

《数据通信与网络》课程笔记

Data Communications and Networking

James

XianIP@163.Com

一、 介绍

1. OSI 模型(OSI Model: Open Systems Interconnection)

应用层(Application):

目的: 直接向用户提供服务

作用: 完成用户想要在网络上完成的各种工作

表示层(Presentation):

目标: 数据格式的转换

作用: 把应用的处理的信息转换成适合网络传输的格式, 或者把下层数据转换为上层能处理的格式。例如图片的编解码, URL 的加解密

会话层(Session):

目标: 把下层进行包装, 给上层提供全自动的自动收发包和寻址

作用: 建立和管理应用程序之间的通信。例如: 验证用户登录、断点续传

传输层(Transport):

目标: 发送大量数据时可能丢包, 要知道丢了哪些包, 让发送者重传

作用: 监控数据传输服务的质量, 保证报文的正确传输。TCP/UDP/进程/端口在这一层

数据名称: user datagram(UDP), segment(TCP)

网络层(Network):

目标: 网络中有台计算机, 找到要发的那台; 中间有多个节点, 选择一个路径并投递

作用: 通过路由选择算法, 为报文(由上一层数据打包而来)通过通信子网选择最恰当的路径。这一层定义的是 IP 地址, 通过 IP 地址寻址, 所以产生了 IP 协议。

数据名称: 报文/数据包(packet)

数据链路层(Data link):

目标: 通过各种控制协议, 把有差错的物理信息变为无差错的, 能可靠传输的数据帧。

作用: 接收来自物理层的比特流, 并封装成帧, 传送到上一层; 同样, 接收从上一层传输来的数据帧, 拆成比特流, 发送到物理层。

数据名称: 帧(frame)。

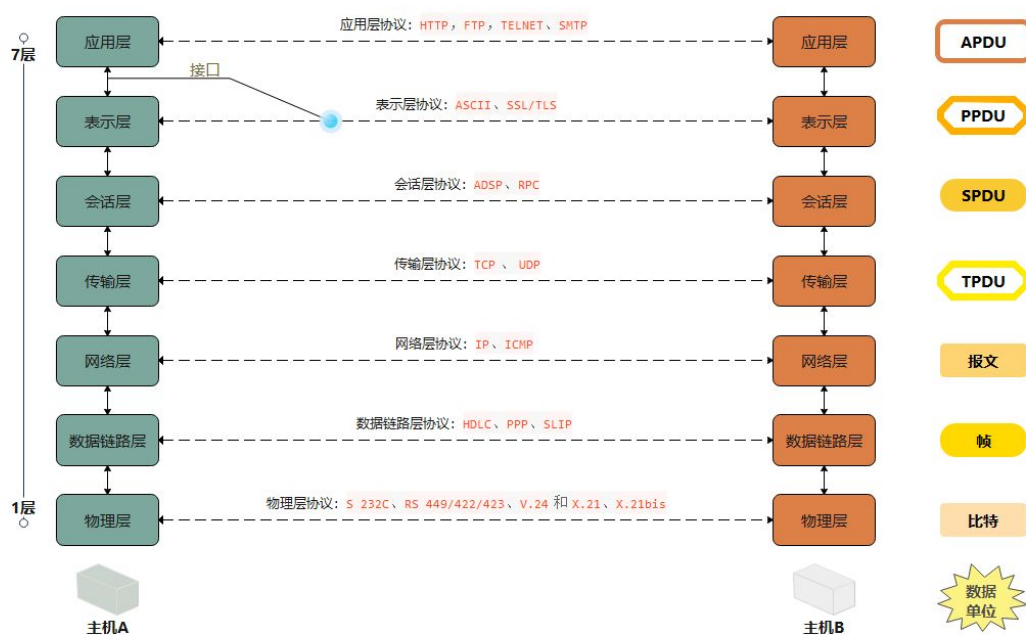
物理层(Physical):

目标: 定义物理设备通信标准, 如网线/光纤接口类型, 传输介质速率, 波长等。

作用: 用于转换和传输比特流, 也就是我们常说的模电转换。

数据名称: 比特(bit)。

OSI参考模型



2. TCP/IP 协议模型

TCP/IP Protocol Suite: Transmission Control Protocol/Internet Protocol

应用层使用户能连接到网络: HTTP/FTP/SMTP/TELNET/etc.

OSI分层	TCP/IP分层	TCP/IP协议栈	本书按5层讲解
应用层	应用层	HTTP FTP SMTP POP3 DNS 等	应用层
表示层			
会话层			
传输层	传输层	TCP UDP	传输层
网络层	网络层	ARP IP ICMP IGMP	网络层
数据链路层	网络接口层	以太网 PPP 帧中继 X.25	数据链路层
物理层			物理层

<https://blog.csdn.net/taotongning>

3. 术语定义

数据通信(Data communications): 两个设备通过某种形式的传输媒介进行数据交换。
(Data communication is the exchange of data between two devices via some form of transmission media)

网络(Network): 一组可以通讯的设备互相之间的连接。(A network is the interconnection of a set of devices capable of communication) 这些设备可以是任何在网络中充当主机的设备，例如电脑、手机、交换机等等

4. 五大组件

发送者(Sender): The sender is the **device that sends the data** message. It can be a computer, workstation, a telephone handset and so on.

接收者(Receiver): The receiver is the **device that receives the message**. It can be a computer, workstation, a telephone handset and so on.

消息(Message): The message is the **information (data) to be communicated**. Forms of information include text, numbers, pictures, audio and video.

传输媒介(Transmission medium): The transmission medium is the **physical path** by which a message travels from sender to receiver. Some examples of transmission media include twisted-pair wire, coaxial cable, fiber-optic cable and radio waves.

协议(Protocol): A protocol is a set of **rules that govern data communications**. It represents an agreement between the communicating devices.

5. 数据流(Data flow)

单向(Simplex): 数据只能从一个设备到另一个设备。例如键盘。

半双工(Half-Duplex): 数据可双向流动，但不可同时发送和接收。例如对讲机(Walkie-talkie)

全双工(Full-Duplex): 数据可同时双向流动。例如手机。

6. 网络好坏的标准(Network Criteria)

性能: 承载力和延时 Performance is often evaluated by i) throughput and ii) delay.

可靠: 错误率，恢复时间，灾难中的稳健型 Reliability is often measured by i) the frequency of failure, ii) the time it takes to recover from a failure and iii) the network's robustness in a catastrophe.

安全: 阻止未授权的访问和损坏；有无从违规和数据丢失中恢复的策略。 Security include i) protecting data from unauthorized access and from damage and ii) implementing policies and procedures for recovery from breaches and data losses.

7. 连接的类型

a. Point-to-Point (点对点连接)

两个设备直接连接，这个连接的容量只能被这两个设备使用

b. Multi-Point (多点连接)

多个设备共享一个连接，这个连接的容量被多个设备共享

8. 网络拓扑(Network topology)

a. 网状 (Mesh)

每两个设备之间都有一个专用连接。因此 n 个设备需要 $n(n-1)/2$ 条连接

优点：

- 1) 专用连接：链接的总容量只给两个设备用
- 2) 健壮：整个系统不会因为一个不可用的连接而无法使用
- 3) 隐私和安全：只有预定的收件人才能在专用连接上收到消息
- 4) 易于识别个隔离故障：可以对流量进行路由，避免出现问题连接

缺点：贵

b. 星型 (Star)

每个设备都有一个专用连接连接到中央控制器。设备之间不能直接通信，必须靠中心控制器交换。

优点：

- 1) 成本：比网状拓扑更省钱
- 2) 健壮：整个系统不会因为一个不可用的连接而无法使用
- 3) 易于识别个隔离故障：可以对流量进行路由，避免出现问题连接

缺点：整个网络依赖于一个中央控制器

c. 总线型 (Bus, 唯一的 Multipoint)

一根长电缆充当骨干网，其他设备都连接到这根电缆上。早期网络的常见拓扑之一。

优点：

- 1) 容易安装
- 2) 与网状和星型相比，布线更少

缺点：

- 1) 单点故障(single point failure): 骨干网(backbone)
- 2) 难以隔离故障(difficult fault isolation)

d. 环形 (Ring)

每个设备都有专用点对点连接，每个设备只连接到两个设备。

优点：

- 1) 容易安装和重新配置：每个设备只连接到它的邻居
- 2) 简化的故障和隔离：信号始终在环上，指定时间内无信号则报警

缺点：

- 1) 单点故障(single point failure): 环断掉会导致整个网络失效

二、 网络模型

1. TCP/IP 协议模型

Application: enables the users to access the network: HTTP, FTP, SMTP, Telnet, etc.

Transport: responsible for the process-to-process delivery of the entire message: process-to-process communication - User Datagram Protocol / Transmission Control Protocol. UDP: Best effort delivery of user datagrams. TCP: flow, error (retransmission/reordering) and congestion control of segments.

Network: responsible for the host-to-host (source-to-destination) delivery of a packet / datagram across multiple network links: host-to-host communication, routing.

Data link: responsible for delivering frames from one station to the next without errors: Data Link Control (DLC) sublayer: framing, error detection and correction of frames/bits; Medium Access Control (MAC) sublayer: physical hardware address, medium access control.

Physical: coordinates the functions required to transmit a bit over a transmission medium: bit representation, type of encoding. Not the physical transmission mediums (twisted pair, coaxial, radio wave).

2. 术语定义

协议(Protocol): 定义了发送者和接受者以及全部中间设备必须遵循的规则, 这些规则使得通信更有效率

A **protocol** defines the rules that both the sender and receiver and all intermediate devices need to follow to be able to communicate effectively.

当通讯比较简单时, 我们可以用一个简单的协议。当通讯比较复杂, 每一层我们都需要一个协议, 或者**协议分层**(称为模块化)

When communication is simple, we may need only one simple protocol; when the communication is **complex**, we **need a protocol at each layer, or protocol layering** (referred to as modularity).

3. 注意

1. 交换机没有 Network 这一层, 路由器才有这一层, 这层用来做路由的
2. 路由器有两个到多个 Data link 和 Physical 层, 连接不同子网

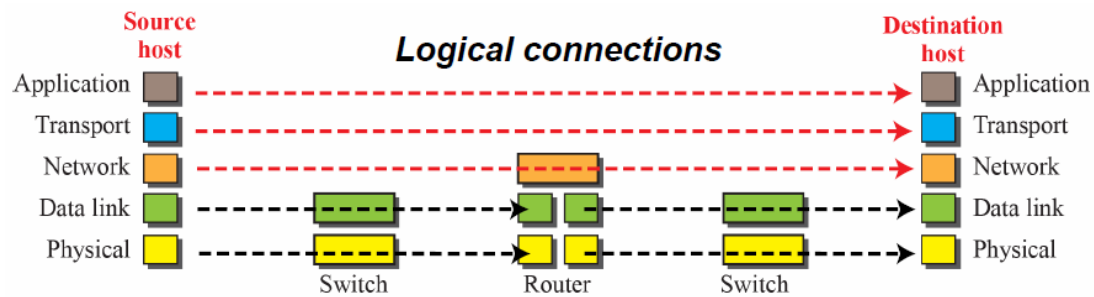
4. 协议层的两个原则

1. 每一层都要能执行两个相反的任务, 即每个方向一个任务(加密解密/收发)
2. 每一层下面的两个对象(数据单位)应该是相同的

5. 逻辑连接

Duty of the application, transport and network layers is **end-to-end**.

Duty of the data link and physical layers is **hop-to-hop**.



6. 封装/解封装(Encapsulation and Decapsulation)

数据每向下一层，就会在数据前面加一个 Header

7. 寻址(Addressing)

Packet names	Layers	Addresses	
Message	Application layer	Names	▪ Eg. www.bcit.ca
Segment / User datagram	Transport layer	Port numbers	▪ Eg. Port 80 (http)
Packet / Datagram	Network layer	Logical addresses	IP address
Frame	Data-link layer	Link-layer addresses	MAC ID
Bits	Physical layer		

三、数据与信号

- 术语定义

模拟信号(analog data): 连续的信息。在一段时间内有无穷多个强度等级。

数字信号(digital data): 有离散状态的信息。只有有优先个已定义的数值。

周期信号(periodic signal): 在一个可测量的时间范围内(period)完成一种模式,并在后续的相同时间范围内重复这一模式。一个完整模式的实现称为循环。

非周期信号(nonperiodic signal): 变化不会随着时间的变化出现重复的模式或循环。

峰值振幅(peak amplitude): 信号最高强度的绝对值,与携带能量成正比,单位是**伏特**。

周期(period): 信号完成一个循环所需的时间,以秒为单位。与频率 f 互为倒数。

频率(frequency): 1 秒以内的周期数。与周期互为倒数。

相位(phase): 波形相对于时间 0 的位置。指明了第一个循环的状态。

时域图(time-domain plot): 横轴为时间,纵轴为振幅。

频域图(frequency-domain plot): 横轴为频率,纵轴为振幅。

带宽(bandwidth): 最高频率与最低频率之差。

比特率(bps): 每秒传输的 bit 的个数

位长(bit length): 位长 = 传播速度 * 位持续的时间

- 两个极端

信号不变化: 频率为 0。因为信号永远无法完成一个循环。

信号瞬变: 频率值为无穷大。因为其周期是 0, 频率为 $1/0$ 。

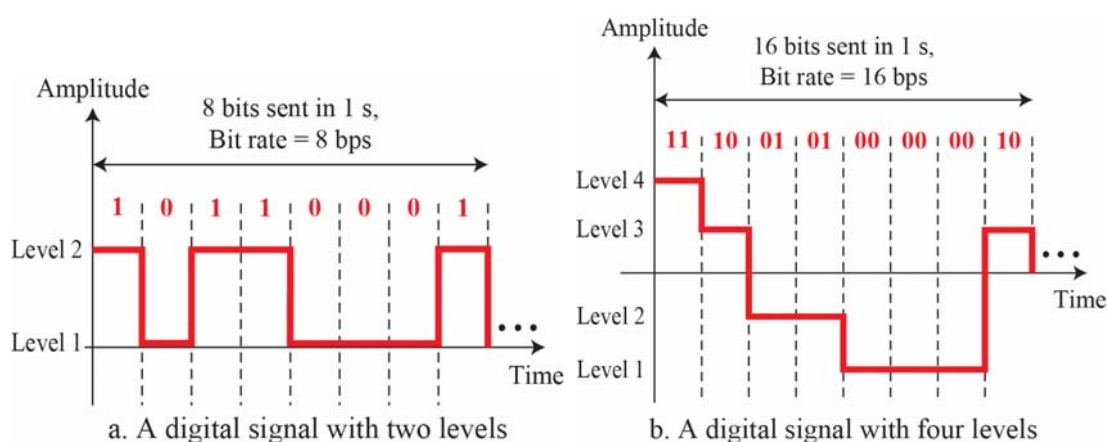
- 波长的计算:

波长 = 传播速度 * 周期 = 传播速度/频率

$$\lambda = c / f = c * T$$

- 电平

一个电平(电压值)可以发送一个到多个位



一个数字信号有 8 个点平, 每个电平可以传输 $\log_2 8 = 3$ 个位, 如果是 9 个点平, 则略大于 3 个, 向上取整, 每个电平要传输 4 个位。

- 数字信号的传输

1. 基带传输 (baseband)

通过通道发送数字信号，而数字信号是无穷大带宽的复合模拟信号

低通通道(low-pass channel): 传输通道的频率最低为 0，其实无法实现。可被视作最低频率为 0 的带通通道

2. 宽带传输

把数字信号**调制(modulate)**成模拟信号再传输

带通通道(bandpass channel): 带宽不从 0 开始的通道，比低通通道更可用

- 传输减损(Transmission Impairment)

3. 衰减(attenuation): 信号经过介质，失去能量需要放大信号

衰减的值用分贝(decibel)来计量，其中 P1 为原始信号，P2 为传输后的信号

$$\text{dB} = 10\log_{10}(P2/P1)$$

常见的数值：-3dB 为衰减了一半

3. 失真(distortion): 信号改变了形态或丢失，产生在由不同频率成分组的符合信号中

3. 噪声(noise): 电缆中的电子移动产生的随机信号

信噪比 (SNR) = 平均信号功率 / 平均噪声功率

SNR 也用分贝作为单位来描述

$$\text{SNR}_{\text{db}} = 10\log_{10}\text{SNR} \rightarrow \text{也可反过来求 SNR}$$

无噪声通道的 SNR 和 SNRdb 都是无穷大，但是现实不可能有这样的比例

- 数据速率限制

三要素: 1. 有效带宽（可用频率范围） 2. 使用的信号电平数 3. 通道质量(噪声电平)

奈奎斯特(Nyquist bit rate)比特率公式

适用于无噪声的通道，通过**增加电平**来增加发送速度，但是电平数越多，接收方很难区分各个电平，**减少了系统的可靠性**。

$$\text{比特率} = 2 * \text{带宽} * \log_2 L$$

这里的带宽是频率带宽，结果是比特率

香农容量定理(Shannon Capacity)

$$\text{通道容量(比特率)} = \text{带宽} * \log_2(1 + \text{SNR})$$

例：信号很弱，噪声非常大，SNR 是 0，那么算出来通道容量也是 0，无法传输数据

- 性能(Performance)
 1. **带宽(bandwidth):** 以赫兹衡量的带宽, 以及用每秒比特数衡量的带宽
 2. **吞吐量(throuput):** 真实传输速度永远小于通道的理论速速
 3. **延迟(delay):** 由如下 4 部分相加得来
 - 传播时间(propagation time): 距离 / 传播速度
 - 传输时间(transmission time): 报文的 bit 数 / 带宽
 - 排队时间(queue time): 设备处理报文前, 保持报文的文件 (网络越繁忙排队越久)
 - 处理延迟(processing delay): 目标收到报文后, 处理报文需要的时间
 4. **带宽与延迟的乘积:** 充满链路的位的个数。如果*2 则是全双工, 发送前必须等确认
- 抖动(jitter)

排队靠后的报文, 比排队较前的报文更早到达, 就会出现抖动

四、数字信号转数字信号

- 术语定义

数据元素(data element): 一块信息的最小实体, 即 bit。是需要被发送的。

信号元素(signal element): 数字信号的最小单元。是我们能发送的。

波特率(baud rate): 一秒钟发送的信号元素的数量, 也被称为 pulse rate 或 modulation rate

基线(baseline): 接收方计算的收到的信号的运行平均值

基线偏移(baseline wandering): 连续接收到都是 0 或者 1 时导致计算平均值偏高或偏低

自同步(self-synchronizing): 数字信号中含有定时信息 (提示接收方起始/中间/结束)

- 如何用数据速率计算信号速率(波特率)

$$S = c * N * 1/r \text{ (baud)}$$

其中

r = 数据元素 / 信号元素

N 是数据的速率(bps)

C 是情形因子, 会根据实际情况改变, 平均是 1/2

S 是信号元素的速率 (波特率)

- 计算最小带宽 (频率范围): $B_{min} = c * N * 1/r$
- 计算最大带宽 (数据速率): $N_{max} = 1/c * B * r$ (此公式与奈奎斯特公式结果一致)

- 线路编码方案

1. 单极编码(unipolar)

所有的信号都在时间轴的一边, 要么在上面要么在下面。

不归零(NRZ): 正电平定义为 1, 零电平定义为 0, 中间信号不会回到 0

缺点: 成本很高, 标准功率 (每个单元线路阻抗发送 1 位的功率) 是极性 NRZ 方案两倍

2. 极性编码(polar)

电平在时间轴的两边。0 的电平可能是正, 1 的反而是负。

下列两方案波特率都是 $N/2$ (bps 的一半):

NRZ-L(NRZ 电平编码): 信号电平决定了位的值。缺点是可能有基线偏移, 且极性改变会导致接收方解释数据错误。

NRZ-I(NRZ 反向编码): 没有跳变时为 0, 有跳变为 1。

NRZ 编码的缺陷: 时钟可能不同步

归零编码(RZ): 在正负切换之间先跳到 0 保持一段时间, 使得接收方知道信号变化。缺点是要用三个电平, 使得生成和辨别很困难。

双向编码 (biphase)

曼彻斯特(Manchester): 位的持续时间被两等分, 前半部分如果是一个电平, 后半部分就切换到另一个电平。克服了 NRZ-L 的一些缺陷。

差分曼彻斯特(Differential Manchester): 位的持续时间也被两等分, 两个位交接处有跳变, 则下一个位是 0, 没有跳变则是 1。

上两个编码的最小带宽是 NRZ 的两倍, 即 N (与 bps 相同)

3. 双极性方案(bipolar)

一个数据元素的电平是 0, 另一个则是在正负之间交替。

交替传号反转编码(AMI/alternate mark inversion): 0 用 0 电平表示, 正负电平是 1

伪三元编码(pseudoternary): AMI 的变体, 1 用 0 电平表示, 正负电平是 0

这种方案可以使得振幅始终在 0 附近, 不会产生基线偏移

五、模拟信号转数字信号

- 术语定义

脉冲码调制(PCM/pulse code): 把模拟信号转成数字信号

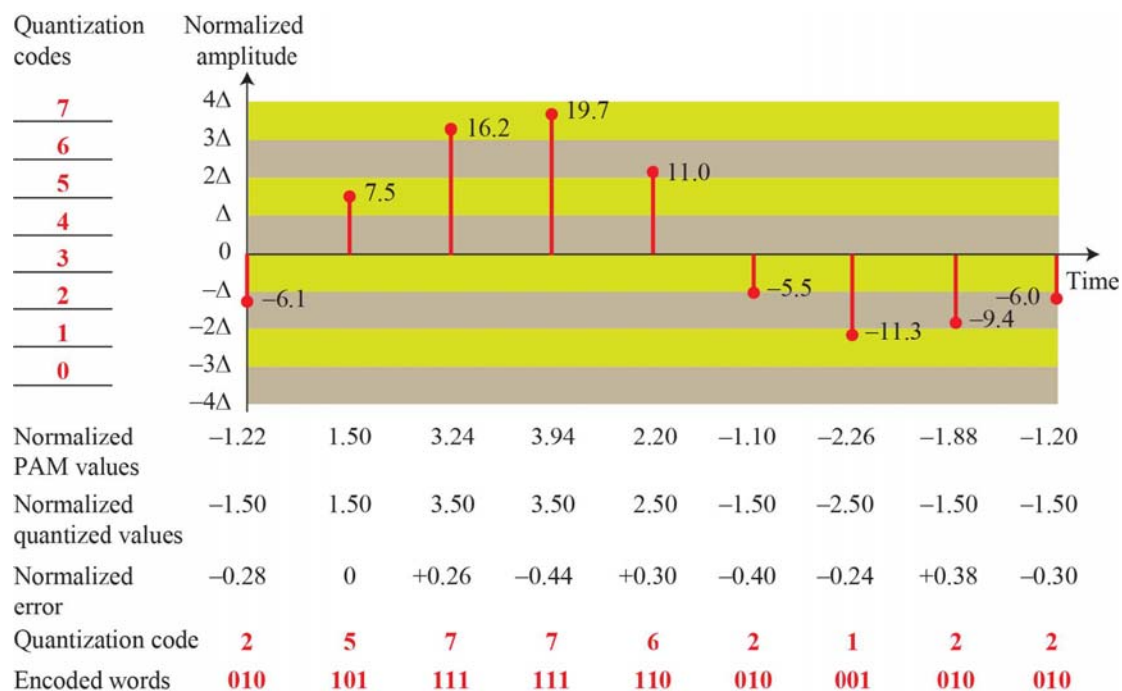
- PCM

采样(sampling): 每隔 T_s 秒对模拟信号进行采样, T_s 是样本间隔或周期

采样率(sampling rate/frequency): 是 T_s 的倒数, 定义为 $f_s = 1 / T_s$

奈奎斯特定理(Nyquist theorem): 采样率至少是模拟信号最高频率的 2 倍

量化(quantizing): 把采集到的振幅转换为特定的数字



量化误差取决于量化级别 L 或每个样本的位数 N_b (反过来也可 SNR_{db} 求位数)

$SNR_{db} = 6.02 * N_b + 1.76 \text{ dB}$ 教科书里用 $SNR_{db} = 6 * N_b - 1.25 \text{ dB}$

这里的 N_b 是编码后的 bit 位数, 也就是 $N_b = \log_2 L$, 所以它也可以求电平数

● 编码(Encoding)

比特率(bit rate)计算公式: $\text{比特率 } N(\text{bps}) = \text{采样速率} * \text{每个样本的位数} = f_s * N_b$

数字信号的频率带宽: 模拟信号的 N_b (bit 位数)倍。

例如一个 4kHz 的信号每个样本 8 位则至少需要 32kHz 的通道。

PCM 最小带宽: $B_{min} = N_b * B_{analog}$

通道最大速率: 比特率 $N_{max} = 2 * \text{带宽} * \log_2 L$

通道最小带宽: $B_{min} = N / (2 * \log_2 L)$

六、数字信号转模拟信号

- 术语定义

数字到模拟的转换(digital-to-analog conversion): 根据数字数据中的信息改变模拟信号的某种特征的过程

载波信号(carrier signal): 又称为载波频率。由发送设备产生的高频信号, 作为基础来承载信号信息。接收设备要把自己的收听频率改到与发送设备的载波信号频率一致。

调制(modulation): 数字信息通过改变载波信号的一个或者多个特性(振幅、频率、相位)来调制载波信号。又称为**移动键控(shift keying)**。

数字转模拟的优点: 可以把发生带宽平移的可用带宽

- 比特率(bite rate)和波特率(baud)

$$S = N * 1/r \text{ (baud)}$$

这里的 N 是 bps, r 是一个信号元素携带的元素个数, r 是 $\log_2 L$

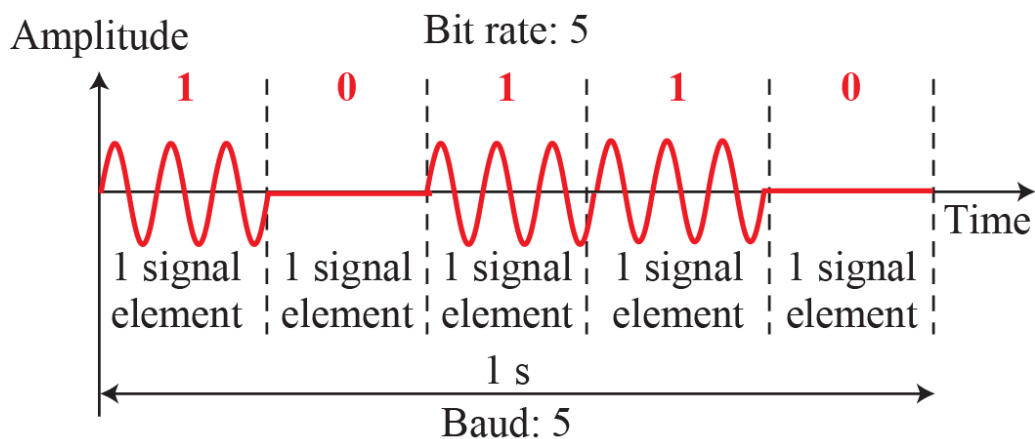
注意: 这里的 L 不是电平个数, 而是信号元素的类型的个数

- 幅移键控 (ASK, amplitude shift keying)

仅通过改变载波信号的振幅来生成信号元素。

二进制 ASK (BASK)

尽管我们可以有多种不同振幅的信号元素, 但是 **BASK** 只用 2 种不同的振幅。ASK 又称为开关键控(on-off keying, OOK)。一个信号的振幅是 0, 另一个跟载波频率一样。



ASK 的带宽

$$B = (1 + d) * S$$

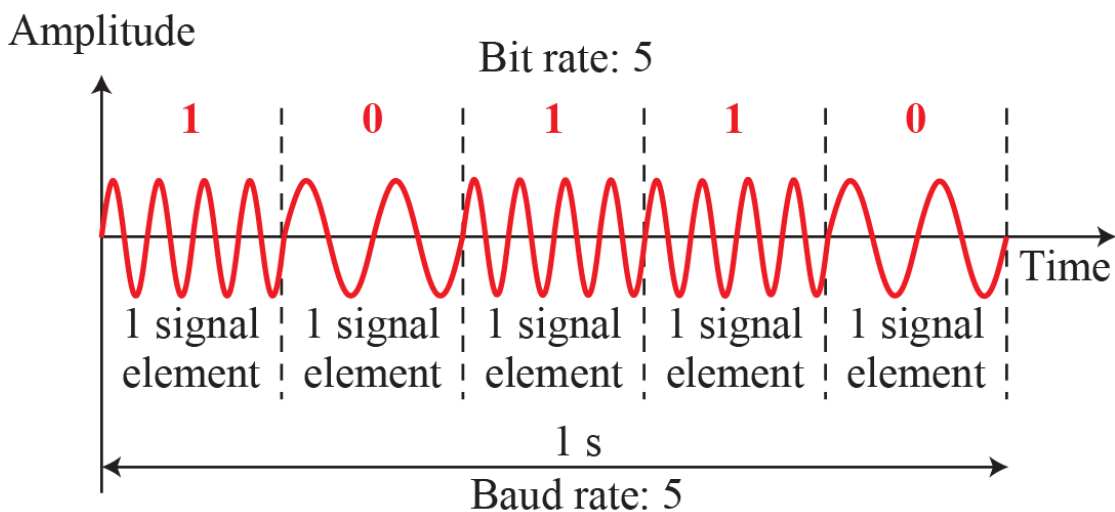
S 是波特率(载波频率), B 是频率带宽, d 在 0-1 之间, 因此带宽在 S 到 2S 之间。一般把载波频率 f_c 选到这个带宽中间。

- 频移键控(FSK, frequency shift keying)

仅通过改变载波信号的频率来生成信号元素, 频率在一个信号元素持续的周期内恒定。

二进制 FSK (BFSK)

选用两个载波频率 f_1 和 f_2 ，一个表示 0，另一个表示 1。有两个带宽，中点频率分别为 f_1 和 f_2 ，这两个频率之差为 Δf



BFSK 的带宽

$$B = (1 + d) * S + 2 \Delta f$$

S 是波特率(载波频率), B 是频率带宽($f_2 - f_1$), d 在 0-1 之间, 其中 $2 \Delta f$ 至少要等于 S

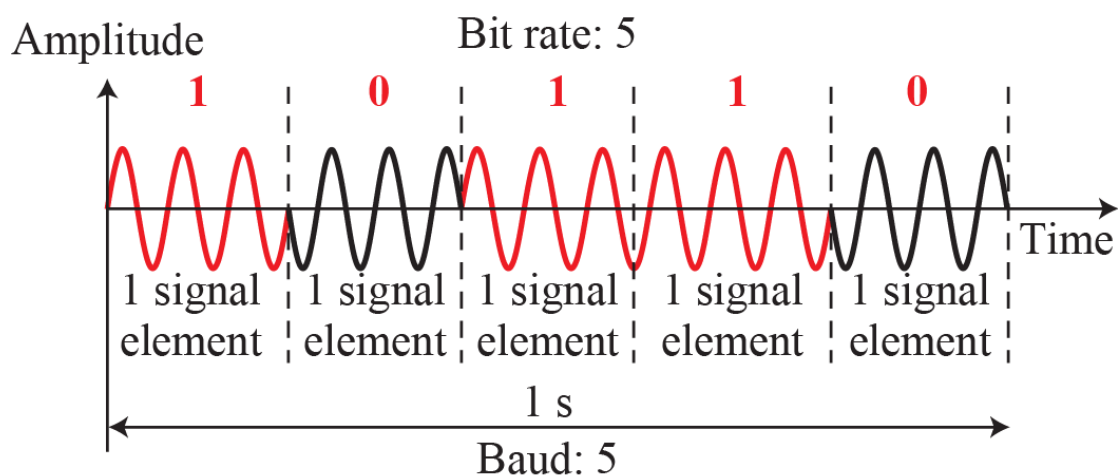
实现

非相干: 一个信号元素结束的位置与下一个信号元素开始的位置不连续。可看作两个 ASK, 用两个载波频率实现。

相干: 一个信号元素结束的位置与下一个信号元素开始的位置连续。用压控振荡器 (voltage-controlled oscillator, VCO) 实现, 压控振荡器根据输入的数字信号电平改变频率。

- 相移键控 (phase shift keying, PSK)

通过改变载波的相位来表示两个或更多不同的信号元素, 峰值和频率保持不变。



二进制 PSK (BPSK)

只用两个信号元素，一个相位是 0 度，另一个相位是 180 度，如上图。

PSK 优点

1) PSK is less susceptible to noise than ASK (noise changes amplitude easier than it can change the phase) 比 ASK 不受噪音影响，因为看相位不看振幅。带宽与 ASK 一致。

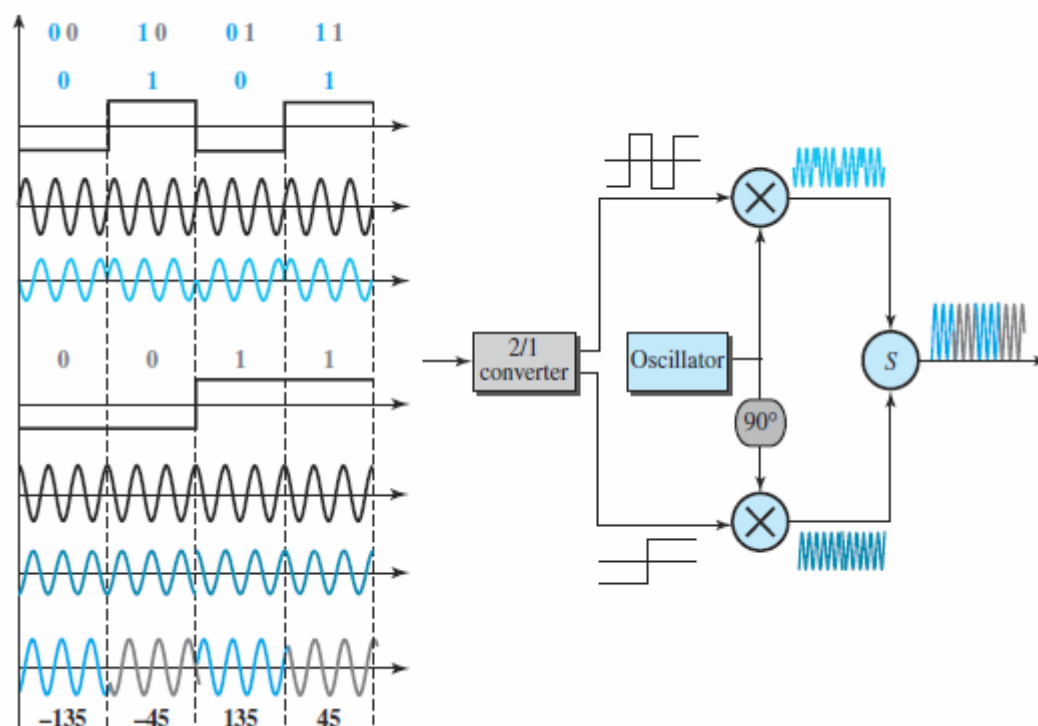
2) PSK does not require multiple carrier signals as compared to FSK. 比 FSK 少用载波信号，减少带宽占用和浪费。

PSK 缺点

1) PSK requires more sophisticated hardware to be able to distinguish between phases. 需要复杂的硬件设备来区分相位

正交 PSK (QPSK, quadrature phase shift keying)

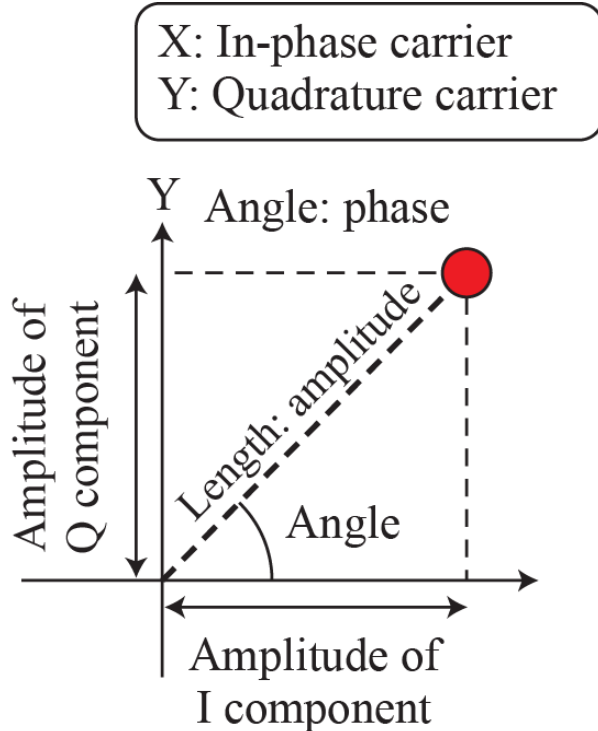
每个信号元素承载 2 个位，减少了波特率和所需带宽。用两个 BPSK 进行调制：一个是相位相同(in-phase)的，一个是相位不同的(out-of-phase)。数字信号的两个位分别发到两个 BPSK 调制后叠加，分别产生相位为 45 度、-45 度、135 度和-135 度其中之一，一共 4 种信号元素。其频率只用原来数字频率的一半。



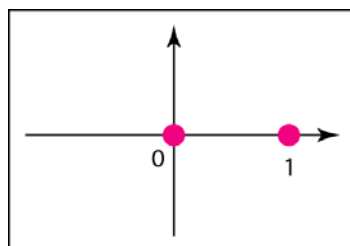
- 星座图 (constellation diagram)

星座图可定义信号元素的振幅和相位，尤其对同相和正交特别有用。

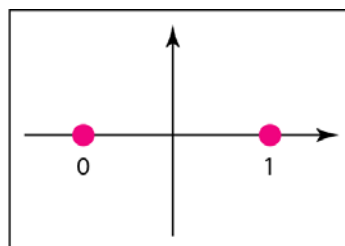
一个元素用一个点表示，携带的位或位组合写在点的旁边。



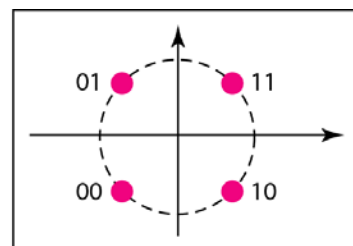
1. X 轴与同相载波相关
2. Y 轴与异相载波相关
3. 点在 X 轴上的投影是同相成分的峰值振幅
4. 点在 Y 轴上的投影是正交成分的峰值振幅
5. 点到原点的连线长度是该信号元素的峰值振幅(X 与 Y 的组合)
6. 连线与 X 轴之间的角度是信号元素的相位



a. ASK (OOK)

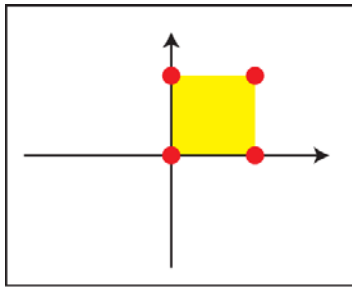


b. BPSK

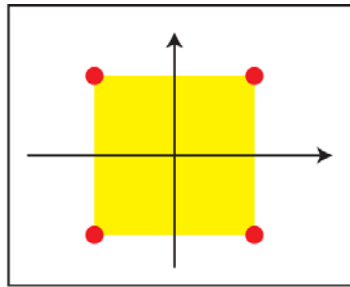


c. QPSK

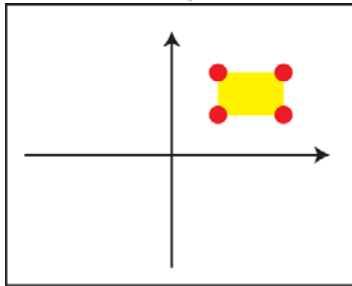
- 正交振幅调制 (QAM, quadrature amplitude modulation)



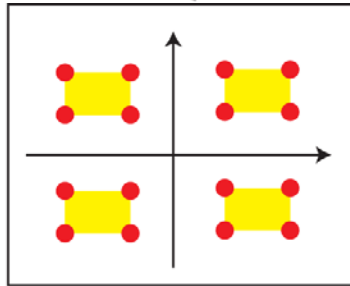
a. 4-QAM



b. 4-QAM



c. 4-QAM



d. 16-QAM

- a. 用单极 NRZ 调制：数字信号分别为 0 和 1，两两组合，跟 ASK(OOK 一致)
- b. 用极性 NRZ 调制：数字信号分别是 -1 和 1，两两组合，跟 QPSK 一致
- c. 用 2 个正电平的信号调制（ 2×2 总共 4 个信号元素）
- d. 用 4 个正电平 4 个负电平，共 8 个电平调制（ 4×4 总共 16 个信号元素）

需要电平个数的算法：点数开方再除以 2

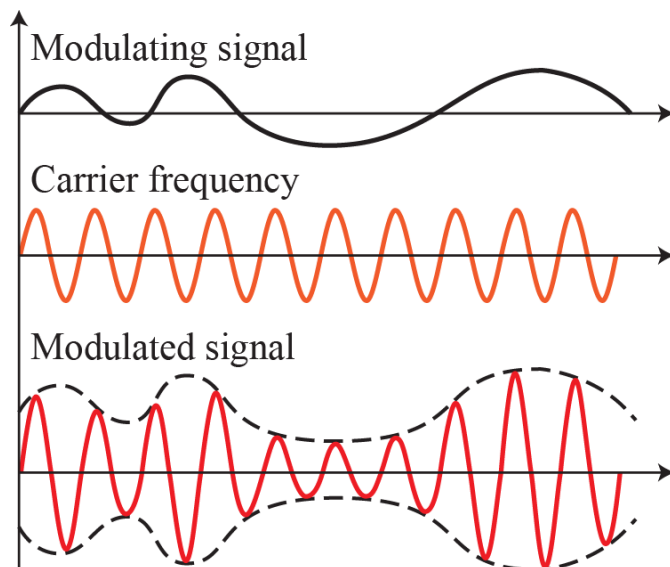
QAM 所需的最小带宽与 ASK 和 PSA 的最小带宽相同。

对于 ASK 来说，QAM 和 PSK 有同样的优点。

六、模拟信号到模拟信号转换 (analog-to-analog conversion)

为了实现将低通信号平移, 可以用 AM, FM 和 PM 三种方法实现

调幅(AM, amplitude modulation)



原始信号称为调制信号。

载波信号的带宽是调制信号带宽的两倍。

载波信号的频率和相位不变, 只有振幅随着调制信号的振幅而改变, 一般用乘法器实现。

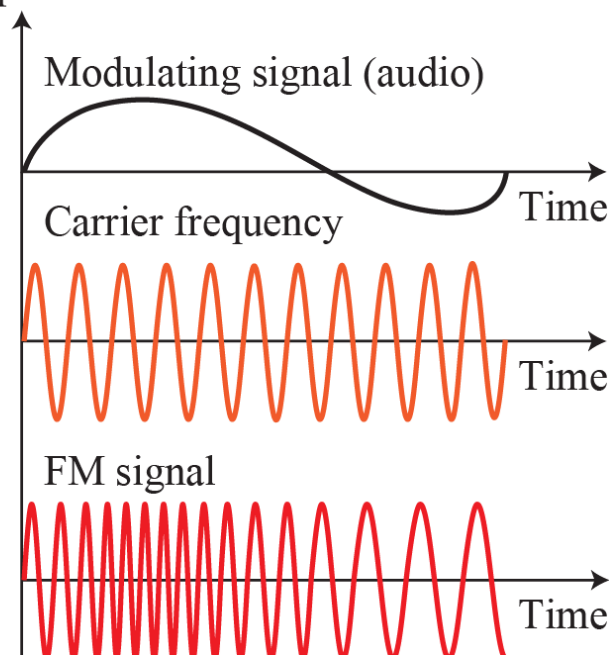
调频(FM, frequency modulation)

载波信号的频率随着调制信号振幅的改变而调整, 振幅越大, 频率越高。

用 FSK 的压控振荡器实现, 载波信号的峰值和相位保持不变。

带宽很难确定, 从经验上看是模拟信号的若干倍, 或者 $2(1+\beta)B$, β 一般是 4

Amplitude

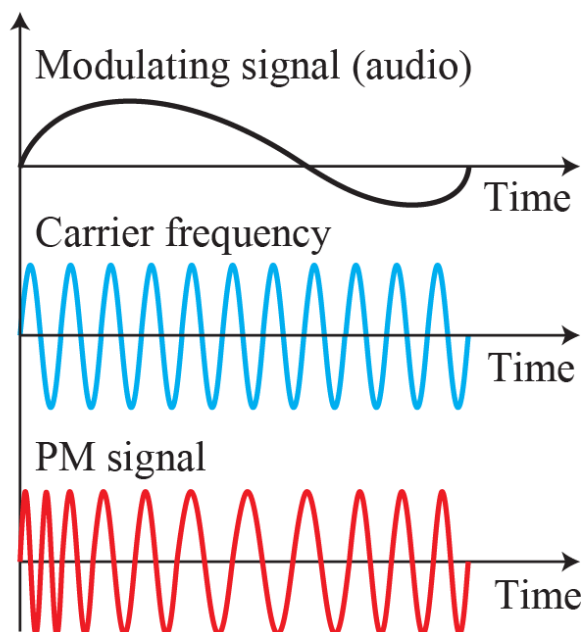


调相(PM, phase modulation)

载波信号的相位根据调制信号的振幅的导数成正比，调制信号变化越快（坡越陡），载波信号的频率越高。

用 FSK 的压控振荡器实现和导数发生器来实现，载波信号的峰值和相位保持不变。带宽很难确定，从经验上看是模拟信号的若干倍，或者 $2(1+\beta)B$ ， β 一般是 1 到 3

Amplitude

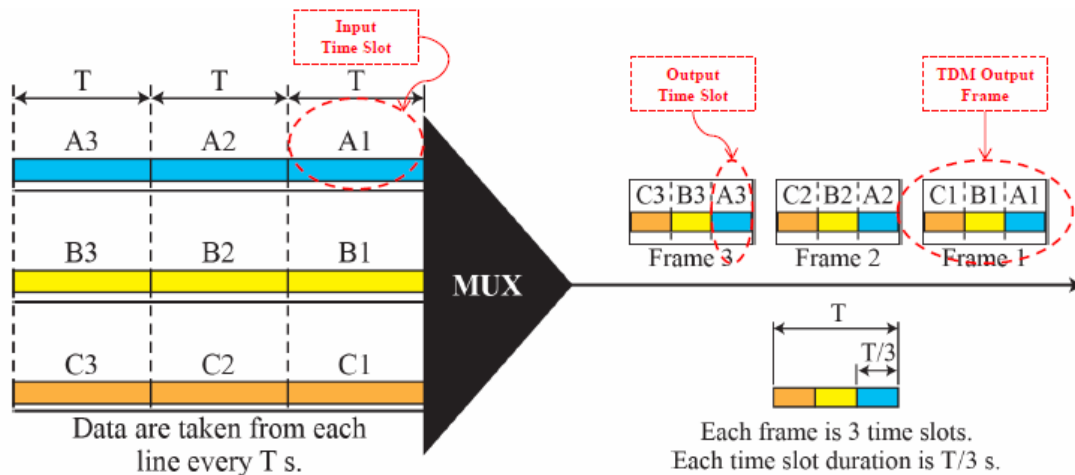


七、带宽利用

- 复用 (multiplexing): 允许同时通过一条数据链路传输多个信号的一组技术
N 条线路共享一个链路的带宽，把传输流量送到**复用器(multiplexer, MUX)**，复用器把流量组成一个单独的传输流。在接收端，这个传输流量被**分离器(demultiplexer, DEMUX)**接收，并分解成原来的几个独立的传输流。**通道(channel)**是复用器和分离器之间的链路的部分。一个链路可能有 N 的通道。
- 频分复用(frequency-division multiplexing, FDM)
不同的设备信号，用不同频率的载波进行调制，然后再合成一个复合信号，然后发送出去。
 - 例1. 无防护带宽，所有频率全部加到一起
语音通道 4kHz，三条合并到一条变成 12kHz，平移通道上占用 20-32kHz
 - 例2. 有防护带宽，频率之间有相隔
5 个通道，每个带宽 100kHz，每个通道之间 10kHz 防护，最小带宽需要 $5 \times 100 + 4 \times 10 = 540\text{kHz}$
- 波分复用(wavelength-division multiplexing, WDM)
跟彩虹一样，在发送端用复用端合并多个光信号，到接收端的用棱镜完成分离。

- 时分复用(time-division multiplexing, TDM, 用于数字信号)
分为同步时分复用(synchronous TDM)和统计时分复用(statistical TDM)

同步时分复用



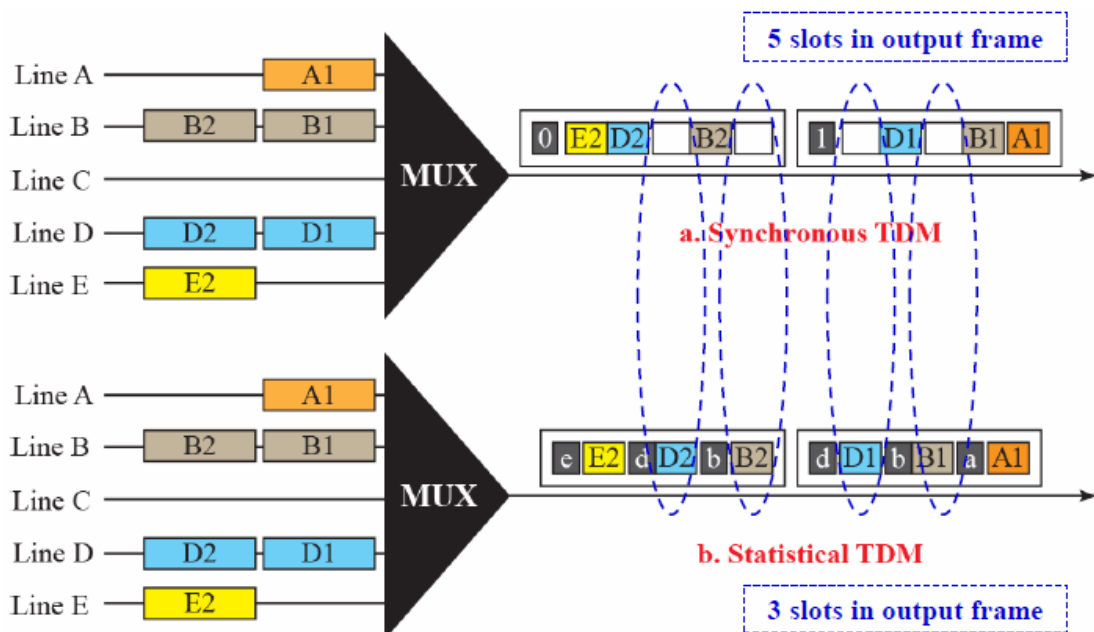
输入前每个单元(slot)持续的时间为 T , 输入后在 frame 的 slot 里变成了 T/n , 这里的 n 是输入的线路数。

有时候, 每个 frame 之间会加一个同步位(synchronous bit), 此时计算各种值需要有改变。

缺点: 不够有效率, 如果一个输入线路没有输入, 那么输出线路也会为它空出位置。

原文: If a source does not have data to send, the corresponding slot in the output frame is empty.

统计时分复用

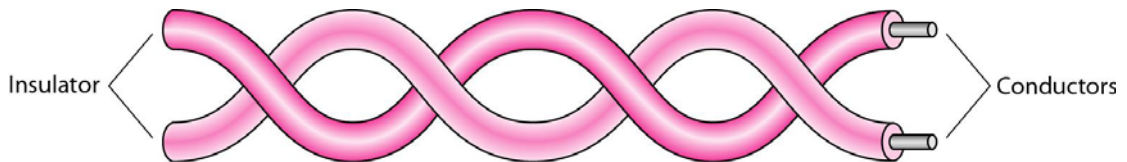


需要寻址, 每个数据都会同时发送目标地址。不需要同步, 所以没有单独的同步位。

八、传输介质

- 传输介质(transmission medium): 能从源端传送信息到目的端的任何东西, 处于物理层之下, 可称为第 0 层
- 有向介质(guided media)
在设备之前提供通路(conduit)的介质, 包括双绞线(twisted-pair cable)、同轴电缆(coaxial cable)和光缆(fiber-optic cable)。沿着这类介质传播的信号, 其传输方向和传播范围受介质的物理边界限制。

1. 双绞线: 一根传输信号, 另一根做地线, 接收方使用两根线的差值。传输中一条导线离噪声源近, 但是在下一个绞合段, 情况恰好相反, 所以可以让两条线受的干扰相同。Suppose in one twist, one wire is closer to the noise source and the other is farther; in the next twist, the reverse is true. The receiver, which calculates the difference between the two wires receives no unwanted signals as the unwanted signals are mostly cancelled out.



非屏蔽双绞线(unshielded twisted-pair, UTP)

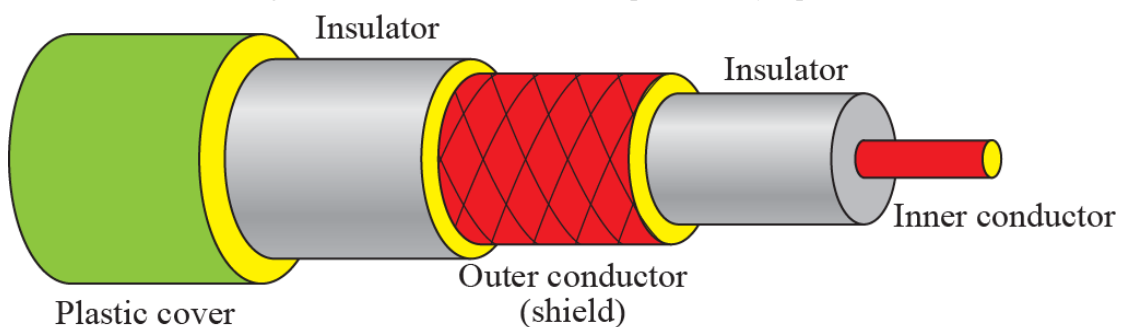
屏蔽双绞线(shielded twisted-pair, STP), IBM 生成, 除 IBM 外很少使用。

连接器(connectors): RJ-45 Male / Female (RJ: Registered Jack 注册插头)

性能: 可通过较宽的频率范围, 振幅越大直径越大, 但当频率超 100kHz 时, 衰减突升

2. 同轴电缆: 可以传输更高的频率范围

Coaxial cable carries signals of higher frequency ranges than those in twisted-pair cable. Instead of having two wires, coaxial cable has a central core conductor of solid or stranded wire (usually copper) enclosed in an insulating sheath, which is encased in an outer conductor of metal foil or braid (or a combination of the two). The outer metallic wrapping serves both as a shield against noise and as the second conductor, which completes the circuit. The outer conductor is also enclosed in an insulating sheath, and the whole cable is protected by a plastic cover.



连接器(connector)

BNC Connector: 连接到设备, 比如电视机; BNC T Connector 用于以太网引出分支连计算机或其他设备; BNC Terminator 用于电缆末端防止信号反射。

性能：频率带宽比双绞线高，但是信号衰减速度高于双绞线，所以要经常使用中继器。

3. 光缆：用玻璃或塑料构成，可以传输光信号。原理：光穿透密度不同的物质时，随着入射角(angle of incidence)的变化，当入射角小于临界角(critical angle)会产生折射(refract)，大于临界角会产生反射(reflect)。

传播模式(propagation mode)

多模(multi mode)：来自光源的多个光束沿着不同的路径传过纤芯

1. 多模阶跃折射率光纤(multimode step-index fiber)：纤芯的密度从中心到边缘保持不变，在纤芯和包层交界处，突然遇到低密度的包层导致光束传播角度的变化，不同光束角度不同，因此传播速度不同，到达目的地时重组会有失真
2. 多模渐变折射率光纤(multimode graded-index fiber)：光纤密度在纤芯中心最高，逐渐降低，到边缘最低，使得各光束传播速度差异减小，可以减少失真。

单模(single mode)：使用阶跃折射率光纤并使用更好的官员，把光束限制在更小的角度内，几乎接近水平

单模光纤(single-mode fiber)：直径比多模小得多，且密度(折射率)非常低，使得临界角几乎接近 90 度，光束的传播基本是水平的，不同光束的传播几乎相同，可以忽略延时，到目的端时可以重组为失真很小的信号。

组成：

1. 外部封套(outer jacket)用 PVC 或特氟龙
2. 封套(DuPont Kevlar for strength)高强度纤维(可用做防弹背心)
3. 塑料缓冲层(plastic buffer)
4. 光纤：由包层和纤芯构成(Gladding and Glass/plastic core)

连接器：

用户通道连接器(subscriber channel, SC)用于有线电视，用推拉式固定

直插式连接器(straight-tip, ST)用于连接网络设备，用卡口固定

MT-RJ 连接器(Mechanical Transfer Registered Jack)，与 RJ-45 规格相同

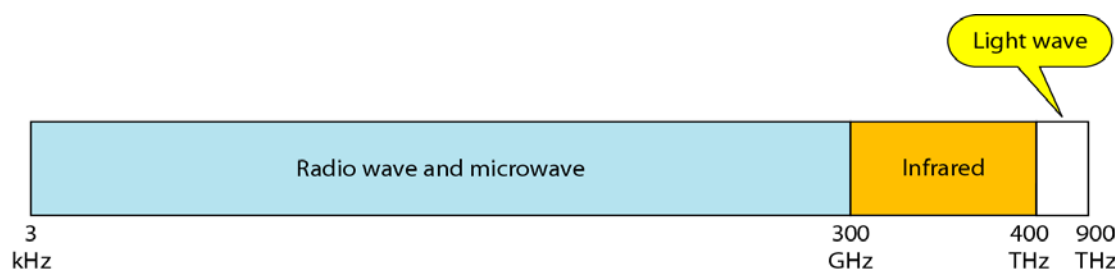
性能：

衰减极少，因此只需要很少的中继器(repeater)

● 无向介质 (unguided media)

不用物理导体传输电磁波，通常是指无线通信(wireless communication)。信号可以被任何有接收器的人接收到。

常用电磁频谱



传播方式：地表传播(ground propagation)，天空传播(sky propagation)，视线传播(line-of-sight propagation) 具体见 PPT Lecture 06 的 19 页。

无线电波(radio wave)

频率为 3-1GHz，全方向，可能收到同频率或同波段的其他信号干扰。

可以传播很远，尤其是以天空模式传播的，例如 AM。

可以穿透物体，尤其是中低频率的点播

波段比较窄（在 1GHZ 以下），导致通信速率低

各个波段都受政府机构管制（美国是 FCC）

需要**全向天线**(omni directional antennas)

微波(microwaves)

1G-300GHZ，单向，发送和接收天线要对准，对准后不受干扰。

属于视线传播，天线塔要互相可见，相距很远天线塔要高，长距离要中继

不能穿透墙体

波段相对比较宽，可以分配更多子波段，且通信速率更高

使用波段中的某个部分，需要得到政府许可

需要**定向天线**(unidirectional antennas)：**碟形天线**(parabolic dish antenna)和**喇叭天线**(horn antenna)

红外波(infrared wave)

300GHz-400THz，波长 1mm-770nm，用于短距离视线通信

不能穿透墙体，短距离内不会影响其他房间的系统

无法进行长距离通信

无法在室外使用，因为太阳射线包含可能干扰通信的红外波

波段接近 400THz，有高速传输数据的巨大潜力

九、交换

● **术语解释**

交换(switch)：一种让网络中的设备可以一对一通讯的方案，一个交换网络由一系列互联的交换机(节点)构成。在大规模的网络里，网状或星型拓扑不切实际且浪费，总线拓扑会超出介质和设备的容量。

交换机(switch)：能够在链接到交换机的两台或多台设备之间建立临时的连接。

● **三种交换方式**

电路交换网，分组交换网(含数据报网和虚拟电路网)，报文交换网络(不作讨论)

● **电路交换网络(circuit-switched network)**

由**物理链路专用通道**链接的一组交换机组成，每条链路划分成 N 个通道。在电路交换中，建立阶段(setup phase)必须预留资源，作为整个数据传输期间的专用资源，直到拆除阶段(teardown phase)。即：要真实的预留物理连接，作为两点间专用连接。

三个阶段

1. 连接建立阶段(connection setup): 通信前建立一条专用电路(通道)
2. 数据传输阶段(data transfer): 专用电路(通道)建立后, 双方可传输数据
3. 拆除阶段(connection teardown): 当任何一方要撤销连接, 就向路径中的 每台交换机发送一个信号来释放资源

效率

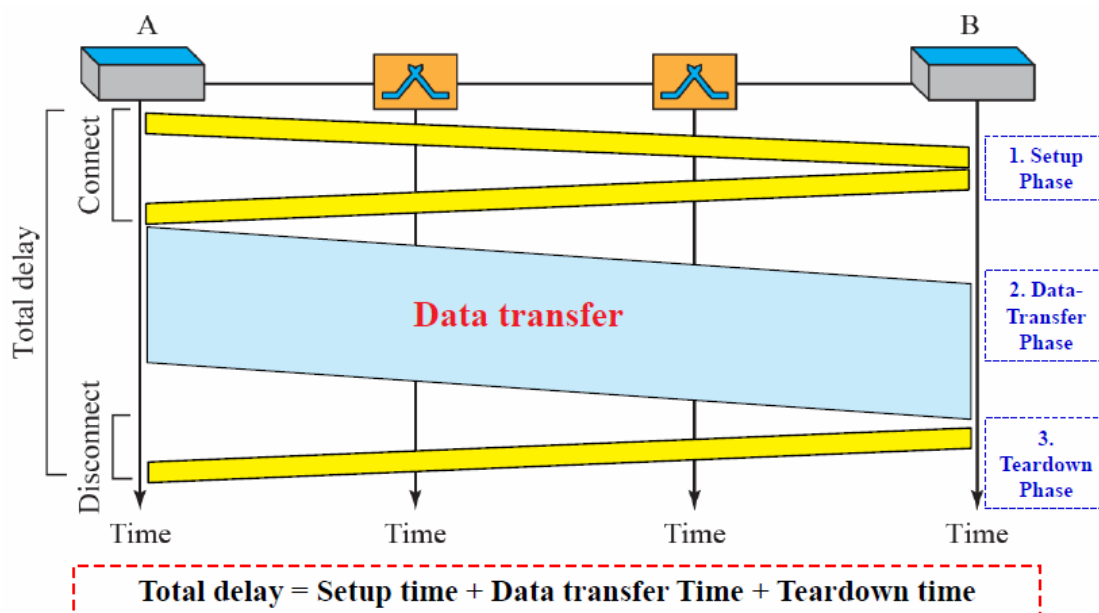
因为资源在整个连接期间都被占用, 且不能被其他连接所用, 所以电路交换没有其
他两种网络类型有效率。

延迟

虽然效率不行, 但是延迟最小。每台交换机没有等待时间。

延迟由三个部分组成:

1. 源计算机请求的传播时间 + 源计算机请求的传输时间 +
目的计算机确认的传播时间 + 目的计算机确认的传输时间
2. 数据传播时间 + 数据传输时间
3. 拆除请求传播时间 + 拆除电路所需时间

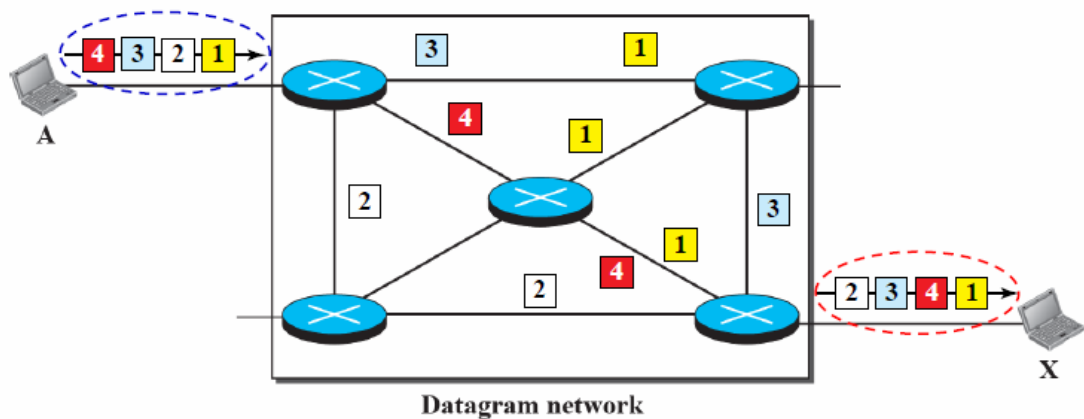


● 分组交换(Packet switching)

报文(message)要被划分为一些固定长度或者可变长度的分组, 分组长度由网络和控制协议决定。对分组不存在资源分配, 链路没有预留的带宽, 对每个分组没有安排预定的处理时间, 资源按需分配(allocated on-demand), 基于先来先服务(FCFS: first com, first serve)的原则。

- **数据报网络(datagram network)**

每个分组单独处理，与其他分组无关，即使某个分组是多分组传输的一部分，网络处理它好像是它单独存在一样，这种方法的分组称为数据报(datagram)



从上图可看出，每个分组可能走的路径都不相同，到目标节点时，顺序可能是错乱的。也可能由于资源的缺乏，导致分组丢失。

数据报网有时也被称为无连接网络(connectionless network): 交换机不保存有关连接状态的信息，不需要连接建立阶段，也不需要拆除阶段。每个分组不管源端和目的端，由交换机同样处理。

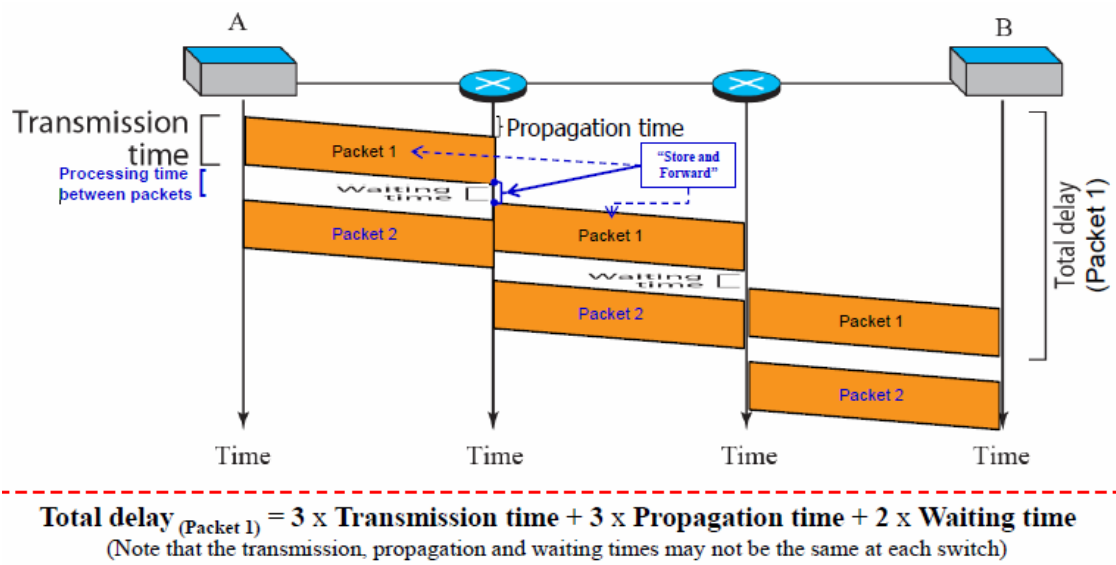
路由表

每个交换机都有一个基于目的地址的、动态、定期更新的路由表，表中记录目的地址和转发端口。数据报网分组的头部含有目的地址，在分组传送期间保持不变。

效率

比电路交换高。仅当有传输的分组时，才分配资源。在另一个分组可发送前，存在很少时间延迟。

延迟(计算方法见下图)



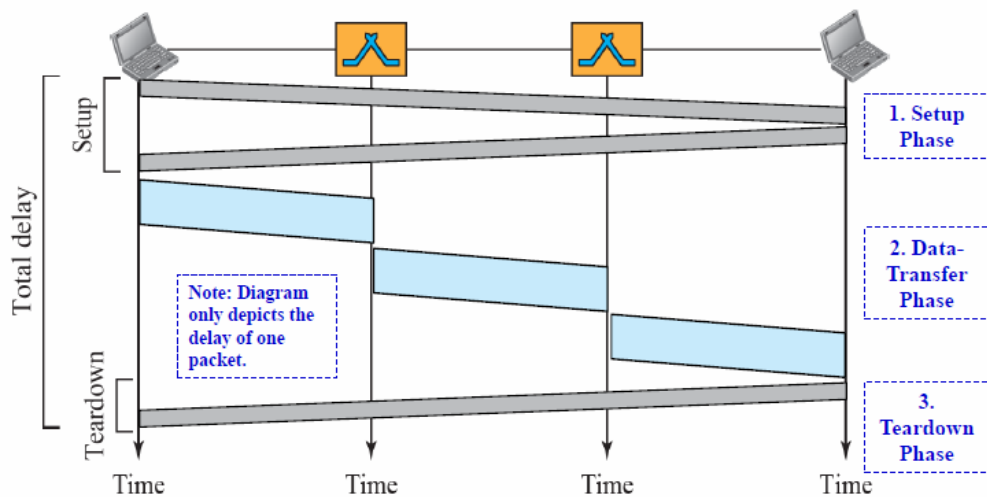
比虚电路交换网长。虽然没有建立阶段和拆除阶段，但是每个分组在转发前，在交换机处会有等待。另外，因为报文中所有分组不需要通过同一个交换机传送，所以分组的延迟不一致。

- 虚拟电路网络(virtual-circuit network)

结合了电路交换网络和数据报网络的产物，有如下特征：

1. 在数据传输阶段，有建立阶段和拆除阶段
2. 按需在建立阶段期间分配资源
3. 数据被划为分组，每个分组头部含有地址，它有本地权限(知道下一个交换机和传送分组的通道是什么，而不是终端到终端的权限)
4. 所有分组沿着连接期间建立的路径传送
5. 虚拟电路网络在数据链路层实现，而电路交换网络在物理层实现，数据报网络在网络层实现

延迟(计算方法见下图)



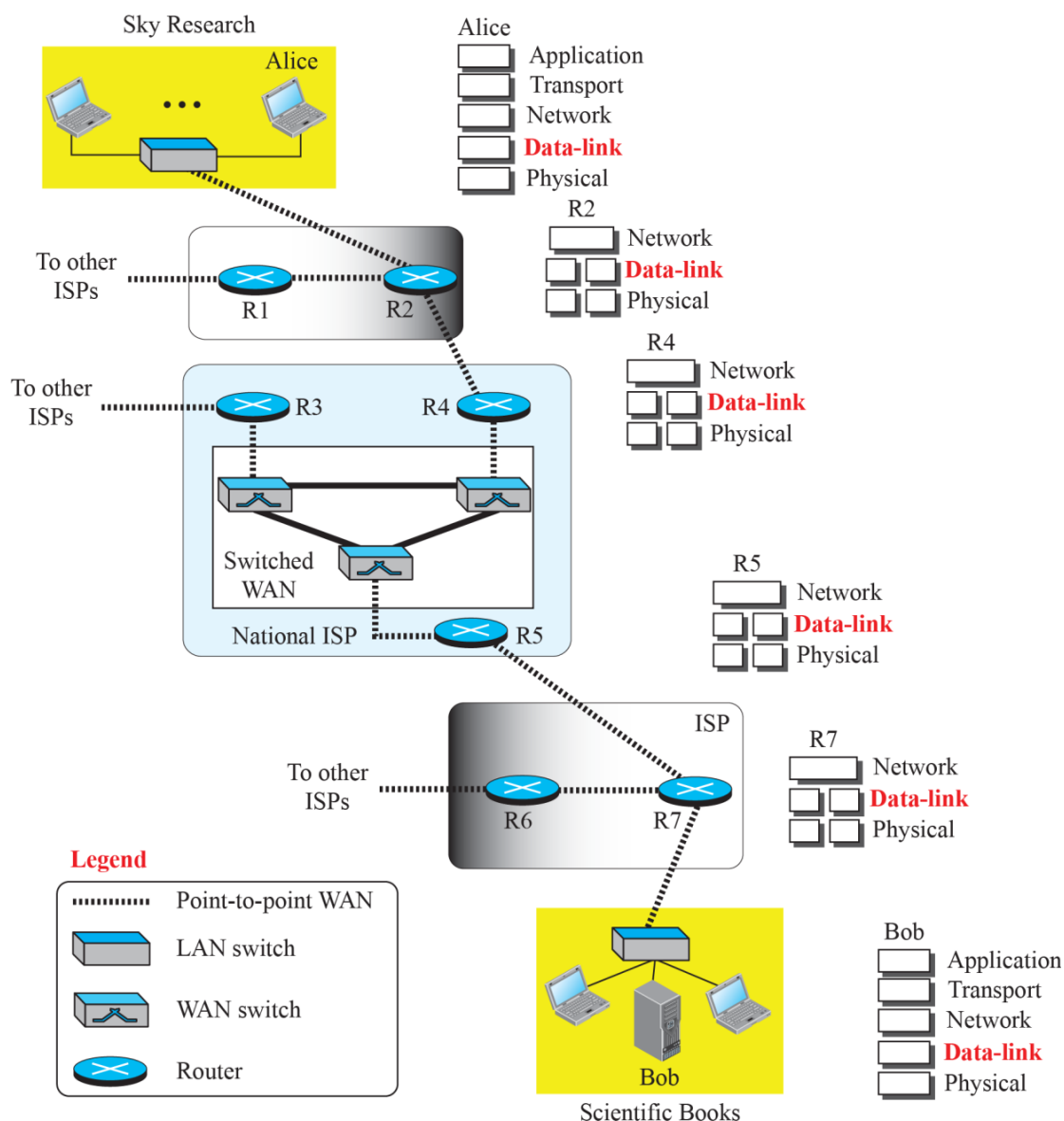
$$\text{Total delay} = \text{Setup time} + 3 \times \text{Transmission time} + 3 \times \text{Propagation time} + \text{Teardown time}$$

(Assumption: resource reservation is made during the setup phase, otherwise include additional delay for individual packets at each switch)

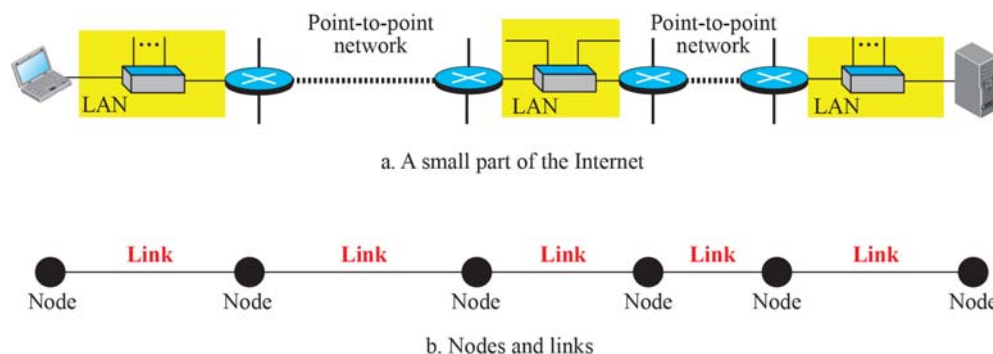
十、数据链路层介绍

● 简介

互联网是通过连接设备（路由器或交换机）粘合在一起的网络组合。如果数据包要从一台主机传输到另一台主机，它需要通过这些网络。



数据链路层的通信是节点到节点的。来自 Internet 中某一点的数据单元需要通过许多网络（LAN 和 WAN）才能到达另一点。这些 LAN 和 WAN 由路由器连接。通常将两端主机和路由器称为节点，将其间的网络称为链路。



- 数据链路层提供的服务

成帧：将从网络层接收到的数据包封装在一个帧中，然后再发送到下一个节点。从逻辑信道上接收的帧中解封数据包。不同的数据链路层具有不同的成帧格式。

流量控制：如果产生帧的速率高于处理帧的速率，接收端的帧需要在等待处理时进行缓冲。如果接收端的缓冲区已满 → 丢帧/反馈到发送节点。

错误控制：帧容易出错，部分原因是电磁信号容易出错。错误需要被检测到，然后 (i) 纠正或 (ii) 丢弃并由发送节点重新传输。

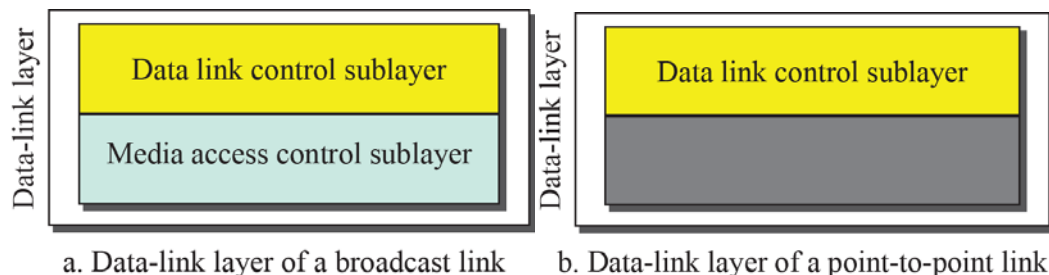
拥塞控制：当链路被帧拥塞时，一些广域网可能会使用拥塞控制。主要由网络层和传输层处理。

- 两种链路

数据链路层控制如何使用介质。共有两种方式：

1. 使用介质的全部容量
链路属于两个设备专有
点对点连接（比如用传统网络打电话）
2. 使用介质的部分容量
链路被多对设备共享
广播链接（比如用手机打电话，空气这个介质被共享）

- 两个子层

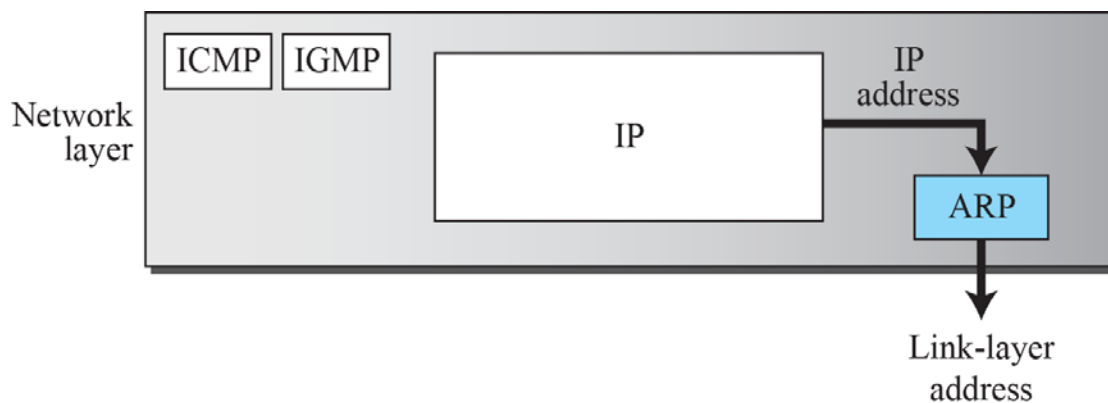


数据链路控制层(DLC, Data link control): 处理点对点和广播链路的所有问题

媒体访问控制层(MAC, media access control): 仅仅处理广播链路的问题

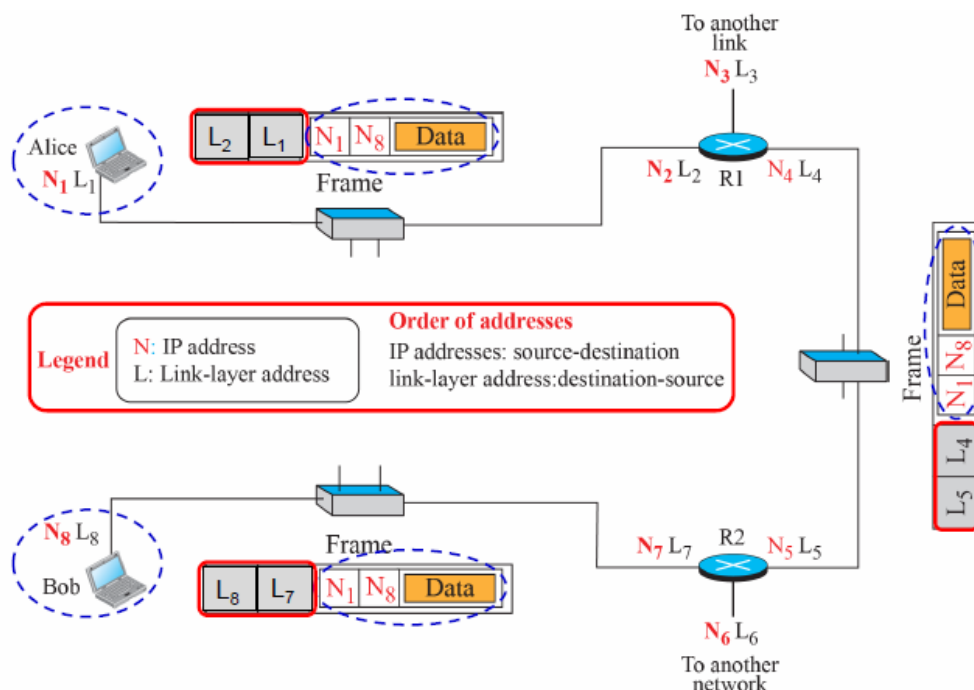
- ARP 协议

Address Resolution Protocol (ARP)在网络层被定义，它把 IP 地址解析为 MAC 地址，并传送给数据链路层。



链路层寻址

在无连接的互联网络(比如因特网)上, 数据包仅有 IP 地址是无法达到目的地的。原地地址和目标地址并没有定义报文走什么路径。相同源主机发给相同目的主机的报文, 走的路径也可能各不相同。



如上图, 报文里面的 IP 地址(N1/N8)不应被修改。数据链路层的寻址机制会把报文封装成帧并添加数据链路层地址(MAC 地址, 如 L2/L1)。每当这个帧从一个链路移动到另一个链路时, 帧里的 MAC 地址就会改变。

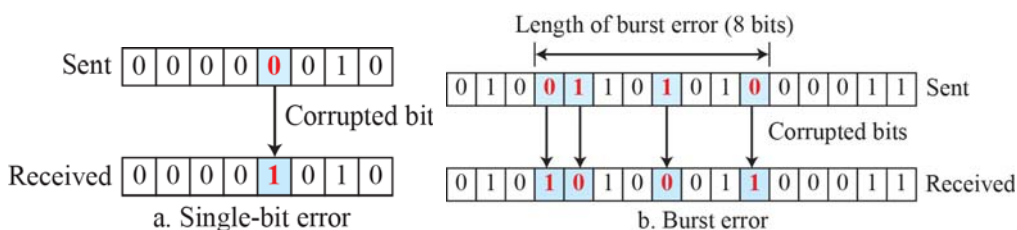
十一、检错与纠错

● 错误发生的原因

比特从一个点流向另一个点时, 会因为干扰(interference)发生不可预测的变化, 这种干扰可能改变信号的形状。被影响的位数取决于数据的速率和噪声持续的时间。

● 差错的类型

1. 单个位差错(single-bit error): 给定的数据单元(一个字节、字符或分组)中仅一位出错
2. 突发性差错(burst error): 数据单元中有两位或更多位发生错误



突发性差错更容易发生, 因为噪音持续的时间往往大于单个位持续的时间, 例如:

1kbps 的传输, 如果噪声持续 1/100 秒, 则影响 10 个位

1Mbps 的传输, 如果噪声持续 1/100 秒, 则影响 10,000 个位

- 冗余
为了检错和纠错，除了发送数据位，还要发送一些额外的位。这些冗余的位由发送方添加，并由接收方去除。它们的存在允许接收方检测或纠正被破坏的位。
- 块编码
把消息划分成块，每个块有 k 位，增加 r 和冗余位，使得长度为 $n = k + r$ ，形成的这个 n 位的块称为码字(codewords)，码率(code rate)是 k / n

差错检测

需要满足两个条件，接收方就能检测出原来码字的一个差错：

1. 接收方有(或能找到)有效的码字列表
2. 原来的码字已变为无效的码字

如果错误的码字是一个有效的码字，那么接收方将无法检测

- 线性块编码(LBC, linear block codes)
任何两个有效码字的异或 (XOR)，生成另一个有效码字，例如：000，011，101 和 110

$$011 \oplus 101 = 110; 011 \oplus 110 = 101; 101 \oplus 110 = 011$$

最小汉明距离：具有最小 1 的个数的非 0 有效码字中 1 的个数。上面这个是 2

简单奇偶校验编码(simple parity check code)

$n=k+1$ ， k 位数据(data word)加上 1 位奇偶位(parity bit)变成 n 位码字，使得码字中 1 的总个数为偶数或奇数。例如下面这个最小汉明距离为 2 的码字表

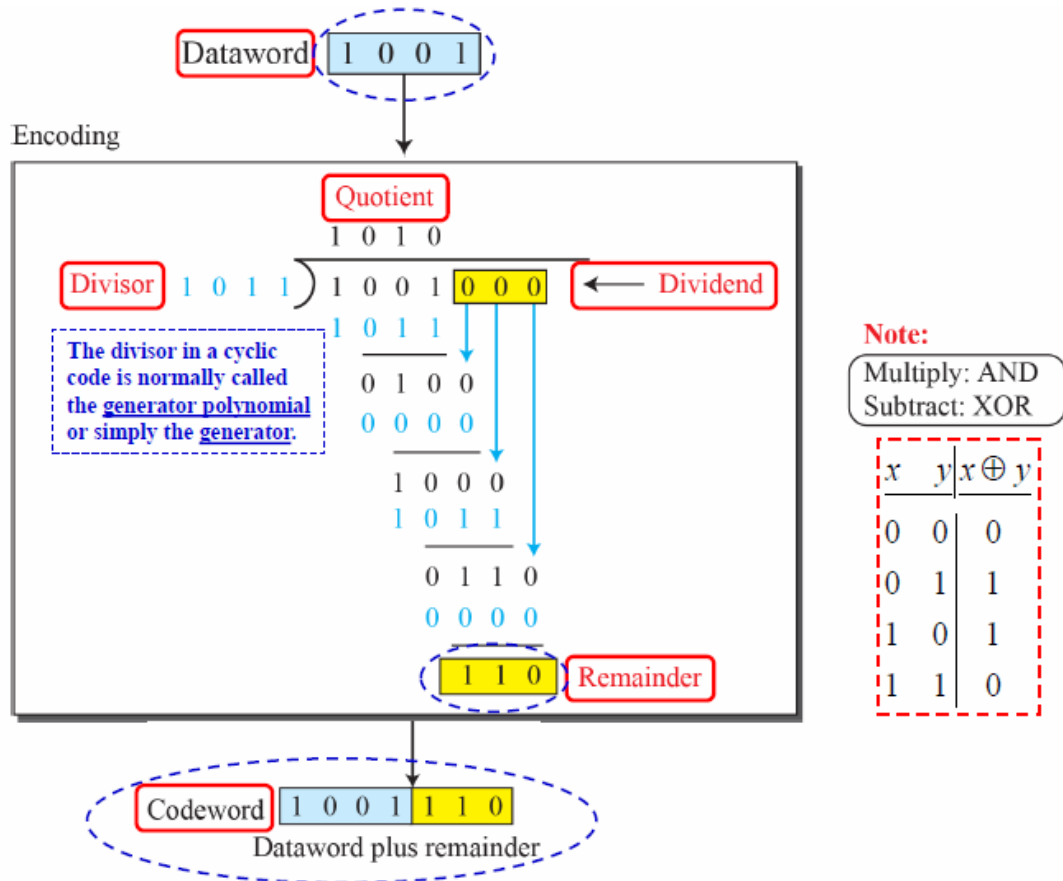
<i>Datawords</i>	Codewords	<i>Datawords</i>	Codewords
0000	00000	1000	10001
0001	00011	1001	10010
0010	00101	1010	10100
0011	00110	1011	10111
0100	01001	1100	11000
0101	01010	1101	11011
0110	01100	1110	11101
0111	01111	1111	11110

- 循环编码(cyclic codes)
一种页数的线性编码。可以假想码字是一个环，可以循环左移或右移。例如：1011000
左移得到 0110001。

- 循环冗余校验(CRC, cyclic redundancy check)

广泛被用于网络，比如 LAN 和 WAN 中。下图是一个 n 为 7， k 为 4 的例子，同时具有线性特性和循环特性。

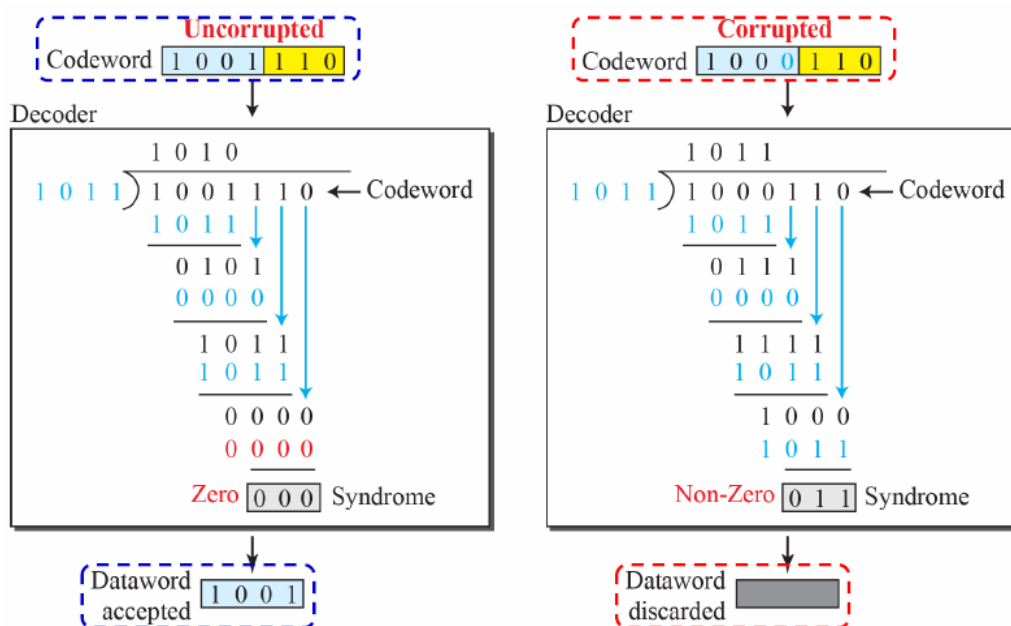
编码器



1. 除数是两方约定好的，长度是 $n-k+1$
2. 首先要给被除数加上 $n-k$ 个 0
3. 每一步的目的都是让被除数的第一位变成 0
4. 过程不用减法，用异或

译码器

用 codeword 执行相同的除法操作，如果最后余数是 0，则校验成功，否则校验失败



● 校验和(checksum)

一个检错方法，可以用于任意长度的报文(message)。多用于网络层和传输层，而不是数据链路层。

例如发送 5 个 4 位的数字 {7, 11, 12, 0, 6}，那么我们多发送一个数字，是这 5 个数字的和，也就是 36。接收方对其进行校验，成功则抛弃 36，接收前 5 个数字，否则认为校验出错。

然而，36 是一个 6 位的数字，4 位无法存储，这时候我们用反码(one's complement)算法，具体操作是：把左边多出来的几位取出来，加到右边去。

$$100100_2 \Rightarrow 10_2 + 0100_2 = 0110_2 = 6_{10}$$

接收方用相同算法来校验是否能得到相同结果。

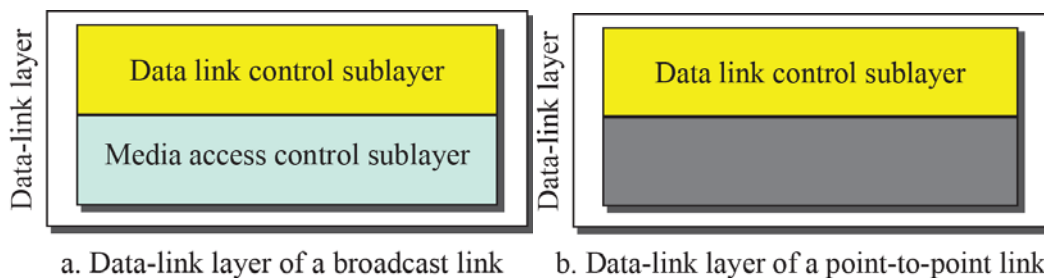
● 因特网校验和

发送方	接收方
1. 报文划分为 16 位的字	1. 报文(含校验和)划分成 16 位字
2. 校验和的值设置为 0	2. 用反码将所有字相加并反码运算
3. 所有字包括校验和用反码运算相加	3. 对该校验和求反生成新校验和
4. 对这个和求反变成校验和	4. 如果此校验和是 0，则通过校验，否则丢弃
5. 校验和随着数据一起发送	

例如发送 5 个 4 位的数字 {7, 11, 12, 0, 6}，那么先全相加，再求反，得 6，再用补码 $15-6=9$ ，然后发送 {7, 11, 12, 0, 6, 9}。接收方把他们全相加得到 45，再求反得到 15，再求补码得 0

十二、数据链路控制

● 两个子层



数据链路控制层(DLC, Data link control)

处理点对点和广播链路的所有问题(成帧, 流量控制, 错误控制)

媒体访问控制层(MAC, media access control):

仅仅处理广播链路的问题(介质访问控制, 共享链路)

● 成帧

把位打包成帧, 使得帧互不相同。帧可以是固定或可变大小的, 每个帧被加上了发送方和目标地址。目标地址定义了分组要去的地方, 发送方的地址帮助接收方确认接收。

固定大小成帧

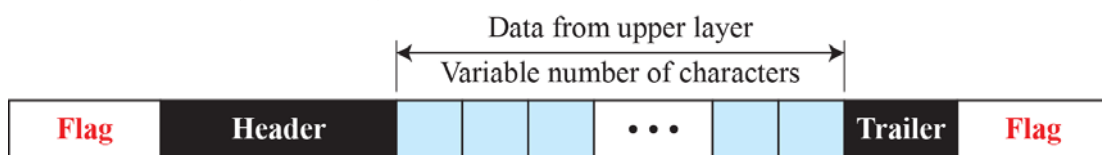
不需要定义帧的边界, 长度本身即可作为分隔符。

可变长度成帧

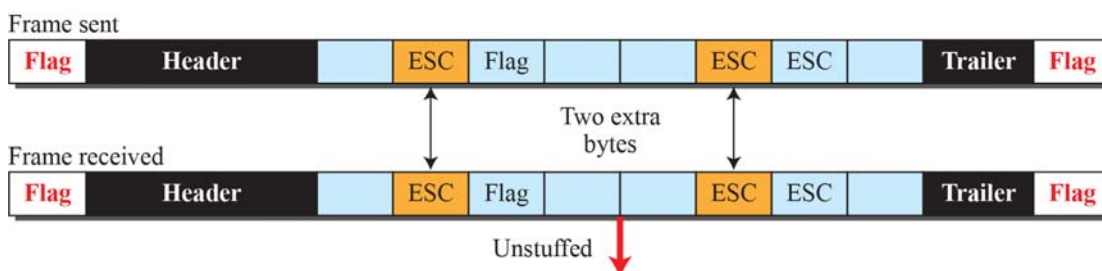
需要定义一个帧的结束, 以及下一个帧的开始。包括“面向字符协议”和“面向位协议”

面向字符协议

数据采用 8 位传输(比如 ASCII), 头尾添加**标记**, 如下图



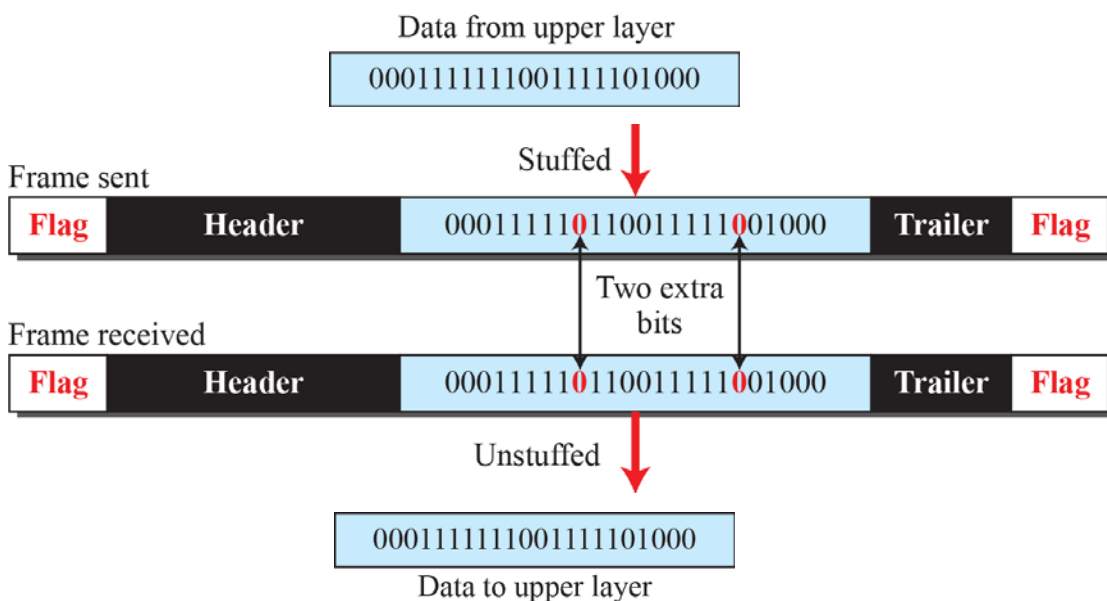
当只传送文本时, 可以用非文本的字符来做标记, 但是传输影音等其他信息时, 可能会把信息的内容当做起止标记, 所以要使用**字节填充(byte-stuffing)**, 这个字节被成为**换义字符(escaped character, ESC)**



但如今的编码系统(例如 Unicode)是 16 和 32 位字符, 会与 8 位产生冲突, 因此“面向位协议”是大势所趋。

面向位协议

数据是一连串的位，大部分的协议用一个特殊的 8 位标志做起止，比如 01111110 作为分界(delimiter)。为了防止数据中出现标志，数据中出现一个 0 后面连续跟着 5 个 1 时，全自动补一个 0，如下图所示：

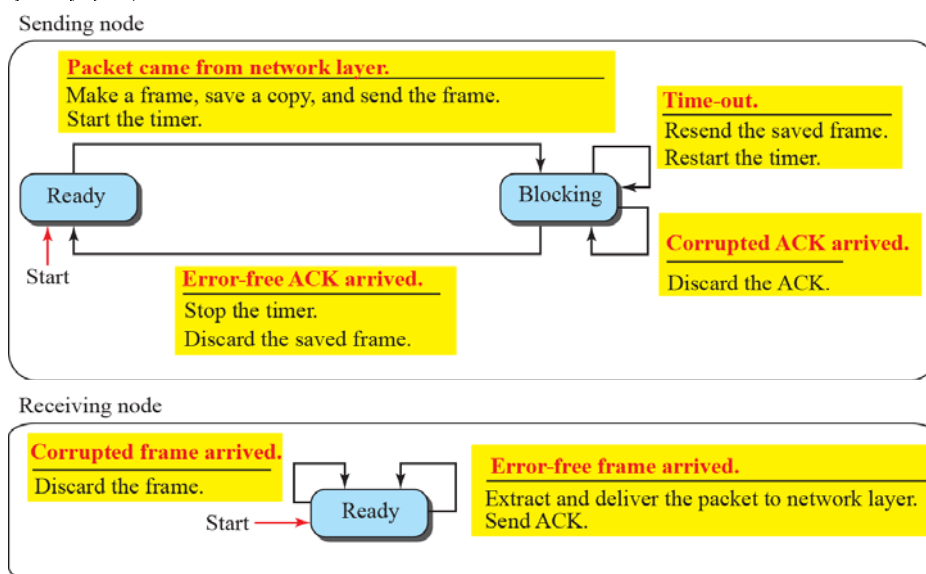


● 数据链路层协议

停止等待协议(stop-and-wait protocol)

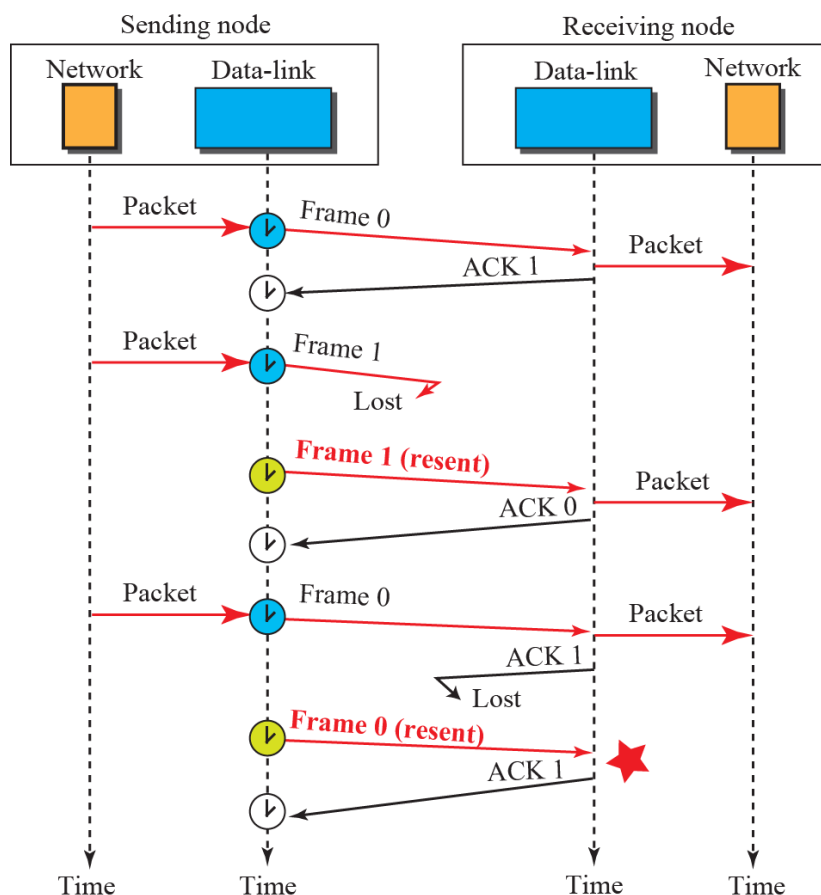
1. 发送方每次发一个帧，并且在发送下一个帧之前等待接收方发来的确认。
2. 为检测错误帧，每个数据帧都会加上 CRC 校验
3. 每次发送一个帧，发送方都会启动一个倒计时
 - a) 如果倒计时结束前收到回应，发送方将发送下一个帧
 - b) 如果定时器超时，则发送方认为数据丢失或损坏，并再次发送这个帧
4. 在通道在任意时刻，仅有一个帧和一个确认消息

状态机如下



发送方何时重发

1. 如果发送的帧丢失或损坏，则接收方不会发送 ACK 确认
 2. 如果接收方收到了帧，但是发回的 ACK 丢失、损坏、或超时到达
- 此时发送方再次发送的帧会导致数据被重复接收和处理，所以要在帧上加上队列编号，并在 ACK 回应上加上回应编号，可防止帧数据被重复处理



十三、媒体访问控制

● 随机访问(random access, 又称为竞争 contention)

1. 每个站点都是平等的，无法控制其他站点。
2. 一个有数据要发送的站点，要用协议定义的程序(procedure)来决定是否发送
3. 是否发送的决定却绝育介质的状态是忙碌还是空闲
4. 当一个以上的站点想要发送时，会产生冲突(collision)，会有帧被损坏或修改

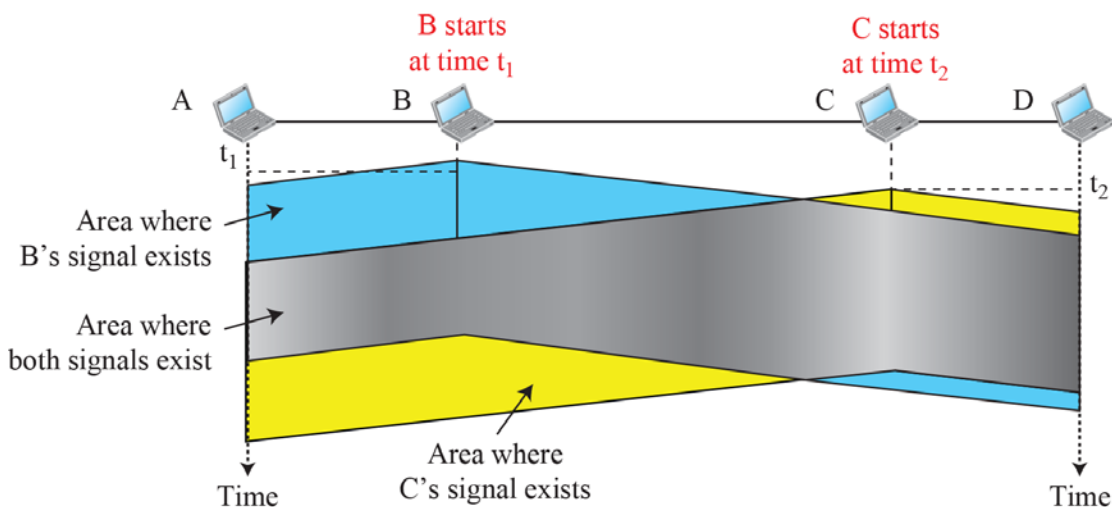
为什么叫做随机访问

1. 站点的传输是随机的，没有时间表
2. 为访问介质，站点展开竞争

● 载波侦听多路访问(CSMA, Carrier sense multiple access)

为减少冲突的发生从而提高性能，我们采用 CSMA 方法。站点在传送之前对介质进行侦听(a senses the medium before trying to use it), 基于“传输前侦听”或“先听再说话”原理。

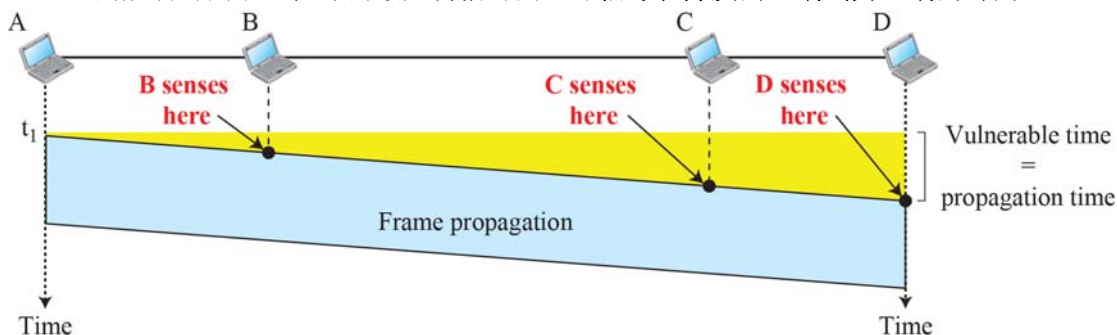
CSMA 可以减少冲突的概率，但不能消除冲突，因为传播有延时。当站点发送一帧时，它的第一个位到达每一个站点，以及站点检测到它的信号，都需要一定时间。



上图中 t_1 时，B 发送帧，到 t_2 时还没传播到 C，于是 C 发送帧

脆弱时间(vulnerable)

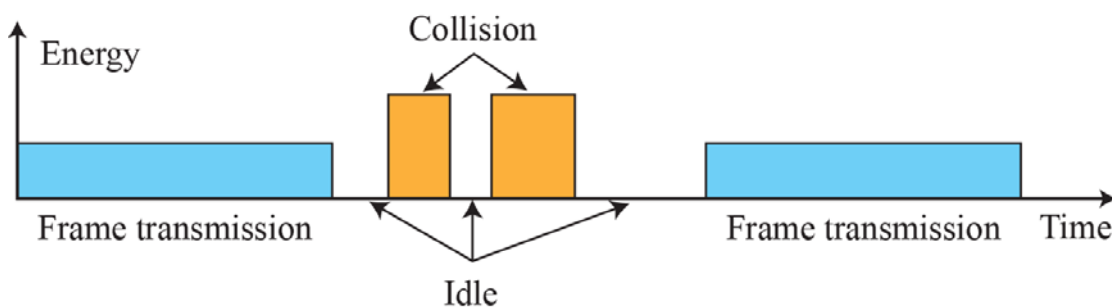
CSMA 的脆弱时间就是信号的最大传播时间，即信号从介质的一端到另一端的时间



● CSMA/CD (Collision Detection)

任何站点都可以发送帧，然后监控介质看传输是否成功；如果发送成功，则完成发送；如果检测到冲突，则需要重新发送此帧。

能量级别(Energy Level, 有三个值)

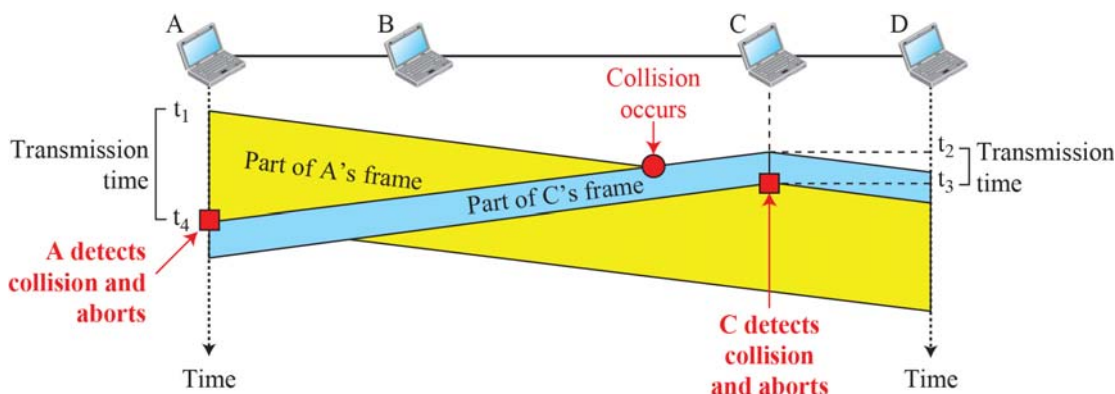


级别 0: 通道空闲

正常级别: 通道忙，一个站点占用了通道且在发送一个帧

不正常级别: 通道发生了冲突，且能量级别是正常级别的两倍

对冲突的处理图



t1: A 开始传输帧

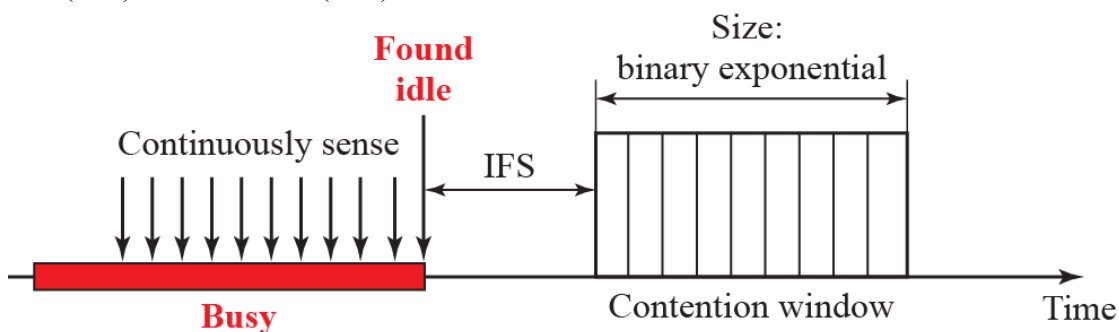
t2: C 没有收到来自 A 帧的第一位，于是也开始传输，在 t_2 后的某个时间，冲突发生在半路

t3: C 检测到了来自 A 帧的第一位，立即放弃传输，总传输时间 $t_3 - t_2$

t4: A 检测到了来自 C 帧的第一位，立即放弃传输，总传输时间 $t_4 - t_1$

● CSMA/CA(Collision Avoidance)

是为无线网络发明的。有线网络的信号能量等级一致，所以检测能量很容易；但是无线网络能量很小，难以检测。所以 CSMA/CA 用三种策略来避免冲突：帧间间隔(IFS, the interframe space), 竞争窗口(contention window)和确认(acknowledgements), 还有 Ready to Send(RTS)和 Clear to Send(CTS)帧。



帧间间隔(IFS)

用推迟传输来避免冲突：即使通道是空闲的，站点也继续等待，因为其他的站点可能已经开始传输。如果在 IFS 时间后，通道依然空闲，此时再发送。

竞争窗口(Contention Window)

竞争窗口把时间分割为时隙。准备传输的站点随机选择一个时隙(slot)数量作为等待的时间，每次站点在 IFS 时间后无法检测到空闲通道时，时隙的数量要翻番(binary exponential)。每次时隙后都要侦听通道，如果通道繁忙，不重启程序，只是暂停定时器并当通道空闲时，再重启定时器。

确认(Acknowledgement)

“肯定”的确认和定时器超时将有助于确保接收方收到帧。

十四、网络层

- 网络层服务(network-layer services)
 1. 打包(packetizing): 在发送方封装来自上层的数据, 在接收方解封来自下层的数据
 2. 路由(routing): 发送方和目的地之间可能有多个路由, 网络层用策略和协议来选择最佳路径并为每个路由器创建路由表
 3. 转发(forwarding): 路由器从一个连接收到数据后, 查转发表转给连到它的其他连接
- 数据包交换(packet switching)

路由器在输入输出端口之间创建了连接, 就像电路交换器连接输入输出端口使电流流动
- IPv4 寻址

IPv4 地址是一个 32 位地址, 它唯一地与通用地定义了一个连接在新特网上的设备(如主机或路由器)。IP 地址是连接的地址, 不是设备的地址。
- 地址空间

地址空间是该协议能用的地址总个数, IPv4 的地址空间是 2^{32} 。
IP 地址可以用二进制、点分十进制和十六进制标记(notated)

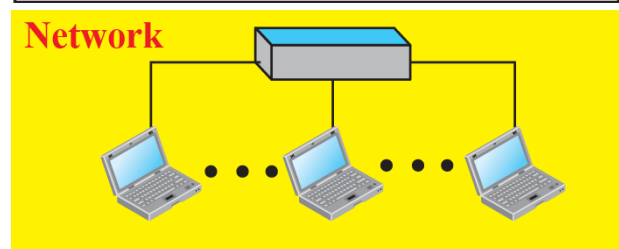
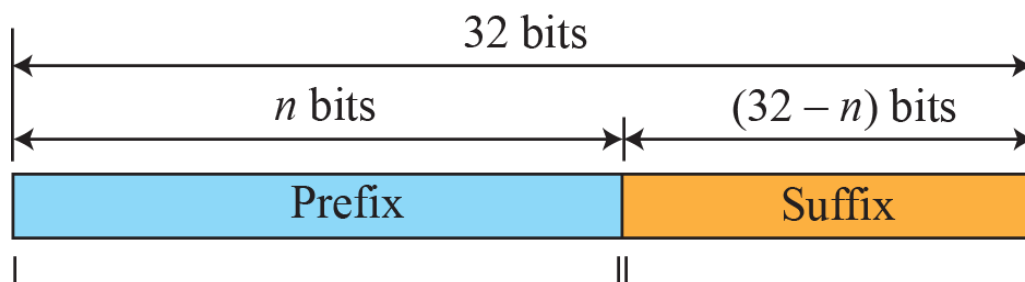
Binary 10000000 00001011 00000011 00011111

Dotted decimal 128 • 11 • 3 • 31

Hexadecimal 80 0B 03 1F

- 一个 IPv4 地址被分成两个部分

第一部分称为前缀(prefix), 拥有固定或可变的长度, 用于定义网络
第二部分称为后缀(suffix), 定义了连接到节点的连接



- 分类寻址

CIDR (Class Interdomain Routing)例子: 12.24.76.8/8

例: 在 167.199.170.82/27 中有多少个 IP 地址, 首位 IP 地址分别是?

$2^{(32-27)} = 32$ 个

81 = 01010010, 其中后 5 位是给予网用的, 子网从 64 (01000000) 开始到 95 结束

路由器收到数据包后, 会把目的地的 IP 地址与子网掩码进行与运算, 然后找到与其最高位匹配得最多的 1 那个接口并发送

十五、网络层协议

- IPv4

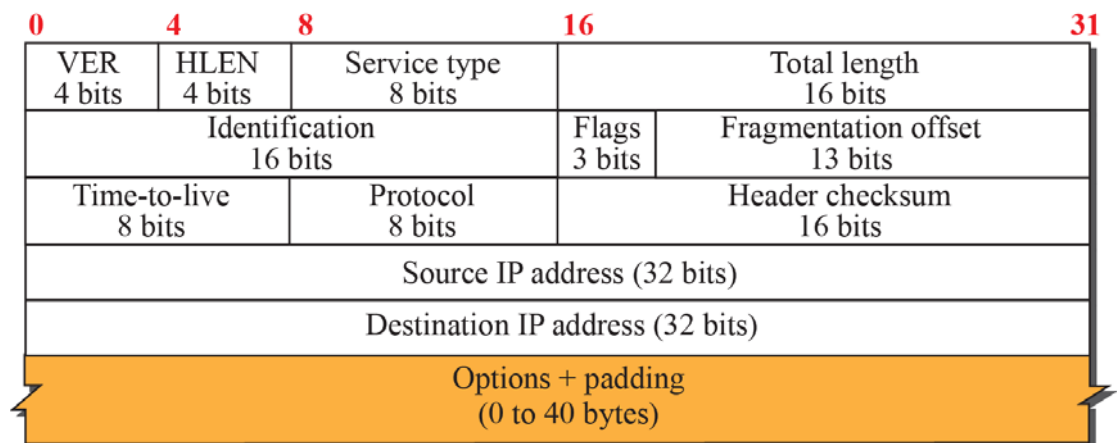
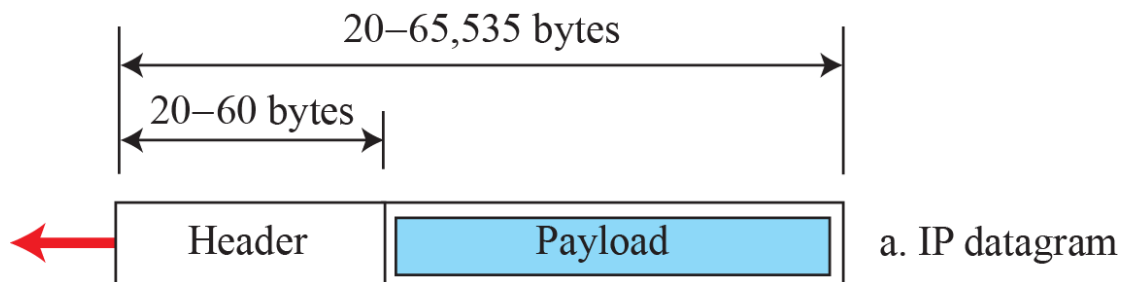
负责打包、转发、传输一个数据包, 它是不可靠的、无连接的协议

- 互联网控制消息协议(Internet Control Message Protocol version 4, ICMPv4)

是一个网络层协议, 配合 IPv4 处理一些可能在传输中发生的错误

- IP 使用的数据包(packets)叫做数据包(datagrams)

长度可变。分为头部(header)和数据(payload), 头部为 20-60 字节, 含有路由和传输的关键信息。TCP/IP 习惯把 IP 头以 32 位的形式来展现。**HLEN 是包头的 DWORD 数(*4B)**



Legend

VER: version number
HLEN: header length
byte: 8 bits

ICMP: 01 UDP: 17
IGMP: 02 OSPF: 89
TCP: 06

Some protocol values

- 校验和
分割为 16 位，以因特网校验和来计算

An example of a checksum calculation for an IPv4 header without options is shown:

16 bits			16 bits		
4	5	0	28		
49 153			0	0	
4	17	0			
10.12.14.5					
12.6.7.9					

4, 5, and 0	→	4	5	0	0
28	→	0	0	1	C
49153	→	C	0	0	1
0 and 0	→	0	0	0	0
4 and 17	→	0	4	1	1
0	→	0	0	0	0
10.12	→	0	A	0	C
14.5	→	0	E	0	5
12.6	→	0	C	0	6
7.9	→	0	7	0	9
Sum	→	1	3	4	4 E
Wrapped sum	→	3	4	4	F
Checksum	→	C	B	B	0

The header is divided into 16-bit sections. All the sections are added and the sum is complemented after wrapping the leftmost digit. The result is inserted in the checksum field.

Note that the calculation of wrapped sum and checksum can also be done as follows in hexadecimal:

$$\begin{aligned} \text{Wrapped Sum} &= \text{Sum mod FFFF}_{16} \\ \text{Checksum} &= \text{FFFF}_{16} - \text{Wrapped Sum} \end{aligned}$$