

# 计算机网络原理 串 讲

崔 勇

清华大学计算机系



计算机网络  
教案社区



# 绪论目标

---



清华大学  
Tsinghua University



计算机网络教案社区

1. 掌握典型交换方式及其优缺点
2. 掌握分层结构和网络协议（核心内容）
3. 掌握网络参考模型（核心内容）
4. 掌握计算机网络主要度量的含义



# 典型交换方式的比较



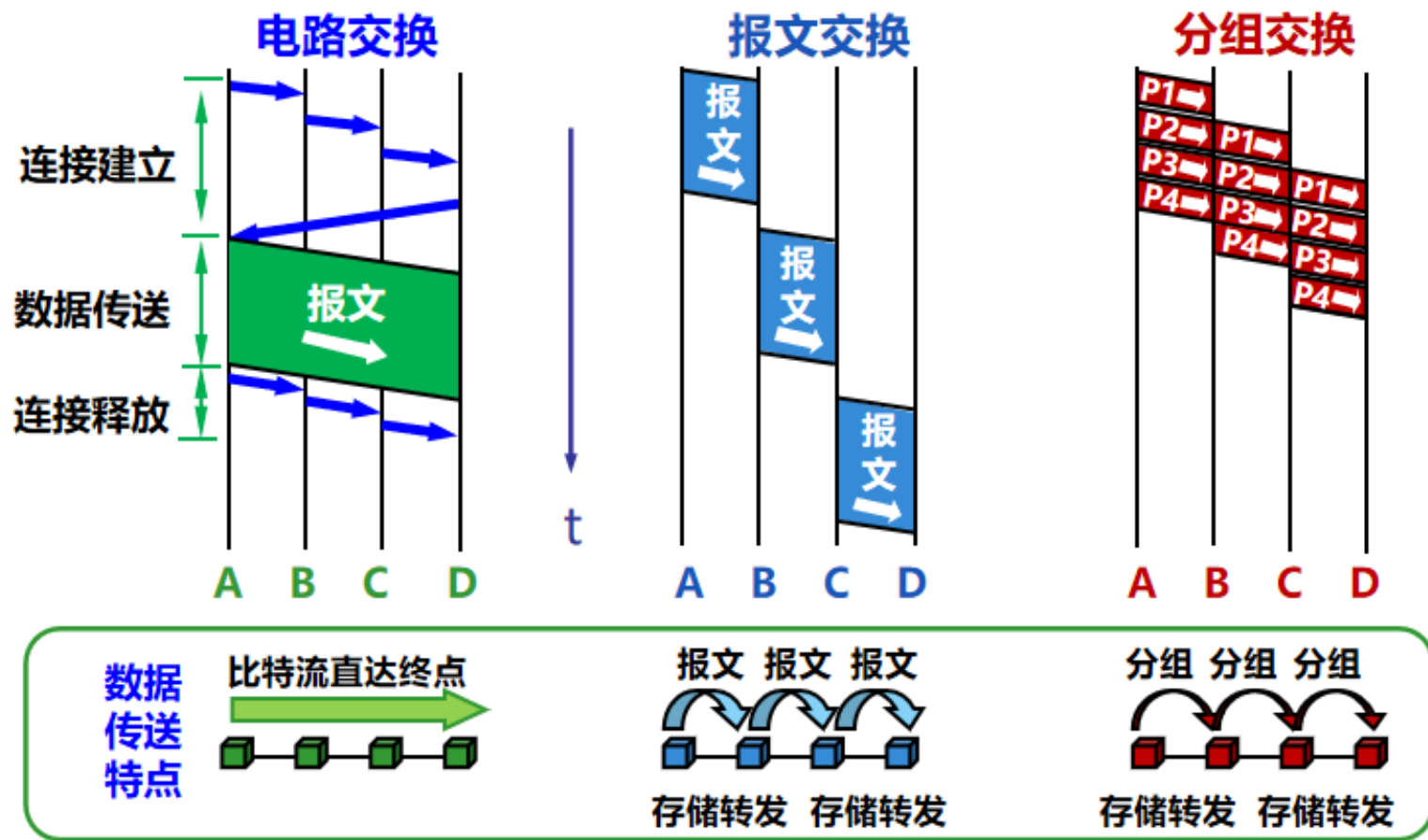
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 三种交换的比较

- 电路交换需要建立连接并预留资源，难以实现灵活复用
- 报文交换和分组交换较灵活，抗毁性高，在传送**突发数据**时可提高网络利用率
- 由于分组长度小于报文长度，分组交换比报文交换的**时延小**，也具有更好的**灵活性**



分组交换适合有大量突发数据传输需求的互联网



# 协议分层结构

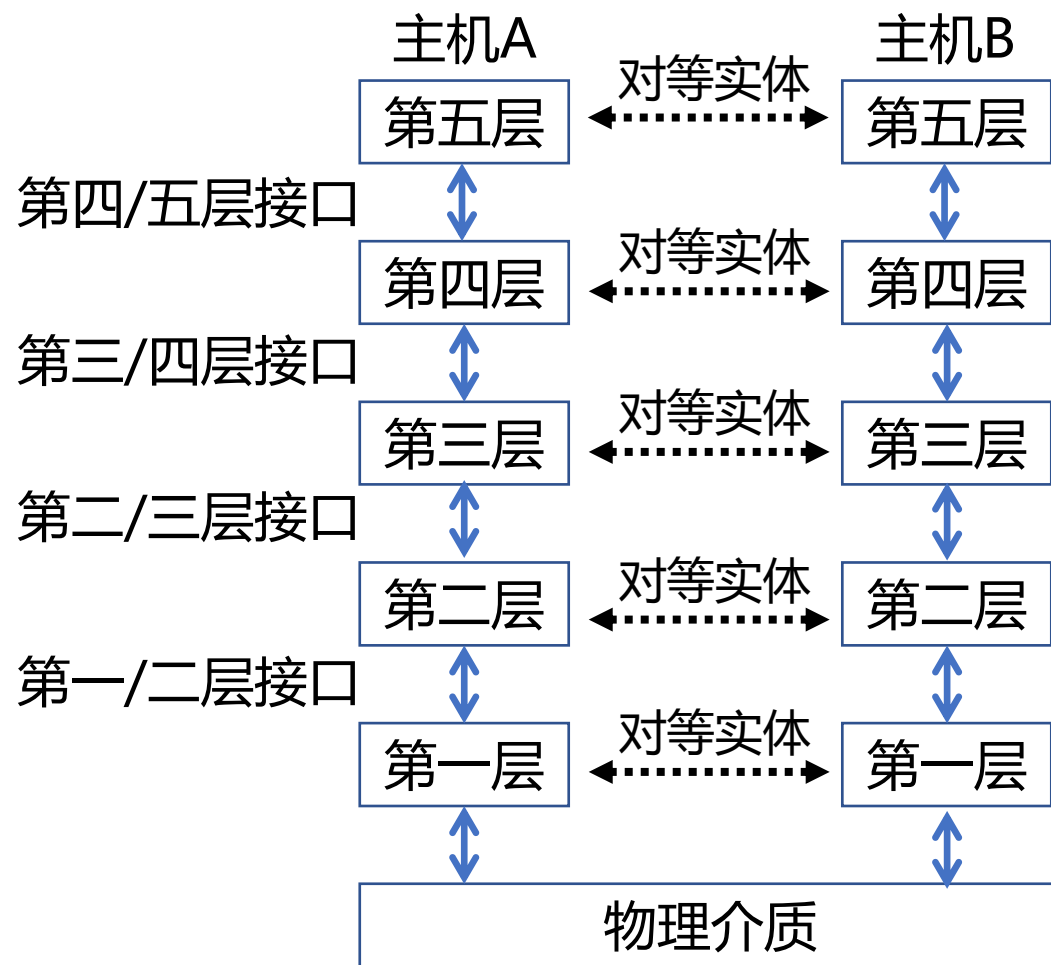


清华大学  
Tsinghua University



计算机网络教案社区

- 层次栈
  - 为降低网络设计的复杂性，各层独立，加速技术演进
  - 每一层都使用其下一层所提供的服务，并为上层提供自己的服务
- 对等实体
  - 不同机器上构成相应层次的实体成为对等实体
- 接口
  - 在每一对相邻层次之间的是接口；接口定义了下层向上层提供哪些**服务原语**
- 网络体系结构
  - 层和协议的集合





# 服务与协议的关系

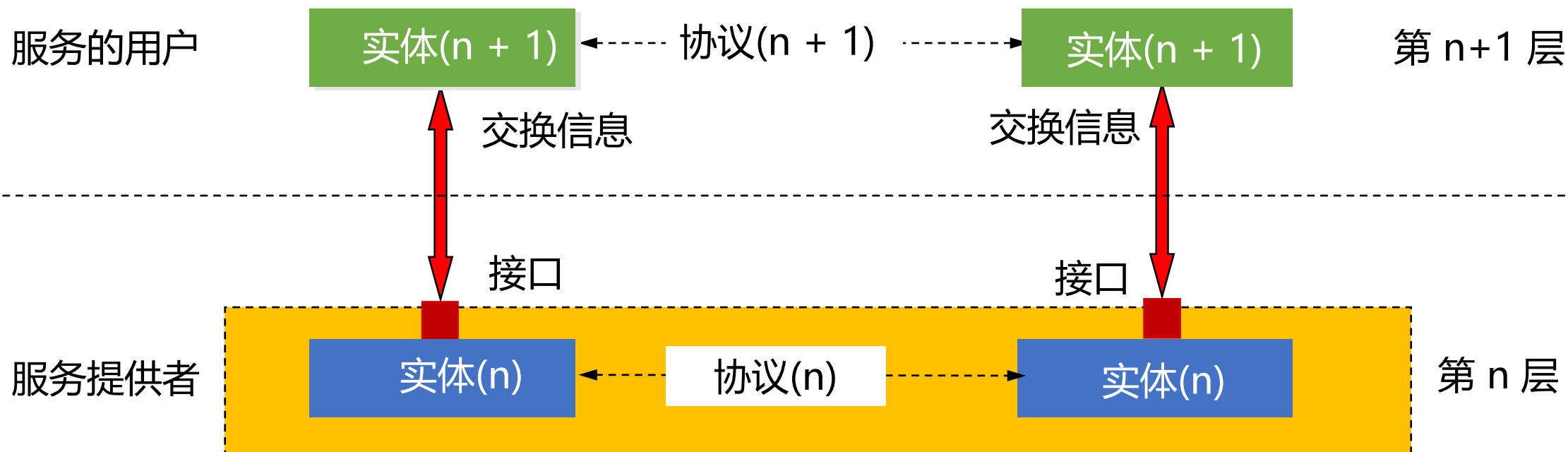


清华大学  
Tsinghua University



计算机网络教案社区

- 协议是“水平”的，服务是“垂直”的
- 实体使用协议来实现其定义的服务
- 上层实体通过接口使用下层实体的服务





# 本教程内容的分层组织



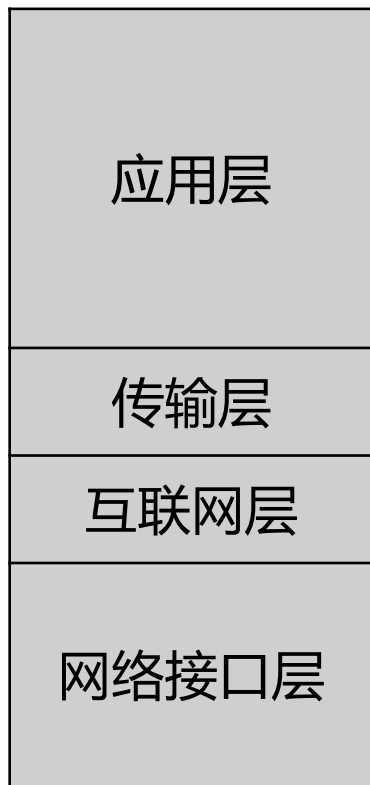
清华大学  
Tsinghua University



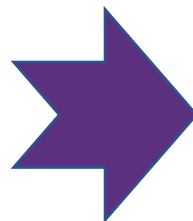
计算机网络教案社区



OSI 7层模型



TCP/IP 4层模型



本教程的分层组织

- 突出核心概念
- 区分接口与分层
- 体现完整性
- 体现通用性
- 简化分层，易于教学



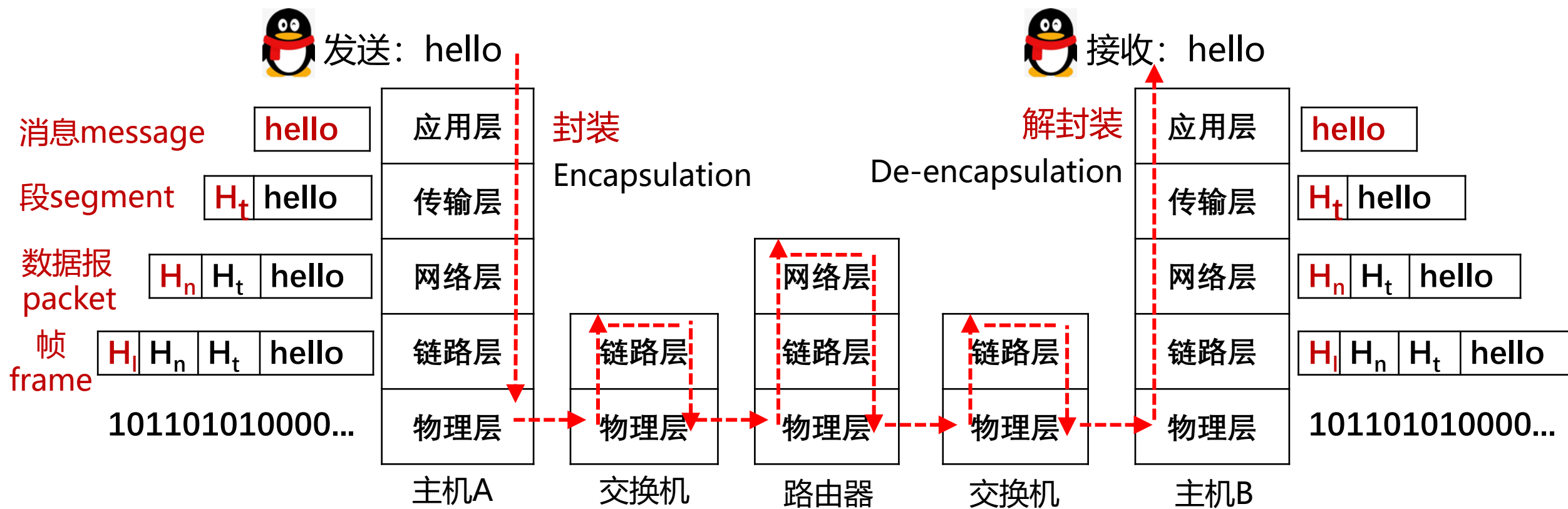
# 分层模型与网络实例



清华大学  
Tsinghua University



计算机网络教案社区



- 流: 五元组
- 每个设备工作在哪层?

- 数据面/控制面
- 分用/复用与封装



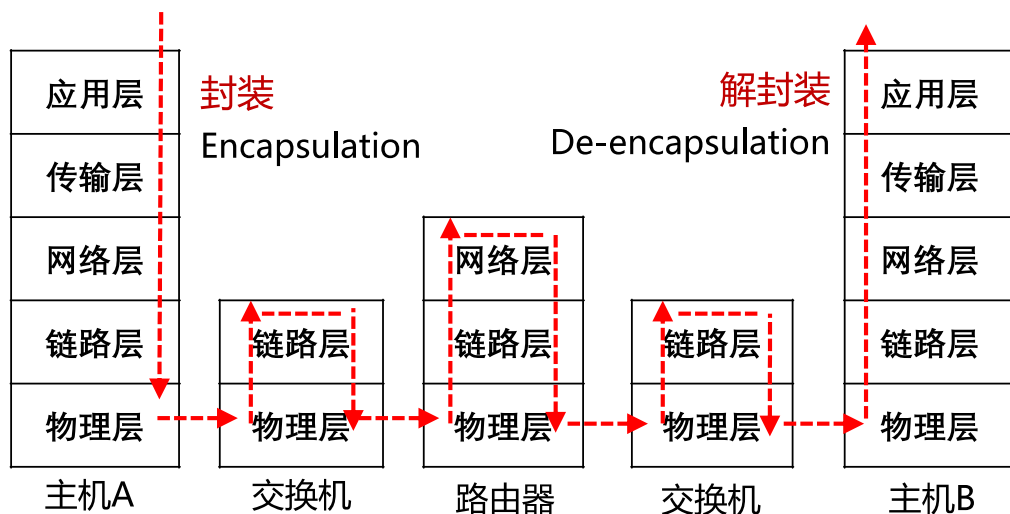
# 分层参考模型



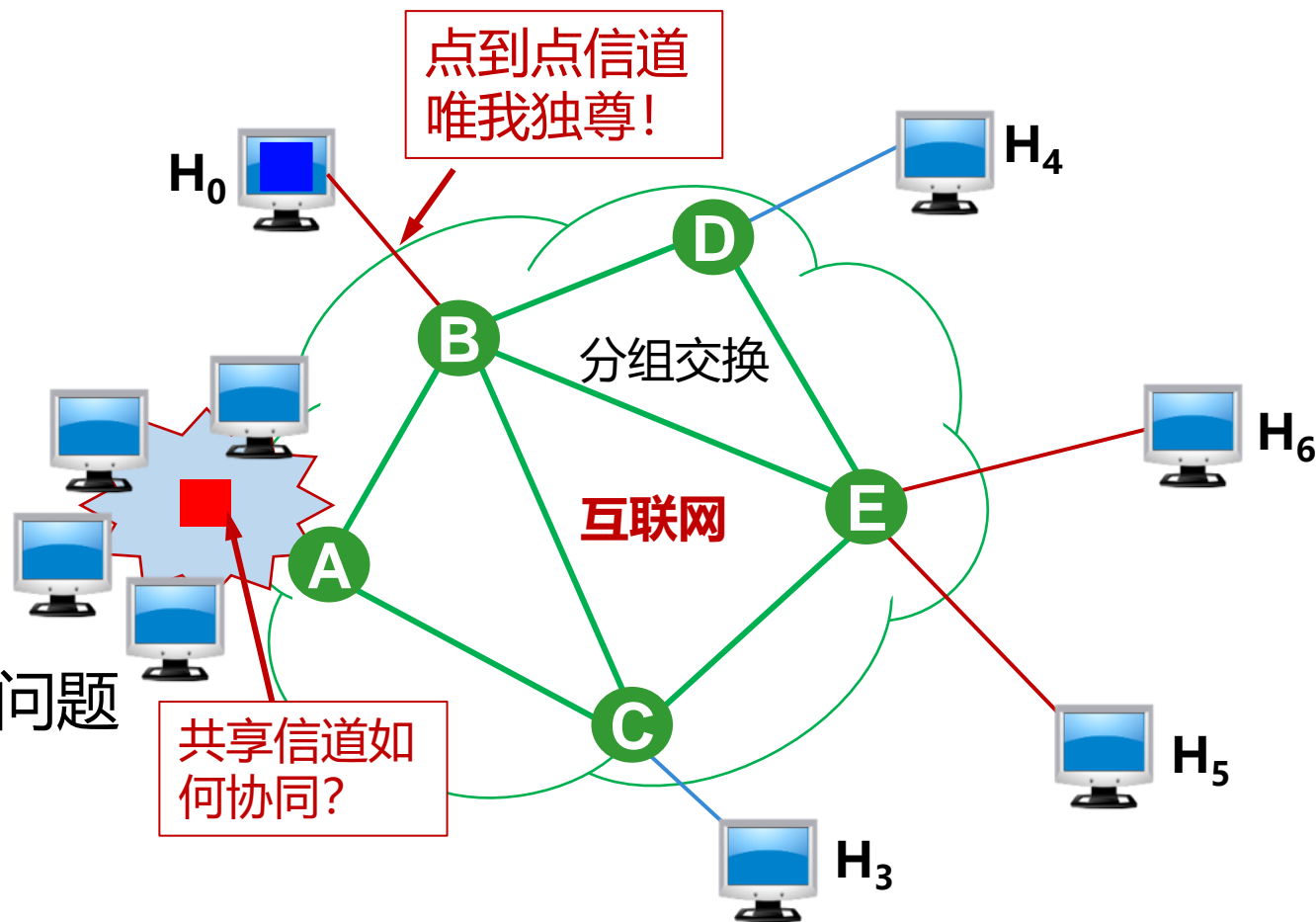
清华大学  
Tsinghua University



计算机网络教案社区



- 每层的主要功能
- 主要协议及其应用场景和解决的问题
- 主要协议的创新思路和优缺点
- 不用协议之间如何相互配合







# TCP/IP参考模型



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ TCP/IP参考模型

- 摒弃电话系统中“笨终端&聪明网络”的设计思路
- **端对端原则**：采用**聪明终端&简单网络**，由**端系统**负责丢失恢复等等，简单网络大大提升了可扩展性
- 实现了建立在简单的、不可靠部件上的可靠系统

## ➤ IP分组交换的特点

- 可在各种底层物理网络上运行(**IP over everything**)
- 可支持各类上层应用(**Everything over IP**)
- 每个IP分组携带各自的目的地址，网络核心功能简单（通过路由表转发分组），适应爆炸性增长



TCP/IP的沙漏模型



# 网络度量单位



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 常见的数据传输速率描述

- **比特率**：数字信道上的数据传送速率
- **带宽**：网络中某通道传送数据的能力
- **包转发率**：以包为单位的转发速率

## ➤ 常见的数据发送后的统计指标

- **吞吐量**：单位时间内通过某个网络或信道的数据量（有效吞吐量goodput）
- **丢包率**：所丢失数据包的数量占所发送数据包的比率
- **利用率**：某信道有效利用的比例

## ➤ 常见的时延相关指标

- **时延**：数据从一端到另一端所需的时间
- **时延带宽积**、往返时间 **RTT**、时延**抖动**
- 流完成时间**FCT**

- 传输时延(transmission delay)  
又称为发送时延
- 传播时延(propagation delay)
- 处理时延(processing delay)
- 排队时延(queueing delay)



# 物理层目标



清华大学  
Tsinghua University



计算机网络教案社区

- 理解物理层基本概念和四个基本特性：**机械、电气、功能、过程**
- 掌握数据通信的基本术语、不同传输方式的特点
  - 数字/模拟，串行/并行，点到点/点到多点，单工、半双工和全双工
  - 异步/同步、基带/频带
  - 数据编码技术：不归零制码、曼彻斯特码、差分曼彻斯特码
  - 三种调制技术：调幅、调频、调相
- 了解不同传输介质的特点
  - 导引型传输介质、非导引型传输介质的**典型代表及其特点**
- 了解无线与卫星通信的特点和属性
- 掌握不同多路复用技术的基本原理
  - 频分、时分、统计时分复用和码分复用



# 链路层目标



清华大学  
Tsinghua University



计算机网络教案社区

- 了解数据链路层在网络体系结构中的位置及基本功能和服务
- 掌握差错检测和纠正的基本原理和典型的编码方法（核心内容）
- 掌握无错信道和有错信道上停等协议的设计和实现方法（核心内容）
- 理解停等协议的性能问题及滑动窗口协议的基本思想
- 掌握回退N和选择重传两种典型滑动窗口协议的工作机制（核心内容）
- 了解点到点链路层协议PPP，不要求PPPoE



# 成帧 (Framing)



## ➤ 关键问题：如何标识一个帧的开始？

- 接收方必须能从物理层接收的比特流中明确区分出一帧的开始和结束，这个问题被称为帧同步或帧定界
- 关键：选择何种定界符？定界符出现在数据部分如何处理？

## ➤ 成帧 (framing) 的方式

- 字节计数法 ( Byte count )
- 带字节填充的定界符法 ( Flag bytes with byte stuffing )
- 带比特填充的定界符法 ( Flag bits with bit stuffing )
- 物理层编码违例 ( Physical layer coding violations )



# 差错检测和纠正



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 目标

- 保证一定差错检测和纠错能力的前提下，如何减少冗余信息量？

## ➤ 常用的检错码

- **奇偶检验** (Parity Check): 1位奇偶校验是最简单、最基础的检错码
- **校验和** (Checksum): 主要用于TCP/IP体系中的网络层和传输层
- **循环冗余校验** (CRC): 数据链路层广泛使用的校验方法

## ➤ 典型纠错码

- **海明码**: 以奇偶校验为基础，如何找到出错位置



# 数据链路层协议



清华大学  
Tsinghua University



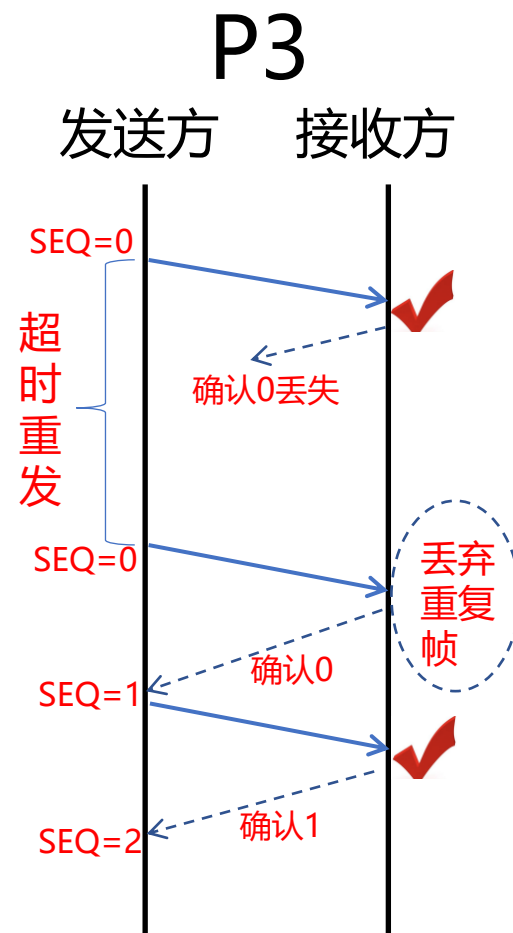
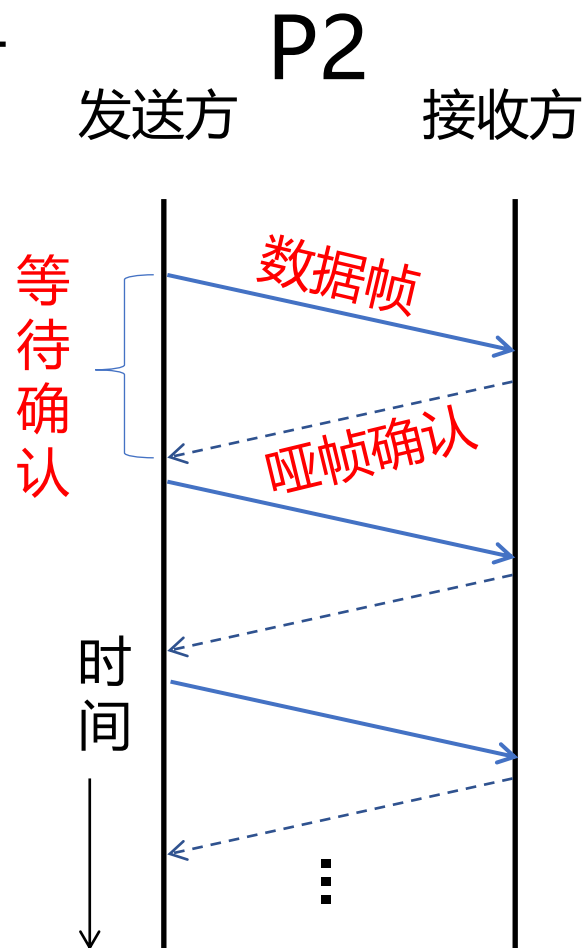
计算机网络教案社区

## ➤ 循序渐进、步步深入的协议设计

- P1 乌托邦式单工协议
- P2 无错信道单工停止-等待协议
- P3 有错信道单工停止-等待协议
- P4 一比特滑动窗口：收发二合一
- P5 回退N的滑动窗口协议
- P6 选择重传协议

## ➤ 理解其差异和优缺点

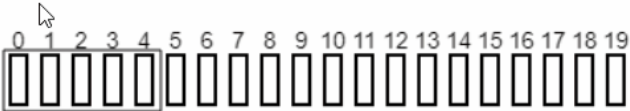
- 场景和解决的主要问题
- 使用的技术手段（需要的buffer等资源）





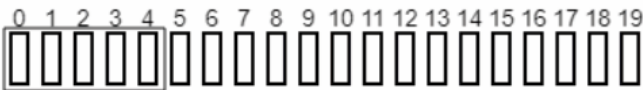
# 累计确认和选择重传的比较

Send New Pause Faster Slower Kill Packet/Ack Reset



Sender (Send Window Size = 5)  
base = 0  
nextseqnum = 0

Send New Pause Faster Slower Kill Packet/Ack Reset



Sender (Send Window Size = 5)  
base = 0  
nextseqnum = 0



Receiver (Send Window Size = 1)



Receiver (Send Window Size = 5)

Packet Received Ack Ack Received Selected

Packet Received Ack Ack Received Selected Buffered

累计确认在ack丢失时效率高

选择重传在ack丢失时效率低





# MAC子层目标



清华大学  
Tsinghua University



计算机网络教案社区

1. 了解MAC子层的位置和功能
2. 掌握两种ALOHA协议的原理和性能
3. 掌握CSMA以及CSMA/CD的工作原理及性能区别
4. 了解以太网发展史和广泛应用的原因
5. 掌握经典以太网的拓扑和帧结构
6. 掌握最小帧长与冲突域和带宽的关系
7. 掌握交换式以太网特征
8. 掌握数据链路层交换的原理
9. 掌握MAC地址表的维护
10. 了解无线局域网的拓扑结构和组织方式
11. 掌握CSMA/CA协议和其主要问题及应对方案



# 三大类多路访问协议



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 随机访问协议：努力减少冲突

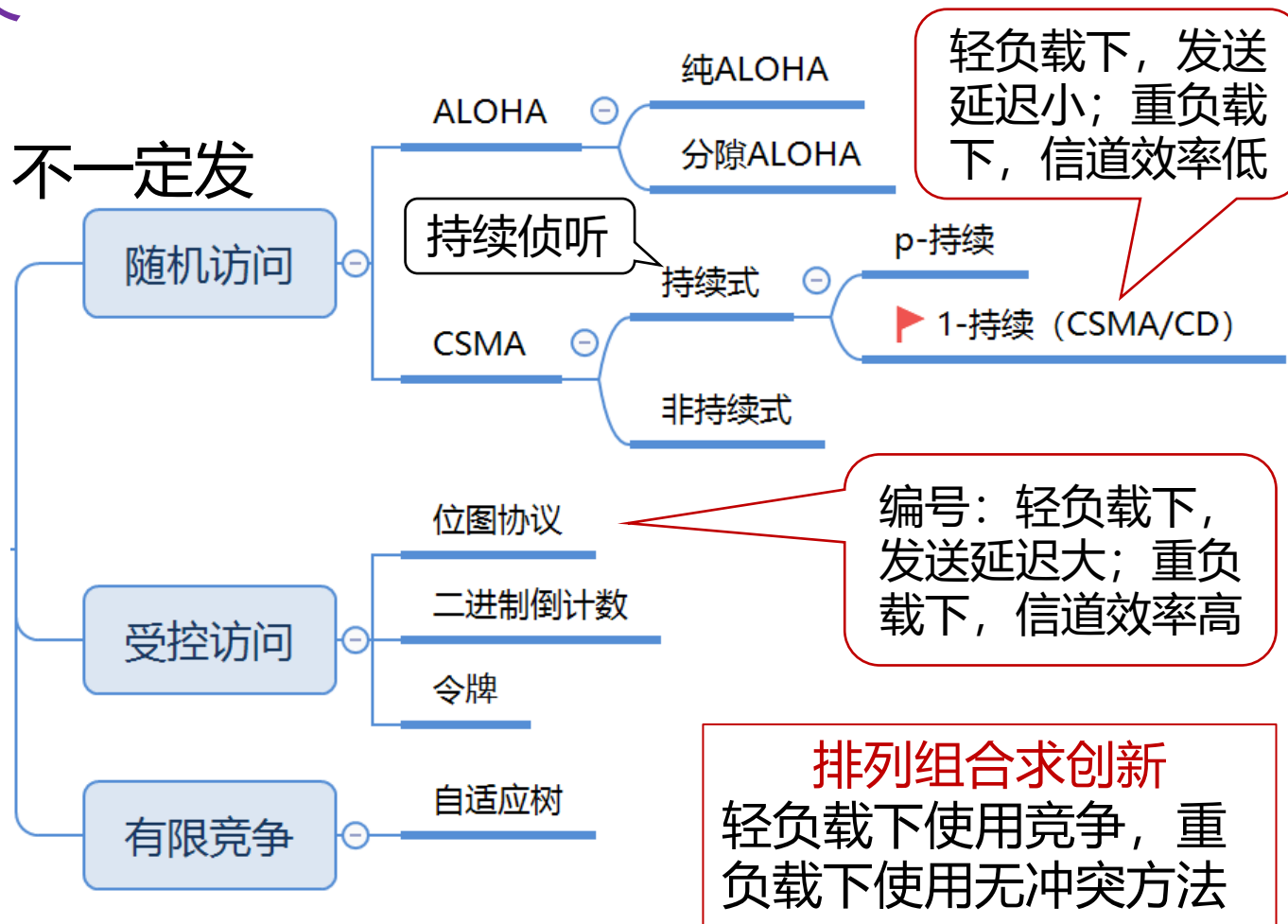
- Aloha：任性->时隙
- CSMA：先听后发；不坚持听，不一定发
- p-持续：持续侦听，概率p发送
- 时延增大，冲突不可避免

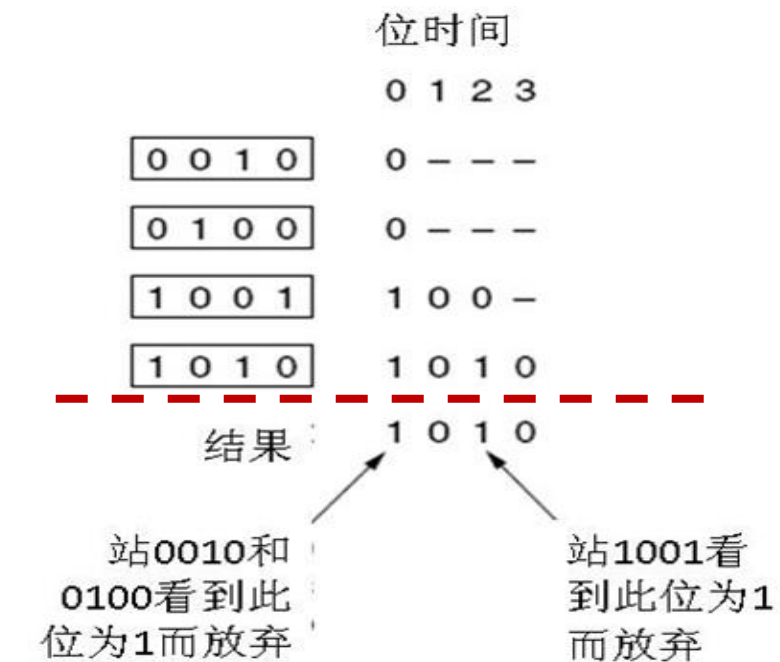
## ➤ 受控访问协议：避免冲突

- 天黑请闭眼，究竟谁要发？
- 令牌环：得令牌者发数据

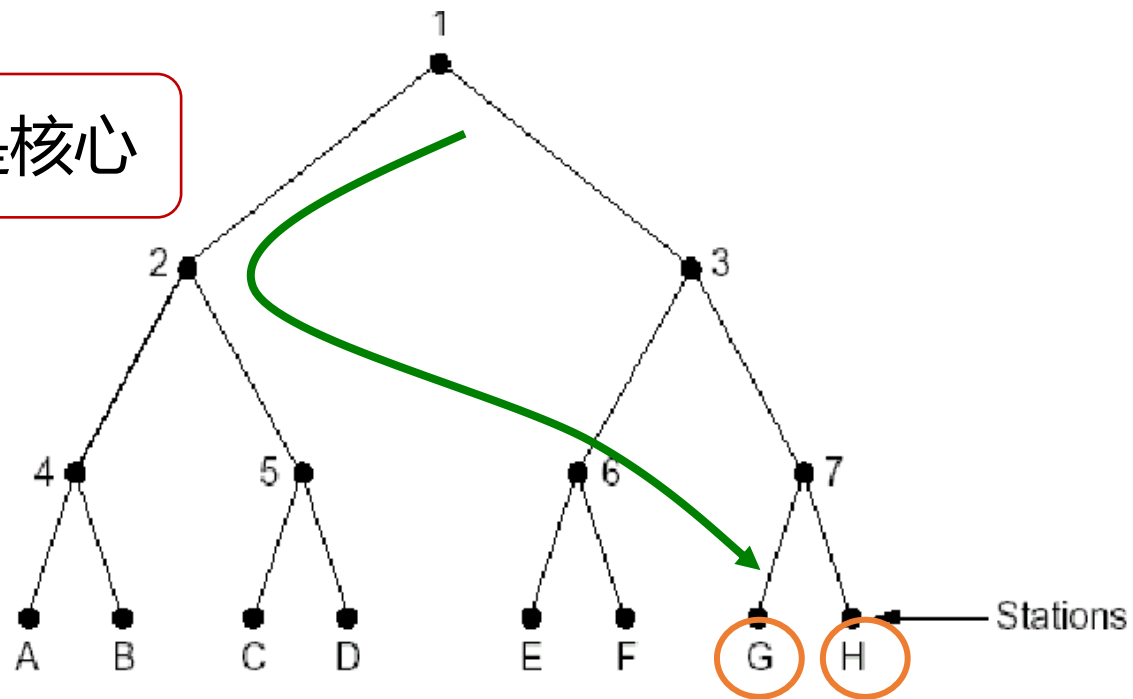
## ➤ 有限竞争协议

- 谁退避谁发送？
- 自适应树搜索协议





## 编号是核心





# 冲突窗口



清华大学  
Tsinghua University



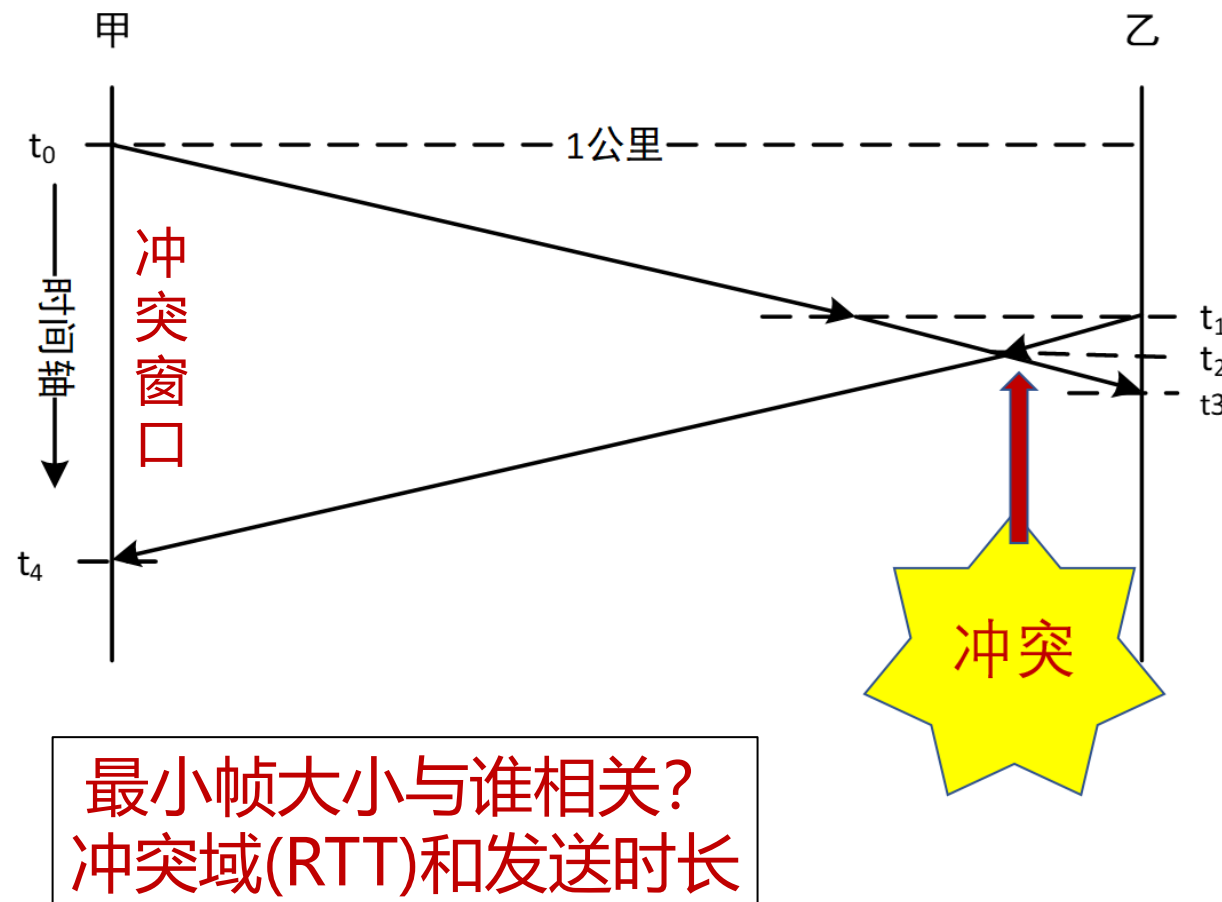
计算机网络教案社区

## ➤ 冲突窗口

- 即发送站发出帧后能检测到冲突（碰撞）的最长时间

## ➤ 冲突窗口是一个时间区间

- 可能侦听到发出的帧遭到冲突（碰撞）
- 等于最远两站传播时间的两倍，即 $2D$ （ $D$ 是单边延迟）
- $2D$ 相当于来回传播延迟 $RTT$ ：  
Round Trip Time





# 以太网



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 规模可扩展的以太网

- 冲突域太大效率低，谦让的CSMA无能为力，如何划分多个冲突域？
- **扩展性1**：集线器->交换机，**无配置并自适应拓扑**（少做无用转发）？
- 交换机：逆向学习+转发、过滤和洪泛
- **扩展性2**：冗余拓扑导致广播风暴、重复帧、MAC表震荡
- 生成树协议：根桥、根端口、指定端口，解决了问题但不够优化
- 源路由网桥：能够寻找最优路径
- **扩展性3**：大规模以太网的时延
- 链路层交换机的三种转发模式（时延v.s.错误）
- 存储转发、直通模式、无碎片模式
- **扩展性4**：局域网规模再大点？VLAN
- 交换机隔离冲突域，但广播帧呢？要隔离不同的局域网
- 类型：端口、MAC地址、协议、子网
- 如何区分不同VLAN且主机不感知：哪些端口属于哪个VLAN，标记的添加/剥除

规模扩展引入的  
不同问题



# 采用交换机隔离冲突域



清华大学  
Tsinghua University

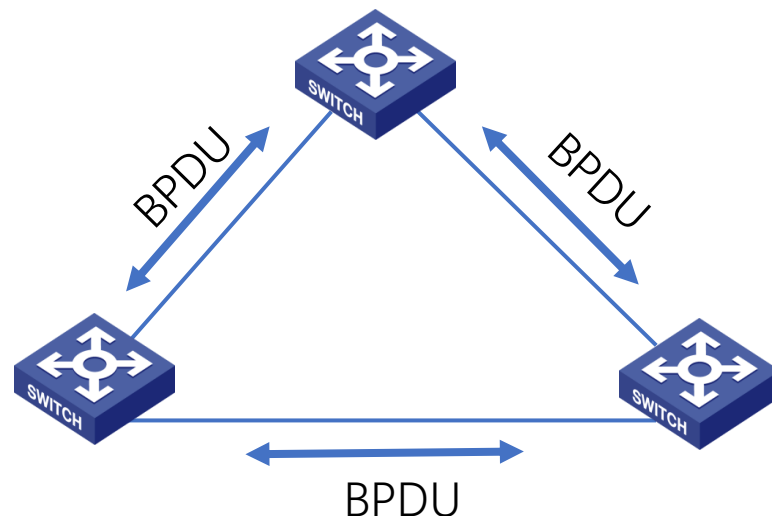


计算机网络教案社区

## ➤ 透明网桥与生成树协议

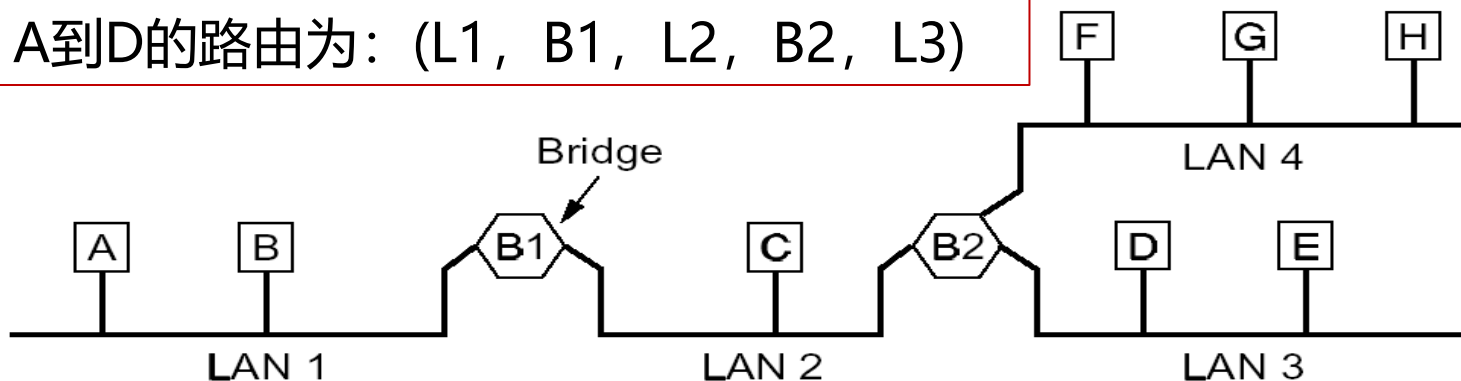
- 问题：广播风暴、重复帧、MAC地址表震荡
- 参与的交换机（网桥）：**收发**桥协议数据单元BPDU
- 选举产生**根桥**、**根端口**、**指定端口**，形成生成树

哪些端口工作？



## ➤ 源路由网桥

A到D的路由为：(L1, B1, L2, B2, L3)



如何实现透明网桥？  
无配置且  
避免不必要的转发



# 无线局域网



清华大学  
Tsinghua University



计算机网络教案社区

- 组网方式：基础架构模式 & 自组织架构
- CSMA/CD失效的2个原因：冲突检测困难，无线传输范围
- 无线介质访问与CSMA/CA协议原理
  - 差错检测与确认重传
  - 基于帧间间隔的优先级控制：SIFS、PIFS、DIFS
  - 隐藏终端、暴露终端问题：RTS/CTS握手
- 无线局域网构建与管理
  - 基础架构模式关联到AP的三个阶段：扫描、认证、关联
  - 站点漫游与睡眠管理
- IEEE 802.11协议
  - 管理（Beacon帧、认证、关联）、控制（RTS/CTS/ACK）、数据帧，四个地址域
- 家庭WLAN配置和问题分析
  - 5G路由，频道设置，信号穿墙



# 隐藏终端与暴露终端问题



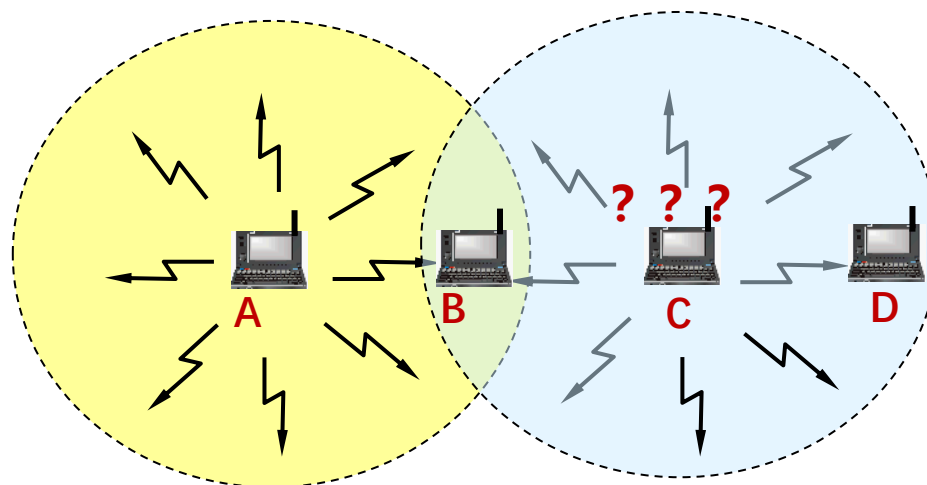
清华大学  
Tsinghua University



计算机网络教案社区

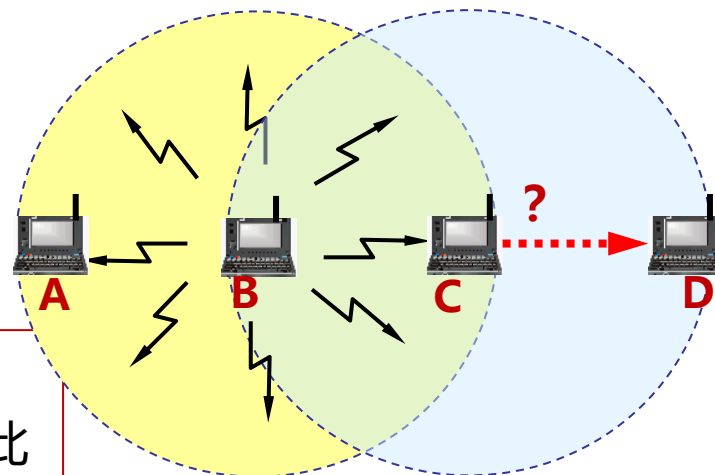
## ➤ RTS-CTS机制 (可选机制)

- 目的: 通过信道预约, 避免数据帧(长帧)冲突
- 发送端发送RTS (request to send)
- 接收端回送CTS (clear to send)
- RTS和CTS中的持续时间 (Duration) 中指明传输所需时间 (数据+控制)



C能发给D吗?

隐藏终端  
A向B发送RTS, B向A发送CTS, 由于C听到B发的CTS, 因此C静默



C能发给D吗?

暴露终端  
C没有听到CTS, 因此C可以给D发送数据





# IEEE 802.11介质访问控制



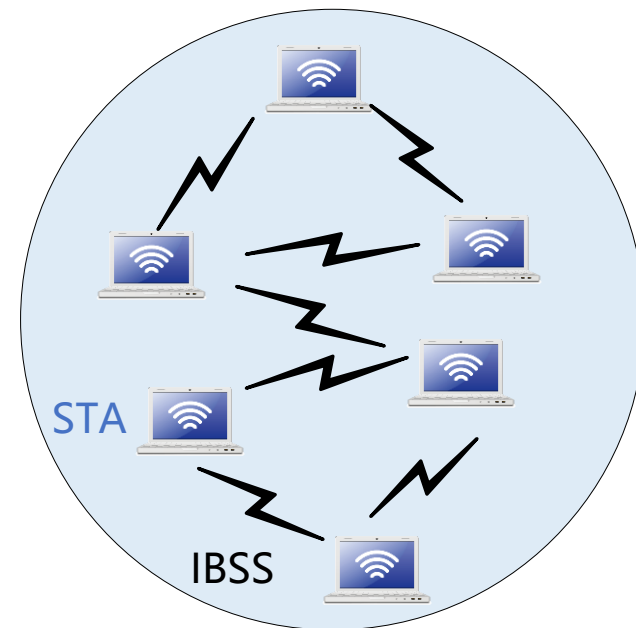
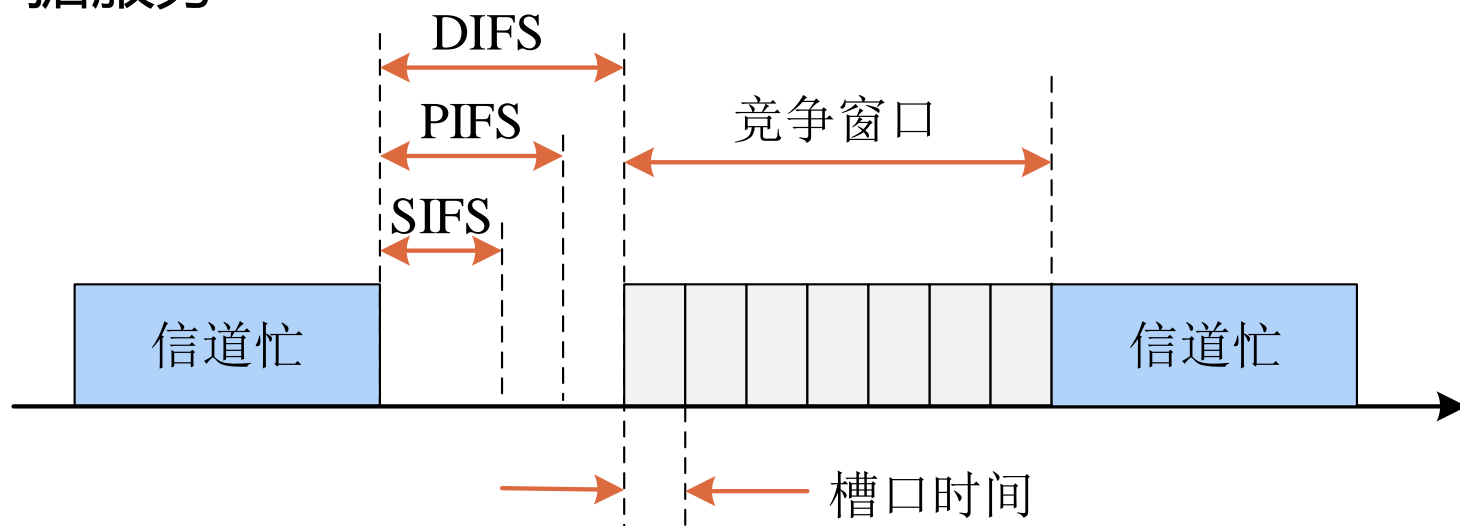
清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 不同帧间隙控制优先级

- SIFS (Short IFS) : 最高优先级, 用于Ack, CTS, 轮询响应等
- PIFS (点协调功能PCF IFS) : 中等优先级 (SIFS+1槽口时间), 轮询服务
- DIFS (分布式协调DCF IFS) : 最低优先级 (SIFS+2槽口时间), 异步数据服务





# 网络层目标



清华大学  
Tsinghua University



计算机网络教案社区

1. 掌握Internet网络层协议：IPv4/IPv6, ICMP, DHCP, NAT, ARP
2. 掌握链路状态、距离矢量等路由算法，掌握层次路由结构
3. 掌握Internet路由协议：OSPF、RIP、BGP，了解MPLS
4. 了解路由器工作原理：控制层和数据层，报文转发机制，交换结构
5. 了解网络拥塞及拥塞控制思想
6. 了解网络服务质量及设计思想
7. 了解IPv6相关机制



# 网络层关键功能



清华大学  
Tsinghua University



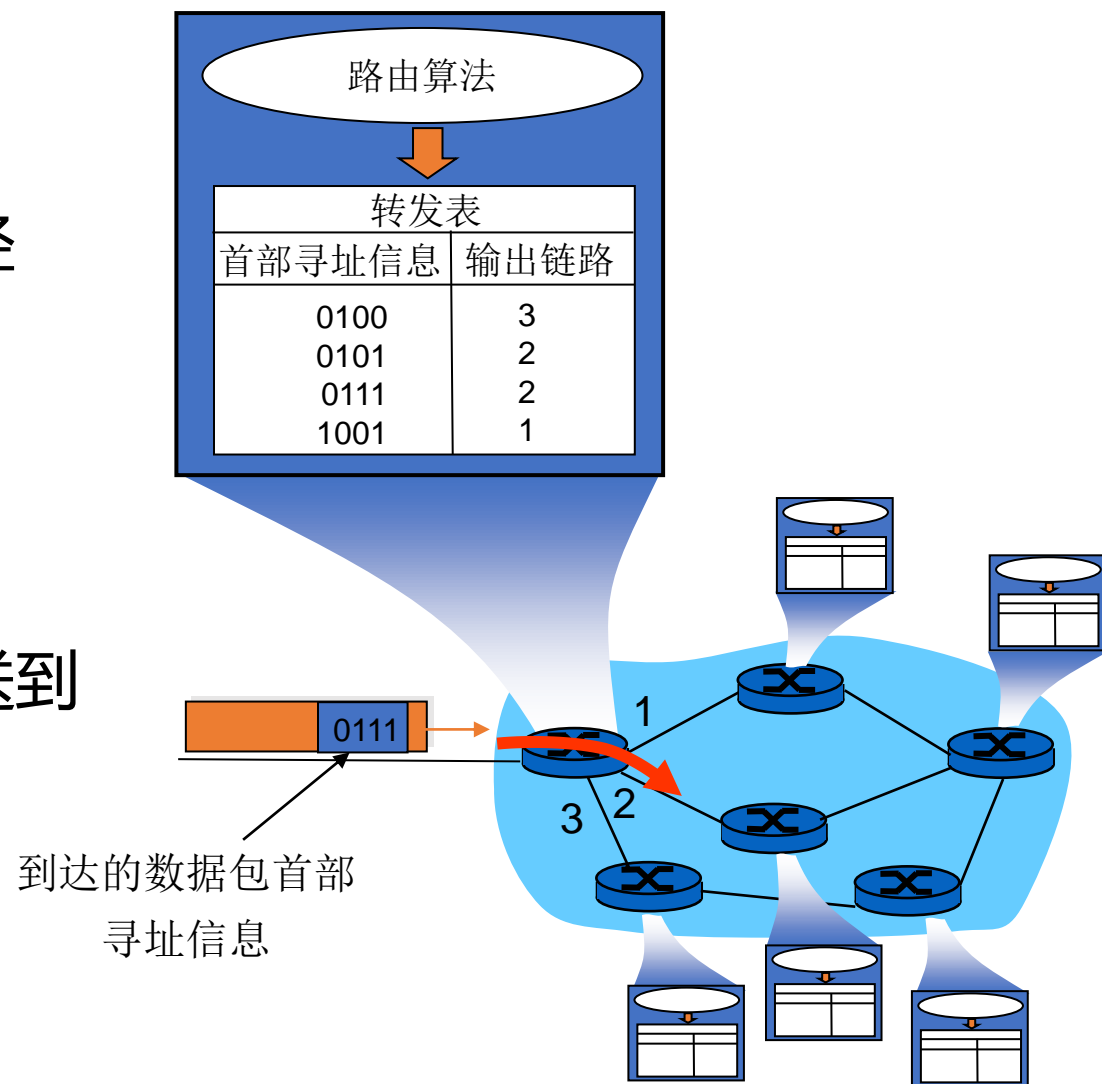
计算机网络教案社区

## ➤ 路由（控制面）

- 选择数据报从源端到目的端的路径
- 核心：路由算法与协议

## ➤ 转发（数据面）

- 将数据报从路由器的输入接口传送到正确的输出接口





# Internet网际协议



清华大学  
Tsinghua University



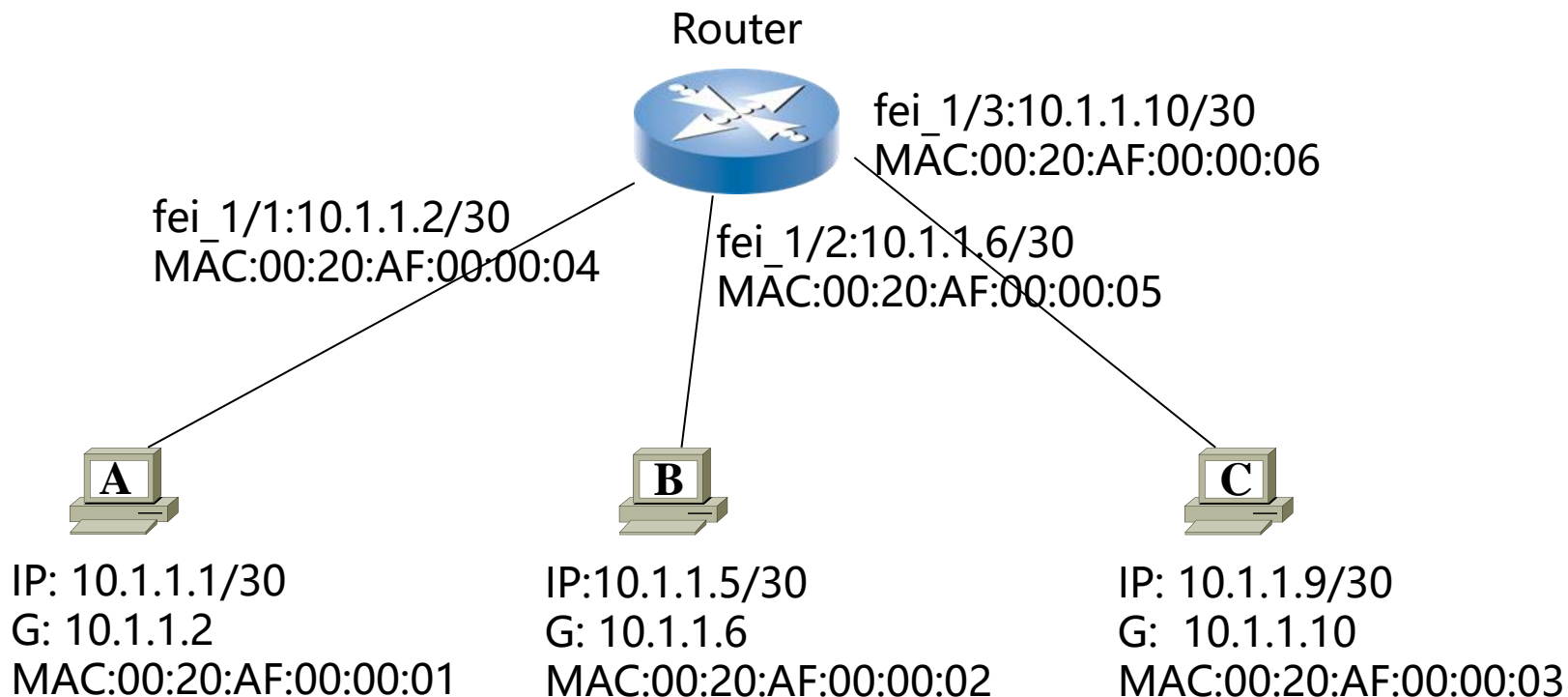
计算机网络教案社区

1. IP编址与地址聚合
2. IPv4协议
3. 地址分配DHCP
4. 地址解析协议ARP
5. Internet控制  
报文协议ICMP

## 地址配置与转发

直接交付：与目的主机在同一个IP子网内

间接交付：与目的主机不在同一个IP子网内





# 路由算法



清华大学  
Tsinghua University

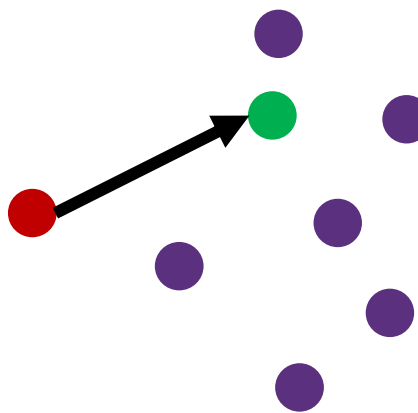


计算机网络教案社区

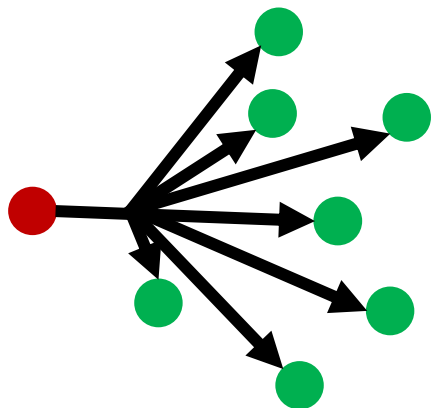
## ➤ 路由算法的分类与概念

- 链路状态路由 (Dijkstra算法) v.s. 距离向量路由 (Bellman-Ford算法)
- 层次路由: 提升互联网的可扩展性和可管理性
- 广播-单播-组播-任意播
- 洪泛: 目标、思路 and 可能的问题

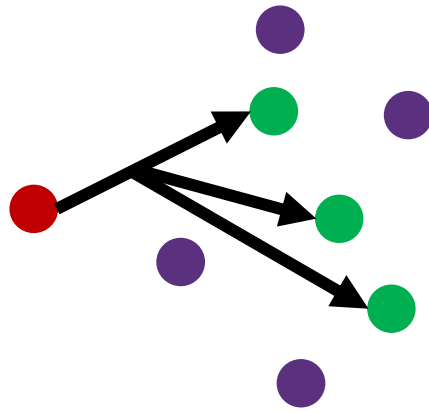
集中式v.s.分布式  
算法v.s.协议



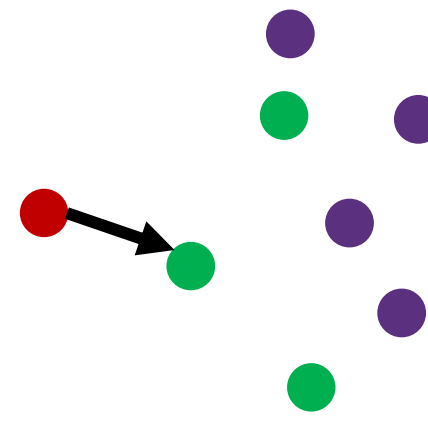
单播



广播



组播



选播



# 路由协议



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 路由协议的基本需求

- 控制平面：路由协议产生路由表
- 数据平面：基于路由表进行分组转发
- 高效收敛且无回路，尽量简单，可扩展，支持多种策略

如何一步步提升  
可扩展性

## ➤ 域内：OSPF/RIP/MPLS；域间BGP

- RIP：距离向量，基于UDP组播；无确认/30s重复发；16跳&水平分裂
- OSPF：链路状态，基于IP组播；描述、请求、发送、确认；序列号
- BGP：**路径**向量，基于TCP；可靠的下层传输协议；序列号
- MPLS：借鉴电路交换，面向连接，标签交换，更强控制，实现VPN和TE



# 思考与发明

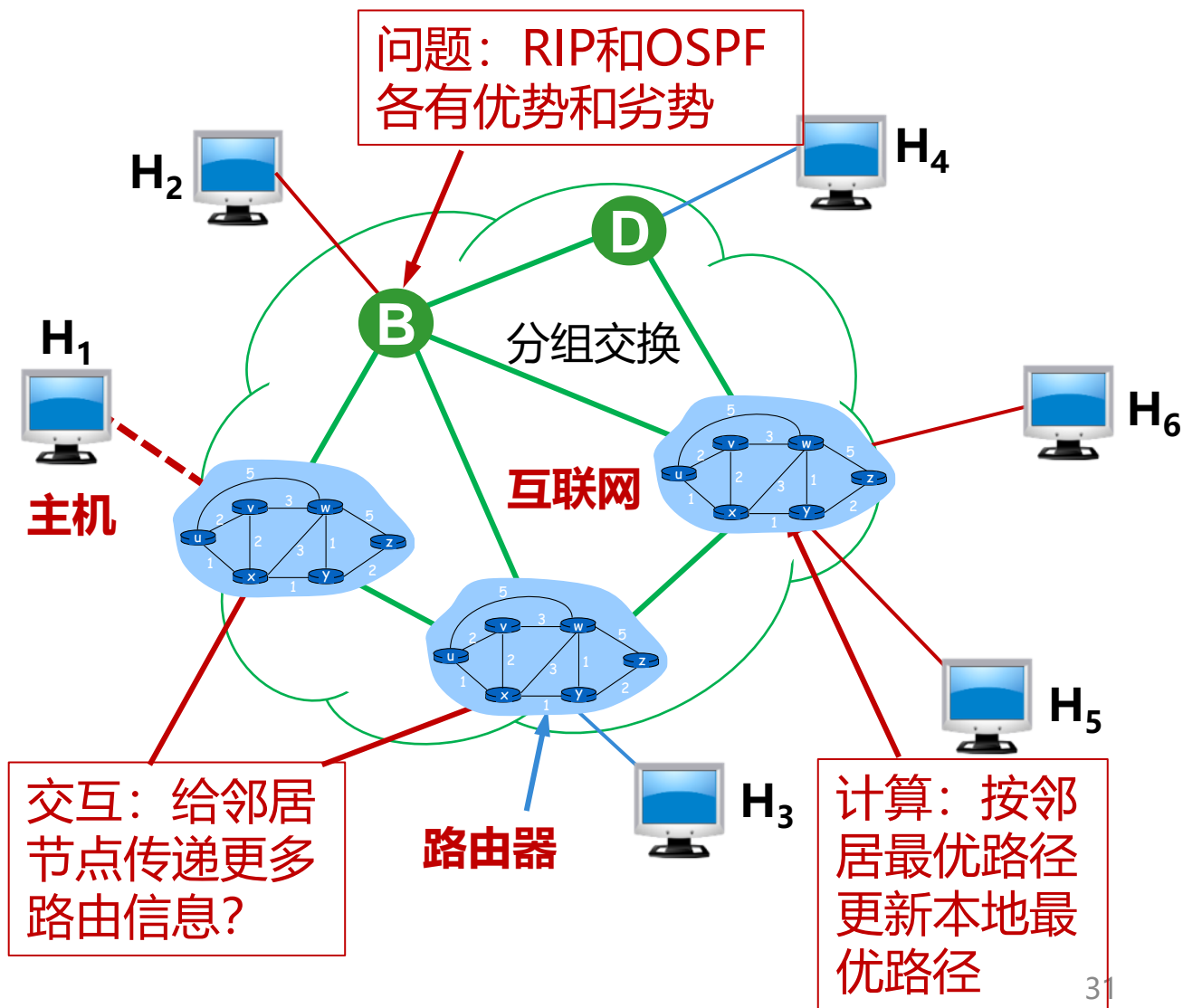


清华大学  
Tsinghua University



计算机网络教案社区

- 全球互联？巨人呢？
  - OSPF独立计算导致计算压力大
  - RIP计算互相帮助但有回路
  - 参考OSPF分层区域概念
  - 可靠传输仅在变化时计算
- 扩展性全球路由设计思路
  - 网络分层：推广区域概念，将网络看成节点
  - 收敛问题：基于距离向量利用更多信息？
  - 无回路的可靠分层RIP？





# 路由器控制平面

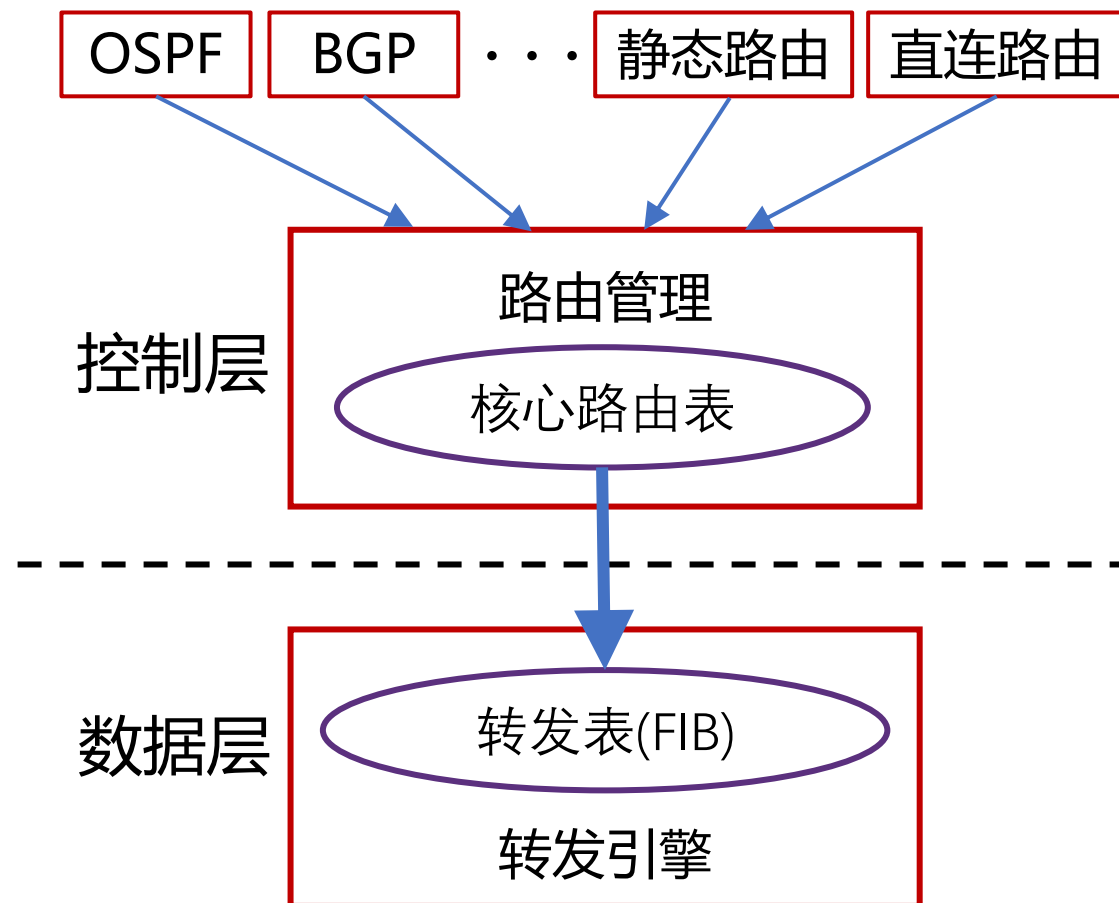


清华大学  
Tsinghua University



计算机网络教案社区

- 路由器可同时运行多个路由协议
- 路由器也可不运行任何路由协议，只使用静态路由和直连路由
- 路由管理根据路由优先级，选择最佳路由，形成核心路由表
- 路由管理将核心路由表的信息，再提供给各个路由协议，实现控制面闭环
- 控制层将核心路由表下发到数据层，形成转发表（FIB）







# 网络层扩展性：NAT到IPv6技术



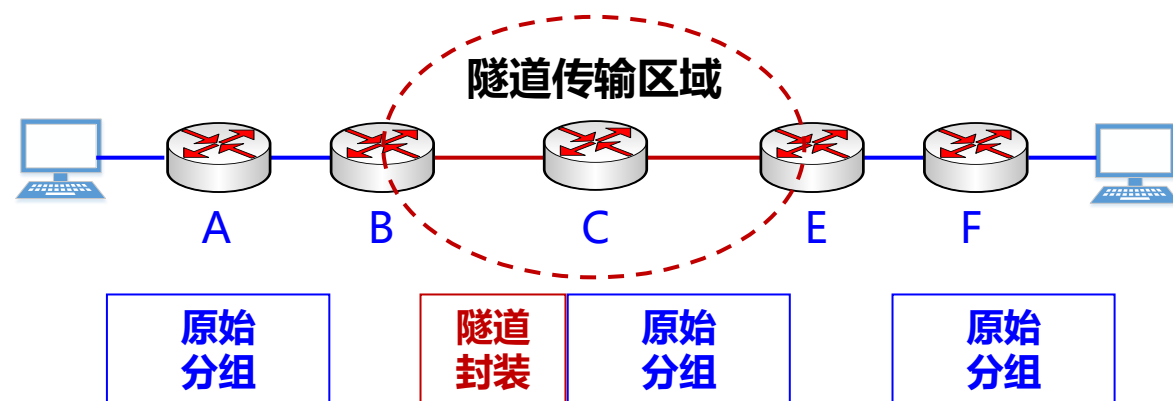
清华大学  
Tsinghua University



计算机网络教案社区

- 相似的出发点：IPv4地址不够用？
  - 在IPv4的框架中优化：NAT技术
  - 拓展地址位数，新设计：IPv6技术
- NAT技术
  - NAT转换表：内网(IP, 端口) <-> 外网(IP, 端口)
- IPv6技术
  - 地址空间：32 -> 128位
  - IP协议族和路由协议的IPv6设计
  - IPv4/IPv6共存：翻译、隧道

网络层基础：IPv6  
动态地址分配：DHCPv6  
单跳网络处理：邻居发现  
基础管理和控制：ICMPv6  
距离向量路由：RIPng  
链路状态路由：OSPFv3  
外部网关路由：MP-BGP





# 网络层服务质量控制



清华大学  
Tsinghua University



计算机网络教案社区

- 如何用上更好的网络服务？
  - 更好？网络服务质量QoS（带宽、时延、抖动、丢包率）
- 网络层QoS
  - 突发流量 <-> 流量整形
    - 限制突发流量产生，避免节点缓存波动导致的丢包
  - 缺少差异化服务 <-> 综合服务、区分服务
    - IntServ逐流维护状态，DiffServ按照每类处理优先级
  - 流量超出带宽 <-> 拥塞控制
    - 流量感知路由：通过调度流量绕开热门区域，疏解流量
    - 端网协同的流量调节：抑制包（显示拥塞通告ECN），逐跳反压机制
    - 网络节点的随机早期检测：未雨绸缪，缓解即将到来的拥塞



# 传输层原理与基础



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 传输服务

- 提供进程到进程的服务
- Berkeley 套接字

最最最  
核心的  
TCP

## ➤ 传输协议的要素

- 寻址、连接建立与释放、差错控制与流量控制、多路复用等

## ➤ 拥塞控制

- 引起拥塞的原因：输入 > 输出持续存在，大量数据在转出前被丢弃
- 解决方法：使发送方以合适的速度发送

## ➤ UDP协议：无连接、不可靠但简单、快速的传输层协议

- 提供了：进程到进程的服务；复用和分用；差错检测
- 未提供：拥塞控制；可靠传输





# 传输协议TCP整体情况



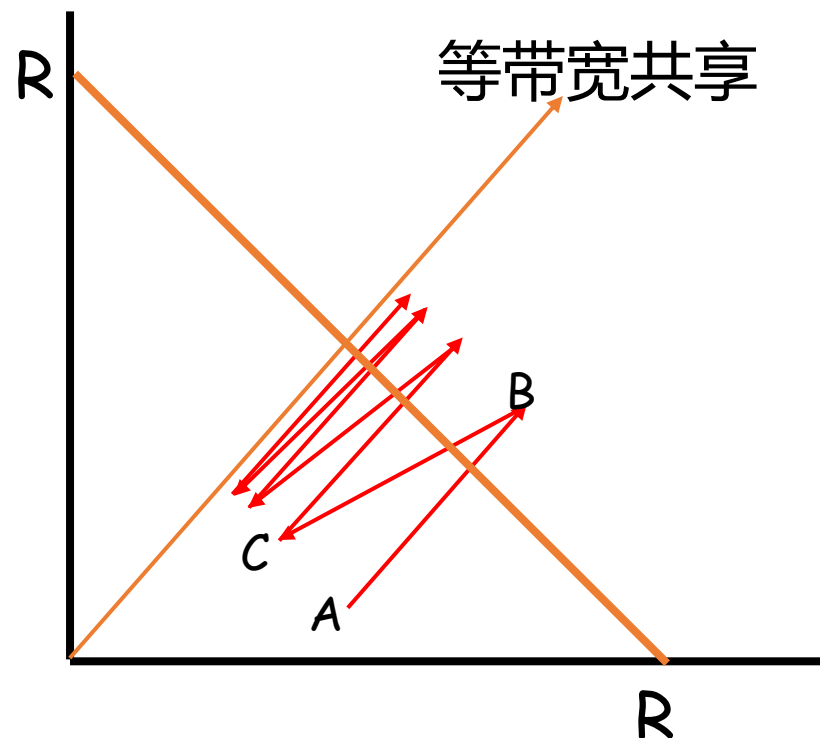
清华大学  
Tsinghua University



计算机网络教案社区

- TCP提供的服务
  - 面向连接的、可靠的、端到端的、基于字节流的传输协议
- 字节流：流入或流出进程的字节序列
  - TCP保证字节流的可靠交付（保序）
- TCP的确认与重传机制
  - 确认时钟
  - 基于ACK的滑动窗口
- TCP拥塞控制
  - AIMD机制的公平性

AI鼓励低速  
MD惩罚高速





# TCP连接管理有限自动机

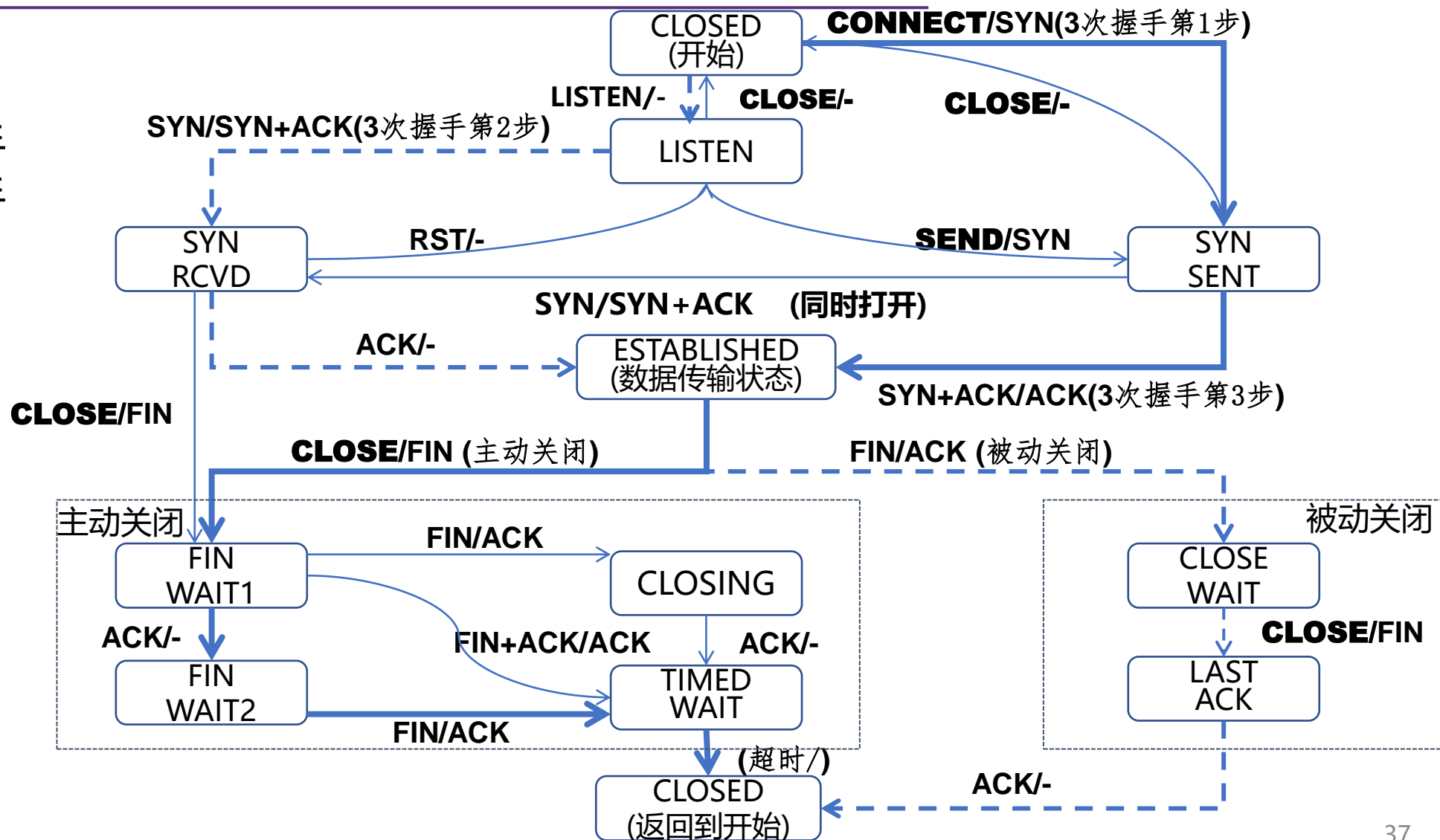


清华大学  
Tsinghua University

 计算机网络教案社区

# TCP连接管理

连接建立：三次握手  
连接释放：四次握手





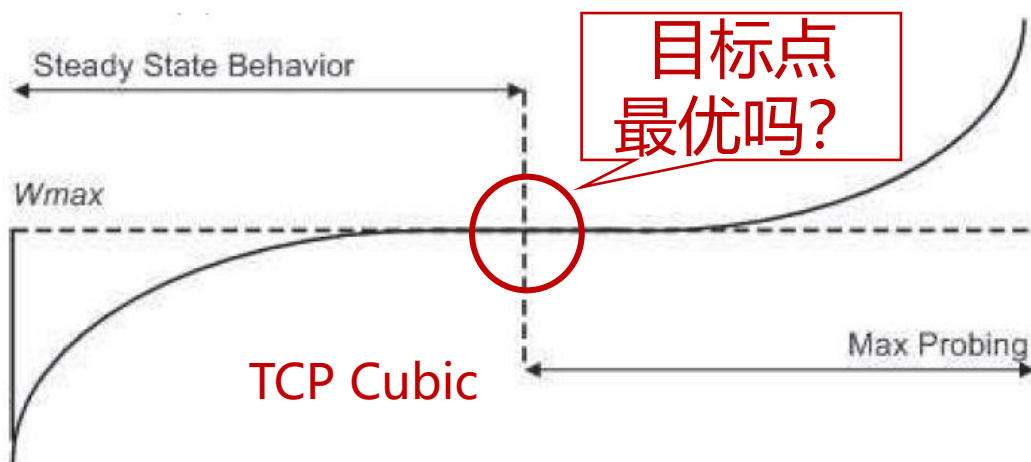
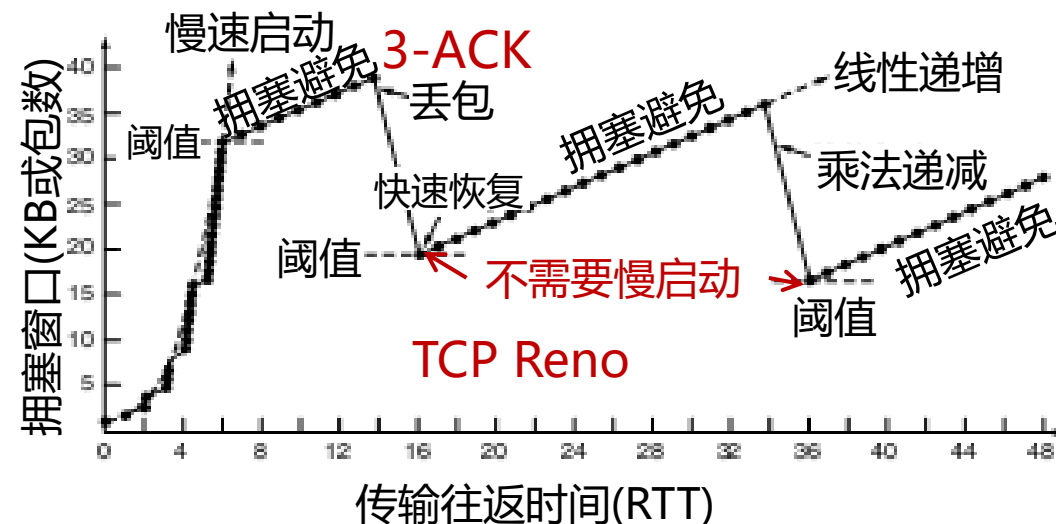
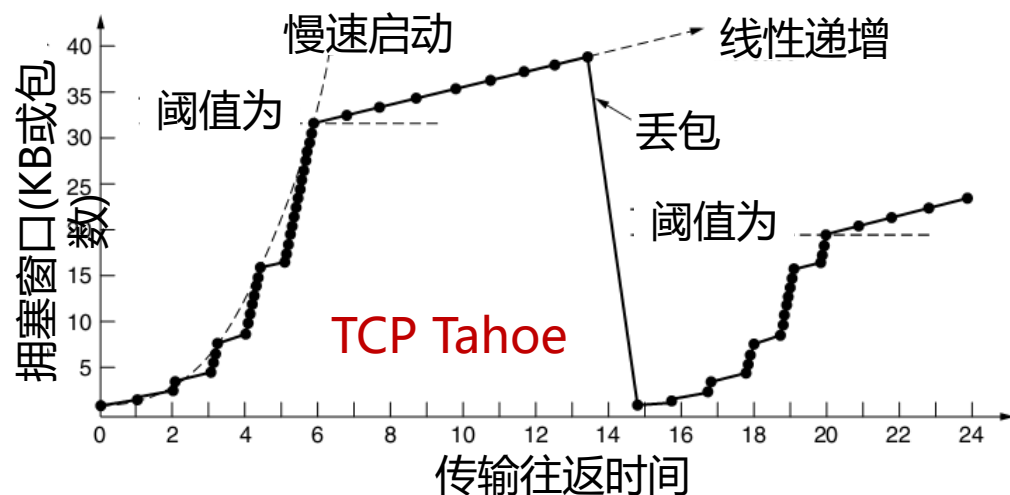
# TCP拥塞控制



清华大学  
Tsinghua University



计算机网络教案社区



如何充分利用带宽?  
快速抵达并保持在拥塞窗口  
适配网络动态性!

如何应对无线网络随机丢失?



# 传输层进阶

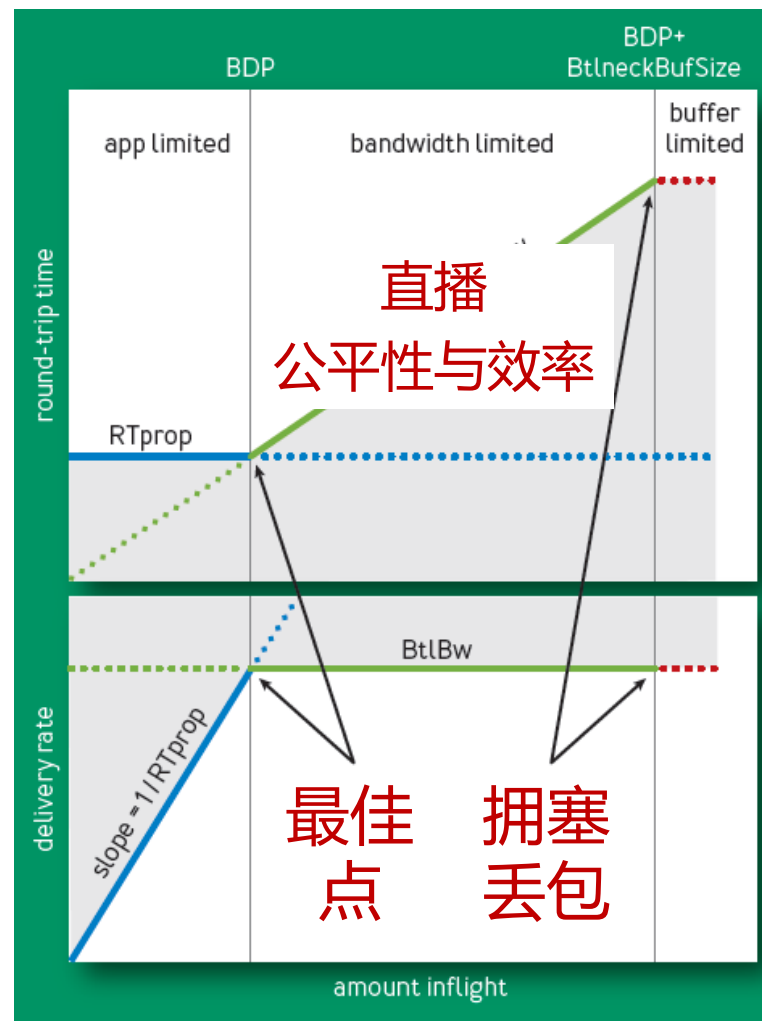


清华大学  
Tsinghua University



计算机网络教案社区

- 拥塞控制算法的发展
  - BBR: 测量瓶颈带宽而忽略丢包
- QUIC协议
  - TCP存在的问题: 优化难度高, 握手时延大, 队头阻塞问题
  - QUIC协议: 基于UDP, 用户态, 建连快, 包序号持续递增, 弱化队头阻塞
- 多径MPTCP
  - 将多径上的多个TCP子流进行协同
- 端网协同的数据中心网络DCTCP





# 应用层目标



清华大学  
Tsinghua University



计算机网络教案社区

1. 掌握应用**进程通信**方式以及**服务进程**工作模式（传输协议选择）
2. 掌握**域名系统DNS**基本原理和工作机制
3. 掌握**电子邮件系统**体系结构及基本工作原理
4. 了解**WWW系统**结构及其应用技术，掌握**HTTP协议**及其工作原理
5. 了解**流媒体**基本概念、数字音视频与编码、流式存储媒体、直播与实时音视频、流媒体动态自适应传输

## 不同应用之间的对比学习

- 应用层协议采用ASCII文本：DNS, Email, http, ...
- 高速缓存作用：DNS、http代理（陈旧问题）
- 安全问题和解决方案





# 应用进程通信方式



- 不同主机上应用进程间的通信和协同
  - 两台主机通信实际是其上对应的两个应用进程(process)在通信
  - **应用进程**: 为解决具体应用问题而彼此通信的进程
  - **每个应用层协议是为了解决某一类应用问题**
- 应用层: 规定应用进程之间在通信时所遵循的协议

客户/服务器 (**C/S**, Client/Server) 方式

浏览器/服务器 (**B/S**, Browser/Server) 方式

对等 (**P2P**, Peer to Peer) 方式



# 层次结构的域名系统及其解析过程

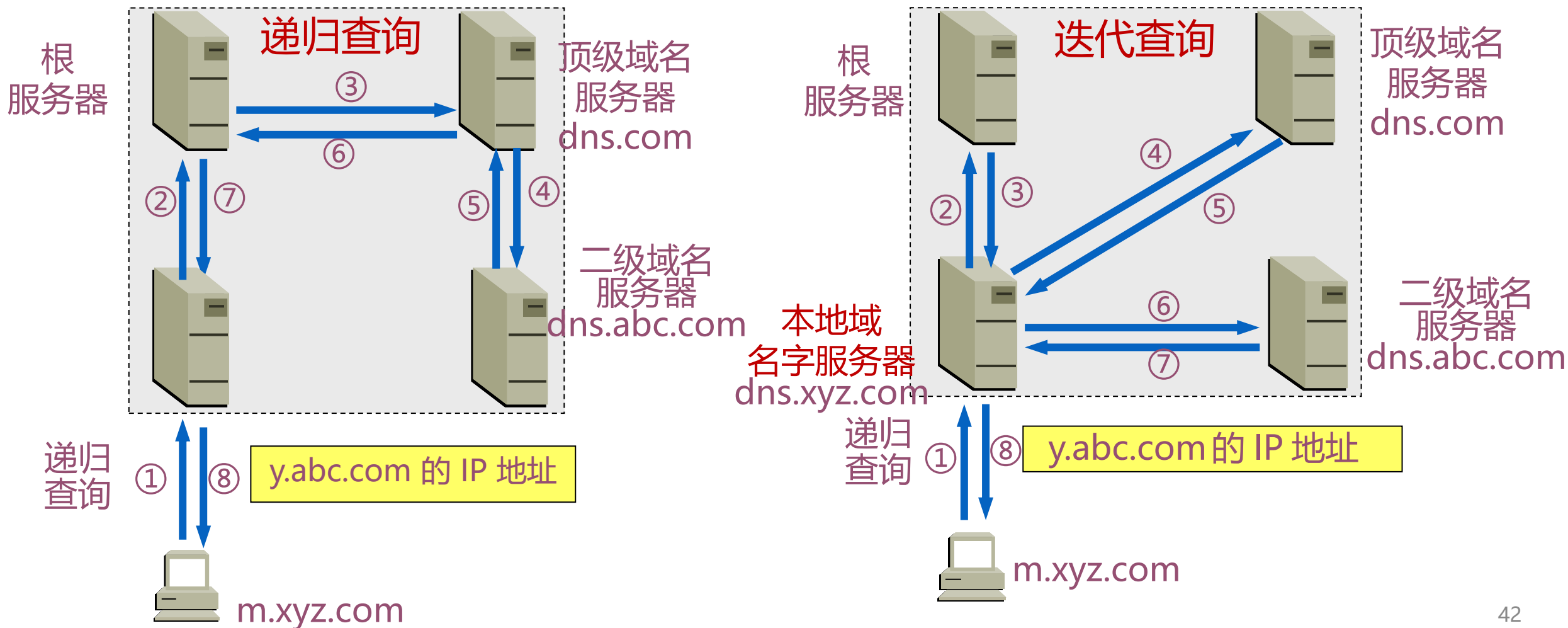


清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 分布式的树状层次结构域名系统（DNS协议：UDP端口号53）





# 邮件传递过程使用的协议

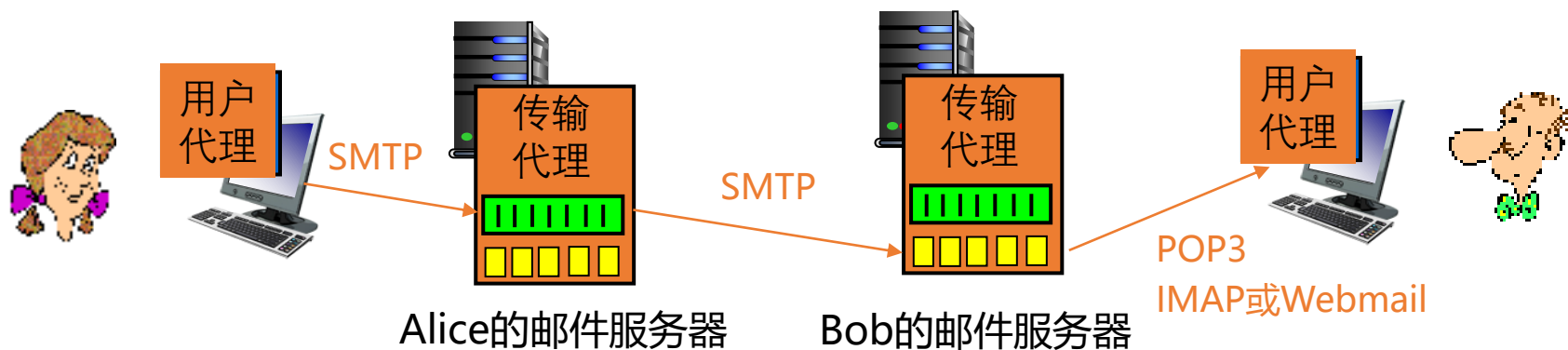


清华大学  
Tsinghua University



计算机网络教案社区

- 用户采用SMTP发送邮件
- 最终交付（邮件访问）协议: 从邮件服务器的邮箱中获取邮件
  - POP3: Post Office Protocol-Version 3, 第三版邮局协议
  - IMAP: Internet Message Access Protocol, Internet邮件访问协议
  - Webmail (HTTP) : 基于Web的电子邮件
- 收发邮件的核心区别是什么?





# 万维网与HTTP



清华大学  
Tsinghua University



计算机网络教案社区

## ➤ 统一资源定位器URLs (Uniform Resource Locators)

➤ 例如: <http://www.phdcomics.com:8000/comics.php>

协议类型

主机名即服务器

端口

路径和文件名

## ➤ HTML文档与静态网页

## ➤ 超文本传输协议HTTP (HyperText Transfer Protocol)

➤ CS模式: 请求与响应报文, 使用TCP协议, 缺省为80端口

➤ **HTTP为无状态协议**, 服务器端不保留之前请求的状态信息, 如何跟踪购物车里的内容?

➤ 非持久连接和持久连接



# 流式音频和视频



清华大学  
Tsinghua University



计算机网络教案社区

- 流媒体概述：点播、直播和RTC交互式
- 数字音视频与编码
  - 图像组(GOP)与IBP帧 (I帧越密越好? )
- 流式存储媒体
  - 如何应对网络传输的抖动特性? HTTP
- 直播与实时音视频：RTSP ( TCP/UDP )
- 流媒体动态自适应传输DASH
  - 自适应码率选择ABR
  - 可扩展视频编码 SVC

前后结合  
对比学习

# 天行健，君子以复习为乐

崔 勇

清华大学·计算机系



计算机网络  
教案社区