

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(43) 国际公布日
2015年4月23日 (23.04.2015)

(10) 国际公布号
WO 2015/054902 A1

- (51) 国际专利分类号:
H04L 12/28 (2006.01)
- (21) 国际申请号: PCT/CN2013/085496
- (22) 国际申请日: 2013年10月18日 (18.10.2013)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 李振斌 (LI, Zhenbin); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR,

CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。

- (54) Title: METHOD, CONTROLLER, FORWARDING DEVICE, AND NETWORK SYSTEM FOR FORWARDING PACKETS
- (54) 发明名称: 转发报文的方法、控制器、转发设备和网络系统

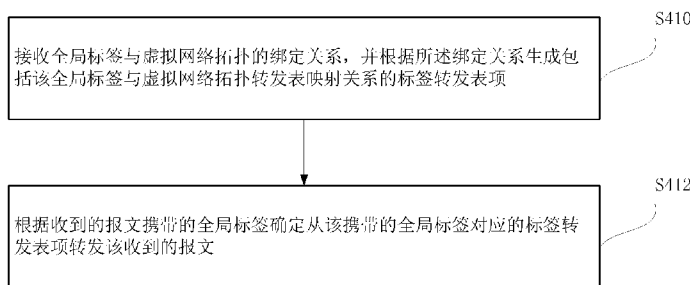


图4 / FIG. 4

S410 GENERATION, ON THE BASIS OF THE LINK BINDING RELATIONS, OF THE LABEL FORWARDING ENTRY COMPRISING THE MAPPING RELATIONS OF THE LINK GLOBAL LABELS AND THE VIRTUAL LINK FORWARDING INFORMATION

S412 DETERMINATION, ON THE BASIS OF THE GLOBAL LABEL CARRIED BY THE RECEIVED PACKET, OF FORWARDING OF THE RECEIVED PACKET FROM THE LABEL FORWARDING ENTRY CORRESPONDING TO THE GLOBAL LABEL CARRIED

(57) Abstract: A network system comprising at least one controller and multiple forwarding devices. The forwarding devices compose multiple virtual network topologies. The controller is used for assigning global labels for each virtual network topology in the multiple virtual network topologies and for forwarding binding relations of the global labels and the virtual network topologies to the forwarding devices, where the global labels assigned by the controller to the different virtual network topologies in the virtual network topologies are different. The forwarding devices are used for receiving the binding relations and for generating, on the basis of the binding relations, a label forwarding entry comprising mapping relations of the global labels and the virtual network topologies. When a packet received by the forwarding devices carries a global label, the forwarding devices look up, on the basis of the global label, a forwarding table of the virtual network topology corresponding to the global label and determines to forward the packet via the forwarding table of the virtual network topology.

(57) 摘要:

[见续页]



WO 2015/054902 A1

一种网络系统，包括至少一个控制器和多个转发设备，该转发设备组成多个虚拟网络拓扑，该控制器，用于为该多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，并将该全局标签与该虚拟网络拓扑的绑定关系发送给该转发设备；其中该控制器分配给该虚拟网络拓扑中的不同虚拟网络拓扑的全局标签不同；该转发设备，用于接收该绑定关系，根据该绑定关系生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；当该转发设备收到的报文中携带有全局标签时，该转发设备根据该全局标签查找到该全局标签对应的虚拟网络拓扑的转发表，确定通过该虚拟网络拓扑的转发表转发该报文。

转发报文的方法、控制器、转发设备和网络系统

技术领域

本发明涉及一种转发报文的方法、控制器、转发设备和系统。

5

背景技术

最大冗余树 (Maximally Redundant Trees, MRT) 的原理: 在基于最短路径优先 (Shortest Path First, SPF) 拓扑进行逐跳 (hop-by-hop) 转发的网络中, 如图 1A 所示, 假定某个网络中包括节点 A、B、C、D、E、F、G、H、I、J、R 等, 针对某个公共的根 Root 节点 (比如图 1A 中的节点 H), 可以计算出两个不相交的冗余树 (如图 1B 中的拓扑和图 1C 中的拓扑), 图 1B 中的冗余树和图 1C 中的冗余树都包含所有节点 A、B、C、D、E、F、G、H、I、J、R, 且具有相同的 Root 节点 H, 但是有向链路 link 不重合。

以节点 A、B、C、D 组成的网络为例, 通过 MRT 可以计算出三个拓扑:

15 缺省拓扑 (参考图 2A): 节点 A->节点 B->节点 C 和节点 A->节点 D->节点 C;

第一拓扑 (参考图 2B): 节点 A->节点 B->节点 C->节点 D->节点 A;

第二拓扑 (参考图 2C): 节点 A->节点 D->节点 C->节点 B->节点 A。

对于缺省拓扑中的路由, 可以通过第一拓扑或第二拓扑中的备份路由来保护。如图 2D 中缺省拓扑中的路由被 MRT 计算得到的图 2C 所示的第二拓扑保护, 当从节点 B 到节点 C 的有向链路发生故障的时候, 流量会切换到第二拓扑中的节点 B 到节点 A 的有向链路, 再经由节点 D 到达目的节点 C。

20 在 MRT FRR (MRT Fast ReRoute, MRT 快速重路由) 中有两种转发机制, 一种是通过标签分发协议 (Label Distribution Protocol, LDP) 多拓扑转发, 也就是在第一拓扑和第二拓扑中为路由同样分发标签, 通过多协议标签交换 (Multi-Protocol Label Switching, MPLS) 标签实现在不同的拓扑中

转发；一种是通过网络协议（Internet Protocol, IP）隧道转发，这种方法不使用 MPLS 技术，通过在第一拓扑或第二拓扑中的 IP 隧道到达目的地址。

发明人在研究过程中发现，MRT FRR 中采用的转发机制存在如下问题：

通过 IP 隧道转发需要预先进行多拓扑地址的规划，并且在网络中建立
5 和维护大量的 IP 隧道，使得网络的管理和维护变得复杂。

使用 LDP MT（LDP Multi-Topology, LDP 多拓扑）来支持 MRT FRR 转发也存在挑战：1. 基于 LDP MT 支持 MRT FRR 对于同一个前缀，需要在三个拓扑中分别分配标签，需要使用大量的标签。对于纯 IP 网络，需要增量部署 LDP MT 来支持 MRT FRR。Internet 骨干网的路由前缀数量可以
10 达到 40 - 50 万的规模，如果完全采用 LDP MT，标签的使用数量会超过 2^{20} ，也就是说现有的标签空间（使用 20bit 来表示标签）可能无法满足应用的需求。2. 对于纯 IP 网络，为了网络的运维简单，运营商不希望再引入新的 MPLS 协议（如 LDP），这样也会使得基于 LDP MT 的 MRT FRR 难以得到实际的部署。

15

发明内容

本发明实施例的目的在于提供一种转发报文的方法、设备和网络系统，以改善 MRT FRR 的维护效率。

为达到上述目的，本发明的实施例提供了如下技术方案。

20 第一方面，本发明实施例提供一种网络系统，包括至少一个控制器和多个转发设备，所述多个转发设备组成多个虚拟网络拓扑，

所述至少一个控制器，用于为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，并将所述全局标签与所述多个虚拟网络拓扑的绑定关系发送给所述多个转发设备；其中所述控制器分配给所述多个虚拟网络
25 拓扑中的不同虚拟网络拓扑的全局标签不同；

所述多个转发设备，用于接收所述全局标签与虚拟网络拓扑的绑定关系，根据该绑定关系生成包括该全局标签与虚拟网络拓扑转发表映射关系

的标签转发表项；当该转发设备收到的第一报文中携带有第一全局标签时，该转发设备根据该第一全局标签查找到该第一全局标签对应的第一虚拟网络拓扑的转发表，确定通过该第一虚拟网络拓扑的转发表转发该第一报文；

所述第一虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

5

在第一方面的第一种可能的实施方式中，所述多个转发设备还用于：

当所述转发设备收到的第二报文中没有携带全局标签，并且该第二报文将要从第二虚拟网络拓扑转发，则该转发设备根据该第二虚拟网络拓扑对应的第二全局标签，为该第二报文增加该第二全局标签的封装，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；；所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

结合第一方面或第一方面的第一种可能的实施方式，在第一方面的第二种可能的实施方式中，所述控制器包括标签分配模块和标签发送模块，其中

15 所述标签分配模块用于：为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，为所述多个虚拟网络拓扑中的不同虚拟网络拓扑分配的标签不同；

所述标签发送模块用于：将所述全局标签与该全局标签对应的虚拟网络拓扑的绑定关系发送给所述多个转发设备。

20 结合第一方面、第一方面的第一种或第二种可能的实施方式，在第一方面的第三种可能的实施方式中，所述转发设备包括接收模块、标签转发表项模块和转发模块，其中

所述接收模块，用于接收所述控制器发送的所述全局标签与该全局标签对应的虚拟网络拓扑的绑定关系；

25 所述标签转发表项生成模块，根据所述全局标签与虚拟网络拓扑的绑

定关系，生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；

所述转发模块，用于根据收到的第一报文携带的第一全局标签确定该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该第一报文。

结合第一方面的第三种可能的实施方式，在第一方面的第四种可能的实施方式中，所述转发设备还包括标签增加模块，

当所述接收模块收到的第二报文中没有携带全局标签，并且该第二报文将要从第二虚拟网络拓扑转发，则所述转发模块根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，为该第二报文增加该第二全局标签，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

结合第一方面或第一方面的第一种可能的实施方式，在第一方面的第五种可能的实施方式中，所述控制器包括第一处理器和第一接口，其中

所述第一处理器用于：为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，为所述多个虚拟网络拓扑中的不同虚拟网络拓扑分配的全局标签不同；

所述第一接口用于：将所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系发送给所述多个转发设备。

结合第一方面、第一方面的第一种或第五种可能的实施方式，在第一方面的第六种可能的实施方式中，所述转发设备包括第二接口、第二处理器和第三处理器，其中

所述第二接口，用于接收所述控制器发送的所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系；

所述第二处理器，用于根据所述全局标签与所述虚拟网络拓扑的绑定

关系，生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；

所述第三处理器，用于根据收到的第一报文中携带的第一全局标签确定该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该第一报文。

结合第一方面的第六种可能的实施方式，在第一方面的第七种可能的实施方式中，该转发设备还包括第四处理器；

如果该第二接口收到的第二报文中没有携带全局标签，并且该第二报文将要从第二虚拟网络拓扑转发，则该第三处理器根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，由该第四处理器为该第二报文增加该第二全局标签该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

结合第一方面或第一方面的第一种或第二种或第三种或第四种或第五种或第六种或第七种可能的实施方式，在第一方面的第八种可能的实施方式中，所述控制器为内部网关协议 IGP 服务器。

结合第一方面或第一方面的第一种或第二种或第三种或第四种或第五种或第六种或第七种或第八种可能的实施方式，在第一方面的第九种可能的实施方式中，所述全局标签包括两个或两个以上的标签。

20

第二方面，本发明实施例提供一种控制器，该控制器应用于包括该控制器和多个转发设备的网络系统中，所述多个转发设备组成多个虚拟网络拓扑，

该控制器包括标签分配模块和标签发送模块，其中

所述标签分配模块用于：为所述多个虚拟网络拓扑中的每个虚拟网络

拓扑分配全局标签，为所述多个虚拟网络拓扑中的不同虚拟网络拓扑分配的标签不同；

所述标签发送模块用于：将所述全局标签与该全局标签对应的虚拟网络拓扑的绑定关系发送给所述多个转发设备。

5 在第二方面的第一种可能的实施方式中，所述全局标签包括两个或两个以上的标签。

第三方面，本发明实施例提供一种控制器，该控制器应用于包括该控制器和多个转发设备的网络系统中，所述多个转发设备组成多个虚拟网络
10 拓扑，

所述控制器包括第一处理器和第一接口，其中

所述第一处理器用于：为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，为所述多个虚拟网络拓扑中的不同虚拟网络拓扑分配的全局标签不同；

15 所述第一接口用于：将所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系发送给所述多个转发设备。

在第三方面的第一种可能的实施方式中，所述全局标签包括两个或两个以上的标签。

20 第四方面，本发明实施例提供一种转发设备，该转发设备应用于包括多个该转发设备和至少一个控制器的网络系统中，所述多个转发设备组成多个虚拟网络拓扑，

所述转发设备包括接收模块、标签转发表项模块和转发模块，其中

所述接收模块，用于接收所述控制器发送的所述全局标签与该全局标

签对应的虚拟网络拓扑的绑定关系；

所述标签转发表项生成模块，根据所述全局标签与虚拟网络拓扑的绑定关系，生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；

- 5 所述转发模块，用于根据收到的第一报文携带的第一全局标签确定该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该第一报文。

在第四方面的第一种可能的实施方式中，所述转发设备还包括标签增加模块，

- 10 当所述接收模块收到的第二报文中没有携带全局标签，并且该第二报文将要从第二虚拟网络拓扑转发，则所述转发模块根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，为该第二报文增加该第二全局标签，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

- 15 结合第四方面、第四方面的第一种可能的实施方式，在第四方面的第二种可能的实施方式中，所述全局标签包括两个或两个以上的标签。

- 20 第五方面，本发明实施例提供一种转发设备，该转发设备应用于包括多个该转发设备和至少一个控制器的网络系统中，所述多个转发设备组成多个虚拟网络拓扑，

所述转发设备包括第二接口、第二处理器和第三处理器，其中

所述第二接口，用于接收所述控制器发送的所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系；

- 25 所述第二处理器，用于根据所述全局标签与所述虚拟网络拓扑的绑定关系，生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表

项；

所述第三处理器，用于根据收到的第一报文中携带的第一全局标签确定该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该第一报文。

5 在第五方面的第一种可能的实施方式中，该转发设备还包括第四处理器；

如果该第二接口收到的第二报文中没有携带全局标签，并且该第二报文将要从第二虚拟网络拓扑转发，则该第三处理器根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，由该第四处理器为该第二报文增加该第
10 二全局标签，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

结合第五方面、第五方面的第一种可能的实施方式，在第五方面的第
二种可能的实施方式中，所述全局标签包括两个或两个以上的标签。

15

第六方面，本发明实施例提供一种转发报文的方法，该方法应用于包括至少一个控制器和多个转发设备的网络系统，所述多个转发设备组成多个虚拟网络拓扑，

接收全局标签与所述虚拟网络拓扑的绑定关系，根据该绑定关系生成
20 包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；所述绑定关系是由所述控制器为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，并将所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系发送给所述多个转发设备的，所述控制器分配给所述虚拟网络拓扑中的不同虚拟网络拓扑的全局标签不同；

25 当该转发设备收到的第一报文中携带有第一全局标签时，该转发设备

根据该第一全局标签查找到该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的第一标签转发表项转发该第一报文；

所述第一虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

- 5 在第六方面的第一种可能的实施方式中，所述至少一个控制器转发所述绑定关系给所述转发设备，具体为：所述至少一个控制器通过内部网关协议IGP扩展发送所述绑定关系给IGP域内的转发设备。

结合第六方面，在第六方面的第二种可能的实施方式中，如果所述转发设备收到的第二报文中没有携带全局标签，并且该第二报文要从第二虚拟网络拓扑转发，则该转发设备根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，为该第二报文增加该第二全局标签，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

结合第六方面、第六方面的第一种可能的实施方式，在第六方面的第
15 二种可能的实施方式中，所述全局标签包括两个或两个以上的标签。

本发明实施例的控制器、转发设备、转发报文的方法，通过 IGP 扩展为 IGP 多拓扑分配 MPLS 全局标签实现 MPLS 网络虚拟化，该方法可以用于 MRT FRR 的转发，达到如下有益效果：通过使用全局标签实现 IP 网络的 MRT FRR 转发，只需要扩展 IGP 即可，不需要部署 LDP。通过使用全
20 局标签，只需要使用标识各个虚拟网络拓扑的少量全局标签就可以实现，简化了网络虚拟化及 MRT FRR 的部署。

附图说明

为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对
25 实施例或现有技术描述中所需要使用的附图作一简单地介绍，显而易见地，下面描述中的附图是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。

图 1A 为一种现有技术的网络的缺省拓扑图；

图 1B 为图 1A 所示的网络的一种有向拓扑图；

图 1C 为图 1A 所示的网络的另一种有向拓扑图；

图 1D 为图 1A 所示的网络中的缺省拓扑中链路故障切换示意图；

5 图 2A 为一种现有技术的网络的缺省拓扑图；

图 2B 为图 2A 所示的网络的一种有向拓扑图；

图 2C 为图 2A 所示的网络的另一种有向拓扑图；

图 2D 为图 2A 所示的网络中的缺省拓扑中链路故障切换示意图；

图 3 为本发明实施例的一种网络系统的结构示意图；

10 图 4 为本发明实施例的一种转发报文的方法的流程示意图；

图 5 为为本发明实施例的另一种网络系统的结构示意图；

图 6 为本发明实施例的一种转发报文的方法的流程示意图；

图 7A 为本发明实施例的一种 TLV 的格式的示意图；

图 7B 为本发明实施例的一种标签映射 TLV 的格式的示意图；

15 图 7C 为本发明实施例的一种 MPLS 多拓扑子 TLV 的格式示意图；

图 7D 为本发明实施例的一种 Global Label Sub-TLV 的格式示意图；

图 8 为本发明实施例的虚拟网络拓扑转发表的组成示意图；

图 9A 为一种现有技术的网络的缺省拓扑图；

图 9B 为图 9A 所示的网络的一种有向拓扑图；

20 图 9C 为图 9A 所示的网络的另一种有向拓扑图；

图 9D 为图 9A 所示的网络中的缺省拓扑中链路故障切换示意图；

图 10 为本发明实施例的另一种报文转发方法的流程示意图；

图 11 为本发明实施例的一种 MPLS 转发表项的组成示意图；

图 12 为本发明实施例的另一种网络系统的结构示意图；

5 图 13 为本发明实施例的又一种网络系统的结构示意图。

具体实施方式

为使本发明实施例的目的、技术方案和优点更加清楚，下面将结合本发明实施例中的附图，对本发明实施例中的技术方案进行描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

本发明实施例通过扩展内部网关协议（Interior Gateway Protocol, IGP）实现多协议标签交换（Multi-Protocol Label Switch, MPLS）网络虚拟化，从而解决 MRT FRR（最大冗余树快速重路由）转发机制存在的 IP（Internet Protocol）地址规划、标签使用量大等问题。

IGP 是一种专用于一个自治网络系统（比如：某个当地社区范围内的一个自治网络系统）中网关间交换数据流转通道信息的协议。网络协议（Internet Protocol, IP）或者其他的网络协议常常通过这些通道信息来确定怎样传送数据流。IGP 包括：路由信息协议（Routing Information Protocol, RIP）、开放最短路径优先（Open Shortest Path First, OSPF）、内部网关路由选择协议（Interior Gateway Routing Protocol, IGRP）、增强型内部网关路由协议（Enhanced Interior Gateway Routing Protocol, EIGRP）和中间系统到中间系统（Intermediate System to Intermediate System, IS-IS）等。

25

如图 3 所示，本发明实施例的一种网络系统，包括至少一个控制器和

多个转发设备，所述多个转发设备组成多个虚拟网络拓扑。

所述至少一个控制器，用于为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，并将全局标签与所述虚拟网络拓扑的绑定关系发送给所述多个转发设备。例如，有 N 个转发设备组成 X 个虚拟网络拓扑，则
 5 控制器需要为每个网络拓扑分配一个全局标签，并且，分配给不同虚拟网络拓扑的全局标签不同。即，控制器分配给拓扑 ID 为 TID₁ 的虚拟网络拓扑的标签为 L₁，分配给拓扑 ID 为 TID₂ 的虚拟网络拓扑的全局标签为 L₂，L₁ 与 L₂ 不相同。控制器为虚拟网络拓扑分配全局标签后，将该虚拟网络拓扑的拓扑 ID 与分配给该虚拟网络拓扑的全局标签绑定或关联起来，形成如
 10 表 1 所示的绑定关系表。如表 1 所示，绑定关系表中的每个绑定关系表项包括虚拟网络拓扑的拓扑 ID 和分配给该虚拟网络拓扑的全局标签。

表 1 绑定关系表

虚拟网络拓扑的拓扑 ID	分配的全局标签
TID ₁	L ₁
TID ₂	L ₂
TID ₃	L ₃
.....
TID _n	L _n

所述多个转发设备，用于接收所述全局标签与所述虚拟网络拓扑的绑定关系，根据该绑定关系生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项。当转发设备收到的报文中携带有全局标签（比如第一全局标签）时，该转发设备根据该第一全局标签查找到该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该报文。
 15

如果所述转发设备收到的报文中没有携带全局标签，并且该报文需要从某个虚拟网络拓扑（比如第二虚拟网络拓扑）转发，则该转发设备根据该第二虚拟网络拓扑的拓扑 ID（比如 TID₂）对应的第二全局标签 L₂，为该报文增加该第二全局标 L₂，该第二全局标签在下一跳节点索引到对应该第
5 二虚拟网络拓扑的转发表。

在 MRT FRR 场景中，进入网络的报文在缺省网络拓扑中以网络协议（Internet Protocol, IP）转发的方式转发，而且转发设备中缺省拓扑的 FIB（Forwarding Information Base）的转发表项有在某个虚拟网络拓扑（比如第三虚拟网络拓扑）中的备份路径。当转发设备在缺省拓扑中接收到 IP 报文，
10 而且该报文因链路或节点故障无法被转发到下一跳时，该转发设备会根据缺省拓扑的转发表项确定该备份路径所在的第三虚拟网络拓扑的拓扑 ID（比如 TID₃），并为该报文增加该拓扑 ID（比如 TID₃）对应的第三全局标签（比如 L₃），并根据该第三全局标签 L₃ 对应的第三虚拟网络拓扑中的备份路径转发该报文。

15

本发明实施例中提到的“转发设备”、“客户端”均支持多拓扑（multi-topology）技术，可以为多个拓扑安装 FIB（Forwarding Information Base），即，每个“转发设备”/“客户端”上有多个转发实例（forwarding instance）。每个转发实例对应一个虚拟网络拓扑（virtual network topology）。

20 如表 2 所示，该标签转发表项可以包括全局标签和该全局标签对应的虚拟网络拓扑转发表。

当转发设备收到报文时，如果该收到的报文携带有全局标签，比如 L₁，则根据该全局标签查找该全局标签对应的标签转发表项，从该全局标签对应的标签转发表项中找到对应该虚拟网络拓扑中的路由转发表，比如 FIB₁，
25 再根据 FIB₁ 转发该报文。

如果所述转发设备收到的报文中没有携带全局标签，并且该报文需要从某个虚拟网络拓扑（比如第二虚拟网络拓扑）转发，则该转发设备根据

该第二虚拟网络拓扑的拓扑 ID（比如 TID₂）对应的第二全局标签 L₂，为该报文增加该第二全局标 L₂，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表。

5 在 MRT FRR 场景中，进入网络的报文在缺省网络拓扑中以 IP 转发的方式转发，而且转发设备中缺省拓扑的 FIB（Forwarding Information Base）的转发表项有在虚拟网络拓扑中的备份路径。当转发设备在缺省拓扑中接收到 IP 报文，而且该报文因链路或节点故障无法被转发到下一跳时，，转发设备会根据缺省拓扑的转发表项确定备份路径所在的虚拟网络拓扑的拓扑 ID（比如 TID₁）为该报文增加该拓扑 ID（比如 TID₁）对应的全局标签
10 （比如 L₁），并根据该增加的全局标签对应的虚拟网络拓扑中的备份路径转发该报文。

表 2 标签转发表

全局标签	虚拟网络拓扑转发表
L ₁	FIB ₁
L ₂	FIB ₂
L ₃	FIB ₃
.....
L _n	FIB _n

15 可选地，所述全局标签可以是所有转发设备共享的全局标签空间中选择的单个的全局标签。

可选地，所述全局标签可以是多个本地标签组成的，即，每个全局标签包括两个或多个本地标签。控制器为每个虚拟网络拓扑分配两个或两个以上的本地标签，分配给某个虚拟网络拓扑的两个或两个以上本地标签

组成一个全局标签，组成的全局标签用来标识该虚拟网络拓扑。

可选地，所述转发设备为路由器或者交换机。

可选地，所述控制器可以是服务器或者路由器。

5 本发明实施例的一种转发报文的方法如图4所示，该方法应用于包括至少一个控制器和多个转发设备的网络系统，所述多个转发设备组成多个虚拟网络拓扑，该方法包括：

S410、接收全局标签与虚拟网络拓扑的绑定关系，并根据所述绑定关系生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项，
10 所述绑定关系是由所述控制器为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，并将所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系发送给所述多个转发设备的。例如，有N个转发设备组成X个虚拟网络拓扑，则控制器需要为每个网络拓扑分配一个全局标签，并且，控制器分配给所述虚拟网络拓扑中的不同虚拟网络拓扑的全局标签不同。

15 S412、根据收到的报文携带的全局标签确定从该携带的全局标签对应的标签转发表项转发该收到的报文。

当转发设备收到的报文中携带有全局标签（比如第一全局标签）时，并根据该第一全局标签查找到该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的第一标签转发表项
20 转发该报文。

如果所述转发设备收到的报文中没有携带全局标签，并且该报文需要从某个虚拟网络拓扑（比如第二虚拟网络拓扑）转发，则该转发设备根据该第二虚拟网络拓扑的拓扑ID（比如TID₂）对应的第二全局标签L₂，为该报文增加该第二全局标L₂，该第二全局标签在下一跳节点索引到对应该第
25 二虚拟网络拓扑的转发表。

在MRT FRR场景中，进入网络的报文在缺省网络拓扑中以IP转发的方

式转发，而且转发设备中缺省拓扑的FIB（Forwarding Information Base）的转发表项有在某个虚拟网络拓扑（比如第三虚拟网络拓扑）中的备份路径。当转发设备在缺省拓扑中接收到IP报文，而且该报文因链路或节点故障无法被转发到下一跳时，该转发设备会根据缺省拓扑的转发表项确定该备份路径所在的第三虚拟网络拓扑的拓扑ID（比如TID₃），并为该报文增加该拓扑ID（比如TID₃）对应的第三全局标签（比如L₃），并根据该第三全局标签L₃对应的第三虚拟网络拓扑中的备份路径转发该报文。

可选地，所述全局标签可以是所有转发设备共享的全局标签空间中选择的单个的全局标签。

10 可选地，所述全局标签可以是多个本地标签组成的，即，每个全局标签包括两个或多个本地标签。控制器为每个虚拟网络拓扑分配两个或两个以上的本地标签，分配给某个虚拟网络拓扑的两个或两个以上本地标签组成一个全局标签，组成的全局标签用来标识该虚拟网络拓扑。

15 可选地，所述至少一个控制器转发所述绑定关系给所述多个转发设备，具体为：所述至少一个控制器通过IGP扩展发送所述绑定关系给IGP域内的IGP客户端。

以下实施例以控制器和转发设备位于IGP域内为例进行说明。

20 如图5所示，本发明一个实施例的网络架构包括一个或多个IGP服务器（server）以及一个或多个IGP客户端（client）。IGP服务器作为控制器，IGP客户端为IGP域中的转发设备，该转发设备可以是具有转发功能的设备，比如路由器、交换机等。所述一个或多个IGP server和所述多个IGP客户端位于同一个IGP域内，所述多个IGP客户端组成多个虚拟网络拓扑，所述至少一个IGP服务器，用于为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，并将每个全局标签与所述IGP域内对应该全局标签的虚拟网络拓扑的绑定关系发送给所述多个IGP客户端。例如，有N个转

发设备组成 X 个虚拟网络拓扑，则控制器需要为每个网络拓扑分配一个全局标签，并且，控制器分配给不同虚拟网络拓扑的全局标签不同。

所述多个 IGP 客户端，用于接收所述全局标签与虚拟网络拓扑的绑定关系，根据该绑定关系生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；当转发设备收到的报文中携带有全局标签（比如第一全局标签）时，并根据该第一全局标签查找到该第一全局标签对应的第一标签转发表项，从而确定通过该该第一全局标签对应的第一虚拟网络拓扑的转发表转发该报文。

如果所述转发设备收到的报文中没有携带全局标签，并且该报文需要从某个虚拟网络拓扑（比如第二虚拟网络拓扑）转发，则该转发设备根据该第二虚拟网络拓扑的拓扑 ID（比如 TID₂）对应的第二全局标签 L₂，该第二全局标签 L₂ 该全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表。

在 MRT FRR 场景中，进入网络的报文在缺省网络拓扑中以 IP 转发的方式转发，而且转发设备中缺省拓扑的 FIB（Forwarding Information Base）的转发表项有在某个虚拟网络拓扑（比如第三虚拟网络拓扑）中的备份路径。当转发设备在缺省拓扑中接收到 IP 报文，而且该报文因链路或节点故障无法被转发到下一跳时，该转发设备会根据缺省拓扑的转发表项确定该备份路径所在的第三虚拟网络拓扑的拓扑 ID（比如 TID₃），并为该报文增加该拓扑 ID（比如 TID₃）对应的第三全局标签（比如 L₃），并根据该第三全局标签 L₃ 对应的第三虚拟网络拓扑中的备份路径转发该报文。该虚拟网络拓扑标签转发表项的组成和实现可以参考表 2 所示的内容。

可选地，所述全局标签可以是所有转发设备共享的全局标签空间中选择的单个的全局标签。

可选地，所述全局标签可以是多个本地标签组成的，即，每个全局标签包括两个或多个本地标签。控制器为每个虚拟网络拓扑分配两个或两个以上的本地标签，分配给某个虚拟网络拓扑的两个或两个以上本地标签

组成一个全局标签，组成的全局标签用来标识该虚拟网络拓扑。

基于图 5 所示的网络架构，本发明实施例的一种转发报文的方法如图 6 所示，包括：

5 S610、IGP 服务器为 IGP 域内的多虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，为不同虚拟网络拓扑分配的全局标签不同。

IGP 域内可以包括一个或多个物理拓扑，在这些物理拓扑上，可以根据需要部署多个逻辑拓扑（虚拟网络拓扑）。多拓扑路由（Multi Topology Routing, MTR）是指在自治系统中独立运行多个虚拟网络拓扑，每个虚拟网络拓扑都会单独维护路由表并转发流量，每个路由可以属于多个虚拟网络拓扑。多拓扑路由实现了拓扑级别的流量调度，报文转发时先选择拓扑再选择路径。

IGP 支持多拓扑功能可以通过 IGP 在配置的每个虚拟网络拓扑单独实现 IGP 路由功能。IGP 支持多拓扑时，每个虚拟网络拓扑相当于一个路由域，保证拓扑内的状态信息同步，进行拓扑路由计算，支持路由扩展功能在每个拓扑内正常执行。

IGP MTR 扩展特性将 IGP 基本功能都进行多拓扑划分，形成了 IGP 进程内的多拓扑模型。每个拓扑就像是一个独立的 IGP 进程，可以独立地维护链路状态信息，独立地计算最短路径并更新 IGP 路由表，最终与其它协议发布的路由一起进行优选得到拓扑路由表。

20 IGP 服务器为 IGP 域内的每个虚拟网络拓扑分配全局标签，即，使用一个全局标签来标识 IGP 域内的多拓扑中的一个虚拟拓扑。

其中 IGP 服务器可以是一个，也可能是多个 IGP 服务器组成的服务器组，多个 IGP 服务器之间可以进行负载分担。

25 S612、该 IGP 服务器将该全局标签与该全局标签对应的虚拟网络拓扑的绑定关系发送给该 IGP 客户端。

IGP 域内的控制器可以通过 IGP 扩展发送全局标签和该全局标签对应

的虚拟网络拓扑的绑定关系给 IGP 域内的转发设备。

图 7A 是 IGP 扩展实现 MPLS 网络虚拟化的一个实例，这里使用一个扩展的标签-长度-值 (Tag-Length-Value, TLV) 来表示 IGP 多拓扑和 MPLS 全局标签的绑定关系，这里的拓扑 ID (Topology ID, TID) 用于标识 IGP 多拓扑中的每个虚拟网络拓扑，Label 用于表示 MPLS 全局标签值。

IGP 服务器为 IGP 域内的每个虚拟网络拓扑分配全局标签，即，IGP 域内的每个虚拟网络拓扑的拓扑 ID 对应一个全局标签。

具体地，由于传统的 MPLS LDP 不支持分配全局标签。这里的全局标签分配时，IGP 服务器收集所有 IGP 客户端的标签能力 (即该 IGP 客户端可分配的标签值范围)，然后计算出所有 IGP 客户端共享的标签空间，可以将所有 IGP 客户端共享的标签空间预留出来作为全局标签分配空间。

可选地，如图 7D 所示，也可以使用现有的 MPLS 标签值范围，即 $0 \sim 2^{20} - 1$ ，为一个虚拟网络拓扑分配多个标签，将分配的多个标签组合起来的一个标签栈来构成一个全局标签，并使用该全局标签来标识该虚拟网络拓扑。比如为拓扑 ID 为 TID_n 的虚拟网络拓扑分配多个本地标签 L_{n1} 、 L_{n2} 、 L_{n3} ，则可以使用多个本地标签 $L_{n1}+L_{n2}+L_{n3}$ 组成的全局标签 L_n 来标识该拓扑 ID 为 TID_n 的虚拟网络拓扑，相应地，该虚拟网络拓扑的拓扑 ID 与分配该虚拟网络拓扑的全局标签之间的绑定关系表项可以为 $\langle TID_n, L_{n1}+L_{n2}+L_{n3} \rangle$ 。

IGP 服务器向 IGP 客户端发送标签分配消息，该标签分配消息中包括图 7A 或图 7D 所示的 TLV，该 TLV 表示 IGP 多拓扑和 MPLS 全局标签的绑定关系，这里的拓扑 ID (Topology ID, TID) 用于标识 IGP 多拓扑中的每个虚拟网络拓扑，Label 用于表示 MPLS 全局标签值。

当网络中的控制器和转发设备之间运行 OSPF 协议时，控制器为 OSPF 域内的多个虚拟网络拓扑中的每个拓扑分配一个全局标签，并将全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系发给转发设备。控制器可以通过构造第一链路状态广播数据 LSA (Link State Advertisement, LSA) 发送该全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系给转发

设备。该第一 LSA 包括 OSPF 域内的多个虚拟网络拓扑的全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系。

5 可选地，控制器将全局标签与虚拟网络拓扑的绑定关系通过第一 LSA 中的 TLV 发送给 OSPF 域内的转发设备。转发设备收到该第一 LSA 时，根据全局标签与该全局标签对应的虚拟网络拓扑的绑定关系，生成与包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项，该标签转发表项可以由全局标签索引到特定的一个虚拟网络拓扑的转发表，从而，当转发设备收到的报文中携带有全局标签时，可以根据携带的全局标签确定对应的虚拟网络拓扑转发表转发报文。该 TLV 可以是如图 7A 或图 7D 所示的
10 TLV。

当网络中的控制器和转发设备之间运行 ISIS 协议时，控制器为 ISIS 域内的多个虚拟网络拓扑中的每个拓扑分配一个全局标签，并将多拓扑与全局标签之间的绑定关系发给转发设备。控制器可以通过构造第一链路状态报文 LSP (Link-state PDU) 发送该全局标签与该全局标签对应的虚拟网络
15 拓扑之间的绑定关系给转发设备。该第一 LSP 包括 ISIS 域内的多个虚拟网络拓扑的全局标签与该全局标签对应的虚拟网络拓扑的绑定关系。

可选地，控制器将全局标签与该全局标签对应的虚拟网络拓扑的绑定关系通过第一 LSP 中的 TLV 发送给 ISIS 域内的转发设备。转发设备收到该第一 LSP 时，获取全局标签与该全局标签对应的虚拟网络拓扑的绑定关系，
20 根据该绑定关系生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项，该标签转发表项可以由全局标签索引到特定的一个虚拟网络拓扑的转发表。从而，当转发设备收到的报文中携带有全局标签时，可以根据携带的全局标签确定对应的虚拟网络拓扑转发表转发报文。该 TLV 可以是如图 7A 或图 7D 所示的 TLV。该 TLV 也可以是图 7B 所示的标签映射
25 TLV。如图 7B 所示，该标签映射 TLV 包括一个或多个标签子 TLV (Label Sub-TLV) 和一个或多个转发等价类子 TLV (FEC Sub-TLV)。该 Label Sub-TLV 包括控制器为虚拟网络拓扑分配的全局标签值，该 FEC Sub-TLV 包括为所述全局标签分配的转发等价类。该 FEC Sub-TLV 可以是图 7C 所示的 MPLS 多拓扑子 TLV，其中 Multi-Topology ID 包括分配的全局标签对

应的虚拟网络拓扑的 ID。

本发明实施例中提到的“全局标签”可以用如图 7D 所示的全局标签子 TLV (Global Label Sub-TLV) 中携带的多个标签组成的标签栈来表示。该标签栈用以标识一个虚拟网络拓扑。如图 7D 所示, 该 Global Label Sub-TLV 5 包括两个或两个以上标签, 该标签的标签值位于现有标签范围 $0 - 2^{20} - 1$ 内, 该标签可以是本地标签或全局标签。该标签域 (Label Field) 包括的多个标签组成标签栈, 每个标签域包括长度为 24 比特的标签值, 高 20bit 表示标签值, 低 4 比特位为栈底标识, 最后一个标签域的低 4 比特值置为 1。该 Global Label Sub-TLV 中的标签栈组合成的全局标签值可能大于 2^{20} 个, 即 10 扩大了现有 MPLS 标签范围。如果只有一个标签域, 则可以与现有的 MPLS 标签分配一致, 即 $16 \sim 2^{20} - 1$ 。

S614、IGP 客户端根据收到的全局标签和 IGP 内的多个虚拟网络拓扑的绑定关系, 生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项。

15 IGP 域中的转发设备根据收到的全局标签和虚拟网络拓扑的绑定关系, 生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项, 该标签转发表项由全局标签索引到特定的一个虚拟网络拓扑的转发表。如图 8 所示, 该标签转发表项包括: 全局标签和该全局标签对应的虚拟网络拓扑的 FIB。比如第一全局标签 L_1 对应的拓扑 ID 为 TID_1 的第一虚拟网络拓扑, 20 该转发设备上安装该第一虚拟网络拓扑的转发表项 FIB_1 , 当该转发设备收到的报文中携带有第一全局标签 L_1 时, 该转发设备查找该全局标签 L_1 对应的第一标签转发表项, 确定该全局标签 L_1 对应的第一虚拟网络拓扑转发表 FIB_1 , 则按照所述第一虚拟网络拓扑的 FIB_1 转发该报文。

如果所述转发设备收到的报文中没有携带全局标签, 并且该报文需要 25 从某个虚拟网络拓扑 (比如第二虚拟网络拓扑) 转发, 则该转发设备根据该第二虚拟网络拓扑的拓扑 ID (比如 TID_2) 对应的第二全局标签 L_2 , 为该报文增加该第二全局标签 L_2 , 该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表。

在 MRT FRR 场景中，进入网络的报文在缺省网络拓扑中以网络协议（Internet Protocol, IP）转发的方式转发，而且转发设备中缺省拓扑的 FIB（Forwarding Information Base）的转发表项有在某个虚拟网络拓扑（比如第三虚拟网络拓扑）中的备份路径。当转发设备在缺省拓扑中接收到 IP 报文，而且该报文因链路或节点故障无法被转发到下一跳时，该转发设备会根据缺省拓扑的转发表项确定该备份路径所在的第三虚拟网络拓扑的拓扑 ID（比如 TID₃），并为该报文增加该拓扑 ID（比如 TID₃）对应的第三全局标签（比如 L₃），并根据该第三全局标签 L₃ 对应的第三虚拟网络拓扑中的备份路径转发该报文。

10 S616、IGP 客户端收到报文后，根据该报文中携带的全局标签，查找该报文中写道的全局标签对应的标签转发表项，确定该全局标签对应的虚拟网络拓扑的 FIB，并通过确定的虚拟网络拓扑的 FIB 转发该报文。

IGP 客户端收到报文时，根据该报文中携带的全局标签转发表项，从标签转发表项中查询到该报文携带的全局标签对应的虚拟网络拓扑的 FIB，并通过确定的虚拟网络拓扑的 FIB 转发该报文。

15 如果所述转发设备收到的报文中没有携带全局标签，并且该报文需要从某个虚拟网络拓扑（比如第二虚拟网络拓扑）转发，则该转发设备根据该第二虚拟网络拓扑的拓扑 ID（比如 TID₂）对应的第二全局标签 L₂，为该报文增加该第二全局标签 L₂，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表。

在 MRT FRR 场景中，进入网络的报文在缺省网络拓扑中以网络协议（Internet Protocol, IP）转发的方式转发，而且转发设备中缺省拓扑的 FIB（Forwarding Information Base）的转发表项有在某个虚拟网络拓扑（比如第三虚拟网络拓扑）中的备份路径。当转发设备在缺省拓扑中接收到 IP 报文，而且该报文因链路或节点故障无法被转发到下一跳时，该转发设备会根据缺省拓扑的转发表项确定该备份路径所在的第三虚拟网络拓扑的拓扑 ID（比如 TID₃），并为该报文增加该拓扑 ID（比如 TID₃）对应的第三全局标签（比如 L₃），并根据该第三全局标签 L₃ 对应的第三虚拟网络拓扑中的备

份路径转发该报文。

以节点 A、B、C、D 组成的网络为例，通过 MRT 可以计算出三个拓扑：

缺省拓扑（参考图 9A）：节点设备 A->节点设备 B->节点设备 C 和节点设备 A->节点设备 D->节点设备 C；

5 第一虚拟网络拓扑（参考图 9B）：节点设备 A->节点设备 B->节点设备 C->节点设备 D->节点设备 A；

第二虚拟网络拓扑（参考图 9C）：节点设备 A->节点设备 D->节点设备 C->节点设备 B->节点设备 A。

如图 9A 至 9D 所示，通过为 IGP 多拓扑分配全局标签的实施例，可以
10 为 MRT FRR 中的第一虚拟网络拓扑（参考图 9B）和第二虚拟网络拓扑（参考图 9C）分配全局标签来标识（图 9B 中用标签 200 来标识第一虚拟网络拓扑，在图 9C 中用全局标签 300 来标识第二虚拟网络拓扑），这样在转发报文的时候可以通过报文中携带的全局标签，查找转发设备的标签转发表项，从而得到该报文携带的全局标签对应的虚拟网络拓扑的 FIB，将报文引
15 导到该全局标签对应的虚拟网络拓扑的 FIB 中转发。比如，节点设备 A 接收第一报文，该第一报文携带的全局标签为 200，根据全局标签 200 查询节点设备 A 的标签转发表项，确定全局标签 200 对应的 FIB 为 FIB₂₀₀，则节点设备 A 根据全局标签 200 对应的第一虚拟网络拓扑的路由表 FIB₂₀₀ 来转发该第一报文。类似地，如果该第一报文携带的全局标签为 300，根据全局
20 标签 300 查询节点设备 A 的标签转发表项，发现全局标签 300 对应的 FIB 为 FIB₃₀₀，则节点设备 A 根据全局标签为 300 对应的第二虚拟网络拓扑的 FIB₃₀₀ 来转发该第一报文。

对于图 9A 中的缺省拓扑中的报文的路由，可以通过图 9B 所示的第一虚拟网络拓扑或图 9C 所示的第二虚拟网络拓扑中的备份路由来保护。如图
25 9D 中缺省拓扑中的路由被 MRT 计算得到的图 9C 所示的第二虚拟网络拓扑保护。当某一 IP 报文，比如 IP 报文 P 被从节点设备 A 发往节点设备 D 时，IP 报文 P 按照缺省拓扑的转发表从节点设备 A 到达节点设备 B，节点设备 B 发现节点设备 B 与节点设备 C 之间的有向路径（节点设备 B->节点设备 C）

发生故障，则节点设备 B 为 IP 报文 P 增加全局标签 300，从而使得流量切换到第二虚拟网络拓扑中的节点设备 B 到节点设备 A 的有向链路，再经由节点设备 D 到达目的节点设备 C，即报文 P 沿着以下路径被发送到节点设备 C：节点设备 A->节点设备 B->节点设备 A->节点设备 D->节点设备 C。

5 可选地，该实施例中，所述本地标签可以是 MPLS 标签。

可选地，所述全局标签可以是所有转发设备共享的全局标签空间中选择的单个的全局标签。

10 可选地，所述全局标签可以是多个本地标签组成的，即，每个全局标签包括两个或多个本地标签。控制器为每个虚拟网络拓扑分配两个或两个以上的本地标签，分配给某个虚拟网络拓扑的两个或两个以上本地标签组成一个全局标签，组成的全局标签用来标识该虚拟网络拓扑。

基于图 5 所示的网络架构，本发明实施例的另一种报文转发方法如图 10 所示，该方法包括：

15 S1000、IGP 服务器为 IGP 域内的多拓扑中的每个虚拟网络拓扑分配全局标签，IGP 服务器为不同虚拟网络拓扑分配的全局标签不同。

IGP 服务器为 IGP 域内的每个虚拟网络拓扑分配一个全局标签，即，使用一个全局标签来标识 IGP 域内的多拓扑中的一个虚拟网络拓扑。其中 IGP 服务器可以是一个，也可能是多个 IGP 服务器组成的服务器组，多个 IGP 服务器之间可以进行负载分担。

20 S1002、IGP 服务器将所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系发送给 IGP 客户端。

S1004、IGP 客户端根据该绑定关系生成 MRT FRR 的转发表项。

25 IGP 客户端根据收到的全局标签与 IGP 域内多个虚拟网络拓扑的绑定关系，通过 MRT 为缺省拓扑中的 IP 路由计算备份路径，生成 MRT FRR 的转发表项。所述备份路径用于当缺省拓扑中的主路径故障或中断时，主路径上正常转发的最后一个节点设备（IGP 客户端）为该报文加上标识备份路

径所在的虚拟网络拓扑的全局标签，从而该报文能够被切换到备份路径所在的虚拟网络拓扑，备份路径上接收到该报文的 IGP 客户端根据该报文携带的全局标签确定该全局标签标识的虚拟网络拓扑的 FIB，按照确定的 FIB 转发该报文。该备份路径通过图 9B 所示的第一虚拟网络拓扑或者图 9C 所示的第二虚拟网络拓扑转发。在生成路由转发表项的时候，对于特定的 IP 前缀，位于缺省拓扑中的主路径采用普通的 IP 路由转发。对于位于另一虚拟网络拓扑中的备份路径，则根据 IGP 服务器发送的全局标签与虚拟网络拓扑的绑定关系，生成该全局标签对应的虚拟网络拓扑的标签转发表项，该标签转发表项包括：全局标签和该全局标签对应的虚拟网络拓扑的 FIB。如图 11 所示，该虚拟网络拓扑的标签转发表项可以包括全局标签、出接口和下一跳信息。该全局标签用于标识备份路径所在的第一虚拟网络拓扑或第二虚拟网络拓扑，具体取决于备份路径在第一虚拟网络拓扑还是第二虚拟网络拓扑中。如果备份路径在第一虚拟网络拓扑中，则该全局标签用于标识该第一虚拟网络拓扑；如果备份路径在第二虚拟网络拓扑中，则该全局标签用于标识该第二虚拟网络拓扑。出接口和下一跳则是 MRT 计算得到的对应的备份路径中用于备份的出接口和下一跳。

通过 MRT 计算还会生成两个虚拟网络拓扑（第一虚拟网络拓扑和第二虚拟网络拓扑）的路由转发表（FIB，每个 FIB 表项将每个 IP 前缀对应到一个标识本虚拟网络拓扑的全局标签及出接口和下一跳。

20 S1006、根据生成的转发表项和路由表转发报文

1) 转发设备接收到 IP 报文时，根据缺省拓扑的路由表转发该 IP 报文。当转发设备发现故障发生的时候，会为该报文加上标识备份路径所在的虚拟网络拓扑的全局标签，使用备份路径所在的虚拟网络拓扑转发表项转发该 IP 报文，实现流量保护倒换。这时转发设备会根据备份路径所在的虚拟网络拓扑的转发表在该 IP 报文中加入标识备份路径所在的虚拟网络拓扑的全局标签，并通过备份路径所在的虚拟网络拓扑的备份的出接口和下一跳，将该报文发送给下一跳节点。

2) 下一跳节点接收到该报文（此时的报文是 MPLS 报文）后，首先会

根据该报文中的 MPLS 全局标签，查找 MPLS 全局标签对应的标签转发表项，通过该标签转发表项索引到对应的备份拓扑（比如第二虚拟网络拓扑），在对应的备份拓扑（即第二虚拟网络拓扑）的 FIB 中查找路由转发表项，根据查找到的路由转发表项对该报文压入标识该备份路径所在的虚拟网络
5 拓扑的全局标签，并通过出接口和下一跳信息继续发送到下一跳节点。

3) 下一跳节点重复 2) 中的转发处理，直至在备份路径所在的虚拟网络拓扑中将该报文转发到目的节点。

可选地，所述全局标签可以是所有转发设备共享的全局标签空间中选择的单个的全局标签。

10 可选地，所述全局标签可以是多个本地标签组成的，即，每个全局标签包括两个或多个本地标签。控制器为每个虚拟网络拓扑分配两个或两个以上的本地标签，分配给某个虚拟网络拓扑的两个或两个以上本地标签组成一个全局标签，组成的全局标签用来标识该虚拟网络拓扑。

本发明实施例的控制器、转发设备、转发报文的方法，通过 IGP 扩展
15 为 IGP 多拓扑分配 MPLS 全局标签或标签栈实现 MPLS 网络虚拟化，该方法可以用于 MRT FRR 的转发，达到如下有益效果：通过使用全局标签或标签栈实现 IP 网络的 MRT FRR，只需要扩展 IGP 即可，不需要部署 LDP。通过使用全局标签，只需要使用标识各个虚拟网络拓扑的少量全局标签就可以实现，简化了网络虚拟化及 MRT FRR 的部署。

20

如图 12 所示，图 3 所示的网络系统中，控制器包括标签分配模块 1200 和标签发送模块 1202，其中

标签分配模块 1200 用于：为所述多个拓扑中的每个虚拟网络拓扑分配全局标签；例如，有 N 个转发设备组成 X 个虚拟网络拓扑，则控制器需要
25 为每个网络拓扑分配全局标签，并且，标签分配模块 1200 分配给不同虚拟网络拓扑的全局标签不同。

标签发送模块 1202 用于：将所述全局标签与该全局标签对应的虚拟网

络拓扑的绑定关系发送给所述多个转发设备。

如图 12 所示，图 3 所示的网络系统中，转发设备包括接收模块 1204、标签转发表项生成模块 1206 和转发模块 1208，其中

接收模块 1204，用于接收控制器发送的所述全局标签与所述虚拟网络
5 拓扑的绑定关系；所述接收模块 1204，还用于接收报文。

标签转发表项生成模块 1206，用于根据接收模块 1204 收到的所述全局
标签与虚拟网络拓扑的绑定关系，生成与该包括该全局标签与虚拟网络拓
扑转发表映射关系的标签转发表项。

转发模块 1208，用于根据接收模块 1204 收到的第一报文携带的第一全
10 局标签，在标签转发表项生成模块 1206 生成的标签转发表项中查到该第一
全局标签对应的第一虚拟网络拓扑的转发表，并通过该第一虚拟网络拓扑
的转发表转发该第一报文。

可选地，该转发设备还包括标签增加模块 1210。如果接收模块 1204 收
到的报文中没有携带全局标签，并且该报文需要从某个虚拟网络拓扑（比
15 如第二虚拟网络拓扑）转发，则转发模块 1208 根据该第二虚拟网络拓扑的
拓扑 ID（比如 TID_2 ）对应的第二全局标签 L_2 ，为该报文增加该第二全局标
 L_2 ，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发
表。

在 MRT FRR 场景中，进入网络的报文在缺省网络拓扑中以网络协议
20 （Internet Protocol, IP）转发的方式转发，而且转发设备中缺省拓扑的 FIB
（Forwarding Information Base）的转发表项有在某个虚拟网络拓扑（比如第
三虚拟网络拓扑）中的备份路径。当接收模块 1204 在缺省拓扑中接收到 IP
报文，而且该报文因链路或节点故障无法被转发到下一跳时，该接收模块
1204 会根据缺省拓扑的转发表项确定该备份路径所在的第三虚拟网络拓扑
25 的拓扑 ID（比如 TID_3 ），并由标签增加模块 1210 为该报文增加该拓扑 ID
（比如 TID_3 ）对应的第三全局标签（比如 L_3 ），并由转发模块 1208 根据
该第三全局标签 L_3 对应的第三虚拟网络拓扑中的备份路径转发该报文。

可选地，所述虚拟网络拓扑转发表可以是如表 2 所示的标签转发表或者如图 8 所示的标签转发表。

可选地，所述全局标签可以是所有转发设备共享的全局标签空间中选择的单个的全局标签。

5 可选地，所述全局标签可以是多个本地标签组成的，即，每个全局标签包括两个或多个本地标签。控制器为每个虚拟网络拓扑分配两个或两个以上的本地标签，分配给某个虚拟网络拓扑的两个或两个以上本地标签组成一个全局标签，组成的全局标签用来标识该虚拟网络拓扑。

可选地，所述转发设备为路由器或者交换机。

10 可选地，所述控制器可以是服务器或者路由器。

可选地，所述控制器可以是 IGP 服务器。

可选地，所述模块可以是 ASIC (Application Specific Integrated Circuit)，或者现场可编程门阵列 (Field-Programmable Gate Array, FPGA)，或者可编程逻辑控制器 (Programmable Logic Controller, PLC)，也可以是可编程
15 网络的一个网元或者软件定义网络的一个网元。

图 12 及图 12 对应的上下文中所示的控制器及转发设备之间的互动，可以参考图 4~11 及其对应的说明书的说明。图 12 及图 12 对应的上下文中所述的控制器也可以是图 5 及图 5 对应的上下文所示的 IGP 服务器，相应地，图 12 及图 12 对应的上下文中所述的转发设备也可以是图 5 及图 5
20 对应的上下文所示的 IGP 客户端。

如图 13 所示，图 3 所示的网络系统中，控制器可以包括第一处理器 1300 和第一接口 1302，其中

第一处理器 1300，用于：为所述多个拓扑中的每个虚拟网络拓扑分配
25 全局标签；例如，有 N 个转发设备组成 X 个虚拟网络拓扑，则控制器需要

为每个网络拓扑分配全局标签，并且，第一处理器 1300 分配给不同虚拟网络拓扑的全局标签不同。

第一接口 1302，用于：将所述全局标签与该全局标签对应的虚拟网络拓扑的绑定关系发送给所述多个转发设备。

5 如图 13 所示，图 3 所示的网络系统中，转发设备包括第二接口 1304、第二处理器 1306 和第三处理器 1308，其中

第二接口 1304，用于接收控制器发送的全局标签与虚拟网络拓扑的绑定关系。

10 第二处理器 1306，用于根据第二接口 1304 收到的所述全局标签与所述虚拟网络拓扑的绑定关系生成与该包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项。

15 第三处理器 1308，用于根据接收到的第一报文携带的第一全局标签，在第二处理器 1306 生成的标签转发表项中查到该第一全局标签对应的第一虚拟网络拓扑的转发表，并通过该第一虚拟网络拓扑的转发表转发该第一报文。

20 可选地，该转发设备还包括第四处理器 1310。如果第二接口 1304 收到的报文中没有携带全局标签，并且该报文需要从某个虚拟网络拓扑（比如第二虚拟网络拓扑）转发，则第三处理器 1308 根据该第二虚拟网络拓扑的拓扑 ID（比如 TID_2 ）对应的第二全局标签 L_2 ，由第四处理器 1310 为该报文增加该第二全局标 L_2 ，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表。

25 在 MRT FRR 场景中，进入网络的报文在缺省网络拓扑中以网络协议（Internet Protocol, IP）转发的方式转发，而且转发设备中缺省拓扑的 FIB（Forwarding Information Base）的转发表项有在某个虚拟网络拓扑（比如第三虚拟网络拓扑）中的备份路径。当转发设备在缺省拓扑中接收到 IP 报文，而且该报文因链路或节点故障无法被转发到下一跳时，该转发设备会根据缺省拓扑的转发表项确定该备份路径所在的第三虚拟网络拓扑的拓扑 ID

(比如 TID_3)，并为该报文增加该拓扑 ID (比如 TID_3) 对应的第三全局标签 (比如 L_3)，并根据该第三全局标签 L_3 对应的第三虚拟网络拓扑中的备份路径转发该报文。

5 可选地，所述虚拟网络拓扑转发表可以是如表 2 所示的虚拟网络拓扑转发表或者如图 8 所示的虚拟网络拓扑转发表。

可选地，所述全局标签可以是所有转发设备共享的全局标签空间中选择的单个的全局标签。

10 可选地，所述全局标签可以是多个本地标签组成的，即，每个全局标签包括两个或多个本地标签。控制器为每个虚拟网络拓扑分配两个或两个以上的本地标签，分配给某个虚拟网络拓扑的两个或两个以上本地标签组成一个全局标签，组成的全局标签用来标识该虚拟网络拓扑。

可选地，所述转发设备为路由器或者交换机。

可选地，所述控制器可以是服务器或者路由器。

15 可选地，所述控制器可以是 IGP 服务器。

20 可选地，所述第一处理器、第二处理器或第三处理器可以是 ASIC (Application Specific Integrated Circuit)，或者现场可编程门阵列 (Field-Programmable Gate Array, FPGA)，或者可编程逻辑控制器 (Programmable Logic Controller, PLC)，也可以是可编程网络的一个网元或者软件定义网络的一个网元。第三处理器也可以是网络处理器 (Network Processor, NP)。

25 图 13 及图 13 对应的上下文中所示的控制器及转发设备之间的互动，可以参考图 4~11 及其对应的说明书的说明。图 13 及图 13 对应的上下文中所述的控制器也可以是图 5 及图 5 对应的上下文所示的 IGP 服务器，相应地，图 13 及图 13 对应的上下文中所述的转发设备也可以是图 5 及图 5 对应的上下文所示的 IGP 客户端。

本发明实施例的控制器、转发设备、转发报文的方法，通过 IGP 扩展

为 IGP 多拓扑分配 MPLS 全局标签或标签栈实现 MPLS 网络虚拟化，该方法可以用于 MRT FRR 的转发，达到如下有益效果：通过使用全局标签或标签栈实现 IP 网络的 MRT FRR，只需要扩展 IGP 即可，不需要部署 LDP。通过使用全局标签或标签栈，只需要使用标识各个虚拟网络拓扑的少量全局标签或标签栈就可以实现，简化了网络虚拟化及 MRT FRR 的部署。

另外，本发明所有实施例所述的“绑定关系”即“映射关系”或关联关系。

在本申请所提供的几个实施例中，应该理解到，所揭露的系统，装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

所述作为分离部件说明的单元可以是或者也可以不是物理上分开的，作为单元显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部单元来实现本实施例方案的目的。

另外，在本发明各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能单元的形式实现。比如，检测器、发送器、接收器、获得单元都可以通过通用中央处理器 CPU 或专用集成电路（Application Specific Integrated Circuit, ASIC）或现场可编程门阵列（Field-Programmable Gate Array, FPGA）来实现。

上述以软件功能单元的形式实现的集成的单元，可以存储在一个计算机可读取存储介质中。上述软件功能单元存储在一个存储介质中，包括若干指令用以使得一台计算机设备（可以是个人计算机，服务器，或者网络设备等等）执行本发明各个实施例所述方法的部分步骤。而前述的存储介质

包括：U 盘、移动硬盘、只读存储器（Read-Only Memory，简称 ROM）、随机存取存储器（Random Access Memory，简称 RAM）、磁碟或者光盘等各种可以存储程序代码的介质。

5 最后应说明的是：以上实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述各实施例所记载的技术方案进行修改，或者对其中部分技术特征进行等同替换；而这些修改或者替换，并不使相应技术方案的本质脱离本发明各实施例技术方案的精神和范围。

权利要求

1、一种网络系统，其特征在于，包括至少一个控制器和多个转发设备，所述多个转发设备组成多个虚拟网络拓扑，

5 所述至少一个控制器，用于为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，并将所述全局标签与所述多个虚拟网络拓扑的绑定关系发送给所述多个转发设备；其中所述控制器分配给所述多个虚拟网络拓扑中的不同虚拟网络拓扑的全局标签不同；

所述多个转发设备，用于接收所述全局标签与虚拟网络拓扑的绑定关系，根据该绑定关系生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；当该转发设备收到的第一报文中携带有第一全局标签时，
10 该转发设备根据该第一全局标签查找到该第一全局标签对应的第一虚拟网络拓扑的转发表，确定通过该第一虚拟网络拓扑的转发表转发该第一报文；

所述第一虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

2、如权利要求 1 所述的网络系统，其特征在于，所述多个转发设备还
15 用于：

当所述转发设备收到的第二报文中没有携带全局标签，并且该第二报文将要从第二虚拟网络拓扑转发，则该转发设备根据该第二虚拟网络拓扑对应的第二全局标签，为该第二报文增加该第二全局标签的封装，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

20 所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

3、如权利要求 1 或 2 所述的网络系统，其特征在于，所述控制器包括标签分配模块和标签发送模块，其中

所述标签分配模块用于：为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，为所述多个虚拟网络拓扑中的不同虚拟网络拓扑分配的
25 的标签不同；

所述标签发送模块用于：将所述全局标签与该全局标签对应的虚拟网络拓扑的绑定关系发送给所述多个转发设备。

4、如权利要求 1-3 中任一所述的网络系统，其特征在于，所述转发设备包括接收模块、标签转发表项模块和转发模块，其中

5 所述接收模块，用于接收所述控制器发送的所述全局标签与该全局标签对应的虚拟网络拓扑的绑定关系；

所述标签转发表项生成模块，根据所述全局标签与虚拟网络拓扑的绑定关系，生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；

10 所述转发模块，用于根据收到的第一报文携带的第一全局标签确定该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该第一报文。

5、如权利要求 4 所述的网络系统，其特征在于，所述转发设备还包括标签增加模块，

15 当所述接收模块收到的第二报文中没有携带全局标签，并且该第二报文将要从第二虚拟网络拓扑转发，则所述转发模块根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，为该第二报文增加该第二全局标签，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

20 6、如权利要求 1 或 2 所述的网络系统，其特征在于，所述控制器包括第一处理器和第一接口，其中

所述第一处理器用于：为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，为所述多个虚拟网络拓扑中的不同虚拟网络拓扑分配的全局标签不同；

25 所述第一接口用于：将所述全局标签与该全局标签对应的虚拟网络拓

扑之间的绑定关系发送给所述多个转发设备。

7、如权利要求 1 或 2 或 6 所述的网络系统，其特征在于，所述转发设备包括第二接口、第二处理器和第三处理器，其中

5 所述第二接口，用于接收所述控制器发送的所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系；

所述第二处理器，用于根据所述全局标签与所述虚拟网络拓扑的绑定关系，生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；

10 所述第三处理器，用于根据收到的第一报文中携带的第一全局标签确定该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该第一报文。

8、如权利要求所述 7 的网络系统，其特征在于，该转发设备还包括第四处理器；

15 如果该第二接口收到的第二报文中没有携带全局标签，并且该第二报文需要从第二虚拟网络拓扑转发，则该第三处理器根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，由该第四处理器为该第二报文增加该第二全局标，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

20 9、如权利要求 1~8 中任一所述的网络系统，其特征在于，所述控制器为内部网关协议 IGP 服务器。

10、如权利要求 1~9 中任一所述的网络系统，其特征在于，所述全局标签包括两个或两个以上的标签。

25 11、一种控制器，其特征在于，该控制器应用于包括该控制器和多个

转发设备的网络系统中，所述多个转发设备组成多个虚拟网络拓扑，

该控制器包括标签分配模块和标签发送模块，其中

所述标签分配模块用于：为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，为所述多个虚拟网络拓扑中的不同虚拟网络拓扑分配的标签不同；

所述标签发送模块用于：将所述全局标签与该全局标签对应的虚拟网络拓扑的绑定关系发送给所述多个转发设备。

12、如权利要求 11 所述的控制器，其特征在于，所述全局标签包括两个或两个以上的标签。

10 13、一种控制器，其特征在于，该控制器应用于包括该控制器和多个转发设备的网络系统中，所述多个转发设备组成多个虚拟网络拓扑，

所述控制器包括第一处理器和第一接口，其中

所述第一处理器用于：为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，为所述多个虚拟网络拓扑中的不同虚拟网络拓扑分配的全局标签不同；

所述第一接口用于：将所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系发送给所述多个转发设备。

14、如权利要求 13 所述的控制器，其特征在于，所述全局标签包括两个或两个以上的标签。

20

15、一种转发设备，其特征在于，该转发设备应用于包括多个该转发设备和至少一个控制器的网络系统中，所述多个转发设备组成多个虚拟网络拓扑，

所述转发设备包括接收模块、标签转发表项模块和转发模块，其中

所述接收模块，用于接收所述控制器发送的所述全局标签与该全局标签对应的虚拟网络拓扑的绑定关系；

所述标签转发表项生成模块，根据所述全局标签与虚拟网络拓扑的绑定关系，生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；

所述转发模块，用于根据收到的第一报文携带的第一全局标签确定该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该第一报文。

16、如权利要求 15 所述的转发设备，其特征在于，所述转发设备还包括标签增加模块，

当所述接收模块收到的第二报文中没有携带全局标签，并且该第二报文将从第二虚拟网络拓扑转发，则所述转发模块根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，为该第二报文增加该第二全局标签，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

15 所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

17、如权利要求 15 或 16 所述的转发设备，其特征在于，所述全局标签包括两个或两个以上的标签。

18、一种转发设备，其特征在于，该转发设备应用于包括多个该转发设备和至少一个控制器的网络系统中，所述多个转发设备组成多个虚拟网络拓扑，

所述转发设备包括第二接口、第二处理器和第三处理器，其中

所述第二接口，用于接收所述控制器发送的所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系；

所述第二处理器，用于根据所述全局标签与所述虚拟网络拓扑的绑定关系，生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表

项；

所述第三处理器，用于根据收到的第一报文中携带的第一全局标签确定该第一全局标签对应的第一标签转发表项，从而确定从该第一全局标签对应的第一虚拟网络拓扑的转发表转发该第一报文。

5 19、如权利要求 18 所述的转发设备，其特征在于，还包括第四处理器；

如果该第二接口收到的第二报文中没有携带全局标签，并且该第二报文需要从第二虚拟网络拓扑转发，则该第三处理器根据该第二虚拟网络拓扑的拓扑 ID 对应的第二全局标签，由该第四处理器为该第二报文增加该第二全局标，该第二全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

10

所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

20、如权利要求 18 或 19 所述的转发设备，其特征在于，所述全局标签包括两个或两个以上的标签。

15 21、一种转发报文的方法，其特征在于，该方法应用于包括至少一个控制器和多个转发设备的网络系统，所述多个转发设备组成多个虚拟网络拓扑，

接收全局标签与所述虚拟网络拓扑的绑定关系，根据该绑定关系生成包括该全局标签与虚拟网络拓扑转发表映射关系的标签转发表项；所述绑定关系是由所述控制器为所述多个虚拟网络拓扑中的每个虚拟网络拓扑分配全局标签，并将所述全局标签与该全局标签对应的虚拟网络拓扑之间的绑定关系发送给所述多个转发设备的，所述控制器分配给所述虚拟网络拓扑中的不同虚拟网络拓扑的全局标签不同；

20

当该转发设备收到的第一报文中携带有第一全局标签时，该转发设备根据该第一全局标签查找到该第一全局标签对应的第一标签转发表项，从

25

而确定从该第一全局标签对应的第一虚拟网络拓扑的第一标签转发表项转发该第一报文；

所述第一虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

22、如权利要求21所述的方法，其特征在于，所述至少一个控制器转
5 发所述绑定关系给所述转发设备，具体为：

所述至少一个控制器通过内部网关协议IGP扩展发送所述绑定关系给
IGP域内的转发设备。

23、如权利要求 21 所述的方法，其特征在于，

10 如果所述转发设备收到的第二报文中没有携带全局标签，并且该第二
报文要从第二虚拟网络拓扑转发，则该转发设备根据该第二虚拟网络拓扑
的拓扑 ID 对应的第二全局标签，为该第二报文增加该第二全局标，该第二
全局标签在下一跳节点索引到对应该第二虚拟网络拓扑的转发表；

所述第二虚拟网络拓扑为所述多个虚拟网络拓扑中的一个。

15 24、如权利要求21-23中任一所述的方法，其特征在于，所述全局标签
包括两个或两个以上的标签。

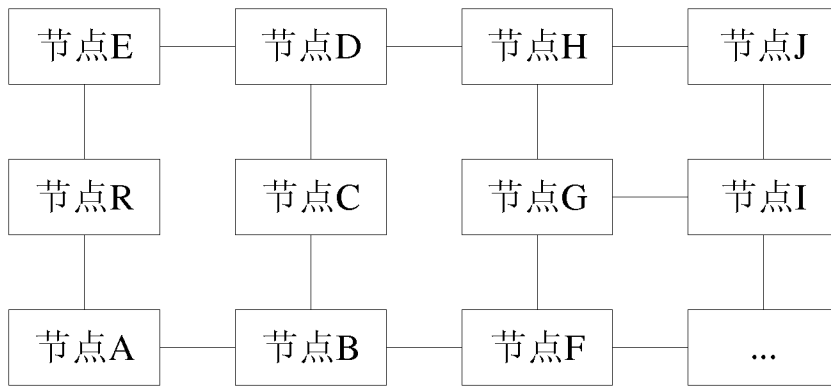


图 1A

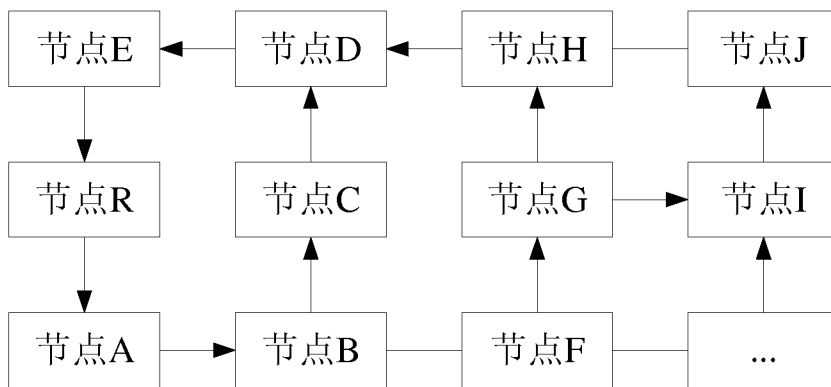


图 1B

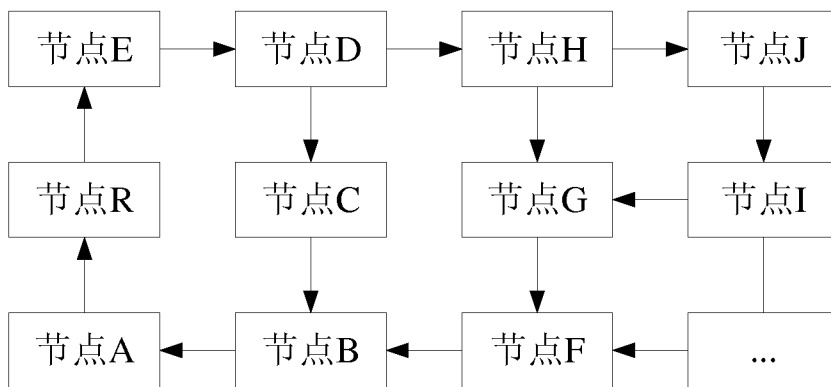


图 1C

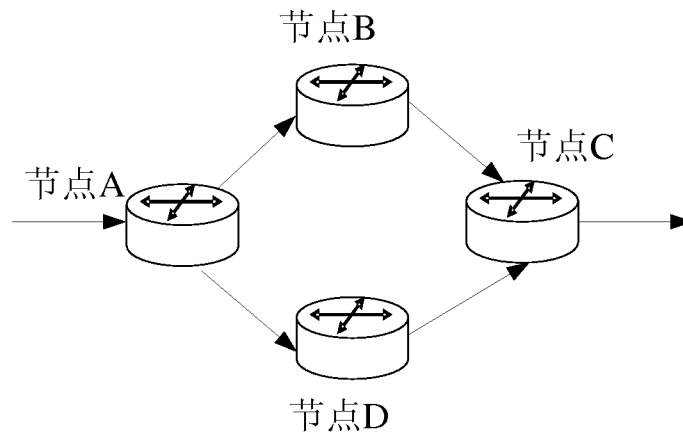


图 2A

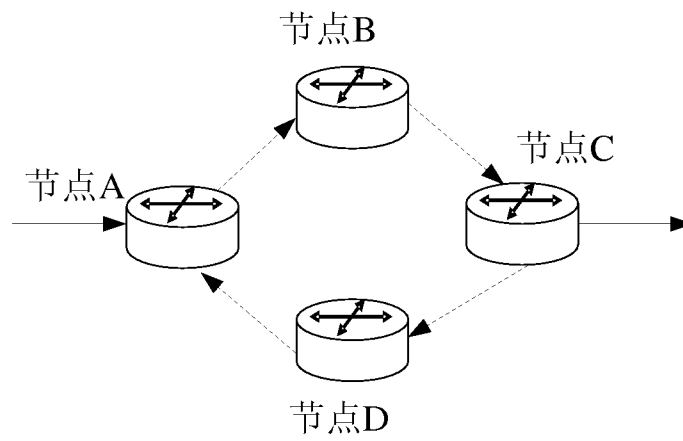


图 2B

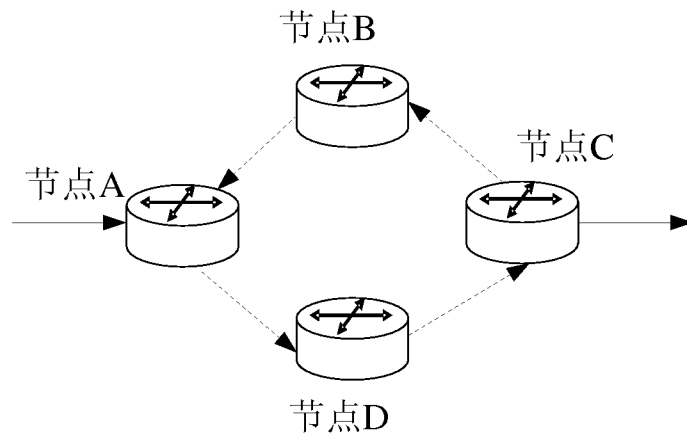


图 2C

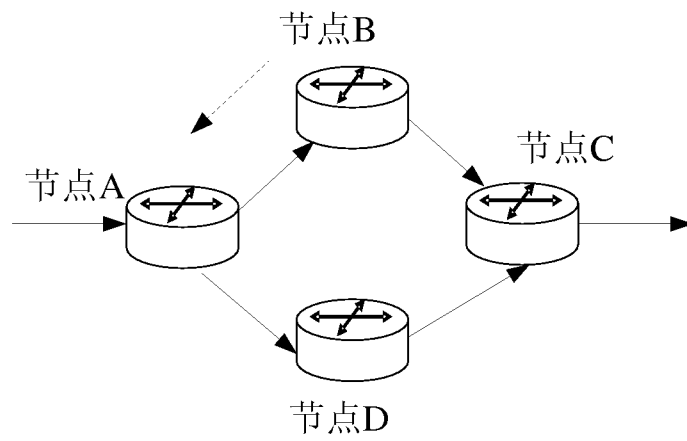


图 2D

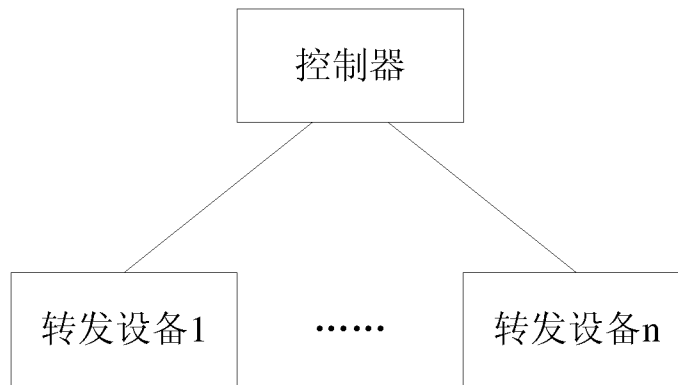


图 3

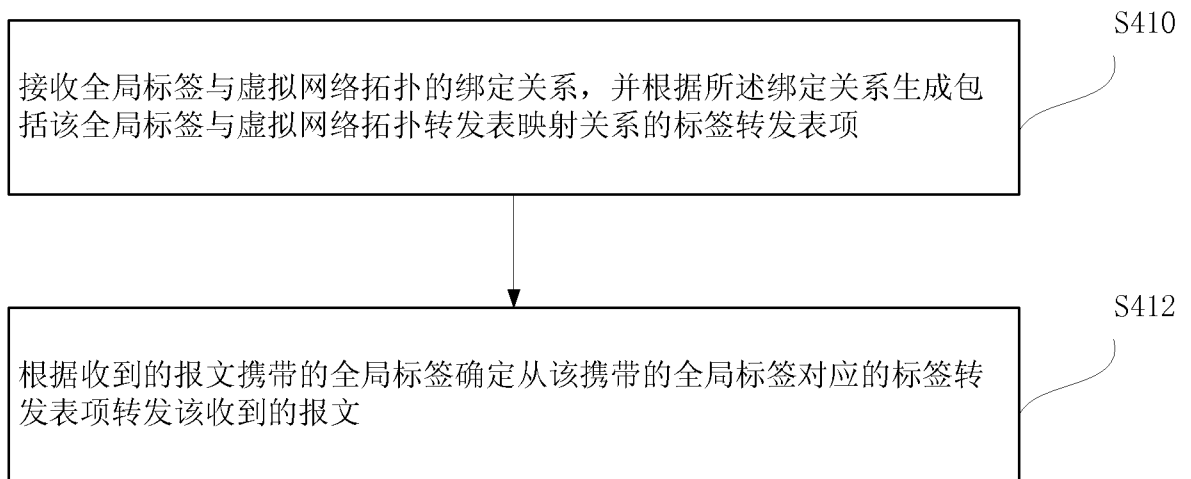


图 4



图 5

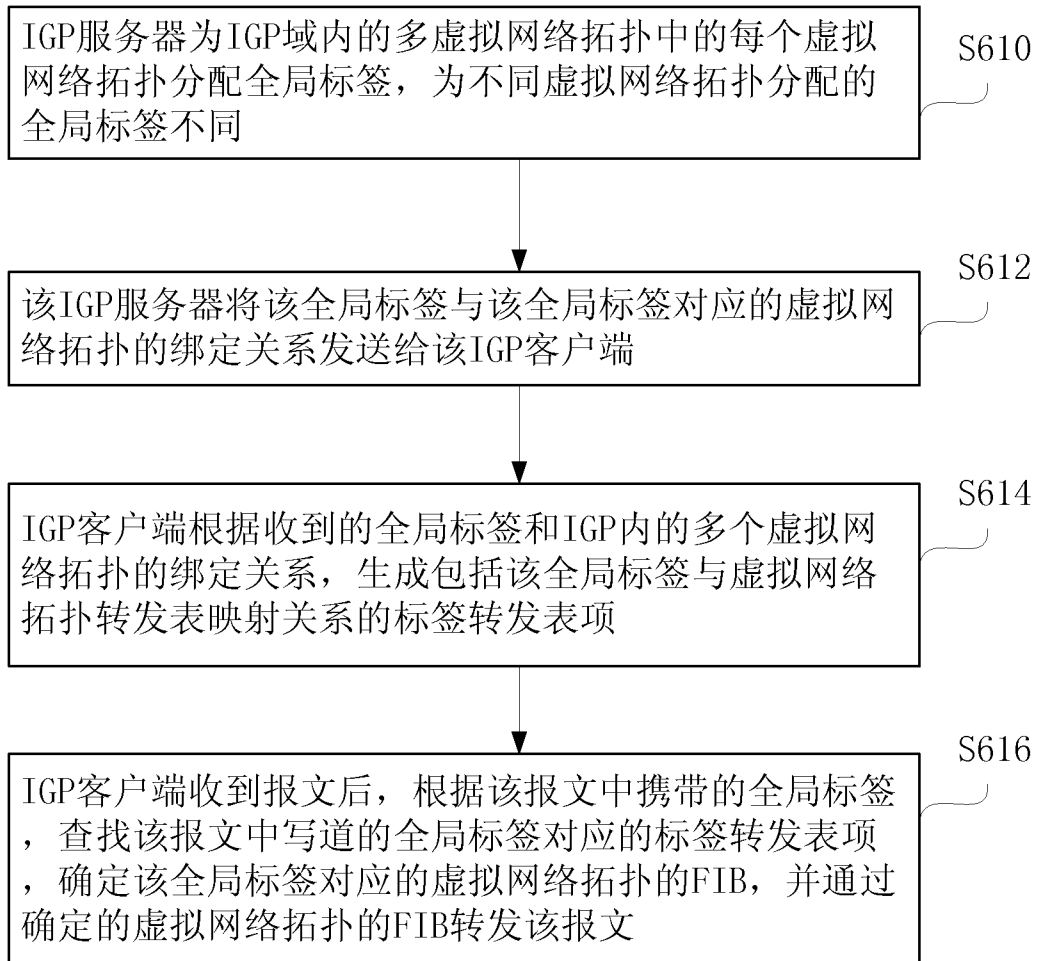


图 6

6/11

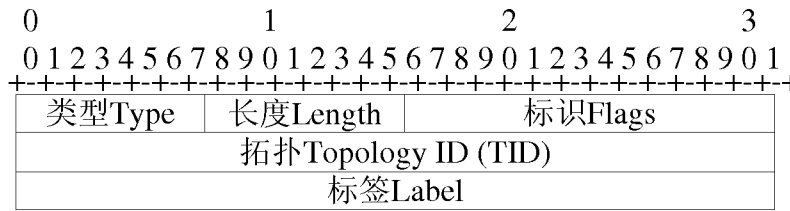


图 7A

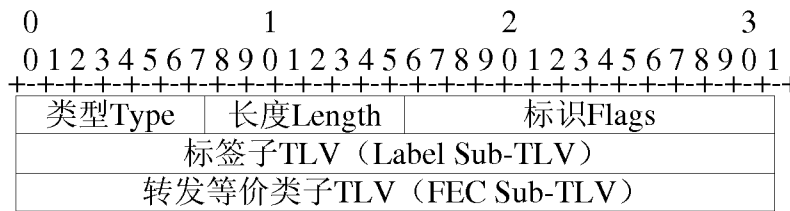


图 7B

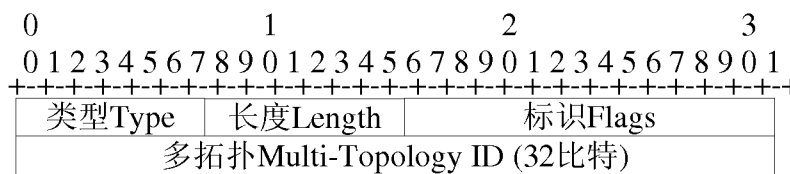


图 7C

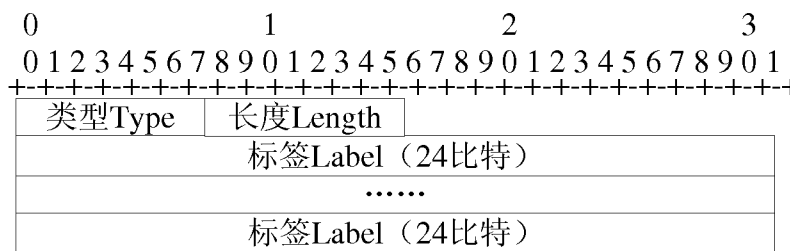


图 7D

虚拟网络拓扑转发表项	全局标签
FIB1	L1
FIB2	L2
FIB3	L3
.....
FIBn	Ln

图 8

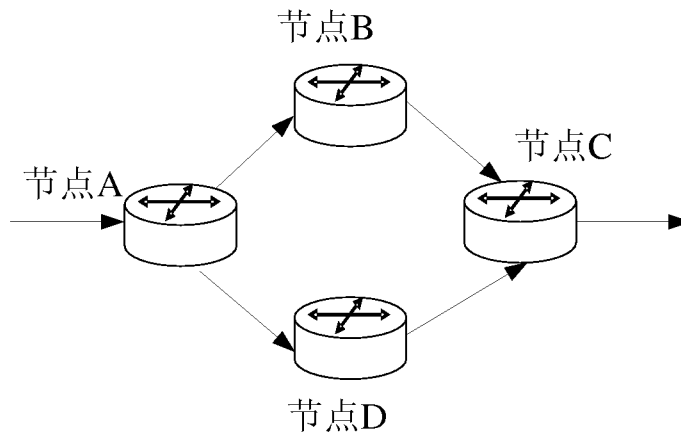


图 9A

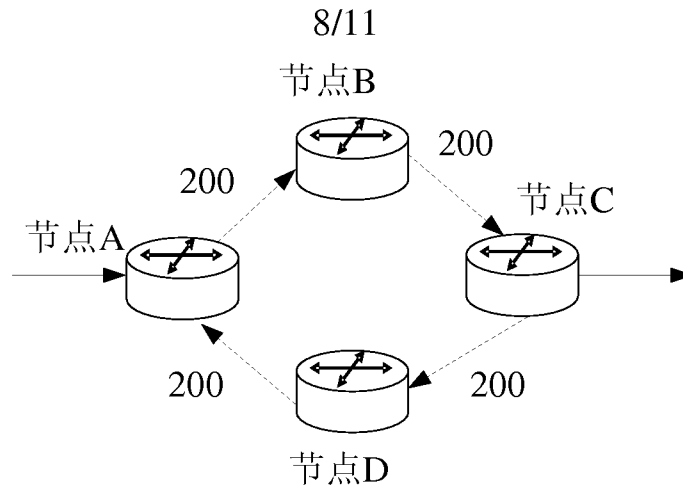


图 9B

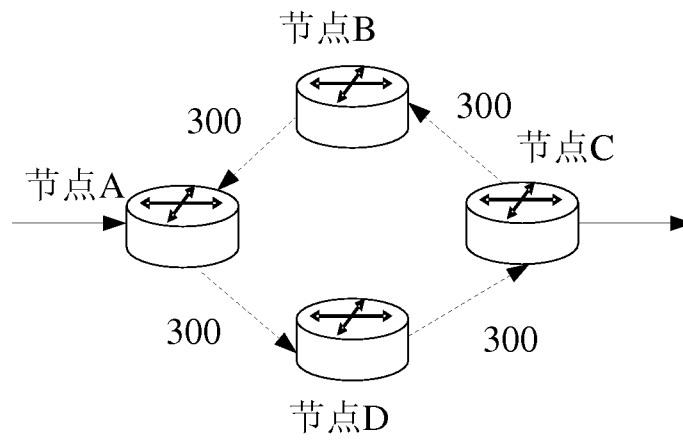


图 9C

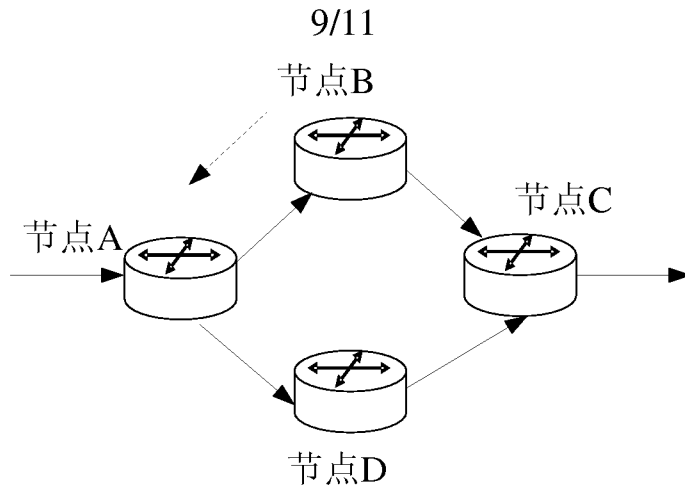


图 9D

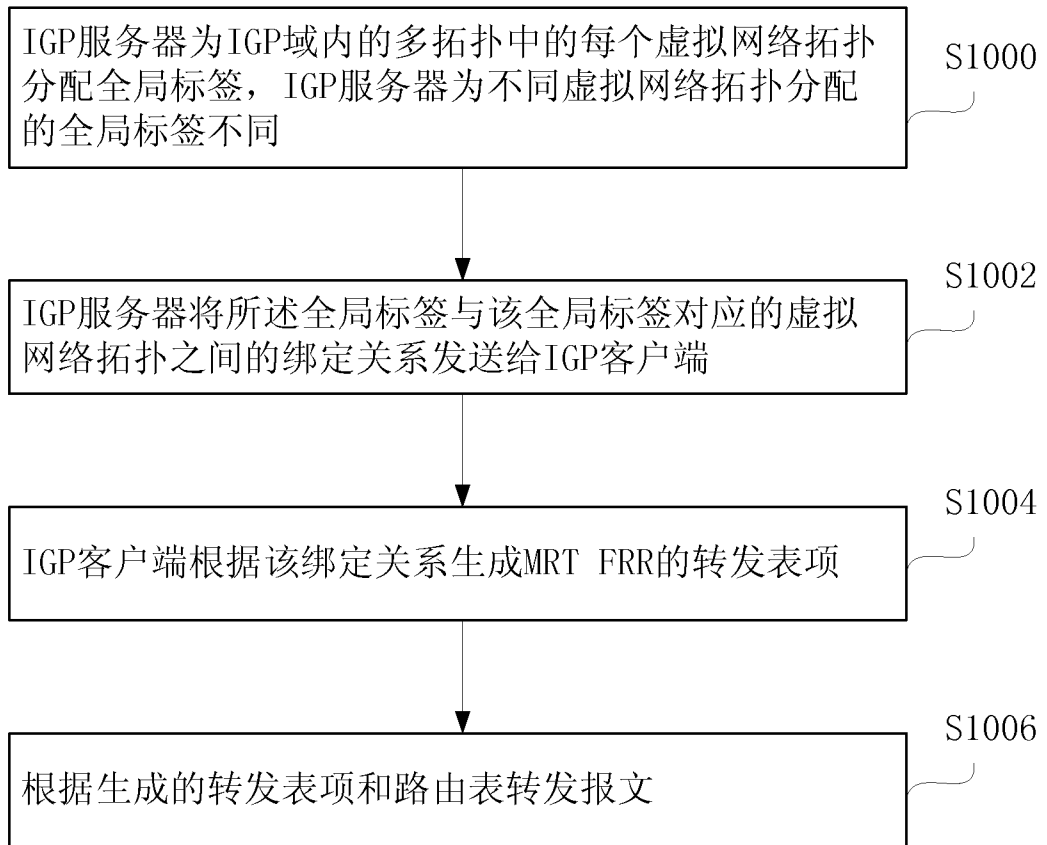


图 10

FIB	全局标签	出接口	下一跳信息
FIB ₁	L ₁	出接口 1	节点 1
FIB ₂	L ₂	出接口 2	节点 2
FIB ₃	L ₃	出接口 3	节点 3
.....
FIB _n	L _n	出接口 n	节点 n

图 11

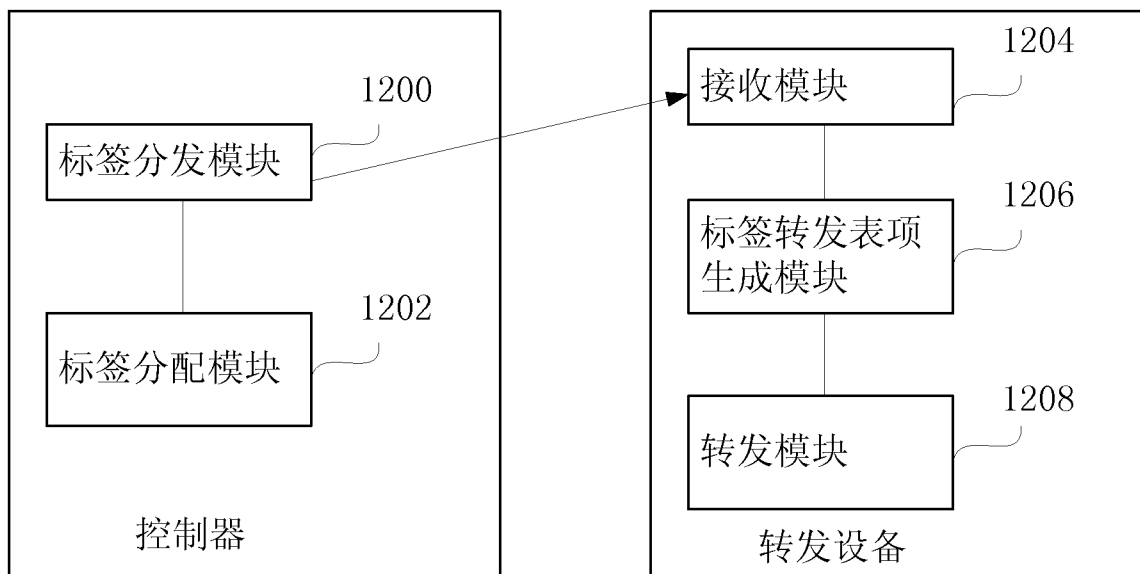


图 12

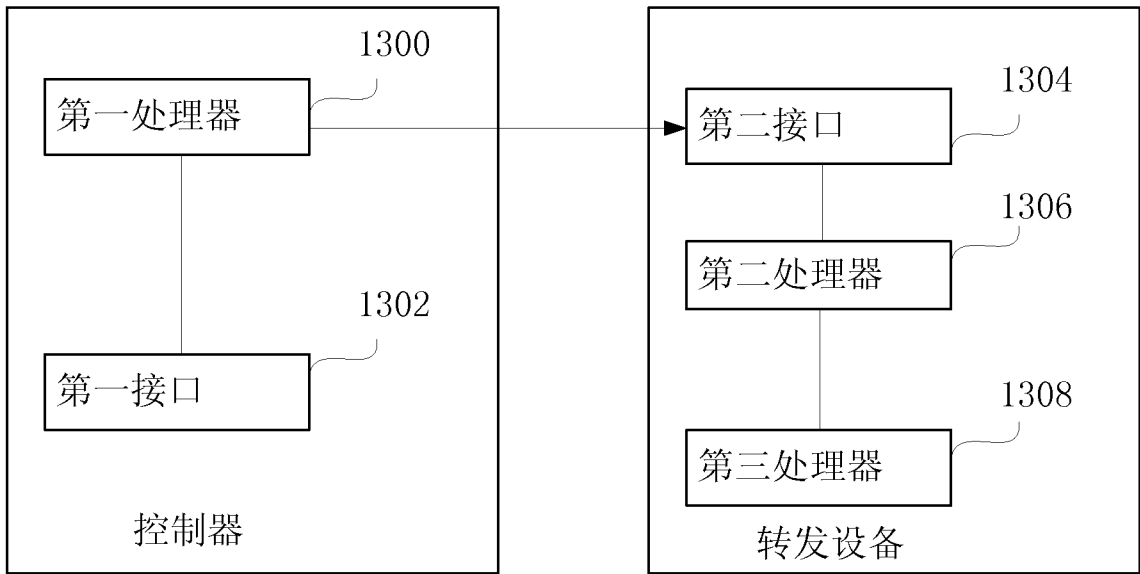


图 13

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/085496

A. CLASSIFICATION OF SUBJECT MATTER

H04L 12/28 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04L; H04W; H04Q; G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, WPI, EPODOC, IEEE, CNKI, GOOGLE: multiple, virtual, network, virtual network topology, topology, flag, identifier, mark, ID, transfer, table, route, item

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 1703016 A (MICROSOFT CORP [US]) 0 011 2005 (0.011.2005) the abstract, claims 1 to 26, page 6, line 1 to page 11 to line 19	1-24
A	CN 103348638 A (NEC CORP.) 09 October 2013 (09.10.2013) the whole document	1-24
A	CN 103297262 A (HITACHI LTD) 11 September 2013 (11.09.2013) the whole document	1-24
A	US 2012054346 A1 (FUTUREWEI TECHNOLOGIES, INC. et al.) 01 March 2012 (01.03.2012) the whole document	1-24

Further documents are listed in the continuation of Box C. See patent family annex.

<p>* Special categories of cited documents:</p> <p>“A” document defining the general state of the art which is not considered to be of particular relevance</p> <p>“E” earlier application or patent but published on or after the international filing date</p> <p>“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)</p> <p>“O” document referring to an oral disclosure, use, exhibition or other means</p> <p>“P” document published prior to the international filing date but later than the priority date claimed</p>	<p>“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention</p> <p>“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone</p> <p>“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art</p> <p>“&” document member of the same patent family</p>
---	---

Date of the actual completion of the international search 07 July 2014	Date of mailing of the international search report 16 July 2014
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer LAI, Yi Telephone No. (86-10) 62413077

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/085496

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 1703016 A	30 November 2005	AU 2004200484 A1	23 September 2004
		PL 365512 A1	20 September 2004
		CA 2456952 A1	06 September 2004
		US 2004210623 A1	21 October 2004
		EP 1455483 A2	08 September 2004
		JP 4838342 B2	14 December 2011
		EP 1455483 A3	05 January 2011
		RU 2382398 C2	20 February 2010
		JP 2011259455 A	22 December 2011
		JP 2011218405 A	27 October 2011
		JP 2004272905 A	30 September 2004
		JP 2011239452 A	24 November 2011
		MY 149582 A	13 September 2013
		RU 2004106718 A	10 August 2005
		JP 2009303275 A	24 December 2009
		BR 0400549 A	28 December 2004
		MX PA04001431 A	17 June 2005
		JP 4444695 B2	31 March 2010
		AU 2004200484 B2	11 March 2010
		BR PI0400549 A	28 December 2004
KR 101143648 B1	09 May 2012		
ZA 200400836 A	23 August 2004		
KR 20040079318 A	14 September 2004		
CN 103348638 A	09 October 2013	EP 2675119 A1	18 December 2013
		WO 2012108382 A1	16 August 2012
		US 2013308462 A1	21 November 2013
CN 103297262 A	11 September 2013	US 2014010109 A1	09 January 2014
		JP 2013175999 A	05 September 2013

INTERNATIONAL SEARCH REPORTInternational application No.
PCT/CN2013/085496

		SG 193069 A1	30 September 2013
US 2012054346 A1	01 March 2012	CN 103109505 A	15 May 2013
		WO 2012025061 A1	01 March 2012
		EP 2569963 A1	20 March 2013
		EP 2569963 A4	07 August 2013
		US 2012054347 A1	01 March 2012
		EP 2569922 A1	20 March 2013
		CN 103069783 A	24 April 2013
		WO 2012025063 A1	01 March 2012
		EP 2569922 A4	07 August 2013

国际检索报告

国际申请号

PCT/CN2013/085496

<p>A. 主题的分类</p> <p>H04L 12/28 (2006.01)i</p> <p>按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类</p>																											
<p>B. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>H04L; H04W; H04Q; G06F</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))</p> <p>CNPAT, WPI, EPDOC, IEEE, CNKI, GOOGLE: 多个, 虚拟, 网络, 虚拟网络拓扑, 拓扑, 标签, 标识, 标记, 标识符, 转发, 表, 路由, 表项, multiple, virtual, network, topology, flag, identifier, mark, ID, transfer, table, route, item</p>																											
<p>C. 相关文件</p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>X</td> <td>CN 1703016A (微软公司) 2005年 11月 30日 (2005 - 11 - 30) 摘要, 权利要求1-26, 说明书第6页第1行-第11页第19行</td> <td>1-24</td> </tr> <tr> <td>A</td> <td>CN 103348638A (日本电气株式会社) 2013年 10月 09日 (2013 - 10 - 09) 全文</td> <td>1-24</td> </tr> <tr> <td>A</td> <td>CN 103297262A (株式会社日立制作所) 2013年 9月 11日 (2013 - 09 - 11) 全文</td> <td>1-24</td> </tr> <tr> <td>A</td> <td>US 2012054346A1 (FUTUREWEI TECHNOLOGIES, INC.等) 2012年 3月 01日 (2012 - 03 - 01) 全文</td> <td>1-24</td> </tr> </tbody> </table> <p><input type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p> <p>* 引用文件的具体类型:</p> <table border="0"> <tr> <td>“A” 认为不特别相关的表示了现有技术一般状态的文件</td> <td>“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件</td> </tr> <tr> <td>“E” 在国际申请日的当天或之后公布的在先申请或专利</td> <td>“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性</td> </tr> <tr> <td>“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)</td> <td>“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性</td> </tr> <tr> <td>“O” 涉及口头公开、使用、展览或其他方式公开的文件</td> <td>“&” 同族专利的文件</td> </tr> <tr> <td>“P” 公布日先于国际申请日但迟于所要求的优先权日的文件</td> <td></td> </tr> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	X	CN 1703016A (微软公司) 2005年 11月 30日 (2005 - 11 - 30) 摘要, 权利要求1-26, 说明书第6页第1行-第11页第19行	1-24	A	CN 103348638A (日本电气株式会社) 2013年 10月 09日 (2013 - 10 - 09) 全文	1-24	A	CN 103297262A (株式会社日立制作所) 2013年 9月 11日 (2013 - 09 - 11) 全文	1-24	A	US 2012054346A1 (FUTUREWEI TECHNOLOGIES, INC.等) 2012年 3月 01日 (2012 - 03 - 01) 全文	1-24	“A” 认为不特别相关的表示了现有技术一般状态的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件	“E” 在国际申请日的当天或之后公布的在先申请或专利	“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性	“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)	“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性	“O” 涉及口头公开、使用、展览或其他方式公开的文件	“&” 同族专利的文件	“P” 公布日先于国际申请日但迟于所要求的优先权日的文件	
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																									
X	CN 1703016A (微软公司) 2005年 11月 30日 (2005 - 11 - 30) 摘要, 权利要求1-26, 说明书第6页第1行-第11页第19行	1-24																									
A	CN 103348638A (日本电气株式会社) 2013年 10月 09日 (2013 - 10 - 09) 全文	1-24																									
A	CN 103297262A (株式会社日立制作所) 2013年 9月 11日 (2013 - 09 - 11) 全文	1-24																									
A	US 2012054346A1 (FUTUREWEI TECHNOLOGIES, INC.等) 2012年 3月 01日 (2012 - 03 - 01) 全文	1-24																									
“A” 认为不特别相关的表示了现有技术一般状态的文件	“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件																										
“E” 在国际申请日的当天或之后公布的在先申请或专利	“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性																										
“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)	“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性																										
“O” 涉及口头公开、使用、展览或其他方式公开的文件	“&” 同族专利的文件																										
“P” 公布日先于国际申请日但迟于所要求的优先权日的文件																											
<p>国际检索实际完成的日期</p> <p>2014年 7月 07日</p>	<p>国际检索报告邮寄日期</p> <p>2014年 7月 16日</p>																										
<p>ISA/CN的名称和邮寄地址</p> <p>中华人民共和国国家知识产权局(ISA/CN) 北京市海淀区蓟门桥西土城路6号 100088 中国</p> <p>传真号 (86-10)62019451</p>	<p>受权官员</p> <p>赖异</p> <p>电话号码 (86-10)62413077</p>																										

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2013/085496

检索报告引用的专利文件	公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN 1703016A	2005年 11月 30日	AU 2004200484A1	2004年 9月 23日
		PL 365512A1	2004年 9月 20日
		CA 2456952A1	2004年 9月 06日
		US 2004210623A1	2004年 10月 21日
		EP 1455483A2	2004年 9月 08日
		JP 4838342B2	2011年 12月 14日
		EP 1455483A3	2011年 1月 05日
		RU 2382398C2	2010年 2月 20日
		JP 2011259455A	2011年 12月 22日
		JP 2011217405A	2011年 10月 27日
		JP 2004272905A	2004年 9月 30日
		JP 2011239452A	2011年 11月 24日
		MY 149582A	2013年 9月 13日
		RU 2004106718A	2005年 8月 10日
		JP 2009303275A	2009年 12月 24日
		BR 0400549A	2004年 12月 28日
		MX PA04001431A	2005年 6月 17日
		JP 4444695B2	2010年 3月 31日
		AU 2004200484B2	2010年 3月 11日
		BR PI0400549A	2004年 12月 28日
KR 101143648B1	2012年 5月 09日		
ZA 200400836A	2004年 8月 23日		
KR 20040079318A	2004年 9月 14日		
CN 103348638A	2013年 10月 09日	EP 2675119A1	2013年 12月 18日
		WO 2012108382A1	2012年 8月 16日
		US 2013308462A1	2013年 11月 21日
CN 103297262A	2013年 9月 11日	US 2014010109A1	2014年 1月 09日
		JP 2013175999A	2013年 9月 05日
		SG 193069A1	2013年 9月 30日
US 2012054346A1	2012年 3月 01日	CN 103109505A	2013年 5月 15日
		WO 2012025061A1	2012年 3月 01日
		EP 2569963A1	2013年 3月 20日
		EP 2569963A4	2013年 8月 07日
		US 2012054347A1	2012年 3月 01日
		EP 2569922A1	2013年 3月 20日
		CN 103069783A	2013年 4月 24日
		WO 2012025063A1	2012年 3月 01日
EP 2569922A4	2013年 8月 07日		

表 PCT/ISA/210 (同族专利附件) (2009年7月)