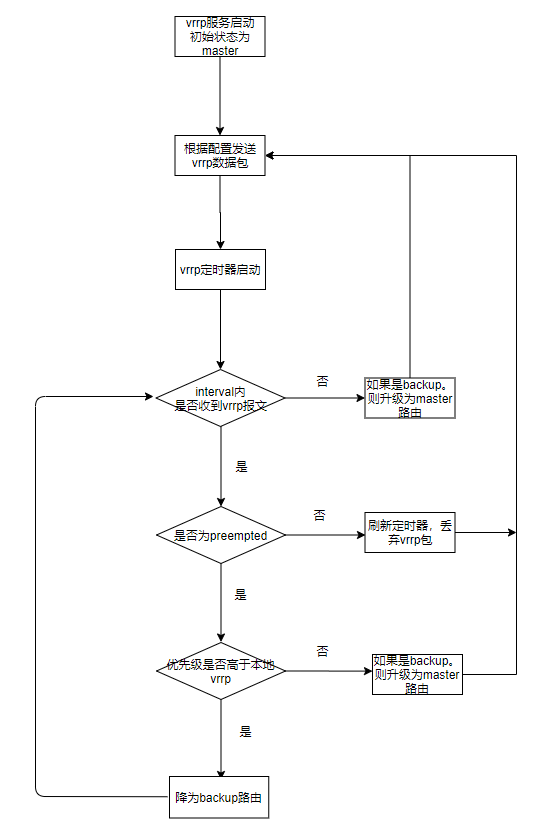
vrrp设计说明文档

1、实现原理及流程

vrrp目前内部实现路程图如下，主要逻辑是通过定时器去控制主备切换的时机，接下来具体描述一下vrrp切换的原理。

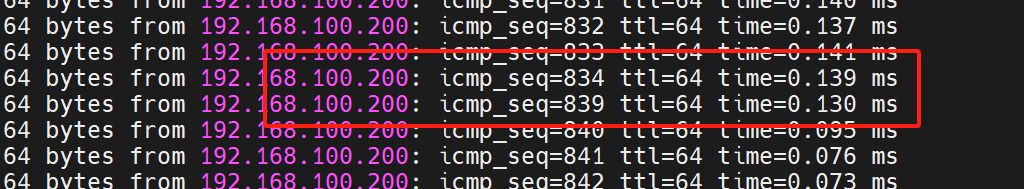


1. A、B两台路由配置启动vrrp服务之后，默认会将自身状态设置为master路由然后启动定时器单播或者组播发送自身的vrrp配置信息。
2. A、B同时会收到对端的vrrp数据报文，然后进行解析，查看vrrp的优先级，当A发送的优先级高于B自身优先级时，再查看配置是否为preempted模式，如果是，那么B路由降为backup状态。
3. B路由降为backup状态之后会停止发送vrrp数据包的定时器，同时会根据vrrp的interval来启动一个接收vrrp数据包的定时器。
4. 在interval区间如果没有接收到新的vrrp数据包，那么B路由则认定对端A路由链路故障，重新将自身升级为master，然后进行第（1）步骤流程。
5. 在interval区间如果收到了新的vrrp数据包，那么B会将解析之后的优先级和自身优先级进行比对，如果自身优先级高于A发送来的优先级，那么也重新将自身升级为master，然后进行第（1）步骤流程。
6. 在interval区间如果收到了新的vrrp数据包，但是优先级高于B自身优先级，那么则刷新定时器以及老化时间，此时对端路由状态为active并且优先级高于自身。

2、vrrp切换量化展示

通过ping命令一直ping虚ip，间隔1ms，手动down master路由的接口，查看vrrp主备切换的时间。



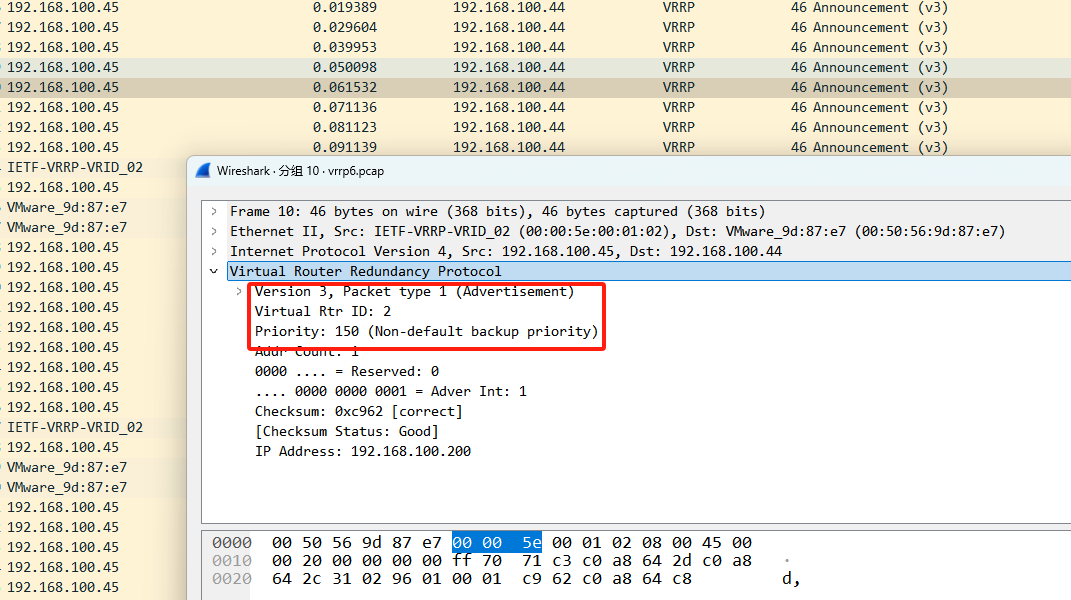


可以看到，手动down master路由之后，丢失了5个ping包，也就是切换间隔为5ms。

经过多次测试，发现这一值平均在5ms之内，所以vrrp切换的极限为5ms。

3、vrrp心跳交互报文

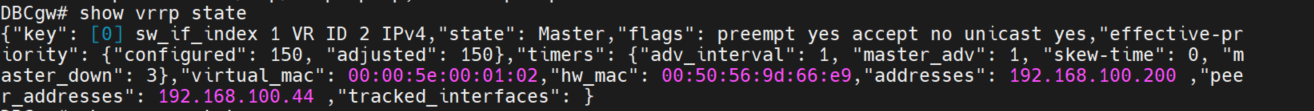
主备路由之间通过vrrp报文交互，使用version为3，master路由向对端单播自己的vrrp协议信息，accept\_mode模式下对端接收到vrrp报文信息之后进行优先级比对，如果优先级高于此，那么通过preempted进行抢占成为master路由，如果优先级低于vrrp报文优先级，那么降为backup路由。





4、vrrp state相关信息说明

启动vrrp服务之后可以实时查看当前路由的vrrp状态，在vpp上进行show vrrp state可以查看具体的内容，如下图所示。



可以看到当前的详细信息。

key：VRRP实例的键值，用于标识特定的VRRP实例。

sw\_if\_index：接口索引，表示VRRP实例所在的网络接口。

VR ID：虚拟路由器的标识符，用于区分不同的VRRP实例。

IPv4：表示VRRP实例的IP版本，这里是IPv4。

state：VRRP实例的当前状态，可以是以下之一：

Master：表示此设备是主路由器，负责转发数据包。

Backup：表示此设备是备份路由器，处于待命状态，准备在主路由器故障时接管。

flags：VRRP实例的标志位，表示一些特定的行为：

preempt：抢占模式，如果设置为yes，表示此设备可以抢占成为主路由器。

accept：接受模式，如果设置为no，表示此设备不接受来自其他路由器的VRRP通告。

unicast：单播模式，如果设置为yes，表示此设备使用单播而不是多播来发送VRRP通告。

effective-priority：VRRP实例的有效优先级：

configured：配置的优先级，表示在配置文件中设置的优先级。

adjusted：调整后的优先级，表示实际使用的优先级，可能因某些条件（如跟踪接口的状态）而调整。

timers：VRRP实例的计时器设置：

adv\_interval：通告间隔，表示主路由器发送VRRP通告的频率（秒）。

master\_adv：主路由器通告计时器，表示主路由器发送通告的时间间隔（秒）。

skew-time：偏移时间，用于计算主路由器的故障检测时间。

master\_down：主路由器故障计时器，表示在检测到主路由器故障后，备份路由器成为主路由器的等待时间（秒）。

virtual\_mac：虚拟MAC地址，VRRP实例的虚拟MAC地址，用于标识虚拟路由器。

hw\_mac：硬件MAC地址，设备的实际MAC地址。

addresses：虚拟IP地址，VRRP实例的虚拟IP地址。

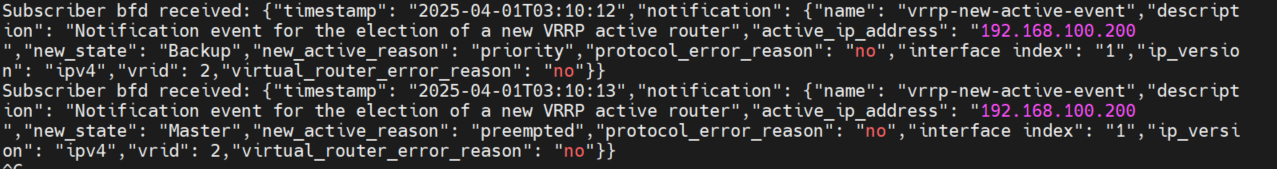
peer\_addresses：对等体IP地址，与VRRP实例通信的对等体的IP地址。

tracked\_interfaces：被跟踪的接口，表示此VRRP实例正在跟踪的网络接口。如果这些接口的状态发生变化，可能会影响VRRP实例的优先级或状态。

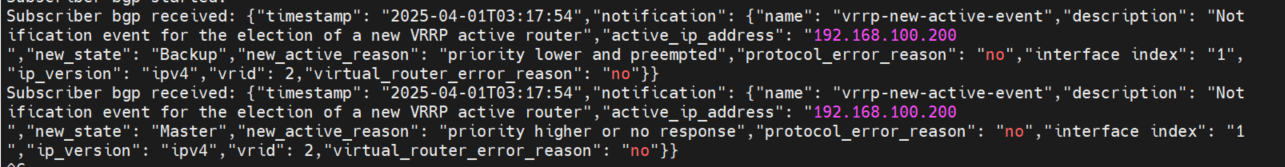
5、vrrp notification机制以及对应的参数说明

其他业务进程如果依赖于vrrp那么会对vrrp服务进行订阅，当vrrp状态切换的时候会发布notification数据到所有订阅的业务或者进程中。目前使用的是zmq机制进行这种消息的订阅-发布，端口为5555。

例：启动了两个业务进程，bfd和bgp，全部订阅vrrp服务，当vrrp切换的时候查看收到的notification信息。

bfd：  


bgp：



两个业务进程分别收到了vrrp发布的状态切换信息，然后解析之后进行各自的业务处理。

timestamp：事件发生的时间戳，格式为ISO 8601，例如2025-04-01T03:17:54。

name：事件的名称，例如vrrp-new-active-event，表示VRRP新活动路由器选举的事件。

description：事件的描述，提供关于事件的详细信息，例如Notification event for the election of a new VRRP active router，表示这是一个关于新VRRP活动路由器选举的通知事件。

active\_ip\_address：虚拟路由器的IP地址，例如192.168.100.200，表示当前活动的虚拟路由器的IP地址。

new\_state：VRRP实例的新状态，例如Master，表示实例现在处于主路由器状态。

new\_active\_reason：状态变为活动的原因，例如priority higher or no response，表示因为优先级更高或没有响应而成为活动路由器。

protocol\_error\_reason：协议错误的原因，例如no，表示没有协议错误。

interface index：接口索引，表示VRRP实例所在的网络接口，例如1。

ip\_version：IP版本，例如ipv4，表示VRRP实例的IP版本。

vrid：虚拟路由器的标识符，例如2，用于区分不同的VRRP实例。

virtual\_router\_error\_reason：虚拟路由器错误的原因，例如no，表示没有虚拟路由器错误