



(12) 发明专利

(10) 授权公告号 CN 108702328 B

(45) 授权公告日 2021.04.09

(21) 申请号 201680081901.5  
 (22) 申请日 2016.04.01  
 (65) 同一申请的已公布的文献号  
 申请公布号 CN 108702328 A  
 (43) 申请公布日 2018.10.23  
 (30) 优先权数据  
 62/295,452 2016.02.15 US  
 (85) PCT国际申请进入国家阶段日  
 2018.08.15  
 (86) PCT国际申请的申请数据  
 PCT/IB2016/051890 2016.04.01  
 (87) PCT国际申请的公布数据  
 WO2017/141079 EN 2017.08.24  
 (73) 专利权人 瑞典爱立信有限公司  
 地址 瑞典斯德哥尔摩  
 (72) 发明人 乌马·S·丘杜里  
 阿马纳特·安米雷迪  
 叶夫根尼·坦楚拉  
 (74) 专利代理机构 北京德琦知识产权代理有限公司 11018  
 代理人 严芬 宋志强

(51) Int.Cl.  
 H04L 12/721 (2006.01)  
 H04L 12/723 (2006.01)  
 H04L 12/707 (2006.01)  
 H04L 12/717 (2006.01)  
 (56) 对比文件  
 US 2014/0169370 A1,2014.06.19  
 US 2013/0077625 A1,2013.03.28  
 CN 104378292 A,2015.02.25  
 CN 104243311 A,2014.12.24  
 US 2014/0355419 A1,2014.12.04  
 US 7664877 B1,2010.02.16  
 CN 104980350 A,2015.10.14  
 CN 104718730 A,2015.06.17  
 US 2015/0109904 A1,2015.04.23  
 C.Filsfils.Segment Routing with MPLS data plane.《IEEE草案》.2015,  
 S.Previdi.IS-IS Extensions for Segment Routing.《IEEE草案》.2015,  
 C.Filsfils.Segment Routing interoperability with LDP.《IEEE草案》.2015,

审查员 孙亚娜

权利要求书3页 说明书18页 附图10页

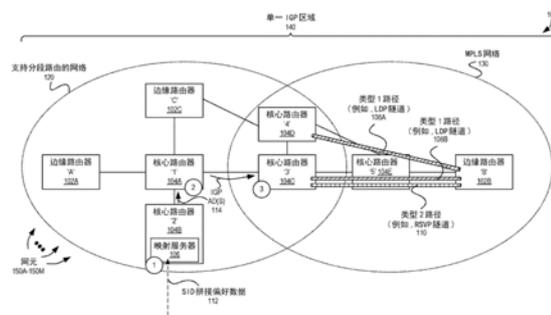
(54) 发明名称

用于穿越分段路由和MPLS网络的业务的灵活路径拼接和选择的IS-IS扩展

(57) 摘要

路由器在利用中间系统到中间系统 (IS-IS) 的网络的分段路由 (SR) 网络部分和多协议标签交换 (MPLS) 网络部分中操作。路由器接收由映射服务器发起的IS-IS通告消息,该IS-IS通告消息包括子类型长度值 (子TLV) 元素,该子类型长度值 (子TLV) 元素识别要由路由器从SR网络部分接收的可识别业务集合的跨MPLS网络部分的路径的偏好类型。路由器至少部分地基于该子TLV元素来识别用于可识别业务集合的跨MPLS网络部

分的多个可用路径中的一个路径,以及配置其转发平面以利用所识别的一个路径用于可识别业务集合。IS-IS通告消息可以是IS-ISTLV,例如SID/标签绑定TLV 149或多拓扑SID/标签绑定TLV150。



CN 108702328 B

1. 一种在实现路由器网元(104C)的设备(602/604/606)中的方法,其中,所述路由器网元是包括分段路由SR网络部分(120)和多协议标签交换MPLS网络部分(130)在内的网络(140)的一部分,其中,所述路由器网元是所述SR网络部分和所述MPLS网络部分的一部分,以及所述网络还利用中间系统到中间系统IS-IS协议来交换路由信息,所述方法包括:

在所述路由器网元处接收由映射服务器(106)发起的IS-IS通告消息(114),其中,所述IS-IS通告消息包括子类型长度值子TLV元素(200),所述子TLV元素识别用于要从所述SR网络部分接收的可识别业务集合的跨所述MPLS网络部分的路径的偏好类型,并且其中,所述路径的偏好类型是资源预留协议RSVP隧道、静态MPLS路径或标签边界网关协议BGP路径;

由所述路由器网元至少部分地基于所述子TLV元素来识别用于所述可识别业务集合的跨所述MPLS网络部分的多个可用路径(108B/110)中的一个路径(110);以及

由所述路由器网元配置所述路由器网元的转发平面(626)以利用所识别的一个路径来转发能够被确定为属于所述可识别业务集合的从所述SR网络部分接收的任何业务。

2. 根据权利要求1所述的方法,还包括:

在所述路由器网元处从在所述SR网络部分中而不在所述MPLS网络部分中操作的另一网元(104A)接收属于所述可识别业务集合的一个或多个分组;以及

由所述路由器网元根据在所述MPLS网络部分上所识别的一个路径来发送所述一个或多个分组。

3. 根据权利要求1所述的方法,其中,所述子TLV元素由IS-IS SID/标签绑定TLV(400)来承载。

4. 根据权利要求1所述的方法,其中,所述子TLV元素由IS-IS多拓扑SID/标签绑定TLV(450)来承载。

5. 根据权利要求1所述的方法,其中:

所述子TLV元素包括分段标识符SID拼接偏好值(206),所述SID拼接偏好值识别所述路径的偏好类型而不唯一地识别特定路径;以及

识别所述一个路径是基于所述SID拼接偏好值。

6. 根据权利要求1所述的方法,其中:

所述子TLV元素包括LSP值(254)和隧道标识符ID值(256);

所述子TLV元素识别跨所述MPLS网络部分的特定路径;以及

识别所述一个路径是基于所述LSP值和所述隧道ID值。

7. 一种具有指令(620/650)的非暂时性计算机可读存储介质(618/648),所述指令在由电子设备(602/604/606)的一个或多个处理器(612/642)执行时,使所述电子设备实现路由器网元,所述路由器网元执行操作,并且所述路由器网元是包括分段路由SR网络部分和多协议标签交换MPLS网络部分在内的网络的一部分,其中,所述路由器网元要作为所述SR网络部分和所述MPLS网络部分的一部分,以及所述网络还利用中间系统到中间系统IS-IS协议来交换路由信息,所述操作包括:

接收由映射服务器发起的IS-IS通告消息,其中,所述IS-IS通告消息包括子类型长度值子TLV元素,所述子TLV元素识别用于要从所述SR网络部分接收的可识别业务集合的跨所述MPLS网络部分的路径的偏好类型,并且其中,所述路径的偏好类型是资源预留协议RSVP隧道、静态MPLS路径或标签边界网关协议BGP路径;

至少部分地基于所述子TLV元素来识别用于所述可识别业务集合的跨所述MPLS网络部分的多个可用路径中的一个路径;以及

配置所述路由器网元的转发平面以利用所识别的一个路径来转发能够被确定为属于所述可识别业务集合的从所述SR网络部分接收的任何业务。

8. 根据权利要求7所述的非暂时性计算机可读存储介质,其中,所述操作还包括:

从在所述SR网络部分中而不在所述MPLS网络部分中操作的另一网元接收属于所述可识别业务集合的一个或多个分组;以及

根据在所述MPLS网络部分上所识别的一个路径来发送所述一个或多个分组。

9. 根据权利要求7所述的非暂时性计算机可读存储介质,其中,所述子TLV元素由IS-ISSID/标签绑定TLV来承载。

10. 根据权利要求7所述的非暂时性计算机可读存储介质,其中,所述子TLV元素由IS-IS多拓扑SID/标签绑定TLV来承载。

11. 根据权利要求7所述的非暂时性计算机可读存储介质,其中:

所述子TLV元素包括分段标识符SID拼接偏好值,所述SID拼接偏好值识别所述路径的偏好类型但不唯一地识别特定路径;以及

识别所述一个路径是基于所述SID拼接偏好值。

12. 根据权利要求7所述的非暂时性计算机可读存储介质,其中:

所述子TLV元素包括LSP值和隧道标识符ID值;

所述子TLV元素识别跨所述MPLS网络部分的特定路径;以及

识别所述一个路径是基于所述LSP值和所述隧道ID值。

13. 一种用于实现路由器网元的设备,包括:

一个或多个处理器;以及

具有指令的非暂时性计算机可读存储介质,所述指令在由所述一个或多个处理器执行时,使所述设备实现将作为包括分段路由SR网络部分和多协议标签交换MPLS网络部分在内的网络的一部分的路由器网元,其中,所述路由器网元要作为所述SR网络部分和所述MPLS网络部分的一部分,以及所述网络要利用中间系统到中间系统IS-IS协议来交换路由信息,所述路由器网元适于:

接收由映射服务器发起的IS-IS通告消息,其中,所述IS-IS通告消息包括子类型长度值子TLV元素,所述子TLV元素用于识别对要从所述SR网络部分接收的可识别业务集合的跨所述MPLS网络部分的路径的偏好类型,并且其中,所述路径的偏好类型是资源预留协议RSVP隧道、静态MPLS路径或标签边界网关协议BGP路径;

至少部分地基于所述子TLV元素来识别用于所述可识别业务集合的跨所述MPLS网络部分的多个可用路径中的一个路径;以及

配置所述路由器网元的转发平面以利用所识别的一个路径来转发能够被确定为属于所述可识别业务集合的从所述SR网络部分接收的任何业务。

14. 根据权利要求13所述的设备,其中,所述路由器网元还适于:

从在所述SR网络部分中而不在所述MPLS网络部分中操作的另一网元接收属于所述可识别业务集合的一个或多个分组;以及

根据在所述MPLS网络部分上所识别的一个路径来发送所述一个或多个分组。

15. 根据权利要求13所述的设备,其中,所述子TLV元素由ISIS SID/标签绑定TLV来承载。

16. 根据权利要求13所述的设备,其中,所述子TLV元素由ISIS多拓扑SID/标签绑定TLV来承载。

17. 根据权利要求13所述的设备,其中:

所述子TLV元素包括分段标识符SID拼接偏好值,所述SID拼接偏好值识别所述路径的偏好类型但不唯一地识别特定路径;以及

识别所述一个路径是基于所述SID拼接偏好值。

18. 根据权利要求13所述的设备,其中:

所述子TLV元素包括LSP值和隧道标识符ID值;

所述子TLV元素识别跨所述MPLS网络部分的特定路径;以及

识别所述一个路径是基于所述LSP值和所述隧道ID值。

19. 一种具有计算机程序逻辑的计算机程序产品,所述计算机程序逻辑被配置为实施权利要求1至6中任一项所述的方法。

20. 一种用于实现路由器网元(104C)的设备(570),所述路由器网元(104C)要作为包括分段路由SR网络部分和多协议标签交换MPLS网络部分在内的网络的一部分,其中,所述路由器网元要作为所述SR网络部分和所述MPLS网络部分的一部分,以及所述网络要利用中间系统到中间系统IS-IS协议来交换路由信息,所述设备包括:

模块(572),适于接收由映射服务器发起的IS-IS通告消息,其中,所述IS-IS通告消息包括子类型长度值子TLV元素,所述子TLV元素识别用于要从所述SR网络部分接收的可识别业务集合的跨所述MPLS网络部分的路径的偏好类型,并且其中,所述路径的偏好类型是资源预留协议RSVP隧道、静态MPLS路径或标签边界网关协议BGP路径;

模块(574),适于至少部分地基于所述子TLV元素,识别用于所述可识别业务集合的跨所述MPLS网络部分的多个可用路径中的一个路径;以及

模块(576),适于配置所述路由器网元的转发平面以利用所识别的一个路径来转发能够被确定为属于所述可识别业务集合的从所述SR网络部分接收的任何业务。

21. 一种用于实现路由器网元的设备,所述路由器网元要作为包括分段路由SR网络部分和多协议标签交换MPLS网络部分在内的网络的一部分,其中,所述路由器网元要作为所述SR网络部分和所述MPLS网络部分的一部分,以及所述网络要利用中间系统到中间系统IS-IS协议来交换路由信息,所述设备适于:

接收由映射服务器发起的IS-IS通告消息,其中,所述IS-IS通告消息包括子类型长度值子TLV元素,所述子TLV元素识别用于要从所述SR网络部分接收的可识别业务集合的跨所述MPLS网络部分的路径的偏好类型,并且其中,所述路径的偏好类型是资源预留协议RSVP隧道、静态MPLS路径或标签边界网关协议BGP路径;

至少部分地基于所述子TLV元素来识别用于所述可识别业务集合的跨所述MPLS网络部分的多个可用路径中的一个路径;以及

配置所述路由器网元的转发平面以利用所识别的一个路径来转发能够被确定为属于所述可识别业务集合的从所述SR网络部分接收的任何业务。

## 用于穿越分段路由和MPLS网络的业务的灵活路径拼接和选择的IS-IS扩展

[0001] 相关申请的交叉引用

[0002] 本申请要求于2016年2月15日提交的美国临时申请No.62/295,452的权益,其通过引用并入本文中。

### 技术领域

[0003] 本发明的实施例涉及计算机联网领域;并且更具体地涉及用于穿越分段路由(SR)和多协议标签交换(MPLS)网络的业务的灵活路径拼接和选择的中间系统到中间系统(IS-IS)扩展。

### 背景技术

[0004] SR节点借助于将SR报头前置于分组而通过受控指令集(称为分段)来控制分组。分段可以表示任何指令(拓扑的或基于服务的)。SR允许流被强制通过任何拓扑路径和服务链,同时仅在通向SR域的入口节点处维持针对每个流的状态。

[0005] SR架构可被直接应用于MPLS数据平面,而不会改变转发平面。互联网工程任务组(IETF)互联网草案(“与LDP的分段路由互操作性”、“draft-ietf-spring-segment-routing-ldp-interop-00”,日期为2015年10月14日)描述了SR可以在部署标签分发协议(LDP)的网络中以及在支持SR和不支持SR的节点共存的情况下如何操作。

[0006] 一种用于使SR域和LDP域互操作的技术利用被称为映射服务器的功能。映射服务器可以被配置为通告只能由支持SR的路由器所理解的映射信息,并使支持SR的路由器在MPLS数据平面中安装相关节点分段,就像该节点分段已经由节点它们自己通告过一样。

### 发明内容

[0007] 描述了在实现路由器网元的设备中的方法。路由器网元是包括SR网络部分和MPLS网络部分在内的网络的一部分。路由器网元是SR网络部分和MPLS网络部分的一部分,并且网络还利用内部网关协议(IGP)来交换路由信息。IGP是中间系统到中间系统(ISIS)。该方法包括:在路由器网元处接收由映射服务器发起的IGP通告消息。IGP通告消息包括:子类型长度值(子TLV)元素,其识别用于要从SR网络部分接收的可识别业务集合的跨MPLS网络部分的路径的偏好类型。该方法还包括:由路由器网元至少部分地基于所述子TLV元素来识别用于所述可识别业务集合的跨MPLS网络部分的多个可用路径中的一个路径。该方法还包括:由路由器网元来配置路由器网元的转发平面以利用所识别的一个路径来转发能够被确定为属于可识别业务集合的从SR网络部分接收的任何业务。

[0008] 在一些实施例中,该方法还包括:在路由器网元处从在SR网络部分中而不在MPLS网络部分中操作的另一网元接收属于可识别业务集合的一个或多个分组;以及由所述路由器网元根据在所述MPLS网络部分上所识别的一个路径来发送所述一个或多个分组。

[0009] 在一些实施例中,路径的偏好类型包括RSVP隧道、静态MPLS路径或标签边界网关

协议 (BGP) 路径。

[0010] 在一些实施例中,该子TLV元素由IS-IS SID/标签绑定TLV来承载,并且在一些实施例中,子TLV元素由IS-IS多拓扑SID/标签绑定TLV来承载。

[0011] 在一些实施例中,子TLV包括分段标识符 (SID) 拼接偏好值,其识别路径的偏好类型但不唯一地识别特定路径,并且对一个路径的识别是基于SID拼接偏好值的。

[0012] 在一些实施例中,子TLV包括LSP (标签交换路径) 值和隧道标识符 (ID) 值,子TLV识别跨MPLS网络部分的特定路径,并且对一个路径的识别是基于LSP值和隧道ID值的。

[0013] 非暂时性计算机可读存储介质具有指令,该指令在由电子设备的一个或多个处理器执行时,使得电子设备实现执行操作且要成为包括SR网络部分和MPLS网络部分在内的网络的一部分的路由器网元。路由器网元要作为SR网络部分和MPLS网络部分的一部分 (例如,在其内操作)。该网络还利用IGP来交换路由信息。IGP是IS-IS。操作包括接收由映射服务器发起的IGP通告消息。IGP通告消息包括:子类型长度值 (子TLV) 元素,其识别用于要从SR网络部分接收的可识别业务集合的跨MPLS网络部分的路径的偏好类型。该操作还包括:至少部分地基于子TLV元素来识别用于可识别业务集合的跨MPLS网络部分的多个可用路径中的一个路径。该操作还包括:配置路由器网元的转发平面以利用所识别的一个路径来转发能够被确定为属于可识别业务集合的从SR网络部分接收的任何业务。

[0014] 设备包括一个或多个处理器和非暂时性计算机可读存储介质。该非暂时性计算机可读存储介质具有指令,该指令在由一个或多个处理器执行时,使得设备实现执行操作且要成为包括SR网络部分和MPLS网络部分在内的网络的一部分的路由器网元。路由器网元要作为SR网络部分和MPLS网络部分的一部分 (例如,在其内操作)。该网络还利用IGP来交换路由信息。IGP是IS-IS。操作包括接收由映射服务器发起的IGP通告消息。IGP通告消息包括:子类型长度值 (子TLV) 元素,其识别用于要从SR网络部分接收的可识别业务集合的跨MPLS网络部分的路径的偏好类型。该操作还包括:至少部分地基于子TLV元素来识别用于可识别业务集合的跨MPLS网络部分的多个可用路径中的一个路径。该操作还包括:配置路由器网元的转发平面以利用所识别的一个路径来转发能够被确定为属于可识别业务集合的从SR网络部分接收的任何业务。

## 附图说明

[0015] 通过参考用于说明本发明实施例的以下描述和附图,可以最好地理解本发明。在附图中:

[0016] 图1是示出了根据一些实施例的利用用于穿越SR和MPLS网络的业务的灵活路径拼接和选择的在支持SR的网络和MPLS网络内通信耦合的路由器网元的系统的高级框图。

[0017] 图2示出了根据一些实施例的非特定偏好子TLV元素、用于OSPF的特定偏好子TLV元素以及用于IS-IS的特定偏好子TLV,其可以用于实现用于穿越SR和MPLS网络的业务的灵活路径拼接和选择。

[0018] 图3示出了根据一些实施例的OSPF扩展前缀范围TLV,其可以用于承载如图2所示的非特定偏好子TLV或用于OSPF的特定偏好子TLV。

[0019] 图4示出了根据一些实施例的专用IS-IS TLV 149和专用IS-IS TLV 150,其可以用于承载如图2所示的非特定偏好子TLV或用于IS-IS的特定偏好子TLV。

[0020] 图5A示出了根据一些实施例的用于实现用于穿越分段路由和MPLS网络的业务的灵活路径拼接和选择的两个操作流程。

[0021] 图5B示出了根据一些实施例的核心路由器104C的高级框图568,其在一些实施例中可以执行图5A的流程500。

[0022] 图5C示出了根据一些实施例的映射服务器106的高级框图578,其在一些实施例中可以执行图5A的流程550。

[0023] 图6A示出了根据一些实施例的示例网络内的网络设备 (ND) 之间的连接性以及ND的三种示例实现。

[0024] 图6B示出了根据一些实施例的用于实现专用网络设备的示例方式。

[0025] 图6C示出了根据一些实施例的可以耦合虚拟网元 (VNE) 的各种示例方式。图6D示出了根据一些实施例的在每个ND上具有单一网元 (NE) 的网络,且在该直观方案中,对比了(由传统路由器常用的)传统分布式方案与用于维护可达性和转发信息(也被称为网络控制)的集中式方案。

[0026] 图6E示出了根据一些实施例的每个ND实现单个NE的简单情况,但是集中控制平面将不同ND中的多个NE抽象为(以表示)虚拟网络之一中的单一NE。

[0027] 图6F示出了根据一些实施例的多个VNE在不同ND上实现且彼此耦合的情况,并且在该情况中,集中控制平面对这些多个VNE进行抽象使得它们在虚拟网络之一内表现为单一VNE。

## 具体实施方式

[0028] 在以下描述中,阐述了大量的具体细节,例如逻辑实现、操作码(opcode)、用于指定操作数的手段、资源分区/共享/复制实现、系统组件的类型和相互关系、以及逻辑分区/整合选择,以提供对本发明的更全面的理解。然而,本领域技术人员将意识到:本发明可以在没有这些具体细节的情况下实施。在其它实例中,并未详细示出控制结构、门级别电路和全软件指令序列,以不使本发明模糊。在使用所包括的描述的情况下,本领域普通技术人员将能够在不进行过度试验的情况下实现恰当的功能。

[0029] 说明书中对“一个实施例”、“实施例”、“示例实施例”等的引用指示了所描述的实施例可以包括特定特征、结构、或特性,但是每个实施例可以不必包括该特定特征、结构、或特性。此外,这种短语不必参考同一实施例。此外,当结合实施例来描述具体特征、结构或特性时,应认为结合其它实施例(不管是否是显式描述的)来实现这种特征、结构或特性是在本领域技术人员的知识内的。

[0030] 在本文中,括号中的文本和具有虚线边界(例如,长划线虚线、短划线虚线、点线虚线、以及点)的框可以用于示出向一些实施例添加附加特征的可选操作。然而,这种标注不应当被视为意味着:在某些实施例中,这些是仅有的选项或可选操作,和/或具有实线边界的框不是可选的。

[0031] 在以下描述和权利要求中,可以使用术语“耦合”和“连接”以及它们的派生词。应当理解:这些术语不意在作为彼此的同义词。“耦合”用于指示两个或更多个元素可以或不彼此直接物理或电学接触、彼此协作或交互。“连接”用于指示在彼此耦合的两个或更多个元素之间建立通信。

[0032] 如背景技术所示,存在通过被称为映射服务器功能来对区域中的SR域和LDP域进行处理的技术,其在IETF互联网草案“与LDP的分段路由互操作性”(draft-ietf-spring-segment-routing-ldp-interop-00)中描述,日期为2015年10月14日。

[0033] 通过映射服务器分段标识符(SID)来通告的映射服务器前缀(例如,开放最短路径优先(OSPF)中的扩展前缀范围TLV)表示区域/域中的非SR节点的标签信息。这使得能够在支持和提供SR和LDP的边界节点处将SR标签“拼接”到LDP标签。

[0034] 然而,仅描述了映射服务器前缀用于提供SR网络部分和LDP网络部分之间的互操作性。这在现实世界环境中产生问题,因为边界节点可能具有规定的其它MPLS客户端(即,除LDP之外),例如静态MPLS、标签BGP/bgp-3107、RSVP等。由于没有针对除LDP之外的任何事物来定义的用于处理映射服务器前缀的权威机制,因此不同的边界节点可能无法将SR路径拼接到特定的MPLS路径(即使它可以使用)。此外,没有机制供运营商通过当前定义的映射服务器SR扩展来指示在边界节点处要拼接哪个MPLS路径(或MPLS路径的类型)的偏好。

[0035] 因此,以下描述记载了用于实现用于穿越SR和MPLS网络的业务的灵活路径拼接和选择的技术。在一些实施例中,可以在IGP通告(例如OSPF或IS-IS消息)内指示路径类型偏好。在被SR网络部分和MPLS网络部分内工作的网元接收到时,路径类型偏好可以在针对从SR网络部分接收到的业务来确定通过MPLS网络部分的路径时使用路径类型偏好来选择遵循路径类型偏好的路径。因此,可以适应路径偏好以适当地将SR和MPLS网络之间的路径“拼接”在一起。

[0036] 在一些实施例中,路径类型偏好可以识别通过MPLS网络的路径或隧道的类型,例如通过指示对资源预留协议(RSVP)路径、静态MPLS路径、标签BGP路径等的偏好。

[0037] 在一些实施例中,路径类型偏好可以由SR网络部分中使用的映射服务器来提供,并且可以使运营商能够对MPLS网络部分和SR网络部分之间流动的业务行使其偏好。因此,在一些实施例中,可以使用IGP扩展以及SR架构的映射服务器功能来增强支持SR的网络部分与传统MPLS网络部分之间的互操作性(例如,使用RSVP、标签BGP、静态MPLS标签等)。此外,在一些实施例中,在与现有的与LDP网络进行拼接的方式相比时,不会导致向后兼容性问题。

[0038] 在一些实施例中,利用新的子TLV。子TLV可以用于OSPF SR扩展中的映射服务器TLV(IETF互联网草案“用于分段路由的OSPF扩展”(draft-ietf-ospf-segment-routing-extensions-05)),或者可以用于在IS-IS SR扩展中定义的映射服务器TLV(IETF互联网草案“用于分段路由的IS-IS扩展”(draft-ietf-isis-segment-routing-extensions-06))。新的子TLV可以提供偏好(如果有的话)以在SR域中提供映射服务器前缀时选择要提供的特定MPLS路径。这可以允许维持要求(例如,来自服务等级协议(SLA)),例如用于特定目的地的特定传输隧道的端到端路径上的带宽和/或延迟。

[0039] 因此,一些实施例可以使得能够在现有MPLS网络中无缝地部分部署SR而在多供应商部署环境中没有任何互操作性问题,并且可以使运营商能够维持某些前缀的SLA网络性能保证。

[0040] 图1是示出了根据一些实施例的包括在支持SR的网络120和MPLS网络130内通信耦合的路由器网元150A~150M在内的系统100的高级框图,该路由器网元150A~150M利用针对穿越SR和MPLS网络的业务的灵活路径拼接和选择。在该示例中,多个网元150A~150M包

括使用所示链路来耦合的三个边缘路由器102A~102C和多个核心路由器104A~104E。

#### [0041] SR到RSVP映射

[0042] 在该示例中,假设要在边缘路由器‘A’ 102A和边缘路由器‘B’ 102B之间建立端到端虚拟专用网络 (VPN) 服务 (L3/L2)。如图所示,只有网络的一部分是支持SR的120:边缘路由器‘A’ 102A、边缘路由器‘C’ 102C、核心路由器‘1’~‘4’ 104A~104D。另外,网络的不同部分具有提供的“传统的”LDP/RSVP:核心路由器‘3’~‘5’ 104C~104E和边缘路由器‘B’ 102B,即,MPLS网络130。因此,核心路由器‘3’ 104C和核心路由器‘4’ 104D能够支持SR和LDP/RSVP这二者。

[0043] 还假设:提供从SR网络120到MPLS网络130的边缘路由器“B” 102B的第一类型(例如,LDP隧道)的两个路径108A~108B,并且提供从SR网络120到MPLS网络130的边缘路由器‘B’ 102B的第二类型(例如,RSVP隧道)的一个路径110。

[0044] 在该示出的示例中,核心路由器‘2’ 104B被提供为用作映射服务器,并且因此在单一IGP区域140(例如,所有节点都参与的OSPF域)中通告核心路由器‘5’ 104E和边缘路由器‘B’ 102B的SR节点SID(也称为“标签”),因为这些网元不支持SR。

[0045] 首先,描述边缘路由器‘A’ 102A学习来自边缘路由器‘B’ 102B的服务标签的情况。为了从边缘路由器‘A’ 102A向边缘路由器‘B’ 102B传递业务,第一阶段包括使用从边缘路由器‘A’ 102A直到核心路由器‘3’ 104C为止的SR路径。例如,在边缘路由器‘A’ 102A处,边缘路由器‘B’ 102B的下一跳可以是核心路由器‘1’ 104A,即,边缘路由器‘B’ 102B的节点索引+核心路由器‘1’ 104A的SRGB(分段路由全局块)。

[0046] 当核心路由器‘3’ 104C接收到绑定(通过在SR OSPF扩展中定义的绑定SID)时,有可能用在其转发路径中可用的LDP标签和RSVP标签这二者来拼接朝向边缘路由器‘B’ 102B的传出(outgoing)路径。

[0047] 然而,映射服务器TLV(在OSPF扩展中定义的OSPF扩展范围)用于LDP拼接,且因此SR路径将被拼接到LDP路径。

[0048] 根据使用这里详述的扩展的一些实施例,可以由运营商根据所选择的路径来指定偏好—例如,可以说在映射服务器SID子TLV中的“偏好RSVP”(例如,并且因此,类型2路径110),且核心路由器‘3’ 104C因此可以将传入(incoming)的SR路径拼接到RSVP隧道(例如,类型2路径110)。

[0049] 当核心路由器‘4’ 104D接收到边缘路由器‘B’ 102B的绑定SID时,并且因为核心路由器‘4’ 104D处可用的仅有的MPLS路径是LDP路径(例如,类型1路径108A),SR路径将被拼接到LDP。因此,本文所述的“偏好RSVP”的偏好不会具有任何影响。类似的情况在边缘路由器‘A’ 102A处产生,因为不存在对于目的地转发等价类(FEC)边缘路由器‘B’ 102B可用的非SR LSP路径。

[0050] 接下来,描述了由边缘路由器‘B’ 102B学习来自边缘路由器‘A’ 102A的服务标签的情况。在核心路由器‘3’ 104C(其在SR网络部分120和MPLS网络部分130这二者中操作)处,将不存在来自其相邻核心路由器‘1’ 104A的可用于边缘路由器“102A”的任何LDP或RSVP绑定。然而,朝向边缘路由器‘A’ 102A的LDP和RSVP“入标签(in-label)”这二者都可以用由边缘路由器‘A’ 102A(其在核心路由器‘3’ 104C处可用)通告的SR SID/标签被拼接到朝向边缘路由器‘A’ 102A的最短路径下一跳。

[0051] 在核心路由器‘4’104D处,LDP绑定不可用于从相邻边缘路由器‘C’102C朝向边缘路由器‘A’102A,但是SR节点分段可用于边缘路由器‘A’102A并且用于将LDP拼接到SR路径。

#### [0052] SR到静态MPLS映射

[0053] 如果静态提供的MPLS路径可用(例如,在核心路由器‘4’104D处),则以上描述(关于SR到RSVP映射)同样适用。

#### [0054] SR到标签BGP隧道映射

[0055] 在无缝MPLS场景中,从SR路径到标签BGP(iBGP+标签)的拼接可以在区域边界路由器(ABR)处发生。因此,如果映射服务器106用SR标签来通告这些前缀,则可以将这些在ABR处映射到标签BGP。

#### [0056] IGP扩展

[0057] 图2示出了根据一些实施例的非特定偏好子TLV元素200、用于OSPF的特定偏好子TLV元素250以及用于IS-IS的特定偏好子TLV 275,其可以用于实现用于穿越SR和MPLS网络的业务的灵活路径拼接和选择。

[0058] 当在IGP区域中提供多个映射服务器的情况下在IGP中通告非特定偏好子TLV元素200时,所有映射服务器通告前缀的相同类型的“偏好”,其与从多个映射服务器通告时先前存在的关于标签值的约束没有区别(即,它必须相同)。非特定偏好子TLV元素200包括类型202、长度204和SID拼接偏好206。类型202可以是将非特定偏好子TLV元素200唯一地识别为非特定偏好子TLV元素的值,且因此可以由标准组织(standards body)等设置。在一些实施例中(例如,在OSPF实施例中),类型202的大小可以是2个字节,而在其它实施例中(例如,在ISIS实施例中),类型202的大小可以是1个字节。

[0059] 非特定偏好子TLV元素200还包括:长度204值,其识别TLV所承载的“值”的大小—这里是SID拼接偏好206。在一些实施例中(例如,在OSPF实施例中),长度204的大小可以是2个字节,但是在其它实施例中(例如,在IS-IS实施例中),长度204的大小可以是1个字节。

[0060] 非特定偏好子TLV元素200还包括SID拼接偏好206,在一些实施例中,SID拼接偏好206的大小为2个字节。非特定偏好子TLV元素200中的SID拼接偏好206包括可以一般地识别要使用的路径类型的值,且因此在一些实施例中不识别特定路径本身。在所描绘的实施例中,SID拼接偏好206的值208可以如下:

值	偏好
0x1	偏好PSVP
0x2	偏好静态MPLS
0x3	偏好标签BGP
0x4-0xffff	保留的

[0062] 当然,这些值和特定偏好仅是示例性的并且在一个实施例中使用,并且在不同的实施例中可以使用许多不同的值和偏好。

[0063] 图2还示出了用于OSPF的特定偏好子TLV元素250和用于IS-IS的特定偏好子TLV 275。与包括指示路径的偏好类型的值在内的非特定偏好子TLV元素200相反,这些子TLV元素250和275可以用于识别特定路径本身。

[0064] 因此,通过利用这些子TLV元素250和275,映射服务器可以指定特定的LSP名称并超控(override)以上针对特定目的地描述的通用偏好。这些中的多个可以存在于系统中,

但不能有任何重复。

[0065] 在所示实施例中,用于OSPF的特定偏好子TLV元素250包括2字节类型202、2字节长度204、2字节SID拼接偏好206、4字节目的地路由器ID 252、2字节LSP ID 254以及2字节隧道ID 256。在该实施例中,目的地路由器ID 252可以包括目的地的4个八位字节的OSPF路由器id,并且LSP ID 254可以包括在IETF请求注释(RFC) 3209(题为“RSVP-TE:用于LSP隧道的针对RSVP的扩展”)的第4.6.2.1节(“LSP\_TUNNEL\_IPv4发送方模板对象”)中定义的16比特“LSP ID”标识符。在一些实施例中,当尚未发信号通知LSP时使用值0。

[0066] 类似地,在一些实施例中,隧道ID 256可以包括:在IETF RFC 3209的第4.6.1.1节(“LSP\_TUNNEL\_IPv4会话对象”)中定义的16比特“隧道ID”值,其可以在隧道的寿命期间保持恒定。

[0067] 类似地,在利用IS-IS的网络中,可以使用用于IS-IS的特定偏好子TLV 275。用于IS-IS的特定偏好子TLV 275包括1字节类型202、1字节长度204、2字节SID拼接偏好206、6字节目的地系统ID 277、2字节LSP ID 254和2字节隧道ID 256。同样,类型202可以是表示子TLV确实是用于IS-IS的特定偏好子TLV 275的唯一标识符,并且长度204和SID拼接偏好206可以与如上所述类似。目的地系统ID 277可以包括例如6个八位字节的IS-IS系统ID,其长度为“ID长度”,如[国际ISO/IEC标准10589或“ISO 10589”]中所定义的。类似地,LSP ID 254可以包括在IETF RFC 3209的第4.6.2.1节(“LSP\_TUNNEL\_IPv4发送方模板对象”)中定义的16比特“LSP ID”标识符,并且如果LSP尚未被发信号通知则使用值0。另外,隧道ID 256可以包括:在IETF RFC 3209中的第4.6.1.1节(“LSP\_TUNNEL\_IPv4会话对象”)中定义的16比特“隧道ID”值,其中,隧道ID可以在隧道的寿命期间保持恒定。

[0068] 在一些实施例中,使用该附加特定子TLV 250/275允许运营商针对特定边界节点和所表示的特定LSP类型来指定拼接偏好。这可以对于超控通用偏好特别有用,并且可以根据在运行时期间在特定节点处出现的LSP的动态特性通过控制器来发信号通知。

[0069] 图3示出了根据一些实施例的OSPF扩展前缀范围TLV 300,其可以用于承载如图2所示的非特定偏好子TLV或用于OSPF的特定偏好子TLV。OSPF扩展前缀范围TLV 300可以是在扩展前缀LSA(链路状态通告)中的IETF草案“draft-ietf-ospf-segment-routing-extensions-05”中定义的版本,其被修改为包括如图所示的非特定偏好子TLV元素200或用于OSPF的特定偏好子TLV元素250。例如,这可以在OSPFv2环境中使用。

[0070] 尽管本文未示出,但是对于OSPFv3,这些子TLV可以由OSPFv3扩展前缀范围TLV来承载,如IETF草案“draft-ietf-ospf-ospfv3-segment-routing-extensions-04”中所定义的。对于OSPFv3,扩展前缀范围TLV是IETF互联网草案“ietf-ospf-ospfv3-lsa-extend”中定义的以下LSA的新的顶级TLV。值得注意的是:包括非特定偏好子TLV元素200或用于OSPF的特定偏好子TLV元素250可以是可选的,并且当它不存在时,相应的映射服务器前缀可以替代地用于LDP。

[0071] 图4示出了根据一些实施例的专用IS-IS“TLV 149”(SID/标签绑定TLV) 400和专用IS-IS“TLV 150”(多拓扑SID/标签绑定TLV) 450,其可以如图2所示用于承载非特定偏好子TLV 200或用于IS-IS的特定偏好子TLV 275。这些TLV 400/450可以基于IETF互联网草案“draft-ietf-isis-segment-routing-extensions-06”中定义的那些,但是被修改为包括本文描述的子TLV 200/275。在一些实施例中,包括该子TLV 200/275在IS-IS中是可选的,

以及如果不存在,则相应的映射服务器前缀可以用于LDP。

[0072] 图5A示出了根据一些实施例的用于实现用于穿越SR和MPLS网络的业务的灵活路径拼接和选择的两个操作流程500/550。将参照其它附图的示例实施例来描述流程图中的操作。然而,应当理解:流程图中的操作可以由除了参照其它附图描述的那些实施例之外的实施例来执行,且参照这些其它附图讨论的实施例可以执行与参照流程图来讨论的那些实施例不同的操作。

[0073] 例如,第一流程500包括可以由在支持SR的网络120部分和MPLS网络130部分中操作的路由器网元(例如,图1的核心路由器‘3’ 104C)执行的操作,其中,支持SR的网络120部分和MPLS网络130部分位于单一IGP区域140中。路由器网元可以由各种类型的设备实现,包括(但不限于)专用网络设备602、通用(常见的现货或“COTS”)计算(或网络)设备604、混合网络设备606等,将参考图6A及以下进一步讨论。

[0074] 流程500包括:在框505处,(例如,在与SR网络部分和MPLS网络部分耦合的路由器处)接收由映射服务器发起的IGP通告消息。IGP通告消息包括子TLV,其识别用于从SR网络部分接收的可识别业务集合的跨MPLS网络部分的路径的偏好类型。在一些实施例中,IGP通告消息包括OSPF或IS-IS消息(例如,OSPFv2或OSPFv3 LSA、诸如TLV类型149或150之类的IS-IS TLV等)。在一些实施例中,子TLV包括图2的非特定偏好子TLV元素200、用于OSPF的特定偏好子TLV元素250或用于IS-IS的特定偏好子TLV元素275。在一些实施例中,子TLV包括识别路径的偏好类型(例如RSVP隧道、静态MPLS路径,标签BGP路径等)的值(例如,SID拼接偏好206)。

[0075] 流程500还包括:在框510处,基于由子TLV识别的路径的偏好类型,识别用于可识别业务集合的跨MPLS网络部分的多个可用路径中的一个路径。在一些实施例中,框510包括:根据SID拼接偏好206来识别路径的偏好类型,以及确定是否存在通向目的地的具有相同类型的路径。在一些实施例中,因此将存在的具有该相同类型的一个路径识别为用于可识别业务集合的跨MPLS网络部分的路径。在一些实施例中,在确定了不存在具有与路径的偏好类型相同类型的这种路径时,识别另一种类型的路径。

[0076] 在框515处,流程500包括:配置路由器的转发平面以利用所识别的一个路径来转发属于可识别业务集合的从SR网络部分接收的任何业务。框515可以包括:安装或以其它方式修改例如转发平面的一个或多个线路卡中的转发/流表条目。在其它实施例中,框515可以包括本领域技术人员已知的用于实现业务转发的其它操作,包括(但不限于)修改存储器中存储的供软件路由器等使用的转发/流表条目。

[0077] 可选地,如框520的虚线边界所示,流程500还包括:从SR网络部分接收属于可识别业务集合的一个或多个业务分组。可选地,如框525的虚线边框所示,流程500包括:根据MPLS网络部分上的所识别的一个路径来发送一个或多个分组,其可以包括使用(例如,线路卡的)物理网络接口、虚拟网络接口(例如,在由同一计算设备实现的两个软件路由器网元之间)等发送一个或多个分组。

[0078] 图5A所示的第二流程550包括:可以由映射服务器106(例如,图1的映射服务器106)执行的操作,其可由诸如专用网络设备602、通用(COTS)计算(或网络)设备604、混合网络设备606等的电子设备执行,将参考图6A及以下进一步讨论。在一些实施例中,映射服务器106可以在SR网络部分120内操作。

[0079] 所示流程550在框555处包括:接收偏好数据,该偏好数据识别用于可识别业务集合的跨MPLS网络的路径的偏好类型。路径的偏好类型可以是RSVP隧道、静态MPLS路径或标签BGP路径。当然,在其它实施例中,可以使用路径的更多、更少和/或不同类型。

[0080] 所示流程550在框560处包括:发送包括子TLV在内的IGP通告消息,该子TLV识别跨MPLS网络的路径的偏好类型。在一些实施例中,IGP通告消息包括OSPF或IS-IS消息(例如,OSPFv2或OSPFv3 LSA、诸如TLV类型149或150之类的IS-IS TLV等)。在一些实施例中,子TLV包括图2的非特定偏好子TLV元素200、用于OSPF的特定偏好子TLV元素250或用于IS-IS的特定偏好子TLV元素275。

[0081] 图5B示出了根据一些实施例的核心路由器104C的高级框图568,其在一些实施例中可以执行图5A的流程500。

[0082] 本文描述的每个模块不必严格地必须实现为物理上分离的单元。一些或所有模块可以在物理单元中组合。而且,模块不需要严格地以硬件实现。可以设想:可以通过硬件和软件的组合来实现这些单元。例如,电子设备570可以包括:一个或多个中央处理单元,执行非暂时性存储介质中或在固件中存储的程序指令以执行模块的功能。

[0083] 电子设备570可以实现核心路由器104C,并且包括接收模块754、识别模块574和/或配置模块576。

[0084] 接收模块572可以适于接收由映射服务器发起的IGP通告消息(例如,IS-IS通告消息)。IGP通告消息(例如,IS-IS通告消息)包括子TLV,其中,子TLV识别用于从分段路由网络部分接收的可识别业务集合的跨MPLS网络部分的路径的偏好类型。在一些实施例中,路径的偏好类型是RSVP隧道、静态MPLS路径或标签BGP路径。在一些实施例中,接收模块572可以执行图5A的流程500的框505。

[0085] 识别模块574可以适于基于由子TLV识别的路径的偏好类型,来识别用于可识别业务集合的跨MPLS网络部分的多个可用路径中的一个路径。在一些实施例中,识别模块574可以执行图5A的流程500的框510。

[0086] 配置模块576可以适于配置核心路由器的转发平面以利用所识别的一个路径来转发属于可识别业务集合的从分段路由网络部分接收的任何业务。在一些实施例中,配置模块576可以执行图5A的流程500的框515。

[0087] 在一些实施例中,电子设备570可以包括其它未示出的模块(诸如用于执行图5A的流程500的框520和/或525的一个或多个模块),并且可以包括用于执行本文所公开的其它操作的其它未示出的模块。因此,所描绘的实施例是说明性的而非限制性的。

[0088] 图5C示出了根据一些实施例的映射服务器106的高级框图578,映射服务器106在一些实施例中可以执行图5A的流程550。

[0089] 本文描述的每个模块不必严格地实现为物理上分离的单元。一些或所有模块可以在物理单元中组合。而且,模块不需要严格地以硬件实现。可以设想:可以通过硬件和软件的组合来实现这些单元。例如,电子设备580可以包括:一个或多个中央处理单元,其执行非暂时性存储介质中或在固件中存储的程序指令以执行模块的功能。

[0090] 电子设备580可以实现映射服务器106,并且包括接收模块582和/或发送模块584。

[0091] 接收模块582可以适于接收识别用于可识别业务集合的跨MPLS网络的路径的偏好类型的偏好数据,其中,路径的偏好类型是RSVP隧道、静态MPLS路径或标签BGP路径。在一些

实施例中,接收模块582可以执行图5A的流程550的框555。

[0092] 发送模块584可以适于发送包括子TLV在内的IGP通告消息(例如,IS-IS通告消息),其中,子TLV识别跨MPLS网络的路径的偏好类型。在一些实施例中,发送模块584可以执行图5A的流程550的框560。

[0093] 在一些实施例中,电子设备580可以包括其它未示出的模块以执行本文公开的其他操作。因此,所描绘的实施例是说明性的而非限制性的。

[0094] 如本文所述,一些实施例使用电子设备。电子设备可以使用机器可读介质(也被称为计算机可读介质)来(内部和/或通过网络使用其它电子设备)存储和/或发送代码(其由软件指令构成,且有时被称为计算机程序代码或计算机程序)和/或数据,机器可读介质是例如机器可读存储介质(例如,磁盘、光盘、只读存储器(ROM)、闪存设备、相变存储器)和机器可读传输介质(也被称为载体)(例如,电、光、无线电、声或其它形式的传播信号—例如载波、红外信号)。从而,许多电子设备(例如,计算机或“计算设备”)可以包括硬件和软件,例如一个或多个处理器的集合,其耦合到用于存储在该处理器集合上执行的代码和/或用于存储数据的一个或多个机器可读存储介质。

[0095] 例如,电子设备可以包括包含代码在内的非易失性存储器,因为即使当电子设备关闭时(当电源被移除时),非易失性存储器也能够保持代码/数据。当电子设备被接通时,由该电子设备的处理器执行的代码的该部分通常从该电子设备的较慢的非易失性存储器复制到易失性存储器(例如动态随机存取存储器(DRAM)、静态随机存取存储器(SRAM))中。许多电子设备还包括用于与其它电子设备建立网络连接(以使用传播信号来发送和/或接收代码和/或数据)或以其它方式与其它电子设备通信的一个或多个物理网络接口的集合。可以使用软件、固件和/或硬件的不同组合来实现本文描述的各种实施例的一个或多个部分。

[0096] 网络设备(ND)是在一个或多个网络中通信互连其它电子设备(例如,其它网络设备、终端用户设备)的电子设备。一些网络设备是为多个联网功能(例如,路由、桥接、交换、层2聚合、会话边界控制、服务质量、和/或订户管理)提供支持、和/或为多应用服务(例如,数据、语音、和视频)提供支持的“多服务网络设备”。

[0097] 图6A示出了根据一些实施例的示例网络内的网络设备(ND)之间的连接性以及ND的三种示例实现。图6A示出了ND 600A~600H以及借助于600A~600B、600B~600C、600C~600D、600D~600E、600E~600F、600F~600G和600A~600G之间以及600H与600A、600C、600D和600G中的每一个之间的线路示出了它们的连接性。这些ND是物理设备,且这些ND之间的连接性可以是无线的或有线的(经常被称为链路)。从ND 600A、600E和600F延伸(且并未示出为在终点处结束)的附加线示出了:这些ND担当网络的入口点和出口点,且因此这些ND有时称为“边缘”ND,而其它ND可以被称为“核心”ND。

[0098] 图6A中的两个示例ND实现是:1)专用网络设备602,其可以包括定制的专用集成电路(ASIC)和专用操作系统(OS);以及2)通用网络设备604,其可以使用常见的现货供应(COTS)组件(例如,处理器)和可能的“标准”OS。

[0099] 专用网络设备602包括联网硬件610,联网硬件610包括计算资源612(通常包括一个或多个处理器的集合)、转发资源614(其通常包括一个或多个ASIC和/或网络处理器)、以及物理网络接口(NI)616(有时被称为物理端口)、以及其中存储有联网软件620的非瞬时机

器可读存储介质618。

[0100] 在一些实施例中, 联网软件620可以包括: IGP模块代码690A, 用于实现可以执行本文公开的操作的IGP模块633A, 该操作包括图5A的流程500/550的操作以及在整个说明书中讨论的操作。

[0101] 物理NI是ND中的硬件, 通过该硬件进行网络连接(例如, 通过无线网络接口控制器(WNIC)以无线方式或者通过将电线插入连接到网络接口控制器(NIC)的物理端口), 例如由ND 600A~600H之间的连接所示出的那些网络连接。

[0102] 在操作期间, 联网软件620可以由联网硬件610来执行, 以实例化一个或多个联网软件实例622的集合。每个联网软件实例622以及联网硬件610的执行该网络软件实例的部分(如果其是专用于该联网软件实例的硬件和/或由该联网软件实例与其它联网软件实例622按时间共享的硬件的时间片)可以形成(单独的)虚拟网元630A~630R。每个虚拟网元(VNE) 630A~630R包括控制通信和配置模块632A~632R(有时被称为本地控制模块或控制通信模块)和转发表634A~634R, 使得给定的虚拟网元(例如, 630A)包括控制通信和配置模块(例如, 632A)、一个或多个转发表的集合(例如, 634A)、以及联网硬件610的执行虚拟网元(例如, 630A)的部分。

[0103] 在一些实施例中, 虚拟网元(VNE) 630A~630R的控制通信和配置模块(例如, 632A)中的一个或多个可以包括通过IGP模块代码690A的执行而实例化或实现的IGP模块633A。

[0104] 专用网络设备602经常在物理上和/或在逻辑上被视为包括: 1) ND控制平面624(有时被称为控制平面), 包括执行控制通信和配置模块632A~632R的计算资源612; 以及2) ND转发平面626(有时被称为转发平面、数据平面、或媒体平面), 包括利用转发表634A~634R的转发资源614以及物理NI 616。作为ND是路由器(或实现路由功能)的示例, ND控制平面624(执行控制通信和配置模块632A~632R的计算资源612)通常负责参与控制如何路由(例如, 数据的下一跳和该数据的输出物理NI)数据(例如, 分组)并负责在转发表634A~634R中存储该路由信息, 以及ND转发平面626负责在物理NI 616上接收该数据并基于转发表634A~634R将该数据转发出物理NI 616中的恰当物理NI。

[0105] 图6B示出了根据一些实施例的用于实现专用网络设备602的示例方式。图6B示出了包括卡638(通常是可热插拔的)在内的专用网络设备。尽管在一些实施例中, 卡638具有两种类型(作为ND转发平面626来工作的一个或多个卡(有时被称为线路卡)、以及进行操作以实现ND控制平面624的一个或多个卡(有时被称为控制卡)), 备选实施例可以将功能结合到单一卡上和/或包括附加卡类型(例如, 一种附加类型的卡被称为服务卡、资源卡、或多应用卡)。服务卡可以提供特殊处理(例如, 层4到层7服务(例如, 防火墙、网际协议安全(IPsec)、安全套接字层(SSL)/传输层安全(TLS)、入侵检测系统(IDS)、对等(P2P)、基于IP的语音(VoIP)会话边界控制器、移动无线网关(网关通用分组无线电服务(GPRS)支持节点(GGSN)、演进分组核心(EPC)网关))。作为示例, 服务卡可以用于端接IPsec隧道, 并执行伴随的认证和加密算法。这些卡通过被示出为背板636的一个或多个互联机制耦合在一起(例如, 耦合线路卡的第一全网状网络以及耦合所有卡的第二全网状网络)耦合。

[0106] 返回图6A, 通用网络设备604包括硬件640, 硬件640包括一个或多个处理器642(其经常是COTS处理器)的集合和网络接口控制器644(NIC, 也被称为网络接口卡)(其包括物理NI 646)、以及其中存储有软件650的非瞬时机器可读存储介质648。

[0107] 软件650可以包括：IGP模块代码690B，其可以由处理器642执行以实现IGP模块（未示出），其可以是应用664A~664R的一部分，以便执行本文所描述的操作。

[0108] 在操作期间，处理器642执行软件650以实例化一个或多个应用664A~664R的一个或多个集合。虽然一个实施例不实现虚拟化，但是备选实施例可以使用不同形式的虚拟化。例如，在一个这样的备选实施例中，虚拟化层654表示操作系统的内核（或在基本操作系统上执行的垫片（shim）），其允许创建被称为软件容器的多个实例662A~662R，每个实例可以用于执行一个（或多个）应用664A~664R集合；其中多个软件容器（也称为虚拟化引擎、虚拟专用服务器或监牢（jail））是用户空间（通常是虚拟存储器空间），用户空间彼此分离并与运行操作系统的内核空间分离；以及除非明确允许，否则在给定用户空间中运行的应用集合不能访问其它进程的存储器。在另一个这样的备选实施例中，虚拟化层654表示管理程序（hypervisor，有时被称为虚拟机监视器（VMM））或在主机操作系统之上执行的管理程序，并且每个应用664A~664R集合在实例662A~662R内的访客操作系统之上运行，实例662A~662R被称为在管理程序之上运行的虚拟机（在某些情况下其可以被认为是一种紧密隔离形式的软件容器）—访客操作系统和应用可能没有意识到它们正在虚拟机上运行而不是在“裸机（bare metal）”主机电子设备上运行，或者出于优化目的，通过超虚拟化（para-virtualization），操作系统和/或应用可能意识到虚拟化的存在。在其它备选实施例中，应用中的一个、一些或全部被实现为单核，其可以通过直接用应用仅编译（例如，来自包括OS收信方（sendee）的驱动程序/库在内的库操作系统（LibOS）的）有限的库集合来生成，该有限的库集合提供应用所需的特定OS服务。由于可以实现单内核以直接在硬件640上运行，直接在管理程序上运行（在这种情况下，单内核有时被描述为在LibOS虚拟机内运行），或者在软件容器中运行，所以实施例可以完全以以下方式实现：直接在由虚拟化层654表示的管理程序上运行的单内核、在由实例662A~662R表示的软件容器内运行的单内核、或者作为单内核和上述技术的组合（例如，单内核和虚拟机都直接在管理程序上运行、单内核和应用的集合在不同的软件容器中运行）。

[0109] 一个或多个应用664A~664R的一个或多个集合的实例化以及虚拟化（如果实现的话）统称为软件实例652。每个应用664A~664R集合、相应的虚拟化构造（例如，实例662A~662R）（如果实现的话）以及硬件640的执行它们的那部分（如果专用于该执行的硬件和/或时间上共享的硬件的时间片）形成单独的虚拟网元660A~660R。

[0110] 虚拟网元660A~660R执行与虚拟网元630A~630R类似的功能—例如，类似于控制通信和配置模块632A以及转发表634A（硬件640的该虚拟化有时被称为网络功能虚拟化（NFV））。从而，NFV可以用于将很多网络设备类型统一到工业标准大容量服务器硬件、物理交换机、和物理存储器，它们可以位于数据中心、ND、和客户住宅设备（CPE）中。尽管将本发明的实施例示出为每个实例662A~662R对应于一个VNE 660A~660R，但是备选实施例可以在更精细级别粒度上实现该对应关系（例如，线路卡虚拟机虚拟化线路卡，控制卡虚拟机虚拟化控制卡，等等）；应当理解：本文参考实例662A~662R与VNE的对应关系来描述的技术同样适用于使用这种更精细级别粒度和/或单内核的实施例。

[0111] 在某些实施例中，虚拟化层654包括提供与物理以太网交换机类似的转发服务的虚拟交换机。具体地，该虚拟交换机在实例662A~662R和NIC 644之间转发业务，以及可选地在实例662A~662R之间转发业务；此外，该虚拟交换机可以在根据策略不被允许彼此通

信的VNE 660A~660R之间强制执行网络隔离(例如,通过执行虚拟局域网(VLAN))。

[0112] 图6A中的第三示例ND实现是混合网络设备606,其在单一ND或ND内的单一卡中包括定制ASIC/专用的OS和COTS处理器/标准OS。在这种混合网络设备的某些实施例中,平台VM(即,实现专用网络设备602的功能的VM)可以向混合网络设备606中存在的联网硬件提供超虚拟化。

[0113] 不管ND的上述示例实现如何,当考虑由ND实现的多个VNE中的单一一个VNE时,或者在NE当前仅实现单一VNE的情况下,缩写的术语网元(NE)有时被用于指代该VNE。同样在全部上述示例实现中,每个VNE(例如,VNE 630A~630R、VNE 660A~660R、以及混合网络设备606中的那些)在物理NI(例如,616、646)上接收数据并将该数据转发出物理NI(例如,616、646)中的恰当物理NI。例如,实现IP路由器功能的VNE基于IP分组中的一些IP报头信息来转发IP分组;其中IP报头信息包括源IP地址、目的地IP地址、源端口、目的地端口(其中“源端口”和“目的地端口”在本文中指代协议端口,而不是ND的物理端口)、传输协议(例如,用户数据报协议(UDP)、传输控制协议(TCP)和差异化服务(DSCP)值)。

[0114] 图6C示出了根据本发明一些实施例的可以耦合VNE的各种示例方式。图6C示出了在ND 600A中实现的VNE 670A.1~670A.P(以及可选的VNE 670A.Q~670A.R)以及ND 600H中的VNE 670H.1。在图6C中,VNE 670A.1~P在以下意义上是彼此分离的:它们可以从ND 600A外部接收分组并向ND 600A外部转发分组;VNE 670A.1与VNE 670H.1耦合,且从而它们在其各自ND之间传输分组;VNE 670A.2~670A.3可以可选地在它们之间转发分组,而不向ND 600A外部转发分组;以及VNE 670A.P可以可选地是VNE链中的第一个,该VNE链包括VNE 670A.Q,之后包括VNE 670A.R(有时这被称为动态服务链接,其中,一系列VNE中的每个VNE提供不同服务—例如,一个或多个层4~7网络服务)。尽管图6C示出了VNE之间的各种示例关系,备选实施例可以支持其它关系(例如,更多/更少的VNE、更多/更少的动态服务链、具有一些公共VNE和一些不同VNE的多个不同动态服务链)。

[0115] 图6A的ND例如可以形成互联网或专用网络的一部分;以及其它电子设备(未示出;例如终端用户设备,包括工作站、膝上型计算机、上网本、平板电脑、掌上型计算机、移动电话、智能电话、平板手机、多媒体电话、基于网际协议的语音(VoIP)电话、终端、便携式媒体播放器、全球定位系统(GPS)单元、可穿戴设备、游戏系统、机顶盒、支持互联网的家用电器)可以耦合到网络(直接或通过诸如接入网之类的其它网络),以通过网络(例如,互联网或覆盖(例如,隧道传输)在互联网上的虚拟专用网络(VPN))彼此通信(直接或通过服务器)和/或访问内容和/或服务。这种内容和/或服务通常由属于服务/内容提供商的一个或多个服务器(未示出)或参与对等(P2P)服务的一个或多个终端用户设备(未示出)来提供,且可以包括例如公开网页(例如,自由内容、商店前端、搜索服务)、私有网页(例如,提供电子邮件服务的用户名/密码访问网页)、和/或基于VPN的公司网络。例如,终端用户设备可以耦合(例如,通过(无线或有线)耦合到接入网的客户住宅设备)到边缘ND,边缘ND耦合(例如,通过一个或多个核心ND)到其它边缘ND,该其它边缘ND耦合到担当服务器的电子设备。然而,通过计算和存储虚拟化,作为图6A中ND来操作的一个或多个电子设备还可以托管一个或多个这种服务器(例如,在通用网络设备604的情况下,软件实例662A~662R中的一个或多个软件实例可以作为服务器操作;这对于混合网络设备606来说也将是正确的;在专用网络设备602的情况下,一个或多个这种服务器还可以运行在由计算资源612执行的虚拟化层上);

在该情况下,服务器被称为是与该ND的VNE同处一地。

[0116] 虚拟网络是提供网络服务(例如,L2和/或L3服务)的物理网络(例如,图6A中的那个物理网络)的逻辑抽象。虚拟网络可以实现为在底层网络(例如,L3网络,如使用隧道(例如,通用路由封装(GRE)、层2隧道传输协议(L2TP)、IPSec)来创建覆盖网络的网际协议(IP)网络上提供网络服务(例如,层2(L2,数据链路层)和/或层3(L3,网络层)服务)的覆盖网络(有时被称为网络虚拟化覆盖)。

[0117] 网络虚拟化边缘(NVE)位于底层网络的边缘处,并参与实现网络虚拟化;NVE的面向网络侧使用底层网络向其它NVE隧道传输帧以及隧道传输来自其它NVE的帧;NVE的面向外部侧向网络外部的系统发送数据和从网络外部的系统接收数据。虚拟网络实例(VNI)是NVE上虚拟网络的特定实例(例如,ND上的NE/VNE、ND上NE/VNE的一部分,其中,通过仿真将该NE/VNE分为多个VNE);可以在NVE上实例化一个或多个VNI(例如,作为ND上的不同VNE)。虚拟接入点(VAP)是NVE上用于将外部系统连接到虚拟网络的逻辑连接点;VAP可以通过逻辑接口标识符(例如,VLAN ID)来识别的物理或虚拟端口。

[0118] 网络服务的示例包括:1)以太网LAN仿真服务(类似于IETF MPLS或以太网VPN(EVPN)服务的基于以太网的多点服务),其中,外部系统通过基于底层网络的LAN环境跨网络互连(例如,NVE针对不同的这种虚拟网络提供分离的L2 VNI(虚拟交换实例),以及提供跨底层网络的L3(例如,IP/MPLS)隧道传输封装);以及2)虚拟化IP转发服务(从服务定义角度来说,类似于IETF IP VPN(例如,BGP/MPLS IPVPN),其中,外部系统通过基于底层网络的L3环境跨网络互连(例如,NVE针对不同的这种虚拟网络提供分离的L3 VNI(转发和路由实例),以及提供跨底层网络的L3(例如,IP/MPLS)隧道传输封装)。网络服务还可以包括服务质量能力(例如,业务分类标记、业务调节(conditioning)和调度)、安全能力(例如,用于保护客户住宅免受源自网络的攻击的过滤器,以避免有缺陷的路由公告)、以及管理能力(例如,完全检测和处理)。

[0119] 图6D示出了根据一些实施例的在图6A的每个ND上具有单一网元的网络,且在该直观方案中,对比了(传统路由器常用的)传统分布式方案与用于维护可达性和转发信息(也被称为网络控制)的集中式方案。具体地,图6D示出了具有与图6A的ND 600A~600H相同连接性的网元(NE) 670A~670H。

[0120] 图6D示出了分布式方案672将用于生成可达性和转发信息(的责任分布在NE 670A~670H上;换言之,邻居发现和拓扑发现的过程是分布式的。

[0121] 例如,在使用专用网络设备602的情况下,ND控制平面624的控制通信和配置模块632A~632R通常包括可达性和转发信息模块以实现一个或多个路由协议(例如,外部网关协议,如BGP、IGP(例如,OSPF、IS-IS、路由信息协议(RIP)、LDP、RSVP(包括RSVP-业务工程(TE):用于LSP隧道的RSVP扩展和广义多协议标签交换(GMPLS)信令RSVP-TE)),其与其它NE进行通信以交换路由,然后基于一个或多个路由度量选择那些路由。从而,NE 670A~670H(例如,执行控制通信和配置模块632A~632R的计算资源612)通过分布式地确定网络内的可达性并计算其各自的转发信息来执行其参与控制如何路由(例如,数据的下一跳以及该数据的输出物理NI)数据(例如,分组)的责任。路由和邻接关系被存储在ND控制平面624上的一个或多个路由结构中(例如,路由信息库(RIB)、标签信息库(LIB)、一个或多个邻接关系结构)。ND控制平面624使用基于路由结构的信息(例如,邻接关系和路由信息)对ND转发

平面626编程。例如,ND控制平面624将邻接关系和路由信息编程到ND转发平面626上的一个或多个转发表634A~634R(例如,转发信息库(FIB)、标签转发信息库(LFIB)、以及一个或多个邻接关系结构)中。对于层2转发,ND可以存储一个或多个桥接表,该一个或多个桥接表用于基于该数据中的层2信息来转发数据。尽管上面的示例使用了专用网络设备602,相同的分布式方案672可以实现在通用网络设备604和混合网络设备606上。

[0122] 图6D示出了集中式方案674(也被称为软件定义的联网(SDN)),其将做出与从底层系统(该底层系统向所选目的地转发业务)发送业务的位置有关的决定的系统加以去耦合。所示出的集中式方案674具有用于在集中式控制平面676(有时被称为SDN控制模块、控制器、网络控制器、开放流控制器、SDN控制器、控制平面节点、网络虚拟化机构、或管理控制实体)中生成可达性和转发信息,并从而使得邻居发现和拓扑发现的过程集中化。集中式控制平面676具有针对数据平面680(有时被称为基础架构层、网络转发平面、或转发平面(其不应当与ND转发平面相混淆))的南向接口(south bound interface)682,该数据平面680包括NE 670A~670H(有时被称为交换机、转发元件、数据平面元件、或节点)。集中式控制平面676包括网络控制器678,网络控制器678包括确定网络内的可达性并通过南向接口682向数据平面680的NE 670A~670H分发转发信息的集中式可达性和转发信息模块679(其可以使用开放流协议)。从而,在通常与ND分离的电子设备上执行的集中式控制平面676中将网络智能集中化。

[0123] 例如,在数据平面680中使用专用网络设备602的情况下,ND控制平面624的每个控制通信和配置模块632A~632R通常包括提供南向接口682的VNE侧的控制代理。在该情况下,ND控制平面624(执行控制通信和配置模块632A~632R的计算资源612)通过与集中式控制平面676通信以从集中式可达性和转发信息模块679接收转发信息(以及在一些情况下的可达性信息)的控制代理来执行其参与控制如何路由(例如,数据的下一跳和该数据的输出物理NI)数据(例如,分组)的责任(应当理解:在本发明的一些实施例中,除了与集中式控制平面676通信之外,控制通信和配置模块632A~632R还可以在确定可达性和/或计算转发信息时扮演某个角色—尽管不像分布式方案情况下那么多;这种实施例一般被视为落入集中式方案674中,但也可以被视为是混合方案)。

[0124] 尽管上述示例使用了专用网络设备602,可以使用通用网络设备604和混合网络设备606来实现同样的集中式方案674(例如,每个VNE 660A~660R通过与集中式控制平面676通信以从集中式可达性和转发信息模块679接收转发信息(以及在一些情况下,可达性信息)来执行其控制如何路由(例如,该数据的下一跳和该数据的输出物理NI)数据(例如,分组)的责任;应当理解:在本发明的一些实施例中,除了与集中式控制平面676通信之外,VNE 660A~660R还可以在确定可达性和/或计算转发信息时扮演某个角色—尽管不像分布式方案情况下那么多)。事实上,对SDN技术的使用可以增强通常在通用网络设备604或混合网络设备606实现中使用的NFV技术,因为NFV能够通过提供其上可以运行SDN软件的基础架构来支持SDN,且因为NFV和SDN都打算利用商品服务器硬件和物理交换机。

[0125] 图6D还示出了集中式控制平面676具有针对应用层686的北向接口(north bound interface)684,在应用层686中驻留有应用688。集中式控制平面676具有针对应用688来形成虚拟网络692(有时被称为逻辑转发平面、网络服务、或覆盖网络(且数据平面680的NE670A~670H是底层网络))的能力。从而,集中式控制平面676维护所有ND和所配置的NE/

VNE的全局视图,且其将虚拟网络高效地映射到底层ND(包括当物理网络通过硬件(ND、链路、或ND组件)故障、添加、或移除来改变时维护这些映射)。

[0126] 尽管图6D示出了与集中式方案674不同的分布式方案672,在某些实施例中,网络控制的工作可以用不同方式分布或者这二者可以结合。例如:1) 实施例一般可以使用集中式方案(SDN) 674,但是具有委托给NE的某些功能(例如,分布式方案可以用于实现故障监视、性能监视、保护交换、以及用于邻居和/或拓扑发现的基元(primitive)中的一项或多项);或者2) 本发明的实施例可以经由集中式控制平面和分布式协议这二者来执行邻居发现和拓扑发现,且比较结果以在它们不一致的地方提出异常。这种实施例一般被视为落入集中式方案674中,但也可以被视为是混合方案。

[0127] 尽管图6D示出了每个ND 600A~600H实现单一NE 670A~670H的简单情况,但应当理解:参照图6D描述的网络控制方案同样适用于一个或多个ND 600A~600H实现多个VNE(例如,VNE 630A~630R、VNE 660A~660R、在混合网络设备606中的那些)的网络。备选地或附加地,网络控制器678还可以对在单一ND中实现多个VNE加以仿真。具体地,取代(或除了)在单一ND中实现多个VNE(之外),网络控制器678(还)可以将单一ND中的VNE/NE的实现呈现为虚拟网络692中的多个VNE(在同一个虚拟网络692中的全部、不同虚拟网络692中的每一个、或某种组合)。例如,网络控制器678可以使得ND在底层网络中实现单一VNE(NE),然后对集中式控制平面676内的该NE的资源进行逻辑分割,以在虚拟网络692中呈现不同的VNE(其中,覆盖网络中的这些不同VNE共享底层网络中ND上的单一VNE/NE实现的资源)。

[0128] 另一方面,图6E和6F分别示出了NE和VNE的示例抽象,网络控制器678可以将其作为不同的虚拟网络692的一部分来呈现。图6E示出了根据本发明一些实施例的以下简单情况:ND 600A~H中的每一个实现单一NE 670A~670H(参见图6D),而集中式控制平面676将不同ND中的多个NE(NE 670A~C和G~H)抽象为(以表示)图6D的虚拟网络692之一中的单一NE 670I。图6E示出了:在该虚拟网络中,NE 670I耦合到NE 670D和670F,它们都依然耦合到NE 670E。

[0129] 图6F示出了根据本发明一些实施例的以下情况:在不同ND(ND600A和ND 600H)上实现多个VNE(VNE 670A.1和VNE 670H.1)并且该多个VNE彼此耦合,以及集中式控制平面676将这些多个VNE加以抽象,使得它们表现为图6D的虚拟网络692之一中的单一VNE 670T。从而,NE或VNE的抽象可以跨多个ND。

[0130] 网络接口(NI)可以是物理或虚拟的;以及在IP的上下文中,接口地址是向NI指派的IP地址,不论是物理NI还是虚拟NI。虚拟NI可以与物理NI相关联、与另一虚拟接口相关联、或者是独立的(例如,环回接口、点对点协议接口)。NI(物理或虚拟)可以是有编号的(具有IP地址的NI)或无编号的(不具有IP地址的NI)。环回接口(及其环回地址)是经常用于管理用途的NE/VNE(物理或虚拟)的特定类型的虚拟NI(和IP地址);其中,这种IP地址被称为节点环回地址。向ND的NI指派的IP地址被称为该ND的IP地址;在更多粒度级别上,向ND上实现的NE/VNE指派的NI指派的IP地址可被称为该NE/VNE的IP地址。

[0131] 路由系统针对给定目的地的下一跳选择可以解析为一个路径(即,路由协议可以生成最短路径上的一个下一跳);但是如果路由系统确定存在多个可行的下一跳(即,路由协议生成的转发方案提供了与最短路径有关的多于一个下一跳—多个相等成本下一跳),使用某些附加标准—例如在无连接网络中,可以使用相等成本多路径(ECMP)(也被称为相

等成本多寻路、多路径转发和IP多路径) (例如,通常的实现使用特定报头字段作为标准,以确保特定分组流的分组始终在相同的下一跳上被转发,以保留分组流排序)。为了多路径转发的目的,将分组流定义为共享排序约束的分组集合。作为示例,特定TCP传输序列中的分组集合需要按顺序到达,否则TCP逻辑将把乱序传输解释为拥塞并将TCP传输速率变慢。

[0132] 与虚拟连接和虚拟信道同义的虚拟电路 (VC) 是通过分组模式通信传送的面向连接的通信服务。虚拟电路通信类似于电路交换,因为两者都是面向连接的,这意味着在两种情况下都以正确的顺序传送数据,并且在连接建立阶段期间需要信令开销。虚拟电路可以存在于不同层。例如,在层4,面向连接的传输层数据链路协议 (例如传输控制协议 (TCP)) 可以依赖于无连接分组交换网络层协议 (例如IP),其中不同的分组可以在不同的路径上被路由,因此可以不按顺序被传送。在底层的不可靠和无连接的IP协议之上用TCP建立可靠的虚拟电路的情况下,虚拟电路由源和目的地网络套接字地址对来识别,即发送方和接收方IP地址和端口号。然而,虚拟电路是可能的,因为TCP包括在接收方侧的分段编号和重新排序以防止无序传送。虚拟电路也可以在层3 (网络层) 和层2 (数据链路层);这种虚拟电路协议基于面向连接的分组交换,意味着数据总是沿着相同的网络路径传送,即通过相同的NE/VNE传送。在这样的协议中,分组不被单独路由,并且在每个数据分组的报头中不提供完整的寻址信息;每个分组中仅需要小的虚拟信道标识符 (VCI);并且在连接建立阶段期间,路由信息被传输到NE/VNE;交换仅涉及在表中查找虚拟信道标识符而不是分析完整地址。网络层和数据链路层虚拟电路协议的示例,其中数据始终在相同的路径上传递:X.25,其中VC由虚拟信道标识符 (VCI) 识别;帧中继,其中VC由VCI识别;异步传输模式 (ATM),其中电路由虚拟路径标识符 (VPI) 和虚拟信道标识符 (VCI) 对识别;通用分组无线业务 (GPRS);以及多协议标签交换 (MPLS),其可以用于虚拟电路上的IP (每个电路由标签识别)。

[0133] 某些ND (例如,某些边缘ND) 使用电路的层级结构。电路层级结构的叶节点是订户电路。订户电路具有层级中的父电路,该父电路通常表示多个订户电路的聚合,并且因此用于提供这些终端用户设备到ND的接入网络连接的网络分段和网元。这些父电路可以表示订户电路的物理或逻辑聚合 (例如,虚拟局域网 (VLAN)、永久虚拟电路 (PVC) (例如,用于异步传输模式 (ATM))、电路组、信道、伪线、ND的物理NI和链路聚合组。电路组是虚拟构造,其允许将各个电路集合组合在一起以用于配置目的,例如聚合速率控制。伪线是层2点对点面向连接的服务的仿真。链路聚合组是为了带宽聚合和冗余的目的而合并多个物理NI的虚拟构造。因此,父电路在物理上或逻辑上封装订户电路。

[0134] 每个VNE (例如,虚拟路由器、虚拟桥接器 (其可以担当虚拟专用LAN服务 (VPLS) 中的虚拟交换机实例) 通常是可独立管理的。例如,在多个虚拟路由器的情况下,每个虚拟路由器可以共享系统资源,但是与其管理域、AAA (认证、授权和计费) 名称空间、IP地址和路由数据库有关的其它虚拟路由器是分开的。可以在边缘ND中采用多个VNE,以便为服务和/或内容提供商的订户提供直接网络接入和/或不同类别的服务。

[0135] 在某些ND内,独立于物理NI的“接口”可以被配置为VNE的一部分,以提供更高层协议和服务信息 (例如,层3寻址)。除了其它订户配置要求之外,AAA服务器中的订户记录还识别相应订户应该在ND内绑定到哪个上下文 (例如,哪个VNE/NE)。如本文所使用的,绑定形成物理实体 (例如,物理NI、信道) 或逻辑实体 (例如,诸如订户电路或逻辑电路 (具有一个或多个订户电路的集合) 之类的电路) 和上下文的接口之间的关联,在该接口上,为该上下文配

置网络协议(例如,路由协议、桥接协议)。当配置某个较高层协议接口并与该物理实体相关联时,订户数据在物理实体上流动。

[0136] 一些ND提供对实现VPN(虚拟专用网络)(例如,层2VPN和/或层3VPN)的支持。例如,在提供商的网络和客户的网络相耦合处的ND分别称为PE(提供商边缘)和CE(客户边缘)。在层2VPN中,转发通常在VPN的任一端上的CE上执行,并且跨网络(例如,通过由其它ND耦合的一个或多个PE)发送业务。在CE和PE之间配置层2电路(例如,以太网端口、ATM永久虚拟电路(PVC)、帧中继PVC)。在层3VPN中,路由通常由PE执行。举例来说,支持多个VNE的边缘ND可以被部署为PE;并且VNE可以配置有VPN协议,且因此该VNE被称为VPN VNE。

[0137] 一些ND提供对VPLS(虚拟专用LAN服务)的支持。例如,在VPLS网络中,终端用户设备通过耦合到CE来访问通过VPLS网络提供的内容/服务,CE通过由其它ND耦合的PE来耦合。VPLS网络可以用于实现三重播放网络应用(例如,数据应用(例如,高速互联网接入)、视频应用(例如,诸如IPTV(网际协议电视)之类的电视服务、VoD(视频点播)服务)和语音应用(例如、VoIP(基于网际协议的语音)服务))、VPN服务等。VPLS是可用于多点连接的一类层2VPN。VPLS网络还允许在单独的地理位置与CE耦合的终端用户设备跨广域网(WAN)相互通信,就像它们在局域网(LAN)中彼此直接附接一样(称为模拟局域网)。

[0138] 在VPLS网络中,每个CE通常可能通过接入网络(有线和/或无线)经由附接电路(例如,CE和PE之间的虚拟链路或连接)附接到PE的桥接模块。PE的桥接模块通过仿真LAN接口附接到仿真LAN。每个桥接模块通过维护将MAC地址映射到伪线和附接电路的转发表来担当“虚拟交换机实例”(VSI)。PE基于那些帧中包括的MAC目的地地址字段将(从CE接收的)帧向目的地(例如,其它CE、其它PE)转发。

[0139] 尽管附图中的流程图示出了本发明的某些实施例执行的特定操作顺序,但是应当理解:这种顺序是示例性的(例如,备选实施例可以以不同的顺序执行操作、组合某些操作、某些操作相同等)。

[0140] 虽然在若干实施例的意义上描述了本发明,本领域技术人员将会认识到:本发明不限于所描述的实施例,而是可利用在所附权利要求的精神和范围内的修改和改变来实现。本描述因此被视为是说明性的,而非限制性的。

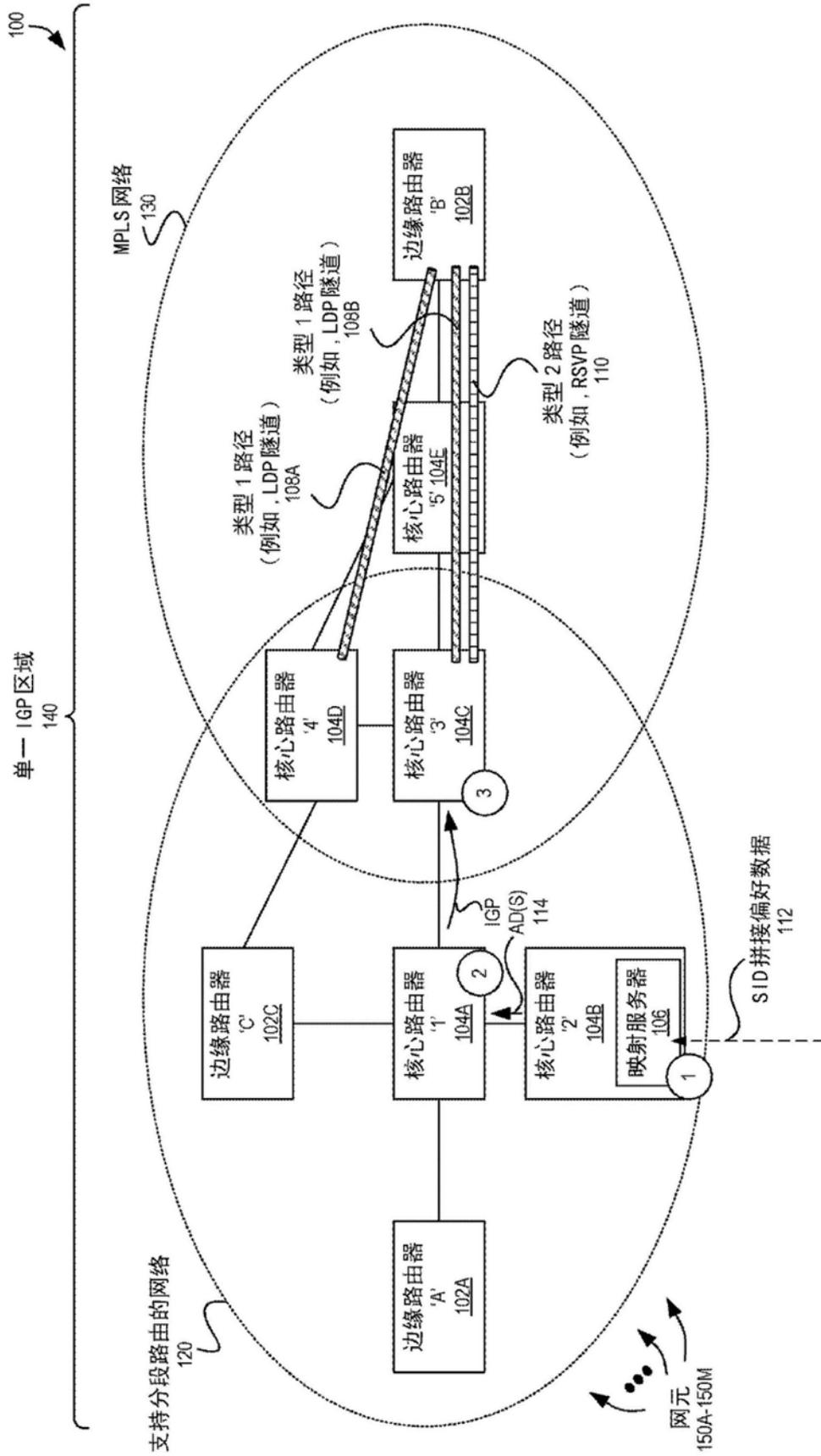


图1

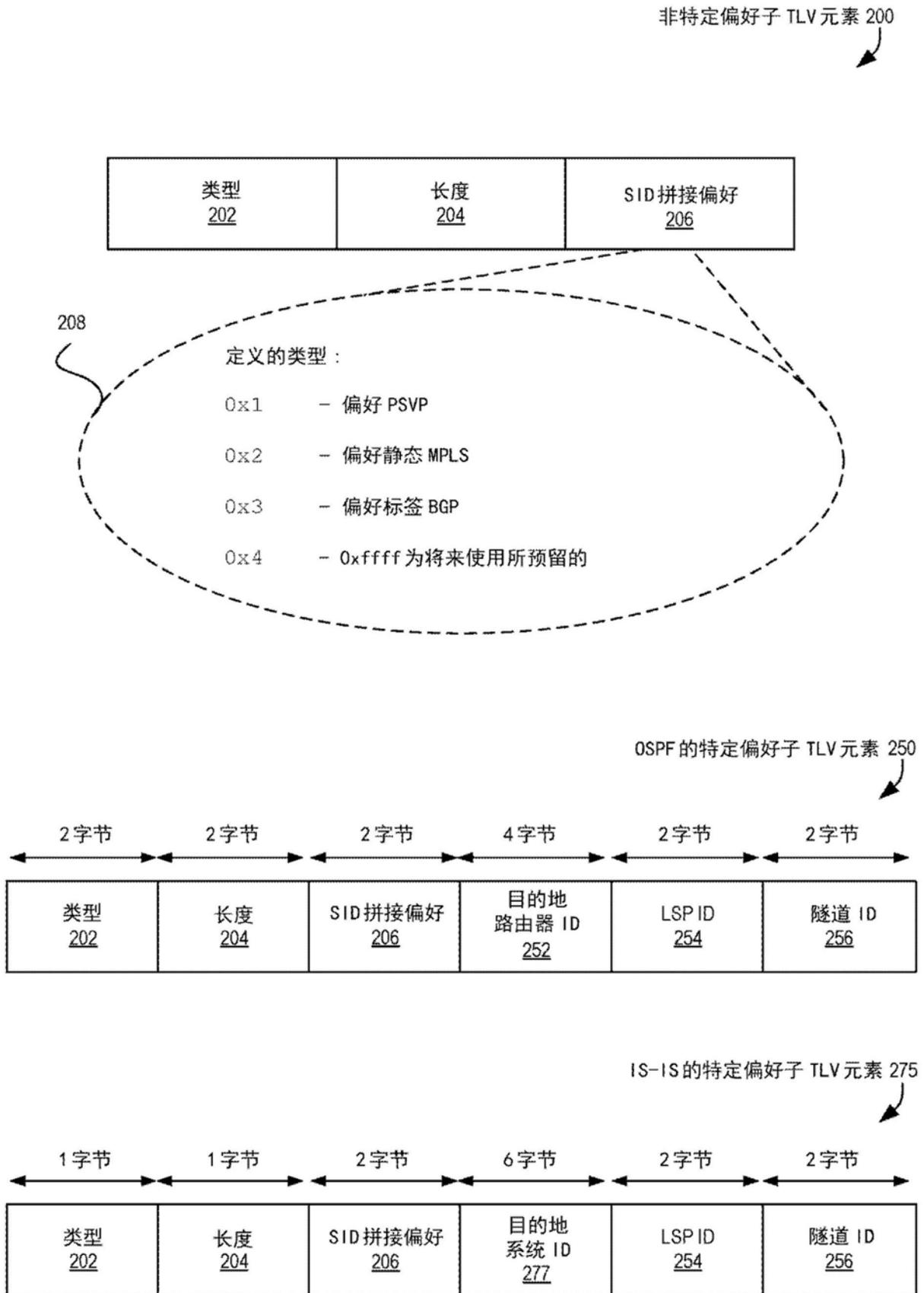


图2

OSPF 扩展前缀范围 TLV 300

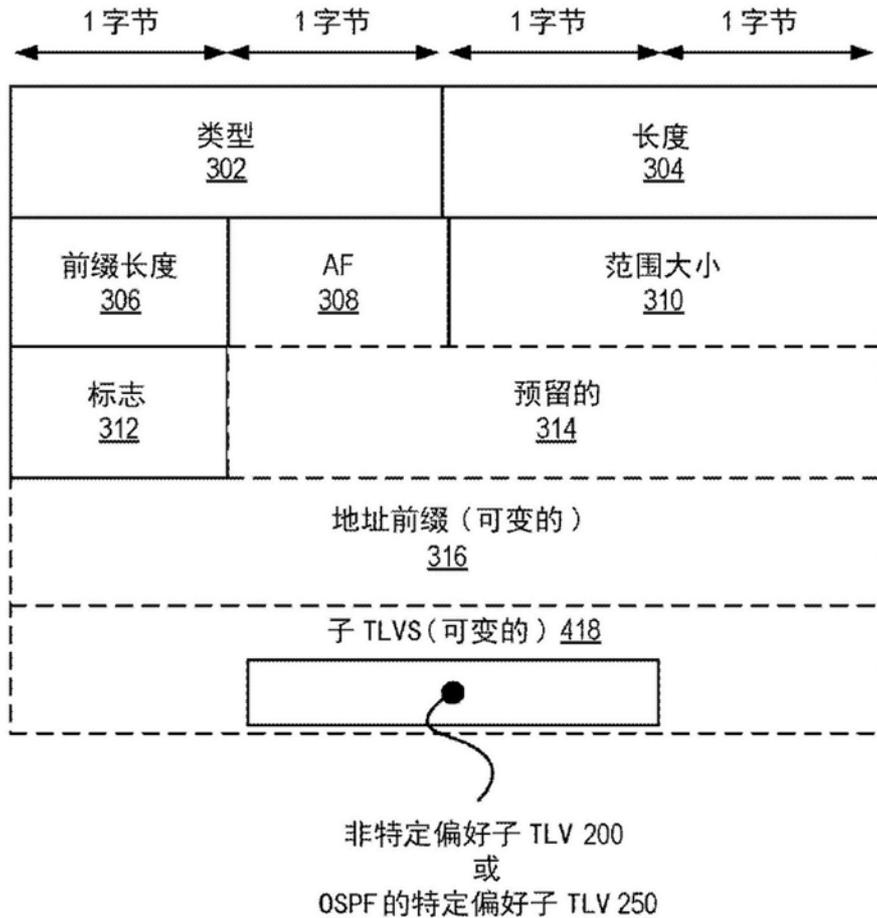
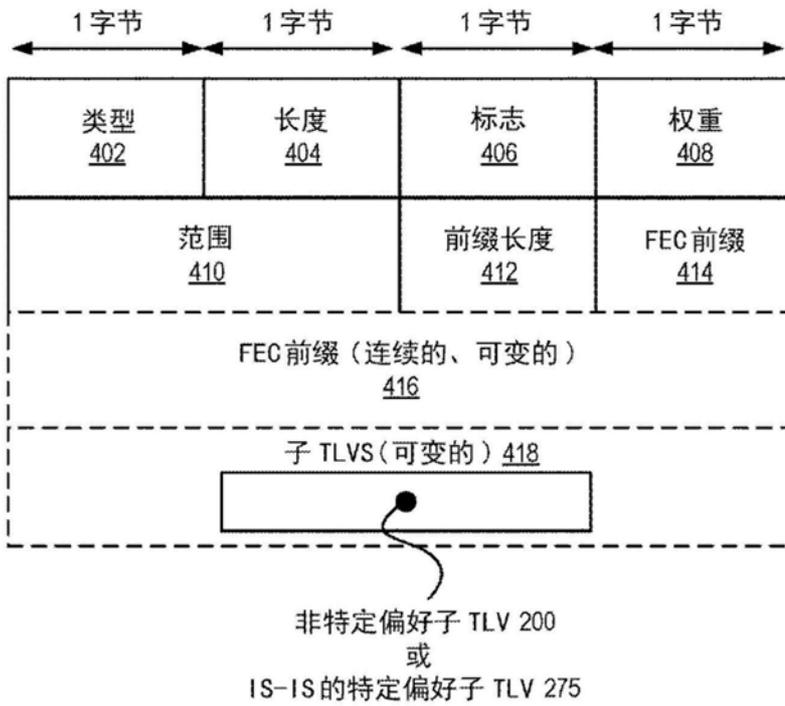


图3

专用 IS-IS TLV 149 (SID/ 标签绑定 TLV) 400



专用 IS-IS TLV 150 (多拓扑 SID/ 标签绑定 TLV) 450

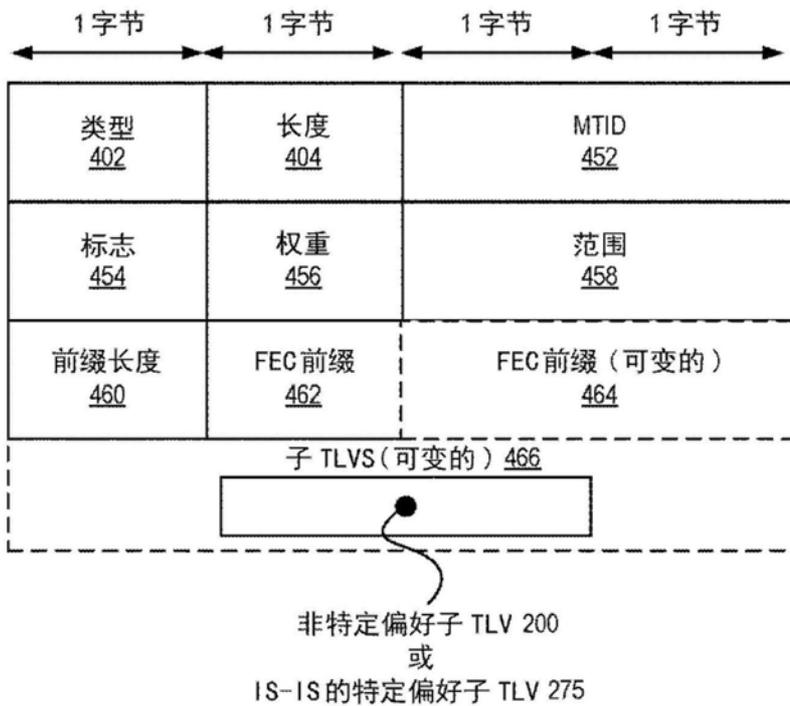


图4

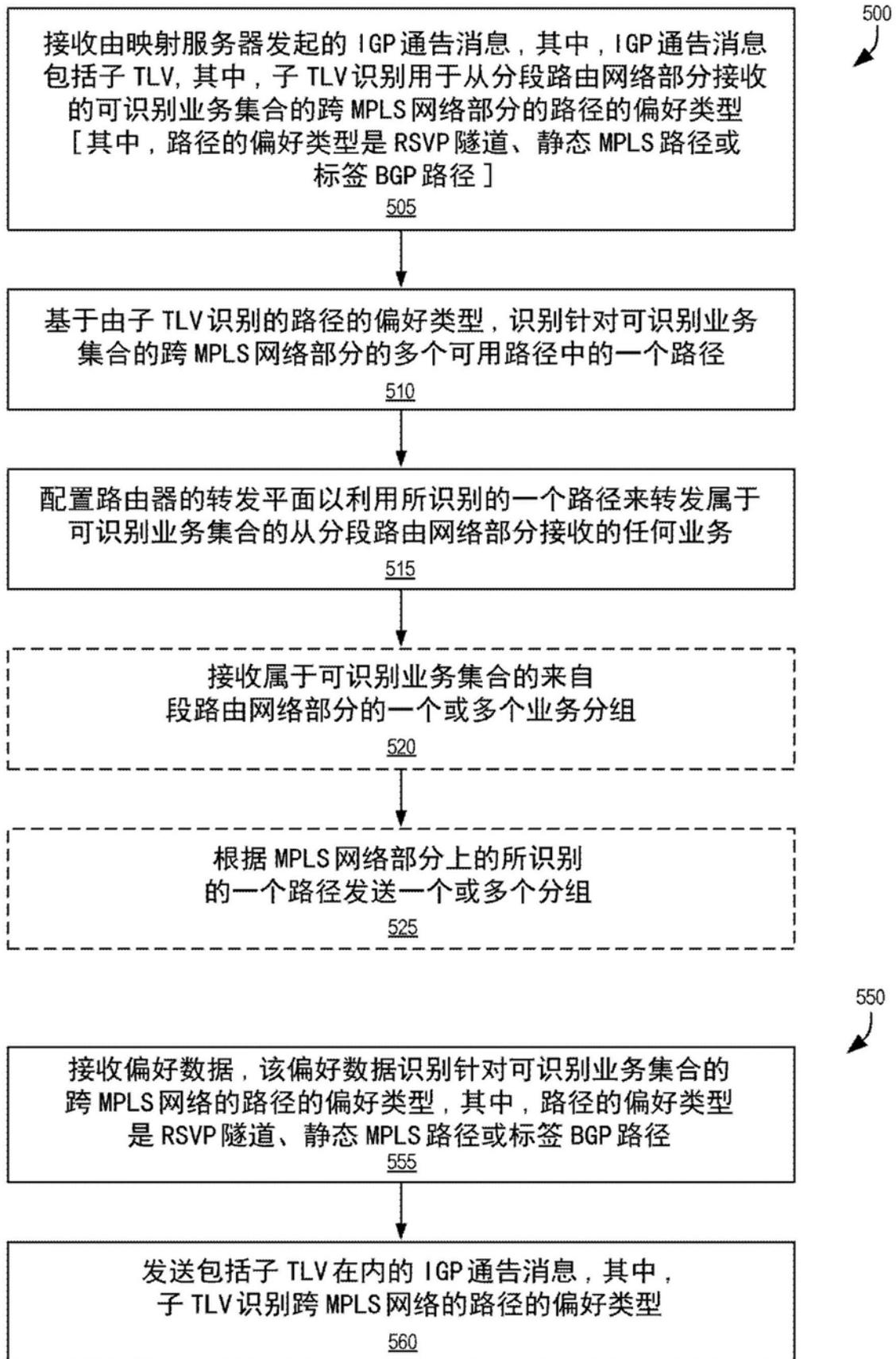


图5A

568

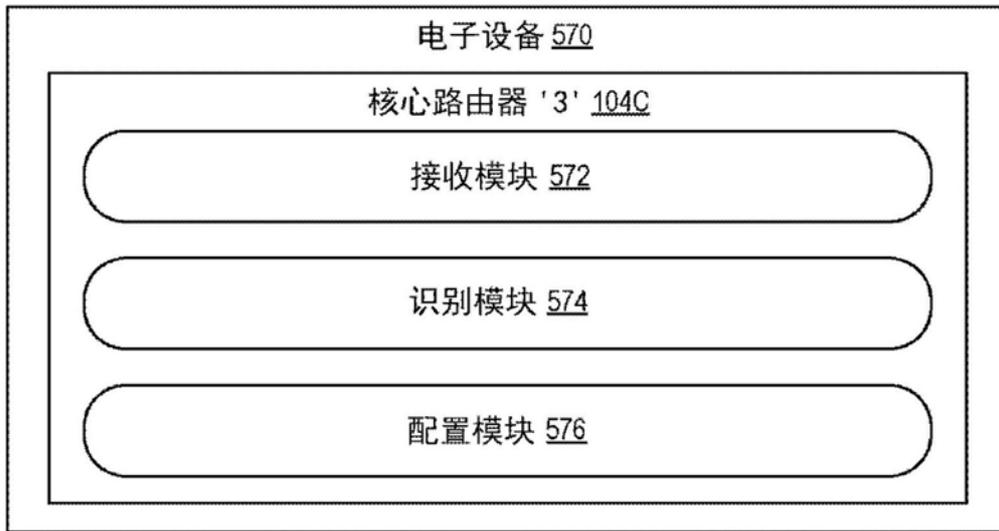


图5B

578



图5C

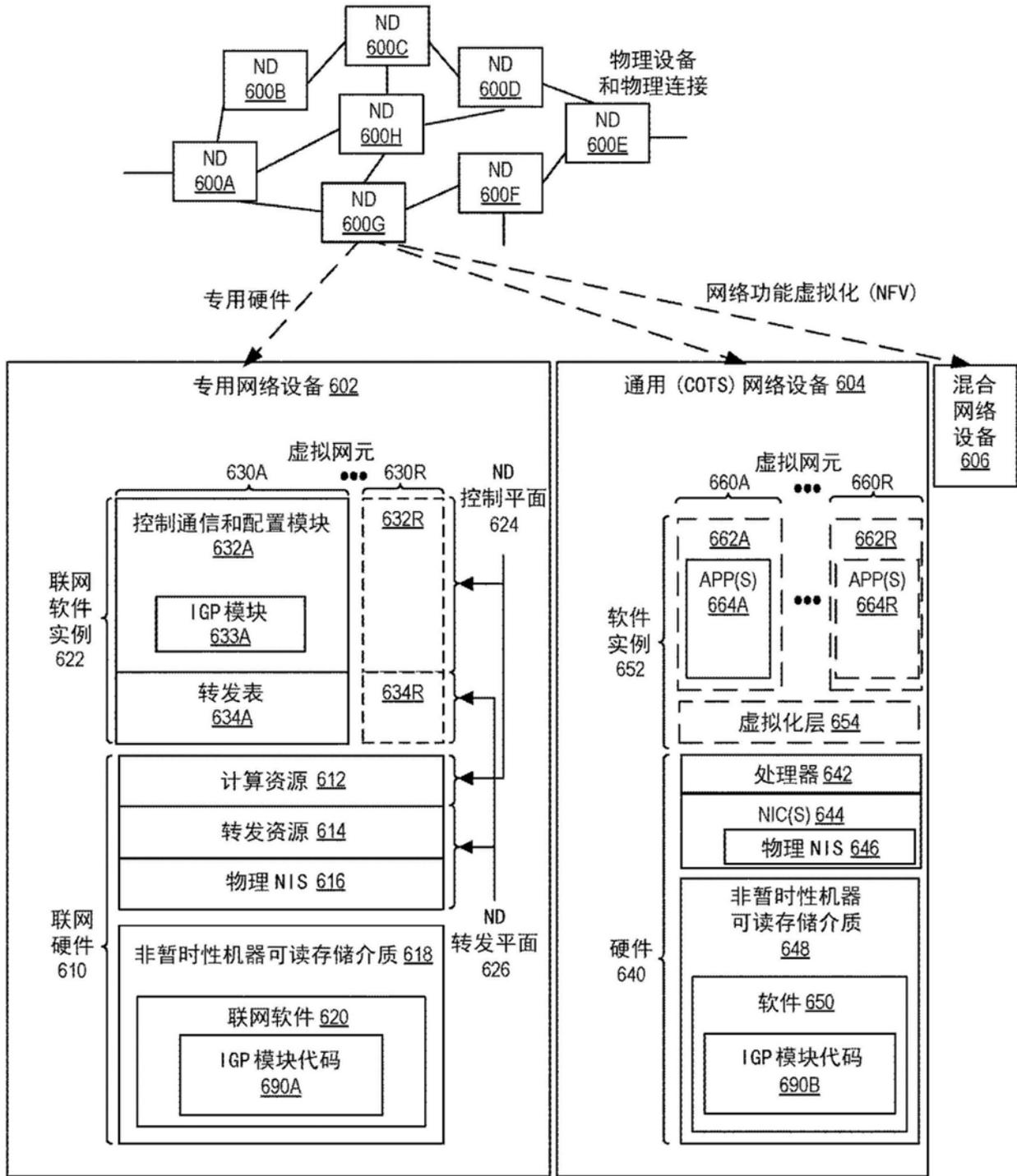


图6A

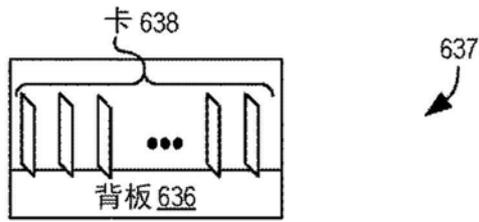


图6B

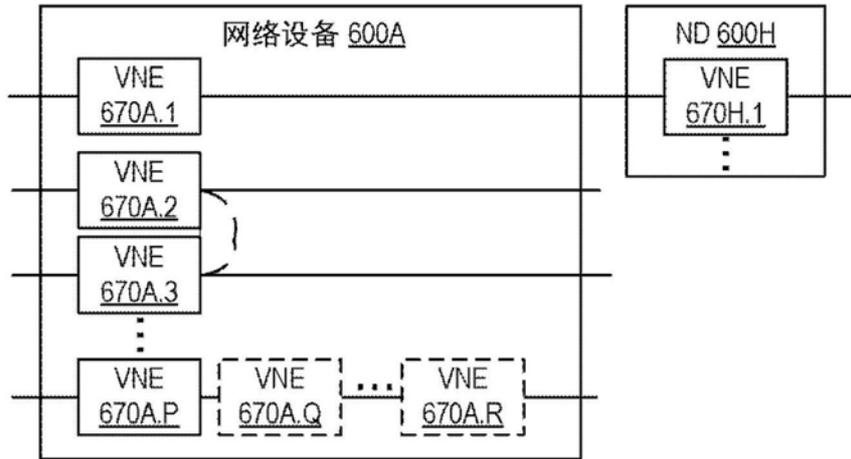


图6C

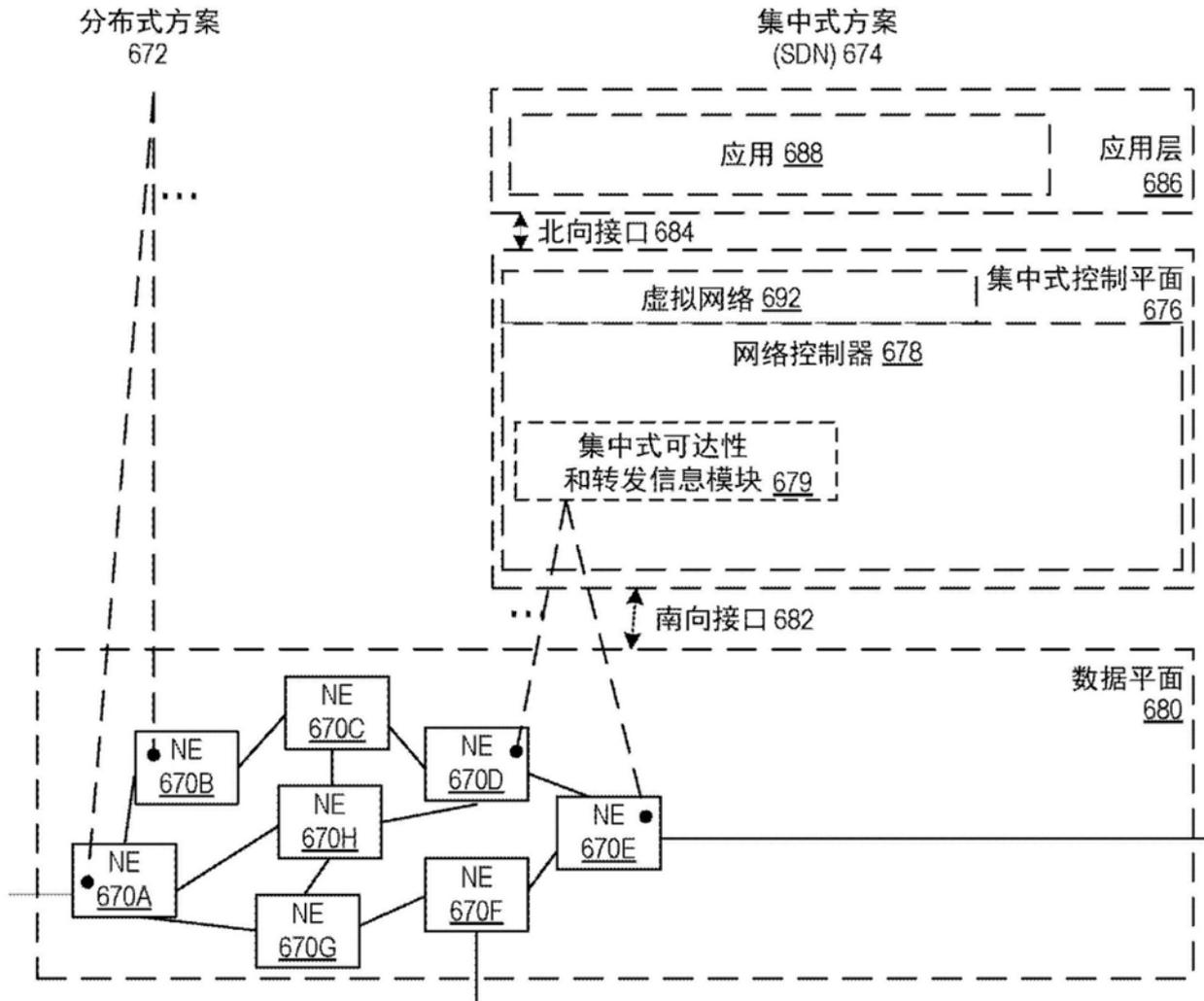


图6D

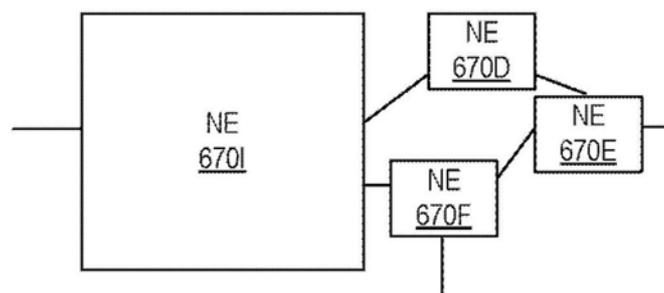


图6E

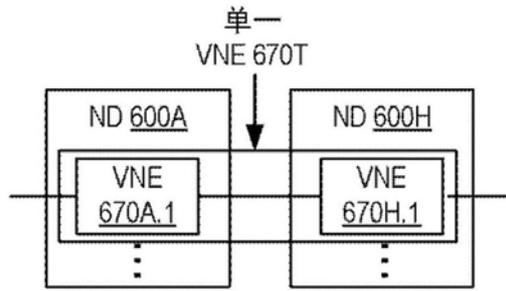


图6F