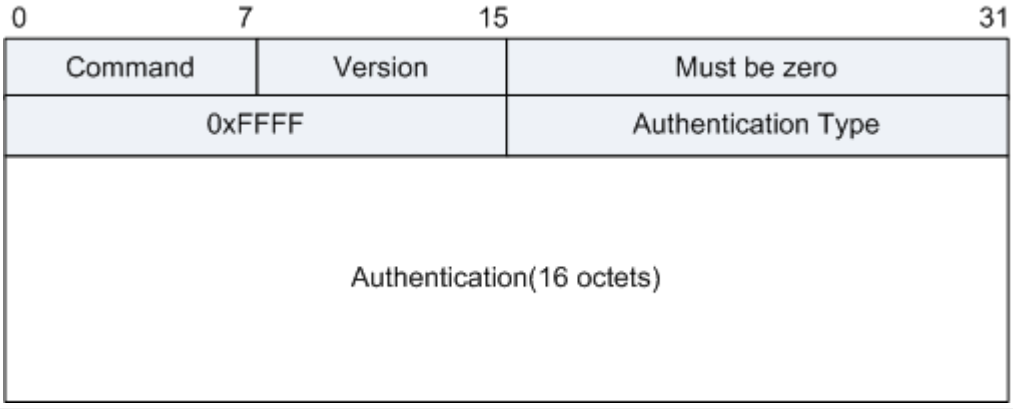


字段名	长度	含义

字段名	长度	含义
Command	8 比特	标识报文的类型： • 1：Request 报文，向邻居请求全部或部分路由信息； • 2：Reponse 报文，发送自己全部或部分路由信息，一个 Response 报文中最多包含 25 个路由表项。
Version	8 比特	RIP 的版本号： • 1：RIP-1 • 2：RIP-2
Must be zero	16 比特	必须为零字段。
AFI （Address Family Identifier）	16 比特	地址族标识，其值为 2 时表示 IP 协议。对于 Request 报文，此字段值为 0。
Route Tag	16 比特	外部路由标记。
IP Address	32 比特	该路由的目的 IP 地址，可以是自然网段的地址，也可以是子网地址或主机地址。
Subnet Mask	32 比特	目的地址的掩码。
Next Hop	32 比特	提供一个更好的下一跳地址。如果为 0.0.0.0，则表示发布此路由的路由器地址就是最优下一跳地址。
Metric	32 比特	路由的开销值。对于 Request 报文，此字段为 16。

RIP-2 为了支持报文验证，使用第一个路由表项（Route Entry）作为验证项，并将 AFI 字段的值设为 0xFFFF 作为标识。

图 3 RIP-2 的验证报文格式



字段名	长度	含义
Command	8 比	标识报文的类型：

字段名	长度	含义
	特	• 1：Request 报文，向邻居请求全部或部分路由信息； • 2：Reponse 报文，发送自己全部或部分路由信息，一个 Response 报文中最多包含 25 个路由表项。
Version	8 比特	RIP 的版本号： • 1：RIP-1 • 2：RIP-2
Must be zero	16 比特	必须为零字段。
0xFFFF	16 比特	验证项标识，表示整个路由报文需要验证。
Authentication Type	16 比特	验证类型： • 2：明文验证； • 3：MD5 验证。
Authentication	16 字节	验证口令，当使用明文验证时该字段才会包含密码信息。

参考标准

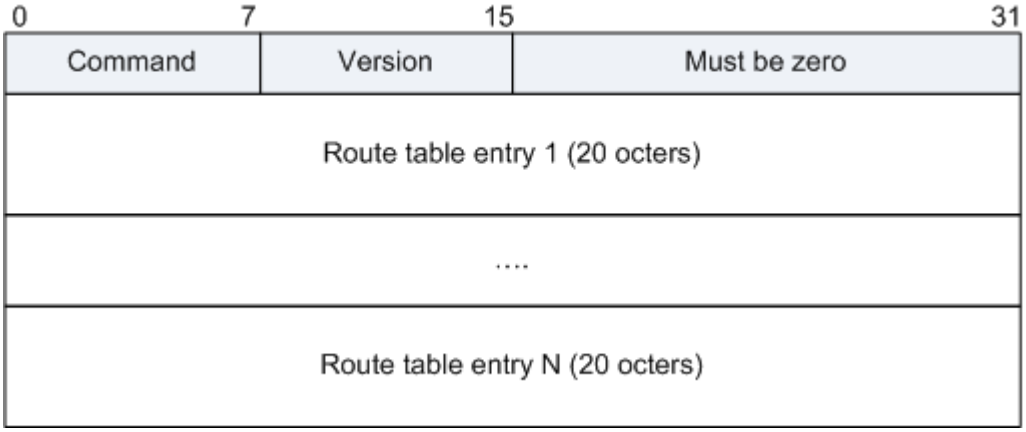
文档编号	描述
RFC 1058	Routing Information Protocol
RFC 1723	RIP Version 2 Carrying Additional Information
RFC 1721	RIP Version 2 Protocol Analysis
RFC 1722	RIP Version 2 Protocol Applicability Statement
RFC 1724	RIP Version 2 MIB Extension
RFC 2082	RIP-2 MD5 Authentication

文档编号	描述
RFC 2453	RIP Version 2

6.16 RIPng 的报文格式

报文格式

RIPng 报文由头部（Header）和多个路由表项 RTEs（Route Table Entry）组成。在同一个 RIPng 报文中，RTE 的最大数目根据接口的 MTU 值来确定。



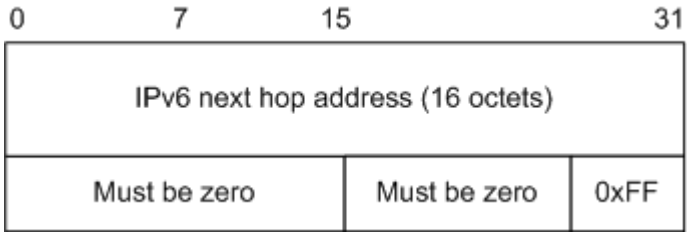
RIPng 报文各字段解释如下表：

字段名	长度	含义
Command	8 比特	标识报文的类型： • 1：Request 报文，向邻居请求全部或部分路由信息 • 2：Reponse 报文，发送自己全部或部分路由信息。路由条目与链路 MTU 有关。
Version	8 比特	RIPng 的版本号：其值为 1。
Must be zero	16 比特	必须为零字段。
RTE（Route table entry）	20 字节	路由表项。

RIPng 有两类 RTE，分别是：

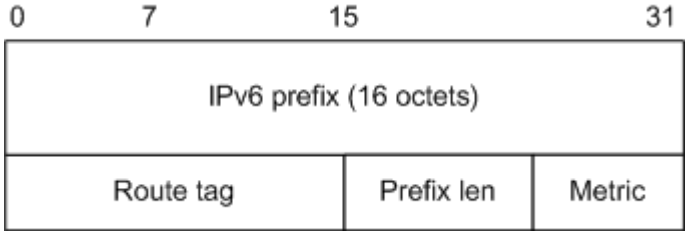
- 下一跳 RTE：位于一组具有相同下一跳的“IPv6 前缀 RTE”的最前面，它定义了下一跳的 IPv6 地址。

图 1 下一跳 RTE 格式



- IPv6 前缀 RTE：位于某个“下一跳 RTE”的后面，同一个“下一跳 RTE”的后面可以有多个不同的“IPv6 前缀 RTE”。它描述了 RIPng 路由表中的目的 IPv6 地址及开销。

图 2 IPv6 前缀 RTE 格式



字段名	长度	含义
IPv6 prefix	16 字节	目的 IPv6 地址的前缀。
Route tag	16 比特	路由标记，用来区分外部路由。
Prefix len	8 比特	IPv6 地址的前缀长度。
Metric	8 比特	路由的度量值（开销）。

6.17 NTP 报文格式

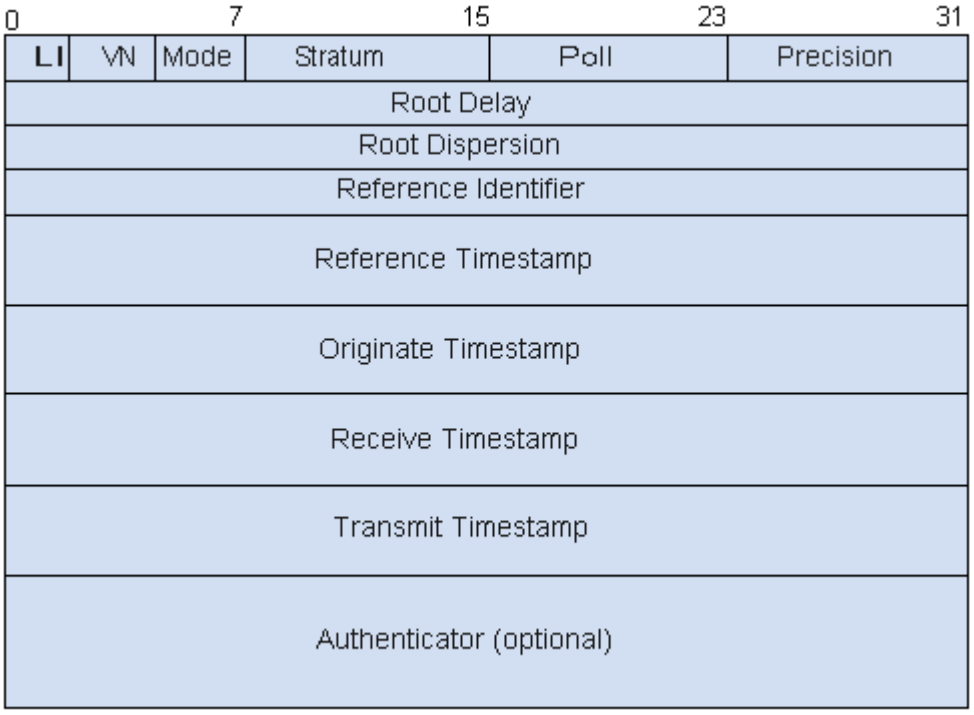
NTP 是从时间协议（Time Protocol）和 ICMP 时间戳报文（ICMP TimeStamp Message）演变而来，在准确性和健壮性方面进行了特殊的设计，理论上精度可达十亿分之一秒。

NTP 协议应用于分布式时间服务器和客户端之间，实现客户端和服务器的时间同步，从而使网络内所有设备的时钟基本保持一致。

NTP 协议是基于 UDP 进行传输的，使用端口号为 123。

报文格式

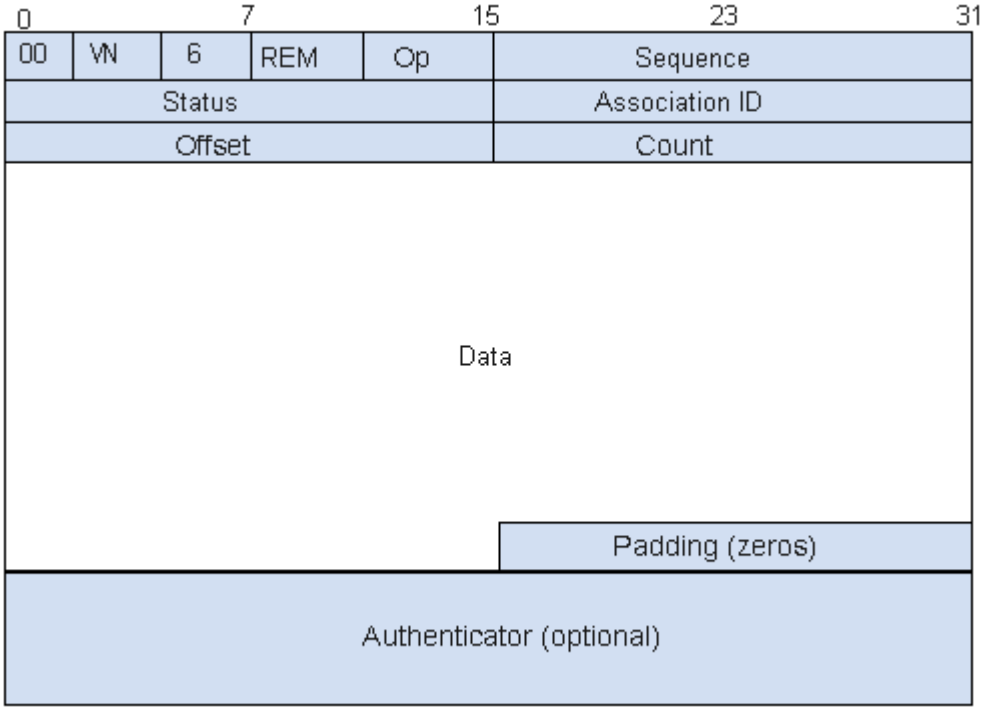
图 1 NTP 数据报文格式



字段名	长度	含义
LI （ Leap Indicator）	2 比特	这是一个两位的代码，表示在 NTP 时间标尺中将要插入的下一跳情况。值为“11”时表示告警状态，时钟不能被同步。
VN （ Version Number）	3 比特	NTP 的版本号。
Mode	3 比特	NTP 的工作模式。不同值表示的含义如下：0：reserved，保留。1：symmetric active，主动对等体模式。2：symmetric passive，被动对等体模式。3：client，客户模式。4：server，服务器模式。5：broadcast，广播模式。6：reserved for NTP control messages，NTP 控制报文。7：reserved for private use，内部使用预留。

字段名	长度	含义
Stratum	8 比特	时钟的层数，定义了时钟的准确度。层数为 1 的时钟准确度最高，从 1 到 15 依次递减。
Poll Interval	8 比特	轮询时间，即发送报文的最小间隔时间。
Precision	8 比特	时钟的精度。
Root Delay	32 比特	到主参考时钟的总往返延迟时间。
Root Dispersion	32 比特	本地时钟相对于主参考时钟的最大误差。
Reference Identifier	32 比特	标识特定参考时钟。
Reference Timestamp	64 比特	本地时钟最后一次被设定或更新的时间。如果值为 0 表示本地时钟从未被同步过。
Originate Timestamp	64 比特	NTP 报文离开源端时的本地时间。
Receive Timestamp	64 比特	NTP 报文到达目的端的本地时间。
Transmit Timestamp	64 比特	目的端应答报文离开服务器端的本地时间。
Authenticator	96 比特	（可选）验证信息。

图 2 NTP 控制报文格式



字段名	长度	含义
0	2 比特	保留位。NTP 本身不做处理。
VN （ Version Number）	3 比特	NTP 的版本号，目前值为 3。
6	3 比特	表明是控制报文。
REM	3 比特	R：0 表示命令，1 表示响应。E：0 表示发送正常响应，1 表示发送错误响应。M：0 表示最后一个分片，1 表示其他。
Op	5 比特	操作码，表明命令的类型。
Sequence	16 比特	发送或接受到报文的顺序号。
Status	16 比特	表明当前系统的状态。
Association ID	16 比特	连接标示。

字段名	长度	含义
Offset	16 比特	偏移量。
Count	16 比特	数据域的长度。
Data	最大 468 比特	包括发送报文或接受报文中的数据信息。
Padding	16 比特	填充字段。
Authenticator	96 比特	（可选）验证信息。

报文示例

图 3 NTP 报文（Broadcast）

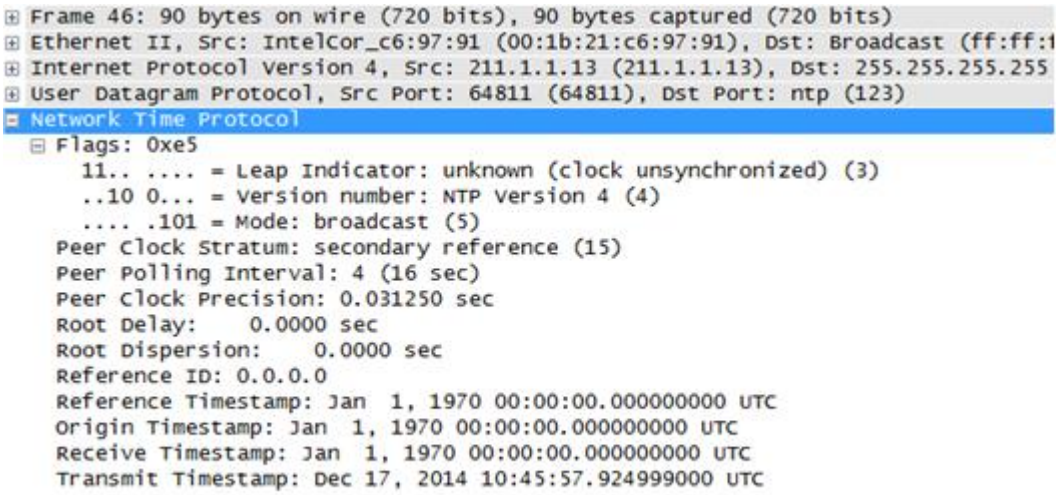


图 4 NTP 报文（client）

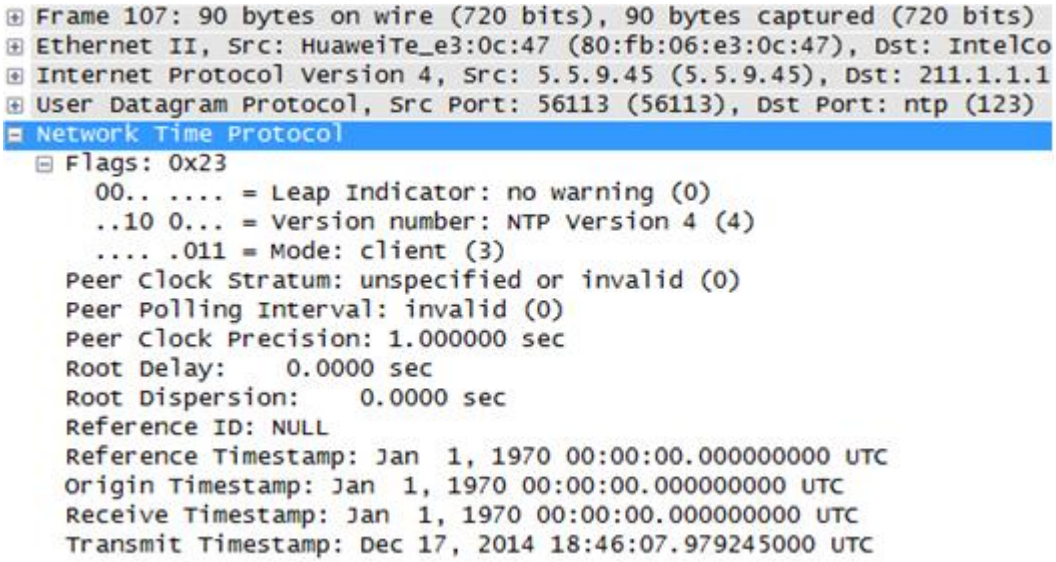


图 5 NTP 报文（server）

```
⊞ Frame 100: 90 bytes on wire (720 bits), 90 bytes captured (720 bits)
⊞ Ethernet II, Src: IntelCor_c1:df:df (00:1b:21:c1:df:df), Dst: IntelC
⊞ Internet Protocol Version 4, Src: 211.1.1.6 (211.1.1.6), Dst: 211.1.
⊞ User Datagram Protocol, Src Port: ntp (123), Dst Port: 53264 (53264)
⊟ Network Time Protocol
  ⊟ Flags: 0x1c
      00.. .... = Leap Indicator: no warning (0)
      ..01 1... = Version number: NTP Version 3 (3)
      .... .100 = Mode: server (4)
    Peer Clock Stratum: primary reference (1)
    Peer Polling Interval: invalid (0)
    Peer Clock Precision: 0.015625 sec
    Root Delay:    0.0000 sec
    Root Dispersion:   10.3870 sec
    Reference ID: uncalibrated local clock
    Reference Timestamp: Dec 17, 2014 02:00:44.156124000 UTC
    Origin Timestamp: Dec 17, 2014 10:55:29.457999000 UTC
    Receive Timestamp: Dec 17, 2014 10:55:29.453125000 UTC
    Transmit Timestamp: Dec 17, 2014 10:55:29.453125000 UTC
```

参考标准

文档编号	描述
RFC 1095	Network Time Protocol (Version 1)
RFC 1119	Network Time Protocol (Version 2)
RFC 1305	Network Time Protocol (Version 3)
RFC 5905	Network Time Protocol Version 4: Protocol and Algorithms Specification
RFC 5906	Network Time Protocol Version 4: Autokey Specification

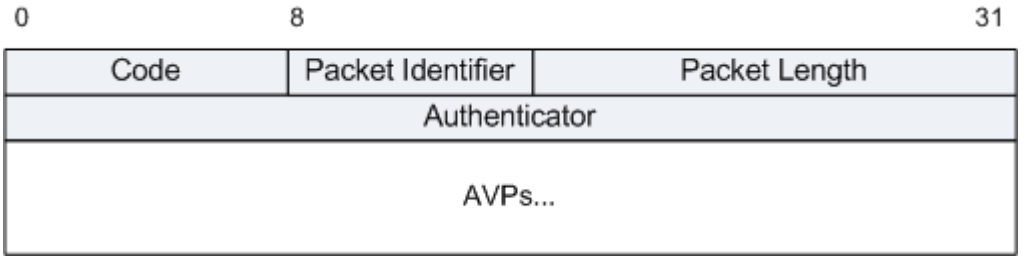
6.18 RADIUS 报文格式

RADIUS（Remote Authentication Dial In User Service）远程用户拨号认证系统，是目前应用最广泛的 AAA 协议。

RADIUS 是一种分布式的、客户端/服务器（Client/Server）结构的信息交互协议，规定了客户端与服务器之间传递用户信息和计费信息的过程和报文格式。其用途是完成用户的认证、授权、计费功能。

报文格式

图 1 RADIUS 报文格式



字段	长度	描述
Code	1 byte	用来标识 RADIUS 报文类型。 • Code = 1: Access-Request，接入请求报文 • Code = 2: Access-Accept，接入成功回应报文 • Code = 3: Access-Reject，接入拒绝回应报文 • Code = 4: Accounting-Request，计费请求报文 • Code = 5: Accounting-Response，计费回应报文 • Code = 11: Access-Challenge，接入挑战报文 • Code = 12: Status-Server (experimental)，服务器状态报文（试验） • Code = 13: Status-Client (experimental)，客户端状态报文（试验） • Code = 255: Reserved，保留
Packet Identifier	1 byte	用于匹配请求和回应报文。果在一个很短的时间内接收到相同的源 IP 地址、源 UDP 端口号和相同的 Identifier 域的请求报文，RADIUS 服务器就可以认为是重复的请求报文。
Packet Length	2 bytes	包含了报文中的 Code 域，Identifier 域，Length 域，Authenticator 域和属性域的总长度。如果收到的报文实际长度超过 Length，超过部分被当做填充内容忽略掉，如果实际长度小于 Length，报文被丢弃。
Authenticator	16 bytes	用于认证来自服务端的响应，也用于用户密码的加密处理中。
AVPs	变长	属性字段，承载认证、授权、计费以及配置等信息。采用 TLV 格式： • 类型（Type）： 占位一个字节。具体类型值对应的属性名请参见下表。 • 长度（Length）域：占位一个字节，表示包括 Type、Length、Value 域在内的属性的长度。 • 值（Value）域：占位零个或者更多字节，它包含了属性信息的详细描述。值域的格式和长度是由属性的类型和长度决定的。需要指出的是，在 RADIUS 中没有任何类型的属性值是以 NULL（十六进制的 0x00）结束的。

表 1 Radius 属性

字段	长度	描述

属性编号	属性名
1	User-Name
2	User-Password
3	CHAP-Password
4	NAS-IP-Address
5	NAS-Port
6	Service-Type
7	Framed-Protocol
8	Framed-IP-Address
9	Framed-IP-Netmask
10	Framed-Routing
11	Filter-Id
12	Framed-MTU
13	Framed-Compression
14	Login-IP-Host
15	Login-Service
16	Login-TCP-Port
18	Reply-Message

字段	长度	描述
19		Callback-Number
20		Callback-Id
22		Framed-Route
23		Framed-IPX-Network
24		State
25		Class
26		Vendor-Specific
27		Session-Timeout
28		Idle-Timeout
29		Termination-Action
30		Called-Station-Id
31		Calling-Station-Id
32		NAS-Identifier
33		Proxy-State
34		Login-LAT-Service
35		Login-LAT-Node
36		Login-LAT-Group
37		Framed-AppleTalk-Link

字段	长度	描述
38		Framed-AppleTalk-Network
39		Framed-AppleTalk-Zone
40		Acct-Status-Type
41		Acct-Delay-Time
42		Acct-Input-Octets
43		Acct-Output-Octets
44		Acct-Session-Id
45		Acct-Authentic
46		Acct-Session-Time
47		Acct-Input-Packets
48		Acct-Output-Packets
49		Acct-Terminate-Cause
50		Acct-Multi-Session-Id
51		Acct-Link-Count
52		Acct-Input-Gigawords
53		Acct-Output-Gigawords
55		Event-Timestamp
60		CHAP-Challenge

字段	长度	描述
61		NAS-Port-Type
62		Port-Limit
63		Login-LAT-Port
64		Tunnel-Type
65		Tunnel-Medium-Type
66		Tunnel-Client-Endpoint
67		Tunnel-Server-Endpoint
68		Acct-Tunnel-Connection
69		Tunnel-Password
70		ARAP-Password
71		ARAP-Features
72		ARAP-Zone-Access
73		ARAP-Security
74		ARAP-Security-Data
75		Password-Retry
76		Prompt
77		Connect-Info
78		Configuration-Token

字段	长度	描述
79		EAP-Message
80		Message-Authenticator
81		Tunnel-Private-Group-ID
82		Tunnel-Assignment-ID
83		Tunnel-Preference
84		ARAP-Challenge-Response
85		Acct-Interim-Interval
86		Acct-Tunnel-Packets-Lost
87		NAS-Port-Id
88		Framed-Pool
89		Chargeable-User-Identity
90		Tunnel-Client-Auth-ID
91		Tunnel-Server-Auth-ID
94		Originating-Line-Info
95		NAS-IPv6-Address
96		Framed-Interface-Id
97		Framed-IPv6-Prefix
98		Login-IPv6-Host

字段	长度	描述
99	Framed-IPv6-Route	
100	Framed-IPv6-Pool	
101	Error-Cause	
192-223	保留给实验用	
224-240	保留给特定实现用	
241-255	预留的，而且不应该使用它们	

报文示例

```
⊞ Frame 1: 98 bytes on wire (784 bits), 96 bytes captured (768 bits)
⊞ Ethernet II, Src: IETF-VRRP-VRID_81 (00:00:5e:00:01:81), Dst: HuaweiTe_54:87:26 (0(
⊞ Internet Protocol, Src: 10.137.162.163 (10.137.162.163), Dst: 10.137.109.108 (10.1:
⊞ User Datagram Protocol, Src Port: tdp-suite (1814), Dst Port: icl-twobase1 (25000)
⊟ Radius Protocol
    Code: Access-Request (1)
    Packet identifier: 0x65 (101)
    Length: 56
    Authenticator: e1671797c52e15f763380b45e841ec32
  ⊟ Attribute Value Pairs
    ⊟ AVP: l=6  t=NAS-Port(5): 1814
        NAS-Port: 1814
    ⊟ AVP: l=6  t=NAS-IP-Address(4): 10.135.14.126
        NAS-IP-Address: 10.135.14.126 (10.135.14.126)
    ⊟ AVP: l=6  t=User-Name(1): test
        User-Name: test
      AVP: l=18  t=User-Password(2): Encrypted
  [Packet size limited during capture: RADIUS truncated]
```

参考标准

标准	描述
RFC 2865	Remote Authentication Dial In User Service (RADIUS)
RFC 2866	RADIUS Accounting
RFC 2867	RADIUS Accounting Modifications for Tunnel Protocol Support
RFC 2868	RADIUS Attributes for Tunnel Protocol Support
RFC 2869	RADIUS Extensions

标准	描述
RFC 3162	RADIUS and IPv6
RFC 3576	Dynamic Authorization Extensions to Remote Authentication Dial In User Service (RADIUS)
RFC 4372	Chargeable User Identity

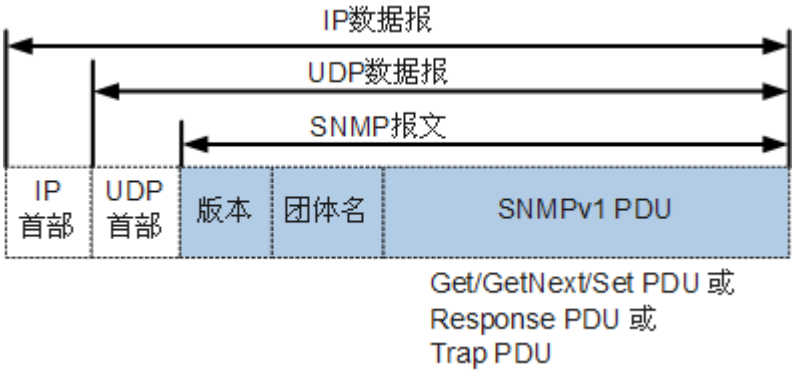
6.19 SNMP 报文格式

- SNMPv1 Packet FormatSNMPv1 报文格式
- SNMPv2c Packet FormatSNMPv2c 报文格式
- SNMPv3 报文格式

父主题： 应用层

6.19.1 SNMPv1 Packet FormatSNMPv1 报文格式

报文格式



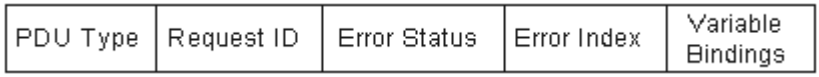
SNMPv1 报文主要由版本、团体名、SNMP PDU 三部分构成。

字段	描述
版本	表示 SNMP 的版本，版本字段的值是报文版本号减 1，如果是 SNMPv1 报文则对应字段值为 0。
团体名	用于在 Agent 与 NMS 之间完成认证，字符串形式，常用的是 6 个字符“public”。团体名包括“可读”和“可写”两种，执行 Get、GetNext 操作时，采用“可读团体名”进行认证；执行 Set 操作时，则采用“可写团体名”认

字段	描述
	证。
SNMPv1 PDU	包含 PDU 类型、请求标识符、变量绑定列表等信息，可以为 GetRequest PDU、GetNextRequest PDU、SetRequest PDU、Response PDU 或 Trap PDU 几种类型。

其中，PDU 的格式如下：

图 1 SNMPv1 PUD 格式



字段	描述
PDU Type	协议数据单元的类型。PDU（Protocol Data Unit）共有 5 种类型： • GetRequest-PDU • GetNextRequest-PDU • GetResponse-PDU • SetRequest-PDU • Trap-PDU
Request ID	请求标示字段，唯一的标示一个请求报文。
Error Status	错误状态标示字段，SNMPv1 中错误码包括： • noSuchName：指定了一个代理不知道的对象。 • tooBig：代理不能一次把请求的结果放入到一个 PDU 中。 • badValue：进行 set 操作时候把变量修改为一个无效的值。 • genErr：除以上错误外的其他错误。
Error Index	错误索引字段。
Variable Bindings	变量绑定字段。

报文示例

图 2 SNMPv1 GetNextRequest 报文

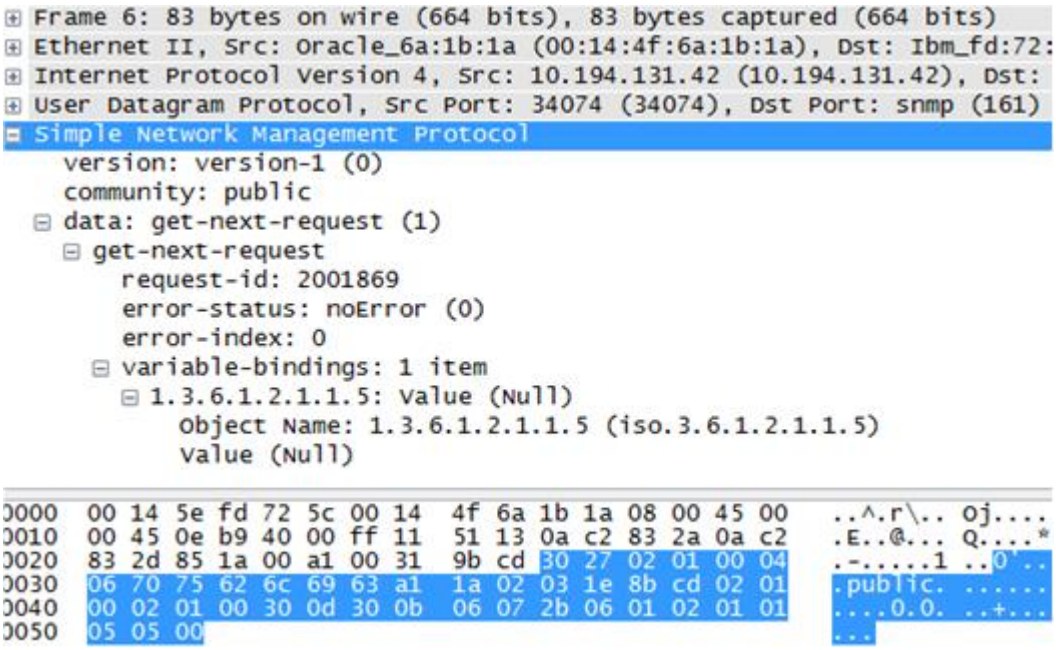


图 3 SNMPv1 GetResponse 报文

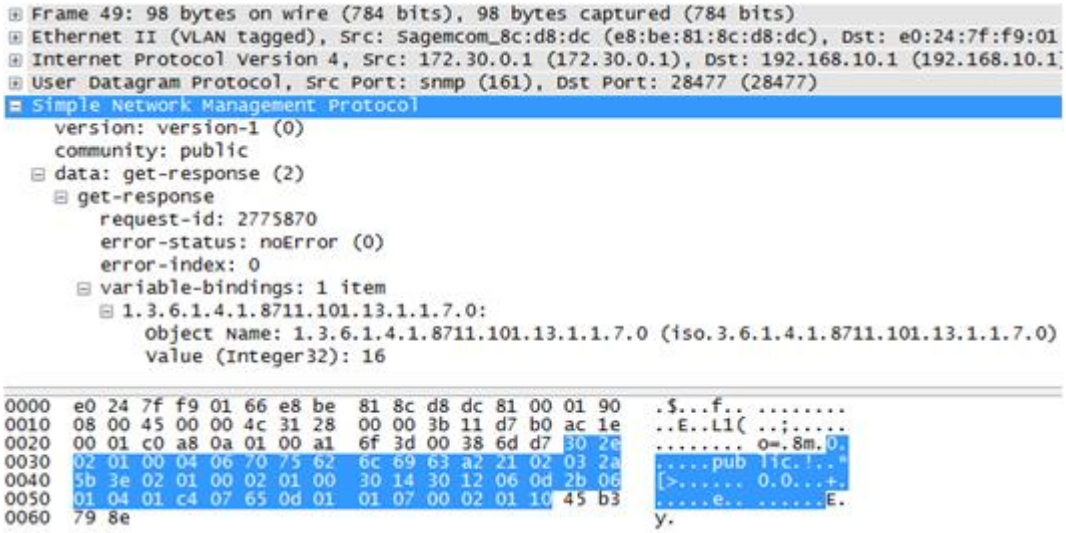
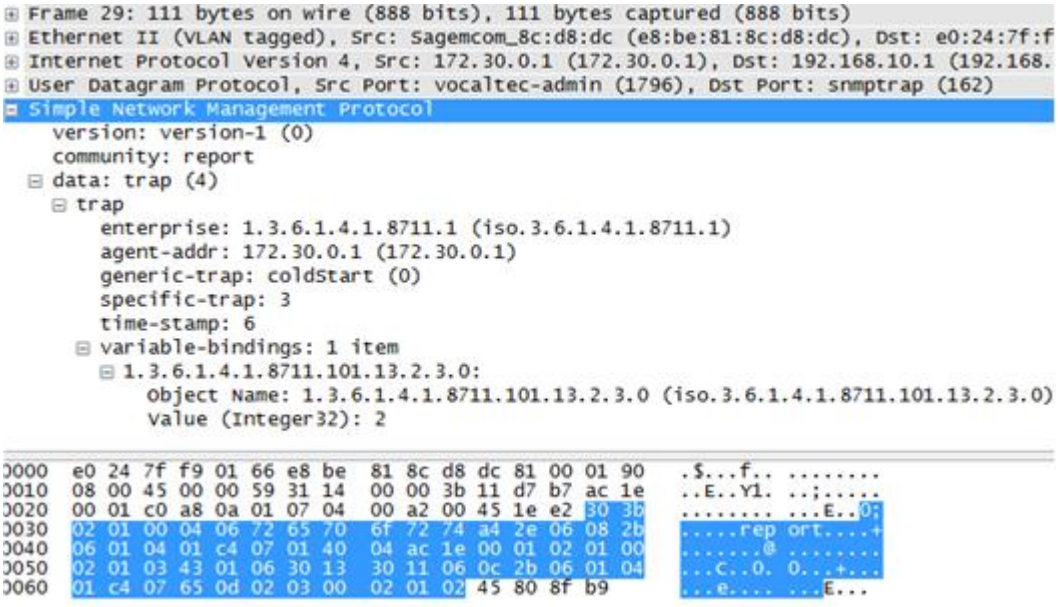


图 4 SNMPv1 Trap 报文



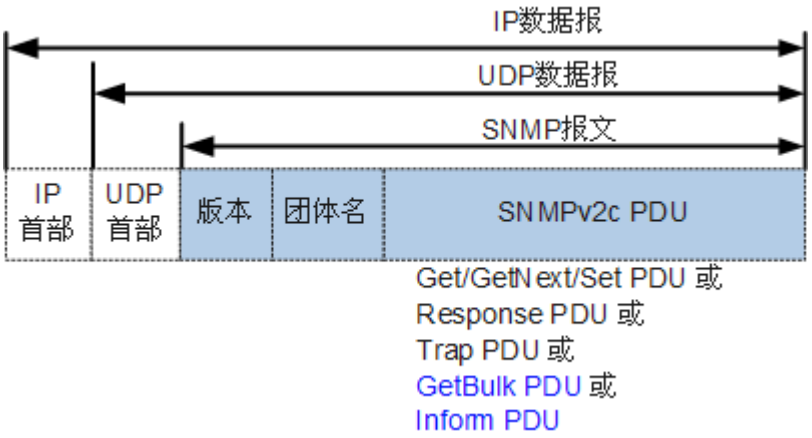
参考标准

标准	描述

标准	描述
RFC 1157	A Simple Network Management Protocol (SNMP)

6.19.2 SNMPv2c Packet FormatSNMPv2c 报文格式

报文格式

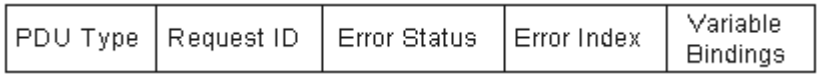


与 SNMPv1 PDU 类似，SNMPv2c PDU 也包括 GetRequest PDU、GetNextRequest PDU、SetRequest PDU、Response PDU、Trap PDU，并新增了 GetBulk PDU 和 Inform PDU 两种类型。

字段	描述
版本	表示 SNMP 的版本，版本字段的值是报文版本号减 1，如果是 SNMPv2 报文则对应字段值为 1。
团体名	用于在 Agent 与 NMS 之间完成认证，字符串形式，常用的是 6 个字符“public”。团体名包括“可读”和“可写”两种，执行 Get、GetNext 操作时，采用“可读团体名”进行认证；执行 Set 操作时，则采用“可写团体名”认证。
SNMPv2c PDU	包含 PDU 类型、请求标识符、变量绑定列表等信息，可以为 GetRequest PDU、GetNextRequest PDU、SetRequest PDU、Response PDU、Trap PDU、GetBulk PDU 和 Inform PDU 几种类型。

其中，PDU 的格式如下：

图 1 SNMPv2c PUD 格式



字段	描述
PDU Type	协议数据单元的类型。PDU（Protocol Data Unit）共有 5 种类型： • GetRequest-PDU • GetNextRequest-PDU • GetBulk • GetResponse-PDU • SetRequest-PDU • Trap-PDU
Request ID	请求标示字段，唯一的标示一个请求报文。
Error Status	错错误状态标示字段。SNMPv2c 中错误码包括： • wrongValue：进行 set 操作时候把变量修改为一个无效的值 • wrongEncoding：进行编码字段的值，与其他的字段不一致 • wrongType：进行 set 操作时候把变量修改为一个无效的类型 • wrongLength：进行 set 操作时候把一个变量值设置成与它长度不一致的值 • inconsistentValue：把一个变量设置为其他的情况下有效的值，当前情况下无效 • noAccess：试图设置一个不可访问的值 • notWritable：试图修改一个存在，但不能修改的值 • noCreation：试图修改一个存在，但不能创建的值 • inconsistentName：试图设置一个当前不存在且当前不能创建的变量 • resourceUnavailable：设置过程中申请某些资源失败 • commitFailed：set 操作失败 • undoFailed：进行 set 操作失败，有些赋值无法回复 • genErr：除以上错误外的其他错误
Error Index	错误索引字段。
Variable Bindings	变量绑定字段。

报文示例

图 2 SNMPv2c GetRequest 报文

```
⊞ Frame 1485: 84 bytes on wire (672 bits), 84 bytes captured (672 bits)
⊞ Ethernet II, Src: Oracle_6a:19:f2 (00:14:4f:6a:19:f2), Dst: Ibm_fd:72
⊞ Internet Protocol Version 4, Src: 10.194.131.41 (10.194.131.41), Dst:
⊞ User Datagram Protocol, Src Port: 60830 (60830), Dst Port: snmp (161)
⊟ Simple Network Management Protocol
    version: v2c (1)
    community: public
  ⊟ data: get-request (0)
    ⊟ get-request
        request-id: 206355
        error-status: noError (0)
        error-index: 0
      ⊟ variable-bindings: 1 item
        ⊟ 1.3.6.1.2.1.1.5.0: Value (Null)
            Object Name: 1.3.6.1.2.1.1.5.0 (iso.3.6.1.2.1.1.5.0)
            Value (Null)

0000  00 14 5e fd 72 5c 00 14  4f 6a 19 f2 08 00 45 00   ..^.r\.. Oj....
0010  00 46 a9 5b 40 00 ff 11  b6 70 0a c2 83 29 0a c2   .F.[@... .p...)
0020  83 2d ed 9e 00 a1 00 32  92 1f 30 28 02 01 01 04   .-.....2 ..0(....
0030  06 70 75 62 6c 69 63 a0  1b 02 03 03 26 13 02 01   .public. ....&...
0040  00 02 01 00 30 0e 30 0c  06 08 2b 06 01 02 01 01   ....0.0. ..+.....
0050  05 00 05 00                                       ....
```

图 3 SNMPv2c GetNextRequest 报文

```
⊞ Frame 1540: 83 bytes on wire (664 bits), 83 bytes captured (664 bits)
⊞ Ethernet II, Src: Oracle_6a:1b:1a (00:14:4f:6a:1b:1a), Dst: Ibm_fd:72
⊞ Internet Protocol Version 4, Src: 10.194.131.42 (10.194.131.42), Dst:
⊞ User Datagram Protocol, Src Port: 34074 (34074), Dst Port: snmp (161)
⊟ Simple Network Management Protocol
    version: v2c (1)
    community: public
  ⊟ data: get-next-request (1)
    ⊟ get-next-request
        request-id: 1999563
        error-status: noError (0)
        error-index: 0
      ⊟ variable-bindings: 1 item
        ⊟ 1.3.6.1.2.1.1.5: Value (Null)
            Object Name: 1.3.6.1.2.1.1.5 (iso.3.6.1.2.1.1.5)
            Value (Null)

0000  00 14 5e fd 72 5c 00 14  4f 6a 1b 1a 08 00 45 00   ..^.r\.. Oj....
0010  00 45 8b 43 40 00 ff 11  d4 88 0a c2 83 2a 0a c2   .E.C@... .....*..
0020  83 2d 85 1a 00 a1 00 31  a3 cf 30 27 02 01 01 04   .-.....1 ..0'....
0030  06 70 75 62 6c 69 63 a1  1a 02 03 1e 82 cb 02 01   .public. ........
0040  00 02 01 00 30 0d 30 0b  06 07 2b 06 01 02 01 01   ....0.0. ..+.....
0050  05 05 00                                          ...
```

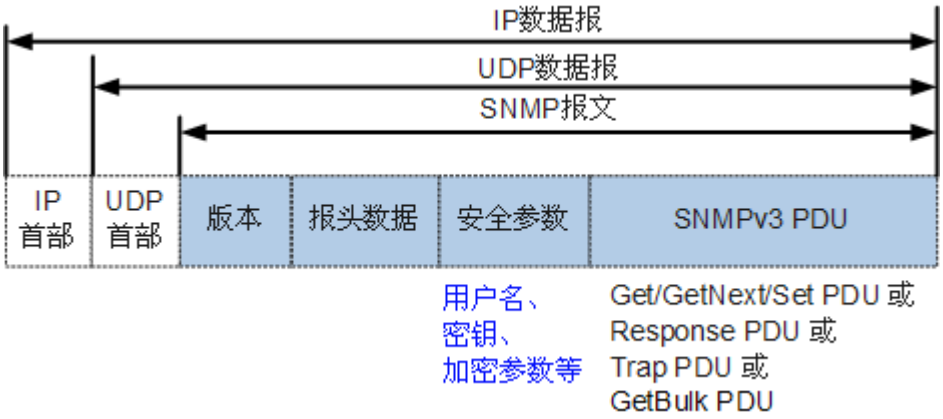
参考标准

标准	描述
RFC 1901	Introduction to Community-based SNMPv2
RFC 1902	Structure of Management Information for Version 2 of the Simple Network Management Protocol (SNMPv2)
RFC 1903	Textual Conventions for Version 2 of the Simple Network Management Protocol (SNMPv2)

标准	描述
RFC 1904	Conformance Statements for Version 2 of the Simple Network Management Protocol (SNMPv2)
RFC 1905	Protocol Operations for Version 2 of the Simple Network Management Protocol (SNMPv2)
RFC 1906	Transport Mappings for Version 2 of the Simple Network Management Protocol (SNMPv2)
RFC 1907	Management Information Base for Version 2 of the Simple Network Management Protocol (SNMPv2)
RFC 1909	An Administrative Infrastructure for SNMPv2

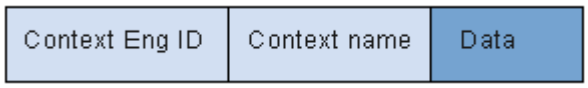
6.19.3 SNMPv3 报文格式

报文格式

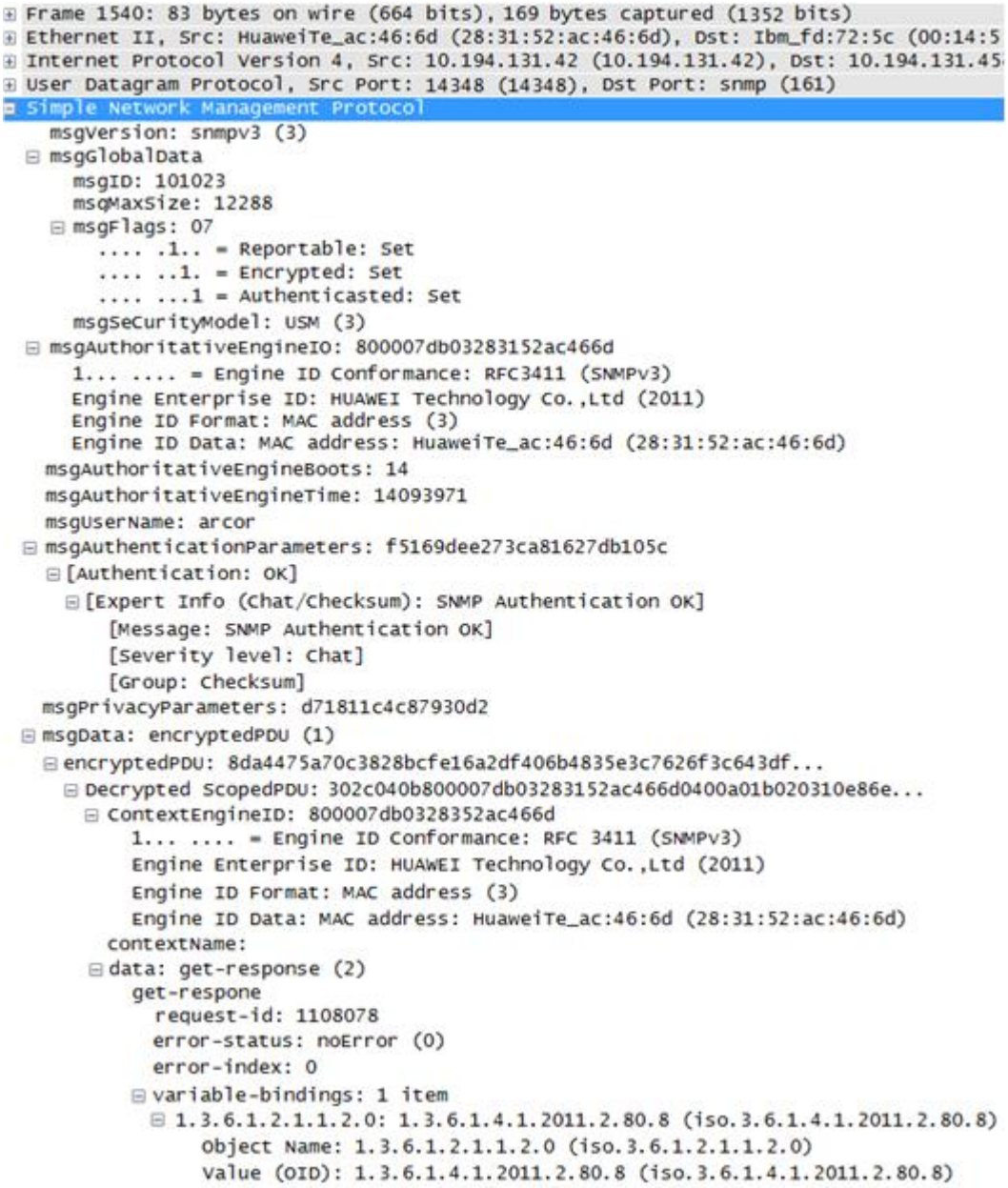


SNMPv3 报文结构从功能上来说，与 SNMPv1、SNMPv2c 的区别主要增加了报头数据和安全参数。

字段	描述
版本	表示 SNMP 的版本，版本字段的值是报文版本号减 1，如果是 SNMPv3 报文则对应字段值为 3。
报头数据	主要包含消息发送者所能支持的最大消息尺寸、消息是否进行加密/认证、采用的安全模式等描述内容。 格式如下： 

字段	描述
	• Msg ID：可以使请求和应答相互关联，响应报文中的 Msg ID 和发送报文中的值相同。 • Msg Max size：消息发送者支持的最大的消息尺寸。 • Msg Sec Model：指明了发送方采用的安全模式。 • Msg Flag：请求报文指定是否要求回应 report 消息，消息是否进行了加密和认证。
安全参数	包含用户名、密钥、加密参数等安全信息。 格式如下：  • Auth Engin ID：唯一的标识一个认证。 • Auth Engin Boots：从配置认证引擎到现在，认证引擎重新启动的次数。 • Auth Engin Time：从配置认证引擎到现在的时间。 • User Name：用户名。 • Auth Para：认证参数值。 • Priv Para：加密后的参数值。
SNMPv3 PDU	包含 PDU 类型、请求标识符、变量绑定列表等信息，可以为 GetRequest PDU、GetNextRequest PDU、SetRequest PDU、Response PDU、Trap PDU、GetBulk PDU 等几种类型。 格式如下：  • Context Egine ID：SNMP 唯一标识符，和 PDU 类型一起决定应该发往那个应用程序。 • Context Name：指明上下文之间的关系，由应用程序决定。 • Data：报文的数据内容。

报文示例



参考标准

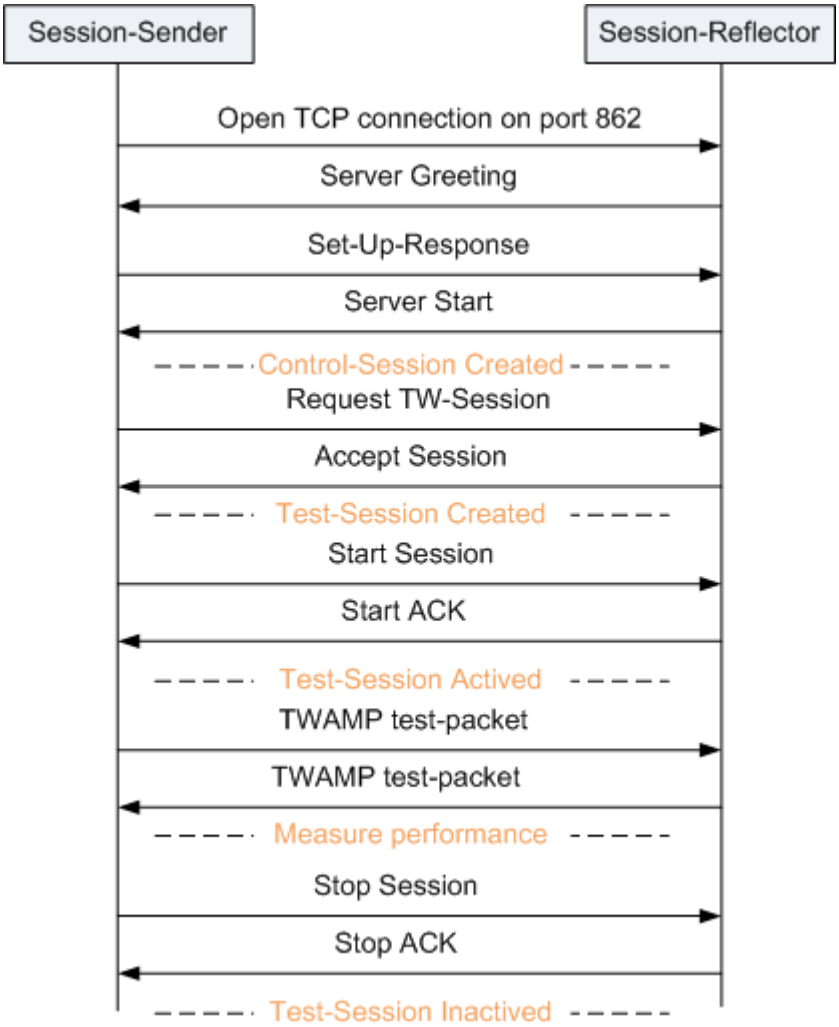
标准	描述
RFC 2570	Introduction to Version 3 of the Internet-standard Network Management Framework
RFC 2574	User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)
RFC 3414	User-based Security Model (USM) for version 3 of the Simple Network Management Protocol (SNMPv3)
RFC 3584	Coexistence between Version 1, Version 2, and Version 3 of the Internet-standard Network Management Framework

6.20 TWAMP 报文格式

TWAMP：Two-way Active Measurement Protocol(TWAMP)，双向测量协议，用于对丢包、时延和抖动等进行性能监控。

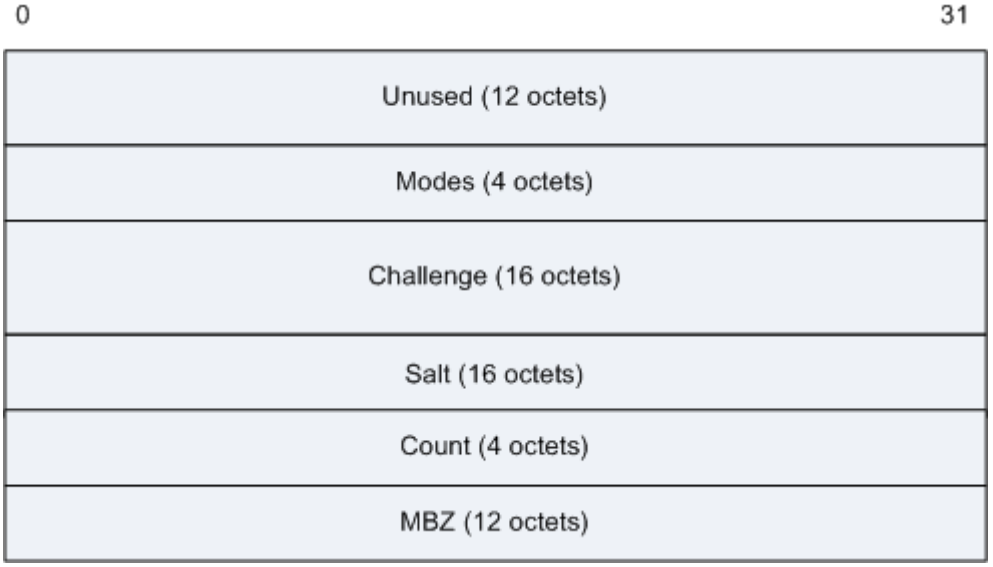
TWAMP 是基于 TCP 连接进行协商和利用 UDP 报文进行测量。TWAMP 的端口号可配置。

TWAMP 工作时序图



TWAMP 控制报文格式

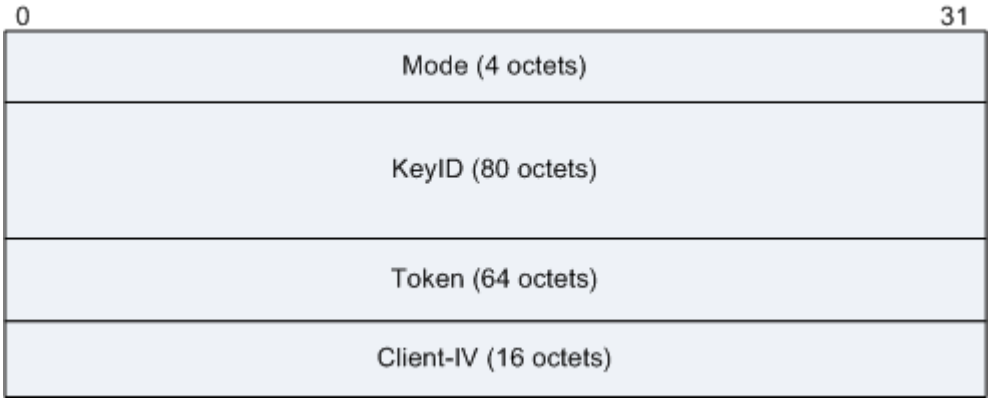
图 1 Server Greeting 消息格式



字段	长度	描述

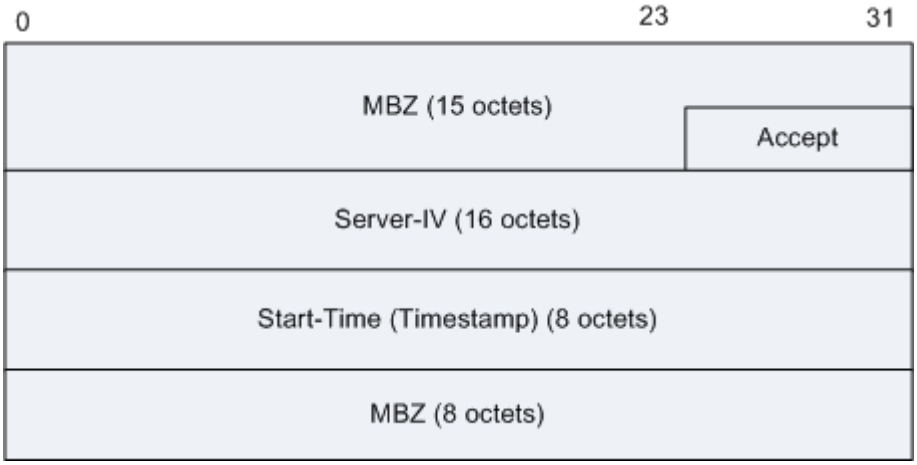
字段	长度	描述
Modes	4 bytes	• 1：不认证 • 2 ：认证 • 4：加密 如果值为 0，表示服务器不希望和客户端交互，可能会立刻关闭连接。
Challenge	16 bytes	服务器生成的随机数，用于接收者对共享秘钥的处理中。
Salt	16 bytes	用于从共享秘钥中提出秘钥的一个参数。
Count	4 bytes	用于从共享秘钥中提出秘钥的一个参数，2 的次幂
MBZ	12 bytes	置 0，接收时忽略。

图 2 Set-Up-Response 消息格式



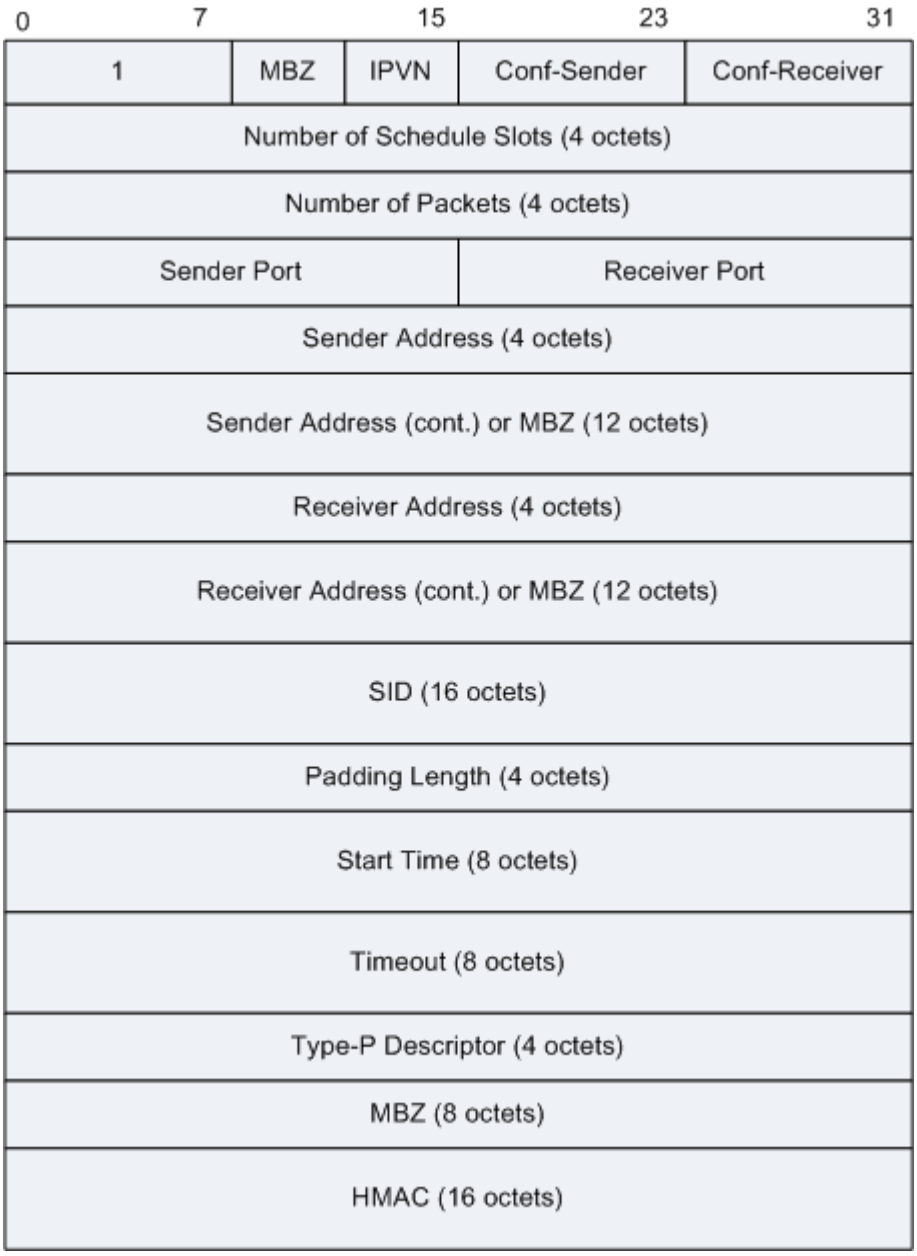
字段	长度	描述
Modes	4 bytes	认证模式
KeyID	80 bytes	用于认证或加密模式中。不认证模式中，不使用此字段。
Token	64 bytes	用于认证或加密模式中。不认证模式中，不使用此字段。
Client-IV	16 bytes	用于认证或加密模式中。不认证模式中，不使用此字段。

图 3 Server Start 消息格式



字段	长度	描述
MBZ	15 bytes	置 0，接收时忽略。
Accept	1 byte	表示服务器意愿继续交互，0 表示服务器接受认证，愿意后续交互。非 0 表示服务器不接受认证。
Server-IV	16 bytes	Server-IV 是服务器随机产生的，Server-IV 不用于不认证模式。
Start-Time	8 bytes	时间戳，代表服务器当前操作开始的时间。
MBZ	8 bytes	置 0，接收时忽略。

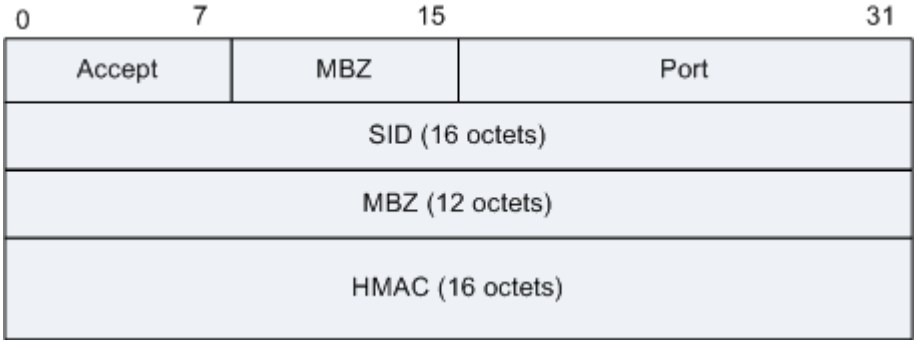
图 4 Request TW-Session 消息格式



字段	长度	描述
MBZ	4 bits	置 0，接收时忽略。
IPVN	4 bits	发送和接收者的 IP 版本号，当前有效值为 4 或 6。
Conf-Sender	1 byte	客户端设置为 0 或 1。服务器将所有非 0 值解析为 1。如果值为 1，表示要求服务器配置对应的代理。
Conf-Receiver	1 byte	客户端设置为 0 或 1。服务器将所有非 0 值解析为 1。如果值为 1，表示要求服务器配置对应的代理。
Number of Schedule	4 bytes	表示在两块 HMAC 之间的槽位记录数量，发送者用来确认发送测试报文的时间。

字段	长度	描述
Slots		
Number of Packets	4 bytes	在 TWAMP 测试会话中发送的主动测量报文的数量。
Sender Port	2 bytes	如果 Conf-Receive 不置位，发送端口是指发送 TWAMP 测试报文的 UDP 端口。
Receiver Port	2 bytes	如果 Conf-Receive 不置位，接收端口是指接收 TWAMP 测试报文的 UDP 端口。
Sender Address	4 bytes	TWAMP 测试会话的发送者 IP 地址。
Receiver Address	4 bytes	TWAMP 测试会话的接收者 IP 地址。
SID	16 bytes	会话 ID，只有 Conf-Receiver 为 0 才有意义。
Padding Length	4 bytes	普通 TWAMP 测试报文的 Padding 字节数。
Start Time	8 bytes	会话发起的时间。格式与 TWAMP-Test 的时间戳相同。
Timeout	8 bytes	超时或时延阈值，时间戳格式。
Type-P Descriptor	4 bytes	前两比特如果为 00，后面 6 个比特指定了发送的 TWAMP 测试报文的 DSCP 值（RFC2474 定义的 DSCP）。 前两比特如果为 01，后面 16 个比特标识要求的 PHB Identification Code (PHB ID)(RFC2836 定义的)。
MBZ	8 bytes	置 0，接收时忽略。
HMAC	16 bytes	TWAMP 使用的 HMAC 是 HMAC-SHA1，128 比特，所以 HMAC 字段为 16 字节。

图 5 Accept Session 消息格式



字段	长度	描述
Accept	1 byte	表示服务器意愿继续交互，0 表示服务器接受认证，愿意后续交互。非 0 表示服务器不接受认证。
MBZ	1 byte	置 0，接收时忽略。
Port	2 bytes	回应消息中 Port 的含义取决于请求消息中的 Conf-Sender 和 Conf-Receiver 的值。如果两者都置位，则 Port 字段是不使用的。如果只是 Conf-Sender 置位，则 Port 表示接受 TWAMP-Test 报文的端口。如果只是 Conf-Receiver 置位，Port 表示 TWAMP-Test 报文发送的端口。
SID	16 bytes	如果只发送了 Conf-Sender，回应消息里的 SID 字段是不使用的。否则，SID 唯一标识一个会话。
MBZ	12 bytes	置 0，接收时忽略。
HMAC	16 bytes	TWAMP 使用的 HMAC 是 HMAC-SHA1，128 比特，所以 HMAC 字段为 16 字节。

图 6 Start Session 消息格式

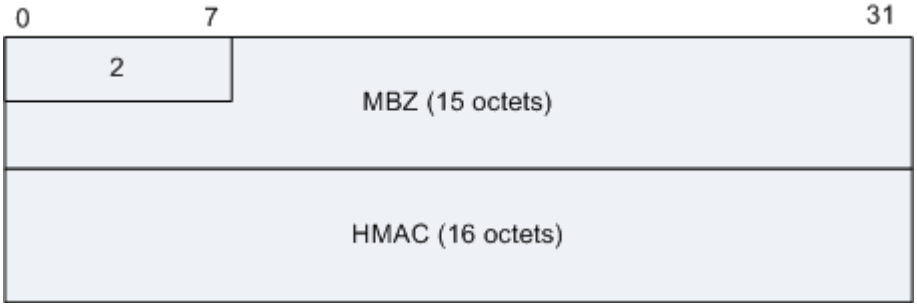
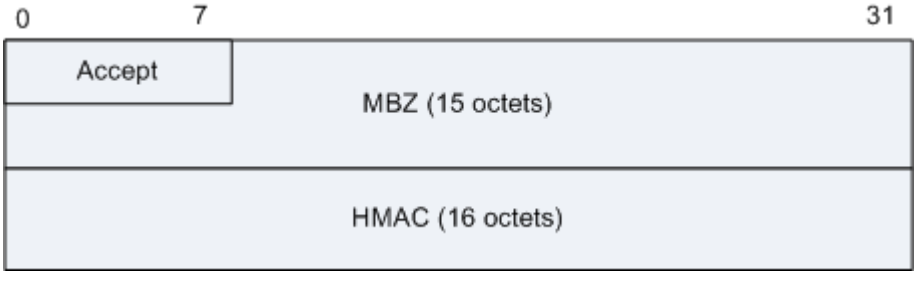
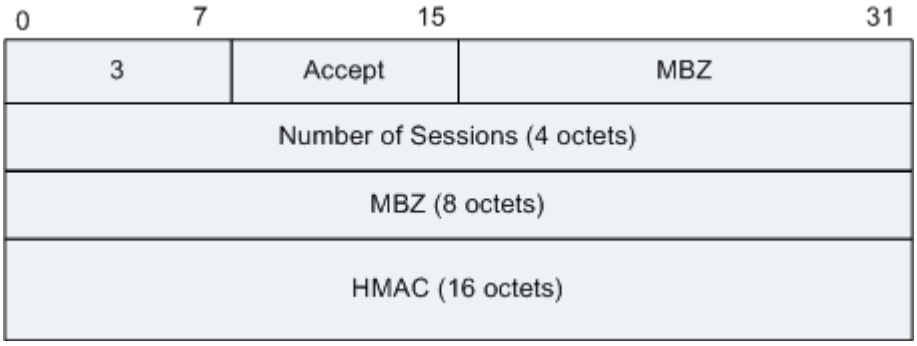


图 7 Start ACK 消息格式



字段	长度	描述
Accept	1 byte	如果是非 0 值，Start-Sessions 请求将被拒绝。0 表示接受。
MBZ	15 bytes	置 0，接收时忽略。
HMAC	16 bytes	TWAMP 使用的 HMAC 是 HMAC-SHA1，128 比特，所以 HMAC 字段为 16 字节。

图 8 Stop Session 消息格式



字段	长度	描述
Accept	1 byte	如果是非 0 值，表示故障，0 表示正常。
MBZ	2 bytes	置 0，接收时忽略。
Number of Sessions	4 bytes	表示 Control-Client 即将停止的会话的数量。
MBZ	8 bytes	置 0，接收时忽略。
HMAC	16 bytes	TWAMP 使用的 HMAC 是 HMAC-SHA1，128 比特，所以 HMAC 字段为 16 字节。

TWAMP 测量报文的格式

图 9 Sender-test

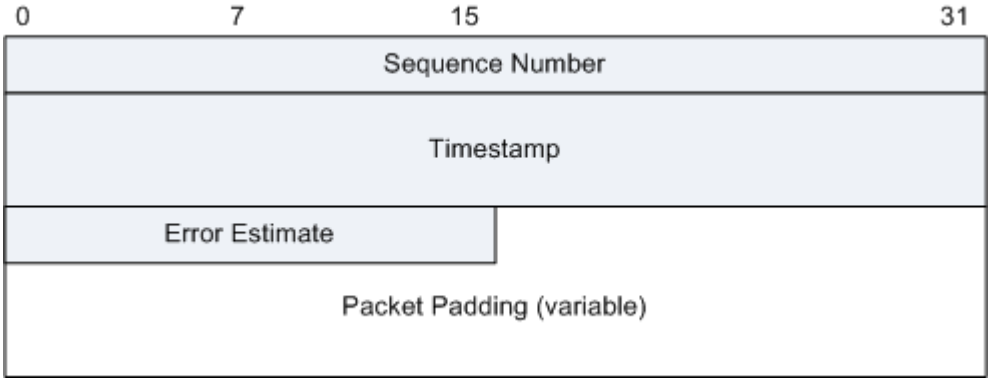
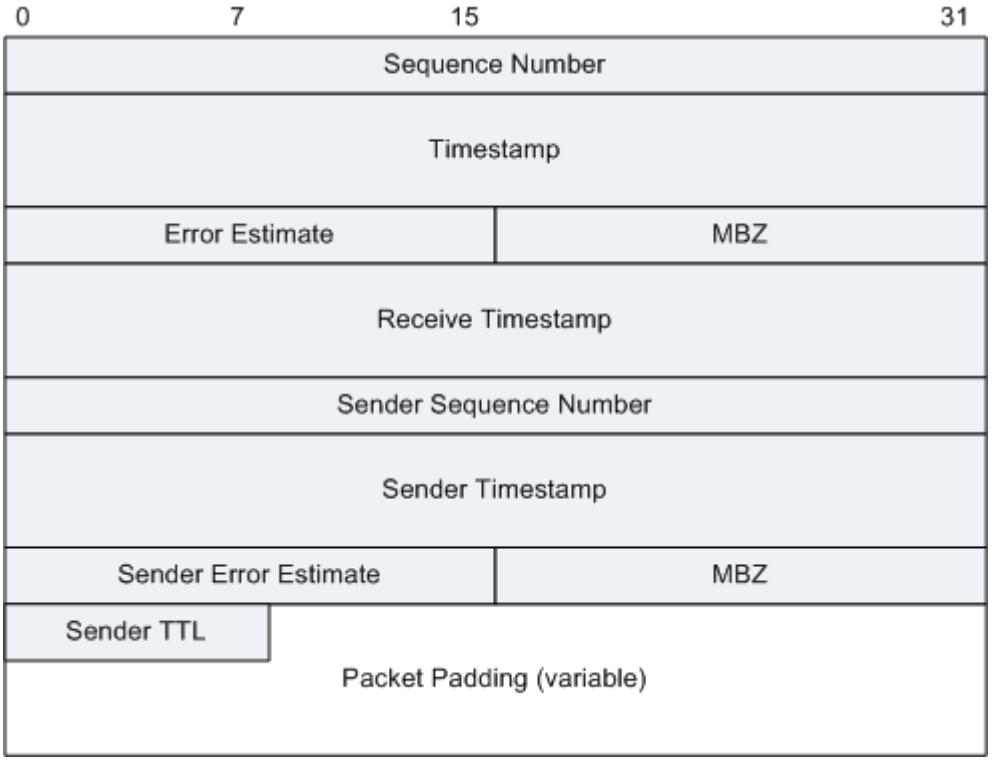
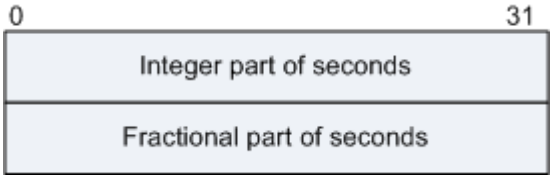


图 10 Reflector test



字段	含义
Sequence Number	根据传输顺序生成的报文序列号。从 0 开始并逐包增加，每个报文分配一个序号。Session-Reflector 生成的报文序列号与到达的报文的序列号无关。
MBZ	必须填为 0，客户端忽略此字段。
Timestamp	时间戳字段表示 Session-Reflector 反射的检测报文所打上的传输时间戳。其定义和格式请参见 OWAMP 标准[RFC4656]的第 4.1.2 章节。 时间戳的格式与 OWAMP [RFC4656]标准定义的时间戳格式相同，如下： **图 11** 时间戳字段的格式

字段	含义
	 与 RFC1305 标准定义的时间戳格式相同，前 32 比特表示从 1900 年 1 月 1 日 0 时至今的秒数的整数部分，后 32 比特表示小数部分。
Error Estimate	错误检测字段表示 Session-Reflector 反射的错误检测，其定义和格式请参见 OWAMP 标准[RFC4656]的第 4.1.2 章节。错误检测字段指定了错误的检测和同步。其格式如下： **图 12** Error Estimate 字段格式  • S：如果产生时间戳的设备的时钟与使用外部源的 UTC 同步，则 S 位置 1。例如，如果使用 GPS 硬件，此比特必须置位，表示需要时钟源的当前位置和时间，或者使用了 NTP，此比特也必须置位，表示与外部时钟源同步。如果没有外部时钟源同步需求，则此比特置 0。 • Z：与 MBZ 字段相同，必须置 0，接收端忽略此字段。 • Scale 和 Multiplier 字段都是无符号整数，Error Estimate = Multiplier*2^(-32)*2^Scale，单位是秒。Multiplier 字段不能为 0，如果为 0，表示为错误报文必须被丢弃。
Sender Timestamp 和 Sender Error Estimate	这两个字段是从 Session-Sender 的检测报文的对应字段里复制过来的。
Sender TTL	Session-Sender 发送的检测报文的 Sender TTL 设置为 255。Session-Reflector 发送的检测报文里的 Sender TTL 设置为 IP 报文头的 TTL 值。
Receive Timestamp	表示 Session-Reflector 接收到检测报文的时间。
Sender Sequence Number	发送序列号是从 Session-Sender 发送的报文的序列号复制的。
HMAC	TWAMP 测量报文的 HMAC 字段包含了 AES（Advanced Encryption Standard）加密的字段，所以在认证模式下，HMAC 包含了第 1 个块（16 字节）。在加密模式下，HMAC 包含了前 6 个块（96 字节）。TWAMP 测量报文中，HMAC 字段不能被加

字段	含义
	密。
Packet Padding	TWAMP 测量报文的报文填充字段。该字段不能被加密。 TWAMP 测量报文的数据段在非认证模式时最小长度为 41 字节，认证模式或加密模式下最小长度为 104 字节。

参考标准

标准	描述
RFC 5357	Two-Way Active Measurement Protocol
RFC 4656	A One-way Active Measurement Protocol (OWAMP)