



第 五 章 网络层 (2)

课程名称： 计算机网络

主讲教师： 姚焱

课程代码： 1020100-1

第31-32讲

E-MAIL : yaoye@nwpu.edu.cn

2020 – 2021 学年第一学期



本节内容

- 1. IP协议
 - IP 分组格式
 - IP 分片与组装
 - IP 分组转发机制
- 2. ARP协议
- 3. ICMP协议



引言-因特网的网际协议 IP

- 网际协议IP是TCP/IP集中最主要协议之一, 主要包括以下功能:
 - 分组生成, 发送、接收和处理;
 - IP分片与组装(分片对象是上层协议数据单元);
 - IP分组转发: 每个IP分组(或IP分片)独立路由
- 网络层其他四个协议:
 - 网际控制报文协议ICMP(Internet Control Message Protocol)
 - 因特网组管理协议IGMP(Internet Group Management Protocol)
 - 地址解析协议 ARP (Address Resolution Protocol)
 - 逆地址解析协议 RARP(Reverse Address Resolution Protocol)



本节内容提要

- 1. IP协议
 - 1.1 IP分组格式
 - 1.2 分组的分片与组装
 - 1.3 分组转发(路由选择)
 - 1.4 选项
- 2. ARP协议
- 3. ICMP协议

- 一个IP 分组由首部和数据两部分组成，首部由固定部分和可变部分组成。
- 固定部分，长度固定，共20字节，每个IP分组必须具有。
- 可变部分：长度可变（0~40B），可选字段和填充字段组成。

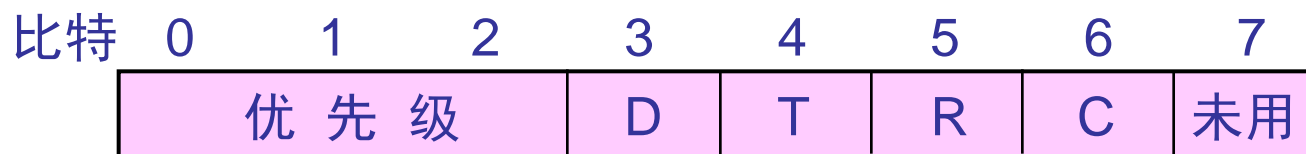




- IP版本：占 4 bit，指IP协议的版本
 - 目前的IP协议版本号为 4（即 IPv4）：0100

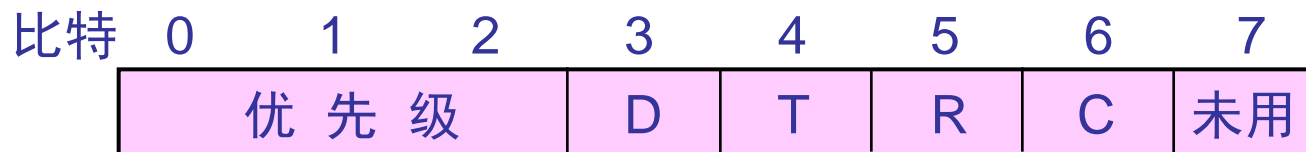


- IP首部长度：占 4 bit，可表示的最大数值是 15 个单位（一个单位为 4 字节）。
 - IP 的首部长度的最大值是60字节，其中固定长度20字节不变；
 - 可变部分长度：0~40字节；
 - IP首部长度范围：20~60字节；（0101-1111）

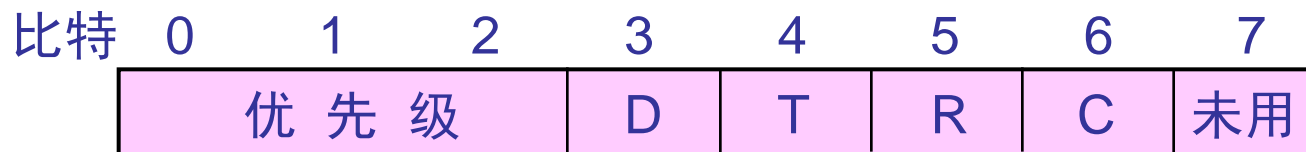


■ 服务类型 (ToS: Type of Service)

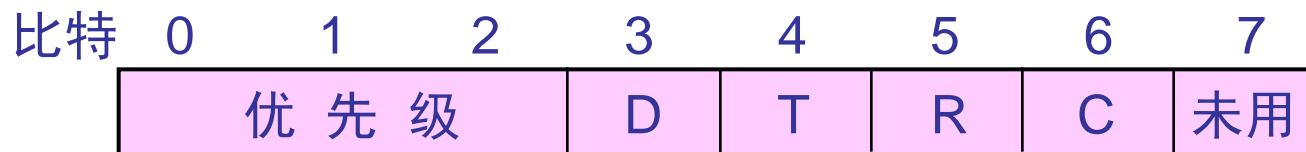
- IP分组优先级: 3 bit; 8个优先级, 0最小, 7最大.
- D: 低延迟 (delay); T: 高吞吐量 (throughput); R: 高可靠性 (reliability: 被路由丢弃概率较小); C: 选择代价更小路由.
- ToS只是用户要求, 对网络并不是强制, 路由器进行路由选择等处理时仅仅作参考。



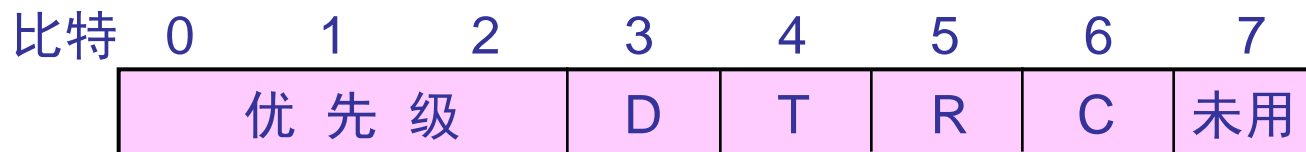
- **总长度:** 占 16 bit, 指首部+数据的长度 (578-1500B) ,
 - 单位为字节, 因此分组的最大长度为 65535 字节。
 - 总长度必须不超过最大传送单元 MTU。
 - 目前该长度足够, 在将来高速网络中, 数据帧MTU有可能大于65535 字节; 每个分片总长度仅仅指该分片首部+数据长度



- 标识 (identification): 占 16 bit.
 - 源节点每产生一个新分组, 计数器+1, 作为该分组的标识, 而**不是序号**, IP提供无连接通信, 不存在按序接收.
 - 注意: 一个分组不同分片具有相同标识.
 - 与分片与组装有关的字段: 标识+标志+片偏移



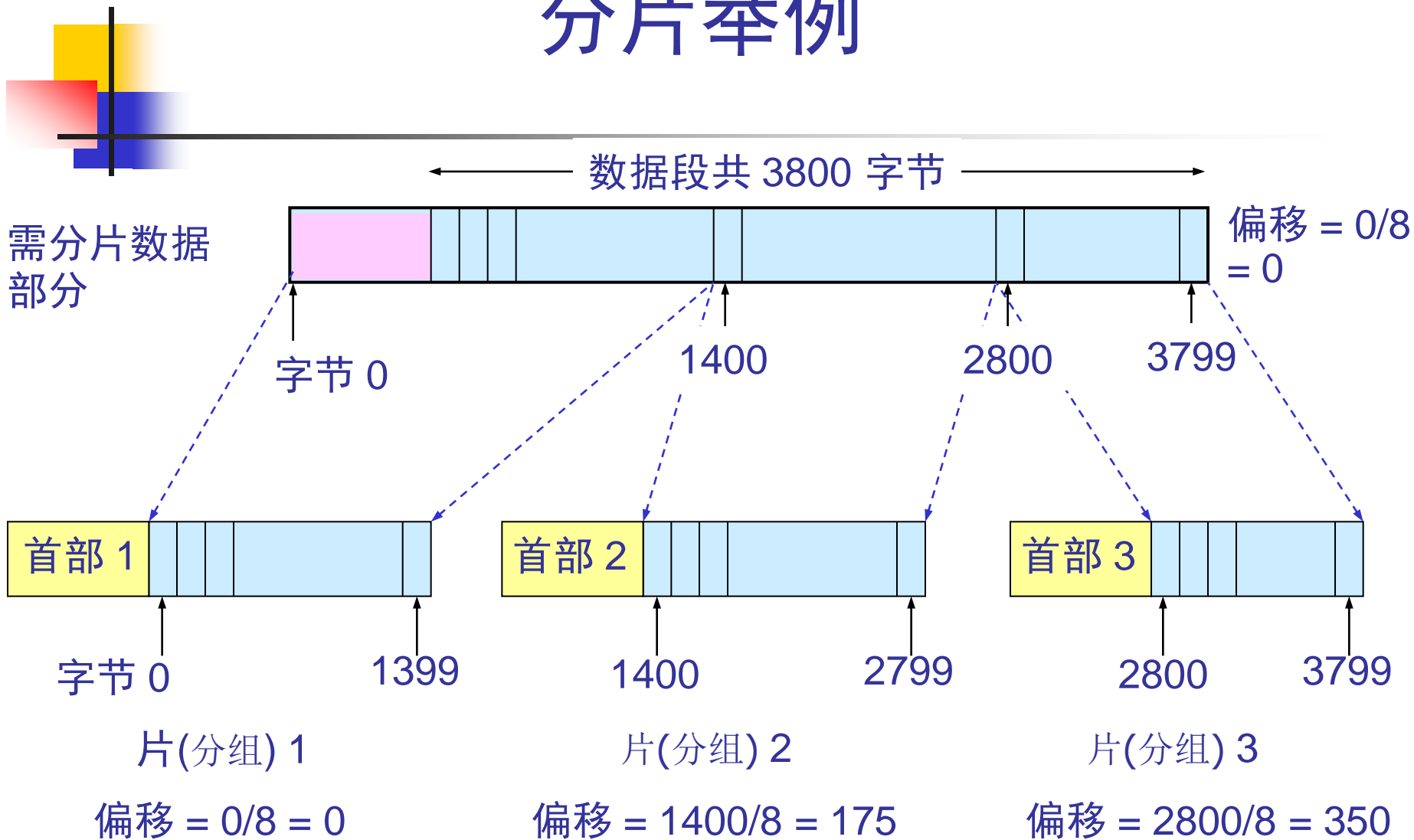
- 标志(Flag): 占 3 bit, 目前只使用两个bit: 0+DF+MF.
 - DF (Don't Fragment), DF=1表示"不允许分片", DF=0表示"允许分片".
 - MF (More Fragment), MF=1表示: 这是一个分片后面有分片; MF=0表示: 如果分片这最后一个分片; 或者这是一个分组(没有分片)。

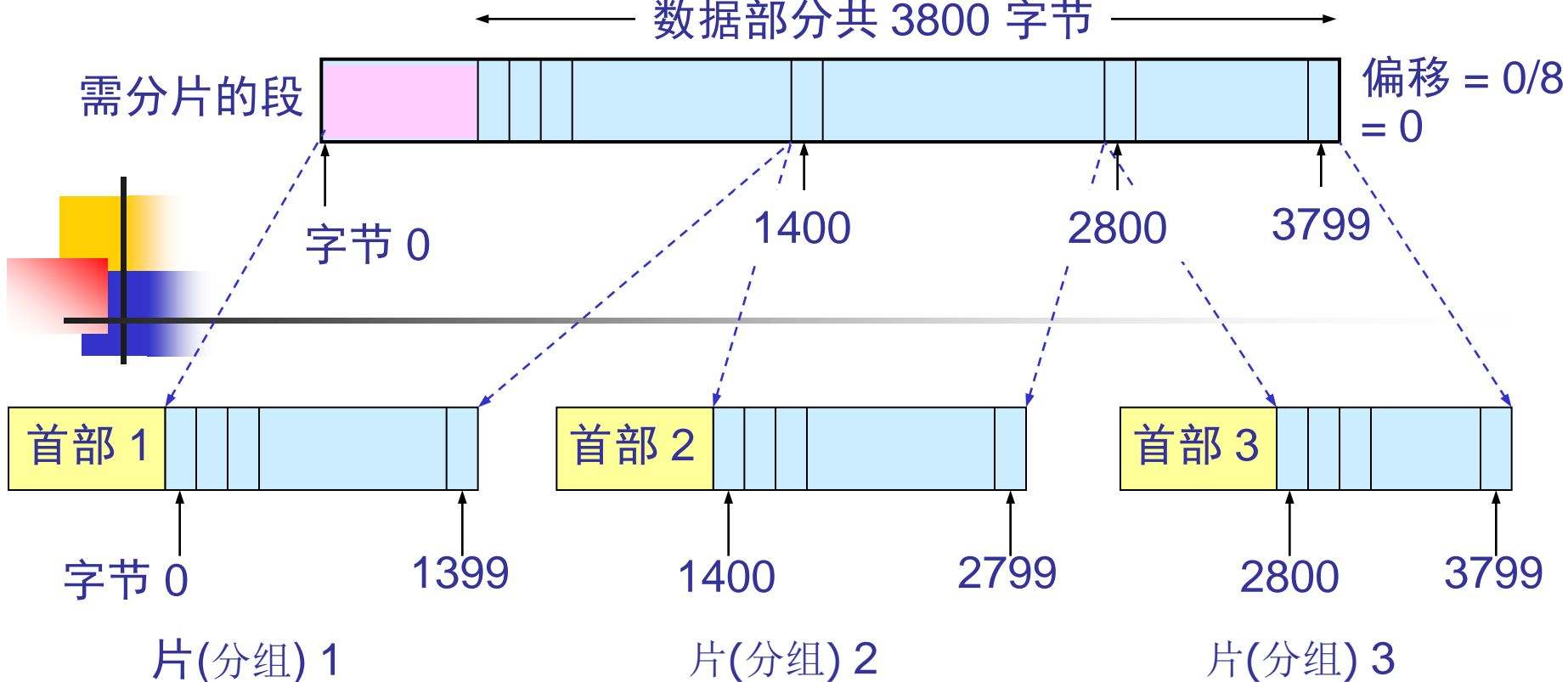


■ 片偏移：12 bit

- 针对较大上层协议数据单元，主要指数据部分在分片后，某一个分片(也是一个分组)的数据部分第一个字节在原始数据中的绝对位置。
- 片偏移以8个字节为偏移单位。

分片举例



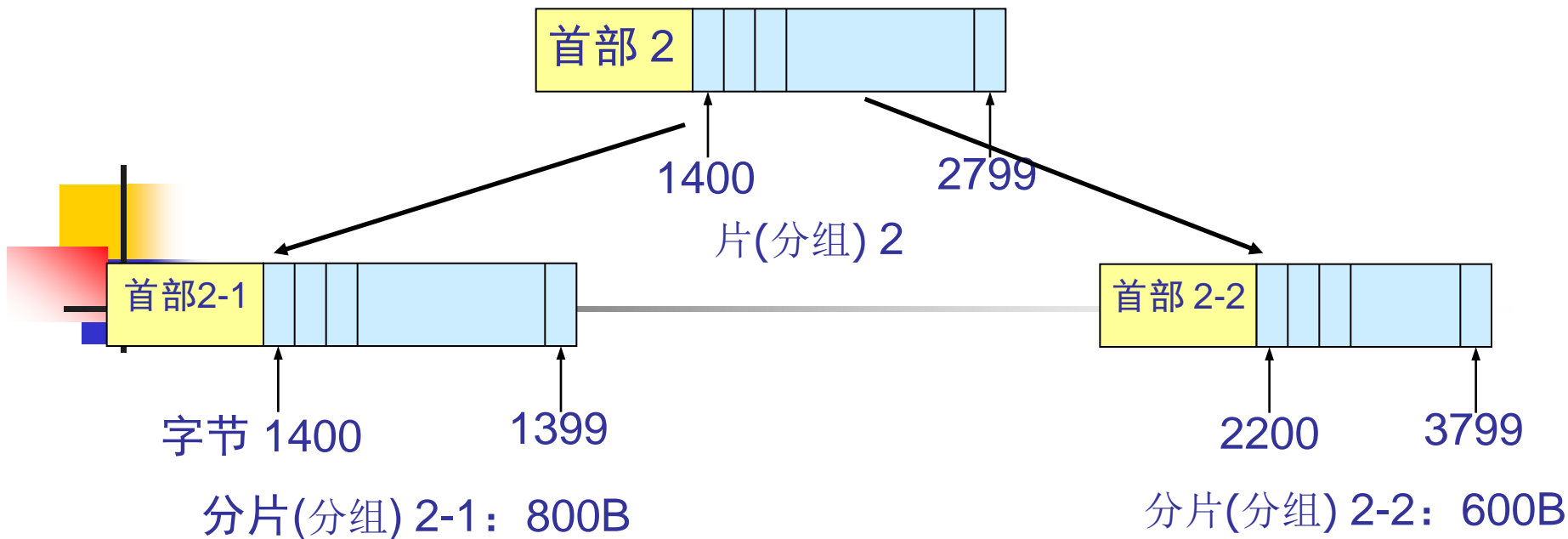


偏移 = $0/8 = 0$

偏移 = $1400/8 = 175$

偏移 = $2800/8 = 350$

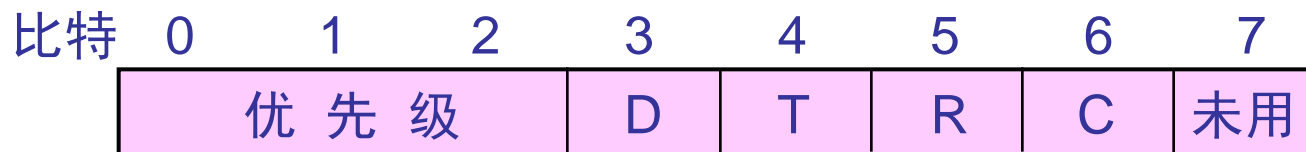
	总长度	标识	MF	DF	片偏移
原始段	3820	12345	0	0	0
分片1	1420	12345	1	0	0
分片2	1420	12345	1	0	175
分片3	1020	12345	0	0	350



偏移 = $1400/8 = 175$

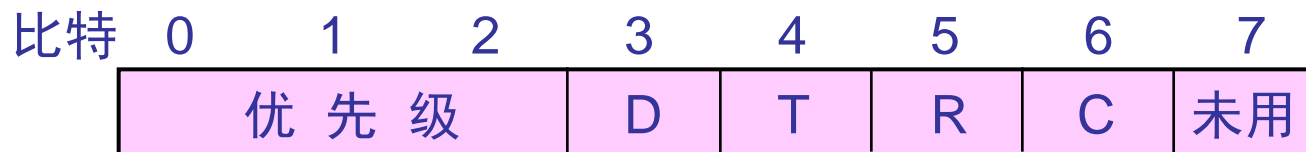
偏移 = $2200/8 = 275$

	总长度	标识	MF	DF	片偏移
原始段	3800	12345	0	0	0
分组1	1420	12345	1	0	0
分组2-1	820	12345	1	0	175
分组2-2	620	12345	1	0	275
分组3	1020	12345	0	0	350



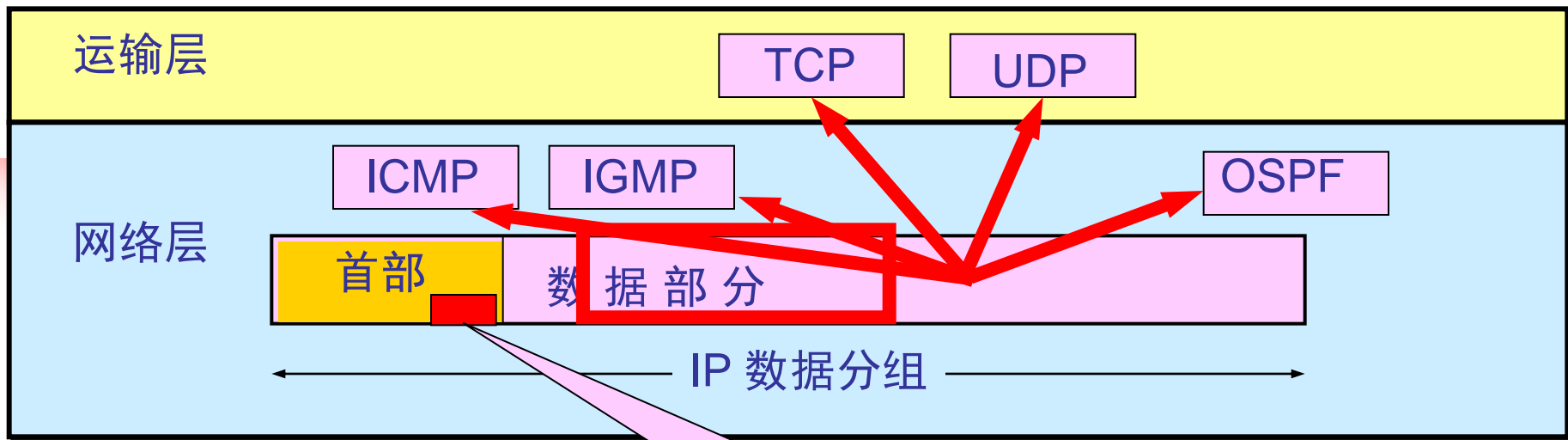
■ 生存时间 (TTL : Time To Live) : 8 bit

- 目的：记录IP分组网络中转发时间长度（或者转发次数）。
- 发送源一般TTL设置为64，最大255，每经过一个路由器，在转发前TTL减1，为0丢弃，同时向源端发送一个ICMP超时信息。
- 情况：网络存在路由环路，或找不到确切路由，一直使用默认路由。



■ 协议字段：8 bit

- 指出此分组携带的数据来自何种上层协议，以便目的主机在IP层将数据部分上交给上层相应协议进程处理。
- 协议号由国际组织IANA负责分配。



TCP	6	EGP	8
UDP	17	IGP	9
ICMP	1	OSPF	89
IGMP	2	IPV6	41

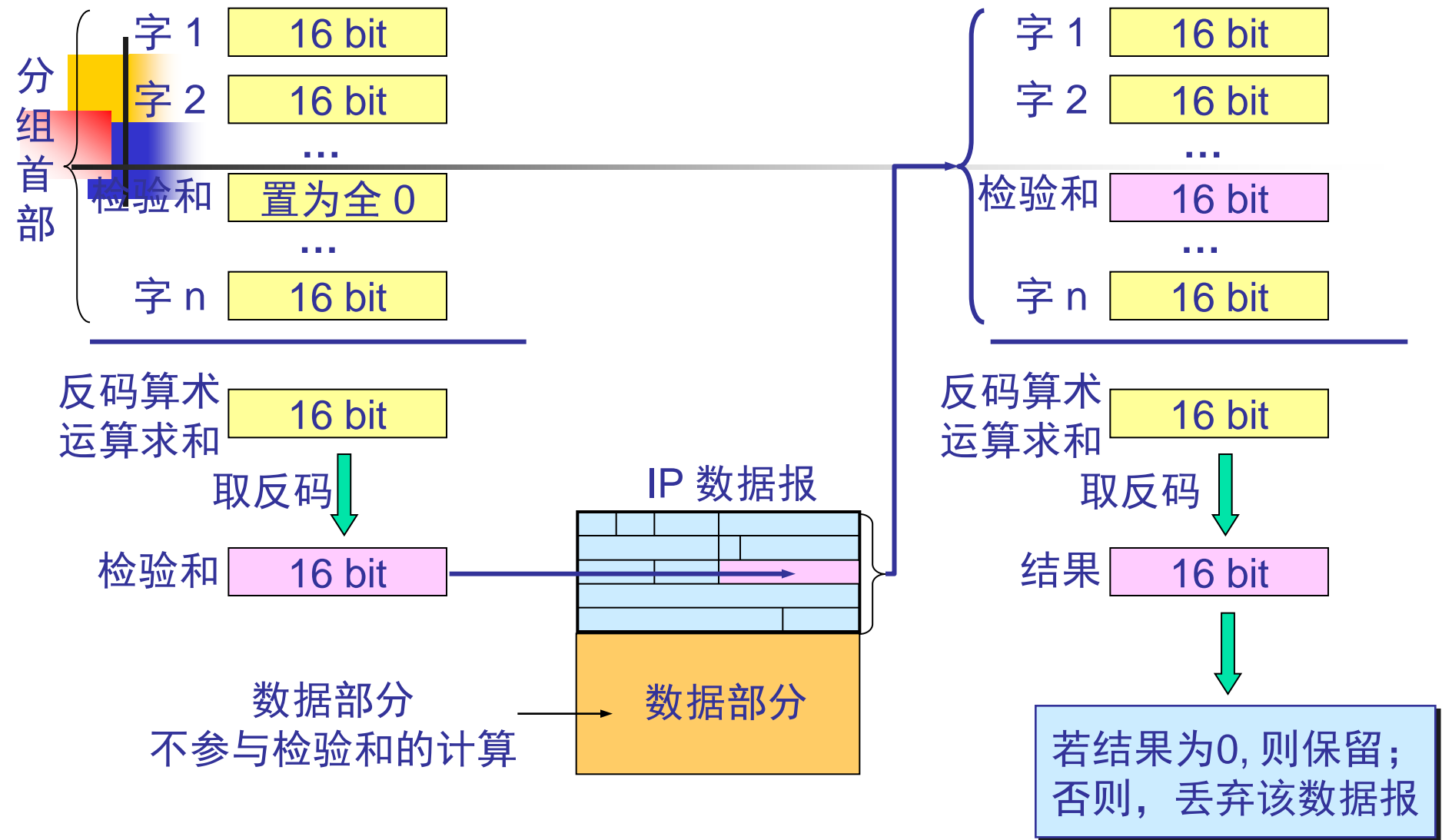
比特	0	1	2	3	4	5	6	7
	优 先 级			D	T	R	C	未用

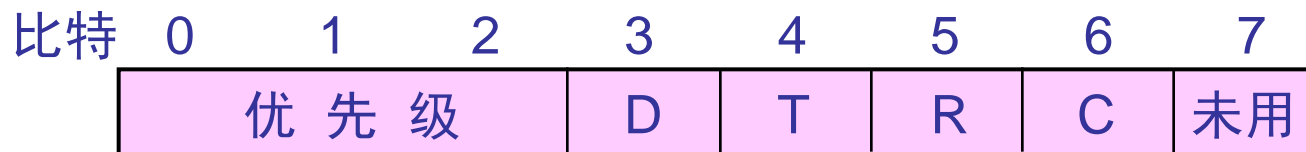


- **首部校验和字段： 16 bit**
 - 只校验IP**分组首部**, **不包括数据部分**, 减少IP层处理分组时间, 数据差错控制由端系统传输层完成。
 - 不采用CRC检验码, 而采用简单的计算方法, 减少路由器计算负荷。
 - 当分组从一个路由器转发给另一个路由器需前要重新计算首部校验和 (至少TTL、片偏移, 总长度等字段发生变化);

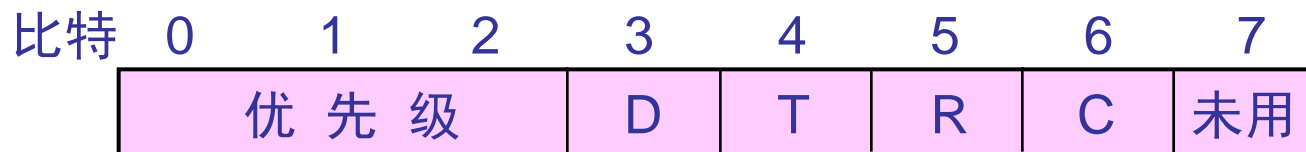
发送端

接收端

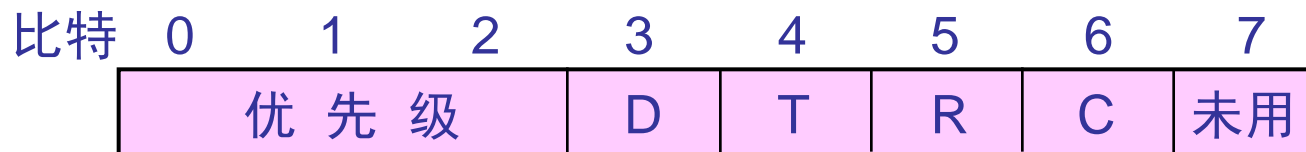




- 源IP地址和目的IP地址：各占 4 字节；
- 在转发路径上，这两个字段的值一直保持不变（除了 NAT 技术外）。



- 选项字段: 0个字节到40个字节不等; 用于控制、测量和安全。
 - 提供各种选服务: 如时间戳选项, 源路由选项等;
 - 可变部分增加IP协议功能的同时, 也使得首部长度成为可变, 增加了每一个路由器处理分组开销, 实际上这些选项现在很少被使用。



- 填充字段：
 - 可变长度, 保证IP分组首部以32比特位边界对齐。



本节内容提要

- 1. IP协议
 - 1.1 IP分组格式
 - 1.2 分组的分片与组装
 - 1.3 分组转发(路由选择)
 - 1.4 选项
- 2. ARP协议
- 3. ICMP协议

1.2 分组的分片与组装

- 各种实际物理网络中, 数据链路层帧的数据字段都有不同的最大长度限制, 称为最大数据传送单元MTU(Maximum Transfer Unit)

物理网络	MTU(字节)
Ethernet	1500
Token Ring(4Mb/s)	4464
Token Ring(16Mb/s)	17914
FDDI	4352
x.25	576
PPP	296
Hyperchannel	65535



1.2 分组的分片与组装

- 为了使分组以适当的大小在物理网络上传输，IP协议首先要根据物理网络所允许的最大帧长对上层协议提交的协议数据单元长度检查，必要时通过分片，分成若干个分组发送。
- 特别强调：分片对象对**分组的数据部分**，而不包括IP首部。
- 分片与组装有关字段：
 - **标识 (ID)**：每个分片具有唯一标识。
 - **标志 (Flag)**：如果是**无分片的IP数据分组**，**MF=0**；如果是有分片的IP数据分组，除了最后一个分片IP数据分组MF=0外，MF=1。
 - **分段偏移**：每一个分片（分组）要表明数据部分第一个字节在原始数据中的位置，用8字节倍数来表示。
- 每个分片独立发送，独立路由，可能在接收方出现乱序；



1.2 分组的分片与**组装**

- 分片组装条件

- 在进行数据分组重新组装时，每个IP分片具有相同标识，按偏移排序。
- 具有相同的上层协议号、源IP地址和目的IP地址。
- 在一定的时间内要全部到齐。

- 问题：

- 如果个别分组（分片）首部有错误，有个别分组没有在规定时间内到，发生什么事情？
- 会造成分片组装错误，这时由传输层重发相应的数据段。

- 分组组装

- 按分片偏移字段大小，将数据组装成一个完整的原始数据段；
- 最后将组装好的原始数据段按上层协议号提交给上层协议进程。
- 如果传输层检测到段数据有错误，发送方传输层需要重发该数据段。



本节内容提要

- 1. IP协议
 - 1.1 IP分组格式
 - 1.2 数据段的分片与组装
 - 1.3 选项字段
 - 1.4 分组转发(路由选择)
- 2. ARP协议
- 3. ICMP协议

1.3 IP分组选项字段



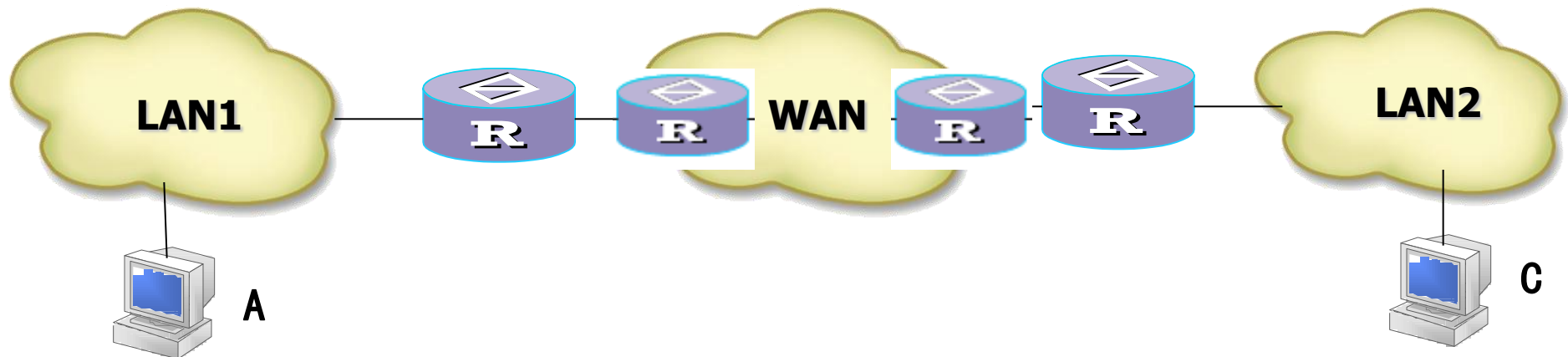


1.3 IP分组选项字段

- IP首部头中选项字段组成：代码+长度+数据
- IPV4定义了5种选项
 - 限制源路由选项(strict route)
 - 自由源路由选项(loose route)
 - 记录路由选项(record route)
 - 时间戳选项(timestamp)
 - 安全选项(Security)

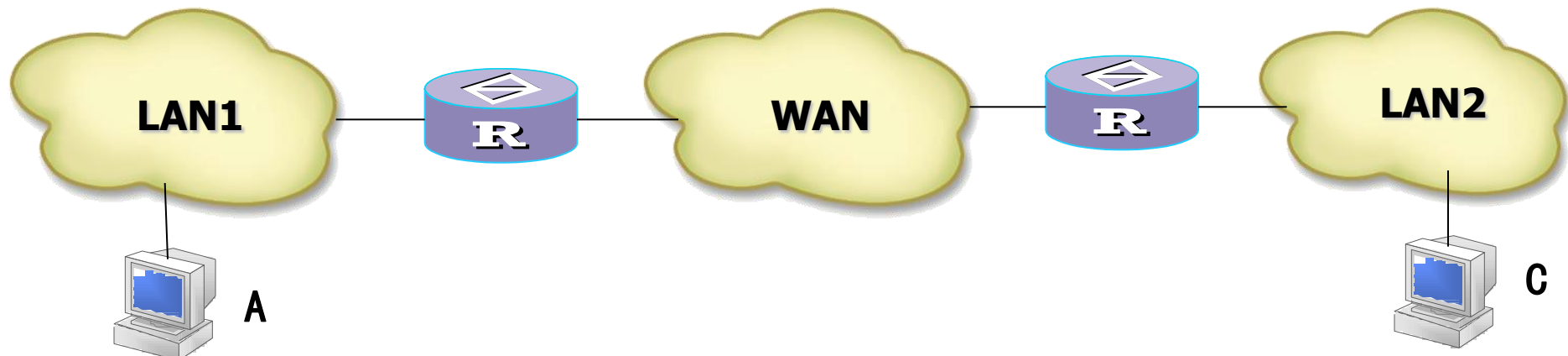
1.3 IP分组选项字段

- 源路由选项：指转发分组的路由信息是由源主机指定，而不是由IP协议通过路由表确定的；
- 源路由选项有两类：
 - 限制源路由：一条完整的路径信息。
 - 自由源路由：不完整的路径信息。
 - 目的：源路由选项可以用于测试某特定网络路径性能，或使分组尽可能绕开出错的网络段等。



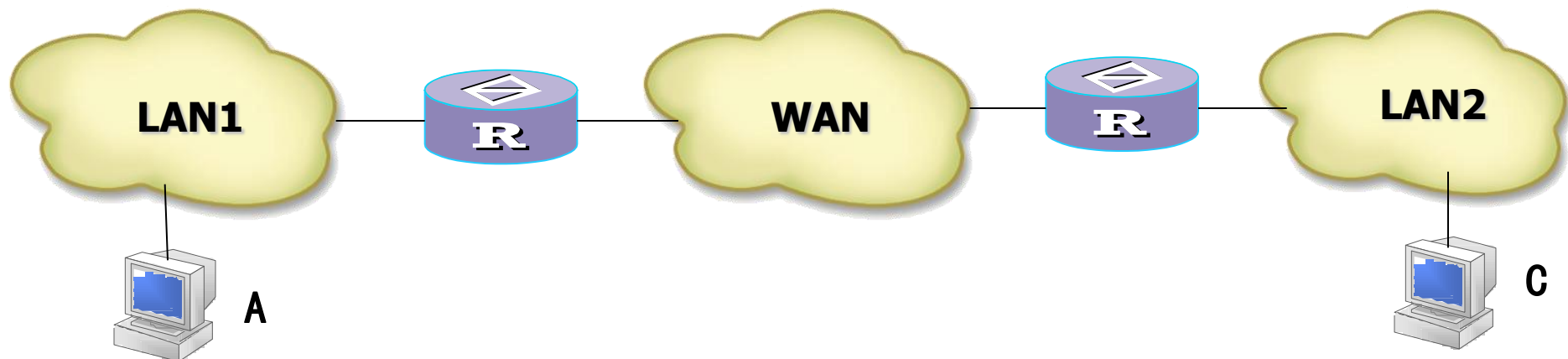
1.3 IP分组选项字段

- 记录路由选项：记录下分组从源到目的所经过路径上各个路由器IP地址。
 - 条件：源主机和目的主机双方都同意。
 - 具体内容：记录下分组从源主机到目的主机所经过路径上各个路由器的IP地址。
 - 特例：记录地址的区域大小是由源主机预先分配并初始化的，如果预先分配的区域不足以记录下全部路径，则IP协议将放弃记录余下地址。
 - 用途：利用该选项可得到源路由信息。（最多可记录10个IP地址）



1.3 IP分组选项字段

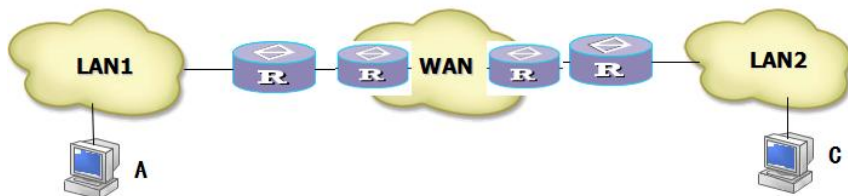
- 时间戳选项：记录下分组从源到目的所经过各个路由器的时间。
 - 条件：源主机和目的主机双方都同意。
 - 具体内容：记录分组从源主机到目的主机经过路径上各个路由器的时间。
 - 使用：选择时间戳选项时，可以设置成时间戳选项和记录路由选项同时使用，或时间戳选项和源路由信息选项同时使用。
 - 作用：分析网络吞吐量、负载情况。





本节内容提要

- 1. IP协议
 - 1.1 IP分组格式
 - 1.2 分组的分片与组装
 - 1.3 选项
 - 1.4 IP分组转发
- 2. ARP协议
- 3. ICMP协议



1.4 IP分组转发(路由选择)

■ IP分组转发

- 当发送节点IP协议进程转发分组时, $TTL-1=0$? 判断; 校验和检查, 排队缓存。
- 进行处理, 如果上层协议已**指定了源路由信息**, 则按指定源路由转发。
- 如果上层协议未指定源路由信息, IP协议则以IP分组中目的IP地址为与**路由表中子网掩码**, 按照**网络地址**查找路由表。
- 调用ARP协议, 得到下一跳IP对应MAC地址, 转发到转发接口, 构建帧。

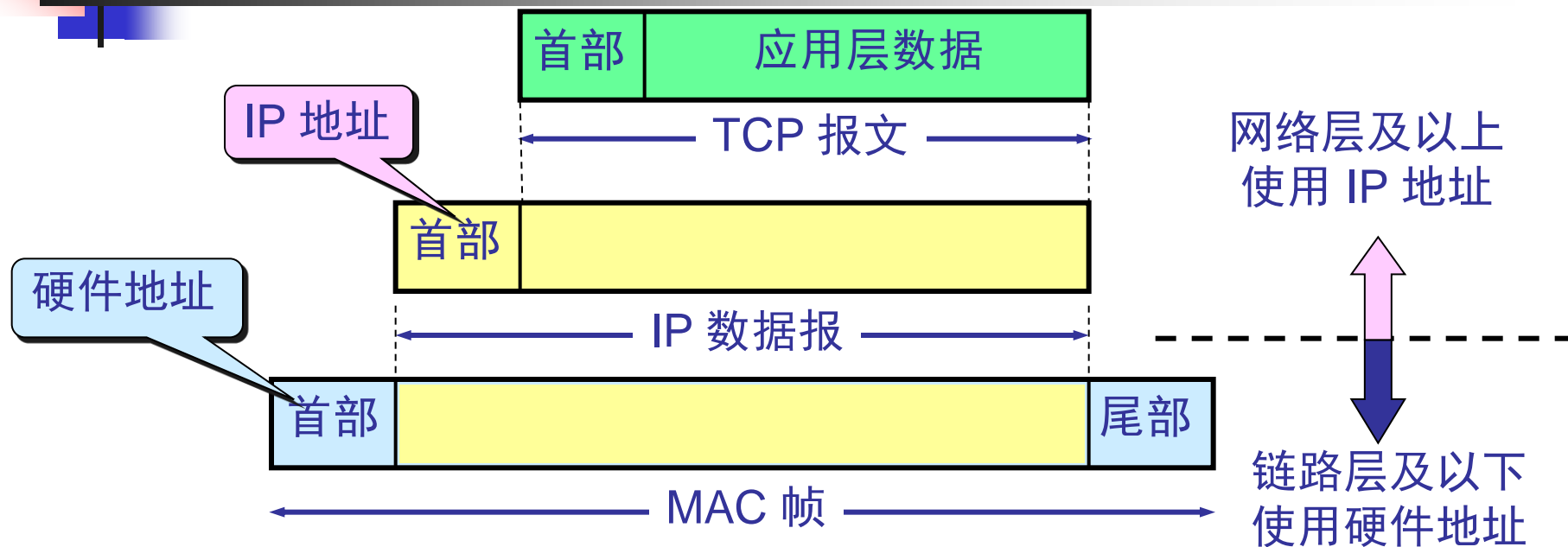
目的地址	子网掩码	下一跳	代价 (Metric)	优先 (Preference)
11. 0. 0. 0	255. 0. 0. 0	8. 8. 8. 9	30	100 (RIP)
11. 168. 0. 0	255. 255. 0. 0	12. 8. 8. 9	10	40 (OSPF)
11. 168. 1. 0	255. 255. 255. 0	13. 8. 8. 9	40	20 (静态)
11. 168. 2. 0	255. 255. 255. 0	Direct	20	0
0. 0. 0. 0	0. 0. 0. 0	14. 8. 8. 9	200	(默认)



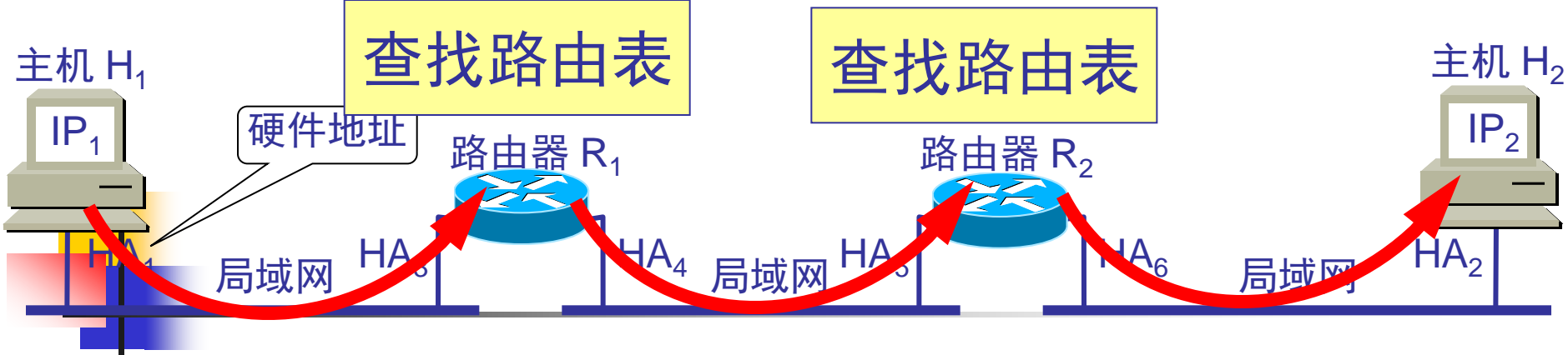
本节内容提要

- 1. IP协议
 - 1.1 IP分组格式
 - 1.2 数据段的分片与组装
 - 1.3 选项
 - 1.4 分组转发(路由选择)
- 2. ARP协议
- 3. ICMP协议

(1) IP 地址与硬件地址



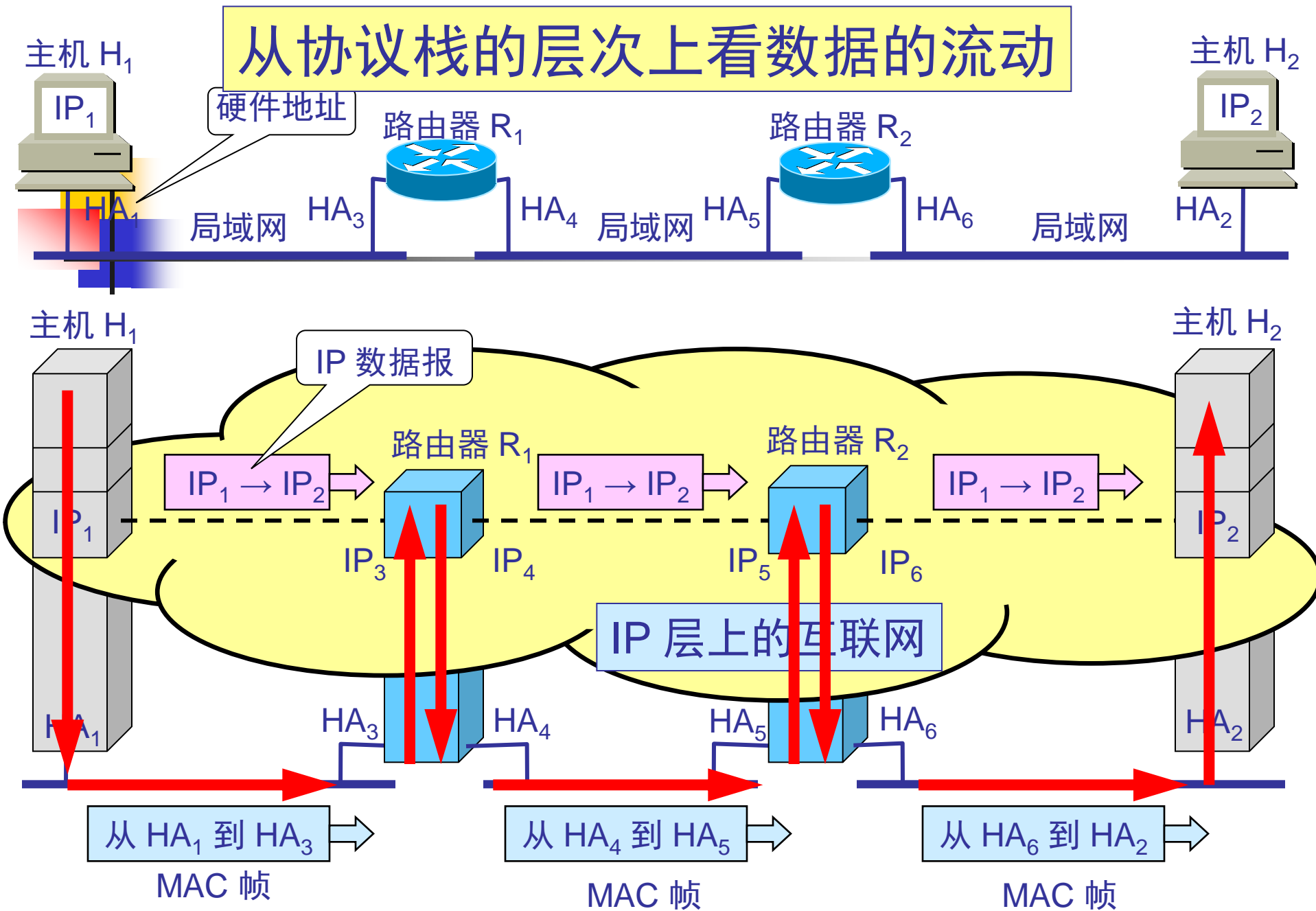
- IP地址（源、目的）在IP首部，网络层使用；
- MAC地址（源、目的）在数据帧首部，数据链路层使用；
- 当一个IP分组封装为数据帧时，在数据链路层看不到分组IP地址。



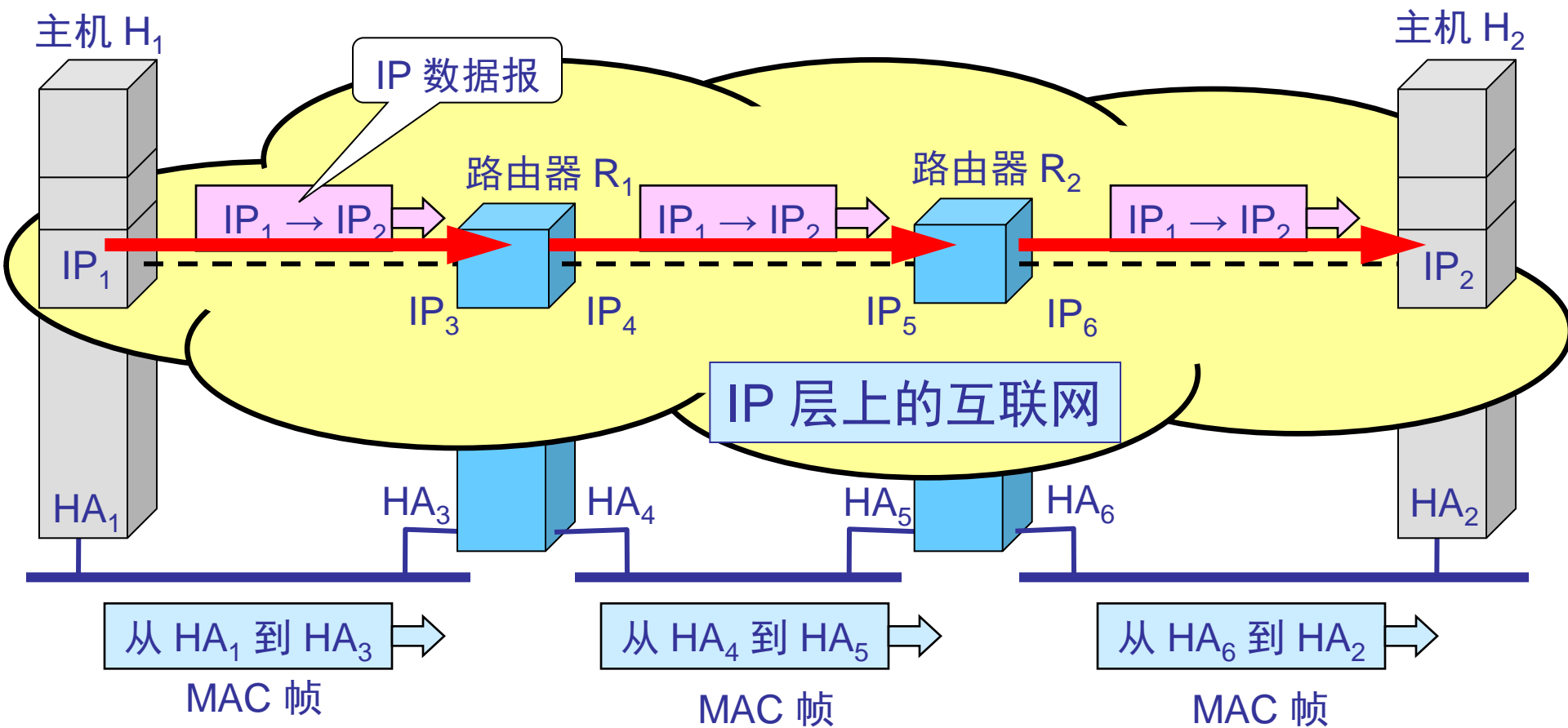
通信的路径

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

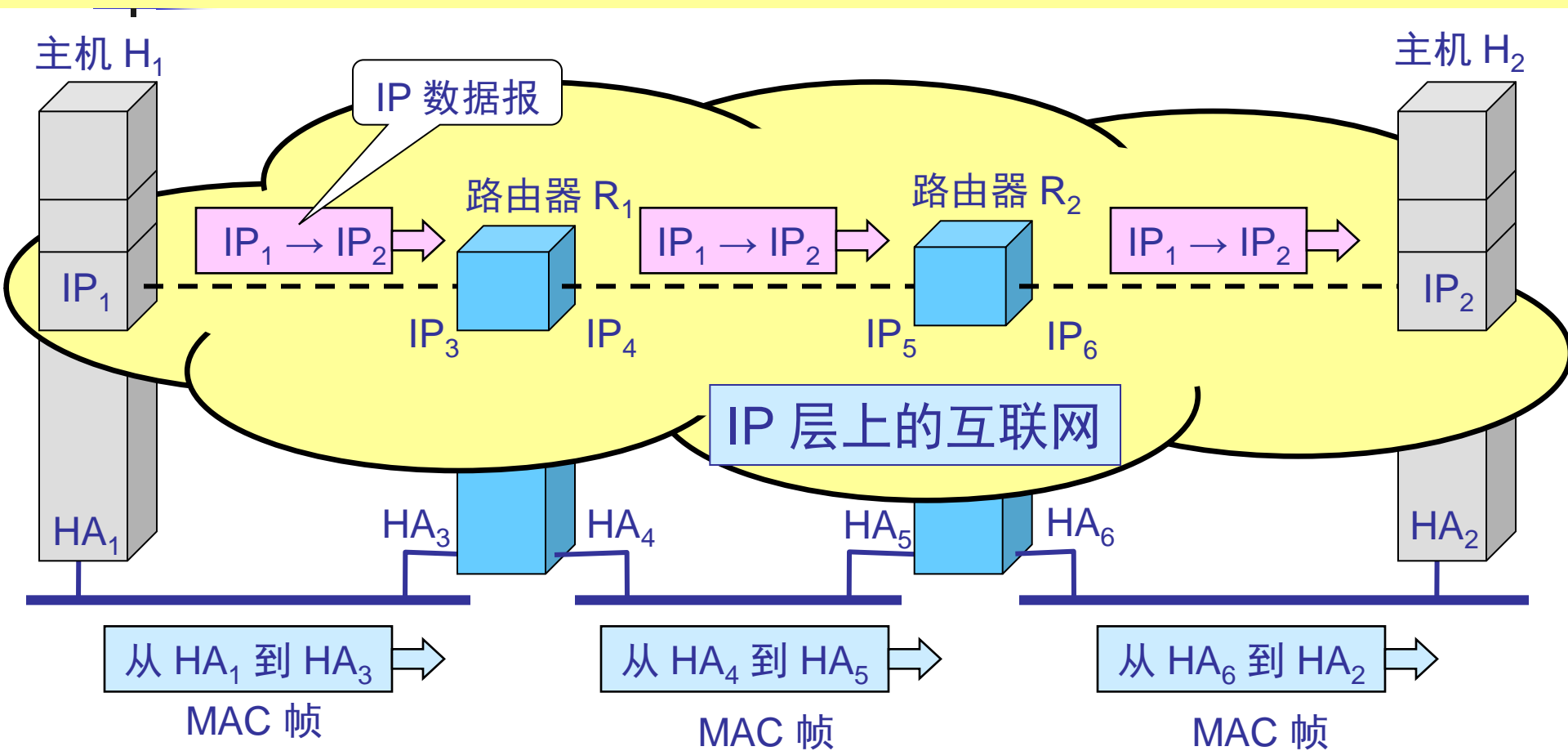
从协议栈的层次上看数据的流动



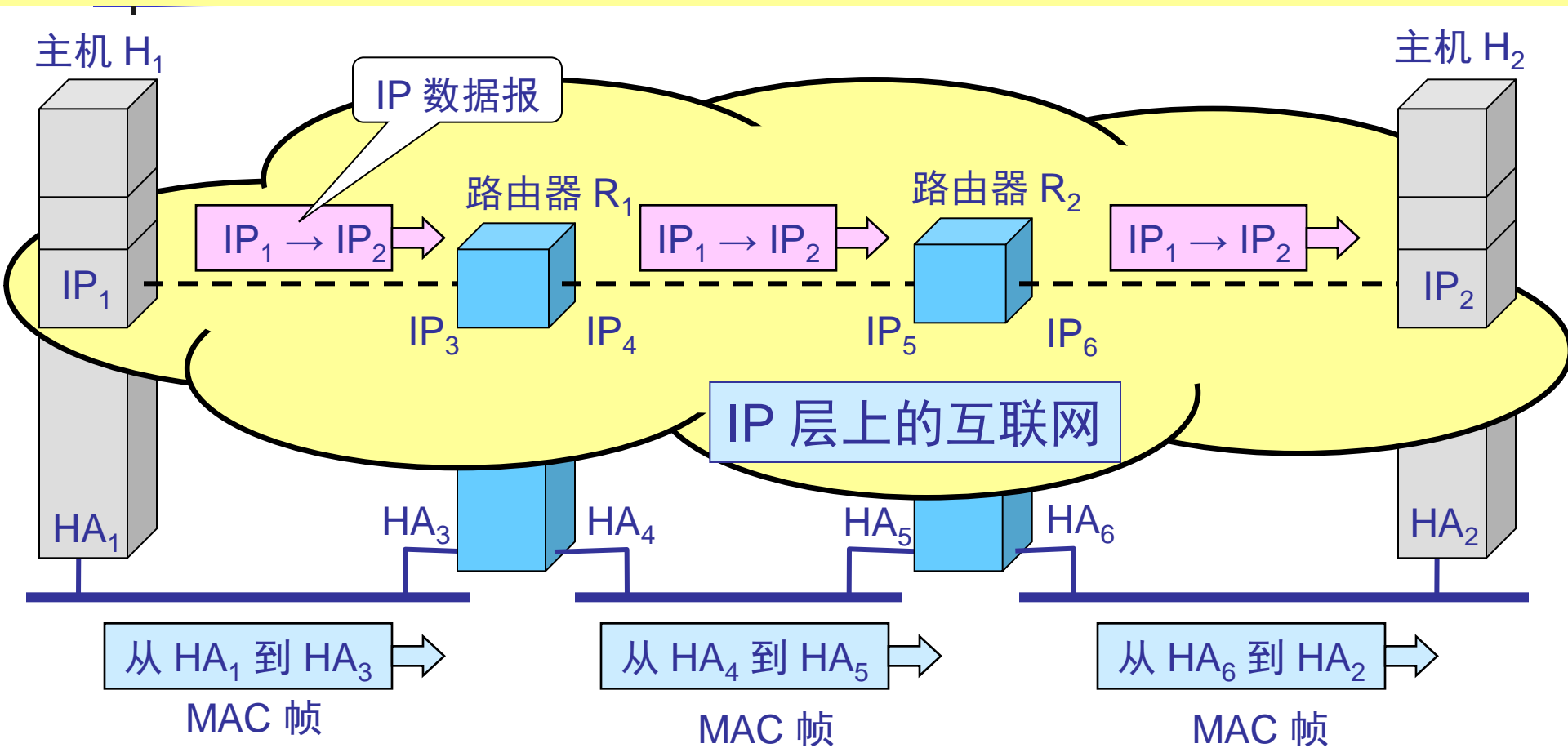
从虚拟的 IP 层上看 IP 数据报的流动



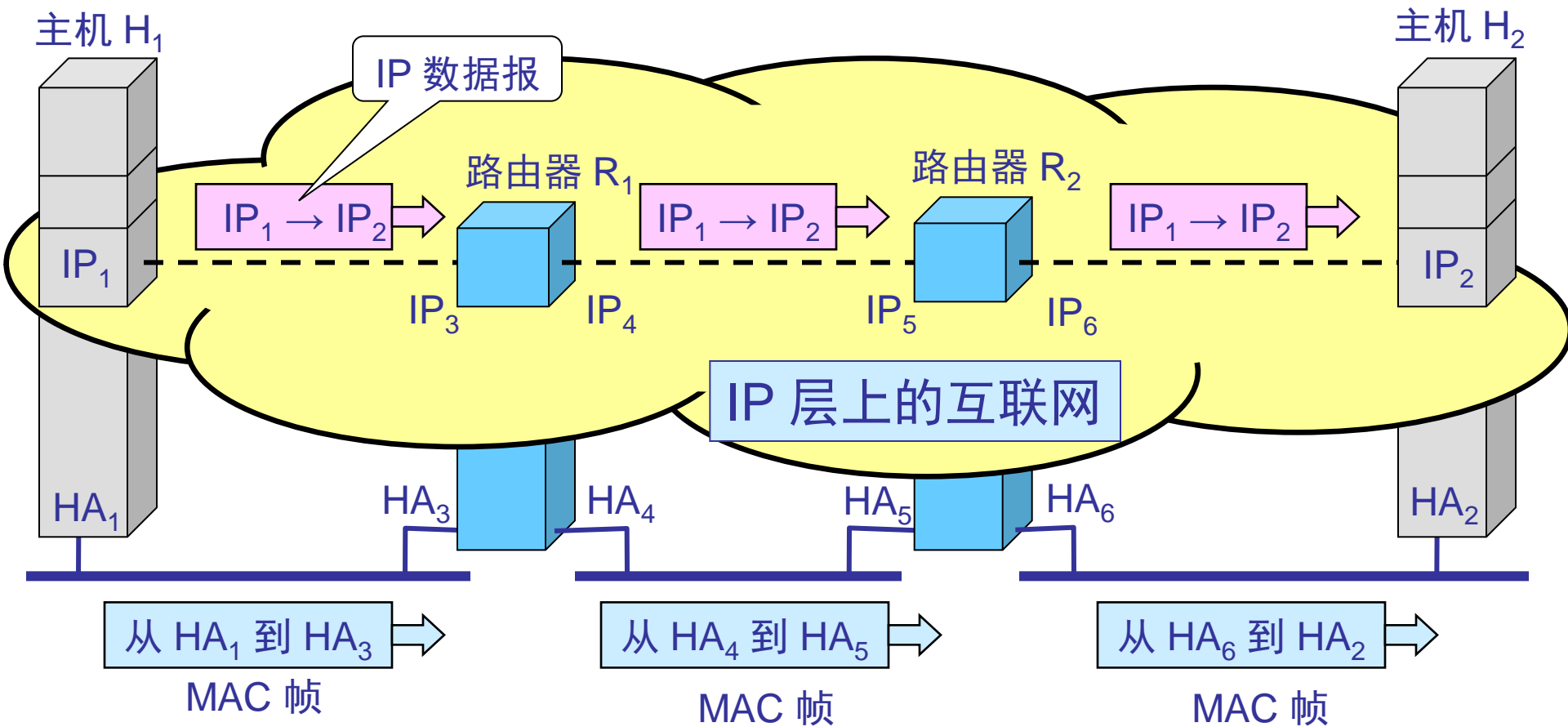
在 IP 层抽象的互联网上只能看到 IP 数据报
图中的 $IP_1 \rightarrow IP_2$ 表示从源地址 IP_1 到目的地址 IP_2
两个路由器的 IP 地址并不出现在 IP 数据报的首部中



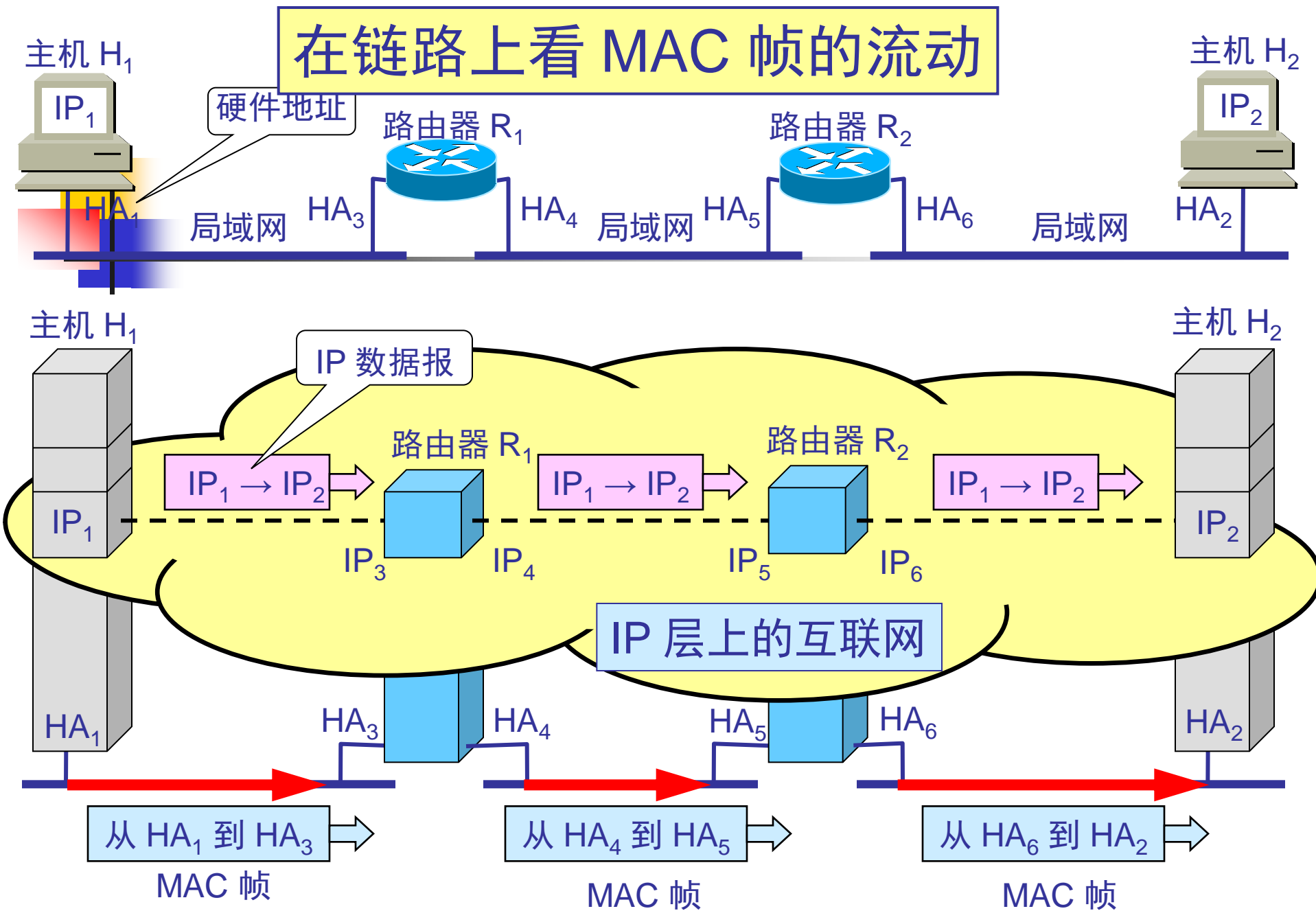
路由器只根据目的站的 IP 地址的网络号进行路由选择



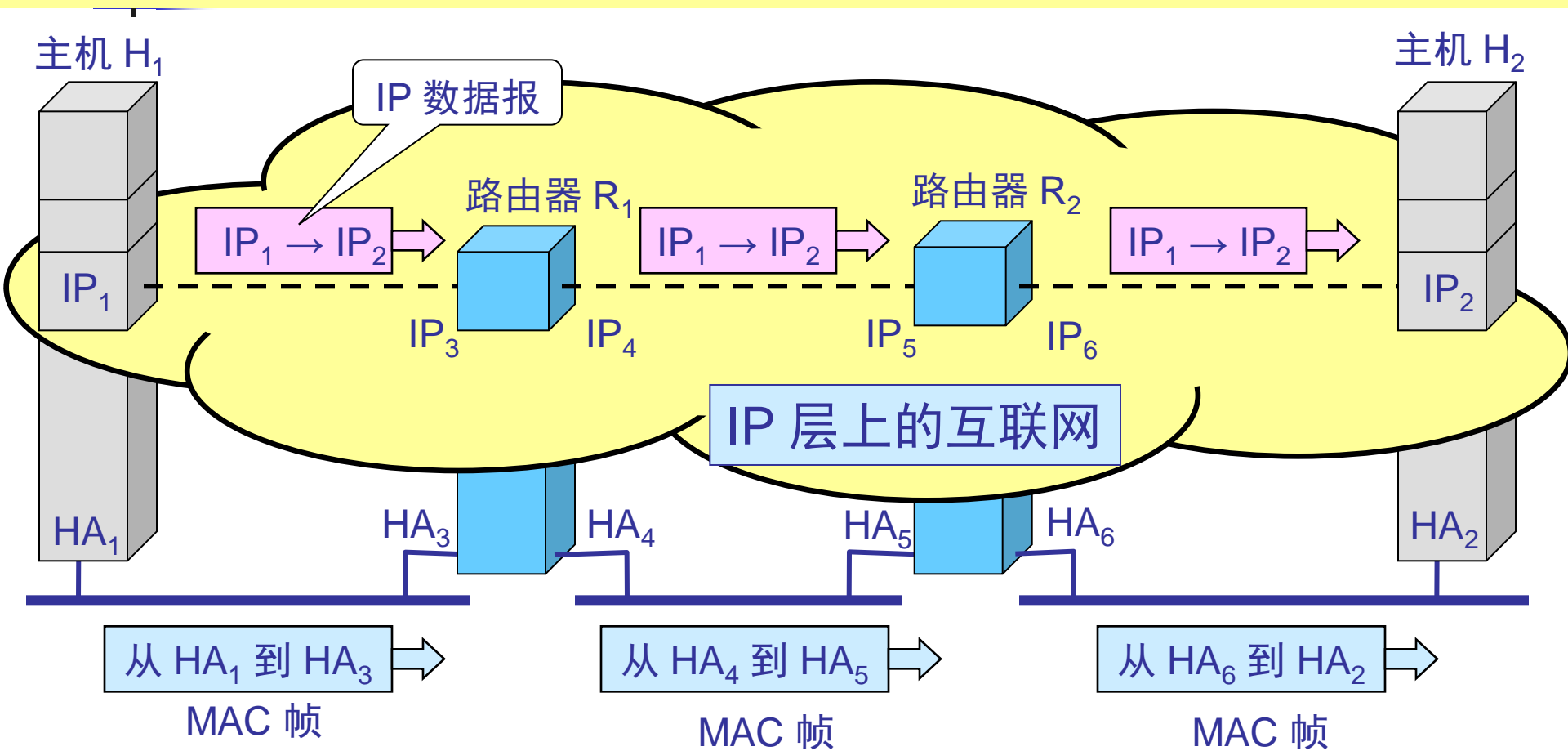
IP层抽象的互联网屏蔽了下层很复杂的细节
在抽象的网络层上讨论问题，就能够使用
统一的、抽象的 IP 地址
研究主机和主机或主机和路由器之间的通信

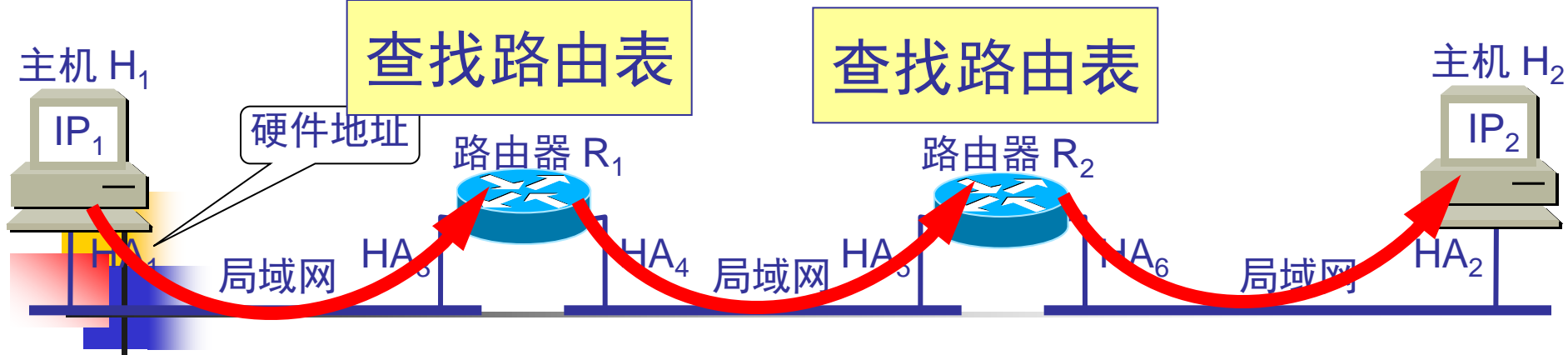


在链路上看 MAC 帧的流动



在具体物理网络链路层，只能看见 MAC 帧而看不见 IP 分组；
MAC地址（源、目的）在不同物理网络上发生了变化；说明：
(1) 源发送数据帧时，需要封装MAC地址（源、目的）； (2)
路由器在转发分组时，需要改变MAC地址（源、目的）。





通信的路径

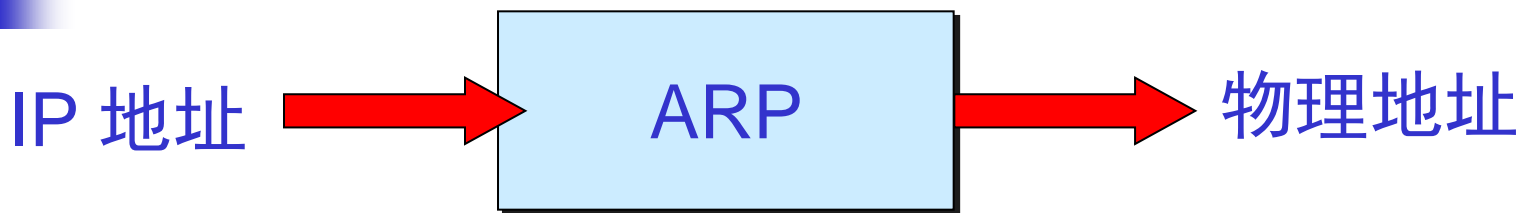
$H_1 \rightarrow$ 经过 R_1 转发 \rightarrow 再经过 R_2 转发 $\rightarrow H_2$

有两个重要问题需要解决：

(1) 主机和路由器如何知道在帧首部该填入什么样的MAC地址？

(2) 路由器中路由表是如何形成的？

(2) 地址解析协议 ARP 和 逆地址解析协议 RARP



- 逆地址解析协议 RARP在无盘工作站时代曾起到很重要作用；现在DHCP协议已经包含了RARP协议功能，本节不再单独讲解。

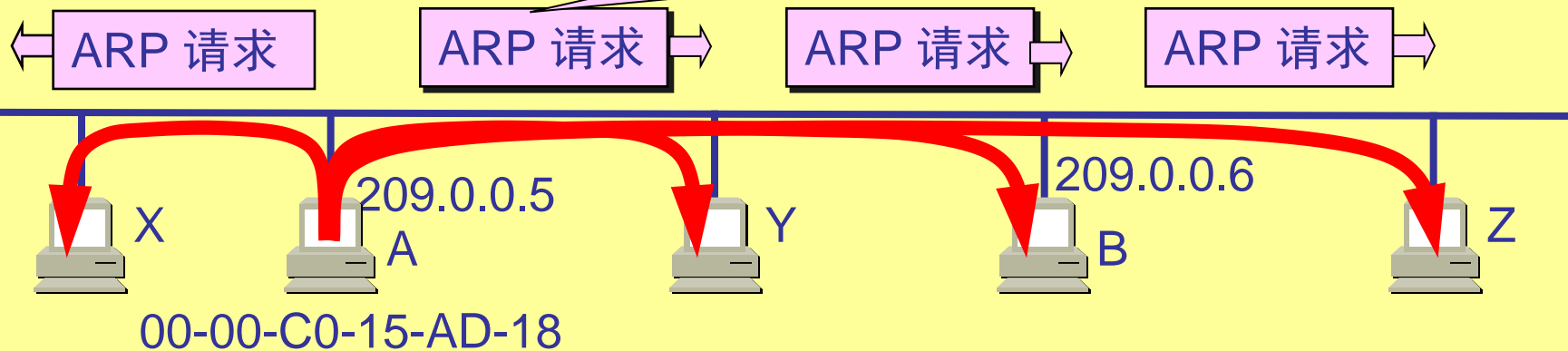


(2) 地址解析协议 ARP

- 不管网络层使用什么协议，在实际链路上传送数据帧时，最终还是必须使用MAC地址。
 - 每一个主机都设有一个 ARP 高速缓存 (ARP cache)，保存有所在的局域网上的各主机和路由器的 **IP 地址到硬件地址的映射**，每一条记录都有生存期。
 - 每个路由器也设有一个 **ARP 高速缓存**，保存所知道的所有IP地址到MAC地址映射，每一条记录都有生存期。
- 当主机 A 欲向本局域网上的某个主机 B 发送 IP 分组时，就先在其 ARP 高速缓存中查看有无主机 B 的 IP 地址-MAC 地址映射；
 - 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。
 - 否则，启动ARP协议，获得B的IP对应MAC地址。

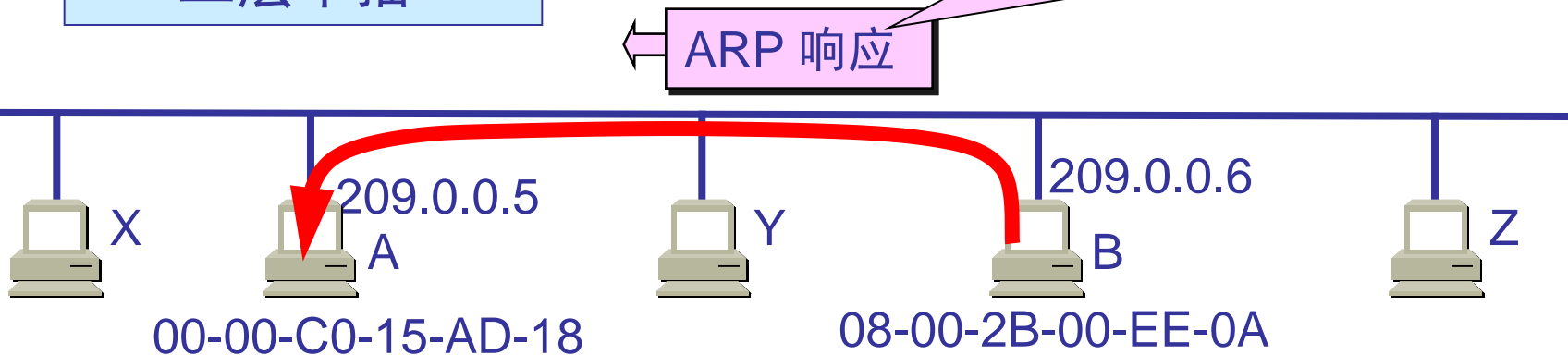
主机 A 二层广播发送
ARP 请求分组

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



主机 B 向 A 发送
ARP 响应分组
二层单播

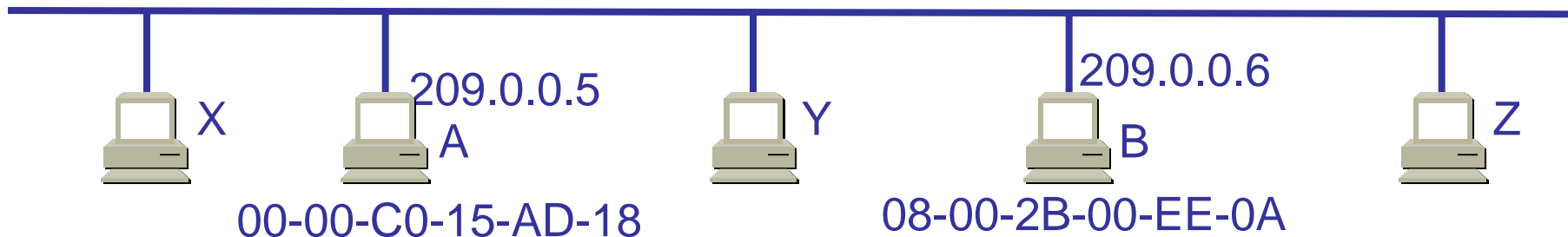
我是 209.0.0.6
硬件地址是 08-00-2B-00-EE-0A



如何提高ARP协议效率

■ ARP高速缓存作用

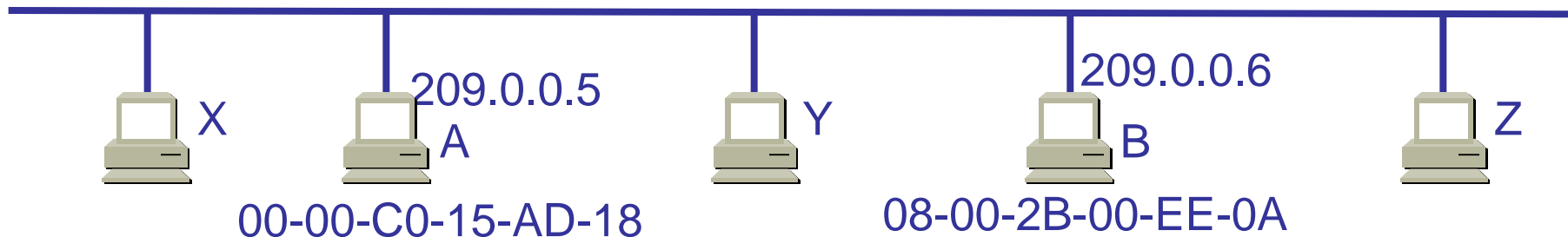
- 为了减少网络上的通信量，主机 A 在发送其 ARP 请求分组时（二层广播），就将自己的 IP 地址到硬件地址的映射写入 ARP 请求分组。
- 当主机 B 收到 A 的 ARP 请求分组时，就将主机 A 的这一地址映射写入主机 B 自己的 ARP 高速缓存中（具有生存期），主机 B 以后向 A 发送报文时就更方便了；防止接收方为解析源主机物理地址而发送ARP请求。
- 命令：arp -a 查看主机ARP缓存内容.



如何提高ARP协议效率

■ 如何提高ARP协议效率

- 在ARP请求报文中放入源主机地址映射(IP地址+MAC地址), 防止接收方为解析源主机物理地址而启动ARP协议。
- 源主机**广播**ARP请求(源IP地址+MAC地址)网络上所有主机要缓存该映射关系(具有一定生命期)。
- 新主机加入网络时, 最好主动广播自己地址映射关系, 以避免其他主机对它地址进行解析。

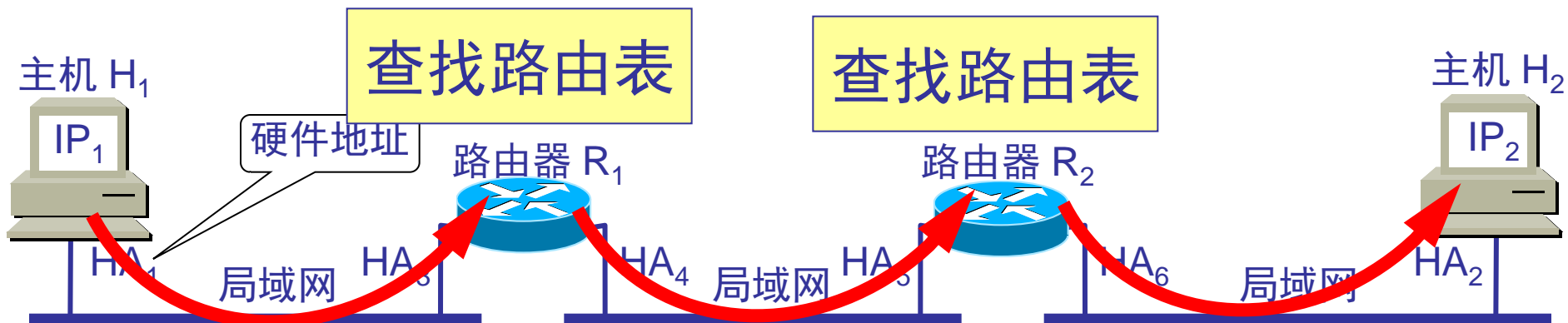




(2) 地址解析协议ARP

- 当源主机向目的主机发送IP分组时，只有知道与对方IP地址相对应的物理地址后才能在物理网络上进行帧传输；这种地址解析服务是由ARP协议提供。
- 目的节点和源节点在同一个网段
 - 源主机向本网段中广播一个ARP请求报文，报文中包含有“目的主机”的IP地址。
 - 本网段所有的主机都能接收到ARP请求报文，只有IP地址与ARP请求报文中目的IP地址相符合的主机才发送一个ARP响应报文，告知源节点自己的物理地址。
- 如果目的节点和源节点不在同一个网段, 如何处理?

举例2：源和目的节点在不同网段

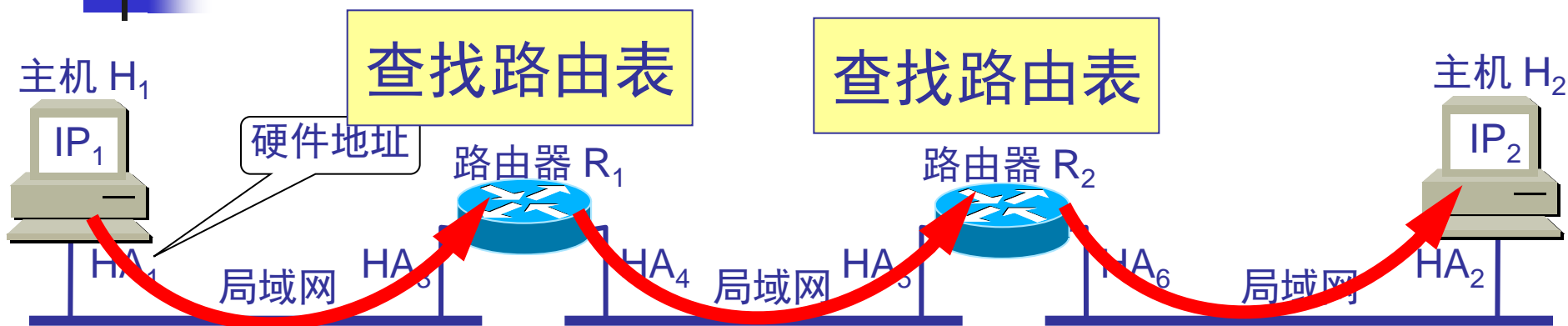


ARP请求 源到目的	MAC1	1...1	MAC4	1...1	MAC6	1...1	二层 广播帧
ARP应答 源到目的	MAC3	MAC1	MAC5	MAC4	MAC2	MAC6	二层 单播帧
数据分组 源到目的	IP1	IP2	IP1	IP2	IP1	IP2	
	MAC1	MAC3	MAC4	MAC5	MAC6	MAC2	

源：H1与H2不在一个网段，采用ARP；

路由器：H2与路由表提供IP地址不在一个网段，采用ARP

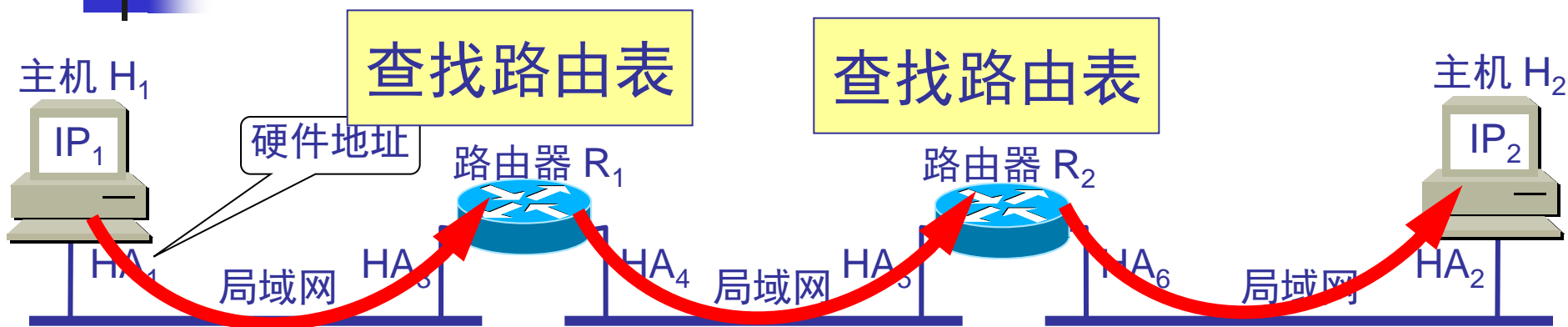
使用ARP的四种典型情况



(1) 发送方是主机，要把IP数据报发送到本网络上的另一个主机。这时用 ARP 找到目的主机的硬件地址。

(2) 发送方是主机，要把 IP 数据报发送到另一个网络上的一个主机。这时用 ARP 找到本网络上的一个路由器的硬件地址。剩下的工作由这个路由器来完成。

使用 ARP 的四种典型情况



(3) 发送方是路由器，要把 IP 数据报转发到另一个网络上的一个主机。这时用 ARP 找到本网络上的一个路由器的硬件地址。剩下的工作由这个路由器来完成。

(4) 发送方是路由器，要把 IP 数据报转发到本网络上的一个主机。这时用 ARP 找到目的主机的硬件地址。

ARP/RARP报文格式

- 物理网络类型字段：2个字节，表示发送方主机的物理网络类型；如 “1”代表以太网。
- 协议类型字段：2个字节，表示发送方使用ARP获取物理地址的高层协议类型；如 “0x0800”代表IP协议。
- 物理地址长度字段：1个字节，用于规定物理地址字段的长度；通常物理地址字段占6个字节（48位地址）。
- IP地址长度字段：1个字节，用于规定IP地址字段的长度；通常IP地址字段占4个字节（IP v4版本）。

物理网络类型
协议类型
物理地址长度
IP地址长度
操作
发送方物理地址
发送方IP地址
目的方物理地址
目的方IP地址



ARP/RARP报文格式

- 操作字段为2个字节，表示报文类型；其中：
 - 1 : ARP请求报文；
 - 2 : ARP响应报文；
 - 3 : RARP请求报文；
 - 4 : RARP响应报文。

物理网络类型

协议类型

物理地址长度

IP地址长度

操作

发送方物理地址

发送方IP地址

目的方物理地址

目的方IP地址



ARP/RARP报文格式

- 发送方物理地址字段:6个字节, 用于存放发送方的物理地址。
- 发送方IP地址字段:4个字节, 用于存放发送方的IP地址。
- 目的方物理地址字段:6个字节, 用于存放目的方的物理地址, 对于ARP请求报文, 该字段为空。
- 目的方IP地址字段:4 个字节, 用于存放目的方的IP地址。

物理网络类型
协议类型
物理地址长度
IP地址长度
操作
发送方物理地址
发送方IP地址
目的方物理地址
目的方IP地址



为什么我们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；
- 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。



ARP协议要点

- 发送方在发送ARP请求报文时，要除目的方物理地址字段外的其它字段均填写。
- 目的方通过发送ARP响应报文报告自己的物理地址。
- 发送与接收方在同一网段中，发送方的ARP请求报文可直接广播发送给本网段中目的主机，并直接获得目的IP地址对应的MAC地址。
- 如果不在同一网段，发送方实际上是想获取“下一跳”接口的物理地址。
- 从IP地址到硬件地址的解析是自动进行的，主机的用户对这种地址解析过程并不知道，即是透明的。



RARP协议

- 实质：无IP地址的主机可以通过RARP协议利用自己MAC地址来获取相应的IP地址。
- 工作原理：
 - 如果一个主机初始化后只有自己的物理地址而没有IP地址，则可通过RARP协议发送广播请求报文来请求RARP服务器告知自己的IP地址。
 - 当发送方以广播方式发送RARP请求报文时，在发送方物理地址字段和目的方物理地址字段上都填入本机物理地址。
 - RARP服务器接收到请求报文后，根据MAC+IP地址映射表，给发送方回送一个RARP响应报文，从目的方IP地址字段中带回发送方的IP地址。
- 应用场合：RARP协议主要用于无盘工作站来获取自己的IP地址；
- 注意：RARP的报文格式与ARP的相同。



本节内容提要

- 1. IP协议
 - 1.1 IP分组格式
 - 1.2 数据段的分片与组装
 - 1.3 分组转发(路由选择)
 - 1.4 选项
- 2. ARP协议
- 3. ICMP协议



引言

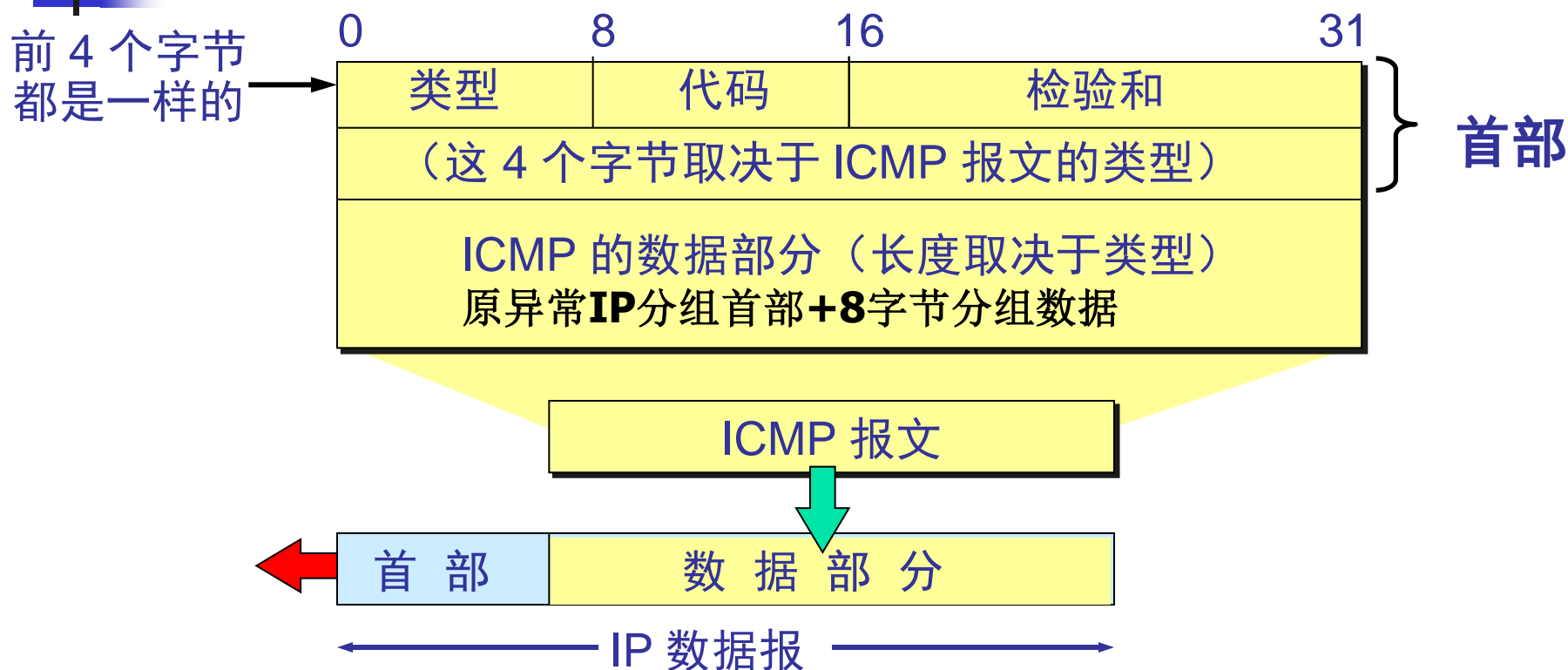
- IP协议提供了无连接的分组转发服务-**尽最大努力交付**。
 - 不提供可靠性;
 - 没有差错报告;
 - 也没有差错纠正。
- 在传送过程中, 如果发生差错或意外情况, 如:
 - 分组目的地址不可达。
 - 分组在网络中的滞留时间超过其生存期。
 - 转发节点或目的节点主机因缓冲区不足而无法接收分组。
- 通过一种通信机制, 向源节点报告异常情况, 以便源节点对差错和拥塞进行处理。
- 网际控制报文协议ICMP(Internetwork Control Message Protocol)正是提供这类**差错报告**服务、一定**拥塞控制**、**路由优化和网络信息查询**服务的协议, 以满足IP层的需求。



本节内容提要

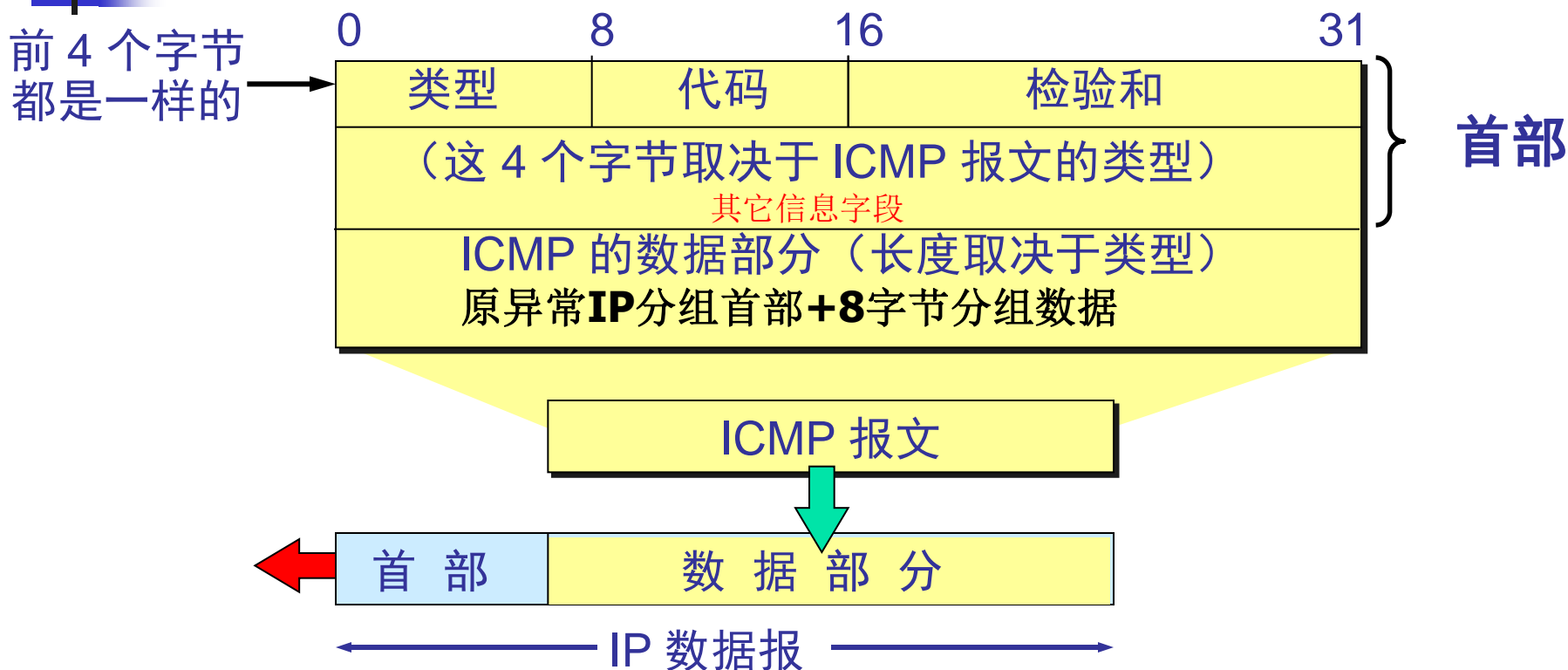
- 3. ICMP协议
 - 3.1 ICMP报文格式
 - 3.2 ICMP差错报告报文
 - 3.3 ICMP控制报文
 - 3.4 ICMP查询报文

3.1 ICMP报文格式



- 每个ICMP报文都是作为IP分组的形式在网际间转发。
- Type字段为1个字节，指出ICMP报文的类型。
- Code字段也是1个字节，提供关于报文类型的进一步信息，每种类型下子类型。

3.1 ICMP报文格式



- 校验和范围为：ICMP首部+数据；而IP分组只对首部检验。
- 数据字段：包含出错分组IP头+该分组前比特64位数据（传输层头重要信息，特别是端口号，发送序号（TCP））。

I CMP报文类型

种类	TYPE	含义
差错报告报文	3	目的不可达到
	11	超时
	12	参数出错
控制报文	4	源抑制
	5	路由重定向
查询报文	8/0	回送请求/应答
	13/14	时间戳请求/应答
	17/18	掩码请求/应答
	10/19	路由器请求/通告

- 差错报告报文和控制报文为单方向：路由器（目的节点）到源节点；
- 查询报文为双方向：发送节点与目的节点（或路由器）之间。
- 掩码请求/应答和路由器请求/通告不再使用[COME06]



本节内容提要

- 3. ICMP协议
 - 3. 1 ICMP报文格式
 - 3. 2 ICMP差错报文
 - 3. 3 ICMP控制报文
 - 3. 4 ICMP查询报文



3.2 ICMP差错报告报文

- ICMP最基本的功能：提供差错报告。
- ICMP的差错报告都是采用路由器（或目的主机）向源主机报告模式。
 - 当路由器发现IP分组差错后（数据），使用ICMP报文向该IP分组的源主机报告其差错；同时，发生差错的IP分组被丢弃掉，不再向前转发。
- 对于差错报告处理方式，ICMP协议没有严格的规定。
 - 事实上，源主机IP层收到ICMP差错报文后，还需要上层协议配合，才能决定所采取差错处理方式。
- ICMP差错报文分类：
 - 目的不可达报文：主机或路由器无法交付分组时向源发送目的不可达
 - 超时报文：路由器转发分组前TTL-1，如果为TTL=0，丢弃，向源发送
 - IP头参数出错报文：主机或路由器收到IP分组首部某字段有错误时，向源发送该报文。

3.2 ICMP差错报告报文

■ 目的不可达差错报告报文

- 路由器的主要功能是为分组选择路由并转发分组, 当从路由表上查不出分组目的IP地址对应的路由时, 则会发生目的不可达的错误。
- 路由器要向源主机发送目的不可达的ICMP报文。
- 目的不可达ICMP报文类型(Type)为 3, 并进一步细分成 13 种子类, 用代码(Code)来标识, 其它信息字段未用, 为全 0。

Type	Code	含义	发送者
3	0	网络不可达, 路由器找不到目的网络路由。	路由器
	1	主机不可达, 到达目的网络后, 路由器发送ARP请求报文不成功。	路由器
	2	协议不可达, IP分组携带的数据属于高层协议, 目的主机高层协议进程没有运行。	目的主机
	3	端口不可达, 目的主机应用进程没有运行	目的主机

3.2 ICMP差错报告报文

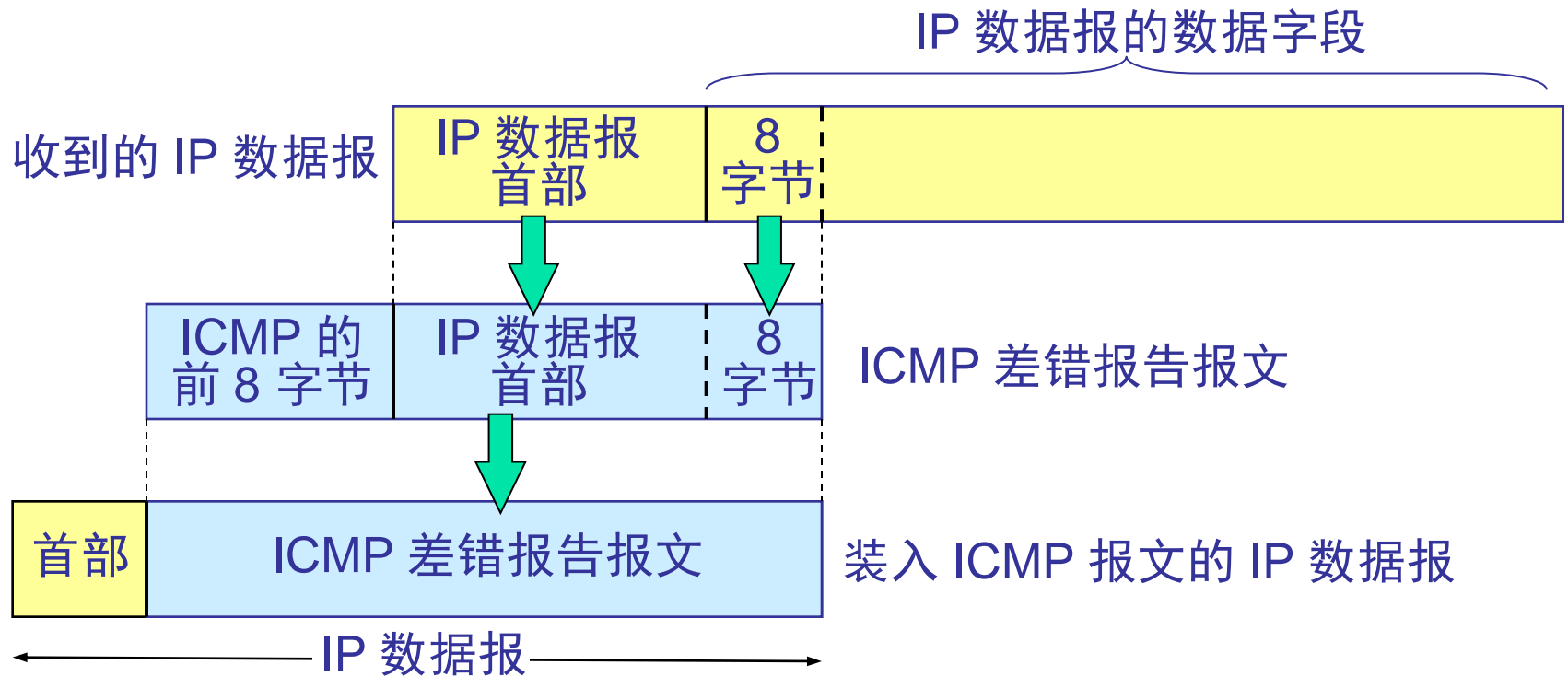
■ 超时差错报告报文

- 分组每经过一个路由器时，其生存期TTL都要减1，如果TTL递减为0时，路由器则会丢弃该分组，并向源主机发送Type=11、Code=0 的ICMP报文，报告该分组丢弃原因超时。
- 当目的主机在对分组进行组装的过程中发生组装超时，将丢弃现已收到的各个分片分组，并在第 1 个分片分组到达后向源主机节点发Type=11、Code=1的ICMP报文。
- 举例：如何找到一条从当前节点到目的节点的路径信息？
tracert www.nwpu.edu.cn (TTL=1, 2, ...)

■ 参数出错差错报文

- 当路由器或目的主机在对收到的IP分组进行处理时，如果发现在IP分组头参数中含有错误时，则将该分组丢弃，并向源主机发Type=12、Code=0的ICMP报文；
- 在ICMP报文的其它信息字段中用1个字节作为指针来指出在分组头中的差错位置（以字节为单位）。
- 原差错IP分组头和分组前64比特（8B）数据字段信息填写在数据字段。

ICMP 差错报告报文的数据字段的内容





注意事项

- 对于携带ICMP差错报告报文的IP分组，不再产生ICMP差错报文；
- 对第一个分片的所有后续分片都不发送 ICMP 差错报告报文。
- 对于组播的IP报文，不产生ICMP差错报文；
- 对于特殊地址（127. x. y. z; 0. 0. 0. 0）的IP报文，不产生ICMP差错报告报文。
- 所有ICMP差错报告报文中，只有参数出错ICMP报文包括数据字段，该字段包括源IP分组首部+源IP分组数据字段前8个字节（64比特）；
 - 这前8个字节数据提供了关于UDP和TCP端口号的信息，源主机根据协议号将差错情况报告给TCP/UCP，再通过端口号报告给应用进程。

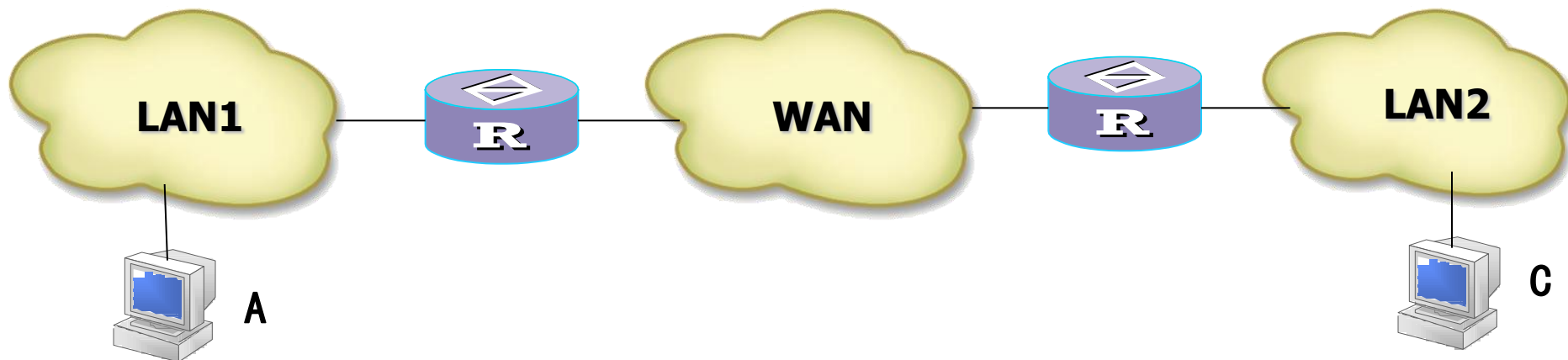


本节内容提要

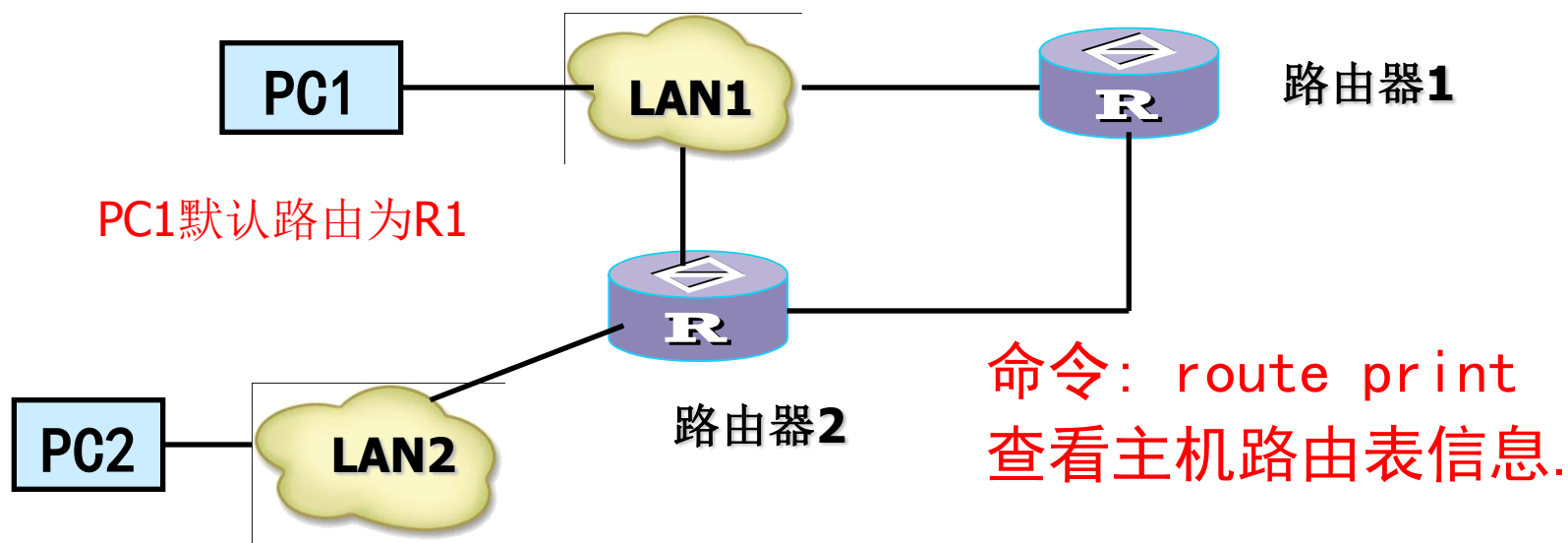
- 3. ICMP协议
 - 3.1 ICMP报文格式
 - 3.2 ICMP差错报告报文
 - 3.3 ICMP控制报文
 - 3.4 ICMP查询报文

3.3 ICMP控制报文

- ICMP控制报文主要用于拥塞控制和路由重定向。
- (1) 源抑制报文: 用于拥塞控制, 抑制源主机节点发送数据报文的速率。
 - 网络节点由于缺乏缓冲区空间而无法接收新分组时, 处理办法是丢弃(载荷脱落机制), 并向源主机节点发送Type=4、Code=0的源抑制ICMP报文。
 - 当源主机节点收到源抑制ICMP报文后, 降低其分组发送速率, 直到不再收到源抑制ICMP报文为止; 然后源主机节点又逐渐增加分组发送速率, 直到再次接收到源抑制ICMP报文为止。



- (2) 重定向报文：提供了一种路由优化控制机制，使源主机能以动态方式寻找最短路径，通常ICMP重定向报文只能在同一段中的源主机与路由器之间使用。
 - 假设PC1的默认网关为R1，PC1向PC2发送分组？
 - 当路由器R1从处于同一子网的主机PC1收到一个需转发的IP分组时，R1将检查自身的路由表信息，它选定了下一个路由器R2继续转发该分组；
 - 如果这时R1判别确认R2和PC1也处于同一子网时，R1就向PC1发送重定向ICMP报文，通知PC1将分组直接发给R2将会是一条较短的转发路径。
 - 在重定向报文的**其它信息字段**中要填入重定向的路由器(如R2) IP地址。





本节内容提要

- 3. ICMP协议
 - 3.1 ICMP报文格式
 - 3.2 ICMP差错报文
 - 3.3 ICMP控制报文
 - 3.4 ICMP查询报文



3.4 ICMP查询报文

- (1) 回送请求/应答报文:主要用于测试网络目的节点（或路由器某个接口）的可达性。
 - 源节点使用ICMP回送请求报文向某一特定的目的主机发送请求，目的节点收到请求后必须使用ICMP回送响应报文来响应对方。
 - 在许多TCP/IP实现中，提供的一种用户命令Ping便是利用这种ICMP回送请求/响应报文来测试目的可达性的。
 - 命令: `Ipconfig /all`; 查看主机网络配置信息.
 - 命令: `ping 127.0.0.1`; 测试TCP/IP协议
 - 命令: `ping 本机IP`; 测试本机网络配置信息.
 - 命令: `ping 网关IP`;

3.4 ICMP查询报文

- (2) 时戳请求/应答报文：估算源和目的节点间的分组往返时间。
 - 在报文中使用了三个时戳字段：**初始时戳字段**是源节点发送时戳请求报文的时间；**接收时戳字段**是目的节点接收到时戳请求报文的时间；**发送时戳字段**是目的节点发送时应答报文的时间。
 - 源节点首先发送时戳请求报文，然后等待目的节点返回其响应报文，并根据这三个时戳字段来估算两个节点间的报文往返时间。
- (3) **掩码请求/应答报文**：主要用于主机获取所在网络的IP地址掩码信息。
 - 主机在发送掩码请求报文时，将IP分组头中的源和目的IP地址字段的**网络号**部分设为0。
 - 网络上的目的节点（通常为路由器）接收到该请求后，填写好网络的掩码向源节点回送应答报文。
- (4) **路由器询问/通告报文**：用于查找连接到源主机或路由器上的正常工作的路由器情况，并从中选择一个作为源主机（或路由器）的默认路由。



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。
- 工作原理
 - ping 使用了 ICMP 回送请求与回送回答报文。
 - 源向目的发送ICMP 回送请求，目的接收到后给源一个回送回答报文
 - 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。
- 返回结果信息
 - 发送报文数、接收报文数、丢失报文数；
 - 往返时间最小值、最大值和平均值。



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
```



ICMP的应用举例

Tracert命令

- Tracert: 跟踪一个分组从源到目的的路径。
- 工作原理: 源主机向目的发送一连串的IP分组, 分组中封装的是无法交付的UDP用户数据报。
 - 源首先向目的发送第1个UDP用户数据报, 并且TTL=1; 当该分组到达路由器R1时, R1先接收, 然后 $TTL-1=0$, 在不再转发, 直接丢弃, 并向源发送一个**ICMP超时报文**, 源节点记下R1的入口IP地址;
 - 源首先向目的发送第2个UDP用户数据报, 并且TTL=2; 当该分组到达路由器R2时 (TTL=1), R2先接收, 然后 $TTL-1=0$, 在不在转发, 直接丢弃, 并向源发送一个ICMP超时报文, 源节点记录下R2的入口IP地址;
 - 依次类推, 当最后一个分组到达目的节点时, TTL=1, **主机不转发, TTL也不减1**, 由于IP分组中封装的**UDP用户数据报无法交付**, 这时, 目的主机向源节点发送一个**“目的不可达”差错报告报文**。



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.
```



本节小结

- 1. IP协议
- 2. ARP协议
- 3. ICMP协议



复习&预习&作业

- 复习

教材PP171-178

- 预习

教材PP178-185

- 作业

教材PP199: 6, 7, 9