这是我考试前自己总结的，在考试前我基本能够记下这里面的所有内容，到了这个程度，其实考试已经非常轻松了。

# chap1

历史：

1. 1960 – 1964：分组交换网络。Baran：自适应系统
2. 1969年，ARPANET诞生
3. 1970年代：X.25 分组交换网，专用的网络体系结构：SNA，DNA；telnet
4. 1979年，TCP/IP成熟
5. 1980年代：ISO/OSI，LAN，Internet 初具规模；FTP，NFS，Email，USENET
6. 1990年代：局域网：Novell, TCP/IP，X.25广域网及其应用，Internet在中国开始大规模发展，CERNET；WWW
7. 今天的互联网：健壮性、适应性和互联程度都下降了（例如NAT）

IETF负责标准的制订、维护和协调。RFC由IETF提交。不是所有的RFC都是Internet标准。1969 年ARPANET 时就开始发布RFC

# chap2

计算机网络的相关定义：

1. 计算机网络：一批独立自洽的计算机系统的互连集合体
   1. 对比：通信网络：重点研究通信终端（电话等）与通信网络
2. 资源子网：服务器，客户计算机；通信子网：通信线路（或称信道），网络互连设备（路由器、交换机、HUB等）
3. 基本通信方式：交换式（主要用于城域网和广域网，关键技术是路由选择），广播式（主要用于局域网，关键技术是通道分配）
4. 计算机网络的体系结构：对计算机网络及其部件所完成功能的比较精确的定义，即从功能的角度描述计算机网络的结构，是层次和层间关系的集合
   1. 仅定义功能，不定义协议实现和各层协议间的接口
5. 对等进程：位于不同计算机上进行对话的第N层通信各方可分别看成是一种进程，称为对等（同等）进程
6. 协议：计算机网络同等层次中，通信双方进行信息交换时必须遵守的规则
   1. 由语法，语义，定时关系组成
   2. 目的主机第N层收到的报文与源主机第N层发出的报文相同
7. 服务：同一实体上下层间交换信息时必须遵守的规则
8. 接口：定义了下层向上层提供的原语操作和服务
   1. 服务和接口的区别在于，服务是双向的，接口是单向的
9. SAP：服务访问点。任何层间服务是在接口的SAP上进行的。每个SAP有唯一的识别地址，每个接口可以有多个SAP。
10. IDU = ICI + SDU。IDU是通过SAP进行传送的层间信息单元。
11. PDU = PCI + SDU（或SDU的分段）。PDU是第N层实体通过网络传送给它的对等实体的信息单元。涉及分段和重组
12. 服务原语：服务在形式上是由一组接口原语（或操作）来描述的
    1. 包括请求，指示，响应，确认。

ISO/OSI模型

1. 物理层（The Physical Layer）：在物理线路上传输原始的二进制数据位
2. 数据链路层（The Data Link Layer）：在有差错的物理线路上提供无差错的数据传输（Frame）
3. 网络层（The Network Layer）：控制通信子网提供源点到目的点的数据传送（Packet）
4. 传输层（The Transport Layer）：为用户提供端到端的数据传送服务
5. 会话层（The Session Layer）：为用户提供会话控制服务（安全认证），例如令牌管理和同步
6. 表示层（The Presentation Layer）：为用户提供数据转换和表示服务
7. 应用层（The Application Layer）

TCP/IP模型

1. Host-to-Network = 物理层 + 数据链路层
2. Internet层（网络层）：控制通信子网提供源点到目的点的 IP 包传送，实现异构网络互联
3. 传输层：提供端到端的数据传送服务。TCP 和 UDP
4. 应用层：提供各种 Internet 管理和应用服务功能

ISO/OSI模型晚于TCP/IP模型

NetWare是Novell公司开发的PC上的网络操作系统，client-server结构。基本思想：文件共享。

Novell NetWare：

1. 网络层协议IPX ：不可靠无连接协议，与IP类似，但地址长度为10位
2. 传输层协议NCP，SPX ：面向连接的协议

X.25 分组交换网的特点：面向连接，支持交换虚电路和永久虚电路。

1. 物理层协议：X.21，X.3 / X.28 / X.29

2. 数据链路层协议：LAP, LAPB

3. 网络层协议：PLP

DTE: Digital Terminal Equipment；DCE: Digital Circuit Terminating Equipment；PAD: Packet Assembler and Disassembler

B-ISDN 和 ATM：

多种网络共存，电信公司想统一成一个网络B-ISDN（宽带综合业务数字网）。技术基础是异步传输模式ATM。

ATM（Asynchronous Transfer Mode）：面向连接。异步传输，没有主时钟。传输单元是短的、定长的包，称为信元。速率主要有两种：155M，622M。采用虚电路交换。

# chap3

频谱是一个信号包含的频率的范围。绝对带宽等于频谱的宽度（上界-下界）。信号的主要能量集中在相对窄的频带内，这个频带被称为有效带宽，或简称带宽。

信号的信息承载能力与带宽有直接关系，带宽越宽，信息承载能力越强。

信道有截止频率fc，0 ~ fc的振幅衰减较弱，fc以上的振幅衰减厉害。

通过信道的谐波次数越多，信号越逼真。一次谐波的频率f1用传输速率直接计算（若8位为一个周期Ｔ，则一次谐波的频率是：f1 = B/8 Hz），能通过信道的最高次谐波数目为：N = fc / f1，要求N>=1, 有限的带宽限制了数据的传输速率。

奈魁斯特：最大传输速率=2Hlog2(V) bps。

香农：最大传输速率=Hlog2(1+S/N) bps，信噪比为10log10(S/N) db。

各种传输方式分类

1. 点到点传输/点到多点传输
2. 单工、半双工和全双工传输：注意单工传输中，监视信号也可以回送
3. 同步传输/异步传输：
   1. 同步：以报文为单位。以同步字符使收发双方同步，信息中的同步字符需要透明传输处理。接收方必须知道每一位信号的开始及其持续时间
   2. 异步：以字符为单位，需要辅助位，每个字符可能需要用10位或11位才能传送。效率低，用于字符终端和计算机之间的通信。
   3. 基于字符的信息传送中，可以采用异步传输，也可以采用同步传输
   4. 基于位的传输，一般采用同步传输
4. 数字传输/模拟传输：
   1. 模拟数据 <-电话系统-> 模拟信号

数字数据 <-调制解调器MODEM -> 模拟信号

模拟数据 <-编码解码器CODEC -> 数字信号

数字数据 <-数字编码解码器->数字信号

* 1. 基带传输：在传输时直接使用基带信号。基带：基本频带，是原始信号所固有的频带。使用于低速高速的各种情况。限制：因基带信号所带的频率成分很宽，所以对传输线有一定的要求
     1. NRZ：0就是0，1就是1。难以分辨位的结束和开始。发送方和接收方必须有时钟同步。连续的0/1会导致信号直流分量累加。
     2. 曼彻斯特码：每一位中间跳变。0->1表示0，1->0表示1
     3. 差分曼彻斯特码：每一位中间跳变表示时钟。开始时有跳变表示0，无跳变表示1。时钟、数据分离，便于提取
     4. NRZ1 / NRZ0：在每位开始时，逢1/0跳变，否则不跳变
  2. 频带传输：数字信号的模拟传输。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（控制信号的参数）。
     1. 调幅ASK，调频FSK，调相PSK
  3. 脉冲代码调制PCM：模拟数据的数字传输。根据Nyquist原理进行采样
     1. 将模拟信号振幅分成多级（2^n），每一级用 n 位表示。T1 载波将模拟信号分成128级，每次采样用7位二进制数表示
     2. 差分脉冲代码调制：根据前后两个采样值的差进行编码
     3. delta调制：根据前后两个采样值的差进行编码，但每次只输出0/1表示下降/上升。编码速度跟不上变化太快的信号。

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。

多路复用技术：TDM，FDM，WDM。T1载波（1.544Mbps）使用TDM，分成 24 个信道

交换技术：动态地接通或断开通信线路

1. 电路交换：直接利用可切换的物理通信线路，连接通信双方。
   1. 建立电路，传输数据，拆除电路。建立电路时间较长，数据传送延迟较短。
   2. 一般采用时分复用。时间被分为帧（frame），帧被分为时槽（slot）。非永久会话需要动态绑定时槽到一个会话
2. 报文交换：信息以报文（逻辑上完整的信息段）为单位进行存储转发
   1. 线路利用率高，但延迟长，要求缓冲大。
3. 分组交换：信息以分组为单位进行存储转发。分组：比报文小的信息段，可定长，也可变长。采用统计复用，按需分配资源。
   1. 数据报交换（这个名字并不好记…）：每个分组均带有网络地址，可走不同的路径
      1. 每个分组都要做路由选择
      2. 可拓展性更好
   2. 虚电路交换：来自同一流的分组通过一个预先建立的路径（虚电路）传输
      1. 建立连接时做一次路由选择
      2. 建立时发带有全称网络地址的呼叫分组，但分组头不需要包含完整的地址信息
      3. 路由器需要维护虚电路的状态信息
      4. 确保分组的顺序

对各类通信子网定义下列参数：

N = 两个给定站点之间所经过的段数；

L = 报文长度（L为分组大小P的整数倍），单位：位；

B = 所有线路上的数据传输速率，单位：位/秒；

P = 分组大小（P <= L），单位：位；

H = 每个分组的分组头，单位：位；

S1 = 线路交换的呼叫建立时间，单位：秒；

S2 = 虚电路的呼叫建立时间，单位：秒；

D = 各段内的传播延迟，单位：秒。

计算延迟：

1. 电路交换：S1+L/B+ND
2. 报文交换：N(L/B+D)。每一级都必须收到完整报文后才传递给下一级
3. 数据报交换：((P+H)/P)(L/B)+ND+(N-1)(P+H)/B。传输时需要考虑头的overhead。
4. 虚电路分组交换：S2+L/B+ND+(N-1)P/B。头的overhead可以忽略。

3和4的最后一项都可看作单个包在传输了一级之后的后续耗时，因为第一级传输已经包含在第一项中了。

交换结构（switch fabric）：crossbar 交换，空分交换，时分交换

# chap4

物理层提供机械、电气、功能、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。

1. 机械特性：主要定义物理连接的边界点
2. 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
3. 功能特性：定义各条物理线路的功能。包括数据，控制，定时，地。
4. 规程特性：定义各条物理线路的工作规程和时序关系

双绞线：既可用于模拟传输，也可用于数据传输。带宽依赖于线的类型和传输距离。

同轴电缆：

1. 基带同轴电缆：50Ω，数据传输
2. 宽带同轴电缆：75Ω，模拟传输

光纤：

1. 单模光纤：半径小，只有一个角度的光可以进入，适于长距离
2. 多模光纤：半径大，多个角度的光可以进入，适于短距离
3. 都支持同时传输几个波长的光，支持波分复用
4. 850nm -> 1310nm -> 1550nm：衰减逐渐减小，价格逐渐提高
5. 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）

SONET与SDH只有微小区别。SONET主要用于北美和日本，SDH主要用于欧洲和中国。

SONET/SDH采用TDM技术，同步传输，由主时钟控制。

SONET路径：路径（path），线路（line），段（section）

复用是基于字节的

移动电话网络：

单方向的寻呼系统：需要很小的带宽

蜂窝电话

第一代：模拟蜂窝电话，只能传送话音

第二代：数字蜂窝电话，主要传送话音，GSM，CDMA

3G / 4G：可以传送话音和数据

模拟蜂窝电话：

早期用于军事通信，push-to-talk system，一个信道，半双工

60年代，IMTS，双频，全双工

# chap5

数据链路层的目的是为了提供功能和规程上的方法，以便建立、维护和释放网络实体间的数据链路

物理层中对应的描述是“数据链路实体”间的物理连接，连接都是高一层的实体之间的

数据链路层协议定义了一条链路的两个相邻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

结点（node）：网络中的主机和网络设备（路由器、交换机等）

链路（link）：通信路径上连接相邻结点的通信信道

点到点（point to point）通信：一条链路上两个相邻结点间的通信

端到端（end to end）通信：从源结点到目的结点的通信，通信路径可能由多个链路组成

实际数据通路：网络层->数据链路层->物理层（图上是物理层的下方？）->另一方的物理层->…

虚拟数据通路：网络层->数据链路层->另一方的数据链路层->…

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

数据链路层为网络层提供三种服务：无确认无连接（用于误码率低/实时业务），有确认无连接（误码率高，如无线网），有确认有链接

组帧方法：往往使用字符计数法和一种其它方法的组合

1. 字符计数法
2. 带字符填充的字符定界法。缺点：局限于8位字符和ASCII字符传送
   1. DLE:Data Link Escape；STX:Start of Text；ETX:End of Text
3. 带位填充的标记定界法
4. 物理层编码违例法：适用于物理层编码有冗余的网络

错误检测和纠正

错误特点：随机，连续突发

海明码：m + r + 1<=2^r

为了检查出d个错，可以使用海明距离为 d + 1 的编码。为了纠正d个错，可以使用海明距离为 2d + 1 的编码。海明码码距为3。

CRC：需要注意的点有：

1. deg(G) = G的位数减1，所以初始时只能在原始数据后补G的位数减1个0，最终传输的数据也只比原始数据多G的位数减1位
2. 加减法都等同于异或，没有进位退位，乘除法是基于这样的加减法进行的
3. 检错能力：
   1. 可以检查单bit
   2. 可以检查两bit错(x^i+x^j)的条件是x^k+i不能被G整除，k<=i-j
   3. G(x)选(x + 1)的倍数就能查出奇数bit错
   4. 能查出所有<=r的突发错。查不出=r+1的突发错的概率为1/2^(r-1)。查不出>r+1的突发错的概率为1/2^r。

对于误码率低的链路，链路层协议可以不实现可靠传输功能

六个传输协议：

协议2是纯停等的。

协议3开始有序号。

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）

协议4开始有捎带回复。单工->全双工。

协议4若双方同时开始发送，则会有一半重复帧。

滑动窗口的意义：

1. 发送方：大小不固定，表示已发送但尚未确认的帧的序号表。下界最小的…的帧，上界是下一个要发送的序号。发送帧时，序号取上界，上界+1；收到ack帧时，下界连续加1（累计确认）直到下界=ack+1。
2. 接收方：大小固定，表示可以接受的序号的范围（上界不算，所以它和发送方窗口一样都是左闭右开）。收到帧时，丢弃窗口外的，如果序号等于下界，则上下界均+1，并产生一个响应帧；其他窗口内的帧虽然接受，但是不改变窗口范围。

协议5：退后n帧重传：发送窗口>1，接受窗口=1。

1. 发送窗口<序号个数，即发送窗口<=MaxSeq。
2. 由于有多个未确认帧，需要设多个计时器
3. 做题时要记住：超时时，将窗口上界置成窗口下界，重传原窗口内的所有帧
   1. 必须将窗口上界置成窗口下界：按照定义窗口上界是下一个帧的编号，但是现在要重传之前的帧了
   2. 尽管重传完之后，窗口上界还是会回到原来的值，但是窗口上界的变化仍然是可观测的，因为一般题目的观测点在发出/收到帧的时候

协议6：选择重传：发送窗口>1，接受窗口>1。

1. 保证接收窗口移动前后，接收窗口内的帧没有重叠
   1. 若发送窗口尺寸>=接收窗口尺寸，则发送窗口<=MaxSeq，接收窗口<=(MaxSeq + 1) / 2
   2. 若发送窗口尺寸<接收窗口尺寸，则发送窗口尺寸+接收窗口尺寸<=序号个数=MaxSeq+1
2. 发送方和接收方的缓冲区大小应等于各自窗口大小，与序号个数无关
3. 增加确认计时器，解决两个方向负载不平衡带来的阻塞问题
4. 可随时发送否定性确认帧NAK
   1. 收到不等于接收窗口下界的包时发NAK（如果这个包在接受窗口内，它会被存下来，但还是要提醒发送方之前的包丢了）
   2. Checksum错误时发NAK
5. 做题时要记住：
   1. 接受窗口的大小一定是不变的。下界变化的时候，上界要跟着变。回退N协议中没有这个问题，因为题目不会考它的发送窗口上界（永远等于下界+1）
   2. 收到不等于接收窗口下界的包时发NAK，序号是窗口下界
   3. ACK timeout的时候发ACK，序号是窗口下界
   4. Timeout的时候不修改发送窗口（而回退N需要将窗口上界置成窗口下界），只重传，重传的帧是发送编号为发送窗口下界的那一帧
      1. 那一帧还包含了一个ack的序号，但是没有关系，就用那一帧
   5. 收到NAK的时候重传的也是发送编号为NAK编号的那一帧

协议工程：

1. 协议说明：既定义一个协议实体提供给它的用户的服务，又定义该协议实体的内部操作
2. 协议验证：验证协议说明是否完整正确
   1. 用于系统实现前的设计阶段，避免可能出现的设计错误
   2. 以协议说明为基础，涉及逻辑证明，原则上验证协议所有可能的状态
   3. 可达性分析是一种常用的验证方法：利用图论知识可以解决状态的可达性问题。可达性分析能够用来解决协议的不完整性、死锁和无关变迁等问题
3. 协议实现：用硬件和软件实现协议说明中规定的功能
4. 协议测试：用测试的方法来检查协议实现是否满足要求
   1. 协议实现是否与协议说明一致（一致性测试）
   2. 协议实现之间的互操作能力（互操作性测试）
   3. 协议实现的性能（性能测试）

常用的形式化方法：

1. 有限状态机FSM。扩展：EFSM
2. 形式化语言模型：LOTOS，Estelle，SDL。都有相应扩展
3. Petri网。扩展：时间Petri网，随机Petri网，高级Petri网
4. 进程代数（Process Algebra）。扩展：随机进程代数

HDLC、X.25 LAPB：LAPB是HDLC的子集

面向比特

地址域：用来区分终端，可用来区分命令和响应。地址=接受方：命令；地址=发送方：响应

控制域：序号：滑动窗口

帧分为：信息帧，监控帧，无序号帧

校验和：CRC校验

站由计算机（路由器）和终端组成，分为主站（发送命令），次站（接受命令），组合站（同时具有主、次站功能）

链路构型分为平衡型（主次-主次，组合-组合，适用于计算机-计算机）和非平衡型（主-次，主-多个次，适用于终端-计算机·）

基本操作模式分为正规响应模式 NRM（适用于非平衡型），异步响应模式 ARM（适用于主-次和主次-主次），异步平衡模式 ABM（适用于组合-组合）

PPP协议：由SLIP协议改进而来

面向字符

通常不使用滑动窗口，但是也具有利用HDLC帧进行可靠传输的可选功能。

以帧为单位发送，而不是原始IP包

协议包括两部分：

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

# chap6

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。

特点：高传输率，短距离，低出错率。使用广播信道。

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。

MAC子层的功能：数据帧封装、发送和接收；介质访问管理。

静态分配：FDM/TDM，都只适用于用户少且固定，各用户通信量大，无法灵活地适应站点数及其通信量的变化。

动态分配：

1. ALOHA协议
   1. 纯：效率Ge^(-2G)，最大0.184
   2. 分槽：效率Ge^(-G)，最大0.368
2. 载波监听多路访问协议CSMA（Carrier Sense Multiple Access）
   1. 发送帧之前监听信道有无载波
   2. 1-坚持：监听到信道空闲时立即发送。冲突时随机等待。
   3. 非坚持：信道忙时随机等待，只有一开始信道就空闲才发送。冲突时随机等待。
   4. p-坚持（适用于分槽信道）：若信道忙，则等待下一个时槽。一开始信道就空闲时，以概率p发送数据。冲突时随机等待。
      1. 1-坚持可以看作p-坚持p=1的特例，非坚持不能看做p-坚持p=0的特例
3. 带冲突检测的载波监听多路访问协议CSMA/CD
   1. 与2的区别在于冲突在发送过程中检测。
   2. 最坏情况下，2倍电缆传输时间后检测到冲突。
4. 无冲突协议
   1. 分为传输周期，竞争周期，空闲周期
   2. 基本位图协议：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。竞争结束后，按照站序号发送。
      1. 轻负载下效率d/(d+N)，重负载下效率d/(d+1)。其中数据帧由d个时间单位组成
      2. 序号大的站得到的服务好。
   3. 二进制下数法：所有站都有一个二进制长度相等的序号。竞争log2N轮，从高位到低位，每轮所有站发送自己序号对应位的值，如果一个站自身这一位是0而所有位或起来的结果是1，则它不再参与竞争，剩余站继续。最终胜利者是序号最大的站。
      1. 效率d / (d + log2N)
      2. 序号大的站得到的服务好
5. 有限竞争协议
   1. 适应树搜索协议
      1. 站点组织成二叉树
      2. 第0轮全部站可竞争信道，只有一个站要使用信道则发送；有冲突则在第1轮内半数站L参与竞争。如其中之一获得信道，本帧后的下一轮竞争只由另一半R参与；如发生冲突，继续在L中折半搜索。
      3. 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
6. 无线局域网协议MACA，MACAW
   1. 特点：
      1. 冲突只能被接收站点发现。
      2. 某一时刻，信道上可以有多个有效数据帧
   2. 问题：隐藏站点：发送者距离接收者附近的竞争者太远，导致发送者不能发现竞争者；暴露站点：非竞争者距离发送者太近，导致发送者不能发送。
   3. MACA（Multiple Access with Collision Avoidance）：发送者发送RTS（Request To Send），发送者周围的站点一定时间内不能发送。接受者收到RTS后发送CTS（Clear To Send），接收者周围的站点一定时间内不能发送。之后发送者开始发送。
      1. 若发生冲突，采用二进制指数后退算法等待随机时间
   4. MACAW：改进的MACA
      1. 对每个成功传输的数据帧，都要产生确认帧。（有确认的无连接服务？）
      2. 增加了发送站点的载波监听
      3. 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
      4. 发生拥塞时，站点间交互信息

LLC（Logical Link Control）子层：位于MAC和网络层之间。由IEEE 802.2定义。

MAC子层不提供确认机制和流量控制，有些情况下网络层需要确认和流控（不包括IP协议），所以需要LLC提供确认机制和流量控制

LLC也提供三种服务：不可靠数据报服务，有确认数据报服务，可靠的面向连接的服务。（因为LLC属于数据链路层，而数据链路层提供这三种服务，所以LLC也提供这三种服务）

LLC帧头基于HDLC协议

IEEE 802.3和Ethernet

曼彻斯特编码

电缆标准：<速率/Mbps><介质种类><最大长度/100m>。例如10Base5表示10Mbps，基带传输（base），500m。

收发器（transceiver）：处理载波监听和冲突检测。

中继器：物理层设备，放大信号。可以拓展网段长度。两个收发器之间最多使用4个中继器，最长2500米。

帧中的地址：IEEE 802.3是2或6个字节，以太网是6个字节。

1. 目的地址第1位（最低位）为0表示单播，为1表示组播。目的地址全为1表示广播。源地址第一位总为0。
2. 第二位为0表示全球地址，为1表示本地地址

帧长度域：2字节，取值在0-1500之间

校验和：CRC，四个字节

最短帧长：10Mbps LAN => 64字节。两倍的信号往返最长时间（考虑中继器延时），为51.2ms，51.2ms \* 10Mbps / 8 = 64B。

发生冲突时采用二进制指数后退算法，将冲突发生后的时间划分为长度为51.2微秒的时槽，第n次冲突后在0..2^n-1间随机选择。16次冲突后放弃。

交换式802.3 LAN：目的：减少冲突。

快速以太网：比特时间100ns -> 10ns。

xByT / xByB：y个传输周期表示x个二进制数。T是线上传输三进制数·。用频率\*x/y得到传输速率。

千兆以太网：在一个冲突域内，只允许一个repeater。

万兆以太网：10GE只工作在全双工方式，不使用CSMA/CD协议，传输距离大大提高。使用单模或多模光纤。

网桥：工作在数据链路层的一种网络互连设备，它在互连的LAN之间实现帧的存储转发

中继器不能隔离冲突域，网桥/交换机可以隔离冲突域。网桥有助于安全保密。

连接k个不同LAN的网桥具有k个MAC子层和k个物理层。

三种不同的LAN互连共有九种组合。（3^2=9）

互连时需要解决的问题：

1. 帧格式的转换
2. 不同的LAN速率不同，网桥要有缓存能力
3. 高层协议的计时器设置
4. 最大帧长度不同。解决办法：丢弃无法转发的帧

分类：

1. 透明网桥
   1. 工作在混杂（promiscuous）方式，接收所有的帧
   2. 采用逆向学习(backward learning)算法收集MAC地址。分析帧的源MAC地址和端口的对应关系。注意每收到一个包就有一次学习的机会。
   3. 转发策略：
      1. 目的LAN与源LAN相同，则丢弃帧
      2. 目的LAN与源LAN不同，则转发帧
      3. 目的LAN未知，则洪泛帧（向除了源端口外的所有端口发送）
2. 生成树网桥：多个网桥并行可能产生回路
   1. 生成树的构造：
      1. 每个桥广播自己的桥编号，号最小的桥称为生成树的根
      2. 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
      3. 生成树需要动态变化
3. 源路由网桥
   1. 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（不包含主机编号，因为信息传递并不经过主机）
   2. 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
   3. 对带宽进行最优的使用。但网桥的插入对于网络是不透明的，需要人工干预
   4. 它可以看作是面向连接的

IEEE 802.5 令牌环

差分曼彻斯特编码

环不是广播介质，而是不同的点到点链路组成的环。

各个站点是公平的，获得信道的时间有上限，避免冲突发生。

1比特物理长度：1比特的传输耗时\*信号传播速率

环接口引入了1比特的传输延迟。

令牌（Token）是一种特殊的比特组合模式。环接口有两种操作模式：监听模式和传输模式。监听模式下经过1比特延迟后令牌继续传递。传输模式即站有数据要发送，它将令牌移出环，然后重新生成令牌，并转入监听模式。

重负载下，效率接近100%。

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。

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

环上存在一个监控站，负责环的维护，通过站的竞争产生。职责：

1. 保证令牌不丢失
2. 处理环断开情况
3. 清除坏帧，检查无主帧

FDDI

使用多模光纤作为传输介质。MAC协议与 Token Ring 类似。4B5B编码。

FDDI有两个环，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。

DPT/RPR

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。

SRP（Spatial Reuse Protocol）是一个媒介无关的MAC层协议，用来实现DPT在光纤环上的功能。

# chap7

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。

通信子网的最高层，端到端传输的最底层。

关键问题：了解通信子网的拓扑结构。

网络层提供哪种服务？

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，将复杂的功能放在网络层(通信子网)
2. 无连接服务：互联网的观点：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，将复杂的功能放在传输层

路由算法：目的是找出并使用汇集树（从所有源结点到一个目的结点的最优路由的集合）

最优化原则：如果路由器 J 在路由器 I 到 K 的最优路由上，那么从J 到 K 的最优路由会落在同一路由上。

1. 洪泛算法：属于静态路由算法。
2. 选择性洪泛算法：洪泛法的一种改进，将接收的每个分组仅发送到与正确方向接近的线路上。健壮性好。
3. 基于流量的路由算法：属于静态路由算法。
   1. 既考虑拓扑结构，又兼顾网络负载
   2. 前提：每对结点间平均数据流相对稳定可预测。需要的信息：网络拓扑，通信量矩阵，带宽矩阵
   3. 提前离线计算路由。可用于流量工程
4. 距离向量算法（DV）：用于RIP
   1. 水平分裂：从邻居结点学到的到X的距离不向邻居结点报告，使坏消息传播的也快
   2. 缺点：
      1. 没有考虑链路带宽
      2. 收敛慢
      3. 存在无穷计算问题（和收敛慢可以作为两个理由）
      4. 路由报文开销大（不是增量更新，每次发布全部信息）
      5. 不适合用于大规模网络（RIP协议最大支持15跳）
5. 链路状态路由算法（LS）：用于OSPF，IS-IS
   1. 发现邻居结点，并学习它们的网络地址：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同一个LAN时，引入人工结点来代表他们（称为代表路由器DR, Designated Router）
   2. 测量到每个邻居结点的延迟或开销，可以用回复延迟或带宽
   3. 将所有学习到的内容封装成一个分组
      1. 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
         1. 32位序号使得很少回绕
         2. 年龄每秒减1，到0后丢弃，从而解决序号出错和路由器崩溃后序号重置的问题
      2. 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
      3. 链路状态分组定期创建或发生重大事件时创建
   4. 将这个分组发送给所有其它路由器（不只是邻居）（发布增量信息）
      1. 链路状态分组需要应答。分组中包含ack信息和send信息，指示这个路由器学会后应该向谁回复以及再教给谁
   5. 每个路由器中根据完整的拓扑，使用dijkstra算法计算最短路径
   6. 缺点：可能存在路由振荡。如link cost = amount of carried traffic
6. 路径向量算法（PV）：用于BGP

LS和DV的健壮性: 如果路由器不能正常工作会发生什么?

LS：结点会广播错误的链路开销，每个结点只计算自己的路由表（所以呢？？？）

DV：结点会广播错误的路径开销，每个结点的路由表被别的结点使用，错误会传播到全网

分层路由：书上分层后路由表项的计算方法是sigma(每层的组数-1)+自身所在组的路由器总数。即自身在路由表中，但自身所在组不在。

移动主机的路由：

移动用户，家乡位置（每个移动用户都有一个永久的家乡位置，用地址来标识），家乡代理（记录不在家的移动用户），外部代理（记录正在访问该区域的移动用户）

过程：三角路由

1. 移动用户向外部代理注册
2. 外部代理与移动主机的家乡代理联系
3. 当一个分组发给移动用户时，首先被转发到用户的家乡局域网
4. 被家乡代理接收，家乡代理查询移动用户的新位置和与其对应的外部代理的地址
5. 家乡代理采用隧道技术，将收到的分组作为净负荷封装到一个新分组中，发给外部代理
   1. 可选：家乡代理告诉发送方，后续分组直接发给外部代理
6. 外部代理收到分组后，将净负荷封装成数据链路帧发给移动用户

拥塞控制算法

拥塞：网络上有太多的分组时，性能会下降。

拥塞控制是全局性的，流控制是局部问题。

闭环控制：基于反馈

开环控制：通过设计避免控制，不考虑网络当前状态

1. 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
2. 漏桶算法，令牌桶算法：都可以用于定长分组（如ATM）和变长分组
   1. 令牌桶高速维持时间=桶大小/(流量速率-积累速率)

虚电路子网中的拥塞控制：

1. 流说明：描述发送数据流的模式和希望得到的服务质量的数据结构
   1. 子网和接收方可以做出三种回复：同意、拒绝、其它建议
2. 准入控制：根据流说明和网络资源分配情况，进行准入控制
   1. 可以在解决拥塞前不允许建立虚电路，也可以绕开拥塞地区
3. 子网根据协议在虚电路上为连接预留资源

抑制分组/逐跳抑制分组：前者只对发送方起作用。后者对它经过的每个路由器都起作用，要求上游路由器要求有更多的缓冲区。

公平队列算法：路由器的每个输出线路有多个队列，路由器循环扫描各个队列，发送队头的分组。可以改成按字节轮询。加权公平队列：优先级高的队列在一个轮询周期内获得更多的时间片。

负载丢弃：文件传输：丢弃新的，wine；多媒体服务：丢弃老的，milk。

网络互联

互连网络（internet）：两个或多个网络构成互连网络

网络互连设备：

1. 中继器：物理层
2. 网桥：数据链路层：在局域网之间存储转发帧，网桥可以改变帧格式
3. 多协议路由器：网络层：在网络之间存储转发分组。必要时，做网络层协议转换
4. 传输网关
5. 应用网关

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：每个分组单独路由，提高网络利用率。

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。（无线主机的路由中用到）

分片：透明（目的主机负责重组）/不透明（出口网关负责重组）。

标记片段的方法：

1. 树形标记。例子：分组0分成三段，分别标记为0.0, 0.1, 0.2，片段0.0构成的分组被分成三片，分别标记为0.0.0, 0.0.1, 0.0.2。
   1. 段标记域要足够长，分片长度前后要一致
2. 偏移量+最后片段指示位

防火墙：防止信息泄露或不好的信息渗透。组成：两个路由器中间夹着一个应用网关。

网络层协议

网络之间互连的纽带是IP（Internet Protocol）协议

1. IP
   1. 20个字节的固定部分和0-40个字节的变长部分。IHL，Version，Type of Service，Total length，Identification，MF，DF，Fragment offset，TTL，Protocol，Checksum，Source address，Destination address
   2. 服务类型（Type of Service）：3个优先级位，3个标志位：D（Delay）、T（Throughput）、R（Reliability）。目前，很多路由器都忽略服务类型域
   3. 分段的时候需要且只需要考虑IP头对偏移量的影响。IP规定offset必须是8的倍数，并且字段里填的是除8后的结果。所以，除了最后一个分段之外，其它分段的载荷大小（不包括IP头）都必须是8的倍数。
   4. 变长部分举例：安全性，严格/松散源路由，路由记录，时间戳
   5. IP地址：网络号+主机号。一个IP 地址并不真正指向一台主机，而是指向一个网络接口
      1. ABCDE的标志是0,10,110,1110,11110，ABC的主机号都按字节对齐，D没有网络号 + 主机号，表示多播，E是保留
      2. 主机号全0表示自身，全1表示广播。
      3. 127.x.x.x用于loopback
   6. 子网：为了便于管理和使用，可以将网络分成若干供内部使用的部分，对外仍表现为一个网络
2. ICMP：主要用来报告错误和测试，报文封装在IP分组中（不经过TCP和UDP）
3. ARP：解决网络层地址（IP地址）与数据链路层地址（MAC地址）的映射问题
   1. 若目的主机在同一子网内，用目的IP地址在ARP表中查找，否则用缺省网关的IP地址在ARP表中查找
   2. 若未找到，则发送广播分组，目的主机收到后给出应答，ARP表增加一项
   3. ARP攻击：攻击者发出伪造的ARP响应，存在于局域网
4. RARP：解决数据链路层地址（MAC地址）与网络层地址（IP地址）的映射问题
   1. 用于无盘工作站的启动
   2. 缺点：由于路由器不转发广播帧 ，RARP服务器必须与无盘工作站在同一子网内
   3. 一种替代协议BOOTP，它使用UDP（上面提到的这几个，ICMP，ARP，RARP，都不经过TCP和UDP）
5. RIP：属于内部网关协议IGP
   1. 基于UDP
   2. 如果180秒（6个路由声明周期）内没有收到来自邻居的路由声明，则认为邻居/链路失效。新的路由声明发往其他邻居，链路失效的信息迅速传播到全网
   3. 使用poison reverse避免ping-pong loops
6. OSPF：属于内部网关协议IGP
   1. 不经过TCP和UDP
   2. 支持多种距离衡量尺度，如距离，带宽，延迟
   3. 支持基于服务类型的路由，支持负载平衡，支持分层路由，适量的安全措施，支持隧道技术
   4. 自治系统可以划分成区域，每个AS有一个主干（backbone）区域，称为区域0。所有其它区域都与主干区域相连，其它区域之间不能相连。
   5. 四类路由器，允许重叠：
      1. 完全在一个区域内的内部路由器
      2. 连接多个区域的区域边界路由器ABR
      3. 主干路由器（就是主干区域的内部路由器）
      4. 自治系统边界路由器ASBR
7. BGP：属于外部网关协议EGP
   1. 基于TCP
   2. 采用路径向量（path vector）算法，路由信息中记录路径的轨迹
      1. 每个BGP网关向邻居广播所有通往目的地的路径
   3. 域间路由跨越不同管理域，要控制流量如何路由，策略比性能更重要
8. IPv6
   1. 减少路由表的大小，简化协议
   2. 更多的关注服务类型，特别是实时数据
   3. IP头简化，由13个域减少为7（好像是8个？）个域：定长40字节。取消IHL ，Protocol替换成next header，取消分片相关的域（Identification，DF，MF，Fragment offset），取消checksum。Type of service替换为Priority和Flow label。
      1. Version，值为6
      2. Priority，区分源端可以/不能流控的分组，值越小优先级越低
      3. Flow label，用来允许源和目的建立一条具有特殊属性和需求的伪连接
      4. Payload length，用来指示IP分组中40字节分组头后面的长度，只统计载荷
      5. Next header，指示扩展分组头，若是最后一个分组头，则指示传输协议类型（TCP/UDP）
         1. hop-by-hop header，用来指示路径上所有路由器必须检查的信息
         2. routing header，列出路径上必须要经过的路由器
         3. fragmentation header，与IPv4相似，扩展头中分组括IP分组标识号、片段号和是否还有片段的位，但只有源主机可以分片
      6. Hop limit，IP分组的最大跳数
      7. Source address，destination address，16字节定长地址
   4. IPv6地址：每个数字开头的0可省略，连续的0可以被一对冒号替代，但是一对冒号只能出现一次

从IPv4往IPv6的过渡

1. 双栈：实现IPv4/v6两套协议栈，主机根据DNS返回的结果或对方发来报文的版本号决定采用哪个协议，路由器根据收到IP分组的版本号决定采用哪个协议

2. 翻译：实现IPv4/v6两套协议栈，在两套之间进行协议翻译和地址翻译

3. 隧道：IPv6的报文作为IPv4报文的净负荷在IPv4网络中传输

CIDR

将剩余的C类地址分成大小可变的地址空间。将世界分成4个区，每个区分配一块连续的C类地址空间。虽然现在用于C类地址，但CIDR思想可用于所有IP地址，没有A、B、C类之分。

路由表中增加一个32位的掩码域。最长前缀匹配原则：路由查找时，若多个路由表项匹配成功，选择掩码最长（1比特数多）的路由表项。

# chap8

引入传输层的原因

1. 消除网络层的不可靠性（传输层不存在有无确认的服务之分，只有有无连接的服务之分）
2. 提供从源端主机的进程到目的端主机的进程的可靠的、与实际使用的网络无关的数据传输

传输实体（transport entity）：完成传输层功能的硬软件。传输层实体利用网络层提供的服务向上层提供有效、可靠或尽力而为的服务。

缓存：发送方为了重传必须有缓存，接受方可以有也可以没有。

缓存分类：固定大小缓存，可变大小缓存，缓存池。

拆除连接方式：

1. 不对称方式：任何一方都可以关闭双向连接。可能丢失数据。
2. 对称方式：每个方向的连接单独关闭（单独一方的关闭也是有意义的)，双方都执行DISCONNECT才能关闭整条连接。不能通过确定的N次握手保证不会丢失数据。

Berkeley Sockets

连接释放是对称的。

服务端：socket->bind->listen（非阻塞）->accept（阻塞）->派生进程去处理->close

客户端：socket->connect->close

双方使用send和receive完成数据的全双工发送

传输服务访问点TSAP（Transport Service Access Point）：IP+端口

端口号的获取：预先约定，或从名称服务器/目录服务器获取。ppt的例子上获取过程是面向连接的。

进程服务器：监听多个端口，需要时唤起对应的服务进程。

TCP

面向连接的、可靠的、端到端的、基于字节流（消息边界传输中不能得到保留）的传输协议

不支持组播和广播

256以下的端口号被标准服务保留。如FTP（21），TELNET（23），HTTP（80）

对于应用程序发来的数据，TCP可以立即发送，也可以缓存一段时间以便一次发送更多的数据。为了强迫数据发送，可以使用PUSH标记

对于紧急数据，可以使用URGENT标记

按字节分配序号，每个字节有一个32位的序号

传输实体之间使用的TPDU称为段（segment）

段的大小必须满足65535字节的IP包数据净荷长度限制，还要满足数据链路层最大传输单元（MTU）的限制，例如以太网的MTU为1500字节

TCP实体使用滑动窗口协议，确认序号等于接收方希望接收的下一个字节序号

TCP头：定长20字节+变长

1. 源和目的端口：各16位
2. 序号和确认号：以字节为单位编号，各32位
3. TCP头的长度：4位，长度单位为32位字，包含选项域
4. 6位的标识位
   1. URG：紧急指针是否有效
   2. ACK：确认号是否有效
   3. PSH：指示发送方和接收方将数据不做缓存，立刻发送或接收
   4. RST：不可恢复的错误重置连接
   5. SYN：连接建立指示
   6. FIN：连接释放指示
5. 窗口大小：用于基于可变滑动窗口的流控，指示发送方从确认号开始可以再发送窗口大小的字节流
6. 校验和：TCP头，数据和IP伪头的校验和
7. 紧急指针：指向从当前序号开始找到紧急数据的字节偏移量，URG=1时有效

345占4byte，67占4byte

TCP连接：三次握手

A –SYN，SEQ=x-> B

B –SYN+ACK，SEQ=y，ACK=x+1->A

A- ACK，SEQ=x+1，ACK=y+1->B

此后可以开始正常通信，一开始使用的SEQ分别是x+1和y+1。

注：TCP中，ACK表示下一个希望接受的字节的序号，与滑动窗口定义差1。所以，第2/3次握手中，ACK分别位x+1和y+1，表示的分别是对序号位x和y的字节的ACK。这字节就是SYN包中的一个字节，而最后一个ACK包的字节数为0，所以接下来A还是使用序号x+1。

若B机器上并没有进程监听A请求的端口，则收到A的请求后会回答RST置位的TCP段。

若两个主机同时试图建立彼此间的连接，则只能建立一条连接。

TCP释放：三次握手+计时器

A-DR->B

B-DR->A

A-ACK->B

释放连接时，发出FIN位置1的TCP段并启动定时器，在收到确认后关闭连接。若无确认并且超时，也关闭连接。

一方关闭连接的时刻一定是它的最后一个动作的时刻。对于DR的发起方，是它在发出ACK后；对于另一方，是它在接收到ACK后。

如果A的DR到达了B，但A迟迟没有收到B发来的DR，一段时间后AB都会因超时而关闭连接。

如果A的DR没有到达B，一段时间后A会因超时而关闭连接，但B不会。这可能导致数据丢失。

TCP的窗口管理机制

基于确认和可变窗口大小

正常情况下，窗口大小为0时发送方不能再发TCP段，但有两个例外：紧急数据可以发送；为防止死锁，发送方可以发送1字节的TCP段，以便让接收方重新声明确认号和窗口大小

优化策略：

1. 发送方缓存应用程序的数据，等到形成一个比较大的段再发出
2. 在没有可能进行“捎带”的情况下，接收方延迟发送确认段
3. Nagle算法：当应用程序向传输实体发出一个字节时，传输实体只发出第一个字节并缓存所有其后的字节直至收到对第一个字节的确认，然后将已缓存的所有字节组段发出并对再收到的字节缓存，直至收到下一个确认
   1. 可以避免多次发送小包，但是不利于实时性强的应用。可能需要禁用Nagle算法
4. 使用Clark算法解决傻窗口症状
   1. 傻窗口症状：当接收方一次从传输层实体读出一个字节时，传输层实体会产生一个一字节的窗口更新段，使得发送方只能发送一个字节
   2. 解决办法：限制收方只有在具备一半的空缓存或最大段长的空缓存时，才产生一个窗口更新段

TCP拥塞控制

TCP按流量控制窗口（由滑动窗口协议维护）和拥塞窗口的最小值发送数据

连接建立时拥塞窗口初始值为最大段长MSS（别忘了初值不是0），阈值为64K

发出一个最大段长的TCP段，若收到正确确认，则拥塞窗口变为两个最大段长。这里需要考虑累计确认，即收到一个ack时可能增加多次congwin，而不是每个ack增加一次。

发出（ 拥塞窗口/最大段长）个最大长度的TCP段，若都得到确认，则拥塞窗口加倍

重复上一步，直至发生丢包产生的超时事件，或拥塞窗口大于阈值

当拥塞窗口大于阈值时，拥塞窗口开始线性增长，每收到拥塞窗口个包，拥塞窗口才增长1。直至发生丢包产生的超时事件

超时事件发生后，阈值设置为当前拥塞窗口大小的一半，拥塞窗口重新设置为一个最大段长

执行慢启动算法

快速重传算法：不必等到计时器超时才判定丢包。如果某个包丢失，则接收方发回的ACK会卡在同一个值。不过需要考虑偶然重复，所以连续收到3个重复确认后，不再等待计时器超时。

指数上升阶段，理论上应该是t=n时有一个陡峭的上升，上升完后congwin=2^n。虽然陡峭，也还是要一定时间，所以是在t=n后一点才变成2^n。但是绘制congwin-t图的时候，可以认为是t=n-1~n间连续地变成了2^n，t=n时就是2^n。

各种拥塞控制的方法：

系数a为加系数，b为乘系数

MIAD：公平性不收敛也不稳定。有效性不收敛，在x1=x2=bIaD/(1-bI)时稳定。

AIAD：公平性不收敛但稳定。有效性不收敛，在aI+aD=0时稳定。

MIMD：公平性不收敛但稳定。有效性收敛。条件是平凡的。

AIMD：公平性收敛。有效性收敛。

UDP

优点：不需要建立连接，延迟小；简单，没有连接状态；报文头小；没拥塞控制，可以尽快的发送

“best effort”服务，UDP 报文可能会丢失、乱序。每个UDP报文的处理都独立于其他报文。

经常用于流媒体应用 ：可以容忍丢包，速率敏感

其他应用范围：RIP，DNS，SNMP，BOOTP。

基于UDP的可靠传输：应用程序自己定义错误恢复。

UDP头包含src/dst端口，udp length，checksum。

# chap9

用户代理（user agent）是指用户和网络应用程序间的接口。比如web浏览器，流媒体播放器等。

API ：定义应用程序和传输层之间的接口

客户/服务器模型是网络应用的基础。客户/服务器分别指参与一次通信的两个应用实体，客户方主动地发起通信请求，服务器方被动地等待通信的建立，并提供服务。服务器软件一般分为两部分：一部分用于接受请求并创建新的进程或线程，另一部分用于处理实际的通信过程。

DNS

基于UDP

典型的客户/服务器交互系统

互联网的顶级域名分为组织结构和地理结构两种。每个域对它下面的子域和机器进行管理。

域名是“.”分的字符、数字串组成的，大小写无关，最长255，每个部分最长63。

在DNS的数据库中用资源记录来表示主机和子域的信息。资源记录是一个五元式：

Domain\_name Time\_to\_live Type Class Value

1. Type=A，Name：hostname，Value：IP地址
2. Type=MX，Value：与name对应的邮件服务器的主机名（hostname）
3. Type=NS，Name：域名（例如，edu.cn），Value：该域权威域名服务器的IP地址
4. Type=CNAME，Name：规范名称（ canonical name）的别名，Value：规范名称

DNS将域名空间划分为许多无重叠区域，每个区域覆盖了域名空间的一部分，区域的边界划分是人工设置的。每个区域有一个主域名服务器和若干个备份域名服务器。

域名解析过程：应用->区域域名服务器-（若找不到）->根域名服务器-（若找不到）->authoritative域名服务器

根域名服务器可能不知道 authoritative域名服务器，但知道中间的域名服务器，而中间域名服务器知道如何与 authoritative域名服务器联系

Recursive query：问一个人就是把任务完全交给它

Iterated query：可以从别人那里得到额外信息。别人：“我不知道，但可以问它”

SNMP

基于UDP

性能管理、故障管理、配置管理、记账管理和安全管理

SNMP模型：

1. 被管理节点：运行SNMP代理程序（SNMP agent），维护一个本地数据库，描述节点的状态和历史，并影响节点的运行
2. 管理工作站：运行专门的网络管理软件（manager），使用管理协议与被管理节点上的SNMP代理通信，维护管理信息库
3. 管理信息：每个站点使用一个或多个变量描述自己的状态，这些变量称为“对象”，所有的对象组成管理信息库MIB（Management Information Base）。
4. 管理协议（SNMP）：管理协议用于管理工作站查询和修改被管理节点的状态，被管理节点可以使用管理协议向管理站点产生“陷阱（trap）”报告

SNMP使用SMI首先将变量定义为“对象”（object），相关的对象被集合成“组”（group），组最后被汇集成“模块”（module）。SNMP的MIB包含10个组。

抽象语法表示法1（ASN.1）：标准的对象定义语言。分为数据描述定义和传输语法定义两部分。可以作为异种计算机设备之间“对象”描述和传输的表示方法。

对象命名树使用编码，唯一地确定每个标准中的对象。

ASN.1的传输语法：需要传输标识符（type/tag），长度域，数据域。

标识符：高两位：00~11分别为Universal，Application，Context specific，Private。第三位：0为primitive，1为constructed。当tag值在0 ~ 30之间时，用低5位表示；当tag值大于30时，低5位为“11111”，用后面字节表示。每个标识字节包括7个数据位，最后一个字节高位为“1”，其它字节高位为“0”。

长度域：当长度 < 128字节时，用一个字节表示长度，高位为“0”。当长度>=128字节时，第一个字节高位为“1”，低7位表示后面表示长度的字节个数，后面的若干个（<=127）字节表示长度。

数据域：BIT STRING：编码表示不变，长度域表示字节个数（需考虑表示…不用的位数的额外字节），在传位串前先传一个字节表示位串最后一个字节不用的位数。例，位串“010011111”传输时变为“07 4f 80”（十六进制），长度为3

电子邮件

基于TCP

用户代理：允许用户阅读和发送电子邮件，一般为用户进程

消息传输代理：将消息从源端发送至目的端，一般为系统的后台进程

电子邮件的组成：

1. 信封：接收方的信息，如名字、地址、邮件的优先级和安全级别
2. 信件内容：由信头和信体组成，信头包含了用户代理所需的控制信息，信体是真正的内容

SMTP负责发送，服务器间转发，POP3和IMAP负责从服务器收取。加密电子邮件协议：PGP与PEM协议

消息传输代理和用户代理两个方向的发送完成后，释放TCP连接(QUIT)。属于持久（Persistent）方式。

WWW

WWW(World Wide Web)是用于访问遍布于互联网上的相互链接在一起的超文本的一种结构框架。

Web页面（网页）：由对象（object）组成，用URL表示地址。

HTTP

使用TCP，80端口

无状态协议：Web服务器不保存客户信息

HTTP1.0：非持久连接，每个object取得至少要两个RTT（因为TCP连接至少需要一个RTT来建立），每个object的传输都要经历慢启动

HTTP1.1：持久连接，客户端一旦得到基本的HTML文件就发出请求索取全部object，较少的RTT和慢启动时间

Web缓存：响应时间短，减轻通往远端服务器的连接的流量

FTP

客户/服务器模式：由客户端发起文件传输

客户端连接到ftp服务器TCP的21号端口，然后建立两个并行的TCP连接:

1. 控制: 在客户端和服务器之间交换命令、响应(采用telnet)
2. 数据: 传递文件数据(可双向，不必始终存在)

Ftp服务器维护状态：当前目录，身份认证