各类数据帧格式及协议内容的总结

1．HDLC 协议

HDLC 协议的全称是高级链路控制协议（High Level Data Link Control），是一种在网上同步传输数据，面向比特的数据链路层协议，广泛用于公用数据网，支持全双工或半双工传输，使用后退 N 帧 ARQ 流控方案。HDLC 定义了 3 种类型的站（主站、从站、复合站），两种链路配置（不平衡配置、平衡配置），3 种数据传输方式（NRM、ABM、ARM）。

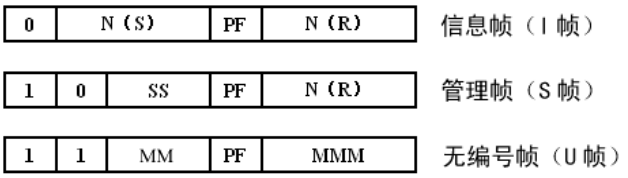
HDLC 帧格式

帧标志 F：HDLC 用一种特殊的位模式 01111110 作为标志以确定帧的边界，采用位填充技术来区分是标志字段还是数据字段，发送站的数据比特序列一旦发现 0 后有 5 个 1，则在第 7 位插入 0。

地址字段 A:地址字段用于标识从站的地址，用在点对多点的链路中，地址通常是 8 位长。

控制字段 C：帧编号 N(S),捎带的肯定应答序号 N(R),PF 位，P 询问、F 终止

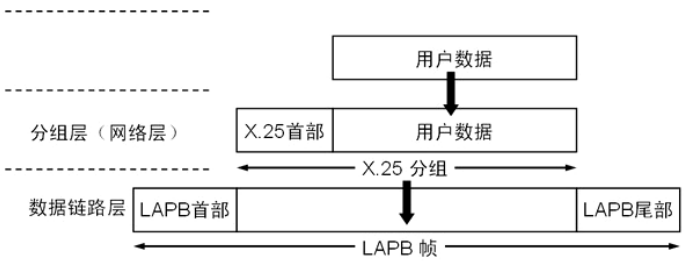
信息帧：主要用于传输数据
监控帧：流量控制和差错控制
无编号帧：链路控制，UDP则承载数据



只有信息帧和无编号帧才有信息字段用于承载用户数据无编号帧仅针对UDP才承载数据

帧校验序列 FCS：含有除标志字段之外的所有其他字段的校验和。通常使用 16 比特的 CRC-CCITT（$G(x)=X^{16}+X^{12}+X^5+1$）标准产生校验序列，有时也采用 CRC-32 产生 32 位的校验序列。

2．X．25 的帧格式及协议



（1）协议概述

X.25 是 CCITT 公布的用于连接数据终端至分组交换数据网络的推荐标准，X.25 是一个面向连接的接口，采用虚电路传递数据分组至网络上的适当终点处。在 X.25 的网络中，用户的计算机终端设备将与分组/拆装设备（PAD）连接，负责完成分割分组、寻址、重组装分组的工作，而不同的 X.25 网络之间则要

使用 X.75 协议互联。X.25 是一个基于分组交换技术构建的网络，分组交换本身是适于无连接业务的，要为用户提供面向连接的接口服务，则必须借助虚拟电路技术（VC），虚电路服务具有两种形式，一种是交换虚电路 SVC，一种是永久虚电路 PVC。最常见的 X.25 协议支持的最大传输速率为 64Kb/s。

（2）X.25 的三层结构

X.25 层次结构	对应 OSI 层	相应标准
分组层	网络层	X.25 PLP 通过建立虚拟连接，提供点对点、面向连接服务。X.25 PLP 层采用后退 N 帧 ARQ 流控协议。PLP 协议把用户数据分成一定大小的块 ，一般为 128 字节，再加上 24 位或 32 的分组头组成数据分组
链路访问层	数据链路层	使用平衡式链路访问规程 LAPB，LAPB 是源于 HDLC 的一种面向位的协议，实际上是平衡的异步方式类别下的 HDLC。LAPB 是 HDLC 的一个子集
物理层	物理层	X.21，但可以使用 RS-232C 和 V.35 代替

***相关知识点**

选择重发 ARQ 协议（有噪声环境双工）：滑动窗口协议与自动请求重发技术的结合，当收到否定应答（NAK）时，只重发出错的帧。$W_{发}=W_{收}\leqslant 2^{K-1}$。

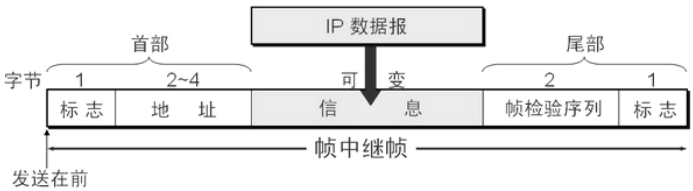
后退 N 帧 ARQ 协议（有噪声环境双工）：滑动窗口协议与自动请求重发技术的结合，当收到否定应答（NAK）时，将从出错处重发已发出过的 N 个帧。接收窗口 $W_{收}=1$，同时 $W_{收}\leqslant 2^{K}-1$。（K 为帧编号的位数）

3. 帧中继的帧格式

（1）协议概述

帧中继是综合业务数字网络（ISDN）的一个产物，没有专门定义物理层接口（可以使用 X.21，V.35 等接口协议），帧中继在第二层建立虚电路，因而第三层被简化掉了，FR 的帧层也比 HDLC 操作简单，只做检错，不再重传，没有滑动窗口式的流控，只有拥塞控制，把复杂的检错丢给高层去处理。帧中继使用的核心协议是 LAPD，它比 LAPB 简单，省去了控制字段。帧中继是基于分组（帧）交换的透明传输，可以承载 IP 数据报；可提供面向连接的服务，支持交换虚电路（SVC）和永久虚电路（PVC）；帧长可变，长度可达 1600～4096 字节，可以承载各种局域网的数据帧；可以应付突发的数据传输，可以提供 2～45Mb/s 的数据率；帧中继不适于延迟较敏感的应用（音频和视频），无法保证可靠提交。

（2）Frame Relay 的帧格式



DLCI（高位）			C/R	EA=0
DLCI（低位）	FECN	BECN	DE	EA=1

标志字段：LAPD 的帧头和帧尾都是一个字节的帧标志字段，编码为 01111110，与 HDLC 一样。

地址字段：

• **EA：**地址扩展比特。该比特为 0 时表示地址向后扩展一个字节，为 1 时表示最后一个字节。

• **C/R：**命令/响应比特。协议本身不使用这个比特，用户可以用这个比特区分不同的帧。

• **FECN：**向前拥塞比特。若网络置该位为 1，则表示在帧的传送方向上出现了拥塞，该帧到达接收端后，接收方可根据此调整发送方的数据率。

• **BECN：**向后拥塞比特。若网络置该位为 1，则表示在帧传送相反的方向上出现了拥塞，该帧到达发送端后，发送方可据此调整发送数据速率。

• **DE：**优先丢弃比特。当网络发生拥塞时，DE 位置 1 的帧会优先丢弃。

• **DLCI：**数据链路连接标识符。帧中继使用虚拟电路的方式提供面向连接的服务，在帧头中包括 DLCI 字段，每个 DLCI 都标识一个虚电路，其中 DLCI0 用于信令传输。

信息字段：信息字段长度可变，1600 是默认最大长度。

帧校验序列：与 HDLC 相同。

（3）帧中继的拥塞控制

在帧中继承载业务中，使用显式信令和隐式信令来避免拥塞的发生。显示信令利用 FECN 和 BECN 比特位置 1 来向端用户发出拥塞警告，以避免拥塞的发生。隐式信令是指上层协议对网络拥塞的监控，当网络开始丢帧时，上层协议就自动降低发送速率，以便网络从拥塞中恢复正常运行。帧中继还可以利用 CLLM（强化链路层管理）的方法，缓解拥塞。

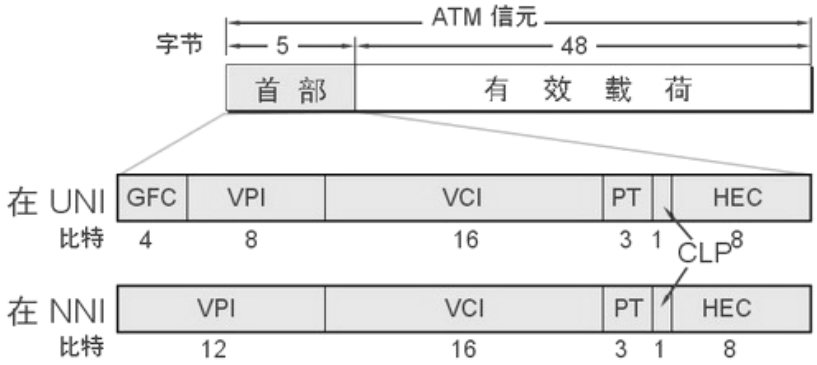
4. ATM 问题

ATM 是一种可以将局域网功能、广域网功能、语音、视频和数据，集成进一个统一的协议设计。ATM 标准最早是作为 B-ISDN 标准的一部分而出现的，它在 QoS 方面有突出表现。在 ATM 传输中，ATM 把用户数据组成 53B 的信元作为分组交换的信息单位，采用统计时分复用模式，提供面向连接的虚电路服务。ATM 连接可以是点到点的连接，也可以是点到多点的连接，分为 PVC 和 SVC 两种虚电路。ATM 通常是在光纤的基础上建立的，典型的数据速率为 155.5Mb/s，因此它是不提供应答的，将少量的错误交给高层处理。ATM 的目的是实现实时通信，对于偶然的信元错误是不重传的，对于要重传的信息由高层处理。

（1）ATM 的分层体系结构

层 次	子 层	功 能	与 OSI 对应
高　层		对用户数据的控制	高　层
ATM 适配层（AAL）	汇聚子层（CS）	为高层数据提供统一接口	第四层
	拆装子层（SAR）	分割和合并用户数据	
ATM 层		VPI 和 VCI 的管理；信元头的组装和拆分； 信元的多路复用；流量控制	第三层
物理层	传输汇聚子层（TC）	信元校验和速率控制；数据帧的组装和拆分	第二层
	物理介质子层（PMD）	比特定时；物理网络接入	第一层

（2）ATM 信元头结构



• **流控标志（GFC）**：用于主机和网络之间的流控或优先级控制。

• **虚通路标识符（VPI）**：8 位（UNI）或 12 位（NNI），常用是 8 位，因此一个主机上的虚通路数 256 个。

• **虚信道标识符（VCI）**：16 位，理论上每个主机上的虚通路可以包含 65536 个虚信道，不过部分信道是用于控制的，并不传送用户数据。

*在 ATM 中，虚电路有两级：虚通路（VP）和虚信道（VC），虚通路是由多条虚信道捆绑在一起形成的。在 ATM 逻辑通道中，是使用 VPI+VCI 的组合来标识连接的，在做 VP 交换或交叉连接时，只需交换 VP，无需改变 VCI 的值。

• **负载类型（PTI）**：区分不同的拥塞信息。

• **信元丢失优先级（CLP）**：这一位用于区分信息的优先级，如果出现拥塞，交换机优先丢弃 CLP 被置 1 的信元。

• **头校验和（HEC）**：它支队信元头进行校验，采用的是 X^8+X^2+X+1 的 8 位 CRC 校验。

（3）ATM 适配层协议及服务

• **AAL1**：恒定比特率，面向连接业务，端到端定时，检错（A 类业务）

• **AAL2**：面向连接的，可变比特率的实时数据流业务，端到端定时，不检错（B 类业务）

• **AAL3/4**：面向连接或无连接，可变比特率，对信元错误和丢失敏感（C 类、D 类业务）

• **AAL5**：面向连接或无连接，可变比特率，在 ATM LANE 中有重要应用（C 类、D 类业务）

－－－

CBR（固定比特率业务）：交互式语音和视频流

RT-VBR（实时性变化比特率业务）：交互式压缩视频信号

NRT-VBR（非实时性变化比特率业务）：多媒体电子邮件

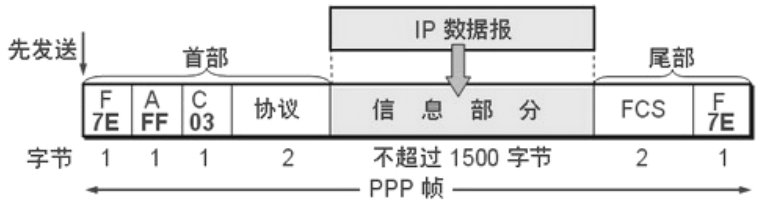
ABR（有效比特率业务）：突发式业务

UBR（不定比特率业务）：IP 分组传送

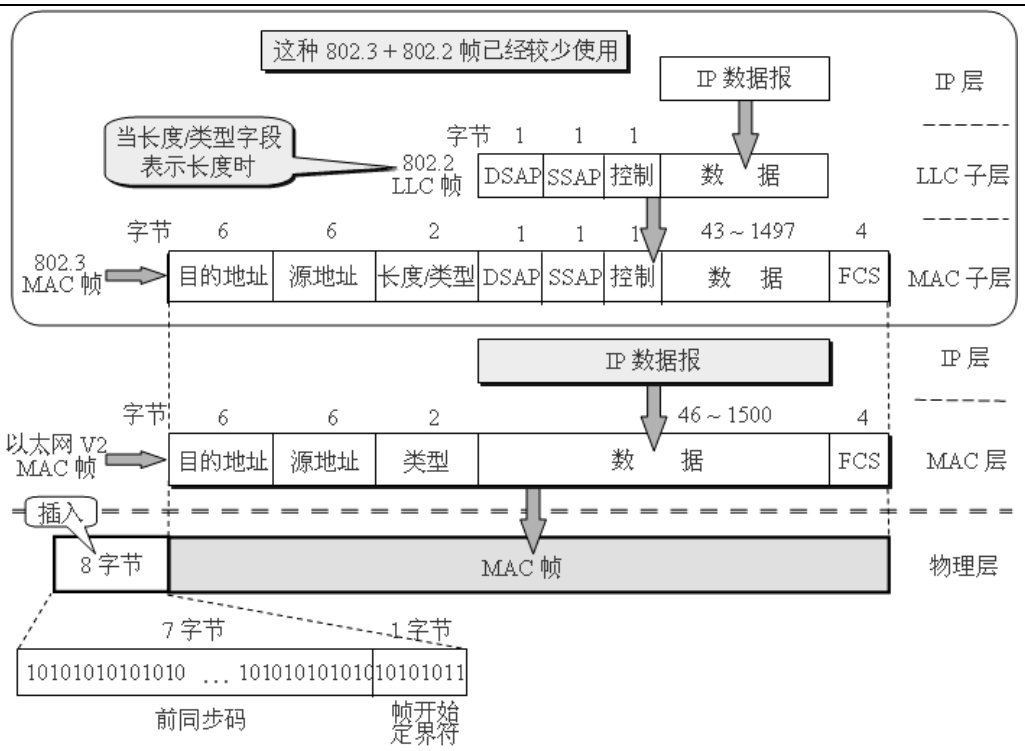
－－－

HDLC、X.25、FR、ISDN、ATM 构成了 5 种常见的广域网通信技术

5. PPP 的帧格式

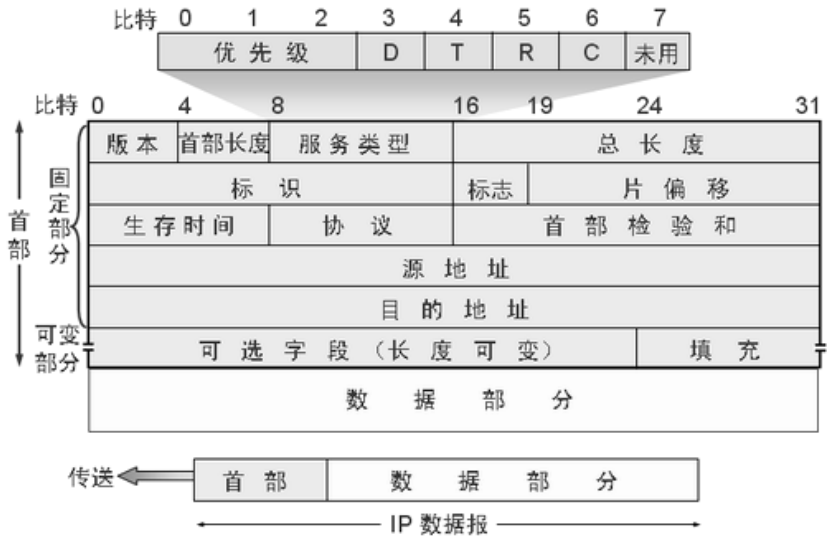


6. 局域网的帧格式



7. IPv4 协议

（1）IPv4 数据报的格式



• **版本号：**占 4 比特，指 IP 协议的版本，目前广泛使用 IPv4。

• **首部长度（IHL）：**IP 头长度，占 4 比特，最大值 15 个单位（1 个单位 4 字节）。

• **服务类型：**该字段包括一个 3 比特的优先级子字段（现在已废弃不用），还包括一个 4 比特的 ToS 子字段，最后 1 比特必须置 0。ToS 中的 4 比特分别代表：最小时延（D）、最大吞吐率（T）、最高可靠性（R）和最小费用（C）,只能有 1 比特置 1。如果所有 4 比特均为 0，那就是一般服务。

• **标识符：**由主机指定同样的标识符。当原主机对数据分段时，对同一上层协议数据单元划分出的各个数据报指定同样的标识符，目标主机上层协议用这个字段进行重装配。

• **标志：**包括三个标志位。一个标志位没有使用；M 标志用于分段和重装配；D 标志为禁止分段标志。

• **段偏置值：**指明该段处于原来数据报中的位置，已 8 字节为单位。

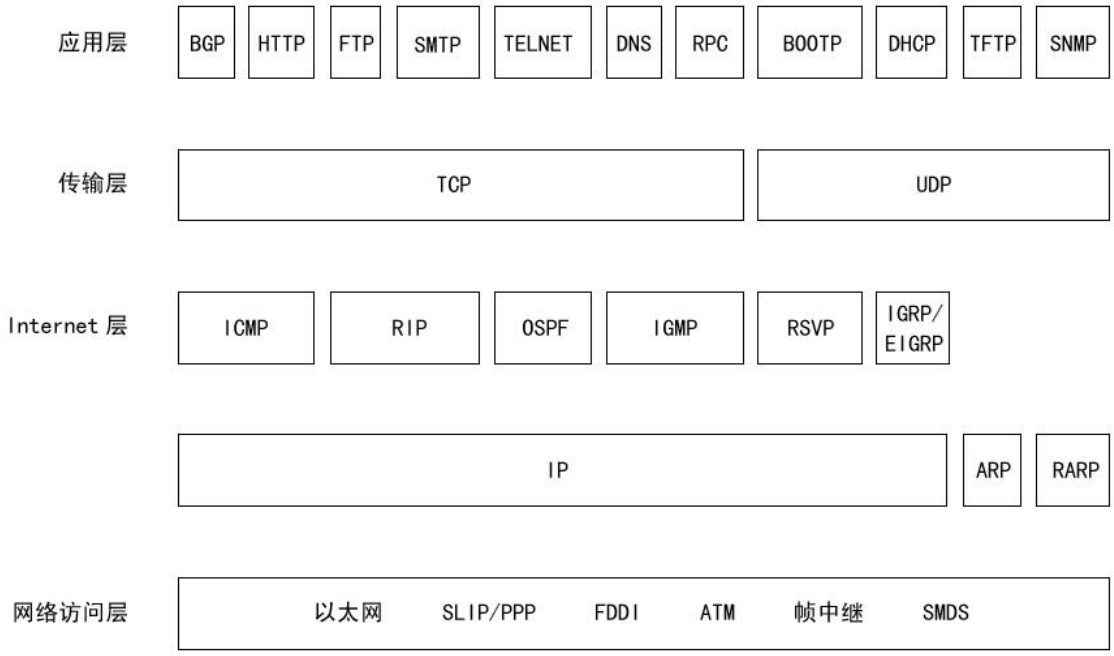
• **生存期（TTL）：**用经过的路由器个数表示，源站设置一个数（32 或 64），每经过一个路由器减 1。如果某个路由器发现 TTL 字段为 0，则丢弃该数据报，不再转发。

• **协议：**上层协议（TCP 或 UDP）。

• **头检验和：**对 IP 头的检验序列。

• **任选数据：**可变长，包含发送者想要发送的控制数据。

（2）IP 协议簇

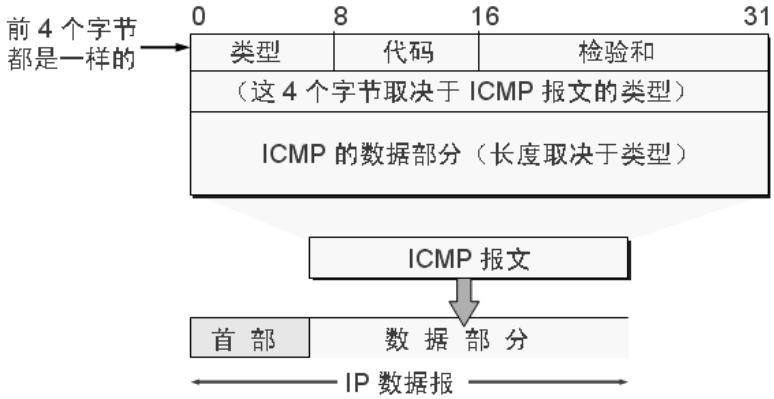


8. ICMP 协议

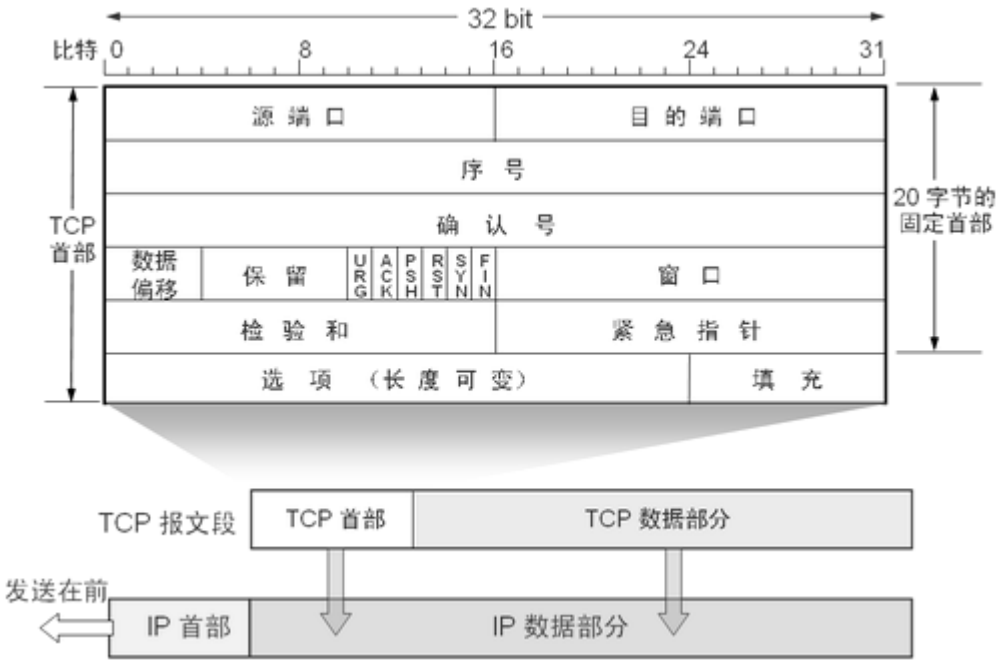
（1）ICMP 协议概述

ICMP（Internet Control Message Protocol）与 IP 协议同属于网络层，封装在 IP 数据报中传输，传送有关网络层通信问题的信息。ICMP 常见应用有：报告访问失效（报告源主机网络不可达）；报告网络拥塞（发送源抑制报文给源主机，降低发送速率）；帮助排错（利用 ICMP 回声功能，ping 工具）；声明报文超时（TraceRoute 工具，利用较小的 TTL 值发现中间设备）。

（2）ICMP 报文格式



9. TCP 格式



（1）TCP 报文格式

源端口和目的端口：都是 16 个比特，分别表示发送方和接收方的端口号。端口号和 IP 地址构成套接字(socket)地址的主要内容。源端和目的端的套接字合起来唯一地表示一条连接。网络应用程序在通信时直接向套接字发送和接收数据。

序列号和确认号：都是 32 位的无符号整数，可以表示 0-4G(232)字节的范围。其中，序列号表示数据部分第一个字节的序列号，而确认号表示该数据报的接收者希望对方发送的下一个字节的序号(即序号小于确认号的数据都已正确地被接收)。

头长度(HLEN)：表示 TCP 报文头的长度。长度以 32-bit 为单位来计算。所以如果选项部分的长度不是 4 个字节的整数倍，则要加上填充(padding)。

保留域：紧接在头长度字段后有 6 个比特，应该把它设置为 0。

再后则是 6 个标志位。标志位特定的含义：

URG(urgent)为紧急数据标志。如果它为 1，则表示本数据报中包含紧急数据。此时紧急数据指针表示的值有效。它表示在紧急数据之后的第一个字节的偏侈值(即紧急数据的总长度)。

ACK(acknowledge)为确认标志位。如果 ACK 为 1，则表示报文中的确认号是有效的。否则，报文中的确认号无效，接收端可以忽略它。

PSH(push)标志位。被置位后，要求发送方的 TCP 协议软件马上发送该数据报，接收方在收到数据后

也应该立即上交给应用程序，即使其接收缓冲区尚未填满。

RST(reset)标志位。用来复位一条连接。RST 标志置位的报文称为复位报文。一般情况下，如果 TCP 收到的一个报文明显不是属于该主机上的任何个连接，则向远端发送一个复位报文。

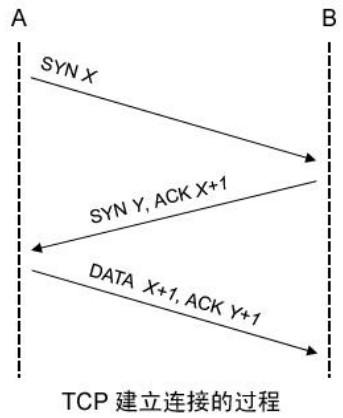
SYN(synchronous)标志位。用来建立连接，让连接双方同步序列号。如果 SYN=1 而 ACK=0，则表示该数据报为连接请求，如 SYN=1 而 ACK=1 则表示是接受连接。

FIN(finish)标志位。表示发送方已经没有数据要传输了，希望释放连接。

窗口(window)字段。窗口表示的是从被确认的字节开始，发送方最多可以连续发送的字节的个数。接收方通过设置该窗口值的大小，可以调节源端发送数据的速度，从而实现流控。

校验和(checksum)域。是 TCP 协议提供的一种检错机制。与我们在前面的章节中学过的 UDP 协议类似，在计算校验和时不仅要计算 TCP 报文自身(报文头和数据)，还要增加一些额外的信息内容 - 12 个字节的“伪包头”。

（2）TCP 三次握手过程

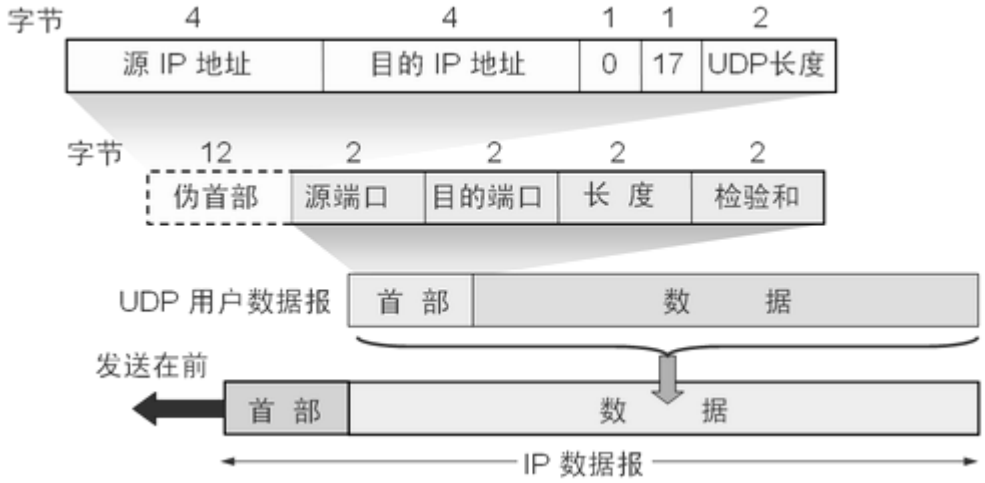


TCP 建立连接的过程

TCP 采用三次握手过程建立连接，首先是发起方发送一个 SYN 标志置位的段，其中的发送顺序号为某个值 X，称为初始顺序号 ISN（Initial Sequence Number），接收方以 SYN 和 ACK 标志置位的段响应，其中的应答顺序号应为 X+1（表示期望从第 X+1 个字节处开始接收数据），发送顺序号为某个值 Y（接收端指定的 ISN）。这个段到达发起端后，发起端以 ACK 标志置位，应答顺序号为 Y+1 的段回答，连接就正式建立了，连接建立的同时发起方还可以发送数据。

TCP 采用的流控方式与数据链路层的流控方式不同，属于可变大小的滑动窗口协议，也叫信贷（Credit）滑窗协议，它更适合于两个相距遥远的主机在无连接的网络上实现流量控制。

10. UDP 格式



（1）UDP 报文格式

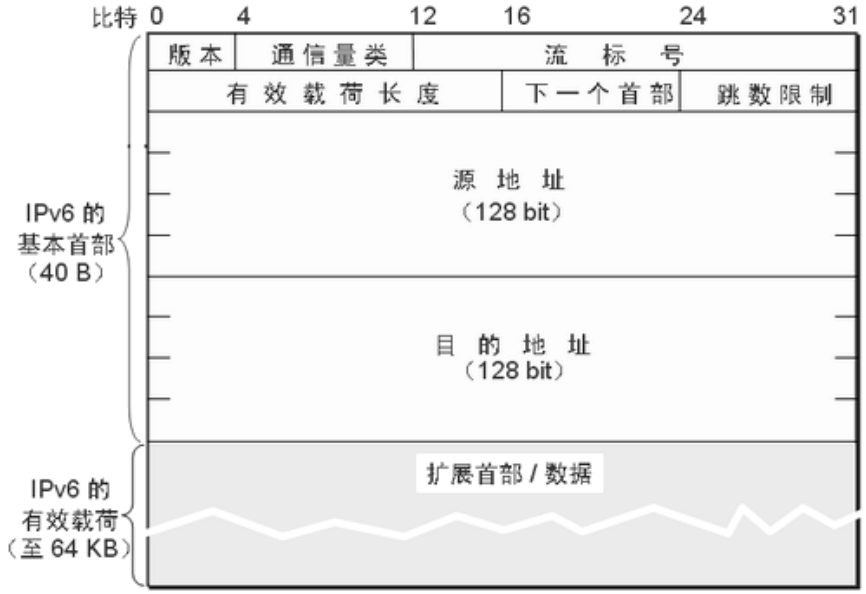
源端口(Source Port)和目的端口(Destination Port)字段包含了 16 比特的 UDP 协议端口号，它使得多个应用程序可以多路复用同一个传输层协议 - UDP 协议，仅通过不同的端口号来区分不同的应用程序。

长度(Length)字段记录了该 UDP 数据包的总长度(以字节为单位)，包括 8 字节的 UDP 头和其后的数据部分。最小值是 8(即报文头的长度)，最大值为 65,535 字节。

UDP 检验和(Checksum)的内容超出了 UDP 数据报文本身的范围，实际上，它的值是通过计算 UDP 数据报及一个伪包头而得到的。但校验和的计算方法与通用的一样，都是累加求和。

所谓“伪首部”是因为这种伪首部并不是 UDP 用户数据报的真正首部。只是在计算检验和时，临时和 UDP 用户数据报连接在一起，得到一个过渡的 UDP 用户数据报。检验和就是按照这个过渡的 UDP 用户数据报来计算的。伪首部既不向下传送也不向上递交，而仅仅是为了计算检验和。

11. IPv6 报文格式



（1）IPv6 数据报的格式

IPv6 包头长度固定为 40 字节，去掉了 IPv4 中一切可选项，只包括 8 个必要的字段，因此尽管 IPv6 地址长度为 IPv4 的四倍，IPv6 包头长度仅为 IPv4 包头长度的两倍。

其中的各个字段分别为：

• **Version（版本号）**：4 位，IP 协议版本号，值= 6。

• **Traffice Class（通信类别）**：8 位，指示 IPv6 数据流通信类别或优先级。功能类似于 IPv4 的服务类型（TOS）字段。

• **Flow Label（流标记）**：20 位，IPv6 新增字段，标记需要 IPv6 路由器特殊处理的数据流。该字段用于某些对连接的服务质量有特殊要求的通信，诸如音频或视频等实时数据传输。在 IPv6 中，同一信源和信宿之间可以有多种不同的数据流，彼此之间以非“0”流标记区分。如果不要求路由器做特殊处理，则该字段值置为“0”。

• **Payload Length（负载长度）**：16 位负载长度。负载长度包括扩展头和上层 PDU，16 位最多可表示 65，535 字节负载长度。超过这一字节数的负载，该字段值置为“0”，使用扩展头逐个跳段（Hop-by-Hop）选项中的巨量负载（Jumbo Payload）选项。

• **Next Header（下一包头）**：8 位，识别紧跟 IPv6 头后的包头类型，如扩展头（有的话）或某个传输层协议头（诸如 TCP，UDP 或着 ICMPv6）。

• **Hop Limit（跳段数限制）**：8 位，类似于 IPv4 的 TTL（生命期）字段。与 IPv4 用时间来限定包的生命期不同，IPv6 用包在路由器之间的转发次数来限定包的生命期。包每经过一次转发，该字段减 1，减到 0 时就把这个包丢弃。

• **Source Address（源地址）**：128 位，发送方主机地址。

• **Destination Address（目的地址）**：128 位，在大多数情况下，目的地址即信宿地址。但如果存在路由扩展头的话，目的地址可能是发送方路由表中下一个路由器接口。

• **扩展首部**：IPv6 包头设计中对原 IPv4 包头所做的一项重要改进就是将所有可选字段移出 IPv6 包头，置于扩展头中。由于除 Hop-by-Hop 选项扩展头外，其他扩展头不受中转路由器检查或处理，这样就能提高路由器处理包含选项的 IPv6 分组的性能。

通常，一个典型的 IPv6 包，没有扩展头。仅当需要路由器或目的节点做某些特殊处理时，才由发送方添加一个或多个扩展头。与 IPv4 不同，IPv6 扩展头长度任意，不受 40 字节限制，以便于日后扩充新增选项，这一特征加上选项的处理方式使得 IPv6 选项能得以真正的利用。 但是为了提高处理选项头和传输层协议的性能，扩展头总是 8 字节长度的整数倍。

目前，RFC 2460 中定义了以下 6 个 IPv6 扩展头：Hop-by-Hop（逐个跳段）选项包头、目的地选项包

头、路由包头、分段包头、认证包头和 ESP 协议包头。

（2）从 IPv4 向 IPv6 过渡的策略

两种向 IPv6 过渡的策略，即使用双协议栈和隧道技术。**双协议栈（dual stack）**是指在完全过渡到 IPv6 之前，使一部分主机（或路由器）装有两个协议栈，一个 IPv4 和一个 IPv6。因此双协议栈主机（或路由器）既能够和 IPv6 的系统通信，又能够和 IPv4 的系统进行通信。双协议栈的主机（或路由器）记为 IPv6/ IPv4，表明它具有两种 IP 地址：一个 IPv6 地址和一个 IPv4 地址。向 IPv6 过渡的另一种方法是**隧道技术（tunneling）**。这种方法的要点就是在 IPv6 数据报要进入 IPv4 网络时，将 IPv6 数据报封装成为 IPv4 数据报（整个的 IPv6 数据变成了 IPv4 数据报的数据部分）。然后 IPv6 数据报就在 IPv4 网络的隧道中传输。当 IPv4 的数据报离开 IPv4 网络中的隧道时再将其数据部分（即原来的 IPv6 数据报）交给主机的 IPv6 协议栈。

网络新技术的总结

1．IP 交换技术

所谓 IP 交换技术是指利用第二层交换技术传送 IP 分组的一组协议和机制，它利用交换机的高带宽和低延迟优势尽可能快地传送分组通过网络。由于 IP 是无连接的协议，对每个分组都必须单独选择路由，因此路由器的转发速度是比较慢的。IP 交换的目的是在快速交换硬件上获得最有效的 IP 实现，并非连接的 IP 和面向连接的 ATM 的优点互补。Ipsilon 公司开发的 IP 交换机提供了快捷通道（Cut Through），使得 IP 路由器的转发能力提高了 5 倍。IP 交换机之间的信令使用流管协议 IFMP（Ipsilon Flow Management Protocol）和通用交换机管理协议 GSMP（General Switch Management Protocol）。

IFMP（RFC1953）的功能是建立结点间之间的邻接关系，并把一个第二层标记绑定到一个特殊的 IP 数据流上。所谓“流”，是指具有相同源地址和目标地址、共同的上层协议（UDP、TCP）和服务类型的一个分组序列。利用标记可以实现对 IP 流进行分类，并更有效地访问有关数据流的路由信息，带有标记的分组无需经过后继结点的第三层转发，而是通过第二层交换快速传输。IFMP 报文包装在 IPv4 分组中广播出去。

GSMP（RFC1987）是一个通用的 ATM 交换机控制协议。GSMP 的功能是建立和释放连接，在组播通信中增加和删除叶子节点，管理交换机的端口，获取配置信息和统计数据等。可变长度的 GSMP 报文封装在 AAL5 协议数据单元中。

IP 交换的转发过程：一个流一旦被识别出来，IP 交换机就通知上游的结点使用新的虚电路（VC）传送这个流，同样的信令也会从下游结点传送过来。当 IP 流通过指定的 VC 传送时，就不再通过路由表转发，而是直接使用 ATM 交换硬件进行处理，同时把第二层标记附加在每个分组的头部，以便加快路由缓冲区的查找。

一台 IP 交换机主要由三个模块组成：ATM 交换模块、IP 交换控制器和专用的管理协议组成。

2．MPLS：Multi-Protocol Label Switching

MPLS 是一种可以在多种第二层媒质上进行标记交换的网络技术。这一技术结合了第二层的交换和第三层路由的特点，将第二层的基础设施和第三层的路由有机地结合起来。第三层的路由在网络的边缘实施，而在 MPLS 的网络核心采用第二层交换，可见 MPLS 相当于 2.5 层协议。

MPLS 通过在每一个节点的标签交换来实现包的转发。它不改变现有的路由协议，并可以在多种第二层的物理媒质上实施，目前有 ATM、FR(帧中继)、Ethernet 以及 PPP 等媒质。

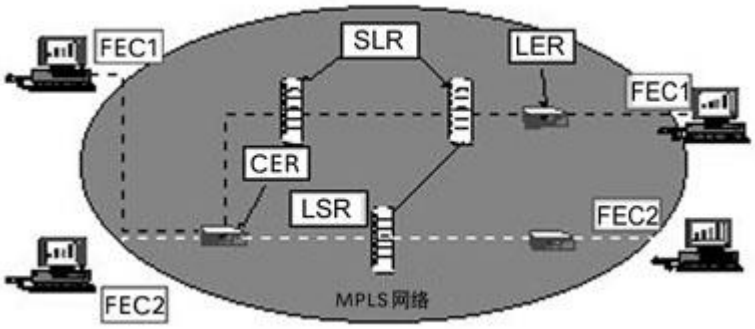
Ethernet/PPP header	Label	Layer 3 data
ATM header	Label	Layer 3 data
VPI/VCI		Layer 3 data

通过 MPLS，第三层的路由可以得到第二层技术的很好补充，充分发挥第二层良好的流量设计管理以及第三层 “Hop-By-Hop（逐跳寻径）”路由的灵活性，以实现端到端的 QoS 保证。

MPLS 作为一种分类转发技术，将具有相同转发处理方式的分组归为一类，称为**转发等价类 FEC**（Forwarding Equivalence Class）。相同转发等价类的分组在 MPLS 网络中将获得完全相同的处理。转发等价类的划分方式非常灵活，可以是源地址、目的地址、源端口、目的端口、协议类型、VPN 等的任意组合。例如，在传统的采用最长匹配算法的 IP 转发中，到同一个目的地址的所有报文就是一个转发等价类。

标签是一个长度固定、只具有本地意义的短标识符，用于唯一标识一个分组所属的转发等价类 FEC。在某些情况下，例如要进行负载分担，对应一个 FEC 可能会有多个标签，但是一个标签只能代表一个 FEC。标签由报文的头部所携带，不包含拓扑信息，只具有局部意义。标签的长度为 4 个字节。

MPLS 是一种特殊的转发机制，它为进入网络中的 IP 数据包分配标记，并通过对标记的交换来实现 IP 数据包的转发。标记作为 IP 包头在网络中的替代品而存在，在网络内部 MPLS 在数据包所经过的路径沿途通过交换标记(而不是看 IP 包头)来实现转发；当数据包要退出 MPLS 网络时，数据包被解开封装，继续按照 IP 包的路由方式到达目的地。



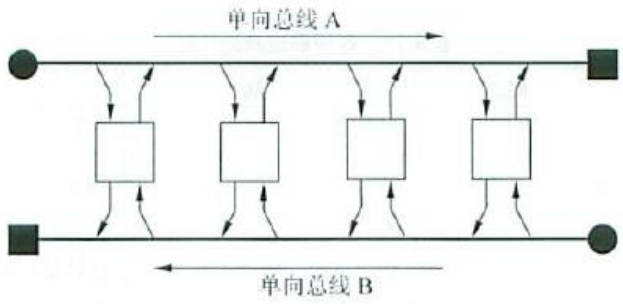
如图所示，MPLS 网络包含一些基本的元素。在网络边缘的节点就称做标记边缘路由器(LER)，而网络的核心节点就称做为标记交换路由器(LSR)。LER 节点在 MPLS 网络中完成的是 IP 包的进入和退出过程;LSR 节点在网络中提供高速交换功能。在 MPLS 节点之间的路径就叫做标记交换路径。一条 LSP 可以看做是一条贯穿网络的单向隧道。

MPLS 术语的缩写

- LDP（Label Distribution Protocol），标记分配协议
- LSP（Label Switched Path），标记交换路径
- FEC（Forwarding Equivalence Class），转发等价类
- LSR（Label Switching Router），标记交换路由器
- LER（Label Edge Router），标记边缘路由器
- CR-LDP（Constraint Route Label Distribution Protocol），限制路由的标记分配协议

3．DQDB 网络

IEEE802.6 城域网采用了分布式队列双总线（Distributed Queue Dual Bus, DQDB）协议。这种双总线一般采用光纤介质，如图所示，在这种配置中，每个站同时连接到两根总线上。一个站要发送数据时必须选择一根总线，使接收站成为它的下游站。



光纤双总线网结构图

图中的黑圆点表示 A 总线和 B 总线的端头，它们不停地产生固定长度为 53 字节的时槽。当时槽沿沿着总线流动到达末端时由终端匹配器（黑方块）吸收。结点可以从忙时槽中读数据，也可以向空时槽中写数据。总线的运行由周期为 125μs 的时钟控制，一个时钟周期内端头可以产生多个时槽，时槽数量决定

了总线的实际速率。

时槽分为两类，一类叫作排队仲裁（Queue-Arbitrated）时槽，用 QA 表示，用于分组交换业务；另一类叫作预仲裁（Pre-Arbitrated）时槽，用 PA 表示，由电路交换业务使用，可提供等时服务。时槽由 1 个控制字节和 52 字节长的段组成，段头 4 字节，信息实际占 48 字节。

DQDB 是一个很有效的协议，可与 802.3 和 802.5 媲美。在重负载下，CD 的值很小，甚至为 0，空时槽是很富裕的。这样就像 CSMA/CD 协议一样，可以很快的访问信道，几乎没有延迟。在重负载下，实际上每个时槽都被等待发送的站利用，信道利用率达到 100%，性能 一点不比令牌环网差。轻负载下的快速访问和重负载下可预见的排队系统的奇妙结合使得 DQDB 成为最适合 MAN 的协议。

4．LANE：LAN Emulated

ATM 论坛开发了 ATM LAN 仿真标准（LAN Emulated, LANE），用以解决不同局域网上的端系统相互作用问题，使得现有的共享介质网络上的主机也可以通过 ATM 网络进行通信。ATM-LAN 转换器对 ATM 信源流和 MAC 帧进行转换。ATM 论坛提议使用 AAL5 对 MAC 帧进行分段和重装配。

在概念上 LANE 结构包括客户机和服务器两种成分：

● LAN 仿真客户机（LEC）:是指传统的网络设备，它按照原有的 MAC 协议进行操作，并具有唯一的 ATM 地址。例如 LAN 交换机就可以被看成 LEC，它有一个 ATM 地址，同时又通过各个端口访问对应的 MAC 网卡。

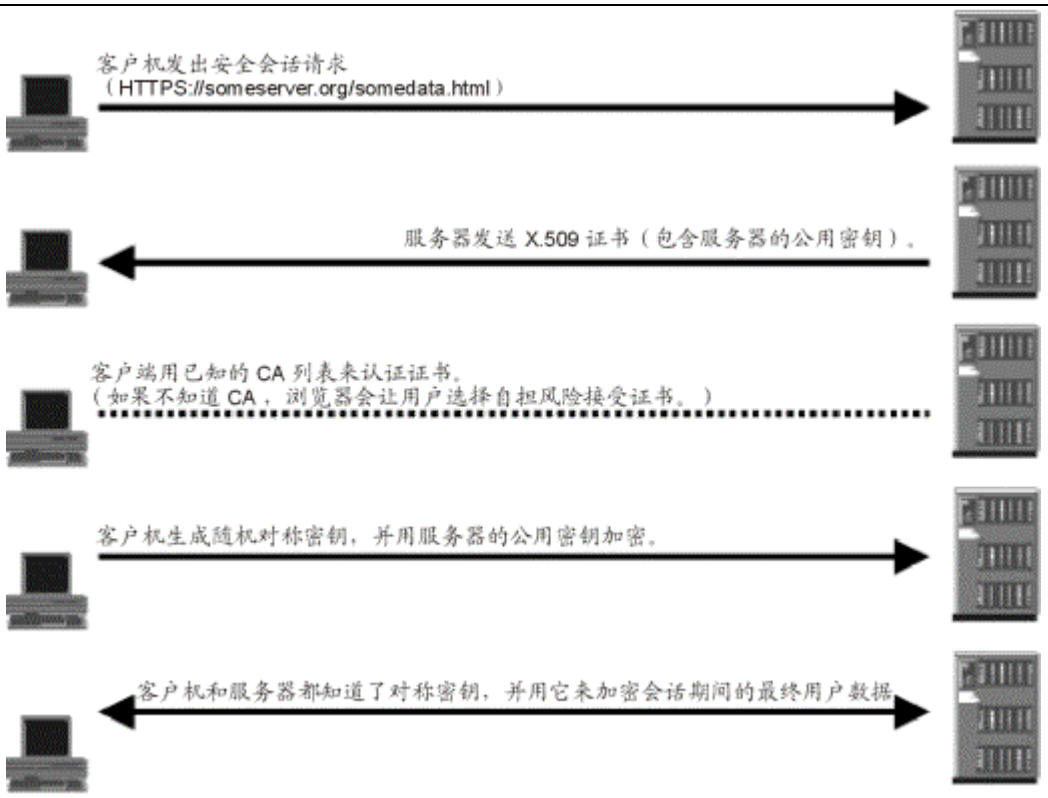
●LAN 仿真服务器（LES）:具有唯一的 ATM 地址，它提供控制功能，处理 LEC 的连接请求，并为它管理的一组 LEC 建立 MAC 地址与 ATM 地址的映像表。

●LAN 仿真配置服务器（LECS）:其作用是根据配置数据库的信息和 LEC 的请求把 LEC 分配个特定的 LES，每个管理域中只有一个 LECS。

●广播和未知服务器（BUS）：第一个作用是实现广播和组播的传输，所有目标地址为广播地址（全 1）的数据帧都被发送给 BUS 传送。它的另外一个作用是利用广播功能实现未知地址的解析，即建立 ATM 地址与 MAC 地址的映像。

几种比较重要的认证方式

1.SSL 认证过程

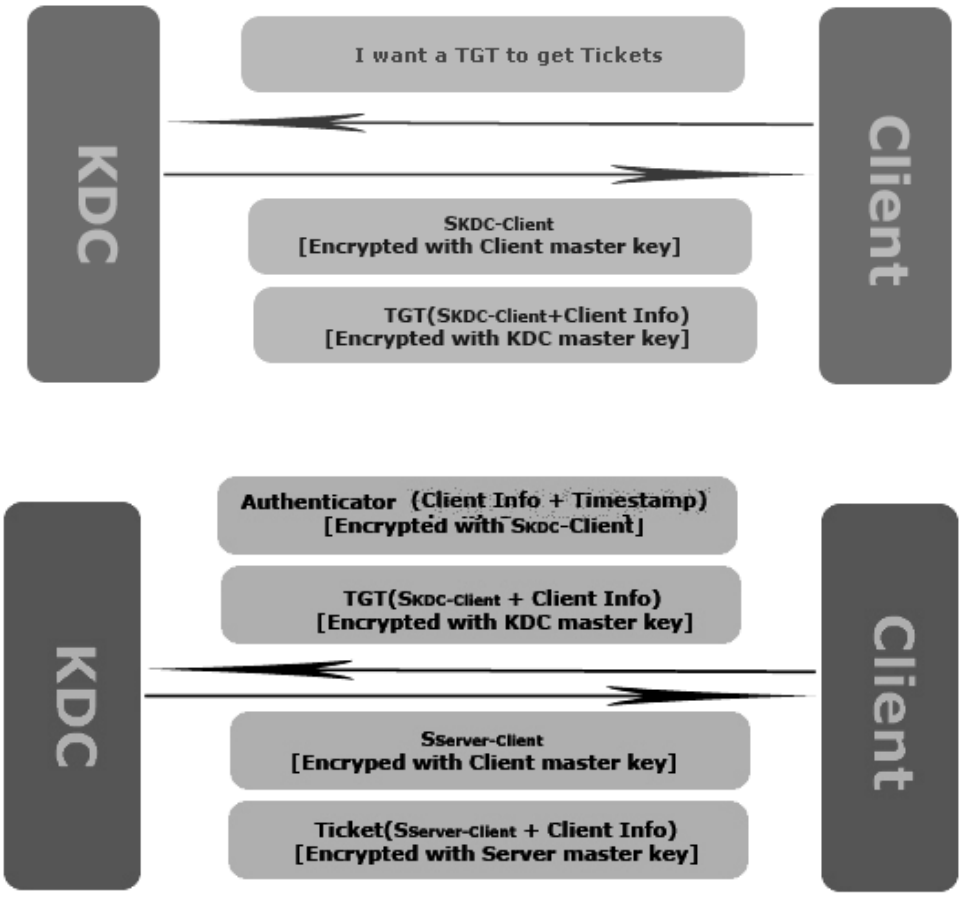


安全套接字协议 SSL（Secure Socket Layer）工作在应用层和传输层之间，提供身份认证和保密通信功能。SSL 所包含的协议有 SSL 握手协议、SSL 修改密文协议、SSL 警告协议和 SSL 记录协议。SSL 握手协议负责身份认证和密钥生成。SSL 记录协议负责接收应用层报文，并将数据划分成可管理的块（214 个字节），选择性地压缩数据，应用报文认证码（MAC）对数据进行加密，并增加首部，通过 TCP 报文段传输数据；接收者将数据进行解密、验证、解压，重装配成应用报文，然后交付更高级的用户。

在客户端与服务器间传输的数据是通过使用对称算法（如 DES 或 RC4）进行加密的。公用密钥算法（通常为 RSA）是用来获得加密密钥交换和数字签名的，此算法使用服务器的 SSL 数字证书中的公用密钥。有了服务器的 SSL 数字证书，客户端也可以验证服务器的身份。SSL 协议的版本 1 和 2 只提供服务器认证。版本 3 添加了客户端认证，此认证同时需要客户端和服务器的数字证书。

SSL 安全连接建立在 TCP443 端口，统一资源定位器使用 HTTPS://开头。

2. Kerberos 认证



整个流程大体上包含以下 3 个子过程：

1. Client 向 KDC 申请 TGT（Ticket Granting Ticket）。
2. Client 通过获得 TGT 向 KDC 申请用于访问 Server 的 Ticket。
3. Client 最终向为了 Server 对自己的认证向其提交 Ticket。

3．WLAN 接入认证方式

WLAN 开放的传输介质使得只要符合协议要求的无线系统均可能在信号覆盖范围内收到所有信息，为达到和有线网络同等的安全性能，IEEE802.11 采取了认证和加密措施。

认证程序控制 WLAN 接入的能力，这一过程被所有无线终端用来建立自己的合法接入到 AP 的身份标志，如果 AP 和工作站之间无法完成相互间的认证，那么它们之间就不能建立有效的连接。IEEE802.11 协议支持多个不同的认证过程，并且允许认证方案扩充。

IEEE802.11 提供的加密方式采用 WEP 机制，WEP 对数据的加密和解密使用同样的算法和密钥。它包括“共享密钥”认证和数据加密两个过程。“共享密钥”认证使得那些没有正确 WEP 密钥的用户无法访问网络，而加密则要求网络中所有数据的发送和接收都必须使用密钥加密。

认证采用了一个标准的询问和响应帧格式。执行过程中，AP 根据 RC4 算法运用共享密钥对 128 字节的

随机序列进行加密后作为询问帧发给用户，用户将收到的询问帧进行解密后以正文形式响应 AP,AP 将正文与原始序列进行比较，如果两者一致，则通过认证。

路由协议的总结

协议名称	协议概述	报文类型	传输协议	维护与更新	路由配置要点
BGP-4	BGP 是一种不同自治系统的路由器之间进行通信的外部网关协议。BGP 的主要功能是控制路由策略，BGP 系统与其他 BGP 系统之间交换网络可到达信息。这些信息包括数据到达这些网络所必须经过的自治系统 AS 中的所有路径。这些信息足以构造一幅自治系统连接图。然后，可以根据连接图删除选路环，制订选路策略。BGP 是一个距离向量协议。支持无类别的域间路由（CIDR）	1.建立（open）：建立邻居关系 2. 保持活动状态（keepalive）：对 open 报文的应答/周期的确认邻居关系 3.更新：发送新的路由信息 4.通告：报告检测到的错误	BGP 邻居之间通过 TCP 连接交换路由信息，使用端口号为 179	BGP 通过定期发送 keepalive 报文给其邻站来检测 TCP 连接对端的链路或主机失败。初始连接建立时要发送全部路由信息，以后只发送改变了的路由信息。BGP 路由器不需要进行周期性路由更新。	
RIP	RIP 协议的全称是路由信息协议，它是一种内部网关协议（IGP），用于一个自治系统（AS）内的路由信息的传递。RIP 协议是基于距离矢量算法（Bellman-ford）的，它使用“跳数”，衡量到达目标地址的路由距离。RIP 使用非常广泛，简单可靠便于配置。RIPv2 支持 CIDR、VLSM 和不连续子网，使用组播地址（224.0.0.9）而不是广播传播路由更新报文，并且采用了触发更新机制来加速路由收敛。RIPv2 支持认证，使用经过散的口令字来限制更新信息的传播。RIP 只适用小型同构网络，允许的最大跳数为 15，任何超过 15 个站点的目的地均被标记为不可达。		RIP 使用 UDP 作为其传输层协议，端口为 520	RIP 的更新是经过定时广播实现的，在默认情况下，路由器每隔 30 秒向相联的网络广播自己的路由表，收到广播的路由器将收到的信息添加到自身的路由表，每个路由器都如此广播，最终网络上的路由器将得知全网的路由信息。正常情况下，路由器每 50 秒就可以得到一条路由的信息确认，经过 180 秒 6 个更新周期一个路由项没有被确认，路由器就认为该路由器失效，若经过 240 秒路由项没有得到确认，就将该路由器从路由表中删除。	**router rip** **network** *network*（相连网络，即相连的网段号） **version** 1/2 **show ip route** **show ip route rip** **no logging console**（防止大量端口状态变化和报警信息对配置过程的影响）
IGRP	内部网关路由协议（IGRP）是 Cisco 公司 20 世纪 80 年代开发的，是一种动态的、长跨度（最大可支持 255 跳）的路由协议，使用度量（向量）来确定到达一个网络的最佳路由，由延时、带宽、可靠性和负载等来计算最优路由，它在同个自治系统内具有高跨度，适合复杂的网络，但本质上讲，IGRP 还是一种距离矢量路由协议。		IGRP 使用 UDP 发送路由表项	默认情况下，IGRP 路由器每隔 90s 更新一次路由信息，如果 270s 内没有收到某路由器的回应，则认为该路由器不可到达；如果 630s 内仍未收到应答，则 IGRP 进程将从路由表中删除该路由。	**router igrp** *AS number*（创建 IGRP 路由进程） **network** *network*（相连网络） **clockrate**（DCE 端串口配置时钟信号，用于同步） **bandwidth**（指定相应端口带宽）

	Cisco IOS 允许路由器管理员对 IGRP 的网络带宽、延时、可靠性和负载进行权重设置,以影响度量的计算。IGRP 不支持 VLSM 和不连续子网。				**no keepalive**（使以太网接口不监测 keepalive 信号，从而在不连接任何设备情况下，可以激活此端口） **show ip route** **show ip route igrp**
OSPF	开放式最短路径优先（OSPF）是一种链路状态选择协议，是由 IETF 开发的内部网关路由协议，基于 Dijkstra 算法。OSPF 的链路状态信息通过链路状态公告（LSA）发布到网上的每台路由器,每台路由器通过 LSA 建立一个关于网络的拓扑数据库。在一个区域（Area）中的路由器（区域边界路由器除外），都应具有相同的链路状态数据库。OSPF 是一种层次化的路由选择协议,区域 0（也称主干区域）是 OSPF 网络中必须具有的区域，其他所有区域要求与区域 0 互连到一起。OSPF 采用触发更新，支持 VLSM 及 CIDR，对跳数没有限制	1.Hello 数据包：用于建立和维护邻居关系，在广播网络中 Hello 分组还用于动态发现邻居路由器 2.链路状态更新数据包：向邻居路由器发送链路状态公告（LSA） 3.链路状态应答数据包：对链路状态更新数据包的应答 4.数据库描述数据包：描述一个路由器 OSPF 链路状态数据库的内容 5.链路状态请求数据包：请求相邻路由器发送其链路数据库中的具体条目	OSPF 路由信息利用 IP 数据报直接传送,IP 数据报的报头中“协议”字段的值为 89 （即协议号为 89）	OSPF 路由器以固定的时间间隔，通常为 10 秒，发送 Hello 数据包建立和维护邻居路由器间的关系。如果 40 秒没有从特定邻居收到 Hello 分组，路由器就认为那个邻居不存在了，并且产生声明该邻居丢失的 LSA。Hello 定时器的值可以改变，但是在一个网段中所有路由器的定时器必须保持一致，在稳定状态下，大的链路状态更新分组 30 分钟才传送一次。每一个区域都具有该区域专用的链路状态数据库。一个区域的网络拓扑结构在区域外是不可见的。同样，每一个区域内路由器对区域外的网络结构也不了解，也就是说，区域内的 LSA 广播被区域边界挡住了，这样就减少了网络中的广播数据包，也减少了链路状态数据库的大小。随着区域概念的引入,AS 内的所有路由器不再具有相同的链路状态数据库，而是只具有所在区域的链路状态数据库。区域边界路由器则具有与其相连的所有区域的链路状态数据库。	**router ospf** *process-id*（指定使用 OSPF 协议，进程号只在路由器内部起作用，不同路由器的进程号可以不同） **network** *address wildcard-mask* **area** *area-id*（指定与该路由器相连的网络，区域号为十进制数，0 为主干区域） **show ip route** **show ip route ospf**
EIGRP	EIGRP 是增强型的 IGRP 协议，是	1.hello 包：用于邻	EIGRP 使用 RTP（可	EIGRP 通过使用 hello 数	**router eigrp** *AS*

	典型的平衡混合路由选择协议，融合了距离矢量和链路状态两种路由选择协议的优点，使用弥散修正算法（DUAL）快速收敛，采用不定期更新（触发更新）以减少带宽消耗。EIGRP 支持 VLSM 及不连续子网。EIGRP 最大的跳数限制为 224。支持对自动路由汇总功能的设定，支持多种网络层协议，支持 IP、IPX、AppleTalk、Novell 等。	居发现与恢复，组播方式发送（224.0.0.10） 2.更新包：当路由器发现新邻居时使用更新包（单播） 3.确认包：对更新包的确认 4.查询包：当 EIGRP 路由器想从特定邻居或所有邻居那里获得特定的信息，EIGRP 路由器使用查询包（单播或多播） 5.响应包：对查询包的响应	靠的传输层协议）传输数据包，不依赖 TCP/IP 协议。RTP 支持多播、组播和单播	据包与邻居路由器建立联系，缺省情况下，hello 数据包每间隔 5 秒发送一次。	*number*（创建 EIGRP 路由进程） **network** *address wildcard-mask*（EIGRP 的网段声明中，如果是没有划分子网的 A、B、C 类主网地址，只需输入此网络地址；如果网络划分了子网，则必须在网络地址后面写入反掩码） **no auto-summary**（关闭 EIGRP 协议的路由汇总功能，默认配置是自动汇总生效。在处理 VLSM 时，通常需要关闭该功能）