



(12) 发明专利申请

(10) 申请公布号 CN 105337867 A

(43) 申请公布日 2016. 02. 17

(21) 申请号 201410370601.1

(22) 申请日 2014.07.30

(71) 申请人 国际商业机器公司
地址 美国纽约

(72) 发明人 聂玖星 李福辉 姚军 陆静

(74) 专利代理机构 北京市中咨律师事务所
11247

代理人 于静 张亚非

(51) Int. Cl.

H04L 12/741(2013.01)

H04L 12/701(2013.01)

H04L 12/767(2013.01)

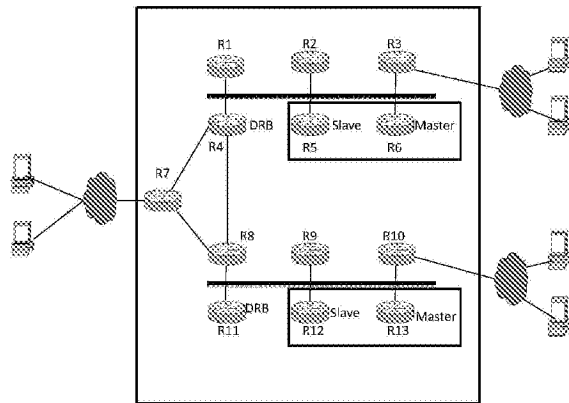
权利要求书2页 说明书9页 附图6页

(54) 发明名称

TRILL 网络中分发非单播路由信息的方法和对应的装置

(57) 摘要

本发明公开了一种在 TRILL 网络中处理非单播路由信息的方法及对应的 RBridge，其中该 RBridge 为与其它 RBridge 在该 TRILL 网络中具有邻居关系，并且该 RBridge 为所述邻居关系中的计算非单播路由信息的 RBridge，该方法中该 RBridge 执行如下步骤：向所述其它 RBridge 发送确认消息，其中所述确认消息确认该 RBridge 为计算非单播路由信息的 RBridge；获得该 TRILL 网络中与所述非单播路由信息相关的网络拓扑；根据所述网络拓扑计算所述非单播路由信息；以及将算出的所述非单播路由信息分发到所述其它 RBridge。该方法和对应的 RBridge 能够减少 TRILL 网络中的 RBridge 的 CPU 资源的消耗。



1. 一种在多链接透明互联 TRILL 网络中处理非单播路由信息的方法,其中具有路由转发特性的网桥 RBridge 与其它 RBridge 在该 TRILL 网络中具有邻居关系,并且该 RBridge 为所述邻居关系中的计算非单播路由信息的 RBridge,该方法中该 RBridge 执行如下步骤:

向所述其它 RBridge 发送确认消息,其中所述确认消息确认该 RBridge 为计算非单播路由信息的 RBridge;

获得该 TRILL 网络中与所述非单播路由信息相关的网络拓扑;

根据所述网络拓扑计算所述非单播路由信息;以及

将算出的所述非单播路由信息分发到所述其它 RBridge。

2. 根据权利要求 1 所述的方法,其中所述其它 RBridge 不计算非单播路由信息。

3. 根据权利要求 1 或 2 所述的方法,其中该 RBridge 与所述其它 RBridge 位于该 TRILL 网络的一个链路 Link 上。

4. 根据权利要求 1 所述的方法,其中该 RBridge 是被所述邻居关系中指定的具有路由转发特性的网桥 DRB 指定为计算非单播路由信息的 RBridge,并且该指定被所述 DRB 编码在报文中传播到所述其它 RBridge。

5. 根据权利要求 1 所述的方法,其中该 RBridge 是通过接收命令被指定为计算非单播路由信息的 RBridge,并且该指定被该 RBridge 编码在报文中传播到所述其它 RBridge。

6. 根据权利要求 4 或 5 所述的方法,其中该 RBridge 不是边界 RBridge。

7. 根据权利要求 4-6 之一所述的方法,其中该 TRILL 网络中,响应于该邻居关系中包括至少 3 个 RBridge,则所述邻居关系中包括的该 RBridge 为主计算 RBridge,所述邻居关系中还包括从计算 RBridge 和所述其它 RBridge,该从计算 RBridge 计算所述非单播路由信息,并且在主计算 RBridge 工作的情况下,不向所述其它 RBridge 发送非单播路由信息。

8. 根据权利要求 7 所述的方法,其中该 TRILL 网络中,响应于该从计算 RBridge 检测到该主计算 RBridge 不可用,该从计算 RBridge 向所述其它 RBridge 发送非单播路由信息。

9. 根据权利要求 8 所述的方法,其中该 TRILL 网络中,该主计算 RBridge 和该从计算 RBridge 的可用性都被检测,响应于仅该主计算 RBridge 或仅从计算 RBridge 不可用,指定新的从计算 RBridge,响应于该主计算 RBridge 和该从计算 RBridge 都不可用,指定新的主计算 RBridge 和新的从计算 RBridge。

10. 一种在多链接透明互联 TRILL 网络中处理非单播路由信息的具有路由转发特性的网桥 RBridge,其中该 RBridge 与其它 RBridge 在该 TRILL 网络中具有邻居关系,并且该 RBridge 为所述邻居关系中的计算非单播路由信息的 RBridge,该 RBridge 包括:

确认消息发送装置,被配置为向所述其它 RBridge 发送确认消息,其中所述确认消息确认该 RBridge 为计算非单播路由信息的 RBridge;

获得装置,被配置为获得该 TRILL 网络中与所述非单播路由信息相关的网络拓扑;

计算装置,被配置为根据所述网络拓扑计算所述非单播路由信息;以及

分发装置,被配置为将算出的所述非单播路由信息分发到所述其它 RBridge。

11. 根据权利要求 10 所述的 RBridge,其中所述其它 RBridge 不计算非单播路由信息。

12. 根据权利要求 10 或 11 所述的 RBridge,其中该 RBridge 与所述其它 RBridge 位于该 TRILL 网络的一个链路 Link 上。

13. 根据权利要求 10 所述的 RBridge,其中该 RBridge 是被所述邻居关系中指定的具

有路由转发特性的网桥 DRB 指定为计算非单播路由信息的 RBridge, 并且该指定被所述 DRB 编码在报文中传播到所述其它 RBridge。

14. 根据权利要求 10 所述的 RBridge, 其中该 RBridge 是通过接收命令被指定为计算非单播路由信息的 RBridge, 并且该指定被该 RBridge 编码在报文中传播到所述其它 RBridge。

15. 根据权利要求 13 或 14 所述的 RBridge, 其中该 RBridge 不是边界 RBridge。

16. 根据权利要求 13-15 之一所述的 RBridge, 其中该 TRILL 网络中, 响应于该邻居关系中包括至少 3 个 RBridge, 则所述邻居关系中包括的该 RBridge 为主计算 RBridge, 所述邻居关系中还包括从计算 RBridge 和所述其它 RBridge, 该从计算 RBridge 计算所述非单播路由信息, 并且在该主计算 RBridge 工作的情况下, 不向所述其它 RBridge 发送非单播路由信息。

17. 根据权利要求 16 所述的 RBridge, 其中该 TRILL 网络中, 响应于该从计算 RBridge 检测到该主计算 RBridge 不可用, 该从计算 RBridge 向所述其它 RBridge 发送非单播路由信息。

18. 根据权利要求 17 所述的 RBridge, 其中该 TRILL 网络中, 该主计算 RBridge 和该从计算 RBridge 的可用性都被检测, 响应于仅该主计算 RBridge 或仅从计算 RBridge 不可用, 指定新的从计算 RBridge, 响应于该主计算 RBridge 和该从计算 RBridge 都不可用, 指定新的主计算 RBridge 和新的从计算 RBridge。

TRILL 网络中分发非单播路由信息的方法和对应的装置

技术领域

[0001] 本发明涉及通信网络,更具体地,涉及一种 TRILL 网络中分发非单播路由信息的方法和对应的 RBridge。

背景技术

[0002] 多链接透明互联 TRILL (Transparent Interconnection of Lots of Links) 是一种在二层网络上基于链路状态算法的路由协议,使用该协议组建的 TRILL 网络由具有路由转发特性的网桥 RBridge (Routing Bridge) 组成,每个 RBridge 通过运行扩展的中间系统到中间系统的路由选择协议 IS-IS 协议 (Intermediate System to Intermediate System Routing Protocol) 获取整个网络拓扑,并且通过最短路径优先算法 SPF (Short Path First) 算法计算路由信息。

[0003] 图 2 示出了一个示意性的 TRILL 网络的结构,根据图 2,其通信的基本过程是:每个 RBridge 通过 SPF 算法计算出的路由信息组成了路由表,以备报文转发时的需要。当一个 RBridge RB-A 从主机 Host1 收到一个数据包且需要通过 TRILL 网络转发到主机 Host 2 时,将该报文封装一个 TRILL 报头,TRILL 报头中包含进入 TRILL 网络的入口 RBridge (Ingress RBridge),即图 2 中的 RB-A,和离开 TRILL 网络的出口 RBridge (Egress RBridge),即图 2 中的 RB-C,也称目的 RBridge,用于在 TRILL 网络中查找路由。TRILL 报文在 RB-A 根据路由信息被转发到 RB-B 后,在 RB-B 根据路由信息被进一步转发到 RB-C;在 RB-C 发现本地 RBridge 与出口 RBridge 相同,确定 TRILL 报文已经到达转发路径上最后一跳的 RBridge,此时需要剥离 TRILL 报头,执行传统二层转发,最终将数据包转发给主机 Host 2。

[0004] 上述通信过程中,在建立如图 2 所示的 TRILL 网络的物理连接后,每个 RBridge 都要获取全网的链路状态信息,得到网络拓扑,并且通过 SPF 算法计算路由信息;当 TRILL 网络结构发生变化时,每个 RBridge 都要重新获取全网的网络拓扑,并且通过 SPF 算法重新计算路由信息;甚至,即使网络没有变化,每个 RBridge 都要定期通过 SPF 算法重新计算路由信息。这些计算将消耗 RBridge 的很多 CPU 资源,可能导致运行在 RBridge 的各种业务受到影响。

发明内容

[0005] 根据本发明的一个方面,提供了一种在多链接透明互联 TRILL 网络中处理非单播路由信息的方法,其中具有路由转发特性的网桥 RBridge 与其它 RBridge 在该 TRILL 网络中具有邻居关系,并且该 RBridge 为所述邻居关系中的计算非单播路由信息的 RBridge,该方法中该 RBridge 执行如下步骤:

[0006] 向所述其它 RBridge 发送确认消息,其中所述确认消息确认该 RBridge 为计算非单播路由信息的 RBridge;

[0007] 获得该 TRILL 网络中与所述非单播路由信息相关的网络拓扑;

[0008] 根据所述网络拓扑计算所述非单播路由信息;以及

- [0009] 将算出的所述非单播路由信息分发到所述其它 RBridge。
- [0010] 根据本发明的另一个方面,提供了一种在多链接透明互联 TRILL 网络中处理非单播路由信息的具有路由转发特性的网桥 RBridge,其中该 RBridge 与其它 RBridge 在该 TRILL 网络中具有邻居关系,并且该 RBridge 为所述邻居关系中的计算非单播路由信息的 RBridge,该 RBridge 包括:
- [0011] 确认消息发送装置,被配置为向所述其它 RBridge 发送确认消息,其中所述确认消息确认该 RBridge 为计算非单播路由信息的 RBridge;
- [0012] 获得装置,被配置为获得该 TRILL 网络中与所述非单播路由信息相关的网络拓扑;
- [0013] 计算装置,被配置为根据所述网络拓扑计算所述非单播路由信息;以及
- [0014] 分发装置,被配置为将算出的所述非单播路由信息分发到所述其它 RBridge。
- [0015] 本发明提供的方法和和设备能够减少 TRILL 网络中 RBridge CPU 资源的消耗。

附图说明

- [0016] 通过结合附图对本公开示例性实施方式进行更详细的描述,本公开的上述以及其它目的、特征和优势将变得更加明显,其中,在本公开示例性实施方式中,相同的参考标号通常代表相同部件。
- [0017] 图 1 示出了适于用来实现本发明实施方式的示例性计算机系统 / 服务器 12 的框图;
- [0018] 图 2 示出了一个示意性的 TRILL 网络的结构图;
- [0019] 图 3 示出了在 TRILL 网络中处理非单播路由信息方法的方法流程图;
- [0020] 图 4 示出了一种 TRILL 网络的多个 RBridge 组成邻居关系的示意图;
- [0021] 图 5 示出了一种将指定的计算 RBridge 传播到邻居关系中的其它 RBridge 的 TLV 格式;
- [0022] 图 6 示出了图 4 中 R6 的通知 TLV 消息内容;
- [0023] 图 7 示出了计算 RBridge 被指定为计算 RBridge 后发送的确认 TLV 格式;
- [0024] 图 8 示出了 R6 发送的确认消息;
- [0025] 图 9 示出了 R5 发送的确认消息;
- [0026] 图 10 示出了作为分发非单播路由信息的消息格式的 LSP 报文中的 TLV 格式;
- [0027] 图 11 示出了图 10 中的邻居信息包含的内容;
- [0028] 图 12 示出了 R6 发送的路由信息;以及
- [0029] 图 13 示出了一种 TRILL 网络中 RBridge 的结构框图。

具体实施方式

[0030] 下面将参照附图更详细地描述本公开的优选实施方式。虽然附图中显示了本公开的优选实施方式,然而应该理解,可以以各种形式实现本公开而不应被这里阐述的实施方式所限制。相反,提供这些实施方式是为了使本公开更加透彻和完整,并且能够将本公开的范围完整地传达给本领域的技术人员。

- [0031] 图 1 示出了适于用来实现本发明实施方式的示例性计算机系统 / 服务器 12 的框

图。图 1 显示的计算机系统 / 服务器 12 仅仅是一个示例,不应对本发明实施例的功能和使用范围带来任何限制。

[0032] 如图 1 所示,计算机系统 / 服务器 12 以通用计算设备的形式表现。计算机系统 / 服务器 12 的组件可以包括但不限于:一个或者多个处理器或者处理单元 16,系统存储器 28,连接不同系统组件(包括系统存储器 28 和处理单元 16)的总线 18。

[0033] 总线 18 表示几类总线结构中的一种或多种,包括存储器总线或者存储器控制器,外围总线,图形加速端口,处理器或者使用多种总线结构中的任意总线结构的局域总线。举例来说,这些体系结构包括但不限于工业标准体系结构 (ISA) 总线,微通道体系结构 (MAC) 总线,增强型 ISA 总线、视频电子标准协会 (VESA) 局域总线以及外围组件互连 (PCI) 总线。

[0034] 计算机系统 / 服务器 12 典型地包括多种计算机系统可读介质。这些介质可以是任何能够被计算机系统 / 服务器 12 访问的可用介质,包括易失性和非易失性介质,可移动的和不可移动的介质。

[0035] 系统存储器 28 可以包括易失性存储器形式的计算机系统可读介质,例如随机存取存储器 (RAM) 30 和 / 或高速缓存存储器 32。计算机系统 / 服务器 12 可以进一步包括其它可移动 / 不可移动的、易失性 / 非易失性计算机系统存储介质。仅作为举例,存储系统 34 可以用于读写不可移动的、非易失性磁介质(图 1 未显示,通常称为“硬盘驱动器”)。尽管图 1 中未示出,可以提供用于对可移动非易失性磁盘(例如“软盘”)读写的磁盘驱动器,以及对可移动非易失性光盘(例如 CD-ROM, DVD-ROM 或者其它光介质)读写的光盘驱动器。在这些情况下,每个驱动器可以通过一个或者多个数据介质接口与总线 18 相连。存储器 28 可以包括至少一个程序产品,该程序产品具有一组(例如至少一个)程序模块,这些程序模块被配置以执行本发明各实施例的功能。

[0036] 具有一组(至少一个)程序模块 42 的程序 / 实用工具 40,可以存储在例如存储器 28 中,这样的程序模块 42 包括——但不限于——操作系统、一个或者多个应用程序、其它程序模块以及程序数据,这些示例中的每一个或某种组合中可能包括网络环境的实现。程序模块 42 通常执行本发明所描述的实施例中的功能和 / 或方法。

[0037] 计算机系统 / 服务器 12 也可以与一个或多个外部设备 14(例如键盘、指向设备、显示器 24 等)通信,还可与一个或者多个使得用户能与该计算机系统 / 服务器 12 交互的设备通信,和 / 或与使得该计算机系统 / 服务器 12 能与一个或多个其它计算设备进行通信的任何设备(例如网卡,调制解调器等等)通信。这种通信可以通过输入 / 输出 (I/O) 接口 22 进行。并且,计算机系统 / 服务器 12 还可以通过网络适配器 20 与一个或者多个网络(例如局域网 (LAN),广域网 (WAN) 和 / 或公共网络,例如因特网)通信。如图所示,网络适配器 20 通过总线 18 与计算机系统 / 服务器 12 的其它模块通信。应当明白,尽管图中未示出,可以结合计算机系统 / 服务器 12 使用其它硬件和 / 或软件模块,包括但不限于:微代码、设备驱动器、冗余处理单元、外部磁盘驱动阵列、RAID 系统、磁带驱动器以及数据备份存储系统等。

[0038] SPF 算法是一种通用的计算路由信息的算法,该算法需要算出是从本节点到目的节点有哪些路可以走,代价是多少,可能存在多条路径;然后在这些多条路径信息中择优,选择一个最佳的出接口和下一跳,作为路由信息。可见,SPF 计算路由是不区分流量类型的。

[0039] 在 TRILL 网络使用 SPF 算法计算路由过程中,对于不同的流量类型,会计算出不同

同的路由表,即对于单播流量类型可以得到单播路由表、对组播和广播流量类型,会得到非单播路由表。例如计算单播路由表时,每个 RBridge 是针对将自己作为根节点组成的网络拓扑使用 SPF 算法计算得到路由信息的,因此,每个 RBridge 计算中使用的拓扑网络是不同的,因此各 RBridge 得到的单播路由表也是不同的。但是对于非单播路由表的计算,每个 RBridge 使用的网络拓扑的根节点都是选举出来的若干 RBridge,网络拓扑是这些根节点为根的若干棵树,因此,在选举所覆盖的所有 RBridge 范围内,其使用的网络拓扑是相同的,计算出的路由信息也是相同的,因此,没有必要如现有技术那样,让每一个 RBridge 都计算出若干个非单播路由信息为组播和广播服务,可以只由特定的 RBridge 计算非单播路由信息,然后通知选举范围内的所有 RBridge,这就是本发明的基本思想。通过这样计算组播和广播路由信息,可以使得相当多的 RBridge 省却非单播路由信息的计算,将节省大量的 CPU 资源。

[0040] 根据本发明的一种实施方式,公开了一种在多链接透明互联 TRILL 网络中处理非单播路由信息的方法,其中具有路由转发特性的网桥 RBridge 与其它 RBridge 在该 TRILL 网络中具有邻居关系,并且该 RBridge 为所述邻居关系中的计算非单播路由信息的 RBridge,图 3 示出了在 TRILL 网络中处理非单播路由信息的方法的方法流程,根据图 3,该方法中该 RBridge 执行如下步骤:

[0041] 在步骤 S301,向所述其它 RBridge 发送确认消息,其中所述确认消息确认该 RBridge 为计算非单播路由信息的 RBridge;

[0042] 在步骤 S302,获得该 TRILL 网络中与所述非单播路由信息相关的网络拓扑;

[0043] 在步骤 S303,根据所述网络拓扑计算所述非单播路由信息;以及

[0044] 在步骤 S304,将算出的所述非单播路由信息分发到所述其它 RBridge。

[0045] 上述方法中,所述与该 RBridge 具有邻居关系的所述其它 RBridge 就可以从该 RBridge 中获取非单播路由信息,就不必计算非单播路由信息,就可以节省其 CPU 资源。

[0046] 在一种实施方式中,邻居关系可以根据 TRILL 网络的连接关系进行定义,并且同一个网络可以定义不同的邻居关系。图 4 示出了一种 TRILL 网络的多个 RBridge 组成邻居关系的示意图,在图 4 中,R1、R2、R3、R4、R5 和 R6 是位于同一个链路(link)上,在一种实施方式中,具有邻居关系的 RBridge 位于 TRILL 网络的一个链路(Link)上。也就是说,R1-R6 组成邻居关系。同样,R8-R13 也组成邻居关系。当然,邻居关系也可以由网络管理员定义,例如,可以将 R1-R13 定义成邻居关系。

[0047] 在现有技术中,邻居关系中的 RBridge 是通过交互 hello 报文来发现 TRILL 网络中的邻居 RBridge 的,每个 RBridge 都知道它是同一个 link 上所有其他 RBridge 的邻居。Hello 报文主要用于 TRILL 邻居的发现和保活,它通常包含形成邻居的一些必备和可选条件,所有的信息都使用 TLV 形式编码在 Hello 报文中。所有使能了 TRILL 的接口的 RBridge 都会发送 Hello 报文。邻居之间会定期互相发送 Hello 报文。TLV 是一种在报文中表达信息的方式,TLV 分别代表类型 (Type),长度 (Length),值 (Value)。发送方和接收方约定报文中的信息是以 TLV 形式编码的,接收方收到报文后将以 TLV 的格式解析报文。

[0048] 在一种实施方式中,该 RBridge 是由邻居关系中的指定的具有路由转发特性的网桥 (Designated RBridge,以下简称 DRB,) 指定为计算非单播路由信息的具有路由转发特性的网桥 (以下简称计算 RBridge) 的。在现有技术中,具有邻居关系的多个 RBridge 中,可

以通过选举来获得 DRB。然后在本发明中由选举的 DRB 来指定邻居关系中的计算非单播路由信息的 RBridge。具体如何选举在 RFC 中已经规定,就是通过比较各个 RBridge 接口优先级,将接口优先级高的 RBridge 选举为 DRB,这属于现有技术,这里不再赘述。以下描述中,假设 R1-R6 的链路组成邻居关系中,R6 被选举为 DRB。在另一种实施方式中,具有邻居关系的多个 RBridge 中,也可以通过系统管理员指定 DRB。然后由 DRB 来指定计算 RBridge,指定的方式与上述方式相同。另外,该指定被所述 DRB 编码在报文中传播到所述其它 RBridge。如何传播后面会详细描述。

[0049] 在又一种实施方式中,该 RBridge 是通过接收命令被指定为计算 RBridge 的。例如系统管理员在某个 RBridge 中使用命令就可以将该计算 RBridge 指定为计算 RBridge。该指定被该 RBridge 编码在报文中传播到所述其它 RBridge。如何传播后面会详细描述。

[0050] 在上述的实施方式中,无论是由 DRB 指定计算 RBridge,还是在 RBridge 中通过接收命令被指定为计算 RBridge,响应于邻居关系中包括超过 3 个 RBridge,在邻居关系中指定两个计算 RBridge,其中一个为主计算 RBridge,另一个为从计算 RBridge。通常,主计算 RBridge 负责计算非单播路由信息并向邻居关系中的其它 RBridge(不包括从计算 RBridge)发送非单播路由信息;从计算 RBridge 也计算非单播路由信息,但是在主计算 RBridge 工作的情况下,不向邻居关系中的邻居 RBridge 发送非单播路由信息。如果主计算 RBridge 不可用了,例如出现故障,从计算 RBridge 将代替主计算 RBridge 承担计算和发送非单播路由信息的工作。

[0051] 在 DRB 指定计算 RBridge 的实施方式中,DRB 可以将指定的计算 RBridge 传播到邻居关系中的其它 RBridge,一种传播方式可以利用 hello 报文,具体来说,指定的计算 RBridge 信息被 DRB 编码在 Hello 报文的 TLV 中。图 5 示出了一种将指定的计算 RBridge 传播到邻居关系中的其它 RBridge 的 TLV 格式,图 5 中,Type 表示 TLV 的类型,可以定义为 Role-Appointment,即角色定义,Length 表示该 TLV 的长度,Master compute node appointment 和 slave compute node appointment 就是 TLV 中的值,分别表示主计算 RBridge 和从计算 RBridge 的别名(nickname)。

[0052] 针对在图 4 中,R4 是 DRB,R4 指定 R6 为主计算 RBridge,R5 为备计算 RBridge。其它 RBridge,包括 R1,R2,R3 和 R4 都是非计算 RBridge,图 6 示出了图 4 中 R6 的通知 TLV 消息内容。

[0053] 在通过命令指定计算 RBridge 的实施方式中,系统管理员可以使用命令“Compute-node role<Master,Slave>”指定主计算 RBridge 和从计算 RBridge,并且在主计算 RBridge 上通过命令“Slave nickname xxxx”同步从计算 RBridge 信息,在从计算 RBridge 上通过命令“Master nickname yyyy”同步主计算 RBridge 信息。例如在图 4 的 R1-R6 组成的邻居关系中,可以在 R6 上配置如下命令:Compute-node role Master 以及 Slave nickname R5 来告诉 R6 其为主计算 RBridge 并且其从计算 RBridge 为 R5,可以在 R5 上配置类似的命令。另外,主计算 RBridge 被指定为计算非单播路由信息的 RBridge 被该主计算 RBridge 编码在 Hello 报文中传播到邻居关系的其它 RBridge 中,使用的编码格式与图 5 描述的格式相同,传播方式也与 DRB 传播的方式相同,这里不再赘述。

[0054] 当然,本领域技术人员可以知道,上述使用 Hello 报文的 TLV 传播任命信息到邻居关系的其它 RBridge 中只是一种具体的实施方式,本领域技术人员完全可以参照上述实施

方式,定义使用其它报文类型的传播方式,例如使用 LSP 报文或者 SNP 报文等,甚至本领域技术人员可以定义单独的报文,来传播上述信息,这里不再赘述。

[0055] 在一种实施方式中,被指定为计算非单播路由信息的 RBridge 不是边界 (edge) RBridge,而是中间 (transit) RBridge,例如,上述 R1-R6 的邻居关系中,R4 和 R3 为边界 RBridge,不能为计算非单播路由信息的 RBridge,计算非单播路由信息的 RBridge 只能为 R1、R2、R5 和 R6 中的至少一个。

[0056] 在一种实施方式中,应该检测主计算 RBridge 和 / 或从计算 RBridge 的可用性。一旦检测到主计算 RBridge 和 / 或从计算 RBridge 不可用,可以启动一个定时器用来等待计算 RBridge 恢复。如果无法恢复,响应于仅主计算 RBridge 或仅从计算 RBridge 不可恢复,可以指定一个新的从计算 RBridge;响应于主计算 RBridge 和从计算 RBridge 都不可恢复,可以指定新的主计算 RBridge 和从计算 RBridge。上述步骤可以由 DRB 执行,或者由系统管理员通过命令指定某个 RBridge 执行。

[0057] 在一种实施方式中,从计算 RBridge 应该检测主计算 RBridge 可用性。响应于从计算 RBridge 检测到主计算 RBridge 不可用,从计算 RBridge 会将计算的非单播路由信息发送到邻居的 RBridge 中。

[0058] 在一种实施方式中,主计算 RBridge 和 / 或从计算 RBridge 被指定为计算 RBridge 后,在步骤 S301 会发送主或者从角色确认消息给邻居关系中的邻居 RBridge。确认消息也可以采用 hello 报文的角色确认子 TLV 或者其他类型的报文,例如使用 LSP 报文或者 SNP 报文等,甚至可以单独定义确认报文,图 7 示出了计算 RBridge 被指定为计算 RBridge 后发送的确认 TLV 格式。图 7 中,TLV 的类型为 Role-Ack,表示角色确认;Compute node role flag 可以为 1 或者 0,例如可以定义 1 是主计算 RBridge,0 是从计算 RBridge,或者相反。由于上述消息编码在 Hello 报文中,Hello 报文本身是有别名标识的,邻居 RBridge 收到 Hello 报文就知道是哪个 RBridge 发来的,解析出来的消息当然也知道是谁在发角色确认消息。

[0059] 针对在图 4 中,R6 为主计算 RBridge,R5 为从计算 RBridge。其他 RBridge,包括 R1,R2,R3 和 R4,都是非计算 RBridge,则 R6 发送的确认消息如图 8 所示。R5 发送的确认消息如图 9 所示。

[0060] 在步骤 S302 和步骤 D303,计算 RBridge 获得该 TRILL 网络的与非单播路由信息相关的网络拓扑以及根据所述网络拓扑计算非单播路由信息,都可以采用现有技术中的方法,其中该 TRILL 网络中与非单播路由信息相关的网络拓扑是以全网选举的若干节点为根节点的树,根据获得的网络拓扑计算非单播路由信息可以采用现有技术中的 SPF 算法。例如对于图 4 的 R5 和 R6,假设需要计算 3 棵树,R5 和 R6 都会计算这 3 棵树,但只有 R6 向邻居 RBridge 发送路由信息。R5 计算路由信息但不发送。

[0061] 在步骤 S304 将算出的非单播路由信息分发到所述其它 RBridge 中,一种实施方式中,主计算 RBridge 分发非单播路由信息可以采用 LSP 报文分发,LSP 是 TRILL 协议中的链路状态报文,主要用于向全网通告本节点的链路状态信息,从而利于其他 RBridge 计算到本 RBridge 的路由。LSP 的具体格式是由 LSP 报文头和众多 TLV 组成的。具体来说,可以采用图 10 所示的 LSP 报文中的 TLV 格式,来作为分发非单播路由信息的信息格式。

[0062] 图 10 中,TLV 类型 (type) 定义为 ROUTES-INFO;图 11 示出了图 10 中的邻居信息 (Neighbor Information) 包含的内容,即包含该 RBridge 的邻居别名 (Nickname),邻居

RBridge 的邻居个数 (NeighborCount), 邻居 RBridge 的邻居的别名 (NeighborNickname)。针对在图 4 中, R6 发送的路由信息如图 12 所示。

[0063] 邻居关系中的其它 RBridge, 例如图 4 中的 R1-R4 接收到 R6 分发的图 12 所示的非单播路由信息, 解析针对自己的非单播路由信息, 例如, R1 只需解析图 12 的第 3 行信息, R2 只需解析图 12 的第 4 行信息, 以此类推, 就可以无需计算而获得非单播路由信息, 节省其 CPU 资源。

[0064] 本领域技术人员可以知道, 主计算 RBridge 分发非单播路由信息也可以采用其它类型的报文分发, 例如也可以采用 Hello 报文、或者 SNP 报文等, 甚至可以单独定义分发报文, 实施方式类似, 这里不再赘述。

[0065] 在同一个发明构思下, 本发明还公开了一种在多链接透明互联 TRILL 网络中处理非单播路由信息的具有路由转发特性的网桥 RBridge, 其中该 RBridge 与其它 RBridge 在该 TRILL 网络中具有邻居关系, 并且该 RBridge 为所述邻居关系中的计算非单播路由信息的 RBridge, 图 13 示出了一种 TRILL 网络中 RBridge1300 的结构框图, 根据图 13, 该 RBridge 包括: 确认消息发送装置 1301, 被配置为向所述其它 RBridge 发送确认消息, 其中所述确认消息确认该 RBridge 为计算非单播路由信息的 RBridge; 获得装置 1302, 被配置为获得该 TRILL 网络中与所述非单播路由信息相关的网络拓扑; 计算装置 1303, 被配置为根据所述网络拓扑计算所述非单播路由信息; 以及分发装置 1304, 被配置为将算出的所述非单播路由信息分发到所述其它 RBridge。

[0066] 在一种实施方式中, 所述其它 RBridge 不计算非单播路由信息。

[0067] 在一种实施方式中, 该 RBridge 与所述其它 RBridge 位于该 TRILL 网络的一个链路 Link 上。

[0068] 在一种实施方式中, 该 RBridge 是被所述邻居关系中指定的具有路由转发特性的网桥 DRB 指定为计算非单播路由信息的 RBridge, 并且该指定被所述 DRB 编码在报文中传播到所述其它 RBridge。

[0069] 在一种实施方式中, 该 RBridge 是通过接收命令被指定为计算非单播路由信息的 RBridge, 并且该指定被该 RBridge 编码在报文中传播到所述其它 RBridge。

[0070] 在一种实施方式中, 该 RBridge 不是边界 RBridge。

[0071] 在一种实施方式中, 在该 TRILL 网络中, 响应于该邻居关系中包括至少 3 个 RBridge, 则所述邻居关系中包括的该 RBridge 为主计算 RBridge, 所述邻居关系中还包括从计算 RBridge 和所述其它 RBridge, 该从计算 RBridge 计算所述非单播路由信息, 并且在该主计算 RBridge 工作的情况下, 不向所述其它 RBridge 发送非单播路由信息。

[0072] 在一种实施方式中, 在该 TRILL 网络中, 响应于该从计算 RBridge 检测到该主计算 RBridge 不可用, 该从计算 RBridge 向所述其它 RBridge 发送非单播路由信息。

[0073] 在一种实施方式中, 在该 TRILL 网络中, 该主计算 RBridge 和该从计算 RBridge 的可用性都被检测, 响应于仅该主计算 RBridge 或仅从计算 RBridge 不可用, 指定新的从计算 RBridge, 响应于该主计算 RBridge 和该从计算 RBridge 都不可用, 指定新的主计算 RBridge 和新的从计算 RBridge。

[0074] 本发明可以是系统、方法和 / 或计算机程序产品。计算机程序产品可以包括计算机可读存储介质, 其上载有用于使处理器实现本发明的各个方面的计算机可读程序指令。

[0075] 计算机可读存储介质可以是可以保持和存储由指令执行设备使用的指令的有形设备。计算机可读存储介质例如可以是一—但不限于—电存储设备、磁存储设备、光存储设备、电磁存储设备、半导体存储设备或者上述的任意合适的组合。计算机可读存储介质的更具体的例子（非穷举的列表）包括：便携式计算机盘、硬盘、随机存取存储器（RAM）、只读存储器（ROM）、可擦式可编程只读存储器（EPROM 或闪存）、静态随机存取存储器（SRAM）、便携式压缩盘只读存储器（CD-ROM）、数字多功能盘（DVD）、记忆棒、软盘、机械编码设备、例如其上存储有指令的打孔卡或凹槽内凸起结构、以及上述的任意合适的组合。这里所使用的计算机可读存储介质不被解释为瞬时信号本身，诸如无线电波或者其他自由传播的电磁波、通过波导或其他传输媒介传播的电磁波（例如，通过光纤电缆的光脉冲）、或者通过电线传输的电信号。

[0076] 这里所描述的计算机可读程序指令可以从计算机可读存储介质下载到各个计算/处理设备，或者通过网络、例如因特网、局域网、广域网和/或无线网下载到外部计算机或外部存储设备。网络可以包括铜传输电缆、光纤传输、无线传输、路由器、防火墙、交换机、网关计算机和/或边缘服务器。每个计算/处理设备中的网络适配卡或者网络接口从网络接收计算机可读程序指令，并转发该计算机可读程序指令，以供存储在各个计算/处理设备中的计算机可读存储介质中。

[0077] 用于执行本发明操作的计算机程序指令可以是汇编指令、指令集架构（ISA）指令、机器指令、机器相关指令、微代码、固件指令、状态设置数据、或者以一种或多种编程语言的任意组合编写的源代码或目标代码，所述编程语言包括面向对象的编程语言—诸如 Java、Smalltalk、C++ 等，以及常规的过程式编程语言—诸如“C”语言或类似的编程语言。计算机可读程序指令可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络—包括局域网（LAN）或广域网（WAN）—连接到用户计算机，或者，可以连接到外部计算机（例如利用因特网服务提供商来通过因特网连接）。在一些实施例中，通过利用计算机可读程序指令的状态信息来个性化定制电子电路，例如可编程逻辑电路、现场可编程门阵列（FPGA）或可编程逻辑阵列（PLA），该电子电路可以执行计算机可读程序指令，从而实现本发明的各个方面。

[0078] 这里参照根据本发明实施例的方法、装置（系统）和计算机程序产品的流程图和/或框图描述了本发明的各个方面。应当理解，流程图和/或框图的每个方框以及流程图和/或框图中各方框的组合，都可以由计算机可读程序指令实现。

[0079] 这些计算机可读程序指令可以提供给通用计算机、专用计算机或其它可编程数据处理装置的处理器，从而生产出一种机器，使得这些指令在通过计算机或其它可编程数据处理装置的处理器执行时，产生了实现流程图和/或框图中的一个或多个方框中规定的功能/动作的装置。也可以把这些计算机可读程序指令存储在计算机可读存储介质中，这些指令使得计算机、可编程数据处理装置和/或其他设备以特定方式工作，从而，存储有指令的计算机可读介质则包括一个制品，其包括实现流程图和/或框图中的一个或多个方框中规定的功能/动作的各个方面的指令。

[0080] 也可以把计算机可读程序指令加载到计算机、其它可编程数据处理装置、或其它

设备上,使得在计算机、其它可编程数据处理装置或其它设备上执行一系列操作步骤,以产生计算机实现的过程,从而使得在计算机、其它可编程数据处理装置、或其它设备上执行的指令实现流程图和 / 或框图中的一个或多个方框中规定的功能 / 动作。

[0081] 附图中的流程图和框图显示了根据本发明的多个实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上,流程图或框图中的每个方框可以代表一个模块、程序段或指令的一部分,所述模块、程序段或指令的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。在有些作为替换的实现中,方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如,两个连续的方框实际上可以基本并行地执行,它们有时也可以按相反的顺序执行,这依所涉及的功能而定。也要注意的,框图和 / 或流程图中的每个方框、以及框图和 / 或流程图中的方框的组合,可以用执行规定的功能或动作的专用的基于硬件的系统来实现,或者可以用专用硬件与计算机指令的组合来实现。

[0082] 以上已经描述了本发明的各实施例,上述说明是示例性的,并非穷尽性的,并且也不限于所披露的各实施例。在不偏离所说明的各实施例的范围和精神的情况下,对于本技术领域的普通技术人员来说许多修改和变更都是显而易见的。本文中所用术语的选择,旨在最好地解释各实施例的原理、实际应用或对市场中的技术的技术改进,或者使本技术领域的其它普通技术人员能理解本文披露的各实施例。

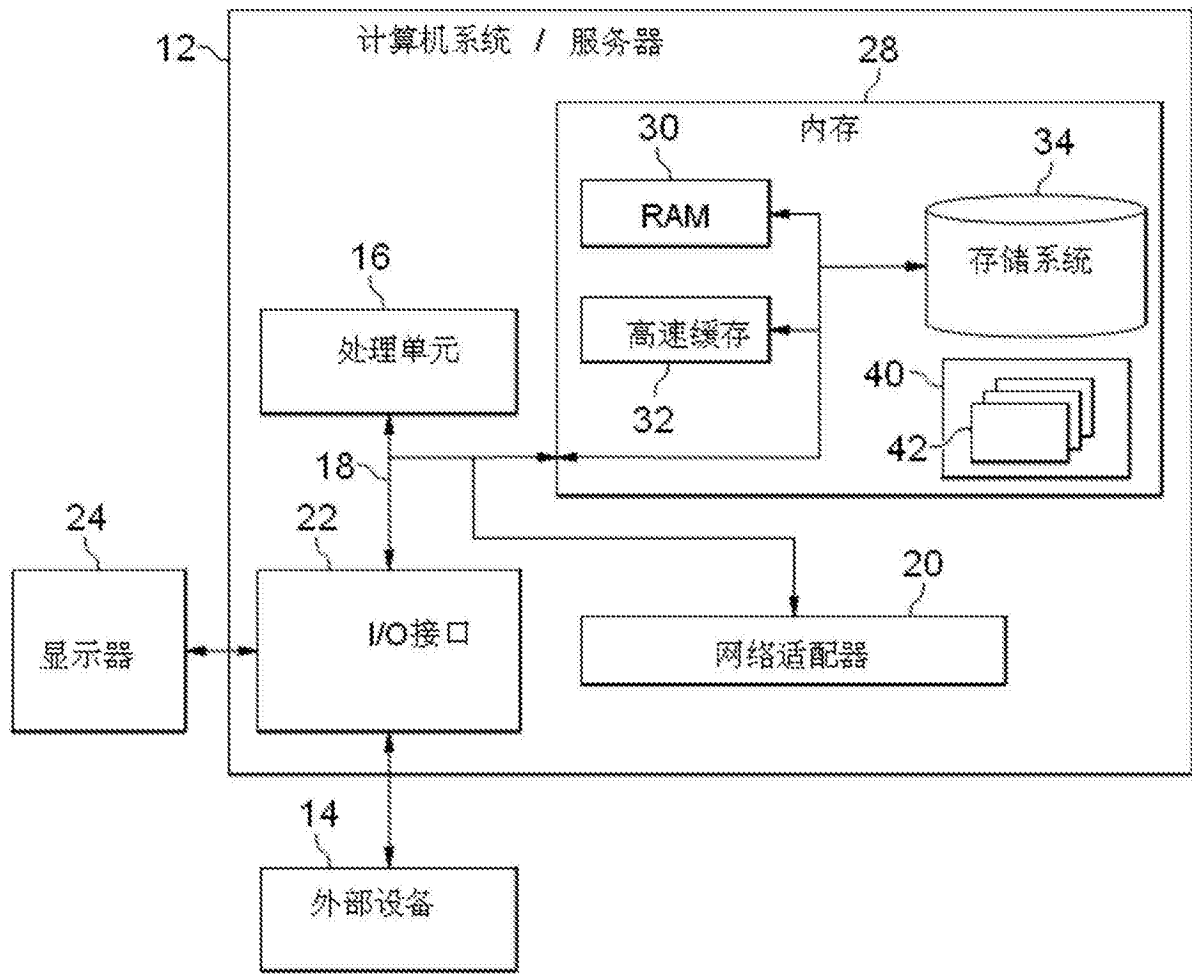


图 1

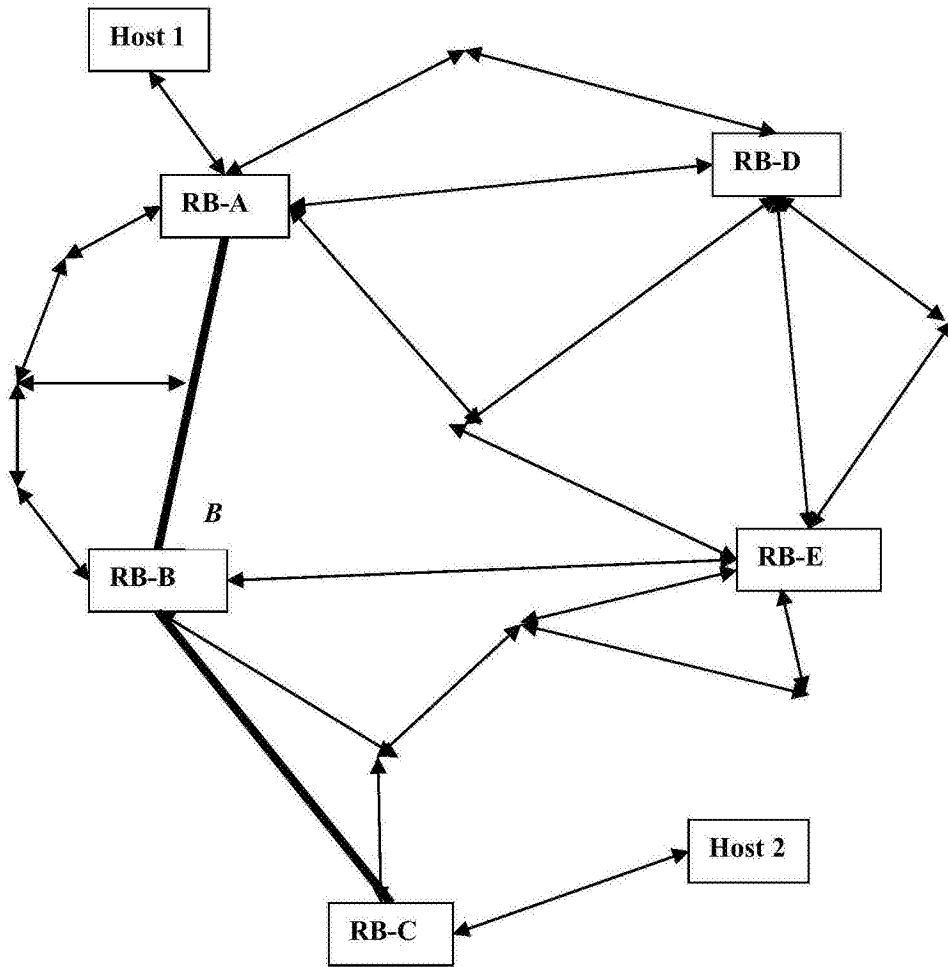


图 2

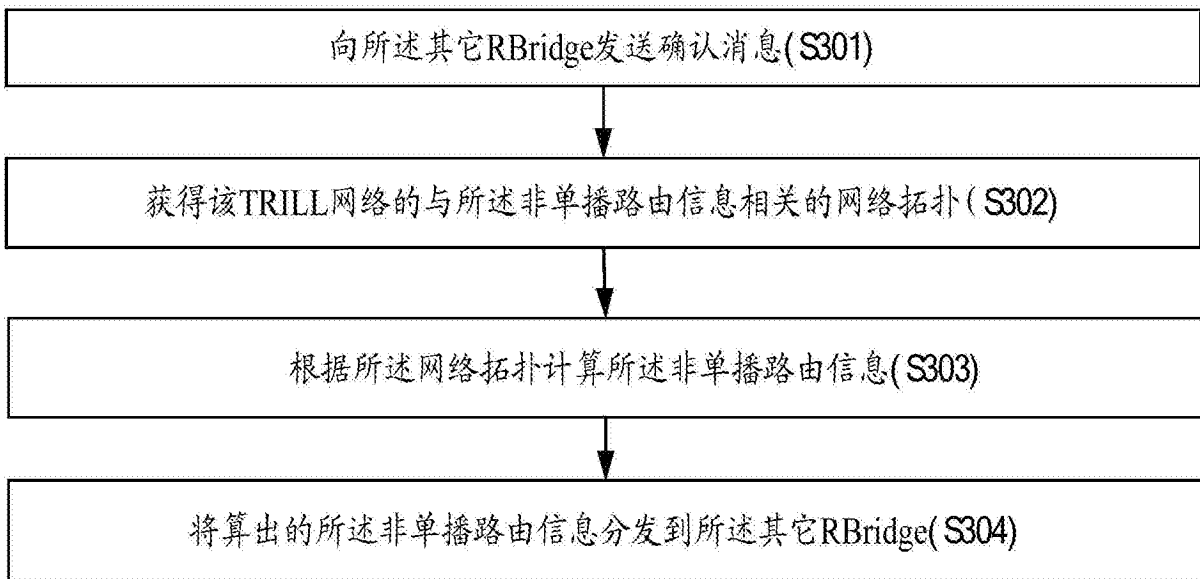


图 3

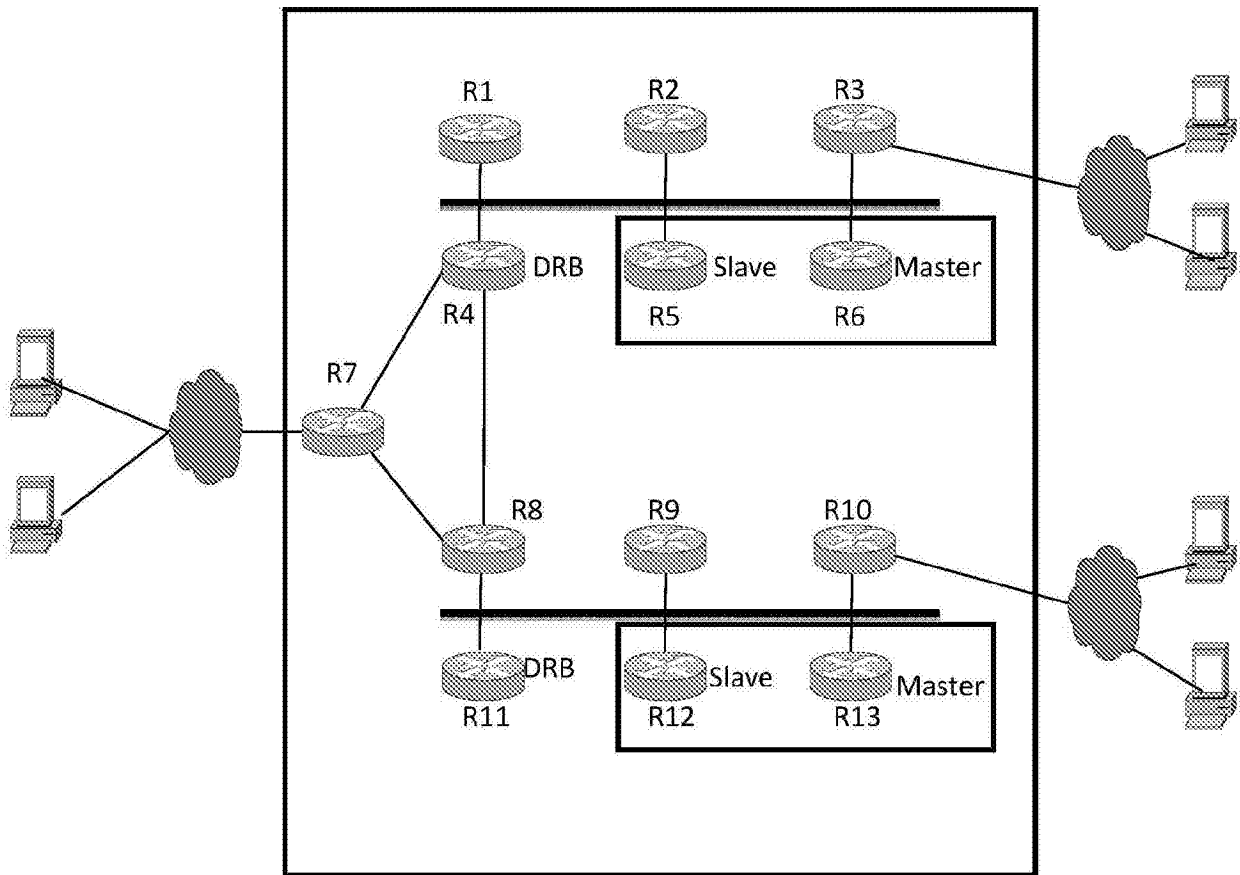


图 4

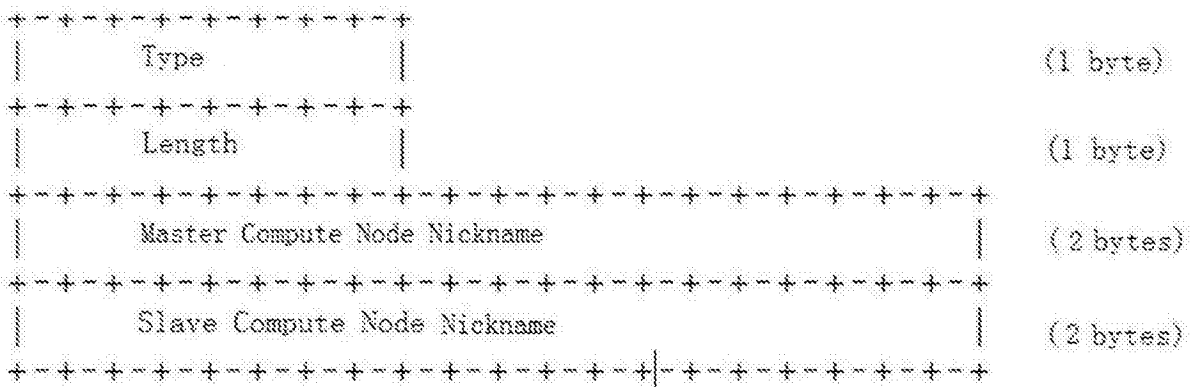


图 5

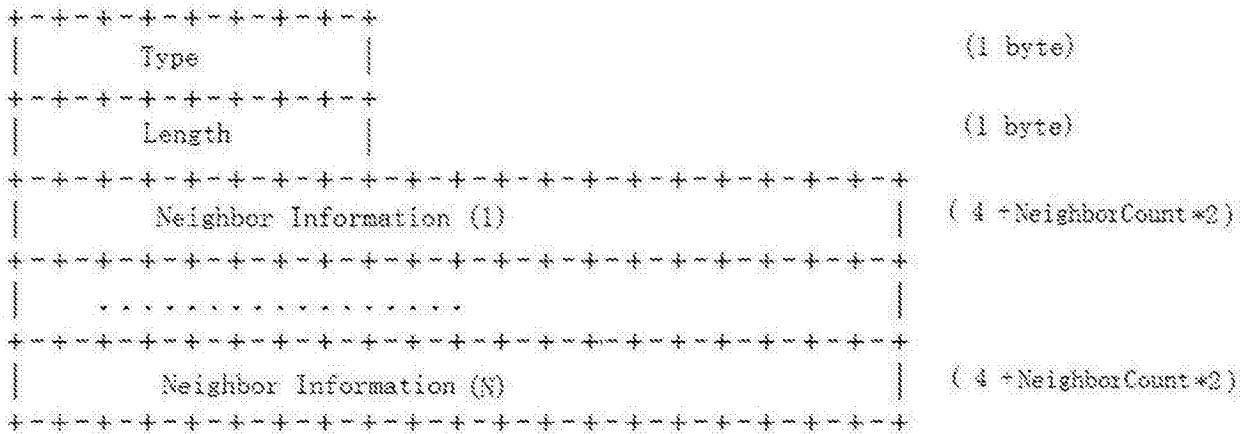


图 10

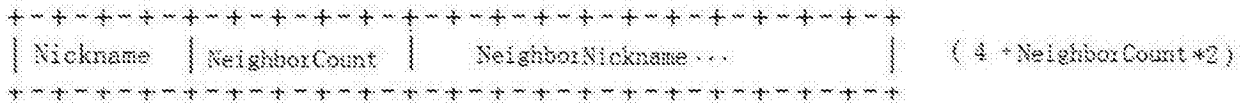


图 11

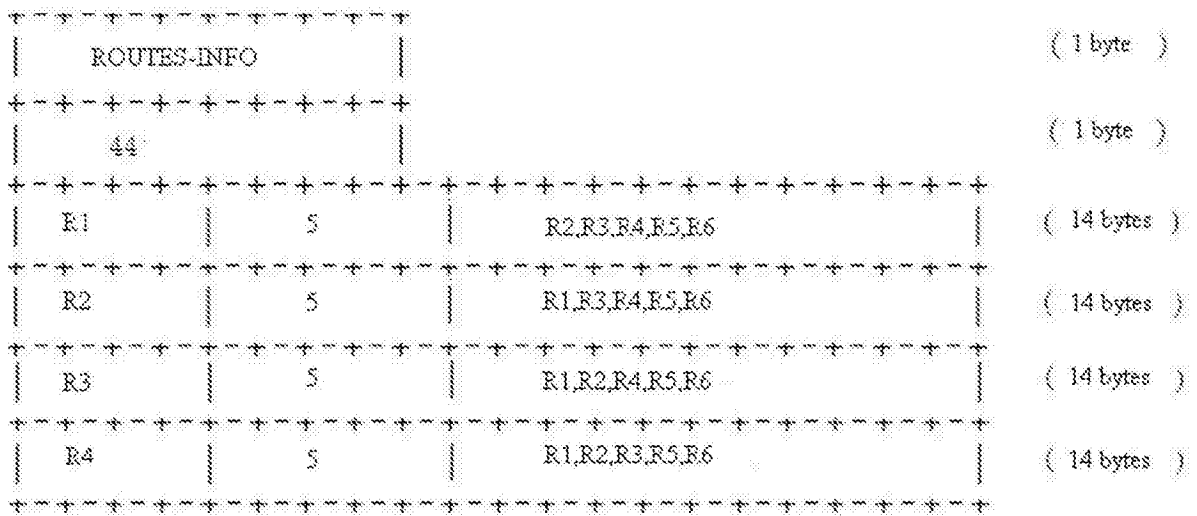


图 12

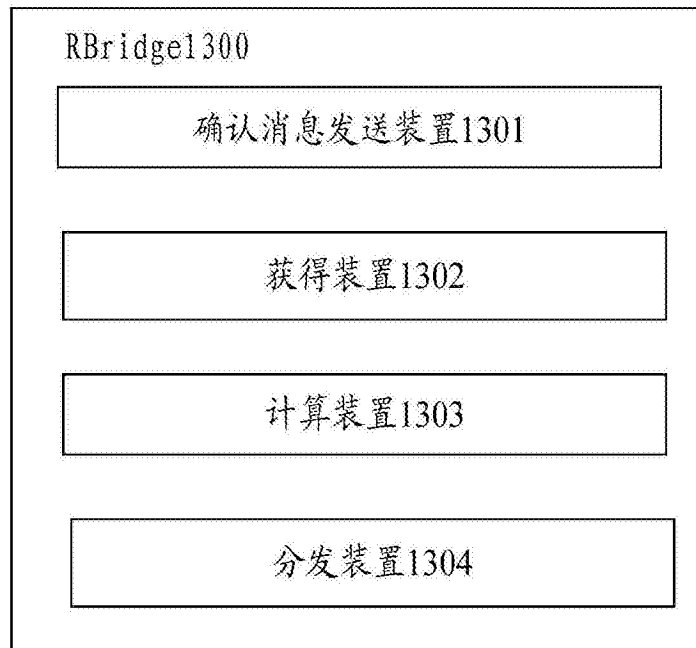


图 13