



(12) 发明专利申请

(10) 申请公布号 CN 103986649 A

(43) 申请公布日 2014. 08. 13

(21) 申请号 201410234199. 4

(22) 申请日 2014. 05. 28

(71) 申请人 杭州华三通信技术有限公司

地址 310052 浙江省杭州市滨江区长河路  
466 号

(72) 发明人 樊超 王伟 王海

(74) 专利代理机构 北京德琦知识产权代理有限  
公司 11018

代理人 谢安昆 宋志强

(51) Int. Cl.

H04L 12/703(2013. 01)

H04L 12/707(2013. 01)

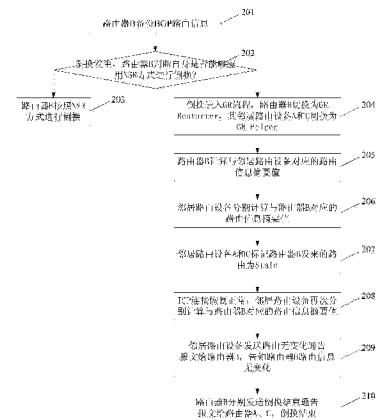
权利要求书2页 说明书9页 附图4页

(54) 发明名称

一种边界网关协议平滑重启方法及路由设备

(57) 摘要

本发明公开了一种边界网关协议平滑重启方法，该方法应用于第一路由设备，所述第一路由设备使能 BGP 不间断路由，备份自身 BGP 路由信息；发生倒换时，如果 TCP 连接异常断开，第一路由设备切换为 GR Restarter，计算 TCP 连接断开期间与邻居路由设备对应的路由信息摘要值；如果 TCP 连接断开期间路由信息发生变化，则在 TCP 连接恢复后通过更新的 BGP 路由信息再次计算与邻居路由设备对应的路由信息摘要值，如果上述两次计算的信息摘要值不相等，则所述第一路由设备发送 TCP 连接恢复后的路由信息给邻居路由设备，使其更新自身记录的路由信息，如果相等，则发送倒换结束通告报文给邻居路由设备。



1. 一种边界网关协议平滑重启方法,其特征在于,该方法应用于第一路由设备,所述第一路由设备与至少一台邻居路由设备建立邻居关系,所述第一路由设备和所述至少一台邻居路由设备具备计算路由信息的信息摘要值能力,所述第一路由设备使能边界网关协议 BGP 不间断路由 NSR,该方法包括:

所述第一路由设备备份自身 BGP 路由信息;

所述第一路由设备发生倒换,如果 TCP 连接异常断开,将自身切换为平滑重启设备 GR Restarter,进入平滑重启 GR Restart 模式;

所述第一路由设备通过备份的 BGP 路由信息计算与邻居路由设备对应的路由信息摘要值 MD5-1;

如果 TCP 连接断开期间路由信息发生变化,所述第一路由设备更新备份的 BGP 路由信息,在 TCP 连接恢复后通过更新后的备份 BGP 路由信息再次计算与邻居路由设备对应的路由信息摘要值 MD5-2,如果所述 MD5-2 与所述 MD5-1 不相等,则所述第一路由设备发送更新后的路由信息给邻居路由设备,使其更新自身记录的路由信息,如果所述 MD5-2 与所述 MD5-1 相等,则所述第一路由设备发送倒换结束通告报文给邻居路由设备,使其删除自身路由信息上的短暂失效标记 Stale,退出平滑重启辅助 GR Help 模式。

2. 根据权利要求 1 所述的方法,其特征在于,该方法进一步包括:

如果 TCP 连接断开期间路由信息无变化,所述第一路由设备发送倒换结束通告报文给邻居路由设备,使其删除自身路由信息上的短暂失效标记 Stale,退出平滑重启辅助 GR Help 模式。

3. 根据权利要求 1 所述的方法,其特征在于,判断所述 TCP 连接断开期间路由信息发生变化的方法为:

所述第一路由设备通过本地路由信息发生变化判定 TCP 连接断开期间路由信息发生变化,和 / 或所述第一路由设备接收到邻居路由设备发送的路由更新信息判定所述邻居路由设备的路由信息发生变化,所述路由更新信息是所述邻居路由设备在 TCP 连接恢复前后分别计算与所述第一路由设备对应的路由信息摘要值,且判定 TCP 连接恢复前后两次计算的信息摘要值不相等时发送给所述第一路由设备的。

4. 根据权利要求 2 所述的方法,其特征在于,判断所述 TCP 连接断开期间路由信息无变化的方法为:

所述第一路由设备的本地路由信息无变化,且所述第一路由设备接收到所有邻居路由设备发送的路由无变化通告报文,判定所述 TCP 连接断开期间路由信息无变化。

5. 根据权利要求 1 所述的方法,其特征在于,该方法进一步包括:

如果所述 TCP 连接正常,则所述第一路由设备按照 BGP NSR 方式执行倒换。

6. 一种路由设备,其特征在于,该设备与至少一台邻居路由设备建立邻居关系,且该路由设备和所述至少一台邻居路由设备具备计算路由信息的信息摘要值能力,该路由设备使能边界网关协议 BGP 不间断路由 NSR,该设备包括:

存储模块,用于备份自身 BGP 路由信息;

处理模块,用于发生倒换时,如果 TCP 连接异常断开,将自身切换为平滑重启设备 GR Restarter,进入平滑重启 GR Restart 模式;

计算模块,用于通过备份的 BGP 路由信息计算与邻居路由设备对应的路由信息摘要值

MD5-1；

判断模块，用于判断 TCP 连接断开期间路由信息是否发生变化；

如果 TCP 连接断开期间路由信息发生变化，所述存储模块还用于更新备份的 BGP 路由信息；

如果判断模块判断 TCP 连接断开期间路由信息发生变化，所述计算模块还用于在 TCP 连接恢复后通过更新后的备份 BGP 路由信息再次计算与邻居路由设备对应的路由信息摘要值 MD5-2；

所述判断模块还用于判断所述 MD5-2 与所述 MD5-1 是否相等；

发送模块，用于所述判断模块判断 MD5-2 与所述 MD5-1 不相等时，发送 TCP 连接恢复后的路由信息给邻居路由设备，使其更新自身记录的路由信息；还用于所述 MD5-2 与所述 MD5-1 相等时，发送倒换结束通告报文给邻居路由设备，使其删除自身路由信息上的短暂失效标记 Stale，退出平滑重启辅助 GRHelp 模式。

7. 根据权利要求 6 所述的设备，其特征在于，如果所述判断模块判断 TCP 连接断开期间路由信息无变化，所述发送模块进一步用于，发送倒换结束通告报文给邻居路由设备，使其删除自身路由信息上的短暂失效标记 Stale，退出平滑重启辅助 GR Help 模式。

8. 根据权利要求 6 所述的设备，其特征在于，

该设备还包括接收模块，用于接收邻居路由设备发送的路由更新信息，发送第一通知报文给判断模块，所述路由更新信息是所述邻居路由设备在 TCP 连接恢复前后分别计算与所述第一路由设备对应的路由信息摘要值，且判定 TCP 连接恢复前后两次计算的信息摘要值不相等时发送给所述第一路由设备的；

所述判断模块进一步用于，通过本地路由信息发生变化判定 TCP 连接断开期间路由信息发生变化，和 / 或接收到所述接收模块发送的第一通知报文判定 TCP 连接断开期间路由信息发生变化。

9. 根据权利要求 7 所述的设备，其特征在于，

所述接收模块还用于，接收到所有邻居路由设备发送的路由无变化通告报文，发送第二通知报文给判断模块；

所述判断模块进一步用于，接收所述接收模块发送的第二通知报文，判定本地路由信息无变化。

10. 根据权利要求 6 所述的设备，其特征在于，所述处理模块进一步用于，如果所述 TCP 连接正常，则按照 BGP NSR 方式执行倒换。

## 一种边界网关协议平滑重启方法及路由设备

### 技术领域

[0001] 本发明涉及网络通信传输技术，特别是一种边界网关协议平滑重启方法及路由设备。

### 背景技术

[0002] 边界网关协议 (Border Gateway Protocol, BGP) 平滑重启 (Graceful Restart, GR) 是一种在主备倒换或 BGP 协议重启时保证转发业务不中断的机制。GR 有两个角色：

[0003] 平滑重启设备 GR Restarter：发生主备倒换或协议重启，且具有 GR 能力的设备。

[0004] 平滑重启辅助设备 GR Helper：和 GR Restarter 具有邻居关系，协助完成 GR 流程的设备。GR Helper 也具有 GR 能力。

[0005] 设备既可以作为 GR Restarter，又可以作为 GR Helper。设备的角色由该设备在 BGP GR 过程中的作用决定。

[0006] BGP GR 的工作过程为：

[0007] (1) GR Restarter 和 GR Helper 通过 Open 消息交互 GR 能力。只有双方都具有 GR 能力时，建立起的 BGP 会话才具备 GR 能力。

[0008] (2) 建立具备 GR 能力的 BGP 会话后，GR Restarter 进行主备倒换或 BGP 协议重启时，GR Restarter 不会删除路由信息库 (Routing Information Base, RIB) 和转发信息库 (Forwarding Information Base, FIB) 表项，仍然按照原有的转发表项转发报文。GR Helper 发现 GR Restarter 进行主备倒换或 BGP 协议重启后，GR Helper 不会删除从该 GR Restarter 学习到的路由，而是将这些路由标记为短暂失效 (Stale) 路由，仍按照这些路由转发报文，从而确保在 GR Restarter 进行主备倒换或 BGP 协议重启的过程中，报文转发不会中断。

[0009] (3) GR Restarter 主备倒换或 BGP 协议重启完成后，它会重新与 GRHelper 建立 BGP 会话。如果在 GR Restarter 通告的 BGP 会话重建时间内没有成功建立 BGP 会话，则 GR Helper 会删除标记为失效的路由。

[0010] (4) 如果在 GR Restarter 通告的 BGP 会话重建时间内成功建立 BGP 会话，则 GR Restarter 和 GR Helper 在建立的 BGP 会话上进行路由信息交互，以便 GRRestarter 恢复路由信息、GR Helper 根据学习到的路由删除路由的失效标记。

[0011] (5) BGP 会话建立后，在 GR Restarter 和 GR Helper 上都会启动路由信息库结束 (End-Of-RIB) 标记等待定时器 (定时器的值通过 graceful-restart timer wait-for-rib 命令配置)，该定时器用来控制路由信息收敛的速度。如果定时器超时没有完成路由信息的交互，则 GR Restarter 不再接收新的路由，根据已经学习到的 BGP 路由信息更新 RIB 表项，删除老化的 RIB 表项；GR Helper 则删除标记为失效的路由。

[0012] (6) 如果在 RIB 路由老化定时器超时时没有完成路由信息的交互，则 GRRestarter 会强制退出 GR 过程，根据已经学习到的 BGP 路由信息更新 RIB 表项，删除老化的 RIB 表项。

[0013] 按照请求评议 RFC(Request For Comments)4724 的规定，GR 重建邻居后 Restart

端需要重新发送全部路由给 Helper。

[0014] 而 BGP 不间断路由 (Nonstop Routing, NSR) 是一种通过在 BGP 协议主备进程之间备份必要的协议状态和数据 (如 BGP 邻居信息和路由信息), 使得 BGP 协议的主进程中断时, 备份进程能够无缝地接管主进程的工作, 从而确保对等体感知不到 BGP 协议中断, 保持 BGP 路由, 并保证转发不会中断的技术。

[0015] BGP NSR 与 BGP GR 具有如下区别 :

[0016] 对 BGP 对等体的要求不同 : 使用 BGP NSR 功能时, BGP 对等体不会感知本地设备发生了 BGP 进程的异常重启或主备倒换等故障, 因为自己备份了 BGP 路由, 及 TCP 连接, 因此不要 BGP 对等体协助恢复 BGP 路由信息。

[0017] BGP GR 要求 BGP 对等体具有 GR 能力, 并且在 BGP 会话中断恢复时, BGP 对等体能够作为 GR helper 协助本地设备恢复 BGP 路由信息。

[0018] 信息摘要算法第五版 (Message Digest Algorithm5.0, MD5) 的典型应用是对一段信息 (Message) 产生信息摘要 (Message-Digest), 以防止被篡改。比如, 在 UNIX 下有很多软件在下载的时候都有一个文件名相同, 文件扩展名为 .md5 的文件, 在这个文件中通常只有一行文本, 大致结构如 :

[0019] MD5(tanajiya.tar.gz) = 0ca175b9c0f726a831d895e269332361

[0020] 这就是 tanajiya.tar.gz 文件的数字签名。MD5 将整个文件当作一个大文本信息, 通过其不可逆的字符串变换算法, 产生了这个唯一的 MD5 信息摘要。如果在以后传播这个文件的过程中, 无论文件的内容发生了任何形式的改变 (包括人为修改或者下载过程中线路不稳定引起的传输错误等), 只要你对这个文件重新计算 MD5 时就会发现信息摘要不相同, 由此可以确定你得到的只是一个不正确的文件。如果再有一个第三方的认证机构, 用 MD5 还可以防止文件作者的 “抵赖”, 这就是所谓的数字签名应用。

## 发明内容

[0021] 有鉴于此, 本发明提出了一种边界网关协议平滑重启方法及路由设备, 有效解决了现有技术中发生倒换时某些设备 TCP 不支持 NSR 或者 TCP 连接意外断开而不能应用 BGP NSR 方式进行倒换的问题, 以及 BGP GR 倒换方式中发送大量不必要的路由导致的带宽占用、路由收敛过慢的问题。本发明提出的技术方案是 :

[0022] 一种边界网关协议平滑重启方法, 该方法应用于第一路由设备, 所述第一路由设备与至少一台邻居路由设备建立邻居关系, 所述第一路由设备和所述至少一台邻居路由设备具备计算路由信息的信息摘要值能力, 所述第一路由设备使能边界网关协议 BGP 不间断路由 NSR, 该方法包括 :

[0023] 所述第一路由设备备份自身 BGP 路由信息 ;

[0024] 所述第一路由设备发生倒换, 如果 TCP 连接异常断开, 将自身切换为平滑重启设备 GR Restarter, 进入平滑重启 GR Restart 模式 ;

[0025] 所述第一路由设备通过备份的 BGP 路由信息计算与邻居路由设备对应的路由信息摘要值 MD5-1 ;

[0026] 如果 TCP 连接断开期间路由信息发生变化, 所述第一路由设备更新备份的 BGP 路由信息, 在 TCP 连接恢复后通过更新后的备份 BGP 路由信息再次计算与邻居路由设备对应

的路由信息摘要值 MD5-2,如果所述 MD5-2 与所述 MD5-1 不相等,则所述第一路由设备发送更新后的路由信息给邻居路由设备,使其更新自身记录的路由信息,如果所述 MD5-2 与所述 MD5-1 相等,则所述第一路由设备发送倒换结束通告报文给邻居路由设备,使其删除自身路由信息上的短暂失效标记 Stale,退出平滑重启辅助 GR Help 模式。

[0027] 上述方案中,该方法进一步包括:

[0028] 如果 TCP 连接断开期间路由信息无变化,所述第一路由设备发送倒换结束通告报文给邻居路由设备,使其删除自身路由信息上的短暂失效标记 Stale,退出平滑重启辅助 GR Help 模式。

[0029] 上述方案中,判断所述 TCP 连接断开期间路由信息发生变化的方法为:

[0030] 所述第一路由设备通过本地路由信息发生变化判定 TCP 连接断开期间路由信息发生变化,和 / 或所述第一路由设备接收到邻居路由设备发送的路由更新信息判定所述邻居路由设备的路由信息发生变化,所述路由更新信息是所述邻居路由设备在 TCP 连接恢复前后分别计算与所述第一路由设备对应的路由信息摘要值,且判定 TCP 连接恢复前后两次计算的信息摘要值不相等时发送给所述第一路由设备的。

[0031] 上述方案中,判断所述 TCP 连接断开期间路由信息无变化的方法为:

[0032] 所述第一路由设备的本地路由信息无变化,且所述第一路由设备接收到所有邻居路由设备发送的路由无变化通告报文,判定所述 TCP 连接断开期间路由信息无变化。

[0033] 上述方案中,该方法进一步包括:

[0034] 如果所述 TCP 连接正常,则所述第一路由设备按照 BGP NSR 方式执行倒换。

[0035] 一种路由设备,该设备与至少一台邻居路由设备建立邻居关系,且该路由设备和所述至少一台邻居路由设备具备计算路由信息的信息摘要值能力,该路由设备使能边界网关协议 BGP 不间断路由 NSR,该设备包括:

[0036] 存储模块,用于备份自身 BGP 路由信息;

[0037] 处理模块,用于发生倒换时,如果 TCP 连接异常断开,将自身切换为平滑重启设备 GR Restarter,进入平滑重启 GR Restart 模式;

[0038] 计算模块,用于通过备份的 BGP 路由信息计算与邻居路由设备对应的路由信息摘要值 MD5-1;

[0039] 判断模块,用于判断 TCP 连接断开期间路由信息是否发生变化;

[0040] 如果 TCP 连接断开期间路由信息发生变化,所述存储模块还用于更新备份的 BGP 路由信息;

[0041] 如果判断模块判断 TCP 连接断开期间路由信息发生变化,所述计算模块还用于在 TCP 连接恢复后通过更新后的备份 BGP 路由信息再次计算与邻居路由设备对应的路由信息摘要值 MD5-2;

[0042] 所述判断模块还用于判断所述 MD5-2 与所述 MD5-1 是否相等;

[0043] 发送模块,用于所述判断模块判断 MD5-2 与所述 MD5-1 不相等时,发送 TCP 连接恢复后的路由信息给邻居路由设备,使其更新自身记录的路由信息;还用于所述 MD5-2 与所述 MD5-1 相等时,发送倒换结束通告报文给邻居路由设备,使其删除自身路由信息上的短暂失效标记 Stale,退出平滑重启辅助 GRHelp 模式。

[0044] 上述方案中,如果所述判断模块判断 TCP 连接断开期间路由信息无变化,所述发

送模块进一步用于，发送倒换结束通告报文给邻居路由设备，使其删除自身路由信息上的短暂失效标记 Stale，退出平滑重启辅助 GR Help 模式。

[0045] 上述方案中，该设备还包括接收模块，用于接收邻居路由设备发送的路由更新信息，发送第一通知报文给判断模块，所述路由更新信息是所述邻居路由设备在 TCP 连接恢复前后分别计算与所述第一路由设备对应的路由信息摘要值，且判定 TCP 连接恢复前后两次计算的信息摘要值不相等时发送给所述第一路由设备的；

[0046] 所述判断模块进一步用于，通过本地路由信息发生变化判定 TCP 连接断开期间路由信息发生变化，和 / 或接收到所述接收模块发送的第一通知报文判定 TCP 连接断开期间路由信息发生变化。

[0047] 上述方案中，

[0048] 所述接收模块还用于，接收到所有邻居路由设备发送的路由无变化通告报文，发送第二通知报文给判断模块；

[0049] 所述判断模块进一步用于，接收所述接收模块发送的第二通知报文，判定本地路由信息无变化。

[0050] 上述方案中，所述处理模块进一步用于，如果所述 TCP 连接正常，则按照 BGP NSR 方式执行倒换。

[0051] 综上所述，本发明提出了一种边界网关协议平滑重启方法及路由设备，当路由设备发生倒换时，如果该路由设备传输控制协议 (Transmission Control Protocol, TCP) 支持 NSR 或 TCP 连接正常，由于已经备份了该路由设备上的路由信息，倒换期间 BGP 邻居不会中断，正常收发数据，以 BGP NSR 方式进行倒换，对等体感知不到 BGP 协议中断，保持 BGP 路由；如果发生倒换的路由设备不支持 NSR 或者 TCP 连接中断，则该路由设备切换为 GR Restarter，其邻居路由设备切换为 GR Helper，GR Restarter 保存了备份路由，GR Helper 保存了 Stale 路由，通过增加 MD5 摘要算法，判断因 TCP 连接断开期间路由是否发生变化来决策是否需要重新发送路由信息，从而加快了路由收敛，减少带宽及处理路由更新的 CPU 占用。

## 附图说明

[0052] 图 1 为方法实施例一和方法实施例二的组网结构图；

[0053] 图 2 为方法实施例一的流程图；

[0054] 图 3 为方法实施例二的流程图；

[0055] 图 4 为本发明实施例的路由设备结构图。

## 具体实施方式

[0056] 按照请求评议 RFC(Request For Comments)4724 的规定，GR 重建邻居后 Restarter 端需要重新发送全部路由给 Helper，实际上其路由都是从 Helper 方恢复的，Helper 端已保存了路由信息的，在大多数情况下，这样的重发机制可能是多余的，会导致 GR 收敛变慢。

[0057] NSR 方式不需要对等体配合路由恢复，控制层面上路由信息需要备份到备板，同时需要协议的 TCP 链接支持 NSR 保证心跳报文，邻居不中断。但是有的路由设备 TCP 不支持

NSR 功能,或者 TCP 异常断开无法应用 NSR 方式进行倒换。

[0058] 基于此,本发明提供了一种边界网关协议平滑重启方法及路由设备。路由设备首先备份自身的 BGP 路由信息,当其发生倒换时,如果 TCP NSR 正常,且 TCP 连接正常,则该路由器设备按照 BGP NSR 方式进行倒换;否则所述路由器设备切换为 GR Restarter,进入 GR 流程。如果以 GR 方式进行倒换,通过增加 MD5 摘要算法,判断因 TCP 连接断开期间路由是否发生变化来决策是否需要重新发送路由信息,从而加快了路由收敛,减少带宽及处理路由更新的 CPU 占用。

[0059] 为使本发明的目的、技术方案和优点表达的更加清楚明白,下面结合附图及具体实施例对本发明再作进一步详细的说明。

[0060] 本发明实施例的技术方案是:

[0061] 所述第一路由设备备份自身 BGP 路由信息;

[0062] 所述第一路由设备发生倒换,如果 TCP 连接异常断开,将自身切换为平滑重启设备 GR Restarter,进入平滑重启 GR Restart 模式;

[0063] 所述第一路由设备通过备份的 BGP 路由信息计算与邻居路由设备对应的路由信息摘要值 MD5-1;

[0064] 如果 TCP 连接断开期间路由信息发生变化,所述第一路由设备更新备份的 BGP 路由信息,在 TCP 连接恢复后通过更新后的备份 BGP 路由信息再次计算与邻居路由设备对应的路由信息摘要值 MD5-2,如果所述 MD5-2 与所述 MD5-1 不相等,则所述第一路由设备发送更新后的路由信息给邻居路由设备,使其更新自身记录的路由信息,如果所述 MD5-2 与所述 MD5-1 相等,则所述第一路由设备发送倒换结束通告报文给邻居路由设备,使其删除自身路由信息上的短暂失效标记 Stale,退出平滑重启辅助 GR Help 模式。

[0065] 方法实施例一

[0066] 以图 1 中的路由组网示意图为本发明技术方案进行说明,图 1 中,路由设备 A 上的本地路由信息记为 a,路由设备 B 上的本地路由信息记为 b,路由设备 C 上的本地路由信息记为 c。路由设备 A 与路由设备 B 建立邻居关系,路由设备 B 与路由设备 C 建立邻居关系。建立邻居关系后各台路由设备分别将本地路由信息发送给其邻居路由,即路由设备 A 将本地路由信息 a 发送给路由设备 B,路由设备 C 将本地路由信息 c 发送给路由设备 B,路由设备 B 将路由信息 b 和 c 发送给路由设备 A,将路由信息 b 和 a 发送给路由设备 C。之后路由设备 A、B、C 上分别记录了路由信息 a、b 和 c。且路由设备 A、B、C 均具有计算路由信息的信息摘要值的能力。本实施例以路由设备 B 发生倒换为例对本发明技术方案进行说明,且假设 TCP 连接断开期间路由设备 A、B、C 的本地路由信息 a、b、c 均没有更新。图 2 为本发明实施例的流程图,如图 2 所示,包括以下步骤:

[0067] 步骤 201:路由设备 B 备份 BGP 路由信息。

[0068] 本步骤中,路由设备 B 将自身记录的本地路由信息 b 和邻居路由信息 a、c 全部备份到备板,与现有技术 NSR 处理方式相同,不作详细介绍。

[0069] 步骤 202:倒换发生,路由设备 B 判断自身是否能够应用 NSR 方式进行倒换,如果是,执行步骤 203,否则执行步骤 204。

[0070] 本步骤中,路由设备 B 如果要成功应用 NSR 方式进行倒换,不仅需要步骤 201 中备份路由信息,还需要 TCP 连接也支持 NSR 并且成功备份,而某些场景下的低端路由设备

的 TCP 是不支持 NSR 的,这就会导致路由设备无法按照 NSR 方式进行倒换;或者路由设备的 TCP 连接支持 NSR,但是 TCP 连接备份时失败导致 TCP 中断,这些情况均使得路由设备无法采用 NSR 方式进行倒换,这种情况下需要执行步骤 204。

[0071] 步骤 203 :路由设备 B 按照 NSR 方式进行倒换。

[0072] 基于步骤 202 的判断结果,路由设备 B 可以按照 NSR 方式进行倒换,则按照 NSR 方式进行倒换,该部分为现有技术,在此不再详述。

[0073] 步骤 204 :倒换进入 GR 流程,路由设备 B 切换为 GR Restarter,其邻居路由设备 A 和 C 切换为 GR Helper。

[0074] 基于步骤 202 的判断结果,路由设备 B 无法按照 NSR 方式进行倒换,则外部采用 GR 方式实现倒换,路由设备 A 和 C 作为 GR Helper 设备协助路由设备 B 进行倒换。

[0075] 步骤 205 :路由设备 B 计算与邻居路由设备对应的路由信息摘要值。

[0076] 步骤 201 中路由设备 B 备份了记录的路由信息 a、b、c,虽然发生倒换,其上仍有备份路由信息,因此可以计算与邻居路由设备对应的路由信息摘要值。所述与邻居路由设备对应的路由信息为路由器 B 应该发送给邻居路由设备的路由信息。

[0077] 与邻居路由设备 A 对应的路由信息为 b、c,则计算得到的信息摘要值为 MD5(BA\_1\_bc);与邻居路由设备 C 对应的路由信息为 a、b,则计算得到的信息摘要值为 MD5(BC\_1\_bc)。

[0078] 步骤 206 :邻居路由设备分别计算与路由设备 B 对应的路由信息摘要值。

[0079] 本步骤中,邻居路由设备 A 和 C 切换为 GR Helper,计算与路由设备 B 对应的路由信息摘要值。所述与路由设备 B 对应的路由信息为邻居路由设备 A 和 C 分别应该发送给路由设备 B 的路由信息。

[0080] 邻居路由设备 A 与路由设备 B 对应的路由信息为 a,则计算得到的信息摘要值为 MD5(AB\_1\_a);邻居路由设备 C 与路由设备 B 对应的路由信息为 c,则计算得到的信息摘要值为 MD5(CB\_1\_c)。

[0081] 步骤 207 :邻居路由设备 A 和 C 标记路由设备 B 发来的路由为 Stale。

[0082] 本步骤中,邻居路由设备 A 和 C 发现路由设备 B 进行倒换,则分别标记倒换发生前路由设备 B 发送来的路由信息为 Stale,即短暂失效,具体的,路由设备 A 将本地保存的路由设备 B 发送来的路由信息 b 和 c 标记为 Stale,路由设备 C 将本地保存的路由设备 B 发送来的路由信息 a 和 b 标记为 Stale。

[0083] 步骤 208 :TCP 连接恢复正常,邻居路由设备再次分别计算与路由设备 B 对应的路由信息摘要值。

[0084] 本步骤中,TCP 连接恢复正常后,邻居路由设备再次分别计算与路由设备 B 对应的路由信息摘要值。

[0085] 邻居路由设备 A 计算的信息摘要值记为 MD5(AB\_2\_a),该信息摘要为路由设备 A 与路由设备 B 对应的路由 a 路由信息的信息摘要;邻居路由设备 C 计算的信息摘要值记为 MD5(CB\_2\_c),该信息摘要为路由设备 C 与路由设备 B 对应的路由 c 路由信息的信息摘要。

[0086] 由于本实施例假设 TCP 连接断开期间路由设备 A、B、C 的本地路由信息 a、b、c 均没有更新,因此必然存在  $MD5(AB_1_a) = MD5(AB_2_a)$ ,  $MD5(CB_1_c) = MD5(CB_2_c)$ 。

[0087] 步骤 209 :邻居路由设备发送路由无变化通告报文给路由设备 B,告知路由设备 B 路由信息无变化。

[0088] 由于步骤 208 中邻居路由设备 A 在 TCP 连接恢复前后计算的对应路由设备 B 的路由信息摘要值相同，则发送路由无变化通告报文给路由设备 B，告知路由设备 B 路由信息无变化；同样，邻居路由设备 C 在 TCP 连接恢复前后计算的对应路由设备 B 的路由信息摘要值也相同，则发送路由无变化通告报文给路由设备 B，告知路由设备 B 路由信息无变化。

[0089] 所述路由无变化通告报文具体可以是 EOR\_127 报文，该 EOR\_127 报文是以 127.0.0.1/32 作为前缀的撤销类型的 UPDATE 报文。具有通告路由发送结束的功能。路由设备 B 接收到该报文后确认发送该报文的邻居路由设备本地路由无更新。

[0090] 步骤 210：路由设备 B 分别发送倒换结束通告报文给路由设备 A、C，倒换结束。

[0091] 本步骤中，由于路由设备 B 接收到所有邻居设备发送来的路由无变化通告报文，确认邻居设备的本地路由信息无变化，且路由设备 B 本地路由信息 b 也无变化，则分别发送倒换结束通告报文给路由设备 A、C，通知各邻居设备收敛结束，使得各邻居设备接收到倒换结束通告报文后去掉步骤 207 中路由设备 B 发送来的路由信息的 Stale 标记。至此，倒换结束。

## [0092] 方法实施例二

[0093] 方法实施例一以 TCP 连接断开期间路由设备 A、B、C 的本地路由信息均无变化为例对本发明技术方案进行说明，本实施例仍采用实施例一中的组网结构与参数设置，假设 TCP 连接断开期间路由设备 A 的本地路由信息 a 更新为 a'，图 3 为本实施例的流程图，包括以下步骤：

[0094] 步骤 301～步骤 307 与步骤 201～步骤 207 相同，在此不再赘述。

[0095] 步骤 308：TCP 连接恢复正常，邻居路由设备再次分别计算与路由设备 B 对应的路由信息摘要值。

[0096] 本步骤中，TCP 连接恢复正常后，邻居路由设备再次分别计算与路由设备 B 对应的路由信息摘要值。

[0097] 邻居路由设备 A 计算的信息摘要值记为 MD5(AB\_2\_a')，邻居路由设备 C 计算的信息摘要值记为 MD5(CB\_2\_c)。

[0098] 由于本实施例假设 TCP 连接断开期间路由设备 A 的本地路由信息 a 更新为 a'，因此必然存在  $MD5(AB_1_a) \neq MD5(AB_2_a')$ ， $MD5(CB_1_c) = MD5(CB_2_c)$ 。

[0099] 步骤 309：路由设备 A 发送更新路由信息 a' 给路由设备 B，路由设备 C 发送路由无变化通告报文给路由设备 B。

[0100] 本步骤中，路由设备 B 接收到路由设备 A 发送的更新路由信息 a'，更新备份的 BGP 路由信息，更新后路由设备 B 上备份的 BGP 路由信息为 a'，b，c。

[0101] 步骤 310：路由设备 B 再次计算对应发送给邻居路由设备的信息摘要值。

[0102] 由于路由设备 B 接收到部分邻居路由设备（本实施例为路由设备 A）发送来的更新路由信息，说明邻居路由信息有变化，此时路由设备 B 需要再次计算对应发送给各邻居路由设备的信息摘要值。

[0103] 对应发送给邻居路由设备 A 的路由信息为 b、c，则计算出的信息摘要值为  $MD5(BA_2_bc)$ ；对应发送给邻居路由设备 C 的路由信息为 a'、b，则计算出的信息摘要值为  $MD5(BC_2_a'b)$ ，则必然存在  $MD5(BA_1_bc) = MD5(BA_2_bc)$ ， $MD5(BC_1_ab) \neq MD5(BC_2_a'b)$ 。

[0104] 步骤 311：路由设备 B 发送倒换结束通告报文给路由设备 A，发送更新路由信息 a'、b 给路由设备 C。

[0105] 由于 TCP 连接恢复前后计算的对应路由设备 A 的路由信息摘要值相同，因此路由设备 A 端的路由信息不用更新，路由设备 B 发送倒换结束通告报文给路由设备 A，通知其路由无变化，使其接收到倒换结束通告报文后去掉路由信息上的 Stale 标记即可；

[0106] TCP 连接恢复前后计算的对应路由设备 C 的路由信息摘要值不相同，因为路由设备 A 的本地路由信息由 a 更新为 a'，而路由设备 C 仍记录的是 a，路由设备 B 需要将更新路由设备 C 端的路由信息，因此发送路由信息 a'、b 给路由设备 C，最后发送一个倒换结束通告报文给路由设备 C，告知路由发送完毕，倒换结束。

[0107] 本实施例中，如果是路由设备 A、C 的本地路由在 TCP 连接断开期间无变化，而路由设备 B 的本地路由信息由 b 更新为 b'，在路由设备 B 接收到路由设备 A、C 发送的路由无变化通告报文后仍需再次计算对应发送给邻居路由设备的信息摘要值，且计算出的摘要值与 TCP 连接恢复前计算得出的摘要值不相等，因此需要将路由信息 a、b' 发送给路由设备 C，将路由信息 b'、c 发送给路由设备 A，处理流程与本实施例相同，不再详述。

[0108] 针对上述方法，本发明还提供一种路由设备，本发明提供的设备如图 4 所示，包括：

[0109] 存储模块 401，用于备份自身 BGP 路由信息。

[0110] 处理模块 402，用于发生倒换时，如果 TCP 连接异常断开，切换为平滑重启设备 GR Restarter，进入平滑重启 GR Restart 模式。

[0111] 计算模块 403，用于通过备份的 BGP 路由信息计算与邻居路由设备对应的路由信息摘要值 MD5-1.

[0112] 判断模块 404，用于判断 TCP 连接断开期间路由信息是否发生变化。

[0113] 如果 TCP 连接断开期间路由信息发生变化，所述存储模块 401 还用于更新备份的 BGP 路由信息。

[0114] 如果所述判断模块 404 判断 TCP 连接断开期间路由信息发生变化，所述计算模块 403 还用于在 TCP 连接恢复后通过更新后的 BGP 路由信息再次计算与邻居路由设备对应的路由信息摘要值 MD5-2.

[0115] 判断模块 404 还用于，判断所述 MD5-2 与所述 MD5-1 是否相等。

[0116] 发送模块 405，用于所述 MD5-2 与所述 MD5-1 不相等时，发送 TCP 连接恢复后的路由信息给邻居路由设备，使其更新自身记录的路由信息；还用于所述 MD5-2 与所述 MD5-1 相等时，发送倒换结束通告报文给邻居路由设备，使其删除自身路由信息上的短暂失效标记 Stale，退出平滑重启辅助 GR Help 模式。

[0117] 如果判断模块 404 判断 TCP 连接断开期间路由信息无变化，所述发送模块 505 进一步用于，发送倒换结束通告报文给邻居路由设备，使其删除自身路由信息上的短暂失效标记 Stale，退出 GR Help 模式。

[0118] 该设备还包括接收模块 405，用于接收邻居路由设备发送的路由更新信息，发送第一通知报文给判断模块，所述路由更新信息是所述邻居路由设备在 TCP 连接恢复前后分别计算与所述第一路由设备对应的路由信息摘要值，且判定 TCP 连接恢复前后两次计算的信息摘要值不相等时发送给所述第一路由设备的。

[0119] 所述判断模块 404 进一步用于,通过本地路由信息发生变化判定 TCP 连接断开期间路由信息发生变化,和 / 或接收到所述接收模块发送的第一通知报文判定 TCP 连接断开期间路由信息发生变化。

[0120] 所述接收模块 406 还用于,接收到所有邻居路由设备发送的路由无变化通告报文,发送第二通知报文给判断模块。

[0121] 所述判断模块 404 进一步用于,接收所述接收模块发送的第二通知报文,判定本地路由信息无变化。

[0122] 所述处理模块 402 进一步用于,如果所述 TCP 连接正常,则按照 BGP NSR 方式执行倒换。

[0123] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。

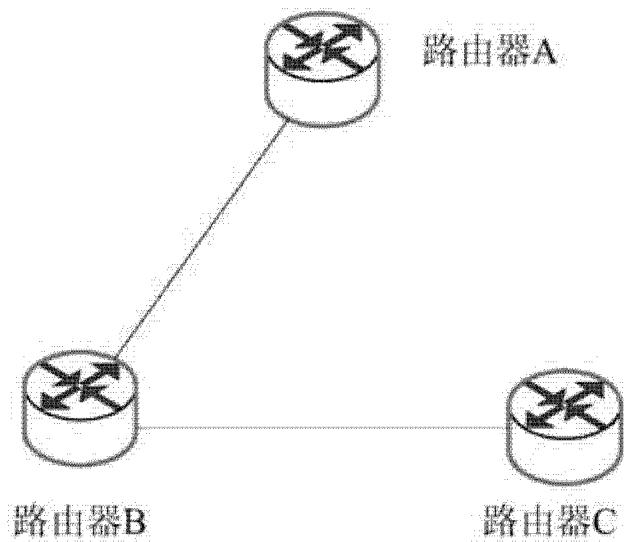
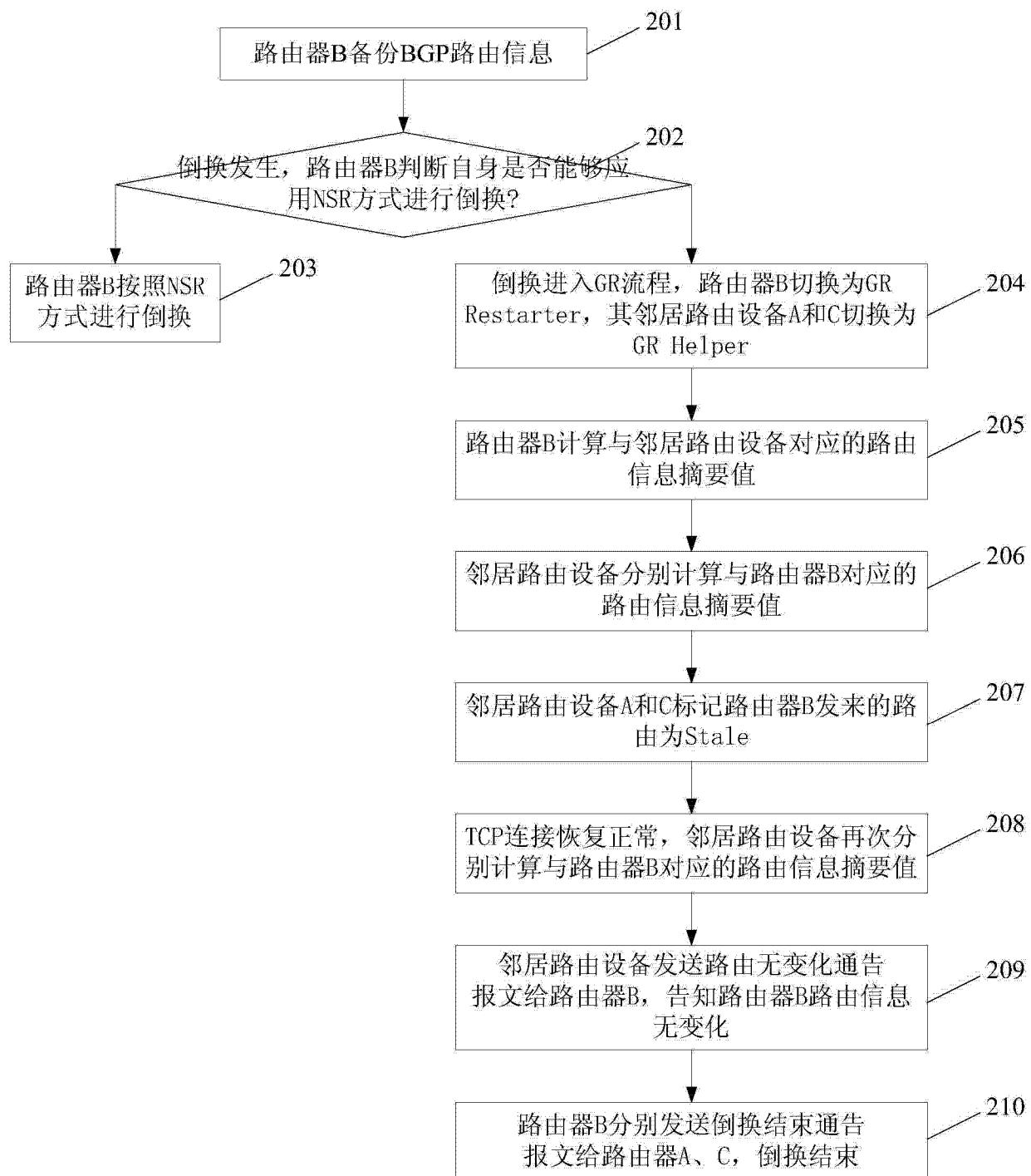


图 1



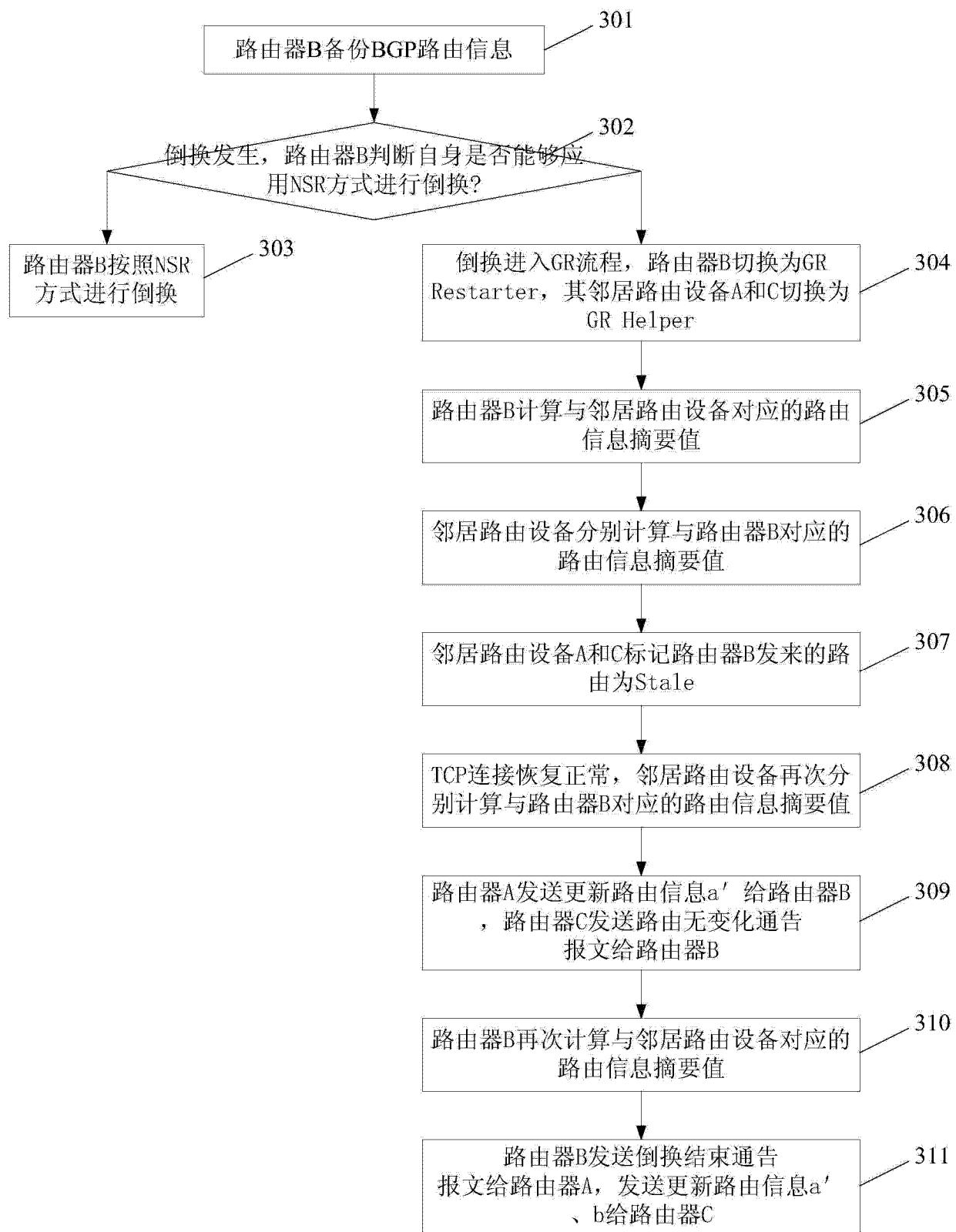


图 3

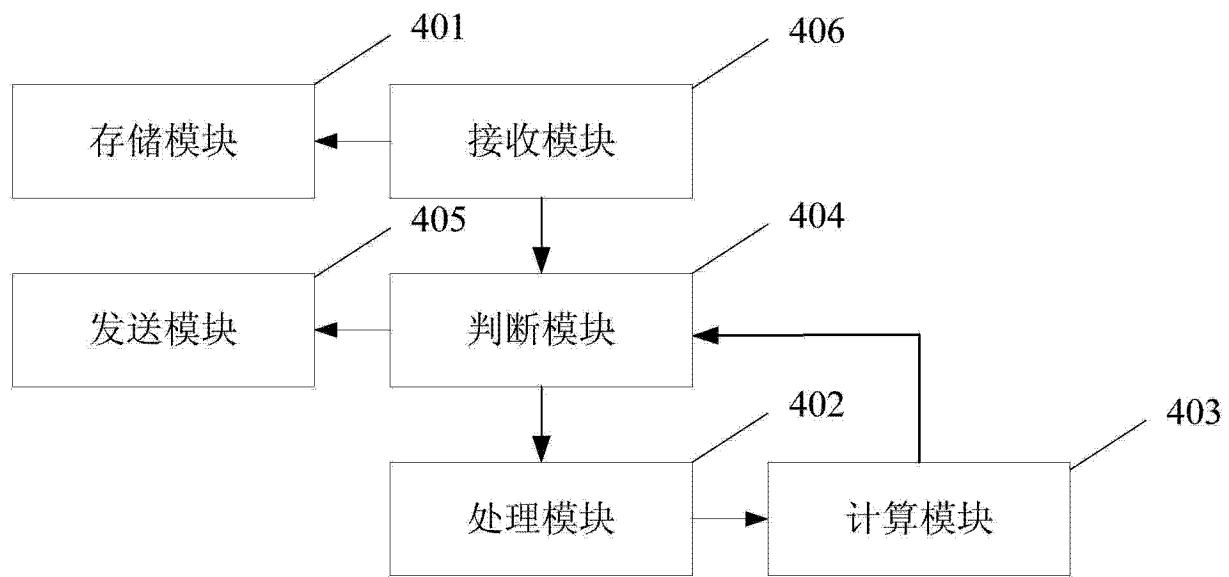


图 4