



第4章

网络层

计算机网络体系结构

OSI 的七层协议体系结构



(a)

TCP/IP 的四层协议体系结构



(b)

五层协议的体系结构



(c)

4.1	网络层的几个重要概念
4.2	网际协议 IP
4.3	IP 层转发分组的过程
4.4	网际控制报文协议 ICMP
4.5	IPv6
4.6	互联网的路由选择协议
4.7	IP 多播
4.8	虚拟专用网 VPN 和网络地址转换 NAT
4.9	多协议标记交换 MPLS
4.10	软件定义网络 SDN 简介

4.1

网络层的几个重要概念

4.1.1

网络层提供的两种服务

4.1.2

网络层的两个层面

4.1.1 网络层提供的两种服务

争论:

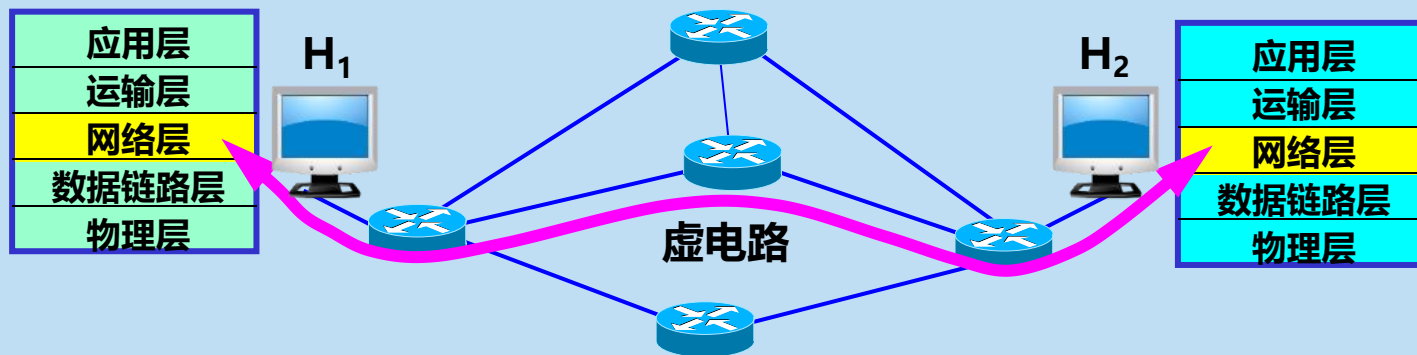
- 网络层应该向运输层提供怎样的服务？**面向连接**还是**无连接**？
- 在计算机通信中，**可靠交付**应当由谁来负责？是**网络**还是**端系统**？

2 种观点:

- 面向连接的可靠交付。
- 无连接的、尽最大努力交付的数据报服务，不提供服务质量的承诺。

一种观点：让网络负责可靠交付

- 通信之前先建立**虚电路 VC** (Virtual Circuit) (即连接)，以保证双方通信所需的一切网络资源。
- 使用**可靠传输**的网络协议，可使所发送的分组无差错按序到达终点，不丢失、不重复。



H₁ 发送给 H₂ 的所有分组都沿着同一条虚电路传送

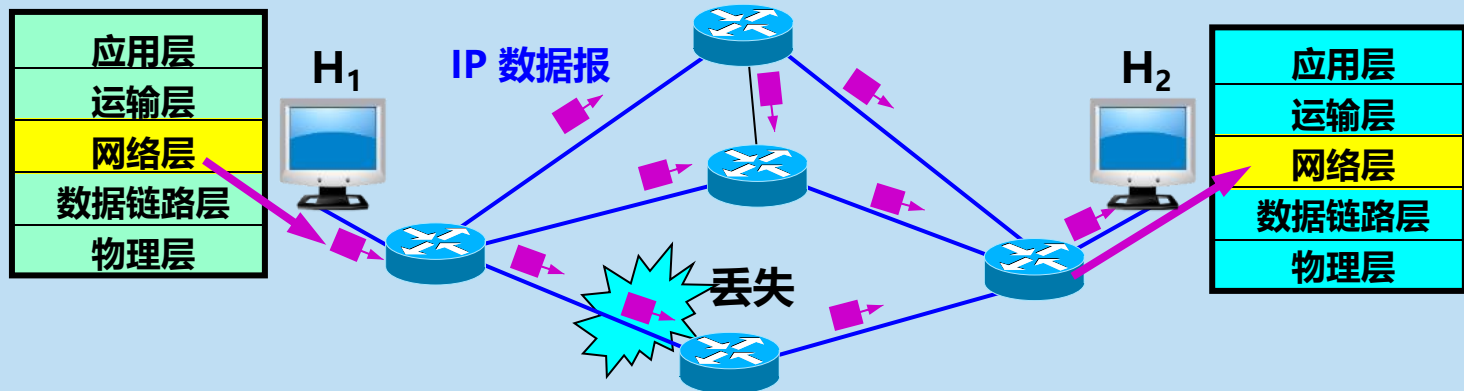
虚电路只是一条**逻辑上的连接**，分组都沿着这条逻辑连接按照存储转发方式传送，**并不是真正建立了一条物理连接**。

另一种观点：网络提供数据报服务

互联网采用的设计思路：

- **网络层**要设计得尽量**简单**，向其上层只提供简单灵活的、**无连接的**、尽最大努力交付的**数据报服务**。
 - ◆ 网络在发送分组时不需要先建立连接。
 - ◆ 每一个分组（即 IP 数据报）独立发送，与其前后的分组无关（不进行编号）。
 - ◆ 网络层不提供服务质量的承诺。即所传送的分组可能出错、丢失、重复和失序（不按序到达终点），也不保证分组传送的时限。
- 由主机中的**运输层**负责可靠的通信。

数据报服务



H₁ 发送给 H₂ 的分组可能沿着不同路径传送

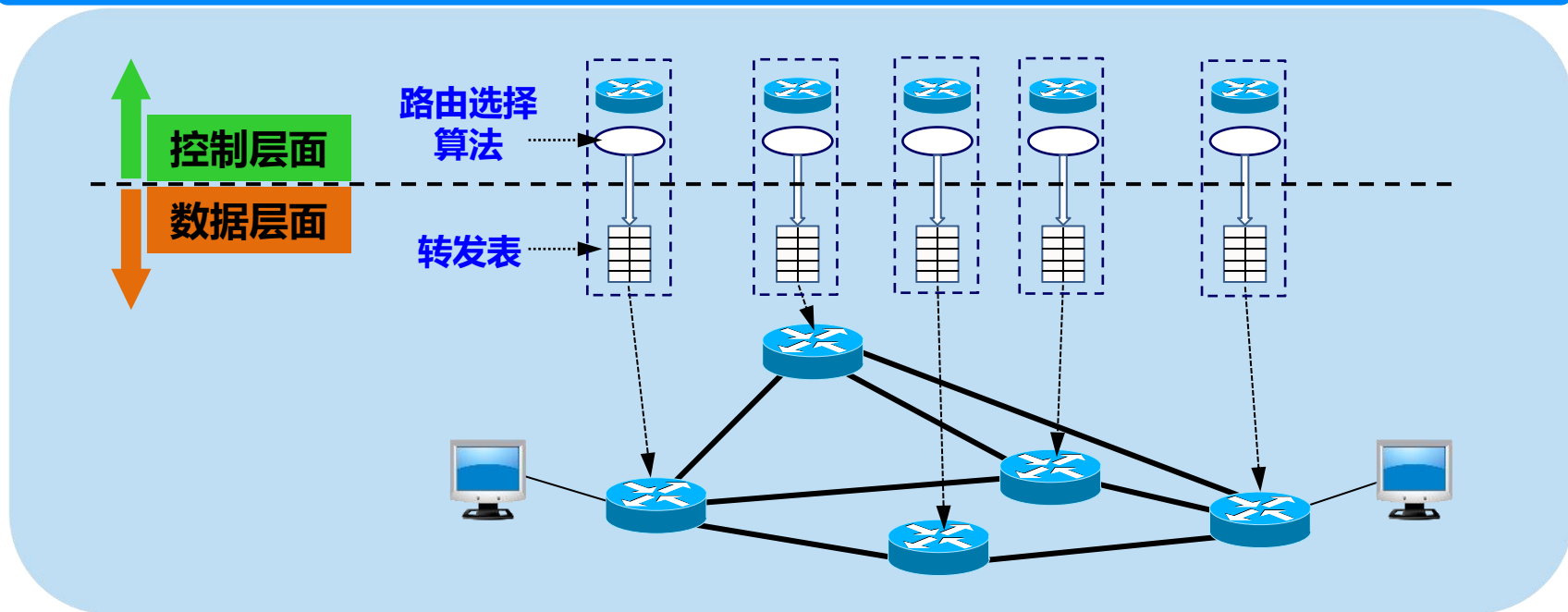
虚电路服务与数据报服务的对比

对比的方面	虚电路服务	数据报服务
思路	可靠通信应当由网络来保证	可靠通信应当由用户主机来保证
连接的建立	必须有	不需要
终点地址	仅在连接建立阶段使用，每个分组使用短的虚电路号	每个分组都有终点的完整地址
分组的转发	属于同一条虚电路的分组均按照同一路由进行转发	每个分组独立选择路由进行转发
当结点出故障时	所有通过出故障的结点的虚电路均不能工作	出故障的结点可能会丢失分组，一些路由可能会发生变化
分组的顺序	总是按发送顺序到达终点	到达终点时不一定按发送顺序
端到端的差错处理和流量控制	可以由网络负责，也可以由用户主机负责	由用户主机负责

4.1.2 网络层的两个层面

- 不同网络中的两个主机之间的通信，要经过若干个路由器转发分组来完成。
- 在路由器之间传送的信息有以下 2 大类：
 1. 数据。
 2. 路由信息（为数据传送服务）。

4.1.2 网络层的两个层面



网络层的 2 个层面：数据层面和控制层面

数据层面和控制层面

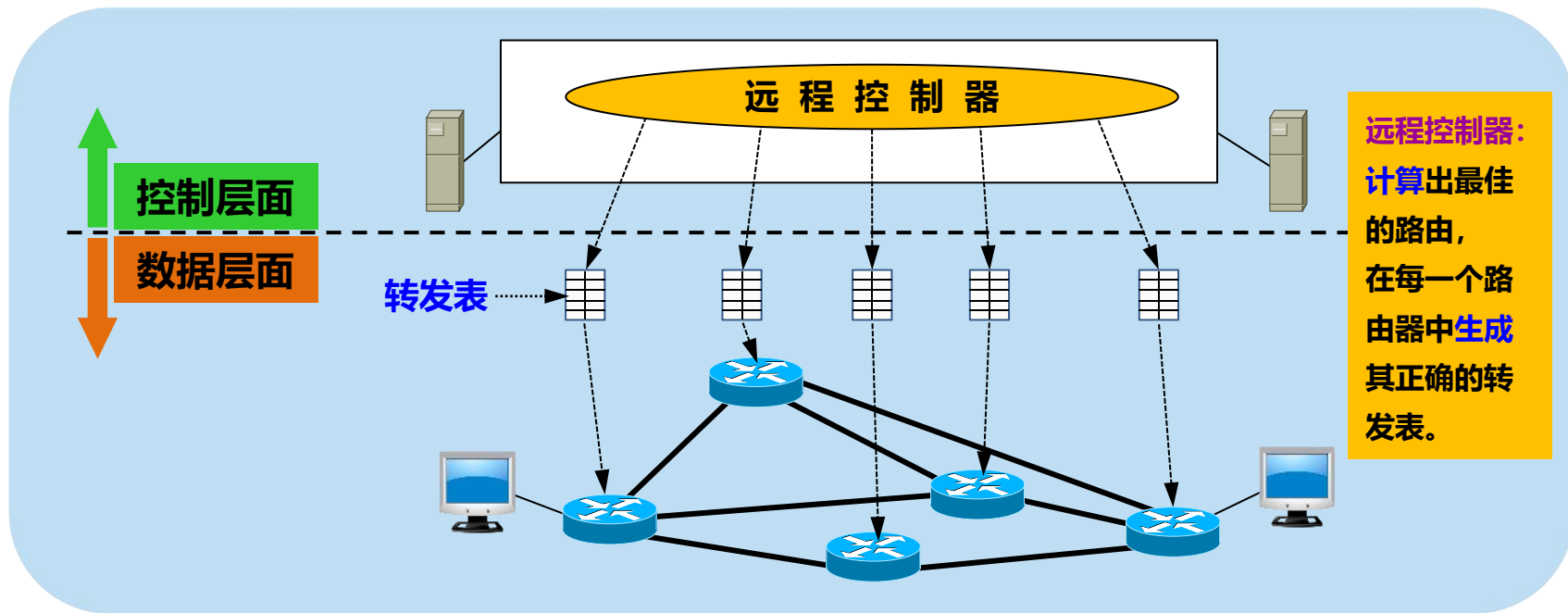
数据层面

- 路由器根据本路由器生成的**转发表**，把收到的分组从查找到的对应接口**转发**出去。
- **独立**工作。
- 采用**硬件**进行转发，快。

控制层面

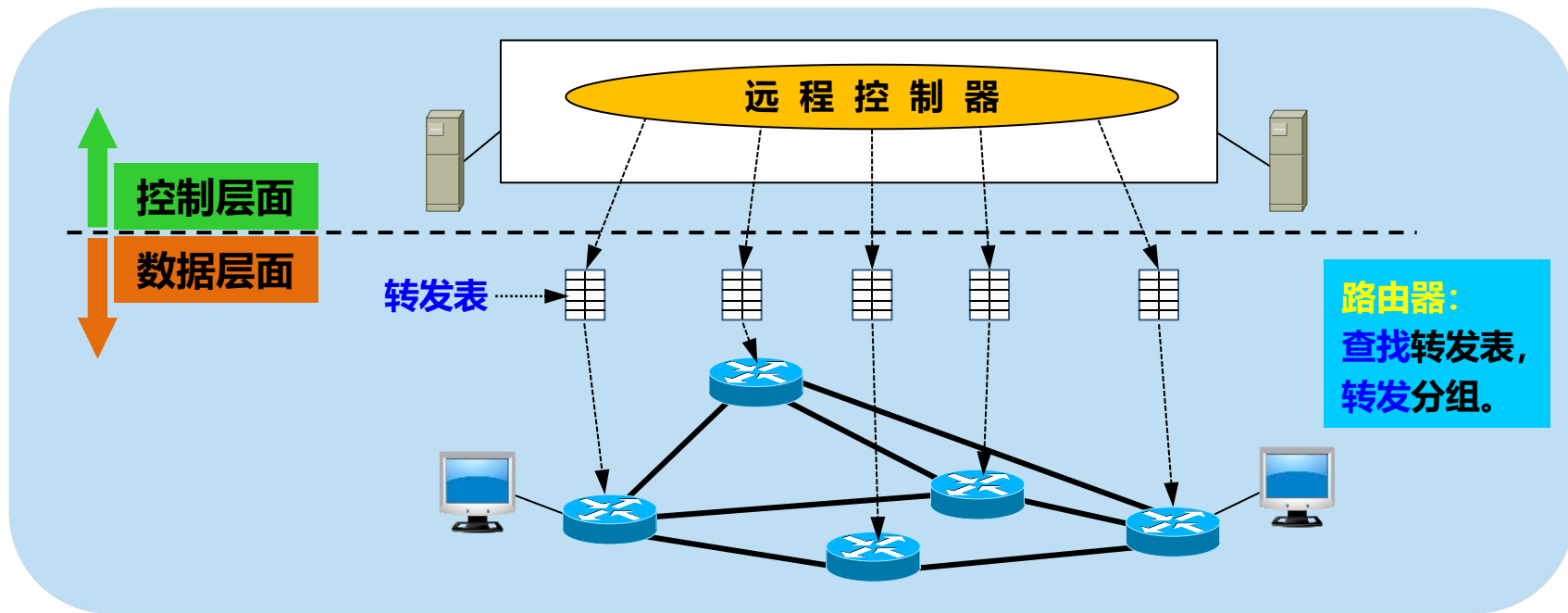
- 根据路由选择协议所用的路由算法**计算路由**，创建出本路由器的**路由表**。
- 许多路由器**协同**动作。
- 采用**软件**计算，慢。

软件定义网络 SDN (Software Defined Network)



软件定义网络 SDN 中的数据层面和控制层面

软件定义网络 SDN (Software Defined Network)



软件定义网络 SDN 中的数据层面和控制层面

4.2

网际协议 IP

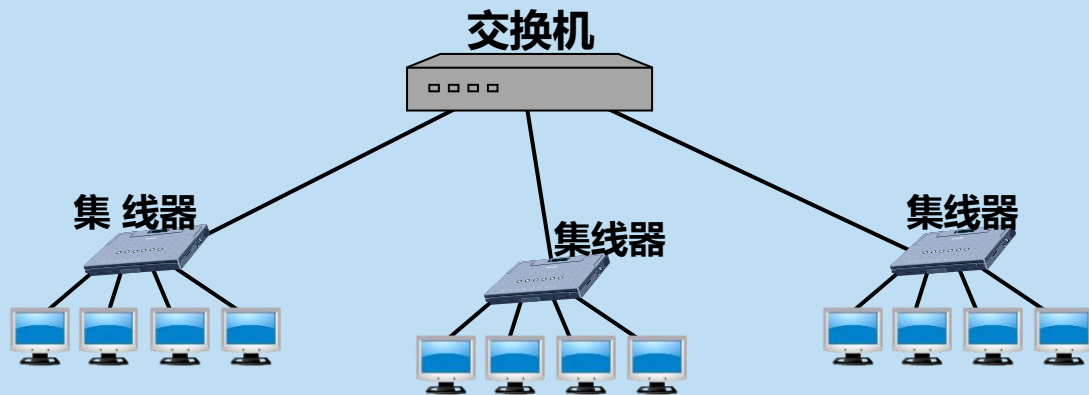
4.2.1

虚拟互连网络

4.2.2

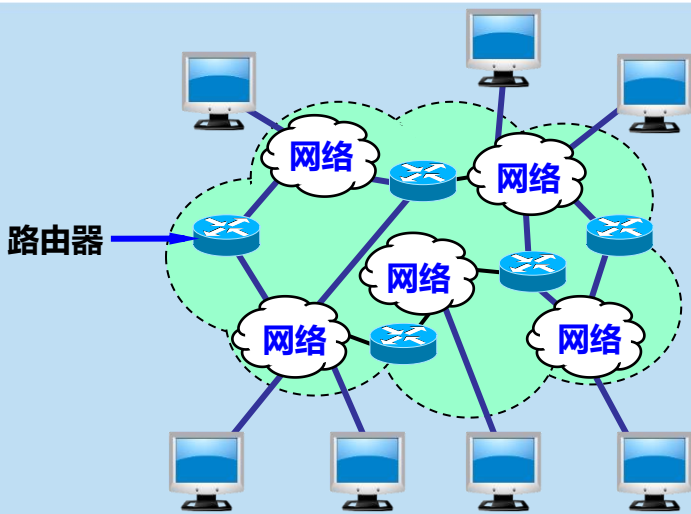
IP 地址

使用转发器或网桥不称为网络互连



转发器、网桥或交换机仅把一个网络扩大了，**仍然是一个网络**

互连网络与虚拟互连网络



(a) 互连网络



(b) 虚拟互连网络

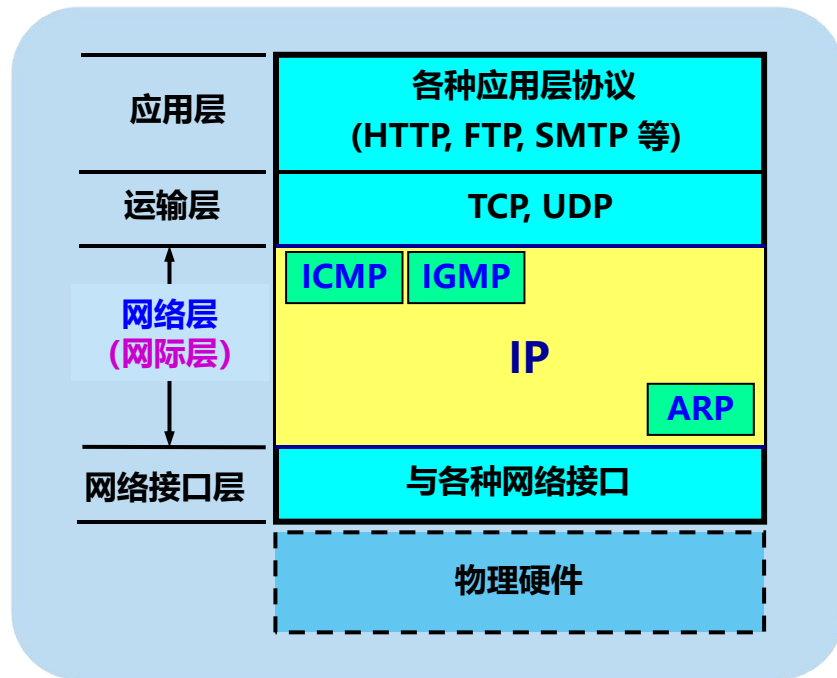
IP 网的概念

- 当互联网上的主机进行通信时，就好像在一个网络上通信一样，看不见互连的**各具体的网络异构细节**。

4.2 网际协议 IP

与网际协议 IPv4 配套的 3 个协议:

1. 地址解析协议 **ARP** (Address Resolution Protocol)
2. 网际控制报文协议 **ICMP** (Internet Control Message Protocol)
3. 网际组管理协议 **IGMP** (Internet Group Management Protocol)



4.2.1 虚拟互连网络

实现网络互连、互通时需要解决许多问题，如以下“不同”：

寻址方案

最大分组长度

网络接入机制

超时控制

差错恢复方法

状态报告方法

路由选择技术

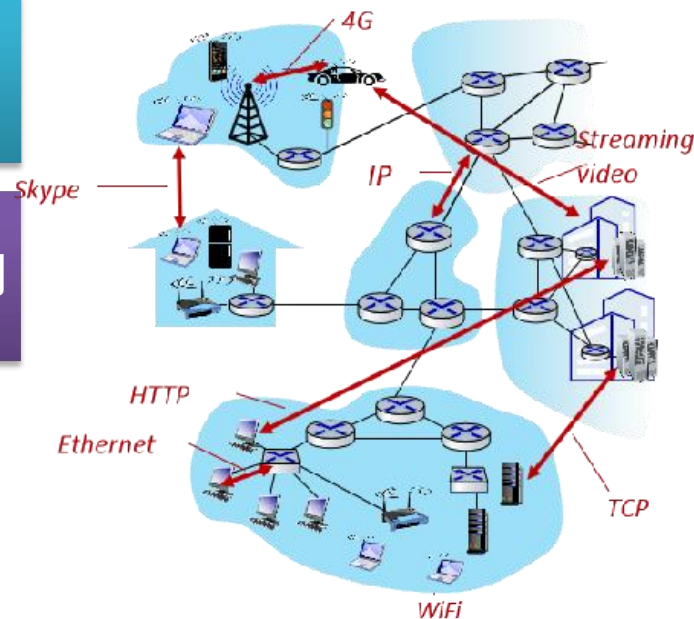
用户接入控制

服务

管理与
控制方式

.....

如何将异构的网络互相连接起来？



实现异构网络的互连互通方法，哪种好？



A. 都使用相同的网络？

- 不能满足不同用户需要。没有一种单一的网络能够适应所有用户的需求。
- 不适应技术发展



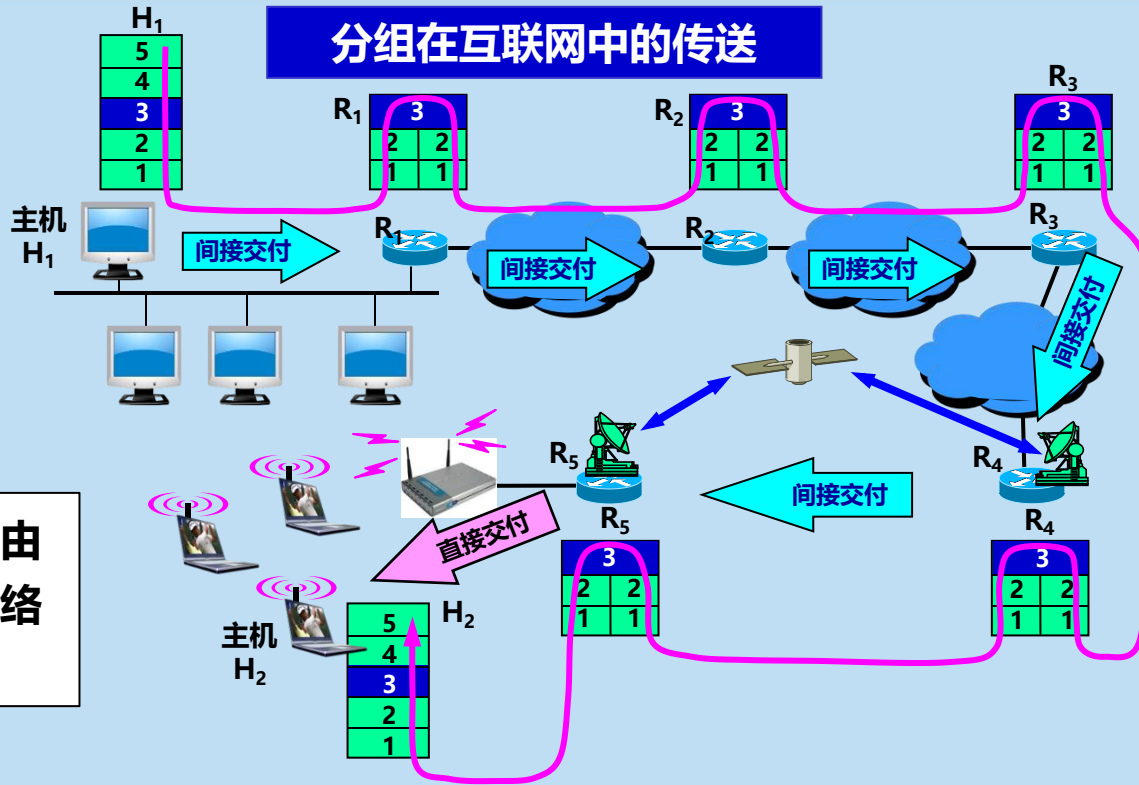
B. 使用中间设备？

- 可以满足不同需求
- 实用

使用中间设备进行互连

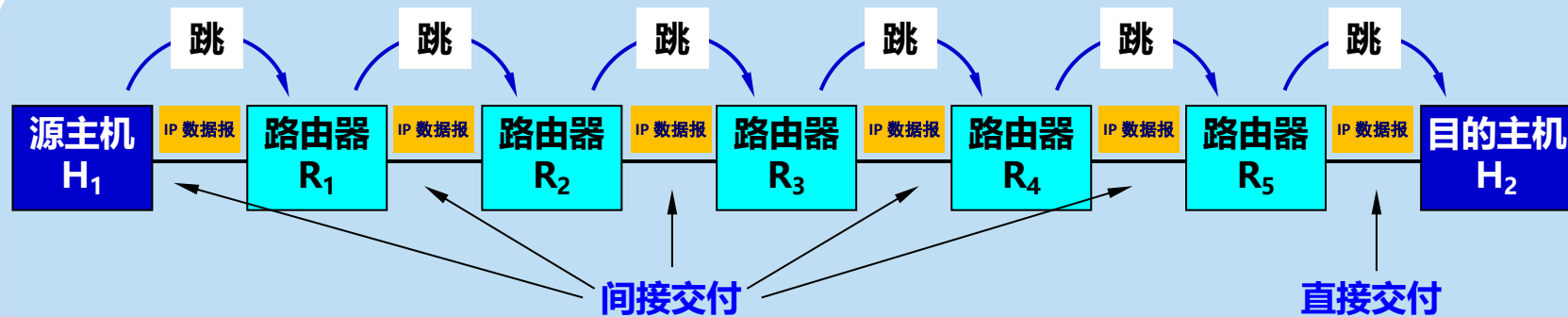
层	中间设备
运输层及以上	网关 (gateway)
网络层	路由器 (router)
数据链路层	网桥或桥接器 (bridge), 交换机 (switch)
物理层	转发器 (repeater)

分组在互联网中的传送



互联网可以由
多种异构网络
互连组成。

分组传输路径



4.2.2 IP 地址

- 在 TCP/IP 体系中，IP 地址是一个最基本的概念。
- 没有IP地址，就无法和网上的其他设备进行通信。
- 本部分重点：
 1. IP 地址及其表示方法
 2. 分类的 IP 地址
 3. 无分类编址 CIDR
 4. IP 地址的特点

1. IP 地址及其表示方法

IP 地址: **32 位**二进制代码

10000000000010110000001100011111

分为每 8 位为一组

将每 8 位的二进制数
转换为十进制数

10000000 00001011 00000011 00011111

128

11

3

31

采用**点分十进制**记法

128.11.3.31

互联网上的每台主机（或路由器）的**每个接口**分配一个在**全世界唯一的 IP 地址**。
由**互联网名字和数字分配机构** ICANN (Internet Corporation for Assigned Names and Numbers) 进行分配。

点分十进制记法举例

32 位二进制数	等价的 点分十进制数
10000001 00110100 00000110 00000000	129.52.6.0
11000000 00000101 00110000 00000011	192.5.48.3
00001010 00000010 00000000 00100101	10.2.0.37
10000000 00001010 00000010 00000011	128.10.2.3
10000000 10000000 11111111 00000000	128.128.255.0

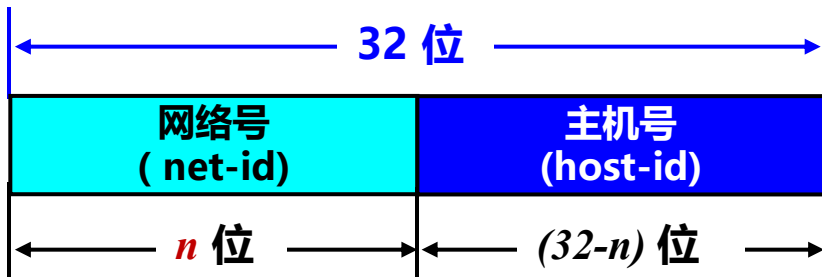
IP 地址采用 2 级结构

2 级结构
2 个字段：网络号和主机号

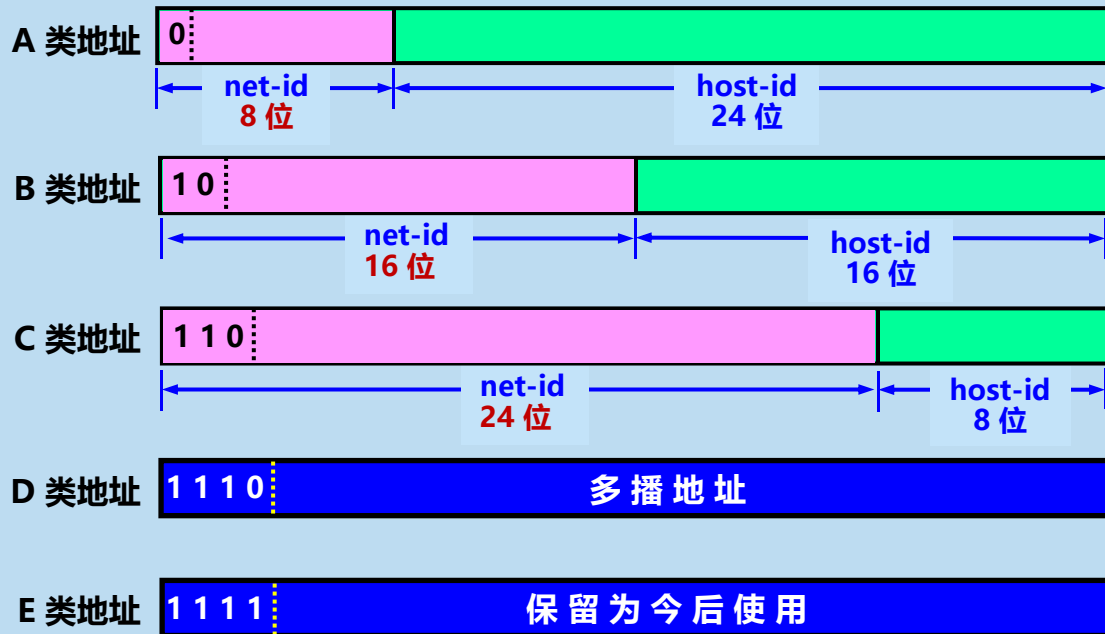
IP 地址 ::= { <网络号>, <主机号> }

IP地址在整个互联网范围内是**唯一**的。
IP 地址指明了连接到某个网络上的一个主机

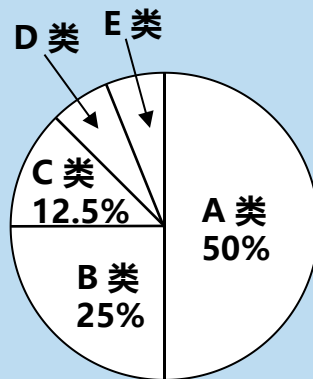
网络号的位数 n 是多少？



2. 分类的 IP 地址



单播地址



各类 IP 地址的指派范围

网络类别	最大可指派的网络数	第一个可指派的网络号	最后一个可指派的网络号	每个网络中最大主机数
A	126 ($2^7 - 2$)	1	126	16777214 ($2^{24} - 2$)
B	16383 ($2^{14} - 1$)	128.1	191.255	65534 ($2^{16} - 2$)
C	2097151 ($2^{21} - 1$)	192.0.1	223.255.255	254 ($2^8 - 2$)

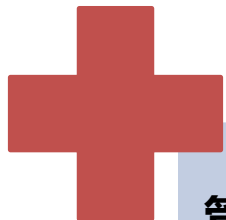
注意：

- A 类网络地址中，网络号 0 和 127 是保留地址，不指派。0 表示“本网络”，127 保留作为本地环回测试地址。
- B 类网络地址中，网络号 128.0 是被 IANA 保留的，不指派。采用无分类编址（CIDR）时可以指派。
- C 类网络地址中，网络号 192.0.0 是被 IANA 保留的，不指派。采用无分类编址（CIDR）时可以指派。
- 指派主机号时，要**扣除**全 0 和全 1。全 0 和全 1 有特殊含义和用途。

一般不使用的特殊的 IP 地址

网络号	主机号	源地址使用	目的地址使用	代表的意思
0	0	可以	不可	在本网络上的本主机 (见 6.6 节 DHCP 协议)
0	X	可以	不可	在本网络上主机号为 X 的主机
全 1	全 1	不可	可以	只在本网络上进行广播 (各路由器均不转发)
Y	全 1	不可	可以	对网络号为 Y 的网络上的所有主机进行广播
127	非全 0 或全 1 的任何数	可以	可以	用于本地软件环回测试

分类的 IP 地址的优点和缺点



管理简单;
使用方便;
转发分组迅速;
划分子网, 灵活地使用。



设计上不合理:
大地址块, **浪费**地址资源;
即使采用划分子网的方法, 也无法解决 IP 地址**枯竭**的问题。

3. 无分类编址 CIDR

- **CIDR** (Classless Inter-Domain Routing) : **无分类域间路由选择。**
- **消除了**传统的 A 类、B 类和 C 类地址以及划分子网的概念, 可以更加有效地分配 IPv4 的地址空间, **但无法解决 IP 地址枯竭的问题。**
- **要点:**
 - (1) 网络前缀
 - (2) 地址块
 - (3) 地址掩码

(1) 网络前缀

2 级结构

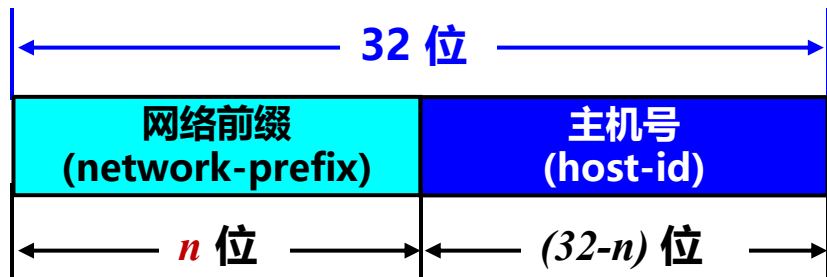
2 个字段：网络前缀和主机号

网络前缀的位数 n 是多少？

最大的区别：

前缀的位数 n 不固定，可以在 0 ~ 32 之间选取任意值。

IP 地址 ::= { <网络前缀>, <主机号> }



CIDR 记法：斜线记法 (slash notation)

a.b.c.d / n ：二进制 IP 地址的前 n 位是网络前缀。

例如：128.14.35.7/20：前 20 位是网络前缀。

(2) 地址块

- CIDR 把**网络前缀都相同的所有连续的** IP 地址组成一个 CIDR 地址块。
- 一个 CIDR 地址块包含的 IP 地址数目，取决于网络前缀的位数。

128.14.32.0/20 组成的地址块 (共 2^{12} 个地址)

最小地址
128.14.32.0

所有地址的
20 位前缀
都是一样的

最大地址
128.14.47.255

10000000 00001110 00100000 00000000
10000000 00001110 00100000 00000001
10000000 00001110 00100000 00000010
10000000 00001110 00100000 00000011
10000000 00001110 00100000 00000100
10000000 00001110 00100000 00000101

...

...

10000000 00001110 00101111 11111011
10000000 00001110 00101111 11111100
10000000 00001110 00101111 11111101
10000000 00001110 00101111 11111110
10000000 00001110 00101111 11111111

可指派的
地址数是
 $2^{12} - 2$ 个

二进制代码表示: 10000000 00001110 0010*

8位无符号二进制整数转十进制数

$$(b_7 b_6 b_5 b_4 b_3 b_2 b_1 b_0)_2 = (b_7 \times 2^7 + b_6 \times 2^6 + b_5 \times 2^5 + b_4 \times 2^4 + b_3 \times 2^3 + b_2 \times 2^2 + b_1 \times 2^1 + b_0 \times 2^0)_{10}$$

8位二进制数的每个位的权值: 128 64 32 16 8 4 2 1

【举例】

$$\begin{aligned}(10101010)_2 &= (1 \times 2^7 + 0 \times 2^6 + 1 \times 2^5 + 0 \times 2^4 + 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0)_{10} \\&= (1 \times 128 + 0 \times 64 + 1 \times 32 + 0 \times 16 + 1 \times 8 + 0 \times 4 + 1 \times 2 + 0 \times 1)_{10} \\&= (170)_{10}\end{aligned}$$

十进制正整数转8位无符号二进制数

除2取余法 (逆序输出)

【举例】

$$(130)_{10} = (10000010)_2$$

$$130 \div 2 = 65 \quad \text{余} 0$$

$$65 \div 2 = 32 \quad \text{余} 1$$

$$32 \div 2 = 16 \quad \text{余} 0$$

$$16 \div 2 = 8 \quad \text{余} 0$$

$$8 \div 2 = 4 \quad \text{余} 0$$

$$4 \div 2 = 2 \quad \text{余} 0$$

$$2 \div 2 = 1 \quad \text{余} 0$$

$$1 \div 2 = 0 \quad \text{余} 1$$

凑值法 (必须记住8位二进制数各位的权值 128 64 32 16 8 4 2 1)

【举例】

$$(171)_{10} = (10101011)_2$$

$$= (1 \times \boxed{128} + 0 \times \boxed{64} + 1 \times \boxed{32} + 0 \times \boxed{16} + 1 \times \boxed{8} + 0 \times \boxed{4} + 1 \times \boxed{2} + 1 \times \boxed{1})_{10}$$

$\begin{matrix} \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\ b_7 & b_6 & b_5 & b_4 & b_3 & b_2 & b_1 & b_0 \end{matrix}$

(2) 地址块

注意:

128.14.35.7/20	是 IP 地址，同时指明了网络前缀为 20 位。该地址是 128.14.32.0/20 地址块中的一个地址。
128.14.32.0/20	是包含有多个 IP 地址的地址块，同时也是这个地址块中主机号为全 0 的 IP 地址。
128.14.35.7	是 IP 地址，但未指明网络前缀长度，不知道其网络地址。
128.14.32.0	不能指明一个网络地址，因为无法知道网络前缀是多少。

(3) 地址掩码 (address mask)

- 又称为**子网掩码** (subnet mask)。
- 位数：**32 位**。
- 目的：让机器从 IP 地址迅速算出网络地址。
- 由一连串 1 和接着的一连串 0 组成，**而 1 的个数就是网络前缀的长度**。

/20 地址块的地址掩码：11111111 11111111 11110000 00000000

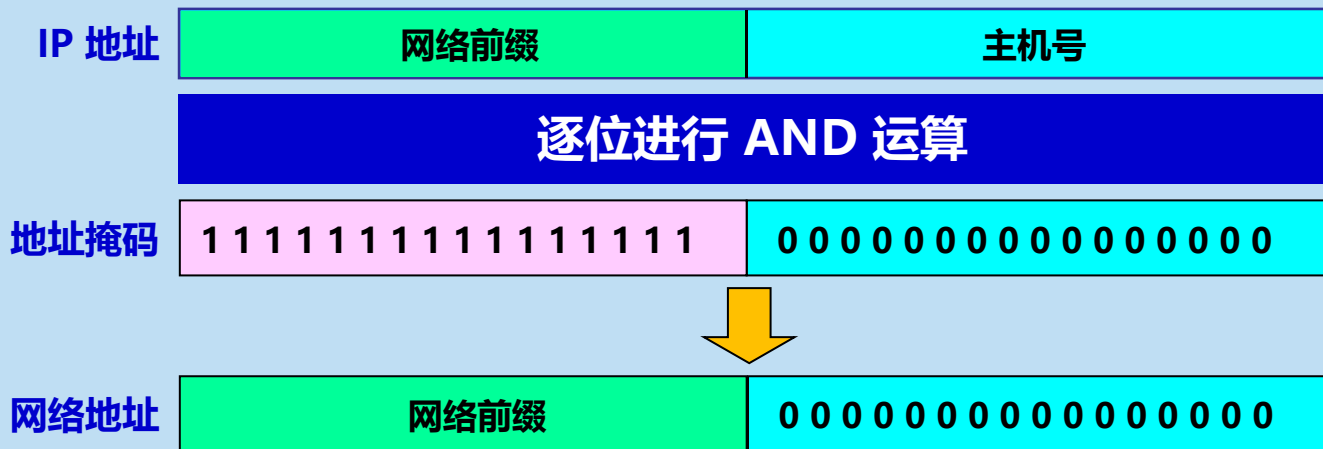
点分十进制记法：255.255.240.0

CIDR 记法：255.255.240.0/20。

默认地址掩码

A类地址	<div> <div>网络号</div> <div>主机号</div> </div> 默认地址掩码 255.0.0.0 <div> <div>11111111</div> <div>000000000000000000000000</div> </div>
B类地址	<div> <div>网络号</div> <div>主机号</div> </div> 默认地址掩码 255.255.0.0 <div> <div>1111111111111111</div> <div>0000000000000000</div> </div>
C类地址	<div> <div>网络号</div> <div>主机号</div> </div> 默认地址掩码 255.255.255.0 <div> <div>11111111111111111111</div> <div>00000000</div> </div>

网络地址 = (二进制的 IP 地址) AND (地址掩码)



网络地址 = (二进制的 IP 地址) AND (地址掩码)

【例】已知 IP 地址是 128.14.35.7/20。求网络地址。

(a) 点分十进制 IP 地址

128 . 14 . 35 . 7

(b) 二进制 IP 地址

1 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 1 0 0 0 1 1 0 0 0 0 0 1 1 1

(c) 地址掩码是 255.255.224.0

1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0

(d) IP 地址与地址掩码按位 AND

1 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0

(e) 网络地址 (点分十进制)

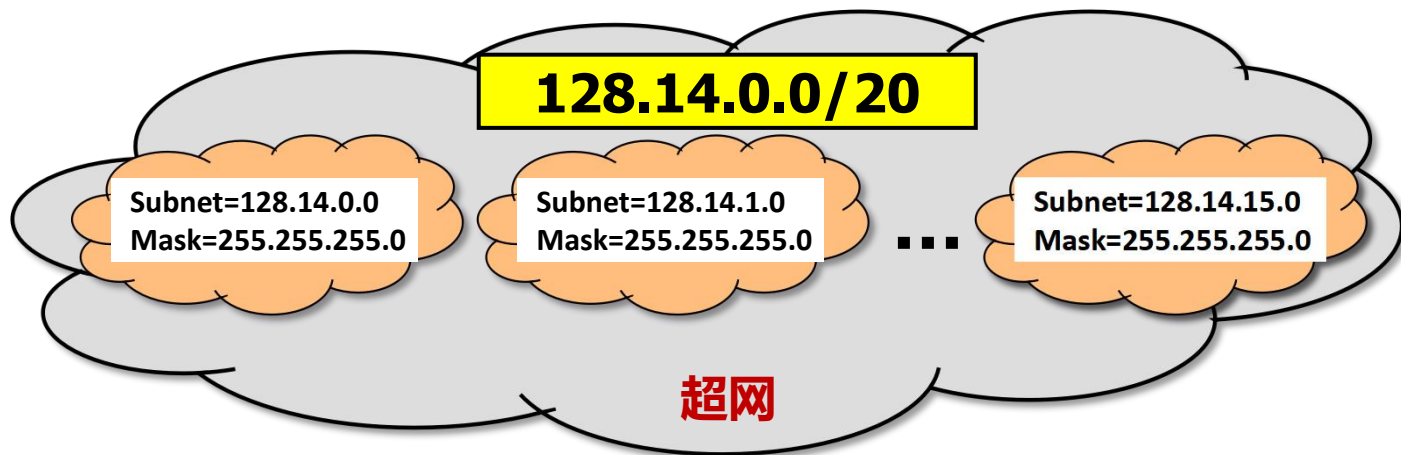
128 . 14 . 32 . 0

常用的 CIDR 地址块

网络前缀长度	点分十进制	包含的地址数	相当于包含分类的网络数
/13	255.248.0.0	512 K	8 个 B 类或 2048 个 C 类
/14	255.252.0.0	256 K	4 个 B 类或 1024 个 C 类
/15	255.254.0.0	128 K	2 个 B 类或 512 个 C 类
/16	255.255.0.0	64 K	1 个 B 类或 256 个 C 类
/17	255.255.128.0	32 K	128 个 C 类
/18	255.255.192.0	16 K	64 个 C 类
/19	255.255.224.0	8 K	32 个 C 类
/20	255.255.240.0	4 K	16 个 C 类
/21	255.255.248.0	2 K	8 个 C 类
/22	255.255.252.0	1 K	4 个 C 类
/23	255.255.254.0	512	2 个 C 类
/24	255.255.255.0	256	1 个 C 类
/25	255.255.255.128	128	1/2 个 C 类
/26	255.255.255.192	64	1/4 个 C 类
/27	255.255.255.224	32	1/8 个 C 类

构造超网

- 每一个 CIDR 地址块中的地址数**一定是 2 的整数次幂**。
- 除最后几行外，CIDR 地址块都包含了多个 C 类地址（是一个 C 类地址的 2^n 倍， n 是整数）。
- 因此在文献中有时称 CIDR 编址为**“构造超网”**。



三个特殊的 CIDR 地址块

网络前缀长度	点分十进制	说 明
/32	255.255.255.255	就是一个 IP 地址。这个特殊地址用于主机路由
/31	255.255.255.254	只有两个 IP 地址，其主机号分别为 0 和 1。 这个地址块用于点对点链路
/0	0.0.0.0	同时 IP 地址也是全 0，即 0.0.0.0/0。用于默认路由。

路由聚合 (route aggregation)

聚合前

16 个 C 类地址,
地址掩码=255.255.255.0,
路由表中需要 16 个路由项目。

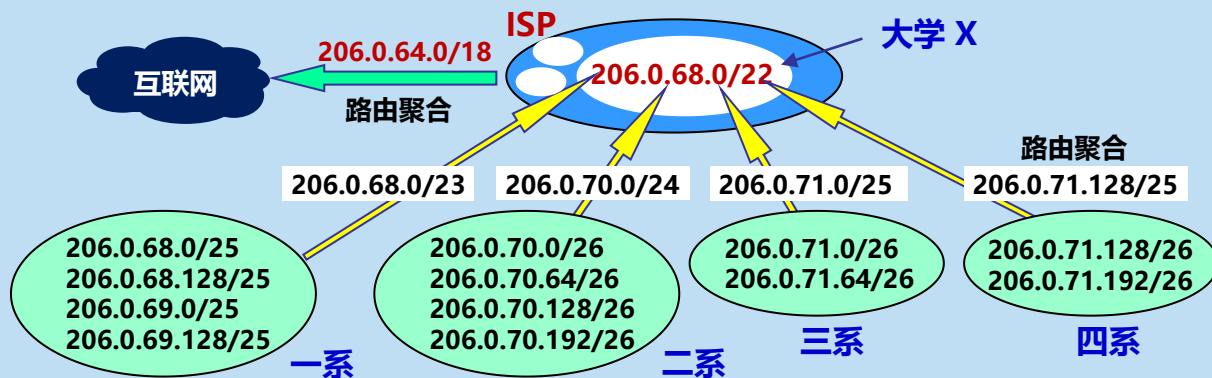
192.24.0
192.24.1
192.24.2
192.24.3
192.24.4
192.24.5
192.24.6
192.24.7
192.24.8
192.24.9
192.24.10
192.24.11
192.24.12
192.24.13
192.24.14
192.24.15

192.24.0.0 / 20

聚合后

聚合为 1 个 地址,
地址掩码=255.255.240.0,
路由表中只需 1 个路由项目。

CIDR 地址块划分举例



单位	地址块	二进制表示的地址块	相当于 C 类网络数
ISP	206.0.64.0/18	11001110.00000000.01*	64
大学	206.0.68.0/22	11001110.00000000.010001*	4
一系	206.0.68.0/23	11001110.00000000.0100010*	2
二系	206.0.70.0/24	11001110.00000000.01000110.*	1
三系	206.0.71.0/25	11001110.00000000.01000111.0*	1/2
四系	206.0.71.128/25	11001110.00000000.01000111.1*	1/2

可见：网络前缀越短，地址块所包含的地址数越多。

CIDR 地址块划分举例



这个 ISP 共有 64 个 C 类网络。如果不采用 CIDR 技术，则在与该 ISP 的路由器交换路由信息的每一个路由器的转发表中，需要有 64 行。采用地址聚合后，转发表中只需要用 1 行来指出到 **206.0.64.0/18** 地址块的下一跳。

在 ISP 内的路由器的转发表中，也仅需用 **206.0.68.0/22** 这 1 个项目，就能把外部发送到这个大学各系的所有分组，都转发到大学的路由器。

单位	地址块	二进制表示的地址块	相当于 C 类网络数
ISP	206.0.64.0/18	11001110.00000000.01*	64
大学	206.0.68.0/22	11001110.00000000.010001*	4
一系	206.0.68.0/23	11001110.00000000.0100010*	2
二系	206.0.70.0/24	11001110.00000000.01000110.*	1
三系	206.0.71.0/25	11001110.00000000.01000111.0*	1/2
四系	206.0.71.128/25	11001110.00000000.01000111.1*	1/2

可见：网络前缀越短，地址块所包含的地址数越多。

【2018年 题38】某路由表有转发接口相同的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，将这4条路由聚合后的目的网络地址为（ C ）。

A. 35.230.0.0/19

B. 35.230.0.0/20

C. 35.230.32.0/19

D. 35.230.32.0/20

解析 路由聚合的方法： 找共同前缀

	将左起第3个十进制数 转换成二进制形式	19比特共同前缀
35.230.32.0/21	→	3 5 . 2 3 0 . 0 0 1 0 0 0 0 0 . 0
35.230.40.0/21	→	3 5 . 2 3 0 . 0 0 1 0 1 0 0 0 . 0
35.230.48.0/21	→	3 5 . 2 3 0 . 0 0 1 1 0 0 0 0 . 0
35.230.56.0/21	→	3 5 . 2 3 0 . 0 0 1 1 1 0 0 0 . 0

路由聚合后的目的网络地址 35.230.32.0 /19

4. IP 地址的特点

(1) 每个 IP 地址都由网络前缀和主机号两部分组成。

(2) IP 地址是标志一台主机（或路由器）和一条链路的接口。

(3) 转发器或交换机连接起来的若干个局域网仍为一个网络

(4) 在 IP 地址中，所有分配到网络前缀的网络都是平等的。

IP 地址是一种**分等级**的地址结构。

- 方便了 IP 地址的分配和管理。
- 实现路由聚合，减小了转发表所占的存储空间，以及查找转发表的时间。

4. IP 地址的特点

(1) 每个 IP 地址都由网络前缀和主机号两部分组成。

(2) IP 地址是标志一台主机（或路由器）和一条链路的接口。

(3) 转发器或交换机连接起来的若干个局域网仍为一个网络

(4) 在 IP 地址中，所有分配到网络前缀的网络都是平等的。

- 当一台**主机**同时连接到**两个网络**上时，该主机就必须同时具有**两个相应的 IP 地址**，其**网络号必须是不同的**。这种主机称为**多归属主机** (multihomed host)。
- 一个**路由器**至少应当连接到两个网络，因此一个路由器至少应当有**两个不同的 IP 地址**。

4. IP 地址的特点

(1) 每个 IP 地址都由网络前缀和主机号两部分组成。

(2) IP 地址是标志一台主机（或路由器）和一条链路的接口。

(3) 转发器或交换机连接起来的若干个局域网仍为一个网络

(4) 在 IP 地址中，所有分配到网络前缀的网络都是平等的。

- 按照互联网的观点，一个网络（或子网）是指具有相同网络前缀的主机的集合。
- 转发器或交换机连接起来的若干个局域网都具有同样的网络号，它们仍为一个网络。
- 具有不同网络号的局域网必须使用路由器进行互连。

4. IP 地址的特点

(1) 每个 IP 地址都由网络前缀和主机号两部分组成。

(2) IP 地址是标志一台主机（或路由器）和一条链路的接口。

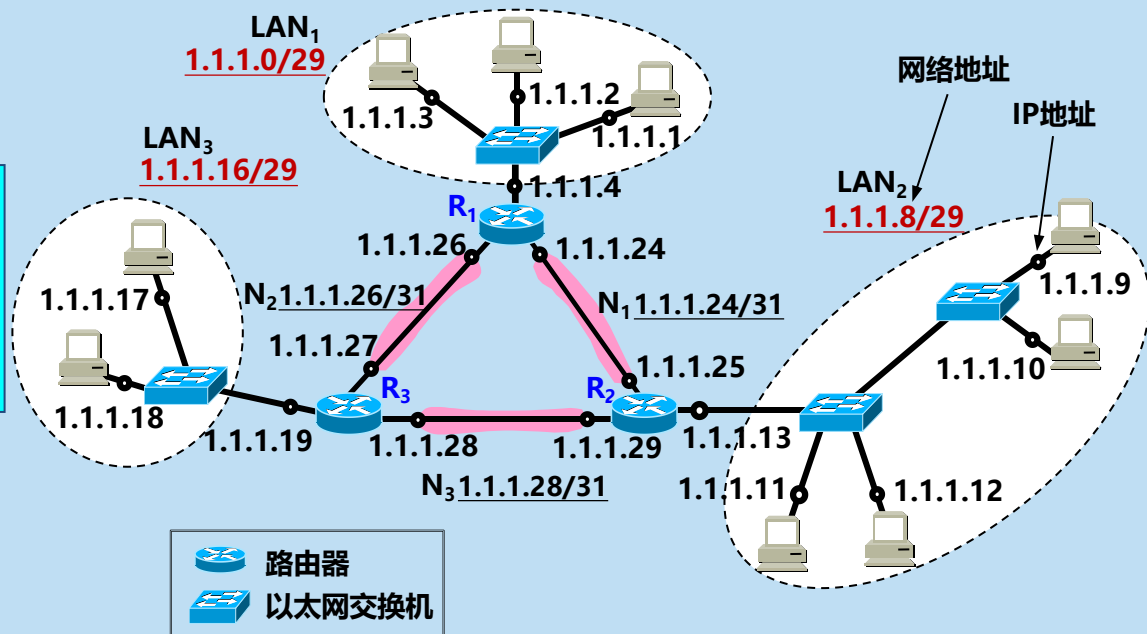
(3) 转发器或交换机连接起来的若干个局域网仍为一个网络

(4) 在 IP 地址中，所有分配到网络前缀的网络都是平等的。

- 互联网**同等对待**每一个 IP 地址，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

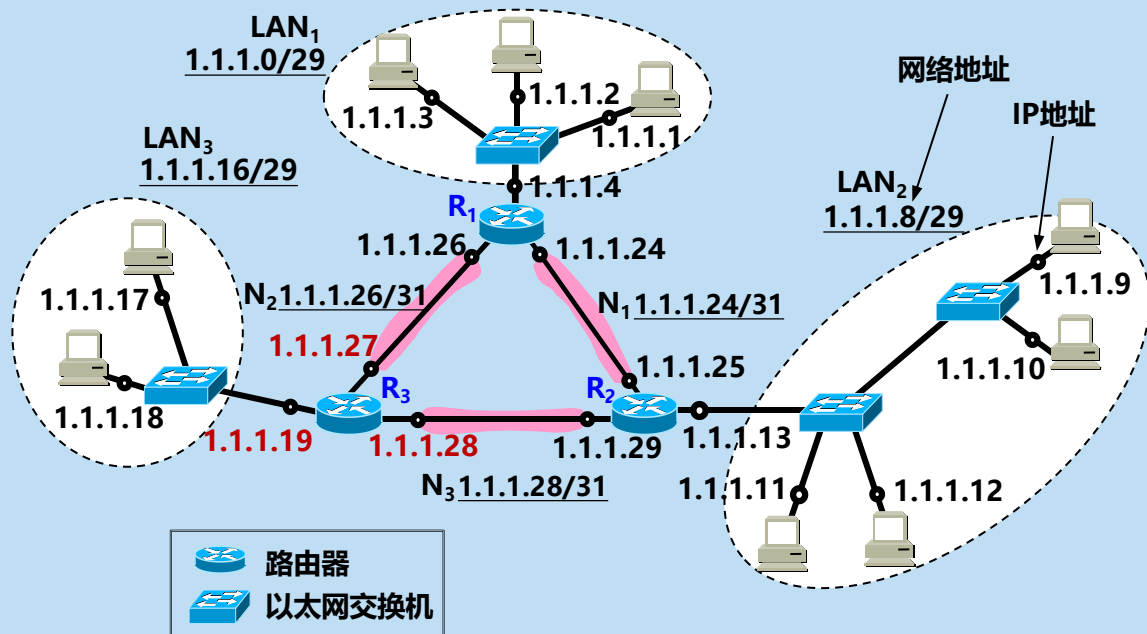
注意:

同一个局域网上的
主机或路由器的
IP 地址中的**网络号**
必须一样。



注意：

路由器的每一个接口都有一个**不同网络号**的 IP 地址。



地址分配、聚合、路由表转发相关内容是本书、本章的重点内容。请同学们课后自己练习，老师抽空课堂上讲解。

CIDR相关课后习题：

**课本P204 4-19、4-22、4-23、4-25、4-31、
4-33、4-47**

本讲小结

掌握：

- IP地址的构成（分类、CIDR）
- IP地址的含义

预习：

- IP地址与MAC地址
- ARP协议
- IP数据报的格式