

计算机网络

第四章 网络层

谢瑞桃

xie@szu.edu.cn

[rtxie.github.io](https://github.com/rtxie)

计算机与软件学院

深圳大学





第四章讲解内容

1. 网络层概述与数据平面
2. 控制平面

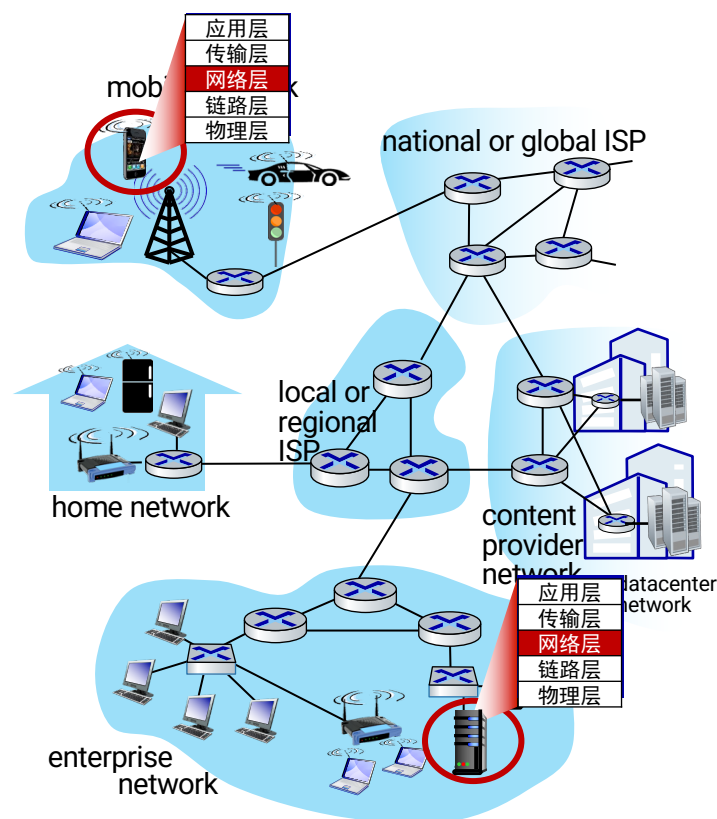


数据平面讲解内容

- 网络层概述与功能
- 路由器体系结构
- IP数据报格式与分片
- IP地址
- 转发
- 划分地址块与地址聚合
- DHCP协议
- 网络地址转换NAT
- IPv6

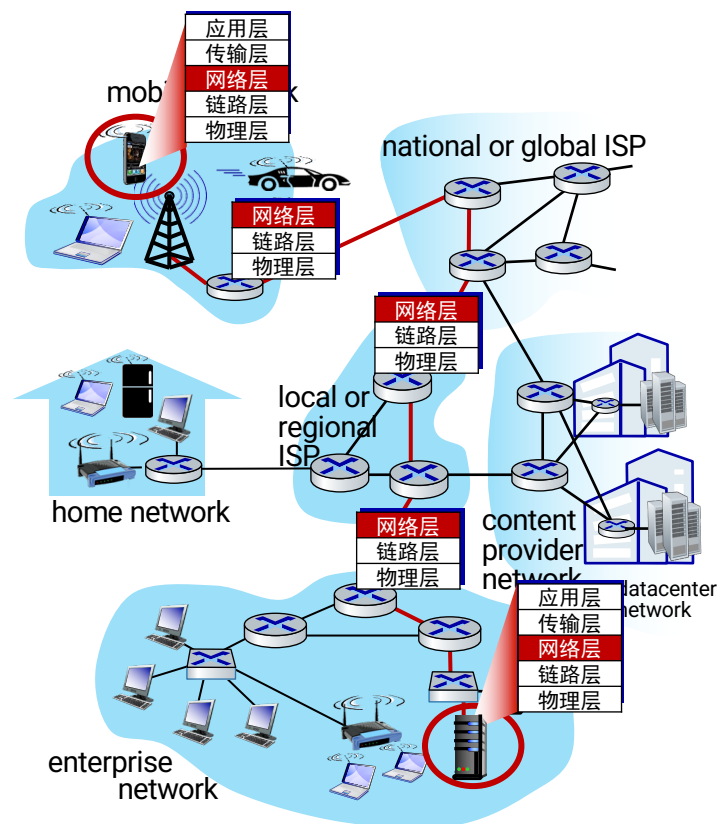
网络层

- 从发送主机向接收主机运送传输层的报文段
- 发送方将其封装**数据报**，接收方解封装



网络层

- 从发送主机向接收主机运送传输层的报文段
- 发送方将其封装**数据报**，接收方解封装
- 网络层协议运行在所有主机和**路由器**上
- 路由器查看IP数据报首部，从而决定如何转发





网络层的功能

■ 转发

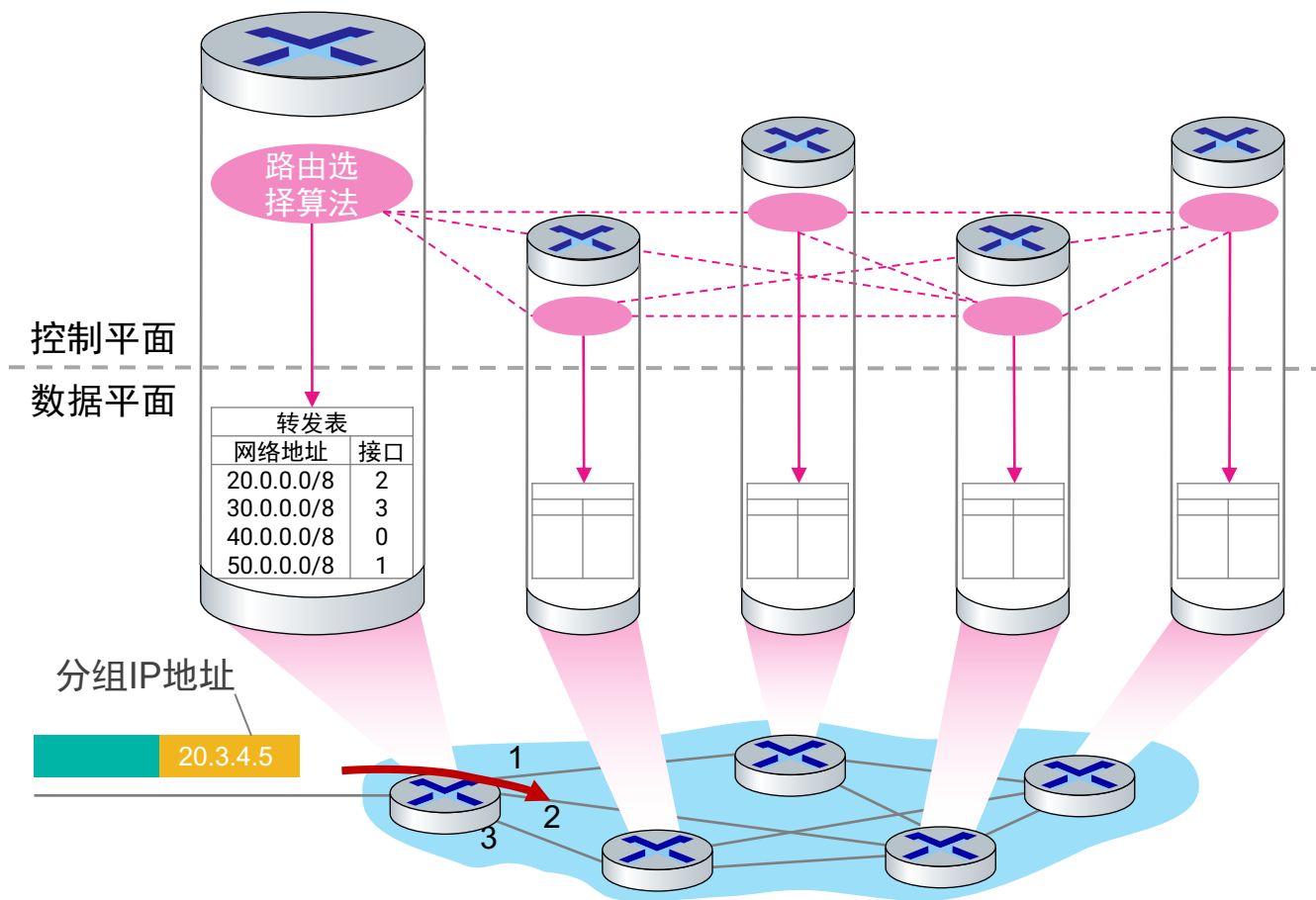
- 路由器将分组由一条输入链路移动到适当的输出链路
- 时间要求是纳秒级别
- 由硬件实现

■ 路由选择

- 决定将分组由源主机移动到目的主机所要经过的路由或路径
- 时间要求是秒级别
- 由软件实现

网络层的功能

- 每个路由器的路由选择组件相互通信，合作生成转发表





第四章知识点汇总

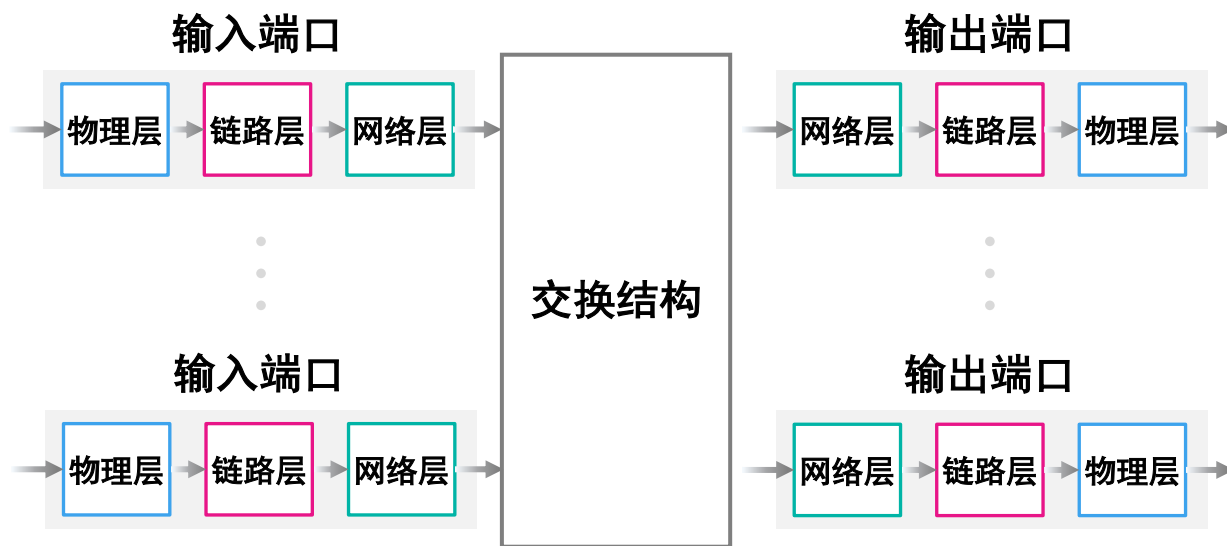
- 理解网络层在协议栈中的作用
- 了解网络层的功能：转发和路由选择



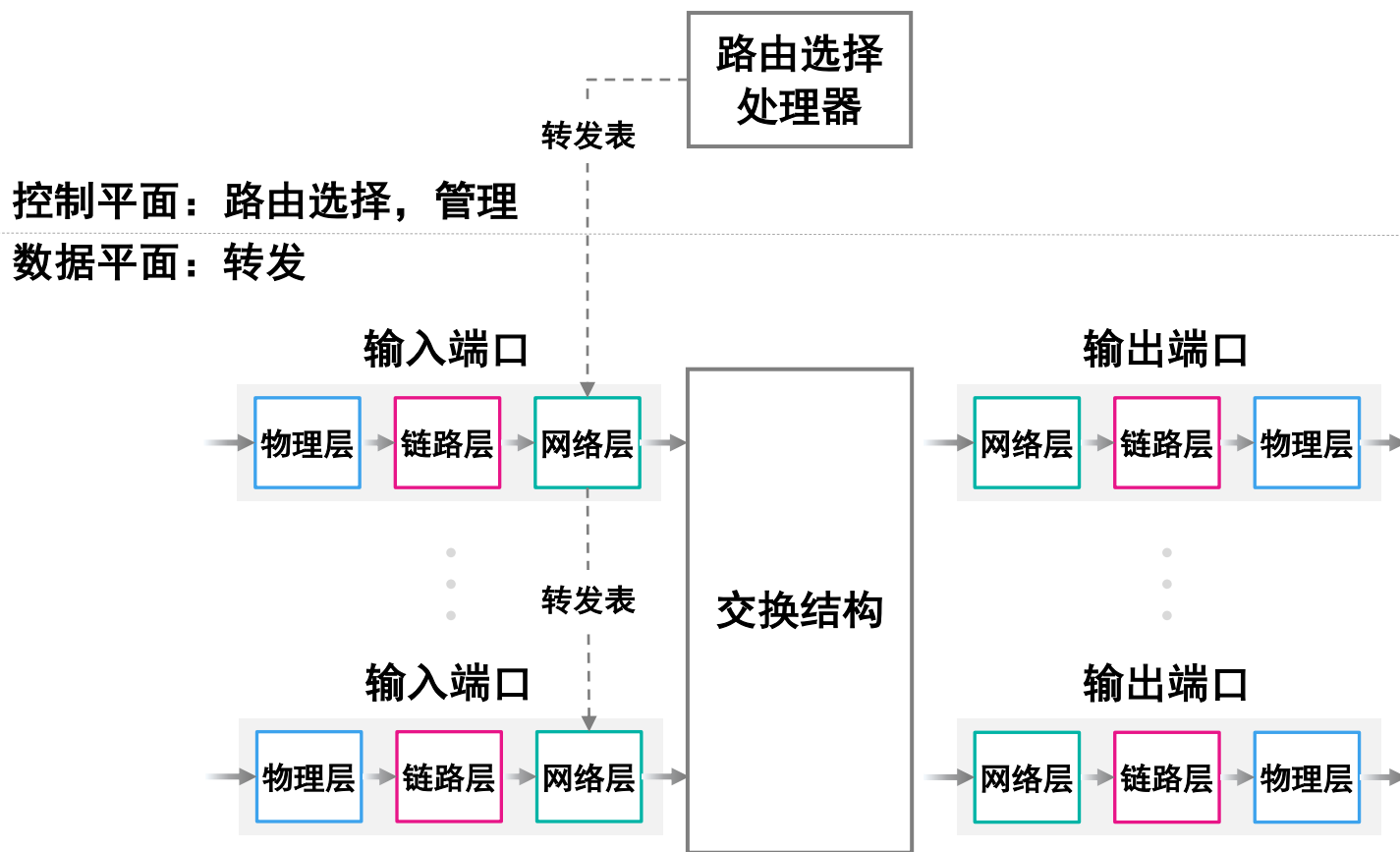
数据平面讲解内容

- 网络层概述与功能
- 路由器体系结构
- IP数据报格式与分片
- IP地址
- 转发
- 划分地址块与地址聚合
- DHCP协议
- 网络地址转换NAT
- IPv6

路由器体系结构



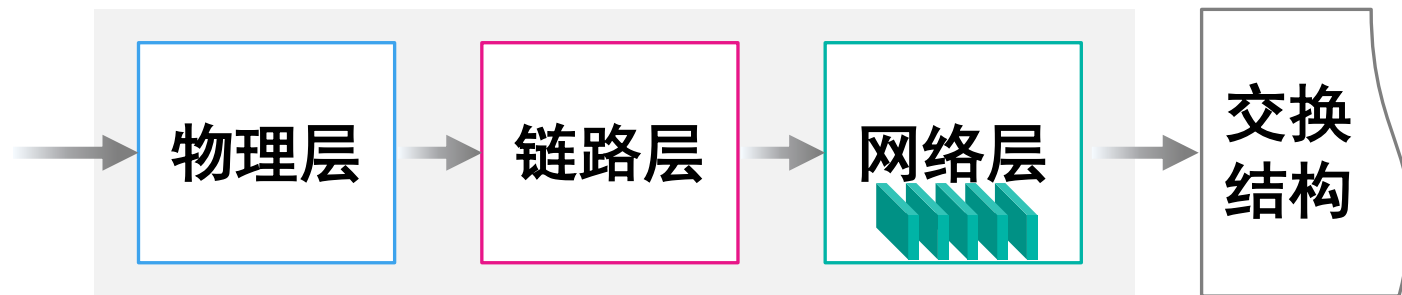
路由器体系结构





输入端口

- 网络层：
 - 查找转发表
 - 转发
 - 如果交换结构不能很快（相对输入链路速度而言）地将到达分组转移到输出接口，则会出现排队

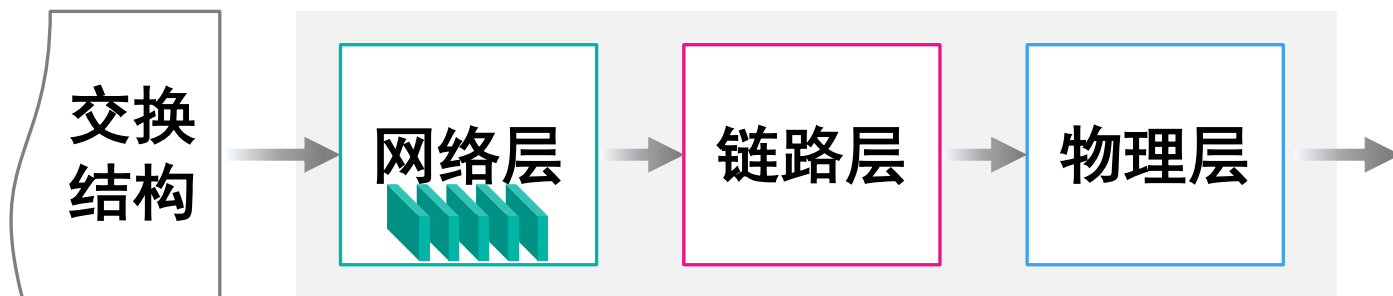




输出端口

■ 网络层:

- 如果数据报到达输出端口的速度快于离开口口的速度（链路发送速度），则需要队列
- 数据报会因为队列溢出而被丢弃
- 丢包和排队时延因此而产生





第四章知识点汇总

- 了解路由器的体系结构
- 了解路由器的输入/输出端口的作用
- 理解丢包与排队时延产生的原因



数据平面讲解内容

- 网络层概述与功能
- 路由器体系结构
- **IP数据报格式与分片**
- IP地址
- 转发
- 划分地址块与地址聚合
- DHCP协议
- 网络地址转换NAT
- IPv6

IP数据报格式



- 首部长度的4位
- 单位是4字节
- 最小值为0101, 十进制为5, 表示 $5 \times 4 = 20$ 字节, 即固定首部的长度
- 最大值为1111, 十进制为15, 表示 $15 \times 4 = 60$ 字节, 即最大的首部长度的

IP数据报格式



- **服务类型：**

区分不同应用的数据报，以提供不同类型的服务

IP数据报分片/组装

- **Time-To-Live：**

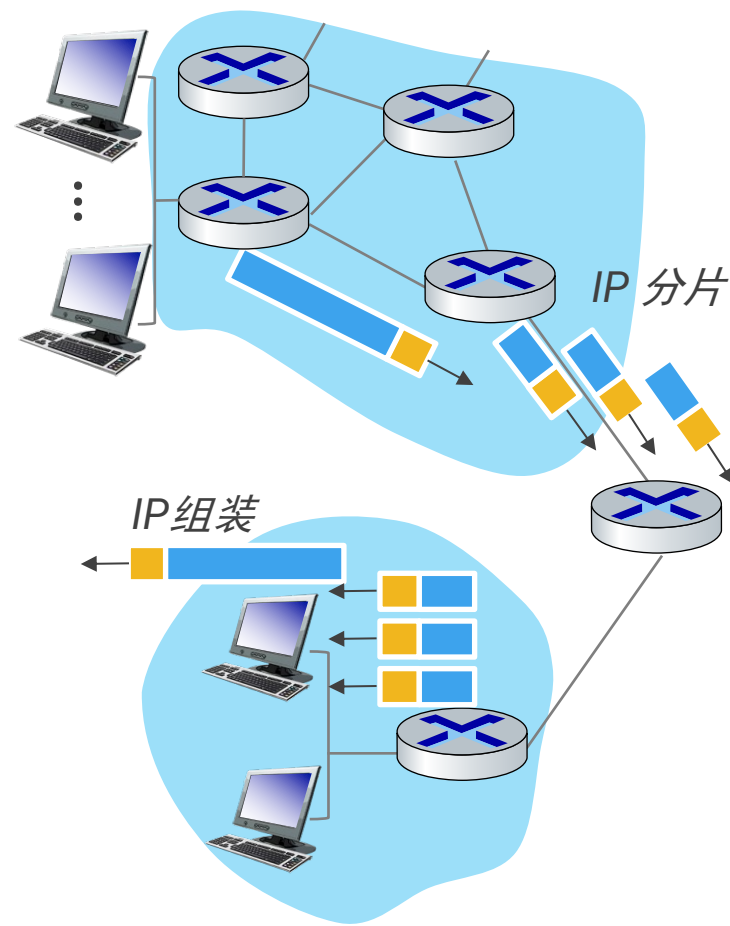
分组寿命，路由器处理完IP数据报后减一，并丢弃TTL为0的数据报

- **上层协议：**

6表示TCP，17表示UDP，指导数据应该交给哪个协议

IP数据报分片

- 网络链路有MTU，限制了IP数据报的长度
 - 不同的链路类型具有不同的MTU
- IP数据报长度 > 链路MTU，被分片
 - 在目的端IP层组装之后才向上交付
 - IP首部的一些比特位用于标识分片



IP数据报分片

| 数据报长度 | 16位标识 | 0 位 | DF 位 | MF 位 | 13位片偏移 |
|-------|----------------|--|---------|---------|------------------------|
| 首部+数据 | 由源主机添加，标识相同的分片 | DF: 0=May Fragment, 1=Don't Fragment. MF: 0=Last Fragment, 1=More Fragments. | | | 数据部分的 偏移，8字节 为单位 |



IP数据报分片

- 举例：数据报4000字节，MTU=1500字节

| | | | | | | |
|----------------------|------------------------|--------------|---|-----------|------------------|--|
| 20字节首部 + 3980字节数据 | 数据报长度 = 4000 | 16位标识 = x | 0 | DF = 0 | MF = 0 | 13位片偏移 = 0 |
| 20字节首部 + 1480字节数据 | 数据报长度 = 1500 | 16位标识 = x | 0 | DF = 0 | MF = 1 | 13位片偏移 = 0 |
| 20字节首部 + 1480字节数据 | 数据报长度 = 1500 | 16位标识 = x | 0 | DF = 0 | MF = 1 | 13位片偏移 = $1480/8 = 185$ |
| 20字节首部 + 1020字节数据 | 数据报长度 = 1040 | 16位标识 = x | 0 | DF = 0 | MF = 0 | 13位片偏移 = 370 |



第四章知识点汇总

- 了解IP数据报格式
- 理解IP数据报分片的原因
- 掌握IP数据报分片的方法



习题

- 假设IPv4数据报长度为1500字节，用TCP发送一个5MB的文件需要（ ）个IP分组。假定所有协议不使用选项，只使用固定长度的头部。(1M=10⁶)
- A. 3334
- B. 3379
- C. 3425
- D. 3473

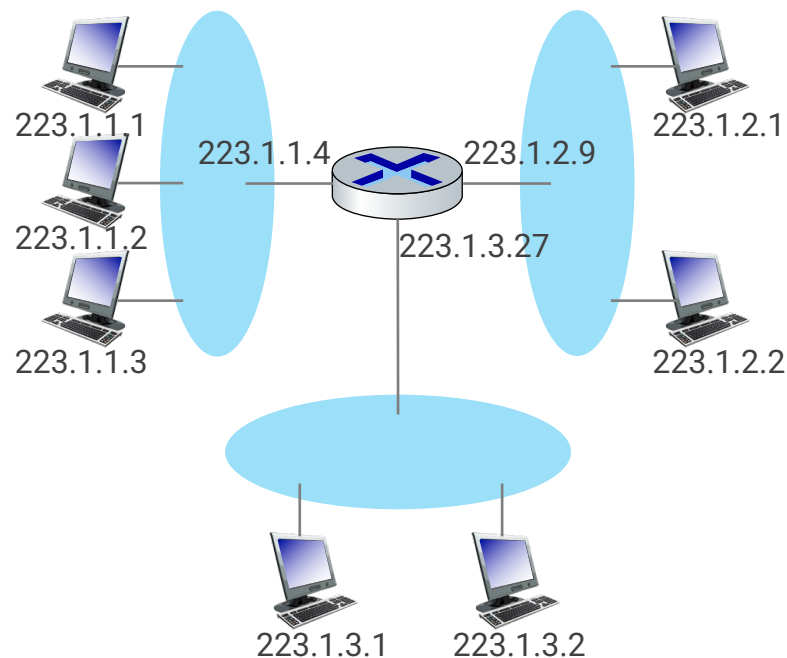


数据平面讲解内容

- 网络层概述与功能
- 路由器体系结构
- IP数据报格式与分片
- **IP地址**
- 转发
- 划分地址块与地址聚合
- DHCP协议
- 网络地址转换NAT
- IPv6

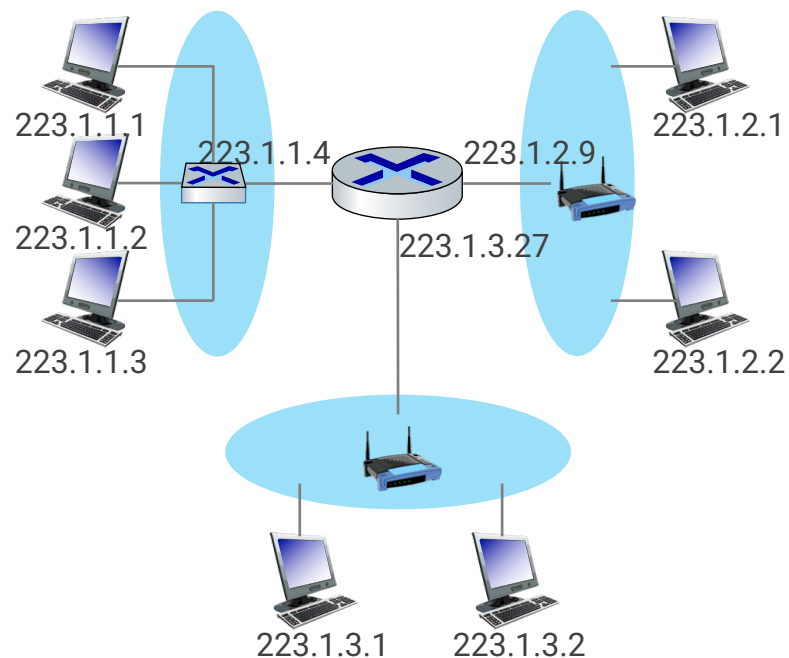
IP地址

- IP地址：主机/路由器接口的标识符
- 接口：
 - 路由器通常有多个接口
 - 主机通常有一个或多个接口（有线以太网、无线802.11）
- IP地址与接口对应



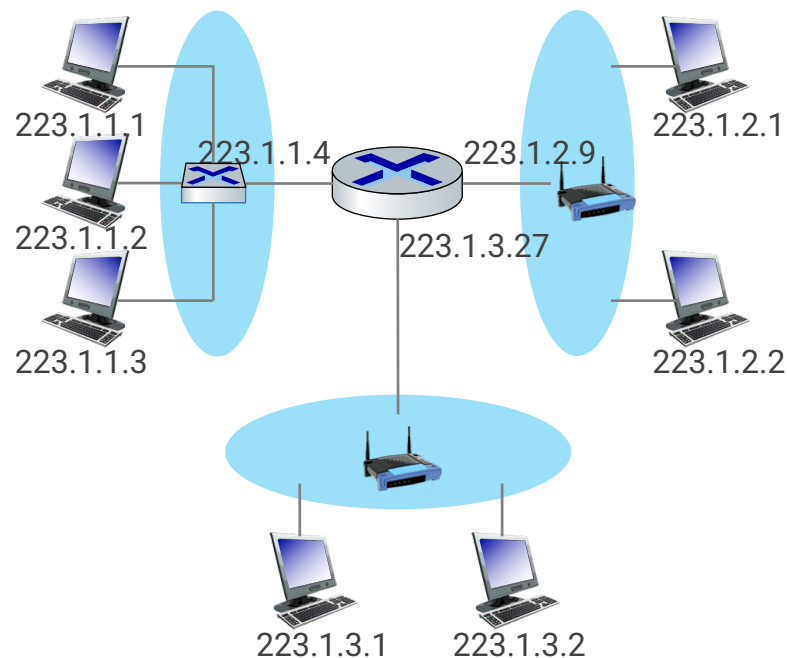
IP地址

- 接口之间如何连接？
 - 通过以太网交换机
 - 通过WiFi接入点
- 将在链路层章节介绍



IP地址

- 接口之间如何连接？
 - 通过以太网交换机
 - 通过WiFi接入点
- 物理上不通过三层网络设备连接的接口形成子网





IP地址

- IP地址
 - 32位，4字节
 - 点分十进制计法

11011111 00000001 00000001 00000001

223 . 1 . 1 . 1



IP地址

- IP地址
 - 32位，4字节
 - 点分十进制计法

11011111 00000001 00000001 00000001

223 . 1 . 1 . 1

子网部分 主机部分

- 两层结构
 - 子网部分：高位，标识子网，IP转发（将分组运送到目的子网）时使用
 - 主机部分：低位，标识子网内部的主机

两部分如何划分呢？



IP地址

- 怎么划分子网部分和主机部分呢？
- 分类编址
- 无类别编址：CIDR



IP地址：分类编址

- 网络部分定长
- 以前的编址方法 [RFC870, 1983]，已不用

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
|----|---|------|------|------|---|---|---|---|------|---|---|---|---|---|---|---|------|---|---|---|---|---|---|---|---|---|------|---|---|---|---|---|--|
| A类 | 0 | 网络部分 | | | | | | | 主机部分 | | | | | | | | | | | | | | | | | | | | | | | | |
| B类 | 1 | 0 | 网络部分 | | | | | | | | | | | | | | 主机部分 | | | | | | | | | | | | | | | | |
| C类 | 1 | 1 | 0 | 网络部分 | | | | | | | | | | | | | | | | | | | | | | | 主机部分 | | | | | | |

IP地址：分类编址

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|---|------|------|------|---|---|---|---|------|---|---|---|---|---|---|---|------|---|---|---|---|---|---|---|------|---|---|---|---|---|---|---|--|--|--|--|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | | | | |
| A类 | 0 | 网络部分 | | | | | | | 主机部分 | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| B类 | 1 | 0 | 网络部分 | | | | | | | | | | | | | | 主机部分 | | | | | | | | | | | | | | | | | | | |
| C类 | 1 | 1 | 0 | 网络部分 | | | | | | | | | | | | | | | | | | | | | 主机部分 | | | | | | | | | | | |

| 类别 | 网络部分长度 | 每个网络的主机个数 (主机部分取值的个数) | 网络个数 (网络部分取值的个数) |
|----|--------|-----------------------------|----------------------------|
| A类 | 1字节 | $2^{24} - 2 = 16777216 - 2$ | $2^7 - 2 = 128 - 2$ |
| B类 | 2字节 | $2^{16} - 2 = 65536 - 2$ | $2^{14} - 2 = 16384 - 2$ |
| C类 | 3字节 | $2^8 - 2 = 256 - 2$ | $2^{21} - 2 = 2097152 - 2$ |



IP地址：分类编址

- 存在的问题：地址浪费严重
- 对于一个机构，C类（最多254个主机地址）太小，而B类（最多65534个主机地址）太大

| 类别 | 网络部分长度 | 每个网络的主机个数 (主机部分取值的个数) | 网络个数 (网络部分取值的个数) |
|----|--------|-----------------------------|----------------------------|
| A类 | 1字节 | $2^{24} - 2 = 16777216 - 2$ | $2^7 - 2 = 128 - 2$ |
| B类 | 2字节 | $2^{16} - 2 = 65536 - 2$ | $2^{14} - 2 = 16384 - 2$ |
| C类 | 3字节 | $2^8 - 2 = 256 - 2$ | $2^{21} - 2 = 2097152 - 2$ |



IP地址: CIDR

- 无类别域间路由选择（**C**lassless **I**nter**d**omain **R**outing CIDR）[1990s首次提出，RFC4632 (2006更新)]
 - 地址格式 $a.b.c.d/x$ ，其中 x 是网络部分（前缀）的位数
 - x 可以是任意 $[0, 32]$ 之间的值

网络部分(前缀)

主机部分

11001000 00010111 00010000 **0** 00000000

200.23.16.**0/23**

- 主机号全零：网络地址
- 主机号全幺：广播地址



子网掩码 subnet mask

- 网络部分（前缀）全幺，主机部分全零
- 举例：
- 网络地址：200.23.16.0/23
- 对应子网掩码：255.255.254.0
- 以前的B类地址：172.16.0.0
- 对应子网掩码：255.255.0.0



子网掩码 subnet mask

- 用途：IP地址与子网掩码做“按位与”操作，提出网络部分，进而判断该IP地址是否属于某个网络
- 网络地址：200.23.16.0/23
- 对应子网掩码：255.255.254.0
- 实例1：200.23.17.66属于上述网络
- 11001000 00010111 00010001 01000010
- 11111111 11111111 11111110 00000000
- 11001000 00010111 00010000 00000000

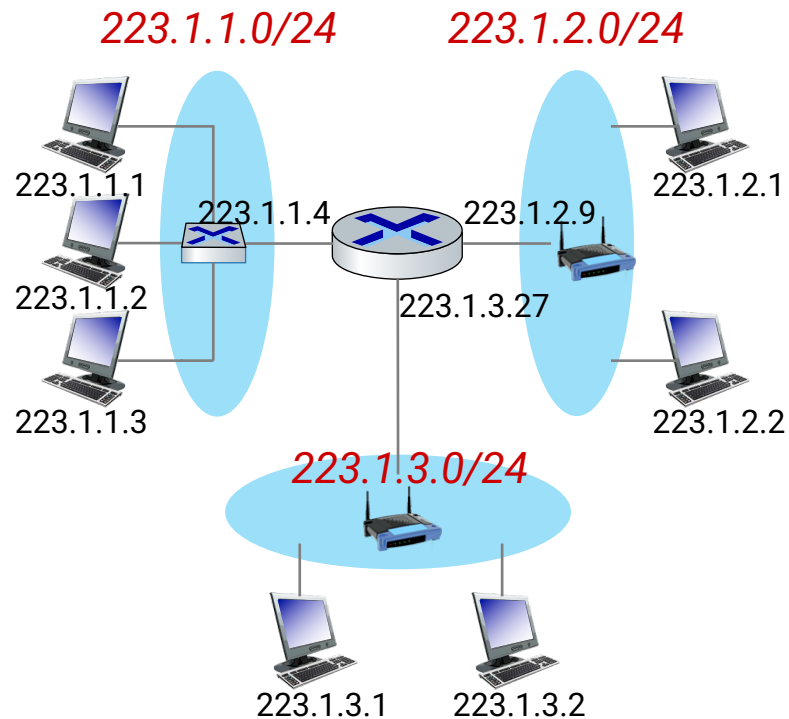


子网掩码 subnet mask

- 用途：IP地址与子网掩码做“按位与”操作，提出网络部分，进而判断该IP地址是否属于某个网络
- 网络地址：200.23.16.0/23
- 对应子网掩码：255.255.254.0
- 实例2：200.23.19.66不属于上述网络
- 11001000 00010111 00010011 01000010
- 11111111 11111111 11111110 00000000
- 11001000 00010111 00010010 00000000

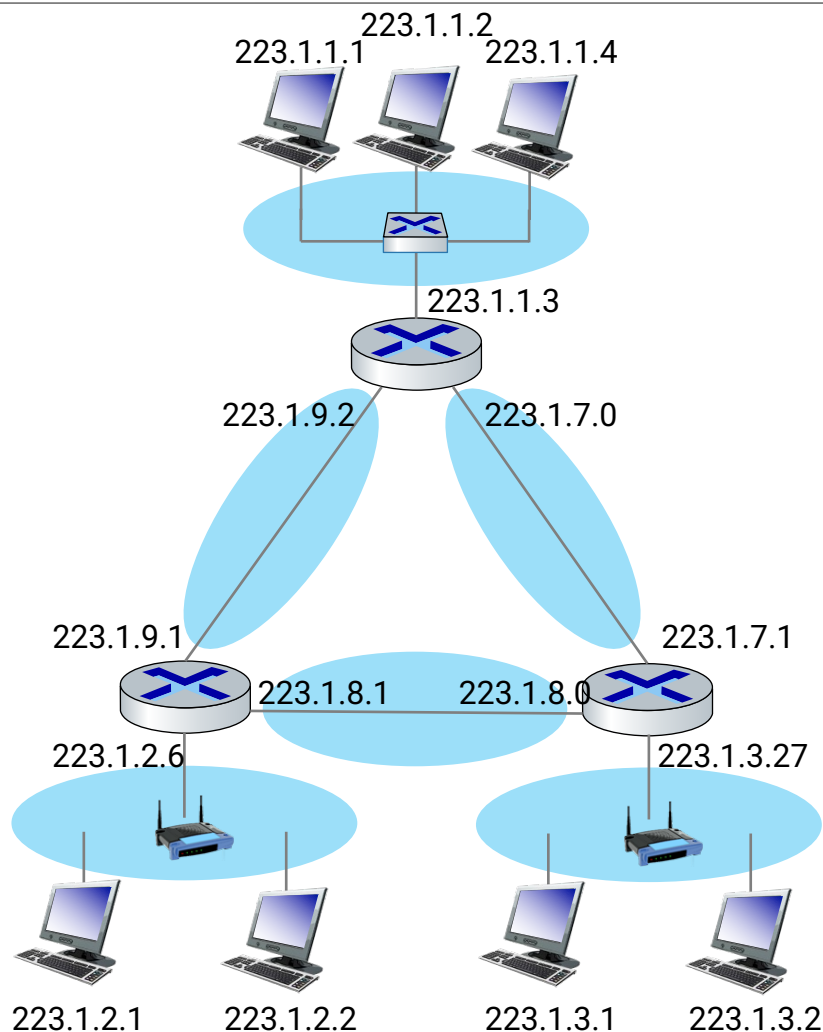
IP地址与子网

- 该网络中含有多少个子网？
- 物理上不通过三层设备连接的接口形成子网
- 每个子网合理的网络地址分别是什么？



子网

- 该网络中含有多少个子网？
- 物理上不通过三层设备连接的接口形成子网



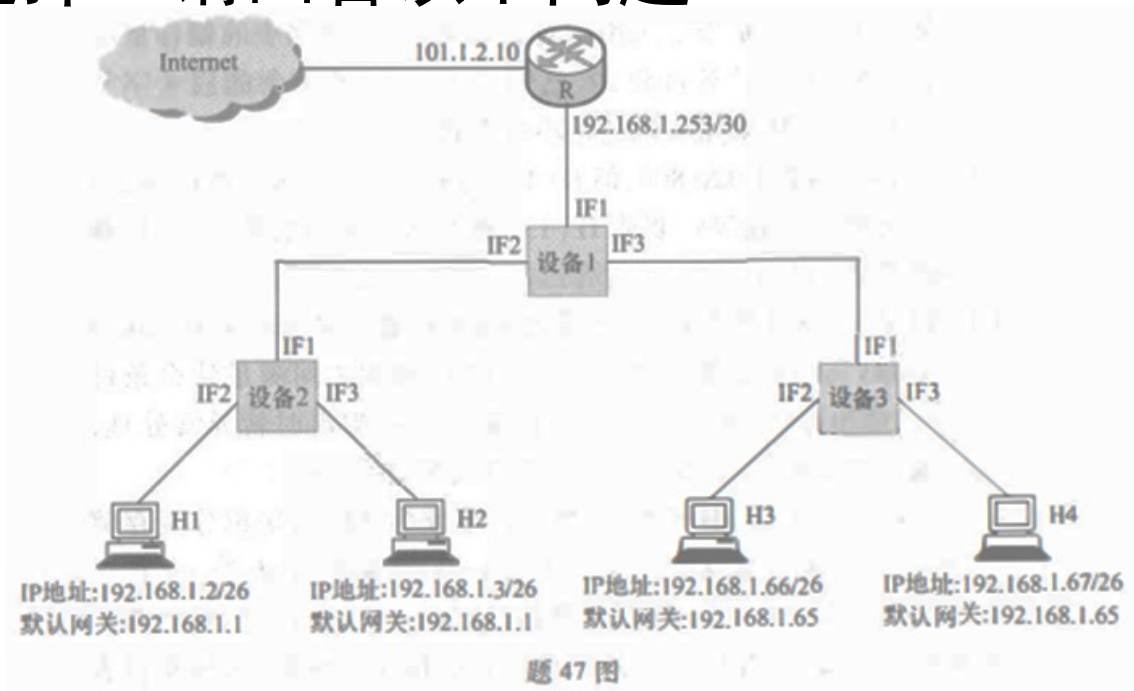


第四章知识点汇总

- 了解IP地址和网络地址
- 了解子网
- 了解CIDR
- 了解以前的A/B/C分类编址方法，以及其缺点

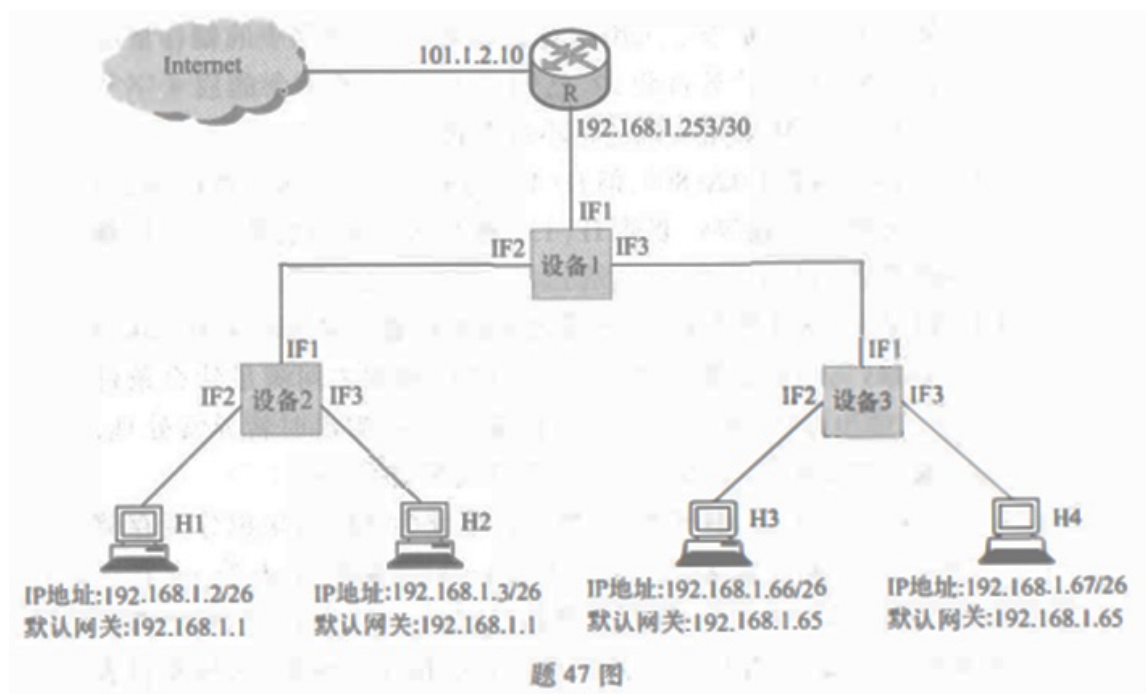
习题

- 【2019年考研47题】某网络拓扑如图所示，其中R为路由器，主机H1-H4的IP地址配置以及R的各接口IP地址配置如图所示。现在若干台以太网交换机（无VLAN功能）和路由器两类网络互连设备可供选择。请回答以下问题：



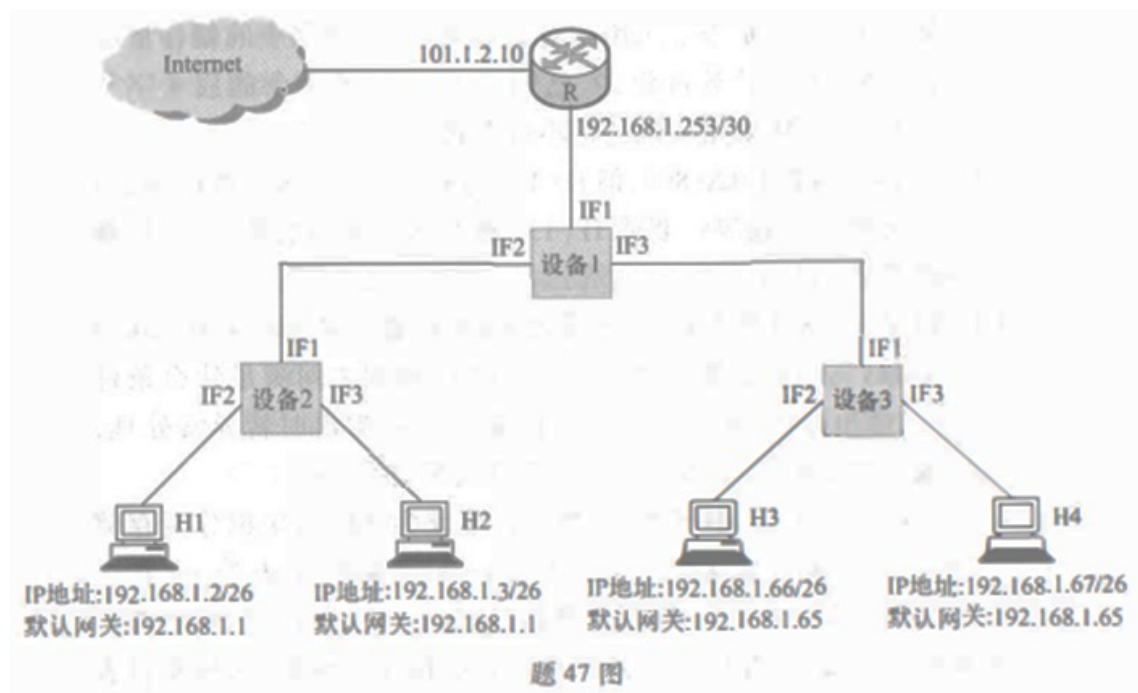
习题

- （1）设备1、设备2和设备3分别应选择什么类型网络设备？



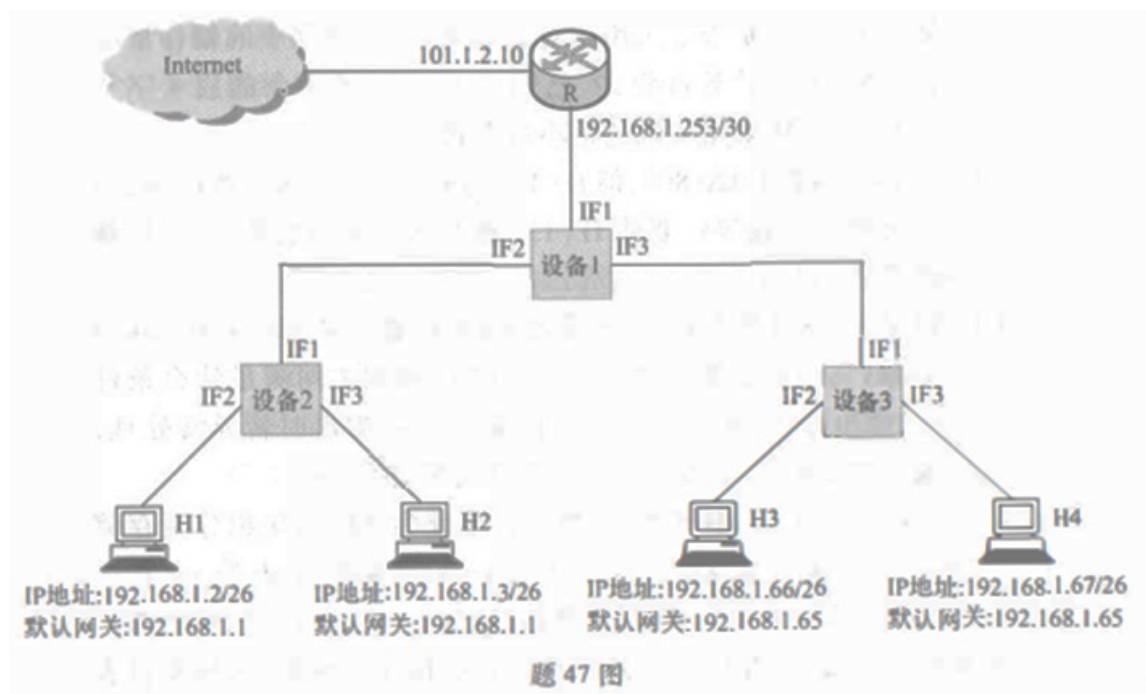
习题

- (2) 设备1、设备2和设备3中，哪几个设备的接口需要配置IP地址？并为对应的接口配置正确的IP地址。



习题

- (4) 若主机H3发送一个目的地址为192.168.1.127的IP数据报，网络中哪几个主机会接收该数据报？





数据平面讲解内容

- 网络层概述与功能
- 路由器体系结构
- IP数据报格式与分片
- IP地址
- 转发
- 划分地址块与地址聚合
- DHCP协议
- 网络地址转换NAT
- IPv6



转发

- 查询转发表，用目的IP地址匹配条目
 - 11001000 00010111 00010110 10100001 转发接口？
 - 11001000 00010111 00011000 10101010 转发接口？

| 目的地址范围 | 链路接口 |
|--|------|
| 11001000 00010111 00010000 00000000 到 11001000 00010111 00010111 11111111 | 0 |
| 11001000 00010111 00011000 00000000 到 11001000 00010111 00011000 11111111 | 1 |
| 11001000 00010111 00011001 00000000 到 11001000 00010111 00011111 11111111 | 2 |
| 其他 | 3 |



转发

- 查询转发表，用目的IP地址匹配条目
 - 11001000 00010111 00010110 10100001 转发接口：0
 - 11001000 00010111 00011000 10101010 转发接口：1

| 目的地址范围 | 链路接口 |
|--|------|
| 11001000 00010111 00010000 00000000 到 11001000 00010111 00010111 11111111 | 0 |
| 11001000 00010111 00011000 00000000 到 11001000 00010111 00011000 11111111 | 1 |
| 11001000 00010111 00011001 00000000 到 11001000 00010111 00011111 11111111 | 2 |
| 其他 | 3 |



转发

- 查询转发表，用目的IP地址匹配条目
 - 11001000 00010111 00010110 10100001 转发接口？
 - 11001000 00010111 00011000 10101010 转发接口？

| 网络地址 | 链路接口 |
|--|------|
| 11001000 00010111 00010000 00000000/21 | 0 |
| 11001000 00010111 00011000 00000000/24 | 1 |
| 11001000 00010111 00011000 00000000/21 | 2 |
| 其他 | 3 |



转发

- 查询转发表，用目的IP地址匹配条目
 - 11001000 00010111 00010110 10100001 转发接口：0
 - 11001000 00010111 00011000 10101010 转发接口：1
- 最长前缀匹配
 - 转发的时候，利用最长前缀匹配目的地址


| 网络地址 | 链路接口 |
|--|------|
| 11001000 00010111 00010000 00000000/21 | 0 |
| 11001000 00010111 00011000 00000000/24 | 1 |
| 11001000 00010111 00011000 00000000/21 | 2 |
| 其他 | 3 |



转发

■ 以上两个转发表等效

| 目的地址范围 | 链路接口 |
|--|------|
| 11001000 00010111 00010000 00000000 到 11001000 00010111 00010111 11111111 | 0 |
| 11001000 00010111 00011000 00000000 到 11001000 00010111 00011000 11111111 | 1 |
| 11001000 00010111 00011001 00000000 到 11001000 00010111 00011111 11111111 | 2 |
| 其他 | 3 |



| 网络地址 | 链路接口 |
|--|------|
| 11001000 00010111 00010000 00000000/21 | 0 |
| 11001000 00010111 00011000 00000000/24 | 1 |
| 11001000 00010111 00011000 00000000/21 | 2 |
| 其他 | 3 |



转发

- 查询转发表，用目的IP地址匹配条目
 - 200.23.22.161 转发接口？
 - 200.23.24.170 转发接口？

| 网络地址 | 链路接口 |
|----------------|------|
| 200.23.16.0/21 | 0 |
| 200.23.24.0/24 | 1 |
| 200.23.24.0/21 | 2 |
| 其他 | 3 |

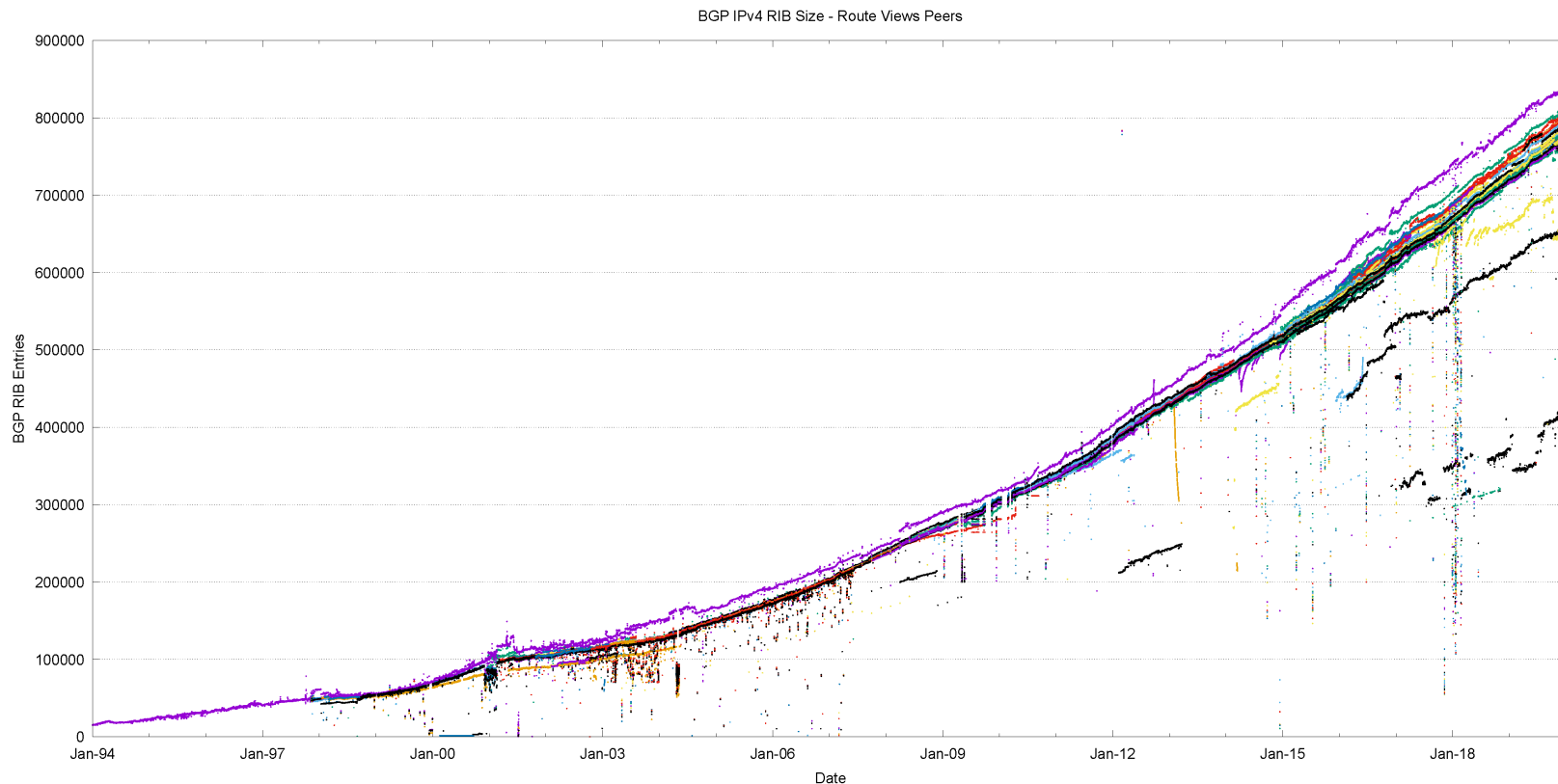


转发

- 路由器的链路速率100 Gbps
- 一个分组512位
- 大约每隔5ns，就有一个分组需要处理
- 查找1,000,000路由条目，在5ns内实现32位值的匹配，如何实现？
- 即使用一个最佳平衡二分搜索结构，也需要20次查找，每次查找要在0.25ns时间内完成。

IPv4路由表

路由表条目的规模接近 10^6 .



<https://www.potaroo.net/ispcol/2020-01/bgp2019.pdf>





第四章知识点汇总

- 理解转发的原理
- 理解最长前缀匹配的原理

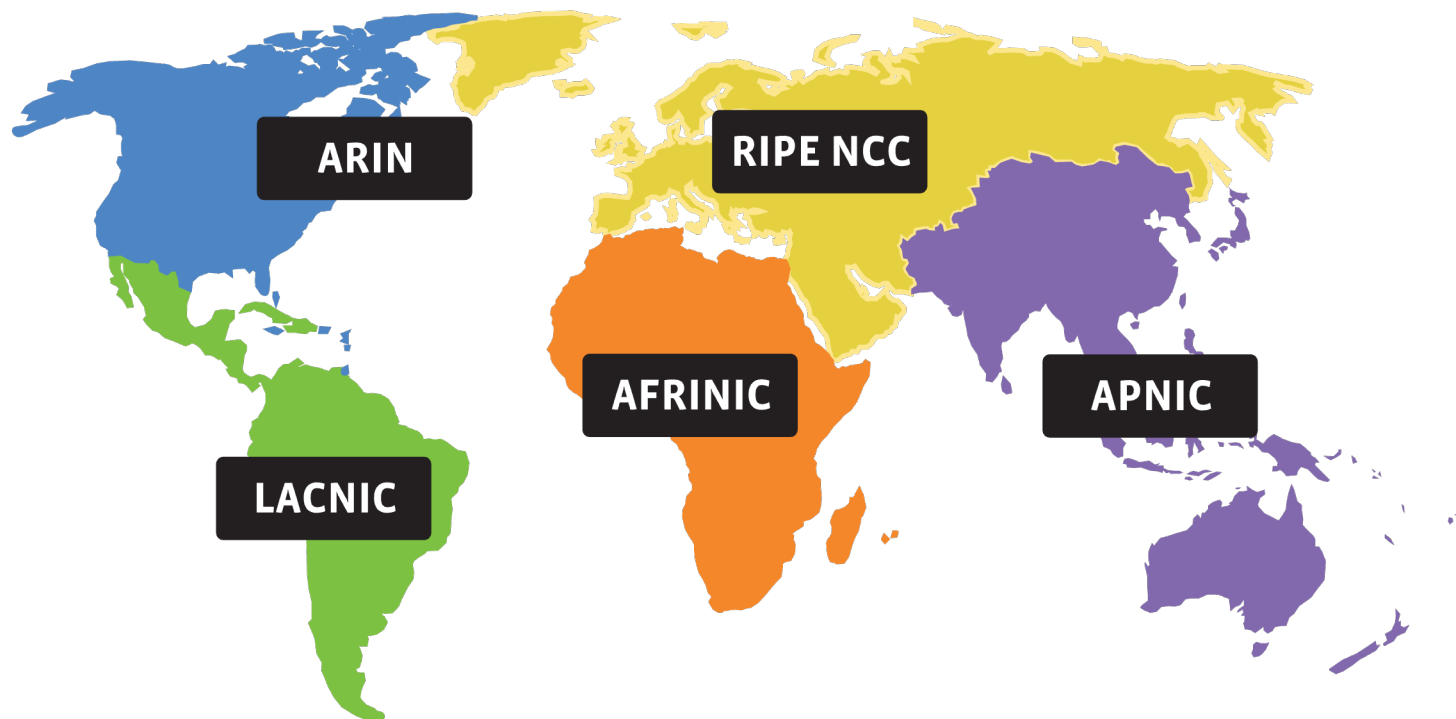


数据平面讲解内容

- 网络层概述与功能
- 路由器体系结构
- IP数据报格式与分片
- IP地址
- 转发
- 划分地址块与地址聚合
- DHCP协议
- 网络地址转换NAT
- IPv6

如何获取IP地址?

- IP地址分配由全球机构IANA ([Internet Assigned Numbers Authority](https://www.iana.org/)) 管理
- 下辖五个Regional Internet Registry



<https://www.iana.org/numbers>



如何获取IP地址?

- ISP向RIR（Regional Internet Registry）申请获得IP地址块
- 公司/学校/机构向ISP申请地址块
- 网络管理员负责为路由器接口分配IP地址



划分地址块

- ISP有一个地址块
- 11001000 00010111 00010000 00000000
200.23.16.0/20
- 有8个机构申请地址块，每个机构需要512个地址
- ISP该如何分配地址块呢？



划分地址块

- ISP有一个地址块

- 11001000 00010111 00010000 00000000 200.23.16.0/20

- 机构地址块：

- 11001000 00010111 00010000 00000000 200.23.16.0/23

11001000 00010111 00010001 11111111

- 11001000 00010111 00010010 00000000 200.23.18.0/23

11001000 00010111 00010011 11111111

- 11001000 00010111 00010100 00000000 200.23.20.0/23

-

- 11001000 00010111 00011110 00000000 200.23.30.0/23



划分地址块

- ISP有一个地址块

- 11001000 00010111 00010000 00000000 200.23.16.0/20

- 机构地址块：

- 11001000 00010111 00010000 00000000 200.23.16.0/23

- 11001000 00010111 00010010 00000000 200.23.18.0/23

- 11001000 00010111 00010100 00000000 200.23.20.0/23

-

- 11001000 00010111 00011110 00000000 200.23.30.0/23

路由广播

机构0

200.23.16.0/23

机构1

200.23.18.0/23

机构2

200.23.20.0/23

⋮

机构7

200.23.30.0/23

我的ISP

别人的ISP

向我发送地址是
200.23.16.0/23的分组
向我发送地址是
200.23.18.0/23的分组
向我发送地址是
200.23.20.0/23的分组
.....

向我发送地址是
200.23.30.0/23的分组



Internet

路由广播

机构0

200.23.16.0/23

机构1

200.23.18.0/23

机构2

200.23.20.0/23

⋮

机构7

200.23.30.0/23

我的ISP

别人的ISP

| 网络地址 | 接口 |
|----------------|----|
| 200.23.16.0/23 | 0 |
| 200.23.18.0/23 | 0 |
| 200.23.20.0/23 | 0 |
| 200.23.22.0/23 | 0 |
| 200.23.24.0/23 | 0 |
| 200.23.26.0/23 | 0 |
| 200.23.28.0/23 | 0 |
| 200.23.30.0/23 | 0 |
| 其他 | 3 |



Internet

地址聚合

■ 聚合路由信息

- 缩小路由表
- 减少路由信息的广播代价

| 网络地址 | 接口 |
|-----------------------|----|
| 200.23.16.0/20 | 0 |
| 其他 | 3 |

机构0

200.23.16.0/23

机构1

200.23.18.0/23

机构2

200.23.20.0/23

⋮

机构7

200.23.30.0/23

我的ISP

向我发送地址是
200.23.16.0/20的分组

别人的ISP

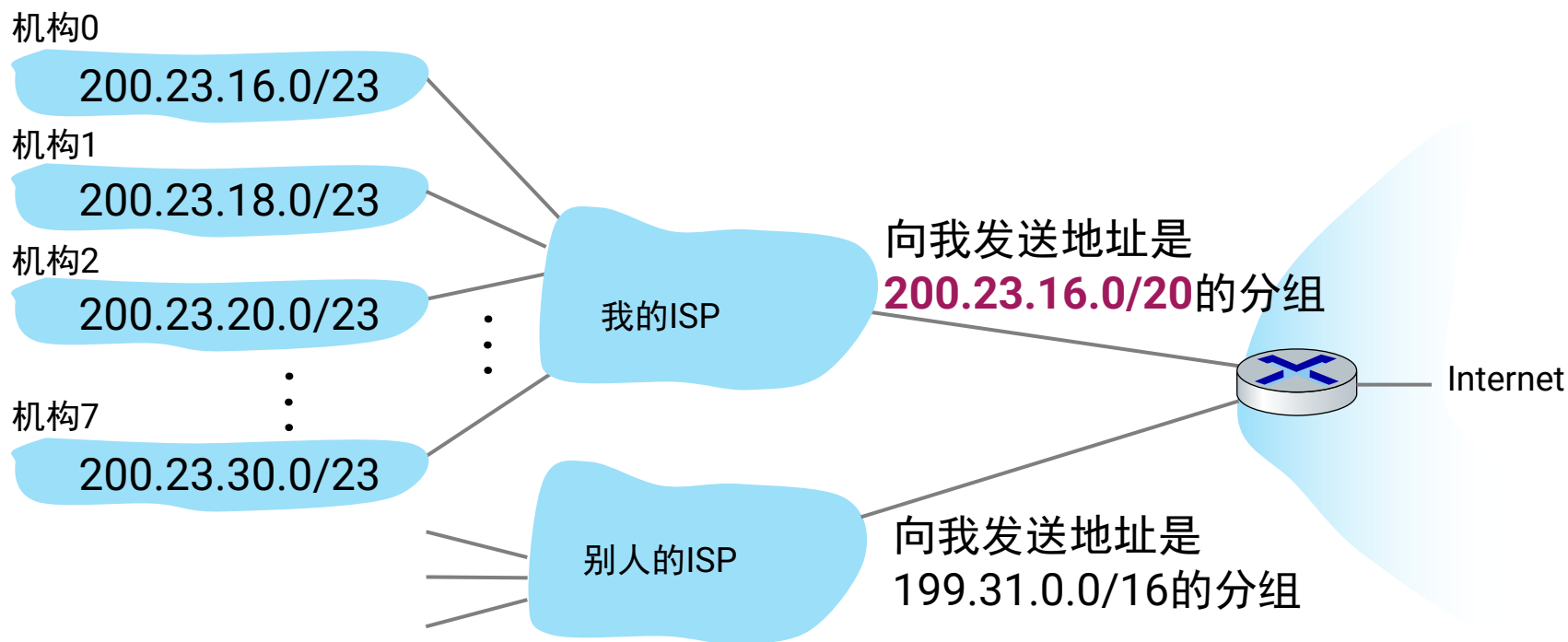


Internet

地址聚合

■ 聚合路由信息

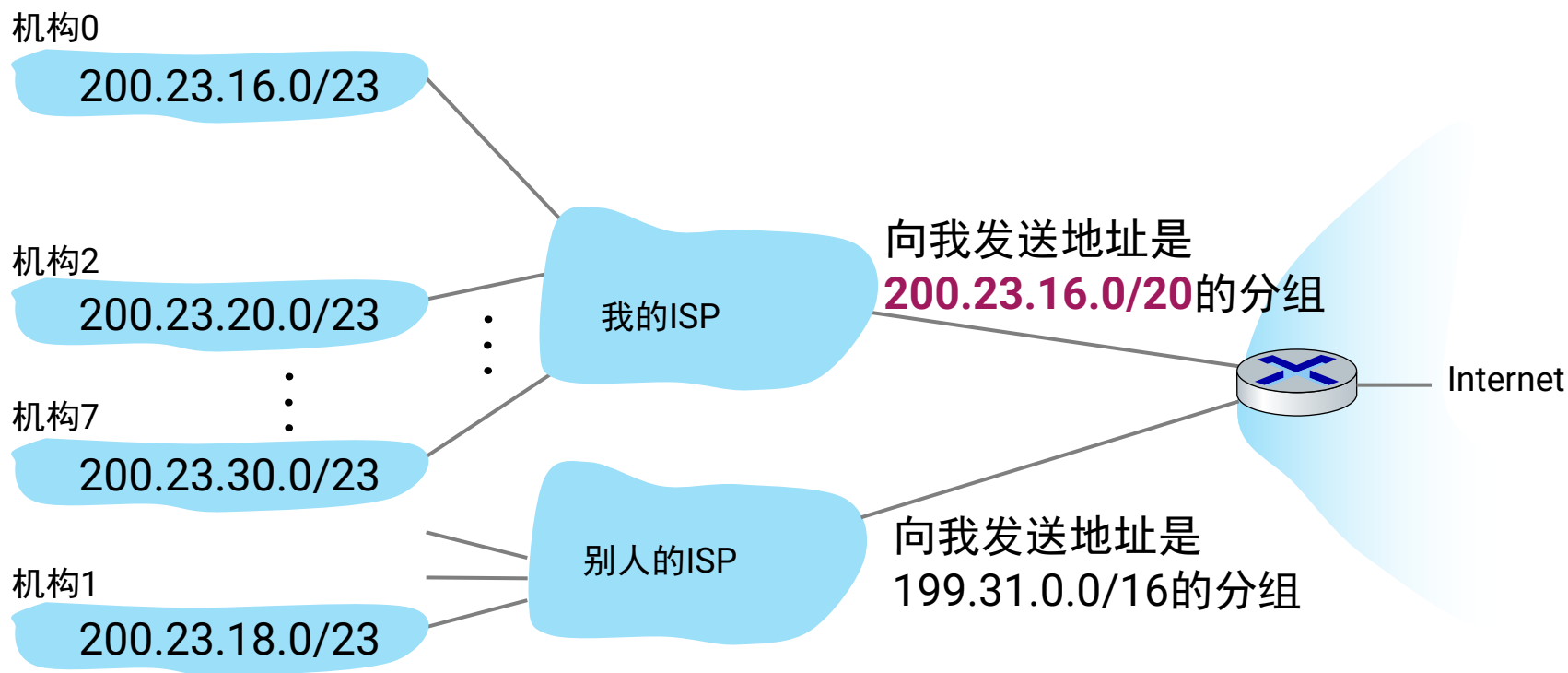
| 网络地址 | 接口 |
|-----------------------|----|
| 200.23.16.0/20 | 0 |
| 199.31.0.0/16 | 2 |
| 其他 | 3 |



地址聚合

- 聚合路由信息
- 需要更改转发表吗？

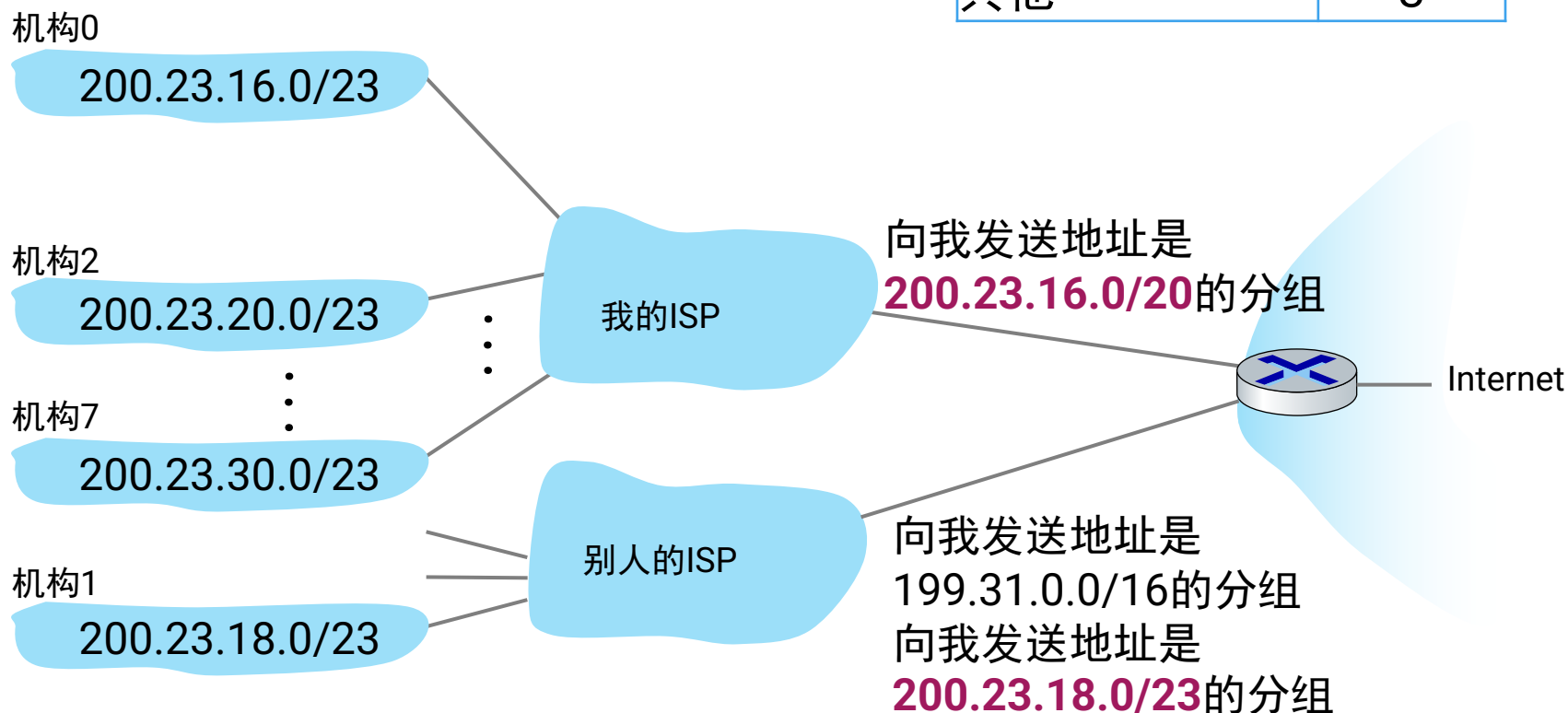
| 网络地址 | 接口 |
|-----------------------|----|
| 200.23.16.0/20 | 0 |
| 199.31.0.0/16 | 2 |
| 其他 | 3 |



地址聚合

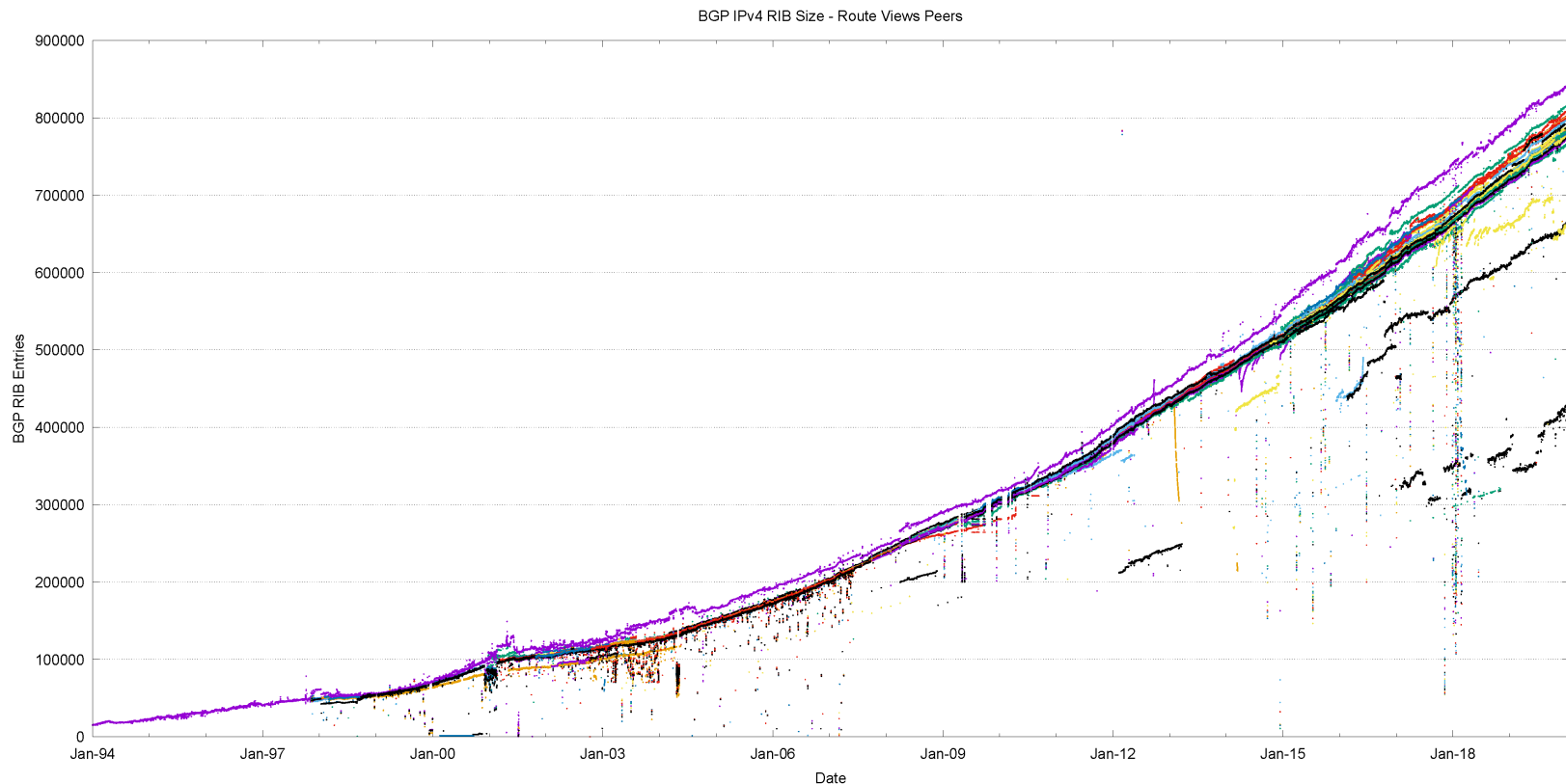
- 聚合路由信息
- 如何更改转发表?

| 网络地址 | 接口 |
|-----------------------|----|
| 200.23.16.0/20 | 0 |
| 200.23.18.0/23 | 2 |
| 199.31.0.0/16 | 2 |
| 其他 | 3 |



路由聚合

■ 真的减少了路由表规模吗？



<https://www.potaroo.net/ispcol/2020-01/bgp2019.pdf>





第四章知识点汇总

- 掌握划分地址块的方法
- 理解地址聚合



习题

- 【2018 年考研 38 题】 某路由表中有转发接口相同的 4 条路由表项，其目的网络地址分别为 35.230.32.0/21、35.230.40.0/21、35.230.48.0/21 和 35.230.56.0/21，将该 4 条路由聚合后的目的网络地址为
 - A. 35.230.0.0/19
 - B. 35.230.0.0/20
 - C. 35.230.32.0/19
 - D. 35.230.32.0/20



数据平面讲解内容

- 网络层概述与功能
- 路由器体系结构
- IP数据报格式与分片
- IP地址
- 转发
- 划分地址块与地址聚合
- **DHCP协议**
- 网络地址转换NAT
- IPv6



如何获取IP地址?

- 主机接口的IP地址可以手工配置
- 也可由DHCP(Dynamic Host Configuration Protocol)协议自动获取



DHCP协议

■ 四个步骤：

- 主机广播 “DHCP Discover” 报文
- DHCP服务器以广播方式响应 “DHCP Offer” 报文
- 主机广播 “DHCP Request” 报文
- DHCP服务器以单播方式响应 “DHCP ACK” 报文

DHCP协议

DHCP服务器：
192.168.2.1

客户端



DHCP

源: 0.0.0.0, 68
目的: 255.255.255.255, 67
Your IP address: 0.0.0.0
Transaction ID: 0x53fbefb5

广播：有没有DHCP服务器

DHCP Offer

源: 192.168.2.1, 67
目的: 255.255.255.255, 68
Your IP address: 192.168.2.178
Transaction ID: 0x53fbefb5
租用时间: 1 day

广播：我是DHCP服务器，
你可以用地址
X.X.X.X

DHCP

源: 0.0.0.0, 68
目的: 255.255.255.255, 67
Your IP address: 0.0.0.0
Transaction ID: 0x53fbefb5
请求地址: 192.168.2.178
DHCP服务器: 192.168.2.1

广播：好啊，我用这个
IP地址

单播：好，这个
地址给你用

DHCP

源: 192.168.2.1, 67
目的: 192.168.2.178, 68
Your IP address: 192.168.2.178
Transaction ID: 0x53fbefb5
租用时间: 1 day

DHCP协议

dhcpcapng

文件(F) 编辑(E) 视图(V) 跳转(G) 捕获(C) 分析(A) 统计(S) 电话(Y) 无线(W) 工具(T) 帮助(H)

应用显示过滤器 ... <Ctrl-/> + http-ieee

| Time | Source | Source Port | Destination | Destination Port | Protocol | Info |
|-----------|---------------|-------------|-----------------|------------------|----------|---|
| 0.000000 | 192.168.2.178 | 68 | 192.168.2.1 | 67 | DHCP | DHCP Release - Transaction ID 0xfdb5b2be |
| 51.627821 | 0.0.0.0 | 68 | 255.255.255.255 | 67 | DHCP | DHCP Discover - Transaction ID 0x53fbefb5 |
| 54.924974 | 192.168.2.1 | 67 | 255.255.255.255 | 68 | DHCP | DHCP Offer - Transaction ID 0x53fbefb5 |
| 54.926967 | 0.0.0.0 | 68 | 255.255.255.255 | 67 | DHCP | DHCP Request - Transaction ID 0x53fbefb5 |
| 54.931202 | 192.168.2.1 | 67 | 192.168.2.178 | 68 | DHCP | DHCP ACK - Transaction ID 0x53fbefb5 |

< >

> Frame 5: 354 bytes on wire (2832 bits), 354 bytes captured (2832 bits) on interface \Device\NPF_{...}

> Ethernet II, Src: ASUSTekC_..., Dst: ASUSTekC_...

> Internet Protocol Version 4, Src: 192.168.2.1, Dst: 192.168.2.178

> User Datagram Protocol, Src Port: 67, Dst Port: 68

> Dynamic Host Configuration Protocol (ACK)

< >

Dynamic Host Configuration Protocol (dhcp), 312 byte(s) | 分组: 5 · 已显示: 5 (100.0%) | 配置: Default

执行ipconfig /release 释放IP地址

执行ipconfig /renew
利用DHCP获取IP地址

传输层用UDP



DHCP协议

- 除了获得IP地址，主机还通过DHCP服务器获得：
 - 网络掩码
 - 网络广播地址
 - DNS服务器
 - 默认网关

```
✓ Option: (1) Subnet Mask (255.255.255.0)
  Length: 4
  Subnet Mask: 255.255.255.0
✓ Option: (28) Broadcast Address (192.168.2.255)
  Length: 4
  Broadcast Address: 192.168.2.255
✓ Option: (6) Domain Name Server
  Length: 4
  Domain Name Server: 192.168.2.1
> Option: (81) Client Fully Qualified Domain Name
> Option: (252) Private/Proxy autodiscovery
✓ Option: (3) Router
  Length: 4
  Router: 192.168.2.1
```



第四章知识点汇总

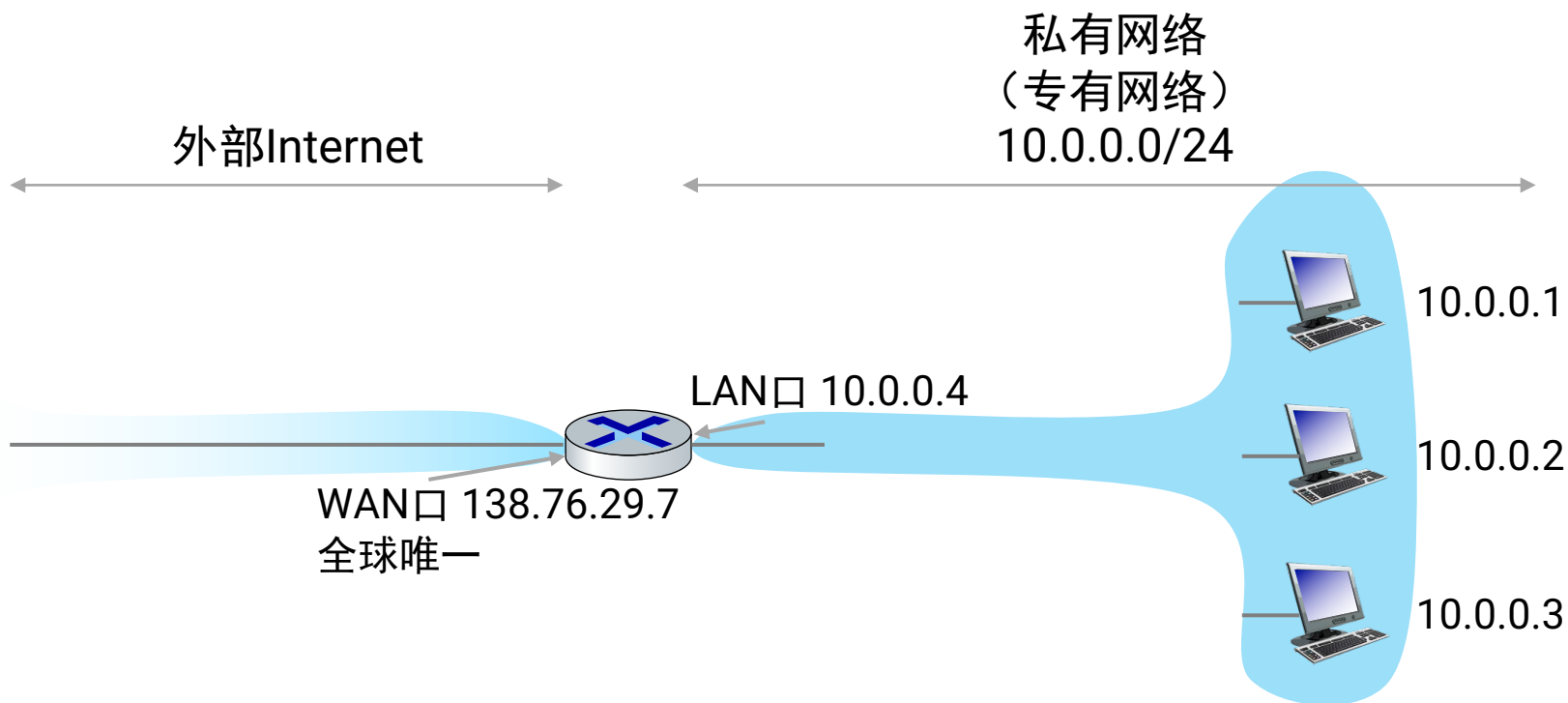
- 理解获取IP地址的方法
- 理解DHCP协议的原理



数据平面讲解内容

- 网络层概述与功能
- 路由器体系结构
- IP数据报格式与分片
- IP地址
- 转发
- 划分地址块与地址聚合
- DHCP协议
- **网络地址转换NAT**
- IPv6

网络地址转换NAT



除了家庭网络，ISP也可能会用NAT



私有网络

- 为私有Internet预留的三块地址空间[RFC1918, 1996]:
- 10.0.0.0 – 10.255.255.255 (10.0.0.0/8)
- 172.16.0.0 – 172.31.255.255 (172.16.0.0/12)
- 192.168.0.0 – 192.168.255.255 (192.168.0.0/16)



网络地址转换NAT

■ Network Address Translation NAT路由器实现

| NAT转换表 | |
|-------------------|----------------|
| WAN端地址 | LAN端地址 |
| 138.76.29.7, 5001 | 10.0.0.1, 3345 |

1. 替换源地址：用端口号区分私有网络内部的设备

(LAN IP地址, 端口号) → (WAN IP地址, 新端口号)

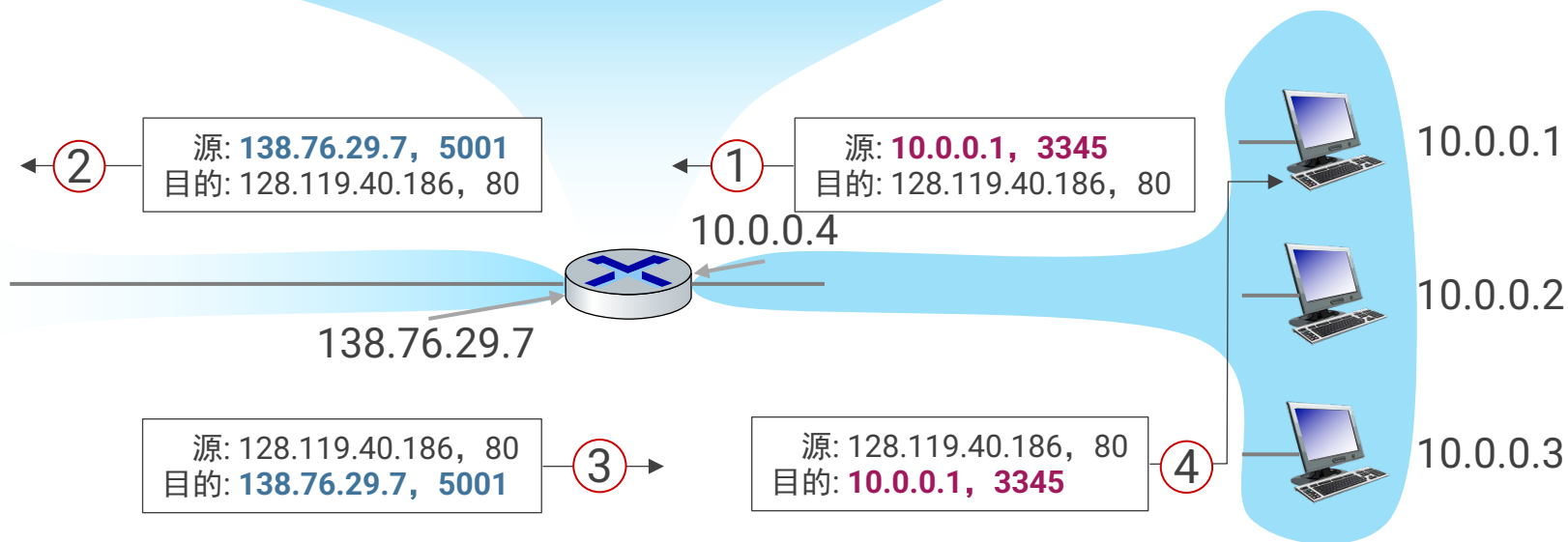
2. 在NAT转换表中添加条目

3. 替换目的地址

(LAN IP地址, 端口号) ← (WAN IP地址, 新端口号)

网络地址转换NAT

| NAT转换表 | |
|-------------------|----------------|
| WAN端地址 | LAN端地址 |
| 138.76.29.7, 5001 | 10.0.0.1, 3345 |





网络地址转换NAT

- 好处：
- 本地网络只需向ISP申请一个IP地址
- 可以随意改变本地网络设备的地址
- 可以更换ISP，而无需改变本地网络设备的地址
- 本地网络内部的设备对外不可见，安全



网络地址转换NAT

- 争议：
- 路由器（网络层）修改端口号（传输层），违反了协议分层原则
- 地址短缺应该用IPv6来解决

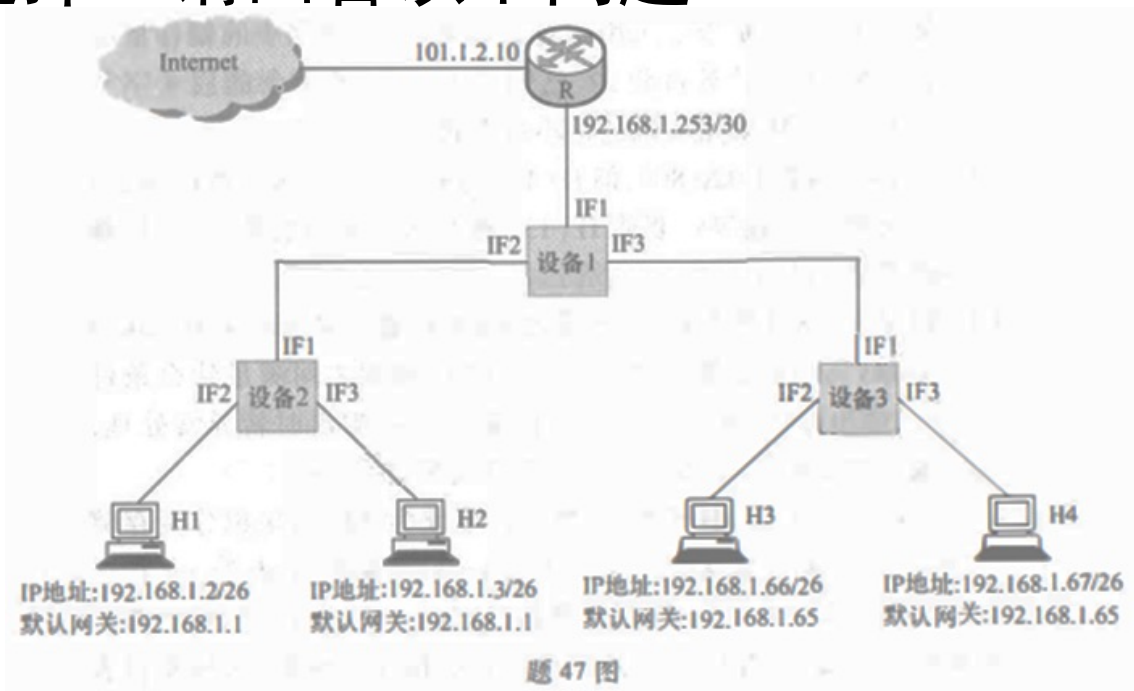


第四章知识点汇总

- 理解NAT的原理
- 了解NAT的好处和存在的争议

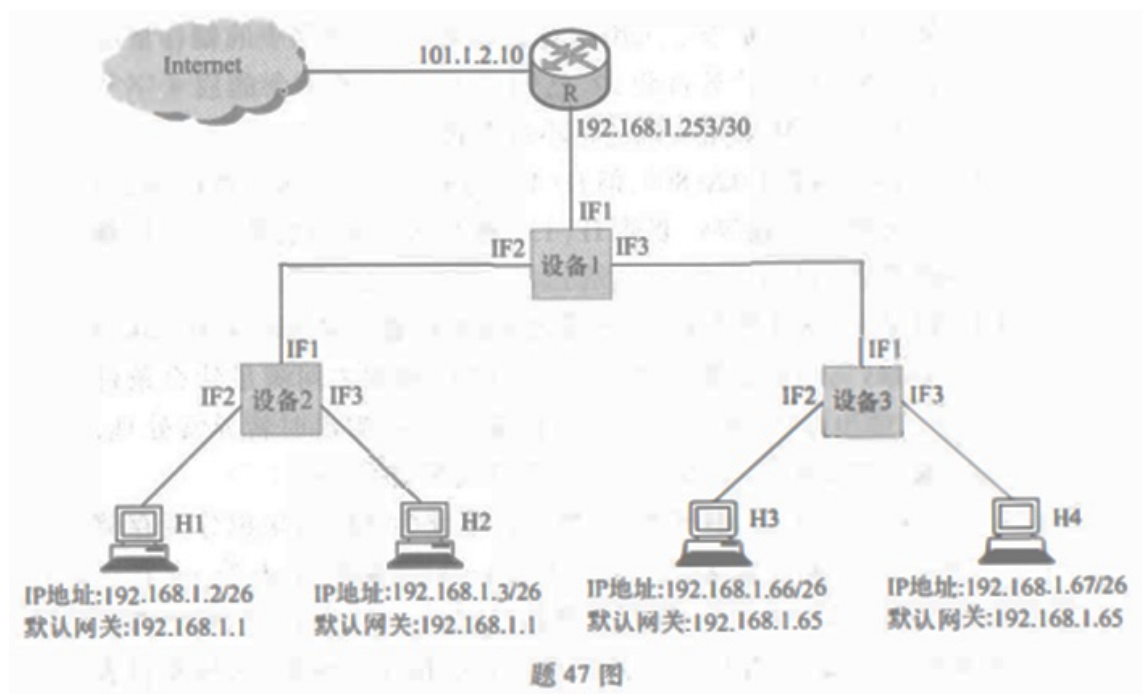
习题

- 【2019年考研47题】某网络拓扑如图所示，其中R为路由器，主机H1-H4的IP地址配置以及R的各接口IP地址配置如图所示。现在若干台以太网交换机（无VLAN功能）和路由器两类网络互连设备可供选择。请回答以下问题：



习题

- (3) 为确保主机H1-H4能够访问Internet, R需要什么服务?





数据平面讲解内容

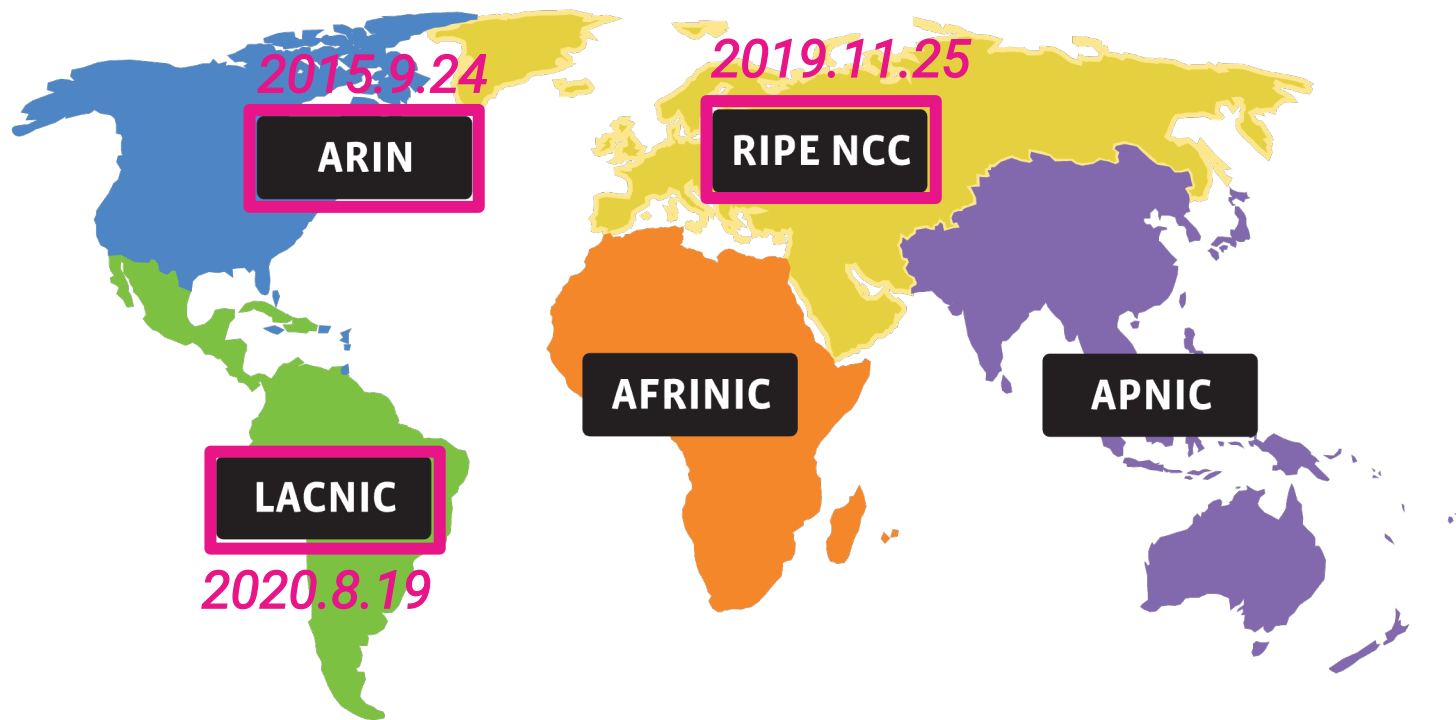
- 网络层概述与功能
- 路由器体系结构
- IP数据报格式与分片
- IP地址
- 转发
- 划分地址块与地址聚合
- DHCP协议
- 网络地址转换NAT
- **IPv6**



IPv6地址[RFC4291, 2006]

- 为了解决IPv4地址短缺的问题
- IPv6将地址长度由32位增加到了128位

IP地址耗尽



<https://www.iana.org/numbers>

<https://www.arin.net/resources/guide/ipv4/>

<https://www.ripe.net/publications/news/about-ripe-ncc-and-ripe/the-ripe-ncc-has-run-out-of-ipv4-addresses>

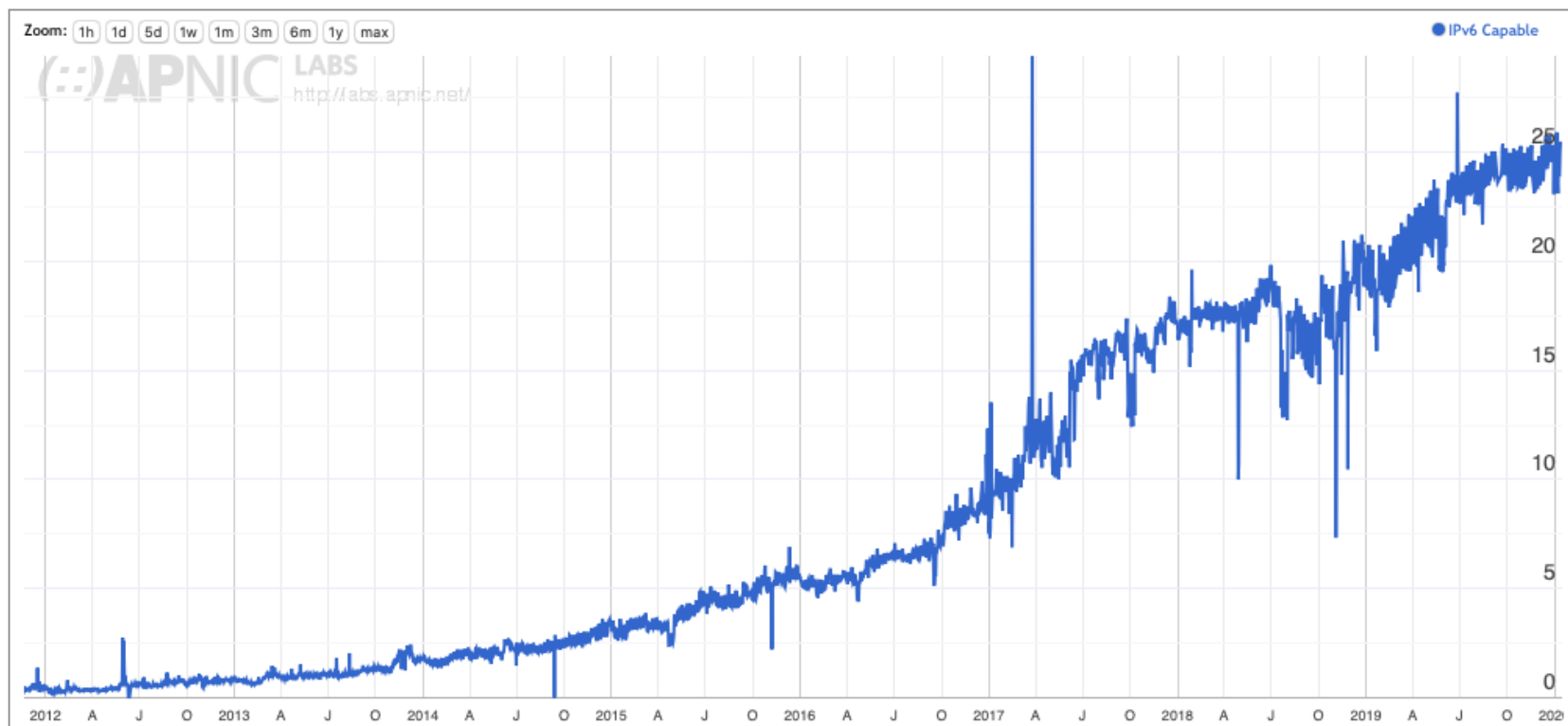
<https://www.apnic.net/manage-ip/ipv4-exhaustion/>

<https://afrinic.net/exhaustion>

<https://www.lacnic.net/1039/2/lacnic/phases-of-ipv4-exhaustion>

IPv6

- 纵轴单位%
- 截止2019年尾，只有25%的用户群支持IPv6.



<https://www.potaroo.net/ispcol/2020-01/bgp2019.pdf>



IPv6地址[RFC4291, 2006]

- IPv6地址表示方法
- 16位一段，分成8段，冒号分隔
- 0010000000000001:0000110110111000:0000000000000000:0000000000
000000:0000000000000001:0000000000000000:0000000000000000:0000
000000000001
- 每段分别用16进制表示
- 2001:db8:0:0:1:0:0:1
- 对于连续的几段零，可以用双冒号代替，且双冒号只能用一次
- 2001:db8::1:0:0:1



IPv6地址[RFC4291, 2006]

- 可以有很多种别的写法：

- 2001:db8:0:0:1:0:0:1

- 2001:db8::1:0:0:1

- 2001:db8::0:1:0:0:1

- 2001:0db8::1:0:0:1

- 2001:db8:0:0:1::1

- 2001:db8:0000:0:1::1

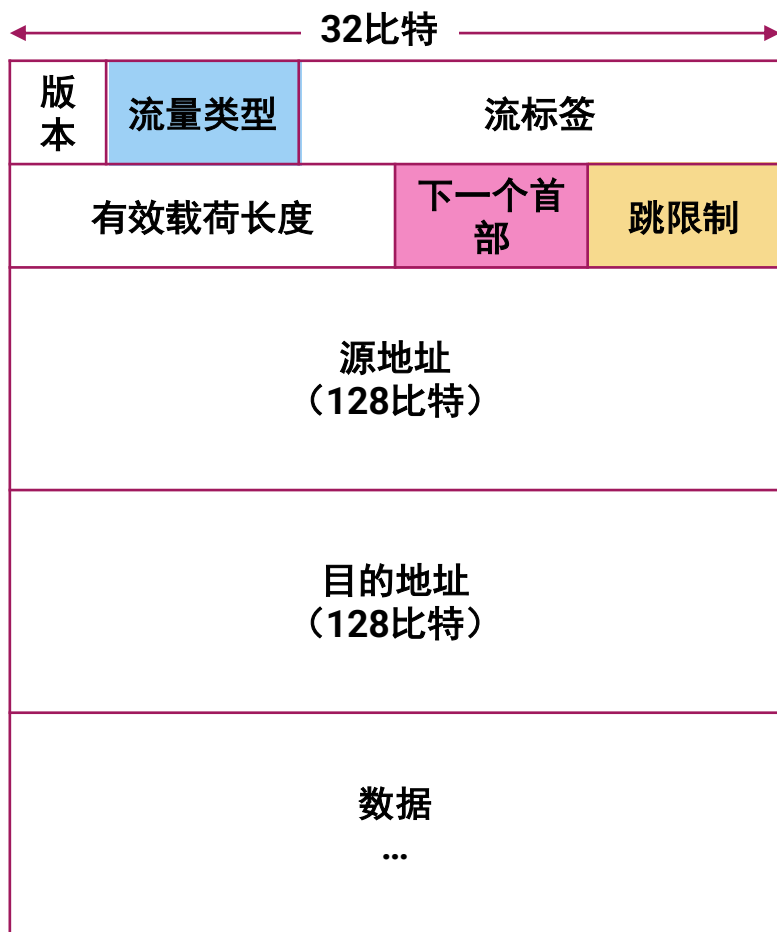
- 2001:DB8:0:0:1::1

-

IPv6地址压缩规则[RFC5952, 2010]

- 一个16位段的起始零必须压缩
 - 错误示范：2001:0db8::1:0:0:1
- “::”的使用
 - 必须用于最大程度的压缩
 - 错误示范：2001:db8::0:1:0:0:1
 - 不能用于一个16位的全零段
 - 如果有多个可以替换的位置，选择可替换段数最多的；如果都一样，就选择第一个
 - 错误示范：2001:db8:0:0:1::1
- 必须使用小写字母
 - 错误示范：2001:DB8::1:0:0:1

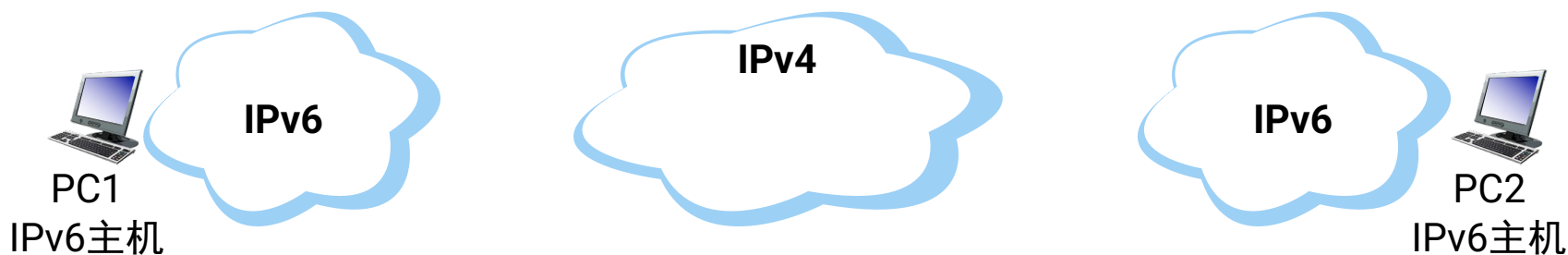
IPv6数据报格式



- 40字节首部
- **流量类型**：与IPv4服务类型相似
- **下一个首部**：类似于IPv4的上层协议
- **跳限制**：类似于IPv4的TTL
- 取消了IP分片
- 取消了首部校验和

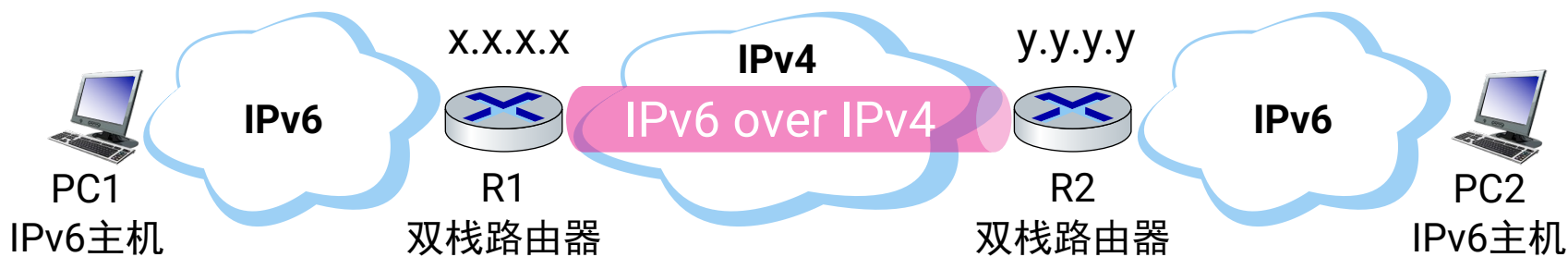
IPv6 over IPv4

- IPv6逐步过渡中，两个IPv6网络之间只有IPv4网络，两个IPv6网络能否利用已有的IPv4网络实现互联呢？



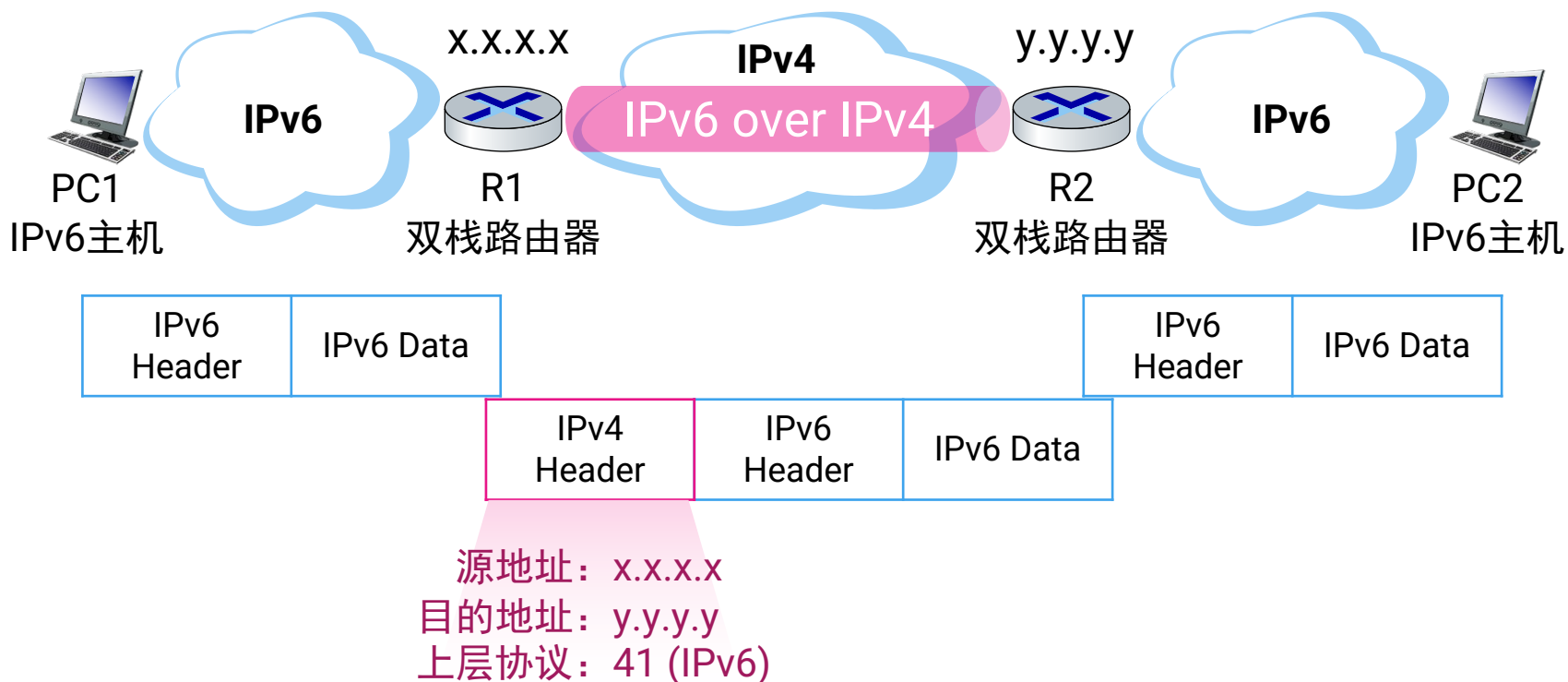
IPv6 over IPv4

- 隧道技术
- 将IPv6数据报封装成IPv4数据报传输



IPv6 over IPv4

- 隧道技术
- 将IPv6数据报封装成IPv4数据报传输





第四章知识点汇总

- 了解IPv6地址的表示方法
- 了解IPv6数据报格式
- 理解IPv6 over IPv4隧道技术的原理

*Asks the Possible to the Impossible,
“Where is your dwelling-place?”
“In the dreams of the impotent,” comes the answer.*

可能问不可能道：
“你住在什么地方呢？”
它回答道：“在那无能为力者的梦境里。”

——*Tagore*