

网络接口层与网络层

第二章

物理层

- 主机依靠物理连接（传输介质）组成网络
 - 有线连接（有线传输介质）
 - 无线连接（无线传输介质）

- 有线连接
 - 同轴电缆 - 技术成熟线的结构复杂
 - 双绞线 - 无屏蔽/屏蔽，使用RJ45接头
 - 光缆
- 无线传输
 - IEEE 802.11系列标准

物理层传输方式

- 串行通信与并行通信
 - 与面向连接/面向消息传输类比
 - 举例：打印机

- 同步传输与异步传输

- 在通信过程中，发送方和接收方必须在时间上保持步调一致，亦即同步，才能准确的传送信息。解决的方法时，要求接收端根据发送数据的起止时间和时钟频率，来校正自己的时间基准与时钟频率。这个过程叫位同步或码元同步。在传送由多个码元组成的字符以及由许多字符组成的数据块时，通信双方也要就信息的起止时间取得一致，这种同步作用有两种不同的方式，因而也就对应了两种不同的传输方式。

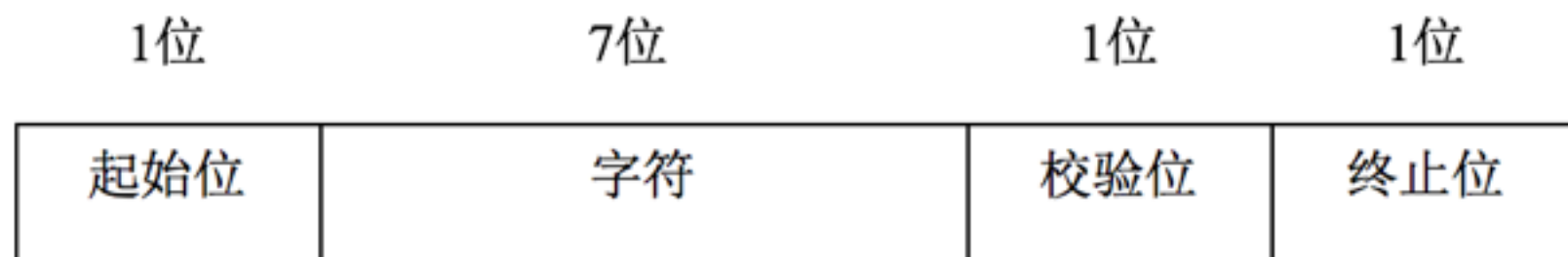
- 同步传输

- 同步传输在传送连续的数据块、大数据块时比异步传输更有效。按这种方式，发送方在发送数据之前先发送一串同步字符SYN（编码为0010110），接收方只要检测到两个以上的SYN字符，就可以确认已进入同步状态，准备接收数据，随后双方以同一频率工作，直到传送完指示数据结束的控制字符。这种方式仅在数据块前加入控制字符SYN，所以效率更高，但实现起来较复杂。在短距离高速数据传输中，多采用同步传输方式。

- 异步传输

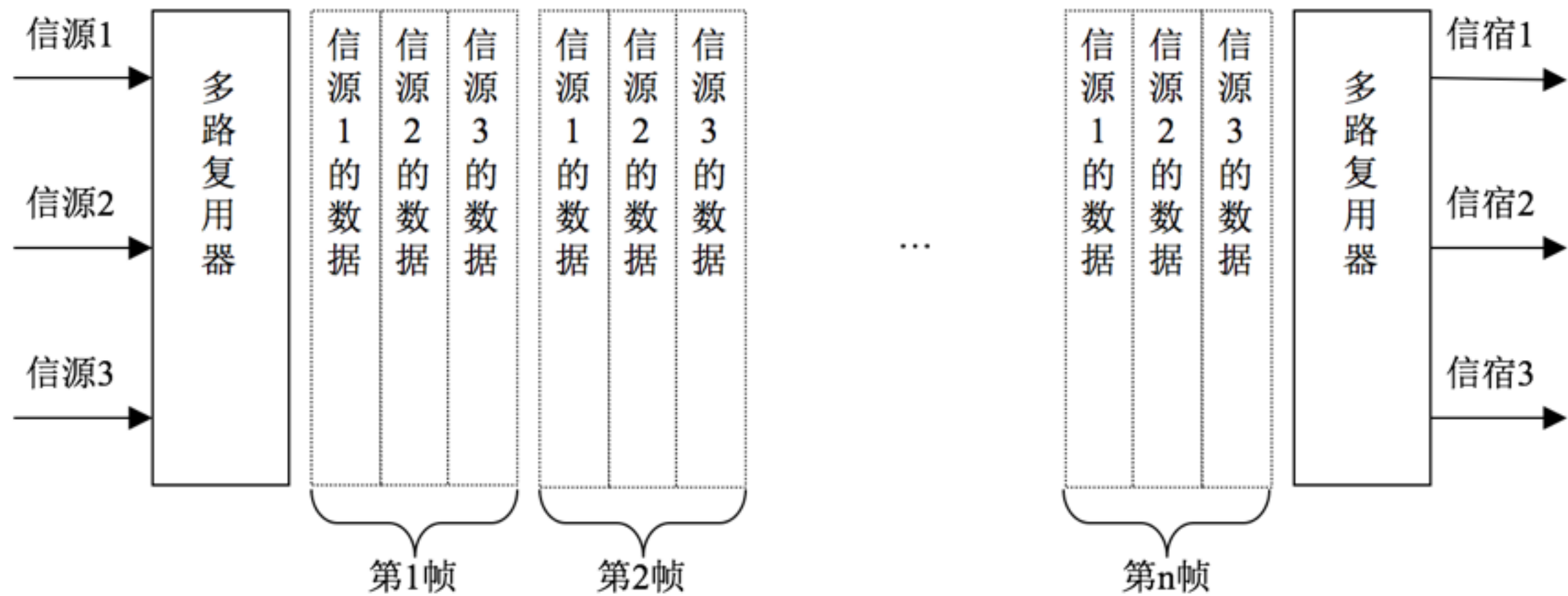
- 异步传输即把各个字符分开传输，字符与字符之间插入同步信息。这种方式也叫起止式，即在组成一个字符的所有位前后分别插入起始位和终止位。

- (类比面向连接的传输, EOF)



信道复用

- 线路容量或者介质带宽是非常宝贵的资源，必须尽量地提高介质利用率
- 为了有效地利用传输系统，可以采用多路复用(Multiplexing) 技术从空间和时间两个方面来解决多路信号同时复用单一介质的问题。多路复用技术包括：
 - 频分多路复用FDM(Frequency Division Multiplexing)、
 - 时分多路复用TDM(Time Division Multiplexing)
 - 码分多路复用CDM (Code Division Multiplexing)
 - 等.....



类比：并行处理

数据链路层

- 基本概念与功能：物理层上的第一层。数据链路层将不可靠的物理链路变成可靠的数据链路，并提供基本的流量控制和差错校验功能
- 数据链路层从功能上来说可以将不可靠的物理链路变成可靠的数据链路，并提供流量控制和差错校验功能。
- 帧传输和帧同步：帧是具有一定长度和格式的信息块，是数据链路层的传输单位，一般由一些字段和标志组成。

- 差错与流量控制：为保证发方发出的所有帧都能够正确有序地交付给收方的网络层，需要对帧的传输进行差错控制和流量控制
- 数据链路管理：数据链路的管理主要是提供各种服务质量参数，包括检测到不可纠正错误的平均时间、漏检差错率、传输延迟和吞吐量等，以及对异常情况的处理

局域网 (LAN)

- 定义 (IEEE)
 - 局域网网络中的通信被限制在中等规模的地理范围内，例如一幢办公楼、一座工厂或一所学校
 - 能够使用具有中等或较高数据速率的物理信道，且具有较低的误码率；局域网是专用的，由单一组织机构所使用。
 - 一种地理范围有限、互连设备有限的计算机网络

- 基本组成
 - 网络服务器
 - 网络工作站
 - 网络适配器
 - 传输介质
- 这些设备在特定网络软件支持下完成特定的网络功能
- 局域网->互联网

- 服务器

- 局域网的一个重要目的是实现资源共享。局域网的许多功能是通过网络服务器来实现的，同时网络服务器也是网络共享资源的管理中心，网络操作系统等软件也主要驻留在网络服务器上。网络服务器管理的常用共享设备有高速打印机、文件系统、数据库、电子邮件、贵重的外围设备等。网络服务器通常放在一个专门的网络节点上，该节点有固定地址，并为网络上的用户所共知，为他们提供服务。

- 提供服务->服务器

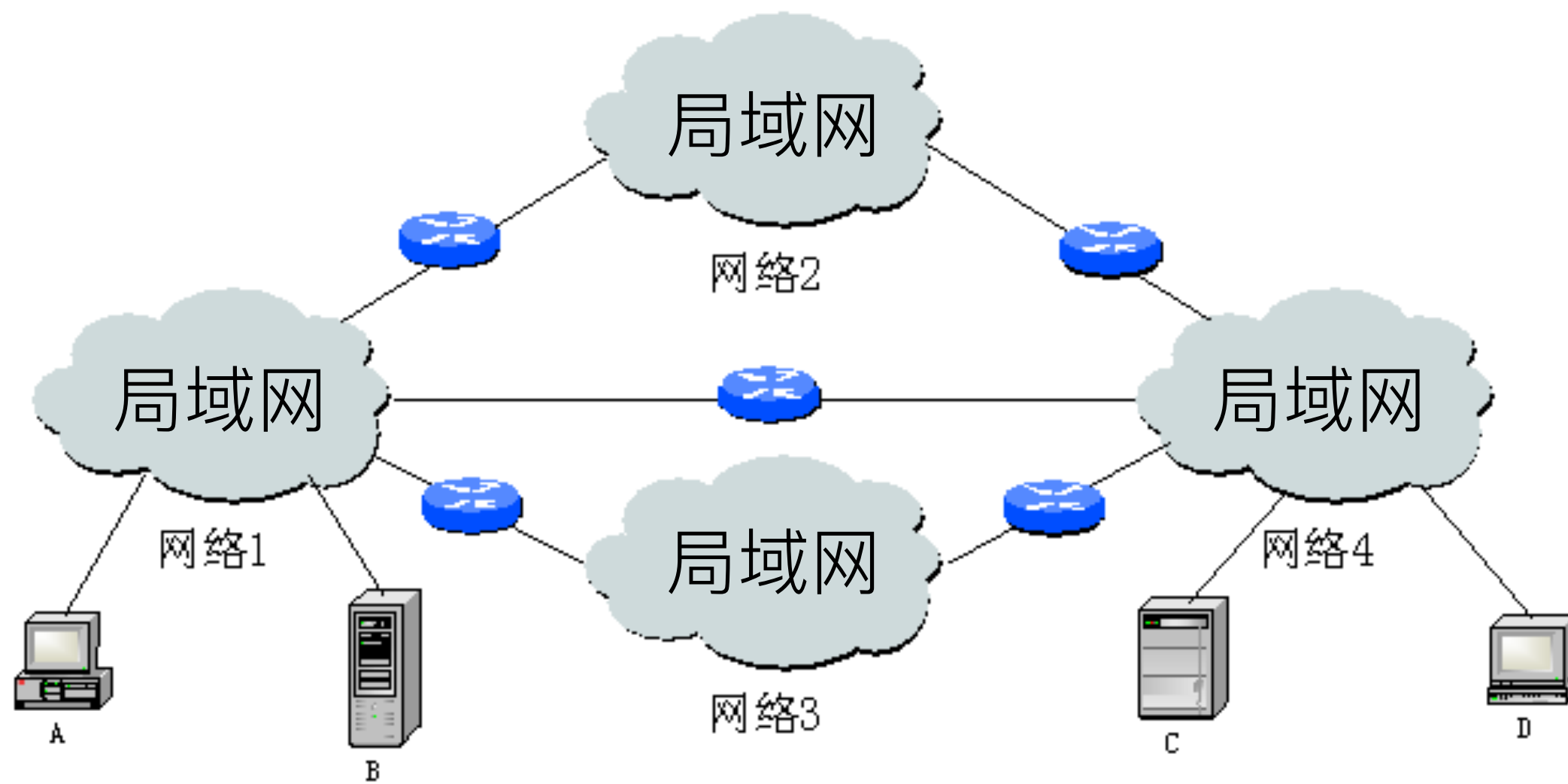
- 工作站
- 网络工作站是指连接到计算机网络上供用户使用的前端窗口，用户通过网络工作站来访问网络的共享资源。在局域网中，用户工作站一般采用微型计算机，如各种PC机

- 网络适配器、NIC (Network Interface Controller)
 - 接口，实现物理连接
 - 发送与接收数据
 - 同一服务器可有多个适配器/NIC (Socket与不同地址绑定?)

- (数据链路层的) 网络软件
 - 工作在数据链路层
 - 协议、驱动程序等
 - 与应用层软件的区别

网络层

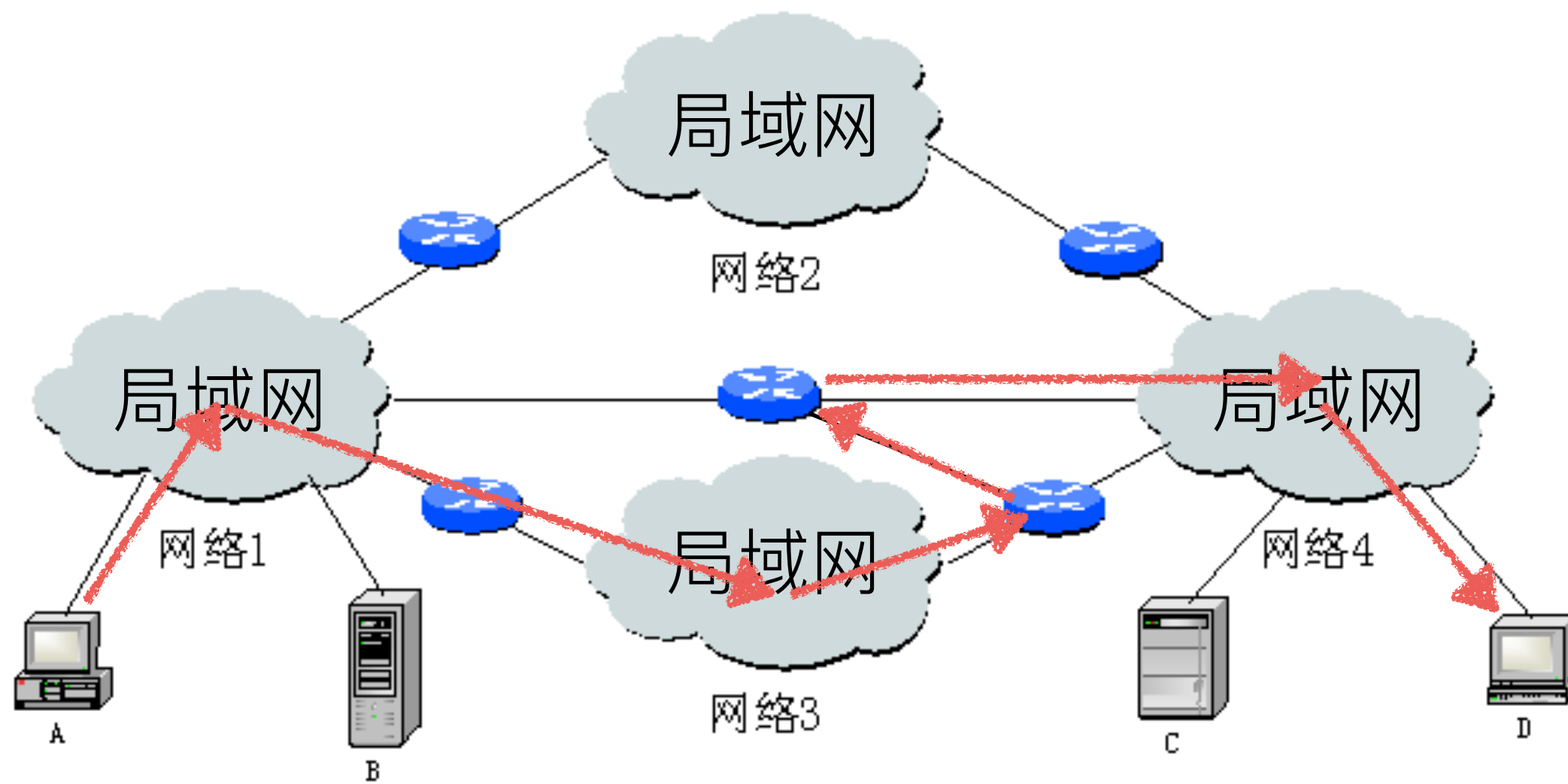
- internetwork and Internet
 - 小写的i - 互联的网络, interconnected networks: 局域网、城域网、片上网络 (Network-on-Chip, NoC)
 - 大写的I - (全球范围内的) 互联网
 - (可以) 通过IP标示每一台 (逻辑上的) 主机
- 网络=IP?



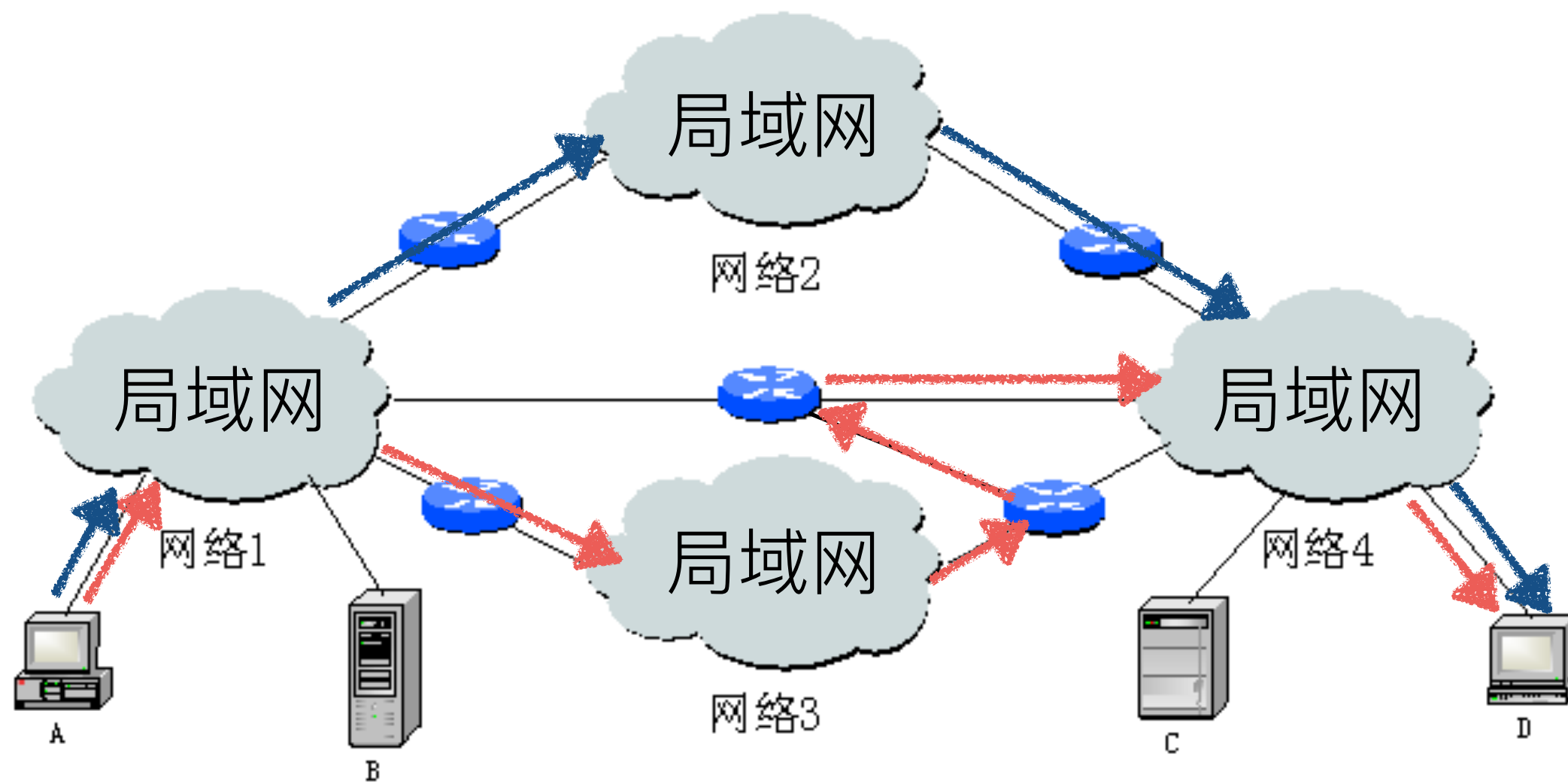
面向连接与无连接的互联

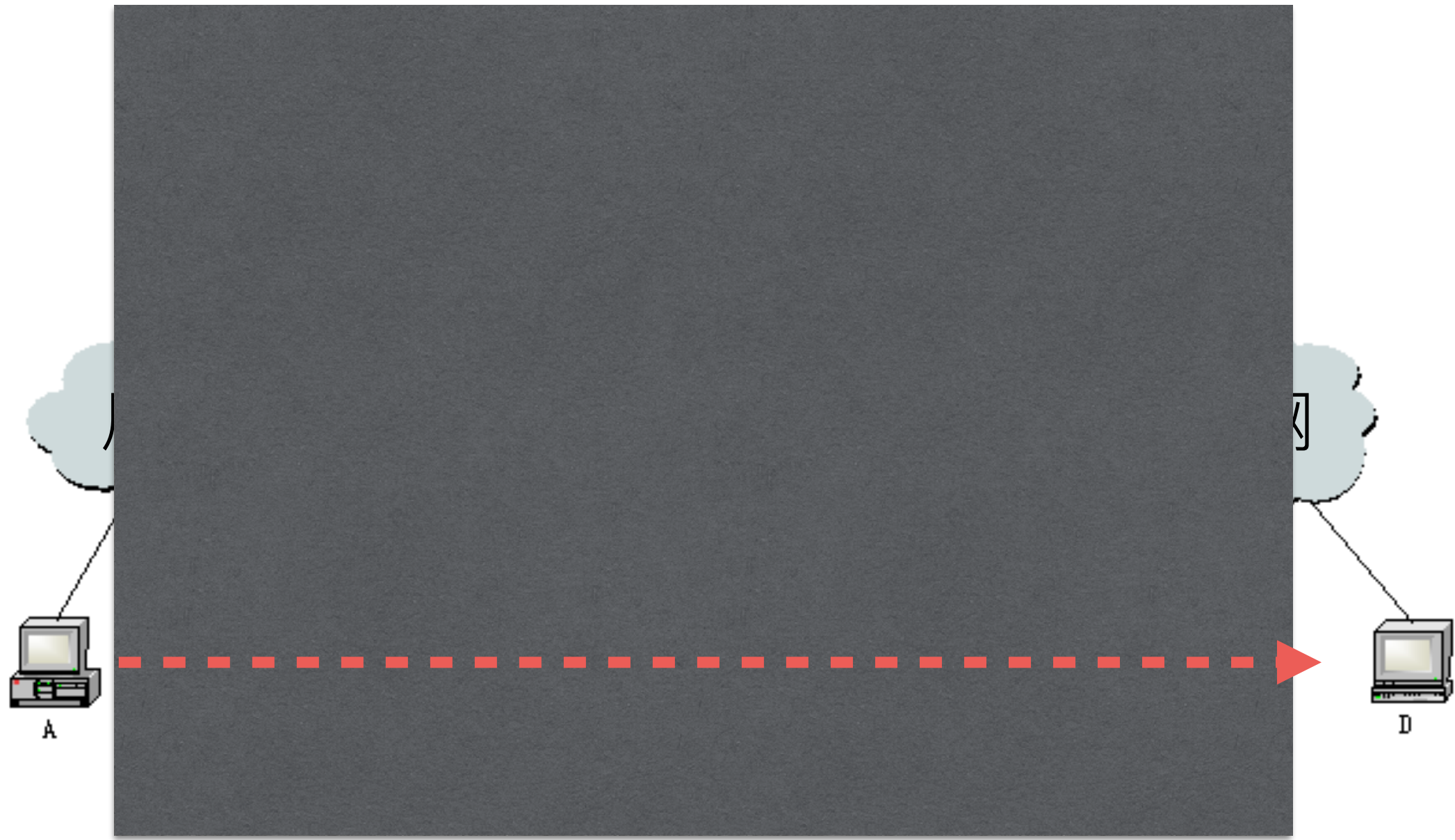
- 网络层可以采用的两种互联方式
 - 面向连接 (connection oriented)
 - 无连接 (connectionless)

- 面向连接
 - 在路由器上建立固定的线路，好像在通信的源和目标之间建立了一条逻辑通路，称为虚电路 (virtual circuit)
 - 接收端收到顺序严格一致的数据分组
 - 缺点：可靠性
 - 例如：X.25 (ITU的标准，比较“电信”，按照OSI模型建立)



- 无连接 (Connectionless)
 - 源端和目标端之间不需要建立一条逻辑通路，而是送到网络中的每个分组带有完整的目标主机地址
 - 收到分组的节点（主要是路由器）根据目标地址和当前网络状况（如通信量等），选择一条合适的线路把分组发送到接近目标端的下一个节点
 - 通过多个节点的转发，最终把分组送达目标节点。





- 换一种说法
 - 面向连接与面向消息
- 类比
 - 传送带与快递

- IP协议实现无连接的通信
- TCP实现了面向连接的通信
- 为什么有如此设计?

网络层互联技术

- 路由器
 - 工作在网络层（对比：局域网）
 - 连接逻辑上分开的异构（Heterogeneous）网络
 - 拓扑结构灵活
 - 不关心子网的硬件（与数据链路层），运行一致的网络层软件（e.g., IP协议）

- 路由器工作原理

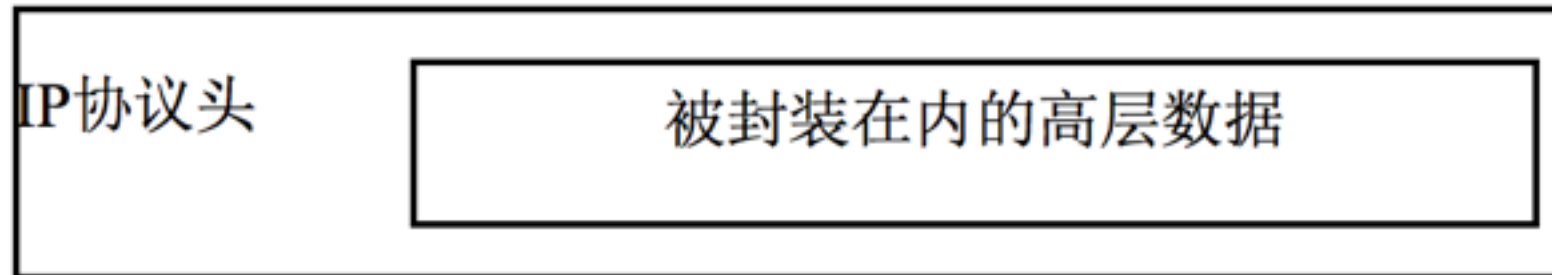
- 当IP子网中的一台主机发送IP报文给同一子网的另一台主机时，它将直接把IP报文送到网络上，对方就能收到
- 而要送给不同IP子网上的主机时，它要选择一个能到达目的子网的路由器，把IP报文送给该路由器，由路由器负责把IP报文送到目的地
- 如果没有找到这样的路由器，主机就把IP报文送给一个称为“默认网关”（default gateway）的路由器上。“默认网关”是每台主机上的一个配置参数，它是接在同一个网络上的某个路由器端口的IP地址。

- 路由器如何判断转发方向？
 - 路由器转发IP报文时，只根据IP报文中目的IP地址的网络号部分，选择合适的端口，把IP报文送出去
 - 每个路由器在收到IP报文后也要判定端口所连接的是否是目的子网
 - 如果是，就直接把报文通过端口送到网络上
 - 否则，也要选择下一个路由器来转发IP报文。
 - 路由器也有它的“默认网关”，用来传送不知道往哪儿送的IP报文。这样，通过路由器把知道如何传送的IP正确转发出去，不知道如何传送的IP报文送给“默认网关”，这样一级级地传送，IP报文最终将送到目的地，送不到目的地的IP报文则被网络丢弃了

- 路由器在网络层实现的功能
 - 为网际间通信选择最合理的路由。
 - 根据需要可对数据包进行拆分或组装。
 - 对于支持多通信协议的路由器，可以将两个使用不同协议的网络连接起来。
 - 网络安全功能，目前大部分路由器都具备一定的防火墙功能，能够屏蔽内部网络的IP地址，自由设定IP地址、通信端口过滤等，使网络更加安全。

IP协议互联的原理

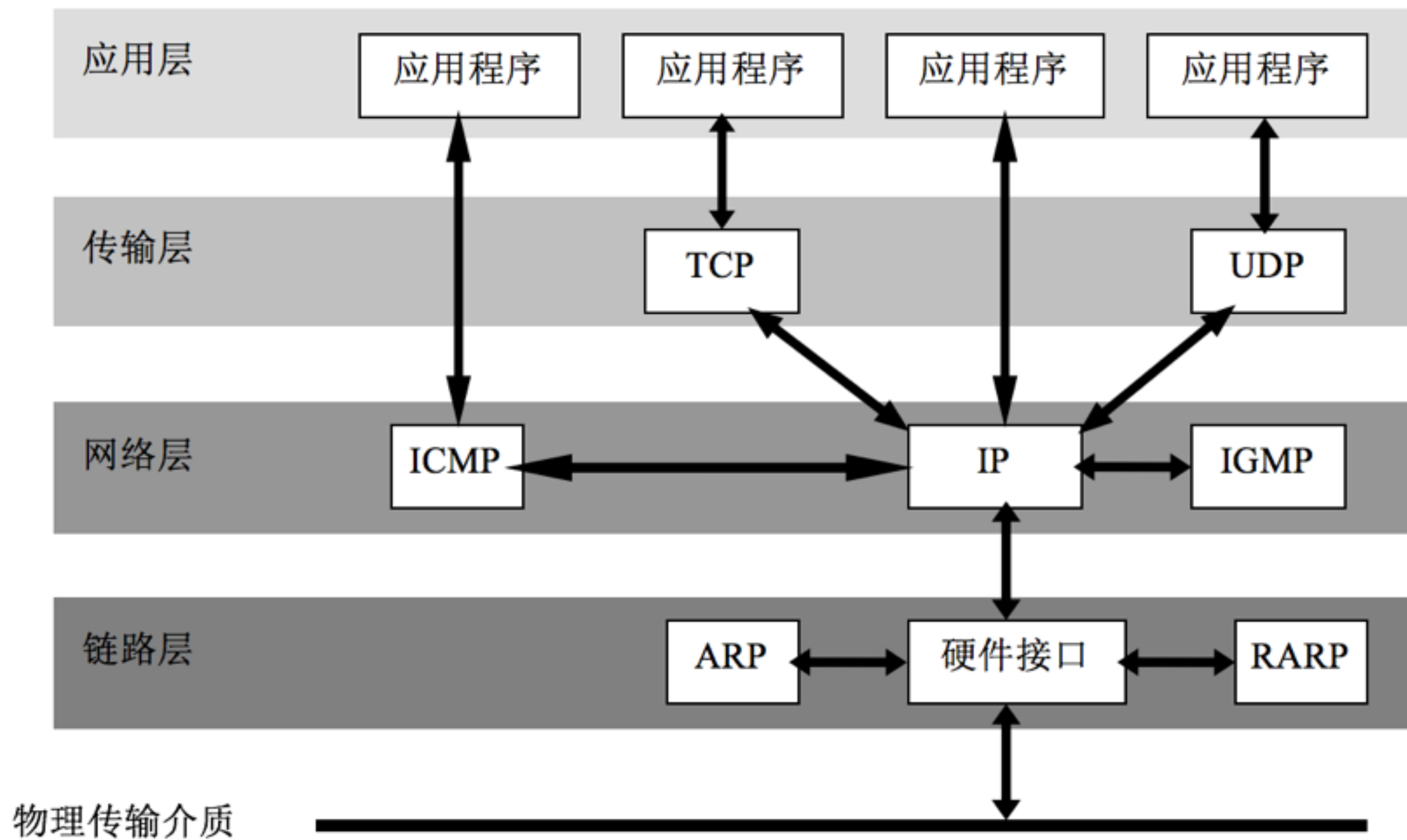
- 数据链路层是对物理层的“透明”化，使得节点组成各种局域网网络
- 网络层如何使结构各异的局域网网络“透明”化？
 - 使数据帧（Frame）携带IP地址等信息，形成IP数据报（Datagram），使整个IP网络中数据通信有统一模式
 - 加入地址信息：IP地址
 - 路由器识别数据报中的信息，如目标地址



- IP协议是TCP/IP协议簇中最重要的协议之一
- 从该层（网络层）开始提供了一个整体统一的互联网
- 思考：
 - 分层协议中从上层到底层，数据是被包装了还是被解包了

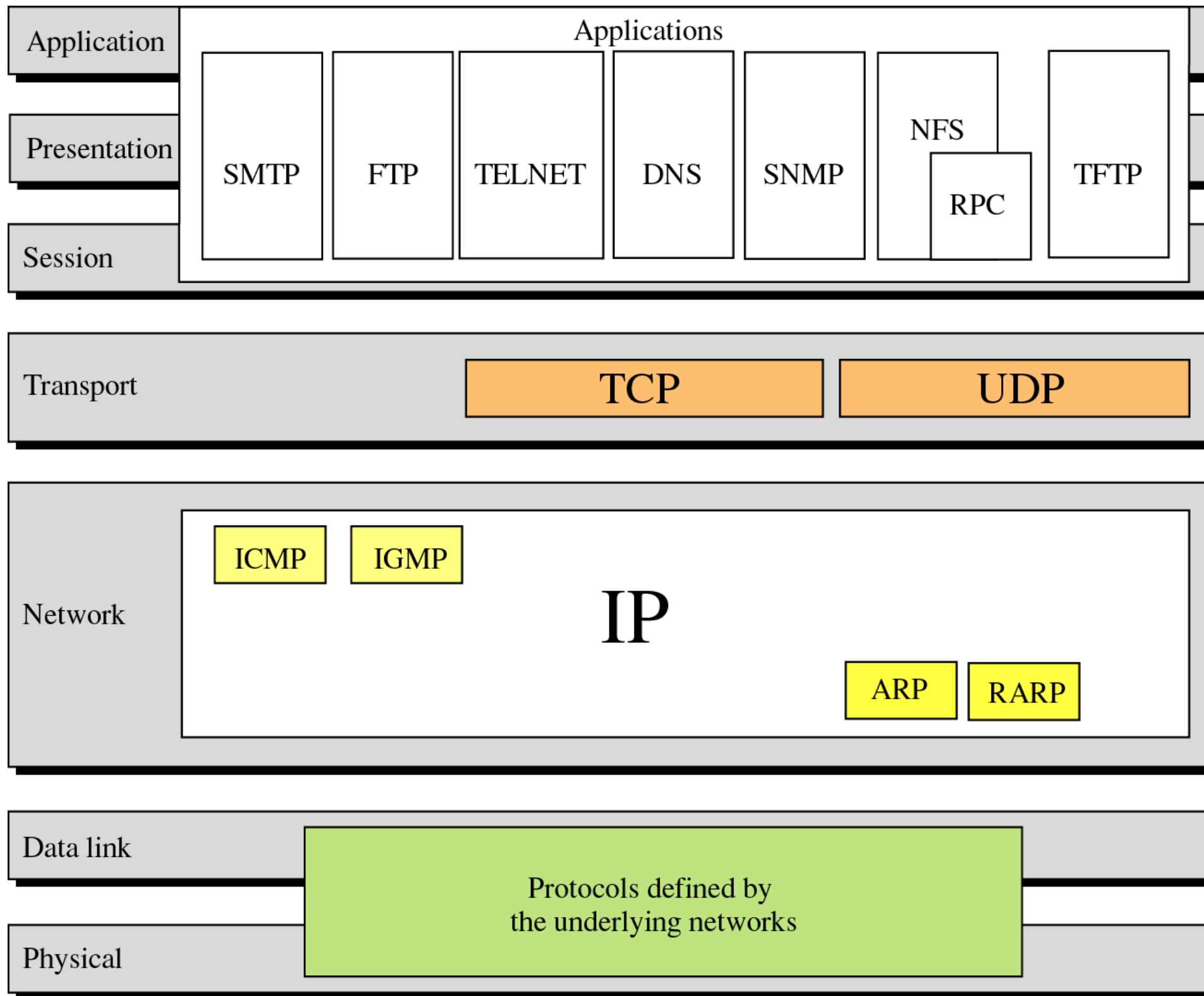
IP协议在TCP/IP协议簇中的地位

- “瓶颈”与“分界”
 - 互联网中所有的数据报都要经过IP协议来传输
 - 通信网络与互联网高层协议的分界？
 - 题外话：所谓计算机类专业与通信类专业的区别



IP协议在TCP/IP协议簇中的特点

- IP协议是一种无连接(connectionless)不可靠(unreliable)的数据报传输协议
- 不可靠：IP协议不能保证数据报能正确的传输到目标主机。它负责数据报在网络中的传输，而不管传输的正确与否，不进行数据报的确认，也不能保证数据报按正确的顺序到达（即先发的不一定先到达）
- 同时它也是“尽最大努力”（Best effort）的传输数据，绝不随便丢弃传输中的数据报



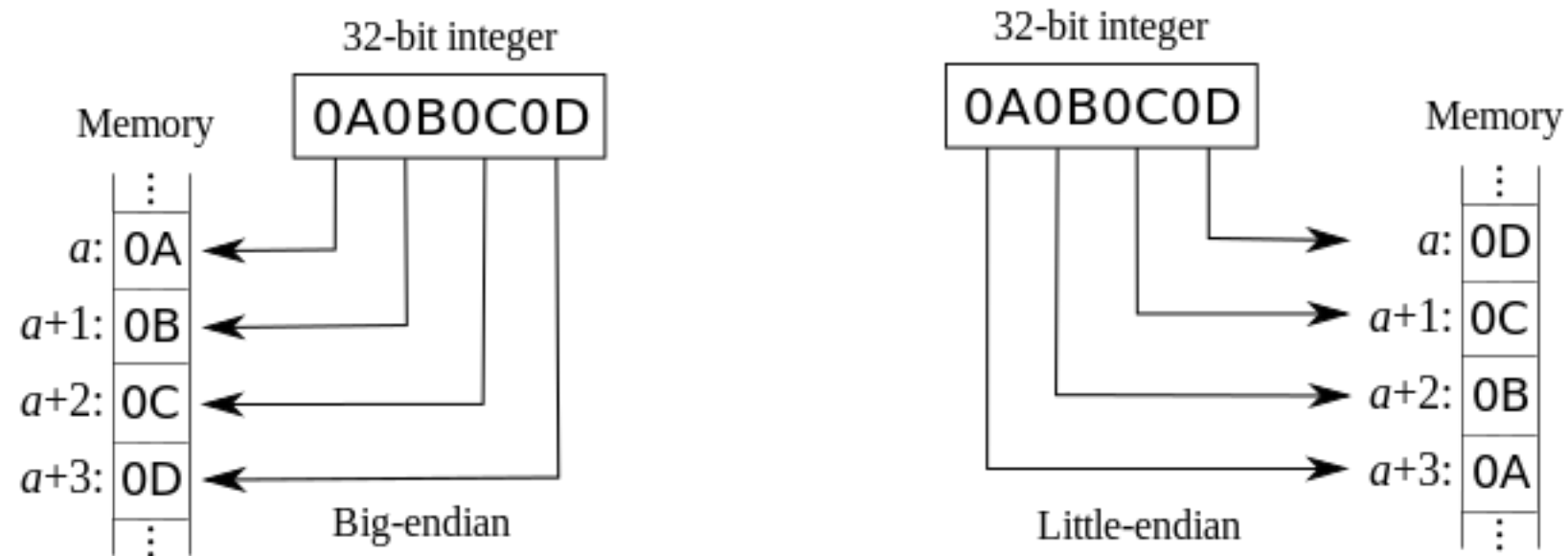
- IP网络的特点
 - IP互联网中的计算机没有主次之分，所有主机地位平等（因为惟一标识它们的是IP地址），从逻辑上来说，IP互联网络所有网络不管规模大小也没有主次之分
 - IP互联网没有确定的拓扑结构
 - 在IP互联网中的任何一台主机，都至少有一个独一无二的IP地址，有多个网络接口卡（NIC）的计算机每个接口可以有一个IP地址，这样一台主机可能就有多个IP地址，有多个IP地址的主机叫多宿主主机（multi-home host）
 - 在互联网中有IP地址的设备，不一定就是一台计算机，如IP路由器、网关等，因为与互联网有独立连接的设备都要有IP地址

IP数据报格式

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

字节序 (endianness) 问题

- 字节序问题：大端 (Big endian) 与小端 (Little endian)
- 数据在内存中的存储方向问题



- 大端和小端，哪个好？

字节序-以IP数据报格式为例

0

31

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 最高位在左记为第0位；最低位在右记为第31位
- 网络传输时，先传输0-7位.....最后传输24~31位

- TCP/IP协议首部中所有的二进制数在网络中传输时都要求以此次序进行，因此把它称为网络字节序
- 在进行程序设计时以其他形式存储的二进制数，必须在传输数据之前，把首部转换成网络字节顺序，否则可能出现传输错误

- 举例处理字节序转换问题
 - X86 architecture采用Little endian格式
 - 主机 (host) 与网络 (network) 传输需进行转换
 - htonl(), htons(), ntohs(), ntohl()
 - 很多通信为主的函数均自动实现了转换

- 封装与拆封来实现的。IP数据报的格式分为报头区和数据区两大部分
- 数据区包括高层协议需要传输的数据，报头区是为了正确传输高层数据而加的各种控制信息。

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- Socket软件开发程序里实现一个类似报文结构?

IP数据报格式-版本号

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)		标志(3位)	片偏移 (13位)	
生存时间TTL (8位)	上层协议标识 (8位)	头部校验和 (16位)		
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 占用4位二进制数
- 表示该IP数据报使用的是哪个版本的IP协议

IP数据报的格式-头长度

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)		标志(3位)	片偏移 (13位)	
生存时间TTL (8位)	上层协议标识 (8位)	头部校验和 (16位)		
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 4位。描述IP数据报头的长度，因IP包头中有变长的可选部分
- 该部分单位为32bit (4个字节)。IPv4包头最小长度为20字节。则对于标准IPv4报头，这个字段的值是_____?

IP数据报的格式-总长度

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 长度16bit
- 以字节为单位
- 因此，一个IP数据包的最大长度为_____?

IP数据报的格式-服务类型

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)		标志(3位)	片偏移 (13位)	
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 规定了对本数据报的处理方式，如服务质量、优先权等等。
- 此字段的意义现今有一些变化

IP数据报的格式-服务类型

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 传统：前3位表示优先权（Precedence），用于拥塞控制等
- 后5为为服务类型ToS，规定了延迟、吞吐量、可靠性、路由开销等等

IP数据报的格式-服务类型

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 演进：前6位Differentiated Services Code Point (DSCP),
- 后2位Explicit Congestion Notification (ECN)

IP数据报的格式-分片相关字段

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 由于IP包会分片传输，由无连接传输特性，导致目标地址不一定按顺序接收，有的数据包会丢失
- Identification, 标识。唯一标识一个发送的数据包（未分片的）

IP数据报的格式-分片相关字段

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- Fragment offset
- 由于IP包会分片传输，由无连接传输特性，导致目标地址不一定按顺序接收，有的数据包会丢失

IP数据报的格式-分片相关字段

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)		标志(3位)	片偏移 (13位)	
生存时间TTL (8位)	上层协议标识 (8位)		头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- Flag
- 0XY格式，X位表示不分片，Y标记MF (More Fragments)

IP数据报的格式-片偏移

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)			标志(3位)	片偏移 (13位)
生存时间TTL (8位)		上层协议标识 (8位)	头部校验和 (16位)	
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- TTL=Time to Live
- 每个沿途的路由器会将IP包的TTL值减少1。如果TTL减少为0，则该IP包会被丢弃

IP数据报格式-上层协议标识

十进制编号	协 议	说 明
0	无	保留
1	ICMP	网际控制报文协议
2	IGMP	网际组管理协议
3	GGP	网关-网关协议
4	无	未分配
5	ST	流
6	TCP	传输控制协议
8	EGP	外部网关协议
9	IGP	内部网关协议
11	NVP	网络声音协议
17	UDP	用户数据报协议

IP数据报分片与重组

- 最大传输单元-MTU (Maximum Transmission Unit)
 - IP数据报在互联网上传输，可能经过多个物理网络数据
 - 不同的网络由于链路层和介质的物理特性不同，因此在进行数据传输时，对数据帧的最大长度都有一个限制，这个限制值叫最大传输单元MTU
 - 分片问题一般只存在于具有不同MTU值的互联网中
 - 所以分片工作通常由路由器进行

- 最大传输单元-MTU
 - 如果两台主机通信要通过多个具有不同MTU值的网络时，重要的是两台通信主机路径上的最小MTU值，它被称作路径MTU
 - 两台主机之间的路径MTU不一定是个常数，它取决于数据报所经过的路径
 - 由于路由不一定是对称的，路径MTU在两个方向上不一定是一致的

- IP数据报的分片 (fragmentation)
 - 当一个IP数据报要通过链路层进行传输时，如果IP数据报的长度比链路层MTU的值大，那么IP层就需要对将要发送的IP数据报进行分片
 - 被分片后的各个IP数据报，可能经过不同的路径到达目标主机。
 - 分片后的IP数据报和原来没有分片的IP数据报结构是相同的，即也是由IP头部和IP数据区两部分组成

- IP数据报的重组
 - 当分了片的IP数据报在目标主机要需要重新进行组装，此过程叫IP数据报的重组
 - 在IP数据报头部唯一地标识主机发送的每一份数据报
 - 通常在源端保持着一个计数器，每发送一个数据报，头部的标识的值就会加1
 - 而在一个数据报被分成片（fragments）时，复制相同的标识。
 - 目标端主机根据标志字段的片未完MF（More fragments）子字段可以确定分片是不是最后一个分片；根据偏移量就可以确定分片在原数据报中的位置。

网络名称	MTU (单位: 字节)
以太网	1500
IEEE802.3/802.2	1492
FDDI	4352
ATM(信元)	48
X.25	576
点到点(低延时)	296
令牌环网 (IBM 16MB/s)	17914
令牌环网 (IEEE802.5 IBM 16MB/s)	4464

ICMP

- Internet Control Message Protocol
- 所处的位置：介于网络层和传输层之间，可认为作为IP协议之补充而存在
- 主要有两类消息：
 - Error, 错误信息
 - Informational, 消息/咨询性质

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)		标志(3位)	片偏移 (13位)	
生存时间TTL (8位)	上层协议标识 (8位)	头部校验和 (16位)		
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- ICMP报文由ICMP报文头区和数据区两部分组成。
- 封装在IP数据报中通过链路层在网络进行传输的
- 这与高层其它协议（如TCP、UDP等）相似，它在IP数据报头中的协议标识是1

- 但一般我们不认为ICMP是“高层协议”，原因是：
 - 仅传送差错与控制
 - 处理的均为IP层的信息
 - 差错处理与控制都需要IP层来执行

- ICMP报文的特点
 - ICMP差错报文都是由路由器最终发送到源主机的
 - ICMP差错报文不规定对各类差错应采取什么样的处理措施
 - ICMP差错报文不享受任何优先权
 - ICMP差错报文是伴随着抛弃出错的IP数据报而产生的

- ICMP报文的特点 (续)

版本号 (4位)	头长度 (4位)	服务类型TOS (8位)	总长度 (16位)	
标识 (16位)		标志(3位)	片偏移 (13位)	
生存时间TTL (8位)	上层协议标识 (8位)	头部校验和 (16位)		
源IP地址 (32位)				
目标IP地址 (32位)				
选项				
数据				

- 当路由器发送一份ICMP差错报文时，ICMP报文数据区始终包含产生ICMP差错报文的IP数据报的首部和其数据区的前8个字节 (64位)
- 为防止广播风暴，发生差错不一定会产生ICMP差错报文

- ICMP消息并非不常见，在网络层很重要
 - Echo
 - Dest. unreachable
 - Time exceeded
 - Source quench
 - Redirect

IP地址

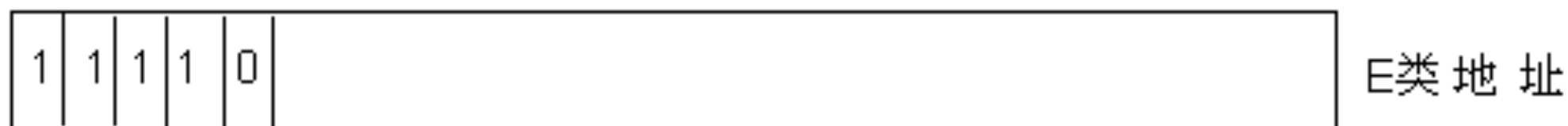
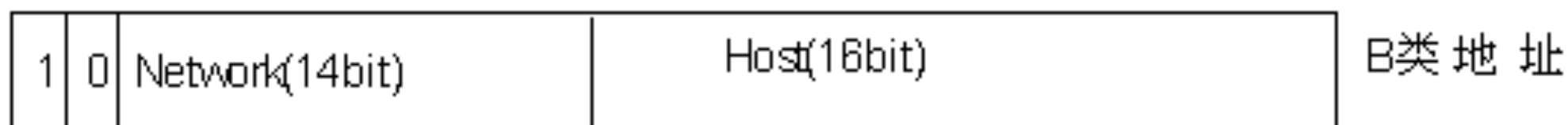
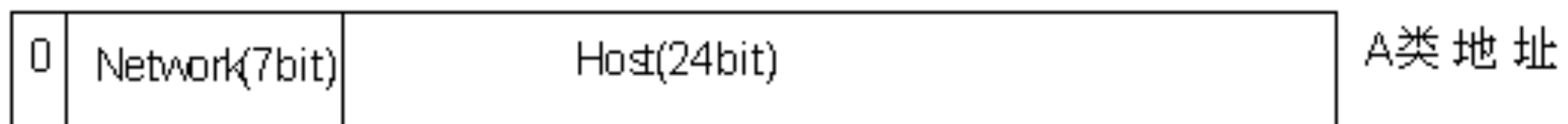
- 互联网是由很多网络连接而成的，互联网中的数据报有些是在本网内主机之间传输的，有些是要送到互联网中其它网络中主机的数据报，因此，IP地址不但要标识在本网内的主机号，还要标识在互联网中的网络号。
- 一个IP地址由网络号和主机号两部分组成，网络号标识互联网中的一个特定网络，主机号标识该网络中的一台特定主机。这样给定一个IP地址，就可以很方便的知道它是哪个网络上的哪一台主机。

表示方法	举 例	说 明
二进制	1000011000011000000010000 1000010	计算机内部使用
十进制	2249721922	很少使用
十六进制	0X86180842	较少使用
点分十进制	134.24.8.66	最常用

- IP地址的分类

- 在Internet发展的初期，人们用IP地址的前8位用来定义所在的网络，后24位用来定义该主机在当地网络中的地址。这样互联网中最多只能有255（应该有256个，但全1的IP地址用于广播）个网络。后来由于这种方案可以表示的网络数太少，而每个网络中可以连入的主机又非常多，这不符合互联网的实际情况。
- 于是，人们设计了一种新的编码方案，该方案中用IP地址高位字节的若干位来表示不同类型的网络，以适应大型、中型、小型网络对IP地址的需求。这种IP地址分类法把IP地址共分为A、B、C、D和E共五类，用IP地址的高位来区分。

类别	类标识	第一字节	网络地址长度	主机地址长度	最大网络数	最大主机数	选用范围
A类	0	1~126	1字节	3字节	126	16777214	大型网络
B类	10	128~191	2字节	2字节	16382	65534	中型网络
C类	110	192~223	3字节	1字节	2097150	254	小型网络
D类	1110	224~239	-	-	-	-	多点播送
E类	11110	240~247	-	-	-	-	保留地址



子网

- “两层IP地址结构”(网络+主机地址)产生网络容量的不公平
- 随着计算机网络技术发展，大量的个人用户和小型局域网接入互联网，即使分配一个C类网络地址仍然会造成IP地址的很大浪费
- 进一步把每个网络划分成若干个子网 (subnet)

- 子网掩码举例
 - 子网掩码用32位二进制数表示，常用点分十进制数格式来书写
 - 掩码中用于标识网络号和子网号的位置为1，主机位为0
 - 以一个C类地址为例
 - 取主机号的两位为子网号，则掩码为11111111.11111111.11111111.11000000(255.255.255.192)，子网可以产生64个可能的主机地址，实际上只有62个地址是可用的，另外两个地址，一个用于识别子网自身，另一个用于子网的广播，因此得到子网内最大可用的主机数时总要减去2。如2位的子网号数学上的组合为00、01、10和11共四种，第一种和最后一种组合有特殊用处，只剩下01和10可用于识别子网，得到2个可用的子网地址。

- 特殊的IP地址
 - 网络号固定+主机号全0：网络地址
 - 网络号固定+主机号全1：直接广播地址
- 例如，当互联网中的一台主机如果使用157.34.255.255为目标地址发送数据报时，则网络号为157.34.0.0的网络中的所有主机都能收到该数据报

网络部分	主机部分	地址类型	用途
Any	全“0”	网络地址	代表一个网段
Any	全“1”	广播地址	特定网段的所有节点
127	any	环回地址	环回测试
	全“0”	所有网络	通常 用于指定默认路由
	全“1”	广播地址	本网段所有节点

路由表

- 网络层进行路由选择的依据是保存在路由器中的路由表。
- 路由表保存了通过该路由器可能到达的目标网络以及如何到达该目标网络的信息。路由表中的基本表项包括目标网络地址、子网掩码以及到达目标网络的下一跳路由器地址。目标地址一般为目标网络地址，因为同一个网络中的主机网络号相同，这样做的优点是减少路由表的规模，提高对路由表的搜索速度。子网掩码可以使路由器以网络中的子网为目标网络（该例中没有划分子网）。下一站路由器是一个与该路由器连接在同一个网络上的路由器，为了把路由器收到的数据报送到目标网络，路由器必须把数据报传输到该地址指定的接口，下一站路由器并不一定与目标网络直接相连，但是通过它可以把数据报转发到最终目的主机。对与该路由器直接相连的网络则可直接投递。
- 路由表的目标地址主要是目标网络地址，但也可以包含两种特殊的路由表项。一种是默认路由，它可以进一步简化路由表，当路由表中没有表项能与目标网络相匹配时，就把数据报发送到默认路由。

```
[Yangs-MacBook-Pro:~ yangzhang$ netstat -r  
Routing tables
```

Internet:

Destination	Gateway	Flags	Refs	Use	Netif	Expire
default	10.138.127.254	UGSc	458	0	en0	
10.138.64/18	link#4	UCS	12	0	en0	
10.138.95.48	link#4	UHLWII	1	0	en0	
10.138.96.191	link#4	UHLWII	1	0	en0	
10.138.104.235	link#4	UHLWII	1	0	en0	
10.138.106.4	link#4	UHLWII	1	0	en0	
10.138.106.16	link#4	UHLWII	1	1	en0	
10.138.106.31	2c:5b:b8:4b:bd:59	UHLWII	1	2	en0	365
10.138.106.35	link#4	UHLWII	1	0	en0	
10.138.106.106	link#4	UHLWII	1	2	en0	
10.138.106.172	link#4	UHLWII	1	0	en0	
10.138.106.183/32	link#4	UCS	1	0	en0	
10.138.113.60	link#4	UHLWII	1	0	en0	
10.138.127.254/32	link#4	UCS	2	0	en0	
10.138.127.254	0:1a:1e:0:2f:70	UHLWIIr	459	32	en0	793
10.138.127.255	link#4	UHLWbI	1	77	en0	
127	localhost	UCS	1	0	lo0	
localhost	localhost	UH	10	2007167	lo0	
169.254	link#4	UCS	1	0	en0	

Internet6:

Destination	Gateway	Flags	Netif	Expire
default	fe80::1a:1e00:c800	UGc	en0	
localhost	localhost	UHL	lo0	
2001:250:4004:6008	link#4	UC	en0	
2001:250:4004:6008	6c:40:8:ac:2b:a8	UHL	lo0	
2001:250:4004:6008	6c:40:8:ac:2b:a8	UHL	lo0	
fe80::%lo0	fe80::1%lo0	UcI	lo0	
fe80::1%lo0	link#1	UHLI	lo0	
fe80::%en0	link#4	UCI	en0	
fe80::1a:1e00:c800	0:1a:1e:0:2f:70	UHLWIir	en0	
yangs-macbook-pro.	6c:40:8:ac:2b:a8	UHLI	lo0	
fe80::%awdl0	link#8	UCI	awdl0	
yangs-macbook-pro.	9a:de:1:ad:24:a4	UHLI	lo0	
ff01::%lo0	localhost	UmCI	lo0	
ff01::%en0	link#4	UmCI	en0	
ff01::%awdl0	link#8	UmCI	awdl0	
ff02::%lo0	localhost	UmCI	lo0	
ff02::%en0	link#4	UmCI	en0	
ff02::%awdl0	link#8	UmCI	awdl0	

Yangs-MacBook-Pro:~ yangzhang\$

Yangs-MacBook-Pro:~ yangzhang\$

路由选择算法

- 路由器从收到的IP数据报中取出目标IP地址；
- 搜索路由表，寻找能与目的IP地址完全匹配的表目（网络号和主机号都要匹配）。如果找到，则把报文发送给该表目指定的下一站路由器或直接连接的网络接口，路由选择过程结束；如果找不到，则执行下一步操作；
- 否则，搜索路由表，把目标IP地址与子网掩码逐位求“与”后，寻找能与目的网络号相匹配的表目。如果找到，则把报文发送给该表目指定的下一站路由器或直接连接的网络接口，路由选择过程结束；如果找不到，则执行下一步操作；
- 否则，搜索路由表，寻找标为“默认（default）”的表目。如果找到，则把报文发送给该表目指定的下一站路由器；如果找不到，则本次选路失败，发送一个“目标不可达”的ICMP报文给源主机。

路由表的建立与刷新-静态路由

- 网络管理员，根据网络的拓扑结构和连接方式人工建立的
- 静态路由在目的地址和路由器端口之间的关系配置好之后，到达一个特定目标网络的**IP**数据报的路径就固定了
- 特点：安全、资源消耗少

路由表的建立与刷新-动态路由

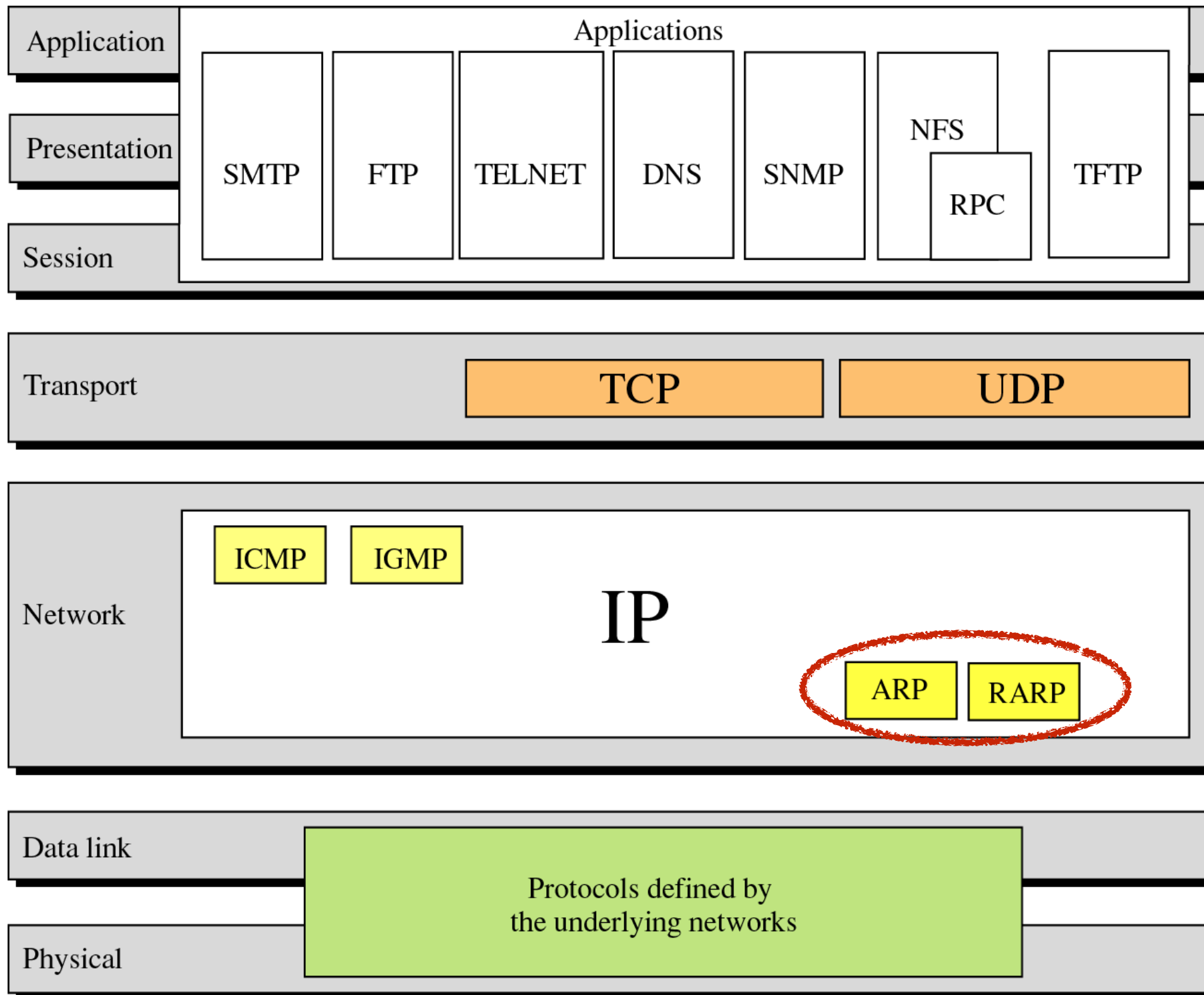
- 动态路由指互联网中的路由器可以通过自身的学习，自动修改和刷新路由表
- 通过相邻路由器之间的通信，以告知对方每个路由器当前所连接的网络，收到路由信息的路由器就用这些信息对自己的路由进行更新
- 路由守护程序（Routing daemon），根据一定选路策略进行（RIP、OSPF等）

- RIP - Routing Information Protocol
 - 采用距离-向量 (Distance-vector) 算法
 - 每个节点将自己与其它节点/网络的路由距离 (跳数) 广播出去, 其它节点根据收到的路由信息计算自己的路由信息, 并广播
 - 周期性广播
 - 优点: 算法简单
 - 缺点: 网络结构发生变化时收敛慢、广播数据过多

- OSPF - Open Shortest-Path First
 - 每个路由器主动地测试与其相连链路的状态，并将这些信息发送给其它相邻的路由器，这样利用扩散，每个路由器可以告诉它所在区域的其它路由器，与它相邻的路由器及其开销
 - 每个节点利用Dijkstra最短路径算法建立一颗路由的树

ARP协议

- Address Resolution Protocol, 地址转换协议
- 将IP地址与物理地址绑定：知道IP地址，根据IP地址获得（以太网上的）物理地址，使数据在以太网上传播
- ARP Cache：缓存了ARP地址，可以用ARP命令来检查和修改ARP高速缓存中的表项
- RARP：ARP的逆过程协议



- 以太网中ARP报文的结构

以太网目标地址(6字节)	以太网源地址(6字节)	帧类型(2字节)	ARP报文(28字节)
--------------	-------------	----------	-------------

硬件类型 (16位)		协议类型 (16位)
硬件地址长度 (8位)	协议地址长度 (8位)	操作代码 (16位)
发送方硬件地址 (以太网为6字节)		
发送方协议地址 (IP地址为4字节)		
目标方硬件地址 (以太网为6字节)		
目标方协议地址 (IP地址为4字节)		

- ARP报文的格式
- “操作代码”字段：区分ARP和逆ARP

- 若ARP高速缓存中没有主机B的映射项，主机自动运行ARP进程来寻找主机B的硬件地址：
 - ARP进程广播发送一个ARP请求分组，内含主机B的IP地址
 - 所有主机上运行的ARP进程都收到此ARP请求分组
 - 机B在ARP请求分组中见到自己的IP地址，就向主机A发送一个ARP响应分组，内含主机B的硬件地址
 - 主机A收到主机B的ARP响应分组后，就在ARP高速缓存中写入主机B的IP地址到硬件地址的映射。
 - （安全问题？）

- ARP高速缓存
 - 每台主机上都有一个ARP高速缓存，存放了最近IP地址到硬件地址之间的映射记录。高速缓冲区中表项建立的方法是：
 - 请求主机收到ARP应答后，主机就把获得的IP地址与物理地址的映射关系存入ARP表中。
 - 由于ARP请求报文是广播发送的，所有收到ARP请求报文的主机都可以把其中发送方的物理地址和IP地址映射存入自己的高速缓存中，以备将来使用。
 - 网络中的主机在启动时，可以主动广播自己的IP地址和物理地址的映射关系，以免其它主机对它提出ARP请求（这也使一台主机在启动时，就可以知道自己的IP地址与网络中其它主机的IP地址有没有冲突）。

```
yangzhang — -bash — 80x24
[Yangs-MacBook-Pro:~ yangzhang$ arp -a
? (10.138.127.254) at 0:1a:1e:0:2f:70 on en0 ifscope [ethernet]
? (10.138.127.255) at (incomplete) on en0 ifscope [ethernet]
? (224.0.0.251) at 1:0:5e:0:0:fb on en0 ifscope permanent [ethernet]
Yangs-MacBook-Pro:~ yangzhang$
```

- 有高速缓存，ARP解析IP地址时首先搜索缓存
- 如没有匹配才会向网络上发送ARP请求报文
- 我们可以用arp命令来检查和修改ARP高速缓存中的表项（静态/动态），如图

IPv6

- IPv4取得了极大的成功
- Internet的快速发展同IPv4地址短缺矛盾越来越突出
- 新技术的应用对IP提出了更多的要求

- IPv4地址资源紧张直接限制了IP技术应用的进一步发展，到1996年已将80%的A类网络地址，50%的B类地址，10%的C类地址全部分配了，曾有专家估计到2010年IPv4地址将全部用完。
- 移动和宽带技术的发展要求更多的IP地址
- CIDR，VLSM，NAT，混合地址等技术只能暂时缓解IPv4地址紧张，但无法根本解决地址问题
- IP地址短缺问题直接加速了IPv4升级的需求

- 最本质的改进：
 - 几乎无限的地址空间，地址长度由32位增加到128位
- 其他（锦上添花）：
 - 简单：简化固定的基本报头，提高处理效率
 - 可扩展：引入灵活的扩展报头，协议易扩展
 - 即插即用：地址配置简化，自动配置
 - 安全：网络层的IPSec认证与加密，端到端安全
 - QoS：新增流标记域
 - 移动：Mobile IPv6

- IPv6发展时间表

- 1992年 IETF成立了IPng工作组
- 1994年 IPng工作组提出下一代IP网络协议（IPv6）的推荐版本
- 1995年 IPng工作组完成IPv6的协议文本
- 1996年 IETF发起成立全球IPv6实验床 - 6BONE
- 1998年 启动面向实用的IPv6教育科研网 - 6REN
- 1999年 完成IETF要求的协议审定和测试
- 1999年 成立了IPv6论坛，开始正式分配IPv6地址，IPv6的协议文本成为标准草案
- 2001年 多数主机操作系统支持IPv6，Windows XP，Linux，Solaris
- 2003年 各主流厂家基本已推出IPv6网络产品
- 2003年 中国启动国家下一代网络示范工程 - CNGI
-

IPv6地址格式举例

- IPv6地址与IPv4地址表示方法有所不同
 - 用十六进制表示，如：FE08:.....
 - 4位一组，中间用“:”隔开，如：2001:12FC:.....
 - 若以零开头可以省略，全零的组可用“::”表示，如： 1:2::ABCD:.....
 - 地址前缀长度用“/xx”来表示，如： 1::1/64
- 以下是同一个地址不同表示法的例子：
 - 0001:0123:0000:0000:0000:ABCD:0000:0001/96
 - 1:123:0:0:0:ABCD:0:1/96
 - 1:123::ABCD:0:1/96

第二章 结束