## IPv6基本认识

The following private IPv4 addresses are defined by RFC 1918:

■ 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)

RFC 6598, IANA-Reserved IPv4 Prefix for Shared Address Space

定义了一个新的私有地址块100.64.0.0/10,用于运营商使用,称为Shared Address Space,常用于CGN(carrier-grade NAT)技术.

-------------

#### NAT遇到的问题:

1.校验和重计算---TCP数据段中的IPv4报头会被修改,意味着IPv4校验和也要重新计算;

2.ICMP伪造:许多ICMP消息（例如Destination Unreachable消息）在其有效负载中嵌入了导致生成ICMP消息的原始IPv4标头。由于原始IPv4地址是由NAT转换的，因此NAT设备也必须转换这些地址;

3.IPsec问题: NAT无法和传输模式下的IPsec一起工作;

4.端到端可达性:若两端都使用NAT及私有地址,则必须有一个公有的中间端服务器;

5.性能:当使用多级NAT时,会造成延时等问题.

6.安全性:使用NAT时,会限制某些安全feature的使用,增加问题排查的难度;通常在具有防火墙功能的设备上使用NAT.

------------

IPv5

Internet Stream Protocol (ST)以及ST2虽然未被正式定义为IPv5,但是当ST协议封装在IPv4报文中时,IP协议字段为5.

IPv6误解

IPV6目前并没有增强QoS,即使IPv6拥有新的流控标签字段Flow Label field,但是现在的很多应用并没有充分使用它.

#### IPv6 Address类型简单介绍

**Global Unicast Address (GUA)全球单播地址**

全球唯一的可路由的IPv6地址,相当于IPv4公有地址, IPv6全球单播地址以十六进制的2/3开头,可作为源/目的地址,如2001:db8:cafe:1::100即为一个GUA地址.

**Link-Local Unicast Address链路本地地址**

链路本地地址是一个单播地址,仅在本地链路上,链路指逻辑上的网络段/子网,该地址仅在特定链路有效,且不可在超出该特定子网的范围路由;

每个IPv6设备必须有一个link-local unicast address,该地址会在启用了IPv6的接口下自动生成,

Link-local地址通常以fe80开头,如fe80::a299:9bff:fe18:50d1.

Link-local地址也可以作为源/目的地址.

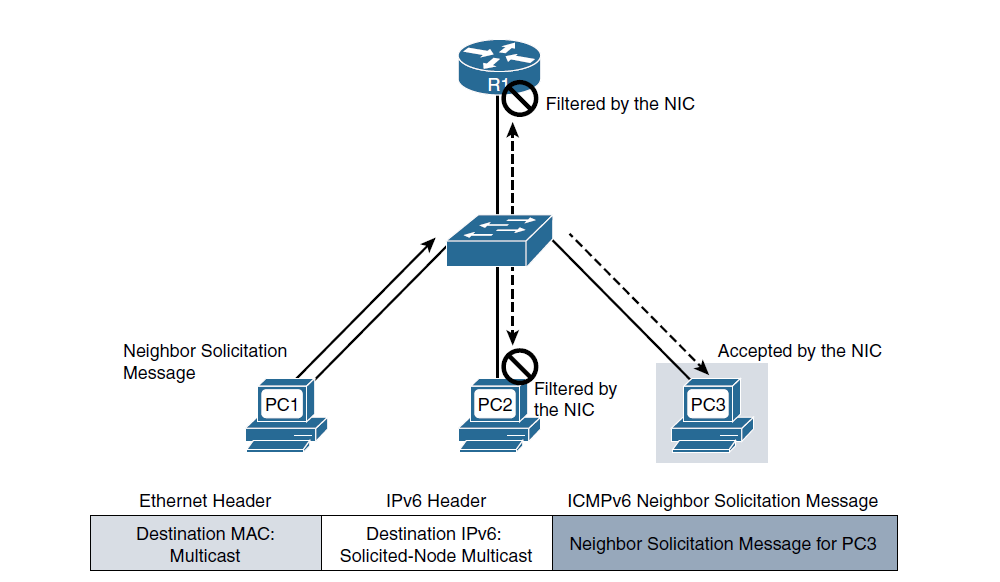
**Unspecified Address未指定地址**

即全0的IPv6地址,表示缺失或匿名地址,仅可以作为源地址,不可被路由器转发.

**Solicited-Node Multicast Address请求节点组播地址**

该地址是一个特殊的IPv6地址,用于某些IPv6协议取代ipv4中的广播地址, IPv6没有广播地址,但是有可以通告所有IPv6设备的地址, 为ff02::1.

#### ICMPv6 NDP(Neighbor Discovery Protocol**)**



NDP协议类似于IPv4中的ARP协议, 其中目的IP地址为一个Solicited-Node Multicast地址, 并映射成一个特殊的目的MAC组播地址 (\*ARP中,以太网帧的目的地址为广播地址, ARP报文中的目的MAC地址为全0地址)

Solicited-Node Multicast地址只能作为目的地址,路由范围不能超出子网,从ff02:0:0:0:0:1:ff/104开始.

IPv6常用术语:

Prefix: 前缀,和IPv4相同,表示网络位;

Prefix-Length: 同IPv4, 对于ipv6,前缀长度范围从/0到/128;

Interface-ID: 同IPv4主机位.

IPv6地址= Prefix + Interface ID

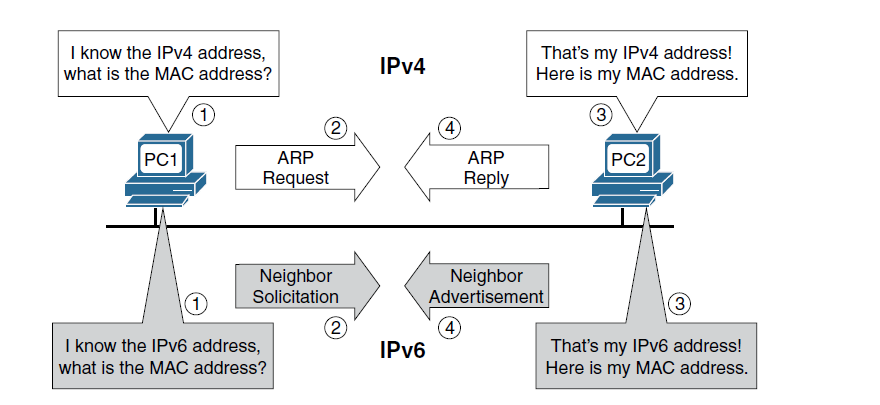
**ICMPv6 NDP(Neighbor Discovery Protocol)**

NDP协议用于网络段(segment/subnet)的设备发现和消息传递, 包含五种消息类型:

Router Solicitation, Router Advertisement, Neighbor Solicitation, Neighbor Advertisement, Redirect Messages.

**Neighbor Solicitation (NS) and Neighbor Advertisement (NA) Messages**

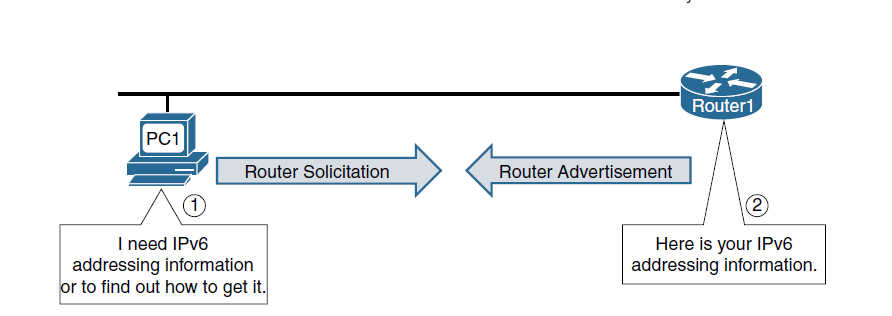
NS和NA消息用于同一子网任意两台设备的通信, 类似于IPv4中ARP的请求和应答消息,



**Router Solicitation (RS) and Router Advertisement (RA) Messages**

RS和RA消息用于设备和路由器在同一子网的通信.

RA消息由路由器发送, 建议IPv6设备动态的获得IPv6地址信息, RS消息由设备发送,用于请求路由器的RA应答.



#### 动态地址分配

IPv4地址获取方式:

静态/手工配置;

从DHCPv4服务器动态获取(地址信息一般包含: IPv4地址,子网掩码,默认网关地址,域名,域名服务器地址).

IPv6地址获取方式:

静态/手工配置;

IPv6路由器周期性的发送RA消息(在IOS中为200s), 通告IPv6地址信息; 或者路由器收到一个RS消息时, 触发RA消息. RA消息的目的地址为全IPv6组播地址ff02::1, 该地址会被同一子网的所有IPv6设备接收, RA消息不会被路由器转发. (IPv6 RA消息包含: 前缀, 前缀长度, 默认网关地址---该地址为路由器出接口的link-local地址,也是RA消息的源地址, 还有3个flag标志位, Autonomous Address Configuration Flag (A flag), the Other Configuration Flag (O flag), and the Managed Address Configuration Flag (M flag), 以及可选信息域名/DNS服务器地址). 故即使没有DHCPv6服务器, IPv6设备也能动态寻址.

**RA消息的三种方式**:

1. Stateless Address Autoconfiguration (SLAAC): IOS默认方式, 设备从RA消息中获取所有地址信息, 会使用IPv6前缀去创建一个global unicast address (64位的interface id的IPv6地址), 使用RA消息的源IPv6地址作为默认网关;
2. SLAAC and stateless DHCPv6 server: 在SLAAC的基础上,从stateless DHCPv6服务器上获取RA消息中没有的地址信息, stateless DHCPv6 server不提供和维护global unicast address, 仅提供所有IPv6设备的公共信息,如 DNS服务器地址等. RA消息并没有定义什么地址信息在stateless DHCPv6 server中是可用的, 仅当路由器和接收RA消息设备都支持RFC 6106时, RA消息才包含DNS地址信息.
3. Stateful DHCPv6 server: 类似于DHCPv4, RA消息通告设备从DHCPv6服务器获取地址信息, 甚至包含global unicast address, 但是设备必须使用方法1/2从RA消息中获取默认网关地址.

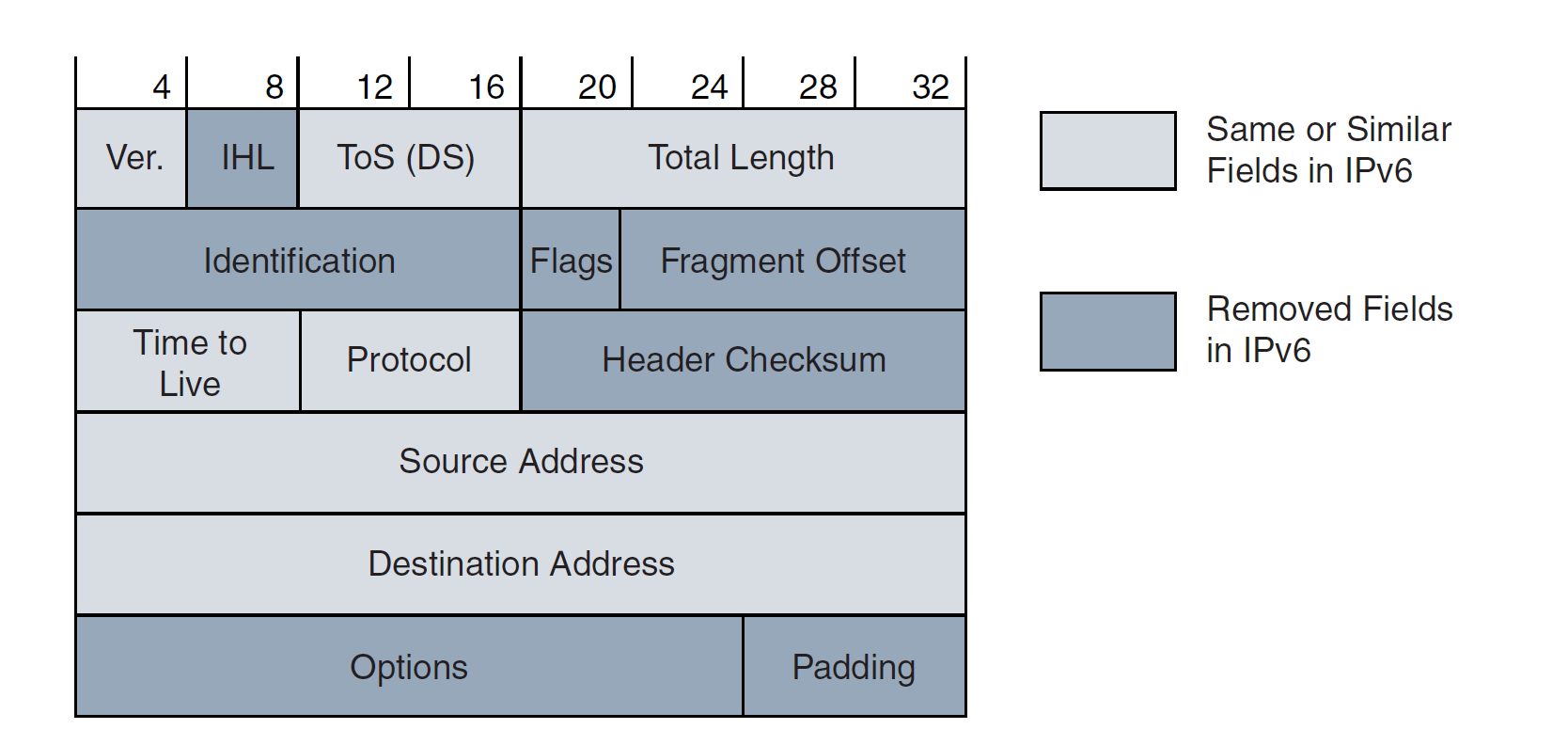
Global Unicast Address地址在SLAAC中的两种生成方式：

1. 随机64位Interface-ID: 系统随机生成64位的interface ID，是windows系统的默认方式；
2. EUI-64（唯一拓展标识符）: 在接口的48位的MAC地址中间插入16 bit fffe, 再将从左到右第7位置位，得到新的Interface ID，是Linux和Mac OS的默认方式，如接口MAC地址为00:0C:85:AB:50:01， 则生成的interface ID为020C:85FF:FEAB:5001。
3. \*第三种方式，使用Cryptographically Generated Addresses (CGA) 方法和Cryptographically Generated Addresses (CGA) 加密哈希出Interface ID。

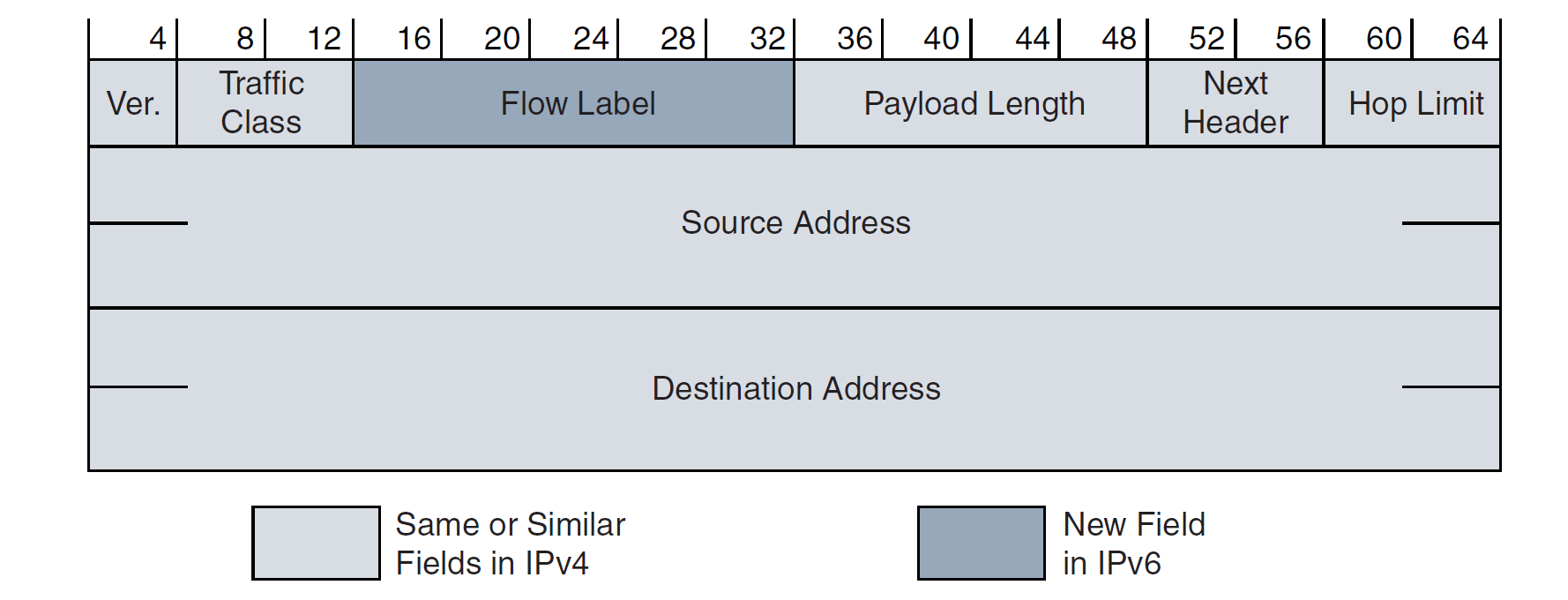
通常一些操作系统会同时使用SLAAC和Stateful DHCPv6方式生成2个IPv6 GUA地址，可以配置RA消息，使其不分配GUA地址。

## IPV4和IPv6报文比较

**IPv4报头**



**IPv6报头---main IPv6 header**



IPv6报头常见为main IPv6 header, 同时包含多种拓展IPv6报头---使用某些feature时，会使用这种报头。

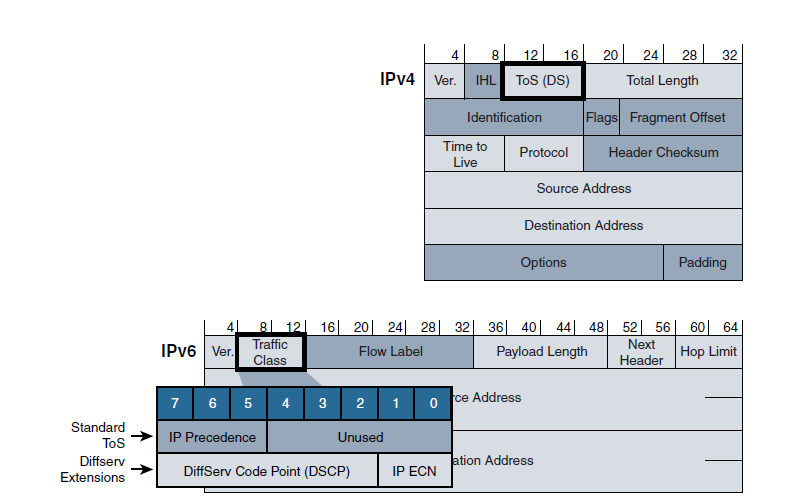
IPv4报头长度为20字节到60字节；IPv6报头字节为固定40字节（main IPv6 header）。

IPv4报头宽度为32bit， 而IPv6为64bit

**version**字段：IPv4为4，IPv6为6，在1990年时， Internet Stream Protocol (ST2) 封装在IPv4中，使用字段5.

**IPv4 Internet Header Length** (IHL)字段：ipv4报头的总长度，包括opthions和padding的长度。范围为20到60字节。

**IPv4 Type of Service (ToS)** 字段和**IPv6 Traffic Class** 字段：相同的字段，仅名字不一样。

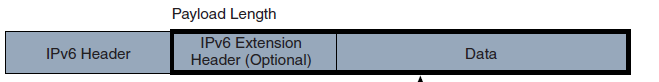


此字段用于QoS特性。

**IPv6 Flow Label**字段：用于标记从一个源节点发往一个或多个目的地节点的IPv6数据包的序列或flow。这个流标签能标记需要路由器特殊处理的数据包序列，如real-time服务。 Flow Label字段用于帮助识别同一流中的所有数据包，以确保所有数据包都接收IPv6路由器的相同类型的处理。Flow label设为0，表示该流量与任何flow都无关。

**IPv4 Total Length**字段：单位为byte，表示整个IPv4包的长度（包括ipv4报头和data长度），字段长度为16bit，故ipv4包的最大长度为65535 bytes。

**IPv6 Payload Length**字段： 字段长度16bit， 不包含固定IPv6报头长度； IPv6 jumbogram特性允许IPv6数据包的payload大于65535 bytes，借助Hop-by-Hop拓展报头，Jumbo Payload可以拓展到32bit，允许IPv6的payload最大到2的32次方-1 bytes大小。



**IPv4 and IPv6 MTU**：整个IP数据包的大小，包括报文和payload data。

IPv4要求最小未分段转发数据包大小为68bytes = 最大60字节报头 + 8字节有效payload，且最小能接收的数据包为576字节（可以是完整数据包，也可以是数据分段）。

IPv6最小的MTU为1280字节，建议设置为1500字节。

**IPv6数据分片**：不同于IPv4，IPv6路由器不会对数据包分段，除非该数据包的源头就是该路由器，所以中间节点的路由器不会对ipv6分片。IPv6固定报头中并没有分片字段，IPv6使用Extension Hearder报头分段。当IPv6路由器收到一个大于其MTU的包时，会直接丢弃该IPv6数据吧并返回一个ICMPv6 Packet Too Big消息。Path MTU (PMTU) feature可以检测到链路上的最小MTU。

**IPv4 Protocol and IPv6 Next Header Fields**

IPv4协议字段和IPv6 Next Header字段：两者类似，但是IPv6多了 Extendsion Header的类型字段。

**IPv4 Time to Live (TTL) and IPv6 Hop Limit Fields**

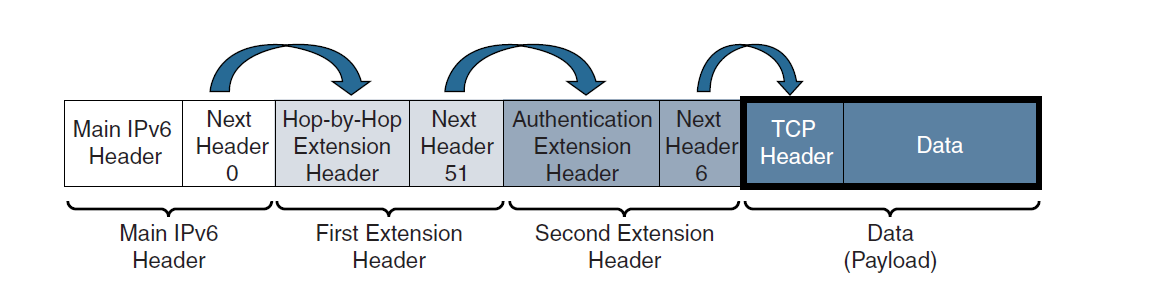
都表示路由器跳数，一跳表示一秒。

**IPv4 Checksum**

校验和占16位，用于检验数据包固定大小的计算，并不同于以太网帧的CRC校验。IPv6没有校验和字段，所以在UDP中可选的16位的checksum字段在IPv6中是强制必须的。

**IPv6 Extension headers**

Extension header跟随在IPv6固定报头后，它也有Next header字段表示后接的是另一个extension还是其他协议的payload。No Next Header是一个特殊的拓展报头，表示后面没有拓展报头或是协议payload。



Extension header功能和IPv4报头的option字段功能类似，当IPv6需要某些拓展或者增强性feature时，会使用extension header， extension header也是64位的位宽，当不足64bit时，会和IPv4报头的padding字段一样（IPv4填补到32位），填补到64位。

## IPv6书写规则（hextet十六进制段表示）

1. 省略前面的0. 仅可以省略十六进制段的开头的0，结尾的0（首数字不为0）不能省略，

如：fe80: 0000: 0000: 0000:a299:9bff:fe18:50d1

可以表示为： fe80:0:0:a299:9bff:fe18:50d1

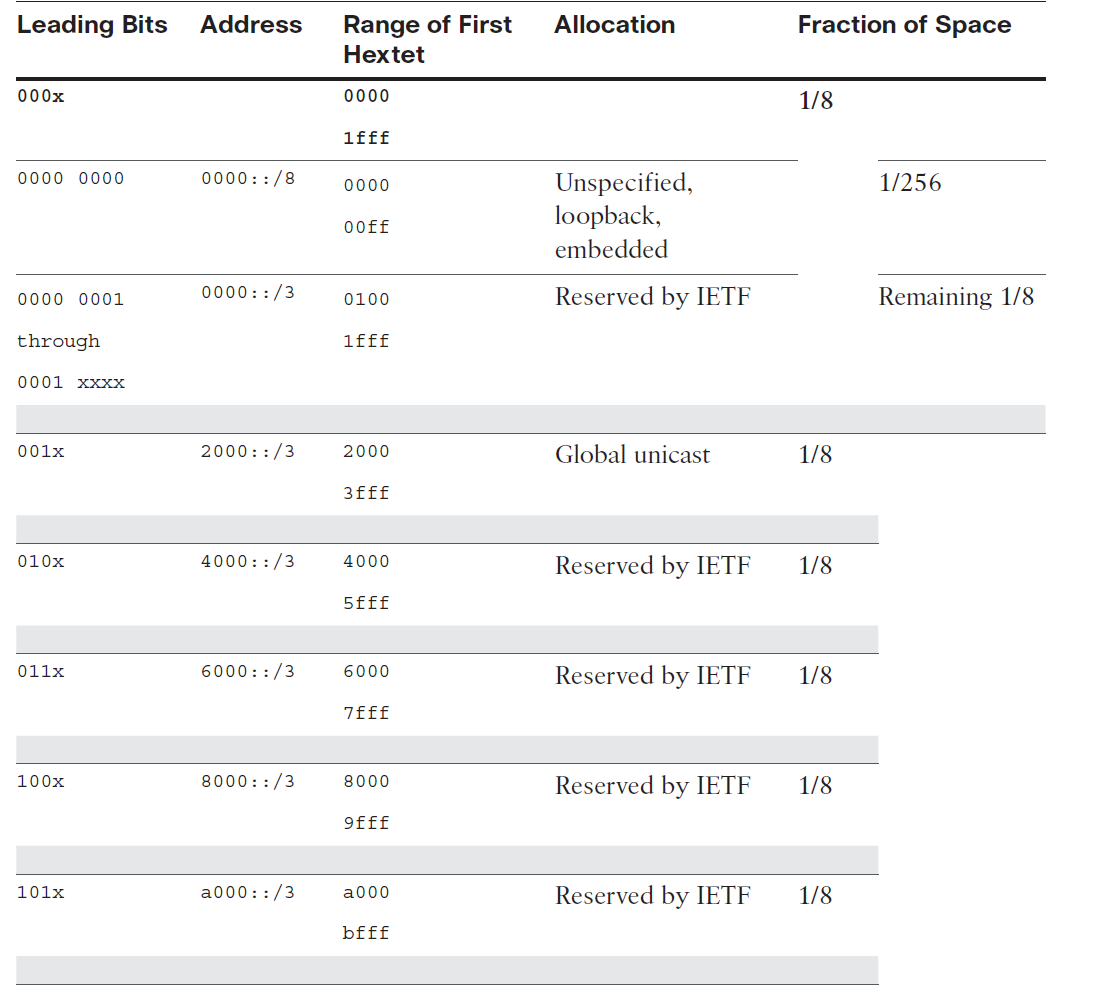
2. 全0的十六进制段可以省略，但是整个IPv6地址中，只能使用一个::符号，且当有多个可以省略的十六进制段

时，只有第一个段可以使用::，

如：0000:0000:0000:0000:0000:0000:0000:0000

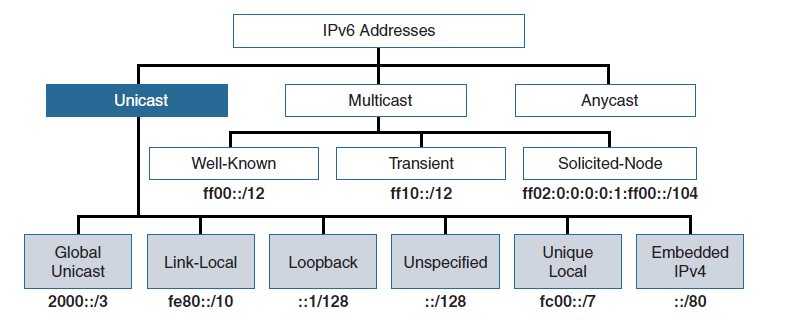
可以表示为 ::

## IPv6分配地址空间

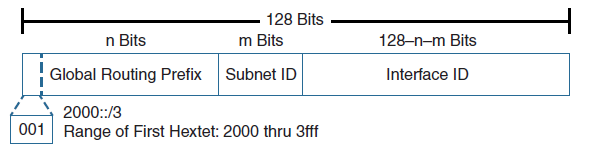




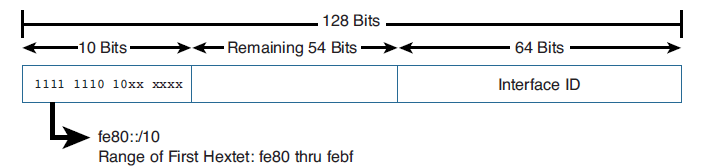
**Unicast Address单播地址**



GUA地址: 即IPv6公网地址, 结构如下,



Link-Local Unicast Address: 该地址由IPv6接口自动且必须生成(也可以手动配置),且**仅在当前链路断/子网唯一**, 故路由器不同接口可以使用同一Link-Local Unicast Address, 但是**每个接口仅能拥有一个Link-Local Unicast Address**. RFC 4291 允许Link-Local Address使用fe80::/10之外的地址, 但是这样可能造成非预期的影响. 设备将路由器的Link-Local Unicast Address作为它们的默认网关. 路由器使用Link-Local Address交换动态路由协议消息, 在路由表中, Link-Local Address也作为动态路由协议的下一跳地址.



Loopback Address环回口地址: 类似于IPv4的127.0.0.1, 用于自身IPv6测试;

Unspecified Address未指定地址: 即全0地址, 只能作为源地址,表示地址信息未指定为空;

Unique Local Address(ULA) : 即IPv6私有地址, 不能在Internet上路由, 以为安全原因, 不建议将NAT和IPv6结合使用. ULA的前缀是fc00::/7, 可以分为:

**fc00::/8---为将来使用预留,**

**fd00::/8---现在使用的有效地址**, 其中第8位称为L flag/ Local flag ,仅当该位置位为1时, 即fd00时, ULA地址才有效.

ULA地址不同于IPv4私有地址的重要一点是, ULA地址通过RFC 4193定义其通过伪随机算法可以是全球唯一的.

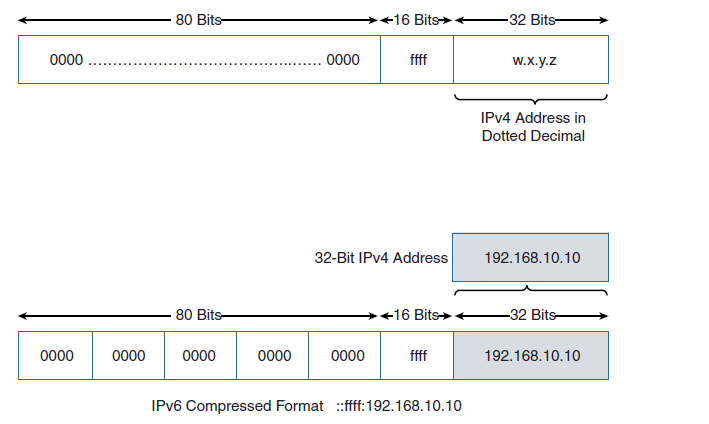
ULA地址替换了已经被弃用的Site-Local Address(RFC 3513).

IPv4 Embedded Address: 将IPv4转化为IPv6地址, 根据IPv6前缀长度的不同, IPv4嵌入的位置也随着变化. 可以参考 <https://tools.ietf.org/id/draft-ietf-behave-address-format-10.html>

RFC 4291定义了两种IPv4 Embedded Address:

IPv4映射IPV6地址---用于双栈场景,需要发送一个IPV6数据包到IPv4-only设备;

IPv4兼容IPv6地址(已弃用).



::ffff:192.168.10.10

通常是/96前缀映射, IPv4地址不必是global unique.