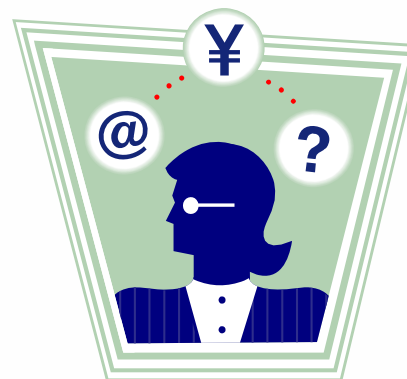


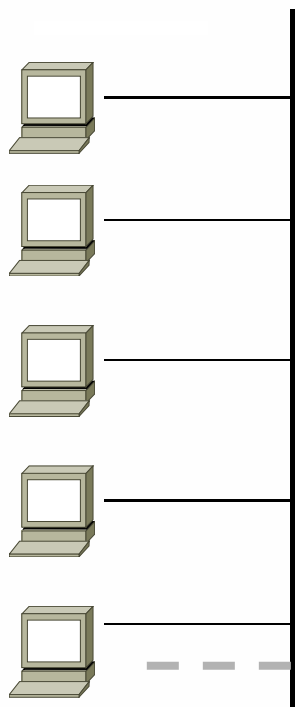


第五章 网络层

课前思考

- 网络层的任务是什么？
- 网络层协议主要涉及哪些内容？
- 报文分组与数据帧有什么关系？
- 报文分组是如何从源主机传输到目的主机的？
- 路由器是如何进行路由选择的？
- 为什么有了MAC地址还需要网络地址？
- 为什么各种异构网络都能接入INTERNET ？



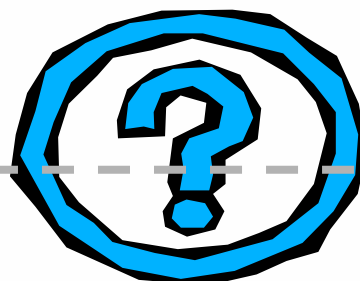


网络1
(以太网)

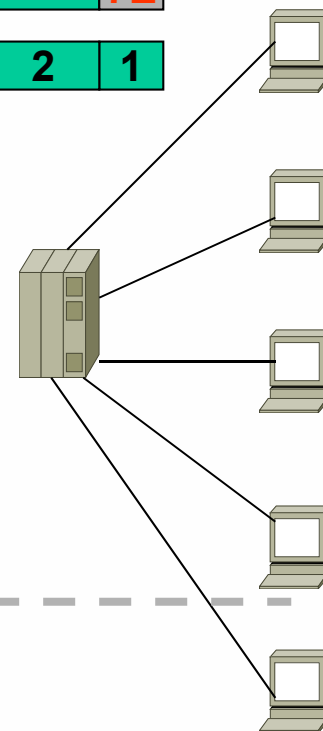


问题一：数据格式不同

问题二：地址不同

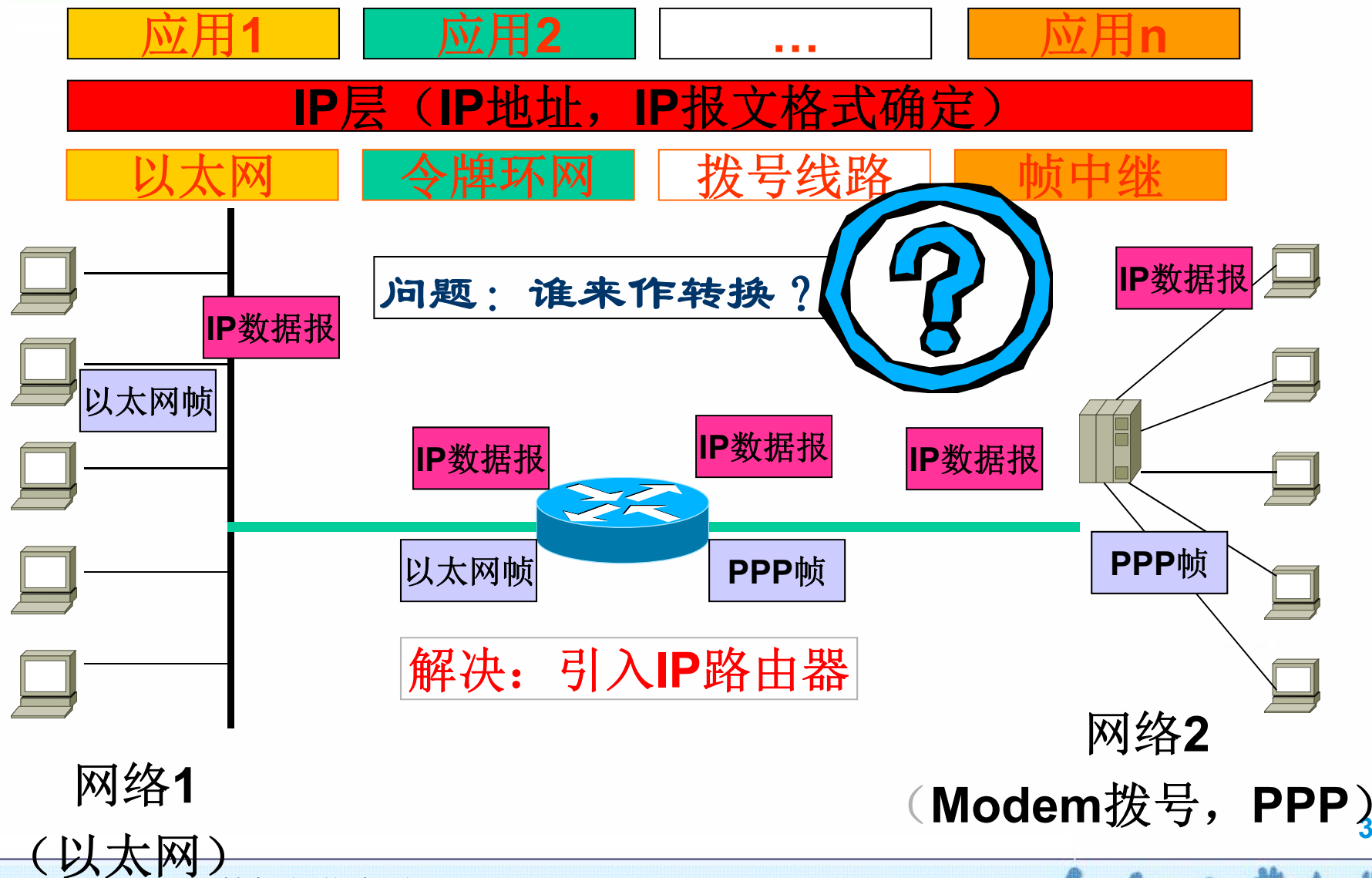


如何通信？



网络2
(Modem拨号, PPP)

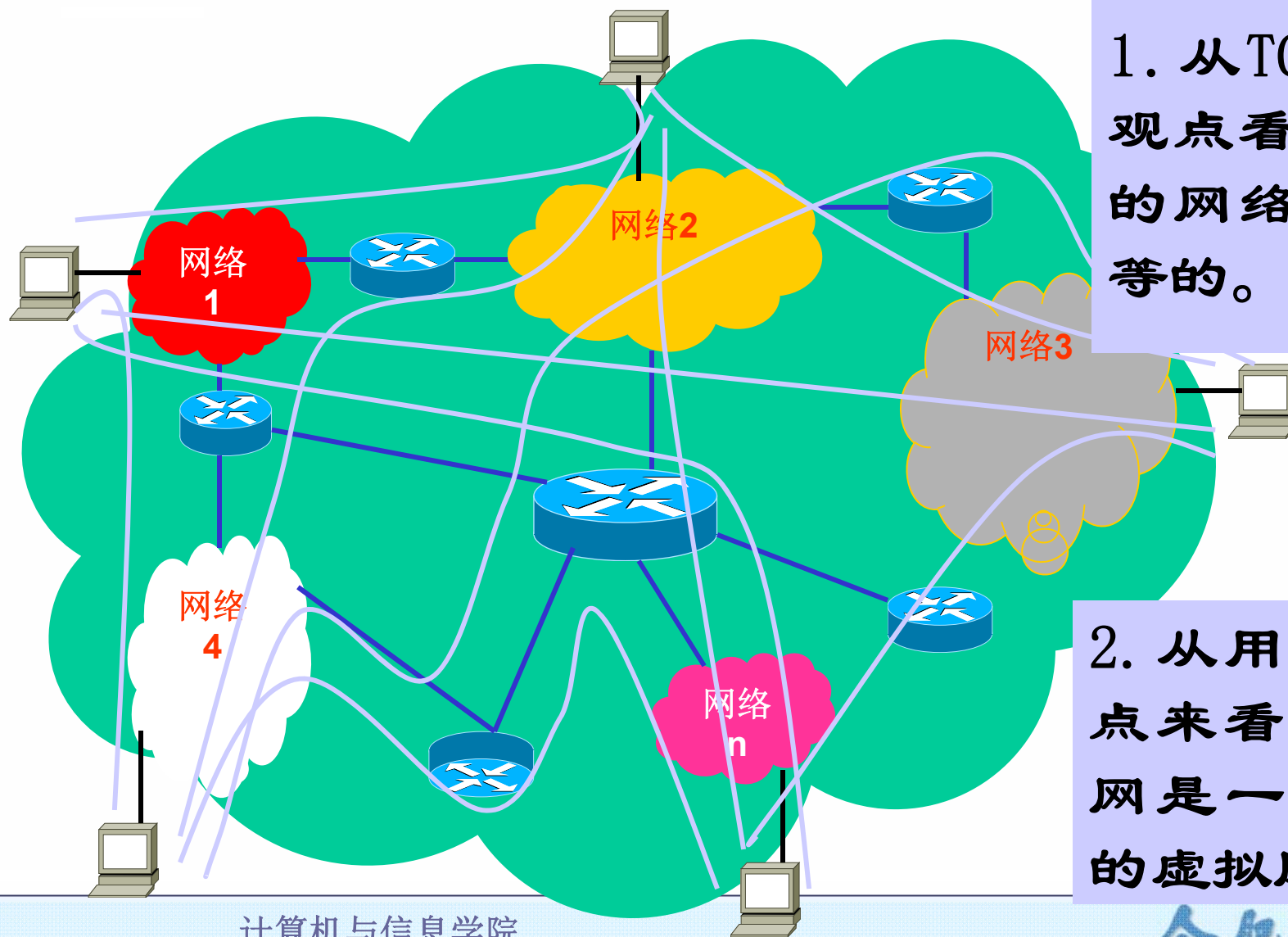






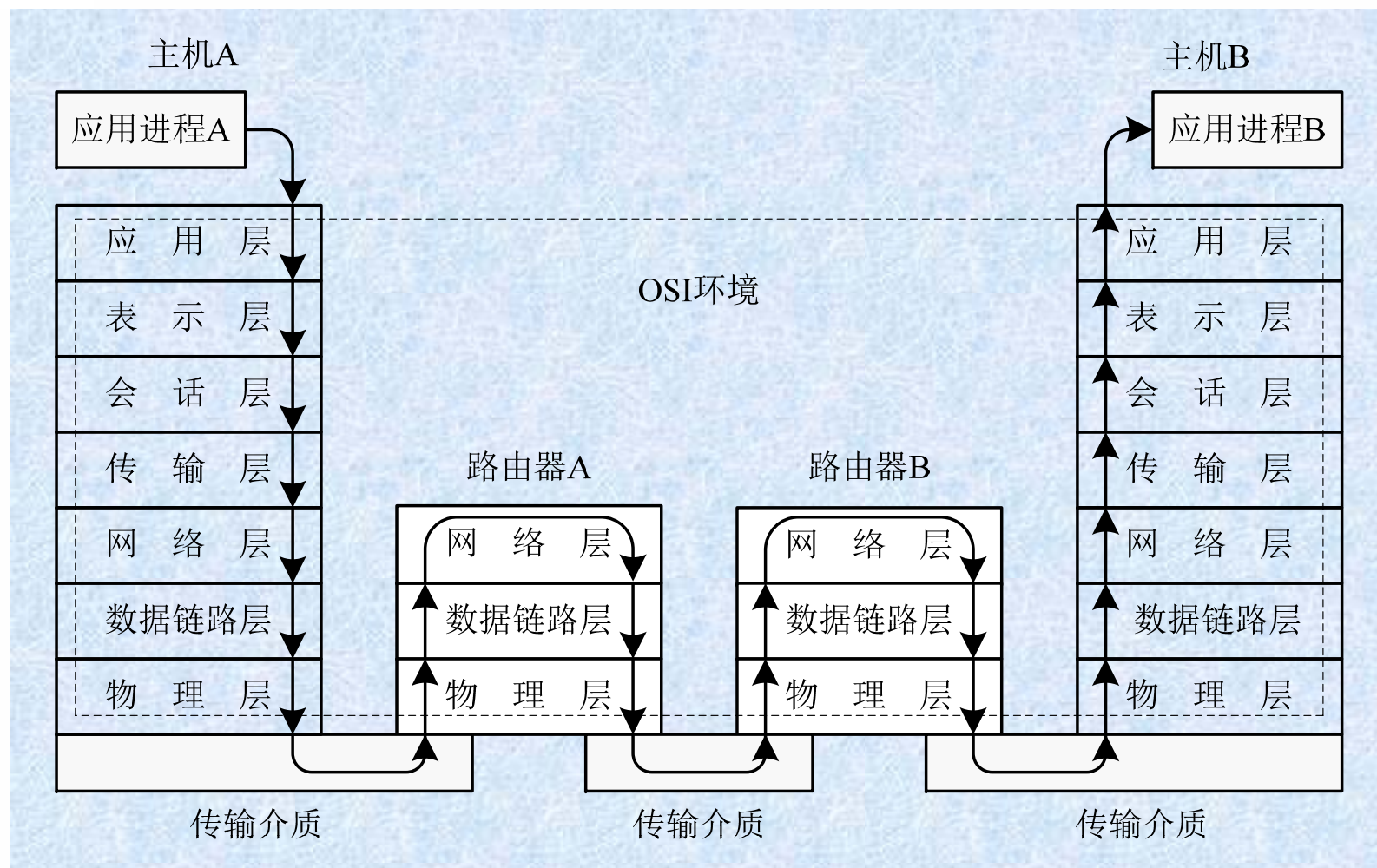
Internet的物理体系结构

用路由器连接起来的多个物理网络



1. 从TCP/IP的观点看，所有的网络都是平等的。

2. 从用户的观点来看，互连网是一个单独的虚拟网络。





本章内容

5.1 概述

5.2 虚电路和数据报

5.3 路由选择算法

5.4 路由协议简介

5.5 IP协议



5.1 概述

- 网络层功能

在数据链路层的支持下，将报文分组从源节点传输到目的节点。

- 网络层为运输层提供服务，它常常是通信子网的边界。

网络层向运输层提供的服务应具有以下特点。

- (1) 服务与通信子网的技术无关。
- (2) 通信子网的数量、类型和拓扑结构对运输层透明。
- (3) 传输层得到的网络地址应该采用统一的编码方式，即使跨越了多个局域网和广域网时也是如此。

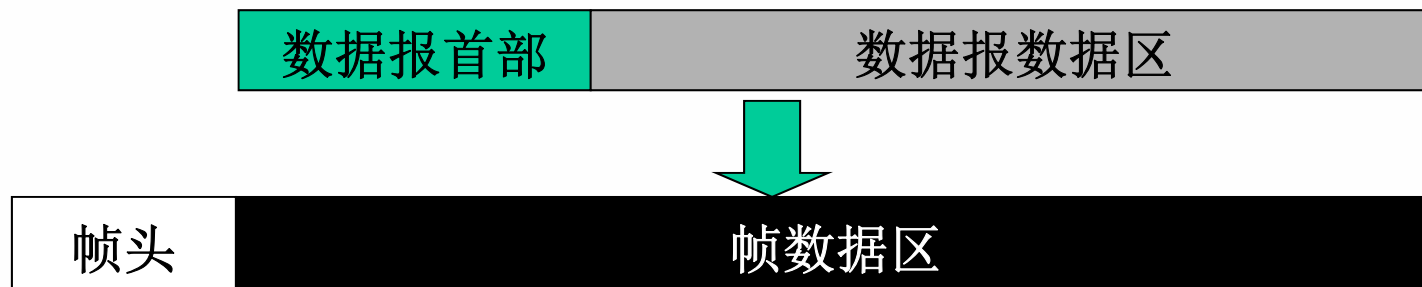


● 报文分组

- 一份完整的信息称为一个报文。
- 报文分组简称“分组”，它是网络层的协议数据单元（NPDU）。
- 分组交换技术将报文划分为若干个较小的格式化信息单位独立传输，这样的格式化信息单位称为报文分组。
- 分组的一般格式：

序号	源地址	目的地址	控制信息	长度	数据
----	-----	------	------	----	----

- 分组作为数据帧的数据部分。





- 网络层协议涉及的主要内容

- 异构网络互联
- 分组格式
- 路由选择
- 拥塞控制
- 网络安全



5.2 分组交换技术

5.2.1 分组交换技术

- 为了提高信道利用率，1964年Baran首次提出分组交换的概念，并于1969年首次在ARPNET上采用分组交换技术。

- 分组交换的基本原理

将一份完整的报文划分为若干个分组，每个分组以存储/转发方式独立的从源结点传输到目的结点，目的结点收到这些分组重新组装成原报文。

- 分组交换技术特征：**化整为零，存储转发。**
- 优点
 - 使多路数据能够复用一条链路，提高链路的利用率。
 - 有利于差错控制。
 - 有利于安全性。
- 缺点
 - 当网络拥塞时，会导致分组传输延迟增加，对流媒体影响较大。
- 分组交换技术的两种实现方式：**虚电路和数据报。**



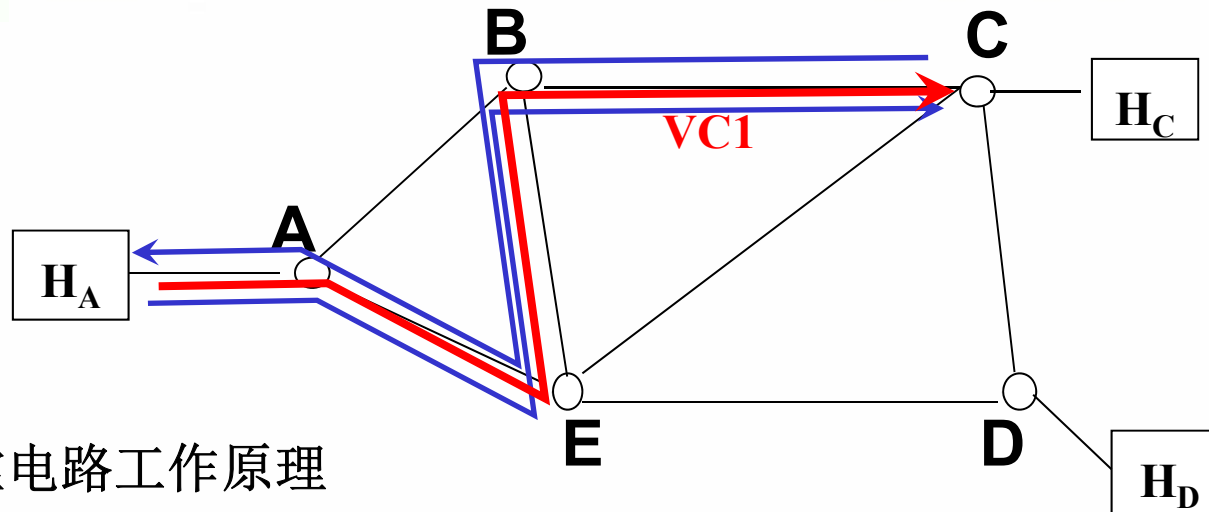
面向连接的虚电路

5.2.2 面向连接服务的虚电路服务

- 面向连接服务就是在数据交换之前，必须先建立连接，当数据交换结束后，则应该终止这个连接。由于面向连接服务和电路交换的许多特性相似，因此面向连接服务在网络层中又称为虚电路服务。
- “虚”的意思是虽然在两个服务用户的通信过程中没有自始至终都占用一条端到端的完整物理电路，但在交换网各个节点中预先申请的资源，从而好像占用了一条完整的物理电路，这些资源连接起来构成了一条虚拟的电路连接。



虚电路的工作原理



- 虚电路工作原理

设：主机 H_A 要向 H_C 发送数据。

- 首先主机 H_A 向 H_C 发一虚呼叫（虚电路连接请求），该虚呼叫选择一条适当的路径传送到 H_C ，记下沿途所经过的路径作为虚电路，如 H_A -A-E-B-C- H_C ，并给其赋一个虚电路号**VC1**。
- 如果沿途所经过的节点及 H_C 准备就绪，则 H_C 发一响应给 H_A ； H_A 收到该响应，则虚电路**VC1**已建立完毕。虚呼叫要向沿途所经过的节点提出请求，任一节点拒绝请求（由于资源不足），均导致虚电路建立失败。
- H_A 和 H_C 的数据交换必须通过该虚电路进行。
- 数据交换完毕，拆除虚电路，释放虚电路占用的资源。



虚电路的特点

● 虚电路的特点

- 通信前, 发送方和接收方之间必须建立连接（虚电路），所以虚电路是面向连接的网络服务。
- 虚电路只是一种**逻辑电路**，而不是真正的物理电路。报文分组在虚电路上传输不像在物理电路上那样中畅通无阻, 而是要经过节点的“存储转发”。
- 一旦虚电路建立完毕，所有分组必须经过该虚电路进行。因此，虚电路能够保证分组的顺序接收。
- 仅当建立虚电路时需要源/目的节点地址，数据分组只需分配一个虚电路号而无需携带源/目的节点地址。



无连接的数据报

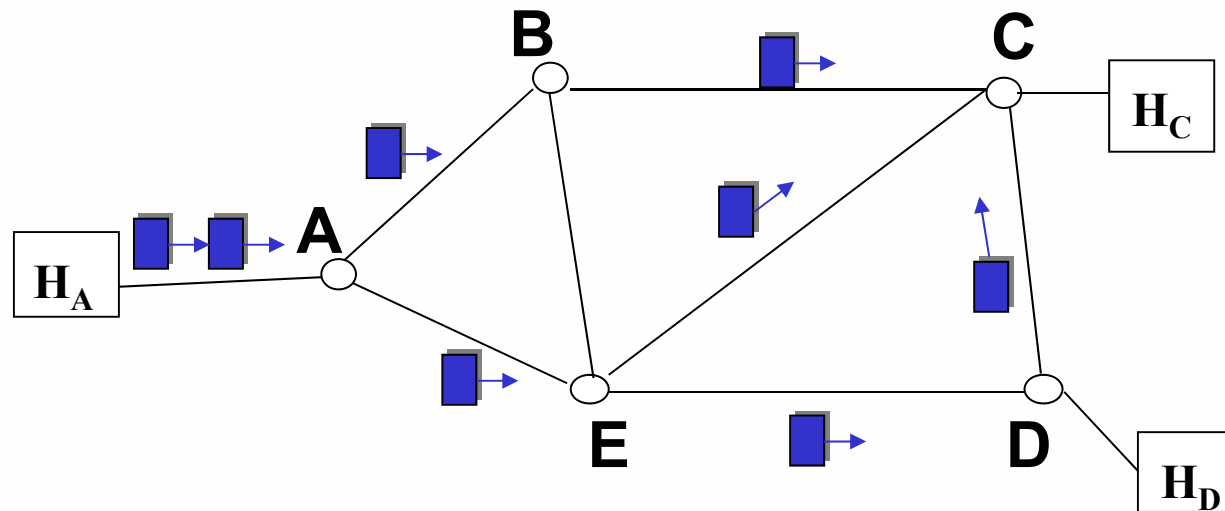
5.2.3 无连接服务的数据报服务

- 在无连接服务的情况下，两个实体之间的通信不需要先建立好一个连接，因此其下层的**有关资源不需要事先进行预定保留**，这些资源是在数据传输时动态地进行分配的。
- 网络中的交换机根据每个数据分组中的目的地址独立地为其临时确定路由。



数据报

- 数据报无需建立连接，每个报文分组携带完整的源/目的地址，独立的选择路径，通过不同的路径到达目的主机





数据报的特点

● 数据报的特点

- 无需建立连接就可传输报文分组，因此数据报称为非面向连接的网络服务。
- 不同的分组可以通过不同的路由到达目的主机，先发出的分组未必先到；因此不能保证分组的顺序到达。
- 每个报文分组携带完整的源/目的地址，独立的选择路径。



虚电路与数据报的对比

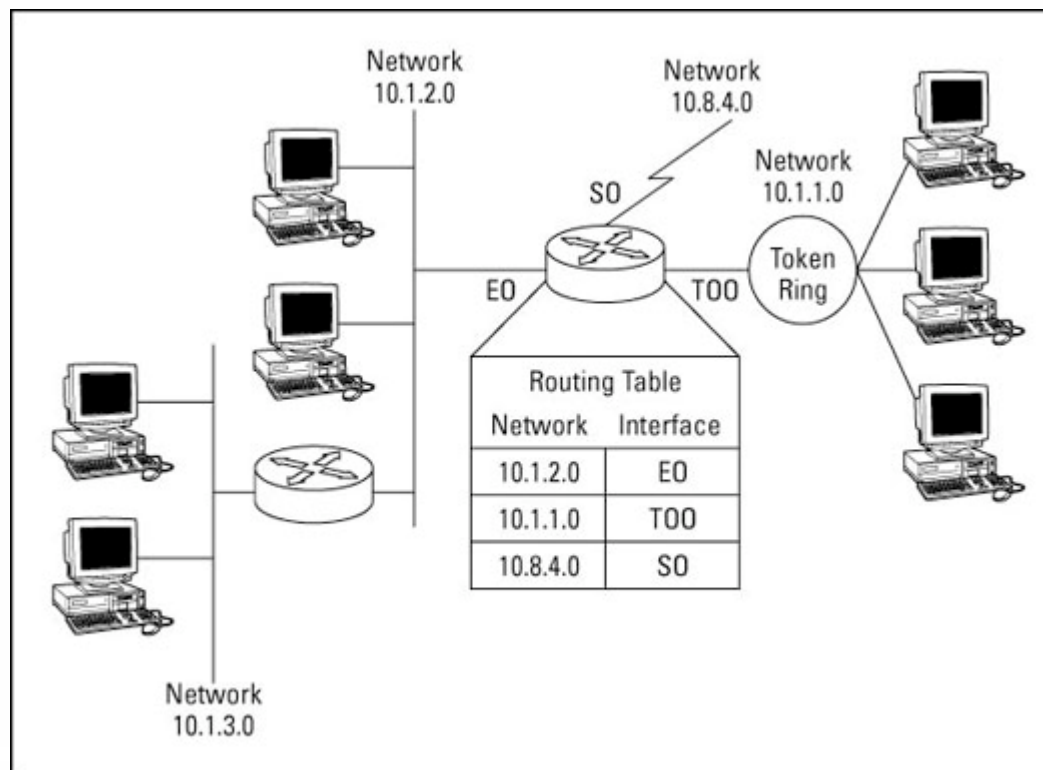
5.2.4 数据报与虚电路的比较

	数据报子网	虚电路子网
延时	分组传输延时	电路建立，分组传输延时
路由选择	每个分组单独选择路由	建立虚电路时选择路由，以后所有分组都使用该路由
状态信息	子网无需保存状态信息	每个结点要保存一张虚电路表
地址	每个分组携带完整的源/目的地址	每个分组分配一个较短的虚电路号
结点失败的影响	除了在崩溃时正在该结点处理的分组都丢失外，无其他影响	所有经过失效结点的虚电路都要被终止
拥塞控制	难	容易



5.3 路由选择

- 在通信子网内部，分组从一个网络节点转移到另外一个网络节点，直至到达目的用户所连接的网络节点，分组在其中的转发过程称为路由选择转发过程。
- 路由选择通过路由表（**routing table**）来实现。路由表是根据一定的路由选择算法得到的，而转发表又是根据路由表构造出的。
- 路由选择协议负责搜索分组从某个节点到目的节点的最佳传输路由，以便构造路由表。





网络地址

● 网络地址

- 网络地址是主机或节点的标识，即给定一个网络地址在整个网络中唯一地标识一个主机或节点。
- 网络地址是结构化地址，由网络号和主机号组成。



IP地址

- **IP网络中的每台主机至少要分配一个IP地址，一台路由器需要分配多个IP 地址（每个端口需分配一个IP地址），每个IP地址只能分配给一台主机或路由器。**
- **Internet上IP地址由网络信息中心（NIC）分配。**
- **IP地址32位，通常用带点的4个十进制数表示；如：
240.45.245.9。**



- A类：共有 $2^7-2=126$ 个网络，每个网络最多有 $2^{24}-2\approx 1.6\times 10^7$ 台主机。
- B类：共有 $2^{14}-2=16382$ 个网络，每个网络最多有 $2^{16}-2=65534$ 台主机。
- C类：共有 $2^{21}-2\approx 200$ 万个网络，每个网络最多有 $2^8-2=254$ 台主机。
- D类：用于组播。
- E类：保留。

● IP地址结构

	0	8	16	24	31	
A类	0	网络号	主机号			1.0.0.0—127.255.255.255
B类	10	网络号		主机号		128.0.0.0—191.255.255.255
C类	110	网络号			主机号	192.0.0.0—223.255.255.255
D类	1110	组播地址				224.0.0.0—239.255.255.255
E类	11110	保留				240.0.0.0—247.255.255.255



IP地址分配

- 连接到Internet的每一个主机（计算机或路由器）至少有一个IP地址。
- IP地址是分配给网络接口的。
- 多归属主机可以有多个IP地址。
- 一个网络接口也可以分配多个IP地址。
- 网桥、Ethernet交换机、集线器Hub属于数据链路层设备，使用MAC地址，不属于网络层设备，不分配IP地址。



特殊的IP地址

● 特殊的IP地址

- 全“0”：表示本机地址，仅当初始启动时使用，以后不再使用。

- 主机号全“1”：广播地址。

例如，192.210.200.255表示192.210.200.0的所有主机。

- 主机号全“0”：表示网络号。

例如，例192.254.252.0表示一个C类网络。

- 127.xxx.yyy.zzz：系统保留作为回路自测。

发到127.xxx.yyy.zzz的分组实际上并不发送到网络上，而是经过内部处理后又返回本机。



● 内部IP地址（私用地址）

- A类：10.0.0.0 （1个网络）。
- B类：172.16.0.0 — 172.31.0.0 （16个网络）。
- C类：192.168.0.0 — 192.168.255.0 （256个网络）。

Internet的路由器不会转发目的地址为上述地址的**IP**分组，即私用地址不能穿越内部网络边界。

类	网络号	总数
A	10.	1
B	172.16~172.31	16
C	192.168.0~192.168.255	256

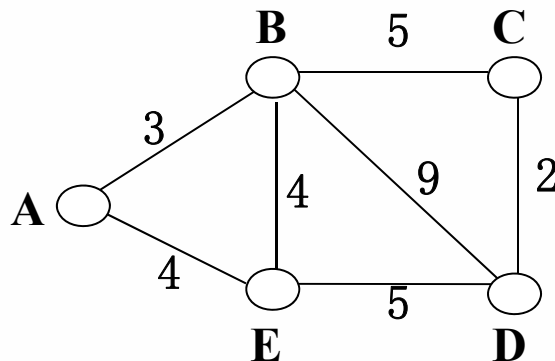


路由表

● 路由表一般结构

目的节点	最短距离	最佳输出链路
------	------	--------	-------

这里的“距离”是广义的，可以是距离、平均流量、延迟、下跳数等。



节点C 的路由表

目的节点	最短距离	最佳输出链路
A	8	C → B
B	5	C → B
C	0	——
D	2	C → D
E	7	C → D



路由选择实现的方法

- 路由器通过路由选择算法，建立并维护一个路由表。
- 路由器根据目的网络号而不是目的IP地址转发IP分组，以减少路由表的信息量。
 - 路由器保存一张路由表，该表保存的主要内容：（目的网络号，最佳输出链路），指示如何到达目的网络，至于到达该网络后，如何到达目的主机则不是本路由器所考虑的。
- 在路由表中包含着目的地址和下一跳路由器地址等多种路由信息。
- 路由表中的路由信息告诉每一台路由器应该把数据包转发给谁，它的下一跳路由器地址是什么。
- 路由器根据路由表提供的下一跳路由器地址，将数据包转发给下一跳路由器。
- 通过一级一级地把包转发到下一跳路由器的方式，最终把数据包传送到目的地。
- 如果路由表中找不到目的网络，则将该分组转发到“缺省链路”。



路由选择及其分类

5.3.1 路由选择及其分类

- 路由选择：根据某种策略，选择一条到达目的主机的最佳路径。
- 路由选择功能由路由器完成；对单个路由器而言，路由选择实质上是选择最佳输出链路（端口），多个路由器协作选择一条最佳路由。
- 无论是虚电路，还是数据报都要进行路由选择。虚电路需要一次路由选择，数据报需要为每个分组选择路由。



路由选择分类

路由选择分类

- 静态路由:

- 按照某种固定的规则进行路由选择，不随网络流量和拓扑结构变化而变化。

- 动态路由:

- 根据当前拓扑结构和流量的变化来动态改变路由，又称为自适应路由。



静态路由

5.3.2 静态路由算法

● 扩散法（洪泛法）

- 当节点收到一个分组后，向除进来的链路外的所有其他链路转发（扩散），其结果是至少有一个分组以最快的速度到达目的节点。
- 问题：扩散过程产生大量重复分组，导致网络无法运行。
- 解决措施：每个分组设置一个下跳数字段，每经过一个结点，下跳数减1，当下跳数为0时，丢弃该分组。



固定式路由

● 固定式路由选择

- 网络管理员为每个路由器配置固定路由表，固定路由表一旦生成，就不再改变，除非网络管理员重新配置。
- 优点：简单，路由算法开销小。
- 缺点：不能适应网络流量和拓扑结构的变化。
- 适用：小规模网络。



动态路由

5.3.3 动态路由算法

- 热土豆算法

- 基本思想

当节点收到一个分组后，选择一条输出队列最短的链路尽快的将其转发出去，而不管目的节点位于何方。

- 优点：提高链路的利用率。
 - 缺点：盲目性。



逆向自学习算法

● 逆向自学习算法

- 每个节点保存一张转发表（路由表），该表主要字段包括：

.....	目的地址	输出接口	时间戳
-------	------	------	-----	-------

- 初始时转发表为空。
- 当数据包到达节点时：
 - 将数据包的源地址视为转发表的目的地地址；如果转发表存在该目的地地址，则刷新该记录；如果转发表不存在该目的地地址，则增加一条新记录。
 - 根据数据包的目的地地址，查询转发表；如果找到，则从指定的输出端口转发；如果找不到，则广播。
- 定期扫描转发表，清除过时的记录。



距离向量法

- 距离向量路由算法（D-V算法）

- 该算法最早在ARPANET中使用，后在Internet及Novell网的IPX中使用，即RIP协议。

- 基本思想

每个节点都保存一张动态路由表，路由表包括目的节点地址，最短距离、最佳输出链路。与固定式路由选择不同的是：相邻结点之间定期交换路由信息（如每隔30秒），并根据最新路由信息，刷新路由表。

- 初始化

- 当节点加入网络时，获取直接相连的链路情况（无需知道网络全局拓扑）；
- 构建初始路由表，初始路由表中的目的节点仅包含直接相连的节点，“距离”值置为0；
- 将初始路由表发给直接相连的节点。



步骤:

- ① **初始化**: 路由器启动时, 对每个直接相连的网络生成一个表项, hop数都为0。
- ② **路由交换**: 路由器周期性向相邻路由器广播自己的整个路由表。(交换信息是 $\langle V, D \rangle$)
- ③ **路由表更新**: 路由器每收到一个邻站的路由表, 即更新自己的路由表。(假设K收到J的路由表)
 - (1) K不知道目的站, 则**加入**
 - (2) 有通过J的更短路, 则**替换**
 - (3) 原下站为J的距离有变化, 则**修改**



例：

目的站	距离	下一跳
网络1	0	直接
网络2	0	直接
网络4	8	路由器L
网络17	5	路由器M
网络24	6	路由器J
网络30	2	路由器Q
网络42	2	路由器J

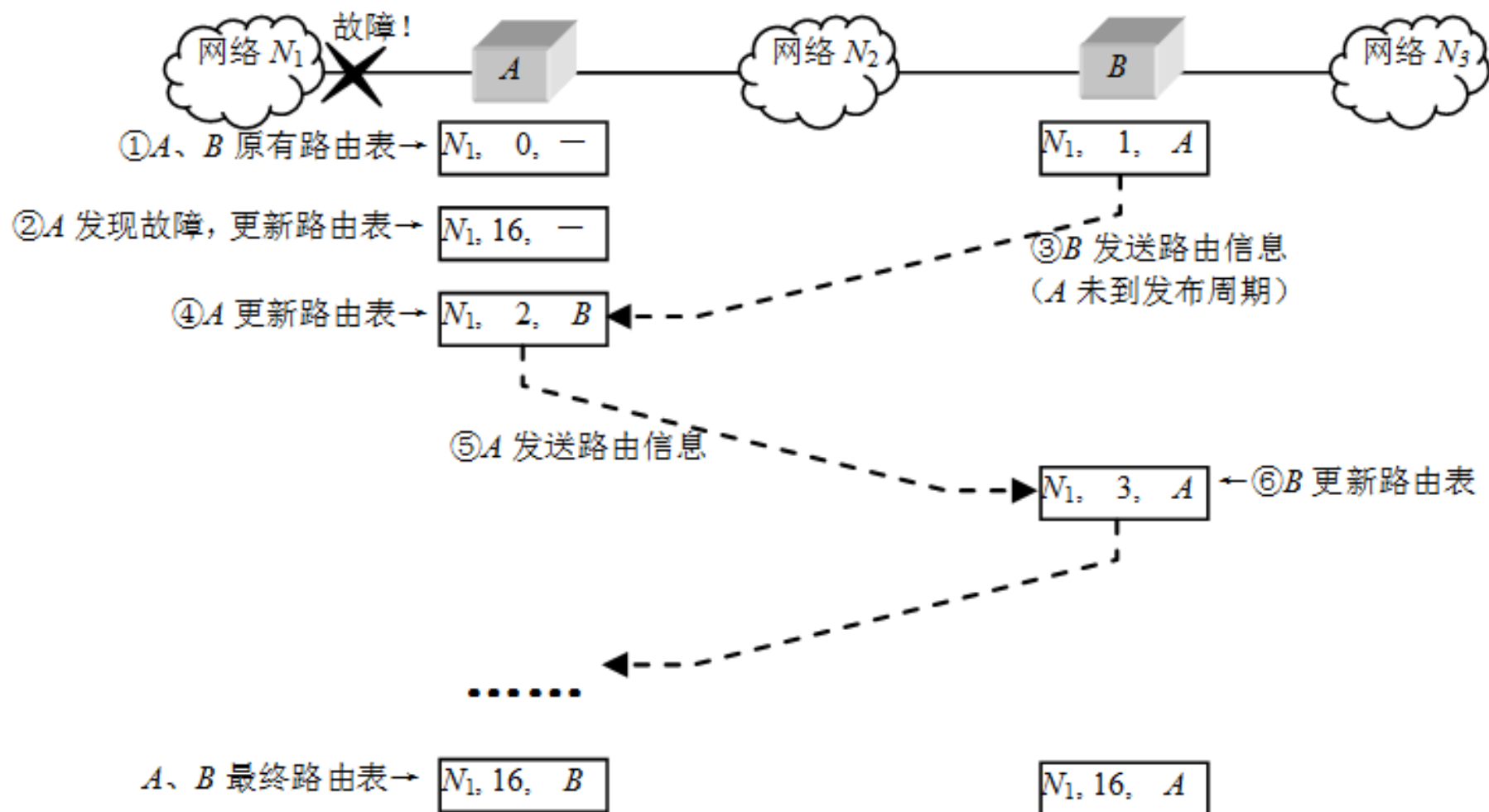
目的站	距离	下一跳
网络1	0	直接
网络2	0	直接
网络4	4	J（替换）
网络17	5	路由器M
网络21	5	J（增加）
网络24	6	路由器J
网络30	2	路由器Q
网络42	4	J（修改）

目的站	距离
网络1	2
网络4	3
网络17	6
网络21	4
网络24	5
网络30	10
网络42	3



●D-V算法优缺点

- 优点：由于仅相邻节点交换路由信息，所以运算量和交换的信息量较小。
- 收敛速度慢，对网络变化需经若干周期才能作出反应。特别是对好消息反应快，对坏消息反应迟钝。





● 无穷计数问题的解决

- 规定一个足够大的数作为 ∞ ，如**RIP**规定为**16**，缺点是限制了网络规模。
- 水平分割法，即不允许将从相邻节点获得的路由信息再提供给该相邻节点。

由于C到A的路由是B提供的，所以B不接受C提供的关于A的路由信息。

	A	B	C	D	E
	○	○	○	○	○
	×				
初始时		1	2	3	4
第1次交换后	∞		2	3	4



链路状态法

- 链路状态路由选择算法（L-S算法）

- 从**1979**年开始，ARPANET以及后来的Internet的内部网关协议由距离向量算法改为链路状态算法。

- 基本思想

所有节点相互交换路由信息，并根据最新路由信息刷新路由表。



● 链路状态路由选择算法（L-S算法）

● L-S算法描述

➤ 发现邻居结点

当一个路由器启动以后，通过向每个端口发送特殊的HELLO分组来发现邻居结点，收到HELLO分组的路由器应返回一个应答来说明它的网络地址。

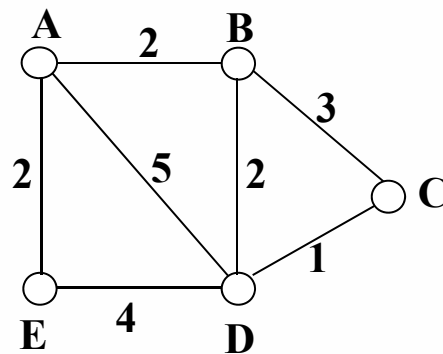
➤ 测量相邻链路开销

通过发送一个特殊的ECHO分组来实现，测量其往返时间，再除以2。



➤ 产生链路状态分组

每个节点实测所有相邻链路的开销，创建链路状态分组（L-S分组）。



D	
序号	
生存期	
A	5
B	2
C	1
E	4

D节点当前L-S分组

D: 表示节点D产生的L-S分组。

序号: IP协议用**32**位表示; 如果一个L-S分组到达, 其序号比最近到达的序号小, 则丢弃, 以保证节点收到是最新的L-S分组。

生存期: 在L-S分组广播过程中, 经过一个结点递减1, 一旦生存期为0, 则丢弃。



➤ 广播L-S分组

每个节点向其他所有节点广播自己的L-S分组。

➤ 刷新路由表

结点获得最新L-S分组后，用最短路径算法计算到其他结点的最短路由，并刷新路由表。

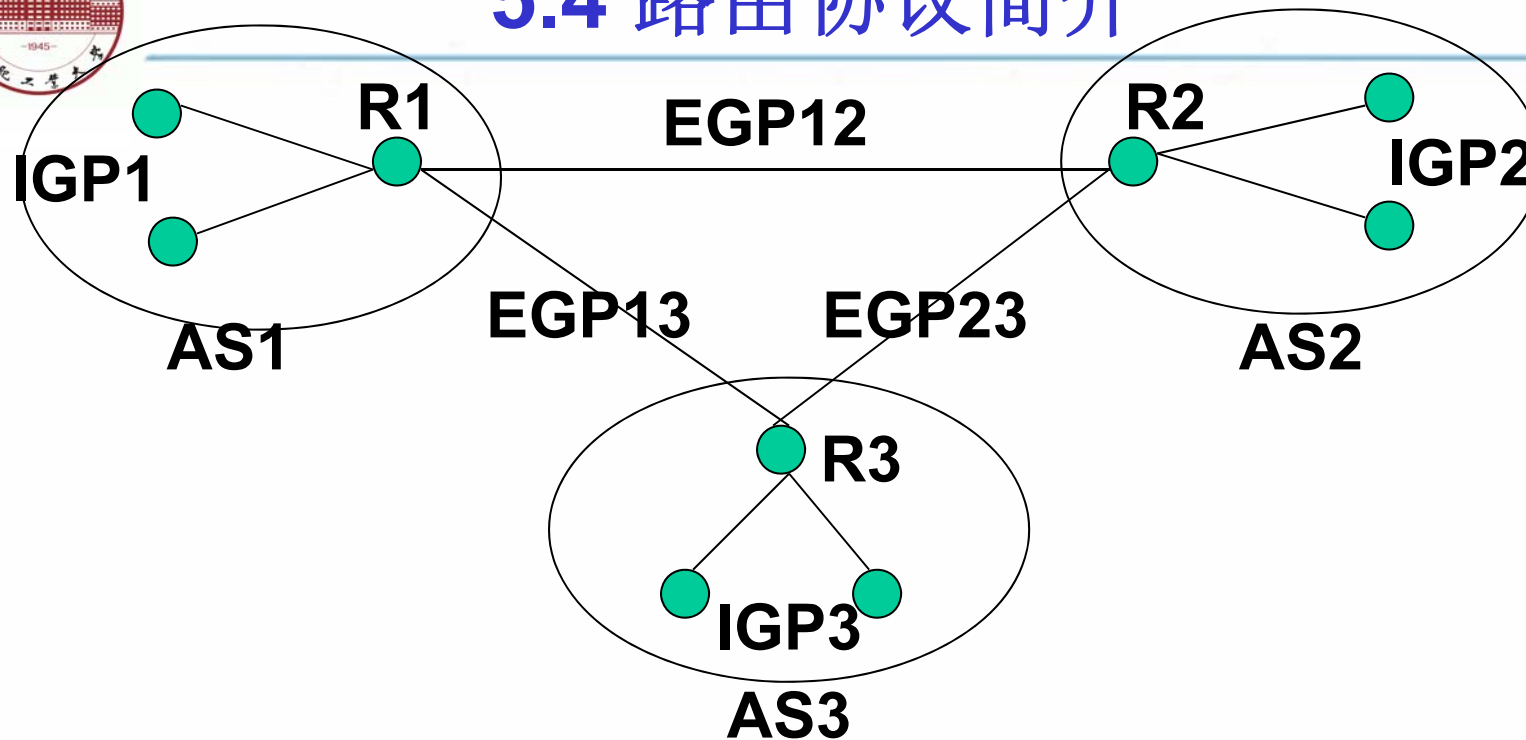
- 优点：对网络变化反应迅速（只需一个周期）。
- 缺点：广播L-S分组占用信道带宽大。
- 应用：Internet的内部网关协议采用L-S算法。

● D-V和L-S算法的比较

	D-V	L-S
交换路由信息	定期	网络拓扑发生改变或定期
交换范围	相邻结点	全网
路由更新	缓慢	迅速
适用范围	变化缓慢的网络	变化较激烈的网络



5.4 路由协议简介



Internet的路由管理模式:

- (1) AS内部: IGP, 比如RIP、OSPF、IS-IS等;
- (2) AS之间: EGP, 最常用的是BGP;
- (3) EGP通常是一种可达性协议。



5.4.1 路由信息协议（RIP）

- **RIP采用D-V算法，用于小规模网络。**
- **技术特点**
 - **距离：下跳数，允许对下跳数加权。**
 - **路由信息交换周期：30秒。**
 - **无穷计数问题：RIP选择16作为 ∞ ；为了加快收敛速度，RIP采用水平分割技术。**
 - **RIP消息交换：通过UDP协议传输，端口号为520。**



OSPF协议

5.4.2 开放最短路由优先协议（OSPF）

- OSPF采用L-S算法，是目前Internet的主要内部网关协议。
- 技术特点
 - 距离：允许选择多种“距离”度量，如延迟，数据率，通信费用，下跳数等。
 - OSPF支持区域概念。
 - OSPF支持认证服务，防止发送假路由信息来愚弄路由器。



边界网关协议 (BGP)

5.4.3 边界网关协议 (BGP)

- BGP采用改进的D-V算法，作为Internet外部网关协议。
- 技术特点
 - 路由表中记录到达目的地的确切路由，而不是“距离”，从而解决“无穷计数”问题。
 - 支持策略路由。



5.5 IP协议

5.5.1 IP概述

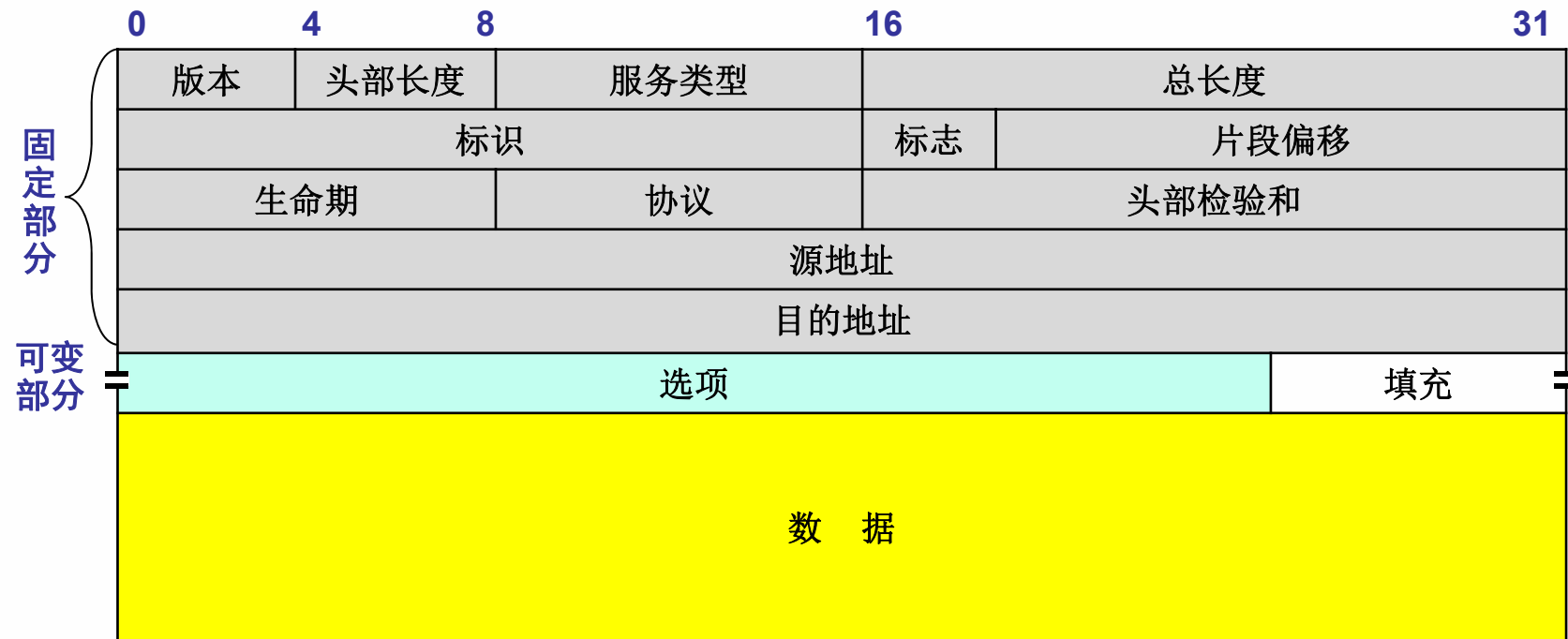
- IP协议是Internet体系结构的核心协议，该协议已成为连接异构网络的工业标准。
- IP提供非面向连接的数据报服务，每个IP分组长度 $\leq 64K$ 字节，传输效率高，但不能保证分组可靠的，按序到达。
- IP协议包括IPv4和IPv6两个版本，本节主要介绍IPv4。



IP分组格式

5.5.2 IP分组格式

- IP分组由IP报头和数据组成。



- 版本：4位，表示IP协议的版本号；4表示IPv4，6表示IPv6。
- 头部长度：4位，给出IP报头的长度，单位为字（32位）；最小值为5（无任何选项），最大值为15（头部最大长度为60字节）。



- **服务类型：**告诉路由器该IP分组想获得何种服务，包括优先级、延迟、吞吐量和可靠性要求。该字段共有**8位**：
 - 前**3位**表示优先权。
 - 中**3位**（D、T、R）表示想获得的服务质量（延迟，吞吐量，可靠性）。
 - 后**2位**保留。
 - 几乎所有路由器都忽略该字段。
- **总长度：**表示整个IP分组长度，包括报头和数据部分；该字段**16位**，所以IP分组最大长度 $2^{16}-1=65535$ 字节（ $\approx 64\text{KB}$ ）。
- **标识：** **16位**，用来标识一个分组，同一分组的不同分段具有相同的标识。
- **标志：** **3位**。
 - 第**1位**：保留
 - 第**2位**：**DF位（Don't Fragment）**。若**DF=1**，则IP分组在传输中不允许分段，如目的主机从远方引导启动机器的映像文件，目的主机无能力重组分组。若**DF=0**，则允许分段。
 - 第**3位**：**MF位（More Fragment）**：若**MF=0**，则该分段是原分组的最后一个分段。



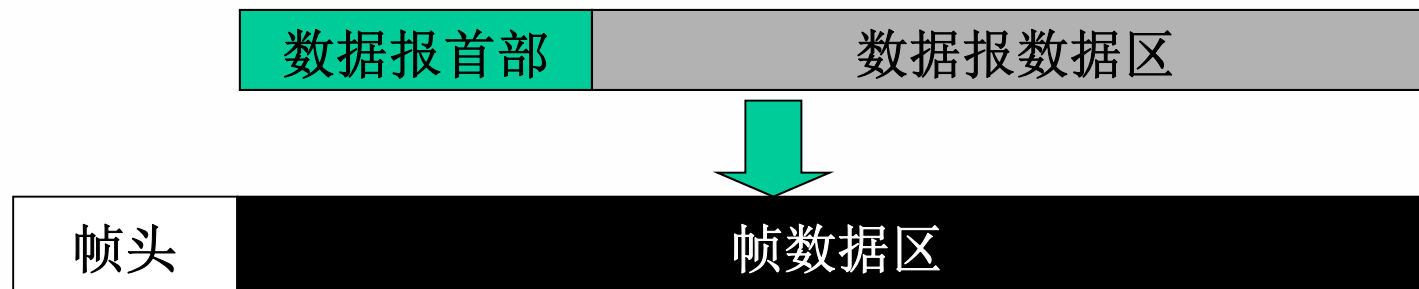
5.5 IP协议

- 分段偏移：表示分段在原分组中所处位置的偏移量，单位为8字节，起始偏移量为0。由于该字段为13位，所以一个分组最多有 $2^{13} = 8192$ 个分段。



数据报的分片控制

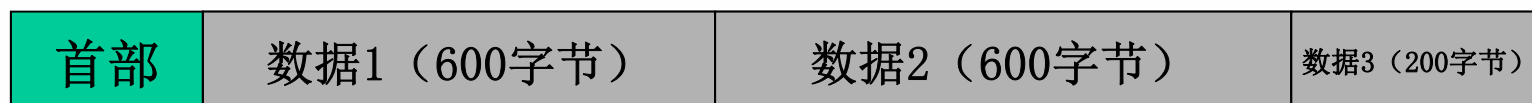
1. 数据报的封装（以太网帧类型0800H）

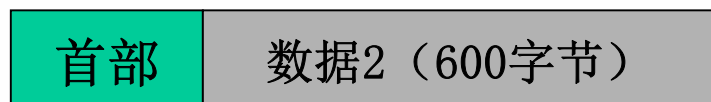
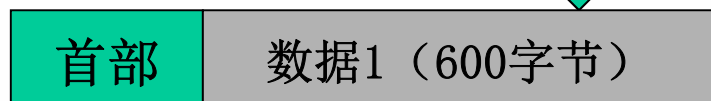
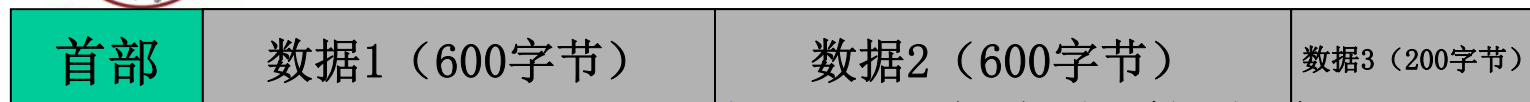


问题：每个网络有固定的MTU，若IP报文长度大于MTU，怎么办？

2. 数据分片

例：IP数据报长度1420字节，网络MTU620字节，如何分片？





问题1: 如何组装分片?

(1) 如何标识同一个数据报的各个分片?

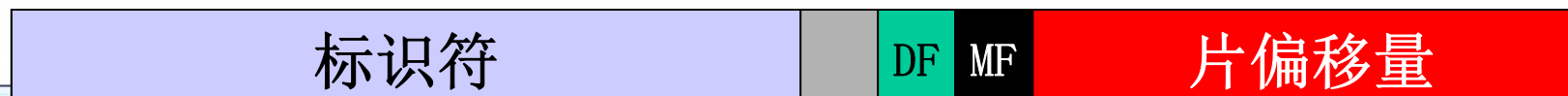
(2) 如何标识同一个数据报分片的顺序?

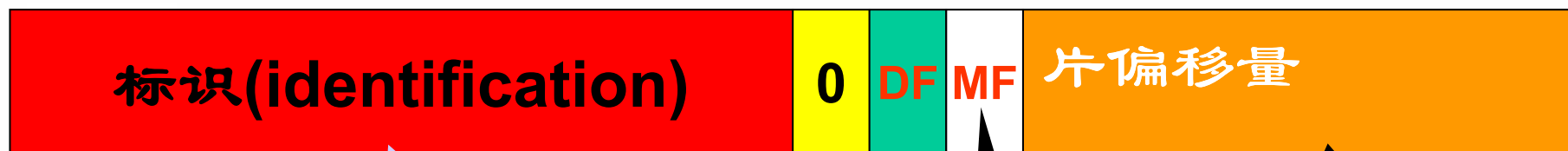
(3) 如何标识同一个数据报分片的结束?

解决: 重新设置**首部**的某些字段

(1) 修改分片标志和片偏移量字段

(2) 首部其它字段复制原来数据报首部的相应字段



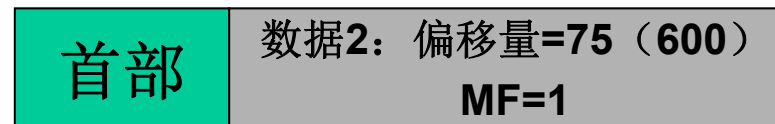
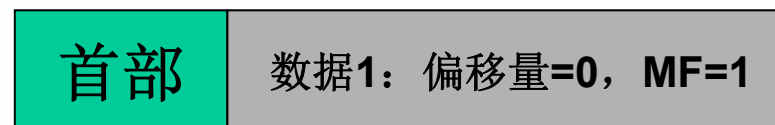


信源机产生，每个数据报唯一
解决了：标识同一数据报的各个分片

标识分片在原来数据报文中的位置
解决了：分片的顺序

DF (Do not Fragment):
不分片位
DF=1，强制不允许分片

MF (More Fragment): 片未完位
MF=0，是最后一片
MF=1，不是最后一片
解决了：分片的结束





问题2：分片首部如何设置？

解答：第一个分片的首部复制原IP数据报的首部，并修改总长度字段，之后的分片首部重新设置。

- (1) 首部长度：若无选项，和原数据报相同，否则重新计算。
- (2) ID、标志域和片偏移量：如问题1所述。
- (3) 总长度：分片首部的长度+分片数据部分的长度。
- (4) 校验和：必须重新计算。
- (5) 选项：记录路由和时间戳选项只能复制到第一个分片中。



问题3：何处重组分片？

答 案：目的主机

优 点：

- (1) 不会反复分片/重组
- (2) 各分片可独立选路
- (3) 路由软件简化

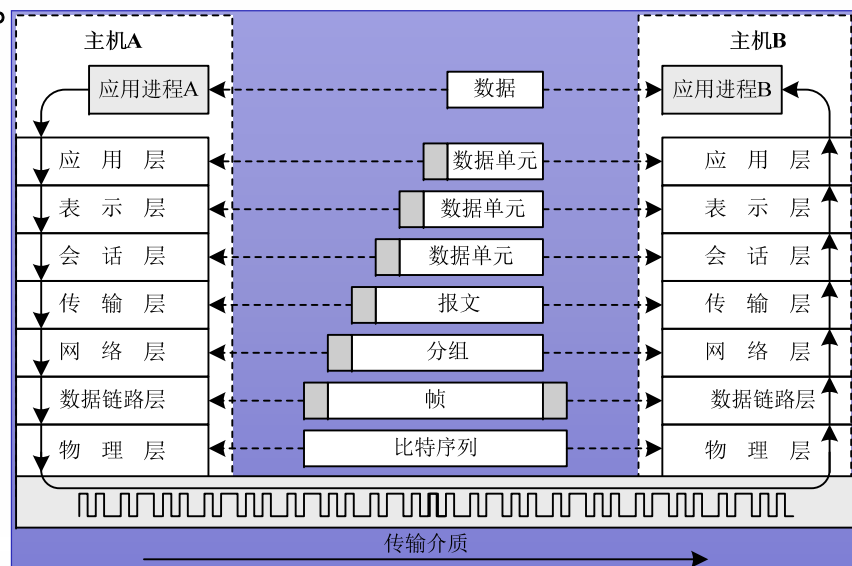
缺点：

- (1) 可能浪费带宽
- (2) 丢失可能性增加



IP分组字段

- **生命期**：该字段限制IP分组在Internet中的生存期，通常用下跳数表示，组每经过一个路由器，则自动减1；当该字段为0时，则丢弃该分组，并向源节点发出ICMP警告信息。
- **协议**：表示IP的上层协议，即IP分组将递交给哪一个高层协议处理（如TCP或UDP）。
- **头部校验和**：将头部所有16位比特字段按二进制补码累加，再取其结果补码，以便接收端对IP头部进行校验。该字段每经过一个路由器都要重新计算，因为至少“生命期”字段会改变。
- **源/目的地址**：源节点和目的节点IP地址（各32位）。



57



● 选项

➤ 安全

告诉路由器如何保证**IP**分组的安全性；例如，该分组不能经过某些路由器。

➤ 严格的源路径

以**IP**地址序列形式给出完整的源/目的路径，分组严格地按照此路径传输。

➤ 宽松的源路径

以**IP**地址序列形式给出到达目的节点必须经过的路由器，但也允许经过其他路由器。



●选项

➤路由记录

告诉沿途经过的路由器将他们的**IP**地址加入到该选项中，以便接收端能够知道该分组经过哪条路由到达。

➤时间标记

类似“路由记录”，但除记录路由器的**IP**地址外，还要记录时间标记。



Internet控制协议

一、为什么需要ICMP？

- (1) 数据报在传送中可能会遇到各种异常；
- (2) IP层也需要控制功能（拥塞、差错控制）。

二、ICMP的作用

传递网络控制信息、提供差错报告。

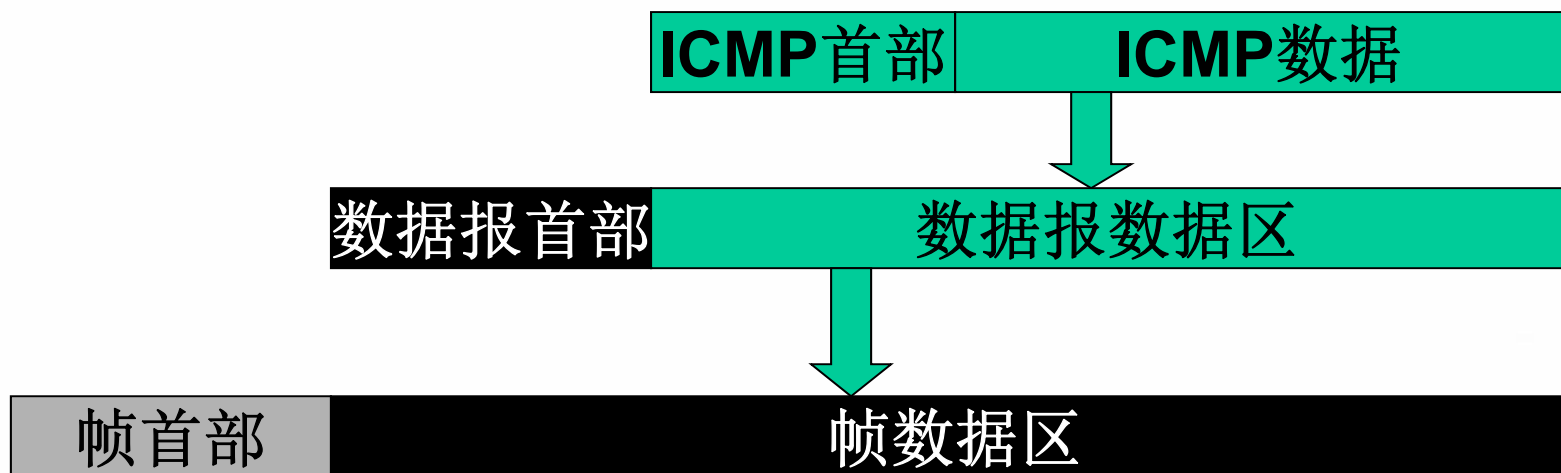
三、ICMP的工作机制

- (1) 路由器（主机）向其它路由器（主机）发送ICMP报文。
- (2) 只能向源端报告差错，不能向中间路由器报告差错。



ICMP

- 这些ICMP消息封装在IP分组中传输，其IP首部“协议”域指明为ICMP报文（代码为1）。因此，ICMP消息可视为“IP控制分组”。
- ICMP与IP同层，只是IP软件中的一个模块





主要ICMP消息

ICMP消息主要有12种， ICMP通常由某个**监测到IP分组中错误的站点产生**。**主要ICMP消息：**

- **目的不可达**：当目的路由器无法找到目的主机，或者**DF**为“**1**”的分组经过**MTU**较小的网络而无法传输时，返回该消息。
- **超时**：当分组的生命期减至**0**时，则自动丢弃该分组，并返回该消息。
- **参数出错**：当**IP**头部参数出现非法值时，返回该消息，说明发送主机或路由器软件有问题。

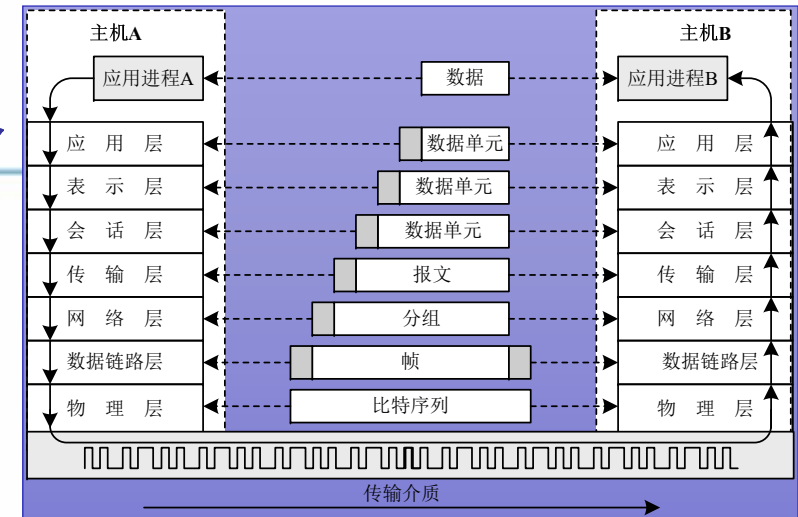


- **源抑制**：源主机收到该消息后，则降低发送速度，以减轻网络拥塞。事实上，该消息几乎并未使用，因为当网络拥塞时，再发送“源抑制”分组等于火上浇油。**Internet**的拥塞控制主要由传输层的**TCP**完成。
- **重导向**：当某路由器发现某分组的路由选择有问题时，向源主机返回该消息。
- **回声请求与应答**：这两个消息用于测试目的地址是否可达和工作正常。
- **时间戳请求与应答**：用于记录探测分组的发出时间和到达时间，通过计算时间差，得到分组的当前延迟。

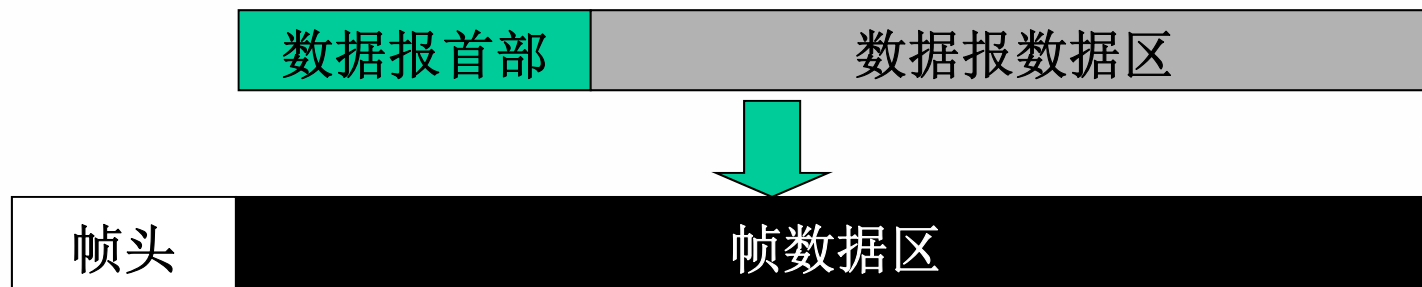


ARP协议

5.5.4 ARP协议

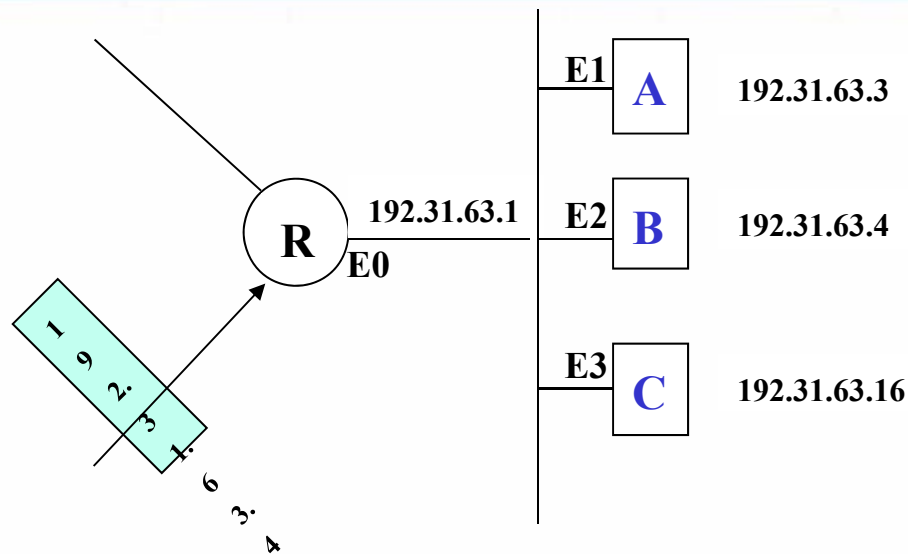


- IP地址是网络层地址，链路层无法识别。当知道目的**IP**地址，目的路由器必须获得对应的目的MAC地址，才能构造MAC帧发送到目的主机。
- ARP协议的功能： **将IP地址转化为相应的MAC地址。**





ARP协议工作原理



➤ R收到目的地址192.31.63.4的分组，知道要从E0端口转发。

➤ 广播询问帧：

.....	全“1”	E0	192.31.63.4
-------	------	----	-------	-------------

询问谁的IP地址是192.31.63.4？

➤ B返回应答帧：

.....	E0	E2	192.31.63.4
-------	----	----	-------	-------------

告诉R，我的IP地址是192.31.63.4，MAC地址是E2。

➤ R收到B的应答帧，知道目的MAC地址为E2，于是为要转发的分组

192.31.63.4

构造数据帧，并从E0端口转发。



总结：广播请求，单播回应！

提高ARP的效率

- ①设置**ARP cache**，存放最近解析出来的IP/MAC对。
- ②请求解析时，**把自己的IP/MAC地址也放在报文中**。
- ③收到ARP请求的所有主机都缓存其中的IP/MAC。
- ④主机入网时，**主动广播**它的IP/MAC。



ARP中软状态使用（缓存信息失效的问题）

问题：一条IP/MAC映射记录失效怎么办？

解决：使用**超时机制**（20分钟），如果在超时到来前都没有使用某条记录，则删除该记录。

优势：

- （1）无需专门的通知机制；
- （2）对硬件可靠性要求低。

缺陷：可能会‘反应迟钝’。



IP子网

5.5.5 IP子网

● 子网

- 将一个网络（A，B或C类）内部分割成若干部分，每个部分称为一个子网。
- 子网对外部不可视。



● 子网划分

● 为什么进行子网划分？

- 充分利用IP地址
- 便于网络管理

IP地址不足的原因：

主机号浪费严重，而网络号又严重缺乏

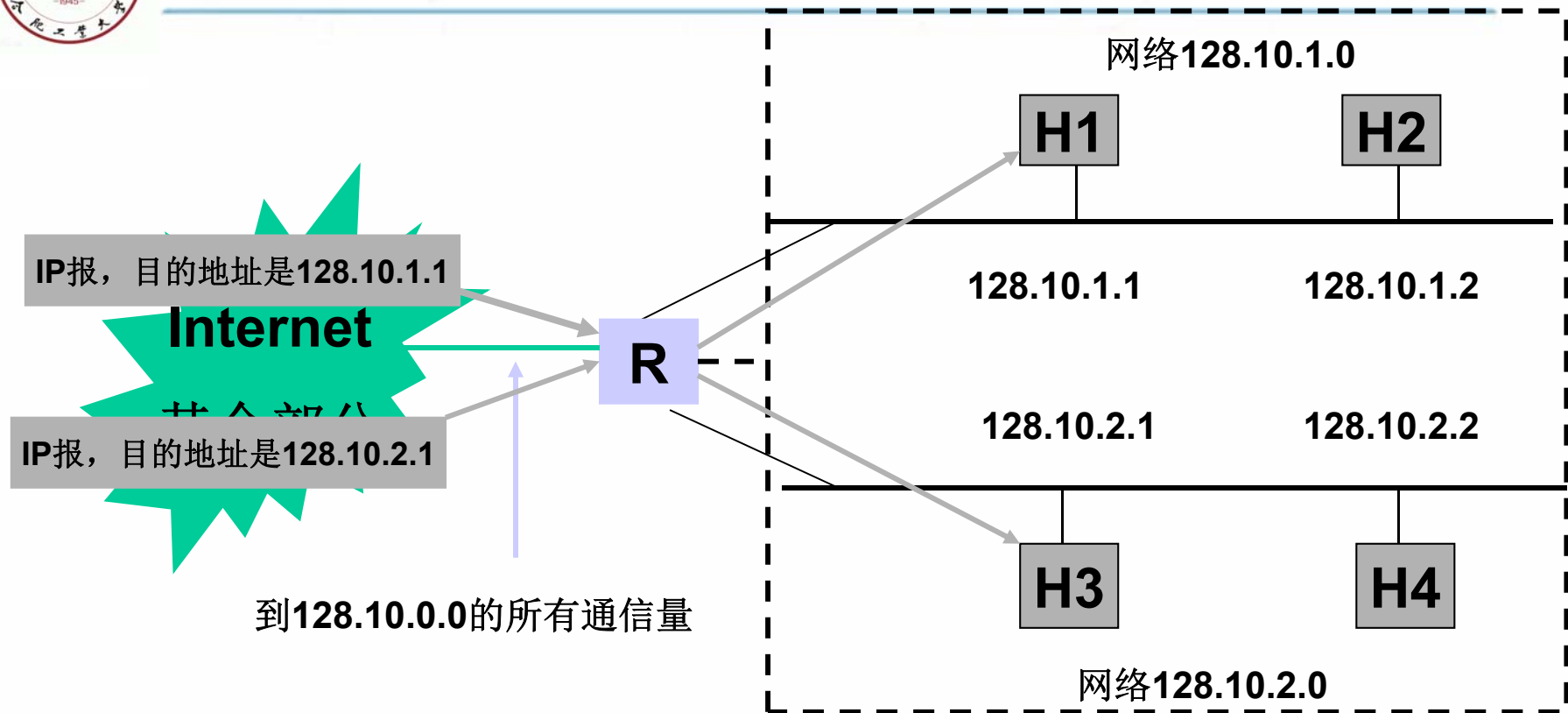
解决思路：

从IP地址的主机部分“借”位，并把它们用在网络部分

IP地址主机号进一步划分为：

子网号 + 主机号





可以让一个网络号跨越多个物理网络

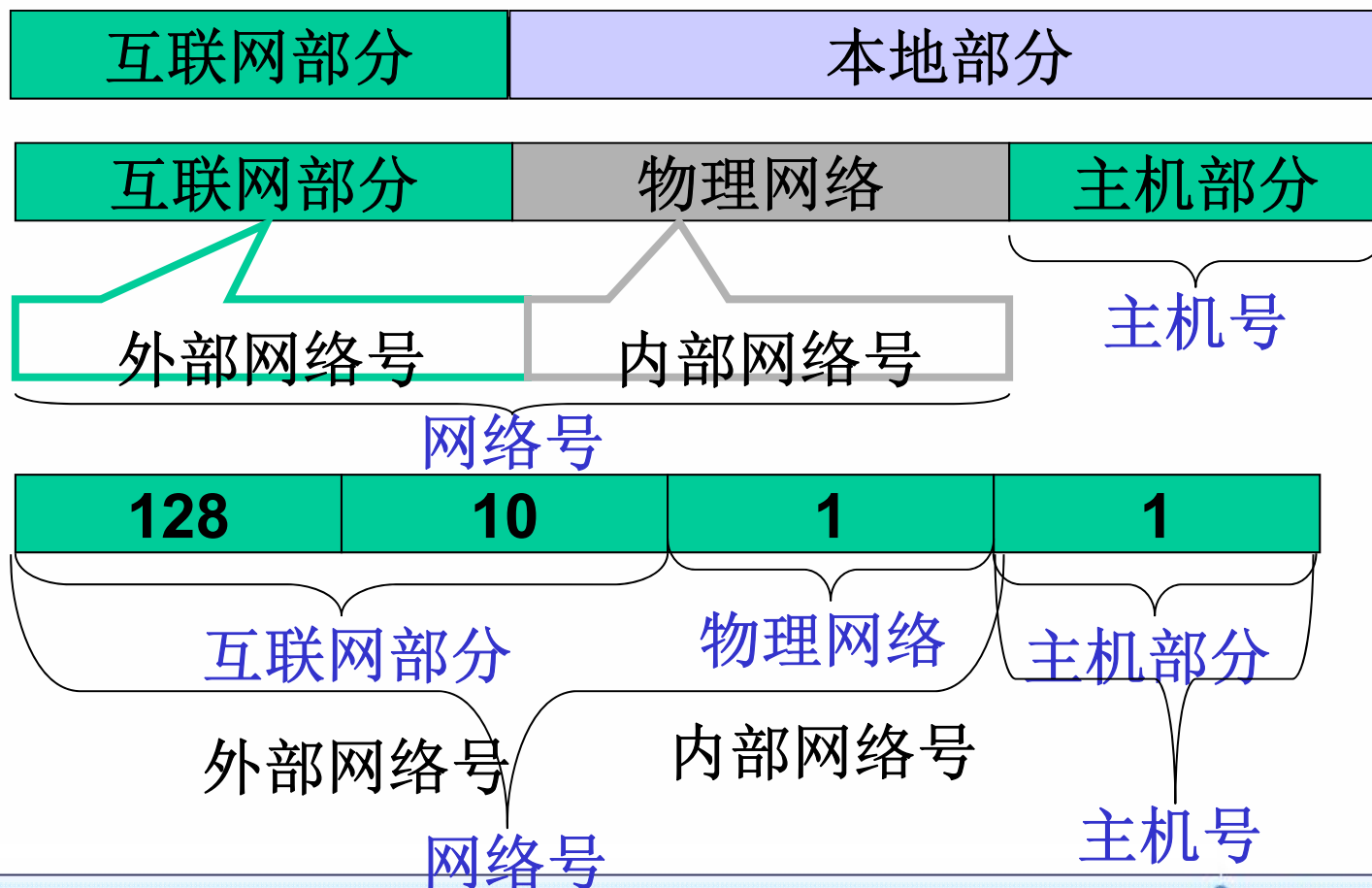
虽然128.10.0.0这个网络号中可以包含多个物理网络，但是互联网的其余部分都只能看到128.10.0.0这个“大网”



子网编址模式和子网掩码

引入子网划分后，IP地址格式有了新的含义：

网络号 + 主机号 \rightarrow 互联网部分 + 本地部分





子网掩码

- 子网掩码

- IP地址由两部分组成：网络号+主机号。
- 在不进行子网划分的情况下，每个网络（A、B或C类）都有一个缺省网络识别方式，分别是：

A类：255.0.0.0

B类：255.255.0.0

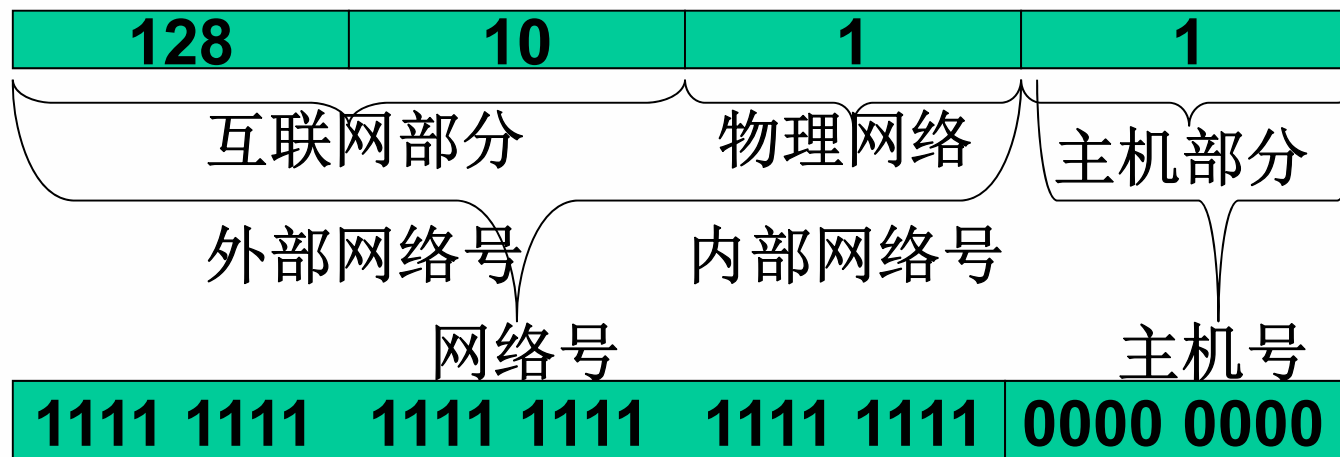
C类：255.255.255.0



问题：如何指示网络部分？

解决：使用**子网掩码**。每一个使用子网的网点都选择一个**32位**的位模式。若位模式中的某位为‘1’，则对应**IP**地址中的某位为网络地址中的一位，否则为主机号中的一位。

子网掩码：32位，1：网络； 0：主机



网络号**128.10.1.0**

主机号**1**

掩码**255.255.255.0**



网络地址转换

(NAT: Network Address Translation)

作用

- 1、解决**IP**地址不足，
多台内部机器共享一个外部**IP**地址
- 2、出于安全考虑，隐藏内部**IP**地址



不予分配的IP地址（保留地址，RFC1918）

•前缀	•最低地址	•最高地址	
•10/8	•10.0.0.0	•10.255.255.255	← 1个A类网号
•172.16/12	•172.16.0.0	•172.31.255.255	← 16个B类网号
•192.168/16	•192.168.0.0	•192.168.255.255	← 256个C类网号
•169.254/16	•169.254.0.0	•169.254.255.255	← 1个B类网号

说明：

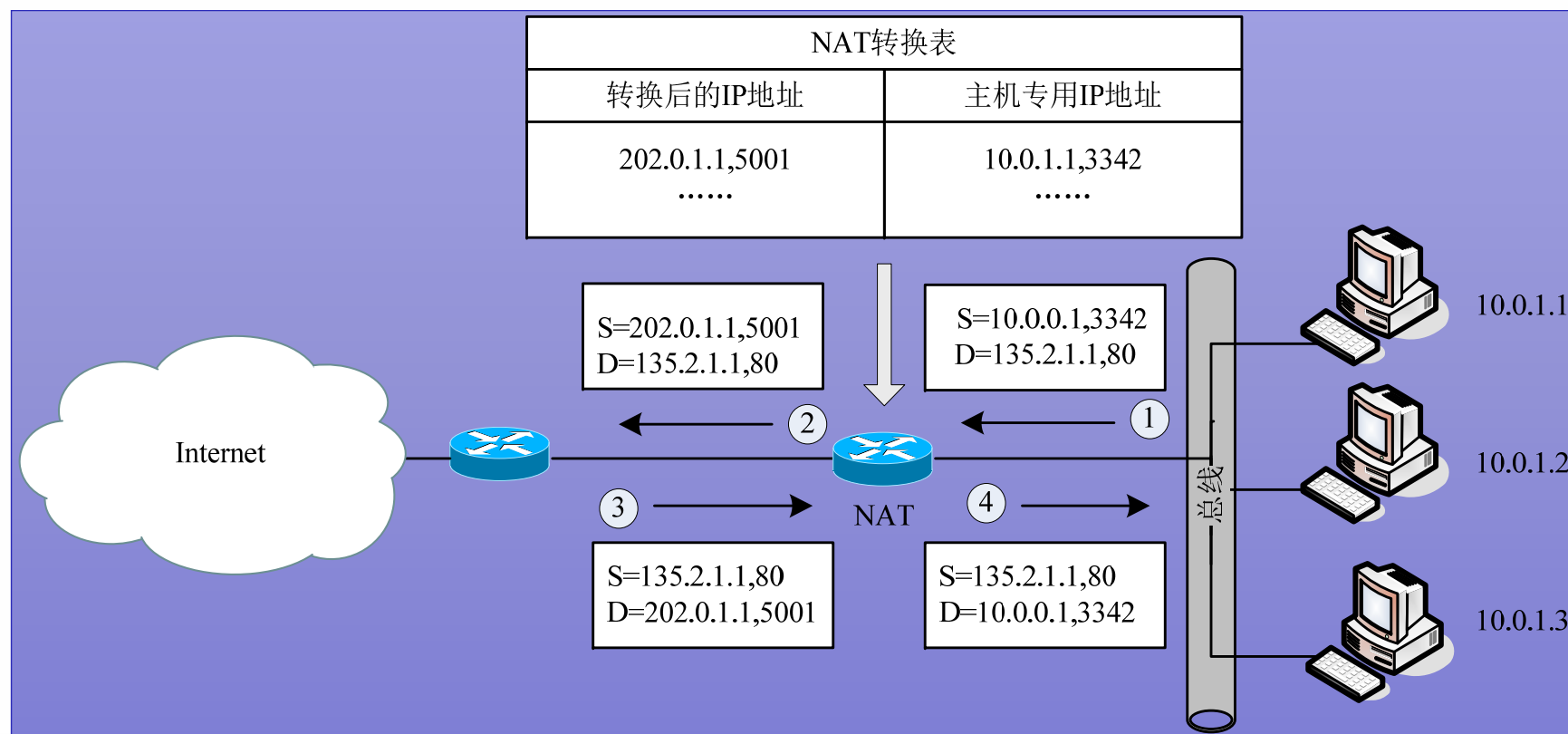
1. 使用保留地址可加强网络安全
2. 使用保留地址的设备要接入Internet，必须使用某种地址转换技术



NAT

NAT工作原理

修改分组源地址（出去的分组）或目的地址（进入的分组）





DHCP

- **Dynamic Host Configuration Protocol**
- **DHCP的前身是BOOTP（引导程序协议）**
- **为什么要DHCP**
 - 网络的用户不固定，反复设置IP信息，太麻烦
 - 网络太大，没法管理IP分配



DHCP示例(1/2)

DHCP配置

本页面可配置DHCP服务器，您可以在这里设置是否启用DHCP服务器及可用的IP地址池。
当本地网络中的主机要求Internet访问的时候本设备可以向它们分配该地址池中的IP地址。
本页面也可以配置DHCP代理，您可以在这里设置DHCP服务器的IP地址。

注意:如果您做完 DHCP 设定，想长期生效，请“[保存/重启!](#)”

本地IP地址: 192.168.1.1 子网掩码: 255.255.255.0

☐ 禁止 ☐ DHCP代理 ☒ DHCP服务器

DHCP服务器设置:

IP地址池: 192.168.1. - 192.168.1.

租约时间: 天 时 分 (均为-1 表示租期不限)

域名:

网关地址:



DHCP示例 (2/2)

静态IP分配列表 - Microsoft Internet Explorer

静态IP分配列表

该页面用来配置基于硬件地址的静态IP地址。您可以分配/删除静态IP。 在主机硬件地址栏，请输入十六进制的字符串。例如“00-d0-59-c6-12-43”。 在分配的IP地址栏，请输入阿拉伯数字的字符串。例如“192.168.1.100”。

主机MAC地址:

预留静态IP地址:

静态IP表:

索引	主机MAC地址	分配的IP地址
<input type="radio"/>	00-14-78-0D-93-CF	192.168.1.10



DHCP

- **DHCP可以简化网络管理工作**

- 动态配置管理主机的地址信息
- 作用是在**TCP/IP**网络中向主机提供配置信息

- **采用Client/Server模式**

- 由**DHCP Client**向**DHCP Server**提出配置申请，**DHCP Server**根据策略返回相应配置信息。



DHCP

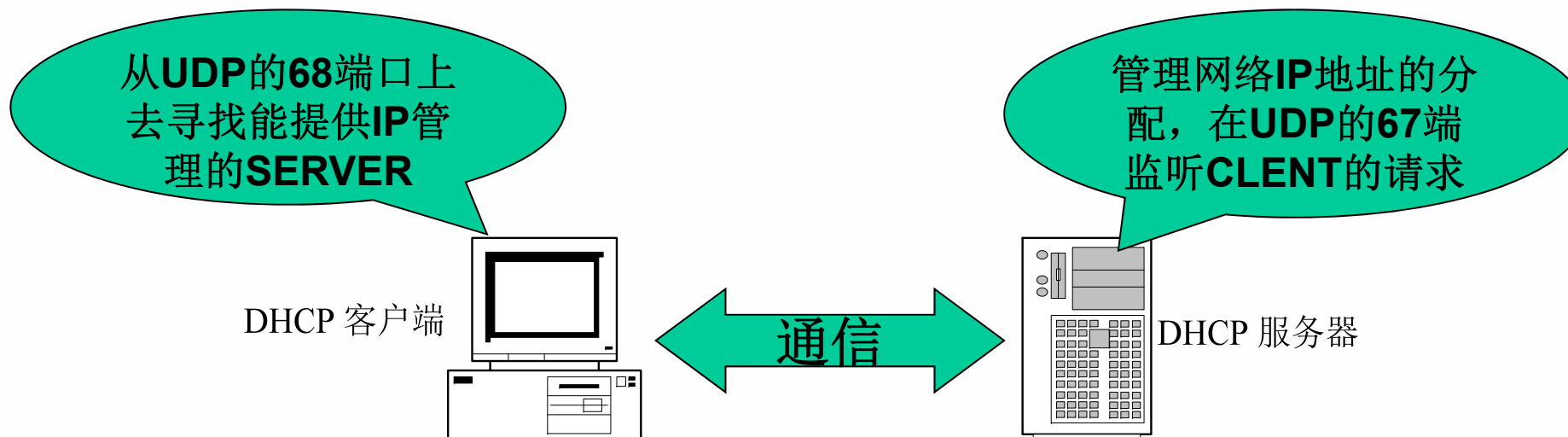
● DHCP协议特点

- 整个配置过程自动实现，**DHCP Client**端无需配置
- 所有配置信息由**DHCP Server**统一管理
- 通过**IP**地址租期管理，提高**IP**地址的使用效率
- 采用广播实现报文交互，报文一般不能跨网段，如果需要跨网段，需要使用**DHCP Relay**技术实现



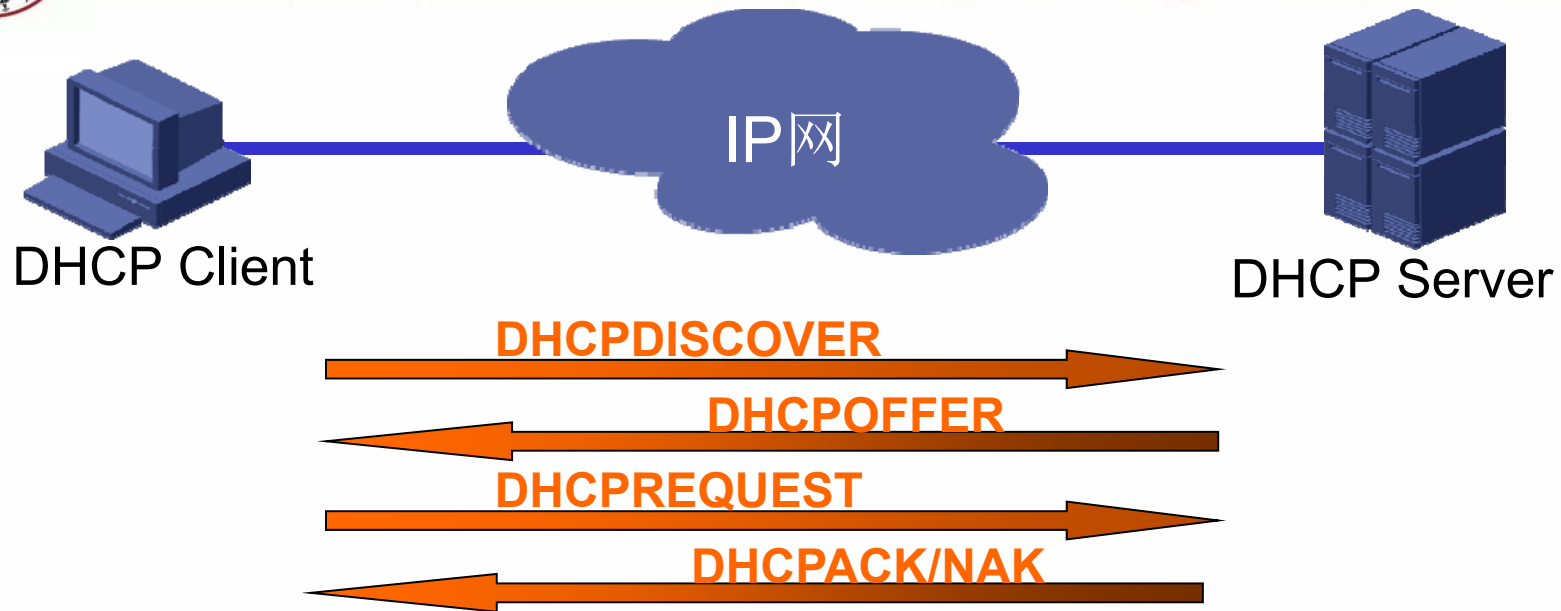
DHCP

工作模型：**C/S**结构
使用的端口：**67 / 68**
使用的传输层协议：**UDP**





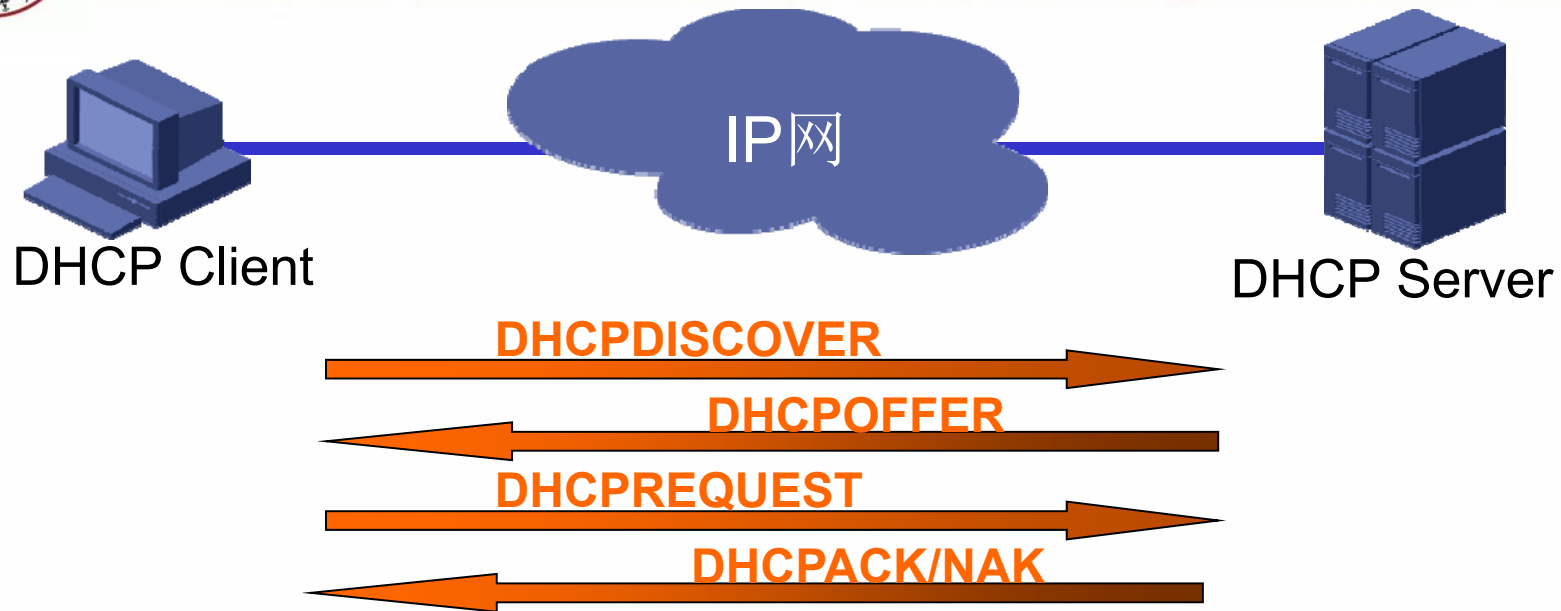
DHCP



1、发现阶段，即**DHCP**客户机寻找**DHCP**服务器的阶段。因为**DHCP**服务器的IP地址对于客户机来说是未知的，所以**DHCP**客户机以广播方式发送**DHCP DISCOVER**发现信息来寻找**DHCP**服务器，即向地址**255.255.255.255**发送特定的广播信息。网络上每一台安装了**TCP/IP**协议的主机都会接收到这种广播信息，但只有**DHCP**服务器才会做出响应。



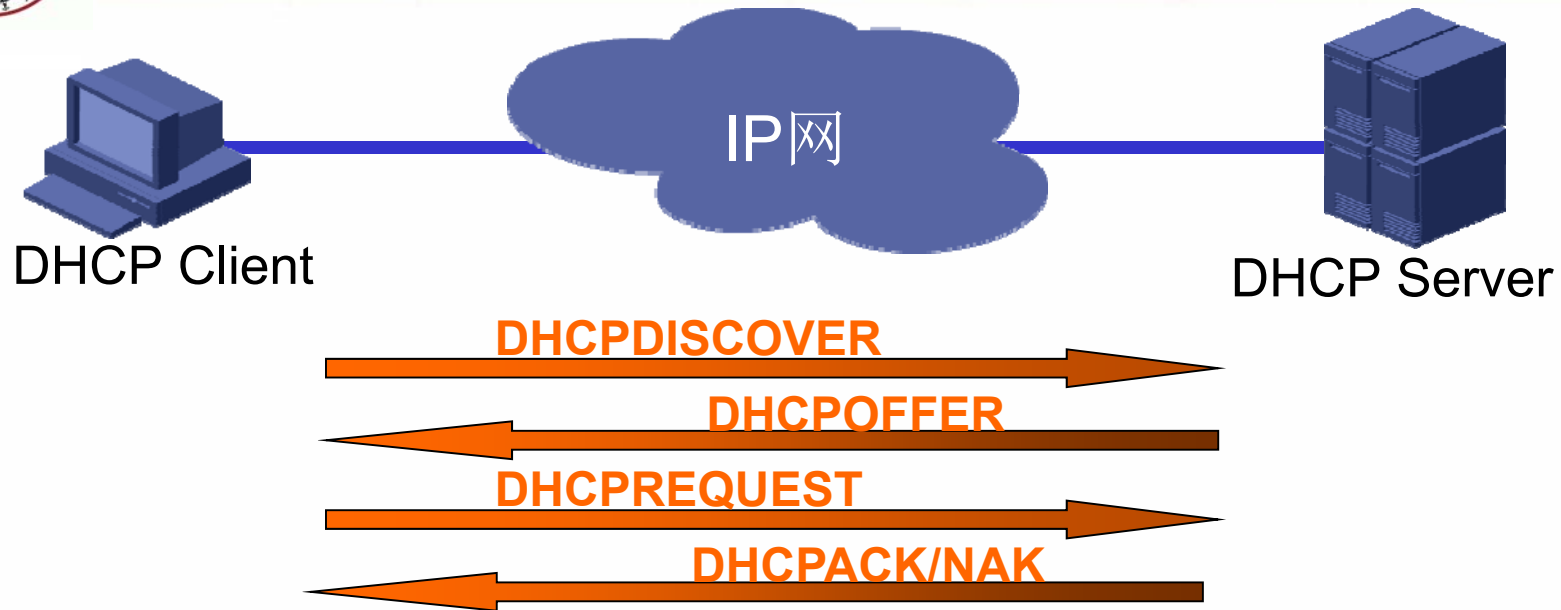
DHCP



2、提供阶段，即**DHCP**服务器提供**IP**地址的阶段。在网络中接收到**DHCP DISCOVER**发现信息的**DHCP**服务器都会做出响应，它从尚未出租的**IP**地址中挑选一个分配给**DHCP**客户机，向**DHCP**客户机发送一个包含出租的**IP**地址和其他设置的**DHCP OFFER**提供信息。



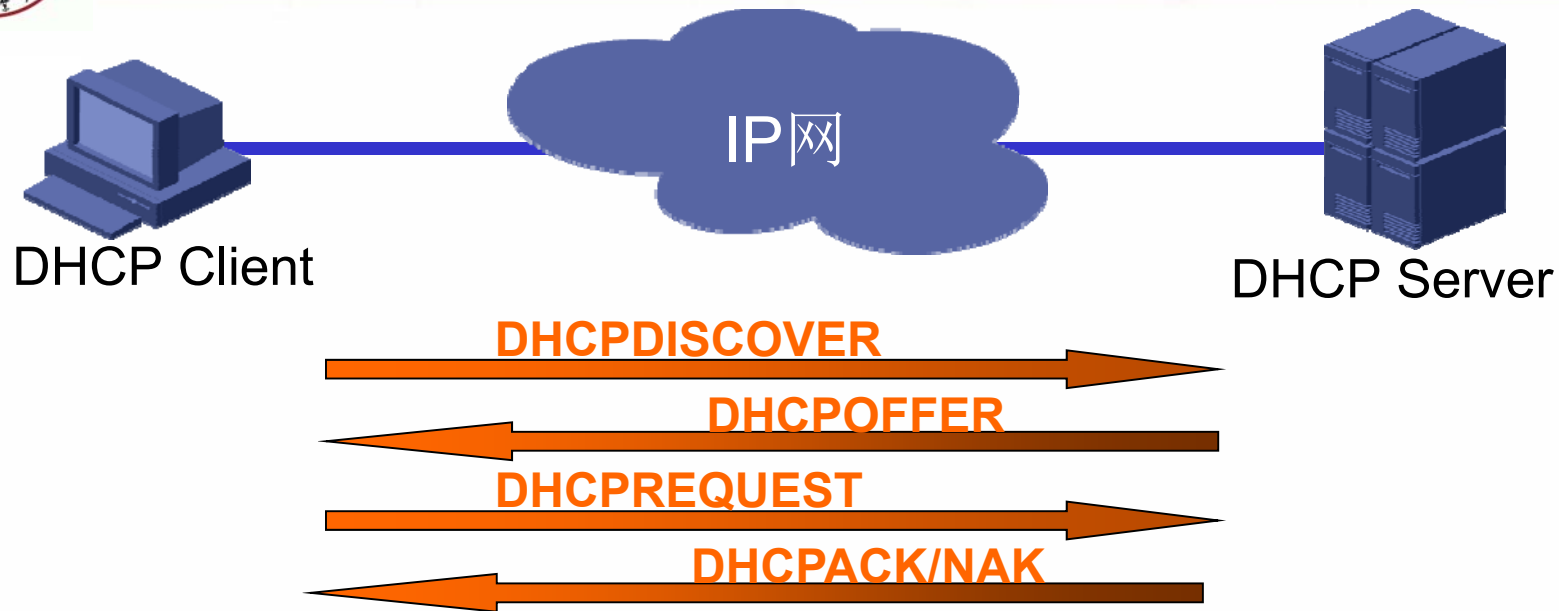
DHCP



3、选择阶段，即**DHCP**客户机选择某台**DHCP**服务器提供的**IP**地址的阶段。如果有多台**DHCP** 服务器向**DHCP**客户机发来的**DHCP OFFER**提供信息，则**DHCP**客户机只接收第一个收到的 **DHCP OFFER**提供信息，然后它就**以广播方式回答一个DHCP REQUEST请求信息**，该信息中包含向它所选定的**DHCP**服务器请求**IP**地址的内容。之所以要以广播方式回答，是为了通知所有的 **DHCP**服务器，他将选择某台**DHCP**服务器所提供的**IP**地址。



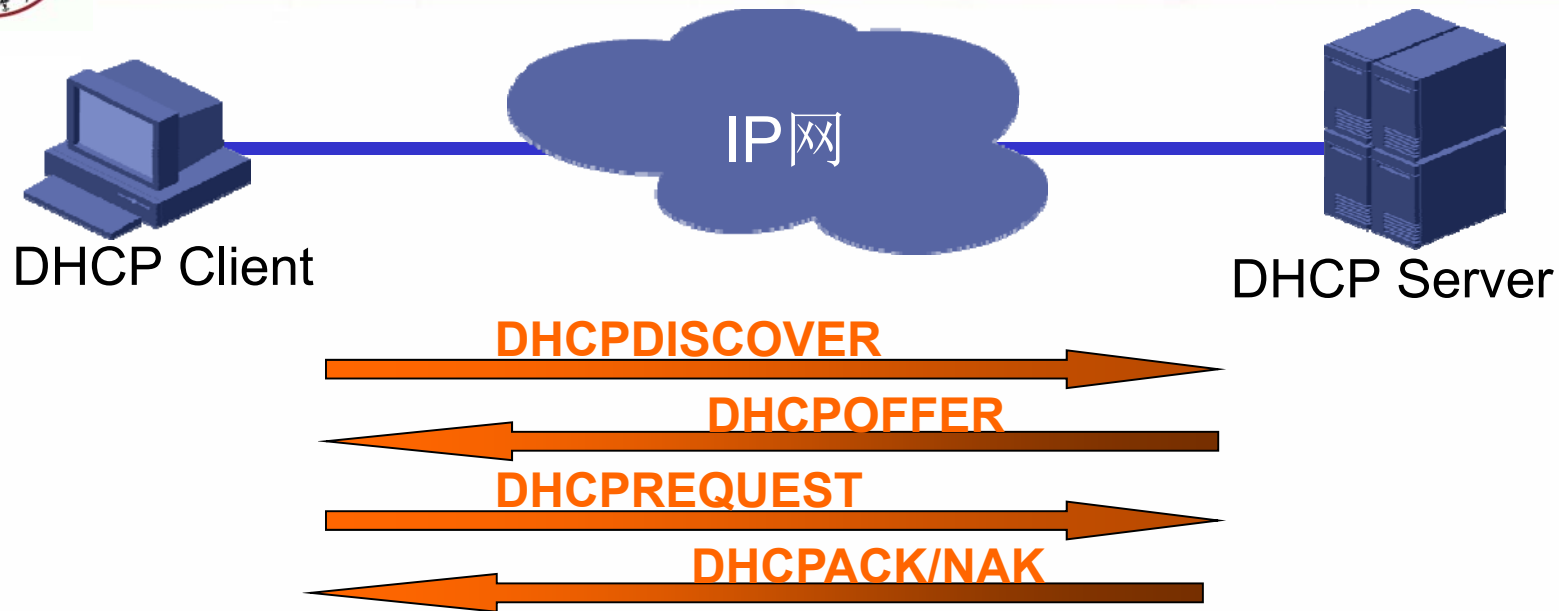
DHCP



4、确认阶段，即**DHCP**服务器确认所提供的**IP**地址的阶段。当**DHCP**服务器收到**DHCP**客户机回答的**DHCP REQUEST**请求信息之后，它便向**DHCP**客户机发送一个包含它所提供的**IP**地址和其他设置的**DHCP ACK**确认信息，告诉**DHCP**客户机可以使用它所提供的**IP**地址。然后**DHCP**客户机便将获取到的**IP**地址与网卡绑定，另外，除**DHCP**客户机选中的服务器外，其他的**DHCP**服务器都将收回曾提供的**IP**地址。



DHCP

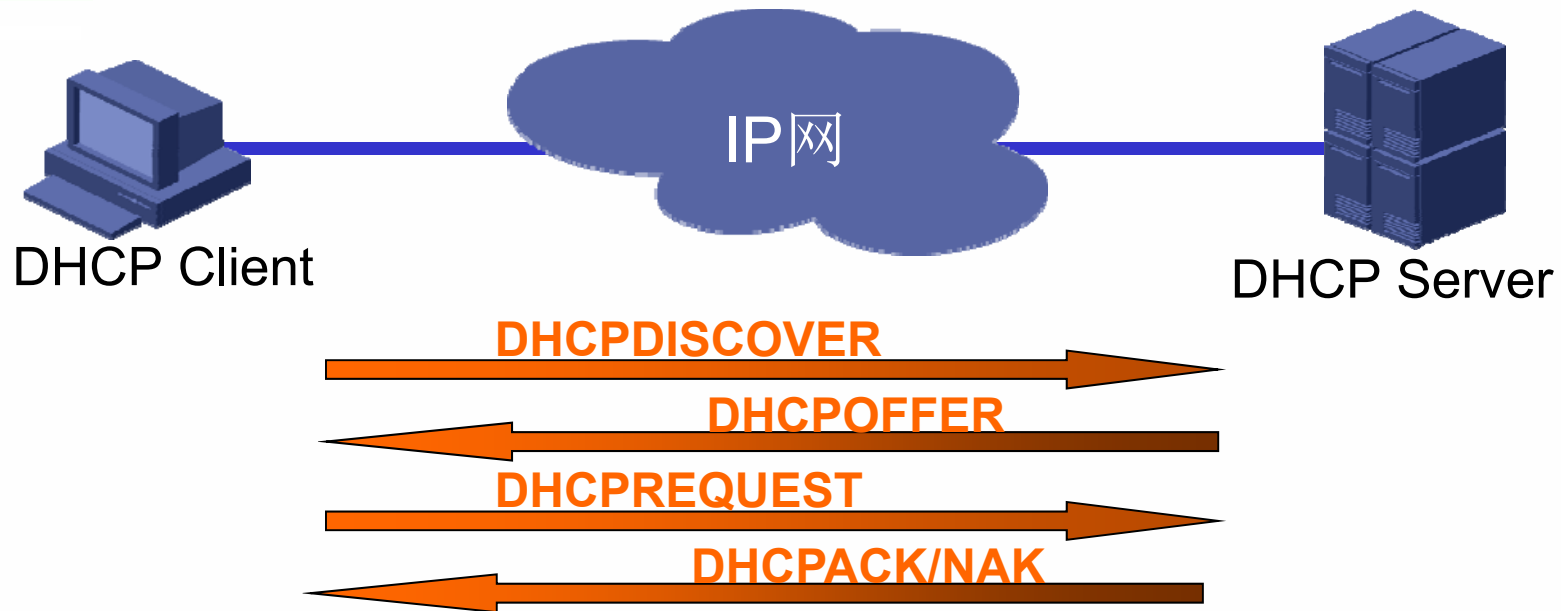


5、**重新登录**。以后**DHCP**客户机每次重新登录网络时，就不需要再发送**DHCP DISCOVER** 发现信息了，而是**直接发送包含前一次所分配的IP地址的DHCP REQUEST请求信息**。当**DHCP**服务器收到这一信息后，它会尝试让**DHCP**客户机继续使用原来的**IP**地址，并回答一个 **DHCP ACK**确认信息。如果此**IP**地址已无法再分配给原来的**DHCP**客户机使用时（比如此**IP**地址已分配给其它**DHCP**客户机使用），则**DHCP**服务器给**DHCP**客户机回答一个**DHCP NAK**否认信息。当原来的**DHCP**客户机收到此**DHCP NAK**否认信息后，它就必须重新发送**DHCP DISCOVER**发现信息来请求新的**IP**地址。

87



DHCP



6、更新租约。 DHCP服务器向DHCP客户机出租的IP地址一般都有一个租借期限，期满后 DHCP服务器便会收回出租的IP地址。如果DHCP客户机要延长其IP租约，则必须更新其IP租约。 **DHCP**客户机启动时和**IP**租约期限过一半时，**DHCP**客户机都会自动向**DHCP**服务器发送更新其**IP** 租约的信息。



DHCP协议分配地址的优先级

- **DHCP Server**数据库中与该**DHCP Client**的**MAC**地址静态绑定的**IP**地址
- 该**DHCP Client**曾经使用过的地址
- 顺序查找**DHCP**地址池中可供分配的**IP**地址，最先找到的可用**IP**地址，优先级高
- 如果未找到可用的**IP**地址，则依次查询超过租期、发生冲突的**IP**地址
- 找不到则报告错误



本章小结

● 内容

主要介绍报文分组在源/目的节点之间传输过程中所涉及的相关技术，包括分组交换技术、数据报和虚电路、路由选择、拥塞控制，以及网络层协议实例IP协议、子网等。

● 重点

- 数据报和虚电路
- 路由选择算法
- IP协议
- 子网