



(12) 发明专利

(10) 授权公告号 CN 101904184 B

(45) 授权公告日 2013. 06. 12

(21) 申请号 200880120296. 3

(22) 申请日 2008. 10. 14

(30) 优先权数据

60/979, 438 2007. 10. 12 US

(85) PCT申请进入国家阶段日

2010. 06. 11

(86) PCT申请的申请数据

PCT/US2008/079803 2008. 10. 14

(87) PCT申请的公布数据

W02009/049307 EN 2009. 04. 16

(73) 专利权人 北方电讯网络有限公司

地址 加拿大魁北克省

(72) 发明人 D·莫罕 P·昂贝哈根 S·基萨拉

(74) 专利代理机构 中国专利代理(香港)有限公司

司 72001

代理人 王岳 蒋骏

(51) Int. Cl.

H04W 4/00(2006. 01)

(56) 对比文件

US 2007/0237156 A1, 2007. 10. 11, 说明书第 24-67 段.

CN 1592487 A, 2005. 03. 09, 全文.

US 2006/0285529 A1, 2006. 12. 21, 说明书第 78-387 段.

审查员 夏刊

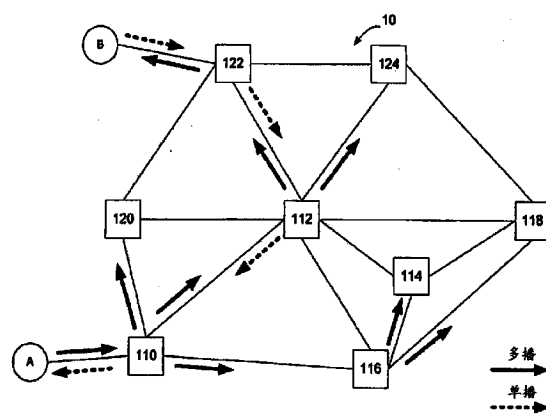
权利要求书1页 说明书12页 附图13页

(54) 发明名称

链路状态控制的以太网中提供的自动 MEP

(57) 摘要

在链路状态协议控制的以太网中自动配置以太网 OAM MEP。在链路状态协议控制的以太网中运行的节点接收包含 TLV 的链路状态 PDU(LSP), 所述 TLV 具有与链路状态协议控制的以太网中的第二节点的以太网 MAC 节点 ID 相关的 MEP, 其中, 第一和第二节点之间的路径包括多个链路。节点更新转发表以指示第二节点的以太网 MAC 节点 ID 与 MEP ID 之间的关联。通过对 Sys-ID 进行散列在链路状态协议控制的以太网中产生以太网 OAM 维护端点以产生 MEP; 将 MEP 存储在 TLV 中; 并在 LSP 中通过链路状态协议控制的以太网来转发 TLV。



CN 101904184 B

1. 一种用于在链路状态协议控制的以太网络上运行的节点中安置转发状态的方法,该方法包括步骤:

由在链路状态协议控制的以太网中运行的第一节点接收包含类型长度值 TLV 的链路状态协议数据单元 LSP,所述 TLV 具有与链路状态协议控制的以太网中的第二节点的以太网介质接入控制 MAC 节点标识符 ID 相关的维护端点标识符 MEP ID,其中,第一和第二节点之间的路径包括链路状态协议控制的以太网中的多个链路;以及

更新转发表以指示链路状态协议控制的以太网中的第二节点的以太网 MAC 节点 ID 与 MEP ID 之间的关联。

2. 权利要求 1 的方法,还包括步骤:

接收操作、管理和维护 OAM 命令;

查询转发表以便从命令中的以太网 MAC 节点 ID 中解析对应于 MEP ID 的 MEP;

在链路状态协议控制的以太网络上朝着 MEP 转发命令。

3. 一种用于在链路状态协议控制的以太网中产生以太网操作、管理和维护 OAM 维护端点 MEP 的方法,包括步骤:

在链路状态协议控制的以太网中的节点中对 Sys-ID 进行散列以产生维护端点标识符 MEP-ID;

将 MEP-ID 存储在类型长度值 TLV 中;

在链路状态协议数据单元 LSP 中通过链路状态协议控制的以太网转发 TLV。

4. 一种用于在链路状态协议控制的以太网络上运行的节点中安置转发状态的设备,该设备包括:

用于由在链路状态协议控制的以太网中运行的第一节点接收包含类型长度值 TLV 的链路状态协议数据单元 LSP 的装置,所述 TLV 具有与链路状态协议控制的以太网中的第二节点的以太网介质接入控制 MAC 节点标识符 ID 相关的维护端点标识符 MEP ID,其中,第一和第二节点之间的路径包括链路状态协议控制的以太网中的多个链路;以及

用于更新转发表以指示链路状态协议控制的以太网中的第二节点的以太网 MAC 节点 ID 与 MEP ID 之间的关联的装置。

5. 权利要求 4 的设备,还包括:

用于接收操作、管理和维护 OAM 命令的装置;

用于查询转发表以便从命令中的以太网 MAC 节点 ID 中解析对应于 MEP ID 的 MEP 的装置;

用于在链路状态协议控制的以太网络上朝着 MEP 转发命令的装置。

6. 一种用于在链路状态协议控制的以太网中产生以太网操作、管理和维护 OAM 维护端点的设备,包括:

用于在链路状态协议控制的以太网中的节点中对 Sys-ID 进行散列以产生维护端点标识符 MEP-ID 的装置;

用于将 MEP-ID 存储在类型长度值 TLV 中的装置;

用于在链路状态协议数据单元 LSP 中通过链路状态协议控制的以太网转发 TLV 的装置。

## 链路状态控制的以太网中提供的自动 MEP

[0001] 相关申请的交叉引用

[0002] 本申请要求题为 PLSB AND IP SHORTCUTS OAM 的 2007 年 10 月 12 日提交的美国临时专利申请序号 60/979, 438 的优先权, 其内容通过引用结合到本文中。

[0003] 本申请涉及题为 IP NETWORK AND PERFORMANCE MONITORING USING ETHERNET OAM 的 2008 年 10 月 12 日提交的共同待决美国专利申请序号 12/249, 941 ; 题为 AUTOMATIC MEPPROVISIONING IN A LINK STATE CONTROLLED ETHERNET NETWORK 的 2008 年 10 月 12 日提交的美国专利申请序号 12/249, 944 ; 和题为 CONTINUITY CHECK MANAGEMENT IN A LINK STATE CONTROLLED ETHERNET NETWORK 的 2008 年 10 月 12 日提交的美国专利申请序号 12/249, 946, 其为 Nortel Networks Limited 共同所有, 并且涉及也为 Nortel Networks Limited 共同所有的题为 IP FORWARDING ACROSS A LINK STATE PROTOCOL CONTROLLED ETHERNET NETWORK、2008 年 5 月 5 日提交的共同待决美国专利申请序号 12/151, 684。

### 技术领域

[0004] 本发明涉及链路状态协议控制的以太网 (link state protocol controlled Ethernet network), 更具体而言, 涉及链路状态协议控制的以太网中的操作、管理、和维护 (OAM)。

### 背景技术

[0005] 数据通信网络可以包括各种计算机、服务器、节点、路由器、交换机、桥接器 (bridge)、集线器、代理、和相互耦合并被配置为相互传递数据的其它网络设备。这些设备在本文中称为“网络元件”。通过利用网络元件之间的一个或多个通信链路, 在网络元件之间, 通过传递诸如网际协议分组、以太网帧、数据单元、段、或数据的位 / 字节的其它逻辑关联等协议数据单元, 经由数据通信网络来传送数据。特定的协议数据单元可以被多个网络元件处理并在其通过网络在其源与其目的地之间行进时穿过多个通信链路。

[0006] 通信网络上的各种网络元件使用在本文中称为协议的预定义规则集来相互通信。使用不同的协议来支配通信的不同方面, 诸如应如何形成用于在网络元件之间传输的信号, 协议数据单元应看起来像的各种方面、应如何由网络元件处理协议数据单元或路由协议数据单元通过网络, 并且应如何在网络元件之间交换诸如路由信息的信息。

[0007] 以太网是已被电气和电子工程师协会 (IEEE) 定义为以太网架构中的标准 802.1 的众所周知的联网协议, 连接到网络的设备为在任何给定时间使用共享电信路径的能力而竞争。在使用多个桥接器或节点来互连网络段的情况下, 常常存在到同一目的地的多个潜在路径。此架构的益处在于其提供桥接器之间的路径冗余并允许以附加链路的形式向网络添加能力。然而, 为了防止形成环路, 常常使用生成树来限制业务在网络上广播或泛洪 (flood) 的方式。生成树的特性是在网络中的任何一对目的地之间仅存在一个路径, 因此, 可以通过观察分组来自哪里而“获悉”与给定生成树相关的连接性。然而, 生成树本身是限制性的, 常常导致在生成树上的链路的过度利用和不是生成树的一部分的链路的未利

用。

[0008] 为了克服实现生成树的以太网中固有的某些局限性,在题为“Provider Link State Bridging”的2006年10月2日提交的申请No. 11/537,775中公开了一种链路状态协议控制的以太网,该申请的内容通过引用结合到本文中。如本申请中更详细地描述的那样,不是通过使用与透明桥接组合的生成树协议(STP)算法在每个节点处利用获悉的网络视图,而是在链路状态协议控制的以太网中,形成网状网络的桥接器交换链路状态广告(advertisement)以使得每个节点能够具有网络拓扑结构的同步视图。这经由链路状态路由系统的被很好理解的机制来实现。网络中的桥接器具有网络拓扑结构的同步视图,具有对必需的单播和多播连接性的知识,能够计算网络中的任何一对桥接器之间的最短路径连接性,并且能够根据所计算的视图个别地填充其转发信息库(FIB)。

[0009] 当所有节点已计算了其在同步视图中的角色并填充其FIB时,网络将具有从对等桥接器集(无论由于什么原因要求到该桥接器的通信的那些)到任何给定桥接器的无环路单播树;以及从任何给定桥接器到在该桥接器处被托管的每个服务实例的同一对等桥接器集或其子集的一致(congruent)和无环路点到多点(p2mp)多播树。结果是给定桥接器对之间的路径并不约束为经过生成树的根桥接器,并且总体结果可以更好地利用网络的连接性的宽度。基本上,每个桥接器是限定到该桥接器的单播连接性和从该桥接器起的多播连接性的一个或多个树的根。

[0010] 当客户业务进入提供商网络时,客户MAC地址(C-MAC DA)被解析成提供商MAC地址(B-MAC DA),以便提供商可以使用提供商MAC地址空间在提供商网络上转发业务。另外,提供商网络上的网络元件被配置为基于虚拟LAN ID(VID)来转发业务,以便可以经由网络通过不同的路径转发被寻址到同一目的地地址但具有不同VID的不同帧。在操作中,链路状态协议控制的以太网可以使一个VID范围与最短路径转发关联,使得可以使用来自该范围的VID来转发单播和多播业务,并且可以在除该最短路径之外的路径上跨越网络产生业务工程路径(traffic engineering path),并使用第二VID范围来转发该业务工程路径。

[0011] 为了向链路状态协议控制的以太网添加真实运营级特征,期望有某些操作、行政、和管理(OAM)特征。当前在IEEE标准802.1ag“Connectivity Fault Management(连接性故障管理)”中定义的以太网OAM定义供在以太网中使用的一组连接性故障管理协议。其包括:连续性检查、链路跟踪、和环回(loopback)协议。802.1ag标准已被扩展成包括性能监视度和消息。此标准在ITU-T SG 13, Y.1731-“Requirements for OAM in Ethernet Networks(对以太网中的OAM的要求)”处有所反映。但是,在这些标准中描述的机制由于标准与链路状态协议以太网之间的寻址及VLAN语义和使用的某些差异而并不可直接适用于链路状态协议以太网。期望的是将用于故障识别、隔离、故障检修、和性能监视目的的OAM特征结合到链路状态协议以太网中。

## 发明内容

[0012] 依照本发明,提供了一种用于在链路状态协议控制的以太网络上运行的节点中安装(install)转发状态(forwarding state)的方法,该方法包括步骤:由在链路状态协议控制的以太网中运行的第一节点接收包含TLV的链路状态PDU(LSP),所述TLV具有与链路状态协议控制的以太网中的第二节点的以太网MAC节点ID相关的MEP,其中,第一和

第二节点之间的路径包括链路状态协议控制的以太网中的多个链路；并且更新转发表以指示链路状态协议控制的以太网中的第二节点的以太网 MAC 节点 ID 与 MEP ID 之间的关联。

[0013] 该方法还包括步骤：接收 OAM 命令；查询转发表以便从命令中的以太网 MAC 节点 ID 中解析 MEP；以及在链路状态协议控制的以太网络上朝着 MEP 点转发命令。

[0014] 依照本发明的另一方面，一种用于在链路状态协议控制的以太网中产生以太网 OAM 维护端点的方法，包括步骤：在链路状态协议控制的以太网中的节点中对 Sys-ID 进行散列以产生 MEP-ID；将 MEP-ID 存储在 TLV 中；以及在 LSP 中通过链路状态协议控制的以太网转发 TLV。

[0015] 本发明还以一种程序产品的方式呈现，所述程序产品包括具有在其中体现的用于存储数据的计算机程序的计算机可读介质。所述计算机程序包括用于在链路状态协议控制的以太网络上运行的节点中安置转发状态的逻辑。用于安置转发状态的逻辑包括：用于由在链路状态协议控制的以太网中运行的第一节点接收包含 TLV 的链路状态 PDU (LSP) 的逻辑，所述 TLV 具有与链路状态协议控制的以太网中的第二节点的以太网 MAC 节点 ID 相关的 MEP，其中，第一和第二节点之间的路径包括链路状态协议控制的以太网中的多个链路；以及用于更新转发表以指示链路状态协议控制的以太网中的第二节点的以太网 MAC 节点 ID 与 MEP ID 之间的关联的逻辑。该计算机程序还包括：用于接收 OAM 命令的逻辑；用于查询转发表以便从命令中的以太网 MAC 节点 ID 中解析 MEP 的逻辑；以及用于在链路状态协议控制的以太网络上朝着 MEP 点转发命令的逻辑。根据本发明，还提供一种程序产品，包括具有在其中体现的用于存储数据的计算机程序的计算机可读介质，该计算机程序在链路状态协议控制的以太网中产生以太网 OAM 维护端点。该计算机程序包括：用于在链路状态协议控制的以太网中的节点中对 Sys-ID 进行散列以产生 MEP-ID 的逻辑；用于将 MEP-ID 存储在 TLV 中的逻辑；以及用于在 LSP 中通过链路状态协议控制的以太网转发 TLV 的逻辑。

## 附图说明

[0016] 在随附权利要求中详细地指出本发明的方面。本发明在以下附图中以举例方式示出，在附图中相同的参考标号指示类似元件。以下附图仅出于说明的目的公开了本发明的各种实施例且并不意图限制本发明的范围。出于明了的目的，并不是在每个图中都对每个组件加以标注。在附图中：

[0017] 图 1 是可以用来实现链路状态协议控制的以太网络的网状网络的功能方框图；

[0018] 图 2 是被配置为在链路状态协议控制的以太网中使用的网络元件 12 的一种实施方式的示意性表示；

[0019] 图 3 是配置的链路状态协议控制的以太网络的示意性表示，其中，诸如 IS-IS 的链路状态协议已执行其发现阶段以使用每个桥接器的 Sys-ID 以无环路结构将桥接器互连，并且随后，多播连接性在作为 ISID 成员的所有节点之间创建 EVPN。

[0020] 图 4 是类似于图 3 的配置的链路状态协议控制的以太网络的示意性表示，其中，现在示出多个服务被映射为离开发现阶段树的叶子 (leaf)。

[0021] 图 5 是由 802.1ag 标准定义的以太网 OAM 维护域的方框图。

- [0022] 图 6 是 802.1ag OAM 分组的方框图。
- [0023] 图 7 是根据本发明的实施例的链路状态协议控制的以太网络的节点处的基础设施级 (infrastructure level) OAM 分组的处理的流程图；
- [0024] 图 8 是根据本发明的实施例在链路状态协议控制的以太网络的节点处执行的基础设施级连续性检查过程的流程图；
- [0025] 图 9 是根据本发明的实施例的链路状态协议控制的以太网络的节点处的服务级 OAM 分组的处理的流程图；
- [0026] 图 10 是根据本发明的实施例在链路状态协议控制的以太网络的节点处执行的服务级连续性检查过程的流程图；
- [0027] 图 11 是根据本发明的实施例的由链路状态协议控制的以太网中的节点进行的 MEP 创建和分发的流程图；
- [0028] 图 12 是根据本发明的实施例的链路状态协议控制的以太网中的节点中的 MEP 接收和转发表更新的流程图；
- [0029] 图 13 是根据本发明的实施例的使用 MEP 查找来从节点 A 向节点 B 发送 OAM 命令的过程的 MEP 的流程图；
- [0030] 图 14 是在链路状态协议控制的以太网上的两个 IP 节点之间执行的 IP “Ping” 命令的示意图。
- [0031] 图 15 是根据本发明的实施例的链路状态协议控制的以太网中的节点处的 IP 级 “Ping” 命令的处理的流程图；
- [0032] 图 16 是根据本发明的实施例的链路状态协议控制的以太网中的节点处的 IP 级 “Traceroute” 命令的处理的流程图；
- [0033] 图 17 是网络的方框图, 其中, 提供商被耦合到具有 IP 电话的客户驻地, 并且所有通信通过链路状态协议控制的以太网而发生。VOIP 网络的性能监视使用依照本发明的以太网 OAM 命令在 IP 层级处发生；
- [0034] 图 18 是根据本发明的实施例的链路状态协议控制的以太网中的节点处的 IP 级性能监视命令的处理的流程图。

### 具体实施方式

[0035] 链路状态协议控制的以太网提供以太网桥接的连接性的等价物, 但经由网络元件转发信息库 (FIB) 的配置而不是通过泛洪和获悉来实现这一点。使用链路状态协议来控制以太网使得能够通过以无环路最短路径转发提供网络容量的更高效利用来将以太网从 LAN 空间缩放到 WAN 或提供商网络空间。不是通过使用与透明桥接组合的生成树协议 (STP) 算法在每个节点处利用获悉的网络视图, 而是在链路状态协议控制的以太网中, 形成网状网络的桥接器交换链路状态广告以使得每个节点能够具有网络拓扑结构的同步视图。这经由链路状态路由系统来的使用来实现。网络中的桥接器具有网络拓扑结构的同步视图, 具有对必需的单播和多播连接性的知识, 能够计算网络中的任何一对桥接器之间的最短路径连接性, 并且能够根据所计算的视图个别地填充其转发信息库 (FIB)。当所有节点已计算其在同步视图中的角色并填充其 FIB 时, 网络将具有从对等桥接器集到任何给定桥接器的无环路单播树; 以及从任何给定桥接器到同一对等桥接器集的一致和无环路点

到多点 (p2mp) 多播树。结果是给定桥接器对之间的路径不限于经过生成树的根桥接器,并且总体结果可以更好地利用网格的连接性的宽度。

[0036] 链路状态协议控制的以太网通常使用对称链路度量,使得任何两个桥接器之间的连接性沿双向遵循同一路径,并对单播和多播连接性使用共同的度量,使得在被多播的分组与被单播的分组之间存在转发的一致性 (congruency)。

[0037] 可以使用 MAC 结构来构造 (略微修改的) 桥接器集之间的最短路径无环路连接性 (出于单播和多播两个目的) 以便向可以使用透明 LAN 服务的 C-MAC (客户 MAC) 层或其它层网络提供透明 LAN 服务。这要求在网络内的链路状态路由协议代替用于相关 VLAN (多个) 的生成树协议的操作和路由系统广告上的 MAC 信息的捎带 (piggyback)。

[0038] 图 1 是链路状态协议控制的以太网的一部分的示例的示意性表示。根据共享网络拓扑结构,每个节点使用最短路径算法来计算到网络中的其它提供商骨干桥接器 (provider backbone bridge, PHB) 或节点的最佳最短路径。在网络上应用最短路径算法的结果和在桥接器中 FIB 的相应的填充提供从每个桥接器通过网格到达网络的成员桥接器的唯一树。

[0039] 与桥接器 (单播和多播) 相关的 MAC 地址对于链路状态协议控制的以太网而言是全局的,并用于基于目的地的转发。这意味着其可以简单地在路由系统广告中泛洪,并且在路由系统的局部收敛 (convergence) 时,可以按照路由系统的指示在本地桥接器转发数据库 (或 FIB) 中实例化 (instantiate)。这样,可以将第 2 层连接性的分布式计算应用于以太网桥接器而不需要不同的信令系统以使连接性与拓扑结构相关联。以其最简单形式,当桥接器已计算到其在两个给定桥接器节点之间的最短路径上时,其简单地将与那些桥接器相关的 MAC 地址安置在 FIB 中,即指向每个感兴趣的桥接器的单播 MAC 地址和从感兴趣的桥接器指向的多播 MAC 地址 (多个)。

[0040] 应理解的是虽然已描述了每个桥接器一个单播 MAC 地址,但绝不排除更细粒度的使用,并且单播 MAC 地址可以指示线路卡、虚拟交换机实例 (VST) 或 UNI 端口。这对于简化目的地桥接器处的流的解复用而言是期望的。

[0041] 在网络中需要环路抑制以保持不稳定时间段 (在拓扑结构变化、由到网络中的所有桥接器的路由系统进行的拓扑结构变化的广告、新拓扑结构的公共视图上的再收敛和转发信息的相应更新之间的时间段) 期间的连接性 (即使以可能劣化的形式)。分布式系统中的不稳定性常常意味着网络的总体视图至少暂时将不是同步的。在网络元件不具有网络的同步视图的情况下,可能形成短暂环路。如本申请中更详细地描述的那样,PLSB 网络可以使用反向路径转发检查来使环路减到最少。可以通过促使诸如以太网桥接器的网络元件通过将包含在分组和分组到达的段中的源 MAC 地址与为与转发数据库中的目的地相同的 MAC 地址配置的值相比较来检查分组而执行 RPF 检查。如果用于源 MAC 地址的已获悉段将修改静态条目,或者不存在静态条目,则该分组被丢弃。在特定实例中,可以根据需要可选地禁用 RPF 检查。

[0042] 链路状态协议控制的以太网可以支持服务实例,其中,任何服务实例仅要求到端口的子集和因此到网络中的桥接器的连接性。可以用来识别与特定服务实例相关的分组的标识符的一个示例是在 IEEE802.1ah 中定义的扩展服务 ID 字段 (I-SID)。发现自己在两个桥接器之间的最短路径上的桥接器安置与每个桥接器相关的单播 MAC 地址 (多个) 和用

于为两个桥接器所共用的所有 I-SID 的多播 MAC 地址。其结果是给定边缘桥接器将具有到所有对等桥接器的单播连接性和为感兴趣的每个 I-SID 识别的团体 (community) 所独有的多播连接性。这将采取以下形式:作为到每个对等体的多点到点 (mp2p) 单播树上的叶子并作为到每个感兴趣的团体的对等节点集的 (S, G) 点到多点 (p2mp) 多播树的根,其中, S 是源的地址而 G 是多播组地址。

[0043] 此外,如整体地通过引用结合到本文中的题为 IP FORWARDING ACROSS A LINK STATE PROTOCOL CONTROLLED ETHERNET NETWORK 的、2008 年 5 月 5 日提交的共同待决美国专利申请序号 12/151,684 所述,链路状态协议控制的以太网能够支持本地 IP。因此,当节点获悉了 IP 地址时,其将该 IP 地址插入其链路状态广告中以对该 IP 地址到网络上的其它节点的可达性进行广告。每个节点将把此 IP 地址添加到其链路状态数据库。如果分组到达入口节点,则该入口节点将读取 IP 地址,确定链路状态协议控制的以太网上的哪个节点知道 IP 地址,并构造 MAC 报头以便将分组转发到正确的节点。MAC 报头的 DA/VID 是对 IP 地址进行广告的节点的节点 MAC (node MAC)。可以实现单播和多播 IP 转发。

[0044] 图 2 是被配置为在链路状态协议控制的以太网中使用的网络元件 12 的可能实施方式的示意性表示。网络元件 12 包括路由系统模块 80,其被配置为使用链路状态路由协议关于网络拓扑结构与网络 10 中的对等体 12 交换包含路由及其它信息的控制消息。路由系统 80 接收到的信息可以被存储在链路状态数据库 90 或以另一种方式存储。如前所述,信息的交换允许网络上的节点生成网络拓扑结构的同步视图,其随后允许路由系统模块 80 计算到网络上的其它节点的最短路径。路由系统 80 计算的最短路径将被编程到 FIB 82 中,FIB 82 被填充有用于基于所计算的最短路径、多播树、业务工程路径条目、并基于其它条目指引业务通过网络的适当条目。

[0045] 路由系统 80 可以交换包含网络层可达性信息的路由更新。网络上的节点所知道的网络层地址将被存储在网络元件 12 上的链路状态数据库 90 中以允许入口节点在网络层分组到达时选择链路状态协议控制的以太网上的正确出口节点。网络层地址的知识还可以通过促使节点安置对同一 IP 多播感兴趣的成对节点之间的转发状态来允许在网络上实现多播转发状态以允许网络上的节点处理网络层多播。

[0046] 网络元件 12 还可以包括诸如反向路径转发检查 (RPFC) 模块 84 的一个或多个其它模块,其可以用来处理输入帧并在 FIB 82 中执行查找以确定在其上接收到帧的端口是否与在用于特定源 MAC 的 FIB 82 中识别的端口一致。在输入端口不与在 FIB 中识别的正确端口一致的情况下,RPFC 模块可以促使消息被删除。

[0047] 如果帧通过 RPFC 84 模块,目的地查找 86 模块根据 FIB 82 来确定应当在其上转发帧的一个或多个端口。如果 FIB 不具有用于 DA/VID 的条目,则该帧被丢弃。

[0048] 还应理解的是所述的模块仅仅是出于说明的目的,并且如本领域的技术人员所理解的那样,可以通过在节点的模块之间组合或分布功能来实现所述的模块。

[0049] 参照图 3,示出了链路状态协议控制的以太网 300,其中,诸如 IS-IS 的链路状态控制协议已执行其发现阶段以使用每个桥接器的 Sys-ID (又叫做节点 MAC 304) 以无环路结构将桥接器 302a ~ h 互连。一旦构成例如 ISID 23 的 ISID,则发送 IS-IS 更新,并且多播连接性在作为 ISID 23 的成员的所有节点之间创建 EVPN。发送不同组的 IS-IS 更新以创建用于 ISID 10 的多播连接性。一旦创建了 ISID,则随后通过在链路状态发现期间使用



Sys-ID 创建的路径完成所有转发。进一步参照图 4, 每个服务是此基础拓扑结构的叶子。如整体地通过引用结合到本文中的题为 IP FORWARDING ACROSS A LINK STATE PROTOCOL CONTROLLED ETHERNET NETWORK 的 2008 年 5 月 5 日提交的共同待决美国专利申请序号 12/151,684 所描述的那样, IP 子网 306 直接映射到 Sys-ID。如整体地通过引用结合到本文中的题为 IMPLEMENTATION OF VPNs OVER LINK STATE PROTOCOL CONTROLLED ETHERNET NETWORK 的 2008 年 6 月 26 日提交的共同待决美国专利申请序号 12/215,350 所描述的那样, VRF 308 经由 ISID 映射。

[0050] 当前在通过引用结合到本文中的 IEEE 标准 802.1ag “Connectivity Fault Management (连接性故障管理)” 中所定义的以太网 OAM 定义供在以太网中使用的一组连接性故障管理协议。其包括: 连续性检查、链路跟踪、和环回协议。802.1ag 标准已被扩展成包括性能监视度和消息。此标准在同样通过引用结合到本文中的 ITU-T SG 13, Y.1731-“Requirements for OAM in Ethernet Networks (对以太网中的 OAM 的要求)” 处有所反映。但是, 这些标准中描述的机制不可直接适用于链路状态协议以太网。依照本发明, 链路状态协议以太网结合了用于故障识别、隔离、故障检修、和性能监视目的的 OAM 特征。

[0051] 802.1ag CFM 消息包括以下各项:

[0052] 连接性检查 - 这些是由维护端点周期性地发出的“心跳 (heartbeat)”消息。其允许维护端点检测它们自己之间的服务连接性的丢失。

[0053] 链路跟踪 - 其由维护端点应管理员的请求传送以跟踪到目的地维护端点的路径 (逐跳)。其允许传送节点发现关于路径的连接性数据。链路跟踪在概念上与 UDP Traceroute 类似。

[0054] 环回 - 其由维护端点应管理员的请求发送以检验到另一维护点的连接性。环回指示目的地是否可到达; 其不允许路径的逐跳发现。其中在概念上与 ICMP Echo (Ping) 类似。

[0055] 维护域

[0056] 在任何给定服务提供商网络内, 如图 5 所示, 以太网 CFM 依赖于由分级维护域组成的功能模型。由管理员对域分配唯一维护级 (在可能的八个之中), 这对定义域的分级关系有用。如果两个域嵌套, 则外面的域必须具有比内部域高的维护级。图 5 所示的是包括提供商域 404、包括 2 个运营商域 406 的客户域 402。维护端点 (正方形) 存在于维护域的边缘, 而维护中间点 (圆圈) 在域内部。因此, 中间点将转发 CFM 分组 (除非其为以该中间点为目的地的环回或链路跟踪), 而端点不转发 CFM 分组, 因为它们必须保持其在域内。其唯一例外情况是当端点也充当用于高级域的中间点时, 在这种情况下, 其将转发 CFM 分组, 只要其为较高级域的一部分。

[0057] 图 5 示出其中服务提供商使用两个运营商的网络来提供服务的示例。服务提供商维护级在 322 处示出。用于运营商 A 和运营商 B 的维护级在 324 处示出。两个特殊情况维护级是客户级 (320) 和物理层级 (326)。客户级允许客户测试连接性 (使用连接性检查) 并隔离出问题 (使用环回和链路跟踪)。另一方面, 物理层级定义可能的最窄维护域: 单个链路域。

[0058] 依照本发明的第一方面, 对以太网 OAM 标准进行修改以便针对基于典型生成树的以太网与链路状态协议控制的以太网之间的差而调整。依照本发明的第二方面, 新型

服务级 OAM 特征利用链路状态协议控制的以太网。依照本发明的第三方面,由 IP 服务通过用于性能监视和控制的链路状态协议控制的以太网使用以太网 OAM。

#### [0059] 基础设施 OAM

[0060] 依照本发明,在建立第一 I-SID 之前,链路状态协议控制的以太网可以在基础设施级处实现 CFM 消息。因此,CFM 消息被图 3 和 4 中的链路层利用,并且在图 5 中的链路 OAM 级处利用。这里,诊断 OAM 可以有助于测试节点之间的连接性,这是在其之间部署服务之前。

[0061] 802.1ag CMF 消息格式在图 6 中示出。依照 802.1ag 标准的某些 CFM 消息、即 LBM 消息采用单播目的地地址。能够将这些 CFM 消息用于诊断目的以检查链路状态控制的以太网络的拓扑结构是有用的。为此,需要链路状态拓扑结构中的节点的适当目的地地址。因此,依照本发明及如图 6 所示,对于采用单播目的地地址的 CFM 消息、即 LBM 和 LBR 消息,使用从目的地节点的 Sys-ID 导出的节点 MAC 地址(例如图 7,400、402、404)。此节点级 MAC 地址在链路状态协议交换时被安置在 FIB 中。

[0062] 诸如 mLBM 和 CCM 的某些 CFM 消息采用唯一广播目的地地址。这些地址与链路状态以太网协议是不相容的,因为 RPF 将中断并将导致环路。因此,在基础设施级,不使用这些消息。

[0063] 进一步依照本发明,对在基础设施级对 LTM CFM 消息进行寻址的方式进行修改。根据标准,LTM 消息采用众所周知的组多播 MAC 地址。然而,在链路状态控制的以太网中,在任何节点 FIB 中不存在多播条目,直至建立第一 I-SID 为止。因此,在此阶段被链路状态控制的以太网节点接收到的标准 LTM 消息将被丢弃。因此,本发明提供对标准实施方式的修改。依照本发明的 LTM 消息对目标目的地节点采用单播目的地地址(图 7,400、406、408)。再次地,采用的目的地地址是从目标目的地节点的 Sys-ID 导出的节点 MAC 地址。由于链路状态控制的以太网被预先配置而不是“泛洪并获悉”,到目的地的路径是已知的,因此,单播 LTM 消息可以遵循到目标节点的预配置路径。

[0064] 现在参照图 8,使用具有 OAM 的链路状态协议控制的以太网提供了用于复核基础设施级的连接性的机会。对于链路状态控制的以太网中的给定一个或多个节点,运营商可以检查链路状态数据库本身以查看链路状态协议生成什么连接。(420,422) 并且,运营商可以从节点起或在成对节点之间运行链路跟踪(424)以检查来查看 LTM 和 LTR 消息是否显示在节点之间存在的实际路径与链路状态协议如 FIB 所反映的那样最初建立的路径相匹配(426~430)。

#### [0065] 服务级 OAM

[0066] 在建立 I-SID 之后,链路状态协议控制的以太网还可以在服务级实现 CFM。以太网 OAM 被设计为在 I-SID 级操作,因此可以使用 802.1ag 和 Y.1733 标准并将其加强以便为链路状态协议控制的以太网提供服务级 OAM 功能。

#### [0067] mLT

[0068] 在典型的泛洪和反向路径获悉以太网中,所有 I-SID 遵循根在单个多播源地址处的同一多播分布路径。但是在链路状态协议控制的以太网中每个服务实例、即 ISID 是多播分布路径的根。因此,如果希望对链路状态协议控制的以太网中的服务实例路径进行故障检修,则作为使用与 ISID 路径不一致的单播 LTM 或基于标准的多播 LTM 的替代,使

用新的替代是有意义的。依照本发明的一个方面,因此在服务级提供新的 OAM 链路跟踪消息。这些链路跟踪消息不是使用图 6 的多播标准以太网 DA,而是使用 I-SID 多播地址作为其 DA(图 9,456)。通过使用 ISID 多播 DA,链路跟踪将遵循优化多播路径,该优化多播路径源于为其发起跟踪的节点而不是源于典型的以太网多播树。

#### [0069] 发现

[0070] 可以出于发现的目的使用服务级 OAM 以确认链路状态协议控制的以太网络的拓扑结构。例如,参照图 9,可以从 ISID 附接到的节点发起“显示 ISID 树 (show ISID tree)”命令 (454)。根据一个选项,可以使用 ISID mDA-不是 802.1ag 标准的 CFM mDA-从 ISID 节点发起 mLBM 命令(通配符 ping(wildcard ping))(458)。或者,对于每个 ISID 端点而言,可以在 ISID 内发起单播 LTM(traceroute)(460)。根据替换选项,可以从 ISID 节点发起前述 mLTM(通配符 traceroute)命令,其将跟踪多播 ISID 树的路径(456)。

[0071] 请注意,链路状态协议已用其网络拓扑结构视图填充网络中的所有节点。因此,例如,如图 10 步骤 500 ~ 506 所示,在链路状态数据库是 IS-IS 数据库的情况下,可以针对附接到给定 ISID 的所有端节点询问 IS-IS 数据库。然后,上述服务级 OAM 链路跟踪可以穿过(runthrough)数据平面以查看数据平面拓扑结构实际上是否是如控制平面指示其应该成为的那样布置的。

[0072] 还可以使用发现来确认网络中的路径。“显示 ISID 路径”命令(图 9,462)可以检验端点之间的路径。例如,为了示出节点 A 和 B 之间的 ISID 101 上的路径,将从 ISID 101 上的节点 A 向节点 B 发起 LTM(traceroute)(464)。再次地,LTM DA 是目的地节点的 Sys-ID(节点 B)的单播 DA-不是基于标准的 CFM DA。

#### [0073] 连接性

[0074] 再次,链路状态协议已用其网络拓扑结构视图填充网络中的所有节点。因此,例如,如图 10 步骤 508 ~ 512 所示,在链路状态数据库是 IS-IS 数据库的情况下,可以针对节点 A 与节点 B 之间的 I-SID 路径从 IS-IS 数据库中的任何节点进行询问。或者,可以针对到另一端节点的路径从 I-SID 上的端节点进行询问-例如从端节点 A 进行询问以显示到端节点 B 的路径。然后,上述服务级 OAM 链路跟踪可以穿过数据平面以查看数据平面拓扑结构实际上是否是如控制平面指示其应该成为的那样布置的。

[0075] 还可以将服务 OAM 用于 I-SID 端点之间和在 I-SID 内的连接性检验和故障检测。可以从附接到 I-SID 的端节点发出等效于 CFM CCM 的 OAM 消息作为连接性检查机制(图 10,514)。再次地,与 CFM-DA 相反,将基于 I-SID mDA 对这些 CCM 消息进行寻址(即,解析为 Sys-ID)。此外,可以在每个服务级发出这些 CCM 消息。IP 子网级 CCM 消息被直接解析为 Sys-ID,而经由 I-SID 解析 IP-VPN、VRF 等。

#### [0076] MEP/MIP 自动生成

[0077] 依照本发明的方面,链路状态协议控制的以太网络允许 MEP 和 MIP 的自动生成。

[0078] 作为链路跟踪协议发现的一部分,链路状态协议控制的以太网络中的每个节点自动地实例化默认 MD 级 802.1ag 逻辑,但是这可以使用被转换成 MAC 地址的 Sys-ID 名来进行。依照本发明,如图 11 所示,在基础设施级,每个节点可以对其 Sys-ID 进行散列(hash)以导出其 MEP 和 / 或 MIP(600),并随后用此信息填充 TLV(602)。然后在链路状态 PDU(LSP)中将 TLV 传送到网络上(604)。在图 12 中,其显示节点何时接收这样的 LSA(610),其使 TLV

中接收到的 MEP 信息与从其接收到 LSA 的端节点相关联。接收节点向其 FIB 添加条目以使 MEP 与从其接收到 LSA 的节点的节点 MAC 相关联以创建 MEP/Sys-ID 绑定。因此,每个节点随后知道对于网络中的每隔一个节点而言 MIP 和 MEP 点是什么。

[0079] 因此,运营商可以从特定节点的观点出发执行基础设施级 OAM 命令。例如,如图 13 步骤 620 ~ 626 所示,运营商选择执行节点 A 和 B 之间的连续性检查。因此,从节点 A 起,运营商执行以太网 OAMLBM- 即“ping”命令。依照本发明,节点 A 针对节点 B 的 MEP 检查其链路状态数据库- 先前在链路状态配置期间被填充。一旦这是已知的,则创建具有节点 B 的目的地址的 LBM 消息。节点 A 的 FIB 指示 LBM 消息应当在其到节点 B 的路上被发送到下一跳 MIP ( 如果的确存在 A 与 B 之间的节点 )。

[0080] 如图 5 所示,不同的维护域与不同的 MEP 和 MIP MD 级相关联。因此,在服务级,指定不同的 MEP 和 MIP 集。链路状态协议控制的以太网络允许在各种服务级根据需要进行 MEP 和 MIP 的动态自动配置。在基础设施级,用于监视链路的端口 MEP 在 802.1ag 标准中阐述的“默认”MAID 级 (MD 级为 0) 处被实例化,并且始终开启 (on)。各种服务级还可以具有用于载送诸如 CCM 的消息的始终开启 MEP。这些 MEP 是根据服务级标识符 I-SID 用 MAID 创建的,并且 MD 级适合于域。可以将 MIP 创建为在链路状态协议发现期间始终开启,MD 级适合于域。

[0081] 链路状态协议以太网络中的 IP OAM

[0082] 如前所述,如通过引用结合到本文中的题为 IP FORWARDING ACROSS A LINK STATE PROTOCOL CONTROLLED ETHERNET NETWORK 的 2008 年 5 月 5 日提交的共同待决美国专利申请序号 12/151,684 中描述的那样,IP 地址可以直接映射到在链路状态协议以太网络中进行转发时使用的 MAC 地址。如其中所解释的,当链路状态协议控制的以太网络中的节点获悉了 IP 地址时,其将把 IP 地址插入其链路状态广告中以便对该 IP 地址到网络上的其它节点的可达性进行广告。每个节点将为此 LSP 添加其通知给其链路状态数据库的 IP 地址。如果分组到达入口节点,则该入口节点将读取 IP 地址,确定链路状态协议控制的以太网络上的哪个节点知道 IP 地址,并构造 MAC 报头以便将分组转发到正确的节点。MAC 报头的 DA/VID 是对 IP 地址进行广告的节点的节点 MAC- 例如,这可以是 Sys-ID。

[0083] 由于 IP 子网可以一致地映射到链路状态协议控制的以太网络,所以 MEP 和 MIP 的自动创建和用于链路状态协议控制的以太网络的增强 OAM 使得能够实现用于 IP 的 OAM 功能,诸如基于以太网 OAM 的 Ping 和 Traceroute 能力。

[0084] 例如,参照图 14,再次示出了链路状态协议控制的以太网络,其中,已如前所述地自动配置了 MEP 和 MIP。具有 Sys-ID San Jose 的节点被示为具有 IP 地址 10.20.0.16/24。具有 Sys-ID Denver 的节点被示为具有 IP 地址 10.20.8.128/24。参照图 15,运营商在节点 San Jose 处输入 IP 命令“Ping 10.20.8.128” (720)。(或者,可能存在经由 DNS 或 IP 到名称转换的某些其它手段被解析的等效 IP 名称。)San Jose 处的节点先前已从宣告 10.20.8.128 附接到其的 Denver 接收到 LSA,因此 San Jose 的数据库将目的地 IP 地址解析成 Denver 的 MAC (722)。IP ping 命令被解析为具有目的地 Denver 的以太网 OAM LBM 命令 (724)。San Jose 节点检查其 FIB 以找到用于 Denver 的 MEP。LBM 被发送到 DA Denver、VID MIP (726)。采用 Denver 与 San Jose 之间的基础设施连续性,LBR 返回到 San Jose。

[0085] 同样地,参照图 16,节点 San Jose 处的运营商可以输入 IP 命令“Traceroute

10.20.8.128”(740)。(再次地,可能存在等效的 IP 名称解析。)San Jose 处的节点先前已从宣告 10.20.8.128 附接到其的 Denver 接收到 LSA,因此 San Jose 的数据库将目的地 IP 地址解析成 Denver 的 MAC(742)。IP traceroute 命令被解析成具有目的地 Denver 的以太网 OAM LT 命令(744)。San Jose 节点检查其 FIB 以找到用于 Denver 的 MEP。LTM 被发送到 Denver、VID MIP(746)。

#### [0086] 性能监视

[0087] 802.1ag 标准已被扩展成包括性能监视度和消息。此标准在通过引用结合到本文中的 ITU-T SG 13, Y.1731-“Requirements for OAM in Ethernet Networks(对以太网网络中的 OAM 的要求)”处有所反映。用适当的 OAM 消息来测量以下性能参数:

[0088] 1) 丢帧率(Frame Loss Ratio,FLR)-FLR 被定义为未递送的服务帧的数目除以时间间隔 T 期间的服务帧总数的比值(表示为百分比),其中,未递送的服务帧的数目是被发送到入口 UNI 的服务帧的数目与在出口 UNI 处接收到的服务帧的数目之间的差。可以进行两种 FLR 测量,双端 LM(丢失测量)和单端 LM。双端 LM 是通过交换包括适当计数的所传送帧和所接收帧的 CCM OAM 帧实现的。这些计数不包括 MEP ME 级的 OAM 帧。双端 LM 使得能够实现 MEG 的每个端处的近端和远端 FLR 二者的主动(proactive)测量。单端 LM 是通过 LMM 和 LMR OAM 帧的应需求交换实现的。这些帧包括适当计数的所传送和所接收帧。单端 LM 仅在发起 LM 请求的端处提供近端和远端 FLR。

[0089] 2) 帧延迟(FD)-FD 被指定为帧的往返行程延迟,其中,FD 被定义为当在帧的目的地节点处执行环回时自源节点开始传输帧的第一位时起直到同一源节点接收到环回帧的最后一位为止所经历的时间。

[0090] 3) 帧延迟变化(FDV)-FDV 是对服务帧对之间的 FD 的变化的度量,其中,服务帧属于点到点以太网连接上的同一 CoS(服务级别)实例。

[0091] 又请注意,在许多情况下,IP 子网一致地映射到链路状态协议控制的以太网络。现在,MEP 和 MIP 的自动创建和用于链路状态协议控制的以太网络的增强 OAM 和性能监视使得能够实现用于迄今尚不可用的 IP over Ethernet 的细粒度的详细“SONET 样式的(SONET-style)”OAM。

[0092] 参照图 17,示出了许多应用之一,其中,以太网性能 OAM 在 IP 应用中是很有价值的。示出了提供商 800 和客户驻地 802。客户具有经由接入箱 806 经由链路状态协议控制的以太网络 808 耦合到提供商 800 的 IP 电话 804。在提供商 800 内,可能存在将网络 808 耦合到向接入箱 806 提供 VOIP 服务并因此向 IP 电话 804 提供 VOIP 服务的服务器 812 的各种桥接器 810。图 18 中的所有设备都是 IP 设备。IP 电话、服务器和桥接器、以及未示出的其它桥接器和设备构成链路状态协议控制的以太网络。同样地,它们每个都具有与之相关联的 Sys-ID。诸如 IS-IS 的链路状态协议已建立网络中的所有元件之间的单播无环路通信路径。IP 电话和服务器已依照在通过引用整体结合到本文中的题为 IP FORWARDING ACROSS A LINK STATE PROTOCOL CONTROLLED ETHERNET NETWORK 的 2008 年 5 月 5 日提交的共同待决美国专利申请序号 12/151,684 中所述的方法建立 IP 通信。简而言之,IP 电话 IP 子网被 IP 电话节点获悉并插入其链路状态广告中以便对 IP 子网到链路状态协议控制的以太网上的其它节点的可到达性进行广告。同样地,服务器的 IP 子网被服务器节点获悉并插入其链路状态广告中以便对其 IP 子网到链路状态协议控制的以太网上的其它节

点的可达性进行广告。如果分组到达入口节点,则该入口节点将读取 IP 地址,确定链路状态协议控制的以太网上的哪个节点知道 IP 地址,并构造 MAC 报头以便将分组转发到正确的节点。MAC 报头的 DA/VID 是对 IP 地址进行广告的节点的节点 MAC。在这种情况下,从 IP 电话到服务器的 IP 流(即 VOIP)将解析到 IP 电话附接到其的节点的 MAC。从服务器到 IP 电话的 IP 流将解析到服务器附接到其的节点的 MAC。

[0093] 如先前关于 CFM OAM 所述,运营商可以执行可以被直接映射到链路状态以太网命令的诸如“ping”和“traceroute”的 IP 级命令。进一步依照本发明,基于链路状态以太网 OAM 命令和反馈来提供 IP 级性能监视功能。

[0094] 例如,参照图 18,期望的是服务器 812 监视延迟和与 IP 电话 804 相关的特定 VOIP 流的抖动。依照本发明,由如下事实来实现此任务:即 VOIP 流通过链路状态协议控制的以太网网络载送,并因此直接利用本文所述的 OAM 功能。例如,运营商可以从服务器节点 812 发起命令“监视在下一小时内 IP 电话的抖动、延迟”(820)。IP 级 OAM 命令将在服务器 812 处被解析成服务器 812 与 IP 电话 804 之间的一系列以太网级 OAM 命令。将在本示例中使用的 OAM 级命令是 FD 和 FDV。首先,通过检查 FIB 来解析具有附接的 IP 电话或电话本身的节点的 MAC 地址(822)。然后,可以将 OAM FD 和 FDV 命令从附接到服务器 812 的节点转发到附接到 IP 电话 804 的节点,在识别的时间段内在服务器 812 的 FIB 中指定的 MIP 桥接器 810 转发路径(826)。因此可以以迄今不可用于以太网上的 IP 流的非常详细的方式为 IP 流收集性能统计。然后,如果需要,可以基于来自 OAM 命令的所产生反馈来调整 VOIP 流(828)。

[0095] 依照用于许多 IP 技术的本发明来实现 IP 性能监视,所述许多 IP 技术包括 IP 电话术、IP TV/ 视频、移动 IP、数据中心等。链路状态协议控制的以太网使得能够实现 IP 性能监视和控制以便将许多不同类型和层级的 IP 域和设备联合。依照本发明直接在 IP 层级利用以太网 OAM 性能监视的能力使得能够实现将容易地导致详细 LSA 的语音、数据和视频的 IP 业务控制级。

[0096] 可以将本发明实现为体现在一个或多个制品上或中的一个或多个计算机可读软件程序。所述制品可以是例如软盘、硬盘、硬盘驱动器、CD-ROM、DVD-ROM、闪存卡、EEPROM、EPROM、PROM、RAM、ROM、或磁带中的任何一个或组合。通常,可以使用任何标准或专有编程或解释语言来产生计算机可读软件程序。此类语言的示例包括 C、C++、Pascal、JAVA、BASIC、Visual Basic、和 Visual C++。软件程序可以被作为源代码、目标代码、解释代码、或可执行代码存储在一个或多个制品上或存储在其中。

[0097] 虽然已参照特定优选实施例示出并描述了本发明,但本领域的技术人员应理解的是在不脱离以下权利要求所定义的本发明的精神和范围的情况下可以进行形式和细节方面的各种改变。

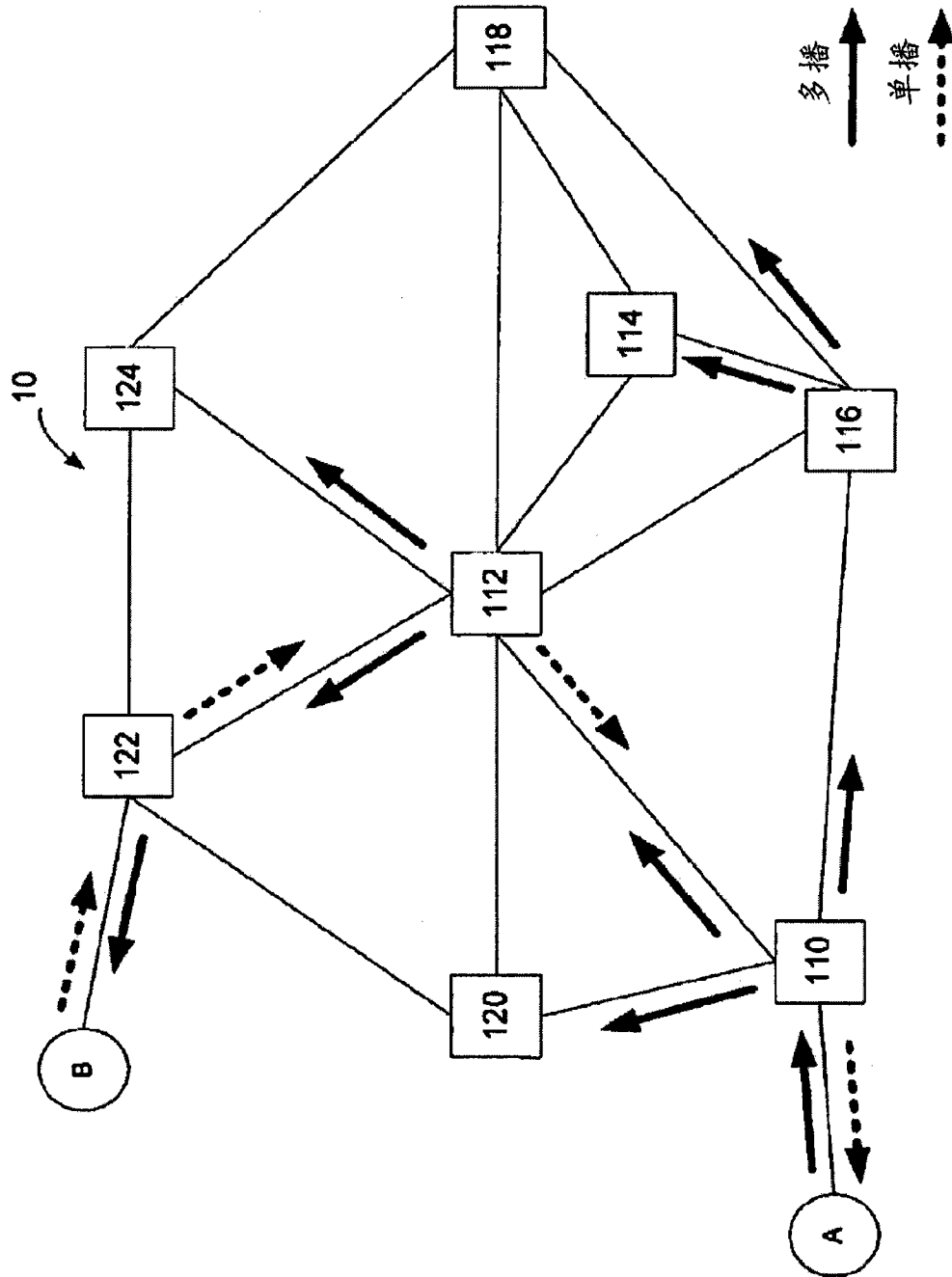


图 1

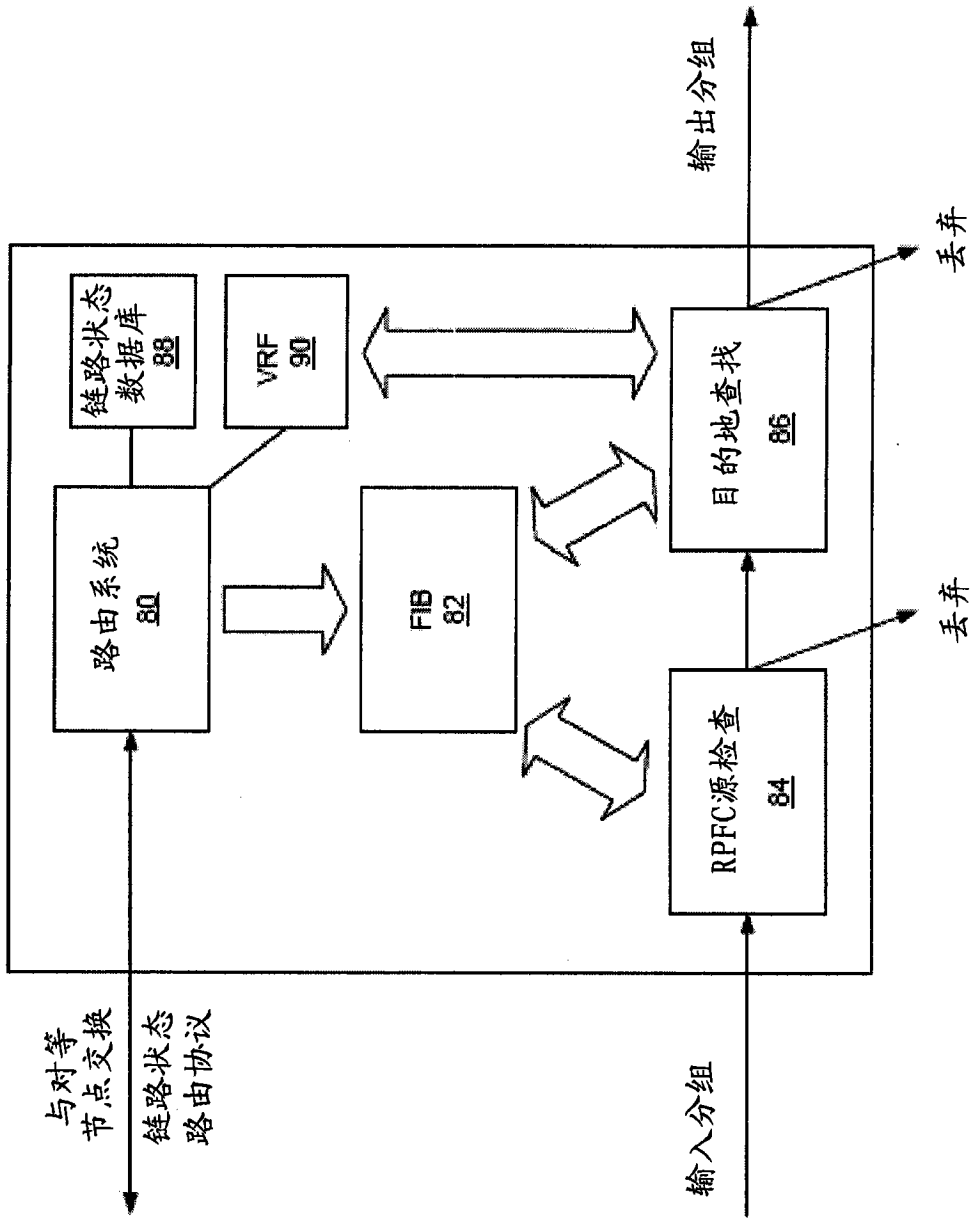


图 2



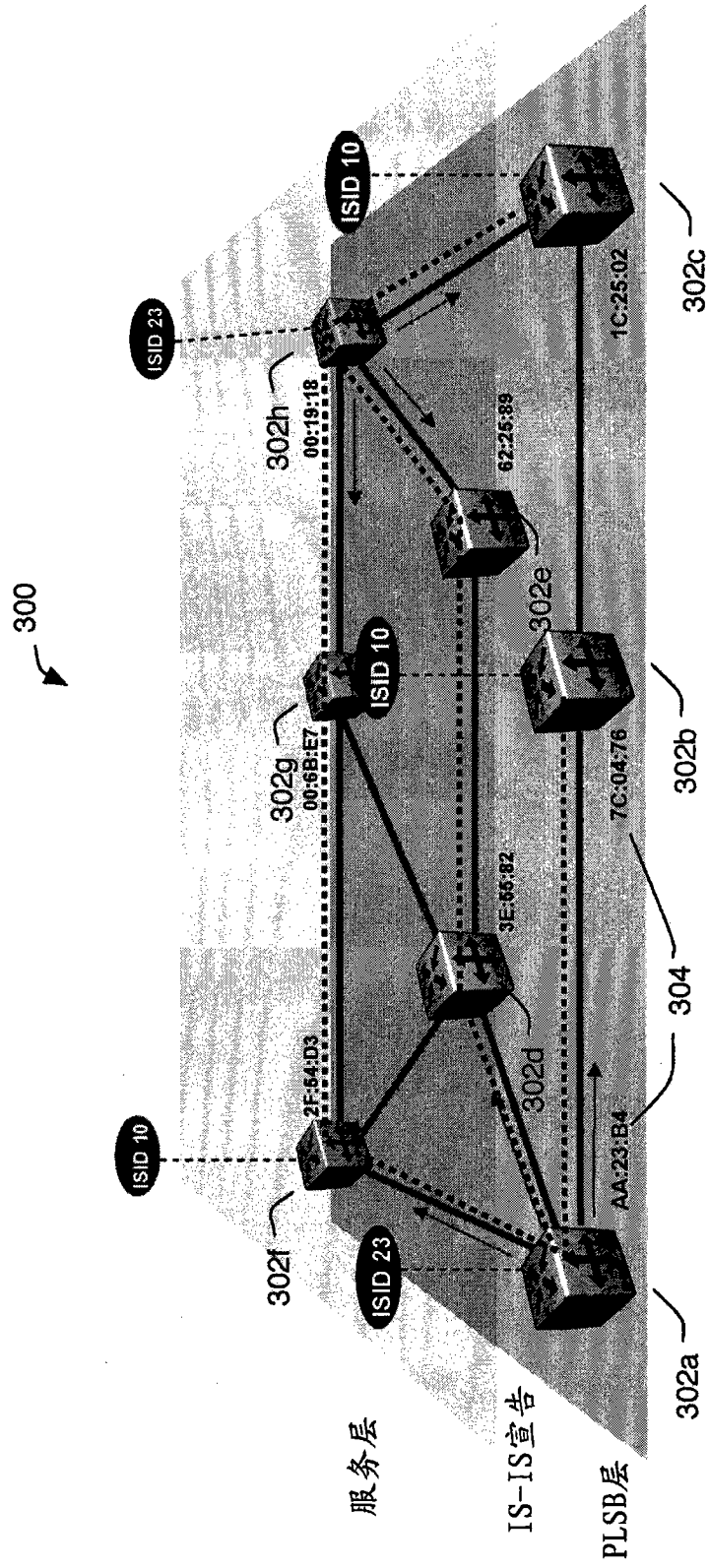


图 3

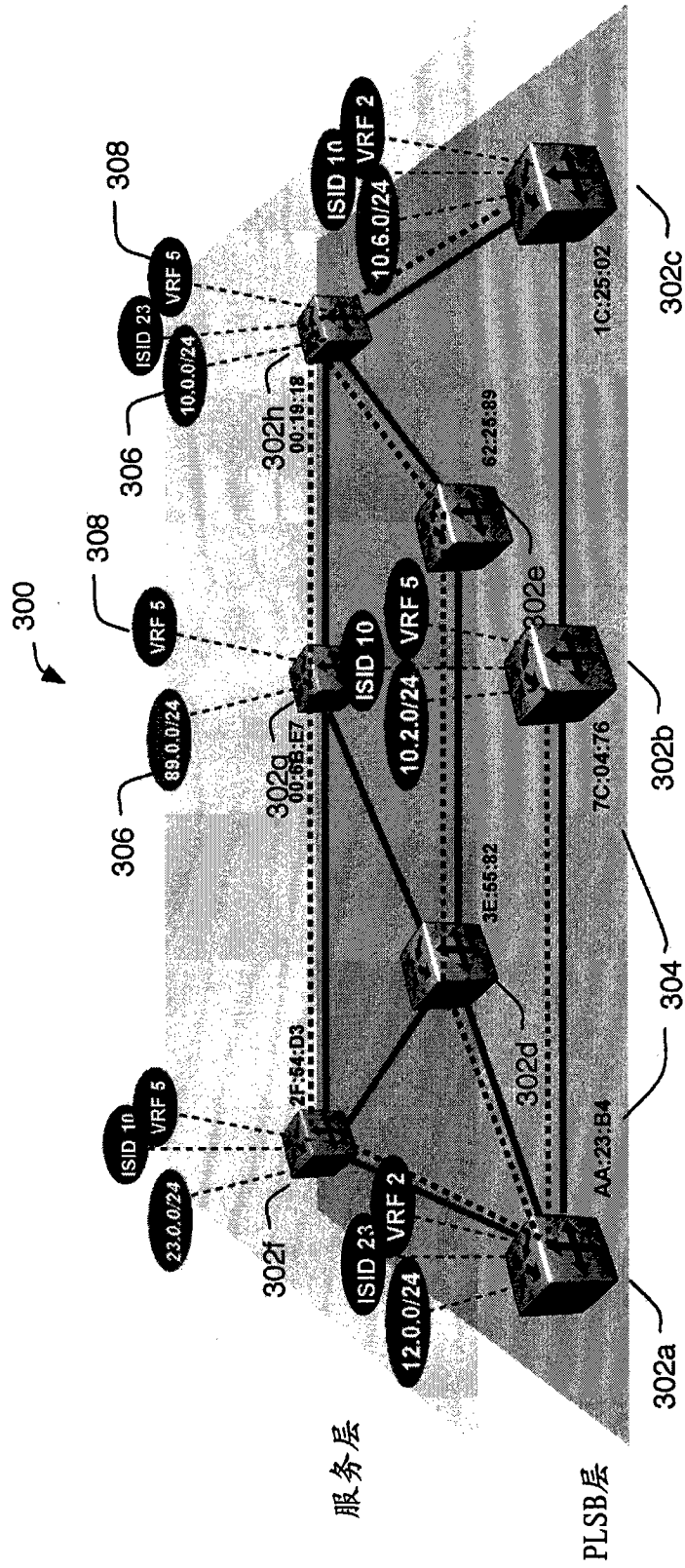


图 4

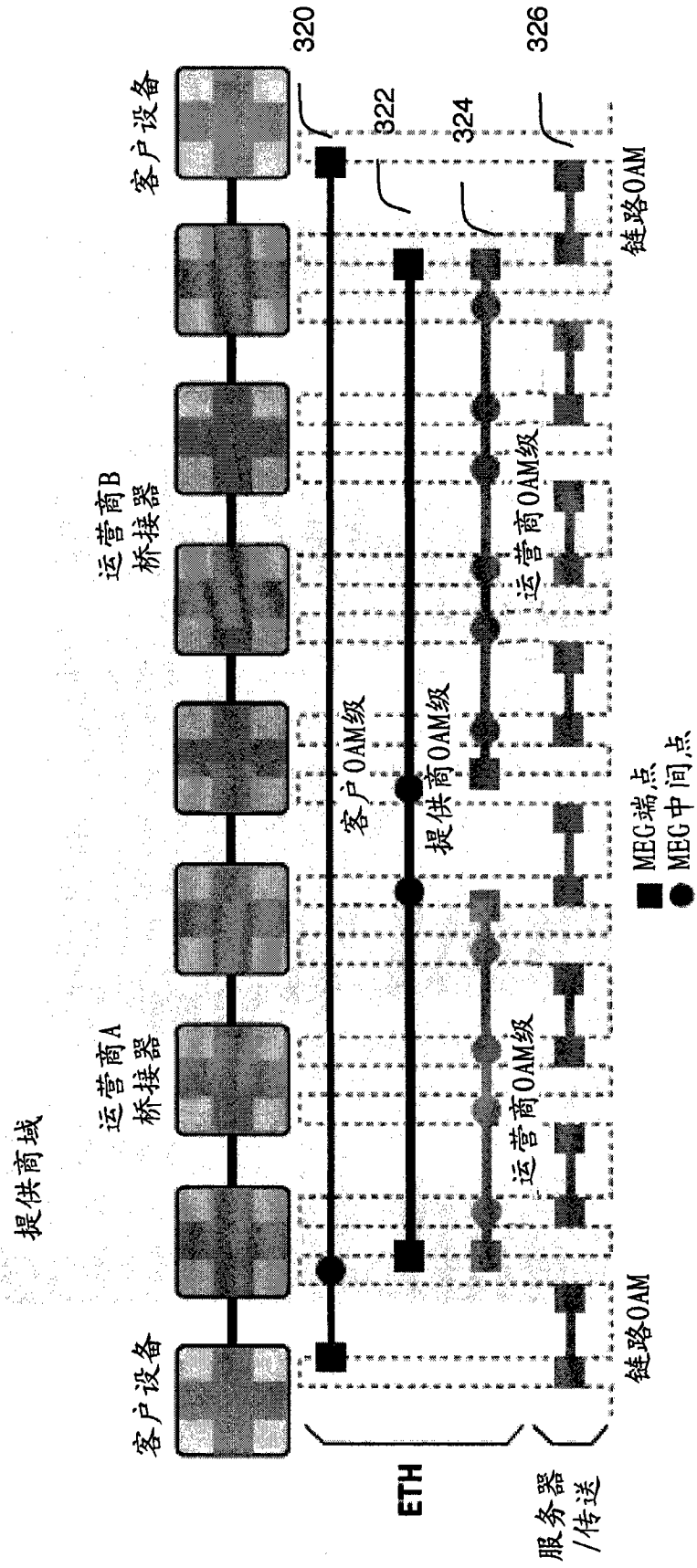


图 5

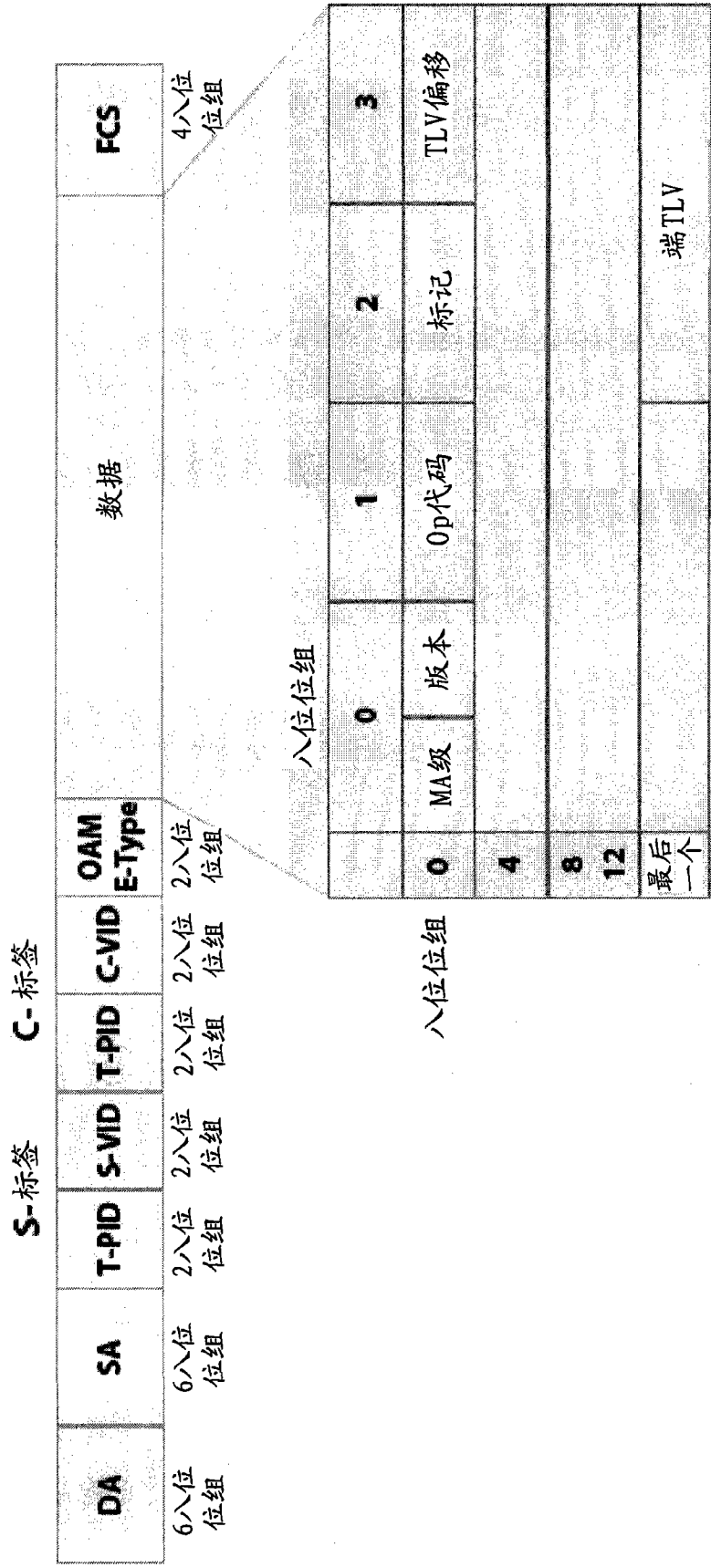


图 6

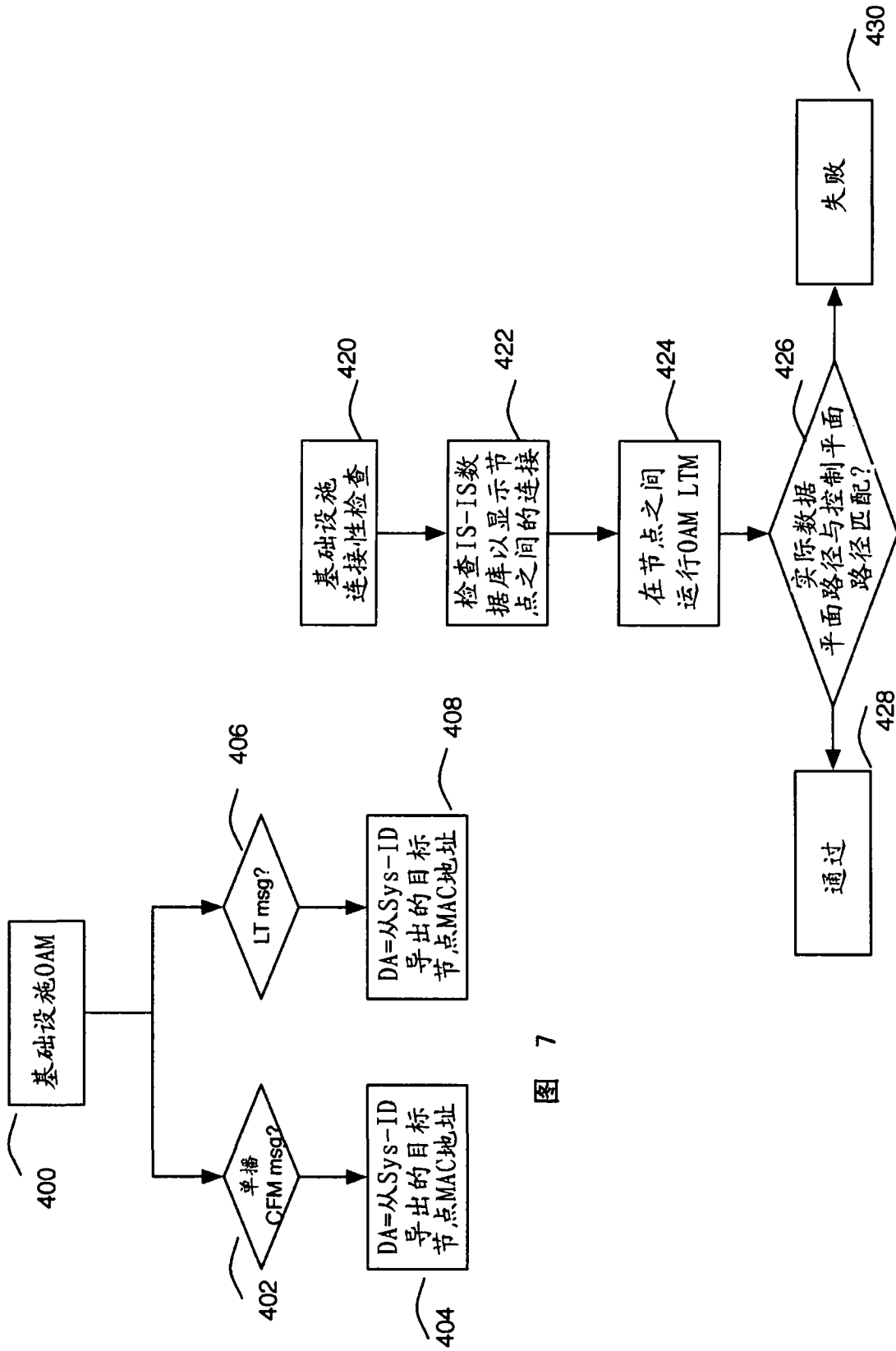


图 7

图 8

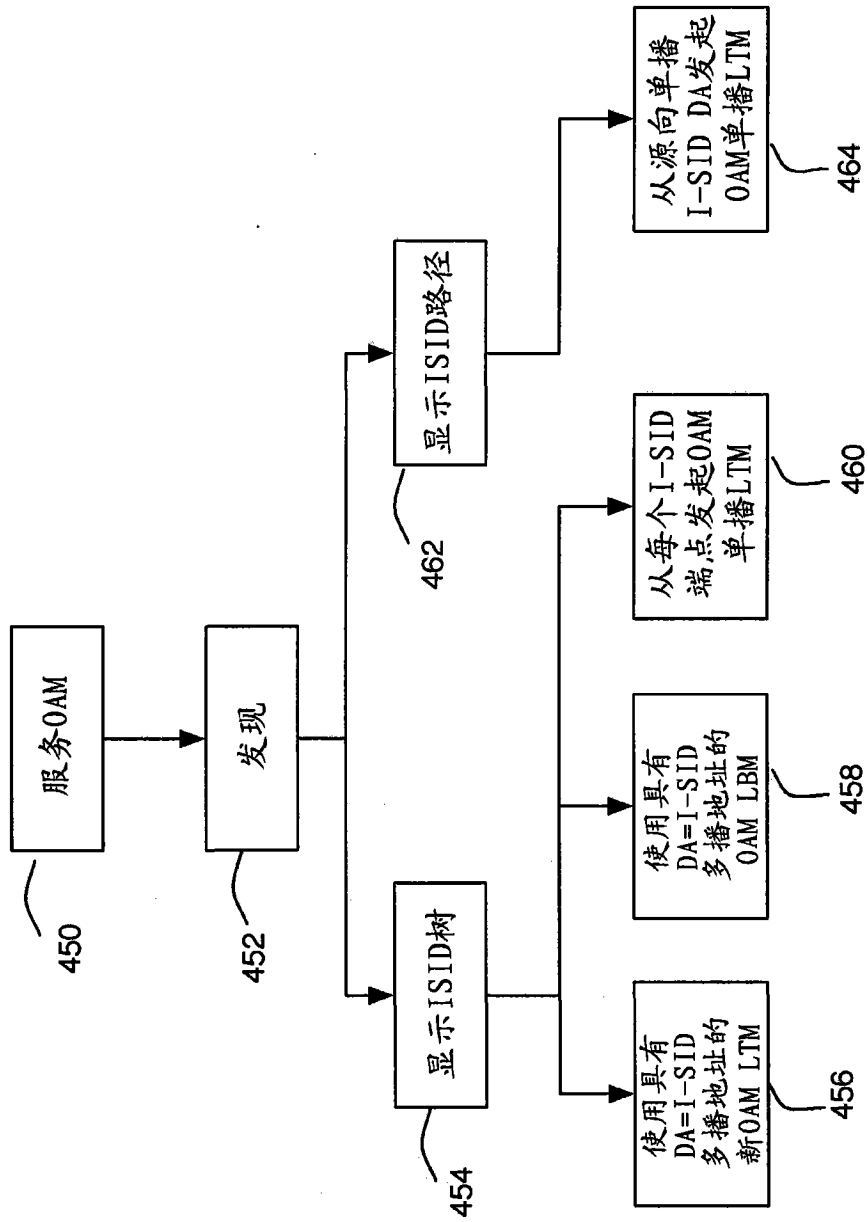


图 9

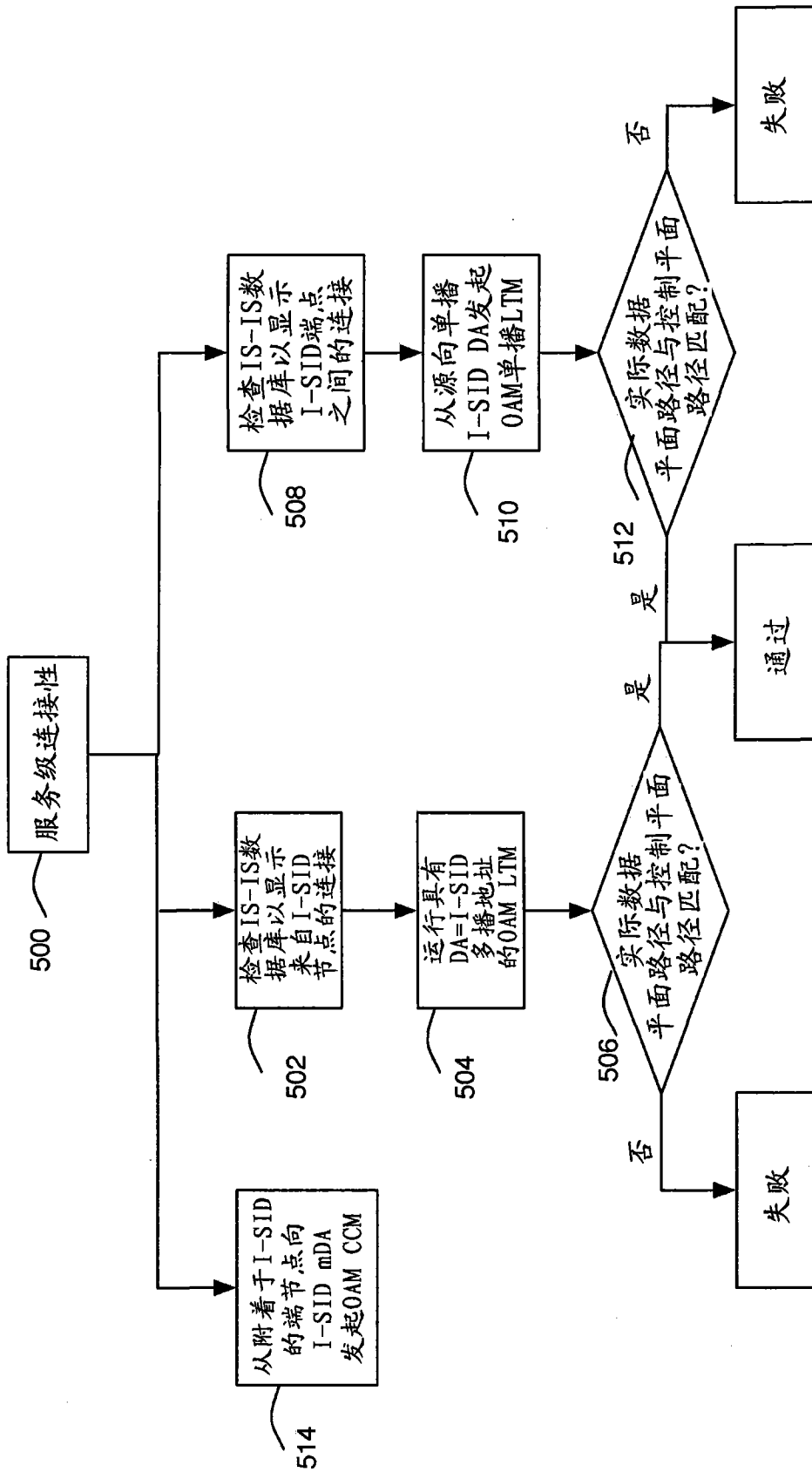


图 10

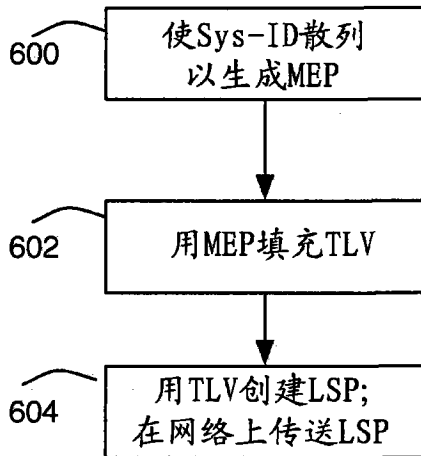


图 11

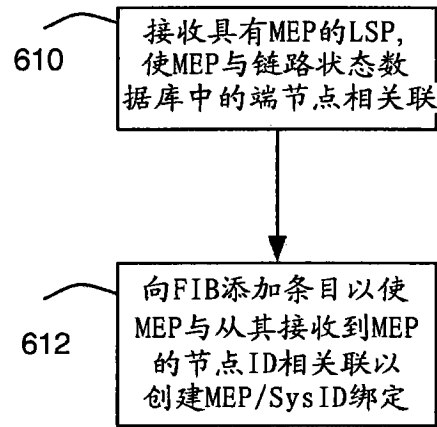


图 12

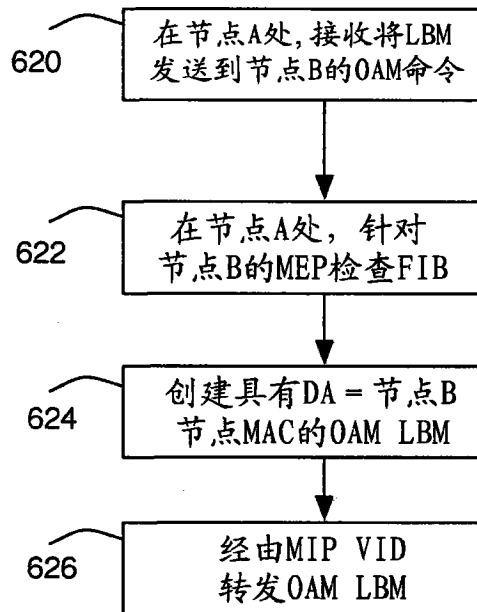


图 13



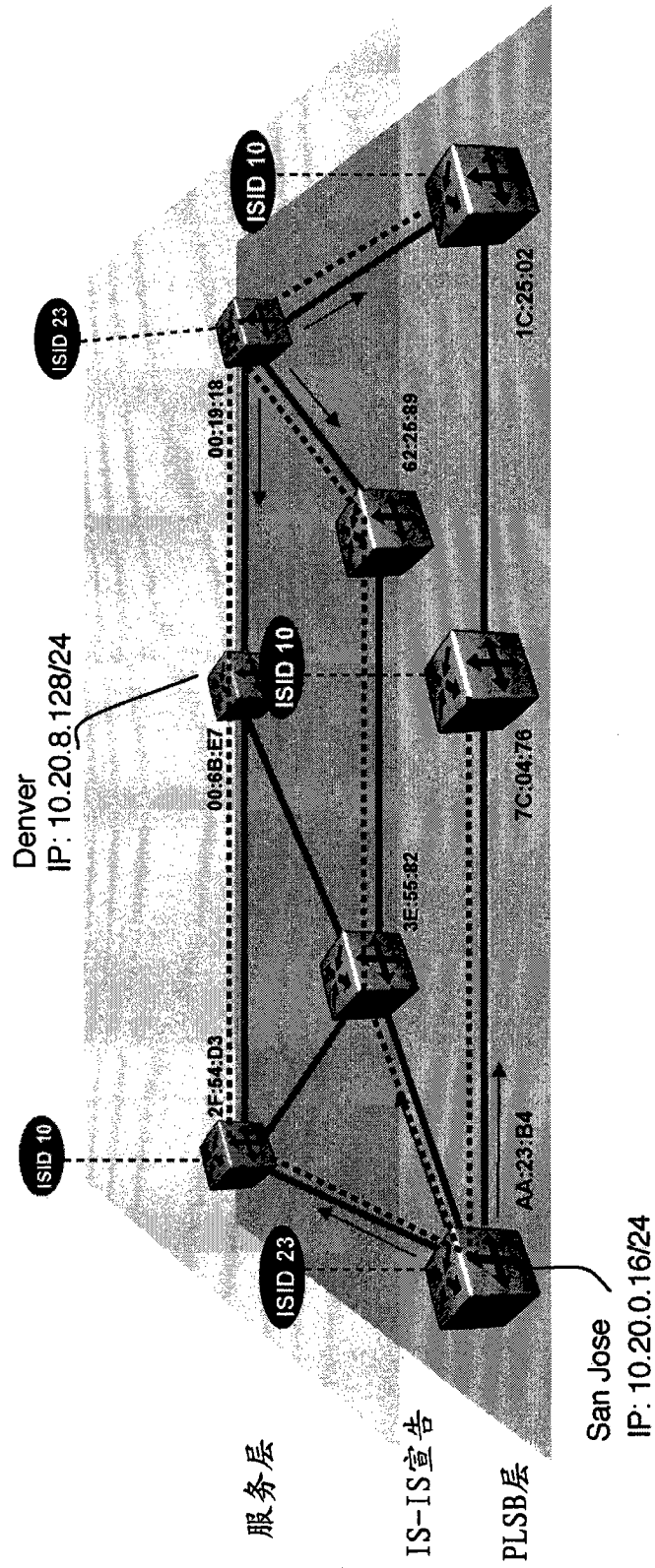


图 14

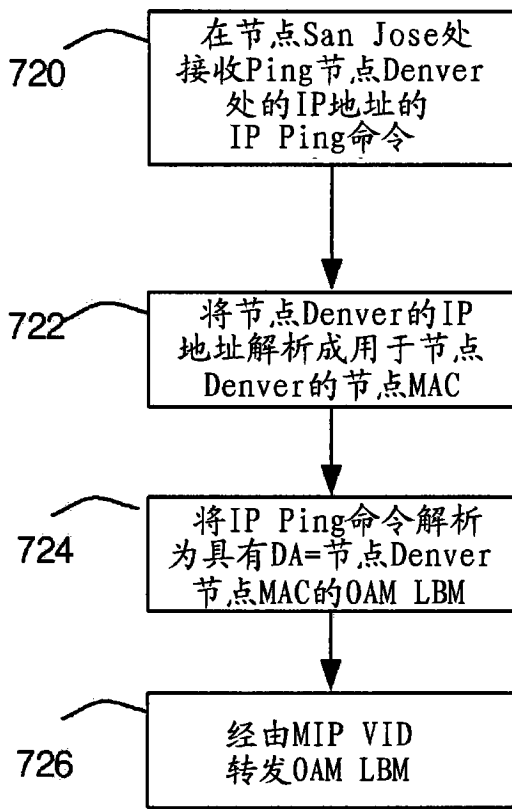


图 15

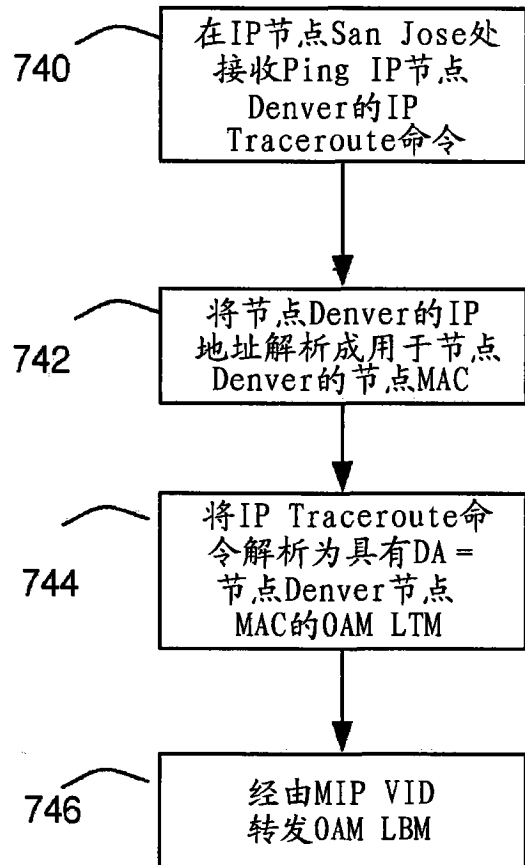


图 16

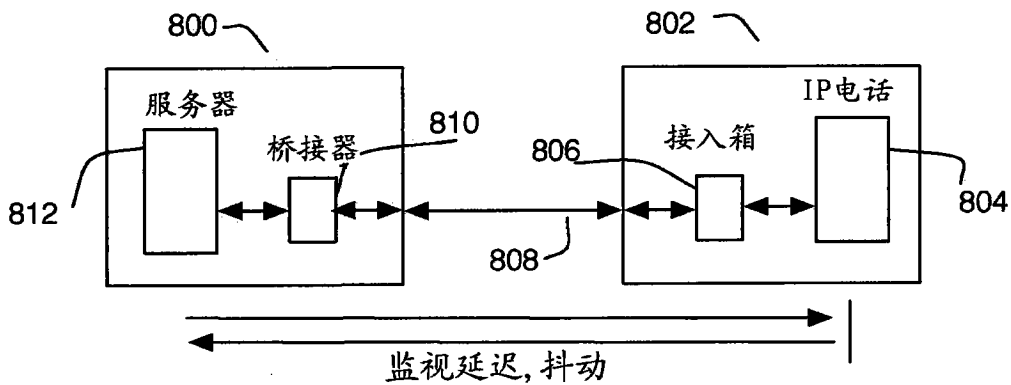


图 17

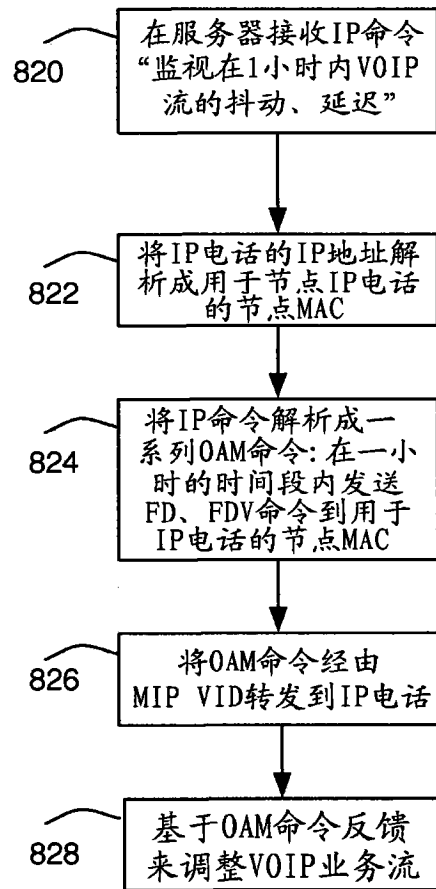


图 18