#### 第二章

进程通信：

1、同一个主机上的两个进程：进程间通信，通常由操作系统决定

2、在不同主机上的两个进程：通过交换信息通信

TCP：面向连接的，可靠的，进行流量控制和拥塞控制

UDP：不可靠传输

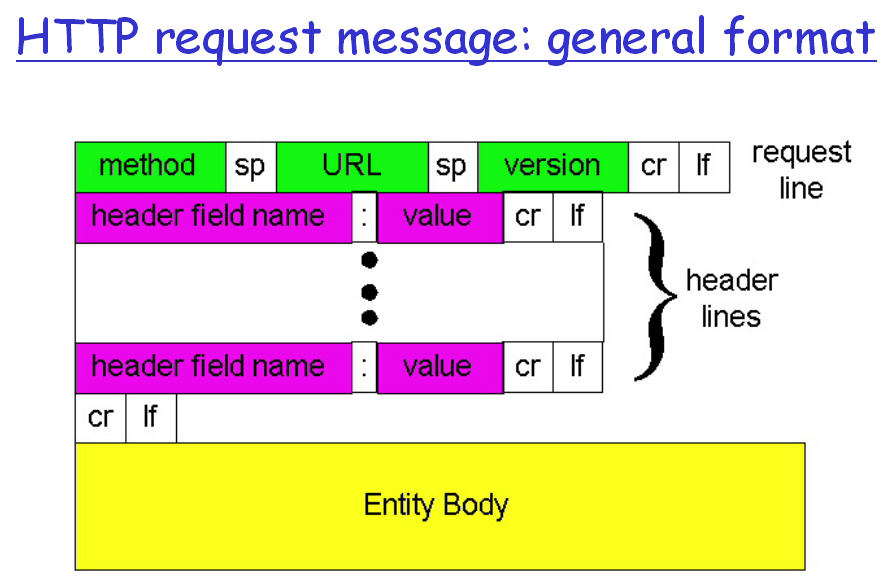
URL：统一资源定位符

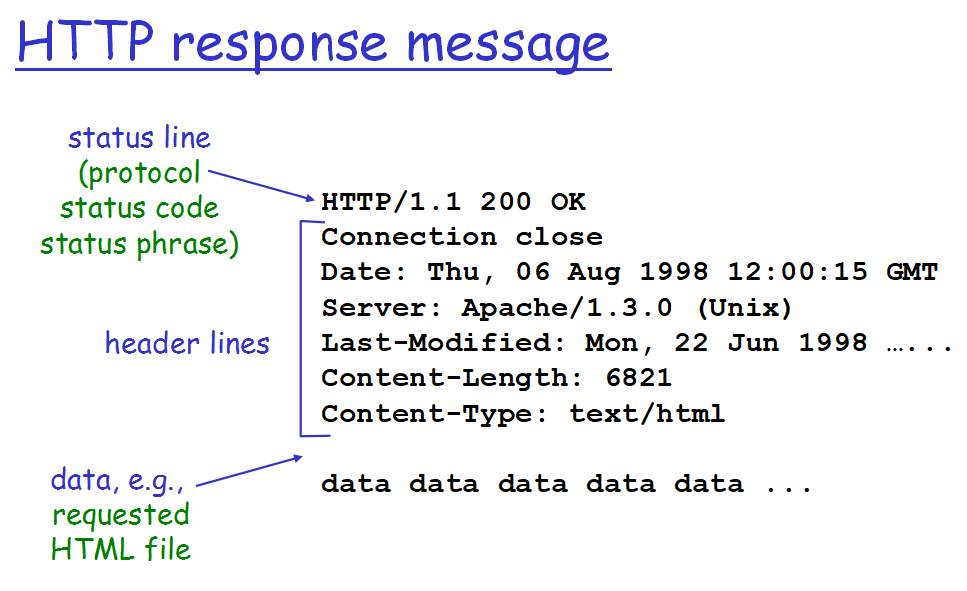
Nonpersistent Http:最多一个对象在TCP连接中

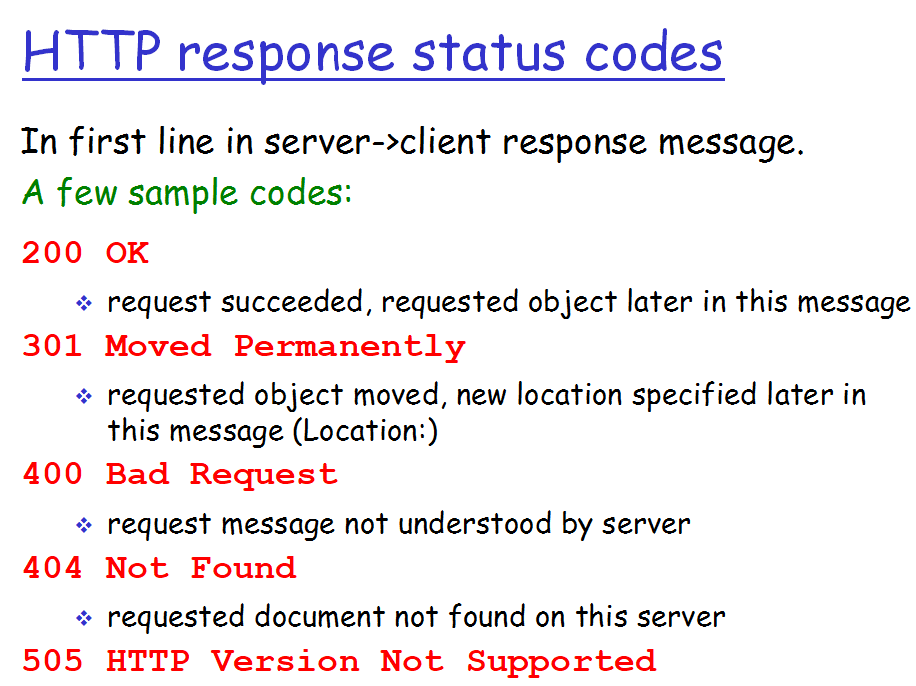
Persistent：同一条TCP连接可以传输多个对象

Post method: 网页经常包括表单输入。输入数据通过entity body上传到服务器

Get method：输入通过请求行上的URL上传到服务器







Cookies：Keeping state

Web缓存(代理服务器）：浏览器的所有http请求都会发送到代理服务器，如果请求的对象在代理服务器的话，那么就返回这个对象副本。如果没有，代理服务器就会向原始服务器发出请求并得到对象，得到对象之后返回到浏览器

代理服务器优点：

1. 降低响应时间
2. 减少原始服务端的流量
3. 让一些性能比较差的内容提供者可以高效地提供内容给客户。

条件GET方法：保证代理服务器不会发送已经在原始服务端已经修改过但没有在代理服务端修改过的对象。

FTP：文件传输协议

控制连接和数据连接：属于TCP连接，控制连接属于带宽外连接，即数据连接和控制连接不属于同一条连接。控制连接一直保持连接。数据连接传输完一个文件之后就关闭。

Email特点：

1. 异步通信媒体
2. 综合媒体

电子邮件三大部分：

1. 用户代理
2. 邮件服务器
3. SMTP简单邮件传输协议

Mail Servers：

邮箱：存储用户的incoming messages

报文队列：outgoing mail messages

SMTP的特点：

1. TCP连接
2. 在发送服务端和接收服务端直接传输
3. 传输包括三个阶段：三次握手、传输报文、关闭
4. 命令行用ASICC码
5. 回复用状态码和短语
6. 报文使用7位的ASCII码

SMTP和HTTP比较：

1. SMTP使用persistent connection
2. SMTP要求报文使用7位ASCII编码
3. SMTP使用点号.来作为报文的结尾
4. Http是拉协议，SMTP是推协议。
5. HTTP和SMTP都有ASCII编码的命令/请求交互，状态码

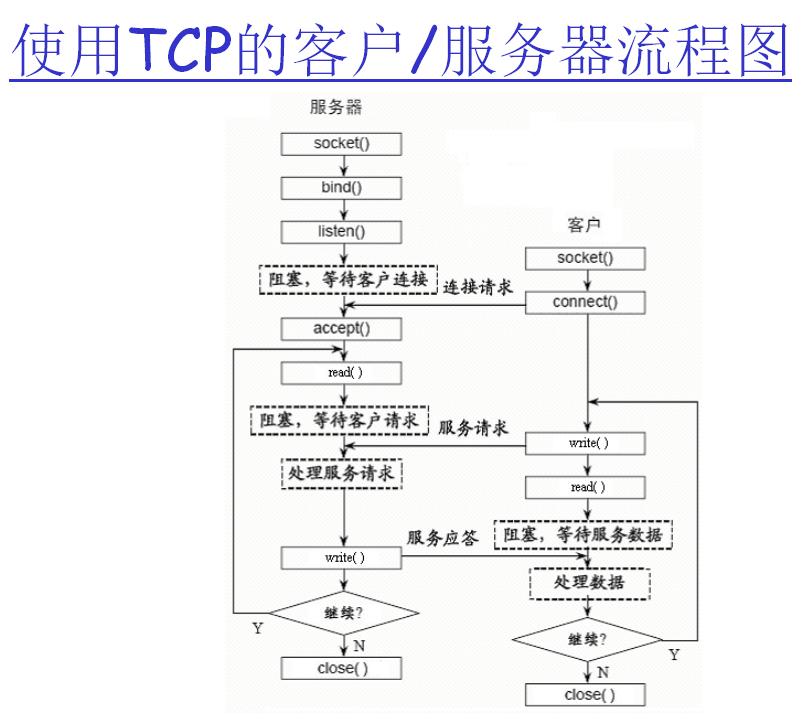
DNS：Domain Name System（域名系统)，用于主机名到IP地址转换的目录服务。

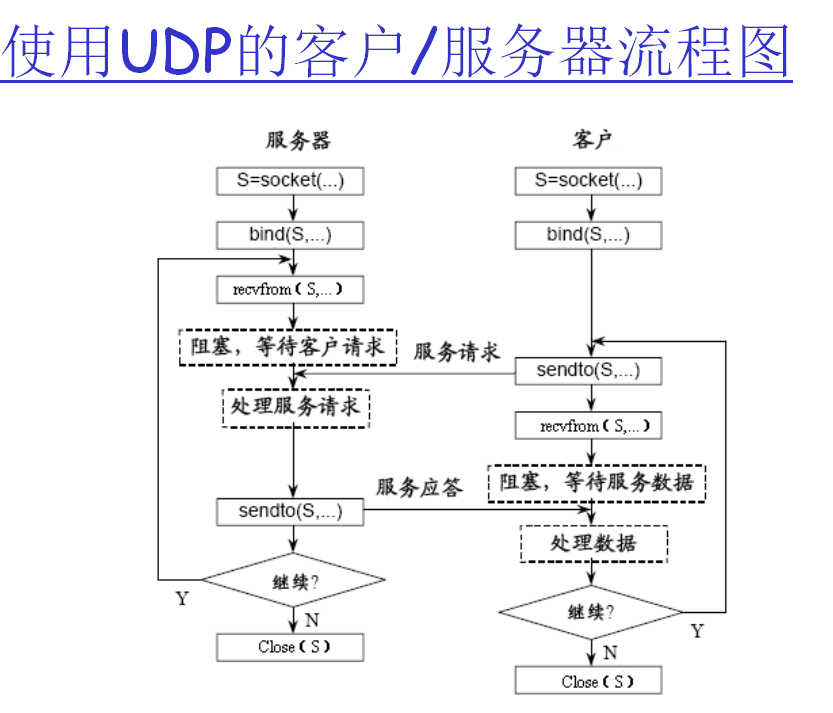
DSN是：

1. 一个由分层的DNS服务器实现的分布式数据库
2. 一个使得主机能够查询分布式数据库的应用层协议

Example Client wants IP for [www.amazon.com;](http://www.amazon.com;)

1. 客户端请求根服务器来寻找





#### 网络层

网络层主要功能：转发、路由选择、建立虚连接（一些网络层）

网络层中的连接和无连接服务：

1. 数据报网络提供网络层无连接服务
2. 虚电路网络提供网络层有连接服务

虚电路的组成：源和目的主机的路径、VC号、路由器中的转发表

属于一条虚电路的分组将在它的首部携带一个VC号。一条链路的VC号可能会改变。

转发表：路由器保存连接状态的信息

信令协议：用于虚电路的连接，数据传输和虚电路的断开。

在虚电路建立阶段，发送运输层与网络层联系，指定接收方地址，等待网络建立虚电路。网络层决定发送方与接收方之间的路径，即该虚电路的所用所有分组要通过的一系列链路与路由器。网络层也为沿着该路径的每条链路决定一个VC号。最后，网络层在沿着路经的每台路由器的每台路由器的转发表中增加一个表项。在虚电路建立期间，网络层还可以预留该虚电路路径上的资源。

数据报网络：无需建立虚电路，路由器不维护任何虚电路的状态信息。

转发表：最长前缀匹配规则

路由器功能：转发和路由选择

交换结构：

1. 经内存交换：最简单的交换结构
2. 经总线交换：总线带宽限制了交换速率
3. 经互联网交换：克服总线带宽的限制。纵横式互联网结构是一种由2N条总线组成的互联网络，但某个时刻仅有一个分组能够发送

输出端口：交换结构、排队（缓存管理）、数据链路层处理（协议封装）、线路端接

输出和输入端口出现排队：交换速率大于输出速率输出端口出现排队；交换速率小于输出速率输入端口出现排队。当缓存不够时就会出现丢包情况。

线路前部阻塞：即在一个输入队列中排队的分组必须等待通过交换结构发送（即使输出端口是空闲的），因为它被位于线路前部的另一个分组所阻塞。

网络层三个组件：

1. IP协议
2. 路由选择协议
3. 报告数据报中的差错和对某些网络层信息请求进行响应的措施的ICMP协议

IPv4数据报格式：

IP数据报分片

MTU：最大传送单元，一个大的数据报需要分成多个小的数据报，每个数据报开头用来识别是否是同一个数据报的，并且它们是按序分开。这些小的数据报都需要在目的组装。每个数据报的开头是由标识、标志和偏移字段放在IP数据报首部中。标识符用于确定是否是同一个数据报。最后一个片的标志位设为0而所有其他片的标志位设为1用于确定是否已经完全收到所有片。偏移字段用于最后组装成数据报。

IPv4编制：IP地址由32位组成，作为主机和路由端口的标识符。

相同子网的主机可以不经过路由直接通信。

无类别域间路由选择：因特网地址分配策略，同一个子网的设备具有相同的子网。

怎样获得一个IP地址：

DHCP：动态主机配置协议。plug-and-play。动态获取临时的IP地址。工作分为四个步骤：

1. DHCP discover
2. DHCP offer
3. DHCP request
4. DHCP ack

NAT（网络地址转换）：家庭网络所有设备流向外界的数据报都有同一个IP地址，而所有外界流向这个家庭网络的数据报都有同一个IP地址。

NAT转换表：源IP地址、端口号与NAT IP地址、端口号之间的转换表。

NAT争议处：妨碍P2P的发展。

解决方法：

1. 连接反转：
2. UPnP：将NAT后面的主机IP地址和端口号映射到NAT公共IP地址和指定端口号，然后通过NAT直接与外界进行TCP连接。
3. 使用Skype中继：外界客户端端和NAT客户端都连接中继，中继桥接这两个连接。

因特网控制报文协议（ICMP）：被主机和路由器用来彼此沟通网络层的信息，最典型的用途是差错报告 。

IPv6：32为IP地址不够用

IPv6报文结构：

流量类型：表示流中各个数据报之间的优先级

流标签：确保多个数据报在同一个流中

下一个首部：标识上层的协议（TCP或者UDP）

IPv6显著变化：

1. 扩大的地址容量
2. 简化高效的40字节首部
3. 流标签和优先级

路由选择算法：

1. 链路状态路由选择（LS）。全局式路由选择算法。Dijkstra算法
2. 距离向量算法（DV）。分散式路由选择算法。Bellman-Ford算法

LS与DV路由选择算法的比较：

1. 报文复杂性：LS算法需要知道网络中每条链路的费用。DV算法只在邻居间交换报文。
2. 收敛速度：LS算法收敛速度比较快。DV收敛时可能会遇到路由选择环路，DV算法会遭遇无穷计数的问题。
3. 健壮性：一个LS结点仅计算自己的转发表，其他结点自行执行类似的计算。这就意味着LS算法下，路由计算在某种程度上是分散的。DV算法的中一个结点的计算会传递到它的邻居，一个不正确结点计算值会扩散到整个网络。

层次路由选择：将多个路由器集合在同一个自治系统（autonomous system），在同一个自治系统的路由器使用同一个路由算法，不同自治系统之间可以使用不同的路由选择算法。

网关路由器：负责本AS之外的目的地转发分组。

自治系统间路由选择协议：从相邻的AS获取可达性信以及向该AS中的所有路由器传播可达性信息。

热土豆路由选择算法：选择达到网关最小费用的路径已确定转发分组的路由网关。

因特网中自治系统内部的路由选择：RIP

RIP：使用的是DV算法，每30秒通过RIP响应报文在路由器邻居之间交换路由选择更新信息。

路由选择表：如果一台路由器经过180秒没有从邻居中听到报文，则该邻居不再被认为是可达的。当这种情况发生时：

1. RIP修改本地路由选择表
2. 本地路由向邻居发送通告
3. 如果邻居路由选择表会更新，邻居向周围邻居发送通告
4. 链路错误信息会传播到整个网络。

RIP响应报文通过UDP协议传播。

OSPF：开放最短路径。使用链路状态算法，Dijkstra算法。使用OSPF时，路由器向自治系统内所有其他路由器广播路由选择信息，而不仅仅是向其相邻路由器广播。

自治系统间的路由选择：BGP（边界网关协议）

BGP作用：

1. 从相邻AS处获得子网可达性信息
2. 向本AS内部所有路由器传播这些可达性信息
3. 基于可达性信息和AS策略，决定达到子网的路由。

广播路由选择：网络层提供了从一种源结点到网络中所有其他结点交付分组的服务。

多播路由选择：使单个源结点能够向其他网络结点的一个子集发送分组的副本。

广播过程中的资源复制效率是很低的。

洪泛：源结点向它的所有邻居发送分组的副本。

广播风暴：导致无休止的广播分组复制，最终导致在该网络中生成大量的广播分组，使得网络变得毫无用处。

受控洪泛：结点不广播以前广播过的分组

最小生成树：避免网络中的环路。

基于中心的方法：建立一棵生成树时，要定义一个中心结点。结点则向中心结点单播加入树报文。加入树报文使用单播路由选择朝着中心结点转发，直到它达到一个已经属于生成树的结点或达到该中心。在任意一种情况下，加入树报文经过的路径定义了发起加入树报文的边缘结点和中心之间的生成树分支。这个新分支能够 认为已被嫁接到现有的生成树上了。

#### 链路层

链路层寻址和ARP

MAC地址：主机和路由器的适配器（即网络接口）具有链路层地址。因此，具有多个网络接口的主机或路由器将具有与之相关联的多个链路层地址。链路层交换机不具有与它们的接口相关联的链路层地址，这是因为链路层交换机的任务是在主机与路由器之间承载数据报，交换机透明地执行该项任务，主机或路由器不必明确的将帧寻址到其间的交换机。

ARP地址解析协议：网络层地址与链路层地址之间的转换协议。发送主机中的ARP模块将取相同局域网上的任何IP地址作为输入，然后返回相应的MAC地址。

ARP协议：发送主机ARP表上没有目的主机MAC地址，发送主机广播ARP请求报文，该报文包含目的主机的IP地址，但目的主机收到该ARP请求报文时，向源主机回复包含其MAC地址的报文，源主机根据这个MAC地址发送数据报。

#### 第一章

协议分层：

1. 应用层：是网络应用程序以及它们的应用层协议残留的地方。位于应用层的信息称为报文。
2. 运输层：在应用程序端点之间传送应用层报文。运输层协议包括TCP和UDP。运输层分组称为报文段
3. 网络层：网络层负责将数据报从一台主机移动到另一台主机。
4. 链路层：将帧从一个网络元素移动到邻近的网络元素
5. 物理层：将帧中的一个一个比特从一个结点移动到下一个结点

运输层

**传输层与网络层的关系**：

**网络层**

功能：主机通信

协议：IP

特点：hop-by-hop

**传输层**

功能：进程通信

协议：UDP和TCP

特点：end-to-end

**多路复用**：在源主机从不同的套接字中收集数据块，并为每个数据块封装上首部信息从而生成报文段，然后将报文段传递到网络层，所有这些工作称为多路复用。

**多路解用**：将运输层报文段中的数据交付到正确的套接字的工作称为多路分解。

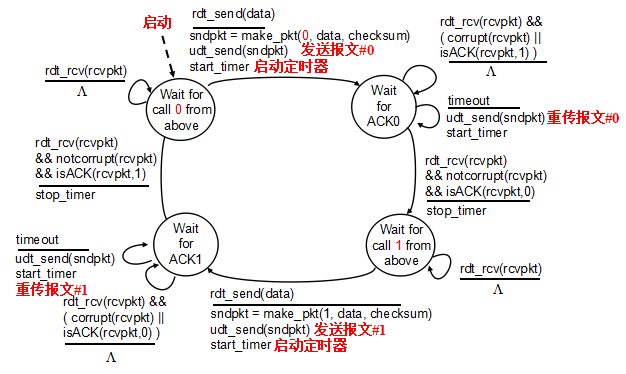
IP和UDP的关系：

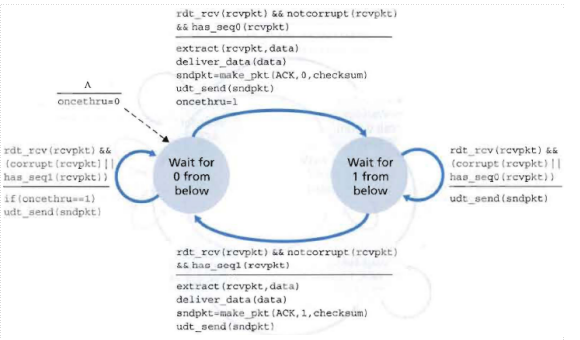


UDP优点：

1. 适用于实时系统
2. 无需建立连接，因此不会有建立连接的延时
3. 无连接状态。TCP需要在端系统中维护连接状态。次连接状态包括接收和发送缓存，拥塞控制参数以及序号与确认号参数
4. 分组首部开销小。
5. 适合多媒体应用开发人员。

TCP可靠数据传输原理：



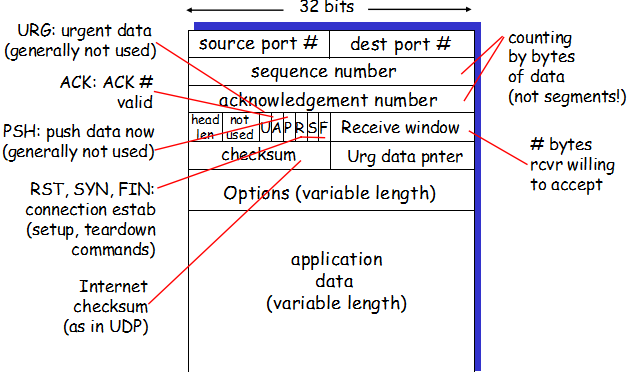


**基于流水线可靠数据传输协议**：

**回退N步**：发送方有N个未确定的包在流水线中，接收方只发送累积的acks，发送方有计时器，如果超时，重发所有未被确认的包。接收方收到失序的包丢掉这个包。

**选择重传**：发送方有N个未确认的包在流水线中，接收方确认单个发送过来的包，不需要累计确认。发送方为每一个发送的包都设置计时器，如果超时，只重传超时的包。在接收方缓存失序的包，当缓存内的包都有序时才交付给上层。

TCP报文结构：



1、32比特的序号字段和确认号字段

2、16比特的接收窗口字段，该字段用于流量控制。

3、4比特的首部长度字段

1. 可选与变长的选项字段，该字段用于发送方与接收方协商最大报文端长度（MSS）时，或在告诉网络环境下用做窗口调节因子时用。
2. 6比特的标志字段。ACK，RST、SYN、FIN、PSH和URG。

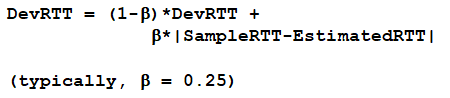
**往返时间的估计与超时**：

往返时间估计：指数加权移动平均。



超时：提供了受限的拥塞控制





TCP的可靠数据传输原理：

1. 基于RDT协议运行在IP不可靠服务之上的数据传输
2. 流水线
3. 累计确认
4. 使用一个重传计时器

重传触发的条件：

1. 超时
2. 收到重复的acks

TCP流量控制：发送端维护一个接收窗口来提供流量控制。保证发送方未被确认的包小于接收端的剩余缓存空间。

**TCP拥塞控制原理**：

**拥塞代价**：

1. 当分组的到达速率接近链路容量时，分组经历巨大的排队时延。
2. 发送方必须要执行重传以补偿因缓存溢出的而丢弃的分组。
3. 发送放在遇到大时延时所进行的不必要重传会引起路由器利用其链路带宽来转发不必要的分组副本
4. 当一个分组沿一条路径被丢弃时，每个上游路由器用于转发该分组到丢弃该分组而使用的传输容量最终被浪费掉。

**TCP拥塞控制方式**：

1、当拥塞控制窗口大小小于慢启动阈值的时候，进入慢启动模式。拥塞窗口大小以指数式增长。

2、当拥塞窗口在慢启动阈值之上时，发送者进入拥塞避免模式。拥塞控制窗口以线性方式增长。

3、当收到3个ACK时，慢启动阈值设为CongWin/2，CongWin设置为慢启动阈值。

4、当超时时，慢启动阈值设置为CongWin/2, ConWin设置为1个MSS。

MSS：最大传输单元。

拥塞控制方法

1. 端到端拥塞控制。在端到端拥塞控制方法中，网络层没有为运输层拥塞控制提供显式支持。
2. 网路辅助的拥塞控制。在网络辅助的拥塞控制中，网路层构件（即路由器）向发送方提供关于网路中拥塞状态的显式反馈信息。

ABR拥塞控制：当网路拥塞时，ABR服务会将其传输速率一直为某些预先确定的最小传输速率；当网络拥塞时，ABR服务会将其传输速率抑制为某些预先确定的最小传输 速率。

资源管理信元：用于传递拥塞控制信息，可用来提供直接网路反馈和经由接收方的网络反馈。

EFCI比特：显式转发拥塞指示比特。某拥塞的网络交换机可把一个数据信元中EFCI比特设置为1来向目的主机发送网络已经拥塞的信令。

CI：拥塞指示比特；严重拥塞时，把CI比特置为1；

NI：无增长比特；轻微拥塞时NI比特置为1；

ER：显式速率。一个拥塞的交换机会降低经过的RM信元中的ER字段包含的值。