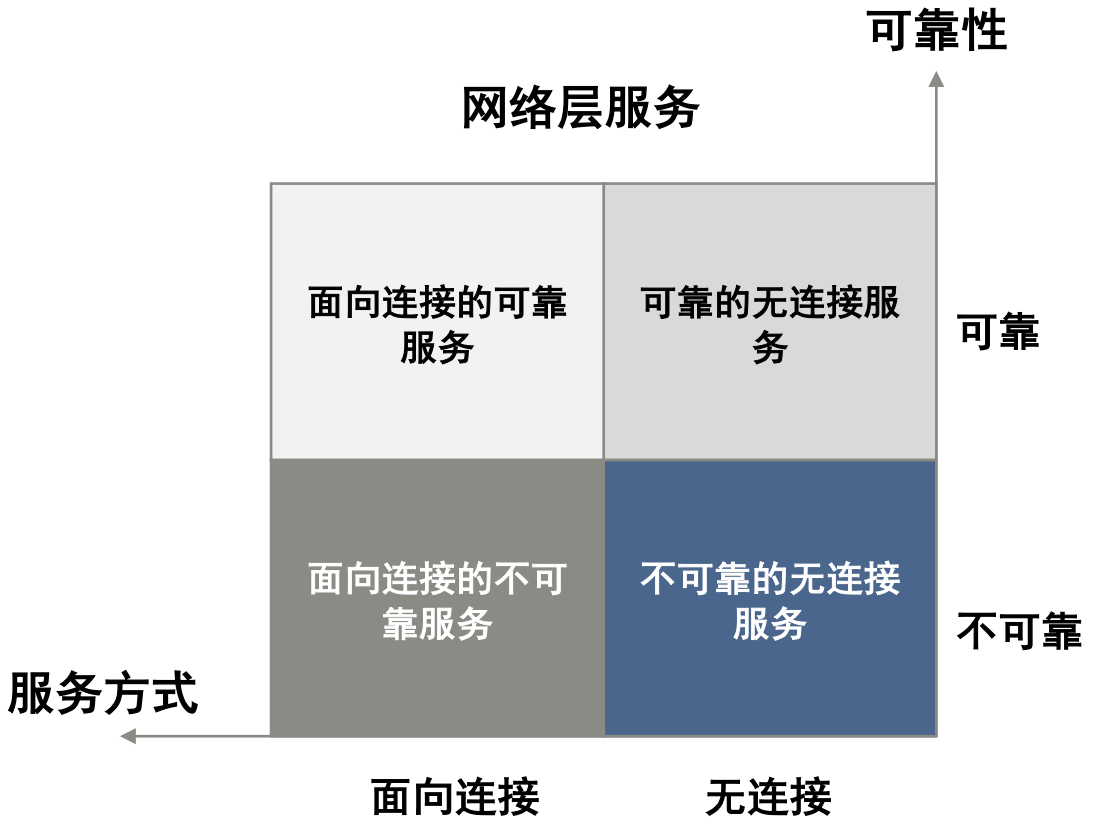
案例学习

TCP/IP之网络层协议

因特网网络体系与协议集

TCP/IP

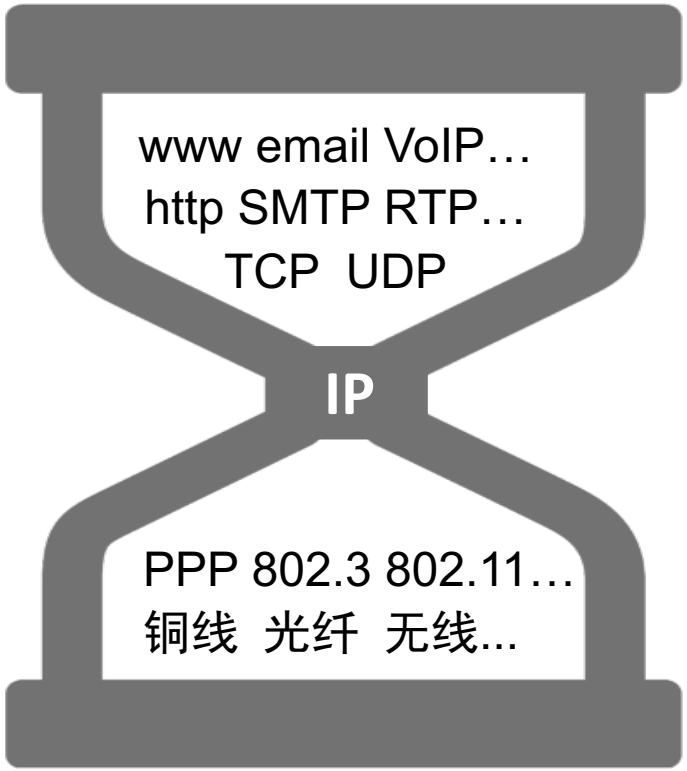
- 网络层作为服务提供者向上面传输层提供一定的服务
- 网络层如何实现服务则是本层协议的目标



因特网网络层协议子集

沙漏模型（细腰结构）

上下两部分很大（扩展性好），所有网络通信都要穿过“细腰”。



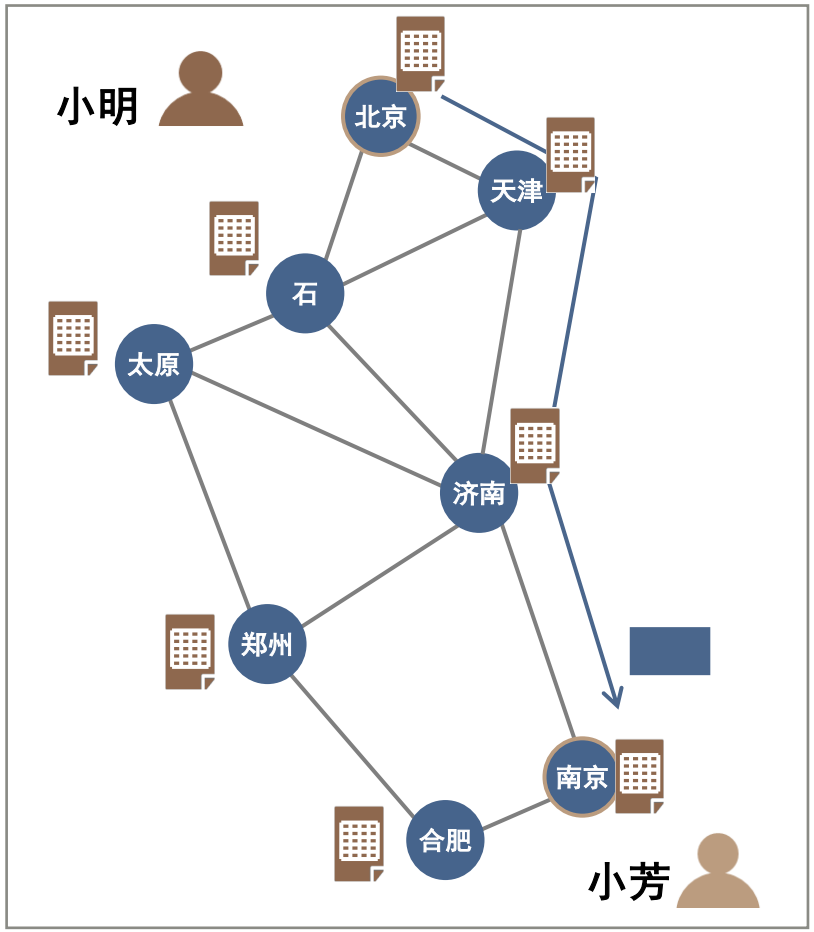
主协议
- IP

包传递有关
- ICMP
- ARP

地址管理有关
- DHCP
- NAT

路由计算有关
- OSPF
- BGP

包传递与数据报子网

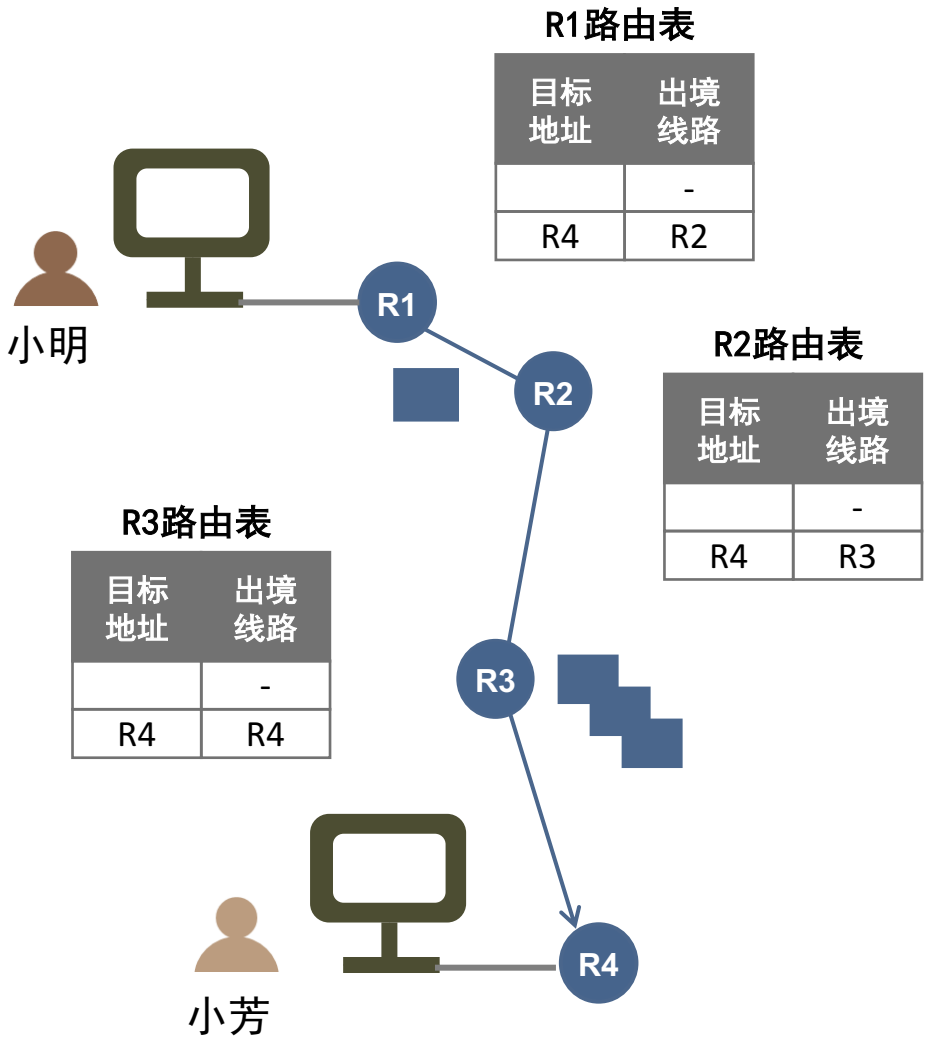


数据包传递

- 包带有完整地址信息
- 途径每个路由器对包进行存储-转发处理
- 转发线路由路由表决定

数据报子网：每个数据报具有完整的地址信息，同一对端系统之间的数据报可走不同的路径。

存储-转发技术与点-点传输



小明发给小芳的包传递过程：

① 小明主机将消息封装在网络层的包中通过网卡发给局域网的路由器R1

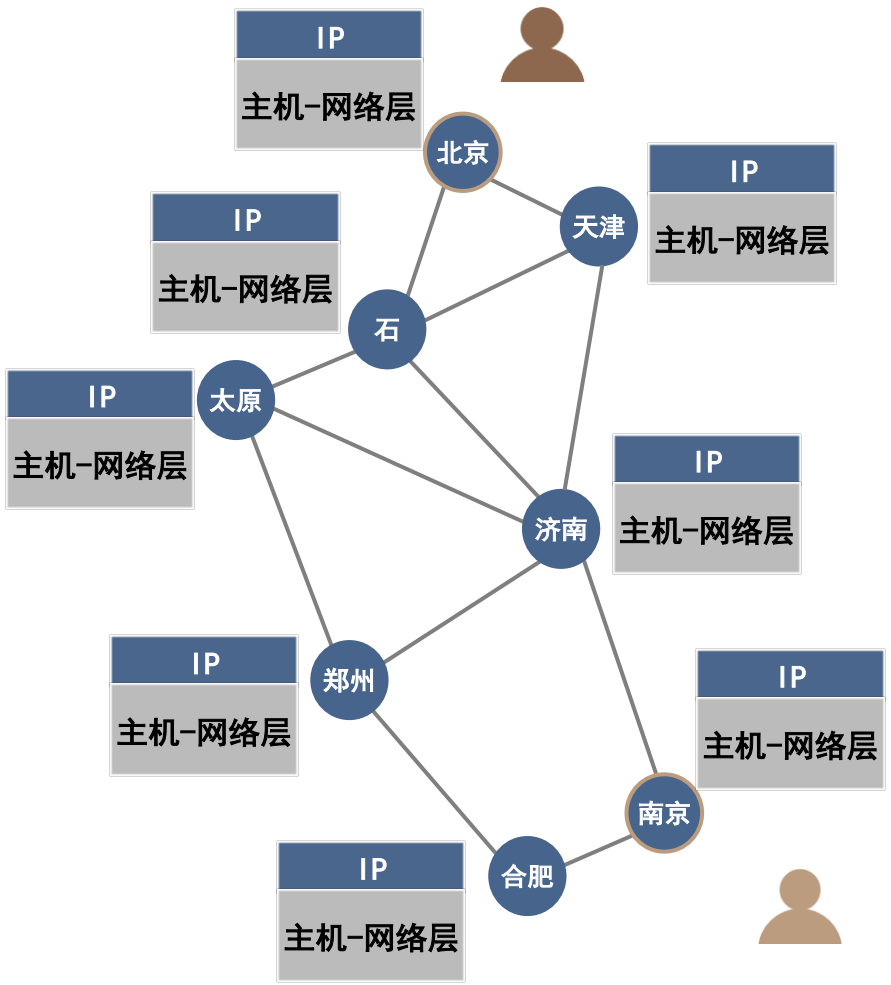
② R1网卡接收该包先存入内存，稍后根据包的目标地址和路由表做出把包发给邻居R2的路由决策，然后相应的网卡将包发到连接R2的线路

③ R2/R3做同样的处理

④ R4网卡接收包，根据包的目的地址得知包已经到达目的地，把包发往局域网，小芳主机就能接收到该包。

- 网络层负责包的前进方向，链路层负责包的实际传输。
- 一次端-端的网络层通信由一系列的点-点传输组成。

互联网络（IP）协议



IP协议：为上层用户提供了尽力而为的无连接不可靠包传递服务。

IP标准

- 全局编址
- 封装和拆封
- 分段和重组

- 每个节点拥有唯一的IP地址
- 规定了如何传递上层数据
- 规定了包在小网络如何传输

IP数据包的作用

硬件帧格式

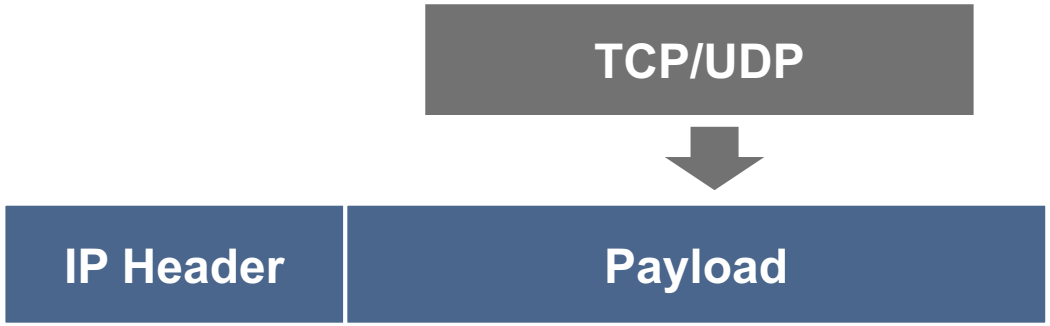
- 路由器要连接异构网络
- 不同类型网络的帧格式不同

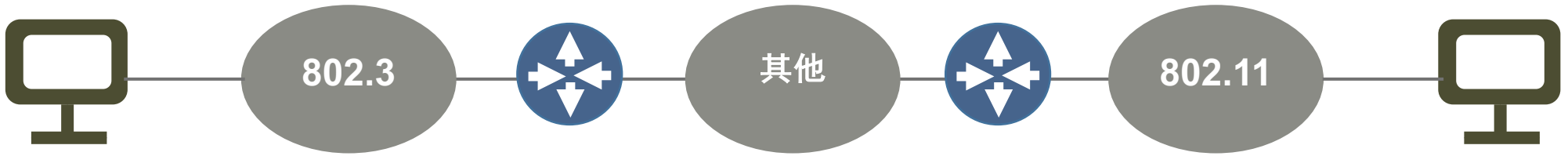
?

用链路层的数据帧传输网络层的用户传输层的数据行不行？

虚拟包

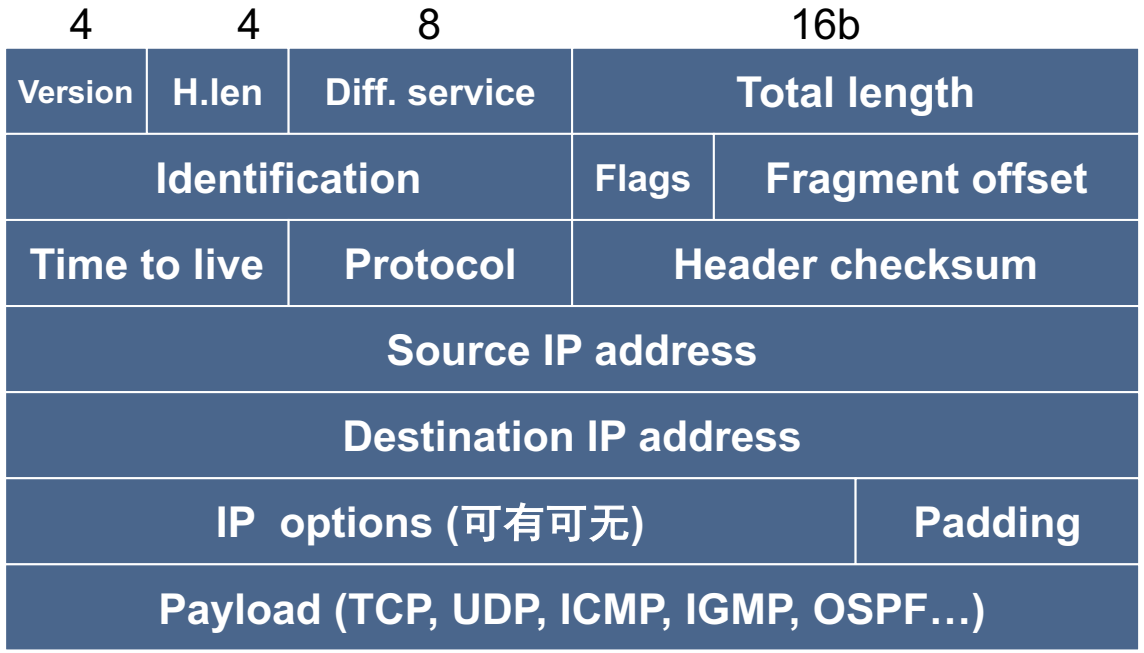
- 一个独立于底层硬件的包格式
- IP数据报/包/分组





IP协议之报文格式

IP数据报格式



- Version：版本号，目前是v4和v6并存
- H.len：IP包头的长度，以字(32位)计数
- Diff service：用于服务质量和拥塞控制
- Total length：整个IP包的总长度
- Identification, Flags, Fragment offset：与分段重组有关

- Time to live：IP包的生存期
- Protocol：指明了IP服务使用者，即上层用户
- Header checksum：校验IP包的头

- Source IP addr./ Destination IP addr：IP包的源端和目标端地址
- IP option：选项，IP的增值服务
- Padding：包头以字计数，必须32位整数
- Payload： 携带上层数据或本层控制消息

IP协议的差错校验

- IP校验和

Header checksum

RFC791
RFC1071
RFC1700

为确保IP包传递到正确地址，必须对地址字段进行再次校验。

计算方法：

① 按16位相加

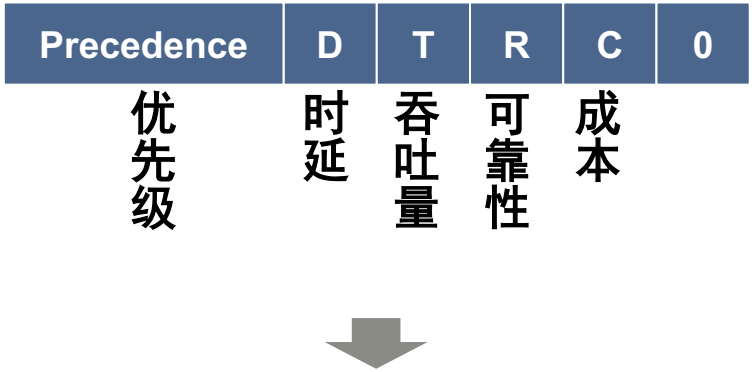
② 结果取反

IP进行软件校验和计算

链路层进行硬件CRC校验和计算

IP协议的服务质量

- 服务类别



RFC1340
RFC1349

应用程序	D	T	R	C
telnet/rlogin	1	0	0	0
ftp/control	1	0	0	0
ftp/data	0	1	0	0
snmp	0	0	1	0
nntp	0	0	0	1
smtp/command	1	0	0	0
Smtp/date	0	1	0	0

- 区分服务

IP协议的用户和包生存期

- 生存期（TTL）

Time to live

- IP包在网络中的生存期限
- 路由器在存储-转发包时将该字段减1
- 一个生存期为0的包将被路由器丢弃

- 协议

Protocol

接收端的网络层把包的有效载荷上交给协议字段标识的上层协议实体。

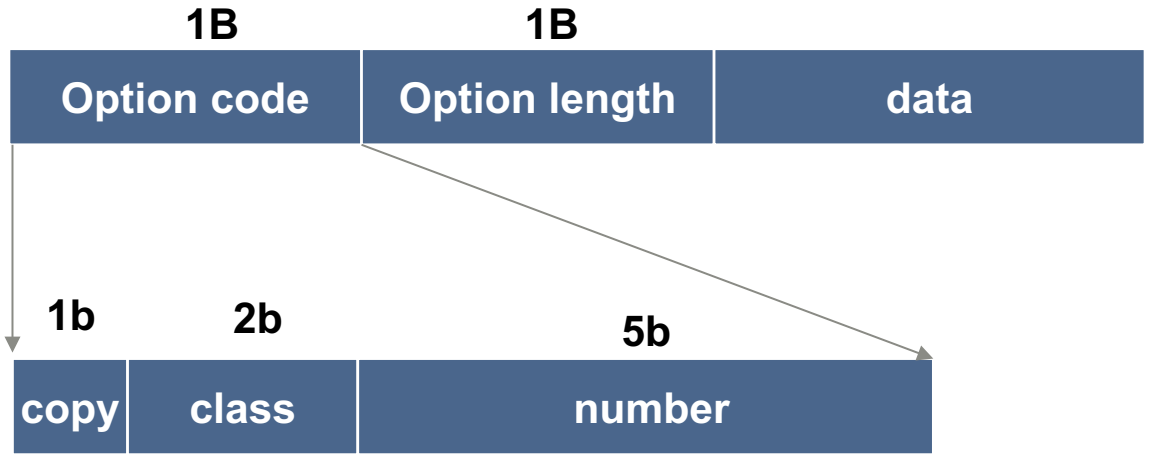
RFC 1700

协议字段定义

0		Reserved
1	ICMP	Internet Control Message
2	IGMP	Internet Group Management
3	GGP	Gateway-to-Gateway
4	IP	IP in IP (encapsulation)
6	TCP	Transmission Control
8	EGP	Exterior Gateway Protocol
17	UDP	User Datagram
29	ISO-TP4	ISO Transport Protocol Class 4
55-60	ISO-IP	ISO Internet Protocol
80	MTP	Multicast Transport Protocol

IP协议的附加功能

- IP选项（可有可无）



copy指示位

- 0：该选项应被拷贝到所有段中
- 1：该选项仅被拷贝到第一段中

Class类别

- 0：数据报/网络控制
- 1：保留
- 2：纠错和度量
- 3：保留

IP包在传递过程中，可以执行某种特殊功能。具体操作由选项码定义。

IP包选项类别

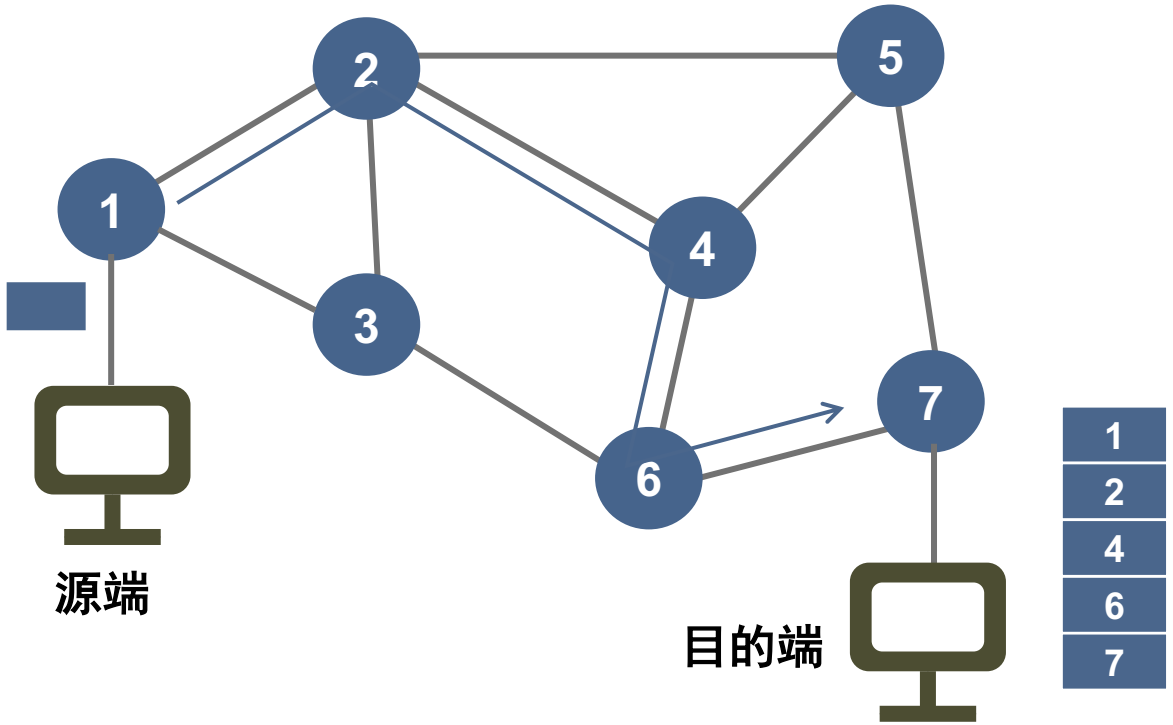
可能的IP选项

class	number	length	描述
0	0	-	选项表结束
0	1	-	无操作
0	2	11	安全和处理限制
0	3	var	松散源路由(指定数据报的路由）
0	7	var	记录路由(用来跟踪路由)
0	8	4	流标识符
0	9	var	严格源路由(指定数据报的路由）
2	4	var	Internet时间戳(记录路由时间)

IP协议的记录路由功能

- 选项中包含一个地址空表
  - 由所有处理过该数据报的路由器把自己IP地址填入表中
  - 路由器在指针所指的位置插入自己的IP地址

(0, 7)	length	Pointer	
First hop IP address			
Second hop IP address			
.....			



包的目标主机从收到的包中获得该包经过的一条路径：1→2 → 4 → 6 → 7

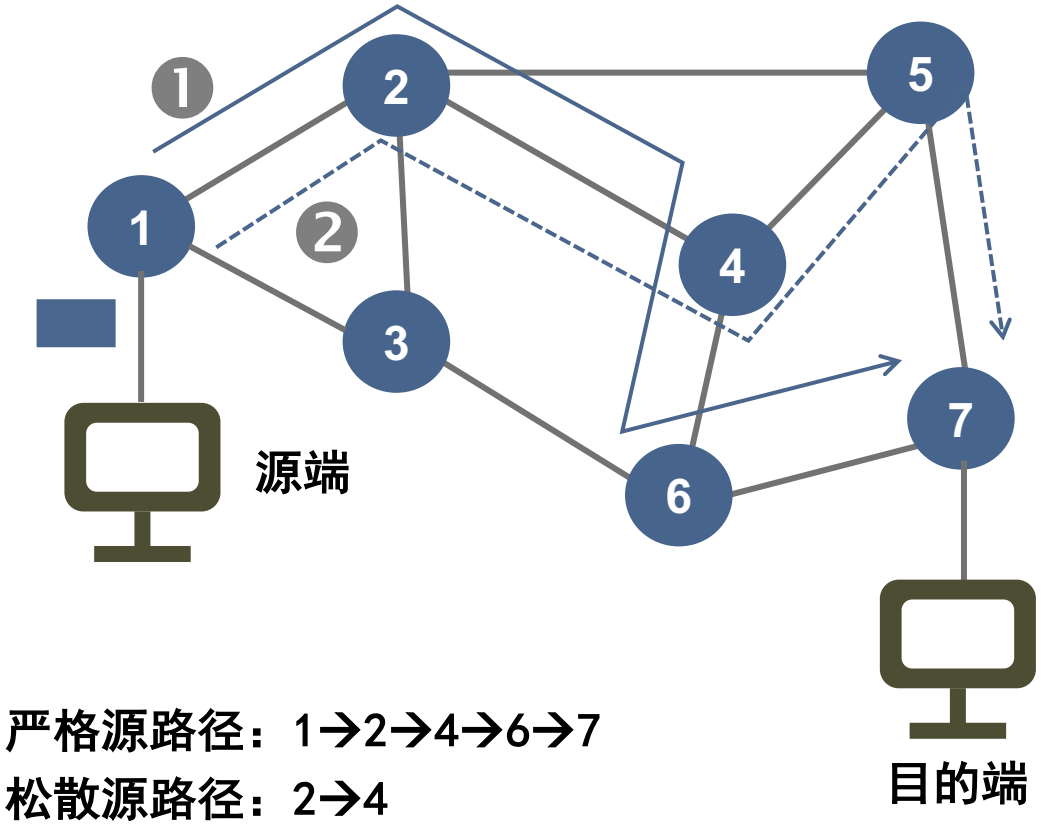
IP协议的源路由功能

- 源路由选项

  包含一个IP地址序列来指定一条路由

  - 严格源路由—两个相邻地址必须处在同一物理网络上
  - 松散源路由—允许相邻两个地址之间跳过多个网络

(0,3)/(0,9)	length	Pointer	
First hop IP address			
Second hop IP address			
.....			



① 严格源路径：1→2→4→6→7

② 松散源路径：2→4

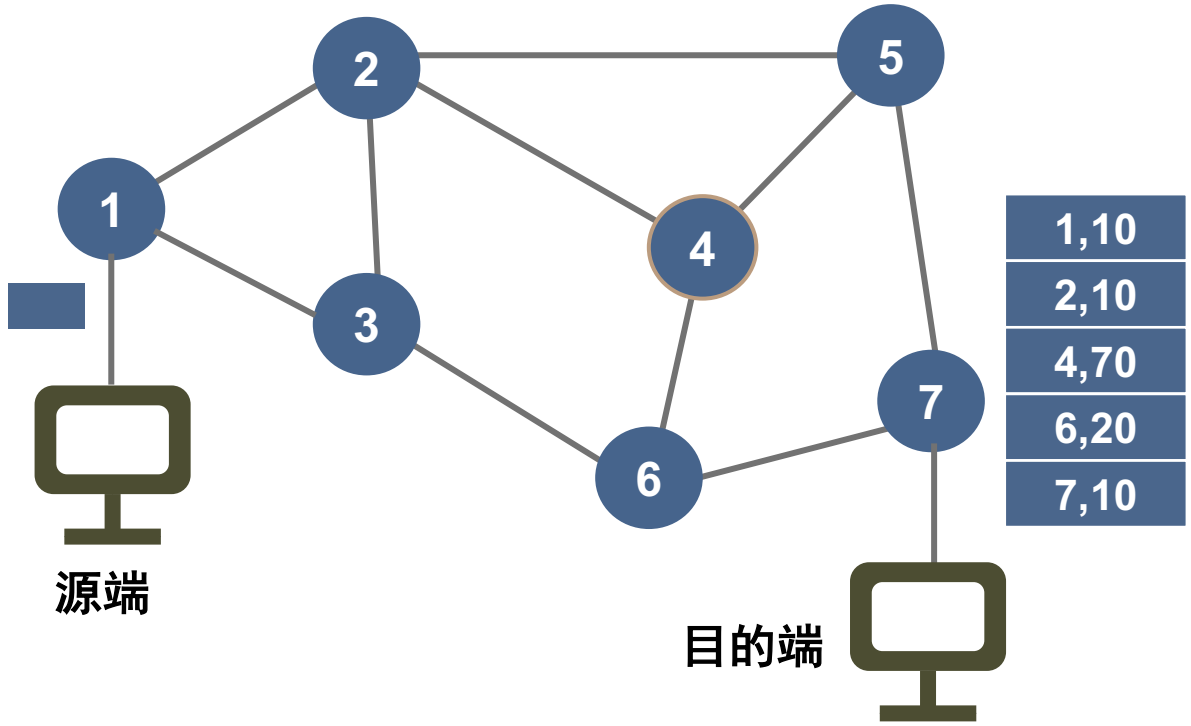
源端可规定包的传输路径

IP协议的时间戳功能

- 选项包含一个记录时间空表

  途径的每个路由器均在表中填入时间

(2,4)	length	Pointer	oflow	flag
First IP address				
First timestamp				
Second IP address				
Second timestamp				



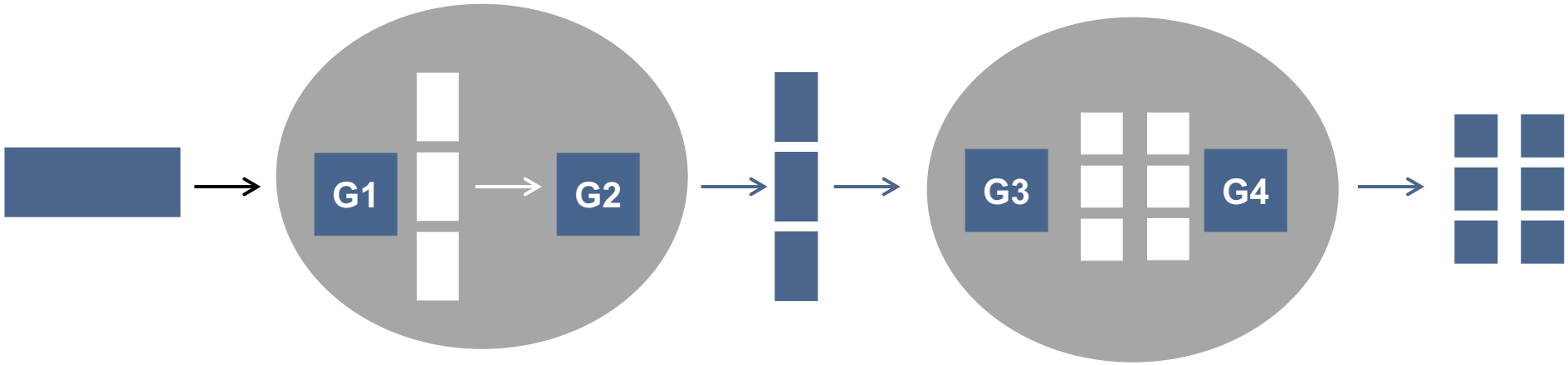
可用来发现路径上发生拥塞的区域.

注意：每一跳的时间是根据时间戳计算得出的。

IP数据报——分段时选项的处理

分段时对选项的处理

- 针对选项功能作不同的处理
- 记录路由选项只拷贝到其中一个段中
- 源路由选项必须拷贝到所有段中

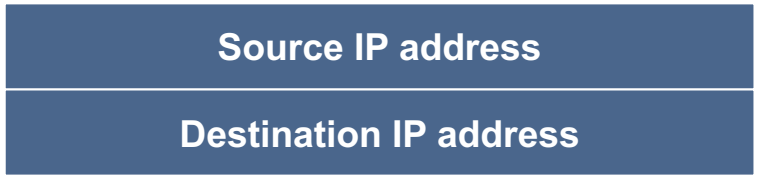


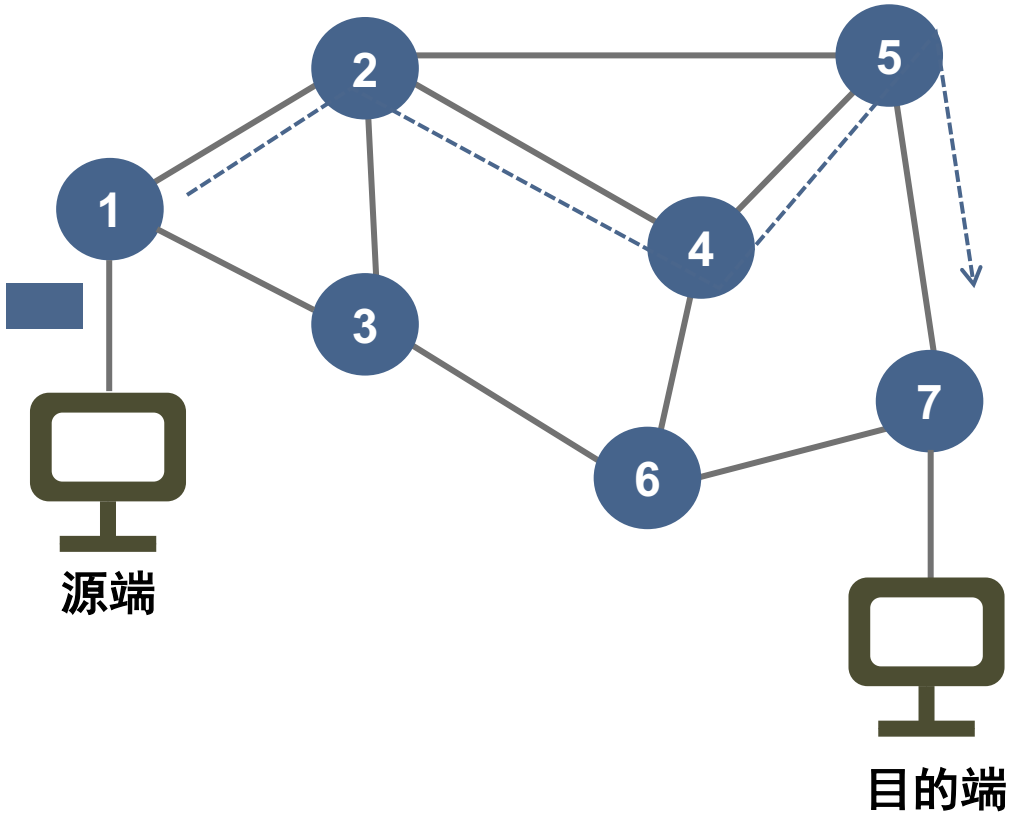
IP协议之网络管理

IP包投递面临的问题

① 如何标识网络中的某个节点

- 因特网中的任何一个节点必须有一个唯一的ID，才能标识IP包的发送方和接收方。
- ID的命名方式要便于网络寻址





② 如何找寻通往那个节点的一条路径

- 每个路由器在存储-转发包时必须有能力将包转发到通向目的地的下一跳
- 路由器的转发线路必须在最好路径上

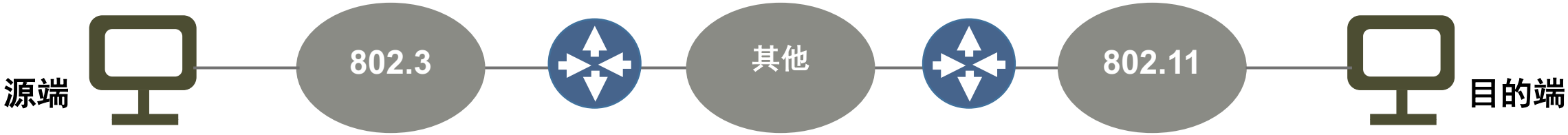
协议地址和IP地址

协议地址：协议软件定义一个与底层物理地址无关的编址方案，给每台主机分配一个唯一的地址。

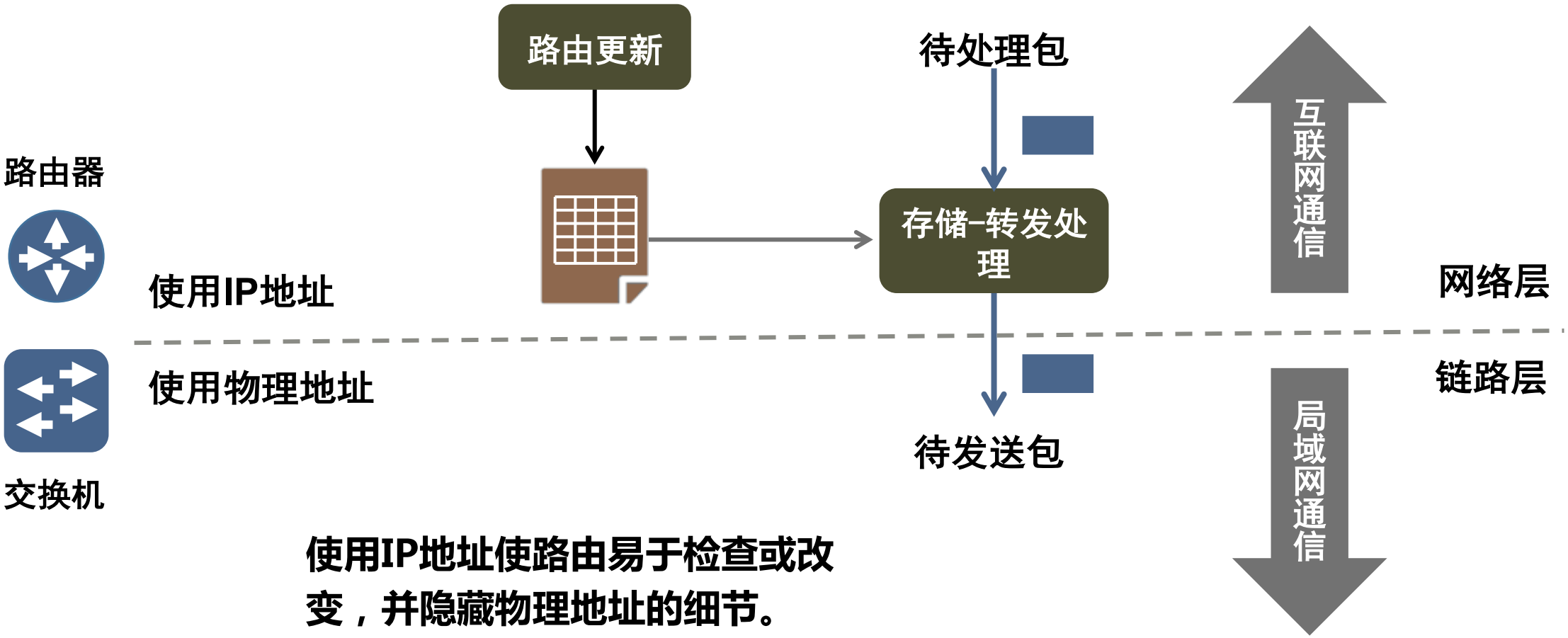
IP规定

- 每台主机有一个32位二进制数作为其Internet地址
- 发送的包必须包括32位的发送方和接收方地址

？ 物理地址能否用于包ID？是否满足寻址要求？



IP地址有利于寻址



使用IP地址使路由易于检查或改变，并隐藏物理地址的细节。

IP地址的层次性

? 假设给一台新入网的计算机任意分配一个未用的地址，会怎样？

IP地址层次结构

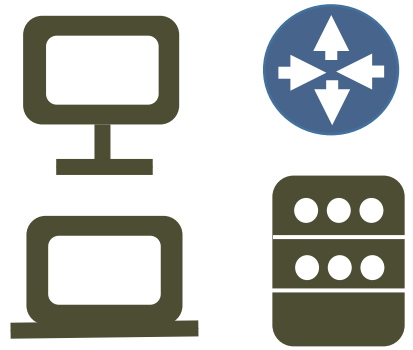
- 前缀：确定计算机所属的物理网络
- 后缀：确定物理网络上一台计算机

IP地址的层次结构使得路由算法针对网络进行路由成为可能。

? 前后缀的比例怎么分

IP地址分类及表示

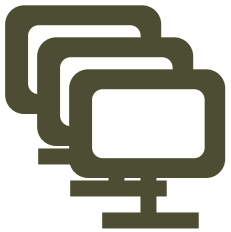
1.0.0.0 ~ 127.255.255.255	0	Network	Host	A
128.0.0.0 ~ 191.255.255.255	10	Network	Host	B
192.0.0.0 ~ 223.255.255.255	110	Network	Host	C



单播地址

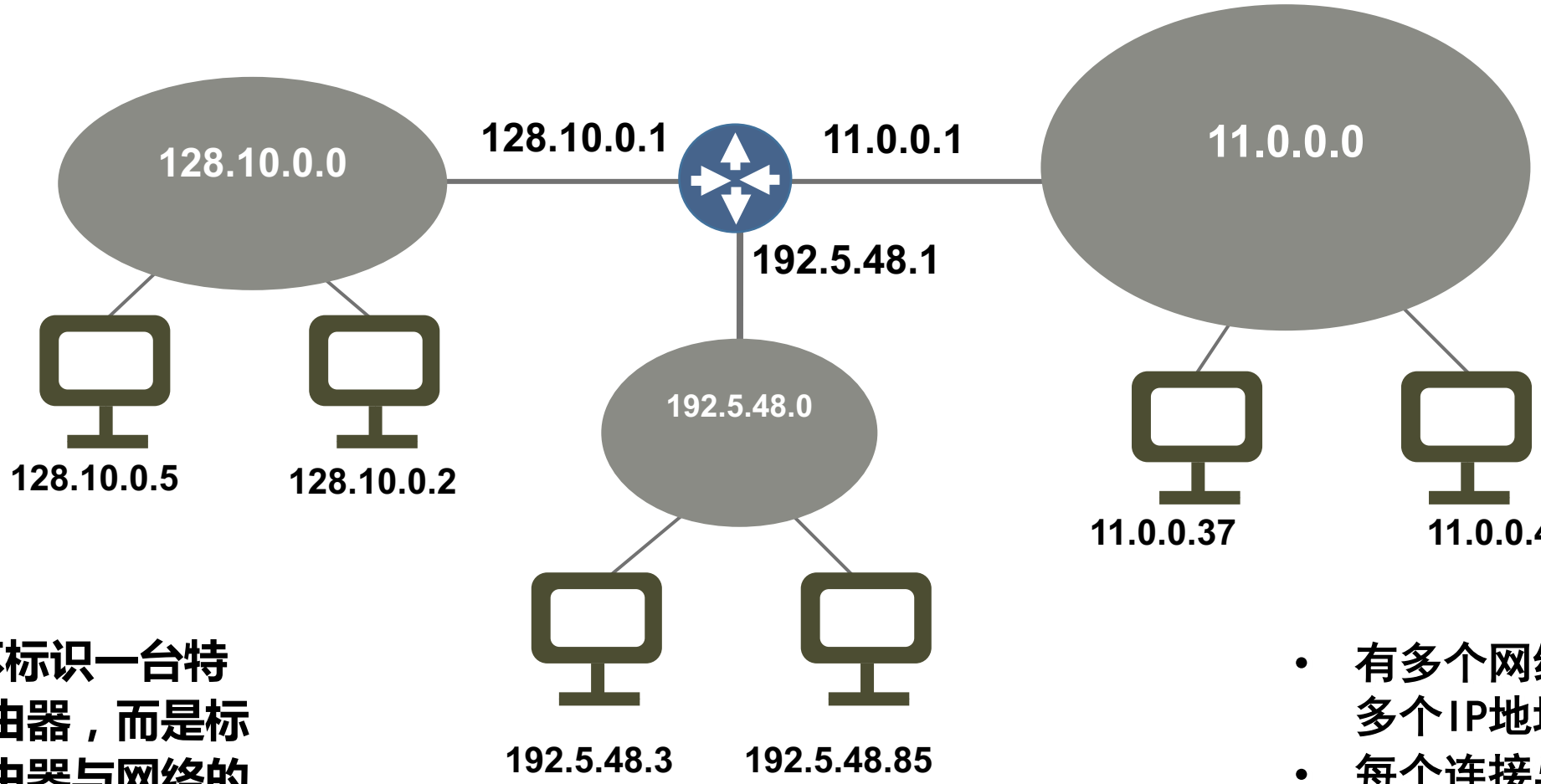
点分十进制:将32位中的每8位为1组，用十进制表示，利用句号分隔各部分。

224.0.0.0 ~ 239.255.255.255	1110	Multicast address	D
240.0.0.0 ~ 247.255.255.255	11110	Reversed for future use	E



组播地址

IP地址层次及分类



IP地址并不标识一台特定主机/路由器，而是标识主机/路由器与网络的一个连接。

- 有多个网络连接就有多个IP地址。
- 每个连接与一个地址关联

谁来管理IP地址

总机构IANA

- 因特网编号管理局(www.iana.org)
  - 北美和南美: 美洲因特网编号注册机构(ARIN)
  - 欧洲: 欧洲IP网络 (RIPE)
  - 亚洲: 亚太网络信息中心 (APNIC)

中国的网络资源分配情况
(www.cnnic.net.cn)

截至2017年6月，我国IPv4地址数量达到3.38亿个、IPv6地址数量达到21283块/32地址，二者总量均居世界第二。
——摘自第40次《中国互联网络发展状况统计报告》

特殊的IP地址

保留地址：一种特殊的地址格式，从不分配给主机。

- 全0的地址用于一个尚未分配到IP地址的主机
- 网络号为0的IP地址标识了本地的一台主机
- 主机号为0的IP地址标识了一个网络
- 主机号为1网络号非0的IP地址标识了一个网络中的全部主机
- 主机号和网络号均为1标识了本地网络中的全部主机
- 127的A类地址仅指本机

IP保留地址

前缀	后缀	类型	用途
全0	全0	本机	启动时使用
全0	主机	主机	本网络中的主机
网络	全0	网络	标识一个网络
网络	全1	直接广播	在指定网络广播
全1	全1	局部广播	在本地网广播
127	任意	回环	测试

IP前缀以及地址掩码

A类网络 （8-bit网络前缀）	B类网络 （16-bit网络前缀）	C类网络 （24-bit网络前缀）
• 前缀: 11.0.0.0/8 • 主机: 11.0.0.0 ~ 11.255.255.255	• 前缀: 128.10.0.0/16 • 主机: 128.10.0.0 ~ 128.10.255.255	• 前缀: 192.5.48.0/24 • 主机: 192.5.48.0 ~ 192.5.48.255
• 掩码：255. 0. 0. 0	• 掩码：255. 255. 0. 0	• 掩码： 255. 255. 255. 0

网络规模太大

网络规模太小

地址分配面临的挑战与应对策略

大型ISP

- 拥有A类地址块
- 很难组织IP地址

小型企业

- 拥有几个C类地址
- 很难管理这么多的前缀

解决方案

- 子网编址(固定大小的子网划分)
- 无类域路由（灵活计算的子网划分）

IP协议之地址管理示例

IP地址管理示例

假设某个办事处申请到一个C类地址202.111.222.0。该办事处分布在5座办公楼，需要设置5个局域网。局域网规模分别是：2个局域网10台主机，2个局域网30台主机，1个局域网2台公共服务器。

试问：

① 采用子网编制技术的子网划分和地址分配方案？

② 采用CIDR技术的子网划分和地址分配方案？

方案一：采用子网编址方法

需求分析

- 最大的子网必须能容纳30台主机，至少需要5位主机号
- 共有5个子网至少需要3位表示子网

分配方案

从202. 111. 222. 0表示主机号的第四个字节中分出3位用来表示子网号，剩余5位表示子网内的主机。

可用地址 **11001010 01101111 11011110 xxxxxxxx**

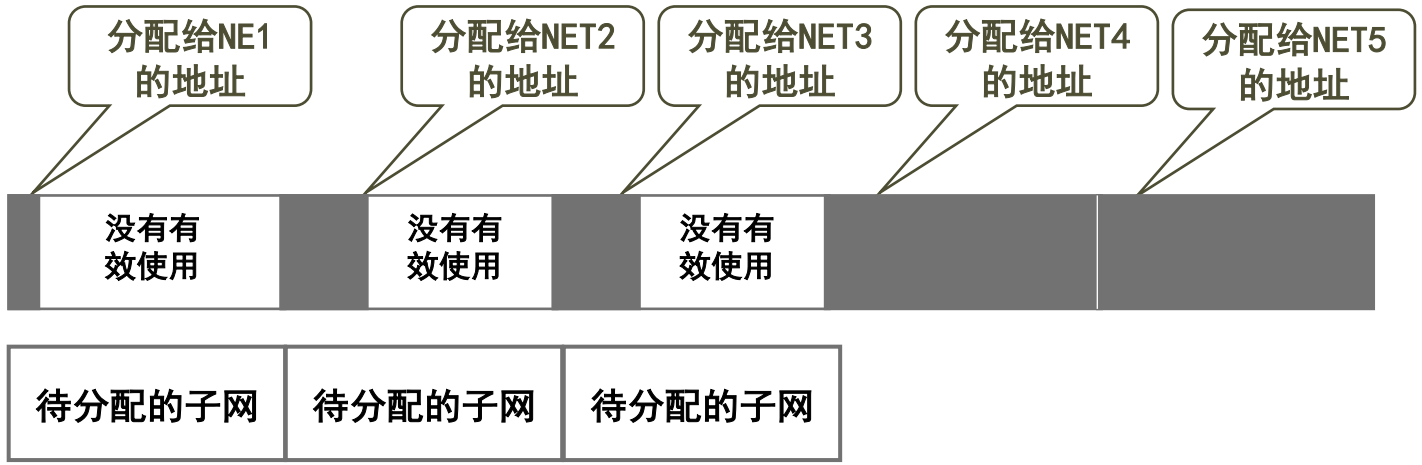
3位子网的掩码 **11111111 11111111 11111111 11100000**

子网编制的地址分配样例

	202.111.222.32	202.111.222.64	202.111.222.96	202.111.222.128	202.111.222.160
	NET1 2台	NET2 10台	NET3 10台	NET4 30台	NET5 30台
子网号/子网掩码	202.111.222.00100000 255.255.255.224	202.111.222.01000000 255.255.255.224	202.111.222.01100000 255.255.255.224	202.111.222.10000000 255.255.255.224	202.111.222.10100000 255.255.255.224

地址资源使用情况

- ••3个子网没有用，可以安排给其他子网
- ••除128和160子网外，其余3个子网浪费了不少地址资源



方案二：采用CIDR方法

需求分析

- ••主机数为2子网，可以用2位表示
- ••主机数为10的子网，可以用4位表示（分别浪费4个/6个地址）
- ••主机数为30的子网，用6位表示（全0全1去掉）

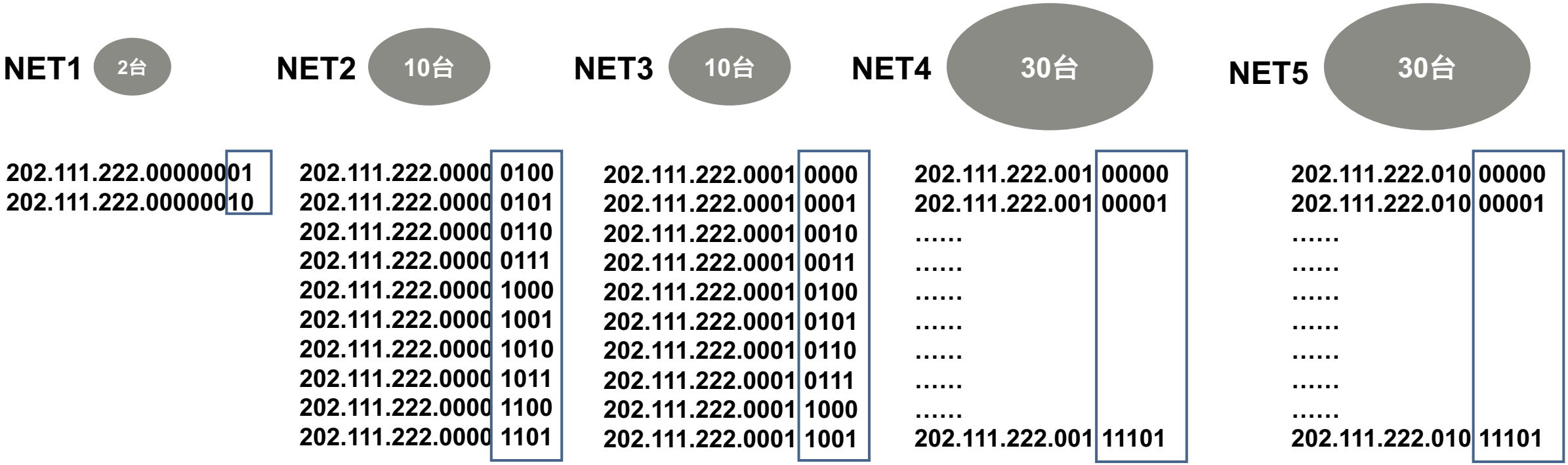
分配方案

从202.111.222.0开始，尽量按连续块分配：

- NET1和NET2分配一个16个地址的块
- NET3分配一个16地址的块
- NET4/NET5分配一个32地址的块

可用地址 **11001010 01101111 11011110 xxxxxxxx**

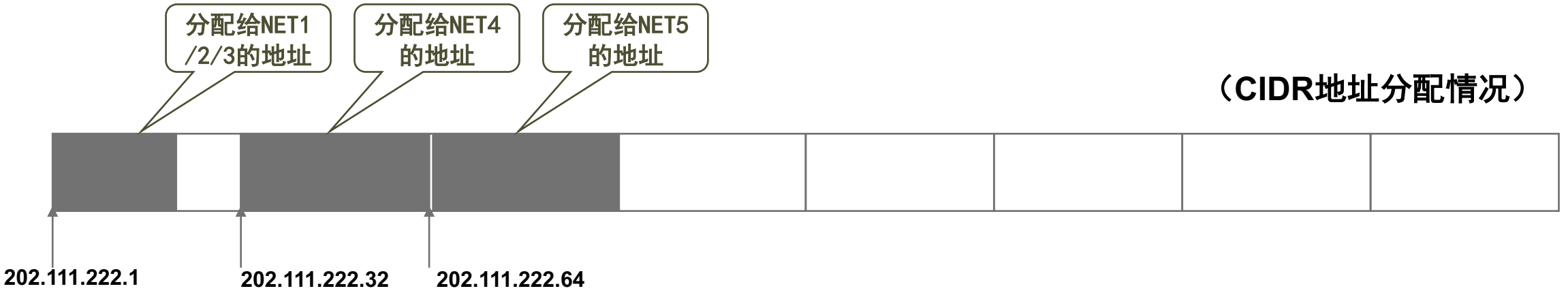
CIDR分配的地址样例



两种地址分配方法比较

地址资源使用情况

- 地址尽量按需大小分配，减少了浪费的地址数
- 便于路由聚合

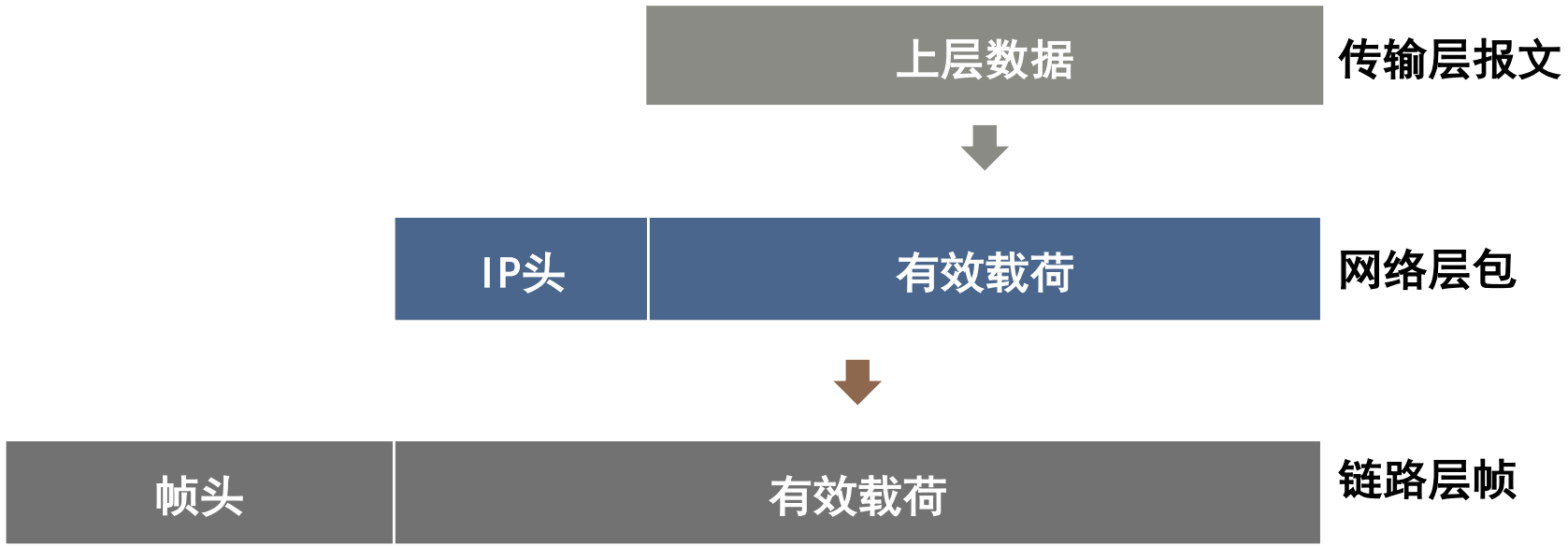


IP协议之分段与重组

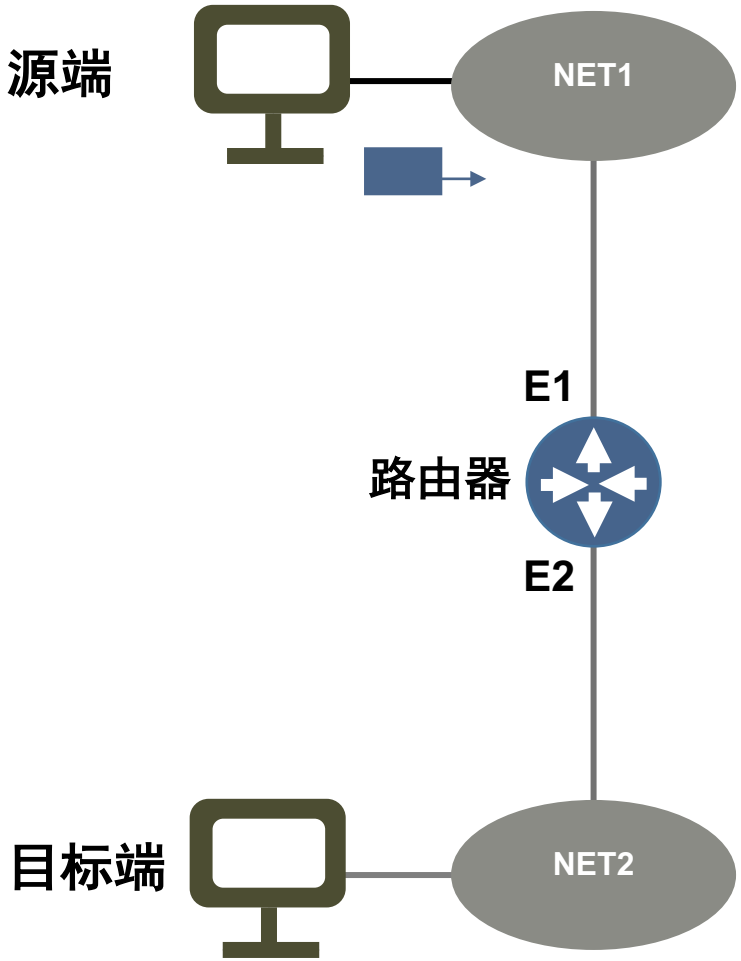
IP包——封装

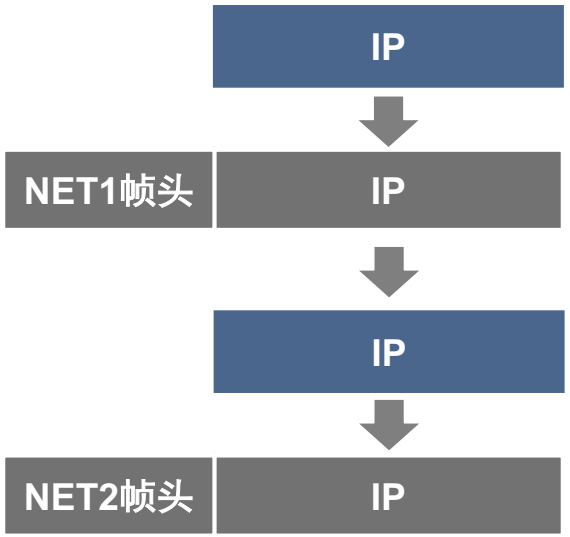
封装：将IP数据报装进一个帧的数据区，网络硬件像对待普通帧一样对待包含着数据报的帧。

IP包封装在哪种类型的数据帧中取决于路由选择的结果和转发接口的MAC协议。



IP包——多次封装





① 主机把IP包封装在网卡发送给本地路由器

② 路由器从网卡E1接收数据帧，提取出其中的IP包做存储-转发处理

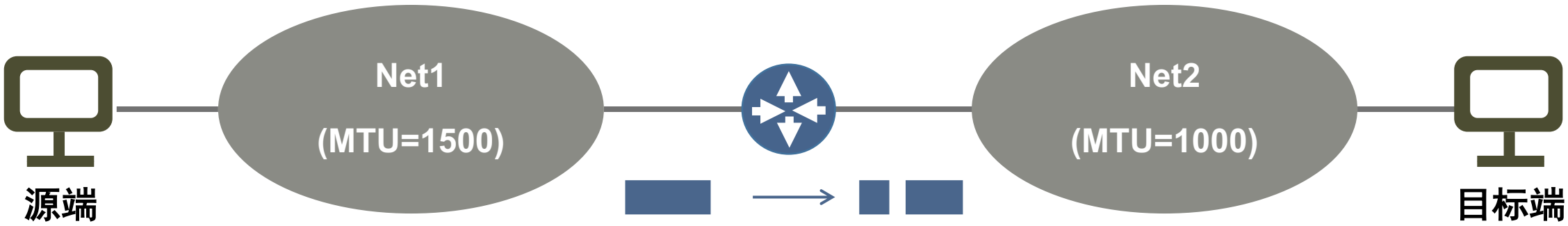
③ 路由器把IP包封装在路由指示的转发接口E2的数据帧发送

？ 为什么要多次封装？

IP包的分段

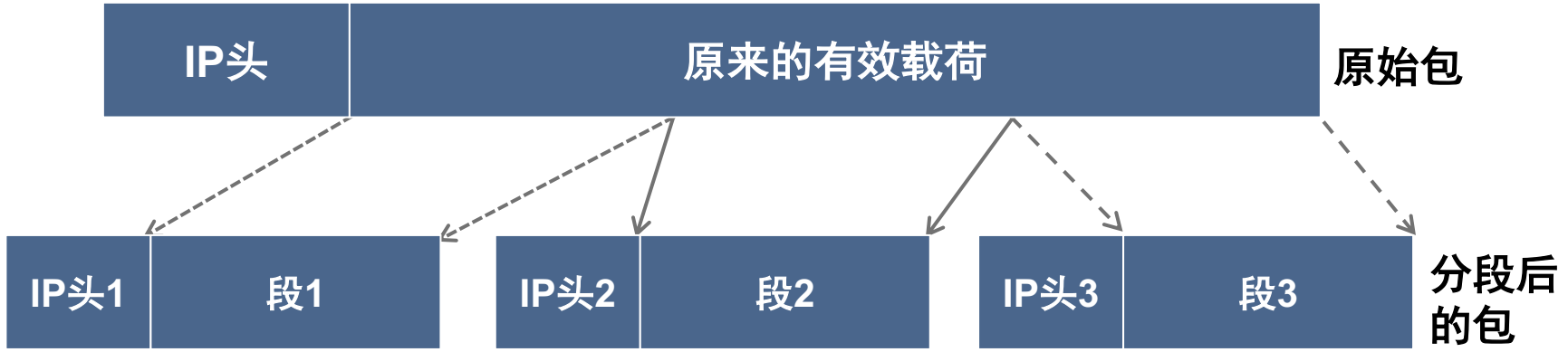
分段：当包的尺寸大于网络的最大传输单元时，路由器将包分成若干个较小部分—称为段。

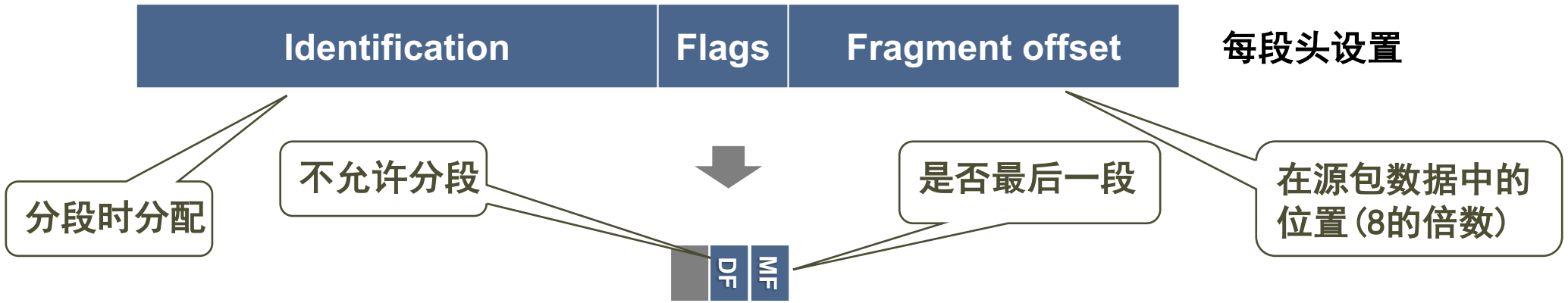
最大传输单元：指特定网络所能传输的最大数据块长度。



IP包分段操作

- 每一段携带取之原数据报的部分数据，具有一个类似于原包的报头
- 分段后包头必须给出用于重组的分段信息





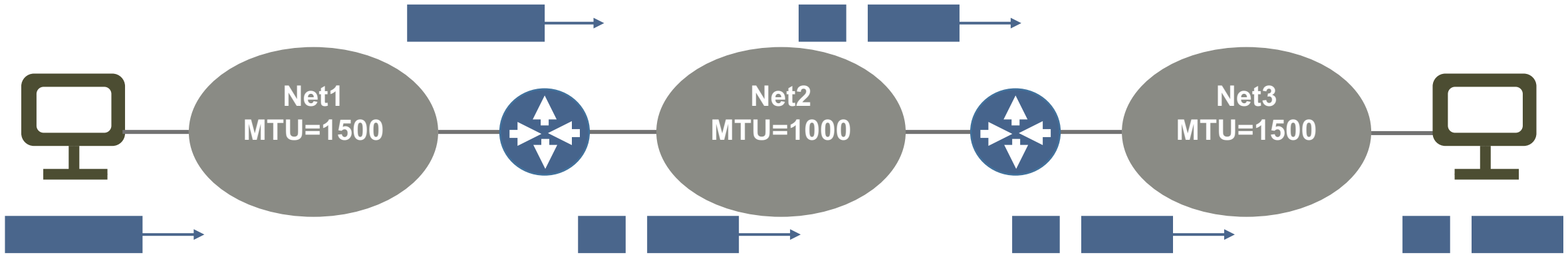
IP包重组

重组：在所有段的基础上重新产生原始数据报的过程。

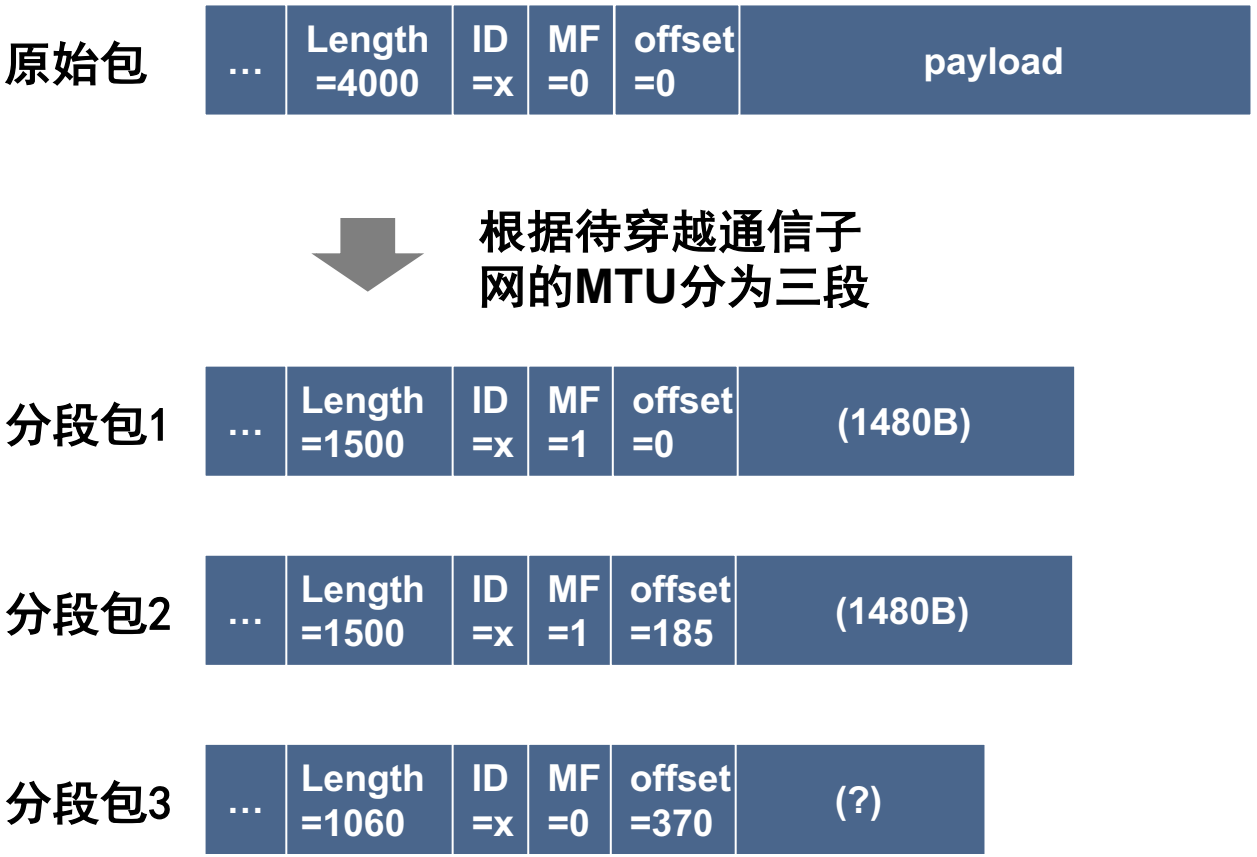
重组时机

- 在每个路由器进行
- 在最终目的地进行

IP标准规定只在最终目的地进行重组



IP分段与重组实例



假设：

- 包长4000字节（包括头）
- 通信子网MTU为1500字节

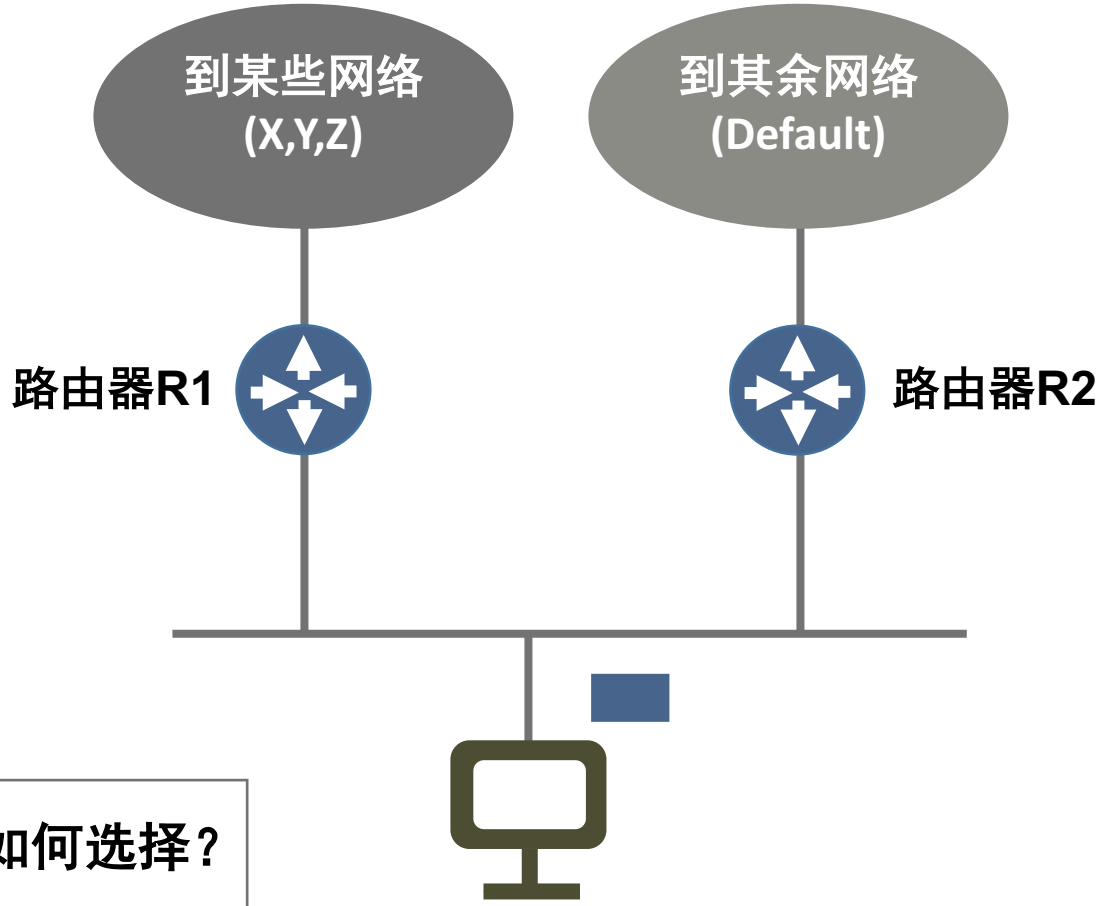
- Identification字段保持不变
- MF设置成1表示后面还有小包
- Offset表示包的有效载荷在原包中的偏移量，必须是8字节的倍数,例如：1480/8=185

IP协议之IP包投递

IP包传递

网络一般配置

- 每个路由器与两个或更多个物理网络直接连接
- 主机通常只与一个物理网络连接

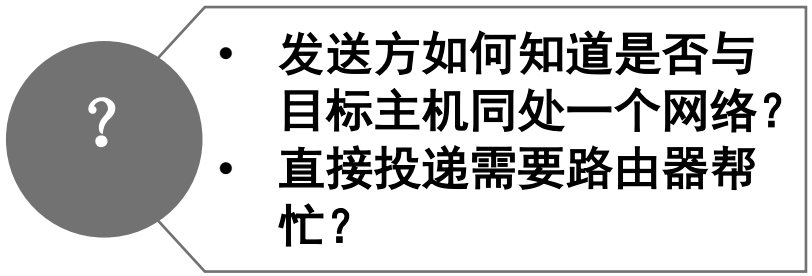


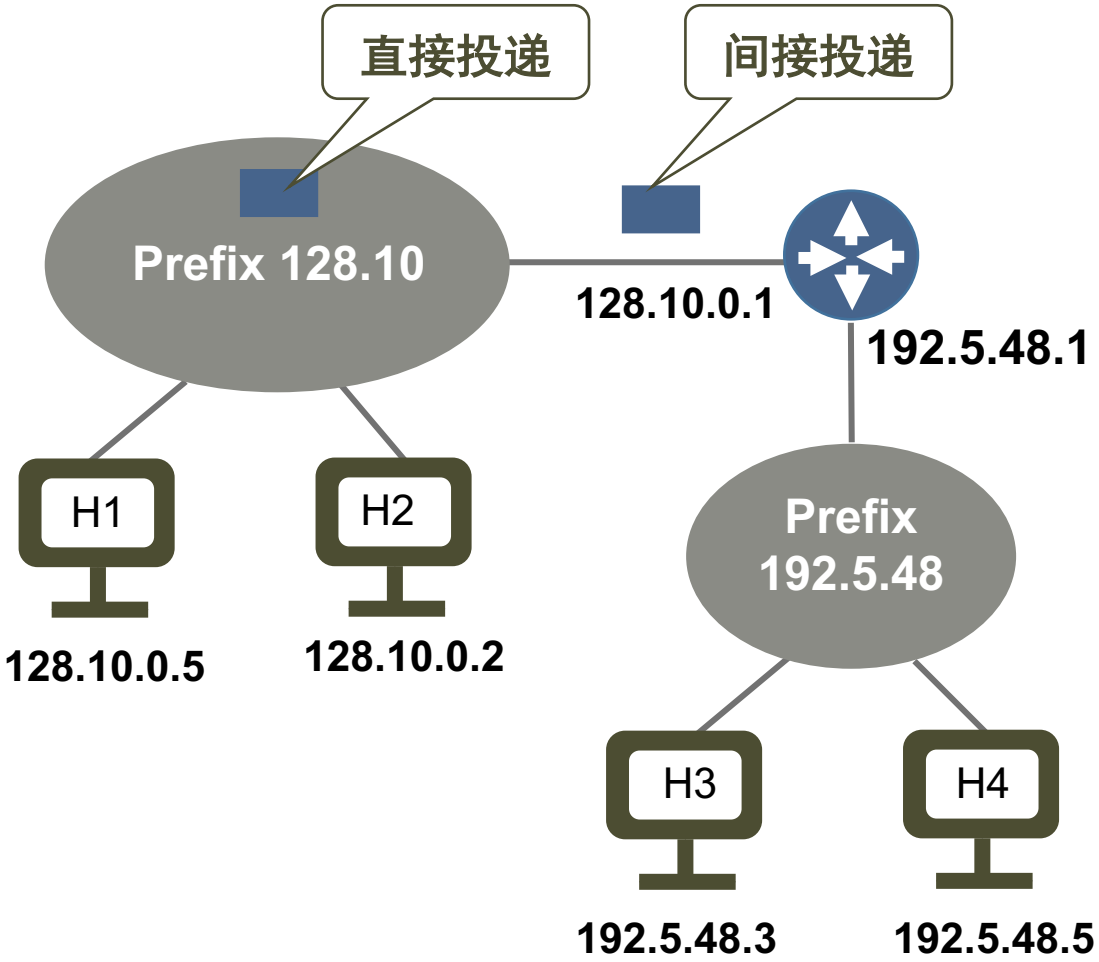
- 主机不需要运行路由协议
- 主机需要选择把数据包发给R1还是R2

? 主机如何选择？

IP包转发——直接投递

直接投递：指在一个物理网络上，数据包从一台机器直接被传送到另一台机器。





IP包转发——间接投递

间接投递：发送方必须把数据包发送给某个路由器，由该路由器把数据包转发到目的网络。

基于路由表的转发

- 决定网络地址的子网掩码
- 目的地的网络地址
- 下一个转发路由器的IP地址

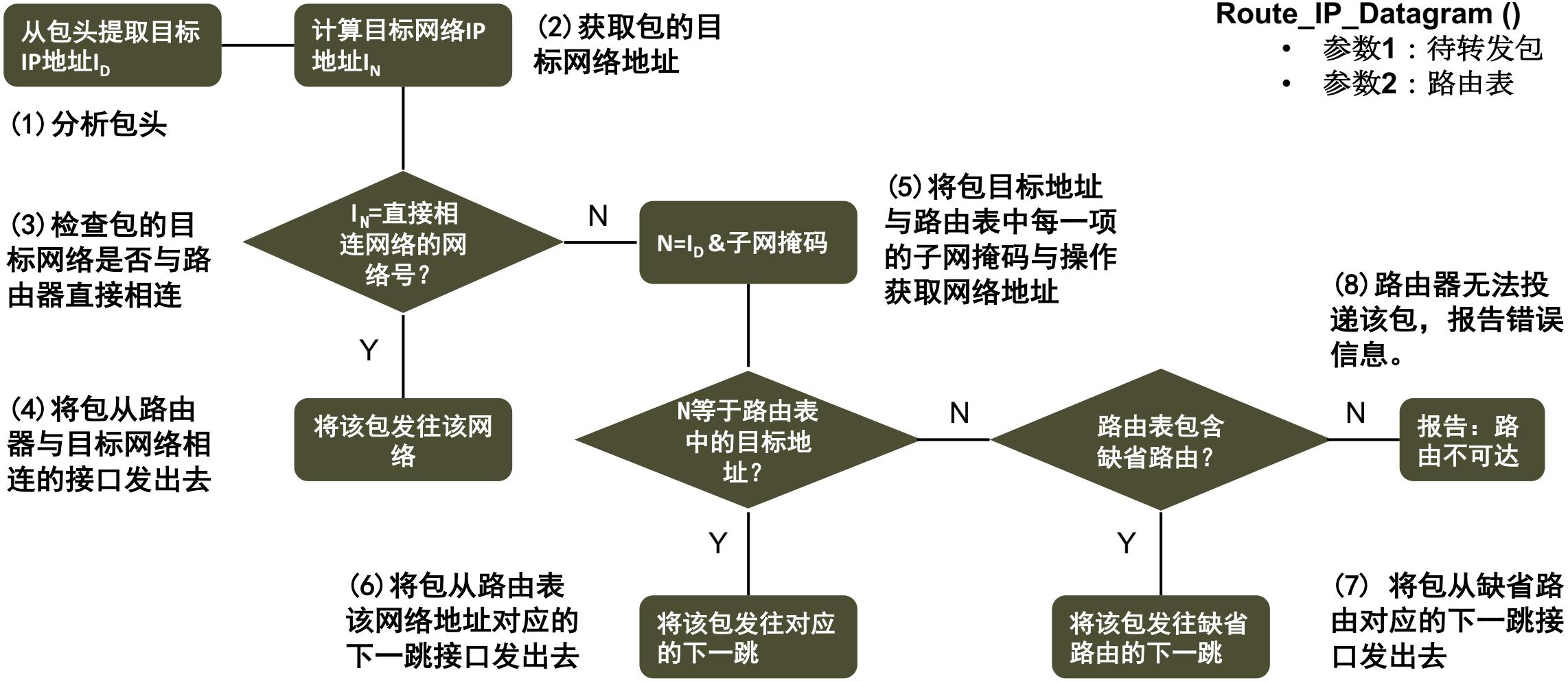
路由表

（目的地址，子网掩码，下一跳地址）

缺省路由适用于本地网络与因特网只有一个连接时的配置。

省缺路由：如果路由表中没有目的地的路由信息，则把数据包发送到一个缺省路由器上。

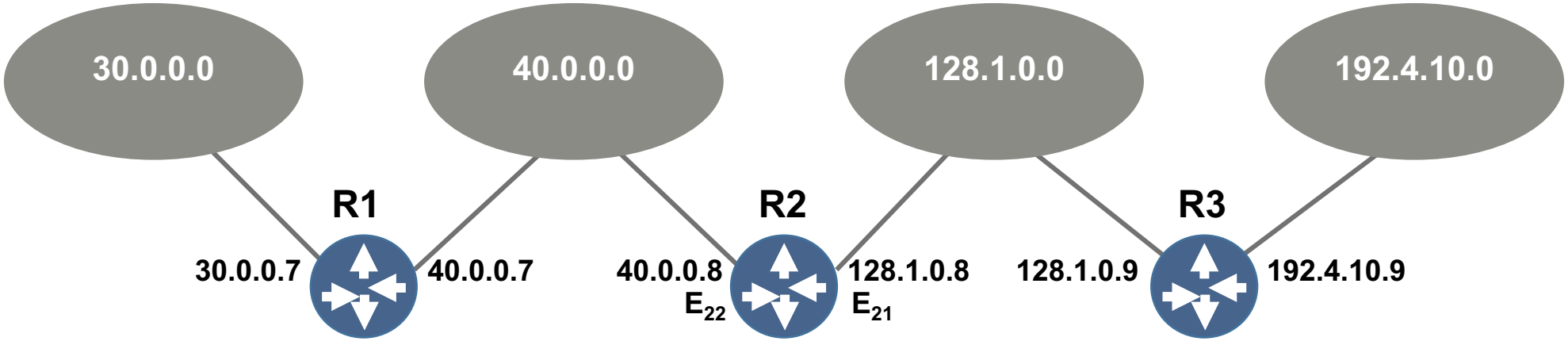
IP包的处理流程



Route_IP_Datagram ()

- 参数**1**：待转发包
- 参数**2**：路由表

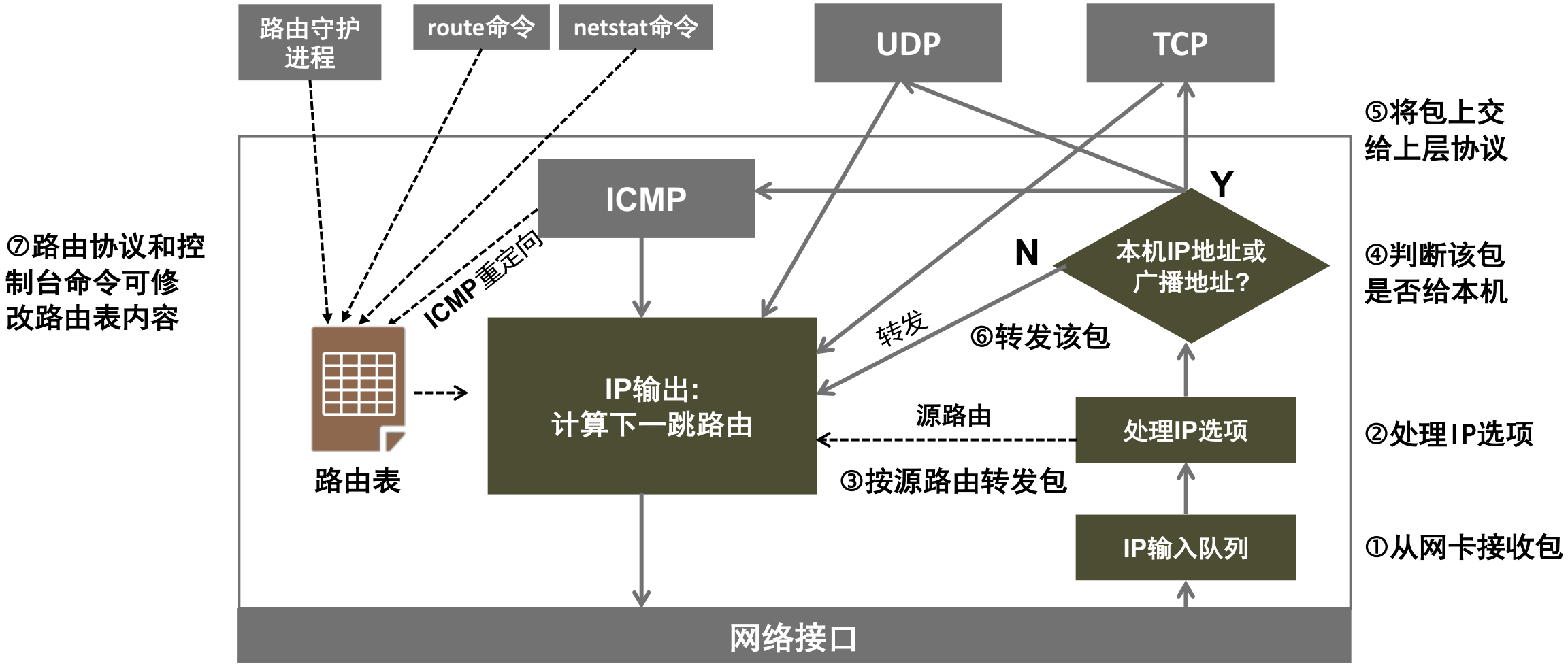
IP包转发示例



R2路由表

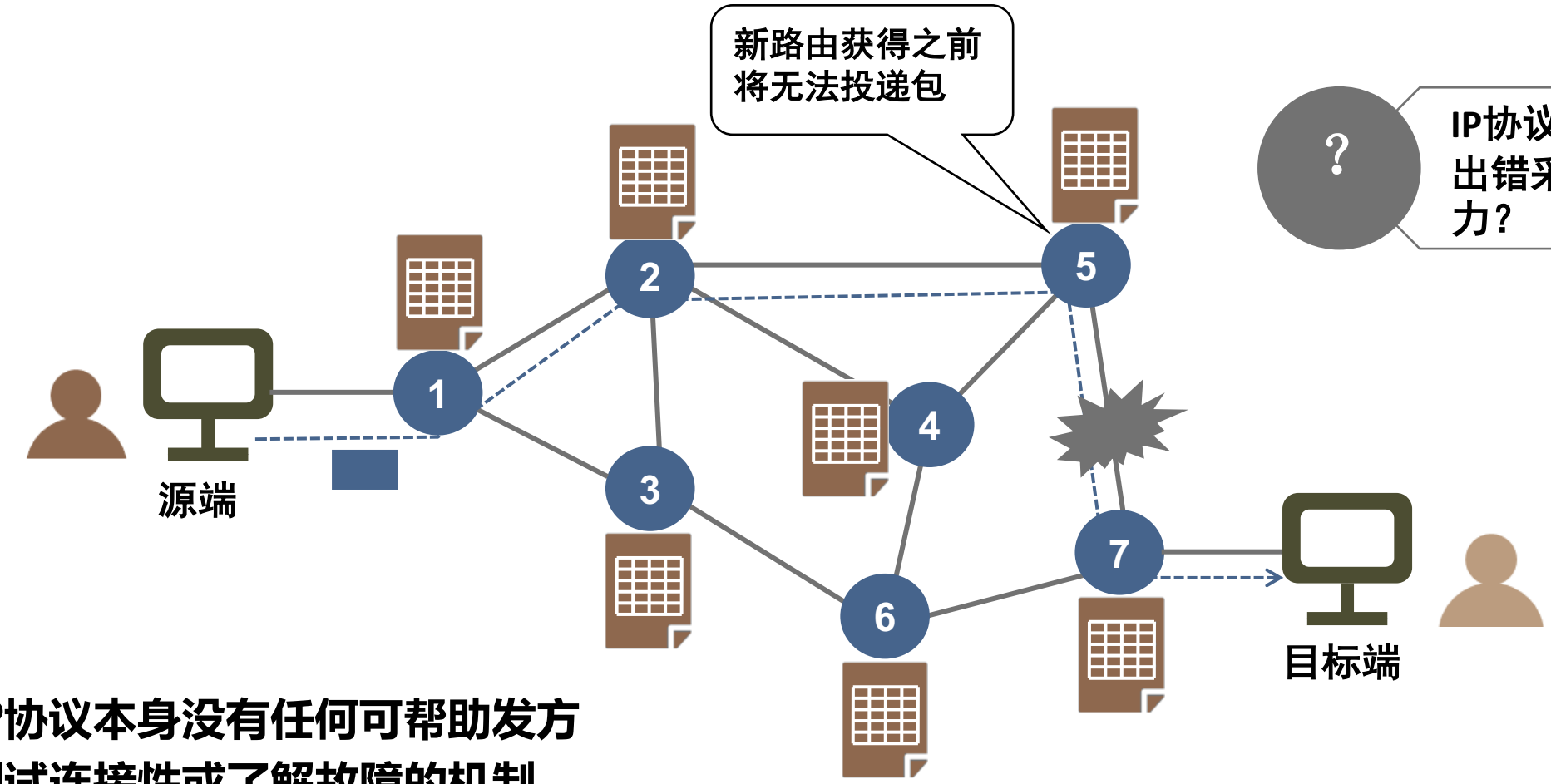
子网掩码	目标网络	下一跳	链路接口
255. 0. 0. 0	30. 0. 0. 0	40. 0. 0. 7	E_{22}
255. 0. 0. 0	40. 0. 0. 0	直接投递	E_{22}
255. 255. 0. 0	128. 1. 0. 0	直接投递	E_{21}
255. 255. 255. 0	192. 4. 10. 0	128. 1. 0. 9	E_{21}

IP层工作过程


路由守护进程
route命令
netstat命令
UDP
TCP
⑤将包上交给上层协议
ICMP
Y
⑦路由协议和控制台命令可修改路由表内容
N
本机IP地址或广播地址?
④判断该包是否给本机
ICMP重定向
转发
⑥转发该包
IP输出:
计算下一跳路由
源路由
处理IP选项
②处理IP选项
路由表
③按源路由转发包
IP输入队列
①从网卡接收包
网络接口
控制流
报文流

案例学习
ICMP协议与包投递错误报告

尽力而为的IP传递服务

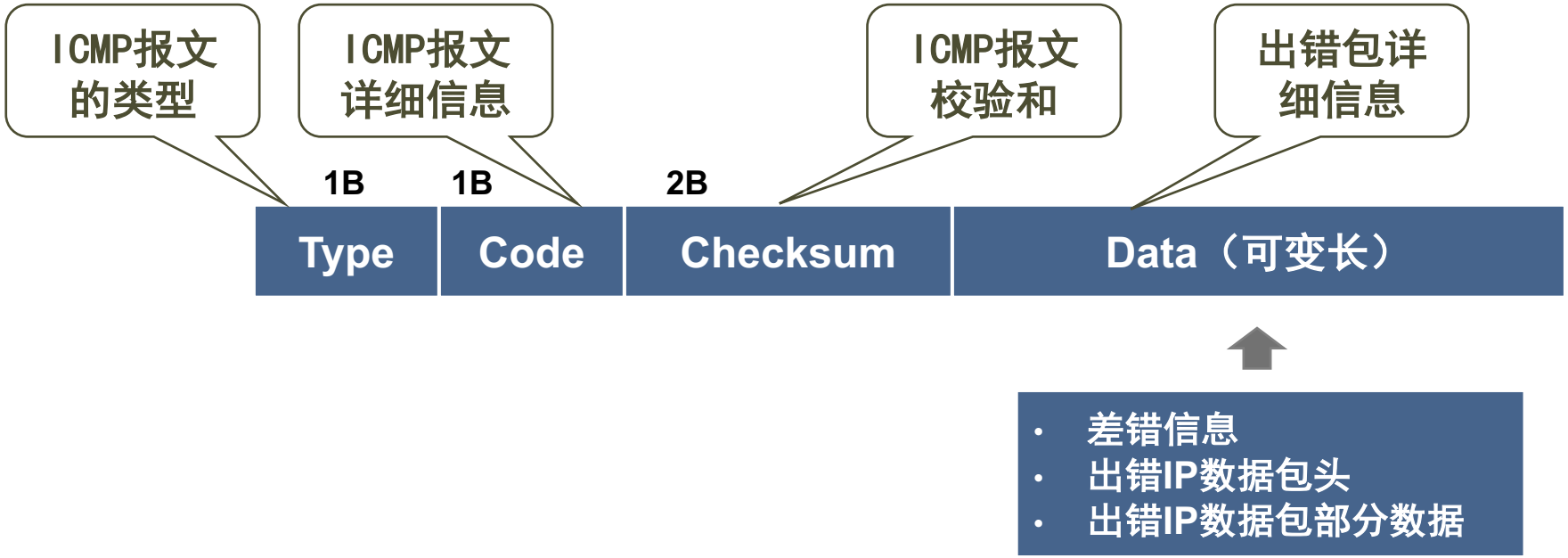


IP协议本身没有任何可帮助发方测试连接性或了解故障的机制。

因特网控制报文协议

ICMP协议是IP协议的辅助协议，负责在包投递过程中出现错误时向包源端报告错误信息。

RFC792



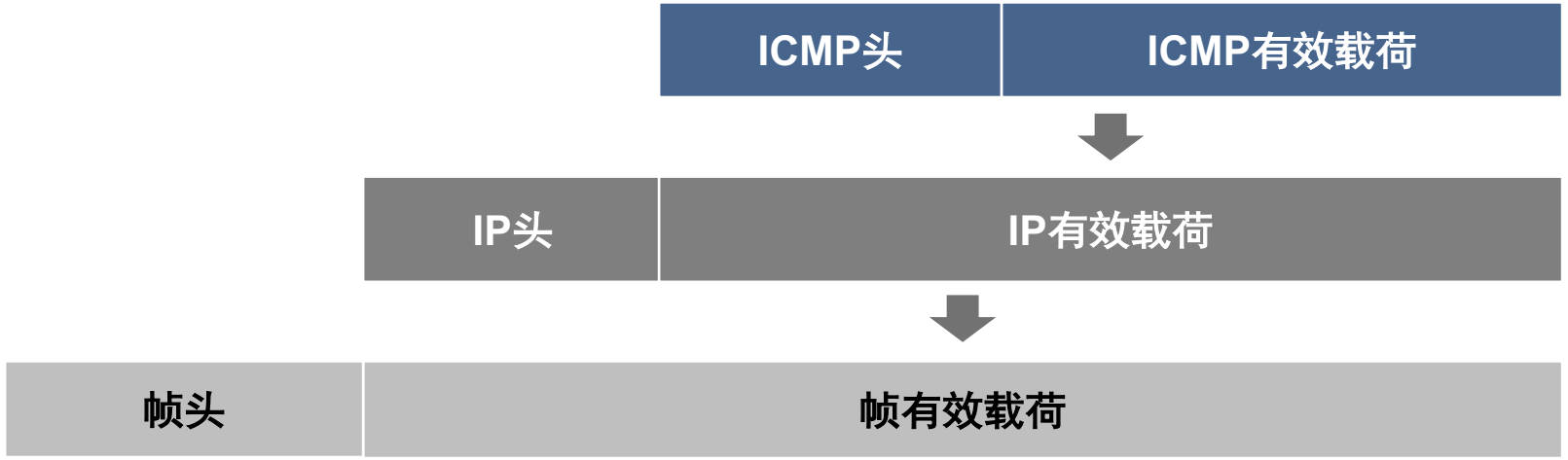
ICMP报文传递与特点

ICMP的特点

- ICMP不是高层协议
- ICMP不具备可靠性和优先级
- 携带ICMP报文的IP包传递出错时不再报告
- 携带ICMP报文的IP包与携带数据的IP包具有完全相同的路由选择

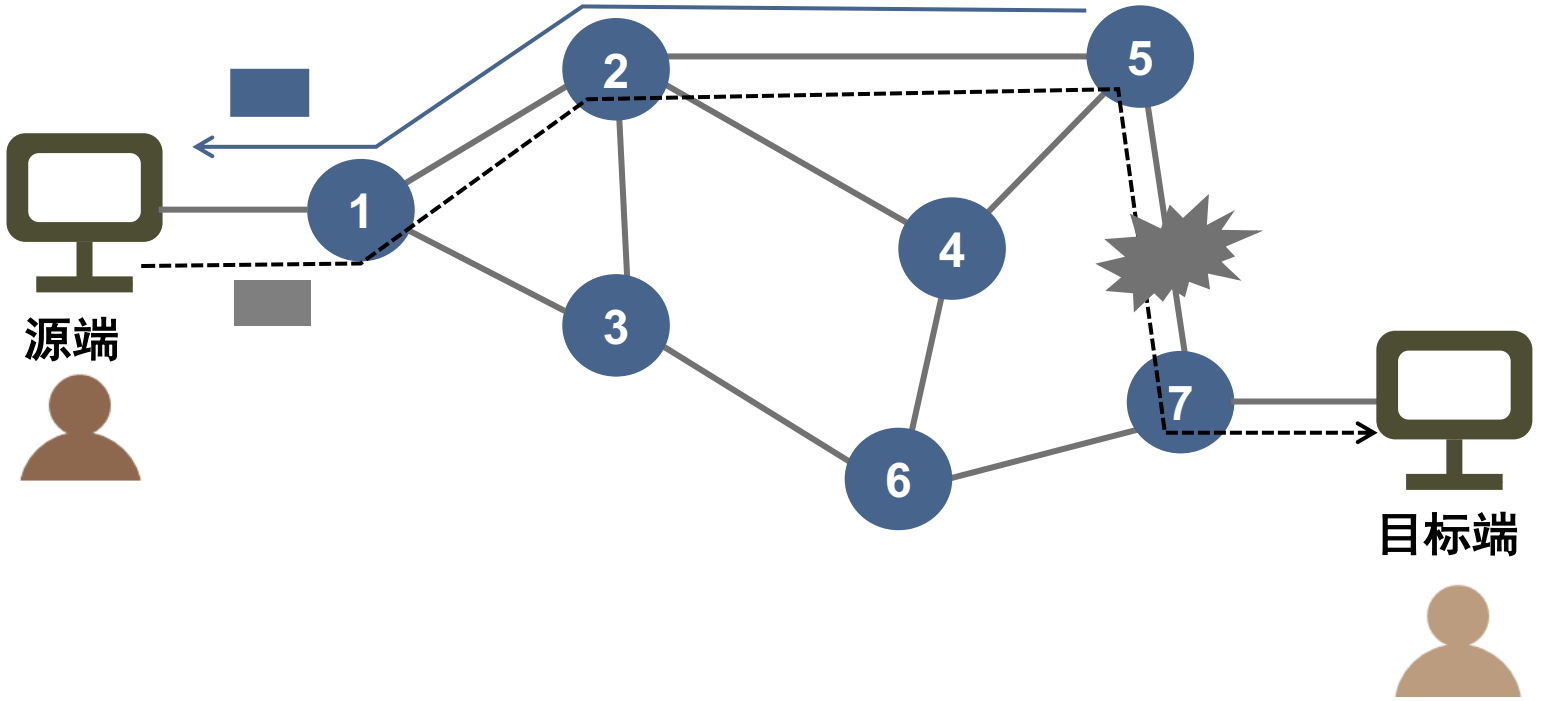
ICMP报文封装在IP数据报的有效负载部分，ICMP报文的最终目的地是处理它的Internet协议软件。

? ICMP报文能否封装在链路层帧内传输？



ICMP工作过程

- 当包投递出现差错时ICMP向包的源端报告差错情况
- 源端必须把差错交给某个应用程序或采取其他措施来纠正



- 数据包头没有历史路径信息
- 路由器不具备全局知识

ICMP协议之基本功能

检测任意节点的可达性与状态

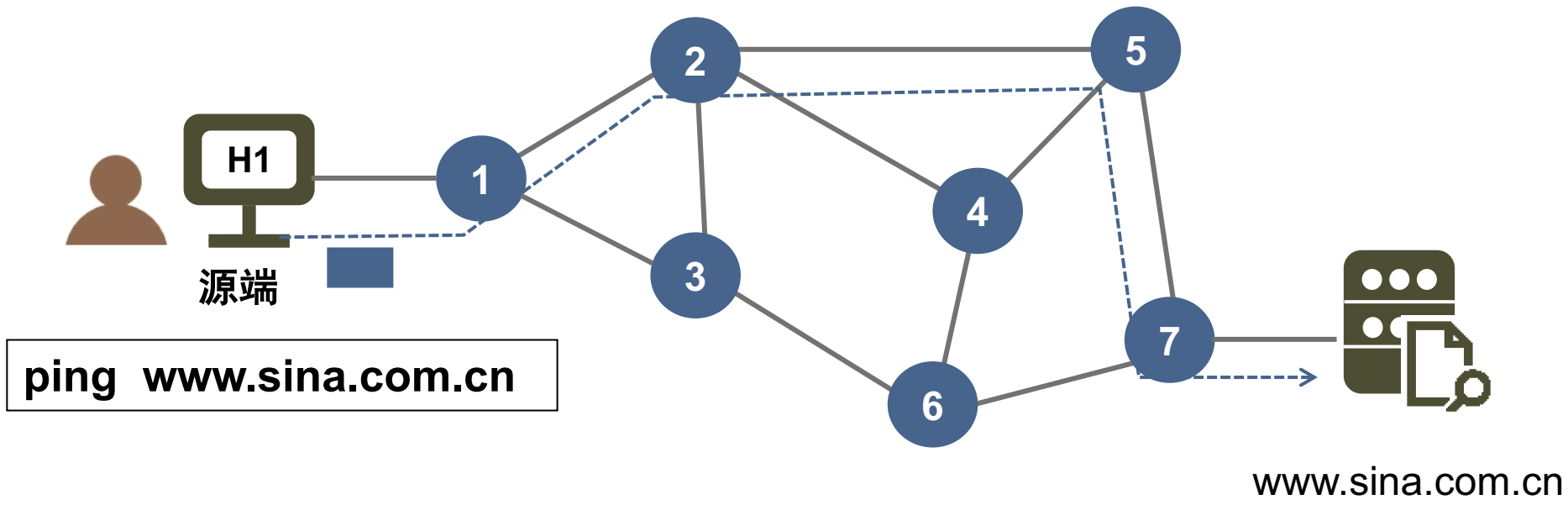
- **检测目标节点是否可达**
  - ① 主机或路由器向指定目标发送ICMP ECHO请求报文，请求报文包含一个可选的数据区
  - ② 收到ECHO请求报文(8)的机器应立即回应一个应答报文(0)，应答报文包含了请求报文中数据的拷贝

? 哪些协议要用到ECHO功能？

Type(8/0)	Code(0)	checksum
Identifier		Sequence No.
Optional data		
.....		

检测可达性示例——PING命令

根据协议，服务器在收到类别是8的ICMP报文后，立即将报文原封不动返回给源端（ICMP的类别是0）。



检测可达性示例——PING命令

路径时延不同

```
>ping www.sina.com.cn

正在 Ping spool.grid.sinaedge.com [58.205.212.206] 具有 32 字节的数据:
来自 58.205.212.206 的回复: 字节=32 时间=4ms TTL=53
来自 58.205.212.206 的回复: 字节=32 时间=12ms TTL=53
来自 58.205.212.206 的回复: 字节=32 时间=12ms TTL=53
来自 58.205.212.206 的回复: 字节=32 时间=6ms TTL=53

58.205.212.206 的 Ping 统计信息:
    数据包: 已发送 = 4，已接收 = 4，丢失 = 0 (0% 丢失)，
往返行程的估计时间(以毫秒为单位):
    最短 = 4ms，最长 = 12ms，平均 = 8ms
```

路径长度

报告目标端不可达报告

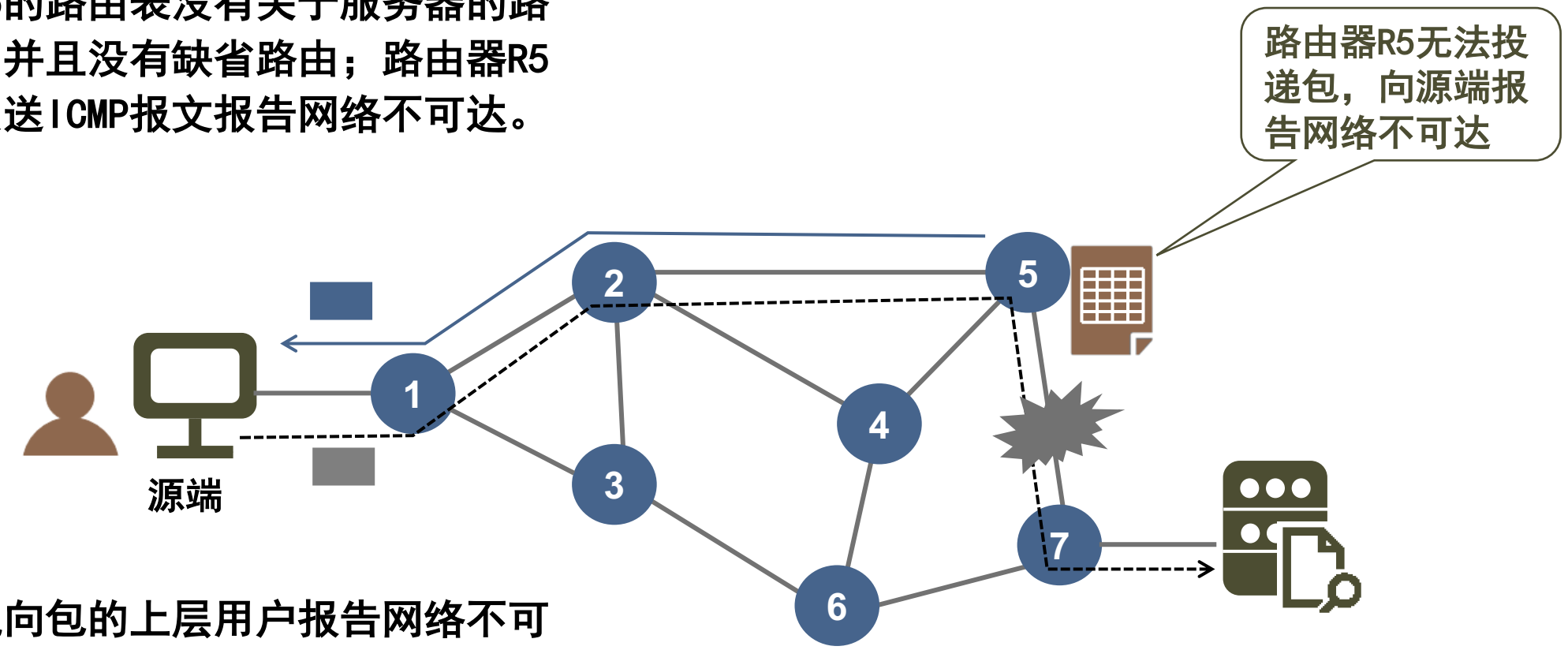
- **当路由器无法投递包时**
  - ① 向源端发回一个目标端不可达报文，并丢弃该包。

Type(3)	Code(0~12)	Checksum
Unused (must be 0)		
IP Header + 64B		
.....		

代码	意义
0	网络不可达
1	主机不可达
2	协议不可达
3	端口不可达
4	需要分段但DF置位
5	源路由失败
6	目的网络未知
7	目的主机未知

报告目标端不可达报告示例

路由器R5的路由表没有关于服务器的路由信息，并且没有缺省路由；路由器R5向源端发送ICMP报文报告网络不可达。



源端主机向包的上层用户报告网络不可达信息，由上层用户决定如何处理。

拥塞控制通知

- 当路由器因缓存溢出不得不丢包时
  - ① 向源端发回一个拥塞报文
  - ② 源端拥塞控制据此采取相应措施

拥塞形成原因

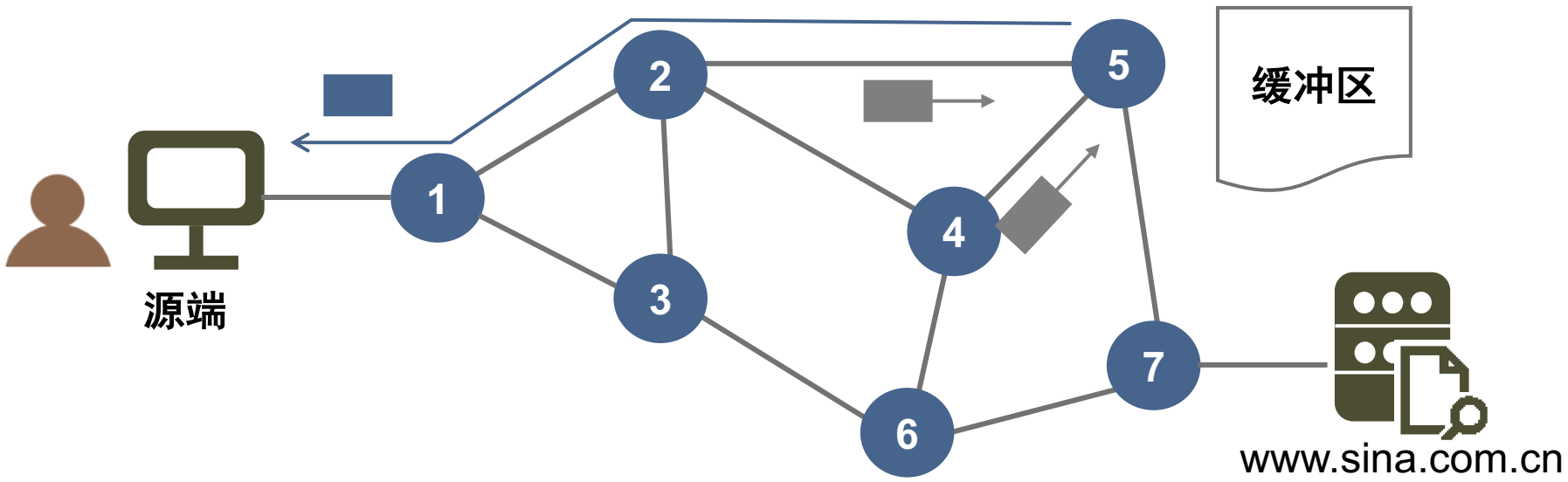
- 高速计算机产生的通信量比网络能传输的包多时
- 许多计算机发送的包同时需要通过某个路由器时

Type (4)	Code (0)	Checksum
Unused (0)		
IP Header + 64B		
.....		

拥塞控制通知示例

假设：R2和R4的路由表将抵达服务器的路径都指向下一跳R5。

发生拥塞的路由器R5为每个丢弃的包发送一个拥塞报文给包源端。



ICMP报文　网络层包

重定向路由

- **主机的路由学习能力**
  - ① 假定路由器知道正确路由
  - ② 主机从最少路由信息开始逐渐从路由器了解新路由信息

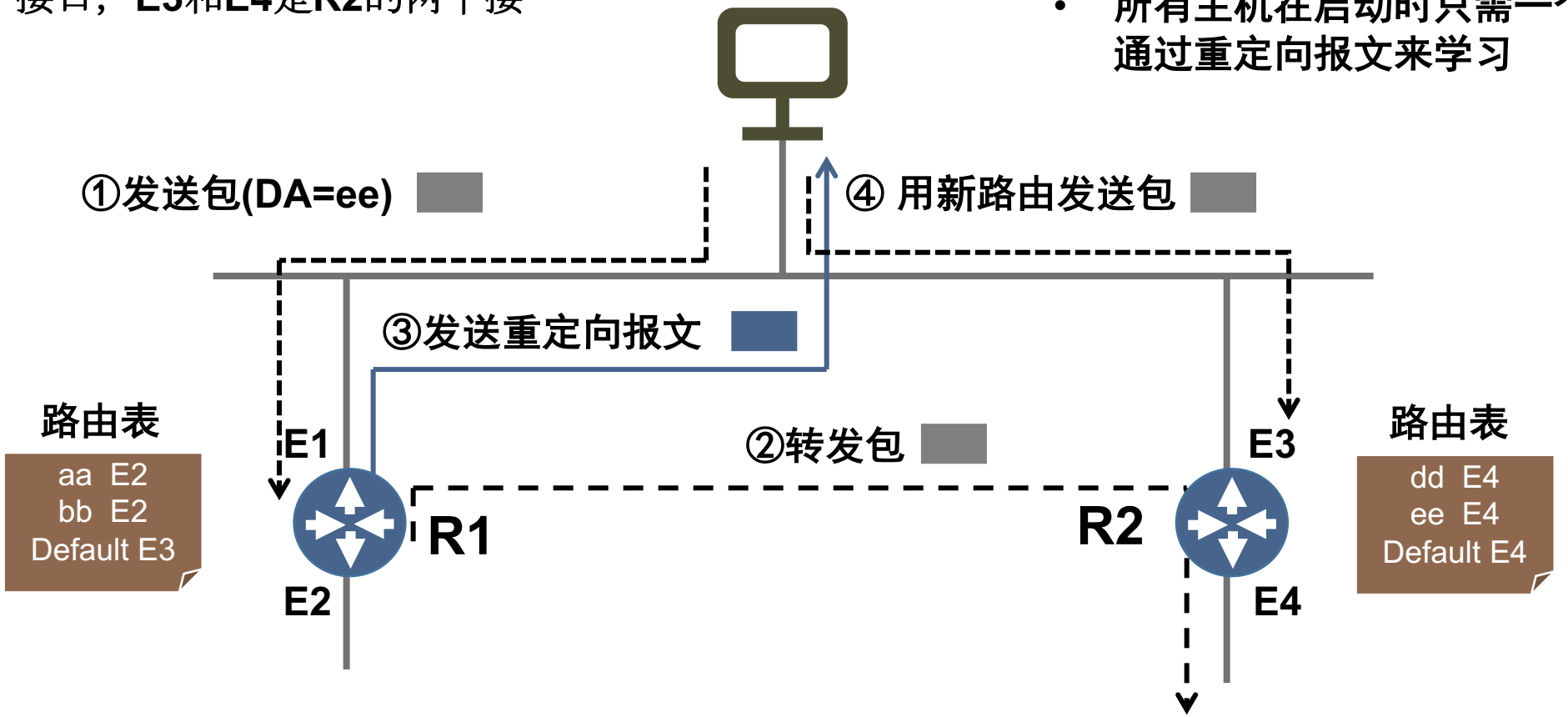
当路由器检测到主机使用了一条非优化路由时就向主机发送一个重定向的ICMP报文，请求主机改变路由，同时转发初始数据报。

Type (5)	Code (0~3)	checksum
Router IP address		
IP Header + 64B		
.....		

重定向路由示例

假设 主机H2给目标地址ee发送一个包，路由器R1和R2的路由表如图所示，E1和E2是R1的两个接口，E3和E4是R2的两个接口。

- 一旦路由发生差错省缺路由器通知主机进行重定向
- 所有主机在启动时只需一个省缺路由，通过重定向报文来学习

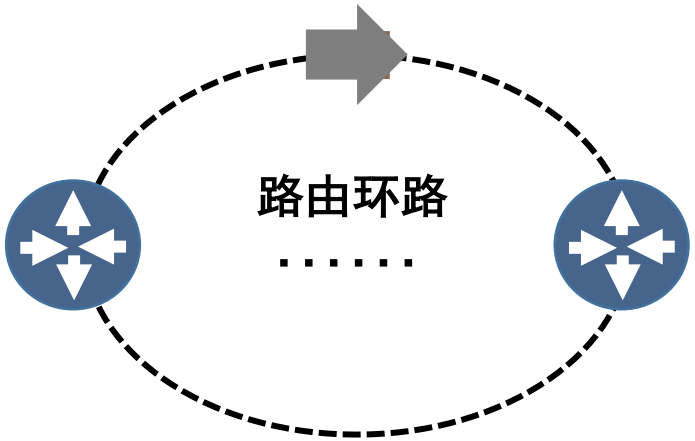


ICMP报文　网络层包

检测循环路由

- **应付错误路由/丢包**
  - ① 一旦路由器因包的TTL为0或主机等待包重组超时而丢弃该包时，向源端发回一个ICMP超时报文。

？ TTL何时改变？
TTL变了影响什么？



Type (11)	Code(0~1)	Checksum
unused(must be 0)		
IP Header + 64B		
……		

0：生存期超时
1：包重组超时

传输时间估计值

- **网络延迟的估算**
  - ① 计算请求报文(13)到目的地、被转换成应答报文(14)及返回所需的时间。

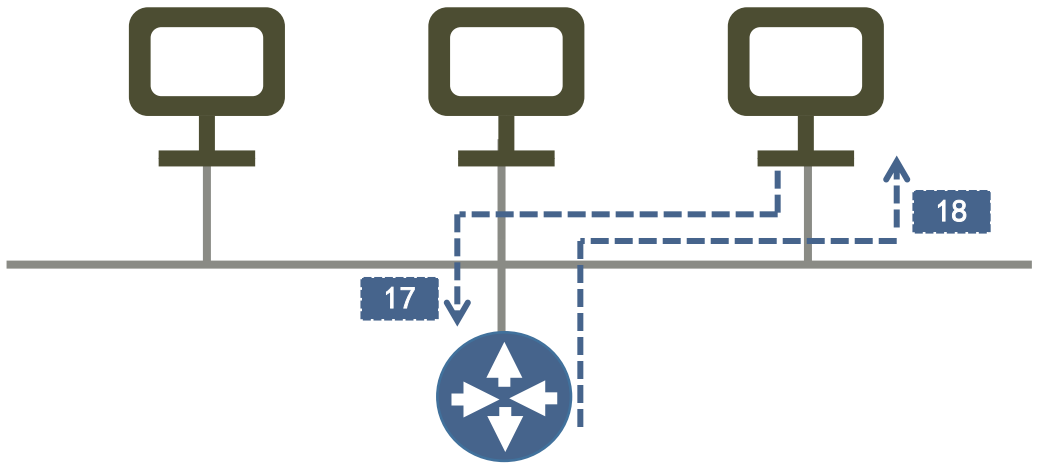
Type(13/14)	Code(0)	Checksum
Identifier		Sequence No.
Send timestamp		
Receive timestamp		
Return timestamp		

请求子网地址掩码

- **子网掩码的获取**

① 为了解本地网络使用的子网掩码，主机可向路由器发出一个地址掩码请求报文（17），并接收一个地址掩码应答报文（18）。

type(17/18)	Code (0)	Checksum
Identifier		Sequence No.
Subnet mask		

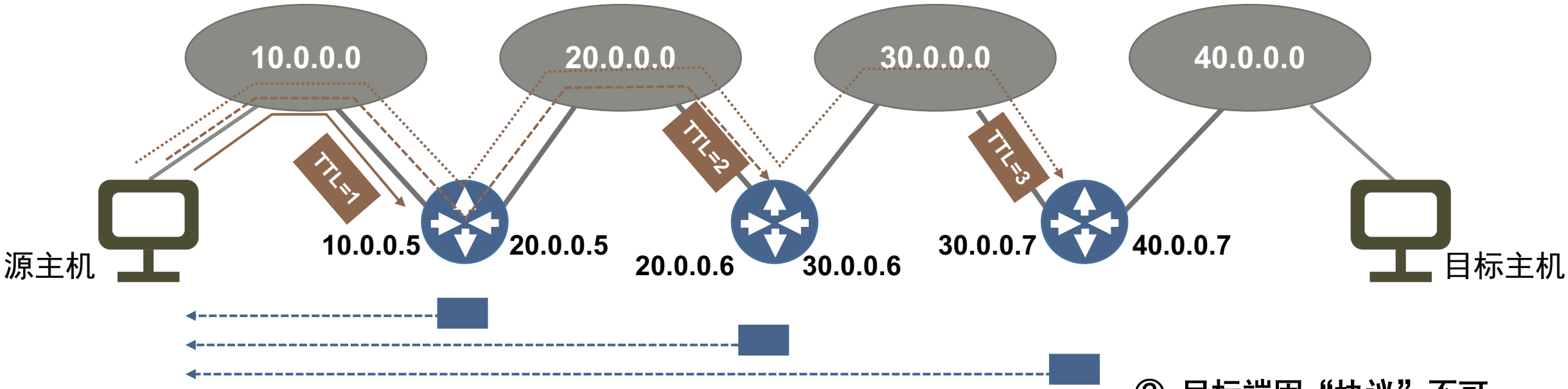


ICMP协议之综合应用示例

ICMP其他应用示例1——跟踪路径

示例1：**trace route**工具

利用**ICMP**超时报文发现到目的地一条路径上的路由器列表。

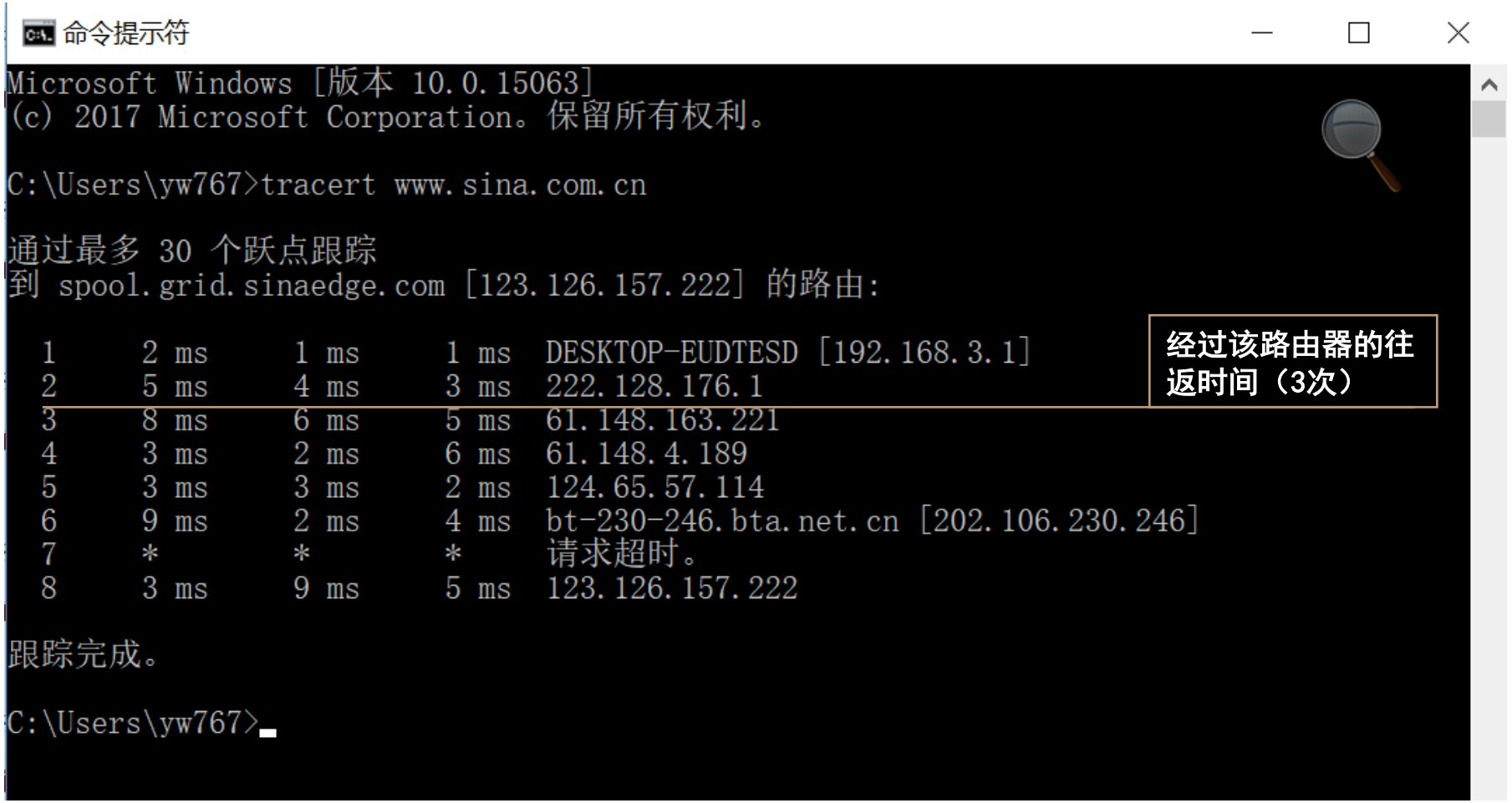


① 源主机依次给目标节点发出TTL递增的包

② 路由器丢弃包，向源端返回ICMP报文

③ 目标端因“协议”不可达，向源端返回ICMP报文，终止探测过程。

发现到某个目标地址的路径



示例1：利用**tracert**命令发现到目标地址的路径

C:> tracert www.sina.com.cn

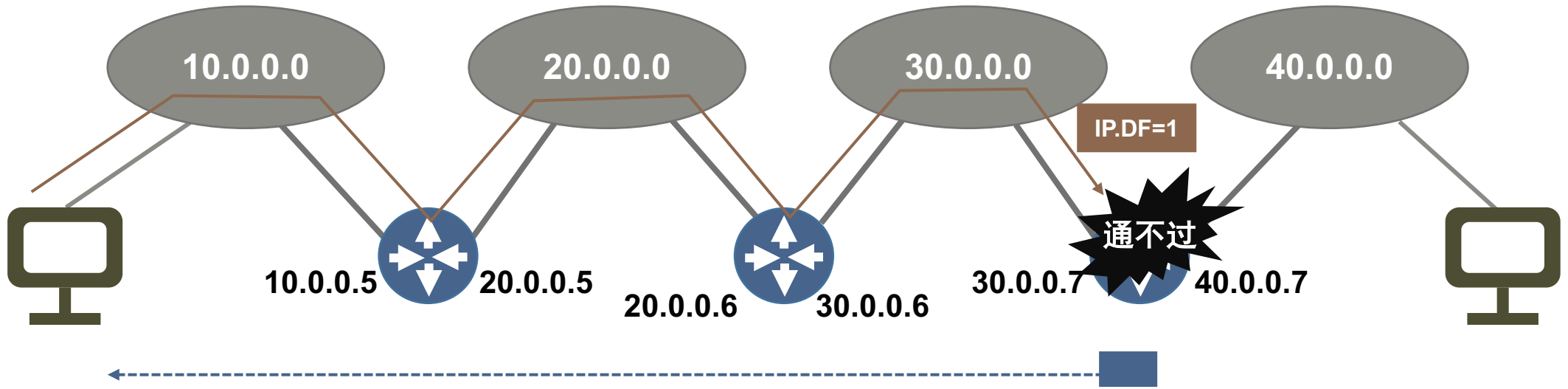
- **本地为内网地址 192.168.1.101**
- **缺省网关 192.168.1.1**

ICMP其他应用示例2——发现路径MTU

示例2：用ICMP发现路径MTU

方法：

- 利用包头中的“不能分(DF)”标志位
- 通过探测的方式发现当前路径上的最小MTU



① 源主机依次给目标节点发出包长度递减的包

② 路由器丢弃超长又不允许分段的包，向源端返回ICMP消息。

ICMP其他应用示例2——发现路径MTU

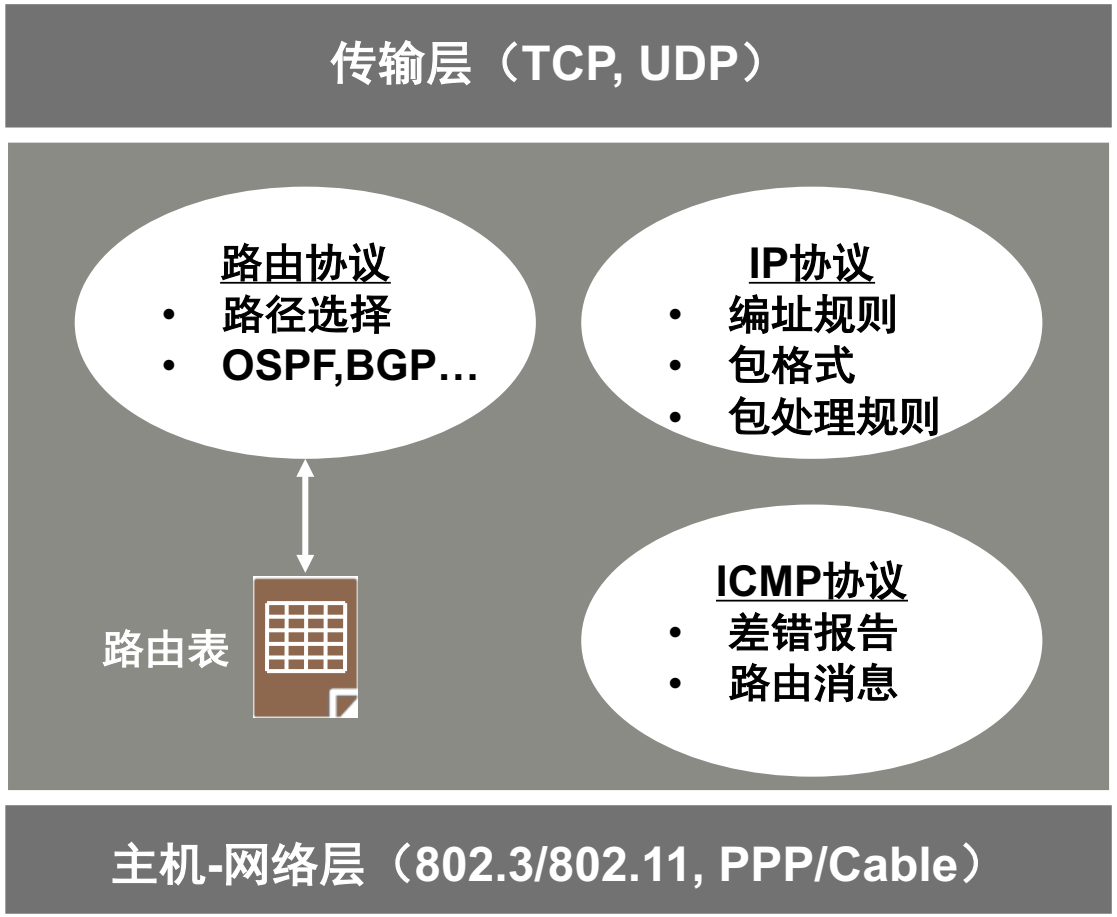
?

- 什么时候需要知道路径MTU？
- 提前知道路径MTU有什么用？

好处

- 途径的路由器无需对包进行分段
- 目的主机无需重组完整包

IP协议实体内部视图


传输层（TCP, UDP）
路由协议
• 路径选择
• OSPF,BGP…
IP协议
• 编址规则
• 包格式
• 包处理规则
路由表
ICMP协议
• 差错报告
• 路由消息
主机-网络层（802.3/802.11, PPP/Cable）

案例学习

ARP协议与包投递

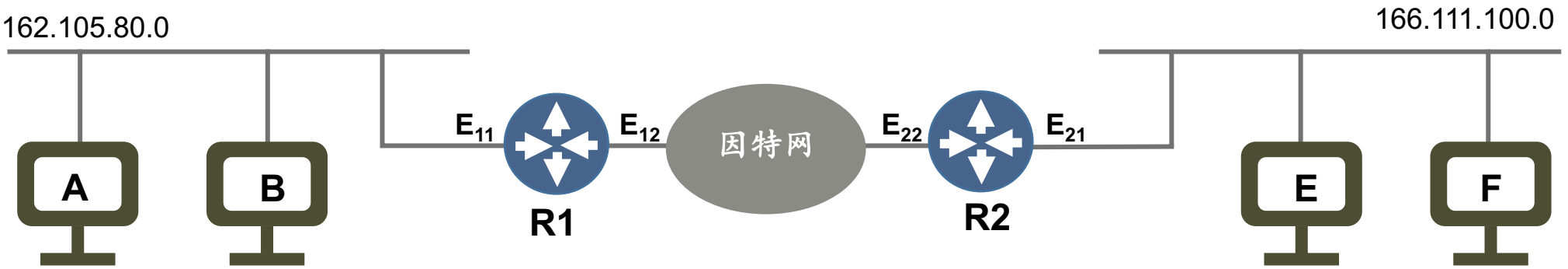
IP地址 vs.MAC地址

不能携带着走

IP地址特点
••IP地址类似于邮政编码，特定于某个地理位置 ••IP地址具有层次特点，取决于节点所连的那个子网

MAC地址特点
••MAC地址类似于身份证号码，可以不受限制地移动 ••MAC地址平面结构，网卡可从一个LAN移动到其他LAN

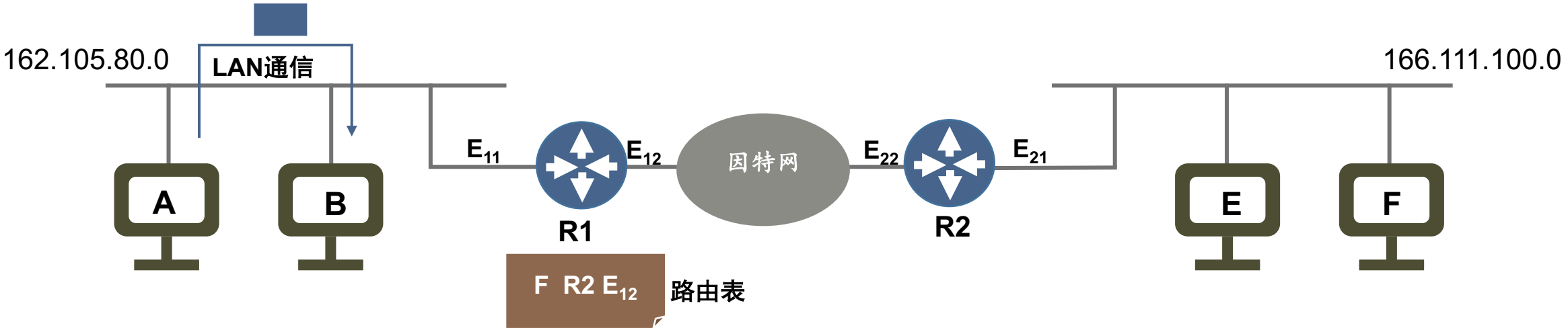
可携带着走



LAN通信中的IP地址与MAC地址

假设：主机A给主机B发送一个IP包，考察该包的完整发送过程。

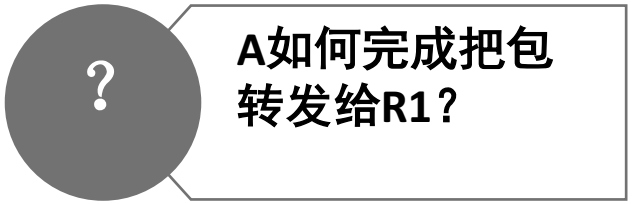
- A获得包的目标网络号为162.105.80.0
- A判断该包属于局域网内部通信
- A把包下交给网卡
- 网卡把包封装成数据帧



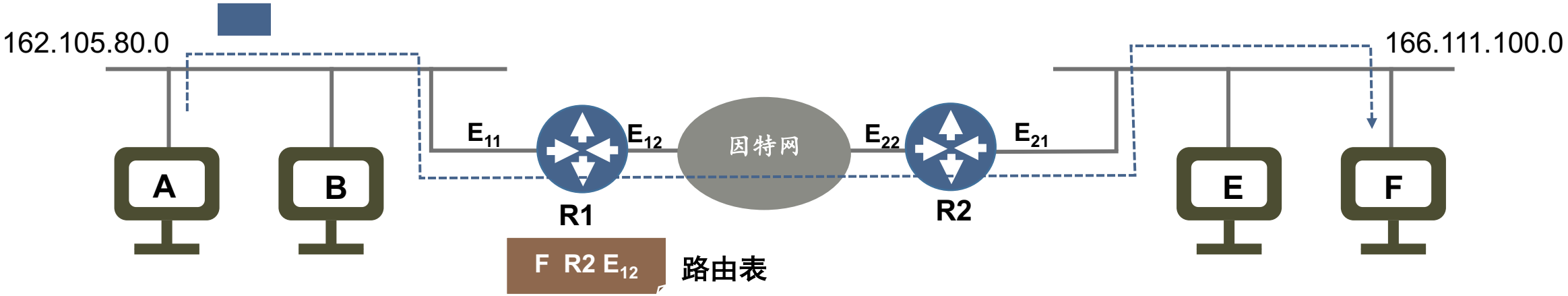
远程通信中的IP地址与MAC地址

假设：主机A给主机F发送一个IP包，考察该包的完整传递过程。

- 主机A把包发给本地路由器R1
- R1根据路由表指示把包转发给R2
- R2把包发给本地网络上的主机F



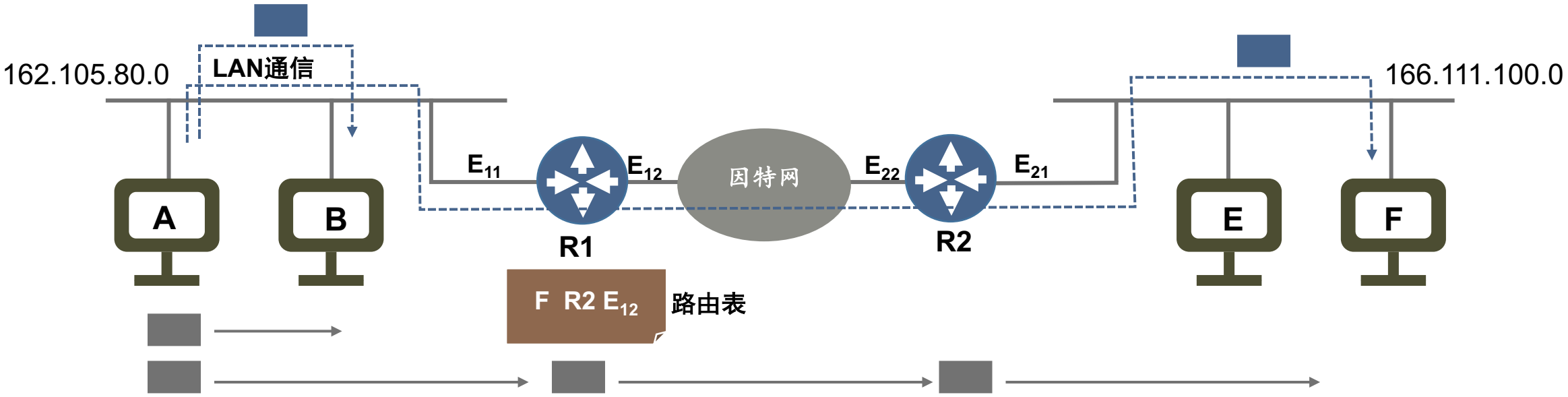
A-F通信由 A-R1 ， R1-R2 ， R2-F 组成。



地址解析在下一跳传递中的作用

- **发送/转发报文时**
  - ① 网卡不知道IP地址后缀（主机号）与特定主机的关系
  - ② 网卡不知道如何用IP地址来定位一台主机

A-F通信由 A-R1 ， R1-R2 ， R2-F 组成。

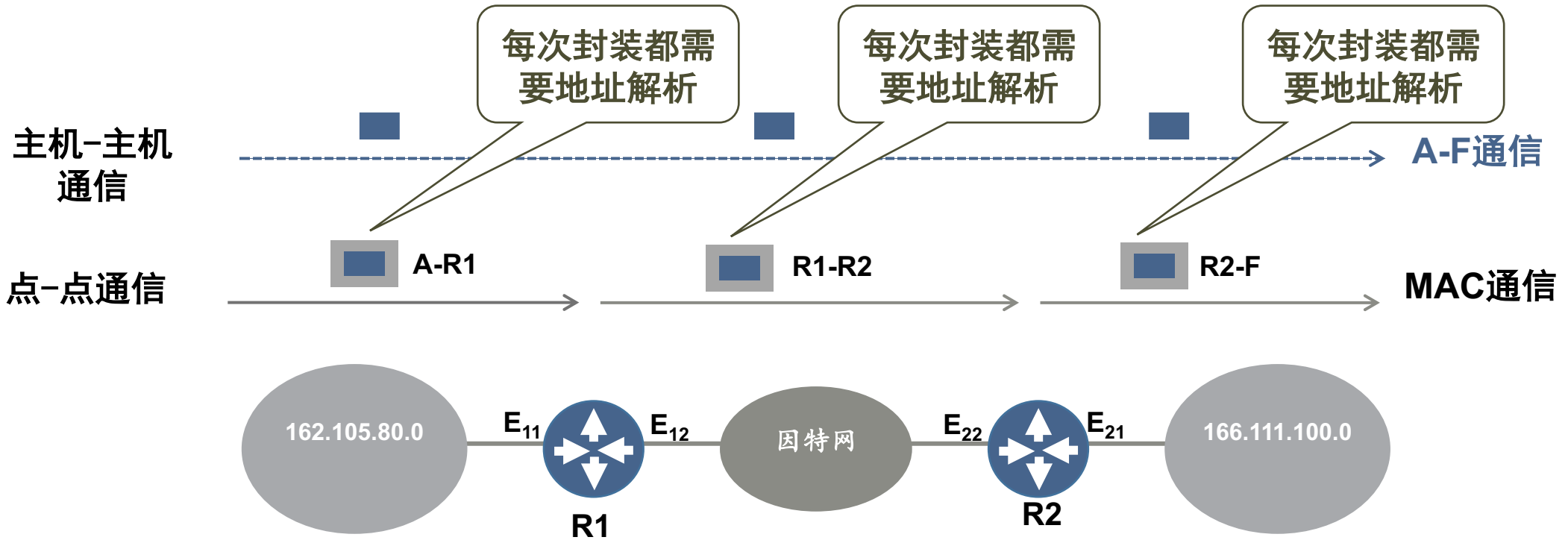


数据帧 网络层包

地址解析技术

地址解析：将IP地址解析成相应的硬件地址的过程。

注意：一台计算机只需要解析连在同一网络上计算机地址.



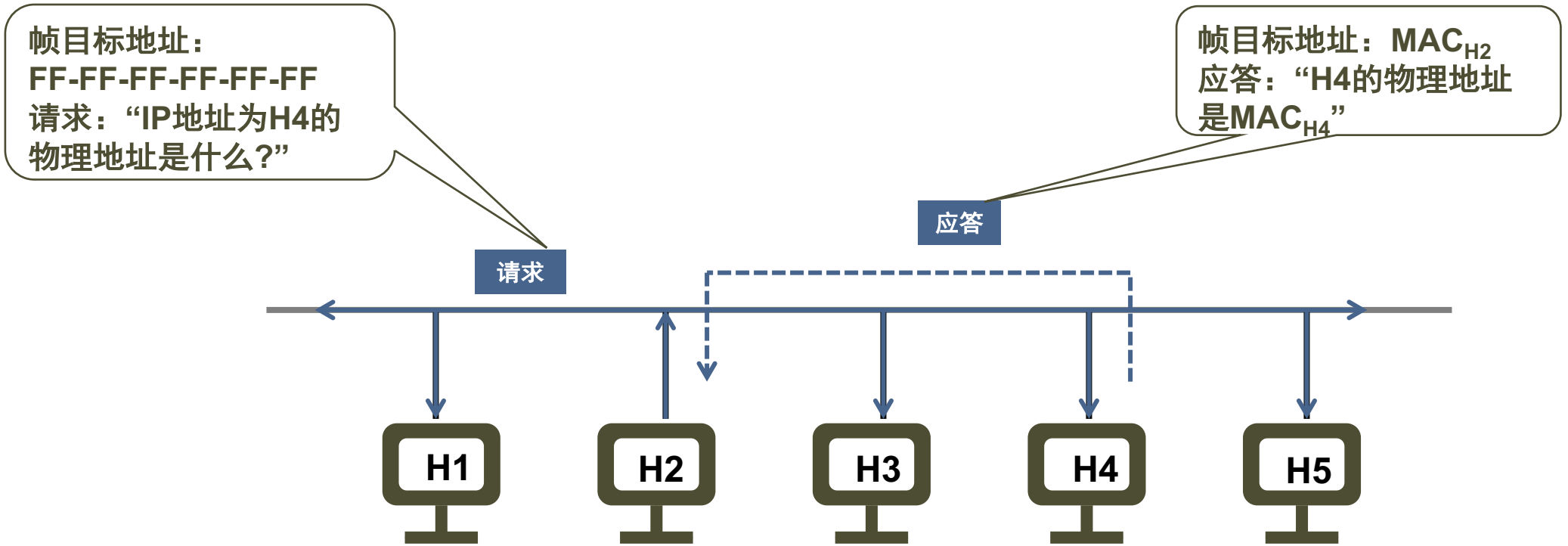
数据帧 网络层包

ARP协议格式与投递

基于动态消息交换的地址解析

动态消息交换法:需要解析地址时通过网络通信获得IP地址对应的物理地址。

例如：H2要给H4发送一个包.



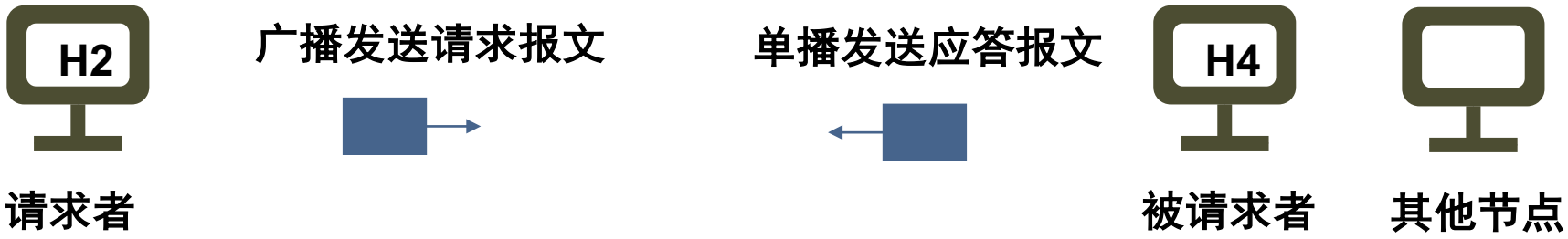
因特网地址解析协议(ARP)

因特网地址解析协议基于动态消息交换法，定义了两类基本消息：请求和应答。

RFC826

ARP协议规定

- 一个ARP报文被放入一个硬件帧后，被广播给网上的所有主机
- 每台主机收到该请求后都会检测其中的IP地址
- 与IP地址匹配的主机发送一个应答报文
- 其他的主机则会丢弃收到的请求，不发任何应答。



ARP报文格式

硬件地址类型：1表示以太网地址
协议地址类型：0x800表示IP地址
硬件地址长度：6表示以太网地址
协议地址长度：4表示IP地址

Hardware Address Type		Protocol Address Type
Length of HA	Length of PA	Operation（Request/Reply）
Sender HD (0~3B)		
Sender HD(4~5B)		Sender PD(0~1B)
Sender PD(2~3B)		Receiver HD (0~1B)
Receiver HD(2~5B)		
Receiver PD (0~3B)		

ARP协议特点

- 为硬件地址引入一个地址长度字段
- 为协议地址引入一个地址长度字段

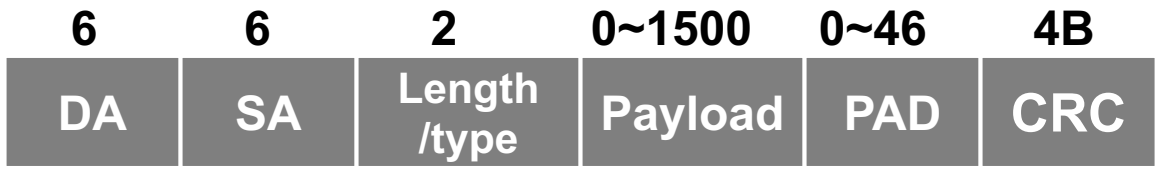
- Hardware Addr. Type：规定了硬件地址类型
- Protocol Addr. Type：规定了协议地址类型
- Length of HA：规定了硬件地址的长度
- Length of PA：规定了协议地址的长度
- Sender HD：ARP报文发送方硬件地址
- Sender PD ARP：报文发送方协议地址
- Receiver HD ARP：报文接收方硬件地址
- Receiver PD ARP：报文接收方协议地址

ARP报文封装

IP包和ARP报文均通过链路层的数据帧携带传输，并且数据帧对于其包括IP包还是ARP报文没有区别对待。

？

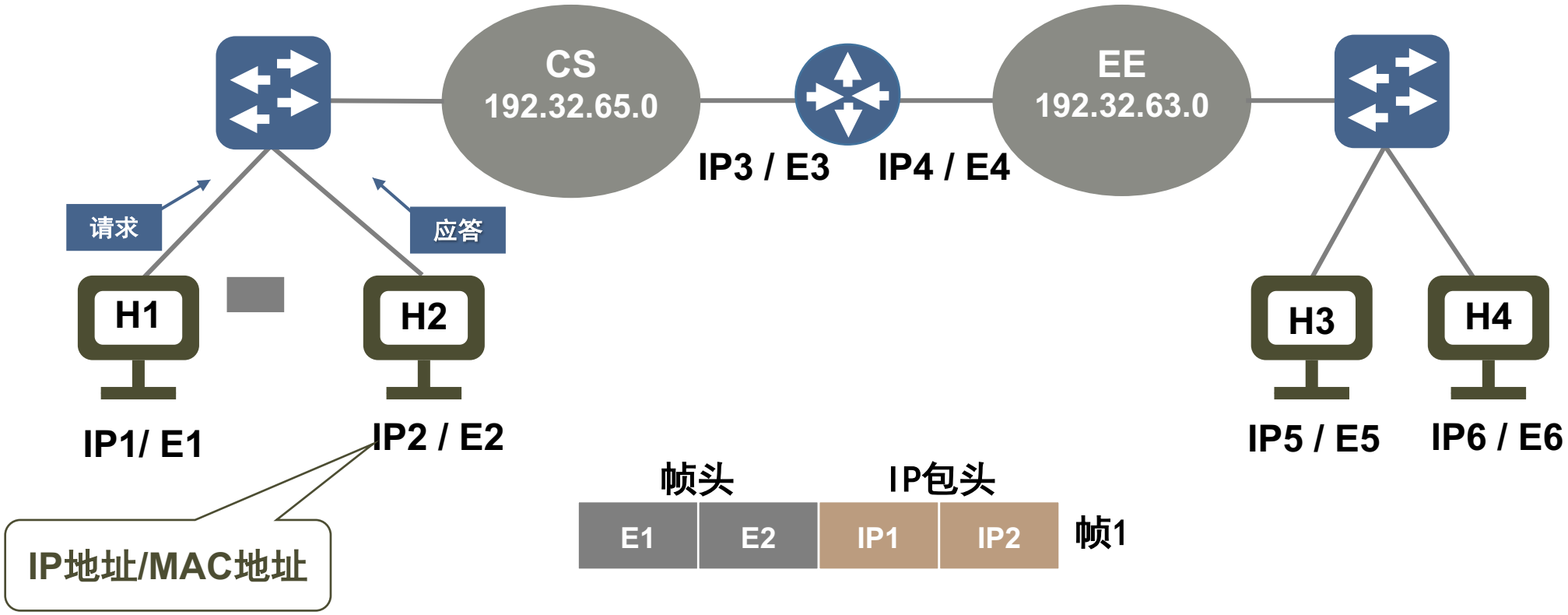
- 如何了解输入帧包含了ARP报文？
- ARP为何不封装在IP包中传递？



以太网/802.3帧头指明了有效载荷包括ARP/IP

ARP应用示例（教材图5-61）

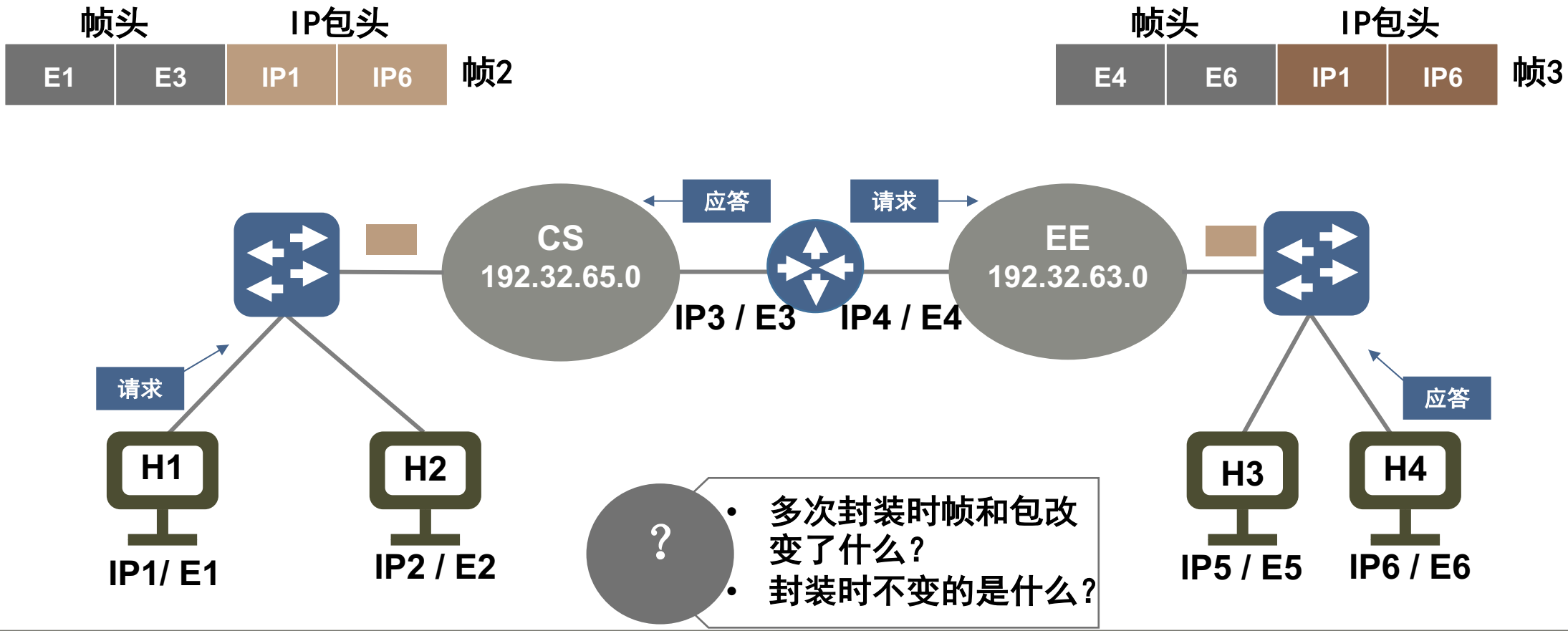
① 当主机1给主机2发送一个IP包



H1发给H2的包

ARP应用示例（教材图5-61）

② 当主机1给主机4发送一个IP包



H1发给H4的包

ARP的缓存技术及优化策略

ARP消息的处理

- 从接收到的消息中取出发送方的地址绑定信息
- 检查消息中的“操作”字段确定收到的是请求/应答消息

ARP的高速缓存

- ARP有一个高速缓存，用来存放最近获得的IP地址与硬件地址绑定信息

ARP的优化策略

- 在回答ARP请求后将请求消息中的发送方地址绑定信息加入自己的高速缓存

案例学习

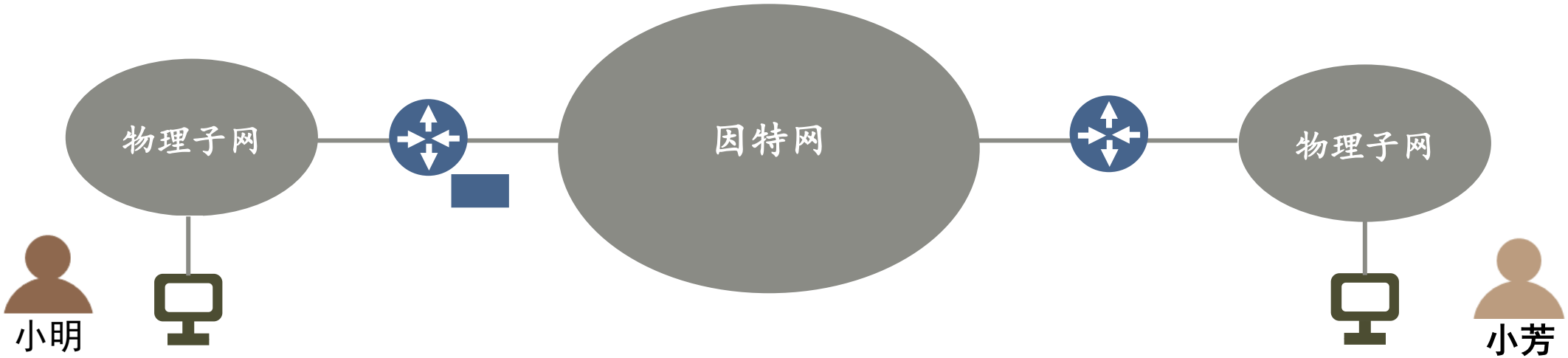
因特网IP地址分配协议

上网的前提是拥有全局唯一的ID

- 用户终端必须拥有一个唯一性IP地址
- 用户发送的IP包必须指明源端和目标端的IP地址
- IP协议为用户通信提供包传递服务

IP地址分配方式

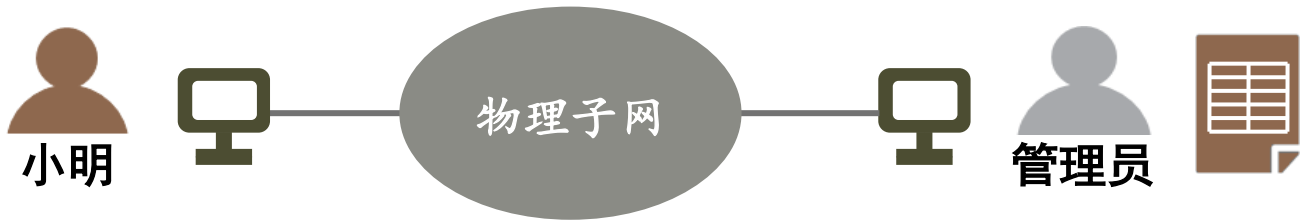
- 管理员手工分配
- 网络自动分配



网络管理员手工配置IP地址

优点	缺点	特点
••简单、易于实现 ••网络无需为分配地址付出额外工作	••不能适应网络动态变化（增加/移除节点都要更新地址列表） ••地址固定不变使得网络容量固定	••管理员拥有一张可用IP地址列表 ••每次增加一个新节点管理员从表中取出一个地址分配给该新节点 ••每次移除一个节点管理员将分配给该节点的地址收回加入地址列表

- 小明在上网前通过其他渠道向管理员申请一个IP地址
- 小明自己配置上网设备的IP地址



- 网络管理员维护一张IP地址分配表

网络协议自动配置IP地址

特点	优点	缺点
••地址分配协议管理一张IP地址列表 ••每当新增加一个节点协议自动从列表中取出一个分给该新节点 ••每当删除一个节点协议自动收回分配给该节点的地址	••适应网络的动态变化 ••能提高地址的利用率 ••网络容量高于固定分配方法	••增加一个协议不仅增加网络节点的工作负荷而且需要额外的网络带宽 ••协议软件要时刻跟踪IP地址使用情况

- 小明上网设备开机后自动向服务器申请一个IP地址
- 小明上网设备自动配置IP地址

- 服务器时刻准备为客户机分配IP地址
- 服务器跟踪客户机状态，必要时收回IP地址

动态主机配置协议（DHCP）

DHCP协议使得用户设备快速并动态地获取IP地址。用户新设备一旦连到网络，立即与地址分配服务器联系并申请一个可用IP地址。

RFC2131
RFC2132

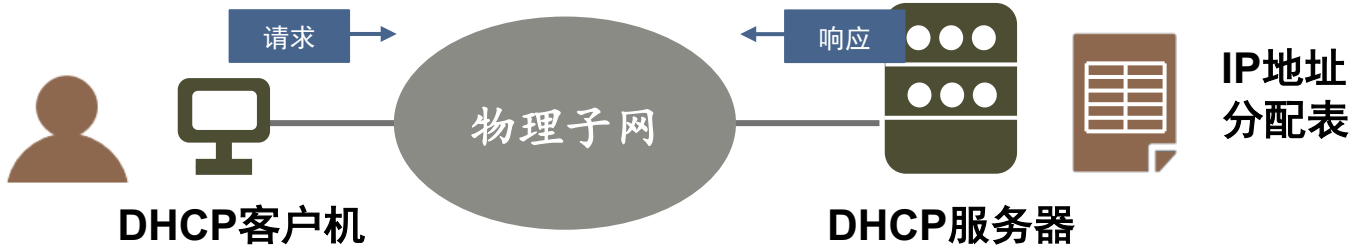
DHCP的临时性

- 服务器将一个地址在有限时间内分配给一个客户机（租赁）
- 服务器在地址分配时指定了租用期

DHCP服务器从管理员指定的地址中选择一个未分配的地址，并将它分配给该计算机。

IP地址分配表

IP地址	客户机MAC地址



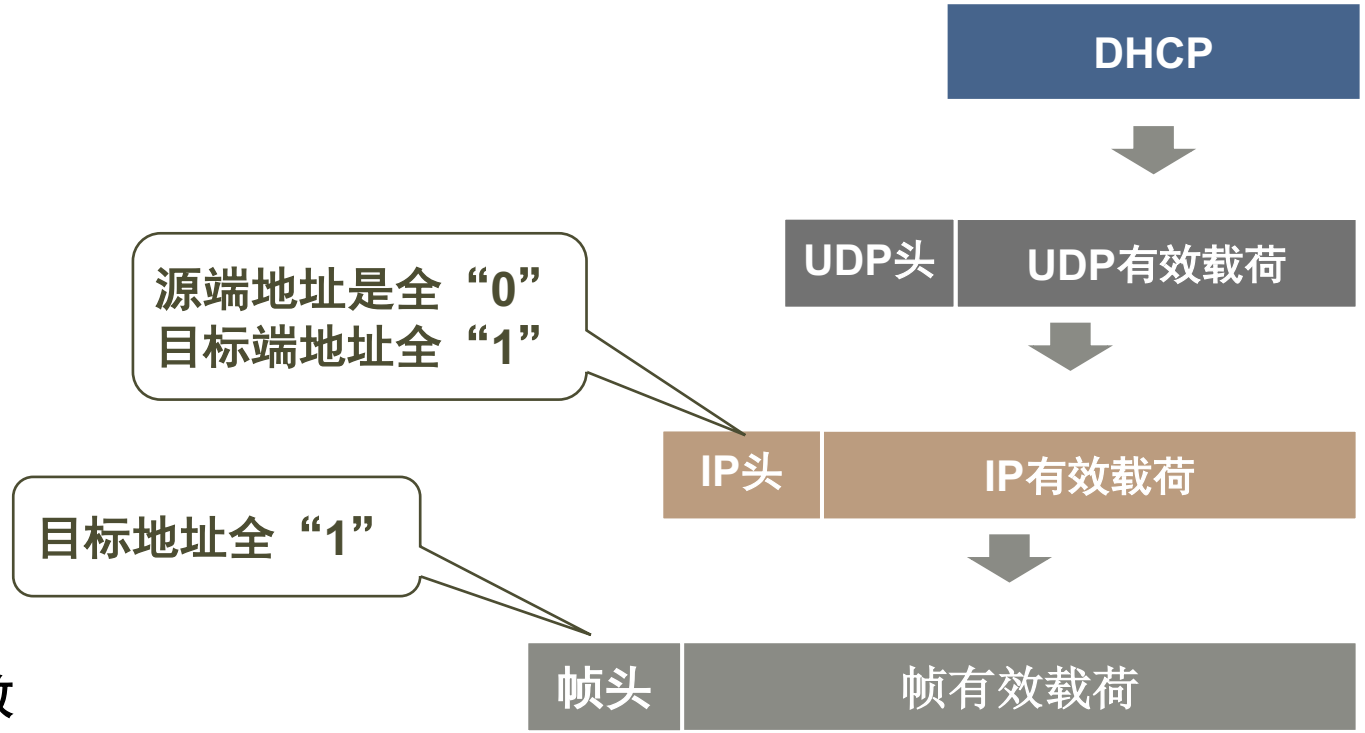
DHCP报文类型

DHCPDISCOVER	••客户机用来发现一个DHCP服务器
DHCPOFFER	••服务器针对DHCPDISCOVER的响应，提供了IP地址和其他参数
DHCPREQUEST	••客户机用来请求某个服务器提供IP地址或请求服务器续租IP地址
DHCPDECLINE	••客户机向服务器报告提供的IP地址已经被占用
DHCPACK	••服务器对DHCPREQUEST的响应，确认客户机请求的IP地址可使用。
DHCPNAK	••服务器指出客户机的IP地址租期已到，或者请求续租的IP地址已经分配给其他用户。
DHCPRELEASE	••客户机向服务器说明不再租用IP地址
DHCPINFORM	••客户机向服务器请求本地配置的参数

DHCP报文投递

DHCP使用UDP协议传递报文：

- 服务器使用UDP的67端口号
- 客户机使用UDP的68端口号



注：UDP协议实体根据端口号把UDP有效载荷部分的数据交给DHCP进程。

DHCP工作过程

DHCP工作过程

- 客户机请求IP地址
- 客户机续租IP地址
- 客户机释放IP地址

IP地址分配表

DHCPDISCOVER

DHCPREQUEST

DHCPRELEASE

DHCPINFORM

DHCPDECLINE

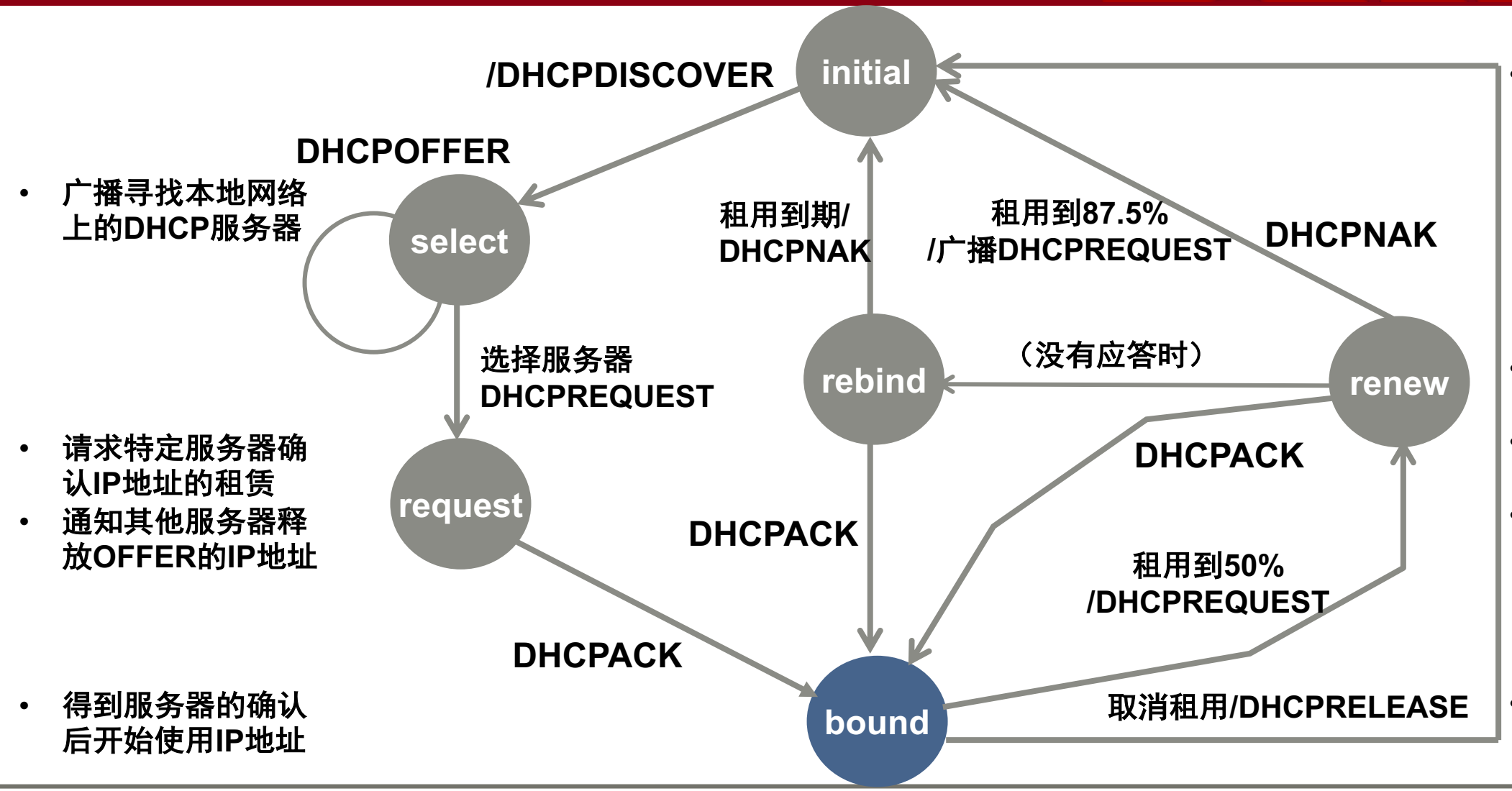
DHCPOFFER

DHCPACK

DHCPNAK

客户机初始时不知道DHCP服务器地址，用广播方式发送请求报文，服务器亦以广播方式回复响应报文。

DHCP状态机



- 广播寻找本地网络上的DHCP服务器

- 请求特定服务器确认IP地址的租赁
- 通知其他服务器释放OFFER的IP地址

- 得到服务器的确认后开始使用IP地址

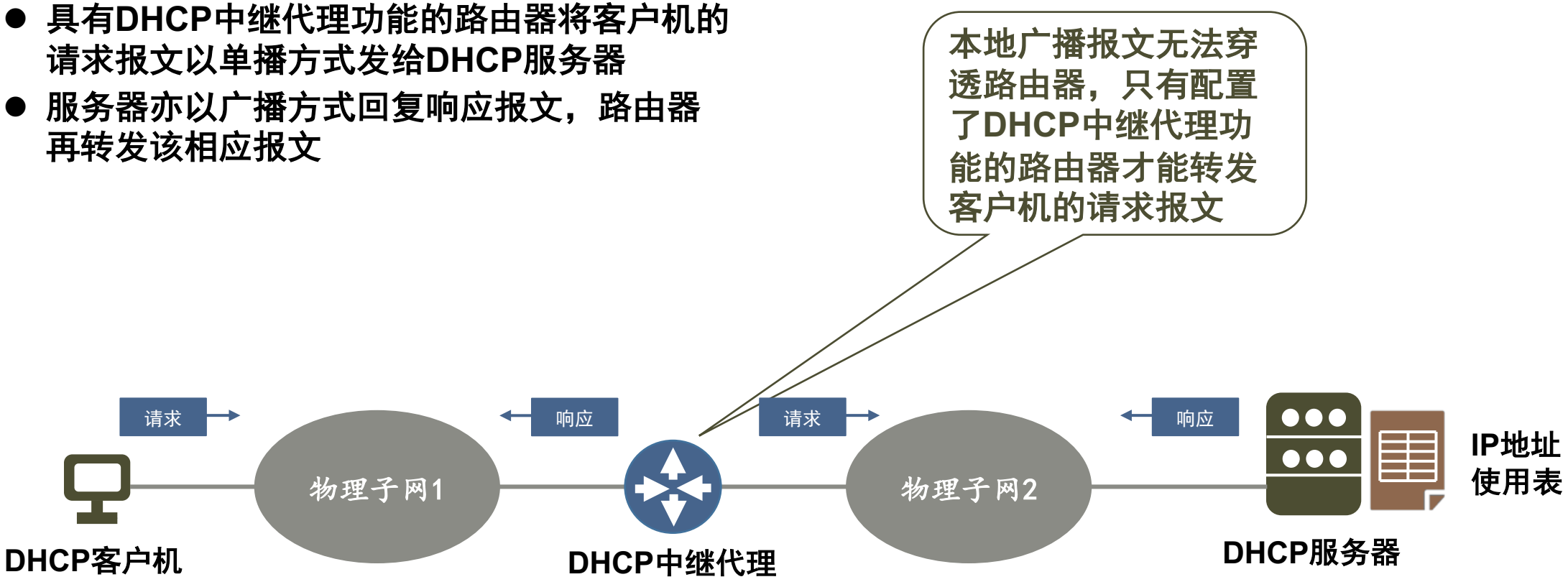
- 当使用租赁时间达50%和87.5%后续租均未果的情况下，地址的租赁期一道立即收回。

- 累计使用租赁时间50%后要续租
- 获得服务器同意后继续使用
- 服务器不同意则重新申请IP地址

- 提前释放IP地址

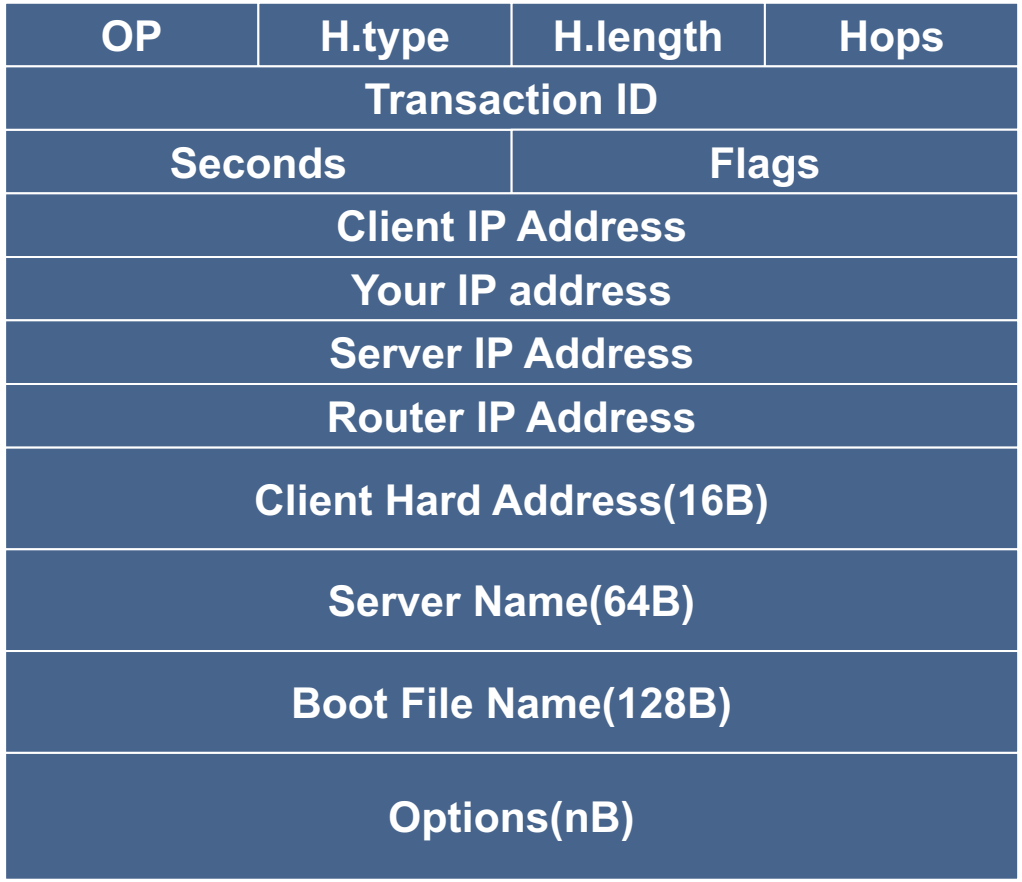
DHCP中继

- 客户机初始时不知道DHCP服务器地址，用广播方式发送请求报文
- 具有DHCP中继代理功能的路由器将客户机的请求报文以单播方式发给DHCP服务器
- 服务器亦以广播方式回复响应报文，路由器再转发该相应报文



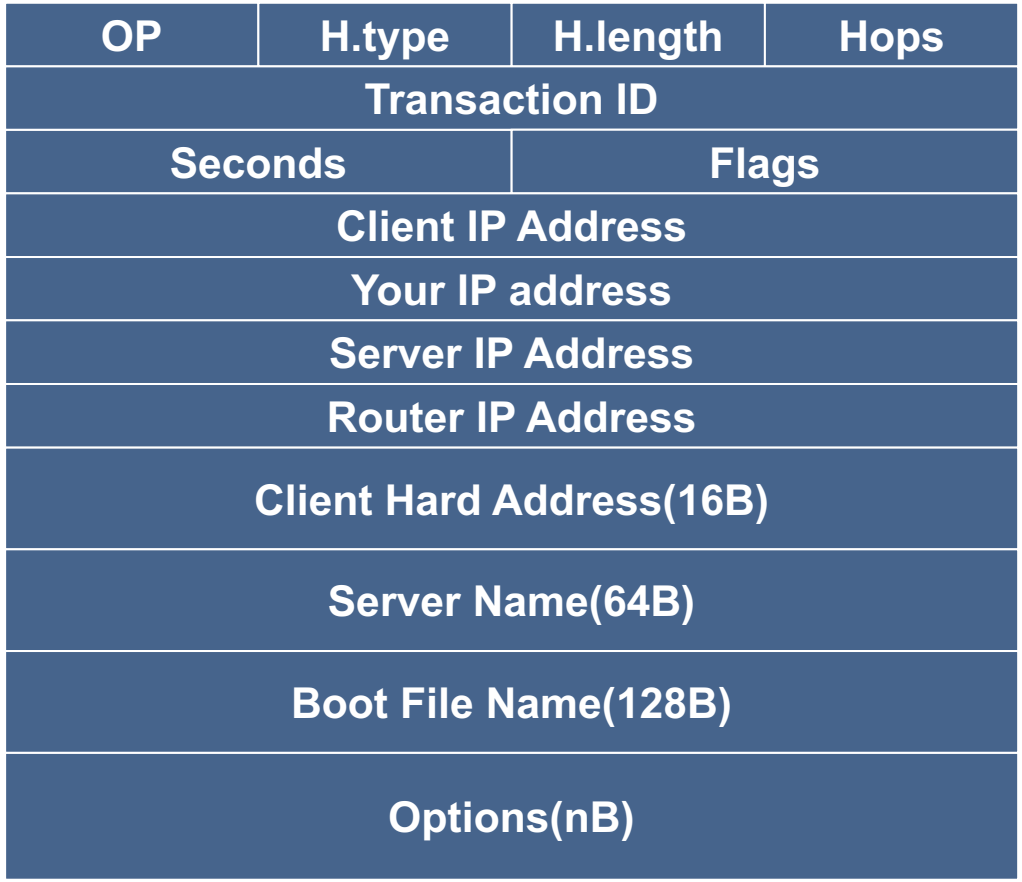
DHCP协议之报文格式

DHCP报文格式



- OP操作码：1表示请求报文，2表示响应报文
- H.type硬件类型：说明物理网络的类型
- H.length硬件地址长度：定义了物理地址的长度（以太网为6）
- Hops跳计数：用于记录报文的转发次数
- TransactionID事务标识：客户机随机生成，用来匹配客户机/服务器的交互
- Seconds秒计数：客户机给出租赁的IP地址已经过多长时间
- Flags标志：高位设置成1其余位设置成0表示客户机要求服务器以广播地址返回响应报文

DHCP报文格式(续)



- 客户机IP地址：未分配IP地址时该字段为0，标志高位设置1；已获得IP地址时填入，并标志高位设置为0
- 您的IP地址：服务器提供给客户机的IP地址
- 服务器IP地址：服务器在DHCPOFFER和DHCPACK中提供的服务器地址
- 路由器IP地址：DHCP中继代理的IP地址
- 客户机硬件地址：客户机的硬件地址
- 服务器名字：可选字段，包括服务器的域名
- 引导文件名字：服务器给出引导文件的全路径名
- 选项：定义了客户机和服务器交互的DHCP报文类型。

DHCP报文选项

选项(≤320B)

Code (53)	Length(1)	Type(1-8)

1	DHCPDISCOVER
2	DHCPOFFER
3	DHCPREQUEST
4	DHCPDECLINE
5	DHCPACK
6	DHCPNAK
7	DHCPRELEASE
8	DHCPINFORM

- 在客户机没获得IP地址之前，服务器通常用IP广播和硬件单播回复响应报文
- 在客户机获得IP地址后，服务器用IP单播回复相应报文

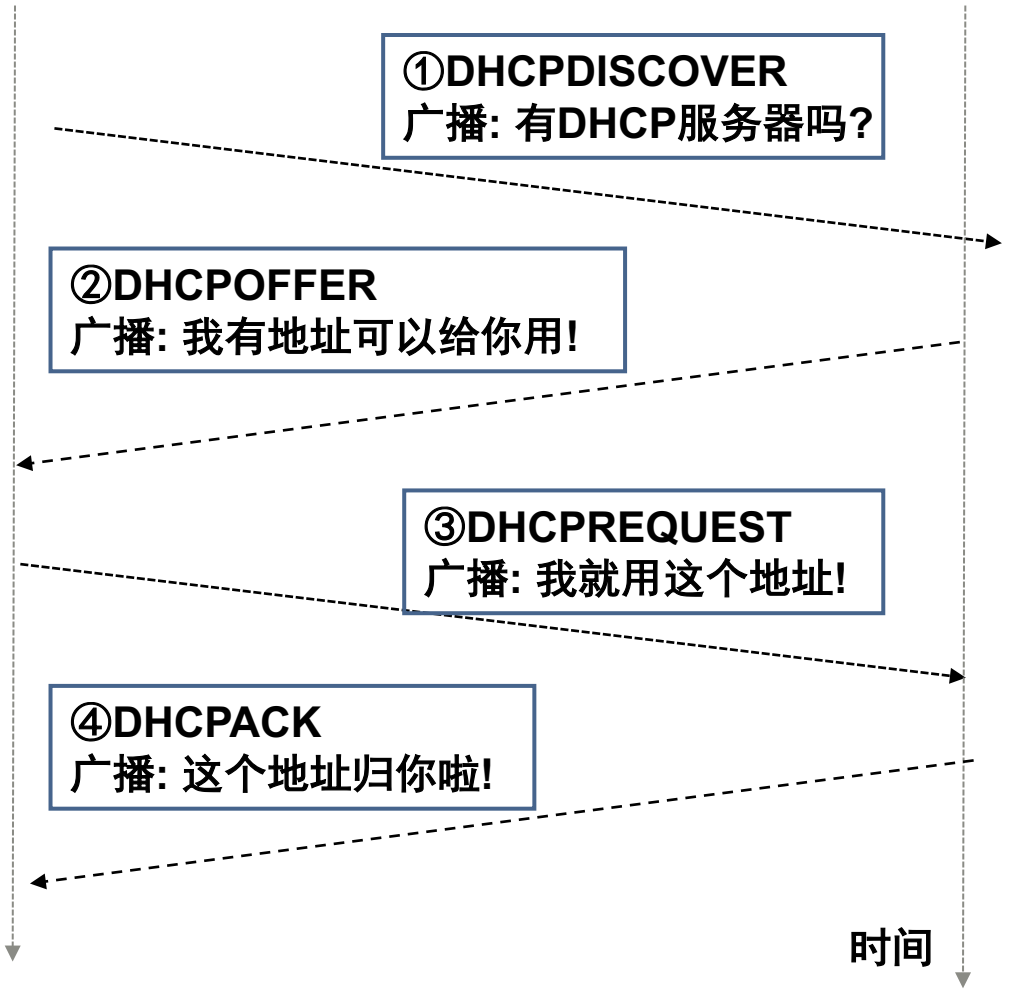
客户机请求IP地址

- 广播发送请求报文，寻找DHCP服务器

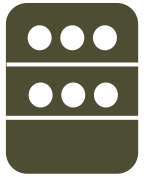
客户机 

- 从所有返回OFFER的服务器中选择一个，广播告之所有服务器将使用的IP地址

- 用ARP本地测试获得的IP地址是否可用，若已有节点使用则向服务器返回DHCPDECLINE报文；否则用该IP地址配置本地网络



- 所有服务器从可用IP地址表中取出一个未分配的IP地址响应

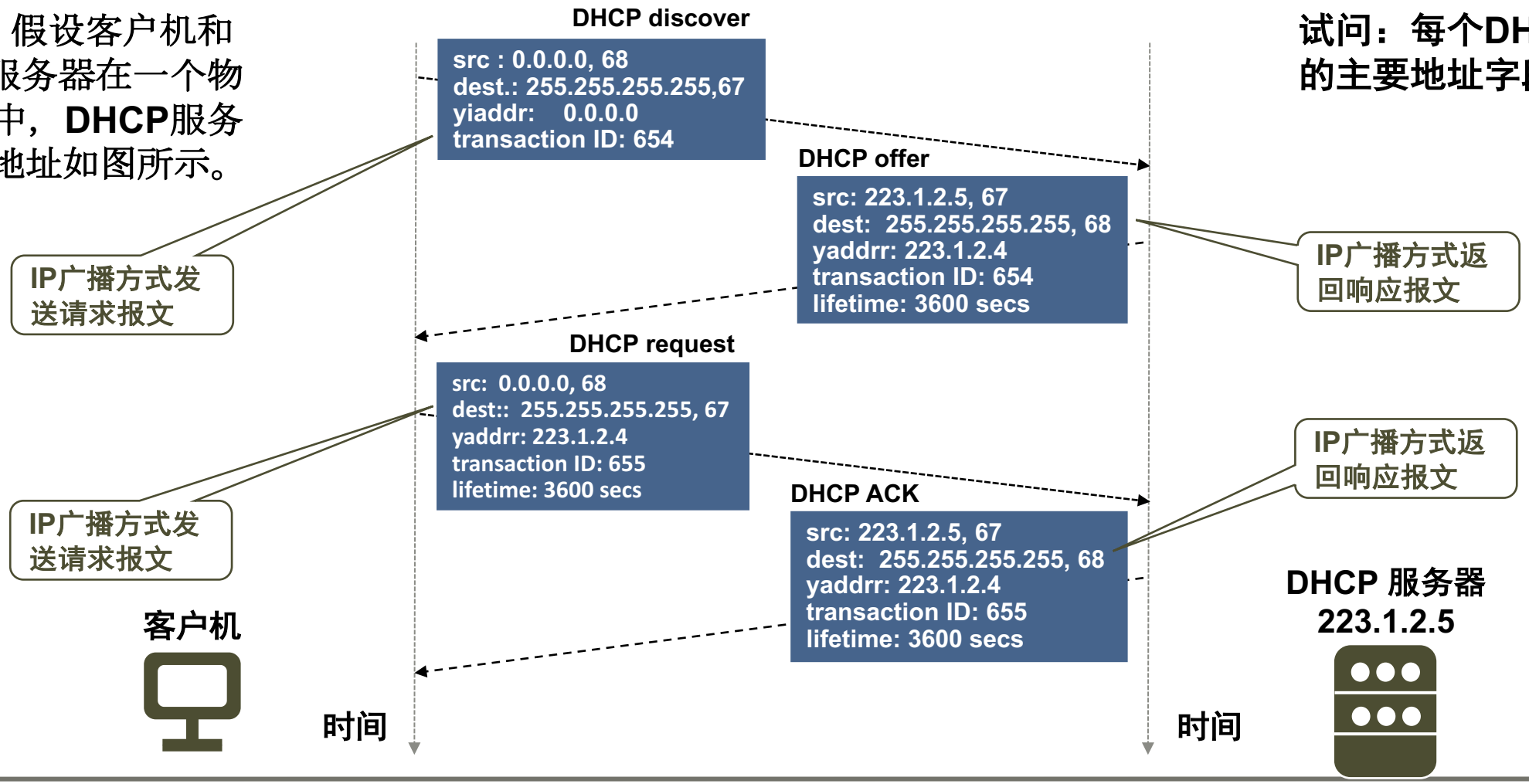
 DHCP 服务器

- 被选服务器确认分配的IP地址并提供其他参数，其余服务器将OFFER的IP地址放回可用IP地址列表中

- 收到DHCPDECLINE报文后将该IP地址设为不可用

客户机请求IP地址——报文示例

示例1：假设客户机和DHCP服务器在一个物理子网中，DHCP服务器的IP地址如图所示。



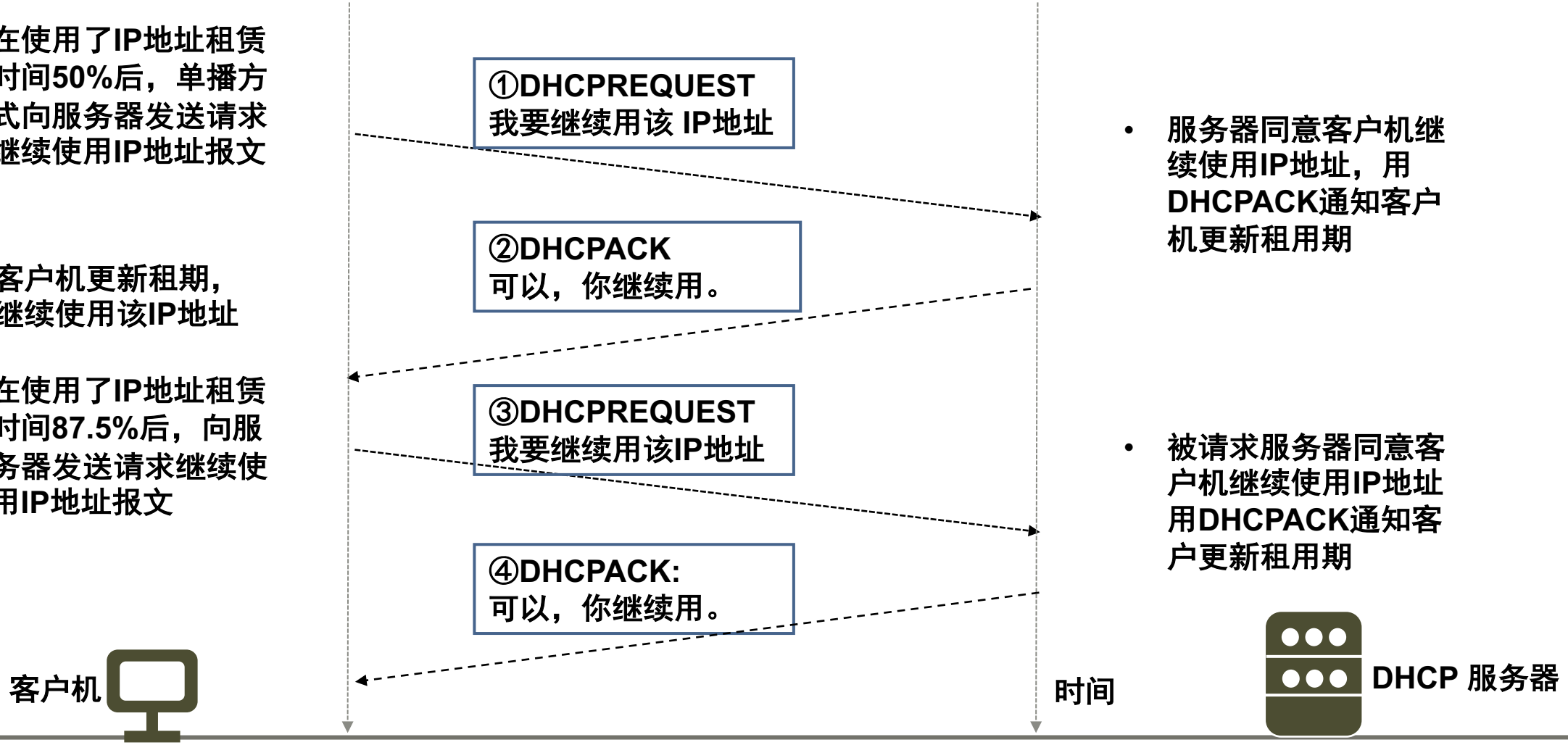
试问：每个DHCP报文的主要地址字段？

客户机续租IP地址

- 在使用了IP地址租赁时间50%后，单播方式向服务器发送请求继续使用IP地址报文

- 客户机更新租期，继续使用该IP地址

- 在使用了IP地址租赁时间87.5%后，向服务器发送请求继续使用IP地址报文

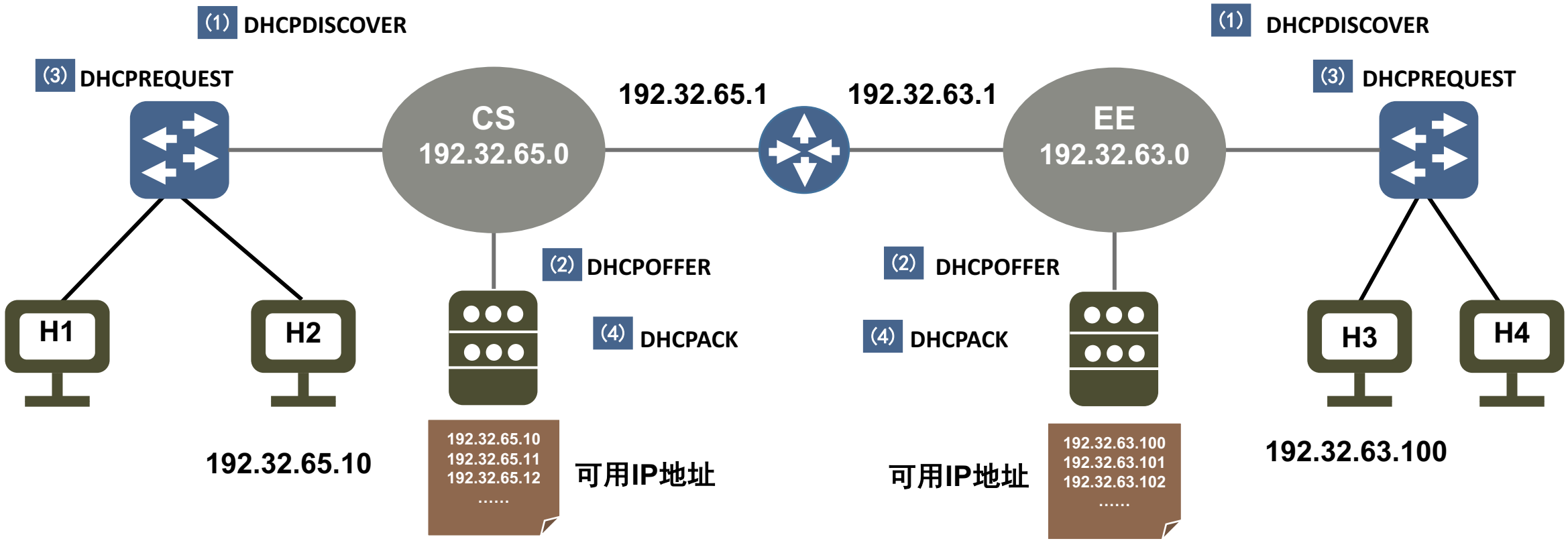


- 服务器同意客户机继续使用IP地址，用DHCPACK通知客户机更新租用期

- 被请求服务器同意客户机继续使用IP地址用DHCPACK通知客户更新租用期

DHCP工作过程示例

示例2：主机H2、H3刚加电，需要获得一个IP地址才能接入网络。



案例学习五

IP地址转换协议与网络扩展

内联网

内联网：不与因特网相连的企业内部网络。

？ 如何对内联网进行编址？

组建内联网好处

- 无需申请全球合法的IP地址
- 网络规模完全自主选择

私有地址

私有地址：不能用在因特网上的内部地址，
路由器将丢弃目标地址是这种地址的IP包。

RFC1918

- 10.0.0.0 ~ 10.255.255.255/8
- 172.16.0.0 ~ 172.31.255.255/12
- 192.168.0.0 ~ 192.168.255.255/16

特点

- 可以任意分配IP地址
- 所用的IP地址仅本地有效
- 所用的IP地址可被不同企业重复使用
- 节点不能与外部因特网上的节点通信

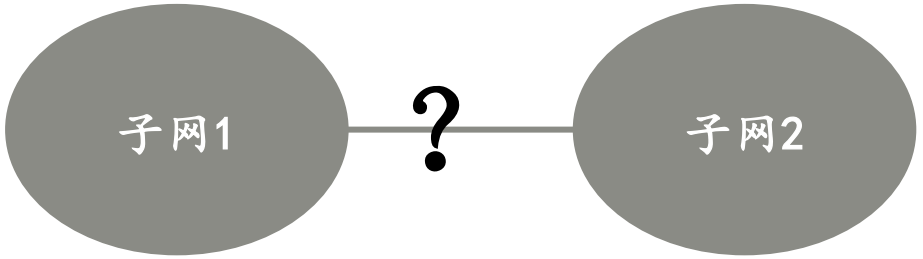
注意：这类地址不能出现在因特网上。

如何互联不同地点的内联网

当网络规模增大，尤其是分布在地理上分散的各地，需要一种方式将各个子网互联起来。

北京公司

如果两个内联子网相距不太远，可以考虑租用专线互联.



北京公司　　上海公司

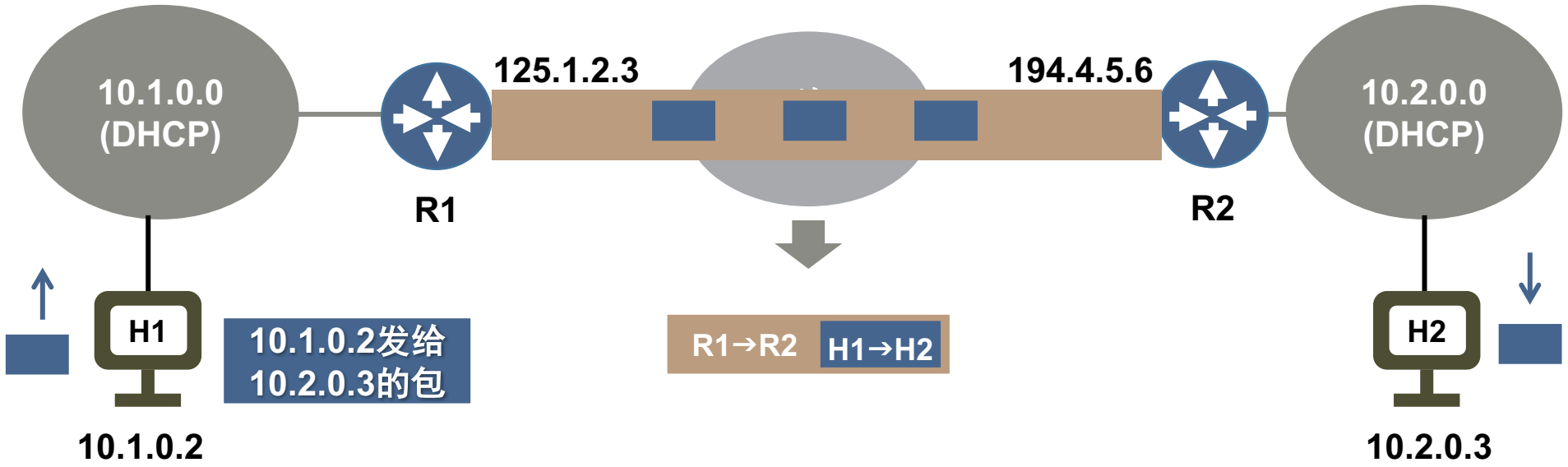
? 如果两个内网相距遥远，且租用专线成本高怎么办

虚拟专用网——内联网的互联

基于隧道的互联

- 前提：每个内联网必须拥有至少一个合法IP地址路由器
- 利用隧道技术将内联网包封装成因特网上的IP包

？ 内联网用户需要访问因特网如何处理

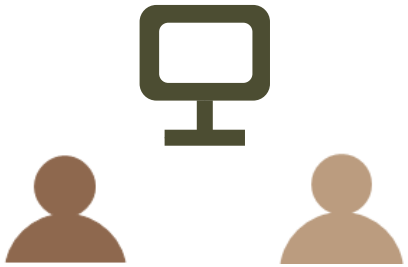


网络资源有限

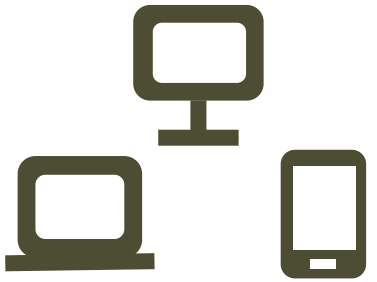
- 截至2017年6月，我国网民规模达到7.51亿，半年共计新增网民1992万人，半年增长率为2.7%。互联网普及率为54.3%。

- 截至2017年6月，我国IPv4地址数量达到3.38亿个、IPv6地址数量达到21283块/32地址，二者总量均居世界第二。

平均2人共用一个IP地址还不够



现状：许多用户有多个上网终端设备



?

- 如何保障每个用户都能上网
- 需要什么技术支撑

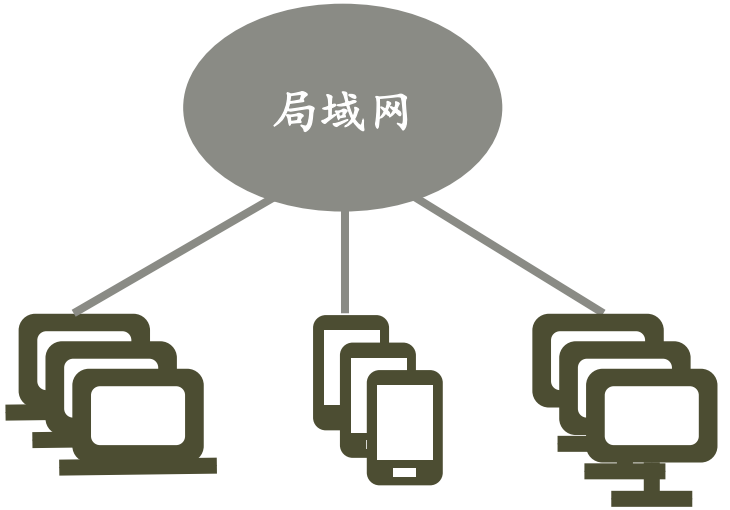
共享理念能否用于网络地址?

共享单车使得人们不必自己拥有一辆自行车就能骑车出行。

只要不是所有用户都需要一直在线，通过DHCP就可以做到一部分用户共享少量的IP地址。

假设：

- 有M个全局可路由的IP地址
- 有N个经常需要上网的设备（N>M）

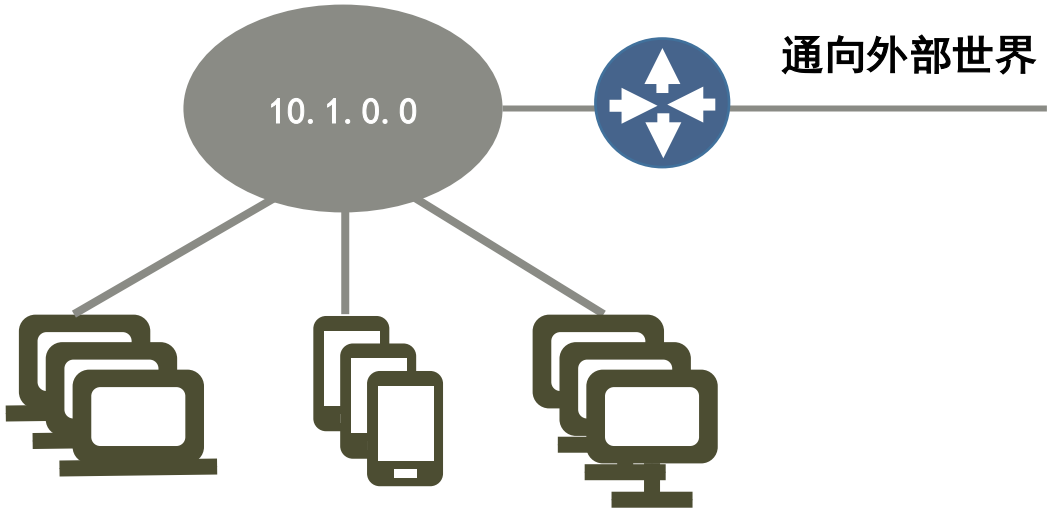


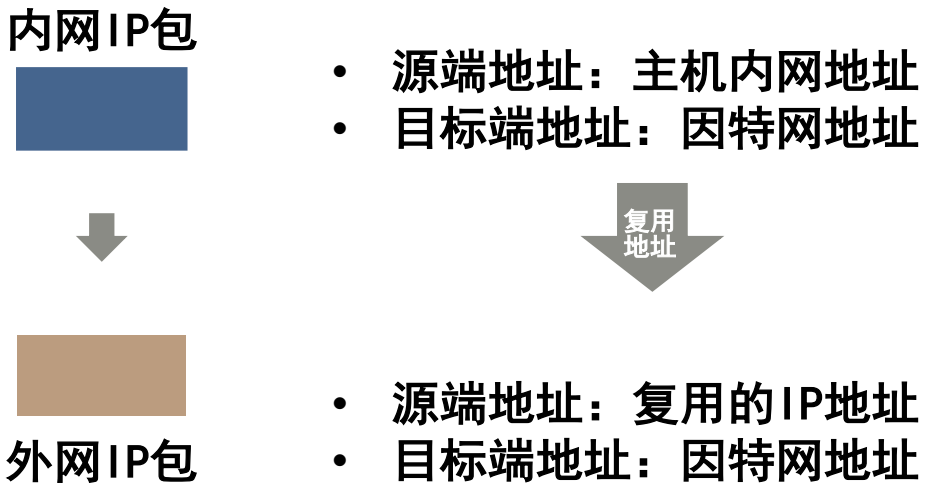
? 如果同时需要上网的用户数多于地址数怎么办

基于共享理念的地址复用技术

地址复用：在源端多个应用程序发送的IP包复用同一个IP地址，返回时将IP包准确分发给相应的应用程序。

- 每当本地产生一个目的地非本地的IP包，路由器就将该包的源地址替换成自己的可路由IP地址
- 接收返回的响应IP包，把包目的地址替换成原始的源地址后转发到本地网络





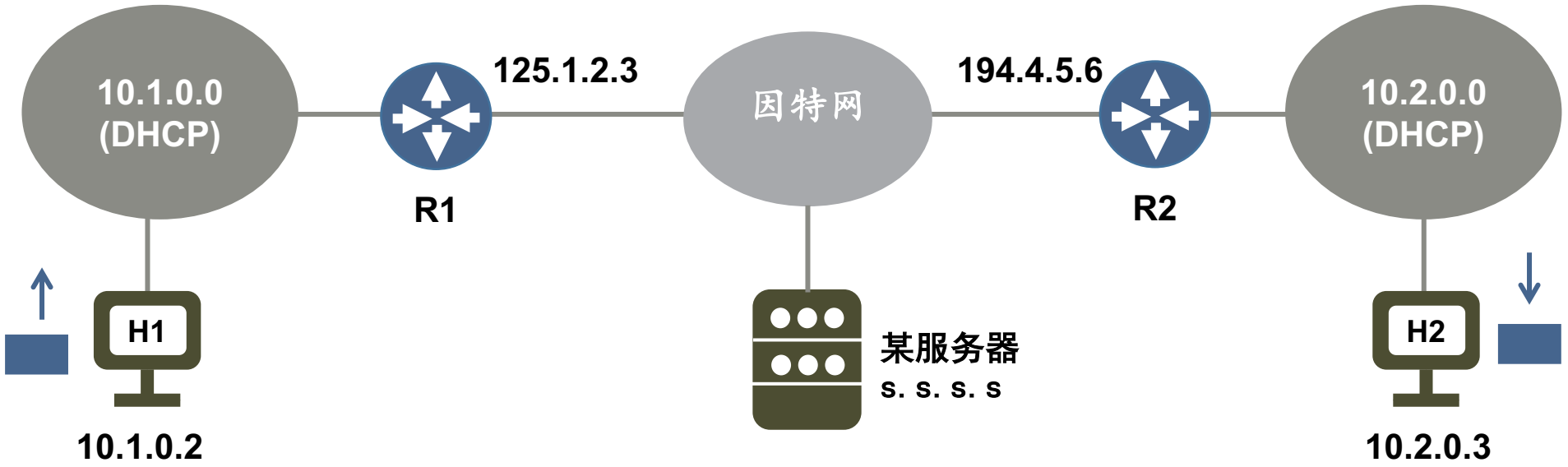
复用地址技术的应用

内联网：

- 内联网主机之间通过虚拟专用网通信
- 内联网主机通过复用地址技术访问因特网服务器

普通局域网：

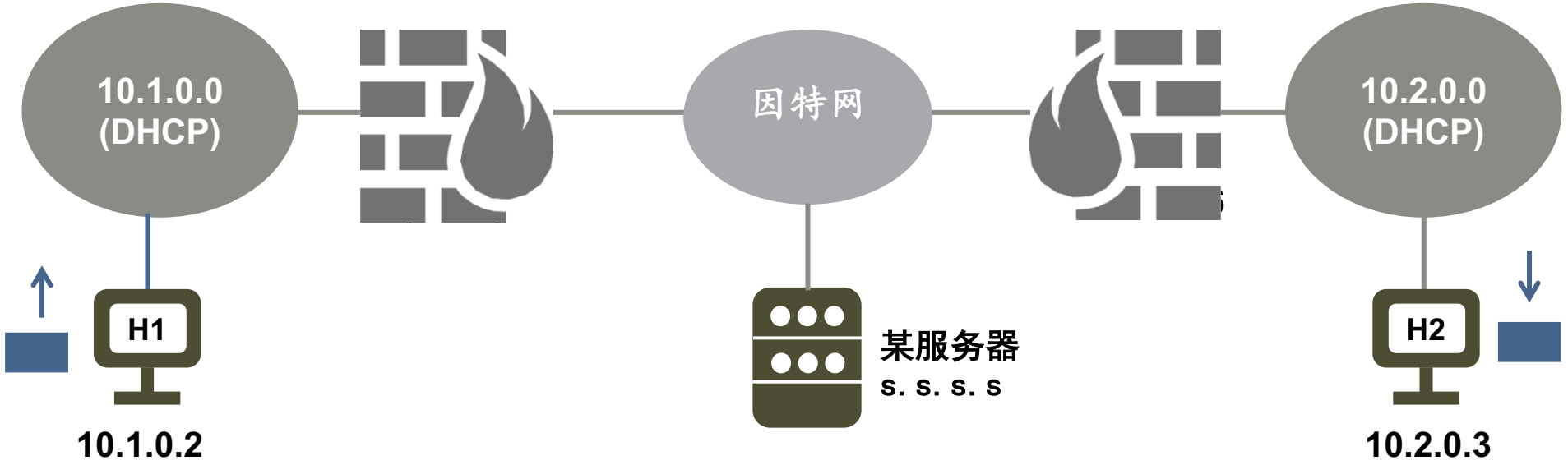
- 主机通过复用地址技术访问因特网
- 同时上网的主机个数不再受限于可路由IP地址数



复用地址增加内网/局域网的安全性

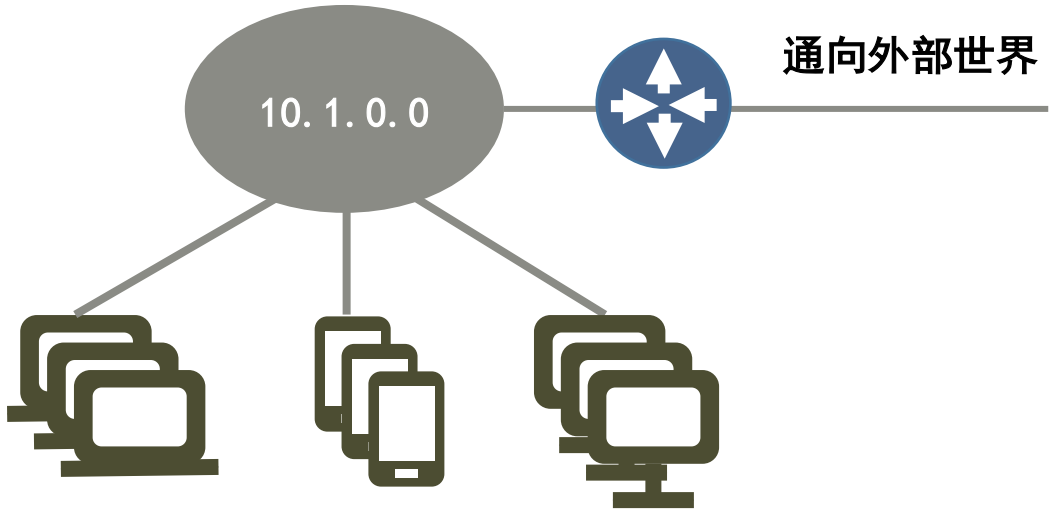
所有进出路由器的IP包都将内网的地址隐藏了起来（甚至对应于应用程序的端口号），使得常规的安全攻击因找不到准确的应用程序而失效。

？ 如何复用可路由的IP地址，将面临什么困难



复用地址面临的困难

目标：必须能够处理同时来自多个主机多个应用程序访问因特网的需求。



- H1的浏览器访问某个门户网站
- H1的微信正在进行语音通话
- H2的FTP正在下载一个文件
- ……

本质上是如何区分内网中不同的应用进程发送的IP包，并且这些应用进程可能来自相同或者不同的主机。

地址转换协议NAT

网络地址转换协议

网络地址转换协议（NAT）：在私有地址和全局可路由地址之间转换的协议。

RFC3022
RFC2993
RFC3235
RFC3027

引入NAT协议的动机

- 增加内部网络的安全性
- 内联网用户需要访问因特网
- DHCP只能部分缓解IP地址资源的不足
- 小型企业和家庭用户需要全程在线连接因特网

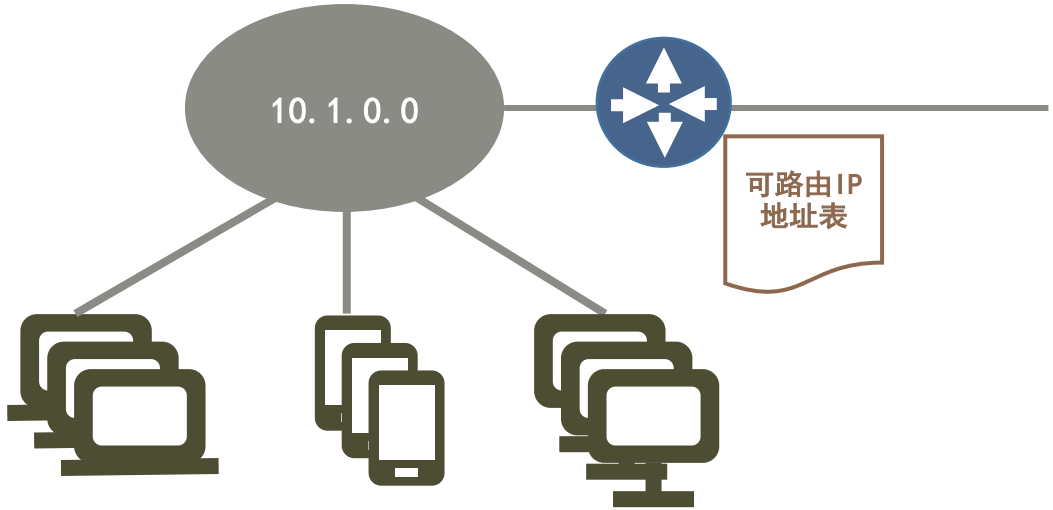
静态和动态NAT

动态NAT

- 动态配置NAT是建立内部本地地址和外部可路由地址的临时映射关系，并且映射关系具有一定的时效性。

静态NAT

- 路由器维护一张可路由IP地址分配表
- 路由器负责建立内网本地地址和外部可路由地址的一对一永久映射关系
- 当外部网络需要通过固定的全局可路由地址访问内部服务器时，静态NAT尤为重要



Router(config)#ip nat inside source static
local-ip global-ip
/*将内网地址映射成外网地址

Router #ip nat inside 10.1.0.2 125.1.2.3

基于端口的NAPT

网络端口地址转换（NAPT），将多个内部地址映射为一个合法的可路由地址，但以不同的协议端口号和内部地址对应于端口号和可路由地址。

<内部地址+内部端口>
vs.
<外部地址+外部端口>

路由器维护一张NAT转换表

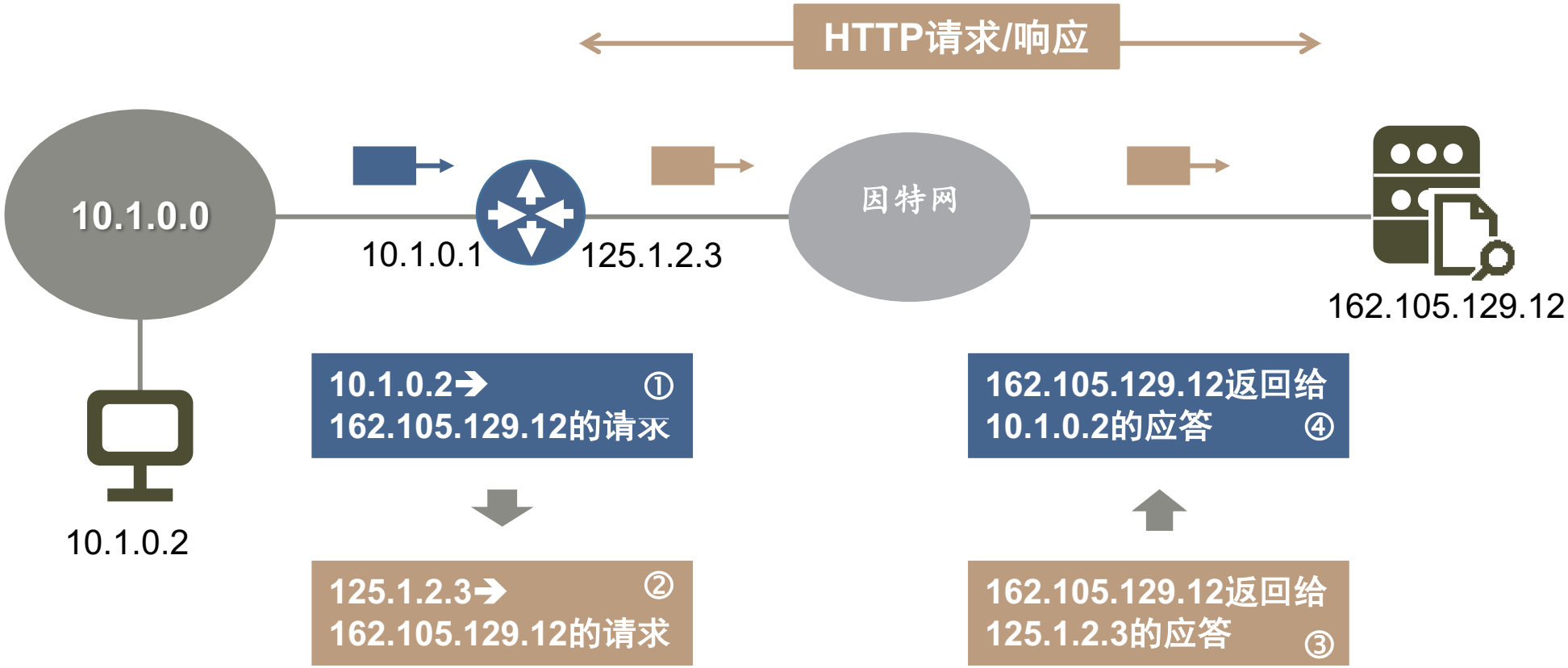
- 通过转换端口号以及地址来提供并发性
- 除了一对源和目的IP地址以外，还包括一对源和目的协议端口号，以及NAT使用的一个协议端口号。

NAT转换表 

NAPT特点	NAPT优势	NAPT缺点
••“多对一”的NAT ••可将中小型网络隐藏在一个合法IP地址后面	能够使用一个全球有效IP地址访问因特网。	仅限于TCP或UDP高层协议

NAPT协议——工作过程

假设：内网主机H访问外部网络web服务器



NAT路由器功能

NAT路由器：负责IP包的源IP地址和目的IP地址转换，地址转换可以静态或者动态设置。

客户机/服务器之间的端-端连接被拆分成两段：

- 客户机-路由器
- 路由器-服务器

针对出境包源地址进行替换

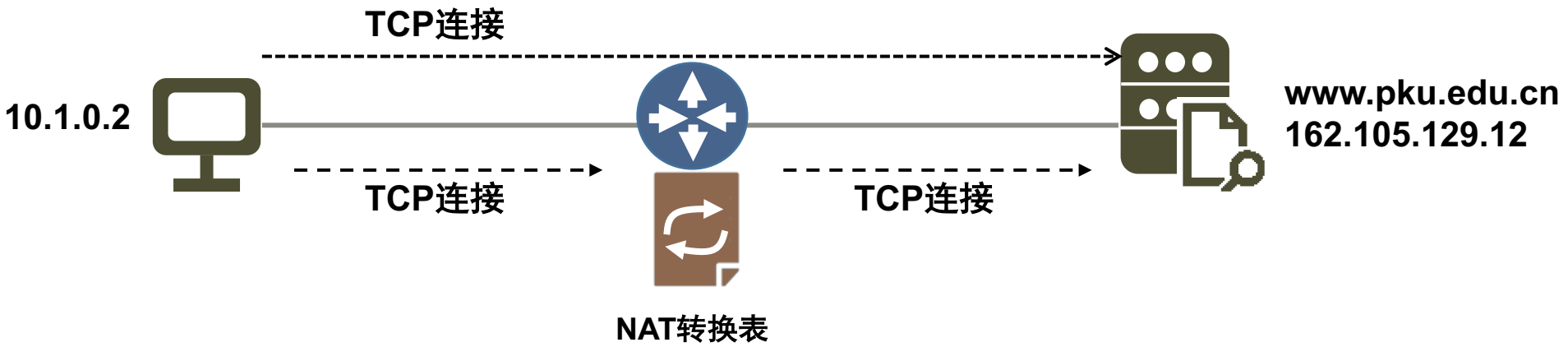
••(源IP地址, port #)→(路由器IP地址, 新port #)

在NAT转换表中记录映射关系

••(源IP地址, port #)→(路由器IP地址, 新port #)

针对入境包目标地址进行替换

••(路由器IP地址, 新port #) →(源IP地址, port #)



NAT实现——地址转换表

NAT转换表

源主机	源IP地址	源Port号	路由器IP地址	NAT指定Port号
H1	10.1.0.2	80	125.1.2.3	102
H2	10.1.0.3	3000	125.1.2.3	103
H3	10.1.0.4	4000	125.1.2.3	104

对于出境包

- 路由器给该包分配一个未用的port号，并用NAT路由器的IP地址和该port号替换包的源IP地址和源port号
- 在NAT地址转换表中添加一项，将源port号源IP地址映射成新分配的port号NAT路由器IP地址

对于入境包

- 路由器以目的port号作为索引查找转换表，以对应的源IP地址和port号置换回去
- 转换表中的有关条目动态在空闲超时后删除

NAT协议——支持/不支持的应用

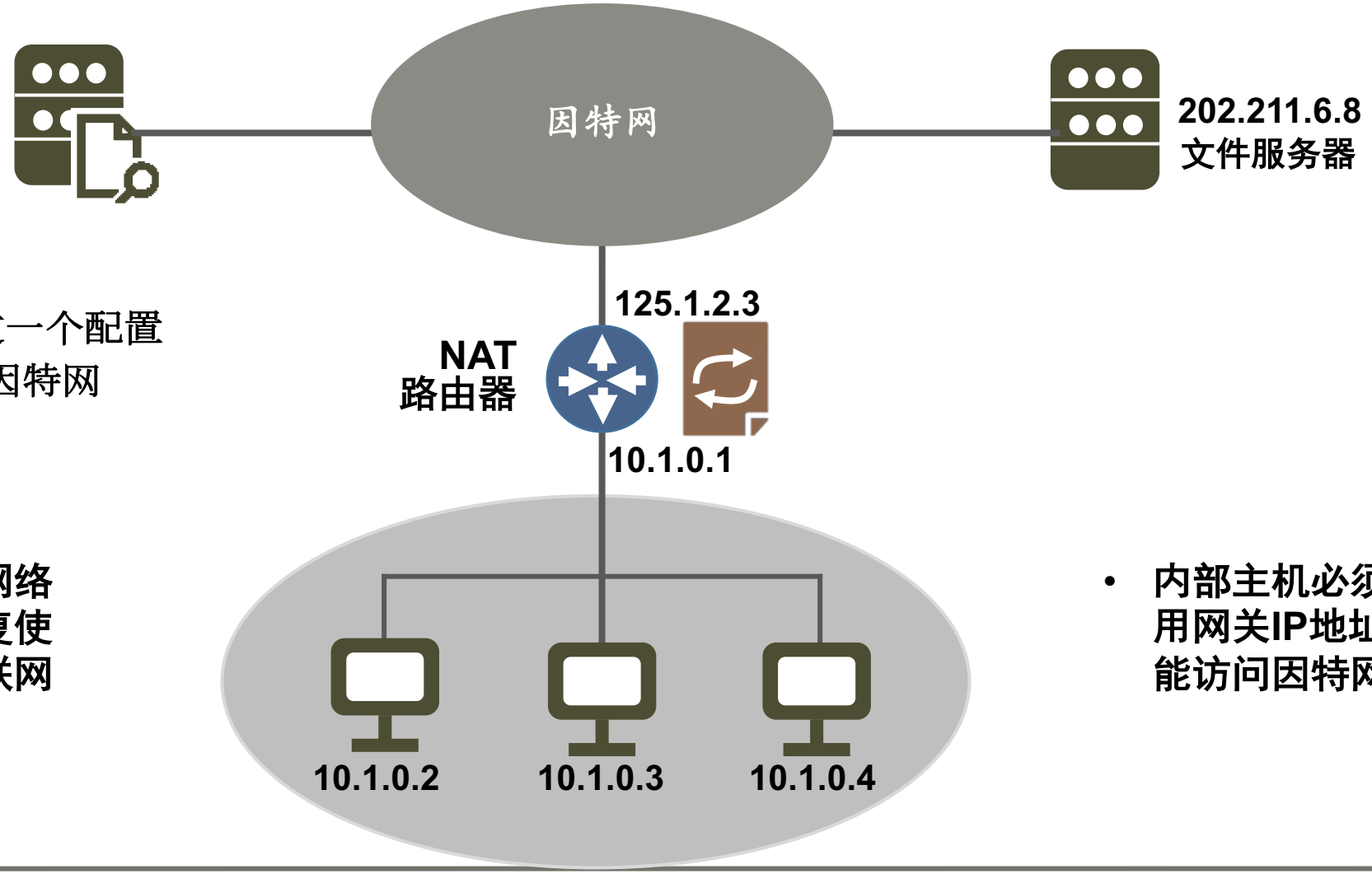
NAT支持的应用

- 数据部分不包含IP地址的TCP/UDP流
- 在数据部分包含IP地址的IP包
- ICMP、FTP
- NetBIOS over TCP
- RealAudio
- CUSeeMe (White Pines)
- Streamworks
- DNS “A” and “PTR” 查询
- H.323 (NetMeeting)
- VDOLive、Vxtreme

NAT不支持

- IP组播
- DNS Zone Transfers
- BOOTP
- Talk, ntalk
- SNMP
- NetShow

NAT典型应用场景——内联网



示例1：内部网络通过一个配置成NAT的路由器访问因特网

- 相同的私有网络地址可被重复使用于多个内联网
- 内部主机必须使用网关IP地址才能访问因特网

NAT协议地址转换示例

内网上的请求包

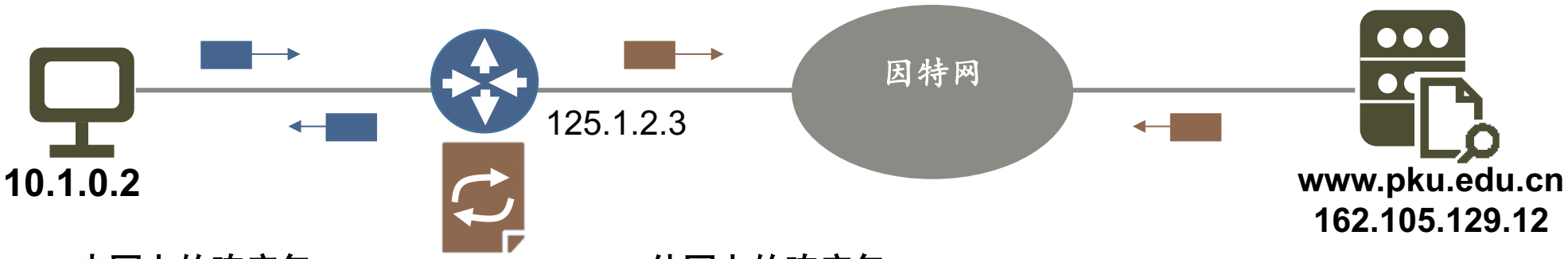
Proto	TCP
Saddr	10.1.0.2
Daddr	162.105.129.12
Sport	1049
Dport	80

外网上的请求包

Proto	TCP
Saddr	125.1.2.3
Daddr	162.105.129.12
Sport	102
Dport	80

示例2：一个内网主机作为客户机访问外网的一个门户网站

试问：IP包在内外网上的地址转换情况？



内网上的响应包

Proto	TCP
Saddr	162.105.129.12
Daddr	10.1.0.2
Sport	80
Dport	1049

外网上的响应包

Proto	TCP
Saddr	162.105.203.12
Daddr	125.1.2.3
Sport	80
Dport	102

NAT协议的不足

内网上的请求包

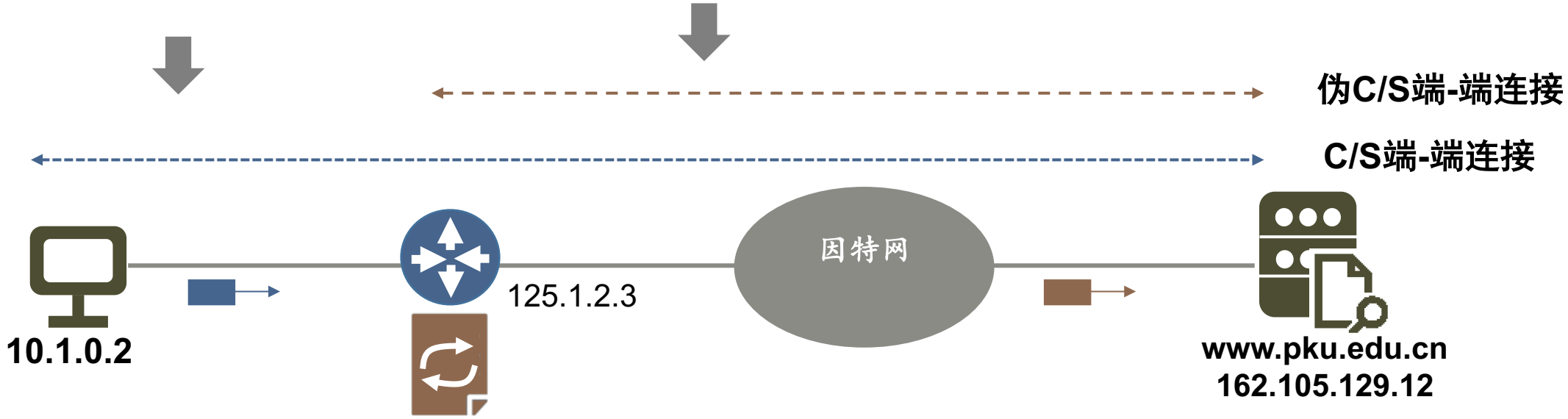
Proto	TCP
Saddr	10.1.0.2
Daddr	162.105.129.12
Sport	1049
Dport	80

外网上的请求包

Proto	TCP
Saddr	125.1.2.3
Daddr	162.105.129.12
Sport	102
Dport	80

NAT协议不足

- 破坏了端-端连接语义
- 没有根本解决地址资源不足



案例学习六

下一代因特网互联协议

IPv4的成功与不足

使用IP的协议
软件自行解决

IPv4的功劳

- 异构性
  - ① 不同系统
  - ② 不同子网
- 扩展性
  - 因特网的规模就是证明

IPv4的不足

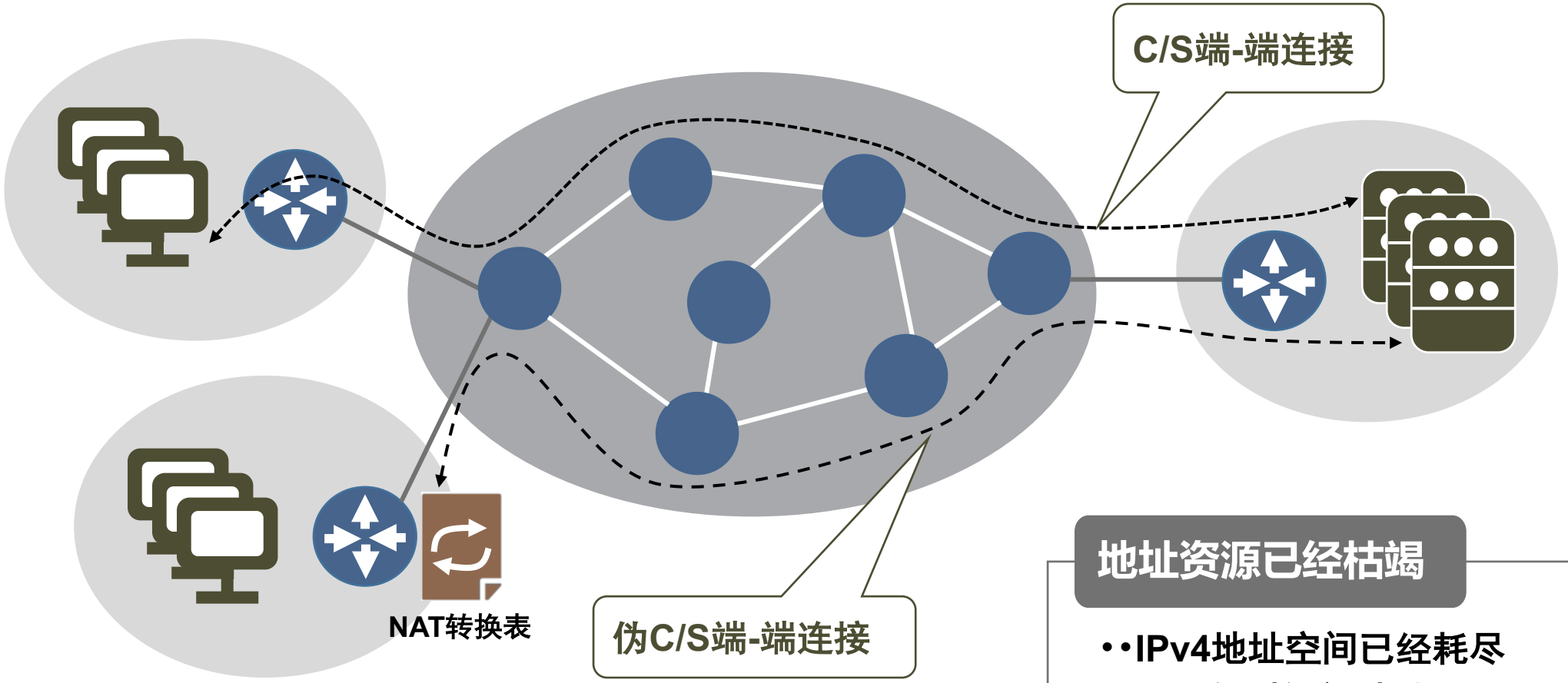
尽力而为的数据包传递服务

- 数据包受损
- 数据包重复
- 延迟或乱序
- 数据包丢失

变革动机

- 有限的地址空间
- 新网络应用对QoS需求
- 复杂寻址和路由能力的需求

为什么要下一代IP协议



地址资源已经枯竭

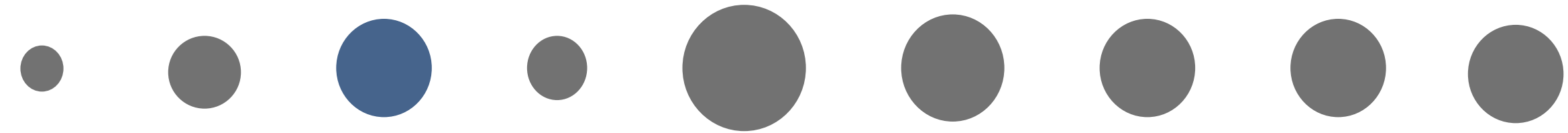
- IPv4地址空间已经耗尽
- NAT暂时缓解地址不足

IPv6具有足够多的地址空间

足够的IPv6地址

- $2^{128} \cong 3.4 \times 10^{38}$
- 可为地球上每平方米提供6×10^{23}个网络地址
- 可为半径为25光年的球体每平方厘米提供一个地址

- 我们最近邻居4.2光年远
- 在25光年距离内30多颗星
- 冥王星轨道直径11.9×10^{14}cm

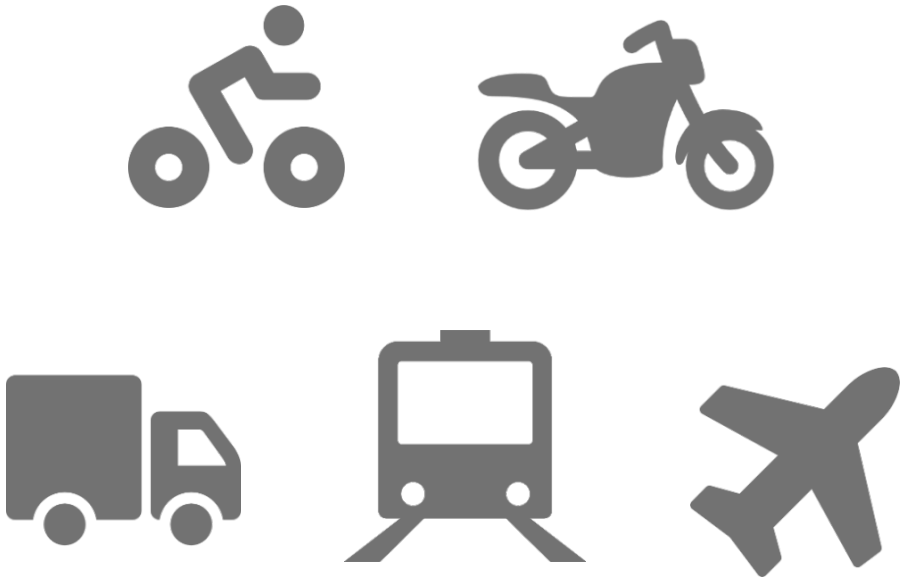


水星 金星 地球 火星 木星 土星 天王星 海王星 冥王星

1 光年 = 9.453×10^{17} cm

IPv6灵活并且安全

●简化的报头和灵活的扩展



●内置的端-端安全认证和加密

- 按需组合不同的IPv6包控制信息头
- 信息加密和用户认证变得必须

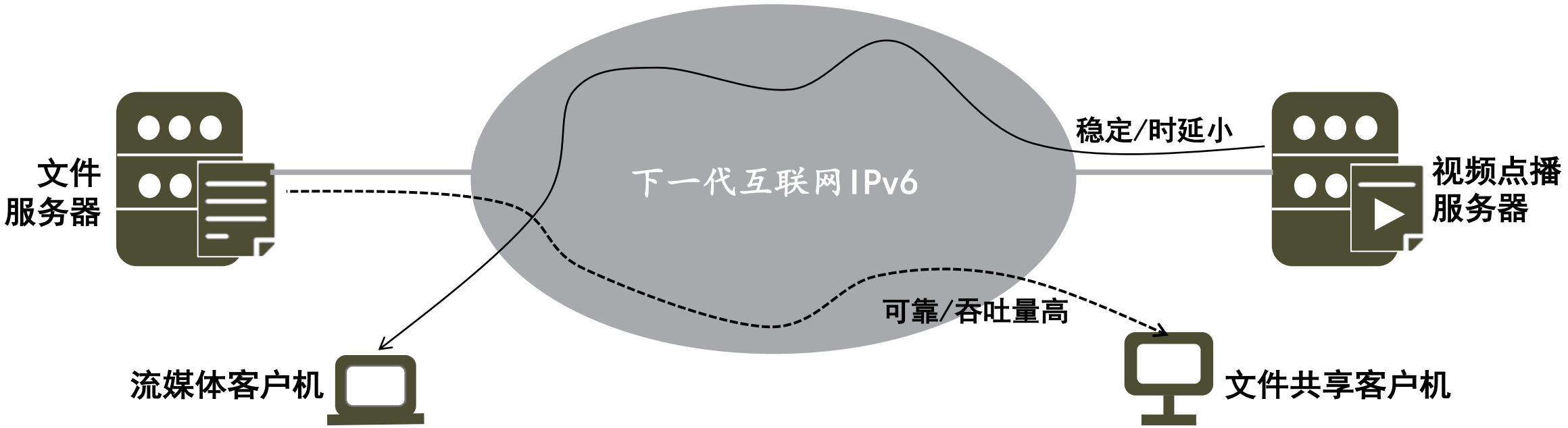
IPv6支持即插即用且保证QoS

支持即插即用

- 本机地址自动配置
- 上网信息自动获取

保障服务质量

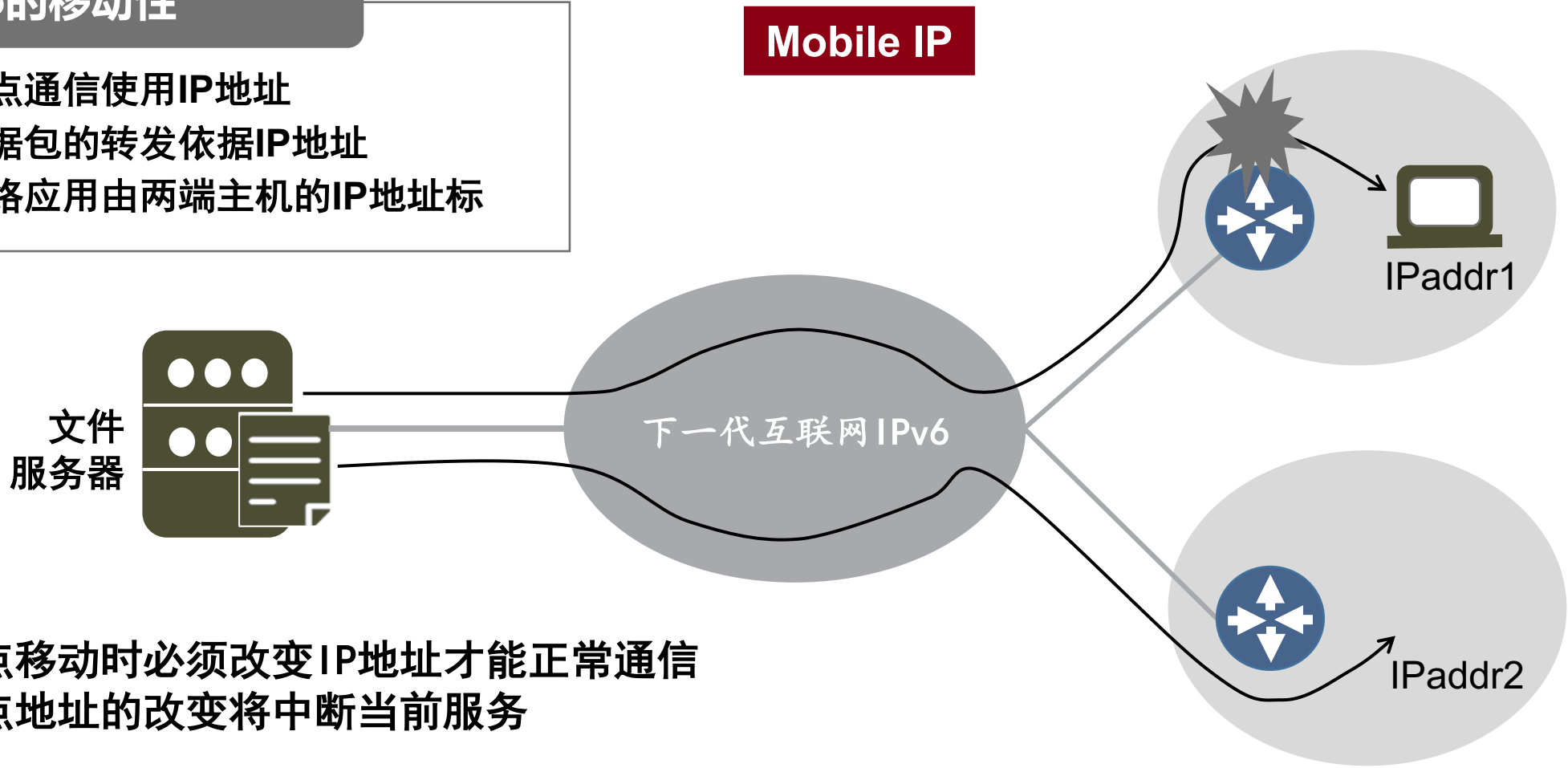
- 提供不同服务质量的服务



IPv6对移动计算的支持

IPv6的移动性

- ··节点通信使用IP地址
- ··数据包的转发依据IP地址
- ··网络应用由两端主机的IP地址标示



- 节点移动时必须改变IP地址才能正常通信
- 节点地址的改变将中断当前服务

IPv6的功能和通信模式

IPv6协议：下一代因特网的互联网络协议。为上层用户提供了不可靠的数据包传递服务，并具有一定的服务质量保障。

单播 (unicast)	组播(multicast)	选播 (anycast)
• “一对一”的通信模式。 • 一个节点给因特网中任何一个节点发送IPv6包。	• “一对多”的通信模式。 • 一个节点可以给一组具有共同特性的节点发送IPv6包	• “一对特定组中一个”的通信模式。 • 一个节点可以给某个组中任意一个节点发送IPv6包

IPv6协议的报文格式

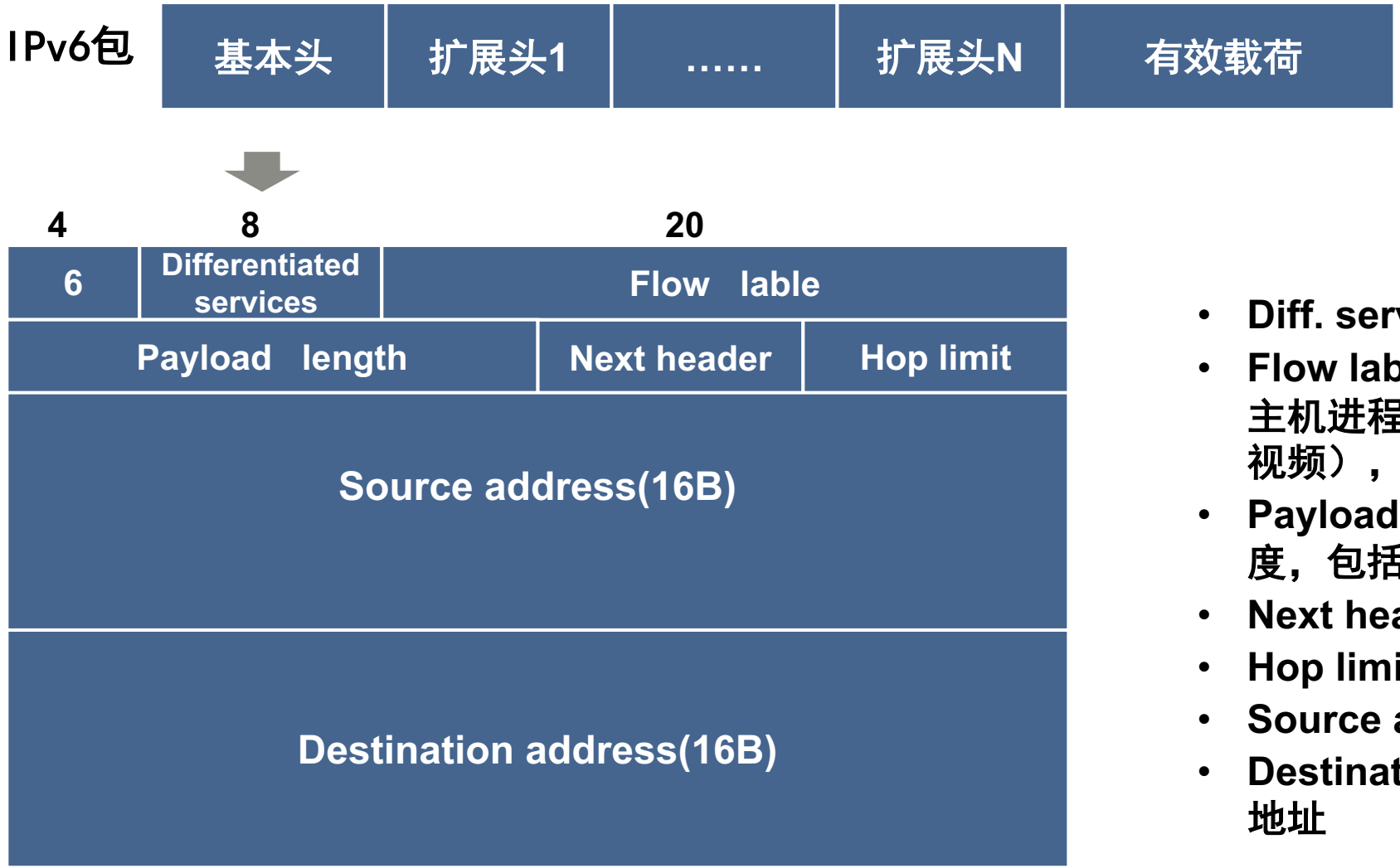
IPv6设计目标

设计目标

- 支持上百亿台主机
- 减小路由表的长度
- 协议简化且可扩展
- 更好的安全性
- 增加服务类型
- 支持组播通信
- 有发展余地
- 新旧协议共存

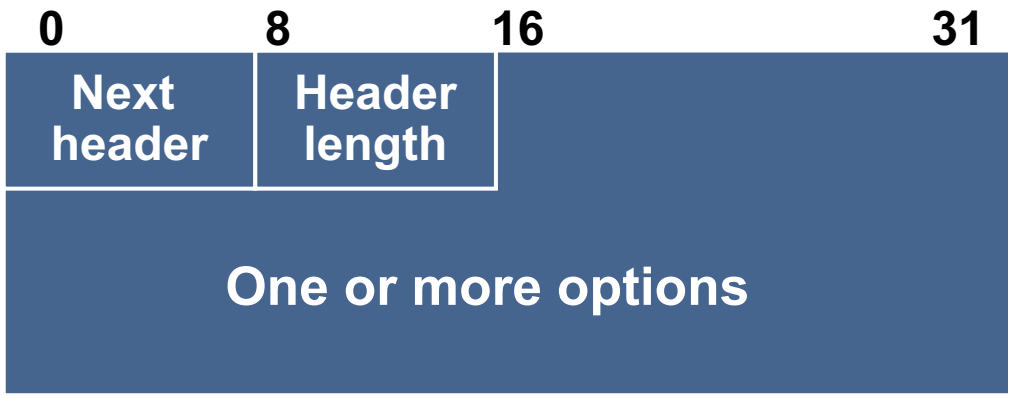
RFC1883
RFC1884
RFC1885
RFC1886
RFC1887

IPv6报文格式



- Diff. services：区分服务，含优先级
- Flow lable：流标记，用于标识一对主机进程之间的媒体数据流（音频，视频），用于服务质量保障
- Payload length：有效载荷的数据长度，包括扩展头
- Next header：下一个头或协议类型
- Hop limit：指明了IP包的生存期限
- Source address：IP包的发送方地址
- Destination address：IP包的接收方地址

IPv6扩展头格式



- Next header指出下一个扩展头的类型或者有效载荷包括的上层报文协议
- Header length指定可变长的扩展头内容长度

下一个扩展头：标识紧接着的下一个头类型或数据块类型。

Next header指定路由器无法处理选项时如何操作：

- 跳过该选项
- 丢弃包
- 丢弃包并返回ICMP包

IPv6扩展头类型

扩展头值	扩展头类型	功能描述
0	逐跳选项	指示路径上每个路由器要检查的选项字段，例如预约资源等。
43	松散路由	指出了一条必须经过指定中间节点的松散路由
44	分段选项	规定主机如何分段IPv6数据报
50	加密安全	指定如何加密有效载荷，缺省加密算法是数据加密标准（DES）
51	认证选项	用于验证发送方的身份，缺省认证算法是MD5
60	目标选项	传递给目标节点的额外信息，仅目标节点检查和处理

- 1 ICMP4
- 2 IGMP4
- 4 IP
- 6 TCP
- 17 UDP
- 41 IPv6固定头
- 46 RSVP
- 58 ICMPv6
- 89 OSPF

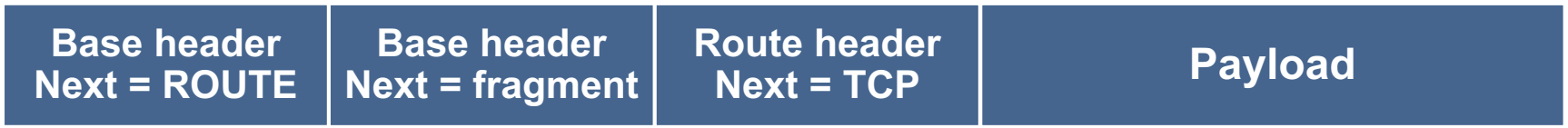
IPv6扩展头示例

假设：源端发送一个**TCP**报文

示例1：IPv6 包含一个基本头和数据

示例2：IPv6 包含一个基本头、一个路由头和数据

示例3：IPv6 包含一个基本头、一个路由头、一个分段头和数据



IPv6引入扩展头的好处

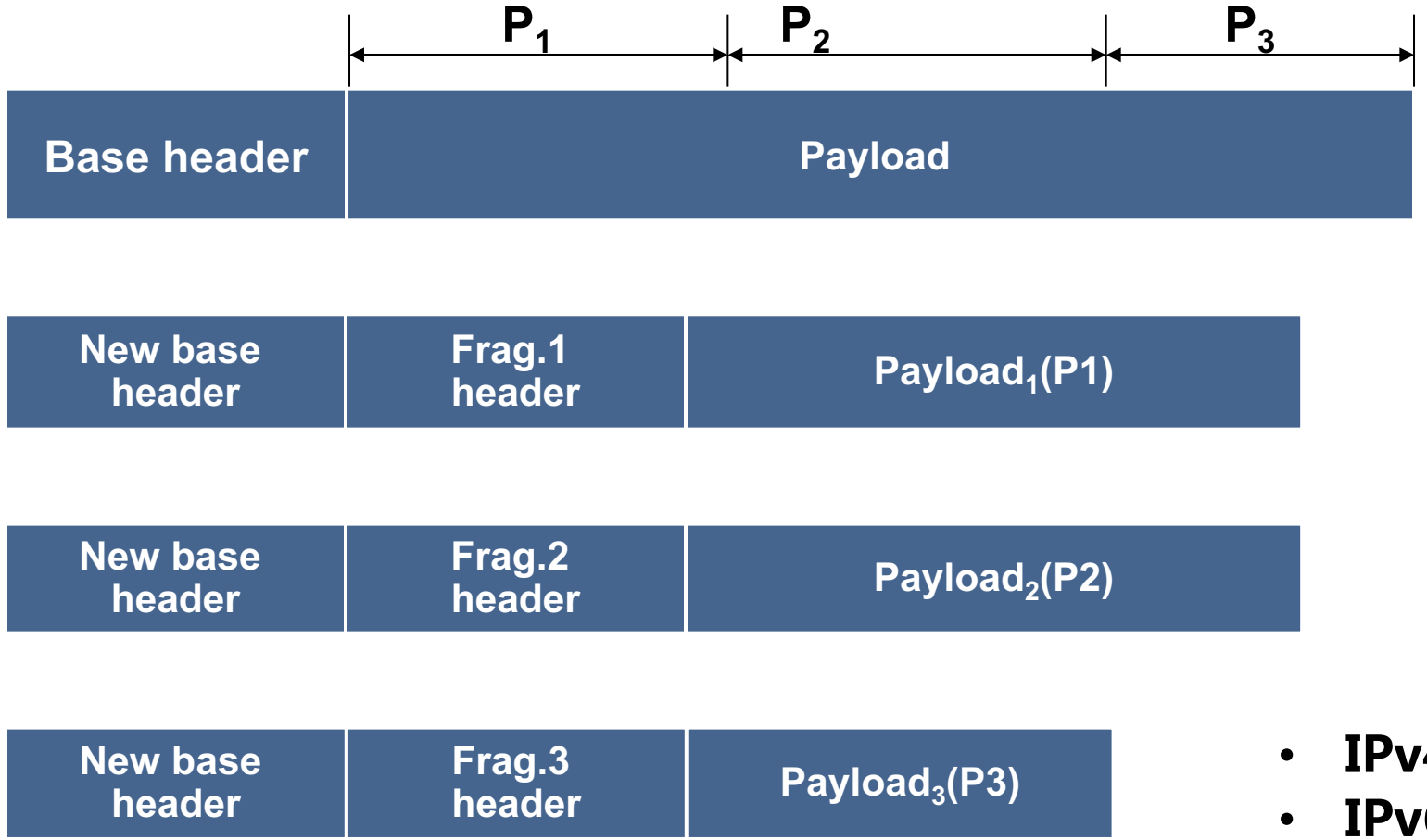
经济性

- 将数据报的功能划分到单独的头可节省空间
- 一个数据报只用所有包头的一个子集
- 按需使用包头更加灵活

扩展性

- 如要为协议增加一个新的功能只要定义一种新的next header类型和格式

IPv6数据报的分段、重组



每段组成

- 新的基本头
- 分段扩展头
- 数据区域

- IPv4：路由器负责分段任务
- IPv6：发送数据报的主机负责分段

IPv6协议的报文传递

IPv6地址表示及特性

IPv6地址表示法

- 每16位为一组，用十六进制表示
- 相邻两组之间用冒分分割

IPv4的点分十进制

105.220.136.100.255.255.255.255.0.0.18.125.140.10.255.255

IPv6的冒分十六进制

69OC:8864:FFFF:FFFF:0:1280:8COA:FFFF

IPv6地址的零压缩特性

- 用两个冒号代替连续的0
- 一个地址只能压缩一次

例如：

- 2007:1022:0001:0000:0000:0000:0000:1234
- 2007:1022:1::1234

IPv4到IPv6地址映射

- 96个零 ＋ IPv4(32位)

例如：

- IPv4地址“192.31.20.46”
- IPv6地址“::192.31.20.46”

IPv6地址分配

RFC3513

地址分配	前缀	IPv6表示
未指定地址	00…0(128位)	::/128
回环地址	00…1(128位)	::1/128
组播地址	11111111	FF00::/8
本地链路组播	1111111010	FE80::/10
本地网点组播	1111111011	FEC0::/10
全球单播地址	其他所有前缀	其他所有地址

- 本地链路组播指连接在同一条本地链路上的组成员
- 本地网点组播指本地私有网络内的组成员

IPv6单播地址——全局地址

全球聚合单播地址

- 类似于IPv4的单播地址
- 即IPv6的因特网地址，俗称公网地址
- 由前缀001标识

RFC3587

提供商ID（48位）	网点(16位)	主机ID(64位)

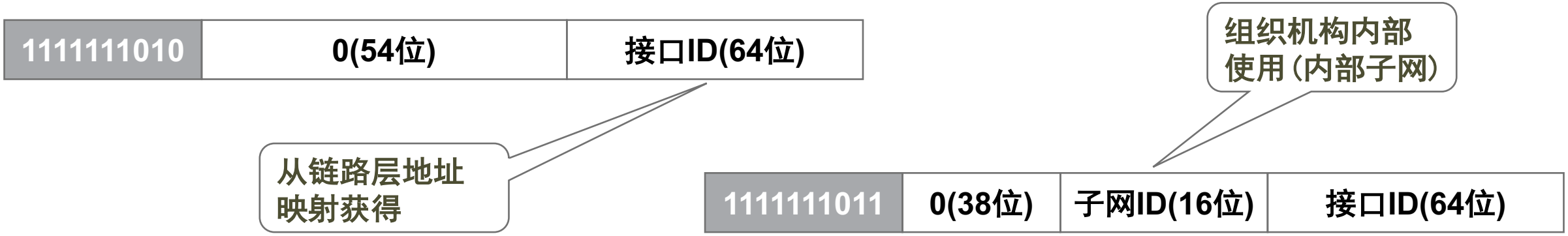
IPv6单播地址——本地地址

本地链路地址

- 应用范围受限的地址类型，只能在连接到同一个本地链路的节点之间使用。
- 当节点启动IPv6协议栈时，每个接口会自动配置一个本地链路地址

本地网点地址

- 应用范围受限的地址类型，只能在某个网点的内部使用（内网）。
- 不能自动生成



IPv6组播地址

组播地址

- 一个源节点发送的报文被多个特定的目标节点接收
- IPv6用特定的前缀标识一个组播地址

范围字段：标识了组播报文的扩散范围

- 0 预留
- 1 本地接口
- 2 本地链路
- 3 本地站点
- 4 本地组织机构
- E 全球范围
- F 预留

RFC2373

11111111	标志(4位)	范围(4位)	组ID(112位)

标志字段：标识组地址的永久性

- 0 永久组播地址
- 1 临时组播地址

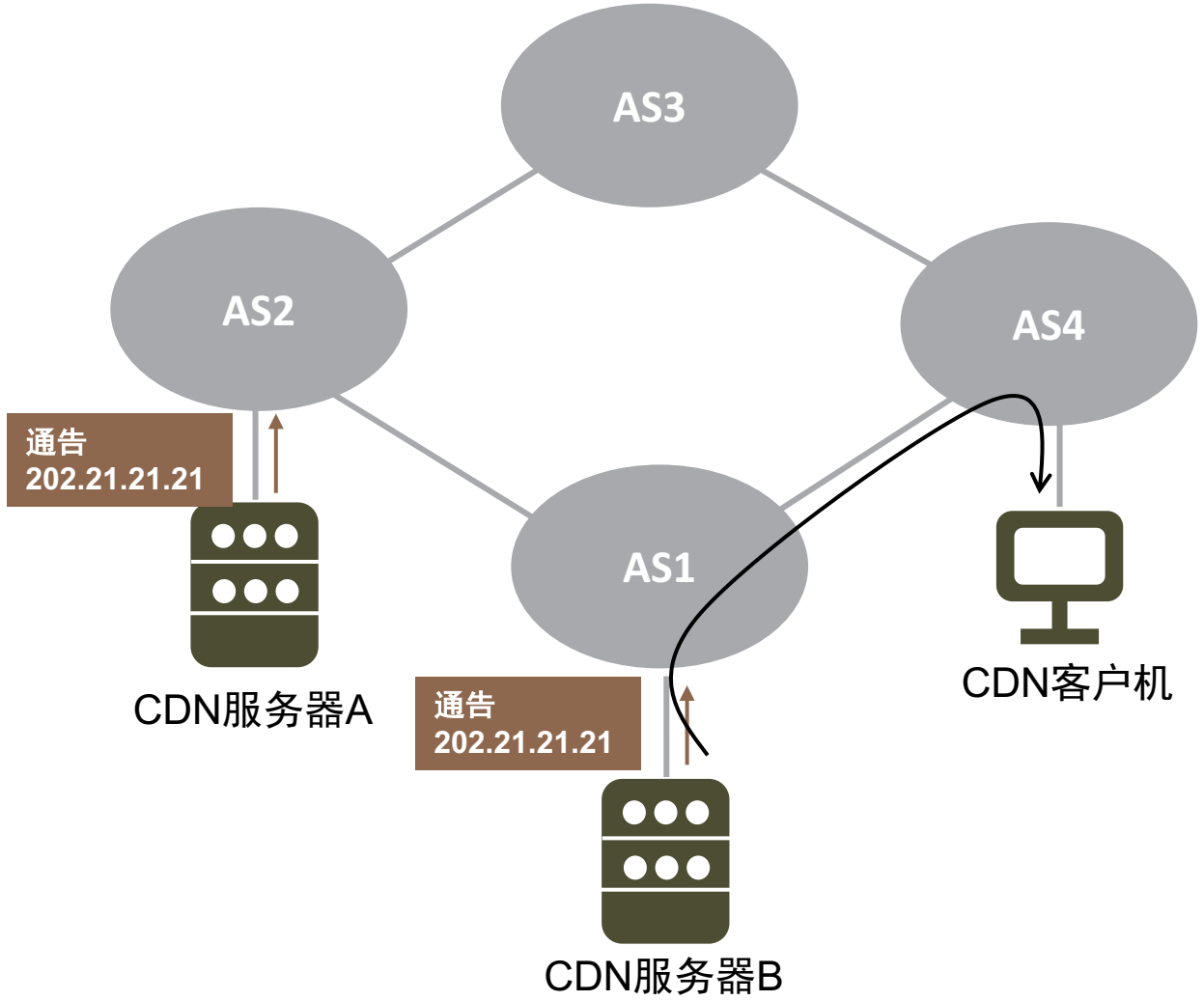
组ID字段：标识一个组播组

- 标准建议使用低32位表示组ID，高80为置0

IPv6选播地址

选播地址

- 目标地址是选播地址的IPv6包发送到最仅的一个组成员
- 与单播地址在形式上没有区别，必须通过显式方式指明



假设：一组同一机构的**CDN**服务器具有相同的**IP**地址

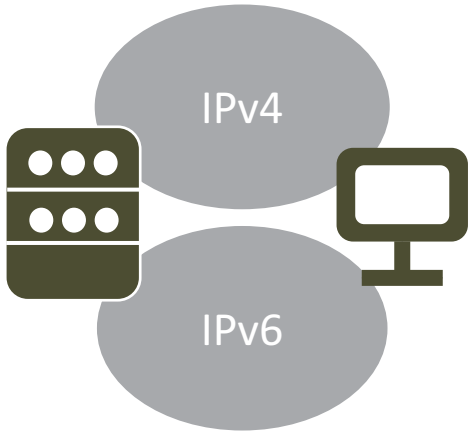
- 所有服务器通告自己的IPv6地址
- 客户机用同一个地址访问服务器
- 网络将客户机请求包路由到最近的服务器

IPv4向IPv6过渡

双协议栈

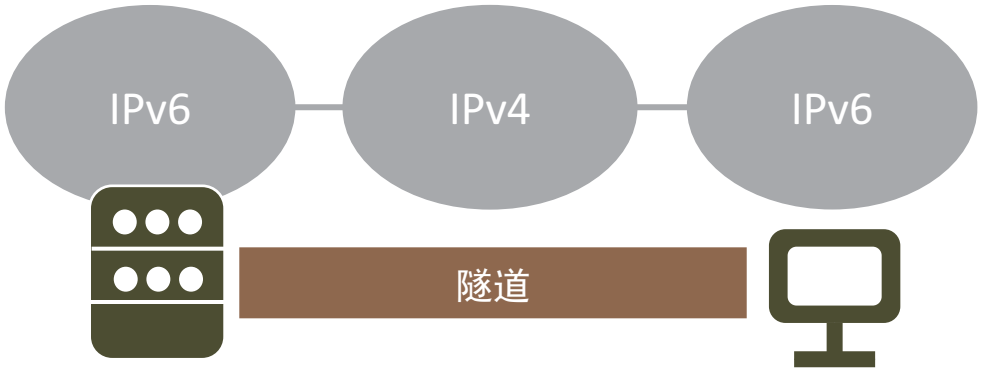
••双栈机制允许IPv4协议和IPv6协议在同一个网络中共存。

••注意：只能提供给相同协议包之间的转发



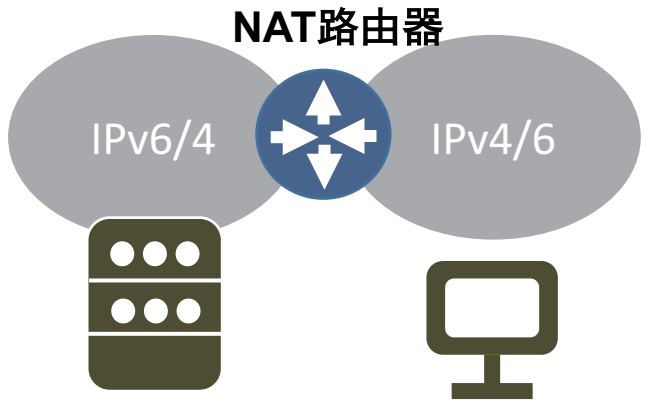
6-in-4隧道

网络边缘的IPv6节点利用隧道技术实现处于因特网边缘的多个IPv6网络的互联。



NAT-PT

在内联网或因特网边缘对IPv4, IPv6地址和报文格式进行转换，从而实现IPv6主机与IPv4主机的双向通信



IPv4与IPv6共存