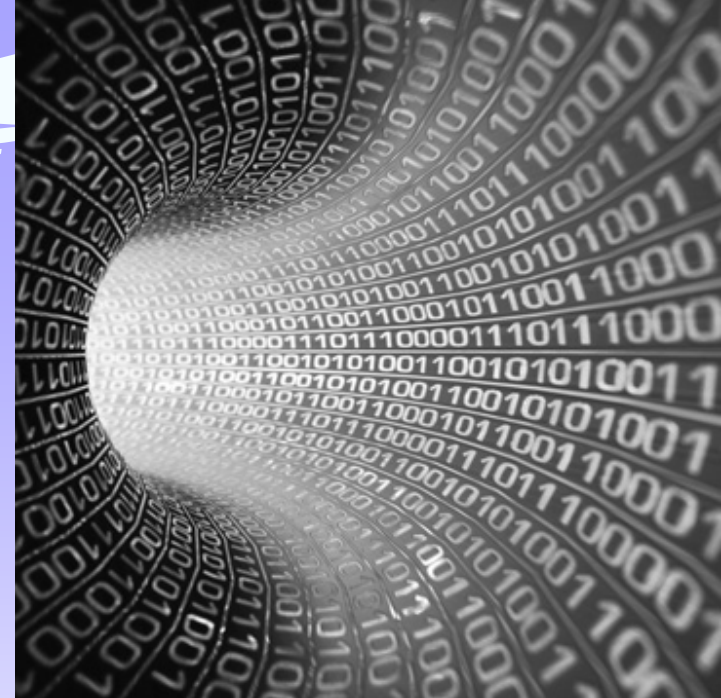


现代计算机 网络理论与技术



Sept. /2010

李芝荣 华中科技大学计算机学院

CERNET华中地区网络中心

课程目标/结构/计划—2010版(32H)

- **梳理本科网络的基本知识**
- **掌握现代网络的理论与技术**
- ◆ **Ch1: 体系结构-4H**
 - 基本概念
 - 传输网基本理论与技术
 - 计算机网络基本理论与技术
- ◆ **Ch2: 拥塞控制-4H**
 - 集成服务与区分服务
 - 拥塞控制
 - 主动队列管理AQM
- ◆ **Ch3: 无线与移动网络-4H**
 - 无线网络
 - 移动网络
- **了解网络问题及研究方法**
- **熟练网络编程方法**
- ◆ **Ch4: P2P网络-6H**
 - 无结构P2P网络
 - 有结构P2P网络
- ◆ **Ch5: 流媒体网络-4H**
 - 音视频压缩原理
 - 实时交互协议RTP/SIP
 - 流媒体分发与监管
- ◆ **Ch6: 下一代互联网-2H**
 - IPv6与CNGI
 - 物联网
- ◆ **Ch7: 网络建模与分析-8H**
 - 排队论
 - 以太网建模分析
 - 网络仿真工具

教学规则与资源

◆ 总成绩

- 总成绩 = 实验**20%** + 考试**80%**
- 考试 = 对投机者说“**不!**”

◆ 学时安排：总计**32**学时/**2**个学分

◆ 资源地址

- 本课件下载地址：<http://ztli.hust.edu.cn>
- 实验服务器地址：**202.224.0.216**

◆ 天将降大任于斯人也，必先苦其心志，劳其筋骨，饿其体肤...

参考书目

1. Computer Networking—A Top-Down Approach Featuring the Internet, Third Edition , James F. Kurose,
2. Computer Network—A System Approach (2) , Larry L, 机械工业出版社, 2000。后续各章作业是此书的作业
3. High Performance Communication Network(2), Jean Walrand. 机械工业出版社, 2000.
4. A. S Computer Networks, Tanenbaum(3) 《计算机网络》 (3) , 熊桂喜等译, 清华大学出版社, 1997.
5. High Speed Networks and Internets Performance and Quality of Service(2), 高速网络与互联网—性能与质量服务 (2) , 电子工业出版社, 2003.
6. 计算机通信与网络教程, 林生, 清华出版.
7. 对等网络: 结构、应用与设计, 陈贵海等, 清华大学出版社, 2007. 9
8. 其它会议和杂志有关网络的论文

Ch1:网络体系结构

1.1 基本问题与概念

- 计算机网络的演进!
- 计算机网络的定义?
- 计算机网络的需求与约束!

1.2 计算机网络的体系结构

- 基本元素
- 基本结构

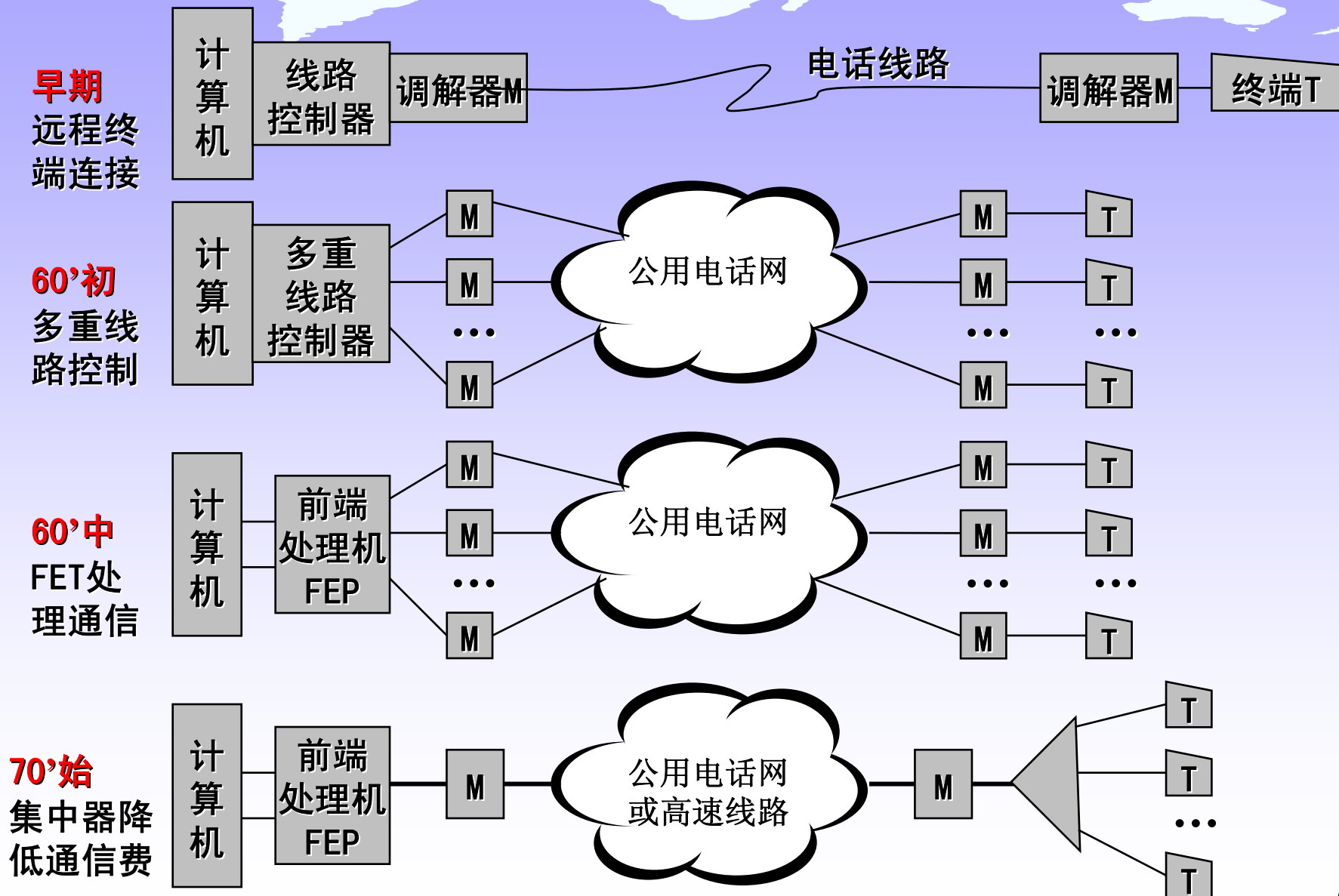
1.3 网络性能测量与评价

- 两个基本指标
- 网络性能指标体系

1.4 怎样操作网络

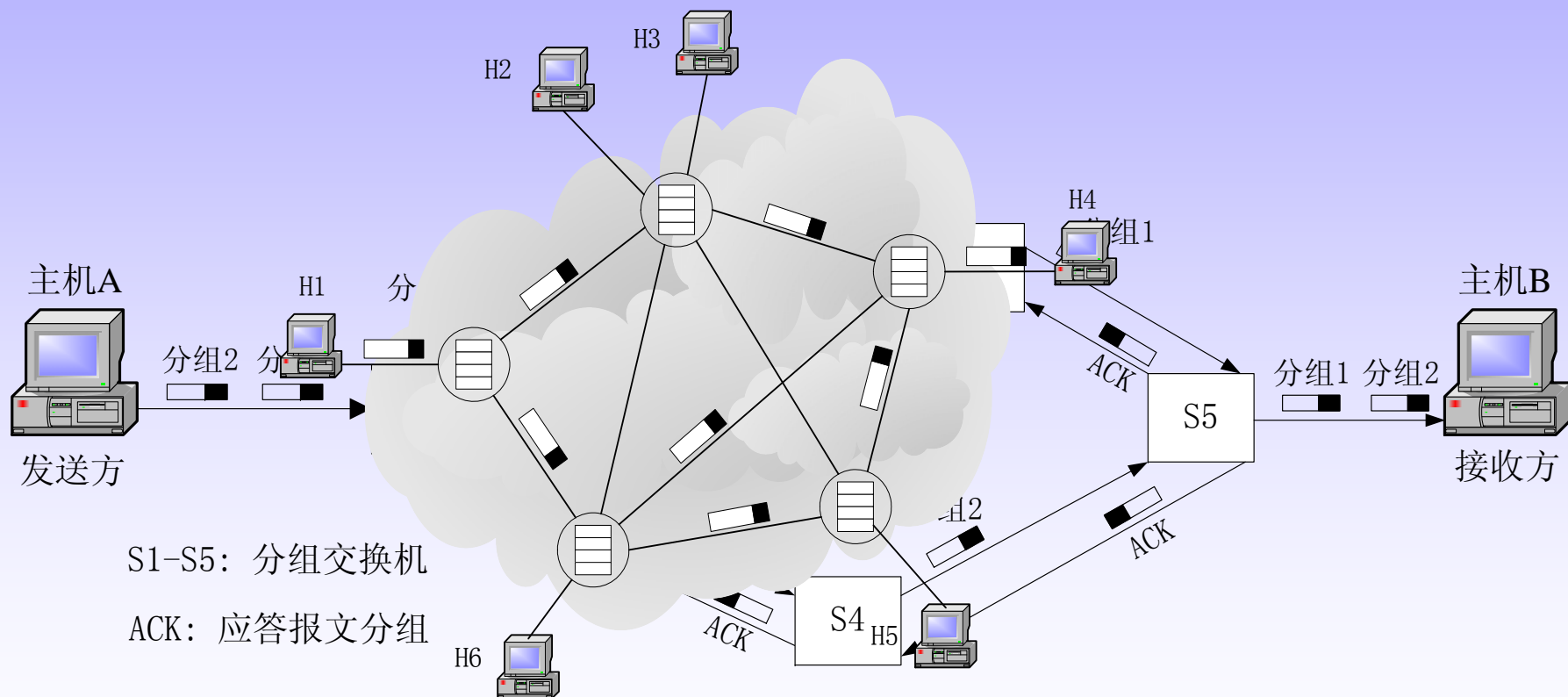
- API接口与Socket编程

计算机网络的演进



◆ 62-65-72' 分组交换--DARPA (Datagram 方式)

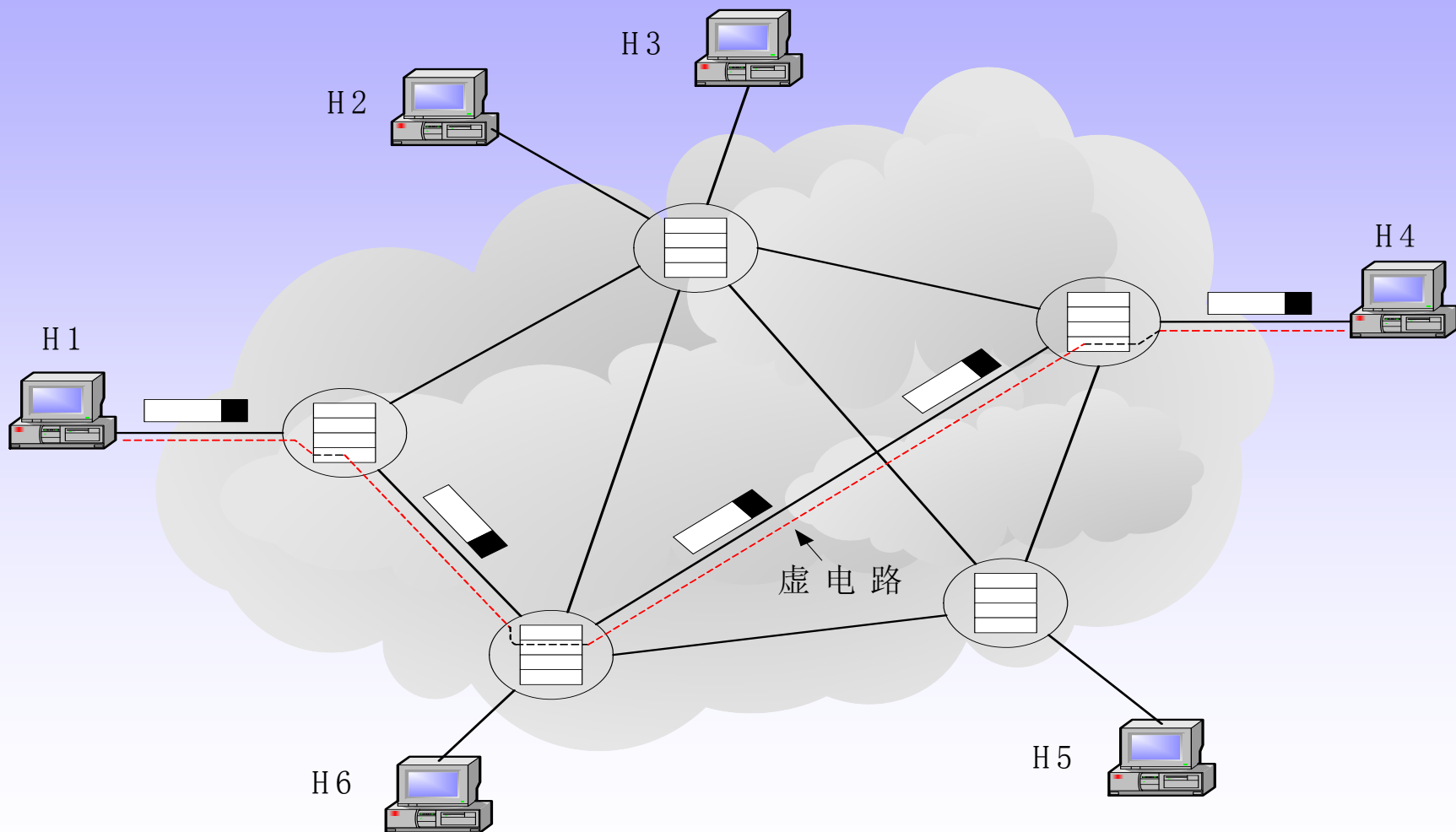
- 各分组**独立地确定路由** (传输路径)
- **不能保证分组按序到达**, 所以目的站点需要按分组编号重新排序和组装



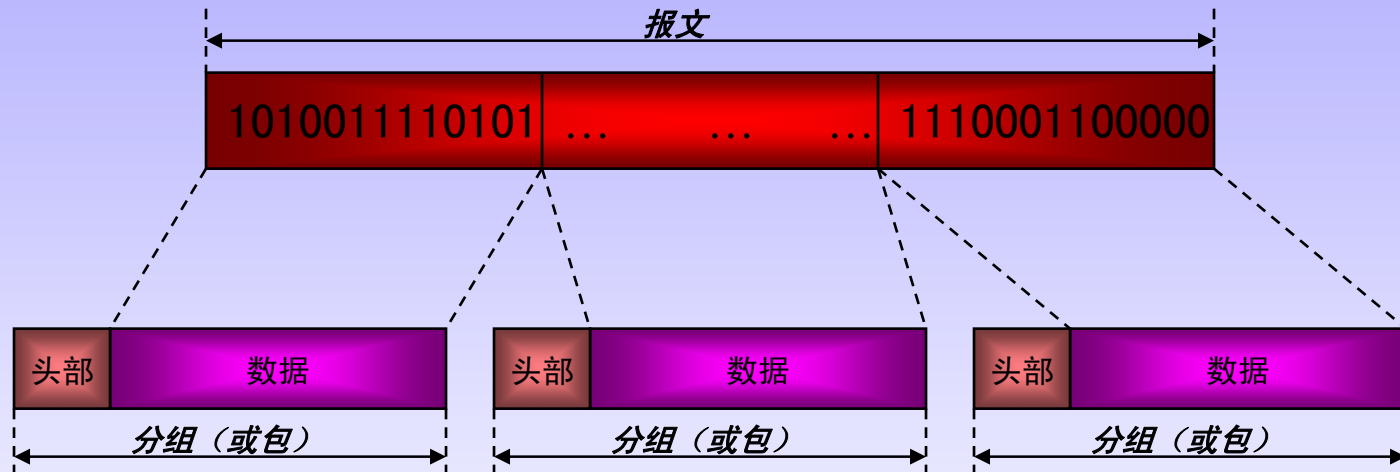
分组数据通过多条路径穿越网络, 不能保证按序到达

◆ 虚电路方式

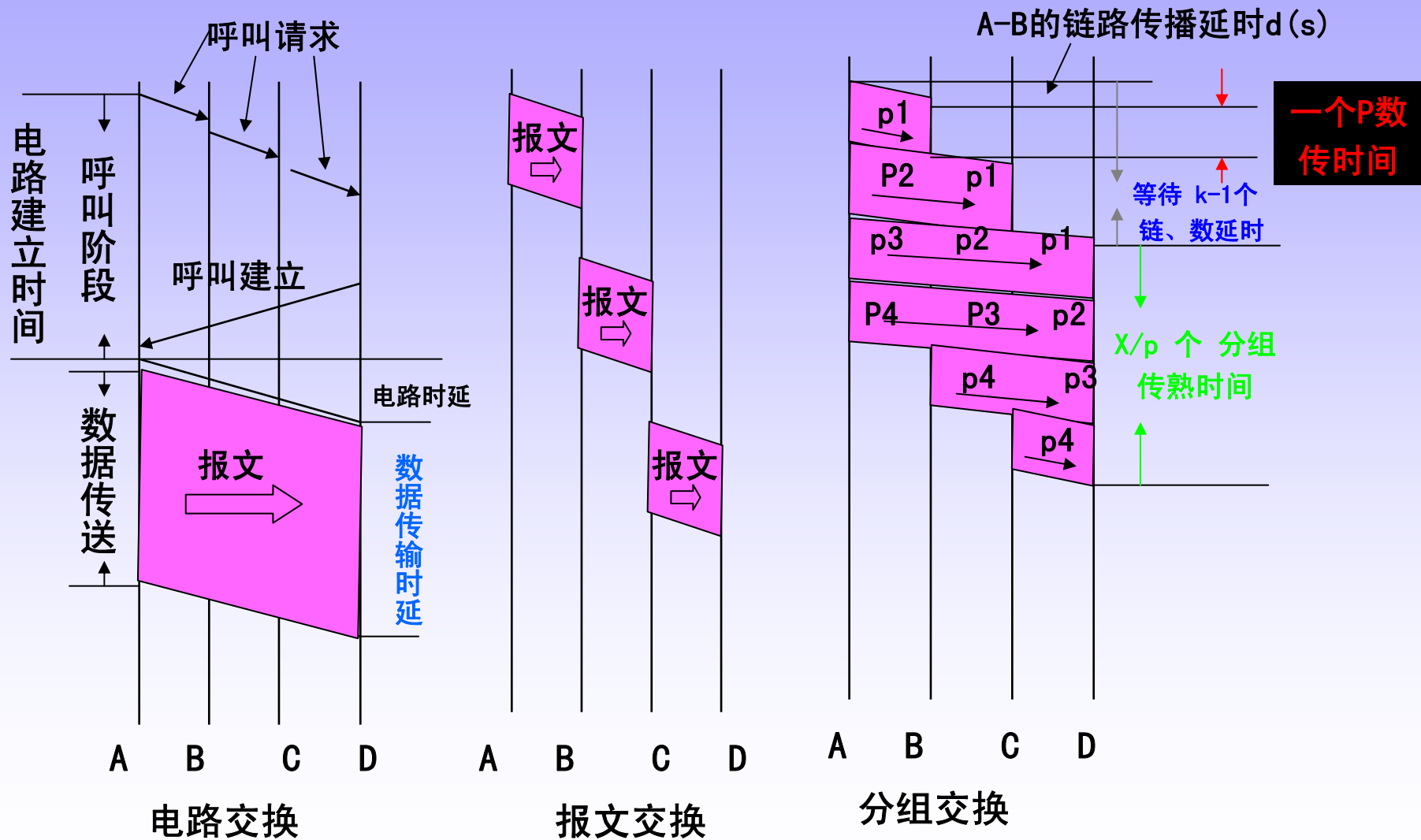
➤ 分组通过预先建立好的虚电路穿越网络



分组交换的概念



电路/报文/分组交换的比较



网络的定义

◆ 互连自治计算机的集合！

➤ 互联：

☞ 经媒质交换数据的能力

➤ 自治：

☞ 面向用户自我管理

☞ 计算机、位置可见

◆ 分布式系统

➤ A可强迫启/停/控B

➤ 面向系统处理/存储均不可见

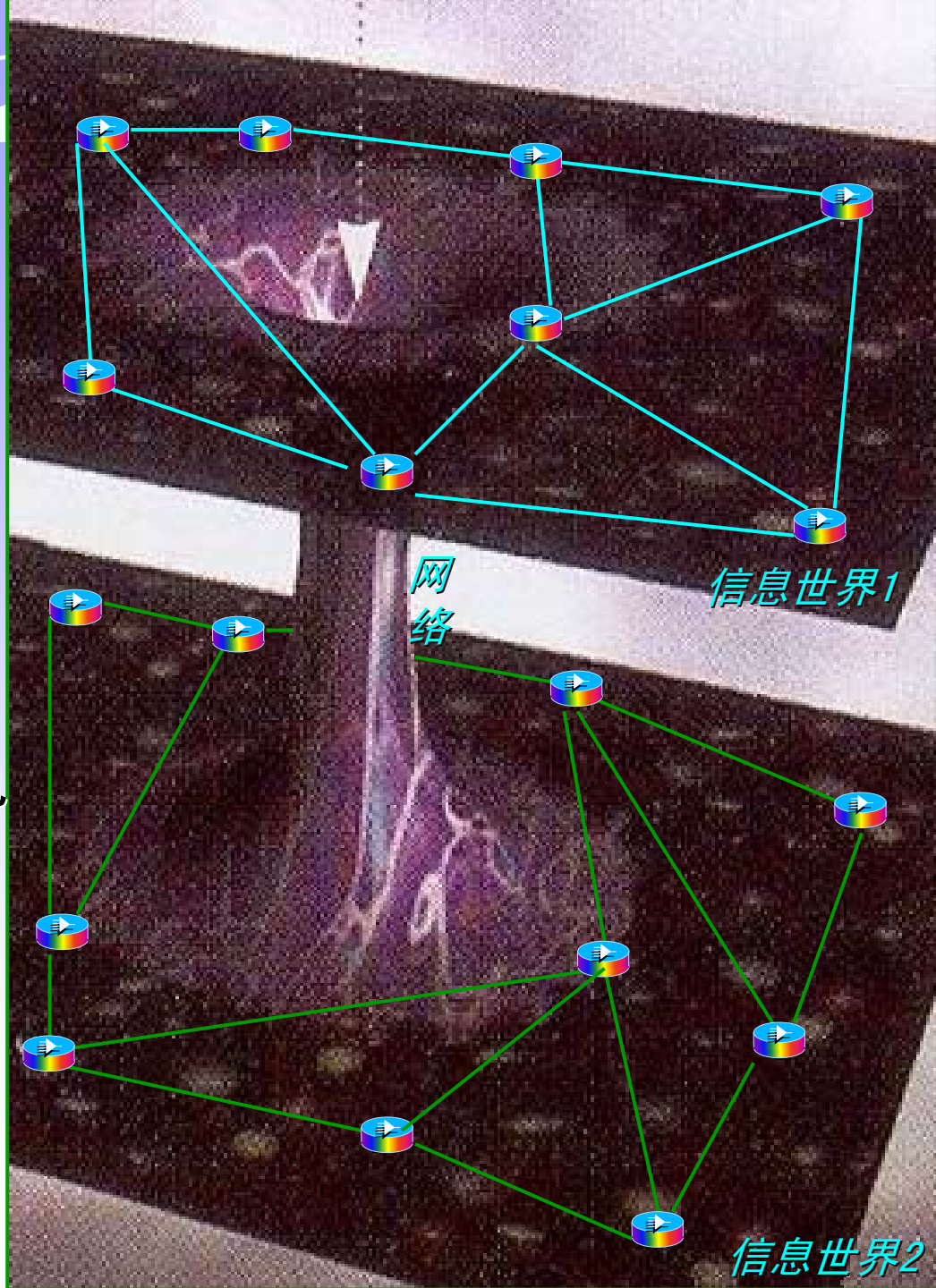
◆ Network is computing!

➤ Computing

➤ Storing

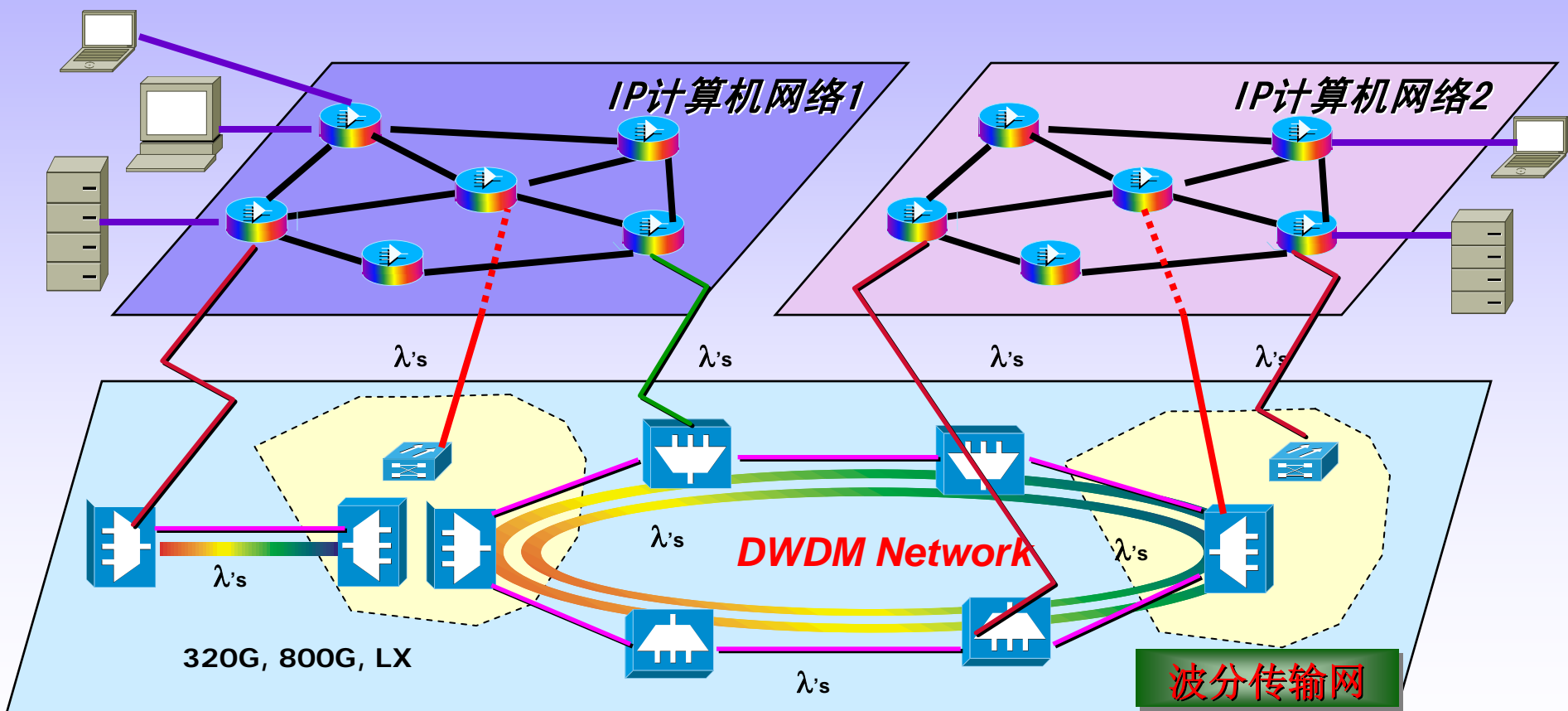
➤ Transforming

◆ 网络空间已经成为世界的第五大战场！

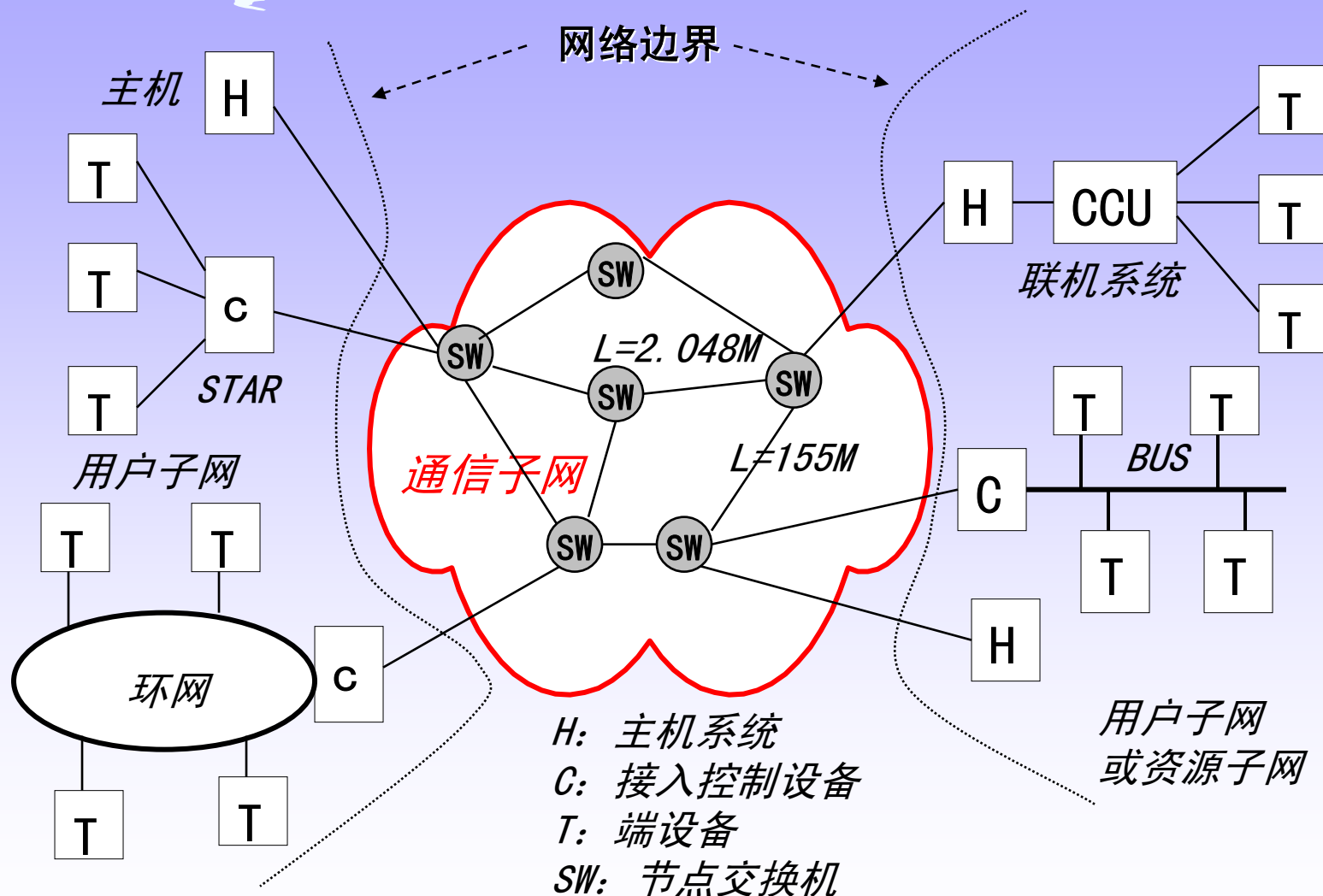


◆网络构成

- 物理：传输网+计算机网
- 逻辑：主干通信+接入资源



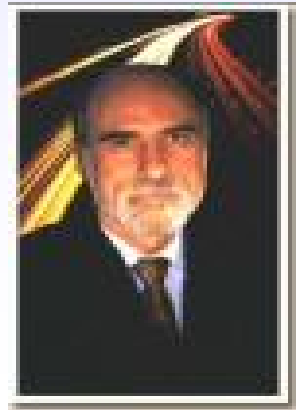
逻辑构成：通信+资源



What is the Internet?

- ◆ The largest network of networks in the world.
- ◆ Uses TCP/IP protocols and packet switching .
- ◆ Runs on any communications substrate.

From Dr. Vinton Cerf,
Co-Creator of TCP/IP



计算机网络的需求与约束

◆ Constraints & Requirements

◆ Application Programmer

- 服务与应用需求
- 如某段时间内需要每个消息无错传输

◆ Network Designer

- 设计的成本—效益特点
- 对不同用户有效分配资源和资源利用率

◆ Network Provider

- 易于管理的系统特征
- 可靠运营/容易隔离故障/计费

◆ Network Supervisor

- 来自检测、监控、保障的需求
- 保障国家或组织的信息安全

- 网络的**设计者**、设备的**制造者**、网络**运营者**和网络**监管者**
4维共同作用于网络

1.2 网络体系结构

网络体系 = 元素+结构

基本元素

◆ 硬件构成元素

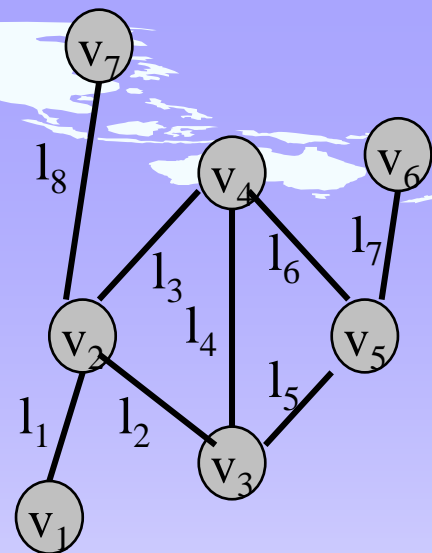
- 点到点链路：光缆、电缆、无线电波
- 通信连接设备：波分传输、路由器
- 端接入设备：路由器、交换机、集线器
- 端设备：计算机、打印机、存储器...

◆ 软件构成元素

- 1—3层（物理、链路和网络IP）的通信协议
- 4—5层（运输、应用）的通信协议
- 各层活动的**进程**（各种系统或应用软件）

网络的拓扑结构

结构 = 元素间相互关系
= **物理**连接+**协议**联接



◆ 网络 = {节点, 链路} = {V, L}

◆ $V = \{v_1, v_2, \dots, v_n\}$, $L = \{l_1, l_2, \dots, l_m\}$

◆ 节点 (Node) :

- 端节点: 自身拥有计算机资源的源宿用户设备
- 转接节点: 支持网络连通并在网络中对数据起交换和转接的节点, 如交换机/路由器/集线器

◆ 链路 (Link) :

- 物理链路: 在物理层连接两个节点的物理介质, 如电话线、同轴电缆、光纤、无线电波

链路、结点、协议和网络云

- 逻辑链路：在两点间通过通信协议的作用建立起来的数据联结通路

◆ 物理连接

- 点一点连接：连接一对结点的物理通路
- 多点连接：多个结点共享一条物理链路

◆ 通路 (Path) :

- 从源点到宿点所经过的一串节点和链路的有序集。或端到端的通路

◆ 协议 (Protocol)

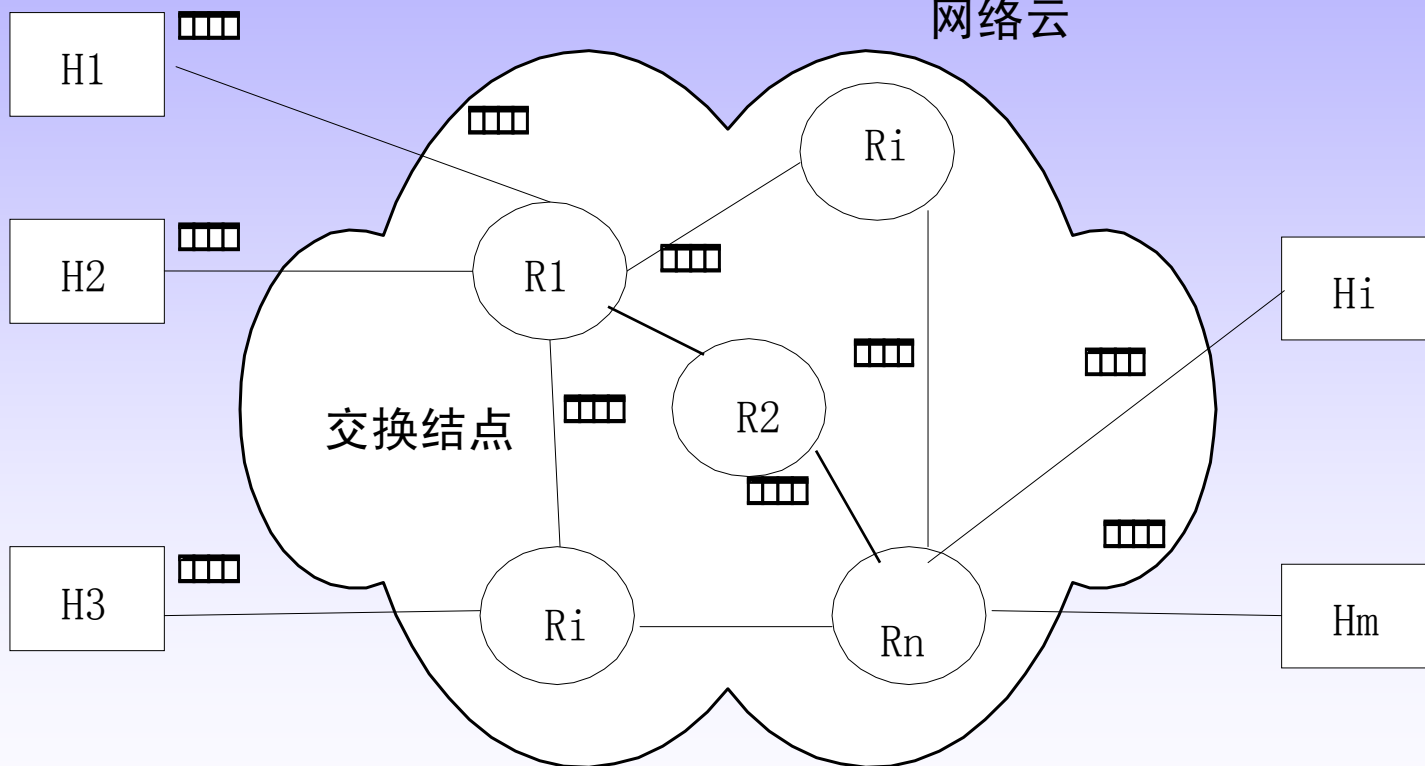
- 多个进程为完成一个任务而共同遵守的动作序列规范
- 三要素：语法、语义、规则(同步时序)

◆ 网络云：

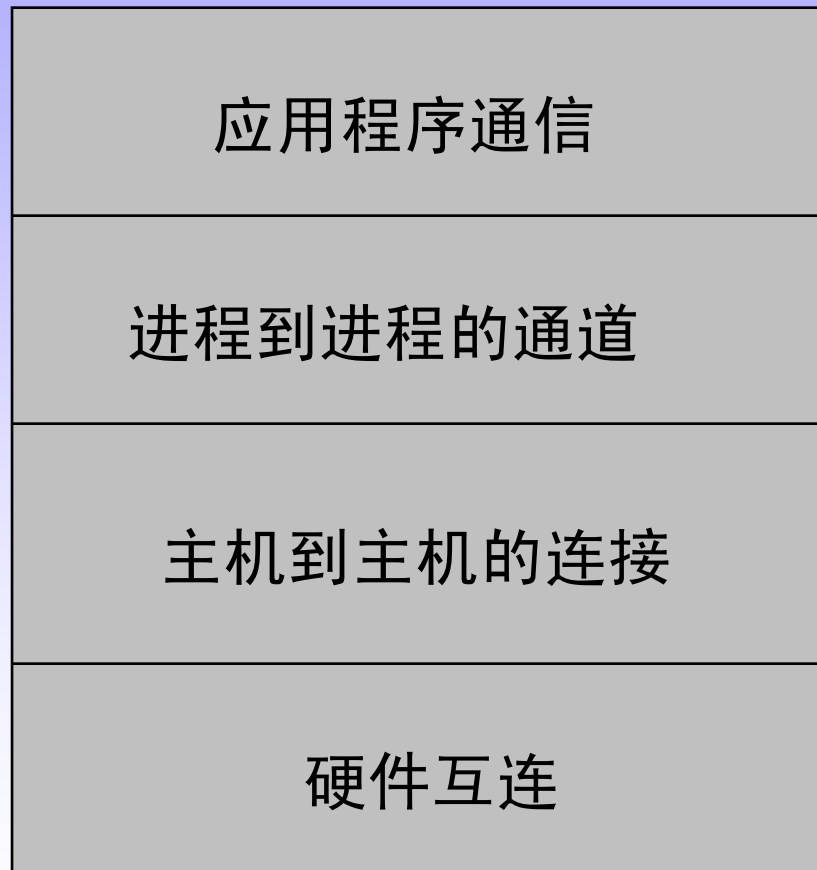
- 区分通信子网交换交换结点（云内）和资源主机结点（云外）的分界线，可表示任何网络（单、多、交换等）

物理网络的抽象表示

主机结点



网络分层结构的通信功能



分层网络系统结构

分层和协议

- ◆ 分层系统在给定层上的功能
- ◆ **协议**：两个或两个以上的参与方为完成特定的任务而采取的一系列确定的步骤。



因特网的大三层结构

◆ 组织性网络：用户级网络：**Stub**

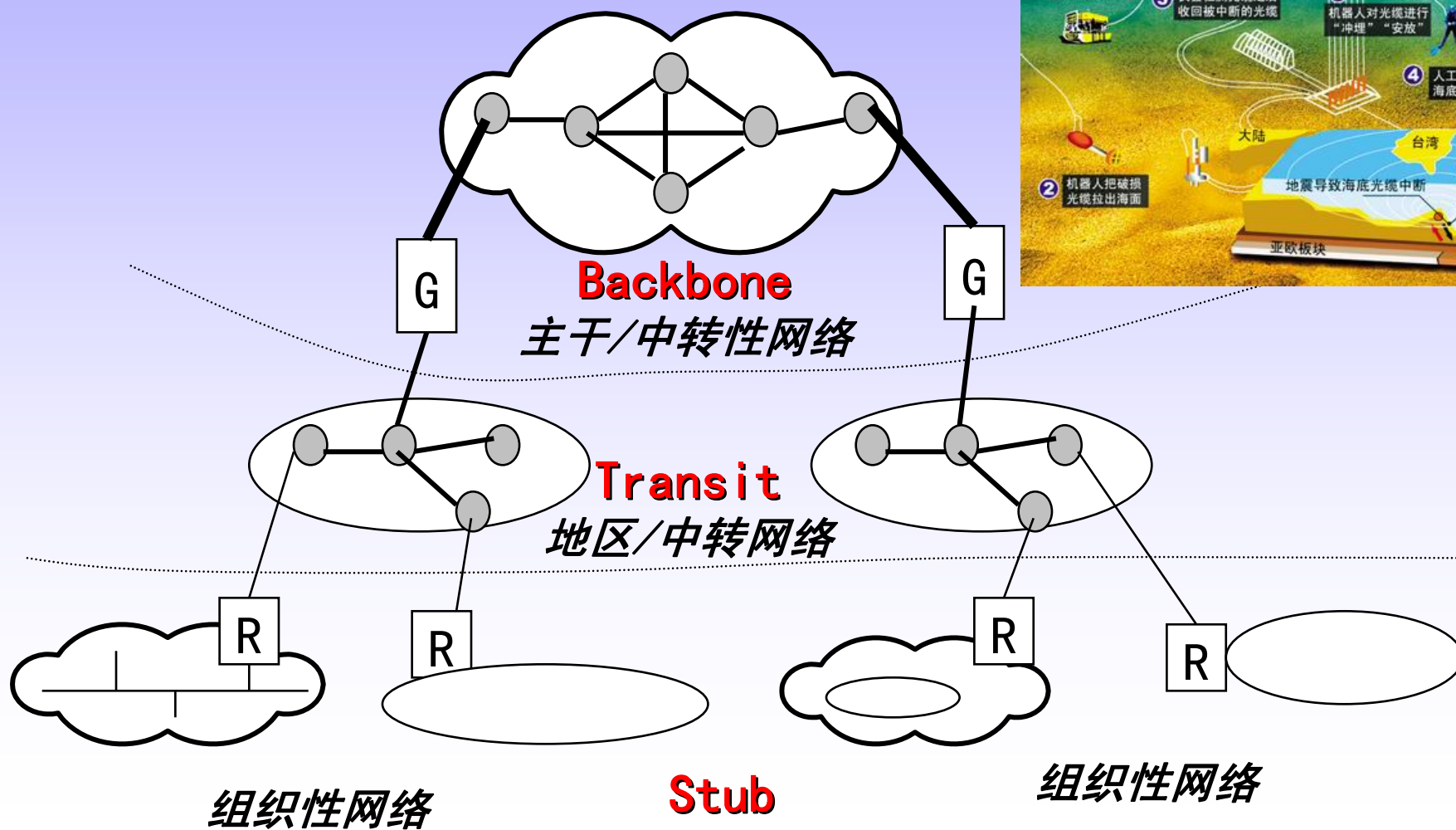
- 是一个自治网络系统，有AS号
- 校园网、企业网、公司网、政府网
- 由外围网关接入地区性网络

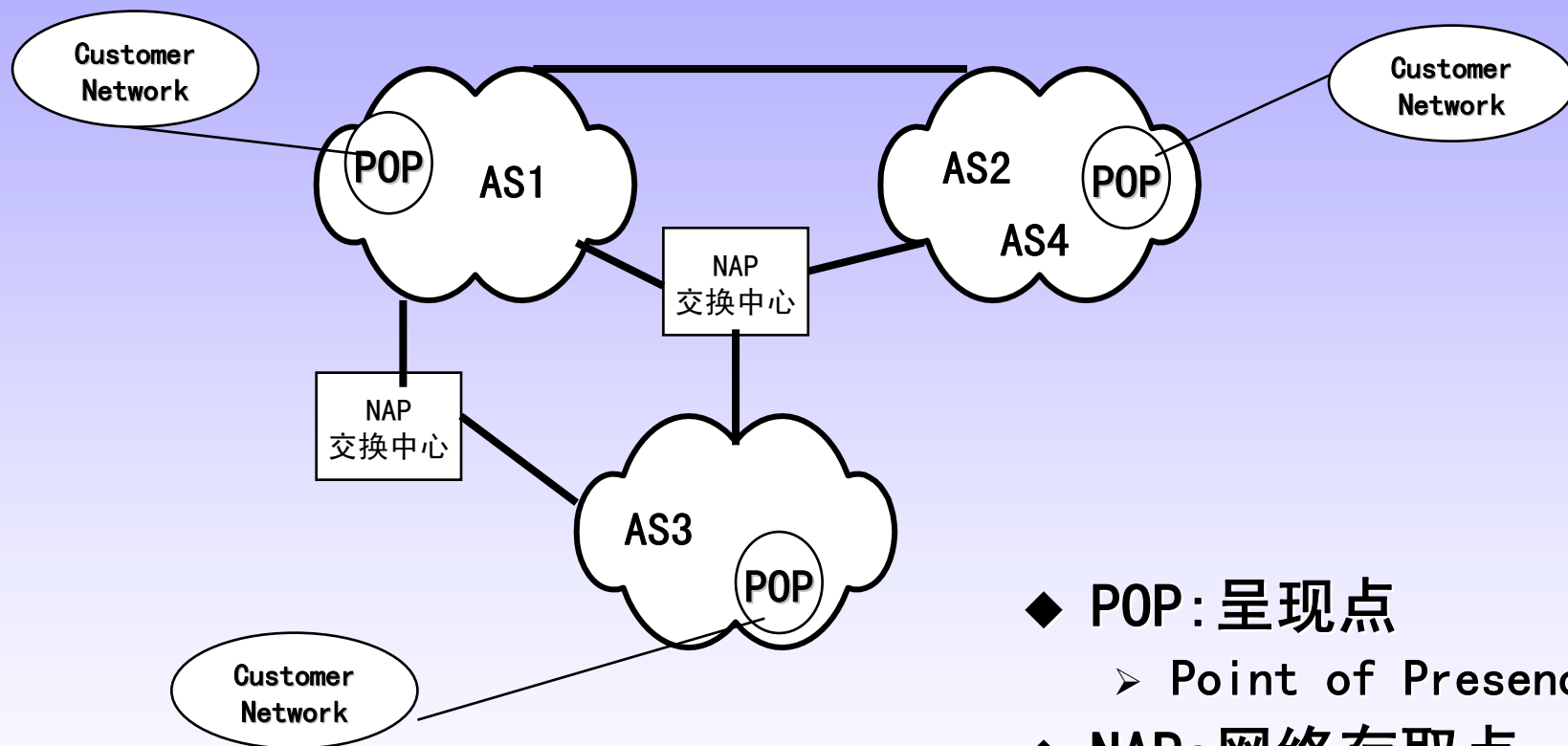
◆ 地区性网络：覆盖省、洲、国家，**Transit**

- 提供到因特网的连接
- 因特网服务提供者ISP

◆ 中转性网络：因特网主干网，**Backbone**

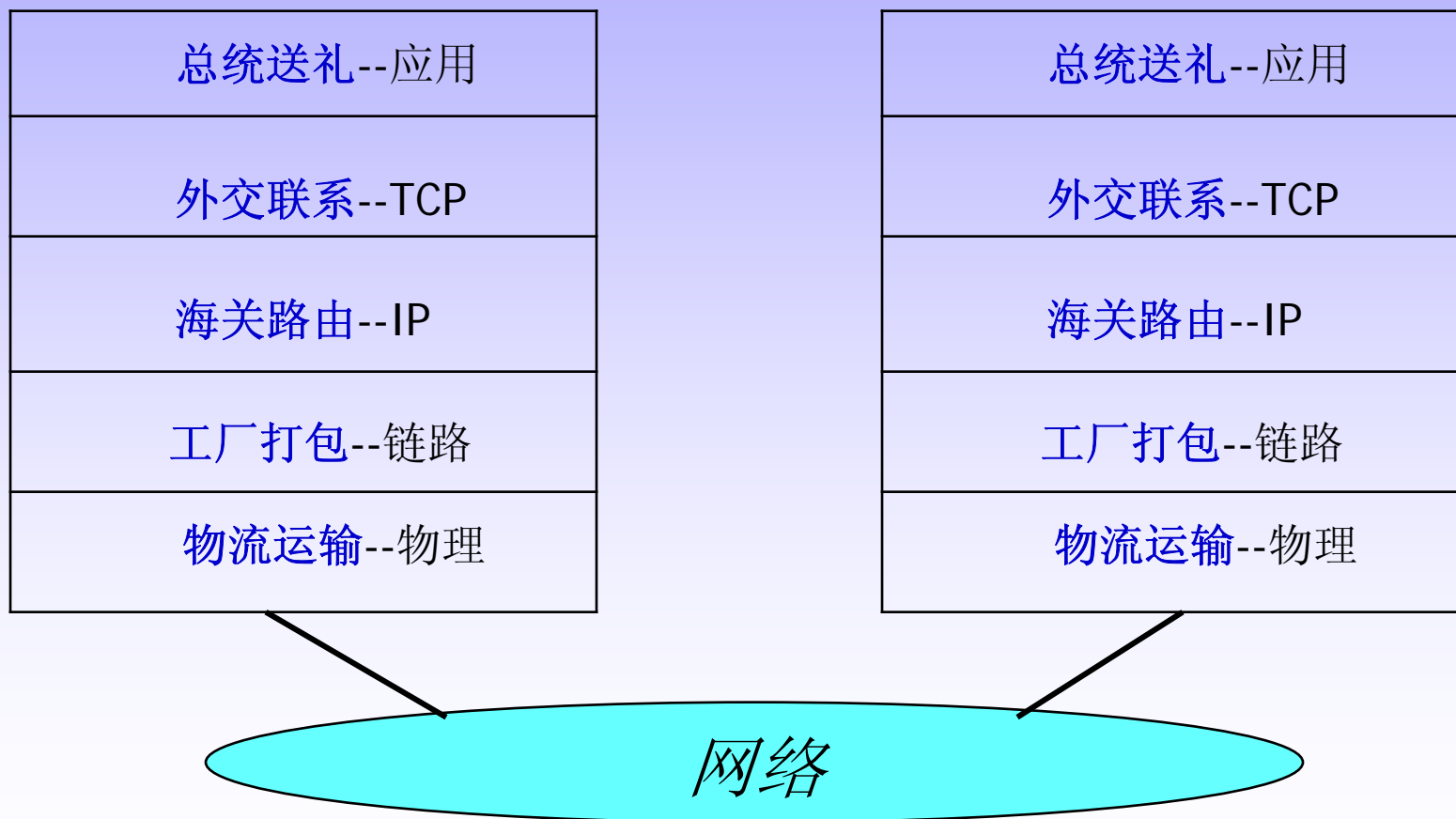
- 美NFSnet/欧Ebone/中Chinanet/CERNET





- ◆ POP: 呈现点
 - Point of Presence
- ◆ NAP: 网络存取点
 - Network Access Point
- ◆ AS: 自治系统
 - Autonomous Systems

网络体系结构：层和协议的集合



1.3 网络性能测量与评价

◆ First get it right and then make it fast——**先对后快!**

◆ 性能测量的两个参数：带宽和延迟

➤ 带宽 (Bandwidth) : Hz, KHz, MHz, GHz, bps

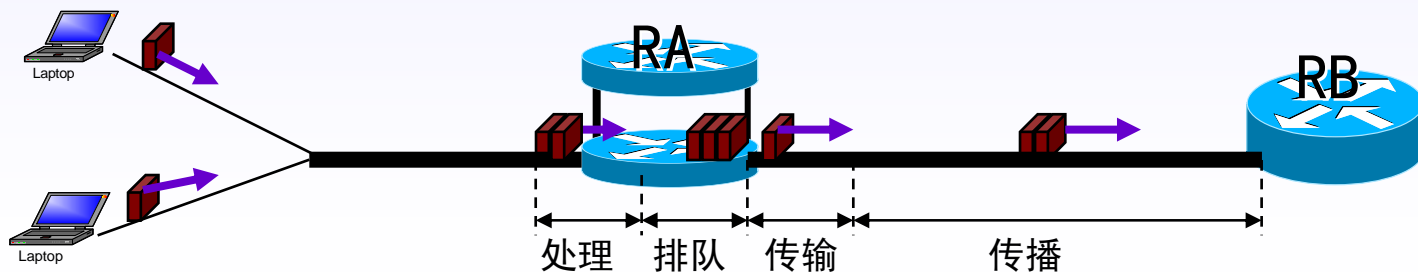
- ☞ 信号带宽：构成一信号的各种**不同频率成分所占据的频率范围**。
如人类声音带宽为：3300Hz—300=声音带宽3000Hz
- ☞ 媒体带宽：通信媒体允许通过的**信号频带范围**
- ☞ 比特率：**某时段内网络上可能传输的比特数**，或传输每比特数据所需的时间宽度。习惯把“带宽”作为数字信道的数据率/比特率
- ☞ 比特率越高，高频分量越多，频率范围越大，信号带宽越高。

➤ 吞吐率 (Throughput) : bps

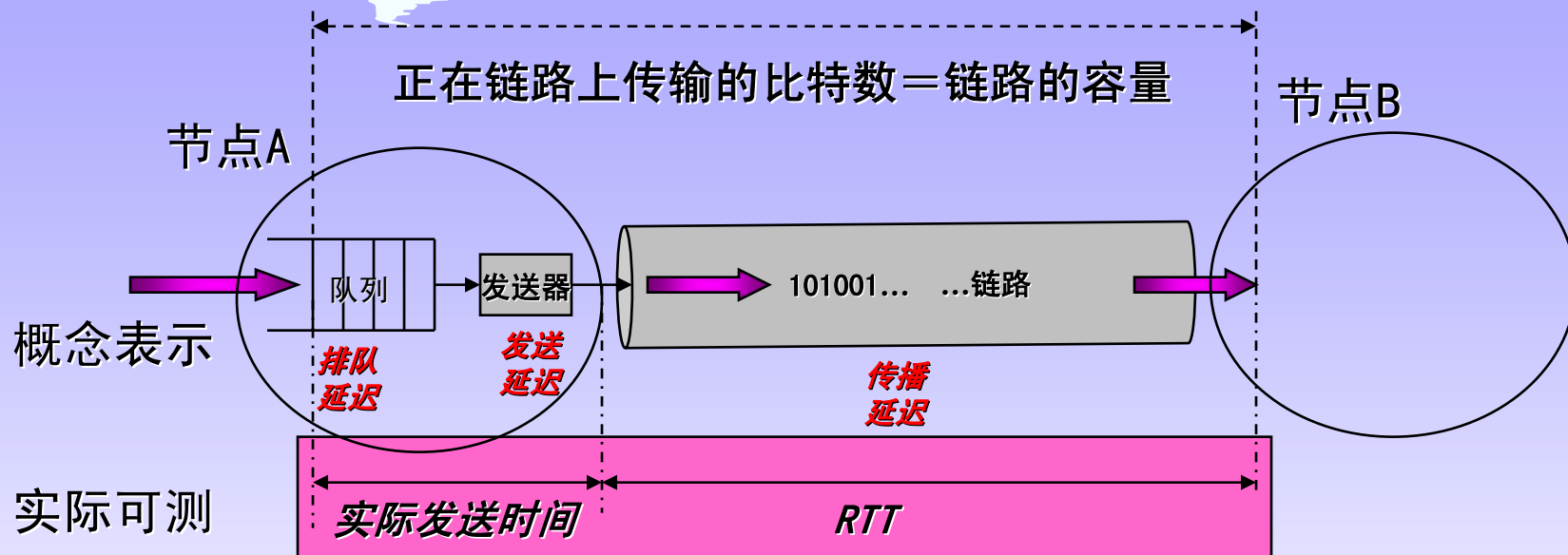
- ☞ 数字信号的发送速率，因此发送带宽也成为吞吐率
- ☞ 由于各种影响10M带宽实际完成2M
- ☞ 吞吐率：链路上**实际每秒传输的比特数**

延时

- ◆ Delay : 把一个报文从网络一端传输到另一端所需的时间 one way, $RTT = \text{two ways}$, **光传播**
 - $3.0 \times 10^8 \text{m/s}$; 光在真空内传播的速率
 - $2.8 \times 10^8 \text{m/s}$; 电在Cable内传播的速率
 - $2.0 \times 10^8 \text{m/s}$; 光在Fiber内传播的速率
- ◆ Round-Trip Time (RTT): 发收来回时间
- ◆ 定义: 延迟 = 处理 + 排队 + 发送 + 传播
 - 处理时延: 检查包首部、决定导向何处; 比特差错检测, 高速R一般在微秒或更低数量级, 接收完整的一个分组的时间 = 包容量 / 链路速率
 - 排队时延: 等待输出链路空闲, 与当时流量和排队规则有关
 - 传输时延: 数据量 / 带宽; 微妙到毫秒级
 - 传播时延: 距离 / 光速 (光缆中1000km传播延迟约5ms) 广域网在毫秒级
- ◆ **传输/传播的比较**
 - 传输时延是路由器把分组发送出去所需时间, 是数据量和链路实际带宽的函数, 与两个路由器间距离无关,
 - 传播时延是一个比特从发送路由器到接收路由器传播所需的时间, 是距离的函数, ms级
- ◆ 例子: 10辆汽车 (P) 在两个收费站 (R) 间的高速公路上行使!



延迟产生的位置



- ◆ 端到端**有效吞吐率** = 实际传输大小/传输时间
- ◆ 实际传输时间 = RTT+ 传输大小/信道带宽
- ◆ RTT: 请求与回答时间

◆ 问题: 信道带宽1Gbps, 端到端时延 $\tau = 10\text{ms}$, TCP发送窗口65535字节。问可能达到的最大吞吐率T ? 信道利用率 ρ ?

◆ 解析: $T = \text{size} / [2\tau + \text{size}/\text{BW}]$

➢ $= \text{size} * \text{BW} / [2\tau * \text{BW} + \text{size}] \text{ bps}$

➢ $= 65535 * 8 / [20 * 10^9 * 10^{-3} \text{s} + 65535 * 8]$

➢ $= 524280 * 10^9 / 20 * 10^6 + 524280$

➢ $= 25.5 \text{ Mbps}, \rho = 25.5\text{M}/1000\text{M} = 2.55\%$

延迟带宽积

- ◆ 一对进程通道间的延迟带宽积：信道管道的体积=所容纳的比特数
- ◆ 一个信道延迟=50ms，带宽45Mbps，则能容纳=50ms x 45Mbps=50X10⁻³ secX45bits/sec=2.25Mbits=280KB



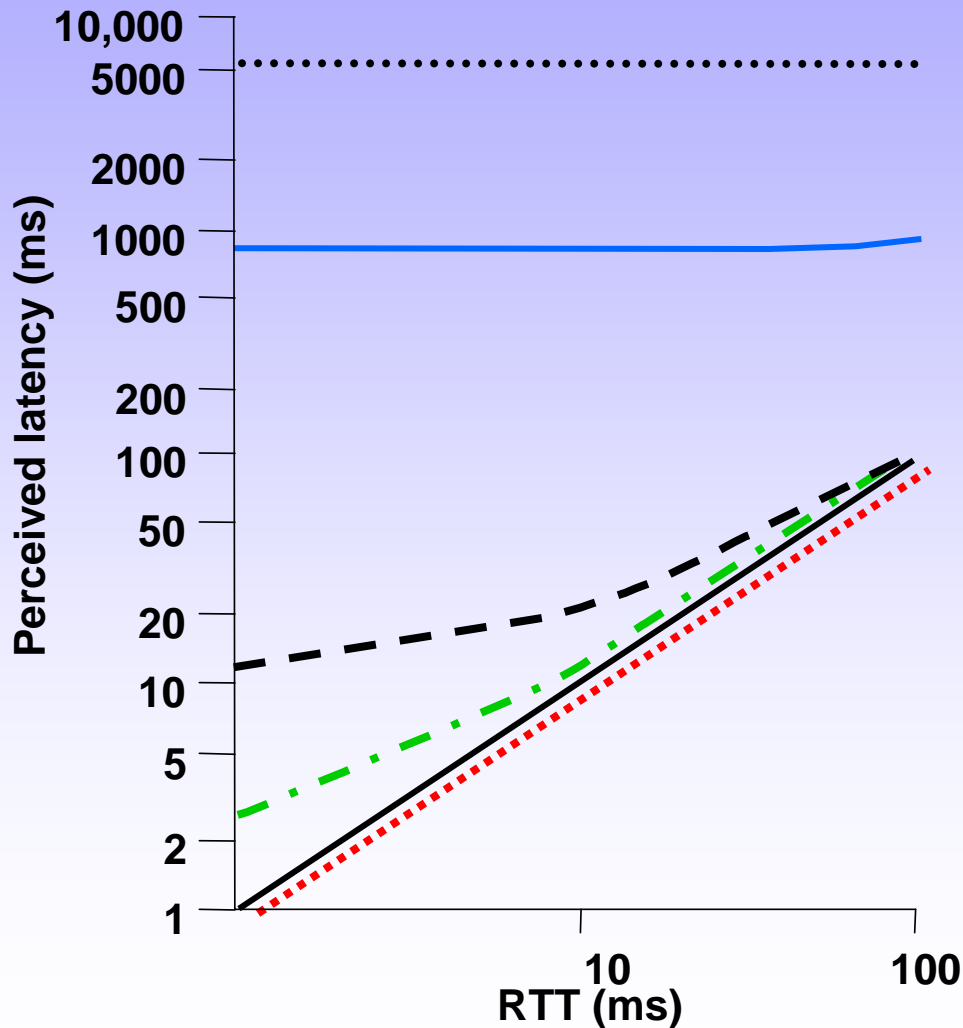
◆ 抖动

- 网络应用看：某段时间内，平均收到数据总量的变化程度
- 某段时间内，相邻包到达时间间隔偏离其均值的大小
- 数据传输速度的变化率，相邻两次比特率的变化率

链路或信道的带宽和延迟

- ◆ 带宽和延迟的组合可描述链路或信道的特征，其相对重要性取决于具体的应用
 - 有些应用**延迟支配带宽**，如端客户发送1字节到服务器并依次收到1字节就是延迟重要，如果响应中没有大量的计算，由于跨洲信道要100ms的RTT，同一房间内仅1msRTT，故其应用完成性能将很不同
 - 数字图书馆来说是**带宽支配延迟**。设带宽10Mbps，则需20s (2.5M/image)，信道是1ms/ 100ms延迟不重要，20.001s和20.1s之差可略

各种对象大小和链路速率的延迟（响应时间）与RTT



- ◆ 对1byte对象（键盘）其延迟同RTT几乎相等，故不能区别1.5M和10M的网络
- ◆ 对2Kb对象（email）链路速率是RTT=1ms时区别很大，然在RTT=100ms时差别又可忽略
- ◆ 对1Mb对象（数字图象），RTT没有造成不同，链路速率支配延迟性能

1-MB object, 1.5-Mbps link
1-MB object, 10-Mbps link ————
2-KB object, 1.5-Mbps link - - -
2-KB object, 10-Mbps link - . -
1-byte object, 1.5-Mbps link
1-byte object, 10-Mbps link ————

因特网应用的不同性能要求

	电子邮件	Telnet	Iphone	组播视频通信
延迟	几分~几小时	几分~几秒	几秒	毫秒~秒
丢失率	0%	0%	小于 25%	允许/平均分布
连续性	无要求	可等待	连续	连续
同步	无要求	无要求	无	有
抖动	无要求	无要求	无串音	不允许

网络性能指标体系

- ◆ 连通性,
- ◆ 吞吐量, 带宽,
- ◆ 时延, 时延抖动
- ◆ 丢包, 丢包率
- ◆ 可靠性、可用性

◆ 基本概念

- 高速并不一定减少时延, 1G链路和1M链路的RTT也许同是100ms, 与距离/拥塞有关
- 端到端有效吞吐率 = 实际传输大小/传输耗时
- 传输耗时 = $RTT + (1/\text{带宽}) \times \text{传输大小}$
- RTT: 请求到回答之间的所有时间

网络性能参数分类

◆ 加性参数

- 时延、抖动、路径长度、路由代价、路由代价...
- 满足可加性：通道特性由沿途各段链路相加决定

◆ 乘性参数

- 可靠性、丢包率
- 满足可乘性：度量为各链路或设备的乘积...

◆ 极性参数

- 带宽、剩余能量、生存时间、吞吐量...
- 满足极值要求（凹性或凸性），度量由通道或设备的瓶颈属性（极值）决定

可靠性

◆ 什么是传输可靠性

- 数据最小颗粒能正确到达
- 数据不同分片能有序到达（当采用分组交换时）

◆ 计算机网络并不是一个完备的世界

- 机器坏/光纤断/电接口受干扰/交换缓存溢出

◆ 网络设计者要考虑三种类型的故障

- 链路上的单比特错，突发连续比特错。

- ☞ 铜缆： 10^6 —— 10^7

- ☞ 光缆： 10^{12} —— 10^{14}

- 包错：一个包在网络上丢失

- ☞ 包中包含有不可纠正的比特错

- ☞ 中继接点因故（缓冲满，TTL）丢掉

- ☞ 输出链路故障等

三种类型的故障

➤ 结点和链路级故障

- ☞ 物理链路断
- ☞ 相连计算机崩溃
- ☞ 电源断电
- ☞ 操作失误

◆ 关键思想：既满足应用的需求又控制在当前技术的极限

◆ 挑战：填补应用希望和当前技术之间的间隙

1.4 如何操作网络？

◆因特网成功的原因之一

- 功能由软件在通用机上提供，加入新功能是很容易的事情

◆应用编程接口API（Sockets）

- Socket interface
- `int socket(int domain, int type, int protocol)` **创建一个插口**，支持任何现行协议，返回socket的一个句柄，以后引用的标识符
- Hook：钩子函数用于在提供的关键点上挂接用户自编的函数

Socket interface

- ◆ Domain:
 - PF_INET ;表示因特网系列
 - PF_UNIX; 表示UNIX系列
- ◆ type :
 - SOCK_DGRAM ;面向报文的服务
 - SOCK_STREAM; 面向字节流服务
- ◆ protocol: 缺省指TCP
- ◆ int bind(int socket, struct addr *address, int addr_len)/给刚创建的插口绑定IP地址和端口结构
- ◆ int listen(int socket, int backlog)/定义在指定插口上有多少待处理的连接
- ◆ int accept(int socket, struct addr *address, int addr_len)/执行被动打开
- ◆ int connect(int socket, struct addr *address, int addr_len)/主动打开, 建立连接
- ◆ int send(int socket, char *message, int message_len, int flags)/在给定套接上发送数据
- ◆ int recv(int socket, char *buffer, int buffer_len, int flags)/在给定套接上接收数据到给定缓冲

练习题

- ◆ 一局域网最大距离为2KM，对于100BYTE的数据包，在带宽为多少时，电磁波的时延等于传输时延？512BYTE的数据包？
- ◆ CIDR地址的利用率分析
 - 划分前网络掩码为m位数，可划分地址 $a+b=32-m$;
 - a是划分后的掩码增加位数；b是划分后的主机位数；
 - 划分前的主机地址数= $2^{a+b}-2$ ；(减去全0全1网络号)
 - 划分后的主机地址数 = 新子网 \times 新子网中的主机数
 - $= (2^a - 2) \times (2^b - 2) = 2^{a+b} - 2 - (2^{a+1} + 2^{b+1} + 2)$
 - 利用率 $r = 1 - (2^{a+1} + 2^{b+1} + 2) / (2^{a+b} - 1)$
 - 对B类地址，设 $a=6, b=10$ ；则 $r = (64 + 1024 + 1) / (32768 - 1)$
 - $1 - 1089 / 32767 = 1 - 0.03323 = 0.9667$
 - 对B类地址， $a+b=16$ ；对C类地址， $a+b=8$ ；

练习题

- ◆ 13. 假设在地球和月球间建立一点对点100MPS的连接，地月距离为385, 000KM，数据传送速率为光速—— 3×10^8 M/S
- A. 计算最小的RTT
 - B. use the RTT as the delay , calculate the delay*bandwidth product for the link
 - C. 在B中计算的delay*bandwidth有何意义？
 - D. 在月球照的照片，并以数字格式存在磁盘上。如果地球上的控制中心想下载25M的最新的图形，那么从发出数据到传输完成所要的最小时间是多少？

练习题

◆ P60--66

◆ 5, 13, 17, 23, 27, 30, 33

Project1:

- 用Socket编程，从服务器得到一个随机数，并把自己的学号用TCP协议送到服务器。----10分。
- 2周时间，10月1日前完成。

Thank you!

