



(12) 发明专利申请

(10) 申请公布号 CN 104518963 A

(43) 申请公布日 2015. 04. 15

(21) 申请号 201410509829. 4

(22) 申请日 2014. 09. 28

(30) 优先权数据

14/042, 015 2013. 09. 30 US

(71) 申请人 聪博网络公司

地址 美国加利福尼亚州

(72) 发明人 K·康佩拉

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 王茂华

(51) Int. Cl.

H04L 12/723(2013. 01)

H04L 12/931(2013. 01)

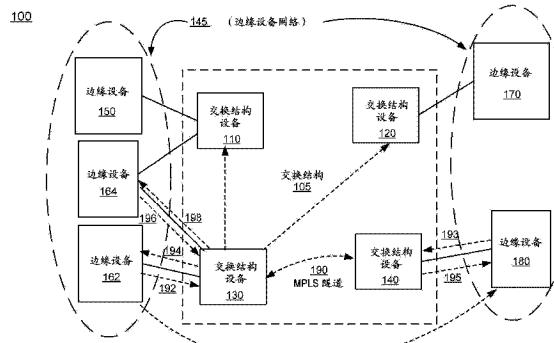
权利要求书3页 说明书14页 附图3页

(54) 发明名称

用于经由交换结构在边缘设备之间实施连接的方法和装置

(57) 摘要

在一些实施例中，一种装置包括经由交换结构可操作地耦合至第二边缘设备的第一边缘设备。第一边缘设备和第二边缘设备共同定义利用基于网络地址的协议来操作的边缘设备网络。该第一边缘设备通过该交换结构、经由多协议标签交换 (MPLS) 与第二边缘设备通信。此外，该第一边缘设备可操作地耦合至该交换结构使得该交换结构的节点能够在没有边缘设备网络的协调的情况下被修改。此外，该第一边缘设备可操作地耦合至第二边缘设备以定义边缘设备网络，使得该边缘设备网络中的边缘设备能够在没有该交换结构的协调的情况下被修改。



1. 一种装置，包括：

第一边缘设备，被配置为经由交换结构可操作地耦合至第二边缘设备，使得所述第一边缘设备和所述第二边缘设备共同定义利用基于网络地址的协议来操作的边缘设备网络，所述第一边缘设备被配置为可操作地耦合至所述第二边缘设备，使得所述第一边缘设备通过所述交换结构、经由多协议标签交换 (MPLS) 隧道与所述第二边缘设备通信，

所述第一边缘设备被配置为可操作地耦合至所述交换结构，使得所述交换结构的节点能够在没有所述边缘设备网络的协调的情况下被修改，

所述第一边缘设备被配置为可操作地耦合至所述第二边缘设备以定义所述边缘设备网络，使得所述边缘设备网络的边缘设备能够在没有所述交换结构的协调的情况下被修改。

2. 根据权利要求 1 所述的装置，其中：

所述第一边缘设备被配置为根据自动发现协议向所述交换结构的交换结构设备发送请求，

所述第一边缘设备被配置为根据所述自动发现协议并且响应于所述请求而接收响应，所述第一边缘设备被配置为向所述第二边缘设备发送被包括在所述响应中的标签，使得所述 MPLS 隧道基于所述标签而被建立在所述第一边缘设备与所述第二边缘设备之间。

3. 根据权利要求 2 所述的装置，其中：

所述自动发现协议是动态主机配置协议 (DHCP)、链路层发现协议 (LLDP) 或地址解析协议 (ARP) 中的至少一个。

4. 根据权利要求 2 所述的装置，其中：

所述第一边缘设备被配置为根据边界网关协议 (BGP) 向所述第二边缘设备发送被包括在所述响应中的所述标签。

5. 根据权利要求 1 所述的装置，其中：

所述第一边缘设备被配置为可操作地耦合至所述交换结构，使得所述第一边缘设备和所述第二边缘设备在没有任何其它协议消息的情况下经由自动发现消息与所述交换结构通信。

6. 根据权利要求 2 所述的装置，其中：

所述交换结构设备是所述交换结构的第一交换结构设备，

所述第一边缘设备被配置为连接至所述第一交换结构设备，使得所述第一交换结构设备根据至少一种 MPLS 协议向所述交换结构的第二交换结构设备发送指示所述标签的广告。

7. 根据权利要求 6 所述的装置，其中：

所述至少一种 MPLS 协议是标签分发协议 (LDP)、资源预留协议 - 流量工程 (RSVP-TE) 或边界网关协议 (BGP) 中的至少一种。

8. 根据权利要求 1 所述的装置，其中：

所述第一边缘设备被配置为经由所述交换结构可操作地耦合至所述第二边缘设备，使得所述边缘设备网络具有第一地址域并且所述交换结构具有不同于所述第一地址域的第二地址域，

来自所述第一地址域的至少一个地址取决于 DHCP 请求并且与来自所述第二地址域的

至少一个地址相对应。

9. 根据权利要求 1 所述的装置，其中：

所述第一边缘设备被配置为可操作地耦合至所述第二边缘设备，使得所述第一边缘设备在没有互联网协议 (IP) 连接的情况下通过所述交换结构、经由所述 MPLS 隧道与所述第二边缘设备通信。

10. 根据权利要求 2 所述的装置，其中：

所述交换结构设备是所述交换结构的第一交换结构设备，

所述第一边缘设备被配置为可操作地耦合至所述第一交换结构设备，使得所述第一交换结构设备根据所述自动发现协议发送响应，所述响应包括熵标签的指示以及与所述交换结构的第二交换结构设备或者所述第二边缘设备相关联的度量值。

11. 一种装置，包括：

交换结构设备，被配置为被包括在交换结构内，所述交换结构可操作地耦合至利用基于网络地址的协议来操作的边缘设备网络，使得所述边缘设备网络中的边缘设备通过所述交换结构、经由多个多协议标签交换 (MPLS) 隧道进行通信，

所述交换结构设备被配置为在没有所述边缘设备网络的协调的情况下被修改，

所述交换结构设备被配置为可操作地耦合至所述边缘设备网络，使得所述边缘设备网络中的边缘设备能够在没有所述交换结构的协调的情况下被修改。

12. 根据权利要求 11 所述的装置，其中：

所述边缘设备是第一边缘设备

所述交换结构设备被配置为耦合至所述第一边缘设备，所述交换结构设备被配置为根据自动发现协议从所述第一边缘设备接收请求，所述交换结构设备被配置为响应于所述请求、根据所述自动发现协议向所述第一边缘设备发送响应，使得来自所述多个 MPLS 隧道中的 MPLS 隧道被建立在所述第一边缘设备与来自所述边缘设备网络的第二边缘设备之间。

13. 根据权利要求 12 所述的装置，其中：

所述交换结构设备被配置为耦合至所述第一边缘设备，使得所述第一边缘设备和所述第二边缘设备在没有其它协议消息的情况下经由自动发现消息进行通信。

14. 根据权利要求 11 所述的装置，其中：

所述交换结构设备是第一交换结构设备，所述第一交换结构设备被配置为根据至少一种 MPLS 协议向来自所述交换结构的第二交换结构设备发送广告，所述广告指示与来自所述多个 MPLS 隧道的 MPLS 隧道相关联的标签，所述 MPLS 隧道与来自所述边缘设备网络的第一边缘设备网络和来自所述边缘设备网络的第二边缘设备相关联。

15. 根据权利要求 11 所述的装置，其中：

所述交换结构设备被配置为耦合至所述边缘设备网络，使得所述边缘设备网络具有第一地址域并且所述交换结构具有不同于所述第一地址域的第二地址域，

来自所述第一地址域的至少一个地址取决于动态主机配置协议 (DHCP) 请求或地址解析协议 (ARP) 请求中的至少一个。

16. 一种装置，包括：

第一边缘设备，被配置为可操作地耦合至交换结构，所述交换结构耦合至第二边缘设备，所述第一边缘设备被配置为通过所述交换结构、经由多协议标签交换 (MPLS) 隧道与所

述第二边缘设备通信，

所述第一边缘设备被配置为可操作地耦合至所述第二边缘设备以定义边缘设备网络，使得所述边缘设备网络中的边缘设备能够独立于所述交换结构而被修改并且所述交换结构的节点能够独立于所述边缘设备网络而被修改。

17. 根据权利要求 16 所述的装置，其中：

所述第一边缘设备被配置为根据自动发现协议向所述交换结构的第一交换结构设备发送请求，

所述第一边缘设备被配置为响应于所述请求而根据所述自动发现协议接收响应，所述第一边缘设备被配置为向所述第二边缘设备发送被包括在所述响应中的标签，使得所述 MPLS 隧道基于所述标签而被建立在所述第一边缘设备与所述第二边缘设备之间。

18. 根据权利要求 17 所述的装置，其中：

所述自动发现协议是动态主机配置协议 (DHCP)、链路层发现协议 (LLDP) 或地址解析协议 (ARP) 中的至少一个。

19. 根据权利要求 16 所述的装置，其中：

所述第一边缘设备被配置为经由所述交换结构可操作地耦合至所述第二边缘设备，使得所述边缘设备网络具有第一地址域并且所述交换结构具有不同于所述第一地址域的第二地址域，

来自所述第一地址域的至少一个地址取决于动态主机配置协议 (DHCP) 请求或地址解析协议 (ARP) 请求中的至少一个。

20. 根据权利要求 16 所述的装置，其中：

所述第一边缘设备被配置为可操作地耦合至所述第二边缘设备，使得所述第一边缘设备在没有互联网协议 (IP) 连接的情况下通过所述交换结构、经由 MPLS 隧道与所述第二边缘设备通信。

21. 根据权利要求 17 所述的装置，其中：

所述第一边缘设备被配置为可操作地耦合至所述交换结构的第一交换结构设备，使得所述第一交换结构设备根据所述自动发现协议发送响应，所述响应包括熵标签的指示以及与所述交换结构的第二交换结构设备或者所述第二边缘设备相关联的度量值。

用于经由交换结构在边缘设备之间实施连接的方法和装置

背景技术

[0001] 这里所描述的一些实施例总体上涉及用于经由交换结构 (switch fabric) 在边缘设备网络中的边缘设备之间实施连接的方法和装置。特别地但并非作为限制,这里所描述的一些实施例涉及用于经由交换结构、在利用基于网络地址的协议进行操作的边缘设备网络中的边缘设备之间实施连接的方法和装置,上述交换结构保持边缘设备与交换结构设备之间的管理、控制平面、数据平面和 / 或地址分离。

[0002] 一些已知的交换结构包括提供服务的边缘设备和在边缘节点之间以及向其它网络提供连接的交换结构设备。这样的已知交换结构通常不会为了管理和 / 或操作而在边缘设备和交换结构设备之间进行分离。例如,边缘设备能够受制于相邻交换结构设备的管理协调。然而,这样的管理和操作边缘设备以及交换结构设备的方法会产生对交换结构造成不利影响的风险并且导致交换结构的安全隐患。此外,边缘设备和交换结构设备经常由不同群体进行管理。

[0003] 因此,需要经由交换结构、在利用基于网络地址的协议进行操作的边缘设备网络中的边缘设备之间实施连接的方法和装置,上述交换结构保持边缘设备与交换结构设备之间的管理、控制平面、数据平面和 / 或地址分离。

发明内容

[0004] 在一些实施例中,一种装置包括经由交换结构可操作地耦合至第二边缘设备的第一边缘设备。第一边缘设备和第二边缘设备共同定义利用基于网络地址的协议来操作的边缘设备网络。该第一边缘设备通过该交换结构、经由多协议标签交换 (MPLS) 与第二边缘设备通信。此外,该第一边缘设备可操作地耦合至该交换结构使得该交换结构的节点能够在没有边缘设备网络的协调的情况下被修改。此外,该第一边缘设备可操作地耦合至第二边缘设备以定义边缘设备网络,使得该边缘设备网络中的边缘设备能够在没有该交换结构的协调的情况下被修改。

附图说明

[0005] 图 1 是根据实施例的经由交换结构连接边缘设备的交换结构系统的示意性图示。

[0006] 图 2 是根据实施例的边缘设备的系统框图。

[0007] 图 3 是根据实施例的交换结构设备的系统框图。

[0008] 图 4 是图示根据实施例的用于经由交换结构而在边缘设备之间的通信的方法的流程图。

具体实施方式

[0009] 在一些实施例中,一种装置包括第一边缘设备,该第一边缘设备经由交换结构可操作地耦合至第二边缘设备。第一边缘设备和第二边缘设备共同定义利用基于网络地址的协议来操作的边缘设备网络。这样的基于网络地址的协议例如可以是内部路由协议

(interior routing protocol, IGP)、标签分发协议 (LDP)、边界网关协议 (BGP)、资源预留协议 – 流量工程 (RSVP-TE) 协议等。在这样的实施例中, 第一边缘设备通过交换结构、经由多协议标签交换 (MPLS) 与第二边缘设备通信。

[0010] 此外, 第一边缘设备可操作地耦合至交换结构使得该交换结构的节点能够在没有边缘设备网络的协调的情况下被修改。类似地, 该第一边缘设备可操作地耦合至第二边缘设备以定义边缘设备网络, 使得该边缘设备网络中的边缘设备能够在没有该交换结构的协调的情况下被修改。

[0011] 在一些情况下, 第一边缘设备可以被配置为根据自动发现协议向交换结构的第一交换结构设备发送请求。这样的自动发现协议例如可以是动态主机配置协议 (DHCP)、地址解析协议 (ARP)、Radius(远程认证拨号用户服务) 协议、Diameter 协议、简单文件传输协议 (TFTP)、链路层发现协议 (LLDP) 等。第一边缘设备可以被配置为根据自动发现协议并且响应于该请求而接收响应。该响应可以包括与第一边缘设备相关联的标签。此外, 第一交换结构设备能够根据至少一种 MPLS 协议向交换结构的第二交换结构设备发送指示该标签的广告。MPLS 协议例如可以是 LDP、RSVP-TE、BGP 等。

[0012] 在接收到该响应之后, 第一边缘设备可以被配置为向第二边缘设备发送该标签, 使得 MPLS 隧道基于该标签而被建立在第一边缘设备与第二边缘设备之间。在一些情况下, 第一边缘设备可以被配置为根据 BGP 向第二边缘设备发送该标签。在一些情况下, 第一边缘设备和第二边缘设备可以在没有任何其它协议消息 (例如, 任何基于网络地址的协议消息) 的情况下经由自动发现消息与交换结构通信。

[0013] 在一些实施例中, 一种装置包括交换结构设备, 该交换结构设备被包括在可操作地耦合至边缘设备网络的交换结构内。该边缘设备网络利用基于网络地址的协议来操作, 使得该边缘设备网络中的边缘设备能够通过该交换结构、经由例如 MPLS 隧道的集合进行通信。在这样的实施例中, 该交换结构设备可以在没有边缘设备网络的协调的情况下被修改。类似地, 该边缘设备网络中的边缘设备能够在没有交换结构的协调的情况下被修改。

[0014] 在一些实施例中, 一种装置包括可操作地耦合至交换结构的第一边缘设备, 该交换结构耦合至第二边缘设备。该第一边缘设备通过该交换结构、经由 MPLS 隧道与该边缘设备通信。在这样的实施例中, 该第一边缘设备还可操作地耦合至第二边缘设备以定义边缘设备网络, 使得该边缘设备网络中的边缘设备能够独立于交换结构而被修改并且该交换结构的 (多个) 节点能够独立于该边缘设备网络而被修改。

[0015] 如本说明书中所使用的, 术语“物理跳 (physical hop)”可以包括两个节点和 / 或设备之间的物理链接。例如, 将第一节点与第二节点可操作地耦合的数据路径可以被称作是物理跳。类似地指出, 物理跳可以物理地链接第一节点与第二节点。

[0016] 如本说明书中所使用的, 术语“单个物理跳”可以包括系统中两个节点和 / 或设备之间的直接物理连接。类似地指出, 单个物理跳可以包括两个节点在没有中间节点的情况下经由其进行耦合的链接。因此, 例如, 如果第一节点经由单个物理跳耦合至第二节点, 则该第一节点能够在不通过中间节点发送数据单元的情况下直接向第二节点发送数据单元 (例如, 数据分组、数据单位等)。

[0017] 如本说明书中所使用的, 模块例如可以是与执行具体功能相关联的可操作地耦合的电学部件的任意组件和 / 或集合, 并且例如可以包括存储器、处理器、电轨线、光学连接

器、(存储在存储器中和 / 或在硬件中执行的) 软件等。

[0018] 如本说明书中所使用的,除非上下文明确另外指示,否则单数形式“一个”(“a”、“an”和“the”)包括复数指代。因此,例如,术语“隧道模块”意在表示单个模块或者共同被配置为执行隧道传输数据的功能的多个模块的组合。

[0019] 图 1 是根据实施例的经由交换结构 105 连接边缘设备(例如,边缘设备 150、162、164、170、180) 的交换结构系统 100 的示意性图示。交换结构系统 100 包括交换结构 105 以及边缘设备网络 145。交换结构 105 包括一个或多个交换结构设备,诸如交换结构设备 110、120、130 和 140。边缘设备网络 145 包括一个或多个边缘设备,诸如边缘设备 150、162、164、170 和 180。在一些实施例中,虽然并未在图 1 中示出,但是交换结构 105 可以包括比图 1 所示的那些更多或更少的交换结构设备;边缘设备网络 145 可以包括比图 1 所示的那些更多或更少的边缘设备。

[0020] 在一些实施例中,交换结构系统 100 可以在各种位置和环境中实施,和 / 或针对不同功能来实施。例如,交换结构系统 100 可以被实施在数据中心,交换结构 105 在那里用作数据中心网络并且边缘设备网络 145 包括数据中心的服务器。对于另一个示例,交换结构系统 100 可以被实施为接入网络,其中交换结构 105 是城域网 (MAN) 并且边缘设备网络 145 包括支持 MPLS 的 DSLAM(数字订户线路接入复用器)、eNodeB(e-utran 节点 B 或演进型节点 B)、OLT(光学线路终端) 等。对于又另一个示例,交换结构系统 100 可以被实施为服务网络,其中交换结构 105 是广域网 (WAN),并且边缘设备网络 145 包括提供商边缘 (PE) 节点。

[0021] 来自边缘设备网络 145 的每个边缘设备耦合至来自交换结构 105 的至少一个交换结构设备。例如,如图 1 所示,交换结构设备 110 耦合至边缘设备 150 和 164;交换结构设备 120 耦合至边缘设备 170;交换结构设备 130 耦合至边缘设备 162 和 164;并且交换结构设备 140 耦合至边缘设备 180。在一些实施例中,虽然并未在图 1 中示出,但是来自交换结构 150 的每个交换结构设备可以(例如,经由 MPLS 隧道)可操作地耦合至来自交换结构 105 的每个其余交换结构设备。因此,来自边缘设备网络 145 的每个边缘设备能够经由交换结构 105 可操作地耦合至来自边缘设备网络 145 的每个其余边缘设备。

[0022] 边缘设备网络 14 的每个边缘设备可以是被配置为可操作地耦合至交换结构 105 的任意设备。这样的边缘设备例如可以是计算节点、服务器、路由器、访问交换机、网关等。在一些实施例中,边缘设备可以是输入节点、输出节点、架顶 (top-of-rack, TOR) 设备等。在一些实施例中,边缘设备可以用作数据传输的源边缘设备和目的地边缘设备。在一些实施例中,虽然并且在图 1 中示出,但是边缘设备网络 145 的边缘设备能够耦合至一个或多个外部处理设备或消费者节点。在这样的实施例中,边缘设备能够将(多个)外部处理设备可操作地耦合至交换结构 105。边缘设备能够往来于交换结构 105 以及往来于所耦合的(多个)外部处理设备发送和接收数据单元(例如,数据分组和 / 或数据单位的流)。

[0023] 能够连接至边缘设备网络 145 的边缘设备的外部处理设备(图 1 中未示出)例如可以包括服务器、存储设备、网关、工作站等。外部处理设备可以使用任意适当的连接而可操作地耦合至边缘设备的一个或多个端口,上述连接例如光学连接(例如,光学线缆和光学连接器)、电连接(例如,电线缆和电连接器)等。类似地指出,边缘设备的每个端口能够提供通信接口,外部处理设备能够通过该通信接口可操作地耦合至该边缘设备。这样,外部

处理设备能够向边缘设备发送数据单元和从边缘设备接收数据单元。在一些情况下，外部处理设备和边缘设备之间的连接可以是直接链接。这样的链接可以被称作单物理跳链接。在其它情况下，外部处理设备可以经由中间节点而可操作地耦合至边缘设备。这样的连接可以被称作多物理跳链接。

[0024] 在一些实施例中，边缘设备网络 145 的边缘设备能够被配置为（例如，向消费者）提供服务，诸如 VPN（虚拟私有网络）、伪线路、互联网接入等。例如在新的消费者加入交换结构系统 100、发起新的服务、引入新的边缘设备以容纳新的消费者等的时候，边缘设备能够为了维护而被升级、配置、重启或关机。此外，在一些实施例中，边缘设备例如可以被称作标签边缘、PE 路由器、提供商边缘节点等。

[0025] 图 2 是根据实施例的边缘设备 200 的框图。边缘设备 200 在结构和功能上可以类似于参考图 1 所示出并描述的边缘设备 150、162、164、170、180。特别地，类似于图 1 所示的边缘设备 145 中的边缘设备，边缘设备 200 可以是任意设备，其被配置为耦合至交换结构的一个或多个交换结构设备（例如，图 1 中的交换结构 105 中的交换结构设备 110、120、130、140）。因此，边缘设备 200 能够经由交换结构可操作地耦合至一个或多个其它边缘设备。

[0026] 如图 2 所示，边缘设备 200 包括通信接口 230、存储器 210 以及包括隧道模块 225 的处理器 220。虽然图 2 中并未示出，但是边缘设备 200 可以包括具有各种功能的其它模块。例如，边缘设备 200 可以包括 DHCP 模块，其被配置为（根据 DHCP）从将边缘设备 200 可操作地耦合至交换结构的交换结构设备获得地址。作为另一个示例，边缘设备 200 可以包括 ARP 模块，其被配置为根据 ARP 与将边缘设备 200 可操作地耦合至交换结构的交换结构设备进行通信。

[0027] 边缘设备 200 的通信接口 230（例如，传输 / 接收数据）、隧道模块 225（例如，封装 / 解封装数据单元）、其它模块（图 2 中未示出）的每种（多种）操作以及存储器 210 上的（多种）操作能够由处理器 220 控制。处理器 220 可以是被配置为运行和 / 或执行处理器 220 中所包括的（多个）模块的任意适当处理器。处理器 220 中的每个模块可以是能够执行与该模块相关联的一种或多种具体功能的基于硬件的模块（例如，现场可编程门阵列（FPGA）、专用集成电路（ASIC）、数字信号处理器（DSP）等）和 / 或基于软件的模块（例如，存储在存储器 210 中和 / 或在处理器 220 执行的计算机代码的模块）的任意组合。

[0028] 如以下所详细描述的，隧道模块 225 可以被配置为使得边缘设备 200 能够通过交换结构而使用隧道往来于耦合至该交换结构的其它边缘设备传输和接收数据。隧道模块 225 能够被配置为使得边缘设备 200 能够通过交换结构、使用 MPLS 与其它边缘设备进行通信。特别地，隧道模块 225 可以被配置为利用与 MPLS 隧道相关联的标签栈对数据进行封装。如以下详细描述的，这样的标签栈可以包括一个或多个标签，诸如识别连接交换结构中的两个交换结构设备的 MPLS 隧道的标签，识别从交换结构设备到耦合至该交换结构设备的边缘设备的路线的标签，等等。

[0029] 边缘设备 200 的通信接口 230 例如可以包括一个或多个端口（图 2 中未示出），其能够被用来在边缘设备 200 与将该边缘设备 200 可操作地耦合至交换结构的（多个）交换结构设备之间实施（多个）连接（例如，有线连接、无线连接）。该（多个）连接例如可以是经由电线缆的双绞线电信令、经由光纤线缆的光纤信令等。这样，边缘设备 200 能够被配置为通信接口 230 的（多个）端口接收数据和 / 或发送数据，通信接口 230 的这些端口与

(多个)交换结构设备的通信接口相连接。此外,在一些实施例中,边缘设备 200 能够通过通信接口 230 的一个端口与交换结构设备之一(例如,图 1 中的交换结构设备 110)实施一种类型的连接,而通过通信接口 230 的另一个端口与另一个交换结构设备(例如,图 1 中的交换结构设备 130)实施不同类型的连接。

[0030] 在一些实施例中,存储器 210 例如可以是随机访问存储器(RAM)(例如,动态 RAM、静态 RAM)、闪存、可移动存储器等。有关与将边缘设备 200 可操作地耦合至交换结构的(多个)交换结构设备进行交互以及通过交换结构建立与其它边缘设备的 MPLS 隧道的控制信息能够被存储在存储器 210 中。例如,如以下所讨论的,被分配给边缘设备 200 的地址和/或识别从交换结构设备(即,将边缘设备 200 耦合至交换结构的交换结构设备)到边缘设备的路线的标签能够被存储在存储器 210 中。作为另一个示例,识别到其它边缘设备的路线的标签和/或识别交换结构内的 MPLS 的标签也能够被存储在存储器 210 中。

[0031] 返回图 1,交换结构 105 的每个交换结构设备(例如,交换结构设备 110、120、130、140)可以是被配置为可操作地耦合至其它(多个)交换结构设备和/或(多个)边缘设备(例如,边缘设备 150、162、164、170、180),使得交换结构 105 能够将边缘设备互相可操作地耦合并且使得能够在该边缘设备之间进行数据传输的任意设备。因此,交换结构 105 的交换结构设备可以被配置为在边缘设备之间以及针对其它网络(图 1 中未示出)提供连接。这样的交换结构设备例如可以是路由设备(例如,L2/L3 路由器)、交换设备(例如,L2/L3 交换机)、网关、工作站、服务器、计算节点等。在一些实施例中,交换结构设备例如可以被称作核心节点、标签交换机、提供商节点、提供商路由器、结构节点等。

[0032] 在一些实施例中,除了提供到边缘设备的连接之外,交换结构 105 的交换结构设备并不提供其它服务。在这样的实施例中,在交换结构设备处出现硬件或软件故障、要在交换结构设备处安装新的线路卡等的时候,执行交换结构设备的维护。如以下详细讨论的,由于边缘设备网络 145 和交换结构 105 之间分离的管理和操作,交换结构设备在边缘设备被修改时并不被打扰。

[0033] 图 3 是根据实施例的交换结构设备 300 的系统框图。交换结构设备 300 在结构和功能上可以类似于关于图 1 所示出并描述的交换结构设备 110、120、130、140。特别地,类似于图 1 中的交换结构 105 的交换结构设备,交换结构设备 300 可以被包括在交换结构(例如,图 1 中的交换结构 105)内并且被配置为将一个或多个边缘设备(例如,图 1 中的边缘设备 150、162、164、170、180)可操作地耦合至该交换结构。

[0034] 如图 3 所示,交换结构设备 300 包括:通信接口 360、包括路由表 352 的存储器 320 以及包括隧道模块 345 的处理器 340。虽然并未在图 3 中示出,但是交换结构设备 300 可以包括具有各种功能的其它模块。例如,交换结构设备 300 可以包括 DHCP 模块,该 DHCP 模块被配置为处理从耦合至交换结构设备 300 的(多个)边缘设备所接收的 DHCP 请求。作为另一个示例,交换结构设备 300 可以包括 ARP 模块,该 ARP 模块被配置为根据 ARP 与耦合至交换结构设备 300 的(多个)边缘设备进行通信。

[0035] 交换结构设备 300 的通信接口 360(例如,传输/接收数据)、隧道模块 345(例如,传输隧道数据)、其它模块(图 3 中未示出)的每种(多种)操作以及存储器 320 上的(多种)操作能够由处理器 340 进行控制。处理器 340 可以是被配置为运行和/或执行处理器 340 中所包括的(多个)模块的任意适当处理器。处理器 340 中的每个模块可以是基于硬

件的模块（例如，CPU、通用处理器、FPGA、ASIC、DSP 等）、基于软件的模块（例如，存储在存储器 320 中和 / 或在处理器 340 执行的计算机代码的模块）和 / 或能够执行与该模块相关联的一种或多种具体功能的基于硬件和软件的模块的组合。

[0036] 隧道模块 345 可以被配置为使得交换结构设备 300 能够往来于其它交换结构设备或耦合至交换结构设备 300 的（多个）边缘设备传输和接收隧道数据。特别地，如以下详细描述的，隧道模块 345 可以被配置为使得能够使用将交换结构设备 300 与交换结构的其它交换结构设备相连接的 MPLS 隧道进行隧道数据的传输。因此，耦合至交换结构设备 300 的（多个）边缘设备能够通过交换结构设备 300、使用 MPLS 隧道与耦合至该交换结构的其它边缘设备进行通信。隧道模块 345 还可以被配置为对数据进行解封装以得到数据中所包括的（多个）标签，并且基于所得到的（多个）标签确定用于传输解封装的数据的路线。在一些实施例中，隧道模块 345 可以被配置为基于路由表 325 中所存储的路由信息来确定用于隧道数据的路线（例如，MPLS 隧道）。

[0037] 交换结构设备 300 的通信接口 360 例如可以包括至少两个端口（图 3 中未示出），它们能够被用来在交换结构设备 300 与（多个）边缘设备之间以及在交换结构设备 300 与该交换结构的（多个）其它交换结构设备之间实施连接。该连接例如可以是经由电线缆的双绞线电信令、经由光纤线缆的光纤信令等。这样，交换结构设备 300 能够被配置为通过通信接口 360 的端口来接收数据和 / 或发送数据，这些端口与（多个）边缘设备和（多个）其它交换结构设备的通信接口相连接。此外，在一些实施例中，交换结构设备 300 能够通过通信接口 360 的一个端口与（多个）边缘设备和（多个）其它交换结构设备之一实施一种类型的连接，同时通过通信接口 360 的另一个端口与另一个边缘设备或交换结构设备实施不同类型的连接。例如，交换结构设备 300 能够与边缘设备实施双绞线电信令连接（例如，经由电线缆），并且与另一个交换结构设备实施光纤信令连接（例如，经由光纤线缆）。

[0038] 在一些实施例中，存储器 320 例如可以是随机访问存储器 (RAM)（例如，动态 RAM、静态 RAM）、闪存、可移动存储器等。有关与耦合至交换结构设备 300 的（多个）边缘设备进行交互的控制信息能够被存储在存储器 320 中。在一些实施例中，有关与其它（多个）交换结构设备隧道传输数据的路由信息能够被存储在路由表 325 中。例如，在一些实施例中，路由表可以被配置为存储标签转发信息库 (LFIB)，其包括与 MPLS 隧道和 / 或边缘设备网络的边缘设备的标签相关联的路由信息。

[0039] 例如，路由表 325 中的 LFIB 能够存储条目，该条目包括识别将交换结构设备 300 与另一个交换结构设备相连接的 MPLS 隧道的标签、以及交换结构设备 300 中与该 MPLS 隧道相关联的端口（被包括在通信接口 360 中）的标识符。因此，响应于接收到包括识别 MPLS 隧道的标签的（多个）数据单元，隧道模块 345 能够基于该标签以及路由表 325 中的 LFIB 中所存储的条目来确定端口。隧道模块 345 随后能够经由该端口发送（多个）数据单元，使得该（多个）数据单元通过 MPLS 隧道被发送到其它交换结构设备。

[0040] 作为另一个示例，路由表 325 中的 LFIB 能够存储条目，该条目包括识别耦合至交换结构设备 300 的边缘设备的标签、以及交换结构设备 300 中与该边缘设备相关联的端口（被包括在通信接口 360 中）的标识符。因此，响应于接收到包括识别该边缘设备的标签的（多个）数据单元，隧道模块 345 能够基于该标签以及存储在路由表 325 中的 LFIB 中的条目来确定该端口。隧道模块 345 随后能够经由该端口发送（多个）数据单元，使得该（多

个)数据单元被发送到该边缘设备。

[0041] 返回图1,在一些实施例中,边缘设备网络145和交换结构105能够被分离地管理、控制和/或操作。也就是说,交换结构系统100可以针对边缘设备网络145中的边缘设备(例如,边缘设备150、162、164、170、180)和针对交换结构中的交换结构设备(例如,交换结构设备110、120、130、140)实施分离的管理、控制和/或操作。在这样的实施例中,边缘设备和交换结构设备之间的分离能够在一个或多个方面中被实施,例如包括管理分离、控制平面分离、数据平面分离、地址分离等。

[0042] 在一些实施例中,可以在边缘设备网络145的边缘设备与交换结构105的交换结构设备之间实施管理分离。特别地,边缘设备和交换结构设备能够由两个(独立)管理员群组分离地管理。例如,边缘设备网络145能够由一个或多个边缘管理员的群组进行管理,而交换结构105则能够由一个或多个结构管理员的另一个群组进行管理。每个边缘管理员或结构管理员例如可以是人、机器(自动执行或由人手工操作)或者(多个)人和(多个)机器的组合,它们针对(多个)边缘设备或(多个)交换结构设备执行管理功能。

[0043] 作为管理分离的结构,当边缘设备网络145或交换结构105中的中一方被管理时(例如,被修改、被配置)时,不需要来自边缘设备网络145或交换结构105中的另一方的协调。例如,在没有来自(多个)边缘管理员的协调的情况下并且在不对边缘设备网络145中的任何边缘设备进行中断、修改或重新配置的情况下,(多个)结构管理员能够对交换结构105中的交换结构设备(例如,交换结构设备110、120、130或140)进行管理(例如,重启、升级、修改配置、改变MPLS协议设置等)。作为另一个示例,类似地,在没有来自(多个)结构管理员的协调的情况下并且在不对交换结构105的任何交换结构设备进行中断、修改或重新配置的情况下,(多个)边缘管理员能够对边缘设备网络145中的边缘设备(例如,边缘设备150、162、164、170或180)进行管理(例如,重启、升级、修改配置等)。

[0044] 此外,作为管理分离和/或其它分离机制的结果,能够针对边缘设备网络145和交换结构105实施即插即用连接。也就是说,(多个)边缘管理员能够在不与(多个)结构管理员进行协调并且不对边缘设备进行中断、修改或重新配置的情况下增加、去除或移动边缘设备。类似地,(多个)结构管理员能够不与(多个)边缘管理员进行协调并且不对交换结构设备进行中断、修改或重新配置的情况下增加、去除或移动交换结构设备。

[0045] 在一些实施例中,可以在边缘设备网络145的边缘设备与交换结构105的交换结构设备之间实施路由协议分离。特别地,在交换结构105内实施的路由结构可以在没有边缘设备网络的协调的情况下得以被执行(例如,信号、消息等的交换)。此外,在边缘设备与交换结构设备之间实施的协议(例如,DHCP、ARP)能够在没有(多个)边缘管理员与(多个)结构管理员之间的协调的情况下得以被执行(例如,信号、消息等的交换)。作为结果,在交换结构系统100中实施的协议能够在没有(多个)结构管理员与(多个)边缘管理员之间的管理协调的情况下得以被执行。

[0046] 在一些实施例中,如图1所示,多个边缘设备(例如,边缘设备162、164)能够耦合至一个交换结构设备(例如,交换结构设备130)。在一些实施例中,如图1所示,能够针对边缘设备网络145中的边缘设备实施有效的多归属机制。特别地,边缘设备能够针对多于一个的交换结构设备进行耦合和多归属。在这样的实施例中,边缘设备与相关联的交换结构设备之间的多个链接能够被用于边缘设备与交换结构105之间所传输的负载平衡业务。结

果,交换结构 105 内的业务也能够被负载平衡。此外,边缘设备与相关联的交换结构设备之间的多个链接能够被用来提高边缘设备与交换结构 105 之间的连接的弹性 (resilience)。作为结果,能够针对边缘设备和交换结构 105 之间的数据传输实现快速恢复。特别地,通过将数据业务指向连接边缘设备与交换结构 105 的(多个)其余链接,能够减少在检测到故障(例如,连接边缘设备和相关联的交换结构设备的链接的故障,耦合至边缘设备的相关联交换结构设备的故障)时的失效备援时间。

[0047] 例如,如图 1 所示,边缘设备 164 能够经由两个分离的链接分别被耦合至交换结构设备 110 和 130。作为结果,在边缘设备 164 与交换结构设备 110、130 之间传输的数据业务能够在这两个链接之间进行负载平衡。此外,在交换结构 105 内往来于交换结构设备 110、130 进行传输的数据业务能够基于这两个链接上的负载而进行负载平衡。此外,这两个链接能够被用来提高弹性并且为边缘设备 164 提供快速恢复。

[0048] 在一些实施例中,可以在边缘设备网络 145 的边缘设备与交换结构 105 的交换结构设备之间实施连接隔离。特别地,来自边缘设备网络 145 的每个边缘设备并不直接与来自边缘设备网络 145 的每个其余边缘设备进行通信。相反,边缘设备网络 145 的边缘设备通过交换结构 105 互相通信。也就是说,来自边缘设备网络 145 的每个边缘设备能够经由交换结构 105 的一个或多个交换结构设备向来自边缘设备网络 145 的每个其余边缘设备发送(多个)数据单元(例如,(多个)数据平面数据单元、(多个)控制平面数据单元),而不是直接从一个边缘设备向另一个边缘设备进行发送。此外,能够在边缘设备与交换结构设备之间进行传输的(多个)数据单元(例如,(多个)数据平面数据单元、(多个)控制平面数据单元)受到限制。例如,(多个)数据单元能够根据给定协议的集合在边缘设备与交换结构设备之间进行传输。换句话说,(多个)数据单元并不根据任何不被包括在给定协议集合中的协议而在边缘设备与交换结构设备之间进行传输。例如,在一些情况下,交换结构 105 可以(例如,经由 MPLS 隧道)向边缘设备网络 145 的边缘设备提供 MPLS 连接,而并不向那些边缘设备提供互联网协议(IP)连接。因此,(多个)数据单元能够根据 MPLS 协议而不是根据 IP 路由协议在边缘设备与交换结构设备之间进行传输。作为连接隔离的结果,使得被盗用的边缘设备无法对交换结构 105 进行攻击。类似地,使得被盗用的边缘设备无法对边缘设备网络 145 进行攻击。

[0049] 在一些实施例中,可以在边缘设备网络 145 的边缘设备与交换结构 105 的交换结构设备之间实施地址分离(或地址独立)。特别地,边缘设备网络 145 和交换结构 105 可以分别使用两个独立的地址域。在一些情况下,这两个独立地址域可以具有(多个)重叠地址但是没有(多个)重叠地址的冲突。在一些情况下,这两个地址域是完全不同的(即,不共享重叠地址)。

[0050] 在一些实施例中,边缘设备网络 145 的边缘设备可以被配置为根据自动发现协议(例如,DHCP、ARP)从耦合至该边缘设备的交换结构设备获得地址。例如,在这样的实施例中,负责管理边缘设备网络 145 的边缘管理员能够为边缘设备网络 145 选择地址空间(例如,地址域)并且向负责管理交换结构 105 的交换管理员通知所选择的地址空间。此外,边缘管理员还能够向结构管理员更新边缘设备网络 145 的地址空间上的任意变化。作为结果,结构管理员能够向交换结构 105 的交换结构设备分配地址空间,该地址空间并不与边缘设备网络 145 的地址空间相重叠。因此,被分配给边缘设备的地址可以不同于被分配给

交换结构设备的地址。

[0051] 例如,边缘设备 162 能够向交换结构设备 130 发送针对地址的请求(例如,DHCP 请求、ARP 请求)。响应于该请求,在交换结构设备 130 处从边缘设备网络 145 的地址空间选择地址并且随后将该地址从交换结构设备 130 发送至边缘设备 162。在一些情况下,被分配给边缘设备 162 的地址也可以被包括在交换结构 105 的地址空间中。因此,该地址将不会被分配给交换结构 105 的任何交换结构设备;作为结果,不会有交换结构设备与边缘设备 162 共享相同的地址。

[0052] 在一些实施例中,交换结构 105 的交换结构设备(例如,交换结构设备 110、120、130、140)可以共同被配置为使得能够在边缘设备网络 145 的边缘设备(例如,边缘设备 150、162、164、170、180)之间进行数据传输。交换结构 105 能够被配置为实施标签交换路径(LSP)的网格,使得交换结构 105 的交换结构设备能够使用 LSP 的网格互相传输数据。能够使用连接交换结构设备的 MPLS 隧道来实施 LSP 的网格,因此形成 MPLS 隧道的网格。在一些实施例中,例如 IGP、LDP、RSVP-TE 等的路由协议能够被用来生成 MPLS 隧道的网格。特别地,例如可以通过在交换结构 105 内执行 LGP 及 LDP、RSVP-TE 和 / 或其它路由协议来建立 MPLS 隧道的网格。

[0053] 在一些实施例中,MPLS 隧道的网格在交换结构 105 内实施且并不延伸至边缘设备网络 145 的边缘设备。也就是说,在交换结构 105 内实施的(多种)路由协议并不在交换结构设备和边缘设备的连接处被实施。相反,自动发现协议(例如,DHCP、ARP、LLDP)能够被用来使得能够在交换结构设备与边缘设备之间进行通信。以下进一步描述使得能够使用 MPLS 隧道在边缘设备之间进行数据传输的细节。

[0054] 如以上所讨论的,交换结构 105 和边缘设备网络 145 均能够具有独立的地址空间。在这样的情况下,交换结构 105 中的每个交换结构设备能够从用于边缘设备网络 145 的地址空间中被分配地址(例如,IP 地址)的块。因此,交换结构设备能够响应于从边缘设备接收到请求(例如,DHCP 请求、ARP 请求)而从分配给该交换结构设备的地址块中向边缘设备分配地址。此外,在一些实施例中,交换结构 105 的交换结构设备能够在它们自己之间保持 IP 连接而并不向边缘设备网络 145 的边缘设备提供 IP 连接。

[0055] 能够参考图 1 而考虑通过交换结构 105 使用 MPLS 隧道在边缘设备(例如,边缘设备 162)与其它边缘设备(例如,边缘设备 164、170、180)之间发起数据传输的方法。最初,边缘设备 162 能够根据自动发现协议(例如,DHCP、ARP、LLDP)向交换结构设备 130 发送请求(在图 1 中示为信号 192)。在一些实施例中,边缘设备 162 能够被配置为例如响应于边缘设备 162 被启动、开机、耦合至交换结构设备 130 等而发送这样的请求。

[0056] 在一些实施例中,该请求可以是针对分配给边缘设备 162 的地址和 / 或标签的请求。例如,该请求可以是针对分配给边缘设备 162 的 IP 地址和 / 或给其它边缘设备用于向边缘设备 162 发送数据的标签的 DHCP 请求。响应于该请求,交换结构设备 130 能够向边缘设备 162 发送响应(在图 1 中示为信号 194)。该响应可以包括分配给边缘设备 162 的地址(例如,IP 地址)。例如,交换结构设备 130 能够向边缘设备 162 分配 IP 地址 10.1.3.1。如以上所讨论的,分配给边缘设备 162 的地址能够从边缘设备网络 145 的地址空间中进行选择,该地址空间独立于交换结构 105 的地址空间。在一些实施例中,边缘设备 145 的地址空间可以与交换结构 105 的地址空间有所重叠。作为结果,分配给边缘设备 162 的地址能

够对应于交换结构 105 的地址空间中所包括（但是还没有被分配）的地址。在这样的实施例中，分配给边缘设备 162 的地址可以被选择为使得该地址并不被分配给交换结构 105 中的任何交换结构设备。

[0057] 在一些实施例中，响应于该请求，交换结构设备 130 能够针对边缘设备 162 定义标签，使得该标签能够被用来将数据路由至边缘设备 162。例如，交换结构 130 能够定义标签 [EL162 → 162, 经由 130]，其指示能够经由交换结构设备 130 到达边缘设备 162。在一些实施例中，分配给边缘设备 162 的地址可以包括在针对边缘设备 162 所定义的标签中。在这样的实施例中，该标签例如可以是 [EL162 → 10.1.3.1, 经由 130]。交换结构设备 130 随后可以被配置为将针对边缘设备 162 的标签存储在例如交换结构设备 130 中的路由表（例如，图 3 中的交换结构设备 300 中的路由表 325）中。在一些实施例中，针对边缘设备 162 的标签例如能够被添加至路由表中的 LFIB。LFIB 中的标签可以与交换结构设备 130 中将交换结构设备 130 连接至边缘设备 162 的端口相关联并且被用来识别该端口。

[0058] 在一些实施例中，交换结构设备 130 能够被配置为（连同所分配的地址一起）向边缘设备 162 发送标签（例如，经由信号 194）。在一些实施例中，在针对边缘设备 162 分配地址并定义标签时，交换结构设备 130 能够被配置为向边缘设备网络 145 中的耦合至交换结构 105 的其它边缘设备（例如，边缘设备 150、164、170、180）、和 / 或向交换结构 105 中的其它交换结构设备声明该地址和标签。特别地，交换结构设备 130 能够被配置为经由交换结构 105 中的交换结构设备向其它边缘设备发送广告（例如，消息），该广告包括边缘设备 162 的路由信息（例如，边缘设备 162 的地址、边缘设备 162 的标签、以及将交换结构设备 130 作为到达边缘设备 162 的“下一跳”的标识符）。例如，该广告可以包括诸如 10.1.3.1、EL162、下一跳 :130 之类的路由信息。作为结果，使得能够在交换结构 105 的其它交换结构设备和 / 或边缘设备网络 145 的其它边缘设备处获得与去往边缘地址 162 的路由相关联的信息（包括边缘设备 162 的地址、标签和下一跳）。

[0059] 在一些实施例中，作为交换结构设备 130 向其它边缘设备发送边缘设备 162 的路由信息的替换形式，边缘设备 162 可以被配置为向边缘设备网络 145 的其它边缘设备发送其路由信息。在这样的实施例中，边缘设备 162 能够根据诸如 BGP 之类的基于网络地址的协议，向其它边缘设备发送路由信息（例如，边缘设备 162 的地址、边缘设备 162 的标签、以及将交换结构设备 130 作为到达边缘设备 162 的“下一跳”的标识符）。在一些实施例中，该 BGP 消息能够在边缘设备网络 145 的控制平面（未示出）内被发送，该边缘设备网络 145 的控制平面独立于交换结构 105 的控制平面（即，不经由交换结构 105）。在这样的实施例中，例如，边缘设备 162 能够在不经过交换结构 105 的情况下向边缘设备 180 发送包括路由信息的 BGP 消息（在图 1 中示为信号 199）。在其它实施例中，该 BGP 消息能够通过交换结构 105 的控制平面被发送（即，经由交换结构 105）。在这样的实施例中，例如，边缘设备 162 能够经由交换结构设备 130、140 向边缘设备 180 发送包括路由信息的 BGP 消息（图 1 中未示出）。

[0060] 如图 1 所示的示例，当边缘设备 180 试图与边缘设备 162 进行通信时，边缘设备 180 能够首先基于所接收的边缘设备 162 的路由信息，将交换结构设备 130 确定为到达边缘设备 162 的下一跳。边缘设备 180 随后能够向交换结构设备 140 发送请求（在图 1 中示为信号 193）。该请求例如可以是用于求解去往交换结构设备 130 的路线的 ARP 请求。响

应于该 ARP 请求,交换结构设备 140 能够进行检查以确定交换结构设备 140 与交换结构设备 130 之间是否存在 MPLS 连接(例如,MPLS 隧道)。在确定交换结构设备 140 与交换结构设备 130 之间存在 MPLS 隧道(例如,图 1 中的 MPLS 隧道 190)之后,交换结构设备 140 能够针对该 MPLS 隧道定义标签。例如,交换结构设备 140 能够针对 MPLS 隧道 190 定义标签 [FL140-130:140 到 130]。交换结构设备 140 随后能够将该标签添加至例如交换结构设备 140 的路由表中的 LFIB。LFIB 中的该标签能够与交换结构设备 140 中经由 MPLS 隧道 190 将交换结构设备 140 连接至交换结构设备 130 的端口相关联、并且被用来识别该端口。

[0061] 在针对 MPLS 隧道 190 定义了该标签之后,交换结构设备 140 能够下将该标签发送至边缘设备 180(图 1 中示为信号 195)。在一些实施例中,交换结构设备 140 例如能够向边缘设备 180 发送交换结构设备 130 的熵标签和 / 或 (多个) 其它交换结构设备的 (多个) 熵标签。这样的熵标签例如可以包括与 MPLS 隧道 190 相关联的负载平衡信息(例如,关于交换结构 105 内的 (多个) 其它 MPLS 隧道的负载平衡信息)。在一些实施例中,交换结构设备 140 能向边缘设备 180 发送与交换结构设备 130、交换结构设备 140 和 / 或 (多个) 其它交换结构设备相关联的 (多个) 其它度量值。因此,边缘设备 180 能够基于 (多个) 熵标签和 / 或与交换结构设备 130、交换结构设备 140 和 / 或 (多个) 其它交换结构设备相关联的 (多个) 其它度量值,决定是否经由交换结构设备 140 和 130 或者任意其它可替换路线(图 1 中未示出)而与边缘设备 162 进行通信。

[0062] 在接收到针对 MPLS 隧道 190 的标签和 / 或附加信息之后,边缘设备 180 能够使用标签栈 <FL140-130, EL162> 来建立与边缘设备 162 的 MPLS 连接。例如,边缘设备 180 能够通过交换结构设备 140、MPLS 隧道 190 和交换结构设备 130 向边缘设备 162 发送数据单元(例如,数据分组、数据单位、数据帧)。特别地,边缘设备 180 的隧道模块(例如,图 2 的边缘设备 200 的隧道模块 225)能够利用标签栈 <FL140-130, EL162> 对该数据单元进行封装。首先,标签 EL162 能够在边缘设备 180 的隧道模块处被用来封装该数据单元以定义该数据单元的第一封装。接下来,标签 FL140-130 能够在边缘设备 180 的隧道模块处被用来对数据单元的第一封装进行封装以定义该数据单元的第二封装。边缘设备 180 能够将该数据单元的第二封装作为封装数据发送至交换结构设备 140。

[0063] 当接收到利用标签栈 <FL140-130> 进行封装的数据单元(即,该数据单元的第二封装)时,交换结构设备 140 的隧道模块(例如,图 3 中的交换结构设备 300 的隧道模块 345)能够从数据单元的第二封装中解封装标签 FL140-130,以得到数据单元的第一封装。基于标签 FL140-130,交换结构设备 140 能够确定经由 MPLS 隧道 190 向交换结构设备 130 发送所解封装的数据(即,该数据单元的第一封装)。例如,交换结构设备 140 能够基于端口的标识符与交换结构设备 140 的路由表中的标签 FL140-130 相关联而确定该端口用于发送该数据单元的第一封装。

[0064] 当接收到利用标签 EL162 进行封装的数据单元(即,该数据单元的第一封装)时,交换结构设备 130 的隧道模块(例如,图 3 中的交换结构设备 300 的隧道模块 345)能够从该数据单元的第一封装中解封装标签 EL162 以得到原始数据单元。基于标签 EL162,交换结构设备 130 能够确定将该数据单元发送至边缘设备 162。例如,交换结构设备 130 能够基于端口的标识符与交换结构设备 130 的路由表中的标签 EL162 相关联而确定该端口用于发送该数据单元。因此,边缘设备 180 能够通过交换结构 105、经由 MPLS 隧道向边缘设备 162

发送该数据单元。

[0065] 作为图 1 所示的另一个示例,当边缘设备 164 尝试与边缘设备 162 进行通信时,边缘设备 164 能够首先基于所接收的边缘设备 162 的路由信息,将交换结构设备 130 确定为到达边缘设备 162 的下一跳。在一些实施例中,类似于以上所描述的边缘设备 180,边缘设备 164 能够向交换结构设备 130 发送 ARP 请求(在图 1 中示为信号 196),以求解去往交换结构设备 130 的路线。响应于该 ARP 请求,交换结构设备 130 能够向边缘设备 164 返回响应(在图 1 中示为信号 198),该响应包括例如空标签。这样的空标签从交换结构设备 130 指示交换结构设备之间的 MPLS 隧道不被用来到达交换设备 130。边缘设备 164 随后能够以类似于以上所描述的边缘设备 180 的方法,使用标签栈<空标签,EL162>对意图在于边缘设备 162 的数据单元(例如,数据分组、数据单位、数据帧)进行封装。边缘设备 164 随后能够以类似于以上所描述的边缘设备 180 的方法,经由交换结构设备 130 将所封装的数据单元发送到边缘设备 162。在一些其它实施例中,如果边缘设备 164 确定其直接耦合至(作为到达边缘设备 162 的下一跳的)交换结构设备 130,则发送 ARP 请求以求解去往交换结构设备 130 的路线的步骤可以被省略或跳过。在这样的实施例中,边缘设备 164 能够使用标签 EL162 对意图在于边缘设备 162 的数据进行封装,并且随后将所封装的数据发送至交换结构设备 130。当从边缘设备 164 接收到(例如,使用标签栈<空标签,EL162>或标签 EL162 所封装的)该封装数据时,交换结构设备 130 的隧道模块能够对该封装数据进行解封装以得到意图在于边缘设备 162 的数据,并且随后将该数据发送至边缘设备 162。

[0066] 图 4 是图示根据实施例的用于经由交换结构在边缘设备之间进行通信的方法 400 的流程图。方法 400 能够在第一边缘设备处执行,该第一边缘设备在结构和功能上类似于关于图 1 和 2 所示出并描述的边缘设备(例如,边缘设备 150、162、164、170、180、200)。特别地,该第一边缘设备能够耦合至交换结构的交换结构设备(例如,图 1 中的交换结构 105 的交换结构设备 110、120、130 或 140、图 3 中的交换结构 300)。该交换结构设备(例如,经由交换结构的(多个)其它交换结构设备)可操作地耦合至第二边缘设备。

[0067] 在一些实施例中,类似于参考图 2 所示出并描述的边缘设备 200,第一边缘设备可以包括处理器(例如,处理器 220)和存储器(例如,存储器 210)。该存储器例如可以是非瞬时处理器可读介质。表示用于执行方法 400 的指令的代码能够被存储在第一边缘设备中所包括的非瞬时处理器可读介质中并且由第一边缘设备的处理器所执行。该代码包括要由该处理器所执行使得第一边缘设备对图 4 中所图示并且如以下所描述的功能进行操作的代码。

[0068] 在 402,第一边缘设备能够在第一时间根据自动发现协议向交换结构的交换结构设备发送请求。如以上所讨论的,该自动发现协议例如可以是 DHCP。该请求(例如,DHCP 请求)可以是针对要分配给第一边缘设备的地址和/或要对第一边缘设备定义的标签(例如,以上参考图 1 所描述的示例中的标签 EL162)的请求。该标签可以被其它边缘设备和/或交换结构设备用来将数据路由至第一边缘设备。

[0069] 在 404,第一边缘设备能够在第一时间之后的第二时间并且响应于该请求、根据自动发现协议从交换结构设备接收响应。例如,第一边缘设备能够从交换结构设备接收 DHCP 响应,该 DHCP 响应包括分配给第一边缘设备的 IP 地址。在一些实施例中,该响应可以包括针对第一边缘设备所定义的标签。在一些实施例中,该响应可以包括与将数据路由至第一

边缘设备相关联的其它路由信息。

[0070] 在 406, 第一边缘设备能够向第二边缘设备发送包括该标签的消息或信号, 该标签被包括在该响应中, 使得基于该标签与第二边缘设备建立 MPLS 隧道。如以上所讨论的, 第一边缘设备例如能够使用例如 BGP 消息向第二边缘设备发送该标签。可替换地, 交换结构设备能够将包括该标签的广告 (例如, 在 (多个) 消息或 (多个) 信号中) 发送至交换结构的其它交换结构设备, 后者进而能够将该标签发送至包括第二边缘设备在内的其它边缘设备。

[0071] 如以上所详细描述的, 在接收到该标签之后, 第二边缘设备能够发起建立 MPLS 隧道以到达第一边缘设备。特别地, 第二边缘设备能够基于该标签而确定针对第一边缘设备的下一跳的交换结构设备 (该示例中的交换结构设备)。第二边缘设备随后能够向相关联的交换结构设备 (即, 将第二边缘设备可操作地耦合至交换结构的交换结构设备) 发送 ARP 请求, 以求解出去往下一跳的交换结构设备的路线。响应于该 ARP 请求, 第二边缘设备能够接收识别将关联于第二边缘设备的交换结构设备与关联于第一边缘设备的交换结构设备相连接的 MPLS 隧道 (例如, 图 1 中将交换结构设备 140 和交换结构设备 130 相连接的 MPLS 隧道 190) 的标签 (例如, 以上关于图 1 所描述的示例中的标签 FL140-130)。

[0072] 在 408, 第二边缘设备能够使用包括第一边缘设备的标签和 MPLS 隧道的标签的标签栈来发送数据单元。基于该标签栈, 该数据单元能够通过与第二边缘设备相关联的交换结构设备、MPLS 隧道以及与第一边缘设备相关联的交换结构设备而以该顺序被传输至第一边缘设备。特别地, 能够在与第二边缘设备相关联的交换结构设备处使用 MPLS 隧道的标签 (例如, FL140-130), 将数据单元从与第二边缘设备相关联的交换结构设备路由至与第一边缘设备相关联的交换结构设备。在与第一边缘设备相关联的交换结构设备处, 能够使用第一边缘设备的标签 (例如, EL162) 来将该数据单元从与第一边缘设备相关联的交换结构设备路由至第一边缘设备。作为第二边缘设备使用 MPLS 隧道将该数据单元发送至第一边缘设备的结果, 第一边缘设备能够在第二时间之后的第三时间通过该 MPLS 隧道从第二边缘设备接收该数据单元。

[0073] 虽然这里关于图 1 示出并描述了边缘设备网络的边缘设备与交换结构的交换结构设备进行交互以得到关联于和其它边缘设备的通信的路由信息 (例如, 识别 MPLS 隧道和其它边缘设备的标签), 但是在其它实施例中, 一个或多个路线反射器能够向边缘设备网络的边缘设备提供路由信息。这样的路线反射器能够在边缘设备、交换结构设备或者与边缘设备网络和交换结构分离的设备处实施。例如, 路线反射器能够被配置为从边缘设备接收 BGP 消息, 该 BGP 消息包括该边缘设备的所更新的路由信息。该路线反射器能够被配置为将该 BGP 消息转发至其它边缘设备。因此, 边缘设备网络中的每个边缘设备保持与路线反射器的 BGP 连接, 而并不保持与其它边缘设备的任何连接。作为另一个示例, 路线反射器能够用作边缘设备网络的 DHCP 服务器和 / 或 ARP 服务器。因此, 每个交换结构设备能够将其相关联的边缘设备所接收的 DHCP 请求和 / 或 ARP 请求转发至路线反射器 (例如, 经由交换结构的其它交换结构设备), 并且从路线反射器接收 DHCP 响应和 / 或 ARP 响应。作为又另一个示例, 路线反射器可以用作中央控制器, 其存储与交换结构相关联的路由信息并且将该路由信息提供至交换结构的交换结构设备。因此, 每个交换结构设备与路线反射器进行通信, 以获得 MPLS 隧道和 / 或边缘设备的标签。交换结构设备随后能够使用从路线反射器

所接收的标签,以通过交换结构路由数据。

[0074] 这里所描述的一些实施例涉及一种具有非瞬时计算机可读介质(也被称作非瞬时处理器可读介质)的计算机存储产品,该非瞬时计算机可读介质在其上具有用于执行各种计算机所实施的操作的指令或计算机代码。就其本身并不包括瞬时传播信号(例如,在诸如空间或线缆的传输介质上承载信息的传播电磁波)的意义而言,计算机可读介质(或处理器可读介质)是非瞬时的。媒介和计算机代码(也被称作代码)可以是出于一个或多个具体目的而设计和构造的那些媒介和计算机代码。非瞬时计算机可读媒介的示例包括但并不局限于:磁性存储媒介,诸如硬盘、软盘和磁带;光学存储媒介,诸如紧致盘/数字视频盘(CD/DVD)、紧致盘只读存储器(CD-ROM)和全息设备;诸如光盘的磁性光学存储媒介;载波信号处理模块;以及被专门配置为存储并执行计算机代码的硬件设备,诸如专用集成电路(ASIC)、可编程逻辑设备(PLD)、只读存储器(ROM)和随机访问存储器(RAM)设备。这里所描述的其它实施例涉及一种计算机程序产品,其例如可以包括这里所讨论的指令和/或计算机代码。

[0075] 计算机代码的示例包括但并不局限于微代码或微指令、诸如由编译器所产生的机器指令、用来产生 web 服务的代码,以及包含由计算机使用解释器执行的更高级指令的文件。例如,实施例可以使用命令编程语言(例如,C、Fortran 等)、功能编程语言(Haskell、Erlang 等)、逻辑编程语言(例如,Prolog)、面向对象编程语言(例如,Java、C++ 等)或者其它适当编程语言和/或开发工具来实施。计算机代码另外的示例包括但并不局限于控制信号、加密代码和压缩代码。

[0076] 虽然以上已经对各个实施例进行了描述,但是应当理解的是,它们仅是作为示例而非限制被给出。在以上所描述的方法指示某些事件以某个顺序进行的情况下,某些事件的排序可以有所修改。此外,某些事件在可能时可以在并行处理中同时执行,以及如以上所描述的那样顺序执行。

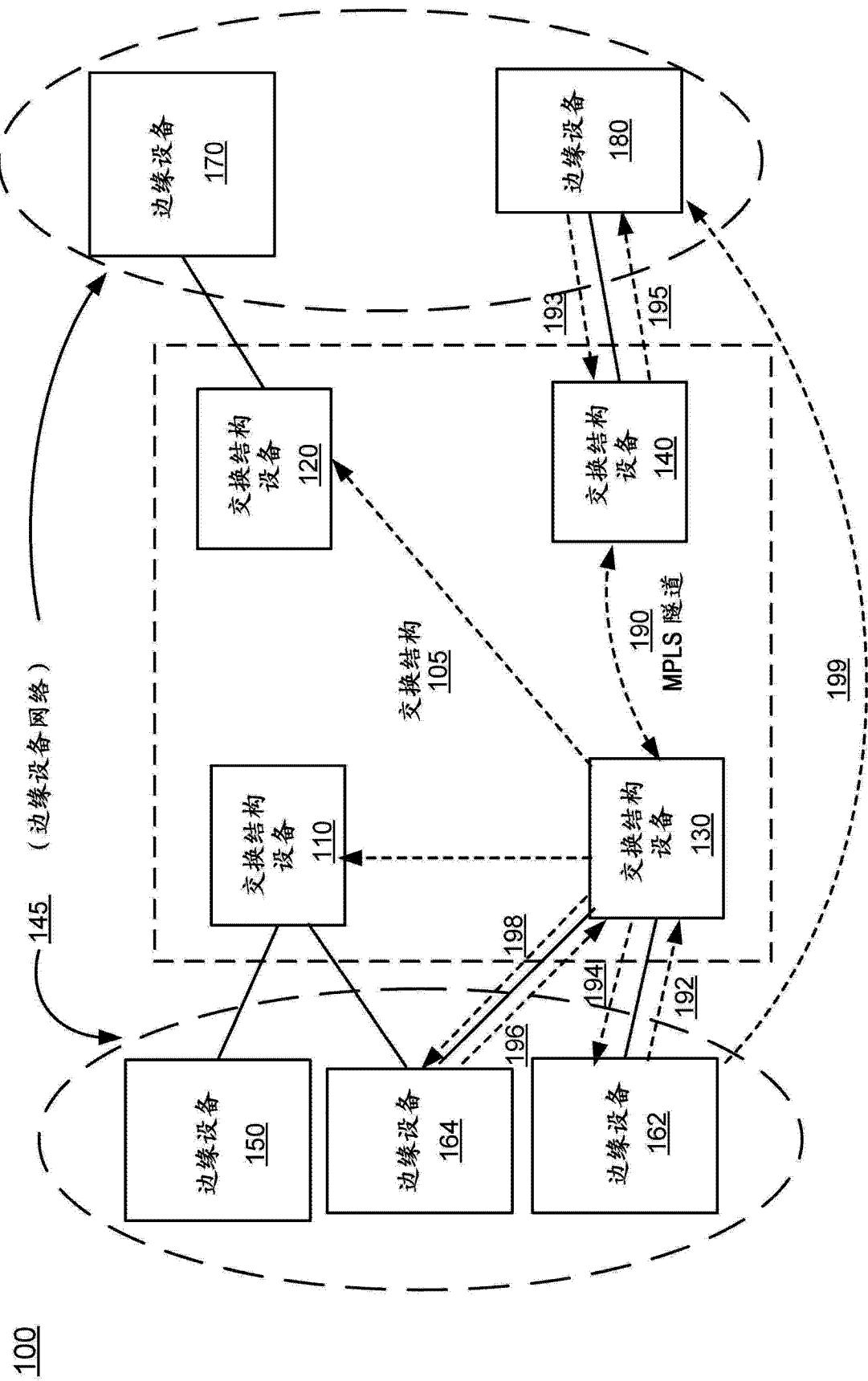


图 1

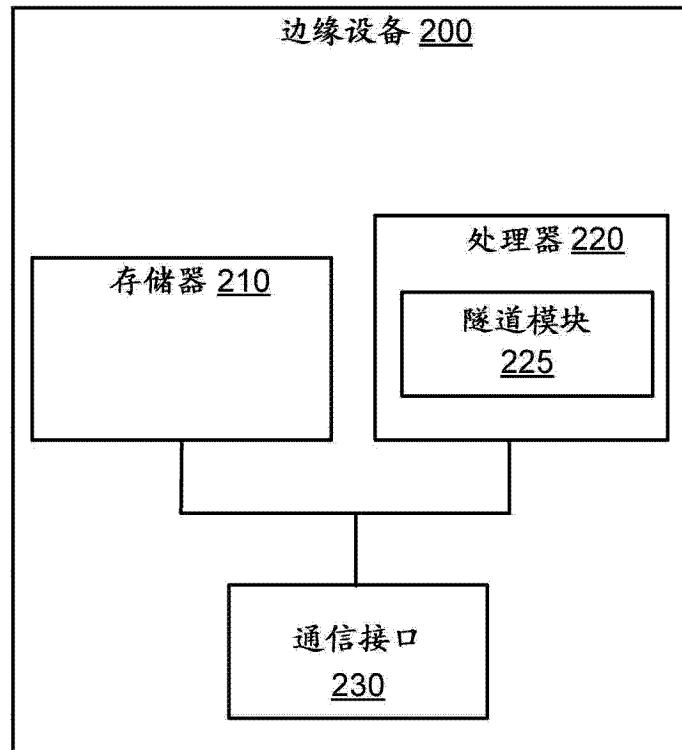


图 2

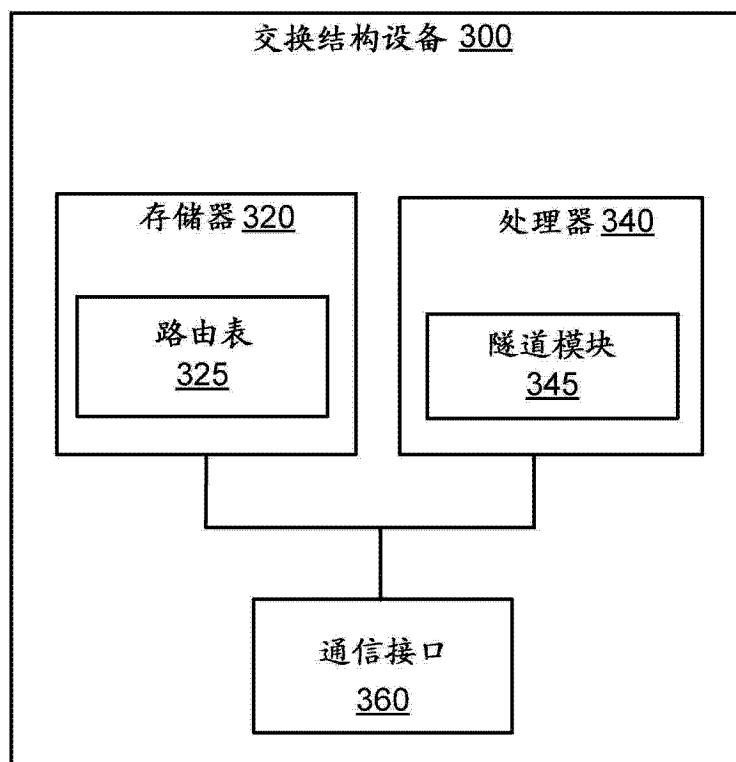


图 3

400

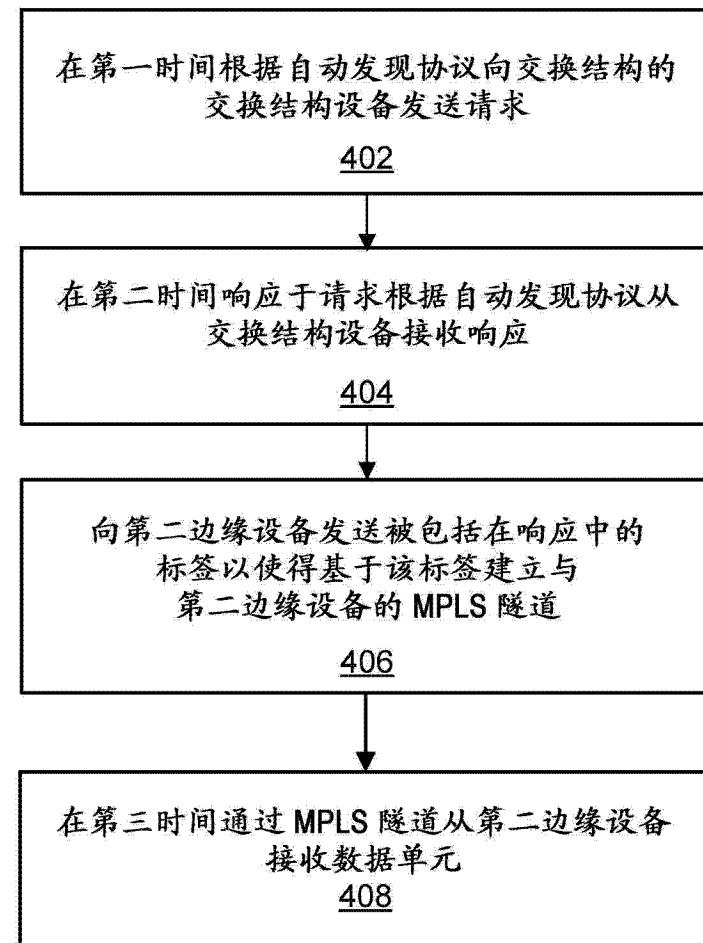


图 4