强机保塔

Mr. Firewall's Talk Show

双机热备篇+出口选路篇7帖合辑

企业网络资料部 防火墙产品部联合出品　　2014第3期总第3期

双机热备来了
VRRP与VGMP的故事
VGMP招式详解
你所不知道的HRP

缺省路由有备无患，明细路由近路建功
多样策略择重选优，定向转发掌控先机

HUAWEI

2014年第3期总第3期
双机热备篇+出口选路篇

序

桌上摆着部门的圣诞礼物平安果，网上铺天盖地的欢度双节大促销，新的一年就这样又来临了。双节来临之际，强叔也给大家送来了一份礼物：第三期《强叔侃墙》电子合辑。

从3月13日强叔第一次在华为企业技术论坛开帖到现在，已经持续了整整9个月了；从介绍什么是防火墙到部署防火墙后网络出口如何选路，总计发布了49篇帖子及9篇案例。可以说，春夏秋冬，强叔一直在努力，笔耕不辍。一路走来，强叔将传统防火墙的基本技术写得差不多了，这里也算是个阶段性总结。

其实提笔的缘起，是想到自己在防火墙碌碌无为多年，抱着也许能广泛普及防火墙技术的想法，卷起袖子、激情澎湃地开始了码字之路。然而，理想是美丽的，现实是残酷的，写作是寂寞的，验证是耗时的，强叔并没有信心能够支撑多久。没有想到的是，贴子一经推出，就受到了众多小伙伴的关注与青睐。9个月来，连载热度一直居高不下，全国各地的小伙伴因讨论防火墙聚集在论坛，与强叔持续互动，给予强叔无数由衷的肯定与赞美，给予强叔很多宝贵的意见和启示。强叔的心也跟着这贴子热了起来。于是老夫聊发少年狂，左调墙，右著章。

由于水平有限，加上时间仓促，贴子存在着疏漏之处。当小伙伴指正一篇五千余字贴子中细节错误时，强叔很意外，也很感动，不敢想象，会有兄弟会逐字逐句研读得如此认真；不敢相信，在如今碎片化信息备受追捧、视频图片刺激感官的时代，仍然有众多兄弟对这冗长的技术文档保持着如此热情。强叔别无所长，在这风云变幻的时代，唯埋头于安全技术中苦修前行。蓦然回首，发现原来还有众多兄弟们与我同行。通过送台历活动，强叔得知，众兄弟身在全国各地，天南海北。然讨论起防火墙及安全技术，则天涯如咫尺。强叔亦因此豁然开朗。莫愁前路无知己，天下谁人不识君。无以为报，唯更加勤勉，筹划新一篇章写作思路去也。

接下来，强叔将重点介绍当前主力销售的下一代防火墙关键技术。再加个好消息，由于之前连载广受欢迎，给强叔投资的人马也开始多了起来，尤其是，强叔侃墙将在明年得以正式出版，还请各位小伙伴届时继续支持强叔。

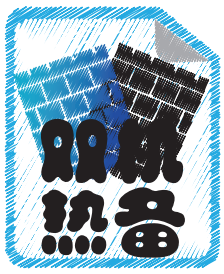
2015年，我们不见不散。

强叔

2014年12月24日于北京环保园

目录

Contents



双机热备来了

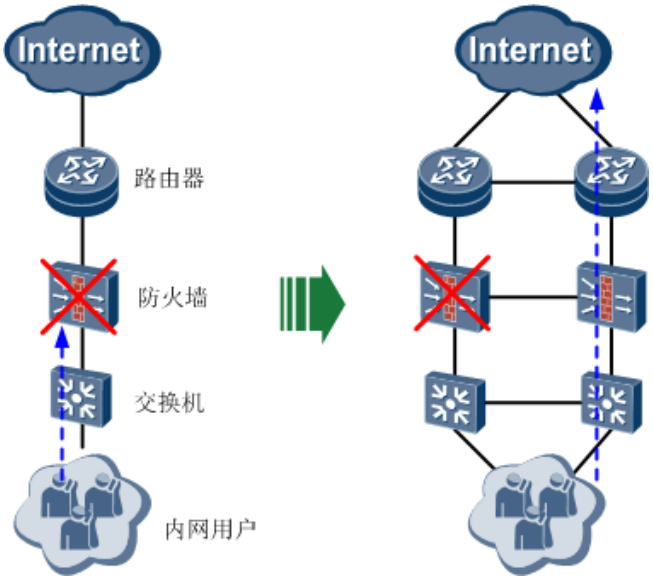
在经历了漫长的学习过程后，强叔终于带大家领略完了防火墙的各种基本功能，想必各位小伙们一定是大有收获的。之前强叔讲到的都是在一台防火墙上配置各种功能，而为了提升网络的可靠性，我们经常需要在两台防火墙上配置相同的功能并使他们相互备份。那么这是如何做到的呢？

这就需要用到强叔本次为大家带来的防火墙一大特色功能——双机热备。双机热备来了，光怪陆离的双机热备真的来了~

双机部署提升网络可靠性

随着移动办公、网上购物、即时通讯、互联网金融、互联网教育等业务蓬勃发展，网络承载的业务越来越多，越来越重要。所以如何保证网络的不间断传输成为网络发展过程中急需解决的一个问题。

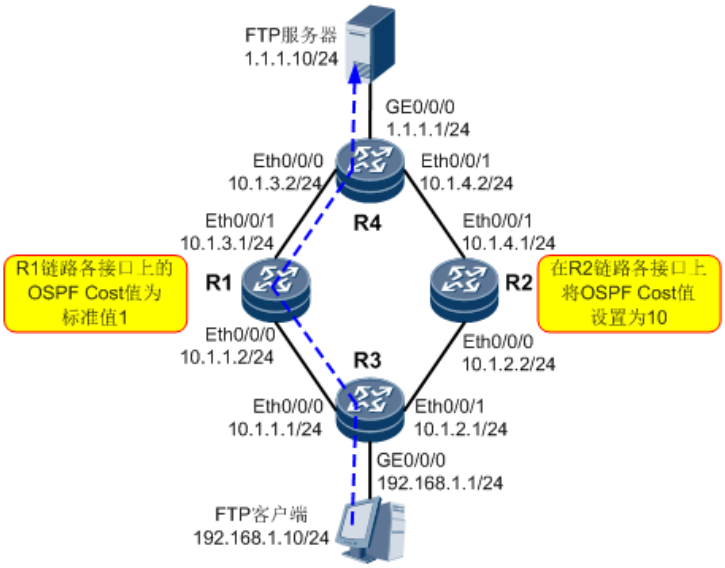
如左下图所示，防火墙部署在企业网络出口处，内外网之间的业务都会通过防火墙转发。如果防火墙出现故障，便会导致内外网之间的业务全部中断。由此可见，在这种网络关键位置上如果只使用一台设备的话，无论其可靠性多高，我们都必然要承受因设备单点故障而导致网络中断的风险。于是，**我们在网络架构设计时，通常会在网络的关键位置部署两台（双机）或多台设备，以提升网络的可靠性。**如右下图所示，当一台防火墙出现故障时，流量会通过另外一台防火墙所在的链路转发，保证内外网之间业务正常运行。



路由器的双机部署只需考虑路由备份

如果是传统的网络转发设备（如路由器、三层交换机），只需要在两台设备上做好路由的备份就可以保证业务的可靠性。因为普通的路由器、交换机不会记录报文的交互状态和应用层信息，只是根据路由表进行报文转发，下面举个例子来说明。

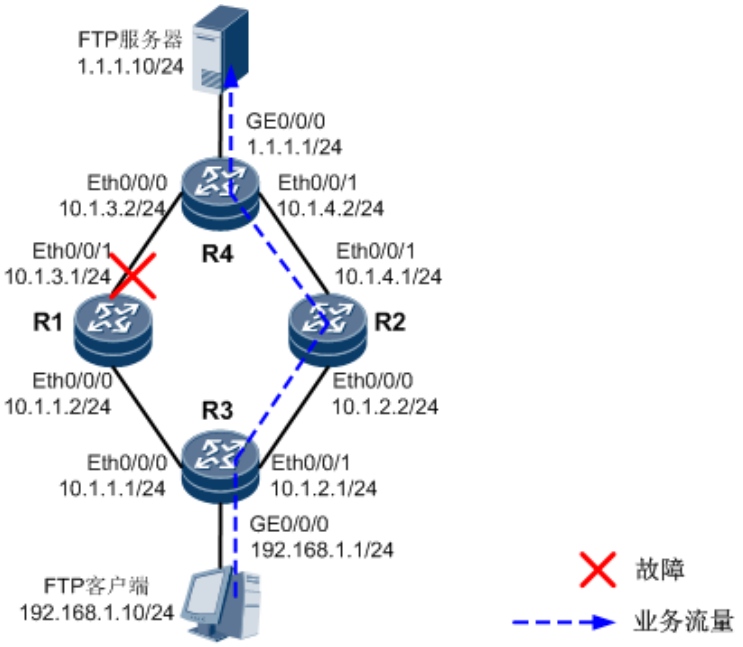
如下图所示，两台路由器 R1 和 R2 与上下行设备 R3 和 R4 之间运行 OSPF 协议。正常情况下，由于以太网接口的缺省 OSPF Cost 值为 1，所以在 R3 上看 R1 所在链路（R3—>R1—>R4—>FTP 服务器）的 Cost 值为 3。而由于我们在 R2 链路（R3—>R2—>R4—>FTP 服务器）的各接口上将 OSPF Cost 值设置为 10，所以在 R3 上看 R2 所在链路的 Cost 值为 21。由于流量只会通过 Cost 值小的链路转发，所以 FTP 客户端与服务器间的业务就都只会通过 R1 转发。



由于系统在选择路由时只会选择最优的路由，所以R3的路由表中只能看到Cost值较小的路由。这样去往FTP服务器（目的地址在1.1.1.0/24网段）的报文只能通过R1（下一跳10.1.1.2）转发。

```
[R3]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 11        Routes : 11

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.0/24  OSPF    10   3           D   10.1.1.2        Ethernet0/0/0
       10.1.1.0/24  Direct  0    0           D   10.1.1.1        Ethernet0/0/0
       10.1.1.1/32  Direct  0    0           D   127.0.0.1       Ethernet0/0/0
       10.1.2.0/24  Direct  0    0           D   10.1.2.1        Ethernet0/0/1
       10.1.2.1/32  Direct  0    0           D   127.0.0.1       Ethernet0/0/1
       10.1.3.0/24  OSPF    10   2           D   10.1.1.2        Ethernet0/0/0
       10.1.4.0/24  OSPF    10   12          D   10.1.1.2        Ethernet0/0/0
      127.0.0.0/8   Direct  0    0           D   127.0.0.1       InLoopBack0
      127.0.0.1/32  Direct  0    0           D   127.0.0.1       InLoopBack0
    192.168.1.0/24  Direct  0    0           D   192.168.1.1     GigabitEthernet0
/0/0
    192.168.1.1/32  Direct  0    0           D   127.0.0.1       GigabitEthernet0
/0/0
```

如下图所示，当 R1 出现故障时，R1 所在链路 Cost 值变成无穷大，而在 R3 上看 R2 所在链路 Cost 值仍为 21。这时网络的路由会重新收敛，流量会根据新的路由被转发到 R2，所以 R2 会接替 R1 处理业务。业务从 R1 切换到 R2 的时间就是网络的路由收敛时间。如果路由收敛时间较短，则正在传输的业务不会中断。



当R1的Eth0/0/1接口故障时，从R3上的路由表可知，去往FTP服务器（目的地址在1.1.1.0/24网段）的报文只能通过R2（下一跳10.1.2.2）转发。

```
[R3]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 10        Routes : 10

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.0/24  OSPF    10   21          D   10.1.2.2        Ethernet0/0/1
       10.1.1.0/24  Direct  0    0           D   10.1.1.1        Ethernet0/0/0
       10.1.1.1/32  Direct  0    0           D   127.0.0.1       Ethernet0/0/0
       10.1.2.0/24  Direct  0    0           D   10.1.2.1        Ethernet0/0/1
       10.1.2.1/32  Direct  0    0           D   127.0.0.1       Ethernet0/0/1
       10.1.4.0/24  OSPF    10   20          D   10.1.2.2        Ethernet0/0/1
      127.0.0.0/8   Direct  0    0           D   127.0.0.1       InLoopBack0
      127.0.0.1/32  Direct  0    0           D   127.0.0.1       InLoopBack0
    192.168.1.0/24  Direct  0    0           D   192.168.1.1     GigabitEthernet0
/0/0
    192.168.1.1/32  Direct  0    0           D   127.0.0.1       GigabitEthernet0
/0/0
```

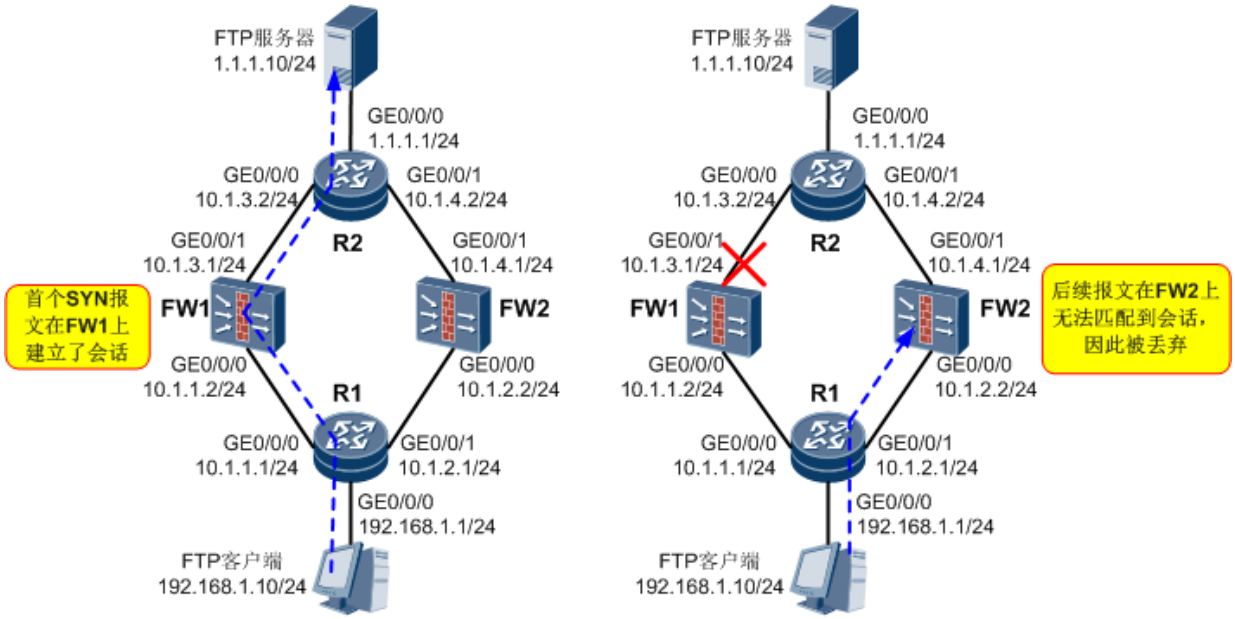
防火墙的双机部署还需考虑会话备份

如果将传统网络转发设备换成状态检测防火墙，情况就大不一样了。回忆一下强叔在“状态检测和会话机制”中讲到的内容：状态检测防火墙是基于连接状态的，他会对一条流量的首包（第

一个报文）进行完整的检测，并**建立会话来记录报文的状态信息**（包括报文的源 IP、源端口、目的 IP、目的端口、协议等）。而这条流量的后续报文只有匹配会话才能够通过防火墙并且完成报文转发，**如果后续报文不能匹配会话则会被防火墙丢弃**。

下面举个例子来说明，两台防火墙 FW1 和 FW2 部署在网络中，与上下行设备 R1 和 R2 之间运行 OSPF 协议。如左下图所示，正常情况下，由于 FW1 所在链路的 OSPF Cost 值较小，所以业务报文都会根据路由通过 FW1 转发（原理同前面的路由器的例子）。**这时 FW1 上会建立会话，业务的后续报文都能够匹配会话并转发。**

如右下图所示，当 FW1 出现故障时，业务会被上下行设备上的路由信息引导到 FW2 上（原理同前面的路由器的例子）。但由于 FW2 上没有会话，**业务报文因为找不到会话而被 FW2 丢弃，从而导致业务中断。**这时用户需要重新发起访问请求（例如重新进行 FTP 下载），触发 FW2 重新建立会话，这样用户的业务才能继续进行。



FW1上存在会话：

```
[FW1]display firewall session table

 Current Total Sessions : 1
  ftp  VPN:public --> public 192.168.1.10:2050-->1.1.1.10:21
```

FW2上不存在会话：

```
[FW2]display firewall session table

 Current Total Sessions :0
```

双机热备出手不凡，解决防火墙会话备份问题

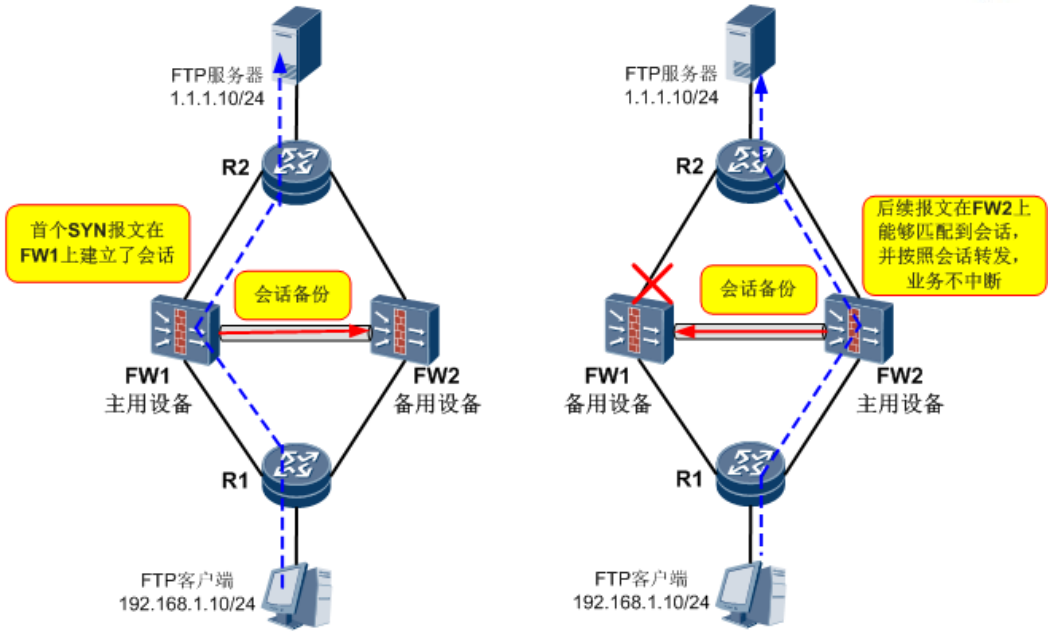
那么如何解决两台防火墙会话备份的问题，使两台防火墙主备状态切换时，保证已经建立的

业务不中断呢？这时**防火墙双机热备功能**就该出手相助了！

如左下图所示，防火墙双机热备功能最大的特点在于提供一条专门的**备份通道（也称为心跳线）**，用于两台防火墙之间协商主备状态，以及备份会话、Server-map 表等重要的状态信息和配置信息。双机热备功能启动后，正常情况下，两台防火墙会根据管理员的配置分别成为主用设备和备用设备。成为**主用设备**的防火墙 FW1 会处理业务，并将设备上的会话、Server-map 表等重要状态信息以及配置信息通过备份通道实时同步给**备用设备** FW2。成为**备用设备**的防火墙 FW2 不会处理业务，只是通过备份通道接收来自主用设备 FW1 的状态信息以及配置信息。

如右下图所示，当主用设备 FW1 发生故障时，两台防火墙会利用备份通道交互报文，重新协商主备状态。这时 FW2 会协商成为新的主用设备，处理业务；而 FW1 会协商成为备用设备，不处理业务。与此同时，业务流量也会被上下行设备的路由信息引导到新的主用设备 FW2 上。由于 FW2 在作为备用设备时已经备份了主用设备上的会话和配置等信息，因此业务报文就能够顺利的匹配到会话从而被正常转发。

以上两点就保证了备用设备 FW2 能够成功接替原主用设备 FW1 处理业务流量，成为新的主用设备，避免了网络业务中断。



FW1上存在会话：

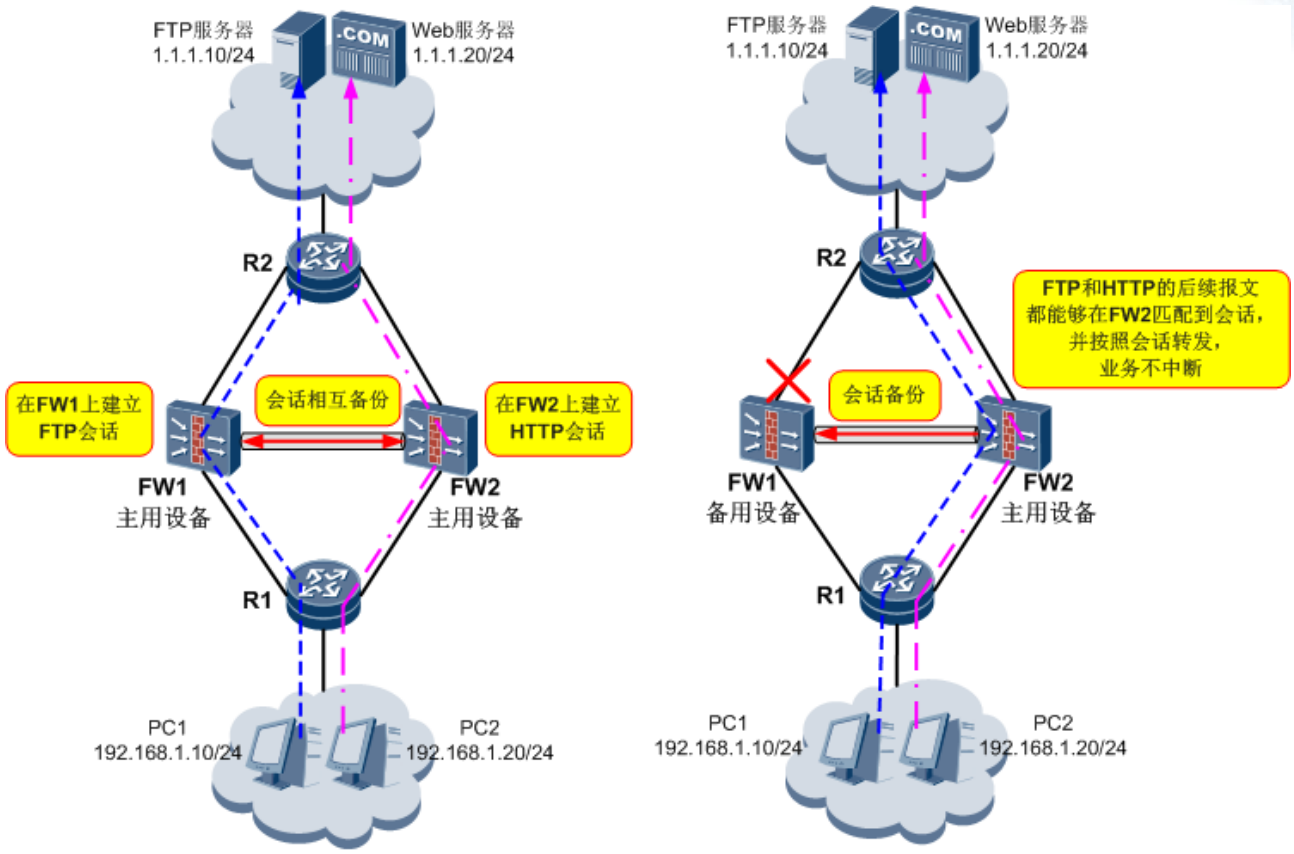
```
[FW1]display firewall session table

 Current Total Sessions : 1
  ftp  VPN:public --> public 192.168.1.10:2050-->1.1.1.10:21
```

FW2上也存在会话：

```
[FW2]display firewall session table

 Current Total Sessions : 1
  ftp  VPN:public --> public 192.168.1.10:2050-->1.1.1.10:21
```

上面介绍的是**主备备份**方式的双机热备。在主备备份场景中，正常情况下备用设备不处理业务流量，处于闲置状态。如果小伙伴们不希望买来的设备闲置，或者只一台设备处理流量时压力较大，可以选择**负载分担**方式的双机热备。

如左下图所示，在**负载分担**场景下，两台防火墙均为主用设备，都建立会话，都处理业务流量。同时两台防火墙又都相互作为对方的备用设备，接受对方备份的会话和配置信息。如右下图所示，当其中一台防火墙故障后，另一台防火墙会负责处理全部业务流量。由于这两台防火墙的会话信息是相互备份的，因此全部业务流量的后续报文都能够在其中一台防火墙上匹配到会话从而正常转发，这就避免了网络业务的中断。



FW1上存在FTP和HTTP会话：

```
[FW1]display firewall session table

 Current Total Sessions : 2
  ftp  VPN:public --> public 192.168.1.10:2050-->1.1.1.10:21
  http VPN:public --> public 192.168.1.20:2080-->1.1.1.20:80
```

FW2上也存在FTP和HTTP会话：

```
[FW2]display firewall session table

 Current Total Sessions : 2
  ftp  VPN:public --> public 192.168.1.10:2050-->1.1.1.10:21
  http VPN:public --> public 192.168.1.20:2080-->1.1.1.20:80
```

总结

最后，强叔再来简单总结下本回所讲的内容。

为了提升网络可靠性，避免单点故障的风险，我们需要在网络关键节点处部署两台网络设备。如果是路由器和交换机，我们只需要做好路由的备份即可。如果是防火墙，我们还必须在两台防火墙之间备份会话表等状态信息。

防火墙的**双机热备**功能提供一条专门的**备份通道**，用于两台防火墙之间协商主备状态，以及会话等状态信息的备份。双机热备主要包括主备备份和负载分担场景。**主备备份是**指正常情况下仅由主用设备处理业务，备用设备空闲；当主用设备接口、链路或整机故障时，备用设

备切换为主用设备，接替主用设备处理业务。**负载分担**也可以称为“互为主备”，即两台设备同时处理业务。当其中一台设备发生故障时，另外一台设备会立即承担其业务，保证原来需要通过这台设备转发的业务不中断。

在本节了解了双机热备功能的由来和概念的基础上，后面强叔将为大家一步步讲解双机热备的实现原理。各位小伙伴敬请期待！

VRRP 与 VGMP 的故事（上）

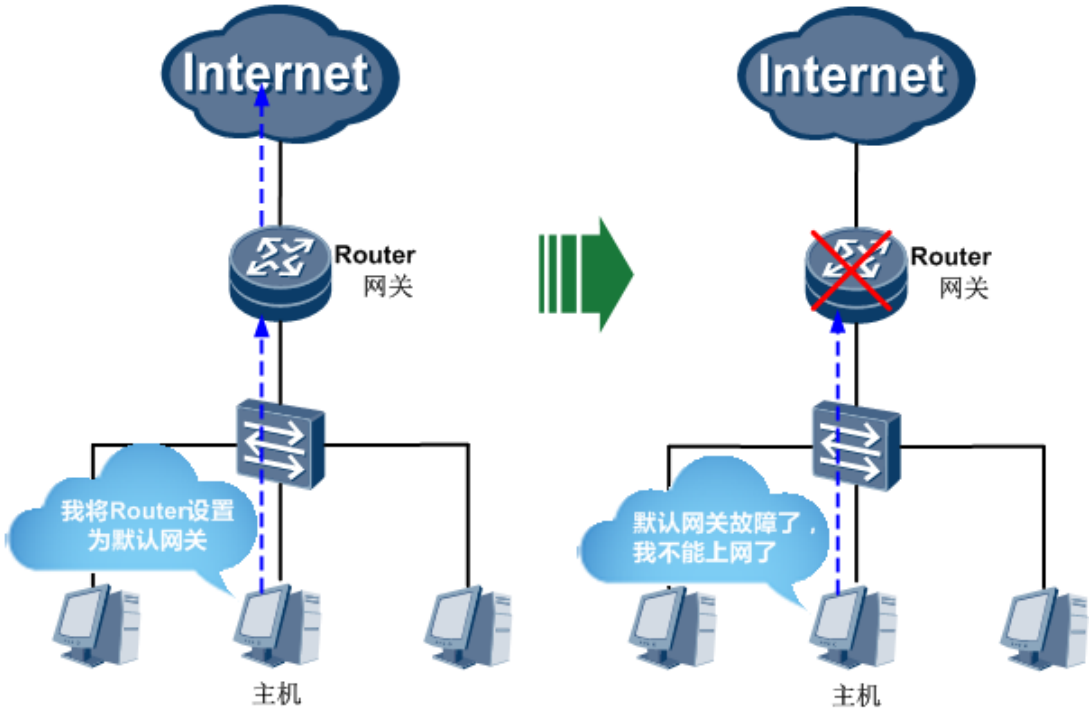
上期强叔为大家讲解了网络可靠性的作用，以及防火墙与路由器可靠性的区别，并初步介绍了防火墙双机热备功能的基本概念。我想小伙伴们一定已经迫不及待地想要了解双机热备功能是如何实现的了。

熟悉路由器和交换机的小伙伴们一提到网络可靠性，首先想到的肯定是 VRRP 协议，其实防火墙的双机热备功能也是在 VRRP 协议的基础上扩展而来的。所以本节我们会首先介绍 VRRP 的基本概念和实现原理,之后我们会讲解 VRRP 在防火墙双机热备应用中遇到的问题，以及如何通过双机热备的核心 VGMP 协议来解决 VRRP 的问题。最后我们会通过 VGMP 报文结构的详解来初步揭开 VGMP 实现原理的面纱，并体会 VRRP 报文在防火墙中的变化。

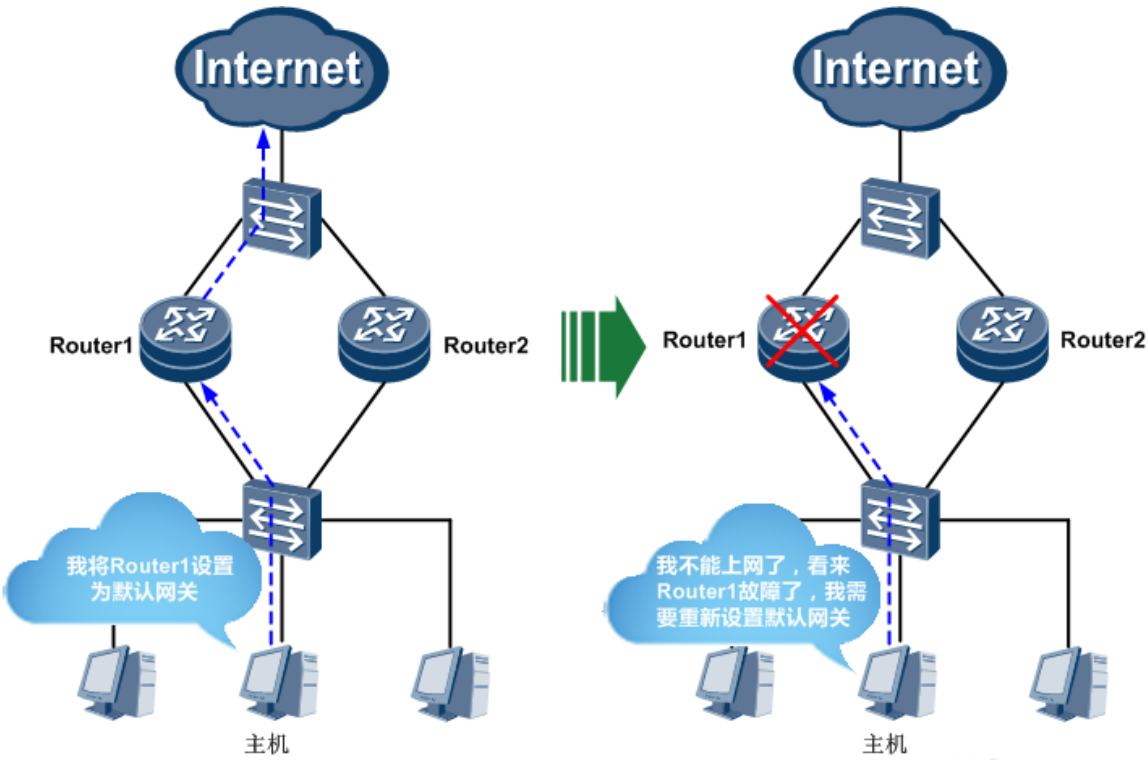
双机部署提升网络可靠性

在上期讲的路由器或防火墙可靠性组网（双机热备）中，流量被引导到主用还是备用设备都是由上下行设备的路由表决定的。这是因为**动态路由可以根据链路状态动态调整路由表，自动将流量引导到正确的设备上。**但如果上下行设备运行的是静态路由呢？静态路由可是无法动态调整的啊。

下面我们就来看一个例子：如下图所示，主机将 Router 设置为默认网关。这样当主机想访问外部网络时，就会先将报文先发送给网关，再由网关传递给外部网络，从而实现主机与外部网络的通信。正常的情况下，主机可以完全信赖网关的工作，但是当网关故障时，主机与外部的通信就会中断。

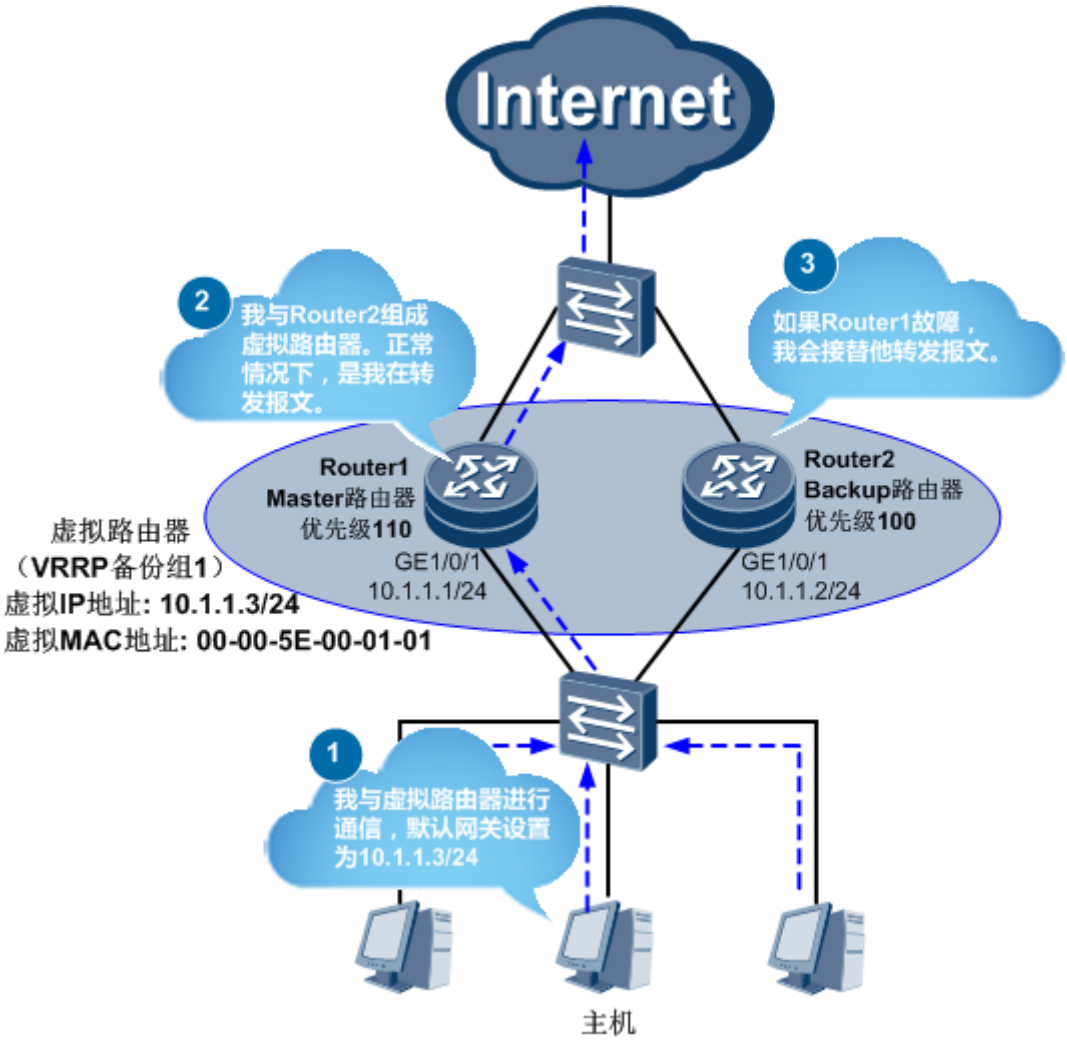


如下图所示,如果想要解决网络中断的问题,我们就需要添加多个网关(Router1 和 Router2)。但一般情况下主机不能配置动态路由，而且只会配置一个默认网关。如果我们把 Router1 设置成默认网关，那么当 Router1 出现故障时，流量无法被自动引导到 Router2 上。这时只有手工调整主机的默认网关为 Router2，才能将主机的流量引导到 Router2 上。但是这样必然会导致主机访问外网的流量中断一段时间，从而影响用户业务的正常运行。而且大型网络中的主机是成百上千的，通过手动调整网络实现网关备份显然是不切实际的。



为了更好地解决由于网关故障引起的网络中断问题，网络开发者提出了 VRRP 协议。**VRRP 是一种容错协议，它保证当主机的下一跳路由器（默认网关）出现故障时，由备份路由器自动代替出现故障的路由器完成报文转发任务，从而保持网络通信的连续性和可靠性。**

如下图所示，我们将局域网内的一组路由器（实际上是路由器的下行接口）划分在一起，形成一个 VRRP 备份组。**VRRP 备份组**相当于一台**虚拟路由器**，这个虚拟路由器有自己的**虚拟 IP 地址和虚拟 MAC 地址**（格式：00-00-5E-00-01-{VRID}，VRID 是 VRRP 备份组的 ID）。所以，局域网内的主机可以将默认网关设置为 VRRP 备份组的虚拟 IP 地址。**在局域网内的主机看来，他们就是与虚拟路由器进行通信的，**然后通过虚拟路由器与外部网络进行通信。



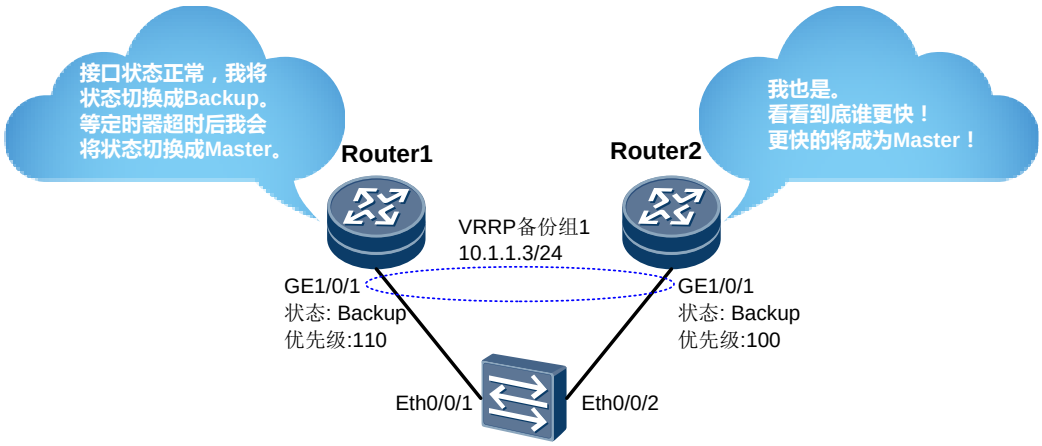
VRRP 备份组中的多个路由器会根据管理员指定的 VRRP 备份组优先级确定各自的 VRRP 备份组状态。优先级最高的 VRRP 备份组状态为 Master，其余 VRRP 备份组状态为 Backup。VRRP 备份组的状态决定了路由器的主备状态。VRRP 备份组状态为 Master 的路由器称为 **Master 路由器**，VRRP 备份组状态为 Backup 的路由器称为 **Backup 路由器**。当 Master 路由器正常工作时，局域网内的主机通过 Master 路由器与外界通信。当 Master 路由器出现故障时，一台 Backup 路由器（VRRP 优先级次高的）将成为新的 Master 路由器，接替转发报

文的工作，保证网络不中断。

图解 VRRP 工作过程

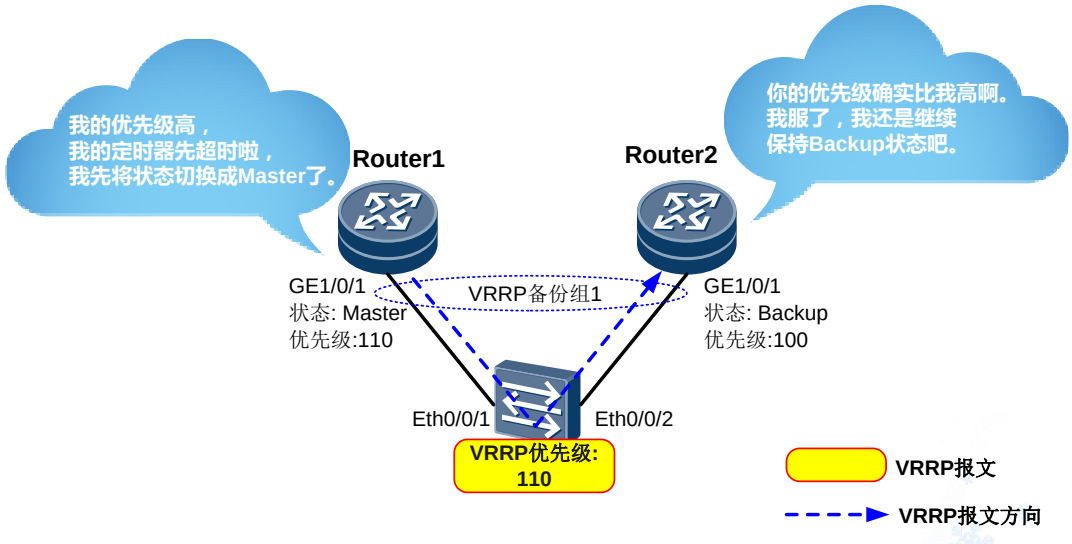
强叔在这里采取一种图说的方式来呈现 VRRP 工作的全流程，借此帮助小伙伴们来理解 VRRP 的实现原理。大家只要看完并记住下面的图，就一定能理解并记忆 VRRP 协议。

1、管理员在路由器上配置完VRRP备份组和优先级后，VRRP备份组会短暂的工作在Initialize状态。当VRRP备份组收到接口Up的消息后，会切换成Backup状态，等待定时器超时后再切换至Master状态。

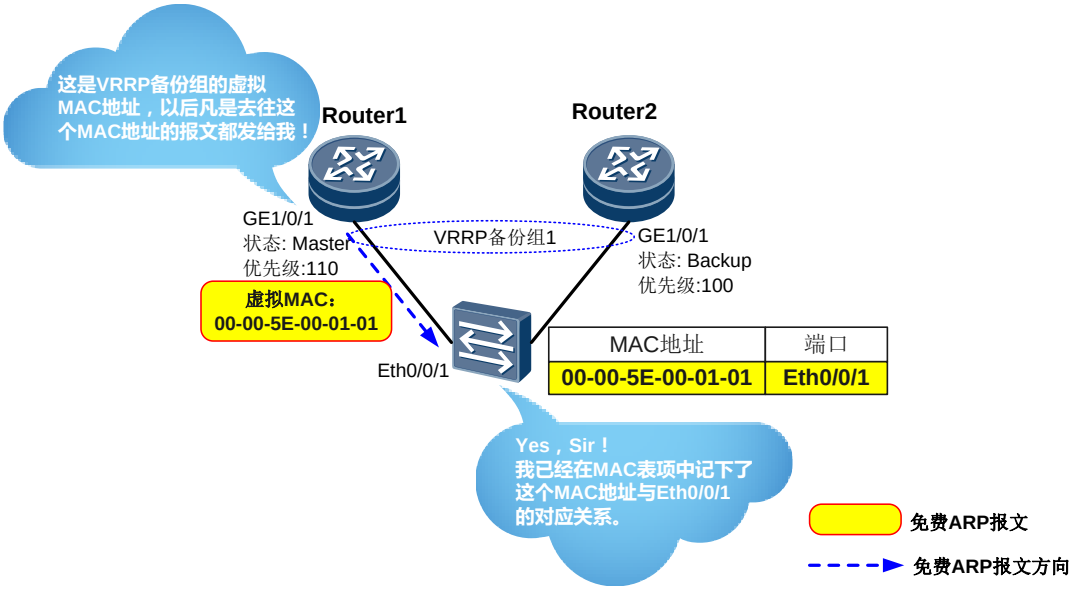


2、在 VRRP 备份组的多个路由器中，**率先将 VRRP 备份组状态切换成 Master 的路由器将会成为 Master 路由器。VRRP 备份组优先级越高的路由器，他的定时器长越短，越容易成为 Master 路由器。**这个根据 VRRP 备份组优先级确定 Master 路由器的过程称为 Master 路由器选举。

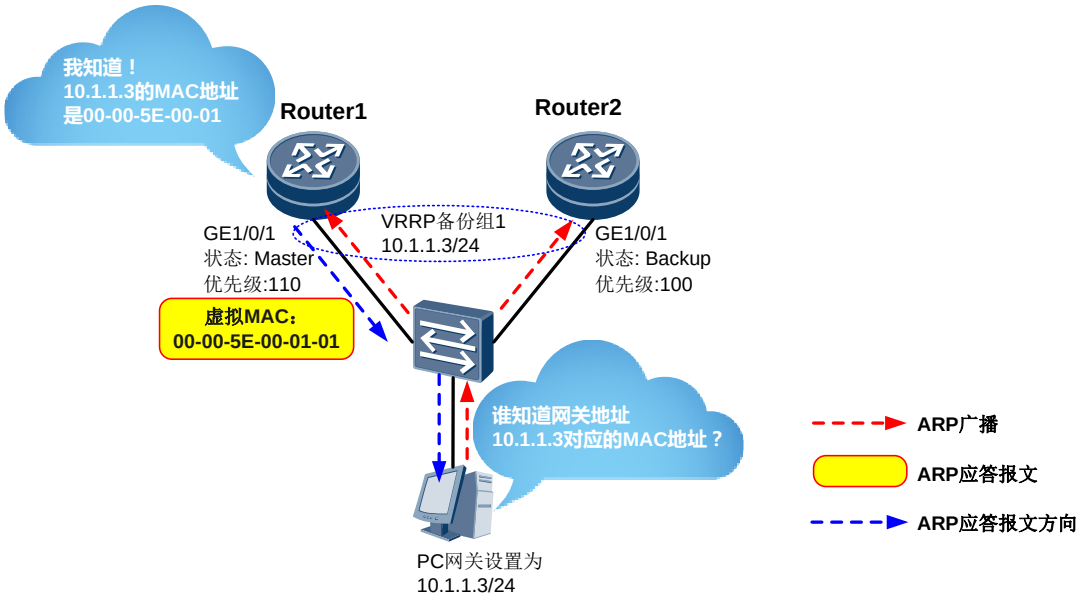
选举成功后，Master 路由器会立即周期性（缺省为 1 秒）地向 VRRP 备份组中内的所有 Backup 路由器发送 VRRP 报文，以通告自己的 Master 状态和优先级。



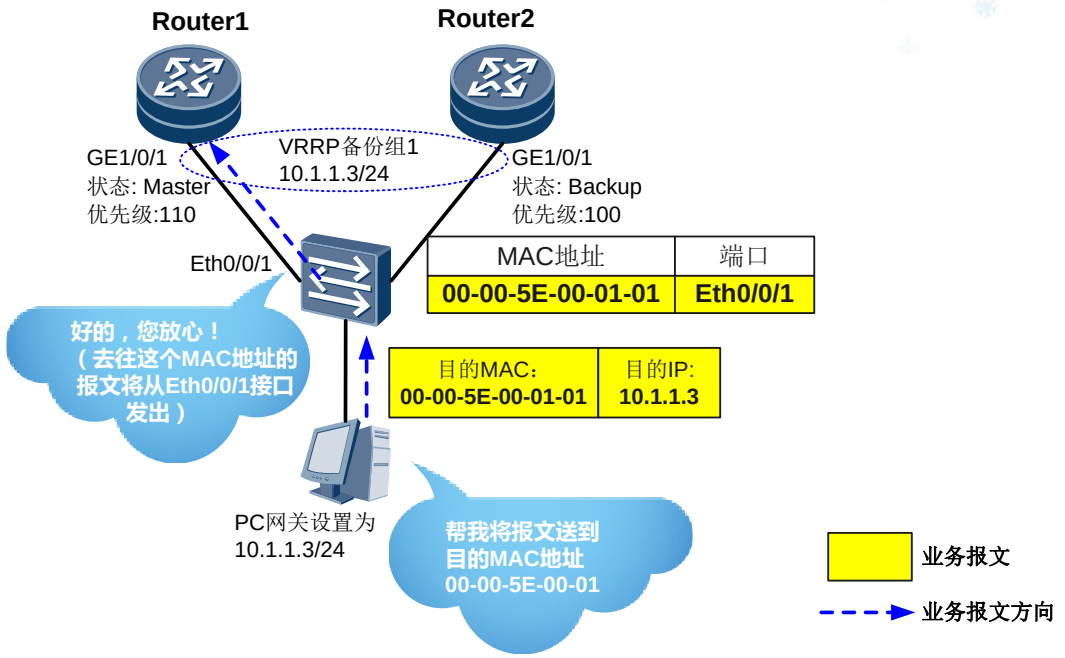
3、同时 Master 路由器会发送免费 ARP 报文，将 VRRP 备份组的虚拟 MAC 地址通知给与它连接的交换机。下行的交换机的 MAC 表项会记录虚拟 MAC 地址与端口 Eth0/0/1 的对应关系。



4、由于内网的 PC 将网关设置为 VRRP 备份组 1 的虚拟 IP 地址，所以当内网 PC 访问 Internet 时，首先会在广播网络中广播 ARP 报文，请求虚拟 IP 地址对应的虚拟 MAC 地址。这时只有 Master 路由器会回应此 ARP 报文，将虚拟 MAC 地址回应给 PC。

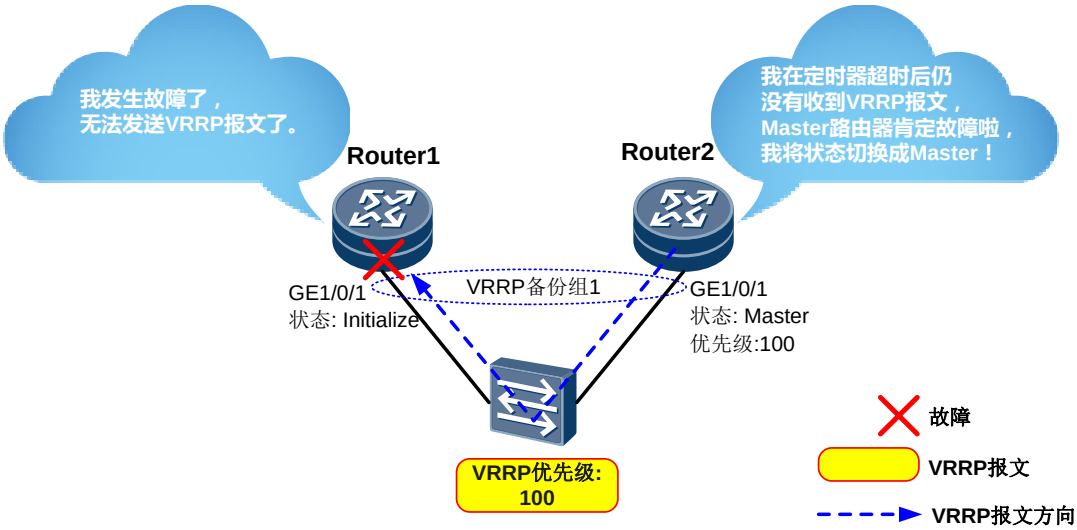


5、PC 使用虚拟 MAC 地址作为目的 MAC 地址封装报文，然后将其发送至交换机。交换机根据 MAC 表记录的 MAC 地址与端口的关系，将 PC 发送的报文通过端口 Eth0/0/1 转发给 Router1。



以上讲的是正常情况下，Master 路由器和 Backup 路由器的状态建立和运行过程。下面将介绍 Master 路由器和 Backup 路由器的状态切换和运行过程。

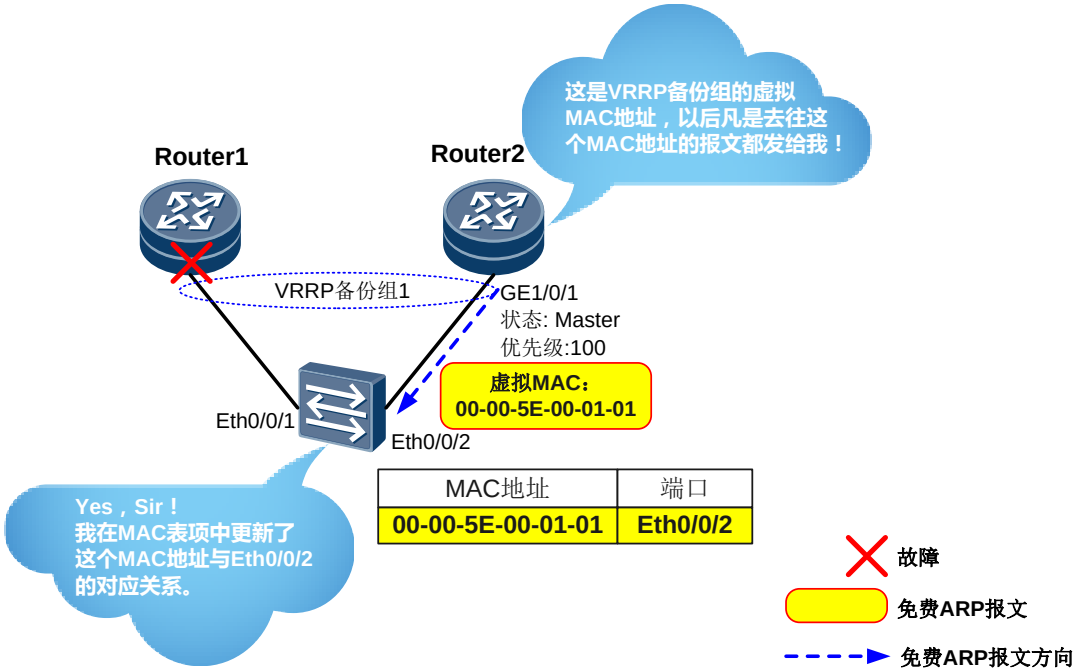
1、当 Master 路由器发生故障（Router1 整机或接口 GE1/0/1 故障）时，他将无法发送 VRRP 报文通知 Backup 路由器。Backup 路由器有这样一个机制：如果 Backup 路由器在定时器超时后仍不能收到 Master 路由器发送的 VRRP 报文，则认为 Master 路由器故障，从而将自身状态切换为 Master。



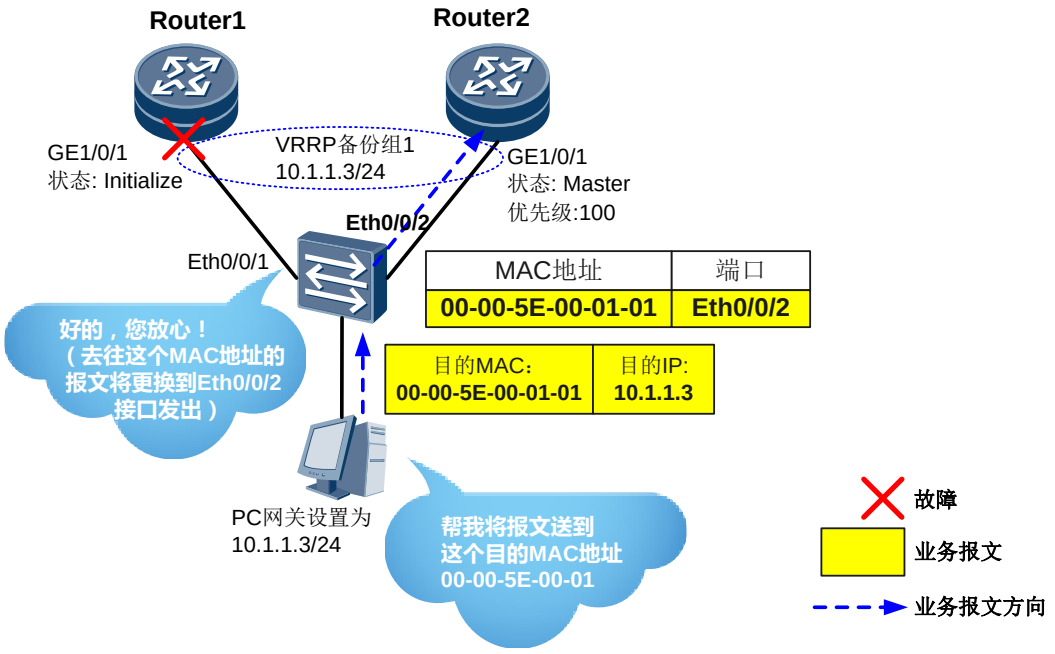
还有一种情况：当 Master 路由器主动放弃 Master 地位（如 Master 路由器退出 VRRP 备份组）时，会立即发送优先级为 0 的 VRRP 报文，使 Backup 路由器快速切换成 Master 路由器。

2、当 VRRP 备份组状态切换完成后，新的 Master 路由器会立即发送携带 VRRP 备份组虚

拟 MAC 地址和虚拟 IP 地址信息的免费 ARP 报文，刷新与它连接的设备（下行交换机）中的 MAC 表项。下行的交换机的 MAC 表项会记录虚拟 MAC 地址与新的端口 Eth0/0/2 的对应关系。

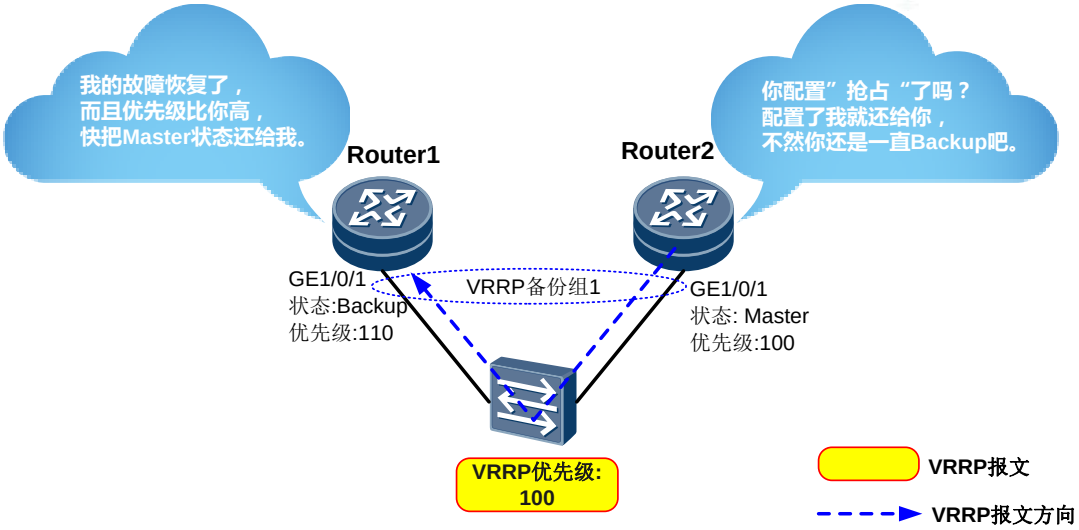


3、当内网 PC 将报文发送给交换机后，交换机会将 PC 发送的报文通过端口 Eth0/0/2 转发给 Router2。这样内网 PC 的流量就都通过新的 Master 路由器 Router2 转发了。这个过程对用户是完全透明的，内网 PC 感知不到 Master 路由器已经由 Router1 切换成 Router2。



4、当原 Master 路由器（现 Backup 路由器）故障恢复后，优先级会高于现在的 Master 路由器。这时如果配置了抢占功能，原 Master 路由器会将状态切换成 Master，重新成为 Master

路由器；如果没有配置抢占功能，原 Master 路由器将仍然保持 Backup 状态。

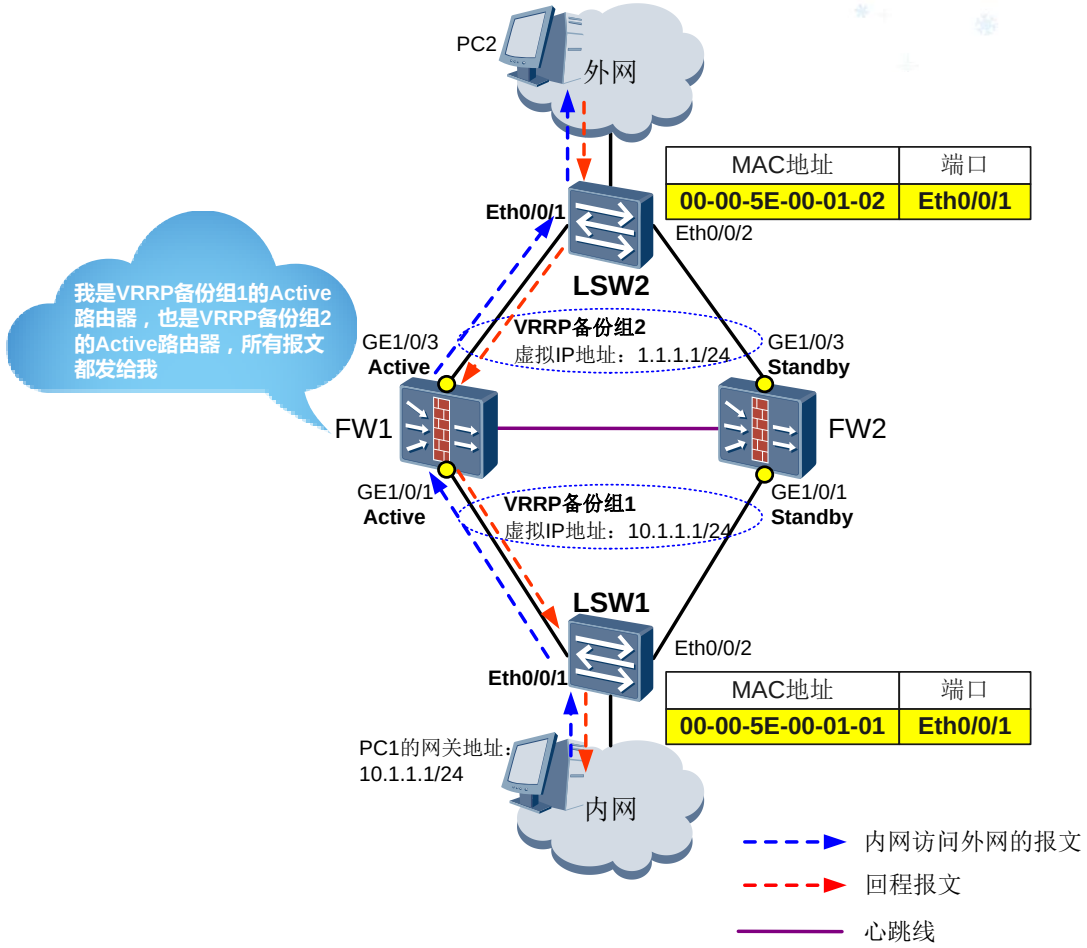


多个 VRRP 状态相互独立产生问题

上面强叔为大家讲解了 VRRP 协议，描述了如何通过 VRRP 实现网络的可靠性。VRRP 的图解看起来确实美如画，那么 VRRP 一定是完美无瑕了吗？其实并不是这样的。VRRP 本身还是存在问题的，而我们下面讲到的 VGMP 正是为了解决 VRRP 的问题而诞生的。

上面讲到通过在网关的下行接口运行 VRRP，可以解决网关的可靠性问题。如果我们在网关的上行和下行接口上同时运行 VRRP，这时情况会是怎样的呢？

现在我们就一起来看一下：如下图所示，两台 FW（作为内外网用户的网关）的下行接口加入 VRRP 备份组 1，上行接口加入 VRRP 备份组 2。正常情况下，FW1 的 VRRP 备份组 1 的状态为 Active，VRRP 备份组 2 的状态为 Active，所以 FW1 是 VRRP 备份组 1 中的 Active 路由器，也是 VRRP 备份组 2 的 Active 路由器。这样由我们上面讲的 VRRP 原理可知，内外网之间的业务报文都会通过 FW1 转发。

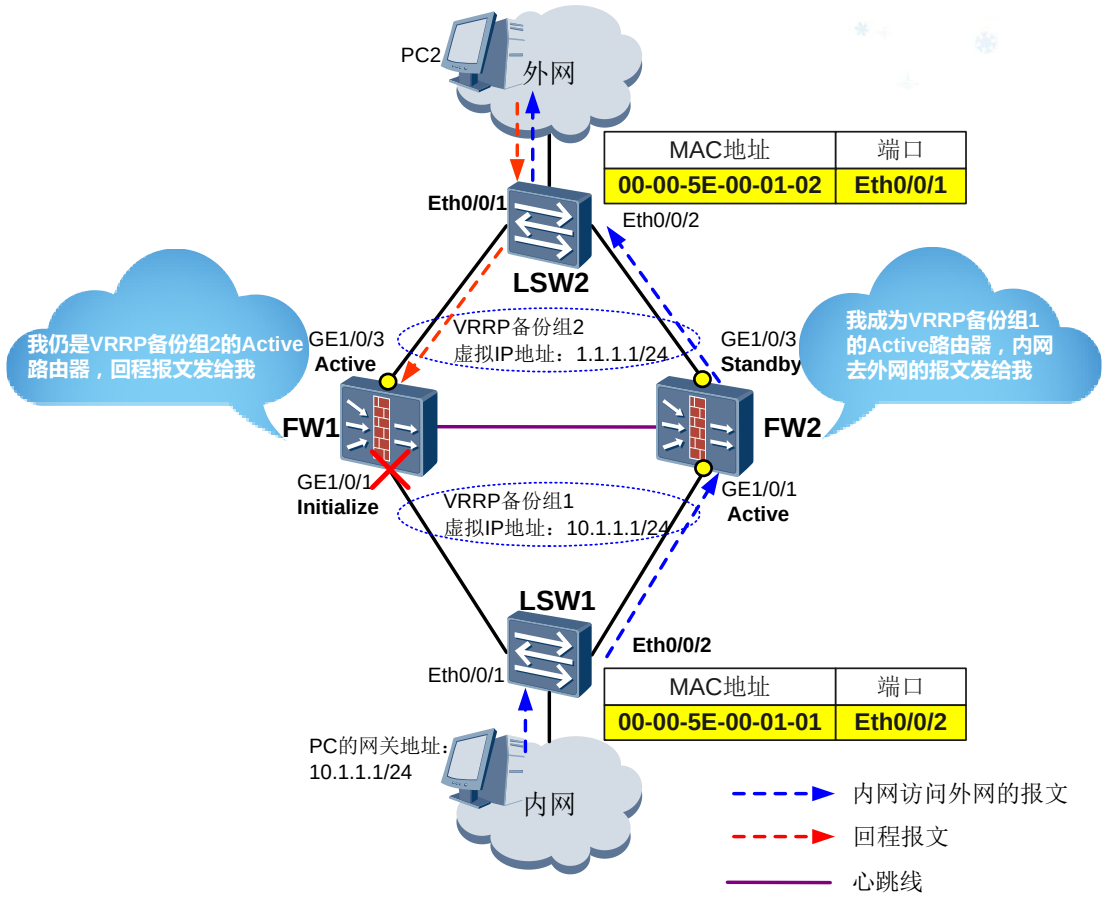


【强叔问答】上面我们在讲解 VRRP 时提到的状态都是“Master”和“Standby”，这里为什么都变成“Active”和“Standby”了呢？

答：防火墙在 NGFW 版本中统一将双机热备（原为“Master”和“Slave”）和 VRRP（原为“Master”和“Standby”）的状态修改为“Active”和“Standby”。所以当你在某些文档中看到这些以前的状态请不要奇怪，按本文描述的“Active”和“Standby”理解即可。

如下图所示，当 FW1 的 GE1/0/1 接口故障时，VRRP 备份组 1 发生状态切换：FW1 的 VRRP 备份组 1 状态切换成 Initialize，FW2 的 VRRP 备份组 1 状态切换成 Active。这样 FW2 成为 VRRP 备份组 1 中的 Active 路由器，并向 LSW1 发送免费 ARP 报文，刷新 LSW1 中的 MAC 表项。这时 PC1 访问 PC2（内网访问外网）的报文就通过 FW2 转发了。

但是由于 FW1 与 LSW2 之间的链路是正常的，所以 VRRP 备份组 2 的状态是不变的，FW1 仍然是 VRRP 备份组 2 中的 Active 路由器，而 FW2 仍是 VRRP 备份组 2 中的 Standby 路由器。因此 PC2 返回给 PC1 的回程报文依然会转发给 FW1，而 FW1 的下行接口 GE1/0/1 是故障的，所以 FW1 只能丢弃此回程报文，这就导致了业务流量的中断。



看完这个过程后，聪明的小伙伴们就会发现 VRRP 问题的所在了：VRRP 备份组之间是相互独立的，当一台设备上出现多个 VRRP 备份组时，他们之间的状态无法同步。

这个问题是 VRRP 无法适应防火墙的致命因素，既然 VRRP 自身无法克服这两点，自然会有高手及时登场。一方唱罢，一方登场，这一点上 IP 江湖跟现实世界没什么不同！

VGMP 的产生解决了 VRRP 的问题

为了解决多个 VRRP 备份组状态不一致的问题，华为防火墙引入 VGMP（VRRP Group Management Protocol）来实现对 VRRP 备份组的统一管理，保证多个 VRRP 备份组状态的一致性。我们将防火墙上的所有 VRRP 备份组都加入到一个 VGMP 组中，由 VGMP 组来集中监控并管理所有的 VRRP 备份组状态。如果 VGMP 组检测到其中一个 VRRP 备份组的状态变化，则 VGMP 组会控制组中的所有 VRRP 备份组统一进行状态切换，保证各 VRRP 备份组状态的一致性。

VGMP 有状态和优先级两个基本属性，并且有三条基本运行原则：

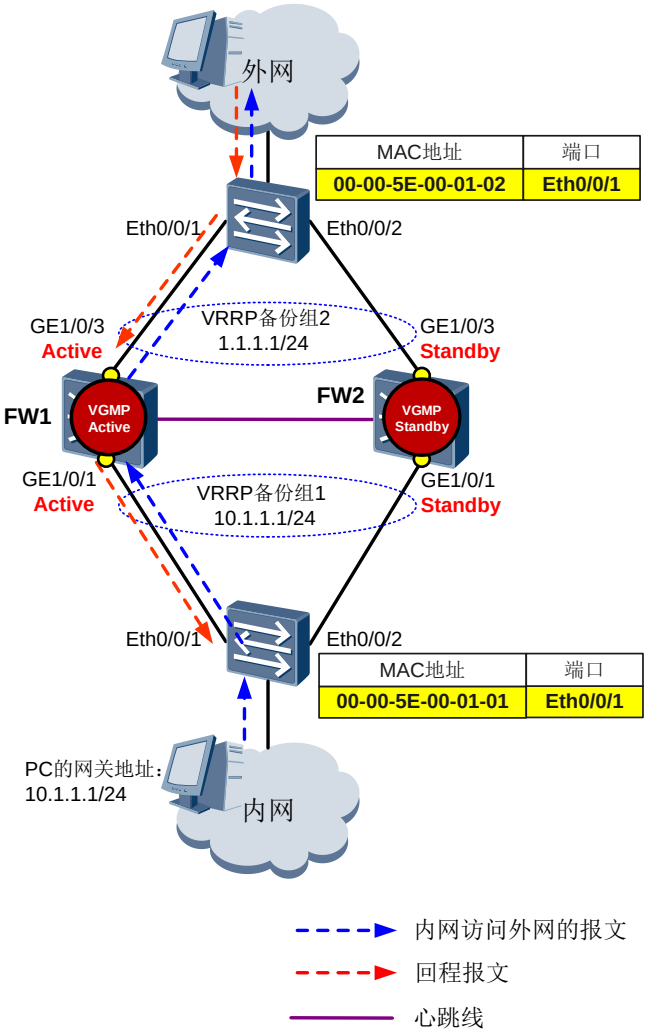
- **VGMP的状态决定了组内VRRP备份组的状态，也决定了防火墙的主备状态。**
- **VGMP组的状态是由两台防火墙的VGMP组通过比较优先级来决定的。优先级高的**

VGMP组状态为Active，优先级低的VGMP组状态为Standby

- **VGMP组会根据组内VRRP备份组的状态变化来更新自己的优先级。每个VRRP备份组的状态变成Initialize， VGMP组的优先级就会降低2。**

了解并熟记了 VGMP 的基本原则后，下面我们一起来看 VGMP 如何解决 VRRP 问题。

如下图所示，我们在 FW1 上将 VRRP 备份组 1 和 VRRP 备份组 2 都加入状态为 Active 的 VGMP 组，在 FW2 上将 VRRP 备份组 1 和 VRRP 备份组 2 都加入状态为 Standby 的 VGMP 组。由于 VGMP 组的状态决定了组内 VRRP 备份组的状态，所以 FW1 上 VRRP 备份组 1 和 2 的状态都为 Active，FW2 上 VRRP 备份组 1 和 2 的状态都为 Standby。这样 FW1 就是 VRRP 备份组 1 和 VRRP 备份组 2 中的 Active 路由器（也就是两台防火墙中的主用设备），而 FW2 就是他们的 Standby 路由器（也就是两台防火墙中的备用设备），所以上下行的业务流量都会被引导到主用设备 FW1 转发。



如下图所示，当 FW1 的接口故障时，VGMP 组控制 VRRP 备份组状态统一切换的过程如下：

1） 当FW1的GE1/0/1接口故障时，FW1上的VRRP备份组1发生状态切换（由Active切换成

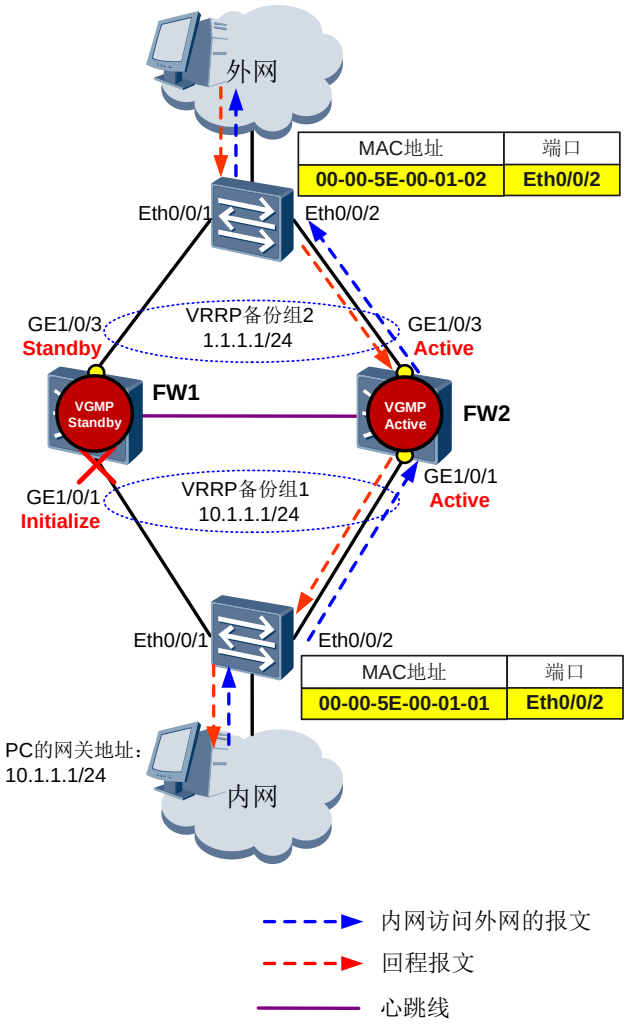
Initialize）。

2）FW1的VGMP组感知到这一故障后，会降低自身的优先级，然后与FW2的VGMP组比较优先级，重新协商主备状态。

3）协商后，FW1的VGMP组状态由Active切换成Standby，FW2的VGMP组状态由Standby切换成Active。

4）同时，由于VGMP组的状态决定了组内VRRP备份组的状态，所以FW1的VGMP组会强制组内的VRRP备份组2由Active切换成Standby状态，FW2的VGMP组也会强制组内的VRRP备份组1和2由Standby切换成Active状态。这样FW2就成为了VRRP备份组1和VRRP备份组2中的Active路由器，也就成了为两台防火墙中的主用设备；而FW1则成为了VRRP备份组1和VRRP备份组2中的Standby路由器，也就成为了两台防火墙中的备用设备。

5）FW2会分别向LSW1和LSW2发送免费ARP，更新他们的MAC转发表，使PC1访问PC2的上行报文和回程报文都转发到FW2。这样就完成了VRRP备份组状态的统一切换，并且保证业务流量不会中断。



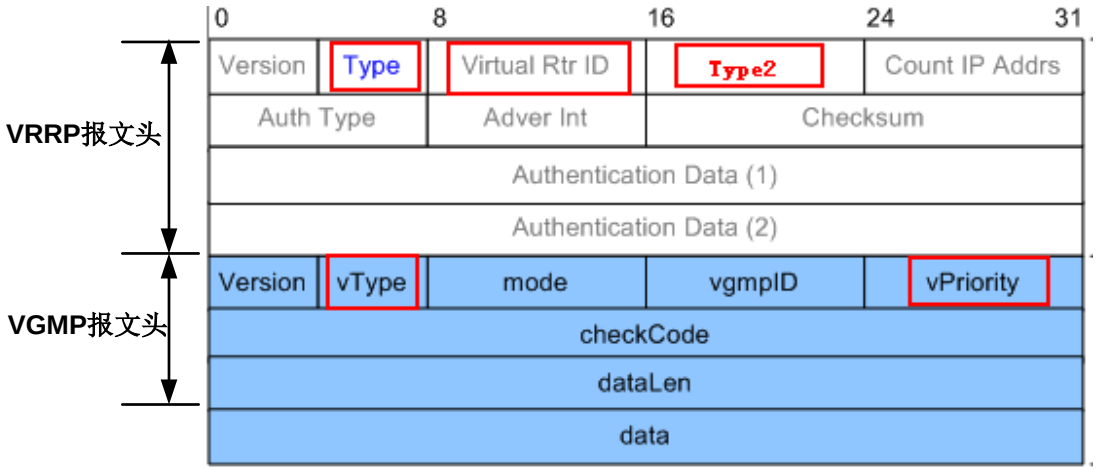
详解 VGMP 报文结构

看到以上内容大家应该明白，VGMP 不仅完成了 VRRP 备份组的统一管理，还借势取代 VRRP 接管了对防火墙主备状态的管理权。那么这时问题来了：

- **两台防火墙的VGMP组是如何传递VGMP的优先级信息的？**
- **两台防火墙的VGMP组的状态协商和切换过程，以及报文交互过程到底是怎样的呢？**

在前面的 VRRP 图解中讲到两台路由器的 VRRP 备份组是通过 VRRP 报文来传递优先级信息的。那么两台防火墙的 VGMP 组还是通过 VRRP 报文和机制来传递优先级信息的吗？这当然不太可能，新的领导自然有新的方法。**两台防火墙的 VGMP 组是通过 VGMP 报文来传递优先级信息的。**VGMP 是华为的私有协议，他为了实现防火墙双机热备功能对 VRRP 报文进行了扩展和修改，并衍生出多种使用 VGMP 报文头封装的报文。理解 VGMP 报文和报文头是理解 VGMP 状态协商和切换的基础，所以让我们先看一下 VGMP 报文的结构。

说明：本篇所讲到的 VGMP 报文结构适用于适用于 USG2000/5000/6000 系列防火墙和 USG9000 系列防火墙的 V1R3 版本。



如上图所示，从 VGMP 报文封装顺序中我们可以发现，VGMP 报文是根植于 VRRP 报文的，是由 VRRP 报文头封装的。但这个 VRRP 报文并不是标准的 VRRP 报文，是经过华为扩展和修改的，具体有以下几点变化：

- 标准VRRP报文的“Type”字段只有“1”一个取值，我们增加了“2”取值。也就是说如果Type=1，就是标准的VRRP报文；如果Type=2，就是我们修改后的VRRP报文。
- 标准VRRP报文的“Virtual Rtr ID”字段代表VRRP备份组ID，而修改后的VRRP报文“Virtual Rtr ID”取值固定为“0”。
- 修改后的VRRP报文中去掉了标准VRRP报文的“IP Address”字段。
- 标准VRRP报文中的“Priority”字段在VRRP报文头中被修改成“Type2”字段。

- 当Type2=1时，报文封装成**心跳链路探测报文**。心跳链路探测报文用于检测对端设备的心跳口能否正常接收本端设备的报文，以确定是否有心跳口可以使用。
- 当Type2=5时，报文封装成**一致性检查报文**。一致性检查报文用于检测双机热备状态下的两台防火墙是否配置了相同的策略。
- 当Type2=2时，VRRP报文才会进一步封装VGMP报文头，并根据VGMP报文头中“vType”字段继续分成以下三种报文：
  - **VGMP报文（VGMP Hello报文）。VGMP Hello报文用于两台防火墙间的VGMP组协商主备状态。这也正是我们问题的答案所在。**
  - **HRP心跳报文（HRP Hello报文）。**HRP心跳报文用于探测对端设备是否处于工作状态。主用设备会每隔一段时间（缺省为1s）向备用设备发送HRP心跳报文，用来通知主用设备处于工作状态。如果备用设备在三个周期内没有收到HRP心跳报文，则认为主用设备故障，而自身切换成主用设备。
  - **HRP数据报文。**我们还需要在VGMP报文头后继续增加HRP报文头，才能封装成HRP数据报文。HRP数据报文用于主备设备之间的数据备份，包括命令行配置的备份和各种状态信息的备份。

看到这里大家是否会问，在防火墙双机热备中 VRRP 报文是用来封装 VGMP 报文的，那标准的 VRRP 报文还存在么，它还有什么作用？答案是**标准 VRRP 报文仍旧存在，它还是用于 VRRP 备份组内部通信。只是其中的优先级字段（Priority）已经为固定值，无法配置，所以标准 VRRP 报文实际上已名存实亡。**优先级字段失去作用导致标准 VRRP 报文已经无法控制 VRRP 备份组的状态选举了，只能在主备防火墙之间通告一下 VRRP 备份组的状态和虚拟 IP 地址了。这跟宪政体制下的“皇帝”的作用相似，保留名号，但没有管理国家的权利。

而 VGMP 想要接管防火墙和 VRRP 备份组的状态管理，就意味着 VGMP 报文必须中要包含 VGMP 组优先级信息。我们再看一下 VGMP 报文头的结构：

- “Mode”字段表示是请求类型报文还是应答类型报文。
- “ID”字段表示VGMP组是Active组还是Standby组。
- **“Priority”字段表示VGMP组的优先级。**

这点表明VGMP备具备接替标准VRRP报文管理VRRP备份组和防火墙状态的物质基础。

综上所述，VGMP 协议修改了标准的 VRRP 报文并定义好了各种使用 VGMP 报文头封装的报文，那么这些报文是通过什么渠道在两台防火墙之间传递的呢？上面我们讲到两台防火墙通过备份通道（心跳线）来传递备份数据，可见 HRP 数据报文是通过备份通道传输的。实际

上以上所讲各种报文（除标准 VRRP 报文），包括 VGMP 报文都是通过备份通道传输的。

另外 USG6000 系列防火墙和 USG2000/5000 系列 V3R1 版本防火墙还支持将以上所讲的各种 VGMP 和 HRP 报文（除标准 VRRP 报文）封装成 UDP 报文，具体结构如下：

UDP Header	VRRP Header	VGMP Header	DATA

那么只使用 VRRP 封装的报文和使用 UDP 封装的报文有什么区别？前者是组播报文，不能跨越网段传输，不受安全策略控制；后者是单播报文，只要路由可达就可以跨越网段传输，但是受安全策略控制。具体点来说就是如果是组播报文，那么两台防火墙的心跳口之间就必须直连或通过二层交换机相连，但是不需要配置安全策略；如果是单播报文，那么两台防火墙的心跳口之间可以通过路由器这种三层设备相连，但是需要配置安全策略允许报文在 local 区域与心跳口所在安全区域间双向通过。

讲完 VGMP 的报文结构大家应该能回答“两台防火墙的 VGMP 组是如何传递 VGMP 的优先级信息的”这个疑问了吧？答案就是**两台防火墙的 VGMP 组通过备份通道（心跳线），利用 VGMP 报文来传递优先级信息**。

这时小伙伴们一定还会关心第二个问题：两台防火墙的 VGMP 组的状态协商和切换过程，以及报文交互过程到底是怎样的呢？。强叔是否能够像图解 VRRP 工作过程一样，图解 VGMP 的状态切换过程呢？欲知详情如何，且听下回分解。

VRRP 与 VGMP 的故事（下）

上篇我们在学习了 VRRP 的实现原理，以及如何通过 VGMP 解决 VRRP 问题后，提出了两个问题。第一个问题是“两台防火墙的 VGMP 组是如何传递 VGMP 的优先级信息的？”。答案是**防火墙的 VGMP 组通过备份通道（心跳线），利用 VGMP 报文来传递优先级信息。**

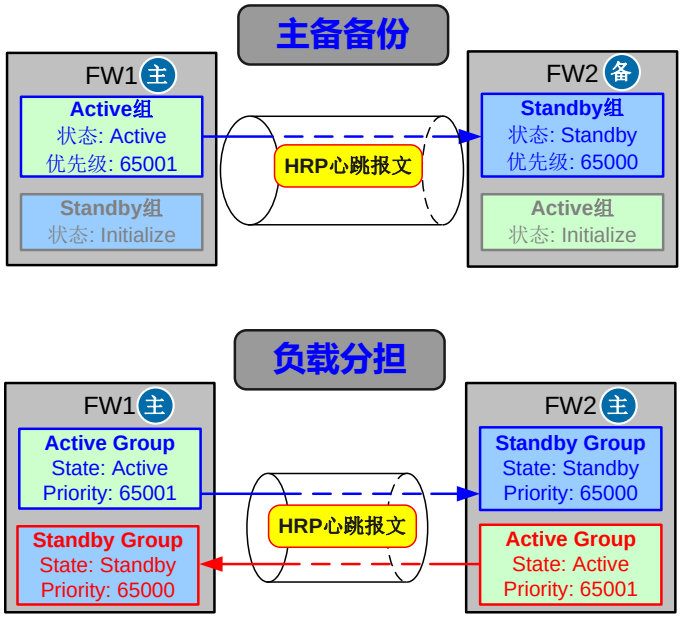
第二个问题是“两台防火墙的 VGMP 组的状态协商和切换过程，以及报文交互过程到底是怎样的呢？”，我们还没有给出答案。

在本篇中，强叔将通过图解的方式完整地给出各种情况下，两台防火墙的 VGMP 组的状态协商和切换过程，以及报文交互过程，保证您有看大片儿的感觉~

防火墙 VGMP 组的缺省情况

在欣赏大片儿前，我们首先来了解下背景知识，也就是防火墙 VGMP 组的缺省情况。

如下图所示，每台防火墙提供两个 VGMP 组：Active 组和 Standby 组。缺省情况下，Active 组的优先级为 65001，状态为 Active；Standby 组的优先级为 65000，状态为 Standby。主备备份情况下，主用设备启用 Active 组，所有成员（例如 VRRP 备份组）加入 Active 组；备用设备启用 Standby 组，所有成员加入 Standby 组。负载分担情况下，两台设备都启用 Active 组和 Standby 组，每台设备上的所有成员分别加入 Active 组和 Standby 组。FW1 的 Active 组和 FW2 的 Standby 组形成一组“主备”，FW2 的 Active 组和 FW1 的 Standby 组形成一组“主备”，两台防火墙互为“主备”，形成负载分担。



以上讲的是中低端防火墙 **USG2000/5000/6000 系列**的情况，由于高端防火墙存在接口板和业务板，所以高端防火墙的缺省优先级与中低端防火墙是不同的：

高端防火墙 Master 组的缺省优先级=45001+1000×（业务板个数＋接口板个数）。

Slave 组的缺省优先级=45000+1000×（业务板个数＋接口板个数）。

说明：这里说明的高端防火墙优先级算法以 USG9000 系列 V1R3 版本为例。而本篇的双机热备配置和状态切换过程以中低防火墙的 USG2000/5000/6000 系列为例。

主备备份双机热备状态形成过程

主备备份方式的双机热备是目前较常用的双机方式，配置和理解也比较简单，因此我们先从主备备份双机热备状态形成的过程来看。

为了让大家能够真实地感受到 VRRP 和 VGMP 在防火墙上的存在，我们下面会先给出防火墙主备备份双机热备的配置，然后描述配置完成后双机热备状态形成的过程。

为了实现主备备份方式的双机热备，我们需要在 FW1 上启用 Active 组，并将 FW1 上的 VRRP 备份组都加入 Active 组；在 FW2 上启用 Standby 组，并将 FW2 上的 VRRP 备份组都加入 Standby 组。实现此操作的命令为 **vrrp vrid** *virtual-router-id* **virtual-ip** *virtual-address* [*ip-mask* | *ip-mask-length*] { **active** | **standby** }。这条命令看似简单，但功能很强大，一条命令配置下去两件事轻松搞定：

- 由于是在接口视图下执行这条命令，所示实际上是将接口加入了VRRP备份组，同时指定了虚拟IP地址和掩码。当接口的IP地址与VRRP备份组的虚拟IP地址不在同一网段时，必须配置虚拟IP地址的掩码。
- “active | standby”参数是将VRRP备份组加入Active或Standby组。

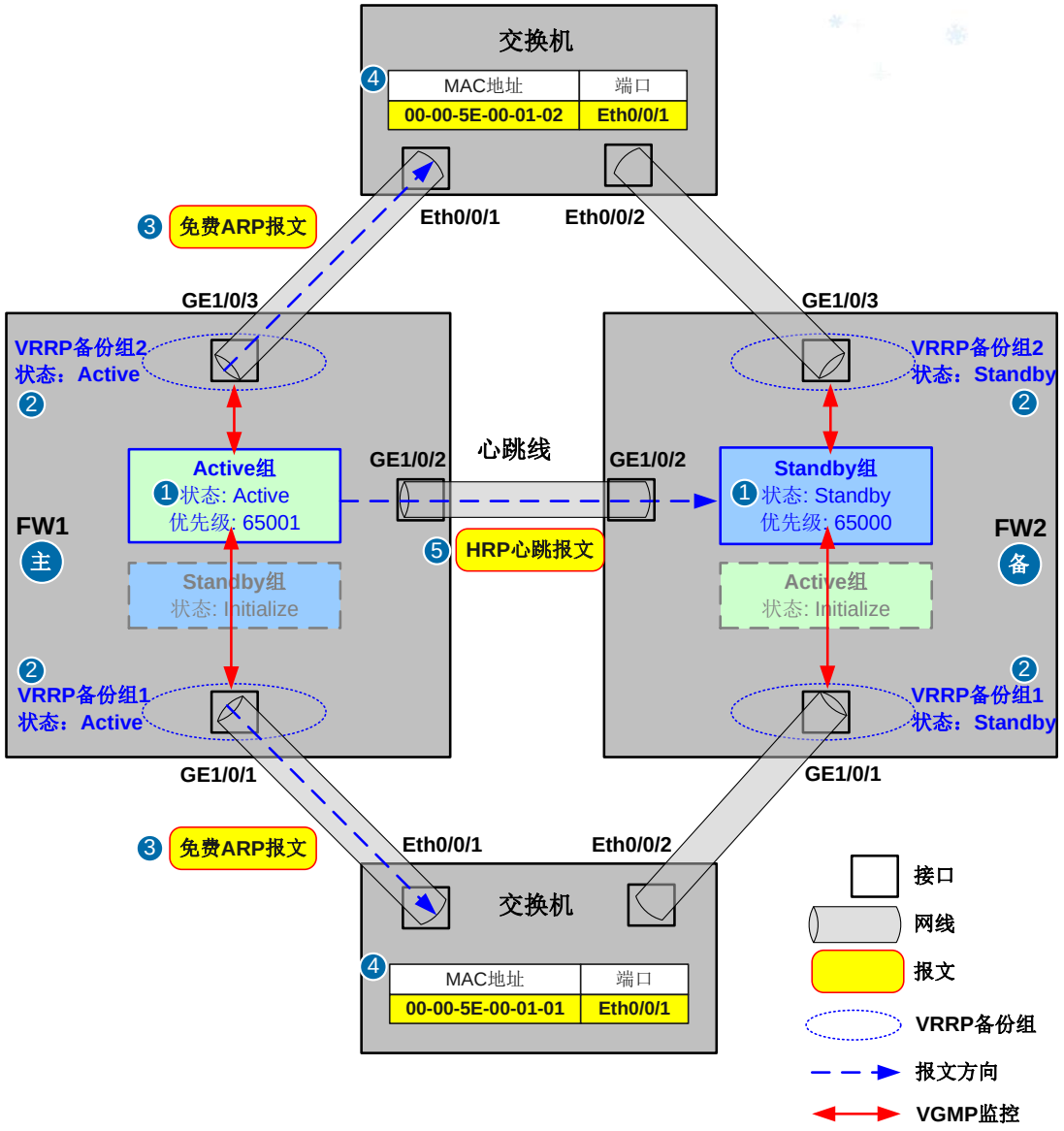
结合下面的图来给出我们建立主备方式双机热备的具体配置，如下表所示：

FW1 的配置	**FW2 的配置**
interface GigabitEthernet 1/0/1 **ip address 10.1.1.2 255.255.255.0** **vrrp vrid 1 virtual-ip 10.1.1.1 255.255.255.0 active** //将接口 GE1/0/1 加入 VRRP 备份组 1，并将 VRRP 备份组 1 加入 Active 组。	**interface GigabitEthernet 1/0/1** **ip address 10.1.1.3 255.255.255.0** **vrrp vrid 1 virtual-ip 10.1.1.1 255.255.255.0 standby** //将接口 GE1/0/1 加入 VRRP 备份组 1，并将 VRRP 备份组 1 加入 Standby 组。
interface GigabitEthernet 1/0/3 **ip address 1.1.1.2 255.255.255.0** **vrrp vrid 2 virtual-ip 1.1.1.1 255.255.255.0 active** //将接口 GE1/0/3 加入 VRRP 备份组 2，并将 VRRP 备份组 2 加入 Active 组。	**interface GigabitEthernet 1/0/3** **ip address 1.1.1.3 255.255.255.0** **vrrp vrid 2 virtual-ip 1.1.1.1 255.255.255.0 standby** //将接口 GE1/0/3 加入 VRRP 备份组 2，并将 VRRP 备份组 2 加入 Standby 组。

FW1 的配置	FW2 的配置
hrp interface GigabitEthernet 1/0/2 //指定心跳口 **hrp enable** //启用双机热备功能	**hrp interface GigabitEthernet 1/0/2** //指定心跳口 **hrp enable** //启用双机热备功能

各种 VGMP 报文和 HRP 报文都是通过心跳口发送的，心跳口可以说是双机热备的“命脉”，要点多多，务必关注：

- 两台设备的心跳口必须加入相同的安全区域。
- 两台设备的心跳口的接口类型和编号必须相同。例如主用设备的心跳接口为GigabitEthernet 1/0/2，那么备用设备的心跳接口也必须为GigabitEthernet 1/0/2。
- 配置心跳口时如果不添加**remote**参数，则两台设备的心跳口需要直接相连或通过二层交换机相连，并且不需要配置安全策略。如果配置心跳口时添加remote参数（例如hrp interface **GigabitEthernet 1/0/2 remote 10.1.1.2**），那么两台设备的心跳口可以通过路由器相连，但是需要配置安全策略。因为不添加**remote**参数时，心跳口发送的报文是用VRRP报文封装的，是一种组播报文。组播报文不能跨越网段传输，且不受安全策略控制。添加**remote**参数后，从心跳口发送的各种报文将封装成UDP报文。UDP报文是一种单播报文，只要路由可达就可以跨越网段传输，但需要受到安全策略控制。安全策略的配置方法是允许报文在local区域与心跳口所在安全区域间双向通过。



如上图所示，配置完成后，主备备份方式的双机热备状态形成过程如下（图中序号与下文序号一致）：

1） 双机热备启用之后，FW1的Active组的状态会由Initialize切换成Active，FW2的Standby组的状态由Initialize切换成Standby。

2） 由于FW1的VRRP备份组都加入了Active组，而Active组的初始状态为Active，所以FW1的VRRP备份组1和VRRP备份组2的状态都为Active。同理FW2的VRRP备份组1和VRRP备份组2的状态都为Standby。

3） 这时FW1的VRRP备份组1和2会分别向上行和下行交换机发送免费ARP报文，将VRRP备份组的虚拟MAC地址通知给他们。其中00-00-5E-00-01-01是VRRP备份组1的虚拟MAC地址，00-00-5E-00-01-02是VRRP备份组2的虚拟MAC地址。

4） 上下行的交换机的MAC表项会分别记录虚拟MAC地址与端口Eth0/0/1的对应关系。这样

当上下行的业务报文到达交换机后，交换机会将报文转发到FW1上，所以FW1成为了主用设备，FW2成为了备用设备。

5） 同时FW1的Active组还会通过心跳线定时向FW2的Standby组发送HRP心跳报文。

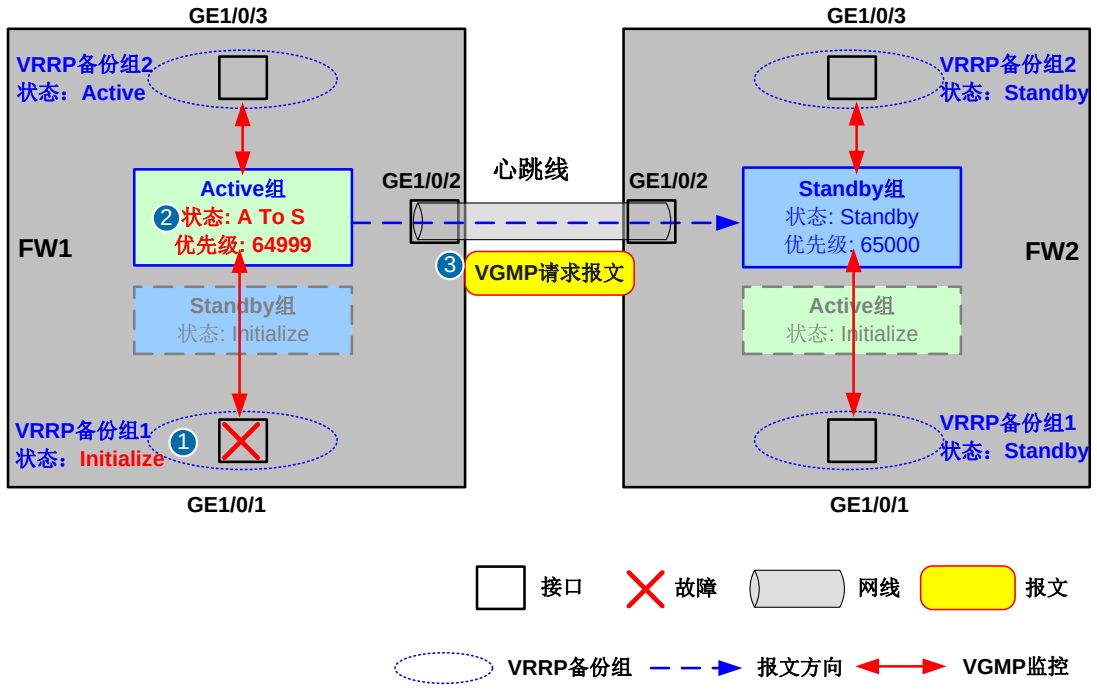
主用设备接口故障后的状态切换过程

两台防火墙形成主备备份状态后，如果主用设备的接口故障，那么两台防火墙会发生主备状态切换，具体过程如下：

1） 如下图所示，当主用设备的接口GE1/0/1故障后，FW1的VRRP备份组1的状态变成Initialize。

2） FW1的Active组会感知到这一变化，将自身的优先级降低2（一个接口故障优先级降低2），并将自身状态切换成Active To Standby（图中简写为A To S）。Active To Standby是一种短暂的中间状态，用户是不可见的。

3） FW1的Active组会向对端发送VGMP请求报文，请求将状态切换成Standby。VGMP请求报文是一种VGMP报文，携带本端VGMP组调整后的优先级64999。

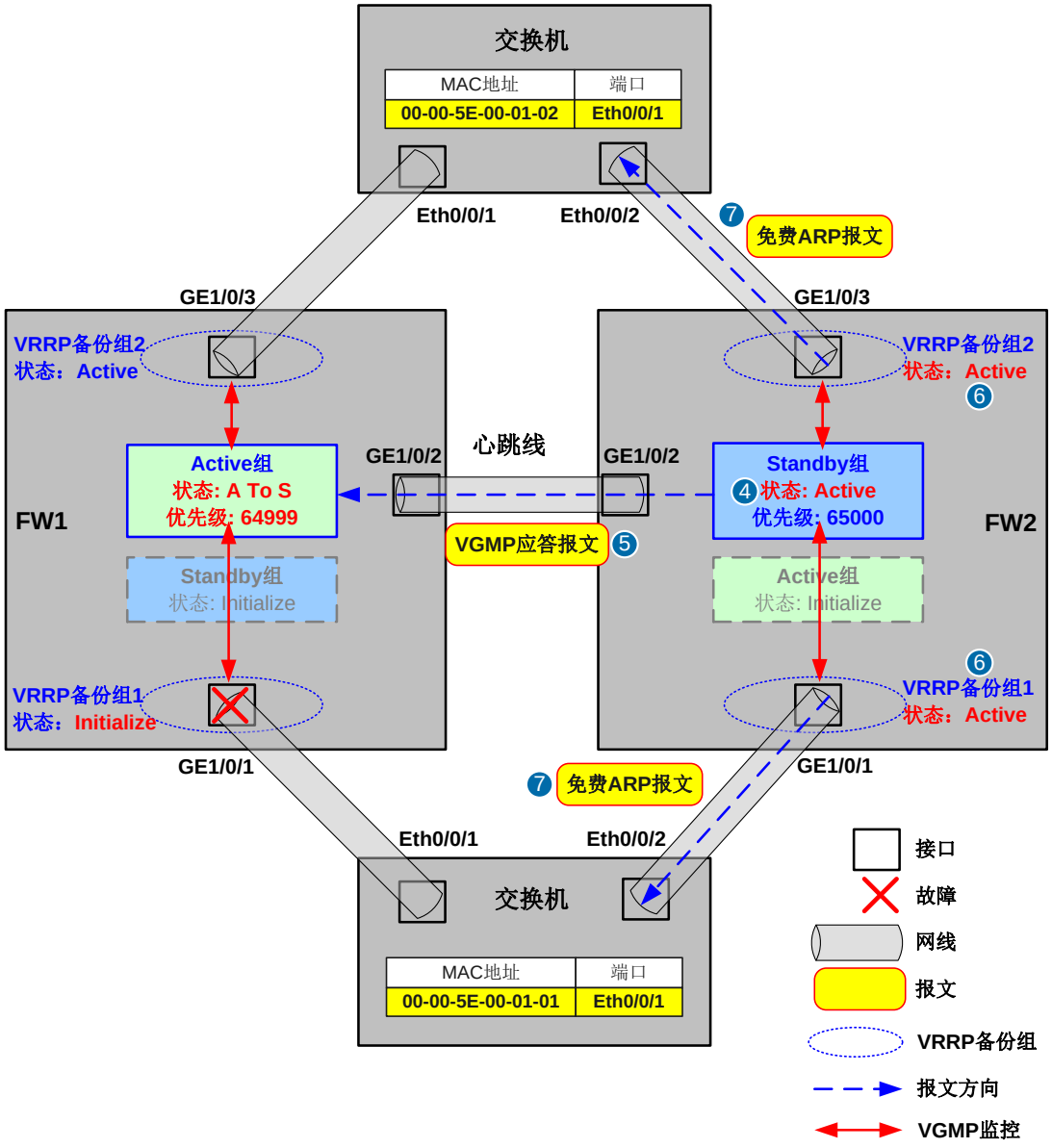


4） 如下图所示，FW2的Standby组收到FW1的Active组的VGMP请求报文后，将会与对端比较VGMP优先级。经过比较后发现本端的优先级65000高于对端的64999，因此FW2的Standby组会将自身状态切换成Active。

5） FW2的Standby组会向对端返回VGMP应答报文，允许对端进行状态切换。

6） 与此同时FW2的Standby组会强制组内的VRRP备份组1和2也将状态切换成Active。

7） FW2的VRRP备份组1和2会分别向下行和上行交换机发送免费ARP报文，更新他们的MAC转发表。



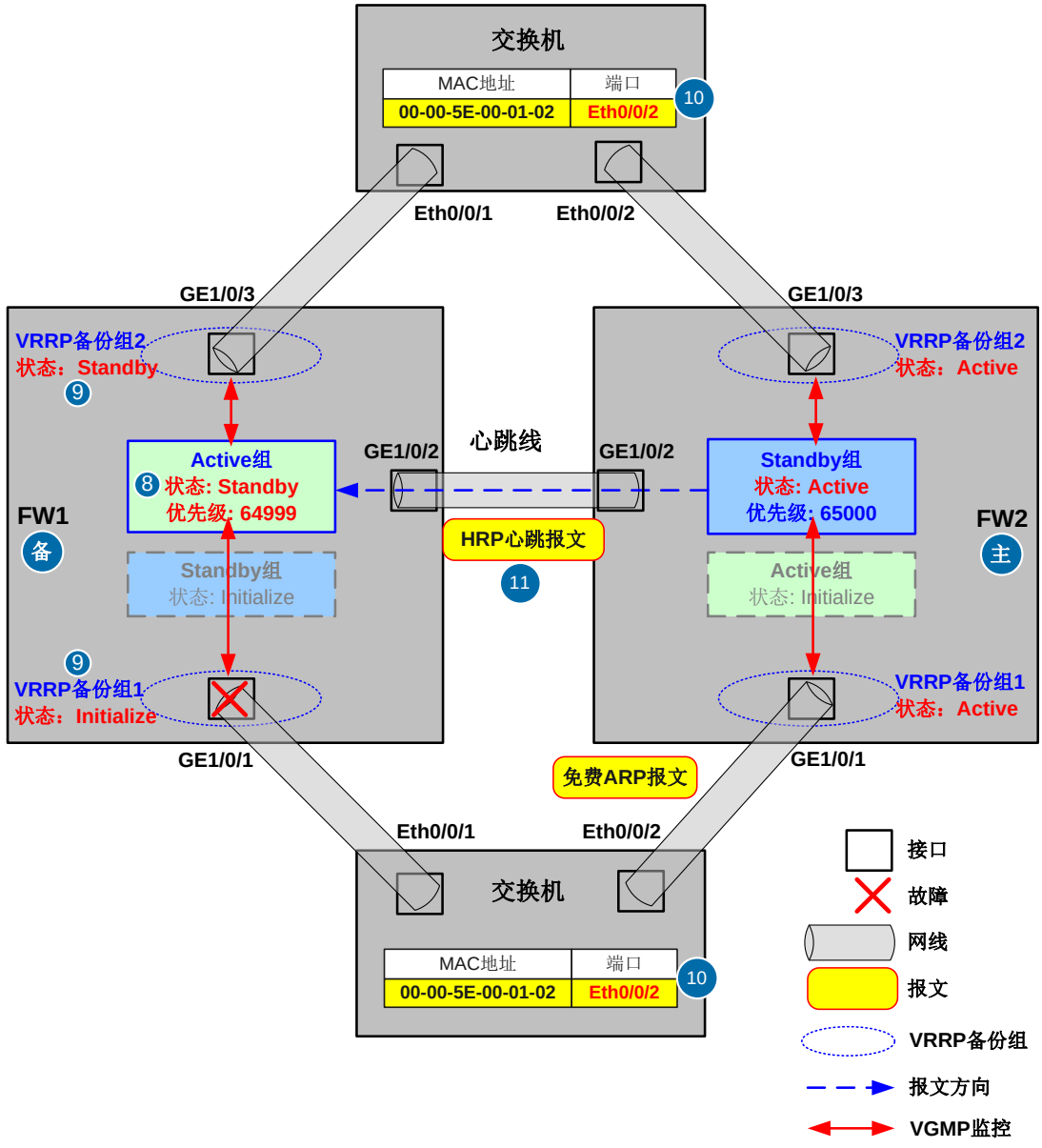
8） 如下图所示，FW1的Active组收到对端的VGMP确认报文后，会将自身状态切换成Standby。

9） FW1的Active组会强制组内的VRRP备份组将状态切换成Standby。由于VRRP备份组1内的接口故障，所以VRRP备份组1的状态为Initialize不变，只有VRRP备份组2的状态切换成Standby。

10）与此同时，上下行交换机收到 FW2 的免费 ARP 报文后会更新 MAC 表项，记录虚拟 MAC 地址与端口 Eth0/0/2 的对应关系。这样当上下行的业务流量到达交换机后，交换机会将

流量转发到 FW2 上。至此两台防火墙的主备状态切换完成，FW2 成为新的主用设备，FW1 成为新的备用设备。

11）主备状态切换完成后，新的主用设备 FW2 会定时向新的备用设备 FW1 发送心跳报文。



主用设备整机故障后的状态切换过程

当主用设备**整机故障**后，主用设备的 VGMP 组将不会再发送 HRP 心跳报文。这时如果备用设备的 VGMP 组连续三次收不到主用设备的 HRP 心跳报文，那么他就会认定对端的 VGMP 组故障，从而将自身切换到主用状态。

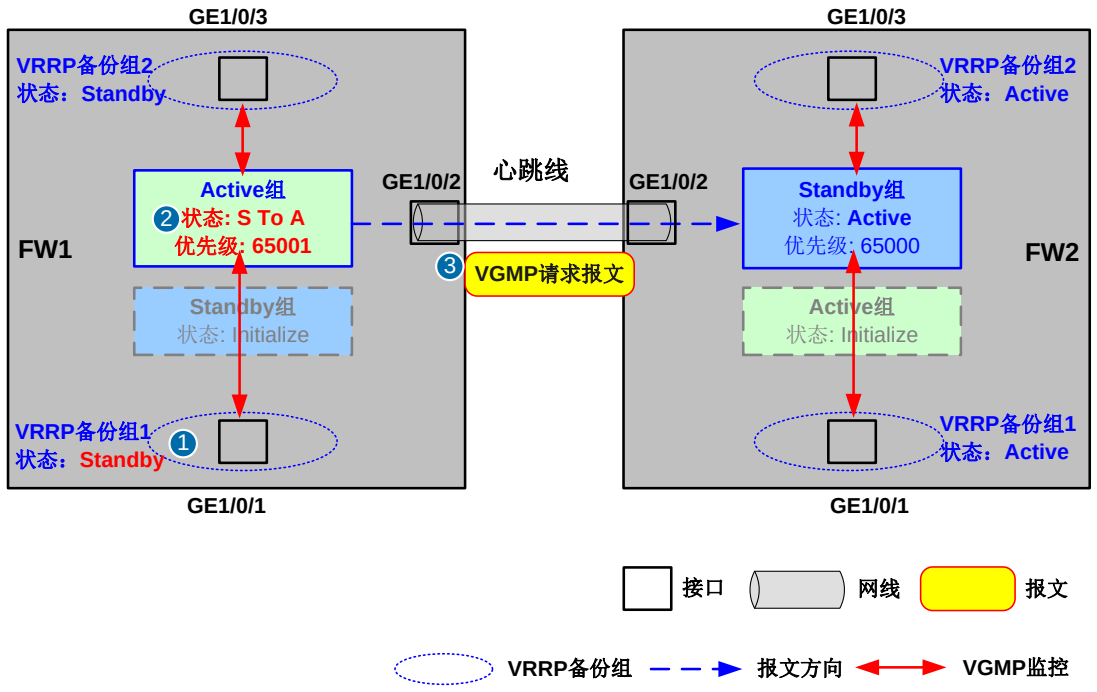
原主用设备故障恢复后的状态切换过程（抢占）

当原主用设备的**故障恢复**后，如果配置了**抢占**功能，那么原主用设备将重新抢占成为主用设备，具体过程如下文所示。如果没有配置抢占功能，则原主用设备依旧保持备份状态。

1） 如下图所示，原主用设备的接口GE1/0/1故障恢复后，VRRP备份组1的状态由Initialize切换成Standby。

2） FW1的Active组感知到这一变化后，会将自身的优先级升高2（一个接口故障恢复优先级升高2），升高到65001。FW1的Active组会与对端比较VGMP优先级，对端的优先级信息是从对端发送的HRP心跳报文中获取的。经过比较后发现本端的优先级65001高于对端的65000。这时如果配置了抢占功能，则会启动抢占延时。抢占延时结束后，FW1的Active组会将状态由Active To Standby切换成Standby To Active（图中简写为S To A）。Standby To Active是一种短暂的中间状态，用户是不可见的。

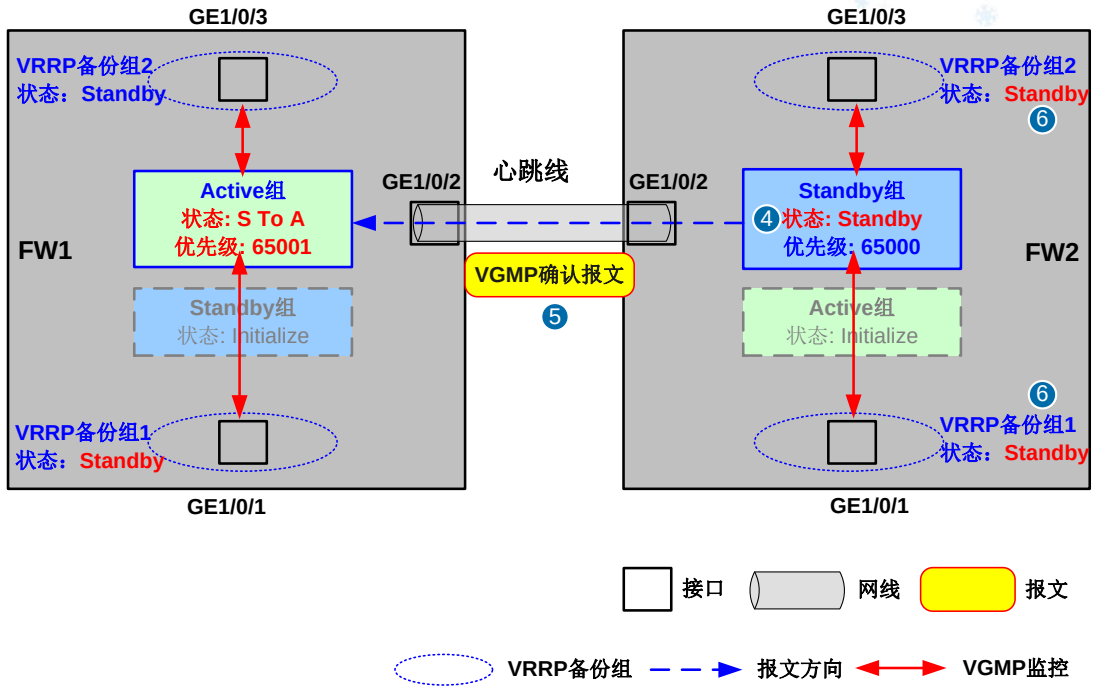
3） FW1的Active组会向对端发送VGMP请求报文，请求将状态切换成Active。VGMP请求报文是一种VGMP报文，携带本端VGMP组调整后的优先级65001。



4） 如下图所示，FW2的Standby组收到FW1的Active组的VGMP请求报文后，将会与对端比较VGMP优先级。经过比较后发现本端的优先级65000低于对端的65001，因此FW2的Standby组会将自身状态切换成Standby。

5） FW2的Standby组会向对端返回VGMP应答报文，允许对端将状态切换成Active。

6） 与此同时FW2的Standby组会强制组内的VRRP备份组1和2也将状态切换成Standby。

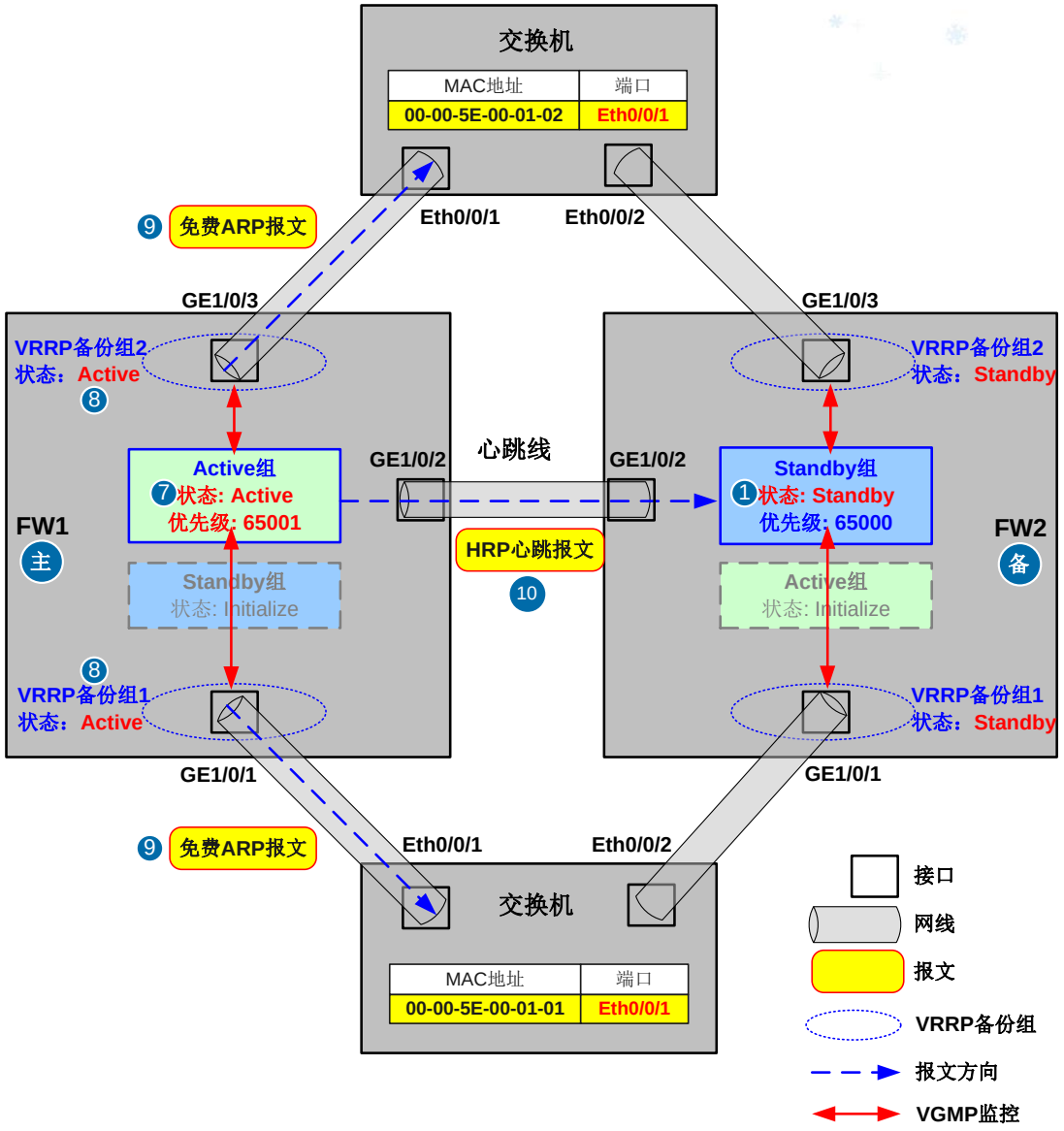


7） 如下图所示，FW1的Active组收到对端的VGMP确认报文后，会将自身状态切换成Active。

8） FW1的Active组会强制组内的VRRP备份组1和2也将状态切换成Active。

9） FW1的VRRP备份组1和2会分别向下行和上行交换机发送免费ARP报文，更新他们的MAC转发表。上下行交换机收到免费ARP报文后会更新MAC表项，记录虚拟MAC地址与端口Eth0/0/1的对应关系。这样当上下行的业务报文到达交换机后，交换机会将报文转发到FW1上。至此两台防火墙的主备状态切换完成，FW1重新抢占成为主用设备，而FW2重新成为备用设备。

10) 主备状态切换完成后，主用设备FW1会定时向备用设备FW2发送心跳报文。



负载分担双机热备状态形成和切换过程

以上我们描述的是主备备份方式的双机热备状态形成和切换过程，下面我们来看负载分担状态的情况。

为了实现负载分担方式的双机热备，我们需要在 FW1 和 FW2 上都启用 Active 组和 Standby 组，使 FW1 的 Active 组与 FW2 的 Standby 组进行通信，构成一组“主备”，FW2 的 Active 组与 FW1 的 Standby 组进行通信，也构成一组“主备”。这样两台 FW 形成互为主备的状态，也就是负载分担状态。

结合下面的图给出负载分担方式的双机热备配置，如下表所示：

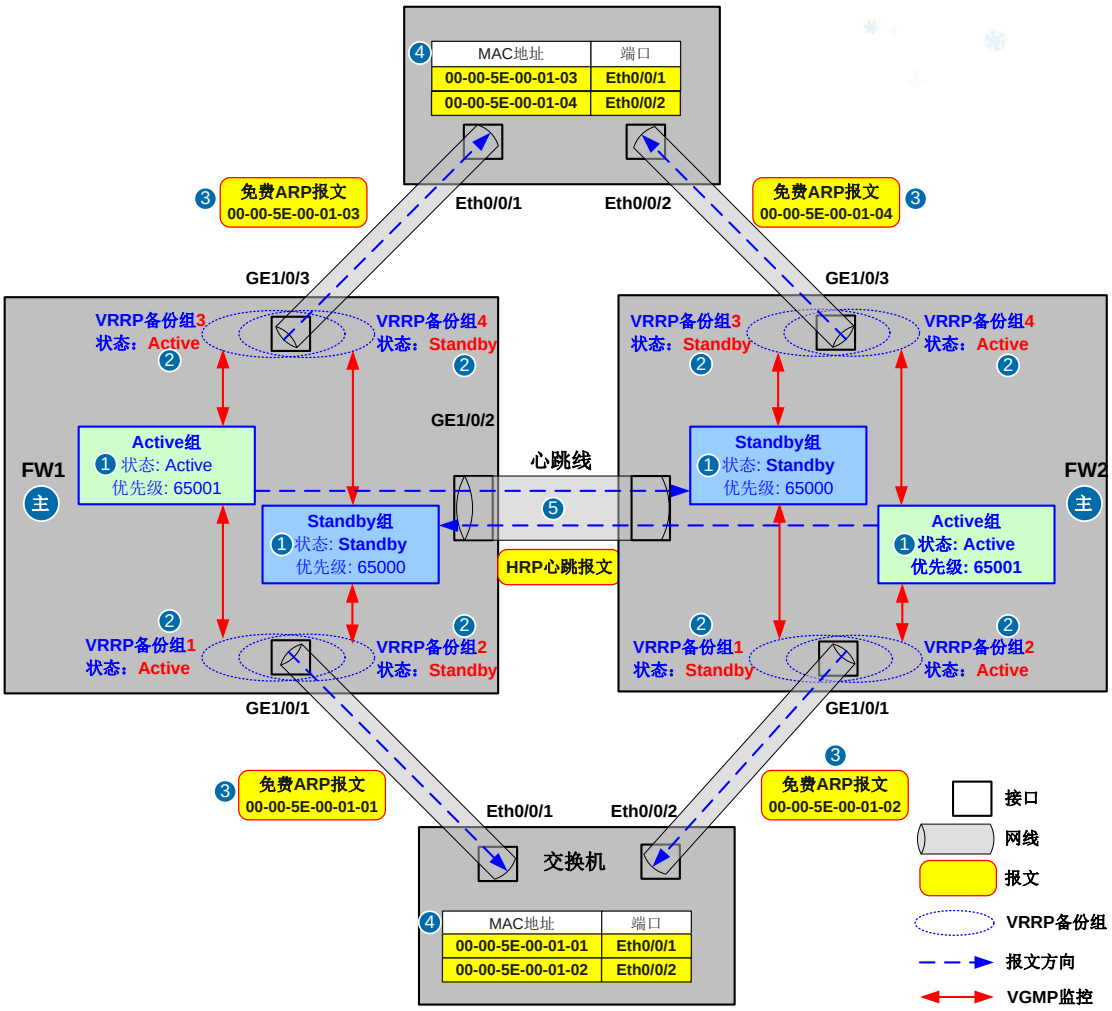
FW1 的配置	FW2 的配置

interface GigabitEthernet 1/0/1
ip address 10.1.1.3 255.255.255.0
vrrp vrid 1 virtual-ip 10.1.1.1 255.255.255.0 active
//将接口 GE1/0/1 加入 VRRP 备份组 1，并将 VRRP 备份组 1 加入 Active 组。
vrrp vrid 2 virtual-ip 10.1.1.2 255.255.255.0 standby
//将接口 GE1/0/1 加入 VRRP 备份组 2，并将 VRRP 备份组 2 加入 Standby 组。
	interface GigabitEthernet 1/0/1
	ip address 10.1.1.4 255.255.255.0
	vrrp vrid 1 virtual-ip 10.1.1.1 255.255.255.0 standby
	//将接口 GE1/0/1 加入 VRRP 备份组 1，并将 VRRP 备份组 1 加入 Standby 组。
	vrrp vrid 2 virtual-ip 10.1.1.2 255.255.255.0 active
	//将接口 GE1/0/1 加入 VRRP 备份组 2，并将 VRRP 备份组 2 加入 Active 组。

interface GigabitEthernet 1/0/3
ip address 1.1.1.3 255.255.255.0
vrrp vrid 3 virtual-ip 1.1.1.1 255.255.255.0 active
//将接口 GE1/0/3 加入 VRRP 备份组 3，并将 VRRP 备份组 3 加入 Active 组。
vrrp vrid 4 virtual-ip 1.1.1.2 255.255.255.0 standby
//将接口 GE1/0/3 加入 VRRP 备份组 4，并将 VRRP 备份组 4 加入 Standby 组。
	interface GigabitEthernet 1/0/3
	ip address 1.1.1.4 255.255.255.0
	vrrp vrid 3 virtual-ip 1.1.1.1 255.255.255.0 standby
	//将接口 GE1/0/3 加入 VRRP 备份组 3，并将 VRRP 备份组 3 加入 Standby 组。
	vrrp vrid 4 virtual-ip 1.1.1.2 255.255.255.0 active
	//将接口 GE1/0/3 加入 VRRP 备份组 4，并将 VRRP 备份组 4 加入 Active 组。

hrp interface GigabitEthernet 1/0/2 //指定心跳口 **hrp enable** //启用双机热备功能	**hrp interface GigabitEthernet 1/0/2** //指定心跳口 **hrp enable** //启用双机热备功能

由上表可以看到：1、负载分担场景下，每个业务接口需要加入两个 VRRP 备份组，且这两个VRRP 备份组要分别加入Active 组和Standby 组。例如GE1/0/1 接口加入了备份组1和2，备份组 1 和 2 分别加入了 Active 组和 Standby 组。2、两台防火墙的相同编号的 VRRP 备份组需分别加入 Active 组和 Standby 组。例如 FW1 的 VRRP 备份组 1 加入了 Active 组，FW2 的 VRRP 备份组 1 加入了 Standby 组。



如上图所示，配置完成后，负载分担方式的双机热备状态形成过程如下：

1) FW1和FW2的Active组的状态会由Initialize切换成Active，Standby组的状态由Initialize切换成Standby。

2) 由于FW1的VRRP备份组1和3加入了Active组，而Active组的初始状态为Active，所以FW1的VRRP备份组1和VRRP备份组3的状态都为Active；由于FW1的VRRP备份组2和4加入了Standby组，而Standby组的初始状态为Standby，所以FW1的VRRP备份组2和VRRP备份组4的状态都为Standby；同理FW2的VRRP备份组1和VRRP备份组3的状态都为Standby，VRRP备份组4和VRRP备份组4的状态都为Active。

3) 这时FW1的VRRP备份组1和3会分别向下行和上行交换机发送送免费ARP报文，将VRRP备份组1和3的虚拟MAC地址通知给他们；FW2的VRRP备份组2和4会分别向下行和上行交换机发送送免费ARP报文，将VRRP备份组2和4的虚拟MAC地址通知给他们。

4) 下行交换机的MAC表项会记录VRRP备份组1的虚拟MAC地址（00-00-5E-00-01-01）与端口Eth0/0/1的对应关系，VRRP备份组2的虚拟MAC地址（00-00-5E-00-01-02）与端口端口Eth0/0/2的对应关系。这样当下行交换机下面的设备请求VRRP备份组1的虚拟MAC

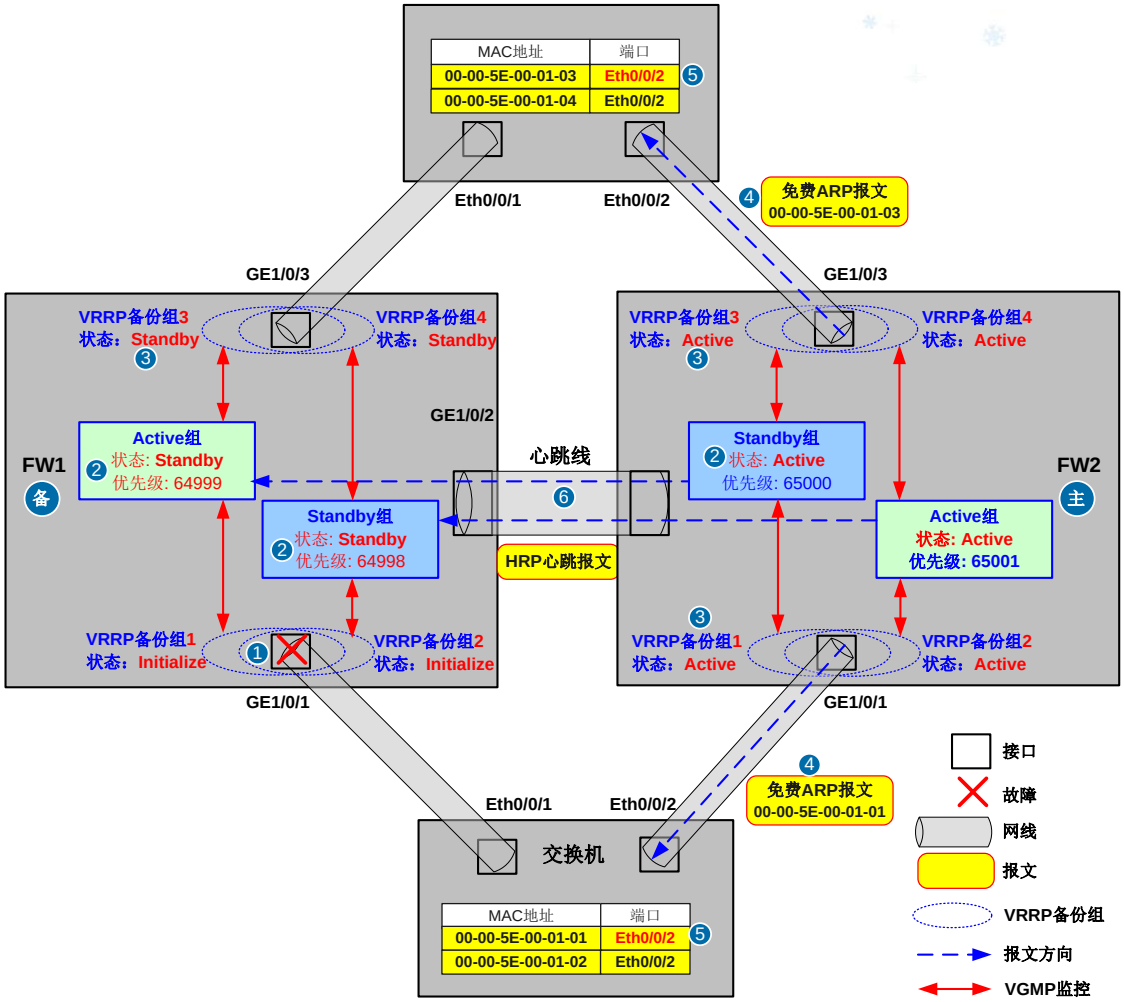
地址（将VRRP备份组1的虚拟IP设置为下一跳）时，交换机会将业务报文转发到FW1上；而请求VRRP备份组2的虚拟MAC地址（将VRRP备份组2的虚拟IP设置为下一跳）时，交换机会将业务报文转发到FW2上。同理上行交换机上面的设备请求VRRP备份组3的虚拟MAC地址（将VRRP备份组3的虚拟IP设置为下一跳）时，交换机会将业务报文转发到FW1上；而请求VRRP备份组4的虚拟MAC地址（将VRRP备份组4的虚拟IP设置为下一跳）时，交换机会将业务报文转发到FW2上。这样FW1和FW2都会转发业务报文，所以FW1和FW2都是主用设备，形成负载分担状态。

5) 负载分担状态形成后，FW1的Active组会定期向FW2的Standby组发送HRP心跳报文，FW2的Active组会定期向FW1的Standby组发送HRP心跳报文。

两台防火墙形成负载分担方式的双机热备后，如果其中一台防火墙的接口故障，那么他们将切换成主备备份状态，具体过程如下：

1) 如下图所示，当FW1的GE1/0/1接口故障时，FW1的VRRP备份组1和2的状态都会变成Initialize。
2) FW1的Active组与Standby组的优先级都会降低2。这时FW1的Active组的优先级变成64999低于FW2的Standby组的优先级65000， FW2的Standby组的优先级变成64998低于FW2的Standby组的优先级65000。这样经过VGMP组之间的状态协商后（具体原理和报文交互请看上面的主备备份方式），FW1的Active组的状态切换成Standby，FW2的Standby组的状态切换成Active。
3) FW1的Active组和FW2的Standby组会强制组内VRRP备份组也进行状态切换，所以FW2的VRRP备份组1和3的状态都切换成Active。
4) FW2的VRRP备份组1和3会分别向下行和上行交换机发送免费ARP报文，更新他们的MAC转发表。
5) 下行交换机收到免费ARP报文后，会更新自身的MAC转发表，将VRRP备份组1的虚拟MAC地址（00-00-5E-00-01-01）修改成与Eth0/0/2对应。同理上行交换机会将VRRP备份组3的虚拟MAC地址（00-00-5E-00-01-03）修改成与Eth0/0/2对应。这样当上下行的业务报文到达交换机后，交换机会将报文都转发到FW2上。至此双机热备状态切换完成，时FW1成为备用设备，FW2成为主用设备，负载分担状态变成主备备份状态。

6) 负载分担切换成主备备份状态后，主用设备FW2会定时向备用设备FW1发送心跳报文。



总结

上面的内容应该可以完美地回答“两台防火墙的 VGMP 组的状态协商和切换过程，以及报文交互过程到底是怎样的呢？”这个问题了。而由上面的例子可以看出，VGMP 在双机热备中主要实现以下三个功能：

- **故障监控：**VGMP组能够监控VRRP备份组状态变化，从而感知到VRRP组内接口的故障和恢复。（私下里嘀咕：VGMP组能不能直接监控接口故障呢，一定要通过VRRP备份组监控接口么？）
- **状态切换：**VGMP的状态切换过程也就是设备主备状态切换的过程。VGMP组感知到VRRP备份组状态变化后，会调整自身的优先级，并与对端的VGMP组重新协商主备状态。（私下里嘀咕：这一点比较清楚了，本篇都是在讲状态如何切换和协商的。）
- **流量引导：**两个VGMP组主备状态建立或者切换后，会强制组内VRRP备份组状态统一切换，然后由状态为Active的VRRP备份组发送免费ARP来引导流量通过自身转发，也就是

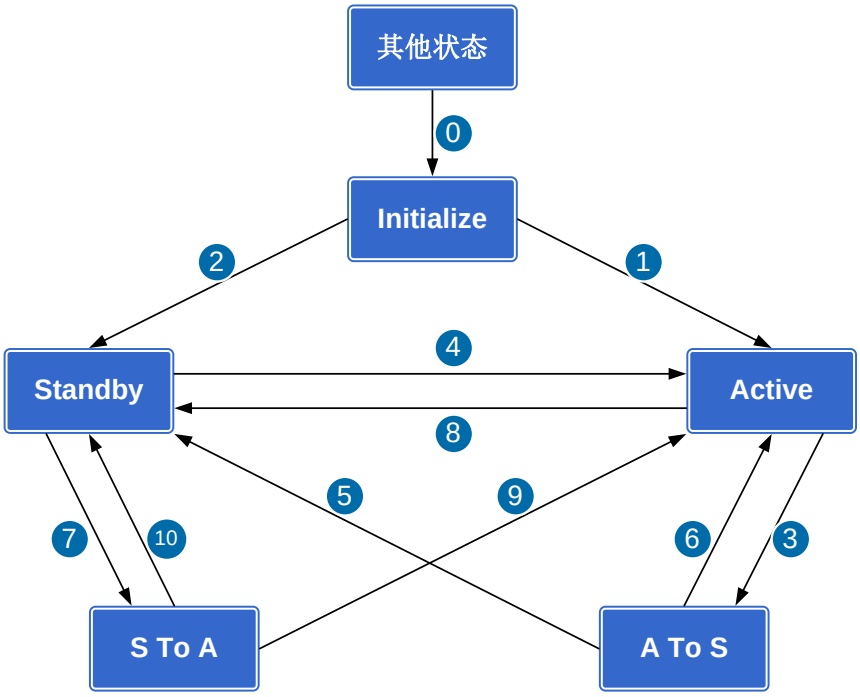
通过主用设备转发。（私下里嘀咕：如果VGMP组能够直接监控接口的话，流量引导方式也是这样么？）

实际上 VGMP 功能很强大，通过监控 VRRP 备份组状态实现防火墙故障监控和流量引导仅仅是 VGMP 的一个招式。这个招式仅仅适用于防火墙上行或下行设备是交换机的场景，因为 VRRP 本身就是为这个场景量身定制的。当防火墙上行或下行设备是路由器的时候，VGMP 就无力应对了么？当然不会的！下节我们就对 VGMP 的更多绝招进行介绍，帮助大家全面了解双机热备功能，做到兵来将挡、水来土掩！

本节附录：VGMP 状态机

前面我们学习了 VGMP 组的各种状态变化过程。下面强叔再通过解释 VGMP 状态机的形式，来帮助各位小伙伴加深对 VGMP 状态变化的理解。难道你不觉得了解了某个功能的状态机后，整个人都变得高大上了吗？！

说明：本篇的 VGMP 状态机目前适用于 USG2000/5000/6000 系列防火墙和 USG9000 系列防火墙的 V1R3 版本。



0) 启用双机热备功能后，各VGMP组进入Initialize（初始化）状态。

1) 启用Active组后，Active组的状态由Initialize切换成Active。

2) 启用Standby组后，Standby组的状态由Initialize切换成Standby。

3) 本端VGMP组监控的接口故障时，状态由Active切换成ActiveToStandby，并发送VGMP请求报文给对端设备的VGMP组。

4) 本端VGMP组收到对端的VGMP请求报文，发现自身优先级高，则将状态由Standby切换成Acitve，并发送VGMP确认报文给对端设备的VGMP组。

5) 本端VGMP组收到对端的VGMP确认报文，确认本端需要进行状态切换，则本端的VGMP组状态由ActiveToStandby切换成Standby。

6) 对端VGMP组确认本端的VGMP组不需要进行状态切换或连续三次没有回应本端的VGMP请报文，则本端的VGMP组状态由ActiveToStandby切换成Active。

7) 本端VGMP组监控的接口故障恢复后,如果本端VGMP组优先级高于对端且配置了抢占功能，则本端VGMP组状态由Standby切换成StandbyToAcitve，并向对端发送VGMP请求报文。

8) 本端VGMP组收到对端的VGMP请求报文，发现对端优先级高，则将状态由Active切换成Standby，并发送VGMP确认报文给对端设备的VGMP组。

9) 本端VGMP组收到对端的VGMP确认报文，确认本端需要进行状态切换，则本端的VGMP组状态由StandbyToAcitve切换成Active，完成抢占过程。

10) 对端VGMP组确认本端的VGMP组不需要进行状态切换或连续三次没有回应本端的VGMP请报文，则本端的VGMP组状态由StandbyToAcitve切换成Standby。

VGMP 招式详解

上篇我们详细介绍了 VGMP 组如何通过 VRRP 备份组来实现接口故障监控、设备状态切换以及上下行流量引导。我们还发现 VGMP 与 VRRP 的配合只适用于防火墙连接二层设备的组网。那么当防火墙连接路由器或防火墙透明接入网络（业务接口工作在二层）时，VGMP 组是使用什么招式来应对的呢？本篇强叔将为您揭秘 VGMP 组的其余保留招式~

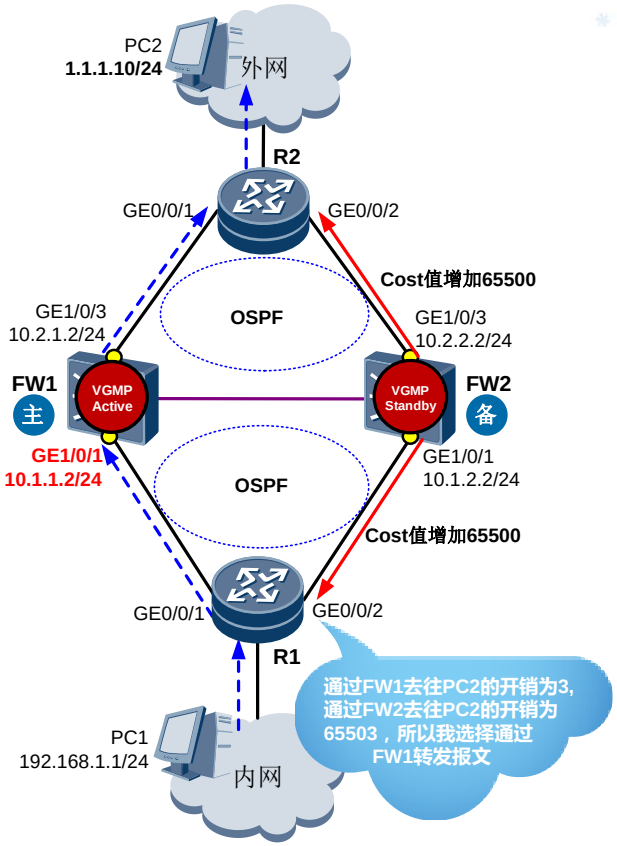
防火墙连接路由器时的 VGMP 招式

如下图所示，两台防火墙上下行业务接口工作在三层，连接路由器。防火墙与路由器之间运行 OSPF 协议。由于上下行不是二层交换机，所以 VGMP 组无法使用 VRRP 备份组。这时 VGMP 组使用的故障监控招式是**直接监控接口状态。**方法是直接将接口加入 VGMP 组。当 VGMP 组中的接口故障时，VGMP 组会直接感知到接口状态变化，从而降低自身的优先级。VGMP 组直接监控接口状态的配置步骤如下（以主备备份方式的双机热备为例）：

FW1 的配置	FW2 的配置
interface GigabitEthernet 1/0/1 ip address 10.1.1.2 255.255.255.0 hrp track active //将接口 GE1/0/1 加入 Active 组。	interface GigabitEthernet 1/0/1 ip address 10.1.2.2 255.255.255.0 hrp track standby //将接口 GE1/0/1 加入 Standby 组。
interface GigabitEthernet 1/0/3 ip address 10.2.1.2 255.255.255.0 hrp track active //将接口 GE1/0/3 加入 Active 组。	interface GigabitEthernet 1/0/3 ip address 10.2.2.2 255.255.255.0 hrp track standby //将接口 GE1/0/3 加入 Standby 组。
hrp interface GigabitEthernet 1/0/2 //指定心跳口 hrp enable //启用双机热备功能	hrp interface GigabitEthernet 1/0/2 //指定心跳口 hrp enable //启用双机热备功能

说明：如果是负载分担方式的双机热备，则需要在每个业务接口上执行 hrp track active 和 hrp track standby，将业务接口分别加入 Active 组和 Standby 组。

【强叔问答】看到这里好奇的小伙伴们或许会问：不是将接口加入 VGMP 组，使 VGMP 组监控接口状态吗？为什么命令行是 **hrp track** 而不是 **vgmp track** 呢？这是因为上节讲到 VGMP 和 HRP 的报文都是由 VRRP 头和 VGMP 头封装的，区别只在于 HRP 报文还需要再封装一个 HRP 报文头。所以当初开发者设计命令时就统一使用了 hrp 这个参数，并流传至今。



在R1上的路由表上可见，去往目的网段1.1.1.0的报文的下一跳为FW1的GE1/0/1的地址10.1.1.2。

```
[R1]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 11        Routes : 11

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.0/24  OSPF    10   3           D   10.1.1.2        GigabitEthernet0/0/1
```

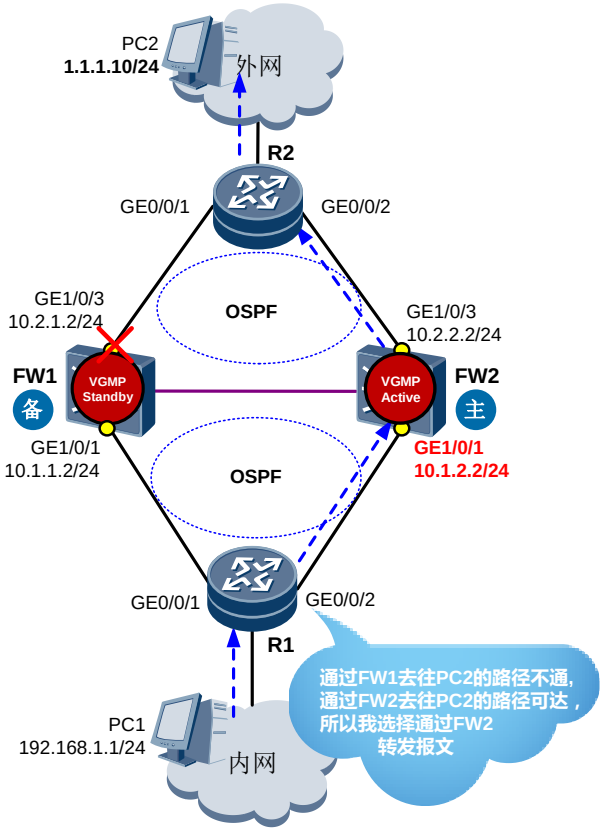
如上图所示，配置完成后，FW1 的 VGMP 组状态为 Active，FW1 成为主用设备；FW2 的 VGMP 组状态为 Standby，FW2 成为备用设备。在双机热备的第一篇中讲到，如果我们希望 PC1 访问 PC2 的流量通过 FW1 转发，那么我们就需要手工将 FW2 所在链路（R1—>FW2—>R2）的 OSPF Cost 值调大。但是如果上下行的路由器 R1 或 R2 我们不方便或不能配置时怎么办呢？这就需要用到防火墙 VGMP 组的流量引导功能，将流量自动引导到主用设备上来。这种组网采用的 VGMP 流量引导招式为**通过自动调整 Cost 值实现流量引导，即防火墙会根据 VGMP 组的状态自动调整 OSPF 的 Cost 值**（命令为 hrp ospf-cost adjust-enable）。启用此功能后，如果防火墙上存在状态为 Active 的 VGMP 组，则防火墙会正常对外发布路由；如果防火墙上的 VGMP 组状态都为 Standby，则防火墙会在发布路由时将 Cost 值增加 65500（此为缺省值，可调整）。

说明：如果是负载分担组网，由于两台防火墙上都存在状态为 Active 的 VGMP 组，所以都

会正常对外发布路由。

例如上图中的主用 FW1（VGMP 组状态为 Active）会正常对外发布路由，备用设备 FW2（VGMP 组状态为 Standby）会在对上下行设备发布路由时将 Cost 值增加 65500。这样在 R1 上来看，通过 FW1 去往 PC2 的 OSPF Cost 值为 1+1+1=3，通过 FW2 去往 PC2 的 OSPF Cost 值为 65501+1+1=65503。因为路由器在转发流量时会选择开销（Cost 值）更小的路径（R1—>FW1—>R2），所以内网 PC1 访问外网 PC2 的流量会通过主用设备 FW1 转发。

如下图所示，当 FW1 的业务接口故障后，两台防火墙的 VGMP 组会进行状态切换，具体切换过程请参见上篇的“主用设备接口故障后的状态切换过程”。状态切换后，FW2 的 VGMP 组状态切换成 Active，FW2 成为主用设备；FW1 的 VGMP 组状态切换成 Standby，FW1 成为备用设备。这时 FW2 正常对外发布路由，FW1 发布的路由 Cost 值增加 65500。而在 R1 上来看，通过 FW1 去往 PC2 的路径不通（因为接口故障），通过 FW2 去往 PC2 的路径可达且 Cost 值为 3.，所以内网 PC1 访问外网 PC2 的流量会通过新的主用设备 FW2 转发。



在R1上的路由表上可见，去往目的网段1.1.1.0的报文的下一跳为FW2的GE1/0/1的地址10.1.2.2。

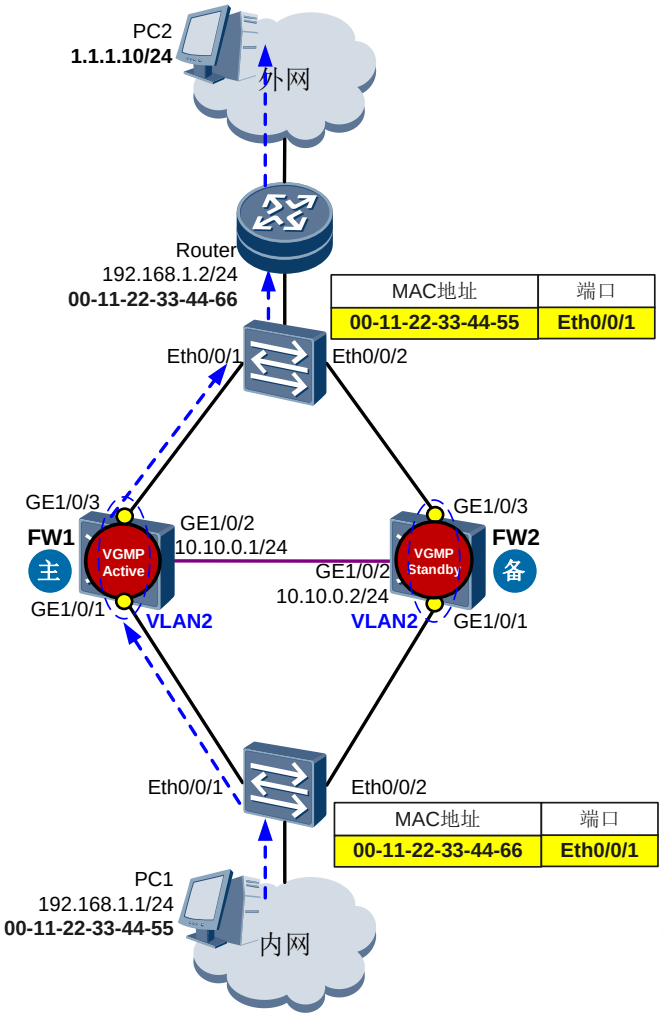
```
[R1]display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 11        Routes : 11

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        1.1.1.0/24  OSPF    10   3           D   10.1.2.2        GigabitEthernet0/0/2
```

防火墙透明接入，连接交换机时的 VGMP 招式

如下图所示，两台防火墙上下行业务接口都工作在二层，连接交换机。由于防火墙的业务接口工作在二层，没有 IP 地址，所以 VGMP 组无法使用 VRRP 备份组或者直接监控接口的状态。这时 VGMP 组使用的故障监控招式是**通过 VLAN 监控接口状态。**方法是将二层业务接口加入 VLAN，VGMP 组监控 VLAN。当 VGMP 组中的接口故障时，VGMP 组会通过 VLAN 感知到其中接口状态变化，从而降低自身的优先级。

VGMP 组通过 VLAN 监控接口状态的配置步骤如下：

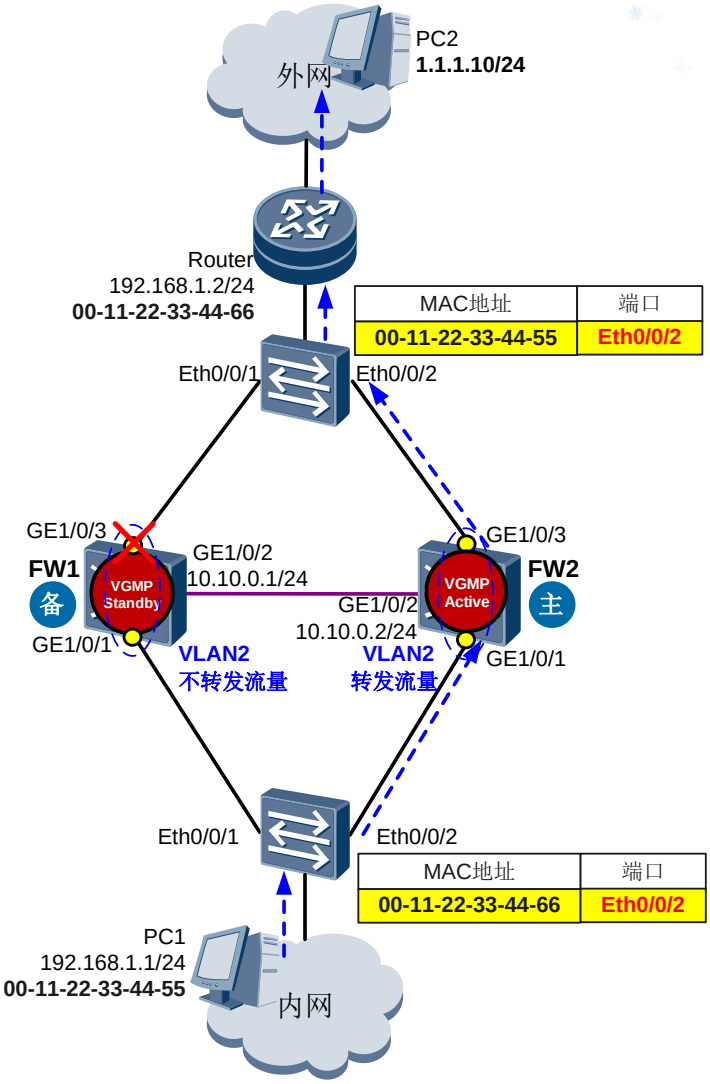
FW1 的配置	FW2 的配置
vlan 2 **port GigabitEthernet 1/0/1** **port GigabitEthernet 1/0/3** //将二层业务接口加入 VLAN2。 **hrp track active** //将 VLAN2 加入 Active 组，由 Active 组监控 VLAN2。	**vlan 2** **port GigabitEthernet 1/0/1** **port GigabitEthernet 1/0/3** //将二层业务接口加入 VLAN2。 **hrp track standby** //将 VLAN2 加入 Active 组，由 Standby 组监控 VLAN2。
hrp interface GigabitEthernet 1/0/2 //指定心跳口 **hrp enable** //启用双机热备功能	**hrp interface GigabitEthernet 1/0/2** //指定心跳口 **hrp enable** //启用双机热备功能



如上图所示，配置完成后，FW1 的 VGMP 组状态为 Active，FW1 成为主用设备；FW2 的 VGMP 组状态为 Standby，FW2 成为备用设备。由于防火墙的业务接口工作在二层，防火墙本身不能运行 OSPF 协议，因此 VGMP 组无法通过控制 OSPF Cost 值的变化来引导上下行流量。这时 VGMP 可以**通过控制 VLAN 是否转发流量**的招式来保证流量引导到主用设备上。当 VGMP 组状态为 Active 时，组内的 VLAN 能够转发流量；当 VGMP 组状态为 Standby 时，组内的 VLAN 被禁用，不能转发流量。VGMP 控制 VLAN 是否转发流量不需要单独配置，只需要按照上表将 VLAN 加入 VGMP 组即可。例如上图中的主用设备 FW1（VGMP 组状态为 Active）的 VLAN 被启用，能够转发流量。备用设备 FW2（VGMP 组状态为 Standby）的 VLAN 被禁用，不能转发流量。因此 PC1 访问 PC2 的流量都从主用设备 FW1 转发。

如下图所示，当 FW1 的业务接口故障后，两台防火墙的 VGMP 组会进行状态切换，具体切换过程请参见上篇的“主用设备接口故障后的状态切换过程”。**当 FW1 的 VGMP 组状态由 Active 切换到 Standby 时，组内 VLAN 中的所有非故障接口都会 Down 然后 Up 一次。这会导致上下行交换机更新自身 MAC 转发表，将目的 MAC 地址改成与端口 Eth0/0/2 的映射，从而将流量引导到 FW2 上。**这时由于 FW2 的 VGMP 组状态已经由 Standby 切换成 Active，所以 FW2 的 VLAN2 能够正常转发流量。

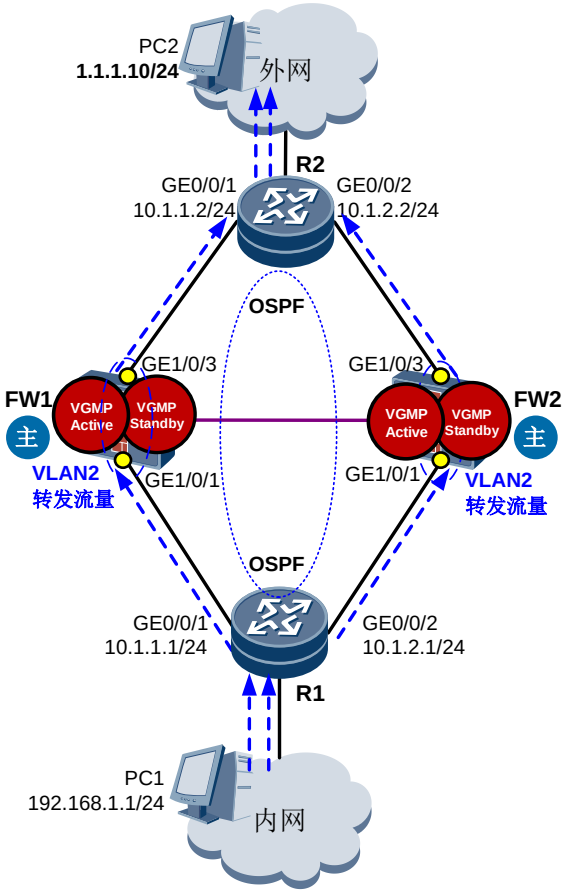
说明：当防火墙的业务接口工作在二层，连接二层设备时，不支持负载分担方式的双机热备。因为如果工作于负载分担方式，则两台设备上的 VLAN 都被启用，都能够转发流量，整个网络就会形成环路。



防火墙透明接入，连接路由器时的 VGMP 招式

如下图所示，两台防火墙上下行业务接口都工作在二层，连接路由器。两台路由器之间运行 OSPF。在此种组网中防火墙的 VGMP 组采用的故障监控和流量引导方式与上一种组网相同，即**通过 VLAN 监控接口状态实现故障监控，通过控制 VLAN 是否转发流量实现流量引导。故障监控的区别之处在于此种组网只支持负载分担方式的双机热备，不支持主备备份方式。**因为如果工作于主备备份方式，备用设备上的 VLAN 被禁用，它的上下行路由器就无法进行通信，无法建立 OSPF 路由。这样当主备切换时，新的主用设备（原备用设备）的 VLAN 被启用，它的上下行路由器才开始新建 OSPF 路由。而 OSPF 路由的新建是需要一定时间的，所以会导致业务的暂时中断。此组网的双机热备配置步骤如下：

FW1 的配置	FW2 的配置
vlan 2 **port GigabitEthernet 1/0/1** **port GigabitEthernet 1/0/3** //将二层业务接口加入 VLAN2。 **hrp track active** //将 VLAN2 加入 Active 组，由 Active 组监控 VLAN2。 **hrp track standby** //将 VLAN2 加入 Standby 组，由 Standby 组监控 VLAN2。	**vlan 2** **port GigabitEthernet 1/0/1** **port GigabitEthernet 1/0/3** //将二层业务接口加入 VLAN2。 **hrp track active** //将 VLAN2 加入 Active 组，由 Active 组监控 VLAN2。 **hrp track standby** //将 VLAN2 加入 Standby 组，由 Standby 组监控 VLAN2。
hrp interface GigabitEthernet 1/0/2 //指定心跳口 **hrp enable** //启用双机热备功能	**hrp interface GigabitEthernet 1/0/2** //指定心跳口 **hrp enable** //启用双机热备功能



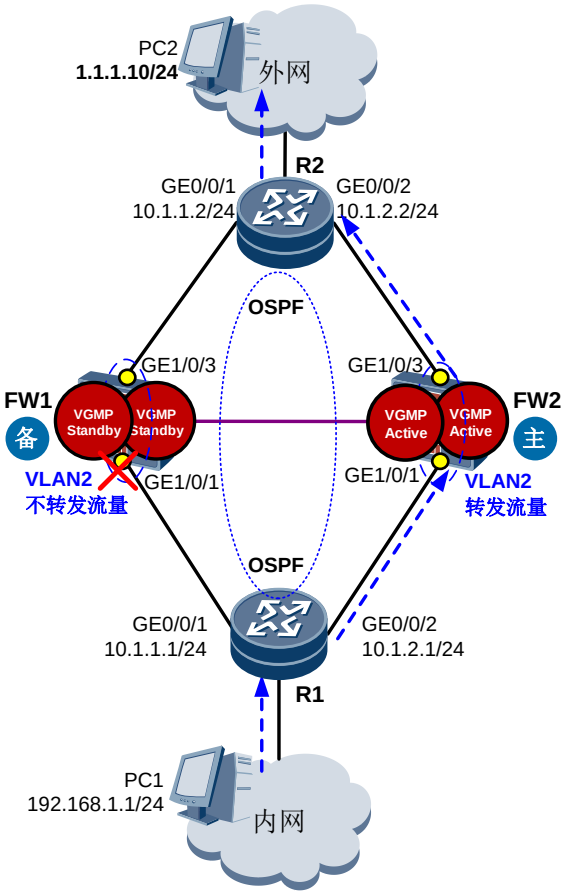
在R1上的路由表上可见，去往目的网段1.1.1.0的报文的下一跳为R2的GE0/0/1的地址10.1.1.2和GE0/0/2的地址10.1.2.2。

```
<R1>display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 14        Routes : 15

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

      1.1.1.0/24    OSPF    10   2           D   10.1.1.2        GigabitEthernet0/0/1
                    OSPF    10   2           D   10.1.2.2        GigabitEthernet0/0/2
```

如上图所示，配置完成后，由于 FW1 和 FW2 上都存在状态为 Active 的 VGMP 组，所以 FW1

和 FW2 都是主用设备，他们的 VLAN2 都转发流量。这时在 R1 的路由表上也可以看到去往 PC2 的流量可以分别通过 FW1 和 FW2 转发。

如下图所示，当 FW1 的业务接口故障后，两台防火墙的 VGMP 组会进行状态切换，双机热备状态也会由负载分担变成主备备份，具体切换过程请参见上篇的“负载分担双机热备状态形成和切换过程”。**当 FW1 的 VGMP 组状态由 Active 切换到 Standby 时，组内 VLAN 中的所有接口都会 Down 然后 Up 一次。这会导致上下行路由器的路由变化并收敛（从下图的路由表中可以看到），从而将流量引导到 FW2 上。**这时由于 FW2 的 VGMP 组状态已经由 Standby 切换成 Active，所以 FW2 的 VLAN2 能够正常转发流量。



在R1上的路由表上可见，去往目的网段1.1.1.0的报文的下一跳变为R2的GE0/0/2的地址10.1.2.2。

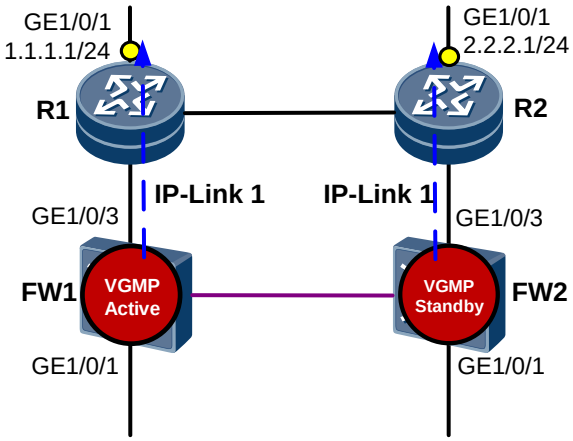
```
<R1>display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public
         Destinations : 10       Routes : 11

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

      1.1.1.0/24    OSPF    10   2           D   10.1.2.2        GigabitEthernet0/0/2
     10.1.2.0/24    Direct  0    0           D   10.1.2.1        GigabitEthernet0/0/2
```

VGMP 组监控远端接口的招式

上面描述的是 VGMP 组应对各种双机热备组网的招式，其中 VGMP 组监控的是防火墙本身的接口。下面我们再来学习两个 VGMP 组监控远端接口的招式。远端接口是指链路上其他设备的接口。需要注意的是这两种 VGMP 监控远端接口的招式只能用于防火墙业务接口工作在三层的组网，因为只有业务接口工作在三层才有 IP 地址，才能对远端设备发送 IP-Link 和 BFD 的探测报文。

- **通过IP-Link监控远端接口状态：**方法是建立IP-Link探测远端接口， 然后VGMP组监控IP-Link状态。当IP-Link探测的接口故障时，IP-Link的状态变成Down，VGMP组感知到IP-Link的状态变化，从而降低自身的优先级。

  如下图所示，我们需要在FW1上使用IP-Link1探测R1的GE1/0/1接口（非直连的远端接口），然后将IP-Link1加入Active组，由Active组监控IP-Link1的状态。

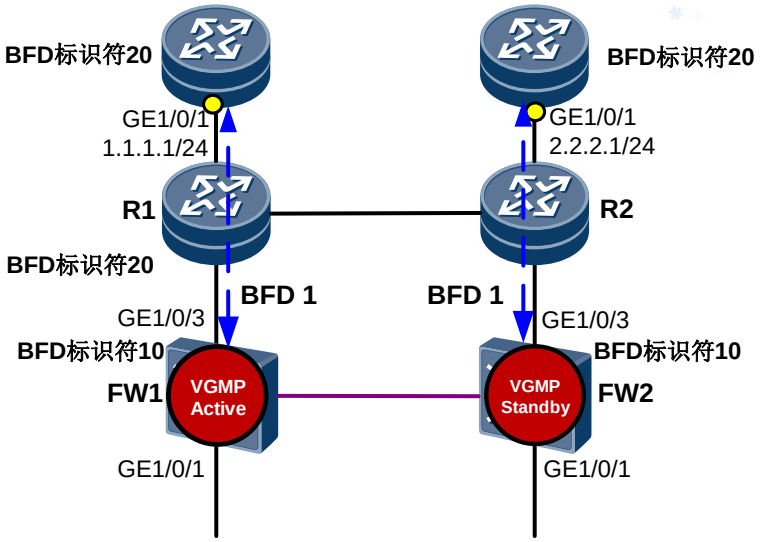


具体配置如下表所示（配置的前提条件是已配置完成双机热备功能）：

FW1 的配置	FW2 的配置
ip-link check enable //启用 IP-Link 功能 ip-link 1 destination 1.1.1.1 interface GigabitEthernet1/0/3 mode icmp //建立 IP-Link1 探测 1.1.1.1 hrp track ip-link 1 active //将 IP-Link1 加入 Active 组	ip-link check enable //启用 IP-Link 功能 ip-link 1 destination 2.2.2.1 interface GigabitEthernet1/0/3 mode icmp //建立 IP-Link1 探测 2.2.2.1 hrp track ip-link 1 standby //将 IP-Link1 加入 Standby 组

- **通过BFD监控远端接口状态：**方法是通过BFD探测远端接口，VGMP组监控BFD状态。当BFD探测的远端接口故障时，BFD的状态变成Down，VGMP组感知到BFD的状态变化，从而降低自身的优先级。

  如下图所示，我们需要在FW1上使用BFD会话10探测R1的GE1/0/1接口（非直连的远端接口），然后将BFD会话1加入Active组，由Active组监控BFD会话1的状态。



具体配置如下表所示（配置的前提条件是已配置完成双机热备功能）：

FW1 的配置	FW2 的配置
bfd 1 bind peer-ip 1.1.1.1 //建立 BFD 会话 1 监控 1.1.1.1 discriminator local 10 //本地标识符为 10 discriminator remote 20/ /对端标识符为 20 hrp track bfd-session 10 active //将 BFD 加入 Active 组	bfd 1 bind peer-ip 2.2.2.1 //建立 BFD 会话 1 监控 2.2.2.1 discriminator local 10 //本地标识符为 10 discriminator remote 20/ /对端标识符为 20 hrp track bfd-session 10 standby //将 BFD 加入 Active 组

总结

综上所述，尽管 VGMP 组监控和流量引导招式五花八门，但都遵循以下两条准则：

- **每当VGMP组监控的一个接口故障时，无论是直接监控还是间接监控，无论是监控防火墙本身的接口还是远端接口，VGMP组的优先级都会降低2。**
- **只有主用设备（VGMP组状态为Active）才会将流量引导到本设备上，备用设备（VGMP组状态为Standby）则是想办法拒绝将流量引导到本设备上。**

最后，我们总结下双机热备各种典型组网与 VGMP 故障监控和流量引导招式的关系，具体如下表所示：

双机热备组网	故障监控招式	流量引导招式
防火墙业务接口工作在三层，连接二层交换机（上篇讲到）	通过VRRP备份组监控接口 通过IP-Link监控接口（可选） 通过BFD监控接口（可选）	主用设备会向连接的交换机发送免费ARP报文，更新交换机的MAC转发表。

双机热备组网	故障监控招式	流量引导招式
防火墙业务接口工作在三层，连接路由器	直接监控接口 通过IP-Link监控接口（可选） 通过BFD监控接口（可选）	主用设备正常对外发布路由，备用设备发布的路由Cost值增加65500。
防火墙业务接口工作在二层（透明模式），连接二层交换机	通过VLAN监控接口	主用设备的VLAN能够转发流量，备用设备的VLAN被禁用。当主用设备切换成备用设备时，主用设备的VLAN中的接口会down然后up一次，触发上下行二层设备更新MAC转发表。
防火墙业务接口工作在二层（透明模式），连接路由器	通过VLAN监控接口	主用设备的VLAN能够转发流量，备用设备的VLAN被禁用。当主用设备切换成备用设备时，主用设备的VLAN中的接口会down然后up一次，触发上下行三层设备的路由收敛。

至此，双机热备的 VGMP 部分全部学习完毕，下篇强叔会为大家带来双机热备功能的最后一块拼图“HRP”，大家敬请期待！

你所不知道的 HRP

强叔在前几篇中讲解了双机热备的核心 VGMP。在介绍 VGMP 报文结构时我们看到了几种 HRP 协议定义的报文，本期强叔就要为大家解析 HRP 协议和这几种 HRP 报文。

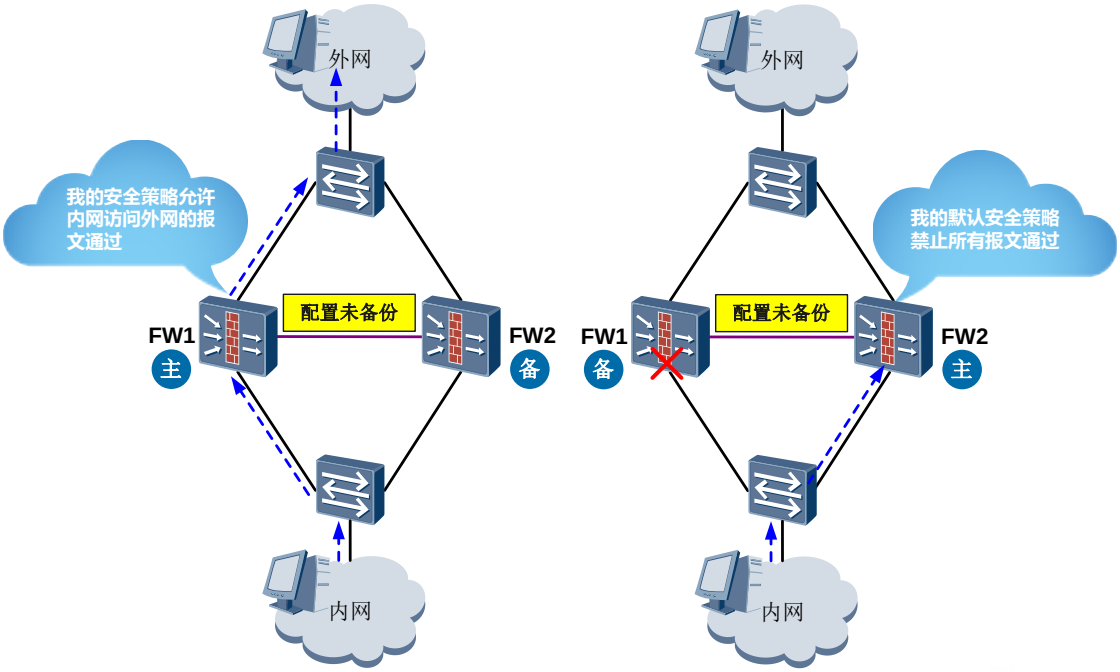
有的小伙伴儿们会说：“HRP 不就是负责双机的数据备份嘛。有什么难度？”其实 HRP 在备份时还是大有文章的，现在强叔就为大家揭秘这些 HRP 鲜为人知的细节。

为什么需要 HRP？

备份配置命令

防火墙通过执行命令（通过 Web 配置实际上也是在执行命令）来实现用户所需的各种功能。如果备用设备切换为主用设备前，配置命令没有备份到备用设备，则备用设备无法实现主用设备的相关功能，从而导致业务中断。

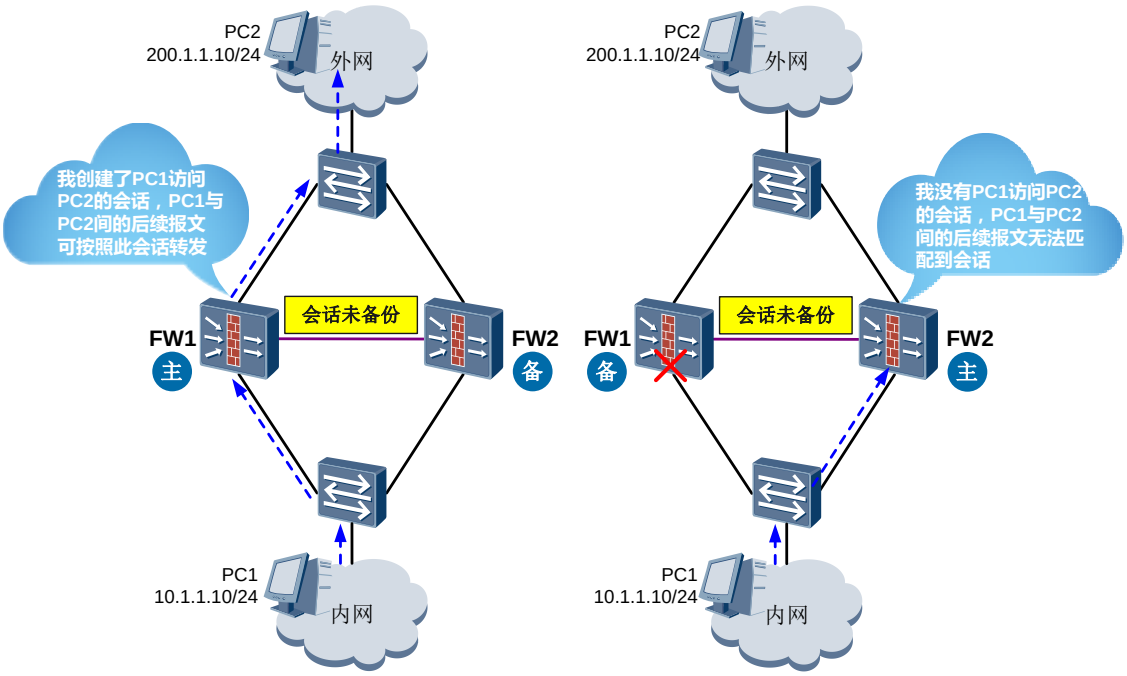
如下图所示，主用设备 FW1 上配置了允许内网用户访问外网的安全策略。如果主用设备 FW1 上配置的安全策略没有备份到备用设备 FW2 上，那么当主备状态切换后，新的主用设备 FW2 将不会允许内网用户访问外网（因为防火墙缺省情况下禁止所有报文通过）。



备份会话

防火墙属于状态检测防火墙，对于每一个动态生成的连接，都有一个会话表项与之对应。主用设备处理业务过程中创建了很多动态会话表项；而备用设备没有报文经过，因此没有创建会话表项。如果备用设备切换为主用设备前，会话表项没有备份到备用设备，则会导致后续业务报文无法匹配会话表，从而导致业务中断。

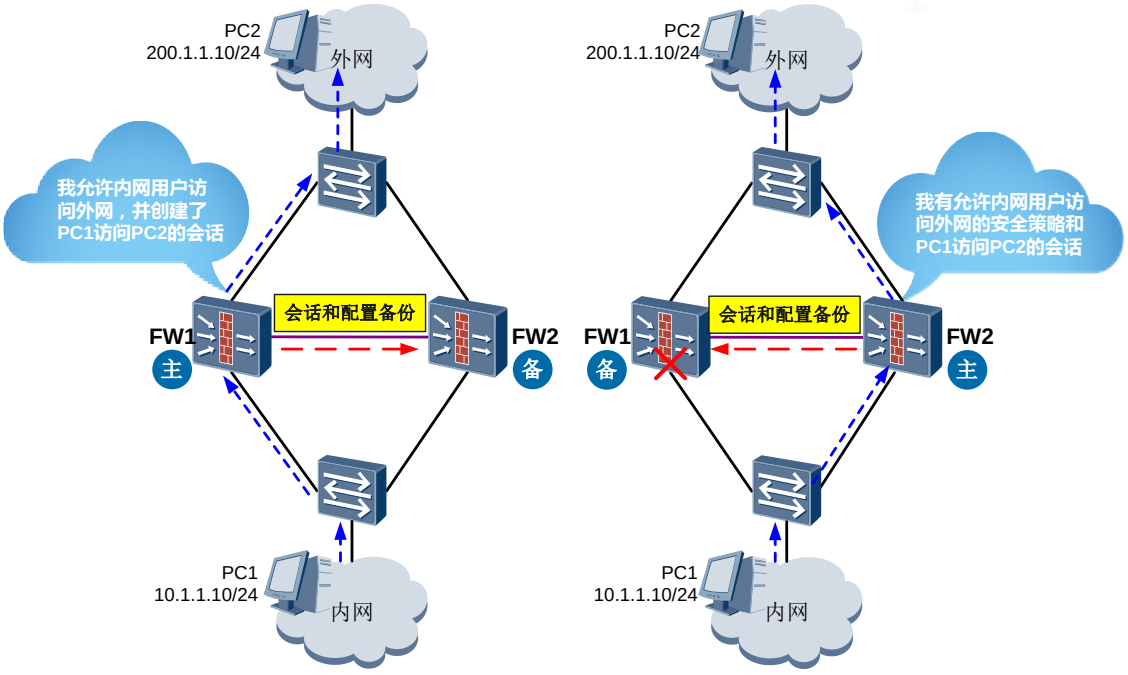
如下图所示，主用设备 FW1 上创建了 PC1 访问 PC2 的会话（源地址为 10.1.1.10，目的地址为 200.1.1.10），PC1 与 PC2 之间的后续报文会按照此会话转发。如果主用设备 FW1 上的会话不能备份到备用设备 FW2 上，那么当主备状态切换后，PC1 访问 PC2 的后续报文在 FW2 上匹配不到会话。这样就会导致 PC1 访问 PC2 的业务中断。



因此为了实现主用设备出现故障时备用设备能平滑地接替工作，必须在主用和备用设备之间备份关键配置命令和会话表等状态信息。为此华为防火墙引入了 HRP（ Huawei Redundancy Protocol）协议，实现防火墙双机之间动态状态数据和关键配置命令的备份。

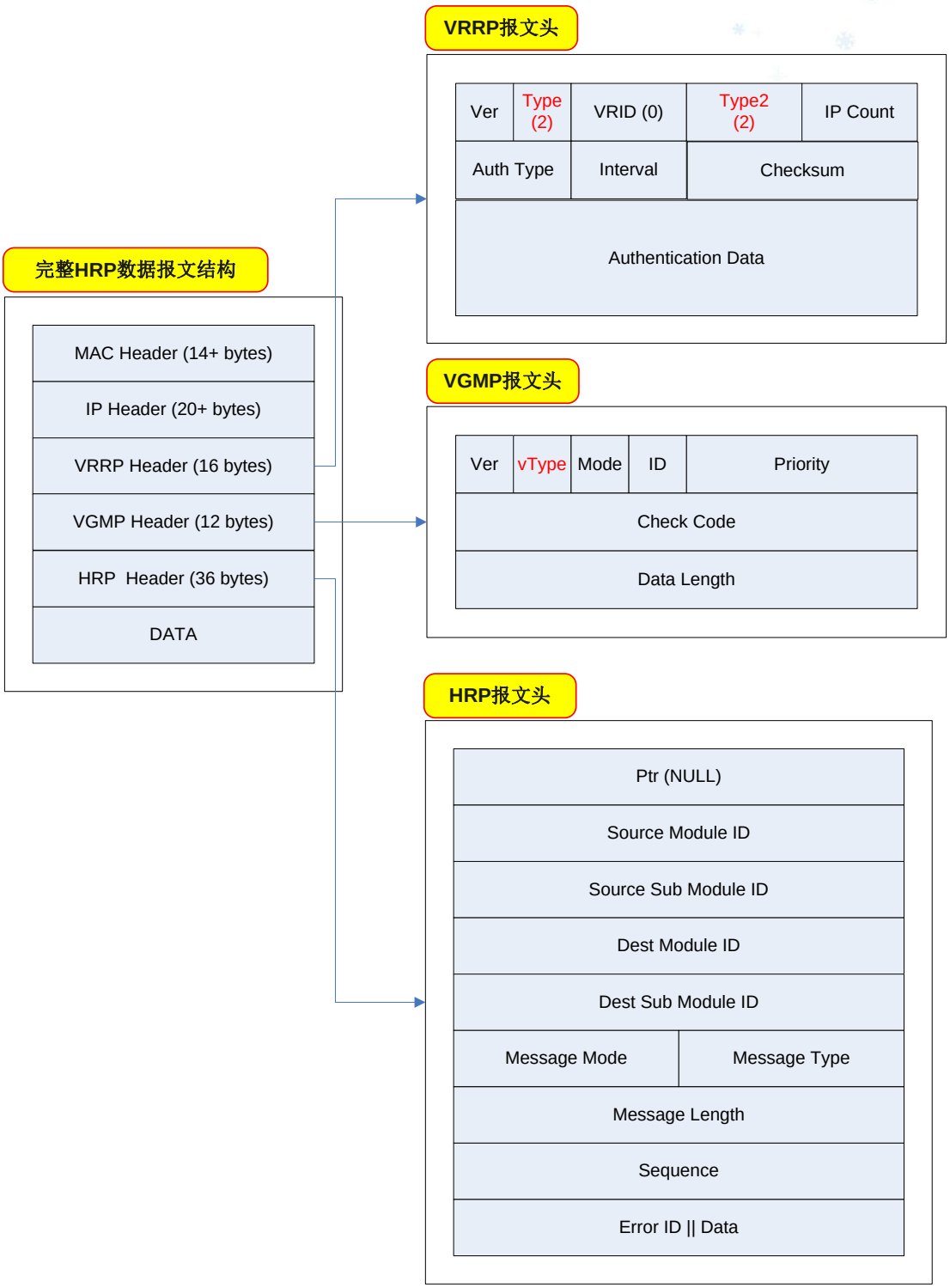
如下图所示，主用设备 FW1 上配置了允许内网用户访问外网的安全策略，所以 FW1 会允许内网 PC1 访问外网 PC2 的报文通过，并且会建立会话。由于在 FW1 和 FW2 上都使用了 HRP 协议（配置了双机热备中的 HRP 功能），因此主用设备 FW1 上配置的安全策略和创建的会话都会备份到备用设备 FW2 上。这样当主备状态切换后，由于备用设备上已经存在允许内网用户访问外网的安全策略以及 PC1 访问 PC2 的会话，所以 PC1 访问 PC2 业务报文不会被

禁止或中断。



HRP 是如何实现备份的？

防火墙通过心跳口（HRP 备份通道）发送和接收 HRP 数据报文来实现配置和状态信息的备份。如下图所示，HRP 数据报文从外到内依次封装了 VRRP 报文头、VGMP 报文头和 HRP 报文头。其中 VRRP 报文头中 Type=2，Type2=2。VGMP 报文头中的“vType”字段对应为“HRP 数据报文”的取值。

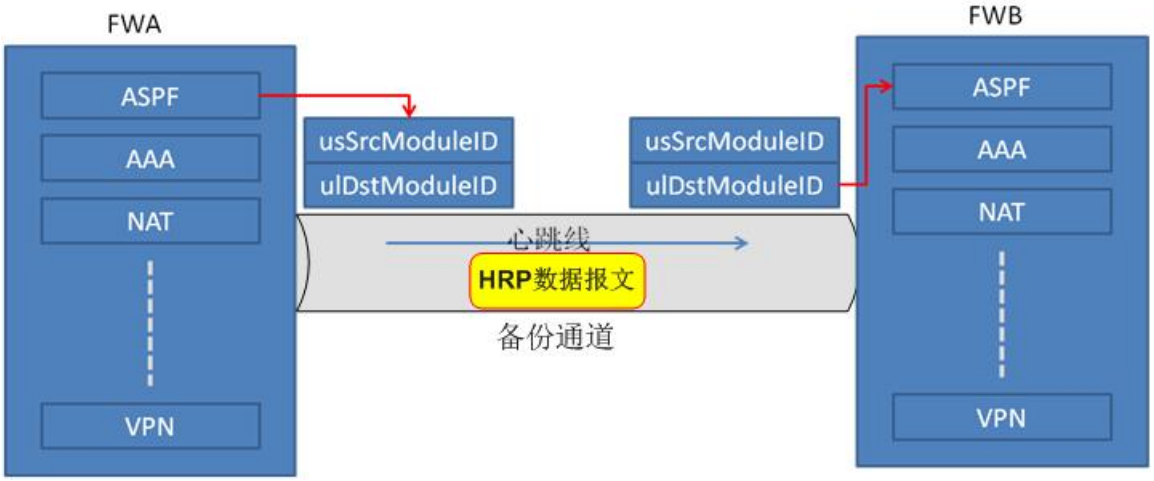


HRP 报文头中的关键参数解释如下：

- Source Module ID和Source Sub Module ID表示本端防火墙哪些特性模块和子模块的数据需要备份。
- Dest Module ID和和Dest Sub Module ID表示需要向对端防火墙的哪些特性模块和子模块备份数据。

HRP 数据备份的过程如下图所示：

- FWA在发送HRP数据报文时，会将ASPF特性模块的ID写入HRP报文头的“Source Module ID”和“Dest Module ID”字段中，并将ASPF模块的配置和表项信息封装到HRP数据报文中。
- FWA将HRP数据报文通过备份通道（心跳线）发送给FWB。
- FWB收到HRP数据报文后，会根据HRP报文头中的“Source Module ID”和“Dest Module ID”字段将报文中的配置和表项信息发送到本端的ASPF特性模块，并进行配置与表项的下发。



在前面双机热备的第二篇中我们讲到 USG6000 系列防火墙和 USG2000/5000 系列 V3R1 版本防火墙还支持将各种 VGMP 报文和 HRP 报文封装成 UDP 报文。当然这里介绍的 HRP 数据报文，以及下面讲的心跳链路探测报文和一致性检查报文都支持这种 UDP 方式的封装。这种方式的封装方法就是在 VRRP 报文头上再封装一个 UDP 头。HRP 数据报文的 UDP 封装结构如下图所示。

IP Header	UDP Header	VRRP Header	VGMP Header	HRP Header	DATA

UDP 封装的好处前面我们也讲过：UDP 封装后的报文是单播报文，只要路由可达就可以跨越网段传输，而且能够被安全策略控制。

HRP 能够备份哪些配置与状态信息？

防火墙能够备份的配置如下（以 HUAWEI USG6000 系列 V100R001 版本为例）：

- 策略：安全策略 、NAT策略 、带宽管理 、认证策略 、攻击防范 、黑名单 、ASPF
- 对象：地址 、地区、服务 、应用 、用户 、认证服务器 、时间段 、URL分类 、关键字组 、邮件地址组 、签名 、安全配置文件（反病毒 、入侵防御 、URL过滤 、文件

过滤 、内容过滤 、应用行为控制 、邮件过滤）

- 网络：新建逻辑接口 、安全区域 、DNS 、IPSec 、SSL VPN 、TSM联动
- 系统：管理员 、日志配置

说明：一般情况下，display、reset、debugging 命令都不支持备份。

根据上面的描述我们可以看到，防火墙的网络基本配置如接口地址和路由等都不能够备份，这些配置需要在双机热备状态成功建立前配置完成。而上面支持备份的配置可以在双机热备状态成功建立后，只在主用设备上配置。

防火墙能够备份的状态信息如下：

- 会话表
- SeverMap表
- IP监控表
- 分片缓存表
- GTP表
- 黑名单
- PAT方式端口映射表
- NO-PAT方式地址映射表

HRP 的备份方向是怎样的？什么是配置主和配置从设备？

在主备备份组网下，配置命令和状态信息都由主用设备备份到备用设备。

在负载分担组网下，配置命令只能由“配置主设备”备份到“配置备设备”。状态信息则是两台设备相互备份的。

那么什么是配置主和配置备设备呢？

在负载分担组网下，两台防火墙都是主用设备（都有状态为 Active 的 VGMP 组）。因此如果允许两台主用设备之间能够相互备份命令，那么可能就会造成两台设备命令相互覆盖或冲突的问题。所以为了方便管理员对两台防火墙配置的统一管理，避免混乱，我们引入配置主和配置从设备的概念。我们定义**负载分担组网下，发送备份配置命令的防火墙称为配置主设备（命令行提示符前有 HRP_A 前缀），接收备份配置命令的防火墙称为配置从设备（命令行提示符前有 HRP_S 前缀）。**

那么在负载分担组网下，如何确定配置主和配置备设备呢？

在负载分担组网下，最先建立双机热备状态的防火墙会成为配置主设备。也就是最先启用双

机热备功能的防火墙会成为配置主设备。

HRP 的几种备份方式各有什么特点？如何使用？

双机热备的 HRP 支持以自动备份、手工批量备份和快速备份三种方式。这三种备份方式的描述和区别下面我们一一来介绍。

自动备份

自动备份功能（命令为 hrp auto-sync [config | connection-status]）缺省为开启状态，能够自动实时备份配置命令和周期性地备份状态信息，适用于各种双机热备组网。

- 启用自动备份功能后，主用（配置主）设备上每执行一条可以备份的命令时，此**配置命令**就会被**立即**同步备份到备用（配置备）设备上。

  如果在主用（配置主）设备上执行不可以备份的命令，则该命令仅在主用（配置主）设备上执行。

  对于可以备份的配置命令，只能在主用（配置主）设备上配置，备用（配置备）设备上不能配置。对于不可以备份的配置命令，备用（配置备）设备上可以配置。关于哪些配置命令可以备份或不可以备份请参见前面的“HRP能够备份哪些配置与状态信息？”。
- 启用自动备份功能后，主用设备会**周期性**的将可以备份的**状态信息**备份到备用设备上。

  即主用设备的状态信息建立后不会立即备份，需要经过一个周期后才会备份到备用设备。

  自动备份不会备份以下类型的会话（只快速会话备份支持）：
  - 到防火墙自身的会话，例如管理员登录防火墙时产生的会话
  - 未完成3次握手的TCP半连接会话
  - 只为UDP首包创建，而不被后续包匹配的会话

手工批量备份

手工批量备份需要管理员手工触发，每执行一次手工批量备份命令（hrp sync [config | connection-status]），主用设备就会立即同步一次配置命令和状态信息到备用设备。因此手工批量备份主要适用于主备设备之间配置不同步，需要手工同步的场景。我们可以在防火墙上配置一致性检查功能，检查两台防火墙的配置是否同步。一致性检查功能需要用到 HRP 一致性检查报文，这个我们最后会讲到。

- 执行手工批量备份命令后，主用（配置主）设备会**立即**同步一次可以备份的配置命令到备用（配置备）设备。
- 执行手工批量备份命令后，主用设备会**立即**同步一次可以备份的**状态信息**到备用设备，而不必等到自动备份周期的到来。

快速备份

快速会话备份功能（命令为 hrp mirror session enable），适用于负载分担的工作方式，以应对报文来回路径不一致的场景。为了保证状态信息的及时同步，快速备份功能只是备份状态信息，不备份配置的命令。配置命令的备份由自动备份功能实现。

启用快速备份功能后，主用设备会**实时**的将可以备份的状态信息（包括上面提到的自动备份不支持的会话）都同步到备用设备上。即在主用设备状态信息建立的时候**立即**将其**实时**备份到备用设备。

综上所示，三种备份方式的使用方式通常是：自动备份（hrp auto-sync [config | connection-status]）默认开启，不要关闭；如果主备设备之间配置不同步，需要执行手工批量备份的命令（hrp sync [config | connection-status]）；如果是负载分担组网，一般需要开启快速会话备份功能（hrp mirror session enable）。

下面来讲解为什么快速会话备份特别适用于负载分担组网？

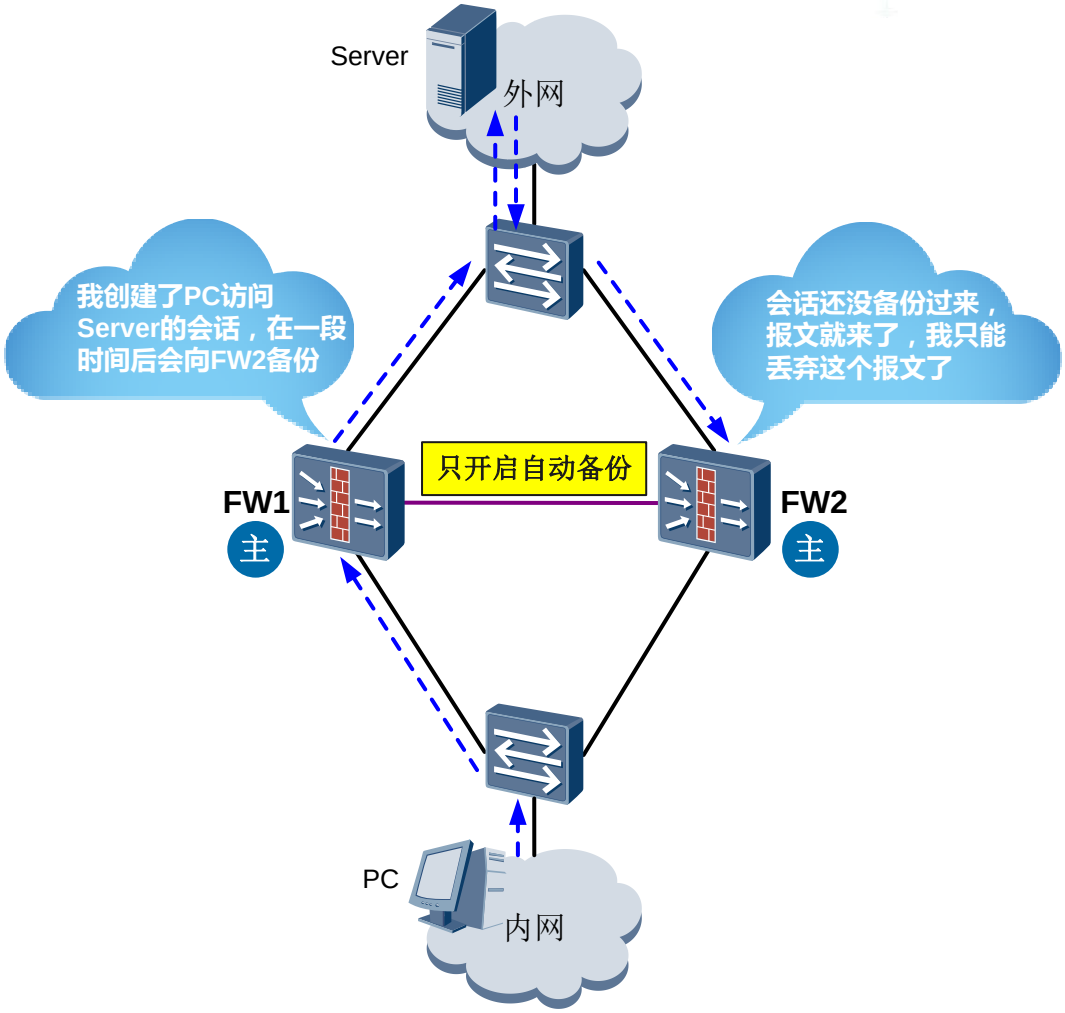
负载分担组网下，由于两台防火墙都是主用设备，都能转发报文，所以可能存在报文的来回路径不一致的情况，即来回两个方向的报文分别从不同的防火墙经过。这时如果两台防火墙的状态信息没有及时相互备份，则回程报文会因为没有匹配到状态信息而被丢弃，从而导致业务中断。

为防止上述现象发生，需要在负载分担组网下配置快速会话备份功能，使两台防火墙能够实时的相互备份状态信息，使回程报文能够查找到相应的状态信息表项，从而保证内外部用户的业务不中断。

下面举个例子来说明。如下图所示，FW1 和 FW2 形成双机热备的负载分担组网。内网 PC 访问外网 Server 的报文通过 FW1 转发，并建立会话。由于来回路径不一致，Server 返回给 PC 的回程报文转发到 FW2。这时如果只启用了自动备份功能，则 FW1 的会话还没有来得及备份到 FW2 上。这就导致回程报文无法在 FW2 上匹配会话而被 FW2 丢弃，从而造成业务中断。

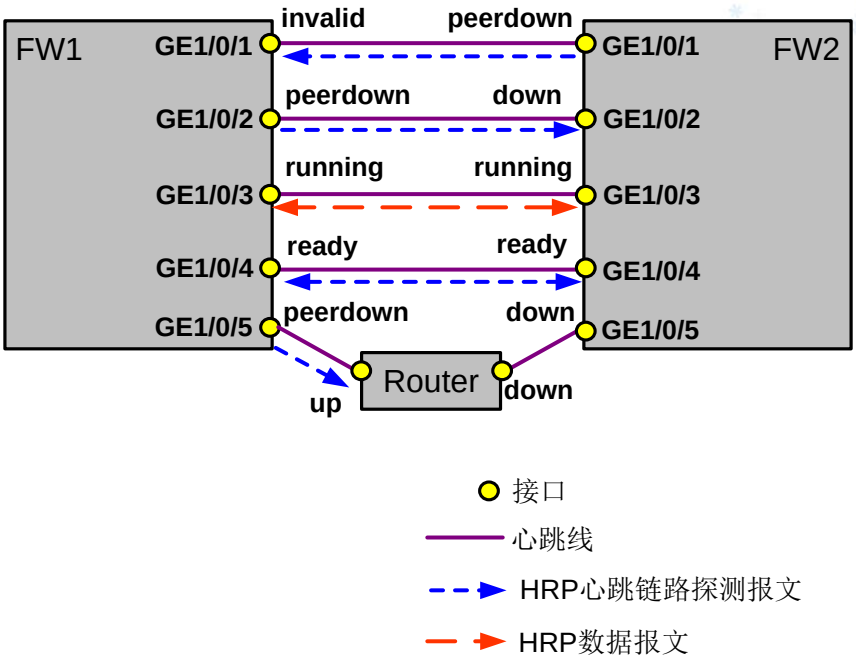
这时如果启用了会话快速备份功能，则 FW1 上产生的会话会立即备份到 FW2 上。这样回程

报文就能在 FW2 上匹配到会话，从而被正常转发到 PC。



什么是心跳链路探测报文？与 HRP 心跳报文有什么区别？

两台防火墙之间备份的数据是通过防火墙的心跳口发送和接收的，是通过心跳链路（备份通道）传输的。心跳口必须是状态独立且具有 IP 地址的接口，可以是一个物理接口 GE，也可以是为了增加带宽，由多个 GE 接口捆绑而成的一个逻辑接口 Eth-Trunk（通常情况下，备份数据流量约为业务流量的 20%～25%，请根据备份数据量的大小选择捆绑 GE 接口的数量）。如下图所示，我们经常会配置多个心跳口（多个 GE 或者多个 Eth-Trunk 接口），从而形成多条心跳链路，以保证备份数据的可靠传输。两台防火墙通过心跳口相互发送心跳链路探测报文，来检测对端设备的心跳口能否正常接收本端设备的报文，以确定是否有心跳口可以使用。心跳链路探测报文也是由 VRRP 报文头封装的，当 VRRP 报文头中 Type=2，Type2=2 时，报文封装成心跳探测链路报文。



如上图所示，心跳接口有五种状态（执行命令 **display hrp interface** 可以查看）：

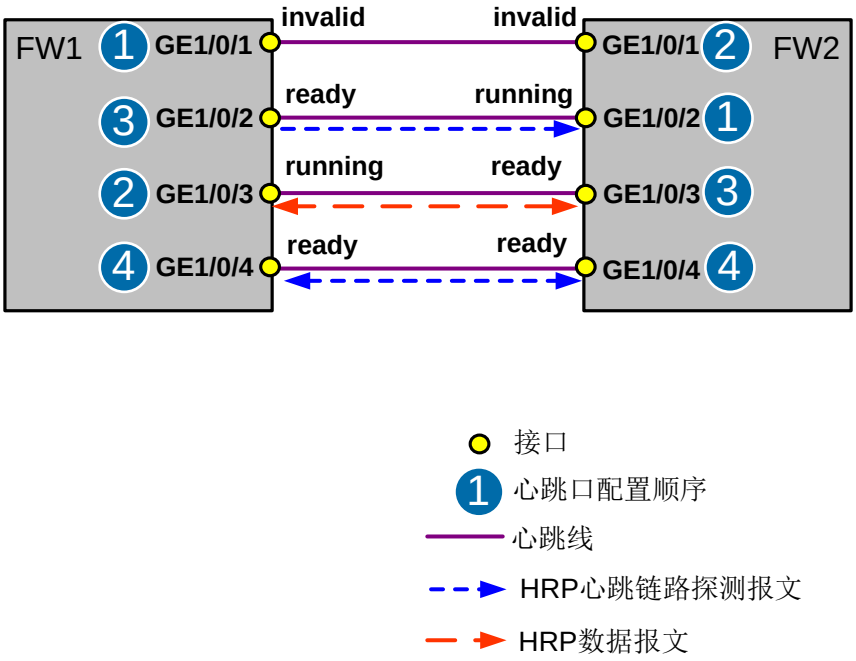
- **invalid**：当本端防火墙上的心跳口配置错误时显示此状态（物理状态up，协议状态down），例如指定的心跳口为二层接口或未配置心跳接口的IP地址。
- **down**：当本端防火墙上的心跳口的物理与协议状态均为down时，则会显示此状态。
- **peerdown**：当本端防火墙上的心跳口的物理与协议状态均为up时，则心跳口会向对端对应的心跳口发送心跳链路探测报文。如果收不到对端响应的报文，那么防火墙会设置心跳接口状态为peerdown。但是心跳口还会不断发送心跳链路探测报文，以便当对端的对应心跳口up后，该心跳链路能处于连通状态。
- **ready**：当本端防火墙上的心跳口的物理与协议状态均为up时，则心跳口会向对端对应的心跳口发送心跳链路探测报文。如果对端心跳口能够响应此报文（也发送心跳链路探测报文），那么防火墙会设置本端心跳接口状态为ready，随时准备发送和接受心跳报文。这时心跳口依旧会不断发送心跳链路探测报文，以保证心跳链路的状态正常。
- **running**：当本端防火墙有多个处于ready状态的心跳口时，防火墙会选择最先配置的心跳口形成心跳链路，并设置此心跳口的状态为running。状态为running的接口负责发送和HRP心跳报文、HRP数据报文、HRP一致性检查报文和VGMP报文。

  这时其余处于ready状态的心跳口处于备份状态，当处于running状态的心跳口或心跳链路故障时，其余处于ready状态的心跳口依次（按配置先后顺序）接替当前心跳口处理业务。

  如上图所示，由于两台防火墙心跳接口的配置顺序与接口编号顺序相同，所以先配置的处于ready状态的心跳口GE1/0/3成为running状态，而后配置的处于ready状态的心跳口

GE1/0/4处于备份状态。

下面我们来看一种特殊情况：两台防火墙的心跳口配置顺序不一致时的情况。如下图所示，两台防火墙的 GE1/0/2、GE1/0/3、GE1/0/4 接口都是能够正常工作的心跳口（处于 ready 状态）。由于 FW1 先配置 GE1/0/3 为心跳口，而 FW2 先配置 GE1/0/2 为心跳口，所以 FW1 的 GE1/0/3 心跳口状态为 running，而 FW2 的 GE1/0/2 心跳口状态为 running。这样 FW1 的 HRP 数据报文会从 GE1/0/3 发送，而 FW2 的 HRP 数据报文会从 GE1/0/2 发送。虽然 HRP 数据报文的发送和接收接口不一致，但这并不会影响双机热备的正常工作。



最后来总结下心跳链路探测报文和 HRP 心跳报文的区别。心跳链路探测报文用于检测对端设备的心跳口能否正常接收本端设备的报文，以确定心跳口是否可用的。只要本端心跳接口的物理和协议状态 up 就会向对端心跳口发送心跳链路探测报文进行探测。HRP 心跳报文是用于探测和感知对端设备（VGMP 组）是否正常工作的。HRP 心跳报文只有主用设备的 VGMP 组通过状态为 running 的心跳口发出。

HRP 一致性检查报文能够检查哪些内容？是如何实现的？

HRP 一致性检查报文用于检测双机热备状态下的两台防火墙的**双机热备配置**是否一致以及**策略配置**是否相同。双机热备配置的一致性检查包括两台防火墙是否监控了相同的业务接口，是否配置了相同的心跳接口等。策略配置一致性检查主要检查两台防火墙是否配置了相同的策略，包括安全策略、带宽策略、NAT 策略、认证策略和审计策略。HRP 一致性检查报文也是由 VRRP 报文头封装的，当 VRRP 报文头中 Type=2，Type2=5 时，报文封装成 HRP 一

致性检查报文。

HRP 一致性检查的实现原理如下：

1) 执行一致性检查命令（**hrp configuration check { all | audit-policy | auth-policy | hrp | nat-policy | security-policy | traffic-policy }**）后，执行此命令的设备会发送一致性检查请求报文给对端，并且同时收集自身的相关模块的配置信息摘要。

2) 对端设备收到请求后，会收集自身相关模块的配置信息摘要，然后封装到一致性检查报文中返回给本端设备。

3) 本端设备会对比自身的配置摘要和对端设备的配置摘要，并记录比较信息。客户可以执行命令**display hrp configuration check**查看一致性检查结果。例如下面的结果表示双机热备配置一致。

```
HRP_A[FWA] display hrp configuration check hrp
Module          State   Start-time              End-time                Result
hrp             finish 2008/09/08 14:21:56 2008/09/08 14:21:56 same configuration
```

至此，双机热备篇圆满结束，感谢各位小伙伴的观看~ 下周强叔将为大家带来本季的收官之作——防火墙**出口选路**篇，大家敬请期待！

就近选路

——缺省路由有备无患，明细路由近路建功

大家好，经过前面这么多篇的积累，相信大家对防火墙的技术知识已经有了一定的了解。从今天开始，强叔将和大家一起探索防火墙出口选路的相关知识，希望能给大家带来一些收获。

在强叔的第一篇贴子中我们提到，防火墙主要部署在网络边界起到隔离的作用，虽然隔离了内外网，但防火墙也承担起内网与外网互联的作用，内网和外网交互的流量都要经过防火墙来转发。然而，在实际场景中，企业出于带宽和可靠性的要求，会向多个 ISP 租用多条 Internet 链路带宽资源，处于出口位置的防火墙就有多个出口链路连接到 Internet，如何给用户流量选择合适的出口链路将是企业网络管理员需要考虑的问题。

为此，强叔也专门总结了防火墙做为企业出口网关时面对多出口环境的几种常用选路方式。

首先，让我们从缺省路由+明细路由的就近选路方式开始说起。

缺省 VS 明细

何为就近选路？顾名思义为选择较近的路，在多出口网络中指的是报文选择离目标网络花销较小的链路进行转发。那报文是如何选择花销较小的链路转发的呢？通过缺省路由和明细路由即可实现。下面，强叔通过解答几个问题的方式先来介绍缺省路由和明细路由的一些基本概念，帮助大家理解。

第一个问题：什么是缺省路由，缺省路由是属于静态路由吗？

其实缺省路由是一种特殊的路由，可以通过静态路由配置，也可以是动态路由生成，如 OSPF 和 IS-IS。所以确切的说，缺省路由不属于静态路由。在路由表中，缺省路由以目的网络为 0.0.0.0、子网掩码为 0.0.0.0 的形式出现。下面为路由表中的缺省路由。

```
[FW] display ip routing-table
Route Flags: R - relay, D - download to fib
------------------------------------------------------------------------------
Routing Tables: Public

         Destinations : 1        Routes : 2

Destination/Mask    Proto   Pre  Cost      Flags NextHop         Interface

        0.0.0.0/0   Static 60   0           RD  10.1.1.2        GigabitEthernet2/2/21
                    Static 60   0           RD  10.2.0.2        GigabitEthernet2/2/17
```

如果数据报文的目的地址不能与任何路由相匹配，那么系统将使用缺省路由转发该数据报文。

第二个问题：什么是明细路由？

强叔认为，明细路由是相对来说的，相对缺省路由，在路由表中的其他路由都属于明细路由，如 10.1.0.0/16、192.168.1.0/24 相对缺省路由都属于明细路由。相对有 10.1.0.0/16 这条汇总路由，10.1.1.0/24、10.1.2.0/24、10.1.3.0/24 这三条路由都属于它的明细路由。明细路由与路由协议类型无关，可以是静态路由配置的，也可以是动态路由协议生成的。

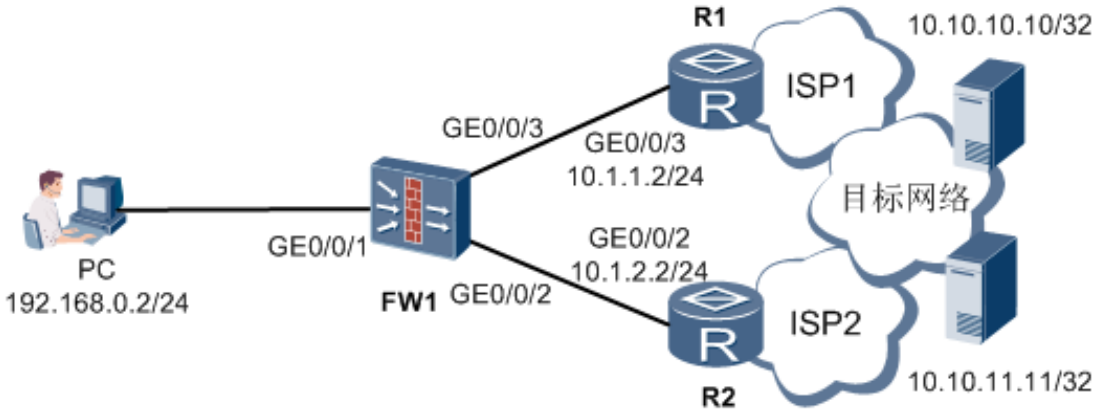
第三个问题：报文是如何查找路由表的？

可能大家都知道，报文查找路由表时是按照最长匹配原则进行查找，什么意思呢？举个例子，路由表中有 10.1.0.0/16、10.1.1.0/24 和 0.0.0.0/0 三条路由，当目的地址为 10.1.1.1/30 的报文查找路由表时，最终匹配的路由将是 10.1.1.0/24 这条路由，因为前面的 24 位与这个报文的目的地址一致，为匹配路由中最长。如果是目的地址为 192.168.1.1/30 的报文查找路由表，则只能匹配 0.0.0.0/0 这条缺省路由了，因为报文的目的地址不能与任何明细路由相匹配，最终系统将使用缺省路由转发该报文。

从上面的问题中我们可以了解到，当路由表中有明细路由时，报文是先匹配明细路由，如果没有明细路由再查找缺省路由。

下面我们来看第四个问题：多条缺省路由间是如何选路的？

我们先来看下面一个组网，在防火墙上我们配置两条缺省路由，一个下一跳指向 R1，一个下一跳指向 R2，现在我们在 PC 上 ping 目标网络上的两个服务器地址。



两条缺省路由在防火墙上配置如下：

```
[FW] ip route-static 0.0.0.0 0 10.1.1.2
[FW] ip route-static 0.0.0.0 0 10.1.2.2
```

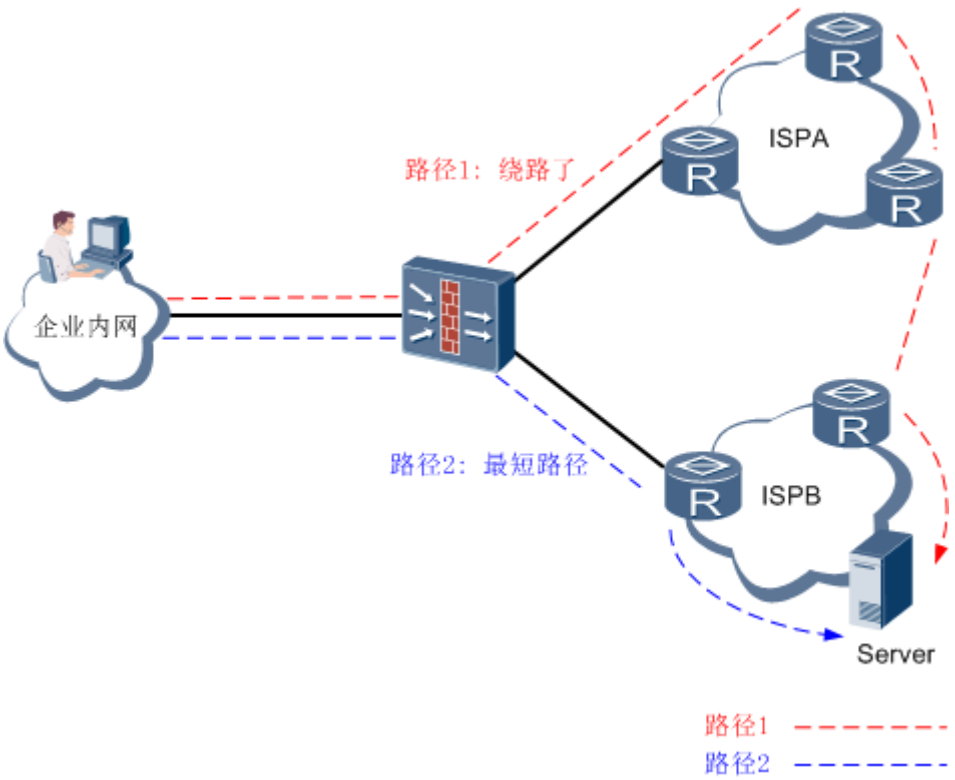
通过在 FW1 的 GE0/0/3 接口上抓包发现报文都是从 GE0/0/3 接口进行转发，截图如下：

No.	Time	Source	Destination	Protocol	Info
1	0.000000	192.168.0.2	10.10.10.10	ICMP	Echo (ping) request (id=0x042f, seq(be/le)=1/256, ttl=127)
2	0.000000	10.10.10.10	192.168.0.2	ICMP	Echo (ping) reply (id=0x042f, seq(be/le)=1/256, ttl=253)

No.	Time	Source	Destination	Protocol	Info
1	0.000000	192.168.0.2	10.10.11.11	ICMP	Echo (ping) request (id=0xfe2f, seq(be/le)=1/256, ttl=127)
2	0.000000	10.10.11.11	192.168.0.2	ICMP	Echo (ping) reply (id=0xfe2f, seq(be/le)=1/256, ttl=253)

为什么出现这种情况呢？两条缺省路由不进行负载分担吗？事实上，多条缺省路由在选路的时候是根据源 IP 地址+目的 IP 地址的 HASH 算法来算出报文具体走哪条链路的，这种算法主要是看报文的源 IP 地址和目的 IP 地址，地址不同，计算出的结果也会不相同。这种算法下，等价缺省路由之间转发报文的机会是均等的。举个例子，如果报文的源 IP 地址相同，目的地址是相邻的，如 10.1.1.1 和 10.1.1.2，那么选路的时候，将会各分担一条流进行转发。然而，由于网络中访问流量的源和目的地址是随机的，所以 HASH 计算结果完全不可控。这个时候虽然多条缺省路由是等价路由，也有可能出现所有报文都从一条链路转发的情况。这个结果也印证了上面举例中报文都从 GE0/0/3 接口转发的原因。

好了，前面说的都是一些基础知识，现在让我们来看看缺省路由+明细路由的就近选路方式是如何就近选路的。先看一个简单的网络环境，如下图所示。



当企业内网用户访问外网服务器 Server 时，报文途经防火墙有两条路径，正常情况下，企业一般会在出口防火墙上配置两条缺省路由，每个 ISP 一条。在前面我们说过缺省路由的选路是通过源 IP+目的 IP 的 HASH 算法来决定数据报文的转发路径，这就有可能导致到 ISPB 的 Server 流量经过 HASH 算法计算后从图中的路径 1 进行转发了，这样从路径 1 到 ISPA，再

经过 ISPA 到 ISPB，绕一大圈后才能到能最终目的地，严重影响的转发效率和用户体验。

那有什么办法让报文不绕道呢？通过配置明细路由即可达到要求，前面我们也说过，报文是优先匹配明细路由的，没有匹配的明细路由再去查找缺省路由的。就上面的组网，我们可以配置到 Server 的明细路由，下一跳指向 ISPB，这样报文匹配到这条明细路由后就不会绕道转发了，而是选择明细路由指定的链路，从图中看，选择的是两条路径中的最短路径，这就是我们所说的就近选路了。我们也可以通过第一个组网验证一下。我们在防火墙上配置如下两条静态路由：

```
[FW] ip route-static 10.10.10.10 255.255.255.255 10.1.1.2(下一跳为 R2 地址)
[FW] ip route-static 10.10.11.11 255.255.255.255 10.1.2.2(下一跳为 R3 地址)
```

通过在 FW1 的 GE0/0/3 接口上抓包发现只有去往 10.10.10.10 的报文。

No.	Time	Source	Destination	Protocol	Info
1	0.000000	192.168.0.2	10.10.10.10	ICMP	Echo (ping) request (id=0x042f, seq(be/le)=1/256, ttl=127)
2	0.000000	10.10.10.10	192.168.0.2	ICMP	Echo (ping) reply (id=0x042f, seq(be/le)=1/256, ttl=253)

在 FW1 的 GE0/0/2 接口上抓包发现有去往 10.10.11.11 的报文。

No.	Time	Source	Destination	Protocol	Info
1	0.000000	192.168.0.2	10.10.11.11	ICMP	Echo (ping) request (id=0xfe2f, seq(be/le)=1/256, ttl=127)
2	0.000000	10.10.11.11	192.168.0.2	ICMP	Echo (ping) reply (id=0xfe2f, seq(be/le)=1/256, ttl=253)

这就证明，报文是优先查找刚配置的两条明细路由。然而在实际组网环境中，Internet 上的 Server 是非常多的，作为管理员在出口网关防火墙上配置那么多的明细路由是不现实的，那有没有一种方便快捷的方法配置明细路由呢？这就要我们的 ISP 路由功能出场了。那什么是 ISP 路由呢？

ISP 路由

ISP 路由，从名字来看有一个 ISP，其实也是它的由来。每个 ISP 都会有自己的公网知名网段，如果把这个 ISP 的所有公网知名网段都像上面说的一样配置成明细路由，那么去往这个 ISP 的所有报文都不会绕路转发了。那如何把 ISP 的公网知名网段变成明细路由呢？

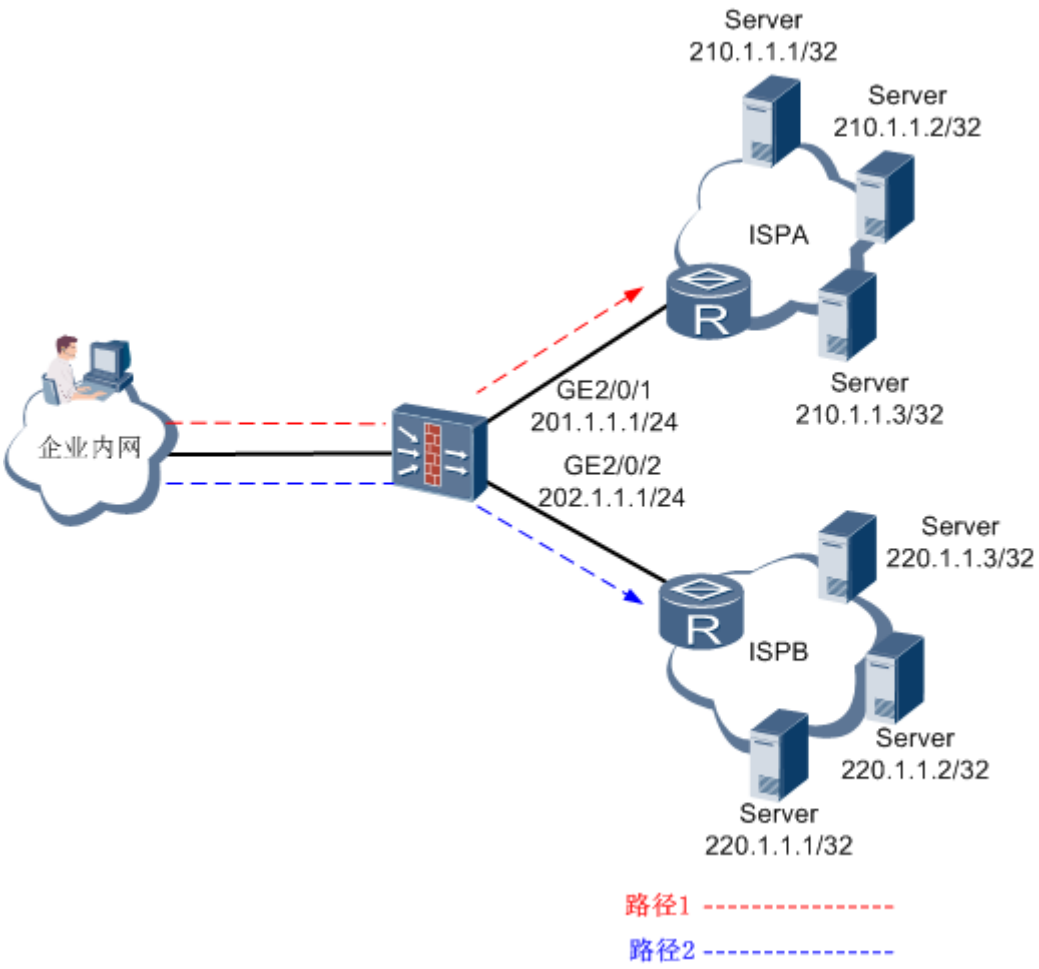
首先管理员需要先收集 ISP 内的所有公网网段（网上都能够搜索到），然后把地址网段编辑到后缀为.csv 的文件中（我们称之为 ISP 地址文件），编辑要求按如下图所示：

A	B
#以#开头的行代表该行是注释行，而非实际数据。	
#提示：导出的格式与导入的格式是相同的， 可以保持文件不变或者修改内容后，再次导入到设备中。	← 填写说明及要求
#目的IP范围(支持单个IP、IP/掩码和IP段类型， 例如：1.1.1.1，1.1.1.0/24,1.1.1.1/255.255.255.0,1.1.1.0-1.1.1.255)。	
##ISP选路	
	目的IP范围
	1.184.0.0-1.185.255.255
	58.116.0.0-58.119.255.255
	58.128.0.0-58.135.255.255
	58.154.0.0-58.155.255.255
	58.192.0.0-58.207.255.255
	59.64.0.0-59.79.255.255
	110.64.0.0-110.65.255.255
	111.114.0.0-111.117.255.255

编辑地址文件完成后，我们需要把后缀为.csv 的 ISP 地址文件上传到防火墙的指定路径上，比如 cfcard 中。上传的方法有很多，比如 SFTP、FTP、TFTP 等。

ISP 地址文件上传到防火墙后，通过设置出接口和下一跳，可以让 ISP 地址文件中的每个 IP 地址网段都生成一条 ISP 路由。

下面我们通过一个实验组网检验下 ISP 路由选路的效果，组网如下图所示：



在此组网中，ISPA 和 ISPB 内的地址网段我们分别编辑在 ispa.csv 与 ispb.csv 文件中。

首先我们通过 SFTP、FTP、TFTP 等方式上传两个 csv 文件到防火墙的指定路径中。其中高端防火墙路径为 cfcard:/isp/；下一代防火墙为 hda1:/isp/。

完成 csv 文件上传后，通过相关命令设置相应出接口和下一跳，启动 ISP 路由功能。以高端防火墙为例配置命令如下：

```
[FW] isp set filename ispa.csv GigabitEthernet 2/0/1 next-hop 201.1.1.2
```

此外，我们还可以通过 Web 配置方式来配置 ISP 路由，这种方式更为简单，上传 csv 文件和导入配置通过一步就可以完成。我们以 USG9000 为例，导入方式如下：

ispb.csv 的导入方法与 ispa.csv 一致，只是出接口和下一跳改成 GE2/0/2 和 202.1.1.2。

导入完成后，防火墙会生存如下路由：

Destination/Mask	Proto	Pre	Cost	Flags	NextHop	Interface
210.1.1.1/32	ISP	60	0	D	201.1.1.2	GigabitEthernet2/0/1
210.1.1.2/32	ISP	60	0	D	201.1.1.2	GigabitEthernet2/0/1
210.1.1.3/32	ISP	60	0	D	201.1.1.2	GigabitEthernet2/0/1
220.1.1.1/32	ISP	60	0	D	202.1.1.2	GigabitEthernet2/0/2
220.1.1.2/32	ISP	60	0	D	202.1.1.2	GigabitEthernet2/0/2
220.1.1.3/32	ISP	60	0	D	202.1.1.2	GigabitEthernet2/0/2

内网用户访问属于 ISPA 的 Server 时，报文匹配路由表后从 GigabitEthernet2/0/1 接口进行转发；同理，访问 ISPB 的 Server 时，会从 GigabitEthernet2/0/2 接口进行转发。这样总是能保证从最短的路径转发到目标网络。

从上面的路由表来看，ISP 路由与静态路由非常相似，在路由表中，除了协议类型为 ISP 外，表中其它内容与静态路由完全一样，且两种路由之间是可以互相覆盖的，如静态路由先配置，后再导入 ISP 路由后，路由表中此条路由的协议类型会从 static 变成 ISP，反之亦然。但在应用中 ISP 路由与静态路由还是有如下区别：

1. 静态路由是手动一条一条配置，配置文件中能够显示出来；ISP 路由只能通过上面所述的方式集体导入，且配置文件中无法显示出 ISP 路由。
2. 静态路由可以逐条删除、增加；ISP 路由只能从 ISP 地址文件中把地址网段删除、增加，而不能通过命令删除或增加单条 ISP 路由。

上面说的是管理员如何构建 ISP 路由的过程，实际上，防火墙在出厂的设置中已经内置了中国移动、中国电信、中国联通和中国教育网 4 个 ISP 的公网知名网段，只需要管理员执行导入即可启动 ISP 路由。

总结就近选路方式，其实就是三种路由的 PK 结果：

- 缺省等价路由让经过防火墙的所以报文都能匹配路由转发，但无法保证报文转发选择最短链路（通过源 IP 地址+目的 IP 地址的 HASH 算法来选择报文转发出口）。
- 明细路由保证访问不同 ISP 服务器的报文都从连接相应 ISP 的链路转发，达到就近访问效果，但是明细路由的手工大批量配置是困扰企业网络管理员的一个难题。
- ISP 路由则填补了明细路由难以手工大批量配置的缺点，分分钟就能搞定一个 ISP 所有地址网段的明细路由配置。

这三种路由各有特点，配合使用方能弥补相互之间的缺陷、发挥出每种路由的优势。配合使用时，明细路由和 ISP 路由用来指导报文近路转发，没有匹配到明细路由的报文通过查找缺省路由完成转发。

然而就近选路方式是以路由为基础的选路方式，大家都知道，查找路由是通过报文目的地址来查找的，那问题就来了，如果管理员希望对内网用户进行区分，让不同优先级的用户从不同链路进行转发；或者管理员想根据不同的应用来区分流量的转发链路，这些都不是我们通过目的地址查找路由能完成的。我们需要更灵活的选路机制，比如通过报文的源 IP 地址、应用协议类型等来区分用户流量，再对不同的用户流量进行区别转发。这个时候就需要我们的策略路由选路出场了，敬请关注下期强叔侃墙！

策略路由选路

——多样策略择重选优，定向转发掌控先机

大家好，通过上一篇我们了解了就近选路的相关原理和应用场景，同时也明确了就近选路的一些使用限制，了解到只凭借目的 IP 地址查找路由的选路方式缺乏灵活，适用的场景比较单一。实际场景中，网络环境复杂，单一的通过目的 IP 地址查找路由的选路方式无法达到网络管理员的要求，所以，有着多样匹配条件的策略路由加入了选路的行列。

说到策略路由选路，强叔先想到了策略路由早期在中国最大的应用莫过于在电信网通互联互通中的应用。电信网通分家之后出现了中国特色的网络环境，就是南电信，北网通（现在合并到联通）。在网络只有单出口的条件下会出现电信用户访问网通的服务较慢，网通用户访问电信的服务也较慢的问题。此时，人们就想到了企业网络双出口方案——网络出口同时接入到电信和网通。双出口方案的普及使得策略路由有了用武之地！通过在企业出口网关设备上配置策略路由，成功的实现了电信流量走电信出口，网通流量走网通出口。

策略路由是如何实现电信、网通流量的科学分流呢？我们还是先从什么是策略路由开始说起。

策略路由的概念

所谓策略路由，顾名思义，即是根据一定的策略进行报文转发。而策略是人为制定的，因此策略路由是一种比传统的按照目的地址选路更灵活的选路机制。在防火墙上配置策略路由后，防火墙首先会根据策略路由配置的规则对接收的报文进行过滤，匹配成功则按照一定的转发策略进行报文转发。其中“配置的规则”即是需要定义匹配条件，一般是通过 ACL 来定义匹配条件；而“一定的转发策略”则是需要根据匹配条件执行相关的动作。由此可以推断策略路由由以下两部分组成，如下：

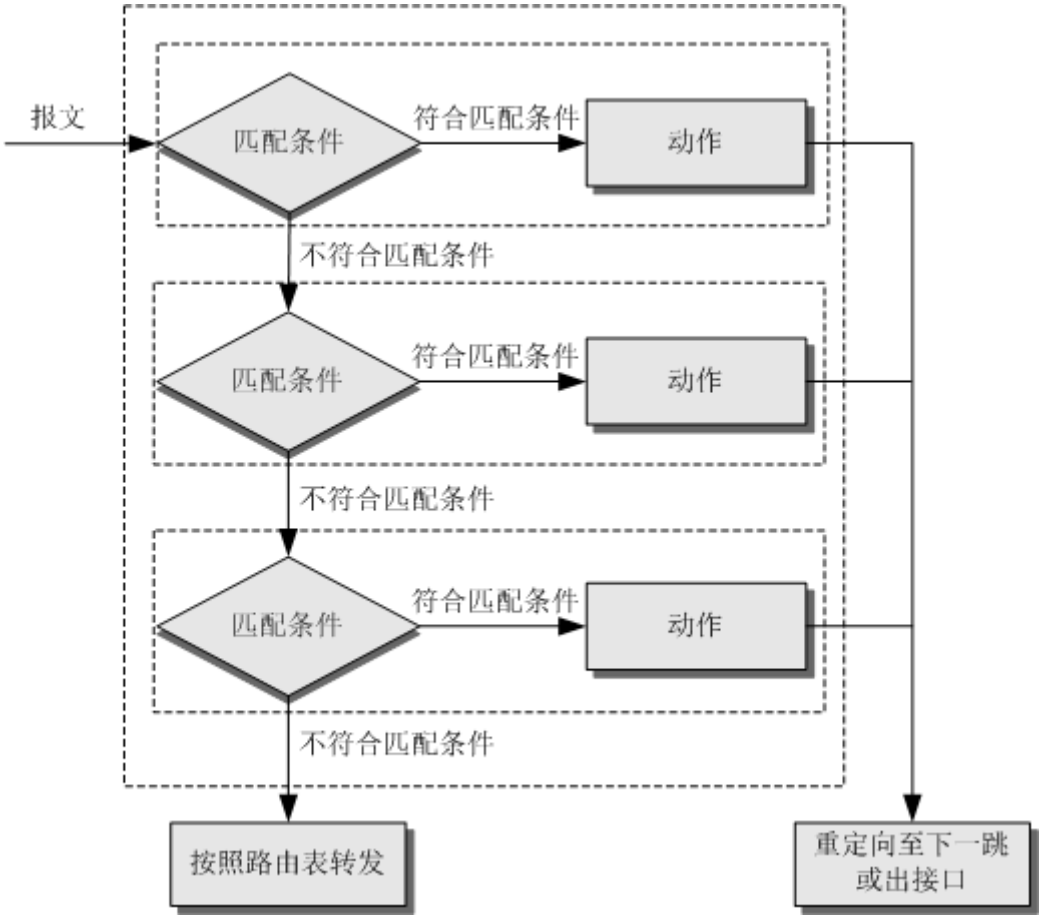
- **匹配条件（通过ACL定义）**

  用于区分将要做策略路由的流量。匹配条件包括：报文源 IP 地址、目的 IP 地址、协议类型、应用类型等，不同的防火墙可以设置的匹配条件略有不同。在一条策略路由规则中，可以包含多个匹配条件，各匹配条件之间是“与”的关系，报文必须同时满足所有匹配条件，才可以执行后续定义的转发动作。

- **动作**

对符合匹配条件的流量采取的动作，包括指定出接口和下一跳。

当有多条策略路由规则时，防火墙会按照匹配顺序，先寻找第一条规则，如果满足第一条策略路由规则的匹配条件，则按照指定动作处理报文。如果不满足第一条规则的匹配条件，则会寻找下一条策略路由规则。如果所有的策略路由规则的匹配条件都无法满足，报文按照路由表进行转发，策略路由的匹配是在报文查找路由表之前完成，也就是说策略路由比路由的优先级高。如下图所示。

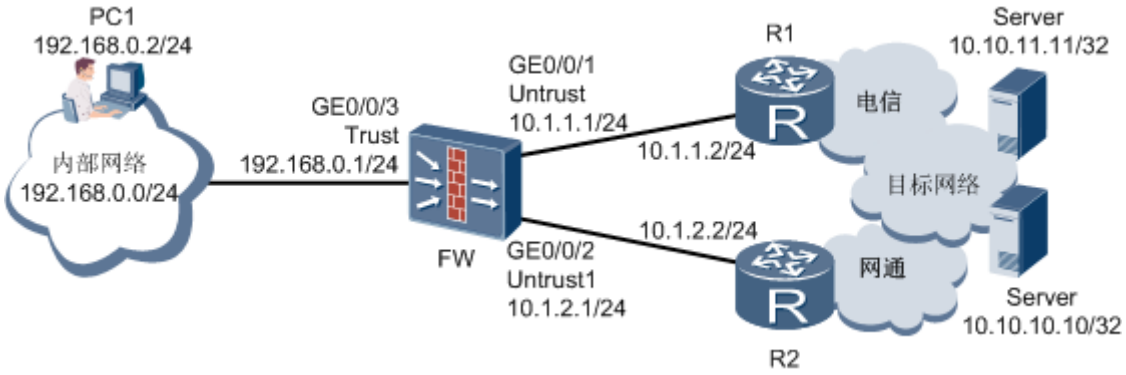


此外，如果策略路由指定的出接口或下一跳 Down 或不可达，那么报文将以报文中的目的地址为依据来查找路由表进行转发。

说完策略路由的基本原理后，现在我们回过头来看看策略路由是如何实现电信流量从电信转发，网通流量从网通转发的？

基于目的 IP 地址的策略路由

我们通过一个组网环境来验证这种策略路由的效果，组网如下：



组网中 FW 为企业出口网关，通过两条链路连接到 Internet，其中经过 R1 的链路为电信的线路；经过 R2 的链路为网通的线路。现在我们要让企业用户访问 10.10.11.11/32 这个 Internet 服务从电信线路转发，而 10.10.10.10/32 这个服务从网通线路转发。

如果在 FW 上配置两条缺省路由，企业用户在访问 10.10.11.11/32 和 10.10.10.10/32 这两个服务时，经过验证发现都是经过 R1 这条链路进行转发的，这个我们在上一期就近选路中就说过，缺省路由的选路是通过源 IP 地址+目的 IP 地址的 HASH 算法来计算报文选择的出口链路，无法控制访问 10.10.11.11/32 的流量从电信线路转发，访问 10.10.10.10/32 的流量从网通线路转发的要求。

下面我们在 FW 上配置策略路由，看看实验效果如何，配置如下(我们以 USG2000/5000 系列防火墙为例)：

1、根据报文目的地址设置匹配条件

```
acl number 3000
 rule 5 permit ip destination 10.10.11.11 0
acl number 3001
 rule 5 permit ip destination 10.10.10.10 0
```

2、配置策略路由

```
policy-based-route test permit node 10
  if-match acl 3000                              //应用匹配条件
  apply ip-address next-hop 10.1.1.2             //配置动作，重定向至电信下一跳
 policy-based-route test permit node 20
  if-match acl 3001                              //应用匹配条件
  apply ip-address next-hop 10.1.2.2             //配置动作，重定向至网通下一跳
```

3、应用策略路由

```
interface GigabitEthernet0/0/3
 ip address 192.168.0.1 255.255.255.0
 ip policy-based-route test                      //在入接口应用策略路由
```

配置完成后，我们在在 PC 上 ping 10.10.11.11 和 10.10.10.10 两个地址，能正常 ping 通。同时在 FW 上查看会话表达的详细信息，显示如下：

```
<FW1>display firewall session table verbose
11:08:22  2014/09/04
 Current Total Sessions : 1
  icmp  VPN:public --> public
  Zone: trust--> trust  TTL: 00:00:20  Left: 00:00:16
  Interface: GigabitEthernet0/0/1  NextHop: 10.1.1.2  MAC: 54-89-98-1d-74-24
  <--packets:1 bytes:60    -->packets:1 bytes:60
  192.168.0.2:54999-->10.10.11.11:2048

<FW1>display firewall session table verbose
11:08:56  2014/09/04
 Current Total Sessions : 1
  icmp  VPN:public --> public
  Zone: trust--> trust  TTL: 00:00:20  Left: 00:00:17
  Interface: GigabitEthernet0/0/2  NextHop: 10.1.2.2  MAC: 54-89-98-ea-53-c9
  <--packets:1 bytes:60    -->packets:1 bytes:60
  192.168.0.2:63959-->10.10.10.10:2048
```

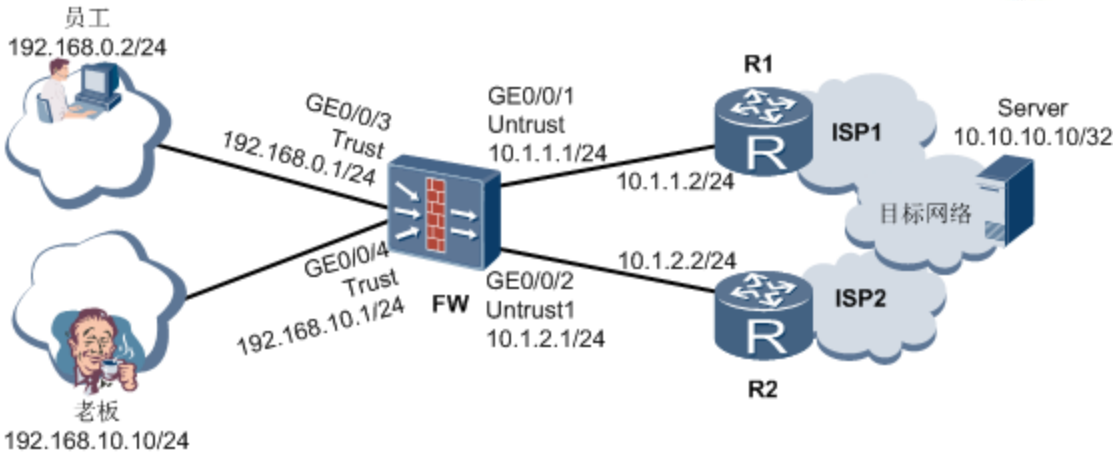
通过显示信息可以看到去往 10.10.11.11 的报文是从 FW 的 GE0/0/1 接口转发的，下一跳为 R1 与 FW 相连的接口地址；而去往 10.10.10.10 的报文是从 FW 的 GE0/0/2 接口转发的，下一跳为 R2 与 FW 相连的接口地址，从而达到访问 10.10.11.11/32 的流量从电信线路转发，而访问 10.10.10.10/32 的流量从网通线路转发的要求。

看到这里，可能大家会说，在上一篇介绍的就近访问也能达到这个要求，对！的确是如此。因为就近选路中缺省路由+明细路由的选路方式是根据目的地址进行报文的转发，而上面配置的策略路由也是以报文目的地址为条件制定转发策略，所以能够完成同样的需求。但策略路由更多的体现在人为的控制方面，而传统的按目的地址路由只能由系统内置的 HASH 算法或者是系统的内部实现来决定报文的转发。

事实上，传统的路由只能根据报文的目的地址为用户提供比较单一的路由方式，它更多的是解决网络报文的转发问题，而不能提供更灵活的服务。策略路由则不同，它使网络管理者不仅能够根据目的地址，而且能够根据报文源 IP 地址、协议类型、应用类型或者其它条件来选择转发路径，所以说策略路由有着比传统路由协议对报文的更强控制能力。

基于源 IP 地址的策略路由

如果说上面的应用与就近访问方式还有一些交集，那我们来看看策略路由选路的另外一个应用。大家都知道，光纤到户是目前网络的发展方向，但光纤的费用在今天的中国并不便宜，于是很多地方都采用了光纤加 ADSL 的方式，然而这样就出现了两条速度不同的线路接入互联网。通过策略路由可以让一部分优先级较高的用户走光纤，另一部分级别低的用户机走 ADSL。我们还是以事实为依据，模拟组网环境如下。



组网中 FW 为企业出口网关，通过从不同的 ISP 连接了两条链路到 Internet，其中经过 R1 的链路带宽速率较高，假设为 10Mbit/s；经过 R2 的链路带宽速率较小，为 2Mbit/s。为保证企业老板访问 Internet 的效果，让其访问流量从经过 R1 的链路进行转发，而员工的访问流量从经过 R2 的链路转发。

想要完成上述要求，通过目的地址查找路由的方式是无法完成的，而通过策略路由设置源 IP 地址为匹配条件很轻松就能解决此问题。在 FW 上的配置如下(我们以 USG2000/5000 系列防火墙为例)：

1、根据报文源 IP 地址设置匹配条件

```
acl number 3000
 rule 5 permit ip source 192.168.10.0 0.0.0.255
acl number 3001
 rule 5 permit ip source 192.168.0.0 0.0.0.255
```

2、配置策略路由

```
policy-based-route boss permit node 10
  if-match acl 3000                                    //应用匹配条件
  apply ip-address next-hop 10.1.1.2                   //配置动作，重定向下一跳为 R1
 policy-based-route employee permit node 10
  if-match acl 3001                                    //应用匹配条件
  apply ip-address next-hop 10.1.2.2                   //配置动作，重定向下一跳为 R2
```

3、应用策略路由

```
interface GigabitEthernet0/0/3
 ip address 192.168.0.1 255.255.255.0
 ip policy-based-route employee                        //在入接口应用策略路由
interface GigabitEthernet0/0/4
 ip address 192.168.10.1 255.255.255.0
 ip policy-based-route boss                            //在入接口应用策略路由
```

配置完成后，分别在老板和员工的 PC 上 ping Internet 上的 Server 地址 10.10.10.10，并在

FW 上查看会话表详细信息，显示如下：

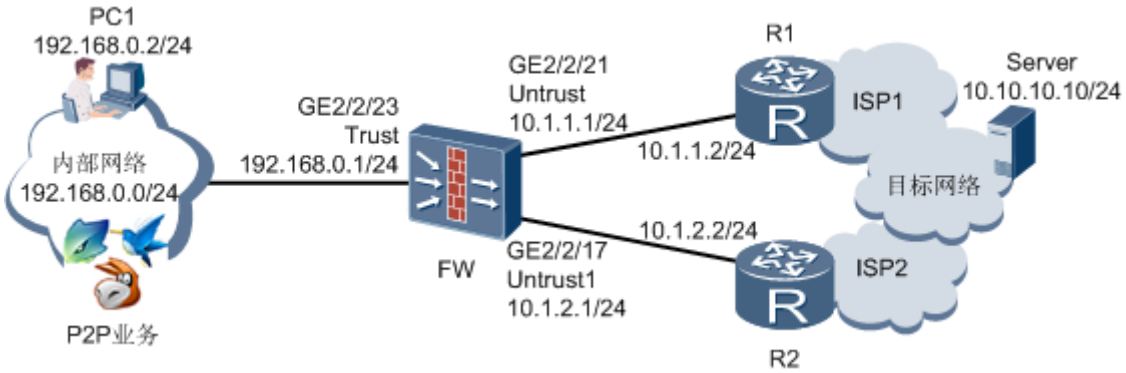
```
<FW1>display firewall session table verbose
10:23:41  2014/09/05
 Current Total Sessions : 1
  icmp  VPN:public --> public
  Zone: trust--> trust  TTL: 00:00:20  Left: 00:00:15
  Interface: GigabitEthernet0/0/1  NextHop: 10.1.1.2  MAC: 54-89-98-1d-74-24
  <--packets:1 bytes:60    -->packets:1 bytes:60
  192.168.10.10:47646-->10.10.10.10:2048

<FW1>display firewall session table verbose
10:24:03  2014/09/05
 Current Total Sessions : 1
  icmp  VPN:public --> public
  Zone: trust--> trust  TTL: 00:00:20  Left: 00:00:14
  Interface: GigabitEthernet0/0/2  NextHop: 10.1.2.2  MAC: 54-89-98-ea-53-c9
  <--packets:1 bytes:60    -->packets:1 bytes:60
  192.168.0.2:53022-->10.10.10.10:2048
```

显示信息中，老板（192.168.10.10）访问 Server 的流量是从 R1（10.1.1.2）的链路进行转发的；而员工（192.168.0.2）访问 Server 的流量是从 R2（10.1.2.2）的链路转发的。从而完成了优先级较高的用户走高速链路，而低级别用户从低速链路转发的需求。

基于应用的策略路由

前面介绍的这些是策略路由选路的一些传统的应用，在现网中，策略路由选路还有一种常用的用法是与应用有关。大家都知道，网络中各种应用层出不穷，其中一些大流量的应用，如 P2P、在线视频等占用了企业大量的出口带宽，严重影响了正常业务流量的转发效果。防火墙的策略路由能与应用识别功能相结合，以流量的应用类型为匹配条件，实现基于应用的策略转发，这就是我们所熟知的基于应用的策略路由。

下面强叔也通过一个实际的组网环境来验证一下基于应用的策略路由的实际效果。



组网环境中，FW 为企业出口网关，通过从 ISP1 和 ISP2 获得两条链路与 Internet 相连，其中 ISP2 提供的链路上下行的带宽对称，链路状态稳定，为企业正常业务流量主要转发链路；

ISP1 提供的链路上下行的带宽不对称，网速较慢，但租用价格低廉，可提供给一些大流量应用（图中为 P2P）转发的链路。

我们通过“比特精灵”工具模拟 P2P 业务，在 Server 上模拟 P2P 服务器，在企业用户 PC1 上模拟 P2P 客户端，同时使用 Ping 模拟正常业务。

现在 FW 上配置基于应用的策略路由，让 P2P 应用的流量从 GE2/2/21 出接口转发，而正常的流量直接通过路由从 GE2/2/17 出接口转发，配置命令如下(我们以 USG9000 系列防火墙为例)：

1、根据报文源 IP 地址设置匹配条件

```
acl number 3000
 rule 5 permit ip source 192.168.0.0 0.0.0.255
```

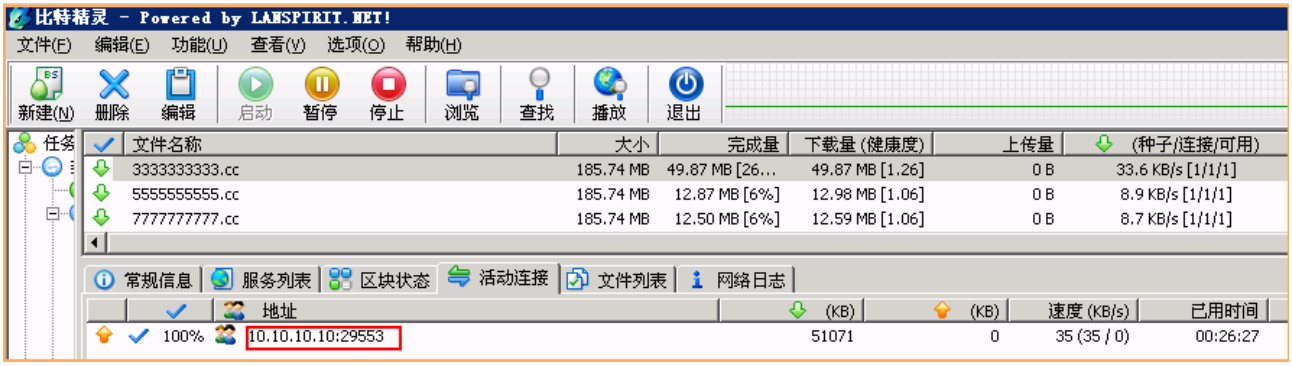
2、配置策略路由

```
traffic classifier p2p operator or
 if-match acl 3000 category p2p                          //对用户的 P2P 应用设置为匹配条件
traffic behavior p2p
 redirect ip-nexthop 10.1.1.2 interface GigabitEthernet2/2/21   //重定向出接口和下一跳
traffic policy p2p
 share-mode
 classifier p2p behavior p2p
```

3、应用策略路由

```
interface GigabitEthernet2/2/23
 ip address 192.168.0.1 255.255.255.0
 traffic-policy p2p inbound                              //在入接口应用策略路由
```

配置完成后，我们在 PC1 上开启“比特精灵”客户端下载功能，截图如下：



然后我们在 FW 上查看会话表，显示如下：

```
[USG9500]display firewall session table verbose
11:49:29  2014/09/09
 Current total sessions: 2
 tcp VPN: public --> public
 Zone: trust --> untrust Slot: 3 CPU: 3 TTL: 00:00:05 Left: 00:00:02
 Interface: GigabitEthernet2/2/21 Nexthop: 10.1.1.2
 <--packets: 0 bytes: 0 -->packets: 2 bytes: 96
 192.168.0.2:1712 --> 10.10.10.10:29553

 tcp VPN: public --> public
 Zone: trust --> untrust Slot: 3 CPU: 3 TTL: 00:00:05 Left: 00:00:02
 Interface: GigabitEthernet2/2/21 Nexthop: 10.1.1.2
 <--packets: 0 bytes: 0 -->packets: 2 bytes: 96
 192.168.0.2:1711 --> 10.10.10.10:29553
```

通过显示信息可以看到，访问的目的地址和端口与“比特精灵”客户端上的显示一致，都是10.10.10.10:29553，并且这部分流量是从出接口为 GE2/2/21、下一跳为 10.1.1.2 的链路进行转发的，与策略路由重定向的出接口和下一跳一致，说明基于 P2P 应用的策略路由应用成功。

下面我们在 PC1 上 ping 10.10.10.10 这个 Server 的地址。

此时，我们再在 FW 上查看会话表，显示如下：

```
[USG9500]display firewall session table verbose
14:25:14  2014/09/09
 Current total sessions: 1
 icmp VPN: public --> public
 Zone: trust --> untrust Slot: 3 CPU: 3 TTL: 00:00:20 Left: 00:00:17
 Interface: GigabitEthernet2/2/17 Nexthop: 10.1.2.2
 <--packets: 4 bytes: 240 -->packets: 4 bytes: 240
 192.168.0.2:768 --> 10.10.10.10:2048
```

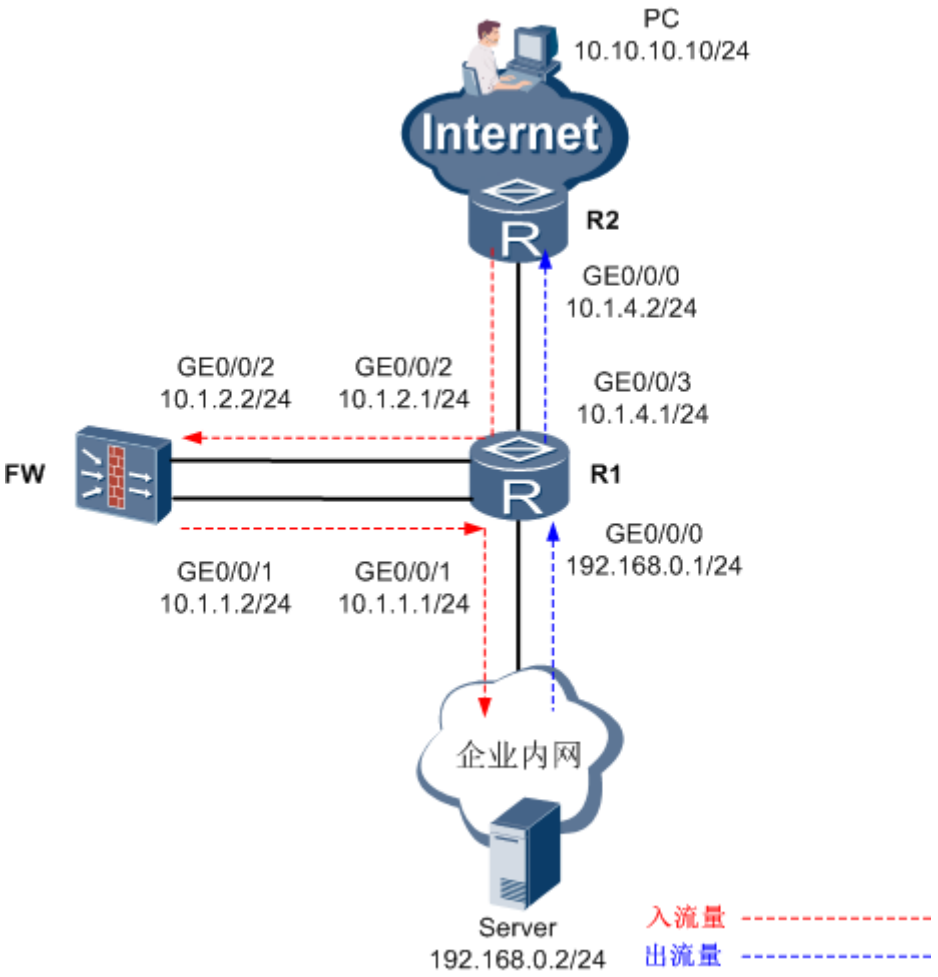
显示信息中显示这部分流量是从出接口 GE2/2/17、下一跳为 10.1.2.2 的链路进行转发的。说明我们的正常业务流量是从 ISP2 提供的链路进行转发的，达到了预期要求。

综合上面几个策略路由选路的应用，可以看到策略路由选路的灵活应用关键在于匹配条件的灵活多样，不同场景匹配不同的条件，如上面的三个应用包含的匹配条件分别为目的 IP 地址、源 IP 地址和应用类型。此外，还有许多匹配条件会比较常用，如用户、协议类型等，因为配置方法基本相同，这里就不一一介绍。

旁路组网下的策略路由选路

在现网中，策略路由的另外一种应用却不得不提，就是防火墙旁挂企业出口时，通过策略路由引流到防火墙进行安全防护。当然此时的策略路由不是在防火墙上进行配置的，但这种应用场景在许多企业出口和数据中心出口都经常用到，为此，我们也对这种应用做一个介绍。

首先我们还是通过实际组网环境来验证这种场景，如下图所示。



组网中企业出口为路由器 R1，防火墙 FW 旁挂在出口路由器 R1 上。当外网用户访问内网服务器时，流量从出口路由器引流到防火墙进行安全防护后，再转发到内网服务器。

此组网中，策略路由是在出口路由器 R1 上配置的，路由器策略路由的配置思路与上面介绍的防火墙的策略路由配置思路一致，都是先定义匹配条件（本地为目的 IP 地址）、设置动作（重定向出接口或下一跳），然后在入接口上应用。此组网中的防火墙除了对引入的流量进行安全防护外，还需要把流量回注到出口路由器 R1 上。安全防护的内容强叔在前面的贴子中已经介绍过，这里主要说下回注的问题。回注其实很简单，这里介绍静态路由和 OSPF 两种回注方法。

- **静态路由配置方法**

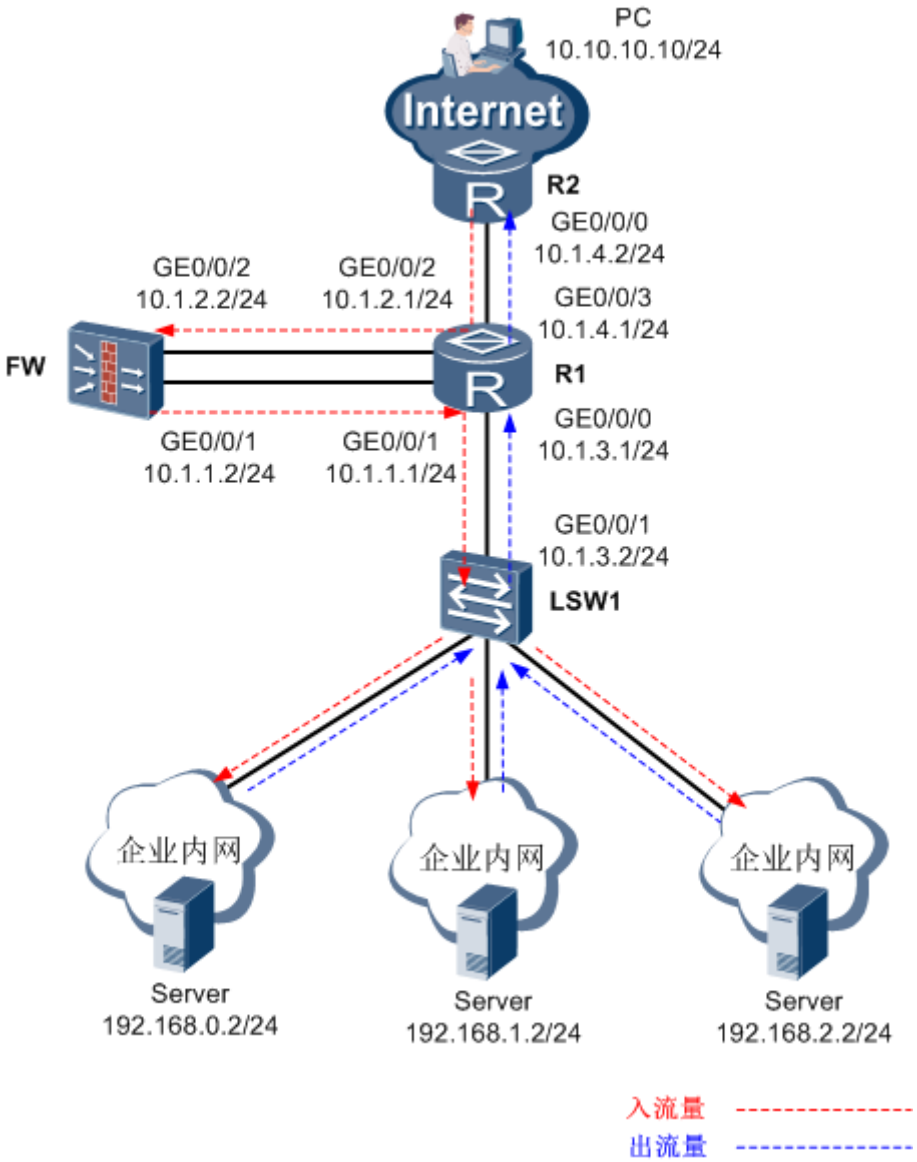
如下表为静态路由的配置方式，其中只列出策略路由和静态路由的配置。

R1	FW
acl number 3000 rule 5 permit ip destination 192.168.0.2 0 policy-based-route in permit node 10	ip route-static 192.168.0.0 255.255.255.0 10.1.1.1

if-match acl 3000 apply ip-address next-hop 10.1.2.2 interface GigabitEthernet0/0/3 ip address 10.1.4.1 255.255.255.0 ip policy-based-route in ip route-static 10.10.10.0 255.255.255.0 10.1.4.2	

- **OSPF路由配置方法：**

当接入用户网络较多时考虑使用此种配置方式，方便管理员维护。



如下表为 OSPF 的配置命令，其中只列出策略路由和 OSPF 的配置。

R1	LSW1	FW
acl number 3000 rule 5 permit ip destination 192.168.0.2 0 policy-based-route in permit node 10 if-match acl 3000 apply ip-address next-hop 10.1.2.2 interface GigabitEthernet0/0/3 ip address 10.1.4.1 255.255.255.0 ip policy-based-route in ospf 1 area 0.0.0.0 network 10.1.1.0 0.0.0.255 network 10.1.3.0 0.0.0.255 ospf 2 area 0.0.0.0 network 10.1.4.0 0.0.0.255 network 10.1.2.0 0.0.0.255	ospf 1 area 0.0.0.0 network 192.168.0.0 0.0.0.255 network 192.168.1.0 0.0.0.255 network 192.168.2.0 0.0.0.255	ospf 1 import-route ospf 2 area 0.0.0.0 network 10.1.1.0 0.0.0.255 ospf 2 import-route ospf 1 area 0.0.0.0 network 10.1.2.0 0.0.0.255

通过 OSPF 回注的配置相对复杂一些，是使用 OSPF 双进程，双进程在 R1 上对上下行流量进行隔离，当流量通过策略路由引流到 FW 后，再在 FW 上通过 OSPF 两个进程互相引入，可以使 OSPF 两个进程中的路由能够互相学习到。

R1 上策略路由完成了流量的引流，FW 上静态路由或 OSPF 完成了流量的回注。配置完成后，通过在外网 PC（10.10.10.10）上 Tracert 内网 Server 的地址 192.168.0.2，可以看到如下路线（以静态路由组网配置为例）：

```
PC>tracert 192.168.0.2

traceroute to 192.168.0.2, 8 hops max
(ICMP), press Ctrl+C to stop
 1  10.10.10.1   16 ms  <1 ms  <1 ms
 2  10.1.4.1     31 ms  31 ms  32 ms
 3  10.1.2.2     140 ms  63 ms  47 ms
 4  10.1.1.1     94 ms  62 ms  47 ms
 5  192.168.0.2   63 ms  78 ms  94 ms
```

路径信息中显示，访问流量经过 FW 后，再回到 R1，最后到目标 Server，达到了预期访问效果。

策略路由选路其实就是对符合匹配条件的流量进行选路，重新选定出接口和下一跳。这就要

求管理员对网络现状有充分的了解，能根据网络现状选择合适的匹配条件。比如清楚的知道多条出口链路的优异，就能让企业重要客户或重要业务的流量从优先级高的链路进行转发。所以说灵活的应用策略路由，为管理员规划网络提供了更多的手段。

至此，防火墙出口选路篇收官结束，也意味着本季的强叔侃墙将告一段落，感谢大家这么长时间的关注和支持！

关于华为防火墙，你还想了解更多吗？请点击【强叔侃墙】汇总贴

如有任何问题，请留言至xuhuiyang.xu@huawei.com

本期策划：徐慧洋

本期作者：卢宏旺、金德胜

美术编辑：潘艺从