

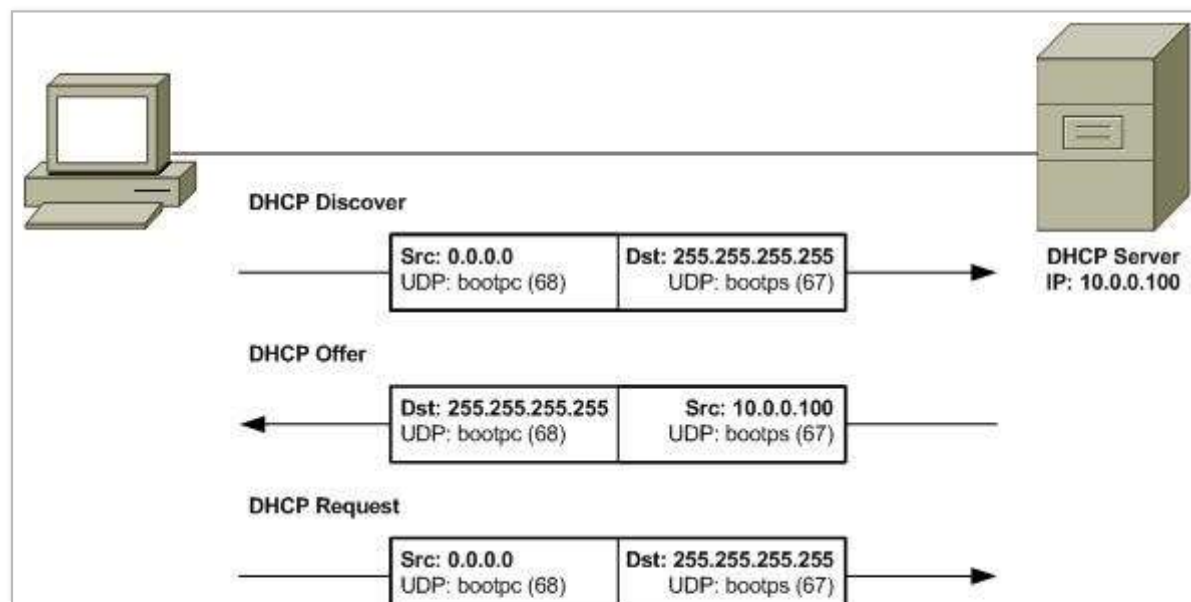
IPv6 --- 动态地址配置

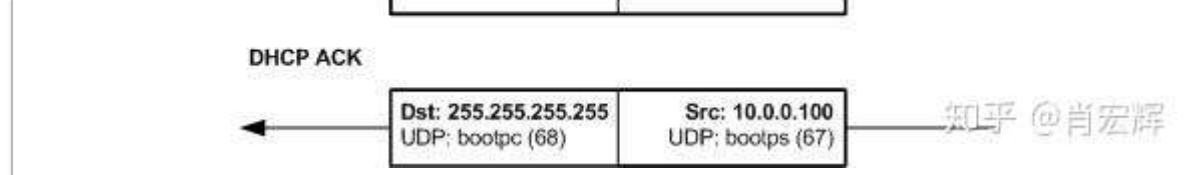
网络地址的配置是网络层管理的重要步骤。最直观的配置方式是静态配置 IP 地址，但是这增加了运维管理成本，并且没有灵活性。因此，大部分场合下，IP 地址都是通过动态配置的。

DHCP

在 IPv4 下，IP 地址的动态配置依赖于 DHCP 协议。DHCP 全称是 Dynamic Host Configuration Protocol，由 RFC2131 定义。DHCP 由 BOOTP 协议（RFC951）发展而来，并且完全兼容 BOOTP。

DHCP 是一个 client-server 模式的协议，Server 管理着 IP 地址池。一般情况下，主机作为 Client，通过下图的方式与 DHCP Server 交互，并请求 IP 地址。IP 地址包含在 DHCP offer 和 ACK 中。





DHCP 对应的中文是动态主机配置协议。所以除了 IP 地址以外，同时也能配置子网掩码长度，主机名，路由，默认网关，DNS 等。所以在 IPv4 中，DHCP 协议就可以完成大部分的主机动态配置。DHCP 协议已经使用了几十年，这里不做过多的介绍。

DHCPv6

对应于 DHCP，IPv6 协议中也有 DHCPv6（RFC8415/RFC3315）。但是就像 IPv6 完全不兼容 IPv4 一样，DHCPv6 除了工作模式与 DHCP 类似以外，与 DHCP 也不兼容，它的协议内容也进行了重新定义。

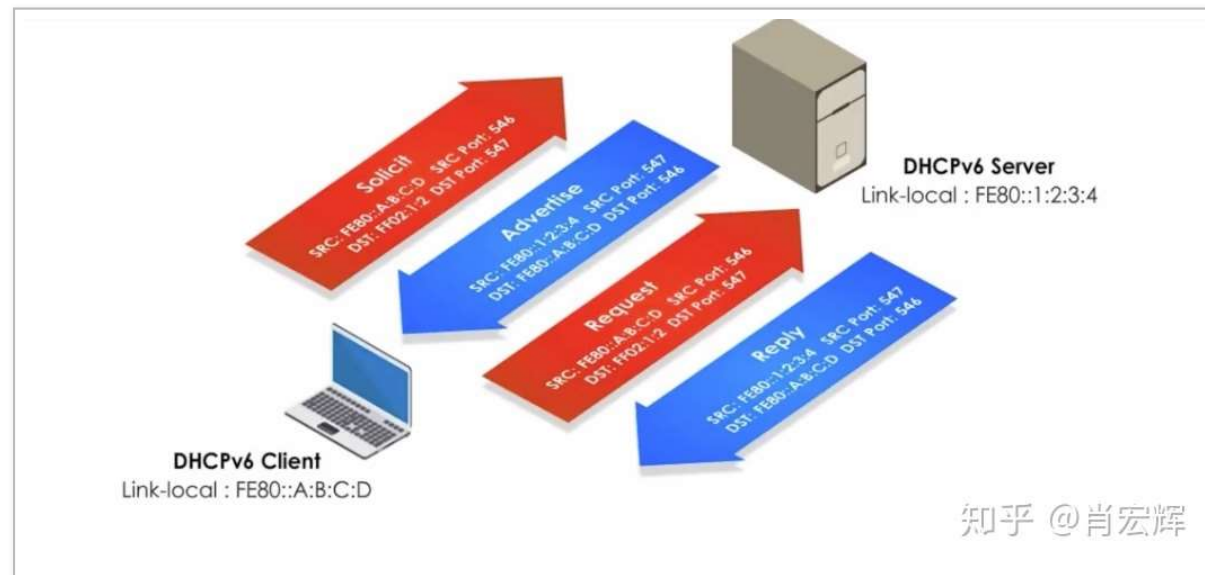
因为 IPv6 中没有广播的概念，所以不像 DHCP 基于广播地址 255.255.255.255 来发现 DHCP Server，DHCPv6 中，有两个保留的组播地址，用来发现网络中的 DHCP Server：

- All_DHCP_Relay_Agents_and_Servers (ff02::1:2)
 - DHCPv6 client 使用这个地址将 DHCP 请求发送给所有的 DHCPv6 relay（中继）agent 和 DHCPv6 server，这样可以发现网络中的 DHCPv6 Server
- All_DHCP_Servers (ff05::1:3)
 - DHCPv6 relay agent 通过这个地址将 DHCP 请求转发给所有的 DHCPv6 server

DHCPv6 仍然是基于 UDP 协议，但是使用的是 UDP 的 547（Server 监听端口）和 546（Client 监听端口）。与 IPv4 协议不同的是，IPv6 协议下，每个网卡都默认带一个 link-local

地址，这个地址是 fe80::/10 的前缀加上（一般情况下）网卡的 MAC 地址生成 [RFC4291]。IPv6 的 link-local 用来在一个二层链路中唯一标识一块网卡，并且可以在有限场景下在一个二层链路中用来通信。

在 DHCP (IPv4) 协议下，DHCP Client 在发起请求的时候，因为还没有 IP 地址，所以源 IP 只能是 Unspecified Address (0.0.0.0)。而 DHCPv6 Client 在发起请求的时候，网卡已经有 IPv6 地址了，所以源 IPv6 地址就是网卡的 link-local 地址，目的地址是 DHCPv6 保留的组播地址 ff02::1:2。



上面就是 DHCPv6 的交互过程。DHCPv6 Server 在收到请求之后，将 IPv6 地址在单播回给 Client 网卡的 link-local 地址。其中 IPv6 地址包含在 Advertise 和 Reply 中，这个过程与 DHCP 类似。

除了上面实现细节上的差别，DHCPv6 与 DHCP 的区别还在于，DHCPv6 不支持传输子网掩码长度，路由，和默认网关。这个区别是由 IPv4 迁移至 IPv6 容易引起错误的地方。

IPv6 动态地址配置模式

动态地址配置对于 IPv6 来说尤其重要，因为在 IPv4 场景下，地址长度，地址数量有限，静态配置还有可行性，但是如果要真正的完全使用 IPv6，通过静态配置地址基本不可能。尽管 IPv6 下面已经有 DHCPv6 协议，尽管 DHCPv6 已经与 DHCP 不兼容，但是 IPv6 和 IPv4 动态地址配置的差别远比 DHCPv6 和 DHCP 的差别大。

两个协议，三种模式

--

IPv6 的动态地址配置主要依赖两个协议，一个是 DHCPv6 (RFC8415)，另一个是 IPv6 Stateless Address Autoconfiguration (RFC4862)。IPv6 的动态地址配置方式客观的说是合理的，并且一定程度结合了 IPv4 动态地址配置的经验（个人猜测）。

在现实场景中，为了让一个网络工作，除了将设备连接起来的线路，交换机，我们主要还需要两个设备，一个是 DHCP Server，主要用来管理、分配地址和配置信息，另一个就是 Router 路由器，主要用来做三层转发，连接当前网络和其他网络。实际中，经常将 DHCP Server 配置在路由器上，或者路由器作为一个 DHCP relay agent。另一方面，如果没有路由器，网络只是一个二层网络，作用有限。因此，这两个设备本身可以只是一个设备，并且它们之中，路由器占主导地位。基于这个背景，IPv6 的动态地址配置有三种模式：

- SLAAC, Stateless Auto Address Configuration
- Stateless DHCPv6
- Stateful DHCPv6

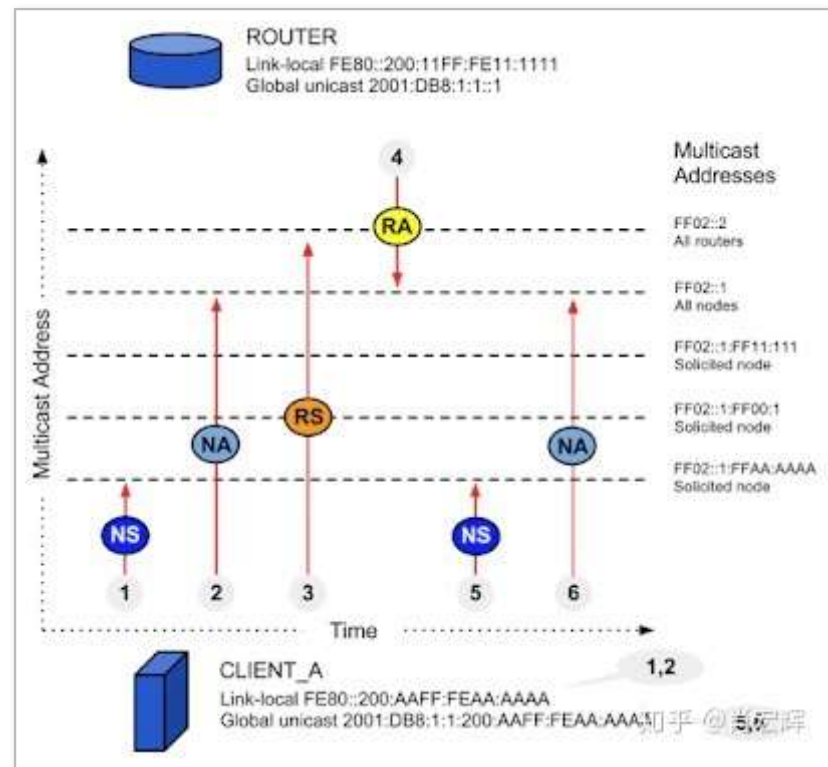
这里的 Stateful，指的就是 DHCP Server 管理的 IP 地址，因为这些地址存在一个分配关系，需要一个程序去管理这个状态。Stateless Address 是指这个地址就是分配给某一个确定的主机使用，没有其他状态。

SLAAC

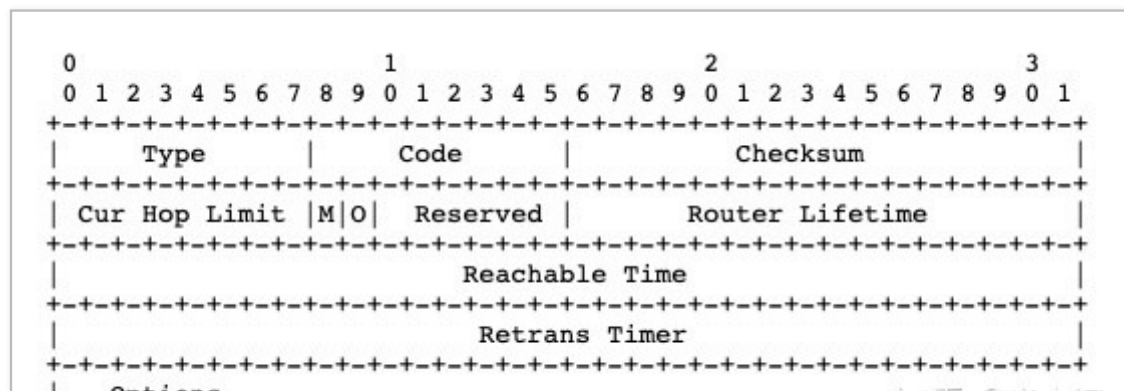
SLAAC 基于协议 RFC4861 和 RFC4862。在 SLAAC 的世界里，没有 DHCPv6。SLAAC 协议由路由器来通告配置 IPv6 地址所需要的信息。具体工作流程是这样：支持 IPv6 的网卡启动的时候会发送一条 RS（Router Solicitation）消息，源 IP 是网卡的 link-local 地址，目的 IP 是 ff02::2。ff02::2 也是保留的组播地址，用来表示所有的路由器。这条消息用来查找当前网络中的路由器。

路由器收到这条消息之后，会回传一条 RA（Router Advertisement）。一般情况下，RA 的源 IP 地址是 Router 的 link-local 地址，目的地址是 ff02::1。ff02::1 也是一个保留的组播地址，用来表示所有的主机。也就是说任意网卡发起的 RS 消息，都会引起路由器将 RA 消息在整个网络中发送给所有的主机。除此之外，就算没有任何 RS 消息，路由器也应当定期向所有主机发送 RA。具体流程如下图的 3，4。





RA 消息是 SLAAC 协议的核心，它的格式内容如下：



RA 可以发送给网络中所有主机的基础就是，不像 DHCP 消息，RA 是无状态的，任何主机接收到了 RA 消息之后，都能根据其中的信息完成 IP 地址配置。RA 的 options 中通常包括：

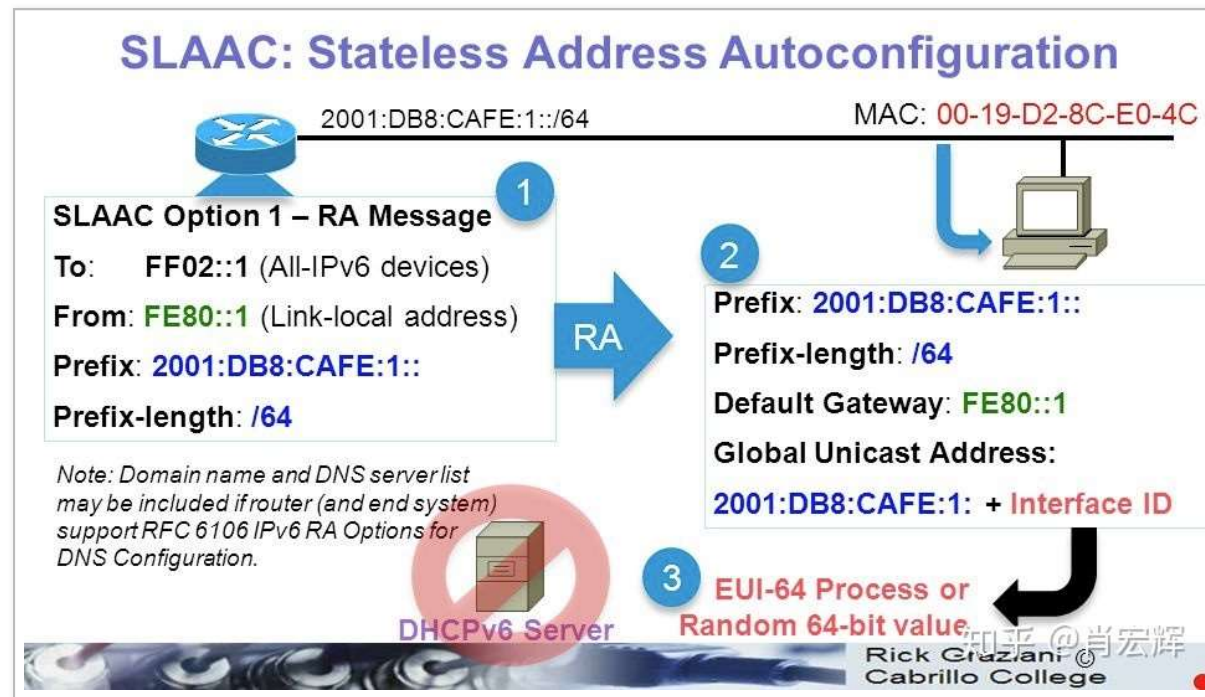
- MTU
 - 主机可以根据这个 MTU 值配置自己的 MTU
- 路由器的 MAC 地址
- 0 或者 N 个 prefix（网段）

如果 RA 中一个 prefix 的 auto 标志位是 1，那相当于告诉网卡，可以自己从这个 prefix 中生成一个 IPv6 的地址。网卡只需要使用一个局域网中唯一的标识符，再加上这个 prefix，就能生成一个 IPv6 地址。网卡在局域网中的天然唯一标识符就是 MAC 地址，因此一般情况下，SLAAC 协议中，网卡是通过 RA 中的 prefix 和自身的 MAC 地址，再根据 EUI-64 格式生成一个 IPv6 地址。SLAAC 协议的基础就是 IPv6 地址长度足够大，能支持这样的配置。

RA 消息还能带来一个潜在的配置，主机会将 IPv6 的默认网关，指向 RA 的源 IP 地址，也就是 Router 的 link-local 地址。生成的默认路由是有时效的，时间是 Router Lifetime。所以需要路由器定时发布 RA（类似于心跳），更新主机的默认路由。这个机制讲道理是合理的，默认路由只有在路由器还活着的时候才有意义。但是这个机制增加了管理的复杂度，尤其在 SDN 场景，虚拟路由器下定时发布 RA。

SLAAC 协议有两个问题，一个是 IPv6 的地址不可控了，一般情况下与主机的 MAC 地址相关；另一个是 RA 格式比较简单，能传递的配置有限，一些复杂的主机配置无法通过 RA 传递。

但是好处是简单，只用路由器，不需要 DHCPv6 的介入，就可以完成简单的 IP 地址配置；另外无状态地址能简化管理，只需要一些简单的程序就能完成 Prefix 分发，不需要集中的管理地址，处理多节点数据同步，这有点像分布式架构。SLAAC 的配置过程如下所示：



Stateless DHCPv6

Stateless DHCPv6 就是结合了 SLAAC 和 DHCPv6。其中 IPv6 地址配置通过 SLAAC 下发，其他配置通过 DHCPv6 下发。这样能弥补 SLAAC 模式下，RA 所能传递配置有限的问题。

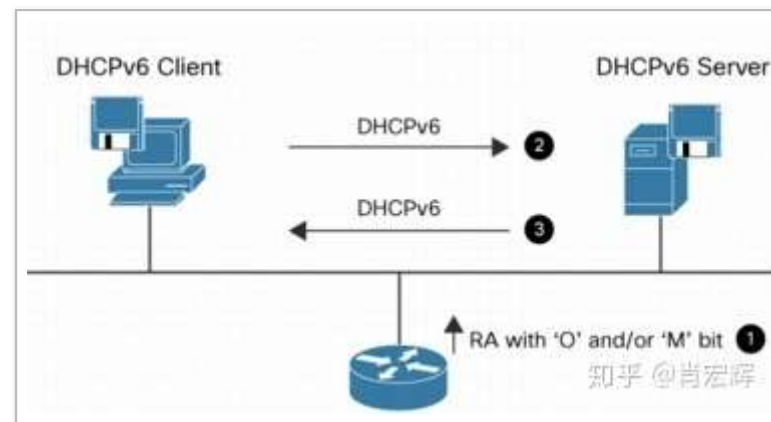
Stateful DHCPv6

这里就是纯纯的 DHCPv6 了，但是前面说过，DHCPv6 不支持子网掩码长度，路由，和默认路

由。这样就导致通过 DHCPv6 获得的 IPv6 地址，因为没有掩码长度，地址的掩码都是 128 位。而路由和默认网关，还是需要依赖 RA。因为在 IPv6 中，动态配置路由和默认网关的唯一方式就是使用 Router Advertisement 消息。所以，就算使用了 DHCPv6，也不能完全摆脱 RA 消息。只是说 DHCPv6 下，IPv6 地址变得可控了。所以，在 IPv6 动态地址配置中，Router Advertisement 是必不可少的，因为它控制了主机路由和默认路由。同时 Router Advertisement 还控制了 IPv6 的动态地址配置模式。在 RA 的数据中，有两个标志位 M (Managed) 和 O (Other)。

- 当 M=0, O=0 时，IPv6 地址工作在 SLAAC 模式下
- 当 M=0, O=1 时，IPv6 地址工作在 Stateless DHCPv6 模式下
- 当 M=1 时，IPv6 地址工作在 Stateful DHCPv6 模式下

总的来说，Stateful DHCPv6 的工作流程如下所示：



IPv6 中，如果需要地址可控，那只能使用 DHCPv6。如果只是需要分配一个 IPv6 地址，对地址无要求，那么使用 Stateless 地址分配方式更为合理。

使用了 全新的简悦词法分析引擎^{beta}，[点击查看详细说明](#)

