



(12)发明专利申请

(10)申请公布号 CN 106973018 A

(43)申请公布日 2017. 07. 21

(21)申请号 201710198697.1

H04L 12/703(2013.01)

(22)申请日 2012.03.26

H04L 12/24(2006.01)

(30)优先权数据

H04L 12/721(2013.01)

61/467,750 2011.03.25 US

H04L 12/753(2013.01)

13/372,596 2012.02.14 US

(62)分案原申请数据

201280013342.6 2012.03.26

(71)申请人 华为技术有限公司

地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 黎仁蔚 陈怀谟

(51)Int.Cl.

H04L 12/751(2013.01)

H04L 12/723(2013.01)

H04L 12/711(2013.01)

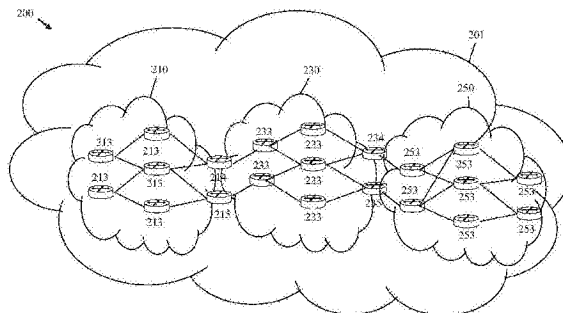
权利要求书3页 说明书15页 附图9页

(54)发明名称

一种在网络通信中实现拓扑透明区的系统和方法

(57)摘要

本申请公开了一种自治系统域,包含拓扑透明区,拥有多个拓扑透明区节点,其中,至少有几个节点是拓扑透明区边缘节点,其中,拓扑透明区节点通过多条内部链路互相连接,而拓扑透明区边缘节点通过多条外部链路和多个相邻的外部节点连接,其中,描述内部链路的链路状态通告(LSA)并没有扩散到相邻的外部节点。



1. 一种自治系统域,其特征在于,包括:

第一拓扑透明区,包括第一组拓扑透明区节点,所述第一组拓扑透明区节点互相连接,从而形成所述第一拓扑透明区的内部拓扑;

第二拓扑透明区,包括第二组拓扑透明区节点,所述第二组拓扑透明区节点互相连接,从而形成所述第二拓扑透明区的内部拓扑;

多个网络节点,连接所述第一拓扑透明区和第二拓扑透明区,从而形成自治系统域的拓扑,其中,所述第一组拓扑透明区节点各节点都包含一个第一链路状态数据库LSDB,所述第一链路状态数据库描述了所述自治系统域的拓扑和所述第一拓扑透明区的拓扑,但不包括所述第二拓扑透明区的拓扑;

其中,所述第一拓扑透明区的各节点利用第一LSDB来计算通往所述自治系统域和第一拓扑透明区内各目的地的最短路径,其中,各路径计算包括将所述自治系统域拓扑中外部TTZ链路的一条链路的度量作为一种特殊类型的度量。

2. 根据权利要求1所述的自治系统域,其特征在于,其中,所述特殊类型的度量比所述第一拓扑透明区的内部链路的度量一个大一个数量级。

3. 一种自治系统域,其特征在于,包括:

第一拓扑透明区,包括第一组拓扑透明区节点,所述第一组拓扑透明区节点互相连接,从而形成所述第一拓扑透明区的内部拓扑;

第二拓扑透明区,包括第二组拓扑透明区节点,所述第二组拓扑透明区节点互相连接,从而形成所述第二拓扑透明区的内部拓扑;

多个网络节点,连接所述第一拓扑透明区和第二拓扑透明区,从而形成自治系统域的拓扑,其中,所述第一组拓扑透明区节点各节点都包含一个第一链路状态数据库LSDB,所述第一链路状态数据库描述了所述自治系统域的拓扑和所述第一拓扑透明区的拓扑,但不包括所述第二拓扑透明区的拓扑;

其中,所述第一拓扑透明区的各节点利用第一LSDB来计算通往所述自治系统域和第一拓扑透明区内各目的地的最短路径,其中,各路径计算包括第一阶段和第二阶段。

4. 根据权利要求3所述的自治系统域,其特征在于,其中,所述第一阶段包括,计算通往所述第一拓扑透明区内所有目的地的最短路径,其中,所述第二阶段包括,在将通往所述第一拓扑透明区内所有目的地的最短路径的成本设为最小度量大约0或1之后,从第一阶段创建的最短路径树开始计算通往所述自治系统域内所有目的地的最短路径。

5. 一种自治系统域,其特征在于,包括:

第一拓扑透明区,包括第一组拓扑透明区节点,所述第一组拓扑透明区节点互相连接,从而形成所述第一拓扑透明区的内部拓扑;

第二拓扑透明区,包括第二组拓扑透明区节点,所述第二组拓扑透明区节点互相连接,从而形成所述第二拓扑透明区的内部拓扑;

多个网络节点,连接所述第一拓扑透明区和第二拓扑透明区,从而形成自治系统域的拓扑;

其中,所述第一组拓扑透明区节点各节点都包含一个第一链路状态数据库LSDB,所述第一LSDB描述了所述自治系统域的拓扑和所述第一拓扑透明区的拓扑,但不包括所述第二拓扑透明区的拓扑;

其中,所述第二组拓扑透明区节点各节点都包含一个第二LSDB,描述了所述自治系统域的拓扑和所述第二拓扑透明区的拓扑,但不包括所述第一拓扑透明区的内部拓扑。

6. 根据权利要求5所述的自治系统域,其特征在于,所述第二组拓扑透明区节点都不包含第一LSDB。

7. 一种位于自治系统域的拓扑透明区,其特征在于,包括:

多个互相连接的拓扑透明区节点,所述节点至少包括几个拓扑透明区边缘节点,其中,拓扑透明区边缘节点通过多条外部链路连接所述自治系统域内的多个相邻的外部节点;

多条内部链路,每条连接一对所述拓扑透明区节点,

其中,所述拓扑透明区节点之间相互交换多个描述所述内部链路的链路状态通告LSA,且无需将任何关于所述拓扑透明区内部拓扑的链路状态信息扩散到任一所述外部节点,

其中,一个拓扑透明区边缘节点创建了一个LSA并将其扩散,所述LSA包括:

一个LSA报头,包括一个链路状态类型字段,用于指定路由器LSA的类型;

一个路由器链路字段,包括多条路由器链路,每条链路描述所述外部链路和连接所述拓扑透明区边缘节点的所述内部链路中一条不相同的链路;

一个链路数量字段,用于指定所述路由器链路字段中路由器链路的数量。

8. 根据权利要求7所述的拓扑透明区,其特征在于,其中,所述多条路由器链路中的第一条链路包括:

一个链路ID字段,用于将所述第一条路由器链路确定为外部链路或内部链路;

一个链路类型字段,包括一个比特标识位和一个第一链路类型字段,用于说明所述第一条路由器链路是否为和另一台路由器的点到点P2P连接、传输网络的连接、末梢网络的连接或虚拟链路。

9. 根据权利要求8所述的拓扑透明区,其特征在于,其中,若所述第一条路由器链路是所述拓扑透明区的内部链路,则所述比特标识位设为1;若所述第一条路由器链路是从所述拓扑透明区向外扩展的外部链路,则所述比特标识位设为0。

10. 根据权利要求7所述的拓扑透明区,其特征在于,其中,所述多条路由器链路中的第一条链路包括:

一个链路ID字段,用于将所述第一条链路确定为外部链路或内部链路;

一个链路类型字段,用于说明所述第一条路由器链路是否为和另一台路由器的点到点P2P连接、传输网络的连接、末梢网络的连接或虚拟链路;

一个TTZ ID字段,其中,若所述第一条路由器链路是所述拓扑透明区的内部链路,则所述TTZ ID字段设为所述拓扑透明区ID;其中,若所述第一条路由器链路是从所述拓扑透明区向外扩展的外部链路,则所述TTZ ID字段设为0。

11. 一种位于自治系统域的拓扑透明区,其特征在于,包括:

多个互相连接的拓扑透明区节点,所述节点至少包括几个拓扑透明区边缘节点,其中,拓扑透明区边缘节点通过多条外部链路连接所述自治系统域内的多个相邻的外部节点;

多条内部链路,每条连接一对所述拓扑透明区节点,

其中,所述拓扑透明区节点之间相互交换多个描述所述内部链路的链路状态通告LSA,且无需将任何关于所述拓扑透明区内部拓扑的链路状态信息扩散到任一所述外部节点,

其中,一个拓扑透明区边缘节点创建了一个LSA并将其扩散,所述LSA包括:

一个LSA报头,包括一个链路状态类型字段,所述链路状态类型字段用于指定路由器LSA的类型;

一个路由器链路字段,包括多条路由器链路,每条链路描述一条不相同的内部链路,且所述内部链路连接所述拓扑透明区内部节点;

一个链路数量字段,用于指定所述路由器链路字段中路由器链路的数量。

12.根据权利要求11所述的拓扑透明区,其特征在于,其中,所述多条路由器链路中的第一条链路包括:

一个链路ID字段,用于标别一条内部链路;

一个链路类型字段,包括一个拓扑透明区比特标识位和一个链路类型1字段,所述链路类型1字段用于说明所述内部链路是否为和另一台路由器的点到点P2P连接、传输网络的连接、末梢网络的连接或虚拟链路。

13.根据权利要求12所述的拓扑透明区,其特征在于,其中,所述比特标识位设为1,从而表示所述路由器链路是所述拓扑透明区的内部链路。

14.根据权利要求12所述的拓扑透明区,其特征在于,其中,所述多条路由器链路中的第一条链路包括:

一个链路ID字段,用于标别一条内部链路;

一个链路类型字段,用于说明所述内部链路是否为和另一台路由器的P2P连接、传输网络的连接、末梢网络的连接或虚拟链路。

一种在网络通信中实现拓扑透明区的系统和方法

技术领域

[0001] 本申请涉及通信网络,尤其涉及一种在网络通信中实现拓扑透明区的系统和方法。

背景技术

[0002] 随着传统通信网络的不断扩大,例如,包含越来越多的节点,开始出现了可扩展性问题以及其他问题。在一个包括单自治系统(Autonomous System,AS)域(例如,如以下图1所示)的网络中,每个节点都需要了解所有其他节点的位置关系(即,邻接),如此一来,所有节点都会创建网络拓扑图。一个节点要可以通过链路状态信息在整个网络洪泛来了解另一个节点的邻接。洪泛是基于一个或多个内部网关协议(Interior Gateway Protocol,IGP)的,例如,开放式最短路径优先(Open Shortest Path First,OSPF)、中间系统(intermediate system,IS)到中间系统(IS-IS)协议等。具体地,涉及IGP的节点可以将描述其自身相关邻接的链路状态通告(Link State Advertisement,LSA)发送给所有与其相邻的节点,这些相邻节点又将收到的LSA转发给所有与它们相邻的节点(除了发送LSA的那个节点)。通过这种方法可以将LSA扩散到整个网络,如此一来,所有网络节点都能了解另一个节点的邻接,从而使得各种节点能够创建拓扑表(例如,链路状态数据库(link state database,LSDB))。LSA洪泛可以在网络初始化,或当网络邻接发生变化时(例如,一个节点被添加/删除,节点/链路故障等)进行。因此,随着添加到网络的节点增多,链路状态扩散开始消耗越来越多网络资源(例如,带宽、处理能力等),结果导致效率越来越低,耗时越来越多。

[0003] 现有主流技术中,为解决大型网络中的可扩展性以及性能问题,定义了IGP的小型本地区域(例如,开放式最短路径优先(Open Shortest Path First,OSPF)区域、中间系统(intermediate system,IS)到中间系统(IS-IS)区域等),试图降低整个网络中LSA洪泛的数量。这一技术(例如,如图2所示)已经在各种刊物中进行了描述,例如,Internet工程任务组(Internet Engineering Task Force,IETF)发布的请求注解(Request For Comments,RFC) 2328,题目为“OSPF版本2”(描述了AS域中的OSPF区域),和IETF发布的RFC1142,题目为“开放系统互连(Open Systems Interconnection,OSI) IS-IS域内路由协议”(描述了AS域内的IS-IS区域),等等。具体地,每个OSPF/IS-IS区域都包括一些互相连接的路由器,包括区域边界路由器(area border router,ABR)以及内部路由器。ABR和内部路由器的区别在于,ABR可以连接外部路由器(例如,连接其他OSPF/IS-IS区域中的路由器或ABR),而内部路由器只能连接该OSPF/IS-IS区域内的其他节点(例如,不连接该OSPF/IS-IS区域范围之外的任一路由器)。在大多数应用中,ABR和内部路由器会在各自的本地OSPF/IS-IS区域范围内执行正常的链路状态扩散(例如,根据IGP),从而使得ABR能够收集汇总那些描述其本地OSPF/IS-IS区域的拓扑信息(例如,构建汇总LSA)。之后,ABR可以将这些汇总LSA扩散给骨干区域的其他ABR(例如,扩散给与之相连的所有其他ABR),从而使得外部的ABR能够完全了解或部分了解围绕骨干区域的各OSPF/IS-IS区域的拓扑。而可否将这些汇总LSA扩散给其

他OSPF/IS-IS区域范围内的内部节点取决于网络配置。

[0004] 因此,虽然现有技术方法降低了整个网络中LSA洪泛的数量,但依然存在一个事实,即,在一个OSPF/IS-IS区域内,将链路状态扩散本地化(例如,描述OSPF/IS-IS区域范围内的内部邻接变化)会导致链路状态扩散到其他区域,从而触发这些区域的路由器重新计算其OSPF或IS-IS路由,更新其路由信息库(Routing Information Base,RIB)以及转发信息库(Forwarding Information Base,FIB)表,或采取其他消耗网络资源(例如,带宽和中央处理器(Central Process Unit,CPU)资源)的行为。

[0005] 在多区域或多域的环境下,要求使用中间的路径计算单元来帮助域间标签交换路径的计算。然而,PCE增加了网络管理的复杂度。如上所说,需要一种简单而有效的方法来解决大型网络中的可扩展性/收敛问题。

发明内容

[0006] 本申请公开了一种自治系统域,包含拓扑透明区,拥有多个拓扑透明区节点,其中,至少包含几个拓扑透明区边缘节点,其中,该拓扑透明区节点通过多条内部链路互相连接,而拓扑透明区边缘节点通过多条外部链路和多个相邻的外部节点连接,其中,描述内部链路的链路状态通告(Link State Advertisement,LSA)并不扩散到相邻的外部节点。

[0007] 本申请还公开了一种在自治系统域中配置拓扑透明区的方法,包括,给连接不同拓扑透明区节点的多条内部链路配置拓扑透明区标识(ID),其中,拓扑透明区节点位于拓扑透明区,且每条内部链路连接一对拓扑透明区节点;通过多条外部链路发送一般的路由器通告给自治系统域中的各相邻外部节点,其中,各相邻外部节点通过至少一条外部链路连接至少一个拓扑透明区节点;基于一般的路由器通告,创建或更新相邻外部节点中的拓扑表,其中,拓扑表将拓扑透明区描述为关联一个路由器ID的单路由器,该路由器ID相当于拓扑透明区ID。

[0008] 本申请还公开了一种自治系统域中的拓扑透明区,包括多个互相连接的拓扑透明区节点,其中,至少包含几个拓扑透明区边缘节点,而拓扑透明区边缘节点通过多条外部链路连接到自治系统域范围内的多个相邻外部节点;多条内部链路,其中,每条链路和一对拓扑透明区节点互相连接,而拓扑透明区节点可以互相交换多个描述内部链路状态的LSA,且无需将任一关于拓扑透明区内部拓扑的链路状态信息扩散到任一外部节点,其中,一般的路由器通告被构建并扩散到相邻的外部节点。

[0009] 本申请还公开了一种自治系统域,包括第一拓扑透明区,其中,包括第一组互相连接的拓扑透明区节点,从而形成第一拓扑透明区的内部拓扑;第二拓扑透明区,其中,包括第二组拓扑透明区内互相连接的拓扑透明区节点,从而形成第二拓扑透明区的内部拓扑;多个和第一拓扑透明区以及第二拓扑透明区互相连接的网络节点,从而形成自治系统域的拓扑,其中,第一组拓扑透明区中的每个节点都包括第一链路状态数据库(link state database,LSDB),描述了自治系统域的拓扑以及第一拓扑透明区的拓扑,而非第二拓扑透明区的拓扑。

[0010] 结合下文的详细描述、附图以及权利要求,能够更清楚地了解上述以及其他特征。

附图说明

[0011] 结合以下附图的简要描述以及详细描述,能够对本发明所公开的内容进行更加全面的理解,其中,类似的附图标记表示类似的部件。

[0012] 图1展示了本发明实施例提供的一种网络,包含一个未划分的AS域。

[0013] 图2展示了本发明实施例提供的一种网络,包含一个被划分为多个OSPF/IS-IS区域的AS域。

[0014] 图3展示了本发明实施例提供的一种网络,包含一个AS域,其中,部分被划分为拓扑透明区(Topology Transparent Zone,TTZ)。

[0015] 图4展示了,从位于TTZ实例中TTZ节点的角度出发得出的网络全景。

[0016] 图5展示了本发明实施例提供的一种TTZ路由器,包含六个接口。

[0017] 图6展示了本发明实施例提供的一种TTZ路由器,包含128个接口。

[0018] 图7展示了本发明实施例提供的一种IGP数据包报头。

[0019] 图8展示了本发明实施例提供的一种链路状态通告(Link State Advertisement,LSA)报头。

[0020] 图9展示了本发明实施例提供的一种路由器LSA。

[0021] 图10展示了本发明实施例提供的一种路由器链路字段。

[0022] 图11展示了本发明实施例提供的一种链路类型字段。

[0023] 图12展示了本发明实施例提供的一种包含新字段TTZ ID的路由器链路。

[0024] 图13展示了本发明实施例提供的一种在AS域中创建TTZ的方法。

[0025] 图14是本发明实施例提供的一种在网络中传输数据的网络单元的示意图。

[0026] 图15是本发明实施例提供的一种通用网络器件的示意图。

具体实施方式

[0027] 首先要注意的是,虽然下文详细说明了一个或多个实施例的实现方式,其公开的系统和/或方法可以通过许多其他已知或现存的技术来实现。本发明并未局限于下文描述的具体实现方式、附图以及技术,包括本文描述说明的较佳设计以及实现方式,但是可以在以下权利要求的范围内结合其对应内容的全部范围进行修改。

[0028] 本发明公开了一种在网络中创建一个或多个拓扑透明区(Topology Transparent Zone,TTZ)的方法和架构,来更好地解决大型网络中的可扩展性和收敛问题。具体地,相较于在网络中实现OSPF/IS-IS区域,实现TTZ能够更有效地降低链路状态扩散,虽然(同时),相较于在网络中实现PCE通信协议(PCE Communication Protocol,PCEP),其中,该协议用于计算跨域的多协议标签交换流量工程LSP(Multi-Protocol Label Switching Traffic Engineering LSP,MPLS TE LSP)的路径,实现TTZ降低了复杂度。和本地OSPF/IS-IS区域类似,TTZ可以包括多个互相连接的节点(例如,TTZ节点),包括连接外部节点的节点(例如TTZ边缘节点)以及只连接TTZ内部节点的节点(例如,TTZ内部节点)。然而,和本地OSPF/IS-IS区域不一样的是,TTZ边缘节点不会将描述TTZ内部拓扑(例如,一对TTZ节点间的邻接)的LSA扩散到TTZ之外,从而降低了整个网络中的链路状态扩散。进一步地,从整体来看,TTZ中的节点了解AS域的拓扑。因此,LSP的路径计算能够不依赖于任何媒介(例如,PCE)。如上所说,虽然本文描述的TTZ技术允许简单计算LSP路径(即,不使用媒介),但仍然可以增强网络的可扩展性(例如,重新收敛更快、性能更高、处理消耗更低)。

[0029] 图1展示了网络100,包含单AS域101,其中,该AS域包括多个节点110到158,这些节点共同参与到一般的链路状态扩散。具体地,节点110到158包括源节点110,多个中间节点112到156,以及目的节点158,各节点互相连接,从而形成AS域101的内部架构(此处称作网络拓扑)。节点110到158可以是任意设备,能够在整个AS域101中以一种确定的方式进行数据交换;且可以包括一个或多个表(例如,LSDB、RIB以及FIB表),其中包含描述AS域101拓扑的链路状态信息。在本实施例中,可以建立横跨AS域101的LSP(实线箭头),从而允许将数据以一种确定的方式从网络100的一端传输到另一端。如图所示,LSP可以通过一个或多个中间节点112到156,从源节点110扩展到目的节点158。

[0030] 随着网络100的扩大(例如,包含越来越多的节点),可扩展性问题随之出现。具体地,可扩展性问题涉及各种网络问题,都是因为无法有效地在大型网络中管理/扩散网络拓扑信息。例如,各节点(例如,节点132)的处理和存储能力/容量会限制网络100的实际规模。此外,及时扩散链路状态信息的能力(例如,收敛)与必须进行LSA交换/洪泛的节点的数量直接相关,因此,越是大型的网络,收敛(例如,用来创建/更新拓扑表的时间段)的时间则越长。正如此处所应用的,拓扑表可以对应LSDB,或其他表,用来描述相应网络、AS域、本地区域或TTZ的拓扑。在大型网络中,重新收敛也是一个问题,如无法及时从一个节点/链路的故障中恢复(或某些其他条件,改变了初始化之后的网络邻接)会中断网络业务(例如,网络中建立的多个LSP)。具体地(当网络存在故障时),需要在整个网络100中进行LSA洪泛,从而通知每个网络节点110到158关于网络拓扑的改变。正如此处所应用的,在网络中扩散拓扑信息(以及更新相关的节点中相应的RIB/FIB)需要的时长可以称作重新收敛期(如果发生在网络初始化期间,则称作收敛期)。

[0031] 重新收敛期会显著影响LSP(例如,LDP LSP)在网络故障时及时恢复的能力,因此,严重影响了网络支持的最高服务质量(Quality of Service,QoS)等级。具体地,网络故障会影响一条或多条LSP,其中,在恢复过程能够完成前,LSP都会一直处于故障状态。例如,节点134(或节点132和134之间的链路)故障会导致LSP(实线箭头)暂时故障。在检测到故障后,一个或多个节点110到158会在整个网络中进行LSA洪泛,从而通知其他节点关于拓扑的变化(即,节点132和134之间的邻接不再存在,或暂时不可用)。接着,会重新计算一个或多个节点110到158,否则,根据各种已知的恢复机制来修复LSP(如虚线所示)。只有在恢复了LSP之后,网络流量传输才能恢复。因此,重新收敛期直接影响LSP在网络故障时及时恢复的能力,因此,严重影响网络满足其各种业务协定对QoS要求的能力。

[0032] 如上所述,现有技术在处理大型网络中的可扩展性/收敛问题时,涉及对网络AS域内OSPF/IS-IS区域的定义。图2展示了网络200,包含AS域201,其中,201被划分为多个本地区域(例如,OSPF/IS-IS区域)。具体地,AS域201可以包括多个本地区域210、230和250。虽然图中只展示了三个本地区域210、230和250,本领域普通技术人员都会知道,AS域201中可以包括的本地区域数量不限(例如,数十个甚至上百个OSPF/IS-IS区域)。本地区域210可以包括多个边界节点214和215,以及多个内部节点213。边界节点214和215连接外部节点(即,位于本地区域210之外的节点),而内部节点213只连接内部节点(即,位于本地区域210之内的节点)。本地区域230可以包括多个边界节点214、215、234和235,以及多个内部节点233,可以参照本地区域210中的相应节点214和215进行配置。本地区域250可以包括多个边界节点234和235,以及多个内部路由器253,可以参照本地区域210中的相应节点214和215进行配

置。

[0033] LSA包括关于本地区域210的本地链路状态信息,且被扩散(例如,洪泛)到本地区域210中的所有节点213到215(但不包括节点233、234和235)。相应地,边界节点214和215会将本地链路状态信息收集汇总到汇总LSA,接着,将汇总LSA转发给相关的外部节点(即,外部节点233以及边界节点234和235)。汇总LSA在本地区域230中洪泛。然后,边界节点234和235会将汇总LSA在本地区域250中洪泛。因此,将AS域201划分为本地区域210、230和250可以降低在AS域201中的链路状态扩散,相较于网络100,节省了网络资源(例如,带宽等)。然而,在本地区域210中的链路状态扩散依然会引起在AS域201范围内至少一些链路状态的扩散,从而导致其他节点(例如,233到235,以及253)重新计算其OSPF或IS-IS路由,更新其RIB以及FIB表,或采取其他消耗网络资源(例如,带宽和CPU资源)的行为。因此,虽然现有技术将AS域划分为多个OSPF/IS-IS区域,部分解决了可扩展性/收敛问题,它依然无法将相关OSPF/IS-IS区域中的本地链路状态信息扩散充分本地化。

[0034] 相对地,本文公开的TTZ方案能够将相关TTZ中的本地链路状态信息扩散充分本地化。图3展示了本发明实施例提供的一种网络300,包含AS域301,其中,采用了示例的TTZ方案。具体地,AS域301可以包括多个节点311到335,以及拓扑透明区(Topology Transparent Zone, TTZ) 360。节点311到335可以参照节点110到158进行配置,互相之间(以及和TTZ 360之间)可以通过相应的IGP(例如,OSPF、IS-IS等)来共享链路状态信息。

[0035] TTZ 360可以包括多个TTZ节点361到373,其中(更确切地说)包括TTZ边缘节点361、363、365和367,以及多个TTZ内部节点371和373。正如此处所应用的,TTZ边缘节点和TTZ内部节点都可以称作TTZ节点。每个TTZ边缘节点361、363、365和367都可以连接至少一个位于TTZ 360之外的节点,例如,TTZ边缘节点361可以连接节点315。TTZ内部节点371和373不可以连接任何一个位于TTZ 360之外的节点,例如TTZ内部节点371只能连接TTZ节点361、363、365、367和373,但是不能连接位于TTZ 360之外的节点。

[0036] 只和TTZ 360的内部拓扑(例如,涉及内部链路或一对TTZ节点之间的邻接)相关的链路状态信息可以只在TTZ 360中洪泛(即,不扩散到TTZ 360之外的节点)。因此,节点311到335不会收到描述TTZ 360的内部拓扑的LSA,因此,无法“看”到(例如,视)TTZ 360的内部拓扑。相反,节点311到335会将TTZ视作一台单路由器/节点。此处节点和路由器两个词可以互换。

[0037] 可以给TTZ 360分配一个单路由器标识(ID)。分配给TTZ(例如TTZ 360)的路由器ID在文中也称作TTZ ID,且可以配置在相关TTZ的每条链路上,如此一来,互相连接的TTZ节点(即,链路两端的节点)知道该链路位于TTZ之内。换言之,给链路配置TTZ ID表示/说明链路是TTZ链路。在本实施例中,若要给链路配置TTZ ID,可以将每个互相连接的节点中的TTZ ID关联给该链路。例如,可以将对应TTZ 360的TTZ ID配置给一条TTZ内部节点371和373之间的链路,从而触发更新每个TTZ内部节点371和373中相应的表(例如,链路表)来关联包含TTZ ID的链路(或对应该链路的端口)。此流程可以从各TTZ内部节点371和373的角度出发,表明所配置的链路位于TTZ 360内。

[0038] 通过使用相同的TTZ ID配置每条TTZ 360内的链路(例如,每条内部链路的两端节点使用TTZ ID关联这条链路),就形成了TTZ 360。例如,要形成一个TTZ 360,可以给TTZ 360内每条链路配置相同的TTZ ID,包括TTZ节点361和365之间的链路、TTZ节点365和367之

间的链路、TTZ节点367和363之间的链路、TTZ节点363和361之间的链路、TTZ节点371和361之间的链路、TTZ节点371和363之间的链路、TTZ节点371和365之间的链路、TTZ节点371和367之间的链路以及TTZ节点371和373之间的链路。一旦给链路配置了对应TTZ 360的TTZ ID,TTZ 360的内部拓扑可以由六个TTZ节点361、363、365、367、371和373,及其链路来定义。

[0039] 从外部节点311到335的角度出发,TTZ 360可以被“看”(例如,视,或识别)作单个节点,拥有单台路由器ID。在某些实施例中,路由器ID可以是全局的TTZ ID,虽然,它还可以是分配给TTZ节点的最小或最大的路由器ID。进一步地,外部节点311到335可以将多个连接TTZ 360的链路“看”(例如,查看,或识别)作从属于单节点的入/出接口。例如,节点315可以将TTZ 360“看”作单节点,包含六个接口,连接节点315、317、323、325、331和329。

[0040] 图4展示了,从位于TTZ实例中TTZ节点的角度出发得出的网络全景400。网络400包含一个AS域401,包括多个外部节点405和465,多个非TTZ实例410、420、440、450和460,以及一个TTZ实例430。可以对每个TTZ 410、420、430、440、450和460进行相似的配置,如此一来,只关联TTZ内部拓扑的链路状态信息就不会扩散到各个TTZ之外。TTZ实例包括多个节点431到439,每个节点都可以“看”到TTZ实例430的内部拓扑,以及AS域401的拓扑。然而,节点431到439不能“看”到任一非TTZ实例410、420、440、450和460的内部拓扑。相反地,从位于TTZ 430的节点431到439的角度出发,非TTZ实例410、420、440、450和460是作为单路由器/节点出现的。

[0041] 此处描述的一种关于TTZ技术的应用,可以将访问点(point of presence,POP)(例如,一层楼或一个房间的路由器)配置为TTZ,其中,整个POP作为单台路由器(即,POP中的内部路由器/链路对外部节点不可见)。在其他实现方式中,此处描述的TTZ方法允许在POP应用中达到更高的可扩展性、更简洁的LSP计算/建立、更快的路由收敛(例如,在故障之后)、更好的性能以及更高的可用性。

[0042] 可扩展性涉及在网络或域范围内快速有效地将链路状态信息扩散给相关节点的能力。例如,一个包括1000台路由器的标准AS域(即,没有TTZ),需要执行至少1000次LSA扩散,从而在每个节点内建立一个链路状态数据库(link state database,LSDB)。相反,将网络划分为多个TTZ能够极大地降低LSA的数量。例如,如果将网络划分为200个TTZ(每个TTZ包括大约五台路由器),仅利用约205个LSA(例如,200个TTZ路由器LSA以及5个位于各TTZ范围内的LSA(同时出现,且在相当小的范围内洪泛))来创建LSDB,而可扩展性提升可以达到约79.5%(可以通过以下公式计算得出: $(1000-205)/1000$)。将网络划分为更少的(例如,更大的)TTZ能够提升更多。例如,如果将包含1000个节点的网络划分为100个TTZ(每个TTZ包括大约10台路由器),可扩展性提升可以达到约89%(可以通过以下公式计算得出: $(1000-110)/1000$), (例如,约一个数量级)。在更大型的网络中,可扩展性提升效果会更加显著。例如,将包括大约200,000台路由器的网络划分为约1000个TTZ(每个TTZ包括约200台路由器),可扩展性提升可以达到约99.4%(可以通过以下公式计算得出: $(200,000-1,200)/200,000$), (例如,大约两个数量级)。

[0043] 本文公开的TTZ方法中,计算和/或建立端到端(end-to-end,E2E)LSP(例如,将LSP从AS域的一端扩展到另一端)可以更快/更简洁。在传统的本地OSPF/IS-IS区域网络中,E2E LSP计算能够非常复杂,经常要用到中间的路径计算单元(path computation entity,PCE)。相反,本文公开的TTZ方案中,E2E LSP计算和建立相对简洁。如图4所示,建立外部路

由器405和外部路由器465之间的LSP (实线箭头)。具体地,针对位于TTZ 410、430和450之外的LSP部分,可以通过传统的手段计算(例如,可以由外部路由器405计算)并建立(例如,用PATH消息等)。例如,LSP可以先在外部节点405和TTZ 410之间建立,然后在TTZ 410和TTZ节点431之间,接着在TTZ节点431和TTZ节点437之间,之后在TTZ节点437和TTZ 450之间,最后在TTZ 450和外部节点465之间。接下来(或结合以上内容),可以计算/建立位于TTZ 410、430和450之内的LSP部分。例如(在TTZ实例430中),LSP可以先在TTZ节点431和TTZ节点435之间建立,然后在TTZ节点435和TTZ 437之间。在TTZ 410和450中,可以通过相似的步骤来建立LSP在TTZ 410和450中的部分。因此,不使用PCE(或其他外部数据库/CPU)也能够建立E2E LSP。在某些实施例中,LSP各部分可以独立计算,同时建立,通过一般的机制(例如,单条PATH消息或一串PATH消息)。例如,位于TTZ 410、430和450之外的LSP部分可以由外部路由器405来进行计算,而相应的跨越各TTZ的LSP部分可以由位于各TTZ 410、430和450之中相应的路由器(例如,TTZ节点431等)来进行计算。

[0044] 此外,本文公开的TTZ方法可以极大地减少重新收敛的时间,从而允许更快的故障恢复(例如,重新收敛时间更短)。例如,如果TTZ节点431和TTZ节点435之间的LSP部分发生故障,能够快速计算并建立TTZ节点431、433和437之间相应的LSP备份部分(虚线箭头),且无需在TTZ 430之外进行任何链路状态扩散。因此,TTZ 430之外的路由器不会关注拓扑变化,(作为添加的益处)并不会重新计算其路由表(即,节约CPU资源,从而增加上述关于其他任务的资源的“可用性”)。当对POP(例如,配置为TTZ)进行重建或改造时,TTZ外部的路由器并不会受到影响。如上所说,将只关联TTZ 430内部拓扑的链路状态扩散本地化不仅极大地缩减了重新收敛的时间,而且将LSP恢复透明化(即,不通知外部节点405、465以及TTZ 410和450)。

[0045] 此外,本文公开的TTZ方法允许在TTZ区域外部实现更高的可用性,且无需消耗CPU资源。具体地,传统的OSPF/IS-IS区域网络将汇总LSA在整个网络中洪泛(例如,先在OSPF/IS-IS ABR之间,然后,在外部OSPF/IS-IS网络范围内),从而消耗了OSPF/IS-IS区域外部的处理资源。相反,本文公开的TTZ方法在TTZ 430中将LSA完全本地化,从而避免了不必要的外部处理资源消耗。

[0046] 本文公开的TTZ方法还提升了网络性能。在某些实施例中,TTZ采用一种TTZ边缘路由器,包含至少两个接口(例如,4个、6个、8个等等)。图5展示了网络500,其中,TTZ路由器包括8个接口。具体地,网络500包括多个TTZ 510、515、520、540、545、550以及一个TTZ实例530,配置和上述讨论的TTZ 410到460类似。TTZ实例530可以包括4个TTZ边缘节点531到539,每个节点包括4个接口。因此,TTZ实例530可以有8个接口,从而将单一的四接口路由器的吞吐率提高大约两倍(例如,8/4)。例如,如果TTZ边缘路由器531到539是40千兆赫(G)的路由器(例如,4个10G的接口),则TTZ530提供的吞吐率能达到约80G。

[0047] 在其他实施例中,TTZ可以采用一种包含更多接口的TTZ边缘路由器,和/或采用更多TTZ边缘路由器。图6展示了网络600,其中,TTZ路由器包含128个接口。具体地,网络600包括一个TTZ实例630(为了表述清晰,此处忽略了非TTZ实例),和上述TTZ430配置相似,可以包括16个TTZ边缘路由器601到616。其中,每个TTZ路由器601到616都可以包括约16个100G的接口(例如,TTZ边缘路由器包括16个接口,其中,8个接口是从接口6011到6018,8个接口是从接口6091到6098,8个接口是从接口6081到6088,8个接口是从接口6161到6168),总吞

吐率为每TTZ边缘路由器601到616,约1.6兆兆字节(1.6T),因此TTZ 630能达到的吞吐率约为12.8T(或者,约为单台1.6T的路由器的8倍)。本领域的普通技术人员都知道,虽然图6所示的实现方式中,TTZ 630包括了16个TTZ边缘路由器以及128个接口,在其他实施例中,包括的接口和/或路由器(例如,路由器吞吐率更高/更低,等)可以更多或更少。

[0048] 图7展示了本发明实施例提供的一种内部网关协议(Interior Gateway Protocol, IGP)数据包报头700,和OSPF数据包报头类似。数据包报头700包括版本号(No.)字段701、类型字段703、数据包长字段705、路由器ID字段707、区域ID字段709、检查和字段711、认证类型(AuType)字段713以及认证字段717。版本号(No.)字段701可以包括OSPF协议的版本号,(在本实施例中)可以设为大约2,表示其兼容OSPF协议的第二个版本。类型字段703可以表示OSPF数据包的类型,可以设为大约1、2、3、4或5,从而表示,相应的OSPF数据包(分别是)是Hello包、数据库描述包、链路状态请求包、链路状态更新包或链路状态确认包。数据包长字段705的值可以表示OSPF协议数据包(包括标准OSPF报头)的长度,单位为字节。路由器ID字段707的值可以表示源路由器(即,首先进行对应数据包报头700的数据包传输的路由器)的路由器ID。区域ID字段709的值可以包括32比特的数字,用来标识不同的区域(例如,OSPF区域),该区域由相应的数据包描述或与其相关联。检查和字段711的值可以表示数据包整体内容的标准IP检查和,始于OSPF数据包报头(但不包括64比特的认证字段717)。AuType字段713的值可以标识针对数据包进行的认证流程,可以是大约0、1或2,从而(分别)表示Null认证、Simple密码或Cryptographic认证。认证字段717可以是一个64比特的字段,用来进行认证。

[0049] 在本实施例中,当TTZ(边缘)节点向其他节点(例如,在TTZ之外)发送一个包括报头700的数据包时,路由器ID字段707的值可以包括适当的TTZ ID。例如,在构建并发送IGP数据包给节点315时(即,TTZ 360之外的一个节点),TTZ边缘节点361(见图3)可以将路由器ID字段707的值设为TTZ ID,对应TTZ 360。在另一个实施例中,当TTZ边缘节点向另一个节点(例如,在TTZ之外)发送一个包括报头700的数据包时,路由器ID字段707的值可以包括TTZ之内路由器ID中最大的路由器ID或最小的路由器ID。例如,假设TTZ节点361、363、365、367、371和373的路由器ID(分别是)10.1.1.161、10.1.1.163、10.1.1.165、10.1.1.167、10.1.1.171和10.1.1.173。向节点315(即,TTZ 360之外的节点)发送数据包时,TTZ边缘节点361可以将路由器ID字段707的值设为TTZ 360中最大的路由器ID(即,10.1.1.173)或最小的路由器ID(即,10.1.1.161)。

[0050] 图8展示了本发明实施例提供的一种LSA报头800,包括链路状态年龄(LS Age)字段801、选项字段803、链路状态类型(LS Type)字段805、链路状态ID字段807、发布路由器字段809、链路状态序列号字段811、链路状态检查和字段813以及长度字段815。在本实施例中,所有LSA都从LSA报头800开始,而链路状态类型(LS Type)字段805、链路状态ID字段807和发布路由器字段809可以用来对LSA进行唯一标识。LS Age字段801的值可以表示自从LSA生成到目前的时间(单位是秒)。选项字段803的值可以表示一个或多个由路由域的相应部分支持的可选能力。在本实施例中,链路状态类型字段805可以表示LSA的类型,可以设为大约1、2、3、4或5,从而(分别)表示,路由器LSA、网络LSA、IP网络汇总LSA、AS边界路由器(AS Border/Boundary Router, ASBR)汇总LSA或AS外部LSA。在同样或其他的实施例中,LS类型字段805可以设为大约6、9、10或11,从而(分别)表示,组成员LSA、不透明LSA(本地链路范

围)、不透明LSA(本地区域范围)或不透明LSA(AS范围)。链路状态ID字段807的值可以标识由LSA描述的互联网环境的一部分。发布路由器字段809可以包括源路由器(即,创建LSA的路由器)的路由器ID。链路状态序列号字段811的值可以是32比特的数字,用来表示LSA的实例。链路状态检查和字段813的值可以包括LSA整体内容的检查和(例如,Fletcher检查和),包括LSA报头800(但不包括LS Age字段801)。长度字段815的值可以表示LSA的长度(包括LSA报头800的长度,单位是字节)。

[0051] 当给TTZ外部的路由器发送LSA时,TTZ(边缘)节点可以将链路状态类型字段805设为大约1,从而表示该LSA为路由器LSA;将链路状态ID字段807以及发布路由器字段809设为相应的TTZ ID,从而表示该LSA对应TTZ。另外的选择是,TTZ(边缘)节点可以将链路状态ID字段807以及发布路由器字段809都设为TTZ 360中最大的路由器ID或最小的路由器ID。

[0052] 图9展示了本发明实施例提供的一种路由器LSA 900,包括LSA报头字段920、标识位字段930、链路数量字段940以及多个路由器链路字段951。路由器LSA 900的LSA报头字段920可以包括和上述LSA报头字段800类似的格式,可以包括链路状态年龄字段901、选项字段903、链路状态类型字段905、链路状态ID字段907、发布路由器字段909、链路状态序列号字段911、链路状态检查和字段913以及长度字段915。在本实施例中,链路状态类型字段905可以设为大约1,从而表示LSA 900是一个路由器LSA。

[0053] 标识位字段930可以表示生成LSA 900的路由器的特征,包括一个虚拟链路(V)比特931、外部(E)比特932、边界(B)比特933。V比特931可以设为大约1,从而表示,源路由器是一个或多个全连接的虚拟链路的端点。E比特932可以设为大约1,从而表示,源路由器是一个AS边界路由器。B比特933可以设为大约1,从而表示,源路由器是一个ABR。链路数量字段940可以是16比特的数字,表示在路由器链路字段950中所描述的路由器链路的数量。路由器链路字段950可以包括路由器链路951,针对源路由器区域(例如,TTZ)中的每条路由器链路。每个路由器链路951可以描述TTZ中一个单独的路由器链路。

[0054] 图10展示了本发明实施例提供的一种路由器链路1000,类似上述的路由器链路951。路由器链路1000可以包括链路ID字段1001、链路数据字段1003、类型字段1011、服务类型数量(NoTOS)字段1013、度量字段1015和多个具体服务类型(TOS)链路度量字段1020到1040。类型字段1011可以表示路由器链路1000中描述的链路的种类,会影响路由器链路1000中所有其他字段1001、1003、1013到1040(即,字段1001、1003、1013到1040的值会变化和/或表达不同的意义,这些都取决于类型字段1011的值)。类型字段1011可以设为大约1、2、3、或4,从而(分别)表示,路由器链路1000描述了和其他路由器的点到点(P2P)连接、和传输网络的连接、和末梢网络的连接或虚拟链路。

[0055] 链路ID字段1001可以标识特定的路由器链路所连接的对象,取决于类型字段1011的值。若类型字段1011的值设为大约1(即,表示所描述的链路是一条P2P链路),则链路ID字段1001可以表示相邻路由器的路由器ID。若类型字段1011的值设为大约2(即,表示所描述的链路是一条传输网络链路),则链路ID字段1001可以表示指定路由器的IP地址。若类型字段1011的值设为大约3(即,表示所描述的链路是一条末梢网络链路),则链路ID字段1001可以表示末梢网络的IP号。若类型字段1011的值设为大约4(即,表示所描述的链路是一条虚拟链路),则链路ID字段1001可以表示相邻路由器的路由器ID。

[0056] 在本实施例中,当通过P2P链路连接到外部路由器(即,TTZ之外的路由器)时,TTZ

(边缘)路由器可以通过P2P链路发送一个LSA,该LSA包括描述P2P链路的路由器链路1000。链路ID字段1001可以表示外部路由器的路由器ID。另外的选择是(若接收LSA的外部路由器是另一个TTZ中的TTZ边缘路由器),源TTZ边缘路由器可以(在某些实例中)通过设置链路ID字段1001,从而表示外部路由器的TTZ ID(或者,TTZ的最大/最小路由器ID),可以从所有外部路由器的角度出发,将其看做TTZ的路由器ID。

[0057] 链路数据字段1003还依赖于类型字段1011的值。若类型字段1011设为大约1、2或4(即,表示所描述的链路是一条P2P链路、传输网络链路或虚拟链路),则链路数据字段1003可以表示路由器链路的IP地址。类型字段1011设为大约3(即,表示所描述的链路是一条末梢网络链路),则链路数据字段1003表示末梢网络的IP地址掩码。

[0058] NoTOS字段1013表示具体TOS度量字段1020到1040的数量,该字段包含在路由器链路1000中。如果没有给所描述的链路指配任何具体TOS度量,则可以将NoTOS字段1013设为大约0,从而表示路由器链路1000不包含任何具体TOS度量。否则,NoTOS字段1013可以包括一个整数,对应具体TOS度量字段1020到1040的数量,该字段包含在路由器链路1000中。度量字段1015表示所描述链路的成本。度量字段1015并不属于任一具体TOS度量字段1020到1040,如此一来,NoTOS字段1013内并不包括Metric 1015。

[0059] 每个具体TOS度量1020-1040都可以采用相同的格式。例如,具体TOS度量1040可以包括TOS字段1041、字段1043(例如,设为大约0)以及TOS度量字段1045。TOS字段1041可以表示所描述的具体TOS度量的业务类型(例如,IP TOS)。TOS字段1041可以设为大约0、2、4、8或16,从而(分别)表示,IP TOS是一种一般的业务、最小费用、最大可靠性、最大速率或最小延迟。TOS度量字段1045的值可以表示链路的成本,该链路用于由TOS字段1041字段标识的IP TOS。

[0060] 图11展示了本发明实施例提供的一种链路类型字段1100,可以包括TTZ(T)比特标识位1111以及类型-1字段1115。链路类型字段1100可以是对图10中链路类型字段1011的扩展。T比特标识位1111可以表示链路为内部链路(即,TTZ之内的链路)或外部链路(即,TTZ之外的链路),而类型-1字段1115可以设为大约1、2、3或4,从而(分别)表示,所描述的链路是和其他路由器的P2P连接、和传输链路的连接、和末梢网络的连接或虚拟链路。

[0061] 在本实施例中,T比特标识位1111设为大约1,从而表示该链路为TTZ的内部链路,或设为大约0,从而表示该链路是外部链路。因此(在此实施例中),若该链路为外部链路(即,TTZ之外的链路),则链路类型字段1100的值可以和链路类型字段1011的一样(例如,因为T比特标识位1111为0),或者,若该链路为内部链路(即,TTZ之内的链路),则该值与链路类型字段1011仅一个比特不同(例如,因为T比特标识位1111为1)。

[0062] 在某个可替代的实施例中,T比特标识位可以设为大约0,从而表示该链路是TTZ的内部链路,或设为大约1,从而表示该链路是外部链路。因此(在此实施例中),若该链路为内部链路,则链路类型字段1100的值可以和链路类型字段1011的一样(例如,因为T比特标识位1111为0),或者,若该链路为外部链路,则该值可与链路类型字段1011仅一个比特不同(例如,因为T比特标识位1111为1)。

[0063] 作为进一步的替换,可以给路由器链路添加一个关于TTZ ID的新字段,表示该链路属于哪个TTZ。新字段可以设为大约0,从而表示相应的链路是连接TTZ之外节点的链路,或设为非0值(例如,一个TTZ ID对应一个TTZ),从而表示该链路属于哪个TTZ。

[0064] 图12展示了本发明实施例提供的一种路由器链路1200,包含一个新字段TTZ ID 1250。路由器链路1200和上述路由器链路900相似。路由器链路1200可以包括链路ID字段1201、链路数据字段1203、类型字段1211、服务类型数量(NoTOS)字段1213、度量字段1215以及多个具体TOS链路度量字段1220到1240。路由器链路1200还可以包括TOS字段1241,字段1243和TOS度量字段1245。这些字段和上述路由器链路900中的相似或一致。路由器链路1200中的字段TTZ ID 1250是一个新字段,表示路由器链路1200中描述的链路属于哪个TTZ。

[0065] 在本实施例中,若新字段TTZ ID 1250的值为大约0,则表示路由器链路1200中描述的相应链路是连接TTZ外部路由器的链路。在另一个实施例中,若新字段的值非0或为TTZ ID,则表示路由器链路1200中描述的相应链路属于由新字段中的TTZ ID标识的TTZ。

[0066] 图13展示了在AS域中配置TTZ的方法1300。方法1300从框1310开始,其中,可以给每条TTZ的内部链路配置TTZ ID。在某些实施例中,TTZ边缘节点和/或TTZ内部节点可以在框1310之前,执行一个中间步骤,即,交换hello包从而了解各自的邻接。接下来,方法1300轮到框1320,即,执行TTZ内部链路状态扩散。具体地,TTZ内部链路状态扩散可以包括LSA的构建和洪泛,其中,LSA描述每个TTZ节点的邻接,包括所有TTZ节点以及所有相邻的外部节点。重要的是,TTZ边缘节点可以在其构建的LSA中描述其外部链路(即,一个TTZ边缘节点和一个外部节点之间的链路),从而允许其他TTZ边缘节点(以及TTZ内部节点)了解其外部链路的拓扑。通过TTZ内部链路状态扩散,每个TTZ节点都能了解TTZ的内部拓扑(即,TTZ节点之间的位置关系,包括内部链路)以及TTZ的邻接(即,TTZ边缘节点和其相邻的外部节点之间的位置关系,包括外部链路的标识)。

[0067] 接下来,方法1300轮到框1330,其中,每个TTZ边缘节点都可以独立构建一个一般的路由器LSA,描述了各外部链路,并将该路由器LSA扩散到与其相连接的相邻外部节点。该路由器LSA可以指定一个TTZ ID(或者,指配给TTZ节点的最大/最小路由器ID)作为TTZ的路由器ID,从而将TTZ作为一台单路由器,这台单路由器包括多个接口,对应多条外部链路。另一个选择是,可以指定TTZ中的一台路由器作为主TTZ路由器,可以是TTZ中路由器ID最小或最大的那台路由器。主TTZ路由器可以构建一个一般的路由器LSA,描述了各外部链路,还可以通过TTZ节点将该路由器LSA扩散到与其相邻的外部节点。

[0068] 接下来,方法1300轮到框1340,其中,TTZ边缘节点可以在整个TTZ中将接收自相邻外部节点的LSA洪泛,从而允许TTZ节点了解AS域的拓扑。在某些实施例中,TTZ边缘节点还可以将接收自相邻外部节点的LSA扩散到其他相邻的外部节点,可以和AS域中实现的IGP保持一致。虽然图13将TTZ边缘节点描述为接收LSA,根据该LSA的描述,AS域的拓扑发生在框1310到1360之后。本领域的普通技术人员都认可,在方法1300执行的过程中可以间歇性地接收此类LSA。

[0069] 接下来,方法1300轮到框1350,其中,可以在每个TTZ节点中创建描述TTZ内部拓扑以及AS域拓扑的LSDB,从而允许TTZ节点了解TTZ内部拓扑以及AS域的拓扑。最后,该方法轮到框1360,其中,各TTZ节点计算各自通往每个目的地的最短路径,该目的地包括在TTZ之内的目的地和在TTZ之外的目的地。

[0070] 以下实施例描述了网络300中的链路状态扩散。在本实施例中,TTZ 360中连接TTZ节点371的各链路可以是P2P链路。如上所说,TTZ节点371可以构建一个路由器LSA并将其洪

泛,该路由器LSA包含5条路由器链路,连接TTZ节点361、363、365、367以及373。第一路由器链路可以描述TTZ节点371和361之间的P2P连接,其中,第一路由器链路的链路ID字段1001指定了TTZ节点361的路由器ID。第二路由器链路可以描述TTZ节点371和363之间的P2P连接,其中,第二路由器链路的链路ID字段1001指定了TTZ节点363的路由器ID。第三路由器链路可以描述TTZ节点371和365之间的P2P连接,其中,第三路由器链路的链路ID字段1001指定了TTZ节点365的路由器ID。第四路由器链路可以描述TTZ节点371和367之间的P2P连接,其中,第四路由器链路的链路ID字段1001指定了TTZ节点367的路由器ID。第五路由器链路可以描述TTZ节点371和373之间的P2P连接,其中,第五路由器链路的链路ID字段1001指定了TTZ节点373的路由器ID。这五条链路中,每条链路都可以包括链路类型字段1011,其中T比特标识位1111设为大约1(表示该链路是内部链路,完全位于TTZ之内),还包括类型-1字段1115,设为大约1(表示所述链路是P2P链路)。

[0071] 在同一个或其他的实施例中,所有连接TTZ节点361的链路都可以是P2P链路。如上所说,TTZ节点361可以构建一个路由器LSA并将其洪泛,该路由器LSA包含4条路由器链路,其中之一连接外部节点315,其余3条连接TTZ节点363、365和371。在接收到LSA后,TTZ节点363、365和371可以将该路由器LSA在TTZ 360中洪泛扩散到其他节点。第一路由器链路可以描述TTZ节点361和315之间的P2P连接,其中,第一路由器链路的链路ID字段1001指定了节点315的路由器ID。第二路由器链路可以描述TTZ节点361和363之间的P2P连接,其中,第二路由器链路的链路ID字段1001指定了TTZ节点363的路由器ID。第三路由器链路可以描述TTZ节点361和365之间的P2P连接,其中,第三路由器链路的链路ID字段1001指定了TTZ节点365的路由器ID。第四路由器链路可以描述TTZ节点361和371之间的P2P连接,其中,第四路由器链路的链路ID字段1001指定了TTZ节点371的路由器ID。第一路由器链路可以包括T比特标识位1111,设为大约0(即,表示该链路在TTZ 360的外部),而第二、第三和第四路由器链路可以包括T比特标识位1111,设为大约1(即,表示该链路在TTZ 360的内部)。这四条链路都可以包括类型-1字段1115,设为大约1(即,表示所述链路为P2P链路)。

[0072] 在同一个或其他的实施例中,TTZ边缘节点361、363、367、365(在TTZ 360之内)和外部节点315、317、323、325、331、329(在TTZ 360之外)之间的所有链路都可以是P2P链路。在此实施例中,TTZ边缘节点361可以构建一个一般的路由器LSA,和由其他TTZ边缘节点发送给外部节点的一般路由器LSA基本或完全一致,但是和TTZ 360之内洪泛的路由器LSA不同。该一般路由器LSA可以包括链路状态ID字段907和发布路由器字段909,指定了路由器ID,相当于关联TTZ 360的TTZ ID(或者,指配给TTZ路由器361到373的最大/最小路由器ID),还包括6条路由器链路,描述了6条外部链路,从TTZ边缘节点361、363、365和367扩展到外部节点315、317、323、325、331和329(即,从TTZ边缘节点361、363、365和367向外扩展到外部节点315、317、323、325、331和329的链路),可以将其发送给相邻节点315。从外部节点315的角度出发,该一般路由器LSA可以是一个正常的路由器LSA,例如,和外部节点315从相邻的外部节点(例如,外部节点317)收到的路由器LSA相似。

[0073] 第一路由器链路可以描述TTZ边缘节点361和外部节点315之间的P2P连接,其中,第一路由器链路的链路ID字段1001指定了外部节点315的路由器ID。第二路由器链路可以描述TTZ边缘节点363和外部节点329之间的P2P连接,其中,第二路由器链路的链路ID字段1001指定了外部节点329的路由器ID。第三路由器链路可以描述TTZ边缘节点365和外部节

点317之间的P2P连接,其中,第三路由器链路的链路ID字段1001指定了外部节点317的路由器ID。第四路由器链路可以描述TTZ边缘节点365和外部节点323之间的P2P连接,其中,第四路由器链路的链路ID字段1001指定了外部节点323的路由器ID。第五路由器链路可以描述TTZ边缘节点367和外部节点325之间的P2P连接,其中,第五路由器链路的链路ID字段1001指定了外部节点325的路由器ID。第六路由器链路可以描述TTZ边缘节点367和外部节点331之间的P2P连接,其中,第六路由器链路的链路ID字段1001指定了外部节点331的路由器ID。这六条路由器链路都可以包括类型字段1011,设为大约1,表示该链路是P2P链路。

[0074] 下面的伪代码描述了路由器LSA RL中所有路由器链路的创建,该路由器LSA RL被发送到外部节点,其中,N是一个变量,用来计算从TTZ向外扩展的外部链路的数量。

```

N = 0;
For each router LSA in the TTZ
{
  For each router link in the router LSA
  {
    If the router link is an external link to a node outside of the TTZ
    {
      N = N + 1;
    }
    Add the router link into router LSA RL as a normal link;
  }
  Else If the router link is a stub link
  {
    N = N + 1;
    Add the router link into LSA RL and set cost to 0;
  }
}

```

[0076] 将路由器LSA RL中链路数量字段933的值设置为N;

[0077] 图11中的链路类型字段1100可以作为对图10中链路类型字段1011的扩展。在本实施例中,T比特标识位1111可以设为大约1,从而表示该链路为内部链路(即,在同一个TTZ之内的两个节点之间的链路),或设为大约0,从而表示该链路为外部链路(即,在TTZ之内的一个节点同TTZ之外的一个外部节点之间的链路)。在此实施例中,若T比特标识位1111设为大约0,以上“if条件”可以为真;若T比特标识位1111设为大约1,则为假。在其他实施例中,T比特标识位1111可以设为大约0,从而表示该链路为内部链路(即,在同一个TTZ中的两个节点之间的链路),或设为大约1,从而表示该链路为外部链路(即,在TTZ之内的一个节点同TTZ之外的一个外部节点之间的链路)。在此实施例中,若T比特标识位1111设为大约1,以上“if条件”可以为真;若T比特标识位1111设为大约0,则为假。无论哪种情况,若“if条件”为真,则将外部链路的路由器链路添加到路由器LSA RL作为一条一般的链路。

[0078] 关于TTZ外部路由器的路由表计算,和RFC 2328中描述的一致。关于TTZ的内部路由器,其流程和RFC 2328中描述的一致,但是扩展了链路的含义以及两端之间的关联。两端

之间的链路能够是TTZ链路,也能够是一般的链路。

[0079] 在TTZ之内的路由器侧,当检查和顶点V相关联的LSA时,针对该LSA中描述的每条链路,假设顶点W是链路的另一端,如果该链路是一般的链路,则能够将顶点W能看做顶点V的相邻顶点;如果该链路为TTZ的内部链路,且LSA由TTZ之内的路由器生成,则可以将顶点W看作顶点V的相邻顶点;如果该链路为TTZ的外部链路,且LSA由TTZ的边缘路由器生成,则能够将顶点W看作顶点V的相邻顶点,且顶点W是TTZ外部链路的另一端,位于TTZ之外。

[0080] 关于TTZ之内的路由器侧的路由表,可以通过以下方式来计算。将TTZ之外的链路(包括TTZ的外部链路)的成本/度量,看作一种特别的度量类型。该度量类型比TTZ之内的链路的度量高一个数量级。即,关于这一特别的类型,其中任一这种特别度量类型的成本,都大于任一TTZ内部路径的成本。通过构建TTZ之内路由器到各目的地的最短路径树,从而计算通往各目的地的路径,包括TTZ之内的目的地和TTZ之外的目的地。

[0081] 另一个选择是,TTZ之内的路由器的路由表计算,可分两个阶段完成。在第一个阶段,计算通往TTZ内所有目的地的路径。接下来,在第二个阶段,在把通往TTZ之内的所有目的地的路径的成本设为最低成本之后,例如大约1或0,从第一个阶段中创建的最短路径树中计算通往TTZ之外的所有目的地的路径。在另一个实施例中,在把通往所有TTZ边缘路由器的路径的成本设为最低成本之后,例如大约1或0,从第一个阶段中创建的最短路径树,可以计算通往TTZ之外的所有目的地的路径。

[0082] 虽然本申请描述的新方面涉及OSPF协议,但是本领域的普通技术人员都了解,这些新方面也能应用于其他IGP协议(例如,IS-IS协议等)。

[0083] 上述网络元件可以在任意一个通用的网络元件中实现,例如计算机或网络元件,拥有充足的处理能力、内存资源以及网络通过能力来处理其负责的相关工作。图14展示了本发明实施例提供的一种典型的、通用的网络元件1400,适用于实现一个或多个由本文公开的、元件的实施例。网络元件1400包括一个处理器1402(可以称作中央处理器或CPU),用来和内存设备进行通信,其中,内存设备包括第二存储器1404、只读存储器(read-only memory,ROM) 1406、随机存取存储器(random access memory, RAM) 1408、输入/输出(input/output, I/O)设备1410、以及网络连接设备1412。处理器1402可以作为一个或多个CPU芯片,也可以作为一个或多个专用集成电路(application specific integrated circuit, ASIC)的一部分。

[0084] 第二存储器1404一般包括一个或多个磁盘驱动器或可擦除可编程只读存储器(erasable programmable ROM, EPROM),用来处理数据的非易失性存储。第二存储器1404可以用来存储加载到RAM 1408的程序,这些程序被选中执行。ROM 1406可以用来存储指令,也可能是数据,即在程序执行过程中被读取的。ROM 1406属于非易失性存储设备,相对于第二存储器1404的大存储容量而言,ROM 1406的存储容量较小。RAM 1408用来存储易失性数据,也可能是指令。访问ROM 1406和RAM 1408的速度一般快于访问第二存储器1404。

[0085] 本文描述的特征/方法,至少有一些可以在网络装置或元件中实现,例如网络节点。例如,本文公开的特征/方法可以利用硬件、固件和/或在硬件上安装运行的软件来实现。网络装饰/元件或节点可以是任意通过网络传输数据帧的设备,网络包括,开关、路由器、桥、服务器等等。如图15所示,网络装置/元件1500可以包括多个如端口或单元,用来接收来自其他节点和逻辑电路的数据帧,从而确定将数据帧发送给哪个节点,还包括多个出

端口或单元,用来向其他节点传输数据帧。如和/或出端口可以包括电气的和/或光学的、传输和/或接收元件。

[0086] 本文公开了至少一个实施例,并且,由本领域普通技术人员对实施例和/或实施例的特征做出的变化、合并和/或修改都在本发明公开的范围。通过合并、融合和/或忽略实施例的某些特征从而生成的实施例也在本发明公开的范围。其中,对于数字的范围或限制都进行了明确的表述,这些范围或限制可以理解为包含反复的范围或相同数量的限制,都是明确表述的范围或限制即,从大约1到大约10包括2、3、4等,大于0.10包括0.11、0.12、0.13等)。例如,当公开了数字范围的一个下限R1和一个上限Ru,则在该范围内指定任意数字。尤其,本文公开了范围内的以下数字: $R=R1+k*(Ru-R1)$,其中,k是一个变量,从1%到100%不定,有1%的增长,即,k是1%、2%、3%、4%、5%,……,70%、71%、72%,……,95%、96%、97%、98%、99%、或100%。此外,具体还公开了由上述定义的两个R数字得出的数字范围。文中“(大)约”表示对接下来的数字进行±10%,除非有明确表示。权利要求的任一要素中所使用的词汇“可选地”,表示该要素是必须的,或者,不是必须的,二者都属于权利要求的范围内。使用广义词汇,例如“包括/包含/有……等”,可以理解为对狭义词汇,例如“由……组成”、“基本由……组成”以及“基本由……构成”的支持补充。相应地,保护范围并不局限于以上描述,而是由下文的权利要求进行限定,该范围包括所有对应权利要求主题的内容。下文进一步公开了权利要求,每一项权利要求的内容都结合在说明书中,且权利要求部分属于本发明公开的实施例。对于文中引用的讨论并非认可它是现有技术,尤其是对那些发布日期在本申请的优先日期之后的引用。本文中引用的所有专利、专利申请以及发表物的内容都是通过引用进行结合,从而给本文提供较佳的、程序上的、以及其他细节的补充。

[0087] 虽然本发明提供了一些实施例,但需要了解的是,其公开的系统和方法还是会以许多其他特定的形式实现,并不脱离本发明的精神和范围。本文的实施例只是用来进行说明,而并非限制,并不应将意图局限于文中所给出的细节。例如,在另一个系统中,可以将各种元素或器件进行合并或融合,也可以忽略或不实现某些特征。

[0088] 此外,关于各种实施例中分别或单独描述说明的技术、系统、子系统、方法,可以和其他系统、技术或方法进行合并或融合,且并不脱离本发明的范围。其他展示或讨论的项,例如耦合、直接耦合、互相通信,都可以是利用一些接口、设备或中间器件,通过电力、机械或其他手段,进行间接耦合或通信。本领域技术人员可以对这些实施例进行修改、替换以及变更,且并不脱离本发明的精神和范围。

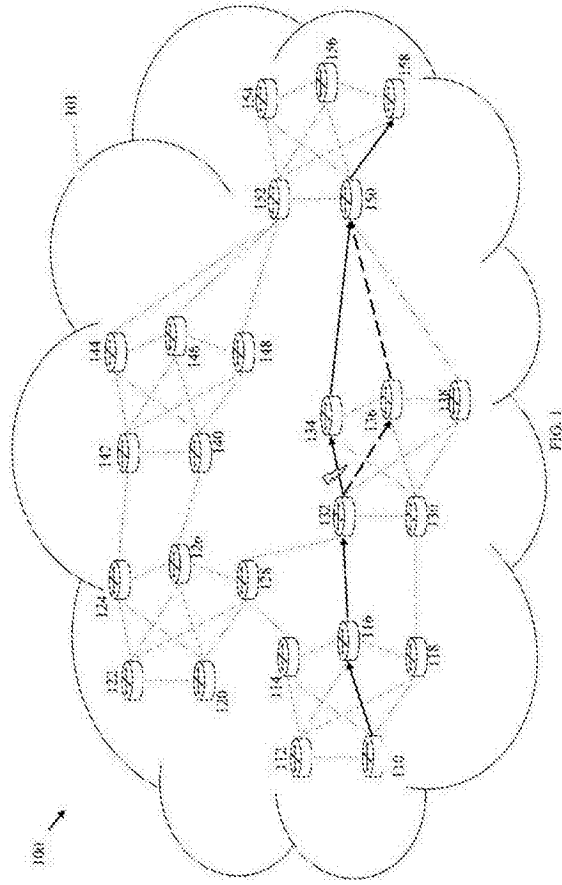


图1

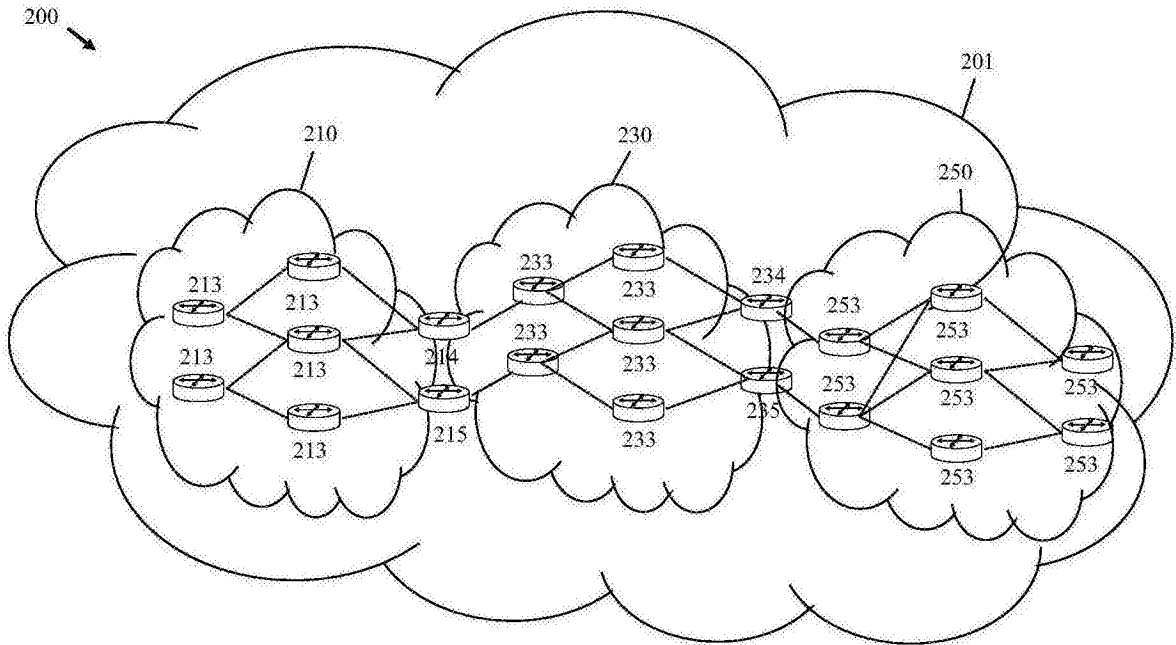
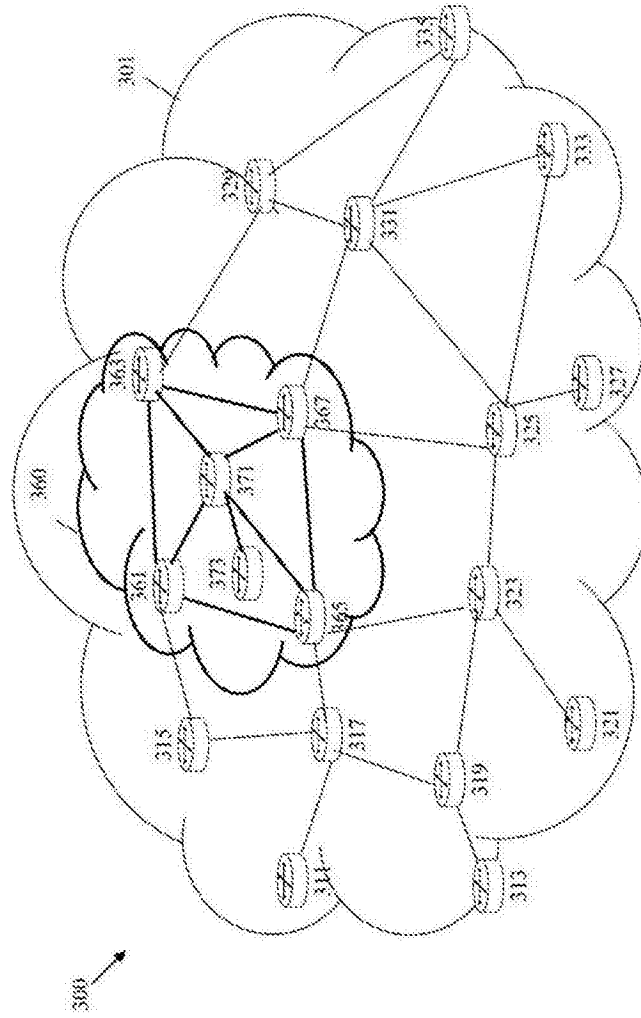


图2



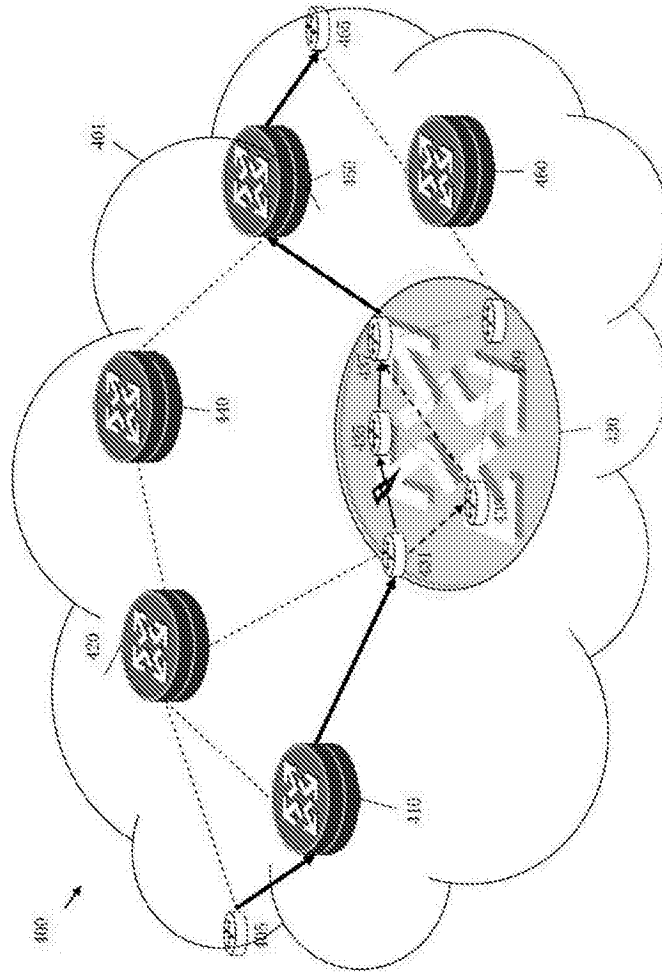


图4

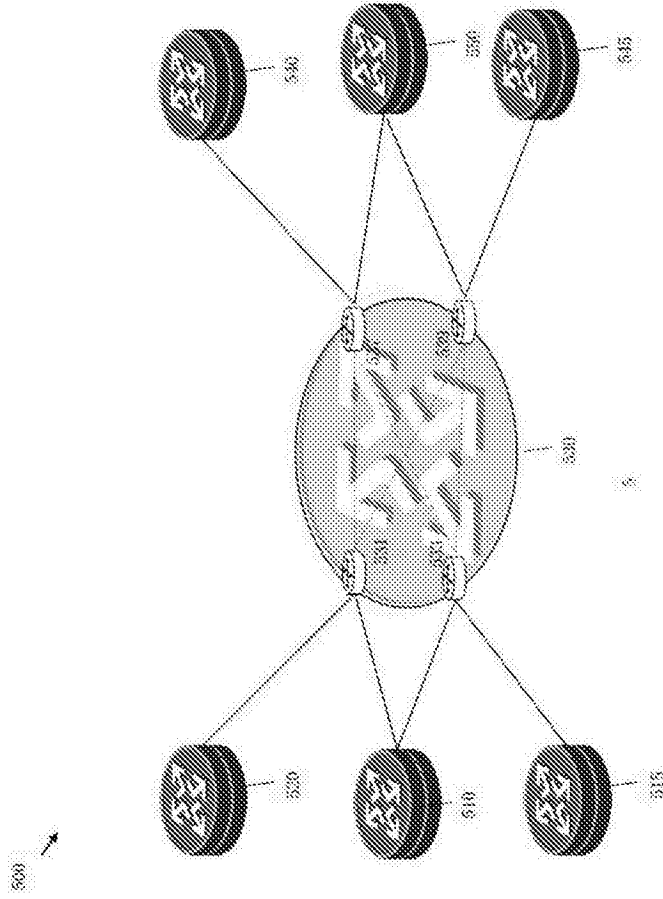


图5

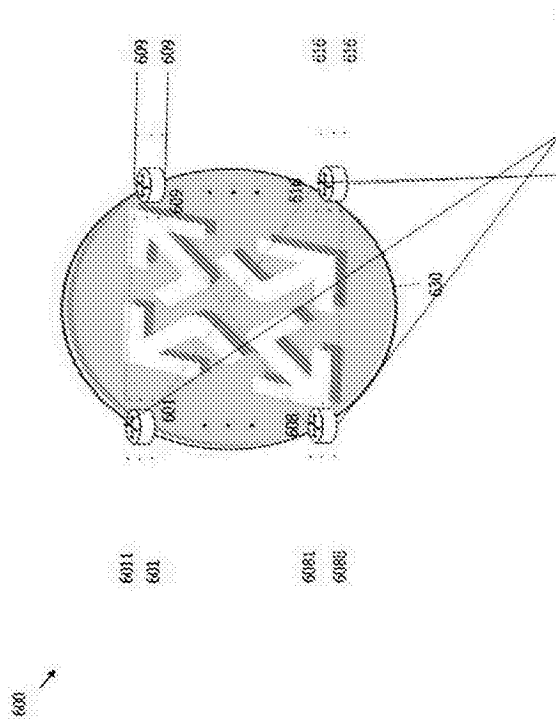


图6

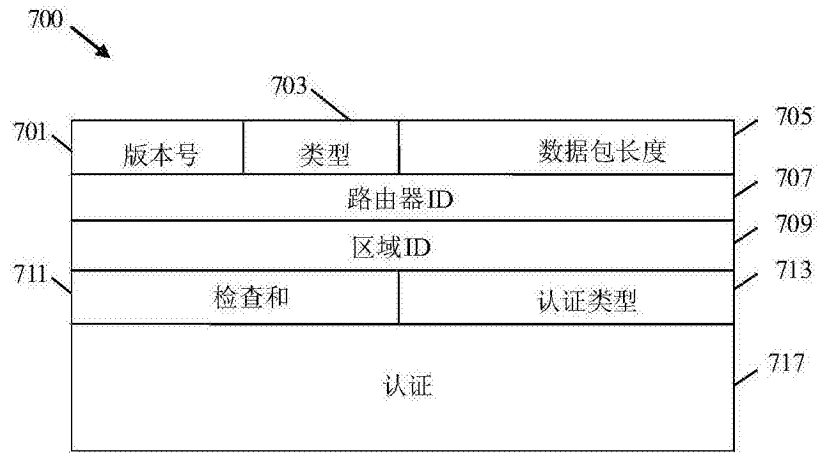


图7

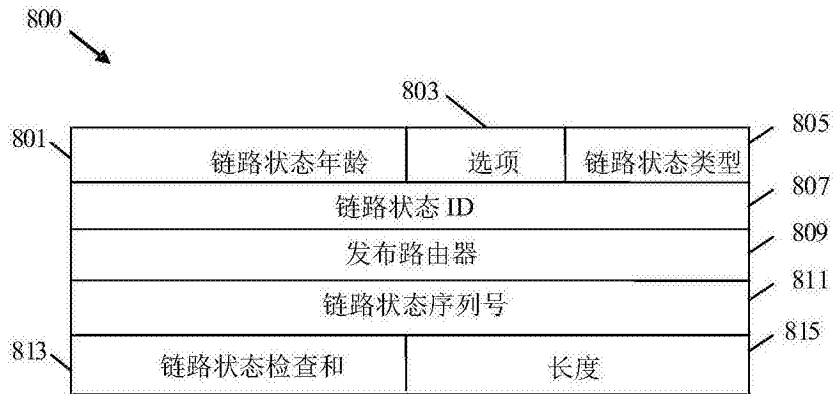


图8

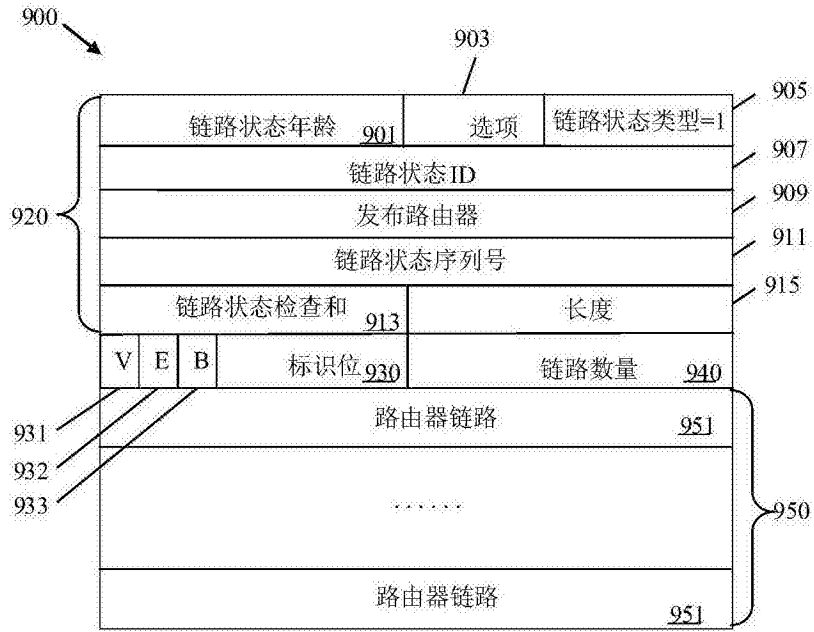


图9

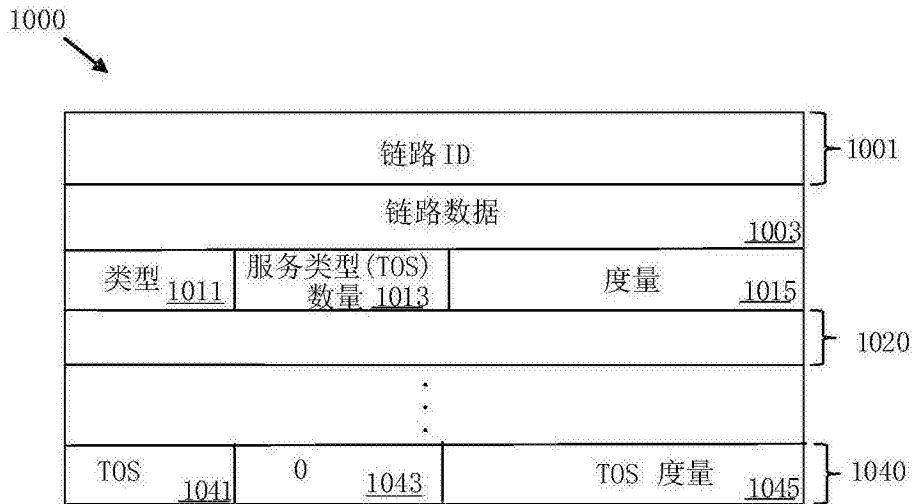


图10

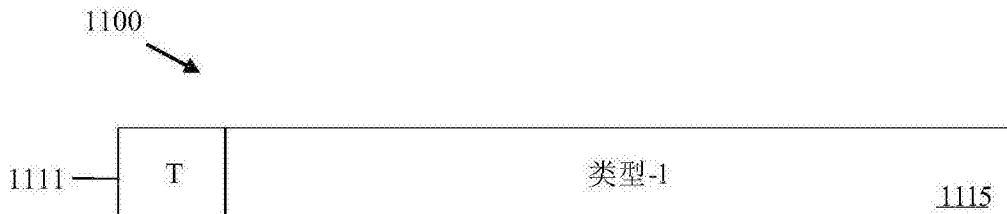


图11

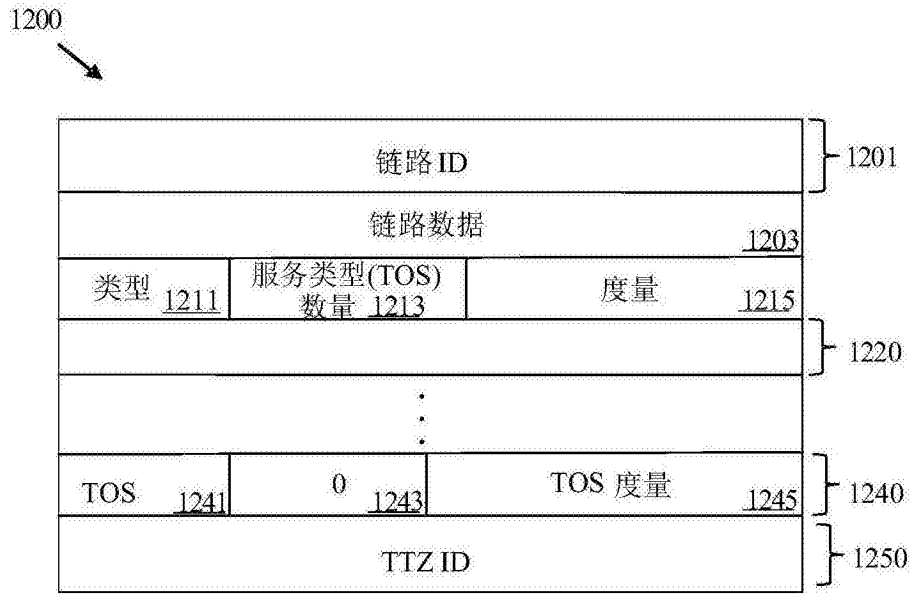


图12

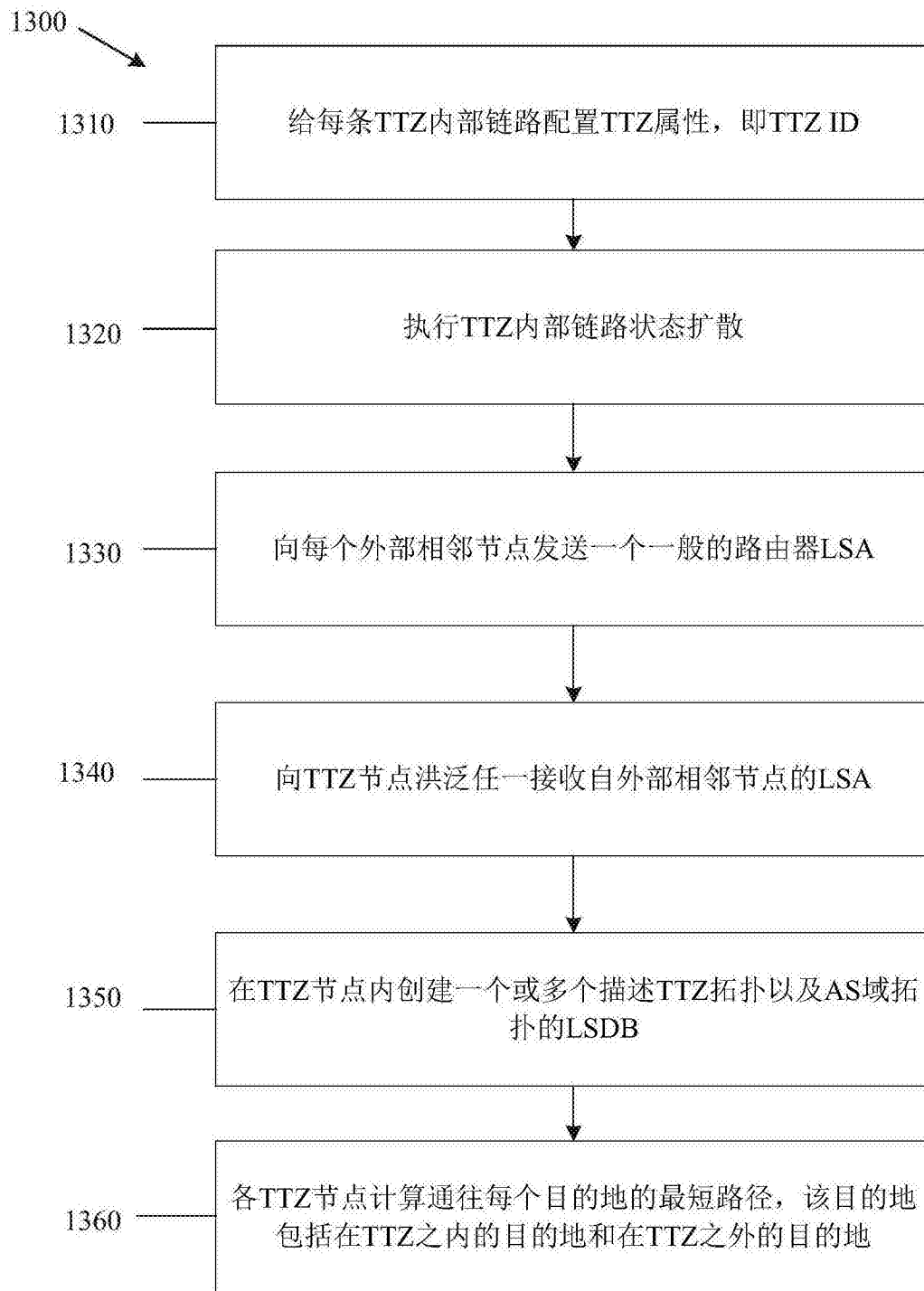


图13

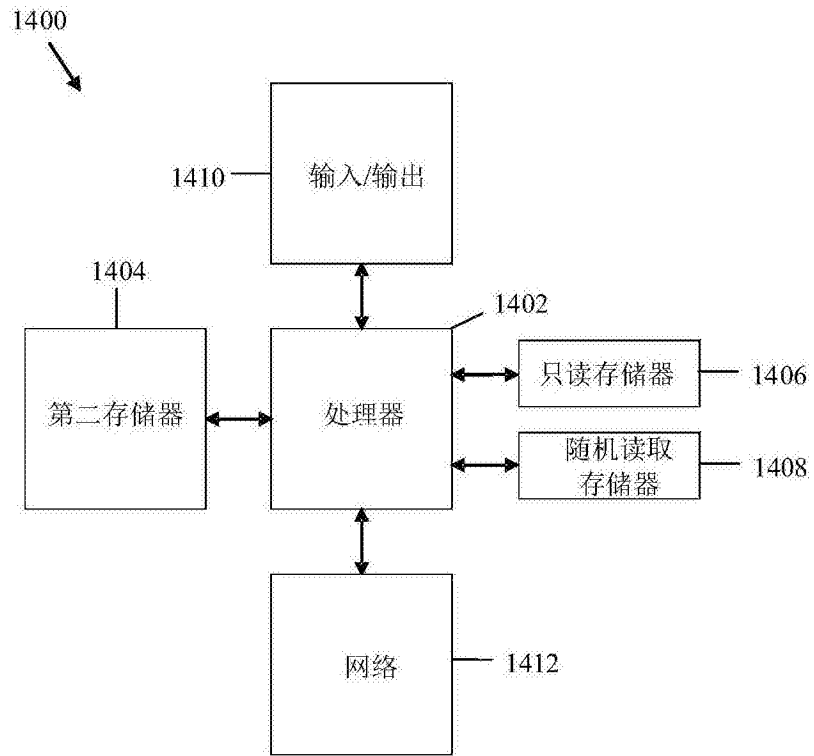


图14

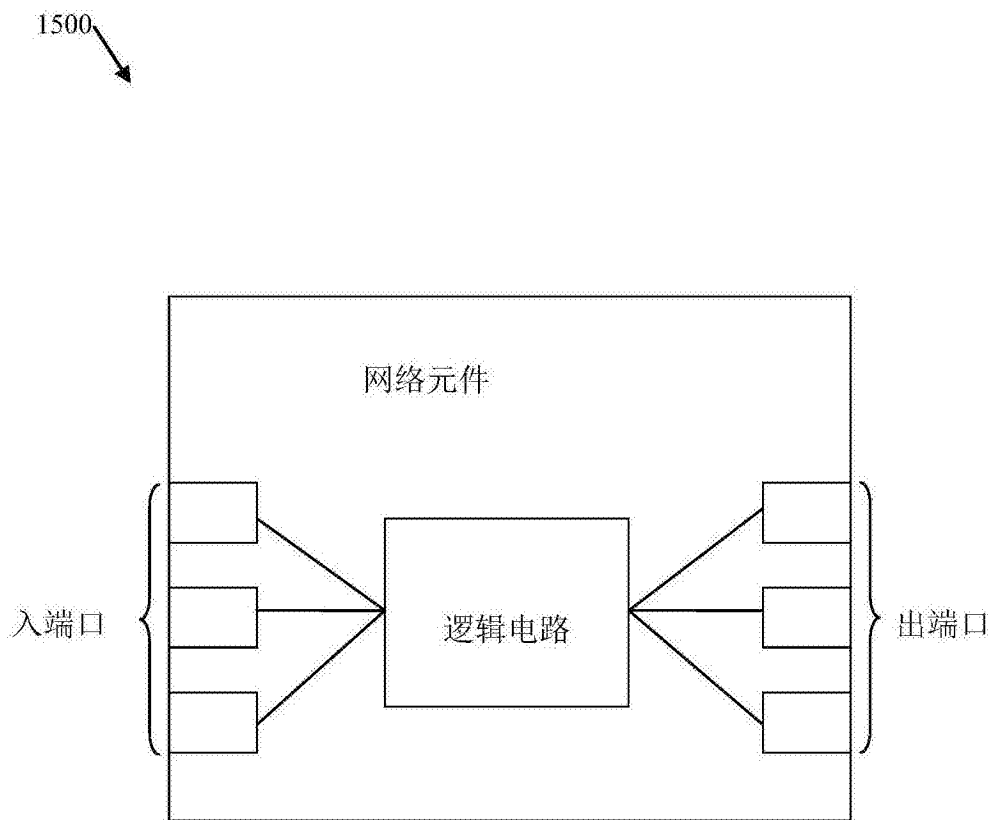


图15