



西南交通大学

第六章

网络层

设计、制作、讲授：谭献海

EMAIL: xhtan@home.swjtu.edu.cn

6.1 网络层概述

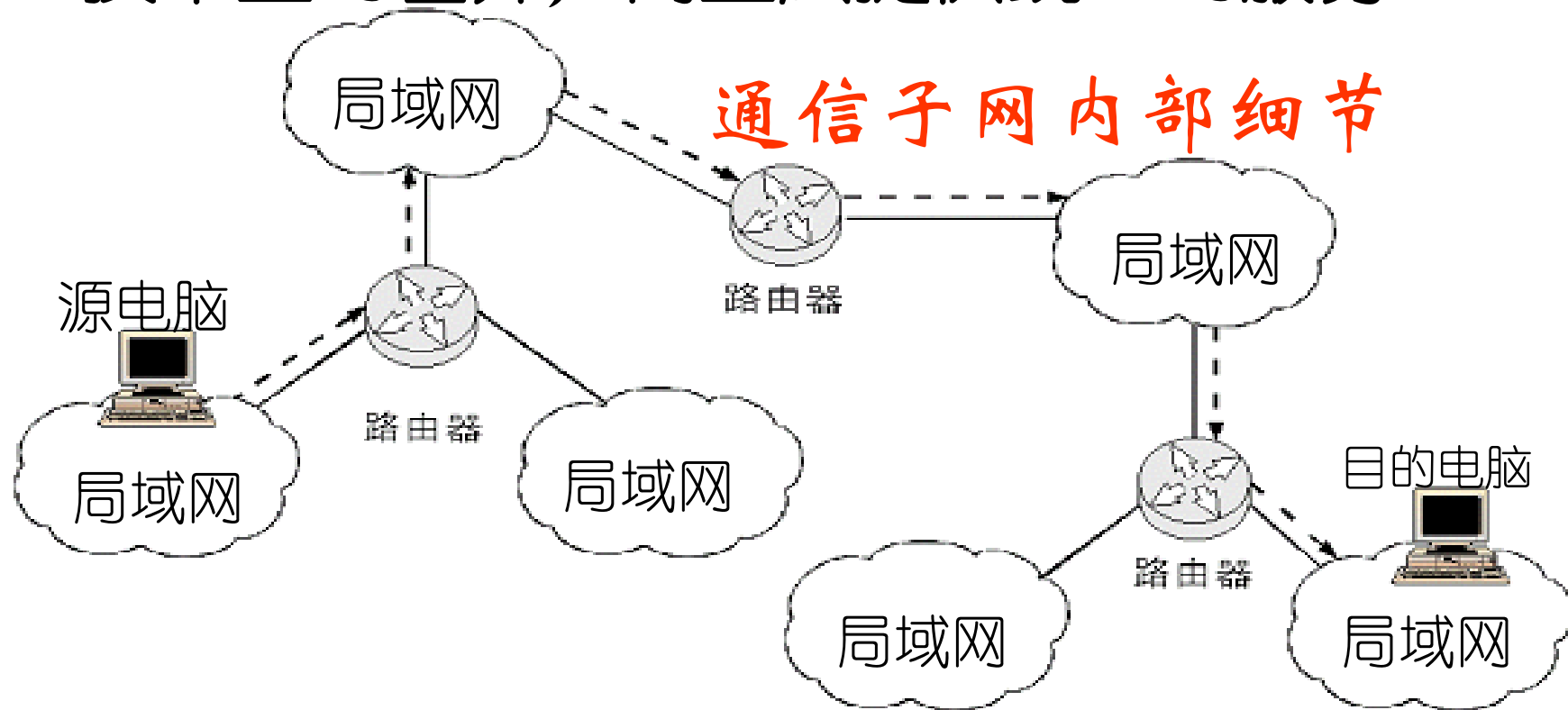


本节主要内容

- 网络层概述
- 网络层提供的服务
- Internet网络层协议
- 拥塞控制方法（自学）

网络层所处的位置及职责

- **基本职责：**实现通信网内端到端（源IMP到目的IMP）的可靠传输，屏蔽不同通信子网技术上的差异，向上层提供统一的服务。

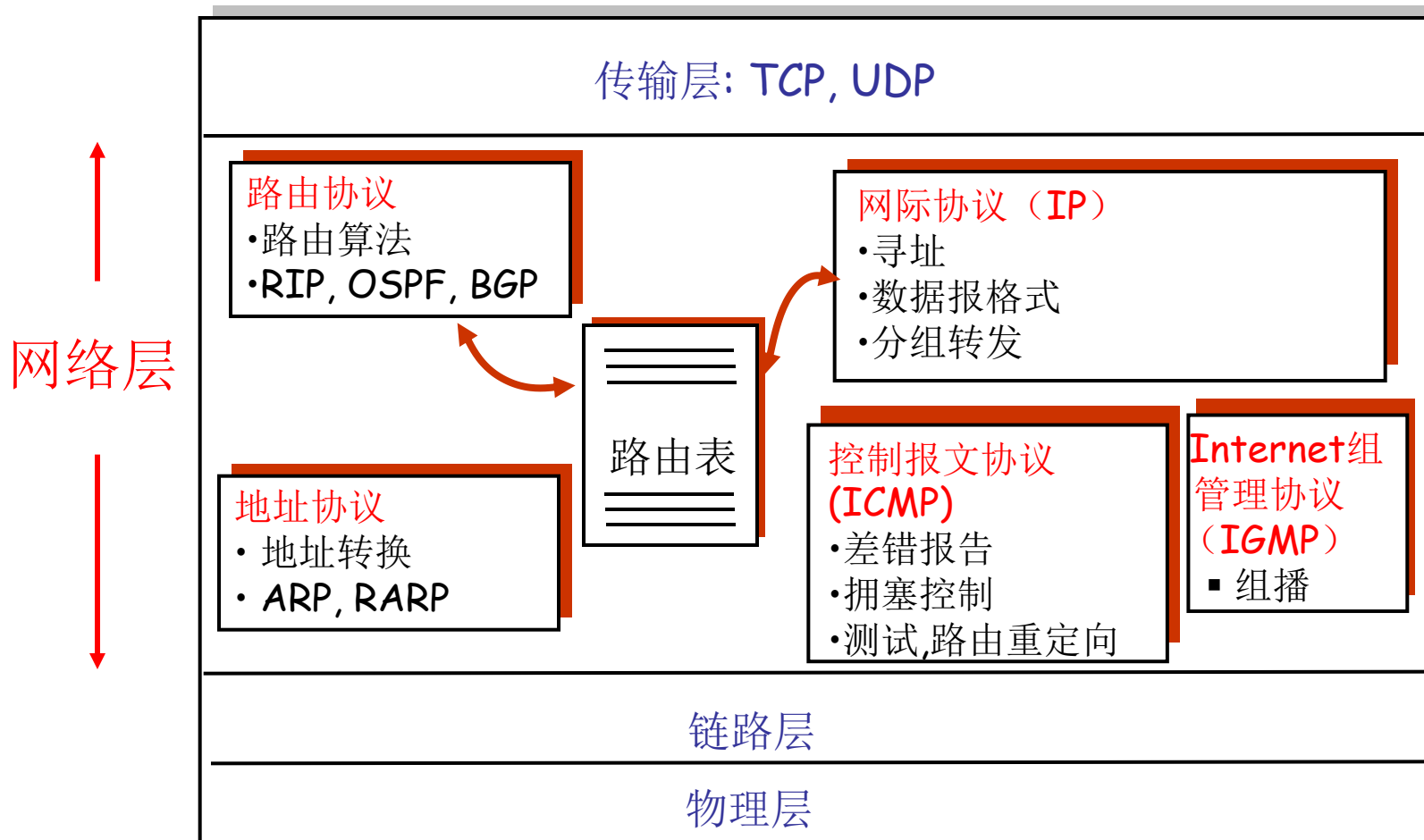


网络层解决的问题



- Network Layer (网络层)
 - 网络服务定义 (VC; DG)
 - 寻址和地址转换 (MAC地址/ IP地址; ARP/RARP/NDP/S)
 - 分组传送 (转发) -- IP
 - 路由 (如何构造路由表; 路由器之间如何交换路由信息; 交换路由信息的实现 (VD/LS; RIP、OSPF、BGP ...))
 - 拥塞控制 --- ICMP
 - 差错控制 --- ICMP
 - 网络互连
(Repeater, HUB, Bridge, switch, router, gateway)
 - 组播 (IGMP)
 - 其它主题 (IP v6/DHCP/Tunneling/MobileIP)

Internet 网络层

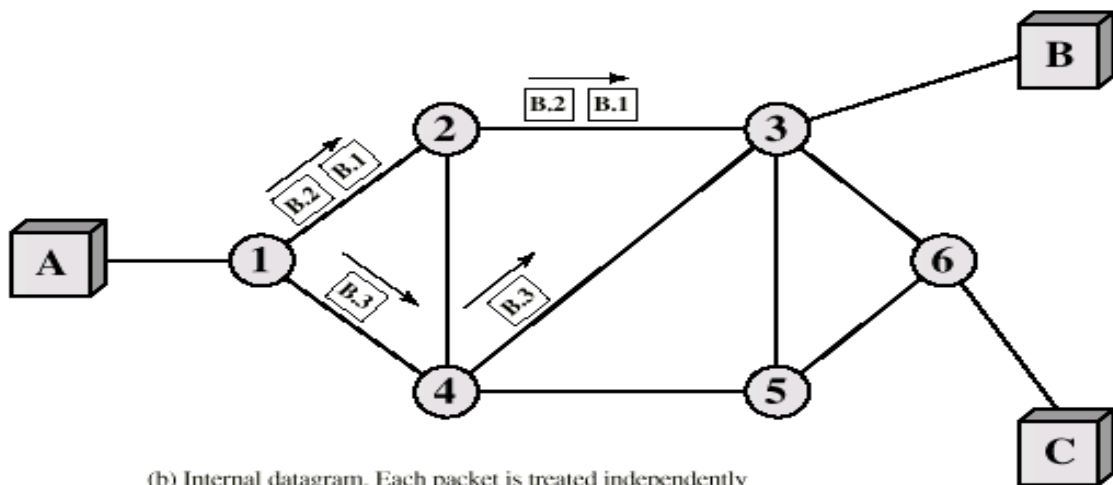
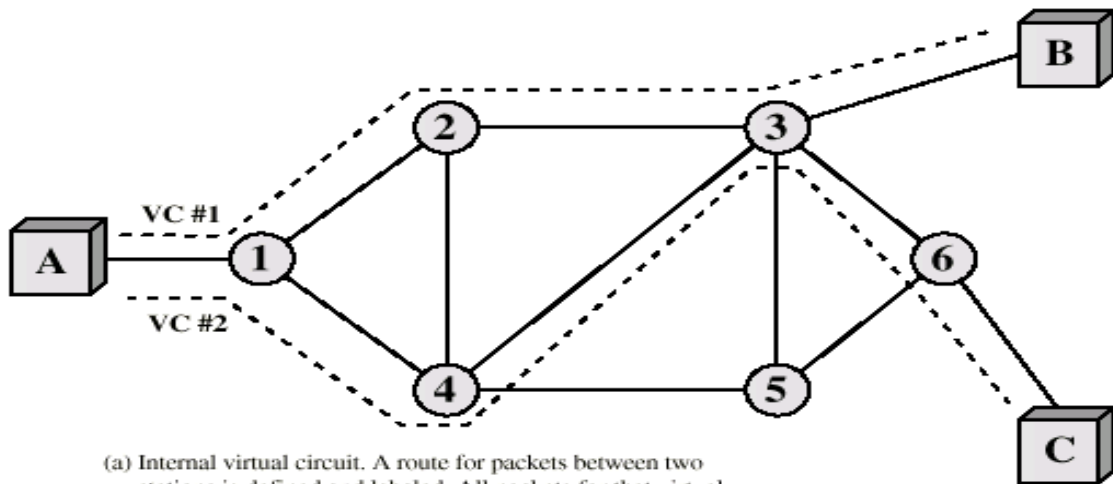


6.1.2 网络层服务

根据 OSI/RM，网络层
向上提供提供两种服务

虚电路
或
数据报？

Internet 只提供数据报服务



两种服务：虚电路和数据报文

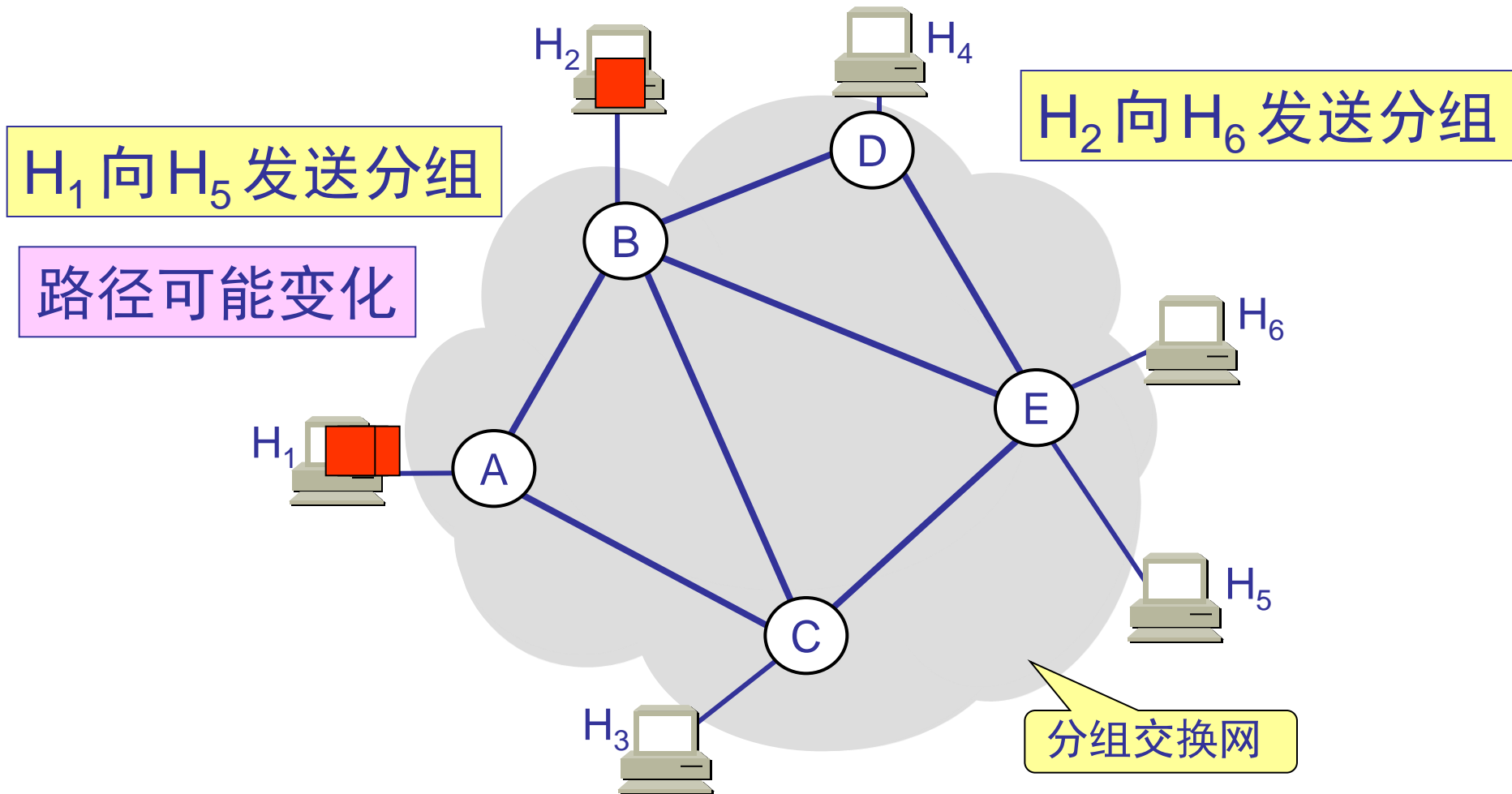
•虚电路（面向连接）

- 在传送数据之前，首先通过呼叫建立一条虚电路
- 所有分组沿同一条路径传送，并且按发出顺序到达
- 类似电路交换
- 建立连接之后，分组中只需要携带较短的连接标识
- 可以在建立连接时协商参数、**QoS**、开销等
- 故障适应能力差，不能平衡网络流量

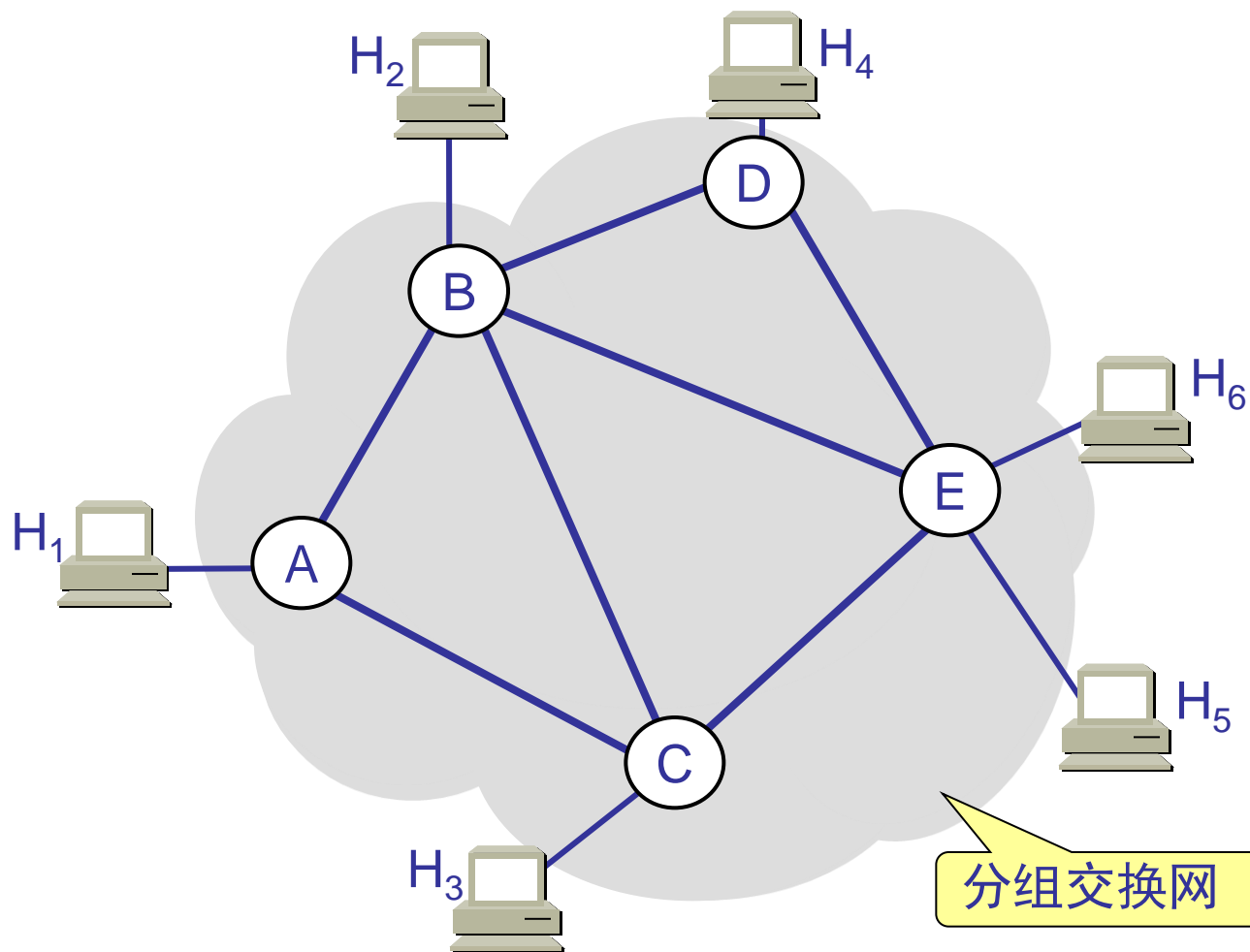
•数据报（无连接）

- 每个分组单独传送
- 网络为每个分组单独选路，路径可能不同
- 分组到达顺序可能与发出顺序可能不同
- 分组中需要携带完整的目的地址（IP地址）
- 故障适应能力强，能平衡网络流量

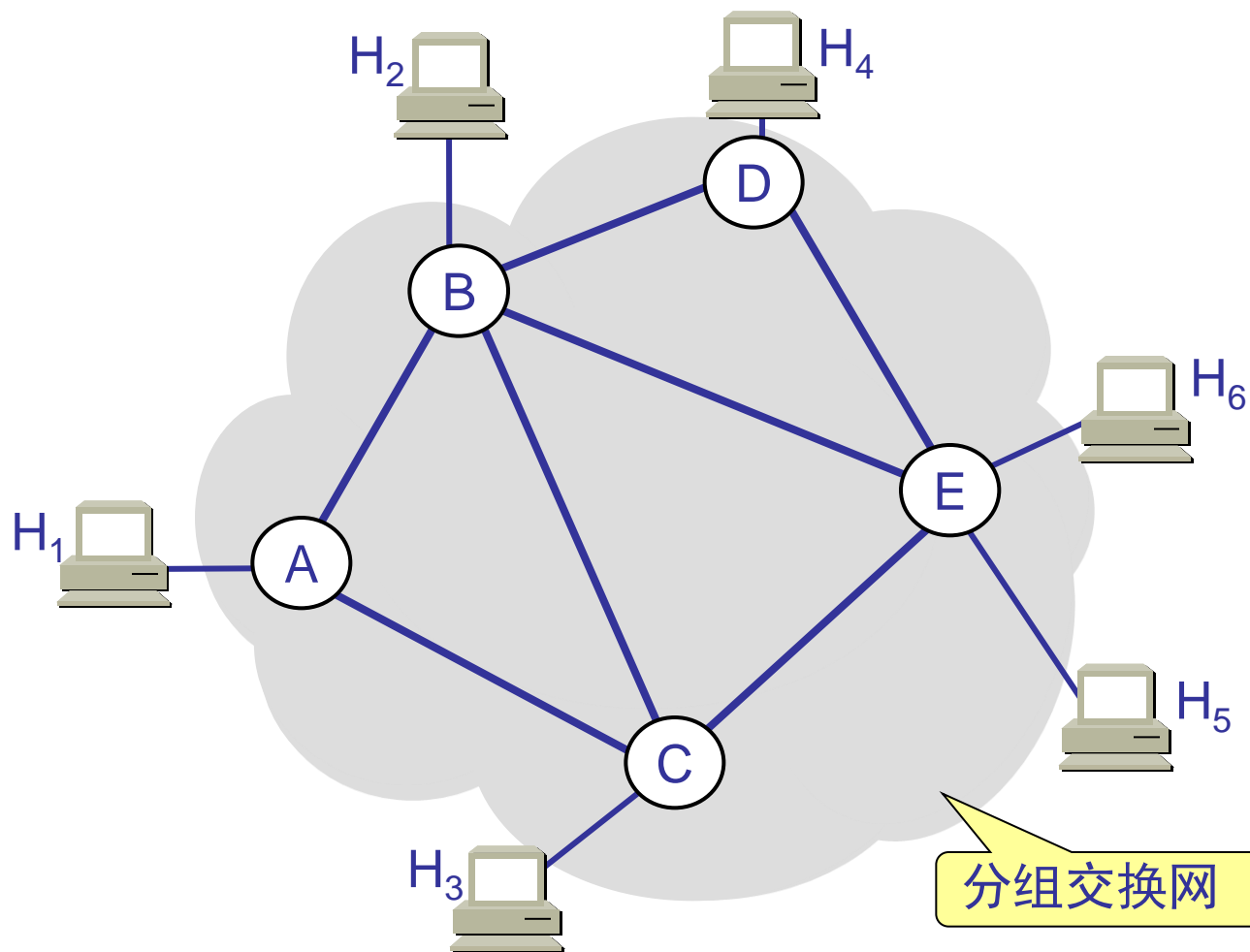
网络随时接受主机发送的分组（即数据报）
网络为每个分组独立地选择路由。



网络尽最大努力地将分组交付给目的主机，
但网络对源主机没有任何承诺。

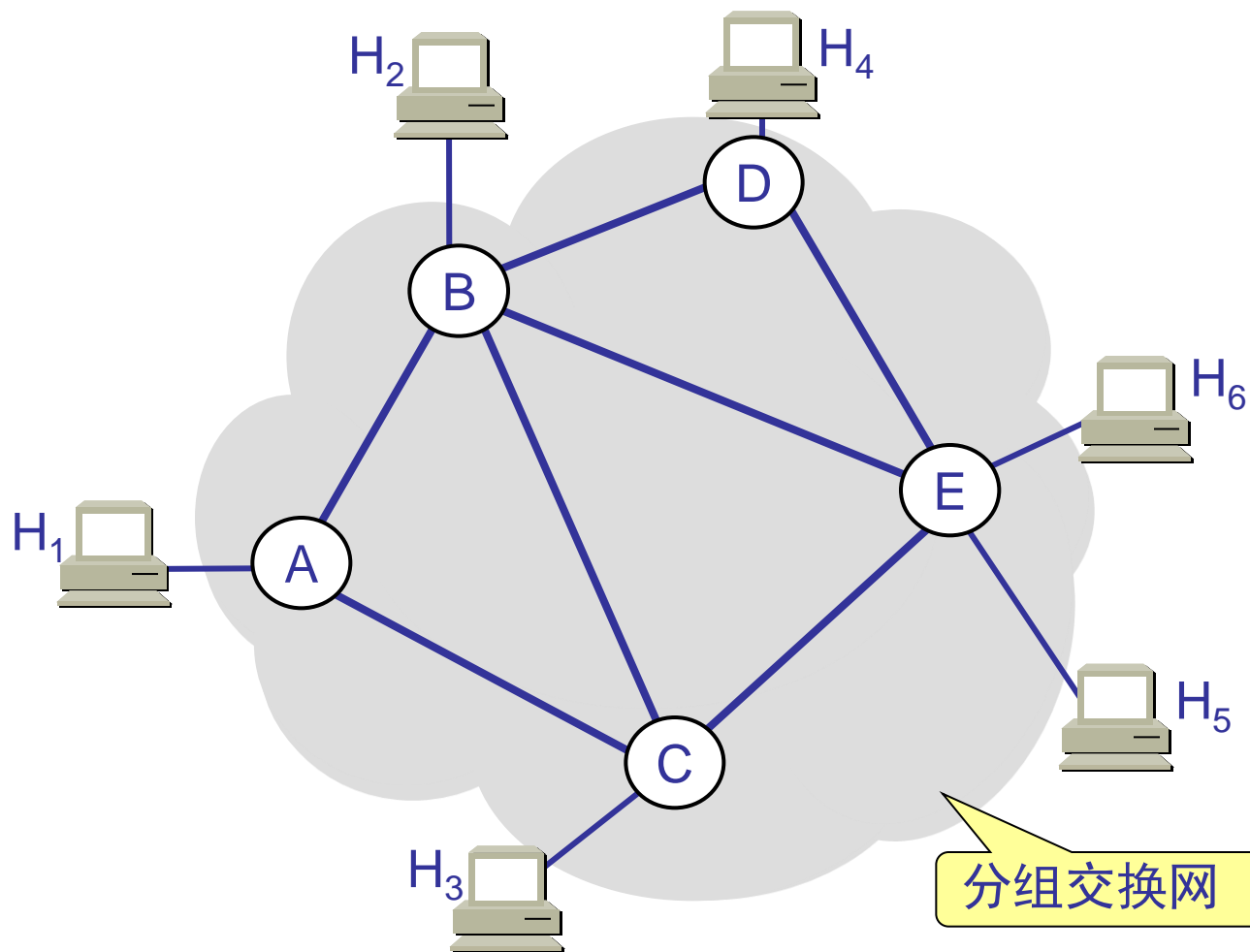


网络不保证所传送的分组不丢失
也不保证按源主机发送分组的先后顺序
以及在时限内必须将分组交付给目的主机

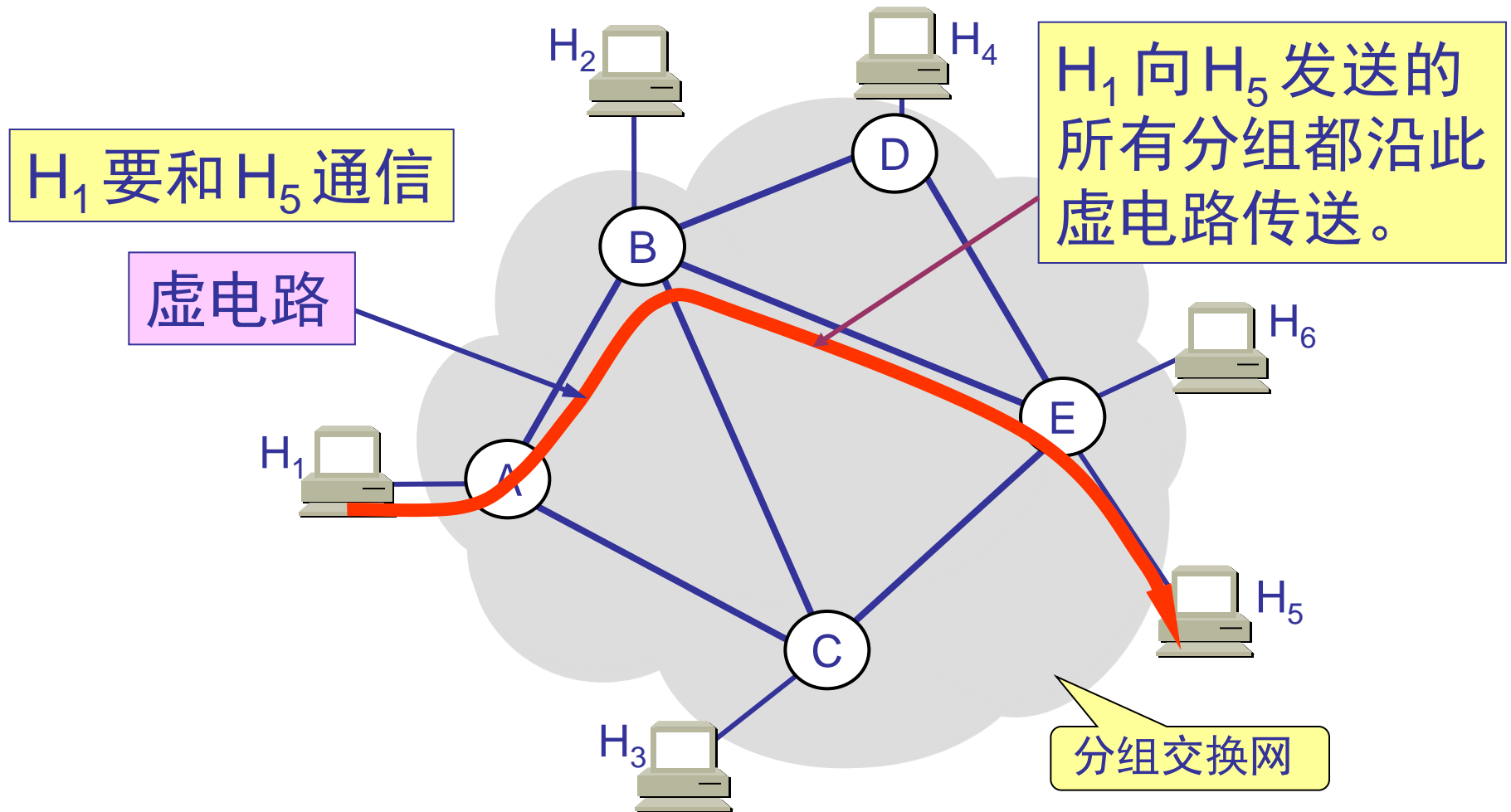


当网络发生拥塞时

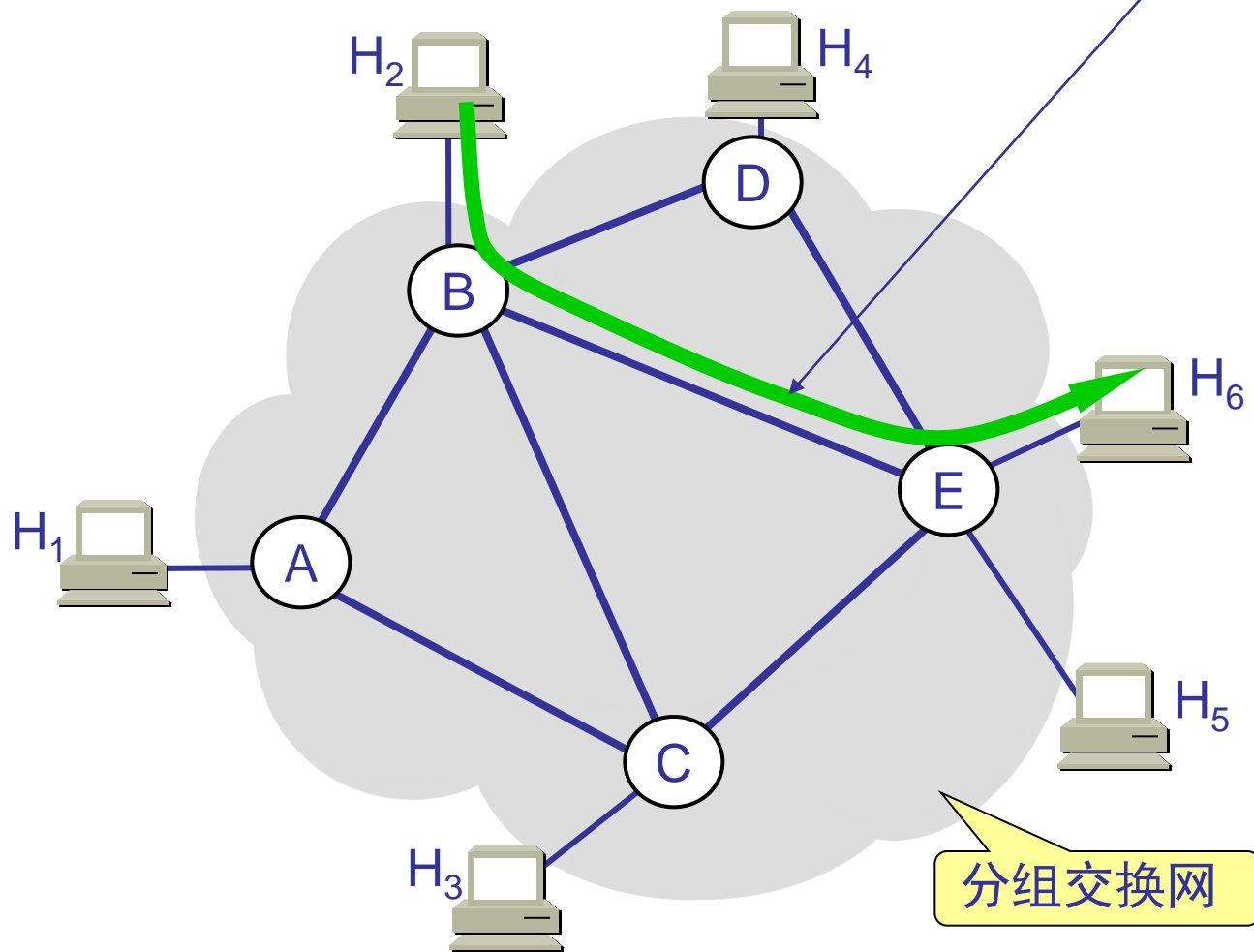
网络中的结点可根据情况将一些分组丢弃



主机 H_1 先向主机 H_5 发出一个特定格式的控制信息分组，要求进行通信，同时寻找一条合适路由。若主机 H_5 同意通信就发回响应，然后双方就建立了虚电路。

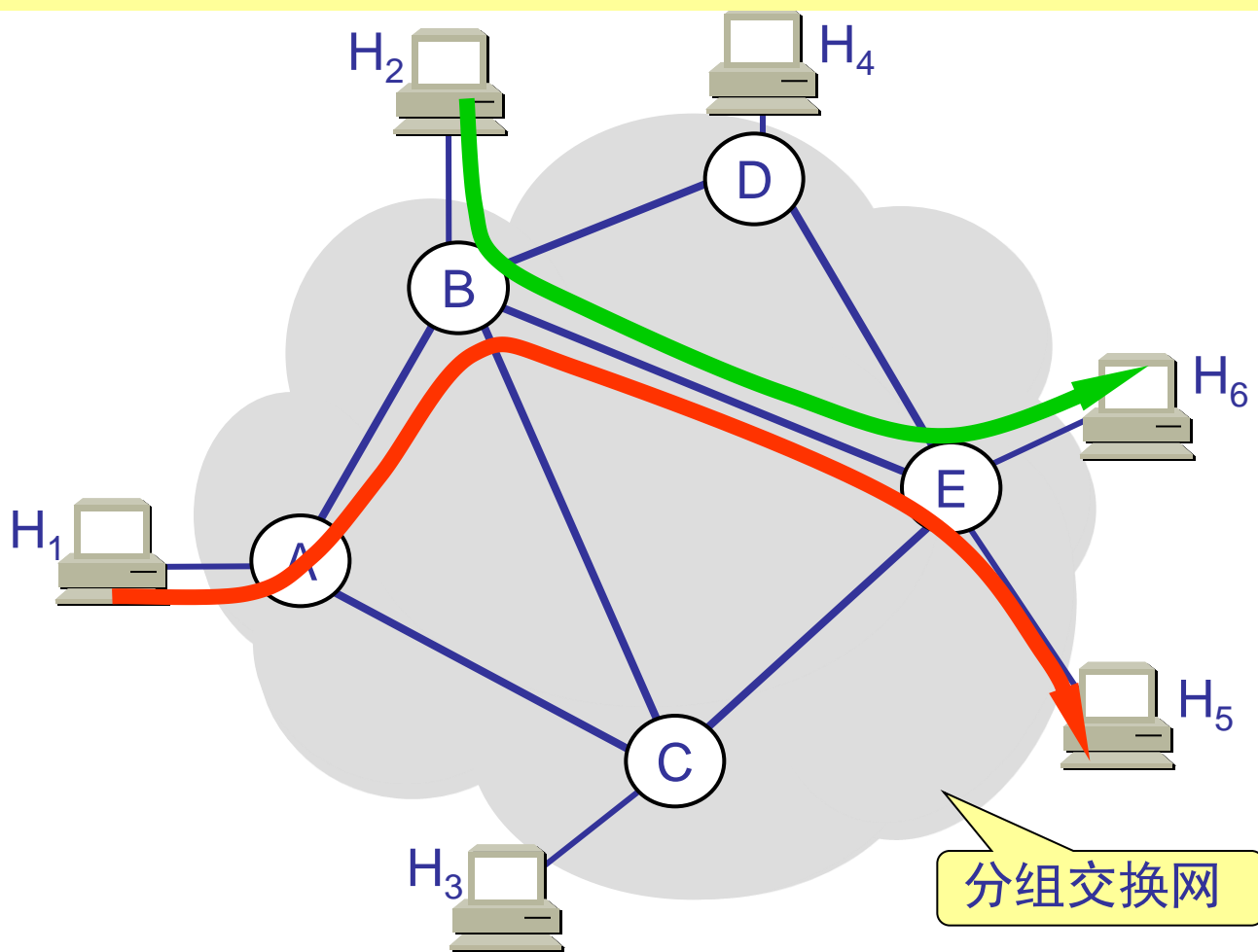


同理，主机 H_2 和主机 H_6 通信之前，也要建立虚电路。

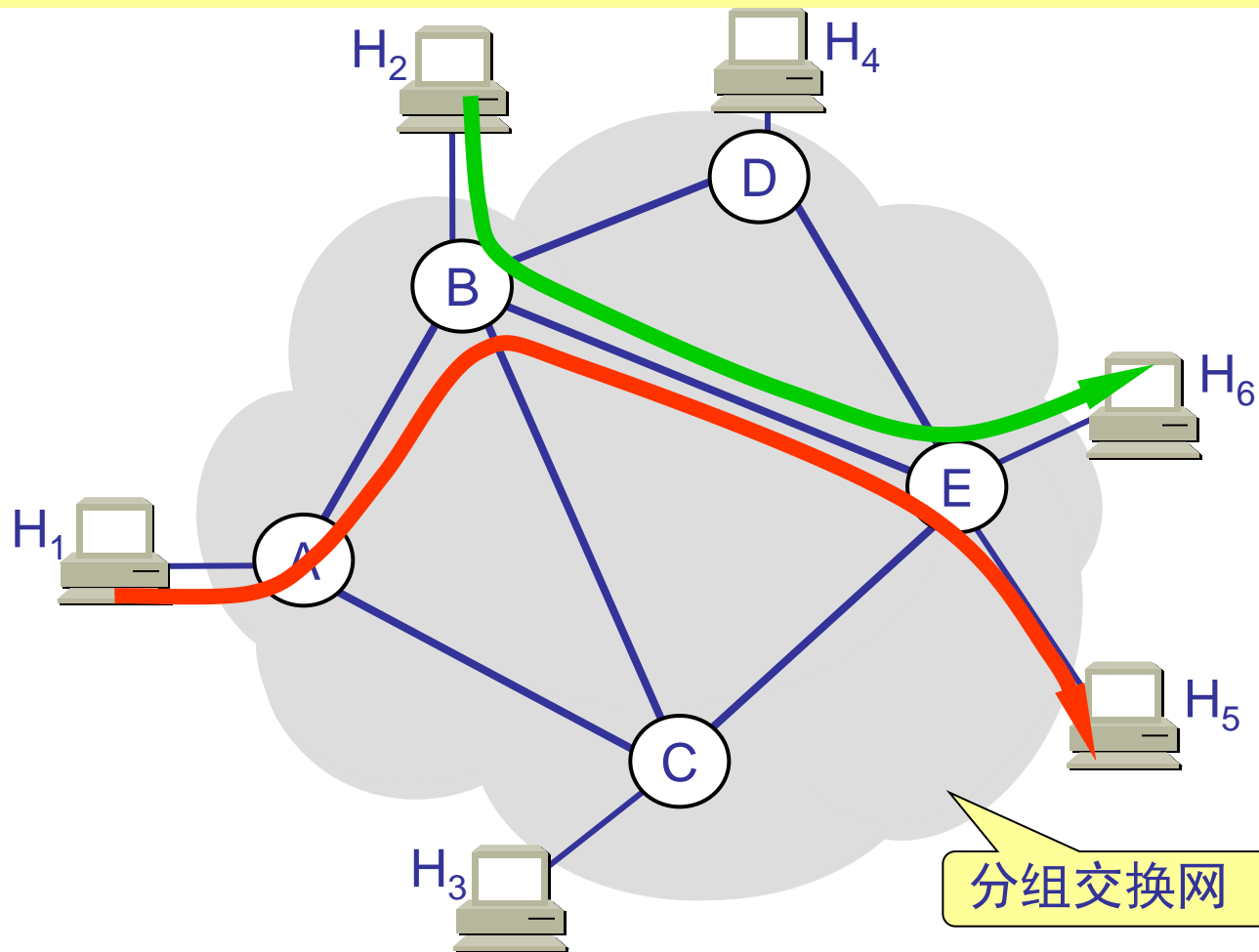


在虚电路建立后，网络向用户提供的服务就好像在两个主机之间建立了一对穿过网络的**数字管道**。

所有发送的分组都按顺序进入管道，然后按照先进先出的原则沿着此管道传送到目的站主机。



到达目的站的分组顺序就与发送时的顺序一致，
因此网络提供虚电路服务对通信的
服务质量 QoS (Quality of Service) 有较好的保证。



虚电路与数据报的比较

对比的方面	数据报	虚电路
是否需要建立连接	不需要	需要
分组中的目的地址	完整地址	vc标识
路由器中的路由表	只需一个很简单的路由表	要为每个虚电路保存一个路由表
选路	每个分组独立选路，路由可能不同	在vc建立时选路，所有分组路由相同
路由器故障的影响	几乎不受影响	所有经过该路由器的vc都将终止
拥塞控制	实现困难	易于实现
差错控制和流量控制	由主机负责	由子网负责

拥塞控制

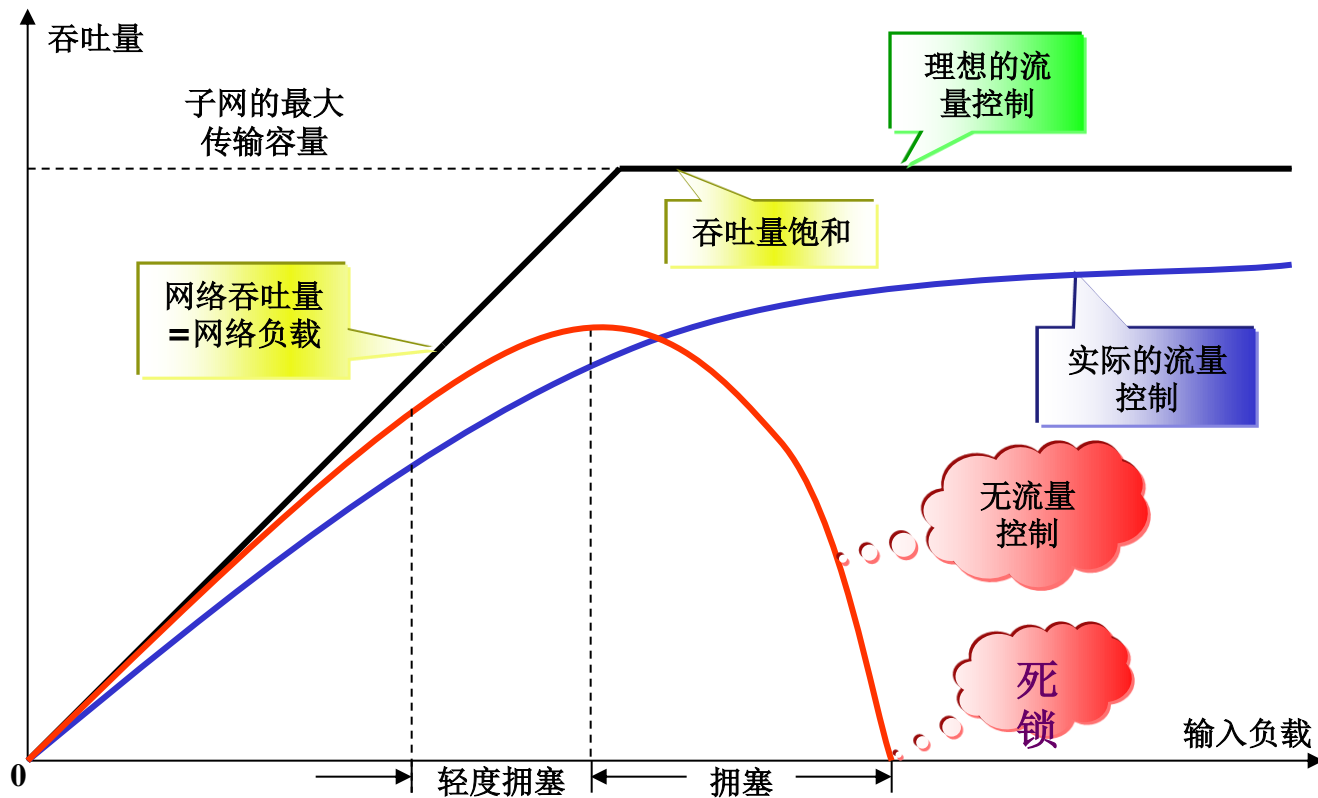
拥塞控制

- 在某段时间，若对网络中某资源的需求超过了该资源所能提供的可用部分，网络的性能就要变坏——产生**拥塞**(congestion)。
- 出现资源拥塞的条件：

对资源需求的总和 $>$ 可用资源

- 若网络中有许多资源同时产生拥塞，网络的性能就要明显变坏，整个网络的吞吐量将随输入负荷的增大而下降。

拥塞控制



当网络负载继续增大到某一数值时, 网络的吞吐量就下降为零, 网络已无法工作. 这就是所谓的“死锁”

图 拥塞控制所起的作用

拥塞控制

死锁主要有两种:一种是直接死锁,另一种重装死锁。

(1)直接死锁: 即由互相占用了对方需要的资源而造成的死锁。

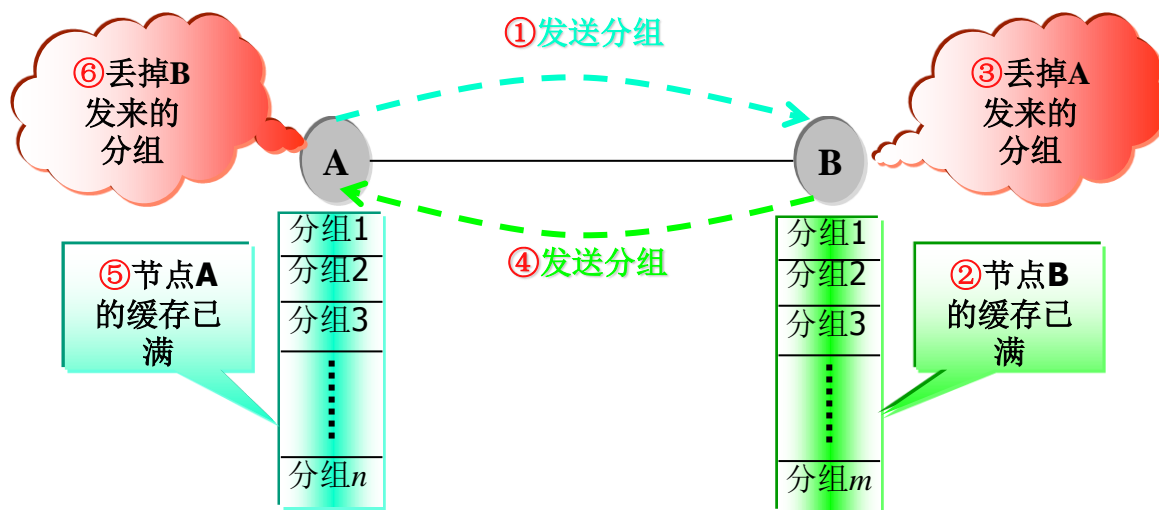


图 直接死锁实例

节点A 保留发给节点B 的分组的副本,等待节点B 的应答. 同样, 节点B 也保留已发送给节点A 的分组的副本, 等待节点A 的应答. 谁也无法成功地发出一个分组.

拥塞控制

(2) 重装死锁: 由于路由器缓存的拥塞而引起的死锁。

设三个报文A, B和C经过路由器P, Q和R发往主机H. 每一个报文由4个分组构成. 又设每个路由器的缓存只能容纳4个分组.

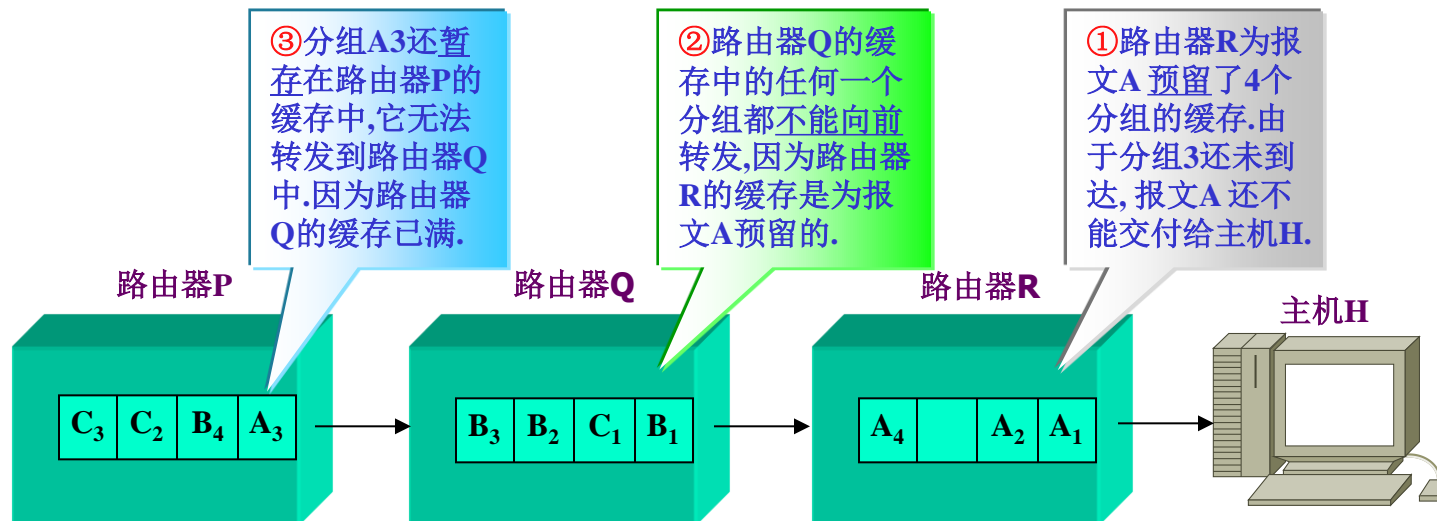


图 重装死锁的例

拥塞控制与流量控制

流量控制与拥塞控制的关系是:

- (1) 拥塞控制是一个全局性的过程,涉及的路由器,以及与降低传输性能有关的应用的前提是网络能够承受现有的网络负荷
- (2) 流量控制是一个局部性的过程,往往在发送端和接收端之间的点到点通信量.流量控制使发送端发送数据的速率不能使接收端来不及接收.流量控制总是存在着从接收端到发送端的某种反馈.接收端知道接收端是处于怎样的状况.

在实施拥塞控制时,拥塞控制算法向发送端发送控制报文,并告诉发送端网络已出现拥塞,必须放慢发送速率.这和流量控制“很像”.

拥塞控制

§ 6.1.2 拥塞控制的一般原理

拥塞控制算法大致可分成开环控制和闭环控制两大类.

(1) 开环控制算法原理

通过良好的网络系统设计来避免拥塞问题的发生,在网络运行过程中,何时接受新分组,何时丢弃分组以及丢弃哪些分组都是事先规划好的,并不考虑当前的网络流量状况.

(2) 闭环控制算法原理

通过反馈机制来调整当前网络流量,使网络流量与网络可用资源相协调,从而使网络拥塞问题得到解决.

拥塞控制

由于闭环控制算法能够根据当前网络状况对流量进行动态控制,具有较高的效率,因此,现代网络系统大都采用闭环控制算法来解决网络拥塞问题.

闭环控制算法中的关键技术是:

①检索技术: 检索机制能够随时发现拥塞问题,判断的依据和参数主要有:因缺少缓冲区空间而丢弃的分组数量;平均分组队列长度;超时重发分组的数量;平均分组延迟时间等. 如果基准参数超过临界值,则意味着可能发生了网络拥塞.

②反馈技术: 反馈机制将发生拥塞的信息从检查点传送到控制点.反馈方式主要有两种:显式反馈和隐式反馈.显式

拥塞控制

反馈采用由检查点向控制点反馈一个警告分组的方式来通告网络已发生了拥塞.隐式反馈采用由控制点(源端)通过观察应答分组返回所用时间进行推断的方式来判断网络是否发生了拥塞.

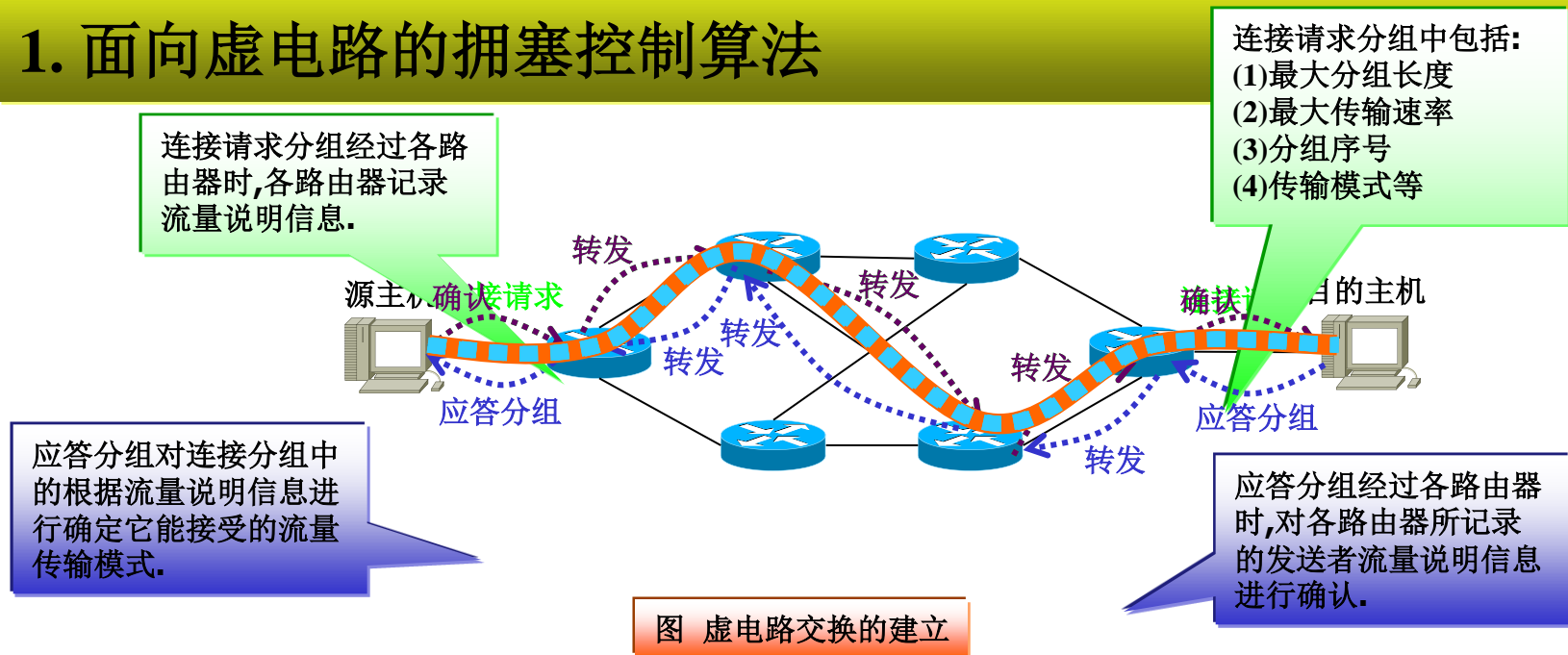
③调整技术: 调整机制是通过检查点和控制点相互协作来共同努力解决拥塞问题.控制点通过降低载荷,即降低分组发送速率来缓解拥塞;检查点通过载荷脱落,即丢弃一些分组来疏导交通,或者通过增加系统资源(如增加附加链路或提高链路带宽)来提高流量通过能力,对于后者,要求系统能够提供后备的资源.

拥塞控制

§ 6.1.3 拥塞控制算法

根据数据交换方式(虚电路或数据报)的不同,拥塞控制算法采用不同的策略.

1. 面向虚电路的拥塞控制算法



拥塞控制

在数据传输过程中,路由器将根据协商好的传输模式对该链路上的交通进行整形.流量整形(Traffic Shaping)是调整数据传输的平均速率,使数据按照传输模式规定的速率进行传输,尽量避免突发性增大通信量,导致产生拥塞问题.交通整形是对传输速率进行调整,而不是调整流量控制中一次可传送的数据量.

流量整形主要采用两种算法:漏桶(Leaky Bucket)算法和令牌桶(Token Bucket)算法.

(1) 漏桶算法

拥塞控制

漏桶算法相当于在路由器内部实现了一个有限长度队列，路由器将以恒定速率从队列中取出分组发送出去，而进入路由器的分组被排到队列的尾部，一旦队列饱和，新来的分组将被丢弃。这种算法实际上是一种具有恒定服务时间的单服务器排队系统。主机系统也可采用该算法来整形分组的发送，将上层应用进程不均匀的数据流整形成均匀的分组流向网络发送，平滑了突发的数据流，大大减少了发生拥塞的机会。

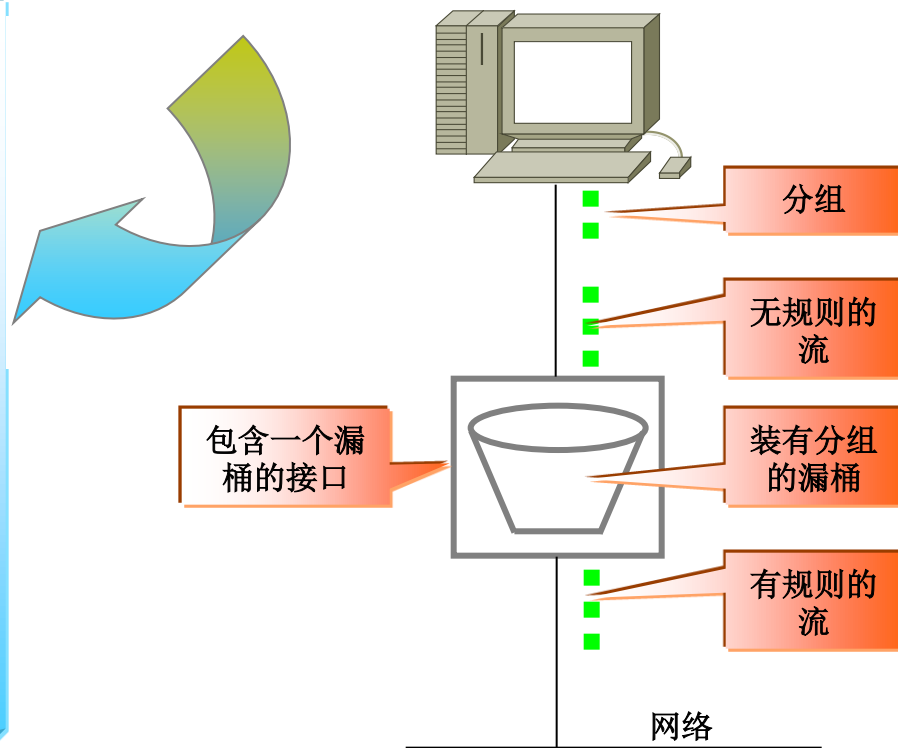


图 一个装水漏桶

图 一个分组的漏桶模型

拥塞控制

(2)令牌桶算法

令牌桶算法与恒定输出速率的漏桶算法有所不同,它允许一定量的突发数据流.该算法以恒定速率产生一个个令牌放入桶内,每发送一个分组都要获得和消耗一个令牌,如果令牌消耗完,则新来的分组就要等待生成新令牌或被丢弃.由于突发性的输入流往往导致拥塞的发生,因此获得令牌的分组将被快速地输出,使突发性的输入流得到迅速的疏导.

在令牌桶算法中,使用一个令牌计数器来计数令牌数量.令牌计数器每隔 Δt 时间加1,表示新增加一个令牌;每发送一个分组,令牌计数器减1,表示已消耗一个令牌.当计数

拥塞控制

器减至0时,表示令牌已消耗完,不能再发送分组了.

这两种交通整形算法既可用于平滑主机向网络发送分组的通信量,也可用于两个路由器之间的交通整形. 在ATM交换机和RSVP协议中,就是应用上述的交通整形算法来解决网络拥塞问题的.

2.面向数据报的拥塞控制算法

数据报服务是一种无连接传输方式。路由器一旦检测到系统可用资源(如链路利用率或队列长度)超过临界值,就会向源主机发送一个抑制分组,警告网络可能发生拥塞. 源端主机定期地侦听抑制分组,如果在侦听期内收到抑制分组,则会减少发送给目的主机的数据量.当减至在侦听期

拥塞控制

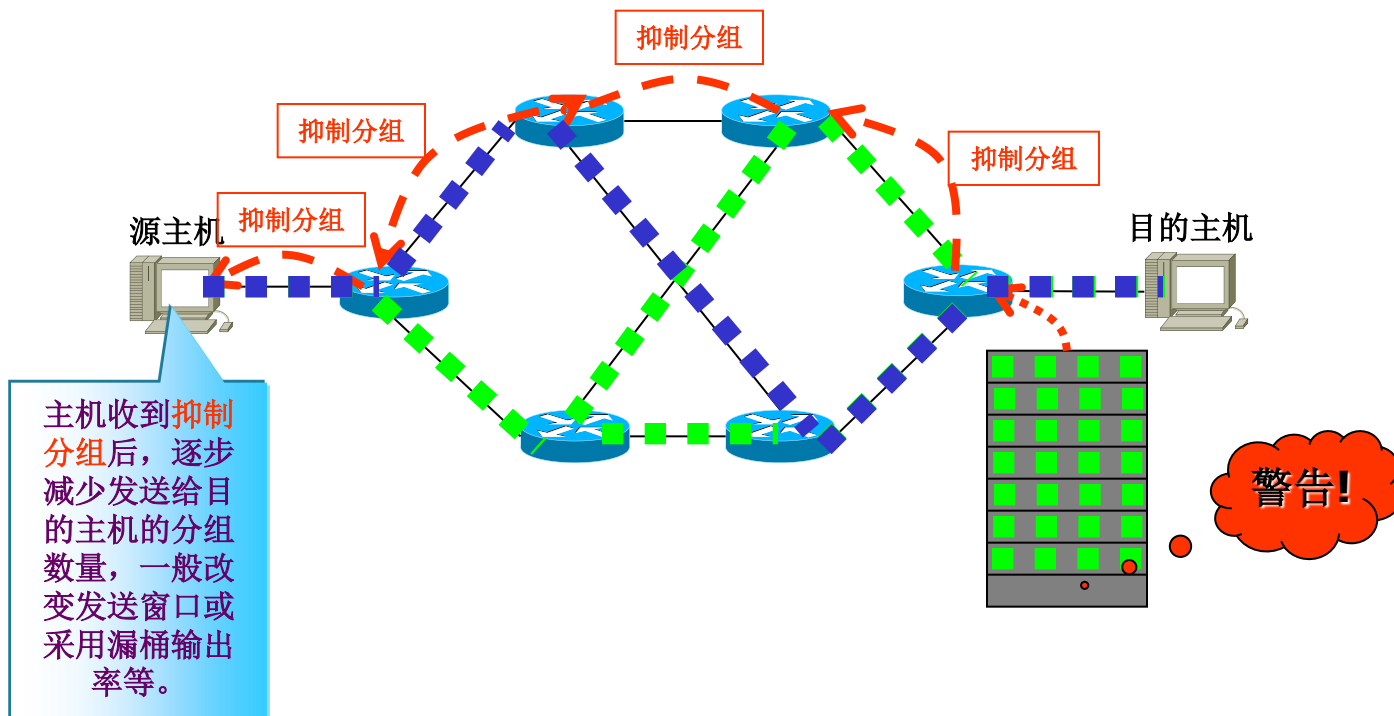


图 基于数据报服务的拥塞控制策略

拥塞控制

在侦听期内不再收到抑制分组后,可以逐渐增加通信量.主机可以通过调整其发送操作的相关参数(发送窗口或漏桶输出率)来减少通信量,路由器通常采用加权公平队列算法来处理分组排队,检测是否超过临界值,以及何时发送抑制分组.

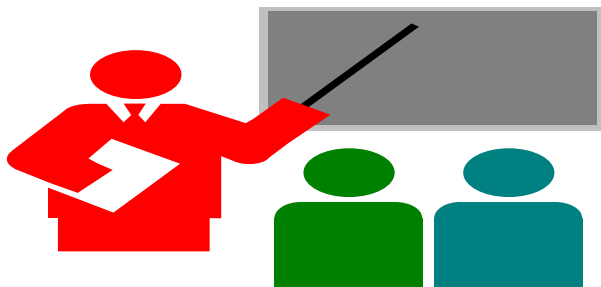
在广域网环境下,采用完全由远程的源端主机减少通信量来缓解拥塞的方法并不是很有效的,因为距离远,反应慢.一种改进的方法是抑制分组途径的各个路由器都要抑制通信量,使通信量得到迅速控制,但中间路由器要提供一定数量的缓冲区空间用于暂存过量的分组.当抑制分组到达源端主机后,由源端主机采取适当措施最终才能使通信

拥塞控制

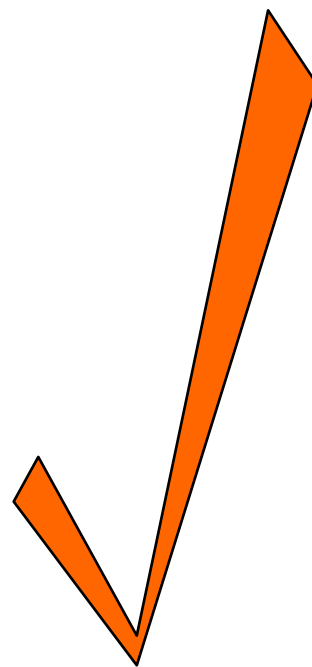
量降下来.在IP协议中,采用了抑制分组的方法来解决网络拥塞问题.

当网络发生严重拥塞时,路由器往往需要采用载荷脱落(Load Shedding),即丢弃分组的方法来疏导交通.路由器一般采用按先来先服务的规则丢弃后到的分组,或按分组优先级的规则丢弃优先级低的分组等.被丢弃的分组要重传. 按分组优先级的规则丢弃分组的方法被很多网络系统或协议所采用,如ATM网络,帧中继网络以及IP协议等.在这些网络系统中,分组(或帧)中均设有一个优先级字段,说明在发生拥塞时,路由器将根据分组中所标记的优先级来选择要丢弃有分组.

Summary



- 网络层概述
- 虚电路
- 数据报
- **Internet网络层**
- 拥塞控制(自学)



Thanks!

