

[19] 中华人民共和国国家知识产权局



# [12] 发明专利申请公布说明书

[21] 申请号 200710073938.6

[51] Int. Cl.

H04L 12/18 (2006.01)

H04L 12/56 (2006.01)

H04L 1/22 (2006.01)

H04L 29/06 (2006.01)

H04L 12/28 (2006.01)

[43] 公开日 2007年9月12日

[11] 公开号 CN 101035009A

[22] 申请日 2007.3.31

[21] 申请号 200710073938.6

[71] 申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为  
总部办公楼

[72] 发明人 王黎阳 刘 珺 赵凤华

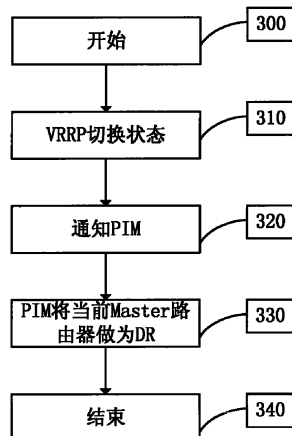
权利要求书 2 页 说明书 11 页 附图 4 页

## [54] 发明名称

一种组播流量冗余保护的方法及设备

## [57] 摘要

本发明公开了一种组播流量冗余保护的方法，发生故障后，虚拟路由冗余协议 VRRP 选举新的主用 Master 路由器，通知协议无关组播路由协议 PIM，所述 PIM 的断言 Assert 机制将所述主用 Master 路由器作为自己的指定路由器 DR。同时，本发明还公开了一种可以提供组播流量冗余保护的三层网络设备。通过本发明可以有效地缩短 Assert 的切换延时，进而避免组播流量短暂中断的问题。



1. 一种组播流量冗余保护的方法,其特征在于,虚拟路由冗余协议VRRP状态切换模块选举新的主用路由器,通知协议无关组播路由协议PIM模块,所述PIM模块选定所述新的主用路由器为指定路由器。

2. 根据权利要求1所述的方法,其特征在于,所述虚拟路由冗余协议VRRP选举新的主用路由器的步骤之前还包括:检测到原主用路由器无法转发组播业务,通知所述虚拟路由冗余协议VRRP选举新的主用路由器。

3. 根据权利要求2所述的方法,其特征在于,所述检测到原主用路由器无法转发组播业务,包括:

通过所述虚拟路由冗余协议VRRP检测到所述原主用路由器无法转发组播业务。

4. 根据权利要求2所述的方法,其特征在于,所述检测到原主用路由器无法转发组播业务,包括:

通过专用故障检测协议检测到所述原主用路由器无法转发组播业务。

5. 根据权利要求4所述的方法,其特征在于,所述专用故障检测协议包括双向转发检测协议BFD。

6. 一种三层网络设备,其特征在于,包括:

虚拟路由冗余协议VRRP状态切换模块:用于选举新的主用路由器;

协议无关组播路由协议PIM模块:用于选定所述新的主用路由器为指定路由器。

其中,当所述VRRP状态切换模块选举出新的主用路由器之后,通知所述PIM模块。

7. 一种三层网络设备,其特征在于,包括:

故障检测模块：用于检测主用路由器是否转发组播业务；

虚拟路由冗余协议VRRP状态切换模块：用于选举新的主用路由器；

协议无关组播路由协议PIM模块：用于选定所述新的主用路由器为指定路由器；

其中，当所述故障检测模块检测到主用路由器出现故障时，通知所述VRRP状态切换模块；当所述VRRP状态切换模块选举出新的主用路由器之后，通知所述PIM模块。

8. 根据权利要求7所述的设备，其特征在于，所述故障检测模块包括VRRP故障检测模块。

9. 根据权利要求7所述的设备，其特征在于，所述故障检测模块包括专用故障检测模块。

10. 根据权利要求9所述的设备，其特征在于，所述专用故障检测模块包括双向转发检测模块。

## 一种组播流量冗余保护的方法及设备

### 技术领域

本发明涉及网络通信技术领域，尤其涉及一种组播流量冗余保护的方法及设备。

### 背景技术

组播技术首先是在网络层中应用和发展起来的。随着Internet网络的不断发展，一方面网络中交互的各种数据、语音和视频信息越来越多，另外新兴的IPTV，电子商务、网上会议、网上拍卖、视频点播、远程教学等服务逐渐兴起。这些服务对信息安全性、有偿性、网络带宽提出了要求。对于这些单点对多点的问题，组播技术有效地解决了该问题，实现了IP网络中点到多点的高效数据传送，能够大量节约网络带宽、降低网络负载。利用网络的组播特性可以方便地提供一些新的增值业务。

IP组播协议主要有用于主机注册的组播组管理协议，和用于组播选路转发的组播路由协议。

在接收者主机和组播路由器之间通常采用组播组管理协议IGMP (Internet Group Membership Protocol)，该协议定义了主机与路由器之间建立和维护组播成员关系的机制。

组播路由器之间运行组播路由协议，组播路由协议用于建立和维护组播路由，并正确、高效地转发组播数据包。组播路由建立了一个从数据源端到多个接收端的无环数据传输路径，即构建分发树结构。按照单播路由一样，组播路由也分为域内和域间两大类：域内组播路由协议，通过发现组播源并构建组播

分发树，将信息传递到接收者，例如：DVMRP (Distance Vector Multicast Routing Protocol)、MOSPF (Multicast Open Shortest Path First)、PIM-DM (Protocol Independent Multicast-Dense Mode) 和 PIM-SM (Protocol Independent Multicast-Sparse Mode) 等协议。域间组播路由协议，发现其他组播域内的组播源 (例如：MSDP, Multicast Source Discovery Protocol)，及在自治系统之间传递组播路由信息 (例如：MBGP, Multicast Border Gateway Protocol)。

PIM (Protocol Independent Multicast) 称为协议无关组播，表示为IP组播提供路由的单播路由协议可以是静态路由、RIP、OSPF、IS-IS、BGP等，组播路由和单播路由协议无关，只要单播路由协议能产生路由表项即可。RPF (Reverse Path Forwarding) 称为逆向路径转发，是组播转发的一种模式。它首先检查收到的数据包的源IP、接口，和单播路由表IP和接口相比较，如果信息包可通过单播路由从收到接口返回到源站点，那么RPF检查成功，就认为这个组播包是从正确路径而来；否则，将组播包作为冗余报文丢弃。借助RPF转发机制，PIM实现了在网络中传递组播信息。为了描述上的方便，我们把由支持PIM协议的组播路由器所组成的网络称为PIM组播域。

如图1所示，在共享网络 (如Ethernet) 中会出现相同报文的重复发送。例如图1中LAN网段上包含多台组播路由器A、B和C，各自都有到组播源S的接收途径。当路由器A、B和C都从上游接收到组播源S发出的组播数据报文后，都会向Ethernet网络上转发该组播报文，这时下游节点HostA就会收到三份完全相同的组播报文。

为了避免这种情况，就需要通过Assert机制来选定一个唯一的转发者，即指定路由器DR路由器 (designate router)。网络中的各路由器通过发送Assert报文选出一条最优的路径，获胜路由器成为该 (S, G) 项的上游邻居，由它负责该 (S, G) 组播报文的转发，而其他落选路由器则剪掉对应的接口以禁止转

发信息。比较项可以为Metric值：路由器上到组播源的cost值；如果Metric值相同，则比较接口上IP地址：IP地址大的优先。

PIM的Assert报文周期缺省为60秒，如果上游邻居路由器出现故障，由PIM的assert进行重新收敛，至少180秒（三个周期），这期间会造成组播流量的中断。

已有技术结合双向转发检测协议（BFD, Bidirectional Forwarding Detection）等故障检测机制，可以尽快发现共享网段PIM路由器故障，触发Assert的重新选举，从而在共享网段中选出新的上游路由器。

但是，BFD发现故障后触发Assert重新收敛，选举出新的指定路由器，仍然存在切换有延时，进而导致组播流量短暂中断的问题。

## 发明内容

本发明解决的技术问题是提供一种组播流量冗余保护的方法及设备，故障发生后，缩短Assert重新收敛导致的切换延时，进而避免组播流量短暂中断的问题。

本发明的目的是通过以下技术方案实现的：

一种组播流量冗余保护的方法，虚拟路由冗余协议VRRP状态切换模块选举新的主用路由器，通知协议无关组播路由协议PIM模块，所述PIM模块选定所述新的主用路由器为指定路由器。

一种三层网络设备，包括：

虚拟路由冗余协议VRRP状态切换模块：用于选举新的主用路由器；

协议无关组播路由协议PIM模块：用于选定所述新的主用路由器为指定路由器。

其中，当所述VRRP状态切换模块选举出新的主用路由器之后，通知所述PIM模块。

一种三层网络设备，包括：

故障检测模块：用于检测主用路由器是否转发组播业务；

虚拟路由冗余协议VRRP状态切换模块：用于选举新的主用路由器；

协议无关组播路由协议PIM模块：用于选定所述新的主用路由器为指定路由器；

其中，当所述故障检测模块检测到主用路由器出现故障时，通知所述VRRP状态切换模块；当所述VRRP状态切换模块选举出新的主用路由器之后，通知所述PIM模块。

因为VRRP的收敛时间是3秒，也就是说切换时间还少与3秒，本实施例借助VRRP的快速切换机制，触发PIM快速切换，将新的Master路由器作为DR路由器，减少Assert重新选举的收敛时间，组播流量可以迅速切换到新的DR路由器上，有效地缩短Assert的切换延时，进而避免组播流量短暂中断的问题。保障了组播数据的可靠性。

进一步，因为本发明实施例中的在单播环境下VRRP机制中的Master路由器与组播下PIM的Assert机制中的DR路由器相同，有效保证PIM组播路径与用户侧的单播路径相同。

进一步，因为实际组网环境中，大部分网络设备都支持VRRP机制，因此本发明实施例在不需更换已有网络设备的情况下，就可以将组播流量迅速切换到新的DR路由器上，有效地缩短Assert的切换延时，进而避免组播流量短暂中断的问题。

## 附图说明

图 1为已有技术中PIM的Assert机制的组网结构示意图；

图 2为本发明实施例一中组播流量冗余保护的组网结构示意图；

图 3为本发明实施例一中组播流量冗余保护的流程图；

图 4为本发明实施例二中组播流量冗余保护的流程图；

图 5为本发明实施例三中组播流量冗余保护的流程图;

图 6为本发明实施例三中组播流量冗余保护的组网结构示意图;

图 7为本发明实施例四中提供组播流量冗余保护的三层网络设备示意图;

图 8为本发明实施例五中提供组播流量冗余保护的三层网络设备示意图;

图 9为本发明实施例六中提供组播流量冗余保护的三层网络设备示意图。

## 具体实施方式

贯穿说明书,示出的该优选实施例和示例应被看作本发明的范例而不受限制。

实施例一:图1中LAN网段上的三台组播路由器A、B和C,及HostA组成的VRRP在PIM路径保护上的组网示意图如图2所示:假设RouterA为Master,负责转发业务;RouterB和RouterC为Backup,为了简化起鉴,图2中在VRRP组网中Master路由器同时作为接收者HostA的上游邻居,实际应用中,Master路由器和HostA之间还可以包含其他网络设备。如图3所示,该方法具体包括:

300:开始;

310:VRRP 的Master路由器无法正常转发业务后, VRRP模块切换状态。具体地,优先级最高的Backup自动变成Master,原Master变为Backup。

虚拟路由冗余协议(VRRP, Virtual Router Redundancy Protocol)将局域网的一组路由器组织成一个备份组,功能上相当于一台虚拟路由器。局域网内的主机仅仅知道这个虚拟路由器的IP地址,并不知道备份组内具体某台设备的IP地址,它们将自己的缺省路由下一跳地址设置为该虚拟路由器的IP地址。于是,网络内的主机就通过这个虚拟路由器与其它网络进行通信。具体地, VRRP机制通过VRRP优先级、IP地址进行比较确定Master,业务一般仅由Master路由器承担。Master路由器每隔一段时间会向组内其他Backup路由器发送VRRP宣告报文来通知自己工作正常,当Backup路由器定时器超时后仍未收到VRRP宣告报



文，则优先级最高的Backup自动变成Master。VRRP的收敛周期是3秒。

尽管在单播的情况下，只有VRRP备份组中的Master负责转发单播业务，但是，在组播情况下，VRRP备份组中的所有路由器都会转发组播业务。因为本发明实施例是针对组播业务的，仅是借助VRRP的快速切换机制以通知PIM的Assert机制来确定DR路由器。

**320:** 通知PIM发生状态切换后的当前Master路由器。具体地，VRRP状态切换模块可以通过参数传递等方式通知PIM模块；告知VRRP发生状态切换后的当前Master；

**330:** PIM的Assert机制将步骤310中的Master路由器作为自己的DR路由器，该DR路由器负责转发组播报文；

Assert确定新的DR路由器之后，还会涉及到组播数据包在网络中的树形转发路径的变化，具体为节点的剪枝或者嫁接操作，即对原Master执行剪枝操作，对新Master执行嫁接操作，因为对本领域技术人员是公知常识，且不涉及本发明实施例的关键点，在此不再累述。

**340:** 结束。

因为VRRP的收敛时间是3秒，也就是说切换时间还少与3秒，本实施例借助VRRP的快速切换机制，触发PIM快速切换，将新的Master路由器作为DR路由器，减少Assert重新选举的收敛时间，组播流量可以迅速切换到新的DR路由器上，有效地缩短Assert的切换延时，进而避免组播流量短暂中断的问题。保障了组播数据的可靠性。

进一步，因为本发明实施例中的在单播环境下VRRP机制中的Master路由器与组播下PIM的Assert机制中的DR路由器相同，有效保证PIM组播路径与用户侧的单播路径相同。

进一步，因为实际组网环境中，大部分网络设备都支持VRRP机制，因此本发明实施例在不需要更换已有网络设备的情况下，就可以将组播流量迅速切换

到新的DR路由器上，有效地缩短Assert的切换延时，进而避免组播流量短暂中断的问题。

步骤310中VRRP切换状态之前，还包括故障机制用于检测当前Master路由器是否发生故障，具体地，该故障检测机制可以是VRRP自己发现故障，也可以是通过BFD等专用故障检测协议发现故障，然后通知VRRP切换状态。以下分别详细说明。

实施例二：实施例一步骤310中VRRP切换状态的步骤之前是VRRP自己发现故障时，如图4所示，在VRRP切换状态之前，还包括步骤VRRP发现故障。VRRP发现故障机制是已有技术，可以为Master每隔一段时间会向组内其他Backup发送VRRP宣告报文来通知自己工作正常，当Backup定时器超时后仍未收到VRRP宣告报文，则认为发生故障。

本实施例借助VRRP的快速故障检测机制，及快速切换机制，触发PIM快速切换，将新的Master路由器作为DR路由器，减少Assert重新选举的收敛时间，组播流量可以迅速切换到新的DR路由器上，有效地缩短Assert的切换延时，进而避免组播流量短暂中断的问题。保障了组播数据的可靠性。

进一步，因为本发明实施例中的在单播环境下VRRP机制中的Master路由器与组播环境下PIM的Assert机制中的DR路由器相同，有效保证PIM组播路径与用户侧的单播路径相同。

进一步，因为实际组网环境中，大部分网络设备都支持VRRP机制，因此本发明实施例在不需更换已有网络设备的情况下，就可以将组播流量迅速切换到新的DR路由器上，有效地缩短Assert的切换延时，进而避免组播流量短暂中断的问题。

实施例三：对于VRRP故障检测机制不够及时的设备中，为使VRRP路由器快速切换，可使用BFD等故障检测技术加以辅助检测触发。图1中LAN网段上的三台组播路由器A、B和C，及HostA组成的BFD+VRRP在PIM路径保护上的组网示意

图如图6所示：假设RouterA为Master，负责转发业务；RouterB和RouterC为Backup，为了简化起鉴，图6中在VRRP组网中Master路由器同时作为接收者HostA的上游邻居，实际应用中，Master路由器和HostA之间还可以包含其他网络设备。同时，在网络上运行BFD机制，用于快速检测故障。如图5所示，该方法具体包括：

500: 开始；

510: BFD发现故障；BFD技术能够在系统之间的任何类型通道上进行故障检测，这些通道包括直接的物理链路，虚电路，隧道，MPLS LSP，多跳路由通道，以及非直接的通道。由于BFD实现故障检测的简单、单一性，致使BFD能够专注于转发故障的快速检测。

BFD用于检测转发引擎之间的通信故障。BFD对两个系统间的、同一路径上的一种数据协议（data protocol）的连通性进行检测，这条路径可以是物理链路或逻辑链路，包括隧道。可以把BFD看作是系统提供的一种服务：上层应用向BFD提供检测地址、检测时间等参数；BFD根据这些信息创建、删除或修改BFD会话，并把会话状态通告给上层应用。上层应用对BFD会话状态变更采取什么措施完全由上层应用自己决定。

具体地，BFD的检测机制为在两个系统建立BFD会话，并沿它们之间的路径周期性发送检测报文，如果一方在既定的时间内没有收到BFD控制报文，则认为路径上发生了故障。BFD可以实现毫秒级别的链路缺陷检测。

520: 通知VRRP切换状态；具体地，BFD模块可以通过参数传递等方式通知VRRP状态切换模块Master无法转发组播业务，需要切换状态；

530: VRRP切换状态。VRRP状态切换模块收到来BFD模块的故障通知之后，VRRP状态切换模块切换状态，具体地，优先级最高的Backup自动变成Master，Master变为Backup；

540: 通知PIM。具体地，VRRP状态切换模块可以通过参数传递等方式通知

PIM模块；告知VRRP发生状态切换后的当前Master；

550: PIM的Assert机制将步骤310中的Master路由器作为自己的DR路由器，该DR路由器负责转发组播报文；

Assert确定新的DR路由器之后，还会涉及到组播数据包在网络中的树形转发路径的变化，具体为节点的剪枝或者嫁接操作，即对原Master执行剪枝操作，对新Master执行嫁接操作，因为对本领域技术人员是公知常识，且不涉及本发明实施例的关键点，在此不再累述。

560: 结束。

以上实施例中仅以BFD为例，实际应用中可以采用其他专用故障检测机制，原理相同，在此不再累述。

因为BFD可以实现毫秒级别的链路缺陷检测，本实施例借助BFD的快速故障检测机制，及VRRP的Master路由器出现故障时，优先级最高的Backup自动变成Master的快速切换机制，触发PIM快速切换，将新的Master路由器作为DR路由器，组播流量可以迅速切换到新的DR路由器上，减少Assert重新选举的收敛时间，有效地缩短Assert的切换延时，进而避免组播流量短暂中断的问题。保障了组播数据的可靠性。

进一步，因为本发明实施例中的在单播环境下VRRP机制中的Master路由器与组播环境下PIM的Assert机制中的DR路由器相同，有效保证PIM组播路径与用户侧的单播路径相同。

实施例四：与实施例一中方法相对应，本发明实施例四提供一种可以提供组播流量冗余保护的三层网络设备，该三层网络设备包括：

VRRP状态切换模块：用于VRRP切换状态，通知PIM模块VRRP机制中的当前Master路由器；

PIM模块：收到来自VRRP状态切换模块的通知，将当前Master路由器作为自己的DR。

与实施例二和实施例三中方法相对应，该三层设备还可以包括故障检测模块，用于检测当前Master路由器是否发生故障，具体地，可以是VRRP故障检测模块自己发现故障，也可以是通过BFD等专用故障检测模块发现故障，发现故障后，通知VRRP切换状态，下面分别详细说明：

实施例五：与实施例二中方法相对应，本发明实施例五提供一种可以提供组播流量冗余保护的三层网络设备，该三层网络设备包括：

VRRP故障检测模块：用于检测Master路由器是否可以正常转发业务，当Master路由器无法正常转发业务时，通知VRRP状态切换模块切换状态；

VRRP状态切换模块：用于VRRP切换状态，通知PIM模块VRRP机制中的当前Master路由器；

PIM模块：收到来自VRRP状态切换模块的通知，将当前Master路由器作为自己的DR。

实施例六：与实施例三中方法相对应，本发明实施例五提供一种可以提供组播流量冗余保护的三层网络设备，该三层网络设备包括：

BFD模块：用于检测Master路由器是否可以正常转发业务，当Master路由器无法正常转发业务时，通知VRRP状态切换模块切换状态；

VRRP状态切换模块：用于VRRP切换状态，通知PIM模块VRRP机制中的当前Master路由器；

PIM模块：收到来自VRRP状态切换模块的通知，将当前Master路由器作为自己的DR。

实施例四、五和六中的三层网络设备可以是路由器或者三层交换机，即可以识别TCP/IP中所规定的第三层信息的网络设备，如IP地址信息。

本实施例借助VRRP/BFD的快速故障检测机制，及VRRP的快速切换机制，触发PIM快速切换，将新的Master路由器作为DR路由器，减少Assert重新选举的收敛时间，组播流量可以迅速切换到新的DR路由器上，有效地缩短Assert的切

换延时，进而避免组播流量短暂中断的问题。保障了组播数据的可靠性。

进一步，因为本发明实施例中的在单播环境下VRRP机制中的Master路由器与组播环境下PIM的Assert机制中的DR路由器相同，有效保证PIM组播路径与用户侧的单播路径相同。

进一步，因为实际组网环境中，大部分网络设备都支持VRRP机制，因此本发明实施例在不需更换已有网络设备的情况下，就可以将组播流量迅速切换到新的DR路由器上，有效地缩短Assert的切换延时，进而避免组播流量短暂中断的问题。

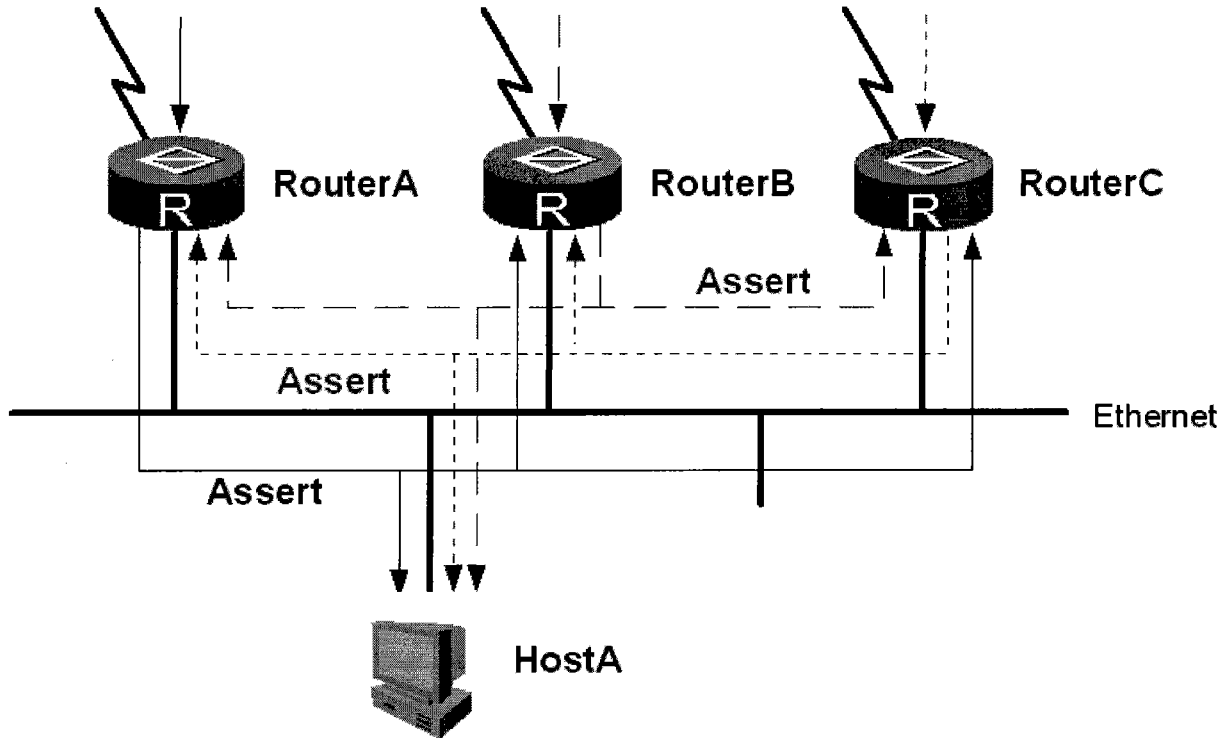


图 1

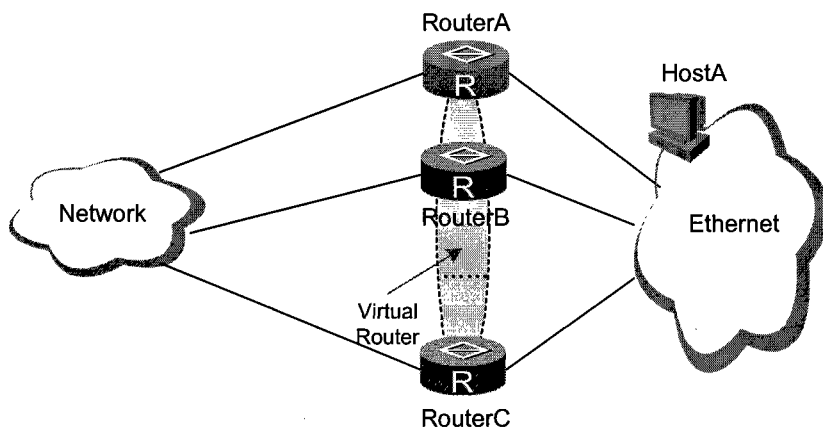


图 2

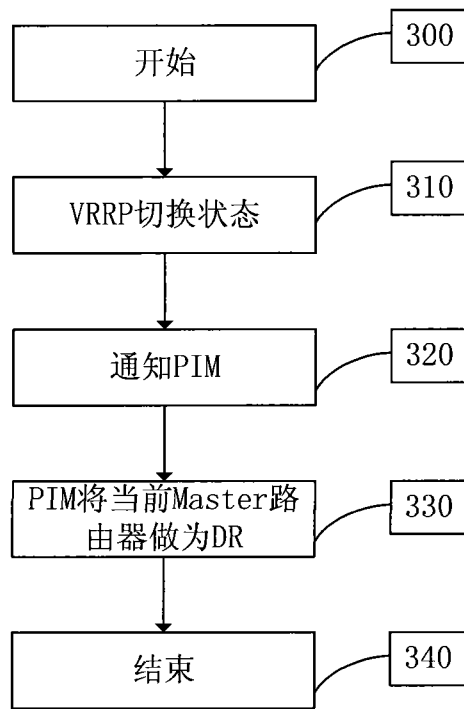


图 3

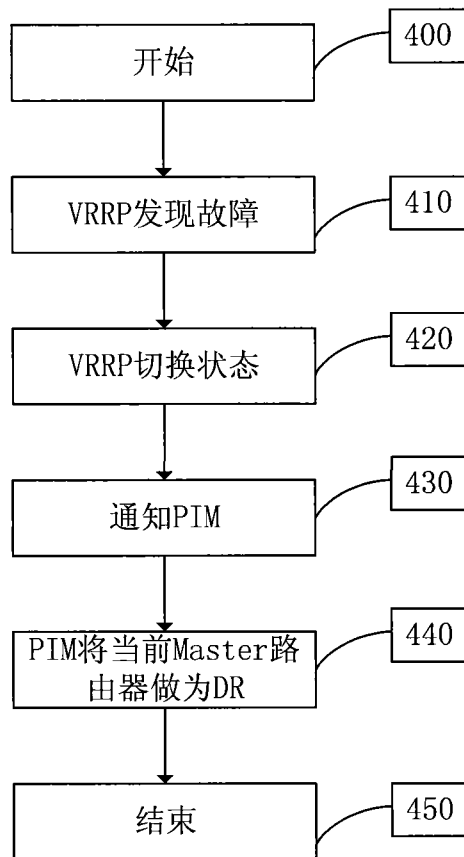


图 4



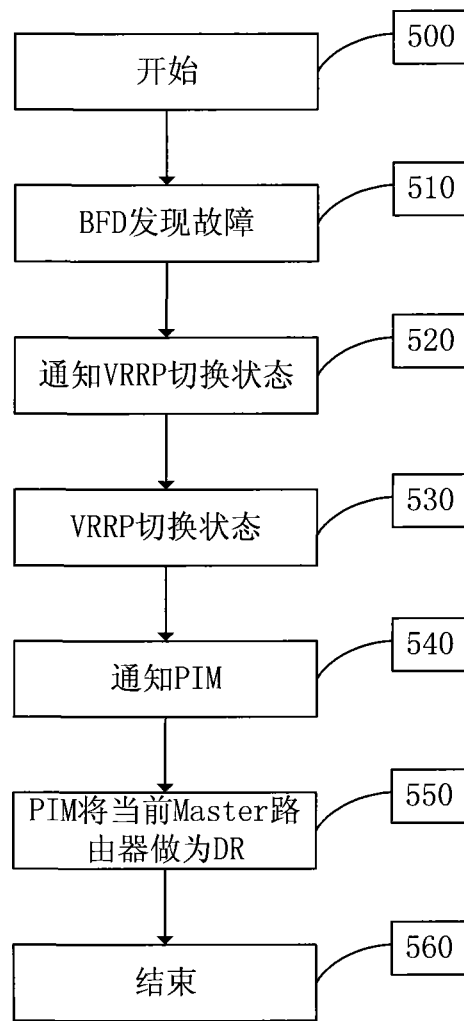


图 5

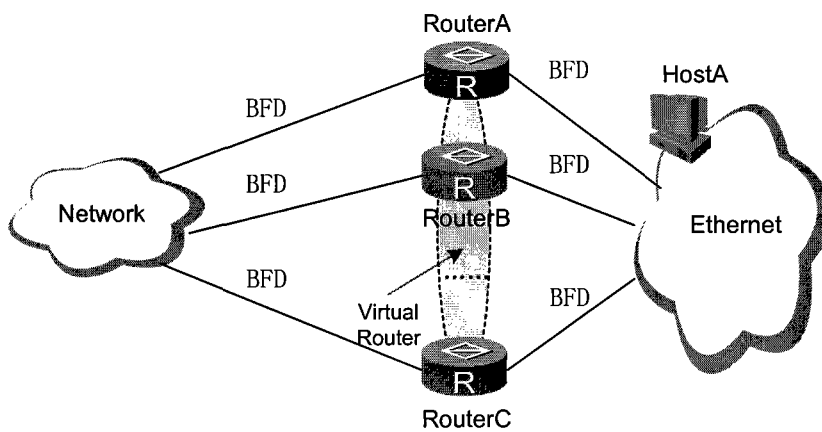


图 6

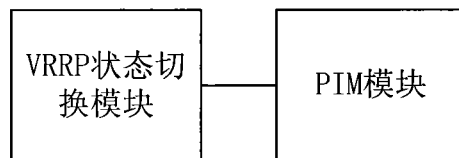


图 7

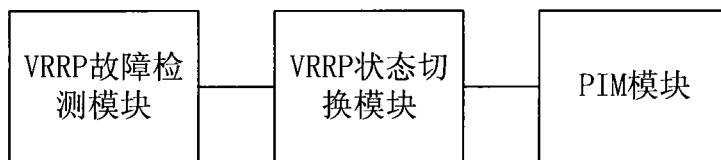


图 8

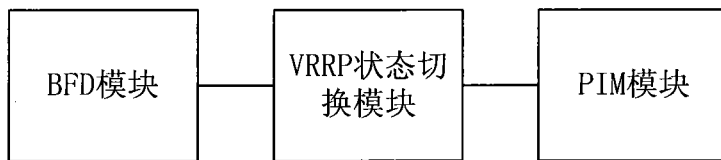


图 9