**5.1 IP 基础知识**

## IP 基本认识

IP 在 TCP/IP 参考模型中处于第三层，也就是**网络层**。

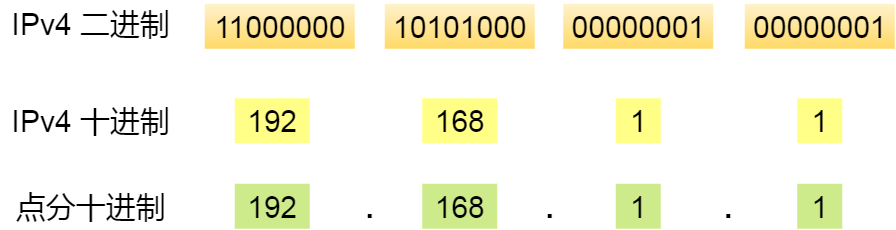
网络层的主要作用是：**实现主机与主机之间的通信，也叫点对点（end to end）通信。**

网络层与数据链路层有什么关系呢？

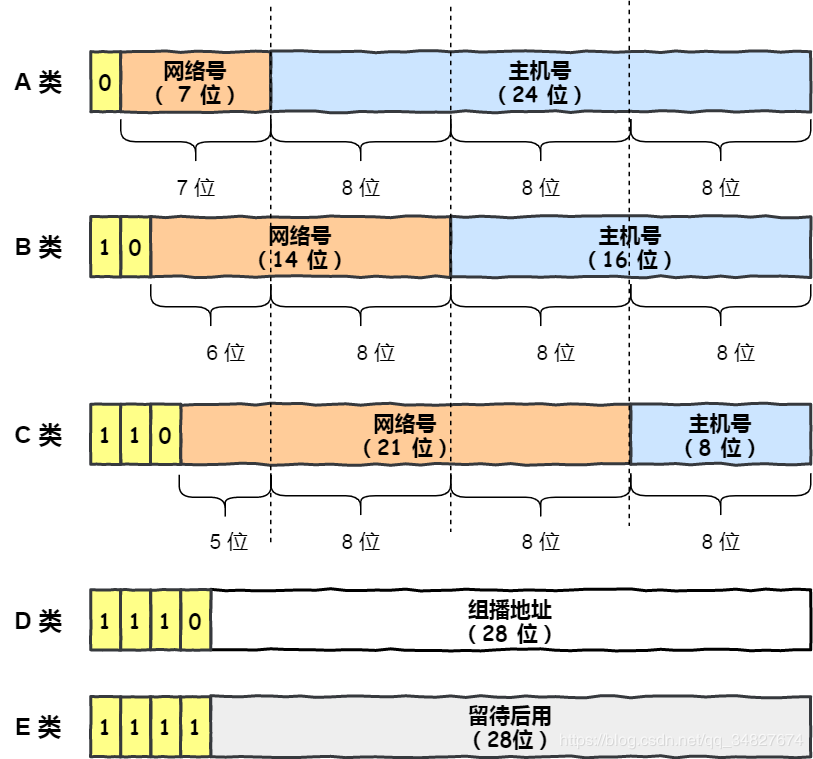
**MAC 的作用则是实现「直连」的两个设备之间通信，而 IP 则负责在「没有直连」的两个网络之间进行通信传输。**

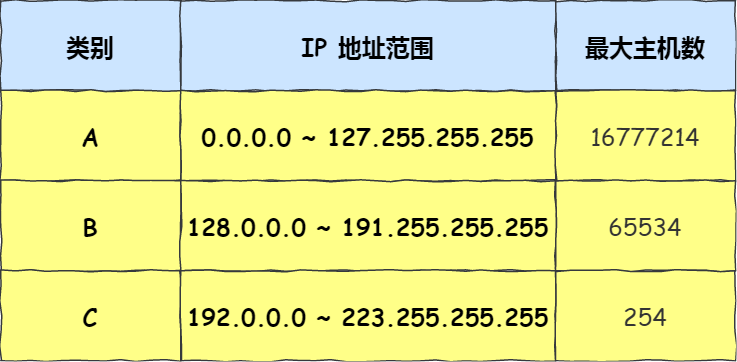
**源IP地址和目标IP地址在传输过程中是不会变化的（前提：没有使用 NAT 网络），只有源 MAC 地址和目标 MAC 一直在变化。**

## IP 地址的基础知识



### IP 地址的分类





A、B、C 分类地址最大主机个数是如何计算的呢？

2的主机号位数次方-2

减 2的原因：**全0和全1不可以用**

* 主机号全为 1 指定某个网络下的所有主机，用于广播
* 主机号全为 0 指定某个网络

广播地址用于什么？

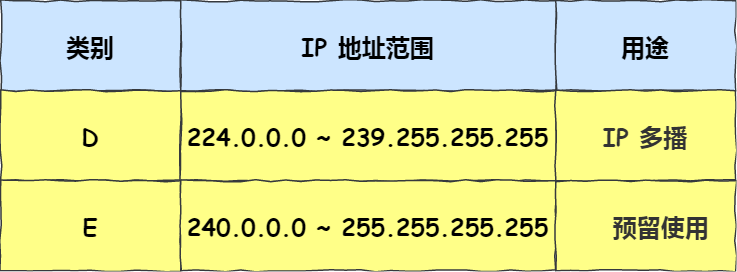
广播地址用于在**同一个链路中相互连接的主机之间发送数据包**

广播地址可以分为本地广播和直接广播两种。

* **在本网络内广播的叫做本地广播**。例如网络地址为 192.168.0.0/24 的情况下，广播地址是 192.168.0.255 。因为**这个广播地址的 IP 包会被路由器屏蔽**，所以不会到达 192.168.0.0/24 以外的其他链路上。
* **在不同网络之间的广播叫做直接广播。**例如网络地址为 192.168.0.0/24 的主机向 192.168.1.255/24 的目标地址发送 IP 包。收到这个包的路由器，将数据转发给 192.168.1.0/24，使得所有 192.168.1.1~192.168.1.254 的主机都能收到这个包（由于直接广播有一定的安全问题，多数情况下会在路由器上设置为不转发）。

什么是 D、E 类地址？

 D 类和 E 类地址是没有主机号的，所以不可用于主机 IP，D 类常被用于**多播（组播）**，E 类是预留的分类，暂时未使用。



多播地址用于什么？

多播用于**将包发送给特定组内的所有主机。广播无法穿透路由**，若想给其他网段发送同样的包，就可以使用**可以穿透路由的多播**。

从 224.0.0.0 ~ 239.255.255.255 都是多播的可用范围，其划分为以下三类：

* 224.0.0.0 ~ 224.0.0.255 为预留的组播地址，**只能在局域网中**，路由器是不会进行转发的。
* 224.0.1.0 ~ 238.255.255.255 为**用户可用的组播地址**，可以**用于 Internet 上**。
* 239.0.0.0 ~ 239.255.255.255 为**本地管理组播地址**，可供内部网在内部使用，仅在**特定的本地范围内有效**。

IP 分类的优点

**简单明了、选路简单**

IP 分类的缺点

缺点一

**同一网络下没有地址层次，缺少地址的灵活性**

缺点二

**不能很好的与现实网络匹配，C类太少，B类太多**

### 无分类地址 CIDR

**无分类地址**不再有分类地址的概念，32 比特的 IP 地址被划分为两部分，前面是**网络号**，后面是**主机号**。

表示形式 a.b.c.d/x，其中 /x 表示前 x 位属于**网络号**， x 的范围是 0 ~ 32，这就使得 IP 地址更加具有灵活性。

**子网掩码：**划分网络号与主机号的形式

**将子网掩码和 IP 地址按位计算 AND，就可得到网络号。**

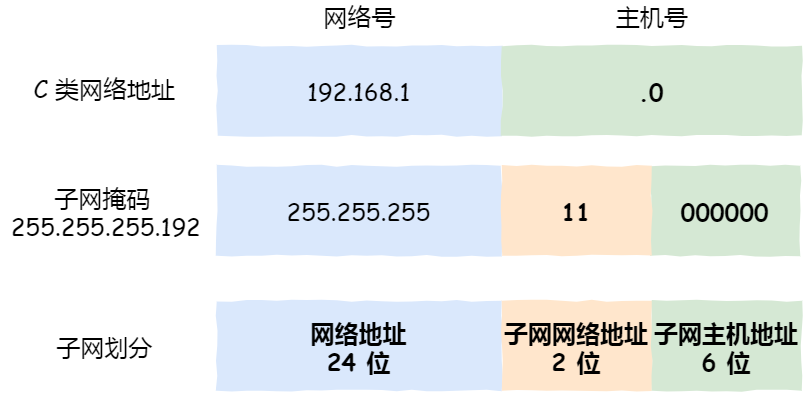
为什么要分离网络号和主机号？

通讯时候用于判断两台计算机是否处于同一个广播域内

划分子网

**子网掩码**的另一个作用：划分子网

网络地址 192.168.1.0，使用子网掩码 255.255.255.192 对其进行子网划分：



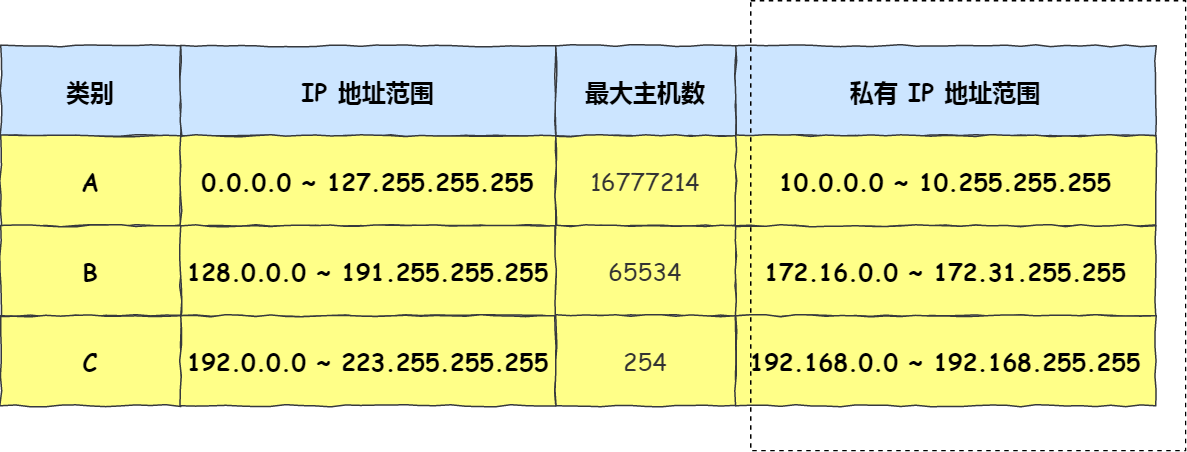
每个子网中，子网主机地址全0和全1也有特殊用途，不能用

全0：这个子网的网络地址

全1：这个子网的广播地址

### 私有 IP 地址

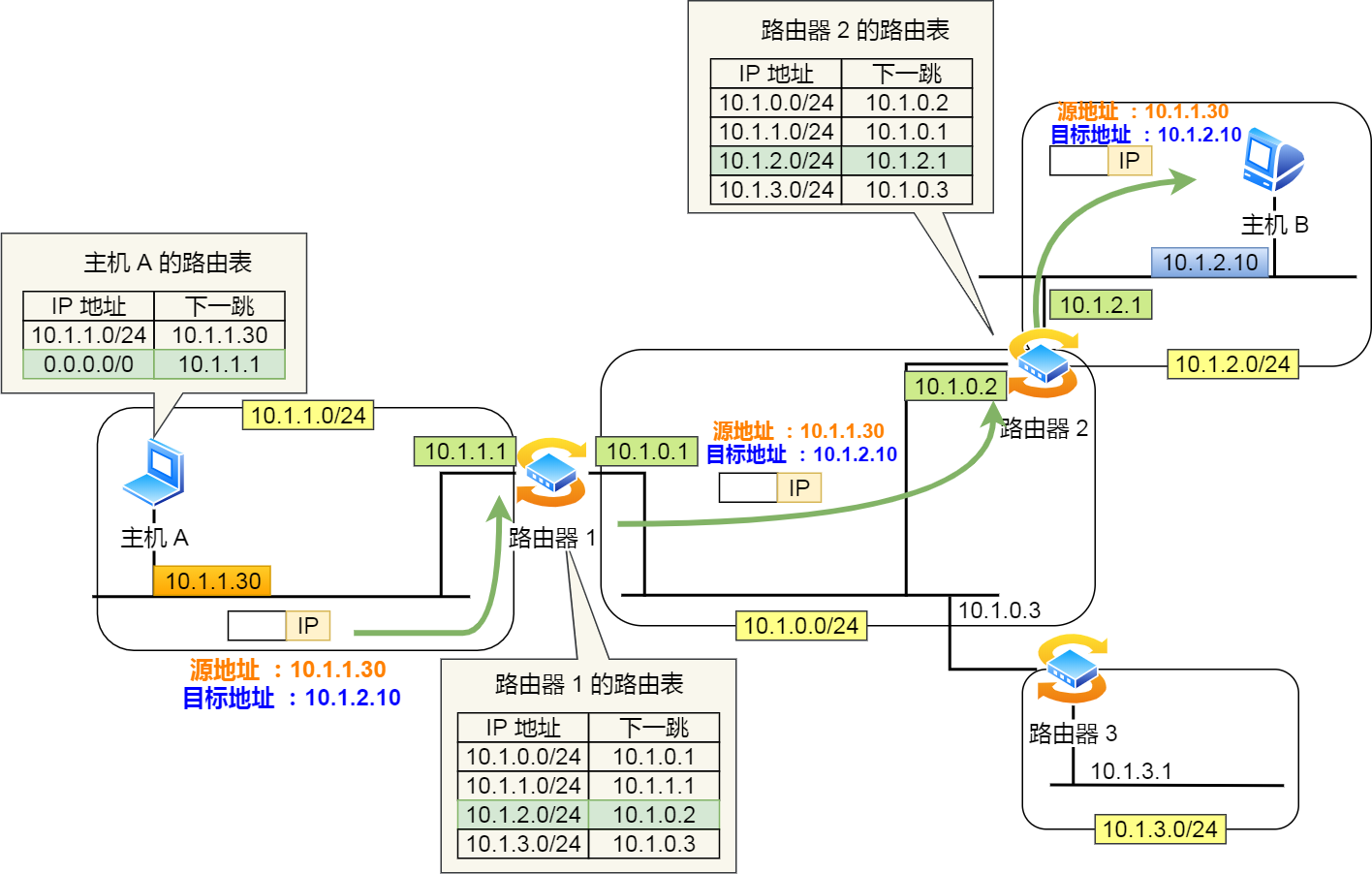
 A、B、C 类地址中，分为公有 IP 地址和私有 IP 地址。允许组织内部的 IT 人员自己管理、自己分配，而且可以重复。



### IP 地址与路由控制

IP地址的**网络地址**这一部分是用于进行路由控制。

发送 IP 包时，首先要确定 IP 包首部中的目标地址，再从路由控制表中找到与该地址具有**相同网络地址**的记录，根据该记录将 IP 包转发给相应的下一个路由器。如果路由控制表中存在多条相同网络地址的记录，就选择相同位数最多的网络地址，也就是最长匹配。



环回地址

计算机使用一个特殊的 IP 地址 **127.0.0.1 作为环回地址**。与该地址具有相同意义的是一个叫做 localhost 的主机名。使用这个 IP 或主机名时，**数据包不会流向网络**。

### IP 分片与重组

每种数据链路的最大传输单元 MTU 都是不相同的

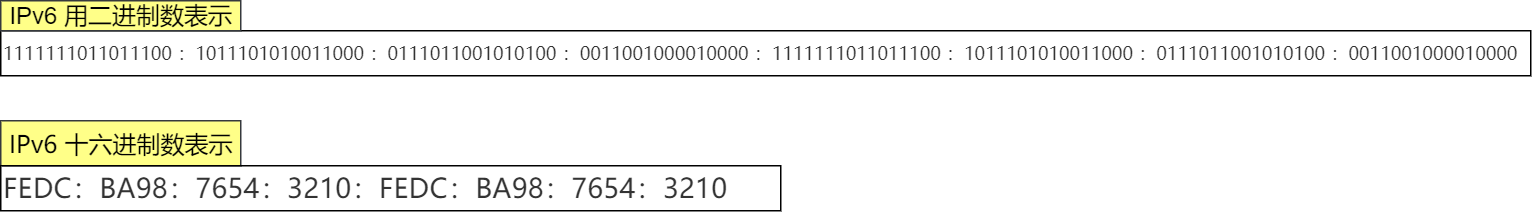
最常见数据链路是**以太网**，它的 MTU 是 1500 字节

 IP 数据包大小大于 MTU 时， IP 数据包就会被分片，由**目标主机**重组。**中间路由只会对 IP 包进行分片（在MTU进一步变小时候），不会对分片进行重组**

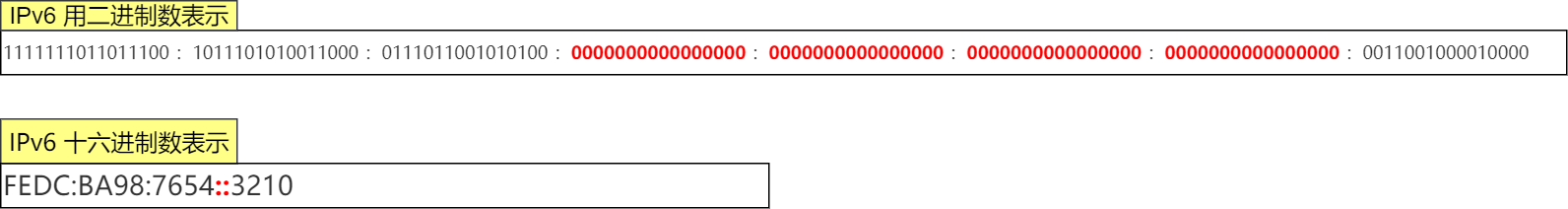
在分片传输中，**一旦某个分片丢失，则会造成整个 IP 数据报作废**，所以 TCP 引入了 MSS 也就是在 TCP 层进行分片不由 IP 层分片，那么对于 **UDP 尽量不要发送一个大于 MTU 的数据报文**。

### IPv6 基本认识

IPv6 的地址是 128 位的，是以每 16 位作为一组，每组用冒号 「:」 隔开

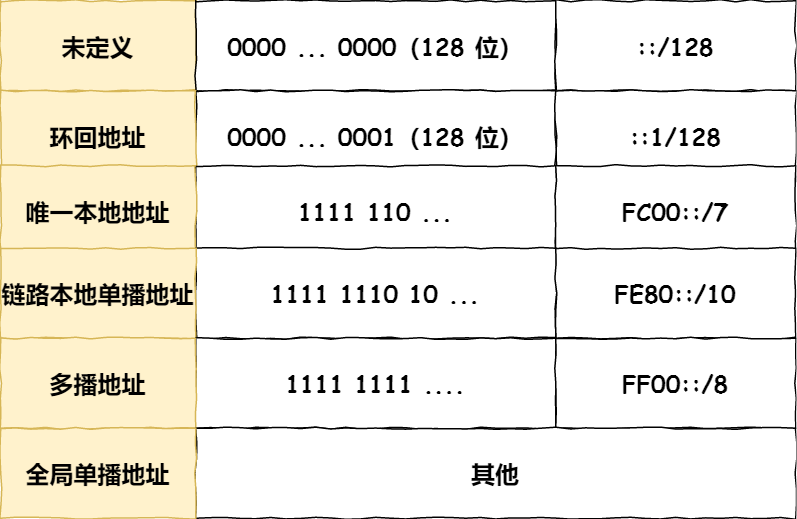


如果出现连续的 0 时还可以将这些 0 省略，并用两个冒号 「::」隔开。但是，一个 IP 地址中只允许出现一次两个连续的冒号



IPv6 的地址主要有以下类型地址：

* 单播地址，用于一对一的通信
* 组播地址，用于一对多的通信
* 任播地址，用于通信最近的节点，最近的节点是由路由协议决定
* **没有广播地址**

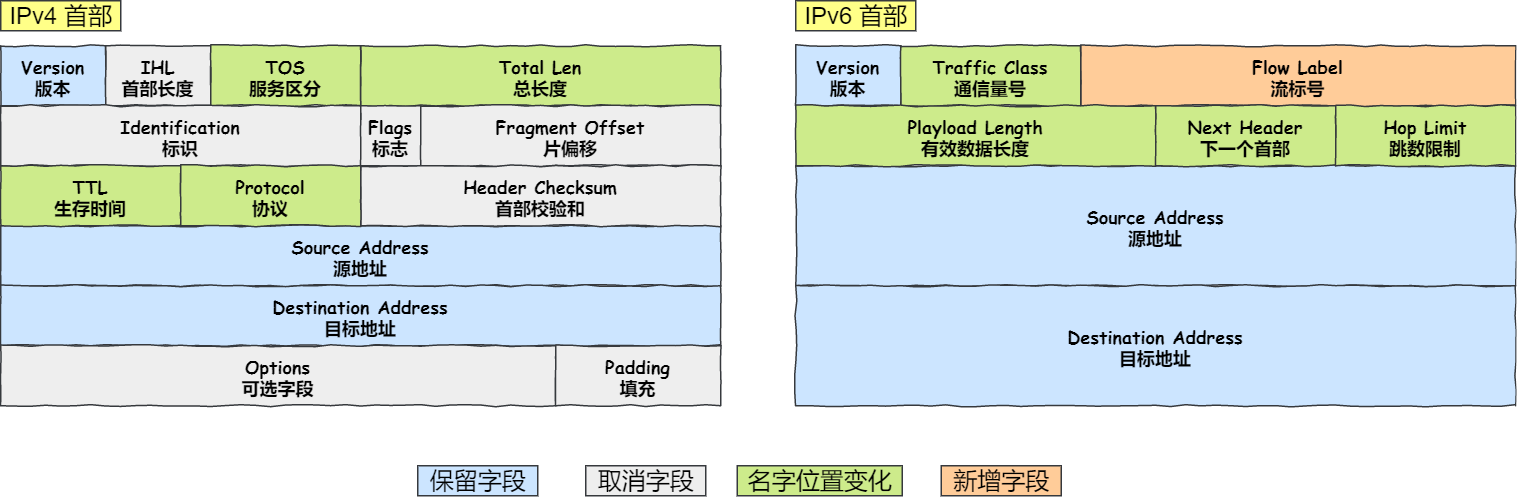


IPv6 单播地址类型

有三类：

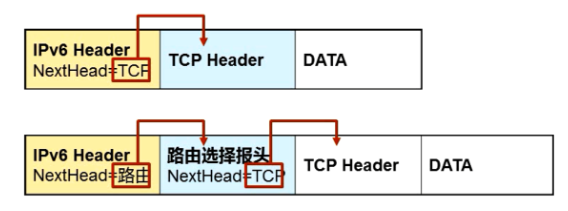
* 在**同一链路单播通信，不经过路由器**，可以使用**链路本地单播地址**，IPv4 没有此类型
* 在**内网里单播通信**，可以使用**唯一本地地址**，相当于 IPv4 的私有 IP
* 在**互联网通信**，可以使用**全局单播地址**，相当于 IPv4 的公有 IP

### IPv4 首部与 IPv6 首部



IPv6 相比 IPv4 的首部改进：

* **取消了首部校验和字段。**
* **取消了分片/重新组装相关字段。** 分片与重组是耗时的过程，IPv6 不允许在中间路由器进行分片与重组（IPv4中有可能分片），这种操作只能在源与目标主机，这将大大提高了路由器转发的速度。
* **取消选项字段。** 选项字段不再是标准 IP 首部的一部分了，但它并没有消失，而是可能出现在 IPv6 首部中的「下一个首部」指出的位置上。删除该选项字段使的 IPv6 的首部成为固定长度的 40 字节。



## IP 协议相关技术

### DNS

**域名解析系统**

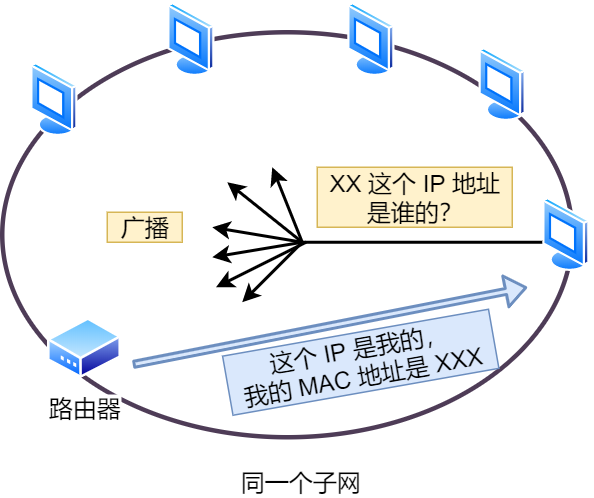
域名解析的工作流程

**浏览器**首先看一下自己的**缓存**里有没有，如果没有就向**操作系统的缓存**要，还没有就检查**本机域名解析文件 hosts**，如果还是没有，就会 DNS 服务器进行查询，查询的过程如下：

1. 客户端首先会发出一个 DNS 请求，问 www.server.com 的 IP 是啥，并发给**本地 DNS 服务器**（也就是客户端的 TCP/IP 设置中填写的 DNS 服务器地址）。
2. 本地域名服务器收到客户端的请求后，如果缓存里的表格能找到 www.server.com，则它直接返回 IP 地址。如果没有，本地 DNS 会去问它的**根域名服务器**：“老大， 能告诉我 www.server.com 的 IP 地址吗？” **根域名服务器是最高层次的，它不直接用于域名解析，但能指明一条道路**。
3. 根 DNS 收到来自本地 DNS 的请求后，发现后置是 .com，说：“www.server.com 这个域名归 .com 区域管理”，我给你 **.com 顶级域名服务器**地址给你，你去问问它吧。”
4. 本地 DNS 收到顶级域名服务器的地址后，发起请求问“老二， 你能告诉我 www.server.com 的 IP 地址吗？”
5. 顶级域名服务器说：“我给你负责 **www.server.com 区域的权威 DNS 服务器**的地址，你去问它应该能问到”。
6. 本地 DNS 于是转向问权威 DNS 服务器：“老三，www.server.com对应的IP是啥呀？” server.com 的权威 DNS 服务器，它是域名解析结果的原出处。为啥叫权威呢？就是我的域名我做主。
7. 权威 DNS 服务器查询后将对应的 IP 地址 X.X.X.X 告诉本地 DNS。
8. **本地 DNS 再将 IP 地址返回客户端**，客户端和目标建立连接。

### ARP

**地址解析协议**



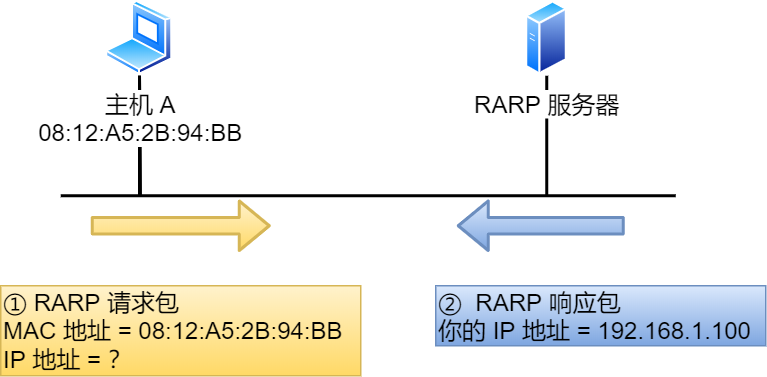
* 主机会通过**广播发送 ARP 请求**，这个包中包含了想要知道的 MAC 地址的主机 IP 地址。
* 当同个链路中的所有设备收到 ARP 请求时，会去拆开 ARP 请求包里的内容，如果 ARP 请求包中的**目标 IP 地址与自己的 IP 地址一致**，那么这个设备就将自己的 MAC 地址塞入 **ARP 响应包**返回给主机。

操作系统通常会把第一次通过 ARP 获取的 MAC 地址缓存起来，超时清除。

RARP 协议

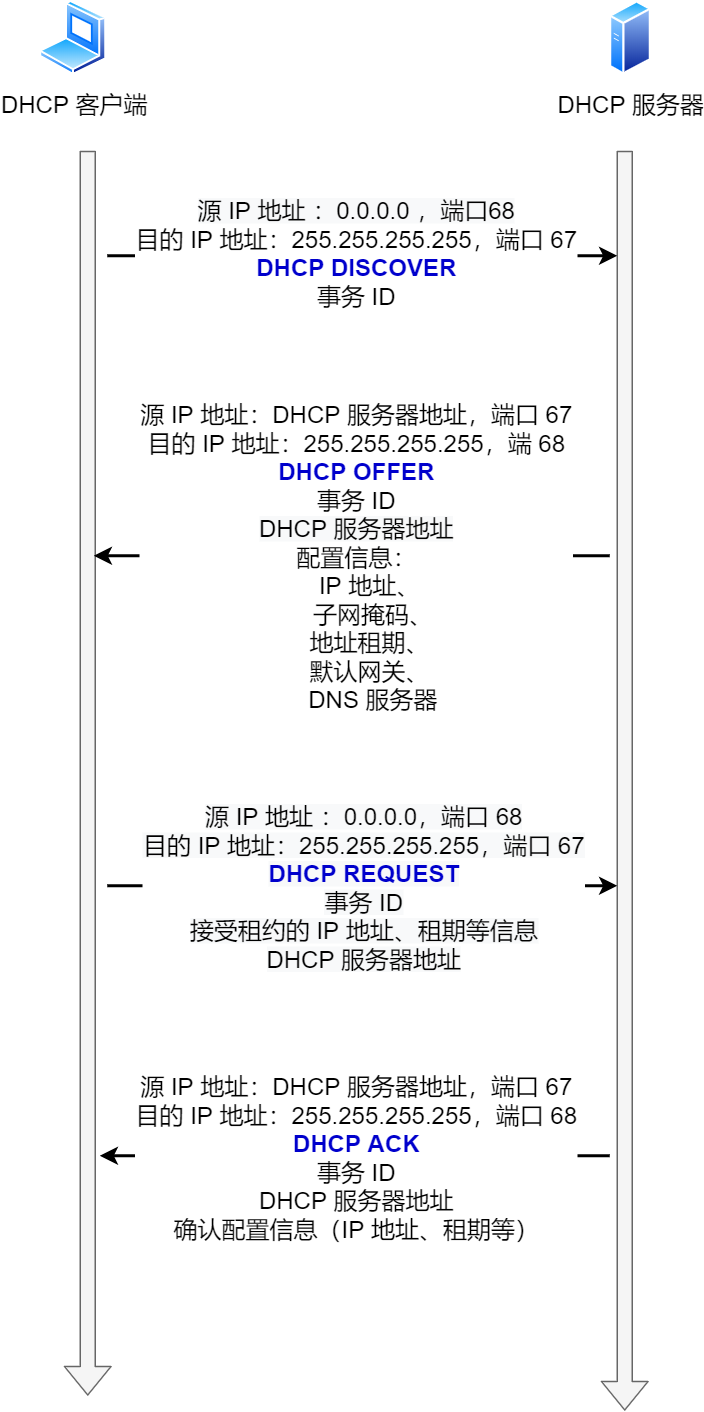
RARP 是**已知 MAC 地址求 IP 地址**。例如将打印机服务器等小型嵌入式设备接入到网络时就经常会用得到。

通常这需要架设一台 RARP 服务器，在这个服务器上注册设备的 MAC 地址及其 IP 地址。然后再将这个设备接入到网络



### DHCP

**动态主机配置协议**



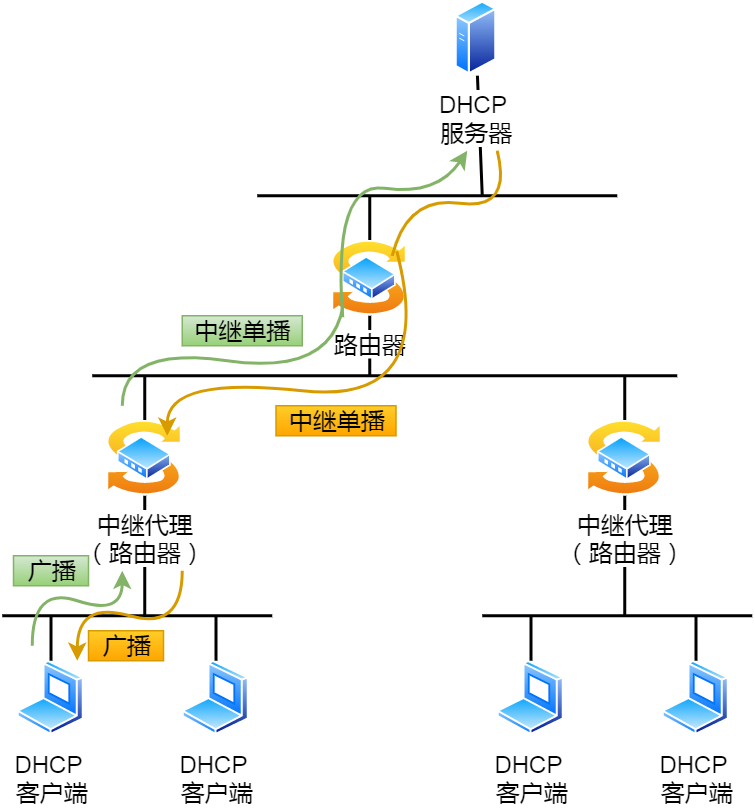
DHCP **客户端**进程监听的是 68 端口号，DHCP **服务端**进程监听的是 67 端口号。

* 客户端首先发起 **DHCP 发现报文（DHCP DISCOVER）** 的 IP 数据报，由于客户端没有 IP 地址，也不知道 DHCP 服务器的地址，所以使用的是 UDP **广播**通信，其使用的广播目的地址是 255.255.255.255（端口 67） 并且使用 0.0.0.0（端口 68） 作为源 IP 地址。DHCP 客户端将该 IP 数据报传递给链路层，链路层然后将帧广播到所有的网络中设备。
* DHCP 服务器收到 DHCP 发现报文时，用 **DHCP 提供报文（DHCP OFFER）** 向客户端做出响应。该报文仍然使用 IP 广播地址 255.255.255.255，该报文信息携带服务器提供可租约的 IP 地址、子网掩码、默认网关、DNS 服务器以及 **IP 地址租用期**。
* 客户端收到**一个或多个**服务器的 DHCP 提供报文后，**从中选择一个服务器**，并向选中的服务器（注意还是用的广播目的地址，不是DHCP服务器地址）发送 **DHCP 请求报文（DHCP REQUEST**进行响应，回显配置的参数。
* 最后，服务端用 **DHCP ACK 报文**对 DHCP 请求报文进行响应，应答所要求的参数。

**全程都是使用 UDP 广播通信**

如果 DHCP 服务器和客户端不是在同一个局域网内

由于路由器不会转发广播包，使用**DHCP 中继代理**，**对不同网段的 IP 地址分配也可以由一个 DHCP 服务器统一进行管理。**

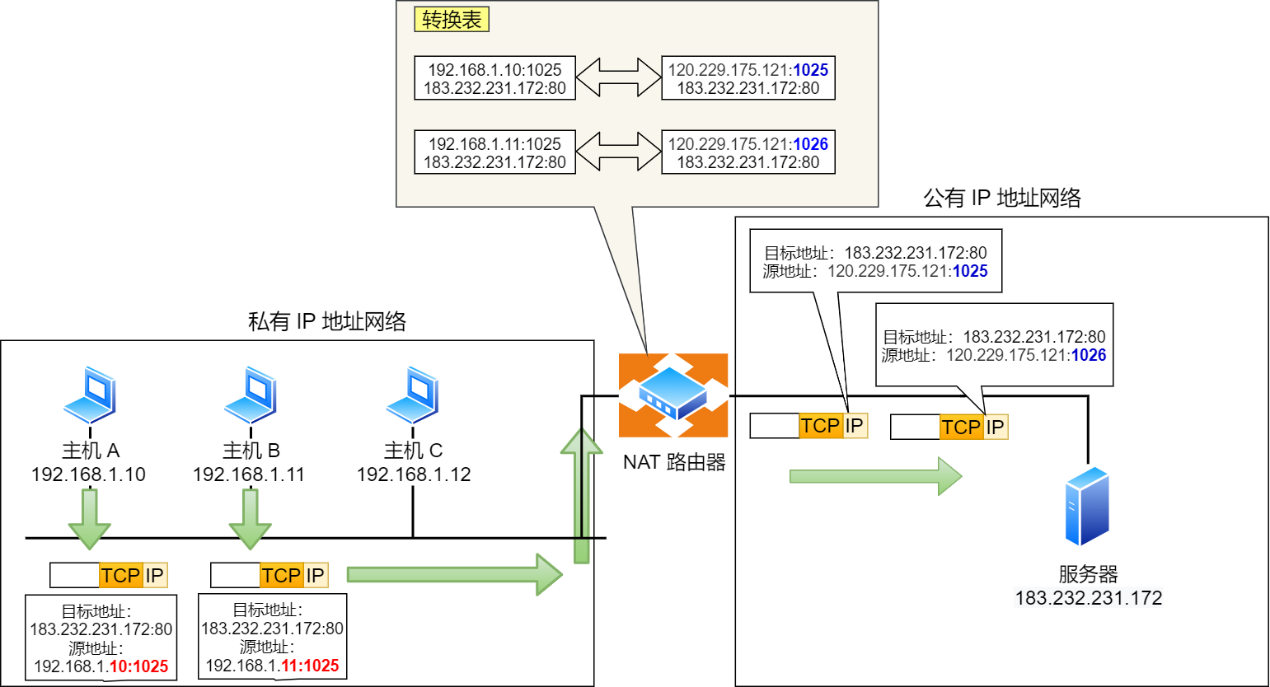


### NAT

**网络地址转换**

NAT 就是同个公司、家庭、教室内的主机对外部通信时，把私有 IP 地址转换成公有 IP 地址。

普通的 NAT 转换没什么意义，可以把 IP 地址 + 端口号一起进行转换，这种转换技术就叫**网络地址与端口转换 NAPT。**



有两个客户端 192.168.1.10 和 192.168.1.11 同时与服务器 183.232.231.172 进行通信，并且这**两个客户端的本地端口都是 1025**。

此时，**两个私有 IP 地址都转换 IP 地址为公有地址 120.229.175.121，但是以不同的端口号作为区分。**生成一个 NAPT 路由器的**转换表**，就可以正确地转换地址跟端口的组合。转换表在 NAT 路由器上自动生成。例如，在 TCP 的情况下，建立 TCP 连接首次握手时的 SYN 包一经发出，就会生成这个表。而后又随着收到关闭连接时发出 FIN 包的确认应答从表中被删除。

NAT缺点

NAT/NAPT 都依赖于自己的转换表：

* 外部无法**主动**与 **NAT 内部服务器**建立连接，因为 NAPT 转换表没有转换记录。
* 转换表的生成与转换操作都会产生性能开销。
* 通信过程中，如果 NAT 路由器重启了，所有的 TCP 连接都将被重置。

如何解决 NAT 潜在的问题

第一种就是改用 IPv6

每台设备都可以配置一个公有 IP 地址

第二种 NAT 穿透技术

在 NAT 穿透技术中，NAT设备后的应用程序处于主动地位，它已经明确地知道 NAT 设备要修改它外发的数据包，于是它主动配合 NAT 设备的操作，主动地建立好映射，这样就不像以前由 NAT 设备来建立映射了。