



(12) 发明专利

(10) 授权公告号 CN 107666438 B

(45) 授权公告日 2021. 10. 22

(21) 申请号 201610601852.5

H04L 12/751 (2013.01)

(22) 申请日 2016.07.27

(56) 对比文件

(65) 同一申请的已公布的文献号

US 2014098711 A1, 2014.04.10

申请公布号 CN 107666438 A

CN 103825818 A, 2014.05.28

(43) 申请公布日 2018.02.06

胡乔林等.BA-BGP:一种基于备份AS通告的域间路由协议.《计算机工程与科学》.2011,(第01期),

(73) 专利权人 中兴通讯股份有限公司

地址 518057 广东省深圳市南山区科技园路55号

审查员 常志沛

(72) 发明人 彭少富 王会来 孙晋松

(74) 专利代理机构 北京康信知识产权代理有限公司 11240

代理人 江舟 董文倩

(51) Int. Cl.

H04L 12/721 (2013.01)

H04L 12/723 (2013.01)

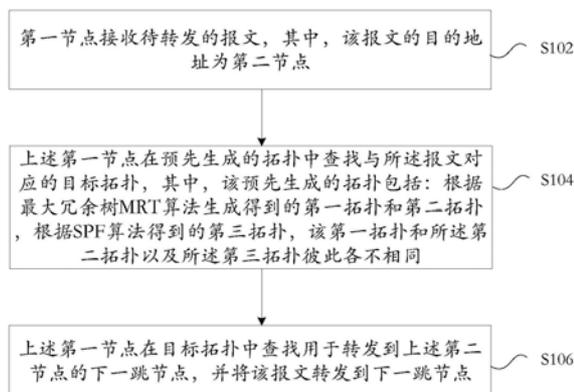
权利要求书3页 说明书19页 附图4页

(54) 发明名称

报文转发方法及装置

(57) 摘要

本发明提供了一种报文转发方法及装置,其中,该方法包括:第一节点接收待转发的报文,其中,该报文的目的地地址为第二节点;上述第一节点在预先生成的拓扑中查找与所述报文对应的目标拓扑,其中,该预先生成的拓扑包括:根据MRT算法生成得到的第一拓扑和第二拓扑,根据SPF算法得到的第三拓扑,该第一拓扑和所述第二拓扑以及所述第三拓扑彼此各不相同;上述第一节点在目标拓扑中查找用于转发到上述第二节点的下一跳节点,并将该报文转发到下一跳节点。通过本发明,实现了分段路由网络与MRT功能结合的目的。



1. 一种报文转发方法,其特征在于,包括:

第一节点接收待转发的报文,其中,所述报文的目的地地址为第二节点;

所述第一节点在预先生成的拓扑中查找与所述报文对应的目标拓扑,其中,所述预先生成的拓扑包括:根据最大冗余树MRT算法生成得到的第一拓扑和第二拓扑,根据最短路径优先SPF算法得到的第三拓扑,所述第一拓扑和所述第二拓扑以及所述第三拓扑彼此各不相同;

所述第一节点在所述目标拓扑中查找用于转发到所述第二节点的下一跳节点,并基于预定转发机制将所述报文转发到所述下一跳节点,其中,所述预定转发机制为基于每拓扑每分段路由全局块SRGB的非隧道嵌套方式的分段路由转发机制。

2. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

所述第一节点按照所述MRT算法从所述第一拓扑中至所述目的地地址的路由和所述第二拓扑中至所述目的地地址的路由中确定用于保护所述第三拓扑中至所述目的地地址的路由的保护路由,并确定所述保护路由对应的拓扑为保护拓扑。

3. 根据权利要求2所述的方法,其特征在于,所述第一节点在预先生成的拓扑中查找与所述报文对应的目标拓扑包括:

所述第一节点判断所述第三拓扑中的用于到达所述第二节点的链路是否出现故障;

在判断没有出现故障的情况下,所述第一节点确定所述第三拓扑为所述目标拓扑;和/或,

在判断出现故障的情况下,所述第一节点确定所述保护拓扑为所述目标拓扑。

4. 根据权利要求1所述的方法,其特征在于,所述第一节点在接收待转发的所述报文之前,所述方法还包括:

所述第一节点根据所述MRT算法生成得到所述第一拓扑和第二拓扑,以及根据所述SPF算法生成得到所述第三拓扑。

5. 根据权利要求4所述的方法,其特征在于,所述第一节点根据所述MRT算法生成得到所述第一拓扑和第二拓扑,以及根据所述SPF算法生成得到所述第三拓扑包括:

所述第一节点确定所述第一节点所在的MRT Island,其中,所述MRT Island是通过在所述第一节点以及与所述第一节点处于同一域area或同一层次level的其他节点上的开放最短路径优先OSPF或者中间系统到中间系统ISIS实例下使能分段路由SR以及最大冗余树配置文件MRT profile后在所述第一节点所在的area或level内由所述第一节点和所述其他节点相互协商形成的;

所述第一节点基于所述MRT Island运行所述MRT算法生成所述第一拓扑和所述第二拓扑,以及,基于所述area或level运行所述SPF算法生成所述第三拓扑。

6. 根据权利要求5所述的方法,其特征在于,所述MRT profile中指定采用所述预定转发机制。

7. 根据权利要求5所述的方法,其特征在于,所述方法还包括以下至少之一:

所述第一节点为所述第一拓扑分配第一分段路由全局块SRGB,并将所述第一SRGB在所述MRT Island所在的域area或层次level内泛洪;

所述第一节点为所述第二拓扑分配第二分段路由全局块SRGB,并将所述第二SRGB在所述MRT Island所在的域area或层次level内泛洪;

所述第一节点为所述第三拓扑分配第三分段路由全局块SRGB,并将所述第三SRGB在所述第一节点所在的所有域area或层次level内泛洪;

所述第一节点接收其他节点上的所述预先生成的拓扑的分段路由全局块SRGB,记录所述其他节点上的所述预先生成的拓扑的SRGB以及将所述其他节点上的所述预先生成的拓扑的SRGB继续通告给除所述其他节点之外的节点。

8.根据权利要求1所述的方法,其特征在于,所述第一节点将所述报文转发到所述下一跳节点包括:

所述第一节点确定所述第一节点的出标签;

所述第一节点将所述第一节点的出标签封装到所述报文上,并将封装后的报文发送到所述下一跳节点。

9.根据权利要求8所述的方法,其特征在于,所述第一节点确定所述第一节点的出标签包括:

所述第一节点判断所述下一跳节点是否为所述第一节点所在的MRT Island外的节点;

在判断结果为否时,所述第一节点确定所述第一节点的对应所述目标拓扑的出标签,其中,所述第一节点的对应所述目标拓扑的出标签由所述第一节点基于所述第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述目标拓扑的SRGB计算得到的;和/或,

在判断结果为是时,所述第一节点确定所述第一节点的对应所述目标拓扑的出标签,其中,所述第一节点的对应所述目标拓扑的出标签由所述第一节点基于所述第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述第三拓扑的SRGB计算得到的。

10.根据权利要求8所述的方法,其特征在于,所述第一节点确定所述第一节点的出标签包括:

所述第一节点判断所述下一跳节点是否为需要将所述报文回归到第三拓扑的区域边界路由器ABR或层次边界路由器LBR;

在判断结果为否时,所述第一节点确定所述第一节点的对应所述目标拓扑的出标签,其中,所述第一节点的对应所述目标拓扑的出标签由所述第一节点基于所述第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述目标拓扑的SRGB计算得到的;和/或,

在判断结果为是时,所述第一节点确定所述第一节点的对应所述目标拓扑的出标签,其中,所述第一节点的对应所述目标拓扑的出标签由所述第一节点基于所述第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述第三拓扑的SRGB计算得到的。

11.根据权利要求10所述的方法,其特征在于,所述第一节点判断所述下一跳节点是否为需要将所述报文回归到第三拓扑的ABR或LBR包括:

所述第一节点判断从所述下一跳节点接收的预先生成的第三拓扑的目的前缀prefix对应的prefix-sid中是否携带有预定标志,其中,所述预定标志表示所述目的prefix是被ABR或LBR从area或level外泄露到本area或level的,并且由ABR或LBR打上所述预定标志,以使得至所述目的prefix的报文在离开ABR或LBR后回归到所述第三拓扑;

在判断结果为从所述下一跳节点接收的所述预先生成的第三拓扑的目的prefix对应的prefix-sid中携带有所述预定标志,并且所述目的prefix在所述第三拓扑内的通告节点为所述下一跳节点时,确定所述下一跳节点为需要将所述报文回归到第三拓扑的ABR或LBR;

在判断结果为从所述下一跳节点接收的所述预先生成的第三拓扑的目的prefix对应的prefix-sid中未携带有所述预定标志,或者在判断结果为携带有所述预定标志但是所述目的prefix在所述第三拓扑内的通告节点不为所述下一跳节点时,确定所述下一跳节点不为需要将所述报文回归到第三拓扑的ABR或LBR。

12. 根据权利要求11所述的方法,其特征在于,所述方法还包括:

当所述第一节点为ABR或LBR时,所述第一节点将所述目的prefix向与所述目的prefix最短路径下一跳节点所处area或level不同的area或level内的邻居通告时,所述目的prefix相应的prefix-sid中将携带有所述预定标志;所述第一节点将所述目的prefix向与所述目的prefix最短路径下一跳节点所处area或level相同的area或level内的邻居通告时,所述目的prefix相应的prefix-sid中将不携带所述预定标志。

13. 根据权利要求7所述的方法,其特征在于,所述第一节点确定所述第一节点的出标签包括:

当所述第一节点和所述第二节点为同一节点时,则所述第一节点基于所述目标拓扑确定的所述下一跳节点为所述第一节点,所述第一节点没有出标签。

14. 根据权利要求8所述的方法,其特征在于,所述第一节点将所述第一节点的出标签封装到所述报文上包括以下至少之一:

当所述报文的报文类型为互联网协议IP报文时,在所述IP报文的IP头上压上所述第一节点的出标签;

当所述报文的报文类型为分段路由SR标签报文时,将所述SR标签报文的标签栈的栈顶标签替换成所述第一节点的出标签。

15. 根据权利要求1所述的方法,其特征在于,所述方法还包括:

当所述第一节点和所述第二节点为同一节点时,包括:所述报文的报文类型为互联网协议IP报文时,所述第一节点将所述报文中送至所述第一节点的控制平面;和/或,当所述报文的报文类型为分段路由SR标签报文时,所述第一节点将所述SR标签报文的标签栈的栈顶标签弹出,并继续基于所述报文的标签栈的下层标签或IP头查表转发。

16. 一种报文转发装置,其特征在于,应用于第一节点中,包括:

接收模块,用于接收待转发的报文,其中,所述报文的目的地地址为第二节点;

查找模块,用于在预先生成的拓扑中查找与所述报文对应的目标拓扑,其中,所述预先生成的拓扑包括根据最大冗余树MRT算法生成得到的第一拓扑和第二拓扑,根据最短路径优先SPF算法得到的第三拓扑,所述第一拓扑和所述第二拓扑以及所述第三拓扑彼此各不相同;

转发模块,用于在所述目标拓扑中查找用于转发到所述第二节点的下一跳节点,并基于预定转发机制将所述报文转发到所述下一跳节点,其中,所述预定转发机制为基于每拓扑每分段路由全局块SRGB的非隧道嵌套方式的分段路由转发机制。

## 报文转发方法及装置

### 技术领域

[0001] 本发明涉及通信领域,具体而言,涉及一种报文转发方法及装置。

### 背景技术

[0002] 最大冗余树(Maximally Redundant Trees,简称为MRT)快速重路由(Fast Re-Route,简称为FRR)是一种较新的FRR技术,该技术中使用两个最大限度不同的转发拓扑,对单点的链路或节点故障能提供100%的保护。MRT架构定义了两种转发机制,即标签分发协议(Label Distribution Protocol,简称为LDP)转发机制和网络协议-隧道(IP-tunnel)转发机制。LDP转发机制通过不同的标签来区分是默认拓扑转发行为还是MRT转发行为,使得转发平面不作任何升级即可支持MRT-FRR。IP-tunnel转发机制则需要浪费专用的MRT loopback地址来支持转发,同样也使得转发平面不作任何升级即可支持MRT-FRR。相比而言,LDP转发机制更加合理,所以MRT架构的默认最大冗余树配置文件default MRT Profile中采用的就是LDP转发机制,目前尚未定义其它MRT Profiles。

[0003] 分段路由技术将使得一个节点可以为报文指定其转发路径,而不是按一般的最短路径转发,通过在报文中附加由段标识Segment ID组成的段列表Segment List相关的信息,不需要在中间节点上为维护每路径的状态信息。分段路由主要扩展IGP以支持通告和学习Segment ID。一般在部署了分段路由的网络中,就不再需要部署LDP与基于流量工程扩展的资源预留协议(Resource ReSerVation Protocol-Traffic Extension,简称为RSVP-TE)了。在分段路由网络中,已知的FRR技术有拓扑无关的无环替换路径(Topology Independent Loop Free Alternate,简称为TI-LFA),但是TI-LFA定义的保护规则十分复杂并且还不成熟。

[0004] 在分段路由网络中引入MRT功能将有重要的意义,但是,迄今为止尚未有文献讨论这方面的内容,因此,在相关技术中,无法实现分段路由网络与MRT功能结合。

### 发明内容

[0005] 本发明实施例提供了一种报文转发方法及装置,以至少解决相关技术中存在的无法实现分段路由网络与MRT功能结合的问题。

[0006] 根据本发明的一个实施例,提供了一种报文转发方法,包括:第一节点接收待转发的报文,其中,所述报文的目的地地址为第二节点;所述第一节点在预先生成的拓扑中查找与所述报文对应的目标拓扑,其中,所述预先生成的拓扑包括:根据最大冗余树MRT算法生成得到的第一拓扑和第二拓扑,根据最短路径优先SPF算法得到的第三拓扑,所述第一拓扑和所述第二拓扑以及所述第三拓扑彼此各不相同;所述第一节点在所述目标拓扑中查找用于转发到所述第二节点的下一跳节点,并基于预定转发机制将所述报文转发到所述下一跳节点,其中,所述预定转发机制为基于每拓扑每分段路由全局块SRGB的非隧道嵌套方式的分段路由转发机制。

[0007] 可选地,所述方法还包括:所述第一节点按照所述MRT算法从所述第一拓扑中至所

述目的地址的路由和所述第二拓扑中至所述目的地址的路由中确定用于保护所述第三拓扑中至所述目的地址的路由的保护路由,并确定所述保护路由对应的拓扑为保护拓扑。

[0008] 可选地,所述第一节点在预先生成的拓扑中查找与所述报文对应的目标拓扑包括:所述第一节点判断所述第三拓扑中的用于到达所述第二节点的链路是否出现故障;在判断没有出现故障的情况下,所述第一节点确定所述第三拓扑为所述目标拓扑;和/或,在判断出现故障的情况下,所述第一节点确定所述保护拓扑为所述目标拓扑。

[0009] 可选地,所述第一节点在接收待转发的所述报文之前,所述方法还包括:所述第一节点根据所述MRT算法生成得到所述第一拓扑和第二拓扑,以及根据所述SPF算法生成得到所述第三拓扑。

[0010] 可选地,所述第一节点根据所述MRT算法生成得到所述第一拓扑和第二拓扑,以及根据所述SPF算法生成得到所述第三拓扑包括:所述第一节点确定所述第一节点所在的MRT Island,其中,所述MRT Island是通过在所述第一节点以及与所述第一节点处于同一域area或同一层次level的其他节点上的开放最短路径优先OSPF或者中间系统到中间系统ISIS实例下使能分段路由SR以及最大冗余树配置文件MRT profile后在所述第一节点所在的area或level内由所述第一节点和所述其他节点相互协商形成的;所述第一节点基于所述MRT Island运行所述MRT算法生成所述第一拓扑和所述第二拓扑,以及,基于所述area或level运行所述SPF算法生成所述第三拓扑。

[0011] 可选地,所述MRT profile中指定采用所述预定转发机制。

[0012] 可选地,所述方法还包括以下至少之一:所述第一节点为所述第一拓扑分配第一分段路由全局块SRGB,并将所述第一SRGB在所述MRT Island所在的域area或层次level内泛洪;所述第一节点为所述第二拓扑分配第二分段路由全局块SRGB,并将所述第二SRGB在所述MRT Island所在的域area或层次level内泛洪;所述第一节点为所述第三拓扑分配第三分段路由全局块SRGB,并将所述第三SRGB在所述第一节点所在的所有域area或层次level内泛洪;所述第一节点接收其他节点上的所述预先生成的拓扑的分段路由全局块SRGB,记录所述其他节点上的所述预先生成的拓扑的SRGB以及将所述其他节点上的所述预先生成的拓扑的SRGB继续通告给除所述其他节点之外的节点。

[0013] 可选地,所述第一节点将所述报文转发到所述下一跳节点包括:所述第一节点确定所述第一节点的出标签;所述第一节点将所述第一节点的出标签封装到所述报文上,并将封装后的报文发送到所述下一跳节点。

[0014] 可选地,所述第一节点确定所述第一节点的出标签包括:所述第一节点判断所述下一跳节点是否为所述第一节点所在的MRT Island外的节点;在判断结果为否时,所述第一节点确定所述第一节点的对应所述目标拓扑的出标签,其中,所述第一节点的对应所述目标拓扑的出标签由所述第一节点基于所述第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述目标拓扑的SRGB计算得到的;和/或,在判断结果为是时,所述第一节点确定所述第一节点的对应所述目标拓扑的出标签,其中,所述第一节点的对应所述目标拓扑的出标签由所述第一节点基于所述第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述第三拓扑的SRGB计算得到的。

[0015] 可选地,所述第一节点确定所述第一节点的出标签包括:所述第一节点判断所述下一跳节点是否为需要将所述报文回归到第三拓扑的区域边界路由器ABR或层次边界路由

器LBR;在判断结果为否时,所述第一节点确定所述第一节点的对应所述目标拓扑的出标签,其中,所述第一节点的对应所述目标拓扑的出标签由所述第一节点基于所述第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述目标拓扑的SRGB计算得到的;和/或,在判断结果为是时,所述第一节点确定所述第一节点的对应所述目标拓扑的出标签,其中,所述第一节点的对应所述目标拓扑的出标签由所述第一节点基于所述第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述第三拓扑的SRGB计算得到的。

[0016] 可选地,所述第一节点判断所述下一跳节点是否为需要将所述报文回归到第三拓扑的ABR或LBR包括:所述第一节点判断从所述下一跳节点接收的预先生成的第三拓扑的目的前缀prefix对应的prefix-sid中是否携带有预定标志,其中,所述预定标志表示所述目的prefix是被ABR或LBR从area或level外泄露到本area或level的,并且由ABR或LBR打上所述预定标志,以使得至所述目的prefix的报文在离开ABR或LBR后回归到所述第三拓扑;在判断结果为从所述下一跳节点接收的所述预先生成的第三拓扑的目的prefix对应的prefix-sid中携带有所述预定标志,并且所述目的prefix在所述第三拓扑内的通告节点为所述下一跳节点时,确定所述下一跳节点为需要将所述报文回归到第三拓扑的ABR或LBR;在判断结果为从所述下一跳节点接收的所述预先生成的第三拓扑的目的prefix对应的prefix-sid中未携带有所述预定标志,或者在判断结果为携带有所述预定标志但是所述目的prefix在所述第三拓扑内的通告节点不为所述下一跳节点时,确定所述下一跳节点不为需要将所述报文回归到第三拓扑的ABR或LBR。

[0017] 可选地,当所述第一节点为ABR或LBR时,所述第一节点将所述目的prefix向与所述目的prefix最短路径下一跳节点所处area或level不同的area或level内的邻居通告时,所述目的prefix相应的prefix-sid中将携带有所述预定标志;所述第一节点将所述目的prefix向与所述目的prefix最短路径下一跳节点所处area或level相同的area或level内的邻居通告时,所述目的prefix相应的prefix-sid中将不携带所述预定标志。

[0018] 可选地,所述第一节点确定所述第一节点的出标签包括:当所述第一节点和所述第二节点为同一节点时,则所述第一节点基于所述目标拓扑确定的所述下一跳节点为所述第一节点,所述第一节点没有出标签。

[0019] 可选地,所述第一节点将所述第一节点的出标签封装到所述报文中包括以下至少之一:当所述报文的报文类型为互联网协议IP报文时,在所述IP报文的IP头上压上所述第一节点的出标签;当所述报文的报文类型为分段路由SR标签报文时,将所述SR标签报文的标签栈的栈顶标签替换成所述第一节点的出标签。

[0020] 可选地,所述方法还包括:当所述第一节点和所述第二节点为同一节点时,包括:所述报文的报文类型为互联网协议IP报文时,所述第一节点将所述报文中送至所述第一节点的控制平面;和/或,当所述报文的报文类型为分段路由SR标签报文时,所述第一节点将所述SR标签报文的标签栈的栈顶标签弹出,并继续基于所述报文的标签栈的下层标签或IP头查表转发。

[0021] 根据本发明的一个实施例,提供了一种报文转发装置,所述装置于第一节点中,包括:第一接收模块,用于接收待转发的报文,其中,所述报文的地址为第二节点;第一查找模块,用于在预先生成的拓扑中查找与所述报文对应的目标拓扑,其中,所述预先生成的

拓扑包括根据最大冗余树MRT算法生成得到的第一拓扑和第二拓扑,根据最短路径优先SPF算法得到的第三拓扑,所述第一拓扑和所述第二拓扑以及所述第三拓扑彼此各不相同;第一转发模块,用于在所述目标拓扑中查找用于转发到所述第二节点的下一跳节点,并基于预定转发机制将所述报文转发到所述下一跳节点,其中,所述预定转发机制为基于每拓扑每分段路由全局块SRGB的非隧道嵌套方式的分段路由转发机制。

[0022] 根据本发明的又一个实施例,还提供了一种存储介质。该存储介质设置为存储用于执行上述步骤的程序代码。

[0023] 通过本发明,在分段路由网络中引入MRT功能,从而实现了分段路由网络与MRT功能结合的目的。

## 附图说明

[0024] 此处所说明的附图用来提供对本发明的进一步理解,构成本申请的一部分,本发明的示意性实施例及其说明用于解释本发明,并不构成对本发明的不当限定。在附图中:

[0025] 图1是根据本发明实施例的报文转发的流程图;

[0026] 图2是根据本发明实施例的MRT Profile选项示意图;

[0027] 图3是根据本发明实施例的ISIS MT SR-Capabilities Sub-TLV的格式图;

[0028] 图4是根据本发明实施例的OSPF MT SID/Label Range TLV的格式图;

[0029] 图5是根据本发明实施例的ISIS Prefix-SID Sub-TLV的格式图;

[0030] 图6是根据本发明实施例的OSPF Prefix SID Sub-TLV的格式图;

[0031] 图7是根据本发明具体实施例一的网络拓扑图;

[0032] 图8是根据本发明具体实施例二的网络拓扑图;

[0033] 图9是根据本发明具体实施例三的网络拓扑图;

[0034] 图10是根据本发明实施例的报文转发装置的结构框图。

## 具体实施方式

[0035] 下文中将参考附图并结合实施例来详细说明本发明。需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互组合。

[0036] 需要说明的是,本发明的说明书和权利要求书及上述附图中的术语“第一”、“第二”等是用于区别类似的对象,而不必用于描述特定的顺序或先后次序。

[0037] 在本实施例中提供了一种报文转发方法,图1是根据本发明实施例的报文转发的流程图,如图1所示,该流程包括如下步骤:

[0038] 步骤S102,第一节点接收待转发的报文,其中,该报文的地址为第二节点;

[0039] 步骤S104,上述第一节点在预先生成的拓扑中查找与所述报文对应的目标拓扑,其中,该预先生成的拓扑包括:根据最大冗余树MRT算法生成得到的第一拓扑和第二拓扑,根据最短路径优先(Shortest Path First,简称为SPF)算法得到的第三拓扑,该第一拓扑和所述第二拓扑以及所述第三拓扑彼此各不相同;

[0040] 步骤S106,上述第一节点在目标拓扑中查找用于转发到上述第二节点的下一跳节点,并将该报文转发到下一跳节点。

[0041] 其中,上述的第一节点可以是MRT Island内的任何节点,上述第二节点可以是SR

Domain内的任何节点,该第二节点可能处于所述MRT Island内或MRT Island外。上述的第一拓扑和第二拓扑可以一个是MRT-red拓扑,一个是MRT-blue拓扑;上述的第三拓扑可以是MT-default拓扑。

[0042] 在上述实施例中,第一节点在将报文转发到下一跳节点时,可以基于预定转发机制进行转发,该预定转发机制为基于每拓扑每分段路由全局块SRGB的非隧道嵌套方式的分段路由转发机制。

[0043] 通过上述步骤,在分段路由网络中引入MRT功能,从而实现了分段路由网络与MRT功能结合的目的。

[0044] 在一个可选的实施例中,上述方法还包括:第一节点按照MRT算法从第一拓扑中至目的地址的路由和第二拓扑中至目的地址的路由中确定用于保护第三拓扑中至目的地址的路由的保护路由,并确定该保护路由对应的拓扑为保护拓扑。在本实施例中,在确定由第一节点至目的地址(即,第二节点)的保护路由时,可以确定第一拓扑中至目的地址的路由作为保护路由,也可以确定第二拓扑中至目的地址的路由作为保护路由,具体确定那个拓扑中至目的地址的路由作为保护路由,需要根据实际情况进行确定。

[0045] 在一个可选的实施例中,上述第一节点在预先生成的拓扑中查找与上述报文对应的目标拓扑包括:上述第一节点判断第三拓扑中的用于到达第二节点的链路是否出现故障;在判断没有出现故障的情况下,上述第一节点确定第三拓扑为目标拓扑;和/或,在判断出出现故障的情况下,上述第一节点确定保护拓扑为目标拓扑。在本实施例中,当链路没有故障时,可以继续按照默认的拓扑进行报文转发,当链路出现故障后,需要使用保护拓扑进行报文转发。

[0046] 在一个可选的实施例中,上述第一节点在接收待转发的报文之前,上述方法还包括:第一节点根据MRT算法生成得到第一拓扑和第二拓扑,以及根据SPF算法生成得到第三拓扑。

[0047] 在一个可选的实施例中,上述第一节点根据MRT算法生成得到第一拓扑和第二拓扑,以及根据SPF算法生成得到第三拓扑包括:第一节点确定第一节点所在的MRT Island,其中,该MRT Island是通过在第一节点以及与第一节点处于同一域area或同一层次level的其他节点上的开放最短路径优先(Open shortest Path First,简称为OSPF)或者中间系统到中间系统(Intermediate system to Intermediate system,简称为ISIS)实例下使能分段路由SR以及最大冗余树配置文件MRT profile后在第一节点所在的area或level内由第一节点和其他节点相互协商形成的;该第一节点基于上述MRT Island运行MRT算法生成上述第一拓扑和第二拓扑,以及,基于上述area或level运行上述SPF算法生成第三拓扑。

[0048] 在一个可选的实施例中,上述MRT profile中指定采用上述预定转发机制,即,采用基于每拓扑每分段路由全局块SRGB的非隧道嵌套方式的分段路由转发机制。在本实施例中,在非隧道嵌套方式中,在报文上封装标签时,该标签将直接指导报文转发至第二节点。

[0049] 在一个可选的实施例中,上述方法还包括以下至少之一:第一节点为第一拓扑分配第一分段路由全局块SRGB,并将该第一SRGB在上述MRT Island所在的域area或层次level内泛洪;第一节点为第二拓扑分配第二分段路由全局块SRGB,并将第二SRGB在上述MRT Island所在的域area或层次level内泛洪;第一节点为第三拓扑分配第三分段路由全局块SRGB,并将第三SRGB在第一节点所在的所有域area或层次level内泛洪;第一节点接收

其他节点上的上述预先生成的拓扑的分段路由全局块SRGB,记录其他节点上的预先生成的拓扑的SRGB以及将其他节点上的预先生成的拓扑的SRGB继续通告给除上述其他节点之外的节点。在本实施例中,各节点都可以为第一拓扑、第二拓扑、第三拓扑分别生成一个SRGB,针对不同的拓扑生成的SRGB是不同的,且不同的节点生成的SRGB是独立的。

[0050] 在一个可选的实施例中,上述第一节点将上述报文转发到下一跳节点包括:第一节点确定上述第一节点的出标签;该第一节点将第一节点的出标签封装到上述报文中,并将封装后的报文发送到下一跳节点。在本实施例中,针对不同类型的报文,标签的封装方式是不同的。

[0051] 在一个可选的实施例中,上述第一节点确定第一节点的出标签包括:第一节点判断下一跳节点是否为第一节点所在的MRT Island外的节点;在判断结果为否时,第一节点确定第一节点的对应目标拓扑的出标签,其中,该第一节点的对应目标拓扑的出标签由第一节点基于第二节点的第三拓扑的前缀段索引prefix-sid与下一跳节点的对应目标拓扑的SRGB计算得到的;和/或,在判断结果为是时,第一节点确定第一节点的对应目标拓扑的出标签,其中,该第一节点的对应目标拓扑的出标签由第一节点基于第二节点的第三拓扑的前缀段索引prefix-sid与下一跳节点的对应第三拓扑的SRGB计算得到的。

[0052] 在一个可选的实施例中,上述第一节点确定第一节点的出标签包括:第一节点判断下一跳节点是否为需要将报文回归到第三拓扑的区域边界路由器(Area Border Router,简称为ABR或层次边界路由器(Level Border Router,简称为LBR);在判断结果为否时,第一节点确定第一节点的对应目标拓扑的出标签,其中,该第一节点的对应目标拓扑的出标签由第一节点基于第二节点的第三拓扑的前缀段索引prefix-sid与下一跳节点的对应目标拓扑的SRGB计算得到的;和/或,在判断结果为是时,第一节点确定第一节点的对应目标拓扑的出标签,其中,该第一节点的对应目标拓扑的出标签由第一节点基于第二节点的第三拓扑的前缀段索引prefix-sid与下一跳节点的对应第三拓扑的SRGB计算得到的。

[0053] 在一个可选的实施例中,上述第一节点判断下一跳节点是否为需要将报文回归到第三拓扑的ABR或LBR包括:第一节点判断从下一跳节点接收的预先生成的第三拓扑的目的前缀prefix对应的prefix-sid中是否携带有预定标志,其中,该预定标志表示目的prefix是被ABR或LBR从area或level外泄露到本area或level的,并且由ABR或LBR打上该预定标志,以使得至目的prefix的报文在离开ABR或LBR后回归到第三拓扑;在判断结果为从所述下一跳节点接收的所述预先生成的第三拓扑的目的prefix对应的prefix-sid中携带有所述预定标志,并且目的prefix在第三拓扑内的通告节点为上述下一跳节点时,确定上述下一跳节点为需要将报文回归到第三拓扑的ABR或LBR;在判断结果为从上述下一跳节点接收的预先生成的第三拓扑的目的prefix对应的prefix-sid中未携带有预定标志,或者在判断结果为携带有上述预定标志但是目的prefix在第三拓扑内的通告节点不为上述下一跳节点时,确定上述下一跳节点不为需要将报文回归到第三拓扑的ABR或LBR。在本实施例中,上述的预定标志可以是Rainbow,当然,也可以是其他的标志。

[0054] 在一个可选的实施例中,当上述第一节点为ABR或LBR时,则它将目的prefix向与目的prefix最短路径下一跳节点所处area或level不同的area或level内的邻居通告时,上述目的prefix相应的prefix-sid中将携带有所述预定标志;它将目的prefix向与目的prefix最短路径下一跳节点所处area或level相同的area或level内的邻居通告时,上述目

的prefix相应的prefix-sid中将不携带所述预定标志。

[0055] 在一个可选的实施例中,上述第一节点确定第一节点的出标签包括:当上述第一节点和第二节点为同一节点时,则第一节点基于目标拓扑确定的下一跳节点为第一节点,该第一节点没有出标签。即,如果上述第一节点就是第二节点,则第一节点基于目标拓扑确定的下一跳节点就是自身,此时没有出标签。

[0056] 在一个可选的实施例中,上述第一节点将第一节点的出标签封装到上述报文上包括以下至少之一:当上述报文的报文类型为互联网协议(Internet Protocol,简称为IP)报文时,在上述IP报文的IP头上压上第一节点的出标签;当上述报文的报文类型为分段路由SR标签报文时,将上述SR标签报文的标签栈的栈顶标签替换成第一节点的出标签。

[0057] 在一个可选的实施例中,上述方法还包括:当第一节点和第二节点为同一节点时,包括:当上述报文的报文类型为互联网协议IP报文时,上述第一节点将报文中送至第一节点的控制平面;和/或,当上述报文的报文类型为分段路由SR标签报文时,上述第一节点将SR标签报文的标签栈的栈顶标签弹出,并继续基于上述报文的标签栈的下层标签或IP头查表转发。

[0058] 下面结合具体实施例对本发明进行说明:

[0059] 在本发明实施例中,将分段路由与MRT技术结合,提供一种基于分段路由转发机制的MRT-FRR方法。在本实施例中为MRT架构新增一种MRT Profile,使用分段路由转发机制,通过分段路由转发机制来区分默认拓扑转发行为和MRT拓扑转发行为。

[0060] 本实施例中的基于分段路由转发机制的MRT-FRR方法包括以下步骤:

[0061] 第一步,定义新的MRT Profile(如图2所示),与default MRT Profile相比,差异主要体现在定义的新的MRT Profile中使用MRT SR-LSP based on multi-SRGB non-tunneling转发机制。SR-LSP是指基于prefix-sid生成的SR(Segment Routing,即分段路由)类型的LSP,与LDP LSP是类似的。MRT SR-LSP based on multi-SRGB non-tunneling转发机制是指MRT拓扑内的SR LSP是基于目的prefix建立的,即发生MRT FRR切换后,在MRT Ingress节点上,默认拓扑内的SR LSP不是隧道嵌套于MRT拓扑的SR LSP中,而是直接粘连,也即将默认拓扑对应的SR标签交换成MRT拓扑对应的SR标签。采用MRT SR-LSP based on multi-SRGB non-tunneling转发机制时,MRT Ingress节点不需要感知MRT Egress节点为目的prefix分配的默认拓扑对应的SR标签。其中,multi-SRGB是指每拓扑每SRGB,即默认拓扑与MRT-red拓扑以及MRT-blue拓扑中的SRGB是不同的。

[0062] 第二步,在IGP area/level内各节点(可以是仅部分节点)上相应IGP实例下使能MRT并且支持上述新的MRT Profile,针对上述新的MRT Profile生成相应的MRT Island。基于该MRT Island运行MRT算法生成相应的MRT-red拓扑与MRT-blue拓扑,相应的MT-ID分别记为MT-red和MT-blue。另外可以将基于SPF算法生成的默认拓扑对应的MT-ID记为MT-default。

[0063] 第三步,使能了分段路由的各个节点上的IGP实例,如果还使能了MRT并且支持上述新的MRT Profile,则不仅为MT-default分配MT-default SRGB,还为MT-red分配MT-red SRGB,以及为MT-blue分配MT-blue SRGB。其中,MT-default SRGB将在该节点所处的每个area或level内泛洪,而MT-red SRGB与MT-blue SRGB仅在上述MRT Island所在的area或level内泛洪。另外可以采取仅为(MT-default,prefix)分配相应的prefix-sid并随(MT-

default,prefix)泛洪,可以不再分别为(MT-red,prefix)和(MT-blue,prefix)分配相应的prefix-sid以及分别随(MT-red,prefix)和(MT-blue,prefix)泛洪。

[0064] 需要说明的是,冗余泛洪(MT-red,prefix)和(MT-blue,prefix)相应的prefix-sid将在一定程度上影响网络性能,并且因为是MRT相关的信息所以也只能在MRT Island所在的area内泛洪,如果该prefix是area外的prefix,那么该prefix在原始通告节点上是不会感知MT-red和MT-blue拓扑的,那么也无从谈起在原始通告节点上分配(MT-red,prefix)和(MT-blue,prefix)相应的prefix-sid并泛洪至MRT-Island.MRT Island内的节点倒是可以通过其它方式比如分段路由映射服务器SRMS来代理分配(MT-red,prefix)和(MT-blue,prefix)相应的prefix-sid,但是这种方法不具备实际可操作性。

[0065] 第四步,MRT Island内的源节点S为Island内或Island外的prefix计算SPF主下一跳以及MRT-blue/red下一跳,即(MT-default,prefix)中将包含SPF主下一跳,(MT-blue,prefix)中将包含MRT-blue下一跳,(MT-red,prefix)中将包含MRT-red下一跳。并且通过MRT算法确定到底是上述MRT-blue下一跳还是MRT-red下一跳保护上述SPF主下一跳。

[0066] 上述SPF主下一跳中的MT-default SR出标签基于(MT-default,prefix)相应的prefix-sid与下一跳节点的MT-default SRGB来计算;上述MRT-blue下一跳中的MRT-blue SR出标签基于(MT-default,prefix)相应的prefix-sid与MRT-blue下一跳节点的MT-blue SRGB计算;上述MRT-red下一跳中的MRT-red SR出标签基于(MT-default,prefix)相应的prefix-sid与MRT-red下一跳节点的MT-red SRGB计算。

[0067] 根据(MT-default,prefix)将生成相应的转发等价类至下一跳标签转发单元(Forwarding Equivalence Class to NHLFE,简称为FTN)表项,主NHLFE包含上述SPF主下一跳及相应的MT-default SR出标签,MRT-FRR备NHLFE包含上述选中用于保护SPF主下一跳的MRT-red下一跳或MRT-blue下一跳及相应的MRT-red SR出标签或MRT-blue SR出标签。根据(MT-default,prefix)也生成相应的ILM表项,其MT-default SR入标签基于(MT-default,prefix)相应的prefix-sid与S节点的MT-default SRGB来计算,NHLFE与上述FTN表项相同。如果prefix为S节点本地或直连prefix,则没有NHLFE信息。

[0068] 第五步,S节点另外还为(MT-blue,prefix)生成相应的入标签映射(Incoming Label Map,简称为ILM)表项,其MT-blue SR入标签基于(MT-default,prefix)相应的prefix-sid与S节点的MT-blue SRGB来计算,NHLFE包含MRT-blue下一跳及相应的MT-blue SR出标签。亦为(MT-red,prefix)生成相应的ILM表项,其MT-red SR入标签基于(MT-default,prefix)相应的prefix-sid与S节点的MT-red SRGB来计算,NHLFE包含MRT-red下一跳及相应的MT-red SR出标签。如果prefix为S节点本地或直连prefix,则没有NHLFE信息。

[0069] 存在以下两个特殊的场景,使得为(MT-blue,prefix)和(MT-red,prefix)生成相应的ILM表项的NHLFE需特别考虑,如下:

[0070] 1) 使用Named Proxy-node方法计算至MRT Island外的MRT路径:

[0071] Named Proxy-node方法可以参考RFC7812.针对某个MRT Island外的prefix,S节点如果是MRT Island的proxy-node attachment router,则S将为该(MT-default,prefix)生成相应的FTN表项与ILM表项,NHLFE只包含MT-default拓扑内的主下一跳及相应的MT-default SR出标签。

[0072] S还将根据自己如果是red proxy-node attachment router (即RFC7811中连接proxy-node的Y节点), 则产生 (MT-red, prefix) 对应的ILM表项, 其MT-red SR入标签基于 (MT-default, prefix) 相应的prefix-sid与S节点的MT-red SRGB来计算, 而NHLFE包含的下一跳转发信息不一定是最短路径转发信息, 如果S是IBR, 则NHLFE中的下一跳指向特定的LFIN (Loop-Free Island Neighbor, 该LFIN的存在使得S被选中作为proxy-node attachment router, 它不一定是S至prefix的最短路径下一跳) 及相应的MT-default SR出标签 (基于 (MT-default, prefix) 相应的prefix-sid与LFIN节点的MT-default SRGB计算得到)。如果S不是IBR, 则NHLFE中包含的下一跳指向引起S通告该prefix的出接口 (该出接口可能离开area/level/AS) 及相应的MT-default SR出标签 (基于 (MT-default, prefix) 相应的prefix-sid与上述出接口对端节点的MT-default SRGB计算得到)。

[0073] 需要说明的是, 特别是attachment router为ABR时, 且目的prefix所在的area2与S在的area1都支持相同的MRT profile, 则ABR节点上为 (MT-red, prefix) 或 (MT-blue, prefix) 产生的SR ILM表项的NHLFE仍然是基于相应MT-red SRGB或MT-blue SRGB计算的, 这本质上是area2内MRT Island内的正常SR ILM表项建立过程, 与area1无关。这种场景下attachment router如何将沿MRT路径转发的流量切换为走默认拓扑请参考以下“2) 跨域转发行为”。

[0074] 类似的, S还将根据自己如果是blue proxy-node attachment router (即RFC7811中连接proxy-node的X节点), 则产生 (MT-blue, prefix) 对应的ILM表项, 在此, 不再赘述。

[0075] 2) 跨域转发行为

[0076] 根据MRT架构描述, 对于OSPF而言, 目的地址在area内的流量在沿MRT-red或MRT-blue路径转发时, 需要呆在MRT-Red或MRT-Blue拓扑内。但是对于离开area的流量, 则希望它离开MRT-Red或MRT-Blue拓扑, 重新回到最短路径转发的默认拓扑。ISIS是类似的。在本实施例中是遵循MRT架构的。

[0077] S如果作为ABR节点, 则对于那些与至目的prefix在同一area内的IGP邻居, ABR将通告如下prefix-sid:

[0078] prefix-sid for (MT-default, prefix)

[0079] 而对于那些与至目的prefix不在同一area内的IGP邻居, ABR将通告如下prefix-sid:

[0080] prefix-sid for (MT-default, prefix) with flag Rainbow

[0081] 邻居节点收到ABR的带rainbow标志 (对应于上述的预定标志) 的prefix-sid通告后, 判断该prefix为area外的prefix且相应的 (MT-blue, prefix) 的MRT-blue下一跳若为prefix的通告节点, 即该ABR节点, 则认为 (MT-blue, prefix) 的MT-blue SR出标签要替换成MT-default SR出标签, 否则无需替换。类似的, 判断该prefix相应的 (MT-red, prefix) 的MRT-red下一跳若为该ABR节点, 则认为 (MT-red, prefix) 的MT-red SR出标签要替换成MT-default SR出标签, 否则无需替换。这样使得跨area的报文在到达相应的ABR前可能走MRT拓扑, 但在该ABR之后开始走默认拓扑。

[0082] 第六步, 故障发生时, 可以按照如下方法沿MRT路径转发IP或SR标签单播流量:

[0083] MRT ingress节点基于 (MT-default, prefix) 相应的FTN表项指导IP报文, 或ILM表项指导SR标签报文转发, 将流量切换至NHLFE中包含的MRT备份路径, 比如MRT-red下一跳,

则将SR标签报文的顶层标签MT-default SR入标签交换成MR-red SR出标签后发往MRT-red 下一跳,或将IP报文直接压上MR-red SR出标签后发往MRT-red下一跳。

[0084] MRT transit节点基于(MT-red,prefix)相应的ILM表项指导报文转发,继续将MRT-red SR入标签交换成MRT-red SR出标签后发往MRT-red下一跳。

[0085] MRT egress节点基于(MT-red,prefix)相应的ILM表项指导报文转发,将MRT-red SR入标签交换成MT-default SR出标签后向NHLFE中给定的下一跳转发或者弹掉MRT-red SR入标签后基于报文IP头转发。

[0086] 上述MRT ingress节点上MRT备份路径为MRT-blue下一跳时是类似的,不再赘述。

[0087] 下面结合附图对技术方案的实施作进一步的详细描述:

[0088] 本具体实施方式中,我们先具体介绍一下本专利引入的MRT Profile以及SR TLV格式。

[0089] 图2是根据本发明实施例的MRT Profile选项示意图,如图2所示的MRT Profile,其与RFC7812中定义的default MRT Profile基本相同,区别是MRT Forwarding Mechanism选项为SR-LSP based on multi-SRGB non-tunneling Option。

[0090] 图3是根据本发明实施例的ISIS MT SR-Capabilities Sub-TLV的格式图,如图3所示的ISIS MT SR-Capabilities Sub-TLV格式,与draft-ietf-isis-segment-routing-extensions中给出的SR-Capabilities Sub-TLV基本相同,区别是多了MT-ID字段,以支持针对不同的拓扑通告不同的SRGB。

[0091] 图4是根据本发明实施例的OSPF MT SID/Label Range TLV的格式图,如图4所示的OSPF MT SID/Label Range TLV格式,与draft-ietf-ospf-segment-routing-extensions中给出的SID/Label Range TLV基本相同,区别是多了MT-ID字段,以支持针对不同的拓扑通告不同的SRGB。

[0092] 图5是根据本发明实施例的ISIS Prefix-SID Sub-TLV的格式图,如图5所示的ISIS Prefix-SID Sub-TLV格式,与draft-ietf-isis-segment-routing-extensions中给出的Prefix-SID Sub-TLV基本相同,区别是在Flags中新增一个标志位:Rainbow-Flag:ABR将(MT-default,prefix)的prefix-sid向另一个area通告时,可以设置Rainbow标志。某个节点收到带Rainbow标志的(MT-default,prefix)的prefix-sid通告后,需要判断该prefix相应的(MT-blue,prefix)的MRT-blue下一跳为通告节点且为ABR,则认为(MT-blue,prefix)的MT-blue SR出标签要替换成MT-default SR出标签,否则无需替换。类似的对该prefix相应的(MT-red,prefix)替换出标签。

[0093] 图6是根据本发明实施例的OSPF Prefix SID Sub-TLV的格式图,如图6所示的OSPF Prefix SID Sub-TLV格式,与draft-ietf-ospf-segment-routing-extensions中给出的Prefix SID Sub-TLV基本相同,区别是在Flags中新增Rainbow-Flag标志位,解释同上。

[0094] 具体实施例一

[0095] 本实施例将描述目的prefix处于MRT Island内的MRT路径转发流程,图7是根据本发明具体实施例一的网络拓扑图,如图7所示,网络中运行OSPF,所有节点均处于同一area内,均在相应的OSPF实例下使能分段路由功能以及使能本专利所定义的MRT Profile。S作为源节点建立至目的节点D的prefix(比如D的某个loopback路由)的MRT路径,然后基于此

MRT路径保护SPF主路径。具体的：

[0096] 步骤S702, S、A、B、D各节点上的OSPF实例下使能SR以及本专利所定义的MRT Profile, 则它们在area内形成一个MRT Island。并且由于MRT Profile中指定使用MRT SR-LSP based on multi-SRGB non-tunneling Option作为转发机制, 则各节点除了分配MT-default SRGB以外, 还额外分配MT-red SRGB与MT-blue SRGB, 比如在S节点上我们将这三种SRGB分别记为: MT\_default\_SRGB\_S[ ], MT\_red\_SRGB\_S[ ], MT\_blue\_SRGB\_S[ ]。

[0097] 各节点上将基于SPF算法得到area内的MT-default拓扑, 以及基于MRT算法得到MT-red和MT-blue拓扑。比如S节点上, 至目的节点D的MT-default路径为S-D, MT-red路径也为S-D, 而MT-blue路径为S-A-B-D。

[0098] 各节点上基于拓扑生成相应的prefix表项, 比如S节点上, MT-default拓扑内至目的节点D的loopback0路由的MT-default下一跳为D, 并且选中MT-blue拓扑中给出的MT-blue路径来保护MT-default下一跳D, 那么相应的MRT-FRR下一跳将拷贝为上述MT-blue中的下一跳即A, 称之为MRT-blue下一跳。假设D的loopback路由对应的 (MT-default, prefix-sid) 为SID\_D, 则上述MT-default下一跳D对应的出标签为MT\_default\_SRGB\_D[SID\_D], 上述MRT-blue下一跳A对应的出标签为MT\_blue\_SRGB\_A[SID\_D]。限于篇幅, 我们选择几个有代表性的节点上的表项罗列如下:

[0099] S节点:

[0100] FTN for (MT-default, D-loopback0)

[0101] 主NHLFE: 下一跳为D, 出标签为MT\_default\_SRGB\_D[SID\_D]

[0102] 备NHLFE: 下一跳为A, 出标签为MT\_blue\_SRGB\_A[SID\_D]

[0103] ILM for (MT-default, D-loopback0)

[0104] 入标签为MT\_default\_SRGB\_S[SID\_D]

[0105] 主NHLFE: 下一跳为D, 出标签为MT\_default\_SRGB\_D[SID\_D]

[0106] 备NHLFE: 下一跳为A, 出标签为MT\_blue\_SRGB\_A[SID\_D]

[0107] A节点:

[0108] ILM for (MT-blue, D-loopback0)

[0109] 入标签为MT\_blue\_SRGB\_A[SID\_D]

[0110] NHLFE: 下一跳为B, 出标签为MT\_blue\_SRGB\_B[SID\_D]

[0111] B节点:

[0112] ILM for (MT-blue, D-loopback0)

[0113] 入标签为MT\_blue\_SRGB\_B[SID\_D]

[0114] NHLFE: 下一跳为D, 出标签为MT\_blue\_SRGB\_D[SID\_D]

[0115] D节点:

[0116] ILM for (MT-blue, D-loopback0)

[0117] 入标签为MT\_blue\_SRGB\_D[SID\_D]

[0118] NHLFE: 无。表示SR-LSP已经终结。

[0119] 步骤S704, 对于发往目的地D-loopback0的报文, 当链路S-D出现故障时, S节点将作为MRT ingress节点将流量切换至事先准备好的MRT-blue下一跳A, 即开始将报文沿MT-blue路径S-A-B-D转发。

[0120] 如果S收到的是MT-default SR标签报文,则它基于ILM for (MT-default,D-loopback0) 表项转发,将入标签MT\_default\_SRGB\_S[SID\_D]交换成出标签MT\_blue\_SRGB\_A[SID\_D];如果S收到的是IP报文,则它基于FTN for (MT-default,D-loopback0) 表项转发,直接在IP头上压上出标签MT\_blue\_SRGB\_A[SID\_D]。

[0121] 步骤S706,A节点收到报文后,根据报文的顶层标签为MT\_blue\_SRGB\_A[SID\_D],匹配到ILM for (MT-blue,D-loopback0) 表项,继续将报文的顶层标签MT\_blue\_SRGB\_A[SID\_D]交换成MT\_blue\_SRGB\_B[SID\_D]后发给B。

[0122] 步骤S708,B节点收到报文后,根据报文的顶层标签为MT\_blue\_SRGB\_B[SID\_D],匹配到ILM for (MT-blue,D-loopback0) 表项,继续将报文的顶层标签MT\_blue\_SRGB\_B[SID\_D]交换成MT\_blue\_SRGB\_D[SID\_D]后发给D。

[0123] 步骤S710,D节点收到报文后,根据报文的顶层标签为MT\_blue\_SRGB\_D[SID\_D],匹配到ILM for (MT-blue,D-loopback0) 表项,本地终结SR-LSP,剥掉标签后基于IP头继续转发,由于IP头为D-loopback0,则报文上送控制平面。

[0124] 根据上述实施例,可知报文沿MRT路径转发时,实际上是沿相应MRT拓扑内的SR-LSP转发。

[0125] 具体实施例二

[0126] 本实施例将描述目的prefix处于MRT Island外的MRT路径转发流程,图8是根据本发明具体实施例二的网络拓扑图,如图8所示,网络中运行OSPF,包含两个area,所有节点均在相应的OSPF实例下使能分段路由功能,其中area1中的S、A、B、C使能本发明所定义的MRT Profile.S作为源节点建立至目的节点D的prefix(比如D的某个loopback路由)的MRT路径,然后基于此MRT路径保护SPF主路径。包括如下步骤:

[0127] 步骤S802,area1与area2内所有节点的OSPF实例下均使能SR。各节点分配MT-default SRGB。

[0128] 步骤S804,area1内的S、A、B、C各节点上的OSPF实例下使能本专利所定义的MRT Profile,则它们在area1内形成一个MRT Island。并且由于MRT Profile中指定使用MRT SR-LSP based on multi-SRGB non-tunneling Option作为转发机制,则S、A、B、C节点还额外分配MT-red SRGB与MT-blue SRGB,比如在S节点上我们将这三种SRGB分别记为:MT\_default\_SRGB\_S[ ],MT\_red\_SRGB\_S[ ],MT\_blue\_SRGB\_S[ ]。

[0129] 各节点上将基于SPF算法得到area内的MT-default拓扑,以及基于MRT算法得到MT-red和MT-blue拓扑。比如S节点上,至目的节点B的MT-default路径为S-C-B,MT-red路径也为S-C-B,而MT-blue路径为S-A-B。

[0130] 各节点上基于拓扑生成相应的prefix表项,比如S节点上,MT-default拓扑内至prefix D-loopback0的MT-default下一跳为C(假设在area1内ABR1作为prefix D-loopback0的通告节点,则使用至目的节点ABR1的MT-default路径来确定下一跳)。对于prefix D-loopback0,假设我们使用Named Proxy-node方法(见RFC7812)计算其MRT路径,假设MRT Island中的A作为blue proxy-node attachment router,B作为red proxy-node attachment router,并且S作为GADAG root,且 $S \ll A \ll B \ll C \ll S$ ,则MT-blue拓扑内至prefix D-loopback0的MRT-blue下一跳为A,MT-red拓扑内至prefix D-loopback0的MRT-red下一跳为C。显然这里我们选中MT-blue拓扑中给出的MRT-blue下一跳A来保护MT-default下一

跳C。假设D的loopback路由对应的(MT-default,prefix-sid)为SID\_D,则上述MT-default下一跳C对应的出标签为MT\_default\_SRGB\_C[SID\_D],上述MRT-blue下一跳A对应的出标签为MT\_blue\_SRGB\_A[SID\_D]。同样,我们选择几个有代表性的节点上的表项罗列如下:

[0131] S节点:

[0132] FTN for (MT-default,D-loopback0)

[0133] 主NHLFE:下一跳为C,出标签为MT\_default\_SRGB\_C[SID\_D]

[0134] 备NHLFE:下一跳为A,出标签为MT\_blue\_SRGB\_A[SID\_D]

[0135] ILM for (MT-default,D-loopback0)

[0136] 入标签为MT\_default\_SRGB\_S[SID\_D]

[0137] 主NHLFE:下一跳为C,出标签为MT\_default\_SRGB\_C[SID\_D]

[0138] 备NHLFE:下一跳为A,出标签为MT\_blue\_SRGB\_A[SID\_D]

[0139] A节点:

[0140] ILM for (MT-blue,D-loopback0)

[0141] 入标签为MT\_blue\_SRGB\_A[SID\_D]

[0142] NHLFE:下一跳为LFIN:ABR3,出标签为MT\_default\_SRGB\_ABR3[SID\_D]

[0143] ABR3节点:

[0144] ILM for (MT-default,D-loopback0)

[0145] 入标签为MT\_default\_SRGB\_ABR3[SID\_D]

[0146] NHLFE:下一跳为ABR2,出标签为MT\_default\_SRGB\_ABR2[SID\_D]

[0147] D节点:

[0148] ILM for (MT-default,D-loopback0)

[0149] 入标签为MT\_default\_SRGB\_D[SID\_D]

[0150] NHLFE:无。表示SR-LSP已经终结。

[0151] 步骤S806,对于发往目的地D-loopback0的报文,当链路S-C出现故障时,S节点将作为MRT ingress节点将流量切换至事先准备好的MRT-blue下一跳A,即开始将报文沿MT-blue路径S-A转发。

[0152] 如果S收到的是MT-default SR标签报文,则它基于ILM for (MT-default,D-loopback0)表项转发,将入标签MT\_default\_SRGB\_S[SID\_D]交换成出标签MT\_blue\_SRGB\_A[SID\_D];如果S收到的是IP报文,则它基于FTN for (MT-default,D-loopback0)表项转发,直接在IP头上压上出标签MT\_blue\_SRGB\_A[SID\_D]。

[0153] 步骤S808,A节点收到报文后,根据报文的顶层标签为MT\_blue\_SRGB\_A[SID\_D],匹配到ILM for (MT-blue,D-loopback0)表项,继续将报文的顶层标签MT\_blue\_SRGB\_A[SID\_D]交换成MT\_default\_SRGB\_ABR3[SID\_D]后发给ABR3。

[0154] 说明:这里ABR3作为A的LFIN节点使得A被选中作为prefix D-loopback0的blue proxy-node attachment router,所以按照RFC7812定义的转发规则,报文将从A节点直接向该LFIN节点转发且进入默认拓扑,所以我们给出的NHLFE中,下一跳指向ABR3,且出标签基于MT\_default\_SRGB计算得到的。

[0155] 步骤S810,ABR3节点收到报文后,根据报文的顶层标签为MT\_default\_SRGB\_ABR3[SID\_D],匹配到ILM for (MT-default,D-loopback0)表项,继续将报文的顶层标签MT\_

default\_SRGB\_ABR3[SID\_D]交换成MT\_default\_SRGB\_ABR2[SID\_D]后发给ABR2。

[0156] 步骤S812,ABR2节点收到报文后,根据报文的顶层标签为MT\_default\_SRGB\_ABR2[SID\_D],匹配到ILM for (MT-default,D-loopback0)表项,继续将报文的顶层标签MT\_default\_SRGB\_ABR2[SID\_D]交换成MT\_default\_SRGB\_D[SID\_D]后发给D。

[0157] 步骤S814,D节点收到报文后,根据报文的顶层标签为MT\_default\_SRGB\_D[SID\_D],匹配到ILM for (MT-default,D-loopback0)表项,本地终结SR-LSP,剥掉标签后基于IP头继续转发,由于IP头为D-loopback0,则报文上送控制平面。

[0158] 根据上述实施例,可知报文沿MRT路径转发时,在MRT Island内实际上是沿相应MRT拓扑内的SR-LSP转发,离开MRT Island后,将沿默认拓扑内的SR-LSP转发。符合RFC7812定义的转发规则。

[0159] 具体实施例三

[0160] 本实施例将描述目的prefix处于MRT Island外的MRT路径转发流程,特别是如何基于SR-LSP实现RFC7812定义的rainbow跨域转发规则。图9是根据本发明具体实施例三的网络拓扑图,如图9所示,网络中运行OSPF,包含两个area,所有节点均在相应的OSPF实例下使能分段路由功能,其中area1中的S、A、B、C使能本专利所定义的MRT Profile,area2中的B、E、D、F也同样使能本专利所定义的MRT Profile。S作为源节点建立至目的节点D的prefix(比如D的某个loopback路由)的MRT路径,然后基于此MRT路径保护SPF主路径。包括如下步骤:

[0161] 步骤S902,area1与area2内所有节点的OSPF实例下均使能SR。各节点分配MT-default SRGB。

[0162] 步骤S904,area1内的S、A、B、C各节点上的OSPF实例下使能本专利所定义的MRT Profile,则它们在area1内形成一个MRT Island。并且由于MRT Profile中指定使用MRT SR-LSP based on multi-SRGB non-tunneling Option作为转发机制,则S、A、B、C节点还额外分配MT-red SRGB与MT-blue SRGB,比如在S节点上我们将这三种SRGB分别记为:MT\_default\_SRGB\_S[ ],MT\_red\_SRGB\_S[ ],MT\_blue\_SRGB\_S[ ]。

[0163] 各节点上将基于SPF算法得到area内的MT-default拓扑,以及基于MRT算法得到MT-red和MT-blue拓扑。比如S节点上,至目的节点B的MT-default路径为S-C-B,MT-red路径也为S-C-B,而MT-blue路径为S-A-B。

[0164] 各节点上基于拓扑生成相应的prefix表项,比如S节点上,MT-default拓扑内至prefix D-loopback0的MT-default下一跳为C(假设在area1内ABR1作为prefix D-loopback0的通告节点,则使用至目的节点ABR1的MT-default路径来确定下一跳)。对于prefix D-loopback0,假设MRT Island中的B作为既作为blue proxy-node attachment router也作为red proxy-node attachment router,并且S作为GADAG root,且S<<A<<B<<C<<S,则MT-blue拓扑内至prefix D-loopback0的MRT-blue下一跳为A,MT-red拓扑内至prefix D-loopback0的MRT-red下一跳为C。显然这里我们选中MT-blue拓扑中给出的MRT-blue下一跳A来保护MT-default下一跳C。假设D的loopback路由对应的(MT-default,prefix-sid)为SID\_D,则上述MT-default下一跳C对应的出标签为MT\_default\_SRGB\_C[SID\_D],上述MRT-blue下一跳A对应的出标签为MT\_blue\_SRGB\_A[SID\_D]。

[0165] 步骤S906,同具体实施例一,area2中的B、E、D、F各节点上的OSPF实例下也使能本

专利所定义的MRT Profile,则它们在area2内也形成一个MRT Island,在area2内生成相应的MT-default拓扑,以及MT-red和MT-blue拓扑。在area2内的各节点上计算各拓扑内至prefix D-loopback2的表项与实施例一中描述的步骤是完全类似的。只不过接下来我们将看到,虽然area2内各节点上生成了MRT拓扑内的ILM表项,但是流量从源节点S跨area发往目的节点D时,按照RFC7812定义转发规则,只会命中area2中生成的默认拓扑内的ILM表项。假设B至D在默认拓扑内的最短路径下一跳为F。

[0166] 步骤S908,B节点作为ABR节点,它在将area2内的MT-default prefix D-loopback0向area1内的邻居泛洪时,相应的prefix-sid中将打上rainbow标志;向area2内的邻居泛洪时则不打。

[0167] 步骤S910,B节点所在area1内的邻居A、C收到带rainbow标志的MT-default prefix D-loopback0的prefix-sid通告后,分别检查相应MRT拓扑内prefix D-loopback0的通告节点以及下一跳节点是否为B,如果是的话,则相应的MT-red (或blue) SR出标签其实都要设置成MT-default SR出标签。具体为:

[0168] A节点上,MT-blue拓扑内至prefix D-loopback0的下一跳为B,而prefix D-loopback0的通告节点也为B(因为B是ABR节点),A节点上将相应ILM for (MT-blue,D-loopback0)的NHLFE的出标签设置为MT\_default\_SRGB\_B[SID\_D]。另外A节点上MT-red拓扑内至prefix D-loopback0的下一跳为S,其相应的MT-red SR出标签无需设置成MT-default SR出标签。

[0169] C节点上,MT-red拓扑内至prefix D-loopback0的下一跳为B,而prefix D-loopback0的通告节点也为B(因为B是ABR节点),C节点上将相应ILM for (MT-red,D-loopback0)的NHLFE的出标签设置为MT\_default\_SRGB\_B[SID\_D]。另外A节点上MT-blue拓扑内至prefix D-loopback0的下一跳为S,其相应的MT-blue SR出标签无需设置成MT-default SR出标签。

[0170] 综上,选取几个有代表性的节点上的表项罗列如下:

[0171] S节点:

[0172] FTN for (MT-default,D-loopback0)

[0173] 主NHLFE:下一跳为C,出标签为MT\_default\_SRGB\_C[SID\_D]

[0174] 备NHLFE:下一跳为A,出标签为MT\_blue\_SRGB\_A[SID\_D]

[0175] ILM for (MT-default,D-loopback0)

[0176] 入标签为MT\_default\_SRGB\_S[SID\_D]

[0177] 主NHLFE:下一跳为C,出标签为MT\_default\_SRGB\_C[SID\_D]

[0178] 备NHLFE:下一跳为A,出标签为MT\_blue\_SRGB\_A[SID\_D]

[0179] A节点:

[0180] ILM for (MT-blue,D-loopback0)

[0181] 入标签为MT\_blue\_SRGB\_A[SID\_D]

[0182] NHLFE:下一跳为B,出标签为MT\_default\_SRGB\_B[SID\_D]

[0183] B节点:

[0184] ILM for (MT-default,D-loopback0)

[0185] 入标签为MT\_default\_SRGB\_B[SID\_D]

- [0186] NHLFE:下一跳为F,出标签为MT\_default\_SRGB\_F[SID\_D]
- [0187] F节点:
- [0188] ILM for (MT-default,D-loopback0)
- [0189] 入标签为MT\_default\_SRGB\_F[SID\_D]
- [0190] NHLFE:下一跳为D,出标签为MT\_default\_SRGB\_D[SID\_D]
- [0191] D节点:
- [0192] ILM for (MT-default,D-loopback0)
- [0193] 入标签为MT\_default\_SRGB\_D[SID\_D]
- [0194] NHLFE:无。表示SR-LSP已经终结。
- [0195] 步骤S912,对于发往目的地D-loopback0的报文,当链路S-C出现故障时,S节点将作为MRT ingress节点将流量切换至事先准备好的MRT-blue下一跳A,即开始将报文沿MT-blue路径S-A-B转发。
- [0196] 如果S收到的是MT-default SR标签报文,则它基于ILM for (MT-default,D-loopback0)表项转发,将入标签MT\_default\_SRGB\_S[SID\_D]交换成出标签MT\_blue\_SRGB\_A[SID\_D];如果S收到的是IP报文,则它基于FTN for (MT-default,D-loopback0)表项转发,直接在IP头上压上出标签MT\_blue\_SRGB\_A[SID\_D]。
- [0197] 步骤S914,A节点收到报文后,根据报文的顶层标签为MT\_blue\_SRGB\_A[SID\_D],匹配到ILM for (MT-blue,D-loopback0)表项,继续将报文的顶层标签MT\_blue\_SRGB\_A[SID\_D]交换成MT\_default\_SRGB\_B[SID\_D]后发给B。
- [0198] 步骤S916,B节点收到报文后,根据报文的顶层标签为MT\_default\_SRGB\_B[SID\_D],匹配到ILM for (MT-blue,D-loopback0)表项,继续将报文的顶层标签MT\_default\_SRGB\_B[SID\_D]交换成MT\_default\_SRGB\_F[SID\_D]后发给F。
- [0199] 步骤S918,F节点收到报文后,根据报文的顶层标签为MT\_default\_SRGB\_F[SID\_D],匹配到ILM for (MT-blue,D-loopback0)表项,继续将报文的顶层标签MT\_default\_SRGB\_F[SID\_D]交换成MT\_default\_SRGB\_D[SID\_D]后发给D。
- [0200] 步骤S920,D节点收到报文后,根据报文的顶层标签为MT\_default\_SRGB\_D[SID\_D],匹配到ILM for (MT-default,D-loopback0)表项,本地终结SR-LSP,剥掉标签后基于IP头继续转发,由于IP头为D-loopback0,则报文中送控制平面。
- [0201] 根据上述实施例,可知报文沿MRT路径转发时,在area1中的MRT Island内实际上是沿相应MRT拓扑内的SR-LSP转发,而离开area1进入area2后,将沿默认拓扑内的SR-LSP转发。符合RFC7812定义的转发规则。
- [0202] 通过以上的实施方式描述,本领域的技术人员可以清楚地了解到根据上述实施例的方法可借助软件加必需的通用硬件平台的方式来实现,当然也可以通过硬件,但很多情况下前者是更佳的实施方式。基于这样的理解,本发明的技术方案本质上或者说对现有技术做出贡献的部分可以以软件产品的形式体现出来,该计算机软件产品存储在一个存储介质(如ROM/RAM、磁碟、光盘)中,包括若干指令用以使得一台终端设备(可以是手机,计算机,服务器,或者网络设备等)执行本发明各个实施例所述的方法。
- [0203] 在本实施例中还提供了一种报文转发装置,该装置用于实现上述实施例及优选实施方式,已经进行过说明的不再赘述。如以下所使用的,术语“模块”可以实现预定功能的软

件和/或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0204] 图10是根据本发明实施例的报文转发装置的结构框图,如图10所示,该装置可以应用于第一节点中,该第一节点可以是MRT Island内的任何节点,该装置包括接收模块102、查找模块104和转发模块106,下面对该装置进行说明:

[0205] 接收模块102,用于接收待转发的报文,其中,该报文的目的地为第二节点;查找模块104,连接至上述接收模块102,用于在预先生成的拓扑中查找与上述报文对应的目标拓扑,其中,该预先生成的拓扑包括根据最大冗余树MRT算法生成得到的第一拓扑和第二拓扑,根据最短路径优先SPF算法得到的第三拓扑,该第一拓扑和所述第二拓扑以及第三拓扑彼此各不相同;转发模块106,连接至上述查找模块104,用于在上述目标拓扑中查找用于转发到第二节点的下一跳节点,并将该报文转发到下一跳节点。在本实施例中,在将报文转发到下一跳节点时,可以基于预定转发机制进行转发,该预定转发机制为基于每拓扑每分段路由由全局块SRGB的非隧道嵌套方式的分段路由转发机制。

[0206] 在一个可选的实施例中,上述装置还包括第一处理模块,用于按照上述MRT算法从第一拓扑中至目的地址的路由和第二拓扑中至所述目的地址的路由中确定用于保护上述第三拓扑中至目的地址的路由的保护路由,并确定该保护路由对应的拓扑为保护拓扑。

[0207] 在一个可选的实施例中,上述查找模块104可以通过如下方式在预先生成的拓扑中查找与报文对应的目标拓扑:判断第三拓扑中的用于到达上述第二节点的链路是否出现故障;在判断没有出现故障的情况下,确定上述第三拓扑为目标拓扑;和/或,在判断出现故障的情况下,确定上述保护拓扑为目标拓扑。

[0208] 在一个可选的实施例中,上述装置还包括第二处理模块,用于在接收待转发的上述报文之前,还包括:根据上述MRT算法生成得到第一拓扑和第二拓扑,以及根据上述SPF算法生成得到第三拓扑。

[0209] 在一个可选的实施例中,上述第二处理模块可以通过如下方式根据MRT算法生成得到第一拓扑和第二拓扑,以及根据SPF算法生成得到第三拓扑:确定第一节点所在的MRT Island,其中,该MRT Island是通过在第一节点以及与第一节点处于同一域area或同一层次level的其他节点上的开放最短路径优先OSPF或者中间系统到中间系统ISIS实例下使能分段路由SR以及最大冗余树配置文件MRT profile后在所述第一节点所在的area或level内由第一节点和其他节点相互协商形成的;基于MRT Island运行MRT算法生成第一拓扑和第二拓扑,以及,基于area或level运行SPF算法生成第三拓扑。

[0210] 在一个可选的实施例中,上述MRT profile中指定采用上述预定转发机制,即,采用基于每拓扑每分段路由全局块SRGB的非隧道嵌套方式的分段路由转发机制。

[0211] 在一个可选的实施例中,上述装置还包括第三处理模块,用于执行以下处理至少之一:为第一拓扑分配第一分段路由全局块SRGB,并将该第一SRGB在上述MRT Island所在的域area或层次level内泛洪;为第二拓扑分配第二分段路由全局块SRGB,并将该第二SRGB在上述MRT Island所在的域area或层次level内泛洪;为第三拓扑分配第三分段路由全局块SRGB,并将该第三SRGB在第一节点所在的所有域area或层次level内泛洪;接收其他节点上的上述预先生成的拓扑的分段路由全局块SRGB,记录其他节点上的预先生成的拓扑的SRGB以及将该其他节点上的预先生成的拓扑的SRGB继续通告给除其他节点之外的节点。

[0212] 在一个可选的实施例中,上述转发模块106可以通过如下方式将上述报文转发到下一跳节点:确定上述第一节点的出标签;将上述第一节点的出标签封装到上述报文上,并将封装后的报文发送到下一跳节点。

[0213] 在一个可选的实施例中,上述转发模块106可以通过如下方式确定第一节点的出标签:判断下一跳节点是否为第一节点所在的MRT Island外的节点;在判断结果为否时,确定该第一节点的对应目标拓扑的出标签,其中,该第一节点的对应所述目标拓扑的出标签由所述第一节点基于第二节点的所述第三拓扑的前缀段索引prefix-sid与所述下一跳节点的对应所述目标拓扑的SRGB计算得到的;和/或,在判断结果为是时,确定该第一节点的对应目标拓扑的出标签,其中,该第一节点的对应目标拓扑的出标签由第一节点基于第二节点的第三拓扑的前缀段索引prefix-sid与下一跳节点的对应第三拓扑的SRGB计算得到的。

[0214] 在一个可选的实施例中,上述转发模块106可以通过如下方式确定第一节点的出标签:判断上述下一跳节点是否为需要将报文回归到第三拓扑的区域边界路由器ABR或层次边界路由器LBR;在判断结果为否时,确定上述第一节点的对应所述目标拓扑的出标签,其中,该第一节点的对应目标拓扑的出标签由第一节点基于第二节点上的第三拓扑的前缀段索引prefix-sid与下一跳节点的对应目标拓扑的SRGB计算得到的;和/或,在判断结果为是时,确定第一节点的对应所述目标拓扑的出标签,其中,该第一节点的对应目标拓扑的出标签由第一节点基于第二节点上的第三拓扑的前缀段索引prefix-sid与下一跳节点的对应第三拓扑的SRGB计算得到的。

[0215] 在一个可选的实施例中,上述转发模块106可以通过如下方式判断下一跳节点是否为需要将报文回归到第三拓扑的ABR或LBR包括:判断从上述下一跳节点接收的预先生成的第三拓扑的目的前缀prefix对应的prefix-sid中是否携带有预定标志,其中,该预定标志表示目的prefix是被ABR或LBR从area或level外泄露到本area或level的,并且由ABR或LBR打上所述预定标志,以使得至目的prefix的报文在离开ABR或LBR后回归到所述第三拓扑;在判断结果为从下一跳节点接收的预先生成的第三拓扑的目的prefix对应的prefix-sid中携带有上述预定标志,并且该目的prefix在第三拓扑内的通告节点为上述下一跳节点时,确定上述下一跳节点为需要将报文回归到第三拓扑的ABR或LBR;在判断结果为从下一跳节点接收的预先生成的第三拓扑的目的prefix对应的prefix-sid中未携带有上述预定标志,或者在判断结果为携带有上述预定标志但是该目的prefix在第三拓扑内的通告节点不为上述下一跳节点时,确定下一跳节点不为需要将报文回归到第三拓扑的ABR或LBR。

[0216] 在一个可选的实施例中,当上述第一节点为ABR或LBR时,上述第三处理模块还用于执行如下处理:将目的prefix向与目的prefix最短路径下一跳节点所处area或level不同的area或level内的邻居通告时,在上述目的prefix相应的prefix-sid中携带上述预定标志;将目的prefix向与目的prefix最短路径下一跳节点所处area或level相同的area或level内的邻居通告时,在上述目的prefix相应的prefix-sid中不携带上述预定标志。

[0217] 在一个可选的实施例中,上述转发模块106可以通过如下方式确定第一节点的出标签:当第一节点和第二节点为同一节点时,则该第一节点基于目标拓扑确定的下一跳节点为第一节点,该第一节点没有出标签。

[0218] 在一个可选的实施例中,上述转发模块106可以通过如下方式至少之一将第一节

点的出标签封装到报文上:当上述报文的报文类型为互联网协议IP报文时,在该IP报文的IP头上压上上述第一节点的出标签;当上述报文的报文类型为分段路由SR标签报文时,将上述SR标签报文的标签栈的栈顶标签替换成第一节点的出标签。

[0219] 在一个可选的实施例中,上述装置还包括第三处理模块,用于当上述第一节点和第二节点为同一节点时,当上述报文的报文类型为互联网协议IP报文时,将上述报文上送至第一节点的控制平面;和/或,当上述报文的报文类型为分段路由SR标签报文时,将上述SR标签报文的标签栈的栈顶标签弹出,并继续基于上述报文的标签栈的下层标签或IP头查表转发。

[0220] 需要说明的是,上述各个模块是可以通过软件或硬件来实现的,对于后者,可以通过以下方式实现,但不限于此:上述模块均位于同一处理器中;或者,上述各个模块以任意组合的形式分别位于不同的处理器中。

[0221] 本发明的实施例还提供了一种存储介质。可选地,在本实施例中,上述存储介质可以被设置为存储用于执行上述步骤的程序代码。

[0222] 可选地,在本实施例中,上述存储介质可以包括但不限于:U盘、只读存储器(Read-Only Memory,简称为ROM)、随机存取存储器(Random Access Memory,简称为RAM)、移动硬盘、磁碟或者光盘等各种可以存储程序代码的介质。

[0223] 可选地,在本实施例中,处理器根据存储介质中已存储的程序代码执行上述各步骤。

[0224] 可选地,本实施例中的具体示例可以参考上述实施例及可选实施方式中所描述的示例,本实施例在此不再赘述。

[0225] 采用本发明实施例中的方法,与相关技术相比,填补了分段路由与MRT技术结合的缺口,为未来网络的演进提供了有价值的探索。

[0226] 显然,本领域的技术人员应该明白,上述的本发明的各模块或各步骤可以用通用的计算装置来实现,它们可以集中在单个的计算装置上,或者分布在多个计算装置所组成的网络上,可选地,它们可以用计算装置可执行的程序代码来实现,从而,可以将它们存储在存储装置中由计算装置来执行,并且在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤,或者将它们分别制作成各个集成电路模块,或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样,本发明不限制于任何特定的硬件和软件结合。

[0227] 以上所述仅为本发明的优选实施例而已,并不用于限制本发明,对于本领域的技术人员来说,本发明可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

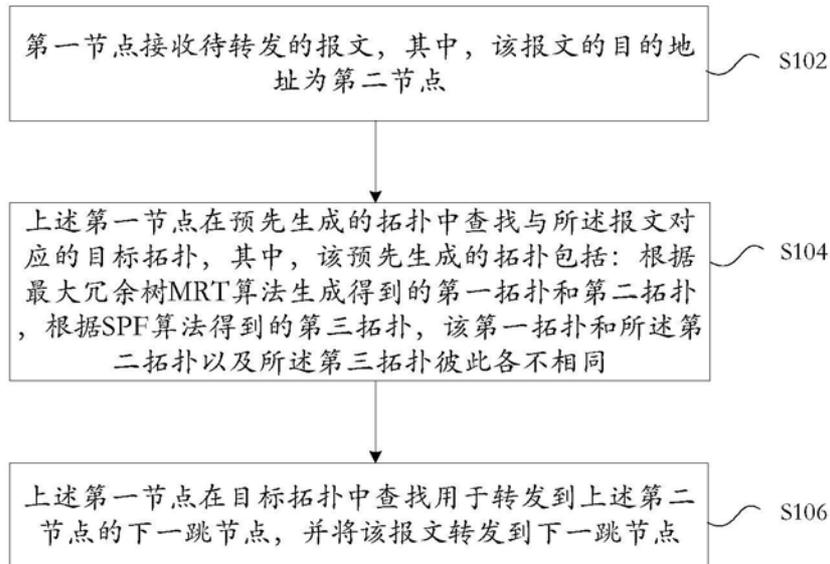


图1

MRT Algorithm	MRT Lowpoint algorithm
MRT-Red MT-ID	从 Multi-Topology Identifiers Registry 分配，待 IANA 分配
MRT-Blue MT-ID	从 Multi-Topology Identifiers Registry 分配，待 IANA 分配
GADAG Root Selection Policy	各路由器通告的 GADAG Root Selection Priority 中选择最小的
MRT Forwarding Mechanism	MRT SR-LSP based on multi-SRGB non-tunneling Option
Recalculation	需等待切换后的 SPT 收敛结束且所有路由器都把流量重新切换到 SPT，才开始计算新的 MRT。
Area/Level Border Behavior	ABR/LBR 需要保证流量在离开 area 时也离开了 MRT-Red 或 MRT-Blue 转发拓扑。

图2

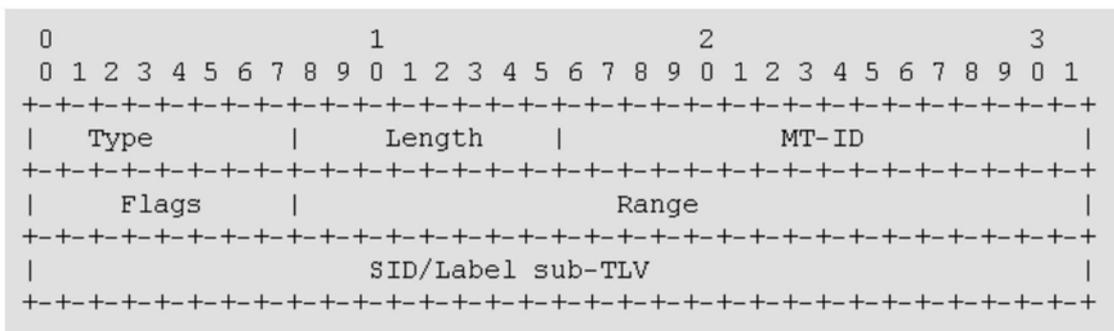


图3

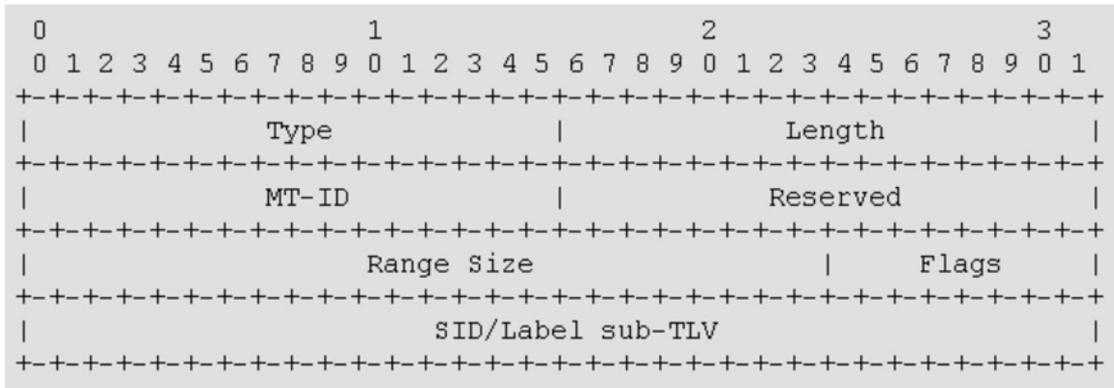


图4

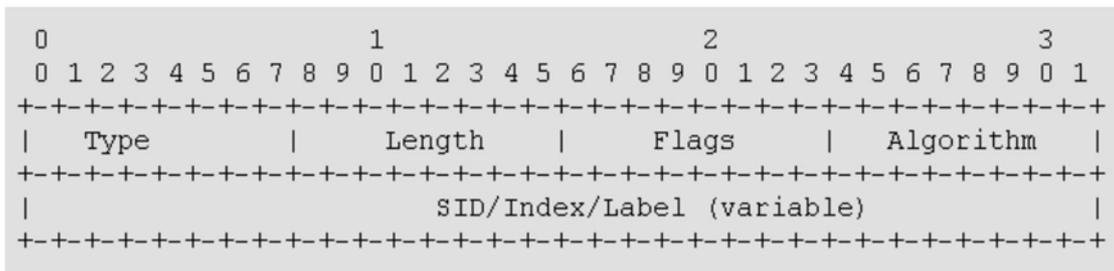


图5

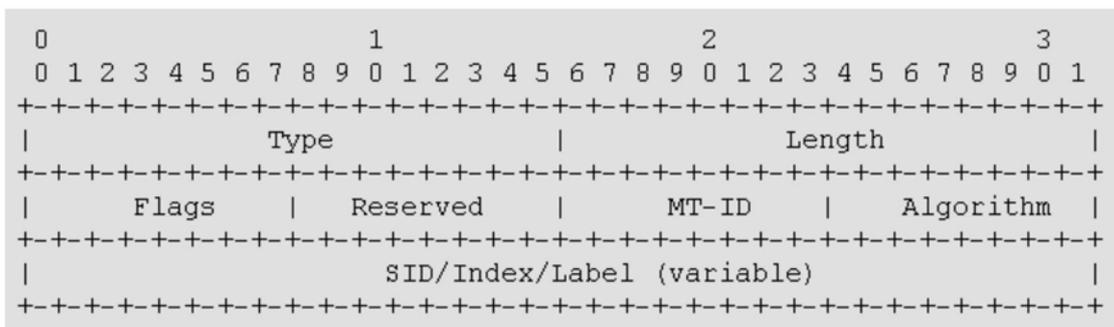


图6

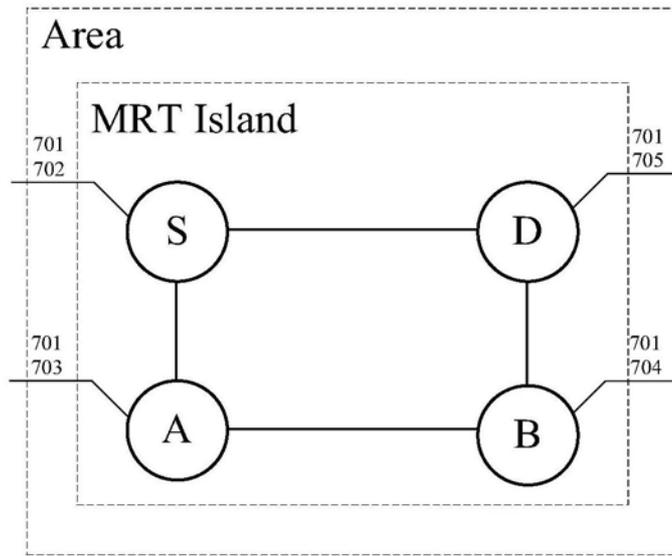


图7

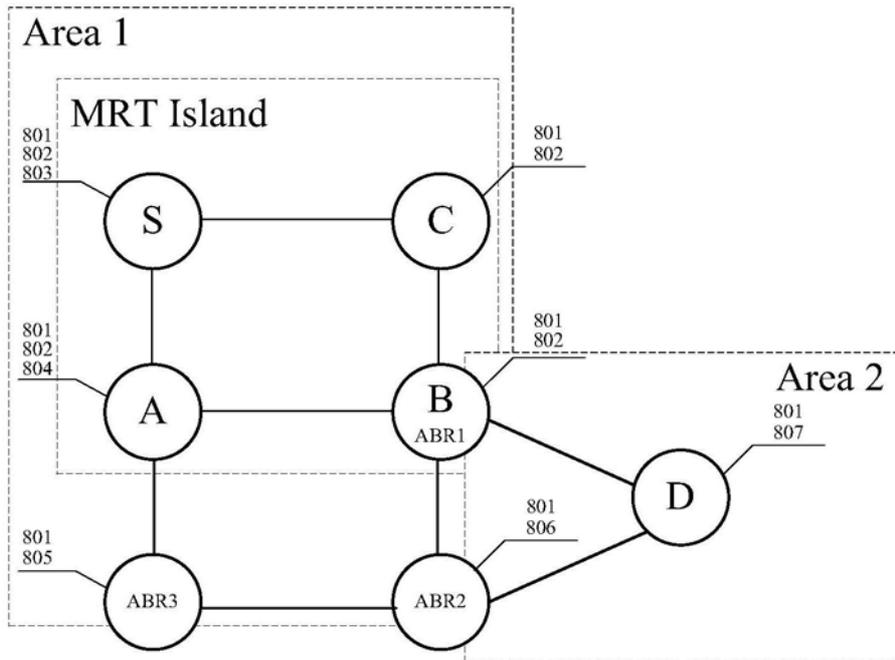


图8

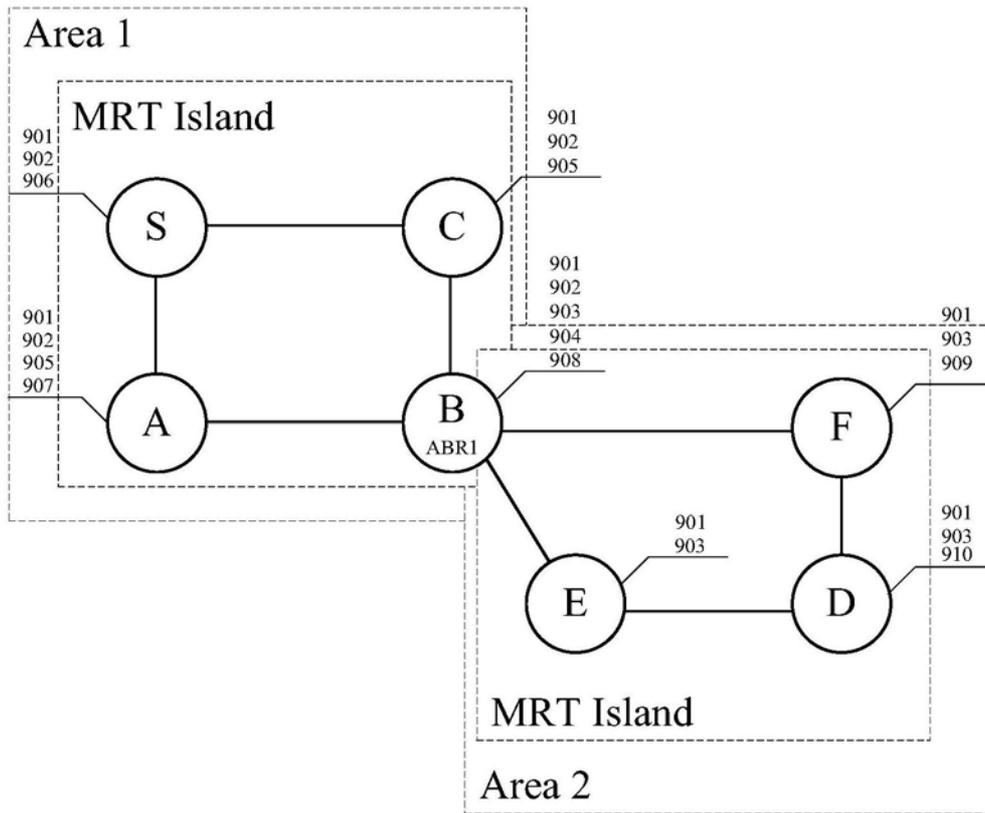


图9



图10