

计算机网络复习

Chap1引言

1.1 使用计算机网络

- 网络的商业应用：用户-服务器模型
- 家庭应用：P2P对等通信各种网络设备通过电缆等连接到internet
- 移动用户：通过移动设备进行语音和文本通信
- 社会问题：

1.2 网络硬件

- 分类的两大因素：**技术和规模**
- **按传输技术分类：点-点式网络、广播式网络**
 - 点对点：每条物理线路连接一对计算机，只有一个发送方和接收方
 - 广播：多台计算机共享一条通信信道。两个以上节点同时发送会冲突。
- **规模分类：个域网、局域网、城域网、广域网、互联网**
 - 广域网：远程网，跨越一个很大地理范围，包含大量的机器称为主机。
 - 互联网：一组相互连接起来的网络。

1.3 网络软件

- 协议层次结构
- 协议：同一层对等实体间交换的信息或分组的格式及含义的一组规则
 - 每对相邻层之间的是接口，接口定义了下层向上层提供哪些原语操作和服务。
 - 分层的好处：
 1. 使每层的设计简单化
 2. 各层独立
 3. 灵活性好
 - 两种接口：上下层之间的服务接口、对等实体之间交换的信息的对等实体接口。
- 服务原语：用户进程通过原语操作来访问该服务。

1.4 参考模型

- OSI(Open System Interconnection)模型
 - 物理层、数据链路层、网络层、传输层、会话层、表示层、应用层
 - 数据链路层：在物理线路上提供无差错的数据传输帧。
 - 网络层：控制通信子网提供源点到目的点的数据传输分组。选择路由和控制拥塞。
 - 传输层：提供端到端的数据传送服务。
 - 会话层：允许不同机器上的用户建立会话关系。
- TCP/IP协议
 - 主机至网络层、互联网层、传输层、应用层。
 - 互联网层：控制通信子网提供源点到目的点的IP分组传送。使用IP地址来寻址和传输数据。本质上无连接的传输。
 - 传输层：端到端的数据传输服务：TCP/UDP
 - 应用层：提供各种Internet管理和应用服务功能。

1.6 网络标准化组织

- ITU (国际电信联盟)：主要用于国与国之间的互联，每个国家内部可能有自己的标准
- ISO (国际标准化组织)：
 - IEEE (电气电子工程师协会)
 - ANSI (美国国家标准研究所)
 - EIA (电子工业协会)
- IAB (Internet体系结构委员会)
 - IETF: RFC标准
 - IRTF
 - ICANN: IP地址和自治域号码分配、域名系统和根域名服务器管理
- W3C
- RFC(Request For Comments)任何人都可以提交，使用数字命名：
 - 因特网草案；
 - 提议标准；
 - 草案标准；
 - 因特网标准。

Chap2 物理层

2.1 数据通信的理论基础

- 傅里叶分析：任何信号都可以看成由不同频率成分组成。
 - 方波的离散函数是连续的
- 带宽是传输介质的物理特性。
- 有限的带宽限制了数据的传输速率。
- 带宽表示：
 - **模拟信道：Hz**
 - **数字信道：bps**

2.1.3 信道的最大数据速率

- 无噪声有限带宽信道的最大数据传输率： **$2B\log_2 V$ (bps)**
 - 任意信号通过一个**带宽为B**的低通滤波器，则**每秒采样2B次**就能完整地重现该信号，**信号电平包括V个离散等级**
- 香农公式：
 - **信噪比 = $10\log_{10}(S/N)$ (db)**
 - **最大数据传输率 (信道容量) = $B\log_2(1 + S/N)$ (bps)**

2.2 有导向传输介质

- 物理传递磁介质：数据率高，延迟大

2.3 双绞线

- 模拟/数字
- 不同电线产生的干扰相互抵消，抵抗共模干扰

2.3 同轴电缆

- 传输特性比双绞线好
- 75Ω-Analog-300~400MHz
- 50Ω-Digital

2.4 电力线

2.5 光纤

- 纤芯、包层、护套
- 通过光信号的有无表示0/1
- 不受电磁干扰
- 高带宽低衰减
- 0.85um,1.30um,1.55um三个窗口

2.5 数字调制和多路复用

- 基带传输
 - 数据比特直接转换成信号
 - 基带：0-到最大值之间的全部频率，是原始信号所固有的频带。
 - 基带传输：在传输时直接使用基带信号。不用调制
 - 通带传输：把基带信号搬移到其他频带传输，通常叫调制
- 通带传输
 - 调制：
- 多路复用
 - 频分复用FDM；时分复用TDM；码分复用CDM。

2.6 公共电话交换网络

2.6.5 交换

- PSTN-电路交换
- Internet-分组交换
- 数据报&虚电路

Chap3 数据链路层

3.1 数据链路层中的设计问题

成帧

- 字符计数法
- 带字节填充的分解符法
- 带位填充的分界标志法
- 差错控制（定时器），流量控制（接收方确认）

3.3 停等协议

3.4 滑动窗口

3.4.1 1位滑动窗口协议

3.4.2 退回N协议

3.4.3 选择重传协议

- 窗口大小为 $(MAX_SEQ + 1) / 2$, $MAX_SEQ = 2^n - 1$
- 即 $2^{(n-1)}$, 为退回N协议等的一半

3.5 数据链路协议实例

- **PPP**--点对点协议, 用于广域网, SONET和ADSL基于PPP
 - **Address和Contrl域可省略**
- **MAC**, 用于局域网, 以太网基于MAC

Chap 4 介质访问控制

- 介质访问控制协议(MAC)是数据链路层协议的一部分。

4.2 多路访问协议

4.2.1 ALOHA协议

- **纯ALOHA:**
 - **直接发送, 发送后监听是否冲突。**若冲突则延后随机时间重发。
 - 效率低, 信道利用率低, 冲突几率大。G=0.5时, 吞吐率S最大, 信道利用率=**0.184**
- **分槽ALOHA:**
 - 把信道时间分成离散时间槽, 槽长为一个帧所需的发送时间。只有在槽开始时可以发送。冲突后延迟随机个时间槽再发送。信道利用率=**0.368** (必须全局时间同步)

4.2.2 载波监听多路访问协议CSMA

监听时间 $2t$, 第一比特到电缆最远端时间 = t

- 使用条件: **传输延迟 << 发送延迟**--局域网传输距离小时有效。
- **载波监听 (发送前监听)**
- **1-坚持性CSMA**
 - 监听至信道空闲, 发送; 发送后冲突, 随机延迟重发。
 - 减少空闲时间, 增加冲突概率。
- **非坚持型CSMA**
 - 监听发现信道忙, 随机等待一段时间后重新监听
 - 减少冲突概率, 增加了空闲时间, 数据发送延迟增大
 - 效率比**1-坚持性CSMA**高, 延迟比**1-坚持性CSMA**大
- **p-坚持性CSMA**
 - **适用于分槽信道**
 - 监听, 发现空闲, 以概率p发送数据, 概率q=1-p延迟到下一时槽。
 - 监听, 发现忙, 等待下一时槽, 重新开始。

- **带冲突检测的CSMA-CSMA/CD**

- 发送期间**边发边监听**，**发现冲突，立刻终止**，并瞬间**发一个干扰**，使所有站点知道发生冲突。

4.2.5 无线局域网协议

- 有线网络同时只能有一个发送。无线网络的多个接受方只要都在对方范围外就可同时发送。
- ****差异****：可能无法给所有站发送帧，也无法接受所有站发送的帧。**冲突发生在接收方**
- **隐藏站**：竞争者离得太远，无法检测到潜在竞争对手。
- **暴露站**：发送站检测到别人的传输过程而影响发送（因为这个传输过程并不会影响接收站，发送站本应正常发送）
- A-->B. A send RTS(包含将要发送的数据帧长度). B send CTS(包含数据帧长度).A received, start transmitting.
- 发送方在期待时间内未听到CTS则重发RTS

4.3 以太网协议

- 经典以太网, CSMA/CD, 3-10Mbps
- 交换式以太网, 100/1000/10000Mbps
- IEEE定义了采用1-坚持型CSMA/CD技术的802.3局域网标准，与以太网协议略有差别。

4.3.1 经典以太网

- **曼彻斯特编码**，同轴电缆作为总线
- 以太网帧标准

以太网帧至少长64bit，头部长18bit，计算时不算前导码。

8	6	6	2	0-1500	0-46	4
前导码	目的地址	源地址	类型	数据	PAD	校验和

- 地址为MAC地址，前三字节为厂商代码，后三字节每个网卡不同
- **以太网地址全1表示广播**
- 类型>600H为以太网协议，<600H为IEEE802.3
- 数据段长度46-1500
- **最短帧长**
- 第一比特到电缆最远端时间 = t
- 帧发送时间 $\geq 2t$
- 网速提高：增大最短帧长/减少站间距离
- 二进制指数后退算法BEB：
 - 发生i次冲突后，在 $0-2^i-1$ 间选一个等待时槽数，开始重传。

4.3.4 交换式以太网

- Hubs集线器，把所有端口集中到一个CSMA/CD冲突域
- Switches交换机隔离每个端口为一个独立冲突域。

4.3.5 快速以太网

- 802.3u, Fast Ethernet 10-100Mbps
- 8B/6T--8位二进制映射到6位三进制上

4.3.6 千兆以太网

- 802.3z, 100-1000Mbps
- 所有配置都是点到点
- 全双工模式：交换机连接所有计算机，不用CSMA/CD
- 半双工模式：计算机连接集线器，使用CSMA/CD
 - 荷载扩充，帧串

4.8 数据链路层交换

- 网桥：工作在数据链路层的一种互联设备（看到的是帧），在互联的LAN之间实现帧的储存和转发。

4.8.1 网桥的使用

- 可以将高负载的LAN分割成几个LAN以减轻负担
- 隔离负载，防止故障损害全网
- 安全保密

4.8.2 学习网桥

- 相当于以太网交换机，接受所有和它相连的LAN上传送帧。
- 一开始地址/端口表为空。采用扩散算法转发帧。
- 通过源地址记录地址与端口对应关系。
- 地址/端口表不断更新，定时检查删除一段时间内未更新的项。
 - 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，并逆向学习。

4.8.3 生成树网桥

- 解决回路问题
- 构造生成树
 1. 每个桥广播自己的序列号，号最小的桥变成生成树的根；
 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. 当某个LAN或网桥发生故障时，要重新计算生成树；
 4. 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。

4.8.4 中继器\集线器\网桥\交换机\路由器\网关

- 物理层：中继器\集线器
- 数据链路层：网桥\交换机
- 网络层：路由器
- 传输层：网关

4.8.5 虚拟局域网

- **将逻辑拓扑结构与物理拓扑结构脱离**
- 在网桥或者交换机中建立配置表，指明了通过那些端口可以访问那些VLAN，一个端口可以被标记上多种VLAN颜色。
- 基于端口的划分、MAC地址、第三层协议或者IP地址的划分
- IEEE 802.1Q：在帧头添加VLAN标签
- 在转发该帧过程中第一个支持VLAN的网桥或者交换机加入VLAN标签，下行路径中最后一个支持VLAN的网桥或者交换机将这些VLAN标签去掉。

Chap5 网络层

5.1 网络层的设计问题

5.1.1 存储转发数据包交换

主机发送分组，会将分组传给最近的路由器，或者他自己LAN的路由器，点到点连接承运商的路由器。

5.1.2 提供给传输层的服务

由用户还是网络解决可靠性问题

- 面向连接的服务：**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)。**
- 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。将复杂的功能放在传输层。**

5.1.3 无连接服务的实现

数据报：每个分组携带完整的目的地址，**独立进行路由**。

5.1.4 面向连接服务的实现

虚电路：先建立虚电路，每个随后的分组都包含一个标示符，指明它属于哪个虚电路。

5.1.5 虚电路与数据报网络的比较

- 数据报：路由器不保留连接状态信息，崩溃时丢失分组，**很难实现服务质量，难以实现拥塞控制。**
- 虚电路：每个分组有一个短的虚电路号，经过失效路由器的虚电路被终止，在**有足够资源条件下可容易实现服务质量和拥塞控制。**

5.2 路由算法

网络层软件的一部分，确定一个进来的分组应该被传送到哪一条输出线路上。

- 1. 根据路由表查进来的分组对应的输出线路
- 2. 根据路由算法，填充和更新路由表
- 特性：正确性、简单性、鲁棒性、稳定性、公平性、最有性
- 分类
 - 静态路由算法：不能根据网络流量和拓扑结构变化更新路由表，使用静态路由表。**简单，开销少，灵活性差。**

- 动态路由算法：可以根据网络流量和拓扑结构变化更新路由表。**开销大，健壮性和灵活性好。**

5.2.1 优化原则

- 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。
- 汇集树：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。
- 路由算法：找出并使用汇集树。

5.2.2 最短路径算法

- Dijkstra算法
- 影响路径长度的因素：节点数量、地理距离、传输延迟、距离，信道带宽等

5.2.3 泛洪算法

- **静态路由算法**
- 把收到的每个分组，向除了分组到来的线路外所有输出线路发送
- 产生大量重复分组
- 解决方法：
 - 分组头包含计数器，每经过一站-1，到0时丢弃。
 - 路由器记录已经扩散过的分组序号。
- 鲁棒性好，很少直接使用，作为测量标准。

5.2.4 距离矢量算法

- **动态路由算法**
- 每个路由器维护一张表，表中给出了当前已知的到每个目的地的最佳距离和线路，并通过与相邻路由器交换距离信息来更新表；
- 问题：对好消息反应迅速，对坏消息反应迟钝
- **无穷计算问题**：无从知道他自己是否在这条路径上

5.2.5 链路状态路由

- 广泛应用于大型网络和Internet
- 1. 发现它的邻居节点，并了解其网络地址
 - 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一个LAN时，引入人工结点
 2. 设置到每个邻居节点的距离或者成本度量值
 - echo，时间/2
 3. 构造一个包含所有刚刚获知的链路信息包
 - 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；
 - 链路状态分组定期创建或发生重大事件时创建
 4. 将这个爆发送给所有其他路由器，并接受来自所有其他路由器的信息包
 - 头序号每次发送新分组时加1
 - 若是新的，则分发；若是重复的，则丢弃；
 - 每秒年龄-1
 5. 计算出到每个其他路由器的最短路径

5.4 服务质量

5.4.2 流量整形

- 定义：指调节进入网络的数据流的平均速率和突发性所采用的技术
- 减少拥塞、可以对业务流量进行监视
- **漏桶算法**：
 - 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包流。
 - **不看负载**，强迫输出按平均速率进行，不灵活。
- **令牌桶算法**：
 - 漏桶存放令牌，每 ΔT 秒产生一个令牌，令牌累积到超过漏桶上界时就不再增加。
- **$C+pS=MS$**
 - 突发时间的长度设为S秒，令牌桶的容量为C字节，令牌到达速率为p字节/秒，以及最大输出速率M字节/秒。

5.5 网络互联

5.5.1 网络如何不同

- 分组在网络接口遇到问题
- 不同网络的差错控制，流量控制和拥塞控制通常不同

5.5.2 何以连接网络

网络协议不同时，路由器需要提取出IP中的内容，转换为另一种网络的协议。

- 网络层：路由器在网络间储存转发分组，面对不同协议的网络，转换分组格式叫多协议路由。
- 传输层、应用层：传输/应用网关。
- 交换机以MAC地址为基础传输，不必理解其中的网络层协议
- 路由器需要提取出分组，利用分组信息决定目标去向，需要理解分组中的网络协议。

5.5.3 隧道

- 源于目的网络类型相同，之间网络的不同，可使用隧道
- 比如用IPV4封装一个IPV6的包

5.5.4 互联网路由

- 两级路由算法：**内部网关协议OSPF**、**外部网关协议BGP**
- 区别：法律不同、收费不同个、服务质量不同
- 独立于其他网络运营的网络（自治系统AS）

5.5.5 数据包分段

- MTU--最大传输单元
- 透明分段：拆分后重组；对网络开销大
- 非透明分段：拆分后不重组；对主机开销大，分组头增加对网络的开销。
- 非透明分组定义一个能通过所有网络的**基本段长度**
- IP协议采用非透明分组。

5.6 Internet的网络层

5.6.1 IPv4协议

- IP数据报=分组头+正文
- 分组头：20字节固定+变长（**长度为4的倍数**，不够则填充，最长40字节）
- TTL：每经过一个路由器-1
- 段偏移量：除最后一个段外所有段长度为8的倍数（基本段长）
- Protocol：使用哪种传输层协议

5.6.2 IP地址

- 32位
- 每个主机和路由器的每个接口都有一个IP地址，IP地址唯一。
- 分类：ABCDE类。
- A:0|7Networks|24Host|
- B:10|14Networks|16Host|
- C:110|21Networks|8Host|
- D:1110|组播地址|
- E:1111|预留|
- **全0：本网络或本主机**
- **全1：广播地址**
- **127开头回路测试**
- 128.208.128.0/17，17代表网络部分的长度
- 网络号：子网号：主机号（子网号源于主机号）
- **计算可用子网/主机数量记得减去全1和全0两种情况**

CIDR

- ip地址即将用完，基于分类的IP地址浪费大量地址
- 将剩余C类地址分成大小可变的地址空间
- 路由表（IP地址，子网掩码，输出线路）
- 多个匹配时选择掩码长的路由表项

NAT

- 每个公司分配一个或少量的IP地址，用于传输Internet流量
- 公司内部，每台计算机有唯一的IP地址，用来传输内部流量。
- 当一个分组离开公司的网络，发向ISP它需要一个地址转换。
- 私有地址：包括这些地址的分组不应出现在Internet上。
- 使用传输层的端口号
- 缺陷：
 - 违反IP的结构模型，重复使用私有地址
 - 将Internet从一个无连接的网络变成一个面向连接的网络；
 - 违反最基本的协议分层规则
 - Internet上层的进程不一定使用TCP或者UDP
 - 有些应用会在正文内容中插入IP地址，例如FTP

5.6.3 IPv6协议

- 支持大量主机
- 减少路由表，简化协议，更好的安全性，更好的扩展功能
- 地址：32-->128bit。8组4个16进制，':'隔开
- IP头简化，定长40Byte.

5.6.4 Internet控制协议

- **ICMP**Internet控制消息协议
 - 提高IP数据报交付成功的机会，发生以外时，意外事件消息由ICMP报告
 - 主要用于报告出错和测试
 - 封装在IP包内
- **ARP**地址解析协议
 - 解决IP与MAC映射问题
 - 向每个主机询问发ARP请求，主机回应自己的IP和MAC
 - 建立ARP表，把结果缓存下来。
 - 主机每次启动广播自己的IP&MAC
- **DHCP**动态主机配置协议
 - 每个网络有一台DHCP服务器负责地址配置。
 - 发出DHCP请求，请求主机IP地址。
 - 每个LAN需要一个DHCP中继代理

5.6.6 OSPF-内部网关路由协议

- 自治系统AS**内**使用的路由算法RIP,OSPF
- OSPF--开放最短路径优先
 - 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，网络结点到路由器结点的弧权为0；
- 分层路由
 - 划分区域，区域互不重叠，不必覆盖所有网络/路由器
 - 主干区域0，所有区域与主干区域相连
 - 区域内部，每台路由器都有同样的链路状态数据库，并运行同样最短路径算法，在一个区域中至少有一台路由器连接到骨干上。
- 三种路由：区域内、区域间、自治系统间。
- 四类路由器，允许重叠：
 - 完全在一个区域内的内部路由器；
 - 连接多个区域的区域边界路由器；
 - 骨干路由器；
 - 自治系统边界路由器
- **工作过程：链路状态路由算法**
 - 邻居之间发送HELLO消息
 - 邻接的路由器之间交换信息
 - 5种消息类型

5.6.7 BGP-外部网关路由协议

- 自治系统AS之间使用的路由算法BGP
- 外部网关与内部网关的区别：路由策略考虑政治、安全、经济因素
- 通过TCP连接传送路由信息；
- 中转流量的三类网络：
 - 末端网络
 - 多连接网络
 - 穿越网络
- 采用路径向量（path vector）算法
 - 与距离矢量法非常类似
 - 每个边界网关向邻居路由器广播到目的节点完整路径

Chap6 传输层

6.1 传输服务

6.1.1 提供给上层的服务

- 最终目标：向它的用户提供高效的、可靠的和成本有效的数据传输服务。
- 面向连接的传输服务：建立连接、数据传输、连接释放。
- 无连接的传输服务
- 消除网络层的不可靠性；网络层--ISP，传输层--用户。
- 一套标准写成的代码可以运行在各种不同的网络上。

6.1.2 传输服务原语

- 通过传输服务原语访问传输服务
- 传输实体间发送的为段
- TPDU:Transport Protocal Data Unit帧(分组(段(负载)))
- 释放连接方法：
 - 不对称方法：任意一方都可以关闭连接；
 - 对称方法：每个方向的连接单独关闭，双方都disconnect才能关闭整条连接；

6.1.3 Berkely套接字

原语	含义
SOCKET	创建一个新通信端点
BIND	将套接字与一个本地地址关联
LISTEN	声明愿意接受连接；给出队列长度
ACCEPT	被动创建一个入境连接
CONNECT	主动创建一个连接
SEND	通过连接发送一些数据
RECEIVE	从连接上接受一些数据

原语	含义
CLOSE	释放连接

- SOCKET规定了之后用到的地址格式，期望的服务类型，返回一个文件描述符
- 调用ACCEPT后程序被阻塞。建立连接的TPDU到来后新建一个SOCKET，创建一个子进程负责此次连接，另一个继续等待连接请求。
- bind：服务器调用，用户端不调用。
- 调用CONNECT后阻塞用户端，连接建立后取消阻塞。

6.1.4 套接字编程实例：Internet文件服务器

6.2 传输协议的要素

- 传输服务需要解决：**差错控制、分组顺序、流量控制**等问题。
- 与数据链路层的区别
 - 传输层需要显式地给出目的端的地址
 - 传输层初始连接复杂
 - 传输层源端和目的端之间的子网具有存储能力
 - 传输层有大量动态变化的连接，和数据链路层的数据缓冲和流量控制具有不同的处理方法

6.2.1 寻址

- 连接时需要确立传输服务访问点TSAP(Transport Service Access Point)
- Internet中TSAP为(IP Address, local port)
- 客户端如何得知服务程序的TSAP
 1. 预先约定
 2. 用户与端口映射器建立连接，询问需要的服务所在的端口。端口映射器返回TSAP地址。
 - 新服务需要向端口映射器注册服务名称和TSAP。
- 初始连接协议
 - 进程服务器（process server）进程(inetd)同时在一组端口上监听；以等待外来的连接请求。
 - 如果用户请求的服务没有在TSAP上监听，连接到进程服务器。进程服务器启动对应的服务进程
 - 客户与对应的服务进程连接，使其继承与用户的连接。进程服务器继续监听。

6.2.2 建立连接

- 对于重复延迟包的解决方法
 1. 使用一次性传输地址
 2. 发起方为每个连接分配一个连接标识符，即序号，序号增一，当一个连接请求到来时，察看记录表，看是否已用过，需要每个传输层实体无限期地维护一定数量的历史信息
- 分组生存期的限定方法
 1. 受限制的子网设计
 2. 在每个分组中放置一个跳计数器
 3. 分每个分组打时间戳
- TCP三次握手
 - A 发出序号为X的CR TPDU；
 - B 发出序号为Y的CC TPDU并确认A的序号为X的CR TPDU；
 - A 发出序号为X的第一个数据TPDU，并确认B的序号为Y的CR TPDU。

6.2.3 连接释放

- 非对称：一方释放连接，存在丢失数据的危险
- 对称式：三次握手+定时器
 1. 主机1：发送DR并启动计时器
 2. 主机2：接受DR，回复DR并启动计时器
 3. 主机1：接受DR，释放连接，回复ACK
 4. 主机2：接受ACK，是否连接。
 - 主机1若未收到DR回复，超时重新发送DR，N次超时后释放连接。
 - 主机2回复DR后未收到ACK，超时后释放连接。

6.2.4 差错控制和流量控制

- 流控：传输层利用**可变滑动窗口协议**来实现流控。
 - 发送方的**发送窗口大小**是由**接收方的实际缓存**情况给出
 - 为避免控制TPDU丢失导致死锁，主机应该周期性的发送TPDU

6.2.5 多路复用

- 向上多路复用：多个服务只有一个网络地址
- 向下多路复用：一个服务使用多个网络地址，提高可靠性。

6.2.6 崩溃恢复

6.4 Internet传输协议：UDP

6.4.1 UDP概述

无连接的端到端传输协议

- 发送经过封装的IP数据报，**不必建立连接**
- 不考虑流控、错误控制，不确认
- 使用UDP的上层协议：RIP，路由信息周期发送；DNS，避免TCP建立连接的延时；SNMP，网络管理协议。

6.5 Internet传输协议：TCP

6.5.1 TCP概述

- **基于字节流**
- 接受本地进程字节流，分割成不超64KB的分片，以单独的IP数据报的形式发送，到另一端时由TCP传输实体重构。

6.5.2 TCP服务模型

- 访问TCP服务通过收发双方创建套接字实现
- **端口：16bit。1024以下为知名端口。**
- 每条连接用（套接字1，套接字2）表示
- TCP连接**全双工，点对点，不支持多播和广播**

- 基于字节流，不保留消息边界
- 可以是立刻发送/缓存发送。
- 强迫发送：PUSH。紧迫数据用URGENT标记。

6.5.3 TCP协议

- **TCP段头**：20Byte
- 段大小满足IP数据包最大 2^{16} byte和链路层以太网MTU：1500byte。
- 使用滑动窗口，确认序号=准备好接受的下一个序号。

6.5.4 TCP段的头

- 定长20Byte。TCP头长4bit。

6.5.5 TCP连接建立

- 三次握手建立链接
 - 服务器：LISTEN & ACCEPT
 - 客户端：**CONNECT (syn=1, ack=0)**
 - 服务器：没有服务进程监听此端口：回RST。有监听：回`syn=1,ack=1`确认连接
 - 客户端：回`syn=0, ack=1`确认连接
- 同时试图建立连接时只能建立一条连接。

6.5.6 TCP连接释放

- 单独释放一对单工连接
- 发出FIN位置的TCP段启动定时器，收到确认后关闭连接。无确认则在超时报关闭连接。

6.5.8 TCP滑动窗口

- 基于确认和可变窗口大小
- 窗口大小为0时不能再发TCP段
 - 可以发紧急数据
 - 可以发1字节的TCP段防止死锁
- Nagle算法：应用程序每向传输实体发出一个字节，传输实体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发出并对再收到的字节缓存，直至收到下一个确认；
- 傻窗口症状：当应用程序一次从传输层实体读出1个字节时，传输层实体会产生一个1字节的窗口更新段，使得发送方只能发送一个字节；
- 解决：Clark算法-->限制接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。

6.5.10 TCP拥塞控制

- TCP假定**分组超时由拥塞引起**
- 网络传输能力造成问题：
 - **发送方维护两个窗口：可变发送窗口和拥塞窗口**，按两个窗口的最小值发送
 - **拥塞窗口**：任何时间都可以往网络发送的字节数。

- 接收方接收能力问题：
 - 在连接建立时声明最大可接收段长度
 - 利用可变滑动窗口协议防止出现拥塞
- 慢启动算法：
 - 连接建立时拥塞窗口（congwin）初始值为该连接允许的最大段长
 - 发出一个最大段长的TCP段，若正确确认，拥塞窗口变为两个最大段长；
 - 发出n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，直至发生丢包超时事件，设置拥塞窗口的大小。

Chap7 应用层

7.1 DNS-域名系统

- 多层次、基于域的命名系统。使用分布式数据库实现
- 将主机名/电子邮件目标映射为IP
- 请求**使用UDP**
- 域名解析时，调用解析器。先向本地DNS查询。若无查询结果，底层DNS服务器暂时变为高层DNS服务器的客户，向上查询，**递归解析**。
- **典型的用户/服务器交互系统。**

7.1.1 DNS名字空间

- **分层次的地址结构**，可以树形表示。
- 分为**200个顶级域**：通用域&国家域
- 域名**最长255字符**，每部分**最长63字符**

7.1.2 域名资源记录

- 每个域名有**与它对应的资源记录**。一般为IP地址
- 资源记录**五元组**：Domain_name(域名)；Time_to_live(生存期)；Class(类别)；Type(类型)；Value(值)。
Class总为IN

7.1.3 名字服务器

- DNS将域名空间划分为许多**无重叠的区域(zone)**，每个区域覆盖了域名空间的一部分并设有域名服务器对这个区域的域名进行管理。
- 每个区域有一个主域名服务器和若干个备份域名服务器
- 区域的**边界划分是人工设置的**
- 域名解析
 - 先查本区域的域名服务器
 - 找不到则本区DNS向对应顶级域名服务器查询
 - 顶级域的域名服务器通过向下的层次查询得到对应的资源记录，返回给该域名服务器
 - 资源记录返回给发起域名解析的机器，在该区域的DNS服务器中缓存。