哈爾濱工業大學
HARBIN INSTITUTE OF TECHNOLOGY
立足航天，服务国防，面向国民经济主战场

第5周 课堂教学-传输层（下）

李全龙

第5周 课堂教学-传输层（下）

- 束广就狭：
  - 3.5 TCP段结构
  - 3.6 TCP连接管理
  - 3.7 TCP的可靠数据传输
  - 3.8 TCP拥塞控制

第5周 课堂教学-传输层（下）

❖ 质疑辨惑：

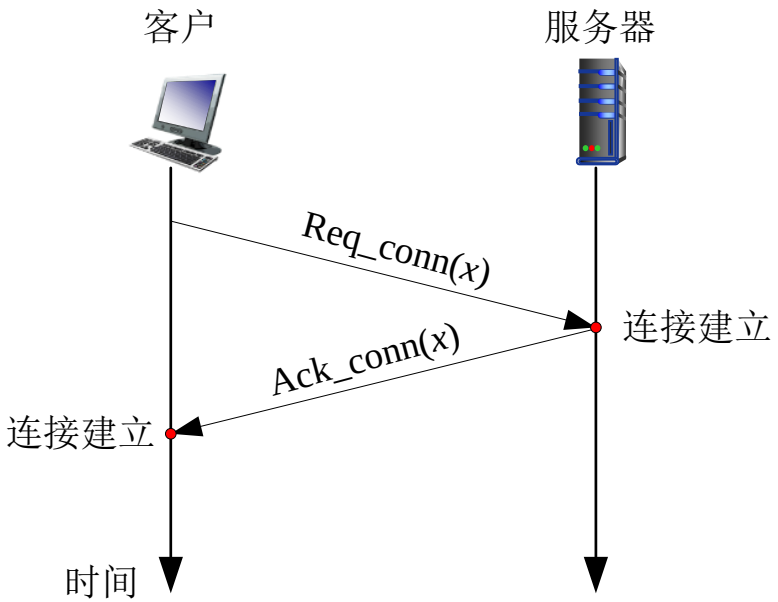
1.TCP如何进行流量控制？可能存在什么问题？

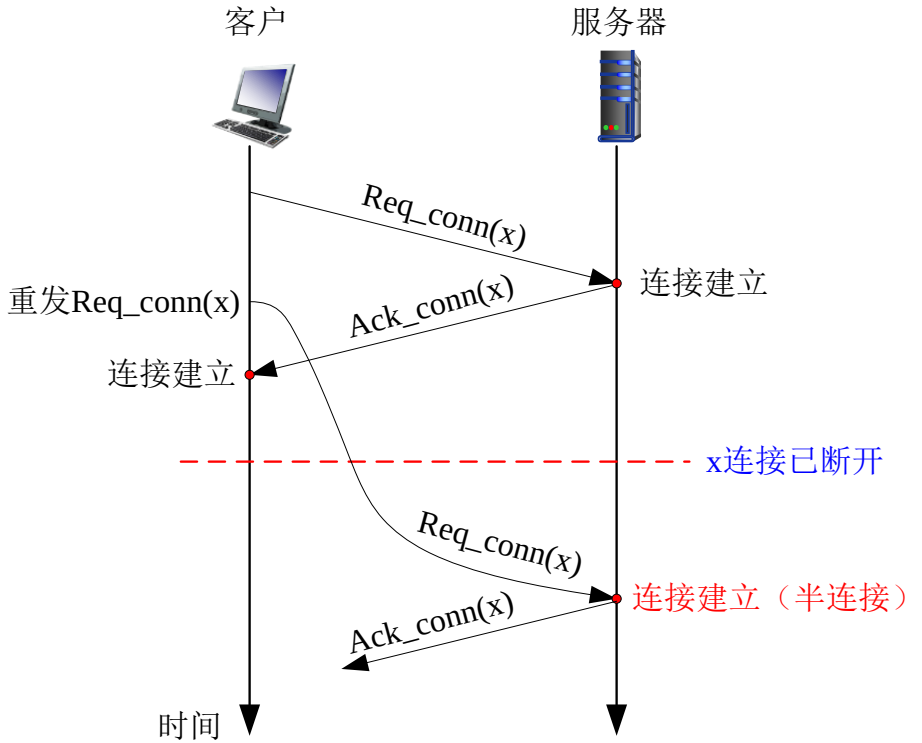
- 糊涂窗口综合症(Silly Window Syndrome)
  - 发送端
    - 解决方法：
      - » **Nagle算法：报文段一定长度再发送**
  - 接收端
    - 解决方法：
      - » **Clark解决方法：0窗口确认**
      - » 延迟确认：阻止发送窗口滑动

质疑辨惑：

2.TCP协议为什么要采用三次握手建立连接？四次挥手断开连接？连接建立与断开过程的序号如何变化？

- 二次握手如何？

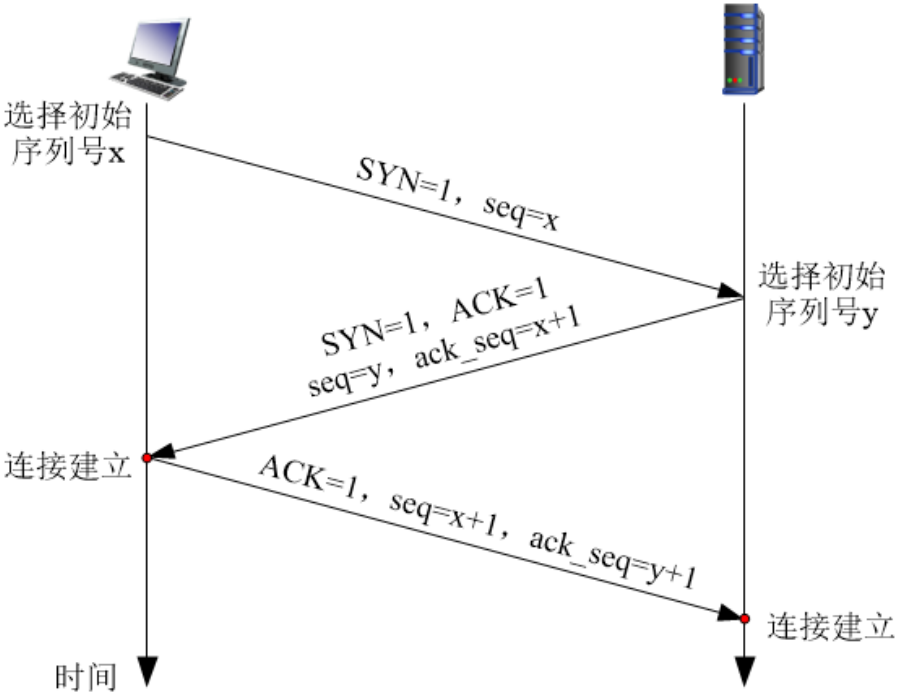


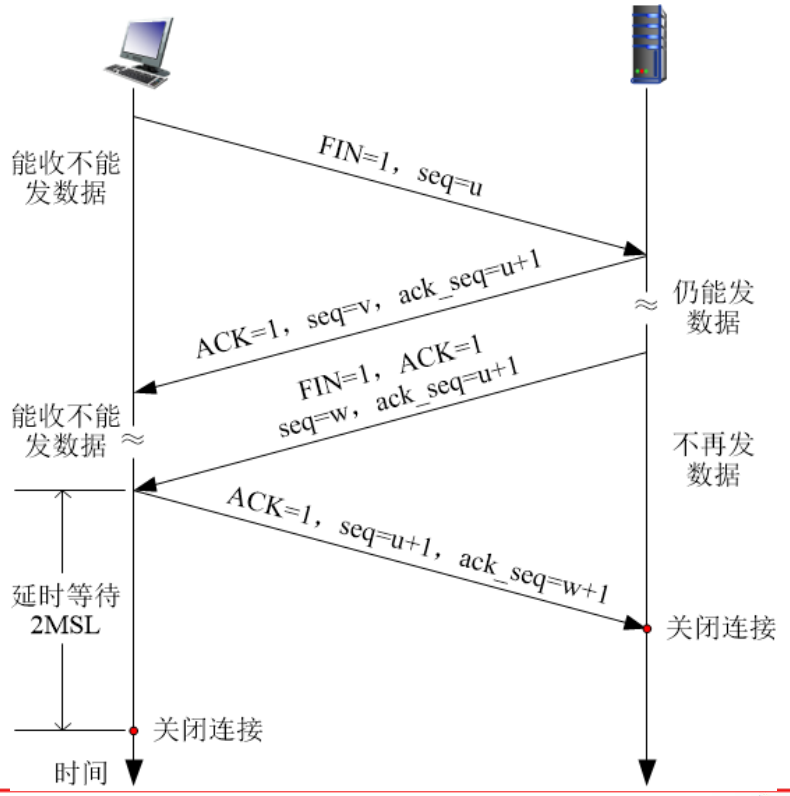


质疑辨惑：

2.TCP协议为什么要采用三次握手建立连接？四次挥手断开连接？连接建立与断开过程的序号如何变化？

- 建连、断连过程的序号变化如何？





第5周 课堂教学-传输层（下）

- 质疑辨惑：

3.网络拥塞控制可以有哪些策略？

- 网络层拥塞控制
  - 流量感知路由
  - 准入控制
  - 流量调节
    - 抑制分组
    - 背压
  - 负载脱落
  - ……

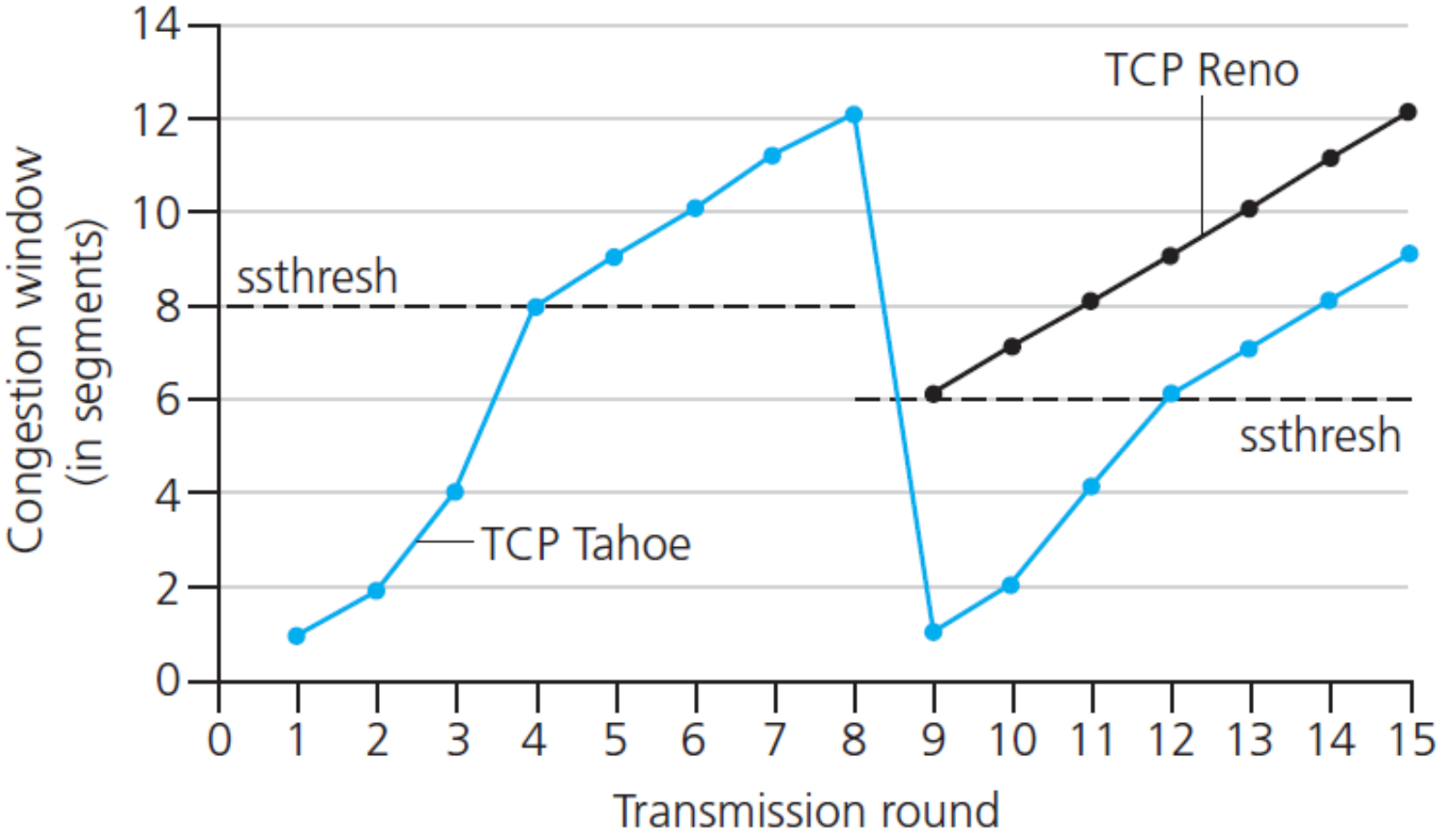
4.如何改进TCP拥塞控制过程以便提高TCP协议吞吐量？

作答

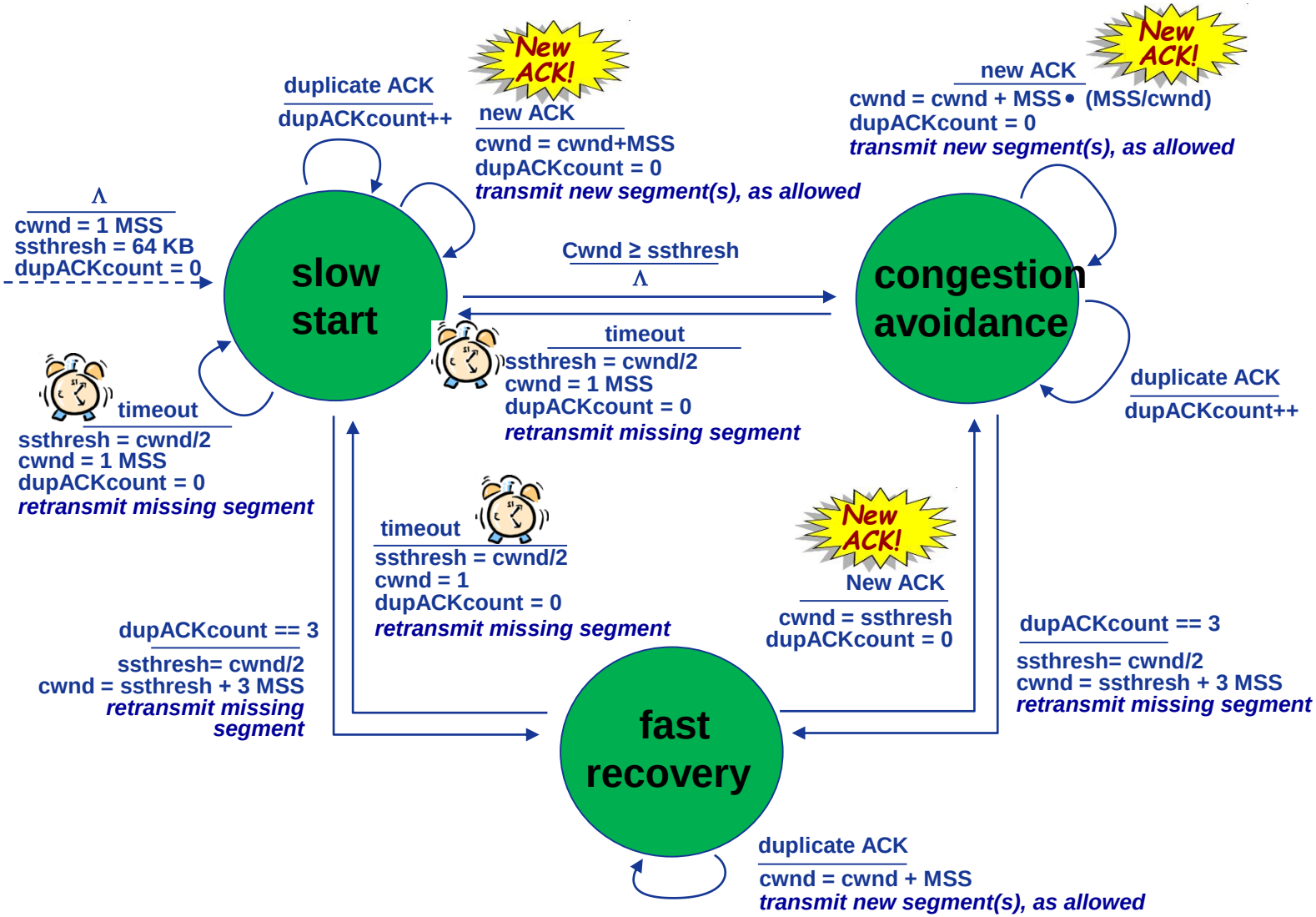
第5周 课堂教学-传输层（下）

- 开疆拓土：
  - TCP拥塞控制的改进

TCP拥塞控制

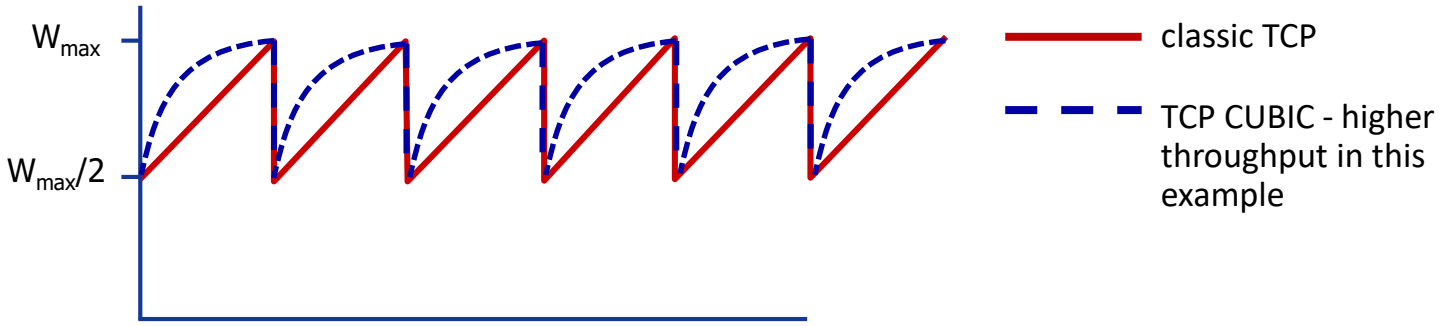

Congestion window
(in segments)
14
12
10
8
6
4
2
0
ssthresh
TCP Reno
ssthresh
TCP Tahoe
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Transmission round

Summary: TCP Congestion Control



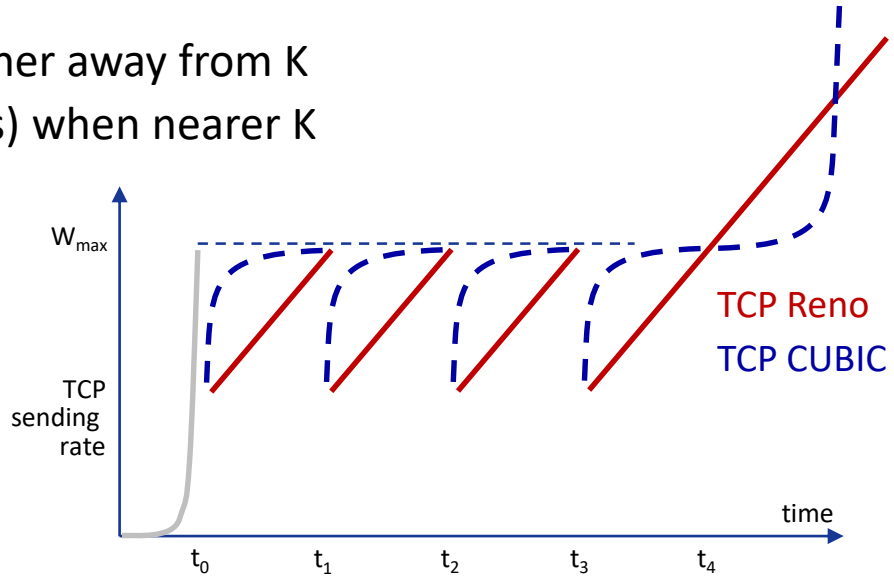
TCP CUBIC

- Is there a better way than AIMD to "probe" for usable bandwidth?
- Insight/intuition:
  - W_{max}: sending rate at which congestion loss was detected
  - congestion state of bottleneck link probably (?) hasn't changed much
  - after cutting rate/window in half on loss, initially ramp to to W_{max} *faster*, but then approach W_{max} more *slowly*



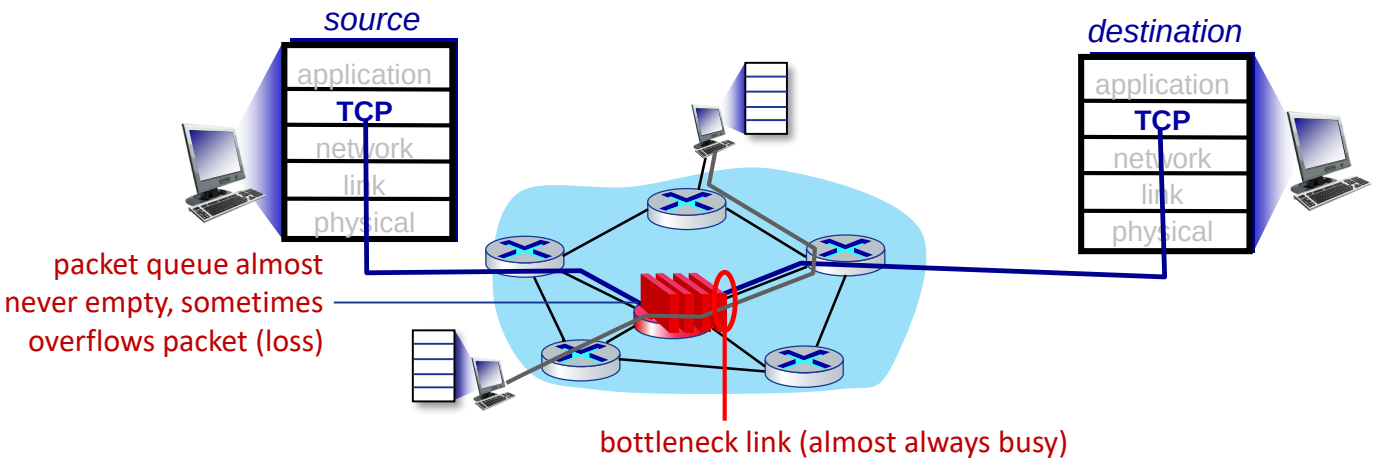
TCP CUBIC

- K: point in time when TCP window size will reach W_{max}
  - K itself is tuneable
- increase W as a function of the *cube* of the distance between current time and K
  - larger increases when further away from K
  - smaller increases (cautious) when nearer K
- TCP CUBIC default in Linux, most popular TCP for popular Web servers



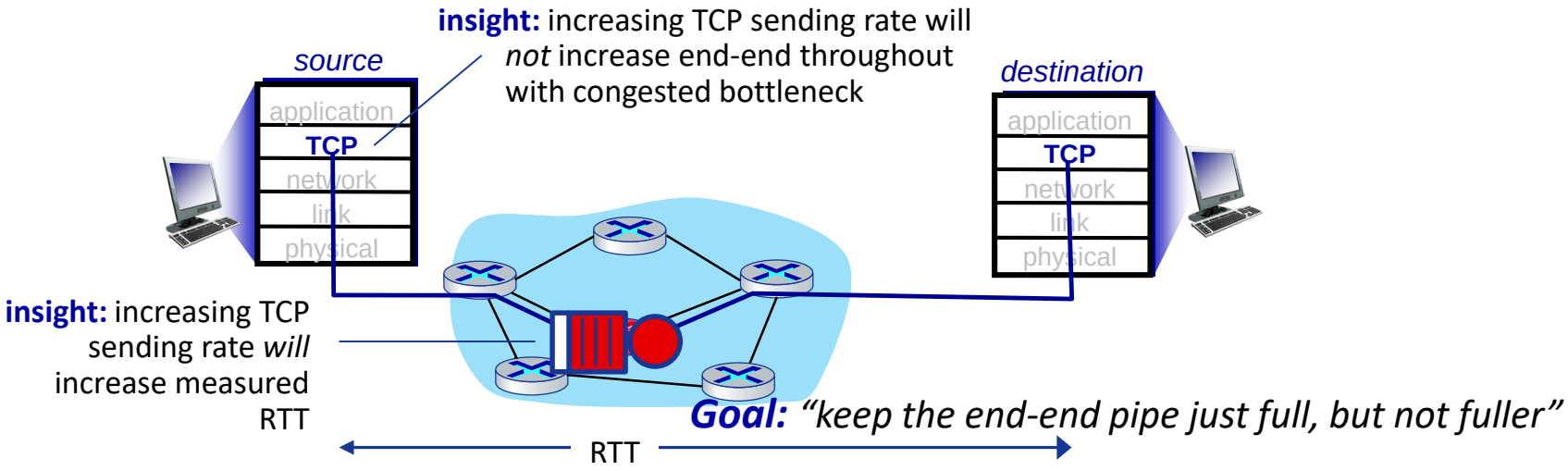
TCP and the congested "bottleneck link"

- TCP (classic, CUBIC) increase TCP's sending rate until packet loss occurs at some router's output: the *bottleneck link*



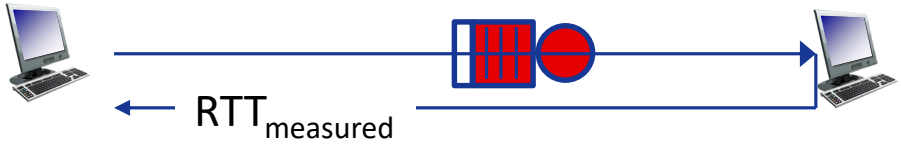
TCP and the congested "bottleneck link"

- TCP (classic, CUBIC) increase TCP's sending rate until packet loss occurs at some router's output: the *bottleneck link*
- understanding congestion: useful to focus on congested bottleneck link



Delay-based TCP congestion control

Keeping sender-to-receiver pipe "just full enough, but no fuller": keep bottleneck link busy transmitting, but avoid high delays/buffering



$$\text{measured throughput} = \frac{\text{\# bytes sent in last RTT interval}}{RTT_{measured}}$$

Delay-based approach:

- RTT_{min} - minimum observed RTT (uncongested path)
- uncongested throughput with congestion window `cwnd` is cwnd/RTT_{min}

```
if measured throughput "very close" to  uncongested throughput
    increase cwnd linearly          /* since path not congested */
else if measured throughput "far below" uncongested throughout
    decrease cwnd  linearly         /* since path is congested */
```

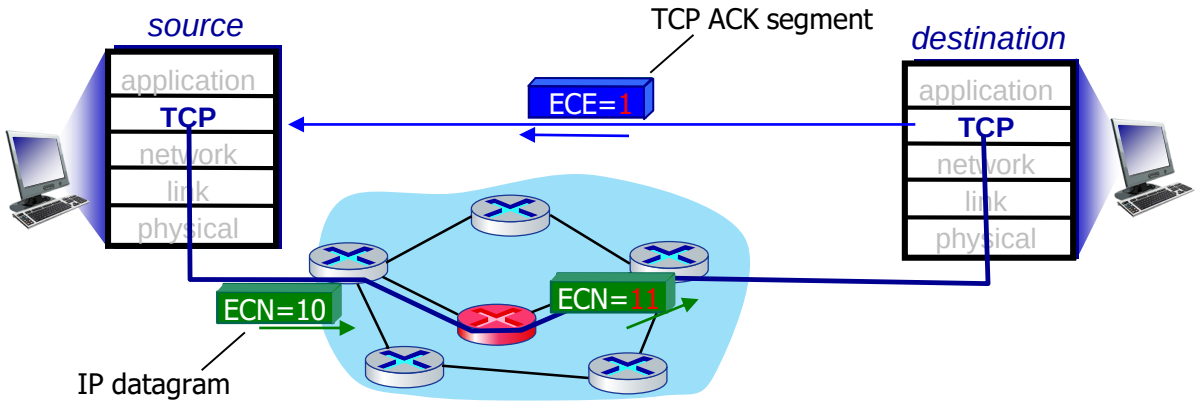
Delay-based TCP congestion control

- congestion control without inducing/forcing loss
- maximizing throughout (“keeping the just pipe full… ”) while keeping delay low (“…but not fuller”)
- a number of deployed TCPs take a delay-based approach
  - BBR deployed on Google’s (internal) backbone network

Explicit congestion notification (ECN)

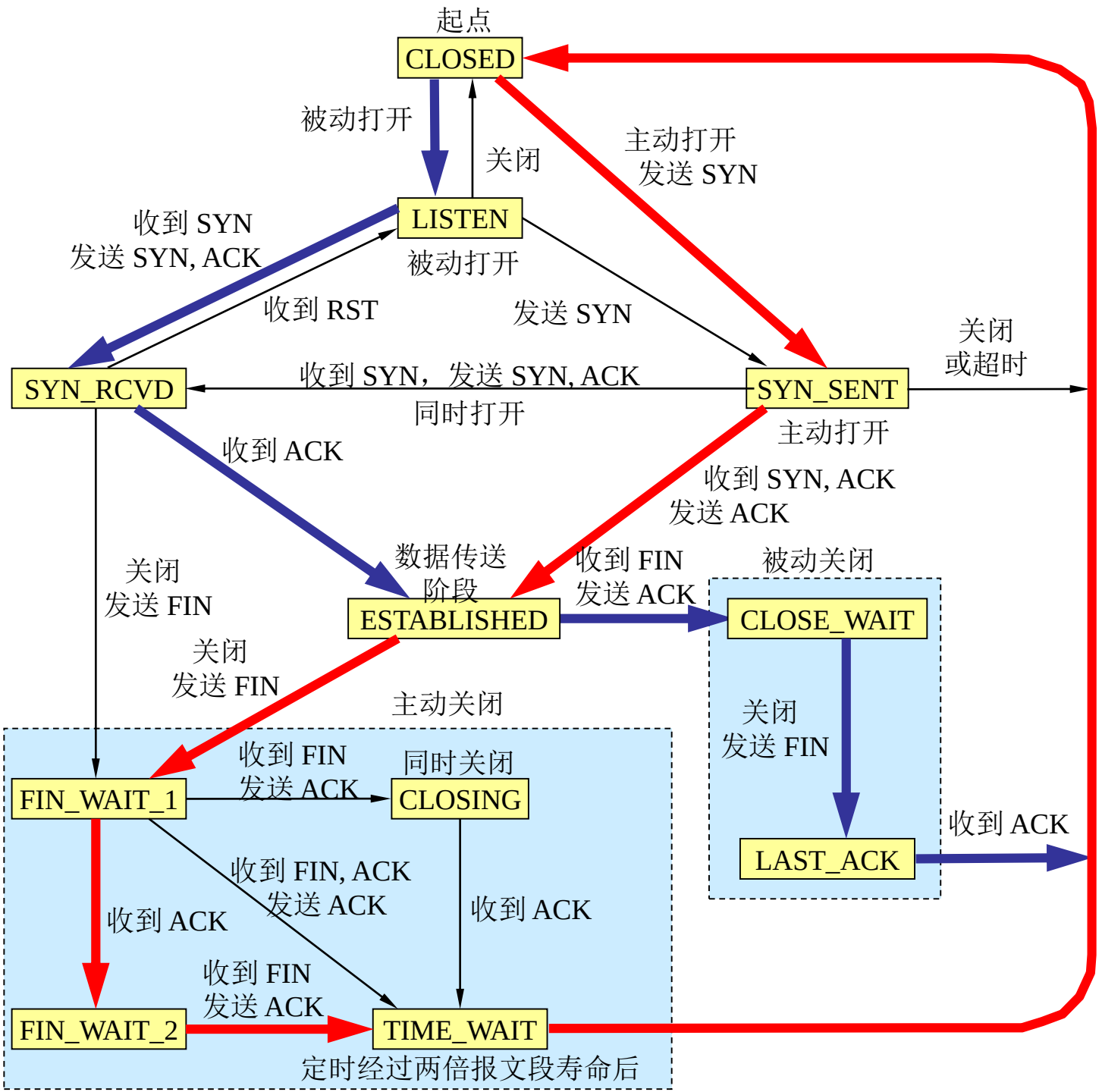
TCP deployments often implement *network-assisted* congestion control:

- two bits in IP header (ToS field) marked *by network router* to indicate congestion
  - *policy* to determine marking chosen by network operator
- congestion indication carried to destination
- destination sets ECE bit on ACK segment to notify sender of congestion
- involves both IP (IP header ECN bit marking) and TCP (TCP header C,E bit marking)



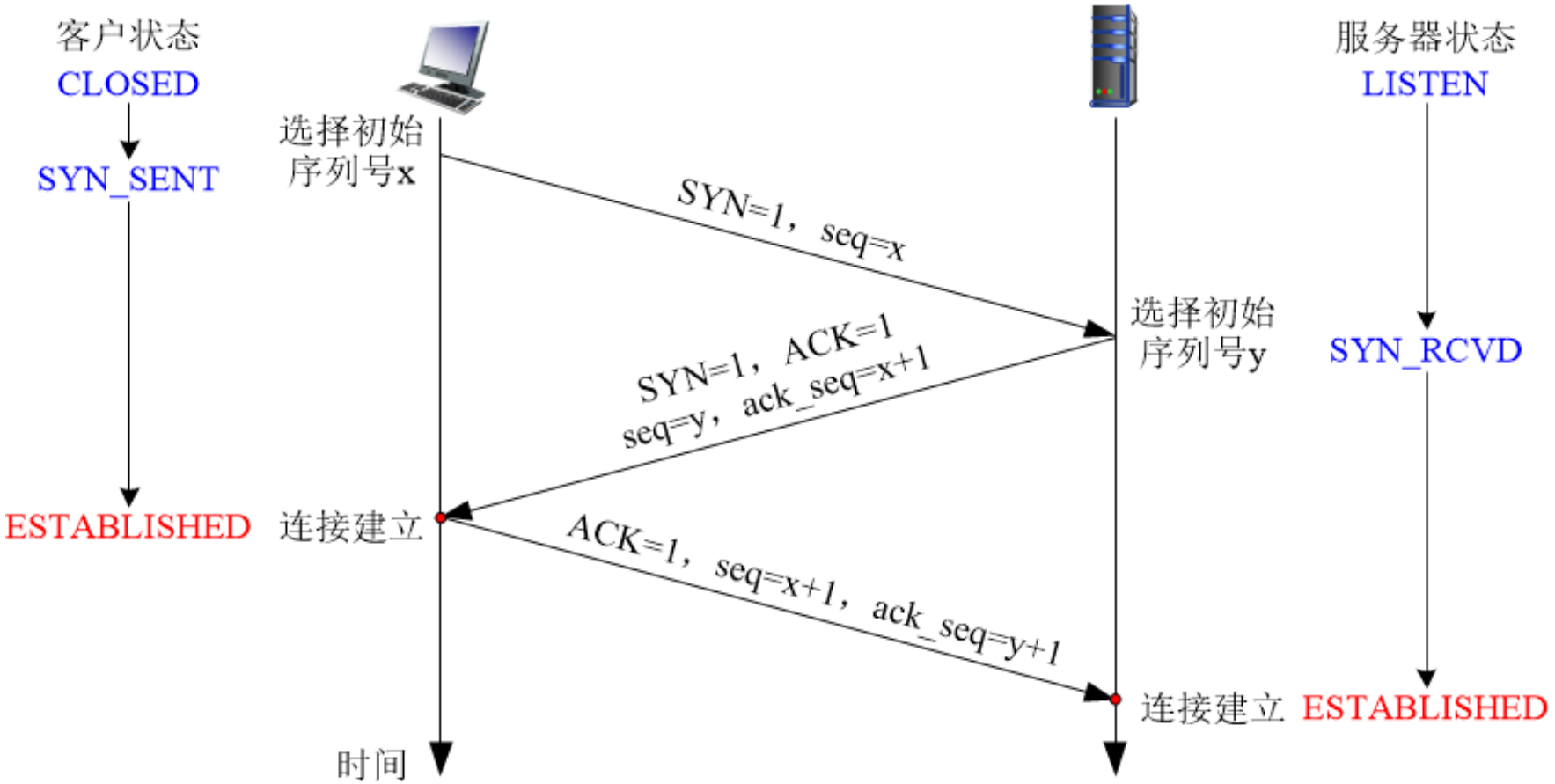
第5周 课堂教学-传输层（下）

- 开疆拓土：
  - TCP的有限状态机

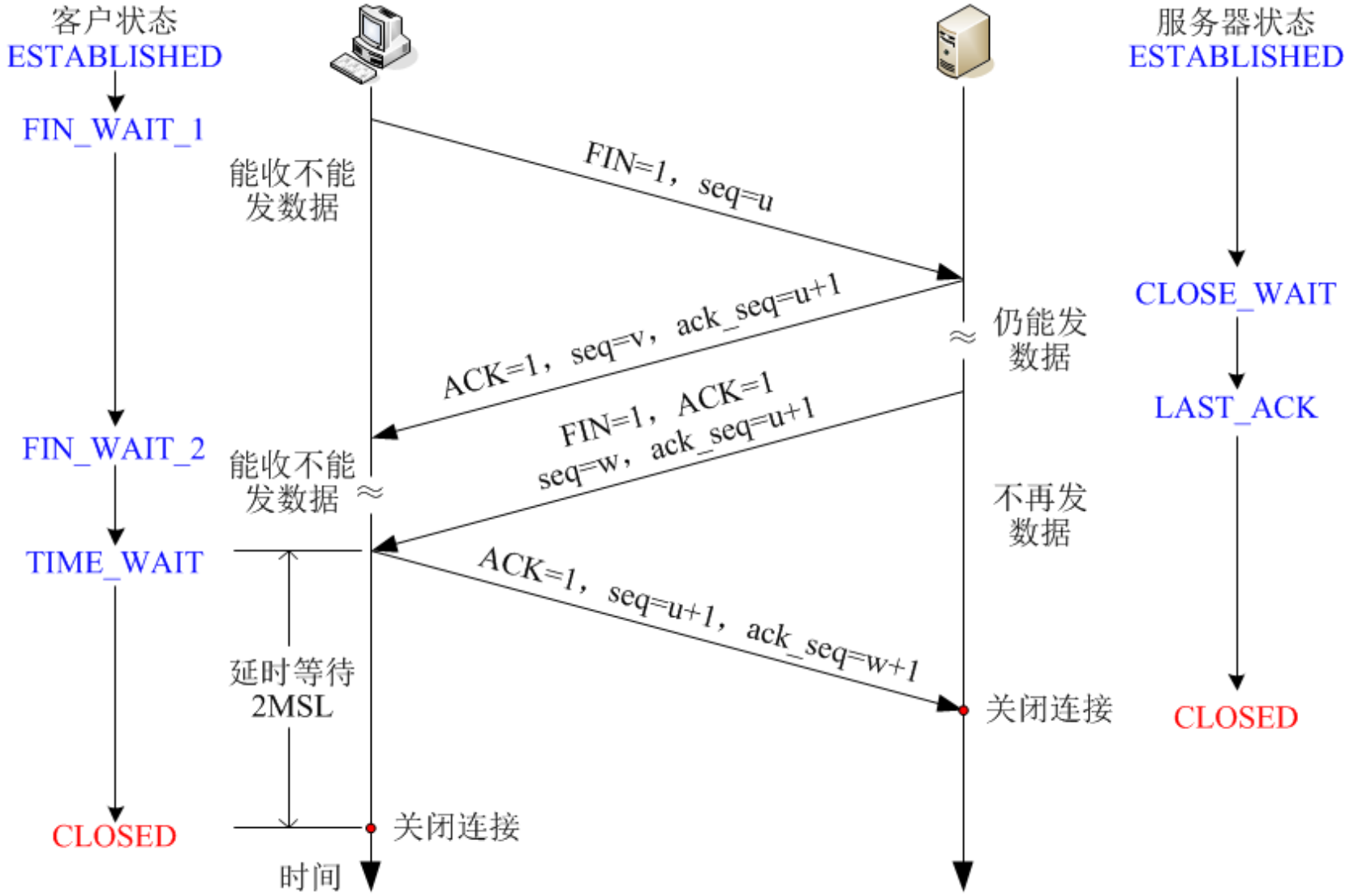

起点
CLOSED
被动打开
关闭
主动打开
发送 SYN
LISTEN
收到 SYN
发送 SYN, ACK
被动打开
收到 RST
发送 SYN
关闭
或超时
SYN_RCVD
收到 SYN，发送 SYN, ACK
同时打开
SYN_SENT
主动打开
收到 ACK
收到 SYN, ACK
发送 ACK
关闭
发送 FIN
数据传送
阶段
收到 FIN
发送 ACK
被动关闭
ESTABLISHED
CLOSE_WAIT
关闭
发送 FIN
主动关闭
关闭
发送 FIN
收到 FIN
发送 ACK
同时关闭
FIN_WAIT_1
CLOSING
关闭
发送 FIN
收到 ACK
收到 FIN, ACK
发送 ACK
收到 ACK
LAST_ACK
收到 ACK
收到 FIN
发送 ACK
FIN_WAIT_2
TIME_WAIT
定时经过两倍报文段寿命后

TCP的有限状态机

TCP建连过程的状态


客户状态
CLOSED
SYN_SENT
ESTABLISHED
选择初始
序列号x
SYN=1，seq=x
SYN=1，ACK=1
seq=y，ack_seq=x+1
ACK=1，seq=x+1，ack_seq=y+1
连接建立
时间
选择初始
序列号y
服务器状态
LISTEN
SYN_RCVD
连接建立
ESTABLISHED

TCP断连过程的状态


客户状态
ESTABLISHED
FIN_WAIT_1
FIN_WAIT_2
TIME_WAIT
CLOSED
能收不能
发数据
FIN=1，seq=u
ACK=1，seq=v，ack_seq=u+1
仍能发
数据
FIN=1，ACK=1
seq=w，ack_seq=u+1
能收不能
发数据
不再发
数据
ACK=1，seq=u+1，ack_seq=w+1
延时等待
2MSL
关闭连接
关闭连接
时间
服务器状态
ESTABLISHED
CLOSE_WAIT
LAST_ACK
CLOSED

第5周 课堂教学-传输层（下）

- 解疑释惑：
  - 1.如何理解TCP协议的点对点特性？
  - 2.TCP是GBN协议还是SR协议？
  - 3.TCP如何处理未按序到达的数据？
  - 4.TCP的确认序号(ack_seq)的意义？
  - 5.TCP如何断开连接？为什么？
  - 6.TCP的发送窗口大小如何确定？
  - 7.TCP拥塞控制的慢启动阶段的拥塞窗口如何变化？
  - 8.TCP拥塞控制的拥塞避免阶段的拥塞窗口如何变化？

第5周 课堂教学-传输层（下）

- 演武修文：
  - 课堂测验

单选题 1分

设置

主机甲与主机乙间已建立一个TCP连接，主机甲向主机乙发送了两个连续的TCP段，分别包含300字节和500字节的有效载荷，第一个段的序列号为200，主机乙正确接收到两个段后，发送给主机甲的确认序列号是

A 500

B 700

C 800

D 1000

提交

单选题 1分

设置

一个TCP连接总是以1 KB的最大段长发送TCP段，发送方有足够多的数据要发送。当拥塞窗口为16 KB时发生了超时，如果接下来的4个RTT（往返时间）时间内的TCP段的传输都是成功的，那么当第4个RTT时间内发送的所有TCP段都得到肯定应答时，拥塞窗口大小是

A 7 KB

B 8 KB

C 9 KB

D 16 KB

提交

单选题 1分

设置

主机甲和主机乙之间已建立了一个TCP连接，TCP最大段长度为1 000字节。若主机甲的当前拥塞窗口为4 000字节，在主机甲向主机乙连续发送两个最大段后，成功收到主机乙发送的对第一个段的确认段，确认段中通告的接收窗口大小为2 000字节，则此时主机甲还可以向主机乙发送的最大字节数是

A 1000

B 2000

C 3000

D 4000

提交

设置

某客户通过一个TCP连接向服务器发送数据的部分过程如下图所示。客户在t_0时刻第一次收到确认序列号ack_seq=100的段，并发送序列号seq=100的段，但发生丢失。若TCP支持快速重传，则客户重新发送seq=100段的时刻是

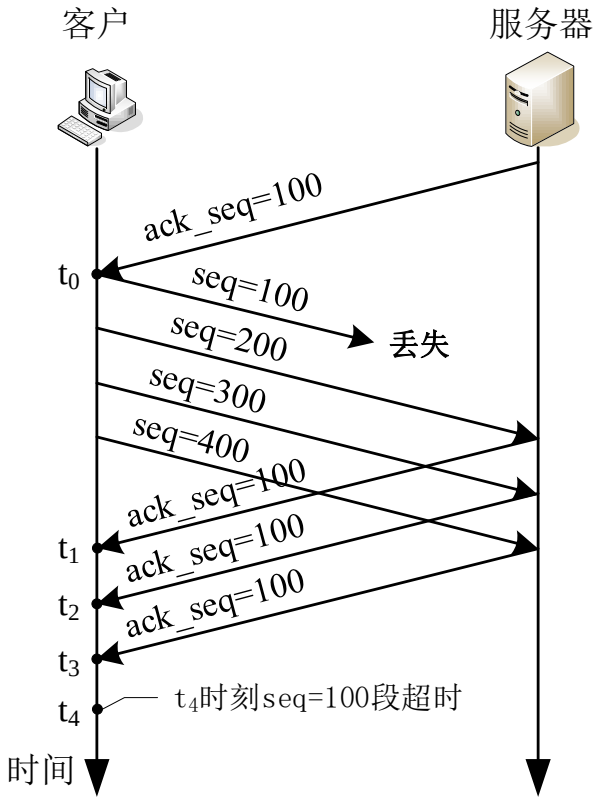
A t_1

B t_2

C t_3

D t_4

提交



主机甲向主机乙发送一个(SYN = 1, seq = 11220)的TCP段，期望与主机乙建立TCP连接，若主机乙接受该连接请求，则主机乙向主机甲发送的正确的TCP段可能是

A (SYN=0, ACK=0, seq=11221, ack_seq=11221)

B (SYN=1, ACK=1, seq=11220, ack _seq=11220)

C (SYN=1, ACK=1, seq=11221, ack _seq=11221)

D (SYN=0, ACK=0, seq=11220, ack _seq=11220)

第6周 课堂教学-网络层（上）

- 束广就狭：（30分钟）第5组总结报告
  - 总结网络层服务，转发与路由，虚电路网络与数据报网络，IP协议与IP数据报，IP地址，IP子网，IP子网划分与子网掩码，CIDR与路由聚合，路由表，DHCP协议。
- 开疆拓土：（15分钟）
  - 路由器结构及各组成部分的主要功能
- 质疑辨惑：（45分钟）
  - 1.虚电路网络有什么特点？数据报网络有什么特点？两者有什么共同点？
  - 2.IP数据报首部长度字段占几位？其单位是什么？
  - 3.IP数据报首部的片偏移量字段占几位？为什么？
  - 4.什么是子网掩码？如何取值？作用是什么？
  - 5.什么是默认网关？作用是什么？
  - ……
- 解疑释惑：（10分钟）
  - 解答疑问
- 演武修文：（10分钟）
  - 课堂测验
  - 讲解

哈爾濱工業大學
HARBIN INSTITUTE OF TECHNOLOGY

立足航天，服务国防，面向国民经济主战场

谢 谢！