



G-CNP v2.0课程

讲师：沈老师





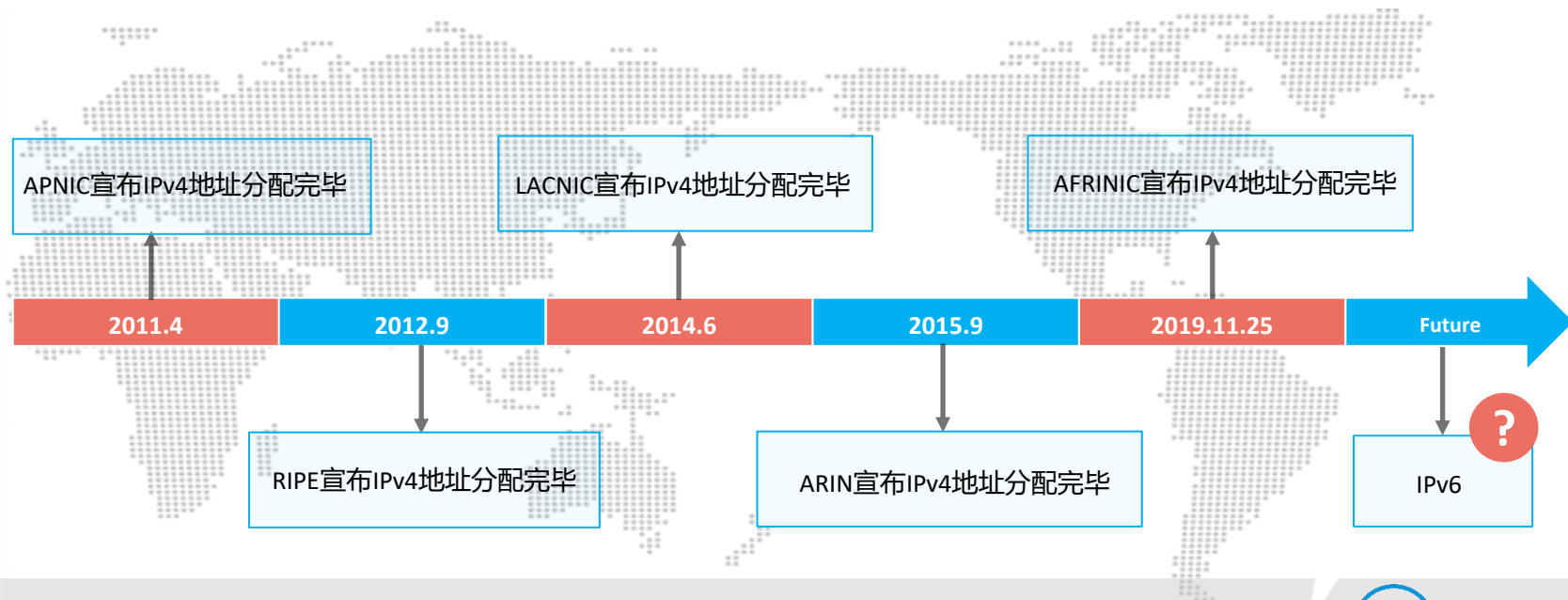
前言

- 20世纪80年代，IETF（Internet Engineering Task Force，因特网工程任务组）发布RFC791，即IPv4协议，标志IPv4正式标准化。在此后的几十年间，IPv4协议成为最主流的协议之一。无数人在IPv4的基础上开发出了各种应用，并且对这个协议做了各种补充和增强，支撑起了今天繁荣的互联网。
- 然而，随着互联网的规模越来越大，以及5G、物联网等新兴技术的发展，IPv4面临的挑战越来越多。IPv6取代IPv4势在必行。
- 本章节描述了为什么需要从IPv4向IPv6进行演进，以及一些关于IPv6的基础知识。



IPv4现状

- 2011年2月3日，IANA（Internet Assigned Numbers Authority，因特网地址分配组织）宣布将其最后的468万个IPv4地址平均分配到全球5个RIR（Regional Internet Registry，区域互联网注册管理机构），此后IANA再没有可分配的IPv4地址。





全球IPv6发展现状

全球IPv6部署率显著增长

截止到2019年10月，综合IPv6部署率在30%以上的国家或地区占了全球地图面积
一半以上

全球IPv6用户数量猛增

根据Google网站监测，至2019年9月，使用IPv6访问Google网站的用户占总用户最高
已超过30%
相比2018年涨幅**35.3%**

主流软件、服务IPv6支持度提高

- 截止到2019年10月10日，全球所有网站中有14.6%的网站支持IPv6访问，排名前100万的网站中有19.2%，排名前1000的网站IPv6支持率接近30%。
- 全球主要的云服务商和CDN运营商Cloudflare、阿卡迈、微软云服务、亚马逊云服务等均已支持IPv6。
- 面向大众用户的操作系统Windows、MacOS已支持IPv6。

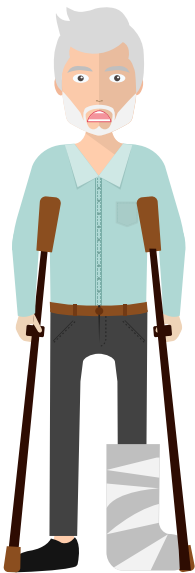
全球网络及域名系统IPv6部署情况

截至2019年10月，在全球1527个顶级域中，有1505个支持IPv6，
占总量的**98.6%**



Why IPv6 ?

IPv4



公网地址枯竭
包头设计不合理
路由表过大，查表效率低
对ARP的依赖，导致广播泛滥

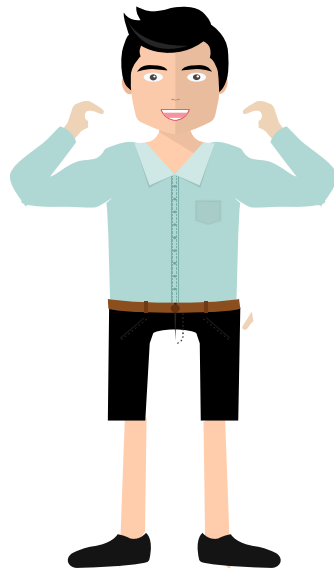
.....

VS

IPv6

“无限”地址
地址层次化分配
即插即用
简化的报文头部
IPv6安全特性
保证端到端通信完整性
对移动性的支持
增强的QoS特性

.....





IPv6优势

“无限”地址空间

地址长度为128 bit，海量的地址空间，满足物联网等新兴业务、有利于业务演进及扩展。

层次化的地址结构

相较于IPv4地址，IPv6地址的分配更加规范，利于路由聚合（缩减IPv6路由表规模）、路由快速查询。

即插即用

IPv6支持无状态地址自动配置（SLAAC），终端接入更简单。

简化的报文头部

简化报文头，提高效率；通过扩展包头支持新应用，利于路由器等网络设备的转发处理，降低投资成本。

安全特性

IPsec、真实源地址认证等保证端到端安全；避免NAT破坏端到端通信的完整性。

移动性

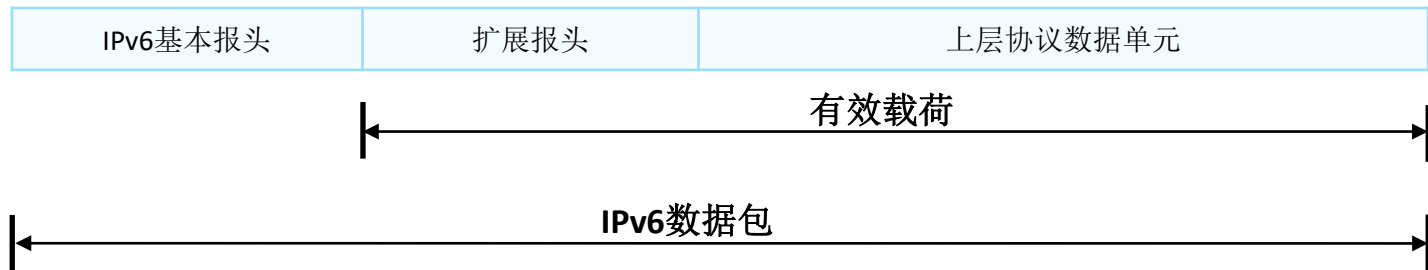
对移动网络实时通信有较大改进，整个移动网络性能有比较大的提升。

增强的QoS特性

额外定义了流标签字段，可为应用程序或者终端所用，针对特殊的服务和数据流，分配特定的资源。



IPv6报文构成



IPv6报文一般由三个部分组成：

- **基本报头：**提供报文转发的基本信息，路由器通过解析基本报头就能完成绝大多数的报文转发任务。
- **扩展报头：**提供一些扩展的报文转发信息，如分段、加密等，该部分不是必需的，也不是每个路由器都需要处理，仅当需要路由器或目的节点做某些特殊处理时，才由发送方添加一个或多个扩展头。
- **上层协议数据单元：**一般由上层协议报头和它的有效载荷构成，该部分与IPv4的上层协议数据单元相似。



IPv6基本报头

IPv4报头 (20Byte ~ 60Byte)

Version	IHL	ToS	Total Length	
Identification			Flags	Fragment Offset
TTL		Protocol	Header Checksum	
Source address				
Destination address				
Options				Padding

IPv6基本报头(40Byte)

Version	Traffic Class	Flow Label	
Payload Length		Next Header	Hop Limit
Source address			
Destination address			

IPv6头部相较于IPv4的改进

- **取消三层校验:** 协议栈中二层和四层的已提供校验, 因此IPv6直接取消了IP的三层校验, 节省路由器处理资源。
- **取消中间节点的分片功能:** 中间路由器不再处理分片, 只在产生数据的源节点处理, 省却中间路由器为处理分片而耗费的大量CPU资源。
- **定义定长的IPv6报头:** 有利于硬件的快速处理, 提高路由器转发效率。
- **安全选项的支持:** IPv6提供了对IPSec的完美支持, 如此上层协议可以省去许多安全选项。
- **增加流标签:** 提高QoS效率。

保留的字段

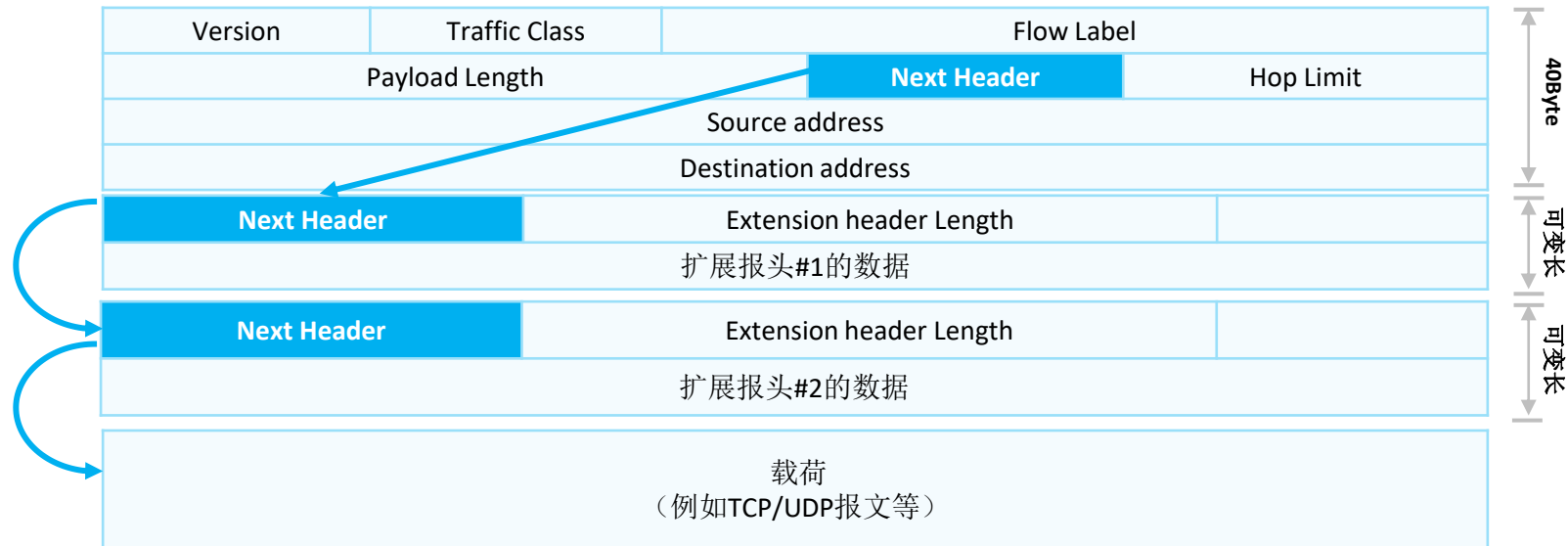
取消的字段

名字/位置变化

新增字段



IPv6扩展报头 (1)



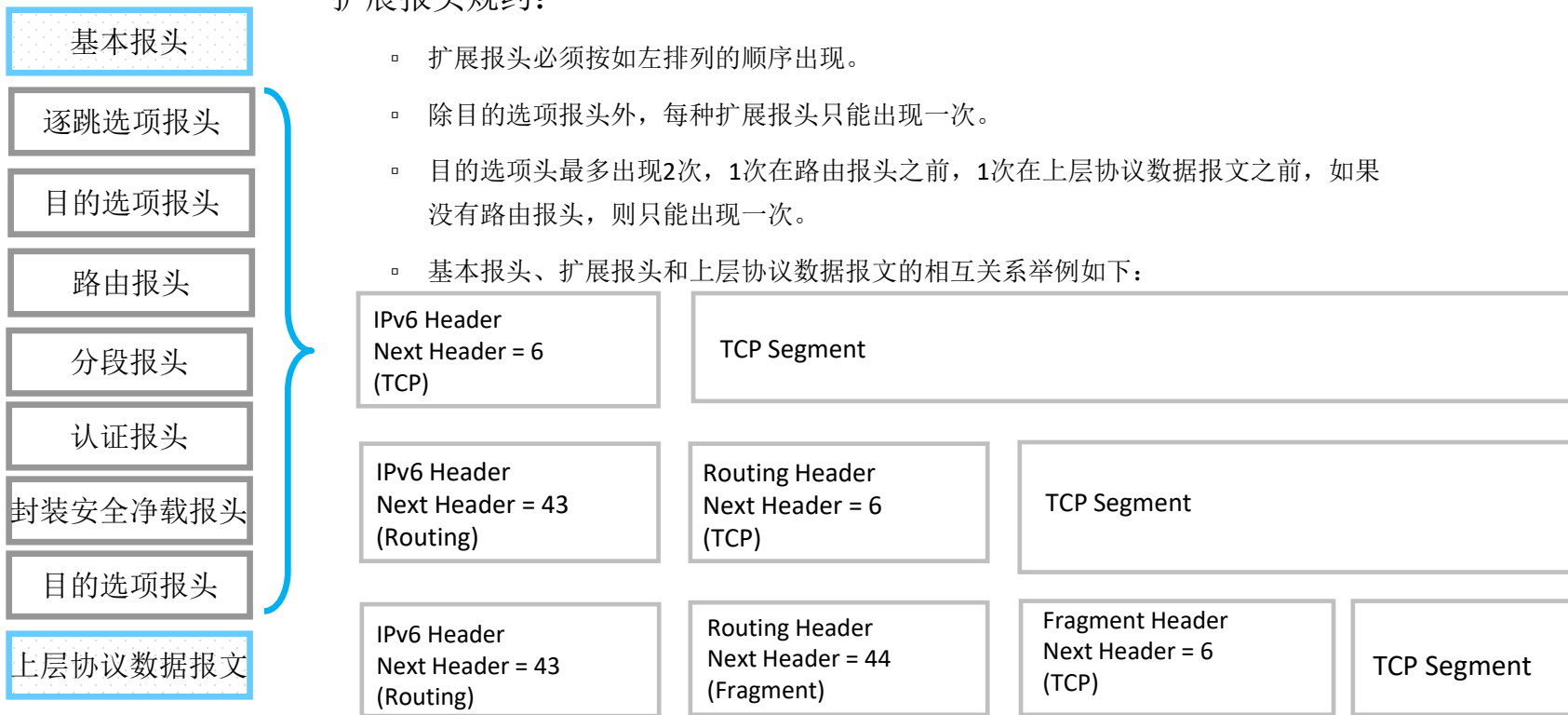
- **Next Header:** 下一报头，长度为8bit。与基本报头的Next Header的作用相同。指明下一个扩展报头（如果存在）或上层协议的类型。
- **Extension Header Length:** 报头扩展长度，长度为8bit。表示扩展报头的长度（不包含Next Header字段）。
- **Extension Header Data:** 扩展报头数据，长度可变。扩展报头的内容，为一系列选项字段和填充字段的组合。



IPv6扩展报头 (2)

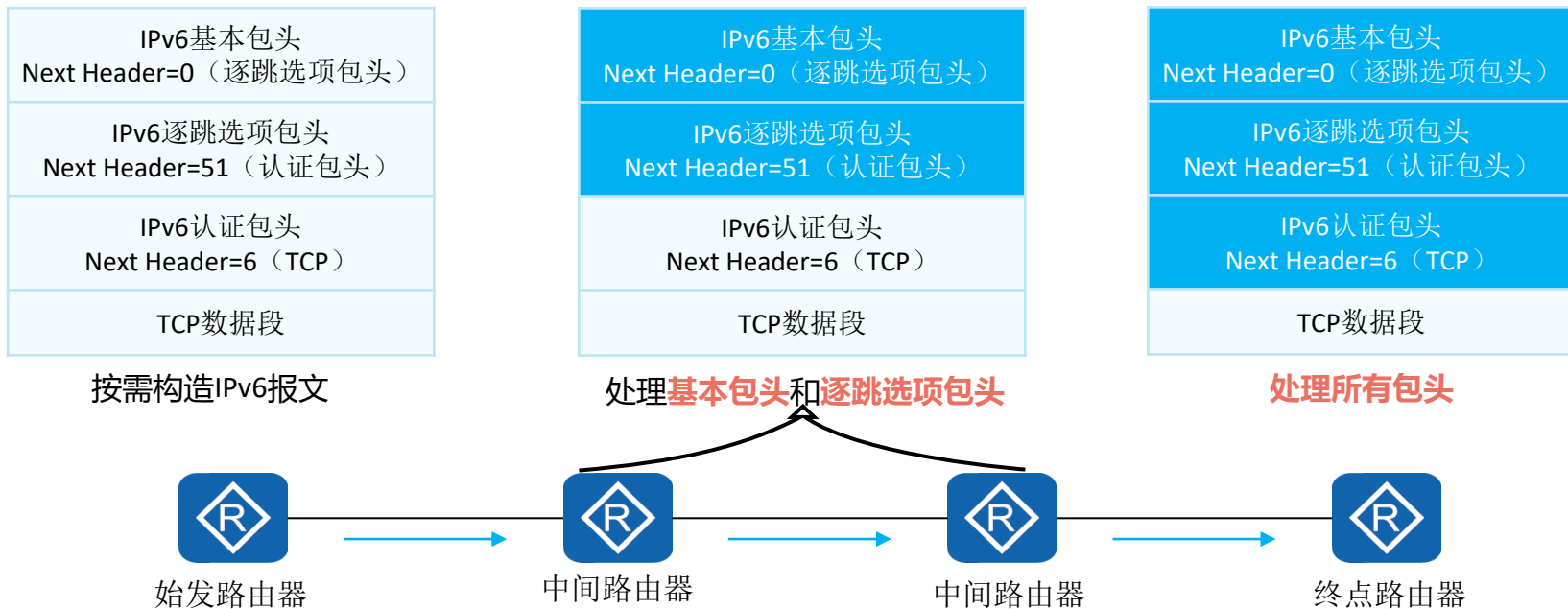
扩展报头规约:

- 扩展报头必须按如左排列的顺序出现。
- 除目的选项报头外，每种扩展报头只能出现一次。
- 目的选项头最多出现2次，1次在路由报头之前，1次在上层协议数据报文之前，如果没有路由报头，则只能出现一次。
- 基本报头、扩展报头 and 上层协议数据报文的相互关系举例如下：





IPv6报文处理机制

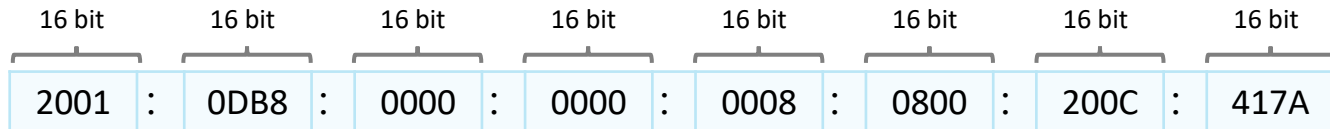


- 基本包头长度固定，提升转发效率！
- 扩展头部实现其他需求，术业有专攻！



IPv6地址

- IPv6地址的长度为128 bit。一般用冒号分割为8段，每一段16 bit，每一段内用十六进制表示。



IPv6地址中的字母大小写不敏感，例如A等同于a。

- 与IPv4地址类似，IPv6也用“IPv6地址/掩码长度”的方式来表示IPv6地址。
 - 例如2001:0DB8:2345:CD30:1230:4567:89AB:CDEF/64

IPv6地址： 2001:0DB8:2345:CD30:1230:4567:89AB:CDEF

子网号： 2001:0DB8:2345:CD30::/64



地址空间

- 为什么IPv6协议的地址长度是128bit？
 - CPU处理字长发展至今分别经历了4bit、8bit、16bit、32bit、64bit等，当数据能用2的指数幂字长的二进制数表示时，CPU对数值的处理效率最高。
 - IPv4地址长度为32bit，原因之一就是当时互联网上的主机CPU字长为32bit。从处理效率和未来网络扩展性上考虑，将IPv6的地址长度定为128bit是十分合适的。
- IPv6的128bit地址是一个什么概念？
 - IPv4有 (2^{32}) = 4,294,967,296个地址。
 - IPv6有 ($2^{128} = 2^{96} \times 2^{32}$) = 340,282,366,920,938,463,463,374,607,431,768,211,456个地址(340万亿万亿个地址)，相当于地球表面每平方米可以分配到67万亿个地址。
 - 夸张的说，地球上每一粒沙子都可以分配到一个IPv6地址。



IPv6地址缩写规范

- 为了书写方便，IPv6可采用以下规则进行缩写。

IPv6地址缩写规范

2001 : 0DB8 : 0000 : 0000 : 0008 : 0800 : 200C : 417A

每组16 bit的单元中的前导0可以省略，但是如果16 bit单元的所有比特都为0，那么至少要保留一个“0”字符；拖尾的0不能被省略。

2001 : DB8 : 0 : 0 : 8 : 800 : 200C : 417A

一个或多个连续的16 bit字符为0时，可用“::”表示，但整个IPv6地址缩写中只允许有一个“::”。

2001 : DB8 : : 8 : 800 : 200C : 417A

若缩写后的IPv6地址出现两个“::”，会导致无法还原为原始IPv6地址。

IPv6地址缩写示例

缩写前 0000:0000:0000:0000:0000:0000:0000:0001

缩写后 ::1

缩写前 2001:0DB8:0000:0000:FB00:1400:5000:45FF

缩写后 2001:DB8::FB00:1400:5000:45FF

缩写前 2001:0DB8:0000:0000:0000:2A2A:0000:0001

缩写后 2001:DB8::2A2A:0:1

缩写前 2001:0DB8:0000:1234:FB00:0000:5000:45FF

缩写后 2001:DB8::1234:FB00:0:5000:45FF

或 2001:DB8:0:1234:FB00::5000:45FF



IPv6地址结构

- 一个IPv6地址可以分为如下两部分：
 - 网络前缀：nbit，相当于IPv4地址中的网络ID。
 - 接口标识：（128-n）bit，相当于IPv4地址中的主机ID。
- IPv6单播地址示例：2001:0DB8:6101:0001:5ED9:98FF:FECA:A298/64。





IPv6地址前缀

- 鉴于IPv4地址在规划和分配上的局限性，IETF对IPv6地址类型进行了精细划分，不同类型的IPv6地址被赋予了不同的前缀，且受地址分配机构的严格管理。
- 现阶段，常用的IPv6地址或前缀有：

IPv6地址或前缀	含义
2001::/16	用于IPv6 Internet，类似于IPv4公网地址
2002::/16	用于6to4隧道
FE80::/10	链路本地地址前缀，用于本地链路范围内的通信
FF00::/8	组播地址前缀，用于IPv6组播
::/128	未指定地址，类似于IPv4中的0.0.0.0
::1/128	环回地址，类似于IPv4中的127.0.0.1
...	...



IPv6地址接口标识

- 接口ID可通过三种方式生成：手工配置、系统自动生成，或基于IEEE EUI-64规范生成。
- 其中，基于IEEE EUI-64规范自动生成接口ID的方式最为常用，该方式将接口的MAC地址转换为IPv6接口标识。

MAC地址（16进制）

08-70-5A-90-1A-01

MAC地址（2进制）

第7bit=0表示该MAC地址是全局管理地址

00001000 - 01110000 - 01011010 - 10010000 - 00011010 - 00000001

2 第7bit取反



1 在这个位置插入FFFE

00001000 - 01110000 - 01011010 - 11111111 - 11111110 - 10010000 - 00011010 - 00000001

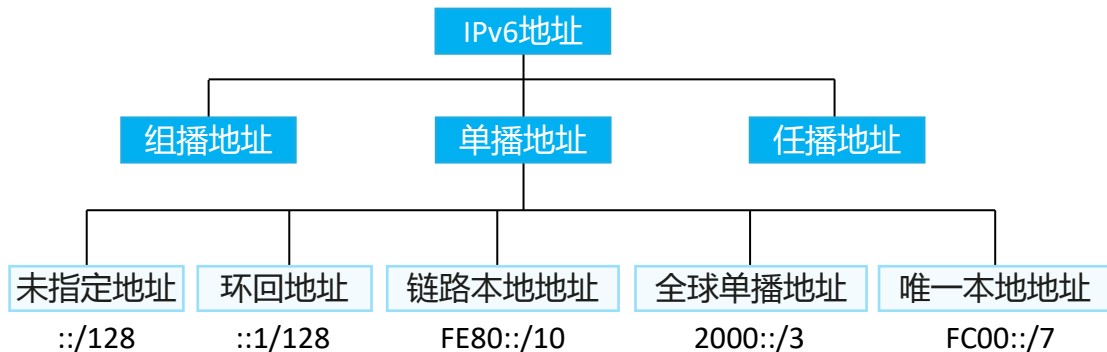
00001010 - 01110000 - 01011010 - 11111111 - 11111110 - 10010000 - 00011010 - 00000001

EUI-64规范的接口ID

0A - 70 - 5A - FF - FE - 90 - 1A - 01



IPv6地址类型

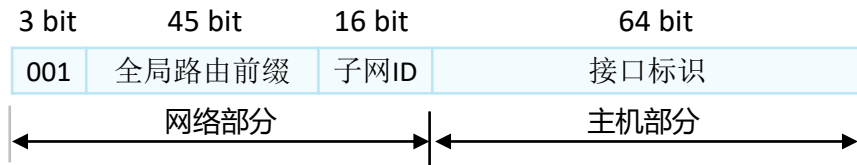


- **单播地址 (Unicast Address)**：标识一个接口，目的地址为单播地址的报文会被送到被标识的接口。在IPv6中，一个接口拥有多个IPv6地址是非常常见的现象。
- **组播地址 (Multicast Address)**：标识多个接口，目的地址为组播地址的报文会被送到被标识的所有接口。只有加入相应组播组的设备接口才会侦听发往该组播地址的报文。
- **任播地址 (Anycast Address)**：任播地址标识一组网络接口（通常属于不同的节点）。目标地址是任播地址的数据包将发送给其中路由意义上最近的一个网络接口。
- **IPv6没有定义广播地址 (Broadcast Address)**

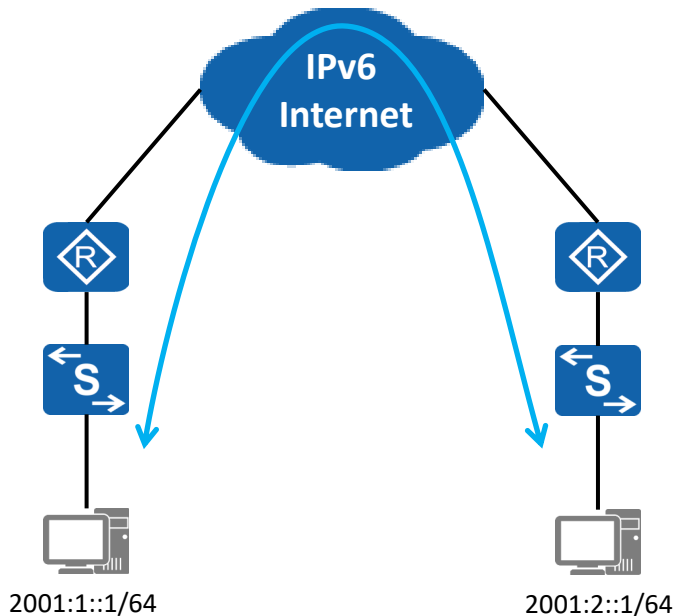


IPv6常见单播地址 - GUA

- GUA（Global Unicast Address，全球单播地址），也被称为可聚合全球单播地址。该类地址全球唯一，用于需要有互联网访问需求的主机，相当于IPv4的公网地址。



- 通常GUA的网络部分长度为64 bit，接口标识也为64 bit。
- 全局路由前缀：由提供商指定给一个组织机构，一般至少为45 bit。
- 子网ID：组织机构根据自身网络需求划分子网。
- 接口标识：用来标识一个设备（的接口）。





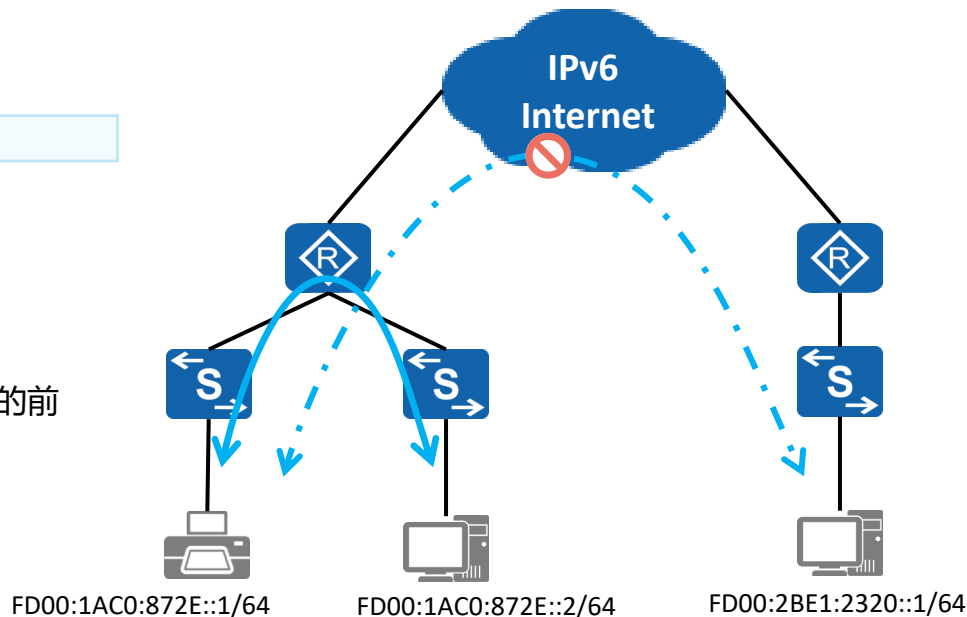
IPv6常见单播地址 - ULA

- ULA (Unique Local Address, 唯一本地地址) 是IPv6私网地址, 只能够在内网中使用。该地址空间在IPv6公网中不可被路由, 因此不能直接访问公网。

8 bit	40 bit	16 bit	64 bit
1111 1101	Global ID	子网ID	接口标识

伪随机产生

- 唯一本地地址使用FC00::/7地址块, 目前仅使用了FD00::/8地址段。FC00::/8预留为以后拓展用。
- ULA虽然只在有限范围内有效, 但也具有全球唯一的前缀 (虽然随机方式产生, 但是冲突概率很低)。



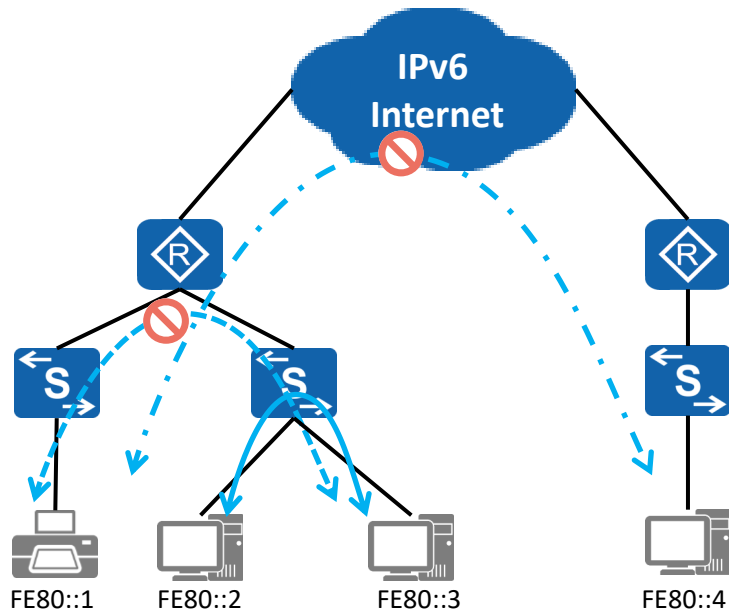


IPv6常见单播地址 - LLA

- LLA（Link-Local Address，链路本地地址）是IPv6中另一种应用范围受限制的地址类型。LLA的有效范围是本地链路，前缀为FE80::/10。

10 bit	54 bit	64 bit
1111 1110 10	0	接口标识
固定为0		

- LLA用于一条单一链路层面的通信，例如IPv6地址无状态自动配置、IPv6邻居发现等。
- 源或目的IPv6地址为链路本地地址的数据包将不会被转发到始发的链路之外，换句话说，链路本地地址的有效范围为本地链路。
- 每一个IPv6接口都必须具备一个链路本地地址。华为设备支持自动生成和手工指定两种配置方式。



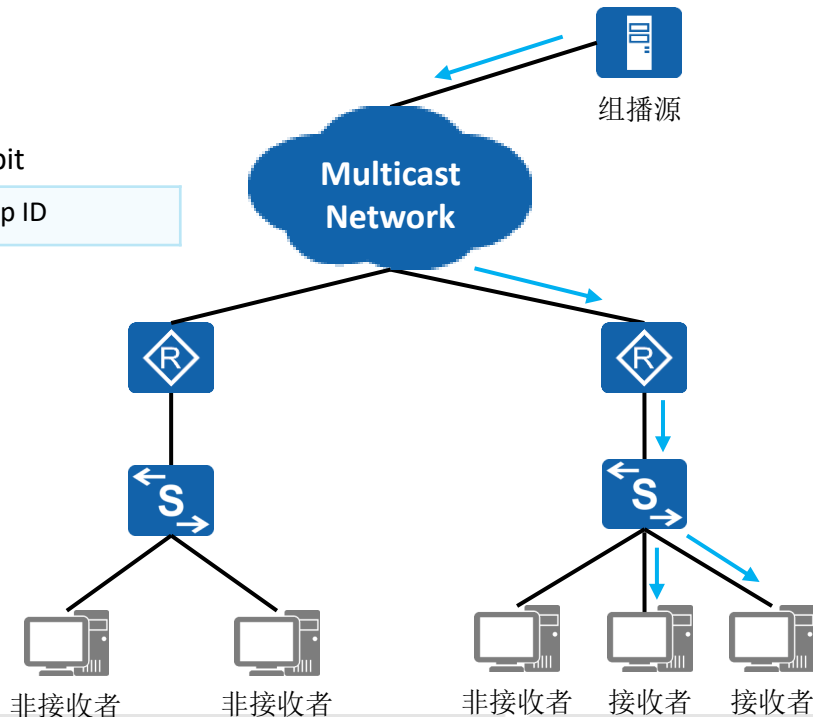


IPv6组播地址

- IPv6组播地址标识多个接口，一般用于“一对多”的通信场景。
- IPv6组播地址只可以作为IPv6报文的目的地址。

8 bit	4 bit	4 bit	80 bit	32 bit
11111111	Flags	Scope	Reserved (必须为0)	Group ID

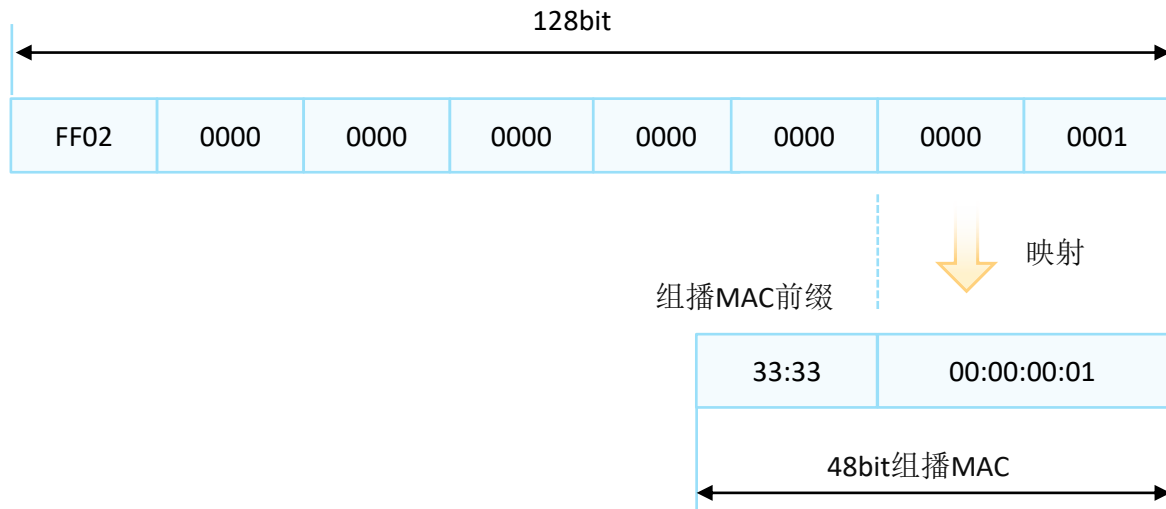
- Flags**
 - 用来表示永久或临时组播地址
 - 0000表示永久组播地址
 - 0001表示临时组播地址
- Scope**
 - 表示组播组的范围
- Group ID**
 - 组播组ID





IPv6组播MAC

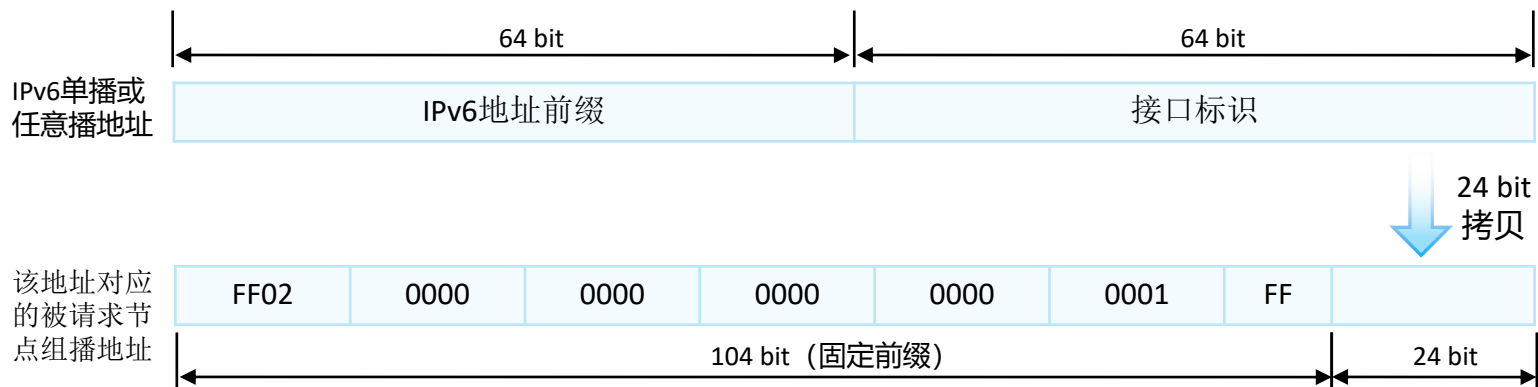
- 组播IPv6报文的目的IP为组播IPv6地址，同样，目的MAC为组播MAC地址。
- 组播MAC的前16bit为“33:33”，是专门为IPv6组播预留的MAC地址前缀。后32bit从组播IPv6地址的后32bit直接映射而来。





被请求节点组播地址

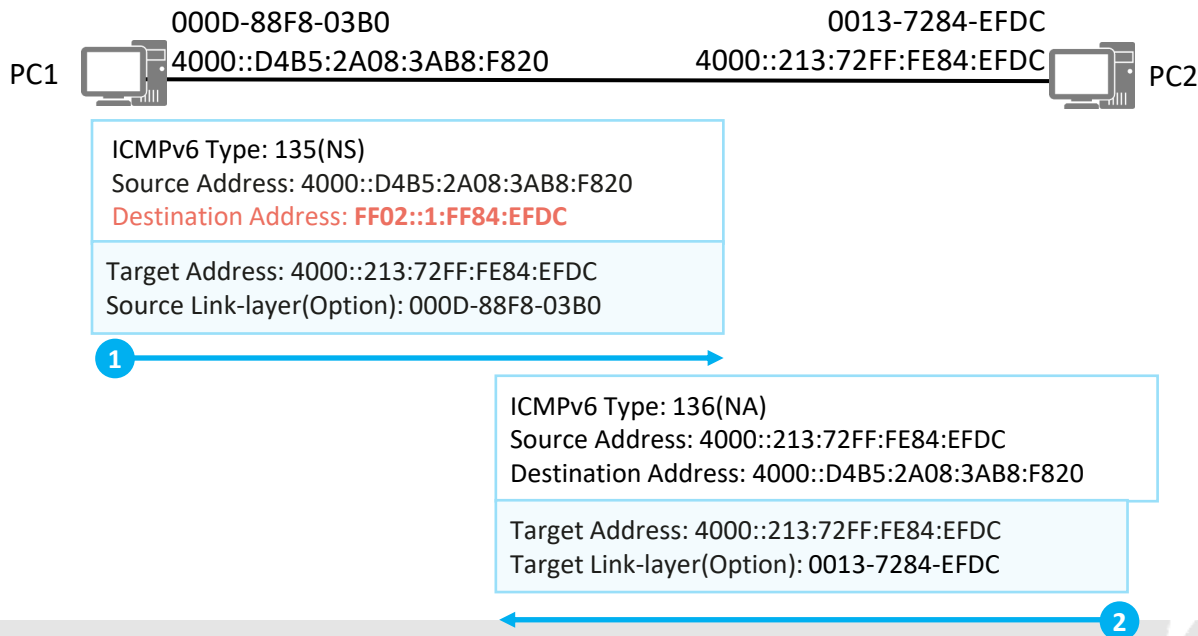
- 当一个节点具有了单播或任播地址，就会对应生成一个被请求节点组播地址，并且加入这个组播组。该地址主要用于**邻居发现机制**和**地址重复检测**功能。被请求节点组播地址的有效范围为本地链路范围。





被请求节点组播地址 - 示例

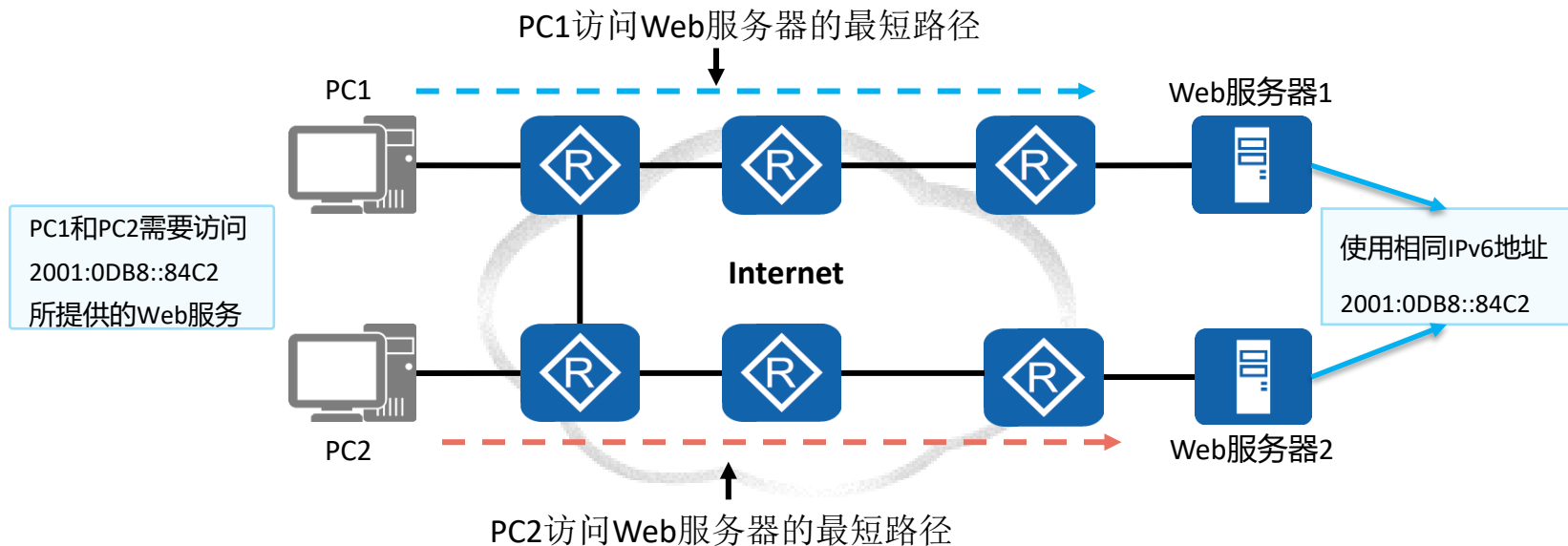
PC1发送数据至PC2前，首先需要获取其MAC地址。PC1将发起类似IPv4中ARP的解析流程，IPv6使用ICMPv6的NS及NA报文来实现地址解析过程，NS报文的目的IPv6地址为目标IPv6单播地址对应的被请求节点组播地址。





IPv6任播地址

任播地址标识一组网络接口（通常属于不同的节点）。任播地址可以作为IPv6报文的源地址，也可以作为目的地址。





IPv6地址和IPv4地址比较

	IPv4	IPv6
地址空间	2^{32}	2^{128}
表示方式	点分十进制	冒号隔开的十六进制
地址类型	单播、组播、广播	单播、组播、任播
其它	A、B、C等主类地址	IPv6中无此概念
	组播地址（224.0.0.0/4）	IPv6组播地址（FF00::/8）
	广播地址	IPv6中无此概念
	未指定的地址0.0.0.0/32	未指定的地址::/128
	环回地址127.0.0.0/8	环回地址是::1/128
	公网IP地址	全球单播地址
	私网IP地址（10.0.0.0/8，172.16.0.0/12以及192.168.0.0/16）	唯一本地地址（FD00::/8）
	APIPA地址（169.254.0.0/16）	链路本地地址（FE80::/10）



主机和路由器的IPv6地址

- 一般情况下，主机和路由器的单播IPv6地址以及加入的组播地址如下所示：



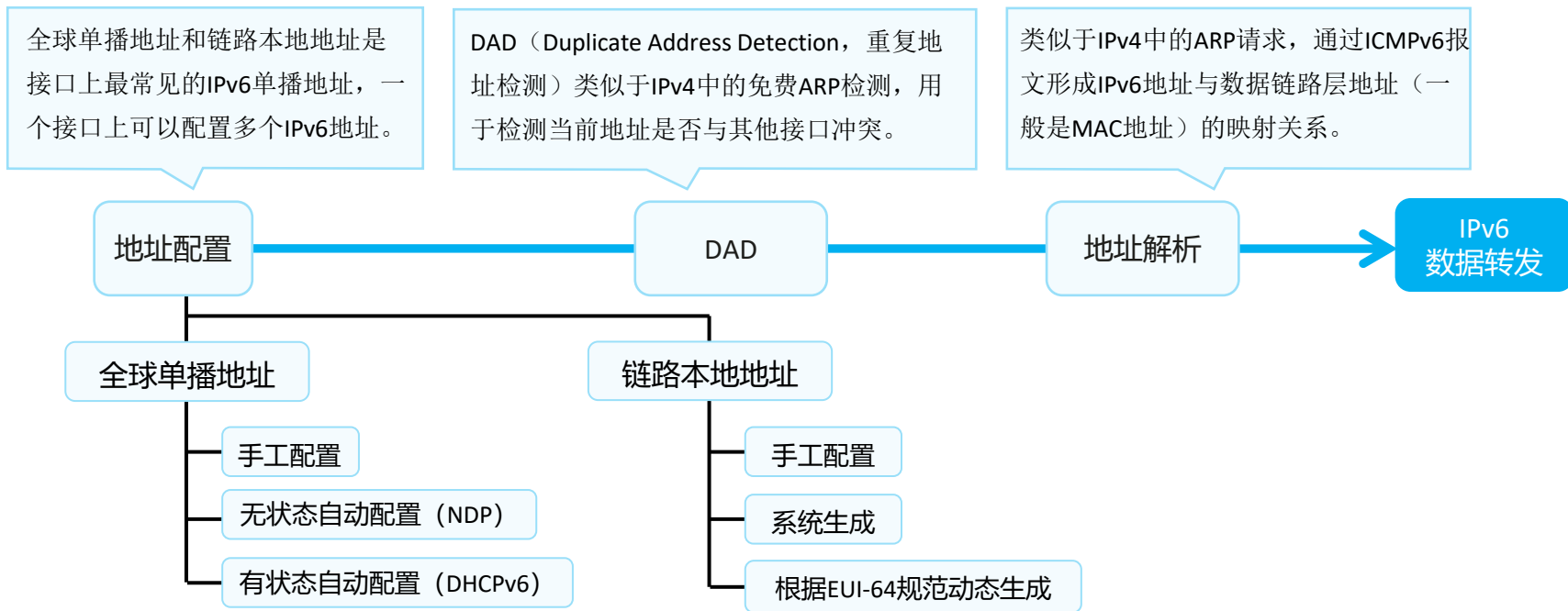
网卡的链路本地地址	FE80::2E0:FCFF:FE35:7287
管理员分配的全球单播地址	2001::1
环回地址	::1
“所有节点”组播地址	FF01::1 及 FF02::1
网卡的每个单播地址对应的被请求节点组播地址	FF02::1:FF35:7287 FF02::1:FF00:1

网卡的链路本地地址	FE80::2E0:FCFF:FE99:1285
管理员分配的全球单播地址	2001::2
环回地址	::1
“所有节点”组播地址	FF01::1 及 FF02::1
“所有路由器”组播地址	FF01::2 及 FF02::2
网卡的每个单播地址对应的被请求节点组播地址	FF02::1:FF99:1285 FF02::1:FF00:2



IPv6单播地址业务流程

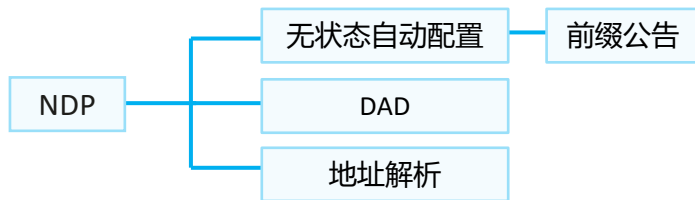
- 一个接口在发送IPv6报文之前要经历地址配置、DAD、地址解析这三个阶段，NDP（Neighbor Discovery Protocol，邻居发现协议）扮演了重要角色。





NDP

- RFC2461定义了NDP，该RFC后来被RFC4861替代。
- NDP使用ICMPv6报文实现其功能。



NDP使用的ICMPv6报文

ICMPv6 Type	报文名称
133	路由器请求 (RS)
134	路由器通告 (RA)
135	邻居请求 (NS)
136	邻居通告 (NA)

机制	RS 133	RA 134	NS 135	NA 136
地址解析			√	√
前缀公告	√	√		
DAD			√	√



IPv6动态地址配置

PC (DHCPv6客户端)



有状态地址配置

DHCPv6服务器端



DHCPv6 Interaction

- 通过DHCPv6报文交互, DHCPv6服务器端自动配置IPv6地址/前缀**及其他**网络配置参数 (DNS、NIS、SNTP服务器地址等参数)。

PC



无状态地址配置

2000::2E0:FCFF:FE35:7287/64

2000::1/64

Router



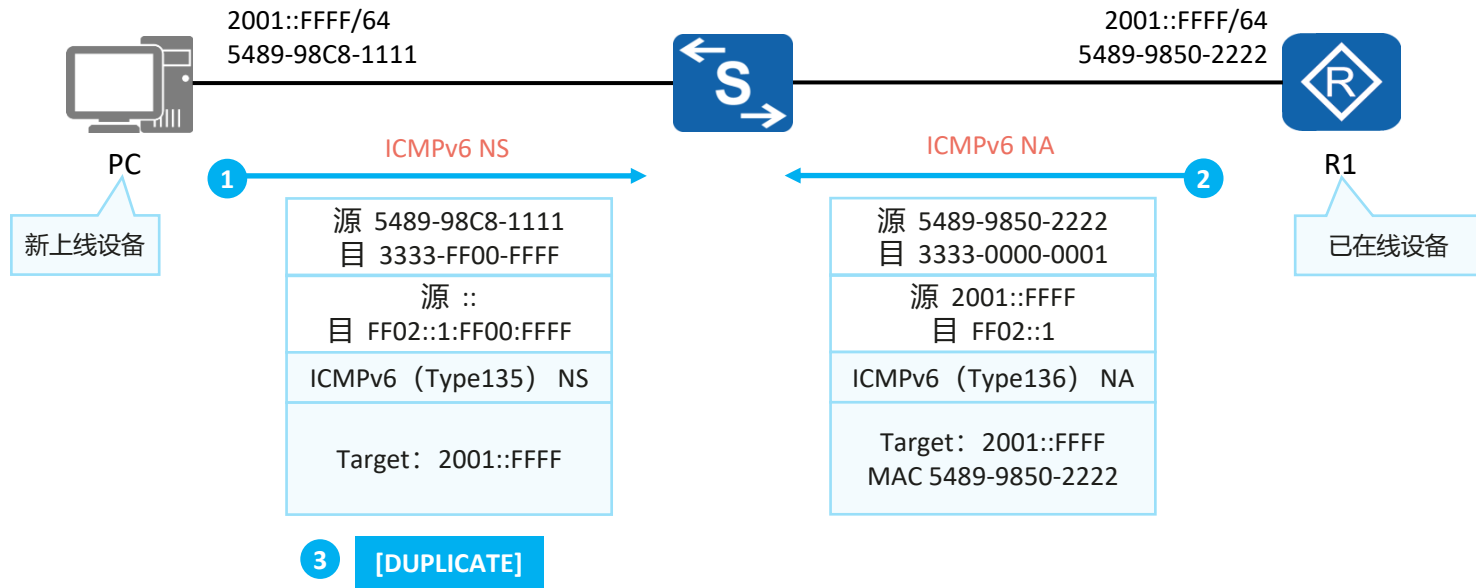
ICMPv6 Router Advertisement
(我的接口地址前缀是2000::/64)

- 主机根据RA中的地址前缀, 并结合本地生成的64 bit接口标识 (例如EUI-64), 生成单播地址。
- 仅可以获得IPv6地址信息, 无法获得NIS、SNTP服务器等参数, 需要配合DHCPv6或者手工配置来获取其他配置信息。



DAD

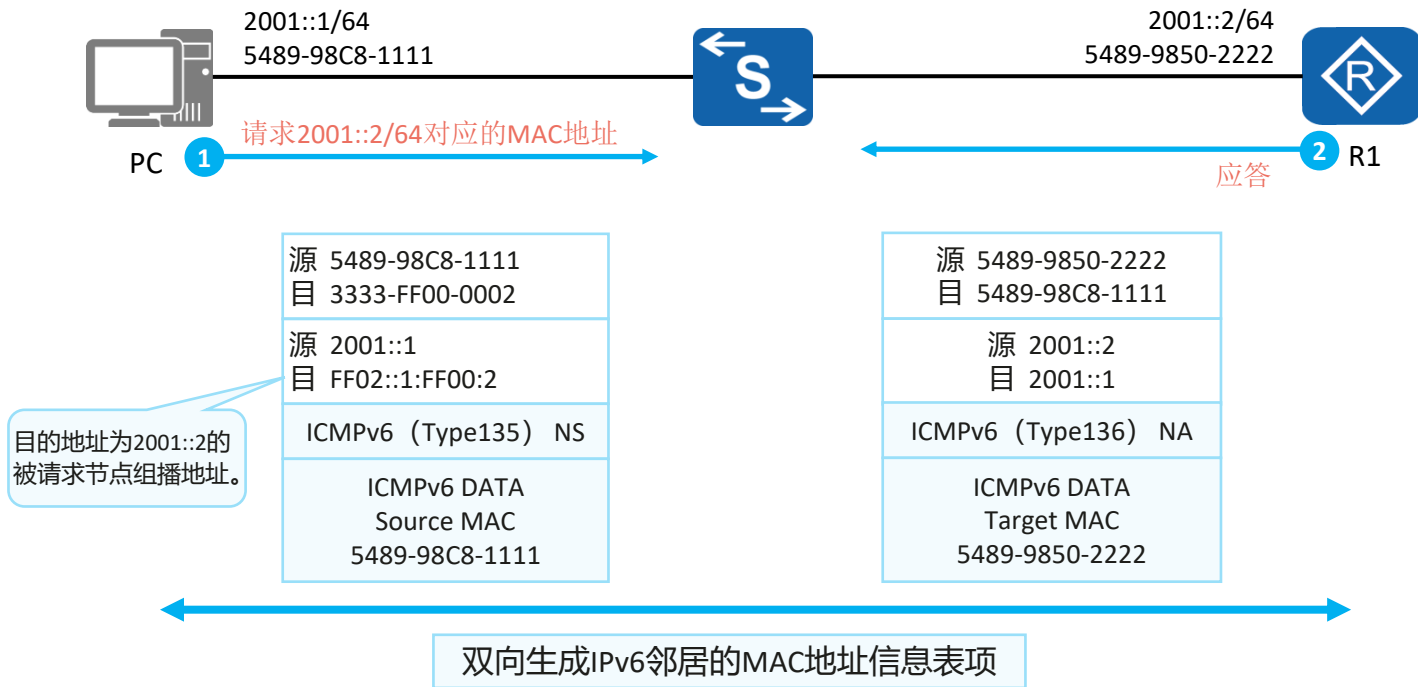
- 无论通过何种方式配置了IPv6单播地址，主机或路由器都会：
 - 通过ICMPv6报文进行DAD
 - 仅当DAD通过之后才会使用该单播地址





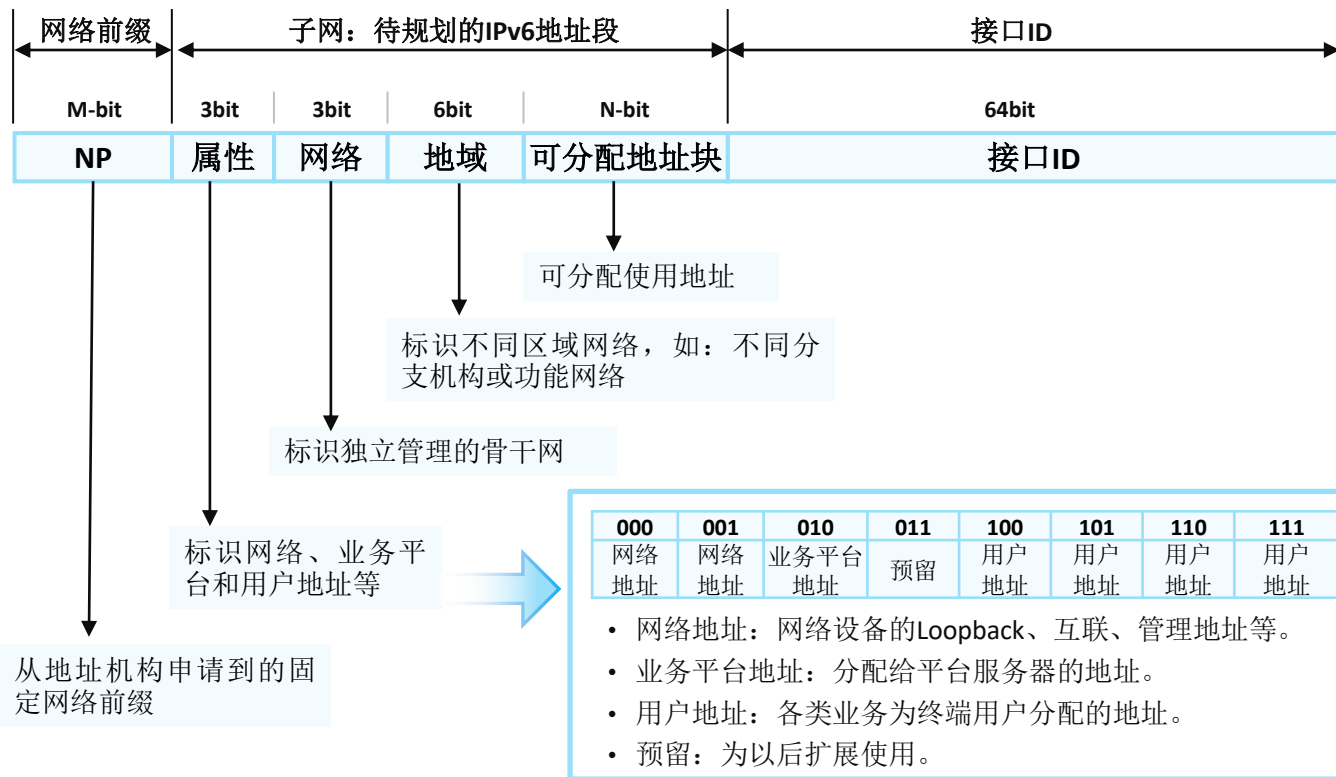
地址解析

- IPv6使用ICMPv6的NS和NA报文来取代ARP在IPv4中的地址解析功能。





IPv6地址规划举例



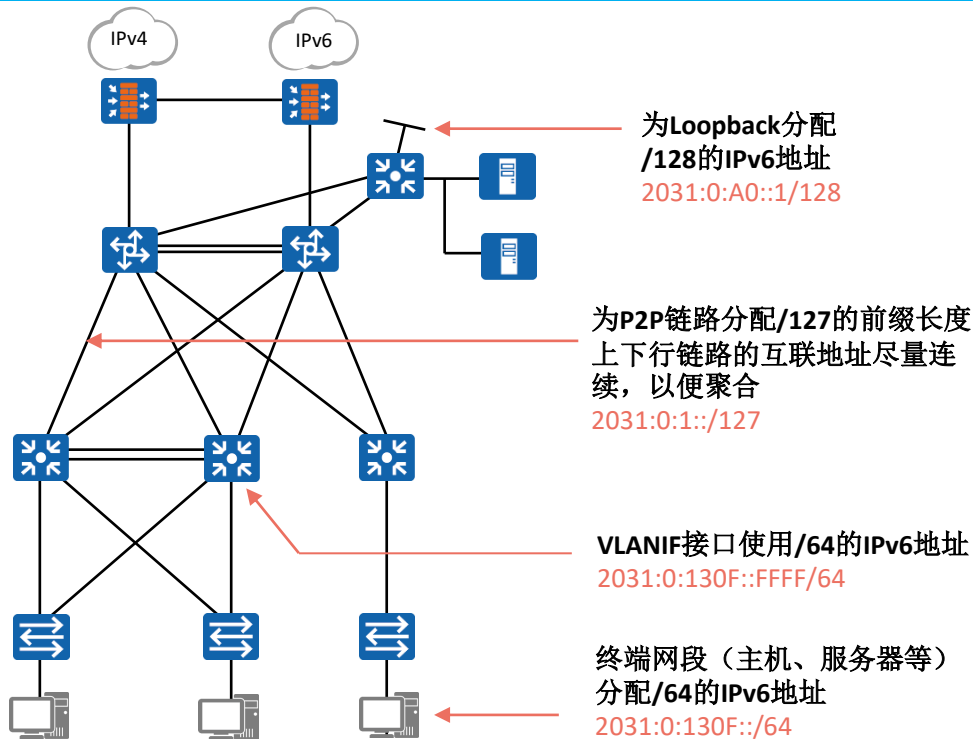


IPv6地址使用建议

地址规划要点

- 一般而言，ISP从地址分配机构获取的IPv6地址至少为/32前缀长度。企业用户获得的前缀长度往往为/48。
- 虽然IPv6地址空间巨大，但是在规划网络时，依然需要对地址进行合理规划。
- 确保IPv6地址的全网唯一性（除某些特殊的应用外，如anycast）。
- 确保IPv6地址规划的连续性、可聚合性、可扩展性。
- 业务地址：可在IPv6地址中规划适当bit用于承载业务信息、VLAN信息或位置信息，利于路由规划、QoS部署。
- 用户地址：考虑为不同业务类型的用户预留连续的地址段。通过特定bit确定用户类型、分布区域等。
- 建议为终端网段（用户主机、服务器等）分配/64的前缀长度。
- 建议为P2P链路分配/127的前缀长度。
- 建议为Loopback接口分配/128的前缀长度。

地址规划实践

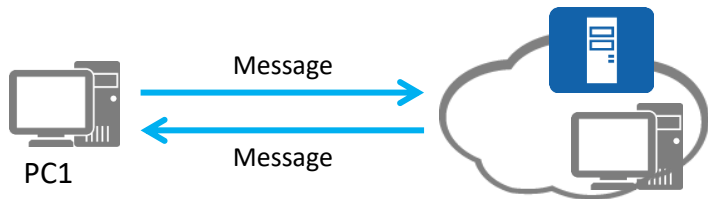




ICMP

Internet控制消息协议ICMP (Internet Control Message Protocol)是IP协议的辅助协议。

ICMP协议用来在网络设备间传递各种差错和控制信息，对于收集各种网络信息、诊断和排除各种网络故障等方面起着至关重要的作用。



以太网头部

IP头部

ICMP报文

以太网尾部

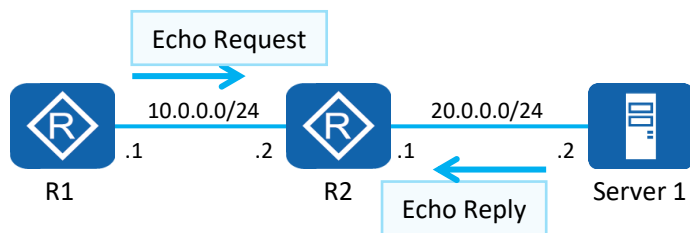
Type	Code	Checksum
ICMP的报文内容		

Type	Code	描述
0	0	Echo Reply
3	0	网络不可达
3	1	主机不可达
3	2	协议不可达
3	3	端口不可达
5	0	重定向
8	0	Echo Request



ICMP差错检测

ICMP Echo消息常用于诊断源和目的地之间的网络连通性，同时还可以提供其他信息，如报文往返时间等。



功能：Ping

Ping是网络设备、Windows、Unix和Linux平台上的一个命令，其实是一个小巧而实用的应用程序，该应用基于ICMP协议。Ping常用于探测到达目的节点的网络可达性。

```
[R1]ping 20.0.0.2
```

```
PING 20.0.0.2: 56 data bytes, press CTRL_C to break
```

```
Reply from 20.0.0.2: bytes=56 Sequence=1 ttl=254 time=70 ms
```

```
Reply from 20.0.0.2: bytes=56 Sequence=2 ttl=254 time=30 ms
```

```
Reply from 20.0.0.2: bytes=56 Sequence=3 ttl=254 time=30 ms
```

```
Reply from 20.0.0.2: bytes=56 Sequence=4 ttl=254 time=40 ms
```

```
Reply from 20.0.0.2: bytes=56 Sequence=5 ttl=254 time=30 ms
```

```
--- 20.0.0.2 ping statistics ---
```

```
5 packet(s) transmitted
```

```
5 packet(s) received
```

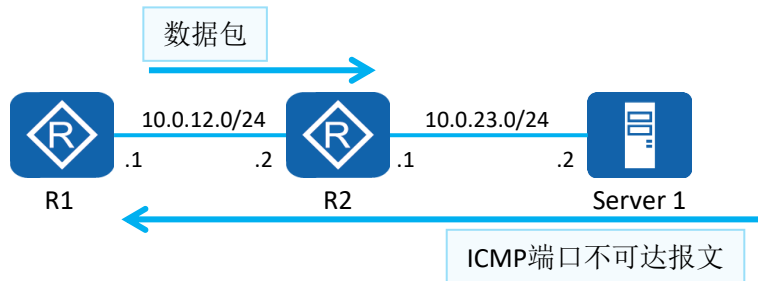
```
0.00% packet loss
```

```
round-trip min/avg/max = 30/40/70 ms
```



ICMP错误报告

ICMP定义了各种错误消息，用于诊断网络连接性问题；根据这些错误消息，源设备可以判断出数据传输失败的原因。



功能：Tracert

Tracert基于报文头中的TTL值来逐跳跟踪报文的转发路径。

Tracert是检测网络丢包和时延的有效手段，同时可以帮助管理员发现网络中的路由环路。

```
[R1]tracert 10.0.23.2
```

```
tracert to 10.0.23.2, max hops: 30 ,packet length: 40,press  
CTRL_C to break
```

1	10.0.12.2	10 ms	80 ms	10 ms
2	10.0.23.2	20 ms	30 ms	30 ms

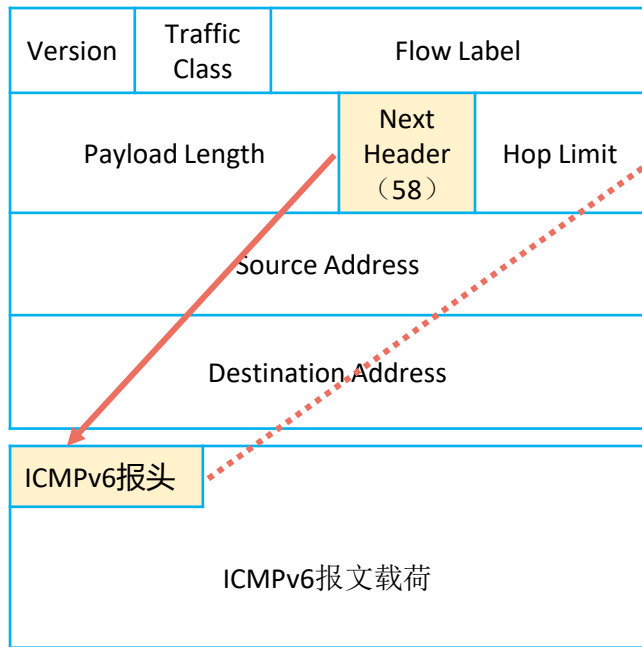


ICMPv6概述

- ICMPv6是IPv6的基础协议之一。
- 在IPv6报文头部中，Next Header字段值为58则对应为ICMPv6报文。
- ICMPv6报文用于通告相关信息或错误。
- ICMPv6报文被广泛应用于其它协议中，包括NDP、Path MTU发现机制等。
- ICMPv6控制着IPv6中的地址自动配置、地址解析、地址冲突检测、路由选择、以及差错控制等关键环节。



ICMPv6报文格式



Type	Code	Checksum
------	------	----------

消息类型	Type	名称	Code
差错消息	1	目的不可达	0 无路由
			1 因管理原因禁止访问
			2 未指定
			3 地址不可达
			4 端口不可达
	2	数据包过长	0
	3	超时	0 跳数到0
			1 分片重组超时
			0 错误的包头字段
			1 无法识别的下一包头类型
			2 无法识别的IPv6选项
信息消息	128	Echo Request	0
	129	Echo Reply	0



ICMPv6报文类型

ICMPv6报文分为两类：差错报文和信息报文。

差错报文

- 差错报文 (Error Messages) , 也称为差错消息, Type字段最高bit为0, 也就是ICMPv6 Type=[0, 127]
- 差错消息用于报告在转发IPv6数据包过程中出现的错误, 如常见的目的不可达、超时等等。

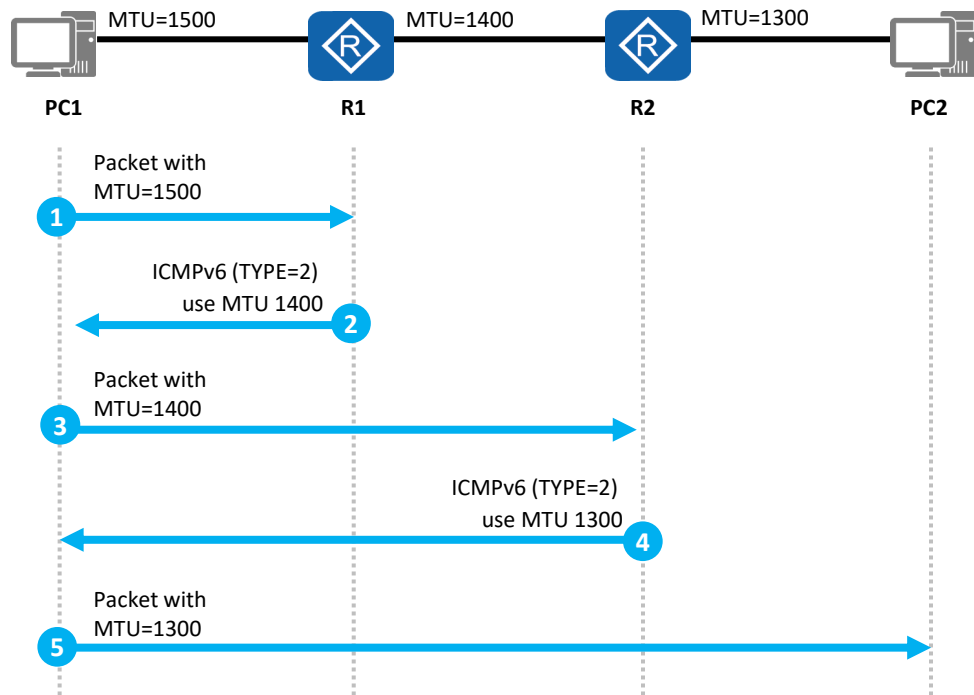
信息报文

- 信息报文 (Information Messages) , 也称为信息消息, Type字段最高bit为1, 也就是ICMPv6 Type=[128, 255]
- 信息报文可以用来实现同一链路上节点间的通信和子网内的组播成员管理等。



ICMPv6差错报文应用 - Path MTU发现

- 在IPv6中，中间转发设备不对IPv6报文进行分片，报文的分片将在源节点进行。
- PMTU（Path MTU）就是路径上的最小接口MTU。
- PMTUD（Path MTU发现机制）的主要目的是发现路径上的MTU，当数据包被从源转发到目的地的过程中避免分段。
- 依赖PMTUD，数据的发送方可以使用所发现到的最优PMTU与目的地节点进行通信，这样可以避免数据包在从源传输到目的的过程之中，被中途的路由器分片而导致性能的下降。

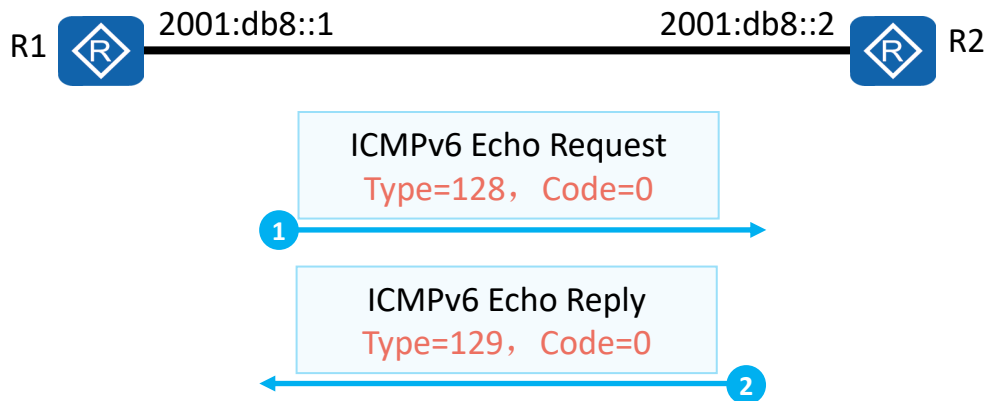




ICMPv6信息报文应用 - Ping

Ping基于ICMPv6信息报文实现

- **Echo Request:** 用于发送到目标节点，以使目标节点立即发回一个Echo Reply应答报文。Echo Request报文的Type字段值为128，Code字段的值为0。
- **Echo Reply:** 当收到一个Echo Request报文时，ICMPv6会用Echo Reply报文响应。Echo Reply报文的Type字段的值为129，Code字段的值为0。





ICMPv6其它常用的报文

- 邻居发现（RFC2461和RFC4861）
 - Type=133 路由器请求（Router Solicitation）
 - Type=134 路由器公告（Router Advertisement）
 - Type=135 邻居请求（Neighbor Solicitation）
 - Type=136 邻居公告（Neighbor Advertisement）
 - Type=137 重定向（Redirect）
- 组播侦听者发现协议（RFC2710和RFC3810）
 - Type=130 查询消息
 - Type=131 报告消息
 - Type=132 离开消息
 - Type=143 MLDv2报告消息



NDP概述

RFC2461定义了IPv6邻居发现协议NDP。NDP是IPv6中非常核心的组件。其主要功能如下：

NDP	路由器发现	发现链路上的路由器，获得路由器通告的信息。
	无状态自动配置	通过路由器通告的地址前缀，终端自动生成IPv6地址。
	重复地址检测	获得地址后，进行地址重复检测，确保地址不存在冲突。
	地址解析	请求目的网络地址对应的数据链路层地址，类似IPv4的ARP。
	邻居状态跟踪	通过NDP发现链路上的邻居并跟踪邻居状态。
	前缀重编址	路由器对所通告的地址前缀进行灵活设置，实现网络重编址。
	重定向	告知其他设备，到达目标网络的更优下一跳。



NDP报文类型及功能

NDP使用以下几种ICMPv6报文：

- RS（Router Solicitation）：路由器请求报文
- RA（Router Advertisement）：路由器通告报文
- NS（Neighbor Solicitation）：邻居请求报文
- NA（Neighbor Advertisement）：邻居通告报文

功能 \ ICMPv6 报文	RS 133	RA 134	NS 135	NA 136	重定向 137
地址解析			●	●	
路由器发现	●	●			
前缀重编址	●	●			
重复地址检测			●	●	
重定向					●



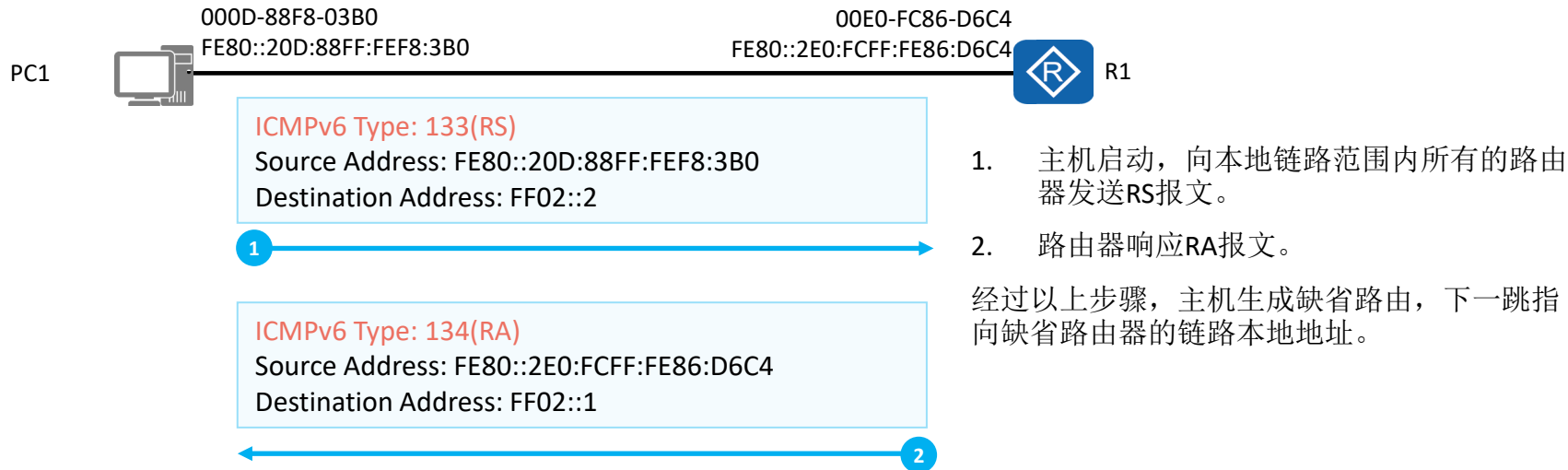
路由器发现

- 路由器发现是指主机发现本地链路上路由器和确定其配置信息的过程。
- 路由器发现可以同时实现以下三个功能：
 - 路由器发现（Router Discovery）：主机定位邻居路由器以及选择哪一个路由器作为缺省网关的过程。
 - 前缀发现（Prefix Discovery）：主机发现本地链路上的一组IPv6前缀的过程，用于主机的地址自动配置。
 - 参数发现（Parameter Discovery）：主机发现相关操作参数的过程，如输出报文的缺省跳数限制、地址配置方式等信息。
- 使用报文：
 - RS 路由器请求
 - RA 路由器通告
- 协议交互主要有两种情况：
 - 主机发送RS触发路由器回应RA
 - 路由器周期发送RA



路由器发现流程 - 主机请求触发

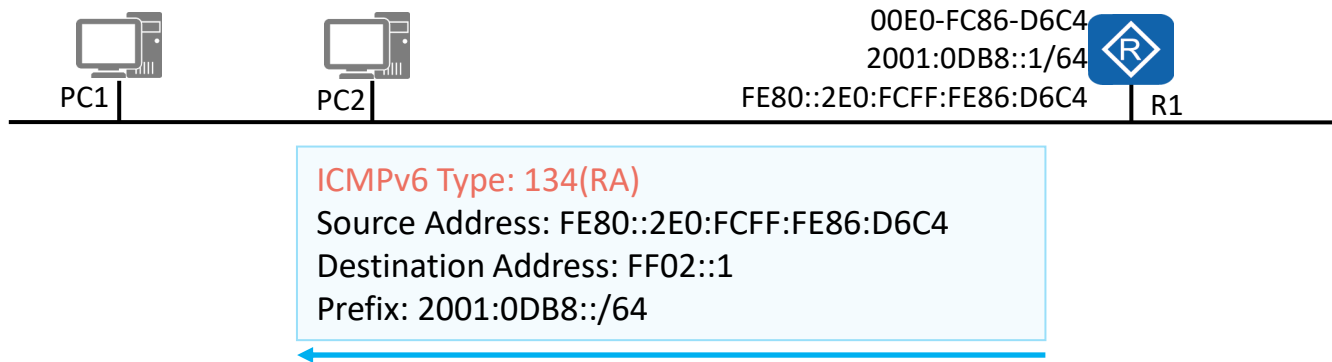
当主机启动时，主机会向本地链路范围内所有的路由器发送RS报文，触发路由器响应RA报文。主机发现本地链路上的路由器后，自动配置缺省路由器，建立缺省路由表、前缀列表和设置其它的配置参数。





路由器发现流程 - 路由器周期性发送

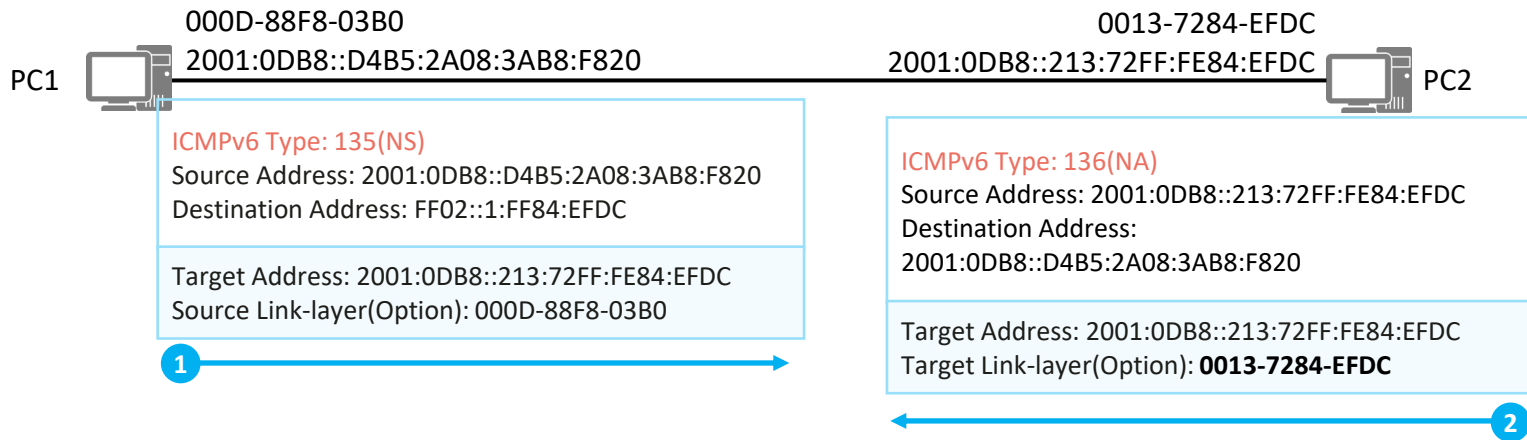
- 路由器周期性的发送RA报文，RA发送间隔是一个有范围的随机值，缺省的最大时间间隔是600秒，最小时间间隔是200秒。
- 对于定期发送的RA报文，其地址有如下要求：
 - Source Address: 必须是发送接口的链路本地地址。
 - Destination Address : FF02::1。





地址解析

- IPv6地址解析通过ICMPv6（NS和NA报文）来实现。
- 在三层完成地址解析，主要带来以下几个好处：
 - 地址解析在三层完成，不同的二层介质可以采用相同的地址解析协议。
 - 可以使用三层的安全机制避免地址解析攻击。
 - 使用组播方式发送请求报文，减少了二层网络的性能压力。





IPv6邻居状态表

IPv6邻居状态表中缓存了IPv6地址与MAC地址的映射，可以通过display ipv6 neighbors命令来查看IPv6邻居状态表。

```
<Huawei>display ipv6 neighbors
```

```
-----  
IPv6 Address   : 2001:DB8::2  
Link-layer     : 00e0-fc23-26e3          State : REACH  
Interface      : GE0/0/0                Age   : 0  
VLAN           : -                      CEVLAN: -  
VPN name       :                        Is Router: TRUE  
Secure FLAG    : UN-SECURE
```

```
IPv6 Address   : FE80::2E0:FCFF:FE23:26E3  
Link-layer     : 00e0-fc23-26e3          State : REACH  
Interface      : GE0/0/0                Age   : 0  
VLAN           : -                      CEVLAN: -  
VPN name       :                        Is Router: TRUE  
Secure FLAG    : UN-SECURE
```

```
-----  
Total: 2    Dynamic: 2    Static: 0
```



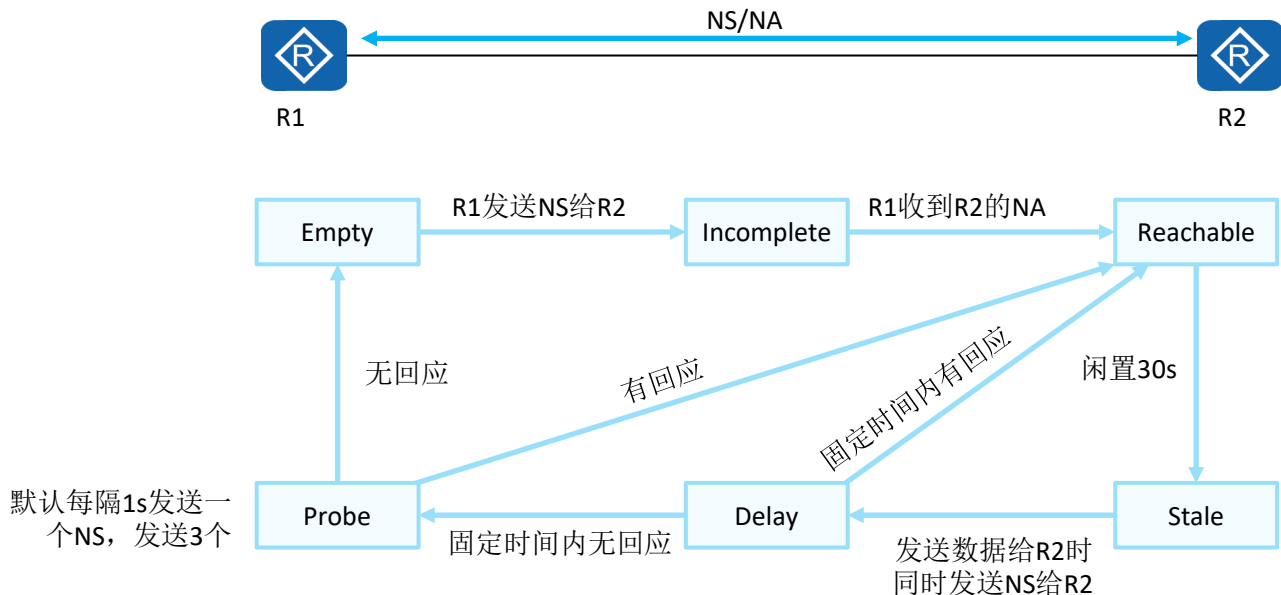
IPv6邻居状态

IPv6节点需要维护一张邻居表，每个邻居都有相应的状态，状态之间可以迁移。5种邻居状态分别是：未完成（Incomplete）、可达（Reachable）、陈旧（Stale）、延迟（Delay）、探查（Probe）。

状态	描述
Incomplete	邻居不可达。正在进行地址解析，邻居的链路层地址未探测到，如果解析成功，则进入Reachable状态。
Reachable	邻居可达。表示在规定时间内（邻居可达时间，缺省情况下是30秒）内邻居可达。如果超过规定时间，该表项没有被使用，则表项进入Stale状态。
Stale	邻居是否可达未知。表明该表项在规定时间内（邻居可达时间，缺省情况下是30秒）内没有被使用。此时除非有发送到邻居的报文，否则不对邻居是否可达进行探测。
Delay	邻居是否可达未知。已向邻居发送报文，如果在指定时间内没有收到响应，则进入Probe状态。
Probe	邻居是否可达未知。已向邻居发送NS报文，探测邻居是否可达。在规定时间内收到NA报文回复，则进入Reachable状态；否则进入Incomplete状态。



邻居状态迁移

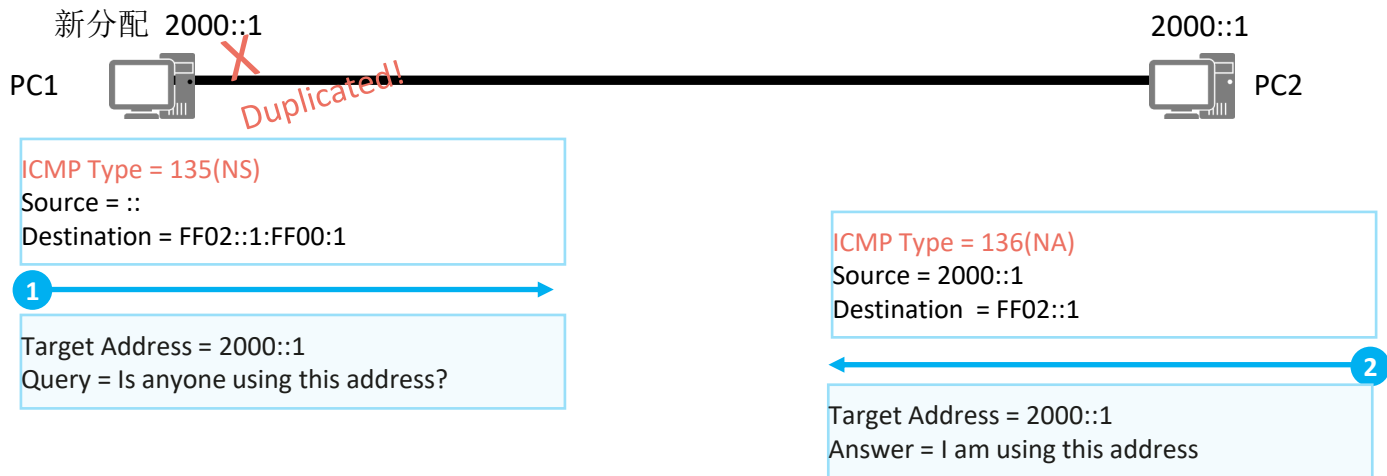


R1的邻居状态



重复地址检测 (1)

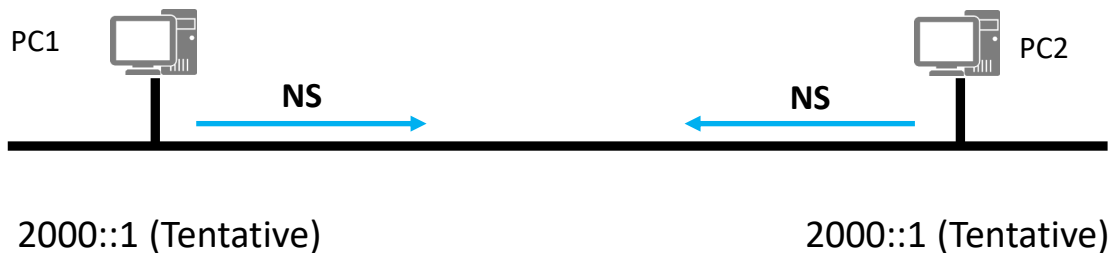
- 重复地址检测(Duplicate Address Detect, DAD)是指接口在使用某个IPv6地址之前, 需要先探测是否有其它的节点使用了该地址, 从而确保网络中没有两个相同的单播地址。
- 接口在启用任何一个单播IPv6地址前都需要先进行DAD, 包括Link-Local地址。





重复地址检测 (2)

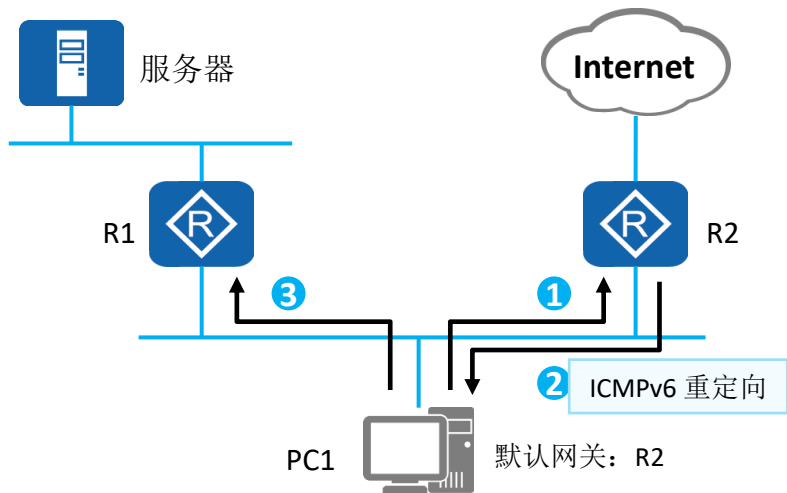
- 一个地址在通过重复地址检测之前称为“tentative地址”，即“试验地址”。此时该接口不能使用这个试验地址进行单播通讯。
- 若2个节点配置相同地址，同时作重复地址检测时，当一方收到对方发出的DAD NS报文，则接收方将不启用该地址。





重定向

重定向是指网关设备发现报文从其它网关设备转发更优，它就会发送重定向报文告知报文的发送者，让报文发送者选择另一个网关设备。

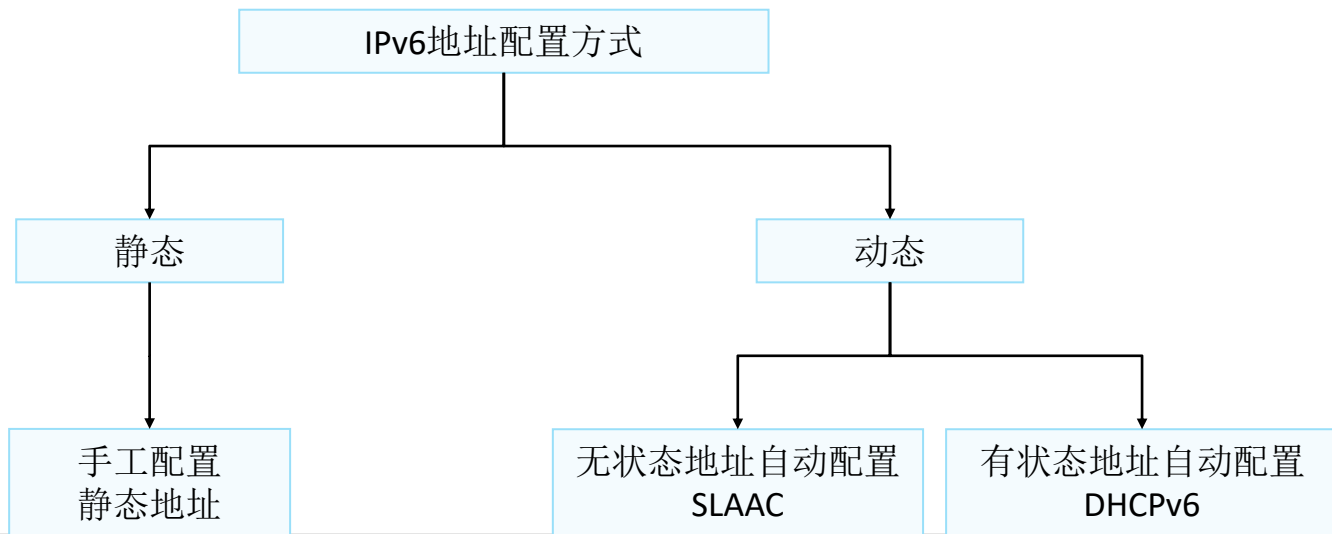


1. PC1希望发送报文到服务器，于是根据配置的默认网关地址向网关R2发送报文。
2. 网关R2收到报文后，检查报文信息，发现报文应该转发到与PC1在同一网段的另一个网关设备R1，此转发路径是更优的路径，于是R2会向PC1发送一个重定向消息，通知PC1去往服务器的报文应直接发给R1。
3. PC1收到重定向消息后，会向R1发送报文，R1再将该报文转发至服务器。



IPv6地址配置方式

IPv6地址配置的方式可以分为静态配置和动态配置。其中，动态地址配置又可以分为无状态地址自动配置（Stateless Address Autoconfiguration, SLAAC）和有状态地址自动配置（Stateful Address Autoconfiguration）。





IPv6地址自动配置的分类

无状态地址自动配置

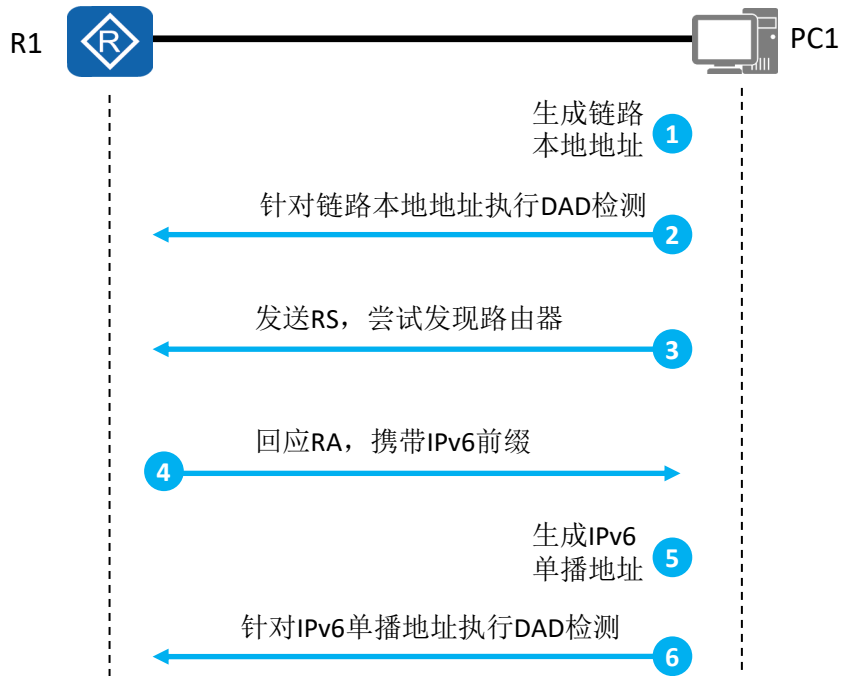
- 不需要IPv6地址分配服务器保存和管理每个节点的状态信息的一种IPv6地址自动配置方式，称之为IPv6无状态地址自动配置。
- 无状态地址自动配置方式基于NDP来实现。

有状态地址自动配置

- IPv6地址分配服务器必须保存每个节点的状态信息，并管理这些保存的信息，这种方式称之为IPv6有状态地址自动配置。
- 有状态地址自动配置基于DHCPv6（Dynamic Host Configuration Protocol for IPv6）来实现。



IPv6地址无状态自动配置过程

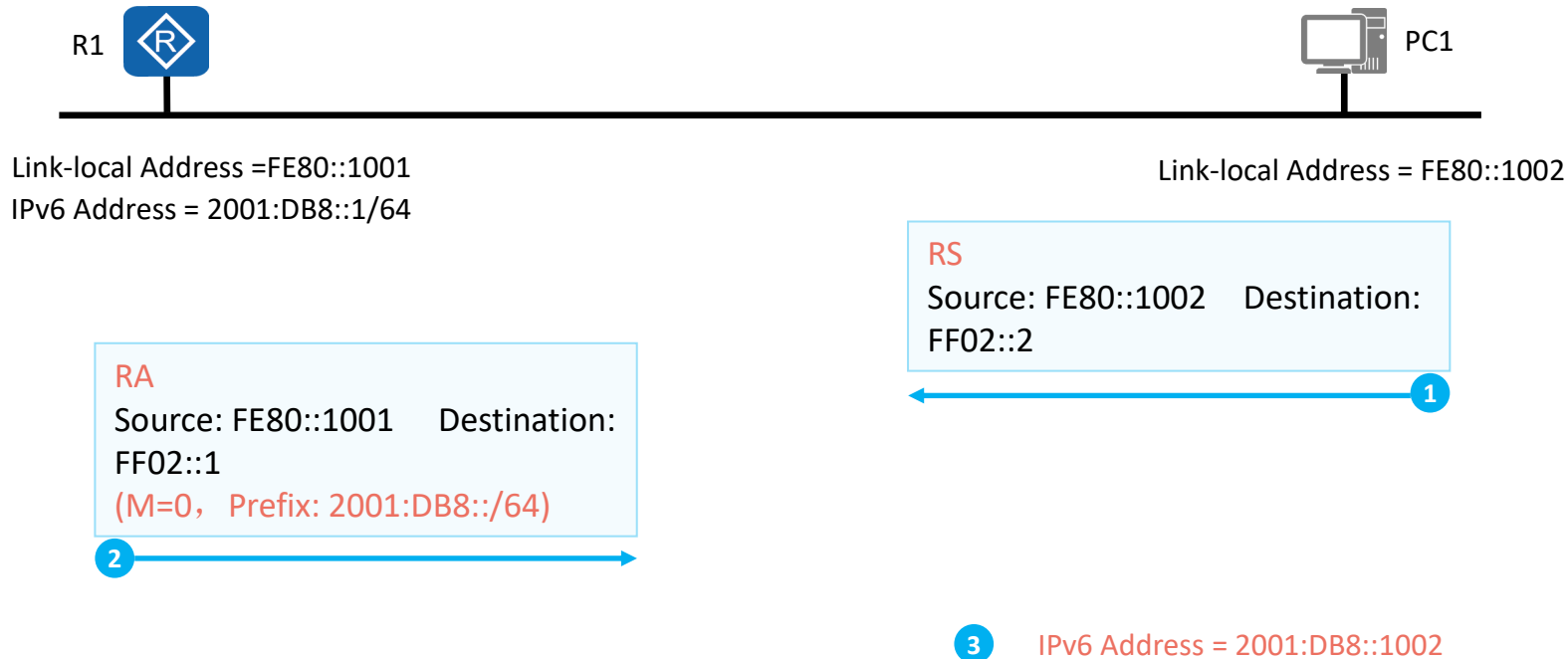


1. PC1根据本地的接口ID自动生成链路本地地址。
2. PC1对该链路本地地址进行DAD检测，如果该地址无冲突则可启用，此时PC1具备IPv6连接能力。
3. PC1发送RS报文，尝试在链路上发现IPv6路由器。
4. R1发送RA报文（携带可用于无状态地址自动配置的IPv6地址前缀。路由器在没有收到RS报文时也能够主动发出RA报文）。
5. PC1解析路由器发送的RA报文，获得IPv6地址前缀，使用该前缀加上本地的接口ID生成IPv6单播地址。
6. PC1对生成的IPv6单播地址进行DAD检测，如果没有检测到冲突，则启用该地址。



IPv6地址无状态自动配置示例

IPv6无状态地址配置通过交互RS和RA报文完成。





RA报文中的Flags字段

Internet Control Message Protocol V6

Type: 134 (Router advertisement)

Code: 0

Checksum: 0x4a68 [Correct]

Cur Hop Limit: 64

Flags: 0x00

0 = Not managed

. 0 = Not Other

. . 0 = Not Home Agent

. . . 0 . . . = Router Preference: Medium

. . . . 0 . . = Not Proxied

Router Lifetime: 1800

Reachable time : 0

Retrans timer: 0

ICMPv6 Option (Source Link-layer address)

ICMPv6 Option (MTU)

ICMPv6 Option (Prefix information)

- M-bit: 管理地址配置标识
 - M-bit默认为0, 此时收到该RA的终端将使用RA中包含的IPv6前缀来构造单播地址, 即采用无状态地址自动配置。
 - M-bit设置为1时, 收到该RA的终端将使用DHCPv6来获取地址及其他信息。当M-bit为1时, 终端可以忽略O-bit。也即终端将使用DHCPv6来获取地址及其他信息(如DNS等)。
- 使用如下命令, 可将M-bit设置为1: `ipv6 nd autoconfig managed-address-flag`
- O-bit: 其它有状态配置标识
 - O-bit默认为0, 此时收到该RA的终端进行无状态自动配置。即路由设备通过RA报文向终端发布除IPv6地址外的其他配置信息。
 - O-bit设置为1时, 收到该RA的终端将使用DHCPv6来获取除地址之外的其他参数。
- 使用如下命令, 可将O-bit设置为1: `ipv6 nd autoconfig other-flag`



RA报文中的可选信息：地址前缀信息

Internet Control Message Protocol V6

Type: 134 (Router advertisement)

Code: 0

...

Retrans timer: 0

田 ICMPv6 Option (Source Link-layer address)

田 ICMPv6 Option (Prefix information)

Type: Prefix information

Prefix-Length: 64

田 Flags: 0xc0

1 = On-Link Flag (L) : Set

. 1 = Autonomously Addr-conf flag (A) : Set

. . 00 0000 = Reserved: 0

Valid lifetime : 2592000

Preferred lifetime: 604800

Reserved

Prefix : 2001:DB8::

IPv6地址前缀长度

IPv6地址前缀

- A-bit, 用于指示终端设备是否能使用该前缀进行无状态地址自动配置。
 - A-bit被设置为0时, 此时终端不能使用该前缀进行无状态地址自动配置。
 - A-bit被设置为1时, 终端可以使用该前缀进行无状态自动配置。
- 使用如下命令将该比特位设置为0: `ipv6 nd ra prefix 2001:DB8:: 64 2592000 604800 no-autoconfig`



RA报文中的可选信息：生存周期

Internet Control Message Protocol V6

Type: 134 (Router advertisement)

Code: 0

...

Retrans timer: 0

田 ICMPv6 Option (Source Link-layer address)

田 ICMPv6 Option (Prefix information)

Type: Prefix information

Prefix-Length: 64

田 Flags: 0xc0

1 = On-Link Flag (L) : Set

. 1 = Autonomously Addr-conf flag (A) : Set

. . 00 0000 = Reserved: 0

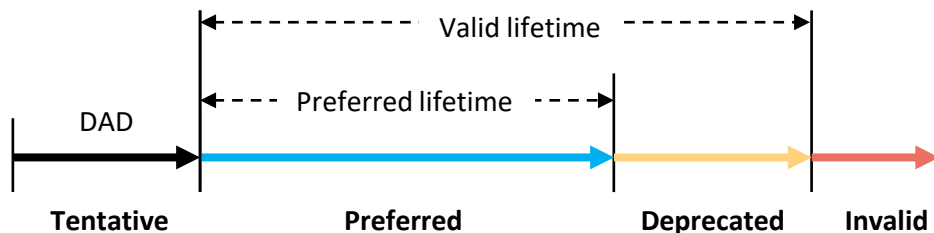
Valid lifetime : 2592000

Preferred lifetime: 604800

Reserved

Prefix : 2001:DB8::

- 当终端获取到前缀并生成IPv6单播地址后，首先进入Tentative状态，在通过DAD后，该地址将进入Preferred状态，并在Preferred lifetime内保持该状态。
- 在Preferred状态，终端可以正常收发报文。Preferred lifetime超时后，地址进入Deprecated状态，并在Valid lifetime内保持该状态。
- 在Deprecated状态，该地址仍然有效，现有的连接可以继续使用该地址，但是无法使用该地址建立新的连接。
- 当Valid lifetime超时后，地址进入Invalid状态，表示该地址无法继续使用。



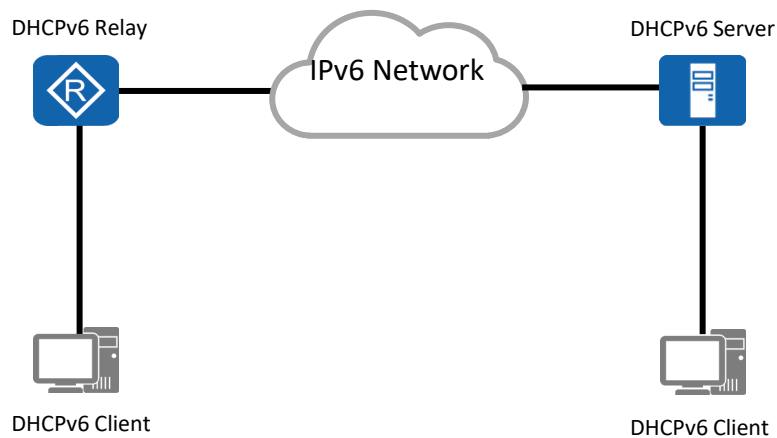


DHCPv6概述

- DHCPv6针对IPv6编址方案设计，支持对客户端分配IPv6前缀、IPv6地址和其他网络配置参数，并记录这些信息，便于网络管理。
- DHCPv6又分为如下三种：
 - DHCPv6有状态自动配置：DHCPv6服务器自动配置IPv6地址/前缀及其他网络配置参数（DNS、NIS、SNTP服务器地址等参数）。
 - DHCPv6无状态自动配置：主机IPv6地址仍然通过路由通告方式自动生成，DHCPv6服务器只分配除IPv6地址以外的配置参数，包括DNS服务器等参数。
 - DHCPv6 PD（Prefix Delegation，前缀代理）自动配置：下层网络路由器不需要再手工指定用户侧链路的IPv6地址前缀，它只需要向上层网络路由器提出前缀分配申请，上层网络路由器便可以分配合适的地址前缀给下层路由器，下层路由器把获得的前缀（前缀一般长度小于64）进一步自动细分成64位前缀长度的子网网段，把细分的地址前缀再通过路由通告(RA)至与IPv6主机直连的用户链路上，实现主机的地址自动配置，从而完成整个IPv6网络的层次化布局。



DHCPv6网络构成



DHCPv6基本协议架构中，主要包括以下三种角色：

- **DHCPv6 Client:** DHCPv6客户端，通过与DHCPv6服务器进行报文交互，获取IPv6地址/前缀和其他网络配置参数，完成自身的网络配置。
- **DHCPv6 Server:** DHCPv6服务器，负责处理来自客户端或中继的地址分配、地址续租、地址释放等请求，为客户端分配IPv6地址/前缀和其他网络配置参数。
- **DHCPv6 Relay:** DHCPv6中继，负责转发来自客户端或服务器的DHCPv6报文，协助DHCPv6客户端和DHCPv6服务器完成地址配置功能。



DHCPv6中的常用概念

有效时间

Valid Lifetime, 地址/前缀的生命周期。

优选时间

Preferred Lifetime, 用于计算续租时间和重绑定时间。

T1

IPv6地址的续租 (Renew) 时间, 默认是Preferred Lifetime的0.5倍。

T2

IPv6地址的重绑定 (Rebind) 时间, 默认是Preferred Lifetime的0.8倍。

IA

Identity association, 即身份联盟, 是使得服务器和客户端能够识别、分组和管理一系列相关IPv6地址的结构, 又分为IA_NA (非临时地址身份联盟) 和IA_PD (代理前缀身份联盟)。

DUID

DHCP Unique Identifier, 即DHCP设备唯一标识符。用来唯一标志一台设备, 每个客户端、服务器、中继都有自己的DUID。

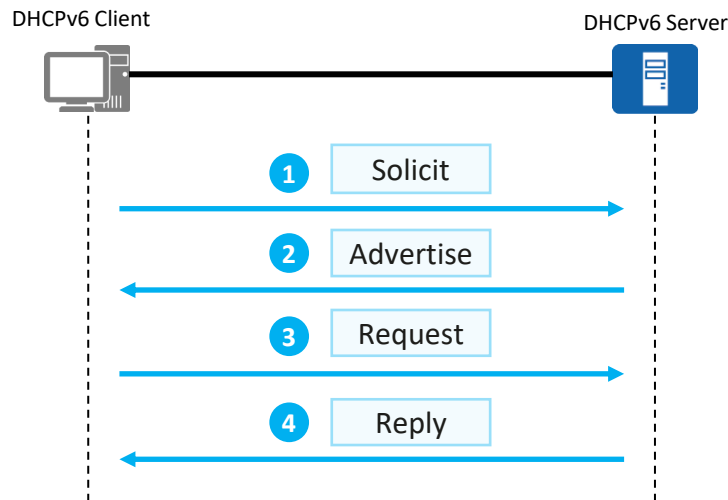
报文类型

Solicit、Advertise、Request、Reply、Information-Request、Renew、Rebind、Release、Confirm、Decline、Reconfigure、Relay-Forward、Relay-Reply等。



DHCPv6有状态自动配置 - 四步交互

四步交互是指DHCPv6客户端与服务器交互四次来完成前缀/地址等参数获取的过程。

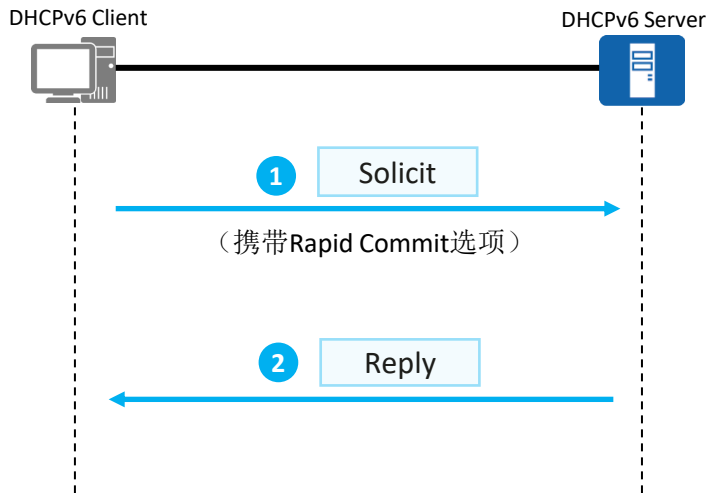


1. DHCPv6客户端发送Solicit消息，请求DHCPv6服务器为其分配IPv6地址/前缀和网络配置参数。
2. DHCPv6服务器回复Advertise消息，通知客户端可以为其分配的地址/前缀和网络配置参数。
3. 如果DHCPv6客户端接收到多个服务器回复的Advertise消息，则根据消息接收的先后顺序、服务器优先级等，选择其中一台服务器，并向该服务器发送Request消息，请求服务器确认为其分配地址/前缀和网络配置参数。
4. DHCPv6服务器回复Reply消息，确认将地址/前缀和网络配置参数分配给客户端使用。



DHCPv6有状态自动配置 - 两步交互

DHCPv6客户端可以在发送的Solicit消息中携带Rapid Commit选项，标识客户端希望服务器能够快速为其分配地址/前缀和网络配置参数。

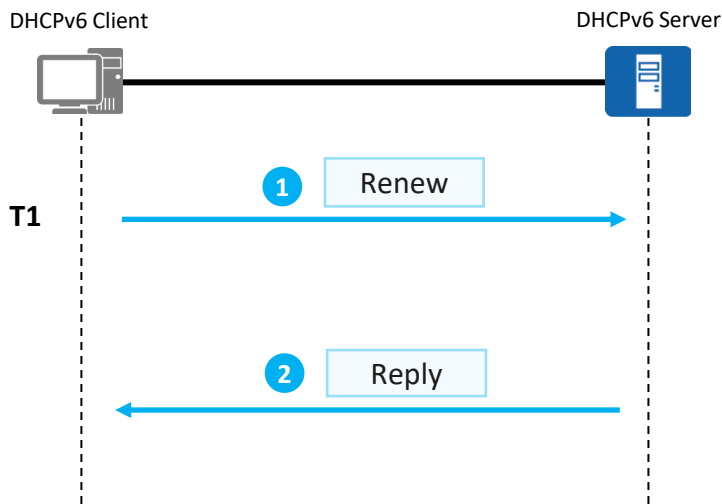


1. DHCPv6客户端发送Solicit报文，携带Rapid Commit选项。
2. DHCPv6服务器接收到Solicit报文后，将会判断，如果DHCPv6服务器支持快速分配，则直接返回Reply报文，为客户端分配IPv6地址/前缀和其他网络配置参数。如果DHCPv6服务器不支持快速分配，则将采用四步交互方式。



地址/前缀租约更新 (1)

DHCPv6服务器分配的IPv6地址/前缀具有有效时间。地址/前缀的租借时间超过有效时间后，DHCPv6客户端不能再使用该地址/前缀。因此，在有效时间超时之前，如果DHCPv6客户端希望继续使用该地址/前缀，则需要更新地址/前缀的租约。

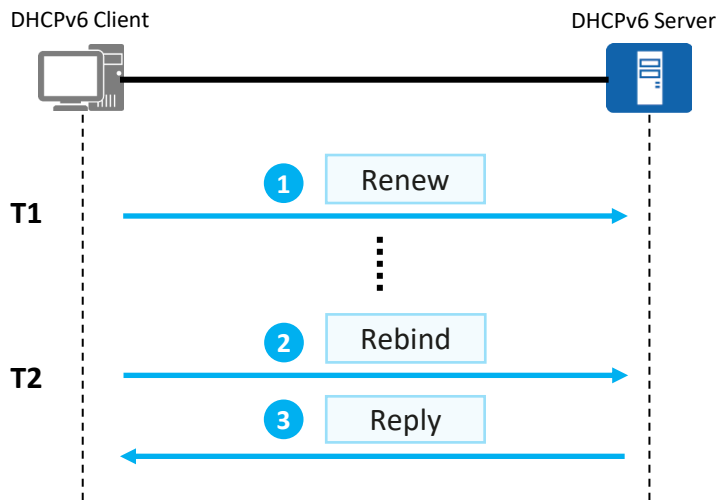


1. DHCPv6 客户端在 T1 时刻（默认为 Preferred Lifetime 的 1/2）发送 Renew 报文进行地址/前缀租约更新请求。
2. 如果 DHCPv6 客户端可以继续使用该地址/前缀，则 DHCPv6 服务器回应续约成功的 Reply 报文，通知 DHCPv6 客户端已经成功更新地址/前缀租约。否则，DHCPv6 服务器回应续约失败的 Reply 报文，通知 DHCPv6 客户端不能获得新的租约。



地址/前缀租约更新 (2)

如果DHCPv6服务器未响应T1时刻DHCPv6客户端发出的Renew请求，则客户端会在T2（默认为Preferred Lifetime的0.8倍）向所有DHCPv6服务器组播发送Rebind请求更新租约。



1. DHCPv6客户端在T1时刻发送Renew请求更新租约，但是没有收到DHCPv6服务器的回应报文。
2. DHCPv6客户端在T2时刻，向所有DHCPv6服务器组播发送Rebind报文请求更新租约。
3. 如果DHCPv6客户端可以继续使用该地址/前缀，则DHCPv6服务器回应续约成功的Reply报文，通知DHCPv6客户端已经成功更新地址/前缀租约。否则DHCPv6服务器回应续约失败的Reply报文，通知DHCPv6客户端不能获得新的租约。



DHCPv6无状态自动配置

DHCPv6服务器为已经具有IPv6地址/前缀的客户端分配除地址/前缀以外的其他网络配置参数，该过程称为DHCPv6无状态自动配置。

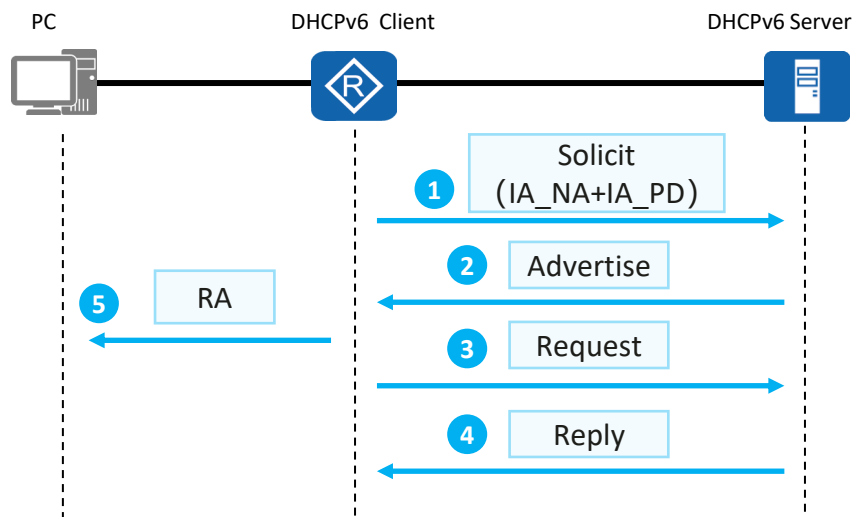


1. DHCPv6客户端以组播的方式向DHCPv6服务器发送Information-request报文，该报文中携带Option Request选项，指定客户端需要从服务器获取的配置参数。
2. 服务器收到该报文后，为客户端分配网络配置参数，并单播发送Reply报文将网络配置参数返回给客户端。客户端检查Reply报文中提供的信息，如果与Information-request报文中请求的配置参数相符，则按照Reply报文中提供的参数进行网络配置；否则，忽略该参数。



DHCPv6 PD自动配置

在一个层次化的网络结构中，不同层次的IPv6地址配置一般是手工指定的。手工配置IPv6地址扩展性不佳，不利于IPv6地址的统一规划管理。DHCPv6 PD可以解决这个问题。

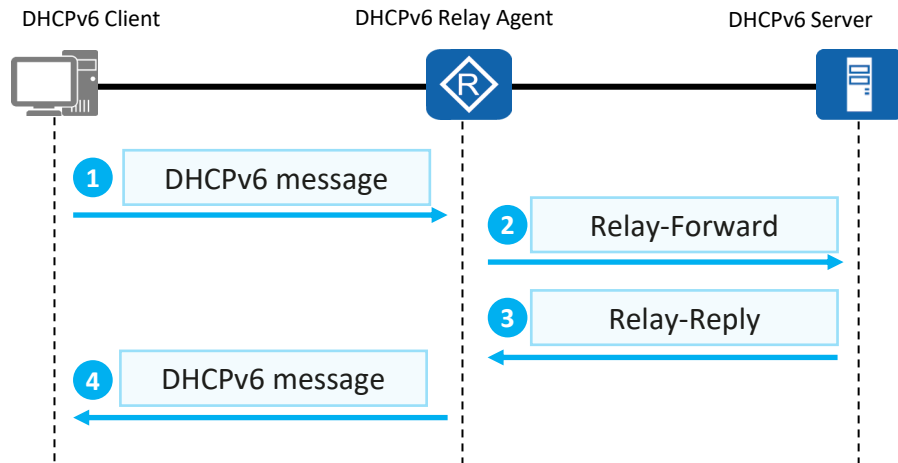


1. DHCPv6客户端发送Solicit报文，请求DHCPv6服务器为其分配IA_NA地址和IA_PD前缀。
2. DHCPv6服务器回复Advertise报文，通知客户端可以为其分配的IPv6地址和前缀。
3. 如果客户端接收到多个服务器回复的Advertise报文，则根据Advertise报文中的服务器优先级等参数，选择优先级最高的一台服务器（若服务器优先级一样，则选择带有该客户端需要的配置参数的Advertise报文），并向其发送Request报文，请求为其分配地址/前缀。
4. DHCPv6服务器回复Reply报文，确认将IPv6地址/前缀分配给DHCPv6客户端。
5. DHCPv6客户端在收到PD前缀后，与终端进行RS/RA报文交互，在RA报文中将携带获取到的PD前缀下发至终端。



DHCPv6中继工作过程

当服务器和客户端不在一个网段时，需要使用到DHCPv6中继来完成IPv6地址/前缀和其他网络配置参数的获取。

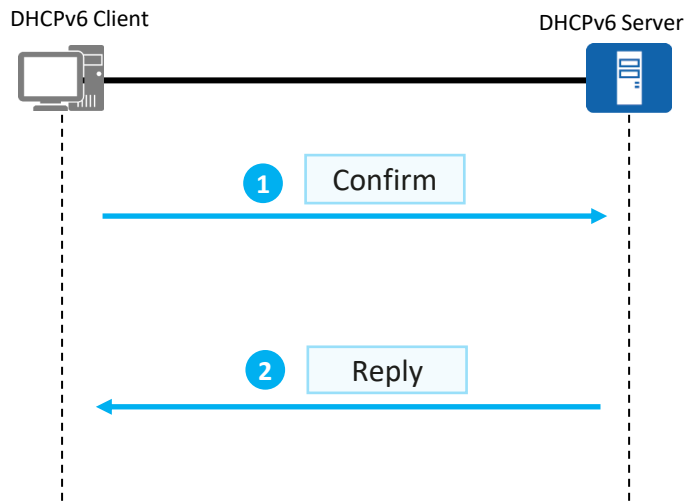


1. DHCPv6客户端向所有DHCPv6服务器和DHCPv6中继的组播地址FF02::1:2发送请求报文。
2. DHCPv6 中继收到客户端的报文后，将其封装在Relay-Forward报文的中继消息选项中，并将Relay-Forward报文发送给DHCPv6服务器或下一跳中继。
3. DHCPv6服务器从Relay-Forward报文中解析出DHCPv6客户端的请求，为DHCPv6客户端选取IPv6地址和其他配置参数，并将Relay-Reply报文发送给DHCPv6中继。
4. DHCPv6中继从Relay-Reply报文中解析出DHCPv6服务器的应答，转发给DHCPv6客户端。



DHCPv6地址确认过程

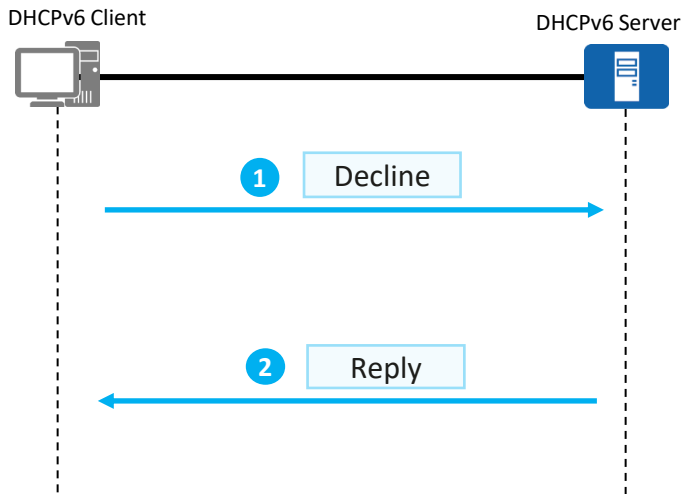
当客户端有断电、掉线、漫游等情况发生时，客户端会发送Confirm报文确认自己的IPv6地址是否可用。如果客户端确认的地址是合法的，则服务器回应；如果没有回应，则客户端需要重新启动地址申请流程。





DHCPv6地址冲突检测过程

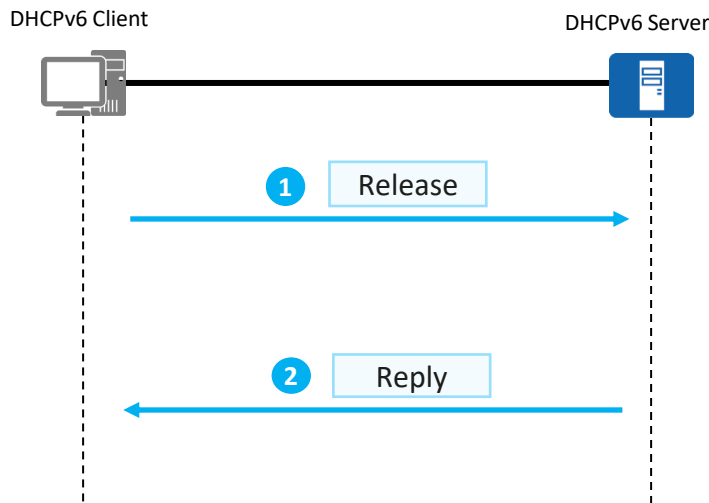
客户端完成地址申请后，会在开始使用该地址前发起DAD探测。如果DAD检测到地址存在冲突，则客户端发送Decline消息通知服务器，并不再使用该地址。





DHCPv6地址释放过程

当客户端不需要再使用某地址时，将发送Release消息至服务器，发起释放地址的交互流程。





DHCPv6报文总结

报文类型	报文作用
Solicit	DHCPv6客户端发送该消息，请求DHCPv6服务器为其分配IPv6地址/前缀和网络配置参数
Advertise	DHCPv6服务器发送Advertise消息，通知客户端可以为其分配的地址/前缀和网络配置参数
Request	如果DHCPv6客户端接收到多个服务器回复的Advertise消息，则根据消息接收的先后顺序、服务器优先级等，选择其中一台服务器，并向该服务器发送Request消息，请求服务器确认为其分配地址/前缀和网络配置参数
Reply	DHCPv6服务器发送Reply消息，确认将地址/前缀和网络配置参数分配给客户端使用
Information-Request	客户端向DHCPv6服务器发送Information-request报文，该报文中携带Option Request选项，指定客户端需要从服务器获取的配置参数
Renew	地址/前缀租借时间到达时间T1时，DHCPv6客户端会向为它分配地址/前缀的DHCPv6服务器单播发送Renew报文，以进行地址/前缀租约的更新
Rebind	如果在T1时发送Renew请求更新租约，但是没有收到DHCPv6服务器的回应报文，则DHCPv6客户端会在T2时，向所有DHCPv6服务器组播发送Rebind报文请求更新租约
Confirm	当有断电、掉线、漫游等情况发生时，客户端会发送Confirm报文确认自己的IP地址是否可用
Decline	当客户端发现地址冲突时，发送Decline通知服务器



地址自动配置比较

地址配置	基于DHCPv6	基于NDP
地址管理	有状态。服务器端存储用户地址或前缀的分配和释放信息	无状态。不保存用户地址分配信息
部署价值	支持128bit地址和不同长度的前缀分配，扩展性强	只支持64bit前缀配置，扩展性差
实现难度	配置复杂	配置简单
安全性	应用层协议，安全性强	安全性较差



IPv6基本配置 (1)

1. 使能设备转发IPv6单播报文的能力

```
[Huawei] ipv6
```

2. 使能接口转发IPv6单播报文的能力

```
[Huawei-GigabitEthernet0/0/0] ipv6 enable
```

3. 手工配置接口的链路本地地址

```
[Huawei-GigabitEthernet0/0/0] ipv6 address ipv6-address link-local
```

4. 自动配置接口的链路本地地址。

```
[Huawei-GigabitEthernet0/0/0] ipv6 address auto link-local
```

5. 手工配置接口的全球单播地址

```
[Huawei-GigabitEthernet0/0/0] ipv6 address { ipv6-address prefix-length | ipv6-address/prefix-length }
```

6. 自动配置接口的全球单播地址

```
[Huawei-GigabitEthernet0/0/0] ipv6 address auto { global | dhcp }
```



IPv6基本配置 (2)

7. 查看接口的IPv6信息

```
[Huawei] display ipv6 interface [ interface-type interface-number | brief ]
```

8. 查看邻居表项信息

```
[Huawei] display ipv6 neighbors
```

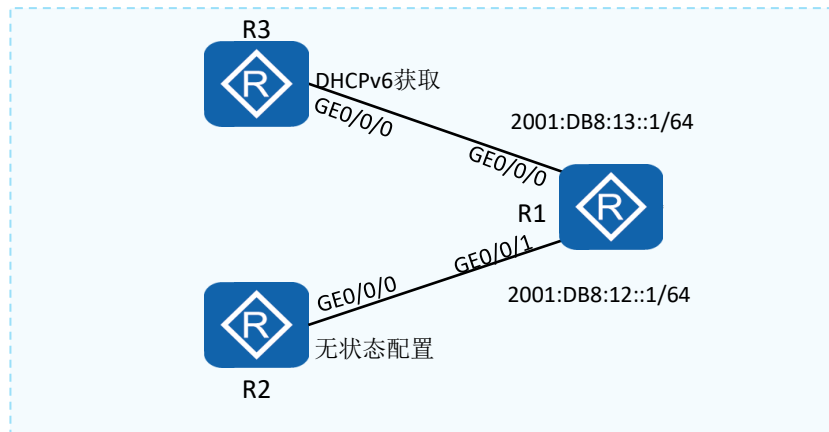
9. 使能接口发布RA报文功能

```
[Huawei-GigabitEthernet0/0/0] undo ipv6 nd ra halt
```

默认情况下，华为路由器接口抑制ICMPv6 RA报文的发送。此时，本网络的主机将不会定期收到更新IPv6地址前缀的信息。若需要周期性的向主机发布RA报文中的IPv6地址前缀和有状态自动配置标志位的信息时，使用undo ipv6 nd ra halt命令使能系统发布RA报文的功能。



IPv6地址配置举例 – 静态配置



配置需求:

- R1的GE0/0/0和GE0/0/1接口采用手工方式配置IPv6地址。
- R2的GE0/0/0接口通过无状态地址自动配置的方式获取IPv6地址。
- R3的GE0/0/0接口通过DHCPv6的方式获取IPv6地址。

1. 在R1、R2、R3使能IPv6，并且让接口自动生成链路本地地址（以R1为例）。

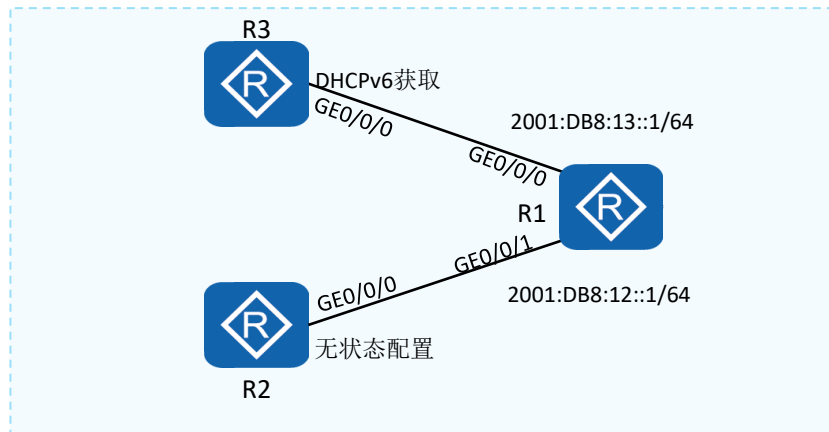
```
[R1]ipv6
[R1]interface GigabitEthernet 0/0/0
[R1-GigabitEthernet0/0/0]ipv6 enable
[R1-GigabitEthernet0/0/0]ipv6 address auto link-local
```

2. 在R1的接口上配置IPv6全球单播地址

```
[R1]interface GigabitEthernet 0/0/0
[R1-GigabitEthernet0/0/0]ipv6 address 2001:DB8:13::1 64 [R1-
GigabitEthernet0/0/0]interface GigabitEthernet 0/0/1
[R1-GigabitEthernet0/0/1]ipv6 address 2001:DB8:12::1 64
```



IPv6地址配置举例 – 无状态自动配置



配置需求:

- R1的GE0/0/0和GE0/0/1接口采用手工方式配置IPv6地址。
- R2的GE0/0/0接口通过无状态地址自动配置的方式获取IPv6地址。
- R3的GE0/0/0接口通过DHCPv6的方式获取IPv6地址。

1. 在R1使能发布RA报文的功能。

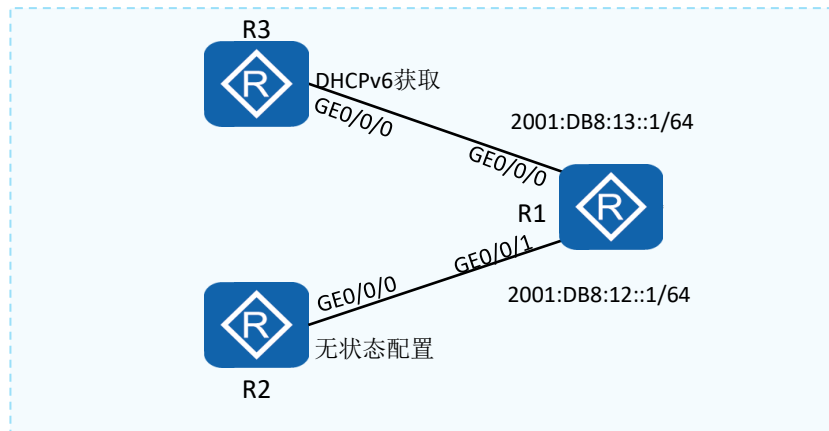
```
[R1]interface GigabitEthernet 0/0/1
[R1-GigabitEthernet0/0/1]undo ipv6 nd ra halt
```

2. 配置R2的GE0/0/0接口自动获取IPv6地址。

```
[R2]interface GigabitEthernet 0/0/0
[R2-GigabitEthernet0/0/0]ipv6 address auto global
```



IPv6地址配置举例 – DHCPv6



配置需求:

- R1的GE0/0/0和GE0/0/1接口采用手工方式配置IPv6地址。
- R2的GE0/0/0接口通过无状态地址自动配置的方式获取IPv6地址。
- R3的GE0/0/0接口通过DHCPv6的方式获取IPv6地址。

1. 在R1上配置DHCPv6服务器功能。

```
[R1]dhcp enable
[R1]dhcpv6 pool pool1
[R1-dhcpv6-pool-pool1]address prefix 2001:DB8:13::/64
[R1-dhcpv6-pool-pool1]excluded-address 2001:DB8:13::1
[R1]interface GigabitEthernet 0/0/0
[R1-GigabitEthernet0/0/0]dhcpv6 server pool1
```

2. 配置R3的GE0/0/0接口通过DHCPv6的方式获取IPv6地址。

```
[R3]dhcp enable
[R3]interface GigabitEthernet 0/0/0
[R3-GigabitEthernet0/0/0]ipv6 address auto dhcp
```



IPv6地址自动配置实现 – 结果验证

验证R2自动配置的IPv6地址:

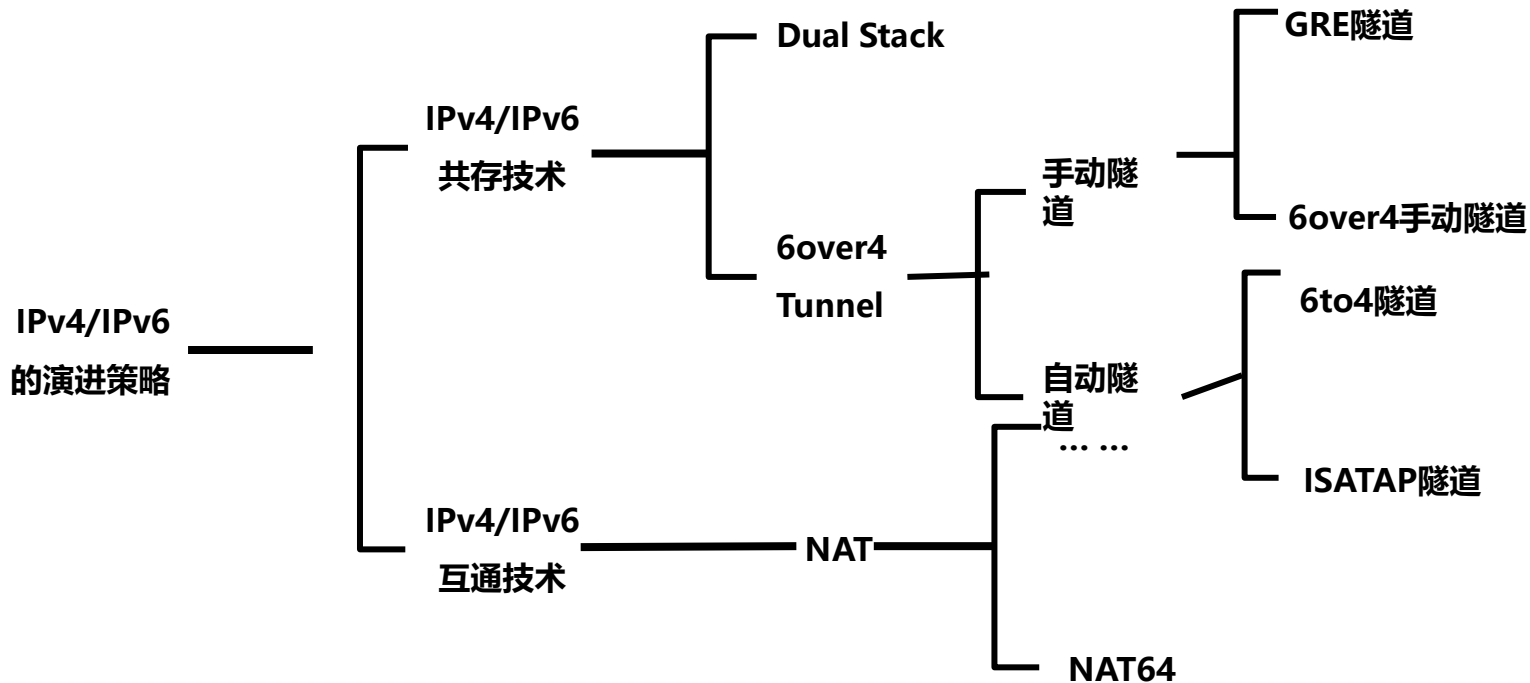
```
[R2]display ipv6 interface GigabitEthernet 0/0/0
GigabitEthernet0/0/0 current state : UP
IPv6 protocol current state : UP
IPv6 is enabled, link-local address is FE80::2E0:FCFF:FEAD:675F
Global unicast address(es):
  2001:DB8:12::2E0:FCFF:FEAD:675F,
  subnet is 2001:DB8:12::/64 [SLAAC]
Joined group address(es):
  FF02::1:FFAD:675F
  FF02::2
  FF02::1
MTU is 1500 bytes
ND DAD is enabled, number of DAD attempts: 1
ND reachable time is 30000 milliseconds
ND retransmit interval is 1000 milliseconds
Hosts use stateless autoconfig for addresses
```

在R3上使用display dhcpv6 client命令查看地址分配信息:

```
[R3]display dhcpv6 client
GigabitEthernet0/0/0 is in stateful DHCPv6 client mode.
State is BOUND.
Preferred server DUID : 0003000100E0FC401312
  Reachable via address : FE80::2E0:FCFF:FE40:1312
IA NA IA ID 0x00000031 T1 43200 T2 69120
  Obtained : 2020-06-17 09:49:58
  Renews : 2020-06-17 21:49:58
  Rebinds : 2020-06-18 05:01:58
  Address : 2001:DB8:13::3
Lifetime valid 172800 seconds, preferred 86400 seconds
Expires at 2020-06-19 09:49:58(172159 seconds left)
```



IPv6过渡技术简介





双栈Dual Stack

- 双栈协议
 - 设备必须支持IPv4/IPv6协议栈。
 - 连接双栈网络的接口必须同时配置IPv4地址和IPv6地址。

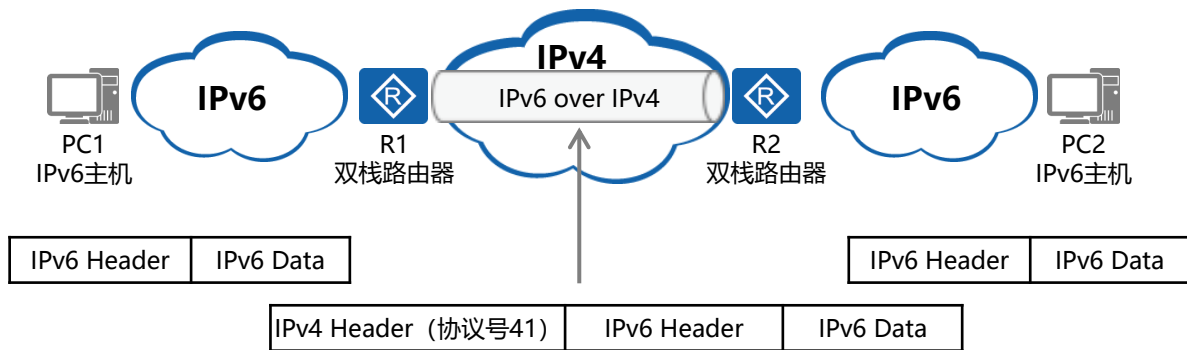


```
[R1]ipv6
[R1]interface GigabitEthernet 0/0/0
[R1-GigabitEthernet0/0/0]ip address 10.1.1.1 24
[R1-GigabitEthernet0/0/0]ipv6 enable
[R1-GigabitEthernet0/0/0]ipv6 address 2001:0001::FFFF 64
```



6over4手动隧道

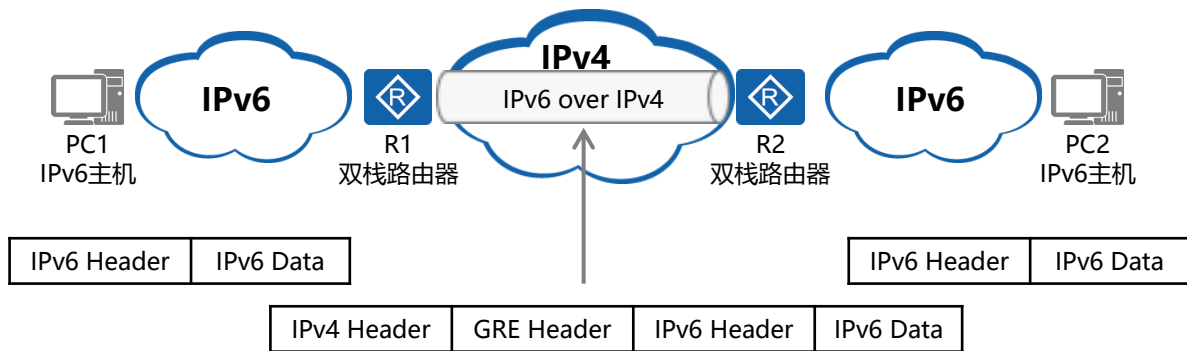
- 6over4手动隧道。
 - 6over4手动隧道的一种；
 - 源地址和目的地址均需手工指定；
 - 用于边界路由器与边界路由器，或者主机与边界路由器之间。





6over4 GRE隧道

- 6over4 GRE隧道。
 - 6over4手动隧道的一种；
 - 手工指定隧道的端点地址；
 - GRE承载IPv6协议。



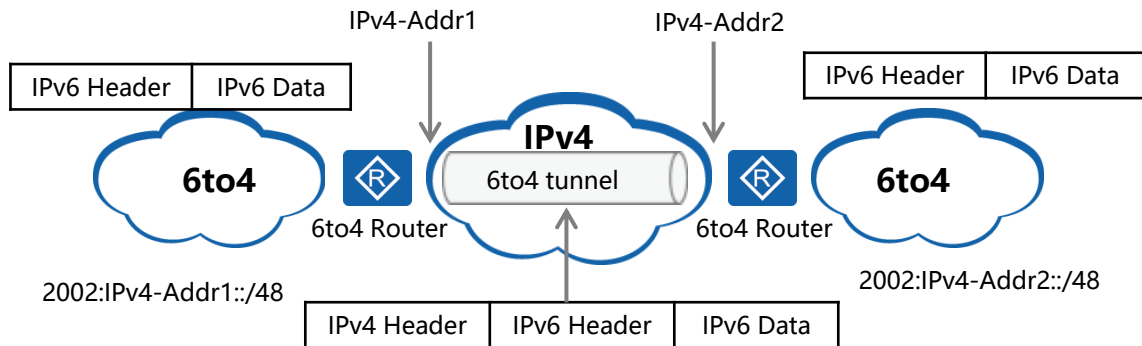


6to4隧道 (1)

- 6to4隧道
 - 6over4自动隧道的一种。
 - 支持Router到Router、Host到Router、Router到Host、Host到Host。
 - 采用6to4专用地址，即2002:IPv4::/48。

地址格式

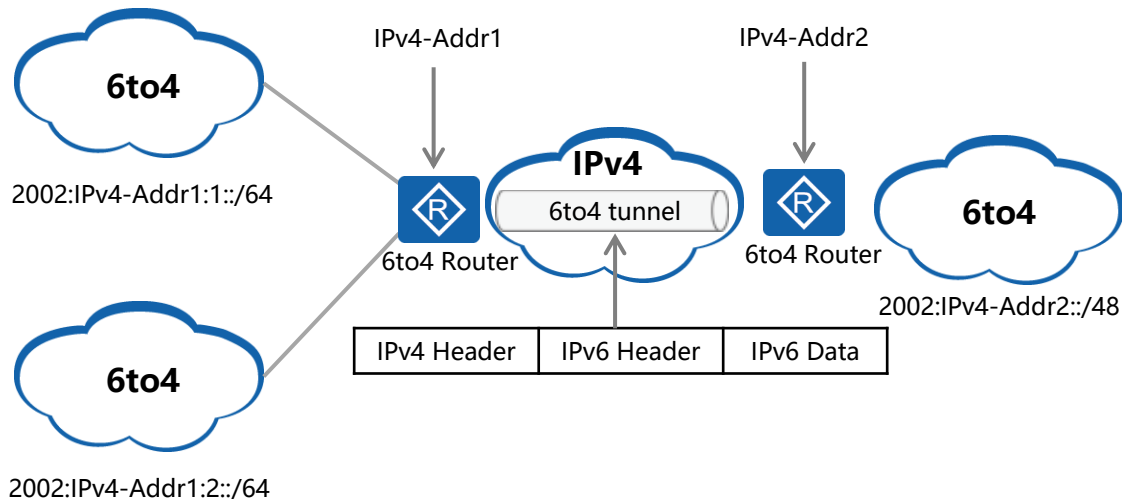
FP	TLA	IPv4 address	SLA ID	Interface ID
----	-----	--------------	--------	--------------





6to4隧道 (2)

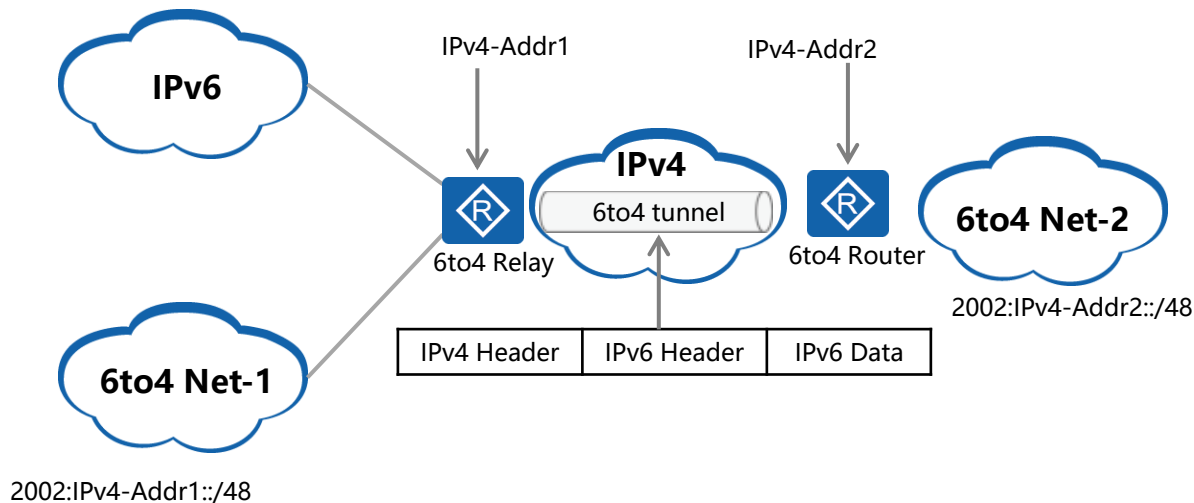
- 6to4隧道
 - 可连接多个6to4网络。
 - 通过SLA ID区分。





6to4隧道 (3)

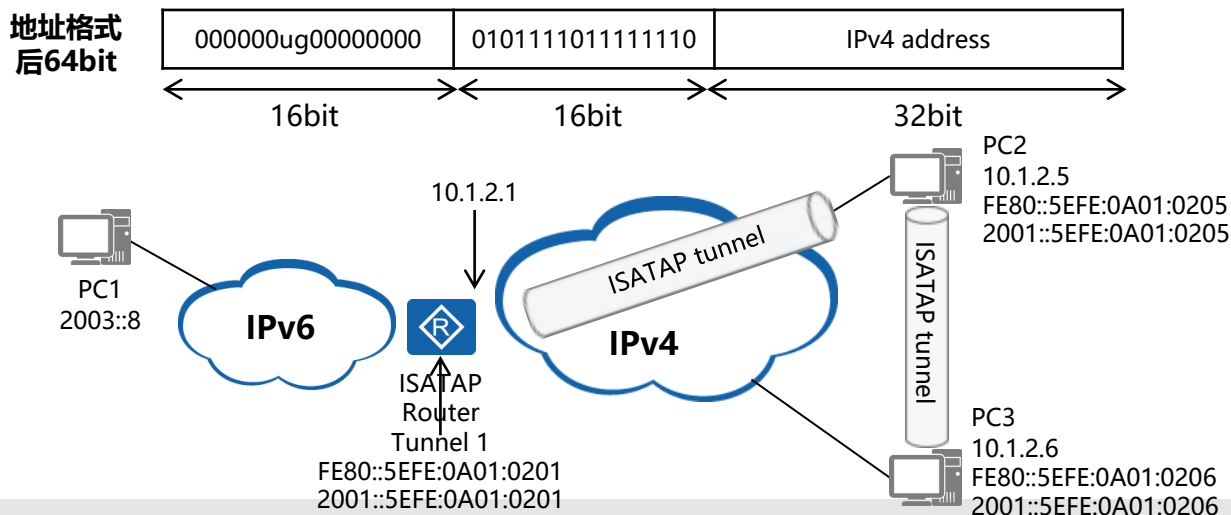
- 6to4中继
 - 实现6to4网络和IPv6普通网络互通。





ISATAP隧道

- ISATAP隧道
 - 6over4自动隧道的一种
 - 支持Host到Router、Router到Host、Host到Host
 - 采用ISATAP隧道专用地址





NAT64

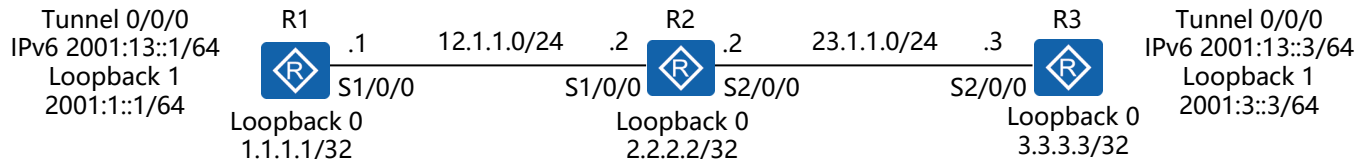
- NAT64技术实际上是一种协议转换技术，能够将分组在V4及V6格式之间灵活转换。
- IPv6过渡中的协议翻译技术就是将IPv6数据包的每个字段与IPv4数据包中的字段建立起一一映射的关系，从而在两个网络的边缘实现数据报文的转换。





配置GRE隧道 (1)

- 公司A网络拓扑如下所示，现根据需求完成如下配置：
 - R1、R2和R3的IPv4地址如图所示，部署在OSPFv2的区域0中，该部分配置已经完成；
 - 所需的IPv6地址已经标出；
 - 采用IPv6 over IPv4 GRE隧道的形式，实现R1与R3的Loopback1之间的互通。



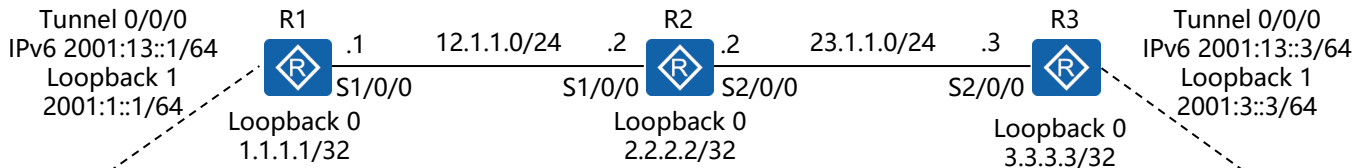


GRE隧道 - 基本配置命令

步骤	操作	命令
1	进入系统视图	system-view
2	创建Tunnel接口	interface tunnel <i>interface-number</i>
3	指定Tunnel为GRE隧道模式	tunnel-protocol gre
4	指定Tunnel的源接口	source { <i>ipv4-address</i> <i>interface-type interface-number</i> }
5	指定Tunnel的目的接口	destination <i>ipv4-address</i>
6	设置对GRE报文头进行校验	gre checksum
7	设置GRE报文头的关键字	gre key <i>key-number</i>
8	设置Tunnel接口的IPv6地址	ipv6 address <i>ipv6-address prefix-length</i>



配置GRE隧道 (2)



```
interface Tunnel0/0/0
ipv6 enable
ipv6 address 2001:13::1/64
tunnel-protocol gre
source LoopBack0
destination 3.3.3.3
#
ospf 1
area 0.0.0.0
network 1.1.1.1 0.0.0.0
network 12.1.1.0 0.0.0.255
#
ipv6 route-static 2001:3:: 64
Tunnel0/0/0
```

```
[R3]display ipv6 interface Tunnel 0/0/0
Tunnel0/0/0 current state : UP
IPv6 protocol current state : UP
IPv6 is enabled, link-local address is FE80::303:303
Global unicast address(es):
  2001:13::3, subnet is 2001:13::/64
Joined group address(es):
  FF02::1:FF03:303
  FF02::2
  FF02::1
  FF02::1:FF00:3
.....
```

```
interface Tunnel0/0/0
ipv6 enable
ipv6 address 2001:13::3/64
tunnel-protocol gre
source LoopBack0
destination 1.1.1.1
#
ospf 1
area 0.0.0.0
network 3.3.3.3 0.0.0.0
network 23.1.1.0 0.0.0.255
#
ipv6 route-static 2001:1:: 64 Tunnel0/0/0
```



ipv6-gre.rar

<R3> ping ipv6 2001:1::1

```
Reply from 2001:1::1 bytes=56 Sequence=1 hop limit=63 time = 81 ms
Reply from 2001:1::1 bytes=56 Sequence=2 hop limit=63 time = 62 ms
Reply from 2001:1::1 bytes=56 Sequence=3 hop limit=63 time = 63 ms
Reply from 2001:1::1 bytes=56 Sequence=4 hop limit=63 time = 63 ms
Reply from 2001:1::1 bytes=56 Sequence=5 hop limit=63 time = 63 ms
```

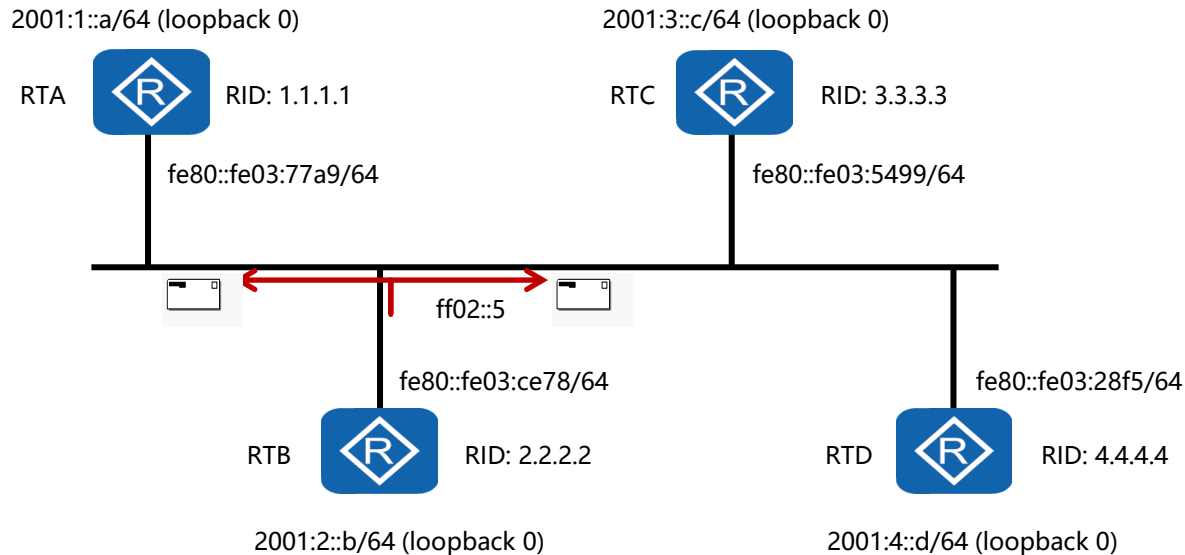



思考题？

- 2001:0DB8:0000:0000:0000:0000:032A:2D70，此IPv6地址压缩到最短是多少？
- IPv6主机无状态地址自动配置的过程是什么？



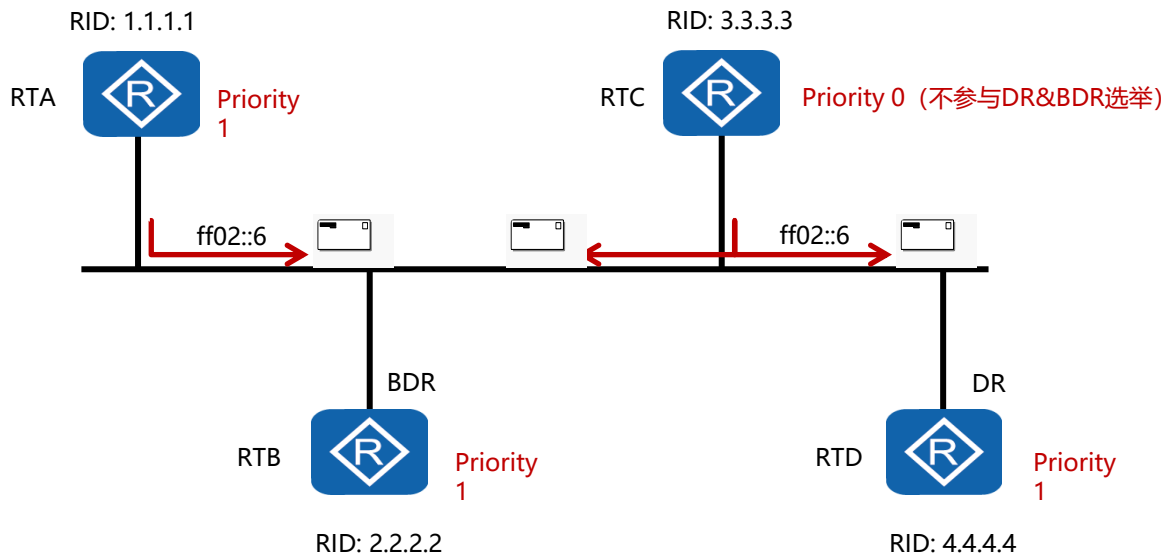
OSPFv3



- ff02::5是为OSPFv3路由协议预留的IPv6组播地址。
- OSPFv3中的路由条目下一跳地址是链路本地地址。



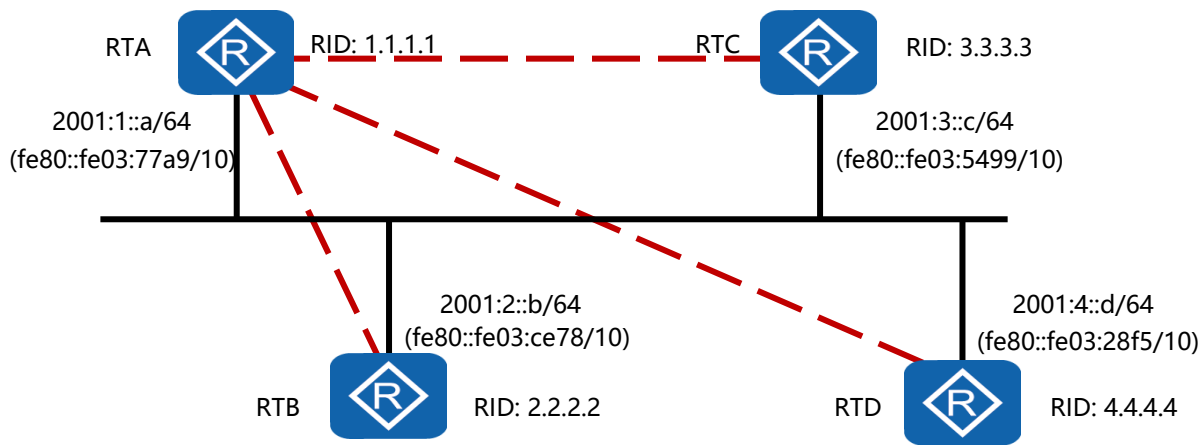
DR&BDR



- **Router ID在OSPFv3中必须手动配置。**
- **在NBMA和广播型网络中OSPFv3选举DR和BDR的过程与OSPFv2相似。**



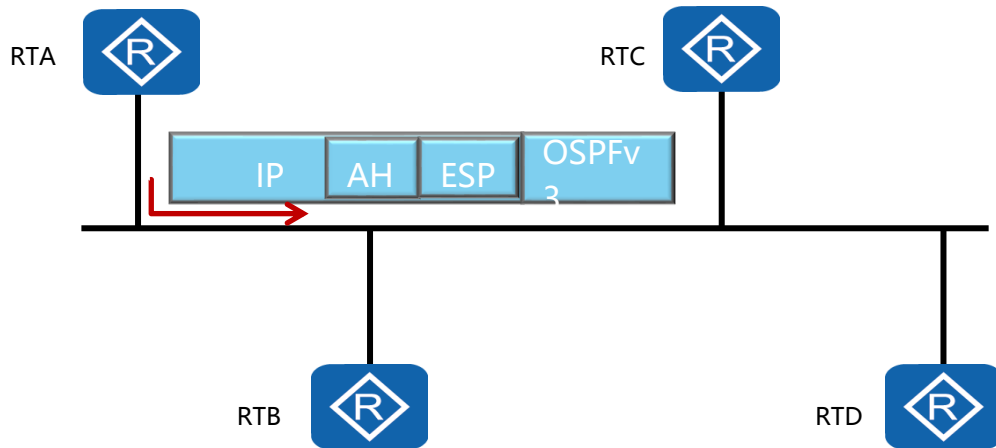
基于链路运行



- OSPFv2是基于网段运行的， OSPFv3的实现是基于链路的。



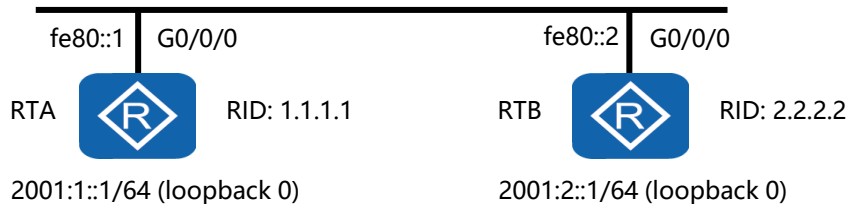
OSPFv3认证



- **OSPFv3**协议本身不提供认证功能，而是通过使用**IPv6**提供的安全机制来保证**OSPFv3**报文的合法性。



OSPFv3配置



```
[RTA]ipv6
[RTA]ospfv3
[RTA-ospfv3-1]router-id 1.1.1.1
[RTA-GigabitEthernet0/0/0]ipv6 enable
[RTA-GigabitEthernet0/0/0]ipv6 address fe80::1 link-local
[RTA-GigabitEthernet0/0/0]ospfv3 1 area 0.0.0.0
[RTA-LoopBack0]ipv6 enable
[RTA-LoopBack0]ipv6 address 2001:1::1/64
[RTA-LoopBack0]ospfv3 1 area 0.0.0.0
```



配置验证

```
[RTA]display ospfv3
Routing Process "OSPFv3 (1)" with ID 1.1.1.1
Route Tag: 0
Multi-VPN-Instance is not enabled
SPF Intelligent Timer[millisecs] Max: 10000, Start: 500, Hold: 2000
LSA Intelligent Timer[millisecs] Max: 5000, Start: 500, Hold: 1000
LSA Arrival interval 1000 millisecs
Default ASE parameters: Metric: 1 Tag: 1 Type: 2
Number of AS-External LSA 0. AS-External LSA's Checksum Sum 0x0000
Number of AS-Scoped Unknown LSA 0. AS-Scoped Unknown LSA's Checksum
Sum 0x0000
Number of FULL neighbors 1
Number of Exchange and Loading neighbors 0
.....
```



IPv6静态路由

- IPv6静态路由与IPv4静态路由类似，也需要管理员手工配置，适合于一些结构比较简单的IPv6网络。
- 在创建IPv6静态路由时，可以同时指定出接口和下一跳，或者只指定出接口或只指定下一跳。
 - 对于点到点接口：指定出接口。
 - 对于广播类型接口：指定下一跳。
- IPv6静态路由负载分担和备份：
 - 在创建相同目的地址的多条IPv6静态路由时，如果指定相同优先级，则可实现负载分担，如果指定不同优先级，则可实现路由备份。



IPv6静态路由的基础配置命令

1. 在公网上配置IPv6静态路由。

```
[Huawei] ipv6 route-static dest-ipv6-address prefix-length { interface-type interface-number [ nexthop-ipv6-address ]  
| nexthop-ipv6-address | vpn-instance vpn-destination-name nexthop-ipv6-address }  
[ preference preference ][ permanent | inherit-cost ] [ description text ]
```

2. 在VPN实例下配置IPv6静态路由。

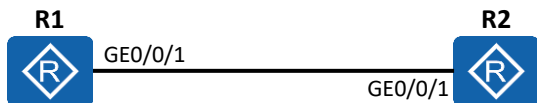
```
[Huawei] ipv6 route-static vpn-instance vpn-instance-name dest-ipv6-address prefix-length { [ interface-type interface-  
number [ nexthop-ipv6-address ] ] | nexthop-ipv6-address [ public ] | vpn-instance vpn-destination-name nexthop-ipv6-  
address } [ preference preference ] [ permanent | inherit-cost ] [ description text ]
```

- 注意事项:

- 静态路由如果不配置优先级，默认优先级为60。
- 如果将目的地址与前缀长度都配置为全0 (::0)，则表示配置的是缺省路由 (::/0)。
- undo命令中配置参数permanent时，只能取消IPv6静态路由永久发布，不能删除IPv6静态路由。



IPv6静态路由配置举例



设备	接口	IPv6地址
R1	GE0/0/1	2001:DB8:2345:12::1/64
	Loopback0	2001:DB8:2345:1::1/128
R2	GE0/0/1	2001:DB8:2345:12::2/64
	Loopback0	2001:DB8:2345:2::2/128

场景描述:

- 某公司网络部署了IPv6网络进行业务测试，在部署初期，通过静态路由实现IPv6网络的互联互通。
- 实现R1和R2的Loopback0接口地址互通。

1、配置各路由器接口的IPv6地址。（略）

2、配置IPv6静态路由。

```
[R1] ipv6 route-static 2001:DB8:2345:2::2 128 2001:DB8:2345:12::2
```

```
[R2] ipv6 route-static 2001:DB8:2345:1::1 128 2001:DB8:2345:12::1
```

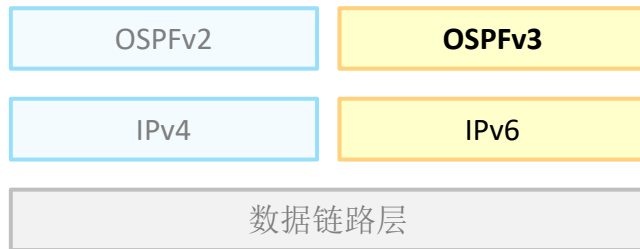
3、测试IPv6网络连通性。

```
[R1] ping ipv6 2001:DB8:2345:2::2
PING 2001:DB8:2345:2::2 : 56 data bytes, press CTRL_C to break
Reply from 2001:DB8:2345:2::2
bytes=56 Sequence=1 hop limit=64 time = 90 ms
Reply from 2001:DB8:2345:2::2
bytes=56 Sequence=2 hop limit=64 time = 20 ms
Reply from 2001:DB8:2345:2::2
bytes=56 Sequence=3 hop limit=64 time = 20 ms
Reply from 2001:DB8:2345:2::2
bytes=56 Sequence=4 hop limit=64 time = 20 ms
Reply from 2001:DB8:2345:2::2
bytes=56 Sequence=5 hop limit=64 time = 20 ms
--- 2001:DB8:2345:2::2 ping statistics ---
5 packet(s) transmitted
5 packet(s) received
0.00% packet loss
round-trip min/avg/max = 20/34/90 ms
```



OSPFv3概述

- OSPF是IETF定义的一种基于链路状态的内部网关路由协议。目前针对IPv4协议使用的是OSPF Version 2（OSPFv2），针对IPv6协议使用OSPF Version 3（OSPFv3）。
- OSPFv3的主要目的是开发一种独立于任何具体网络层的路由协议。为实现这一目的，OSPFv3的内部路由器信息被重新进行了设计。



OSPFv2是运行在IPv4上的IGP路由协议；OSPFv3是运行在IPv6上的IGP路由协议。两者并不兼容。



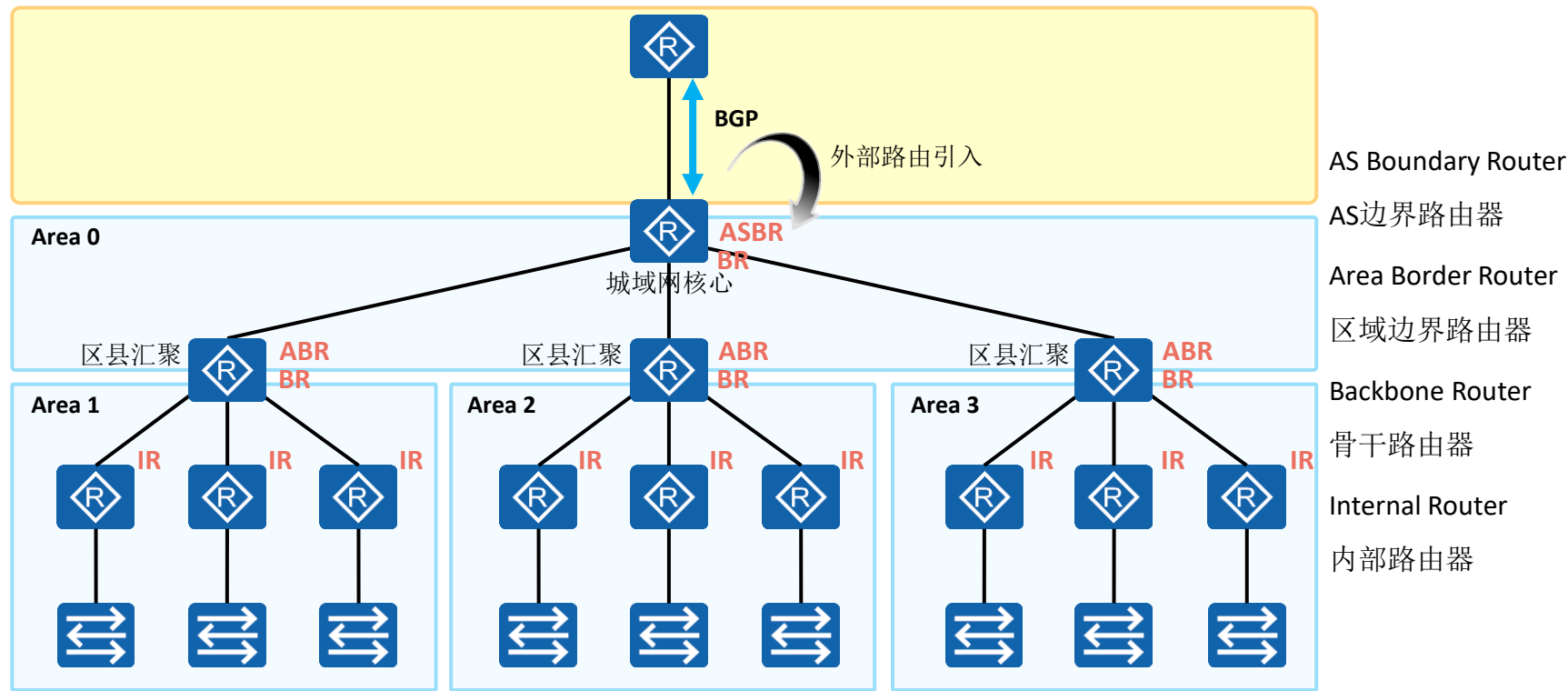
OSPFv3与OSPFv2的相同点

OSPF的基本运行机制没有改变，包括：

- 基本概念：
 - 区域划分及路由器类型
 - 路由计算影响参数：优先级、度量值
 - 支持的网络类型：Broadcast（广播类型）、NBMA、P2P（点到点类型）、P2MP（点到多点类型）
 - 报文类型：Hello报文、DD报文、LSR报文、LSU报文和LSAck报文
- 工作原理：
 - 邻居关系的建立及邻居状态的转换
 - DR与BDR的选举
 - LSA泛洪机制
 - 路由计算过程

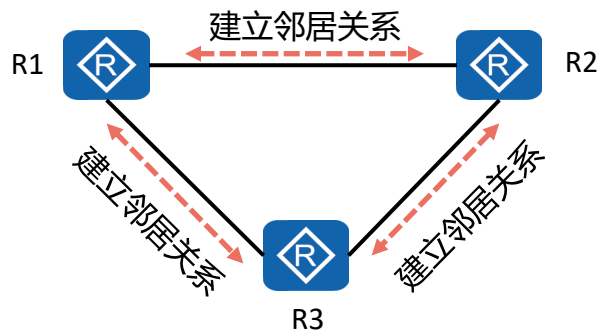


OSPFv3拓扑和路由器类型



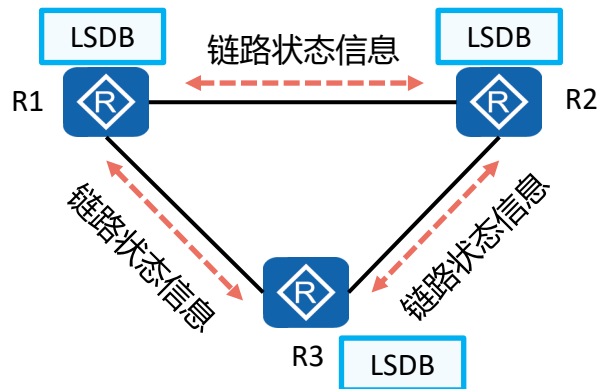


OSPFv3的基本工作原理与OSPFv2相似



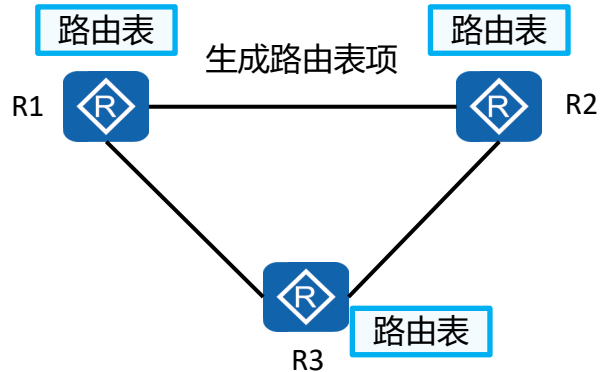
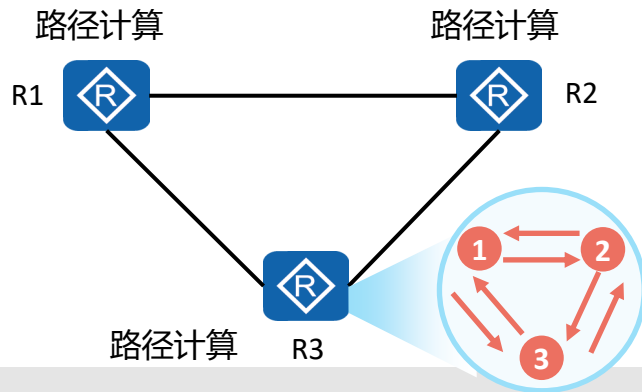
1

2



3

4





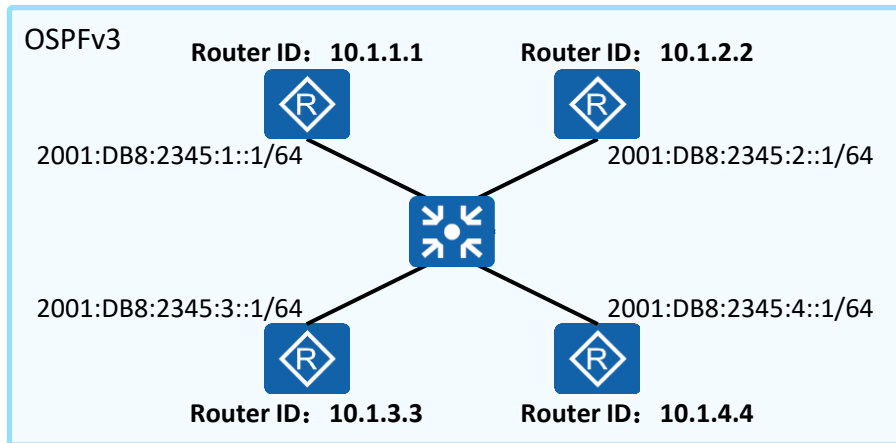
OSPFv3与OSPFv2的不同点

- OSPFv3基于链路运行以及拓扑计算，而不再是网段。
- OSPFv3支持一个链路上多个实例。
- OSPFv3报文和LSA中去掉了IP地址的意义，且重构了报文格式和LSA格式。
 - OSPFv3报文和Router LSA/Network LSA中不包含IP地址。
 - OSPFv3的LSA中定义了LSA的泛洪范围。
 - OSPFv3中创建了新的LSA承载IPv6地址和前缀。
 - OSPFv3邻居不再由IP地址标识，只由Router ID标识。



唯一邻居标识: Router ID

OSPFv3通过Router ID来标识网络设备。



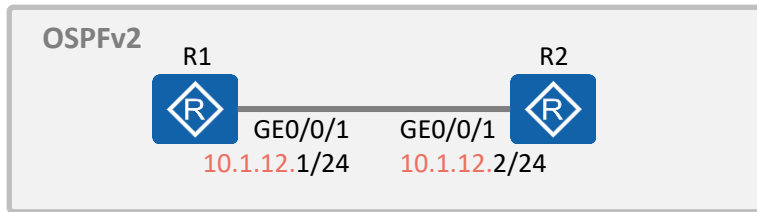
- Router ID是一个OSPFv3设备在自治系统中的唯一标识。如果用户没有指定Router ID，则OSPFv3进程无法运行。
- 设置Router ID时，必须保证自治系统中任意两台设备的Router ID都不相同。
- Router ID长度32bit，本地标识符，与IPv6地址无关，用点分十进制表示法来表示。



OSPFv3基于链路运行

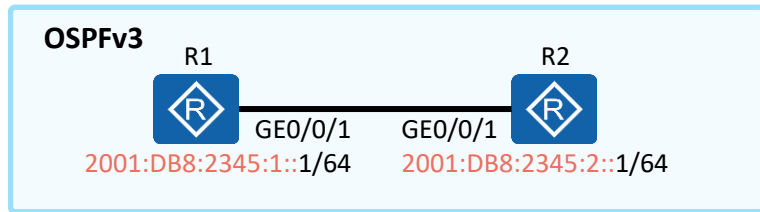
OSPFv3是基于链路运行的，设备只要在同一链路，就可以建立邻居关系。

OSPFv2基于网段运行



- R1与R2之间运行OSPFv2。
- R1的GE0/0/1接口与R2的GE0/0/1接口地址在**同一网段**，R1与R2才可以建立OSPFv2邻居关系。
- OSPFv2运行时，邻居双方必须处于相同网段。

OSPFv3基于链路运行

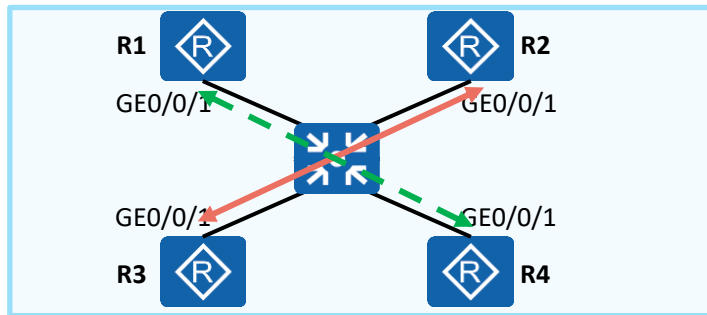


- R1与R2之间运行OSPFv3。
- R1的GE0/0/1接口与R2的GE0/0/1接口地址**不在同一网段**，但是R1与R2仍然可以建立OSPFv3邻居关系。
- OSPFv3运行时，仅需邻居双方处于相同直连链路即可。

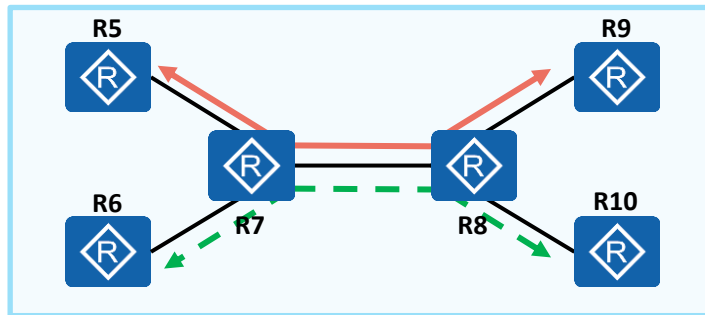


链路支持多实例

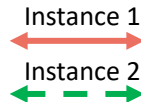
- 一个OSPFv3物理接口可以和多个实例绑定，并用不同的实例标识（Instance ID）区分，即OSPFv3的单个链路支持运行多个OSPFv3实例。
- 这些运行在同一条物理链路上的多个OSPFv3实例，分别与链路对端设备建立邻居及发送报文，且互不干扰，这样可以充分共享同一链路资源。



通过Instance ID可以实现R2和R3建立OSPFv3邻居关系，R1和R4也建立OSPFv3邻居关系。



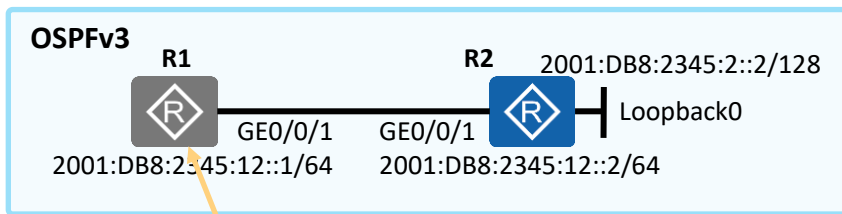
可以在R7和R8各自创建2个OSPFv3实例，并且在一条链路上将两个实例激活，通过Instance ID可以实现两部分邻居关系的建立。





OSPFv3对链路本地地址的使用

- OSPFv3使用链路本地（FE80::/10）地址作为发送报文的源地址和路由的下一跳地址。
 - 使用链路本地地址来维持邻居关系，同步LSA数据库。
 - 在虚连接上，必须使用全球单播地址或者站点本地地址作为OSPFv3协议报文的源地址。
- 优势：
 - 不需要配置IPv6全球单播地址，就可以得到OSPFv3拓扑，实现拓扑与地址分离。
 - OSPFv3报文不会被转发到始发链路范围之外，减少了报文不必要的泛洪，节省了带宽。



```
[R1]display ipv6 routing-table
```

```
.....
```

Destination:	2001:DB8:2345:2::2	PrefixLength : 128
NextHop:	FE80::2E0:FCFF:FE82:522B	Preference : 10
Cost:	1	Protocol : OSPFv3
RelayNextHop:	::	TunnelID : 0x0
Interface:	GigabitEthernet0/0/1	Flags : D

```
.....
```



OSPFv3报文

- OSPFv3与OSPFv2有相同类型的报文：
 - Hello报文、DD报文、LSR报文、LSU报文和LSAck报文。
- OSPFv3与OSPFv2使用相同的协议号89。
 - OSPFv2: IPv4报文头部中的协议号（Protocol）为89。
 - OSPFv3: IPv6报文头部中的下一报头号（Next Header）为89。
- OSPFv3与OSPFv2类似，使用组播地址作为OSPF报文目的地址。
 - OSPFv2使用IPv4组播地址：
 - OSPF IGP Routers: 224.0.0.5; OSPF IGP DR : 224.0.0.6。
 - OSPFv3使用IPv6组播地址：
 - OSPF IGP Routers: FF02::5; OSPF IGP DR: FF02::6。



OSPFv3报文头部

- OSPFv2报文头部:

Version=2	Type	Packet Length
Router ID		
Area ID		
Checksum	Auth Type	
Authentication		

- OSPFv3报文头部:

Version=3	Type	Packet Length
Router ID		
Area ID		
Checksum	Instance ID	0

- 与OSPFv2一样，OSPFv3的五种报文都有同样的报文头，只是报文中的字段有些不同。
- 变化点:
 - **Instance ID:** 1Byte，缺省值为0。允许在一个链路上运行多个OSPFv3的实例。每个实例具有唯一的Instance ID。Instance ID只在本地链路上有意义。
 - **OSPFv3报文头部移除了所有的认证字段:** OSPFv3的认证可以使用IPv6的认证及安全处理，也可以通过OSPFv3自身机制来完成报文认证。



Hello报文

- OSPFv2的Hello报文格式:

Network Mask		
HelloInterval	Options	Rtr Pri
RouterDeadInterval		
Designated Router		
Backup Designated Router		
Neighbor		

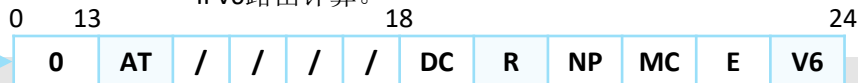
- OSPFv3的Hello报文格式:

Interface ID	
Rtr Pri	Options
HelloInterval	RouterDeadInterval
Designated Router ID	
Backup Designated Router ID	
Neighbor ID	

- 与OSPFv2的Hello报文相比，OSPFv3的Hello报文去掉了Network Mask字段，增加了Interface ID字段，用来标识发送该Hello报文的接口ID。

- 变化点:

- Interface ID:** 4Byte，唯一标识了建立连接的（发送Hello报文的）接口。
- Options:** 扩展到3Byte，可选项。与OSPFv2相比，增加了R位和V6位。
 - R:** 指明始发路由器是否具备转发能力。设置R位为0时，描述该始发节点的路由信息将不参与路由计算。
 - V6:** 如果V6位为0，表示该路由器或链路不会参与IPv6路由计算。





OSPFv3的LSA头部

- OSPFv2

LS Age	Options	LS Type
Link State ID		
Advertising Router		
LS Sequence Number		
LS Checksum	Length	

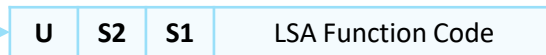
- OSPFv3

LS Age	LS Type
Link State ID	
Advertising Router	
LS Sequence Number	
LS Checksum	Length

- 与OSPFv2的LSA头部相比，OSPFv3的LSA头部去掉了Options字段，并扩充了LS Type字段。OSPFv3同样使用LS Type、Link State ID和Advertising Router三元组唯一地标识一个LSA。

- 变化点：

- LS Type: 扩展到2Byte, Link State Type, 链路状态类型。
 - U: 1bit, 标识对未知LSA的处理方法。
 - S2/S1: 2bit, 标识LSA的泛洪范围。
 - LSA Function Code: 13bit, LSA功能代码, 标识了LSA的类型。





OSPFv3的LSA类型

- OSPFv3与OSPFv2相比，具有类似的LSA名称，但是功能略有区别。
- OSPFv3新增了两类LSA，包括：链路LSA和区域内前缀LSA。

OSPFv2的LSA		OSPFv3的LSA		异同点说明
类型	名称	类型	名称	
1	Router-LSA（路由器LSA）	0x2001	Router-LSA（路由器LSA）	作用类似，但不再描述地址信息，仅描述拓扑结构
2	Network-LSA（网络LSA）	0x2002	Network-LSA（网络LSA）	
3	Network-Summary-LSA（网络汇总LSA）	0x2003	Inter-Area-Prefix-LSA（区域间前缀LSA）	作用类似，名称不同
4	ASBR-Summary-LSA（ASBR汇总LSA）	0x2004	Inter-Area-Router LSA（区域间路由器LSA）	
5	AS-External-LSA（AS外部LSA）	0x4005	AS-External-LSA（AS外部LSA）	作用与名称完全相同
7	NSSA LSA（非完全末梢区域LSA）	0x2007	NSSA LSA（非完全末梢区域LSA）	
		0x0008	Link-LSA（链路LSA）	OSPFv3新增
		0x2009	Intra-Area-Prefix-LSA（区域内前缀LSA）	



Type1: Router-LSA

- OSPFv2

0	V	E	B	0	# Links
Link ID					
Link Data					
Link Type		# Tos		Metric	

- OSPFv3

0	W	V	E	B	Options	
Link Type				0		Metric
Interface ID						
Neighbor Interface ID						
Neighbor Router ID						

- 在OSPFv2中，通过Link Type、Link ID以及Link Data来描述一个接口信息。在OSPFv3中，设备会为每个运行OSPFv3接口所在的区域产生一个LSA，描述了设备的链路状态（Link Type、Interface ID、Neighbor Interface ID和Neighbor Router ID）和开销，在所属的区域内传播。
- 重要字段介绍：
 - Link Type: 1Byte，链路类型。
 - Interface ID: 4Byte，接口ID。
 - Neighbor Interface ID: 4Byte，邻居的接口ID。
 - Neighbor Router ID: 4Byte，邻居的路由器ID。



Type2: Network-LSA

- OSPFv2

Network Mask
Attached Router
Attached Router
.....

- OSPFv3

0	Options
Attached Router	
Attached Router	
.....	

- 与OSPFv2相比，OSPFv3的Network-LSA删除了网络掩码字段，仅用相连的路由器的Router ID来描述本网段的链路状态，由DR产生，在所属的区域内传播。
- 重要字段介绍：
 - Attached Router: 4Byte，相连的路由器，指连接在同一个网段上的所有路由器的Router ID，也包括DR的Router ID。



Type3: Inter-Area-Prefix-LSA

- OSPFv2: Network-Summary-LSA

Network Mask	
0	Metric
.....	

- OSPFv3

0	Metric	
PrefixLength	PrefixOptions	0
Address Prefix		
.....		

- PrefixOptions:

	P	MC	LA	MU
--	---	----	----	----

- 与OSPFv2类似，由ABR产生，描述区域内某个IPv6地址前缀的路由，并通告给其他相关区域。每个IPv6地址前缀，ABR都会单独发送一个Type3的LSA。
- 重要字段介绍（描述前缀的三元组）：
 - PrefixLength: 1Byte，前缀的比特数。
 - PrefixOptions: 1Byte，表示这个前缀的一些特性，以便在各种不同的路由计算时做相应的判断和处理。
PrefixOptions字段格式如下：
 - P位: 1bit，传播位。
 - MC位: 1bit，组播位。
 - LA位: 1bit，本地地址位。
 - NU位: 1bit，非单播位。
 - Address Prefix: 变长，IPv6地址前缀。



Type4: Inter-Area-Router-LSA

- OSPFv2: ASBR-Summary-LSA

Network Mask	
0	Metric
.....	

- OSPFv3

0	Options
0	Metric
Destination Router ID	

- 与OSPFv2类似，由ABR产生，描述到ASBR的路由，通告给除ASBR所在区域的其他相关区域。对于所描述的每一个ASBR，ABR都会单独发送一个Type4的LSA。
- 重要字段介绍：
 - Destination Router ID: 4Byte，LSA中描述的目的路由器的Router ID，即ASBR的Router ID。



Type5: AS-External-LSA

- OSPFv2

Network Mask		
E	0	Metric
Forwarding Address		
External Route Tag		
.....		

- OSPFv3

0	E	F	T	Metric	
PrefixLength		PrefixOptions		Referenced LS Type	
Address Prefix					
Forwarding Address (Optional)					
External Route Tag (Optional)					
Referenced Link State ID (Optional)					

- 与OSPFv2类似，由ASBR产生，描述到达AS外部的一个前缀的路由，通告到所有的区域（除了Stub区域和NSSA区域）。
- 重要字段介绍：
 - Referenced LS Type: 2Byte，引用链路状态类型，表明这个LSA是否需要参考其他LSA。
 - 0: 不参考:
 - 1: 参考Router-LSA
 - 2: 参考Network-LSA



新增Type8: Link-LSA

Link-LSA报文格式:

Rtr Pri		Options	
Link-Local Interface Address			
# Prefixes			
PrefixLength	PrefixOptions	0	
Address Prefix			
.....			
.....			
PrefixLength	PrefixOptions	0	
Address Prefix			
.....			

- 每个设备都会为每个链路产生一个Link-LSA，仅在始发链路内泛洪。
- Link-LSA作用：
 - 向该链路上其他路由器通告本接口的链路本地地址。
 - 向该链路上其他路由器通告本接口的IPv6前缀列表。
 - 向该链路上其他路由器通告本链路始发的Network-LSA中设置的可选项。
- 重要字段介绍：
 - Link-Local Interface Address: 16Byte，路由器与该链路相连的接口上配置的链路本地地址(该地址只出现在Link-LSA中)。



新增Type9: Intra-Area-Prefix-LSA

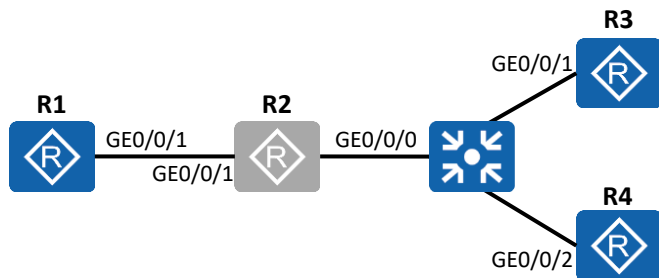
Intra-Area-Prefix-LSA报文格式:

Number of Prefixes		Referenced LS Type
Referenced Link State ID		
Referenced Advertising Router		
PrefixLength	PrefixOptions	Metric
Address Prefix		
.....		
.....		

- 在OSPFv2中, 可以通过Type1和Type2的LSA来描述拓扑信息和网段信息; 而OSPFv3的此两类LSA仅包含拓扑信息, 那OSPFv3如何描述网段信息呢?
- Type9 LSA描述的是网段信息, 只在所属的区域内传播, 它需要依赖于拓扑信息, 才能实现OSPFv3的路由计算。其类型可以分为两种:
 - 每台设备均产生描述与Router-LSA相关联的IPv6前缀地址的Type9 LSA。
 - DR会产生描述与Network-LSA相关联的IPv6前缀地址的Type9 LSA。



OSPFv3的LSA举例



设备	Router ID	接口	接口地址
R1	10.1.1.1	GE0/0/1	2001:DB8:2345:12::1/64
R2	10.1.2.2	GE0/0/1	2001:DB8:2345:12::2/64
		GE0/0/0	2001:DB8:2345:23::2/64
R3	10.1.3.3	GE0/0/1	2001:DB8:2345:23::3/64
R4	10.1.4.4	GE0/0/2	2001:DB8:2345:23::4/64

[R2]display ospfv3 lsdb

* indicates STALE LSA

OSPFv3 Router with ID (10.1.2.2) (Process 1)

Type8 LSA

Link-LSA (Interface GigabitEthernet0/0/0)

Link State ID	Origin Router	Age	Seq#	CkSum	Prefix
0.0.0.3	10.1.2.2	0080	0x80000002	0x1c52	1
0.0.0.4	10.1.3.3	0075	0x80000002	0x2a59	1
0.0.0.5	10.1.4.4	0089	0x80000002	0xe2a3	1

Link-LSA (Interface GigabitEthernet0/0/1)

Link State ID	Origin Router	Age	Seq#	CkSum	Prefix
0.0.0.4	10.1.1.1	0070	0x80000002	0xe971	1
0.0.0.4	10.1.2.2	0069	0x80000002	0x9fdd	1

Router-LSA (Area 0.0.0.0)

.....

Network-LSA (Area 0.0.0.0)

Link State ID	Origin Router	Age	Seq#	CkSum
0.0.0.4	10.1.2.2	0030	0x80000001	0x9564
0.0.0.5	10.1.4.4	0036	0x80000002	0xc810

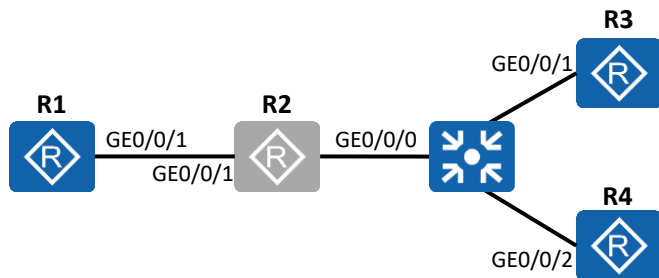
Type9 LSA

Intra-Area-Prefix-LSA (Area 0.0.0.0)

Link State ID	Origin Router	Age	Seq#	CkSum	Prefix	Reference
0.0.0.1	10.1.2.2	0028	0x80000006	0x1527	1	Network-LSA
0.0.0.1	10.1.4.4	0034	0x80000003	0xfc28	1	Network-LSA



Link-LSA举例



设备	Router ID	接口	接口地址
R1	10.1.1.1	GE0/0/1	2001:DB8:2345:12::1/64
R2	10.1.2.2	GE0/0/1	2001:DB8:2345:12::2/64
		GE0/0/0	2001:DB8:2345:23::2/64
R3	10.1.3.3	GE0/0/1	2001:DB8:2345:23::3/64
R4	10.1.4.4	GE0/0/2	2001:DB8:2345:23::4/64

查看R2的Link State ID为0.0.0.5的Link-LSA，输出信息如下所示。

```
[R2]display ospfv3 lsdb link 0.0.0.5
```

OSPFv3 Router with ID (10.1.2.2) (Process 1)

Link-LSA (Interface GigabitEthernet0/0/0)

LS Age: 446
LS Type: Link-LSA
Link State ID: 0.0.0.5
Originating Router: 10.1.4.4
LS Seq Number: 0x80000002
Retransmit Count: 0
Checksum: 0xE2A3
Length: 56

LSA头部

Priority: 1
Options: 0x000013 (-|R|-|E|V6)
Link-Local Address: FE80::2E0:FCFF:FEAC:6C5
Number of Prefixes: 1

Type8 LSA

链路本地地址

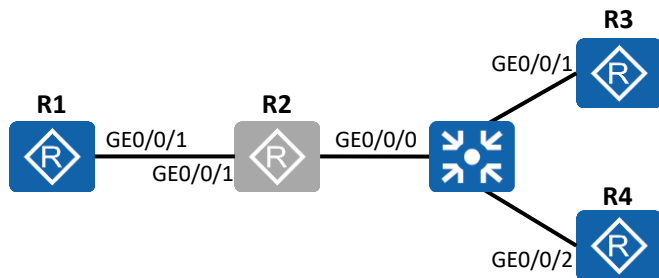
Prefix: 2001:DB8:2345:23::/64

Prefix Options: 0 (-|-|-|-|-)

直连网段的IPv6地址前缀



Intra-Area-Prefix-LSA举例



设备	Router ID	接口	接口地址
R1	10.1.1.1	GE0/0/1	2001:DB8:2345:12::1/64
R2	10.1.2.2	GE0/0/1	2001:DB8:2345:12::2/64
		GE0/0/0	2001:DB8:2345:23::2/64
R3	10.1.3.3	GE0/0/1	2001:DB8:2345:23::3/64
R4	10.1.4.4	GE0/0/2	2001:DB8:2345:23::4/64

查看R2的Intra-Area-Prefix-LSA，输出信息如下所示。

```
[R2]display ospfv3 lsdb intra-prefix
```

OSPFv3 Router with ID (10.1.2.2) (Process 1)

Intra-Area-Prefix-LSA (Area 0.0.0.0)

.....

LS Age: 556

LS Type: Intra-Area-Prefix-LSA

Link State ID: 0.0.0.1

Originating Router: 10.1.4.4

LS Seq Number: 0x80000003

Retransmit Count: 0

Checksum: 0xFC28

Length: 44

LSA头部

Number of Prefixes: 1

Referenced LS Type: 0x2002

Referenced Link State ID: 0.0.0.5

Referenced Originating Router: 10.1.4.4

Type9 LSA

参考由DR (R4) 产生的
Network-LSA

Prefix: 2001:DB8:2345:23::/64

Prefix Options: 0 (-|-|-|-|-)

Metric: 0

由DR产生，路由器直接相连的
传输网段的IPv6地址前缀



OSPFv3的基础配置命令（1）

1. 启动OSPFv3

```
[Huawei] ospfv3 [ process-id ] [ vpn-instance vpn-instance-name ]
```

创建并运行OSPFv3进程，并将创建的OSPFv3进程与VPN实例进行绑定（可选）。

```
[Huawei-ospfv3-1] router-id router-id
```

配置设备在该OSPFv3进程中所使用的Router ID。

注意：如果用户没有指定Router ID，则OSPFv3进程无法运行。

2. 在接口上使能OSPFv3

```
[Huawei-GigabitEthernet0/0/1] ospfv3 process-id area area-id [ instance instance-id ]
```

在接口上使能OSPFv3的进程，并指定所属区域，也可以指定接口所属的实例ID。

注意：配置此命令前，必须先创建OSPFv3进程和使能IPv6功能。



OSPFv3的基础配置命令 (2)

3. (可选) 配置接口的OSPFv3网络类型

```
[Huawei-GigabitEthernet0/0/1] ospfv3 network-type { broadcast | nbma | p2mp [ non-broadcast ] | p2p }  
[ instance instance-id ]
```

缺省情况下，接口的OSPFv3网络类型根据物理接口的数据链路层封装而定。以太网接口的缺省网络类型为Broadcast，串口（封装PPP协议或HDLC协议时）的缺省网络类型为P2P。

4. 进入OSPFv3区域视图

```
[Huawei-ospfv3-1] area area-id
```

区域ID可以采用十进制整数或IPv4地址形式输入，但显示时是IPv4地址形式。



检查OSPFv3基本功能的配置结果

1. 查看OSPFv3的接口信息

```
[Huawei] display ospfv3 [ process-id ] interface [ area area-id ] [ interface-type interface-number ]
```

2. 查看OSPFv3的邻居信息

```
[Huawei] display ospfv3 [ process-id ] [ area area-id ] peer [ interface-type interface-number | neighbor-id ] [ verbose ]
```

neighbor-id: 指定邻居的Router ID号。

3. 查看OSPFv3的LSDB信息

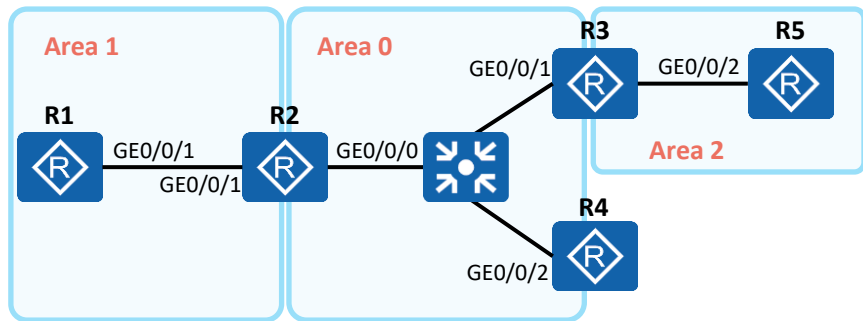
```
[Huawei] display ospfv3 [ process-id ] lsdb [ area area-id ] [ originate-router advertising-router-id | self-originate ]  
[ { router | network | inter-router [ asbr-router asbr-router-id ] | { inter-prefix | nssa } [ ipv6-address prefix-length ]  
| link | intra-prefix } [ link-state-id ] ]
```

4. 查看OSPFv3的路由表信息

```
[Huawei] display ospfv3 [ process-id ] routing [ ipv6-address prefix-length | abr-routes | asbr-routes | intra-routes | inter-routes | ase-routes | nssa-routes | [ statistics ] ]
```



OSPF双栈配置举例



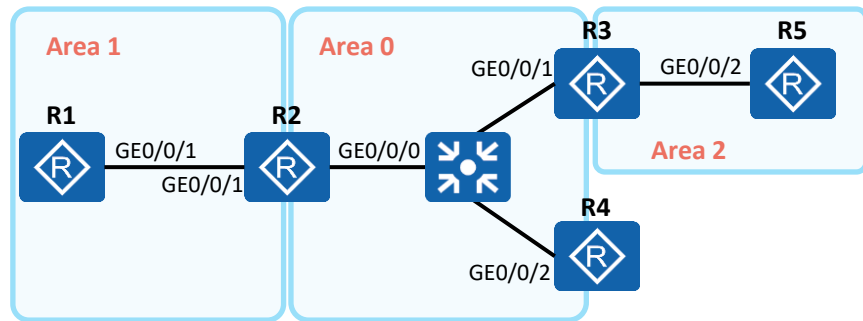
设备	Router ID	接口	IPv4地址	IPv6地址
R1	10.1.1.1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
R2	10.1.2.2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
		GE0/0/0	10.1.23.2/24	2001:DB8:2345:23::2/64
R3	10.1.3.3	GE0/0/1	10.1.23.3/24	2001:DB8:2345:23::3/64
		GE0/0/2	10.1.35.3/24	2001:DB8:2345:35::3/64
R4	10.1.4.4	GE0/0/2	10.1.23.4/24	2001:DB8:2345:23::4/64
R5	10.1.5.5	GE0/0/2	10.1.35.5/24	2001:DB8:2345:35::5/64

场景描述:

- 某公司通过部署OSPFv2实现IPv4网络的互联互通。该公司为了保证未来的业务发展，同时部署了IPv6网络进行业务测试，在该网络中运行OSPFv3实现了IPv6网络的互联互通。
- 所有路由器运行OSPFv2和OSPFv3协议，整个自治系统分为3个区域。配置完成后，每台路由器都应学到AS内的所有网段的IPv4路由和IPv6路由。



部署IPv4网络 (1)



设备	Router ID	接口	IPv4地址	IPv6地址
R1	10.1.1.1	GEO/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
R2	10.1.2.2	GEO/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
		GEO/0/0	10.1.23.2/24	2001:DB8:2345:23::2/64
R3	10.1.3.3	GEO/0/1	10.1.23.3/24	2001:DB8:2345:23::3/64
		GEO/0/2	10.1.35.3/24	2001:DB8:2345:35::3/64
R4	10.1.4.4	GEO/0/2	10.1.23.4/24	2001:DB8:2345:23::4/64
R5	10.1.5.5	GEO/0/2	10.1.35.5/24	2001:DB8:2345:35::5/64

1、配置各路由器接口的IPv4地址。（略）

2、配置OSPF基本功能。

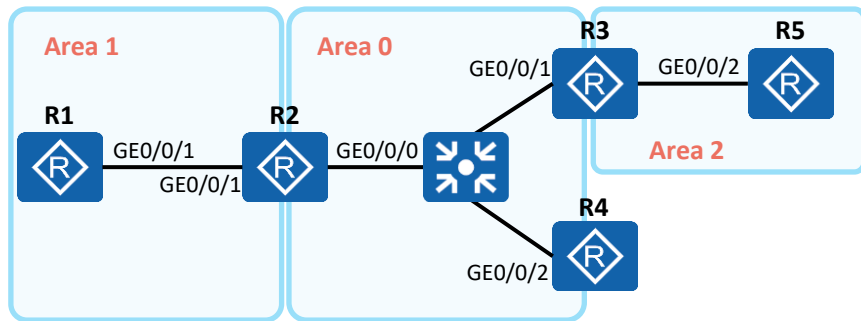
```
[R1] ospf 1 router-id 10.1.1.1
[R1-ospf-1] area 1
[R1-ospf-1-area-0.0.0.0] network 10.1.12.1 0.0.0.0
```

```
[R2] ospf 1 router-id 10.1.2.2
[R2-ospf-1] area 1
[R2-ospf-1-area-0.0.0.1] network 10.1.12.2 0.0.0.0
[R2-ospf-1-area-0.0.0.1] quit
[R2-ospf-1] area 0
[R2-ospf-1-area-0.0.0.0] network 10.1.23.2 0.0.0.0
```

```
[R3] ospf 1 router-id 10.1.3.3
[R3-ospf-1] area 0
[R3-ospf-1-area-0.0.0.0] network 10.1.23.3 0.0.0.0
[R3-ospf-1-area-0.0.0.0] quit
[R3-ospf-1] area 2
[R3-ospf-1-area-0.0.0.2] network 10.1.35.3 0.0.0.0
```



部署IPv4网络 (2)



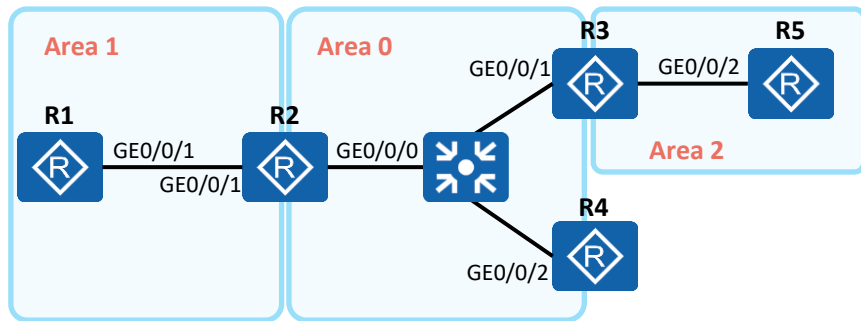
[R4] ospf 1 router-id 10.1.4.4
[R4-ospf-1] area 0
[R4-ospf-1-area-0.0.0.0] network 10.1.23.4 0.0.0.0

[R5] ospf 1 router-id 10.1.5.5
[R5-ospf-1] area 2
[R5-ospf-1-area-0.0.0.2] network 10.1.35.5 0.0.0.0

设备	Router ID	接口	IPv4地址	IPv6地址
R1	10.1.1.1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
R2	10.1.2.2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
		GE0/0/0	10.1.23.2/24	2001:DB8:2345:23::2/64
R3	10.1.3.3	GE0/0/1	10.1.23.3/24	2001:DB8:2345:23::3/64
		GE0/0/2	10.1.35.3/24	2001:DB8:2345:35::3/64
R4	10.1.4.4	GE0/0/2	10.1.23.4/24	2001:DB8:2345:23::4/64
R5	10.1.5.5	GE0/0/2	10.1.35.5/24	2001:DB8:2345:35::5/64



部署IPv6网络 (1)



设备	Router ID	接口	IPv4地址	IPv6地址
R1	10.1.1.1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
R2	10.1.2.2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
		GE0/0/0	10.1.23.2/24	2001:DB8:2345:23::2/64
R3	10.1.3.3	GE0/0/1	10.1.23.3/24	2001:DB8:2345:23::3/64
		GE0/0/2	10.1.35.3/24	2001:DB8:2345:35::3/64
R4	10.1.4.4	GE0/0/2	10.1.23.4/24	2001:DB8:2345:23::4/64
R5	10.1.5.5	GE0/0/2	10.1.35.5/24	2001:DB8:2345:35::5/64

1、全局及接口下使能IPv6功能，配置各路由器接口的IPv6地址。（略）

2、启动OSPFv3功能。

```
[R1] ospfv3 1
[R1-ospfv3-1] router-id 10.1.1.1
```

```
[R2] ospfv3 1
[R2-ospfv3-1] router-id 10.1.2.2
```

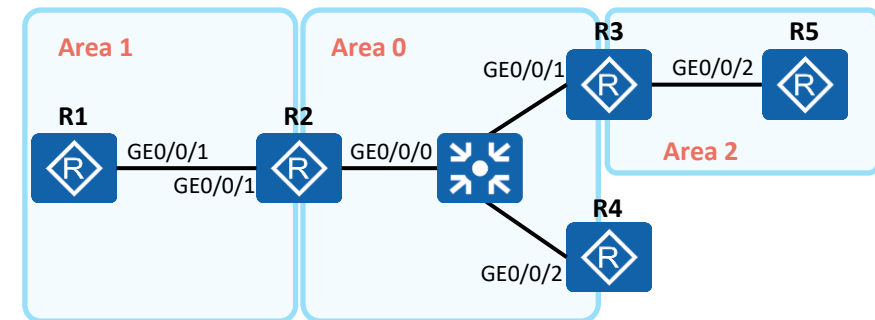
```
[R3] ospfv3 1
[R3-ospfv3-1] router-id 10.1.3.3
```

```
[R4] ospfv3 1
[R4-ospfv3-1] router-id 10.1.4.4
```

```
[R5] ospfv3 1
[R5-ospfv3-1] router-id 10.1.5.5
```



部署IPv6网络 (2)



设备	Router ID	接口	IPv4地址	IPv6地址
R1	10.1.1.1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
R2	10.1.2.2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
		GE0/0/0	10.1.23.2/24	2001:DB8:2345:23::2/64
R3	10.1.3.3	GE0/0/1	10.1.23.3/24	2001:DB8:2345:23::3/64
		GE0/0/2	10.1.35.3/24	2001:DB8:2345:35::3/64
R4	10.1.4.4	GE0/0/2	10.1.23.4/24	2001:DB8:2345:23::4/64
R5	10.1.5.5	GE0/0/2	10.1.35.5/24	2001:DB8:2345:35::5/64

3、接口使能OSPFv3功能。

```
[R1] interface gigabitethernet 0/0/1  
[R1-GigabitEthernet0/0/1] ospfv3 1 area 1
```

```
[R2] interface gigabitethernet 0/0/1  
[R2-GigabitEthernet0/0/1] ospfv3 1 area 1  
[R2-GigabitEthernet0/0/1] quit  
[R2] interface gigabitethernet 0/0/0  
[R2-GigabitEthernet0/0/0] ospfv3 1 area 0
```

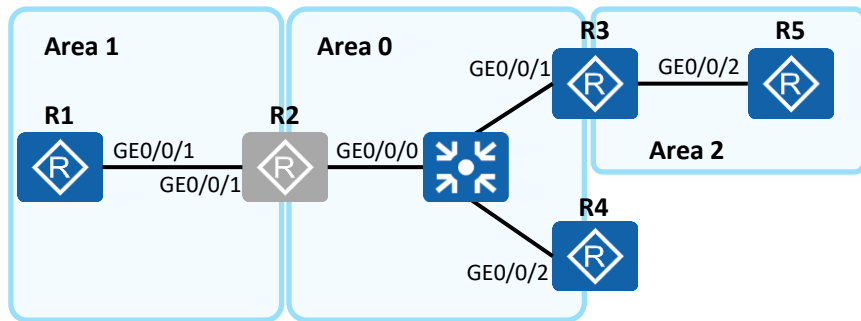
```
[R3] interface gigabitethernet 0/0/1  
[R3-GigabitEthernet0/0/1] ospfv3 1 area 0  
[R3-GigabitEthernet0/0/1] quit  
[R3] interface gigabitethernet 0/0/2  
[R3-GigabitEthernet0/0/2] ospfv3 1 area 2
```

```
[R4] interface gigabitethernet 0/0/2  
[R4-GigabitEthernet0/0/2] ospfv3 1 area 0
```

```
[R5] interface gigabitethernet 0/0/2  
[R5-GigabitEthernet0/0/2] ospfv3 1 area 2
```



查看OSPFv3网络的邻居信息



[R2]display ospfv3 peer

OSPFv3 Process (1)

OSPFv3 Area (0.0.0.0)

Neighbor ID	Pri	State	Dead Time	Interface	Instance ID
10.1.3.3	1	Full/Backup	00:00:31	GE0/0/0	0
10.1.4.4	1	Full/DR	00:00:36	GE0/0/0	0

OSPFv3 Area (0.0.0.1)

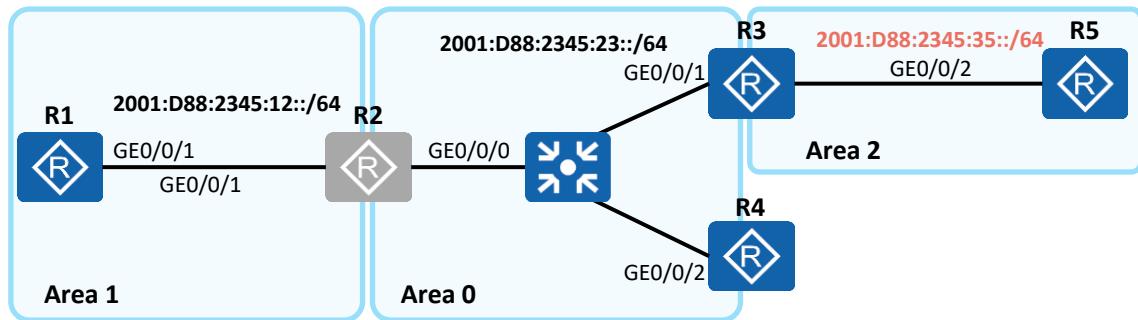
Neighbor ID	Pri	State	Dead Time	Interface	Instance ID
10.1.1.1	1	Full/Backup	00:00:37	GE0/0/1	0

当不指定实例时，默认值为0。

两台设备的直连接口的实例ID需要相同，
否则无法建立OSPFv3邻接关系。



查看OSPFv3网络的路由信息



[R2]display ospfv3 routing

Codes : E2 - Type 2 External, E1 - Type 1 External, IA - Inter-Area, N - NSSA, U - Uninstalled

OSPFv3 Process (1)

Destination	Metric	Next-hop
2001:DB8:2345:12::/64	1	directly connected, GigabitEthernet0/0/1
2001:DB8:2345:23::/64	1	directly connected, GigabitEthernet0/0/0
IA 2001:DB8:2345:35::/64	2	via FE80::2E0:FCFF:FE38:50A0, GigabitEthernet0/0/0

在OSPFv3中，路由的下一跳地址为链路本地地址。
本条路由的下一跳地址为R3的GE0/0/1接口的链路本地地址。



查看OSPFv3网络LSDB信息

[R2]display ospfv3 lsdb

* indicates STALE LSA

OSPFv3 Router with ID (10.1.2.2) (Process 1)

Link-LSA (Interface GigabitEthernet0/0/0)

Link State ID	Origin Router	Age	Seq#	CkSum	Prefix
0.0.0.3	10.1.2.2	0862	0x80000003	0x486d	1
0.0.0.4	10.1.3.3	0854	0x80000003	0xc512	1
0.0.0.5	10.1.4.4	0840	0x80000003	0x4e3c	1

Link-LSA (Interface GigabitEthernet0/0/1)

Link State ID	Origin Router	Age	Seq#	CkSum	Prefix
0.0.0.4	10.1.1.1	0878	0x80000003	0x9d8f	1
0.0.0.4	10.1.2.2	0868	0x80000003	0xcbf8	1

Router-LSA (Area 0.0.0.0)

Link State ID	Origin Router	Age	Seq#	CkSum	Link
0.0.0.0	10.1.2.2	0567	0x8000000f	0xd411	1
0.0.0.0	10.1.3.3	0576	0x8000000e	0xd70c	1
0.0.0.0	10.1.4.4	0577	0x8000000a	0xdd08	1

Network-LSA (Area 0.0.0.0)

Link State ID	Origin Router	Age	Seq#	CkSum
0.0.0.5	10.1.4.4	0577	0x80000006	0xc014

Inter-Area-Prefix-LSA (Area 0.0.0.0)

Link State ID	Origin Router	Age	Seq#	CkSum
0.0.0.1	10.1.2.2	0862	0x80000003	0xceb3
0.0.0.1	10.1.3.3	0782	0x80000004	0x3824

Intra-Area-Prefix-LSA (Area 0.0.0.0)

Link State ID	Origin Router	Age	Seq#	CkSum	Prefix	Reference
0.0.0.1	10.1.4.4	0576	0x8000000a	0xee2f	1	Network-LSA

Router-LSA (Area 0.0.0.1)

Link State ID	Origin Router	Age	Seq#	CkSum	Link
0.0.0.0	10.1.1.1	0819	0x80000007	0x9460	1
0.0.0.0	10.1.2.2	0827	0x80000007	0x8a67	1

Network-LSA (Area 0.0.0.1)

Link State ID	Origin Router	Age	Seq#	CkSum
0.0.0.4	10.1.2.2	0828	0x80000003	0x9166

Inter-Area-Prefix-LSA (Area 0.0.0.1)

Link State ID	Origin Router	Age	Seq#	CkSum
0.0.0.1	10.1.2.2	0570	0x80000001	0x510f
0.0.0.2	10.1.2.2	0596	0x80000001	0xfb76

Intra-Area-Prefix-LSA (Area 0.0.0.1)

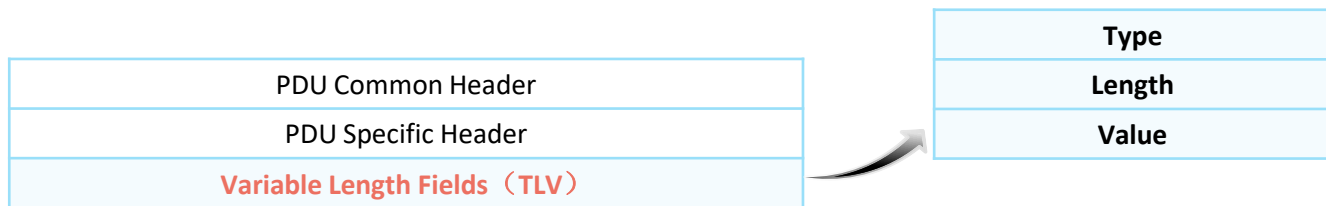
Link State ID	Origin Router	Age	Seq#	CkSum	Prefix	Reference
0.0.0.1	10.1.2.2	0826	0x80000004	0x1925	1	Network-LSA

查看R2的LSDB，有Type1、Type2、Type3、Type8和Type9五种LSA。



IS-IS (IPv6)概述

- IS-IS最初是为OSI网络设计的一种基于链路状态协议的动态路由协议。之后为了提供对IPv4的路由支持，扩展应用到IPv4网络，称为集成IS-IS。
- IS-IS报文有以下几种类型：Hello PDU（Protocol Data Unit）、LSP和SNP。
 - 报文格式为：



- IS-IS报文中的变长字段部分是多个TLV（Type-Length-Value）三元组，使用TLV结构构建报文使IS-IS更具灵活性和扩展性，增加新特性只需要增加新TLV即可。
- 为了支持IPv6路由的处理和计算，IS-IS新增了两个TLV（Type-Length-Value）和一个NLPID（Network Layer Protocol Identifier，网络层协议标识符）。



新增TLV

232号TLV (IPv6 Interface Address)

相当于132号TLV（用于描述IPv4接口地址），只不过把原来的32bit的IPv4地址改为128bit的IPv6地址。

Type=232	Length	Interface Address 1
		Interface Address 1
		Interface Address 1
		Interface Address 1
Interface Address 1		Interface Address 2
.....		

在不同的PDU中，接口地址字段的内容是不同的。

- 对于Hello报文，“接口地址”只能包含发送Hello报文接口的链路本地地址。
- 对于LSP报文，“接口地址”只能包含分配给设备接口的非链路本地地址。

236号TLV (IPv6 Reachability)

相当于128号和130号TLV，通过X-bit来标识内/外部可达性信息。

Type=236	Length	Metric				
Metric		U	X	S	R	Prefix Length
Prefix						
Sub-TLV Len	Sub-TLV					

上述IPv6 Reachability TLV在LSP中可以出现任意次数（包括0次）。其中，链路本地前缀不是用这个TLV来发布。



129号TLV中新增NLPID

- 为了支持IPv6路由的处理和计算，IS-IS在129号TLV中新增了一个NLPID。
- 129号TLV（Protocol Supported）

Type=129	Length	NLPID	NLPID
.....			

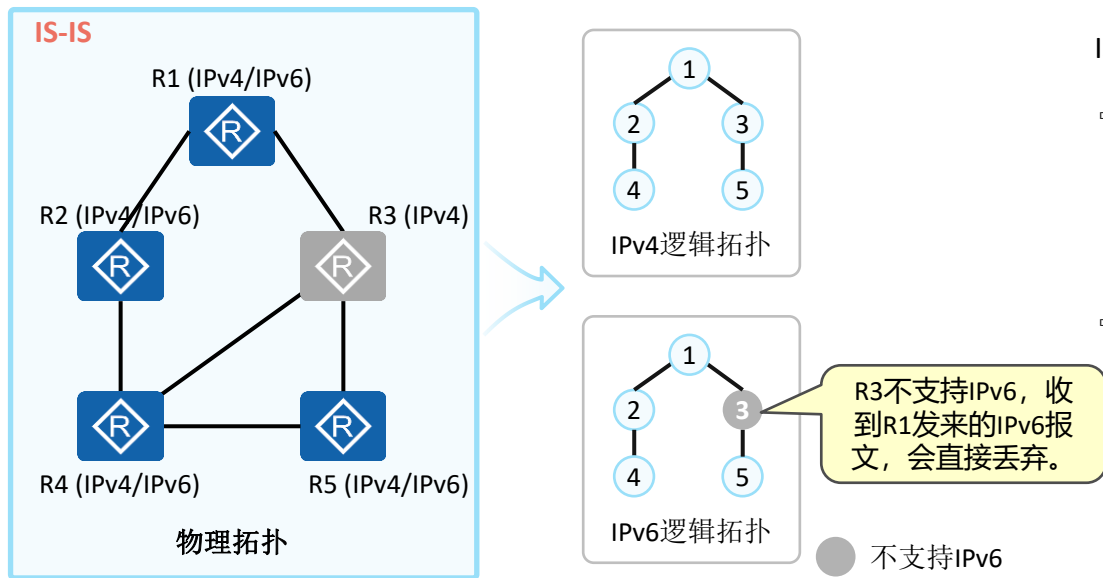
字段解释：

- Type：8bit，TLV类型，此时值为129（0x81），表示支持协议TLV。
 - Length：8bit，TLV的Value部分长度。
 - NLPID：8bit，网络层协议标识符。若支持IPv4，则值为204（0xCC）；若支持IPv6，则值为142（0x8E）
- 如果IS-IS支持IPv6，那么向外发布IPv6路由时必须携带NLPID值。



IS-IS多拓扑技术背景

缺省情况下，在运行IS-IS的网络环境中，IPv4和IPv6的混合拓扑被看成是一个集成的拓扑，IS-IS针对IPv4和IPv6经计算形成相同的最短路径树。



IS-IS单拓扑存在的问题：

- 混合拓扑中的一些路由器和链路不支持IPv6协议，但是支持双协议栈的路由器无法感知到这些路由器和链路，仍然会把IPv6报文转发给它们，这就导致IPv6报文因无法转发而被丢弃。
- 同样，存在不支持IPv4的路由器和链路时，IPv4报文也无法转发。



IS-IS多拓扑概述

- IS-IS多拓扑（Multi-Topology，MT）特性是指在一个IS-IS自治域内运行多个独立的IP拓扑。例如IPv4拓扑和IPv6拓扑，而不是将它们视为一个集成的单一拓扑。这有利于IS-IS在路由计算中根据实际组网情况来单独考虑IPv4和IPv6网络。根据链路所支持的IP协议类型，不同拓扑运行各自的SPF计算，实现网络的相互屏蔽。
- IS-IS多拓扑的实现过程
 - 建立拓扑：通过报文交互建立邻居，从而建立多拓扑。
 - SPF计算：在不同的拓扑上分别进行SPF计算。



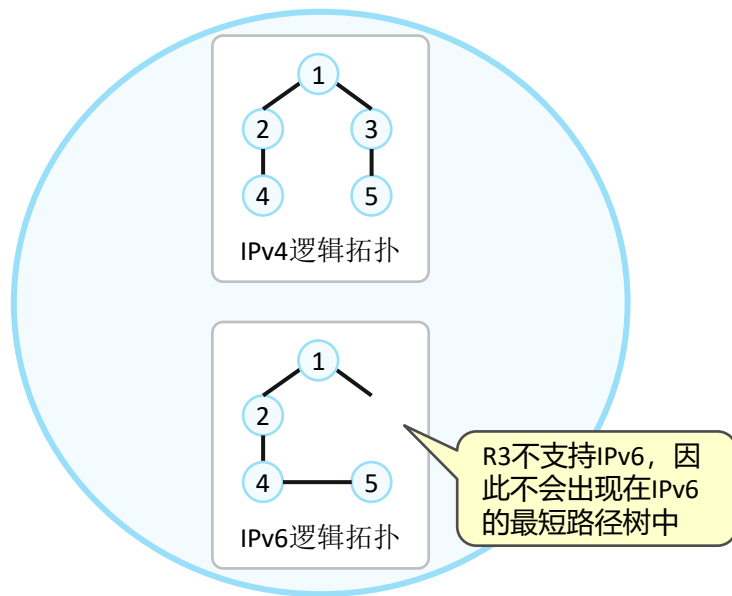
IS-IS多拓扑原理

- IS-IS定义了新的TLV，该TLV中包含接口所属拓扑信息（MT信息）。MT信息的传播，使得网络按不同的拓扑分别进行SPF计算，最终实现拓扑分离。
- 229号多拓扑TLV：

Type=229	Length	O	A	R	MT ID
.....					

- 重要字段解释：

- Type：8bit，TLV类型，此时值为229（0xE5），表示支持多拓扑。
- O：1bit，Overload，超载位。
- A：1bit，Attach，附着位。
- MT ID：12bit，表示该接口属于什么拓扑。



IPv4和IPv6各自计算自己独立的拓扑。



IS-IS (IPv6) 的基础配置命令

1. 使能IS-IS (IPv6) 功能

```
[Huawei-isis-1] ipv6 enable [ topology { ipv6 | standard } ]
```

在IS-IS进程下，使能该进程的IPv6能力。

```
[Huawei-GigabitEthernet0/0/1] isis ipv6 enable [ process-id ]
```

在接口上使能IS-IS的IPv6功能并指定要关联的IS-IS进程号。

注意：配置此命令前，必须先使能接口的IPv6功能。

2. 配置IS-IS接口在IPv6网络中的开销

```
[Huawei-GigabitEthernet0/0/1] isis ipv6 cost { cost | maximum } [ level-1 | level-2 ]
```

缺省情况下，IPv6拓扑中IS-IS接口的链路开销值为10。



检查IS-IS（IPv6）基本功能的配置结果

1. 查看使能了IS-IS（IPv6）的接口信息

```
[Huawei] display isis interface interface-type interface-number [ verbose ]
```

2. 查看IS-IS（IPv6）的邻居信息

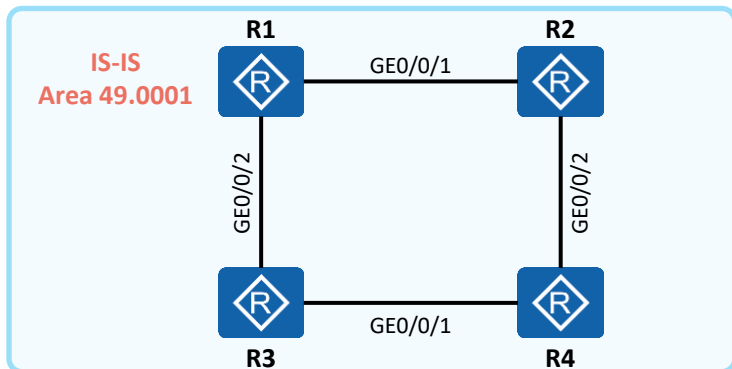
```
[Huawei] display isis process-id peer [ verbose ]
```

3. 查看IS-IS（IPv6）的路由信息

```
[Huawei] display isis route [ process-id | vpn-instance vpn-instance-name ] ipv6 [ verbose | [ level-1 | level-2 ] ] ipv6-address [ prefix-length ] ]
```



IS-IS双栈配置举例



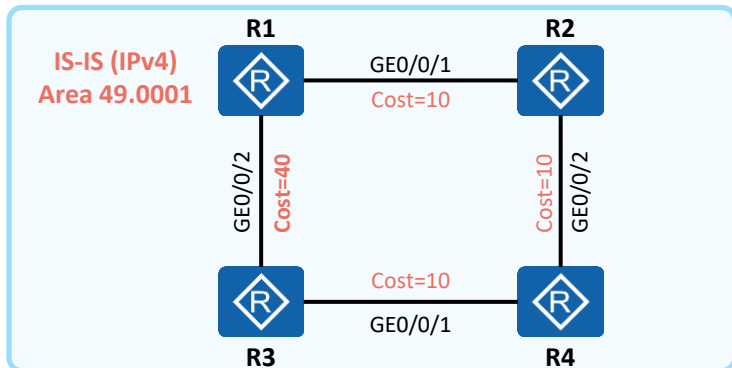
场景描述:

- 某公司通过部署IS-IS实现IPv4网络的互联互通。该公司为了保证未来的业务发展，同时部署了IPv6网络进行业务测试，因此需要在该网络的IS-IS中使能IPv6功能。
- 由于设备性能不同，要求通过修改开销值，让IPv4和IPv6业务优选不同的路径，且IPv6网络单独计算拓扑。
- 所有路由器运行IS-IS协议，整个网络都处于区域49.0001中，且所有路由器均为Level-2路由器。配置完成后，每台路由器都应学到AS内的所有网段的IPv4路由和IPv6路由。

设备	接口	IPv4地址	IPv6地址
R1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
	GE0/0/2	10.1.13.1/24	2001:DB8:2345:13::1/64
R2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
	GE0/0/2	10.1.24.2/24	2001:DB8:2345:24::2/64
R3	GE0/0/1	10.1.34.3/24	2001:DB8:2345:34::3/64
	GE0/0/2	10.1.13.3/24	2001:DB8:2345:13::3/64
R4	GE0/0/1	10.1.34.4/24	2001:DB8:2345:34::4/64
	GE0/0/2	10.1.24.4/24	2001:DB8:2345:24::4/64



部署IPv4网络



设备	接口	IPv4地址	IPv6地址
R1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
	GE0/0/2	10.1.13.1/24	2001:DB8:2345:13::1/64
R2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
	GE0/0/2	10.1.24.2/24	2001:DB8:2345:24::2/64
R3	GE0/0/1	10.1.34.3/24	2001:DB8:2345:34::3/64
	GE0/0/2	10.1.13.3/24	2001:DB8:2345:13::3/64
R4	GE0/0/1	10.1.34.4/24	2001:DB8:2345:34::4/64
	GE0/0/2	10.1.24.4/24	2001:DB8:2345:24::4/64

1、配置各路由器接口的IPv4地址。（略）

2、配置IS-IS基本功能。

```
[R1] isis 1
[R1-isis-1] is-level level-2
[R1-isis-1] network-entity 49.0001.0000.0000.0001.00
#
[R1] interface gigabitethernet 0/0/1
[R1-GigabitEthernet0/0/1] isis enable 1
[R1-GigabitEthernet0/0/1] quit
[R1] interface gigabitethernet 0/0/2
[R1-GigabitEthernet0/0/2] isis enable 1
[R1-GigabitEthernet0/0/2] isis cost 40 level-2
```

R2、R3和R4的配置与R1类似，不再赘述。

其中，它们的网络实体名分别为：

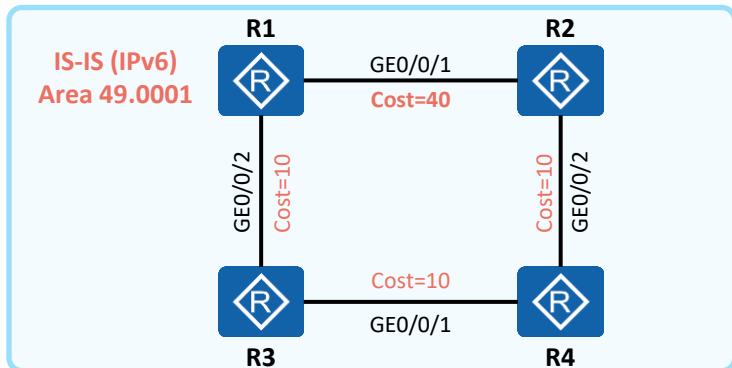
R2: 49.0001.0000.0000.0002.00

R3: 49.0001.0000.0000.0003.00

R4: 49.0001.0000.0000.0004.00



部署IPv6网络



设备	接口	IPv4地址	IPv6地址
R1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
	GE0/0/2	10.1.13.1/24	2001:DB8:2345:13::1/64
R2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
	GE0/0/2	10.1.24.2/24	2001:DB8:2345:24::2/64
R3	GE0/0/1	10.1.34.3/24	2001:DB8:2345:34::3/64
	GE0/0/2	10.1.13.3/24	2001:DB8:2345:13::3/64
R4	GE0/0/1	10.1.34.4/24	2001:DB8:2345:34::4/64
	GE0/0/2	10.1.24.4/24	2001:DB8:2345:24::4/64

1、配置各路由器接口的IPv6地址。（略）

2、使能IS-IS（IPv6）功能。

```
[R1] isis 1
[R1-isis-1] ipv6 enable topology ipv6
#
[R1] interface gigabitethernet 0/0/1
[R1-GigabitEthernet0/0/1] isis ipv6 enable 1
[R1-GigabitEthernet0/0/1] isis ipv6 cost 40 level-2
[R1-GigabitEthernet0/0/1] quit
[R1] interface gigabitethernet 0/0/2
[R1-GigabitEthernet0/0/2] isis ipv6 enable 1
```

R2、R3和R4的配置与R1类似，不再赘述。



查看IS-IS (IPv4)网络路由信息

以R1为根计算最短路径树为例，可以发现IS-IS（IPv4）的逻辑拓扑如右下图所示。

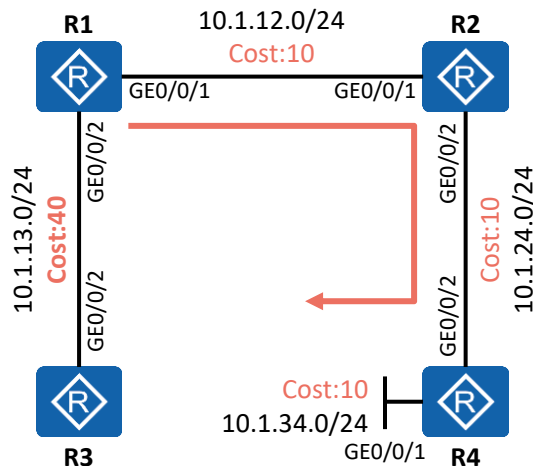
```
[R1]dis ip routing-table protocol isis
Route Flags: R - relay, D - download to fib
-----
Public routing table : ISIS
Destinations : 2    Routes : 2

ISIS routing table status : <Active>
Destinations : 2    Routes : 2

Destination/Mask  Proto  Pre  Cost  Flags NextHop  Interface
10.1.24.0/24     ISIS-L2 15   20    D   10.1.12.2  GigabitEthernet0/0/1
10.1.34.0/24     ISIS-L2 15   30    D   10.1.12.2  GigabitEthernet0/0/1

ISIS routing table status : <Inactive>
Destinations : 0    Routes : 0
```

IS-IS (IPv4) 逻辑拓扑



→ R1访问10.1.34.0/24网段的流量访问路径



查看IS-IS (IPv6)网络路由信息

以R1为根计算最短路径树为例，可以发现IS-IS（IPv6）的逻辑拓扑如右下图所示。

```
[R1]dis ipv6 routing-table protocol isis
```

```
Public Routing Table : ISIS
```

```
Summary Count : 2
```

```
ISIS Routing Table's Status : < Active >
```

```
Summary Count : 2
```

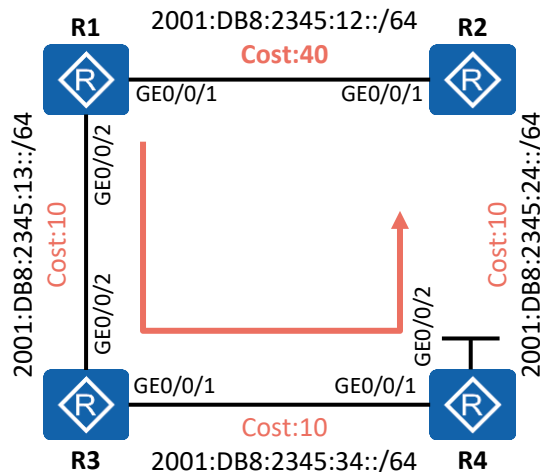
```
Destination : 2001:DB8:2345:24:: PrefixLength : 64
NextHop : FE80::2E0:FCFF:FE90:3D3A Preference : 15
Cost : 30 Protocol : ISIS-L2
RelayNextHop : :: TunnelID : 0x0
Interface : GigabitEthernet0/0/2 Flags : D
```

```
Destination : 2001:DB8:2345:34:: PrefixLength : 64
NextHop : FE80::2E0:FCFF:FE90:3D3A Preference : 15
Cost : 20 Protocol : ISIS-L2
RelayNextHop : :: TunnelID : 0x0
Interface : GigabitEthernet0/0/2 Flags : D
```

```
ISIS Routing Table's Status : < Inactive >
```

```
Summary Count : 0
```

IS-IS (IPv6) 逻辑拓扑



→ R1访问2001:DB8:2345:24::/64网段的流量访问路径



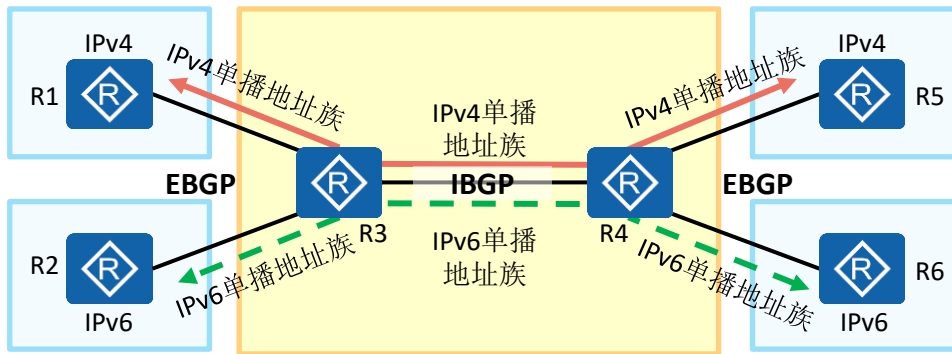
BGP4+概述

- 传统的BGP-4只能管理IPv4单播路由信息，BGP多协议扩展（MultiProtocol BGP，MP-BGP）提供对多种网络层协议的支持。目前的MP-BGP，使用扩展属性和地址族来实现对IPv6、组播和VPN相关内容的支持，BGP协议原有的报文机制和路由机制并没有改变。
- 其中，MP-BGP对IPv6单播网络的支持特性称为BGP4+。BGP4+为IPv6单播网络建立独立的拓扑结构，并将路由信息储存在独立的路由表中，保持单播IPv4网络和单播IPv6网络之间路由信息相互隔离。



MP-BGP支持的地址族

- MP-BGP采用地址族来区分不同的网络层协议，要在BGP对等体之间交互不同类型的路由信息，则需要在正确的地址族视图下激活对等体，以及发布BGP路由。
- MP-BGP支持的地址族有：
 - IPv4单播地址族
 - IPv4组播地址族
 - IPv6单播地址族
 - VPNv4地址族
 - VPNv6地址族
 -

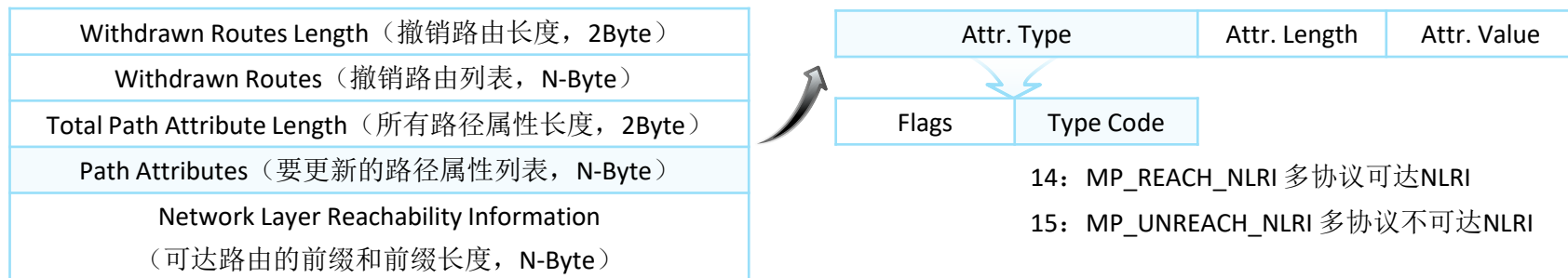


IPv4网络的邻居需要在IPv4单播地址族中使能，而IPv6网络邻居则需要要在IPv6单播地址族中使能。



BGP路径属性

- BGP的Update报文在对等体之间传递路由信息，可以用于发布和撤销路由。
- Update报文格式：



- BGP4+中引入了两个NLRI属性，分别是：
 - MP_REACH_NLRI: Multiprotocol Reachable NLRI，多协议可达NLRI。用于发布可达路由及下一跳信息。
 - MP_UNREACH_NLRI: Multiprotocol Unreachable NLRI，多协议不可达NLRI。用于撤销不可达路由。



NLRI属性

- MP_REACH_NLRI格式:

Address Family Identifier
Subsequent Address Family Identifier
Length of Next Hop Network Address
Network Address of Next Hop
Reserved
Network Layer Reachability Information



- 当传递IPv6路由时
 - AFI=2, SAFI=1 (单播), SAFI=2 (组播)。
 - 下一跳地址长度字段决定了下一跳地址的个数。
 - 长度字段=16, 下一跳地址为下一跳路由器的全球单播地址。
 - 长度字段=32, 下一跳地址为下一跳路由器的全球单播地址和链路本地地址。
 - 保留字段, 恒等于0。
 - NLRI字段, 可变长字段, 表示路由前缀和掩码信息。

- MP_UNREACH_NLRI格式:

Address Family Identifier
Subsequent Address Family Identifier
Withdrawn Routes



- 当撤销IPv6路由时
 - AFI=2, SAFI=1 (单播), SAFI=2 (组播)。
 - Withdrawn Routes 字段代表需要撤回的路由前缀及掩码。



BGP4+的基础配置命令

1. 配置BGP对等体

```
[Huawei-bgp] peer ipv6-address as-number { as-number-plain | as-number-dot }
```

在BGP视图下，创建IPv6对等体。

2. 使能BGP对等体

```
[Huawei-bgp] ipv6-family [ unicast | vpnv6 | vpn-instance vpn-instance-name ]
```

使能BGP的IPv6地址族并进入BGP的各IPv6地址族视图。

```
[Huawei-bgp-af-ipv6] peer ipv6-address enable
```

在IPv6地址族视图下使能与指定对等体之间交换相关的路由信息。

3. 配置BGP路由注入

```
[Huawei-bgp-af-ipv6] network ipv6-address prefix-length [ route-policy route-policy-name ]
```

配置BGP将IPv6路由表中的特定路由注入到BGP路由表中。



检查BGP4+基本功能的配置结果

1. 查看BGP4+的对等体信息

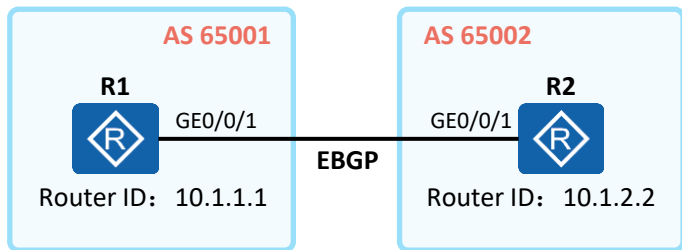
```
[Huawei] display bgp ipv6 peer ipv6-address [ verbose ]
```

2. 查看BGP4+的路由信息

```
[Huawei] display bgp ipv6 routing-table
```




BGP双栈配置举例



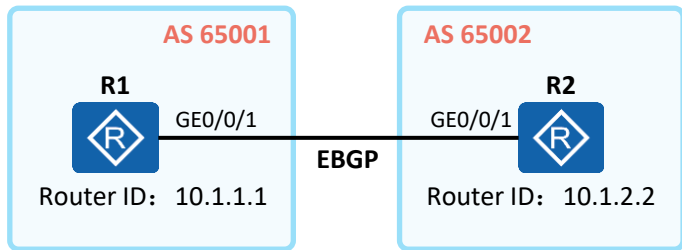
场景描述:

- 某公司两个分部之间通过BGP实现网络互通。该公司为了保证未来的业务发展，在两个分部都部署了IPv6网络进行业务测试，因此还需要部署BGP4+实现IPv6网络互通。

设备	接口	IPv4地址	IPv6地址
R1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
	Loopback0	10.1.1.1/32	2001:DB8:2345:1::1/128
R2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
	Loopback0	10.1.2.2/32	2001:DB8:2345:2::2/128



部署IPv4网络



设备	接口	IPv4地址	IPv6地址
R1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
	Loopback0	10.1.1.1/32	2001:DB8:2345:1::1/128
R2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
	Loopback0	10.1.2.2/32	2001:DB8:2345:2::2/128

1、配置各路由器接口的IPv4地址。（略）

2、配置BGP基本功能：建立EBGP对等体。

```
[R1] bgp 65001
[R1-bgp] router-id 10.1.1.1
[R1-bgp] peer 10.1.12.2 as-number 65002
```

```
[R2] bgp 65002
[R2-bgp] router-id 10.1.2.2
[R2-bgp] peer 10.1.12.1 as-number 65001
```

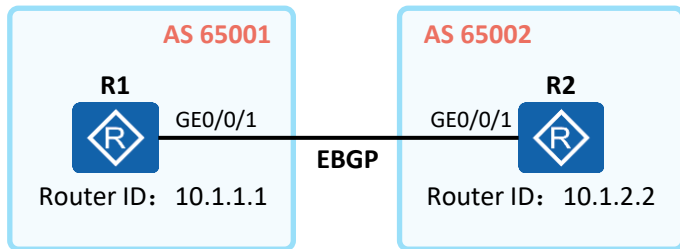
3、发布IPv4路由。

```
[R1-bgp] ipv4-family unicast
[R1-bgp-af-ipv4] network 10.1.1.1 32
```

```
[R2-bgp] ipv4-family unicast
[R2-bgp-af-ipv4] network 10.1.2.2 32
```



部署IPv6网络



设备	接口	IPv4地址	IPv6地址
R1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
	Loopback0	10.1.1.1/32	2001:DB8:2345:1::1/128
R2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
	Loopback0	10.1.2.2/32	2001:DB8:2345:2::2/128

- 1、配置各路由器接口的IPv6地址。（略）
- 2、配置BGP4+基本功能：建立EBGP对等体，并使能对等体。

```
[R1] bgp 65001
[R1-bgp] peer 2001:DB8:2345:12::2 as-number 65002
[R1-bgp] ipv6-family unicast
[R1-bgp-af-ipv6] peer 2001:DB8:2345:12::2 enable
```

```
[R2] bgp 65002
[R2-bgp] peer 2001:DB8:2345:12::1 as-number 65001
[R2-bgp] ipv6-family unicast
[R2-bgp-af-ipv6] peer 2001:DB8:2345:12::1 enable
```

- 3、发布IPv6路由。

```
[R1-bgp-af-ipv6] network 2001:DB8:2345:1::1 128
```

```
[R2-bgp-af-ipv6] network 2001:DB8:2345:2::2 128
```



查看BGP对等体信息

- 通过**display bgp [ipv6] peer**，分别查看IPv4和IPv6网络中的BGP对等体信息。
- 发现除建立对等体的地址不同，其他信息基本一致。

[R1]display bgp peer

BGP

BGP local router ID : 10.1.1.1

Local AS number: 65001

Total number of peers: 1

Peers in established state : 1

Peer	PrefRcv	V	AS	MsgRcvd	MsgSent	OutQ	Up/Down	State	
10.1.12.2		4	65002	11	12	0	00:06:49	Established	1

[R1]display bgp ipv6 peer

BGP4+

BGP local router ID : 10.1.1.1

Local AS number : 65001

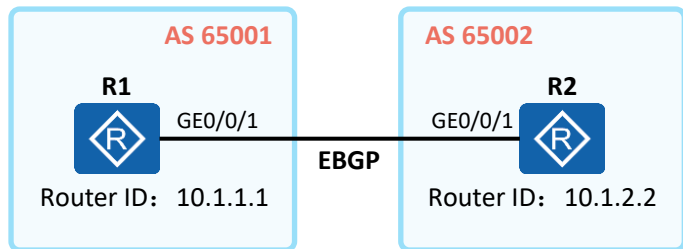
Total number of peers : 1

Peers in established state : 1

Peer	PrefRcv	V	AS	MsgRcvd	MsgSent	OutQ	Up/Down	State	
2001:DB8:2345:12::2	4	65002	14	15	0	00:07:35	Established		1



查看BGP4+路由信息



设备	接口	IPv4地址	IPv6地址
R1	GE0/0/1	10.1.12.1/24	2001:DB8:2345:12::1/64
	Loopback0	10.1.1.1/32	2001:DB8:2345:1::1/128
R2	GE0/0/1	10.1.12.2/24	2001:DB8:2345:12::2/64
	Loopback0	10.1.2.2/32	2001:DB8:2345:2::2/128

[R1]display bgp ipv6 routing-table

BGP Local router ID is 10.1.1.1

Status codes: * - valid, > - best, d - damped,
h - history, i - internal, s - suppressed, S - Stale
Origin : i - IGP, e - EGP, ? - incomplete

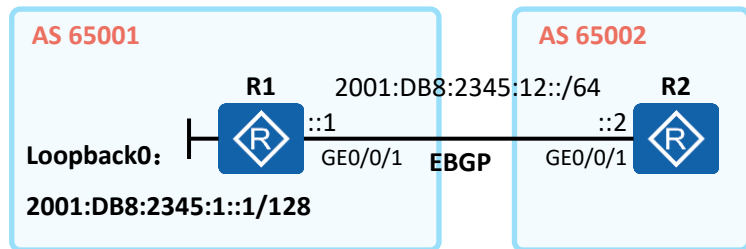
Total Number of Routes: 2

```
*> Network :      2001:DB8:2345:1::1PrefixLen :      128
    NextHop :      ::
    MED :                0
                                0
    Label :
    Path/Ogn :      i
*> Network :      2001:DB8:2345:2::2      PrefixLen :      128
    NextHop :      2001:DB8:2345:12::2    LocPrf :
    MED :                0                PrefVal :      0
    Label :
    Path/Ogn :      65002 i
```

查看BGP4+路由表，发现共有2条路由信息，均为有效路由。



查看BGP4+中的NLRI属性



通过抓包，可以在R1发出的Update报文中，查看到MP_REACH_NLRI属性信息。其中：

- 该可达路由的下一跳地址为：
2001:db8:2345:12::1
- 该可达路由地址前缀及前缀长度为：
2001:db8:2345:1::1/128

UPDATE Message-Border Gateway Protocol

UPDATE Message

Marker: 16 bytes

Length: 85 bytes

Type: UPDATE Message (2)

Unfeasible routes length: 0 bytes

Total path attribute length: 62 bytes

Path attributes

ORIGIN: IGP (4 bytes)

AS_PATH: 65001 (9 bytes)

MULTI_EXIT_DISC: 0 (7 bytes)

MP_REACH_NLRI (42 bytes)

Flags: 0x90 (Optional, Non-transitive, Complete, Extended Length)

Type code: MP_REACH_NLRI (14)

Length: 38 bytes

Address family: IPv6 (2)

Subsequent address family identifier: Unicast (1)

Next hop network address (16 bytes)

Next hop: 2001:db8:2345:12::1 (16)

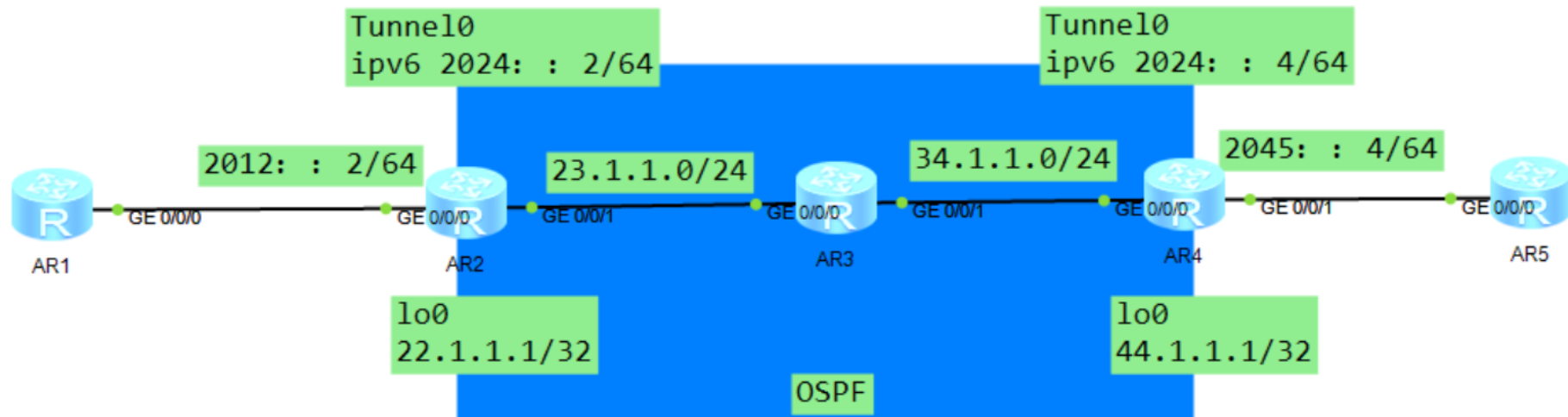
Subnetwork points of attachment: 0

Network layer reachability information (17 bytes)

2001:db8:2345:1::1/128

MP Reach NLRI prefix length: 128

MP Reach NLRI prefix: 2001:db8:2345:1::1



需求:

- 1.R1和R5分别模拟PC, 通过SLAAC获取ipv6地址
- 2.R2/R3/R4运行OSPF (ipv4)
- 3.R2和R4模拟出口路由器, 建立隧道, 实现R1与R5之间的互通

THANK YOU

Ping 通您的梦想 ~

腾讯课堂交流群：17942636

ADD：苏州市干将东路666号和基广场401-402； Tel：0512-8188 8288；

课程咨询QQ：2853771087 ； 官网 :www.51glab.com