

VRRP技术白皮书

关键词：VRRP、虚拟路由器

摘 要：本文介绍了VRRP的基本原理和典型应用，以及H3C公司VRRP特性解决方案的特点和组网情况。

缩略语：

缩略语	英文全名	中文解释
VRRP	Virtual Router Redundancy Protocol	虚拟路由器冗余协议
NQA	Network Quality Analyzer	网络质量分析
BFD	Bidirectional Forwarding Detection	双向转发检测
IRDP	ICMP Router Discovery Protocol	ICMP路由发现协议
VRID	Virtual Router ID	虚拟路由器号

目 录

1 概述	3
1.1 产生背景	3
1.2 技术优点	4
2 VRRP协议介绍	4
2.1 相关术语	4
2.2 虚拟路由器简介	5
2.3 VRRP工作过程	6
2.3.1 Master路由器的选举	6
2.3.2 Master路由器状态的通告	7
2.3.3 认证方式	8
3 Comware实现的技术特色	8
3.1 监视上行链路	8
3.2 Backup监视Master工作状态	9
4 典型组网案例	9
4.1 主备备份	9
4.2 负载分担	10
4.3 Master使用BFD/NQA监视上行链路	11
4.4 Backup使用BFD监视Master状态	12
5 附录	13
5.1 参考文献	13

1 概述

1.1 产生背景

随着Internet的发展，人们对网络可靠性的要求越来越高。特别是对于终端用户来说，能够实时与网络其他部分保持联系是非常重要的。一般来说，主机通过设置默认网关来与外部网络联系，如图1所示：

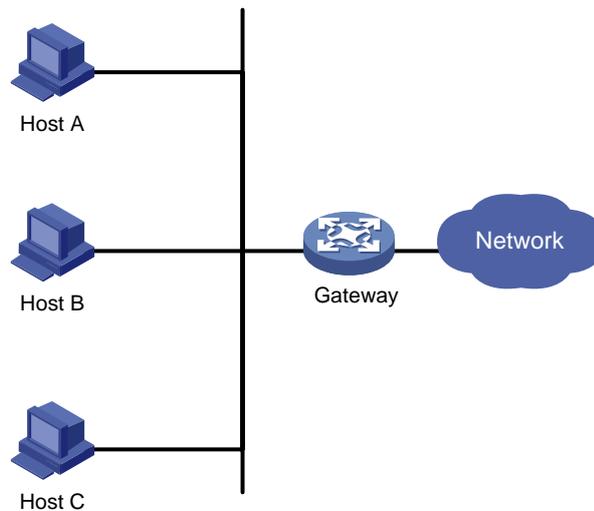


图1 常用局域网组网方案

主机将发送给外部网络的报文发送给网关，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，但是当网关坏掉时，主机与外部的通信就会中断。要解决网络中断的问题，可以依靠再添加网关的方式解决，不过由于大多数主机只允许配置一个默认网关，此时需要网络管理员进行手工干预网络配置，才能使得主机使用新的网关进行通信；有时，人们运用动态路由协议的方法来解决网络出现故障这一问题，如运行RIP、OSPF等，或者使用IRDP。然而，这些协议由于配置过于复杂，或者安全性能不好等原因都不能满足用户的需求。

为了更好地解决网络中断的问题，网络开发者提出了VRRP，它既不需要改变组网情况，也不需要主机上做任何配置，只需要在相关路由器上配置极少的几条命令，就能实现下一跳网关的备份，并且不会给主机带来任何负担。和其他方法比较起来，VRRP更加能够满足用户的需求。

1.2 技术优点

VRRP是一种容错协议，它保证当主机的下一跳路由器出现故障时，由另一台路由器来代替出现故障的路由器进行工作，从而保持网络通信的连续性和可靠性。

VRRP具有如下优点：

- 简化网络管理。在具有多播或广播能力的局域网（如以太网）中，借助 VRRP 能在某台设备出现故障时仍然提供高可靠的缺省链路，有效避免单一链路发生故障后网络中断的问题，而无需修改动态路由协议、路由发现协议等配置信息，也无需修改主机的默认网关配置。
- 适应性强。VRRP 报文封装在 IP 报文中，支持各种上层协议。
- 网络开销小。VRRP 只定义了一种报文——VRRP 通告报文，并且只有处于 Master 状态的路由器可以发送 VRRP 报文。

2 VRRP协议介绍

2.1 相关术语

- 虚拟路由器：由一个 Master 路由器和多个 Backup 路由器组成。主机将虚拟路由器当作默认网关。
- VRID：虚拟路由器的标识。有相同 VRID 的一组路由器构成一个虚拟路由器。
- Master 路由器：虚拟路由器中承担报文转发任务的路由器。
- Backup 路由器：Master 路由器出现故障时，能够代替 Master 路由器工作的路由器。
- 虚拟 IP 地址：虚拟路由器的 IP 地址。一个虚拟路由器可以拥有一个或多个 IP 地址。
- IP 地址拥有者：接口 IP 地址与虚拟 IP 地址相同的路由器被称为 IP 地址拥有者。
- 虚拟 MAC 地址：一个虚拟路由器拥有一个虚拟 MAC 地址。虚拟 MAC 地址的格式为 00-00-5E-00-01-{VRID}。通常情况下，虚拟路由器回应 ARP 请求使用的是虚拟 MAC 地址，只有虚拟路由器做特殊配置的时候，才回应接口的真实 MAC 地址。
- 优先级：VRRP 根据优先级来确定虚拟路由器中每台路由器的地位。

- 非抢占方式：如果 Backup 路由器工作在非抢占方式下，则只要 Master 路由器没有出现故障，Backup 路由器即使随后被配置了更高的优先级也不会成为 Master 路由器。
- 抢占方式：如果 Backup 路由器工作在抢占方式下，当它收到 VRRP 报文后，会将自己的优先级与通告报文中的优先级进行比较。如果自己的优先级比当前的 Master 路由器的优先级高，就会主动抢占成为 Master 路由器；否则，将保持 Backup 状态。

2.2 虚拟路由器简介

VRRP将局域网内的一组路由器划分在一起，形成一个VRRP备份组，它在功能上相当于一台虚拟路由器，使用虚拟路由器号进行标识。以下使用虚拟路由器代替VRRP备份组进行描述。

虚拟路由器有自己的虚拟IP地址和虚拟MAC地址，它的外在表现形式和实际的物理路由器完全一样。局域网内的主机将虚拟路由器的IP地址设置为默认网关，通过虚拟路由器与外部网络进行通信。

虚拟路由器是工作在实际的物理路由器之上的。它由多个实际的路由器组成，包括一个Master路由器和多个Backup路由器。Master路由器正常工作时，局域网内的主机通过Master与外界通信。当Master路由器出现故障时，Backup路由器中的一台设备将成为新的Master路由器，接替转发报文的工作，如图2所示。

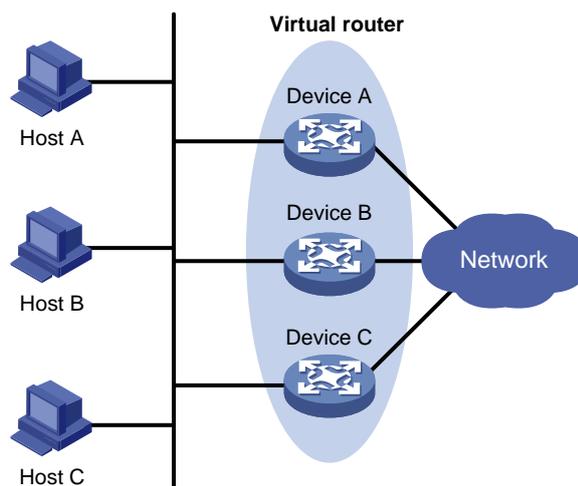


图2 虚拟路由器示意图

2.3 VRRP工作过程

VRRP的工作过程为：

- (1) 虚拟路由器中的路由器根据优先级选举出 **Master**。**Master** 路由器通过发送免费 ARP 报文，将自己的虚拟 MAC 地址通知给与它连接的设备或者主机，从而承担报文转发任务；
- (2) **Master** 路由器周期性发送 VRRP 报文，以公布其配置信息（优先级等）和工作状况；
- (3) 如果 **Master** 路由器出现故障，虚拟路由器中的 **Backup** 路由器将根据优先级重新选举新的 **Master**；
- (4) 虚拟路由器状态切换时，**Master** 路由器由一台设备切换为另外一台设备，新的 **Master** 路由器只是简单地发送一个携带虚拟路由器的 MAC 地址和虚拟 IP 地址信息的免费 ARP 报文，这样就可以更新与它连接的主机或设备中的 ARP 相关信息。网络中的主机感知不到 **Master** 路由器已经切换为另外一台设备。
- (5) **Backup** 路由器的优先级高于 **Master** 路由器时，由 **Backup** 路由器的工作方式（抢占方式和非抢占方式）决定是否重新选举 **Master**。

由此可见，为了保证**Master**路由器和**Backup**路由器能够协调工作，VRRP需要实现以下功能：

- **Master** 路由器的选举；
- **Master** 路由器状态的通告；
- 同时，为了提高安全性，VRRP 还提供了认证功能；

下面将从上述三个方面详细介绍VRRP的工作过程。

2.3.1 Master路由器的选举

VRRP根据优先级来确定虚拟路由器中每台路由器的角色（**Master**路由器或**Backup**路由器）。优先级越高，则越有可能成为**Master**路由器。

初始创建的路由器工作在**Backup**状态，通过VRRP报文的交互获知虚拟路由器中其他成员的优先级：

- 如果 VRRP 报文中 **Master** 路由器的优先级高于自己的优先级，则路由器保持在 **Backup** 状态；

- 如果 VRRP 报文中 Master 路由器的优先级低于自己的优先级，采用抢占工作方式的路由器将抢占成为 Master 状态，周期性地发送 VRRP 报文，采用非抢占工作方式的路由器仍保持 Backup 状态；
- 如果在一定时间内没有收到 VRRP 报文，则路由器切换为 Master 状态。

VRRP优先级的取值范围为0到255（数值越大表明优先级越高），可配置的范围是1到254，优先级0为系统保留给路由器放弃Master位置时候使用，255则是系统保留给IP地址拥有者使用。当路由器为IP地址拥有者时，其优先级始终为255。因此，当虚拟路由器内存在IP地址拥有者时，只要其工作正常，则为Master路由器。

2.3.2 Master路由器状态的通告

Master路由器周期性地发送VRRP报文，在虚拟路由器中公布其配置信息（优先级等）和工作状况。Backup路由器通过接收到VRRP报文的情况来判断Master路由器是否工作正常。

Master路由器主动放弃Master地位（如Master路由器退出虚拟路由器）时，会发送优先级为0的VRRP报文，致使Backup路由器快速切换变成Master路由器。这个切换的时间称为Skew time，计算方式为： $(256 - \text{Backup路由器的优先级}) / 256$ ，单位为秒。

当Master路由器发生网络故障而不能发送VRRP报文的时候，Backup路由器并不能立即知道其工作状况。Backup路由器等待一段时间之后，如果还没有接收到VRRP报文，那么会认为Master路由器无法正常工作，而把自己升级为Master路由器，周期性发送VRRP报文。如果此时多个Backup路由器竞争Master路由器的位置，将通过优先级来选举Master路由器。Backup路由器默认等待的时间称为Master_Down_Interval，取值为： $(3 \times \text{VRRP报文的发送时间间隔}) + \text{Skew time}$ ，单位为秒。

在性能不够稳定的网络中，Backup路由器可能因为网络堵塞而在Master_Down_Interval期间没有收到Master路由器的报文，而主动抢占为Master位置，如果此时原Master路由器的报文又到达了，就会出现虚拟路由器的成员频繁的进行Master抢占现象。为了缓解这种现象的发生，特制定了延迟等待定时器。它可以使得Backup路由器在等待了Master_Down_Interval后，再等待延迟等待时间。如在此期间仍然没有收到VRRP报文，则此Backup路由器才会切换为Master路由器，对外发送VRRP报文。

2.3.3 认证方式

VRRP提供了三种认证方式：

- 无认证：不进行任何 VRRP 报文的合法性认证，不提供安全性保障。
- 简单字符认证：在一个有可能受到安全威胁的网络中，可以将认证方式设置为简单字符认证。发送 VRRP 报文的路由器将认证字填入到 VRRP 报文中，而收到 VRRP 报文的路由器会将收到的 VRRP 报文中的认证字和本地配置的认证字进行比较。如果认证字相同，则认为接收到的报文是合法的 VRRP 报文；否则认为接收到的报文是一个非法报文。
- MD5 认证：在一个非常不安全的网络中，可以将认证方式设置为 MD5 认证。发送 VRRP 报文的路由器利用认证字和 MD5 算法对 VRRP 报文进行加密，加密后的报文保存在 Authentication Header（认证头）中。收到 VRRP 报文的路由器会利用认证字解密报文，检查该报文的合法性。

3 Comware实现的技术特色

3.1 监视上行链路

VRRP网络传输功能有时需要额外的技术来完善其工作。例如，Master路由器到达某网络的链路突然断掉时，主机无法通过此Master路由器远程访问该网络。此时，可以通过监视指定接口上行链路功能，解决这个问题。当Master路由器发现上行链路出现故障后，主动降低自己的优先级（使Master路由器的优先级低于Backup路由器），并立即发送VRRP报文。Backup路由器接收到优先级比自己低的VRRP报文后，等待Skew_Time切换为新的Master路由器。从而，使得能够到达此网络的Backup路由器充当VRRP新的Master路由器，协助主机完成网络通讯。

VRRP可以直接监视连接上行链路的接口状态。当连接上行链路的接口down时，将Master路由器降低指定的优先级。VRRP优先级最低可以降低到1。

VRRP可以利用NQA技术监视上行链路连接的远端主机或者网络状况。例如，Master设备上启动NQA的ICMP-echo探测功能，探测远端主机的可达性。当ICMP-echo探测失败时，它可以通知本设备探测结果，达到降低VRRP优先级的目的。

VRRP也可以利用BFD技术监视上行链路连接的远端主机或者网络状况。由于BFD的精度可以到达10ms，通过BFD能够快速检测到链路状态的变化，达到快速抢占的目的。例如，可以在Master路由器上使用BFD技术监视上行设备的物理状态，在

上行设备坏掉之后，快速检测到该变化，并降低Master路由器的优先级，致使Backup路由器等待Skew time后，抢占成为新的Master路由器。

3.2 Backup监视Master工作状态

Backup路由器在Master路由器坏掉之后，正常情况下需要等待Master_Down_Interval才能切换为新的Master的位置，这段时间内主机将无法正常通信，因为此时没有Master设备替它转发报文。为了解决这个网络故障，Backup设备提供了一个监听Master工作状态的功能，使得Master路由器坏掉之后Backup能够立即切换成为新的Master路由器，维持网络通讯。

Backup路由器监视Master路由器采用的是具有快速检测功能的BFD技术。在Backup设备上使用该技术监视Master路由器的状态，一旦Master路由器发生故障，Backup就可以自动切换成为新的Master路由器，将切换时间缩短到毫秒级。

4 典型组网案例

4.1 主备备份

主备备份方式表示业务仅由Master路由器承担。当Master路由器出现故障时，才会由选举出来的Backup路由器接替它工作。如图3中所示。

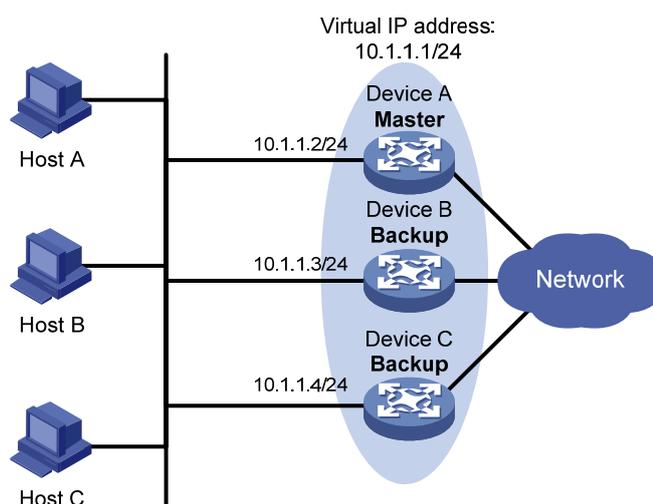


图3 主备备份VRRP

初始情况下，Device A是Master路由器并承担转发任务，Device B和Device C是

Backup路由器且都处于就绪监听状态。如果Device A发生故障，则虚拟路由器内处于Backup状态的Device B和Device C路由器将根据优先级选出一个新的Master路由器，这个新Master路由器继续为网络内的主机转发数据。

4.2 负载分担

在路由器的一个接口上可以创建多个虚拟路由器，使得该路由器可以在一个虚拟路由器中作为Master路由器，同时在其他的虚拟路由器中作为Backup路由器。

负载分担方式是指多台路由器同时承担业务，因此负载分担方式需要两个或者两个以上的虚拟路由器，每个虚拟路由器都包括一个Master路由器和若干个Backup路由器，各虚拟路由器的Master路由器可以各不相同，如图4中所示。

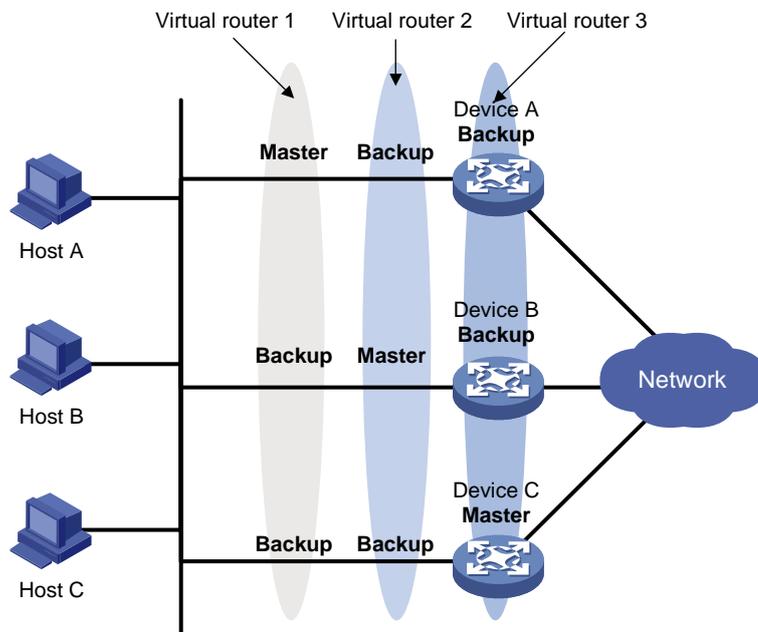


图4 负载分担VRRP

在图4中，有三个虚拟路由器存在：

- 虚拟路由器 1: Device A 作为 Master 路由器，Device B 和 Device C 作为 Backup 路由器。
- 虚拟路由器 2: Device B 作为 Master 路由器，Device A 和 Device C 作为 Backup 路由器。
- 虚拟路由器 3: Device C 作为 Master 路由器，Device A 和 Device B 作为 Backup 路由器。

为了实现业务流量在Device A、Device B和Device C之间进行负载分担，需要将局域网内的主机的默认网关分别设置为虚拟路由器1、2和3。在配置优先级时，需要确保三个虚拟路由器中各路由器的VRRP优先级形成一定的交叉，使得一台路由器尽可能不同时充当2个Master路由器。

4.3 Master使用BFD/NQA监视上行链路

VRRP可以通过BFD或NQA等快速检测协议监视一些上行敏感链路，使得Master路由器快速地发现网络故障，降低自身的优先级，从而保证上行链路工作正常的Backup路由器能够接替它的工作。

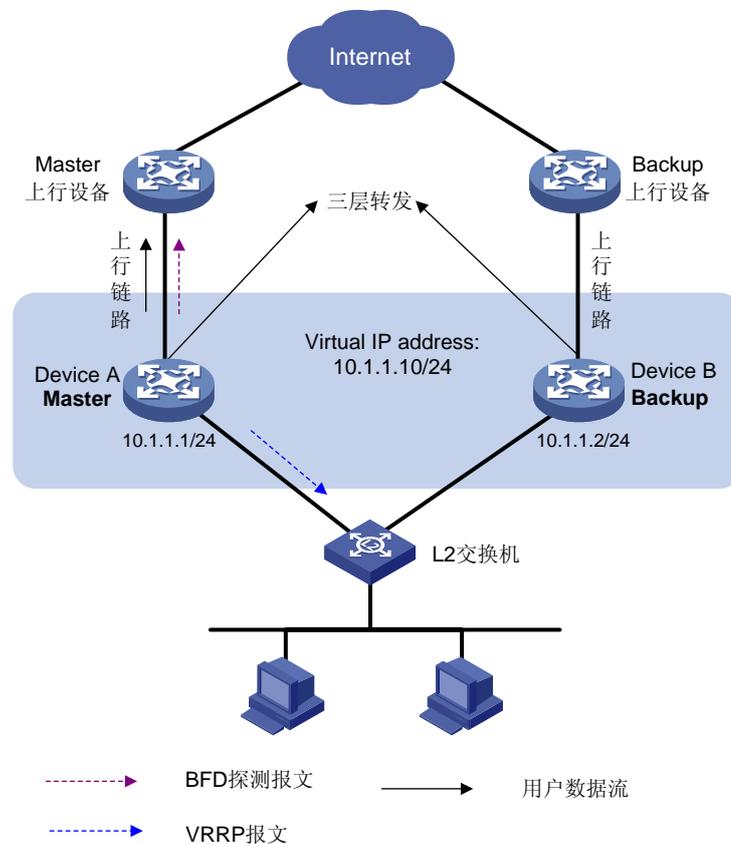


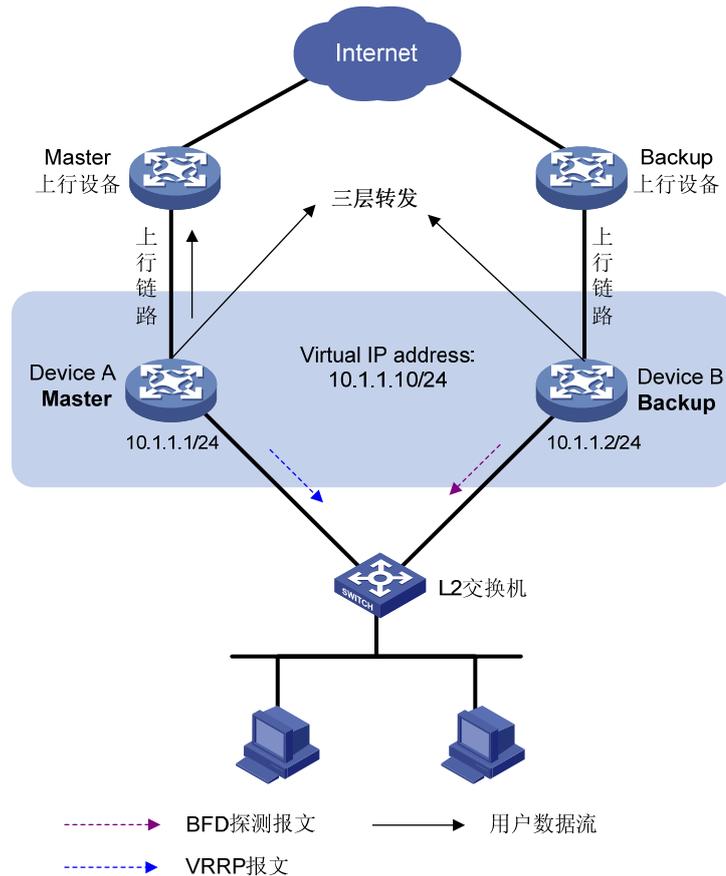
图5 Master监视上行链路

如图5所示，初始情况下，Device A作为Master路由器，承担转发任务；Device B为Backup路由器，处于就绪监听状态。Device A使用BFD监视上行到达Internet的链路状态。如果Device A的上行链路发生故障，Device A可以在毫秒级感知到网络变化，立即发送低优先级的VRRP报文给Device B。如果此时Device B的优先级高于报文中的优先级，那么它将在Skew Time时间之后切换为新的Master路由器，之

后由这个新的Master路由器为网络内的主机转发数据。

4.4 Backup使用BFD监视Master状态

为了保证网络传输的稳定性，可以在Backup设备上使用BFD技术监视Master的状态，使得Master设备发生故障时，Backup设备能够立即切换为新的Master设备。



如图6中所示，初始情况下，Device A作为Master路由器，承担转发任务；Device B是Backup路由器，处于就绪监听状态。Device B使用BFD监视Device A上IP地址10.1.1.1的可达性。如果Device A发生故障，Device B可以立即通过BFD感知到对端的变化，主动切换成为新的Master设备，之后这个新Master路由器将为网络内的主机转发数据。

5 附录

5.1 参考文献

- RFC 3768: Virtual Router Redundancy Protocol (VRRP)

Copyright ©2008 杭州华三通信技术有限公司 版权所有，保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

本文档中的信息可能变动，恕不另行通知。