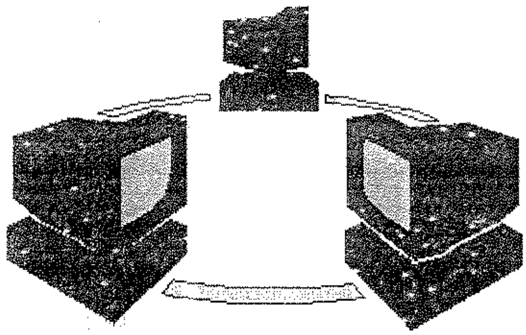
COMPUTER NETWORKS

FOURTH EDITION

2P 网络技术基础（英）



Andrew S. Tanenbaum

1 Introduction

学习目的：了解计算机网络发展历史、计算机网络的分类、网络体系结构（OSI/RM、TCP/IP）、计算机网络相关的基本概念和交换技术等。

学习重点：熟练掌握计算机网络、网络协议、网络体系结构等三个重要概念。

学习难点：掌握三种交换技术之间的异同。

Computer Network: A collection of autonomous computers（自主计算机） interconnected.

Computer Network is composed of **hosts, protocols** and **communication-subnet.**

- **The development of computer network:(additional)**

(1) First, Terminal-oriented network.

Computing technology combine with communication technology.

Circuit-exchange(电路交换): Connection establishment, Data transform, Connection release.

(2) Second, Packet-exchange network: 1969.11, ARPANET

Compare Circuit-exchange, Packet-exchange and Message-exchange

(3) Architecture of network:

ISO: OSI/RM(开放系统互连基本参考模型,Reference model)

Internet: TCP/IP

IBM: SNA(系统网络体系结构)，1974，the first architecture of network

Digital: DNA(Distributed network architecture, 分布式系统体系结构)

1.1 Uses of Computer Networks

1.1.1 Business Application

The goals of the computer network:

(1) Resource sharing（hardware and software 资源共享）

(2) High reliability（高可靠性）

(3) Saving money（节约经费）

(4) Expansion（可扩充性）

(5) Communication medium（通信手段）

1.1.2 Public Application

(1) Access to remote information(访问远程信息)

(2) Person tom person communication(个人间通信)

(3) Interactive entertainment(交互式娱乐)

(4) E－commerce(电子商务)

(5) Mobile users(移动用户)

1.1.3 Social Issues

1.2 Network Hardware

Broadly speaking(从广义上讲)，two types of transmission technology:

1. Broadcast links
2. Point-to-point links

Broadcast network(广播式网络)： broadcasting and multicasting

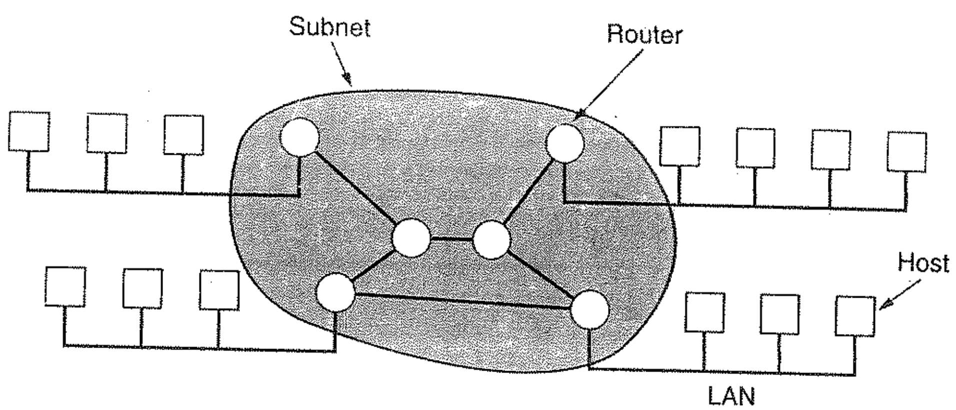
Point-to-point network(点到点网络)： packet—transmission(分组转发) and routing—algorithms（路由选择）

1.2.1 Local Area Networks(局域网)

three characteristics: (1)Size(范围) (2)transmission technology(传输技术) (3)topology(拓扑结构)

1.2.2 Metropolitan Area Networks(城域网)

1.2.3 Wide Area Networks(广域网)



1.2.4 Wireless Networks

(1) System interconnection

(2) Wireless LANs

(3) Wireless WANs

1.3 Network Software

1.3.1 Protocol Hierarchies(协议分层)

To reduce their design complexity, most networks are organized as a stack of **layers** or **levels.**

- **Merit of Protocol Hierarchies:(分层的优点)**

(1) Independence of layers: These layers are relative independence.(各层之间是相对独立的)

(2) If any layer is changed, other layers are not affected, so far as not changing interface. (灵活性好)

(3) Each layer can be realized in right technology.(结构上可以分割开)

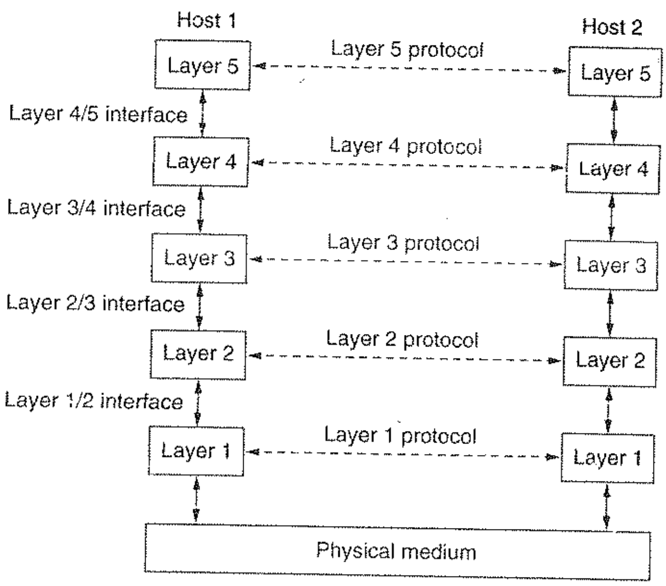
(4) It is easy to realize and maintain. (易于实现和维护)

(5) Standardization (能促进标准化工作)

- **Defect of Protocol Hierarchies:(分层的缺点)**

(1) Some functions may be realized reduplicatively.(功能重复实现，资源浪费)

(2) The number of layer is not certain.(层次数难以确定)



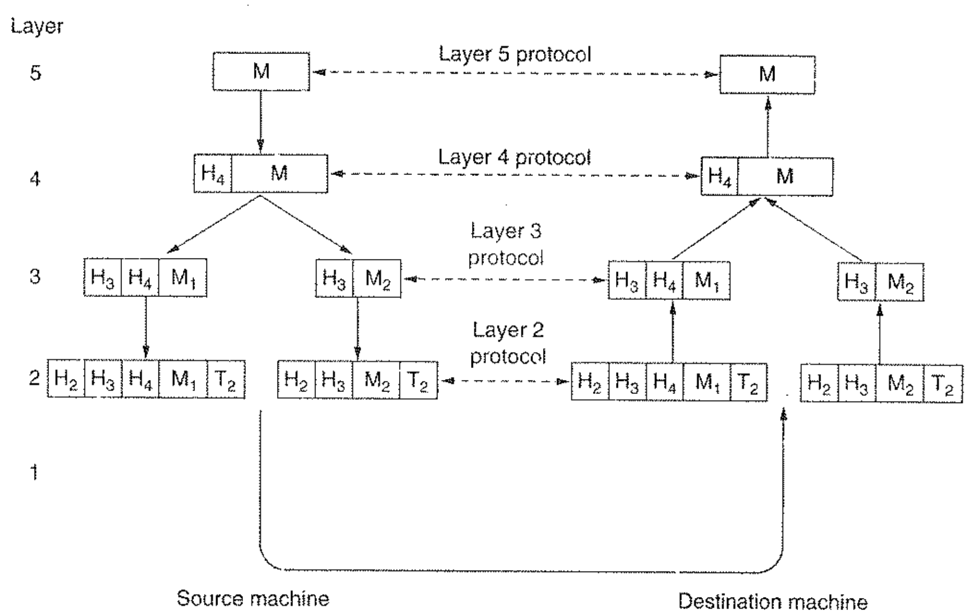
A list of protocols used by a certain system, one protocol per layer, is called a protocol stack.

Protocol is an agreement between the communicating parties on how communication is to proceed.

*****Protocol、Service** and **SAP*****

Network architecture is a set of layers and protocols.

A list of protocols used by a certain system, one protocol per layer, is called a **protocol stack.**



1.3.2 Design Issues for the Layers

(1)A mechanism for identifying senders and receivers.

(2)The rules for data transfer: simplex communication(单工通信)、half-duplex communication(半双工通信)、full-duplex communication(全双工通信).

(3)Error control(差错控制)

(4)Flow control(流量控制)

(5)Disassembling(分割) and gathering（合并）messages

(6)Multiplexing(多路复用) and demultiplexing(多路分用、解多路复用)

(7)Routing(路由选择)

1.3.3 Connection-Oriented and Connectionless Services

Type of connection	Service	Example
Connection-oriented	Reliable message stream	Sequence of pages
	Reliable byte stream	Remote login
	Unreliable connection	Digitized voice
Connectionless	Unreliable datagram	Electronic junk mail
	Acknowledged datagram	Registered mail
	Request-reply	Database query

连接类型	服务类型	应用实例
面向连接的服务	可靠的消息流	页码序列
	可靠的字节流	远程登陆
	不可靠的连接	数字化的声音
无连接服务	不可靠的数据报	电子方式的函件
	有确认的数据报	挂号邮件
	问答	数据查询

1.3.4 Service Primitives(服务原语)

Primitive	Meaning
LISTEN	Block waiting for an incoming connection

CONNECT	**Establish a connection with a waiting peer(对等进程)**
RECEIVE	**Block waiting for an incoming message**
SEND	**Send a message to the peer**
DISCONNECT	**Terminate a connection**

原语	含　义
请求 **request**	一个实体希望得到完成某些操作的服务
指　示 **indication**	通知一个实体，有某个事件发生
响　应 **response**	一个实体希望响应一个事件
证　实 **confirm**	返回对先前请求的响应

1.3.5 The Relationship of Services and Protocols

(1)A service is a set of primitives (operations) that a layer provides to the layer above it.

(2)A protocol is a set of rules governing the format and meaning of the packets, or messages that are exchanged by the peer entities within a layer.

(3)Protocol is horizontal, and service is vertical.

1.4 Reference Models

1.4.1 The OSI Reference Model

The OSI model has seven layers. The partitioned principles:

(1)A layer should be created where a different level of abstraction is needed.

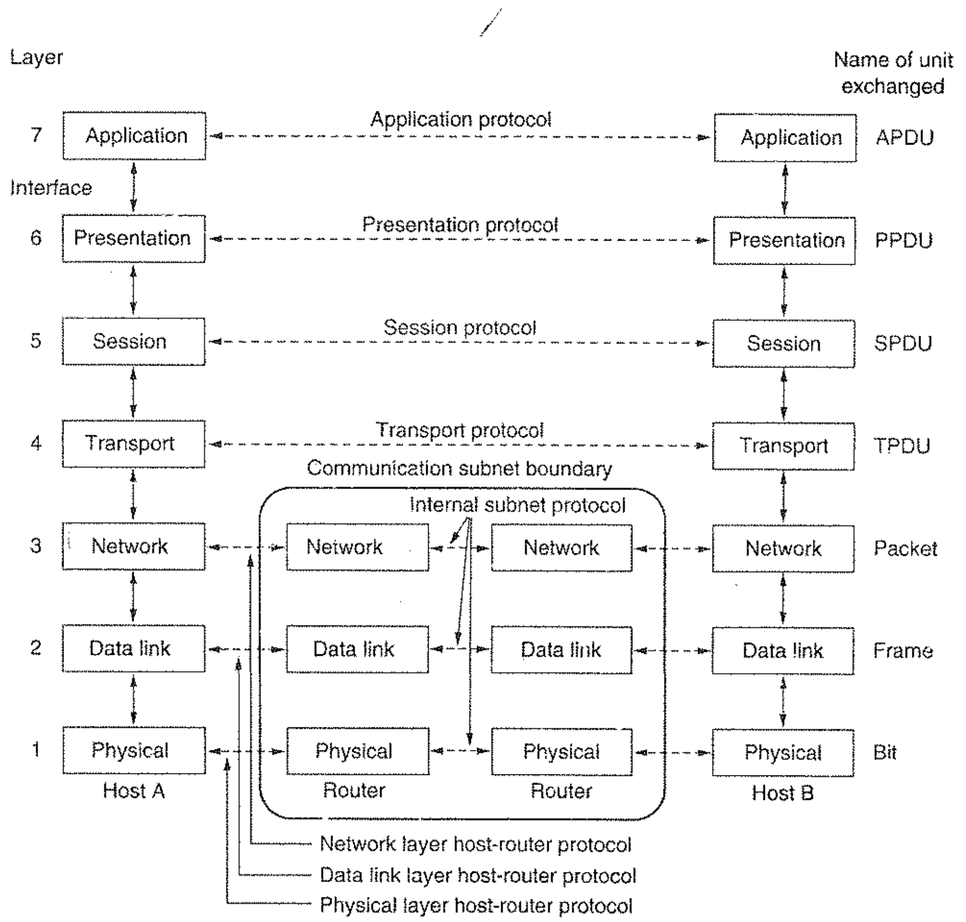
(2)Each layer should perform a well defined function.

(3)The function of each layer should be chosen with an eye toward defining internationally standardized protocols.

(4)The layer boundaries should be chosen to minimize the information flow across the interfaces.

(5)The number of layers should be large enough that distinct functions need not be thrown together in the same layer out of necessity, and small enough that the architecture does not become unwieldy(庞大的).

1. Functions of each layer



(1).**The Physical Layer**: To transmit bits

Such as mechanical characteristic, electrical characteristic, time characteristic and function characteristic.

(2).**The Data Link Layer**: To transform data frames without errors.

(3).**The Network Layer**: how to route from source to destination.(transmit packet)

(4).**The transport Layer**: the most important layer. To accept data from the session layer, split it up into smaller units if need be, pass these to the network layer, and ensure that the pieces all arrive correctly at the other end.

(5).**The Session layer**: To establish sessions between different users.

A session service is token management.(令牌管理)

Another session service is synchronization.（同步）

(6).**The Presentation layer**:

its function contains: (a)data compress and decompress

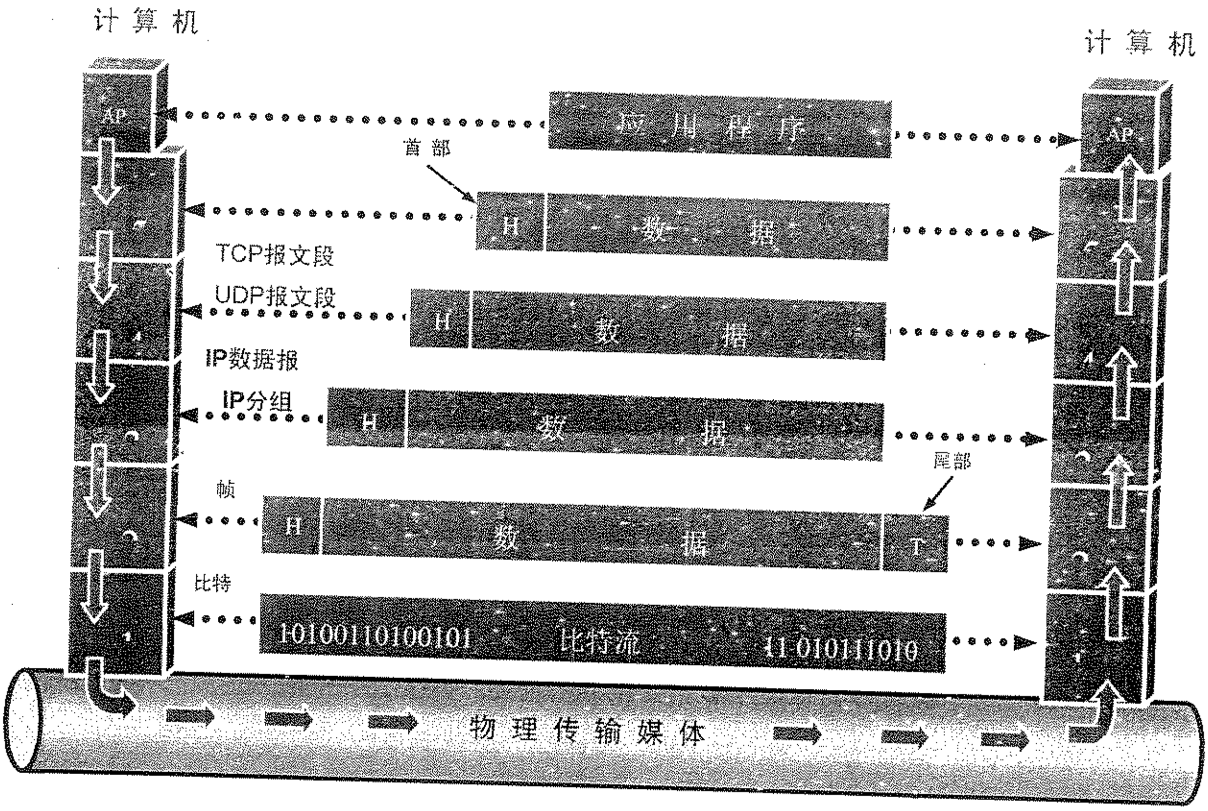
(b)data encrypt and decrypt

(c)data format

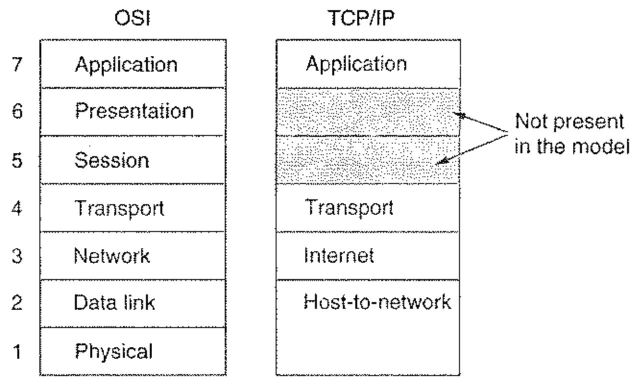
(7).**The Application Layer**: (a)To provide service for application procedure.

(b)file transfer

2.Data Transmission in the OSI Model



1.4.2 The TCP/IP Reference Model



1. The Host-to-Network Layer(主机至网络层): point out that the host has to connect to the network.

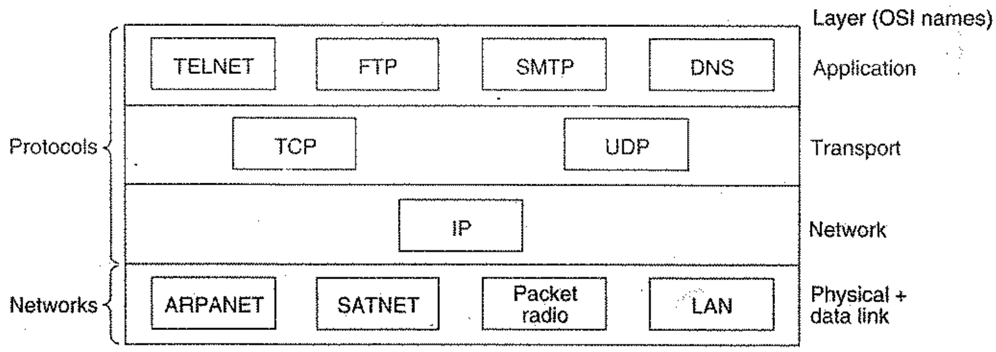
2. **The Internet Layer**（互连层）: Packet routing(分组路由) and avoiding congestion（避免拥塞）.

3. **The Transport Layer(传输层):** end-to-end communication. **(TCP and UDP)**

TCP: an reliable, connection-oriented protocol

UDP: an unreliable, connectionless protocol

4. The Application Layer(应用层):



1.4.3 A Comparison of the OSI and TCP/IP Reference Models

Differences:

1. The OSI/RM defines three concepts: Services, Interfaces and Protocols. But the TCP/IP doesn't.

2. Connection-oriented and connectionless communication.

1.4.4 A Critique of the OSI Model and Protocols

1. Bad timing(糟糕的提出时机)

2. Bad technology（糟糕的技术）

3. Bad implementations（糟糕的实现）

4. Bad politics（糟糕的策略）

1.4.5 A Critique of the TCP/IP

1. The TCP/IP does not clearly distinguish the concepts of service, interface and protocol.

2. The TCP/IP is not at all general and is poorly suited to describing and protocol stack other than TCP/IP.

3. **The host-to-network layer is an interface, but not a layer.**

4. **The TCP/IP does not distinguish the physical and data link layers.**

5. Although, the IP and TCP protocols were carefully thought out and well implemented, many of the other protocols were **ad hoc**(特别)

and have defect.

In summary:

5	Application layer
4	Transport layer
3	Network layer
2	Data link layer
1	Physical layer

1.5 Performance parameters of computer network

1. Bandwidth(带宽): The range of signal frequency

2. Delay(时延)：

- 发送时延＝ 数据块长度/信道带宽
- 传播时延＝ 信道长度/电磁波在信道上的传播速率

  电磁波在自由空间的传播速率是光速，即 3.0×10^5km/s，在铜缆中的传播速率约为 2.3×10^5km/s，在光纤中传播速率约为 2.0×10^5km/s.

- 处理时延：数据在交换结点为存储转发而进行一些必要的处理所花费的时间。
- 时延带宽积

  时延带宽积＝ 传播时延 × 带宽

  从公式可知：时延带宽积又可以称为**以比特为单位的链路长**

度。

- 往返时延：表示从发送端发送数据开始，到发送端收到接收端的确认，总共经历的时延。

Number	Topic
802.1	Overview and architecture of LANs
802.2 ↓	Logical link control
802.3 *	Ethernet
802.4 ↓	Token bus (was briefly used in manufacturing plants)
802.5	Token ring (IBM's entry into the LAN world)
802.6 ↓	Dual queue dual bus (early metropolitan area network)
802.7 ↓	Technical advisory group on broadband technologies
802.8 †	Technical advisory group on fiber optic technologies
802.9 ↓	Isochronous LANs (for real-time applications)
802.10 ↓	Virtual LANs and security
802.11 *	Wireless LANs
802.12 ↓	Demand priority (Hewlett-Packard's AnyLAN)
802.13	Unlucky number. Nobody wanted it
802.14 ↓	Cable modems (defunct: an industry consortium got there first)
802.15 *	Personal area networks (Bluetooth)
802.16 *	Broadband wireless
* 802.17	Resilient packet ring

1.6.1 Who's Who in the Telecommunications World

ISO (International Standards Organization)

ITU (International Telecommunication Union).

IEEE (Institute of Electrical and Electronics Engineers)

IAB (Internet Activities Board)

IRTF (Internet Research Task Force)

IETF (Internet Engineering Task Force)

RFCs (Request For Comments).

1.7 Metric Units

Exp.	Explicit	Prefix	Exp.	Explicit	Prefix
10^{-3}	0.001	milli	10^{3}	1,000	Kilo
10^{-6}	0.000001	micro	10^{6}	1,000,000	Mega
10^{-9}	0.000000001	nano	10^{9}	1,000,000,000	Giga
10^{-12}	0.000000000001	pico	10^{12}	1,000,000,000,000	Tera
10^{-15}	0.000000000000001	femto	10^{15}	1,000,000,000,000,000	Peta
10^{-18}	0.000000000000000001	atto	10^{18}	1,000,000,000,000,000,000	Exa
10^{-21}	0.000000000000000000001	zepto	10^{21}	1,000,000,000,000,000,000,000	Zetta
10^{-24}	0.000000000000000000000001	yocto	10^{24}	1,000,000,000,000,000,000,000,000	Yotta

P78, KB, MB, GB for the xth power of 2

kbps, Mbps, Gbps for the xth power of 10

2 THE PHYSICAL LAYER

学习目的：了解物理层的基本功能，熟练掌握奈氏公式和香农公式，了解各种传输媒体的特性。

学习重点：熟练掌握奈氏公式和香农公式，以及各种传输媒体的特性。

学习难点：熟练掌握曼彻斯特编码和差分曼彻斯特编码。

- mechanical characteristic(机械特性)
- electrical characteristic(电气特性)
- time characteristic(规程特性，时间特性)
- function characteristic(功能特性)

2.1 The Theoretical Basis for Data Communication

2.1.1 Basic Concept

- 信息 **information**：信息是对客观物质的反映，可以是对物质的形态，大小，结构，性能等全部或部分特性的描述，也可以是物质与外部的联系。信息有各种存在的形态，如文字，声音，图像等等。
- 数据 **data**：信息也可以用数字的形式表示，数字化的信息称为数据。数据是装载信息的实体，信息则是数据的内在含义或解释。为了确切的表示信息，数据有时是一些连续值，另一些则取离散值，如声音的强度，灯光的亮度等都可以连续变化，而成绩，名次等的取值都是离散的。连续值的数据叫做模拟数据，离散值的数据叫做数字数据。
- 信号 **signal**：数据通信中的“信号”是指数据的电编码或磁编码。它分为模拟信号和数字信号两种。模拟信号是连续变化的电磁波，数字信号则是一串电压脉冲序列。两种信号在一定技术措施下可

以相互转换。信号可以在各种传输媒介上传送，如双绞线，电话线，同轴电缆，光缆，甚至可以通过卫星以微波的方式传送。

- 噪声 **noise**：信号在传输过程中受到的干扰称为噪声，干扰可能来自外部，也可能由信号传输过程本身产生。噪声过大将影响被传送的信号的真实性或正确性。
- 信道 **channel**：信道是传送信号的一条通路，由传输介质及相应的附属设备组成。同一个传输介质上可以同时存在多条信号通路，即一条传输线路上可以有多个信道。例如一条光缆可以供上千对人同时通话，有上千个电话信道。
- **The rules for data transfer**: simplex communication(单工通信)、half-duplex communication(半双工通信)、full-duplex communication(全双工通信).

2.1.2 The maximum data rate of a channel

Nyquist's theorem:

理想低通信道：maximum data rate= $2H \log_2 V$ bps （H：channel bandwidth）

理想带通信道：maximum data rate= $H \log_2 V$ bps

Notice: difference of Baud(波特) and Bit（比特）

Shannon's theorem:

Maximum number of bps= $H \log_2(1+S/N)$ bps

Notice: H is channel bandwidth

S/N is signal-to-noise ratio(信噪比)

Example：bandwidth is 3kHz, signal-to-noise ratio is 30dB, so,

$10\log_{10}S/N=30$,　$S/N=1000$

Maximum number of bps=$3k\times\log_2(1+1000)\approx 30k$ bps

2.2 Transmission Media

2.2.1 Twisted pair(双绞线)

UTP (Unshielded Twisted Pair): 非屏蔽双绞线

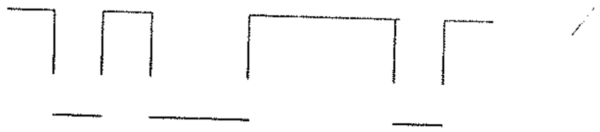
STP (Shielded Twisted Pair): 屏蔽双绞线

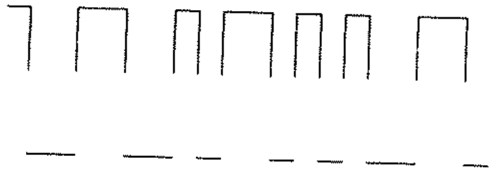
2.2.2 Coaxial Cable(同轴电缆)

(1) 50Ω：digital transmission（BNC）

Notice: Manchester Code(曼彻斯特编码) && Differential Manchester Code(差分曼彻斯特编码)**(Page275)**

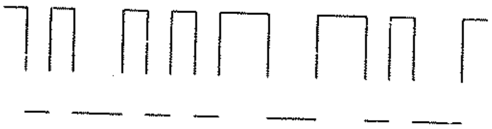
1 0 1 0 0 1 1 1 0 1





Manchester Code

Differential Manchester Code



Example: Bit rate: 10Mbps (Ethernet network)

Baud rate=?

(2) 75 Ω： analog transmission,CATV (AUI)

2.2.3 Fiber Cables(光缆) (F/O)

(1) Multi-mode fiber: 多模光纤

(2) Single-mode fiber: 单模光纤

Light sources:

(1) LED: 发光二极管

(2)Semiconductor laser: 半导体激光管

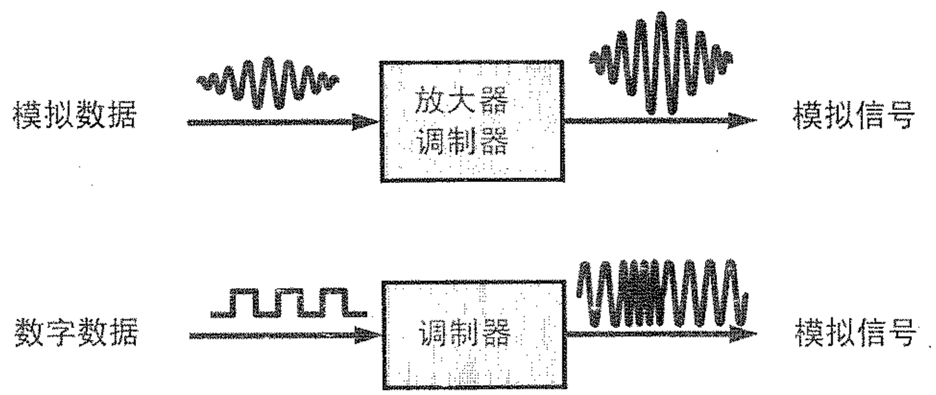
Item	LED	Semiconductor laser
Data rate	Low	High
Fiber type	Multimode	Multimode or single mode
Distance	Short	Long
Lifetime	Long life	Short life
Temperature sensitivity	Minor	Substantial
Cost	Low cost	Expensive

2.3 Wireless Transmission(无线传输)

1. Radio transmission(无线电传输)
2. Microwave transmission(微波传输)
3. Infrared and Millimeter waves(红外线和毫米波)
4. Lightwave transmission(激光传输)

2.4 Analog Transmission and Digital Transmission(模拟传输和数字传输)

2.4.1 Analog Transmission



Modem: 调制解调器

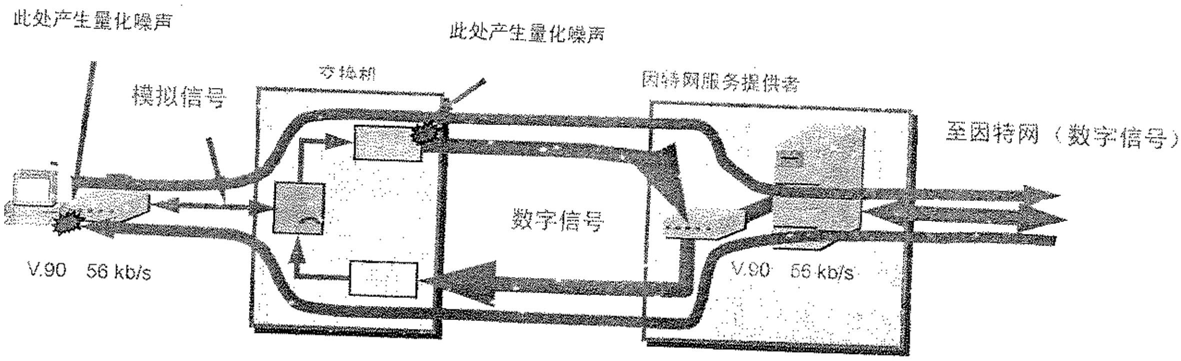
Modulation：调制，digital data ---->analog signal

AM：调幅

FM：调频

PM：调相

Demodulation：解调，analog signal ----> digital data



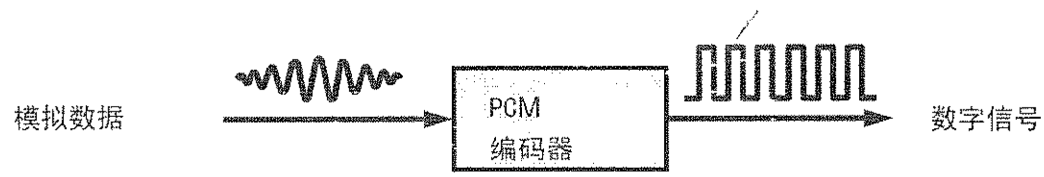
Example： Modem 上行速率：33.6kbps（limit rate：35kbps）

『Bandwidth：3.1kHz，S/N=2500，

Maximum number of bps= H $\log_2(1+S/N)$ bps=35kbps』

Modem 下行速率：56kbps

2.4.2 Digital Transmission



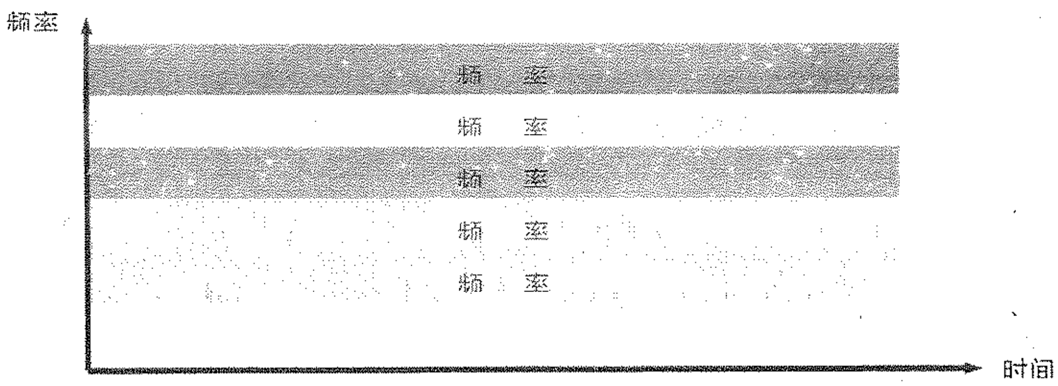
Codec: 编码解码器

Code：编码，analog data---->digital signal

Decode：解码，digital signal ----> analog data

2.5 Multiplexing(多路复用)

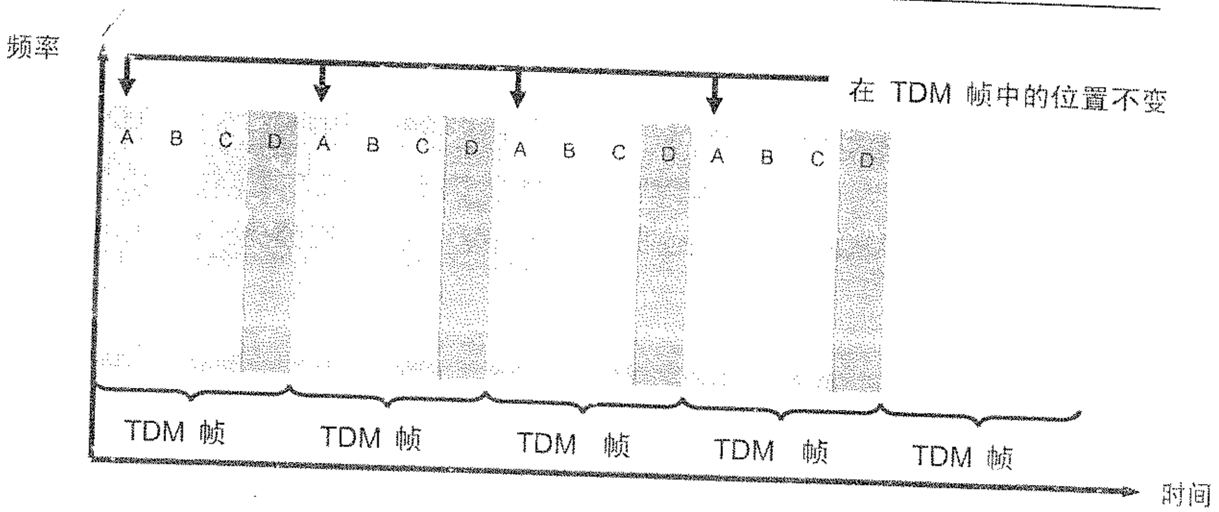
2.5.1 <u>Frequency Division Multiplexing(FDM, 频分多路复用)</u>

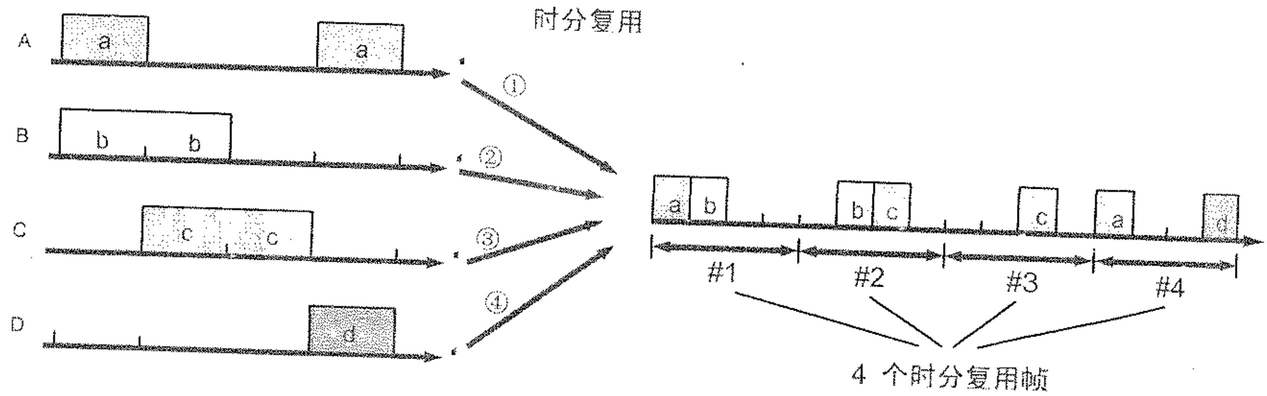


所有用户在同样的时间占用不同的带宽资源。

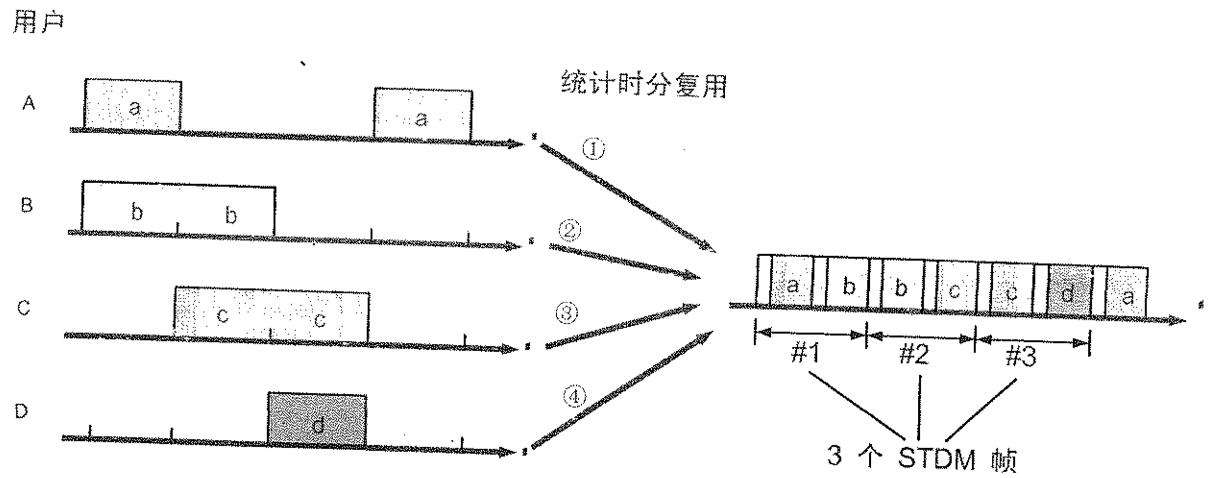
2.5.2 <u>Time Division Multiplexing(TDM, 时分多路复用)</u>

所有用户在不同的时间占用同样的频带宽度。





2.5.3 Statistical TDM(统计时分复用)



2.5.4 Wavelength Division Multiplexing(WDM, 波分复用)

即：光的频分复用

2.5.5 <u>Code Division Multiplexing(CDM, 码分复用)</u>

各用户使用经过特殊挑选的不同码型，因此彼此不会造成干扰。

3 The Data Link Layer

学习目的：了解数据链路层的基本概念，掌握差错控制方法，理解停止等待 ARQ、连续 ARQ 和选择重传 ARQ 协议的基本原理。

学习重点：差错控制方法的应用，三种 ARQ 协议的分析。

学习难点：窗口机制

Link: Physical connect, physical link(物理链路)

Data Link: Hardware + Software + Link, logical link(逻辑链路)

Interconnect: 互连

Intercommunicate: 互通

Inter-manipulate: 互操作

3.1 Data Link Layer Design Issues (Functions of DLL)

1. Services provided to the Network Layer

The Data Link Layer commonly provided three services:

(1) Unacknowledged connectionless service(无确认的无连接服务):

Uses:

the error rate is very low; (误码率低)

real-time voice; (实时语音通信)

most LANS; （大多数局域网使用）

(2) Acknowledged connectionless service(有确认的无连接服务):

Uses:

Wireless systems(无线系统)

(3) Acknowledged connection-oriented service(有确认的面向连接服务)

Uses:

A reliable bit stream(可靠的比特流)

2. Framing(成帧，帧同步)

Four methods:

(1) Character count(字符计数法)：DEC 公司的 DDCMP（数字数据通信报文协议）

(2) Flag bytes with bytes stuffing(带字符填充的首尾界符法)：IBM 公司的 BSC 协议（二进制同步通信）

(3) Starting and ending flags, with bit stuffing（带位填充的首尾标志法）：ISO 的 HDLC（高级数据链路控制规程）

(4) Physical layer coding violations(物理层编码违例法)：如物理层采用曼彻斯特编码中就可采用这种方法

3. Error control(差错控制)

Noise causes errors.(随机热噪声和冲击噪声)

4. Flow control(流量控制)

(1) Feedback-based flow control: 基于反馈机制的流量控制（the data link layer）

(2) Rate-based flow control: 基于速率机制的流量控制（the network layer）

5. Link management(链路管理)：主要用于面向连接服务

(1) Connection establishment

(2) Connection keep

(3) Connection release

3.2 Error Detection and Correction

1. Error-detecting codes: 检错码

(1) Parity detection(奇偶检测)

(2) CRC 码（循环冗余码）

Example: $G(x)=x^4+x^3+1$, $M(x)=1011001$,

$T(x)=?$

2. Error-correcting codes: 纠错码

(1) Hamming code(海明码)

若信息位为 k 位，冗余位为 r 位，则 $2^r>=k+r+1$

3.3 Data Link Protocols(数据链路协议)

3.3.1 Stop-and-wait protocol(ARQ)

1. Unrestricted model: (理想模型)

assume1: no error; (数据传输不会出错，也不会丢失)

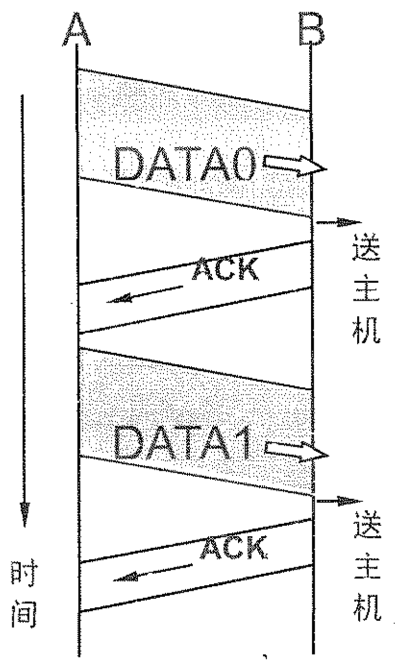
assume2: no flow-control.（不管发方以多快的速率发送数据，收方总是来得及收下）

2. Simple model(简单模型)

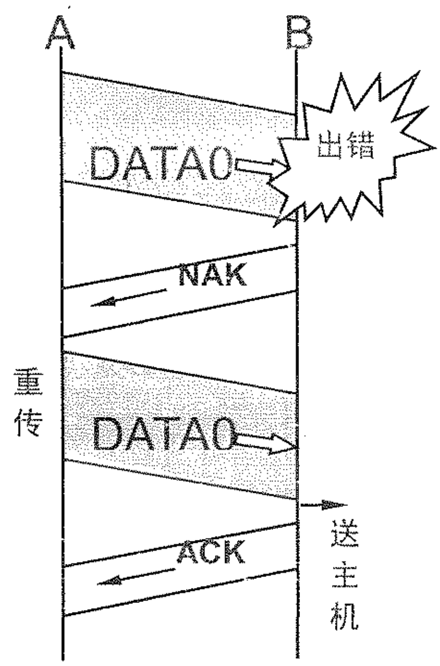
assume1：no error；(数据传输不会出错，也不会丢失)

Flow-control : the receiver provide feed-back to the sender. (由接收方控制发送方的数据流量，实现简单的流量控制功能，这也是计算机网络中流量控制的一个基本方法)

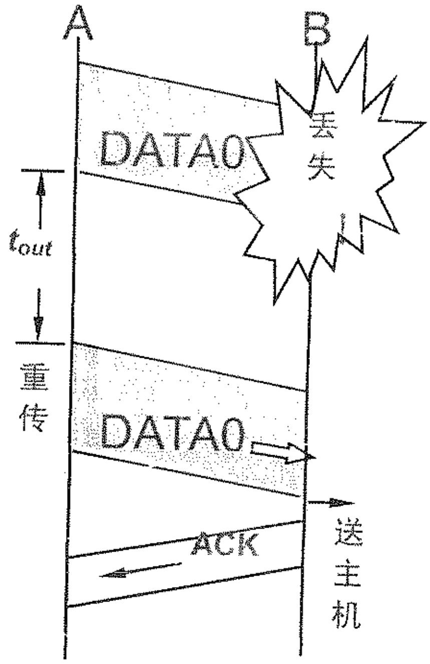
3. Normal model(实际模型)



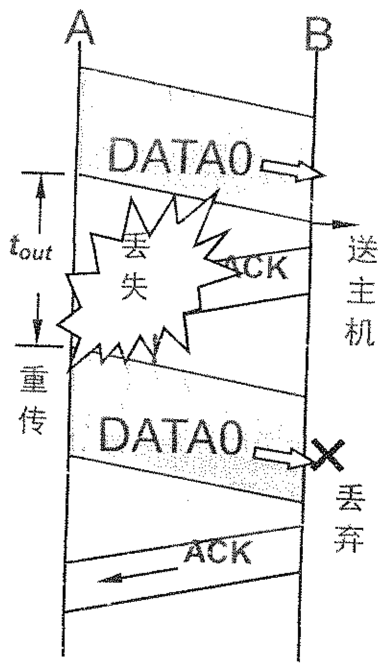
(a) 正常情况



(b) 数据帧出错



(c) 数据帧丢失



(d) 确认帧丢失

T_{out}(超时计时器)：其值（重传时间）一般选为略大于“从发完数据帧

到收到确认帧所需的平均时间”。

3.3.2 Go-back-n protocol(连续 ARQ 协议)

1. Sliding window protocols(滑动窗口协议)

(1) Sending window $\mathbf{W_T}$: 发送窗口，对发送端进行流量控制，其大小代表在还没收到对方确认信息的情况下发送端最多可以发送多少数据帧。

(2) Receiving window $\mathbf{W_R}$: 为了控制可以接收哪些数据帧而不可以接收哪些帧，在连续 ARQ 协议中，其大小为 1 。

2. Go-back-n ARQ protocols(连续 ARQ 协议)

Notice:

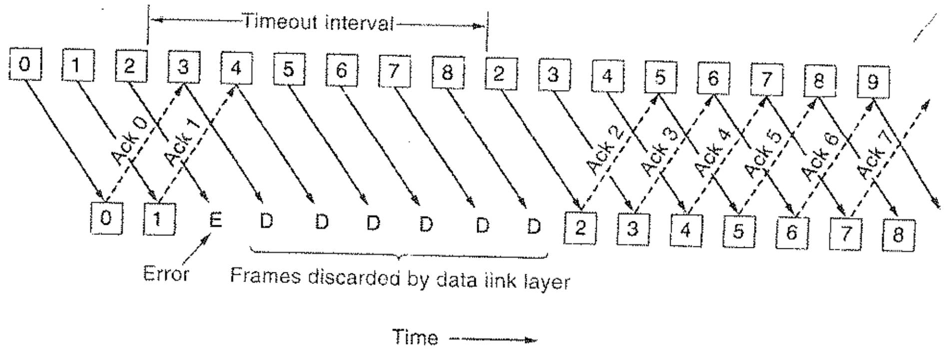
(1) Receiving frames in sequence(按序接收帧)

(2) Starting timeout timer after sending a frame(每发送完一个帧，就要启动一个超时计时器。)

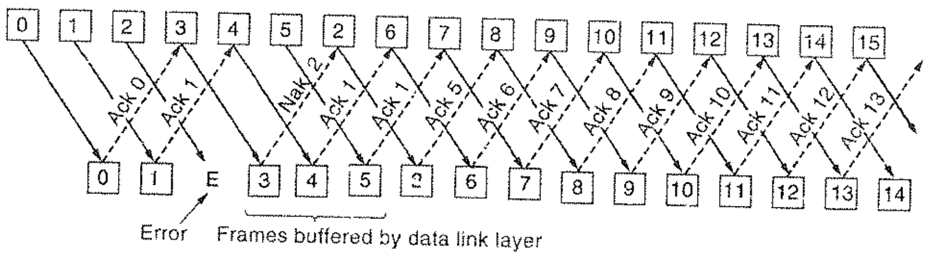
(3) Suppose, n-bit sequence number, $\mathbf{W_T<=2^n-1}$.

(4) Why not $\mathbf{W_T=2^n}$?

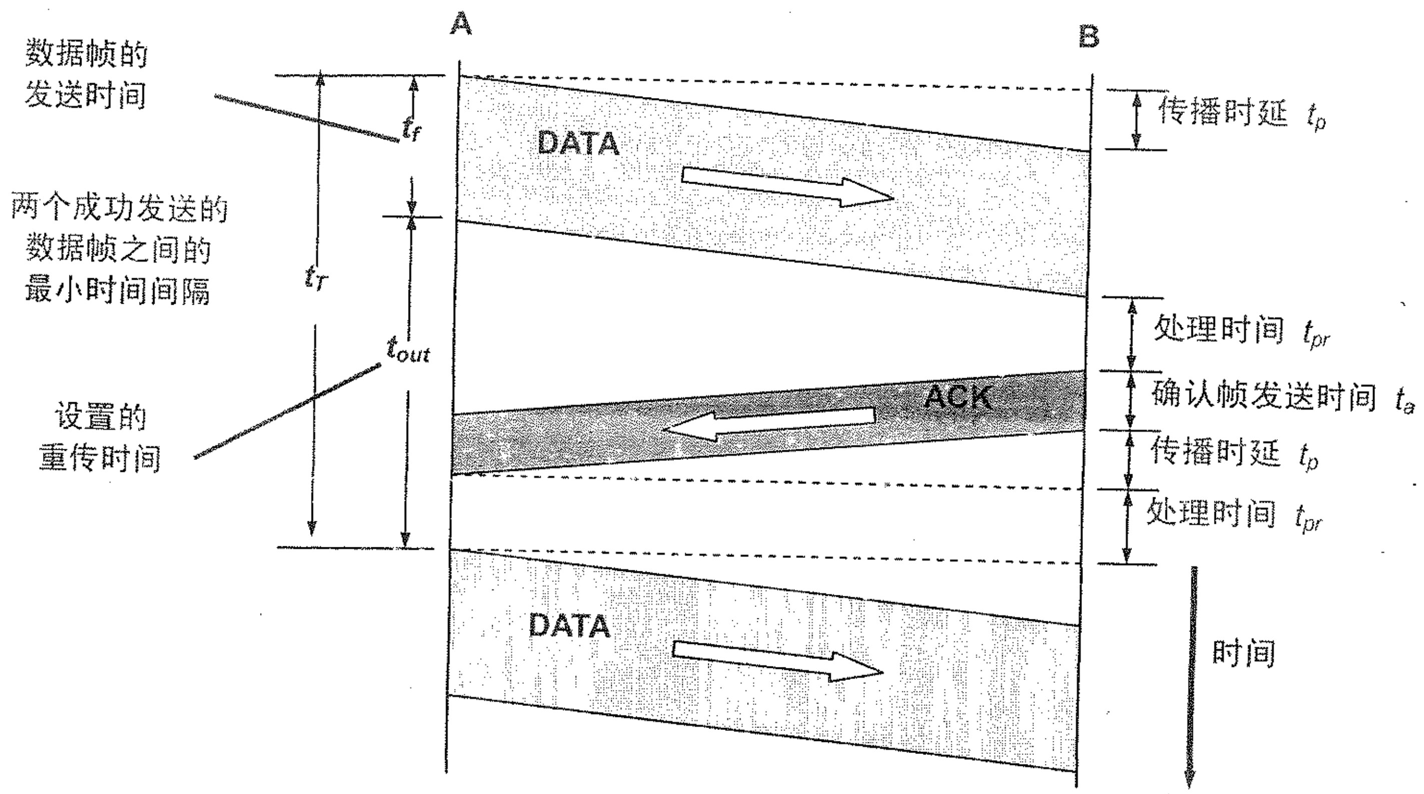
example: n=3， W_T=8.


Timeout interval
0 1 2 3 4 5 6 7 8 2 3 4 5 6 7 8 9
Ack 0
Ack 1
Ack 2
Ack 3
Ack 4
Ack 5
Ack 6
Ack 7
0 1 E D D D D D D 2 3 4 5 6 7 8
Error
Frames discarded by data link layer
Time

(a)


0 1 2 3 4 5 2 6 7 8 9 10 11 12 13 14 15
Ack 0
Ack 1
Nak 2
Ack 1
Ack 1
Ack 5
Ack 6
Ack 7
Ack 8
Ack 9
Ack 10
Ack 11
Ack 12
Ack 13
0 1 E 3 4 5 2 6 7 8 9 10 11 12 13 14
Error
Frames buffered by data link layer

(b)


A
B
数据帧的
发送时间
t_f
两个成功发送的
数据帧之间的
最小时间间隔
t_T
t_{out}
设置的
重传时间
DATA
ACK
DATA
传播时延 t_p
处理时间 t_{pr}
确认帧发送时间 t_a
传播时延 t_p
处理时间 t_{pr}
时间

3.3.3 Selective repeat ARQ protocol(选择重传 ARQ 协议)

Notice:

(1) Suppose, n-bit sequence number, $W_R \le 2^n/2$ & $W_T + W_R \le 2^n$

(2) Why not $W_R > W_T$? & $W_T + W_R \le 2^n$?

example: n=3， $W_R=4$ & $W_T=5$

Summary: (1) Stop-and-wait ARQ protocol: $W_T=1, W_R=1$

(2)Go-back-N ARQ protocol: $W_T \ge 1, W_R=1$

(3)Selective repeat ARQ protocol: $W_T \ge 1, W_R \ge 1$

3.4 Example Data Link Protocols

Bits	8	8	8	≥0	16	8
	01111110	Address	Control	Data	Checksum	01111110

3.4.1 HDLC(面向比特的高级数据链路控制规程)

1. the address field: 地址字段

2. the control field: 控制字段，sequence numbers, acknowledgements, and other purposes.

(1) an information frame: 信息帧

(2) a supervisory frame: 监督帧，no data information. The minimum frame contains three fields and totals **32 bits**, excluding the flags on either end. 比如 ACK 和 NAK 帧

(3) an unnumbered frame: 无编号帧，比如控制帧

Bits	1		3	1	3
(a)	0		Seq	P/F	Next
(b)	1	0	Type	P/F	Next
(c)	1	1	Type	P/F	Modifier

3. the data field: 数据字段 may contain any information

notice: 0-bit stuffing for data transparency: 零比特填充法保证了数据的透明传输。

4. the checksum field: 校验和字段，CRC－CCITT redundancy code（冗余码）

3.4.2 PPP(因特网的点对点协议，面向字节)

The PPP full frame format:

Bytes	1	1	1	1 or 2	Variable	2 or 4	1
	Flag 01111110	Address 11111111	Control 00000011	Protocol	Payload	Checksum	Flag 01111110

4 The MAC Sub-layer

学习目的：理解 CSMA/CD 协议的基本原理，学习 ETHERNET 的相关知识，网络互联知识。

学习重点：CSMA/CD 的基本原理，局域网的体系结构，网络互联。

学习难点：ETHERNET 中最大跨距的计算。

Static channel allocation methods: 信道静态分配方法， such as: FDM、TDM、WDM、CDM

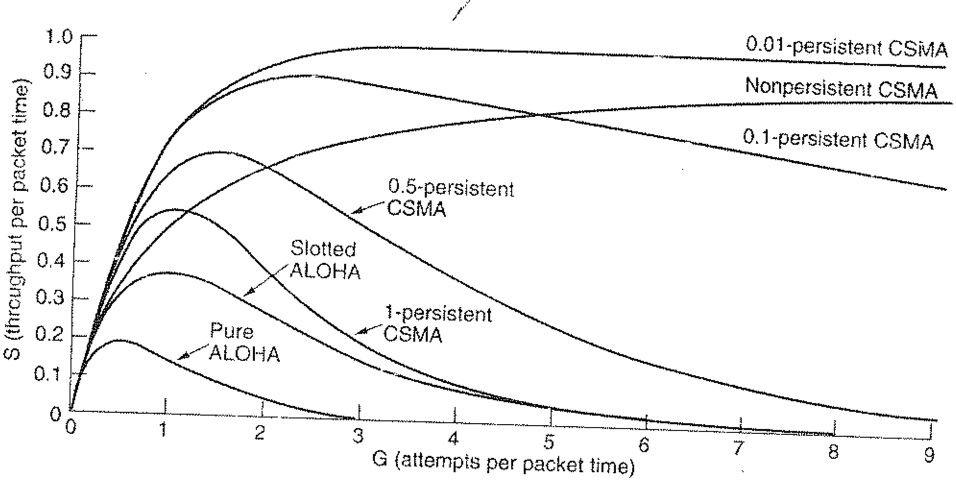
Dynamic channel allocation methods: 信道动态分配方法，

- **random access channel:** 随机接入，Multiple access channel(多点接入)， such as : ALOHA、CSMA、CSMA/CD
- **controlled access channel:** 受控接入， such as : polling(轮询)

4.1 Carrier Sense Multiple Access Protocols(CSMA，载波侦听多路访问协议)

First sense, second send (先听后发)

1. 1-persistent CSMA: 1－坚持 CSMA
2. 0-persistent CSMA: 0－坚持 CSMA
3. p-persistent CSMA: p－坚持 CSMA



Question: 如果一个站在发送数据之前已经监听到信道为空闲，那么信道是否一定空闲呢？

4.2 CSMA/CD (CSMA with Collision Detection, 带冲突检测的 CSMA)

First sense, Second send (先听后发), simultaneously sense and send (边听边发)

1. Collision Detection:

(1) 比较接收到的信号的电压大小。若接收到的信号的电压值超过某一门限值，就可认为是发生了冲突。

(2) 当采用曼彻斯特编码时，电压的过零点是在每一比特的正中央。发生冲突时，过零点的位置将改变。根据过零点位置的变化，也可以判断是否发生了冲突。

(3) 在发送帧时也同时进行接收，将收到的信号逐比特地与发送的比特相比较。若有不符合的，就说明有冲突存在。

2. The Binary Exponential Backoff Algorithm (二进制指数后退算法)

即冲突后随机等待延迟的确定。设基本退避时间等于两站点之间的最大传播时延的两倍。定义 k=min[重传次数，10]，然后从离散的整数集合[0，1，…，2^{k-1}]中随机地取出一个数，记为 r，该站点的等待时延就是 r 倍的基本退避时间。当重传 16 次仍不能成功时，则丢弃该帧。

4.3 IEEE 802 and Ethernet

4.3.1 IEEE 802

Number	Topic
802.1	Overview and architecture of LANs
802.2 ↓	Logical link control
802.3 *	Ethernet
802.4 ↓	Token bus (was briefly used in manufacturing plants)
802.5	Token ring (IBM's entry into the LAN world)
802.6 ↓	Dual queue dual bus (early metropolitan area network)
802.7 ↓	Technical advisory group on broadband technologies
802.8 †	Technical advisory group on fiber optic technologies
802.9 ↓	Isochronous LANs (for real-time applications)
802.10 ↓	Virtual LANs and security
802.11 *	Wireless LANs
802.12 ↓	Demand priority (Hewlett-Packard's AnyLAN)
802.13	Unlucky number. Nobody wanted it
802.14 ↓	Cable modems (defunct: an industry consortium got there first)
802.15 *	Personal area networks (Bluetooth)
802.16 *	Broadband wireless
802.17	Resilient packet ring

LAN Reference Model: 局域网参考模型

<table>
<tr><td rowspan="2">Data link layer</td><td>LLC</td></tr>
<tr><td>MAC</td></tr>
<tr><td colspan="2">Physical layer</td></tr>
</table>

(1) Physical layer: signal code and decode(信号的编码和译码), produce and wipe off front-SYN(产生和去除前同步码), send and receive bits(传输和接收比特)

(2) The MAC sub-layer: framing(成帧)，error control(bits)，locate address(寻址)

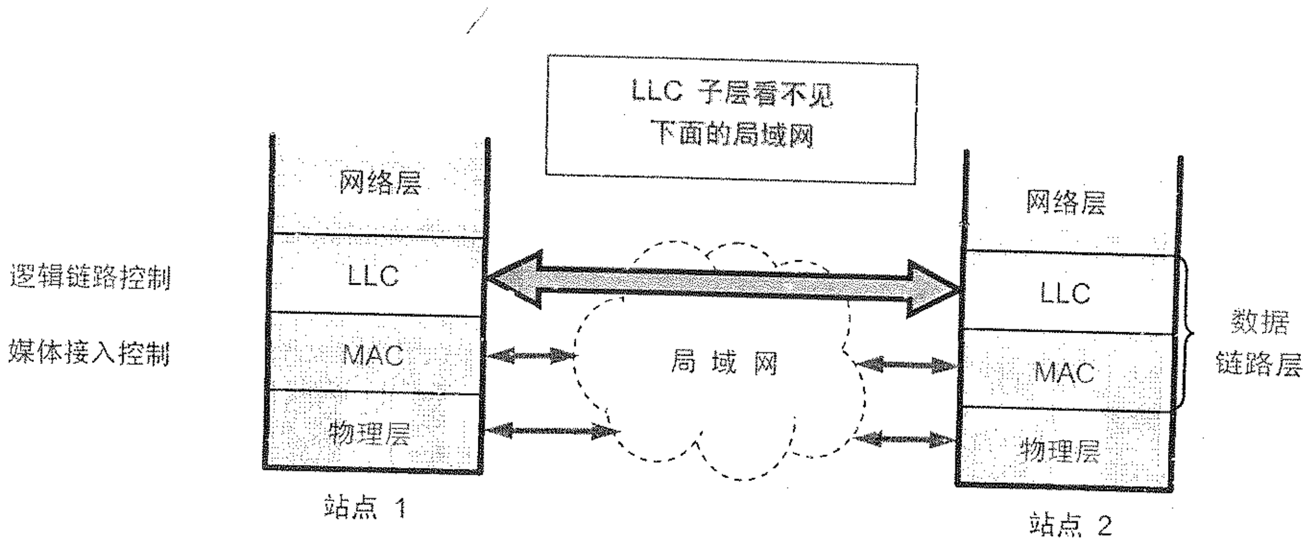
(3) The LLC sub-layer: Link management, provide interface, error control, numbering (编号)

LAN characteristic:

(1) Transmission media: Twisted pair, Coaxial cable, Fiber optics

(2) Topology structure: Star, Bus, Ring, Tree

(3) MAC: 介质访问控制方法, CSMA/CD, Token Ring, Token Bus

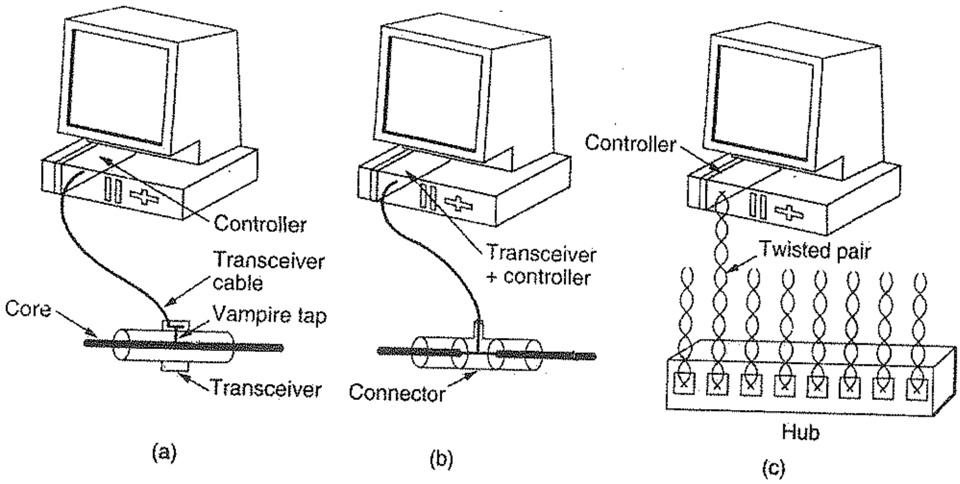


物理媒体	环形	总线	树形	星形
双绞线	★	★		★
基带同轴电缆	★	★		
宽带同轴电缆		★	★	
光纤	★	☆		★
无线媒体	★	★		★

4.3.2 IEEE 802.3 and Ethernet

IEEE 802.3: CSMA/CD, Ethernet: DIX----DEC, Intel, Xerox

Name	Cable	Max. seg.	Nodes/seg.	Advantages
10Base5	Thick coax	500 m	100	Original cable; now obsolete
10Base2	Thin coax	185 m	30	No hub needed
10Base-T	Twisted pair	100 m	1024	Cheapest system
10Base-F	Fiber optics	2000 m	1024	Best between buildings



1. Transceiver: 收发器

（1）Receive/send data：接收/发送数据

（2）Collision detection：检测冲突

（3）Carrier sense：载波侦听

（4）Jabber control：超长控制：当检测到某个数据帧的长度超过上限值时，即认为该站出了故障，接着就自动禁止该站向总线发送数据。

2. Transceiver cable: 收发器电缆，5 STP

（1）Two of the pairs: data in and data out

（2）Two of the pairs: control signals in and out

（3）One of the pairs: power cable

3. Network Interface card：网卡

（1）framing: 成帧，帧的封装和接封

（2）link management: 链路管理

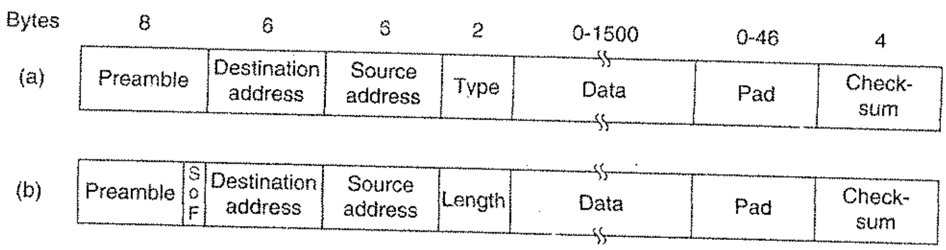
（3）code and decode: 编码和译码，Manchester Code(曼彻斯特编码)

IEEE802.3 10Mbps

	10BASE5	10BASE2	10BASE-T	10BROAD36	10BASE-FP	10BASE-FL	10BASE-FB
传输媒体	同轴电缆（50Ω）	同轴电缆（50Ω）	非屏蔽双绞线	同轴电缆（75Ω）	850mm光纤对	850mm光纤对	850mm光纤对
编码技术	基带（曼彻斯特编码）	基带（曼彻斯特编码）	基带（曼彻斯特编码）	宽带（DPSK）	曼彻斯特/on-of	曼彻斯特/on-of	曼彻斯特/on-of
拓扑结构	总线	总线	星形	总线/树形	星形	星形	/
最大网段长度（m）	500	185	100	1 800	1000	2000	2000
每网段的	100	30	/	/	33	/	/

节点数								
收发器	外置设备	内置芯片	内置芯片	/	内置芯片	内置芯片	内置芯片	
最大跨距媒体段数	2500m/5	925m/5	500m/5	3600m/2	2km/2	4km/2	30km/15	
电缆直径（mm）	10	5	0.4~1.0	0.4~1.0	62.5~125um	62.5~125um	62.5~125um	

4. 802.3 MAC sub-layer



(a) DIX Ethernet　　(b) IEEE 802.3

Notice:

(1) The minimum frame is 64 bytes, and the maximum frame is 1518 bytes.

(2) The Preamble field: **10101010 10101010…………10101011**

(3) The Invalid frame:

★ 数据字段的长度与长度字段的值不一致；

★ 帧的长度不是整数个字节；

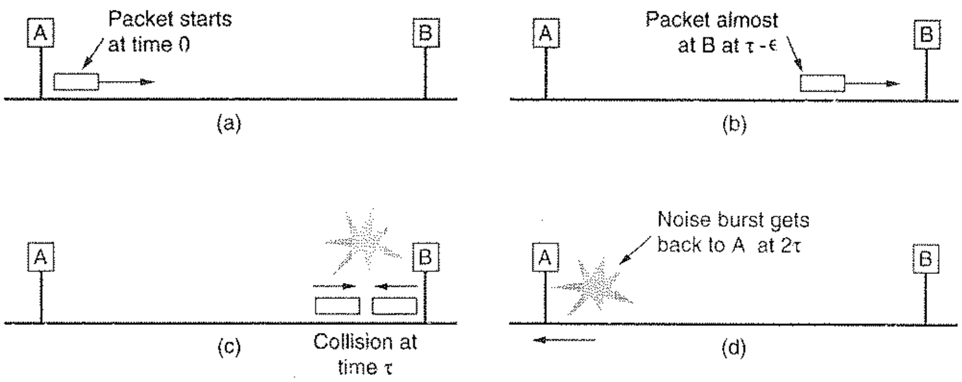
★ 用收到的帧检验序列 FCS 查出有差错；

★ 数据字段的长度不在 46 -- 1500 字节之间。

★ 有效的 MAC 帧长度为 64 -- 1518 字节之间。

★ 对于检查出的无效 MAC 帧就简单地丢弃。以太网不负责重传丢弃的帧

5. 以太网的最大作用距离（跨距）



假设跨距为 L，传播速率为 R1（注：电缆中，R1 一般固定为200m/us），最小帧长为 F，传输速率为 R2，物理层处理时间为 t_{PHY}，使用中继器的数量为 N，一个中继器的延时为 t_r，则**碰撞槽时间**（slot time）为：（注：碰撞槽时间即是在帧发送过程中，发生碰撞时间的上限。即在这段时间中，可能检测到碰撞，而过了这段时间，永远不会发生碰撞，也不会检测到碰撞。）

slot time $\approx 2L/R1+2t_{PHY}+2Nt_r$ ①

slot time=F/R2 ②

由式①、②可知：$L \approx R1（F/R2-2t_{PHY}-2Nt_r）/2$

由上式可知：若传输率 R2 固定，则跨距 L 越大，最小帧长度 F 越大。若最小帧长度 F 不变，则传输率 R2 越高，跨距 L 越小。

4.3.3 Ethernet Performance

Suppose: $a= \tau / T_0$ τ 为单程传播延迟，T_0 为帧的发送时间 = F/C

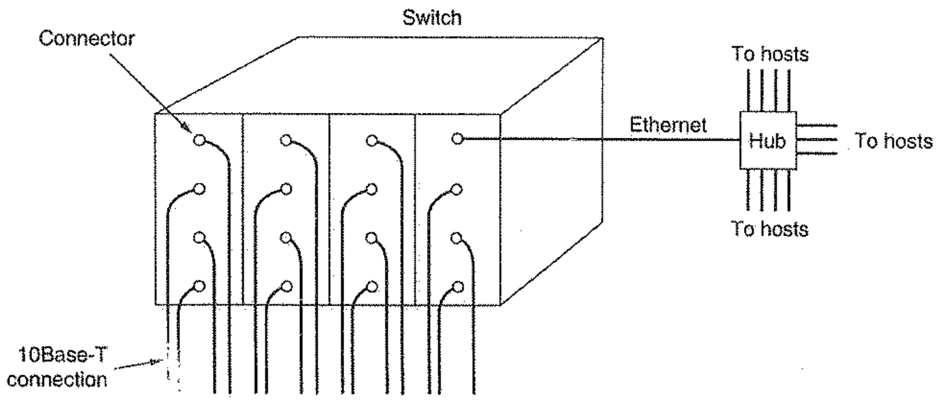
$= \tau C/F$

$S_{max} = 1/(1+4.44a)$ $S= N\lambda / \mu C$

$S_{lim} = T_0/(T_0+\tau)=1/(1+4.44a)$

4.4 Switched Ethernet(交换式以太网)

- ✧ full-duplex communication: 通常采用全双工通信
- ✧ switch and share: 交换和共享的区别



✧ data link layer: 交换机工作在数据链路层

✧ 对于普通 10 Mb/s 的共享式以太网，若共有 *N* 个用户，则每个用户占有的平均带宽只有总带宽(10 Mb/s)的 *N* 分之一。

✧ 使用以太网交换机时，虽然在每个端口到主机的带宽还是 10 Mb/s，但由于一个用户在通信时是独占而不是和其他网络用户共享传输媒体的带宽，因此对于拥有 *N* 对端口的交换机的总容量为 *N*10 Mb/s。这正是交换机的最大优点。

4.5 High-speed Ethernet

4.5.1 Fast Ethernet

Name	Cable	Max. segment	Advantages
100Base-T4	Twisted pair	100 m	Uses category 3 UTP
100Base-TX	Twisted pair	100 m	Full duplex at 100 Mbps (Cat 5 UTP)
100Base-FX	Fiber optics	2000 m	Full duplex at 100 Mbps; long runs

4.5.2 Gigabit Ethernet

Name	Cable	Max. segment	Advantages
1000Base-SX	Fiber optics	550 m	Multimode fiber (50, 62.5 microns)
1000Base-LX	Fiber optics	5000 m	Single (10 μ) or multimode (50, 62.5 μ)
1000Base-CX	2 Pairs of STP	25 m	Shielded twisted pair
1000Base-T	4 Pairs of UTP	100 m	Standard category 5 UTP

4.6 Wireless LANS

MACA: multiple access with collision avoidance，避免冲突的多路访问

MACAW: multiple access with collision avoidance for WLAN，无线环境下的避免冲突的多路访问

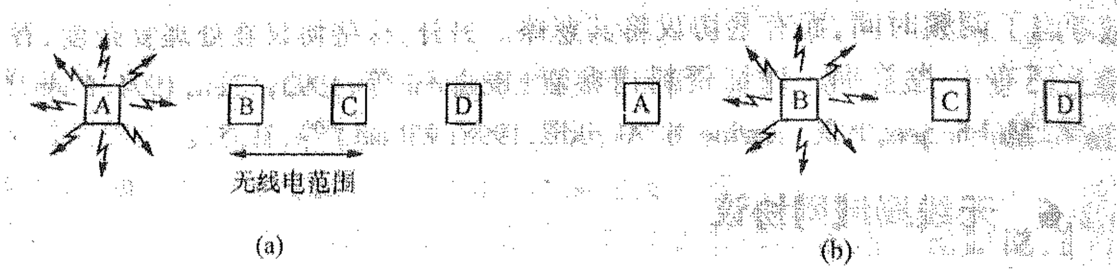


图 4-11　一个无线网

(a) A 发送；(b) B 发送。

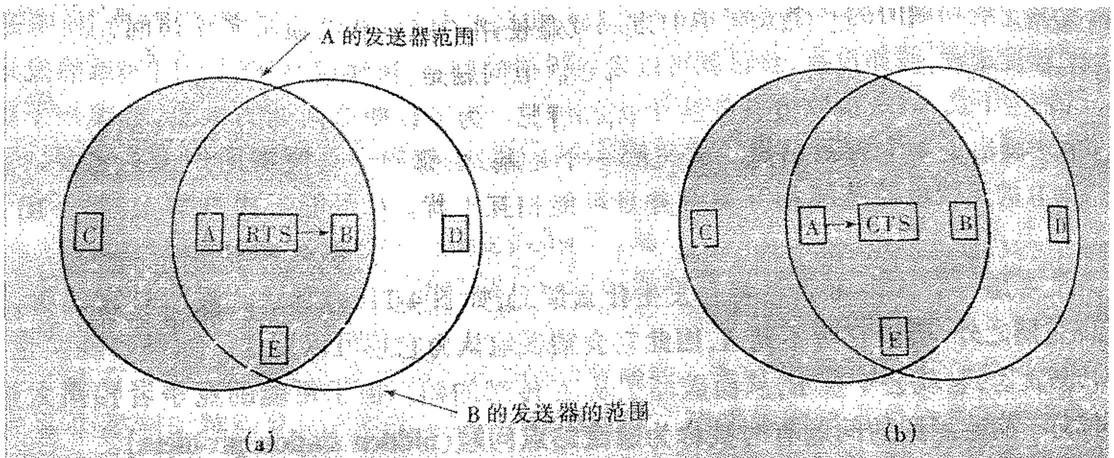


图 4-12　MACA 协议

(a) A 向 B 发送 RTS；(b) B 以 CTS 响应 A。

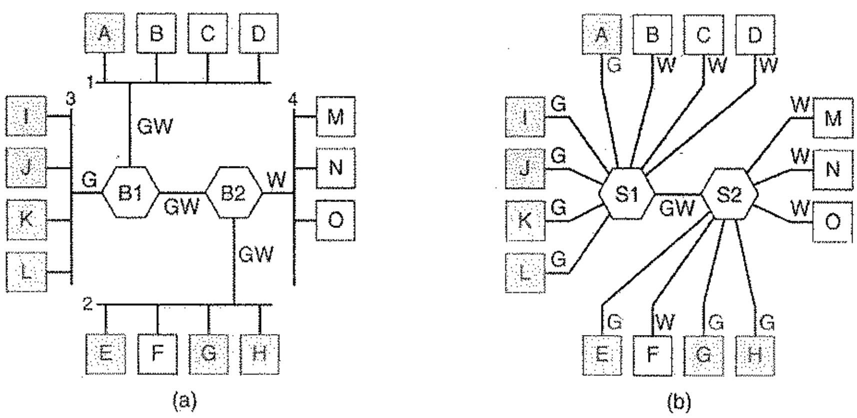
4.7　Virtual LAN（虚拟局域网）

How to assign a VLAN？ Three methods：

1. Every port is assigned a VLAN.(按端口划分 VLAN)

2. Every MAC address is assigned a VLAN. (按 MAC 地址划分 VLAN)

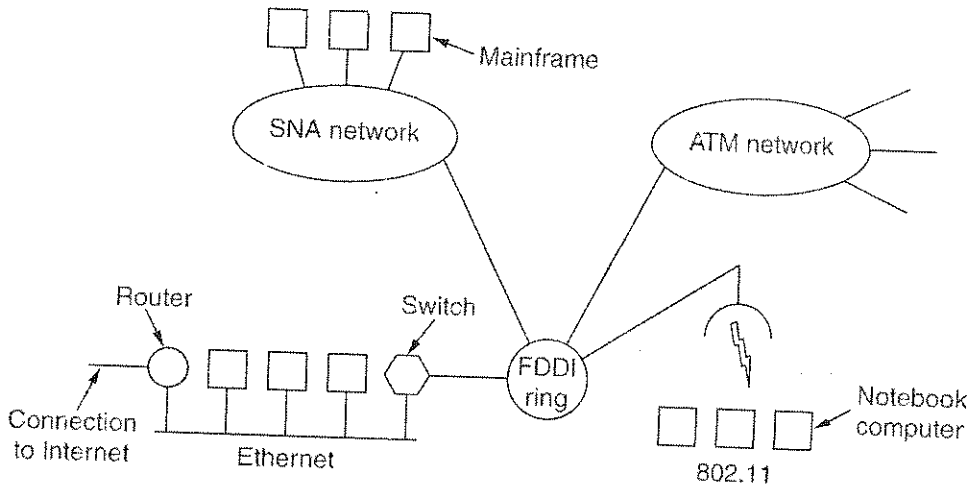
3. Every layer 3 protocol or IP address is assigned a VLAN. (按第 3 层地址或 IP 地址划分 VLAN)



4.8 Internetworking(网络互连)

(1) LAN-LAN
(2) LAN-WAN
(3) WAN-WAN
(4) LAN-WAN-LAN

Example:

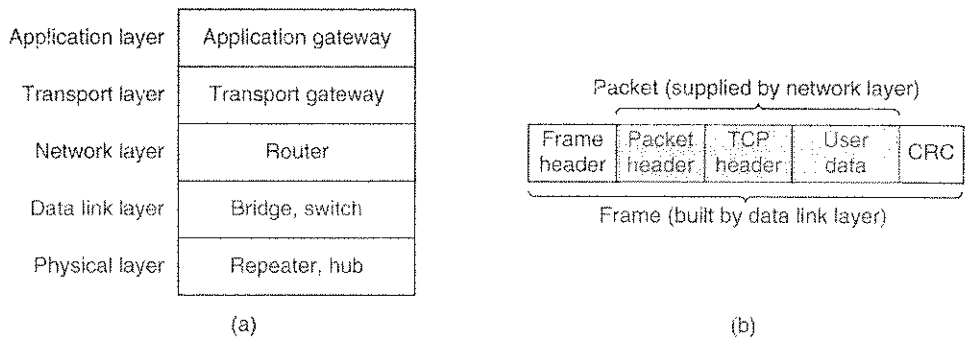


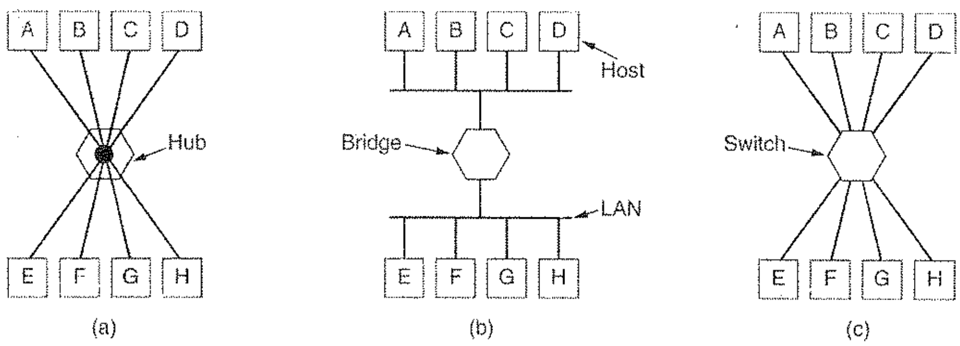
How networks differ?

Item	Some Possibilities
Service offered	Connection oriented versus connectionless
Protocols	IP, IPX, SNA, ATM, MPLS, AppleTalk, etc.
Addressing	Flat (802) versus hierarchical (IP)
Multicasting	Present or absent (also broadcasting)
Packet size	Every network has its own maximum
Quality of service	Present or absent; many different kinds
Error handling	Reliable, ordered, and unordered delivery
Flow control	Sliding window, rate control, other, or none
Congestion control	Leaky bucket, token bucket, RED, choke packets, etc.
Security	Privacy rules, encryption, etc.
Parameters	Different timeouts, flow specifications, etc.
Accounting	By connect time, by packet, by byte, or not at all

4.8.1 Interconnected equipment(互连设备)

Including: repeaters, hubs, bridges, switches, routers, gateways





1. repeater: **physical layer, analog device, volts.** 连接网络线路的一种装置，常用于两个网络节点之间物理信号的双向转发工作。中继器是最简单的网络互联设备，主要完成物理层的功能，负责在两个节点的物理层上按位传递信息，完成信号的复制、调整和放大功能，以此来延长网络的长度。

Example: 5-4-3 规则（适用于 10M 以太网）：

A. 在一个以太网中，最多允许有 5 个网段；

B. 在一个以太网中，最多允许有 4 个中继器；

C. 其中，最多允许有 3 个网段可以有节点；

2. hub: **physical layer.** 集线器（Hub）是中继器的一种形式，区别在于集线器能够提供多端口服务，也称为多口中继器。

3. bridge: **data link layer，frame.**

- Spanning tree bridge（生成树网桥）: transparent bridge,透明网桥。

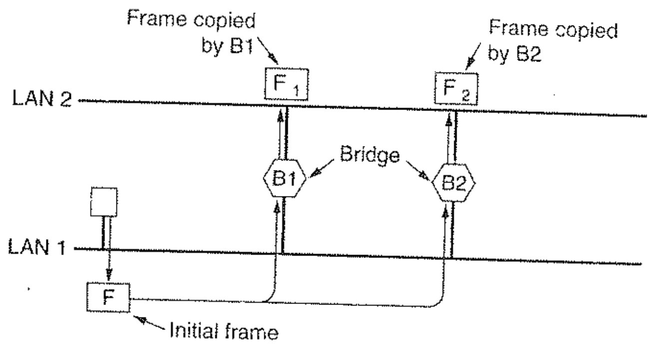
(1) If destination and source LANs are the same, discard the frame(丢弃帧)

(2) If the destination and source LANs are different, forward the

frame(转发帧)

(3) If the destination LAN is unknown, use flooding(洪泛法)

Because of, a problem can be seen in the below figure, the actual topology must be built the spanning tree. So we need the RSTP.



RSTP: IEEE 802.1w, Rapid Spanning Tree Protocol, 快速生成树协议。

RTSP: real-time stream protocol, 实时流协议。（**注意区分**）

- Source route bridge: 源站选路网桥，源站在发送帧时将详细的路由信息放在帧的首部中。

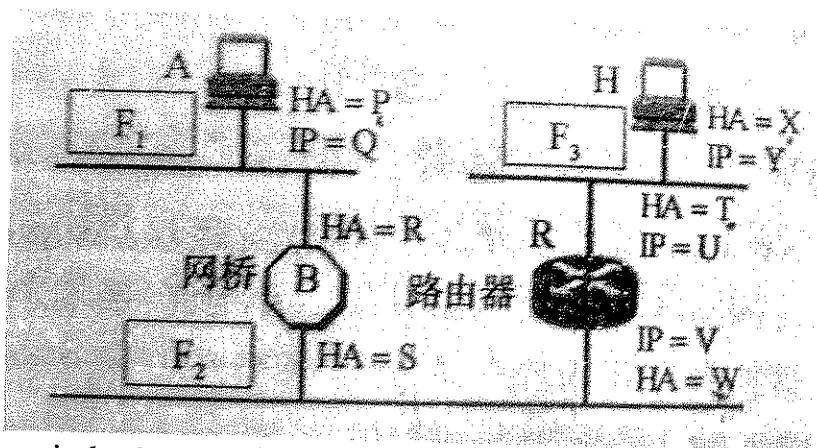
Question: 源站怎样才能知道应当选择什么样的路由？

Answer: 源站以广播方式向欲通信的目的站发送一个**发现帧**（discovery frame）作为探测之用。

Notice: 网桥并不改变它转发的帧的源地址。

Example:

	D-HA	S-HA	D-IP	S-IP
F1	W	P	Y	Q
F2	W	P	Y	Q
F3	X	T	Y	Q



4. switch: **data link layer，frame. 多端口网桥。**

（1）对于普通 10 Mb/s 的共享式以太网，若共有 N 个用户，则每个用户占有的平均带宽只有总带宽(10 Mb/s)的 N 分之一。

（2）使用以太网交换机时，虽然在每个端口到主机的带宽还是 10 Mb/s，但由于一个用户在通信时是独占而不是和其他网络用户共享传输媒体的带宽，因此对于拥有 N 对端口的交换机的总容量为 N10 Mb/s。这正是交换机的最大优点。

（3）store-and-forward switch: 存储转发交换机。

（4）cut-through switch: 直通交换机。

5. router: **network layer, packet.**

作用：（1）选择最佳的转发数据的路径，建立非常灵活的连接，均衡网络负载。

（2）利用通信协议本身的流量控制功能来控制数据传输，有效地解决拥挤问题。

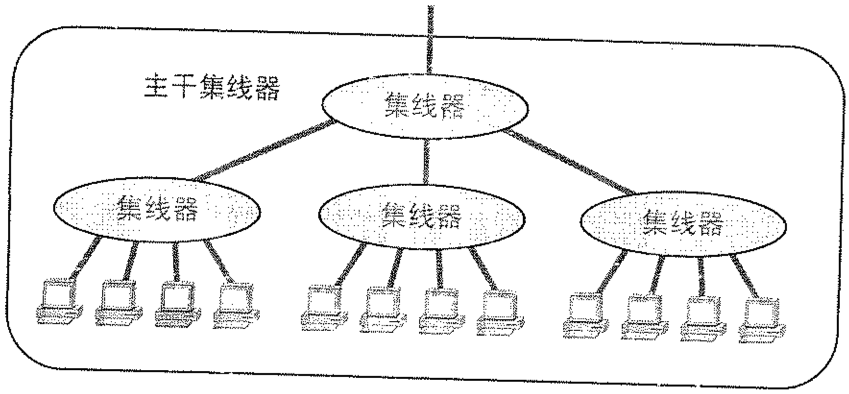
（3）具有判断哪些数据分组需要转发的功能，不仅可根据 LAN 网络地址和协议类型，而且可根据网间地址、主机地址、数据类型（如文件传输、远程登录或电子邮件）等，判断分组是否应该转发。将不该转发的信息（包括错误信息）过滤掉，从而避免了广播风暴。

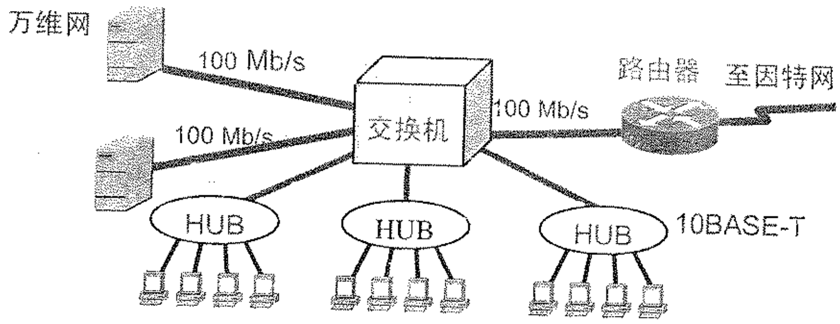
（4）把一个大的网络划分为若干个子网。

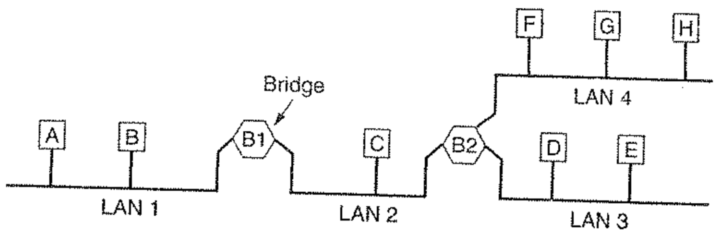
6. gateway: **transport layer and application layer, message.**

互连设备	优点	缺点
中继器	网络互连容易，价格低	互连规模有限，不能隔离不需要流量， 无法控制信息传输
网桥	网络互连容易，提供网络管理功能， 提供安全保密性，协议透明， 隔离不必要流量，交换效率高	具有广播风暴现象，不能决定最佳路径， 管理控制功能有限，错误处理功能不强， 不能完全隔离不必要的流量
路由器	适于大规模的复杂网络，容易互连不同类型局域网，管理控制功能强，安全性、保密性好，充分隔离不需要的流量	网络设置复杂 不支持非路由协议（如 DECnet LAT，Netbios 等） 价格昂贵

4.8.2 Local Internetworking(局域网互连)







5 The Network Layer

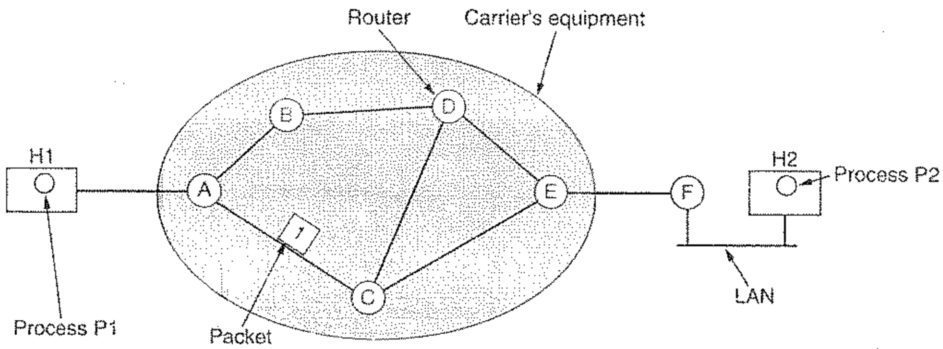
学习目的：了解网络层向高层所提供的服务类型，掌握两种路由选择算法的原理，理解拥塞控制的概念，掌握 **IP** 协议的相关知识。

学习重点：**IP** 协议、**IP** 地址的相关知识。

学习难点：两种路由选择算法的原理和应用。

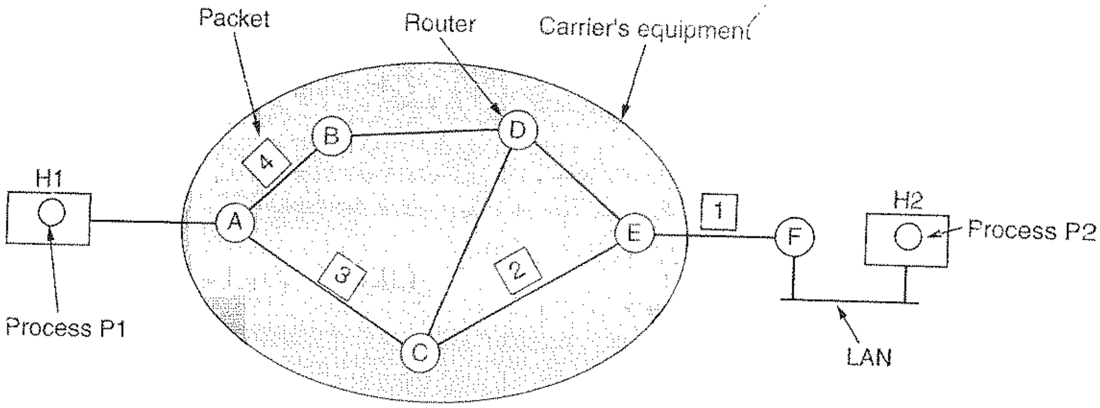
5.1 Network layer design issues

5.1.1 Store-and-forward packet switching(存储转发分组交换)



5.1.2 Services provided to the transport layer

1. Connectionless service(无连接服务)



A's table

initially

Dest.	Line
A	–
B	B
C	C
D	B
E	C
F	C

later

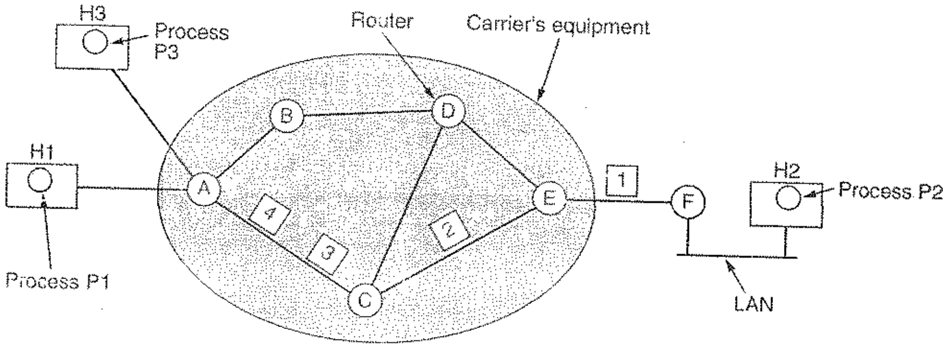
A	–
B	B
C	C
D	B
E	B
F	B

C's table

A	A
B	A
C	–
D	D
E	E
F	E

E's table

A	C
B	D
C	C
D	D
E	–
F	F



A's table

In		Out	
H1	1	C	1
H3	1	C	2

C's table

A	1	E	1
A	2	E	2

E's table

C	1	F	1
C	2	F	2

2. Connection-oriented service(面向连接服务)

3. Comparison of virtual-circuit and datagram subnets

Issue	Datagram subnet	Virtual-circuit subnet
Circuit setup	Not needed	Required
Addressing	Each packet contains the full source and destination address	Each packet contains a short VC number
State information	Routers do not hold state information about connections	Each VC requires router table space per connection
Routing	Each packet is routed independently	Route chosen when VC is set up; all packets follow it
Effect of router failures	None, except for packets lost during the crash	All VCs that passed through the failed router are terminated
Quality of service	Difficult	Easy if enough resources can be allocated in advance for each VC
Congestion control	Difficult	Easy if enough resources can be allocated in advance for each VC

项目类型	数据报子网	虚电路子网
电路设置	不需要	需要
地址	每个分组都有源端和目的端的完整地址	每个分组都含有一个短的虚电路号
状态信息	子网不存储状态信息	建立好的每条虚电路都要求占用子网表空间
路由选择	对每个分组独立选择	当虚电路建好时，路由就已确定，所有分组都经过此路由
路由失败的影响	除了在崩溃时全丢失分组外，无其他影响	所有经过失效路由器的虚电路都要被终止
服务质量	难	如果有足够的资源分配给已经建立的每条虚电路，则容易控制
拥塞控制	难	如果有足够的资源分配给已经建立的每条虚电路，则容易控制

5.2 Routing Algorithms

Properties：

（1）correctness(正确性)

（2）simplicity（简单性）

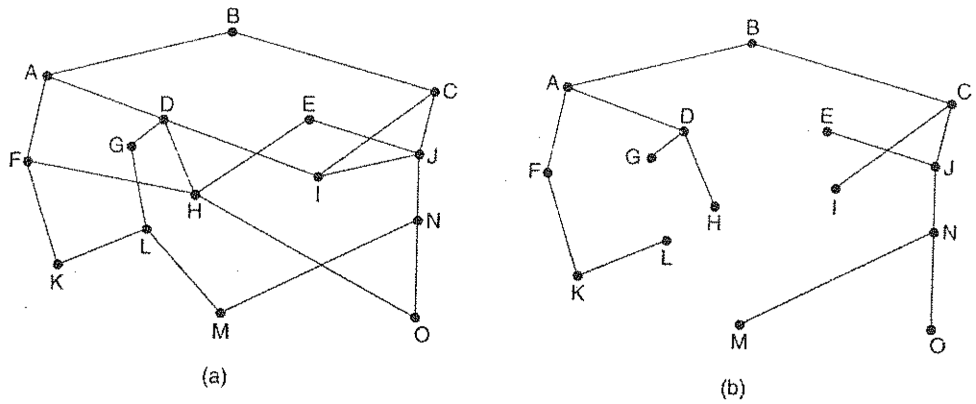
（3）robustness（健壮性）

（4）stability（稳定性）
（5）fairness（公平性）
（6）optimality（最优性）

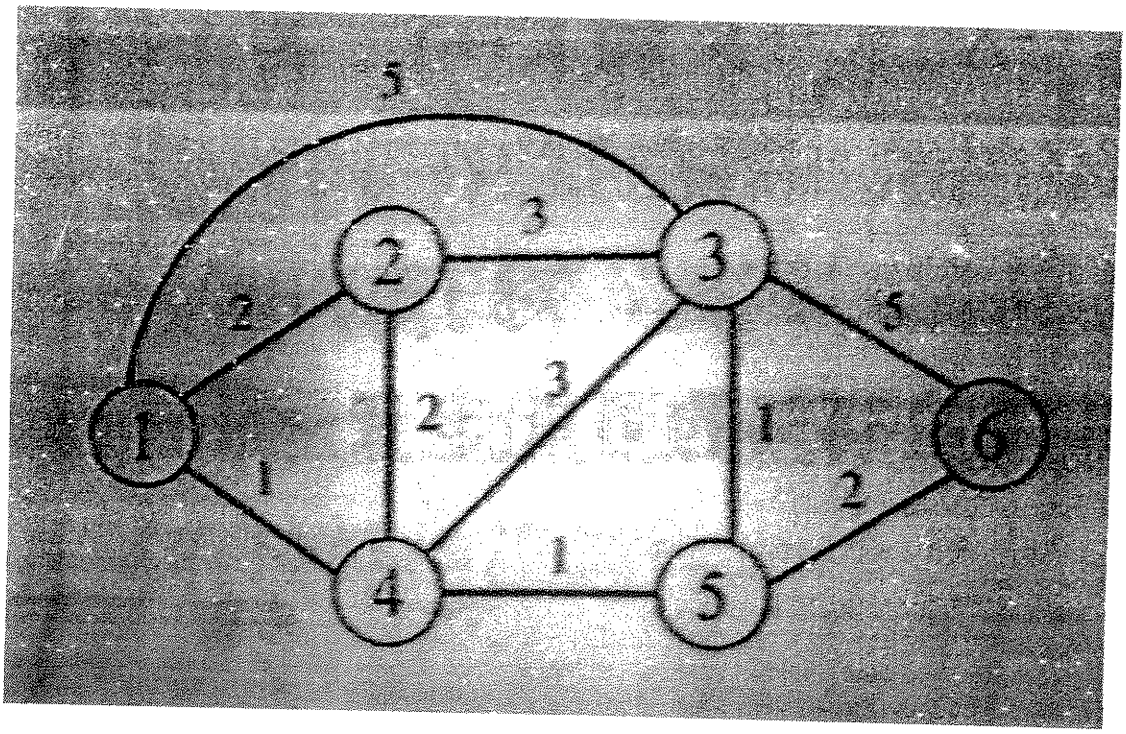
Two major classes：
（1）Nonadaptive algorithms（非自适应算法）：static routing，静态路由选择
（2）adaptive algorithms（自适应算法）：dynamic routing，动态路由选择

5.2.1 The optimality principle（最优化原则）

如果路由器 J 在从路由器 I 到 K 的最佳路由上，那么从 J 到 K 的最佳路由就会在同一路由之中。



5.2.2 Shortest path routing(static routing)

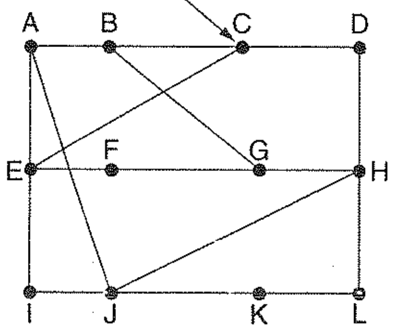


shortest path

Step	N	D(2)	D(3)	D(4)	D(5)	D(6)
Initialization	{1}	2	5	1	∞	∞
1	{1,4}	2	4	1	2	∞
2	{1,4,5}	2	3	1	2	4
3	{1,2,4,5}	2	3	1	2	4
4	{1,2,3,4,5}	2	3	1	2	4
5	{1,2,3,4,5,6}	2	3	1	2	4

5.2.3 Distance vector routing(dynamic routing)距离向量路由选择

1. Example:



(a)

To	A	I	H	K	New estimated delay from J	Line
A	0	24	20	21	8	A
B	12	36	31	28	20	A
C	25	18	19	36	28	I
D	40	27	8	24	20	H
E	14	7	30	22	17	I
F	23	20	19	40	30	I
G	18	31	6	31	18	H
H	17	20	0	19	12	H
I	21	0	14	22	10	I
J	9	11	7	10	0	–
K	24	22	22	0	6	K
L	29	33	9	9	15	K
	JA delay is 8	JI delay is 10	JH delay is 12	JK delay is 6	New routing table for J	

Vectors received from J's four neighbors

(b)

2. The count-to-infinity problem: 无穷计数问题

距离向量算法的缺点：好消息传得快，坏消息传得慢。

A	B	C	D	E	
	•	•	•	•	Initially
	1	•	•	•	After 1 exchange
	1	2	•	•	After 2 exchanges
	1	2	3	•	After 3 exchanges
	1	2	3	4	After 4 exchanges

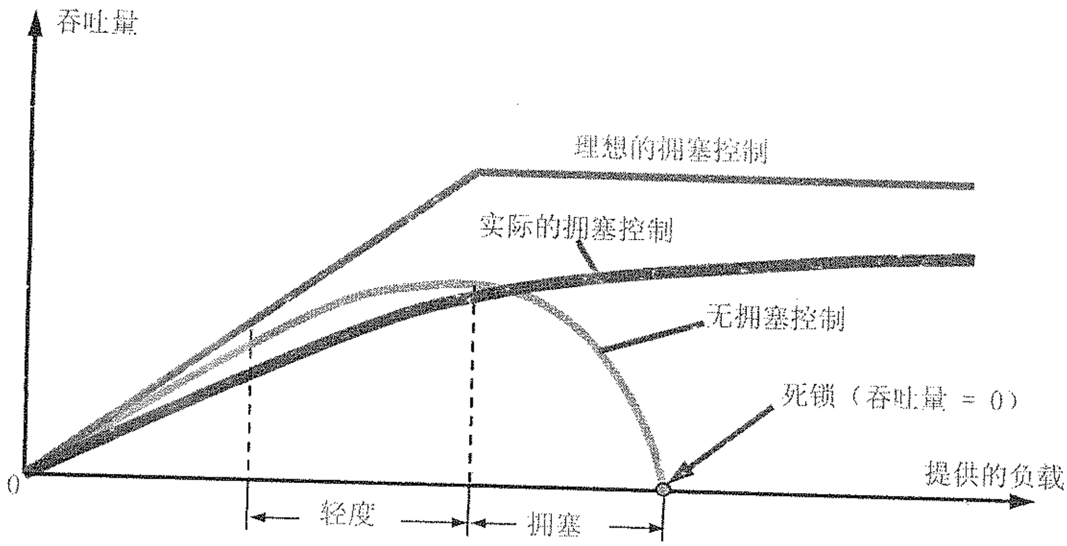
(a)

A	B	C	D	E	
	1	2	3	4	Initially
	3	2	3	4	After 1 exchange
	3	4	3	4	After 2 exchanges
	5	4	5	4	After 3 exchanges
	5	6	5	6	After 4 exchanges
	7	6	7	6	After 5 exchanges
	7	8	7	8	After 6 exchanges
		⋮			
	•	•	•	•	

(b)

5.3 Congestion control algorithms(拥塞控制算法)

Reason: 需求的资源数>网络可利用的资源数



Question： 当网络负载较小时，有拥塞控制的吞吐量反而比无拥塞控制时要小。**为什么？**

举例说明（红绿灯的问题）。

Notice: 拥塞控制与流量控制的关系

(1) 拥塞控制是一个全局性的过程，涉及到所有的主机、所有的路由器，以及与降低网络传输性能有关的所有因素。

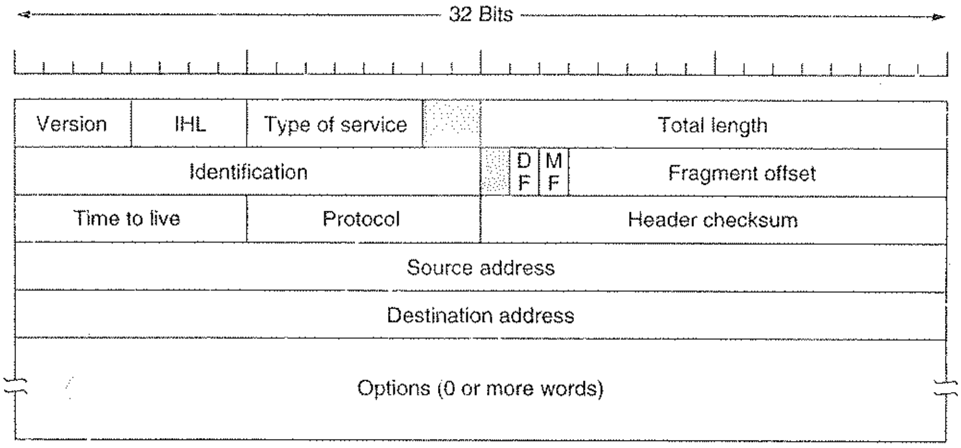
(2) 流量控制往往是指在给定的发送端和接收端之间的点对点通信量。流量控制所要做的就是使发送端发送数据的速率不能使接收端来不及接收。流量控制几乎总是存在着从接收端到发送端的某种直接反馈，使发送端知道接收端是处于怎样的状况。

Example: 一个光纤网络，其传送能力为 1000Gb/s。一台超级计算机想以 1Gb/s 的速度向一台个人计算机发送一个文件。虽然完全不存在拥塞问题，但也需要流量控制来促使超级计算机不断地停发数据，以便给个人计算机处理已收到的数据。

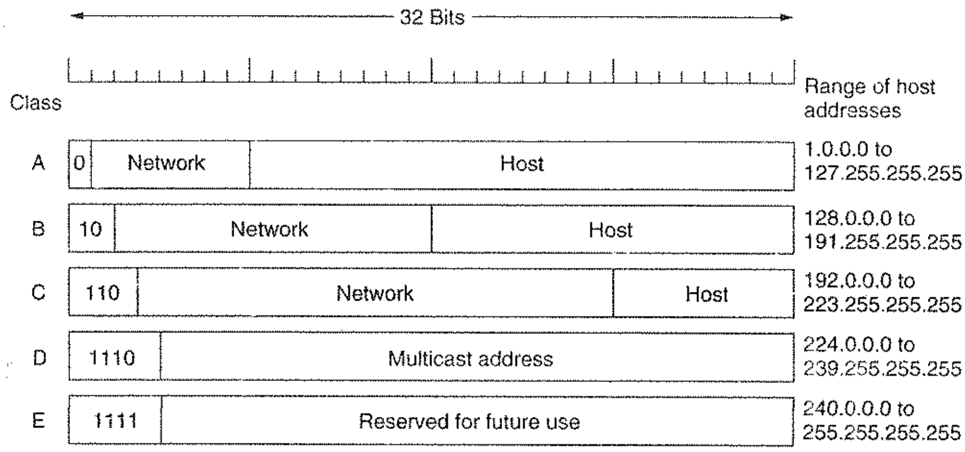
一个具有 1Mb/s 线路和 1000 台计算机的存储一转发网络，其中一半计算机想以 100kb/s 给另一半传送文件，这时，就需要拥塞控制，因为总的通信量可能会超过网络的处理能力。

5.4 The network layer in the internet

5.4.1 The IP protocol



5.4.2 IP Addresses



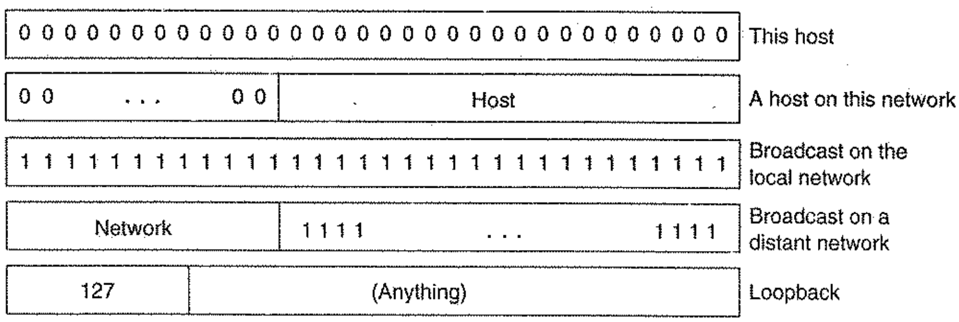
Private IP addresses:

10.0.0.0 --- 10.255.255.255/8 (16777216 hosts)

172.16.0.0 --- 172.31.255.255/12 (1048576 hosts)

192.168.0.0 --- 192.168.255.255/16 (65536 hosts)

Special IP addresses:



IP characteristics:

（1）IP 地址是一种分等级的地址结构。

（2）IP 地址标志一个主机（或路由器）和一条链路的接口。

（3）当一个主机同时具有两个以上的 IP 地址时，这种主机称为多接口主机。

（4）用转发器或网桥连接起来的若干个局域网仍为一个网络。

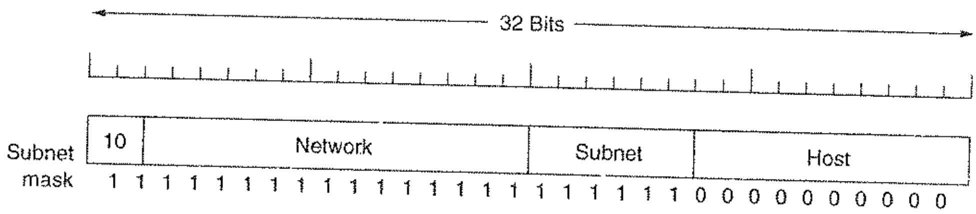
（5）所有分配到网络号 net－id 的网络都是平等的。

ARP/RAPR:地址解析协议、逆向地址解析协议，使用广播方式完成地址转换。

ICMP:因特网控制报文协议，PING 工具

IGMP:因特网组管理协议

Subnets: 子网



Example:

1.若子网掩码为 255.255.255.224（11111111 11111111 11111111 11100000），IP 地址为 140.252.20.68。由于这是一个 B 类地址，因此网络号为 140.252。而 20.68 即为 00010100 01000100，即子网号为前 11 位，而主机号占 5 位。子网号=00010100 010=162，而主机号=00100=4。

2. 若有一个 C 类网，需划分为 6 个子网，每个子网拥有不超过 30 台计算机，请设计该网络的子网掩码，并写出每个子网的主机地址范围。

子网掩码：255.255.255.224

每个子网的主机地址范围： 001 00001~001 11110；

010 00001~010 11110； 011 00001~011 11110；

100 00001~100 11110； 101 00001~101 11110；

110 00001~110 11110；

CIDR (Classless InterDomain Routing). The basic idea behind CIDR, which is described in RFC 1519, is to allocate the remaining IP addresses in variable-sized blocks, without regard to the classes. If a site needs, say, 2000 addresses, it is given a block of 2048 addresses on a 2048-byte boundary.

University	First address	Last address	How many	Written as
Cambridge	194.24.0.0	194.24.7.255	2048	194.24.0.0/21
Edinburgh	194.24.8.0	194.24.11.255	1024	194.24.8.0/22
(Available)	194.24.12.0	194.24.15.255	1024	194.24.12/22
Oxford	194.24.16.0	194.24.31.255	4096	194.24.16.0/20

The routing tables all over the world are now updated with the three assigned entries. Each entry contains a base address and a subnet mask. These entries (in binary) are:

```
Address Mask
C: 11000010 00011000 00000000 00000000 11111111
11111111 11111000 00000000
E: 11000010 00011000 00001000 00000000 11111111
11111111 11111100 00000000
O: 11000010 00011000 00010000 00000000 11111111
11111111 11110000 00000000
```

Now consider what happens when a packet comes in addressed to 194.24.17.4, which in binary is represented as the following 32-bit string

11000010 00011000 00010001 00000100

First it is Boolean ANDed with the Cambridge mask to get

11000010 00011000 00010000 00000000

This value does not match the Cambridge base address, so the original address is next ANDed with the Edinburgh mask to get

11000010 00011000 00010000 00000000

This value does not match the Edinburgh base address, so Oxford is tried next, yielding 11000010 00011000 00010000 00000000

This value does match the Oxford base. If no longer matches are found farther down the table, the Oxford entry is used and the packet is sent along the line named in it.

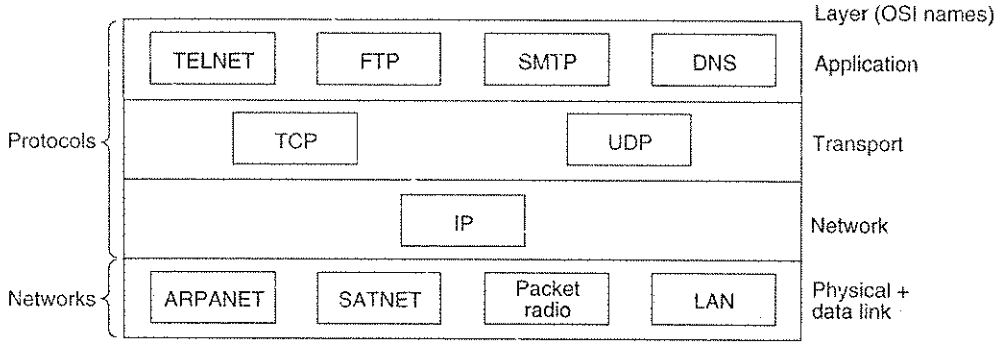
6 The transport layer

学习目的：了解传输层的基本概念，了解 UDP 协议和端口的应用，理解和领会 TCP 协议的基本原理和算法应用。

学习重点：TCP 协议的基本原理

学习难点：TCP 协议的算法应用。

6.1 The transport service



TCP: an reliable, **connection-oriented**, full-duplex communication protocol

UDP: an unreliable, **connectionless protocol**

Question 1: If the transport layer service is so similar to the network layer service, why are there two distinct layers?（既然传输层服务与网络层服务如此类似，那为什么还要将它区分为不同的两层呢？）

Question 2: UDP 用户数据报与 IP 数据报、TCP 的连接与网络层中的虚电路有什么不同？

Application	Application layer protocol	Transport layer protocol
名字转换	DNS	UDP/TCP
文件传送	TFTP	UDP
路由选择协议	RIP	UDP
IP 地址配置	BOOTP, DHCP	UDP
网络管理	SNMP	UDP

远程文件服务器	NFS	UDP
IP 电话	专用协议	UDP
流式多媒体通信	专用协议	UDP
多播	IGMP	UDP
电子邮件	SMTP	TCP
远程终端接入	TELNET	TCP
万维网	HTTP	TCP
文件传送	FTP	TCP

6.2 Socket and port

1. **port:** TSAP, 端口

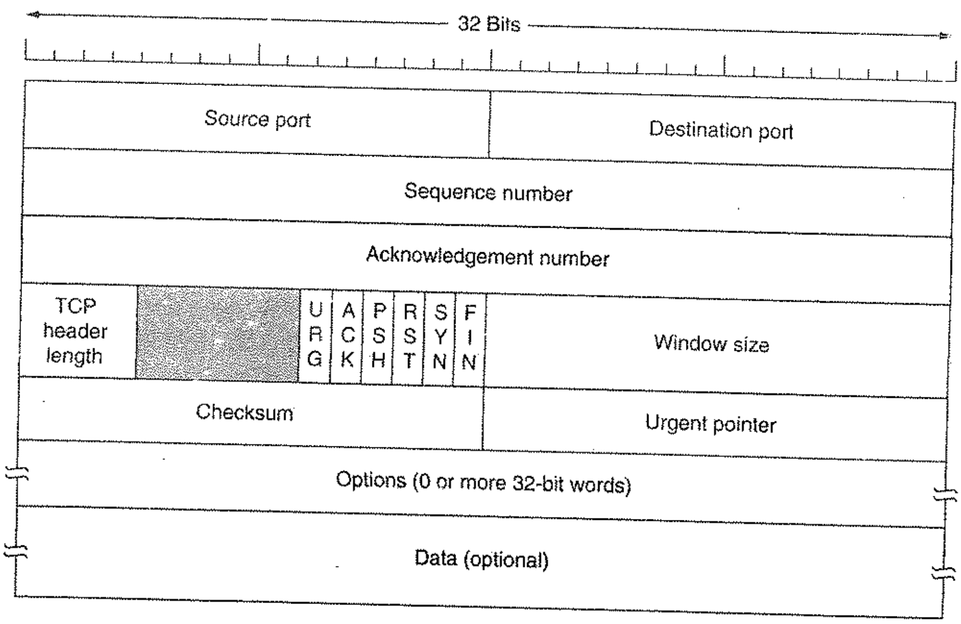
Well－known port: 熟知端口，0～1023

Application	FTP	TELNET	SMTP	DNS	TFTP	HTTP	SNMP
Well－known port	21	23	25	53	69	80	161

2. **Socket:** =(IP, Port)=(202.193.64.33, 80)

6.5 TCP

6.5.1 The TCP protocol

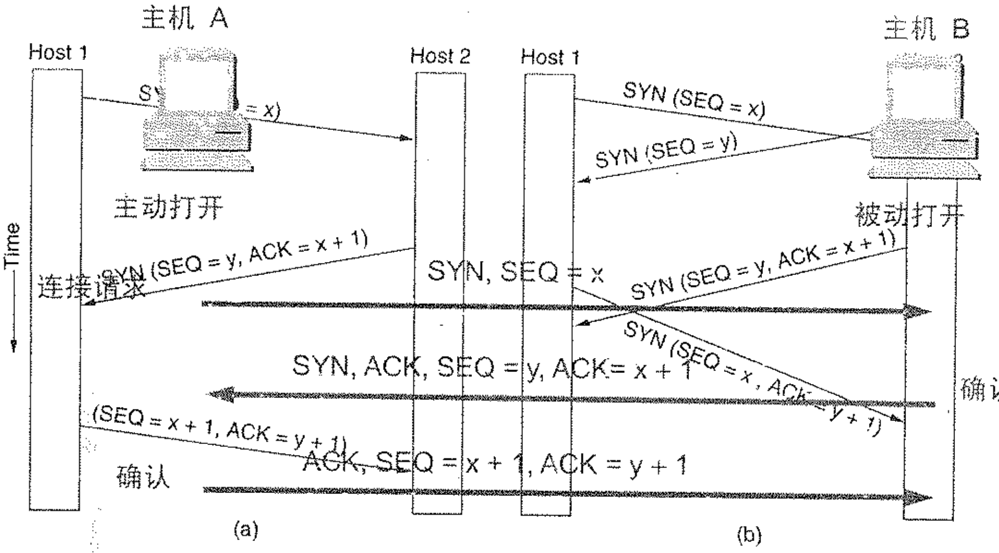


6.5.2 TCP connection establishment and release

1. TCP connection establishment

Three-way handshake: 三次握手

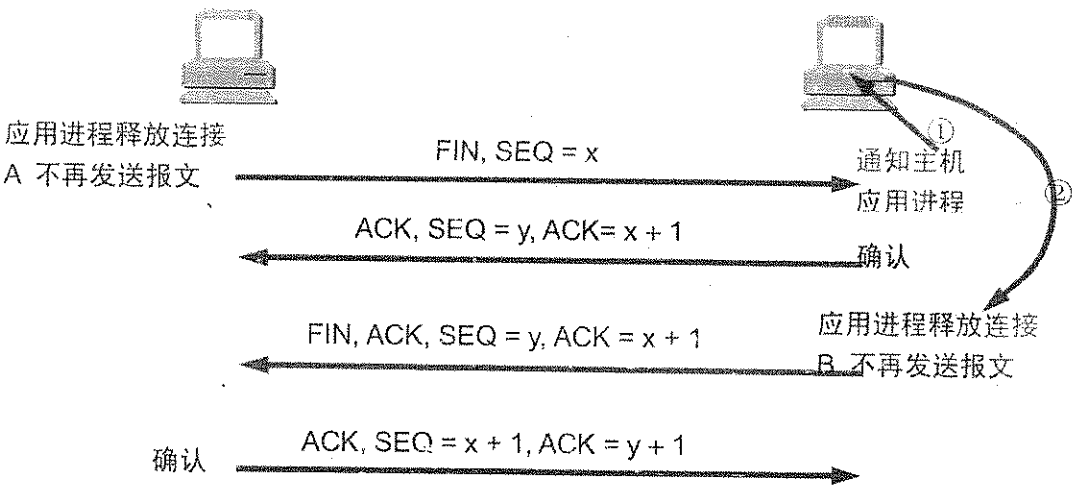
Question: why to use three-way handshake? 主要是为了防止已失效的连接请求报文段突然又传送到了主机 B，因而产生错误。举例说明。



序号 x，表明主机 A 向主机 B 传送数据时的第一个数据字节的序号是 x＋1

序号 y，表明主机 B 向主机 A 传送数据时的第一个数据字节的序号是 y＋1

2. TCP connection release



序号 x 等于主机 A 向主机 B 传送数据时的最后一个字节的序号加 1

序号 y 等于主机 B 向主机 A 传送数据时的最后一个字节的序号加 1

6.5.3 TCP 的编号与确认

1．编号：TCP 对所要传送的报文中的每一个字节编一个序号。在建立连接时，双方要商定初始序号。

2．确认：TCP 的确认是对接收到的数据的最高序号（即收到的数据

流中的最后一个序号）进行确认。但返回的确认序号是已收到的数据的最高序号加 1，即确认序号表示期望下次收到的第一个数据字节的序号。

3．TCP 的确认是对一段报文的确认，而不是对一个字节的确认，这是因为：假设用户只发一个字节，加上 20 字节的首部后，得到 21 字节长的 TCP 报文段。再加上 20 字节的 IP 首部，形成 41 字节长的 IP 数据报。在接收端 TCP 发出确认，构成的数据报是 40 字节长。若用户要求远地主机回送这一字符，则用户仅发一个字符，线路上就需传送总长度为 162 字节共 4 个报文段（包括用户端对回送字符的确认）。当线路带宽并不富裕时，这种传送方法的效率很低。

4．糊涂窗口综合症：若接收端的缓存已满，而交互式的应用进程依次只从缓存中读取一个字符（这样就在缓存产生 1 个字节的空位子），然后想发送端发送确认，并通知窗口为 1 个字节（但发送的数据报是 40 字节长）。接着，发送端又发来 1 个字符（但发来的数据报是 41 字节长）。接收端发回确认，仍然通知窗口为 1 个字节。这样进行下去，使网络的效率很低。

6.5.4 TCP Flow Control

TCP 采用可变发送窗口的方式进行流量控制。窗口大小的单位是字节。在 TCP 报文段首部的窗口字段写入的数值就是当前设定的接受窗口数值。

发送窗口在连接建立时由双方商定。但在通信的过程中，接收端可根据自己的资源情况，随时动态的调整自己的接受窗口（可增大或减小），然后告诉对方，使对方的发送窗口和自己的接收窗口一致。这种由接收端控制发送端的做法，在计算机网络中经常使用。

接 收 端 许 诺 的 发 送 窗 口

收到确认即可前移

1	101	201	301	401	501	601	701	801
100	200	300	400	500	600	700	800	900

已发送并被确认　　已发送但未被确认　　还可继续发送

不可发送

TCP 中的窗口概念

例：如下图利用可变窗口大小进行流量控制。

主机 A　　主机 B

SEQ=1 → A 还能发送

300 字节

SEQ=101 → A 还能发送

200 字节

SEQ=201 丢失 →

← ACK=201，WIN=300 允许 A 再发

送 300 字节

SEQ=301 → A 还能发送

200 字节

SEQ=401 → A 还能发送

100 字节

SEQ=201 → A 超时重

发，但不能再发送

← ACK=501，WIN=200 允许 A 再发

送 200 字节

SEQ=501 → A 还能发送

100 字节

← ACK=601，WIN=0 不允许 A 再

发送

6.5.5 TCP Congestion Control

1. Receiver window: 接收端窗口 rwnd，是来自接收端的流量控制
2. congestion window: 拥塞窗口 cwnd，是来自发送端的流量控制

发送窗口的上限值＝Min【rwnd，cwnd】

7 The Application Layer

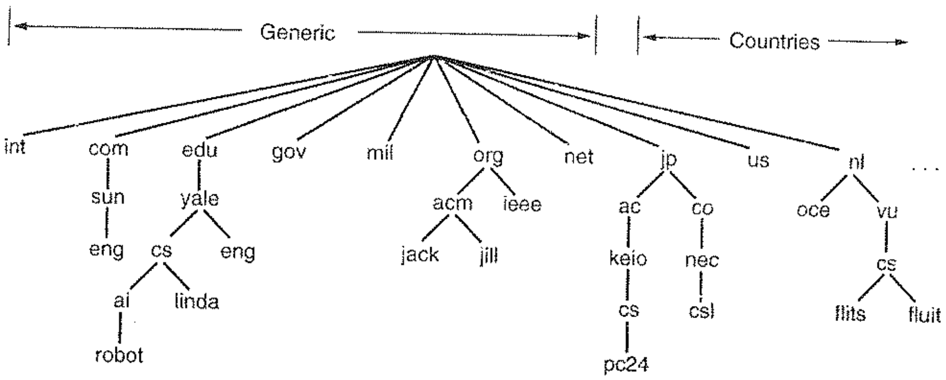
学习目的：了解应用层的基本概念，理解 **DNS**、**E-mail**、**FTP** 等协议的工作原理。

学习重点：**DNS**，**E-mail**、**FTP**、**WWW**

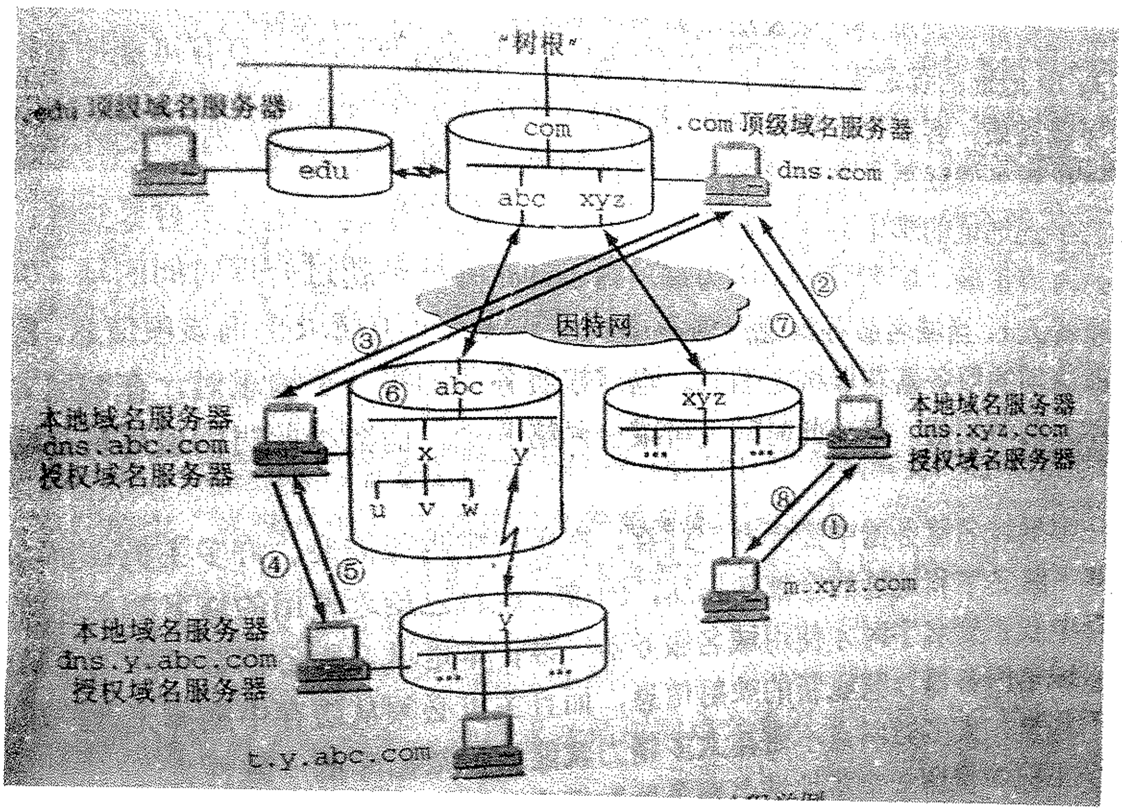
学习难点：无

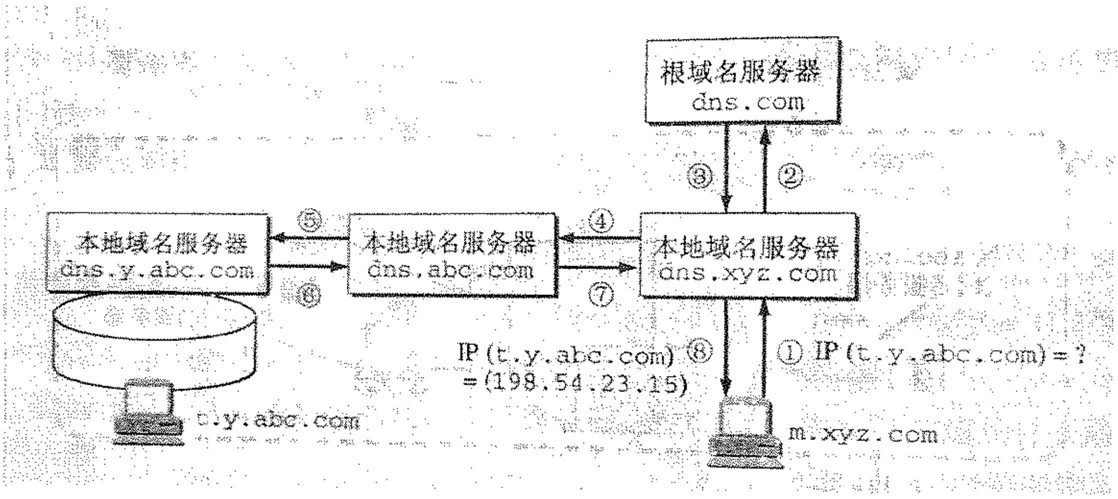
7.1 DNS—The Domain Name System

1. generic: 通用顶级域名，com，edu，gov，int，mil，net，org
2. countries: 国家顶级域名，cn，us，jp，de



3. domain name resolution: 域名解析
   recursive query：递归查询





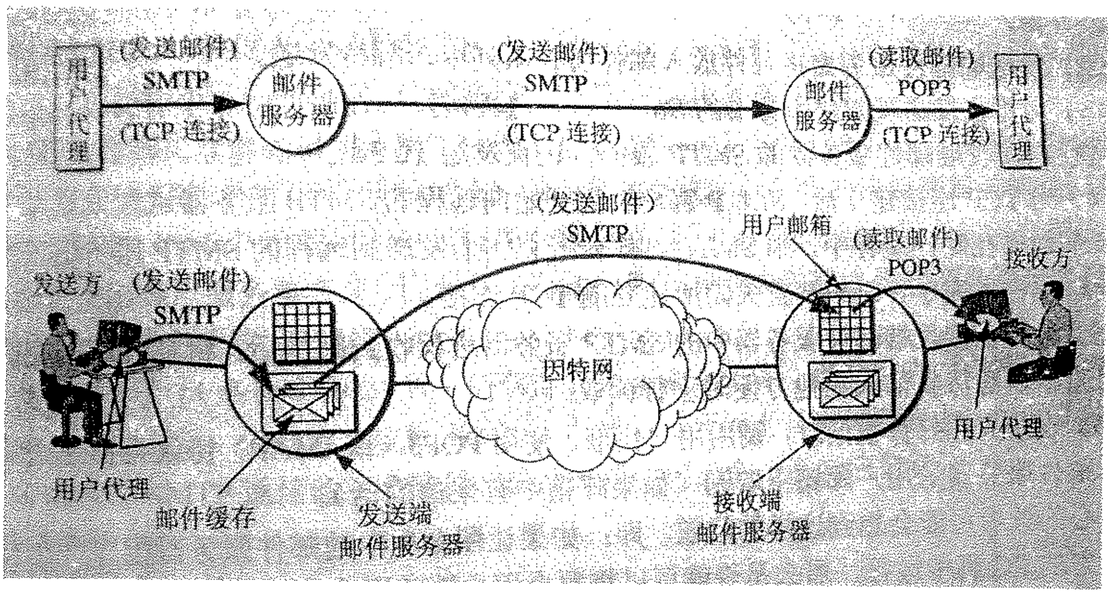
7.2 E－mail

username@mailserver

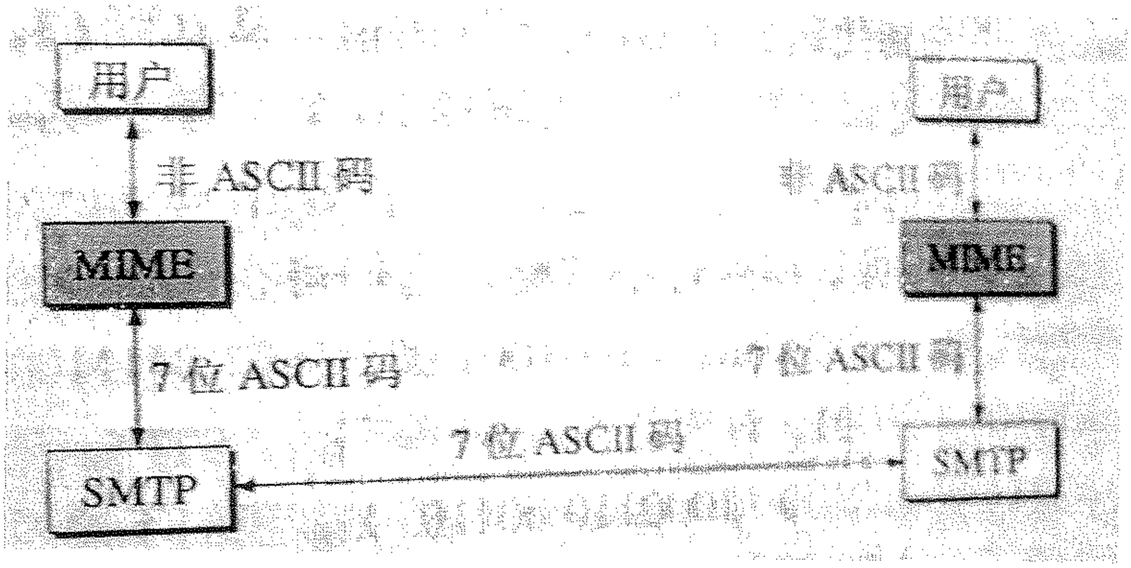
Header	Meaning
To:	E-mail address(es) of primary recipient(s)
Cc:	E-mail address(es) of secondary recipient(s)
Bcc:	E-mail address(es) for blind carbon copies
From:	Person or people who created the message
Sender:	E-mail address of the actual sender
Received:	Line added by each transfer agent along the route
Return-Path:	Can be used to identify a path back to the sender

SMTP: 用于发送，简单邮件传输协议

POP3/IMAP: 用于接收，邮局协议 3/ 因特网报文存取协议



MIME：通用因特网邮件扩充



7.3 The World Wide Web(WWW)

1. hypertext：超文本
2. hypermedia：超媒体＝超文本＋多媒体
3. URL: Uniform Resource Locators,统一资源定位器
   http://www.gliet.edu.cn/index.htm
   <协议类型>://<主机名>:<端口>/<路径及文件名>
4. HTML: the HyperText Markup Language，超文本标记语言
5. Browser：浏览器
6. HTTP：超文本传输协议

7.4 Multimedia：多媒体，非网络主要内容

7.5 FTP/TFTP/NFS

7.6 TELNET

ADSL Asymmetric Digital Subscriber Line
ARP Address Resolution Protocol
ARQ Automatic repeat request
BGP Border Gateway Protocol
CATV Community Antenna Television
CDMA Code Division Multiple Access
CIDR Classless Inter-Domain Routing
CRC Cyclic Redundancy Check
CSMA/CD Carrier Sense Multiple Access with Collision Detect
DNS Domain Name System
FTP File Transfer Protocol
FTTH Fiber to the Home
HTTP Hypertext Transfer Protocol
ICMP Internet Control Message Protocol
IEEE Institute of Electrical and Electronic Engineers
IETF Internet Engineering Task Force
IPv6 Internet Protocol version 6
IRTF Internet Research Task Force
ISO International Organization for Standardization
ISP Internet service provider
ITU International Telecommunication Union
LAN Local Area Network
MAC Media Access Control
OSI/RM Open System Interconnection/Reference Model
OSPF Open Shortest Path First
RIP Routing Information Protocol
SMTP Simple Mail Transfer Protocol
SNMP Simple Network Management Protocol
TCP Transmission Control Protocol
UDP User Datagram Protocol
URL Uniform resource locator
WAN Wide area network
WWW World Wide Web

Page 81

1.1 Imagine that you have trained your St. Bernard, Bernie, to carry a box of three 8mm tapes instead of a

flask of brandy. (When your disk fills up, you consider that an emergency.) These tapes each contain 7

gigabytes. The dog can travel to your side, wherever you may be, at 18 km/hour. For what range of

distances does Bernie have a higher data rate than a transmission line whose data rate (excluding

overhead) is 150 Mbps?

The dog can carry 21 gigabytes, or 168 gigabits. A speed of 18 km/hour

equals 0.005 km/sec. The time to travel distance x km is x /0.005 = 200x sec,

yielding a data rate of 168/200x Gbps or 840/x Mbps. For x < 5.6 km, the

dog has a higher rate than the communication line.

1.18. Which of the OSI layers handles each of the following:

a. (a) Dividing the transmitted bit stream into frames.

b. (b) Determining which route through the subnet to use.

(a) Data link layer. (b) Network layer.

1.20 With n layers and h bytes added per layer, the total number of header bytes

per message is hn, so the space wasted on headers is hn. The total message

size is M + nh, so the fraction of bandwidth wasted on headers is
hn /(M + hn).

A system has an n-layer protocol hierarchy. Applications generate messages of length M bytes. At each

of the layers, an h-byte header is added. What fraction of the network bandwidth is filled with headers?

Page 177

2.4. If a binary signal is sent over a 3-kHz channel whose signal-to-noise ratio is 20 dB, what is the maximum achievable data rate?

A signal-to-noise ratio of 20 dB means S/N = 100. Since log2101 is about 6.658, the Shannon limit is about 19.975 kbps. The Nyquist limit is 6 kbps. The bottleneck is therefore the Nyquist limit, giving a maximum channel capacity of 6 kbps.

2.42 Compare the delay in sending an x-bit message over a k-hop path in a circuit-switched network and in a (lightly loaded) packet-switched network. The circuit setup time is s sec, the propagation delay is d sec per hop, the packet size is p bits, and the data rate is b bps. Under what conditions does the packet network have a lower delay?

With circuit switching, at t = s the circuit is set up; at t = s + x /b the last bit is sent; at t = s + x /b + kd the message arrives. With packet switching, the last bit is sent at t = x /b. To get to the final destination, the last packet must be retransmitted k - 1 times by intermediate routers, each retransmission taking p /b sec, so the total delay is x /b + (k - 1)p /b + kd. Packet switching is faster if s > (k - 1)p /b.

Page 243

3.2. The following character encoding is used in a data link protocol: A: 01000111; B: 11100011; FLAG:

01111110; ESC: 11100000 Show the bit sequence transmitted (in binary) for the four-character frame: A

B ESC FLAG when each of the following framing methods are used:

a. (a) Character count.

b. (b) Flag bytes with byte stuffing.

c. (c) Starting and ending flag bytes, with bit stuffing.

The solution is

(a) 00000100 01000111 11100011 11100000 01111110

(b) 01111110 01000111 11100011 11100000 11100000 11100000 01111110

01111110

(c) 01111110 01000111 110100011 111000000 011111010 01111110

3.3 The following data fragment occurs in the middle of a data stream for which the byte-stuffing algorithm

described in the text is used: A B ESC C ESC FLAG FLAG D. What is the output after stuffing?

After stuffing, we get A B ESC ESC C ESC ESC ESC FLAG ESC FLAG D.

3.17 A channel has a bit rate of 4 kbps and a propagation delay of 20 msec. For what range of frame sizes

does stop-and-wait give an efficiency of at least 50 percent?

Efficiency will be 50% when the time to transmit the frame equals the roundtrip

propagation delay. At a transmission rate of 4 bits/ms, 160 bits takes 40

ms. For frame sizes above 160 bits, stop-and-wait is reasonably efficient.

Page 339

4.4 Ten thousand airline reservation stations are competing for the use of a single slotted

ALOHA channel. The average station makes 18 requests/hour. A slot is 125 μsec. What
is the approximate total channel load?

Each terminal makes one request every 200 sec, for a total load of 50 requests/sec. Hence G = 50/8000 = 1/160.

Page 475

5.9 Consider the subnet of Fig. 5-13(a). Distance vector routing is used, and the following
vectors have just come in to router C: from B: (5, 0, 8, 12, 6, 2); from D: (16, 12, 6,
0, 9, 10); and from E: (7, 6, 3, 9, 0, 4). The measured delays to B, D, and E, are 6, 3,
and 5, respectively. What is C's new routing table? Give both the outgoing line to use
and the expected delay.

Going via B gives (11, 6, 14, 18, 12, 8).
Going via D gives (19, 15, 9, 3, 9, 10).
Going via E gives (12, 11, 8, 14, 5, 9).
Taking the minimum for each destination except C gives (11, 6, 0, 3, 5, 8).
The outgoing lines are (B, B, –, D, E, B).

14. Looking at the subnet of Fig. 5-6, how many packets are generated by a broadcast from
B, using
(b) the sink tree?

(b) The sink tree needs four rounds and 14 packets.

27 A computer on a 6-Mbps network is regulated by a token bucket. The token bucket is
filled at a rate of 1 Mbps. It is initially filled to capacity with 8 megabits. How long can
the computer transmit at the full 6 Mbps?

The correct answer can be obtained by using the formula $S = C/(M - \rho)$. Substituting, we get $S = 8/(6 - 1)$ or 1.6 sec.