



计算机通信与网络

Computer Communications & Networks

第5章 网络层与网络互连

南京邮电大学通信与信息工程学院

“计算机通信与网络” 国家精品课程组

大纲要求

Requirements



- 1** 掌握**IP**协议的原理和工作机制
- 2** 掌握**Internet**编址技术，掌握子网划分技术
- 3** 掌握常用路由算法，掌握路由协议在网络通信中的作用
- 4** 了解**IP**协议的发展，掌握**IPv6**的主要特点

内容纲要

Contents Page



01 网络层概念

02 网络互连

03 差错与控制报文协议(ICMP)

04 子网编址及无分类编址与CIDR

05 因特网的路由选择协议

06 IP组播

07 移动IP

08 专用网络互连(VPN和NAT)

过渡页

Transition Page



01 网络层概念

02 网络互连

03 差错与控制报文协议(ICMP)

04 子网编址及无分类编址与CIDR

05 因特网的路由选择协议

06 IP组播

07 移动IP

08 专用网络互连(VPN和NAT)



为什么需要网络互连



网络互连需要解决的问题



因特网的分层模型

网络互联

5.1.1 为什么需要网络互联

网络互联/网络互连——用线路和互连设备将采用各种不同低层（网络层以下）协议的网络连接起来。

- 互连 (inter-connection) : 强调物理连接
- 互联 (inter-networking) : 强调逻辑连接

网络互连设备:

物理层互连设备: 集线器 (hub) , 转发器(repeater)

数据链路层互连设备: 网桥(bridge)或交换机 (switch)

网络层互连设备: 路由器(router)

传输层及以上功能层互连设备: 网关(gateway)

5.1.1 为什么需要网络互连

在1970年代中期已存在大大小小相互独立的多种物理网络

已有的异种网络客观存在

没有哪种网络技术可以满足所有需求

全球范围的通信需求

各物理网络中的用户希望能够相互通信，不受物理网络边界限制

网络互联的作用

5.1.1 为什么需要网络互联

隐藏底层细节，使互联起来的网络可以看成是单一的虚拟网络，所有计算机都与它相连，而不管实际的物理连接如何。

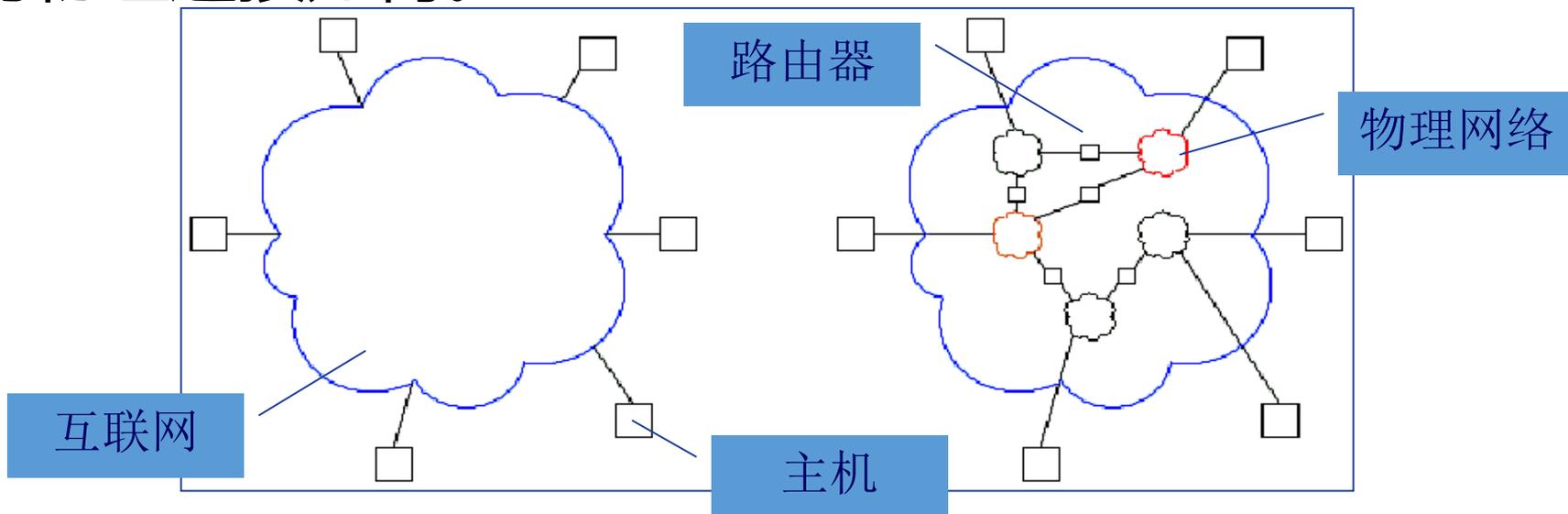


图 5-1 (a) 用户观点的互联网; (b) 互联网的实际连接示例

5.1.2 网络互联需要解决的问题

IP层主要功能是负责为不同物理网络上的主机提供通信服务。需要解决若干问题：

如何屏蔽异种网络的差异

- 统一网络地址

- 统一数据格式

如何实现全网的数据传输和交换

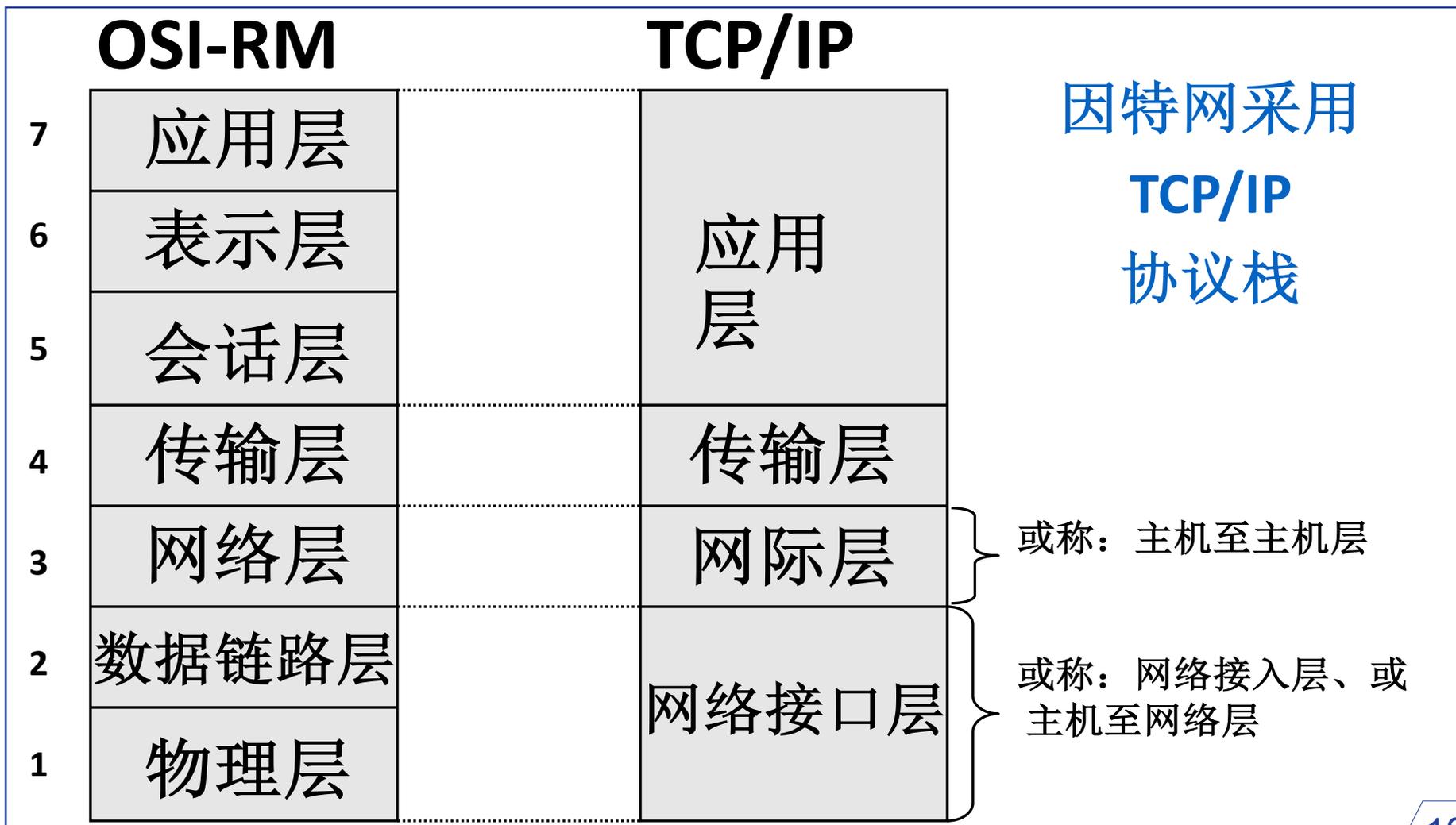
- 网络之间的数据包寻径、转发

- 路由表的产生和动态刷新

- 差错处理

在全网内的组播

5.1.3 因特网的分层模型



5.1.3 因特网的分层模型

OSI_RM

5~7

应用层

Telnet

FTP

SMTP

DNS

SNMP

4

传输层

TCP

UDP

3

IP层

ICMP

IP

ARP

RARP

2

网络接口层

局域网

广域网

其它

1

过渡页

Transition Page



01 网络层概念

02 网络互连

03 差错与控制报文协议(ICMP)

04 子网编址及无分类编址与CIDR

05 因特网的路由选择协议

06 IP组播

07 移动IP

08 专用网络互连(VPN和NAT)



分类的IP地址

IP地址的分配与使用

IP地址到物理地址的映射

逆地址解析协议

IP数据报

无连接的数据报传送



IP协议



- 支持TCP/IP软件隐藏物理网络细节，使构成的互联网看起来是一个统一实体的基础：IP编址和IP数据报。
 1. 分类的IP地址
 2. IP地址的分配与使用
 3. IP数据报
 4. 因特网地址到物理地址的映射
 5. 无连接的数据报传送
 6. 差错与控制报文(ICMP)
 7. 子网编址
 8. 无分类编址与CIDR



网络互连



- 因特网IP层的服务是一个尽最大努力传送的、不可靠、无连接的数据报存储转发系统。
- 支持TCP/IP软件隐藏物理网络细节，使构成的互联网看起来是一个统一实体的基础：IP编址和IP数据报。
- IP层协议包括 IP协议、Internet控制报文协议 ICMP、地址解析/逆地址解析协议 ARP/RARP



5.2.1 分类的IP地址

IP地址：给每个连接在因特网上的主机分配一个在全世界范围是惟一的32 bit 的标识符。

网络地址实际上是表示主机到网络的一个连接。

IP地址由因特网名字与号码指派公司ICANN 分配。

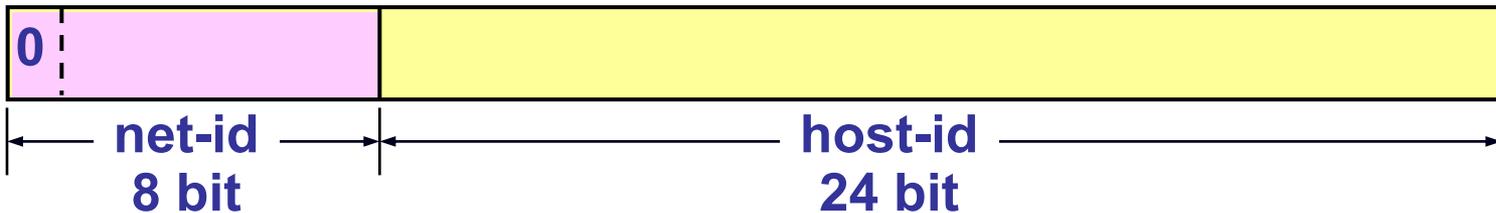
IP地址采用分层的地址结构：**前缀 + 后缀**



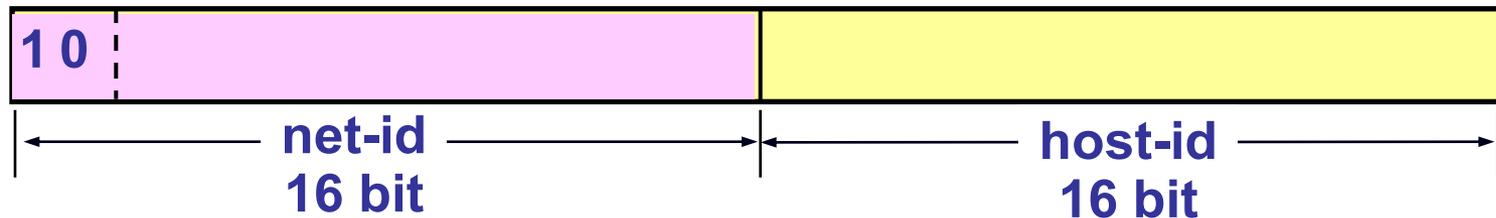
网络号和主机号

5.2.1 分类的IP地址

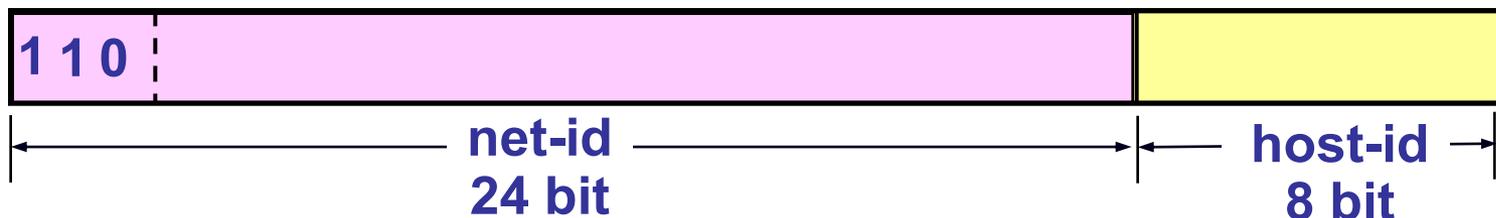
A 类地址



B 类地址



C 类地址



D 类地址



E 类地址





点分十进制记法

5.2.1 分类的IP地址

机器中存放的 IP 地址
是 32 bit 二进制代码

10000000000010110000001100011111

将每 8 bit 的二进制数
转换为十进制数

10000000 00001011 00000011 00011111
128 11 3 31

采用点分十进制
记法提高可读性

128.11.3.31

分类IP地址是**自标识的**(self-identifying), 仅从地址本身就能够确定前后缀之间的边界。



ABC类地址使用范围

5.2.1 分类的IP地址

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	网络最大主机数
A	126 ($2^7 - 2$)	1	126	16,777,214
B	16,384 (2^{14})	128.0	191.255	65,534
C	2,097,152 (2^{21})	192.0.0	223.255.255	254



特殊的IP地址

5.2.1 分类的IP地址

特殊IP地址	含义
[网络号, 0]	指定的网络地址
[网络号, < -1 (全1) >]	定向广播地址
[< -1 >, < -1 >]	有限广播地址
[0 ,0]	本主机
[0 ,主机号]	本网络上的某主机
[127 ,<任意>]	回送地址



IP地址分配

5.2.2 IP地址的分配与使用

IP 地址是一种分等级的地址结构

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。

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和路由查找时间。



IP地址分配

5.2.2 IP地址的分配与使用

IP 地址是标志主机或路由器的一条网络链路的接口。

当一个主机(路由器)同时连接到两个网络上时, 该主机就必须同时具有两个IP 地址, 且其网络号(net-id)必须是不同的。



IP地址分配

5.2.2 IP地址的分配与使用

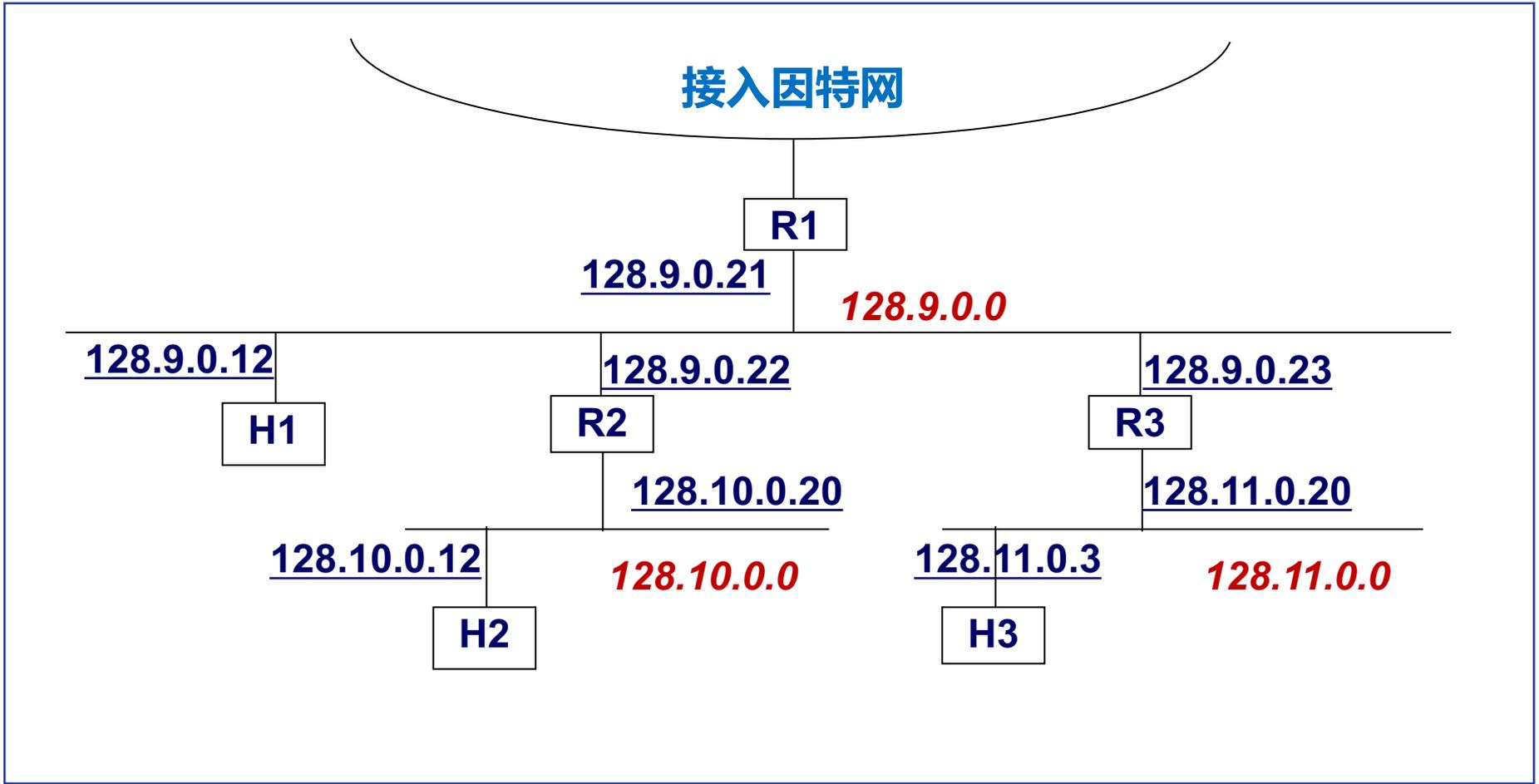
用转发器或网桥(交换机)连接起来的若干个局域网，虽然在物理上是多个网络，但在逻辑上仍为一个网络，这些局域网都具有相同的网络号(net-id)。

所有分配到网络号的网络，不管是范围很小的局域网，还是可能覆盖很大地理范围的广域网，都是平等的。



IP地址分配示例

5.2.2 IP地址的分配与使用





概述

5.2.3 IP地址到物理地址的映射

不管网络层使用的是什么协议，在实际物理网络的链路上上传送数据帧时，最终还是必须使用硬件地址。

地址统一(网络层以上使用IP地址)的代价是需要建立IP地址和物理地址之间的映射。

地址解析协议ARP为网络层(IP)地址 和 数据链路层使用的任何类型的地址提供动态映射。



ARP cache

5.2.3 IP地址到物理地址的映射

每一个主机中都设有一个 **ARP 高速缓存表**(ARP cache), 里面有所在的局域网上各主机和路由器的 IP 地址到硬件地址的映射。

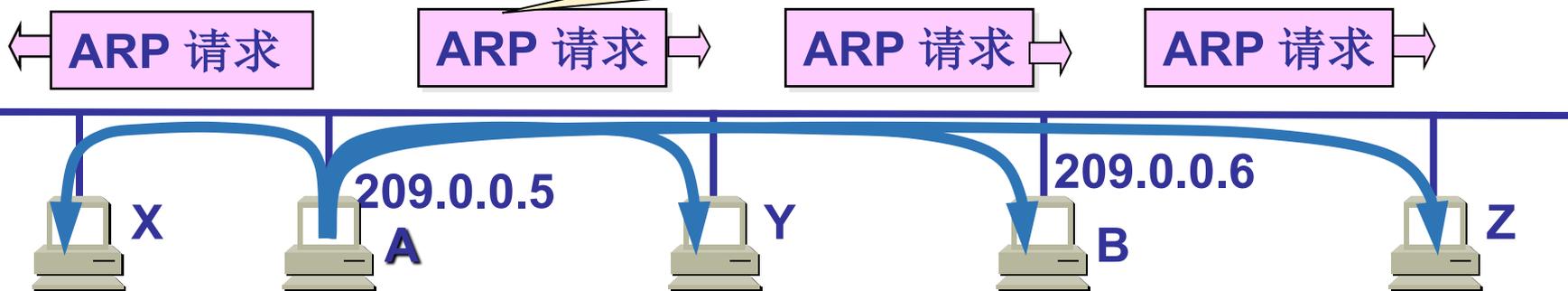
当主机 A 欲向本局域网上的某个主机 B 发送 IP 数据报时, 就**先在其 ARP 高速缓存表中查看**有无主机 B 的 IP 地址。如有, 就可查出其对应的硬件地址, 如果没有, 则广播发送一个**ARP 请求报文**。

以太网上的示例

5.2.3 IP地址到物理地址的映射

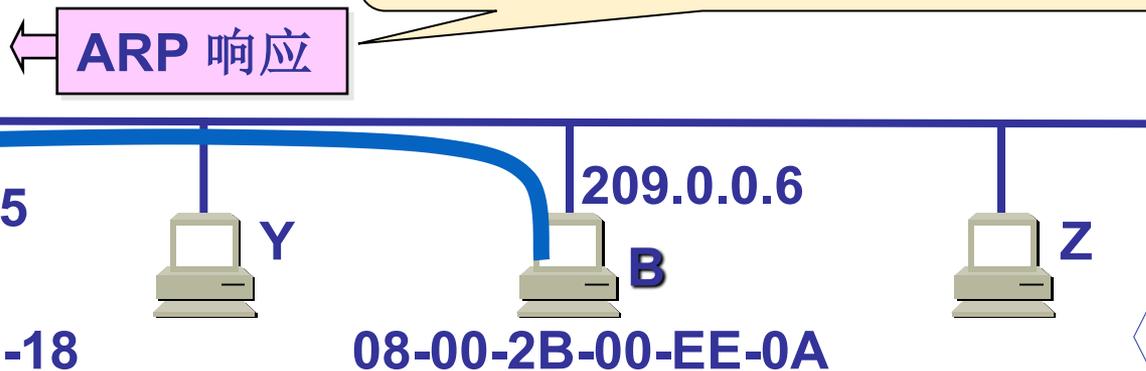
主机 A 广播发送
ARP 请求报文

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 响应报文

我是 209.0.0.6
硬件地址是 08-00-2B-00-EE-0A





5.2.3 IP地址到物理地址的映射

ARP 是解决**同一个物理网络上的**主机或路由器的 IP 地址和硬件地址的映射问题。

如果目的主机和源主机不在同一个局域网，先选路，再通过 ARP 找到下一跳路由器在本 LAN 的硬件地址，然后把分组发送给它，让它把分组转发给下一个网络。

从IP地址到硬件地址的解析是**自动**进行的，是由主机所运行的TCP/IP内核实现的。



ARP报文格式

5.2.3 IP地址到物理地址的映射

2字节 硬件类 型	2字节 协议 类型	1字节 硬件地 址长度	1字节 协议 地址长 度	2字节 操作 类型	发送方 硬件地 址	发送方 协议地 址	目标 硬件地 址	目标 协议地 址
1: 以 太网 6: IEEE 802网 络 15:帧 中继	0x08 00: IP	6: 以太 网硬件 地址长 度	4: IP 地址 长度	1: ARP请求 2: ARP响应 3: RARP请求 4: RARP响应				



RARP

5.2.4 逆地址解析协议

逆地址解析协议RARP使只知道自己硬件地址的主机能够知道其 IP 地址。

其因为较限于IP地址的运用以及其他的一些缺点，因此渐为更新的[BOOTP](#)或[DHCP](#)所取代。因此目前RARP 协议已很少使用。

[RARP](#)由应用软件来实现，运行RARP应用软件的主机称为[RARP服务器](#)。



解析过程

5.2.4 逆地址解析协议

一个新启动的工作站可以在引导程序的控制下，通过网卡广播其以太网地址，并请求获得自己的IP地址。被授权提供RARP服务器处理请求并发送应答。

RARP服务器发现这个请求后，在其配置文件中找到以太网地址，填写目标协议地址字段，并把报文类型从“请求”改为“应答”，直接把应答发给提出请求的计算机。



5.2.5 IP数据报

因特网的基本传送单元是**IP数据报**(或IP分组)。

一个 IP 数据报由首部 (报头) 和数据两部分组成。

首部的前一部分是**固定部分**，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 数据报首部的固定部分和可选字段长度之和最大为**60 字节**。



版本

5.2.5 IP数据报

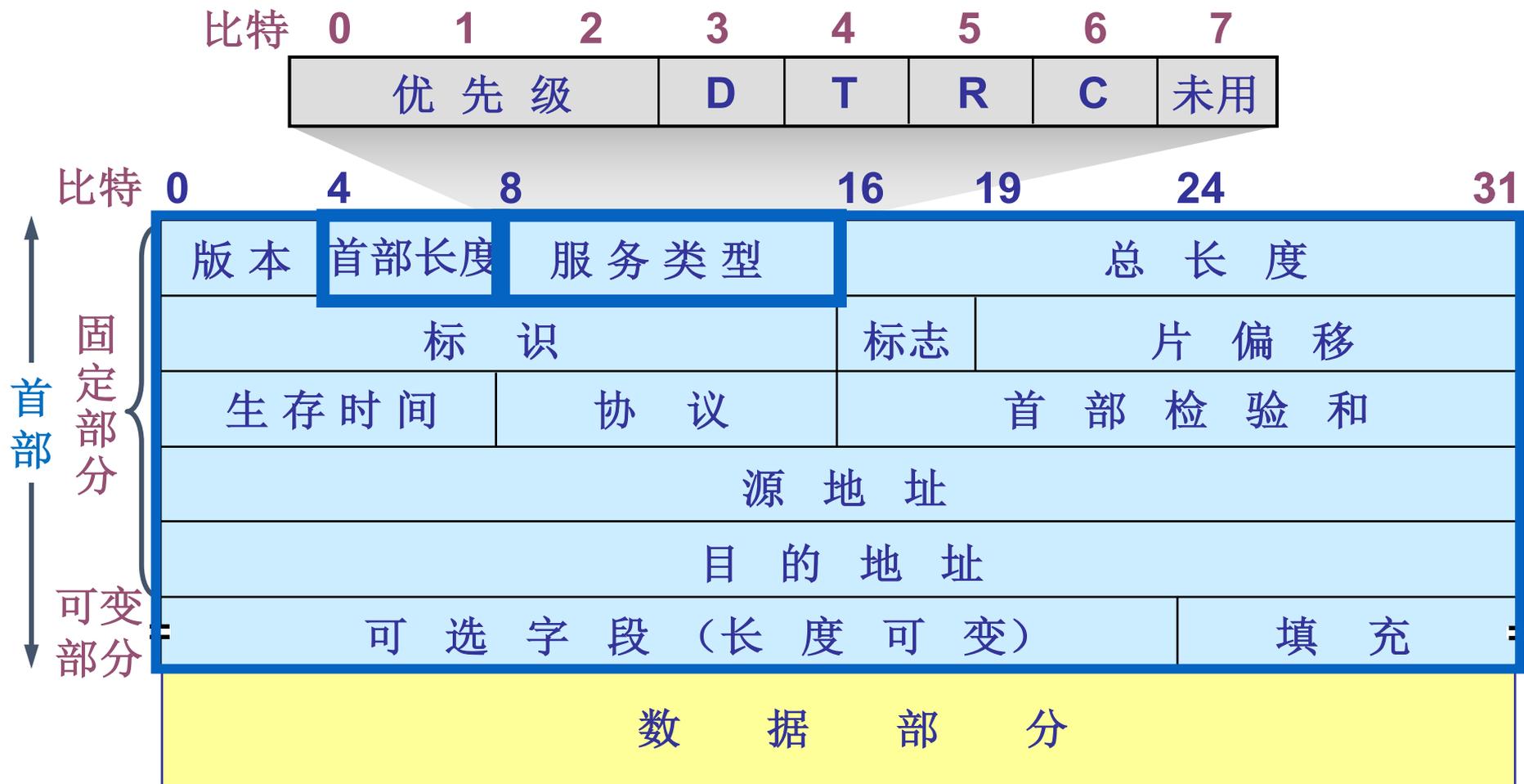


版本——占 4 bit, 指IP协议的版本



首部长度和服务类型

5.2.5 IP数据报



首部长度——占 4 bit, 最大值是60字节



总长度

5.2.5 IP数据报



总长度——首部和数据之和的长度，单位为字节
总长度必须不超过最大传送单元 **MTU**。



标识

5.2.5 IP数据报



标识(identification) ——一个计数器，用来产生数据报的标识。当数据报需要分片时，此标识表示同一个数据报的分片。



标志

5.2.5 IP数据报



标志(flag) 占 3 bit, 1位保留, DF位用来表示数据报是否允许分片, MF位表示是否有后续分片。



片偏移

5.2.5 IP数据报

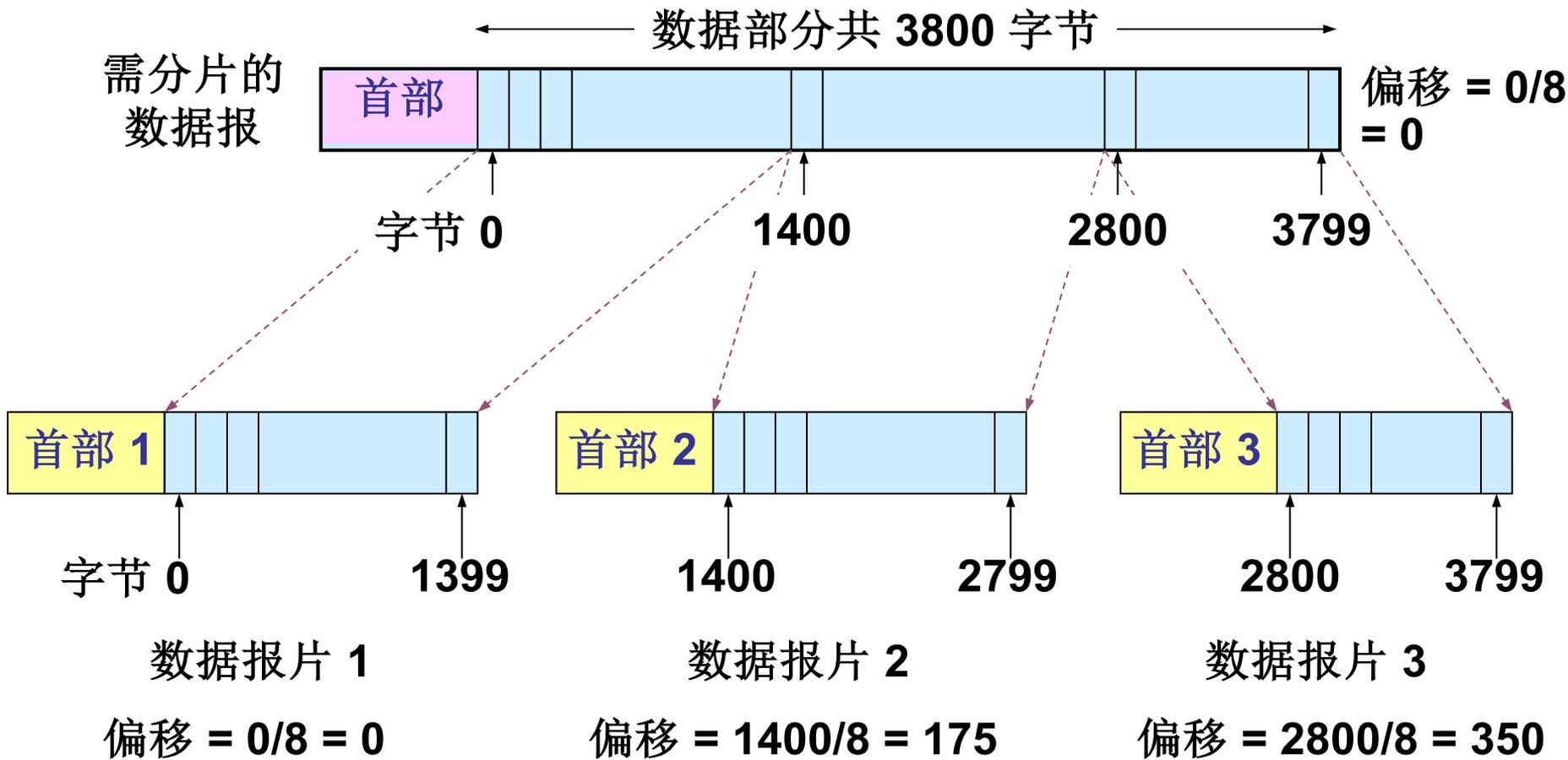


片偏移(13 bit)——分片在原分组中的相对位置，以 8 个字节为偏移单位。



IP数据报分片举例

5.2.5 IP数据报





生存时间

5.2.5 IP数据报

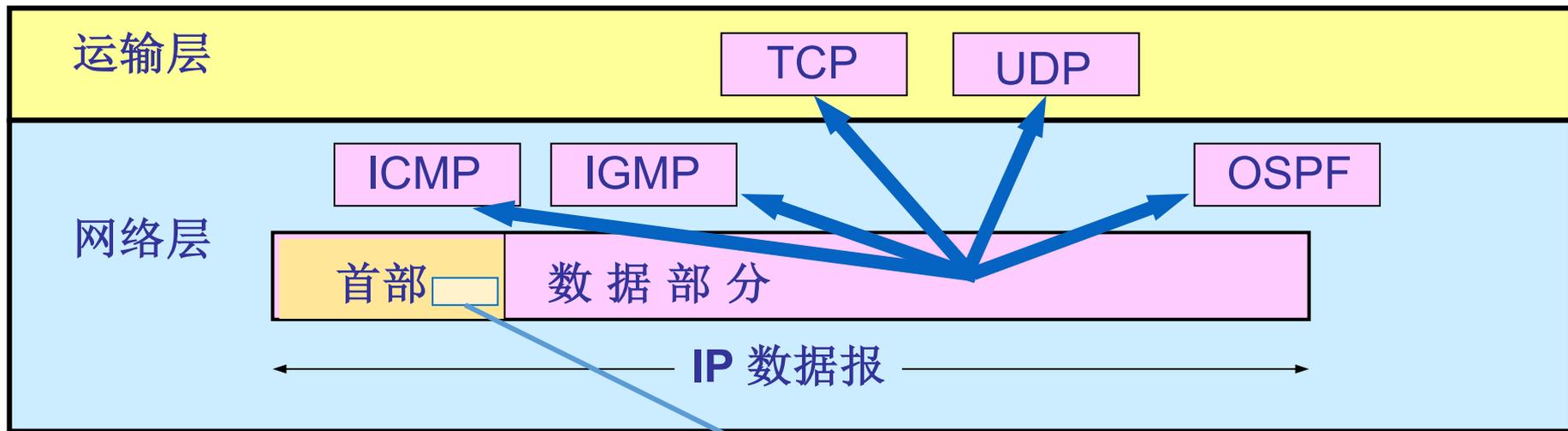


生存时间(Time To Live)——数据报在网络中的寿命，其单位为秒。在目前的实际应用中，以“跳”为单位。



协议

5.2.5 IP数据报



协议(8 bit)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相应处理过程

协议字段值	1	2	3	4	6	8	17	88	89
协议名	ICMP	IGMP	GGP	IP	TCP	EGP	UDP	IGRP	OSPF



首部校验和

5.2.5 IP数据报

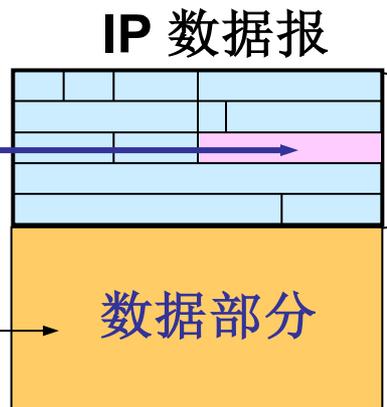
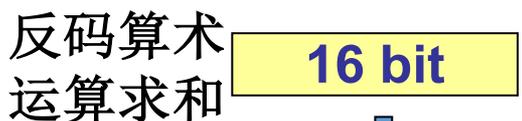


首部校验和(16 bit)——只检验数据报的首部
不包括数据部分，采用简单的计算方法。

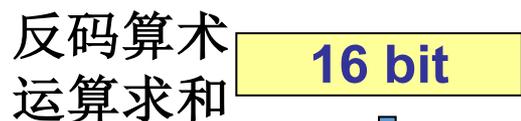
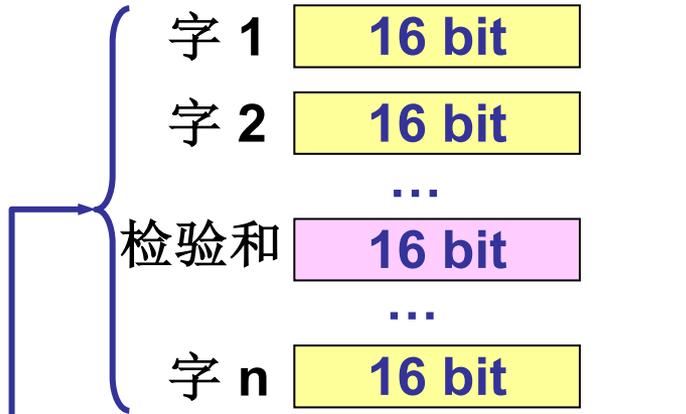


首部校验和

5.2.5 IP数据报



数据部分
不参与校验和的计算



若结果为 0, 则保留;
否则, 丢弃该数据报



地址

5.2.5 IP数据报



源地址和目的地址都各占 4 字节



可选字段

5.2.5 IP数据报

IP 首部的可选字段，用来支持排错、测量以及安全等措施。

选项字段长度可变，从 1 个字节到 40 个字节不等。

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。

这增加了每一个路由器处理数据报的开销。

实际上这些选项较少被使用。



IP包解析举例

5.2.5 IP数据报

例 在IP地址为10.10.1.95的主机上用网络监听工具监测网络流量，获取的一个IP包的前28字节用十六进制表示如下：

45 00 00 47 E6 EE 00 00 67 11

19 2A 75 4E D2 D6 0A 0A 01 5F

A4 CA 0D 4B 00 33 6B 26

请解析IP包各字段。



IP包解析举例

5.2.5 IP数据报

0	4	8	16	19	24	31
版本4	首部长5	服务类型0	总长度 00 47			
标识 E6 EE			标志0	片偏移量0		
生存时间 67		协议 11	首部校验和 19 2A			
源站IP地址 75 4E D2 D6						
目的站IP地址 0A 0A 01 5F						
数据 A4 CA 0D 4B 00 33 6B 26						
...						



路由器功能

5.2.6 无连接的数据报传送

路由器在网际互连中的功能：

网络分段，实现子网的划分

提供不同类型网络的互联

隔离广播风暴

实现子网之间的信息传输

提供安全访问的机制

支持网络层的特殊服务



路由器的组成

5.2.6 无连接的数据报传送

路由器实际上就是一种通信专用的计算机，具有和计算机类似的结构。

处理器CPU：微处理器或RISC微处理器

存储器：ROM(引导软件, 备份IOS)、FLASH(IOS)、NVRAM (Startup-Config)、RAM(内存, 暂时存放OS和数据)

接口：控制台接口

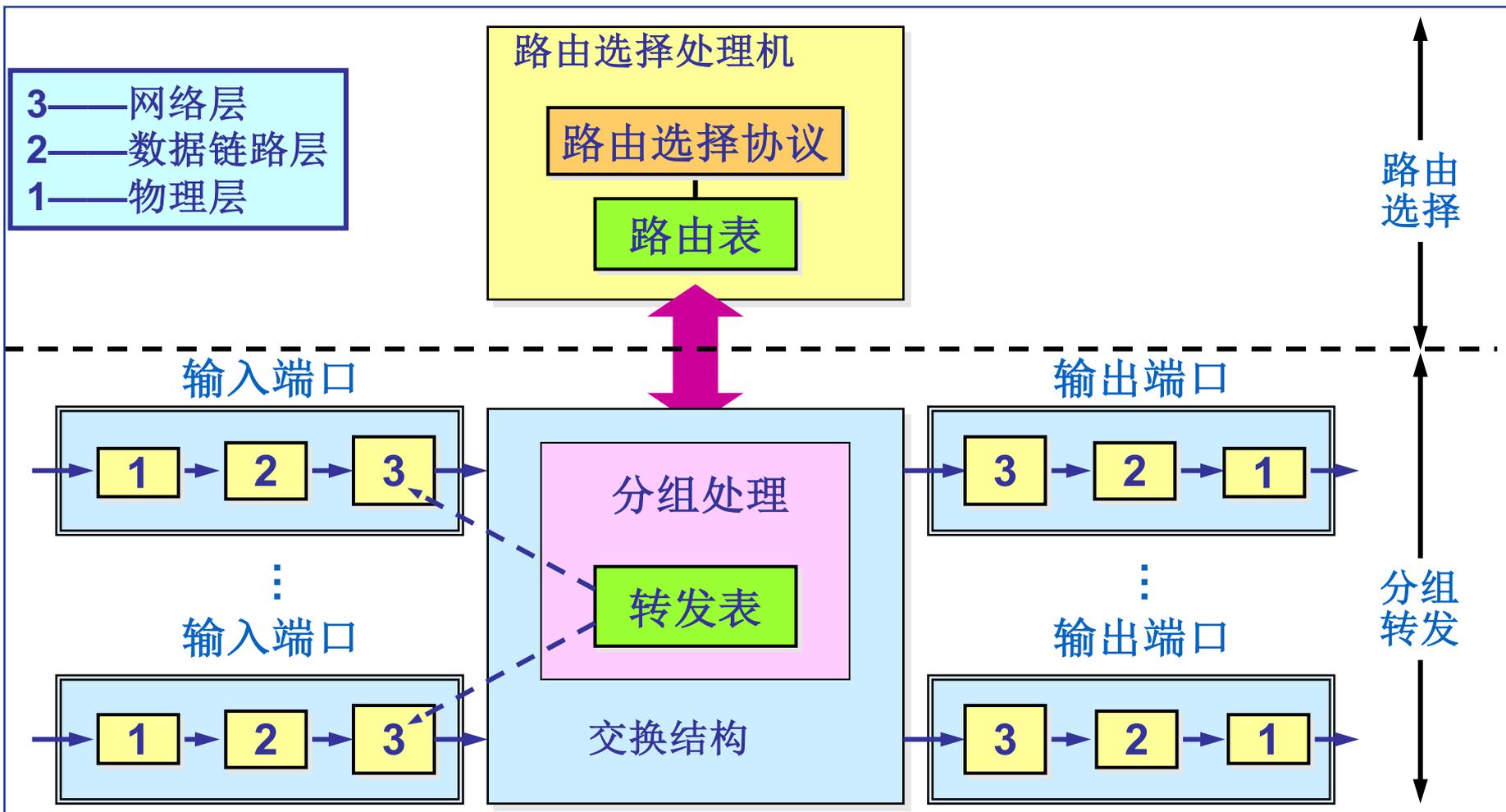
局域网接口：以太网、令牌环网、光纤网

广域网接口：ATM、X25、DDN、FRN、ISDN

逻辑接口：LOOP BACK, NULL, CHANNEL

典型的路由器结构

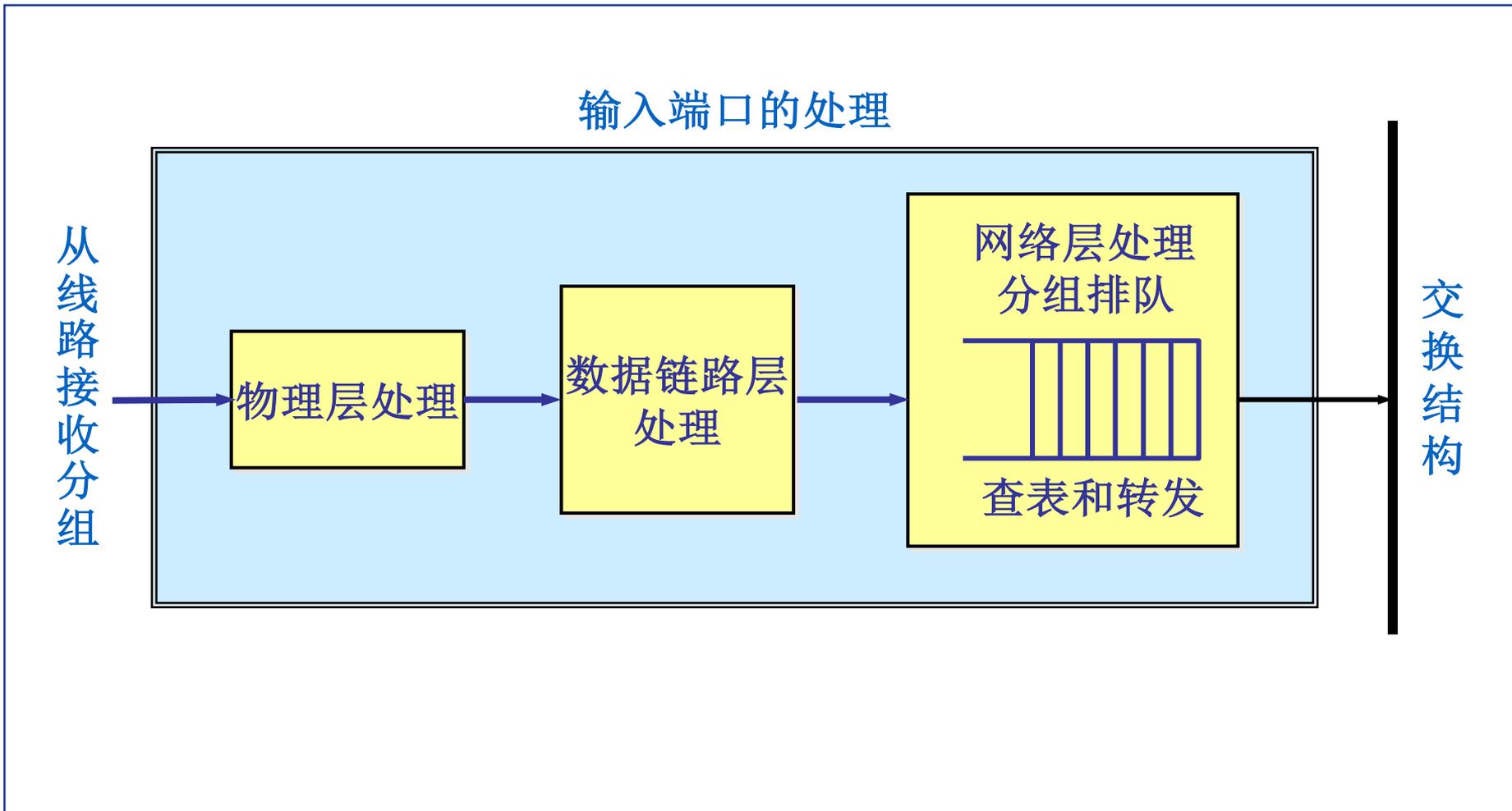
5.2.6 无连接的数据报传送





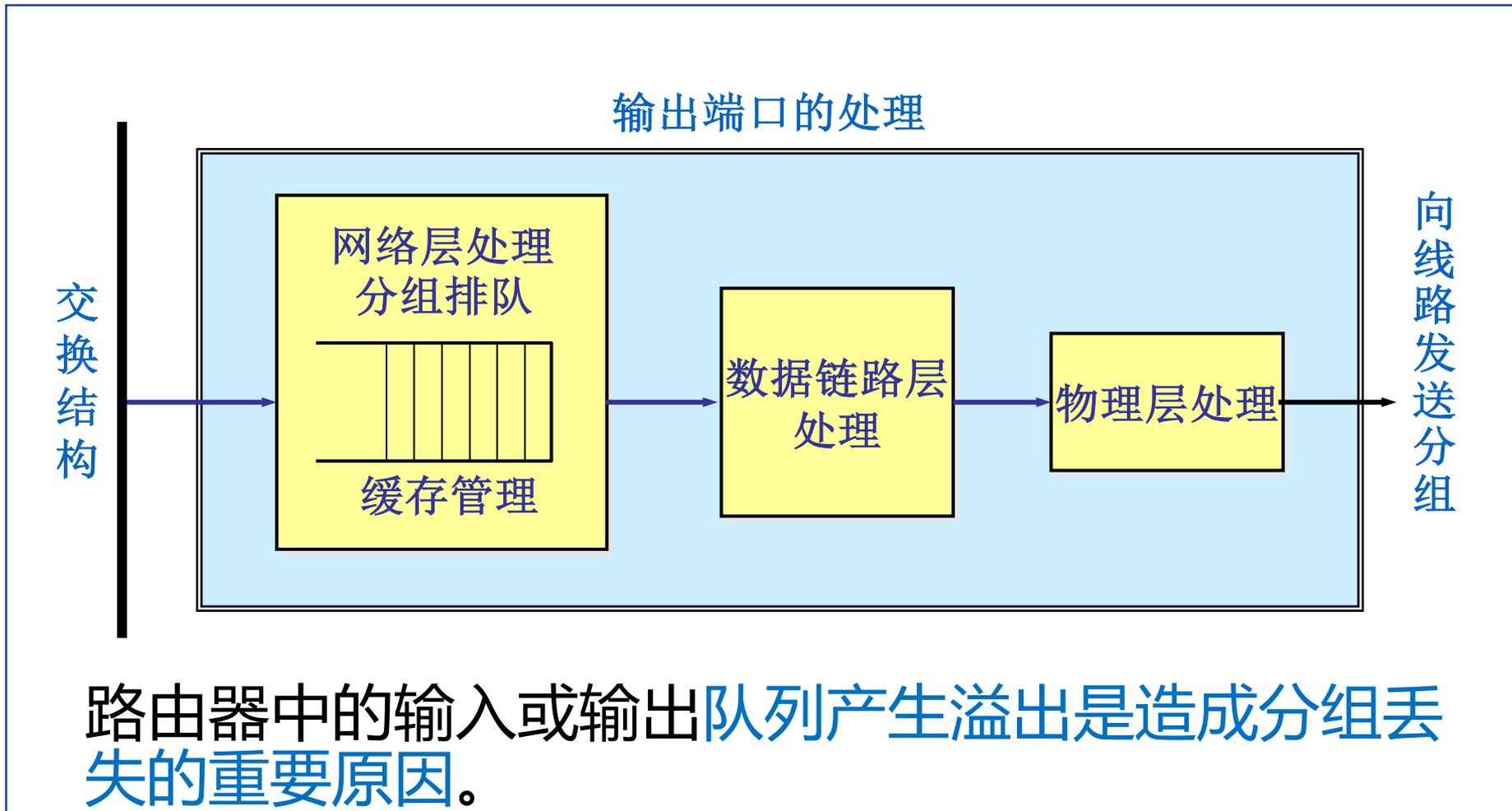
输入端口的处理

5.2.6 无连接的数据报传送



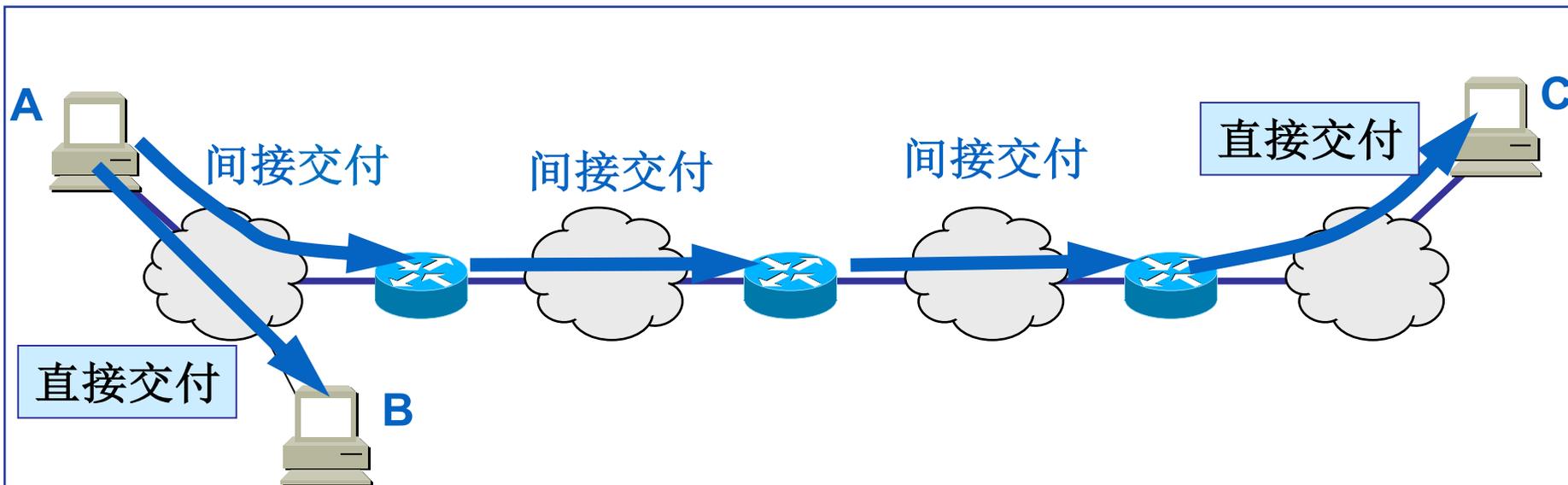
输出端口的处理

5.2.6 无连接的数据报传送



直接与间接交付

5.2.6 无连接的数据报传送



**直接交付不需要使用路由器
而间接交付就必须使用路由器**



IP包转发算法

5.2.6 无连接的数据报传送

采用最初的分类型编址方案时的IP数据报转发（数据报DG，路由表T）

从数据报DG中取出目的站IP地址 I_D ；

if 表T中含有 I_D 的一个特定路由，则

把DG发送到该表项指明的下一跳

return.

else 根据分类地址规则，从 I_D 中提取出网络前缀，得到网络地址N；

if N与任何一个直接相连的网络地址匹配，则

通过该网络把DG直接交付给目的站

else if 表T中包含一个到网络N的路由，则

把DG发送到该表项指明的下一跳

else if 表T中包含一个默认路由，则

把DG发送到该表项指明的下一跳（默认路由器）；

else

向DG的源站发送一个目的不可达差错报



对传入IP包的处理

5.2.6 无连接的数据报传送

主机收到数据报：

从数据报DG中取出目的站IP地址 I_D ;

if I_D 与主机的IP地址(单播或广播地址)匹配, 则

接受DG, 根据DG中的协议指示将DG的数据交给高层协议软件处理;

else

丢弃DG;



对传入IP包的处理

5.2.6 无连接的数据报传送

路由器收到数据报:

从数据报DG中取出目的站IP地址 I_D ;

if (I_D 与路由器的任一个物理网络连接的IP地址匹配) || (I_D 是受限IP广播地址, 或是路由器的某直连网络的定向IP广播地址), 则

接受DG, 根据DG中的协议指示将DG的数据交给相应协议软件做进一步处理;对于定向广播,在指定的网络上广播该数据报

else

把DG首部中的生存时间TTL减1;

if TTL为0,则丢弃DG,向DG的源站发送一个超时差错报告;

else 重新计算校验和, 并转发数据报

过渡页

Transition Page



01 网络层概念

02 网络互连

03 差错与控制报文协议(ICMP)

04 子网编址及无分类编址与CIDR

05 因特网的路由选择协议

06 IP组播

07 移动IP

08 专用网络互连(VPN和NAT)



ICMP报文



目的不可达报文



超时报文



回应请求与应答报文



ICMP的作用

5.3.1 ICMP报文

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 允许主机或路由器**向源站报告差错情况和异常情况**，还包括**提供信息的功能**。

ICMP 不是高层协议，而是 IP 层的协议。



ICMP报文类型

5.3.1 ICMP报文

ICMP报文分为**两大类**：**差错报告报文**和**提供信息的报文**。

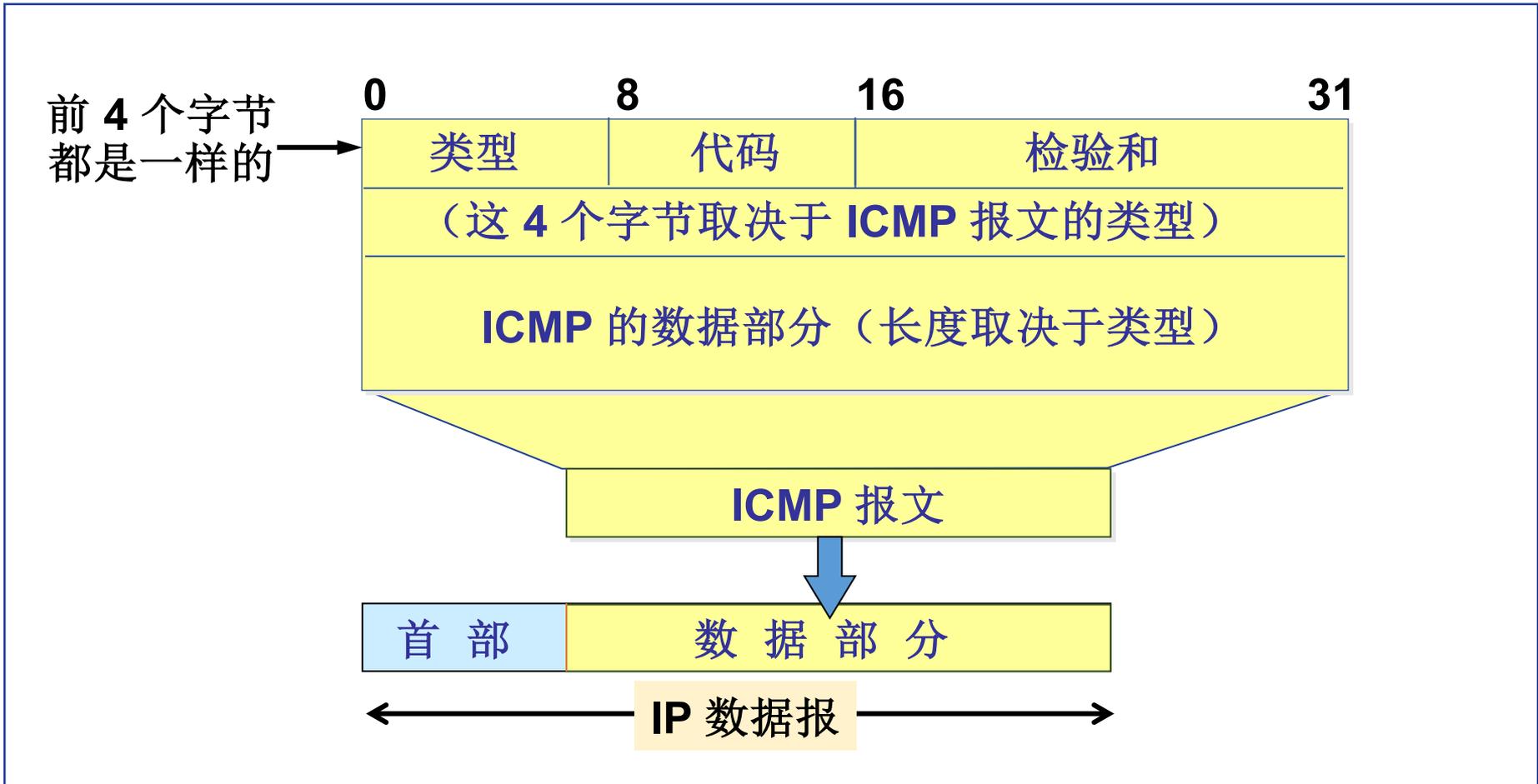
差错报告报文有：目的不可达报文、超时报文、源抑制报文等。

提供信息的报文有：回应请求与应答报文、时间戳请求与应答报文、地址掩码请求与应答报文等。



ICMP报文格式

5.3.1 ICMP报文





ICMP报文的类型字段

5.3.1 ICMP报文

类型字段	ICMP 报文类型
0	Echo Reply
3	Destination Unreachable
4	Source Quench
5	Redirect (change a route)
8	Echo Request
9	Router Advertisement
10	Router Solicitation
11	Time Exceeded for a Datagram
12	Parameter Problem on a Datagram
13	Timestamp Request
14	Timestamp Reply
15	Information Request (obsolete)
16	Information Reply (obsolete)
17	Address Mask Request
18	Address Mask Reply

ICMP报文的代码字段

5.3.1 ICMP报文

代码字段包含了有关本类型的更多信息。如在目的不可达类型的报文中，可能的代码值为：

码值

含义

0
1
2
3
4
5
6
7
8
9
10

网络不可达

主机不可达

协议不可达

端口不可达

需要分片并且**DF**置位

源路由失败

目的网络未知

目的主机未知

源主机被隔离

出于管理的目的禁止了与目的地网络的通信

出于管理的目的禁止了与目的地主机的通信

通常意味着选路失败

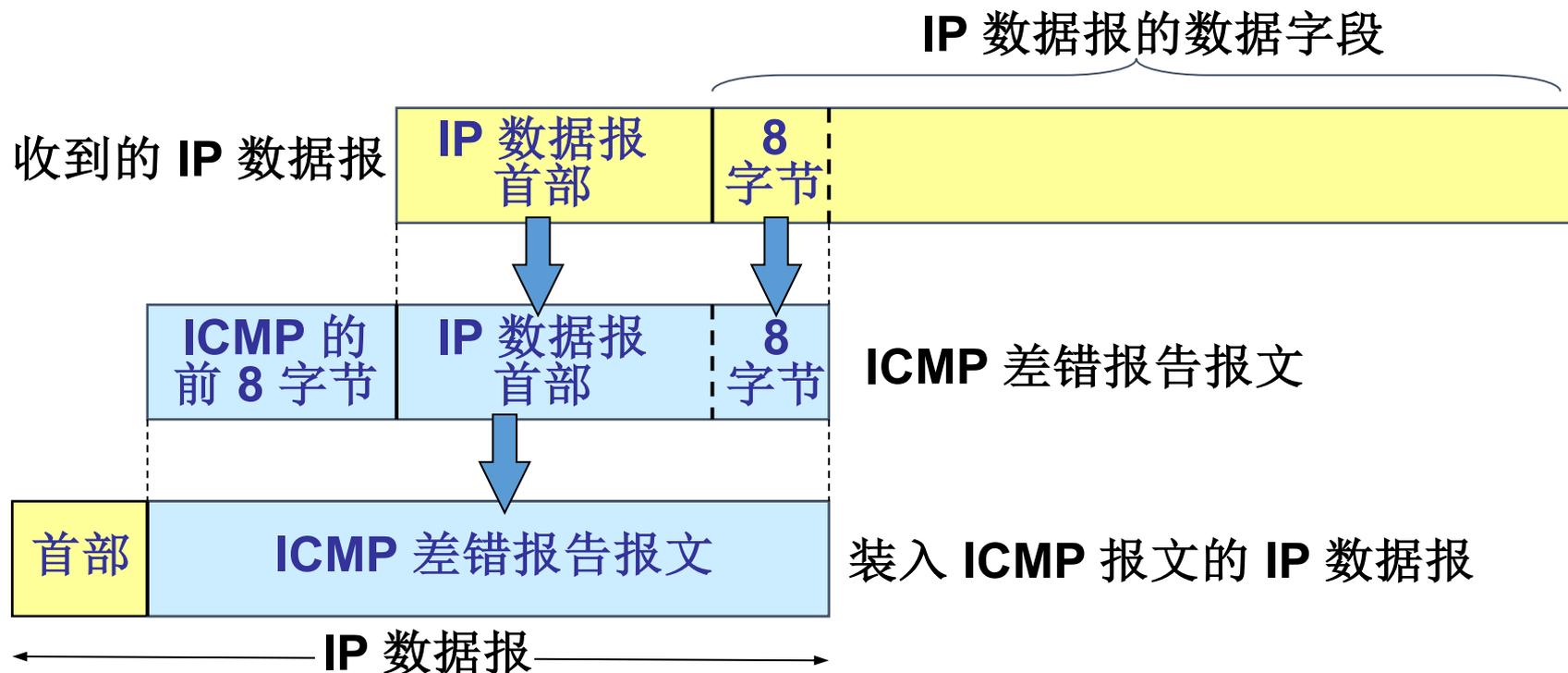
通常意味着交付失败



ICMP报文格式

5.3.1 ICMP报文

ICMP 差错报告报文数据字段的内容





5.3.1 ICMP报文

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差错报告报文。

对第1个分片的数据报片的所有后续数据报片都不发送 差错报告报文。

对具有组播地址的数据报都不发送差错报告报文。

对具有特殊地址（如127.0.0.0或0.0.0.0）的数据报不发送差错报告报文。

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**。
- 当网络中存在网关或防火墙时，由于其防护和数据包过滤功能，连通性测试结果可能不正确。

差错与控制报文

Ping命令的使用（网络连通的数据包捕获实例）

Filter: Expression... Clear Apply

No.	Time	Source	Destination	Protocol	Info
21	5.520353	10.10.138.174	10.10.138.20	ICMP	Echo (ping) request
24	5.521611	10.10.138.20	10.10.138.174	ICMP	Echo (ping) reply
29	6.521016	10.10.138.174	10.10.138.20	ICMP	Echo (ping) request
30	6.521424	10.10.138.20	10.10.138.174	ICMP	Echo (ping) reply
37	7.521991	10.10.138.174	10.10.138.20	ICMP	Echo (ping) request
38	7.522391	10.10.138.20	10.10.138.174	ICMP	Echo (ping) reply
42	8.523014	10.10.138.174	10.10.138.20	ICMP	Echo (ping) request
43	8.523411	10.10.138.20	10.10.138.174	ICMP	Echo (ping) reply
3	0.429674	MS-NLB-PhysServer-	Broadcast	MS NLB	MS NLB heartbeat
4	0.930450	MS-NLB-PhysServer-	Broadcast	MS NLB	MS NLB heartbeat
6	1.431172	MS-NLB-PhysServer-	Broadcast	MS NLB	MS NLB heartbeat

⊕ Frame 21 (74 bytes on wire, 74 bytes captured)
⊕ Ethernet II, Src: AsustekC_27:05:9b (00:15:f2:27:05:9b), Dst: Enterasy_1c:1a:db (00:01:f4:1c:1a:db)
⊕ Internet Protocol, src: 10.10.138.174 (10.10.138.174), dst: 10.10.138.20 (10.10.138.20)
⊕ Internet Control Message Protocol
 Type: 8 (Echo (ping) request)
 Code: 0
 Checksum: 0x4a5c [correct]
 Identifier: 0x0200
 Sequence number: 0x0100
 Data (32 bytes)

差错与控制报文

Ping命令的使用（不通的情况下）

```
Pinging www.xaonline.com [202.100.4.11] with 32 bytes of data:
```

```
Request timed out.  
Request timed out.  
Request timed out.  
_
```

差错与控制报文

ICMP报文的应用举例（路由追踪命令）

```
C:\Documents and Settings\管理员>tracert www.baidu.com
```

```
Tracing route to www.baidu.com [220.181.6.18]  
over a maximum of 30 hops:
```

1	<1 ms	<1 ms	<1 ms	192.168.0.5
2	33 ms	29 ms	30 ms	116.25.176.1
3	29 ms	28 ms	29 ms	113.106.42.101
4	29 ms	29 ms	29 ms	119.145.44.210
5	30 ms	33 ms	30 ms	121.15.141.185
6	29 ms	29 ms	29 ms	202.97.25.6
7	66 ms	65 ms	65 ms	202.97.49.149
8	66 ms	65 ms	69 ms	220.189.16.54
9	70 ms	86 ms	68 ms	220.181.16.138
10	65 ms	65 ms	66 ms	220.181.17.146
11	66 ms	65 ms	65 ms	220.181.6.18

```
Trace complete.
```

过渡页

Transition Page



01 网络层概念

02 网络互连

03 差错与控制报文协议(ICMP)

04 子网编址及无分类编址与CIDR

05 因特网的路由选择协议

06 IP组播

07 移动IP

08 专用网络互连(VPN和NAT)



子网编址

子网转发

无分类编址与CIDR

使用CIDR时的路由查找算法

专用IP地址



存在的问题

5.4.1 子网编址

在 ARPANET 的早期，IP 地址的设计——给每一个物理网络分配一个网络号，确实不够合理：

IP 地址空间的利用率有时很低

会使路由表变得太大，因而使网络性能变坏



划分子网

5.4.1 子网编址

从 1985 年起在 IP 地址中增加了“子网号字段”，使两级的 IP 地址变成为三级的 IP 地址。

使得多个物理网络可以共用一个网络前缀，这种做法叫作划分子网(subnetting)。

划分子网已成为因特网的正式标准协议。



子网编址的实现

5.4.1 子网编址

IP 地址 = [网络号, 子网号, 主机号]

一般从主机域的最左边开始连续借用若干位用作子网号，位数是可变的。

为了反映哪些位用于子网号，采用子网掩码

网络号、子网号对应位为1，主机号对应位为0

子网地址 = 子网掩码 & IP 地址

子网号在网外是不可见的，仅在子网内使用。



子网的规划设计

5.4.1 子网编址

在设计选择子网划分方案时，必须考虑5个问题：

- (1) 该网络内将划分几个子网？
- (2) 每个子网有多少有效主机？
- (3) 在该子网划分中，子网掩码是什么？
- (4) 有效的子网地址是什么？
- (5) 每个子网的广播地址是什么？



子网掩码计算举例

5.4.1 子网编址

设有一个网络地址为 172.168.0.0，要在此网络中划分14个子网，问：需要多少位表示子网？子网掩码的点分十进制数值是多少？每个子网地址是什么？

解：子网数 $\leq 2^x - 2$ ，则 $x = 4$ ，需借用 4 位表示子网，其中减2是指除去子网位全0和全1情形，因为它们默认是无效的；

又172.168是B类网络，网络掩码为255.255.0.0

所以子网掩码为：255.255.240.0 $128 + 64 + 32 + 16 = 240$

子网地址：172.168.16.0, 172.168.32.0, 172.168.48.0,, 172.168.224.0

子网广播地址：172.168.31.255, 172.168.47.255,, 172.168.239.255



子网划分

5.4.1 子网编址

使用子网划分会造成部分主机地址损失，但仍可提高IP地址利用率。

根据需要设计子网掩码划分子网。

主机号全0或全1的IP地址不能分配给主机，因此子网划分需要确保主机号位数至少为2；

如果不允许子网号全0或全1，则子网划分要确保子网号位数至少为2。



子网转发

5.4.2 子网转发

路由表表项：(子网掩码,目的网络地址,下一跳地址)

采用子网编址方案时的IP数据报转发（数据报DG，路由表T）

从数据报DG中取出目的IP地址 I_D ;

for 表T中的每一表项

将 I_D 与表项中的子网掩码按位相“与”，结果为N;

if N等于该表项中的目的网络地址,则

if 下一跳指明应直接交付,则

把DG直接交付给目的站

else

把DG发往本表项指明的下一跳地址

return

end_for

没有找到匹配的表项,向DG的源站发送一个目的不可达差错报告



3个问题

5.4.3 无分类编址与CIDR

划分子网在一定程度上缓解了因特网在发展中遇到的困难。然而在 1992 年因特网仍然面临三个必须尽早解决的问题：

B 类地址在 1992 年已分配了近一半，很快就要全部分配完毕。而一个 C 类地址空间较小。

因特网主干网上的路由表中的项目数急剧增长（从几千个增长到几万个）。

B类地址空间未得到充分利用。整个 IPv4 的地址空间将耗尽。



IP 编址问题的演进

5.4.3 无分类编址与CIDR

变长子网掩码

1987年，RFC 1009 就指明了在一个划分子网的网络中可同时使用几个不同的子网掩码。使用**变长子网掩码 VLSM**可进一步提高IP地址资源的利用率。

无分类编址

在VLSM的基础上又进一步研究出**无分类编址**方法，它的正式名字是无分类域间路由选择**CIDR** (Classless Inter-Domain Routing)。



CIDR主要特点

5.4.3 无分类编址与CIDR

CIDR **消除了**传统的 A 类、B 类和 C 类地址以及划分子网的概念，因而可以更加有效地分配 IPv4 的地址空间。

CIDR使用**各种长度的“网络前缀”** (network-prefix) 来代替分类地址中的网络号和子网号。

IP 地址从三级编址（使用子网掩码）**又回到了两级编址。**



CIDR两级编址

5.4.3 无分类编址与CIDR

无分类的两级编址的记法是：

IP地址 ::= { <网络前缀>, <主机号> }

CIDR 还使用 “斜线记法”：

即在IP地址后面加上一个斜线 “/”，后面写上网络前缀所占的比特数，如20.5.0.0/10，还可简写为 20.5/10

CIDR 将网络前缀相同的连续的 IP 地址组成 “CIDR 地址块”。



CIDR地址块

5.4.3 无分类编址与CIDR

128.14.32.0/20表示的地址块共有 2^{12} 个地址（主机号的比特数 = $32-20$ ）。

该地址块的起始地址是 128.14.32.0。

该地址块的最小地址：128.14.32.0

10000000 00001110 0010 0000 00000000

该地址块的最大地址：128.14.47.255

10000000 00001110 0010 1111 11111111



路由聚合

5.4.3 无分类编址与CIDR

一个 CIDR 地址块可以表示很多地址，这种地址的聚合常称为**路由聚合(route aggregation)**，它使得路由表中的一个项目可以表示很多个（例如上千个）原来传统分类地址的路由。



构成超网

5.4.3 无分类编址与CIDR

前缀长度不超过 23 bit 的 CIDR 地址块都包含了多个 C 类地址。

这些 C 类地址合起来就构成了超网(supernetting)。

CIDR 地址块中的地址数一定是 2 的整数次幂。

网络前缀越短，地址块所包含的地址数就越多。



5.4.4 使用CIDR时的路由查找算法

路由表中的每个项目由 **“网络前缀/掩码”** 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。

应当从匹配结果中选择具有最长网络前缀的路由：
最长前缀匹配(longest-prefix matching)。

网络前缀越长，其地址块就越小，因而路由就越具体。



5.4.4 使用CIDR时的路由查找算法

最长前缀匹配举例

收到的分组的目的地地址 $D = 206.0.71.130$

路由表中的项目：① $206.0.68.0/22$

② $206.0.71.128/25$

解： $D \& 255.255.252.0 = 206.0.68.0$ 与①匹配

$D \& 255.255.255.128 = 206.0.71.128$ 与②匹配

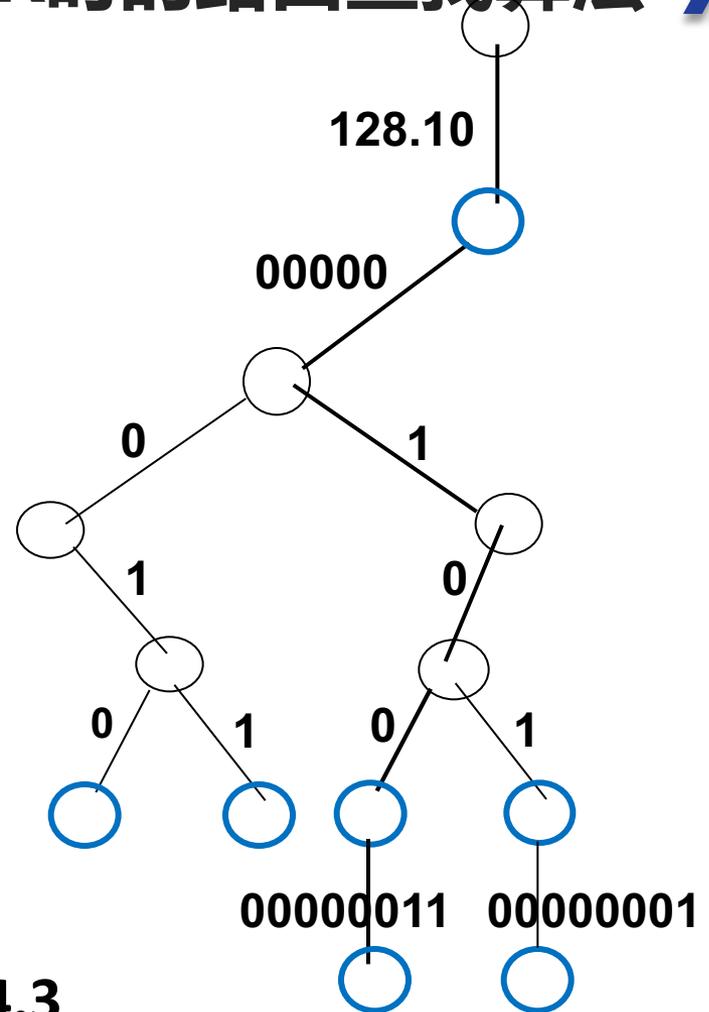
选择两个匹配的地址中更具体的一个，即**选择具有最长网络前缀的路由。**

二叉线索树

5.4.4 使用CIDR时的路由查找算法

- 无分类查找使用二叉线索树

网络前缀/前缀长	下一跳
128.10.0.0/16	10.0.0.2
128.10.2.0/24	10.0.0.4
128.10.3.0/24	10.1.0.5
128.10.4.0/24	10.0.0.6
128.10.4.3/32	10.0.0.3
128.10.5.0/24	10.0.0.6
128.10.5.1/32	10.0.0.3



? 128.10.4.3

? 128.10.4.5



5.4.5 专用IP地址

IP地址资源是有限的，为了节约地址的使用，IANA保留了三块只能用于专用互联网（private internet）内部通信的IP地址空间[RFC1918]

<u>前缀</u>	<u>最低地址</u>	<u>最高地址</u>
10/8	10.0.0.0	10.255.255.255
172.16/12	172.16.0.0	172.31.255.255
192.168/16	192.168.0.0	192.168.255.255