ICMPv6 和 NDP

- 在 IPv4 中，ICMP 允许主机或设备报告差错情况。ICMP 报文作为 IP 报文的数据部分，再封装上 IP 报文首部，组成完整的 IP 报文发送出去。常用的 Ping、Tracert 等命令都是基于 ICMP 实现的。
- IPv6 定义了 ICMPv6（Internet Control Message Protocol for IPv6），除了提供类似 ICMP 的功能外，还有诸多扩展。邻居发现协议（Neighbor Discovery Protocol，以下简称 NDP）便是基于 ICMPv6 实现的，作为 IPv6 的关键协议，NDP 提供了如前缀发现、重复地址检测、地址解析、重定向等功能。
- 本课程详细介绍 ICMPv6 和 NDP。

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。

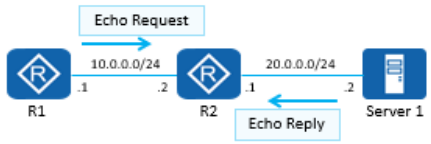
Type	Code	描述
0	0	Echo Reply
3	0	网络不可达
3	1	主机不可达
3	2	协议不可达
3	3	端口不可达
5	0	重定向
8	0	Echo Request

- 为了更有效地转发 IP 数据报文和提高数据报文交互成功的机会，在网络层使用 ICMP 协议。ICMP 允许主机或设备报告差错情况和提供有关异常情况的报告。
- ICMP 消息：
- ICMP 消息封装在 IP 报文中，IP 报文头部 Protocol 值为 1 时表示 ICMP 协议。

- 字段解释：
- ICMP 消息的格式取决于 Type 和 Code 字段，其中 Type 字段为消息类型，Code 字段包含该消息类型的具体参数。
- 校验和字段用于检查消息是否完整。

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。



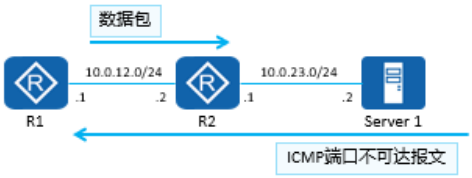
功能：Ping
Ping是网络设备、Windows、Unix和Linux平台上的一个命令，其实是一个小巧而实用的应用程序，该应用基于ICMP协议。Ping常用于探测到达目的节点的网络可达性。

```
[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 的一个典型应用是 Ping。Ping 是检测网络连通性的常用工具，同时也能够收集其他相关信息。用户可以在 Ping 命令中指定不同参数，如 ICMP 报文长度、发送的 ICMP 报文个数、等待回复响应的超时时间等，设备根据配置的参数来构造并发送 ICMP 报文，进行 Ping 测试。

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。




```
[R1]tracert 10.0.23.2

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

1 10.0.12.2              80 ms     10 ms     10 ms

2 10.0.23.2              30 ms     30 ms     20 ms
```

功能：Tracert

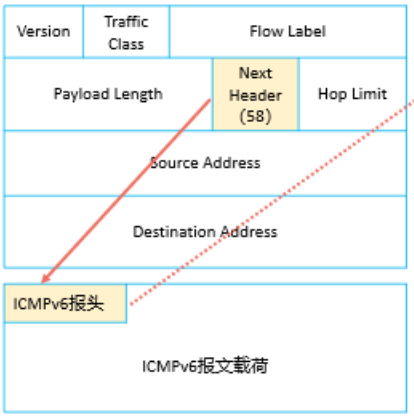
Tracert基于报文头中的TTL值来逐跳跟踪报文的转发路径。

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。

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
	4	参数错误	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 报文载荷由 ICMPv6 报文类型决定，因报文类型的不同而不同。
- **Type**：表明消息的类型。
- **Code**：表示消息类型的细分。
- **Checksum**：表示 ICMPv6 报文的校验和。

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

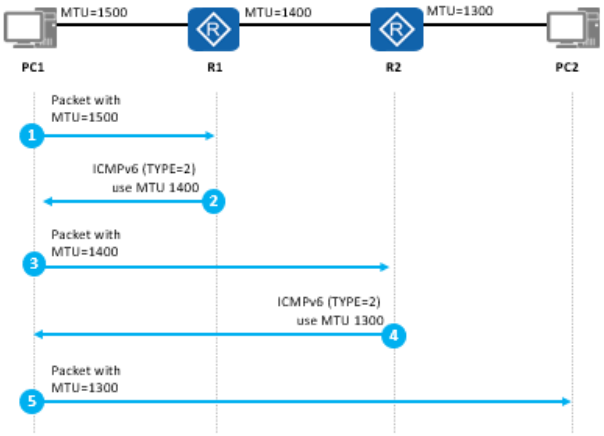
- 差错报文（Error Messages），也称为差错消息，Type字段最高bit为0，也就是ICMPv6 Type=[0，127]
- 差错消息用于报告在转发IPv6数据包过程中出现的错误，如常见的目的不可达、超时等等。

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。

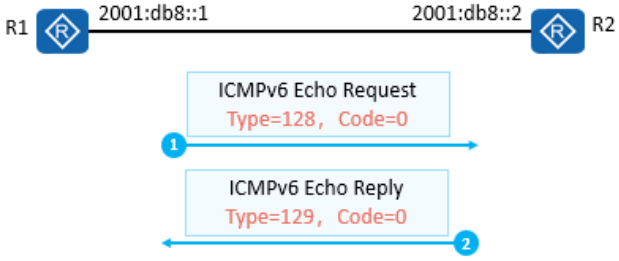


- 首先 PC1 用 1500 字节作为 MTU 向 PC2 发送 IPv6 数据包。
- R1 意识到数据包过大，出站接口 MTU 为 1400 字节，于是回复一个 ICMPv6（Type=2）报文给 PC1，指定 MTU 值为 1400 字节。
- 然后，PC1 开始使用 1400 作为 MTU 发送 IPv6 数据。
- 数据包到达 R2 后，R2 意识到出站接口 MTU 为 1300 字节，于是发送一个 ICMPv6（Type=2）报文给 PC1，指定 MTU 值为 1300 字节。
- PC1 开始使用 1300 作为 MTU 发送 IPv6 数据。

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

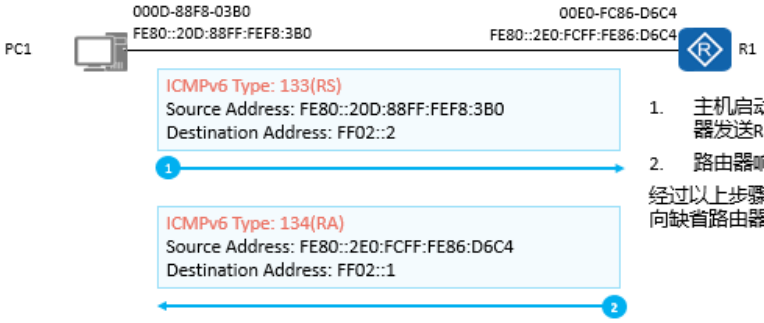
功能 \ ICMPv6报文	RS 133	RA 134	NS 135	NA 136	重定向 137
地址解析			•	•	
路由器发现	•	•			
前缀重编址	•	•			
重复地址检测			•	•	
重定向					•

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。

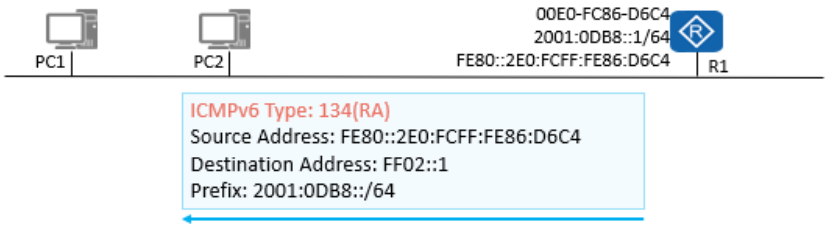


1. 主机启动，向本地链路范围内所有的路由器发送RS报文。
2. 路由器响应RA报文。

经过以上步骤，主机生成缺省路由，下一跳指向缺省路由器的链路本地地址。

路由器发现流程 - 路由器周期性发送

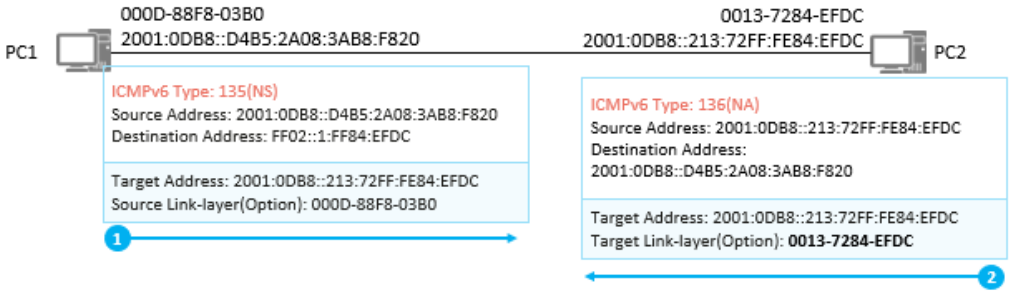
- 路由器周期性的发送RA报文，RA发送间隔是一个有范围的随机值，缺省的最大时间间隔是600秒，最小时间间隔是200秒。
- 对于定期发送的RA报文，其地址有如下要求：
  - Source Address：必须是发送接口的链路本地地址。
  - Destination Address： FF02::1。



- ipv6 nd ra { max-interval maximum-interval | min-interval minimum-interval }命令用来配置发送周期。

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。



- 本例中，当 PC1 要传送数据包到 PC2 时，如果不知道 PC2 的链路层地址，则需要完成以下协议交互过程：
- PC1 发送一个 NS 报文到网络上，目的地址为 PC2 对应

的被请求节点组播地址（FF02::1:FF84:EFDC），选项字段中带上 PC1 的链路层地址 000D-88F8-03B0。

- PC2 侦听到该 NS 报文后，由于报文的目的地址 FF02::1:FF84:EFDC，自己在该组播组，处理该报文；同时，根据 NS 报文的源地址和源链路层地址选项更新自己的邻居缓存表项。
- PC2 发送一个 NA 报文应答 NS，同时在消息的目标链路层地址选项中带上自己的链路层地址 0013-7284-EFDC。
- PC1 接收到 NA 报文后，获悉了 PC2 的链路层地址，创建一个目标节点的邻居缓存表项。
- 这样通过交互后，PC1 和 PC2 就知道了对方的链路层地址，建立其对方的邻居缓存表项（类似于 IPv4 的 ARP 表），就可以相互通信了。

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。

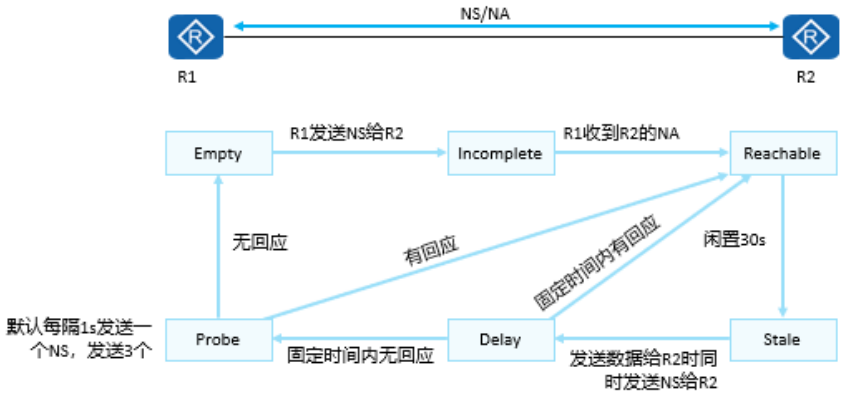
```
<Huawei>display ipv6 neighbors
------------------------------------------------------------------------------
IPv6 Address    : 2001:DB8::2
Link-layer      : 00e0-fc23-26e3              State : REACH
Interface       : GE0/0/0                     Age   : 0
VLAN            : -                           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（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（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



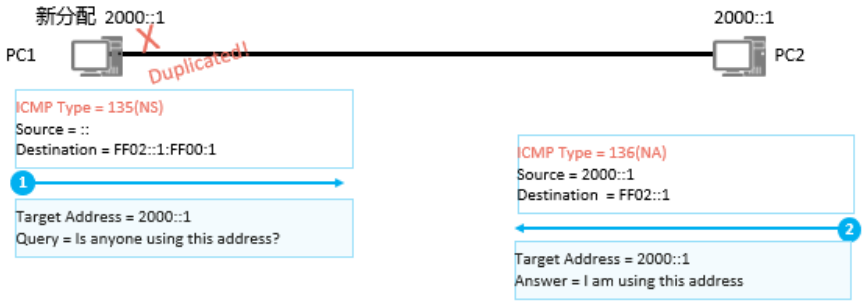
R1的邻居状态

- R1 先发送 NS 报文，并生成缓存条目，此时，邻居状态为 Incomplete。
- 若收到 R2 回复的 NA 报文，则邻居状态由 Incomplete 变为 Reachable，否则固定时间后邻居状态由 Incomplete 变为 Empty。
- 经过邻居可达时间（默认 30s），邻居状态由 Reachable 变为 Stale，即未知是否可达。
- 如果在 Reachable 状态，R1 收到 R2 的非请求 NA 报文，且其中携带的 R2 的链路层地址和表项中不同，则邻居状态马上变为 Stale。

- 在 Stale 状态若 R1 要向 R2 发送数据，则邻居状态由 Stale 变为 Delay，并发送 NS 请求。
- 在经过一段固定时间后，邻居状态由 Delay 变为 Probe，其间若有 NA 应答，则邻居状态由 Delay 变为 Reachable。
- 在 Probe 状态，R1 每隔一定时间间隔（默认 1s）发送单播 NS，发送固定次数后，有应答则邻居状态变为 Reachable，否则邻居状态变为 Empty。

重复地址检测 (1)

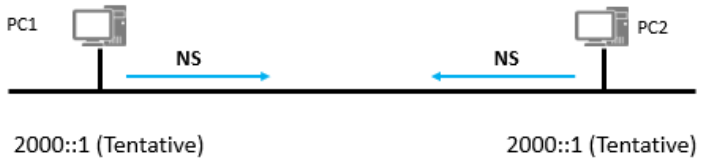
- 重复地址检测(Duplicate Address Detect，DAD)是指接口在使用某个IPv6地址之前，需要先探测是否有其它的节点使用了该地址，从而确保网络中没有两个相同的单播地址。
- 接口在启用任何一个单播IPv6地址前都需要先进行DAD，包括Link-Local地址。



- 重复地址检测是节点确定即将使用的地址是否被另一节点使用的过程。在节点自动配置某个接口的 IPv6 单播地址之前，必须在本地链路范围内验证要使用的地址是唯一的，并且未被其他节点使用过。只要 NS 报文发送到本地链路上（缺省发送一次 NS 报文），如果在规定时间内没有 NA 报文进行应答，则认为这个临时单播地址在本地链路上是唯一的，可以分配给接口；反之，这个临时地址是重复的，不能配置到接口。

重复地址检测 (2)

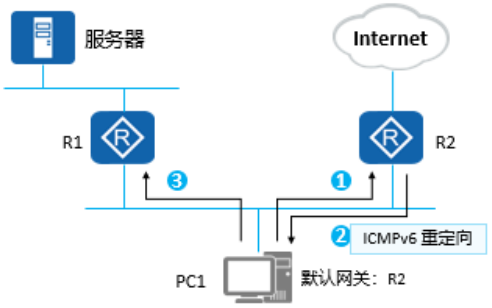
- 一个地址在通过重复地址检测之前称为“tentative地址”，即“试验地址”。此时该接口不能使用这个试验地址进行单播通讯。
- 若2个节点配置相同地址，同时作重复地址检测时，当一方收到对方发出的DAD NS报文，则接收方将不启用该地址。



- 特殊情况：有两台主机同时分配到同一个 IP 地址。假设 PC1 和 PC2 都想使用 2000::1 这个地址，那么进一步假设 PC1 先发送 NS，PC2 收到以后将不会发送 NS 了（当然也不会发送 NA），直接停止使用 2000::1 这个地址，等待其他方式生成新的地址。如果同时收到 NS 报文，则都会放弃使用 2000::1 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。

思考题：

- （多选题）ICMPv6 报文类型分为哪几大类？
- 差错报文

- 信息报文
- 其他报文
- 参数报文
- （多选题）IPv6 地址解析通过以下哪种报文实现？
- RS
- RA
- NS
- NA

答案：

- AB
- CD
-