第1 章概述

1 计算机网络的作用

2 计算机网络的发展过程

分组交换的产生,因特网时代,关于因特网的标准化工作,计算机网络在我国的发展

3 计算机网络的分类

计算机网络的不同定义, 几种不同的分类方法

4 计算机网络的主要性能指标

带宽, 时延,时延带宽积和往返时延

5 计算机网络的体系结构

计算机网络体系结构的形成,划分层次的必要性,计算机网络的原理体系结构, 实体、协议、服务和服务访问点, 面向连接服务和无连接服务,OSI与TCP/IP体系结构的比较

1 计算机网络的作用

21 世纪的一些重要特征就是数字化、网络化和信息化，它是一个以网络为核心的信息时代。

网络现已成为信息社会的命脉和发展知识经济的重要基础。

网络是指“三网”，即电信网络、有线电视网络和计算机网络。

发展最快的并起到核心作用的是计算机网络。

2 计算机网络的发展过程:分组交换的产生

传统的电路交换(circuit switching)的电信网有一个缺点：正在通信的电路中有一个交换机或有一条链路被炸毁，则整个通信电路就要中断。

如要改用其他迂回电路，必须重新拨号建立连接。这将要延误一些时间。

新型网络的基本特点

网络用于计算机之间的数据传送，而不(仅)是为了打电话。

网络能够连接不同类型的计算机，不局限于单一类型的计算机。

所有的网络结点都同等重要，因而大大提高网络的生存性。

计算机在进行通信时，必须有冗余的路由。

网络的结构应当尽可能地简单，同时还能够非常可靠地传送数据。

计算机网络涉及到通信与计算机两个领域。

计算机与通信的相互结合主要有两个方面：

一方面，通信网络为计算机之间的数据传递和交换提供了必要的手段；

另一方面，数字计算技术的发展渗透到通信技术中，又提高了通信网络的各种 性能。

当然，这两个方面的进展都离不开人们在半导体技术(主要是超大规模集成电路VLSI-Very Large ScaleIntegrated circuit技术)上取得的辉煌成就

5 部电话机两两相连，需10 对电

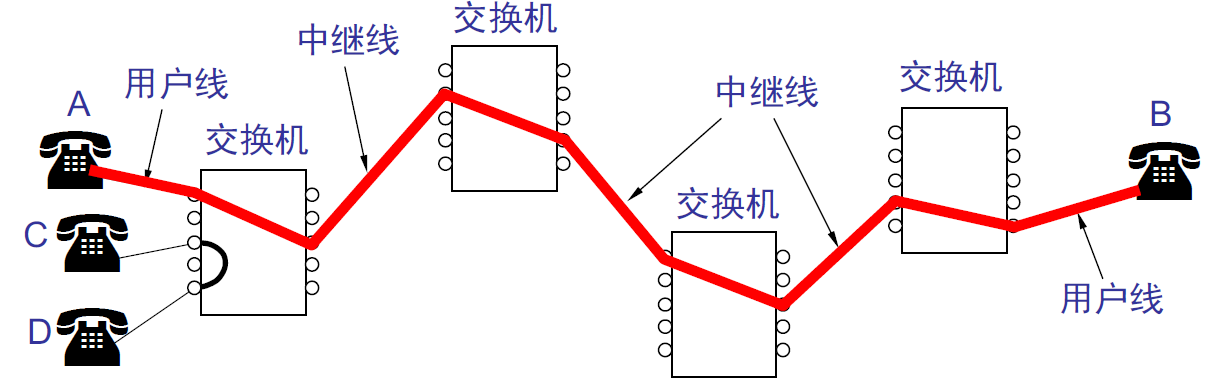
N 部电话机两两相连，需N(N – 1)/2对电线。

当电话机的数量很大时，这种连接方法需要的电线对的数量与电话机数的平方成正比。

电路交换的特点:

电路交换必定是面向连接的。

电路交换的三个阶段：建立连接,通信,释放连接



网络核心: 电路交换

分组交换的原理1): 在发送端，先把较长的报文划分成较短的、固定长度的数据段 2) 每一个数据段前面添加上首部构成分组3) 分组交换网以“分组”作为数据传输单元。依次把各分组发送到接收端（假定接收端在左边）。4) 接收端收到分组后剥去首部还原成报文5) 最后，在接收端把收到的数据恢复成为原来

的报文。\*(我们假定分组在传输过程中没有出现差错，在转发时也没有被丢弃)

分组首部的重要性: 每一个分组的首部都含有地址等控制信息。分组交换网中的结点交换机根据收到的分组的首部中的地址信息，把分组转发到下一个结点交换机。用这样的存储转发方式，最后分组就能到达最终目的地。

分组交换vs.电路交换: 我们以饭店提供的两种经营方式做个比喻:一种是，提供并接受预订服务（人大贤进楼）;另一种是，不要求预订也不接受预订服务（人大留学生楼）。

分组交换: 存储转发的过程:

将报文划分成较小的数据块: “分组（packets）”

存储转发: 交换机等到整个分组到达完毕后, 再进行转发或路由接力

Q: 如果报文以整个的形式发送又将如何? 参考教科书p127,图2-43

分组交换vs. 电路交换: 分组交换使得更多用户可“同时”使用网络!

1 Mb/s链路每个用户: 100Kb/s ，当“激活”时,激活时间为10%

电路交换: 10 用户, 分组交换:对35个用户来说, 概率: > 10个用户同时激活小于.0004

比较:

1)电路交换预先静态地保留所要带宽；而分组交换却是根据需要动态地获得和释放带宽

2)在电路交换中，已分配给线路上未用的带宽只能浪费掉；而分组交换中，未用的带宽可以被别的传输所利用。

3)电路交换对数据传输是完全透明的，而分组交换则要利用分组所携带的参数进行路由转发。

4)分组交换是存储转发的，会增加传输延时；电路交换则是连续的通过物理线路传输。

5)分组交换中，分组到达目的地可能不按原顺序；在电路交换中，不会发生乱序现象。

6)分组交换按传输的字节和连接时间计费；电路交换按时间、距离计费。

从主机为中心到以网络为中心

今日的多级结构的因特网

大致上可将因特网分为以下五个接入级1)网络接入点NAP 2)国家主干网（主干ISP）3)地区ISP 4)本地ISP 5)校园网、企业网或PC 机上网用户

网络时代的三大基本定律: 摩尔定律:CPU性能18个月翻番,10年100倍。所有电子系统（包括电子通信系统，计算机）都适用; 光纤定律：超摩尔定律，骨干网带宽9个月翻番，10年10000倍。带宽需求呈超高速增长的趋势; 梅特卡夫定律：联网定律，网络价值随用户数平方成正比。未联网设备增加N倍，效率增加N倍。联网设备增加N倍，效率增加N2倍

IT业三大定律：

集成电路、通信网络、软件是IT业的三大支柱。

l 三大定律中一条涉及集成电路及计算机技术，两条与通信网络有关。表明了网络在IT中的重要地位。

l 梅特卡夫定律中指的网络是数据通信网或者计算机网，该定律定量描述了单机与联网之间的巨大差别和网络的重要价值。

l 梅特卡夫定律还反映了数据通信网相对于其它通信网络的效率差别。电话网中每部电话的效率是1:1,电视网中每台电视机的效率是1:N,而数据通信网端节点的效率为N2。

l 根据吉尔特定律与摩尔定律得知，通信网络的发展大大快于集成电路与计算机系统的发展。

计算机网络的分类:

最简单的定义：计算机网络是一些互相连接的、自治的计算机的集合。因特网(Internet)是“网络的网络”

几种不同的分类方法

从网络的交换功能进行分类

从网络的作用范围进行分类

从网络的使用者进行分类

从网络的交换功能分类1)电路交换2)报文交换3)分组交换4)混合交换

从网络的作用范围进行分类1)广域网WAN (Wide Area Network) 2)局域网LAN (Local Area Network) 3)城域网MAN (Metropolitan Area Network) 4)接入网AN (Access Network)

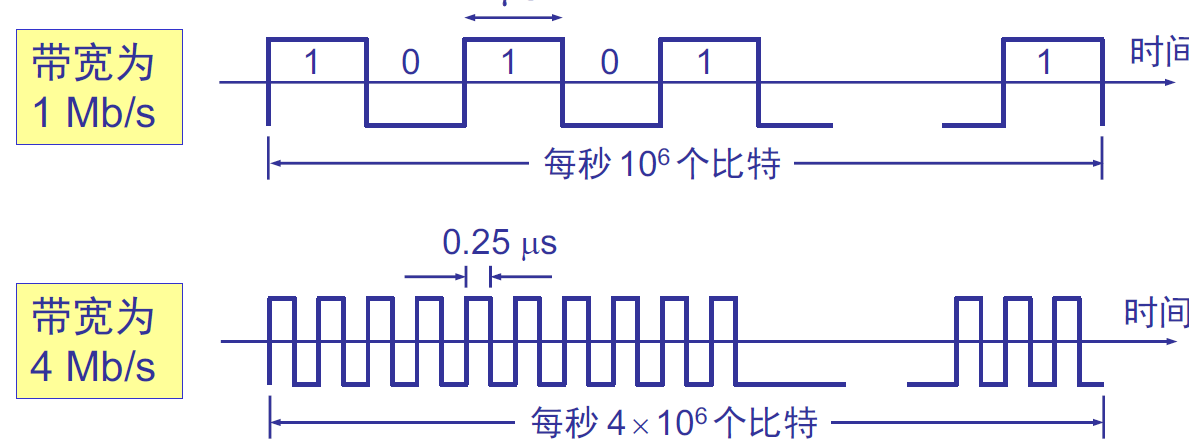
从网络的使用者进行分类1)公用网(public network)2)专用网(private network)

计算机网络的主要性能指标:带宽

“带宽”(bandwidth)本来是指信号具有的频带宽度，单位是赫（或千赫、兆赫、吉赫等）。

现在“带宽”是数字信道所能传送的“最高数据率”的同义语，单位是“比特每秒”，或b/s (bit/s)。

在时间轴上信号的宽度随带宽的增大而变窄

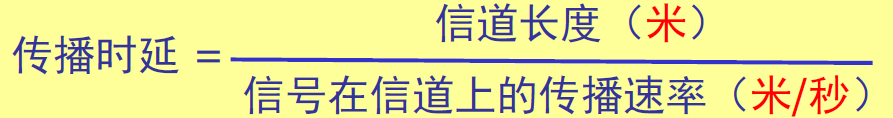


发送时延（传输时延） 发送数据时，数据块从结点进入到传输媒体所需要的时间。

信道带宽数据在信道上的发送速率。常称为数据在信道上的传输速率。

传播时延电磁波在信道中需要传播一定的距离而花费的时间。

信号传输速率（即发送速率）和信号在信道上的传播速率是完全不同的概念。



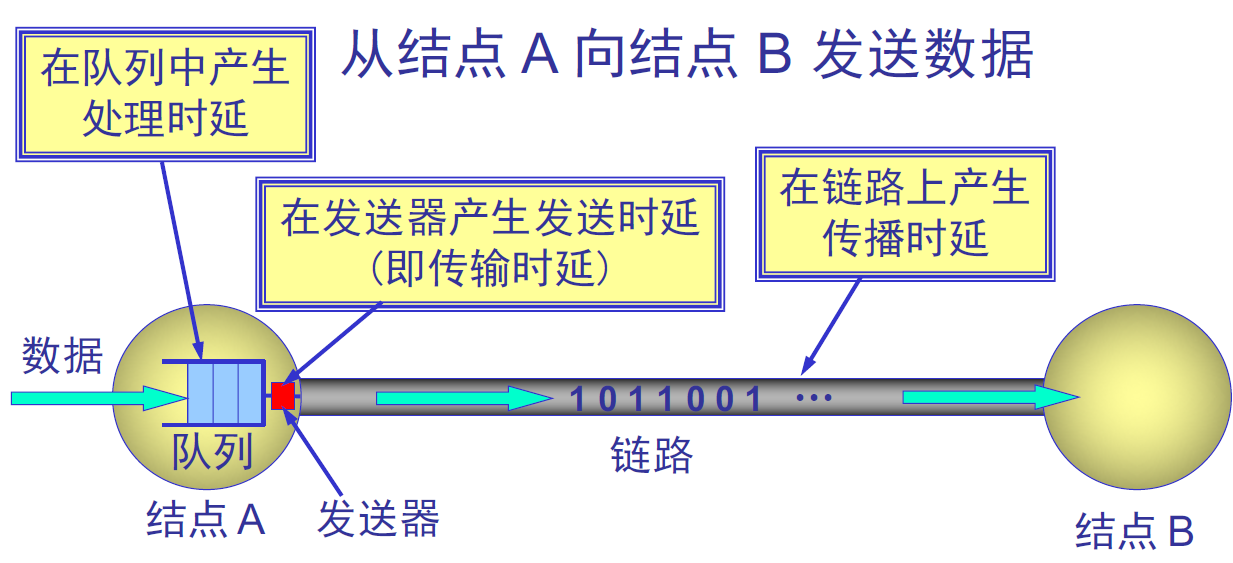
处理时延交换结点为存储转发而进行一些必要的处理所花费的时间。

结点缓存队列中分组排队所经历的时延是处理时延中的重要组成部分。

处理时延的长短往往取决于网络中当时的通信量。

有时可用排队时延作为处理时延。

数据经历的总时延就是发送时延、传播时延和处理时延之和：



容易产生的错误概念:

1)对于高速网络链路，我们提高的仅仅是数据的发送速率而不是比特在链路上的传播速率。2) 提高链路带宽是减小了数据的发送时延。

往返时延RTT (Round-Trip Time) 表示从发送端发送数据开始，到发送端收到来自接收端的确认（接收端收到数据后立即发送确认），总共经历的时延。

关于开放系统互连参考模型:OSI/RM

只要遵循OSI 标准，一个系统就可以和位于世界上任何地方的、也遵循这同一标准的其他任何系统进行通信。

在市场化方面OSI 却失败了。

OSI 的专家们在完成OSI 标准时没有商业驱动力；

OSI 的协议实现起来过分复杂，且运行效率很低；

OSI 标准的制定周期太长，因而使得按OSI 标准生产的设备无法及时进入市场；

OSI 的层次划分并也不太合理，有些功能在多个层次中重复出现。

法律上的(de jure)国际标准OSI 并没有得到市场的认可。

是非国际标准TCP/IP 现在获得了最广泛的应用, TCP/IP 常被称为事实上的(de facto) 国际标准。

为什么要分层?

对于复杂的系统:1)显式的结构使得复杂系统的问题定位和不同组成部分之间的关联讨论称为可能;分层的参考模型(reference model) 可用于讨论

2)模块化简化了系统的维护和升级;某个层次服务实现对系统的其余部分是透明的

分层的做法有没有坏处?

网络协议的组成要素:

语法:数据与控制信息的结构或格式。

语义:需要发出何种控制信息，完成何种动作以及做出何种响应。

同步:事件实现顺序的详细说明。

分层的好处:

各层之间是独立的。灵活性好。结构上可分割开。易于实现和维护,能促进标准化工作。

层数多少要适当:

若层数太少，就会使每一层的协议太复杂。

层数太多又会在描述和综合各层功能的系统工程任务时遇到较多的困难。

计算机网络的体系结构:

计算机网络的体系结构(architecture)是计算机网络的各层及其协议的集合。

体系结构就是这个计算机网络及其部件所应完成的功能的精确定义。

实现(implementation)是遵循这种体系结构的前提下用何种硬件或软件完成这些功能的问题。

体系结构是抽象的，而实现则是具体的，是真正在运行的计算机硬件和软件。

协议分层和数据的封装

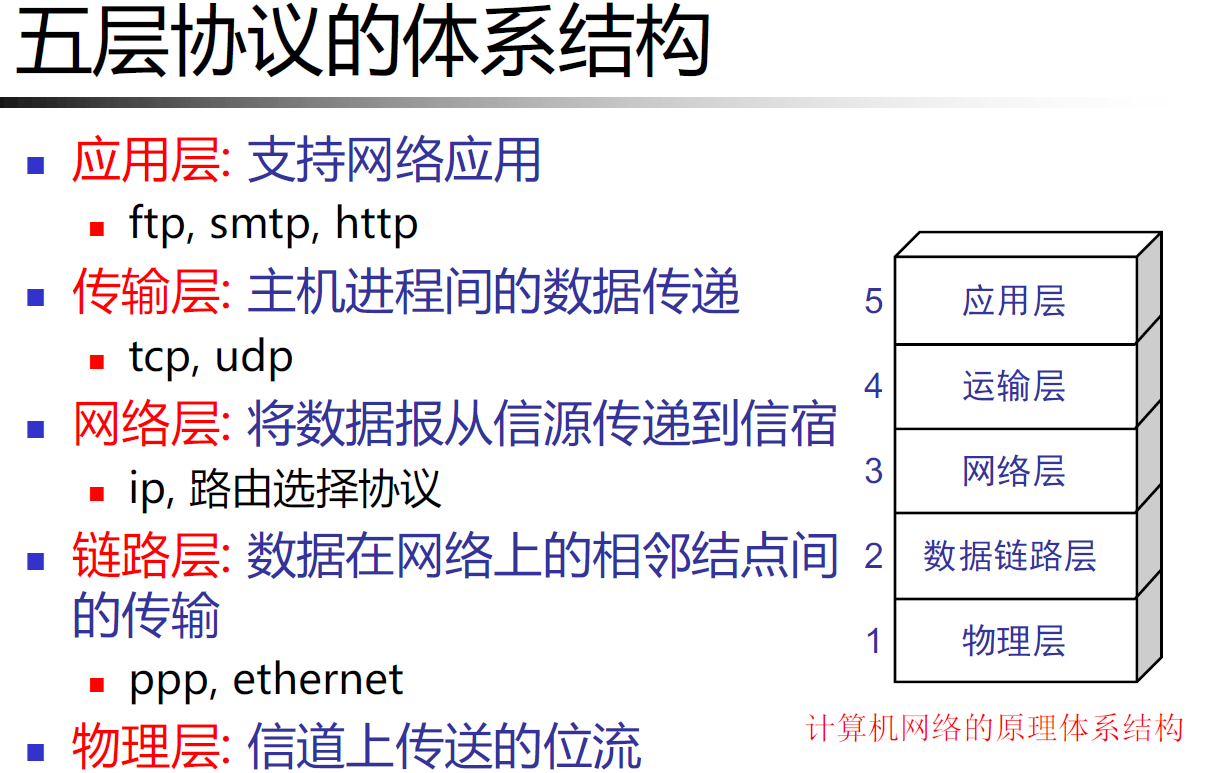
每个层次都从上层取得数据:1)加上首部信息形成新的数据单元2)将新的数据单元传递给下一层次

五层协议的体系结构:

TCP/IP 是四层的体系结构：应用层、运输层、网际层和网络接口层。

最下面的网络接口层并没有具体内容。因此往往采取折中的办法，即综合OSI 和

TCP/IP 的优点，采用一种只有五层协议的体系结构。



计算机1 向计算机2 发送数据:

应用进程数据先传送到应用层,加上应用层首部，成为应用层PDU->

应用层PDU 再传送到运输层,加上运输层首部，成为运输层报文->

运输层报文再传送到网络层,加上网络层首部，成为IP 数据报（或分组）->

IP 数据报再传送到数据链路层,加上链路层首部和尾部，成为数据链路层帧->

数据链路层帧再传送到物理层,最下面的物理层把比特流传送到物理媒体->

电信号（或光信号）在物理媒体中传播,从发送端物理层传送到接收端物理层->

计算机2

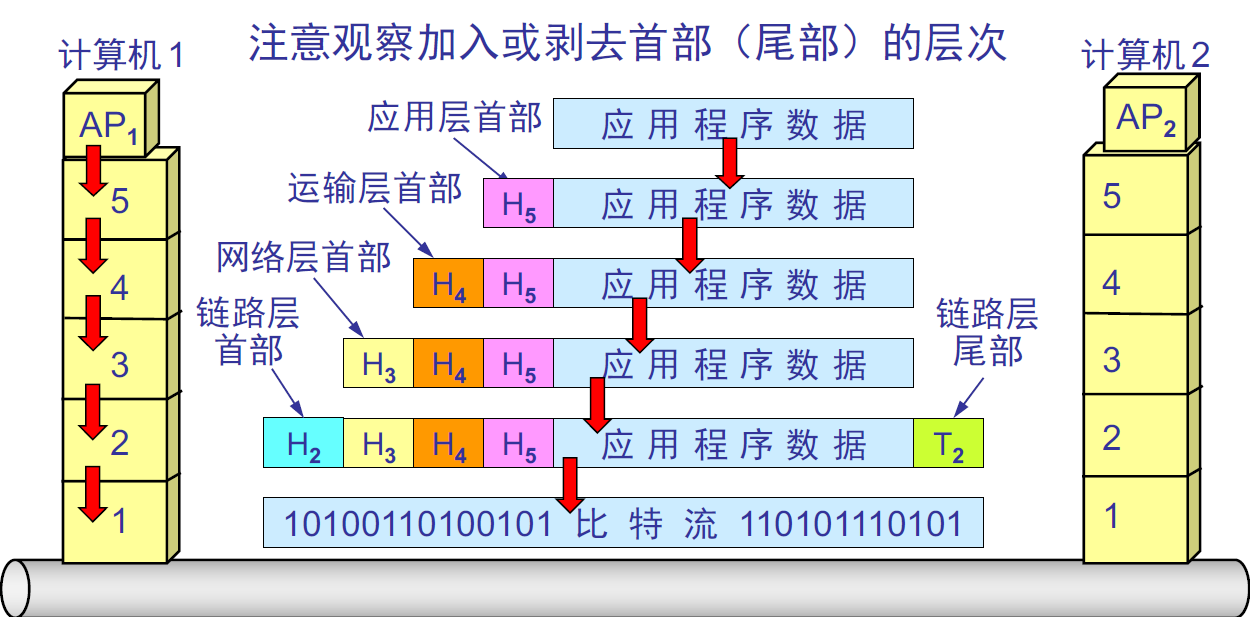
物理层接收到比特流，上交给数据链路层->

数据链路层剥去帧首部和帧尾部,取出数据部分，上交给网络层->

网络层剥去首部，取出数据部分,上交给运输层->

运输层剥去首部，取出数据部分,上交给应用层->

应用层剥去首部，取出应用程序数据,上交给应用进程->





实体、协议、服务和服务访问点:

实体(entity) 表示任何可发送或接收信息的硬件或软件进程;

协议是控制两个对等实体进行通信的规则的集合;

在协议的控制下，两个对等实体间的通信使得本层能够向上一层提供服务。

要实现本层协议，还需要使用下层所提供的服务。

本层的服务用户只能看见服务而无法看见下面的协议。

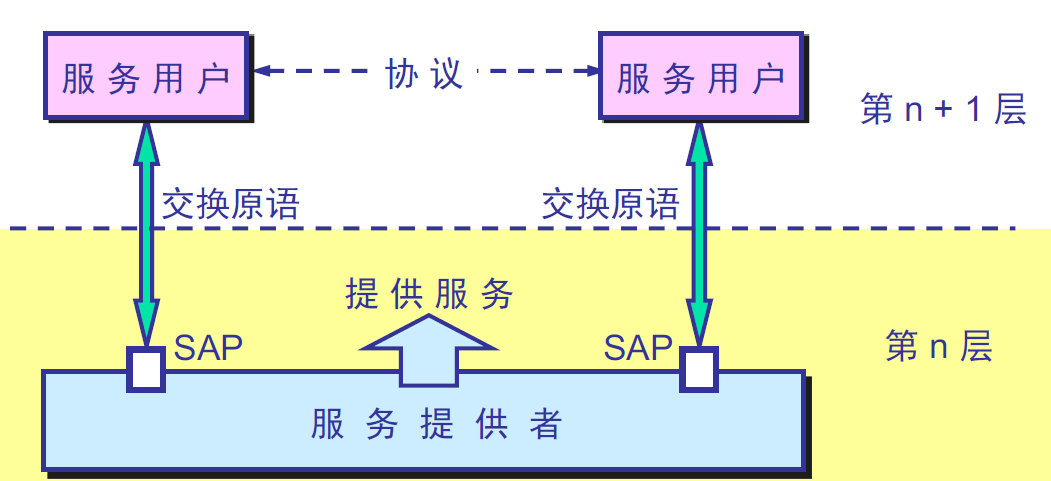
下面的协议对上面的服务用户是透明的。

协议是“水平的”，即协议是控制对等实体之间通信的规则。

服务是“垂直的”，即服务是由下层向上层通过层间接口提供的。

同一系统相邻两层的实体进行交互的地方，称为服务访问点SAP (Service Access

Point)。



协议： 定义网络实体之间信息收发的格式和顺序, 以及信息发送和接收后所需采取的动作（语法、语义、同步或规则）

协议通常很复杂:

协议必须将各种不利的条件事先都估计到，而不能假定一切情况都是很理想和很顺利的。

必须非常仔细地检查所设计协议能否应付所有的不利情况。

应当注意：事实上难免有极个别的不利情况在设计协议时并没有预计到.在出现这种情况时，协议就会失败.因此实际上协议往往只能应付绝大多数的不利情况。

**面向连接服务与无连接服务**

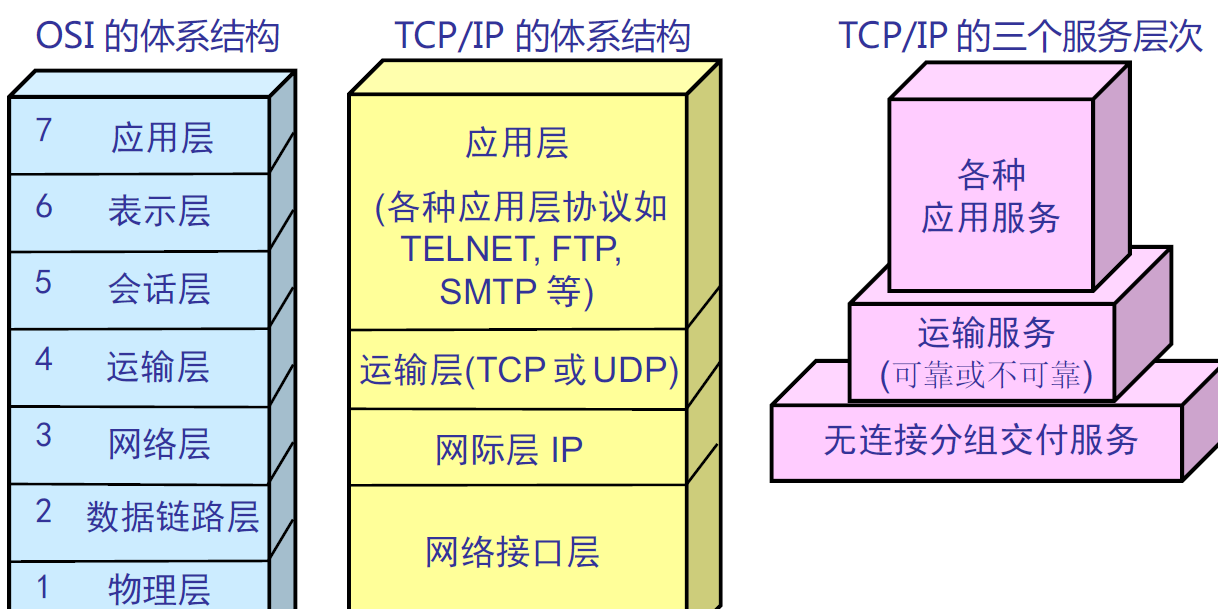
面向连接服务具有连接建立、数据传输和连接释放这三个阶段。

无连接服务(connectionless)

两个实体之间的通信不需要先建立好连接。

是一种不可靠的服务。这种服务常被描述为“尽最大努力交付”(best effort delivery)或“尽力而为”。

**OSI 与TCP/IP体系结构的比较**



本章小结:

计算机网络体系结构

**计算机网络概述**

计算机网络的概念、组成与功能

计算机网络的分类

计算机网络与互联网的发展历史

计算机网络的标准化工作及相关组织

**计算机网络体系结构与参考模型**

计算机网络分层结构

计算机网络协议、接口、服务等概念

ISO/OSI参考模型和TCP/IP模型

**网络性能指标**:带宽、时迟、时延带宽积

第2 章:物理层

2.1 物理层的基本概念

2.2 数据通信的基础知识

数据通信的理论基础

数据通信系统的模型

有关信道的几个基本概念

信道的最高码元传输速率

信道的极限信息传输速率

2.3 物理层下面的传输媒体

导向传输媒体

非导向传输媒体

2.4 模拟传输与数字传输

模拟传输系统

调制解调器

数字传输系统

2.5 信道复用技术

频分复用、时分复用和统计时分复用

波分复用

码分复用

2.1 物理层的基本概念

物理层的定义: ISO/OSI 关于物理层的定义：物理层提供机械的、电气的、功能的和规程的特性，目的是启动、维护和关闭数据链路实体之间进行比特传输的物理连接。这种连接可能通过中继系统，在中继系统内的传输也是在物理层的。

物理层的功能:在两个网络设备之间提供透明的比特流传输。

研究内容:物理连接的启动和关闭，正常数据的传输，以及维护管理;

物理层考虑的是:怎样才能在连接各种计算机的传输媒体上传输数据的比特流，而不是指连接计算机的具体的物理设备或具体的传输媒体。

用于物理层的协议也常称为物理层规程(procedure)。

几点说明: 连接方式（点到点，点到多点）

通信方式（单工，半双工，全双工）

位传输方式（串行，并行）

物理层的四个重要特性1)机械特性(mechanical characteristics) 2)电气特性(electrical characteristics) 3)功能特性(functional characteristics) 4)规程特性(procedural characteristics)

物理层的机械特性: 主要定义物理连接的边界点，即接插装置。规定物理连接时 所采用的规格、引脚的数量和排列情况。

常用的标准接口:

ISO 2110，25芯连接器，EIA RS-232-C，EIA RS-366-A

ISO 2593，34芯连接器，V.35宽带MODEM

ISO 4902，37芯和9芯连接器，EIA RS-449

ISO 4903，15芯连接器，X.20、X.21、X.22

物理层的电气特性:

规定传输二进制位时，线路上信号的电压高低、阻抗匹配、传输速率和距离限制。

早期的标准是在边界点定义电气特性，例如EIA RS-232-C、V.28；最近的标准则说明了发送器和接收器的电气特性，而且给出了有关对连接电缆的控制。

CCITT 标准化的电气特性标准:

CCITT V.10/X.26：新的非平衡型电气特性，EIA RS-423-A

CCITT V.11/X.27 ：新的平衡型电气特性，EIA RS-422-A

CCITT V.28：非平衡型电气特性，EIA RS-232-C

CCITT X.21/EIA RS-449

物理层的功能特性和规程特性:

功能特性:主要定义各条物理线路的功能。

线路的功能分为四大类： 数据,控制,定时,地(保护地和信号地)

规程特性:主要定义各条物理线路的工作规程和时序关系。

2.2 数据通信的基础知识

• 数据通信的理论基础

• 数据通信系统的模型

• 有关信号的几个基本概念

• 信道的最高码元传输速率

• 信道的极限信息传输速率

数据通信的理论基础:

主要内容：信号在通信信道上传输时的数学表示及其所受到的限制。

傅立叶分析：任何一个周期为T的有理周期性函数g(t) 可分解为若干项

（可能无限多项）正弦和余弦函数之和：

信号在信道上传输时的特性：

1)传输设备对不同傅立叶分量的衰减不同，因此引起输出失真；

2)信道有截止频率fc, 0 ~ fc的振幅不衰减， fc以上的振幅衰减厉害，这主要由 信道的物理特性决定， 0 ~ fc称为信道的有效带宽；

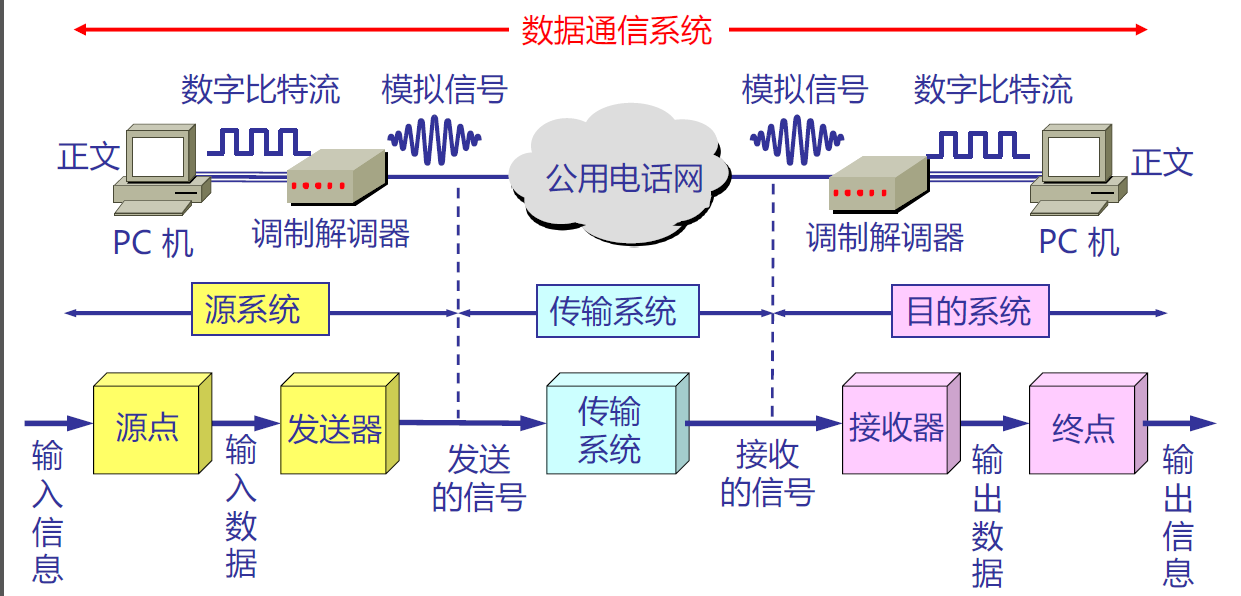
3)实际使用时，可以接入滤波器，限制用户的带宽；

4)通过信道的谐波次数越多，信号越逼真。

例如: 音频线路的截止频率为3000Hz，传输率为b(bps)，发送8比特需要的时间是8/b s，一次谐波的频率是b/8 Hz，最高谐波数：

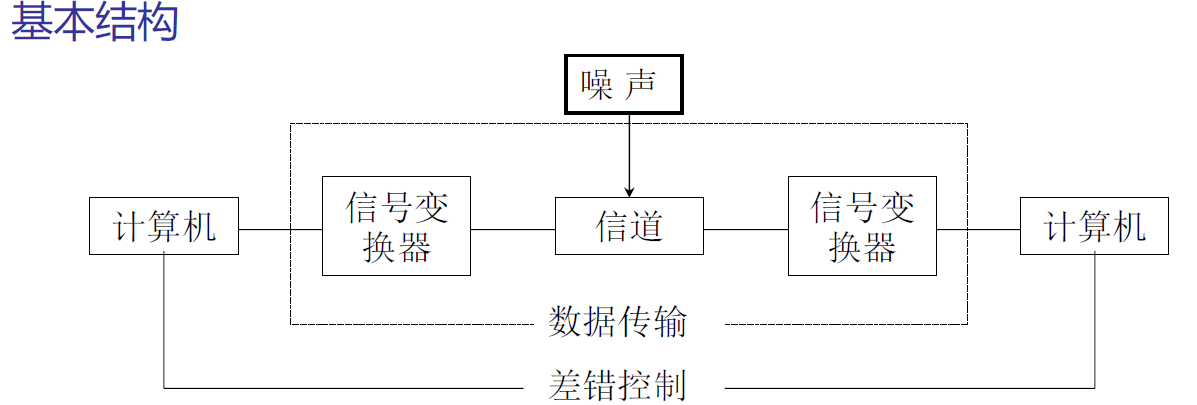
N = fc / f1 = 3000/(b/8) = 24000/b

结论:即使对于完善的信道,有限的带宽限制了数据的传输速率



数据通信系统的基本结构:

主要内容：数据在通信信道上的各种传输方式及其所采用的技术。



信道与电路:

信道和电路并不等同；

电路是用来传输电磁波信号的通路；信道一般都是用来表示向某一个方向传送信息的媒体。因此，一条通信电路至少包含一条发送信道和(或)一条接收信道.一个信道可以看成是一条电路的逻辑部件。

模拟信道和数字信道:

模拟信道：以连续模拟信号形式传递数据的信道

数字信道：以数字脉冲形式传递数据的信道

作为一般的通信系统，信息源产生的信息可能是模拟数据，也可能是数字数据。

如果从通信的发送端所产生的信号形式来看，则信号可以分为以下的两大类：

模拟信号： 即连续的信号，如话音信号和目前的广播电视信号。

数字信号： 即离散的信号,如计算机通信所用的二进制代码1和0组成的信号

数据表示和数据信号:

数据表示:1)模拟数据(Analog Data)—连续值2)数字数据(Digital Data)—离散值

数据信号:

1)模拟信号(Analog Signals)—随时间连续变化的信号，这些信号的某些参量（振幅、相位和频率）可以表示要传送的信息

2)数字信号(Digital Signals)—只取有限个离散值，以某一瞬间的状态表示它们传送的信息。

数据传输方式:

1)模拟信号发送（模拟信道）：数据进入信道前先经过调制，变换为模拟的调制信号。调制信号的频谱较窄，信道的利用率高。模拟信号传输中会衰减，还会受到噪声的干扰，如果用放大器将信号放大，混入的噪声也会放大。

2)数字信号发送（数字信道）：直接传输二进制数据或经过二进制编码的数据。因数字信号只取有限个离散值，在传输过程中即使受到噪声干扰，只要没有畸变到不可辨认的程度，就可以用信号再生的办法进行恢复。传输数字信号所要求的频带要宽的多，其信道利用率低。

几个术语:

1)数据(data)——运送信息的实体。

2)信号(signal)——数据的电气的或电磁的表现。

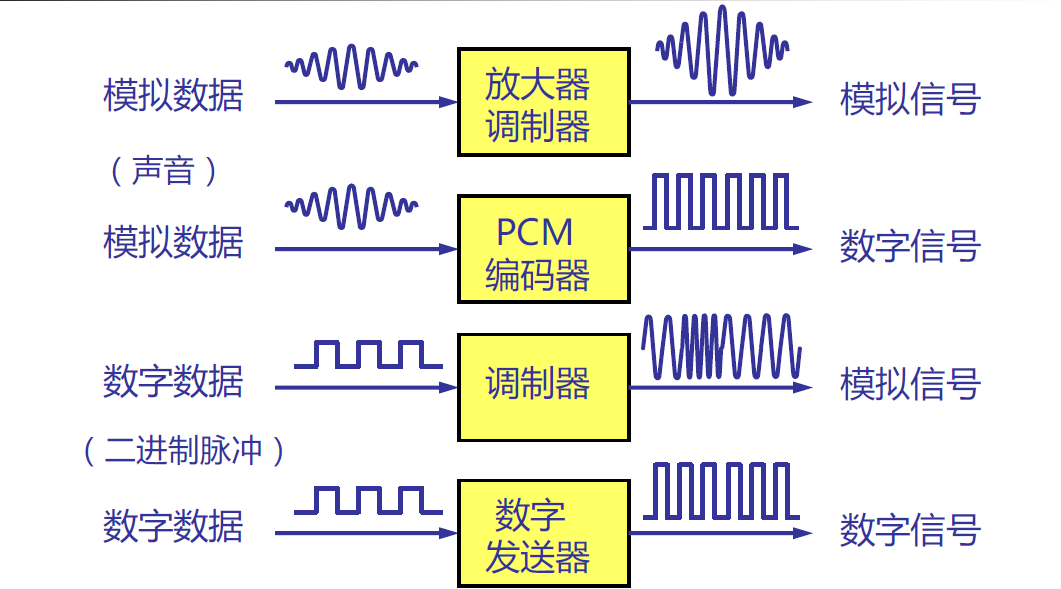
3)“模拟的”(analogous)——连续变化的。

4)“数字的”(digital)——取值是离散数值。

5)调制——把数字信号转换为模拟信号的过程。

6)解调——把模拟信号转换为数字信号的过程

信号发送的4种方式:



数字信号发送的优点是：价格便宜，对噪声不敏感；

缺点是：易受衰减，频率越高，衰减越厉害。

有关信号的几个基本概念:

单向通信（单工通信）——只能有一个方向的通信而没有反方向的交互。

双向交替通信（半双工通信）——通信的双方都可以发送信息，但不能双方同时发送(当然也就不能同时接收)。

双向同时通信（全双工通信）——通信的双方可以同时发送和接收信息。

基带信号和宽带信号:

基带信号(baseband)就是将数字信号1 或0 直接用两种不同的电压来表示，然后送到线路上去传输。

宽带信号(broadband)则是将基带信号进行调制后形成的频分复用模拟信号。

通带信号(passband)被搬移到某个更大频率范围的信号。

信道的最高码元传输速率:

1)任何实际的信道都不是理想的,在传输信号时会产生各种失真以及带来多种干扰。

2)码元传输的速率越高，或信号传输的距离越远,在信道的输出端的波形的失真就越严重

尼氏(Nyquist)准则:

1924年，尼奎斯特（H. Nyquist）推导出理想低通信道的最高码元传输速率= 2H Baud

H 是理想低通信道的带宽，单位为赫(Hz)

1)每赫带宽的理想低通信道的最高码元传输速率是每秒2 个码元。

2)Baud 是波特，是码元传输速率的单位，1 波特为每秒传送1 个码元。

要强调以下三点:

1)实际的信道所能传输的最高码元速率,要明显地低于尼氏准则给出上限数值。

2)当信号电平分为Ｖ级，尼奎斯特（H. Nyquist）无噪声有限带宽信道的最大数据传输率公式表示为：最大数据传输率= 2Hlog2V (bps) ,这说明，任意信号通过一个带宽为Ｈ的低通滤波器，则每秒采样2Ｈ次就能完整地重现该信号。

例如,V=2（只有两级）,无噪声的3k Hz信道传输二进制最高的速率是6000 bps。

要注意:

1)波特(Baud)和比特(bit)是两个完全不同的概念。

波特是码元传输的速率单位（每秒传多少个码元）。码元传输速率也称为调制速率、波形速率或符号速率。

比特是信息量的单位。

2) 信息的传输速率"比特/秒”与码元的传输速率“波特”在数量上却有一定的关系。

3) 若1个码元只携带1 bit 的信息量，则“比特/秒”和“波特”在数值上相等。

若1个码元携带n bit 的信息量，则M Baud的码元传输速率所对应的信息传输速率为M´n b/s。

波特率（baud）和比特率（bit）的关系:

波特率：信号每秒钟变化的次数，也称调制速率。

比特率：每秒钟传送的二进制位数。

波特率与比特率的关系取决于信号值与比特位的关系。

1)例：每个信号值可表示３位，则比特率是波特率的３倍；每个信号值可表示１位，则比特率和波特率相同。

2)例如，有一个带宽为3kHz的理想低通信道，其最高码元传输速率为6000 Baud(即每秒传6000个码元)。若1个码元能携带3bit的信息量，则最高信息传输速率为18000b/s。

信道的最大数据传输速率:

香农(Shannon)用信息论的理论推导出了带宽受限且有高斯白噪声干扰的信道的极限、无差错的信息传输速率。

信道的极限信息传输速率C 可表达为C = H log2(1+S/N) b/s

H 为信道的带宽（以Hz 为单位）；S 为信道内所传信号的平均功率；

N 为信道内部的高斯噪声功率。

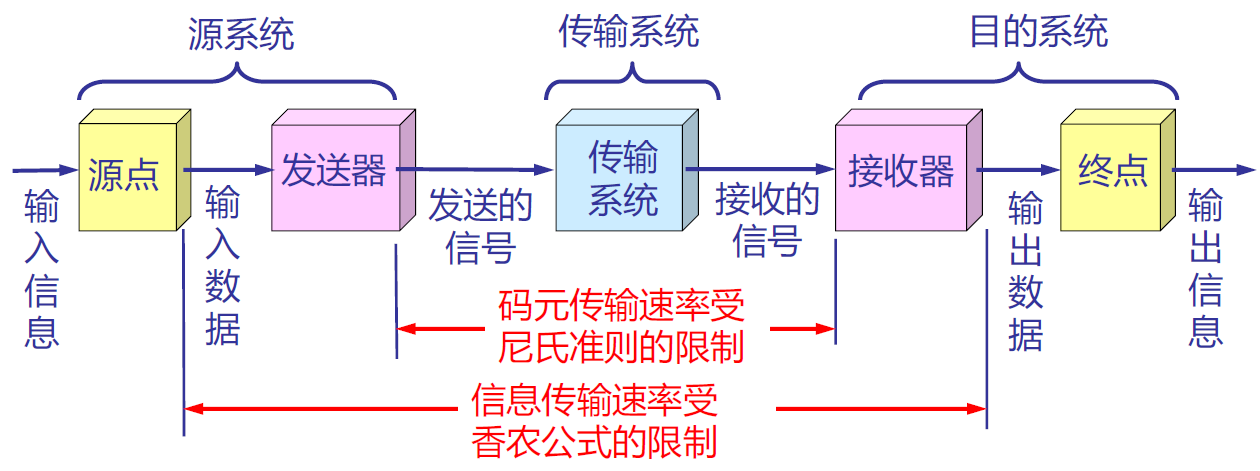
香农公式表明:

信道的带宽或信道中的信噪比越大，则信息的极限传输速率就越高。

只要信息传输速率低于信道的极限信息传输速率，就一定可以找到某种办法来实现无差错的传输。

若信道带宽H 或信噪比S/N 没有上限（当然实际信道不可能是这样的），则信道的极限信息传输速率C 也就没有上限。

实际信道上能够达到的信息传输速率要比香农的极限传输速率低不少。



补充材料：分贝的概念

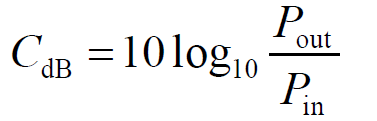
信号强度是任何传输系统中一个重要的参数。当信号沿着传输媒体传播时，其强度会有损耗或衰减。为了补偿这些损耗，可以在不同的地点加入一些放大器，以使信号强度获得一个增益

我们习惯用分贝（decibel）来表示增益、损耗及相对值，原因如下：

1)信号强度通常以指数形式下降，因此，用分贝很容易表示损耗，分贝是一个对数单位。

2)在一个串联的传输通道上，净增益或损耗可以用简单的加减计算。

分贝是对两个信号电平之间的比值的一种度量，分贝增益可用下列公式计算：



如果CdB的值为正的，则表示功率上的一个实际的增益。例如，3dB的增益意味着功率增加一倍。如果CdB的值为负的，则表示功率上的一个实际的损耗。例如，-3dB的增益意味着功率减半，这是功率的损耗。通常，这种情况下就说存在着3dB的损耗。

分贝值指的是相对量值或量值的变化:

例如： 如果在传输线上加入一个功率值为10mW的信号，并且在一定距离之外测得其功率为5mW，那么它的损耗就可以表示为

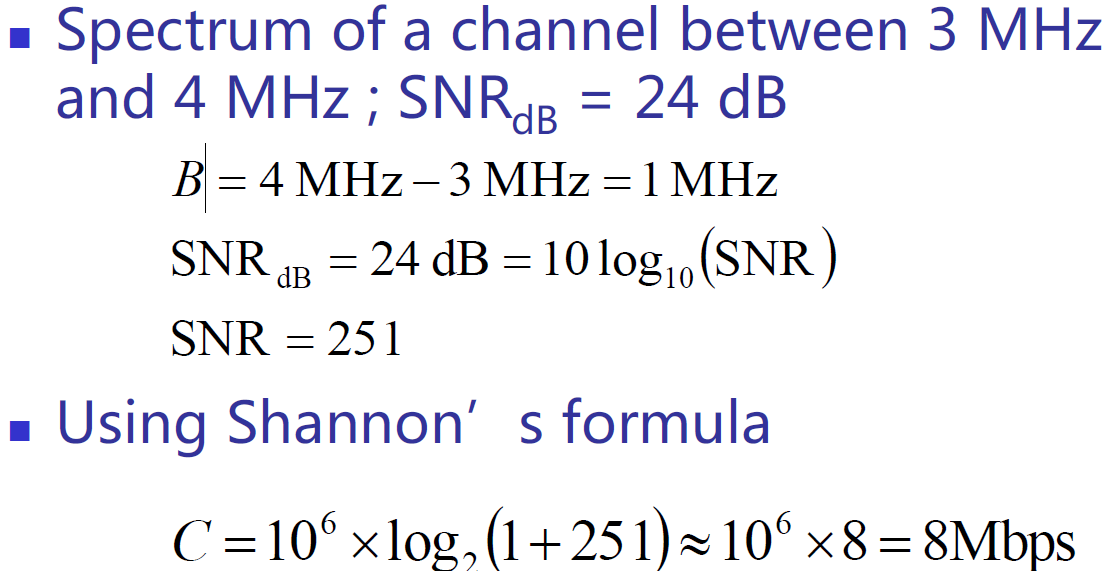
LdB=10lg10/5=10(0.3)=3dB

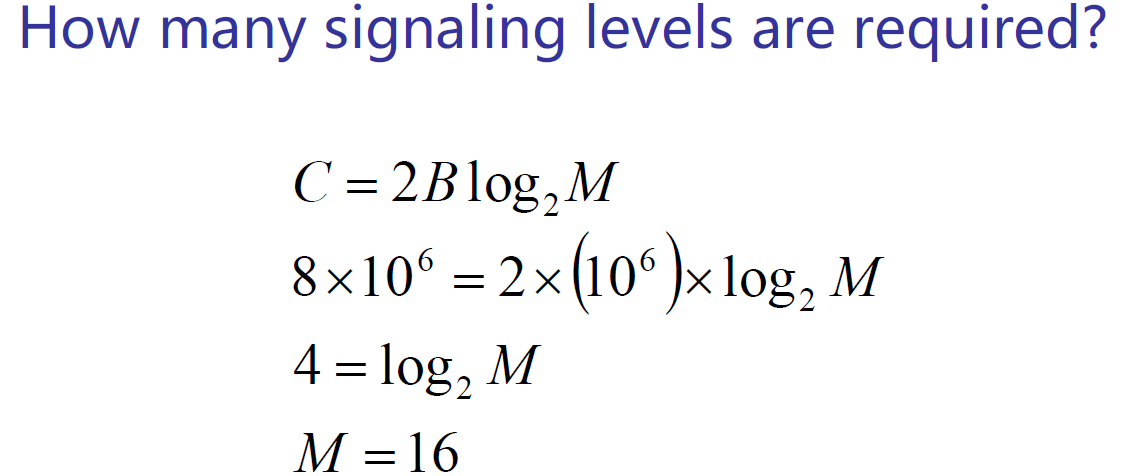
1)注意分贝是一个相对值的度量，而不是绝对差值的度量。从1000mW到500mW的损耗也是3dB。

2) 相对值只说明两个功率（或电压、电流、声强等）的相对关系，无法知道它的绝对值。

尼氏与香农公式举例:

Spectrum of a channel between 3 MHz and 4 MHz ; SNRdB = 24 dB





物理层下面的传输媒体:

1. 导向传输媒体 2)非导向传输媒体

导向传输媒体:

磁介质,双裸线,双绞线:屏蔽双绞线STP (Shielded Twisted Pair):无屏蔽双绞线UTP (Unshielded Twisted Pair),同轴电缆:50W同轴电缆;75W同轴电缆, 光缆

双裸线:

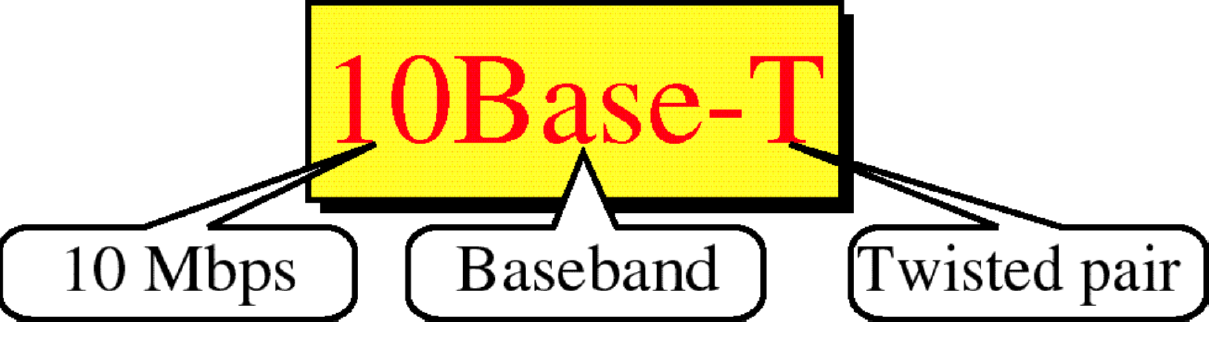
一种最简单的传输介质，也称平行线。

两根裸露在空气中的铜导线彼此隔离地并排放在一起构成图。

串扰严重，速率低，距离短

双绞线(Twisted-pair):

也称双扭线由两根互相绝缘的铜导线用规则的方法扭绞起来构成图; 双绞线的扭绞结构是为了减少相邻导线之间的串扰和消除外界干扰



UTP(Unshielded Twisted-Pair )电缆



同轴电缆(Coaxial Cable):

由单股实心或多股绞合的铜质芯线（内导体）、绝缘层、网状编织的屏蔽层(外导体)以及保护外层所组成

同轴电缆的种类:

①粗同轴电缆

简称粗缆，用于传送基带数字信号. 50欧姆，用于数据传输；

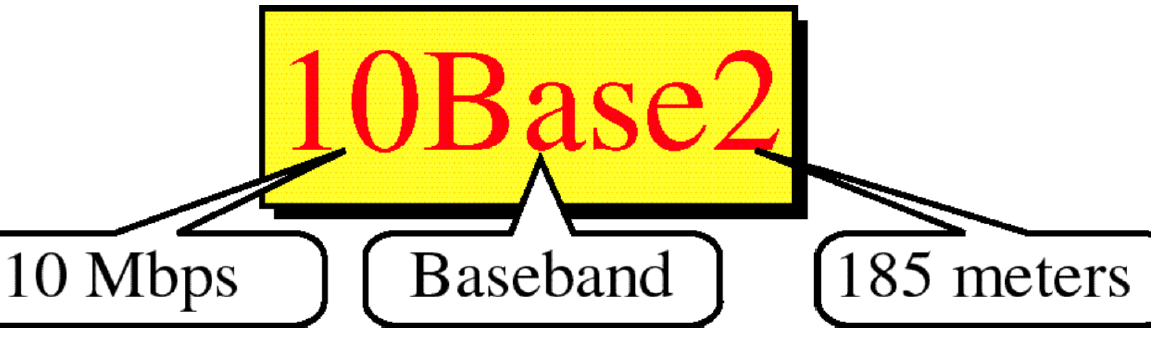
②细同轴电缆

简称细缆，也用于传送基带数字信号。最大传输距离不及粗缆，但因比粗缆便宜而被广泛使用。

③TV电缆

75欧姆，用于模拟传输系统，它是公用天线电视系统(CATV)中的标准传输电缆。其采用的是频分复用的宽带信号，300MHz或450MHz。





双裸线、双绞线、铜轴电缆的优缺点:

• 双裸线构成的线路的串扰现象非常严重。

• 双绞线减少相邻导线间的串扰。主要缺点是存在较强的趋肤效应，随着传输速率的增加，导线中的电流趋向于在导体的外层流动，从而减少可使用的有效截面积，增大导线的电阻和信号的衰减。

• 同轴电缆由于外导体的作用，外来的电磁干扰被有效地屏蔽了，因此具有很好的扰干扰特性，并且因趋肤效应所引起的功率损失也大大减小。同时，与双绞线相比，同轴电缆具有更宽的带宽和更快的传输速率。

传输介质—光纤:

光纤分类：单模光纤和多模光纤

模式（mode）：是一个与很多参数有关的量，可以简单地理解为偏振方向，单模光纤可以传输多种波长，但每个波长只能有一种模式。

常用的三个波长窗口（光纤波段）: Fig. 2-7

0.85um：衰减（attenuation)大，传输速率和距离受限制，但价格便宜；

1.30um：衰减小,无色散（dispersion）补偿、功率放大情况下，最大传40km（最坏情况）；

1.55um：衰减小，无色散补偿、功率放大情况下，最大传80km（最坏情况）

光网络: 组网方式,点到点：四根线（两根用于保护倒换）;环：两根线（一根用于保护倒换）

中继器：光— 电— 光，全光

全光网，光因特网论坛OIF

光纤通信的特点:

传输频带宽，通信容量大; 损耗低，中继距离长;抗干扰能力强;保密性好;适用范围广;重量轻、体积小;节省有色金属，抗腐蚀能力强;

光纤的基本知识:

光 :光是一种电磁能量形式;在真空中传播最快，300000km/s;光速与传输媒质密度成反比

光的自然特性:折射,全反射角,反射（全反射）

**非导向传输媒体**

无线传输所使用的频段很广。

短波通信主要是靠电离层的反射，但短波信道的通信质量较差。

微波在空间主要是直线传播:地面微波接力通信;卫星通信

2.4 模拟传输与数字传输

数据编码技术;模拟传输系统;调制解调器;数字传输系统

数据编码:

数字数据在数字信道上传输时需要数字信号编码

数字数据在模拟信道上传输时需要调制编码

模拟数据在数字信道上传输时需要采样编码

用模拟信号传输模拟数据不需要编码

常用的几种编码方式:

1）不归零制码（NRZ：Non-Return to Zero）

原理：用两种不同的电平分别表示二进制信息“0”和“1”，低电平表示“0”,高电平表示“1”。

缺点：a. 难以分辨一位的结束和另一位的开始b. 发送方和接收方必须有时钟同步；c. 若信号中“0”或“1”连续出现，信号直流分量将累加。

结论：容易产生传播错误。

2）曼彻斯特码（Manchester），也称相位编码

原理：每一位中间都有一个跳变，从低跳到高表示“0”，从高跳到低表示“1”。

优点：克服了NRZ码的不足。每位中间的跳变即可作为数据，又可作为时钟，能够自同步。

3）差分曼彻斯特码（Differential Manchester）

原理：每一位中间都有一个跳变，每位开始时有跳变表示“0”，无跳变表示“1”。位中间跳变表示时钟，位前跳变表示数据。

优点：时钟、数据分离，便于提取，抗干扰性好。需复杂技术。

4）逢“1”变化的NRZ码

原理：在每位开始时,逢“1”电平跳变,逢“0”电平不跳变。

5）逢“0”变化的NRZ码

原理：在每位开始时,逢“0”电平跳变,逢“1”电平不跳变。

数字数据的模拟传输（频带传输）: 采用调制编码

频带传输：指在一定频率范围内的线路上，进行载波传输。用基带信号对载波进行调制，使其变为适合于线路传送的信号。

调制（Modulation）：用基带脉冲对载波信号的某些参量进行控制，使这些参量随基带脉冲变化。

解调（Demodulation）：调制的反变换。

调制解调器MODEM（modulationdemodulation)

模拟数据数字传输——采样编码

解决模拟信号数字化问题，现在的数字传输系统都是采用脉码调制PCM(Pulse Code Modulation)体制，也称为脉冲代码调制PCM。

步骤是：采样-> 量化-> 编码

1. 采样：根据Nyquist原理进行，每隔一定的时间对连续模拟信号采样。

2. 量化：分级过程。把采样所得到的脉冲信号按量级比较，并且取整，这样脉冲序列就成为数字信号了。

3. 编码：用以表示采样序列量化后的量化幅度，用一定位数的二进制码表示。调制解调器的作用:

调制解调器(modem)包括：

调制器(MOdulator)：把要发送的数字信号转换为频率范围在300~3400 Hz 之间的模拟信号，以便在电话用户线上传送。

解调器(DEModulator)：把电话用户线上传送来的模拟信号转换为数字信号。

本书中的调制解调器是指使用在标准的二线模拟话路（3.1 kHz 的标准话路带宽）上的调制解调器。

调制器的主要作用就是个波形变换器，它把基带数字信号的波形变换成适合于模拟信道传输的波形

解调器的作用就是个波形识别器，它将经过调制器变换过的模拟信号恢复成原来的数字信号。

1)若识别不正确，则产生误码。2)在调制解调器中还要有差错检测和纠正的设施。

几种最基本的调制方法:

1）调制就是进行波形变换（频谱变换）

最基本的二元制调制方法有以下几种：

(1)调幅(AM) 即载波的振幅随基带数字信号而变化。

例如，0对应于无载波输出，而1对应于有载波输出。

(2)调频(FM) 即载波的频率随基带数字信号而变化。

例如，0对应于频率f1，而1对应于频率f2。

(3)调相(PM) 即载波的初始相位随基带数字信号而变

化。例如，0对应于相位0o，而1对应于180o。

常用的三种调制技术：根据载波Asin(wt + j)的三个特性：幅度、频率、相位，产生常用的三种调制技术： 幅移键控法Amplitude-shift keying (ASK)；频移键控法Frequency-shift keying (FSK)；相移键控法Phase-shift keying (PSK)；

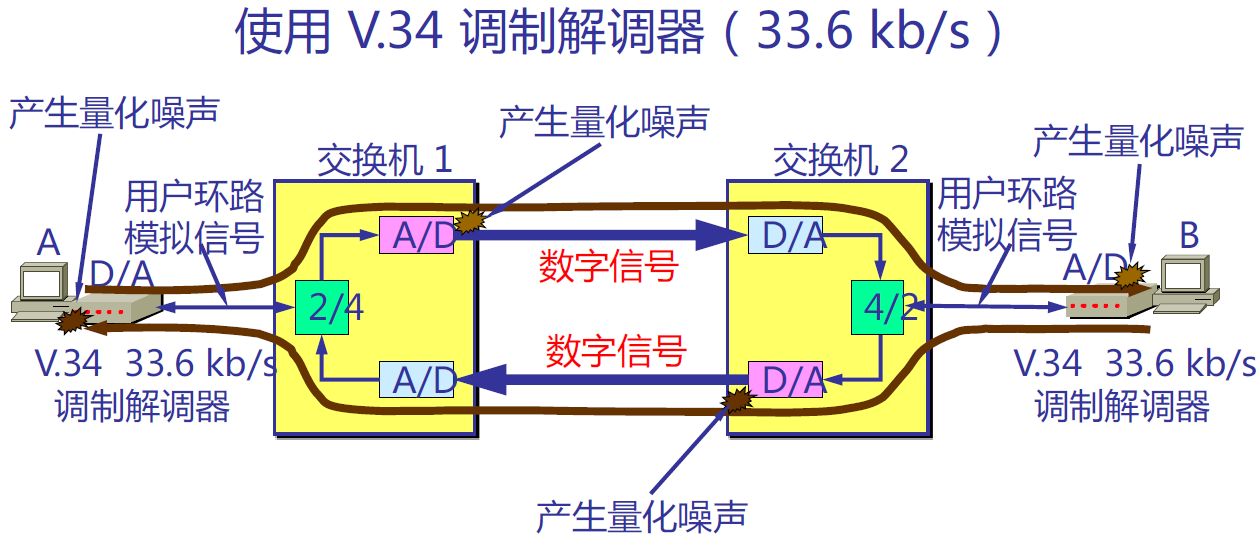
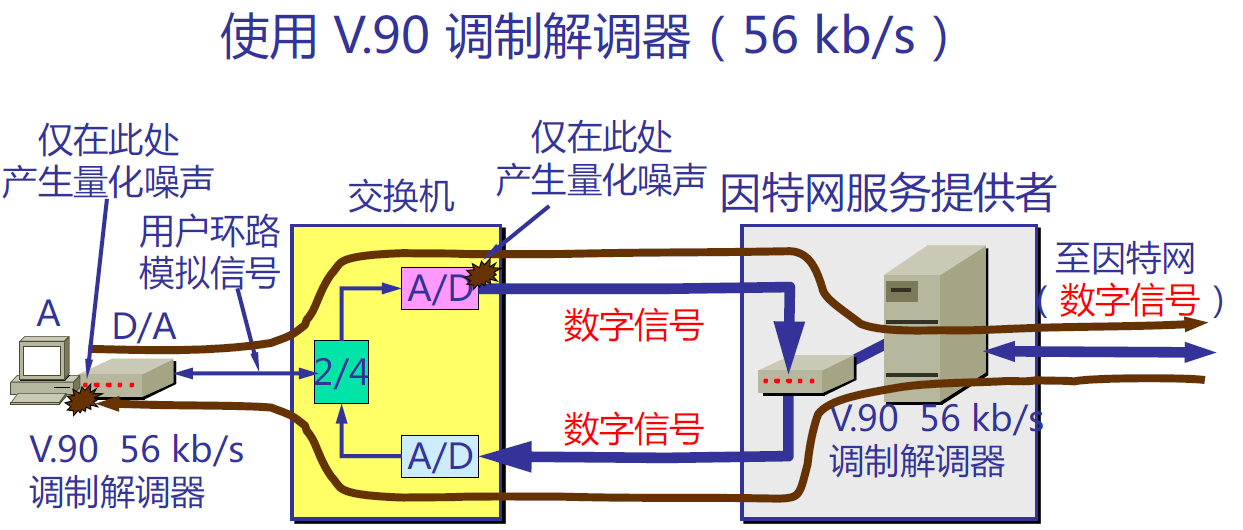
调制解调器的速率：

目前调制解调器的信息传输速率已很接近于香农的信道容量极限了。

要提高信息传输速率，只能设法提高信噪比。

在电话的用户线上，最大的噪声来自模拟到数字的模数转换所带来的量化噪声。

产生量化噪声的地方（经过A/D 变化）：



调制解调器使用异步通信方式：

数据通信可分为同步通信和异步通信两大类：

同步通信要求接收端时钟频率和发送端时钟频率一致。发送端发送连续的比特流。

异步通信时不要求接收端时钟和发送端时钟同步。

发送端发送完一个字节后，可经过任意长的时间间隔再发送下一个字节。

异步通信的通信开销较大，但接收端可使用廉价的、具有一般精度的时钟来进行数据通信。

数字传输系统：

现在的数字传输系统均采用脉码调制PCM (Pulse Code Modulation)体制。

PCM最初并不是为传送计算机数据用的。它是为了使电话局之间一条中继线不是只传送一路电话而是可以传送几十路的电话。由于历史上的原因，PCM有两个互不兼容的国际标准，即北美的24路PCM(简称为T1)和欧洲的30路PCM(简称为E1)。n T1的速率是1.544 Mb/s，E1的速率是2.048Mb/s。

脉码调制PCM：

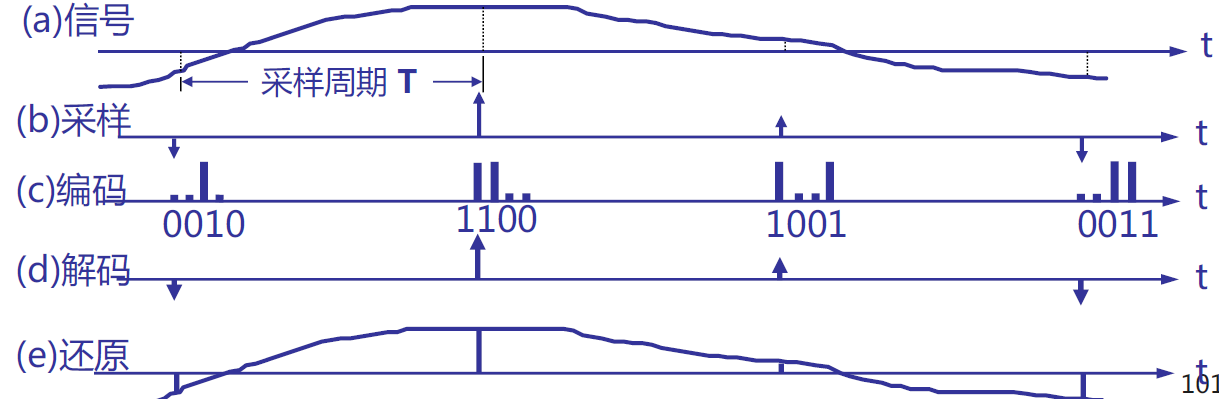
为了将模拟电话信号转变为数字信号，必须先对电话信号进行取样。根据取样定理，只要：取样频率不低于电话信号最高频率的2倍，就可以从取样脉冲信号无失真地恢复出原来的电话信号。标准的电话信号的最高频率为3.4kHz，为方便起见，取样频率就定为8kHz，相当于取样周期为125 μs。

PCM的基本原理：

T为取样周期。连续的电话信号经取样后成为如图中(a)所示的离散脉冲信号，其振幅对应于取样时刻电话信号的数值。

下一步就是编码。为简单起见,图(c)将不同振幅的脉冲编为4位二进制码元。在我国使用的PCM体制中，电话信号是采用8 bit编码,也就是说,将取样后的模拟的电话信号量化为256个不同等级中的一个。模拟信号转换为数字信号后就进行传输。

在接收端进行解码的过程与编码过程相反。经滤波后最后得出恢复后的模拟电话信号图中的(e)。



这样，一个话路的模拟电话信号，经模数变换后，就变成为每秒8000个脉冲信号，每个脉冲信号再编为8位二进制码元。因此一个标准话路的PCM信号速率为64 kb/s。以kb/s的速率是最早制订出的话音编码的标准速率。随着话音编码技术的不断发展，人们可以用更低的数据率来传送同样质量的话音信号。现在已经能够用32kb/s，16kb/s或甚至低到8kb/s(或更低)的数据率来传送一路话音信号。

时分复用：为了有效地利用传输线路，可将多个话路的PCM 信号用时分复用TDM (Time Division Multiplexing)的方法装成时分复用帧，然后发送到线路上。

中国采用欧洲体制，以E1 为一次群。

美国和日本等国采用北美体制，以T1 为一次群。

2.5 信道复用技术

频分复用、时分复用和统计时分复用

波分复用WDM

码分复用CDM

频分复用：所有用户在同样的时间占用不同的带宽资源。

时分复用：所有用户在不同的时间占用同样的频带宽度。

时分复用可能会造成线路资源的浪费

统计时分复用STDM：工作原理-动态地按照需要来分配时隙的多路复用方式：

. 一边是一组I/O线路,另一边是一条高速复用线路,每条I/O线路有一缓冲区;

• 对于n条I/O线路在TDM帧中只有k个时间槽可用;

• 输入时多路复用器的功能是扫描输入缓冲区收集数据直到一帧为满，就发送出去;

• 输出时，多路复用器接收一帧，将各时间槽的数据分发到相应的输出缓冲区。

波分复用WDM: 波分复用就是光的频分复用。

码分复用CDM: 常用的名词是码分多址CDMA(Code Division Multiple Access)。

各用户使用经过特殊挑选的不同码型，因此彼此不会造成干扰。这种系统发送的信号有很强的抗干扰能力，其频谱类

似于白噪声，不易被敌方发现。

每一个比特时间划分为m 个短的间隔，称为码片(chip)。

本章小结:

**通信基础**

信道、信号、宽带、码元、波特、速率等基本概念

奈奎斯特定理与香农定理

信源与信宿

编码与调制

电路交换、报文交换与分组交换

数据报与虚电路

**传输介质**

双绞线、同轴电缆、光纤与无线传输介质

物理层接口的特性

**信道复用技术**: 频分多路复用、时分多路复用、波分多路复用、码分多路复用的概念和基本原理。

**物理层设备** 中继器 **;**集线器

**Chapter 3 Data Link Layer**

**主要内容:**

3.1 定义和功能

1. 定义

2.为网络层提供服务

3.成帧

4.差错控制

5.流量控制

3.2 错误检测和纠正

1. 纠错码

2.检错码

3.3 基本的数据链路层协议

1. 无约束单工协议

2. 单工停等协议

3. 有噪声信道的单工协议

3.4 滑动窗口协议

1. 一比特滑动窗口协议

2. 退后n帧协议

3. 选择重传协议

3.5常用的数据链路层协议

3.1 定义和功能（1）

要解决的问题: 如何在有差错的线路上，进行无差错传输

ISO关于数据链路层的定义:数据链路层的目的是为了提供功能上和规程上的方法，以便建立、维护和释放网络实体间的数据链路。

基本概念:

结点（node）：网络中的主机（host）和路由器（router）称为结点

链路（link）：通信路径上连接相邻结点的通信信道称为链路

数据链路层协议定义了一条链路的两个结点间交换的数据单元格式，以及结点发送和接收数据单元的动作。

端到端（end to end）：从源结点（source node）到目的结点（destination node）的通信称为端到端通信，通信路径（path）可能由多个链路组成。

点到点（point to point）：在相邻结点间的一条链路上的通信称为点到点通信。

虚拟数据通路

实际数据通路

数据链路控制规程: 为使数据能迅速、正确、有效地从发送点到达接收点所采用的控制方式。

数据链路层协议应提供的最基本功能: 数据在数据链路上的正常传输(建立、维护和释放);定界与同步，也处理透明性问题;差错控制;顺序控制;流量控制

数据链路层基本概念说明:

链路(link)是一条无源的点到点的物理线路段，中间没有任何其他的交换结点。

一条链路只是一条通路的一个组成部分。

数据链路(data link) 除了物理线路外，还必须有通信协议来控制这些数据的传输。若把实现这些协议的硬件和软件加到链路上，就构成了数据链路。

现在最常用的方法是使用适配器（即网卡）来实现这些协议的硬件和软件。

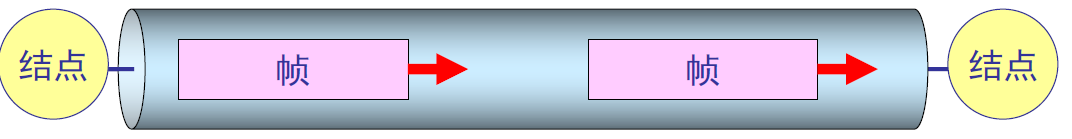
一般的适配器都包括了数据链路层和物理层这两层的功能。

数据链路层像个数字管道:

常常在两个对等的数据链路层之间画出一个数字管道，而在这条数字管道上传输的数据单位是帧

早期的数据通信协议曾叫作通信规程(procedure)。因此在数据链路层，规程

和协议是同义语。



数据链路层的主要功能:

1. 链路管理(2) 帧定界(3) 流量控制(4) 差错控制(5) 将数据和控制信息区分

(6) 透明传输(7) 寻址

2. 设计目标：为网络层提供服务 ,

为网络层提供三种合理的服务1) 无确认无连接服务，适用于误码率很低的线路，错误恢复留给高层；实时业务;大部分局域网;2) 有确认无连接服务，适用于不可靠的信道，如无线网3) 有确认有连接服务

3. 成帧（Framing）:将比特流分成离散的帧，并计算每个帧的校验和。

成帧方法：1)字符计数法2)带字符填充的首尾字符定界法3)带位填充的首尾标记定界法4)物理层编码违例法

注意：在很多数据链路协议中，使用字符计数法和一种其它方法的组合。

字符计数法: 在帧头中用一个域来表示整个帧的字符个数, 缺点：若计数出错，对本帧和后面的帧有影响;

带字符填充的首尾字符定界法: 起始字符DLE STX，结束字符DLE ETX, DLE:Data Link Escape,STX:Start of Text, ETX:End of Text;

字符填充：数据中出现标志字符时插入转义字符

缺点：局限于8位字符和ASCII字符传送。

带位填充的首尾标记定界法: 帧的起始和结束都用一个特殊的位串“01111110”，称为标记(flag);“0”比特插入删除技术

物理层编码违例法: 只适用于物理层编码有冗余的网络802 LAN：曼彻斯特编码或差分曼彻斯特编码用high-low pair/low-high pair表示1/0，highhigh/low-low不表示数据，可以用来做定界符。

4. 差错控制: 一般方法：接收方给发送方一个反馈（响应）。

出错情况1)帧（包括发送帧和响应帧）出错；2)帧（包括发送帧和响应帧）丢失

通过计时器和序号保证每帧最终交给目的网络层仅一次是数据链路层的一个主要功能。

5. 流量控制: 1)基于反馈机制2) 流量控制主要在传输层实现

3.2 错误检测和纠正（1）

差错出现的特点：随机，连续突发（burst）

处理差错的两种基本策略1)使用纠错码：发送方在每个数据块中加入足够的冗

余信息，使得接收方能够判断接收到的数据是否有错，并能纠正错误。

2)使用检错码：发送方在每个数据块中加入足够的冗余信息，使得接收方能够判断接收到的数据是否有错，但不能判断哪里有错。

3.2 错误检测和纠正（2）

1. 纠错码: 码字（codeword）：一个帧包括m个数据位,r个校验位，n = m + r，则此n比特单元称为n位码字。

海明距离（Hamming distance）：两个码字之间不同的比特(对应)位数目。

例：0000000000 与0000011111的海明距离为5

如果两个码字的海明距离为d，则需要d个单比特错就可以把一个码字转换成另一个码字；

对于n位码字的集合，只有2m个码字是有效的。也就是说，通常并未使用所有2n个码字。

检错和纠错: 一种编码的检错和纠错能力取决于编码后码字海明距离的大小。

为了检测出d个比特的错，需要使用海明距离为d+1的编码;例如：数据后加奇偶校验位，编码后的海明距离为2，能检测1比特错。

在任意两个有效码字间找出具有最小海明距离的两个码字，该海明距离便定义为全部码字的海明距离。

为了纠正d个比特的错，必须用距离为2d+1的编码。

3.2 错误检测和纠正（3）

最简单的例子是奇偶校验，在数据后填加一个奇偶位

要求设计仅纠正单比特错的纠错码

要求：m个信息位，r个校验位，当r满足什么条件时，能纠正所有单比特错；

3.2 错误检测和纠正（4）

海明码:码位从左边开始编号，从“1”开始；

位号为2的幂的位是校验位，其余是信息位；

每个校验位使得包括自己在内的一些位的奇偶值为偶数（或奇数）。

为看清数据位k对哪些校验位有影响，将k写成2的幂的和.例：11 = 1 + 2 + 8

3.2 错误检测和纠正（5）

海明码工作过程: 每个码字到来前，接收方计数器清零；

接收方检查每个校验位k (k = 1, 2, 4 …)的奇偶值是否正确；

若第k 位奇偶值不对，计数器加k；

所有校验位检查完后，若计数器值为0，则码字有效；若计数器值为m，则第m位出错。例：若校验位1、2、8出错，则第11位变反。

3.2 错误检测和纠正（6）

使用海明码纠正突发错误: 1）可采用k个码字（n = m + r）组成k ′ n 矩阵，按

列发送，接收方恢复成k ′ n 矩阵 2)kr个校验位，km个数据位，可纠正最多为k个的突发性连续比特错

纠错码与检错码：

在实际通信中使用纠错码好还是检错码好呢？

例如：假设一个信道误码率是10-6，且出错是孤立产生的（即只有单比特错），数据块长度为1000比特，如果采用纠错编码，需要10个校验位（210>1011），

传送1M数据需要10000个校验位；如果采用检错编码，每个数据块只需一个奇偶校验位，传送1M数据只需1000个校验位和一个重传的数据1001位，共需要2001比特的额外开销。

在多数通信中采用检错编码，但在单工信道中需要纠错编码。

3.2 错误检测和纠正（7）

2. 检错码

使用纠错码传数据，效率低，适用于不可能重传的场合；大多数情况采用检错码加重传。

循环冗余码（CRC码，多项式编码）：110001，表示成多项式x5 + x4 + 1

生成多项式G(x)：发方、收方事前商定；生成多项式的高位和低位必须为1；生成多项式必须比传输信息对应的多项式短。

3.2 错误检测和纠正（8）

CRC码基本思想：校验和（checksum）加在帧尾，使带校验和的帧的多项式能被G(x)除尽；收方接收时，用G(x)去除它，若有余数，则传输出错。

校验和计算算法：1）设G(x)为r 阶，在帧的末尾加r 个0，使帧为m+r位，相应多项式为xrM(x)；

2）按模2除法用对应于G(x)的位串去除对应于xrM(x)的位串；

3）按模2减法从对应于xrM(x)的位串中减去余数(等于或小于r位)，结果就是要传送的带校验和的多项式T(x)。

3.2 错误检测和纠正（9）

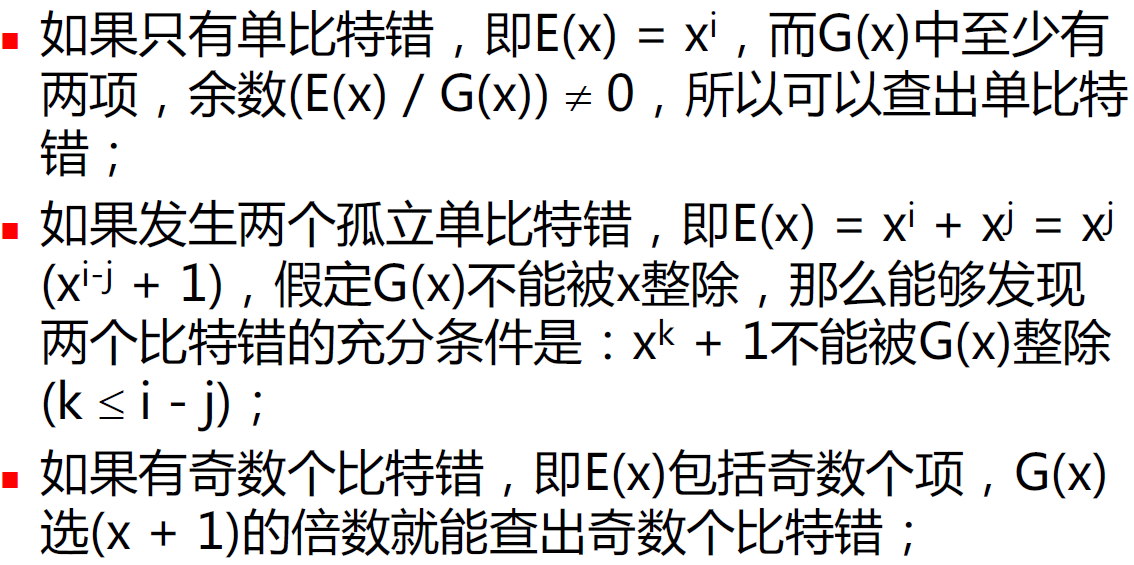
CRC的检错能力：发送：T(x)；接收：T(x) + E(x), E(x) 1 0；

余数((T(x) + E(x)) / G(x)) = 0 + 余数(E(x) / G(x))

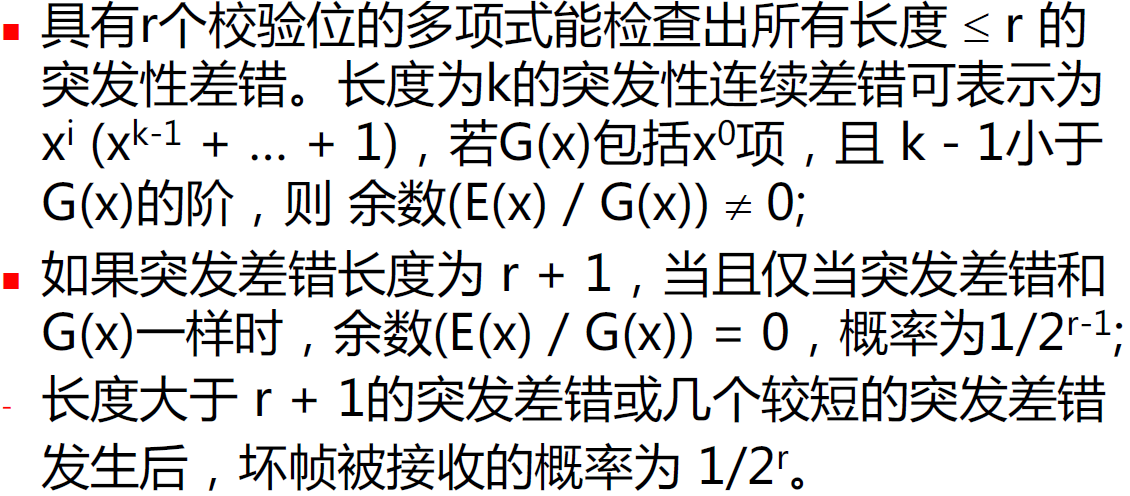
若余数(E(x) / G(x)) = 0，则差错不能发现；否则，可以发现。

3.2 错误检测和纠正（10）

CRC检错能力的几种情况分析：



3.2 错误检测和纠正（11）



帧检验序列FCS：

在数据后面添加上的冗余码称为帧检验序列FCS (Frame Check Sequence)。

循环冗余检验CRC 和帧检验序列FCS并不等同。

CRC 是一种常用的检错方法，而FCS 是添加在数据后面的冗余码。

FCS 可以用CRC 这种方法得出，但CRC 并非用来获得FCS 的惟一方法。

检测出差错：

只要得出的余数R 不为0，就表示检测到了差错。

但这种检测方法并不能确定究竟是哪一个或哪几个比特出现了差错。

一旦检测出差错，就丢弃这个出现差错的帧。

只要经过严格的挑选，并使用位数足够多的除数P，那么出现检测不到的差错的概率就很小很小。

应当注意：

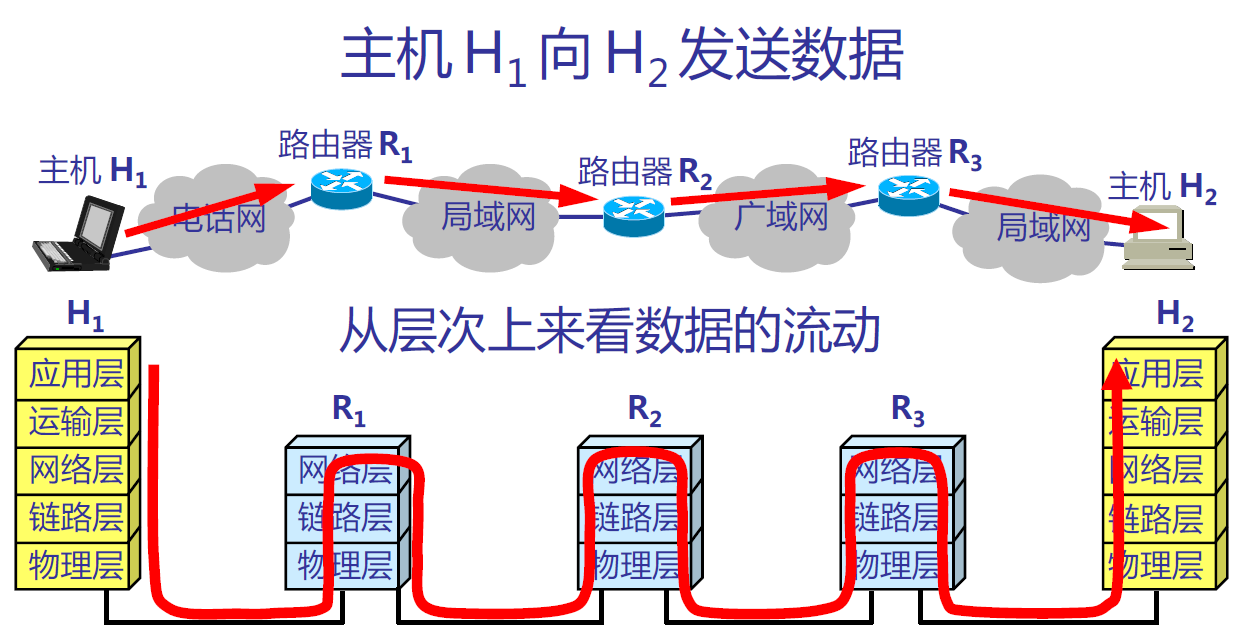
仅用循环冗余检验CRC 差错检测技术只能做到无差错接受(accept)。

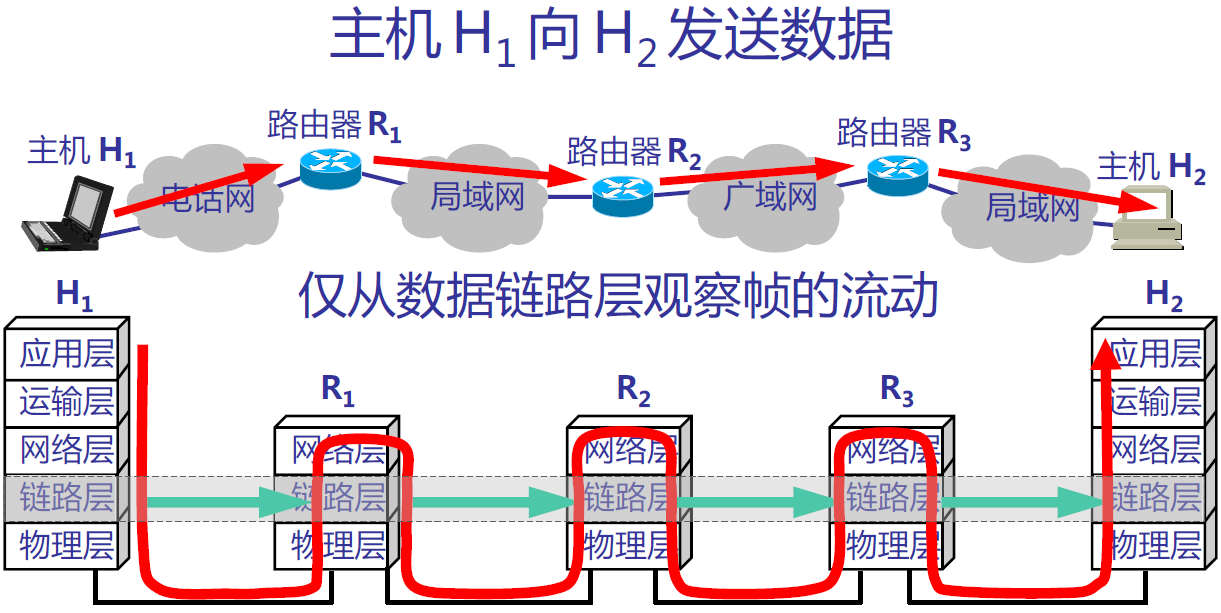
“无差错接受”是指：“凡是接受的帧（即不包括丢弃的帧），我们都能以非常接近于1 的概率认为这些帧在传输过程中没有产生差错”。

也就是说：“凡是接受的帧都没有传输差错”有差错的帧就丢弃而不接受）。

要做到“可靠传输”（即发送什么就收到什么）就必须再加上确认和重传机制。

数据链路层的简单模型：





3.3 基本的数据链路层协议

1. 无约束单工协议(An Unrestricted Simplex Protocol)

工作在理想情况，几个前提：单工传输； 发送方无休止工作（要发送的信息无限多）；接收方无休止工作（缓冲区无限大）；通信线路（信道）不损坏或丢失信息帧

工作过程：发送程序: 取数据，构成帧，发送帧；接收程序：等待，接收帧，送数据给高层

无约束单工协议是一种完全理想化的数据传输

完全理想化的数据传输:基于两个假定：假定1： 链路是理想的传输信道，所传送的任何数据既不会出差错也不会丢失。假定2： 不管发方以多快的速率发送数据，收方总是来得及收下，并及时上交主机。 这个假定就相当于认为：接收端向主机交付数据的速率永远不会低于发送端发送数据的速率。

具有最简单流量控制的数据链路层协议

现在去掉上述的第二个假定。但是，仍然保留第一个假定，即主机A 向主机B传输数据的信道仍然是无差错的理想信道。然而现在不能保证接收端向主机交付数据的速率永远不低于发送端发送数据的速率。

由收方控制发方的数据流，乃是计算机网络中流量控制的一个基本方法。

2. 单工停等协议(A Simplex Stop-and-Wait Protocol)

增加约束条件：接收方不能无休止接收。

解决办法：接收方每收到一个帧后，给发送方回送一个响应。

工作过程： 发送程序：取数据，成帧，发送帧，等待响应帧；接收程序：等待，接收帧，送数据给高层，回送响应帧。

具有最简单流量控制的数据链路层协议算法：

在发送结点：(1) 从主机取一个数据帧。

(2) 将数据帧送到数据链路层的发送缓存。

(3) 将发送缓存中的数据帧发送出去。

(4) 等待。

(5) 若收到由接收结点发过来的信息(此信息的格式与内容可由双方事先商定好)，则从主机取一个新的数据帧，然后转到(2)。

在接收结点：(1) 等待。

(2) 若收到由发送结点发过来的数据帧，则将其放入数据链路层的接收缓存。

(3) 将接收缓存中的数据帧上交主机。

(4) 向发送结点发一信息，表示数据帧已经上交给主机。

(5) 转到(1)。

3. 有噪声信道的单工协议(A Simplex Protocol for a Noisy Channel)

增加约束条件：信道（线路）有差错，信息帧可能损坏或丢失。

解决办法：出错重传。

带来的问题： 什么时候重传—— 定时；响应帧损坏怎么办(重复帧)——发送帧头中放入序号； 为了使帧头精简，序号取多少位—— 1位

4.发方在发下一个帧之前等待一个肯定确认的协议叫做PAR（Positive Acknowledgement with Retransmission）或ARQ（Automatic Repeat reQuest）

注意协议3的漏洞：由于确认帧中没有序号，超时时间不能太短，否则协议失败。因此假设协议3的发送和接收严格交替进行。

协议3亦称为实用的停止等待协议

超时计时器的作用: 结点A发送完一个数据帧时，就启动一个超时计时器(timeout timer)。

计时器又称为定时器。

若到了超时计时器所设置的重传时间tout而仍收不到结点B 的任何确认帧，则结点A 就重传前面所发送的这一数据帧。

一般可将重传时间选为略大于“从发完数据帧到收到确认帧所需的平均时间”。

解决重复帧的问题

使每一个数据帧带上不同的发送序号。每发送一个新的数据帧就把它的发送序号加1。

若结点B 收到发送序号相同的数据帧，就表明出现了重复帧。这时应丢弃重复帧，因为已经收到过同样的数据帧并且也交给了主机B。

但此时结点B 还必须向A 发送确认帧ACK，因为B 已经知道A 还没有收到上一次发过去的确认帧ACK。

帧的编号问题

任何一个编号系统的序号所占用的比特数一定是有限的。因此，经过一段时间后，发送序号就会重复。

序号占用的比特数越少，数据传输的额外开销就越小。

对于停止等待协议，由于每发送一个数据帧就停止等待，因此用一个比特来编号就够了。

一个比特可表示0 和1 两种不同的序号。

帧的发送序号:

数据帧中的发送序号seq以0 和1 交替的方式出现在数据帧中。

每发一个新的数据帧，发送序号就和上次发送的不一样。用这样的方法就可以使收方能够区分开新的数据帧和重传的数据帧了。

可靠传输: 虽然物理层在传输比特时会出现差错，但由于数据链路层的停止等待协议采用了有效的检错重传机制，数据链路层对上面的网络层就可以提供可靠传输的服务

停止等待协议的算法: 通常不使用否认帧（实用的数据链路层协议大都是这样的），而且确认帧带有序号n。

按照习惯的表示法，ACKn 表示“第n – 1 号帧已经收到，现在期望接收第n 号帧”。

ACK1 表示“0 号帧已收到，现在期望接收的下一帧是1 号帧”；

ACK0 表示“1 号帧已收到，现在期望接收的下一帧是0 号帧”。

在发送结点: next\_frame\_to\_send : 发送状态变量;

seq: 发送帧的序号, ack: 确认帧序号

(1) 从主机取一个数据帧，送交发送缓存。

(2) next\_frame\_to\_send←0。

(3) seq← next\_frame\_to\_send 。

(4) 将发送缓存中的数据帧发送出去。

(5) 设置超时计时器。

(6) 等待。{等待以下(7)和(8)这两个事件中最先出现的一个}

(7) 收到确认帧ACK，

若ack = next\_frame\_to\_send，则：从主机取一个新的数据帧，放入发送缓存；

inc(next\_frame\_to\_send)，转到(3)。否则，丢弃这个确认帧，转到(4)。

(8) 若超时计时器时间到，则转到(4)。

在接收结点: Frame\_expected:接收状态变量;seq:接收帧的序号

(1) frame\_expected←0。

(2) 等待。

(3) 收到一个数据帧，用CRC检验有无错误；若seq = frame\_expected ，则执行(4)；否则(传输有错误或重复帧)丢弃此数据帧,然后转到(6)。

(4) 将收到的数据帧中的数据部分送交上层软件也就是数据链路层模型中的主机）。

(5) inc(frame\_expected)。

(6) ack ← 1 - frame\_expected ；发送确认帧ACK，转到(2)。

停止等待协议的要点:

只有收到序号正确的确认帧ACK后，才更新发送状态变量next\_frame\_to\_send一次，并发送新的数据帧。

接收端接收到数据帧时，就要将接收序号seq与本地的接收状态变量frame\_expected相比较。

若二者相等就表明是新的数据帧，就收下，并发送确认。

否则为重复帧，就必须丢弃。但这时仍须向发送端发送确认帧ACK ， 而接收状态变量frame\_expected和确认序号ack 都不变。

连续出现相同发送序号的数据帧，表明发送端进行了超时重传。连续出现相同序号的确认帧，表明接收端收到了重复帧。

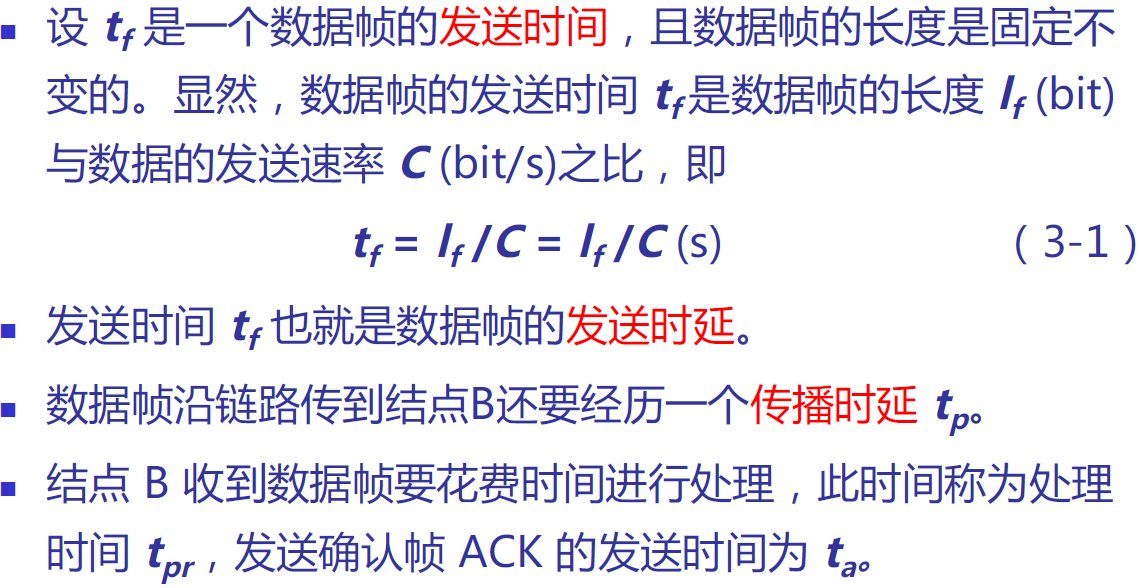
发送端在发送完数据帧时，必须在其发送缓存中暂时保留这个数据帧的副本。这样才能在出差错时进行重传。只有确认对方已经收到这个数据帧时，才可以清除这个副本。

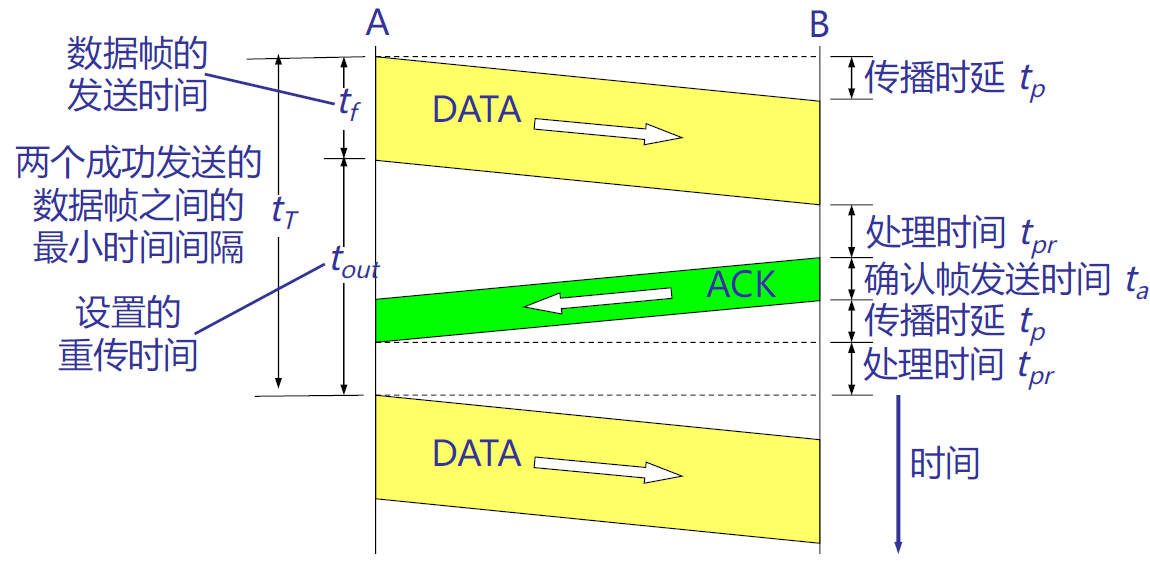
实用的CRC 检验器都是用硬件完成的。

CRC 检验器能够自动丢弃检测到的出错帧。因此所谓的“丢弃出错帧”，对上层软件或用户来说都是感觉不到的。

发送端对出错的数据帧进行重传是自动进行的，因而这种差错控制体制常简称为ARQ(Automatic Repeat reQuest)，直译是自动重传请求，但意思是自动请求重传。

停止等待协议的定量分析





重传时间:

重传时间的作用是：数据帧发送完毕后若经过了这样长的时间还没有收到确认帧，就重传这个数据帧。

为方便起见，我们设重传时间为**tout = tp + tpr+ ta + tp + tpr**

设上式右端的处理时间**tpr** 和确认帧的发送时间t**a**都远小于传播时延**tp**，因此可将重传时间取为两倍的传播时延，即**tout = 2tp**

**简单的数学分析:**

两个发送成功的数据帧之间的最小时间间隔是tT = tf + tout = tf + 2tp

设数据帧出现差错(包括帧丢失)的概率为p，但假设确认帧不会出现差错。

设正确传送一个数据帧所需的平均时间tav:tav = tT ( 1 + 一个帧的平均重传次数)

停止等待协议ARQ 的优缺点:

优点：比较简单。

缺点：通信信道的利用率不高，也就是说，信道还远远没有被数据比特填满。

为了克服这一缺点，就产生了另外两种协议，即连续ARQ(亦称退后n帧协议)和选择重传ARQ(即选择重传协议)。这将在后面进一步讨论。

连续ARQ 协议的工作原理:

在发送完一个数据帧后，不是停下来等待确认帧，而是可以连续再发送若干个数据帧。

如果这时收到了接收端发来的确认帧，那么还可以接着发送数据帧。

由于减少了等待时间，整个通信的吞吐量就提高了。

滑动窗口的引入:

发送端和接收端分别设定发送窗口和接收窗口。

发送窗口用来对发送端进行流量控制。

发送窗口的大小WT 代表在还没有收到对方确认信息的情况下发送端最多可以发送多少个数据帧。

接收端设置接收窗口:

在接收端只有当收到的数据帧的发送序号落入接收窗口内才允许将该数据帧收下。

若接收到的数据帧落在接收窗口之外，则一律将其丢弃

在连续ARQ 协议中，接收窗口的大小WR = 1;

只有当收到的帧的序号与接收窗口一致时才能接收该帧。否则，就丢弃它。

每收到一个序号正确的帧，接收窗口就向前（即向右方）滑动一个帧的位置.同时发送对该帧的确认。

滑动窗口的重要特性:

只有在接收窗口向前滑动时（与此同时也发送了确认），发送窗口才有可能向前滑动。

收发两端的窗口按照以上规律不断地向前滑动，因此这种协议又称为滑动窗口协议。

当发送窗口和接收窗口的大小都等于1时，就是停止等待协议。

发送窗口的最大值：

当用n 个比特进行编号时，若接收窗口的大小为1，则只有在发送窗口的大小WT <=2n -1时，连续ARQ 协议才能正确运行

3.4 滑动窗口协议（1）

单工——> 全双工

捎带/载答（piggybacking）：暂时延迟待发确认，以便附加在下一个待发数据帧的技术。

优点：充分利用信道带宽，减少帧的数目意味着减少“帧到达”中断；

带来的问题：复杂。

本节的三个协议统称滑动窗口协议，都能在实际（非理想）环境下正常工作，区别仅在于效率、复杂性和对缓冲区的要求。

滑动窗口协议（Sliding Window Protocol）工作原理

1)发送的信息帧都有一个序号，从0到某个最大值，0 ~ 2的n次方 - 1，一般用n个二进制位表示；

2)发送端始终保持一个已发送但尚未确认的帧的序号表，称为发送窗口。发送窗口的上界表示要发送的下一个帧的序号，下界表示未得到确认的帧的最小编号发送窗口大小= 上界- 下界，大小可变；

3)发送端每发送一个帧，序号取上界值，上界加1；每接收到一个正确响应帧，下界加1；

4）接收端有一个接收窗口，大小固定，但不一定与发送窗口相同。接收窗口的上界表示允许接收的序号最大的帧，下界表示希望接收的帧；

5）接收窗口容纳允许接收的信息帧，落在窗口外的帧均被丢弃。序号等于下界的帧被正确接收，并产生一个响应帧，上界、下界都加1。接收窗口大小不变。

一比特滑动窗口协议（A One Bit Sliding Window Protocol）

协议特点1）窗口大小：N = 1，发送序号和接收序号的取值范围：0，1；

2）可进行数据双向传输，信息帧中可含有确认信息（piggybacking技术）；

3）信息帧中包括两个序号域：发送序号和接收序号（已经正确收到的帧的序号）

存在问题

能保证无差错传输，但是基于停等方式；

若双方同时开始发送，则会有一半重复帧；

Fig. 3-17（第三版、第四版、第五版翻译书上该图均有误）

效率低，传输时间长

2. 退后n帧协议(A Protocol Using Go Back n)

为提高传输效率而设计

例：卫星信道传输速率50kbps，往返传播延迟500ms，若传1000bit的帧，使用协议4，则传输一个帧所需时间为：发送时间+ 信息信道延迟+ 确认信道延迟（确认帧很短，忽略发送时间）= 1000bit / 50kbps + 250ms +250ms = 520ms

信道利用率= 20 / 520 约等于 4%

一般情况，信道带宽b 比特/秒，帧长度L比特，往返传播延迟R 秒，则信道利用率为: (L/b) / (L/b + R) = L / (L + Rb)

结论：传播延迟大、信道带宽高、帧短时，信道利用率低。

解决办法：连续发送多帧后再等待确认，称为流水线技术（pipelining）。

带来的问题：信道误码率高时，对损坏帧和非损坏帧的重传非常多

两种基本方法：

退后n帧（go back n）

接收方从出错帧起丢弃所有后继帧；

接收窗口为1；

对于出错率较高的信道，浪费带宽。

选择重传（selective repeat）

接收窗口大于1，先暂存出错帧的后继帧；

只重传坏帧；

对最高序号的帧进行确认；

接收窗口较大时，需较大缓冲区。

后退n帧的滑动窗口协议：

协议4的主要问题是信道利用率太低

发送端等待发送下一帧的时间至少是发送端到接收端信号传播时间的两倍，没有充分利用两条信道的传输能力

信道的利用率定义：数据发送时间/从数据开始发送到ACK返回的总耗时

协议5是一个发送管道化（pipelining）的协议；在等待ACK的时间内连续发送

出错后重发必须后退n帧：一旦重发定时器超时，必须将已发的n帧全部重发

可靠性传输;

差错控制：校验、重发和序号;避免帧错误的保证：帧的校验;避免帧丢失的保证：超时和重发;避免帧重复的保证：帧有序号

流量控制：窗口协议;发送方和接收方之间的协调

通过确认和重传机制

传输层协议，如TCP，也提供可靠传输服务

链路层的可靠传输服务通常用于高误码率的连路上，如无线链路。

3.5 常用的数据链路层协议（1）

ISO和CCITT在数据链路层协议的标准制定方面做了大量工作，各大公司也形成了自己的标准。

数据链路层协议分类:面向字符的链路层协议;ISO的IS1745，基本型传输控制规程及其扩充部分（BM和XBM）;IBM的二进制同步通信规程（BSC）;DEC的数字数据通信报文协议（DDCMP）; PPP

面向比特的链路层协议: IBM的SNA使用的数据链路协议SDLC（Synchronous

Data Link Control protocol）； ANSI修改SDLC，提出ADCCP（Advanced Data

Communication Control Procedure）； ISO修改SDLC，提出HDLC（High-level Data Link Control）； CCITT修改HDLC，提出LAP（Link Access Procedure）

作为X.25网络接口标准的一部分，后来改为LAPB。

高级数据链路控制规程HDLC: HDLC的组成 帧结构;规程元素;规程类型

使用HDLC的语法可以定义多种具有不同操作特点的链路层协议。

帧结构:

定界符:1) 01111110 2)空闲的点到点线路上连续传定界符

地址域（Address）: 多终端线路，用来区分终端；

点到点线路，有时用来区分命令和响应。

若帧中的地址是接收该帧的站的地址，则该帧是命令帧；

若帧中的地址是发送该帧的站的地址，则该帧是响应帧。

控制域（Control）: 序号,使用滑动窗口技术，3位序号，发送窗口大小为7

数据域（Data）:任意信息，任意长度（上层协议SDU有上限）,

校验和（Checksum）:CRC校验; 生成多项式：CRC-CCITT

帧类型: 信息帧（Information）;监控帧（Supervisory）;无序号帧（Unnumbered）

控制域: 序号（Seq）;使用滑动窗口技术，3位序号，发送窗口大小为7

捎带确认（Next）:捎带第一个未收到的帧序号，而不是最后一个已收到的帧序号

探询/结束P/F位（Poll/Final）：1）命令帧置“P”，响应帧置“F”。有些协议，P/F位用来强迫对方机器立刻发控制帧；

2）多终端系统中，计算机置“P”，允许终端发送数据；终端发向计算机的帧中，最后一个帧置为“F”，其它置为“P”。

类型（Type）：“0”表示确认帧RR(RECEIVE READY)；

“1”表示否定性确认帧REJ(REJECT)。

“2”表示接收未准备好RNR(RECEIVE NOT READY)

“3”表示选择拒绝SREJ(SELECTIVE REJECT)

无序号帧：可以用来传控制信息，也可在不可靠无连接服务中传数据。

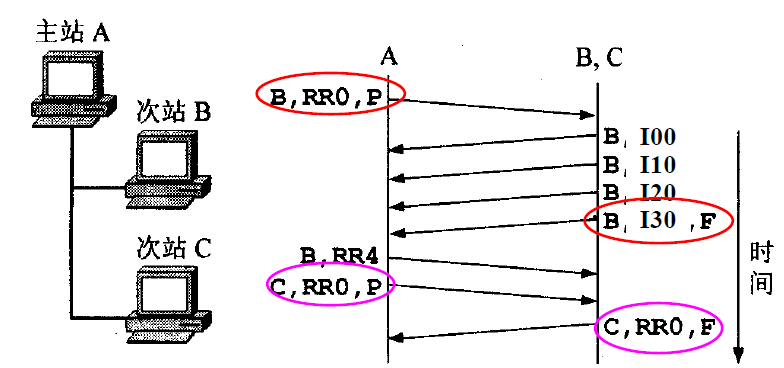
监控帧的第5比特也是P/F比特。

若P/F值为零，则P/F比特并没有任何意义。只有当P/F比特的值为1时才具有意义。

需要注意：在不同的数据传送方式中，P/F比特的用法是不一样的。在说明帧的传送过程中，为了更清楚地表示P/F比特的作用，往往将它写为P比特或F比特。

在非平衡配置的正常响应方式NRM中，次站不能主动向主站发送信息。次站只有收到主站发出的P比特为1的命令帧(S帧或I帧)以后才能发送响应帧。若次站有数据发送，则在最后一个数据帧中将F比特置1。若无数据发送，则应在回答的S帧中将F比特置1。

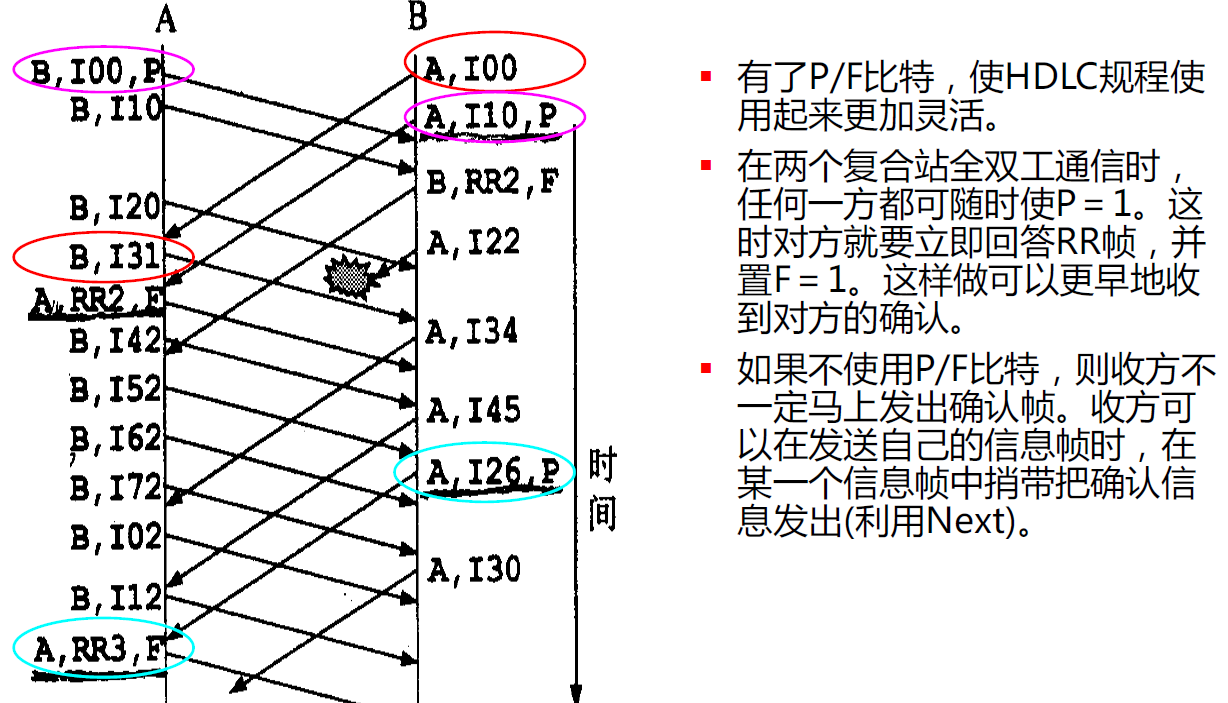
在非平衡配置的异步响应方式ARM或在平衡配置的异步平衡方式ABM中，任何一个站都可以在主动发送的S帧和I帧中将比特置1。对方站收到P = 1的帧后，应尽早地回答本站的状态并将P比特置1，不过此时并不表示数据已发完和不再发送数据了。



图中主站A与次站B和C连成多点链路。

我们将所传送的帧的一些主要参数按照“地址，帧名和序号，P/F”的先后顺序标注在图中。

这里的“地址”是指地址字段中应填入的站地址，“帧名”是指帧的名称，如RR或I， 序号则是指监控帧中的Next或信息帧中的SeqNext(R)。P/F是在当P/F为1时才写上P或F，表明此时控制字段的第5比特为1。



Internet数据链路层协议：

点到点通信的两种主要情形1）路由器到路由器2）通过modem拨号上网，连到路由器或接入服务器

SLIP —— Serial Line IP：

1984年，Rick Adams提出，RFC1055，发送原始IP包，用一个标记字节来定界，采用字符填充技术；

新版本提供TCP和IP头压缩技术，RFC 1144

存在的问题：不提供差错校验；只支持IP；IP地址不能动态分配；不提供认证；多种版本并存，互连困难

点到点协议PPP——Point-to-Point Protocol

RFC 1661，RFC 1662，RFC 1663

与SLIP相比，PPP有很大的提高，提供差错校验、支持多种协议、允许动态分配IP地址、支持认证等。

以帧为单位发送，而不是原始IP包；

包括两部分：1）链路控制协议LCP（Link Control Protocol）可使用多种物理层服务：modem，HDLC串线，SDH/SONET等； 2）网络控制协议NCP（Network Control Protocol）可支持多种网络层协议

PPP总结：具有多协议成帧机制，可以在modem, HDLC bitserial lines, SDH/SONET等物理层上运行，支持差错，检测、选项协商和包头压缩功能，并具有利用HDLC帧进行可靠传输的可选功能。

**本章小结：**

数据链路层的功能

组帧

差错控制：检错编码； 纠错编码

流量控制与可靠传输机制

流量控制、可靠传输与滑轮窗口机制

单帧滑动窗口与停止-等待协议

多帧滑动窗口与后退N帧协议（GBN）

多帧滑动窗口与选择重传协议（SR）

常见数据链路层协议： PPP协议、HDLC协议

Chapter 4 Media Access Control SubLayer

主要内容（1）

4.1 局域网概述

4.2 局域网拓扑结构和传输介质

4.3 局域网技术

4.3.1 信道分配

4.3.2 多路访问协议

4.4 局域网的IEEE 802系列标准

4.4.1 IEEE 802.3 和Ethernet

4.4.2 IEEE 802.4：Token Bus

令牌总线

4.4.3 IEEE 802.5：Token Ring

令牌环

4.4.4 几种局域网的比较

4.4.5 逻辑链路控制LLC

4.5 以太网的MAC 层

4.5.1 MAC 层的硬件地址

4.5.2 两种不同的MAC 帧格式

4.6 网桥技术

4.6.1 连接802.X和802.Y的网桥

4.6.2 透明网桥/生成数网桥

4.6.3 源路由网桥

4.7 高速局域网技术

4.7.1 光纤分布式数据接口FDDI

4.7.2 DPT (Dynamic Packet Transport)

4.7.3 快速以太网

4.7.4 千兆以太网

4.1 局域网概述

局域网产生的原因: 1)80年代，微型机发展迅速，彼此需要相互通信（近距离），共享资源； 2)分布式的网络应用：分布式计算，分布式数据库

定义:局域网是一种将小区域内的各种通信设备互连在一起的通信网络。

局域网最主要的特点是：网络为一个单位所拥有，且地理范围和站点数目均有限。

局域网的基本特点:

高数据传输率（10 ~ 1000 Mbps）

短距离（0.1 ~ 10 km）

低出错率（10-8 ~ 10-11）

局域网具有如下的一些主要优点：

能方便地共享昂贵的外部设备、主机以及软件、数据。从一个站点可访问全网。

便于系统的扩展和逐渐地演变，各设备的位置可灵活调整和改变。

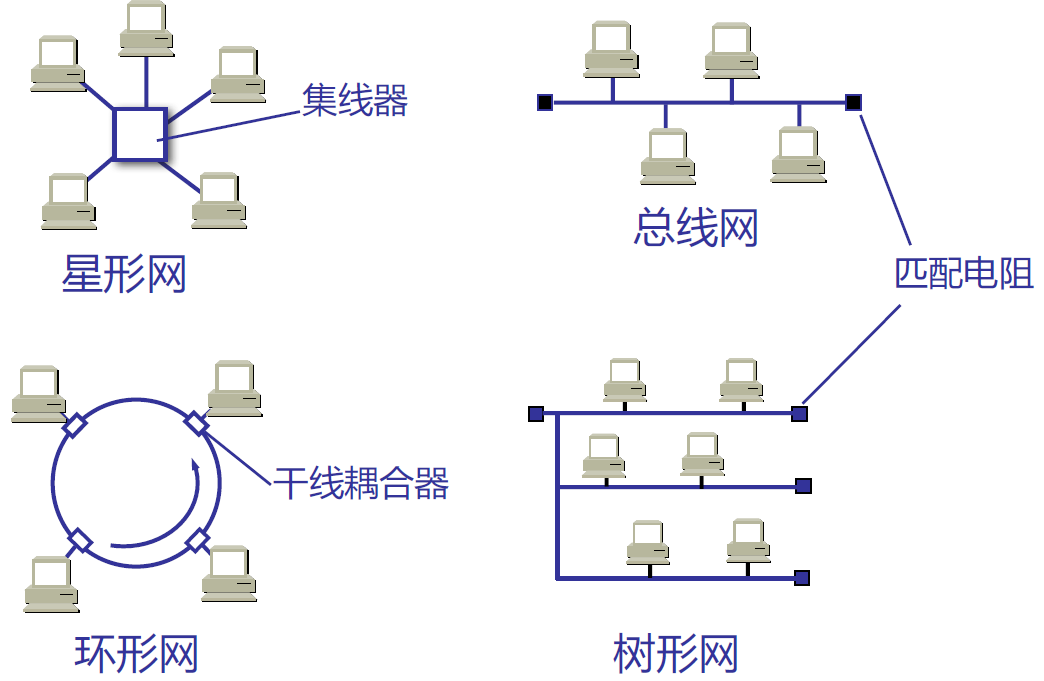
提高了系统的可靠性、可用性和残存性。

局域网发展趋势

高速，10G Ethernet; 移动，无线局域网IEEE 802.11

局域网拓扑结构: 星型结构; 环型结构;总线型结构;树型结构

传输介质:双绞线;基带同轴电缆;光纤;无线



4.3 局域网技术

4.3.1 信道分配

l 计算机网络可以分成两类

. 使用点到点连接的网络—— 广域网

. 使用广播信道（多路访问信道，随机访问信道）的网络——局域网

关键问题：如何解决对信道争用?

解决信道争用的协议称为介质访问控制协议MAC(Medium Access Control)，是数据链路层协议的一部分。

信道分配方法有两种: 静态分配;动态分配

媒体共享技术

静态划分信道:频分复用;时分复用;波分复用;码分复用

动态媒体接入控制（多点接入）:随机接入;受控接入，如多点线路探询(polling)，或轮询。

静态分配

频分多路复用FDM（波分复用WDM）

原理：将频带平均分配给每个要参与通信的用户

优点：适合于用户较少，数目基本固定，各用户的通信量都较大的情况

缺点：无法灵活地适应站点数及其通信量的变化

时分多路复用TDM

原理：每个用户拥有固定的信道传送时槽

优点：适合于用户较少，数目基本固定，各用户的通信量都较大的情况

缺点：无法灵活地适应站点数及其通信量的变化

动态分配

信道分配模型的五个基本假设：

站点模型：每个站点是独立的,并以统计固定的速率产生帧,一帧产生后到被发送走之前,站点被封锁；

单信道假设：所有的通信都是通过单一的信道来完成的，各个站点都可以从信道上收发信息

冲突假设：若两帧同时发出，会相互重叠，结果使信号无法辨认，称为冲突。所有的站点都能检测到冲突，冲突帧必须重发；

连续时间和时间分槽（确定何时发送）；

载波监听和非载波监听（确定能否发送）。

两个标准:

DIX Ethernet V2 是世界上第一个局域网产品（以太网）的规约。

IEEE 的802.3 标准。

DIX Ethernet V2 标准与IEEE 的802.3 标准只有很小的差别，因此可以将802.3 局域网简称“以太网”。

. 严格说来，“以太网”应当是指符合DIX Ethernet V2 标准的局域网

数据链路层的两个子层:

为了使数据链路层能更好地适应多种局域网标准，802 委员会就将局域网的数据链路层拆成两个子层：

. 逻辑链路控制LLC (Logical Link Control)子层

. 媒体接入控制MAC (Medium Access Control)子层。

与接入到传输媒体有关的内容都放在MAC子层，而LLC 子层则与传输媒体无关，不管采用何种协议的局域网对LLC 子层来说都是透明的

很多厂商生产的网卡上就仅装有MAC协议而没有LLC 协议。

4.3 局域网技术-网卡的作用

网络接口板又称为通信适配器(adapter)或网络接口卡NIC (Network Interface Card)，或“网卡;

网卡的重要功能: 进行串行/并行转换;对数据进行缓存;在计算机的操作系统安装设备驱动程序;实现以太网协议

计算机通过网卡和局域网进行通信

早期的以太网:

最初的以太网是将许多计算机都连接到一根总线上。当初认为这样的连接方法既简单又可靠，因为总线上没有有源器件。

以太网的广播方式发送:

总线上的每一个工作的计算机都能检测到B 发送的数据信号。

由于只有计算机D 的地址与数据帧首部写入的地址一致，因此只有D 才接收这个数据帧。

其他所有的计算机（A, C 和E）都检测到不是发送给它们的数据帧，因此就丢弃这个数据帧而不能够收下来。

具有广播特性的总线上实现了一对一的通信。

为了通信的简便以太网采取了两种重要的措施:

采用较为灵活的无连接的工作方式，即不必先建立连接就可以直接发送数据。

以太网对发送的数据帧不进行编号，也不要求对方发回确认:这样做的理由是局域网信道的质量很好，因信道质量产生差错的概率是很小的。

以太网提供的服务:

以太网提供的服务是不可靠的交付，即尽最大努力的交付。

当目的站收到有差错的数据帧时就丢弃此帧，其他什么也不做。差错的纠正由高层来决定。

如果高层发现丢失了一些数据而进行重传，但以太网并不知道这是一个重传的帧，而是当作一个新的数据帧来发送。

4.3 局域网技术（4）

4.3.2 多路访问协议

定义：控制多个用户共用一条信道的协议

（1）ALOHA协议

70年代，Norman Abramson设计了ALOHA协议

目的：解决信道的动态分配，基本思想可用于任何无协调关系的用户争用单一共享信道使用权的系统；

分类：纯ALOHA协议和分槽ALOHA协议;

纯ALOHA协议

基本思想：用户有数据要发送时，可以直接发至信道；然后监听信道看是否产生冲突，若产生冲突，则等待一段随机的时间重发

多用户共享单一信道，并由此产生冲突，这样的系统称为竞争系统；

随机接入的纯ALOHA协议:

随机接入的特点是所有的用户都可以根据自己的意愿随机地发送信息。总线网就属于这种类型：当两个或更多的用户同时发送信息时，就产生了帧的冲突(collision)。冲突又称为碰撞，它导致冲突用户的发送都告失败。现已研究出了多种解决冲突的网络协议。随机接入实际上就是争用接入，争用胜利者才可获得总线(即信道)，从而获得信息的发送权。

随机接入最初是在20世纪70年代初期在夏威夷大学试验成功的。这是为了使地理上分散的用户通过无线电来使用中心计算机。由于无线电信道是一个公用信道，一个站发送的信息可以同时被多个站收到，而每个站又是随机发送的，因此这种系统是一个随机接入系统。

夏威夷大学早期研制的系统称为ALOHA，是Additive Link On-line HAwaii system的缩写，而ALOHA恰好又是夏威夷方言的“你好”。

ALOHA系统中的两个归一化参数;

(1)吞吐量S 这又称为吞吐率，它等于在帧的发送时间T0内成功发送的平均帧数。显然，0≤S≤1，而S＝1是极限情况。

(2)网络负载(offered load)G 从网络的角度看，G等于在t内总共发送的平均帧数。在稳定状态下，吞吐量S与网络负载G的关系为：S＝ G·P[发送成功]这里P[发送成功]是一个帧发送成功的概率，它实际是发送成功的帧在所发送的帧的总数中所占的比例。

分槽ALOHA协议

基本思想：把信道时间分成离散的时间槽，槽长为一个帧所需的发送时间。每个站点只能在时槽开始时才允许发送。其他过程与纯ALOHA协议相同。

信道效率

冲突危险区是纯ALOHA的一半，所以P0=e-G，S=Ge-G；

与纯ALOHA协议相比，降低了产生冲突的概率，信道利用率最高为34.8%。（2）载波监听多路访问协议CSMA（Carrier Sense Multiple Access Protocols）

载波监听（Carrier Sense）

. 站点在为发送帧而访问传输信道之前，首先监听信道有无载波，若有载波，说明已有用户在使用信道，则不发送帧以避免冲突。

多路访问（Multiple Access）:多个用户共用一条线路

1-坚持型CSMA（1-persistent CSMA）

原理：1）若站点有数据发送，先监听信道；2）若站点发现信道空闲，则发送；

3）若信道忙，则继续监听直至发现信道空闲，然后完成发送；

4）若产生冲突，等待一随机时间，然后重新开始发送过程。

优点：减少了信道空闲时间；

缺点：增加了发生冲突的概率；

广播延迟对协议性能的影响：广播延迟越大，发生冲突的可能性越大，协议性能越差。

非坚持型CSMA（nonpersistent CSMA）

. 原理

若站点有数据发送，先监听信道；

若站点发现信道空闲，则发送；

若信道忙，等待一随机时间，然后重新开始发送过程；

若产生冲突，等待一随机时间，然后重新开始发送过程。

. 优点：减少了冲突的概率；

. 缺点：增加了信道空闲时间，数据发送延迟增大

. 信道效率比1-坚持CSMA高，传输延迟比1-坚持CSMA大。

p-坚持型CSMA（p-persistent CSMA）

适用于分槽信道

原理

若站点有数据发送，先监听信道；

若站点发现信道空闲，则以概率p发送数据，以概率

q =1- p 延迟至下一个时槽发送。若下一个时槽仍空闲，重复此过程，直至数据发出或时槽被其他站点所占用；

若信道忙，则等待下一个时槽，重新开始发送

若产生冲突,等待一随机时间,然后重新开始发送

（3）带冲突检测的载波监听多路访问协议CSMA/CD

引入原因

当两个帧发生冲突时，两个被损坏帧继续传送毫无意义，而且信道无法被其他站点使用，对于有限的信道来讲，这是很大的浪费。如果站点边发送边监听,并在监听到冲突之后立即停止发送，可以提高信道的利用率,因此产生了CSMA/CD。

原理:

站点使用CSMA协议进行数据发送；

在发送期间如果检测到冲突，立即终止发送，并发出一个瞬间干扰信号，使所有的站点都知道发生了冲突；

在发出干扰信号后，等待一段随机时间，再重复上述过程。

CSMA/CD说明:

CSMA/CD 表示Carrier Sense Multiple Access with Collision Detection。

“多点接入”表示许多计算机以多点接入的方式连接在一根总线上。

“载波监听”是指每一个站在发送数据之前先要检测一下总线上是否有其他计算机在发送数据，如果有，则暂时不要发送数据，以免发生碰撞。

总线上并没有什么“载波”。因此,”载波监听”就是用电子技术检测总线上有没有其他计算机发送的数据信号。

碰撞检测：

“碰撞检测”就是计算机边发送数据边检测信道上的信号电压大小。

当几个站同时在总线上发送数据时，总线上的信号电压摆动值将会增大（互相叠加）。

当一个站检测到的信号电压摆动值超过一定的门限值时，就认为总线上至少有两个站同时在发送数据，表明产生了碰撞。

所谓“碰撞”就是发生了冲突。因此“碰撞检测”也称为“冲突检测”。

检测到碰撞后：

在发生碰撞时，总线上传输的信号产生了严重的失真，无法从中恢复出有用的信息来。

每一个正在发送数据的站，一旦发现总线上出现了碰撞，就要立即停止发送，免得继续浪费网络资源，然后等待一段随机时间后再次发送。

重要特性：使用CSMA/CD 协议的以太网不能进行全双工通信而只能进行双向交替通信（半双工通信）。

每个站在发送数据之后的一小段时间内，存在着遭遇碰撞的可能性。

这种发送的不确定性使整个以太网的平均通信量远小于以太网的最高数据率。

争用期：最先发送数据帧的站，在发送数据帧后至多经

过时间2t （两倍的端到端往返时延）就可知道发送的数据帧是否遭受了碰撞。

以太网的端到端往返时延2t 称为争用期，或碰撞窗口。

经过争用期这段时间还没有检测到碰撞，才能肯定这次发送不会发生碰撞。

二进制指数类型退避算法(truncated binary exponential type)

发生碰撞的站在停止发送数据后，要推迟（退避）一个随机时间才能再发送数据。

确定基本退避时间，一般是取为争用期2t。

定义重传次数i ，i <=10，即 i = Min[重传次数, 10]

从整数集合[0,1,…, (2的i次方 -1)]中随机地取出一个数，记为r。重传所需的时延就是r 倍的基本退避时间。

当重传达16 次仍不能成功时即丢弃该帧，并向高层报告。

争用期的长度：

以太网取51.2 μs 为争用期的长度。

对于10 Mb/s 以太网，在争用期内可发送512 bit，即64 字节。

以太网在发送数据时，若前64 字节没有发生冲突，则后续的数据就不会发生冲突。

最短有效帧长：

如果发生冲突，就一定是在发送的前64 字节之内。

由于一检测到冲突就立即中止发送，这时已经发送出去的数据一定小于64 字节。

以太网规定了最短有效帧长为64 字节，凡长度小于64 字节的帧都是由于冲突而异常中止的无效帧。

强化碰撞：

当发送数据的站一旦发现发生了碰撞时，除了立即停止发送数据外，还要再继续发送若干比特的人为干扰信号(jamming signal)，以便让所有用户都知道现在已经发生了碰撞。

工作状态：传输周期,竞争周期,空闲周期

一个站点确定发生冲突要花多少时间？最坏情况下，2倍电缆传输时间

（4）无冲突协议（Collision-Free Protocols）

基本位图协议（A Bit-Map Protocol）

工作原理:共享信道上有N个站，竞争周期分为N个时槽，如果一个站有帧发送,则在对应的时槽内发送比特1； N个时槽之后，每个站都知道哪个站要发送帧，这时按站序号发送。

象这样在实际发送信息前先广播发送请求的协议称为预留协议（reservation protocol）

效率:轻负载下，效率为d / (N + d)，数据帧由d个时间单位组成；重负载下，效率为d / (d + 1)。

缺点:1)与站序号有关的不平等性,序号大的站得到的服务好；2)每个站都有1 比特的开销。

二进制下数法（Binary Countdown）: 工作原理

所有站的地址用等长二进制位串表示，若要占用信道，则广播该位串

不同站发的地址中的位做“或”操作，一旦某站了解到比本站地址高位更高的位置被置为“1”，便放弃发送请求.

效率:d / (d + log2N)

（5）有限竞争协议（Limited-Contention Protocols）

占用信道的策略

竞争方法

例，CSMA；

轻负载下，发送延迟小；重负载下，信道效率低

无冲突方法

例，基本位图法；

轻负载下，发送延迟大；重负载下，信道效率高

有限竞争方法

结合以上两种方法，轻负载下使用竞争，重负载下使用无冲突方法。

减少竞争的站的数目可以增加获取信道的概率

基本思路：将站分组，组内竞争；

问题：如何分组？

适应树搜索协议（The Adaptive Tree Walk Protocol）:

工作原理:站点组织成二叉树；

一次成功传输之后，第0槽全部站可竞争信道，只有一个站要使用信道则发送；有冲突则在第1槽内半数站（2以下站）参与竞争。如其中之一获得信道，本帧后的时槽留给3以下的站；如发生冲突，继续折半搜索。

. 当系统负载很重时，从根结点开始竞争发生冲突的概率非常大。为提高效率，可以从中间结点开始竞争。问题：搜索应该从树的哪一级开始？

（6） 无线局域网协议

无线局域网产生背景

. 笔记本电脑的普及促进了无线局域网的发展

. portable ≠ mobile

. 要做到真正的移动，需要使用无线信号进行通信

无线局域网的特点

基于蜂窝（cell）的通信

每个蜂窝内只有一个信道（与蜂窝电话不同）

一个站点发送的信号，只能被它周围一定范围内的站点接收到

短距离传输

无线局域网与有线局域网不同，具有

1)隐藏站点问题（hidden station problem）

由于站点距离竞争者太远，从而不能发现潜在介质竞争者的问题称为隐藏站点问题。

A向B发送数据的过程中，C由于收不到A的数据，也可以向B发送数据，导致B接收发生冲突。

1. 暴露站点问题（exposed station problem）

由于非竞争者距离发送站点太近，从而导致介质非竞争者不能发送数据的问题称为暴露站点问题; B向A发送数据，被C监听到，导致C不能向D发送数据

传统的CSMA协议不适合于无线局域网，需要特殊的MAC子层协议

CSMA

在电缆上，信号传播给所有站点

CSMA只判断本发送站点周围是否有活跃发送站点

冲突被发送站点发现

某一时刻，信道上只能有一个有效数据帧

. 无线局域网

信号只能被发送站点周围一定范围内的站点接收

MAC子层协议需要尽量保证接收站点周围一定范围内只有一个发送站点。

冲突被接收站点发现

某一时刻，信道上可以有多个有效数据帧

MACA（Multiple Access with Collision Avoidance）多重访问避免碰撞

是IEEE 802.11无线局域网标准的基础

基本思想：发送站点刺激接收站点发送应答短帧，从而使得接收站点周围的站点监听到该帧，并在一定时间内避免发送数据

基本过程

. A向B发送RTS(Request To Send)帧，A周围的站点在一定时间内不发送数据，以保证CTS帧返回给A；

. B向A回答CTS(Clear To Send)帧，B周围的站点在一定时间内不发送数据，以保证A发送完数据；

. A开始发送

. 若发生冲突，采用二进制指数后退算法等待随机时间，再重新开始。

MACAW:

对MACA协议做了改进，提高了性能

Ø 主要改进

对每个成功传输的数据帧，都要产生确认帧

增加了（发送站点的）载波监听

发生冲突后，针对每个数据流（相同源和目的地址）执行后退算法，而不是针对每个站点

发生拥塞时，站点间交互信息

**小结:**

关键问题：在MAC子层,如何解决多个站点争用共享

信道？

信道分配方式：静态(FDM、WDM、TDM)，动态

多路访问协议

. ALOHA：纯ALOHA、分槽ALOHA，（冲突危险区）

. CAMA：1-坚持型CSMA、非坚持型CSMA、p-坚持型,CSMA（分槽协议）

. CSMA/CD

无冲突协议、有限竞争协议

无线局域网协议

**4.4 局域网的IEEE 802系列标准**

LAN 指“局域网”(Local Area Network)

MAC（Media Access Control或者Medium Access Control）地址:意译为媒体访问控制，或称为物理地址、硬件地址，用来定义网络设备的位置。

IEEE 802协议:

IEEE 802系列标准定义了若干种LAN，包括对物理层、MAC子层的定义和描述。它的组成如下：

802.1 基本介绍和接口原语定义

802.2 逻辑链路控制（LLC）子层

802.3 采用CSMA/CD技术的局域网

802.4 采用令牌总线（Token Bus）技术的局域网

802.5 采用令牌环（Token Ring）技术的局域网

LAN的参考模型:

逻辑链路控制子层LLC（Logical Link Control）

引入LLC子层的原因：

MAC子层只提供尽力而为的数据报服务，不提供确认机制和流量控制（滑动窗口），有些情况下，这种服务足够，如支持IP协议；当需要确认和流控的时候，这种服务就不能满足，需要LLC。

LLC子层提供确认机制和流量控制；

LLC隐藏了不同802MAC子层的差异，为网络层提供单一的格式和接口；

LLC提供三种服务选项：不可靠数据报服务;有确认数据报服务;可靠的面相连接的服务

LLC帧头基于HDLC协议

介质访问控制子层MAC（Medium Access Control）

数据封装（发送和接收）:成帧（帧定界、帧同步）;寻址（源和目的地址处理）;差错检测.

介质访问管理:介质分配（冲突避免）;冲突解决（处理冲突）

分成两个子层的原因

管理多点访问信道的逻辑不同于传统的数据链路控制；

对于同一个LLC，可以提供多个MAC选择

传统以太网的连接方法:

传统以太网可使用的传输媒体有四种：

铜缆（粗缆或细缆） 铜线（双绞线） 光缆.

这样，以太网就有四种不同的物理层。

网卡的功能:

数据的封装与解封发送时将上一层交下来的数据加上首部和尾部，成为以太网的帧。接收时将以太网的帧剥去首部和尾部，然后送交上一层。

链路管理主要是CSMA/CD 协议的实现。

编码与译码即曼彻斯特编码与译码。

星形网10BASE-T:

不用电缆而使用无屏蔽双绞线。每个站需要用两对双绞线，分别用于发送和接收。

在星形网的中心则增加了一种可靠性非常高的设备，叫做集线器(hub)。

集线器使用了大规模集成电路芯片，因此这样的硬件设备的可靠性已大大提高了。

以太网在局域网中的统治地位:

10BASE-T 的通信距离稍短，每个站到集线器的距离不超过100 m。

这种10 Mb/s 速率的无屏蔽双绞线星形网的出现，既降低了成本，又提高了可靠性。

10BASE-T 双绞线以太网的出现，是局域网发展史上的一个非常重要的里程碑，它为以太网在局域网中的统治地位奠定了牢固的基础。

集线器的一些特点:

集线器是使用电子器件来模拟实际电缆线的工作，因此整个系统仍然像一个传统的以太网那样运行。

使用集线器的以太网在逻辑上仍是一个总线网，各工作站使用的还是CSMA/CD 协议，并共享逻辑上的总线。

集线器很像一个多端口的转发器，工作在物理层。

10Base5含义:10：10Mbps; Base：基带传输（baseband medium）；5：500米

10Base5：粗缆，AUI接口；

10Base2：细缆，BNC接口，T型头；

10Base-T：RJ-45接口

收发器（transceiver）：处理载波监听和冲突检测

布线拓扑结构: 总线形，脊椎形，树形，分段

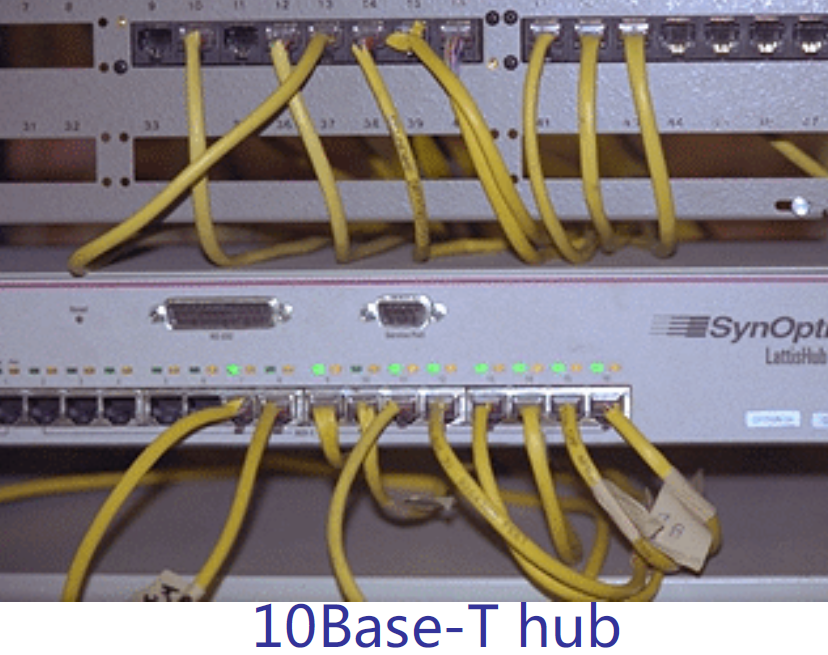
扩展网段长度

中继器：物理层设备，只对信号进行接收、放大和双向重传；

两个收发器之间最多使用4个中继器，最长2500米

802.3的信号编码

由于曼彻斯特编码的简单，所有的802.3基带系统都使用曼彻斯特编码。



802.3的MAC子层帧格式: 前导序列（7个字节10101010）;帧开始标志（1字节， 10101011）

目标地址和源地址: 2 或6个字节，以太网为6个字节(IEEE 802 specifies the use of either 16- or 48-bit addresses, no conformant implementation of IEEE 802.3 uses 16-bit addresses.)

目的地址第一位（LSB: Least Significant Bit）为0，表示单地址（individual address）；为1，表示组地址（group address），支持multicast，目的地址全1，为广播地址。源地址第一位（LSB）为0。

地址中的第二位（LSB）用来区分本地地址和全球地址

帧长度域（2字节，取值在0-1500之间）

数据（0-1500个字节）

填充（0-46字节）

校验和：CRC校验（4个字节）

扩展域：用于千兆以太网

以太网MAC 层的硬件地址:

在局域网中，硬件地址又称为物理地址，或MAC 地址。

802 标准所说的“地址”严格地讲应当是每一个站的“名字”或标识符。

但鉴于大家都早已习惯了将这种48 bit 的“名字”称为“地址”，所以我们也采用这种习惯用法，尽管这种说法并不太严格。

网卡检查MAC 地址:

网卡从网络上每收到一个MAC 帧就首先用硬件检查MAC 帧中的MAC 地址.

如果是发往本站的帧则收下，然后再进行其他的处理。

否则就将此帧丢弃，不再进行其他的处理。

“发往本站的帧”包括以下三种帧：

Ø 单播(unicast)帧（一对一）

Ø 广播(broadcast)帧（一对全体）

Ø 多播(multicast)帧（一对多）

常用的以太网MAC帧格式有两种标准：

. DIX Ethernet V2 标准

. IEEE 的802.3 标准

最常用的MAC 帧是以太网V2 的格式。

以太帧和802.3帧:

在TCP/IP世界中，以太网IP数据报文的封装在RFC 894中定义，IEEE802.3网络的IP数据报文封装在RFC 1042中定义。标准规定：

1）主机必须能发送和接收采用RFC 894（以太网）封装格式的分组；

2）主机应该能接收RFC 1042（IEEE 802.3）封装格式的分组；

3）主机可以发送采用RFC 1042（IEEE 802.3）封装格式的分组。如果主机能同时发送两种类型的分组数据，那么发送的分组必须是可以设置的，而且默认条件下必须是RFC 894（以太网）。

最常使用的封装格式是RFC 894定义的格式，俗称Ethernet或者Ethernet DIX。下面，我们就以Ethernet II称呼RFC 894定义的以太帧，以IEEE802.3称呼RFC 1042定义的以太帧。

类型字段还是长度字段:

除了吉比特以太网之外的所有以太网版本，如果传输数据少于46个字节，应将数据字段填充至46字节。不过，填充字符的个数不包括在长度字段值中。

同时支持以太网和IEEE802.3帧格式的网络接口卡通过这一字段的值区分这两种帧。也就是说，因为数据字段的最大长度为1500字节，所以超过十六进制数05DC的值说明它不是长度字段（IEEE802.3).而是类型字段（Ethernet II）。

当数据字段的长度小于46 字节时，应在数据字段的后面加入整数字节的填充字段，以保证以太网的MAC 帧长不小于64 字节。

无效的MAC 帧:

l 数据字段的长度与长度字段的值不一致；

l 帧的长度不是整数个字节；

l 用收到的帧检验序列FCS 查出有差错；

l 数据字段的长度不在46 ~ 1500 字节之间。

l 有效的MAC 帧长度为64 ~ 1518 字节之间。

l 对于检查出的无效MAC 帧就简单地丢弃。以太网不负责重传丢弃的帧。

最短帧长:

避免帧的第一个比特到达电缆的远端前帧已经发完，帧发送时间应该大于2t；10Mbps LAN，最大冲突检测时间为51.2微秒，最短帧长为64字节；

网络速度提高，最短帧长也应该增大或者站点间的距离要减小。

二进制指数后退算法（binary exponential backoff）:

1)将冲突发生后的时间划分为长度为51.2微秒的时槽

2)发生第一次冲突后，各个站点等待0 或1 个时槽再开始重传；

3)发生第二次冲突后，各个站点随机地选择等待0,1, 2或3个时槽再开始重传；

4)第i 次冲突后，在0 至2i-1 间随机地选择一个等待的时槽数，再开始重传；

5)10次冲突后，选择等待的时槽数固定在0至210-1间；

6)16次冲突后，发送失败，报告上层。

交换式802.3 LAN:

目的：减少冲突；

两种实现方法: 一个卡内是一个802.3LAN，构成自己的冲突域，卡间

并行；使用端口缓存，无冲突发生。

4.4.2 IEEE 802.4：令牌总线Token Bus

令牌总线局域网在物理上是一个总线网，而在逻辑上却是一个令牌网。

这样，令牌总线局域网既具有总线网的“接入方便”和“可靠性较高”的优点，也具有令牌环形网的“无冲突”和“发送时延有确定的上限值”的优点。

在令牌总线局域网中，令牌传递的顺序不是按照站的物理位置，因此必须有一个有效的MAC子层协议来管理网络的令牌。这就使得令牌总线局域网的MAC子层协议非常复杂(802.4标准的原文超过200页)。

4.4.3 IEEE 802.5：令牌环Token Ring

技术产生原因: 1)环实际上并不是一个广播介质，而是不同的点到点链路组成的环，点到点链路有很多技术优势；

2) 各个站点是公平的，获得信道的时间有上限，避免冲突发生；

3) IBM选择Token Ring作为它的LAN技术。

基本思想:

令牌（Token）是一种特殊的比特组合模式，一个站要发送帧时，需要抓住令牌，并将其移出环；

Ø 环本身必须有足够的时延容纳一个完整的令牌，时延由两部分组成：每站的1比特延迟和信号传播延迟。对于短环，必要时需要插入人工延迟；

Ø 环接口有两种操作模式：监听模式和传输模式。

Ø 当一个站点有数据发送时，在令牌通过此站点时，将令牌从环上取下，发送自己的数据，然后重新生成令牌，发送站负责将发出的帧从环上移去，并转入监听模式。

确认：帧内一个比特域，初值为0，目的站收到后，将其变为1；对广播的确认比较复杂；

重负载下，效率接近100%。

环网设计分析的一个主要问题是1 比特的“物理长度”，数据传输速率为R Mbps，典型信号传播速率为200米/微秒，则1 比特的“物理长度”为200/R米；

环接口引入了1比特的传输延迟。

802.5的布线:

屏蔽双绞线，速率为1/4/16M，采用差分曼彻斯特编码传输；

为解决环断裂导致整个环无法工作的问题，使用线路中心（Wire Center）进行布线，线路中心设有旁路中继器。

令牌环MAC子层协议:

协议基本操作：无信息传输时，3字节的令牌在环上循环；有信息要发送时，站获得令牌，并将第二个字节的某一位由0 变成1，将令牌的前两个字节变成帧的起始序列，然后输出帧的其它部分;

开始定界符SD和结束定界符ED标志着帧的开始和结束,使用差分曼彻斯特编码模式(HH和LL，物理层编码违例法)；

Ø 访问控制域AC 包括令牌位、监视位、优先级位和保留位；

Ø 帧控制域FC 用于将数据帧和控制帧区别开来和进行环的维护；帧状态字节FS用于报告帧的传送情况，包括地址位A和拷贝位C，帧经过目的站，A置为“1”，帧被接收，C置为“1”。A、C位提供了自动确认。为增加可靠性，A、C在FS中出现两次。

A = 0，C = 0，目的站不存在或未加电；

A = 1，C = 0，目的站存在但帧未被接收；

A = 1，C = 1，目的站存在且帧被复制。

Ø 令牌持有时间（token-holding time），一般为10毫秒；

Ø 提供优先级控制：访问控制域中的优先级位给出令牌的优先级，只有当要发送的帧的优先级大于等于令牌的优先级时才能获得令牌，站还可以预约某个优先级的令牌。

环的维护:

环上存在一个监控站，负责环的维护，通过站的竞争产生；

- 监控站的职责

• 保证令牌不丢失；

• 处理环断开情况；

• 清除坏帧，检查无主帧。

4.4.4 三种局域网的比较

802.3

Ø 优点

使用最为广泛；

算法简单；

站点可以在网络运行中安装；

使用无源电缆；

轻负载时，延迟为0。

Ø 缺点

使用模拟器件,每个站点在发送的同时要检测冲突；

最短帧长64字节，对于短数据来讲开销太大；

无优先级,发送是非确定性的，不适合于实时工作；

电缆最长2500米（使用中继器）；

速率提高时，帧传输时间减少，竞争时间不变（2t），效率降低；

重负载时，冲突严重。

802.4

优点

发送具有确定性，支持优先级，可处理短帧；

使用宽带电缆，支持多信道；

重负载时，吞吐量和效率较高。

缺点:

使用大量的模拟装置；

协议复杂；

轻负载时，延迟大；

很难用光纤实现。

802.5

优点

使用点到点连接，完全数字化；

使用线路中心，自动检测和消除电缆故障；

支持优先级，允许短帧，但受令牌持有时间限制，不允许任意长的帧；

重负载时，吞吐量和效率较高。

缺点

中央监控；

轻负载时，延迟大。

唯一可做的结论是：在很重的负载下802.3局域网彻底不能用，而基于令牌的局域网则可达到接近于100 %的效率。若负载范围是从轻到中等，则三种局域网都能胜任。

在物理层扩展局域网:用多个集线器可连成更大的局域网

用集线器扩展局域网: 优点

Ø 使原来属于不同碰撞域的局域网上的计算机能够进行跨碰撞域的通信。

Ø 扩大了局域网覆盖的地理范围。

缺点

Ø 碰撞域增大了，但总的吞吐量并未提高。

Ø 如果不同的碰撞域使用不同的数据率，那么就不能用集线器将它们互连起来。

在数据链路层扩展局域网

在数据链路层扩展局域网是使用网桥。

网桥工作在数据链路层，它根据MAC 帧的目的地址对收到的帧进行转发。

网桥具有过滤帧的功能。当网桥收到一个帧时，并不是向所有的端口转发此帧，而是先检查此帧的目的MAC 地址，然后再确定将该帧转发到哪一个端口

4.5 网桥技术

定义：网桥（bridge）是工作在数据链路层的一种网络互连设备，它在互连的LAN之间实现帧的存储和转发。

为什么使用桥？

. 学校和企业的各个部门分别拥有自己独立管理的LAN，为了进行交互，需要使用桥来实现互连；

. 一个企业分布在相隔很远的不同建筑物内，在每个建筑物内组建单独的LAN，并使用桥将这些LAN连接起来，是比较经济的方案；

使用网桥带来的好处:

过滤通信量。

扩大了物理范围。

提高了可靠性。

可互连不同物理层、不同MAC 子层和不同速率（如10 Mb/s 和100 Mb/s 以太网）的局域网。

使用网桥带来的缺点:

存储转发增加了时延。

在MAC 子层并没有流量控制功能。

具有不同MAC 子层的网段桥接在一起时时延更大。

网桥只适合于用户数不太多(不超过几百个)和通信量不太大的局域网，否则有时还会因传播过多的广播信息而产生网络拥塞。这就是所谓的广播风暴。

网桥和集线器（或转发器）不同:

集线器在转发帧时，不对传输媒体进行检测。

网桥在转发帧之前必须执行CSMA/CD 算法。

Ø 若在发送过程中出现碰撞，就必须停止发送和进行退避。

Ø 在这一点上网桥的接口很像一个网卡。但网桥却没有网卡。

由于网桥没有网卡，因此网桥并不改变它转发的帧的源地址。

透明网桥:

目前使用得最多的网桥是透明网桥(transparent bridge)。

“透明”是指局域网上的站点并不知道所发送的帧将经过哪几个网桥，因为网桥对各站来说是看不见的。

透明网桥是一种即插即用设备，其标准是IEEE 802.1D。

网桥按以下算法处理收到的帧和建立转发表:

(1) 从端口x 收到无差错的帧（如有差错即丢弃），在转发表中查找目的站MAC 地址。

(2) 如有，则查找出到此MAC 地址应当走的端口d，然后进行(3)，否则转到(5)。

(3) 如到这个MAC 地址去的端口d = x，则丢弃此帧（因为这表示不需要经过网桥进行转发）。否则从端口d 转发此帧。

(4) 转到(6)。

(5) 向网桥除x 以外的所有端口转发此帧（这样做可保证找到目的站）。

(6) 如源站不在转发表中，则将源站MAC 地址加入到转发表，登记该帧进入网桥的端口号，设置计时器。然后转到(8)。如源站在转发表中，则执行(7)。

(7) 更新计时器。

(8) 等待新的数据帧。转到(1)。

网桥在转发表中登记以下三个信息:

站地址：登记收到的帧的源MAC 地址。

端口：登记收到的帧进入该网桥的端口号。

时间：登记收到的帧进入该网桥的时间。

转发表中的MAC 地址是根据源MAC 地址写入的，但在进行转发时是将此MAC 地址当作目的地址。

Ø 如果网桥现在能够从端口x 收到从源地址A 发来的帧，那么以后就可以从端口x 将帧转发到目的地址A。

源路由网桥:

透明网桥容易安装，但网络资源的利用不充分。

源路由(source route)网桥在发送帧时将详细的路由信息放在帧的首部中。

源站以广播方式向欲通信的目的站发送一个发现帧，每个发现帧都记录所经过的路由。

发现帧到达目的站时就沿各自的路由返回源站。源站在得知这些路由后，从所有可能的路由中选择出一个最佳路由。凡从该源站向该目的站发送的帧的首部，都必须携带源站所确定的这一路由信息。

4. 多端口网桥——以太网交换机

1990 年问世的交换式集线器(switching hub)，可明显地提高局域网的性能。

交换式集线器常称为以太网交换机(switch)或第二层交换机（表明此交换机工作在数据链路层）。

以太网交换机通常都有十几个端口。因此，以太网交换机实质上就是一个多端口的网桥，可见交换机工作在数据链路层。

以太网交换机的特点:

以太网交换机的每个端口都直接与主机相连，并且一般都工作在全双工方式。

交换机能同时连通许多对的端口，使每一对相互通信的主机都能像独占通信媒体那样，进行无碰撞地传输数据。

以太网交换机由于使用了专用的交换结构芯片，其交换速率就较高。

独占传输媒体的带宽:

对于普通10 Mb/s 的共享式以太网，若共有N 个用户，则每个用户占有的平均带宽只有总带宽(10 Mb/s)的N 分之一。

使用以太网交换机时，虽然在每个端口到主机的带宽还是10 Mb/s，但由于一个用户在通信时是独占而不是和其他网络用户共享传输媒体的带宽，因此对于拥有N 对端口的交换机的总容量为N\*10Mb/s。这正是交换机的最大优点。

技术: 将一个负载很重的大LAN分隔成使用网桥互连的几个LAN以减轻负担；

LAN上的两台机器其距离超过2500米，必须使用网桥将这个LAN 分隔以保证网络的正常工作；

网桥可以互连不同类型的LAN；

网桥可以隔离负载，防止出故障的站点损害全网

网桥可以有助于安全保密。

网桥的工作原理:

连接k个不同LAN的网桥具有k个MAC子层和k个物理层。

4.5.1连接802.X和802.Y的网桥

互连时需要解决的相同问题

. 不同LAN帧格式的转换；

. 不同的LAN速率不同，网桥要有缓存能力；

. 高层协议的计时器设置；

. 不同的LAN支持的最大帧长度不同，分别为1500(802.3)，8191(802.4) ， 5000(802.5) 。解决办法：丢弃无法转发的帧。

三种不同的LAN互连共有九种组合:

动作: 格式转换和重新计算校验和; 变换位的顺序;复制优先级; 产生一个虚拟的优先级;放弃优先级;把环排空; 设置A/C位;解决速率快慢问题;处理帧太长的问题

4.5.2 透明网桥/生成树网桥:

工作原理：

网桥工作在混杂（promiscuous）方式，接收所有的帧；

网桥接收到一帧后，通过查询地址/端口对应表来确定是丢弃还是转发；

网桥刚启动时，地址/端口对应表为空，采用洪泛（flooding）方法转发帧；

在转发过程中采用逆向学习(backward learning)算法收集MAC地址。网桥通过分析帧的源MAC地址得到MAC地址与端口的对应关系,并写入地址/端口对应表；

网桥软件对地址/端口对应表进行不断的更新，并定时检查，删除在一段时间内没有更新的地址/端口项；

帧的路由过程

目的LAN与源LAN相同，则丢弃帧；

目的LAN与源LAN不同，则转发帧；

目的LAN未知，则洪泛帧。

多个网桥（并行网桥）可能产生回路

透明网桥使用了支撑树算法: 这是为了避免产生转发的帧在网络中不断地兜圈子

解决多个网桥产生回路的问题:

思想

让网桥之间互相通信，用一棵连接每个LAN的生成树(Spanning Tree）覆盖实际的拓扑结构

构造生成树

每个桥广播自己的桥编号，号最小的桥称为生成树的根；

每个网桥计算自己到根的最短路径，构造出生成树，使得每个LAN和桥到根的路径最短；

当某个LAN或网桥发生故障时，要重新计算生成树；

生成树构造完后，算法继续执行以便自动发现拓扑结构变化，更新生成树。

4.4.3 源路由网桥:

CSMA/CD和Token Bus选择了透明网桥，Token Ring选择了源路由网桥

源路由网桥的原理

. 帧的发送者知道目的主机是否在自己的LAN内；

. 如果不在，在发出的帧头内构造一个准确的路由序列，包含要经过的网桥、LAN的编号。并将发出的帧的源地址的最高位置1；

每个LAN有一个12位的编号,每个网桥有一个4位的编号；

Ø 网桥只接收源地址的最高位为1的帧，判定是转发还是丢弃；

Ø 源路由的产生：每个站点通过广播“发现帧”(discovery frame)来获得到各个站点的最佳路由。

若目的地址未知，源站发送“发现帧”，每个网桥收到后广播，目的站收到后发应答帧，该帧经过网桥时被加上网桥的标识，源站收到后就知道了到目的站的最佳路由。

优点

- 对带宽进行最优的使用。

缺点

- 网桥的插入对于网络是不透明的，需要人工干预。站点要知道网络的拓扑结构。

交换机内的电路让每个计算机位于单独的局域网网段上并与其他网段通过网桥连接

虚拟局域网VLAN:

虚拟局域网VLAN 是由一些局域网网段构成的与物理位置无关的逻辑组。

. 这些网段具有某些共同的需求。

. 每一个VLAN 的帧都有一个明确的标识符，指明发送这个帧的工作站是属于哪一个VLAN。

虚拟局域网其实只是局域网给用户提供的一种服务，而并不是一种新型局域网。

虚拟局域网限制了接收广播信息的工作站数，使得网络不会因传播过多的广播信息(即“广播风暴”)而引起性能恶化。

4.5.2 虚拟局域网使用的以太网帧格式

虚拟局域网协议允许在以太网的帧格式中插入一个4 字节的标识符，称为VLAN 标记(tag)，用来指明发送该帧的工作站属于哪一个虚拟局域网。



4.6 高速以太网

速率达到或超过100 Mb/s 的以太网称为高速以太网。

100BASE-T是在双绞线上传送100 Mb/s 基带信号的星型拓扑以太网，仍使用IEEE 802.3的CSMA/CD 协议。100BASE-T 以太网又称为快速以太网(Fast Ethernet)。

4.6.1光纤分布式数据接口FDDI（Fiber Distributed Data Interface）

特征

使用多模光纤作为传输介质

MAC协议与Token Ring 类似

100M的速率

采用4B5B编码方法

32中组合中的16种表示数据，3种表示定界符，2种表示控制，3种表示硬件信号，8种保留

最大距离200公里

最多1000个站点

通常作为连接LAN的主干网络

FDDI的双环操作

FDDI定义了两类站：A类站连接双环，B类站连接单环。

为提高信道利用率，站点发完数据后立即产生新令牌，环上可能同时存在多个帧；

4.6.2 DPT (Dynamic Packet Transport) (option)

Cisco 的技术，主要用于城域网(MAN)，非国际标准OC-12 (622 Mbps)，OC-48（2.488 Gbps）;

结合了IP带宽利用率高、服务种类丰富的特点和光纤环高带宽、自治愈的特点。

DPT环是双环.每个环都同时用于用户数据和控制数据的传输

Spatial Reuse Protocol (SRP): SRP是一个媒介无关的MAC层协议，用来实现DPT在光纤环情况上的功能。SRP提供基本的寻址，报文封装，带宽控制和控制信息的传输机制

目的地提取报文: 报文被目的节点从环上取下，不继续占用带宽。这样DPT环可以提供空间复用，使得多个不同网段可以同时全速使用带宽

DPT结合了SONET/SDH的处理能力和第二层的管理能力，来实现多层性能监视，错误检查和错误隔离功能

4.6.3 快速以太网（Fast Ethernet）

l 标准

. 1995年，IEEE通过802.3u标准，实际上是802.3的一个补充。原有的帧格式、接口、规程不变，只是将比特时间从100ns缩短为10ns。

l对10 Mbps 802.3 LAN的改进

. 一种方法是改进10Base-5 或10Base-2，采用CSMA/CD，最大电缆长度减为1/10，未被采纳；

. 另一种方法是改进10Base-T，使用HUB，被采纳

100BASE-T4

• 使用4 对UTP 3 类线或5 类线。

100BASE-TX

• 使用2 对UTP 5 类线或屏蔽双绞线STP。

100BASE-FX

• 使用2 对光纤。

100Base-T4:

使用4对ISO/IEC 11801定义的3、4、5类平衡双绞线; 3类非屏蔽双绞线（UTP），使用25MHz的信号（802.3使用20MHz的信号，由于使用Manchester编码，波特率=2\*比特率）；

4对双绞线，1对to the hub，1对from the hub，另外2对根据数据传输方向变换；

8B6T（8 bits map to 6 trits）编码，使用三进制信号（ternary signals），1对双绞线的比特率为25 \* 8/6 =33.3 Mbps，正向100M，反向33.3M。

100Base-TX

使用2对5类平衡双绞线或150W屏蔽平衡电缆，1对to the hub，1对from the hub，全双工；

5类双绞线使用125 MHz的信号；

4B5B编码，5个时钟周期发送4个比特，物理层与FDDI兼容，比特率为125 \* 4/5 = 100 Mbps；

100Base-FX

使用2根多模光纤，全双工

100Base-T4 和100Base-TX 统称100Base-T

两种类型的HUB

共享式HUB，一个冲突域，工作方式与802.3相同，CSMA/CD，二进制指数后退算法，半双工…

交换式HUB，输入帧被缓存，一个端口构成一个冲突域。

4.6.4 千兆以太网（Gigabit Ethernet） 标准：802.3ab

**本章小结:**

**介质访问控制**

**信道划分介质访问控制**

**随机访问介质访问控制**

**ALOHA协议；CSMA协议；CSMA/CD协议；CSMA/CA协议。**

**轮询访问介质访问控制：令牌传递协议**

**局域网**

**局域网的基本概念与体系结构**

**以太网与IEEE 802.3**

**IEEE 802.11**

**数据链路层设备**

**网桥：网桥的概念；透明网桥与生成树算饭；源选径网桥与源选径算法**

**局域网交换机及其工作原理。**

**Chapter 5 Network Layer**

**主要内容:**

**5.1 网络层的设计问题**

**5.2 路由选择算法**

**5.3 拥塞控制算法**

**5.4 服务质量**

**5.5 网络互联**

**5.6 因特网的网络层**

网络层概述:

网络层是处理端到端传输的最低层。

网络层要解决的关键问题是了解通信子网的拓扑结构，选择路由。

网络层设计的有关问题

为传输层提供服务

1)面向连接服务——传统电信的观点：通信子网应该提供可靠的、面向连接的服务。

2)无连接服务——Internet的观点：通信子网无论怎么设计都是不可靠的，因此网络层只需提供无连接服务。

3)IP/ATM

网络层为接在网络上的主机所提供的服务可以有两大类：1)无连接的网络服务（数据报服务datagram ）2) 面向连接的网络服务（虚电路服务virtual circuit ）。

**ISO 定义（网络层）**

网络层为一个网络连接的两个传送实体间交换网络服务数据单元提供功能和规程的方法，它使传送实体独立于路由选择和交换的方式。

广域网的基本概念

当主机之间的距离较远时，例如，相隔几十或几百公里，甚至几千公里，局域网显然就无法完成主机之间的通信任务。这时就需要另一种结构的网络，即广域网。

应当注意

即使是覆盖范围很广的互联网，也不是广域网，因为在这种网络中，不同网络的“互连”才是其最主要的特征。

广域网是单个的网络，它使用结点交换机连接各主机而不是用路由器连接各网络。

结点交换机在单个网络中转发分组，而路由器在多个网络构成的互联网中转发分组。

连接在一个广域网（或一个局域网）上的主机在该网内进行通信时，只需要使用其网络的物理地址即可。

广域网中的分组转发机制:

“转发”(forwarding)和“路由选择”(routing)这两个名词的使用在过去有些混乱。现在的文献倾向于将它们区分开来。

转发是当交换结点收到分组后，根据其目的地址查找转发表(forwarding table)，并找出应从结点的哪一个接口将该分组发送出去。

路由选择是构造路由表(routing table)的过程。

路由表是根据一定的路由选择算法得到的，而转发表又是根据路由表构造出的。“转发”和“路由选择”

“转发” (forwarding)就是路由器根据转发表将用户的IP 数据

报从合适的端口转发出去。

路由选择协议负责搜索分组从某个结点到目的结点的最佳传输路由，它是按照分布式算法，根据从各相邻路由器得到的关于网络拓扑的变化情况，动态地改变所选择的路由，以便构造路由表。

路由表是根据路由选择算法得出的，而转发表是从路由表得出的。从路由表构造出转发分组的转发表，分组是通过转发表进行转发的。

为了使讨论更简单些，可以不严格区分“转发”和“路由选择”，也不一定使用“转发表”这一名词。

在转发分组时可以不是说“查找转发表”而是说“查找路由表”。

局域网采用了平面地址结构(flat addressing)。对局域网，这种结构非常方便。

广域网中一般都采用层次地址结构(hierarchical addressing)。

路由器在网际互连中的作用:

路由器的构成

当主机A 要向另一个主机B 发送数据报时，先要检查目的主机B 是否与源主机A 连接在同一个网络上。

如果是，就将数据报直接交付给目的主机B 而不需要通过路由器。

但如果目的主机与源主机A 不是连接在同一个网络上，则应将数据报发送给本网络上的某个路由器，由该路由器按照转发表指出的路由将数据报转发给下一个路由器。这就叫作间接交付。

数据链路层剥去帧首部和尾部后，将分组送到网络层的队列中排队等待处理。这会产生一定的时延。

当交换结构传送过来的分组先进行缓存。数据链路层处理模块将分组加上链路层的首部和尾部，交给物理层后发送到外部线路。

分组丢弃

若路由器处理分组的速率赶不上分组进入队列的速率，则队列的存储空间最终必定减少到零，这就使后面再进入队列的分组由于没有存储空间而只能被丢弃。

路由器中的输入或输出队列产生溢出是造成分组丢失的重要原因。

5.1.2 为传输层提供的服务

服务目标: 服务与路由器技术无关; 路由器的数量、类型和拓扑结构是隐蔽的;传输层可用的网络地址统一编址

服务规范的两大阵容: Internet组织;电话公司

服务方式:面向连接/无连接;可靠/不可靠

5.1.5 虚电路与数据报网络

创建时间与地址解析时间

虚电路需要在建立连接时花费时间

数据报则在每次路由时过程复杂

路由器内存要求的表空间数量

虚电路方式，路由器需要维护虚电路的状态信息；

数据报方式，每个数据报都携带完整的目的/源地址，浪费带宽

可靠性与拥塞控制

虚电路方式很容易保证服务质量QoS(Quality of Service)，适用于实时操作，但比较脆弱。

数据报不太容易保证服务质量，但是对于通信线路的故障，适应性很强。

两种服务的思路来源不同:

虚电路服务的思路来源于传统的电信网。 电信网负责保证可靠通信的一切措施，因此电信网的结点交换机复杂而昂贵。

数据报服务力求使网络生存性好和使对网络的控制功能分散，因而只能要求网络提供尽最大努力的服务。可靠通信由用户终端中的软件（即TCP）来保证。

数据报服务和虚电路服务都各有一些优缺点:

网络上传送的报文长度，在很多情况下都很短。

用数据报既迅速又经济。

若用虚电路，为了传送一个分组而建立虚电路和释放虚电路就显得太浪费网络资源了。

在使用数据报时，每个分组必须携带完整的地址信息。

在使用虚电路的情况下，每个分组不需要携带完整的目的地址，而仅需要有个很简单的虚电路号码的标志。

这就使分组的控制信息部分的比特数减少，因而减少了额外开销。

在使用数据报时，主机承担端到端的差错控制和流量

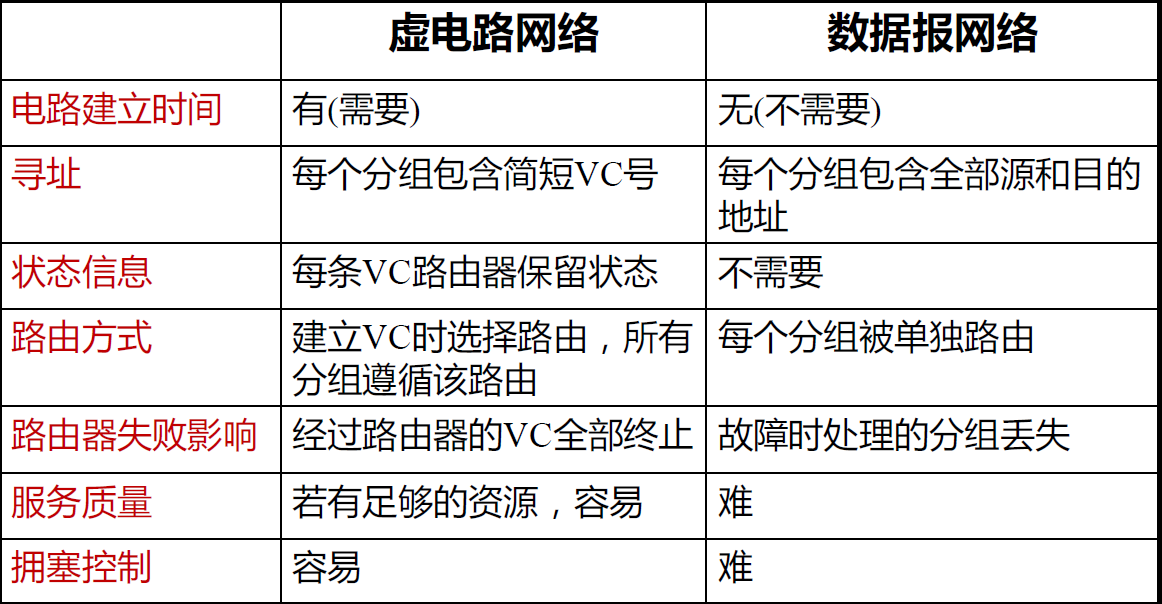
控制。

在使用虚电路时，分组按顺序交付，网络可以负责差错控制和流量控制。

数据报服务对军事通信有其特殊的意义。当某个结点发生故障时，后续的分组就可另选路由，因而提高了可靠性。

但在使用虚电路时，结点发生故障就必须重新建立另一条虚电路。

数据报服务还很适合于将一个分组发送到多个地址(即广播或多播)。



5.2 路由选择算法

优化原则;最短路径算法;泛洪算法;(基于流量的路由选择);距离矢量算法;链路状态路由;分级路由;广播路由;组播路由;选播路由;移动主机路由;自组织网络路由

基本原理: 路由算法是网络层软件的一部分

子网采用数据报方式，每个分组都要做路由选择；

子网采用虚电路方式，只需在建立连接时做一次路由选择。会话路由

路由算法应具有的特性:正确性（correctness）; 简单性（simplicity）;健壮性（robustness）;稳定性（stability）;公平性（fairness）;最优性（optimality）

路由算法分类: 非自适应算法（no adaptive algorithm）：静态路由算法;自适应算法（adaptive algorithm）：动态路由算法

路由与转发;公平与最优的冲突

最优化原则（optimality principle）

若I->J->K 为最佳路由，则其中的J ->K也是最佳路由

汇集树（sink tree）

从所有的源结点到一个给定的目的结点的最优路由的集合形成了一个以目的结点为根的树，称为汇集树；

路由算法的目的是找出并使用汇集树。

静态算法:最短路径算法:最短路径的含义; 泛洪算法:扩散法/选择性扩散法

最短路径（shortest path）算法基本思想:构建子网的拓扑图，图中的每个结点代表一个路由器，每条弧代表一条通信线路。为了选择两个路由器间的路由，算法在图中找出最短路径。

测量路径长度的方法:结点数量;地理距离;传输延迟; 距离、信道带宽等参数的加权函数

Dijkstra算法

1. 每个结点用从源结点沿已知最佳路径到本结点的距离来标注，标注分为临时 性标注和永久性标注；

2. 初始时，所有结点都为临时性标注，标注为无穷大；

3. 将源结点标注为0，且为永久性标注，并令其为工作结点；

4. 检查与工作结点相邻的临时性结点，若该结点到工作结点的距离与工作结点的标注之和小于该结点的标注，则用新计算得到的和重新标注该结点；

5. 在整个图中查找具有最小值的临时性标注结点，将其变为永久性结点，并成为下一轮检查的工作结点；

6. 重复第4、5步，直到目的结点成为工作结点；

洪泛算法 基本思想:向分组到来线路以外的所有线路发送

主要问题:洪泛产生大量重复分组

解决措施:分组头包含传递站点计数器;记录分组扩散路径

洪泛算法的改进:

选择性扩散法（selective flooding）:选择接近正确方向的线路发送

三应用情况:路由器和线路的资源过于浪费，实际很少直接采用；具有极好的健壮性，可用于军事应用；作为衡量标准评价其它路由算法。

动态算法:

距离矢量算法（distance vector routing）

每个路由器维持一张以子网中其他路由器为索引的路由选择表

表中每一项分组括应选择的输出线路及估计到达目标的时间或距离

链路状态路由（link state routing）

距离矢量算法属于动态路由算法，也称Bellman-Ford路由算法和Ford-

Fulkerson算法，最初用于ARPANET，被RIP协议采用。

基本思想

每个路由器维护一张表，表中给出了到每个目的地的已知最佳距离和

线路，并通过与**相邻路由器**交换距离信息来更新表；

以子网中其它路由器为表的索引，表项包括两部分：到达目的结点的

最佳输出线路，和到达目的结点所需时间或距离；

每隔一段时间，路由器向所有邻居结点发送它到每个目的结点的距离

表，同时它也接收每个邻居结点发来的距离表；

邻居结点X发来的表中，X到路由器i的距离为Xi，本路由器到X的距离为m，则路由器经过X到i的距离为Xi + m。根据不同邻居发来的信息，计算Xi + m，并取最小值，更新本路由器的路由表；

注意：本路由器中的老路由表在计算中不被使用。

5.2.6 链路状态路由

距离向量路由算法的主要问题:选择路由时，没有考虑线路带宽； 路由收敛速度慢。

链路状态路由：OSPF、IS-IS

1. 发现邻居节点，了解其网络地址

2. 测量线路开销，到邻居节点的距离或成本度量值

3. 组装链路状态分组

4. 发送链路状态分组，接收其他所有链路状态分组

5. 计算最短路径

6. 完整的拓扑结构和所有的延时都被测量并被测量到各个路由器中，各个路由器可以用Dijkstra算法来找出最短路径。

发现邻居节点:

路由器启动后，通过发送HELLO分组发现邻居结点；

两个或多个路由器连在一个LAN时，引入人工结点；

设置链路成本:

为寻找最短路径，需要每条链路的距离或成本

与链路带宽成反比，1G的成本：1，100M的成本：10

地理上分散的线路延迟估算：往返时间除2: 一种直接的方法是：发送一个要对方立即响应的ECHO分组，来回时间除以2即为延迟。

构造链路状态分组:

分组以发送方的标识符开头，后面是序号、年龄和一个邻居结点列表；

列表中对应每个邻居结点，都有发送方到它们的延迟或开销；

链路状态分组定期创建或发生重大事件时创建:周期性创建;事件驱动：线路断掉、邻居结点down机、重新恢复运行

泛洪法发布: 基本思想: 洪泛链路状态分组，为控制洪泛，每个分组包含一个序号，每次发送新分组时加1。

路由器记录信息对：<源路由器，序号>，当一个链路状态分组到达时，若是新的，则分发；若是重复的，则丢弃；若序号比路由器记录中的最大序号小，则认为过时而丢弃；

存在问题: 序号绕回; 路由器崩溃，序号重新开始，后面的分组会被当作重复分组被拒绝; 序号传送出错

解决方法： 使用32位序号；232 < 137′365′24′60′60， 137年才循环一次; 在每个分组中增加一年龄字段（如初值为60）;每秒减1，路由器减1;为0后该分组将被丢弃，否则不会被认为是过时分组

被广泛应用于实际网络中的链路状态路由协议:IS-IS（DECnet），OSPF（IETF）

网络规模增长带来的问题:路由器中的路由表增大；路由器为选择路由而占用的内存、CPU时间和网络带宽增大。

分层路由

分而治之的思想；

根据需要，将路由器分成域（regions）、簇（clusters）、区（zones）和组（groups）… Fig. 5-14，路由表由17项减为7项。

分层路由带来的问题: 路由表中的路由不一定是最优路由。

最优级数（Kamoun & Kleinrock 1979）:*N*个路由器最优级数为：*lnN*;每个路由器表项数为：*elnN*

广播路由: 广播: 将分组发往所有目的地

广播路由选择算法

1. 重复一一每一个目的地单独发送

2. 多目的地路由选择:分组含有按发送线路分组的目的地清单

3. 泛洪法

4. 逆向路径转发:广播分组来自发往广播源的线路

5. 利用源路由器的生成树:生成树：包括通信子网的所有路由器，但不包含回路

逆向路径传送: 基本原理：当某一广播分组到达路由器时，路由器对它进行检查，如该分组来自通常向广播源发送分组的线路，则将该分组转发到除进线以外的其它线路，否则丢弃

5.2.9 组播（Multicast）路由

组播: 小组中广播;

网络实例: b：汇集树(10);c、d：组播树

算法: 修剪生成树:网络有n个组，每组平均m个成员，共存n′m棵树

核心基本树: 每个组只计算一棵生成树，根成员处于组的中心部位

生成树法: 组中的每个成员都必须以自己作为出发点建立一棵生成树，根据自己所在小组对生成树进行修剪，得到一棵本小组修剪过的生成树，并告诉同组的其他成员

多址传输时仅沿相应的生成树转发: 缺点：存储量大,如一个网络有n个组，每个组平均有m个成员，则要存储m\*n棵生成树

核心树(core-based tree): 每个组只有一棵生成树，它的树根靠近组的中心部位，要发送一个分组给这个组成员，则发给树根，由树根将该分组沿核心树发往

各成员，这样，每组只有一棵核心树，节约了存储空间

选播（Anycast）路由:

数据分组传递模型: 单播、广播、组播

选播: 数据分组传递给最近的一个组成员;目的是获得正确的信息; 报时或内容分发，多信息源, 域名系统

选播路由: 利用普通的距离矢量和链路状态路由

移动主机路由: 需要解决的问题,为了能够将数据分组转发给移动主机，网络必须首先要找到移动的主机。

用户分类: 静态用户;移动用户（mobile users）：包括位置发生变化，通过固定方式或移动方式与网络连接的两类用户:迁移用户;漫游用户

主位置(home location)

主地址（home location）：所有用户都有一个永久的主位置，用一个地址来标识；

主代理（home agent）：每个区域有一个主代理，负责记录主地址在该区域，但是目前正在访问其它区域的用户；

外地代理（foreign agent）：每个区域（一个LAN或一个wireless cell）有一个或多个外地代理，它们记录正在访问该区域的移动用户。

外地代理工作步骤:

移动用户进入一个新区域时，必须首先向外部代理注册

1. 外地代理广播: 外地代理定期广播声明自己的存在和地址的分组，新到达的移动主机接收该信息；

若移动用户未能收到该信息:则移动主机广播分组，询问外地代理的地址；

2. 移动主机申请登录: 移动主机向外地代理注册，告知其主地址、目前的数据链路层地址和一些安全信息；

3. 外地代理向主代理请求确认: 外地代理与移动主机的主代理联系;告知移动主机的目前位置、自己的网络地址和一些安全信息；

4. 主代理回答确认: 主代理检查安全信息，通过，则给外地代理确认；

5. 外地代理通知移动主机:外地代理收到确认后，在登记表中加入一项，并通知移动主机注册成功。

移动用户的路由转发过程:

当一个分组发给移动用户时，首先被转发到用户的主地址局域网；

该分组到达用户的主地址局域网后，被主代理接收，主代理查询移动用户的新位置和与其对应的外地代理的地址；

主代理采用隧道技术，将收到的分组作为净荷封装到一个新分组中，发给外地代理；

主代理告诉发送方，发给移动用户的后续分组作为净荷封装成分组直接发给外地代理；

外地代理收到分组后，将净荷作为数据链路帧发给移动用户；

拥塞控制算法: 拥塞控制基础;拥塞控制的途径;流量感知路由; 准入控制; 流量调节; 负载脱落

拥塞控制基础

拥塞:通信子网中有太多的分组时性能降低，这种情况称为**拥塞。**

拥塞产生的原因:多个输入对应一个输出；慢速处理器；低带宽线路。

解决办法: 针对某个因素的解决方案，只能对提高网络性能起到一点点好处，甚至可能仅仅是转移了影响性能的瓶颈；需要全面考虑各个因素。

拥塞崩溃: 路由器内存：队列溢出与无限内存; 网络瓶颈：线路带宽与低效路由器;拥塞与流量控制;

拥塞控制（congestion control）; 拥塞控制需要确保通信子网能够承载用户提交的通信量，是一个全局性问题，涉及主机、路由器等很多因素；

流量控制（flow control）: 流量控制与点到点的通信量有关，主要解决快速发送

方与慢速接收方的问题，是局部问题，一般都是基于反馈进行控制的。

流量感知路由: 忽略负载，只考虑带宽和传输延迟;

考虑负载:多路径路由;流量缓慢迁移

准入控制:

许可控制: 拥塞后不再建立任何虚电路

变通: 新的虚电路绕过有问题的区域;在删除拥塞后的通信子网上建立虚电路

资源预留:建立虚电路时，主机与子网达成协议，子网根据协议在虚电路上为此连接预留资源。

流量调节:

拥塞避免:监视何时何地发生拥塞: 缺少缓冲区空间而丢失分组的比例;平均队列长度;超时和重发分组的数量; 平均分组延时、标准方差分组延时

将信息传递给造成拥塞的发送方: 拥塞路由器向信流源发送一个分组;分组保留拥塞位或字段;证实分组询问拥塞;

抑制分组: 主机调整策略参数;数据率减半

抑制分组（Choke Packets）:

基本思想1)路由器监控输出线路及其它资源的利用情况，超过某个阈值，则此资

源进入警戒状态；

每个新分组到来，检查它的输出线路是否处于警戒状态；

若是，则向源主机发送抑制分组，分组中指出发生拥塞的目的地址。同时将原分组打上标记（为了以后不再产生抑制分组），正常转发；

源主机收到抑制分组后，按一定比例减少发向特定目的地的流量，并在固定时间间隔内忽略指示同一目的地的抑制分组。然后开始监听，若此线路仍然拥塞，则主机在固定时间内减轻负载、忽略抑制分组；若在监听周期内没有收到抑制分组，则增加负载；

通常采用的流量增减策略是：减少时，按一定比例减少，保证快速解除拥塞；增加时，以常量增加，防止很快导致拥塞。

逐跳抑制分组（Hop-by-Hop Choke Packets）

在高速、长距离的网络中，由于源主机响应太慢，抑制分组算法对拥塞控制的效果并不好，可采用逐跳抑制分组算法；

基本思想

抑制分组对它经过的每个路由器都起作用；能够迅速缓解发生拥塞处的拥塞；上游路由器要求有更多的缓冲区。

调节方法：显示拥塞通知：警告位，应答通知 ； 逐跳抑制

负载脱落：

载荷脱落：葡萄酒策略：文件传输； 牛奶策略：多媒体

优先级：价格，低优先级便宜；通信量整形：空闲时低优先级先发送；虚电路带宽：可超过虚电路建立时确定的极限值，但标为低优先级；随机早期检测：队列超过阀值，随机丢分组 ，隐形通知

流量整形(traffic shaping)：

基本思想：造成拥塞的主要原因是网络流量通常是突发性的；强迫分组以一种可预测的速率发送；在ATM网中广泛使用。

漏桶算法（The Leaky Bucket Algorithm）

将用户发出的不平滑的数据分组流转变成网络中平滑的数据分组流；

可用于固定分组长的协议，如ATM；也可用于可变分组长的协议，如IP，使用字节计数；

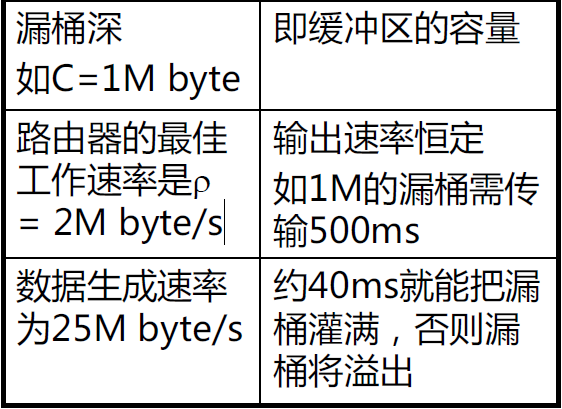
无论负载突发性如何，漏桶算法强迫输出按平均速率进行，不灵活。

平滑输入流量，控制进入网络的流量

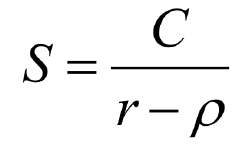
漏桶算法中的主机与网络：

在主机上生成的发送数据具有突发性和间歇性

通过漏桶算法，将使进入网络的数据通信量变得平稳



设C为漏桶容量（Byte, cell数），r为漏桶的输出速率（Byte/s）

则突发速率r与允许突发时间的长度S的关系为：

上例中突发长度的计算允许突发时间的长度S =C/(r-r )=1/(25-2)= 43.5(ms)

漏桶算法的不足之处：漏桶满时，造成数据丢失；漏桶算法强迫输出保持一个固定的平均速率，不能体现通信量的突发

漏桶方法的改进：

增加排队缓冲区，提高系统的吞吐量(增加桶的容量）

许可证控制（令牌桶）

跳跃式窗口：通过反馈通知主机缩小发送窗口的大小

信元标注法：将违约信元加标识，进入网络，如网络较空，则通过，否则丢弃

信元按优先级划分：若网络拥挤，先丢掉优先级低的信元

令牌桶算法（The Token Bucket Algorithm）

漏桶算法不够灵活，因此加入令牌机制；

基本思想：漏桶存放令牌，每DT秒产生一个令牌，令牌累积到超过漏桶上界时就不再增加。包传输之前必须获得一个令牌，传输之后删除该令牌；

漏桶算法与令牌桶算法的区别

流量整形策略不同：漏桶算法不允许空闲主机积累发送权，以便以后发送大的突发数据；令牌桶算法允许，最大为桶的大小。

漏桶中存放的是数据包，桶满了丢弃数据包；令牌桶中存放的是令牌，桶满了丢弃令牌，不丢弃数据包。

令牌桶算法(Token bucket Algorithm)

希望当大的突发流量到来时，输出也能有适当响应；

数据的输出必须持有令牌；令牌的产生是定时的，每隔Δt，产生一个令牌，如令牌没有被及时使用，可以存在令牌桶内，令牌桶满则将被丢弃

当突发数据到达时，如令牌桶内有多个令牌，则突发数据可按令牌数输出

令牌桶突发时间长度的计算:

如：C：为令牌桶的容量;r：为令牌到达速率;M：为最大的输出速率（数据到达速率> M）;则最大的突发时间S为：

C + rS = MS 即S = C/(M - r)

当：令牌桶的容量C = 250K byte,令牌到达速率r = 2M byte/s,最大的输出速率M = 25M byte/s时，如令牌桶已满，则S = 250K/(25M - 2M) (ms) = 250/23 (ms) 约为11 ms

令牌桶算法

参数:突发时间：S秒;令牌桶容量：B字节;令牌到达速率：R字节/秒;最大输出速率：M字节/秒

最大速率发送突发持续时间: B + RS = MS; S = B/(M - R)

漏桶与令牌桶示例:

计算机和网络、路由器最大速率:1000Mbps/8 = 125MBps

路由器工作速率：200Mbps/8 = 25MBps

数据突发输入量：

以最大速率发送16000KB的突发数据:

16000KB/125MBps =16000/(125\*1.024)= 125ms

以25MBps发送6400KB的突发数据

6400KB/25MBps = 6400/(25\*1.024)=250ms

令牌桶：

B=9600KB，R=200Mbps

S = B/(M-R) = 9600KB/(125-25)MBps = 93.75ms

5.5 网络互联:

互连在一起的网络要进行通信，会遇到许多问题需要解决，如：不同的寻址方案;不同的最大分组长度;不同的网络接入机制;不同的超时控制;不同的差错恢复方法;不同的状态报告方法;不同的路由选择技术;不同的用户接入控制;不同的服务（面向连接服务和无连接服务）;不同的管理与控制方式

网络互连使用路由器

当中继系统是转发器或网桥时，一般并不称之为网络互连，因为这仅仅是把一个网络扩大了，而这仍然是一个网络。

网关由于比较复杂，目前使用得较少。

互联网都是指用路由器进行互连的网络。

由于历史的原因，许多有关TCP/IP 的文献将网络层使用的路由器称为网关。

虚拟互连网络的意义:

所谓虚拟互连网络也就是逻辑互连网络，它的意思就是互连起来的各种物理网络的异构性本来是客观存在的，但是我们利用IP 协议就可以使这些性能各异的网络从用户看起来好像是一个统一的网络。

使用IP 协议的虚拟互连网络可简称为IP 网。

使用虚拟互连网络的好处是：当互联网上的主机进行通信时，就好像在一个网络上通信一样，而看不见互连的各具体的网络异构细节。

网络互连设备:

中继器repeater（物理层）:放大或再生弱信号，两个电缆段间复制比特

网桥bridge（数据链路层）:存储-转发设备，在LAN之间存储转发数据链路帧

路由器router（网络层）:在异型网络间转发分组

网关gateway（传输层以上）:在传输层连接两个网络，或在应用层连接两部分应用程序

网络的连接: 两个以太网的连接

1) 交换式连接：不识别分组， MAC

2) 路由连接：路由器从帧中提取分组，分组地址

级联虚电路（Concatenated Virtual Circuits）:

工作过程:级联虚电路工作过程与虚电路子网工作过程相似；

建立连接:当目的主机不在子网内时，则在子网内找一个离目的网络最近的路

由器，与之建立一条虚电路；该路由器与外部网关建立虚电路；该网关与下一个子网中的一个路由器建立虚电路；重复上述操作，直到到达目的主机。

传输数据：相同连接的包沿同一虚电路按序号传输；网关根据需要转换分组格式和虚电路号。

拆除连接

无连接网络互连（Connectionless Internetworking）：

工作过程：无连接网络互连的工作过程与数据报子网的工作过程相似；每个包单独路由，提高网络利用率，但不能保证包按顺序到达；根据需要，连接不同子网的多协议路由器做协议转换，包括包格式转换和地址转换等。

级联虚电路与无连接网络互连的比较：

级联虚电路的优点

路由器预留缓冲区等资源，保证服务质量；包按序号传输；短分组头。

级联虚电路的缺点：路由器需要大量内存，存储虚电路信息；一旦发生拥塞，没有其它路由；健壮性差；如果网络中有一个不可靠的数据报子网，级连虚电路

很难实现。

无连接网络互连的优点：能够容忍拥塞，并能适应拥塞；健壮性好；可用于多种网络互连。

无连接网络互连的缺点：长分组头；分组不能保证按序号到达；不能保证服务质量。

隧道技术：源和目的主机所在网络类型相同，连接它们的是一个不同类型的网络，这种情况下可以采用隧道技术。

分段（Fragmentation）

每种网络都对最大分组长度有限制，有以下原因：硬件，例如TDM 的时槽限制； 操作系统；协议，例如分组长度域的比特个数；与标准的兼容性；希望减少传输出错的概率；希望避免一个分组占用信道时间过长。

大分组经过小分组网络时，网关要将大包分成若干段（fragment），每段作为独立的分组传输。

解决网络最大长度限制：

透明分段(ATM网采用策略)：网关将大分组分段后，每段都要经过同一出口网关，并在那里重组；

不透明分段(IP网采用策略)：中间网关不做重组，而由目的主机做。

分段重组过程对其它网络透明：例如ATM网络；

带来的问题： 出口网关需要知道何时所有分组都到齐；所有分组必须从同一出口网关离开；大分组经过一系列小分组网络时，需要反复地分段重组，开销大。

分段重组过程对其它网络不透明：

带来的问题：对主机要求高，能够重组；每个段都要有一个分组头，网络开销增大。

分段方法

标记段

树型标记法：例，分组0分成三段，分别标记为0.0, 0.1, 0.2，段0.0构

成的分组被分成三段，分别标记为0.0.0, 0.0.1, 0.0.2；

存在的问题：段标记域要足够长；分段长度前后要一致

偏移量法：定义一个基本段长度，使得基本段能够通过所有网络；分组分段时，除最后一个段小于等于基本段长度外，所有段长度都等于基本段长度；一个分组可以包括几个段，分组头中包括：原始分组序号，分组中第一个基本段的偏移量，最后段指示位。

基本分段长度：(a)包含10个字节的分组；(b)经过最大分组长度为8的网络后；(c)经过最大分组长度为5的网络后

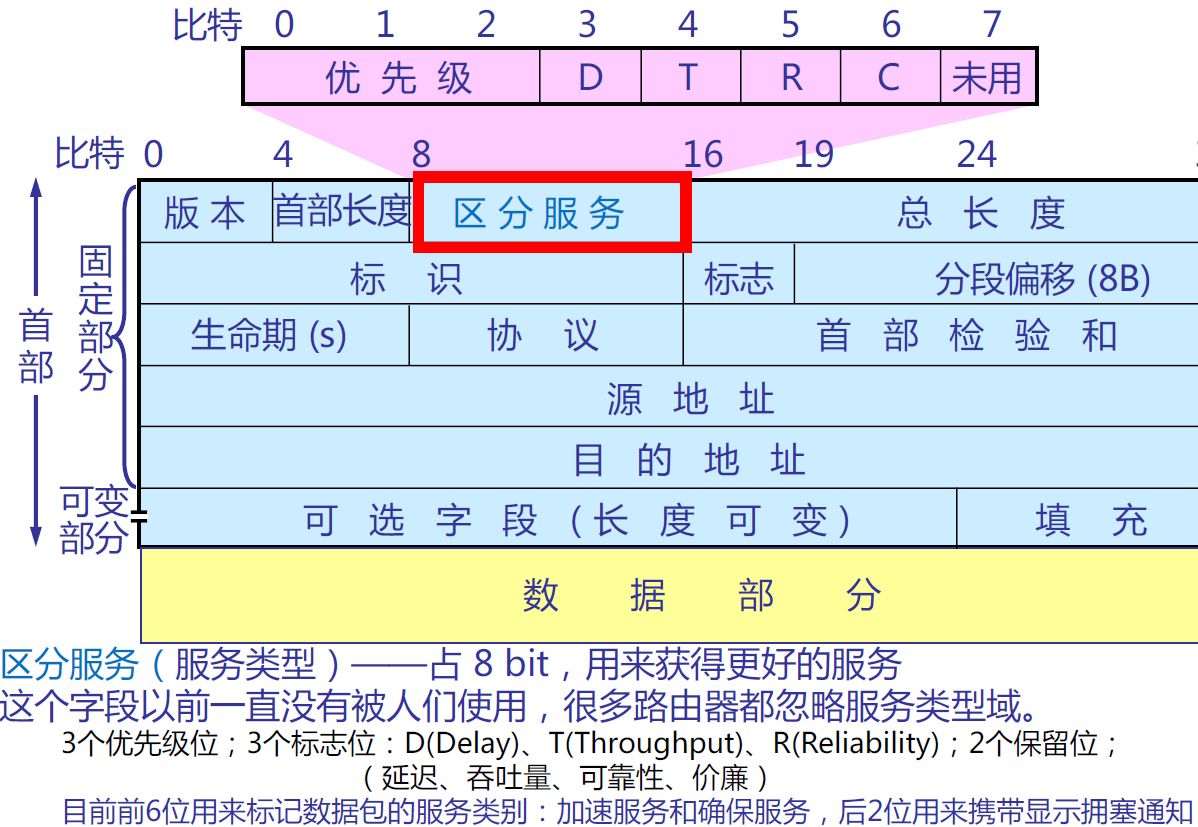
5.6 因特网上的网络层

IPv4协议， IP地址，IPv6协议，因特网控制协议，标签交换与MPLS， OSPF——内部网关路由协议，BGP——外部网关路由协议，因特网组播，移动IP

IPv4协议：一个IP 数据报由首部和数据两部分组成

首部的前一部分是固定长度，共20 字节，是所有IP 数据报必须具有的。

在首部的固定部分的后面是一些可选字段，其长度是可变的



总长度——占16 bit，指首部和数据之和的长度，单位为字节，因此数据报的最大长度为65535 字节。总长度必须不超过最大传送单元MTU。

标识(identification) 占16 bit，它是一个计数器，用来产生数据报的标识。确定新到达的分段属于哪个数据报

标志位占3 bit，第1位未用，第2位DF代表不要分段，第3位MF代表还有分段。

分段偏移(13bit)指出：较长的分组在分段后，某分段在原分组中的相对位置；分段偏移以8 个字节为偏移单位。

生命期(8 bit)记为TTL (Time To Live)数据报在网络中的寿命，其单位为秒。

协议(8 bit)字段指出此数据报携带的数据使用何种协议以便目的主机的IP 层将数据部分上交给哪个处理过程

首部检验和(16 bit)字段只检验数据报的首部，不包括数据部分。这里不采用CRC 检验码而采用简单的计算方法。

二进制反码求和：

0和0相加是0，0和1相加是1，1和1相加是0但要产生一个进位1，加到下一列．若最高位相加后产生进位，则最后得到的结果要加上溢出的进位1(可能是多個1).

IP/ICMP/IGMP/TCP/UDP等协议的校验和算法都是相同的，发送数据时，为了计算IP数据包的校验和，按如下步骤：

（1）把IP数据包的校验和字段置为0；（2）把首部看成以16位为单位的数字组成，依次进行二进制反码求和；（3）把得到的结果存入校验和字段中。

在接收数据时，计算数据包的校验和相对简单，按如下步骤：

（1）把首部看成以16位为单位的数字组成，依次进行二进制反码求和，包括校验和字段；（2）检查计算出的校验和的结果是否等于零（反码应为16个0）；（3）如果等于零，说明被整除，校验是和正确。否则，校验和就是错误的，协议栈要抛弃这个数据包。

计算对IP首部检验和的算法如下：

（1）把IP数据包的校验和字段置为0；（2）把首部看成以16位为单位的数字组成，依次进行二进制求和（注意：求和时应将最高位的进位保存，所以加法应采用32位加法）；（3）将上述加法过程中产生的进位（最高位的进位）加到低16位（采用32位加法时，即为将高16位与低16位相加，之后还要把该次加法最高位产生的进位加到低16位）（4）将上述的和取反，即得到校验和。

5.6.2 IP地址

前缀,子网,CIDR——无类域间路由选择,分类和特殊寻址,NAT—网络地址转换

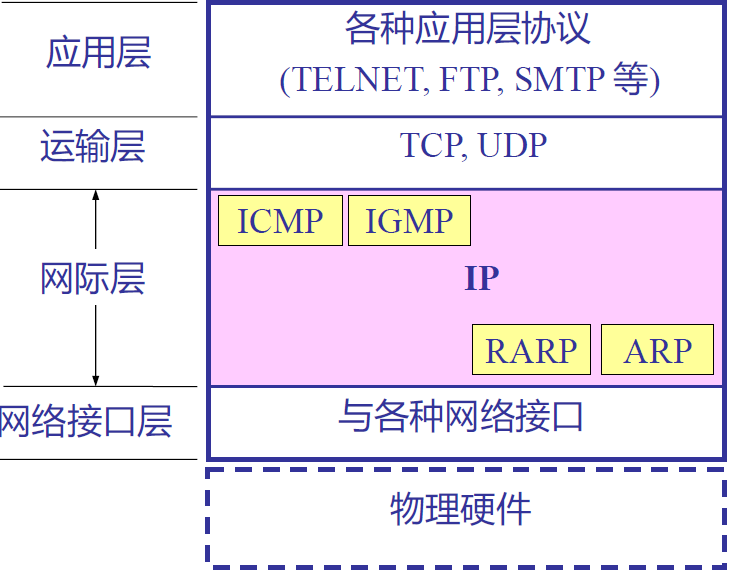
与IP 协议配套使用的四个协议:

地址解析协议ARP(Address Resolution Protocol)

逆地址解析协议RARP(Reverse Address Resolution Protocol)

因特网控制报文协议ICMP(Internet Control Message Protocol)

因特网组管理协议IGMP(Internet Group Management Protocol)



我们把整个因特网看成为一个单一的、抽象的网络。IP 地址就是给每个连接在因特网上的主机（或路由器）分配一个在全世界范围是惟一的32 bit 的标识

符。

IP 地址现在由因特网名字与号码指派公司ICANN(Internet Corporation for Assigned Names and Numbers)进行分配

IP 地址的编址方法:

分类的IP 地址:这是最基本的编址方法，在1981年就通过了相应的标准协

议。

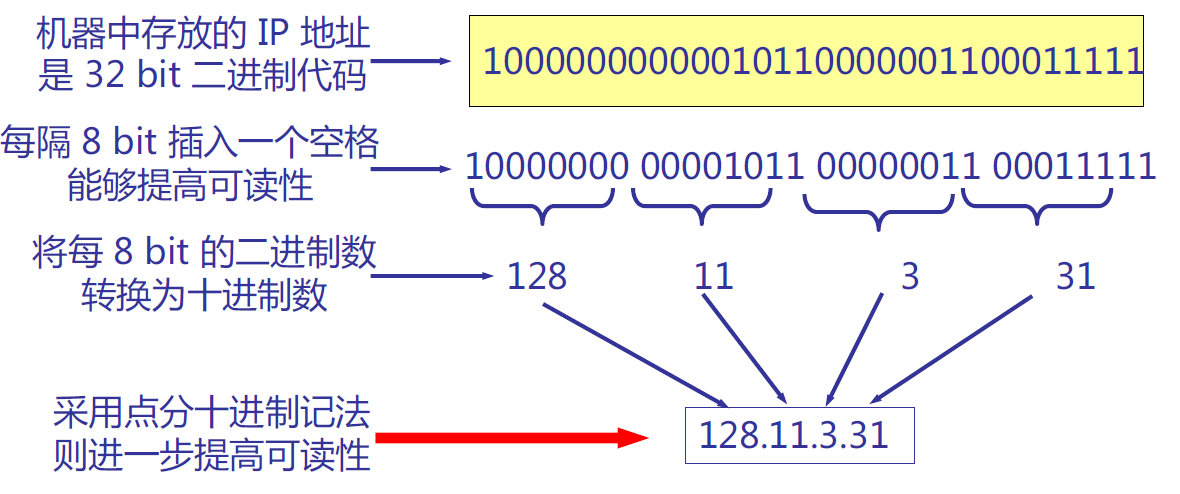
子网的划分: 这是对最基本的编址方法的改进，其标准[RFC 950]在1985年通过。

无类域间路由:这是比较新的无类编址方法。1993年提出后很快就得到推广应用。

前缀:

IPv4地址与书写方式:32位二进制：80D00297,1000 0000 1101 0000 0000 0010 1001 0111;点分十进制：128.208.2.151

子网掩码: 255.255.255.0 ;/24



IP 地址的一些重要特点

1.IP 地址是一种分等级的地址结构，分级的好处：

IP地址管理机构只分配IP地址网络号，主机号则由得到该网络号的单位自行分配。这样就方便了IP 地址的管理。

路由器仅根据目的主机所连接的网络号来转发分组，这样就可以使路由表中的项目数大幅度减少，从而减小了路由表所占的存储空间。

2. IP 地址是标志一个主机（或路由器）和一条链路的接口

当一个主机同时连接到两个网络上时，该主机就必须同时具有两个相应的IP 地址，其网络号net-id 必须是不同的。这种主机称为多接口主机(multihomed host)。

由于一个路由器至少应当连接到两个网络（这样它才能将IP 数据报从一个网络转发到另一个网络），因此一个路由器至少应当有两个不同的IP 地址。

3. 用转发器或网桥连接起来的若干个局域网仍为一个网络，因此这些局域网都具有同样的网络号net-id

4. 所有分配到网络号net-id 的网络，无论范围很小的局域网，还是可能覆盖很大地理范围的广域网，都是平等的

网中的IP 地址

互联网:

在同一个局域网上的主机或路由器的IP 地址中的网络号必须是一样的;

路由器总是具有两个或两个以上的IP 地址。路由器的每一个接口都有一个

不同网络号的IP 地址。

两个路由器直接相连的接口处，可指明也可不指明IP地址。如指明IP 地址，则这一段连线就构成了一种只包含一段线路的特殊“网络” 。现在常不指明IP 地址

子网

子网由来:英特网域名和地址分配机构（ICANN）;ICANN将地址空间授权给区域机构;区域机构再将IP地址分配给ISP和公司 例如：分配“/16”大小的网络

子网划分:均匀、按需（非均匀）

子网划分结构

划分子网的基本思路:

划分子网纯属一个单位内部的事情。单位对外仍然表现为没有划分子网的网络。

从主机号借用若干个比特作为子网号subnet-id，而主机号host-id 也就相应减少了若干个比特。IP地址::= {<网络号>, <子网号>, <主机号>}

凡是从其他网络发送给本单位某个主机的IP 数据报，仍然是根据IP 数据报的目的网络号net-id，先找到连接在本单位网络上的路由器。

然后此路由器在收到IP 数据报后，再按目的网络号net-id 和子网号subnet-id 找到目的子网。

最后就将IP 数据报直接交付给目的主机。

划分子网后变成了三级结构:

当没有划分子网时，IP 地址是两级结构，地址的网络号字段也就是IP 地址的“因特网部分”，而主机号字段是IP 地址的“本地部分”。

划分子网后IP 地址就变成了三级结构。划分子网只是将IP 地址的本地部分进行再划分，而不改变IP 地址的因特网部分。

子网掩码:

从一个IP数据报的首部并无法判断源主机或目的主机所连接的网络是否进行了子网的划分。

使用子网掩码(subnet mask)可以找出IP 地址中的子网部分。

常用的三种类别的IP 地址:

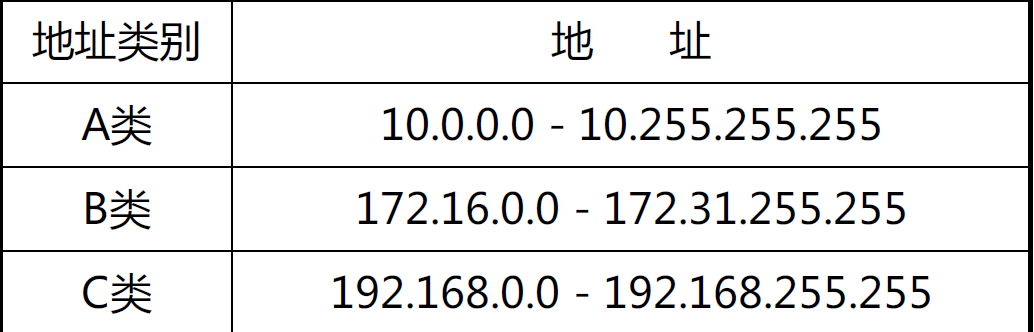
IP 地址的使用范围



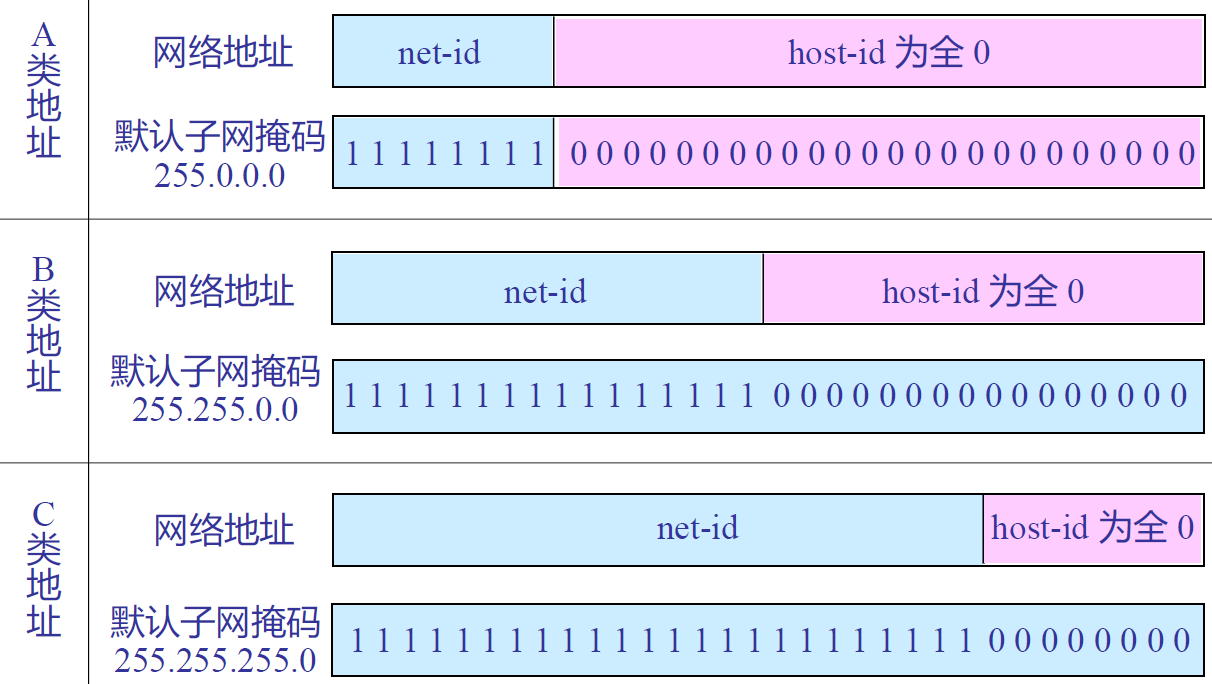
分类和特殊寻址:



IANA（Internet Assigned Number Authority）保留给私有网络的IP地址段



A类B类和C类IP地址的默认子网掩码

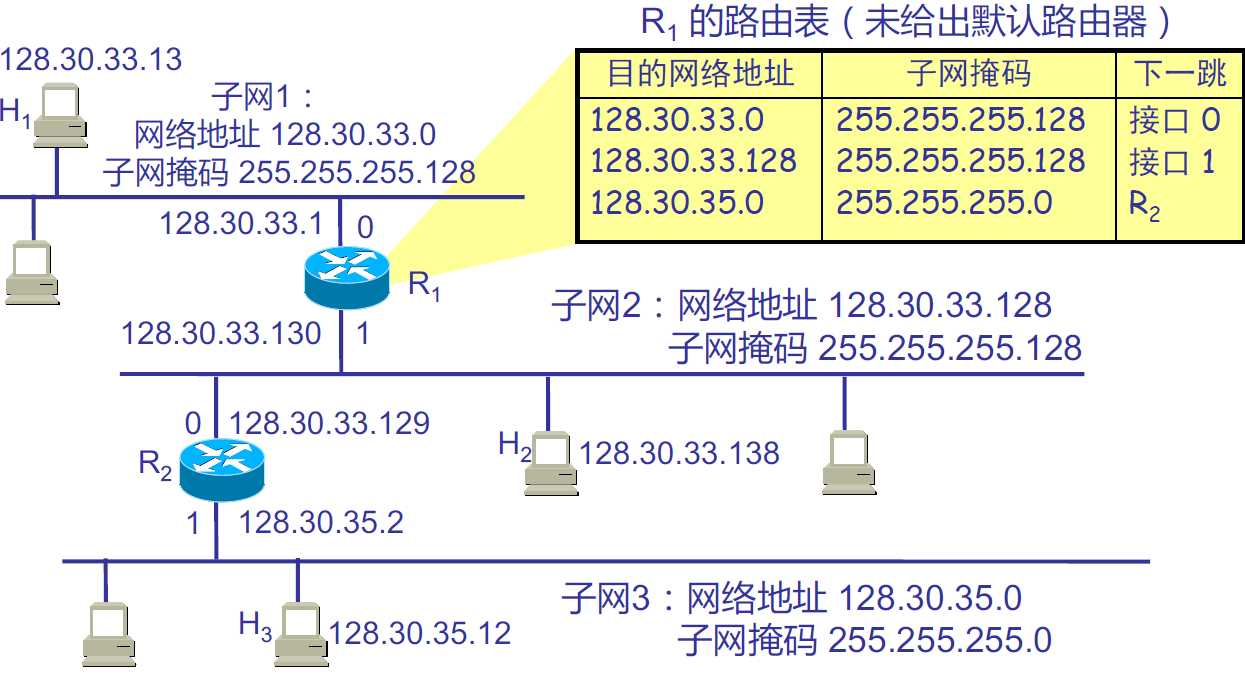


使用子网掩码的分组转发过程:

在不划分子网的两级IP 地址下，从IP 地址得出网络地址是个很简单的事。

但在划分子网的情况下，从IP地址却不能惟一地得出网络地址来，这是因为网络地址取决于那个网络所采用的子网掩码，但数据报的首部并没有提供子网掩码的信息。

因此分组转发的算法也必须做相应的改动。



路由器转发分组的步骤:

先按所要找的IP 地址中的网络号net-id 把目的网络找到。

当分组到达目的网络后，再利用主机号host-id 将数据报直接交付给目的主机。

按照整数字节划分net-id 字段和host-id 字段，就可以使路由器在收到一个分组时能够更快地将地址中的网络号提取出来。

CIDR——无类域间路由

网络数:一个组织：子网路由表项，外网缺省路由;

ISP与骨干网：网络数目100万以上，路由表爆炸

CIDR 路由聚合

一个IP地址，可以对应于/22网络路由，也可以聚合为/20网络路由

地址分块

194.0.0.0~195.255.255.255：欧洲(33,554,432个地址)

194.0.0.xx~195.255.255.xx : 131,072

198.0.0.0~199.255.255.255：北美

200.0.0.0~201.255.255.255：中美和南美

202.0.0.0~203.255.255.255：亚洲和太平洋地区

204.0.0.0~223.255.255.255：保留(335,544,320个地址)

路由聚合(route aggregation):

一个CIDR 地址块可以表示很多地址，这种地址的聚合常称为路由聚合，它使得路由表中的一个项目可以表示很多个（例如上千个）原来传统分类地址的路由。

CIDR虽然不使用子网了，但仍然使用“掩码”这一名词（但不叫子网掩码）。

最长前缀匹配

使用CIDR 时，路由表中的每个项目由“网络前缀”和“下一跳地址”组成。在查找路由表时可能会得到不止一个匹配结果。

应当从匹配结果中选择具有最长网络前缀的路由：最长前缀匹配(longest-prefix matching)。

网络前缀越长，其地址块就越小，因而路由就越具体。

最长前缀匹配又称为最长匹配或最佳匹配。

5.6.3 IPv6协议

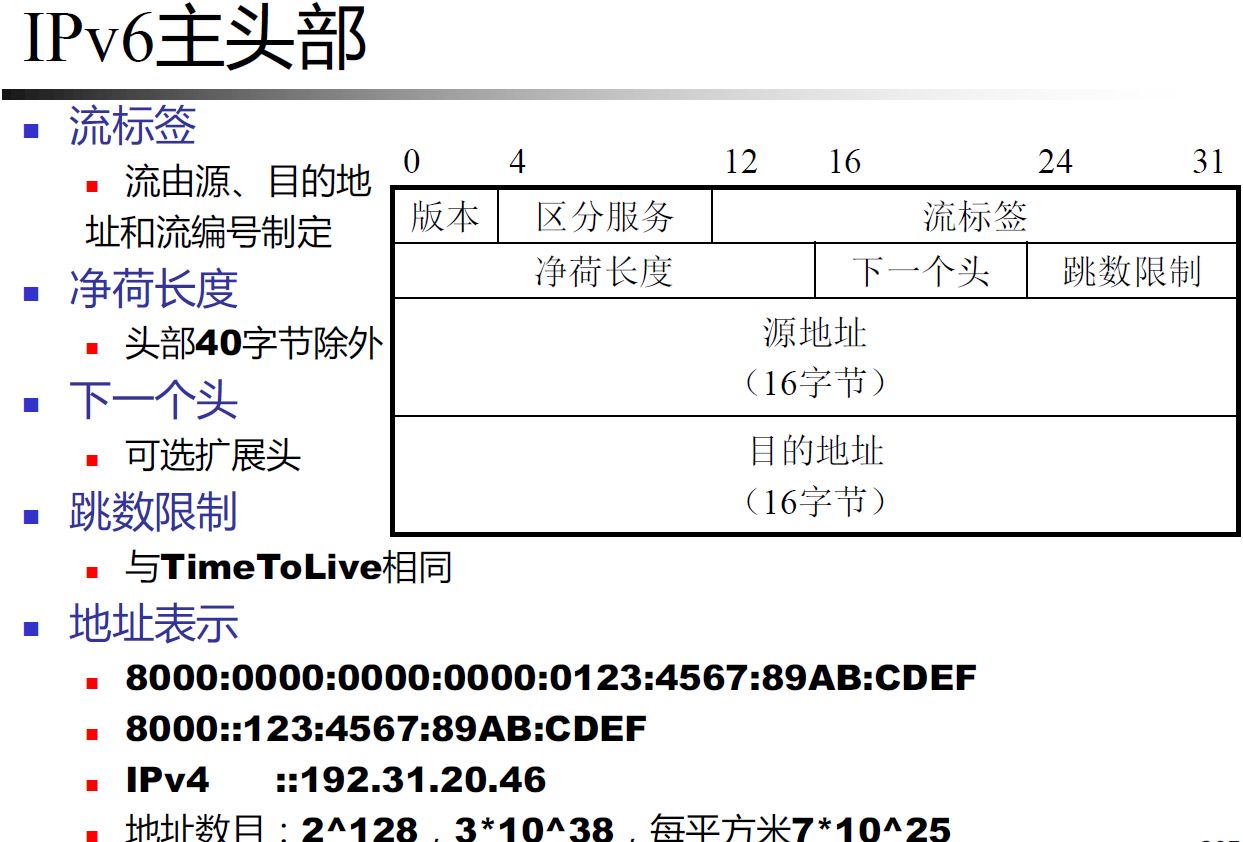
IP新版本的目标

远超几十亿台主机(IPv4地址4,294,967,296); 降低路由表大小;简化协议，加快路由器处理速度;安全性：身份认证与隐私权;关注服务类型，特别是实时数据;组播、漫游;协议演进、共存

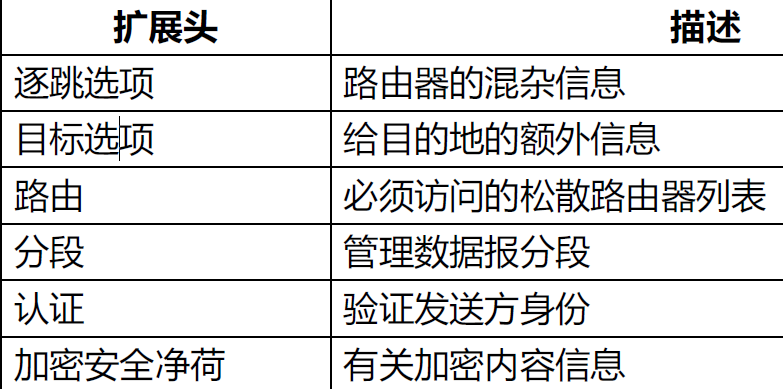
IPv6的改进

16字节地址;头部简化：由IPv4的13个字段变为7个;选项处理：必需字段变

为选项提高效率; 安全性：身份认证与隐私权; 区分服务：最初为流量类型，最低两位用于拥塞



扩展头部



三元组：

<Type，Length，Value>; 逐跳头; 路由扩展头

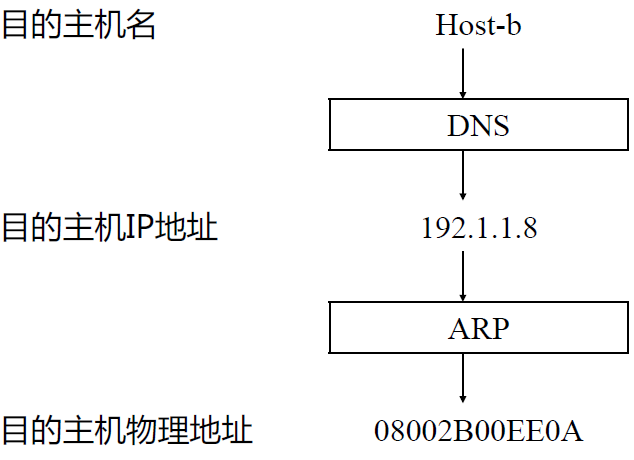
5.6.4 因特网控制协议

ICMP因特网控制报文协议;ARP地址解析协议;DHCP动态主机配置协议

ICMP因特网控制报文协议

可将路由器意外等事件报告给源端;主要用来报告出错和测试，ICMP报文封装在IP包中

ARP地址解析协议



ARP 和RARP

不管网络层使用的是什么协议，在实际网络的链路上传送数据帧时，最终还是必须使用硬件地址。

每一个主机都设有一个ARP 高速缓存(ARP cache)，里面有所在的局域网上的各主机和路由器的IP 地址到硬件地址的映射表。

当主机A 欲向本局域网上的某个主机B 发送IP 数据报时，就先在其ARP 高速缓存中查看有无主机B 的IP 地址。如有，就可查出其对应的硬件地址，再将此硬件地址写入MAC 帧，然后通过局域网将该MAC 帧发往此硬件地址。

DHCP动态主机配置协议

DHCP的操作过程:发送DHCP Discover;中继转发;回复DHCP Offer

标签交换与MPLS(Multi-Protocol Label Switching)

标签位置

标签Lable：索引;QoS：服务类型;S：堆栈字段,涉及多标签，为1表示最底部标签;TTL：生命期

MPLS

MPLS

标签交换路由器(LSR,Label Switched Router)

转发与交换的区别: IP转发使用最长前缀匹配算法;交换则使用标签作为索引，查询转发表，更为简单快速

标签边缘路由器(LER,Label Edge Router); 附加和删除标签

转发等价类(FEC,Forwarding Equivalence Class)

将终止于某特定路由器或LAN的多个流合并成一组，使用同一标签，这些流称为FEC

5.6.6 OSPF—内部网关路由协议

自治系统内部: 域内路由算法

内部网关协议: 距离矢量,RIP路由信息

链路状态: OSPF(Open Shortest Path First),开放最短路径优先

因特网有两大类路由选择协议:

内部网关协议IGP (Interior Gateway Protocol)

即在一个自治系统内部使用的路由选择协议。

目前这类路由选择协议使用得最多，如RIP和OSPF协议。

外部网关协议EGP (External Gateway Protocol)

若源站和目的站处在不同的自治系统中，当数据报传到一个自治系统的边界时，就需要使用一种协议将路由选择信息传递到另一个自治系统中。

这样的协议就是外部网关协议EGP。

在外部网关协议中目前使用最多的是BGP-4。

OSPF需求:开放最短路径优先OSPF

开放，公开发表； 支持多种距离衡量尺度，例如，物理距离、延迟

等； 动态算法； 支持基于服务类型的路由； 负载平衡； 支持分层系统； 适量的安全措施；支持隧道技术。

有向拓扑图

构造有向拓扑图

根据实际的网络、路由器和线路构造有向图；

每个弧赋一个开销值；

两个路由器之间的线路用一对弧来表示，弧权可以不同；

多路访问（multiaccess）网络，网络用一个结点表示，每个路由器用一个结点表示，网络结点与路由器结点的弧权为0；

分层路由

自治系统AS可以划分区域（areas）；

每个AS有一个主干（backbone）区域，称为区域0，所有区域与主干区域相连；

一般情况下，有三种路由

区域内

区域间:从源路由器到主干区域；穿越主干区域到达目的区域；到达目的路由器。

自治系统间

骨干区域: 主干路由器

区域边界路由器:连接多个区域的区域边界路由器

AS边界路由器:自治系统边界路由器

内部路由器:完全在一个区域内的内部路由器

BGP—外部网关路由协议:

外部网关协议与内部网关协议的不同

路由策略涉及三方面:

政治：微软的流量不经过谷歌

安全：五角大楼的流量不经过俄罗斯

经济：教育网路部承载商业流量

边界网关协议BGP（Border Gateway Protocol）

通过TCP连接传送路由信息；

采用路径矢量算法，路由信息中记录路径的轨迹

BGP

距离矢量协议，用政策选择路由

路径矢量协议，BGP跟踪所使用的路径

5.6.8 因特网多播

本地组播地址(永久组地址)

224.0.0.1 LAN上的所有系统

224.0.0.2 LAN上的所有路由器

224.0.0.5 LAN上的所有OSPF路由器

224.0.0.251 LAN上的所有DNS服务器

5.6.9 移动IP

目标要求

每台移动主机固定IP;不修改固定主机软; 不修改路由器软件;发给移动主机数据多数不绕道;移动主机在家无开销

小结

IPv4协议: 分段与重组;IPv4地址

ICMP协议:ICMP报文封装在IP包中

ARP协议: ARP协议工作过程

RARP协议: RARP与BOOTP的区别

RIP: 内部网关协议;距离向量算法; 基于UDP

OSPF:内部网关协议;链路状态算法;分层思想;四种类型路由器

BGP:外部网关协议;路径向量算法;基于TCP

CIDR:CIDR的思想;IP地址分配和掩码配置;最长匹配原则

IPv6:IPv6与IPv4的主要区别;IPv6地址;双栈模式与隧道模式

**本章小结**

**网络层的定义、功能和设计方法**

**分组交换：虚电路和数据报**

**路由算法的分类和基本原理**

**网络互联：Internet 的网络层**

**IP协议;子网划分;RIP、OSPF、BGP; NAT;IPv6**

**路由器的基本原理**