高数帮

课时4 网络层

课时4 网络层 考点分析

考点	重要程度	占分	题型
功能	★		选择题
路由算法	★★		选择题
IPv4	★★★★		选择题、问答题
IPv6	★★		选择题
移动IP	★		选择题
IP组播	★		选择题
路由协议	★★★★		选择题、问答题

网络层的功能

异构网络互联

TCP/IP 体系在网络互联上采用的做法是在网络层（即 IP 层）采用标准化协议，但相互连接的网络可以是异构的图 a 表示用许多计算机网络通过一些路由器进行互联。由于参加互联的计算机网络都使用相同的网际协议，因此可以把互联后的计算机网络视为如图 b 所示的一个虚拟IP网络

(a) 实际互联网络

(b) 虚拟 IP 网络

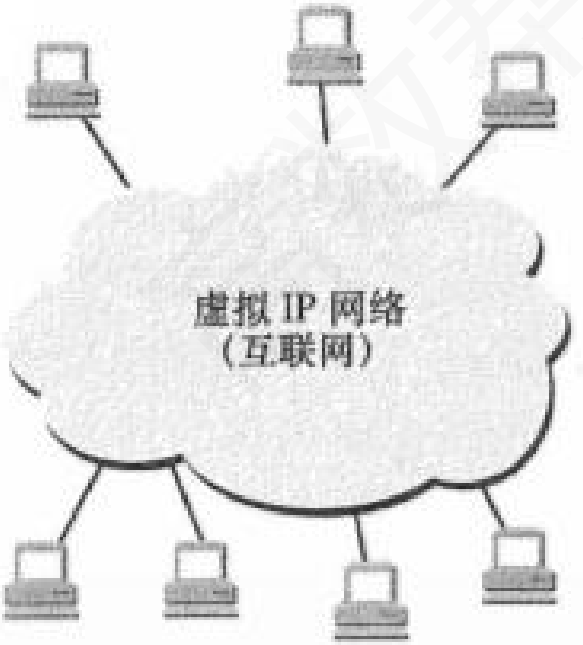
网络层的功能

路由与转发

路由器主要完成两个功能：一是路由选择（确定哪一条路径），二是分组转发（当一个分组到达时所采取的动作）

前者是根据特定的路由选择协议构造出路由表，同时经常或定期地和相邻路由器交换路由信息而不断地更新和维护路由表

(a) 实际互联网络



(b) 虚拟 IP 网络

后者处理通过路由器的数据流，关键操作是转发表查询、转发及相关的队列管理和任务调度等

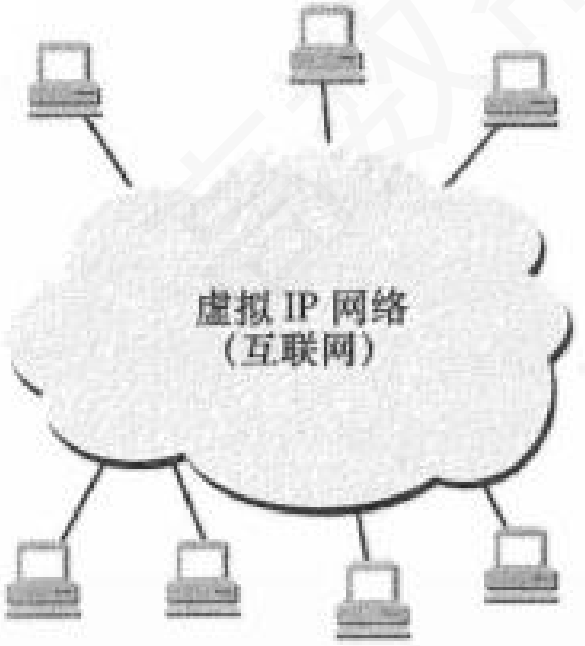
网络层的功能

拥塞控制

拥塞控制的作用是确保子网能够承载所达到的流量，这是一个全局性的过程，涉及各方面的行为

视频讲解更清晰

(a) 实际互联网络



(b) 虚拟 IP 网络

4.2 路由算法

静态路由与动态路由

静态路由算法（又称非自适应路由算法），指由网络管理员手工配置的路由信息，当网络的拓扑结构或链路的状态发生变化时，网络管理员需要手工去修改路由表中相关的静态路由信息

它不能及时适应网络状态的变化，对于简单的小型网络，可以采用静态路由

动态路由算法（又称自适应路由算法）。指路由器上的路由表项是通过相互连接的路由器之间彼此交换信息，然后按照一定的算法优化出来的

这些路由信息会在一定时间间隙里不断更新，以适应不断变化的网络，随时获得最优的寻路效果

4.2 路由算法

距离-向量路由算法

所有结点都定期地将它们的整个路由选择表传送给所有与之直接相邻的结点

这种路由选择表包含：每条路径的目的地（另一结点）；路径的代价（也称距离）

在这种算法中，所有结点都必须参与距离向量交换，以保证路由的有效性和一致性，也就是说，所有的结点都监听从其他结点传来的路由选择更新信息，并在下列情况下更新它们的路由选择表：

1、被通告一条新的路由，该路由在本结点的路由表中不存在，此时本地系统加入这条新的路由

2、发来的路由信息中有一条到达某个目的地的路由，该路由与当前使用的路由相比，有较短的距离

4.2 路由算法

链路状态路由算法

链路状态路由算法要求每个参与该算法的结点都具有完全的网络拓扑信息，它们执行下述两项任务

第一，主动测试所有邻接结点的状态。两个共享一条链接的结点是相邻结点，它们连接到同一条链路，或者连接到同一广播型物理网络

第二，定期地将链路状态传播给所有其他结点（或称路由结点）

视频讲解更清晰

4.2 路由算法

链路状态路由算法主要有三个特征：

1、向本自治系统中所有路由器发送信息，路由器通过所有端口向所有相邻的路由器发送信息。而每个相邻路由器又将此信息发往其所有相邻路由器（但不再发送给刚刚发来信息的那个路由器）

2、发送的信息是与路由器相邻的所有路由器的链路状态，但这只是路由器所知道的部分信息

所谓"链路状态"，是指说明本路由器与哪些路由器相邻及该链路的"度量"

3、只有当链路状态发生变化时，路由器才向所有路由器发送此信息

4.2 路由算法

层次路由

特网将整个互联网划分为许多较小的自治系统（注意一个自治系统中包含很多局域网），每个自治系统有权自主地决定本系统内应采用何种路由选择协议。如果两个自治系统需要通信， 那么就需要一种在两个自治系统之间的协议来屏蔽这些差异

因特网把路由选择协议划分为两大类：

一个自治系统内部所使用的路由选择协议称为内部网关协议（IGP），也称域内路由选择

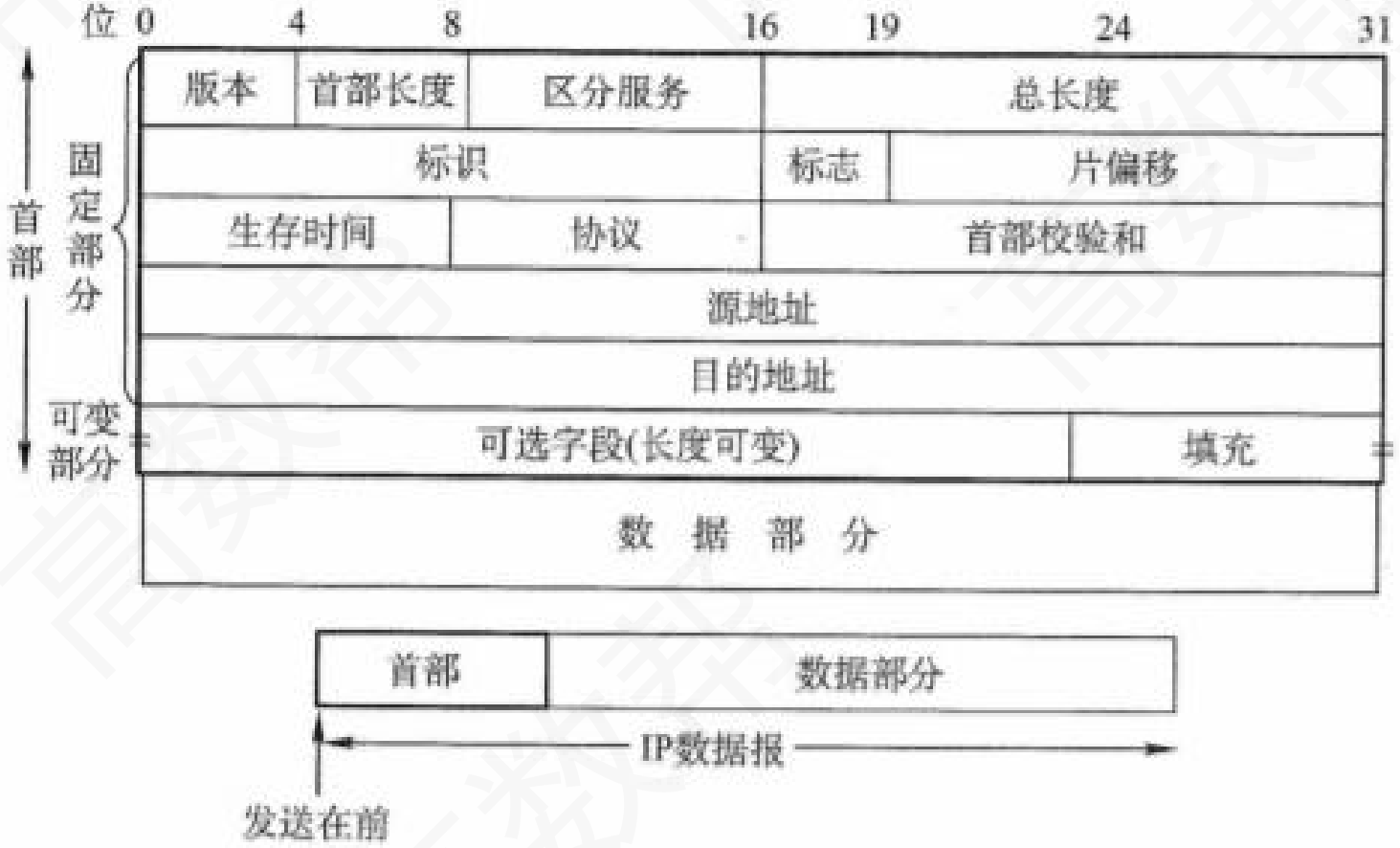
具体的协议有 RIP 和 OSPF 等

自治系统之间所使用的路由选择协议称为外部网关协议（EGP），也称城间路由选择，用在不同自治系统的路由器之间交换路由信息，并负责为分组在不同自治系统之间选择最优的路径

具体的协议有 BGP

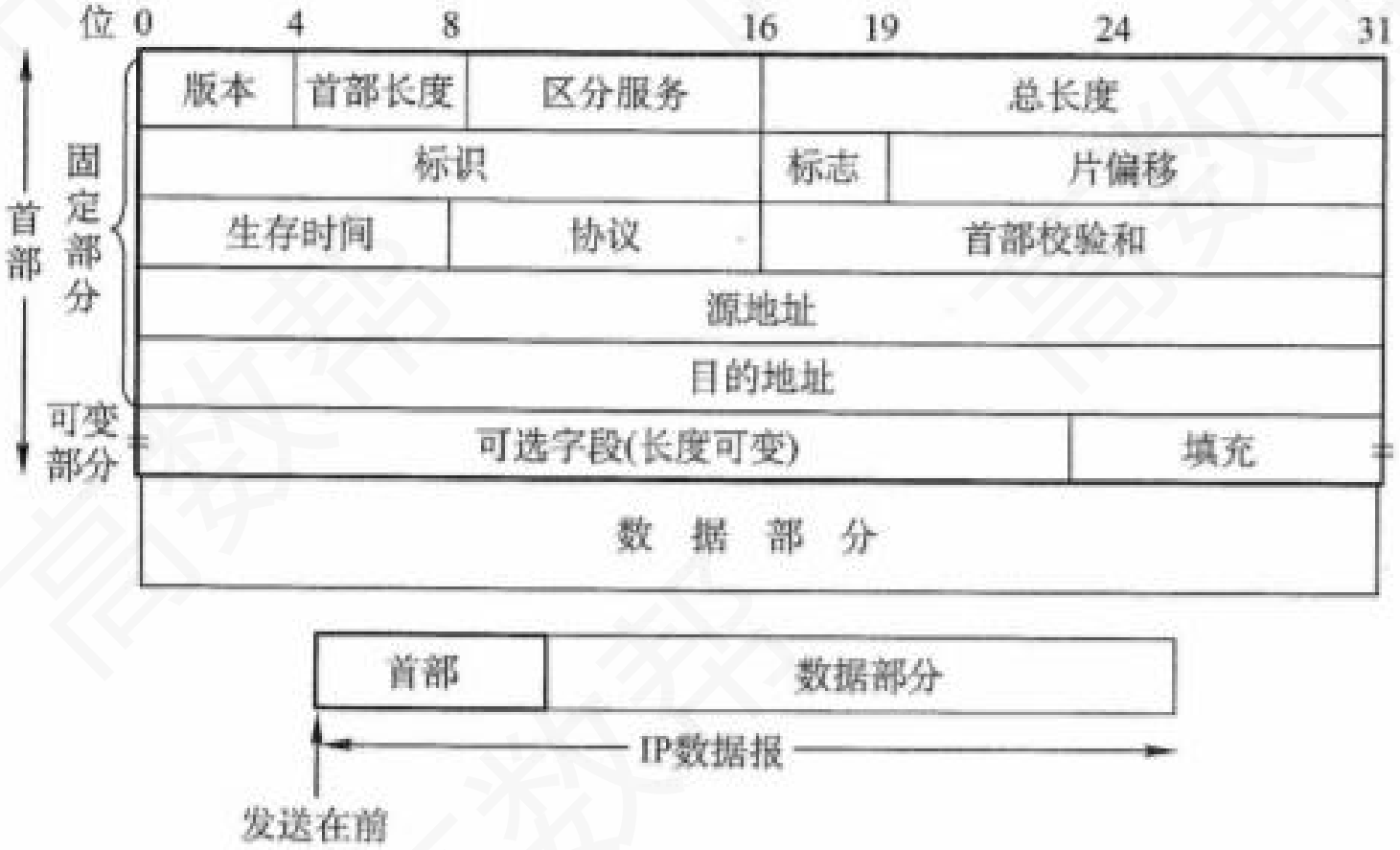
IPv4分组的格式

首部长度。占4位，可以表示的最大十进制数是 15。以32位为单位，最大值为 60B（15×4B）。最常用的首部长度是20B，此时不使用任何选项（即可选字段）



IPv4分组的格式

总长度。占16 位。指首部和数据之和的长度，单位为字节，因此数据报的最大长度为65535B。以太网帧的最大传送单元（MTU）为1500B，因此当一个IP数据报封装成帧时，数据报的总长度（首部加数据）一定不能超过下面的数据链路层的MTU 值



IPv4分组的格式

标识。每产生一个数据报就加1，当一个数据报的长度超过网络的MTU时，必须分片，此时每个数据报片都复制一次标识号，以便能正确重装成原来的数据报

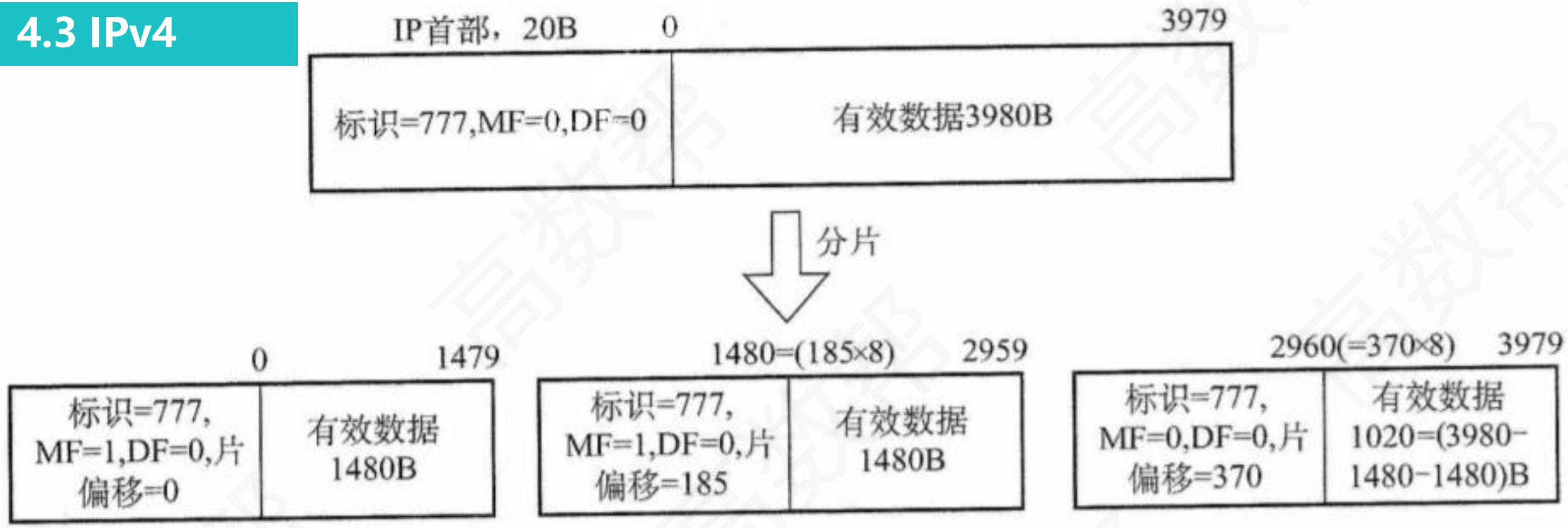
标志。占3位。标志字段的最低位为MF，MF=1表示后面还有分片；标志字段中间的一位是DF，只有当DF=0时才允许分片偏移。

分片偏移。它指出较长的分组在分片后，某片在原分组中的相对位置。片偏移以8 个字节为偏移单位，即每个分片的长度一定是 8B的整数倍

IP 数据报分片

一个数据链路层数据报能承载的最大数据量称为最大传送单元（MTU）。因为IP数据报被封装在数据链路层数据报中，因此数据链路层的 MTU 严格地限制着 IP 数据报的长度，而且在 IP数据报的源与目的地路径上的各段链路可能使用不同的数据链路层协议，有不同的 MTU

IP分片涉及一定的计算。例如，一个长 4000B 的IP数据报（首部 20B，数据部分3980B）到达一个路由器，需要转发到一条 MTU为1500B的链路上



这意味着原始数据报中的3980B数据必须被分配到3个独立的片中（每片也是一个IP 数据报）。假定原始数据报的标识号为77，如图所示。可以看出，由于偏移值的单位是8B，所以除最后一个片外，其他所有片中的有效数据载荷都是8的倍数

IP 数据报分片

网络层转发分组的流程

1、从数据报的首部提取目的主机的IP地址D，得出目的网络地址N

2、若网络N与此路由器直接相连，则把数据报直接交付给目的主机 D，这称为路由器的直接交付；否则是间接交付，执行步骤 3

3、若路由表中有目的地址为D的特定主机路由（对特定的目的主机指明一个特定的路由，通常是为了控制或测试网络，或出于安全考虑才采用的），则把数据报传送给路由表中所指明的下一跳路由器；否则，执行步骤 4。

IP 数据报分片

网络层转发分组的流程

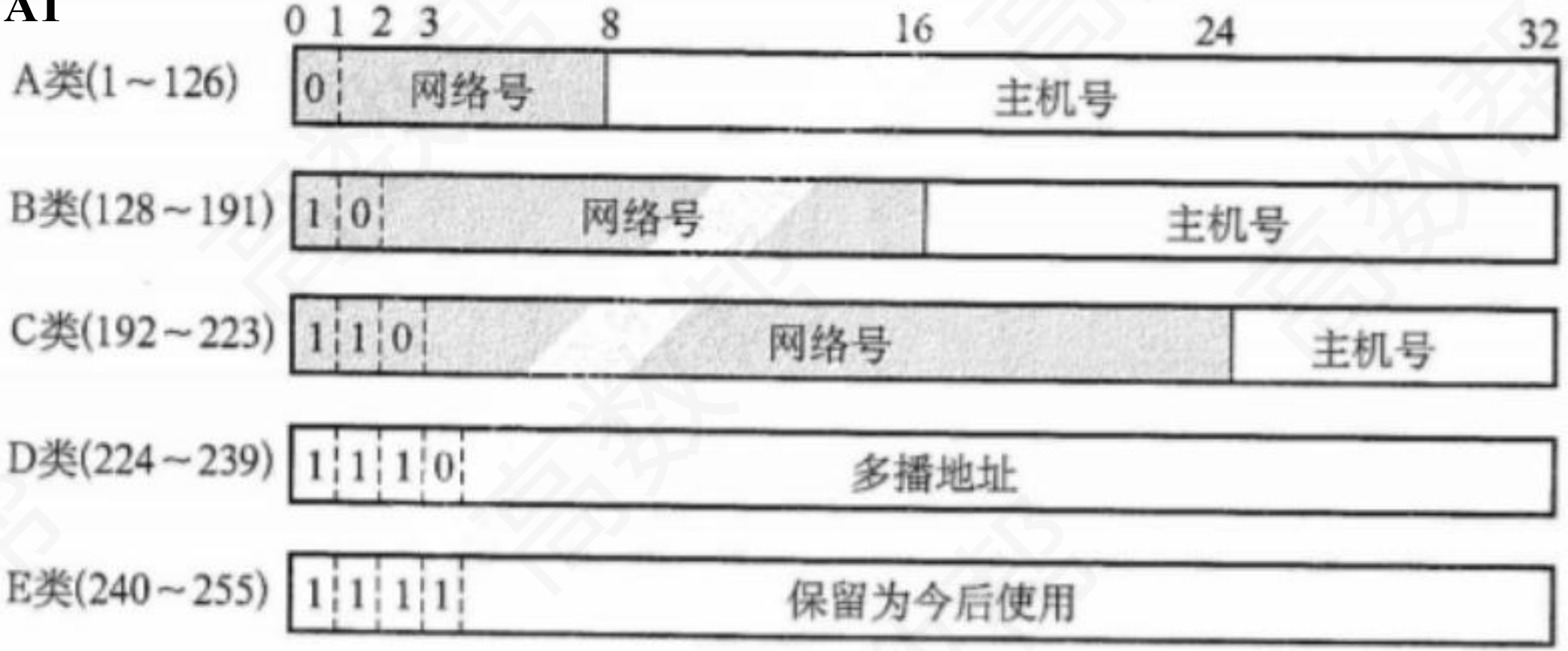
4、若路由表中有到达网络N的路由，则把数据报传送给路由表指明的下一跳路由器：否则，执行步骤 5

5、路由表中有一个默认路由，则把数据报传送给路由表中所指明的默认路由器;否则， 执行步骤6

6、报告转发分组出错

4.3 IPv4

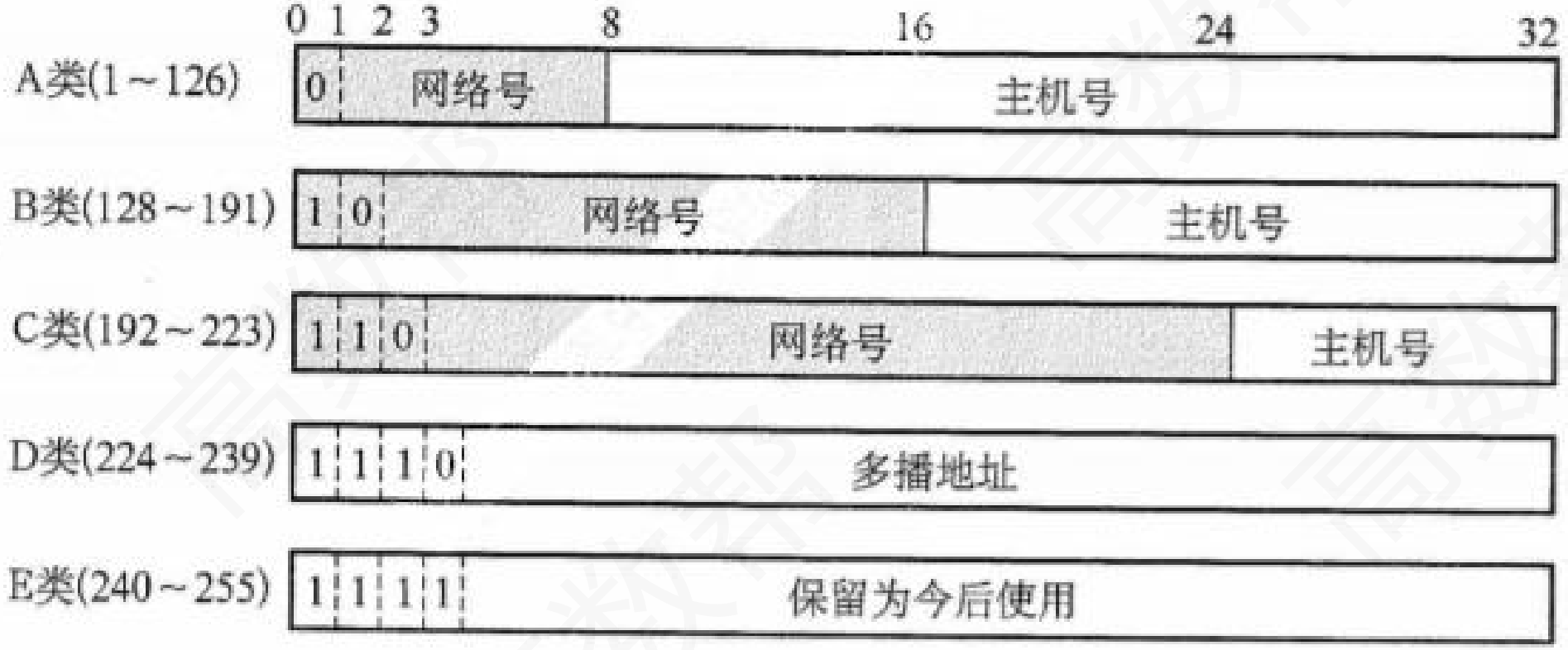
IPv4地址与 NAT



分类的IP 地址如图

主机号全为0表示本网络本身；

主机号全为1表示本网络的广播地址，又称直接广播地址；



127.0.0.0保留为环回自检（LoopbackTest）地址，此地址表示任意主机本身，目的地址为环回地址的 IP 数据报永远不会出现在任何网络上；

32 位全为0，即 0.0.0.0 表示本网络上的本主机；

32 位全为1，即255.255.255.255 表示整个TCP/IP网络的广播地址，又称受限广播地址。 实际使用时，由于路由器对广播域的隔离，255.55255.255 等效为本网络的广播地址

网络地址转换

网络地址转换（NAT）是指通过将专用网络地址转换为公用地址，从而对外隐藏内部管理的IP地址。

私有 IP 地址网段如下：

A类：1个A类网段，即**10**.0.0.0~**10**.255.255.255。

B类：16个B类网段，即**172.16.**0.0~**172.31**.255.255。

C类：256个C类网段，即**192.168.0**.0~**192.168.255**.255。

视频讲解更清晰

4.3 IPv4

用NAT时需要在专用网连接到因特网的路由器上安装 NAT软件，NAT路由器至少有一个有效的外部全球地址

使用本地地址的主机和外界通信时，NAT 路由器使用 NAT 转换表将本地地址转换成全球地址，或将全球地址转换成本地地址

NAT转换表中存放着{本地IP地址：端口）到{全球IP地址： 端口}的映射

通过{ip地址：端口}这样的映射方式，可让多个私有IP地址映射到同一个全球 IP地址

子网划分与子网掩码、CIDR

子网划分

子网划分的基本思路如下：

子网划分纯属一个单位内部的事情。单位对外仍然表现为没有划分子网的网络

从主机号借用若干比特作为子网号，当然主机号也就相应减少了相同的比特。

三级IP地址的结构如下：IP地址 ={<网络号>，<子网号>，<主机号>}

凡是从其他网络发送给本单位某台主机的IP数据报，仍然是根据IP数据报的目的网络号，先找到连接到本单位网络上的路由器。然后该路由器在收到 IP 数据报后，按目的网络号和子网号找到目的子网。最后把IP 数据报直接交付给目的主机

子网划分与子网掩码、CIDR

子网掩码

子网掩码是一个与IP地址相对应的、长32bit的二进制串，它由一串1和跟随的一串0组成

其中，1对应于IP地址中的网络号及子网号，而0对应于主机号

计算机只需将IP地址和其对 应的子网掩码逐位"与"（逻辑 AND运算），就可得出相应子网的网络地址

主机的IP地址为 180.80.77.55，子网掩码为 255.255.252.0。若该主机向其所在子网发送广播分组，则目的地址：

子网划分与子网掩码、CIDR

无分类域间路由选择（CIDR）

消除了传统A、B、C 类地址及划分子网的概念，因而可以更有效地分配IPv4 的地址空间

CIDR使用"网络前缀"的概念代替子网络的概念。因此，IP地址的无分类两级编址为

IP= {<网络前缀>，<主机号>}。

例如，对于 128.14.32.5/20这个地址，它的掩码是 20个连续的1和后续12个连续的0，通过逐位相"与"的方法可以得到该地址的网络前缀（或直接截取前20 位）

逐位与 {
IP = 10000000.00001110.00100000.00000101
掩码 = 11111111.11111111.11110000.00000000
}
网络前缀 = 10000000.00001110.00100000.00000000（128.14.32.0）

使用 CIDR 时，路由表中的每个项目由"网络前缀"和"下一跳地址"组成。在查找路由表时可能会得到不止一个匹配结果。此时，应当从匹配结果中选择具有最长网络前缀的路由，因为网络前缀越长，其地址块就越小，因而路由就越具体

ARP、DHCP 与 ICMP

IP地址与硬件地址

IP 地址是网络层使用的地址，它是分层次等级的

硬件地址是数据链路层使用的地址，在网络层及网络层之上使用IP 地址，IP地址放在IP数据报的首部，而 MAC 地址放在 MAC帧的首部

通过数据封装，把IP数据报分组封装为 MAC帧后，数据链路层看不见数据报分组中的 IP 地址

地址解析协议（ARP）

IP 工作在网络层，其工作原理如下：主机 A欲向本局域网上的某台主机B发送IP数据报时，先在其 ARP高速缓存中查看有无主机B的IP地址

如有，就可查出其对应的硬件地址，再将此硬件地址写入 MAC 帧，然后通过局域网将该 MAC 帧发往此硬件地址

如果没有，那么就通过使用目的 MAC地址为 FF-FF-FF-FF-FF-FF 的帧来封装并广播ARP请求分组，使同一个局域网里的所有主机收到ARP请求

主机B收到该ARP请求后，向主机A发出响应ARP分组，分组中包含主机B的 IP 与MAC 地址的映射关系，主机A在收到后将此映射写入ARP缓存，然后按查询到的硬件地址发送 MAC帧

ARP、DHCP 与 ICMP

动态主机配置协议（DHCP）

动态主机配置协议常用于给主机动态地分配IP 地址，它提供了即插即用的联网机制，这种机制允许一台计算机加入新的网络和获取 IP 地址 而不用手工参与

DHCP 是应用层协议，它是基于 UDP 的

视频讲解更清晰

4.3 IPv4

DHCP 服务器聚合 DHCP 客户端的交换过程如下：

DHCP客户机广播"DHCP发现"消息，试图找到网络中的DHCP服务器，以便从DHCP服务器获得一个 IP地址

DHCP服务器收到"DHCP发现"消息后，向网络中广播"DHCP提供"消息，其中包括提供 DHCP客户机的IP地址和相关配置信息

DHCP客户机收到"DHCP提供"消息，如果接收 DHCP 服务器所提供的相关参数，那么通过广播"DHCP 请求"消息向 DHCP 服务器请求提供IP地址。

DHCP 服务器广播"DHCP确认"消息，将IP地址分配给 DHCP客户机

ARP、DHCP 与 ICMP

网际控制报文协议（ICMP）

ICMP 报文的种类有两种，即 ICMP 差错报告报文和ICMP 询问报文

ICMP 差错报告报文用于目标主机或到目标主机路径上的路由器向源主机报告差错和异常情况

终点不可达。当路由器或主机不能交付数据报时，就向源点发送终点不可达报文

源点抑制。当路由器或主机由于拥塞而丢弃数据报时，就向源点发送源点抑制报文

时间超过。当路由器收到生存时间为零的数据报时，除丢弃该数据报外还要向源点发送时间超过报文

不应发送 ICMP差错报告报文的几种情况如下

对ICMP差错报告报文不再发送 ICMP 差错报告报文

第一个分片的数据报片的所有后续数据报片都不发送ICMP 差错报告报文

对具有组播地址的数据报都不发送 ICMP 差错报告报文

4.4 IPv6

IPv6 的主要特点

更大的地址空间。IPv6将地址从IPv4的32位增大到了128位

支持即插即用（即自动配置）

IPv6 只有在包的源结点才能分片，是端到端的，传输路径中的路由器不能分片

IPv6 首部长度必须是 8B的整数倍，而IPv4 首部是 4B 的整数倍

从根本上解决了IP 地址的耗尽问题

路由信息协议（RIP）

RIP 规定

网络中的每个路由器都要维护从它自身到其他每个目的网络的距离记录

距离也称跳数（Hop Count），规定从一个路由器到直接连接网络的距离（跳数）为1。而每经过一个路由器，距离（跳数）加1

RIP认为好的路由就是它通过的路由器的数目少，即优先选择跳数少的路径

RIP允许一条路径最多只能包含 15个路由器（即最多允许15 跳）
因此距离等于16时，它表示网络不可达

RIP默认在任意两个使用 RIP 的路由器之间每 30秒广播一次RIP 路由更新信息，以便自动建立并维护路由表（动态维护）

路由信息协议（RIP）

RIP 的特点

仅和相邻路由器交换信息

路由器交换的信息是当前路由器所知道的全部信息，即自己的路由表

按固定的时间间隔交换路由信息，如每隔 30 秒

4.5 路由协议

距离向量算法

当原来的路由表中没有目的网络N时，把该项目添加到路由表中

当原来的路由表中有目的网络 N，且下一跳路由器的地址是X时，用收到的项目替换 原路由表中的项目

当原来的路由表中有目的网络N，且下一跳路由器的地址不是X时，如果收到的项目中的距离d小于路由表中的距离，那么就用收到的项目替换原路由表中的项目；否则什么也不做

缺点

网络出现故障时，会出现慢收敛现象（即需要较长时间才能将此信息传送到所有路由器），俗称"坏消息传得慢"，使更新过程的收敛时间长

开放最短路径优先（OSPF）协议

基本特点

向本自治系统中的所有路由器发送信息，这里使用的方法是洪泛法

发送的信息是与本路由器相邻的所有路由器的链路状态，但这只是路由器所知道的部分信息

只有当链路状态发生变化时，路由器才用洪泛法向所有路由器发送此信息，并且更新过程收敛得快

OSPF 是网络层协议，它不使用UDP或 TCP，而直接用 IP数据报传送（其IP数据报首部的协议字段为 89）；而 RIP 是应用层协议，它在传输层使用 UDP

4.5 路由协议

由于各路由器之间频繁地交换链路状态信息，因此所有路由器最终都能建立一个链路状态数据库。这个数据库实际上就是全网的拓扑结构图，它在全网范围内是一致的（称为链路状态数据库的同步）。然后，每个路由器根据这个全网拓扑结构图，使用Dijkstra最短路径算法计算从自己到各目的网络的最优路径，以此构造自己的路由表

边界网关协议（BGP）

是不同自治系统的路由器之间交换路由信息的协议，是一种外部网关协议。边界网关协议常用于互联网的网关之间。路由表包含已知路由器的列表、路由器能够达到的地址及到达每个路由器的路径的跳数

视频讲解更清晰

边界网关协议（BGP）

边界网关协议（BGP）只能力求寻找一条能够到达目的网络且比较好的路由（不能兜圈子），而并非寻找一条最佳路由。BGP采用的是路径向量路由选择协议，它与距离向量协议和链路状态协议有很大的区别。BGP 是应用层协议，它是基于 TCP 的

每个自治系统的管理员要选择至少一个路由器（可以有多个）作为该自治系统的"BGP发言人"

4.5 路由协议

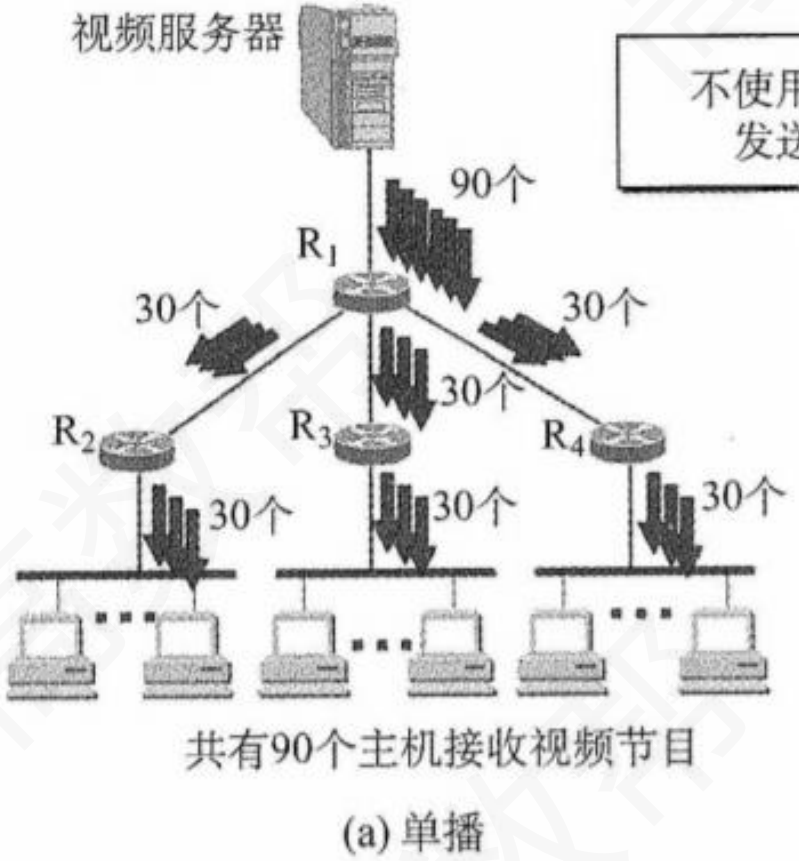
三种路由协议的比较

协　议	RIP	OSPF	BGP	
类型	内部	内部	外部	
路由算法	距离-向量	链路状态	路径-向量	
传递协议	UDP	IP	TCP	
路径选择	跳数最少	代价最低	较好，非最佳	
交换结点	和本结点相邻的路由器	网络中的所有路由器	和本结点相邻的路由器	
交换内容	当前本路由器知道的全部信息，即自己的路由表	与本路由器相邻的所有路由器的链路状态	首次	整个路由表
			非首次	有变化的部分

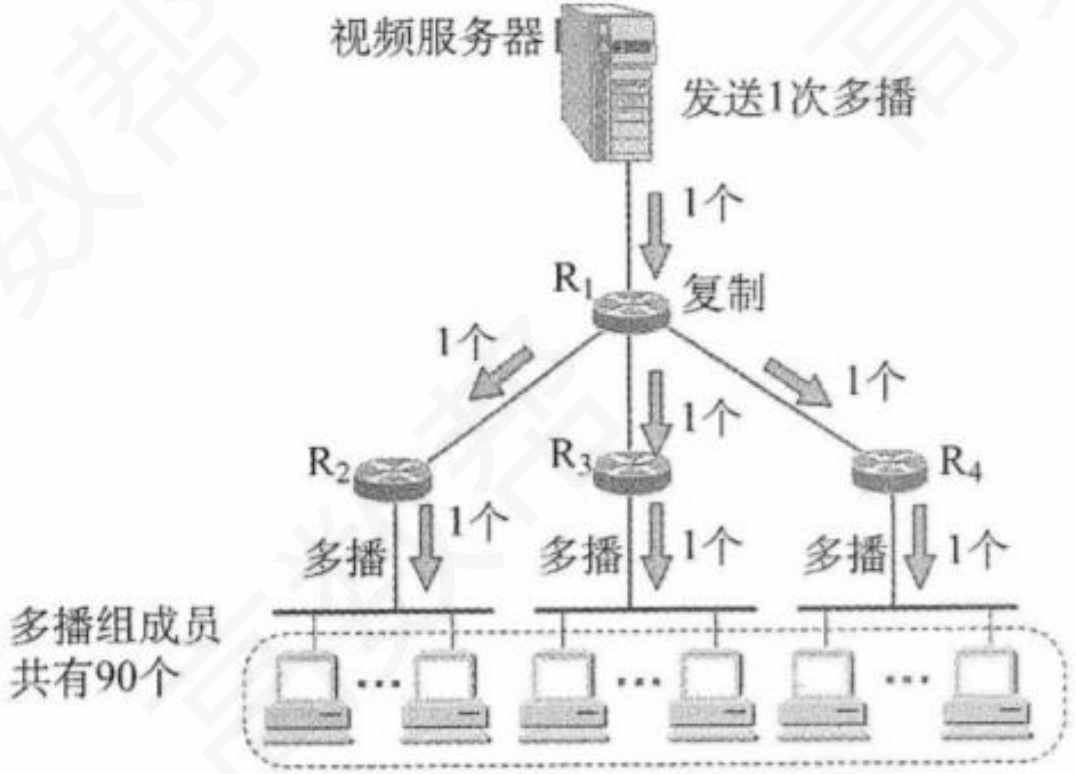
4.6 IP组播

组播的概念

用组播的缘由是，有的应用程序要把一个分组发送给多个目的地主机。不是让源主机给每个目的地主机都发送一个单独的分组，而是让源主机把单个分组发送给一个组播地址，该组播地址标识一组地址



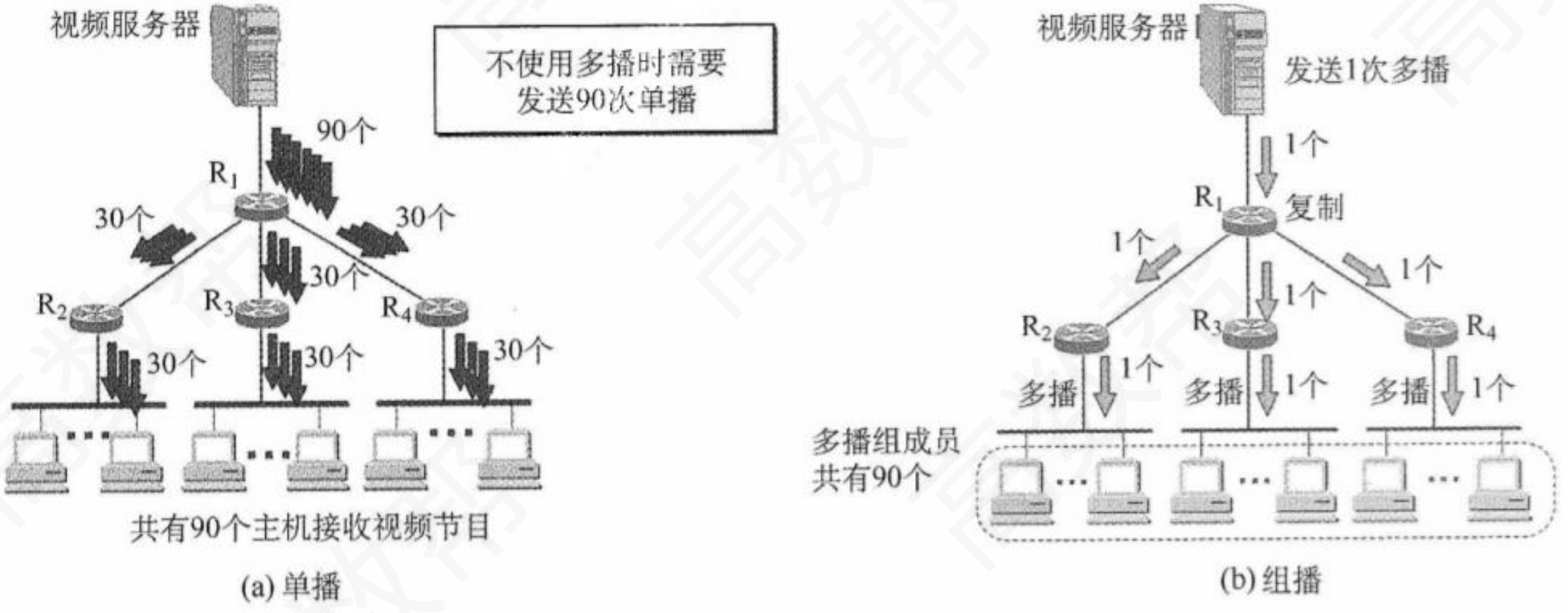
(a) 单播



(b) 组播

4.6 IP组播

主机使用一个称为 IGMP（因特网组管理协议）的协议加入组播组。它们使用该协议通知本地网络上的路由器关于要接收发送给某个组播组的分组的愿望。通过扩展路由器的路由选择和转发功能，可以在许多路由器互联的支持硬件组播的网络上面实现因特网组播



4.6 IP组播

IP 组播地址

P 组播使用D类地址格式。D类地址的前四位是 1110，因此D类地址范围是 224.0.0.0 ~ 239.255.255.255。每个D类IP地址标志一个组播组

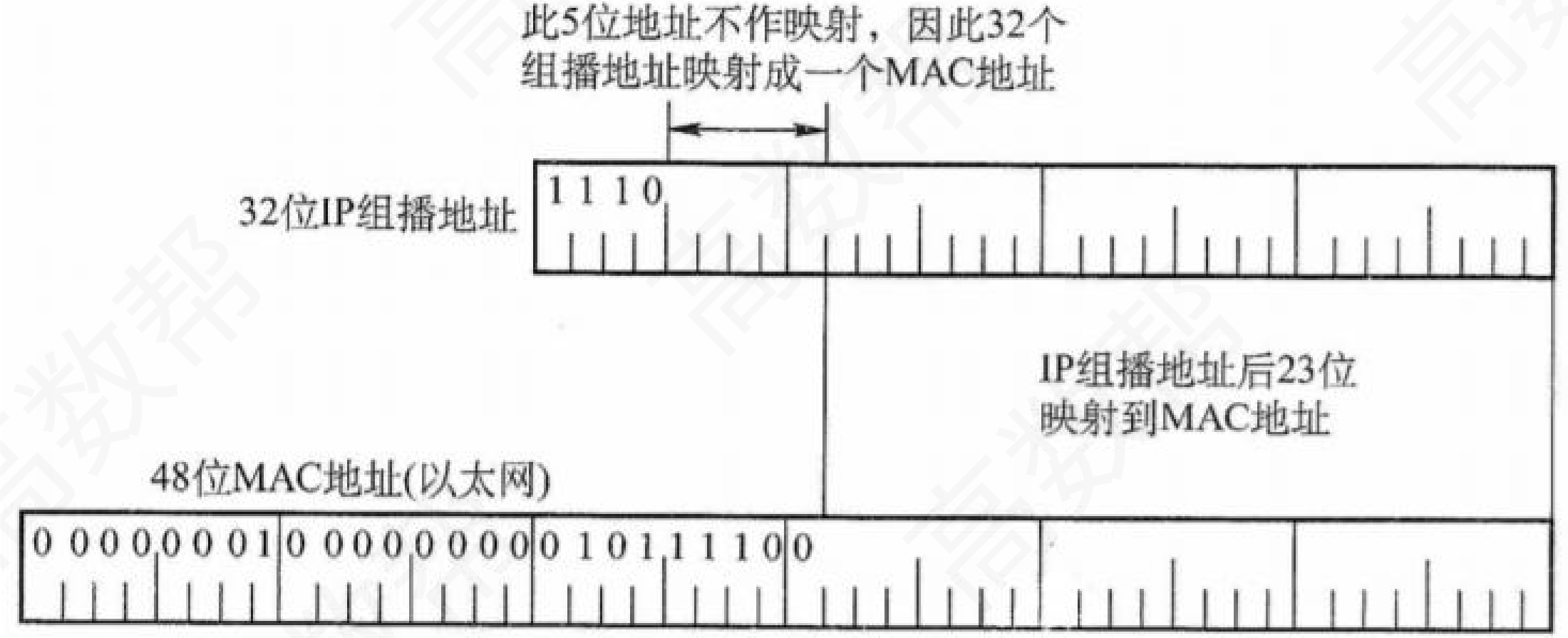
IP 组播可以分为两种：一种只在本局域网上进行硬件组播；另一种则在因特网的范围内进行组播。在因特网上进行组播的最后阶段，还是要把组播数据报在局域网上用硬件组播交付给组播组的所有成员

4.6 IP组播

下面讨论这种硬件组播

D类 IP 地址与以太网组播地址的映射关系：

以太网组播地址的范围是从01-00-5E-00-00-00到 01-00-5E-7F-FF-FF



IGMP 与组播路由算法

第一阶段：当某台主机加入新的组播组时，该主机应向组播组的组播地址发送一个IGMP报文，声明自己要成为该组的成员。本地的组播路由器收到IGMP报文后，将组成员关系转发给因特网上的其他组播路由器

第二阶段：因为组成员关系是动态的，本地组播路由器要周期性地探询本地局域网上的主机，以便知道这些主机是否仍继续是组的成员。只要对某个组有一台主机响应，那么组播路由器就认为这个组是活跃的。但一个组在经过几次的探询后仍然没有一台主机响应时，则不再将该组的成员关系转发给其他的组播路由器

移动IP

概念

基于IPv4的移动 IP 定义三种功能实体：移动结点、归属代理（也称本地代理）和外埠代理（也称外部代理）。归属代理和外埠代理又统称为移动代理。

移动结点。具有永久IP地址的移动结点

本地代理。在一个网络环境中，一个移动结点的永久"居所"被称为归属网络，在归属网络中代表移动结点执行移动管理功能的实体称为归属代理（本地代理），它根据移动用户的转交地址，采用隧道技术转交移动结点的数据包

外部代理。在外部网络中帮助移动结点完成移动管理功能的实体称为外部代理

4.7 移动IP

移动IP 通信过程

基本流程

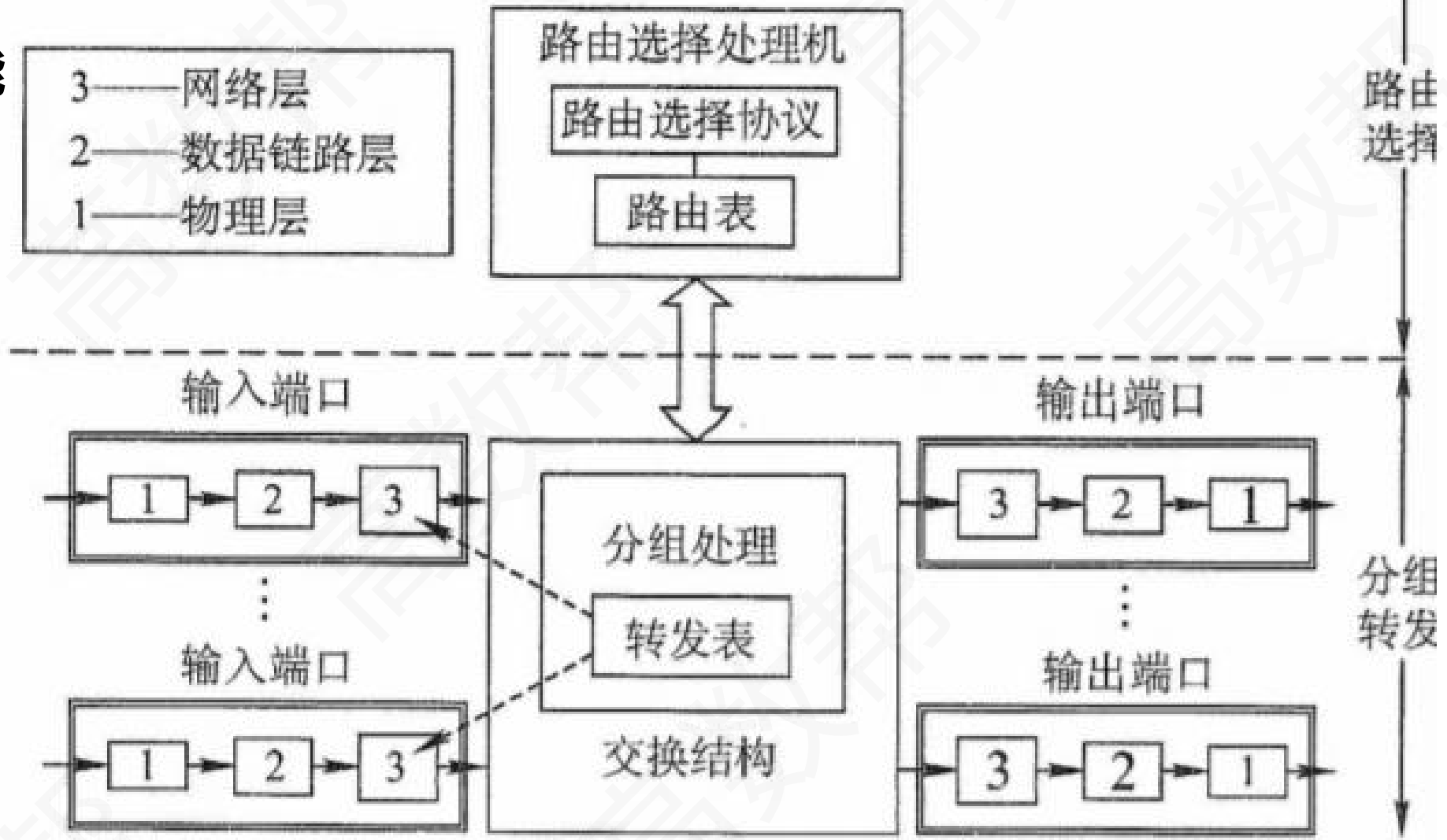
移动结点漫游到一个外地网络时，仍然使用固定的P地址进行通信。为了能够收到通信对端发给它的IP 分组，移动结点需要向本地代理注册当前的位置地址，这个位置地址就是转交地址

本地代理接收来自转交地址的注册后，会构建一条通向转交地址的隧道，将截获的发给移动结点的 IP 分组通过隧道送到转交地址处

在转交地址处解除隧道封装，恢复原始的IP分组，最后送到移动结点，这样移动结点在外网就能够收到这些发送给它的IP分组

4.8 网络层设备

路由器的组成和功能



路由表和路由转发