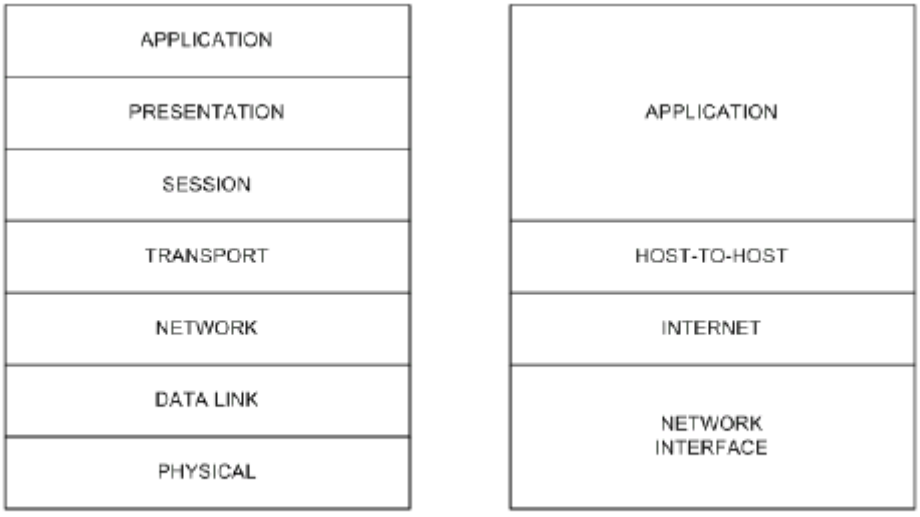
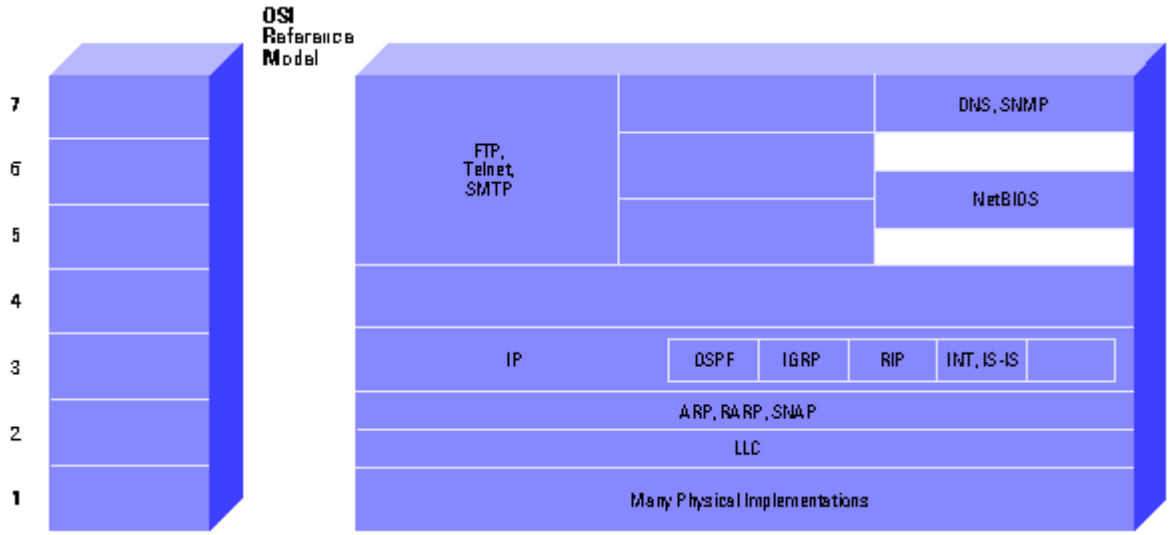
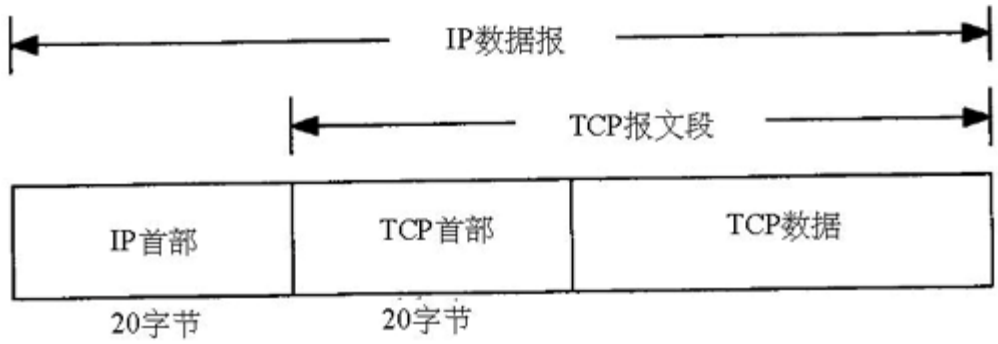
Packet Format Of Internet

NETWORK MODEL





TCP 首部:



16								32 bit
Source port								Destination port
Sequence number								
Acknowledgement number								
Offset	Reserved	U	A	P	R	S	F	Window
Checksum								Urgent pointer
Option + Padding								
Data								

- Source Port - 识别上层源处理器接收 TCP 服务的点。
- Destination Port - 识别上层目标处理器接收 TCP 服务的点。
- Sequence Number - 通常指定分配到当前信息中的数据首字节的序号。在连接建立阶段，该字段用于设别传输中的初始序列号。
- Acknowledgment Number - 包含数据包发送端期望接收的数据下一字节的序列号。一旦连接成功，该值会一直被发送。
- Data Offset - 4 位。TCP 协议头中的 32 位字序号表示数据开始位置。
- Reserved - 6 位。预留以备用，必须设置为 0。
- Control Bits （Flags） - 6 位。传送各种控制信息。控制位可以是：

U （URG）	Urgent pointer field significant.
A （ACK）	Acknowledgment field significant.

 MSN : MingLang.Xu@Hotmail.Com

P （PSH）	Push function.
R （RST）	Reset the connection.
S （SYN）	Synchronize sequence numbers.
F （FIN）	No more data from sender.

- Window - 16 位。指定发送端接收窗口的大小，也就是说，数据可用的八位缓存区大小。
- Checksum - 16 位。指出协议头在传输中是否遭到破坏。
- Urgent Pointer - 16 位。指向数据包中的第一个重要数据字节。
- Option + Padding - 指定各种 TCP 选项。可选项有两种可能形式：单个八位可选类型和八位可选类型，八位可选长度和实际可选数据八位位组。
- Data - 包含上层信息。

Van Jacobson：压缩 TCP 协议

Van Jacobson 是一个压缩的 TCP 协议，改进了 TCP/IP 在低速（ 300~19200bps ）串行连接上的性能，解决了连接框架的制定、地址分配、路由选择、认证以及运行等问题。

由 Van Jacobson 协议提出的压缩与 Thinwire-II 协议在主旨上是类似的。但相比之下，该协议压缩更有效（与 Thinwire-II 的 13 字节相比，其平均压缩头是 3 字节），并且它在实施方面更加简单而有效。 Van Jacobson 压缩是特殊的 TCP/IP 数据包。

协议结构

压缩 TCP 格式如下所示：

	C	I	P	S	A	W	U
Connection Number （C）							
TCP Checksum							
Urgent Pointer （U）							
D Window （W）							
D Ack （A）							
D Sequence （S）							

D IP ID （I）
Data

- C、I、P、S、A、W、U — 改变掩码。识别字段期望改变值和每个数据包实际改变值。
- Connection Number — 用于分配 TCP 连接中最后数据包的保存备份。
- TCP Checksum — 保证终端对终端数据完整性校验仍然有效。
- Urgent Pointer — 如果设置 URG，才发送。
- D Values for Each Field — 描述由源 TCP 转变而来的连接字段数量（改变掩码中对每个字段都有指定）。

UDP:

16	32bit
Source port	Destination port
Length	Checksum
Data	

- Source Port — 16 位。源端口是可选字段。当使用时，它表示发送程序的端口，同时它还被认为是没有其它信息的情况下需要被寻址的答复端口。如果不使用，设置值为 0。
- Destination Port — 16 位。目标端口在特殊因特网目标地址的情况下具有意义。
- Length — 16 位。该用户数据报的八位长度，包括协议头和数据。长度最小值为 8。
- Checksum — 16 位。IP 协议头、UDP 协议头和数据位，最后用 0 填补的信息假协议头总和。如果必要的话，可以由两个八位复合而成。
- Data — 包含上层数据信息。

IPV4:

4	8	16	32bit
Version	IHL	Type of service	Total length
Identification		Flags	Fragment offset
Time to live		Protocol	Header checksum

 MSN : MingLang.Xu@Hotmail.Com

Source address
Destination address
Option + Padding
Data

- Version - 4 位字段，指出当前使用的 IP 版本。
- IP Header Length （IHL） — 指数据报协议头长度，具有 32 位字长。指向数据起点。正确协议头最小值为 5。
- Type-of-Service — 指出上层协议对处理当前数据报所期望的服务质量，并对数据报按照重要性级别进行分配。这些 8 位字段用于分配优先级、延迟、吞吐量以及可靠性。
- Total Length — 指定整个 IP 数据包的字节长度，包括数据和协议头。其最大值为 65,535 字节。典型的主机可以接收 576 字节的数据报。
- Identification — 包含一个整数，用于识别当前数据报。该字段由发送端分配帮助接收端集中数据报分片。
- Flags — 由 3 位字段构成，其中低两位（最不重要）控制分片。低位指出数据包是否可进行分片。中间位指出在一系列分片数据包中数据包是否是最后的分片。第三位即最高位不使用。
- Fragment Offset — 13 位字段，指出与源数据报的起始端相关的分片数据位置，支持目标 IP 适当重建源数据报。
- Time-to-Live — 是一种计数器，在丢弃数据报的每个点值依次减 1 直至减少为 0。这样确保数据包无止境的环路过程。
- Protocol — 指出在 IP 处理过程完成之后，有哪种上层协议接收导入数据包。
- Header Checksum — 帮助确保 IP 协议头的完整性。由于某些协议头字段的改变，如生存期（Time to Live），这就需要对每个点重新计算和检验。Internet 协议头需要进行处理。
- Source Address — 指定发送代码。
- Destination Address — 指定接收代码。
- Options — 允许 IP 支持各种选项，如安全性。
- Data — 包括上层信息。

IPv6/IPng：网际协议第 6 版

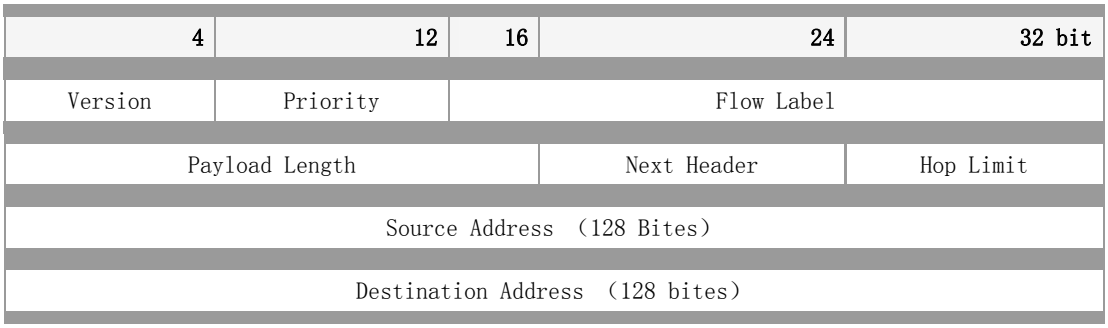
网际协议第 6 版（IPv6）是基于 IPv4 的最新版本， Ipv4 是一个网络层协议，它包含寻址信息和一些控制信息，可使数据包在网络中路由。 IP 协议有两种版本： IPv4 和 IPv6 ，其中 IPv6 也被称之为下一代 IP 或 IPng 。 IPv4 和 IPv6 都是关于介质层的去复用技术。例如，在 IPv6 中，在以太网上传送数据包采用的是 86DD （十六进制）， IPv4 用的是 0800 。本文主要介绍 IPv6 细节，有关 IP 和 IPv4 的内容在个别文件中另作介绍。

IPv6 把 IP 地址从 32 位增至 128 位，可以支持更多的寻址层次，更大数量的节点，以及更简单的地址自动配置，引入了组播地址的可缩放性，又定义了一个叫做“任意播”（anycast） 的新地址类

型， 用于给任意节点组发送数据包。相对于 IPv4 ， IPv6 主要有两个方面的改进：

- 支持扩展和选项的改进—— Ipv6 选项位于 IPv6 头和传输层头之间的单独协议头中。 IP 首部选项编码方式的改变使得传输过程更为高效，选项长度限制更少并且添加新选项更为灵活。扩展的头包括下一跳选项、路由选择、片断、目的选项、认证和封装负荷。
- 数据流标签能力——标签属于不同流量的数据包，用于发送端提出特殊处理请求，比如：非缺省服务质量或者 “ 实时 ” 服务。

协议结构



- Version — 网际协议版本号（IPv6 是第 6 版）；
- Priority — 流量类字段，识别发送数据包的优先权。优先权值的划分根据信息源提供拥塞控制和非拥塞流量控制的流量进行；
- Flow Label — 流标签用于信息源为需要特殊处理 IPv6 路由器的产品进行标签。该流由源地址和非零流标签共同唯一识别；
- 有效负载长度 — 有效负载长度包括协议头；
- Next Header — 迅速识别 IPv6 协议头后面的协议头类型；
- Hop Limit — 每个节点在转发数据包时的消耗。如果 Hop limit 消耗到 0，则取消数据包；
- Source Address — 数据包发送端 128 比特地址；
- Destination Address — 数据包的设定接收端 128 比特地址（未必是终点接收端）。

RIPng：路由选择信息协议下一代（应用于 IPv6）

下一代路由选择信息协议（RIPng ，应用于 IPv6 ）是一种基于 IPv4 网络协议和算法的协议。在国际性网络中，如因特网，拥有很多应用于整个网络的路由选择协议。形成网络的每一个自治系统（AS），都有属于自己的路由选择技术，不同的自治系统，路由选择技术也不同。自治系统内部的路由选择协议称为内部网关协议（IGP）。外部网关协议（EGP）是一种用于在自治系统之间传输路由选择信息的协议。

RIPng 在中等规模的 AS 中被用作 IGP 协议。对于较复杂的网络环境， RIPng 不适用。

RIPng 是一种距离向量（Distance Vector）算法。此协议所用的算法早在 1969 年， ARPANET 就用其来计算路由。然而该协议最初属于 XEROX 网络协议。 PUP 协议通过网关信息协议交换路由选择信息,而 XNS 则采用该协议的更新版本,命名为路由选择信息协议(RIP)实现路由选择信息交换。Berkeley 的路由协议很大程度上与 RIP 相同，即能够处理 IPV4 及其它地址类型的通用地址格式取代了 XNS 地址，同时路由选择每隔 30 秒更新一次。正是因为这种相似性， RIP 既适用于 XNS 协议，也适用于路由类协议。

关于 IPV4 网络，路由选择信息协议即指 RIP / RIP2，具体内容可参照相关说明。本文主要阐述 RIPng 。

协议结构

Command（1 byte）	Version（1 byte）	0（2 bytes）
Route Table Entry 1（20 bytes）		
……		
Route Table Entry N（20 bytes）		

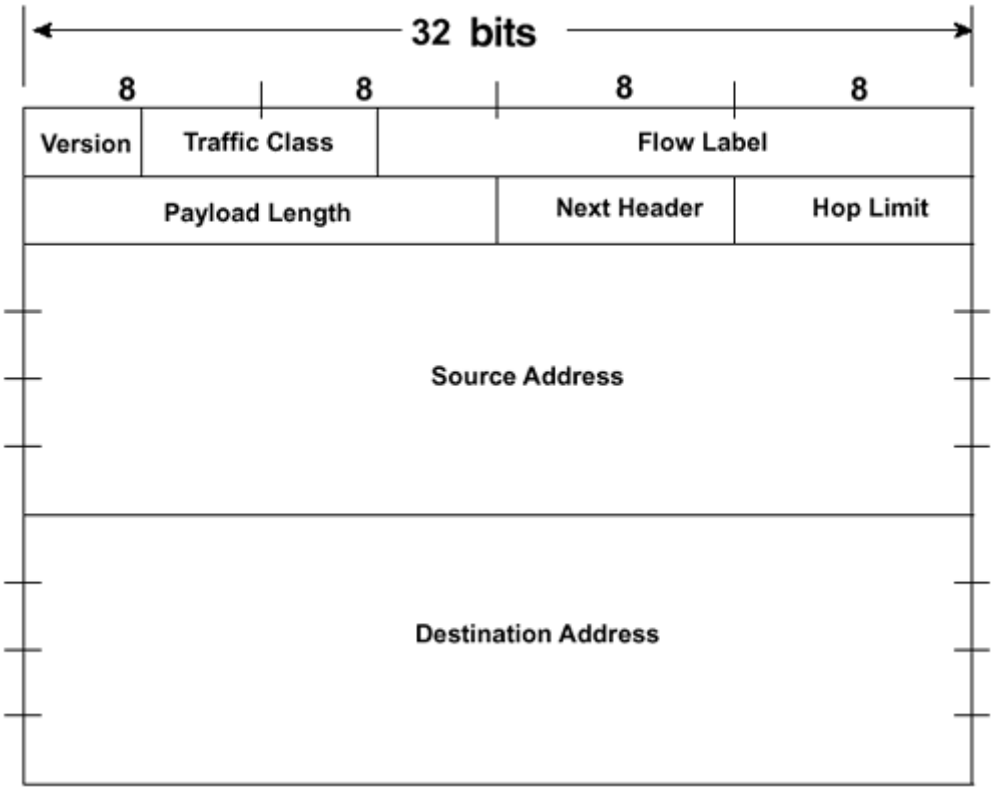
- Command — 两个命令是：

Request	A request for the responding system to send all or part of its routing table
Response	A message containing all or part of the sender’s routing table.

- Version — 协议版本号。当前为 1。
- Route Table Entry — 每个路由表入口包括一个目标前缀、前缀中重要位数目以及到达目的地的所消耗的成本。

MSN : MingLang.Xu@Hotmail.Com

IPv6 Header Format



The fields in the IPv6 header format are defined as follows:

Version indicates the IP version (in this case, 6).

Payload Length is the length of the IP packet, excluding this header, in octets. Extensionheaders, discussed in the next section, are considered part of the payload and are therefore included in this length.

Next Header is the value identifying the header immediately following the IPv6 header. The next header is either an upper-layer header (such as ICMP, TCP, or UDP) or it is an IPv6 extension header, as discussed in the next section.

Hop Limit is decremented by each node the packet traverses. The packet is discarded if the hop limit reaches zero. Some IPv6 functions, such as Router Advertisements, Neighbor Advertisements and Solicitations, and IPv6 Redirects, are used only between devices on a single link. A technique used by IPv6 processes to validate that a packet was not sent by an off-link node (perhaps as an attempt to maliciously redirect traffic) is to require the hop limit to be set to 255, which is the

maximum value for the hop limit. If the packet had traversed a router and was thus sent by an off-link node, the hop limit of the received packet would be something less than 255. An IPv6 node receiving this packet considers it invalid and drops it.

ı*Source Address/Destination Address* are 128-bit fields for the IPv6 source and destination addresses.

ICMP:

8	16	32bit
Type	Code	Checksum
Identifier		Sequence number
Address mask		

- Type — 错误消息或信息消息。错误消息可能是不可获得目标文件，数据包太大，超时，参数问题等。可能的信息消息有：Echo Request、Echo Reply、Group Membership Query、Group Membership Report、Group Membership Reduction。
- Code — 每种消息类型具有多种不同代码。不可获得目标文件正式这样一个例子，即其中可能的消息是：目标文件没有路由，禁止与目标文件的通信，非邻居，不可获得地址，不可获得端口。具体细节请参照相关标准。
- Checksum — 计算校验和时，Checksum 字段设置为 0。
- Identifier — 帮助匹配 Requests/Replies 的标识符，值可能为 0。
- Sequence Number — 帮助匹配 Requests/Replies 的序列号，值可能为 0。
- Address Mask — 32 位掩码地址。

TFTP:

基本 TFTP 协议头结构：

16 bits	String	16 bits	String	16 bits
Opcode	Filename	0	Mode	0

Opcode：操作代码或命令。以下为 TFTP 命令：

Opcode	Command	Description
1	Read Request	Request to read a file
2	Write Request	Request to write to a file
3	File Data	Transfer of file data
4	Data Acknowledge	Acknowledgement of file data
5	Error	Error indication

Filename：传送的字段名称。

Mode：数据模式。协议传输的文件数据格式。可以是 NetASCII，也可以是标准 ASCII，八位二进制数据或邮件标准 ASCII。

DNS:

16	21		28	32 bit
ID	Q	Query	A	T	R	V	B	Rcode

 MSN : MingLang.Xu@Hotmail.Com

Question count	Answer count
Authority count	Additional count

- ID - 用于连接查询和答复的 16bit。
- Q - 识别查询和答复消息的 1 位字段。
- Query - 描述消息类型的 4 位字段：
  0 标准查询（由姓名到地址）；
  1 逆向查询；
  2 服务状态请求
- A - 命令回答：1 位字段。当设置为 1 时，识别由命令名字服务器作出的答复。
- T - 切断。1 位字段。当设置为 1，表明消息已被切断。
- R - 1 位字段。由名字服务器设置为 1 请求递归服务。
- V -1 位字段。由名字服务器设置表示递归服务的实用性。
- B -3 位字段。备用，必须设置为 0。
- Rcode - 响应代码，由名字服务器设置的 4 位字段用以识别查询状态。
- Question count - 16 位字段用以定义问题部分的登陆号。
- Answer count - 16 位字段，用以定义回答部分的资源记录号。
- Authority count - 16 位字段，用以定义命令部分名字服务器的资源记录号。
- Additional count - 16 位字段，用以定义附加记录部分的资源记录号。

HTTP：超文本传输协议

超文本传输协议（HTTP）是应用层协议，由于其简捷、快速的方式，适用于分布式和合作式超媒体信息系统。自 1990 年起， HTTP 就已经被应用于 WWW 全球信息服务系统。

HTTP 允许使用自由答复的方法表明请求目的，它建立在统一资源识别器（URI）提供的参考原则下，作为一个地址（URL）或名字（URN），用以标志采用哪种方法，它用类似于网络邮件和多用途网际邮件扩充协议（MIME）的格式传递消息。

HTTP 也可用作普通协议，实现用户代理与连接其它 Internet 服务（如 SMTP 、 NNTP 、 FTP 、 GOPHER 及 WAIS ）的代理服务器或网关之间的通信，允许基本的超媒体访问各种应用提供的资源，同时简化了用户代理系统的实施。

HTTP 是一种请求 / 响应式的协议。一个客户机与服务器建立连接后，发送一个请求给服务器，请求的格式是：统一资源标识符（URI）、协议版本号，后面是类似 MIME 的信息，包括请求修饰符、客户机信息和可能的内容。服务器接到请求后，给予相应的响应信息，其格式是：一个状态行包括信息的协议版本号、一个成功或错误的代码，后面也是类似 MIME 的信息，包括服务器信息、实体信息和可能的内容。

HTTP 的第一版本 HTTP/0.9 是一种简单的用于网络间原始数据传输的协议。而由 RFC 1945 定义的 HTTP/1.0 ，在原 HTTP/0.9 的基础上，有了进一步的改进，允许消息以类 MIME 信息格式存在，包括请求 / 响应范式中的已传输数据和修饰符等方面的信息。但是， HTTP/1.0 没有充分考虑到分层代理服务器、高速缓冲存储器、持久连接需求或虚拟主机等方面的效能。相比之下， HTTP/1.1 要求更加严格以确保服务的可靠性。关于安全增强版的 HTTP （即 S-HTTP），将在相关文件中再作介绍。

协议结构

HTTP 报文由从客户机到服务器的请求和从服务器到客户机的响应构成。 请求报文格式如下：

请求行	通用信息头	请求头	实体头	报文主体

请求行以方法字段开始，后面分别是 URL 字段和 HTTP 协议版本字段，并以 CRLF 结尾。SP 是分隔符。除了在最后的 CRLF 序列中 CF 和 LF 是必需的之外，其他都可以不要。有关通用信息头，请求头和实体头方面的具体内容可以参照相关文件。

应报文格式如下：

状态行	通用信息头	响应头	实体头	报文主体

状态码元由 3 位数字组成，表示请求是否被理解或被满足。原因分析是对原文的状态码作简短的描述，状态码用来支持自动操作，而原因分析用来供用户使用。客户机无需用来检查或显示语法。有关通用信息头，响应头和实体头方面的具体内容可以参照相关文件。

NNTP：网络新闻传输协议

网络新闻传输协议（NNTP）是一种通过使用可靠的服务器－客户机流模式（如 TCP/IP 端口 119 ）实现新闻文章的发行、查询、修复及记录等过程的协议。借助 NNTP ，新闻文章只需要存储在一台服务器主机上，而位于其它网络主机上的订户通过建立到新闻主机的流连接阅读到新闻文章。 NNTP 为新闻组的广泛应用建立了技术基础。

NNTP 模型在新闻组网络系统（ USENET 新闻系统）后建成，但是 NNTP 对新闻文章的结构、内容及存放只作了很少的要求，因此，它很容易被其他非 USENET 系统采纳。使用 NNTP ，对于交流新闻文章的主机存在一种交互式机制来决定哪些文章需要传送。

主机想要获得新的新闻消息，或想知道哪台机器有新的新闻发送，需要通过 NNTP 联系一个甚至更多的网络邻居。然后主机客户端就会查询哪些新文章已经到达整个新闻组或某几个新闻组，这一过程借助于 NEWNEWS 指令完成。客户端将会从服务器端收到新文章的一个列表并请求传送那些他本身没有并且

想要的文章。最后，客户机告诉服务器它们已收到的文章。服务器会将那些已被拷贝的和哪些需要发送的添加到其收藏夹中，所以只有那些没有重复并且客户机想要的文章能够传输。

协议结构

NNTP 使用命令和响应实现通信。其中命令由命令字构成，在有些情况下带有参数。NNTP 具有很多命令。主要命令有：

- 新闻（信息 ID）：显示信头，空行及特定文章体（文本）。
- 信息 ID：可选域；是文章信息 ID，位于文章信头。如果是空，表示当前的文章是假设的。
- 信头：等同于 ARTICLE 命令，但它只返回文章信头。
- 状态：类似于 ARTICLE 命令，但它不返回文本信息。
- 组（ggg）：必需的参数 ggg 是选定的新闻组的名称。LIST 命令中包含一组有效的新闻组。成功选择响应会返回组中首尾两篇新闻的新闻号以及对存档新闻号估计。
- 新闻体：等同于 ARTICLE 命令，但它只返回新闻文本体。
- 目录：返回一列有效新闻组及相关信息。
- 新闻组：由日期和时间构成的一列新闻组会以和 LIST 命令相同的格式列出。
- NewNews：因为已经列出“日期”，所以特定的新闻组能传送或接收一组新闻信息 IDS。
- 下一个：内部维护的“当前新闻指示器”先进于当前新闻组中的下一个新闻。
- 邮件：如果邮件允许，返回响应代码 340，表示传送的新闻应该发送。
- 停止：服务器程序响应 QUIT 命令，然后关闭对客户机的连接。

S-HTTP：安全超文本传输协议

安全超文本传输协议（S-HTTP）是一种面向安全信息通信的协议，它可以和 HTTP 结合起来使用。S-HTTP 能与 HTTP 信息模型共存并易于与 HTTP 应用程序相整合。

S-HTTP 协议为 HTTP 客户机和服务器提供了多种安全机制，提供安全服务选项是为了适用于万维网上各类潜在用戶。 S-HTTP 为客户机和服务器提供了相同的性能（同等对待请求和应答，也同等对待客户机和服务器），同时维持 HTTP 的事务模型和实施特征。

S-HTTP 客户机和服务器能与某些加密信息格式标准相结合。 S-HTTP 支持多种兼容方案并且与 HTTP 相兼容。使用 S-HTTP 的客户机能够与没有使用 S-HTTP 的服务器连接，反之亦然，但是这样的通讯明显地不会利用 S-HTTP 安全特征。

S-HTTP 不需要客户端公用密钥认证（或公用密钥），但它支持对称密钥的操作模式。这点很重要因为这意味着即使没有要求用户拥有公用密钥，私人交易也会发生。虽然 S-HTTP 可以利用大多现有的认

证系统，但 S-HTTP 的应用并不必依赖这些系统。

S-HTTP 支持端对端安全事务通信。客户机可能“首先”启动安全传输（使用报头的信息），例如它可以用来支持已填表单的加密。使用 S-HTTP，敏感的数据信息不会以明文形式在网络上发送。

S-HTTP 提供了完整且灵活的加密算法、模态及相关参数。选项谈判用来决定客户机和服务器在事务模式、加密算法（用于签名的 RSA 和 DSA 、用于加密的 DES 和 RC2 等）及证书选择方面取得一致意见。

虽然 S-HTTP 的设计者承认他有意识的利用了多根分层的信任模型和许多公钥证书系统，但 S-HTTP 仍努力避开对某种特定模型的滥用。 S-HTTP 与摘要验证（在 [RFC-2617] 中有描述）的不同之处在于，它支持共钥加密和数字签名，并具有保密性。 HTTPS 作为另一种安全 web 通信技术，是指 HTTP 运行在 TLS 和 SSL 上面的实现安全 web 事务的协议。

协议结构

在语法上，S-HTTP 报文与 HTTP 相同，由请求或状态行组成，后面是信头和主体。显然信头各不相同并且主体密码设置更为精密。

正如 HTTP 报文，S-HTTP 报文由从客户机到服务器的请求和从服务器到客户机的响应组成。请求报文的格式如下：

请求行	通用信息头	请求头	实体头	信息主体

为了和 HTTP 报文区分开来，S-HTTP 需要特殊处理，请求行使用特殊的“安全”途径和指定协议“S-HTTP/1.4”。因此 S-HTTP 和 HTTP 可以在相同的 TCP 端口混合处理，例如，端口 80，为了防止敏感信息的泄漏，URI 请求必须带有“*”。

S-HTTP 响应采用指定协议“S-HTTP/1.4”。响应报文的格式如下：

状态行	通用信息头	响应头	实体头	信息主体

注意，S-HTTP 响应行中的状态并不表明展开的 HTTP 请求的成功或失败。如果 S-HTTP 处理成功，服务器会一直显示 2000K。这就阻止了所有请求的成功或失败分析。接受器由压缩数据对其中正确的作出判断，并接受所有的异常情形。

SNMP

SNMP 是一种应用程序协议，封装在 UDP 中。各种版本的 SNMP 信息通用格式如下所示：

Version	Community	PDU

- Version：SNMP 版本号。管理器和代理器必须使用相同版本的 SNMP。需要删除具有不同版本号的信息，并不对它们作进一步的处理。
- Community：团体名称，用于在访问代理器之前认证管理器。
- PDU（协议数据单元）：SNMPv1、v2 和 v3 中的 PDU 类型和格式将在对应文件中作具体介绍

SNMP V1:

SNMP 是一种应用程序协议，封装在 UDP 中。各种版本的 SNMP 信息通用格式如下所示：

Version	Community	PDU

- Version：SNMP 版本号。管理器和代理器必须使用相同版本的 SNMP。需要删除具有不同版本号的信息，并不对它们作进一步的处理。
- Community：团体名称，用于在访问代理器之前认证管理器。
- PDU（SNMPv1）：具有五种不同类型的 PDU：GetRequest、GetNextRequest、GetResponse、SetRequest 和 Trap。有关每部分的详细介绍请参见以下部分：

GetRequest、GetNext Request、GetResponse 和 SetRequest PDUs 格式如下所示：

PDU Type	Request ID	Error Status	Error Index	Object 1,Value 1	Object 2,Value 2	…

- PDU Type：指定传输的 PDU 类型：0 GetRequest；1 GetNextRequest；2 GetResponse；3 SetRequest。
- Request ID：连接 SNMP 请求和响应。
- Error Status：指出一个错误及错误类型。只有响应操作可以设置该字段，其它操作设置该字段为 0。
- Error Index：连接一个错误和一个特殊的对象实例。只有响应操作可以设置该字段，其它操作设置该字段为 0。
- Variable Bindings:用作 SNMPv1 PDU 的数据字段。每个变量绑定在当前值(除 Get 和 GetNext 请求之外，它们中值忽略不计）下都对应一个特殊对象实例

Trap PDU 格式如下所示：

PDU Type	Enterp	Agent Addr	Gen Trap	Spec Trap	Time Stamp	Obj 1,Val 1	Obj 1,Val 1	…

- PDU Type：指定传输的 PDU 类型（Trap＝4）。
- Enterprise：识别管理企业，在其注册权下定义 Trap。
- Agent Address：代理器的 IP 地址，用于进一步的识别。
- Generic Trap Type：描述事件报告字段，以下定义了 7 个值。
- Specific Trap Type：当通用 Trap 成为企业指定类型时，用于识别非通用 Trap。
- Timestamp：SysUpTime 对象值，表示最后一次设置初值和产生对应 Trap 间的时间数量。

SNMP V2:

SNMP 是一种应用程序协议，封装在 UDP 中。各种版本的 SNMP 信息通用格式如下所示：

Version	Community	PDU

- Version：SNMP 版本号。管理器和代理器必须使用相同版本的 SNMP。需要删除具有不同版本号的信息，并不对它们作进一步的处理。
- Community：团体名称，用于在访问代理器之前认证管理器。
- PDU（协议数据单元）：SNMPv1、v2 和 v3 中的 PDU 类型和格式将在对应文件中作具体介绍。

在 SNMPv2 中，Get、GetNext、Inform、Response、Set 和 Trap PDUs 具有以下格式：

PDU Type	Request ID	Error Status	Error Index	Object 1,value 1	Object 2,value 2	…

- PDU Type：识别传输的 PDU 类型（Get、GetNext、Inform、Response、Set 或 Trap）。
- Request ID：连接 SNMP 请求和响应。
- Error Status：指出一个错误及错误类型。只有响应操作可以设置该字段，其它操作设置该字段为 0。
- Error Index：连接一个错误和一个特殊的对象实例。只有响应操作可以设置该字段，其它操作设置该字段为 0。
- Variable Bindings：用作 SNMPv2 PDU 的数据字段（值 1，值 2…）。每个变量绑定在当前值（除 Get 和 GetNext 请求之外，它们中值忽略不计）下都对应一个特殊对象实例。

SNMPv2 GetBulk PDU 格式如下：

 MSN : MingLang.Xu@Hotmail.Com

PDU Type	Request ID	Non Repeaters	Max Repetitions	Obj 1, Val 1	Obj 1, Val 1	…

- PDU Type：识别 PDU 为 GetBulk 操作。
- Request ID：连接 SNMP 请求和响应。
- Non Repeaters：指定变量绑定字段中的对象实例号，并从请求开始多次进行检索。当实例是只有一个变量的标量对象时，使用该字段。
- Max Repetitions：定义除 Non Repeaters 字段指定的变量以外的变量检索次数最大值。
- Variable Bindings：用作 SNMPv2 PDU 数据字段（对象 1，对象 2……）。每个变量绑定在当前值（除 Get 和 GetNext 请求之外，它们中值忽略不计）下都对应一个特殊对象实例。

SNMP V3:

SNMPv3 信息格式：

Msg Processed by MPM (Msg Processing Model)					
Version	ID	Msg Size	Msg Flag	Security Model	
Msg Processed by USM (User Security Module)					
Authoritative Engin ID	Authoritative Boots	Authoritative Engine Time	User name	Authentication parameters	Privacy Parameter
Scoped PDU					
Context engine ID	Context name	PDU			

- Version：SNMPv3（3）。
- ID：用作两个 SNMP 实体间的唯一标识，以调整请求和响应信息。
- Msg Size：信息发送端所支持的八位信息最大值
- Msg Flags：八位的串，包含三个最不重要的标记位：ReportableFlag、PrivFlag、AuthFlag。
- Security Model：标识发送端使用的安全模式，接收端使用该安全模式处理该信息。
- AuthoritativeEngineID：SNMP 的 SnmpEngineID 值包括信息交换。因此，该值涉及 Trap 资源、响应或报告，通过 Get、GetNext、GetBulk、Set 或 Inform 发送至目的地。
- AuthoritativeEngineBoots：SNMP 的 snmpEngineBoots 值包括信息交换。
- AuthoritativeEngineTime：SNMP 的 SnmpEngineTime 值包括信息交换。
- User Name：发生信息交换的用户。
- AuthenticationParameters：如果交换没有被认证，则为空。否则它就是一个认证参数。

 MSN : MingLang.Xu@Hotmail.Com

- PrivacyParameters：不允许私有交换，则为空。否则它就是一个私有参数。
- PDU（Protocol Data Unit）：SNMPv3 中的 PDU 类型与 SNMPv2 中的相同。

RMON

网络层中 RMON 1 和 RMON 2 的主要监控中心：

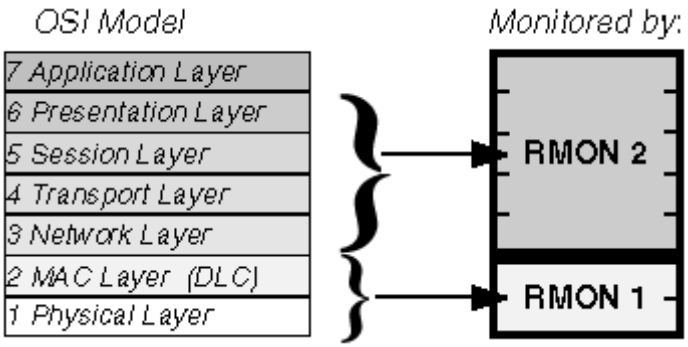


图 2 — 1 RMON 监控层

RMON 1 MIB 组	功能	元素
统计量	包括探测器为该设备每个监控的接口测量的统计值。	数据包丢弃、数据包发送、广播数据包、CRC 错误、大小块、冲突以及计数器的数据包。范围从 64~128、128~256、256~512、512~1024 以及 1024~1518 字节。
历史	定期地收集统计网络值地记录并存储起来以便日后提取。	取样周期、样品数目和项目。提供有关网段流量、错误包、广播包、利用率以及碰撞次数等其他统计信息的历史数据。

告警	定期从探测器的变量选取统计例子。并与前面配的阈值相比较。	告警类型、间隔、阈值上限、阈值下限
主机	包括网络上发现的与每个主机相关的统计值。	主机地址、数据包、接收字节、传输字节、广播传送等。
HostTopN	准备描述主机的表，根据一个统计值排序列表。	统计值、主机、周期的开始和结束、速率基值、持续时间。
真值表	记录关于子网上两个主机之间流量的信息，该信息以矩阵形式存储起来。	源地址和目的地址对、数据包、字节和每一对的错误。
过滤器	允许监视器观测与一过滤器相匹配的数据包。	字节过滤器类型、过滤器表达式等。
捕获包	数据包在流过一个信道之后被捕获。	捕获所有通过过滤器的数据包或简单地记下基于这些数据包的统计。
事件	控制在此处事件的产生和报告.	事件类型、描述、事件最后一个发送的时间
令牌环	支持令牌环	不常使用

RMON 2 MIB 组	功能
协议目录	协议目录是一种简单的便于共同建立 RMON2 应用程序、实现 RMON 代理的途径。这对于应用程序和代理出自不同的提供商的情况尤其重要。
协议分配	将探测器收集的数据转换为正确的协议名，从而可以显示给网络管理者。
地址变换	MAC 层的地址与网络层的地址之间的转换使得读和记忆变得容易。地址转换不仅为网络管理者提供了帮助，而且它支持 SNMP 管理平台并引入了改进的拓扑布局转换。
网络层主机	网络层主机(IP 层)统计值。
网络层真值表	在两个地址之间存储并重新获取网络层主机(IP 层)统计值。
应用层主机	应用层主机统计值。
应用层真值表	在两个地址之间存储并重新获取应用层主机(IP 层)统计值。
用户历史	这一特性使网络管理者能够配置系统中的任何历史记录，例如在指定文件服务器或路由器对路由器的连接上的特殊历史。
探测器配置	RMON2 的这一特性使某提供商的 RMON 应用程序能够配置其他提供商的 RMON 探测器。

 MSN : MingLang.Xu@Hotmail.Com

- Type — 0x11 信息类型（会员查询）
- Max Response Time — 只用于会员查询信息。规定每 1/10 秒中发送响应报告之前的最大允许时间。在所有其它信息中，发送方设置该值为 0，而接收方忽略不计。
- Checksum — 信息差错的校验和。
- Group Address — 当发送一个通用查询时，Group Address 设为 0。当发送一个特定组查询或组及特定源查询时，它被设置为正在查询的 Group Address。在离开组信息的会员报告中，该字段用于保存将要报告或离开的组的 IP 组播组地址。
- RSV — 预留。传输过程中设置为 0，接收方忽略不计。
- QQIC — 查询者的查询间隔代码。
- Number of Source （N） — 信息中源地址的数目。
- Source Address — IP 单播地址向量。

有关其它信息类型的具体内容情参照相关链接中的 RFC 1112、2236 和 3376。

IGMP：Internet 组管理协议

Internet 组管理协议（IGMP）是因特网协议家族中的一个组播协议，用于 IP 主机向任一个直接相邻的路由器报告他们的组成员情况。IGMP 信息封装在 IP 报文中，其 IP 的协议号为 2 。IGMP 具有三种版本，即 IGMP v1 、v2 和 v3 。

- IGMPv1 ：主机可以加入组播组。没有离开信息（leave messages）。路由器使用基于超时的机制去发现其成员不关注的组。
- IGMPv2 ：该协议包含了离开信息，允许迅速向路由协议报告组成员终止情况，这对高带宽组播组或易变型组播组成员而言是非常重要的。
- IGMPv3 ：与以上两种协议相比，该协议的主要改动为：允许主机指定它要接收通信流量的主机对象。来自网络中其它主机的流量是被隔离的。 IGMPv3 也支持主机阻止那些来自于非要求的主机发送的网络数据包。

IGMP 协议变种有：

- 距离矢量组播路由选择协议（DVMRP ： Distance Vector Multicast Routing Protocol）
- IGMP 用户认证协议 （IGAP ： IGMP for user Authentication Protocol）
- 路由器端口组管理协议（RGMP： Router-port Group Management Protocol）

协议结构

IGMP v3 必须实现 5 种基本信息类型且与以前的版本相兼容：

- 0x11：会员查询

- 0x22：第 3 版本会员报告
- 0x12：第 2 版本会员报告
- 0x16：第 2 版本会员报告
- 0x17：第 2 版本离开组

例如，0x11（会员查询）信息格式如下所示：

8		16		32bit
Type		Max Response Time		Checksum
Group address				
RSV	S	QRV	QQIC	Number of Source
Source Address （1）				
……				
Source Address （N）				

- Type — 0x11 信息类型（会员查询）
- Max Response Time — 只用于会员查询信息。规定每 1/10 秒中发送响应报告之前的最大允许时间。在所有其它信息中，发送方设置该值为 0，而接收方忽略不计。
- Checksum — 信息差错的校验和。
- Group Address — 当发送一个通用查询时，Group Address 设为 0。当发送一个特定组查询或组及特定源查询时，它被设置为正在查询的 Group Address。在离开组信息的会员报告中，该字段用于保存将要报告或离开的组的 IP 组播组地址。
- RSV — 预留。传输过程中设置为 0，接收方忽略不计。
- QQIC — 查询者的查询间隔代码。
- Number of Source （N） — 信息中源地址的数目。
- Source Address — IP 单播地址向量。

有关其它信息类型的具体内容情参照相关链接中的 RFC 1112、2236 和 3376

IGMPV1

4	7	15	23	31
Ver	Type	Unused	CheckSum	
Groud Address				

Ver:

Code Version = 1

Type:

1 = Host Membership Query

2 = Host Membership Report

 MSN : MingLang.Xu@Hotmail.Com

Group Address:

Multicast Group Address

• Version

– is the IGMP version and should be “0x1” in IGMPv1.

This field has been merged with the Type field in IGMPv2 and eliminated.

• Type

– is the IGMP message type.

0x1 = Host Membership Query

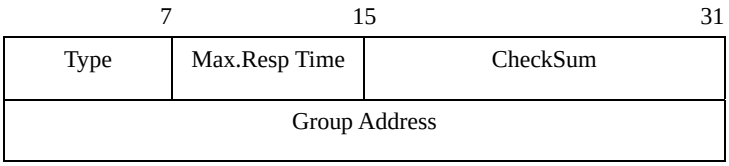
0x2 = Host Membership Report

This field has been expanded into an 8 bit field in IGMPv2

• Group

– is the Multicast Group address being specified for reports.

IGMPV2



Type:

0x11 = Membership Query

0x12 = Version 1 Membership Report

0x16 = Version 2 Membership Report

0x17 = Leave Group

Max. Resp. Time

max. time before sending a responding

report in 1/10 secs. (Default = 10 secs)

Group Address:

Multicast Group Address

(0.0.0.0 for General Queries)

• Type

– In IGMPv2, the old 4 bit “Version” field was merged with the old 4 bit “Type” field to create a new 8 bit “Type” field. By assigning IGMPv2 type codes `0x11` and `0x12` as the Membership Query (V1 & V2) and the “V1 Membership Report” respectively, backwards compatibility of IGMP v1 and v2 packet formats was maintained.

• Max. Response Time

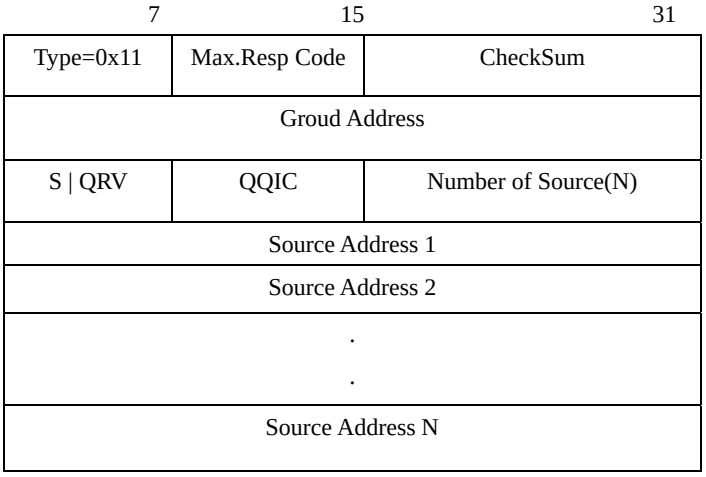
– This new field allows the querying router to specify exactly what the Query Interval Response Time is for this Query. The value (in 1/10 seconds) is used by the IGMPv2 hosts as the upper bound when randomly choosing the value of

their response timers. This helps to control the burstiness of the responses during the Query-Response interval.

• Group Address

– This field is identical to the IGMPv1 version of this field with the exception that it is set to `0.0.0.0` for General Queries.

IGMPV3



Type = 0x11

IGMP Query

Max. Resp. Time

Max. time to send a response

if < 128, Time in 1/10 secs

if > 128, FP value (12.8 - 3174.4 secs)

Group Address:

Multicast Group Address

(0.0.0.0 for General Queries)

S Flag

Suppresses processing by routers

QRV (Querier Robustness Value)

Affects timers and # of retries
QQIC (Querier's Query Interval)
Same format as Max. Resp. Time
Number of Sources (N)
(Non-zero for Group-and-Source Query)
Source Address
Address of Source

• Type
– The same IGMPv2 type code `0x11` is used as the IGMPv3 Membership Query Type code.
• Max. Response Time (1/10 seconds)
– This field has been reformatted to permit longer times to be expressed. If the value is < 128, the the time is absolute (.1- 12.8 seconds). If the value is > 128, it is interpreted as a floating-point number as follows:

```
+-+-+-+-+-+-+-+-+
|1| exp | mant | value = (mant|0x10)<<(exp+3)
+-+-+-+-+-+-+-+-+
```

• Group Address
– This field is identical to the IGMPv2 version of this field. It is set to `0.0.0.0` for General Queries.
• S Flag
– Indicates that the routers that receive this message should not process it.
• QRV (Querier Robustness Value)
– This value causes all hosts to adjust their Robustness Values which in turn affect various timers and retry counts. Increasing this value provides more protocol robustness at the expense of latency.
• QQIC (Querier' Query Interval)
– This field indicates the Query Interval in use by the Querying router. It's format is the same as the Max. Response Time field.
• Number of Sources
– The number of Source Addresses in the Group-and-Source-Specific Query.

CGMP：思科组管理协议

思科组管理协议（CGMP）主要用来限定只向与 IP 组播客户机相连的端口转发 IP 组播数据包。这些客户机自动加入和离开接收 IP 组播流量的组，交换机根据请求动态改变其转发行为。CGMP 主要提供以下服务：

- 允许只将 IP 组播数据包转发到连接 IP 组播客户机的那些端口。
- 通过限制不必要的 IP 组播流量，节省了网络带宽。
- 不需要改变终端主机系统。

- 不会产生为交换网络中的每个组播组创建独立 VLAN 的额外开销。

当启用了 CGMP 后，交换机能自动识别与启用了 CGMP 的路由器相连接的端口。CGMP 默认情况下是启用的，它支持最大为 64 个 IP 组播组注册。支持 CGMP 的组播路由器周期性地发送 CGMP 加入信息（join messages），向网络中的交换机通告自己的存在。交换机保存保存信息，并设置一个等同于路由器保持时间（holdtime）的计时器（timer）。交换机每次接收一个 CGMP 加入信息，定时器就会更新。当最后一个路由器保持时间失效时，交换机删除所有从 CGMP 学得的 IP 组播组。

CGMP 结合 IGMP 信息共同工作实现动态配置 Cisco Catalyst 交换机端口，从而使得 IP 组播流量只被转发给与 IP 组播客户机相连接的那些端口。由于启用了 CGMP 的 IP 组播路由器能够看到所有 IGMP 数据包，因此它可以通知交换机某个主机什么时候加入或离开 IP 组播组。当该路由器接收到一个 IGMP 控制数据包时，它会创建一个包含请求类型（加入或离开）、组播组地址和主机实际 MAC 地址的 CGMP 数据包。然后路由器将 CGMP 数据包发送到所有 Catalyst 交换机都在监听的一个地址上。当交换机接收 CGMP 数据包时，解读该数据包同时更改组播组的转发行为。至此，该组播流量只被发送到与 IP 组播客户机相连接的那些端口。该过程自动实现，无需用户参与。

协议结构

CGMP 信息格式：

1 byte	6 bytes	1 byte	6 bytes	1 byte
Count	Group Destination Address	Type	Unicast Source Address	Version

- Count：无符号 8 位整数；
- Group Destination Address：目标设备的硬件 MAC 地址；
- Type：信息类型；
- Unicast Source Address：单播源设备的硬件 MAC 地址；
- Version：CGMP 版本号。

RGMP：思科路由器端口组管理协议

思科路由器端口组管理协议(RGMP)弥补了 Internet 组管理协议(IGMP:Internet Group Management Protocol）在 Snooping 技术机制上所存在的不足。RGMP 协议作用于组播路由器和交换机之间。通过 RGMP，可以将交换机中转发的组播数据包固定在所需要的路由器中。RGMP 的设计目标是应用于具有多种路由器相连的骨干交换网（Backbone Switched Networks）。

IGMP Snooping 技术的局限性主要体现在：该技术只能将组播流量固定在接收机间经过其它交换机

直接或间接相连的交换端口，在 IGMP Snooping 技术下，组播流量不能固定在至少与一台组播路由器相连的端口处，从而引起这些端口的组播流量扩散。IGMP Snooping 是机制固有的局限性。基于此，路由器无法报告流量状态，所以交换机只能知道主机请求的组播流量类型，而不知道路由器端口接收的流量类型。

RGMP 协议支持将组播流量固定在路由器端口。为高效实现流量固定，要求网络交换机和路由器都必须支持 RGMP 。通过 RGMP，骨干交换机可以知道每个端口需要的组类型，然后组播路由器将该信息传送给交换机。但是路由器只发送 RGMP 信息，而忽视了所接收的 RGMP 信息。当组不再需要接收通信流量时，路由器会发送一个 RGMP 离开信息（Leave Message）。RGMP 协议中网络交换机需要消耗网络端口达到 RGMP 信息并对其进行处理操作。此外，RGMP 中的交换机不允许将接收到的 RGMP 信息转发/扩散到其它网络端口。

RGMP 的设计目标是与支持分配树 Join/Prune 的组播路由选择协议相结合使用。其典型协议为 PIM-SM。RGMP 协议只规定了 IP v4 组播路由选择操作，而不包括 IP v6。

协议结构

RGMP 信息格式与 IGMPv2 相同：

8	16	32 bit
Type	Reserved	Checksum
Group Address		

- Type — 路由器和交换机交互作用中存在四种 RGMP 信息。将 Type 代码定义为最大值(octet)以避免分配的 IGMP 类型代码的再利用：0xFF = Hello；0xFE = Bye；0xFD = Join a group；0xFC = Leave a group。
- Reserved — 信息中的 Reserved 字段必须以 0 传输，且在接收端忽略。
- Checksum — Checksum 包括 RGMP 信息（整个 IPv4 有效载荷）。校验和的算法和处理与 IGMP 中相同。
- Group Address — 在 RGMP Hello 或 Bye 信息中，Group Address 字段被设置为 0。在 RGMP Join 或 Leave 信息中，Group Address 字段持有将加入或离开的组的 IPv4 组播组地址。

DVMRP：距离矢量组播路由选择协议

距离矢量组播路由选择协议 （DVMRP） 是一种互联网路由协议，为互联网络的主机组提供了一种面向无连接信息组播的有效机制。 DVMRP 是一个“ 内部网关议 ”；适合在自治系统内的使用，不适合在不同的自治系统之间使用。当前开发的 DVMRP 不能用于为非组播数据报路由，因此要想一个路由器既能

为多播数据报又能为单播数据报路由，则它必须运行两个不同的路由选择进程。

DVMRP 的开发基于路由选择信息协议（RIP）。 DVMRP DVMRP 整合 RIP 中的许多特性和截断方向路径广播（TRPB ： Truncated Reverse Path Broadcasting）算法。另外，为了试验跨越不支持多播的网络可行性，开发了一种叫“ 隧道 ”的机制。 DVMRP 和 RIP 的主要不同之处在于：RIP 路由和转发数据报到明确的目的地。 DVMRP 的目的是为了跟踪到组播数据报出发地的返回路径。

DVMRP 数据包封装于 IP 数据报中,使用的 IP 协议号为 2 ,这点与 Internet 组管理协议（IGMP）相同。

协议结构

DVMRP 通过 IGMP 交换路由选择数据报。DVMRP 数据报由两部分组成：一个小型定长的 IGMP 头和一个标志数据流。

4	8	16	24	32bit
Version	Type	Sub-Type	Checksum	
DVMRP Data stream				

- Version — 版本号为 1。
  Type — DVMRP 类型为 3。
- Sub-Type — 子类型有：1 = Response，提供一些目的地路线。2 = Request，请求到达目的地的路线。3 = Non-Membership Report，提供非会员报告。4 = Non-Membership Cancellation，取消先前的非会员报告。
- Checksum — Checksum 必须基于传输进行计算并且基于数据包的接收而生效。DVMRP 信息的 Checksum 计算前提是 Checksum 字段设置为 0。

PIMv1 Packet Header

3 7 15 31

Type	Code	CheckSum
Ver	Reserved	

Code:
0 = Router-Query
1 = Register (SM only)
2 = Register-Stop (SM only)
3 = Join/Prune
4 = RP-Reachibility (SM only)
5 = Assert
6 = Graft
7 = Graft-ACK
Ver:
PIM Version = 1
PIMv1 used
IGMP Packets
• PIMv1 messages are multicast to the ALL-Routers (224.0.0.2) group with a TTL of 1.
Note: (PIMv1 packet formats are not shown. Only PIMv2 packets will be given.)

• PIM (v1) Packet Headers
• An IGMP type code of 0x14 indicates the frame is carrying a PIMv1 message - the code field then determines the type of PIM messages.
– PIMv1 messages are multicast to the ALL-Routers (224.0.0.2) multicast group address with a TTL of 1. This means that these control messages are Link-Local in scope.

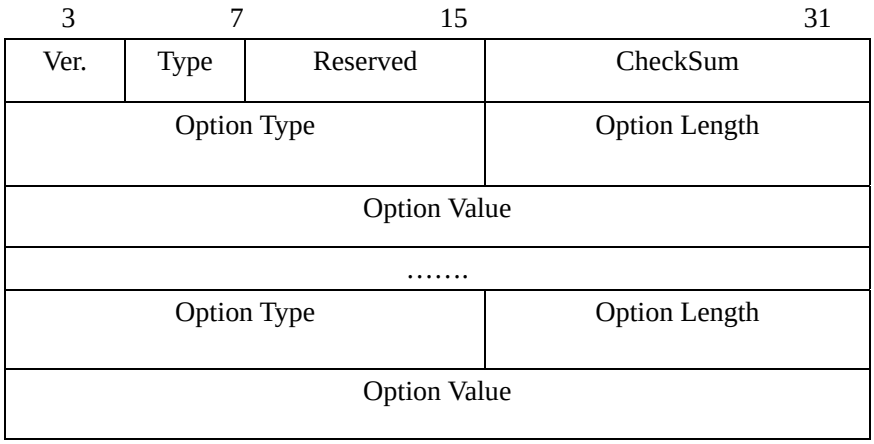
PIMv2 Packet Header

3	7	15	31
Ver.	Type	Reserved	CheckSum

Ver:
PIM Version = 2
Type:
0 = Hello
1 = Register (SM only)
2 = Register-Stop (SM only)
3 = Join/Prune
4 = Bootstrap (SM BSR only)

5 = Assert
6 = Graft (DM only)
7 = Graft-Ack (DM only)
8 = C-RP-Announcement (SM BSR only)
PIMv2 is assigned protocol number 103
• PIMv2 messages are multicast to the ALL-PIM-Routers (224.0.0.13) group with a TTL of 1.
• PIM (v2) Packet Headers
• PIMv2 packets are encoded in their own protocol packets using PIM assigned protocol number of 103. The Type field then determines the type of PIMv2 message.
– PIMv2 messages are multicast to the ALL-PIM-Routers (224.0.0.13) multicast group address with a TTL of 1. This means that these control messages are Link-Local in scope.

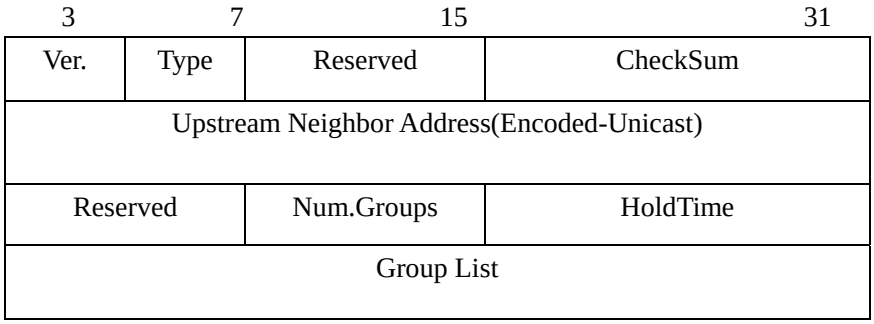
PIM Hello Messages



Option Types:
1 = Holdtime (Period of time in seconds before this PIM neighbor times out.)
19 = DR Priority
20 = Generation ID
• PIMv2 Hello Messages
• PIMv2 Hello messages are used to form and maintain neighbor adjacencies
– They are sent periodically to indicate to the other PIM routers on the network that this PIM router is still present.
• The PIMv2 Hello message format defines numerous Option TLV's which include:
– Holdtime: This specifies the time in seconds that this neighbor is reachable.

A value of 0xffff indicates the neighbor never times out. A value of 0x0000 means the neighbor is immediately timed out.
– DR Priority: This value can be used in the election of the DR for the subnet.
– Generation ID: This is a random 32-bit value that is sent whenever the neighbor actives PIM on the interface. It can be used to determine when the neighbor has been reactivated after a failure.

PIM Join/Prune Packets

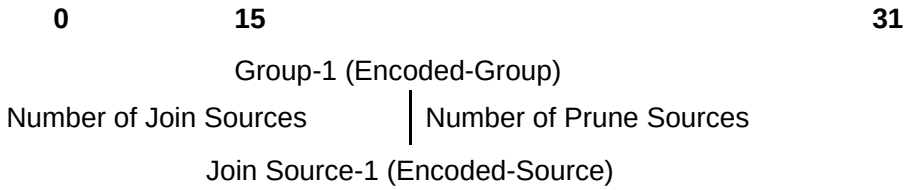


Upstream Neighbor Address:
IP address of RPF of
upstream neighbor
Holdtime:
Period of time in seconds
before this join/prune
times out.
Num. Grps
# of Groups in Group list
Group List:
List (by group) of sources
to Join and/or Prune

• PIM Join/Prune Packets

• The JOIN/PRUNE is a single packet format that contains a list of Joins and a list of Prunes. Either list may be empty (although not both).

PIM Group Lists



Join Source-n (Encoded-Source)

Prune Source-1 (Encoded-Source)

Prune Source-n (Encoded-Source)

Group-2 (Encoded-Group)

Number of Join Sources Number of Prune Sources

Group-*x*

Group IP Address

Number of Join Sources

# of Joins for Group-*x*

Number of Prune Sources

# of Prunes for Group-*x*

Join/Prune Source -*x*

Encoded Source address to be Joined/Pruned

• Group Lists

• Group Lists are used in Join/Prune messages as well as Graft and Graft -Ack messages.

• A Group List is a list of Group entries each beginning with a Group IP Address and Group mask to identify the Multicast Group.

• Each Group List entry contains a list of zero or more sources to Join followed by a list of zero or more sources to Prune.

– Group IP Address

– Number of Join Sources

– Number of Prune Sources

– Join List

– Prune List

• The addresses used in Join and Prune lists use a special encoded format that allows for other protocols besides IPv4. (See next slides.)

Encoded Unicast Addresses

3 7 10 15 31

Addr Family Encoding Type Unicast Address ...

… Unicast Address

Addr Family:

IANA Address Family Identifier (1=IPv4)

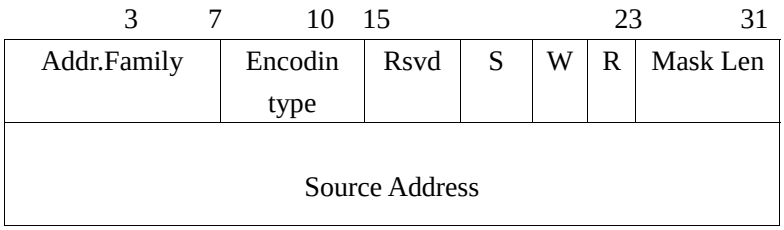
Encoding Type:
Type of encoding within Address Family
Unicast Address :
Unicast Address of the target device.

• Encoded Unicast Addresses

The Unicast Addresses contained in the Join and Prune Lists of a Group List entry are encoded in a special format as shown in the slide above.

• Address Family
– Indicates the IANA Address Family Identifier. For IPv4, this value is 1.
• Encoding Type
– Indicates the encoding type within the Address Family.
• Unicast Address
– IP unicast address of the target device

Encoded Source Addresses



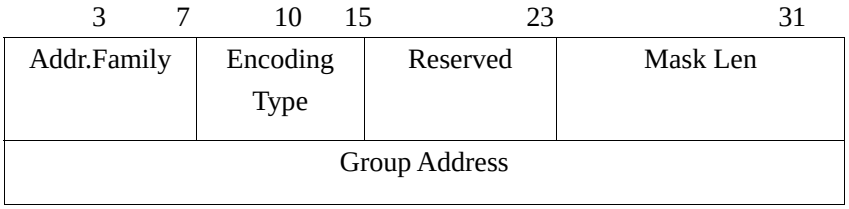
Addr Family:
IANA Address Family Identifier (1=IPv4)
Encoding Type:
Type of encoding within Address Family
S = Sparse Mode bit :
Indicates sparse mode group. Rcvrs must send periodic joins.
W = Wildcard bit :
Indicates join/prune applies to (*, G) entry.
R = RP bit :
Indicates this join/prune should be sent up the Shared Tree towards the RP.
Mask Length:
Number of bits in the prefix of the Group Address.
Source Address :
Address of Multicast Source.

• Encoded Source Addresses

The Source Addresses contained in the Join and Prune Lists of a Group List

entry are encoded in a special format as shown in the slide above.
• Address Family
– Indicates the IANA Address Family Identifier. For IPv4, this value is 1.
• Encoding Type
– Indicates the encoding type within the Address Family.
• S = Sparse Mode bit
– Used by routers on the Shortest-Path Tree (SPT) to indicate the group is a sparse mode group which tells the receiver of this join that it must send periodic Joins toward the source.
• W = Wildcard bit
– Indicates that the Join/Prune applies to the (*, G) entry. If this bit is cleared, it indicates that this applies to an (S, G) entry. Joins and Prunes sent to the RP should have this bit set.
• R = RP bit
– Indicates that this information should be sent up the Shared Tree towards the RP. If this bit is clear, the information should be sent up the Shortest-Path Tree toward the source.
• M. Len
– Mask length in bits.
• Source Address
– IP address of the Source.

Encoded Group Addresses



Addr Family:
IANA Address Family Identifier (1=IPv4)
Encoding Type:
Type of encoding within Address Family
Mask Length:
Number of bits in the prefix of the Group Address.
Group Address :
Multicast Group Address.

• Encoded Group Addresses

Group Addresses contained in the Join and Prune Lists of a Group List entry are encoded in a special format as shown in the slide above.
• Address Family
– Indicates the IANA Address Family Identifier. For IPv4, this value is 1.
• Encoding Type

– Indicates the encoding type within the Address Family.

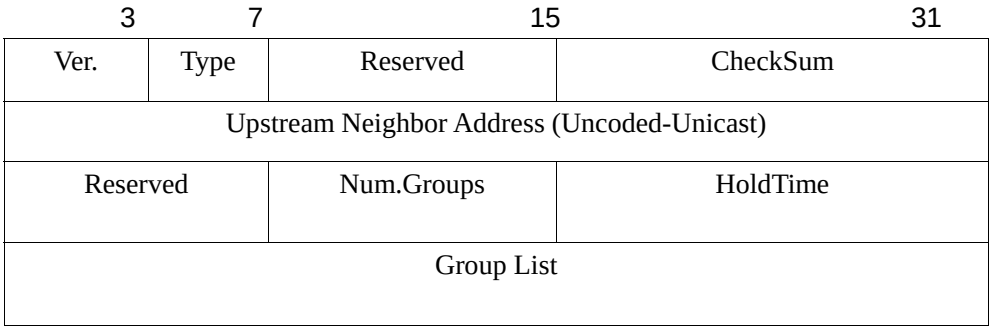
• M. Len

– Mask length in bits.

• Group Address

– IP multicast group address.

PIM Graft/Graft-Ack Packets



Upstream Neighbor Address:

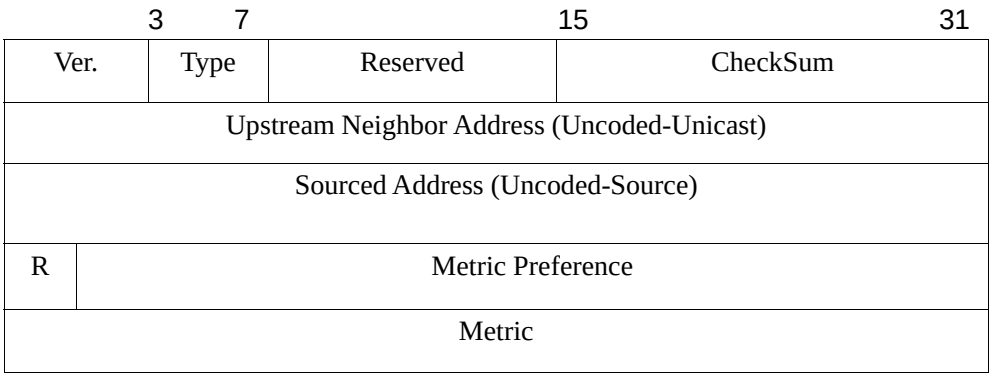
IP address of RPF of upstream neighbor Holdtime: Period of time in seconds before this join/prune times out. Num. Grps # of Groups in Group list Group List: List (by group) of sources to Graft or Graft-Ack

• PIM Graft/Graft-Ack Packets

• Graft/Graft-Ack are used in dense mode for grafting onto the tree

• These are the only PIM messages that are sent reliably (I.e. get an acknowledgement)

PIM Assert Packets



Group Address:

Identifies Group of the Assert

Source Address:

Identifies Source of the Assert

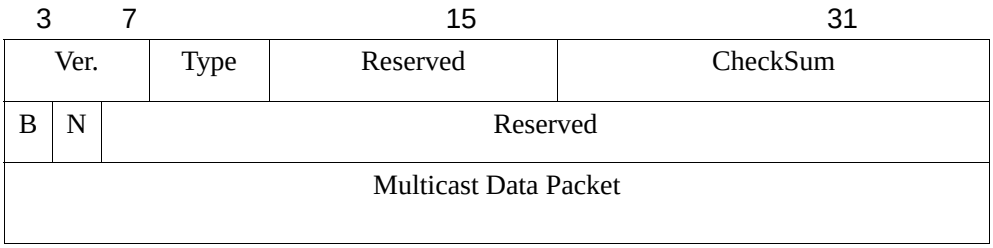
R: (Sparse Mode)

1 = Assert down RP Tree; 0 = Assert Down SPT

Metric Preference:
Admin. Distance of unicast routing protocol
Metric:
Unicast routing protocol metric
• PIM Assert Packets
• Assert messages determine who will be the active forwarder when there is redundancy in the network toward the source
• If the same routing protocol is used between the redundant neighbors, the metric is compared and the best metric wins
• In the case of an equal cost metric with the same routing protocol - the highest IP address neighbor will break the tie
• In the case where dissimilar unicast routing protocols are used, a metric preference is used to weight the preferred order of the routing information of each unicast routing protocol (like administrative distance)

PIM Register Packets

Sparse Mode Only



B = Border Bit:
Indicates DR is a border router performing a proxy-register
N = Null Register Bit:
Indicates DR is sending a Null-Register before expiring its register-suppression timer.
Multicast Data Packet:
The original packet sent by the source. For periodic sending of registers, this part is null.
• PIM Register Packets
• Used in SM by the DR to encapsulate multicast packets and send them to the RP so they may be forwarded down the shared tree.
• Register messages with encapsulated multicast packets continue to be sent to the RP by the DR until a Register-Stop message is received from the RP.

PIM Register-Stop Packets

3	7	15	31
Ver.	Type	Reserved	CheckSum
Group Address (Encoded-Group)			
Source Address (Encoded-Source)			

Group Address:

The group address from the register message.

Source Address:

IP host address of source from multicast data packet in register.

• PIM Register-Stop Packets

• Used in SM by the RP to inform the DR to stop sending Register messages.

This message is sent after the RP has joined the source tree to the DR and is receiving the multicast traffic natively via the SPT.

MSN : MingLang.Xu@Hotmail.Com

RIP/RIPV2

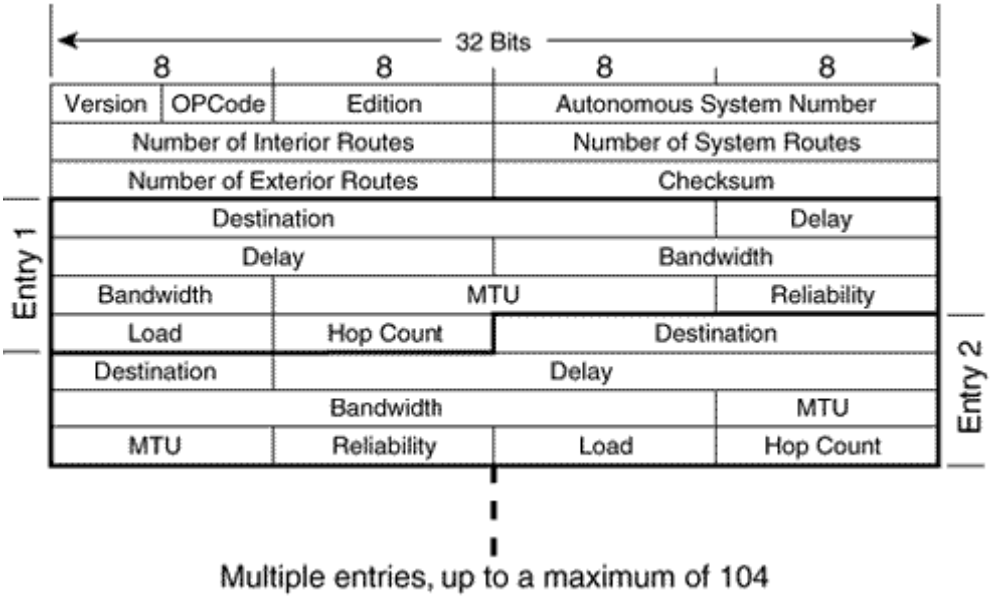
8	16	32bit
Command	Version	Unused
Address Family Identifier		Route Tag （only for RIP2; 0 for RIP）
IP Address		
Subnet Mask （only for RIP2; 0 for RIP）		
Next Hop （only for RIP2; 0 for RIP）		
Metric		

- Command — 该命令字段用来指定数据报用途。命令有五种：Request,Response,Traceon（已经淘汰），Traceoff（已经淘汰）和 Reserved。
- Version — RIP 版本号，当前为 2。
- Address Family Identifier — 指出该入口的地址类型。由于 RIP2 可能使用几种不同协议传送路由选择信息，所以使用到该字段。IP 中的 Address Family Identifier 为 2。
- Route Tag — 路由器指定属性，必须通过路由器保存和重新广告。路由标志是分离内部和外部 RIP 路由线路的一种常用方法（路由选择域内的网络传送线路），该方法在 EGP 或 IGP 都有应用。
- IP Address — 目标 IP 地址。
- Subnet Mask — 应用于 IP 地址，生成非主机地址部分。如果为 0，说明该入口不包括子网掩码。
- Next Hop — 中间下一跳 IP 地址，由路由入口指定的通向目的地的数据包需要转发到该地址。
- Metric — 表示从主机到目的地获得数据报过程中的整个成本。该 Metric 就是与网络相关联的成本总和。

MSN : MingLang.Xu@Hotmail.Com

IGRP

The IGRP packet format.



- Version — IGRP 版本号（当前值为 1）。
- Opcode — 操作码。表示信息类型：1 更新（Update）；2 请求（Request）。
- Edition — 序列号，路由表中发生任何改变时，该值会增加。
- Asystem — 自治系统号。一个网关能够应用于多个自治系统中，其中每个系统执行其自身的 IGRP。对于每个自治系统，都具有完全属于自己的独立路由表。该字段允许网关选择使用什么类型的路由表。
- Ninterior、Nsystem、Nexterior — 表示更新信息中登录这三个部分使用的编号。第一个登录（Ninterior）为内部登录，第二个登录（Nsystem）为系统登录，最后一个登录（Nexterior）是外部登录。
- Checksum — IP 校验和。其计算算法与 UDP 校验和算法相同。

 MSN : MingLang.Xu@Hotmail.Com

EIGRP

8	16	32bit
Version	Opcode	Checksum
Flags		
Sequence number		
Acknowledge number		
Asystem: Autonomous system number		
Type		Length

Version specifies the particular version of the originating EIGRP process. Although two software releases of EIGRP are currently available,[11] the version of the EIGRP process itself has not changed since its release. *Opcode* specifies the EIGRP packet type, as shown in Table 8.2. Although the IPX SAP packet type is included in the table, a discussion of IPX EIGRP is outside the scope of this book.

EIGRP packet types.

Opcode Type

1 Update

3 Query

4 Reply

5 Hello

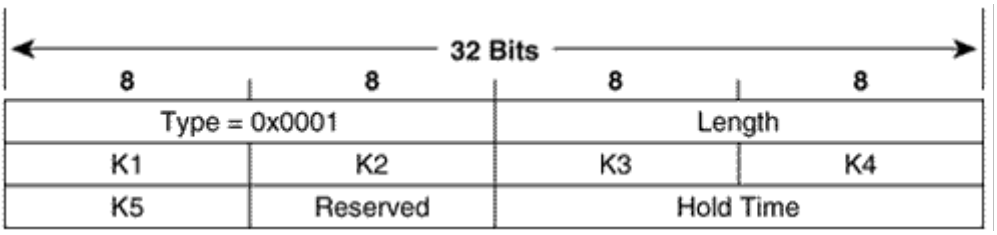
6 IPX SAP

Checksum is a standard IP checksum. It is calculated for the entire EIGRP packet, excluding the IP header.*Flags* currently include just two flags. The right-most bit is *Init*, which when set (0x00000001) indicates that the enclosed route entries are the first in a new neighbor relationship. The second bit (0x00000002) is the Conditional Receive bit, used in the proprietary Reliable Multicasting algorithm. *Sequence*is the 32-bit sequence number used by the RTP.*ACK* is the 32-bit sequence number last heard from the neighbor to which the packet. is being sent. A Hello packet with a nonzero ACK field will be treated as an ACK packet rather than as a Hello. Note that an ACK field will only be nonzero if the packet itself is unicast because acknowledgments are never multicast.*Autonomous System Number* is the identification number of the EIGRP domain.

TLV Fields please consult Routing TCP/IP V1

- Type — Type 字段值有：1、EIGRP 参数；2、预留；3、序列；4、软件版本；5、下一个组播序列。
- Length — 帧长。

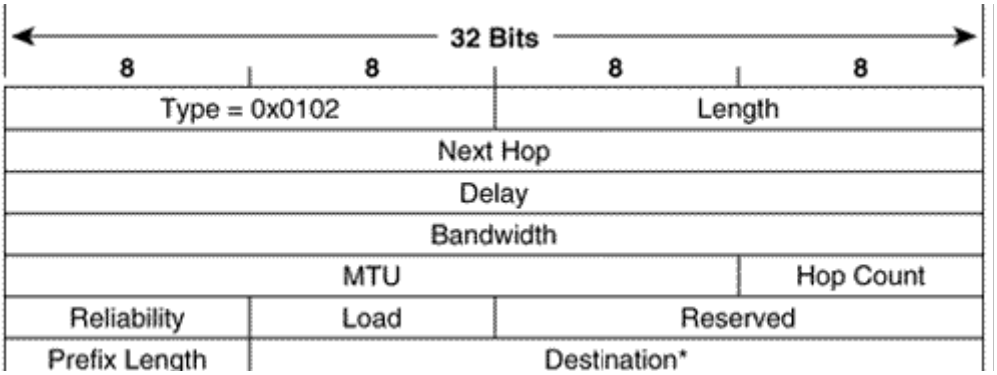
The EIGRP Parameters TLV.



IP-Specific TLV Fields

Each Internal and External Routes TLV contains one route entry. Every Update, Query, and Reply packet contains at least one Routes TLV. The Internal and External Routes TLVs include metric information for the route. As noted earlier, the metrics used by EIGRP are the same metrics used by IGRP, although scaled by 256.

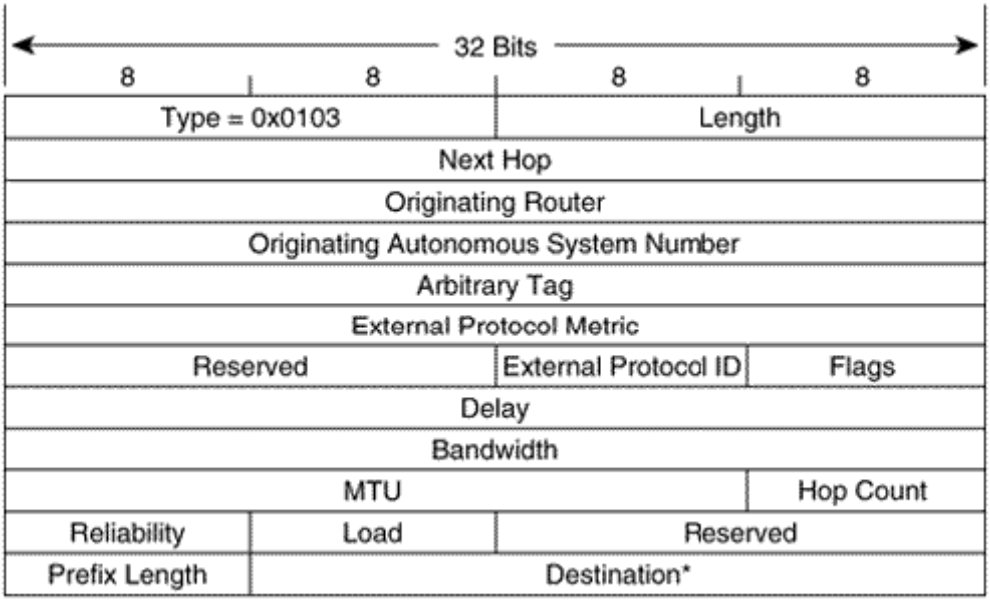
The IP Internal Routes TLV.



Next Hop is the next-hop IP address. This address may or may not be the address of the originating router.
Delay is the sum of the configured delays expressed in units of 10 microseconds. Notice that unlike the
24-bit delay field of the IGRP packet, this field is 32 bits. This larger field accommodates the 256
multiplier used by EIGRP. A delay of 0xFFFFFFFF indicates an unreachable route.
Bandwidth is 256 * $BW_{IGRP(min)}$, or 2,560,000,000 divided by the lowest configured bandwidth of any
interface along the route. Like Delay, this field is also eight bits larger than the IGRP field.
MTU is the smallest Maximum Transmission Unit of any link along the route to the destination. Although

an included parameter, it has never been used in the calculation of metrics.
Hop Count is a number between 0x01 and 0xFF indicating the number of hops to the destination. A router
will advertise a directly connected network with a hop count of 0; subsequent routers will record and
advertise the route relative to the next-hop router.
Reliability is a number between 0x01 and 0xFF that reflects the total outgoing error rates of the interfaces along the route, calculated on a 5-minute exponentially weighted average. 0xFF indicates a 100% reliable link.
Load is also a number between 0x01 and 0xFF, reflecting the total outgoing load of the interfaces along the route, calculated on a 5-minute exponentially weighted average. 0x01 indicates a minimally loaded link.
Reserved is an unused field and is always 0x0000.
Prefix Length specifies the number of network bits of the address mask. *Destination*is the destination address of the route. the field varies with the specific address. For example, if the route is to 10.1.0.0/16, the prefix length will be 16 and the destination will be a two-octet field containing 10.1. If the route is to 192.168.17.64/27, the prefix length will be 27 and the destination will be a four-octet field containing 192.168.16.64. If this field is not exactly three octets, the TLV will be padded with zeros to make it end on a four-octet boundary.

The IP External Routes TLV.

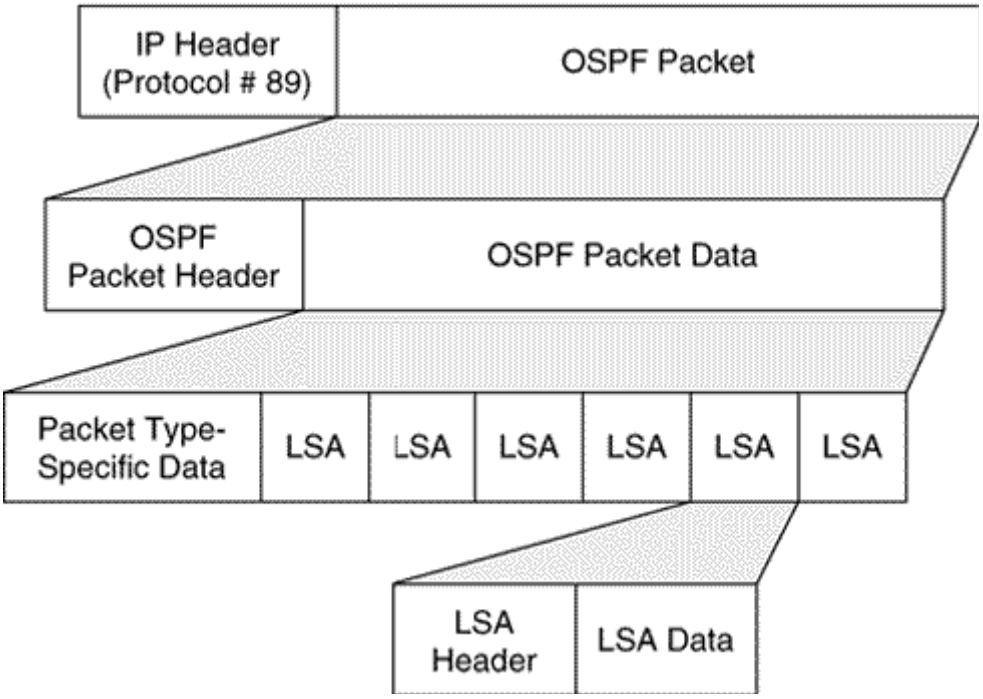


IP External Routes TLV
An external route is a path that leads to a destination outside of the EIGRP autonomous system and that has been redistributed into the EIGRP domain. *Next Hop* is the next-hop IP address. On a multiaccess network, the router advertising the route may not be the best next-hop router to the destination. For example, an

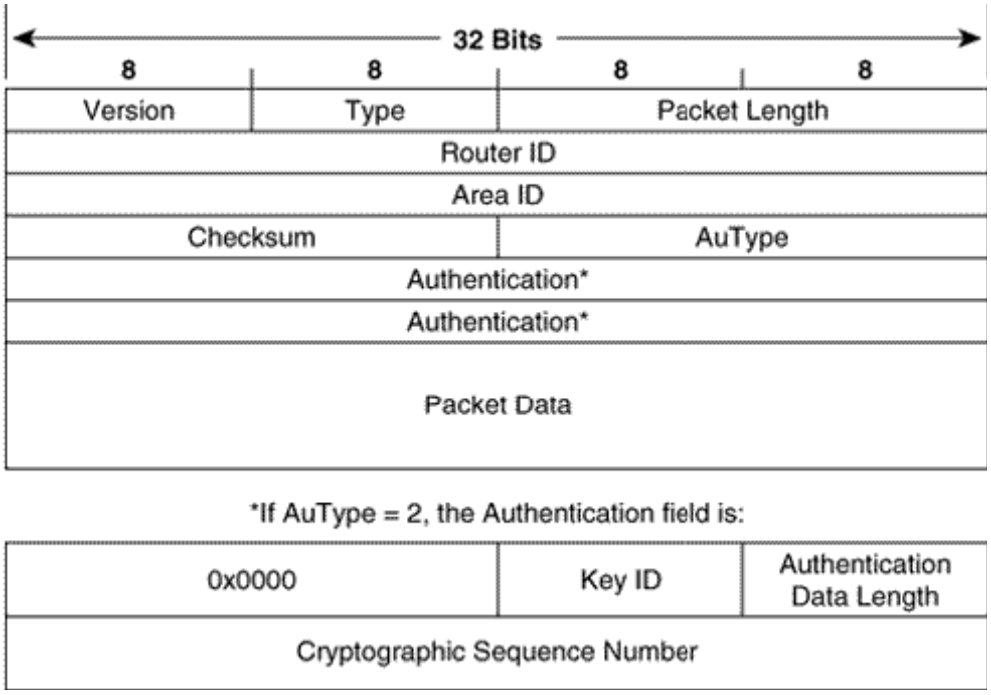
EIGRP-speaking router on an Ethernet link may also be speaking BGP and may be advertising a BGP-learned route into the EIGRP autonomous system. Because other routers on the link do not speak BGP, they may have no way of knowing that the interface to the BGP speaker is the best next-hop address. The Next Hop field allows the "bilingual" router to tell its EIGRP neighbors, "Use address A.B.C.D as the next hop instead of using my interface address." *Originating Router* is the IP address or router ID of the router that redistributed the external route into the EIGRP autonomous system. *Originating Autonomous System Number* is the autonomous system number of the router originating the route.
*Arbitrary Tag*may be used to carry a tag set by route maps. For information on the use of route maps. *External Protocol metric* is, as the name implies, the metric of the external protocol. This field is used, when redistributing with IGRP, to track the IGRP metric.*Reserved* is an unused field and is always 0x0000.

OSPF

An OSPF packet is composed of a series of encapsulations.



The OSPF packet header.



*If AuType = 2, the Authentication field is:

Version is the OSPF version number. As of this writing, the most recent OSPF version number is 2.(The Ospf V3 support IPV6 and add a few new feature).

OSPF packet types.

Type	Code Description
1	Hello
2	Database Description
3	Link State Request
4	Link State Update
5	Link State Acknowledgment

Packet length is the length of the OSPF packet, in octets, including the header.
Router ID is the ID of the originating router.
Area ID is the area from which the packet originated. If the packet is sent over a virtual link, the Area ID will be 0.0.0.0, the backbone Area ID, because virtual links are considered part of the backbone.
Checksum is a standard IP checksum of the entire packet, including the header.
AuType is the authentication mode being used.

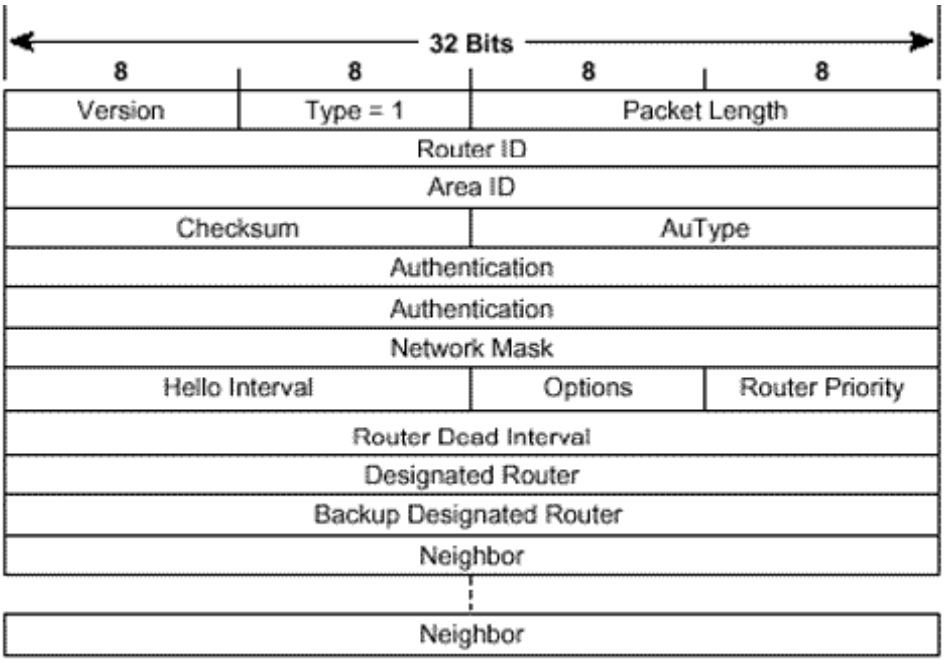
OSPF authentication types.

AuType	Authentication Type
0	Null (no authentication)
1	Simple (clear text) Password Authentication
2	Cryptographic (MD5) Checksum

Authentication is the information necessary for the packet to be authenticated by whatever mode is specified in the AuType field. If AuType = 0, the field is not examined and therefore may contain anything. If AuType = 1, the field will contain a password of up to 64 bits. If AuType = 2, the Authentication field will contain a Key ID, the Authentication Data Length, and a nondecreasing Cryptographic sequence number. The message digest is appended to the end of the OSPF packet, and is not considered part of the packet itself.
Key ID identifies the authentication algorithm and the secret key used to create the message digest.
Authentication Data Length specifies the length, in octets, of the message digest appended to the end of the packet.
Cryptographic Sequence Number is a nondecreasing number used to prevent replay attacks.

 MSN : MingLang.Xu@Hotmail.Com

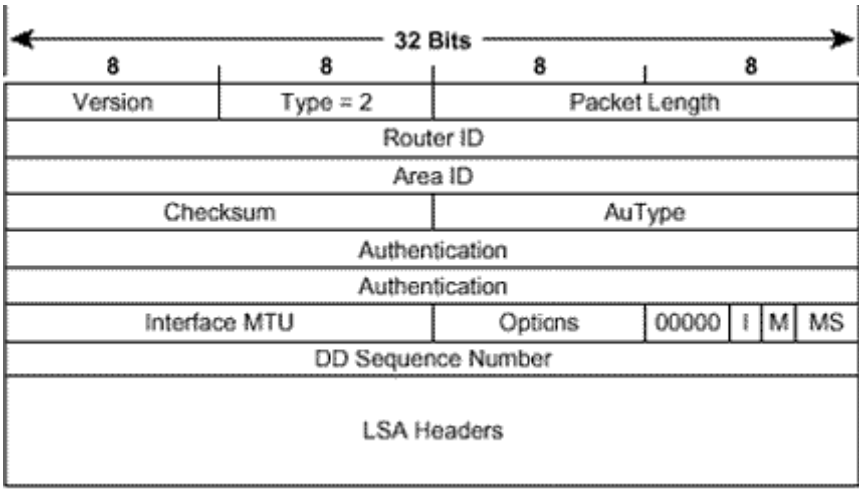
The OSPF Hello packet.



Network Mask is the address mask of the interface from which the packet was sent. If this mask does not match the mask of the interface on which the packet is received, the packet will be dropped. This technique ensures that routers will become neighbors only if they agree on the exact address of their shared network.
Hello Interval, as discussed earlier, is the period, in seconds, between transmissions of Hello packets on the interface. If the sending and receiving routers don't have the same value for this parameter, they will not establish a neighbor relationship.
Options are described in "The Options Field," later in this chapter. This field is included in the Hello packet to ensure that neighbors have compatible capabilities. A router may reject a neighbor because of a capabilities mismatch.
Router Priority is used in the election of the DR and BDR. If set to zero, the originating router is ineligible to become the DR or BDR.
Router Dead Interval is the number of seconds the originating router will wait for a Hello from a neighbor before declaring the neighbor dead. If a Hello is received in which this number does not match
the RouterDeadInterval of the receiving interface, the packet will be dropped. This technique ensures that
neighbors agree on this parameter.
Designated Router is the IP address of the interface of the DR on the network (not its Router ID). During the DR election process, this may only be the originating router's idea of the DR, not the finally elected
DR. If there is no DR (because one has not been elected or because the network type does not require DRs), this field will be set to 0.0.0.0.
Backup DR is the IP address of the interface of the BDR on the network. Again, during the DR election process, this may only be the originating router's idea of the

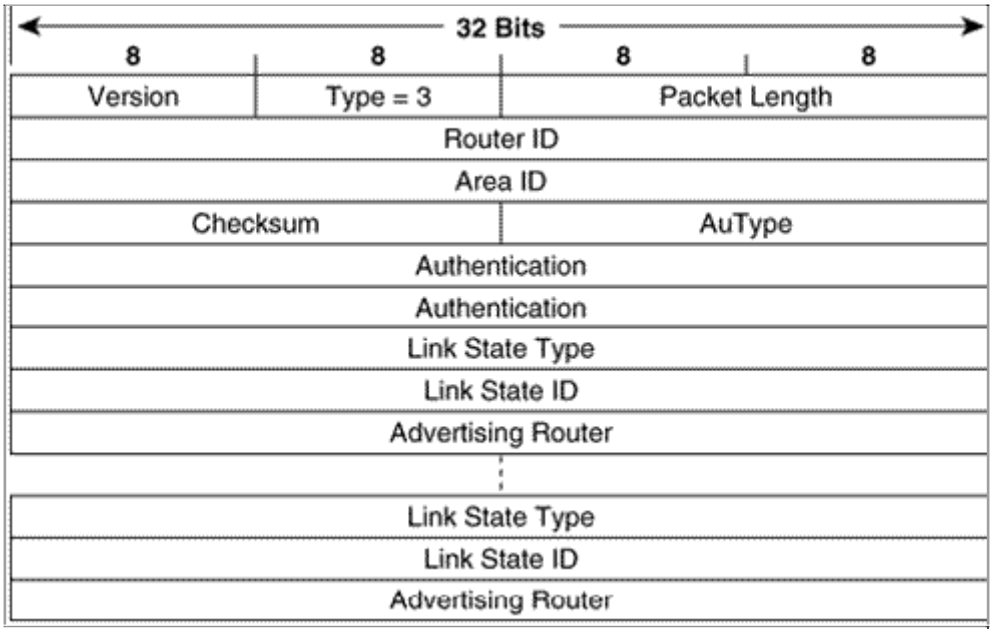
BDR. If there is no BDR, this field is set to 0.0.0.0.
Neighbor is a recurring field that lists all neighbors on the network from which the originating router has received a valid Hello in the past RouterDeadInterval.

The OSPF Database Description packet.



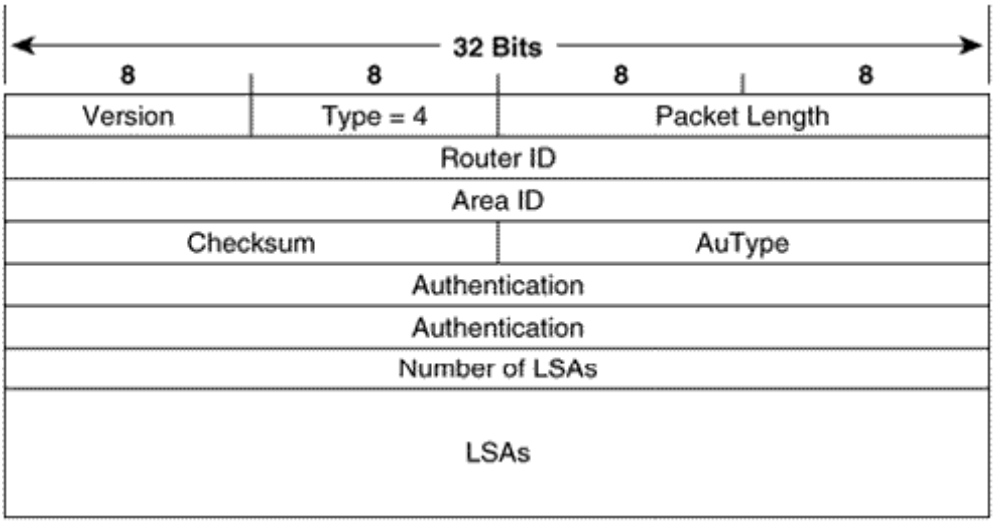
Interface MTU is the size, in octets, of the largest IP packet that can be sent out the originator's interface without fragmentation. This field will be set to 0x0000 when the packet is sent over virtual links.
Options are described in "The Options Field." The field is included in the Database Description packet so that a router may choose not to forward certain LSAs to a neighbor that doesn't support the necessary capabilities.
The first five bits of the next octet are unused and are always set to 00000b.
I-bit, or Initial bit, is set to 1 when the packet is the initial packet in series of DD packets. Subsequent DD packets will have I-bit = 0.
M-bit, or More bit, is set to 1 to indicate that the packet is not the last in a series of DD packets. The last DD packet will have M-bit = 0.
MS-bit, or Master/Slave bit, is set to 1 to indicate that the originator is the master (that is, is in control of
the polling process) during a database synchronization. The slave will have MS-bit = 0.
DD Sequence Number ensures that the full sequence of DD packets are received in the database
synchronization process. The sequence number will be set by the master to some unique value in the first DD packet, and the sequence will be incremented in subsequent packets.
LSA Headers list some or all of the headers of the LSAs in the originator's link state database. See "The Link State Header," for a full description of the LSA header; the header contains enough information to uniquely identify the LSA and the particular instance of the LSA.

The OSPF link state request packet.



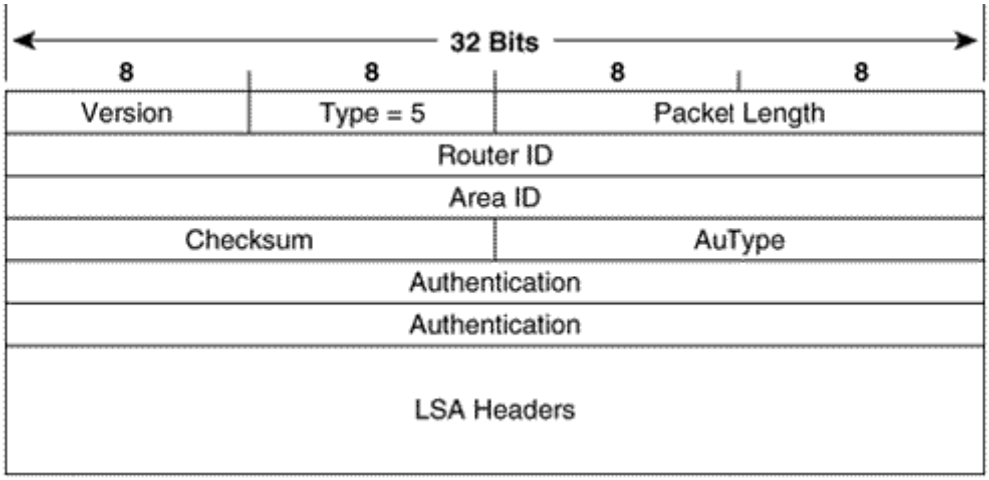
Link State Type is the LS type number, which identifies the LSA as a router LSA, network LSA, and so on.
Link State ID is a type-dependent field of the LSA header. See "The Link State Header" and the LSAspecific sections for a full description of how the various LSAs use this field.
Advertising Router is the Router ID of the router which originated the LSA.

The OSPF Link State Update packet.



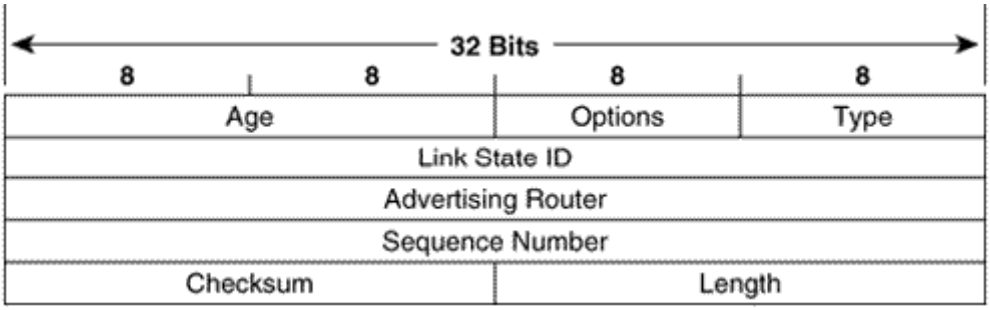
Number of LSAs specifies the number of LSAs included in this packet.
LSAs are the full LSAs as described in OSPF LSA formats. Each update may carry multiple LSAs, up tothe maximum packet size allowed on the link.

The OSPF Link State Acknowledgment packet.



Link State Acknowledgment packets are used to make the flooding of LSAs reliable. Each LSA received by a router from a neighbor must be explicitly acknowledged in a Link State Acknowledgment packet.The LSA being acknowledged is identified by including its header in the LS ACK packet, and multiple LSAs may be acknowledged in a single packet. he LS ACK packet consists of nothing more than an OSPF packet header and a list of LSA headers.

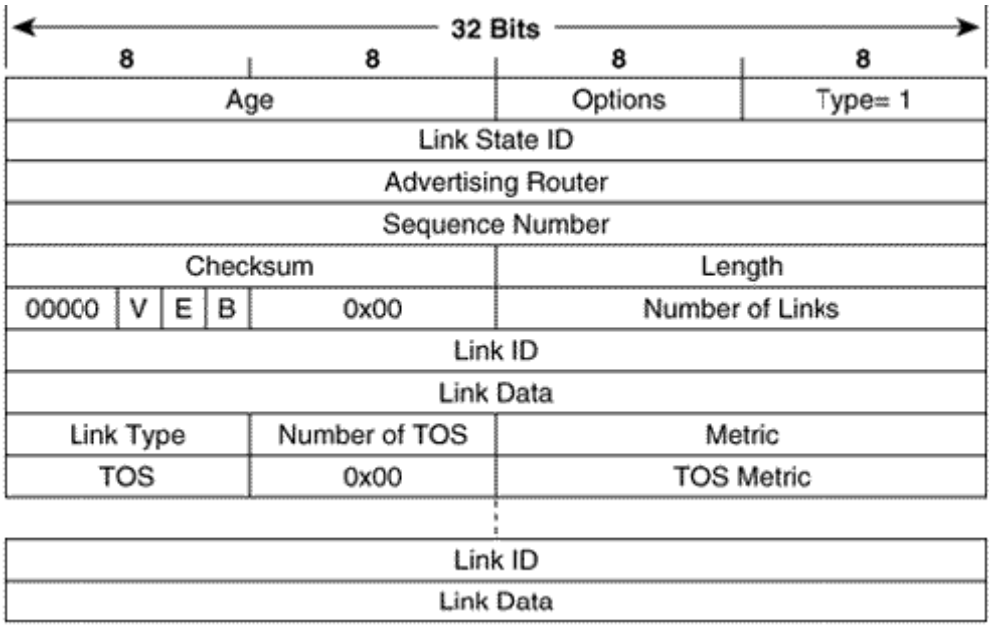
The OSPF LSA header.



Age is the time, in seconds, since the LSA was originated. As the LSA is flooded, the age is incremented
by InfTransDelay seconds at each router interface it exits. The age is also incremented in seconds as it
resides in a link state database.
Options is described in "The Options Field." In the LSA header, the Options field specifies the optional

capabilities supported by the portion of the OSPF domain described by the LSA.
Type is the LSA type. The type codes are shown in Routing TCP/IP V1 Table 9.4.
Link State ID identifies the portion of the OSPF domain being described by the LSA. The specific usage of this field varies according to the LSA type; the descriptions of each LSA include a description of how the LSA uses this field.
Advertising Router is the router ID of the router that originated the LSA.
Sequence Number is incremented each time a new instance of the LSA is originated. This update helps other routers identify the most recent instance of the LSA.
Checksum is the Fletcher checksum of the complete contents of the LSA except the Age field. If the Age field were included, the checksum would have to be recalculated every time the age was incremented.
Length is the number of octets of the LSA, including the header.

The OSPF Router LSA.



Link State ID for router LSAs is the originating router's Router ID.
V, or *Virtual Link Endpoint* bit, is set to one when the originating router is an endpoint of one or more fully adjacent virtual links having the described area as the transit area. *E*, or *External* bit, is set to one when the originating router is an ASBR. *B*, or *Border* bit, is set to one when the originating router is an ABR.
Number of Links specifies the number of router links the LSA describes. The router LSA must describe all of the originating router's links, or interfaces, to the area in which the LSA is flooded.
Subsequent fields in the Router LSA describe each link and appear one or more times, corresponding to the number in the Number of Links field. This discussion covers the Link Type field first, although that field does not appear until after the Link Data field. Understanding link type first is important because the descriptions of the Link ID and Link Data fields vary according to the value of the Link Type field.

 MSN : MingLang.Xu@Hotmail.Com

Link type values.

LinkType	Connection
1	Point-to-point connection to another router
2	Connection to a transit network
3	Connection to a stub network
4	Virtual link

Link ID values.

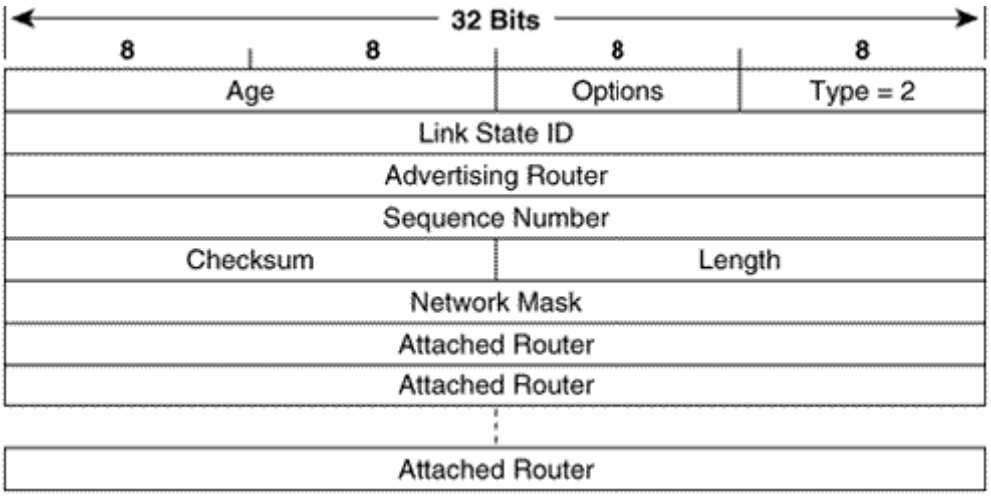
Link Type	Value of Link ID Field
1	Neighboring router's Router ID.
2	IP address of the DR's interface.
3	IP network or subnet address.
4	Neighboring router's Router ID.

Link data values.

Link Type	Value of Link Data Field
1	IP address of the originating router's interface to the network.*
2	IP address of the originating router's interface to the network.
3	Network's IP address or subnet mask.
4	The MIB-II ifIndex value for the originating router's interface.

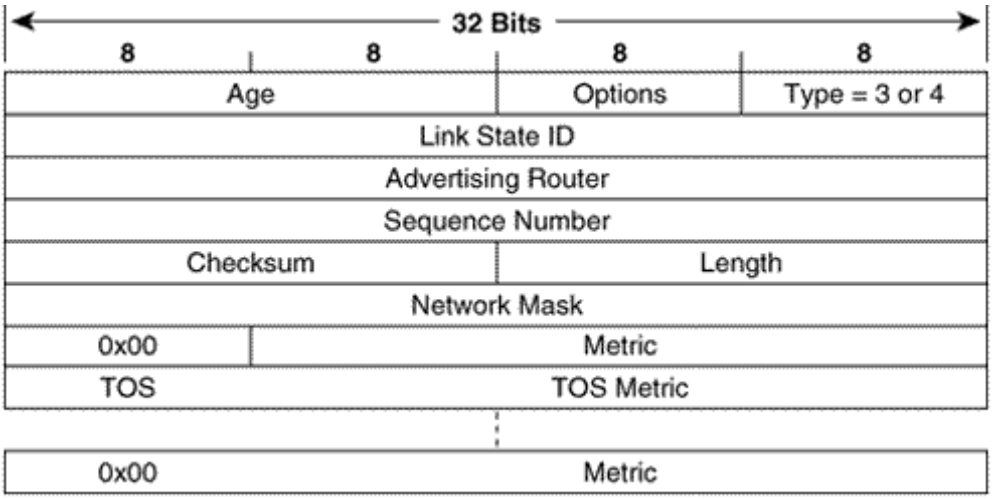
Number of TOS specifies the number of Type of Service metrics listed for this link. Although TOS is no longer supported in RFC 2328, the TOS fields are still included for backward compatibility with earlier OSPF implementations. If no TOS metrics are associated with a link, this field is set to 0x00.
Metric is the cost of the link (interface).
The next two fields are associated with a link corresponding to the number (#) of TOS field. For example, if # of TOS = 3, there will be three 32-bit words containing three instances of these fields. If # of TOS =0, there will be no instances of these fields. Note that Cisco supports only TOS = 0.

The OSPF Network LSA.



Link State ID for Network LSAs is the IP address of the DR's interface to the network.
Network Mask specifies the address or subnet mask used on this network.
Attached Router lists the Router IDs of all routers on the network that are fully adjacent with the DR, and the Router ID of the DR itself. The number of instances of this field (and hence the number of routers listed) can be deduced from the LSA header's Length field.

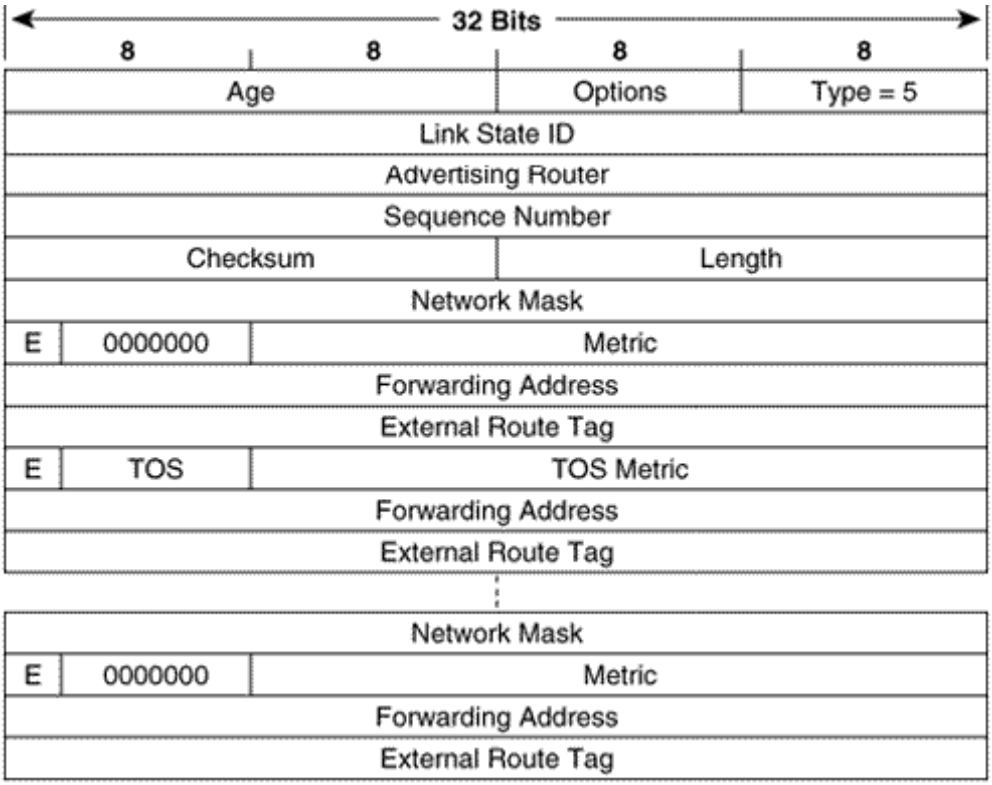
The OSPF Summary LSA. The format is the same for both type 3 and type 4 Summary LSAs.



Link State ID, for type 3 LSAs, is the IP address of the network or subnet being advertised. If the LSA is type 4, the Link State ID is the Router ID of the ASBR being advertised.
Network Mask is the address or subnet mask of the network being advertised in type 3 LSAs. In type 4 LSAs, this field has no meaning and is set to 0.0.0.0.If a type 3 LSA is advertising a default route, both the Link State ID and the Network Mask fields will be 0.0.0.0.
Metric is the cost of the route to this destination.
The TOS and TOS Metric fields are optional and are described in "The Router LSA." Again, Cisco supports only TOS = 0.

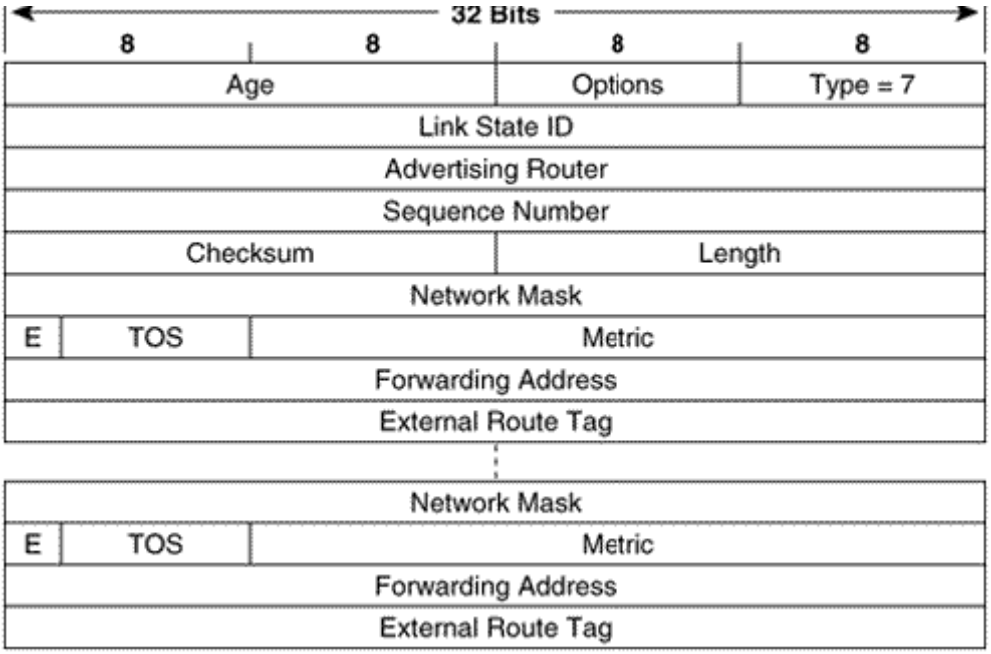
The OSPF Autonomous System External LSA.

 MSN : MingLang.Xu@Hotmail.Com



Link State ID for AS External LSAs is the IP address of the destination.
Network Mask is the address or subnet mask for the destination being advertised.
If the type 5 LSA is advertising a default route, the Link State ID and the network mask will both be 0.0.0.0.*E*, or *External Metric* bit, specifies the type of external metric to be used with this route. If the E-bit is set to 1, the metric type is E2. If the E-bit = 0, the metric type is E1. See the section "Path Types," earlier in this chapter, for more information on E1 and E2 external metrics.
Metric is the cost of the route, as set by the ASBR.
Forwarding Address is the address to which packets for the advertised destination should be forwarded. If the forwarding address is 0.0.0.0, packets will be forwarded to the originating ASBR.
External Route Tag is an arbitrary tag that may be applied to the external route. This field is not used by the OSPF protocol itself, but is instead provided for external route management.
Optionally, TOS fields may be associated with the destination. These fields are the same as discussed previously, except each TOS metric also has its own E-bit, Forwarding Address, and External Route Tag.

The OSPF NSSA LSA.



Forwarding Address, if the network between the NSSA ASBR and the adjacent autonomous system is advertised as an internal route, is the next hop address on the network. If the network is not advertised as an internal route, the forwarding address will be the NSSA ASBR's Router ID.

The OSPF Options field.

The Options field is present in every Hello and Database Description packet and in every LSA. The Options field allows routers to communicate their optional capabilities to other routers.

The asterisk, *, indicates an unused bit, normally set to zero.DC is set when the originating router is capable of supporting OSPF over demand circuits.

EA is set when the originating router is capable of receiving and forwarding External Attributes LSAs.These LSAs are not yet in general usage and are not covered in this book.

N is used only in Hello packets. A router will set N-bit = 1 to indicate support for NSSA External LSAs.If N-bit = 0, the router will not accept or send these LSAs. Neighboring routers with mismatched N-bits will not become adjacent; this restriction ensures that all routers in an area support NSSA capabilities equally. If the N-bit = 1, the E-bit must be 0.

P is used only in NSSA External LSA headers. (For this reason, the N- and P-bit can use the same position.) This bit tells the ABR of a not-so-stubby area to translate type 7 LSAs into type 5 LSAs.

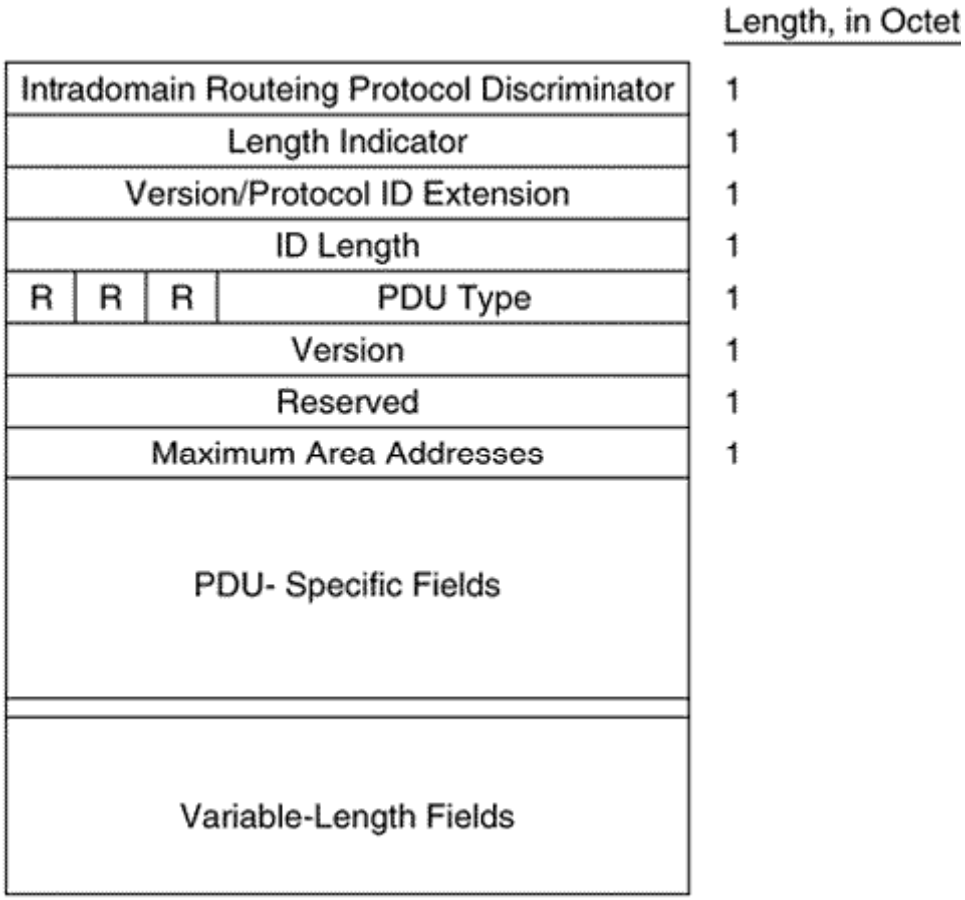
MC is set when the originating router is capable of forwarding IP multicast packets. This bit is used by

MOSPF.
E is set when the originating router is capable of accepting AS External LSAs. It will be set to 1 in all AS External LSAs and in all LSAs originated in the backbone and nonstub areas. E-bit = 0 in all LSAs originated within a stub area. Additionally, the bit is used in the Hello packet to indicate an interface's capability of sending and receiving type 5 LSAs. Neighboring routers with mismatched E-bits will not become adjacent; this restriction ensures that all routers in an area support stub capabilities equally.
T is set when the originating router is capable of supporting TOS.

 MSN : MingLang.Xu@Hotmail.Com

ISIS

The first eight octets of the IS-IS PDUs.



Intradomain Routeing Protocol Discriminator is a constant assigned by ISO 9577[16] to identify NPDUs.All IS-IS PDUs have a value of 0x83 in this field.

[16] International Organization for Standardization, "Protocol Identification in the Network Layer," IDO/IEC TR 9577, 1990.

Length Indicator specifies the length of the fixed header in octets.

Version/Protocol ID Extension is always set to one. *ID Length* describes the length of the System ID field of NSAP addresses and NETs used in this routing domain. This field is set to one of the following values:

An integer between 1 and 8 inclusive, indicating a System ID field of the same length in octets 0, indicating a System ID field of six octets 255, indicating a null System ID field (zero octets) The System ID of Cisco routers must be six octets, so the ID Length field of a Cisco-originated PDU will always be zero.

*PDU Type*is a five-bit field containing one of the PDU type numbers shown in Routing TCP/IP V1 Table 10.1. The preceding three bits (R) are reserved and are always set to zero.

Version is always set to one, just like the Version/Protocol ID Extension in the third octet.

Reserved is always set to all zeroes.

Maximum Area Addresses describes the number of area addresses permitted for this IS area. The number is set to one of the following values:

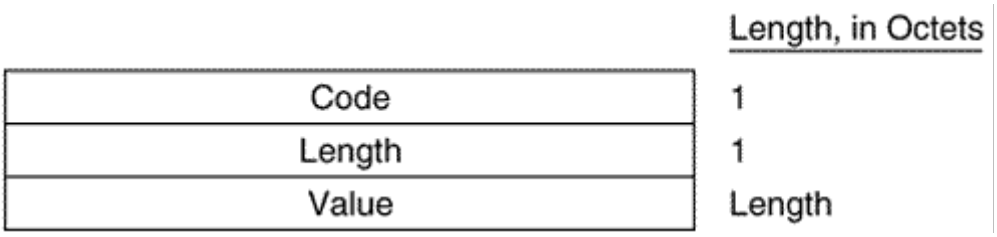
 MSN : MingLang.Xu@Hotmail.Com

An integer between 1 and 254 inclusive, indicating the number of areas allowed
0, indicating that the IS only supports a maximum of three addresses Cisco IOS supports a maximum of three areas, so the Maximum Area Addresses field in IS-IS PDUs originated by Cisco routers will always be zero.

IS-IS PDU types.

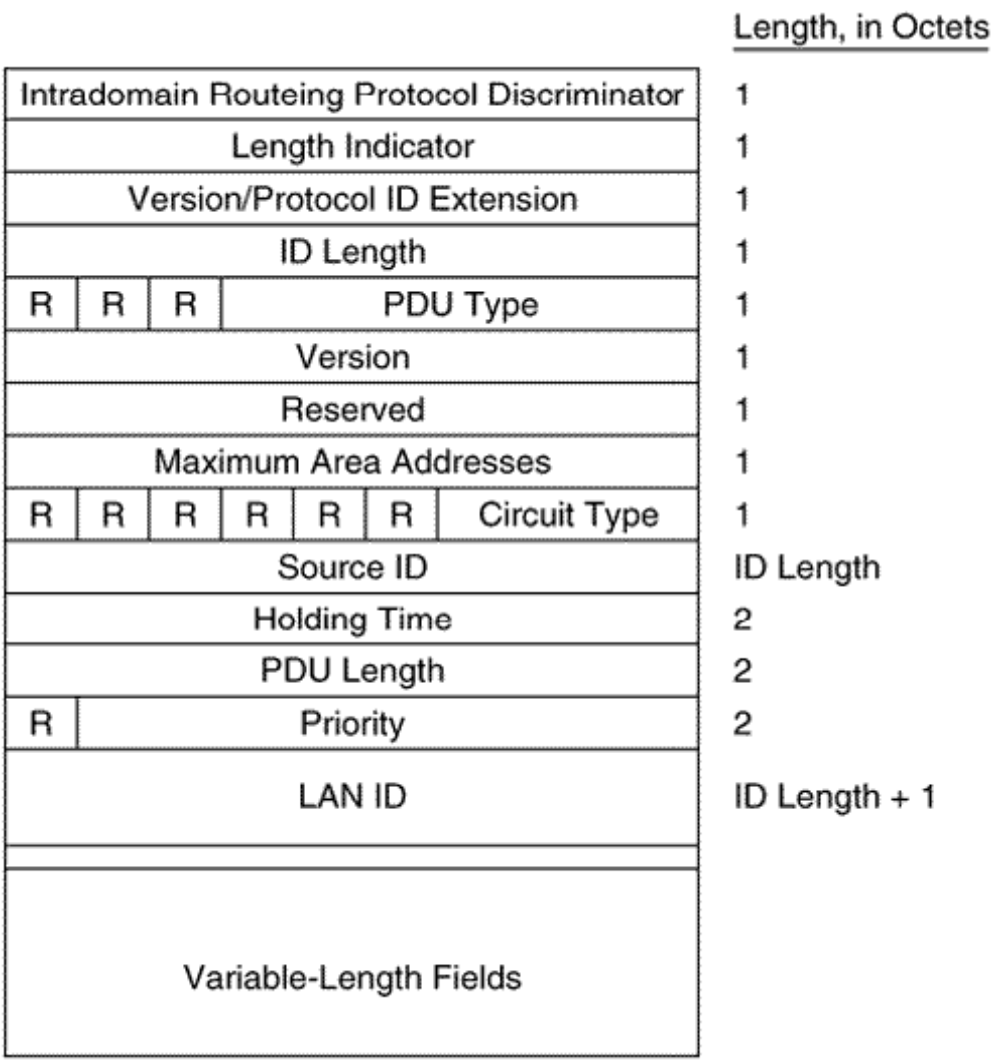
IS-IS PDU	Type Number
Hello PDUs	
Level 1 LAN IS-IS Hello PDU	15
Level 2 LAN IS-IS Hello PDU	16
Point-to-point IS-IS Hello PDU	17
Link State PDUs	
Level 1 LSP	18
Level 2 LSP	20
Sequence Numbers PDUs	
Level 1 CSNP	24
Level 2 CNSP	25
Level 1 PSNP	26
Level 2 PSNP	27

IS-IS Code/Length/Value triplets perform the same function for IS-IS as Type/Length/Valuetriplets perform for EIGRP.



CLV codes used with IS-IS.

The IS-IS LAN Hello PDU format.



*Circuit Type*is a two-bit field (the preceding six bits are reserved and are always zero) specifying whether the router is an L1 (01), L2 (10), or L1/L2 (11). If both bits are zero (00), the entire PDU is ignored.
Source ID is the System ID of the router that originated the Hello.
Holding Time is the period a neighbor should wait to hear the next Hello before declaring the originating router dead.
PDU Length is the length of the entire PDU in octets.
Priority is a seven-bit field used for the election of a DR. The field carries a value between 0 and 127 with the higher number having the higher priority. L1 DRs are elected by the priority in L1 LAN Hellos, and L2 DRs are elected by the priority in L2 LAN Hellos.
LAN ID is the System ID of the DR plus one more octet (the Pseudonode ID) to differentiate this LAN ID from another LAN ID that might have the same DR.
The following CLVs can be used by an IS-IS LAN Hello:[18]

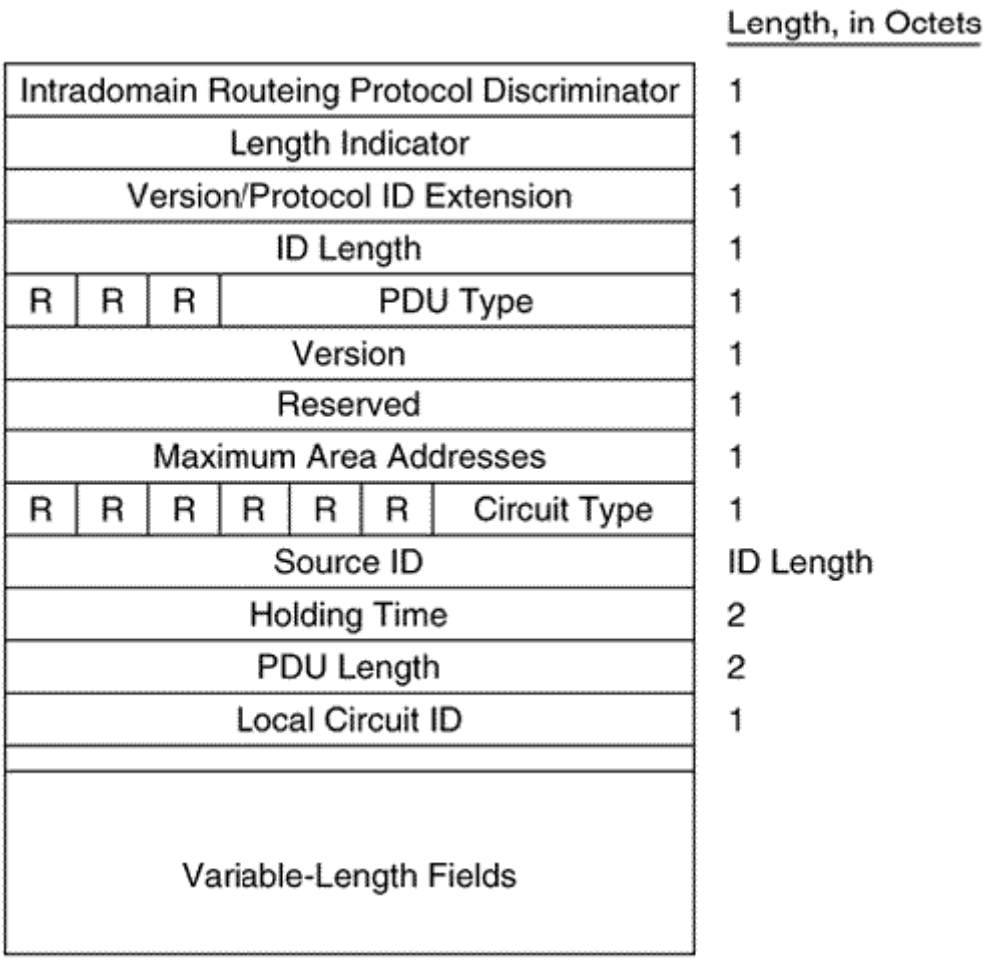
[18] As a reminder, RFC 1195 also specifies an Authentication Information CLV with a type number of 133. Cisco uses the ISO-specified type

number of 10 to identify its Authentication Information CLVs.

Area Addresses (type 1)

Intermediate System Neighbors (type 6)
Padding (type 8)
Authentication Information (type 10)
Protocols Supported (type 129)
IP Interface Address (type 132)

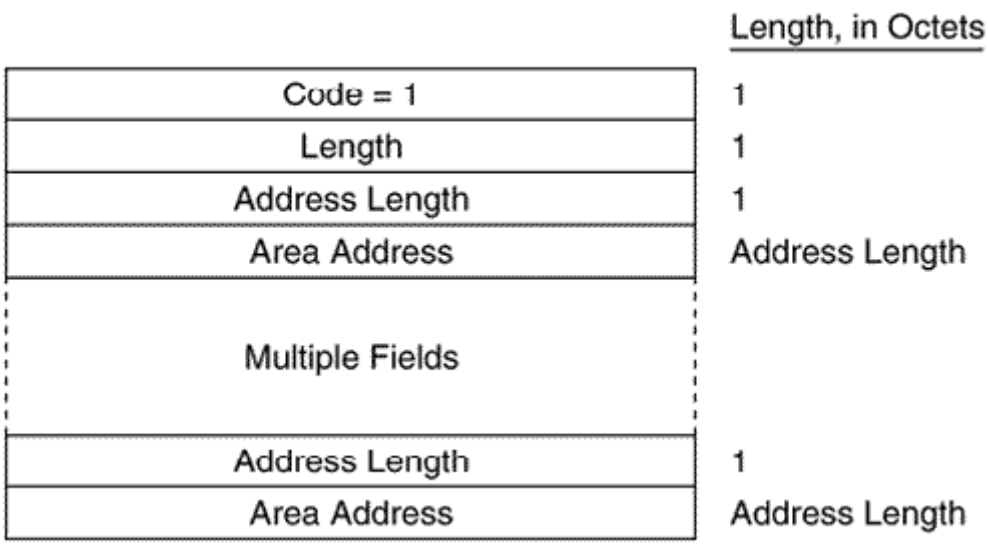
The IS-IS point-to-point Hello PDU.



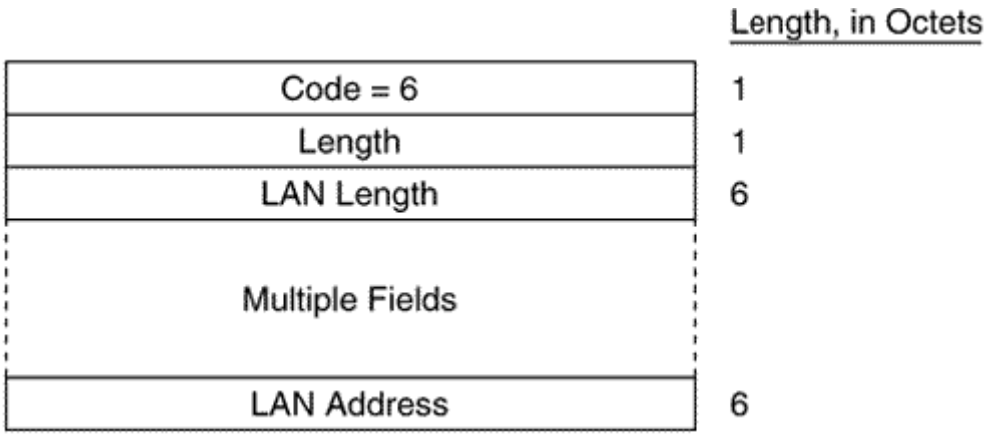
Local Circuit ID is a one-octet ID assigned to this circuit by the router originating the Hello and is unique among the router's interfaces. The Local Circuit ID in the Hellos at the other end of the point-to-point link may or may not contain the same value. The IS-IS point-to-point Hello does not use the IS Neighbors CLV. With that exception, the same CLVs are used as the LAN Hello.

The Area Addresses CLV.

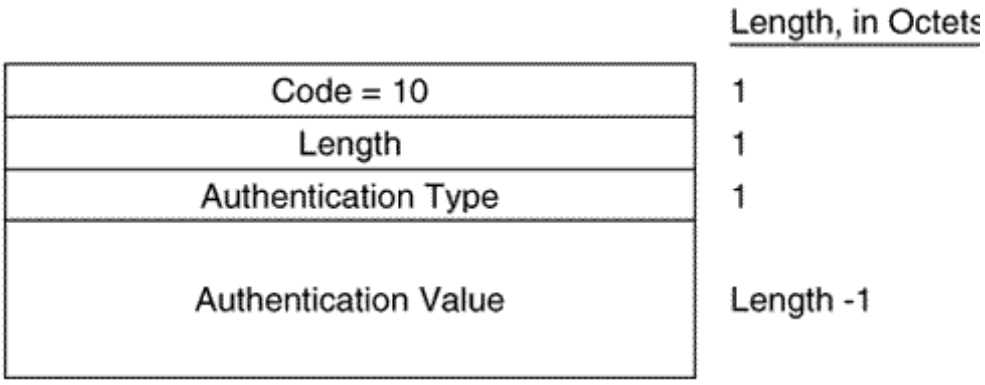
 MSN : MingLang.Xu@Hotmail.Com



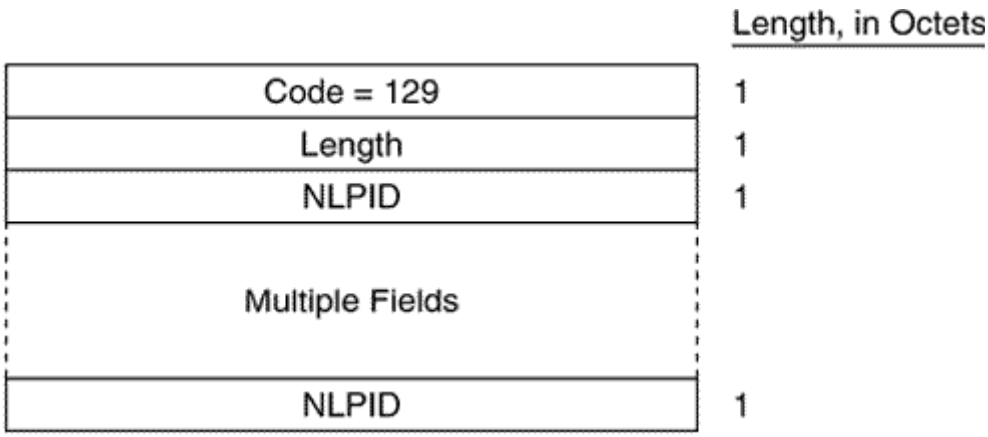
The IS Neighbors CLV for Hello PDUs.



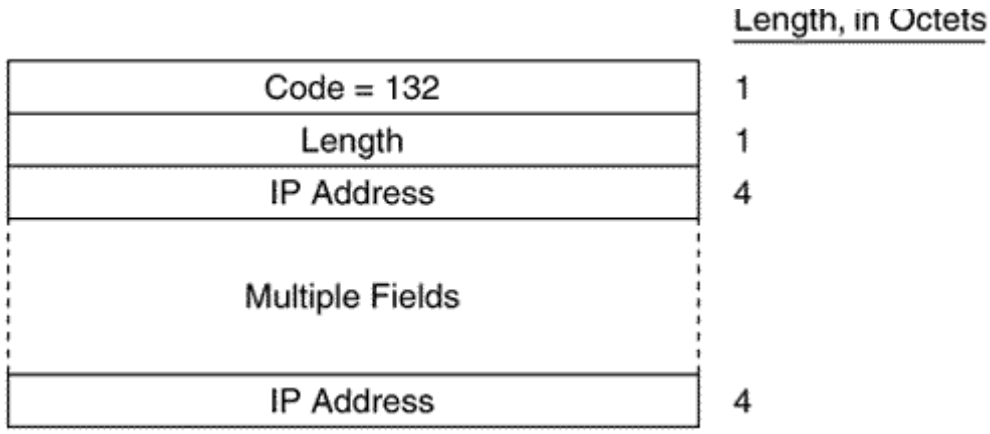
The Authentication Information CLV.



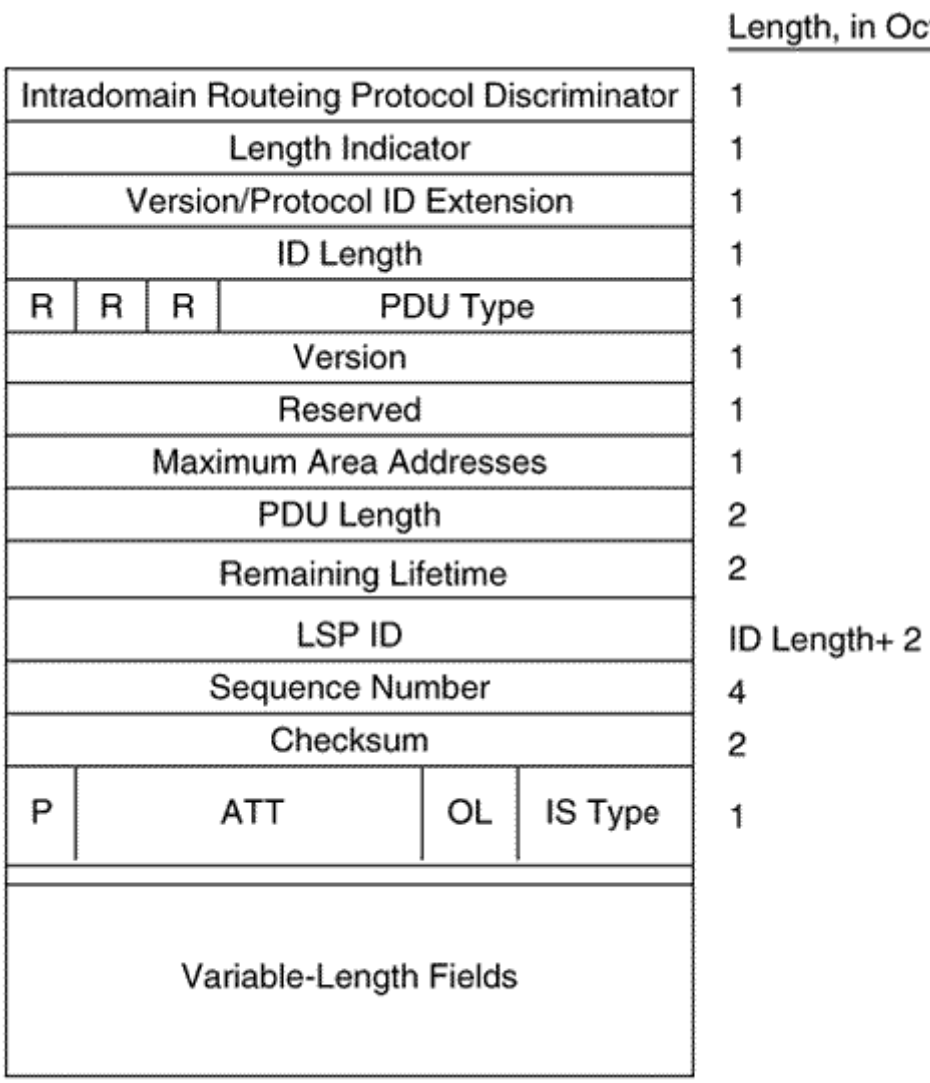
The Protocols Supported CLV.



The IP Interface Address CLV.



The IS-IS LSP format.



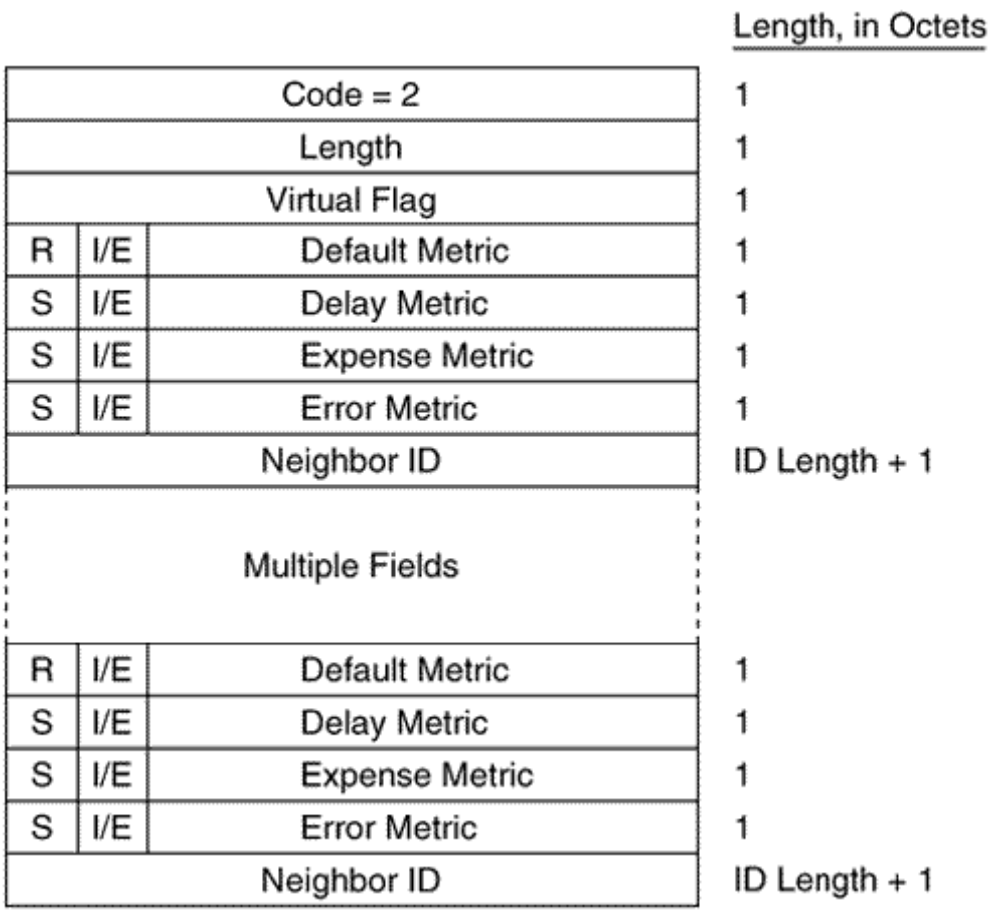
PDU Length is the length of the entire PDU in octets.
Remaining Lifetime is the number of seconds before the LSP is considered to be expired.
LSP ID is the System ID, the Pseudonode ID, and the LSP Number of the LSP. The LSP ID is described in more detail in "The Update Process."
Sequence Number is a 32-bit unsigned integer.
Checksum is the checksum of the contents of the LSP.
P is the Partition Repair bit. Although the bit exists in both L1 and L2 LSPs, it is relevant only in L2 LSPs. When this bit is set, it indicates that the originating router supports the automatic repair of area partitions. Cisco IOS does not support this function, so the P bit of LSPs originated by Cisco routers will always be zero.
ATT is a four-bit field indicating that the originating router is attached to one or more other areas. Although the bits exist in both L1 and L2 LSPs, they are relevant only in L1 LSPs that have been originated by L1/L2 routers. The four bits indicate which metrics are supported by the attachment. Reading from left to right, the bits indicate:
Bit 7: The Error metric
Bit 6: The Expense metric
Bit 5: The Delay metric

 MSN : MingLang.Xu@Hotmail.Com

Bit 4: The Default metric Cisco IOS supports only the default metric, so bits 5 through 7 will always be zero.
OL is the Link State Database Overload bit. Under normal circumstances, this bit will be zero. If the originating router is experiencing a memory overload condition, it will set the OL bit to one. Routers receiving an LSP with the OL bit set will not use the originating router as a transit router, although they will still route to destinations on the originator's directly connected links.
IS Type is a two-bit field indicating whether the originating router is an L1 or an L2:
00 = Unused value
01 = L1
10 = Unused value
11 = L2
An L1/L2 router sets the bits according to whether the LSP is an L1 or an L2 LSP.
The following CLVs can be used by an L1 LSP:
Area Addresses (type 1)
IS Neighbors (type 2)
ES Neighbors (type 3);
Authentication Information (type 10)
IP Internal Reachability Information (type 128)
Protocols Supported (type 129)
IP Interface Address (type 132)
The following CLVs can be used by an L2 LSP:
Area Addresses (type 1)
IS Neighbors (type 2)
Partition Designated Level 2 IS (type 4)
Prefix Neighbors (type 5)
Authentication Information (type 10)
IP Internal Reachability Information (type 128)
IP External Reachability Information (type 130)
Inter-Domain Routing Protocol Information (type 131)
Protocols Supported (type 129)
IP Interface Address (type 132)

The Intermediate System Neighbors CLV for LSPs.

 MSN : MingLang.Xu@Hotmail.Com



Virtual Flag, although eight bits long, has a value of either 0x01 or 0x00. A 0x01 in this field indicates that the link is a level 2 virtual link to repair an area partition. The field is relevant only to L2 routers that support area partition repair; Cisco does not, so the field will always be 0x00 in Cisco-originated LSPs.

R is a reserved bit and is always zero.

I/E, associated with each of the metrics, indicates whether the associated metric is internal or external.

The bit has no meaning in IS Neighbors CLVs because all neighbors are by definition internal to the IS-IS domain. Therefore, this bit is always zero in IS Neighbors CLVs.

Default Metric is the six-bit default metric for the originating router's link to the listed neighbor and contains a value between 0 and 63.

S, associated with each of the optional metrics, indicates whether the metric is supported (zero) or unsupported (one). Cisco does not support any of the three optional metrics, so the bit is always set to one and the associated six-bit metric fields are all zeroes.

Neighbor ID is the System ID of the neighbor, plus one more octet. If the neighbor is a router, the last octet is 0x00. If the neighbor is a pseudonode, the System ID is that of the DR and the last octet is the Pseudonode ID.

 MSN : MingLang.Xu@Hotmail.Com

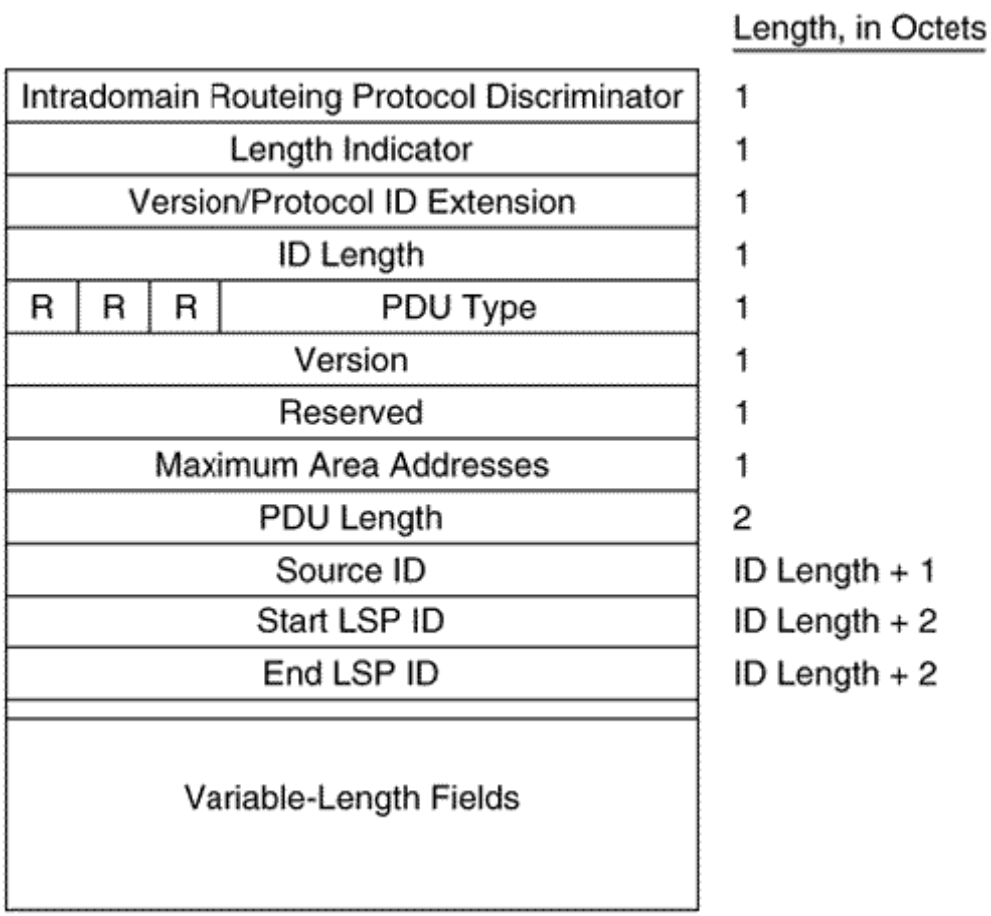
The IP Internal Reachability Information CLV.

			Length, in Octets
Code = 128			1
Length			1
R	I/E	Default Metric	1
S	R	Delay Metric	1
S	R	Expense Metric	1
S	R	Error Metric	1
IP Address			4
Subnet Mask			4
Multiple Fields			
R	I/E	Default Metric	1
S	R	Delay Metric	1
S	R	Expense Metric	1
S	R	Error Metric	1
IP Address			4
Subnet Mask			4

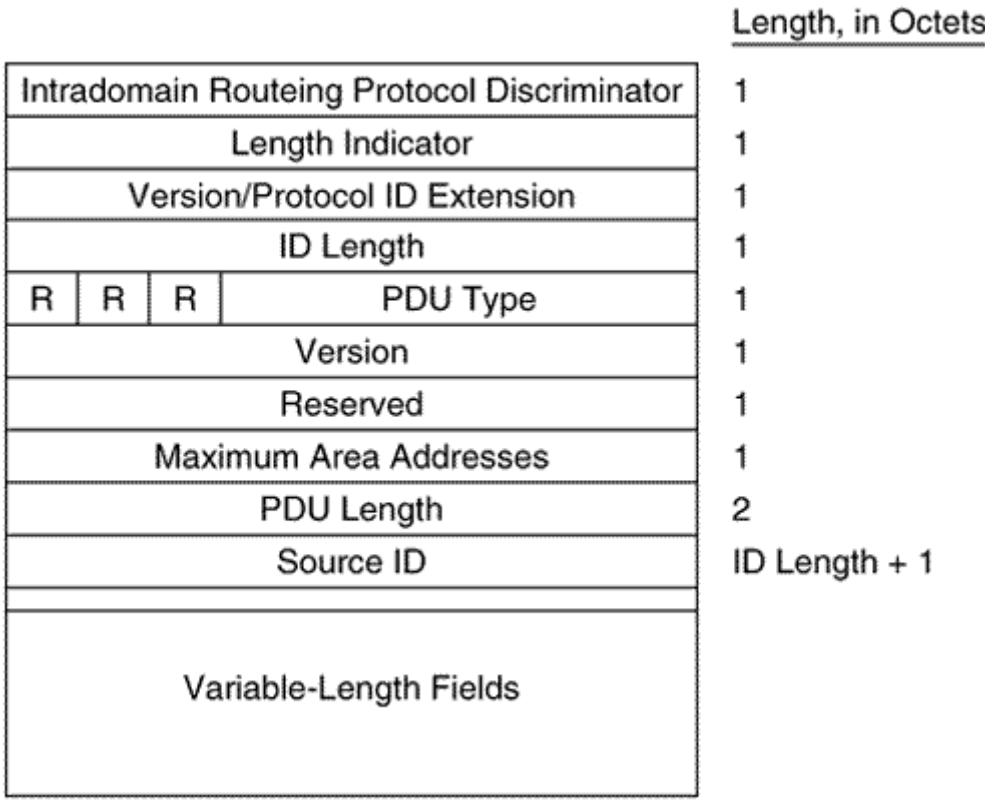
The Inter-Domain Routing Protocol Information CLV.

	Length, in Octets
Code = 131	1
Length	1
Inter-Domain Information Type	1
External Information	Variable

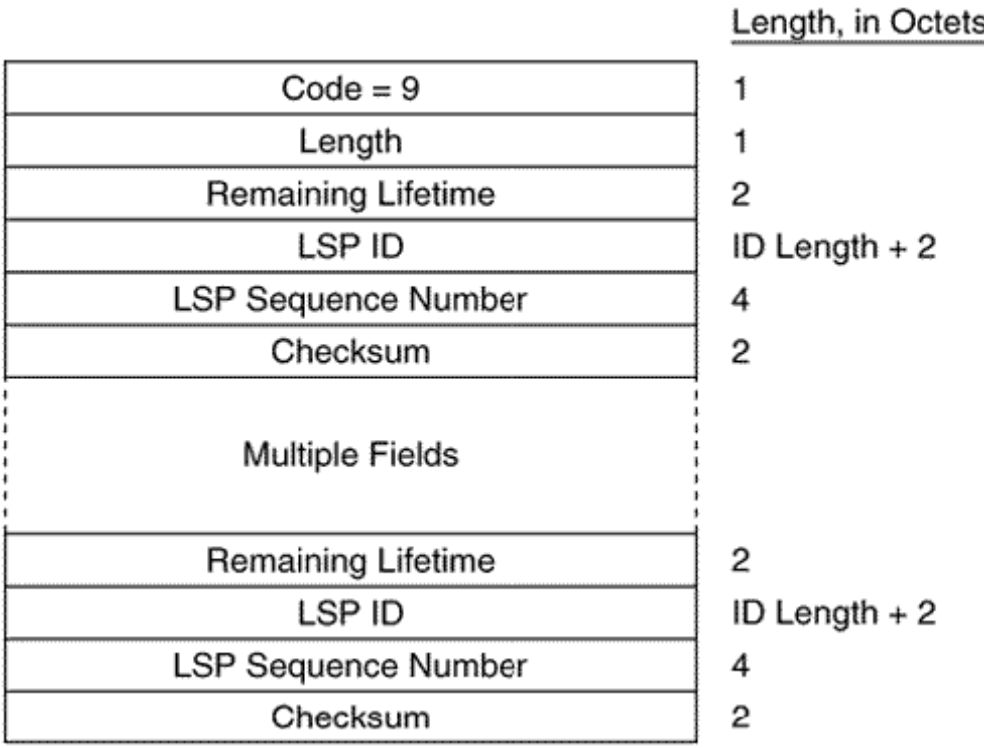
The IS-IS CSNP format.



The IS-IS PSNP format.

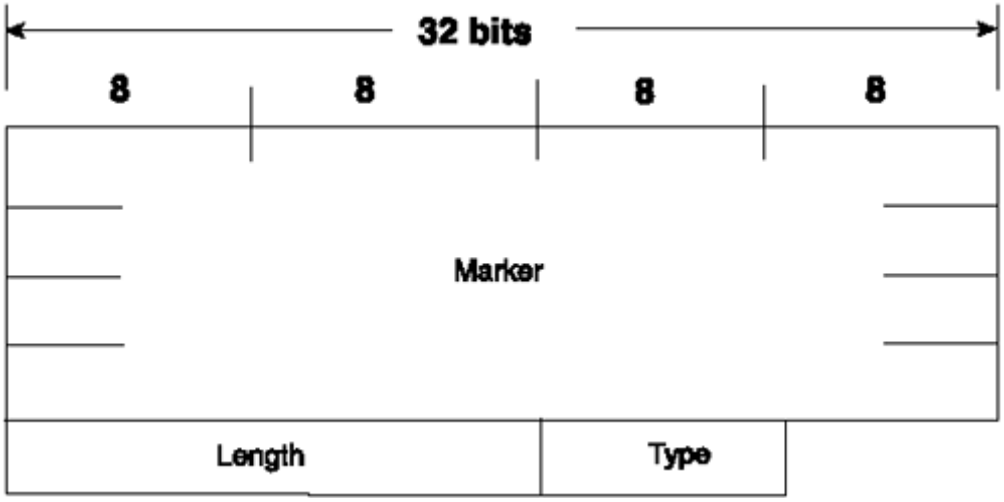


The LSP Entries CLV.



BGP 4

The BGP Message Header



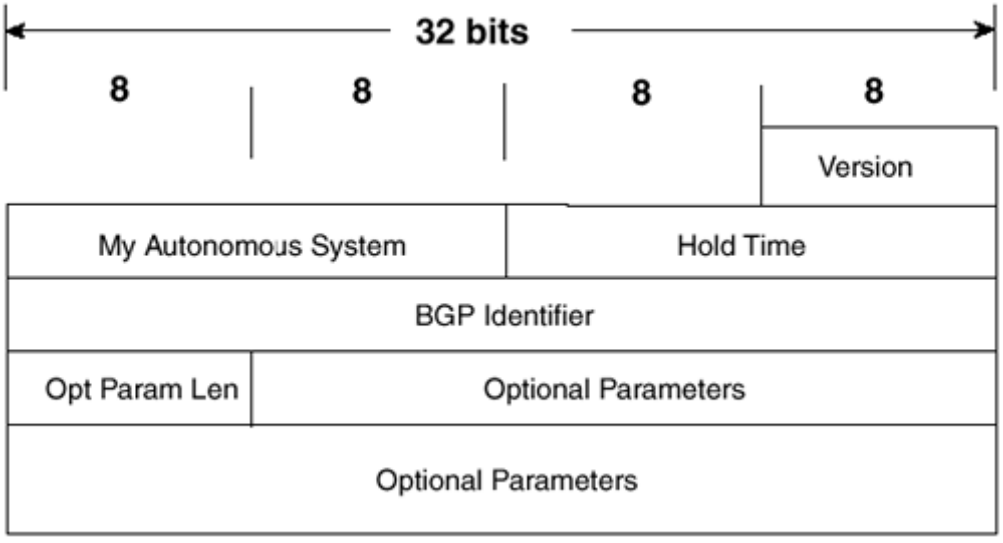
Marker is a 16-octet field that is used to detect loss of synchronization between BGP peers and to authenticate messages when authentication is supported. If the message type is Open or if the Open message contains no authentication information, the Marker field is set to all 1s. Otherwise, the value of the marker can be predicted by some computation as part of the authentication process. *Length* is a 0-octet field that indicates the total length of the message, including the header, in octets.

BGP Type Codes

Code Type

1 Open
2 Update
3 Notification
4 Keepalive

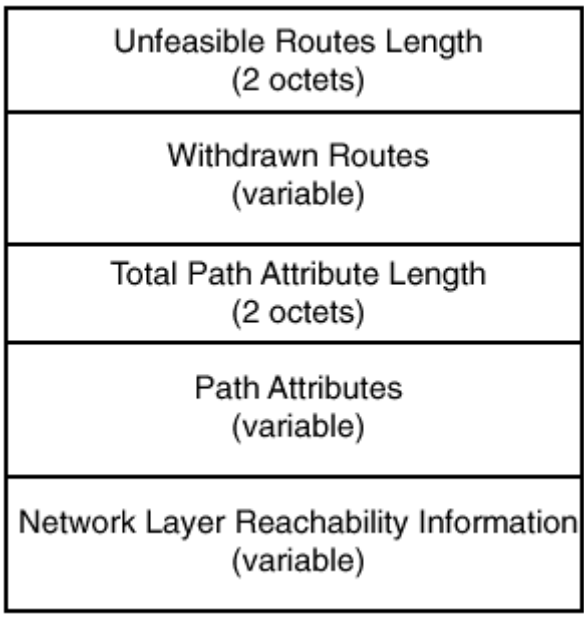
The BGP Open Message Format



The BGP Open message contains the following fields:

- **Version—** A 1-octet field specifying the BGP version running on the originator.
- **My Autonomous System—** A 2-octet field specifying the AS number of the originator.
- **Hold Time—** A 2-octet number indicating the number of seconds the sender proposes for the hold time. A receiver compares the value of the Hold Time field and the value of its configured hold time and accepts the smaller value or rejects the connection. The hold time must be either 0 or at least 3 seconds.
- **BGP Identifier—** The router ID of the originator. A Cisco router sets its router ID as either the highest IP address of any of its loopback interfaces or, if no loopback interface is configured, the highest IP address of any of its physical interfaces.
- **Optional Parameters Length—** A 1-octet field indicating the total length of the following Optional Parameters field, in octets. If the value of this field is zero, no Optional Parameters field in included in the message.
- **Optional Parameters—** A variable-length field containing a list of optional parameters. Each parameter is specified by a 1-octet type field, a 1-octet length field, and a variable-length field containing the parameter value.

The BGP Update Message Format



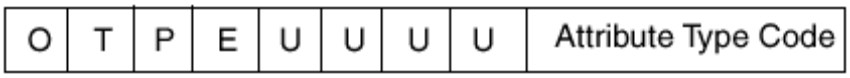
The BGP Update message contains the following fields:

- **Unfeasible Routes Length—** A 2-octet field indicating the total length of the following

Withdrawn Routes field, in octets. A value of zero indicates that no routes are being withdrawn and that no Withdrawn Routes field is included in the message.

- **Withdrawn Routes—** A variable-length field containing a list of routes to be withdrawn from service. Each route in the list is described with a (Length, Prefix) tuple in which the Length is the length of the prefix and the Prefix is the IP address prefix of the withdrawn route. If the Length part of the tuple is zero, the Prefix matches all routes.
- **Total Path Attribute Length—** A 2-octet field indicating the total length of the following Path Attribute field, in octets. A value of zero indicates that attributes and NLRI are not included in this message.
- **Path Attributes—** A variable-length field listing the attributes associated with the NLRI in the following field. Each path attribute is a variable-length triple of (Attribute Type, Attribute Length, Attribute Value). The Attribute Type part of the triple is a 2-octet field consisting of four flag bits, four unused bits, and an Attribute Type code

The Attribute Type Part of the Path Attributes Field



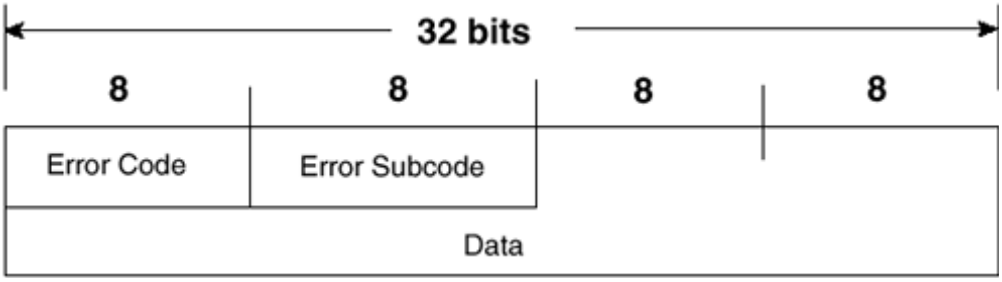
Flag bits
O: Optional bit
0 = Optional
1 = Well-Known
T: Transitive bit
0 = Transitive
1 = Non-Transitive
P: Partial bit
0 = Optional Transitive attribute is partial
1 = Optional Transitive attribute is complete
E: Extended length bit
0 = Attribute Length is one octet
1 = Attribute Length is two octets
U: Unused

Attribute Types and Associated Attribute Values[*]

Attribute Type Code	Attribute Type	Attribute Value Code	Attribute Value
1	ORIGIN	0	IGP
		1	EGP
		2	Incomplete
2	AS_PATH	1	AS_SET
		2	AS_SEQUENCE
		3	AS_CONFED_SET
		4	AS_CONFED_SEQUENCE
3	NEXT_HOP	0	Next-hop IP address

4 MULTI_EXIT_DISC 0 4-octet MED
5 LOCAL_PREF 0 4-octet LOCAL_PREF
6 ATOMIC_AGGREGATE 0 None
7 AGGREGATOR 0 AS number and IP address of aggregator
8 COMMUNITY 0 4-octet community identifier
9 ORIGINATOR_ID 0 4-octet router ID of originator
10 CLUSTER_LIST 0 Variable-length

The BGP Notification Message Format



The BGP Notification message contains the following fields:

- **Error Code—** A 1-octet field indicating the type of error.
- **Error Subcode—** A 1-octet field providing more-specific information about the error. Table 2-8 shows the possible error codes and associated error subcodes.
- **Data—** A variable-length field used to diagnose the reason for the error. The contents of the Data field depend on the error code and subcode.

BGP Notification Message Error Codes and Error Subcodes
Error Code Error Error Subcode Subcode Detail

1 Message Header
Error
1 Connection not synchronized
2 Bad message length
3 Bad message type
2 Open Message
Error
1 Unsupported version number
2 Bad peer AS
3 Bad BGP identifier
4 Unsupported optional parameter
5 Authentication failure
6 Unacceptable hold time
3 Update Message
Error
1 Malformed attribute list
2 Unrecognized well-known attribute

3 Missing well-known attribute
4 Attribute flags error
5 Attribute length error
6 Invalid ORIGIN attribute
7 AS routing loop
8 Invalid NEXT_HOP attribute
9 Optional attribute error
10 Invalid network field
11 Malformed AS_PATH
4 Hold Timer Expired 0 —
5 Finite State
Machine Error 0 —
6 Cease 0 —

MPLS 多协议标签交换

20	23	24	32 bit
Label	Exp	S	TTL

- Label — Label 值传送标签实际值。当接收到一个标签数据包时，可以查出栈顶部的标签值，并且系统知道：A、数据包将被转发的下一跳；B、在转发之前标签栈上可能执行的操作，如返回到标签进栈顶入口同时将一个标签压出栈；或返回到标签进栈顶入口然后将一个或多个标签推进栈。
- Exp — 试用。预留以备试用。
- S — 栈底。标签栈中最后进入的标签位置，该值为 0，提供所有其它标签入栈。
- TTL — 生存期字段（Time to Live），用来对生存期值进行编码。

MPLS 结构协议组包括：

- MPLS：相关信令协议，如 OSPF、BGP、ATM PNNI 等。
- LDP：标签分发协议（Label Distribution Protocol）
- CR-LDP：基于路由受限标签分发协议（Constraint-Based LDP）

RSVP-TE：基于流量工程扩展的资源预留协议（resource Reservation Protocol - Traffic

- Label — Label 值传送标签实际值。当接收到一个标签数据包时，可以查出栈顶部的标签值，并且系统知道：A、数据包将被转发的下一跳；B、在转发之前标签栈上可能执行的操作，如返回到标签进栈顶入口同时将一个标签压出栈；或返回到标签进栈顶入口然后将一个或多个标签推进栈。
- Exp — 试用。预留以备试用。
- S — 栈底。标签栈中最后进入的标签位置，该值为 0，提供所有其它标签入栈。
- TTL — 生存期字段（Time to Live），用来对生存期值进行编码。

MPLS 结构协议组包括：

- MPLS：相关信令协议，如 OSPF、BGP、ATM PNNI 等。
- LDP：标签分发协议（Label Distribution Protocol）
- CR-LDP：基于路由受限标签分发协议（Constraint-Based LDP）
- RSVP-TE：基于流量工程扩展的资源预留协议（resource Reservation Protocol - Traffic Engineering）

下面图形描述了 MPLS 协议栈结构：

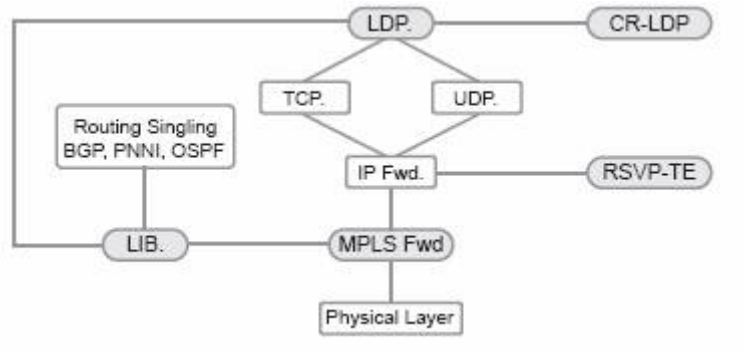


图 2 — 4 MPLS 协议栈结构

每种结构将在各独立文档中具体讲解…… 待续^_^

LAN

以太网：IEEE 802.3 局域网协议

以太网协议是由一组 IEEE 802.3 标准定义的局域网协议集。在以太网标准中，有两种操作模式：半双工和全双工。半双工模式中，数据是通过在共享介质上采用载波监听多路访问/冲突检测（CSMA/CD）协议实现传输的。它的主要缺点在于有效性和距离限制，链路距离受最小 MAC 帧大小的限制。该限制极大的降低了其高速传输的有效性。因此，引入了载波扩展技术来确保千兆位以太网中 MAC 帧的最小长度为 512 字节，从而达到了合理的链路距离要求。

当前定义在光纤和双绞线上的传输速率有四种：

- 10 Mbps － 10Base-T 以太网
- 100 Mbps － 快速以太网
- 1000 Mbps － 千兆位以太网（802.3z）
- 10 千兆位以太网 － IEEE 802.3ae

本文我们主要讨论以太网的总体概况。有关快速以太网、千兆位以太网以及 10 千兆位以太网的具体内容将在其它文档中另作介绍。

以太网系统由三个基本单元组成：1. 物理介质，用于传输计算机之间的以太网信号；2. 介质访问控制规则，嵌入在每个以太网接口处，从而使得计算机可以公平的使用共享以太网信道；3. 以太帧，由一组标准比特位构成，用于传输数据。

在所有 IEEE 802 协议中，ISO 数据链路层被划分为两个 IEEE 802 子层，介质访问控制（MAC）子层和 MAC － 客户端子层。IEEE 802.3 物理层对应于 ISO 物理层。

MAC 子层有两个基本职能：

- 数据封装，包括传输之前的帧组合和接收中、接收后的帧解析 / 差错检测。
- 介质访问控制，包括帧传输初始化和传输失败恢复。

介质访问控制（MAC）－ 客户端子层可能是以下一种：

- 逻辑链路控制（LLC），提供终端协议栈的以太网 MAC 和上层之间的接口，其中 LLC 由 IEEE 802.2 标准定义。
- 网桥实体，提供 LANs 之间的 LAN-to-LAN 接口，可以使用同种协议（如以太网到以太网）和不同的协议（如以太网到令牌环）之间。网桥实体由 IEEE 802.1 标准定义。

以太网上的每台计算机都能独立运行，不存在中心控制器。连接到以太网的所有工作站都接入共享信令系统，又称为介质。要发送数据时，工作站首先监听信道，如果信道空闲，即可以以太帧或数据包格式传输数据。

每帧传输完毕之后，各工作站必须公平争取下一帧的传输机会。对于共享信道的访问取决于嵌入到

每个工作站的以太网接口的介质访问控制机制。该机制建立在载波监听多路访问/冲突检测（CSMA/CD）基础上。

当以太帧发送到共享信道后，所有以太网接口查看它的目标地址。如果帧目标地址与接口地址相匹配，那么该帧就能被全部读取并且被发送到那台计算机的网络软件上。如果发现帧目标地址与它们本身的地址不匹配时，则停止帧读取操作。

信号如何通过组成以太网系统的各个介质段有助于我们掌握系统拓朴结构。以太网的信号拓朴是一种逻辑拓朴，用来区别介质电缆的实际物理布局。以太网的逻辑拓朴结构提供了一条单一信道（或总线）用于传送以太网信号到所有工作站。

多个以太网段可以链接在一起构成一个较大的以太网，这通过一种能够放大信号和重新计时的叫做中继器的设备实现。通过中继器，多段以太网系统可以像“无根分支树”（non-rooted branching tree）一样扩展。“无根”意味着系统在任意方向上都可以生成链接段，且没有特定的根段。最重要的是，各段的连接不能形成环路。系统的每个段必须具有两个终端，这是由于以太网系统在环路路径上不能正确运行。

即使介质段以星形模式物理连接，且许多段都接在中继器上，但是它的逻辑拓朴结构仍就是通过以太网单信道传送信号至所有工作站。

协议结构

10/100 Mbps 以太网中的基本 IEEE 802.3 MAC 数据格式如下：

7	1	6	6	2	46-1500 bytes	4 bytes
Pre	SFD	DA	SA	Length Type	Data unit + pad	FCS

- Preamble（Pre） — 7 字节。Pre 字段中 1 和 0 交互使用，接收站通过该字段知道导入帧，并且该字段提供了同步化接收物理层帧接收部分和导入比特流的方法。
- Start-of-Frame Delimiter（SFD） — 1 字节。字段中 1 和 0 交互使用，结尾是两个连续的 1，表示下一位是利用目的地址的重复使用字节的重复使用位。
- Destination Address（DA） — 6 字节。DA 字段用于识别需要接收帧的站。
- Source Addresses（SA） — 6 字节。SA 字段用于识别发送帧的站。
- Length/Type — 2 字节。如果是采用可选格式组成帧结构时，该字段既表示包含在帧数据字段中的 MAC 客户机数据大小，也表示帧类型 ID。
- Data — 是一组 n（46=< n =<1500）字节的任意值序列。帧总值最小为 64 字节。
- Frame Check Sequence（FCS） — 4 字节。该序列包括 32 位的循环冗余校验（CRC）值，由发送 MAC 方生成，通过接收 MAC 方进行计算得出以校验被破坏的帧。

包含千兆位载波扩展的 MAC 帧：

1000 Base-X 最小帧大小为 416 字节；1000 Base-T 最小帧大小为 520 字节。通过扩展字段可以满

足长度小于最小值的帧需求。

7	1	6	6	2	46=< n =<1500	4 bytes	Variable
Pre	SFD	DA	SA	Length Type	Data unit + pad	FCS	Ext

 MSN : MingLang.Xu@Hotmail.Com

关于以太网 802.3 中，快速以太网帧最小值为 64 字节，最大为 1518 字节

快速以太网（Fast Ethernet）的速度在原 10Base-T 以太网的基础上提高了 10 倍，并保留了其帧格式、MAC 机制和 MTU。同时现有的 10Base-T 的应用程序和网络管理工具同样适用于快速以太网。100Base-T 标准正式定义在 IEEE 802.3u 中。

和以太网一样，100Base-T 仍是基于载波监听多路访问和冲突检测（CSMA/CD）技术。100Mbps 快速以太网标准几种不同的布线方式：

- 100BASE － TX：两对高质量双绞线
- 100BASE － T4：四对普通双绞线
- 100BASE － FX：光缆

快速以太网规范包括传输速度的自动协商机制，这使得供应商提供的双速以太网接口能够自动安装以及运行在 10 Mbps 或 100 Mbps 的速度下。

IEEE 标识符包括三块信息。其一，“100”表示传输速度 100M；其二，“Base”表示“基带”，信令的一种。基带信令可以简单理解为以太网信号是唯一的介质传输信号。

其三是关于网络段类型的阐述。“T4”属于双绞线网段类型，具有四对电话线路级双绞线；“TX”属于双绞线网络段类型，具有两对双绞线，基于由 ANSI （美国国家标准化组织）制定的数据级双绞线物理介质标准；“FX”属于光纤链路网络段类型，基于由 ANSI 制定的光纤双绞线物理介质标准，并使用光纤中的两股。TX 和 FX 介质标准统称为 100BASE － X。

快速以太网中的 100BASE － TX 和 100BASE － FX 介质标准均来自 ANSI 制定的物理介质标准。ANSI 物理介质标准最初是为光纤分布式数据接口（FDDI） LAN 标准（ANSI 标准 X3T9.5）开发的，并且现在广泛应用于 FDDI 局域网中。

协议结构

关于以太网 802.3 中，快速以太网帧最小值为 64 字节，最大为 1518 字节。

7	1	6	6	2	46=< n =<1500	4 bytes
Pre	SFD	DA	SA	Length Type	Data unit + pad	FCS

- Preamble（Pre） － 7 字节。Pre 字段中 1 和 0 交互使用，接收站通过该字段知道导入帧，并且该字段提供了同步化接收物理层帧接收部分和导入比特流的方法。
- Start-of-Frame Delimiter（SFD） － 1 字节。字段中 1 和 0 交互使用，结尾是两个连续的 1，表示下一位是利用目的地址的重复使用字节的重复使用位。
- Destination Address（DA） － 6 字节。DA 字段用于识别需要接收帧的站。
- Source Addresses（SA） － 6 字节。SA 字段用于识别发送帧的站。
- Length/Type － 2 字节。如果是采用可选格式组成帧结构时，该字段既表示包含在帧数据字段中的 MAC 客户机数据大小，也表示帧类型 ID。

 MSN : MingLang.Xu@Hotmail.Com

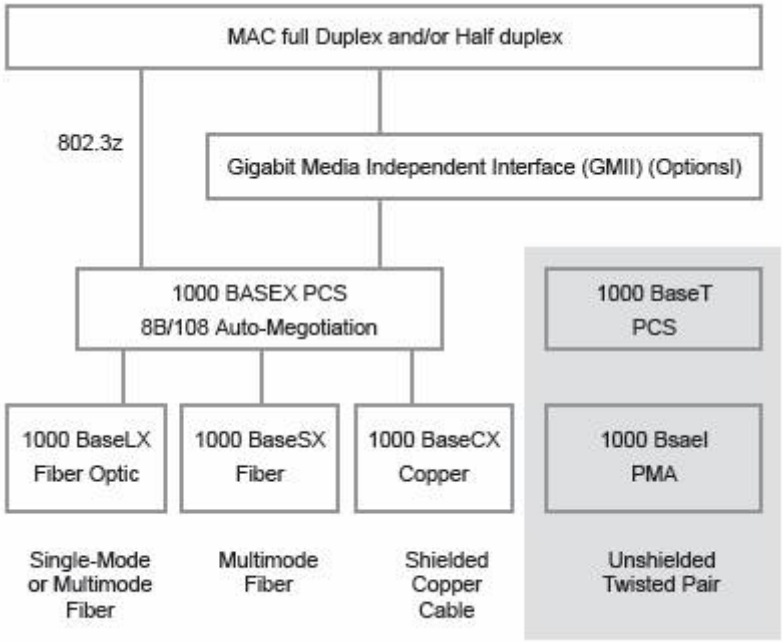
- Data — 是一组 n（46=< n =<1500）字节的任意值序列。帧总值最小为 64 字节。
- Frame Check Sequence（FCS） — 4 字节。该序列包括 32 位的循环冗余校验（CRC）值，由发送 MAC 方生成，通过接收 MAC 方进行计算得出以校验被破坏的帧。

Gigabit Ethernet：千兆位以太网（1000 Mbps）

以太网协议是由一组 IEEE 802.3 标准定义的局域网协议集。千兆位以太网协议基于以太网协议，速度在快速以太网的基础上提高了 10 倍，由于其数据帧较短，所以在其后增加载波扩展。千兆位以太网以 IEEE 802.3z 和 802.3ab 发布，作为 IEEE 802.3 标准的补充。

载波扩展是一种简单的解决方案，但它浪费较多的带宽。包突发机制即指“载波扩展加上数据包突发”。该突发特点是，允许 MAC 发送短帧序列（最大长度大约为 5.4 的帧），而不放弃介质控制。

千兆位以太网标准完全与以太网和快速以太网相兼容。千兆技术仍是基于载波监听多路访问和冲突检测（CSMA/CD）模式，它支持全双工/半双工工作方式、单模/多模光纤、短程同轴电缆以及双绞线等连接方案。千兆位以太网结构如下：



千兆位以太网结构

IEEE 802.3z 中定义了光纤和电缆连接方式的千兆位以太网，以及物理介质标准 1000Base－X（ 1000BaseSX －短波 － 覆盖 500 m 和 1000BaseLX －长波 － 覆盖 5 km）。IEEE 802.3ab 中定义了无屏蔽双绞线千兆位以太网（1000Base-T，覆盖 75m）。

通过千兆位接口转换器（GBIC），网络管理员能够在基于 port-by-port 的基础上配置千兆位端口以适用于短波（SX）、长波（LX）、远程（LH）和铜物理接口（CX）。LH GBIC 将单模光纤距离从 5 km 扩展到 10 km。

协议结构

1000Base-X 最小帧大小为 416 字节；1000Base-T 最小帧大小为 520 字节。扩展字段用于填充足长度小于最小值的帧 。

7	1	6	6	2	46=< n =<1500	4 bytes	Variable
Pre	SFD	DA	SA	Length Type	Data unit + pad	FCS	Ext

- Preamble（Pre） － 7 字节。Pre 字段中 1 和 0 交互使用，接收站通过该字段知道导入帧，并且该字段提供了同步化接收物理层帧接收部分和导入比特流的方法。
- Start-of-Frame Delimiter（SFD） － 1 字节。字段中 1 和 0 交互使用，结尾是两个连续的 1，表示下一位是利用目的地址的重复使用字节的重复使用位。
- Destination Address（DA） － 6 字节。DA 字段用于识别需要接收帧的站。
- Source Addresses（SA） － 6 字节。SA 字段用于识别发送帧的站。

 MSN : MingLang.Xu@Hotmail.Com

- Length/Type — 2 字节。如果是采用可选格式组成帧结构时，该字段既表示包含在帧数据字段中的 MAC 客户机数据大小，也表示帧类型 ID。
- Data — 是一组 n（46=< n =<1500）字节的任意值序列。帧总值最小为 64 字节。
- Frame Check Sequence（FCS） — 4 字节。该序列包括 32 位的循环冗余校验（CRC）值，由发送 MAC 方生成，通过接收 MAC 方进行计算得出以校验被破坏的帧。
- Ext — 扩展。它是一个非数据变量扩展字段，适用于长度小于最小值的帧。

数据包突发模式：

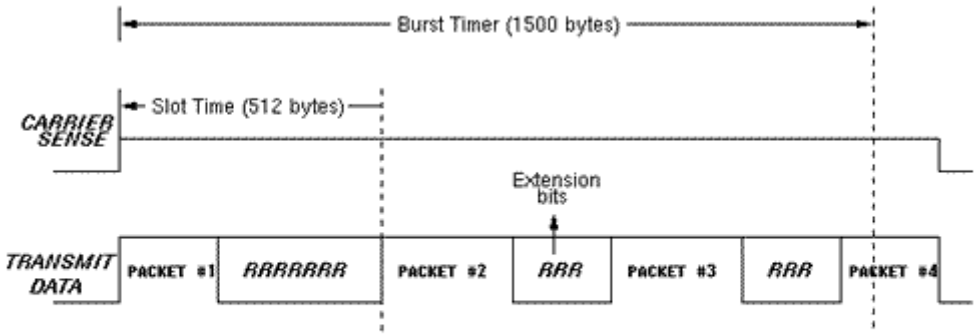
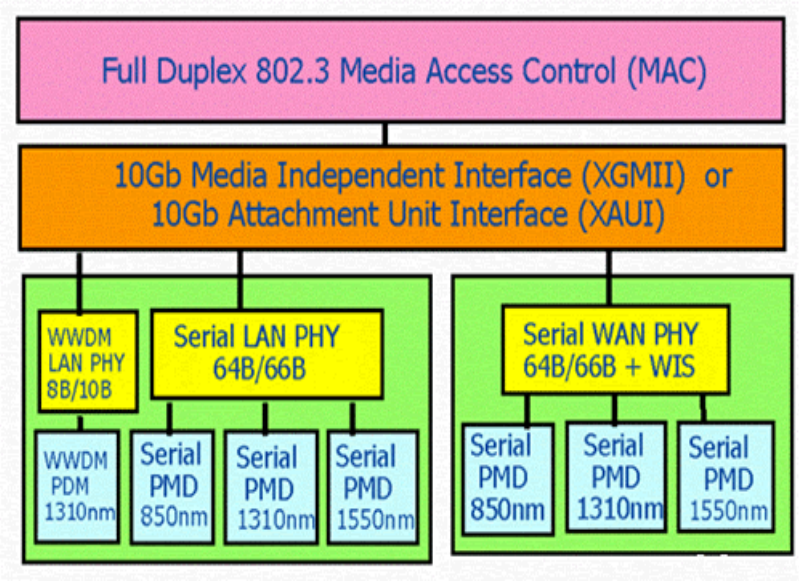


Fig. 2. Packet Bursting

千兆位以太网中的数据包突发模式

10 Gigabit Ethernet：10 千兆位以太网

10 千兆位以太网，定义在 IEEE 802.3ae 中，其数据传输速率达到百亿比特每秒。基于当今广泛应用的以太网技术，10 千兆位以太网提供了与各种以太网标准相似的有利特点。它主要用于局域网（ LAN）、广域网（WAN）以及城域网（MAN）之间的相互连接。它采用大家熟知的以太网介质访问控制协议及其帧格式和帧大小。然而，10 千兆位以太网只支持全双工，而不支持半双工工作模式，并且只工作于光纤上，因此它不再需要其它以太网标准使用的载波监听多路访问和冲突检测（CSMA/CD）协议。10 千兆位以太网结构如下所示：



10 千兆位以太网结构

10 千兆位规范包含在 IEEE 802.3 标准的补充标准 IEEE 802.3ae 中，它扩展了 IEEE 802.3 协议和 MAC 规范使其支持 10Gb/s 的传输速率。除此之外，通过 WAN 界面子层（WIS：WAN interface sublayer），10 千兆位以太网也能被调整为较低的传输速率，如 9.584640 Gb/s （OC-192），这就允许 10 千兆位以太网设备与同步光纤网络（SONET） STS -192c 传输格式相兼容。

10GBASE-SR 和 10GBASE-SW 主要支持短波（850 nm）多模光纤（MMF），光纤距离为 2m 到 300 m 。10GBASE-SR 主要支持“暗光纤”（dark fiber），暗光纤是指没有光传播并且不与任何设备连接的光纤。10GBASE-SW 主要用于连接 SONET 设备，它应用于远程数据通信。

10GBASE-LR 和 10GBASE-LW 主要支持长波（1310nm）单模光纤（SMF），光纤距离为 2m 到 10km （约 32808 英尺）。10GBASE-LW 主要用来连接 SONET 设备时，10GBASE- LR 则用来支持“暗光纤”（dark fiber）。

10GBASE-ER 和 10GBASE-EW 主要支持超长波（1550nm）单模光纤（SMF），光纤距离为 2m 到 40km （约 131233 英尺）。10GBASE-EW 主要用来连接 SONET 设备，10GBASE- ER 则用来支持“暗光纤”（dark fiber）。

最后，还有一种 10GBASE-LX4 介质类型，采用波分复用技术，在单对光缆上以四倍光波长发送信号。10GBASE-LX4 系统运行在 1310nm 的多模或单模暗光纤方式下。该系统的设计目标是针对于 2m 到 300 m 的多模光纤模式或 2m 到 10km 的单模光纤模式。

协议结构

关于以太网 802.3.10 中，千兆位以太网帧最小为 64 字节，最大可达到 1518 字节。

7	1	6	6	2	46=< n =<1500 bytes	4 bytes

 MSN : MingLang.Xu@Hotmail.Com

Pre	SFD	DA	SA	Length Type	Data unit + pad	FCS

- Preamble（Pre） — 7 字节。Pre 字段中 1 和 0 交互使用，接收站通过该字段知道导入帧，并且该字段提供了同步化接收物理层帧接收部分和导入比特流的方法。
- Start-of-Frame Delimiter（SFD） — 1 字节。字段中 1 和 0 交互使用，结尾是两个连续的 1，表示下一位是利用目的地址的重复使用字节的重复使用位。
- Destination Address（DA） — 6 字节。DA 字段用于识别需要接收帧的站。
- Source Addresses （SA） — 6 字节。SA 字段用于识别发送帧的站。
- Length/Type — 2 字节。如果是采用可选格式组成帧结构时，该字段既表示包含在帧数据字段中的 MAC 客户机数据大小，也表示帧类型 ID。
- Data — 是一组 n（46=< n =<1500）字节的任意值序列。帧总值最小为 64 字节。
- Frame Check Sequence（FCS） — 4 字节。该序列包括 32 位的循环冗余校验（CRC）值，由发送 MAC 方生成，通过接收 MAC 方进行计算得出以校验被破坏的帧。

IEEE 802.1P：有关流量优先级的 LAN 第二层 QoS/CoS 协议

IEEE 802.1P 规范使得第二层交换机能够提供流量优先级和动态组播过滤服务。优先级规范工作在媒体访问控制（MAC）帧层（OSI 参考模型第二层）。802.1P 标准也提供了组播流量过滤功能，以确保该流量不超出第二层交换网络范围。

802.1P 协议头包括一个 3 位优先级字段，该字段支持将数据包分组为各种流量种类。IEEE 极力推荐网络管理员实施这些流量种类，但它并不要求强制使用。流量种类也可以定义为第二层服务质量（QoS）或服务类（CoS），并且在网络适配器和交换机上实现，而不需要任何预留设置。802.1P 流量被简单分类并发送至目的地，而没有带宽预留机制。

802.1P 是 IEEE 802.1Q （VLAN 标签技术）标准的扩充协议，它们协同工作。IEEE 802.1Q 标准定义了为以太网 MAC 帧添加的标签。VLAN 标签有两部分：VLAN ID （12 比特）和优先级（3 比特）。 IEEE 802.1Q VLAN 标准中没有定义和使用优先级字段，而 802.1P 中则定义了该字段。

802.1P 中定义的优先级有 8 种。尽管网络管理员必须决定实际的映射情况，但 IEEE 仍作了大量建议。最高优先级为 7，应用于关键性网络流量，如路由选择信息协议（RIP）和开放最短路径优先（OSPF）协议的路由表更新。优先级 6 和 5 主要用于延迟敏感（delay-sensitive）应用程序，如交互式视频和语音。优先级 4 到 1 主要用于受控负载(controlled-load)应用程序,如流式多媒体(streaming multimedia)和关键性业务流量（business-critical traffic） — 例如，SAP 数据 — 以及 "loss eligible" 流量。优先级 0 是缺省值，并在没有设置其它优先级值的情况下自动启用。

 MSN : MingLang.Xu@Hotmail.Com

协议结构

以太网中的 IEEE 802.1Q 标签帧格式 — 在以太网（802.3）帧基础上修订而成：

7	1	6	6	2	2	2	42-1496 bytes	4 bytes
Preamble	SFD	DA	SA	TPID	TCI	Type Length	Data	CRC

- Preamble（Pre） — 7 字节。Pre 字段中 1 和 0 交互使用，接收站通过该字段知道导入帧，并且该字段提供了同步化接收物理层帧接收部分和导入比特流的方法。
- Start-of-Frame Delimiter（SFD） — 1 字节。字段中 1 和 0 交互使用，结尾是两个连续的 1，表示下一位是利用目的地址的重复使用字节的重复使用位
- Destination Address（DA） — 6 字节。DA 字段用于识别需要接收帧的站。
- Source Addresses（SA） — 6 字节。SA 字段用于识别发送帧的站。
- TPID — 值为 8100（hex）。当帧中的 EtherType 也为 8100 时，该帧传送标签 IEEE 802.1Q/802.1P。
- TCI — 标签控制信息字段，包括用户优先级（User Priority）、规范格式指示器（Canonical Format Indicator）和 VLAN ID。

3	1	12 bits
User Priority	CFI	Bits of VLAN ID (VIDI) to identify possible VLANs

  - User Priority：定义用户优先级，包括 8 个（2^3）优先级别。IEEE 802.1P 为 3 比特的用户优先级位定义了操作。
  - CFI：以太网交换机中，规范格式指示器总被设置为 0。由于兼容特性，CFI 常用于以太网类网络和令牌环类网络之间，如果在以太网端口接收的帧具有 CFI，那么设置为 1，表示该帧不进行转发，这是因为以太网端口是一个无标签端口。
  - VID：VLAN ID 是对 VLAN 的识别字段，在标准 802.1Q 中常被使用。该字段为 12 位。支持 4096（2^12） VLAN 的识别。在 4096 可能的 VID 中，VID＝0 用于识别帧优先级。 4095（FFF）作为预留值，所以 VLAN 配置的最大可能值为 4,094。
- Length/Type — 2 字节。如果是采用可选格式组成帧结构时，该字段既表示包含在帧数据字段中的 MAC 客户机数据大小，也表示帧类型 ID。
- Data — 是一组 n（46=< n =<1500）字节的任意值序列。帧总值最小为 64 字节。
- Frame Check Sequence（FCS） — 4 字节。该序列包括 32 位的循环冗余校验（CRC）值，由发送 MAC 方生成，通过接收 MAC 方进行计算得出以校验被破坏的帧。

MSN : MingLang.Xu@Hotmail.Com

FDDI：光纤分布式数据接口

光纤分布式数据接口（FDDI）是由美国国家标准化组织（ANSI）制定的在光缆上发送数字信号的一组协议。FDDI 使用双环令牌，传输速率可以达到 100Mbps。由于支持高宽带和远距离通信网络，FDDI 通常用作骨干网。CCDI 是 FDDI 的一 种变型，它采用双绞铜缆为传输介质，数据传输速率通常为 100Mbps。

FDDI-2 是 FDDI 的扩展协议，支持语音、视频及数据传输。FDDI 的另一个变种，称为 FDDI 全双工技术（FFDT），它采用与 FDDI 相同的网络结构，但传输速率可以达到 200Mbps 。

FDDI 使用双环架构，两个环上的流量在相反方向上传输。双环由主环和备用环组成。在正常情况下，主环用于数据传输，备用环闲置。正如本篇后面所述，使用双环的用意是能够提供较高的可靠性和健壮性。

FDDI 详细阐明了 OSI 参考模型的物理层和介质访问层。实质上 FDDI 并不是单一规范，而是由四个子部分组成，每部分具有各自特定功能。各部分合起来使得 FDDI 能够在上层协议（如 TCP/IP、IPX）和介质（如光缆）间提供高速连接。

FDDI 四个子规范为介质访问控制（MAC）、物理层协议层（PHY）、物理介质相关层（PMD）以及站管理（SMT）。MAC 规定了怎样访问介质，包括协议所需要的帧格式、寻址、令牌处理、循环冗余校验算法（CRC）以及差错恢复机制。PHY 规定了传输编码和解码程序、时钟要求及其它功能；PMD 规定了传输介质应具备的特性，包括光纤链路（fiber-optic link）、功率电平（power level）、误码率（bit-error rate）、光纤器件（optical component）以及连接器（connector）。SMT 规定了 FDDI 站配置、环配置以及环控制等特征，包括站的插入和删除、启动、故障分离和恢复、模式安排及统计集合。

协议结构

2	6	6	0-30	Variable	4 bytes
Frame Control	Destination Address	Source Address	Route Information	Information	FCS

- Frame control — 该字段结构如下：

C	L	F	F	Z	Z	Z	Z

  - C 类别位： 0 异步帧；1 同步帧
  - L 地址长度位：0 16 位（never）；1 48 位（always）
  - FF 帧位
  - ZZZZ 控制位
- Destination Address — 该地址字段结构如下：

I/G	U/L	Address Bits

- Source Address — 该地址字段结构如下：

I/G>	RII	Address Bits

  - I/G 个人/组地址：0 组地址；1 个人地址
  - RII 路由信息指示器： 0 RI 不在；1 RI 在
- Route Information — 路由信息字段结构如下：

3	5	1	6	1	16	16		16
RT	LTH	D	LF	r	RD1	RD2	…	RDn

  - RC — 路由选择控制（16 位）
  - RDn — 路由描述符（16 位）
  - RT — 路由选择类型（3 位）
  - LTH — 长度（5 位）
  - D — 方向位（1 位）
  - LF — 最大帧（6 位）
  - r — 预留（1 位）
- Information — Information 字段可能为 LLC、MAC 或 SMT 协议。
- FCS — 帧校验序列。

ARP 和 InARP：地址转换协议和逆向地址转换协议

地址转换协议（ARP）是用来实现 IP 地址与本地网络认知的物理地址（以太网 MAC 地址）之间的映射。例如，在第四版 IP 中， IP 地址长为 32 位。然而在以太局域网络中，设备地址长为 48 位。有一张表格，通常称为 ARP 缓冲（ARP cache），来维持每个 MAC 地址与其相应的 IP 地址之间的对应

关系。 ARP 提供一种形成该对应关系的规则以及提供双向地址转换。

由于每一类局域网协议细节不同,那么就需要为以太网、帧中继、ATM 、光纤分布式数据接口、HIPPI 以及其它协议等提供独立的 ARP 规范说明。 InARP 是 ARP 的补充协议以支持帧中继环境下的 ARP 。

此外还为不知道自己 IP 地址的主机提供了一种反向地址转换协议(RARP),从而可以从网关的 ARP cache 上请求它们的 IP 地址。有关 RARP 的具体细节可参见个别文件。

协议结构

<table>
<tr><th colspan="2">16</th><th>32 bit</th></tr>
<tr><td colspan="2">Hardware Type</td><td>Protocol Type</td></tr>
<tr><td>HLen</td><td>Plen</td><td>Operation</td></tr>
<tr><td colspan="3">Sender Hardware Address</td></tr>
<tr><td colspan="3">Sender Protocol Address</td></tr>
<tr><td colspan="3">Target Hardware Address</td></tr>
<tr><td colspan="3">Target Protocol Address</td></tr>
</table>

- Hardware Type — 指定一种硬件接口类型，为发送方请求响应所用。
- Protocol Type — 指由发送方提供的高级协议地址类型。
- Hlen — 硬件地址大小。
- Plen — 协议地址大小。
- Operation — 各个值如下表所示:

1	ARP Request
2	ARP Response
3	RARP Request
4	RARP Response
5	Dynamic RARP Request
6	Dynamic RARP Reply
7	Dynamic RARP Error
8	InARP Request
9	InARP Reply

- Sender Hardware Address — HLen 二进制大小
- Sender Protocol Address — PLen 二进制大小
- Target Hardware Address — HLen 二进制大小
- Target Protocol Address — PLen 二进制大小

RARP：反向地址转换协议

反向地址转换协议（RARP）允许局域网的物理机器从网关服务器的 ARP 表或者缓存上请求其 IP 地址。网络管理员在局域网网关路由器里创建一个表以映射物理地址（MAC）和与其对应的 IP 地址。当设置一台新的机器时，其 RARP 客户机程序需要向路由器上的 RARP 服务器请求相应的 IP 地址。假设在路由表中已经设置了一个记录， RARP 服务器将会返回 IP 地址给机器，此机器就会存储起来以便日后使用。

RARP 可以使用于以太网、光纤分布式数据接口及令牌环 LAN 。

协议结构

RARP 协议头结构和 ARP 相同：

<table>
<tr><td colspan="2">16</td><td>32 bit</td></tr>
<tr><td colspan="2">Hardware Type</td><td>Protocol Type</td></tr>
<tr><td>Hlen</td><td>Plen</td><td>Operation</td></tr>
<tr><td colspan="3">Sender Hardware Address</td></tr>
<tr><td colspan="3">Sender Protocol Address</td></tr>
<tr><td colspan="3">Target Hardware Address</td></tr>
<tr><td colspan="3">Target Protocol Address</td></tr>
</table>

- Hardware Type — 指定一种硬件接口类型，为发送方请求响应所用。
- Protocol Type — 指由发送方提供的高级协议地址类型。
- Hlen — 硬件地址大小。
- Plen — 协议地址大小。
- Operation — 各个值如下表所示：

1	ARP Request

2	ARP Response
3	RARP Request
4	RARP Response
5	Dynamic RARP Request
6	Dynamic RARP Reply
7	Dynamic RARP Error
8	InARP Request
9	InARP Reply

- Sender Hardware Address — HLen 二进制大小
- Sender Protocol Address — PLen 二进制大小
- Target Hardware Address — HLen 二进制大小
- Target Protocol Address — PLen 二进制大小

RPC：远程过程调用协议

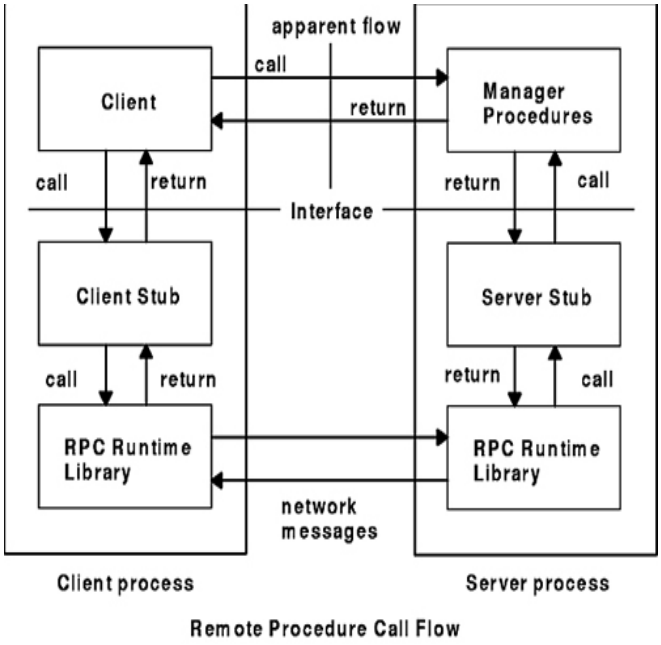
远程过程调用（RPC）是一种通过网络从远程计算机程序上请求服务，而不需要了解底层网络技术的协议。RPC 协议假定某些传输协议的存在，如 TCP 或 UDP，为通信程序之间携带信息数据。在 OSI 网络通信模型中，RPC 跨越了传输层和应用层。RPC 使得开发包括网络分布式多程序在内的应用程序更加容易。

RPC 采用客户机/服务器模式。请求程序就是一个客户机，而服务提供程序就是一个服务器。首先，调用进程发送一个有进程参数的调用信息到服务进程，然后等待应答信息。在服务器端，进程保持睡眠状态直到调用信息的到达为止。当一个调用信息到达，服务器获得进程参数，计算结果，发送答复信息，然后等待下一个调用信息，最后，客户端调用过程接收答复信息，获得进程结果，然后调用执行继续进行。

目前，有多种 RPC 模式和执行。最初由 Sun 公司提出。IETF ONC 宪章重新修订了 Sun 版本，使得 ONC PRC 协议成为 IETF 标准协议。现在使用最普遍的模式和执行是开放式软件基础的分布式计算环境（DCE）。

协议结构

远程过程调用（RPC）信息协议由两个不同结构组成：调用信息和答复信息。信息流程如下所示：



Remote Procedure Call Flow

RPC：远程过程调用流程

RPC 调用信息：每条远程过程调用信息包括以下无符号整数字段，以独立识别远程过程：

- 程序号（Program number）
- 程序版本号（Program version number）
- 过程号（Procedure number）

RPC 调用信息主体形式如下：

```
struct call_body {

  unsigned int rpcvers;

  unsigned int prog;

  unsigned int vers;

  unsigned int proc;

  opaque_auth cred;

  opaque_auth verf;

  1 parameter

  2 parameter . . .
```

};

RPC 答复信息：RPC 协议的答复信息的改变取决于网络服务器对调用信息是接收还是拒绝。答复信息请求包括区别以下情形的各种信息：

- RPC 成功执行调用信息。.
- RPC 的远程实现不是协议第二版，返回 RPC 支持的最低和最高版本号。
- 在远程系统中，远程程序不可用。
- 远程程序不支持被请求的版本号。返回远程程序所支持的最低和最高版本号。
- 请求的过程号不存在。通常是呼叫方协议或程序差错。

RPC 答复信息形式如下：

```
enum reply_stat stat {

MSG_ACCEPTED = 0,

MSG_DENIED   = 1

};
```

HSRP：热备份路由器协议

热备份路由器协议（HSRP） 的设计目标是支持在特定环境下 IP 流量中断后的无损故障转移，并允许主机使用单路由器，以及即使在实际的第一跳路由器当机的情形下仍能保持网络连接。换句话说，当主机不能动态习得第一跳路由器的 IP 地址时，HSRP 协议能够保护第一跳路由不失败。该协议有多个路由器参与，并共同创建了一个虚拟路由器。HSRP 协议确保有且只有一个路由器代表虚拟路由器实现数据包转发过程。终端主机将它们各自的数据包转发到该虚拟路由器上。

负责转发数据包的路由器称之为活动路由器（active router）。一旦活动路由器出现故障，HSRP 将激活备份路由器（standby router）取代主动路由器。根据 所有路由器的 IP 地址，HSRP 协议提供了一种决定使用活动路由器还是备份路由器的机制。一旦活动路由器出现故障，备份路由器就会接管其任务，而不会中断主机的网络连接。

HSRP 运行在 UDP 上，采用端口号 1985。路由器使用它们的实际 IP 地址转发协议数据包，而并非虚拟地址，正是基于这一点， HSRP 路由器间能相互识别。

协议结构

8	16	24	32 bit
Version	Op code	State	Hellotime
Holdtime	Priority	Group	Reserved
Authentication data			
Authentication data			
Virtual IP address			

- Version — HSRP 版本号。当前值为 0。
- Op code — 数据包中包含的信息类型。可能值有：
  - 0 — 发送 Hello，表示路由器正在运行，并有可能成为主动或备份路由器。
  - 1 — 发送 Coup，当路由器希望成为主动路由器时发送。
  - 2 — 发送 Resign，当路由器不再希望成为主动路由器时发送。
- State — 备份组中的每个路由器运行一个状态机器。State 字段描述发送信息的路由器的当前状态。可能值有：0 Initial；1 Learn；2 Listen；4 Speak；8 Standby； 16 Active。
- Hellotime — 指路由器发送的 Hello 信息间的大约周期。如果不在路由器上配置 Hellotime 字段，将从主动路由器的 Hello 信息获知。
- Holdtime — 时间值，指当前 Hello 信息的有效时间（只对 Hello 信息）。
- Priority — 用于选择主动和备份路由器。当比较两个路由器的优先级时，具有较高优先级数字的路由器优先。当两个路由器具有同等优先级时，IP 地址较高的那个路由器优先。
- Group — 识别备份路由器组。对于令牌环，值在 0 到 2 之间为有效。而对于其它媒体，值在 0 到 255 之间为有效。
- Authentication Data — 清除 8 字符再生密码。如果没有配置 Authentication Data 字段，那么推荐缺省值为 0x63 0x69 0x73 0x63 0x6F 0x00 0x00 0x00。
  Virtual IP Address — 该组使用的虚拟 IP 地址。如果路由器中没有配置虚拟 IP 地址，将从主动路由器的 Hello 信息获知。通常地址只在没有配置地址和 Hello 信息通过认证的情况下获知。

GVRP：GARP VLAN 注册协议

GARP VLAN 注册协议（GVRP）是一种 GARP （通用属性注册协议）的应用，提供 802.1Q 兼容的 VLAN

裁剪（VLAN pruning）功能和在 802.1Q 干线端口（trunk port）上建立动态 VLAN。GVRP 定义在 IEEE 802.1P 标准中，允许对 802.1Q VLAN 进行控制。

GVRP 交换机之间能够相互交换 VLAN 配置信息，裁剪不必要的广播和未知单播流量以及在通过 802.1Q 干线连接的交换机上动态创建和管理 VLAN 。

GVRP 中采用了 GID 和 GIP，这两部分分别提供了用于基于 GARP 应用程序的通用状态机制描述和通用信息传播机制。GVRP 只运行在 802.1Q 干线链路上。GVRP 通过剪除干线链路使得只有活动 VLAN 才在干线连接上传输。在 GVRP 为干线添加一个 VLAN 之前，它首先要收到来自交换机的 join 信息。GVRP 更新信息和计时器都是可以改变的。GVRP 端口有多种运行模式，从而控制它们裁剪 VLAN 的方式。GVRP 能够为 VLAN 数据库动态添加和管理 VLAN 。

换句话说，GVRP 支持设备之间 VLAN 信息的传播服务。在 GVRP 中，能够手工配置一台交换机的 VLAN 信息，该网络中的其它所有交换机能够动态了解那些 VLAN 的情况。终端节点可以接入任何交换机并且连接到所需的 VLAN 上。终端要使用 GVRP 就需要安装 GVRP 兼容的网络接口卡（NIC）。GVRP 兼容的 NIC 可以配置加入到所需的 VLAN 或 VLAN，然后接入一个 GVRP-enabled 交换机。NIC 与交换机之间建立通信连接，并在两者之间实现了 VLAN 连通性。

协议结构

GVRP 与 GARP 具有相同的结构。GVRP 特定属性类型可能有：1、VID 组属性类型（Group Attribute

Type）。

GARP PDU 格式：

2 bytes	
Protocol ID	Message

GARP 信息结构：

1 byte				
Attribute Type	Attribute 1	. . .	Attribute n	End Mark

GARP 属性结构：

1 byte	1 byte	1 byte
Attribute Length	Attribute Event	Attribute Value

- Protocol ID — 识别 GARP 协议。
- Identifier — 十进制值，用于匹配 Request 和 Reply 命令。
- Attribute Type — 定义属性。可能值有：1、组属性（Group Attribute）；2、服务请求属性（Service Requirement Attribute）。
- Attribute Length — 属性长。
- Attribute Event — Attribute Event 字段值可能为：0 — Leave_all；1 — Join_Empty Operator；2 — Join_In Operator；3 — Leave_Empty Operator；4 — Leave_In Operator；5 — Empty Operator。
- Attribute Value — 该字段编码与 Attribute Type 规范相一致。
- End Mark — 编码为 0。

VLAN：虚拟局域网和 IEEE 802.1Q

虚拟局域网（VLAN）是指位于一个或多个局域网的设备经过配置能够像连接到同一个信道一样进行

通信，而实际上它们分布在不同的局域网段中。由于 VLAN 基于逻辑连接而不是物理连接，所以它可以提供灵活的用户/主机管理、带宽分配以及资源优化等服务。

从技术角度讲，VLAN 一般有以下四种划分方法：

- 基于端口的 VLAN：为每个物理交换端口配置一个访问列表标示其 VLAN 成员关系。
- 基于 MAC 地址的 VLAN:为交换机配置一个访问列表,将各个 MAC 地址映射到 VLAN 成员关系。
- 基于协议的 VLAN：为交换机配置一个访问列表，将第三层协议类型映射到 VLAN 成员关系，这样就可以过滤掉来自于像使用 IPX 的终端站的 IP 流量。
- ATM VLAN：使用 LAN 仿真（LANE）协议将以太网数据包映射到 ATM 信元上，并通过以太网 MAC 地址映射到 ATM 地址的方式将数据包发送到目的地。

IEEE 802.1Q 规范为标识带有 VLAN 成员信息的以太帧建立了一种标准方法。IEEE 802.1Q 标准定义了 VLAN 网桥操作，从而允许在桥接局域网结构中实现定义、运行以及管理 VLAN 拓朴结构等操作。802.1Q 标准主要用来解决如何将大型网络划分为多个小网络,如此广播和组播流量就不会占据更多带宽的问题。此外 802.1Q 标准还提供更高的网络段间安全性。

IEEE 802.1Q 完成以上各种功能的关键在于标签。支持 802.1Q 的交换端口可被配置来传输标签帧或无标签帧。一个包含 VLAN 信息的标签字段可以插入到以太帧中。如果端口有支持 802.1Q 的设备（如另一个交换机）相连，那么这些标签帧可以在交换机之间传送 VLAN 成员信息，这样 VLAN 就可以跨越多台交换机。但是，对于没有支持 802.1Q 设备相连的端口我们必须确保它们用于传输无标签帧，这一点非常重要。很多 PC 和打印机的 NIC 并不支持 802.1Q ，一旦它们收到一个标签帧，它们会因为读不懂标签而丢弃该帧。在 802.1Q 中，用于标签帧的最大合法以太帧大小已由 1，518 字节增加到 1，522 字节，这样就会使网卡和旧式交换机由于帧“尺寸过大”而丢弃标签帧。

协议结构

以太网中的 IEEE 802.1Q 标签帧格式：

7	1	6	6	2	2	2	42-1496 bytes	4 bytes
Preamble	SFD	DA	SA	TPID	TCI	Type Length	Data	CRC

- Preamble（Pre） — 7 字节。Pre 字段中 1 和 0 交互使用，接收站通过该字段知道导入帧，并且该字段提供了同步化接收物理层帧接收部分和导入比特流的方法。
- Start-of-Frame Delimiter（SFD） — 1 字节。字段中 1 和 0 交互使用，结尾是两个连续的 1，表示下一位是利用目的地址的重复使用字节的重复使用位。
- Destination Address（DA） — 6 字节。DA 字段用于识别需要接收帧的站。
- Source Addresses（SA） — 6 字节。SA 字段用于识别发送帧的站。
- TPID — 值为 8100（hex）。当帧中的 EtherType 也为 8100 时，该帧传送标签 IEEE 802.1Q/802.1P。
- TCI — 标签控制信息字段，包括用户优先级（User Priority）、规范格式指示器（Canonical Format Indicator）和 VLAN ID。

- User Priority：定义用户优先级，包括 8 个（2^3）优先级别。IEEE 802.1P 为 3 比特的用户优先级位定义了操作。
- CFI：以太网交换机中，规范格式指示器总被设置为 0。由于兼容特性，CFI 常用于以太网类网络和令牌环类网络之间，如果在以太网端口接收的帧具有 CFI，那么设置为 1，表示该帧不进行转发，这是因为以太网端口是一个无标签端口。
- VID：VLAN ID 是对 VLAN 的识别字段，在标准 802.1Q 中常被使用。该字段为 12 位。支持 4096（2^12） VLAN 的识别。在 4096 可能的 VID 中，VID＝0 用于识别帧优先级。 4095（FFF）作为预留值，所以 VLAN 配置的最大可能值为 4094。

- Length/Type — 2 字节。如果是采用可选格式组成帧结构时，该字段既表示包含在帧数据字段中的 MAC 客户机数据大小，也表示帧类型 ID。
- Data — 是一组 n（46=< n =<1500）字节的任意值序列。帧总值最小为 64 字节。
- Frame Check Sequence（FCS） — 4 字节。该序列包括 32 位的循环冗余校验（CRC）值，由发送 MAC 方生成，通过接收 MAC 方进行计算得出以校验被破坏的帧。

VLAN：虚拟局域网和 IEEE 802.1Q

虚拟局域网（VLAN）是指位于一个或多个局域网的设备经过配置能够像连接到同一个信道一样进行通信，而实际上它们分布在不同的局域网段中。由于 VLAN 基于逻辑连接而不是物理连接，所以它可以提供灵活的用户/主机管理、带宽分配以及资源优化等服务。

从技术角度讲，VLAN 一般有以下四种划分方法：

- 基于端口的 VLAN：为每个物理交换端口配置一个访问列表标示其 VLAN 成员关系。
- 基于 MAC 地址的 VLAN:为交换机配置一个访问列表,将各个 MAC 地址映射到 VLAN 成员关系。
- 基于协议的 VLAN：为交换机配置一个访问列表，将第三层协议类型映射到 VLAN 成员关系，这样就可以过滤掉来自于像使用 IPX 的终端站的 IP 流量。
- ATM VLAN：使用 LAN 仿真（LANE）协议将以太网数据包映射到 ATM 信元上，并通过以太网 MAC 地址映射到 ATM 地址的方式将数据包发送到目的地。

IEEE 802.1Q 规范为标识带有 VLAN 成员信息的以太帧建立了一种标准方法。IEEE 802.1Q 标准定义了 VLAN 网桥操作，从而允许在桥接局域网结构中实现定义、运行以及管理 VLAN 拓朴结构等操作。802.1Q 标准主要用来解决如何将大型网络划分为多个小网络,如此广播和组播流量就不会占据更多带宽的问题。此外 802.1Q 标准还提供更高的网络段间安全性。

IEEE 802.1Q 完成以上各种功能的关键在于标签。支持 802.1Q 的交换端口可被配置来传输标签帧或无标签帧。一个包含 VLAN 信息的标签字段可以插入到以太帧中。如果端口有支持 802.1Q 的设备（如另一个交换机）相连，那么这些标签帧可以在交换机之间传送 VLAN 成员信息，这样 VLAN 就可以跨越多台交换机。但是，对于没有支持 802.1Q 设备相连的端口我们必须确保它们用于传输无标签帧，这一点非常重要。很多 PC 和打印机的 NIC 并不支持 802.1Q ，一旦它们收到一个标签帧，它们会因为读不懂标签而丢弃该帧。在 802.1Q 中，用于标签帧的最大合法以太帧大小已由 1，518 字节增加到 1，522 字节，这样就会使网卡和旧式交换机由于帧“尺寸过大”而丢弃标签帧。

 MSN : MingLang.Xu@Hotmail.Com

协议结构

以太网中的 IEEE 802.1Q 标签帧格式：

7	1	6	6	2	2	2	42-1496 bytes	4 bytes
Preamble	SFD	DA	SA	TPID	TCI	Type Length	Data	CRC

- Preamble（Pre） — 7 字节。Pre 字段中 1 和 0 交互使用，接收站通过该字段知道导入帧，并且该字段提供了同步化接收物理层帧接收部分和导入比特流的方法。
- Start-of-Frame Delimiter（SFD） — 1 字节。字段中 1 和 0 交互使用，结尾是两个连续的 1，表示下一位是利用目的地址的重复使用字节的重复使用位。
- Destination Address（DA） — 6 字节。DA 字段用于识别需要接收帧的站。
- Source Addresses（SA） — 6 字节。SA 字段用于识别发送帧的站。
- TPID — 值为 8100（hex）。当帧中的 EtherType 也为 8100 时，该帧传送标签 IEEE 802.1Q/802.1P。
- TCI — 标签控制信息字段，包括用户优先级（User Priority）、规范格式指示器（Canonical Format Indicator）和 VLAN ID。
  - User Priority：定义用户优先级，包括 8 个（2^3）优先级别。IEEE 802.1P 为 3 比特的用户优先级位定义了操作。
  - CFI：以太网交换机中，规范格式指示器总被设置为 0。由于兼容特性，CFI 常用于以太网类网络和令牌环类网络之间，如果在以太网端口接收的帧具有 CFI，那么设置为 1，表示该帧不进行转发，这是因为以太网端口是一个无标签端口。
  - VID：VLAN ID 是对 VLAN 的识别字段，在标准 802.1Q 中常被使用。该字段为 12 位。支持 4096（2^12） VLAN 的识别。在 4096 可能的 VID 中，VID＝0 用于识别帧优先级。 4095（FFF）作为预留值，所以 VLAN 配置的最大可能值为 4094。
- Length/Type — 2 字节。如果是采用可选格式组成帧结构时，该字段既表示包含在帧数据字段中的 MAC 客户机数据大小，也表示帧类型 ID。
- Data — 是一组 n（46=< n =<1500）字节的任意值序列。帧总值最小为 64 字节。
- Frame Check Sequence（FCS） — 4 字节。该序列包括 32 位的循环冗余校验（CRC）值，由发送 MAC 方生成，通过接收 MAC 方进行计算得出以校验被破坏的帧。

802.1X：WLAN 认证&密钥管理

IEEE 802.1X 是一种为受保护网络提供认证、控制用户通信以及动态密钥分配等服务的有效机制。802.1X 将可扩展身份认证协议（EAP）捆绑到有线和无线局域网介质上，以支持多种认证方法，如令牌

（token card）、Kerberos、一次性口令（one-time password）、证书（certificate）以及公开密钥认证（public key authentication）等。

802.1x 结构主要有三部分组成：1. 申请者（supplicant）：想得到认证的用户或客户；2. 认证服务器（authentication server）：通常为 RADIUS 服务器；3. 认证系统（authenticator）：如无线接入点。

802.1x 中的主要协议是 LAN 上的 EAP 封装协议（EAPOL），当前它被定义来用于 Ethernet-like LAN，包括 802.11 无线局域网、令牌环网（包括 FDDI）。802.1X 中操作过程如下：

1. 申请者（如客户机无线网卡）发送一个“EAP 响应/身份认证” 数据包给认证系统（如 802.11 接入点），然后传送至认证服务器（RADIUS 服务器，位于接入点有线端）。
2. 认证服务器发回一个验证给认证系统，认证系统将此验证在 IP 层解包并重新打包在 EAPOL，然后再发送给申请者。
3. 申请者响应认证系统发送来的验证，并将响应通过认证系统传送给认证服务器。认证服务器使用特定认证算法来检验客户身份，这可以通过数字证书或其它 EAP 认证类型实现。

如果申请者提供的身份正确，认证服务器便响应一个成功信息，并将该信息返回给申请者。认证系统根据认证服务器返回的信息属性打开端口为申请者能够访问 LAN。

不管是否实施 802.11 WEP 密钥机制或完全没有提供加密技术，802.1X（EAPOL）协议都提供了有效的认证机制。如果 802.1X 服务器配置为实现动态密钥交换，那么它便能将接收信息连同会话密钥一起发送接入点。发送接入点使用会话密钥建立、标记和加密 EAP 密钥信息，并在发送完成功信息后立即将其发送给客户机。客户机通过密钥信息内容来定义应用加密密钥。

802.1X（EAPOL）实际上是一种传送机制，而不提供实质的认证机制。当采用 802.1X 时，必须选择某种 EAP 类型，如传输层安全协议 （EAP-TLS）或 EAP 隧道传输层安全协议（EAP-TTLS），它们定义认证如何发生。特定类型的 EAP 位于认证服务器中或客户机操作系统或应用软件里。接入点作为 802.1X 信息的“通过”路径，这意味着在支持 802.1X 的接入点不需要升级的情况下，可以指定使用任意类型的 EAP 。

协议结构

802.3/Ethernet 中的 EAPOL 帧格式如下所示：

2 bytes	1 byte	1 byte	2 bytes	Variable
PAE Ethernet Type	Protocol Version	Packet Type	Packet Body Length	Packet Body

- PAE Ethernet Type — PAE（端口访问实体） Ethernet Type 包括 PAE 分配使用的以太网类型值。
- Protocol Version — 无符号二进制数，为 EAPOL 协议的版本。
- Packet Type — 无符号二进制数，该字段值用于决定数据包类型：a、EAP-packet；b、

EAPOL-Start；c、EAPOL-Logoff；d、EAPOL-Key；e、EAPOL-Encapsulated-ASF-Alert。

- Packet Body Length — 无符号二进制数，该字段用于定义 Packet Body 字段的长度（八位字节）。
- Packet Body — 如果 Packet Type 包括 EAP-Packet、EAPOL-Key 或 EAP-Encapsulated-ASF-Alert，那么需使用该字段，否则不予使用。

Token Ring /FDDI 中的 EAPOL 帧格式：

8 bytes	1 byte	1 byte	2 bytes	Variable
SNAP Ethernet Type	Protocol version	Packet type	Packet Body length	Packet Body

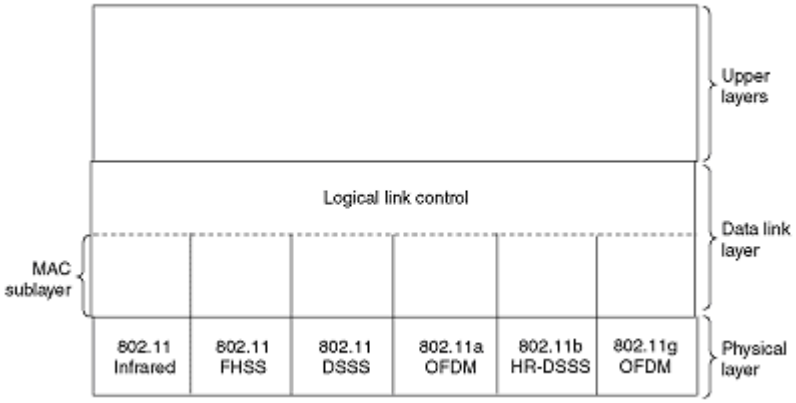
SNAP Ethernet Type — 包括在 SNAP 格式中的 SNAP-encoded Ethernet Type 编码类型：1-3 字节传送标准 SNAP 头，4-6 字节传送 SNAP PID；7-8 字节传送 PAE Ethernet Type 值。

WLAN：无线局域网

无线局域网（WLAN）技术定义在 IEEE 802.11 规范系列中。目前该系列包含以下四种规范：802.11、802.11a、802.11b 以及 802.11g。所有这四种协议都采用以太网协议和载波监听多路访问/冲突避免技术（CSMA/CA，替代了 CSMA/CD）来实现信道共享。

- 802.11 — 应用于无线局域网，使用 2.4 GHz 波段，传输速率为 1 Mbps 或 2 Mbps；既支持跳频技术（FHSS）也支持直序列展频（DSSS）。
- 802.11a — 802.11 的扩展，使用 5GHz 波段，传输速率为 54 Mbps；802.11a 支持正交频分复用（OFDM）编码方式，而不支持 FHSS 或 DSSS。802.11a 应用于无线 ATM 系统并用于接入集线器（access hub）。
- 802.11b（又称为 802.11 高速率或 Wi-Fi） — 802.11 的扩展，使用 2.4 GHz 波段，传输速率为 11 Mbps（也可能降低为 5.5 Mbps、2 Mbps 或 1 Mbps）。802.11b 只支持 DSSS。802.11b 是原 802.11 标准的修订版，用无线方式实现类似以太网功能。
- 802.11g — 支持短距离无线传输，使用 2.4 GHz 波段，传输速率为 20 Mbps 到 54 Mbps。802.11g 支持 OFDM 编码方式。

802.11 中一直采用相移键控（PSK）调制方式。802.11b 中采用的调制方式为补码键控（CCK）， 该方式支持更高的数据传输速率并且不易受多路传播的干扰。802.11a 采用的调制方式为正交频分复用（OFDM），该方式下，数据传输速率可以达到 54 Mbps，但在大多数情况下，通信时的传输速率为 6 Mbps、12 Mbps 或 24 Mbps。802.11 协议栈结构如下：



802.11 栈结构

协议结构

801.11 协议族 MAC 帧结构：

2	2	6	6	6	2	6	0-2312 bytes	4 bytes
Frame Control	Duration	Address 1	Address 2	Address 3	Seq	Address 4	Data	Check Sum

- 帧控制结构：

2	2	4	1	1	1	1	1	1	1	1
Version	Type	Subtype	To DS	From DS	MF	Retry	Pwr	More	W	O

- Protocol Version — 表示 IEEE 802.11 标准版本。
- Type — 帧类型：管理、控制和数据。
- Subtype — 帧子类型：认证帧（Authentication Frame）、解除认证帧（Deauthentication Frame）、连接请求帧（Association Request Frame），连接响应帧（Association Response Frame）、重新连接请求帧（Reassociation Request Frame）、重新连接响应帧（Reassociation Response Frame）解除连接帧（Disassociation Frame）、信标帧（Beacon Frame）、Probe 帧（Probe Frame）、Probe 请求帧（Probe Request Frame）或 Probe 响应帧（Probe Response Frame）。
- To DS — 当帧发送给 Distribution System（DS）时，该值设置为 1。
- From DS — 当帧从 Distribution System（DS）处接收到时，该值设置为 1。
- MF — More Fragment 表示当有更多分段属于相同帧时该值设置为 1。

  - Retry — 表示该分段是先前传输分段的重发帧。
  - Pwr — Power Management，表示传输帧以后，站所采用的电源管理模式。
  - More — More Data，表示有很多帧缓存到站中。
  - W — WEP，表示根据 WEP（Wired Equivalent Privacy） 算法对帧主体进行加密。
  - O — Order 表示利用严格顺序服务类发送帧的顺序。
- Duration/ID（ID）
- 站 ID 用于 Power-Save Poll 信息帧类型。
- Duration 值用于网络分配向量（NAV）计算。
- Address Fields（1-4） — 包括 4 个地址（源地址、目标地址、发送方地址和接收方地址），取决于帧控制字段（ToDS 和 FromDS 位）。
- Sequence Control — 由分段号和序列号组成。用于表示同一帧中不同分段的顺序，并用于识别数据包副本。
- Data — 发送或接收的信息。
- CRC — 包括 32 位的循环冗余校验（CRC）。

LLC：逻辑链路控制（IEEE 802.2）

逻辑链路控制（LLC）是一种 IEEE 802.2 LAN 协议，规定了数据链路层中 LLC 子层的实现。 IEEE 802.2 LLC 应用于 IEEE802.3 （以太网）和 IEEE802.5（令牌环） LAN ，以实现如下功能：

- 管理数据链路通信
- 链接寻址
- 定义服务接入点 Service Access Points （SAP）
- 排序

LLC 为上层提供了处理任何类型 MAC 层的方法，例如，以太网 IEEE 802.3 CSMA/CD 或者令牌环 IEEE 802.5 令牌传递（Token Passing）方式。

LLC 是在高级数据链路控制（HDLC ： High-Level Data-Link Control）的基础上发展起来的，并使用了 HDLC 规范的子集。LLC 定义了三种数据通信操作类型：

类型 1：无连接。该方式不保证发送的信息一定可以收到。

类型 2：面向连接。该方式提供了四种服务：连接的建立、确认和数据到达响应、差错恢复（通过请求重发接收到的错误数据实现）以及滑动窗口（系数：128）。滑动窗口用来提高数据传输速率。

类型 3：无连接应答响应服务。

类型 1 的 LLC 无连接服务中规定了一种静态帧格式，并允许在其上运行网络协议。使用传输层协议的网络协议通常会使用服务类型 1 方式。

 MSN : MingLang.Xu@Hotmail.Com

类型 2 的 LLC 面向连接服务支持可靠数据传输,运用于不需要调用网络层和传输层协议的局域网环境 。

协议结构

逻辑链路控制层（LLC）头：

8	16	24 or 32 bit	Variable
DSAP	SSAP	Control	LLC Information

DSAP — 目标服务访问点（Destination Service Access Point）结构如下：

1	8 bit
I/G	Address Bits

I/G：个人/组地址可能为：0 表示个人 DSAP；1 表示组 DSAP

SSAP — 源服务访问点（Source Service Access Point）字段结构如下：

1	8 bit
C/R	Address Bits

C/R：命令/响应： 0 命令；1 R 响应

Control — 控制（Control）字段结构如下：

	1	8			9	16 bit
Information	0	N(S)			P/F	N(R)
Supervisory	1	0	SS	XXXX	P/F	N(R)
Unnumbered	1	1	MM	P/F	MMM	

N(S) — 发送方发送序列号 Transmitter send sequence number。

N(R) — 发送方接收序列号 Transmitter receive sequence number。

P/F — Poll/final 位。命令 LLC PDU 传输/响应 LLC PDU 传输。

- S — 监督功能位。
- 00 — 准备接收（RR）。
- 01 — 拒绝（REJ）。
- 10 — 未准备接收（RNR）。

X — 预留，值为 0。

M — 修正功能位。

LLC Information — LLC 数据或高层协议。

SNAP：子网络访问协议

子网络访问协议（SNAP）规范了在 IEEE802 网络上传输 IP 数据报的标准方法。换句话说，IP 数据报可以封装在 802.2 LLC，SNAP 数据链路层和 802.3、802.4 或 802.5 网络物理层中，然后在 IEEE802 网络上发送。

SNAP 包含于逻辑链路控制（LLC IEEE 802.2）协议头中，主要用来在 IEEE 802 网络上封装 IP 数据包、地址解析协议（ARP）的请求和答复。SNAP 协议头位于 LLC 协议头后并且包含了组织代码，该组织代码表示接下来 16 位的以太类（EtherType）代码。通常情况下，人们采用 802.2 类型 1 实现所有通信过程。但同样位于 IEEE 802 网络的系统（Consenting systems）的两信点 在经过检验后都支持可以使用 802.2 类型 2，该过程可以通过 802.2 XID 机制实现。但目前仍然推荐使用类型 1 方案而且所有实施必须支持该服务类型。

通过地址解析协议(ARP)的动态发现过程,可以将 32 位 Internet 地址映射为 16 位或 48 位 IEEE 802 地址。IEEE 802 网络具有 16 位或 48 位物理地址。SNAP 中可以使用任意一种。

在 SNAP 中，IP 数据报的传输并不依赖于下层 LAN 技术（各种以太网和令牌环网类型）的传输速率，它们具有各种不同的传输速率（从 1 Mbps 到 20 Mbps）。

协议结构

LLC 头结构：

8	16	24 or 32 bit

DSAP	SSAP	Control

有关 LLC 头结构具体细节，请参照 LLC 页面。

SNAP 头结构：

24	40 bit
Organization Code	EtherType

当前为 SNAP 协议时，包含在 LLC 头结构中的 DSAP 和 SSAP 各字段值为 170（十进制），控制（Control）字段值为 3（无编号信息）。

Organization Code — 设置为 0。

EtherType — 规定封装在 IEEE 802 网络中的协议：IP = 2048，ARP = 2054。

STP：生成树协议 —IEEE 802.1D

生成树协议（Spanning Tree）定义在 IEEE 802.1D 中，是一种链路管理协议，它为网络提供路径冗余同时防止产生环路。为使以太网更好地工作，两个工作站之间只能有一条活动路径。网络环路的发生有多种原因，最常见的一种是有意生成的冗余 — 万一一个链路或交换机失败，会有另一个链路或交换机替代。

STP 允许网桥之间相互通信以发现网络物理环路。该协议定义了一种算法，网桥能够使用它创建无环路（loop-free）的逻辑拓朴结构。换句话说，STP 创建了一个由无环路树叶和树枝构成的树结构，其跨越了整个第二层网络。

生成树协议操作对终端站透明，也就是说，终端站并不知道它们自己是否连接在单个局域网段或多网段中。当有两个网桥同时连接相同的计算机网段时，生成树协议可以允许两网桥之间相互交换信息，这样只需要其中一个网桥处理两台计算机之间发送的信息。

网桥之间通过桥接协议数据单元（Bridge Protocol Data Unit — BPDU）交换各自状态信息。生成树协议通过发送 BPDU 信息选出网络中根交换机和根节点端口，并为每个网段（switched segment）选出根节点端口和指定端口。

网桥中的程序能够决定如何使用生成树协议，这称为生成树算法，该算法能够避免网桥环路，并确

保在多路径情形下网桥能够选择一条最有效的路径。如果最佳路径失败，可以使用该算法重新计算网络路径并找出下一条最佳路径。

利用生成树算法可以决定网络（哪台计算机主机在哪个区段），并通过 BPDU 信息交换以上数据。该过程主要分为以下两个步骤：

- 步骤 1：通过评估它所接收到的所有配置信息和选择最优选项，来决定一个网桥可发送的最佳信息。
- 步骤 2：一旦选定某网桥发送的信息，网桥将该信息与来自无根（non-root）连接的可能配置信息相比较。如果步骤 1 中选择的最佳选项并不优于可能配置信息，便删除该端口。

协议结构

网桥协议数据单元（BPDU）：

Protocol ID (2)	Version (1)	Type (1)	Flags (1)	Rood ID (8)	Root Path (4)
Sender BID (8)	Port ID (2)	M-Age (2)	Max Age (2)	Hello (2)	FD (2 Bytes)

- Protocol ID — 恒为 0。
- Version — 恒为 0。
- Type — 决定该帧中所包含的两种 BPDU 格式类型（配置 BPDU 或 TCN BPDU）。
- Flags — 标志活动拓朴中的变化，包含在拓朴变化通知（Topology Change Notifications）的下一部分中。
- Root BID — 包括有根网桥的网桥 ID。会聚后的网桥网络中，所有配置 BPDU 中的该字段都应该具有相同值（单个 VLAN）。NetXRay 可以细分为两个 BID 子字段：网桥优先级和网桥 MAC 地址。
- Root Path Cost — 通向有根网桥（Root Bridge）的所有链路的积累资本。
- Sender BID — 创建当前 BPDU 的网桥 BID。对于单交换机（单个 VLAN）发送的所有 BPDU 而言，该字段值都相同，而对于交换机与交换机之间发送的 BPDU 而言，该字段值不同）
- Port ID — 每个端口值都是唯一的。端口 1/1 值为 0×8001，而端口 1/2 值为 0×8002。
- Message Age — 记录 Root Bridge 生成当前 BPDU 起源信息的所消耗时间。
- Max Age — 保存 BPDU 的最长时间，也反映了拓朴变化通知（Topology Change Notification）过程中的网桥表生存时间情况。
- Hello Time — 指周期性配置 BPDU 间的时间。
- Forward Delay — 用于在 Listening 和 Learning 状态的时间，也反映了拓朴变化通知（Topology Change Notification）过程中的时间情况。

 MSN : MingLang.Xu@Hotmail.Com

ISL & DISL：思科交换机内链路协议和动态 ISL 协议

交换机内链路协议（ISL），是思科专有协议，主要用于当网络流量流经交换机和路由器时保持 VLAN 信息。

ISL 标签(tagging)能与 802.1Q 干线执行相同任务，只是所采用的帧格式不同。ISL 干线(trunk)是 Cisco 专有的，指两设备间（如交换机）的一条点对点连接线路。“交换机内链路协议”名称中即包含了这层含义。ISL 帧标签采用一种低延迟（low-latency）机制为单个物理路径上的多 VLAN 流量提供复用技术。ISL 主要用于实现交换机、路由器以及各节点（如服务器所使用的网络接口卡）之间的连接操作。为支持 ISL 功能特征，每台连接设备都必须配置采用 ISL 。配置有 ISL 的路由器支持 VLAN 间通信服务。无 ISL 配置的设备，接收由 ISL 封装的以太帧（Ethernet frame）后，会由于格式和大小的不同就将这些接收的帧归因于协议差错。

和 802.1Q 一样，ISL 作用于 OSI 模型第 2 层。所不同的是，ISL 协议头和协议尾封装了整个第 2 层的以太帧。正因为此，ISL 与协议无关，是一种能在交换机间传送任何类型的第 2 层数据帧或上层协议帧。ISL 所封装的帧可以是令牌环网（token ring）或快速以太网（Fast Ethernet），从发送端传输到接收端时保持不变。ISL 具有以下特征：

- 由专用集成电路执行（ASIC ： application-specific integrated circuit）
- 不干涉客户端；客户端不会看到 ISL 协议头
- 为交换机与交换机、路由器与交换机、交换机与带有 ISL 支持网卡的服务器之间的运行提供高效性能。

动态交换机内链路协议（DISL），也属于思科协议。它简化了两台相互连接的快速以太网设备上 ISL 干线的创建过程。快速以太通道技术是将两个全双工快速以太网链路聚合起来，用于高性能中枢连接。由于 DISL 中只需将一个链路终端配置为干线，所以 DISL 简单化了 VLAN 干线配置过程。

协议结构

ISL 头结构：

40	4	4	48	16	8	24	15	1	16	16 bits
DA	Type	User	SA	Len	AAA03	HSA	VLAN	BPDU	Index	Resv

- DA — 40 位组播目的地址。
- Type — 各种封装帧（Ethernet (0000)、Token Ring (0001)、FDDI (0010) 和 ATM (0011)）的 4 位描述符。

 MSN : MingLang.Xu@Hotmail.Com

- User — Type 字段使用的 4 位描述符扩展或定义 Ethernet 优先级。该二进制值从最低优先级开始 0 到最高优先级 3。
- SA — 传输 Catalyst 交换机中使用的 48 位源 MAC 地址。
- LEN — 16 位帧长描述符减去 DA、type、user、SA、LEN 和 CRC 字段。
- AAAA03 — 标准 SNAP 802.2 LLC 头。
- HAS — SA 的前 3 字节（厂商的 ID 或组织唯一 ID）。
- VLAN — 15 位 VLAN ID。低 10 位用于 1024 VLAN。
- BPDU — 1 位描述符，识别帧是否是生成树网桥协议数据单元（BPDU）。如果封装帧为思科发现协议（CDP）帧，也需设置该字段。
- INDEX — 16 位描述符，识别传输端口 ID。用于诊断差错。
- RES — 16 位预留字段，应用于其它信息，如令牌环和分布式光纤数据接口帧（FDDI），帧校验（FC）字段。

DTP：思科动态中继协议

思科动态中继协议（DTP），是 VLAN 协议组中思科专有协议，主要用于协商两台设备间链路上的中继及中继封装（如 802.1Q）类型。

中继协议有很多不同类型。如果一个端口可以成为 trunk 端口，那么该端口也可能具有自动中继功能，在某些情况下，甚至具有协商哪种中继类型的功能。这种与其它设备之间进行的协商中继方法的过程被称之为动态中继技术。

第一个问题是，中继电缆（trunk cable）两端最好都能理解它们是中继端口，否则它们将中继帧视为正常帧。终端工作站无法理解信息帧头里另外添加的标签信息，其驱动程序栈也无法识别该标签信息，从而导致终端系统锁定或当机。为解决这个问题，思科推出了用于交换机的协议以实现通信目的。推出的第一版本是 VTP，即 VLAN 中继协议，它与 ISL 共同工作。最新推出的版本，即动态中继协议（DTP），也可与 802.1q 共同工作。

其次是创建 LAN 。一个交换机的配置 VLAN ，需要做很多工作并且容易引起较多矛盾，如在一台交换机上 VLAN 100 属于工程部，而在另一台交换机上 VLAN100 可能被配置成属于财务部。这就使在故障排除工作中引起混乱，也会破坏精心设计的 VLAN 安全模式。该问题可通过 VTP/DTP 解决。在某台交换机上创建或删除一个 VLAN ，该信息自动传播到相同管理控制区域下的所有交换机上，这些交换机就是一个 VTP 域。

协议结构

关于基于 Catalyst 设置的交换机，其建立中继链路的语法如下所示：

set trunk mod_num/port_num [on | desirable | auto | nonegotiate] [isl | dot1q | negotiate] [vlan_range]

通过命令设置特定端口或中继端口。关键字的首次设置主要负责管理 DTP 模式：

Mode	模式
on	支持永久中继链路，即使相邻设备不同意。
off	支持永久不中继链路，即使相邻设备不同意。
desirable	引发端口成为中继，受相邻设备同意。
auto	引发端口被动转换中继。只有设置相邻设备或需要时端口才会中继。这是缺省模式。注意自动到自动（所有终端缺省）链路将不会成为中继链路。
nonegotiate	引发端口永久中继但不发送 DTP 帧。当 DTP 帧混淆邻近（非思科）802.1q 交换机时，使用该命令。必须手动将邻近交换机设置为中继。

关键字的第二次设置主要负责管理 VLAN 标签使用类型：ISL、802.1q 或协商使用。

VTP：思科 VLAN 中继协议

思科 VLAN 中继协议（VTP）是思科第 2 层信息传送协议，主要控制网络内 VLAN 的添加、删除和重命名。VTP 减少了交换网络中的管理工作。用户在 VTP 服务器上配置新的 VLAN，该 VLAN 信息就会分发到所有交换机，这样可以避免到处配置相同的 VLAN 。VTP 是思科专有协议，它支持大多数的 Cisco Catalyst 系列产品。

通过 VTP，其域内的所有交换机都清楚所有的 VLAN 情况。然而 VTP 会产生不必要的网络流量。这时，所有未知的单播和广播在整个 VLAN 内进行扩散，使得网络中的所有交换机接收到所有广播，即使 VLAN 中没有几个连接用户，情况也不例外。而 VTP pruning 技术正可以消除该多余流量。

缺省方式下，所有 Cisco Catalyst 交换机都被配置为 VTP 服务器。这种情形适用于 VLAN 信息量小且易存储于任意交换机 （NVRAM）上的小型网络。对于大型网络，由于每台交换机都会进行 NVRAM 存储操作，但该操作对于某些点是多余的，所以在这些点必须设置一个“判决呼叫”（judgment call）。基于此，网络管理员所使用的 VTP 服务器应该采用配置较好的交换机，其它交换机则作为客户机使用。选择作为 VTP 服务器的交换机数量应能提供网络所需的冗余。

 MSN : MingLang.Xu@Hotmail.Com

到目前为止，VTP 具有三种版本。其中 VTPv2 与 VTPv1 区别不大，主要区别在于：VTPv2 支持令牌环 VLAN，而 VTPv1 不支持。通常只有在使用 Token Ring VLAN 时，才会使用到 VTPv2，否则一般情况下并不使用 VTPv2。

VTPv3 不能直接处理 VLAN 事务，它只负责管理域（administrative domain）内不透明数据库的分配任务。与前两版相比，VTPv3 具有以下改进：

- 支持扩展 VLAN 。
- 支持专用 VLAN 的创建和通告。
- 改进的服务器认证性能。
- 避免“错误”数据库进入 VTP 域。
- 与 VTP v1 和 VTP v2 交互作用。
- 支持每端口（on a per-port basis）配置。
- 支持传播 VLAN 数据库和其它数据库类型。

协议结构

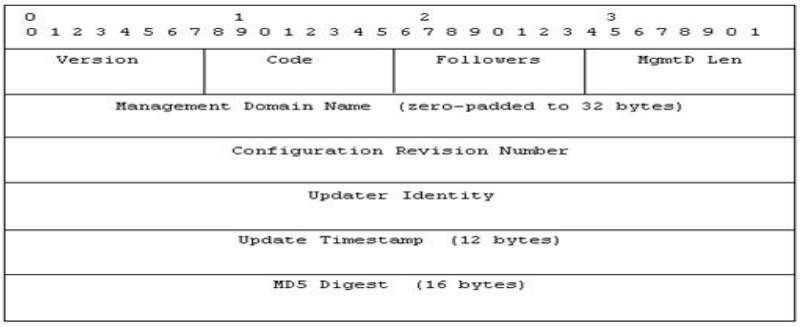
VTP 头结构格式可以改变，这主要取决于 VTP 信息类型。但是它们都包括以下字段：

- VTP 协议版本：1、2 或 3。
- VTP 信息类型：
  - Summary advertisements
  - Subset advertisement
  - Advertisement requests
  - VTP join messages
- 管理域大小
- 管理域名称

Summary Advertisements

当交换机接收到一个 summary advertisement 数据包时，它将该数据包的 VTP 域名称与其自己的 VTP 域名称相比。如果名称不同，那么该交换机忽略该数据包。如果名称相同，那么再比较两者的配置修订（configuration revision）。如果该交换机自己的配置修订高于或等于发送的 summary advertisement 数据包的修订，那么忽略该数据包。反之，就发送一个广告请求（advertisement request）。

Summary Advert Packet Format:



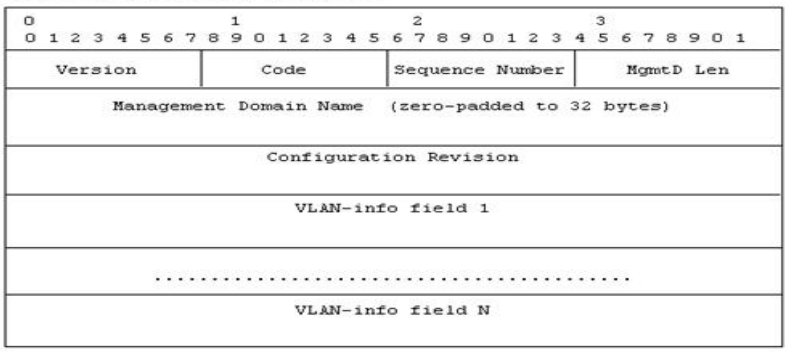
Summary Advert Packet Format

- Followers 表示该数据包后面跟随一个 Subset Advertisement 数据包。
- updater identity 表示最后一个增加配置修订的交换机的 IP 地址。
- Update timestamp 指配置修订的最后增量的日期和时间。
- Message Digest 5 (MD5)，在配置密码的情况下，用于传送 VTP 密码；还用于认证 VTP 更新的有效性。

Subset Advertisements

当交换机中需要添加、删除或改变一个 VLAN 时，发生改变的服务器交换机会增加配置修订并发送出一个 summary advertisement，其后跟随的是一个或多个 subset advertisement。一个 subset advertisement 包括一列 VLAN 信息。如果有多个 VLAN，那么需要提供更多 subset advertisement 来广告所有 VLAN。

Subset Advert Packet Format:



Summary Advert Packet Format

下面的表格表示每个不同的 VLAN 信息字段包含的信息（ISL VLAN ID 值最小最先发生）：

V-info-len	Status	VLAN-Type	VLAN-name Len
ISL VLAN-id		MTU Size	
802.10 index			
VLAN-name (padded with zeros to multiple of 4 bytes)			

不同的 VLAN 信息字段包含的信息

该数据包中的大多数字段比较容易理解。主要说明以下两点：

- Code — ubset advertisement 中的该字段值为 0x02。
- Sequence number — 指 summary advertisement 后的数据包流中的数据包序列。序列号从 1 开始。

Advertisement Requests

在以下情况下，交换机需要 VTP advertisement request：

- 交换机被重新设置；
- VTP 域名被改变；
- 交换机接收到高于自己配置修订的 VTP summary advertisement。

一旦接收到 advertisement request，VTP 设备便发送一个 summary advertisement，接着是一个或多个 subset advertisement。实例如下：

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1			
Version	Code	Rsvd	MgmtD Len
Management Domain Name (zero-padded to 32 bytes)			
Start-Value			

Advertisement Requests

- Code — advertisement request 中的该字段值为 0x03。
- Start Value — 适用于有多个 subset advertisement 的情况。如果已经接收到第一个（N）subset advertisement，而其后的一个（N+1）尚未接收到时，Catalyst 只发送来自第（N+1）个 subset advertisement 的 advertisement 请求。

WAN

PPPoE：以太网上的 PPP

PPPOE 使得一个网络上的计算机可以通过简单桥接访问设备连接到远端接入设备。在这个模型下，每个用户主机利用自身的 ppp 堆栈，并且用户使用熟悉的界面。访问控制、计费、服务类型等都可以针对每个用户来进行，而不是每个站点。

为了提供以太网上的点到点连接，每一个 PPP 会话必须知道远程通信对方的以太网地址，同时建立一个唯一的会话标识符。PPPoE 包含一个（以太网地址）发现协议来提供这个功能。

PPPoE 过程分为两个不同的阶段，即 Discovery （地址发现）阶段和 PPP 会话阶段。当某个主机希望发起一个 PPPoE 会话时，它必须首先执行 Discovery 来确定对方的以太网 MAC 地址并建立起一个 PPPoE 会话标识符（SESSION_ID）。虽然 PPP 定义的是端到端的对等关系，Discovery 却是一种客户端 - 服务器关系。在 Discovery 过程中，主机（作为客户端）发现某个访问集中器（Access Concentrator，作为服务器），根据网络的拓扑结构，可能主机能够发现多个访问集中器。 Discovery 阶段允许主机发现所有的访问集中器并从中选择一个。当 Discovery 阶段成功完成之后，主机和所选择的访问集中器两者都具备了用于在以太网上建立点到点连接所需的所有信息。

Discovery 阶段保持无状态（stateless）直到建立起一个 PPP 会话。一旦 PPP 会话建立，主机和访问集中器两者都必须为一个 PPP 虚拟接口分配资源。

协议结构

PPPoE 的以太网有效载荷显示如下：

4	8	16	32 bit
Ver	Type	Code	Session-ID
Length			Payload

- VER — PPPOE 版本。必须设置为 0x1。
- TYPE — 必须设置为 0x1。
- CODE — Discovery 和 PPP Session 阶段有定义。
- SESSION_ID — 无符号值。Discovery 数据包中有该字段定义。对于特定的 PPP Session 而言，该值为固定值。实际上，该字段定义了包括以太网 SOURCE_ADDR 和 DESTINATION_ADDR 的 PPP。0xffff 作为预留值，不作使用。

- LENGTH — 表示 PPPoE 有效载荷长。不包括以太网或 PPPoE 头的长度。

PPPoA：基于 ATM AAL5 的 PPP

PPPoA 使用 ATM 适配第 5 层（AAL5）分帧 PPP 封装的包。

PPP 为基于点对点连接的多协议数据包的传输提供了一个标准方法。

ATM AAL5 主要为连接到相同网络的终端站提供虚拟连接。这些连接提供了一个数据包发送服务，包括差错检测，但不包括差错修正。

目前大多数 PPP 使用 ISO 3309 HDLC 为其帧式化。

当 ATM 网络被配置为用于点对点连接时，PPP 就使用 AAL5 作为帧式化机制。

PPP 层将底层 ATM AAL5 层服务作为位同步点对点连接。在这种情况下，PPP 链路对应于 ATM AAL5 虚拟连接。该虚拟连接必须是全双工点对点连接方式，它可能是专用的，也可能是可交换的。基于 AAL5 上的 LLC 封装的 PPP 技术是多元 VC PPP 技术的另一种选择。

当在 AAL5 上传输一个 PPP 负载时，其具体实现如下：

1. 通过两终端的相互配置或协商，必须支持虚拟电路多元 PPP 负载，正如下面第五部分描述的一样。该技术也称为“多元 VC PPP”。

2. 通过两终端的相互配置或协商，必须支持 PVCs 上的 LLC 封装的 PPP 负载，如下面第六部分描述的一样。该技术也称为“LLC 封装 PPP”。

3. 为设置 SVC，必须通过 Q.2931 [9] 附件 C 实现协商过程，同时对宽带低层接口（B-LLI）信息元素进行编码，从而指出是多元 VC PPP 还是 LLC 封装 PPP 。

协议结构

AAL5 上的多元 VC PPP。AAL5 PDU 格式如下所示：

AAL5 CPCS-PDU 格式。

1 byte	0-47 bytes	1 byte	1 byte	2 bytes	4 bytes
CPCS-PDU	PAD	CPCS-UU	CPI	Length	CRC
			CPCS-PDU Trailer		

AAL5 CPCS-PDU 有效载荷字段编码如下：

- LLC Header 字段：2 字节，指定被路由的 OSI PDU（值为 0xFE 0xFE）的源 SAP（Source SAP）和目标 SAP（Destination SAP），其后是一个无编号信息（UI）的帧类型（Frame Type）（值为 0x03）。
- 表示 PPP（值为 0xCF）的网络层协议标识符（NLPID）。
- PPP 协议标识符字段，可以为 8 字节或 16 字节长。
- 后面是 PPP 信息（PPP Info）字段。

Destination SAP	Source SAP	Frame type	LLC Header
NLPID = PPP			
Protocol ID	PPP Info	Padding	PPP Payload
PAD (0 - 47 bytes)			
CPCS-UU	CPI	Length	CRC	CPCS-PDU Trailer

LCP：PPP 链路控制协议

LCP 用于就封装格式选项自动达成一致，处理数据包大小限制的变化，探测环路链路和其他普通的配置错误，以及终止链路。LCP 提供的其他可选功能有：认证链路中同等单元身份，和当链路功能正常或链路失败时的作出相应决定。PPP 中的 LCP 功能全面，适用于大多数环境。

LCP 包有 3 类：

1. 链路配置包，用于建立和配置链路（Configure-Request、Configure-Ack、Configure-Nak 和 Configure-Reject）。

2. 链路中止包被用于断开一个链路（Terminate-Request 和 Terminate-Ack）

3. 链路维护包被用于管理和调试一个链路（Code-Reject、Protocol-Reject、Echo-Request、Echo-Reply 和 Discard-Request）。

为了简单化，LCP 包里没有版本字段。一个正确运作的 LCP 将总是对带有可以简单识别的 LCP 包的未知协议和代码进行响应，因此需要为其他版本的实现提供一个确定性的可靠机制。

不管启用哪种配置选项，都得发送所有的 LCP 链路配置，链路终止和代码 – 拒绝包（代码 1 到 7），就像没有协商配置选项一样，而且每个配置选项都指定缺省值。这就保证了 LCP 包总可以被识别，甚至当链路的一个终端错误地认为该链路已经开放。

确切的说一个 LCP 包被封装在 PPP 信息字段中，该 PPP 协议字段表示类型为十六进制 c021（链路控制协议）。

协议结构

8	16	32 bit	variable
Code	Identifier	Length	Data

- Code — 十进制值，表示 LCP 数据包类型。
  - 1 — Configure-Request
  - 2 — Configure-Ack
  - 3 — Configure-Nak
  - 4 — Configure-Reject
  - 5 — Terminate-Request
  - 6 — Terminate-Ack
  - 7 — Code-Reject
  - 8 — Protocol-Reject
  - 9 — Echo-Request
  - 10 — Echo-Reply
  - 11 — Discard-Request
  - 12 — Link-Quality Repor
- Identifier — 十进制值，表示匹配 Request 和 Reply。
- Length — LCP 数据包长度，包括 Code、Identifier、Length 和 Data 字段。
- Data — 可变长字段，可能包括一或多个配置选项。

NCP：网络核心协议

网络核心协议(NCP)管理对 NetWare 服务器资源的访问。NCP 向 NetWare 文件共享协议(即 NFSP: NetWare File Sharing Protocol）发送过程调用消息，处理 NetWare 文件和打印资源请求。 NCP 是用于 NetWare 服务器和客户机之间传输信息的主要协议。

NCP 主要负责处理登入请求以及其它文件系统和打印系统请求。NCP 是一种基于客户机/服务器的

LAN 协议。工作站建立 NCP 请求并通过 IPX 在网络上发送这些请求服务。服务器端负责接收、拆包（unpacked）并解读 NCP 请求。

NCP 服务包括：文件访问、文件锁定（file locking）、安全性、资源分配跟踪（tracking of resource allocation）、事件通知（event notification）、与其它服务器同步、连接和通信、打印服务，以及队列和网络管理。

NCP 使用的是底层互联网分组交换协议 （即 IPX ： Internetwork Packet Exchange Layer Services）。目前许多最新版的 NetWare（继 NetWare 5.0 之后 ）也都支持 TCP/IP 协议。

协议结构

NCP 请求头格式如下所示：

8	16bit
Request type	
Sequence number	connection number low
Task number	connection number high
Request code	

- Request type — 识别数据包类型：

  1111H. 分配 slot 请求；

  2222H 文件服务器请求；

  3333H 文件服务器答复；

  5555H 取消分配 slot 请求；

  7777H 突发模式数据包 （BMP）；

  9999H 肯定确认；

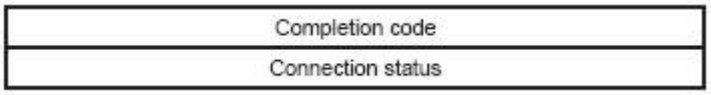
  H 表示十六进制符号。

- Sequence number — 工作站和文件服务器通过该字段识别发送和接收的数据包。
- Connection number low — 分配给工作站的低连接 ID 号 。
- Task number — 识别操作系统 ，如 DOS ， task 。
- Connection number high — 分配给工作站的高连接 ID 号 。只用于 1000 用户 NetWare 版本，其它版本上的该值都为 0 。

 MSN : MingLang.Xu@Hotmail.Com

- Request code — 识别特定请求功能代码。

NCP 答复头结构和请求头结构相同，但 Connection Number High 后的最后 2 字节不同 ，如下所示：



- Completion code — completion code 字段表示客户机请求是否成功。 Completion Code 字段值为 0 表示请求成功，否则表示请求出错。
- Connection status — 如果在 console prompt 处输入 DOWN ，那么 该字节中的第四位为 1，表示关闭服务器 。

SLIP：串行线路 IP

串行线路 IP （SLIP）用于运行 TCP/IP 的点对点串行连接。 SLIP 通常专门用于串行连接，有时候也用于拨号，使用的线路速率一般介于 1200bps 和 19.2Kbps 之间。 SLIP 允许主机和路由器混合连接通信（主机 - 主机、主机 - 路由器、路由器 - 路由器都是 SLIP 网络通用的配置），因而非常有用。

SLIP 只是一个包组帧协议，仅仅定义了在串行线路上将数据包封装成帧的一系列字符。它没有提供寻址、包类型标识、错误检查 / 修正或者压缩机制。

SLIP 定义了两个特殊字符：END 和 ESC 。END 是八进制 300（十进制 192），ESC 是八进制 333（十进制 219）。发送分组时， SLIP 主机只是简单地发送分组数据。如果数据中有一个字节与 END 字符的编码相同，就连续传输两个字节 ESC 和八进制 334 （十进制 220）。如果与 ESC 字符相同，就连续传输两个字节 ESC 和八进制 335（十进制 221 。当分组的最后一个字节发出后，再传送一个 END 字符。

因为没有“标准的” SLIP 规范，也就没有 SLIP 分组最大长度的实际定义。可能最好是接受 Berkeley UNIX SLIP 驱动程序使用的最大分组长度： 1006 字节，其中包括 IP 头和传输协议头（但不含分帧字符）。

压缩串行线路 IP （CSLIP）在传送出的 IP 分组上执行 Van Jacobson 头部压缩。这个压缩过程显著提高了交互式会话吞吐量。

如今，点对点协议（PPP）广泛替代了 SLIP ，因为它有更多特性和更灵活。

Frame Relay：帧中继

帧中继是一种局域网互联的 WAN 协议，它工作在 OSI 参考模型的物理层和数据链路层。它为跨越多个交换机和路由器的用户设备间的信息传输提供了快速和有效的方法。

帧中继是一种数据包交换技术，与 X.25 类似。它可以使终端站动态共享网络介质和可用带宽。帧中继采用以下两种数据包技术：1）可变长数据包；2）统计多元技术。它不能确保数据完整性，所以当出现网络拥塞现象时就会丢弃数据包。但在实际应用中，它仍然具有可靠的数据传输性能。

帧中继帧通过“虚电路”传输到其目的地， 帧中继的虚电路是源点到目的点的逻辑链路，它提供终端设备之间的双向通信路径，并由数据链路连接标识符（DLCI）唯一标识。 帧中继采用 复用技术，将大量虚电路复用为单一物理电路以实现跨网络传输。这种能力可以降低连接终端的设备和网络的复杂性。虚电路能够通过任意数量的位于帧中继数据包转换网络上的中间交换机。

帧中继网络提供的业务有两种：永久虚电路（PVC）和交换虚电路（SVC）。永久虚电路由网络管理器建立用来提供专用点对点连接；交换虚电路建立在呼叫到呼叫（call-by-call）的基础上，它采用与建立 ISDN 相同的信令。

由于其高带宽和高可靠性，在局域网互连中，帧中继可以作为专线和 X.25 网络的一个有吸引力的替代方案。

协议结构

帧中继（基于 LAPF Q.922）帧结构如下所示：

1 byte	2 bytes	Variable	2 bytes	1 byte
Flags	Address	Data	FCS	Flags

- Flags — 划定帧的起始和结束。该字段值不变，并表示为十六进制数 7E 或二进制数 01111110。
- Address — 包含以下信息：

6	7	8	12	13	14	15	16 bit
DLCI	C/R	E	DLCI	FECN	BECN	DE	EA

 MSN : MingLang.Xu@Hotmail.Com

- DLCI — 数据链路连接标识符字段表示帧地址并与 PVC 相对应。
- C/R — 指明帧是命令还是响应。
- EA — 扩展地址字段，表示帧中继头中附加的两个字节。
- FECN — 前向显式拥塞通知（参见下面的 ECN）。
- BECN — 后向显式拥塞通知（参见下面的 ECN）。
- DE — 丢弃指示。

- Data — 包括封装上层数据。可变长字段中的每帧包括一个用户数据，或者有效载荷字段长将变为 16,000 Octets。该字段通过帧中继网络用于传输高层协议数据包（PDU）。
- Frame Check Sequence — 确保传输数据的完整性。通过源设备计算该字段值，通过接收方校验该值以确保传输的完整性。

帧中继帧结构遵循于 LMI 规范，它由以下各字段构成：

1 byte	2 bytes	1 byte	1 byte	1 byte	1 byte
Flags	LMI DLCI	I-Indicator	Protocol Dis	Call Ref	M-Type
Information Elements (Variable)			FCS		Flags

- Flags — 划定帧的起始和结束。
- LMI DLCI — 帧被识别为 LMI 帧，替代基本帧中继帧。LMI 协会规范中的特定 LMI DLCI 值为 DLCI = 1023。
- Unnumbered Information Indicator — 将 Poll/Final 位设置为 0。
- Protocol Discriminator — 总包含一个代表 LMI 帧的值。
- Call Reference — 总包含 0。当前该字段不作任何使用。
- Message Type — 将帧标签为以下其中一种信息类型：
  - Status-Inquiry Message — 允许用户设备查询网络状态。
  - Status Message —响应 Status-Inquiry Messages 信息。Status Messages 包括 Keepalive 和 PVC Status Message 等信息。
- Information Elements — 包括个人信息元素（IE）的可变量。IE 由以下字段构成：
  - IE Identifier — 唯一识别 IE。
  - IE Length — 表示 IE 的长度。
  - Data — 由一个或多个字节构成，其中包括封装上层数据。
- Frame Check Sequence（FCS） — 确保传输数据的完整性。

ATM：异步传输模式

异步传输模式（ATM）在 ATM 参考模式下由一个协议集组成，用来建立一个在固定 53 字节的数据包(信元)流上传输所有通信流量的机制。固定大小的包可以确保快速且容易地实现交换和多路复用。ATM 是一种面向连接的技术，也就是说，两个网络系统要建立相互间的通信，需要通知中间介质服务需求和

流量参数。

ATM 参考模式分为三层：ATM 适配层（AAL）、ATM 层和物理层。 AAL 连接更高层协议到 ATM 层，其主要负责上层与 ATM 层交换 ATM 信元。当从上层收到信息后， AAL 将数据分割成 ATM 信元；当从 ATM 层收到信息后， AAL 必须重新组合数据形成一个上层能够辨识的格式，上述过程即称之为分段与重组（SAR）。不同的 AAL 用于支持在 ATM 网络上使用的不同的流量或服务类型。

ATM 层主要负责将信元从 AAL 转发给物理层便于传输和将信元从物理层转发给 AAL 便于其在终端系统的使用。 ATM 层能够决定进来的信元应该被转发至哪里；重新设置相应的连接标识符并且转发信元给下一个链接、缓冲信元以及处理各种流量管理功能，如信元丢失优先权标记、拥塞标注和通用流控制访问。此外 ATM 层还负责监控传输率和服从服务约定（流量策略）。

ATM 的物理层定义了位定时及其它特征，将数据编码并解码为适当的电波或光波形式，用于在特定物理媒体上传输和接收。此外它还提供了帧适配功能，包括信元描绘、信头错误校验（HEC）的生成和处理、性能监控以及不同传输格式的负载率匹配。物理层通常使用的介质有 SONET 、DS3 、光纤、双绞线等。

协议结构

ATM 信元格式：

<table>
<tr><td rowspan="5">HEADER</td><td colspan="2">GFC or VPI</td><td>VPI</td></tr>
<tr><td colspan="2">VPI</td><td>VCI</td></tr>
<tr><td colspan="3">VCI</td></tr>
<tr><td>VCI/td></td><td>PT（3 Bit）</td><td>CLP</td></tr>
<tr><td colspan="3">HEC</td></tr>
<tr><td>IE</td><td colspan="3">Cell Payload（48 Bytes）</td></tr>
</table>

- Header — （5 字节）通用流控制 VPI/VCI 和其它控制头。
- IE — （48 字节）信元有效载荷。

物理层规范说明 — 专用 UNI：

帧格式	比特率/线路速率	媒体
Cell Stream	25.6 Mbps/32 Mbaud	UTP-3
STS-1	51.84 Mbps	UTP-3
FDDI	100 Mbps/125 Mbaud	Multimode Fiber

 MSN : MingLang.Xu@Hotmail.Com

STS-3c STM-1	155.52 Mbps	UTP-5
STS-3c STM-1	155.52 Mbps	Single-Mode Fiber, Multimode Fiber, Coax pair
Cell Stream	155.52 Mbps/ 194.4Mbaud	Multimode Fiber, STP
STS-3c STM-1	155.52 Mbps	UTP-3
STS-12, STM-4	622.08 Mbps	SMF, MMF

物理层规范说明 — 公用 UNI：

帧格式	比特率/线路速率	媒体
DS1	1.544 Mbps	Twisted pair
DS3	44.736 Mbps	Coax pair
STS-3c, STM-1	155.520 Mbps	Single-mode Fiber
E1	2.048 Mbps	Twisted pair, Coax pair
E3	34.368 Mbps	Coax Pair
J2	6.312 Mbps	Coax Pair
N × T1	N × 1.544 Mbps	Twisted Pair

PPP：点对点协议

点对点协议（PPP）为在点对点连接上传输多协议数据包提供了一个标准方法。PPP 最初设计是为两个对等节点之间的 IP 流量传输提供一种封装协议。在 TCP-IP 协议集中它是一种用来同步调制连接的数据链路层协议（OSI 模式中的第二层），替代了原来非标准的第二层协议，即 SLIP。除了 IP 以外 PPP 还可以携带其它协议，包括 DECnet 和 Novell 的 Internet 网包交换（IPX）。

PPP 主要由以下几部分组成：

封装：一种封装多协议数据报的方法。PPP 封装提供了不同网络层协议同时在同一链路传输的多路复用技术。PPP 封装精心设计，能保持对大多数常用硬件的兼容性。

链路控制协议：PPP 提供的 LCP 功能全面，适用于大多数环境。LCP 用于就封装格式选项自动达成一致，处理数据包大小限制，探测环路链路和其他普通的配置错误，以及终止链路。LCP 提供的其他可选功能有：认证链路中对等单元的身份，决定链路功能正常或链路失败情况。

网络控制协议：一种扩展链路控制协议，用于建立、配置、测试和管理数据链路连接。

 MSN : MingLang.Xu@Hotmail.Com

配置：使用链路控制协议的简单和自制机制。该机制也应用于其它控制协议，例如：网络控制协议（NCP）。

为了建立点对点链路通信，PPP 链路的每一端，必须首先发送 LCP 包以便设定和测试数据链路。在链路建立，LCP 所需的可选功能被选定之后，PPP 必须发送 NCP 包以便选择和设定一个或更多的网络层协议。一旦每个被选择的网络层协议都被设定好了，来自每个网络层协议的数据报就能在链路上发送了。

链路将保持通信设定不变，直到有 LCP 和 NCP 数据包关闭链路，或者是发生一些外部事件的时候（如，休止状态的定时器期满或者网络管理员干涉）。

协议结构

8	16	24	40 bits	Variable…	16-32 bits
Flag	Address	Control	Protocol	Information	FCS

- Flag — 表示帧的起始或结束，由二进制序列 01111110 构成。
- Address — 包括二进制序列 11111111，标准广播地址（注意：PPP 不分配个人站地址）
- Control — 二进制序列 00000011，要求用户数据传输采用无序帧。
- Protocol — 识别帧的 Information 字段封装的协议。
- Information — 0 或更多八位字节，包含 Protocol 字段中指定的协议数据报。
- FCS — 帧校验序列(FCS)字段，通常为 16 位。PPP 的执行可以通过预先协议采用 32 位 FCS 来提高差错检测效果。

SDLC：同步数据链路控制

同步数据链路控制（SDLC）协议是一种 IBM 数据链路层协议，适用于系统网络体系结构（SNA）。

通过同步数据链路控制（SDLC）协议，数据链路层为特定通信网络提供了网络可寻址单元（NAUs：Network Addressable Units）间的数据差错释放（Error-Free）功能。信息流经过数据链路控制层由上

层往下传送至物理控制层。然后通过一些接口传送到通信链路。SDLC 支持各种链路类型和拓朴结构。应用于点对点和多点链接、有界（Bounded）和无界（Unbounded）媒体、半双工（Half-Duplex）和全双工（Full-Duplex）传输方式，以及电路交换网络和分组交换网络。

SDLC 支持识别两类网络节点：主节点（Primary）和次节点（Secondary）。主节点主要控制其它节点（称为次节点：Secondaries）的操作。主节点按照预先确定的顺序选择次节点，一旦选定的次节点已经导入数据，那么它即可进行传输。同时主节点可以建立和拆除链路，并在运行过程中控制这些链路。主节点支配次节点，也就是说，次节点只有在主节点授权前提下才可以向主节点发送信息。

SDLC 主节点和次节点可以在四种配置中建立连接：

- 点对点（Point-to-Point）：只包括两个节点：一个主节点，一个次节点。
- 多点（Multipoint）：包括一个主节点，多个次节点。
- 环（Loop）：包括一个环形拓朴：连接起始端为主节点，结束端为次节点。通过中间次节点相互之间传送信息以响应主节点请求。
- 集线前进（Hub Go-Ahead）：包括一个 Inbound 信道和一个 Outbound 信道。主节点使用 Outbound 信道与次节点进行通信。次节点使用 Inbound 信道与主节点进行通信。通过每个次节点，Inbound 信道以菊花链（Daisy-Chained）格式回到主节点。

为适应不同环境，SDLC 具有一些派生类：

- HDLC，一种 ISO 协议，适用于 x.25 网络；
- LAPB，一种 ITU-T 协议，适用于 ISDN 网络；
- LAPF，一种 ITU-T 协议，适用于帧中继（Frame Relay）网络；
- IEEE 802.2，通常指 LLC，具有三种类型，适用于局域网（Local Area Network）；
- QLLC，适用于在 X.25 网络上传输 SNA 数据。

协议结构

1 byte	1-2 bytes	1-2 bytes	Variable	2 bytes	1 byte
Flag	Address field	Control field	Data	FCS	Flag

- Flag — 启动和终止差错校验。
- Address — 包括次站 SDLC 地址，表明帧来自于主站还是次站。
- Control — 使用 3 种不同格式，取决于使用的 SDLC 帧类型：
  - Information（I）frame — 传递上层信息和一些控制信息。
  - Supervisory （S）frame — 提供控制信息。S 帧可以请求和挂起传输、报告状态、确认 I 帧接收。S 帧不包含信息帧（information field）。
  - Unnumbered （U）frame — 支持控制目标，无编号。U 帧用于启动次站。取决于 U 帧，其控制字段可能为 1 字节也可能为 2 字节。有些 U 帧包含信息字段。
- Data — 包含路径信息单元（PIU）或交换识别（XID）信息。

 MSN : MingLang.Xu@Hotmail.Com

- Frame check sequence （FCS)) — 优于结束标签分隔符，通常指循环冗余校验（CRC）计算余数。

HDLC：高级数据链路控制

高级数据链路控制（HDLC）协议是基于的一种数据链路层协议，促进传送到下一层的数据在传输过程中能够准确地被接收（也就是差错释放中没有任何损失并且序列正确）。HDLC 的另一个重要功能是流量控制，换句话说，一旦接收端收到数据，便能立即进行传输。HDLC 具有两种不同的实现方式：高级数据链路控制正常响应模式即 HDLC NRM（又称为 SDLC）和 HDLC 链路访问过程平衡（LAPB）。其中第二种使用更为普遍。HDLC 是 X.25 栈的一部分。

HDLC 是面向比特的同步通信协议，主要为全双工点对点操作提供完整的数据透明度。它支持对等链路，表现在每个链路终端都不具有永久性管理站的功能。另一方面，HDLC NRM 具有一个永久基站以及一个或多个次站。

HDLC LAPB 是一种高效协议，为确保流量控制、差错监测和恢复它要求额外开销最小。如果数据在两个方向上（全双工）相互传输，数据帧本身就会传送所需的信息从而确保数据完整性。

帧窗口是用于在接收第一个帧已经正确收到的确认之前发送复帧。这就意味着在具有长“turn-around”时间滞后的情况下数据能够继续传送，而不需要停下来等待响应。例如在卫星通信中会发生这种情形。

通常，帧分为三种类型：

- 信息帧：在链路上传送数据，并封装 OSI 体系的高层；
- 管理帧：用于实现流量控制和差错恢复功能；
- 无编号帧：提供链路的初始化和终止操作。

协议结构

1 byte	1-2 bytes	1 byte	variable	2 bytes	1 byte
Flag	Address field	Control field	Information	FCS	Flag

- Flag — 该字段值恒为 0x7E。
- Address Field — 定义发送帧的次站地址，或基站发送帧的目的地。该字段包括服务访问点（6 比特）、命令/响应位（表示帧是否与节点发送的信息帧有关或帧是否被节点接收）、地址扩展

位（通常设置为 1 字节长）。当设置错误时，表示一个附加字节。

- Extended Address — HDLC 为基本格式提供了另一种扩展。通过多方协定，Address Field 可以被扩展为多个字节。
- Control Field — 识别帧类型。另外，根据帧类型划分，该字段还包括序列号、控制特性和差错跟踪。
- FCS — 帧校验序列（FCS）字段通过许可传输帧数据的完整性，使高层物理差错控制可以被校验。

LAP-D：ISDN 链路访问协议（D 信道）

LAP-D 是 ISDN 协议集中的第二层协议，与 X.25 LAP-B 协议几乎相同。ISDN 的三种逻辑数字通信信道执行如下功能：

- B 信道 — 传送用户业务信息包括数字数据、视频和语音；
- D 信道 — 在用户和网络间传送信令和数据包；
- H 信道 — 执行与 B 信道相同的功能但其运行速率超过 DS-0 （64 Kbps）。

LAP-D 完成 ISDN 链路建立过程具体如下：

- 终端端点（TE）和网络交换准备接受帧（RR），等待初始化一个连接。
- 终端端点（TE）发送一个未编号的信息（UI）帧，其中 SAPI（业务接入点标识）的值为 63（管理协议、查询网络），TEI （终端终点标识） 的值为 127（建立广播链路）。
- 网络分配一个可使用 TEI（范围为 64 — 126 ）。
- TE 发送一个 SABME 帧，其 SAPI 值为 0（呼叫控制），TEI 值为网络分配的值。
- 网络发出未编号响应（UA）帧，SAPI=0，TEI 值为分配值 。

LAPD 由 CCITT Q.920/921 定义，其工作于平衡式异步模式（ABM）下，该模式是完全平衡的（也就是没有主从关系）。任何时候每个站都可以进行初始化、监督、错误恢复和发送帧操作。 DTE 和 DCE 在该协议中是同等的。

协议结构

标准 LAPD 帧格式如下所示：

Flag	Address Field	Control Field	Information	FCS	Flag

Flag — Flag 字段的值恒为 0x7E。采用“Bit Stuffing”技术以确保帧分隔符标志的位模式不会

出现在帧的数据字段。

Address Field — 帧中头标志后面的起始两个字节为 Address Field。Address Field 格式如下所示：

1	2	3	4	5	6	7	8
SAPI (6 bits)						C/R	EA0
TEI (7 bits)							EA1

- SAPI（服务访问点标识符），6 位（如下所示）；
- C/R（命令/响应）位表示帧是命令还是响应；
- EA0（地址字段扩展）位表示是否是地址最后八位字节；
- TEI（终端终点标识符），7 位设备标识符（如下所示）；
- EA1（地址字段扩展）位，与 EA0 相同。

Control Field — Control Field 字段在 Address Field 字段后面，用于识别帧类型。另外，它还包括序列号、控制特征和差错追踪，这主要取决于帧类型。下面提供了 LAPD 中定义的管理帧类型：

RR	信息帧确认和指示以接收更多信息。
REJ	请求重发指定序列号后面的所有帧。
RNR	表示临时占有站的状态（如全窗口）。

LAPD 支持的一些未编号的帧类型有 DISC（请求断开）、UA（确认帧）、DM （DISC 的应答，表示断开模式）、FRMR （帧拒绝）、SABM、SABME、UI 和 XID。

FCS — 通过检查所传输帧数据的完整性， 帧校验序列（FCS）可以进行高级别的物理差错控制。首先，传输方使用基于帧中所有位的值的算法得出序列号。然后接收方对接收到的帧采用相同的算法，将得到的值与 CRC 作比较。

Window size — LAPD 支持扩展的窗口大小（模数为 128），确认帧的可能值是从 8 扩展到 128。这种扩展值通常应用于确认时延比帧传输时间大得多的卫星传输系统中。链路初始帧类型决定会话模数，并且在基本帧类型名称中增加了“E”（如由 SABM 变为 SABME）。

BISDN：宽带综合业务数字网

宽带综合业务数字网（BISDN 或者 Broadband ISDN）处理高宽带应用程序。当前，BISDN 采用基于

SONET 传输电路的 ATM 技术提供从 155Mbps 到 622 Mbps 及以上的数据传输率。这一点与窄带 ISDN（或 N-ISDN）形成明显的对比，窄带 ISDN 只能支持从 64 kps 到最大 2 Mbps 的数据率。

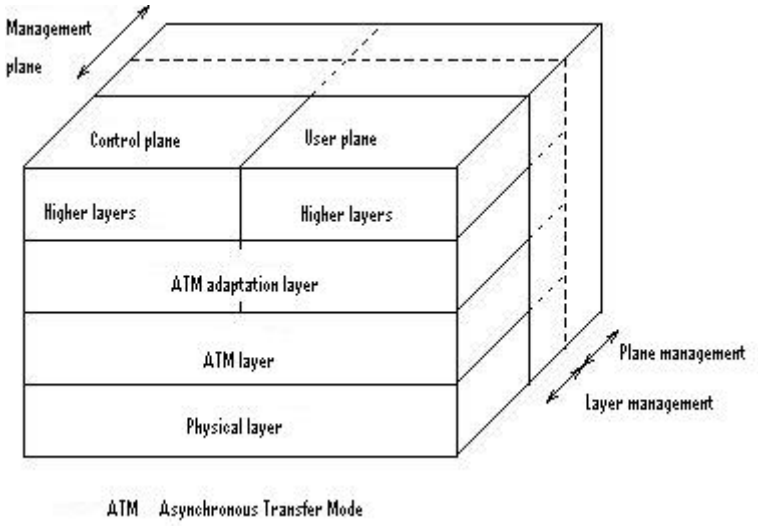
BISDN 所支持的服务主要有以下几类：

- 对话服务，诸如电话类服务，窄带 ISDN 也支持该服务。由于带宽的增加，BISDN 支持视频电话、视频会议以及一些高速数据传输服务。
- 信息传送服务：即指存储－转发（ store-and-forward ）服务。应用程序包括：语音和视频邮件、多媒体邮件和传统的电子邮件等。
- 检索服务：提供对信息存储器的访问，按照用户要求发送访问信息。
- 无用户控制表示：例如 TV，用户可以随意选择看和不看。
- 用户控制表示：应用于广播信息，用户可以部分控制。

B-ISDN 既支持面向连接服务也支持无连接服务。在两种情况下，都是通过异步传输模式（ATM），使用端对端逻辑连接或虚拟电路实现宽带信息的传输。B-ISDN 使用的是带外（out-of-band）信令（与 N-ISDN 相同）。N-ISDN 中采用 D 信道实现信令传输，而 B-ISDN 采用的是一种特殊虚拟电路信道实现该过程。但到目前为止，B-ISDN 尚未得到普遍应用。

协议结构

宽带 ISDN 协议参考模型基于 ATM 参考模型：



ATM 参考模型

ATM 适配层（AAL）。该层负责将 ATM 提供的服务映射给高层需要的服务上。它具有两个子层：

- ATM 层，该层独立于传输需要的物理媒体。具有以下功能：通用流控制 （GFC）功能、信元头生成和提取、信元复用和解复用。
- 物理层，由两个子层构成：传输会聚（TC）和物理媒体（PM） 。

管理面有两种功能用于完成层管理和面管理。面管理不像其它层一样采用分层方法，这是因为面管理需要通过系统各方面信息为整个系统提供管理设施。层管理为每个独立层中的协议实体提供信息和控制设施，包括每层的运行和维护（OAM）功能。

控制面负责连接的监督和管理，包括呼叫建立、呼叫释放和维护。

用户面提供用户信息传输。也包括执行差错恢复、流控制功能等机制。

Q.931：信令网络层协议

Q.931，作为电信体系的网络层协议，主要为 ISDN 提供呼叫建立及维护和终止两设备间的逻辑网络连接。Q.931 是电信体系网络层（第三层）协议之一，由 ITU Q 系列 Q.930-931 文件详细说明。

在第三层呼叫建立期间，有三方参与发送和接收信息：

- 呼叫方，
- ISDN 交换机，
- 接收方。

下面是一个关于呼叫设置步骤的例子：

1. 呼叫方发送一个建立呼叫信息（SETUP）给交换机；
2. 如果 SETUP 通过，交换机发送一个 CALL PROCeeding 信息给呼叫方，并发送一个 SETUP 信息给接收方；
3. 接收方收到 SETUP 。如果该信息正常，它就振铃电话并发送一个 ALERTING 信息给交换机；
4. 交换机转发该 ALERTING 信息给呼叫方；
5. 当接收方应答呼叫后，就发送一个 CONNECT 信息给交换机；
6. 交换机转发该 CONNECT 信息给呼叫方；
7. 呼叫方发送一个 CONNECT 响应信息给交换机；
8. 交换机转发该 CONNECT 响应信息给接收方；
9. 呼叫方发送 CONNECT ACKnowledge 信息到交换机；
10. 交换机转发该信息到接受方；
11. 完成。连接建立成功。

ISDN 设备具有的服务和特征在可选字段 — 业务预置文件 ID （SPID）中规定，但它们只能在呼叫建立之前的设备初始化时期被访问。SPID 的一般格式是 ISDN 线路的 10 位数字电话号码，并具有前缀和后缀以识别在线特征，但其格式也可由电信公司决定。

协议结构

Information Field Structure — Information Field 是可变长字段，包括 Q.931 协议数据：

1	2	3	4	5	6	7	8
Protocol Discriminator							
0	0	0	0	Length of CRV			
Call Reference Value (1 or 2 octets)							
0	Message Type						
Mandatory & Optional Information Elements (variable)							

- Protocol Discriminator （1 octet） — 识别第 3 层协议。如果是 Q.931 头，该值恒为 0816。
- Length （1 octet） — 表示下一字段即 CRV 的长度。
- Call Reference Value （CRV）（1 或者 2 octet） — 唯一地识别用户网络接口上的每个呼叫。在呼叫开始时分配该字段值。当该呼叫清除后，该字段值可以用于其它呼叫过程。
- Message Type （1 octet） — 识别信息类型（也就是 SETUP、CONNECT 等）。该字段决定需要并许可哪些其它信息。
- Mandatory and Optional Information Elements （variable length） — 可选项，主要取决于 Message Type。

LAPB：链路访问过程平衡

（LAPB：Link Access Procedure Balanced for x.25）

链路访问过程平衡（LAPB）是数据链路层协议，负责管理在 X.25 中 DTE 设备与 DCE 设备之间的通信和数据包帧的组织过程。LAPB 是源于 HDLC 的一种面向位的协议，它实际上是 BAC （平衡的异步方式类别）方式下的 HDLC。LAPB 能够确保传输帧的无差错和正确排序。

LAPB 与 SDLC 和 HDLC 共享相同的帧格式、帧类型和字段功能，但与后两者不同的是，LAPB 受 ABM 传输模式的限制且只适用于组合站。LAPB 电路可由 DTE 或 DCE 建立。启动呼叫的站称为主站，响应的另一站称为次站。此外 LAPB 所使用的 P/F 比特位其它协议不同。

在 LAPB 中，由于没有主从关系，发送端使用 Poll 比特位来要求立即响应。在响应帧中，这个比特位变成接收端的 Final 比特位。接收端总是打开 Final 比特位去响应来自发送端 Poll 比特位的命令。由于确认响应可能会丢失并导致任何一端无法确保帧是否正确排序，就会采用 P/F 比特位，同时需要重建参考点。

LAPB 帧类型：

- 信息帧（I- 帧）传送高层协议信息和一些控制信息，主要功能是排序、控制流量、错误监测及恢复，它携带发送和接收序号。
- 监控帧（S- 帧）传送控制信息，主要功能是请求和挂起传输、报告状态信息及确认接收到 I- 帧，它只携带接收序号。
- 非数字帧（U- 帧）携带控制信息，主要功能是建立和终止链路以及报告错误，它不携带序号。

协议结构

LAPB 帧格式如下：

1 byte	1 byte	1-2 bytes	Variable	2 bytes	1 byte
Flag	Address field	Control field	Information	FCS	Flag

- Flag — 该字段值恒为 0x7E。为确保帧分隔符标志的位模式（Bit Pattern）不出现在帧的数据字段，通常在发送方和接收方利用 Bit Stuffing 技术。.
- Address Field — 在 LAPB 中，由于协议工作在点对点模式下，所有 Address Field 没有实际意义。DTE 网络地址在第三层数据包中由描述。
- Control Field — 识别帧类型。另外，根据帧类型划分，该字段还包括序列号、控制特性和差错跟踪。
- Modes of Operation — LAPB 工作于异步平衡模式（ABM）。该模式完全平衡（也就是说没有主/从关系）且采用 SABM（E）帧格式表示。任何时候各站都有可能进行初始化、监督管理、差错恢复及发送帧等操作。DTE 和 DCE 一律同等对待。
- FCS — 帧校验序列（FCS）字段通过许可传输帧数据的完整性，使高层物理差错控制可以被校验。
- Window Size — LAPB 支持扩展窗口大小（模数为 128），确认帧的大小可能从 8 扩展到 128。

 MSN : MingLang.Xu@Hotmail.Com

ISDN：综合业务数字网

综合业务数字网（ISDN）是一种数字化电话连接系统。几十年来，电话通信使用的一直是模拟连接方式，而 ISDN 是第一部定义数字化通信的协议，该协议支持标准线路上的语音、数据、视频、图形等的高速传输服务。ISDN 的承载信道（B 信道）负责同时传送各种媒体，占用带宽为 64 kb/s （有些交换机将带宽限制为 56 kb/s）。数据信道（D 信道）主要负责处理信令，传输速率从 16 kb/s 到 64 kb/s 不定，这主要取决于服务类型。ISDN 不仅限于公共电话网络，其传输还可以通过分组交换网络、电报网或有线电视网络等完成。ISDN 有两种基本服务类型：

- 基本速率接口（BRI：Basic Rate Interface）：由两个 64 kb/s 的 B 信道和一个 16 kb/s 的 D 信道构成，总速率为 144 kb/s。该服务主要适用于个人计算机用户。Telco 提供的 U 接口的 BRI 支持双线、传输速率为 160 kb/s 的数字连接。通过回波消除操作降低噪音影响。各种数据编码方式（北美使用 2B1Q，欧洲国家使用 4B3T）可以为单线本地环路提供更高的数据传输率。
- 主要速率接口（PRI：Primary Rate Interface）：能够满足用户的更高要求。典型的 PRI 由 23 个 B 信道和一个 64kb/s 的 D 信道构成，总速率为 1536 kb/s。在欧洲，PRI 由 30 个 B 信道和一个 64kb/s 的 D 信道构成，总速率为 1984kb/s。通过 NFAS（Non-Facility Associated Signaling），PRI 也支持具有一个 64 kb/s D 信道的多 PRI 线路。

CCITT（现在为 ITU-T）研究组于 1984 年首次推出了一组 ISDN 推荐协议。在此之前，各个地区都有各自的 ISDN 版本。通过代码设置机制可以启用国家专用信息单元（nation-specific information elements）以支持各个地区使用数据结构内自己的信息单元。其中通用的一种国家 ISDN 专用单元为 National ISDN，由 Bellcore 提出，适用于美国。National ISDN 中包含四种专用网络信息类型，它不包含任何单个八位字节信息单元。与其它信息单元相比，National ISDN 中增加了 SEGMENT、FACILITY 和 REGISTER 信息类型以及分段信息和扩展功效信息单元，此外也改变了一些字段值意思，并增加了一些新字段值。

受带宽和服务的限制，传统的 ISDN 被称之为窄带 ISDN，与如今的宽带 ISDN（BISDN）相对。

协议结构

ISDN 帧的通用结构如下：

<table>
<tr><th>8</th><th>7</th><th>6</th><th>5</th><th>4</th><th>3</th><th>2</th><th>1</th></tr>
<tr><td colspan="8">Protocol discriminator</td></tr>
<tr><td>0</td><td>0</td><td>0</td><td>0</td><td colspan="4">Length of Reference cCall Value</td></tr>
<tr><td>Flag</td><td colspan="7">Call Reference Value</td></tr>
<tr><td>0</td><td colspan="7">Message Type</td></tr>
<tr><td colspan="8">Other Information Elements as Required</td></tr>
</table>

 MSN : MingLang.Xu@Hotmail.Com

- Protocol Discriminator — 它用于将用户和网络呼叫控制信息与本建议范围内的其它消息相区别。
- Length of Call Reference Value — 规定下一字段的长度。Call Reference 可能为 8 到 16 个字节长，这取决于正在编码的值大小。
- Flag — 由分配 Call Reference Value 的一方发送信息时，该字段设置为 0，否则为 1。
- Call Reference Value — 在指定会话期间分配的专用值。用于识别呼叫维护设备和 ISDN 交换机间的呼叫。
- Message Type — 定义帧主要功能目标。Message Type 可以为 8 字节或 16 字节（网络特定信息）。当大于 8 字节时，第一个 8 字节采用 8 个 0 进行编码。Message Type 的完整列单请见下面的 ISDN Message Types。
- ISDN Information Elements — 包括两种信息元素：单八位字节和可变长。

单八位字节的信息元素如下所示：

8	7	6	5	4	3	2	1
1	Information Element Identifier				Information Element		

- Variable Length Information Elements — 可变长信息元素的格式如下：

8	7	6	5	4	3	2	1
0	Information Element Identifier						
Length of Information Elements							
Information Elements (Multiple bytes)							

Information Element Identifier 用于识别已选择元素，并且在特定的 Codeset 中是唯一的。

Length of the Information Element 字段使得接收方知道每个信息元素后八位字节的数量。

- ISDN Message Types — 可能的 ISDN 信息类型有：呼叫建立、呼叫信息阶段、呼叫清除和混合。
- Codeset — 定义了三种主要的 Codeset。在每个 Codeset 中，通过关联协议变量定义了部分 Information Elements 字段。

Codeset 0	The default code, referring to the CCITT set of information elements.

Codeset 5	The national specific Codeset.
Codeset 6	The network specific Codeset.

CPE — 用户基地设备。涉及连接在用户站点的所有 ISDN 可兼容设备。该设备的具体例子包括电话、PC、电报、传真等。NT1 中的 FCC 是个例外。FCC 将 NT1 视为一种 CPE，这是因为 FCC 处于用户站点，而 CCITT 将 NT1 视为网络的一部分。因此，网络边界的网络参考点主要取决于使用的变量类型。

ISDN Channels B、D 和 H — ISDN 中的三种逻辑数字通信信道，各执行以下功能：

B 信道	传送用户服务信息，包括数字数据、视频和语音。
D 信道	传送用户和网络间的信号和数据包。
H 信道	与 B 信道执行相同的功能，但运行速率超过 DS-0（64 Kbps）。具体执行为 H0（384 kb/s；6 B 信道），H10（1472 kb/s； 23 B 信道），H11（1536 kb/s；24 B 信道）以及 H12（1920 kb/s；只用于国际 E1）。

X.25：分组交换

X.25 是 ISO 和 ITU-T 为广域网（WAN）通信所建议的一种包交换数据网络协议，它定义数据终端设备（DTE）和数据电路终端设备（DCE）之间的数据以及控制信息的交换。

无论连接到网络的系统类型是什么，X.25 都具有高效的使用性能。X.25 通常用于分组交换网络上，如电话行业。它是根据订户使用的网络进行收费。X.25 是面向连接的业务从而确保数据包的顺序传输。

当一台 DTE 设备向另一台 DTE 发出通信会话连接请求时，就建立 X.25 会话。接收请求的 DTE 设备端可以同意也可以拒绝该连接。如果同意请求，那么两个系统便开始进行全双工通信传输；任意一台 DTE 设备可以终止该连接。一旦会话终止，任何后续的通信都需要建立一个新会话。X.25 采用虚电路数据包通信方式，可使用交换虚电路和永久虚电路。

X.25 协议集有三层，与 OSI 协议栈的底三层相关联。

物理层：描述物理环境接口。该组包括三种协议： 1）X.21 接口运行于 8 个交换电路上；2）X.21bis 定义模拟接口，允许模拟电路访问数字电路交换网络；3）V.24 使得 DTE 能在租用模拟电路上运行以连接到包交换结点或集中器。

链路层：负责 DTE 和 DCE 之间的可靠通信传输。包括四种协议：1）LAPB 源自 HDLC，具有 HDLC 的所有特征，使用较为普遍，能够形成逻辑链路连接。2）链路访问协议（LAP）是 LAPB 协议的前身，如今几乎不被使用；3）LAPD 源自 LAPB，用于 ISDN，在 D 信道上完成 DTE 之间，特别是 DTE 和 ISDN 节

点之间的数据传输；4）逻辑链路控制（LLC）一种 IEEE 802 LAN 协议，使得 X.25 数据包能在 LAN 信道上传输。

分组层（PLP）协议：描述网络层（第三层）中分组交换网络的数据传输协议。PLP 负责虚电路上 DTE 设备之间的分组交换。PLP 能在 LAN 和正在运行 LAPD 的 ISDN 接口上运行逻辑链路控制（LLC）。PLP 实现五种不同的操作方式：呼叫建立（call setup）、数据传送（data transfer）、闲置（idle）、呼叫清除（call clearing）和重启（restarting）。

- call setup 方式用于在 DTE 设备间建立 SVC ；
- data transfer 方式用于在虚电路上的两个 DTE 设备间传送数据；
- idle 方式用于虚电路已经建立但没有进行数据传输的情况；
- call clearing 方式用于结束 DCE 设备间的通信会话并终止 SVC ；
- restarting 方式用于在 DCE 设备与本地连接的 DCE 设备之间同步传输。

X.75 是 X.25 的信令协议，定义了 PDN 间的信令系统。X..75 实质上是一种网间接口（NNI）。

这里我们主要讨论 X.25 PLP ，其它协议在个别文件中再作讲解。

协议结构

X.25 PLP 包括很多控制信息。控制数据包，以及所有 X.25 数据包，都是以 3 字节头开始。字节 1,2 包括组（Group）和信道（Channel）字段，两者共同形成 12 位虚拟电路号。每个信息的附加信息都不相同。

1、控制包：

1	2	3	4	8	16	23	24bit
0	0	0	1	Group	Channel	Type	C
Additional Information (Variable)							

2、呼叫请求数据包的附加信息如下：

4 bits	4 bits	Variable	2 bits	6 bits	Variable
Length Calling address	Length Called address	Calling & Called address	00	Facility length	Facilities
Data (Variable)					

其它控制包：

- 如果可以接收呼叫，CALL ACCEPTED 数据包由被呼叫方 DTE 发出。
- 发送 CLEAR REQUEST 信息有多种原因。数据包第 4 字节指明连接清除原因。通过 CLEAR REQUEST CONFIRMATION 数据包进行确认。
- INTERRUPT 数据包允许短信号（32 字节）被发送出序列。通过 INTERRUPT CONFIRMATION 数据包进行确认。
- 在没有反向流量的位置，RECEIVE READY（RR）数据包用来发送独立确认信息。PPP 字段（type 字段的前 3 位）通告需要的下一个数据包。
- RECEIVE NOT READY（RNR）数据包允许 DTE 通知其它方暂时停止发送数据包给它。
- REJECT 数据包允许 DTE 请求重发数据包系列。PPP 字段提供需要的第一个序列号。
- RESET 和 RESTART 数据包用于不同程度的故障恢复。通过 RESET CONFIRMATION 和 RESTART CONFIRMATION 进行相对确认。
- DIAGNOSTIC 数据包为用户提供故障通知。

3、数据包格式如下所示：

1	2	4	8	16	23	24	31	32bit
Q	D	Modulo	Group	Channel	Piggyback	M	sequence	C
Data (Variable)								

LAPF：数据链路层帧方式接入协议 — ITU Q.922

数据链路层帧方式接入协议（LAPF）定义在 ITU Q.922 中，是一种在帧中继网络中为帧方式业务提供拥塞控制性能的增强版 LAPD （Q.921） 。其主要功能有：

- 帧界定、队列、标志透明度；
- 虚电路复用技术和解除复用技术；
- 八位字节队列：0 比特位之前插入整数八位字节；
- 检验帧大小最小值和最大值；
- 差错监测、序列和无复制；
- 拥塞控制；

LAPF 的作用是为 ISDN 用户 — 网络接口的 B、D 或 H 通路上为帧方式承载业务，在用户平面上的数据链路(DL)业务用户之间传递数据链路层服务数据单元(SDU)。帧方式承载链接业务通过在 Q.933 推荐中说明的程序建立，或者通过（用于永久虚电路）预订方式建立。LAPF 可以使用物理层业务，并允

许兼容 HDLC 程序在 ISDN 用户 — 网络接口的 B、D 或 H 通路上为一个或多个帧方式承载链接业务提供统计复用技术。

LAPF 应用于点对点信令方式的帧中继网络。

协议结构

LAPF 类似于 LAPD，其地址格式如下：

6	7	8	12	13	14	15	16bit
DLCI	C/R	EA	DLCI	FECN	BECN	DE	EA

- DLCI — 数据链路连接标识符字段表示帧地址并与 PVC 相对应。
- C/R — 指明帧是命令还是响应。
- EA — 扩展地址字段，表示帧中继头中附加的两个字节。
- FECN — 前向显式拥塞通知（参见下面的 ECN）。
- BECN — 后向显式拥塞通知（参见下面的 ECN）。
- DE — 丢弃指示。

LAPF 控制字段格式：

Control field bits (Modulo 128)	8	7	6	5	4	3	2	1
I Format	N(S)							0
	N(R)	P/F						
S Format	X	X	X	X	Su	Su	0	1
	N(R)							P/F
U Format	M	M	M	P/F	M	M	1	1

- N(S) — 发送方发送序列号。
- N(R) — 发送方接收序列号。
- P/F — 用作命令时，作为 Poll 位；用作响应时，作为 Final 位。
- X — 预留，值为 0。
- Su — 管理功能位。
- M — 修正功能位。

MSN : MingLang.Xu@Hotmail.Com

Q.2931：ATM 信令用户网络接口

信令是一个供 ATM 用户和网络交换控制信息、请求网络资源使用或者协商电路参数使用的进程。Q.2931 基于 Q.931，是信令协议的 ITU 版本。它定义了在 B-ISDN 用户网络接口上建立、维护和释放网络连接的过程。UNI 3.1 规范基于 Q.2931，该过程根据所交换的信息而定义。

Q.2931 信令支持的基本能力如下：

- 信道连接需求（交换虚拟）；
- 点对点交换信道连接；
- 对称或非对称带宽连接；
- 单连接（点对点）呼叫；
- 协议信息、信息元素和过程的基本信令功能；
- X 类、 A 类和 C 类 ATM 传输服务；
- 信令参数的请求和指示；
- VCI 协商；
- 所有的信令消息在指定的带外通路上传输；
- 差错恢复；
- 用公共 UNI 寻址方式唯一标识 ATM 终端；
- 支持端到端兼容性参数的识别；
- N-ISDN 信令网络和 N-ISDN 服务规定；
- 转发兼容性。

Q.2931 的信息类型与 UNI 中的大致相同，除了点对多点信息（3.0/3.1 版本不支持）、建立应答和版本 4.0 中的 INFORMATION 等以外。Q.2931 新添加的信令信息有：ALERTING、PROGRESS、SETUP ACKNOWLEDGE、INFORMATION 和 NOTIFY 。

一旦信令交换成功，就分配 VPI/VCI 和请求带宽。信息的发送在信令 ATM 适配层（SAAL）上，这样可以确保发送的可靠性。

协议结构

下面的协议栈图表支持用户网络连接控制信令：

<table>
<tr><td></td><td>User-Network Signaling</td></tr>
<tr><td rowspan="5">SAAL</td><td></td></tr>
<tr><td>UNI SSCF</td></tr>
<tr><td></td></tr>
<tr><td>SSCOP</td></tr>
<tr><td></td></tr>
<tr><td>AAL Type 5 Common Part</td></tr>
<tr><td></td><td>ATM Layer</td></tr>
</table>

 MSN : MingLang.Xu@Hotmail.Com

	Physical Layer

Q.2931 信令信息：

8	7	6	5	4	3	2	1	bit/Octet
Protocol discriminator								1
0	0	0	0	Length of Call Reference Value				2
Flag	Call reference value							3
Call Reference value (continued)								4
								5
Message Type								6
Message Type (continued)								7
Message Length								8
Message length (continued)								9
Variable Length Information Elements as Required								etc

- Protocol Discriminator — 区别用户网络呼叫控制信息和其它信息（Q.2931 信息为 9）。
- Call Reference — 每个 ATM 连接的唯一编号，通过它链接相同连接中的所有信令信息。该字段由 Call Reference 值和 Call Reference 标志构成。Call Reference 标志表示分配 Call Reference 值的对象。
- Message Type — 连接控制信息类型。
- Message Length — 信息内容长度。
- Information Elements — 具有多种信息元素。其中一些在信息中只出现一次；其他一些出现不止一次。有些信息元素是强制的，有些是可选择的，这主要取决于信息类型。信息元素的顺序与信令协议无关。有关 Q.2931 中的信息元素如下所示：
    - 被呼叫方号码（Called Party Number）
    - 被呼叫方子地址（Called Party Sub-Address）
    - 转接网络选择（Transit Network Selection）
    - 重启指示器（Restart Indicator）
    - 窄带低层兼容性（Narrow-Band Low Layer Compatibility）
    - 窄带高层兼容性（Narrow-Band High Layer Compatibility）
    - 宽带锁定转换（Broadband Locking Shift）
    - 宽带非锁定转换（Broadband Non-Locking Shift）
    - 宽带完全发送（Broadband Sending Complete.）
    - 宽带重复指示器（Broadband Repeat Indicator）
    - 呼叫方号码（Calling Party Number）
    - 呼叫方子地址（Calling Party Sub-Address）

 MSN : MingLang.Xu@Hotmail.Com

- ATM 适配层参数（ATM Adaptation Layer Parameters）
- ATM 流量描述符（ATM Traffic Descriptor）
- 连接标识符（Connection Identifier）
- OAM 流量描述符（OAM Traffic Descriptor）
- 服务质量参数（Quality of Service Parameter）
- 宽带承载容量（Broadband Bearer Capability）
- 宽带低层信息（Broadband Low Layer Information (B-LLI)）
- 宽带高层信息（Broadband High Layer Information (B-HLI)
- 终端对终端转接时延（End-to-End Transit Delay）
- 通知指示器（Notification Indicator）
- 呼叫状态（Call State）
- 进程指示器（Progress Indicator）
- 窄带承载容量（Narrow-Band Bearer Capability）
- 事件起因（Cause）

MPOA：ATM 上的多协议

MPOA 的目的是在 LANE 环境中有效地传输子网间的单播数据。MPOA 集成了 LANE 和 NHRP 以保留 LANE 的优点，同时在数据路径不需要路由器的情况下，允许 ATM VCC 上的子网和 internet 网络层协议间的通信。此外 MPOA 提供了一个框架，能在有不同协议、网络技术和 IEEE 802.1 虚拟 LAN 环境下，将桥接和路由与 ATM 有效的结合起来。MPOA 能够同时使用网桥和路由信息来定位 ATM 环境最佳出口。MPOA 允许网络层路由记算和数据转发在物理位置上分离，这称为虚拟路由技术。

基于 ATM UNI 信令、LAN 仿真和下一跳解析协议(NHRP),MPOA 定义了两部分 — MPOA 客户机(MPC)和 MPOA 服务器（MPS） — 以及通信和接收服务所需的相关协议。

MPS 是路由器的一部分，只应用于具有下一跳服务器（NHS）和一个或多个 LEC 接口的路由器。从路由器出发通过 LEC 到达 LANE 的数据和控制路径保持不变。但是 MPS 与路由器、LEC、NHS 及其它 MPOA 成分是相互作用相互影响的，单一 MPS 对应联系一个 LEC。

MPOA 使用基于下一跳解析协议(NHRP)的协议来管理高速缓存以及建立捷径。它主要执行以下操作：

- 配置 —获取正确的配置信息；
- 发现 —MPC 和 MPS 相互知道对方存在；
- 目标解析 — 决定目标与出口 ATM 地址间的映射，可选标记符和参数组用来建立 VCC 捷径，由此在子网边缘转发数据包；
- 连接管理 —VCC 的创建、维护和终止，目的是传输控制信息和数据；
- 数据传输 — 通过捷径转发 internet 网络层数据 。

MPOA 必须支持所有 PDU 的 LLC/SNAP 的封装。缺省状况下，VCC 必须使用 LLC 封装的信号。MPOA 必须能够建立、接收和维护 VCC，该 VCC 通向符合连接管理程序的任何实体，不管实体是否属于 MPOA 组

成部分。

协议结构

MPOA 标签封装格式：

0	LLC- X "AA"	LLC-X "AA"	LLC X "03"	OUI-X "00"
4	OUI-X "00"	OUI-X "00"	Frame-Type = 0x884C	
8	MPOA Tag			
12-n	Internetwork Layer PDU (up to 2^16 - 13 octets)			

MPOA Control Frame — MPOA 标签封装格式：

0	LLC- X "AA"	LLC-X "AA"	LLC X "03"	OUI-X "00"
4	OUI-X "00"	OUI-X "5E"	Frame-Type = 0x0003	
8 -n	MPOA PDU (up to 2^16 - 9 octets)			

缺省状态下，MPOA 通过 LLC 封装管理 [NHRP] 中定义的所有控制流，它采用与 NHRP 数据包相同的固定头格式，如下所示：

0	ar$afn		ar$pro.type	
4	ar$pro.snap			
8	ar$pro.snap	ar$hopcnt	ar$pkstz	
12	ar$chksum		ar$extoff	
16	ar$op.version	ar$op.type	ar$shtl	ar$sstl

- ar$afn — 定义传送的链路层地址类型。
- ar$pro.type — 协议类型。该字段是 16 位无符号整型数据。
- ar$pro.snap — 当 r$pro.type 字段等于 0x0080，ar$pro.snap 字段的一种 snap 编码扩展，用来编码协议类型。缺省状态下，该字段值设为 0。
- ar$hopcnt — 跳数：MPOA 数据包中允许经过的最大 NHS 数目。
- ar$pktsz — MPOA 数据包的总长（octet）。
- ar$chksum — 整个 MPOA 数据包上的标准 IP 16 位校验和。
- ar$extoff — 该字段用于识别 MPOA 扩展的存在和位置。
- ar$op.version — 通用地址映射和管理协议的版本，设置为 X "01" NHRP。

- ar$op.type — MPOA 数据包类型。具有以下类型值：

128	MPOA Cache Imposition Request.	129	MPOA Cache Imposition Reply.
130	MPOA Egress Cache Purge Request.	131	MPOA Egress Cache Purge Reply.
132	MPOA Keep-Alive.	133	MPOA Trigger.
134	MPOA Resolution Request.	135	MPOA Resolution Reply.
136	MPOA Error Indicator		

- ar$shtl — 源 NBMA 地址的类型和长度。
- ar$sstl — 源 NBMA 子地址的类型和长度。

ATM PNNI：ATM 专用网间接口

专用网间接口（PNNI）是一种 ATM 网间信令协议，主要提供了一种机制 — 支持可扩展的基于 QoS 的 ATM 路由和交换机到交换机的交换虚拟连接（SVC）间的协作性。

PNNI 是一种分层式的动态链路状态路由协议。它支持大规模的 ATM 网络。PNNI 协议为其信息使用 VPI/VCI 0.18。此外在多个网络情况下，PNNI 通过信令信息建立网络连接。PNNI 基于 UNI 4.0 和 Q.2931，UNI 4.0 中加入某些特定信息元素用于支持 PNNI 的路由处理。PNNI 信令包含了动态建立、维护和清除 ATM 连接的过程，该连接存在于 2 个 ATM 网络或 2 个 ATM 网络结点间的专用网到网络接口或网络结点接口上。PNNI 信令协议基于 ATM 论坛 UNI 规范和 Q.2931。

PNNI 信息包括：发信号、呼叫进行、连接、安装、释放、完全释放、通报、状态、状态查询、请求、请求承认响应、状态、添加部分、添加部分承认响应、部分发信号、添加部分拒绝、结束部分、结束部分承认响应。

协议结构

PNNI 头结构如下所示：

2	2	1	1	1	1

Packet type	Packet length	Prot ver	Newest ver	Oldest ver	Reserved

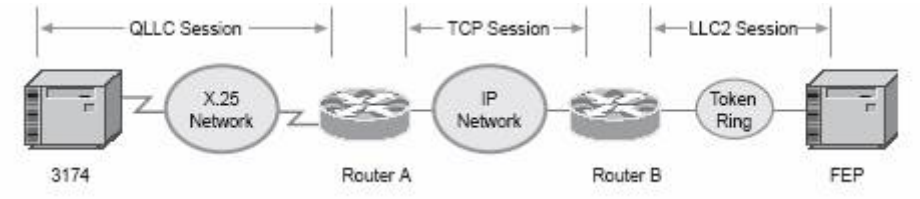
- Packet Type：定义了以下几种数据包类型
    1. Hello — 每个节点发送该数据包，以识别同一对等组中的邻居节点。
    2. PTSP — PNNI 拓朴状态数据包。在各组之间传送拓朴信息。
    3. PTSE — PNNI 拓朴状态元素（请求和确认）。传送拓朴参数，如活动链路，可用带宽等。
    4. Database Summary — 用于相邻对等结构间的源数据库交换期间。
- Packet Length — 数据包长度。
- Prot ver — 协议版本。根据数据包采用的版式决定。
- Newest ver/Oldest ver — 最新版本支持/最旧版本支持。最新版本支持和最旧版本支持字段依次包含在该结构中，这样节点可以协商最近的协议版本，进行特殊数据包类型交换的两个节点都能理解这种版本。

QLLC：限定式逻辑链路控制

限定式逻辑链路控制(QLLC)的由 IBM 定义的一种数据链路层协议，它支持 X.25 网络上的 SNA 数据传输。当 SNA 用于 X.25 时，在 X.25 数据包头通过 Q-bit 表示特殊链路控制信息。该信息与两个系统间互相通信的 SNA 控制 相关，而与 X.25 链路控制无关。通过以上这些限定式数据包，SNA 能决定两个通信系统间的主叫方与被叫方以及最大信息大小。

X.25 数据包中的 QLLC 命令通过 Q-bit 实现。X.25 数据包包括 QLLC 原语，该包通常为 5 字节长，有时为 X.25 数据包头加上 2 字节的 QLLC 控制信息。一旦 QLLC 连接成功，便可以利用 X.25 连接的虚拟电路来转发数据流量。逻辑链路控制(LLC：Logical Link Control)是高级数据链路控制 (HDLC：High Level Data Link Control)的子集，另外同步数据链路控制(SDLC：Synchronous Data Link Control)和 QLLC 也是 HDLC 的子集。

QLLC 典型网络体系结构如下所示：



QLLC 网络结构图

QLLC 支持以下 X.25 可选功效：

- Modulo 8/128 数据包序列号
- 关闭的用户组
- 认可的私有操作代理
- 网络用户识别
- 被叫计费（Reverse Charging）数据包大小协商
- 数据包大小协商
- 窗口大小协商
- 吞吐量（Throughput）类协商

协议结构

QLLC 帧结构与 HDLC 相同，具有以下帧类型：

- QRR — 准备接收
- QDISC — 无链接
- QUA — 无编号确认
- QDM — 无连接模式
- QFRMR — 帧拒绝
- QTEST — 测试
- QRD — 请求无连接
- QXID — 交换识别
- QSM — 设置模式

LLC：逻辑链路控制（IEEE 802.2）

逻辑链路控制（LLC）是一种 IEEE 802.2 LAN 协议，规定了数据链路层中 LLC 子层的实现。 IEEE 802.2 LLC 应用于 IEEE802.3 （以太网）和 IEEE802.5（令牌环） LAN ，以实现如下功能：

- 管理数据链路通信
- 链接寻址
- 定义服务接入点 Service Access Points （SAP）
- 排序

LLC 为上层提供了处理任何类型 MAC 层的方法，例如，以太网 IEEE 802.3 CSMA/CD 或者令牌环

IEEE 802.5 令牌传递（Token Passing）方式。

LLC 是在高级数据链路控制（HDLC ： High-Level Data-Link Control）的基础上发展起来的，并使用了 HDLC 规范的子集。LLC 定义了三种数据通信操作类型：

类型 1：无连接。该方式不保证发送的信息一定可以收到。

类型 2：面向连接。该方式提供了四种服务：连接的建立、确认和数据到达响应、差错恢复（通过请求重发接收到的错误数据实现）以及滑动窗口（系数：128）。滑动窗口用来提高数据传输速率。

类型 3：无连接应答响应服务。

类型 1 的 LLC 无连接服务中规定了一种静态帧格式，并允许在其上运行网络协议。使用传输层协议的网络协议通常会使用服务类型 1 方式。

类型 2 的 LLC 面向连接服务支持可靠数据传输,运用于不需要调用网络层和传输层协议的局域网环境 。

协议结构

逻辑链路控制层（LLC）头：

8	16	24 or 32 bit	Variable
DSAP	SSAP	Control	LLC Information

DSAP — 目标服务访问点（Destination Service Access Point）结构如下：

1	8 bit
I/G	Address Bits

I/G：个人/组地址可能为：0 表示个人 DSAP；1 表示组 DSAP

SSAP — 源服务访问点（Source Service Access Point）字段结构如下：

1	8 bit
C/R	Address Bits

C/R：命令/响应： 0 命令；1 R 响应

 MSN : MingLang.Xu@Hotmail.Com

Control — 控制（Control）字段结构如下：

	1			8	9	16 bit
Information	0	N(S)			P/F	N(R)
Supervisory	1	0	SS	XXXX	P/F	N(R)
Unnumbered	1	1	MM	P/F	MMM	

N(S) — 发送方发送序列号 Transmitter send sequence number。

N(R) — 发送方接收序列号 Transmitter receive sequence number。

P/F — Poll/final 位。命令 LLC PDU 传输/响应 LLC PDU 传输。

- S — 监督功能位。
- 00 — 准备接收（RR）。
- 01 — 拒绝（REJ）。
- 10 — 未准备接收（RNR）。

X — 预留，值为 0。

M — 修正功能位。

LLC Information — LLC 数据或高层协议。

SONET/SDH：同步光纤网络和同步数字层级

同步光纤网络（SONET）和同步数字层级（SDH），是一组有关光纤信道上的同步数据传输的标准协议，常用于物理层构架和同步机制。SONET 是由美国国家标准化组织(ANSI)颁布的美国标准版本。SDH 是由国际电信同盟（ITU）颁布的国际标准颁布。

SONET/SDH 可以应用于 ATM 或非 ATM 环境。SONET/SDH（POS）上的数据包利用点对点协议（PPP），将 IP 数据包映射到 SONET 帧负载中。在 ATM 环境下，SONET/SDH 线路连接方式可能为多模式、单模式或 UTP。SONET 是基于传输在基本比特率是 51.840 Mbps 的多倍速率，或 STS-1。而 SDH 是基于 STM-1，数据传输率为 155.52Mbps，与 STS-3 相当。目前常用 SONET/SDH 数据传输率列表如下：

SONET 信号	比特率	SDH 信	SONET 性能	SDH 性能

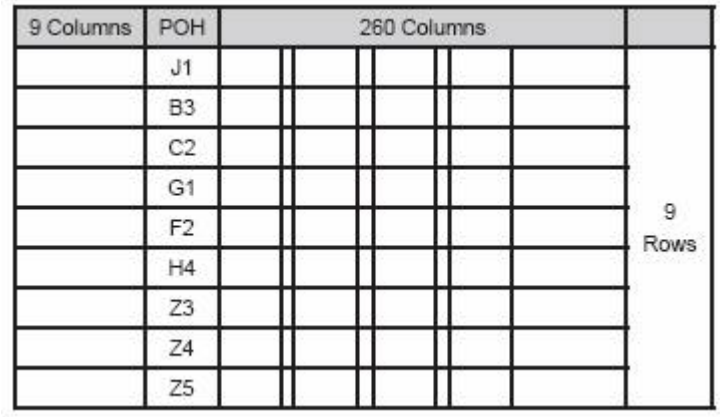
 MSN : MingLang.Xu@Hotmail.Com

	(Mbps)	号		
STS－1 和 OC－1	51.840	STM－0	28 DS－1s 或 1 DS－3	21 E1s
STS－3 和 OC－3	155.520	STM－1	84 DS－1s 或 3 DS－3s	63 E1s 或 1 E4
STS－12 和 OC－12	622.080	STM－4	336 DS－1s 或 12 DS－3s	252 E1s 或 4 E4s
STS－48 和 OC－48	2,488.320	STM－16	1,344 DS－1s 或 48 DS－3s	1,008 E1s 或 16 E4s
STS－192 和 OC－192	9,953.280	STM－64	5,376 DS－1s 或 192 DS－3s	4,032 E1s 或 64 E4s
STS-768 和 OC-768	39,813,120	STM-256	21,504 DS－1s 或 768 DS－3s	16,128 E1s 或 256 E4s

另外一些速率定义，如 OC-9、OC-18、OC-24、OC-36、OC-96 及 OC-768，可参照相关标准文档，但它们使用并不普遍。其它更高的传输速率供未来使用。

协议结构

STS 和 STM 的帧结构不同。这里我们只对 STS-1 帧结构作具体介绍。STS-1 帧结构具有 9 行 90 列。前 3 列为传输开销 （TOH）并包括帧结构、差错监控、管理和有效载荷指针信息。剩余的 87 列供数据（有效载荷）使用，其中第 1 列为通道开销 （POH）。 TOH 中的指针用于识别有效载荷的开始，这里可以参考 同步净荷包 （SPE）。



SONET/SDH：同步光纤网络和同步数字层级

- SOH — 段开销。
  - A1 和 A2 — 帧定位。该字段包括值 0xF628。接收方在导入比特流中搜索这些值。这些值是规则的。
  - C1 — STS-1 识别。OC-3c 和 STM-1 包括 3 个 STS-1 流，相对而言这 3 个 C1 字节包括 0x01、 0x02 和 0x03。
  - B1 — 段差错监控。前帧中包括了所有位的 BIP-8 并采用偶校验方法。

- LOH — 线路开销。
  - B2 — 线路差错监控。包括前帧偶校验时，线路开销的所有位上计算的 BIP-24。
  - H1（bits 1-4） — 新数据标志（指针改变时指定该标志），通道 AIS。
  - H1 和 H2（bits 7-16） — 指针值，通道 AIS。该字段指定了指针和第一个有效载荷字节间的偏移量。至少连续接收 3 次接收，该值中的改变可以忽略。
  - H1* 和 H2* — 级联指示，通道 AIS。
  - H3 — 指针作用（用于频率调整），通道 AIS。
  - K2（bits 6-8） — 线路 AIS、线路 FERF、线路 FERF 的移位。
  - Z2 — 线路 FEBE。该字段包括早先时间检测到的 B2（BIP-24）差错数目。
- POH — 通道开销
  - J1 — STS 通道追踪。传输 64 字节的固定串，使得通道中的接收终端可以校验与发送方间的继续连接。其内容尚未指定。
  - B3 — 通道差错监控。包括前帧有效载荷的所有位上的通道 BIP-8，不规则性之前采用偶校验方法。
  - C2 — 通道信号电平指示器。包括其中一种代码：
    代码 0：表示 STS 未装备有效载荷：无通道源装备。
    Code 1：表示 STS 装备有效载荷：无需进一步进行区别的有效载荷，没有指定有效载荷。
  - G1（bits 1-4） — 通道 FEBE。允许对复杂通道的任意点上的全部全双工通道进行监控。
  - G1（bit 5） — 通道黄色警告。通道 RDI（远端缺陷—降质指示）。

AAL0 - AAL5：ATM 适配层类型 0-5，为可变比特率视频传输提供的预留

ATM 层位于 ATM 参考模式的第二层，提供了 ATM 适配层（AAL）和物理层之间的接口。该层主要负责将信元从 AAL 转发给物理层，如用于传输的 SONET；以及从物理层转发给 AAL，便于其在终端系统使用。ATM 层能够决定到达的信元应该被转发至哪里；重新设置通讯连接标识符并且转发信元给下一个链接、缓冲导入/导出信元以及处理各种流量管理功能，如信元丢失优先权标记、拥塞标注和通用流控制接入。此外 ATM 层还负责监控传输率和服务合同（流量策略）监控的一致性。

在 ATM 头字段定义了 ATM 层功能。ATM 信元头有两种不同格式：其一是用户网络接口（UNI），其二是网络结点接口（NNI），它主要用于内部网络。

协议结构

UNI 结构：

8	7	6	5	4	3	2	1

GFC	VPI	
VPI	VCI	
VCI		
VCI	PT (3 Bit)	CLP
HEC		
Information (48 Bytes)		

- GFC — 通用流控制。限制拥塞控制期间进入网络的数据数量。
- VPI — 虚路径标识符。
- VCI — 虚通道标识符。VPI 和 VCI 共同组成路由选择字段，称之为 VPCI。VPCI 通过一个特殊的通道或电路联接每个信元。VCI 是一个单通道标识符。VPI 支持不同的 VCI 组成 VCI 组，并支持该组作为一个实体共同交换。但是 VPI 和 VCI 只作用于本地链路，路由选择字段的内容随着信元从一个链路到另一个链路的改变而改变。UNI 中该字段，可以支持高达 160 万位用户加入网络会话。
- PT — 有效载荷类型。
- CLP — 信元丢失优先权.
- HEC — 信头错误控制

NNI 结构：

8 7 6 5 4 3 2 1		
VPI		
VPI	VCI	
VCI		
VCI	PT (3 Bit)	CLP
HEC		
Information (48 Bytes)		

- VPI — 虚路径标识符。
- VCI — 虚通道标识符。请参照以上 VPCI 的具体介绍。该字段描述了 ATM 信元中从网络到节点的路由选择信息，它能够支持 26.8 百万的 NNI 会话。
- PT — 有效载荷类型。
- CLP — 信元丢失优先权。
- HEC — 信头错误控制。

 MSN : MingLang.Xu@Hotmail.Com

LANE UNI：ATM LAN 仿真 UNI

ATM LAN 仿真 UNI（LANE）定义了 ATM 网络如何充分仿真一组现有 LAN 技术（如以太网、令牌环等）的介质访问控制（MAC）服务，这样不需要修改即可以使用高层网络协议。一个仿真 LAN （ELAN），以在 ATM 交换网络上的以太网或令牌环网的形式出现，由 LE 客户机集和一组协作服务实体组成： LAN 仿真配置服务器（LECS）、仿真服务器（LES）、广播和未知服务器（BUS）及可选择组播服务器（SMS）。

LAN 仿真 UNI （LUNI）定义了在仿真客户机（客户机）和仿真服务之间的协议及交互作用，包括初始化、注册、地址解析和数据传输过程。通过 LUNI，每一个 LE 客户机可能连接到单一 LES 及 BUS ，但也可能连接到单一 LECS 或多个 SMS。

LE 客户机间、LE 客户机和 LE 服务间的通信都在 ATM 虚拟信道连接（VCC）上完成，而所有 LE 客户机必须在控制和数据 VCC 上实现与 LE 服务通信。LANE 假定点对点和点对多点的交换虚拟电路（SVC）可用。组播转发和控制分发流都是在点对多点 VCC 上所传输的；直接数据、直接控制、直接配置、缺省组播发送和可选择组播发送都是在点对点 VCC 上完成的。除了直接数据流是多元流以外，其它所有的都是非多元流。

LAN 仿真包含了以太网和令牌环仿真。在以太网仿真中，LAN 仿真组件需要仅仅通过检查数据帧的目标 MAC 地址，来指引帧到达最终目的地；但在令牌环仿真中，LAN 仿真组件必须要通过由数据帧路由信息字段（RIF）中提取出来的“路由描述符”来正确指引 LAN 仿真上的数据帧。

在 ATM-to-legacy LAN 桥和 ATM 终端系统中，大多数 LAN 仿真服务都是作为网络层下的设备驱动程序实现的。在 LANE 中，通常使用“可用比特率”（ABR）服务支持带宽管理能力。

协议结构

LE 数据帧：

1、关于 802.3（Ethernet）Frame — 非复用数据帧

0	LE Header	Destination Address
4	Destination Address	
8	Source Address	
12	Source Address	Type / Length
16 and on	User Info	

2、关于 802.5（Token Ring） Frame — 非复用数据帧：

0	LE Header	AC PAD	FC

 MSN : MingLang.Xu@Hotmail.Com

4	Destination Address	
8	Destination Address	Source Address
12	Source Address	Type / Length
16-46	Routing Information Field	
	User Info	

- LE Header — LAN 仿真头，包括 LAN 仿真客户机识别值和发送客户机或 X "0000"。

LE 控制帧：直接 VCC 除了 LLC 复用数据外，所有 LAN 仿真控制帧，如 LE_FLUSH_REQUEST、READY_IND 和 READY_QUERY，采用的格式如下：

0	MARKER = X "FF00"		PROTOCOL = X "01"	VERSION = X "01"
4	OP-CODE		STATUS	
8	TRANSACTION-ID			
12	REQUESTER-LECID		FLAGS	
16	SOURCE-LAN-DESTINATION			
24	TARGET-LAN-DESTINATION			
32	SOURCE-ATM-ADDRESS			
52	LAN-Type	MAX-Frame-Size	Number-TLVS	ELAN-Name-Size
56	TARGET-ATM-ADDRESS			
76	ELAN-NAME			
108	TLVs BEGIN			

- OP-CODE — （2 字节）控制帧操作类型。部分 OP 代码如下：

OP-CODE Value	OP-CODE Function
X "0001" & X "0101"	LE_CONFIGURE_REQUEST & LE_CONFIGURE_RESPONSE
X "0002" & X "0102"	LE_JOIN_REQUEST & LE_JOIN_RESPONSE
X "0003" & X "0103"	READY_QUERY & READY_IND
X "0004" & X "0104"	LE_REGISTER_REQUEST & LE_REGISTER_RESPONSE
X "0005" & X "0105"	LE_UNREGISTER_REQUEST & LE_UNREGISTER_RESPONSE

 MSN : MingLang.Xu@Hotmail.Com

X "0006" & X "0106"	LE_ARP_REQUEST & LE_ARP_RESPONSE
X "0007" & X "0107"	LE_FLUSH_REQUEST LE_FLUSH_RESPONSE
X "0008" & X "0108"	LE_NARP_REQUEST & Undefined
X "0009" & X "0109"	LE_TOPOLOGY_REQUEST & Undefined
X "000A" & X "010A"	LE_VERIFY_REQUEST & LE_VERIFY_RESPONSE

- Status — （2 字节）控制帧操作状态。定义的部分状态代码如下：

Code (dec)	Name	Code (dec)	Name
0	Success	1	Version Not Supported
2	Invalid request parameters	4	Duplicate LAN Destination registration
5	Duplicate ATM address	6	Insufficient resources to grant request
7	Access denied	8	Invalid REQUESTOR-ID
9	Invalid LAN Destination	10	Invalid ATM Address
20	No Configuration	21	LE_CONFIGURE Error
22	Insufficient Information	24	TLV Not Found

- TLV — 类型/长度/值编码参数。

LANE LLC-Multiplexed Frame — 具有 12-Octet LLC 复用头：

0	LLC- X "AA"	LLC-X "AA"	LLC X "03"	OUI-X "00"
4	OUI-X "A0"	OUI-X "3E"	Frame-Type	
8	ELAN-ID			
12-28/ 58	LANE Data Frame Header (802.3 or 802.5)			
	User Info			

LLC 字段为 3 Octet，包括定值 X "AAAA03"，后面是 OUI。

OUI 字段为 3 Octet，包括定值 X "00A03E"，表示“ATM Forum”。

后面的两个字节是 FRAME-TYPE 字段，包括 IEEE 802.3 数据帧值为 X "000C"，IEEE 802.5 数据

帧值为 X ″000D″，LANE LLC-Multiplexed READY_IND 和 READY_QUERY 控制帧值为 X ″000E″。

ELAN-ID 字段识别该数据帧的仿真 LAN。

LANE NNI：ATM LAN 仿真 NNI

ATM LAN 仿真 NNI（LANE）使得能在 ATM 网络上实现仿真 LAN 操作。一个仿真 LAN 在其所有用户间提供了用户数据帧通信，这类似于实际的 LAN。一个或者更多仿真 LAN 可以运行在相同的 ATM 网络上，但是每一个仿真 LAN 相互之间在逻辑上都是独立的。仿真 LAN 间的通信需要一定的互连设备（网桥、路由器等），即使在某些环境下明确地允许在两个仿真 LAN 间直接的 ATM 连接。LAN 仿真 LUNI 定义了 LAN 仿真客户机（LE 客户机）和 LAN 仿者服务之间的协议和交互作用。通过 LUNI，每一个 LE 客户机能够连接到单一 LES 及 BUS ，但也可以连接单一 LECS 或多个 SMS 。

LAN 仿真 NNI（LNNI）定义了相互之间能够看到的 LANE 服务组件行为，并定义了提供分布和可靠 LAN 仿真服务所需的程序。单一 ELAN 可以由多个 LECS、LES、BUS 和 SMS 服务（效力）。一个 LES、BUS 或 SMS 只服务单一 ELAN，而一个 LECS 则可服务多个 ELAN。LANE 服务组件和多个 VCC 互相连接，以便于配置、状态、数据库同步、控制和数据转发。LNNI 规范提供了服务于单个 ELAN 的组件之间的多供应商互操作性，这样客户可以实现不同供应商提供的 LANE 服务的混合和匹配使用。

LANE 服务有四个主要的组成部分：

1. 局域网仿真客户机（LEC）——安装在 ATM 终端系统上，实施 LUNI 接口，作为 LAN 系统代理用于执行数据转发和地址解析，并为高层软件提供一个 MAC 级别的仿真以太网 /IEEE 802.3 或 IEEE 802.5 服务接口。
2. 局域网仿真服务器（LES）——支持地址解析协议（LE-ARP），用于决定负责确定 MAC 地址的目的 ATM 地址。一个 LE 客户机只能连接一个 LE 服务器，它向 LE 服务器注册 LAN 目的地和/或要接收的组播地址。同时 LE 客户机通过查询 LE 服务器来解析 MAC 地址或 ATM 地址的路径描述符。
3. 广播/未知服务器（BUS）——处理转发给 LEC 的组播流量。 LE 客户机可以看到单个广播和未知服务器。
4. 可选择组播服务器（SMS）——用于从 BUS 上分担许多组播处理，同时也需要转发广播帧和无法解析目的地址的帧，以有效转发组播帧。

服务于单个 ELAN 的多个 LANE 服务实体需要互相协作和通信，从而提供一个分层且可靠的仿真服务。 LNNI 所需的通信方式划分如下：

a）控制面板

- 配置和状态通信——从配置直接 VCC 上的 LECS ， LES 和 SMS 获得配置信息。在相同的连接上 LECS 可以获得 LES 和 SMS 状态信息。
- LANE 控制通信——每个 LES 主要负责从本地 LE 客户机到本地 LE 客户机，以及其它 LES 服务器为未注册目标地址分发 LE_ARP 请求。 LES 同时也需要转发回 LE_ARP 响应给原发送者。

另外，LES 还必须转发 LE_FLUSH 响应和 LE_TOPOLOGY 请求给正确的目标地址。

b）同步面板

- LECS 同步－－一个特定的 LECS 也许不能直接接收所有服务器组件状态信息，因此， LECS 相互之间必须交换 LES 和 SMS 状态信息。为了分发这种状态信息， ELAN 中的所有 LECS 必须对网络中所有其它 LECS 维护一个 LECS 同步 VCC 。
- LES-SMS 数据库同步－－ LES 和 SMS 通过 SCSP 使其数据库保持同步。

c）数据面板

- BUS 数据通信－－逻辑上，每个 BUS 和 LES 是成对的，并且 BUS 有权访问 LES 维护的注册数据库，其中包括所有 BUS 的 ATM 地址。在 BUS 和 LES 之间没有定义任何协议。
- SMS 数据通信－－每个 SMS （和 LES ）都能通过 SCSP 为整个 ELAN 获得一份完整的注册数据库拷贝，所以每一个 SMS 都会知道其它每一个 SMS 和 BUS 。当 LE 客户机要解析组播地址时，如果 SMS 可用，则 LES 应该将客户机分配到 SMS 作为发送者，否则 LE_ARP 响应中会返回 BUS 的 ATM 地址。 ELAN 以及其中的所有 SMS 既可以工作在分布模式下，也可以工作在单机模式下，这主要由网络管理员决定。

协议结构

LANE 数据帧：

LNNI 控制帧格式如下所示：

	0	LLC = X ″AAAA03″			OUI
	4	OUI	Frame Type		
	8	ELAN-ID			
LANE Control Frame	12	REQUESTER-LECID		FLAGS	
	16	SOURCE-LAN-DESTINATION			
	24	TARGET-LAN-DESTINATION			
	32	SOURCE-ATM-ADDRESS			
	52	LAN-Type	MAX-Frame-Size	Number-TLVS	ELAN-Name-Size
	56	TARGET-ATM-ADDRESS			
	76	ELAN-NAME			
	108	TLVs BEGIN			

MSN : MingLang.Xu@Hotmail.Com

- LLC — 逻辑链路控制：控制并列 VCC 都封装在 LLC 中。
- OUI — 组织性的唯一标识符= X "00A03E"，表示 ATM 论坛。
- FRAME TYPE = X "000F"。
- ELAN-ID — Emulated LAN ID 。
- OP-CODE（2 Bytes） — 控制帧操作类型。

下面列出了一些已定义的 OP 代码：

OP-CODE Value	OP-CODE Function
X "000b"	LNNI_CONFIGURE_TRIGGER
X "000C"	LNNI_LECS_SYNC_REQUEST
X "000d"	LNNI_KEEP_ALIVE_REQUEST
X "010d"	LNNI_KEEP_ALIVE_RESPONSE
X "000e"	LNNI_VALIDATE_REQUEST
X "010e"	LNNI_VALIDATE_RESPONSE

- Status — （2 字节）控制帧操作状态。
-

TLV — 类型/长度/值编码参数。

LNNI TLV 事例如下：

项目	类型	长度	描述
ServerId	00-A0-3E-14	2	ELAN 中服务器的唯一标识符
ServerGroupId	00-A0-3E-15	2	唯一连接到 ELAN-ID。供 SCSP 使用。
SynchronizationPeerServer	00-A0-3E-16	20	利用 SCSP 进行 DB 同步的 ES 或 SMS 多元 ATM 地址
SmsModeOfOperation	00-A0-3E-19	1	表示 SMS 可行性方式 0 = STAND_ALONE 1 = DISTRIBUTED

IPCP 和 IPv6CP：IP 控制协议和 IPv6 控制协议

IP 控制协议（IPCP）和 IPv6 控制协议（IPv6CP）是一种网络控制协议，用于建立和配置 PPP 上的 IP 或 IPv6 ，它提供一种方法，通过 PPP 协商和使用 Van Jacobson TCP/IP 头部压缩。

IPCP 主要负责在点对点链接终端双方上配置，可用及停用 IP 协议模块。IPCP 使用与链接控制协议（LCP）相同的包交换机制。但只有在 PPP 达到网络层协议阶段时， IPCP 包才可以被交换。在达到这种阶段前接收的 IPCP 包需要被丢弃掉。

在任意 IP 包可能被传送之前， PPP 必须达到网络层协议阶段，并且 IP 控制协议必须处于公开状态。

Van Jacobson TCP/IP 头部压缩将 TCP/IP 协议头的大小减少至 3 字节。这对于低速串行线路，特别是交互式通信来说都是非常重要的。

IP 压缩协议配置选项指出了接收压缩包的能力。如果需要双向压缩，那么每一个链接终端必须分别请求该选项。

IPv6CP 主要负责在点对点链接终端双方上配置，可用及停用 IPv6 协议模块。 IPv6CP 使用与链接控制协议（LCP）相同的包交换机制。但只有在 PPP 达到网络层协议阶段时， IPv6CP 包才可以被交换。在达到这种阶段前接收的 IPv6CP 包需要被丢弃掉。

协议结构

IPCP 和 IPv6CP 配置选项数据包头：

8	16	32 bit
Type	Length	Configuration Option

- Type — 类型 1：IP 地址；类型 2：IP 压缩协议；类型 3：IP 地址
- Length >= 4
- Configuration Option — 该字段为 16 位，可以选择以下一种类型：

对于 IPCP：

- Type 1：IP 地址
- Type 2：IP 压缩协议
- Type 3：IP 地址

对于 IPv6CP：

- Type 1：接口标识符
- Type 2：IPv6 压缩协议

IPCP 和 IPv6CP 头结构：

8	16	32 bit
Code	Identifier	Length
Data （variable）		

- Code — 规定实现的功能
- Identifier — 匹配请求和答复
- Length — 数据包大小（包含头）
- Date — Length 字段规定的 0 或多字节数据。该字段可能有一个或多个选项。

IPX：互联网分组交换协议

互联网分组交换协议（IPX）是 Novell NetWare 操作系统所支持的在互联网络中路由数据包的早期网络协议。IPX 是一种面向无连接通信的数据报协议 — 类似于 TCP/IP 协议组中的网际协议(即 IP)。其高层协议，如 SPX 和 NCP ，主要提供差错恢复服务。

为了选择最佳路径,IPX 使用动态距离矢量(distance vector)路由选择协议,如路由信息协议(RIP：Routing Information Protocol）或链路状态协议（NLSP：NetWare Link-State Protocol）。

Novell IPX 网络地址是唯一的，以十六进制表示。它由两部分组成：网络号和节点号。IPX 网络号由网络管理员分配，地址长 32 位。节点号，通常是系统网络接口卡（NIC）的介质访问控制（MAC）地址，地址长 48 位。通过 MAC 地址作为节点号，系统可以通过发送节点以判断数据链路使用的 MAC 地址。

Novell NetWare IPX 支持单路由器接口的四种封装模式：

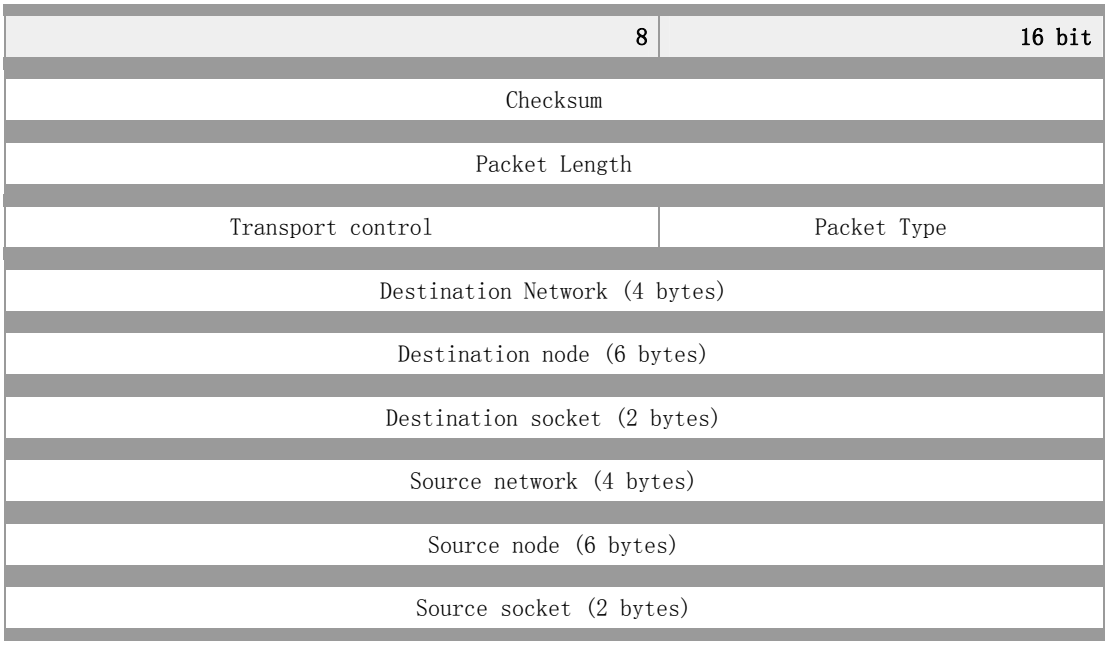
- Novell 私有（ Novell Proprietary ） —也称为原始 802.3 或者 Novell Ethernet_802.3 ，Novell 私有 是 Novell 公司最初使用的封装模式。
- 802.3 — 也称为 Novell_802.2，802.3 是 IEEE 802.3 的标准帧格式。
- Ethernet v 2 — 也称为 Ethernet-II 或 ARPA，Ethernet v2 包括标准 Ethernet v2 协议头，它由目标地址和源地址字段构成，后面是 EtherType 字段。
- SNAP — 也称为 Ethernet_SNAP，SNAP 在原 IEEE 802.2 协议头增加了 type 代码，与 Ethernet v2 中定义的 EtherType 类似。

IPX 数据包数据部分的长最小为 30 字节 （只有头部分 ） ，最大值不定，这主要取决于使用的低

层 MAC 协议（ 以太网或令牌环 ）。

协议结构

NetWare IPX 数据包头：



8	16 bit
Checksum	
Packet Length	
Transport control	Packet Type
Destination Network (4 bytes)	
Destination node (6 bytes)	
Destination socket (2 bytes)	
Source network (4 bytes)	
Source node (6 bytes)	
Source socket (2 bytes)	

- Checksum — 表示当 16 位字段全设置为 1（FFFF）时，不使用校验和。
- Packet length — 指定完整 IPX 数据报的字节长。IPX 数据包大小任意，可以达到媒体最大传输单元（MTU）大小（不允许数据包分片）。
- Transport control — 指明数据包传送过程中经过的路由器数量。当该值为 16 时，假设发生路由回路，并丢弃数据包。
- Packet type — 指定哪个上层协议应该接收该数据包信息。通常包括两个值：
  - 5 — 表示序列分组交换（SPX）
  - 17 — 表示 NetWare 核心协议（NCP）
- Destination network、Destination node、and Destination socket — 表示目标信息。
- Source network、Source node and Source socket — 表示源信息。

DECnet 及其协议

DECnet 是由数字设备公司（Digital Equipment Corporation）推出并支持的一组协议集合。目前市面上的 DECnet 有多种版本。最初的 DECnet 支持两台直接相连的小型机之间的通信。后来推出的版本在原 DECnet 功能基础上另外提供了对附加所有者和标准协议的支持。当前使用较为广泛的两种 DECnet 版本分别为： DECnet Phase IV 和 DECnet phase V。现在 DECnet 已成为 HP 开放 VMS 的一部分。

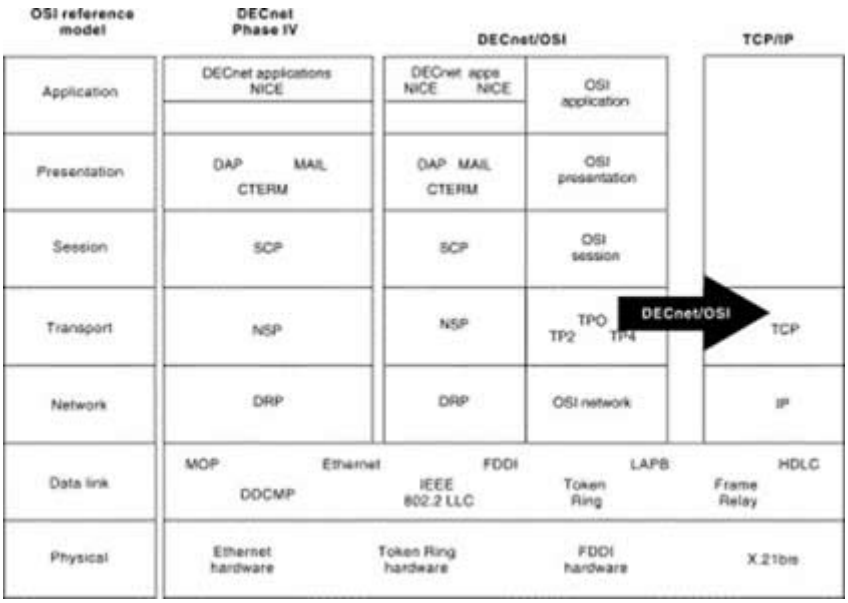
DECnet 是一种基于数字网络体系结构（DNA：Digital Network Architecture）的较为全面的分层网络体系结构，它支持大量的所有者和标准协议。

DECnet Phase IV DNA 类似于 OSI 体系结构,同样采用了分层结构,只是它被分为八层。DECnet Phase IV DNA 规定了上面四层提供用户交互服务、网络管理能力、文件传输和会话管理等功能。上面四层分别为：用户层（user layer）、网络管理层（network management layer）、网络应用层（network application layer）以及会话控制层（session control layer）。

DECnet phase V（可以表示为 DECnet Plus 或 DECnet/OSI）的分层模式可以实现三种协议组：OSI、DECnet 和 TCP/IP。DECnet plus 遵循七层 OSI 参考模型并支持众多的标准 OSI 协议，提供了与 DECnet Phase IV 的反向兼容性（backward compatibility）且支持多个所有者数字协议。此外 DECnet plus 还支持应用层、表示层及会话层各功能，它支持低层 TCP/IP 协议，并能够实现在 TCP 传输协议上的 DECnet 流量传输。

主要协议

OSI 模型中的 DECnet DNA phase IV 和 V OSI 以及与 TCP/IP 组的比较如下：



OSI 模型中的 DECnet DNA phase IV 和 V OSI 以及与 TCP/IP 组的比较

DECnet 协议组包括的主要协议有：

应用层（Application）

- NICE:网络信息和控制交换协议 （NICE:Network Information and Control Exchange protocol）

表示层（Presentation）

- DAP：数据访问协议 （DAP：Data Access Protocol）
- CTERM：命令终端 （CTERM：Command Terminal）

会话层（Session）

- SCP：会话控制协议 （SCP：Session Control Protocol）

传输层（Transport）

- NSP：网络服务协议 （NSP：Network Service Protocol）

网络层（Network）

- DRP：DECnet 路由选择协议 （DRP：DECnet Routing Protocol）

数据链路控制（Data Link Control）

- MOP：维护操作协议 （MOP：Maintenance Operation Protocol）
- DDCMP：数字数据通信报文协议 （DDCMP：Digital Data Communications Message Protocol）

MSN : MingLang.Xu@Hotmail.Com

RSVP-TE：基于流量工程扩展的资源预留协议

基于流量工程扩展的资源预留协议（RSVP-TE）作为 RSVP 协议的一个补充协议，用于为 MPLS 网络建立标签交换路径。这个 RSVP 扩展协议主要用于在有或者没有资源预留的情况下支持明确传送 LSP 的实例。同时它也支持 LSP 的平滑重新路由、优先权及环路监测。

RSVP 协议定义的会话指有明确目的地址及传输层协议的数据流。但当 RSVP 与 MPLS 相结合时，流或者会话的定义具有较大的灵活性和一般性。LSP 的入口结点使用许多方法来决定给一些数据包分配一个特定的标签。一旦某个标签被分配给一组包，这个标签将会有效地定义通过 LSP 的流。我们将这样的 LSP 看作 LSP 隧道，是因为通过它的流量对沿着标签交换路径的中间结点是不透明的。新的 RSVP 会话、发送方及过滤器说明对象，被称之为 LSP 隧道 IPv4 和 LSP 隧道 IPv6 ，已被用来支持 LSP 隧道特征。从标签交换路径结点的角度来看，这些对象语义上是指基于包独立识别的 LSP 隧道流量，其中这些包是从具有上流发送方结点分配的特定标签值的 PHOP 到会话获取的。事实上，出现在对象名字中的 IPv4（v6）只表示目的地址是一个 IPv4 （IPv6）地址。一般情况下，当涉及到这些对象时，需要用到限定词“ LSP 隧道”。

在某些应用程序中，连接 LSP 隧道组是很有用的，诸如，在重新路由操作期间或者传播流量在复合路径上时，这样的集合被称为 TE 隧道。为了能够鉴定和连接 LSP 隧道，需要携带两个标识符。隧道 ID 是会话对象的一部分，会话对象唯一地定义了一个流量工程隧道。发送方和过滤器说明对象携带一个 LSP ID，发送方（或者过滤器说明）对象结合会话对象唯一定义一个 LSP 隧道。

协议结构

除 RSVP 中列出的信息类型外，还包括：

Value	Message Type
14	Hello

此外，还具有以下 Protocol Object Types：

Value	Object Type
16	Label
19	Optical
20	Explicit Route
21	Record Route
22	Hello
207	Attribute Session

 MSN : MingLang.Xu@Hotmail.Com

RSVP：资源预留协议

资源预留协议（RSVP）是一种用于互联网上质量整合服务的协议。 RSVP 允许主机在网络上请求特殊服务质量用于特殊应用程序数据流的传输。路由器也使用 RSVP 发送服务质量（QOS）请求给所有结点（沿着流路径）并建立和维持这种状态以提供请求服务。通常 RSVP 请求将会引起每个节点数据路径上的资源预留。

RSVP 只在单方向上进行资源请求，因此，尽管相同的应用程序，同时可能既担当发送者也担当接受者，但 RSVP 对发送者与接受者在逻辑上是有区别的。 RSVP 运行在 IPV4 或 IPV6 上层，占据协议栈中传输协议的空间。 RSVP 不传输应用数据，但支持因特网控制协议，如 ICMP、IGMP 或者路由选择协议。正如路由选择和管理类协议的实施一样， RSVP 的运行也是在后台执行，而并非在数据转发路径上。

RSVP 本质上并不属于路由选择协议， RSVP 的设计目标是与当前和未来的单播（unicast）和组播（multicast）路由选择协议同时运行。 RSVP 进程参照本地路由选择数据库以获得传送路径。以组播为例，主机发送 IGMP 信息以加入组播组，然后沿着组播组传送路径，发送 RSVP 信息以预留资源。路由选择协议决定数据包转发到哪。 RSVP 只考虑根据路由选择所转发的数据包的 QOS 。为了有效适应大型组、动态组成员以及不同机种的接收端需求，通过 RSVP ，接收端可以请求一个特定的 QOS[RSVP93] 。QOS 请求从接收端主机应用程序被传送至本地 RSVP 进程，然后 RSVP 协议沿着相反的数据路径，将此请求传送到所有节点（路由器和主机），但是只到达接收端数据路径加入到组播分配树中时的路由器。所以， RSVP 预留开销是和接受端的数量成对数关系而非线性关系。

协议结构

4	8	16	32bit
Version	Flags	Message Type	RSVP Checksum
Send TTL		Reserved	RSVP Length

- Version — 协议版本号，当前为 1。
- Flags — 还没有定义标志位。
- Message Type — 可能值有：1 Path，2 Resv，3 PathErr，4 ResvErr，5 PathTear，6 ResvTear，7 ResvConf。
- RSVP Checksum — 用于信息差错的校验和。
- Send TTL — 信息发送时的 IP TTL 值。
- RSVP Length — RSVP 信息二进制形式下的总长，包括通用头和可变长对象。

xDSL：数字用户线路技术（DSL、IDSL、ADSL、HDSL、SDSL、VDSL、G.Lite）

数字用户线路（xDSL）是在家庭和企业的普通铜质电话线路（POTS）上提供宽带数据访问的一种调制解调器技术。xDSL 是所有 DSL 类型的统称，诸如 ADSL、G.Lite、HDSL、SDSL、IDSL 和 VDSL 等。xDSL 有时称为“最后一英里”技术，这是因为它们只用于电话交换站和家庭或办公室之间的连接，而不是交换站之间的连接。

xDSL 与 IDSL 的相似处有：它们两者在现有电话网络（POTS）上运行，都采用复杂调制模式；都需要到中心电信局的短距离连接（一般小于 20，000 英尺）。但是 xDSL 提供了更高的速度 — 上行流量达到 32 Mbps，下行流量从 32 Kbps 到 1 Mbps 以上。

各种 DSL 支持一些相关调制技术：

- 离散多音复用技术（DMT）；
- 简单线路代码（SLC）；
- 无载波幅度和相位调制（CAP）；
- 多虚拟数字用户线（MVL）；
- 离散小波多频调制（DWMT）。

为了多个 DSL 用户和高速中枢网络之间相互连接，电话公司采用数字用户线接入复用器（DSLAM）来实现。DSLAM 集合所有访问 DSL 线路上的数据传输，然后将它们连接到异步传输模式（ATM）网络。在每个传输的另一终端，DSLAM 非多路复合信令然后将它们转发给适当的个人 DSL 连接。

大多数 DSL 技术要求在客户处安装一个信号分离器，但也可以在中心办公室远程管理分离器，也就是无分离 DSL、“DSL Lite”、G.Lite 或者通用 ADSL。

协议结构

各种 DSL 规范摘要列表如下所示：

类型	描述	数据速率	模式	距离	应用
IDSL	ISDN 数字用户线	128 kbps	双工	18k ft on 24 gauge wire	ISDN 服务于语音和数据通信
HDSL	高速率数字用户线	1.544 Mbps 到 42.048 Mbps	双工	12k ft on 24 gauge wire	T1/E1 服务于 WAN、LAN 访问和服务器访问
SDSL	单线对数字用户线	1.544 Mbps 到 2.048 Mbps	双工	12k ft on 24 gauge wire	与 HDSL 应用相同，另外为对称服务提供场所访问。
ADSL	非对称数字用户线	1.5 到 9 Mbps 16 到 640 kbps	下行	Up to 18k ft on 24 gauge wire	Internet 访问，视频点播、单一视频、远程 LAN 访问、交互多媒体
DSL Lite	“Splitterless”	1.544 Mbps 到	下	18k ft on 24	标准 ADSL；在用户场所无

 MSN : MingLang.Xu@Hotmail.Com

(G.Lite)	DSL	6 Mbps 16 到 640 kbps	行	gauge wire	需安装 splitter。
VDSL	甚高数据速率数字用户线	13 到 52 Mbps 1.5 到 2.3 Mbps	下行	1k to 4.5k ft depending on data rate	与 ADSL 相同，另外可以传送 HDTV 节目。

MAIL

SMTP：简单邮件传输协议

SMTP 是一种提供可靠且有效电子邮件传输的协议。 SMTP 是建模在 FTP 文件传输服务上的一种邮件服务，主要用于传输系统之间的邮件信息并提供来信有关的通知。

SMTP 独立于特定的传输子系统，且只需要可靠有序的数据流信道支持。 SMTP 重要特性之一是其能跨越网络传输邮件，即“ SMTP 邮件中继”。通常，一个网络可以由公用互联网上 TCP 可相互访问的主机、防火墙分隔的 TCP/IP 网络上 TCP 可相互访问的主机，及其它 LAN/WAN 中的主机利用非 TCP 传输层协议组成。使用 SMTP ，可实现相同网络上处理机之间的邮件传输，也可通过中继器或网关实现某处理机与其它网络之间的邮件传输。

在这种方式下，邮件的发送可能经过从发送端到接收端路径上的大量中间中继器或网关主机。域名服务系统（DNS）的邮件交换服务器可以用来识别出传输邮件的下一跳 IP 地址。

协议结构

SMTP 命令是发送于 SMTP 主机之间的 ASCII 信息，可能命令如下所示：

命令	描述
DATA	开始信息写作
EXPN <string>	在指定邮件表中返回名称
HELO <domain>	返回邮件服务器身份
HELP <command>	返回指定命令中的信息
MAIL FROM <host>	在主机上初始化一个邮件会话
NOOP	除服务器响应确认以外，没有引起任何反应
QUIT	终止邮件会话
RCPT TO <user>	指明谁收到邮件
RSET	重设邮件连接
SAML FROM <host>	发送邮件到用户终端和邮箱
SEND FROM <host>	发送邮件到用户终端
SOML FROM <host>	发送邮件到用户终端或邮箱
TURN	接收端和发送端交换角色

 MSN : MingLang.Xu@Hotmail.Com

VRFY <user>	校验用户身份

POP & POP3：邮局协议(邮局协议第 3 版)

POP 协议允许工作站动态访问服务器上的邮件，目前已发展到第三版，称为 POP3 。 POP3 允许工作站检索邮件服务器上的邮件。 POP3 传输的是数据消息，这些消息可以是指令，也可以是应答。

创建一个分布式电子邮件系统有多种不同的技术支持和途径： POP （邮局协议）、 DMSP （分层式电子邮件系统协议）和 IMAP （因特网信息访问协议）。其中， POP 协议创建最早因此也最为人们了解；DMSP 具有较好的支持“无连接”操作的性能，但其很大程度上仅限于单个应用程序（PCMAIL ）； IMAP 提供了 POP 和 DMSP 的扩展集并提供对远程邮件访问的三种支持方式：离线、在线和无连接。

POP 协议支持“离线”邮件处理。其具体过程是：邮件发送到服务器上，电子邮件客户端调用邮件客户机程序以连接服务器，并下载所有未阅读的电子邮件。这种离线访问模式是一种存储转发服务，将邮件从邮件服务器端送到个人终端机器上，一般是 PC 机或 MAC 。一旦邮件发送到 PC 机或 MAC 上，邮件服务器上的邮件将会被删除。

POP3 并不支持对服务器上邮件进行扩展操作，此过程由更高级的 IMAP4 完成。 POP3 使用 TCP 作为传输协议。

协议结构

POP3 是发送在客户机和服务器间的 ASCII 信息。POP3 命令摘要：

命令	描述
USER	用户名
PASS	用户密码
STAT	服务器上的邮件信息
RETR	获取的信息数
DELE	删除的信息数
LIST	显示的信息数
TOP <messageID> <nombredelignes>	从头开始（包含协议头）打印 X 行信息

QUIT	退出 POP3 服务器

可选 POP3 命令：

APOP name digest　　AUTHORIZATION 状态有效；

TOP msg n　　TRANSACTION 状态有效；

UIDL [msg]

POP3 Replies：

+ OK

- ERR。

IMAP & IMAP4：因特网信息访问协议

因特网信息访问协议(IMAP)用于访问存储在邮件服务器系统内的电子邮件和电子公告板信息。IMAP 允许用户邮件程序如同操作本机系统一样访问远程消息存储器。可通过台式电脑远程操作保存在 IMAP 服务系统内的邮件，而不需要在计算机之间来回传输消息或文档。

创建一个分布式电子邮件系统有多种不同的技术和途径：其中有 POP （邮局协议）、 DMSP （分层式电子邮件系统协议）和 IMAP （因特网信息访问协议）。这三者中， POP 协议创建最早因此也最为人们了解； DMSP 具有较好的支持“无连接”操作的性能，但其很大程度上仅限于单个应用程序（ PCMAIL ）； IMAP 提供了 POP 和 DMSP 的扩展集并提供对远程邮件访问的三种支持方式：离线、在线和无连接。

在线方式下， IMAP 用户不用一次性地从共享服务器上收取邮件然后删除。 IMAP 采用的是交互式客户机—服务器方式，用户可以向服务器请求特定邮件的信头或主体，或者请求服务器搜索满足一定条件的邮信。收件箱中的信件标有各种状态标志（如“删除”或“已回复”），它们一直保留直到用户真正删除。在 IMAP 系统中，用户可以像在本地一样远程操作管理邮箱。根据 IMAP 客户端实现方式及系统管理员设计的邮件系统结构，用户可以在本地机器上保存邮件，或在服务器上保存邮件，也可以两者选一。

IMAP 包括了一系列操作：邮箱的建立、删除及重命名、检查新邮件、永久删除邮件、设置和清除标志、基于服务器和 MIME 的分析和搜索、有效并有选择的取回邮件属性、文本和部分内容。 IMAP 允许用户从多台计算机上访问邮件（新邮件或保存过的邮件）。对于保证电子邮件的可靠性和同时使用多台计算机的环境来说，这一特性尤其重要。

IMAP4 是 IMAP 的最新版本，其主要特征如下：

1 与因特网消息标准完全兼容，如： MIME ；

2 允许多台计算机同时访问和管理邮件；

3 允许通过低效率的文件访问协议进行访问；

4 提供对“在线”、“离线”和“无连接”三种访问方式的支持；

5 支持共享邮箱的并发访问；

6 客户端软件不需要了解服务器上的文件存储格式 。

协议结构

IMAP 主要命令:

APPEND	AUTHENTICATE	CAPABILITY	CHECK	CLOSE
COPY	CREATE	DELETE	DELETEACL	EXAMINE
EXPUNGE	FETCH	GETACL	GETQUOTA	GETQUOTAROOT
LIST	LISTRIGHTS	LOGIN	LOGOUT	LSUB
MYRIGHTS	NOOP	RENAME	SEARCH	SELECT
SETACL	SETQUOTA	STARTTLS	STATUS	STORE
SUBSCRIBE	UID	UNSELECT	UNSUBSCRIBE	X<atom>

MIME/S-MIME：多用途网际邮件扩充协议

多用途网际邮件扩充协议（MIME）是 Multipurpose Internet Mail Extensions 的缩写，说明了如何安排消息格式使消息在不同的邮件系统内进行交换。 MIME 的格式灵活，允许邮件中包含任意类型的文件。 MIME 消息可以包含文本、图象、声音、视频及其它应用程序的特定数据。具体来说， MIME 允许邮件包括：

- 单个消息中可含多个对象；
- 文本文档不限制一行长度或全文长度；
- 可传输 ASCII 以外的字符集，允许非英语语种的消息；

- 多字体消息 ；
- 二进制或特定应用程序文件 ；
- 图象、声音、视频及多媒体消息。

MIME 复合消息的目录信头设有分界标志，这个分界标志绝不可出现在消息的其它位置，而只能是在各部之间以及消息体的开始和结束处。

MIME 的安全版本 S/MIME (Secure/Multipurpose Internet Mail Extensions) 设计用来支持邮件的加密。基于 MIME 标准， S/MIME 为电子消息应用程序提供如下加密安全服务：认证、完整性保护、鉴定及数据保密等。

传统的邮件用户代理（MUA）可以使用 S/MIME 来加密发送邮件及解密接收邮件。然而， S/MIME 并不仅限于邮件的使用，它也能应用于任何可以传送 MIME 数据的传输机制，例如 HTTP 。同样， S/MIME 利用 MIME 的面向对象特征允许在混合传输系统中交换安全消息。

此外， S/MIME 还可应用于消息自动传送代理，它们使用不需任何人为操作的加密安全服务，例如软件文档签名、发送到网上的 FAX 加密等。

协议结构

MIME 邮件头字段定义如下：

实体头：=[目录 CRLF]

[编码 CRLF]

[ID CRLF]

[描述 CRLF]

*（MIME 扩展字段 CRLF）

MIME 消息头：=实体头

字段

CRLF 版本

；在 BNF 定义中声明的消息头字

；段顺序应该忽略；

MIME 局部头：=实体头

[字段]

；不从“目录”开始的任何字段

；没有具体含义，可忽略。

；在 BNF 定义中声明的消息头字

；段顺序应该忽略；

消息格式和 S/MIME 实现过程在相关文件中可以查阅。

Secuity

IPsec：IP 层协议安全结构

IPsec 在 IP 层提供安全服务，它使系统能按需选择安全协议，决定服务所使用的算法及放置需求服务所需密钥到相应位置。 IPsec 用来保护一条或多条主机与主机间、安全网关与安全网关间、安全网关与主机间的路径。

IPsec 能提供的安全服务集包括访问控制、无连接的完整性、数据源认证、拒绝重发包（部分序列完整性形式）、保密性和有限传输流保密性。因为这些服务均在 IP 层提供，所以任何高层协议均能使用它们，例如 TCP 、 UDP 、ICMP 、 BGP 等等。

这些目标是通过使用两大传输安全协议，头部认证（AH） 和封装安全负载 （ESP），以及密钥管理程序和协议的使用来完成的。所需的 IPsec 协议集内容及其使用的方式是由用户、应用程序、和/或站点、组织对安全和系统的需求来决定。

当正确的实现、使用这些机制时，它们不应该对不使用这些安全机制保护传输的用户、主机和其他英特网部分产生负面的影响。这些机制也被设计成算法独立的。这种模块性允许选择不同的算法集而不影响其他部分的实现。例如：如果需要，不同的用户通讯可以采用不同的算法集。

定义一个标准的默认算法集可以使得全球因英特网更容易协同工作。这些算法辅以 IPsec 传输保护和密钥管理协议的使用为系统和应用开发者部署高质量的因特网层的加密的安全技术提供了途径。

协议结构 — IPsec：IP 网络安全结构

IPsec 结构包括众多协议和算法。这些协议之间的相互关系如下所示：

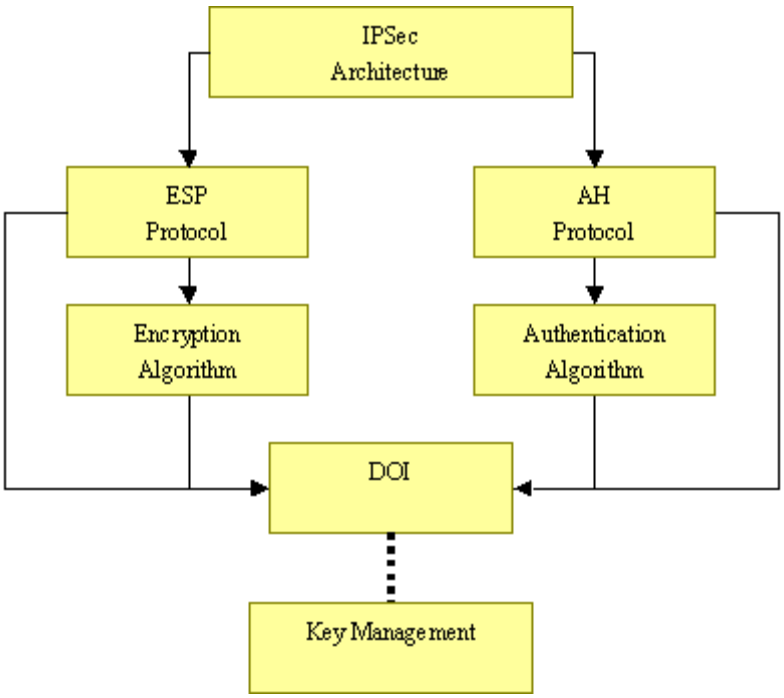


图 2 — 5 IPsec：IP 层协议安全结构

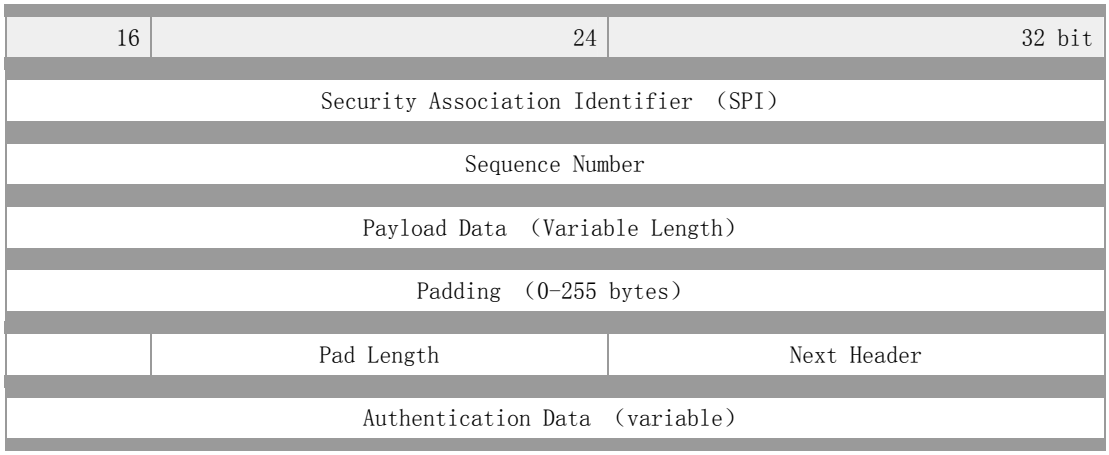
IPsec ESP：IPsec 封装安全负载

IPsec 封装安全负载（IPsec ESP）是 IPsec 体系结构中的一种主要协议，其主要设计来在 IPv4 和 IPv6 中提供安全服务的混合应用。 IPsec ESP 通过加密需要保护的数据以及在 IPsec ESP 的数据部分放置这些加密的数据来提供机密性和完整性。根据用户安全要求，这个机制既可以用于加密一个传输层的段（如： TCP、UDP、ICMP、IGMP），也可以用于加密一整个的 IP 数据报。封装受保护数据是非常必要的，这样就可以为整个原始数据报提供机密性。

ESP 头可以放置在 IP 头之后、上层协议头之前 （传送层），或者在被封装的 IP 头之前 （隧道模式）。IANA 分配给 ESP 一个协议数值 50 ，在 ESP 头前的协议头总是在“ next head ”字段（IPv6）或“协议”（IPv4）字段里包含该值 50 。 ESP 包含一个非加密协议头，后面是加密数据。该加密数据既包括了受保护的 ESP 头字段也包括了受保护的用户数据，这个用户数据可以是整个 IP 数据报，也可以是 IP 的上层协议帧（如： TCP 或 UDP） 。

ESP 提供机密性、数据源认证、无连接的完整性、抗重播服务（一种部分序列完整性的形式） 和有限信息流机密性。所提供服务集由安全连接（SA）建立时选择的选项和实施的布置来决定，机密性的选择与所有其他服务相独立。但是，使用机密性服务而不带有完整性 / 认证服务（在 ESP 或者单独在 AH 中）可能使传输受到某种形式的攻击以破坏机密性服务。数据源验证和无连接的完整性是相互关联的服务，它们作为一个选项与机密性 （可选择的）结合提供给用户。只有选择数据源认证时才可以选择抗重播服务，由接收方单独决定抗重播服务的选择。

协议结构



- Security Association Identifier — 一个伪随机值，用于识别数据报的安全联接（Security Association）。
- Sequence Number — 包含无变化的增长计数器值,该值是强制性的,即使接收端不为特定 SA 提供 Anti-Replay 服务，它仍然存在。
- Payload Data — 一个可变长字段，包括 Next Header 字段中描述的数据。
- Padding — 供加密使用。
- Pad Length — 指出 Pad 字节前的号码。
- Next Header — 识别包含在有效负载数据字段中的数据类型。如 IPv6 中的扩展头或上层协议标识符。
- Authentication Data — 一个可变长字段，包括在 ESP 数据包上计算的减去 Authentication Data 的完整校验值（ICV）。

IPsec AH：IPsec 认证头协议

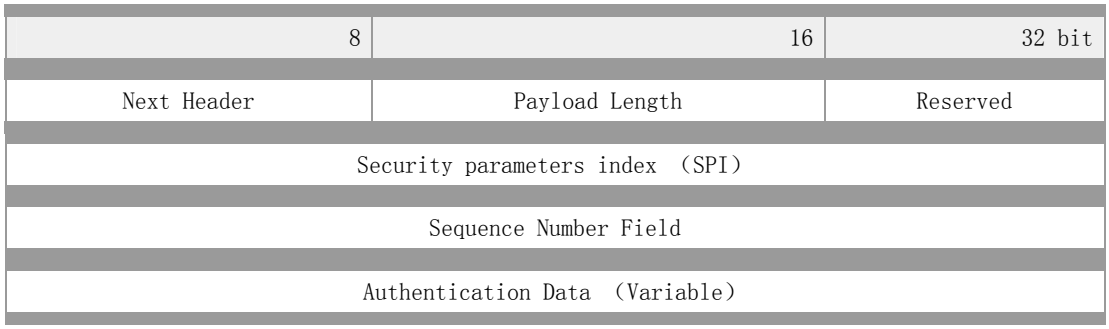
IPsec 认证头协议（IPsec AH）是 IPsec 体系结构中的一种主要协议，它为 IP 数据报提供无连接完整性与数据源认证，并提供保护以避免重播情况。一旦建立安全连接，接收方就可能会选择后一种服务。 AH 尽可能为 IP 头和上层协议数据提供足够多的认证。但是，在传输过程中某些 IP 头字段会发生变化，且发送方无法预测当数据包到达接受端时此字段的值。 AH 并不能保护这种字段值。因此， AH 提供给 IP 头的保护有些是零碎的。

AH 可被独立使用，或与 IP 封装安全负载（ESP）相结合使用，或通过使用隧道模式的嵌套方式。在通信主机与通信主机之间、通信安全网关与通信安全网关之间或安全网关与主机之间可以提供安全服务。 ESP 提供了相同的安全服务并提供了一种保密性（加密）服务，而 ESP 与 AH 各自提供的认证其

根本区别在于它们的覆盖范围。特别地，不是由 ESP 封装的 IP 头字段则不受 ESP 保护。有关在不同网络环境下如何使用 AH 和 ESP 的详细内容，可参见相关文件。

通常，当用与 IPv6 时，AH 出现在 IPv6 逐跳路由头之后 IPv6 目的选项之前。而用于 IPv4 时，AH 跟随主 IPv4 头。

协议结构



- Next Header — 识别认证头面后的下一个有效负载的类型。
- Payload Length — 规定 AH 的长（32 位字，4-字节单元），减去“2”）
- SPI — 专有 32 位值，与目的 IP 地址和安全协议（AH）相结合，唯一识别数据报的安全联接（Security Association）。
- Sequence Number — 包含无变化的增长计数器值，该值是强制性的，即使接收端不为特定 SA 提供 Anti-Replay 服务，它仍然存在。
- Authentication Data — 一个可变长字段，包括在 ESP 数据包上计算的减去 Authentication Data 的完整校验值（ICV）。

IPsec IKE：Internet 密钥交换协议

Internet 密钥交换（IPsec IKE）是 IPsec 体系结构中的一种主要协议。它是一种混合协议，使用部分 Oakley 和部分 SKEME， 并协同 ISAKMP 提供密钥生成材料和其它安全连系，比如用于 IPsec DOI 的 AH 和 ESP 。

ISAKMP 只对认证和密钥交换提出了结构框架，但没有具体定义。 ISAKM 与密钥交换相独立，支持多种不同的密钥交换。 IKE 是一系列密钥交换中的一种，称为“模式”。

 MSN : MingLang.Xu@Hotmail.Com

IKE 可用于协商虚拟专用网（VPN），也可用于远程用户（其 IP 地址不需要事先知道）访问安全主机或网络，支持客户端协商。客户端模，式即为协商方不是安全连接发起的终端点。当使用客户模式时，端点处身份是隐藏的。

IKE 的实施必须支持以下的属性值：

- DES 用在 CBC 模式，使用弱、半弱、密钥检查。
- MD5[MD5] 和 SHA[SHA] 。
- 通过预共享密钥进行认证。
- 缺省的组 1 上的 MODP 。

另外， IKE 的实现也支持 3DES 加密；用 Tiger[TIGER] 作为 hash ；数字签名标准， RSA[RSA] ，使用 RSA 公共密钥加密的签名和认证；以及使用组 2 进行 MODP 。IKE 实现可以支持其它的加密算法，并且可以支持 ECP 和 EC2N 组。

只要实现了 IETF IPsec DOI ， IKE 模式就必须实施。其它 DOI 也可使用这里描述的模式。

协议结构

IKE 信息是由 ISAKMP 头和 SKEME 以及 Oakley 字段联合构成。其特定格式取决于信息状态和模式。具体细节，请参照相关链路文档。

IPsec ISAKMP：Internet 安全连接和密钥管理协议

（ISAKMP：Internet Security Association and Key Management Protocol）

Interne 安全连接和密钥管理协议（ISAKMP）是 IPsec 体系结构中的一种主要协议。该协议结合认证、密钥管理和安全连接等概念来建立政府、商家和因特网上的私有通信所需要的安全。

因特网安全联盟和密钥管理协议（ISAKMP）定义了程序和信息包格式来建立，协商，修改和删除安全连接（SA）。 SA 包括了各种网络安全服务执行所需的所有信息，这些安全服务包括 IP 层服务（如头认证和负载封装）、传输或应用层服务，以及协商流量的自我保护服务等。 ISAKMP 定义包括交换密钥生成和认证数据的有效载荷。这些格式为传输密钥和认证数据提供了统一框架，而它们与密钥产生技术，加密算法和认证机制相独立。

 MSN : MingLang.Xu@Hotmail.Com

ISAKMP 区别于密钥交换协议是为了把安全连接管理的细节从密钥交换的细节中彻底的分离出来。不同的密钥交换协议中的安全属性也是不同的。然而，需要一个通用的框架用于支持 SA 属性格式，谈判，修改与删除 SA ， ISAKMP 即可作为这种框架。

把功能分离为三部分增加了一个完全的 ISAKMP 实施安全分析的复杂性。然而在有不同安全要求且需协同工作的系统之间这种分离是必需的,而且还应该对 ISAKMP 服务器更深层次发展的分析简单化。

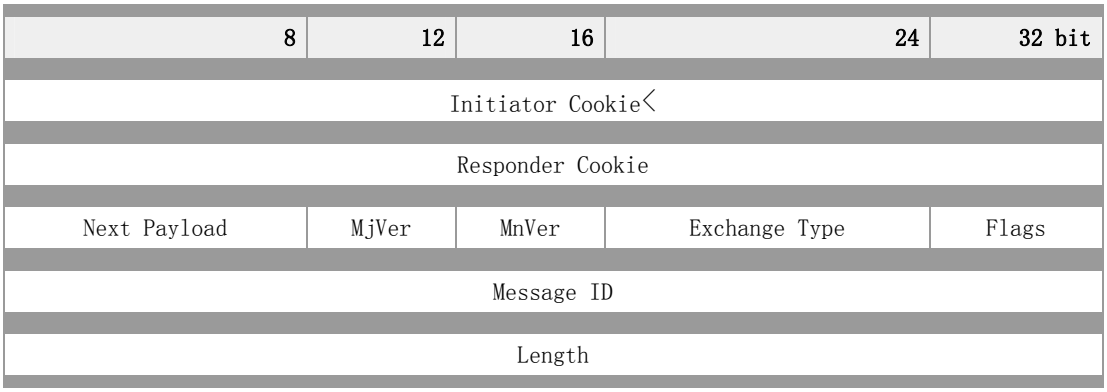
ISAKMP 支持在所有网络层的安全协议 （如： IPSEC、TLS、TLSP、OSPF 等等）的 SA 协商。 ISAKMP 通过集中管理 SA 减少了在每个安全协议中重复功能的数量。 ISAKMP 还能通过一次对整个栈协议的协商来减少建立连接的时间。

ISAKMP 中，解释域（DOI）用来组合相关协议，通过使用 ISAKMP 协商安全连接。共享 DOI 的安全协议从公共的命名空间选择安全协议和加密转换方式，并共享密钥交换协议标识。同时它们还共享一个特定 DOI 的有效载荷数据目录解释，包括安全连接和有效载荷认证。

总之， ISAKMP 关于 DOI 定义如下方面：

- 特定 DOI 协议标识的命名模式；
- 位置字段解释；
- 可应用安全策略集；
- 特定 DOI SA 属性语法；
- 特定 DOI 有效负载目录语法；
- 必要情况下，附加密钥交换类型；
- 必要情况下，附加通知信息类型。

协议结构



- Initiator Cookie — Initiator Cookie：启动 SA 建立、SA 通知或 SA 删除的实体 Cookie。
- Responder Cookie — Responder Cookie：响应 SA 建立、SA 通知或 SA 删除的实体 Cookie。
- Next Payload — 信息中的 Next Payload 字段类型。
- Major Version — 使用的 ISAKMP 协议的主要版本。
- Minor Version — 使用的 ISAKMP 协议的次要版本。

 MSN : MingLang.Xu@Hotmail.Com

- Exchange Type — 正在使用的交换类型。
- Flags — 为 ISAKMP 交换设置的各种选项。
- Message ID — 唯一的信息标识符，用来识别第 2 阶段的协议状态。
- Length — 全部信息（头＋有效载荷）长（八位）。

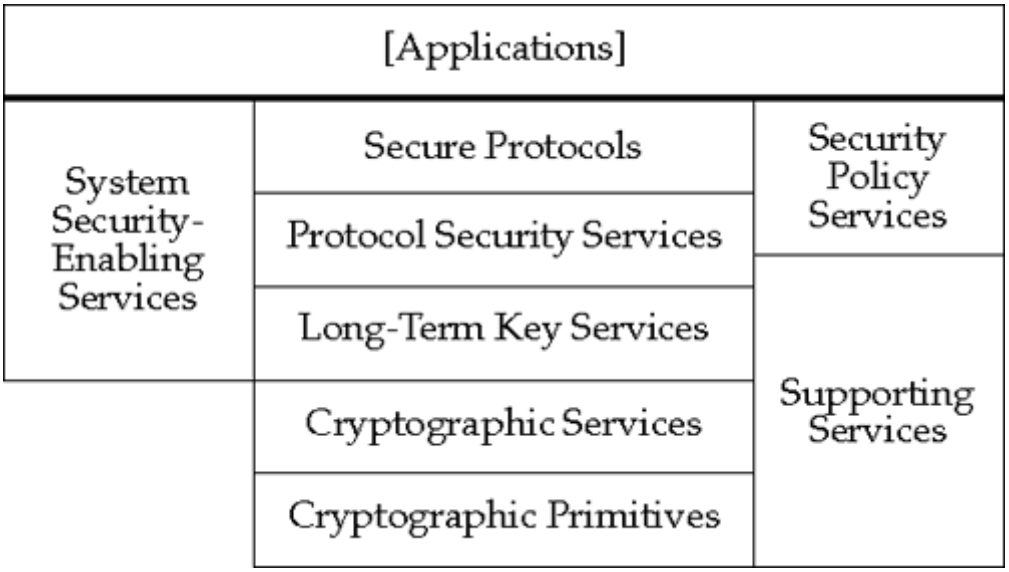
公共密钥基础设施

公共密钥基础设施（PKI）是基于公共密钥加密（Public Key Cryptography）概念、提供公共密钥创建和管理的系统，它支持用户高效实现数据加密和密钥交换过程。

PKI 体系由 Internet 标准组和美国国家标准与技术协会（NIST）制定。在 PKI 结构中，包括一个关键的认证机构，即认证中心 CA（属于第三方组织），它是一个负责发放和管理数字证书的权威机构。PKI 体系主要支持以下功能部分：

- 系统安全增强服务 — 允许建立用户身份并与他们在 PKI 系统中操作联系起来。
- 加密原语和服务 — 支持加密功能，公共密钥安全性正是建立在该功能基础之上，包括私人密钥原语，如国际数据加密算法（IDEA）。
- 长期密钥服务 — 允许用户管理自己的长期密钥和认证，以及检索和校验其他主要认证的有效性。
- 协议安全服务 — 提供安全功能，如数据起源认证、数据完整性保护、数据私有保护以及认可。
- 安全协议 — 增强未知安全性或已知有限安全性的应用程序内通信的安全性。
- 安全策略服务 — 通过安全策略信息，增强访问控制，并指导访问控制，策略强制校验已知安全应用程序的效能。
- 支持服务 — 安全操作（但不是安全策略强制功能）方面的功能

PKI 体系：



PKI 体系

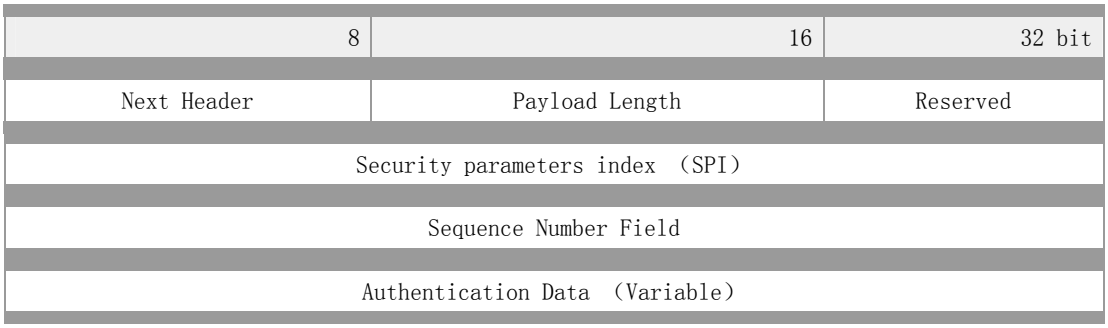
IPsec AH：IPsec 认证头协议

IPsec 认证头协议（IPsec AH）是 IPsec 体系结构中的一种主要协议，它为 IP 数据报提供无连接完整性与数据源认证，并提供保护以避免重播情况。一旦建立安全连接，接收方就可能会选择后一种服务。 AH 尽可能为 IP 头和上层协议数据提供足够多的认证。但是，在传输过程中某些 IP 头字段会发生变化，且发送方无法预测当数据包到达接受端时此字段的值。 AH 并不能保护这种字段值。因此， AH 提供给 IP 头的保护有些是零碎的。

AH 可被独立使用，或与 IP 封装安全负载（ESP）相结合使用，或通过使用隧道模式的嵌套方式。在通信主机与通信主机之间、通信安全网关与通信安全网关之间或安全网关与主机之间可以提供安全服务。 ESP 提供了相同的安全服务并提供了一种保密性（加密）服务，而 ESP 与 AH 各自提供的认证其根本区别在于它们的覆盖范围。特别地，不是由 ESP 封装的 IP 头字段则不受 ESP 保护。有关在不同网络环境下如何使用 AH 和 ESP 的详细内容，可参见相关文件。

通常，当用与 IPv6 时， AH 出现在 IPv6 逐跳路由头之后 IPv6 目的选项之前。 而用于 IPv4 时，AH 跟随主 IPv4 头。

协议结构



- Next Header — 识别认证头面后的下一个有效负载的类型。
- Payload Length — 规定 AH 的长（32 位字，4-字节单元），减去“2”）
- SPI — 专有 32 位值，与目的 IP 地址和安全协议（AH）相结合，唯一识别数据报的安全联接（Security Association）。
- Sequence Number — 包含无变化的增长计数器值，该值是强制性的，即使接收端不为特定 SA 提供 Anti-Replay 服务，它仍然存在。
- Authentication Data — 一个可变长字段，包括在 ESP 数据包上计算的减去 Authentication Data 的完整校验值（ICV）。

 MSN : MingLang.Xu@Hotmail.Com

网络安全包括了网络通信安全、信息在网络传输中的保密性和完整性、控制访问受限网域与敏感信息以及在公共网络如因特网上使用隐秘通讯。为了解决这些问题，各大组织及技术供应商纷纷推出了各种网络和信息安全技术。其技术概要如下 ：

AAA ：授权、认证和计费是一种提供网络资源访问智能控制、执行策略、使用审核以及必要的服务信息费用等的技术。认证提供识别用户的方式，即在允许访问之前，用户要提供有效的用户名和有效密码。授权处理主要决定用户是否有访问特定信息或者一些网络子域的权限。计费即计算用户消耗的网络资源，包括系统时间长短或用户在一个会话中发送、接收数据的数量，这些可以用于授权控制、记帐、趋向分析、资源利用以及容量计划等。一个专门的 AAA 服务器或执行这些功能的程序通常可以提供：授权、认证和计费服务。

VPN ：虚拟专用网络是一种允许企业或个人私有通信的技术，例如，远程访问公司网络或者使用公共电信网络，如因特网。虚拟专用网也可以是一个在公共互联网络上配置出来的仅供一个组织使用的网络。现在可以通过各种网络隧道技术如 L2TP 实现这一目标。使用像 IPsec 等加密技术可以提高公用或虚拟专用网络上的信息保密性。

防火墙：防火墙可以是一个软件程序或硬件设备，用于过滤通过因特网联接进入到内部网络或计算机系统的信息。防火墙使用三种方法中的一种或更多来控制进出网络的通信。

- 包过滤－－数据包根据过滤器进行分析。通过过滤波的包发送至请求系统，其它的则丢弃。
- 代理服务－－防火墙获取来自因特网的信息，并发送至请求系统，反之亦然。
- 状态检查－－将某些传送包的主要部分与可靠信息数据库作比较。监视防火墙中外出信息的特征，然后将进入信息与这些特征作比较。如果比较结果满足合理的匹配，就允许信息通过，否则丢弃。

主要协议

AAA 认证、授权、计费 （Authentication、Authorization、Accounting）

- Kerberos：网络认证协议（Kerberos：Network Authentication Protocol）
- RADIUS：远程用户拨入认证系统（RADIUS：Remote Authentication Dial In User Service）
- SSH：安装外壳协议（SSH：Secure Shell Protocol）

隧道技术 （Tunneling）

- L2F：第二层转发协议（L2F：Level 2 Forwarding protocol）
- L2TP：第二层隧道协议（L2TP：Layer 2 Tunneling Protocol）

 MSN : MingLang.Xu@Hotmail.Com

- PPTP：点对点隧道协议（PPTP：Point-to-Point Tunneling Protocol）

安全路由选择 （Secured Routing）

- DiffServ：区分服务体系结构（DiffServ：Differentiated Service）
- GRE：通用路由封装（GRE：Generic Routing Encapsulation）
- IPsec：IP层协议安全结构（IPsec：Security Architecture for IP network）
- IPsec AH：IPsec认证头协议（IPsec AH：IPsec Authentication Header）
- IPsec ESP：IPsec封装安全负载（IPsec ESP：IPsec Encapsulating Security Payload）
- IPsec IKE：Internet密钥交换（IPsec IKE：Internet Key Exchange Protocol）
- IPsec ISAKMP：Internet安全连接和密钥管理协议（IPsec ISAKMP：Internet Security Association and Key Management Protocol）
- TLS：安全传输层协议（TLS：Transport Layer Security Protocol

RADIUS：远程用户拨号认证系统

RADIUS 是一种在网络接入服务器（Network Access Server）和共享认证服务器间传输认证、授权和配置信息的协议。 RADIUS 使用 UDP 作为其传输协议。此外 RADIUS 也负责传送网络接入服务器和共享计费服务器间的计费信息。

RADIUS 主要特征如下：

- 客户 / 服务器模式：网络接入服务器作为 RADIUS 的客户端，负责将用户信息传递给指定的 RADIUS 服务器，然后根据返回信息进行操作。 RADIUS 服务器负责接收用户连接请求，认证用户后，返回所有必要的配置信息以便客户端为用户提供服务。 RADIUS 服务器可以作为其他 RADIUS 服务器或认证服务器的代理。
- 网络安全：客户端与 RADIUS 记帐服务器之间的通信是通过共享密钥的使用来鉴别的，这个共享密钥不会通过网络传送。此外，任何用户口令在客户机和 RADIUS 服务器间发送时都需要进行加密过程，以避免有人通过嗅探非安全网络可得到用户密码。
- 灵活认证机制： RADIUS 服务器支持多种用户认证方法。当用户提供了用户名和原始口令后，RADIUS 服务器可支持 PPP PAP 或 CHAP, UNIX 登录和其它认证机制。
- 协议的可扩充性：所有的事务都是由不同长度的“属性－长度－值”的三元组构成的。新的属性值的加入不会影响到原有协议的执行。

 MSN : MingLang.Xu@Hotmail.Com

协议结构

8	16	32 bit
Code	Identifier	Length
Authenticator （16 bytes）		

- Code — 信息类型如下所述：1、请求访问（Access-Request）；2、接收访问（Access-Accept）；3、拒绝访问（Access-Reject）；4、计费请求（Accounting-Request）；5、计费响应（Accounting-Response）；11、挑战访问（Access-Challenge）；12、服务器状况（Status-Server — Experimental）；13、客户机状况（Status-Client — Experimental）；255、预留（Reserved）
- Identifier — 匹配请求和响应的标识符。
- Length — 信息大小，包括头部。
- Authenticator — 该字段用来识别 RADIUS 服务器和隐藏口令算法中的答复

SSH：安全外壳协议

安全外壳协议（SSH）是一种在不安全网络上提供安全远程登录及其它安全网络服务的协议。SSH 主要有三部分组成：

传输层协议［SSH-TRANS］提供了服务器认证，保密性及完整性。此外它有时还提供压缩功能。SSH-TRANS 通常运行在 TCP/IP 连接上，也可能用于其它可靠数据流上。SSH-TRANS 提供了强力的加密技术、密码主机认证及完整性保护。该协议中的认证基于主机，并且该协议不执行用户认证。更高层的用户认证协议可以设计为在此协议之上。

用户认证协议［SSH-USERAUTH］用于向服务器提供客户端用户鉴别功能。它运行在传输层协议 SSH-TRANS 上面。当 SSH-USERAUTH 开始后，它从低层协议那里接收会话标识符（从第一次密钥交换中的交换哈希 H ）。会话标识符唯一标识此会话并且适用于标记以证明私钥的所有权。SSH-USERAUTH 也需要知道低层协议是否提供保密性保护。

连接协议［SSH-CONNECT］将多个加密隧道分成逻辑通道。它运行在用户认证协议上。它提供了交互式登录话路、远程命令执行、转发 TCP/IP 连接和转发 X11 连接。

一旦建立一个安全传输层连接，客户机就发送一个服务请求。当用户认证完成之后，会发送第二个服务请求。这样就允许新定义的协议可以与上述协议共存。连接协议提供了用途广泛的各种通道，有标准的方法用于建立安全交互式会话外壳和转发（“隧道技术”）专有 TCP/IP 端口和 X11 连接。

协议结构

安全外壳协议（SSH）拥有很多信息，每个信息可能具有不同的格式，关于具体的信息格式，请参照相关文档。

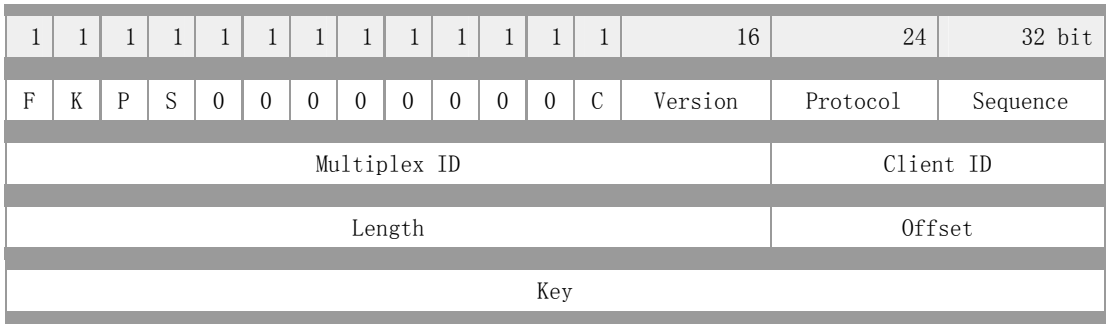
L2F：第二层转发协议

第二层转发协议（L2F）用于建立跨越公共网络（如因特网）的安全隧道来将 ISP POP 连接到企业内部网关。这个隧道建立了一个用户与企业客户网络间的虚拟点对点连接。

第二层转发协议（L2F）允许高层协议的链路层隧道技术。使用这样的隧道，使得把原始拨号服务器位置和拨号协议连接终止与提供的网络访问位置分离成为可能。

L2F 允许在其中封装 PPP/SLIP 包。ISP NAS 与家庭网关都需要共同了解封装协议，这样才能在因特网上成功地传输或接收 SLIP/PPP 包。

协议结构



1	1	1	1	1	1	1	1	1	1	1	1	16	24	32 bit
F	K	P	S	0	0	0	0	0	0	0	0	C	Version	Protocol	Sequence
Multiplex ID													Client ID		
Length													Offset		
Key															

- Version — 用于创建数据包的 L2F 软件的主修版本。
- Protocol — 协议字段，规定 L2F 数据包中传送的协议。
- Sequence — 当 L2F 头部的 S 位设置为 1 时的当前序列号。
- Multiplex ID — 数据包 Multiplex ID 用于识别一个隧道中的特殊链接。
- Client ID — Client ID （CLID）支持解除复用隧道中的终点。
- Length — 整个数据包的长度大小（八位形式），包括头、所有字段以及有效负载。
- Offset — 该字段规定通过 L2F 协议头的字节数，协议头是有效负载数据起始位置。如果 L2F 头部的 F 位设置为 1 时，就会有该字段出现。
- Key — Key 字段出现在将 K 位设置在 L2F 协议头的情况。这属于认证过程。
- Checksum — 数据包的校验和。Checksum 字段出现在 L2F 协议头中的 C 位设置为 1 的情况。

L2TP：第二层隧道协议

第二层隧道协议（L2TP）是用来整合多协议拨号服务至现有的因特网服务提供商点。 PPP 定义了多协议跨越第二层点对点链接的一个封装机制。特别地，用户通过使用众多技术之一（如：拨号 POTS、ISDN、ADSL 等）获得第二层连接到网络访问服务器（NAS），然后在此连接上运行 PPP 。在这样的配置中，第二层终端点和 PPP 会话终点处于相同的物理设备中（如： NAS）。

L2TP 扩展了 PPP 模型，允许第二层和 PPP 终点处于不同的由包交换网络相互连接的设备来。通过 L2TP ，用户在第二层连接到一个访问集中器（如：调制解调器池、 ADSL DSLAM 等），然后这个集中器将单独得的 PPP 帧隧道到 NAS 。这样，可以把 PPP 包的实际处理过程与 L2 连接的终点分离开来。

对于这样的分离，其明显的一个好处是，L2 连接可以在一个（本地）电路集中器上终止，然后通过共享网络如帧中继电路或英特网扩展逻辑 PPP 会话，而不用在 NAS 上终止。从用户角度看，直接在 NAS 上终止 L2 连接与使用 L2TP 没有什么功能上的区别。 L2TP 协议也用来解决“多连接联选组分离”问题。多链接 PPP ，一般用来集中 ISDN B 通道，需要构成多链接捆绑的所有通道在一个单网络访问服务器（NAS）上组合。因为 L2TP 使得 PPP 会话可以出现在接收会话的物理点之外的位置，它用来使所有的通道出现在单个的 NAS 上，并允许多链接操作，即使是在物理呼叫分散在不同物理位置的 NAS 上的情况下。

L2TP 使用以下两种信息类型，即控制信息和数据信息。控制信息用于隧道和呼叫的建立、维持和清除。数据信息用于封装隧道所携带的 PPP 帧。控制信息利用 L2TP 中的一个可靠控制通道来确保发送。当发生包丢失时，不转发数据信息 。

协议结构

L2TP 命令头：

<table>
<tr><th colspan="12"></th><th>16</th><th>32 bit</th></tr>
<tr><td>T</td><td>L</td><td>X</td><td>X</td><td>S</td><td>X</td><td>O</td><td>P</td><td>X</td><td>X</td><td>X</td><td>X</td><td>VER</td><td>Length</td></tr>
<tr><td colspan="13">Tunnel ID</td><td>Session ID</td></tr>
<tr><td colspan="13">Ns （opt）</td><td>Nr （opt）</td></tr>
<tr><td colspan="13">Offset Size （opt）</td><td>Offset Pad （opt）</td></tr>
</table>

- T — T 位表示信息类型。若是数据信息，该值为 0；若是控制信息，该值为 1。
- L — 当设置该字段时，说明 Length 字段存在，表示接收数据包的总长。对于控制信息，必须设置该值。
- X — X 位为将来扩张预留使用。在导出信息中所有预留位被设置为 0，导入信息中该值忽略。
- S — 如果设置 S 位，那么 Nr 字段和 Ns 字段都存在。对于控制信息，S 位必须设置。
- O — 当设置该字段时，表示在有效负载信息中存在 Offset Size 字段。对于控制信息，该字段值设为 0。
- P — 如果 Priority （P）位值为 1，表示该数据信息在其本地排队和传输中将会得到优先处理。

- Ver — Ver 位的值总为 002。它表示一个版本 1 L2TP 信息。
- Length — 信息总长，包括头、信息类型 AVP 以及另外的与特定控制信息类型相关的 AVPs。
- Tunnel ID — 识别控制信息应用的 Tunnel。如果对等结构还没有接收到分配的 Tunnel ID，那么 Tunnel ID 必须设置为 0。一旦接收到分配的 Tunnel ID，所有更远的数据包必须和 Tunnel ID 一起被发送。
- Call ID — 识别控制信息应用的 Tunnel 中的用户会话。如果控制信息在 Tunnel 中不应用单用户会话（例如，一个 Stop-Control-Connection-Notification 信息），Call ID 必须设置为 0。
- Nr — 期望在下一个控制信息中接收到的序列号。
- Ns — 数据或控制信息的序列号。
- Offset Size & Pad — 该字段规定通过 L2F 协议头的字节数，协议头是有效负载数据起始位置。Offset Padding 中的实际数据并没有定义。如果 Offset 字段当前存在，那么 L2TP 头 Offset Padding 的最后八位字节后结束。

PPTP：点对点隧道协议

点对点隧道协议（PPTP）是一种支持多协议虚拟专用网络的网络技术。通过该协议，远程用户能够通过 Microsoft Windows NT? 工作站、 Windows? 95 和 Windows 98 操作系统以及其它装有点对点协议的系统安全访问公司网络，并能拨号连入本地 ISP ，通过 Internet 安全链接到公司网络。

PPTP 可以用于在 IP 网络上建立 PPP 会话隧道。在这种配置下， PPTP 隧道和 PPP 会话运行在两个相同的机器上，呼叫方充当 PNS 。 PPTP 使用客户机－服务器结构来分离当前网络访问服务器具备的一些功能并支持虚拟专用网络。PPTP 作为一个呼叫控制和管理协议，它允许服务器控制来自 PSTN 或 ISDN 的拨入电路交换呼叫访问并初始化外部电路交换连接。

PPTP 只能通过 PAC 和 PNS 来实施，其它系统没有必要知道 PPTP 。拨号网络可与 PAC 相连接而无需知道 PPTP 。标准的 PPP 客户机软件可继续在隧道 PPP 链接上操作。

PPTP 使用 GRE 的扩展版本来传输用户 PPP 包。这些增强允许为在 PAC 和 PNS 之间传输用户数据的隧道提供低层拥塞控制和流控制。这种机制允许高效使用隧道可用带宽并且避免了不必要的重发和缓冲区溢出。 PPTP 没有规定特定的算法用于低层控制，但它确实定义了一些通信参数来支持这样的算法工作。

协议结构

16	32 bit
Length	PPTP Message Type
Magic Cookie	

Control Message Type	Reserved 0
Protocol Version	Reserved 1
Framing Capability	
Bearing Capability	
Maximum Channels	Firmware Revision
Host Name （64 Octets）	
Vendor String （64 Octets）	

- Length — 该 PPTP 信息的八位总长，包括整个 PPTP 头。
- PPTP Message Type — 信息类型。可能值有：1、控制信息；2、管理信息。
- Magic Cookie — Magic Cookie 以连续的 0x1A2B3C4D 进行发送，其基本目的是确保接收端与 TCP 数据流间的正确同步运行。
- Control Message Type — 可能值有：1、开始－控制－链接－请求（Start-Control-Connection-Request）；2、开始－控制－链接－答复（Start-Control-Connection-Reply）；3、停止－控制－链接－请求（Stop-Control-Connection-Request）；4、停止－控制－链接－答复（Stop-Control-Connection-Reply）；5、回音－请求（Echo-Request）；6、回音－答复（Echo-Reply）；
- Call Management－－可能值有：1、导出－呼叫－请求（Outgoing-Call-Request）；2、导出－呼叫－答复（Outgoing-Call-Reply）；3、导入－呼叫－请求（Incoming-Call-Request）；4、导入－呼叫－答复（Incoming-Call-Reply）；5、导入－呼叫－链接（Incoming-Call-Connected）；6、呼叫－清除－请求（Call-Clear-Request）；7、呼叫－断开链接－通告（Call-Disconnect-Notify）；8、广域网－错误－通告（WAN-Error-Notify）。
- PPP Session Control — 设置－链路－信息（Set-Link-Info）。
- Reserved 0 & 1 — 必须设置为 0。
- Protocol Version — PPTP 版本号。
- Framing Capabilities — 指出帧类型，该信息发送方可以提供：1、异步帧支持（Asynchronous Framing Supported）；2、同步帧支持（Synchronous Framing Supported）。
- Bearer Capabilities — 指出承载性能，该信息发送方可以提供：1、模拟访问支持（Analog Access Supported）；2、数字访问支持（Digital access supported）。
- Maximum Channels — 该 PAC 可以支持的个人 PPP 会话总数。
- Firmware Revision — 若由 PAC 出发，则包括发出 PAC 时的固件修订本编号；若由 PNS 出发，则包括 PNS PPTP 驱动版本。
- Host Name — 包括发行的 PAC 或 PNS 的 DNS 名称。
- Vendor Name — 包括特定供应商字串，指当请求是由 PNS 提出时，使用的 PAC 类型或 PNS 软件类型。

DiffServ：区分服务体系结构

区分服务体系结构（DiffServ）定义了一种可以在互联网上实施可扩展的服务分类的体系结构。一种“服务”，是由在一个网络内，在同一个传输方向上，通过一条或几条路径传输数据包时的某些重要特征所定义的。这些特征可能包括吞吐率、时延、时延抖动，和/或丢包率的量化值或统计值等，也可能是指其获取网络资源的相对优先权。服务分类要求能适应不同应用程序和用户的需求，并且允许对互联网服务的分类收费。

DiffServ 体系结构由许多在网络节点上实现的功能要素组成，包括每一跳转发小集合，数据包归类功能，和交通调节功能。其中，交通调节功能又包含测量、标记、整形、和监察策略四部分。在本体系结构，只在网络的边界节点上实现复杂的分类和调节功能，并且，通过在 IPv4 和 IPv6 包头的 DS 段做适当的标记 [DSFIELD]，聚合流量，然后根据所做的标记，采取不同的每一跳转发策略。因此，本体系结构具备可扩展性。“每一跳行为”保证了在互相竞争资源的数据流中为每个网络节点分配缓冲区和带宽资源时，有一个合理的处理力度。在核心网络节点上，无需维护每个应用程序流或每个用户转发状态。

分类服务体系结构基于这样一个简单模型：进入网络的流量在网络边缘处进行分类和可能的调节，然后被分配到不同的行为集合中去。每一个行为集合由唯一的 DS 编码点标识。在网络核心处，数据包根据 DS 编码点对应的每一跳行为转发。在本节中，我们讨论在分类服务区域中的关键组件，流量分类和调节功能，以及分类服务是如何通过流量调节和基于 PHB 的转发而实现的。

协议结构

在 DiffServ 中，定义了一个替换头字段，称为 DS 字段，用来取代现有的 IPv4 TOS(Octet)和 IPv6 Traffic Class（Octet）。其格式如下所示：

6	8 bit
DSCP	CU

- DSCP — 即区分服务代码点，用于选择 PHB。
- CU — 当前尚未使用。

GRE：通用路由封装

（GRE: Generic

Routing
Encapsulation)

通用路由封装（GRE）定义了在任意一种网络层协议上封装任意一个其它网络层协议的协议。

在大多数常规情况下，系统拥有一个有效载荷（或负载）包，需要将它封装并发送至某个目的地。首先将有效载荷封装在一个 GRE 包中，然后将此 GRE 包封装在其它某协议中并进行转发。此外发协议即为发送协议。当 IPv4 被作为 GRE 有效载荷传输时，协议类型字段必须被设置为 0x800 。当一个隧道终点拆封此含有 IPv4 包作为有效载荷的 GRE 包时， IPv4 包头中的目的地址必须用来转发包，并且需要减少有效载荷包的 TTL 。值得注意的是，在转发这样一个包时，如果有效载荷包的目的地址就是包的封装器（也就是隧道另一端），就会出现回路现象。在此情形下，必须丢弃该包。当 GRE 包被封装在 IPv4 中时，需要使用 IPv4 协议 47 。

GRE 下的网络安全与常规的 IPv4 网络安全是较为相似的，GRE 下的路由采用 IPv4 原本使用的路由，但路由过滤保持不变 。包过滤要求防火墙检查 GRE 包，或者在 GRE 隧道终点完成过滤过程。在那些这被看作是安全问题的环境下，可以在防火墙上终止隧道。

协议结构

1	13	16	32 bit
C	Reserved 0 & 1	Ver	Protocol Type
Checksum （optional）			Reserved

- C — 当前校验和。
- Reserved 0 & 1 — 预留以备后用。
- Ver — 版本号，当前为 0。
- Protocol Type — 包括有效载荷数据包的协议类型。
- Checksum — 包括 GRE 头和有效负载数据包中所有 16 位字的 IP 校验和总数

IPsec ISAKMP：Internet 安全连接和密钥管理协议

Interne 安全连接和密钥管理协议（ISAKMP）是 IPsec 体系结构中的一种主要协议。该协议结合认证、密钥管理和安全连接等概念来建立政府、商家和因特网上的私有通信所需要的安全。

因特网安全联盟和密钥管理协议（ISAKMP）定义了程序和信息包格式来建立，协商，修改和删除安全连接（SA）。 SA 包括了各种网络安全服务执行所需的所有信息，这些安全服务包括 IP 层服务（如头认证和负载封装）、传输或应用层服务，以及协商流量的自我保护服务等。 ISAKMP 定义包括交换密钥生成和认证数据的有效载荷。这些格式为传输密钥和认证数据提供了统一框架，而它们与密钥产生技

术，加密算法和认证机制相独立。

ISAKMP 区别于密钥交换协议是为了把安全连接管理的细节从密钥交换的细节中彻底的分离出来。不同的密钥交换协议中的安全属性也是不同的。然而，需要一个通用的框架用于支持 SA 属性格式，谈判，修改与删除 SA ， ISAKMP 即可作为这种框架。

把功能分离为三部分增加了一个完全的 ISAKMP 实施安全分析的复杂性。然而在有不同安全要求且需协同工作的系统之间这种分离是必需的,而且还应该对 ISAKMP 服务器更深层次发展的分析简单化。

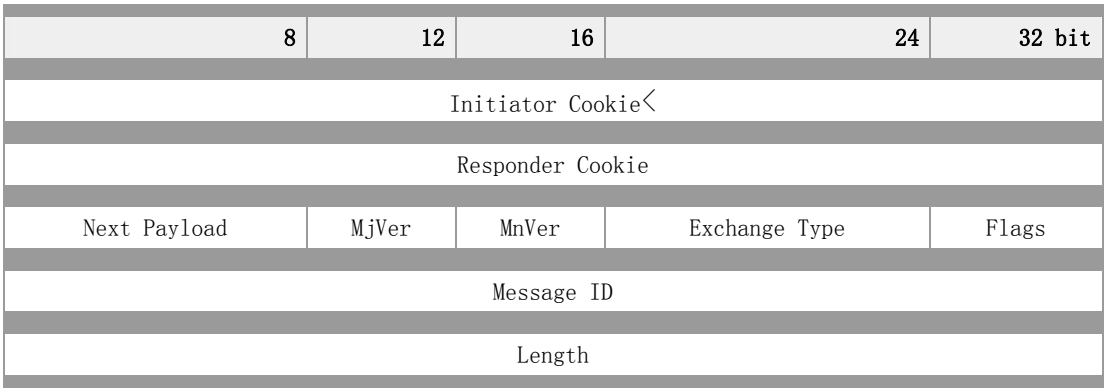
ISAKMP 支持在所有网络层的安全协议 （如： IPSEC、TLS、TLSP、OSPF 等等）的 SA 协商。 ISAKMP 通过集中管理 SA 减少了在每个安全协议中重复功能的数量。 ISAKMP 还能通过一次对整个栈协议的协商来减少建立连接的时间。

ISAKMP 中，解释域（DOI）用来组合相关协议，通过使用 ISAKMP 协商安全连接。共享 DOI 的安全协议从公共的命名空间选择安全协议和加密转换方式，并共享密钥交换协议标识。同时它们还共享一个特定 DOI 的有效载荷数据目录解释，包括安全连接和有效载荷认证。

总之， ISAKMP 关于 DOI 定义如下方面：

- 特定 DOI 协议标识的命名模式；
- 位置字段解释；
- 可应用安全策略集；
- 特定 DOI SA 属性语法；
- 特定 DOI 有效负载目录语法；
- 必要情况下，附加密钥交换类型；
- 必要情况下，附加通知信息类型。

协议结构



- Initiator Cookie — Initiator Cookie：启动 SA 建立、SA 通知或 SA 删除的实体 Cookie。
- Responder Cookie — Responder Cookie：响应 SA 建立、SA 通知或 SA 删除的实体 Cookie。
- Next Payload — 信息中的 Next Payload 字段类型。

- Major Version — 使用的 ISAKMP 协议的主要版本。
- Minor Version — 使用的 ISAKMP 协议的次要版本。
- Exchange Type — 正在使用的交换类型。
- Flags — 为 ISAKMP 交换设置的各种选项。
- Message ID — 唯一的信息标识符，用来识别第 2 阶段的协议状态。
- Length — 全部信息（头＋有效载荷）长（八位）。

TLS：安全传输层协议

安全传输层协议(TLS)用于在两个通信应用程序之间提供保密性和数据完整性。该协议由两层组成：TLS 记录协议（TLS Record）和 TLS 握手协议（TLS Handshake）。较低的层为 TLS 记录协议，位于某个可靠的传输协议（例如 TCP）上面。 TLS 记录协议提供的连接安全性具有两个基本特性：

- 私有——对称加密用以数据加密（DES 、RC4 等）。对称加密所产生的密钥对每个连接都是唯一的，且此密钥基于另一个协议（如握手协议）协商。记录协议也可以不加密使用。
- 可靠——信息传输包括使用密钥的 MAC 进行信息完整性检查。安全哈希功能（ SHA、MD5 等）用于 MAC 计算。记录协议在没有 MAC 的情况下也能操作，但一般只能用于这种模式，即有另一个协议正在使用记录协议传输协商安全参数。

TLS 记录协议用于封装各种高层协议。作为这种封装协议之一的握手协议允许服务器与客户机在应用程序协议传输和接收其第一个数据字节前彼此之间相互认证，协商加密算法和加密密钥。 TLS 握手协议提供的连接安全具有三个基本属性：

- 可以使用非对称的，或公共密钥的密码术来认证对等方的身份。该认证是可选的，但至少需要一个结点方。
- 共享加密密钥的协商是安全的。对偷窃者来说协商加密是难以获得的。此外经过认证过的连接不能获得加密，即使是进入连接中间的攻击者也不能。
- 协商是可靠的。没有经过通信方成员的检测，任何攻击者都不能修改通信协商。

TLS 的最大优势就在于：TLS 是独立于应用协议。高层协议可以透明地分布在 TLS 协议上面。然而，TLS 标准并没有规定应用程序如何在 TLS 上增加安全性；它把如何启动 TLS 握手协议以及如何解释交换的认证证书的决定权留给协议的设计者和实施者来判断。

协议结构

TLS 协议包括两个协议组——TLS 记录协议和 TLS 握手协议——每组具有很多不同格式的信息。

在此文件中我们只列出协议摘要并不作具体解析。具体内容可参照相关文档。

TLS 记录协议是一种分层协议。每一层中的信息可能包含长度、描述和内容等字段。记录协议支持信息传输、将数据分段到可处理块、压缩数据、应用 MAC 、加密以及传输结果等。对接收到的数据进行解密、校验、解压缩、重组等，然后将它们传送到高层客户机。

TLS 连接状态指的是 TLS 记录协议的操作环境。它规定了压缩算法、加密算法和 MAC 算法。

TLS 记录层从高层接收任意大小无空块的连续数据。密钥计算：记录协议通过算法从握手协议提供的安全参数中产生密钥、 IV 和 MAC 密钥。 TLS 握手协议由三个子协议组构成，允许对等双方在记录层的安全参数上达成一致、自我认证、例示协商安全参数、互相报告出错条件。

- 改变密码规格协议
- 警惕协议
- 握手协议

SOCKS：防火墙安全会话转换协议

SOCKS 协议提供一个框架,为在 TCP 和 UDP 域中的客户机/服务器应用程序能更方便安全地使用网络防火墙所提供的服务。这个协议从概念上来讲是介于应用层和传输层之间的 “中介层（shim-layer）”，因而不提供如传递 ICMP 信息之类的网络层网关服务。

利用网络防火墙将组织内部的网络结构与外部网络如 INTERNET 中有效地隔离开来，这种方法正变得逐渐流行起来。这些防火墙系统通常以应用层网关的形式工作在网络之间，提供受控的 TELNET 、FTP 、 SMTP 等的接入。 SOCKS 提供一个通用框架来使这些协议安全透明地穿过防火墙。

SOCKSv5 为这些协议穿越提供了有力的认证方案，而 SOCKSv4 为 TELNET 、FTP 、HTTP 、WAIS 和 GOPHER 等基于 TCP 协议的客户/服务器程序仅仅提供了一个不安全防火墙穿越。新的协议 SOCKS v5 在 SOCKSV4 基础上作了进一步扩展，从而可以支持 UDP ，并对其框架规定作了扩展，以支持安全认证方案。同时它还采用地址解析方案 （addressing scheme） 以支持域名和 IPV6 地址。

为了实现这个 SOCKS 协议，通常需要重新编译或者重新链接基于 TCP 的客户端应用程序以使用 SOCKS 库中相应的封装程序。

协议结构

SOCKS v5 具有以下一些不同格式的信息：

版本标识符/信息选择方法：

1 byte	1 byte	1-255 byte
Version	NMethods	Methods

SOCKS 请求信息：

1 byte	1 byte	Value of 0	1 byte	Variable	2 bytes
Version	CMD	Rsv	ATYP	DST Addr	DST Port

信息选择方法：

1 byte	1 byte
Version	Method<

答复信息：

1 byte	1 byte	Value of 0	1 byte	Variable	2 bytes
Version	REP	RSV	ATYP	BND Addr	BND Port

UDP 请求头：

2byte	1 byte	1 byte	Variable	2	Variable
RSV	FRAG	ATYP	DST Addr	DST Port	Data

PAP：密码认证协议

密码认证协议（PAP），是 PPP 协议集中的一种链路控制协议，主要是通过使用 2 次握手提供一种对等结点的建立认证的简单方法，这是建立在初始链路确定的基础上的。

完成链路建立阶段之后，对等结点持续重复发送 ID/ 密码给验证者，直至认证得到响应或连接终止。

PAP 并不是一种强有效的认证方法，其密码以文本格式在电路上进行发送，对于窃听、重放或重复尝试和错误攻击没有任何保护。对等结点控制尝试的时间和频度。所以即使是更高效的认证方法（如 CHAP），其实现都必须在 PAP 之前提供有效的协商机制。

该认证方法适用于可以使用明文密码模仿登录远程主机的环境。在这种情况下，该方法提供了与常规用户登录远程主机相似的安全性。

协议结构

密码认证协议的配置选项格式：

8	16	32 bit
Type	Length	Authentication-Protocol

- Type — 3
- Length — 4
- Authentication-Protocol — C023（Hex））

PAP 数据包格式：

8	16	32 bit	variable
Code	Identifier	Length	Data

- Code — Code 字段为 8 字节，用于识别 PAP 数据包类型。PAP Code 字段分配如下：1、Authenticate - Request；2、Authenticate - Ack；3、Authenticate-Nak 。
- Identifier — Identifier 字段为 8 字节，用于匹配 Request 和 Reply。
- Length — Length 字段为 16 字节，表示 PAP 数据包的长，包括 Code、Identifier、Length 和 Data 字段。Length 字段外的八位位组用作数据链路层间隙，且在接收方忽略。
- Data — Data 字段为 0 或更多字节。Data 字段格式取决于 Code 字段。

CHAP：PPP 挑战握手认证协议

挑战握手认证协议（CHAP）通过三次握手周期性的校验对端的身份，在初始链路建立时完成，可以在链路建立之后的任何时候重复进行。

1. 链路建立阶段结束之后，认证者向对端点发送“challenge”消息。

2. 对端点用经过单向哈希函数计算出来的值做应答。

3. 认证者根据它自己计算的哈希值来检查应答，如果值匹配，认证得到承认；否则，连接应该终止。

4. 经过一定的随机间隔，认证者发送一个新的 challenge 给端点，重复步骤 1 到 3 。

通过递增改变的标识符和可变的挑战值，CHAP 防止了来自端点的重放攻击，使用重复校验可以限制暴露于单个攻击的时间。认证者控制验证频度和时间。

该认证方法依赖于只有认证者和对端共享的密钥，密钥不是通过该链路发送的。

虽然该认证是单向的，但是在两个方向都进行 CHAP 协商，同一密钥可以很容易的实现相互认证。

由于 CHAP 可以用在许多不同的系统认证中，因此可以用 NAME 字段作为索引，以便在一张大型密钥表中查找正确的密钥，这样也可以在一个系统中支持多个 NAME/ 密钥对，并可以在会话中随时改变密钥。

CHAP 要求密钥以明文形式存在，无法使用通常的不可回复加密口令数据库。

CHAP 在大型网络中不适用，因为每个可能的密钥由链路的两端共同维护。

协议结构

CHAP 的配置选项格式如下：

8	16	32	40 bit
Type	Length	Authentication-Protocol	Algorithm

- Type — 3
- Length — 5
- Authentication-Protocol — 对于 CHAP，为 C223（Hex）。
- Algorithm — Algorithm 字段为八位字节，表示使用的认证方法。

CHAP 数据包结构如下所示：

8	16	32 bit	Variable
Code>	Identifier	Length	Data . . .

- Code — 识别 CHAP 数据包类型。CHAP 代码具有以下几种：1、Challenge；2、Response；3、Success；4、Failure。
- Identifier — 用于匹配 Challenges、Responses 和 Replies 信息。
- Length — CHAP 数据包的长度，包括 Code、Identifier、Length 和 Data 字段。
- Data — 0 或更多八位字节。该字段格式取决于 Code 字段。对于 Success 和 Failure，Data 字段包括一个独立执行的可变信息字段。

EAP：PPP 的扩展认证协议

PPP 扩展认证协议（EAP）用于 PPP 认证，可以支持多种认证机制。EAP 并不在链路控制阶段指定认证方法，而是把这个过程推迟到认证阶段。这样认证方就可以在要求更多的信息以后再决定使用什么认证方法。这种机制就允许使用一台“后端“服务器来真正执行认证机制，而 PPP 认证方只是传递认证交换信息。

1. 在链路建立阶段完成以后，认证方向对端发送一个或多个请求报文去认证结点。在请求报文中有一个类型字段用来指明认证方所请求的信息，例如是 ID、MD5 的挑战字、一次密码（OTP）以及通用令牌卡等。MD5 的挑战字对应于 CHAP 认证协议的挑战字。通常认证方首先发送一个初始的 ID 请求随后再发送其他的请求信息。当然，这个 ID 请求报文并不是必须的，在对端身份是已知的情况下（如租用线、拨号专线等）可以跳过这个步骤。

2. 端点对每一个请求报文回应一个应答包。和请求报文一样，应答报文中也包含一个类型字段，对应于所回应的请求报文中的类型字段。

3. 认证方通过发送一个成功或者失败的报文来结束认证过程。

EAP 可以支持多种认证机制，而无需在 LCP 阶段预协商过程中指定一种认证机制。某些设备（例如：网络接入服务器）不需要了解每一个请求报文的类型，而是作为一个代理把认证报文直接传递给后端的认证服务器。设备只需关心认证结果是成功还是失败，然后结束认证阶段。

EAP 需要在 LCP 中增加一个新的认证类型，这样现有的 PPP 要想使用 EAP 就必须进行修改。同时，使用 EAP 也和先前的在 LCP 协商阶段指定认证方法的 PPP 认证模型不一致。

协议结构

提供 EAP 认证协议的配置选项格式如下所示：

8	16	32 bit	Variable
Type	Length	Authentication-Protocol	Data

- Type — 3
- Length — 4
- Authentication-Protocol — 对于 PPP 中的 EAP，该字段为 C227（十六进制）。

一个 PPP EAP 数据包封装在 PPP 数据链路层帧的 Information 字段，其中的 Protocol 字段表示类型为十六进制 C227（PPP EAP）。EAP 数据包格式如下所示：

8	16	32 bit	Variable
Code	Identifier	Length	Data

- Code — Code 字段识别 EAP 数据包类型。
- EAP 代码分配如下：1、请求（Request）；2、响应（Response）；3、成功（Success）；4、失败（Failure）
- Identifier — Identifier 字段用于匹配响应和请求信息。
- Length — Length 字段表示 EAP 数据包的长度，包括 Code、Identifier、Length 和 Data 字段
- Data — Data 字段的格式取决于 Code 字段。

Kerberos：网络认证协议

Kerberos 是一种网络认证协议，其设计目标是通过密钥系统为客户机 / 服务器应用程序提供强大的认证服务。该认证过程的实现不依赖于主机操作系统的认证，无需基于主机地址的信任，不要求网络上所有主机的物理安全，并假定网络上传送的数据包可以被任意地读取、修改和插入数据。在以上情况下， Kerberos 作为一种可信任的第三方认证服务，是通过传统的密码技术（如：共享密钥）执行认证服务的。

认证过程具体如下：客户机向认证服务器（AS）发送请求，要求得到某服务器的证书，然后 AS 的响应包含这些用客户端密钥加密的证书。证书的构成为： 1) 服务器 “ticket” ； 2) 一个临时加密密钥 （又称为会话密钥 “session key”） 。客户机将 ticket （包括用服务器密钥加密的客户机身

份和一份会话密钥的拷贝）传送到服务器上。会话密钥可以（现已经由客户机和服务器共享）用来认证客户机或认证服务器，也可用来为通信双方以后的通讯提供加密服务，或通过交换独立子会话密钥为通信双方提供进一步的通信加密服务。

上述认证交换过程需要只读方式访问 Kerberos 数据库。但有时，数据库中的记录必须进行修改，如添加新的规则或改变规则密钥时。修改过程通过客户机和第三方 Kerberos 服务器（Kerberos 管理器 KADM）间的协议完成。有关管理协议在此不作介绍。另外也有一种协议用于维护多份 Kerberos 数据库的拷贝，这可以认为是执行过程中的细节问题，并且会不断改变以适应各种不同数据库技术。

协议结构

Kerberos 信息：

- 客户机/服务器认证交换

信息方向	信息类型
客户机向 Kerberos	KRB_AS_REQ
Kerberos 向客户机	KRB_AS_REP 或 KRB_ERROR<

- 客户机/服务器认证交换

信息方向	信息类型
客户机向应用服务器	KRB_AP_REQ
[可选项] 应用服务器向客户机	KRB_AP_REP 或 KRB_ERRORR

- 票证授予服务（TGS）交换

信息方向	信息类型
客户机向 Kerberos	KRB_TGS_REQ
Kerberos 向客户机	KRB_TGS_REP 或 KRB_ERROR

- KRB_SAFE 交换
- KRB_PRIV 交换

- KRB_CRED 交换

TACACS＆TACACS+：终端访问控制器访问控制系统

终端访问控制器访问控制系统（TACACS）通过一个或多个中心服务器为路由器、网络访问服务器以及其它网络处理设备提供了访问控制服务。 TACACS 支持独立的认证（authentication）、 授权（authorization）和计费（accounting）功能。

TACACS 允许客户机接受用户名和口令，并发送查询指令到 TACACS 认证服务器（又称之为 TACACS daemon 或 TACACSD）。在通常情况下，该服务器运行在主机上的一个程序。该主机决定是否接受或拒绝，然后返回一个响应。 TIP 根据响应类型，判断是否允许访问。在上述过程中，判断处理过程是“公开（opened up）”的，并且对应的算法和数据在运行 TACACS daemon 的主机的完全控制之下。此外 TACACS 扩展协议支持更多类型的认证请求和响应代码。

当前 TACACS 具有三种版本，其中第三版 TACACS+ 与前两版不兼容。

协议结构

4	8	16	24	32 bit
Major	Minor	Packet type	Sequence no.	Flags
Session ID				
Length				

- Major Version — 主要 TACACS+ 版本号。
- Minor Version — 次要 TACACS+ 版本号。当需要维持后向兼容性时，允许修订 TACACS+ 协议。
- Packet Type — 可能值包括：
  TAC_PLUS_AUTHEN: = 0x01 （认证）；
  TAC_PLUS_AUTHOR: = 0x02 （授权）；
  TAC_PLUS_ACCT: = 0x03 （计费）。
- Sequence Number — 当前会话中的数据包序列号。会话中的第一个 TACACS+ 数据包序列号必须为 1，其后的每个数据包序列号逐次加 1。因此客户机只发送奇序列号数据包，而 TACACS+ Daemon 只发送偶序列号数据包。
- Flags — 该字段包括各种位图格式的标志（flag）。Flag 值表明数据包是否进行加密。
- Session ID — 该 TACACS+ 会话的 ID。

- Length — TACACS+ 数据包主体总长（不包括头部）。