

# 第四章 网络层

## 4.1 网络层提供的两种服务

1.虚电路服务。2.数据报服务。

网络层向上只提供简单灵活的、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。  
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）。

## 4.2 网际协议 IP

网际协议 IP 是 TCP/IP 体系中两个最主要地协议之一。  
与 IP 协议配套使用的还有三个协议：  
地址解析协议 ARP 、网际控制报文协议 ICMP、网际组管理协议 IGMP

### 4.2.1 虚拟互联网络

将网络相互连接起来的中间设备有以下四种：

- 1.物理层中继系统：转发器。
- 2.数据链路层使用的中间设备：网桥或桥接器。
- 3.网络层适应的中间设备：路由器。
- 4.网络层以上使用的中间设备：网关。

中继系统是转发器或网桥时，一般不称之为网络互连，因为仅仅是把一个网络扩大了，二者任然是一个网络。  
网络互连都指用路由器进行网络互连和路由选择。

### 4.2.2 分类的 IP 地址

在 TCP/IP 体系中，IP 地址是一个最基本的概念。  
IP 地址的编址方式（三个历史阶段）：  
分类的 IP 地址、子网划分、构成超网。

分类 IP 地址：  
每一类地址都由两个固定长度的字段组成，其中一个字段是网络号net-id，另一个字段时主机号 host-id，他标志该主机（或路由器）。主机号在它前面的网络号所指明的网络范围内必须是唯一的。

A类地址：8（0...）：24  
B类地址：16（10...）：16  
C类地址：24（110...）：8  
D类地址：（1110...）多播地址  
E类地址：（1111...）保留为今后使用

网络类别	最大可指派网络数	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
------	----------	------------	-------------	------------

网络类别	最大可指派网络数	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(        )	1	126	16777214
B	16383(        )	128.1	191.255	65534
C	2097151(        )	192.0.1	223.255.255	254

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	代表意思
0	0	可以	不可	在本网络上的本主机
0	host-id	可以	不可	在本网络上的某台主机
全 1	全 1	不可	可以	只在本网络上进行广播
net-id	全 1	不可	可以	对net-id上所有网络进行广播
127	非全 0 或非全 1的任何数	可以	可以	用于本地软件环回测试
net-id	全 0	可以	可以	路由器

### 4.2.3 IP 地址与硬件地址

IP 地址放在 IP 数据报的首部，而硬件地址放在 MAC 帧的首部。

硬件地址是数据链路层和物理层使用的地址

IP 地址是网络层和以上各层使用的地址，是一种逻辑地址

### 4.2.4 地址解析协议 ARP

通信时使用了两个地址：

IP 地址、MAC 地址。

地址解析协议 ARP 的作用：从网络层使用的 IP 地址，解析出在数据链路层使用的硬件地址。

每一个主机都设有一个 ARP 高速缓存，里面有所在的局域网上的各主机和路由器的 IP 地址到硬件地址的映射表。

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，再将此硬件地址写入 MAC 帧，然后通过局域网将该 MAC 帧发往此硬件地址。如没有，ARP 进程在本局域网上广播发送一个 ARP 请求分组。收到 ARP 响应分组后，将得到的 IP 地址到硬件地址的映射写入 ARP 高速缓存。

ARP 用于解决同一个局域网上的主机或路由器的 IP 地址和硬件地址的映射问题。

如果所要找的主机和源主机不在同一个局域网，那么就要通过 ARP 找到一个位于本局域网上的某个路由器的硬件地址，然后把分组发送给这个路由器，让这个路由器把分组转发给下一个网络。剩下的工作就由下一个网络来做。

### 4.2.5 IP 数据报的格式

一个 IP 数据报由首部和数据两部分组成。

首部的前一部分是固定长度，共 20 字节，是所有 IP 数据报必须具有的，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中的各字段：

版本（4位）：IP协议的版本。

首部长度（4位）：可表示的最大数值是15个单位（1个单位4个字节，说明总长最大为60字节）。

分区服务（8位）：一般不用。

总长度（16位）：首部和数据之和的长度。

标识（16位）

标志（3位）：

片偏移（13位）：较长的分组在分片后

某片在原分组中的相对位置。片偏移以 8 个字节为偏移单位。

生存时间（8位）：TTL，指示数据报在网络中可通过的路由器数的最大值

协议（8位）：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那个处理过程。

首部检验和（16位）：只检验数据报的首部，不检验数据部分。这里不采用 CRC 检验码而采用简单的计算方法。IP 数据报首部检验和的计算采用 16 位二进制反码求和算法

源地址、目的地址（各4字节）

IP 数据报首部的可变部分：

长度可变：1~40字节。实际上这些选项很少被使用。

## 4.2.6 IP 层转发分组的流程

按主机所在的网络地址来制作路由表

查找路由表：根据目的网络地址就能确定下一跳路由器，所以一定能找到目的主机所在网络上的路由器；只有到达最后一个路由器时，才试图向目的之进行直接交付。

路由器分组转发算法：

- 1.从数据报的首部提取目的主机的 IP 地址 D, 得出目的网络地址为 N。
- 2.若网络 N 与此路由器直接相连，则把数据报直接交付目的主机 D；否则是间接交付，执行 (3)。
- 3.若路由表中有目的地址为 D 的特定主机路由，则把数据报传送给路由表中所指明的下一跳路由器；否则，执行 (4)。
- 4.若路由表中有到达网络 N 的路由，则把数据报传送给路由表指明的下一跳路由器；否则，执行 (5)。
- 5.若路由表中有一个默认路由，则把数据报传送给路由表中所指明的默认路由器；否则，执行 (6)。
- 6.报告转发分组出错。

## 4.3 划分子网和构造超网

### 4.3.1 划分子网

划分子网纯属一个单位内部的事情。单位对外仍然表现为没有划分子网的网络

从主机号借用若干个位作为子网号 subnet-id，而主机号 host-id 也就相应减少了若干个

1.从两级 IP 地址搭配三级 IP 地址

划分子网：在 IP 地址中增加了一个“子网号字段”，使两级的IP地址变成了三级的 IP 地址。

从主机号借用若干位作为子网号，而主机号也就相应减少了若干位，

凡是从其他网络发送给本单位某个主机的 IP 数据报，仍然是根据 IP 数据报的目的网络号 net-id，先找到连接在本单位网

路上的路由器。

然后此路由器在收到 IP 数据报后，再按目的网络号 **net-id** 和子网号 **subnet-id** 找到目的子网。

最后就将 IP 数据报直接交付目的主机。

划分子网只是把 IP 地址的主机号 **host-id** 这部分进行再划分，而不改变 IP 地址原来的网络号 **net-id**。

优点：

减少了 IP 地址的浪费

使网络的组织更加灵活

更便于维护和管理

划分子网纯属一个单位内部的事情，对外部网络透明，对外仍表现为没有划分子网的一个网络。

## 2.子网掩码

用子网掩码可以找出 IP 地址中的子网部分。

规则：

子网掩码长度 = 32 位

子网掩码左边部分的一连串 1，对应于网络号和子网号

子网掩码右边部分的一连串 0，对应于主机号

(IP 地址) AND (子网掩码) = 网络地址

路由器的路由表中的每一个项目，除了要给出目的网络地址外，还必须同时给出该网络的子网掩码。

划分子网增加了灵活性，但却减少了能够连接在网络上的主机总数。

## 4.3.2 使用子网时分组的转发

在划分子网情况下路由器转发分组的算法：

- 1.从收到的分组的首部提取目的 IP 地址 **D**。
- 2.先用各网络的子网掩码和 **D** 逐位相“与”，看是否和相应的网络地址匹配。若匹配，则将分组直接交付。否则就是间接交付，执行(3)。
- 3.若路由表中有目的地址为 **D** 的特定主机路由，则将分组传送给指明的下一跳路由器；否则，执行 (4)。
- 4.对路由表中的每一行，将子网掩码和 **D** 逐位相“与”。若结果与该行的目的网络地址匹配，则将分组传送给该行指明的下一跳路由器；否则，执行 (5)。
- 5.若路由表中有一个默认路由，则将分组传送给路由表中所指明的默认路由器；否则，执行 (6)。
- 6.报告转发分组出错。

## 4.3.3 无分类编址 CIDR

网络前缀：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。

无分类的两级编址：网络前缀+主机号。

CIDR 使用“斜线记法”(slash notation)，它又称为 CIDR 记法，即在 IP 地址后面加上一个斜线“/”，然后写上网络前缀所占的位数（这个数值对应于三级编址中子网掩码中 1 的个数）。例如： 220.78.168.0/24

128.14.32.0/20 表示的地址块共有 一个地址。

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，也成为构成超网。

路由聚合有利于减少路由器之间的路由选择信息的交换，从而提高了整个互联网的性能。

构成超网：前缀长度不超过23位的 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。CIDR 的一个好处是：可以更加有效地分配 IPv4 的地址空间，可根据客户的需要分配适当大小的 CIDR 地址块。

最长前缀匹配：

使用 CIDR 时，路由表中的每个项目由“网络前缀”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。

使用二叉线索查找路由表：  
（类似平衡树写法，二分查找）

## 4.4 网际控制报文协议

ICMP 允许主机或路由器报告差错情况和提供有关异常情况的报告。ICMP 不是高层协议，而是 IP 层的协议。

ICMP 报文的种类：ICMP 差错报告报文、ICMP 询问报文。

差错报告报文：终点不可达、时间超过、参数问题、改变路由。

询问报文：回送请求和回答、时间戳请求回答。

ICMP 应用举例：

PING 用来测试两个主机之间的连通性

tracert 用来跟踪一个分组从源点到终点的路径。

## 4.5 互联网的路由选择协议

### 有关路由选择的几个基本概念

算法必须是正确的和完整的。

算法在计算上应简单。

算法应能适应通信量和网络拓扑的变化，这就是说，要有自适应性。

算法应具有稳定性。

算法应是公平的。

算法应是最佳的。

静态路由选择策略（非自适应） 动态路由选择策略（自适应）

采用分层次的路由选择协议：

1.互联网规模非常大，路由表会巨大，处理起来太花时间。2.许多单位不愿意外界了解自己单位网络的不拘细节和本部门采用的路由选择协议。

自治系统AS：

定义：在单一技术管理下的一组路由器，而这些路由器使用一种自治系统内部的路由选择协议和共同的度量。

一个 AS 对其他 AS 表现出的是一个单一的和意志的路由选择协议。

路由选择协议划分为两类：

1.内部网关协议 IGP（RIP 和 OSPF 协议）

2.外部网关协议 EGP（BGP-4）

自治系统之间的叫域间路由选择，自治系统内部的叫域内自由选择。

内部网关协议 RIP

RIP是一种基于距离向量的路由选择协议。  
从一路由器到直接连接的网络的距离定义为 1， 经过路由器则加 1。  
RIP 允许一条路径最多只能包含 15 个路由器，因此距离为16相当于不可达，所以 RIP 只适用于小型互联网。  
RIP 选择一条具有最少路由器的路由。

RIP 是应用层协议，封装于 UDP 数据报。  
RIP 的首部占 4 个字节，每个路由信息占 20 个字节，最多 25 个路由，即最大长度为 504 字节，超过则再用一个 RIP 报文传送。

RIP 协议特点：好消息传播得快，坏消息传播得慢。

优点：实现简单、开销小。

RIP 与 OSPF 对照：

	和谁交换信息	交换什么信息	什么时候交换
RIP	相邻路由器	全部信息	固定间隔时间
OSPF	自治系统所有路由器	部分信息	链路状态发生变化时（洪泛法发送）

4.5.3 内部网关协议 OSPF

开放最短优先路径，洪泛法

4.5.4 外部网关协议 BGP

是不同自治系统的路由器之间交换路由信息的协议

4.5.5 路由器的构成

路由器的作用：  
1.连通不同的网络  
2.选择信息传送的线路，选择畅通快捷的近路，能大大提高通信的速度，减轻网络系统通信负荷，节约网络系统资源，提高网络系统畅通率，从而让网络系统发挥更大的效率。

4.6 IPv6

4.6.1 IPv6的基本首部

- 1.更大的地址空间：128 位
- 2.扩展的地址层次结构
- 3.灵活的首部格式

- 4.改进的选项
- 5.允许协议继续扩充
- 6.支持即插即用
- 7.支持资源的预分配
- 8.IPv6 首部改为 8 字节对齐。

IPv6 数据报有两大部分组成，即基本首部和后面的有效载荷。有效载荷允许有零个或多个扩展首部  
首部为固定的 40 字节。

六种扩展首部：

- 1.zhu't