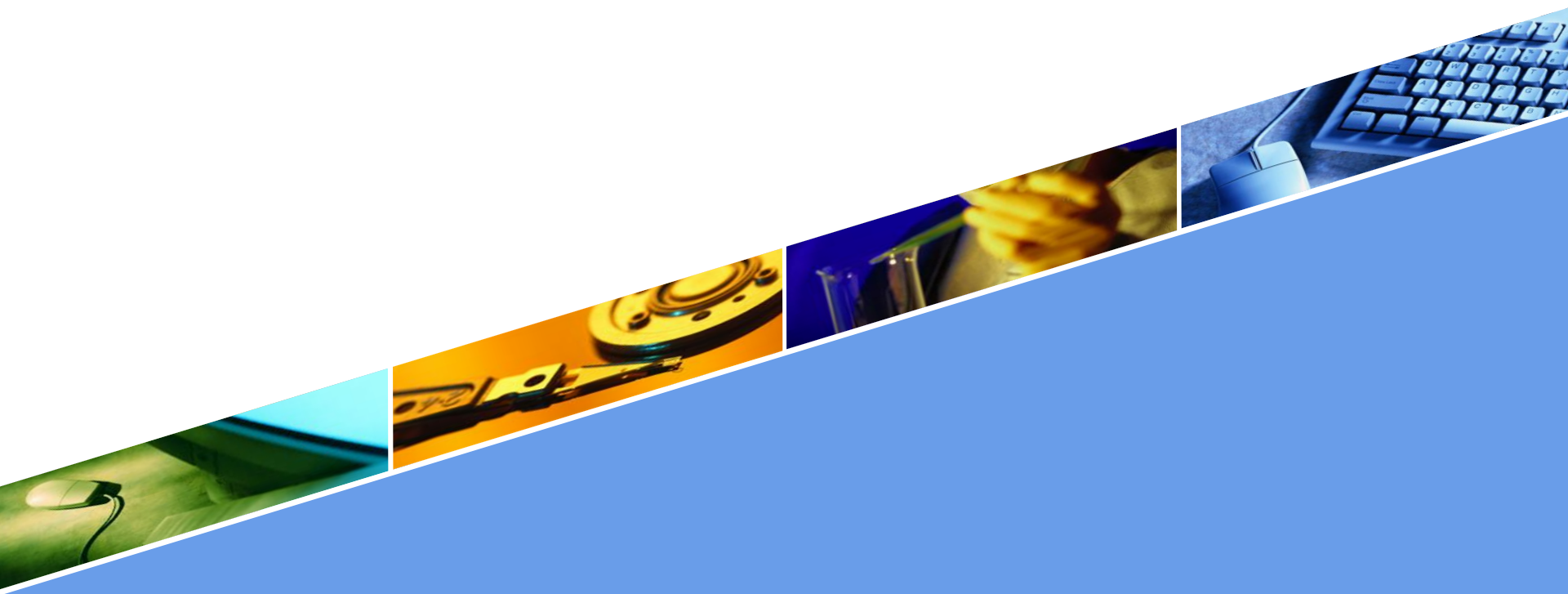


Internet网际互连及网络层协议



主题 1



1 网络层概述

2 网际互连与网络设备

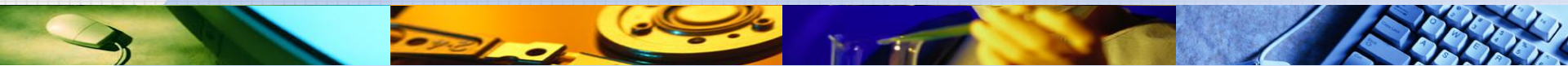
3 IP地址

4 地址解析协议

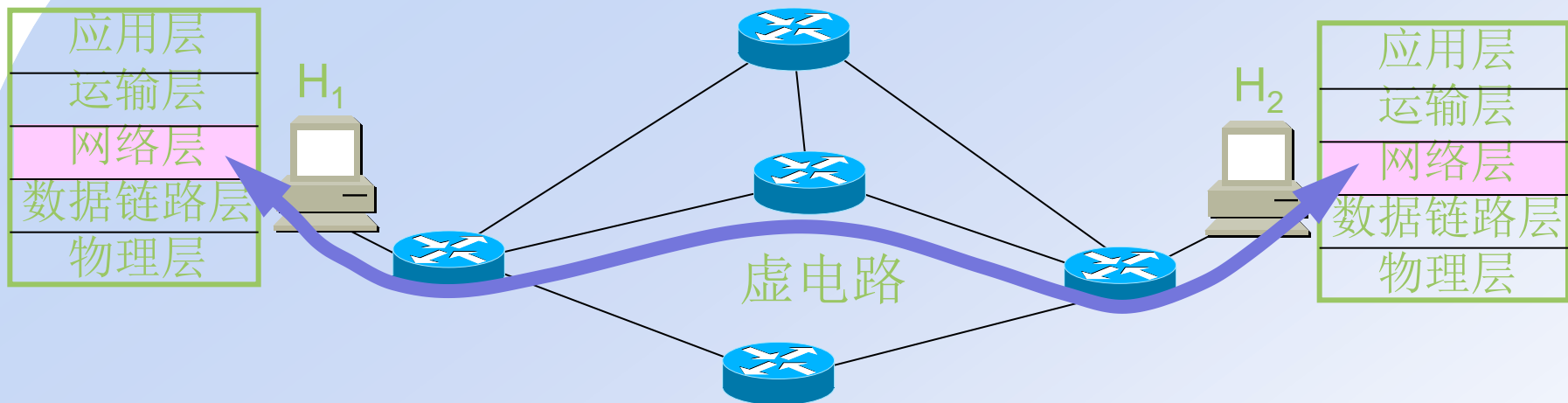
5 IP协议

6 ICMP协议

网络层提供的两种服务

- 
- ❖ 在计算机网络领域，网络层应该向运输层提供怎样的服务（“面向连接”还是“无连接”）曾引起了长期的争论。
 - ❖ 争论焦点的实质就是：在计算机通信中，可靠交付应当由谁来负责？是网络还是端系统？
 - ❖ 计算机网络层的设计包括两种服务方式：
 - 面向连接的虚电路服务：
 - 无连接的数据报服务

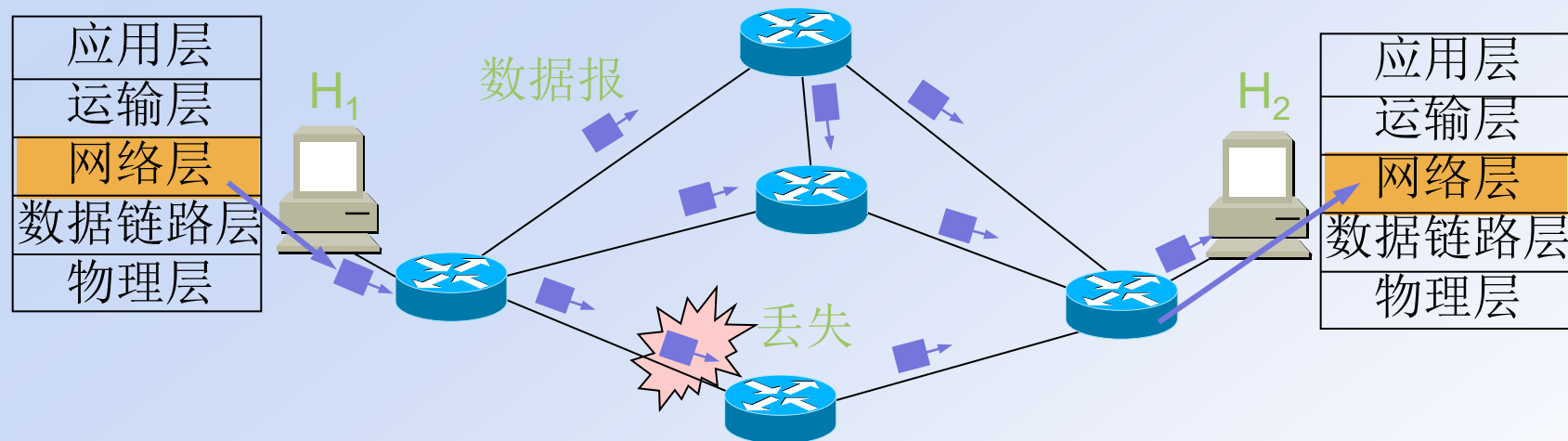
虚电路服务



H_1 发送给 H_2 的所有分组都沿着同一条虚电路传送

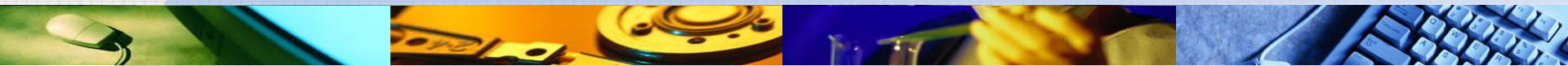
- ❖ 虚电路表示这只是一条**逻辑上的连接**，分组都沿着这条逻辑连接按照存储转发方式传送，而并不是真正建立了一条物理连接。
- ❖ 请注意，电路交换的电话通信是先建立了一条**真正的连接**。因此分组交换的虚连接和电路交换的连接只是类似，但并不完全一样。

数据报服务



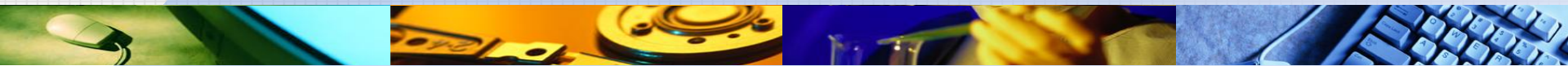
H_1 发送给 H_2 的分组可能沿着不同路径传送

因特网采用的设计思路

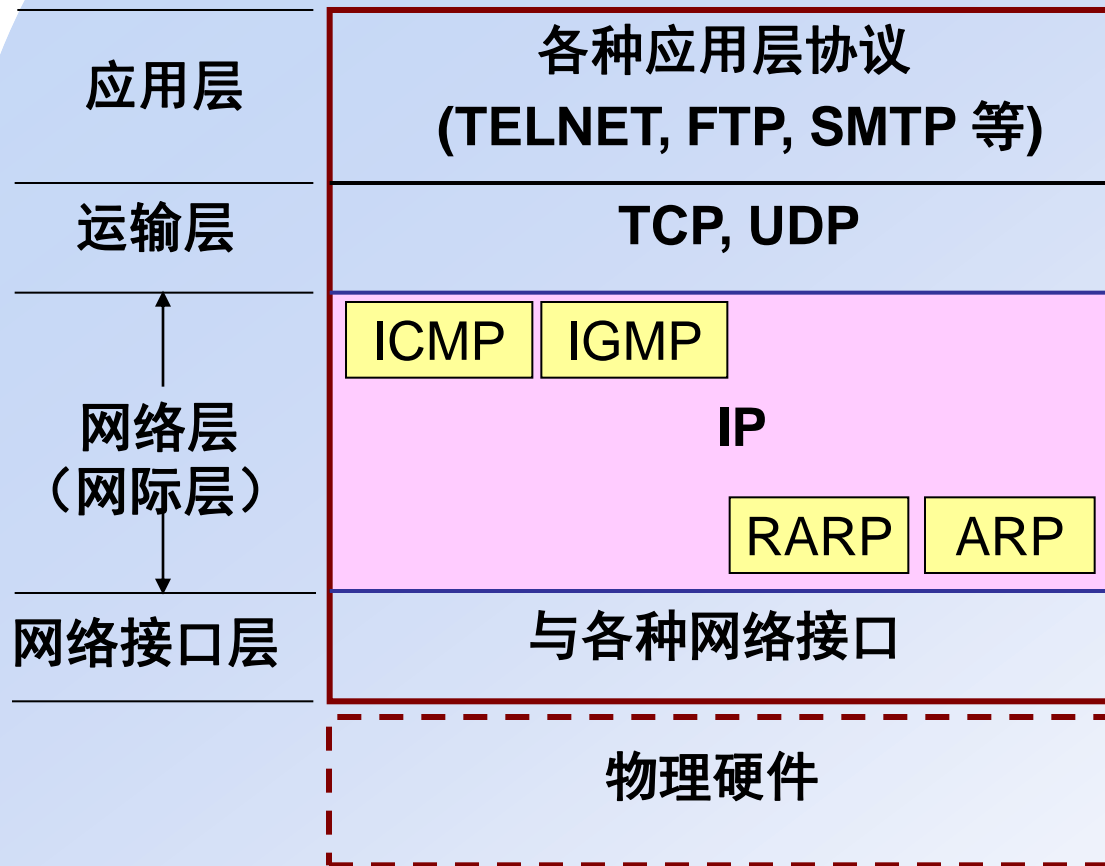


- ❖ 网络层向上只提供简单灵活的、无连接的、尽最大努力交付的数据报服务。
- ❖ 网络在发送分组时不需要先建立连接。每一个分组（即 IP 数据报）独立发送，与其前后的分组无关（不进行编号）。
- ❖ 网络层不提供服务质量的承诺。即所传送的分组可能出错、丢失、重复和失序（不按序到达终点），当然也不保证分组传送的时限。
- ❖ 如果主机（即端系统）中的进程之间的通信需要是可靠的，那么就由网络的主机中的运输层负责（包括差错处理、流量控制等）

尽最大努力交付的好处

- 
- ❖ 网络中的路由器可以做得比较简单，而且价格低廉（与电信网的交换机相比较）。
 - ❖ 网络的造价大大降低，运行方式灵活，能够适应多种应用。
 - ❖ 因特网能够发展到今日的规模，充分证明了当初采用这种设计思路的正确性。

网络层（网际层）协议



网际协议 IP 是 TCP/IP 体系中两个最主要的协议之一。与 IP 协议配套使用的还有四个协议：

- 地址解析协议 ARP
(Address Resolution Protocol)
- 逆地址解析协议 RARP
(Reverse Address Resolution Protocol)
- 因特网控制报文协议 ICMP
(Internet Control Message Protocol)
- 因特网组管理协议 IGMP
(Internet Group Management Protocol)

主题 2



1 网络层概述

2 网际互连与网络设备

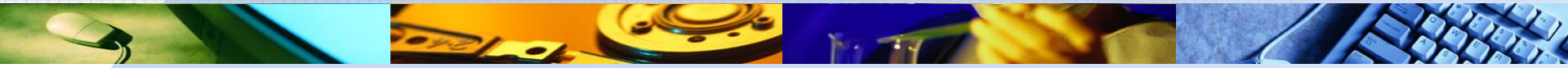
3 IP地址

4 地址解析协议

5 IP协议

6 ICMP协议

网际互连的复杂性和问题



❖ 互连在一起的各种异构的网络要进行通信，会遇到许多问题需要解决，如：

- 不同的寻址方案
- 不同的最大分组长度
- 不同的网络接入机制
- 不同的超时控制
- 不同的差错恢复方法
- 不同的状态报告方法
- 不同的路由选择技术
- 不同的用户接入控制
- 不同的服务（面向连接服务和无连接服务）
- 不同的管理与控制方式

利用 统一的**IP** 协议构建全球互联网（**Internet**），异构网络的通信作为网络接口层协议：可以使这些性能各异的网络从用户看起来好像是一个统一的网络——**虚拟互联网**

网络设备：统称，网际层设备仅路由器

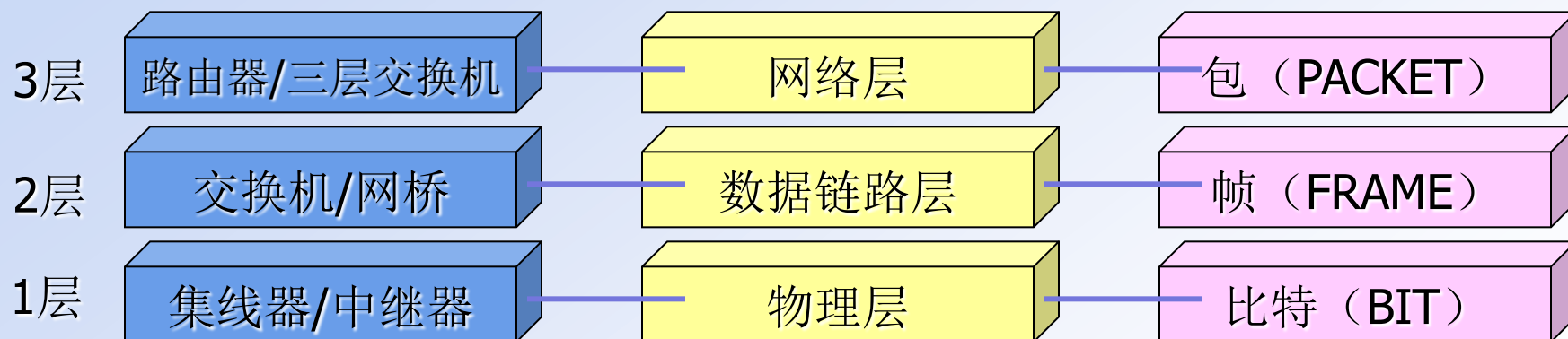


❖ **网络设备**：将网络互相连接起来的设备，也称为中继（relay）系统，根据中继系统所处的层次的不同，可有以下五种中继系统：

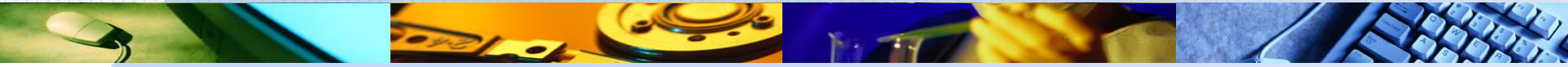
- **物理层中继系统：转发器（repeater）。**
- **数据链路层中继系统：网桥（bridge）、交换机（switch）。**
- **网络层中继系统：路由器（router）。**
- **网桥和路由器的混合物：桥路器（brouter）、三层交换机。兼有网桥和路由器的功能。**
- **在网络层以上的中继系统：网关（gateway）。主要实现高层协议的转换。**

很少有运输层之上的
中继设备，通常作为
安全控制的一部分

主要网络设备

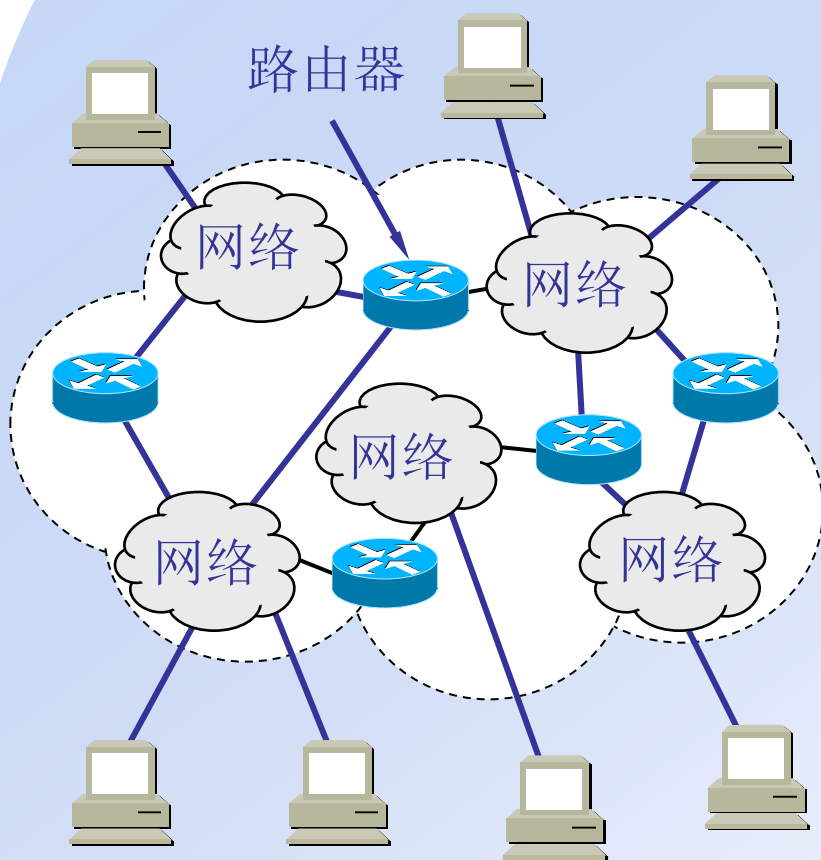


网络互连或网际互连



- 当中继系统是转发器或网桥时，一般并不称之为网络互连，因为这仅仅是把一个网络扩大了，而这仍然是一个网络。
- 互联网都是指用路由器进行互连的网络。
- 由于历史的原因，许多有关 TCP/IP 的文献将网络层使用的路由器称为网关。

互连网络与虚拟互连网络



(a) 互连网络



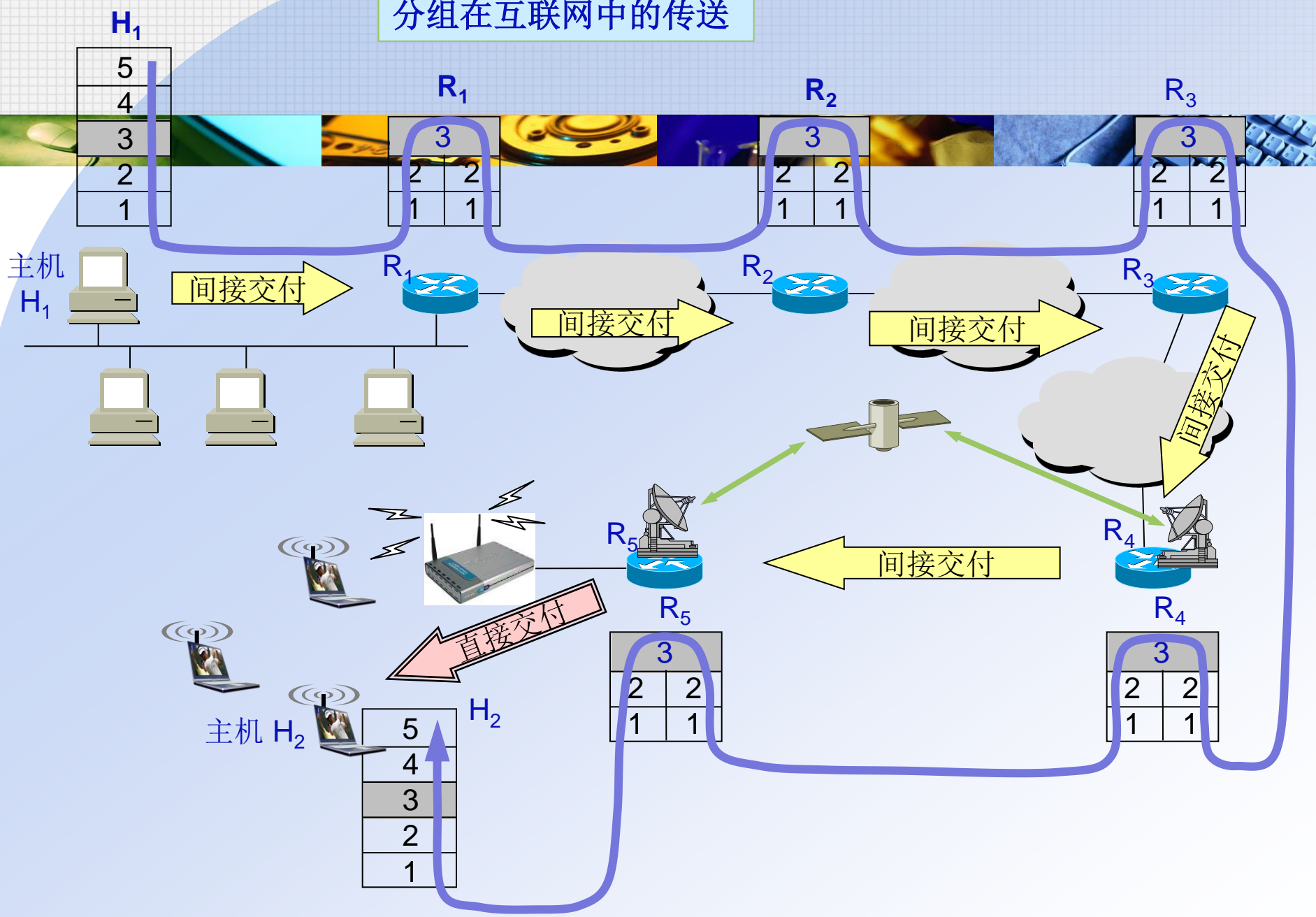
(b) 虚拟互连网络

虚拟互连网络的意义



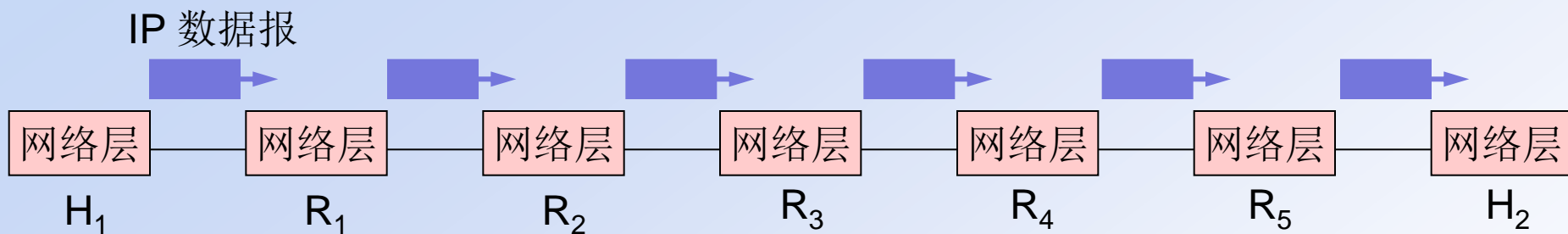
- ❖ 所谓虚拟互连网络也就是逻辑互连网络，它的意思就是互连起来的各种物理网络的异构性本来是客观存在的，但是我们利用 IP 协议就可以使这些性能各异的网络从用户看起来好像是一个统一的网络。
- ❖ 使用 IP 协议的虚拟互连网络可简称为 IP 网。
- ❖ 使用虚拟互连网络的好处是：当互联网上的主机进行通信时，就好像在一个网络上通信一样，而看不见互连的各具体的网络异构细节。

分组在互联网中的传送

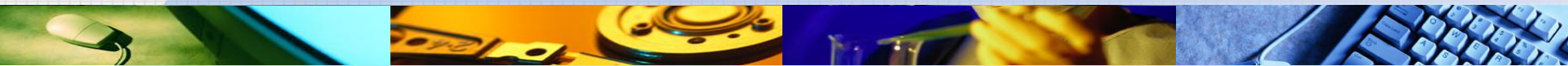


从网络层看 IP 数据报的传送

❖ 如果我们只从网络层考虑问题，那么 IP 数据报就可以想象是在网络层中传送。



主题 3



1 网络层概述

2 网际互连与网络设备

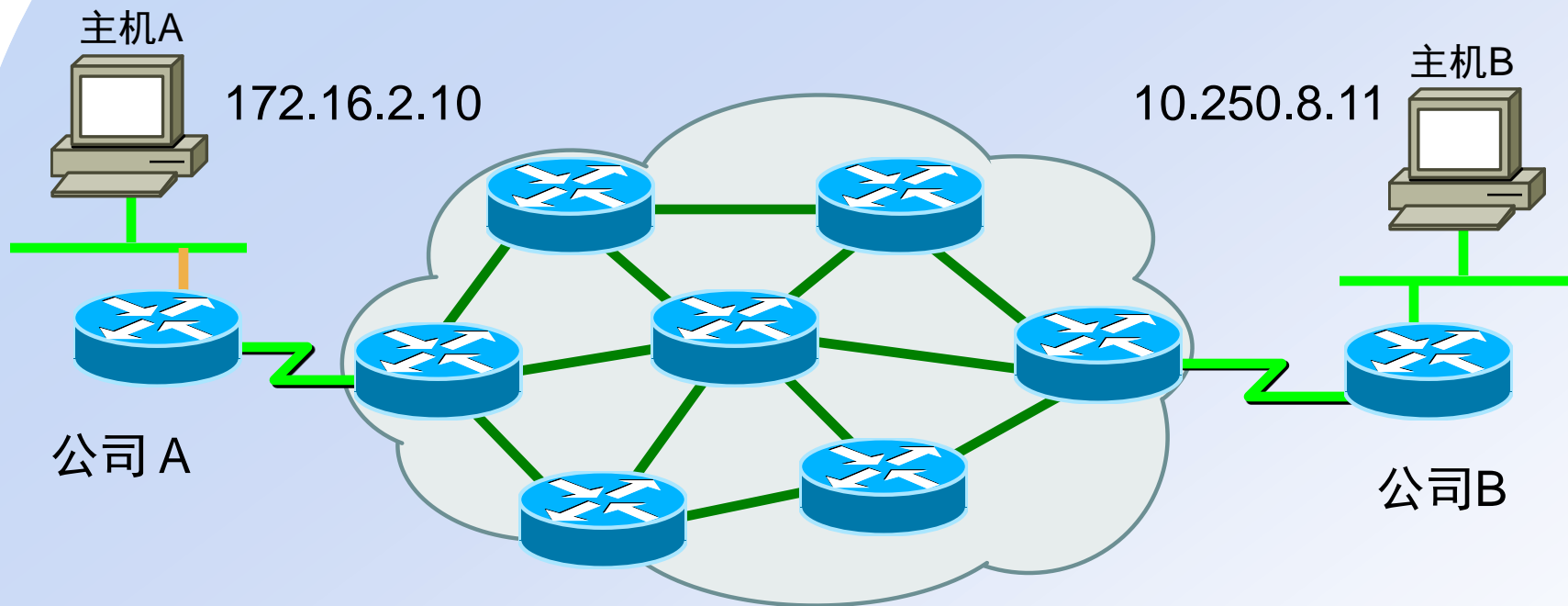
3 IP地址

4 地址解析协议

5 IP协议

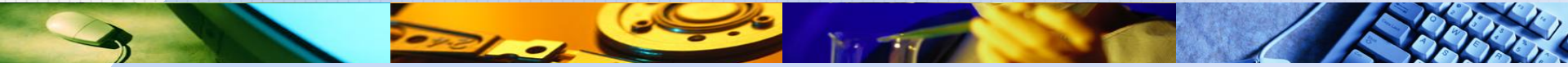
6 ICMP协议

IP地址表示和重要性



- 唯一编址使得端系统间的通信成为可能。

IP地址类型



- 把整个因特网看成为一个单一的、抽象的网络。IP 地址就是给每个连接在因特网上的主机（或路由器）分配一个在全世界范围是惟一的 32 bit 的标识符。
- IP地址现在由因特网名字与号码指派公司ICANN进行分配。
- IP 地址的编址方法：
 - ❖ 分类的 IP 地址，这是最基本的编址方法。〔1981年〕
 - ❖ 子网的划分，这是对最基本的编址方法的改进。〔1985年—RFC950〕
 - ❖ 构成超网，这是比较新的无分类编址方法。〔1993年〕

分类的IP地址



❖ 每一类地址都由两个固定长度的字段组成，其中一个字段是**网络号 net-id**，它标志主机（或路由器）所连接到的网络，而另一个字段则是**主机号 host-id**，它标志该主机（或路由器）。

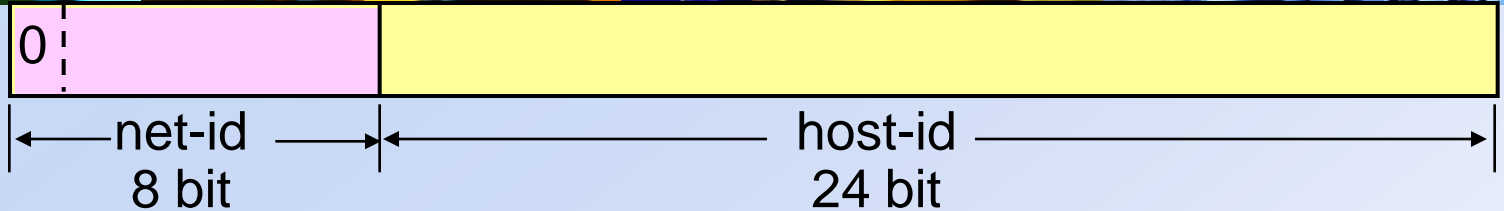
❖ 两级的 IP 地址可以记为：

IP 地址 ::= { <网络号>, <主机号> }

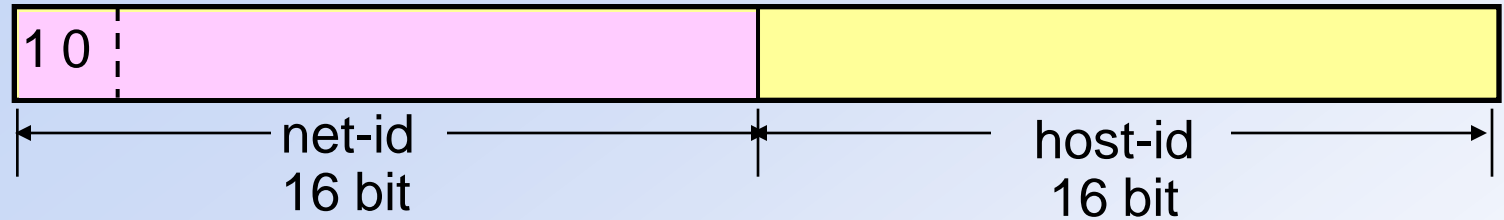
::= 代表“定义为”

分类的IP地址

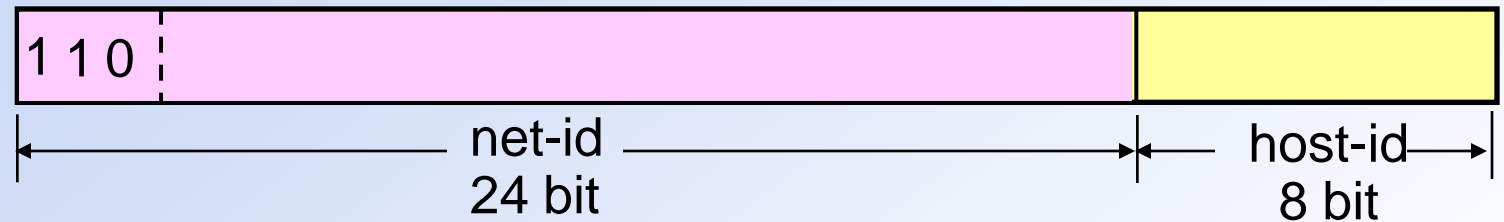
A 类地址



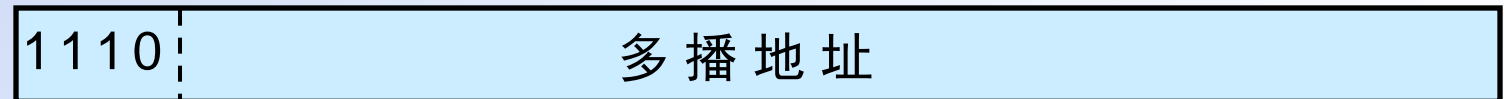
B 类地址



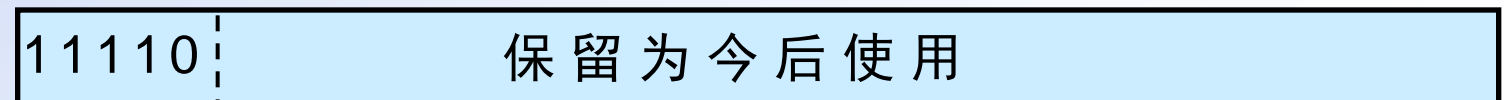
C 类地址



D 类地址



E 类地址



点分十进制记法

机器中存放的 IP 地址
是 32 位 二进制代码

10000000000010110000001100011111

每隔 8 位插入一个空格
能够提高可读性

10000000 00001011 00000011 00011111

将每 8 位的二进制数
转换为十进制数

128

11

3

31

采用点分十进制记法
则更加便于使用

128.11.3.31

常用的三类类别的 IP 地址

IP 地址的使用范围

网络类别	最大网络数	第一个可用的网络号	最后一个可用的网络号	每个网络中最大的主机数
A	$126 (2^7 - 2)$	1	126	16,777,214
B	$16,383 (2^{14} - 1)$	128.1	191.255	65,534
C	$2,097,151 (2^{21} - 1)$	192.0.1	223.255.255	254

一般不使用的特殊IP地址



net-id	host-id	作源地址	作目的地址	含 意
0	0	可以	不可	在本网络上的本主机
0	host-id	可以	不可	在本网络上的某个主机
全1	全1	不可	可以	只在本网络上进行广播
net-id	全1	不可	可以	对net-id 上的所有主机广播
127	任意	可以	可以	用作本地软件回送测试 (loopback test)之用

特殊地址说明



a. 回送地址

IP地址中以127开始的IP地址作为保留地址，被称为“回送地址”。回送地址用于网络软件的测试，以及本地进程的通信。

任何程序一旦接到使用了回送地址为目的地址的数据，则该程序将不再转发数据，而是将其立即回送给源地址。

例如，使用“ping 127.0.0.1”可以通过ping软件测试本地网卡进程之间的通信。

TCP/IP协议规定：

含网络号为127的分组不能出现在任何网络上；
主机和路由器不能为该地址广播任何寻址信息。

b. 直接广播地址 (*Directed Brordcasting*)

将主机号各位全为“1”的IP地址称为直接广播地址。该地址主要用于广播，在使用时，用来代表该网络上所有的主机，例如，202.112.144是一个C类的网络标识，该网络的广播地址就是202.112.144.255；当该网络中的某台主机需要发送广播时，就可以使用这个地址向该网络上的所有主机发送报文

特殊地址说明

c. 有限广播地址 (*Limiting Brordcasting*)

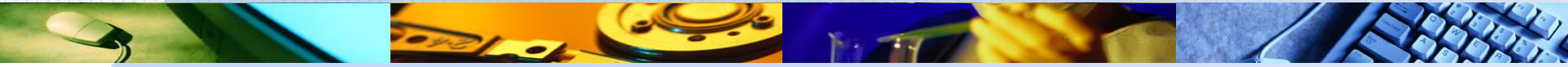
TCP/IP协议规定，32比特位全为“1”的IP地址（255.255.255.255）为“有限广播地址”，这个地址主要用来进行本网广播。当需要在本网内广播，又不知道本网的网络号时，即可使用“有限（受限）广播地址”

- 某主机可以用来将一个分组以广播方式发送给本网的所有主机；
- 分组将被本网的所有主机将接受该分组，路由器则阻挡该分组通过。

d. 本网地址

- 本网地址：将IP地址中主机地址位全为“0”的IP地址叫做本网地址。这个地址用来表示“本主机所连接的网路”。例如，用“128.16.0.0”表示“128.16”这个B类网路；用“202.112.144.0”表示“202.112.144”这个C类网路。本网地址又被称为“0”地址。

本地地址和全球地址



◆ [RFC 1918]指明的专用地址

- 10.0.0.0 到 10.255.255.255
- 172.16.0.0 到 172.31.255.255
- 192.168.0.0 到 192.168.255.255

➤ 这些地址只能用于一个机构的内部通信，而不能用于和因特网上的主机通信。

➤ 专用地址只能用作本地地址而不能用作全球地址。在因特网中的所有路由器对目的地址是专用地址的数据报一律不进行转发。

IP 地址的一些重要特点

(1) IP 地址是一种分等级的地址结构。分两个等级的好处是：

- ❖ 第一，IP 地址管理机构在分配 IP 地址时 **只分配网络号**，而剩下的主机号则由得到该网络号的单位自行分配。这样就方便了 IP 地址的管理。
- ❖ 第二，路由器仅 **根据目的主机所连接的网络号来转发分组**（而不考虑目的主机号），这样就可以使路由表中的项目数大幅度减少，从而减小了路由表所占的存储空间。

IP 地址的一些重要特点



(2) 实际上 IP 地址是标志一个主机（或路由器）和一条链路的接口。

- ❖ 当一个主机同时连接到两个网络上时，该主机就必须同时具有两个相应的 IP 地址，其网络号 net-id 必须是不同的。这种主机称为多归属主机 (multihomed host)。
- ❖ 由于一个路由器至少应当连接到两个网络（这样它才能将 IP 数据报从一个网络转发到另一个网络），因此一个路由器至少应当有两个不同的 IP 地址。

IP 地址的一些重要特点



- (3) 用转发器、网桥或交换机连接起来的若干个局域网仍为一个网络，因此这些局域网都具有同样的网络号 net-id。
- (4) 所有分配到网络号 net-id 的网络，无论范围很小的局域网，还是可能覆盖很大地理范围的广域网，都是平等的。

主题 4



1 网络层概述

2 网际互连与网络设备

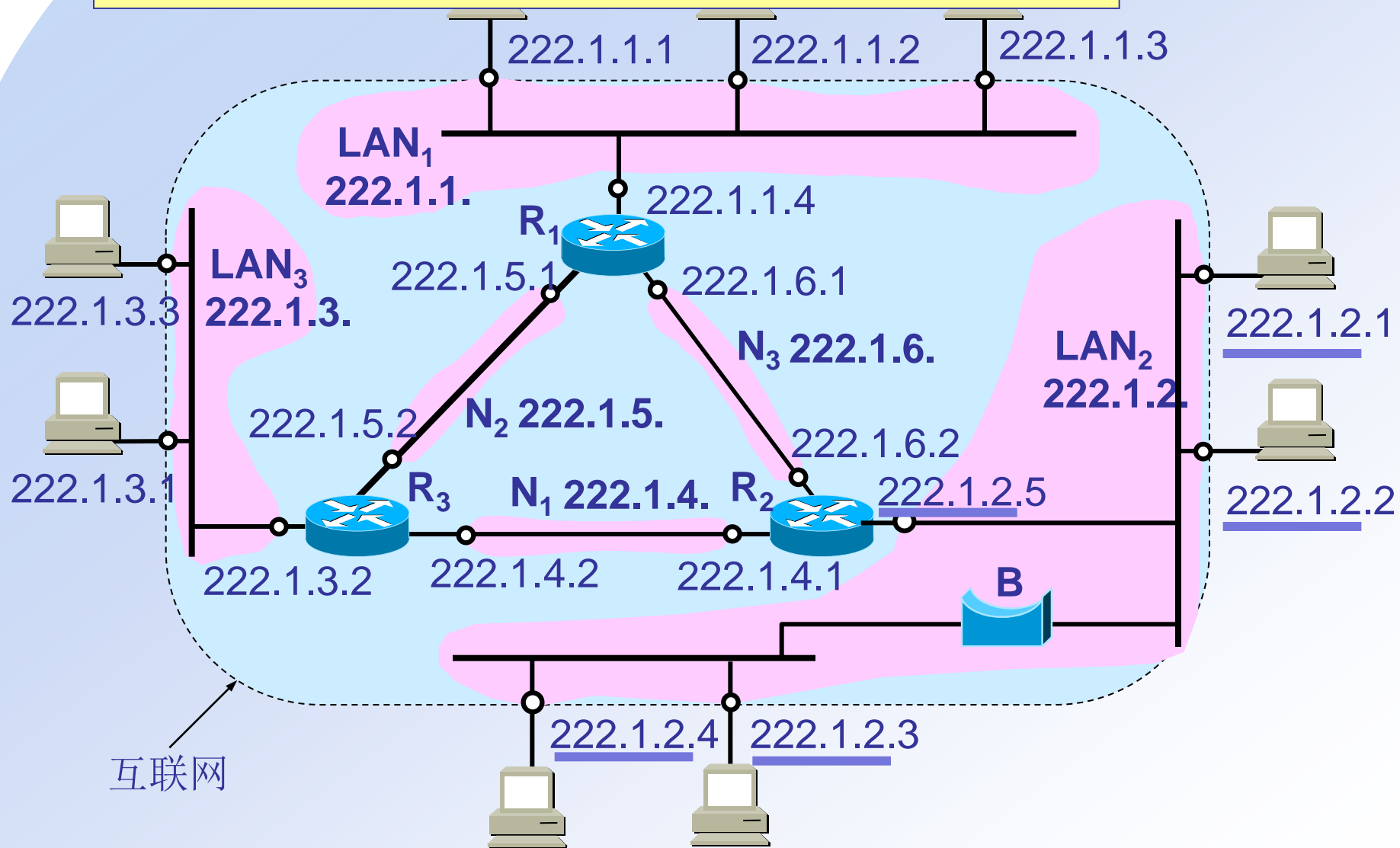
3 IP地址

4 地址解析协议

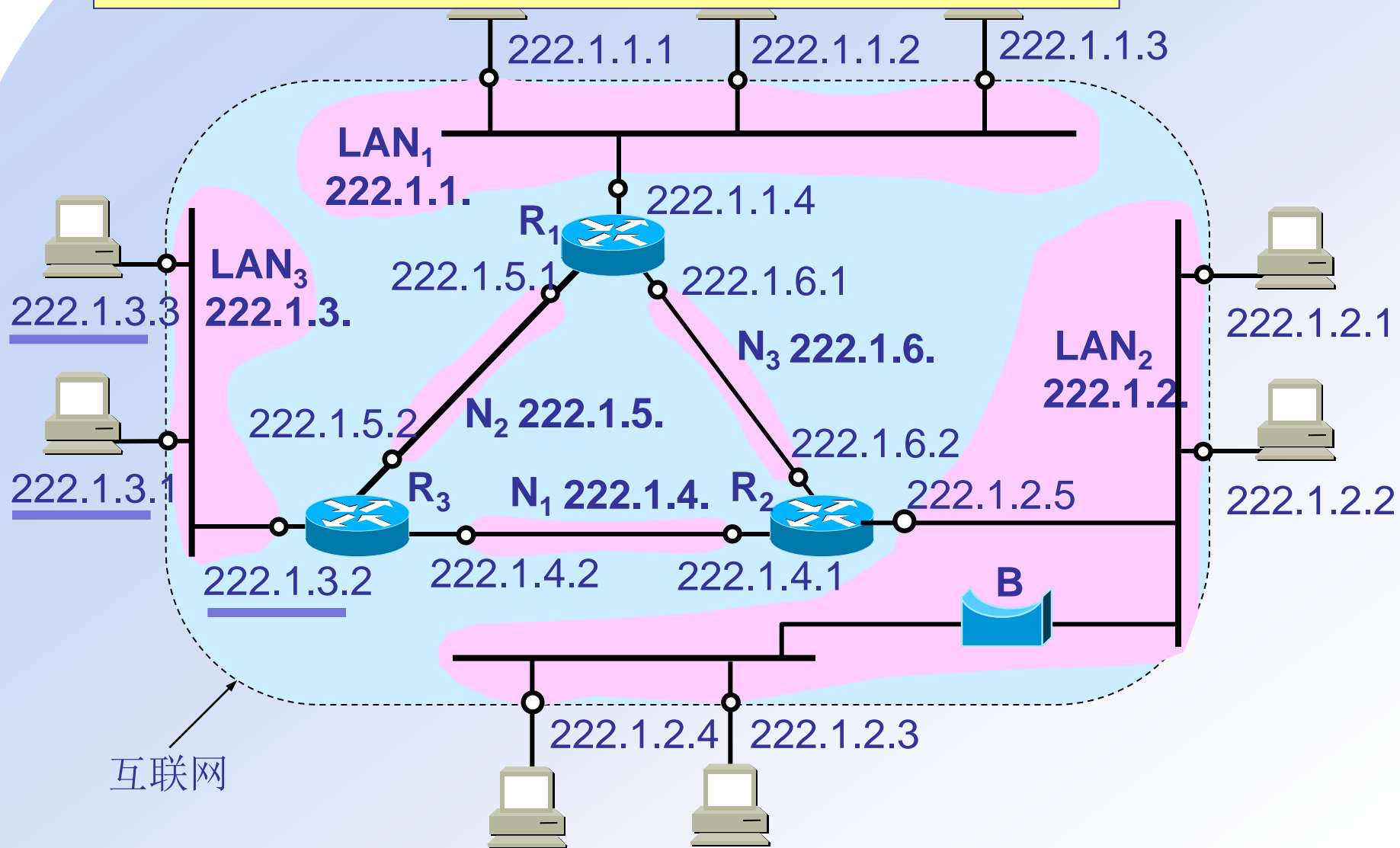
5 IP协议

6 ICMP协议

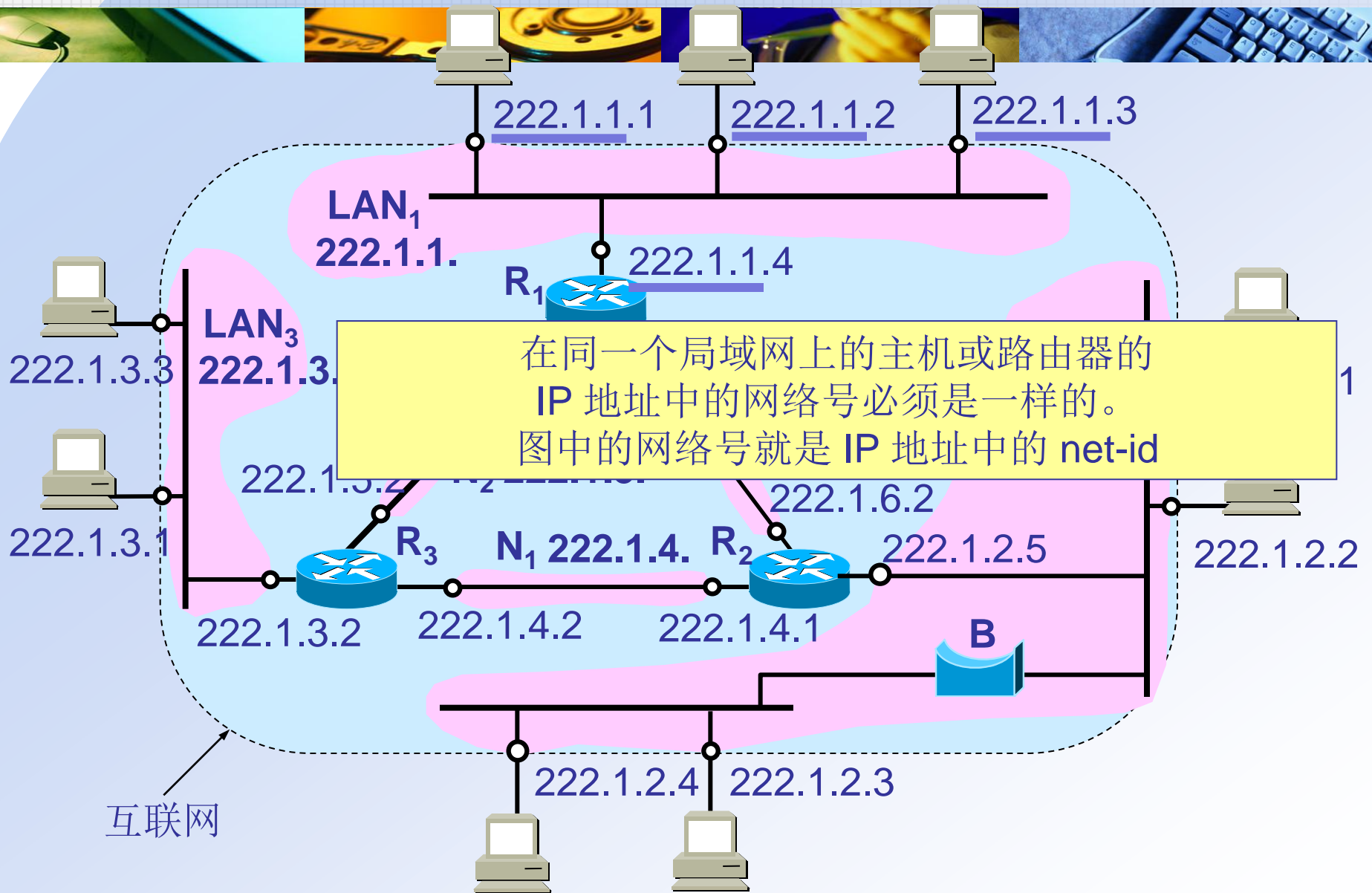
在同一个局域网上的主机或路由器的
IP 地址中的网络号必须是一样的。
图中的网络号就是 IP 地址中的 net-id



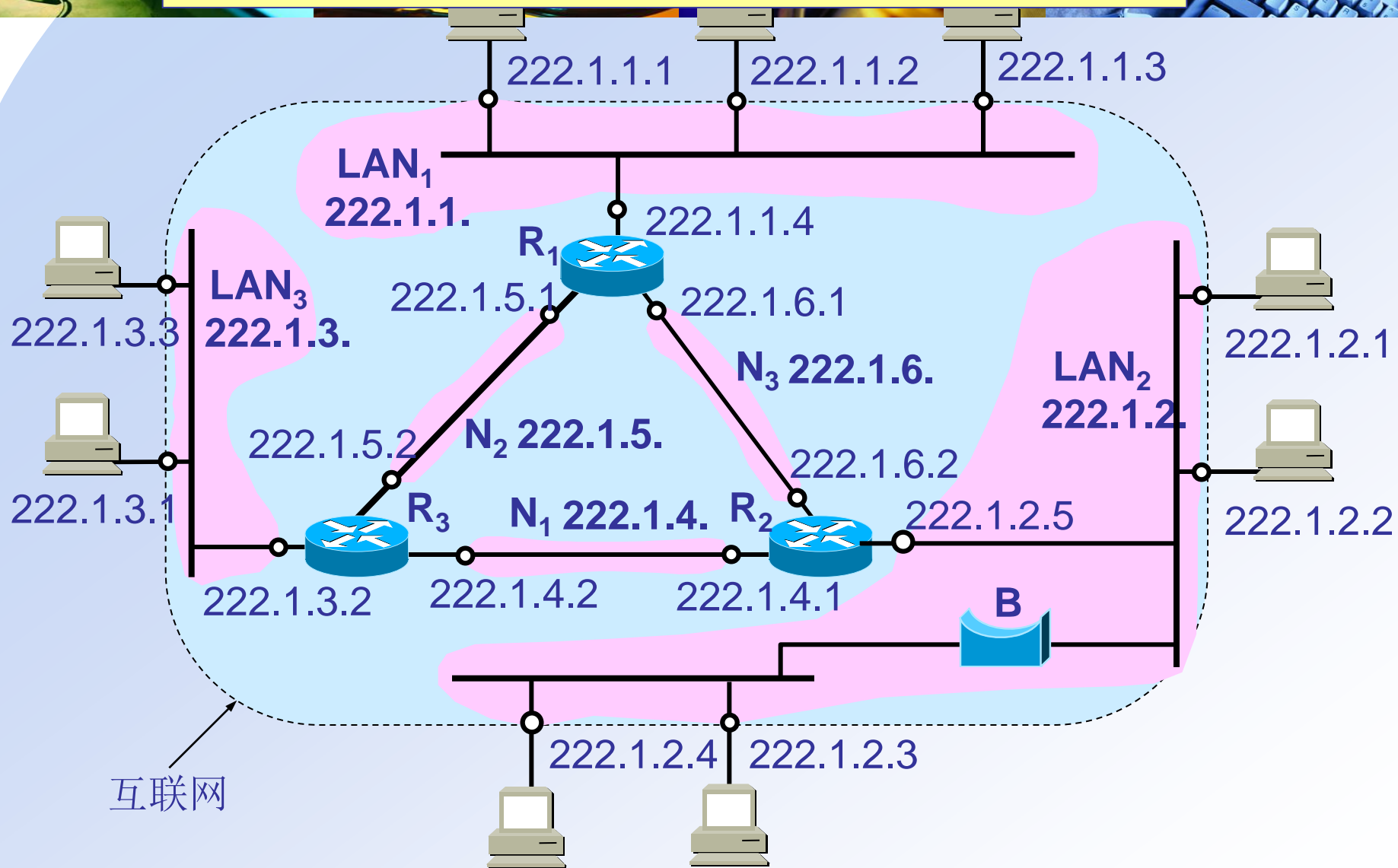
在同一个局域网上的主机或路由器的
IP 地址中的网络号必须是一样的。
图中的网络号就是 IP 地址中的 net-id



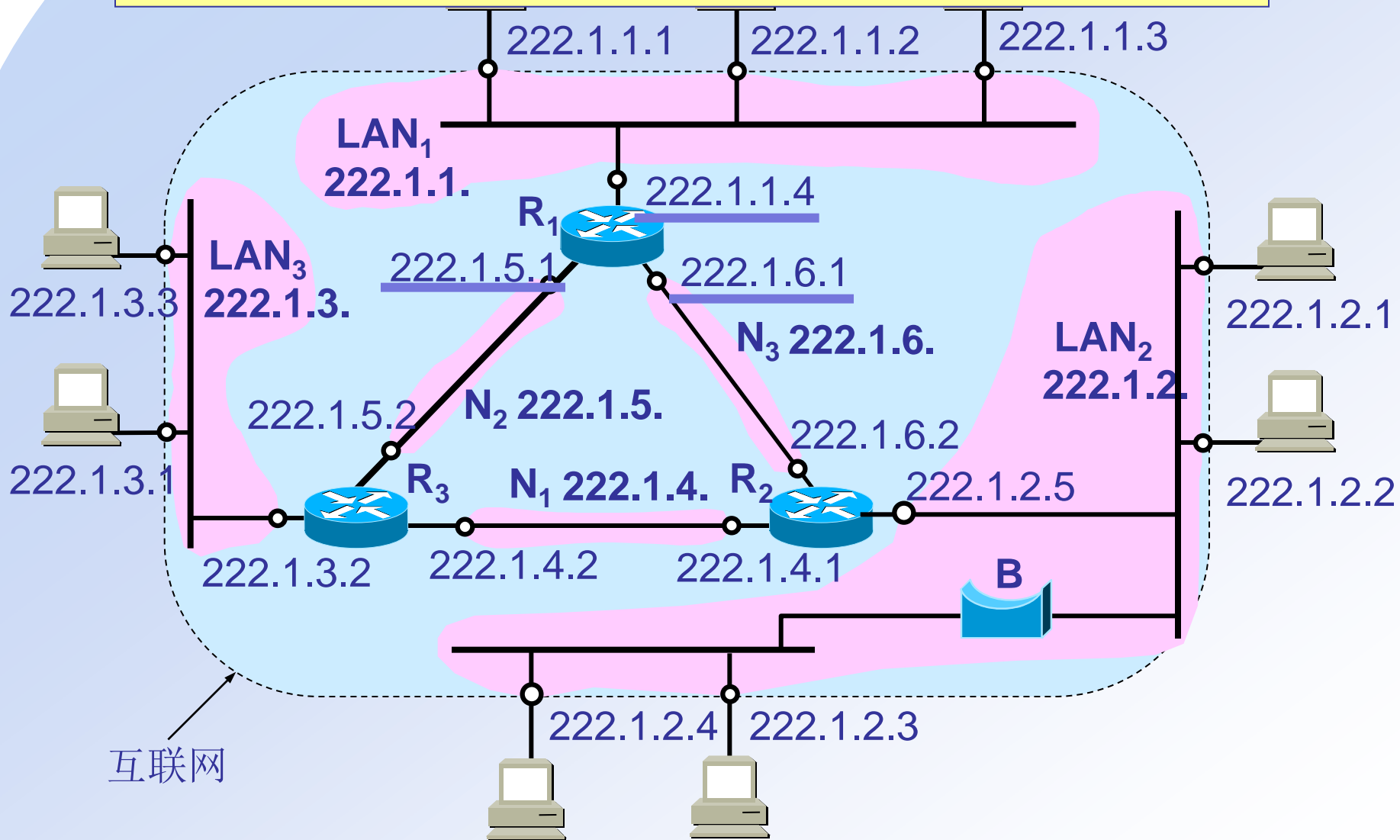
互联网中的 IP 地址



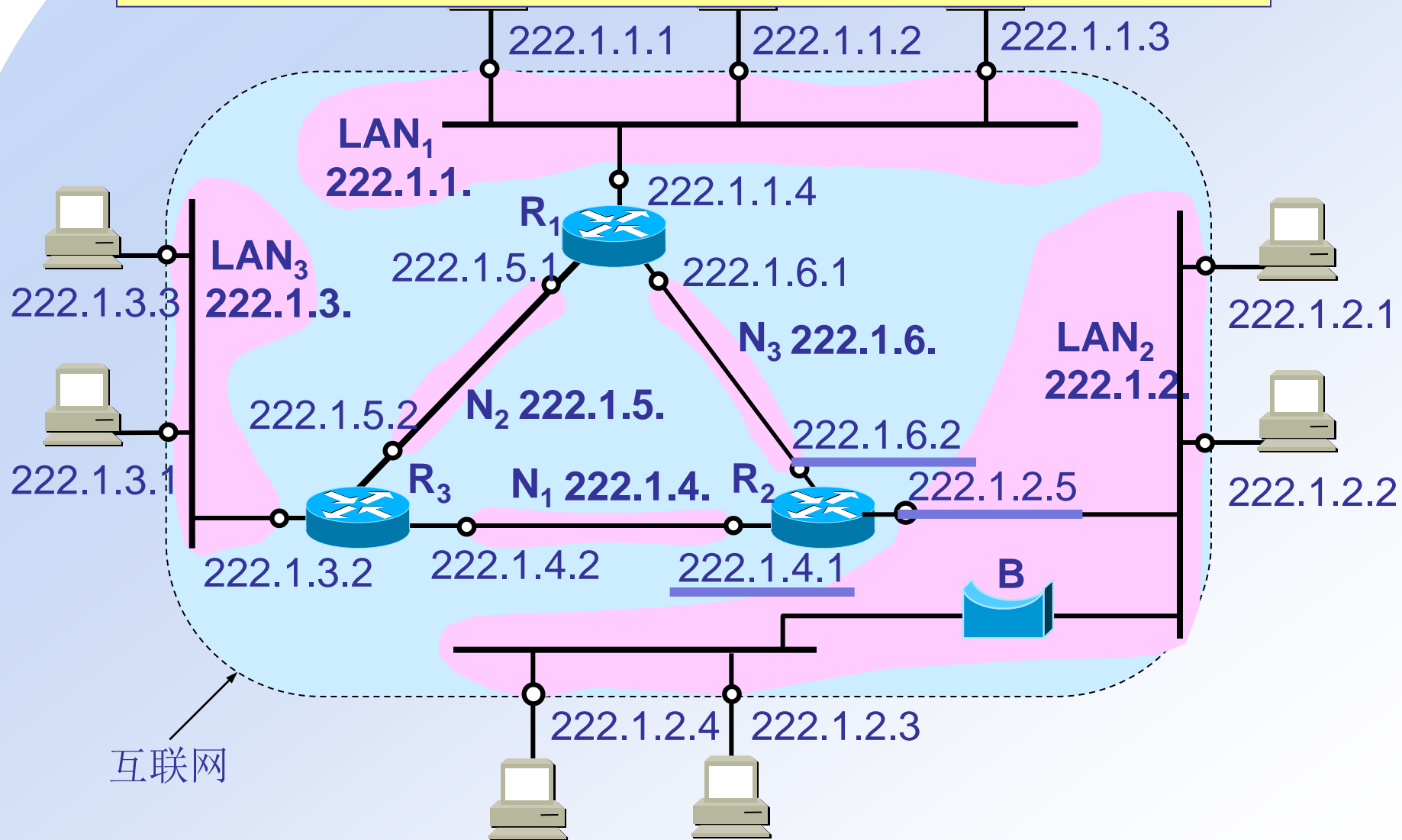
在同一个局域网上的主机或路由器的
IP 地址中的网络号必须是一样的。
图中的网络号就是 IP 地址中的 net-id



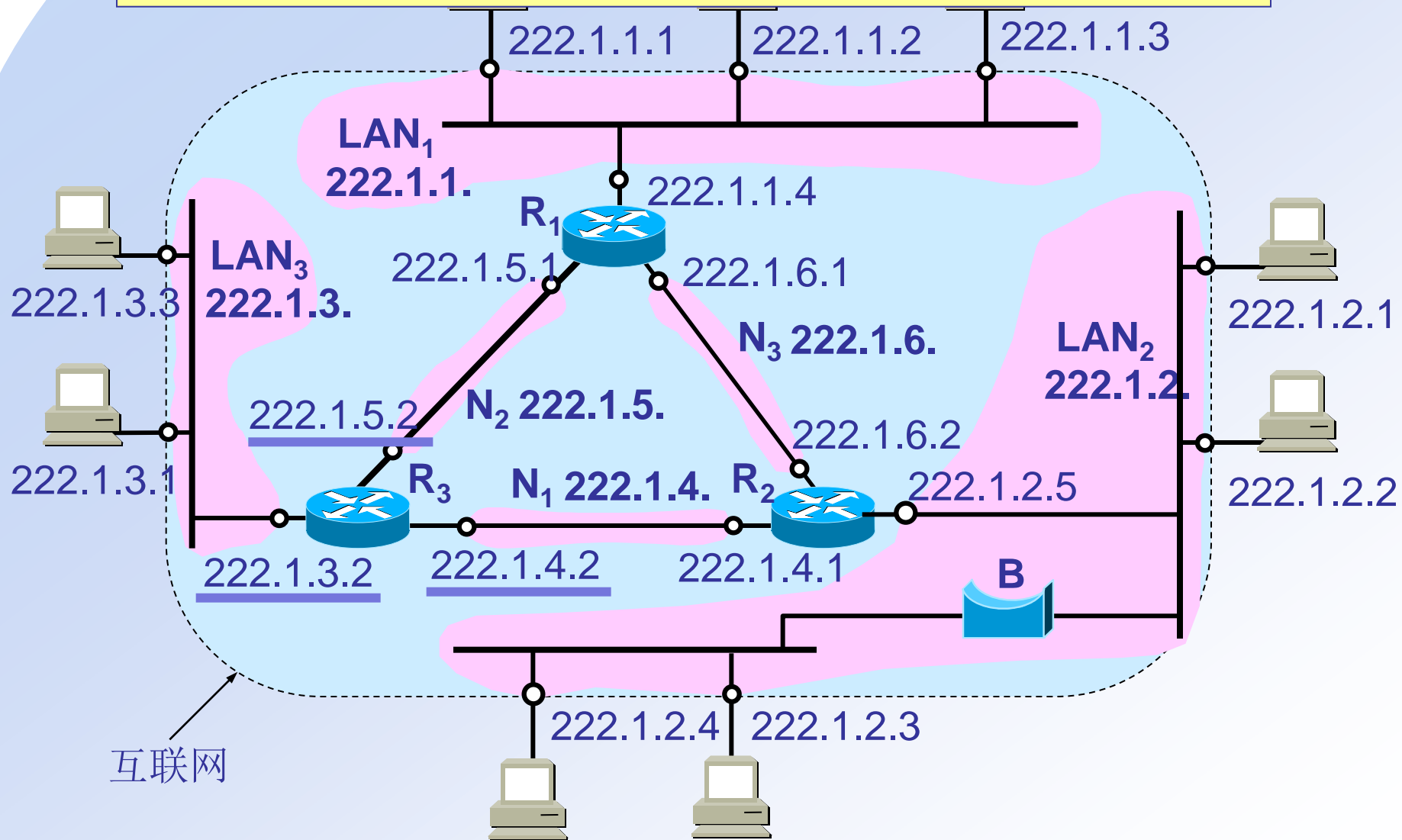
路由器总是具有两个或两个以上的 IP 地址。
路由器的每一个接口都有一个
不同网络号的 IP 地址。



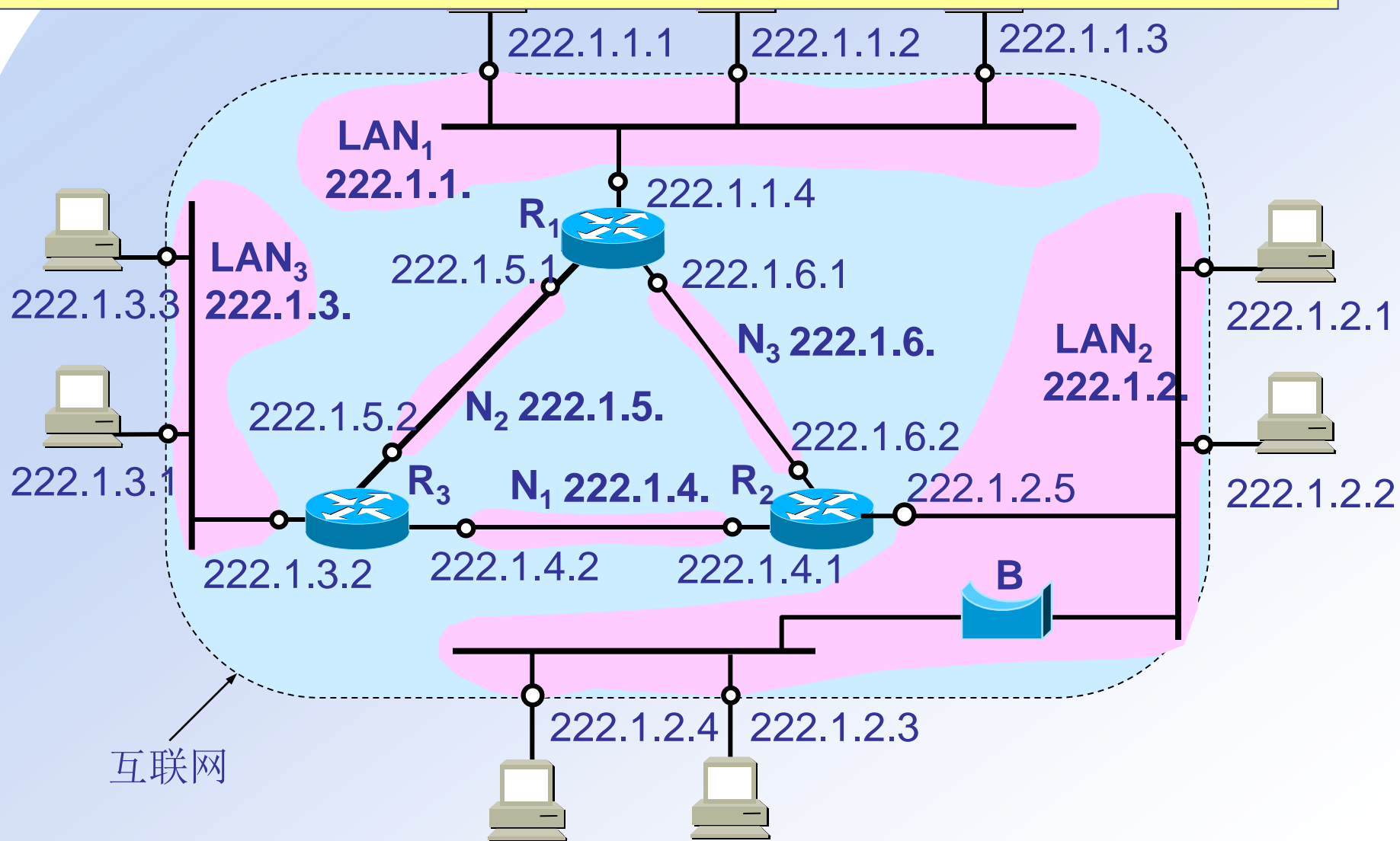
路由器总是具有两个或两个以上的 IP 地址。
路由器的每一个接口都有一个
不同网络号的 IP 地址。



路由器总是具有两个或两个以上的 IP 地址。
路由器的每一个接口都有一个
不同网络号的 IP 地址。

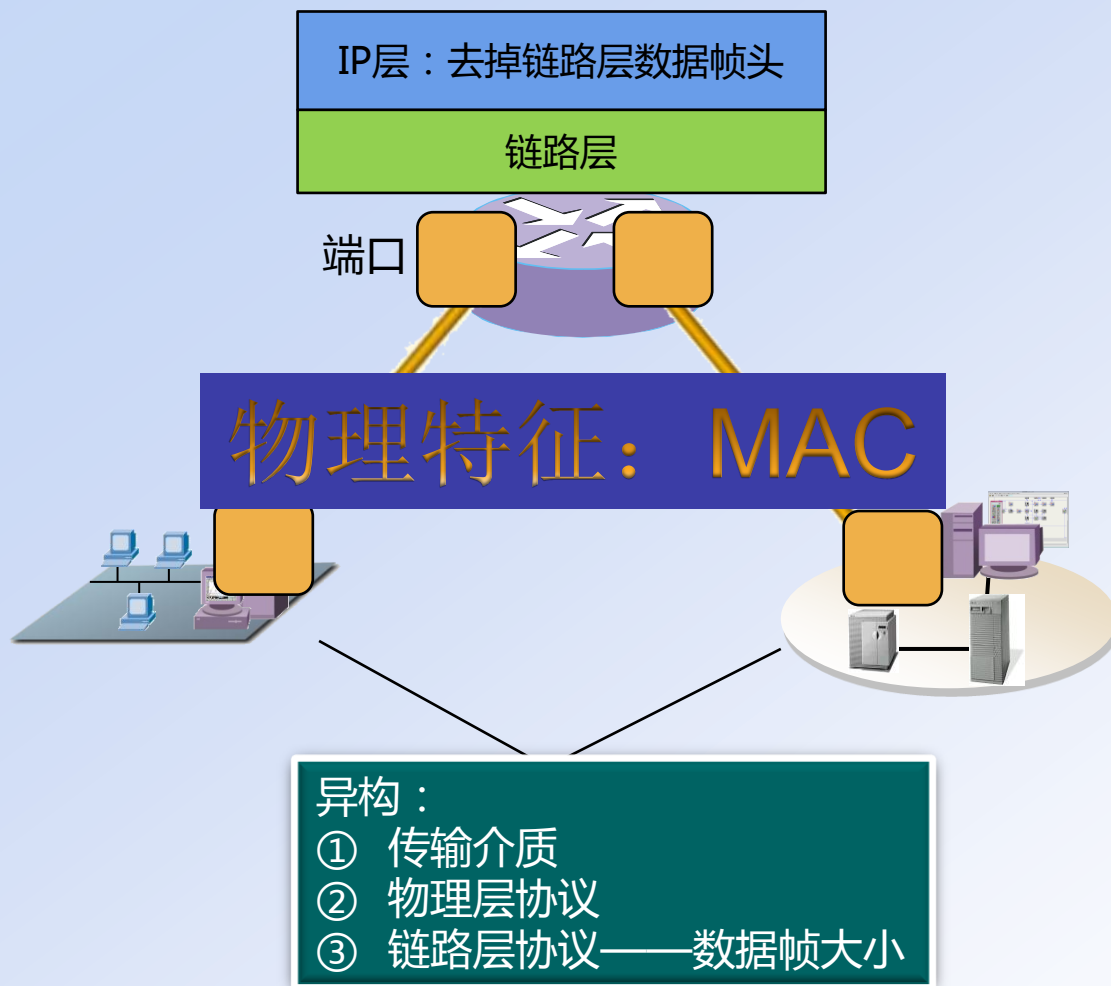


两个路由器直接相连的接口处，可指明也可不指明 IP 地址。如指明 IP 地址，则这一段连线就构成了一种只包含一段线路的特殊“网络”。现在常不指明 IP 地址。

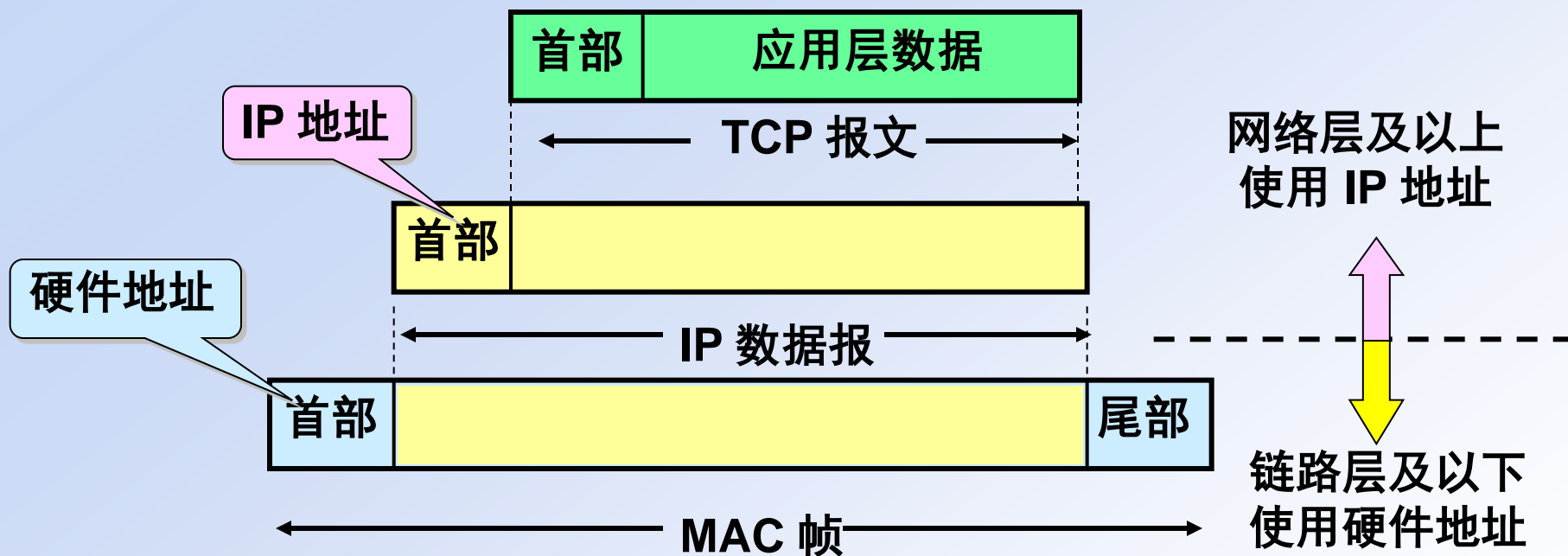


IP地址与硬件地址（MAC地址）

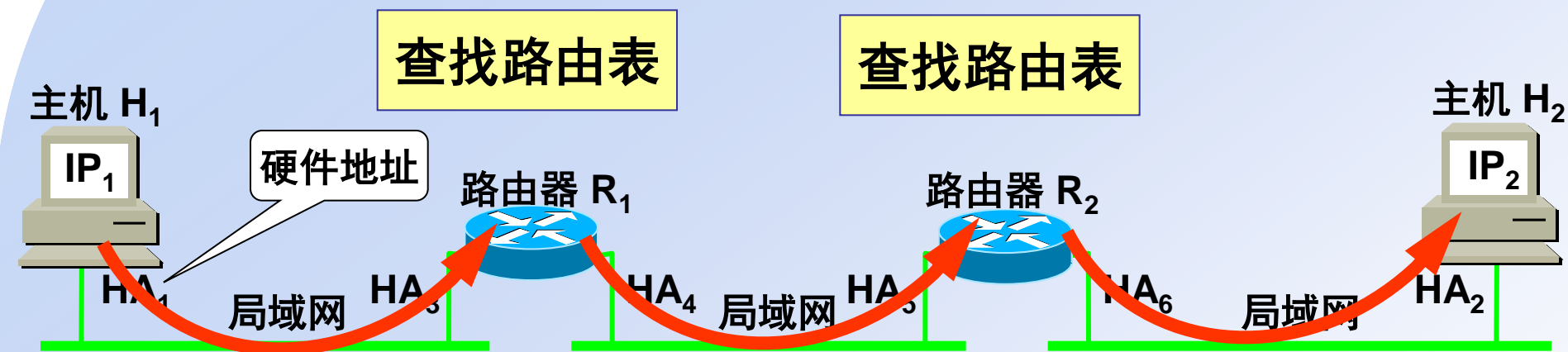
- ① 屏蔽局部物理细节，寻址：看到的是IP
- ② 去掉数据帧头、分组大小、IP地址包头符合规则



IP地址与MAC地址



IP地址与MAC地址

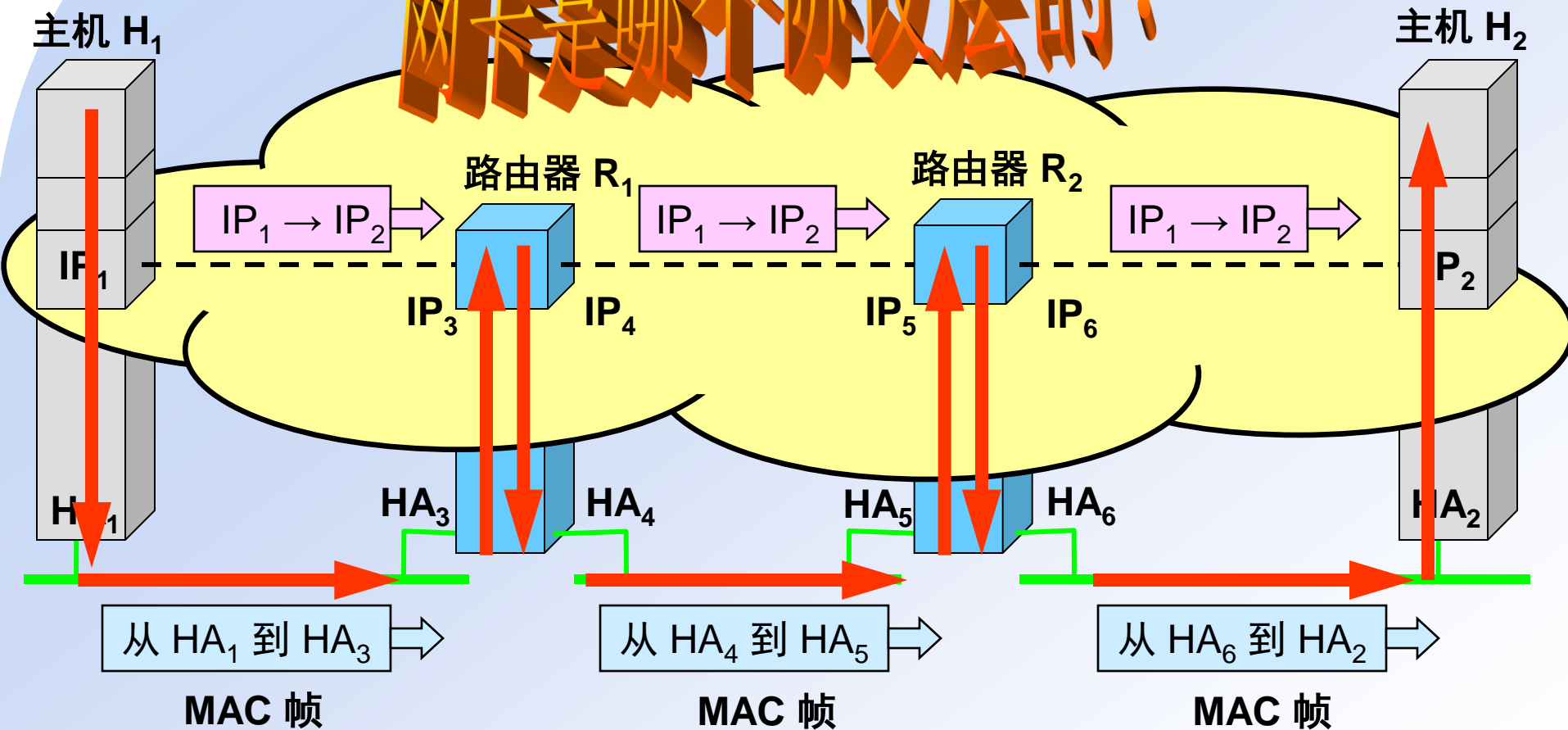


通信的路径：

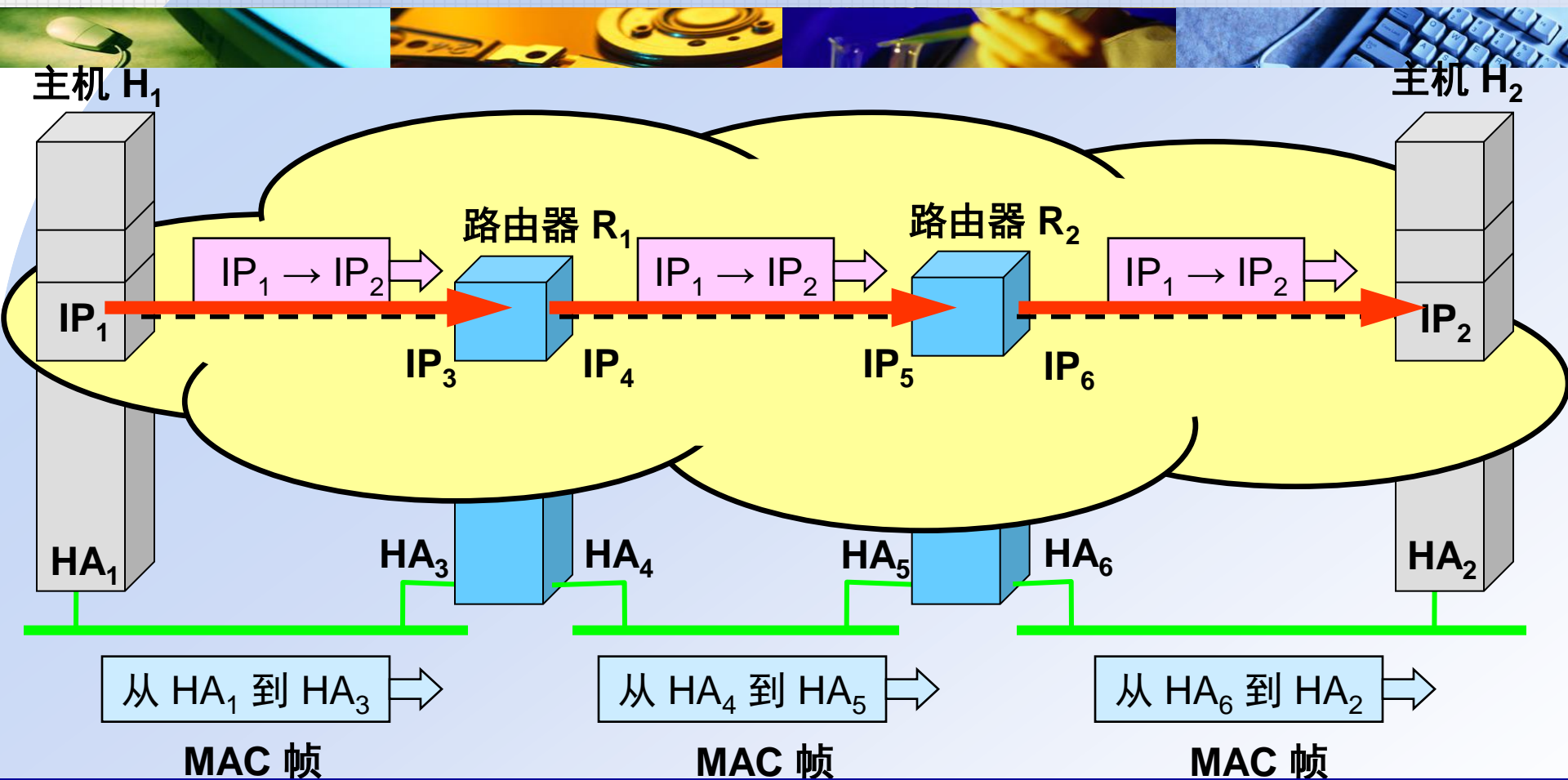
$H_1 \rightarrow$ 经过 R_1 转发 \rightarrow 再经过 R_2 转发 $\rightarrow H_2$

从协议栈的层次上看数据的流动

网卡是哪个协议层的?



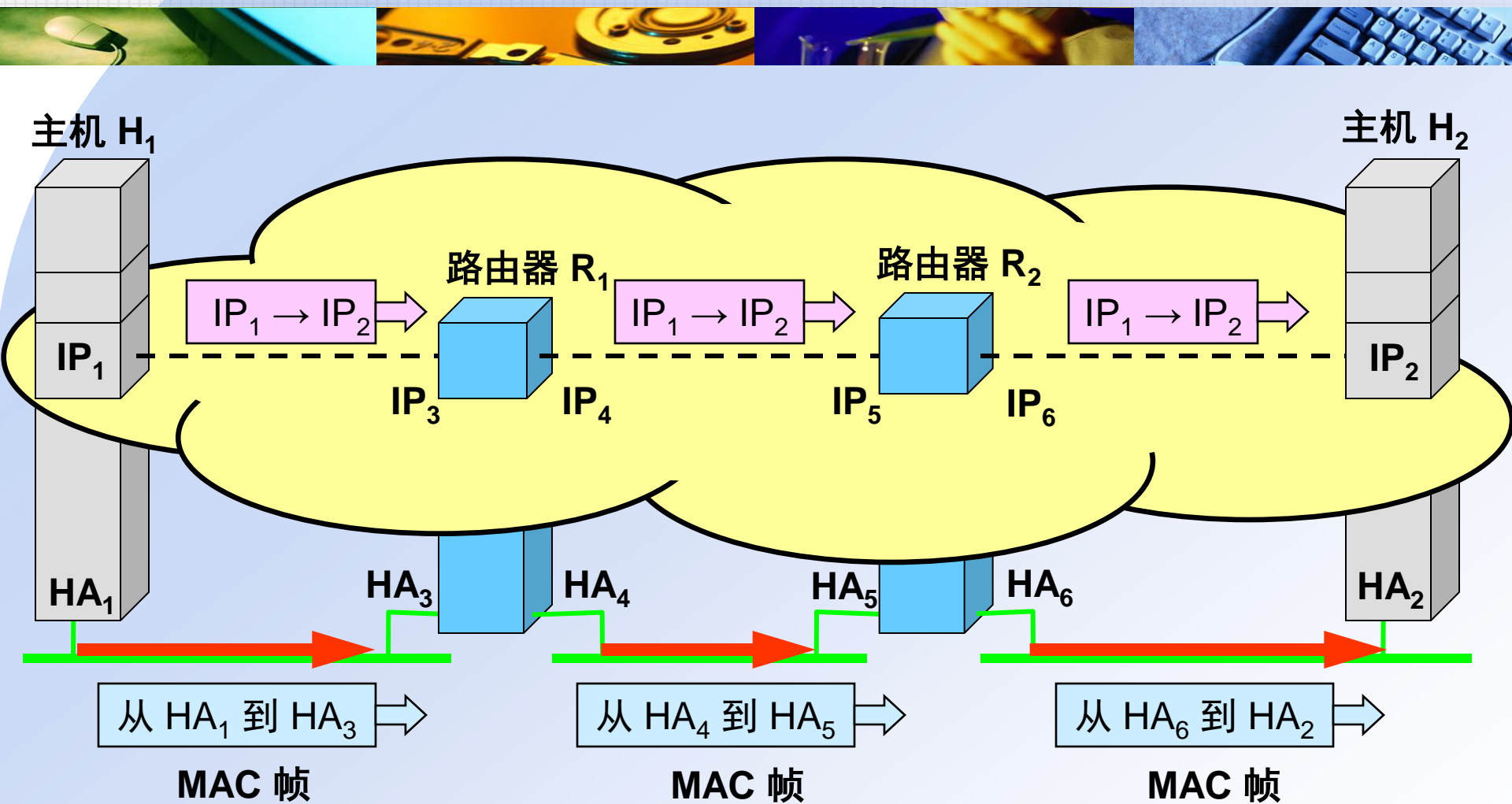
IP层看数据传输



在 IP 层抽象的互联网上只能看到 IP 数据报， $IP_1 \rightarrow IP_2$ 表示从源地址 IP_1 到目的地址 IP_2 ，两个路由器的 IP 地址并不出现在 IP 数据报的首部中。

路由器只根据目的站的 IP 地址的网络号进行路由选择

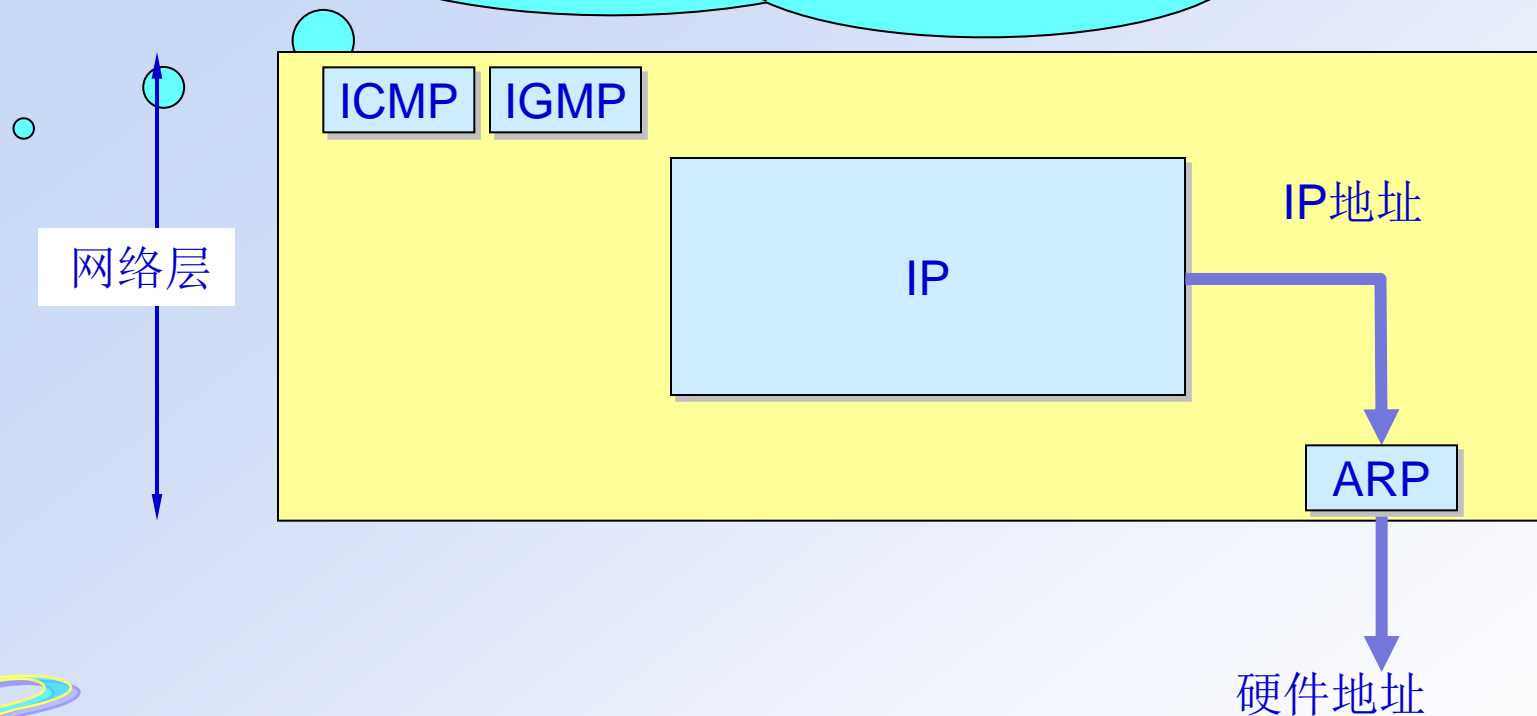
数据链路层看数据传输



在具体的物理网络的链路层只能看见 MAC 帧而看不见 IP 数据报。

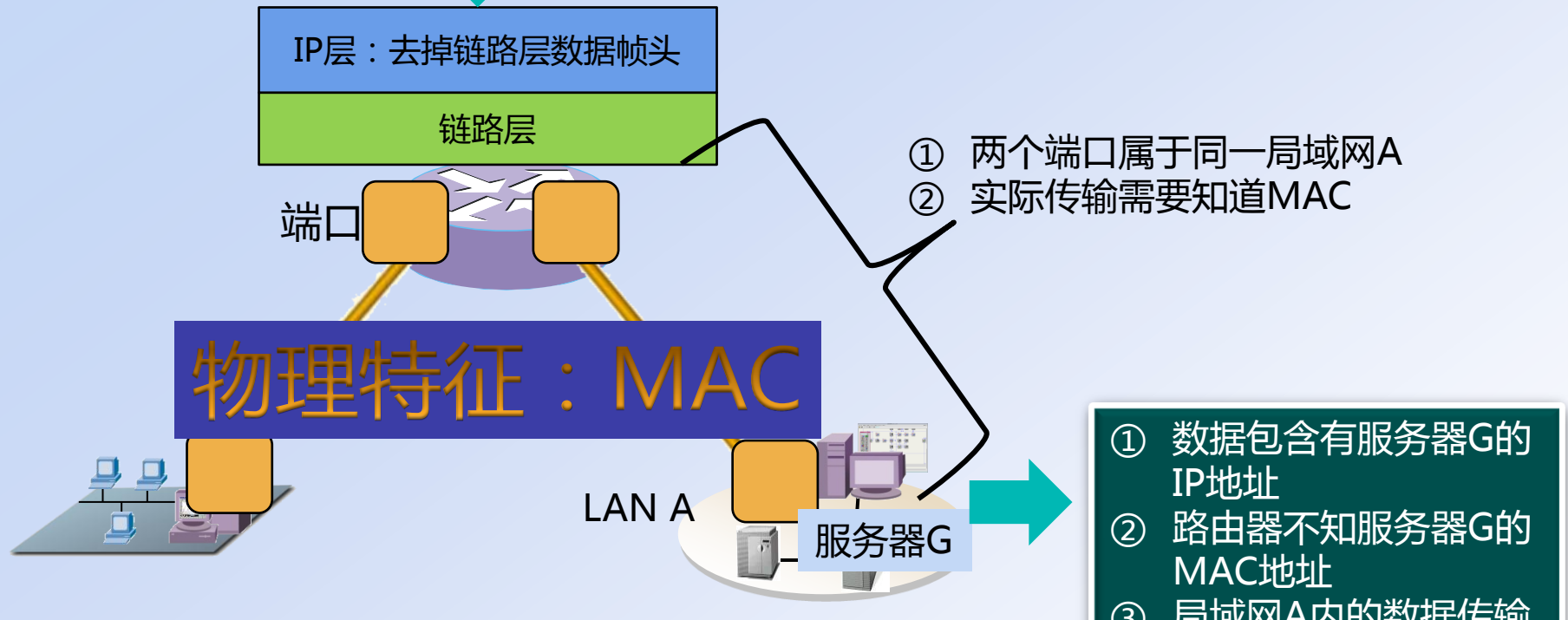
地址解析协议——ARP

主机或路由器怎样知道应当在MAC帧的首部填入什么样的硬件地址？



地址解析需求场景

数据包：寻址到局域网A



局域网A内需要有协议，将服务器G的IP地址转换为MAC地址

地址解析：局域网通信中IP地址的作用



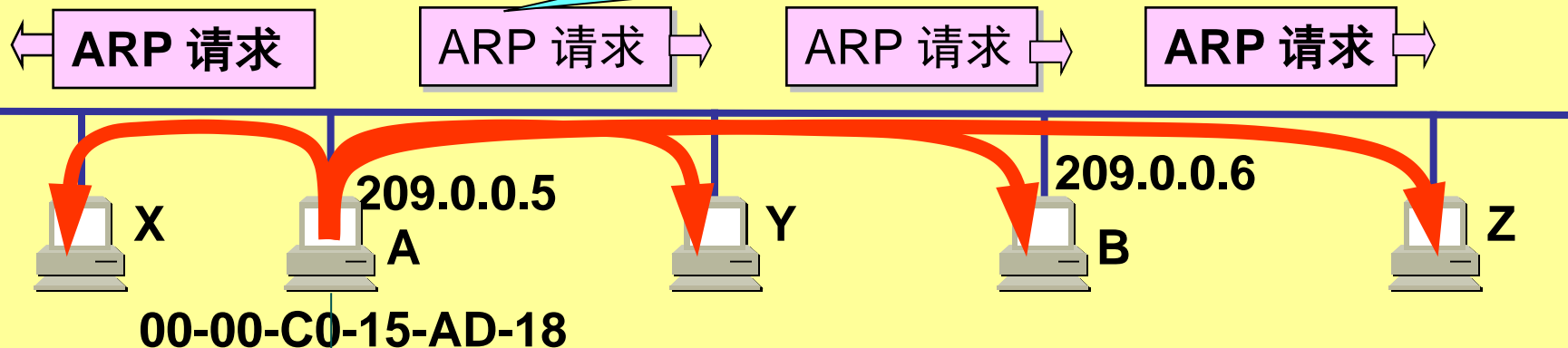
MAC地址获取和写入的时机

- ① 构造完IP首部后，明确源IP和目标IP——怎么获取？从应用层
- ② 查询ARP缓存，确定目标IP是否局域网地址
- ③ 是，就可查出其对应的硬件地址，再将此硬件地址写入 MAC 帧
- ④ 否，则查出路由器对内端口MAC地址，再将此硬件地址写入 MAC 帧
- ⑤ 通过局域网将该 MAC 帧发往目标硬件地址MAC

ARP协议的执行：如何知道其他主机的IP和MAC

主机 A 广播发送
ARP 请求分组

我是 209.0.0.5，硬件地址是 00-00-C0-15-AD-18
我想知道主机 209.0.0.6 的硬件地址



ARP协议使得每个接入网络的设备很“开朗热情”

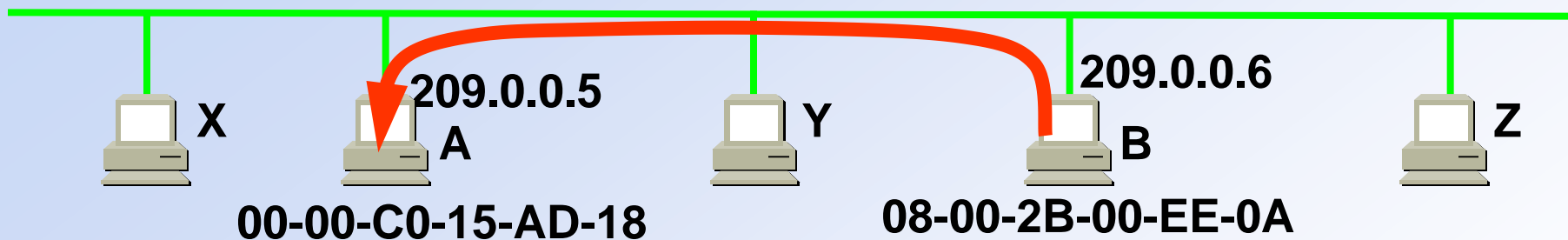
- ① 接入局域网时，通过DHCP或手工方式获取IP
- ② 将IP保存到ARP缓存
- ③ IP与自身MAC形成对应关系
- ④ ARP协议开始作用.....

ARP协议的执行：如何知道其他主机的IP和MAC

主机 B 向 A 发送
ARP 响应分组

我是 209.0.0.6
硬件地址是 08-00-2B-00-EE-0A

← ARP 响应



ARP协议的执行



- 为了减少网络上的通信量，主机 A 在发送其 ARP 请求分组时，就将自己的 IP 地址到硬件地址的映射写入 ARP 请求分组。
- 当主机 B 收到 A 的 ARP 请求分组时，就将主机 A 的这一地址映射写入主机 B 自己的 ARP 高速缓存中。这对主机 B 以后向 A 发送数据报时就更方便了。

ARP 总结

■ ARP 是解决

- 1) 同一局域网内

- 2) IP 地址和硬件地址的映射问题。

- 3) 若主机和源主机不在同一个局域网上，则

- ① 就要通过 ARP 找到一个本局域网上的某个路由器的硬件地址

- ② 然后把分组发送给这个路由器，让该路由器寻址

- ③ 如此持续——找到目标主机所在的网络

- 4) 地址解析是自动进行的（对用户透明）

- 只要有本地通信，ARP 协议就会自动地将该 IP 地址解析为链路层所需要的MAC地址。

为什么不直接使用硬件地址进行通信？



- 由于全世界存在着各式各样的网络，它们使用不同的硬件地址。要使这些异构网络能够互相通信就必须进行非常复杂的硬件地址转换工作，因此几乎是不可能的事。
- 连接到因特网的主机都拥有统一的 IP 地址，它们之间的通信就像连接在同一个网络上那样简单方便，因为调用 ARP 来寻找某个路由器或主机的硬件地址都是由计算机软件自动进行的，对用户来说是看不见这种调用过程的。

逻辑地址的优越性
网络接口的作用

主题 5



1 网络层概述

2 网际互连与网络设备

3 IP地址

4 地址解析协议

5 IP协议

6 ICMP协议

IP数据包的设计

- ❖ 网络互连传输数据面临的问题
- ❖ 数据报结构和格式
- ❖ 分组转发过程
- ❖ 通信示例
- ❖ 无连接协议

网际互连的问题

如何思考

系统

目标：满足需求

结构和运动

控制

目标：寻址，要关注上下层的需求

结构：上 - TCP；
下层：MAC包头

运动-行为：分片和组片

运动-状态：状态和时间、状态监督（质量指标）

主机

H₁



分组交换网

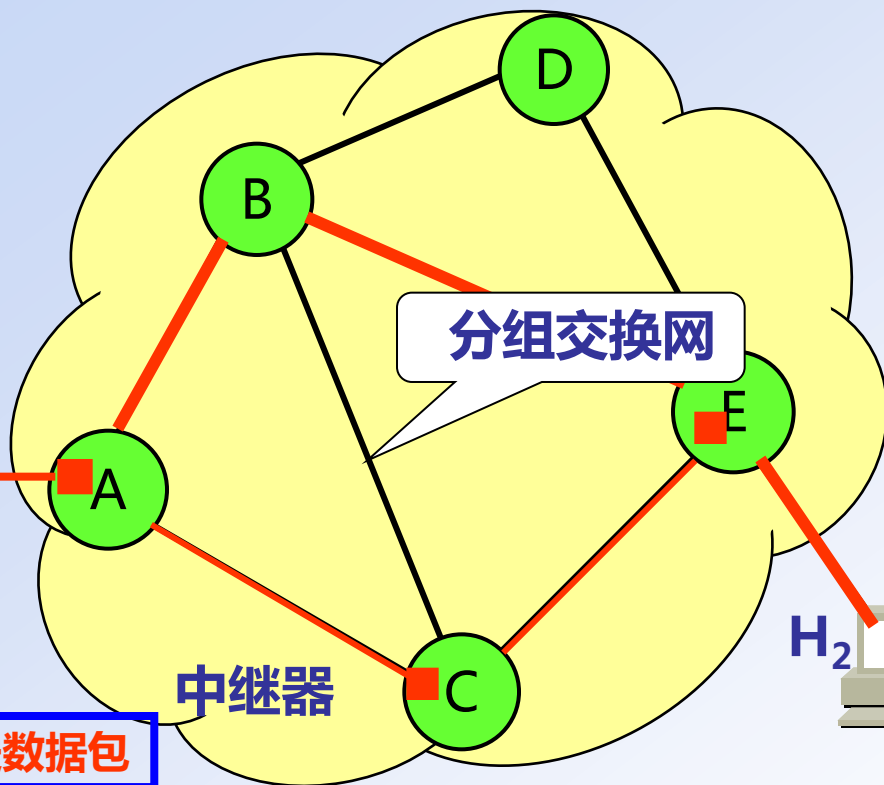
中继器

H₂



分组按照不同的路径到达目的地

H₁ 向 H₂ 发送数据包

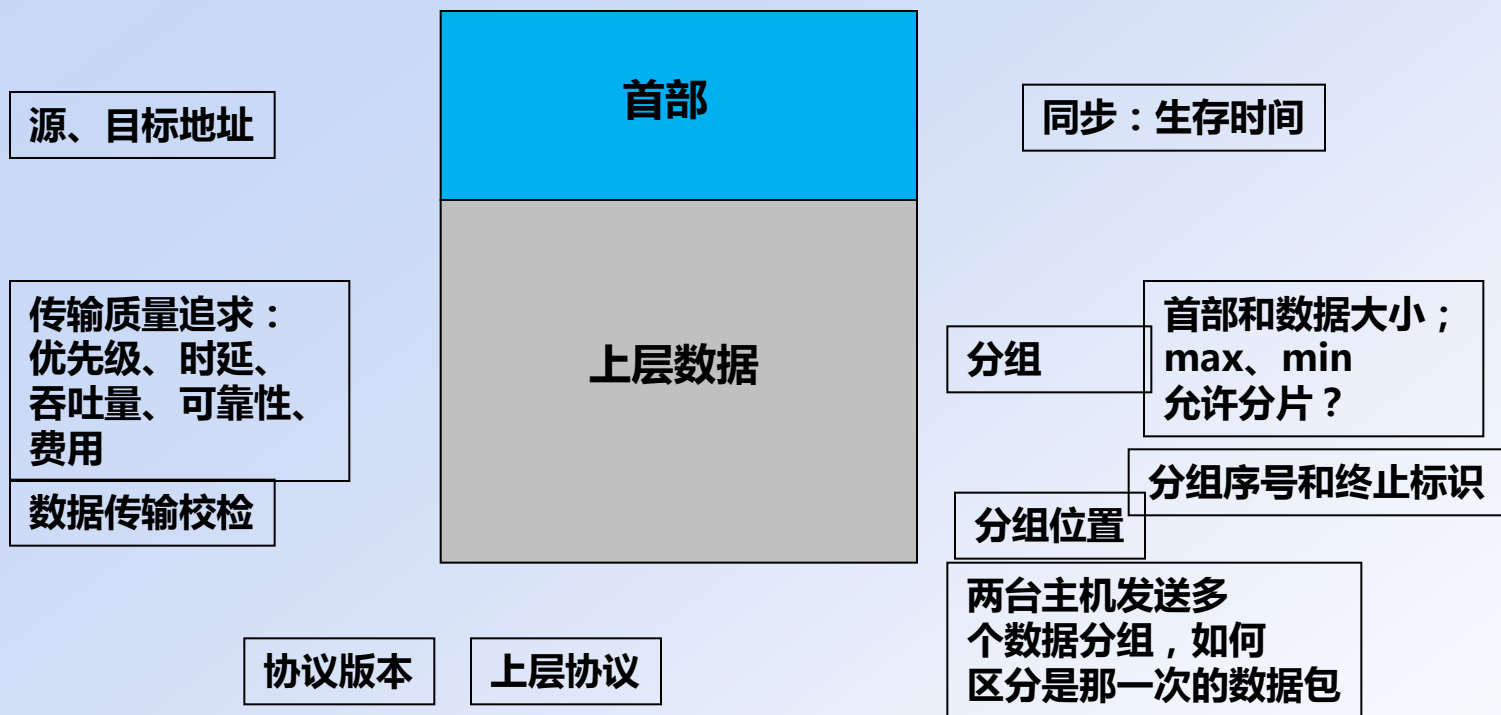


网际互连的问题

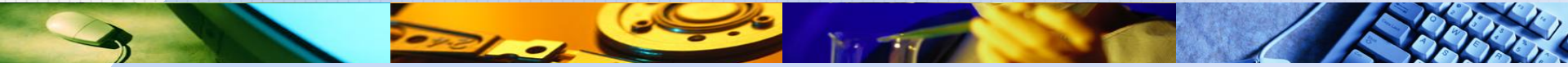


网络互连传输数据面临的问题

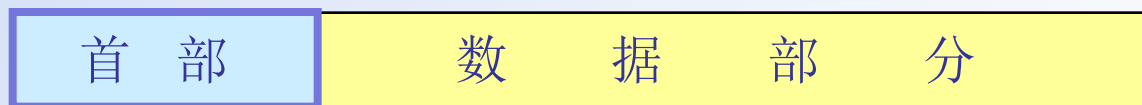
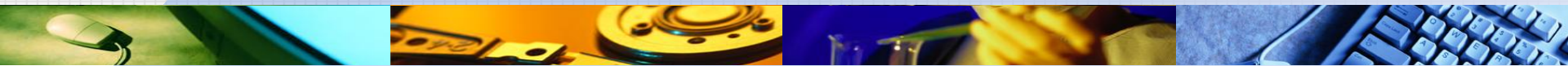
需要了解？



数据包结构

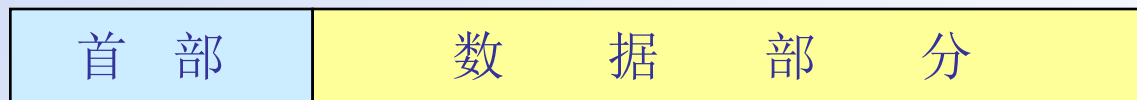
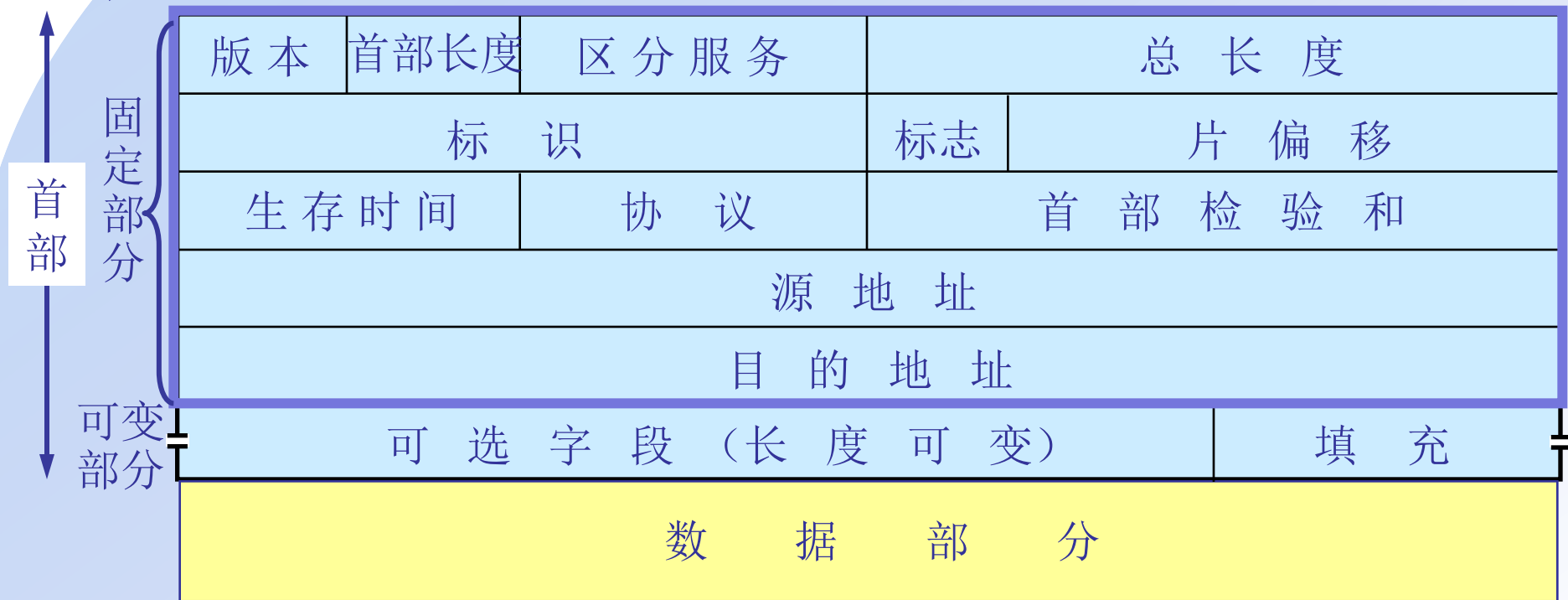
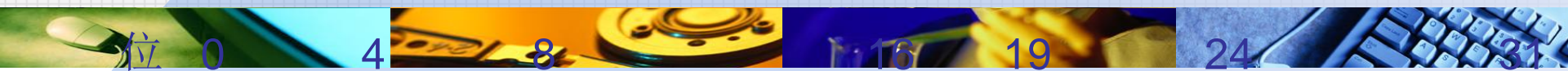


- 一个 IP 数据报由首部和数据两部分组成。
- 首部的前一部分是固定长度，共 20 字节，是所有 IP 数据报必须具有的。
- 在首部的固定部分的后面是一些可选字段，其长度是可变的。



IP 数据报

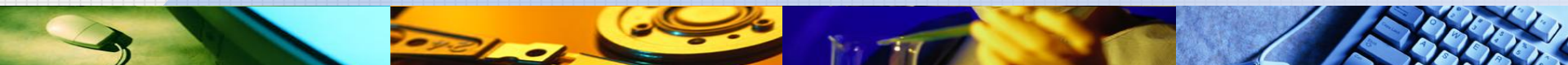
发送在前



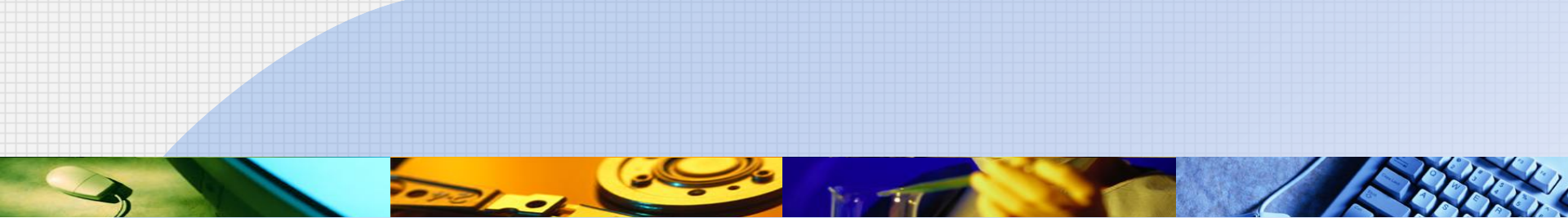
IP 数据报

发送在前

1. IP 数据报首部的固定部分中的各字段



版本——占 4 位，指 IP 协议的版本
目前的 IP 协议版本号为 4 (即 IPv4)



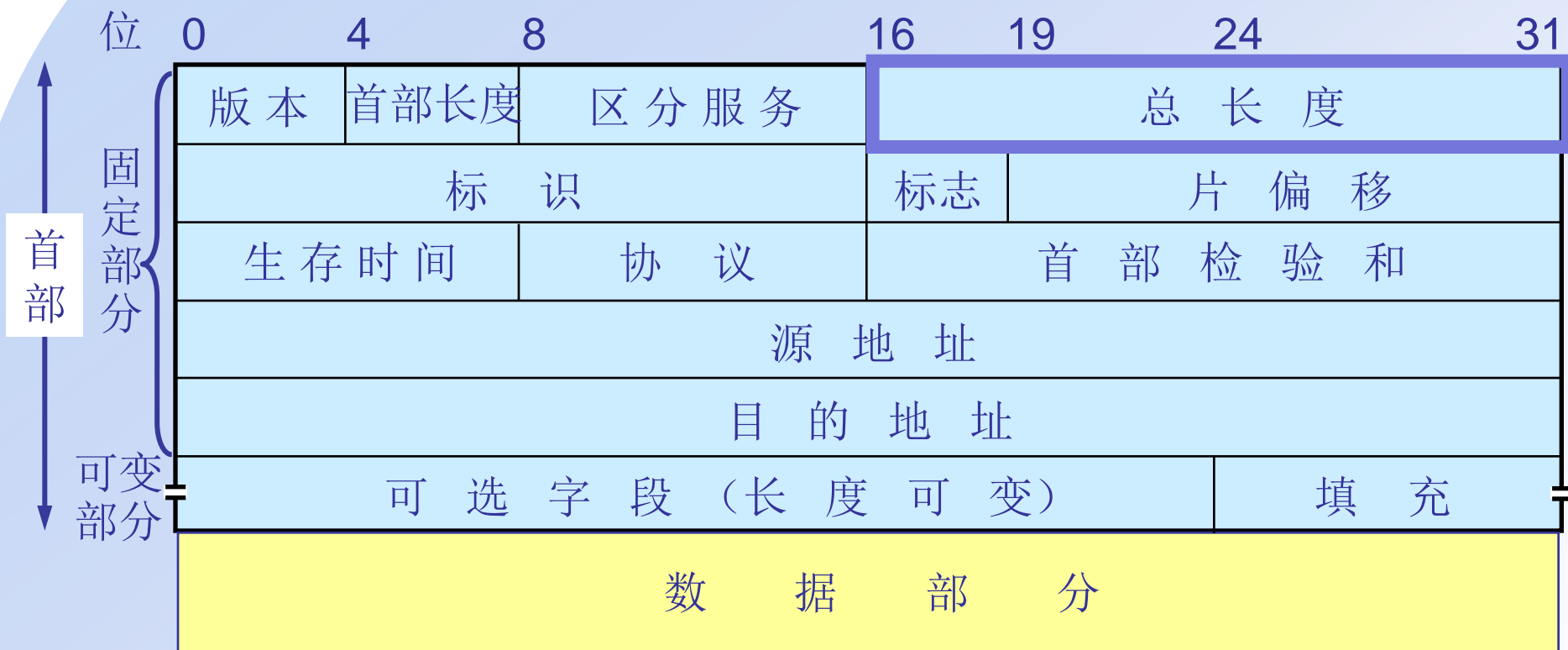
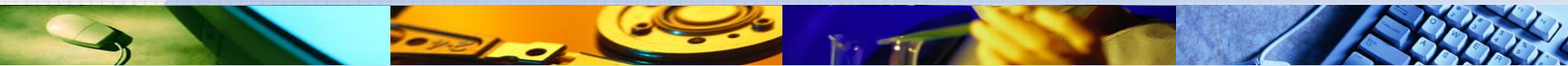
首部长度——占 4 位，可表示的最大数值是 15 个单位(一个单位为 4 字节)
因此 IP 的首部长度的最大值是 60 字节。



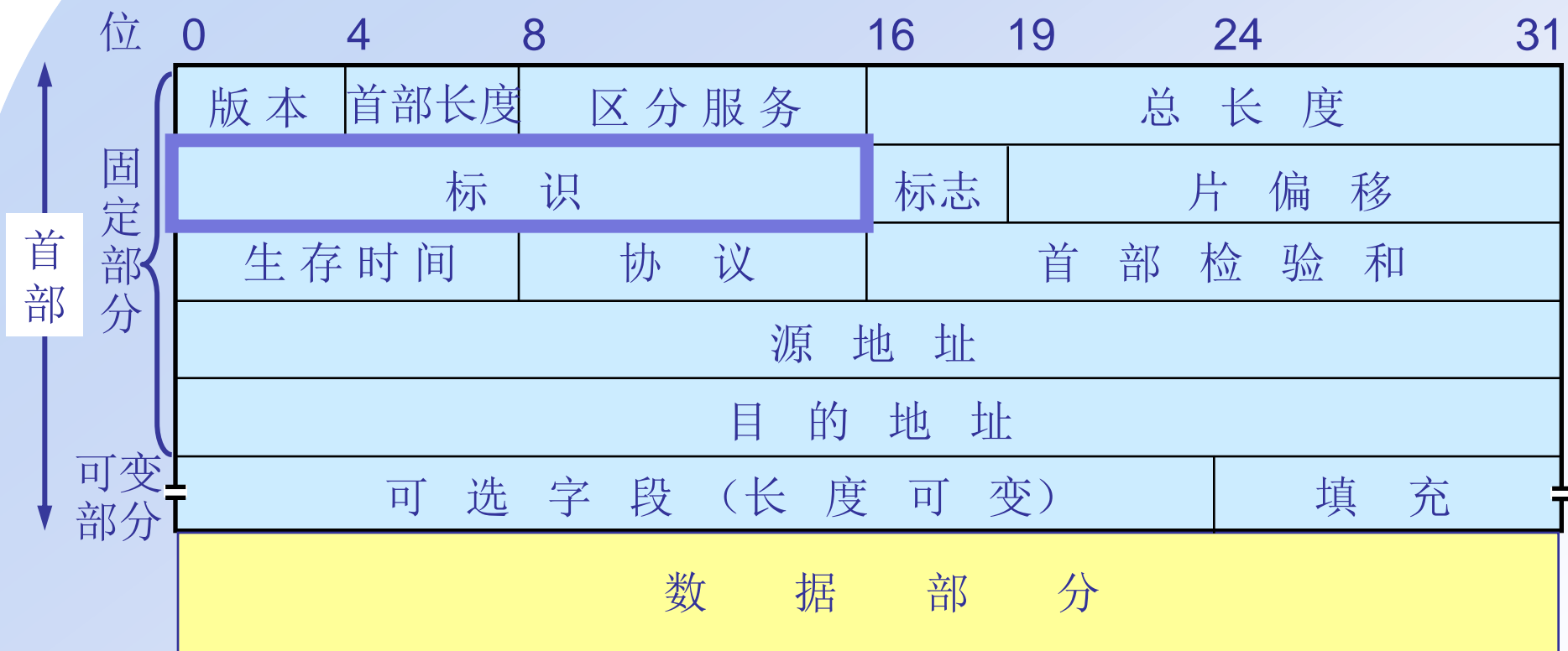
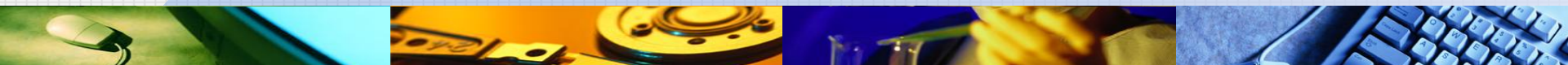
区分服务——占 8 位，用来获得更好的服务
在旧标准中叫做服务类型，但实际上一直未被使用过。

1998 年这个字段改名为区分服务。

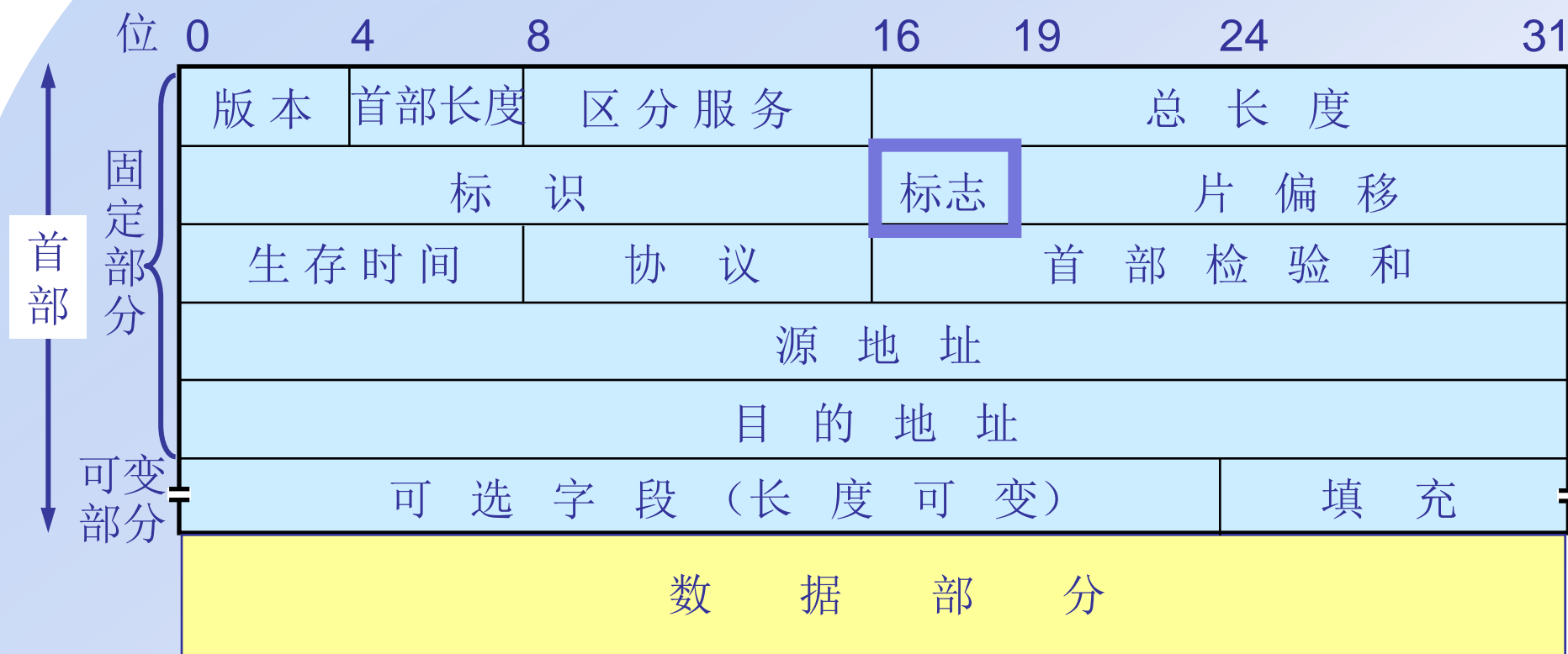
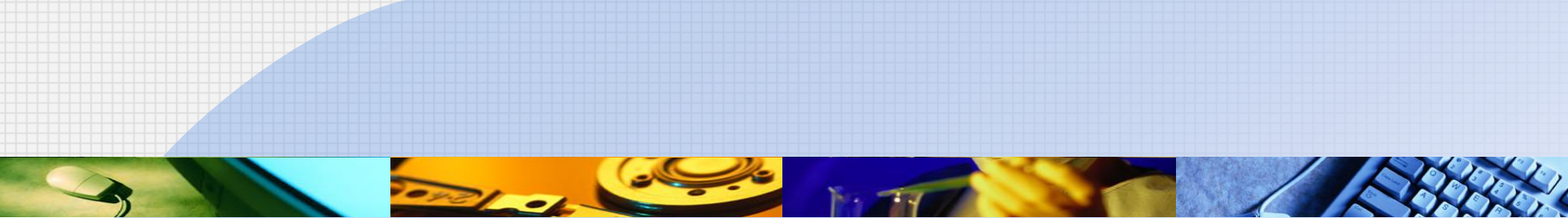
只有在使用区分服务（DiffServ）时，这个字段才起作用。
在一般的情况下都不使用这个字段



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 **65535** 字节。
总长度必须不超过最大传送单元 **MTU**。



标识(identification) 占 16 位，
信源主机在产生 IP 分组时，需要给每个 IP 分组分配一个唯一的标识符，
用来区分该主机发送的不同分组以及分片重组时属于哪个分组。



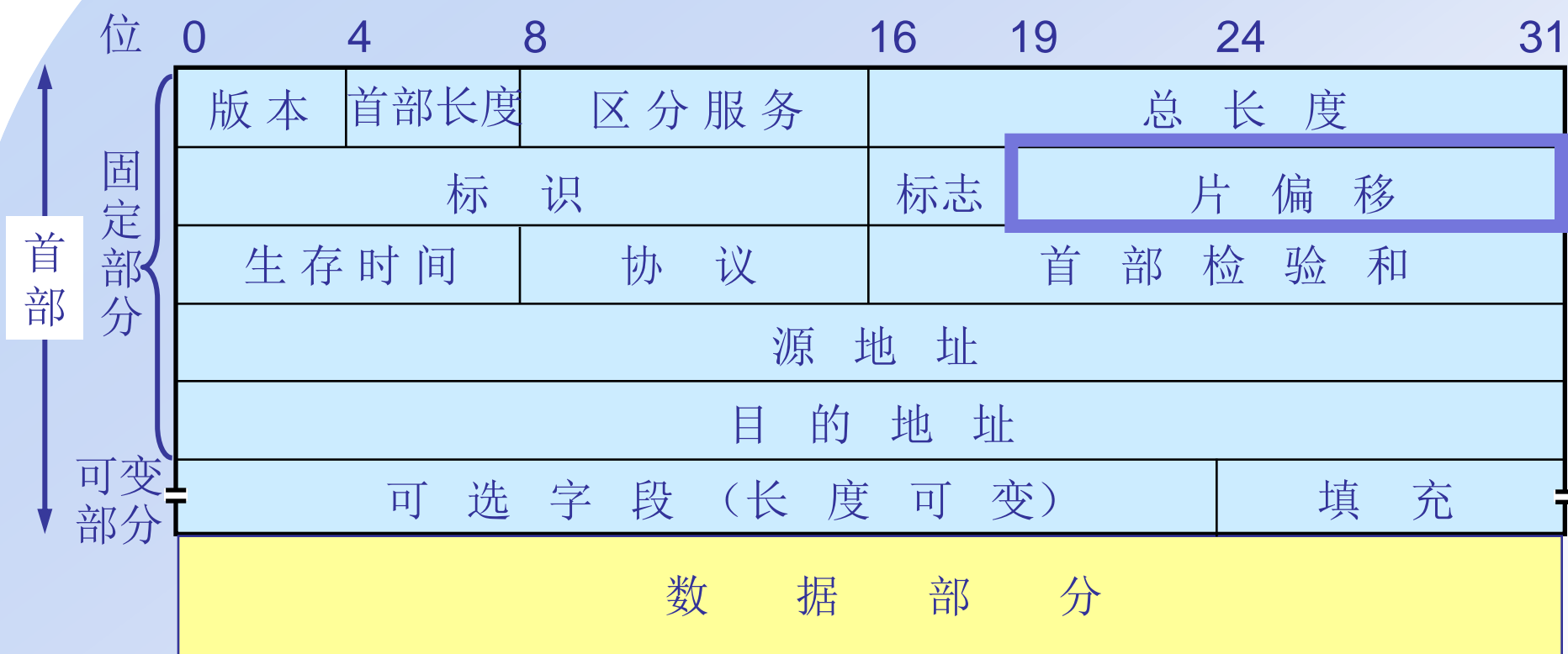
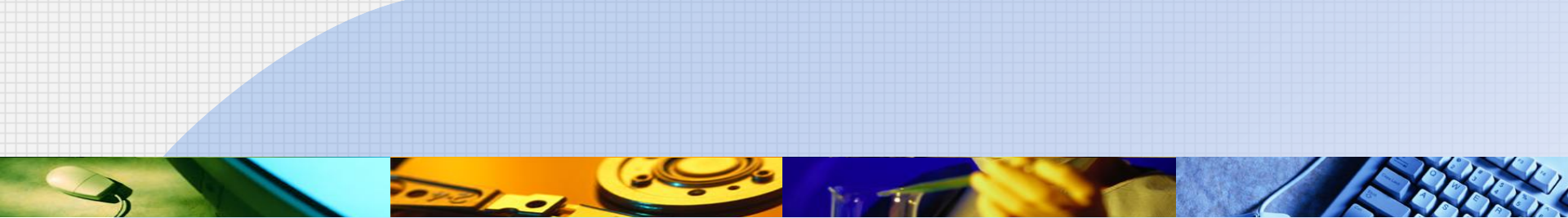
标志(flag) 占 3 位，目前只有前两位有意义。

标志字段的最低位是 **MF** (More Fragment)。

MF = 1 表示后面“还有分片”。**MF = 0** 表示最后一个分片。

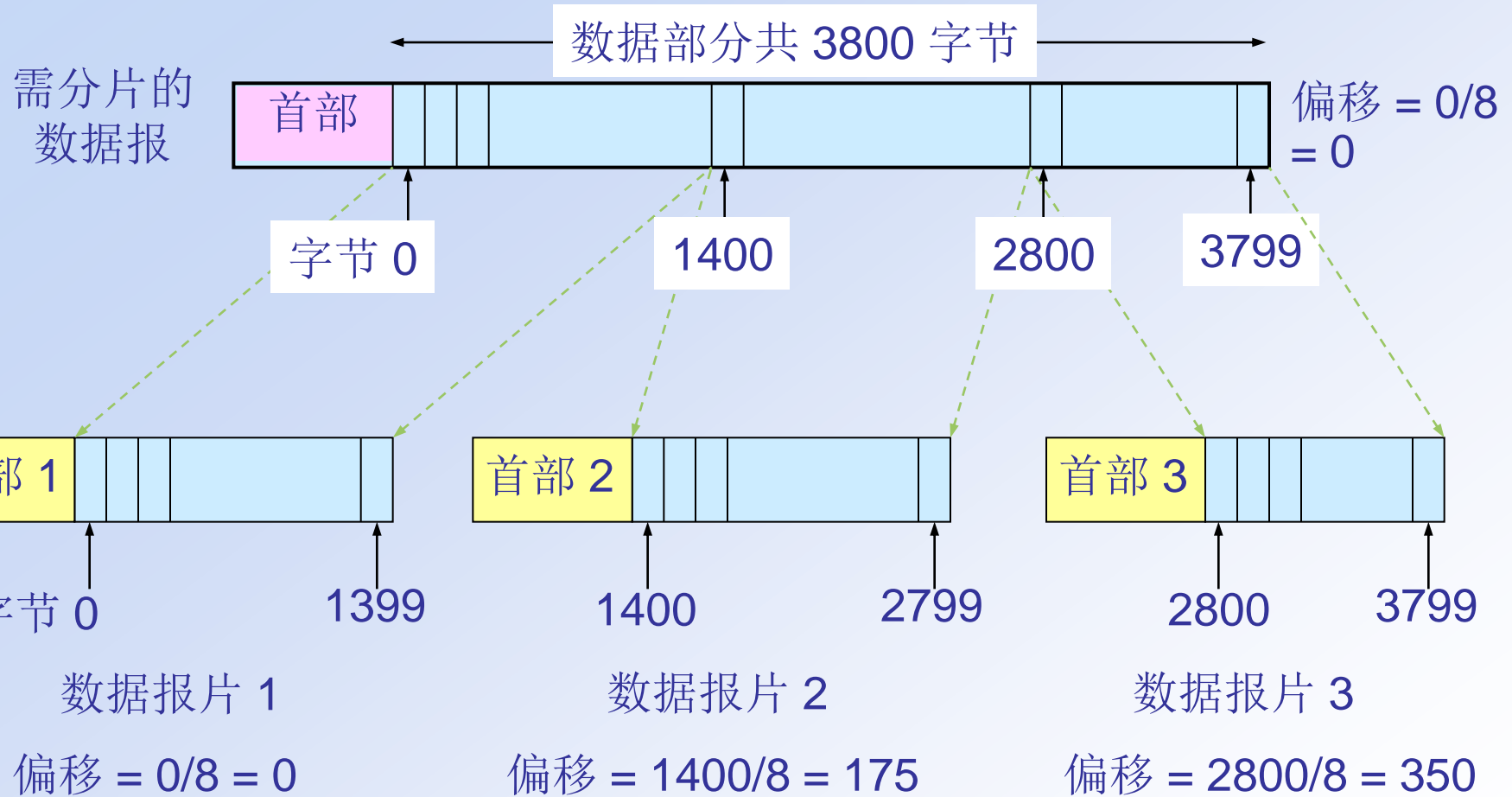
标志字段中间的一位是 **DF** (Don't Fragment) 。

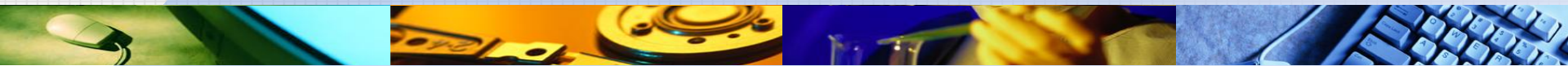
只有当 **DF = 0** 时才允许分片。



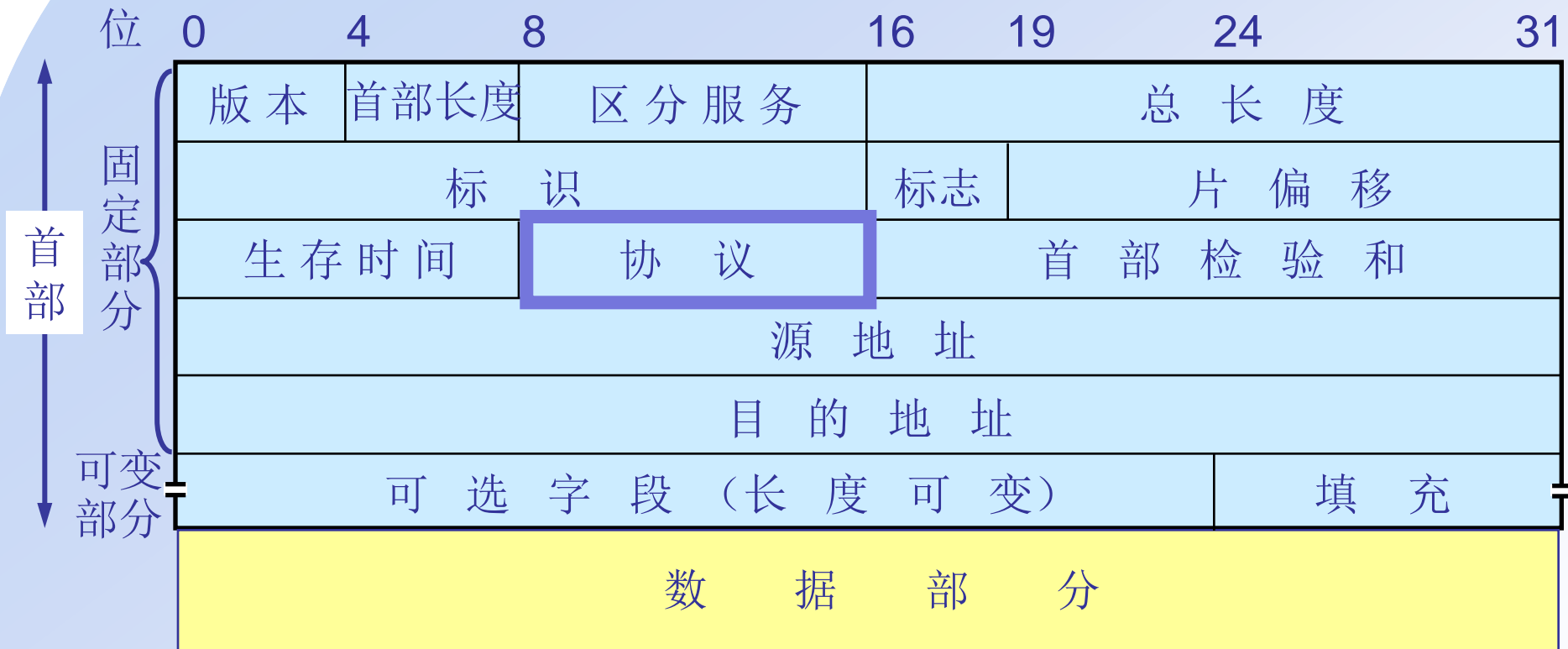
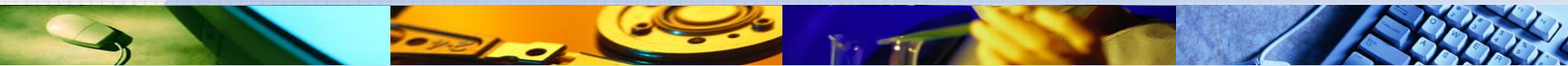
片偏移(13 位)指出：较长的分组在分片后
某片在原分组中的相对位置。
片偏移以 8 个字节为偏移单位。

IP 数据报分片示例

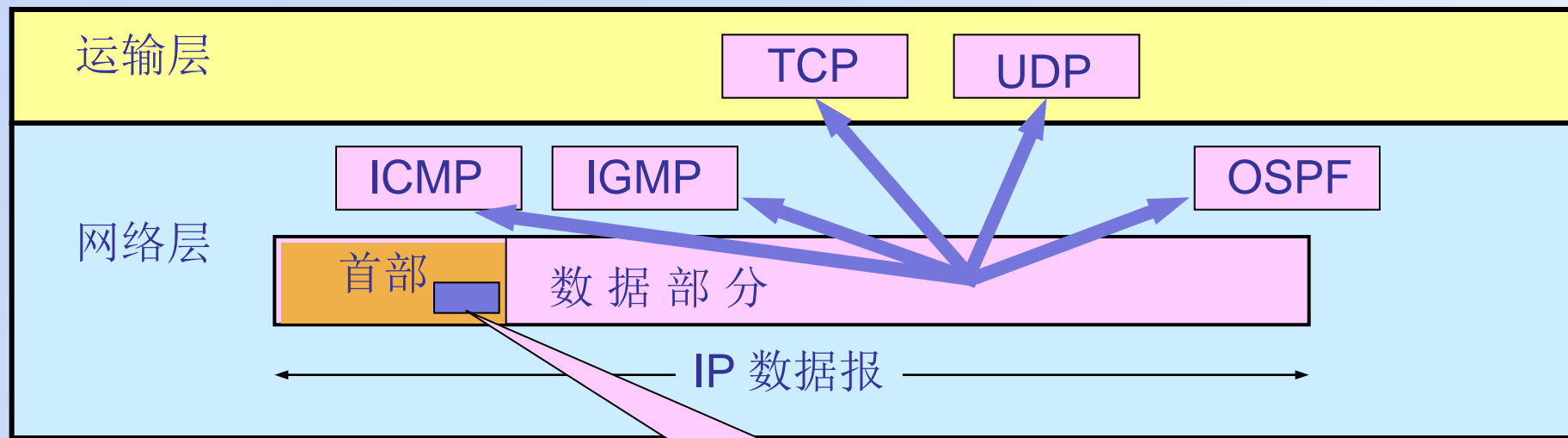
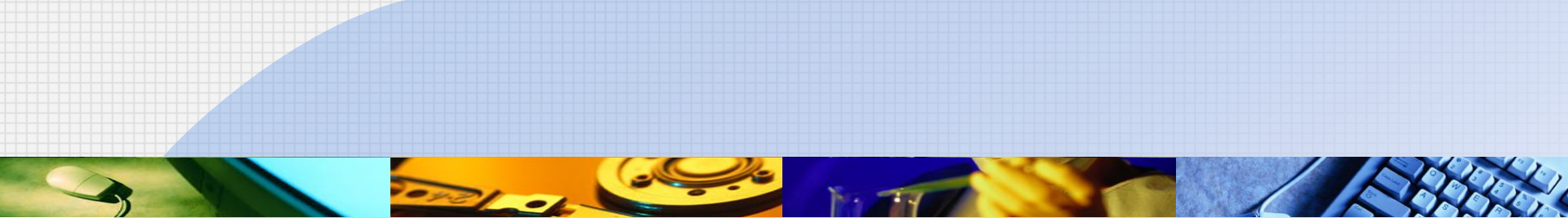




生存时间(8 位)记为 TTL (Time To Live)
数据报在网络中可通过的路由器数的最大值。



协议(8 位)字段指出此数据报携带的数据使用何种协议以便目的主机的 IP 层将数据部分上交给哪个处理过程



协议字段指出应将数据部分交给哪一个进程



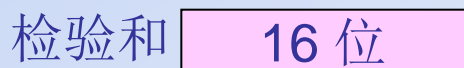
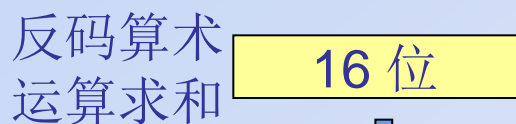
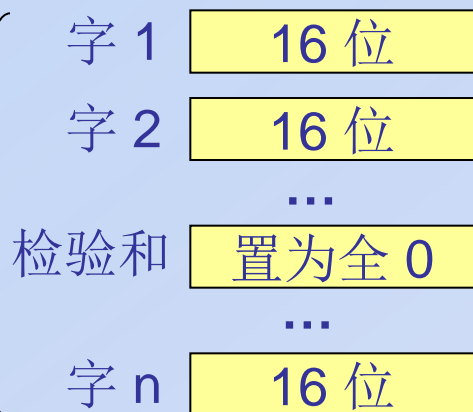
首部检验和(16 位)字段只检验数据报的首部
不检验数据部分。

这里不采用 **CRC** 检验码而采用简单的计算方法。

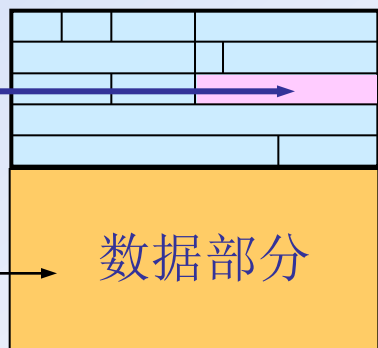
发送端

接收端

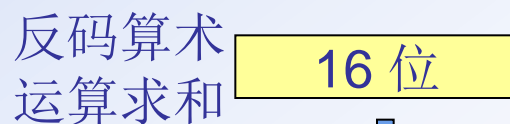
数据报首部



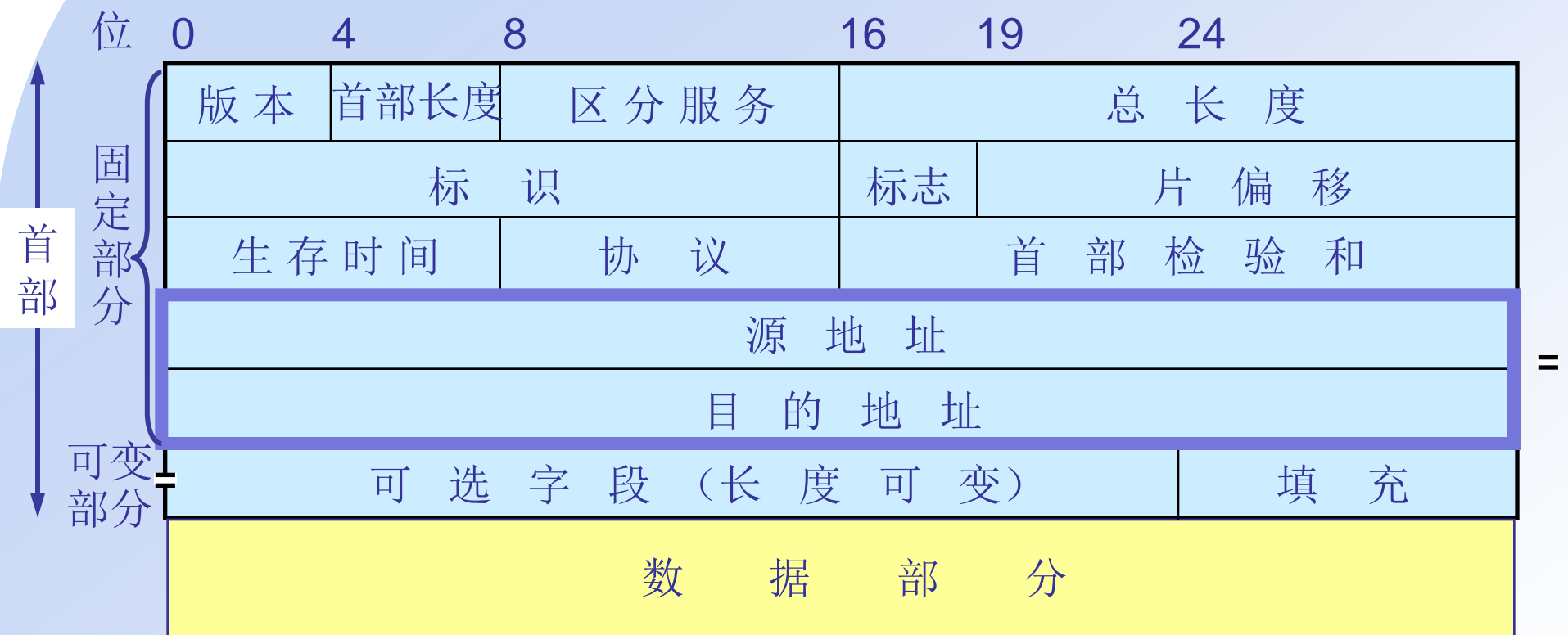
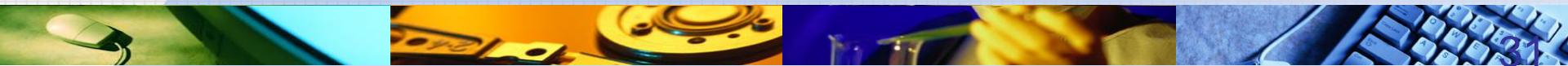
IP 数据报



数据部分
不参与检验和的计算



若结果为 0, 则保留;
否则, 丢弃该数据报



源地址和目的地址都各占 4 字节

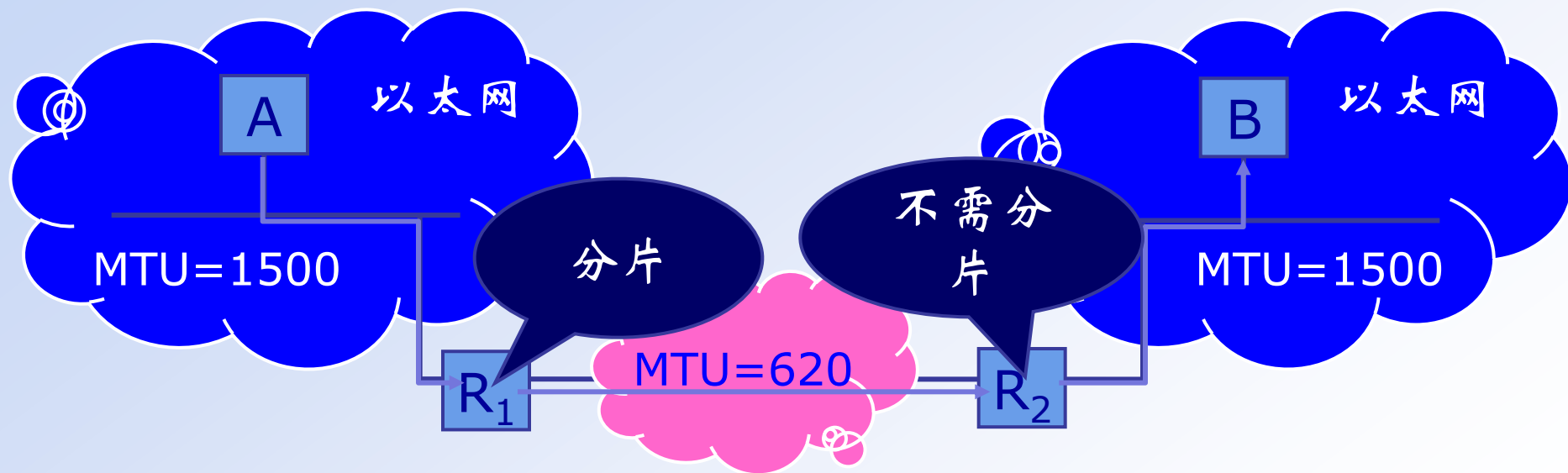
2. IP 数据报首部的可变部分

- ❖ IP 首部的可变部分就是一个选项字段，用来支持排错、测量以及安全等措施，内容很丰富。
- ❖ 选项字段的长度可变，从 1 个字节到 40 个字节不等，取决于所选择的项目。
- ❖ 增加首部的可变部分是为了增加 IP 数据报的功能，但这同时也使得 IP 数据报的首部长度成为可变的。这就增加了每一个路由器处理数据报的开销。
- ❖ 实际上这些选项很少被使用。

分组转发过程一分片

❖ 对分组进行分片的地点：

- 在分组的传输通路上，分片操作只能出现在两个 MTU 不同的网络的交界处，也就是出现在路由器上；
- 进入一个新网络时，若新网络的 MTU 小于原有网络的 MTU，则可能需要进行分片；若新 MTU 值不小于原有 MTU 就不必进行分片。



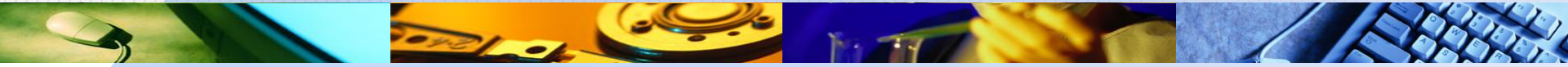
分组转发过程—组装



■ 重组过程

- 片重组 (reassembly) 是分片的逆过程, 所有的片重组操作都在目的主机上进行;
- 当目的主机收到 $MF=1$ 的分片时, 首先将其进行缓存;
- 当收到 $MF=0$ 的分片时, 计算原始 IP 分组的长度;
- 利用原始分组长度检查是否收到全部分片; 若是, 则按照各片的偏移值重新组装成 IP 分组, 提交给高层软件; 否则将等待, 直到收到该分组的全部分片;
- 为了防止无限等待, 在接收端设置重组定时器; 当接收到分组的第一片时启动该定时器, 如果在指定的时间内未能完成分组重组, 将放弃整个重组、释放资源.

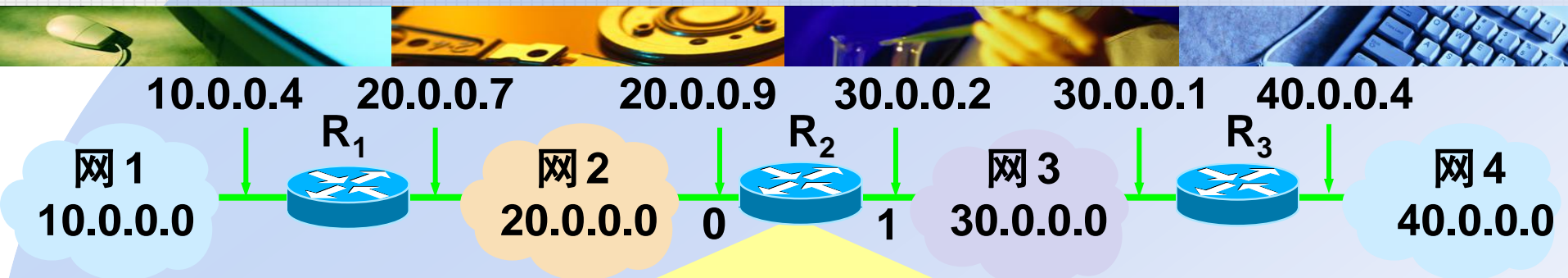
分组转发过程



路由器和结点交换机有些区别：

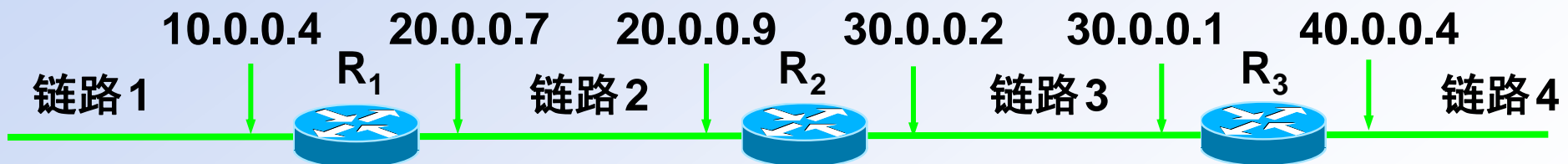
- 路由器是用来连接不同的网络，而结点交换机只是在一个特定的网络中工作。
- 路由器是专门用来转发分组的，而结点交换机还可接上许多个主机。
- 路由器使用统一的 IP 协议，而结点交换机使用所在网的特定协议。
- 路由器根据目的网络地址找出下一个路由器，而结点交换机则根据目的站所接入的交换机号找出下一跳（即下一个结点交换机）。

分组转发过程



路由器 R₂ 的路由表

目的主机所在的网络	下一跳路由器的地址
20.0.0.0	直接交付, 接口 0
30.0.0.0	直接交付, 接口 1
10.0.0.0	20.0.0.7
40.0.0.0	30.0.0.1



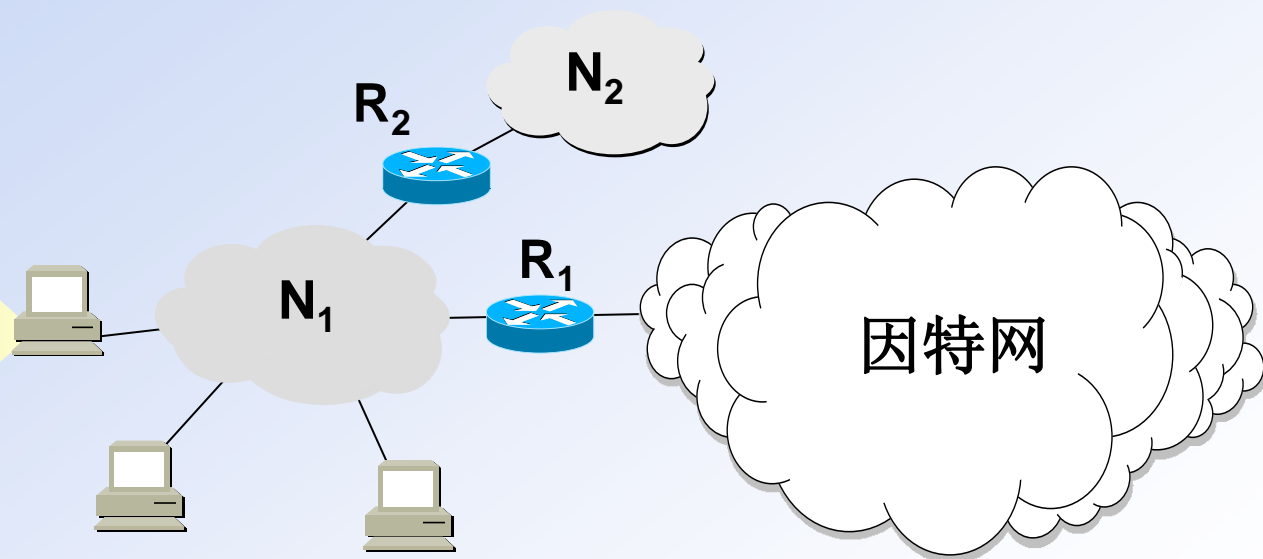
在路由表中, 对每一条路由, 最主要的是 (目的网络地址, 下一跳地址)

默认路由的使用

只要目的网络不是 N_1 和 N_2 ，
就一律选择默认路由，
把数据报先间接交付路由器 R_1 ，
让 R_1 再转发给下一个路由器。

路由表

目的网络	下一跳
N_1	直接
N_2	R_2
默认	R_1

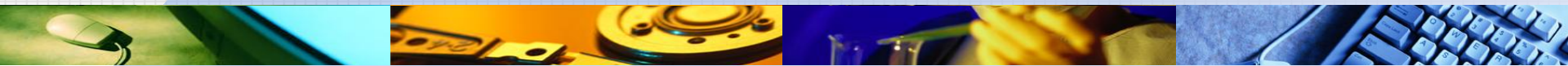


分组转发过程一算法



- (1) 从数据报的首部提取目的站的 IP 地址 D , 得出目的网络地址为 N 。
- (2) 若网络 N 与此路由器**直接相连**, 则直接将数据报交付给目的站 D ; 否则是间接交付, 执行(3)。
- (3) 若路由表中有目的地址为 D 的**特定主机路由**, 则将数据报传送给路由表中所指明的下一跳路由器; 否则, 执行(4)。
- (4) 若路由表中有**有到达网络 N 的路由**, 则将数据报传送给路由表指明的下一跳路由器; 否则, 执行(5)。
- (5) 若路由表中有一个**默认路由**, 则将数据报传送给路由表中所指明的默认路由器; 否则, 执行(6)。
- (6) 报告转发分组出错。

分组转发过程一算法



- IP 数据报的首部中没有地方可以用来指明下一跳路由器的 IP 地址”。

路由器IP地址不在数据包中

- 当路由器收到数据报，不是将下一跳路由器的 IP 地址填入IP数据报，而是送交下层的网络接口软件。

- 网络接口软件使用 ARP 负责将下一跳路由器的 IP 地址转换成硬件地址，并将此硬件地址放在链路层的 MAC 帧的首部，然后根据这个硬件地址找到下一跳路由器。

主题 6



1 网络层概述

2 网际互连与网络设备

3 IP地址

4 地址解析协议

5 IP协议

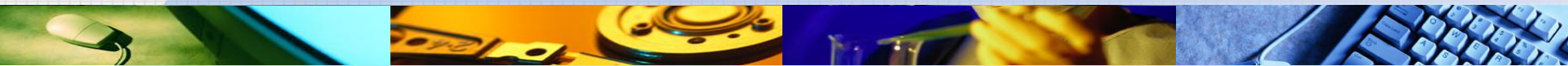
6 ICMP协议

ICMP



- ❖ 网际网控制报文协议 (Internet Control Message Protocol)
- ❖ IP层大致相当于ISO七层模型中的网络层，相对来说，IP 层控制功能最为复杂，它所提供的主要控制功能包括：
 - 差错控制
 - 拥塞控制
 - 路由控制
- ❖ 控制功能的实现包括信息收集和控制两部分。
- ❖ IP层控制功能的实现依赖于 ICMP 协议。

ICMP



- 为了提高 IP 数据报交付成功的机会，在网际层使用了因特网控制报文协议 ICMP (Internet Control Message Protocol)。
- ICMP 协议用于主机、路由器、网关之间交换网络层信息。
- ICMP 允许主机或路由器报告差错情况和提供有关异常情况的报告。
- ICMP 不是高层协议，而是 IP 层的协议。
- ICMP 报文作为 IP 层数据报的数据，加上数据报的首部，组成 IP 数据报发送出去。

ICMP具体功能

❖ 最初设计的 ICMP 协议主要用于差错的报告：

- 在 IP 网络中，各路由器独立完成分组的选路和传输，不需要信源主机参与。
- 出现传输错误时，IP 协议内部没有机制来获取差错信息并进行相应的控制。
- ICMP的设计使路由器发现传输错误时，可以通过 ICMP 协议向信源主机发送差错报告，由信源主机采取相应的差错处理措施。

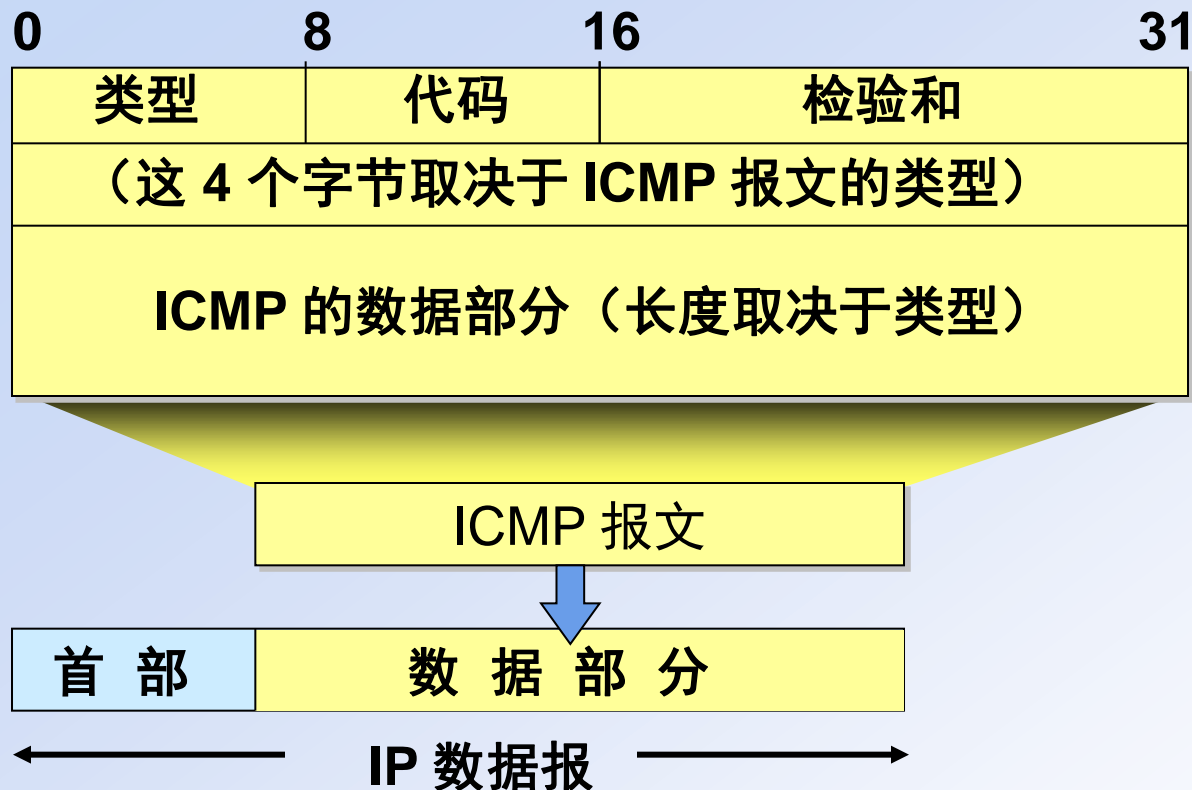
❖ ICMP 协议中又逐渐添加了拥塞控制、路由控制等功能。

❖ ICMP还提供一种双向的交互机制，网络节点可以用“请求+应答”的形式来获取网络中的某些有用的信息。以便进行故障诊断和网络控制。

❖ ICMP协议通过ICMP报文完成上述功能。

不是对数据内容自身差错的判断，那是TCP的工作

ICMP报文格式



- ICMP 报文的种类有两种，**差错报告**报文和**询问报文**。
- ICMP 报文的前 4 个字节是统一的格式，共有三个字段：即类型、代码和检验和。接着的 4 个字节的内容与 ICMP 的类型有关。

ICMP 报文的类型



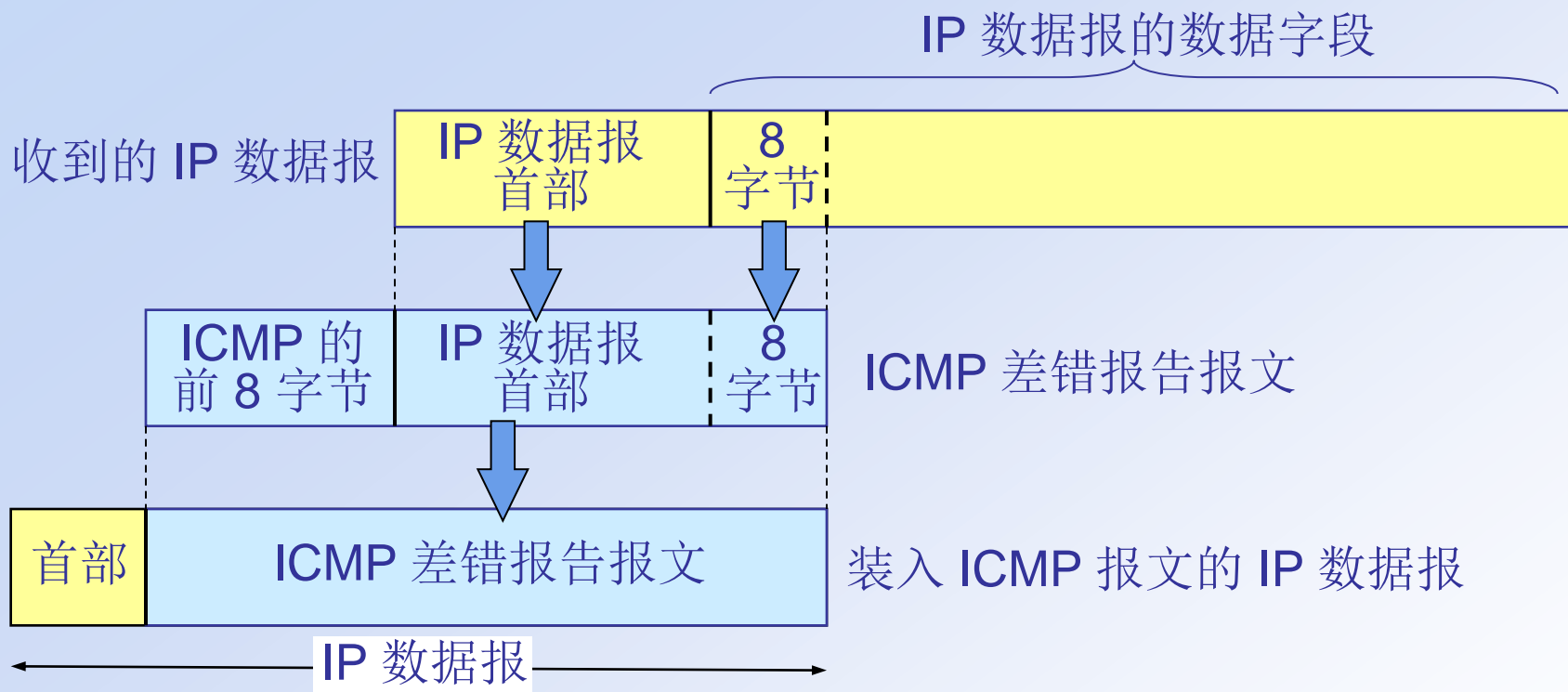
ICMP报文种类	类型值	ICMP报文的类型
差错报告报文	3	终点不可达
	4	源点抑制
	11	时间超时
	12	参数问题
	5	改变路由
询问报文	8或0	回送请求或回答
	13或14	时间戳请求或回答

ICMP—差错报告报文（5种）



- **终点不可达**：网络不可达、主机不可达、协议不可达、端口不可达、需要分片但DF已置为1，源路由失败等六种情况。
- **源站抑制**：网络拥塞时，主机/路由器丢弃数据报，向源站发送抑制报文，通知源站放慢发送速度。
- **时间超过**：路由器收到TTL为0的数据报时，向源站发送该报文。
- **参数问题**：当路由器/主机收到的报文的首部字段有错时，就向源站发送参数问题报文。
- **改变路由（重定向）**：路由器将改变路由报文发送给主机，让主机知道下次应将数据报发送给另外的路由器。

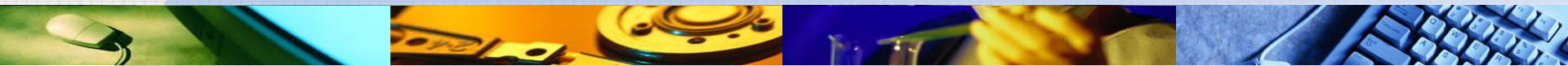
ICMP 差错报告报文的数据字段的内容



不应发送 ICMP 差错报告报文的几种情况

- ❖ 对 ICMP 差错报告报文不再发送 ICMP 差错报告报文。
- ❖ 对第一个分片的数据报片的所有后续数据片都不发送 ICMP 差错报告报文。
- ❖ 对具有多播地址的数据报都不发送 ICMP 差错报告报文。
- ❖ 对具有特殊地址（如127.0.0.0 或 0.0.0.0）的数据报不发送 ICMP 差错报告报文。

ICMP 询问报文有两种

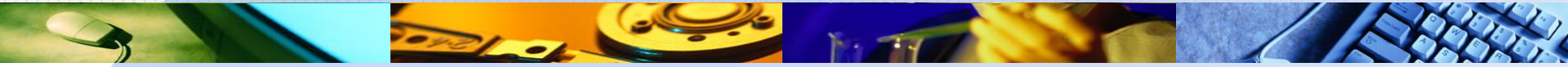


- ❖ 回送请求和回答报文
- ❖ 时间戳请求和回答报文

下面的几种 ICMP 报文不再使用

- ❖ 信息请求与回答报文
- ❖ 掩码地址请求和回答报文
- ❖ 路由器询问和通告报文

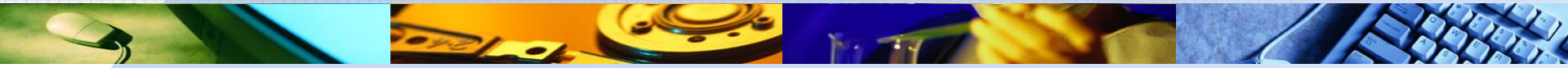
ICMP的应用举例



PING (Packet InterNet Groper)

- ❖ PING 用来测试两个主机之间的连通性。
- ❖ PING 使用了 ICMP 回送请求与回送回答报文。
- ❖ PING 是应用层直接使用网络层 ICMP 的例子，它没有通过运输层的 TCP 或UDP。

PING 的应用举例



```
C:\Documents and Settings\XXR>ping mail.sina.com.cn
```

```
Pinging mail.sina.com.cn [202.108.43.230] with 32 bytes of data:
```

```
Reply from 202.108.43.230: bytes=32 time=368ms TTL=242
```

```
Reply from 202.108.43.230: bytes=32 time=374ms TTL=242
```

```
Request timed out.
```

```
Reply from 202.108.43.230: bytes=32 time=374ms TTL=242
```

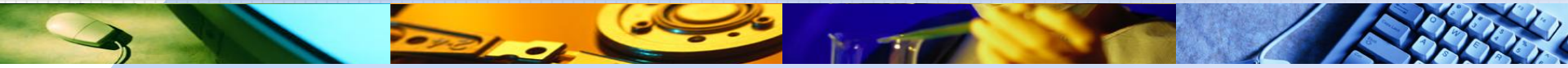
```
Ping statistics for 202.108.43.230:
```

```
    Packets: Sent = 4, Received = 3, Lost = 1 (25% loss),
```

```
Approximate round trip times in milli-seconds:
```

```
    Minimum = 368ms, Maximum = 374ms, Average = 372ms
```

Traceroute 的应用举例



```
C:\Documents and Settings\XXR>tracert mail.sina.com.cn
```

```
Tracing route to mail.sina.com.cn [202.108.43.230]  
over a maximum of 30 hops:
```

1	24 ms	24 ms	23 ms	222.95.172.1
2	23 ms	24 ms	22 ms	221.231.204.129
3	23 ms	22 ms	23 ms	221.231.206.9
4	24 ms	23 ms	24 ms	202.97.27.37
5	22 ms	23 ms	24 ms	202.97.41.226
6	28 ms	28 ms	28 ms	202.97.35.25
7	50 ms	50 ms	51 ms	202.97.36.86
8	308 ms	311 ms	310 ms	219.158.32.1
9	307 ms	305 ms	305 ms	219.158.13.17
10	164 ms	164 ms	165 ms	202.96.12.154
11	322 ms	320 ms	2988 ms	61.135.148.50
12	321 ms	322 ms	320 ms	freemail43-230.sina.com [202.108.43.230]

```
Trace complete.
```