



(12)发明专利

(10)授权公告号 CN 104025508 B

(45)授权公告日 2018.01.05

(21)申请号 201280065228.8

(22)申请日 2012.12.11

(65)同一申请的已公布的文献号
申请公布号 CN 104025508 A

(43)申请公布日 2014.09.03

(30)优先权数据
13/341,726 2011.12.30 US

(85)PCT国际申请进入国家阶段日
2014.06.27

(86)PCT国际申请的申请数据
PCT/US2012/068996 2012.12.11

(87)PCT国际申请的公布数据
W02013/101453 EN 2013.07.04

(73)专利权人 思科技术公司
地址 美国加利福尼亚州

(72)发明人 劳伦斯·克里格
萨拉齐·内里卡尔
苏达尔莎娜·坎达彻尔·司理得哈

拉·姚
尤达亚玛卡尔·斯瑞尼瓦桑
阿奴瑞格·米塔尔
里连·西尔维娅·费尔南德斯

(74)专利代理机构 北京东方亿思知识产权代理
有限责任公司 11258

代理人 李晓冬

(51)Int.Cl.
H04L 12/18(2006.01)
H04L 12/46(2006.01)

(56)对比文件
US 2008089244 A1,2008.04.17,
CN 101765827 A,2010.06.30,
CN 102067533 A,2011.05.18,
US 7924742 B2,2011.04.12,
CN 103026347 A,2013.04.03,
CN 102045385 A,2011.05.04,
CN 102244653 A,2011.11.16,

审查员 潘智慧

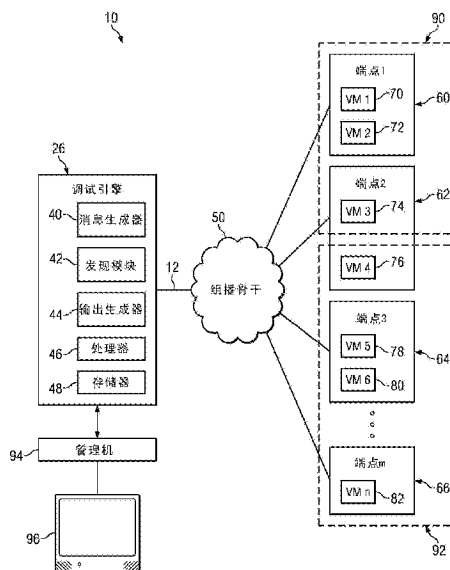
权利要求书2页 说明书11页 附图5页

(54)发明名称

用于在网络环境中发现多点端点的方法、设备及装置

(57)摘要

一种示例方法被提供并且可以包括:在覆盖网络中组播发现分组,所述覆盖网络包括越过第3层网络的第2层方案;以及基于各个端点对所述发现分组的响应来识别所述端点,其中所述端点跨越组播骨干被耦合。在更具体的实施例中,所述方法可以包括基于缺少来自自己断开端点的响应识别所述覆盖网络中的已断开端点。



1. 一种用于在网络环境中发现多点端点的方法,包括:

在覆盖网络中向跨越组播骨干被耦合的端点组播发现分组,其中,所述端点中的每个端点包括虚拟以太网模块(VEM),并且其中每个端点与所述覆盖网络的多于一个的段相关联,其中所述发现分组包括与所述多于一个的段中选定的段相对应的唯一的段标识符,其中所述覆盖网络包括越过第3层网络的第2层方案,并且其中所述发现分组包括被配置为发现特定段中的端点的链路层分组;以及

基于各个端点对所述发现分组的响应来识别所述端点。

2. 根据权利要求1所述的方法,进一步包括:

基于缺少来自已断开端点的响应识别所述覆盖网络中的已断开端点。

3. 根据权利要求1所述的方法,其中所述端点被配置成支持一个或多个虚拟机(VM)。

4. 根据权利要求1所述的方法,其中所述发现分组被以类型长度值(TLV)格式提供并且包括所述覆盖网络的所述特定段的标识符。

5. 根据权利要求1所述的方法,其中所述覆盖网络是一组网络中的选择的一个,所述一组网络包括:

a) 虚拟可扩展局域网(VXLAN);

b) 网络虚拟化通用路由封装(NV-GRE)网络;以及

c) 媒体访问控制(MAC)-in-MAC网络。

6. 根据权利要求1所述的方法,其中所述覆盖网络是包括由对应的VXLAN段标识符(VNI)来识别的一个或多个段的多租户VXLAN。

7. 根据权利要求1所述的方法,进一步包括:

将所述响应聚合成多点信息;以及

显示所述多点信息。

8. 根据权利要求1所述的方法,其中所述发现分组是思科发现协议(CDP)分组和链路层发现协议(LLDP)分组中的选择的一个。

9. 一种用于在网络环境中发现多点端点的设备,包括:

用于在覆盖网络中向跨越组播骨干被耦合的端点组播发现分组的装置,其中,所述端点中的每个端点包括虚拟以太网模块(VEM),并且其中每个端点与所述覆盖网络的多于一个的段相关联,其中所述发现分组包括与所述多于一个的段中选定的段相对应的唯一的段标识符,其中所述覆盖网络包括越过第3层网络的第2层方案,并且其中所述发现分组包括被配置为发现特定段中的端点的链路层分组;以及

用于基于各个端点对所述发现分组的响应来识别所述端点的装置。

10. 根据权利要求9所述的设备,进一步包括:

用于基于缺少来自已断开端点的响应识别所述覆盖网络中的已断开端点的装置。

11. 根据权利要求9所述的设备,其中所述端点被配置成支持一个或多个虚拟机(VM)。

12. 根据权利要求9所述的设备,其中所述发现分组被以类型长度值(TLV)格式提供并且包括所述覆盖网络的所述特定段的标识符。

13. 根据权利要求9所述的设备,其中所述覆盖网络是一组网络中的选择的一个,所述一组网络包括:

a) 虚拟可扩展局域网(VXLAN);

- b) 网络虚拟化通用路由封装 (NV-GRE) 网络; 以及
- c) 媒体访问控制 (MAC)-in-MAC 网络。

14. 根据权利要求9所述的设备, 其中所述覆盖网络是包括由对应的VXLAN段标识符(VNI)来识别的一个或多个段的多租户VXLAN。

15. 根据权利要求9所述的设备, 进一步包括:
用于将所述响应聚合成多点信息的装置; 以及
用于显示所述多点信息的装置。

16. 根据权利要求9所述的设备, 其中所述发现分组是思科发现协议(CDP)分组和链路层发现协议(LLDP)分组中的选择的一个。

17. 一种用于在网络环境中发现多点端点的装置, 包括:
调试引擎;
存储器, 其用于存储指令; 以及

处理器, 其可操作来执行所述指令使得所述装置被配置用于:

在覆盖网络中向跨越组播骨干被耦合的端点组播发现分组, 其中, 所述端点中的每个端点包括虚拟以太网模块(VEM), 并且其中每个端点与所述覆盖网络的多于一个的段相关联, 其中所述发现分组包括与所述多于一个的段中选定的段相对应的唯一的段标识符, 其中所述覆盖网络包括越过第3层网络的第2层方案, 并且其中所述发现分组包括被配置为发现特定段中的端点的链路层分组; 以及

基于各个端点对所述发现分组的响应来识别所述端点。

18. 根据权利要求17所述的装置, 其中所述装置被进一步配置用于:
基于缺少来自已断开端点的响应识别所述覆盖网络中的已断开端点。

19. 根据权利要求17所述的装置, 其中所述装置被进一步配置用于:
将所述响应聚合成多点信息; 以及
显示所述多点信息。

20. 根据权利要求17所述的装置, 其中所述发现分组被以类型长度值(TLV)格式提供并且包括所述覆盖网络的所述特定段的标识符。

用于在网络环境中发现多点端点的方法、设备及装置

技术领域

[0001] 本公开一般地涉及通信领域,并且更特别地,涉及用于在网络环境中发现多点端点的系统和方法。

背景技术

[0002] 信息技术(IT)界中的虚拟化趋势正在驱使许多虚拟化技术的开发。一个这样的技术涉及服务器虚拟化。服务器虚拟化已在诸如在数据中心内的许多网络环境中赢得突出优势。服务器虚拟化能够掩盖实际服务器用户的服务器资源(例如,掩盖各个物理服务器、处理器以及操作系统(OS)的数目和身份)。服务器管理员使用软件应用来将一个物理服务器划分成一个或多个隔离的虚拟环境(常常被称为虚拟机)。服务器虚拟化能够有效地被用来消除服务器蔓延、帮助灾难恢复、用于测试和开发、更高效地利用服务器资源、提供服务器可用性、以及使服务器管理集中化。

附图说明

[0003] 为了提供对本公开及其特征和优点的更透彻理解,对结合附图进行的以下描述进行参考,其中同样的附图标记表示同样的部分,在附图中:

[0004] 图1是图示根据一个实施例用于在网络环境中发现多点端点的系统的简化框图;

[0005] 图2是图示依照一个实施例的系统的示例细节的简化框图;

[0006] 图3是图示可以与系统的实施例相关联的示例操作的简化流程图;

[0007] 图4是可以与系统的实施例相关联的简化示例显示屏幕;以及

[0008] 图5是图示可以与系统的实施例相关联的示例操作的简化流程图。

具体实施方式

[0009] 概述

[0010] 提供了示例方法,该方法可以包括:在覆盖网络中组播发现分组,所述覆盖网络包括第3层网络之上的第2层方案;以及基于各个端点对发现分组的响应来识别端点,其中端点跨越组播骨干被耦合。在更具体的实施例中,该方法可以包括基于缺少来自自己断开端点的响应在覆盖网络中识别断开的端点。

[0011] 端点可以是配置成支持一个或多个虚拟机(VM)的虚拟以太网模块(VEM)。发现分组能够被以类型长度值(TLV)格式提供并且包括识别覆盖网络的段的标识符。发现分组可以是思科发现协议(CDP)分组、链路层发现协议(LLDP)分组等。

[0012] 覆盖网络可以是虚拟可扩展局域网(VXLAN);网络虚拟化通用路由封装(NV-GRE)网络;或媒体访问控制(MAC)-in-MAC网络。覆盖网络可以是包括由对应的VXLAN段标识符(VNI)来识别的一个或多个段的多租户VXLAN。在更具体的实施例中,该方法可以包括将响应聚合为多点信息;以及显示所述多点信息。

[0013] 示例实施例

[0014] 转向图1,图1是图示用于在网络环境中发现多点端点的系统10的简化框图。在特定实施方式中,图1的架构被配置成提供在用于为网络问题的后续故障诊断的组播覆盖环境中发现端点的框架。图1包括将一个或多个交换机14连接到多个服务器16a、16b以及16c的覆盖网络12。术语“覆盖网络”简单地指的是被提供在另一网络之上的计算机或通信网络,诸如提供在第3层网络之上的第2层覆盖方案。这样的覆盖网络的示例能够包括(但不限于)虚拟可扩展局域网(VXLAN)、网络虚拟化通用路由封装(NV-GRE)、媒体访问控制(MAC)-in-MAC等。覆盖网络12可以支持一个或多个组播网络,其提供封装方案以致力于在存在虚拟机(VM)的情况下(例如,在多租户环境中)的第2层和第3层数据中心网络基础设施。在特定实施方式中,服务器16a、16b以及16c可以分别提供有虚拟以太网模块(VEM) 17a、17b以及17c。VEM 17a、17b以及17c分别将高级联网能力提供给一个或多个VM 18、20以及22。

[0015] 交换机14可以提供有虚拟管理机模块(VSM) 24,其将一个或多个VEM(例如,VEM 17a-c)控制为一个模块化交换机。VSM 24被配置成支持分别在物理服务器16a-c内部的软件中所提供的多个VEM 17a-c。VEM 17a-c的配置可以通过VSM 24来提供并且自动地传播到VEM。VEM 17a、17b以及17c利用来自VSM 24的配置信息,执行第2层交换,并且提供高级联网功能,诸如端口通道、服务质量(QoS)、安全(例如,专用虚拟局域网(VLAN)、端口安全等)以及监控(例如,网络流、交换机端口分析器(SPAN)、封装式远程SPAN等)。网络管理员能够根据耦合到VSM 24的单个接口来定义关于所有VEM 17a-c的配置。VSM 24可以提供有调试引擎26,所述调试引擎26被配置成检查覆盖网络12上的链路连通性。在图1中所图示的实施例中,VSM 24可以与服务器28集成在一起,服务器28提供单个控制台来操作并且管理VSM 24。

[0016] 注意,分配给图1的元件的数字和字母符号表示不暗示任何类型的层次;符号表示是任意的并且已被仅用于教导的目的。这样的符号表示不应该被以任何方式解释成限制它们在可能受益于系统10的特征的潜在环境中的能力、功能性或应用。为了便于描述,仅两个代表性服务器场被图示。替换地,可以在覆盖网络12中连接任何数目的服务器场和交换机。

[0017] 出于图示系统10的技术和能力的目的,重要的是理解给定系统中的通信。以下基本信息可以被视为本公开内容可以根据其被适当地说明的基础。这样的信息被真实地提供仅用于说明的目的,并且因此,不应该被以任何方式解释成限制本公开内容及其潜在应用的广泛范围。在组播网络中,若干节点(还被称为主机或端点)可以通过虚拟或物理网络中的网际协议(IP)基础设施来互连以使得能实现一对多和多对多网络递送服务。“节点”可以是任何电子设备(例如,机器设备或移动设备)、客户端、服务器、对等体、服务、应用、VEM、VM,或能够在网络中通过通信通道来发送、接收或者转发信息的其它对象。如本文所用的“端点”包含在网络流中发起数据分组的任何节点,和/或作为网络流中的数据分组的目的地任何节点。因此,端点能够与客户端、客户、手持式设备或希望经由一些网络在系统10中启动通信的终端用户相关联。术语“端点”包括用来启动通信的设备,诸如接收机、计算机、路由器、网关、网络装置、专有设备、机顶盒、因特网无线电设备(IRD)、手机、任何种类的智能电话、平板电脑、个人数字助理(PDA)、iPhone、iPad、谷歌Droid或任何其它设备、组件、元件,或能够在系统10内启动语音、音频、视频、媒体或数据交换的对象。端点还可以包括适于人类用户的接口,诸如显示器、键盘、触摸板、远程控制或其它终端设备。端点还可以是寻求代表另一实体或元件来启动通信的任何设备,诸如程序、数据库或任何其它组件、设备、元件,或能够在系统10内启动交换的对象。如在本文档中本文所用的那样,数据指的是任何

类型的数字、语音、视频、媒体或脚本数据,或任何类型的源代码或目标代码,或可以被从一个点传递到另一点的具有任何适当的格式的任何其它适合的信息。

[0018] 端点还可以包括跨越组播骨干(例如,包括由物理或虚拟通信链路所连接的主机、路由器、交换机(聚合交换机、接入交换机、核心交换机等)以及其它网络设备的网络的虚拟网络骨干)连接到彼此的VEM。在这种意义上,组播骨干包含任何这样的骨干可能性,其能够用作网络中的端点和各种其它网络元件的链路。组播骨干中的端点和网络元件典型地用组播协议来启用以用于数据同时地到多个接收机的高效递送。

[0019] 组播能够为诸如视频会议之类的应用并且为应该被同时地递送到数个其它主机的音频提供一对多和多对多网络递送服务。如本文所用的那样,“组播”指的是(例如,同时地在单个传输中)将消息(例如,数据分组)从发送者递送到一组接收者。如本文所用的那样,术语“发送者”包含将数据分组发送到接受者的节点,以及术语“接收者”包含从发送者接收数据分组的节点。组播通过不需要许多接收者如何存在的现有知识按大的接收者群体缩放。即使数据分组将被递送到大量接收者,组播也通过从发送者发送数据分组仅一次来高效地使用网络基础设施。

[0020] 在组播网络中,发送者通常不知道谁接收数据分组。发送者仅仅将数据分组发送到组播地址(例如,224.0.0.0、239.255.255.255、01:00:0c:cc:cc:cc、01:80:c2:00:00:0e、01:80x2:00:00:03、01:80:c2:00:00:00等)。组播地址典型地与一组感兴趣的接收者相关联。组播骨干中的中间路由器做出数据分组的拷贝并且将它们发送到已加入组播组的适当的接收者。当主机希望加入具有特定组播地址的组播组时,主机发出因特网组管理协议(IGMP)请求。然后用于该子网的组播路由器能够通知其它路由器,使得到所述组的组播数据分组可以最终被放置在该主机的局域网(LAN)上。

[0021] 另外,服务器虚拟化使传统的(或物理的)服务器访问网络设计的数个假设无效。例如,一个这样的假设是每个网络访问端口都对应于运行单个映像的单个物理服务器,并且映像与网络之间的关系是静态的。服务器虚拟化高效地使得OS映像能够变得移动。这种级别的移动性对网络的结果不是无价值的,并且它们的影响可以延伸超越仅仅接入层。

[0022] 在虚拟化使得多个OS能够共享单个物理服务器的情况下,可能期望或者甚至有必要对于虚拟化服务器重复传统的安全和分割,因为虚拟化可以创建含糊的且不确定的安全边界。虚拟化网络可能不得不支持诸如商业或租户的多个线路之类的许多特征:其中的每一个都需要受保护且信任的虚拟计算环境。其它特征能够包括工作负荷弹性,其中虚拟服务器被实例化成解决增加的工作负荷并且周期性地停止或者去除以支持数据中心效率和绿色计算。进一步其它的特征能够解决对于高可用性和可伸缩性的升高的要求。

[0023] 在虚拟化环境中,组播网络中的端点(例如,主机)能够被一般地分布在网络上。存在这样的技术,其中相同的L2域中的端点能够使用覆盖网络横跨数据中心。像通用路由封装(GRE)和最近的VXLAN这样的一些技术将组播用作覆盖来展开第2层(L2)域。将组播用作覆盖,能够潜在地创建许多隧道,其中端点可以是不可知论者。在这样的环境中识别实际的端点可能是一种挑战。此外,网络中的故障诊断问题(由于组播问题而导致)可能是有问题的。

[0024] 现今存在的组播查验和mtrace协议发现组播网络中的端点。例如,终端主机能够查验组播组,并且该组播组的成员可以对查验回应。然而,这种方法无法提供应用级可见

性,因为相同的组播组能够被用于像GRE、VXLAN这样的不同应用或通用组播业务,因为组播空间是有限的。此外,如果仅属于特定GRE密钥/VXLAN网络标识符(VNI)等的终端主机将被识别则这样的方法无法工作。

[0025] 关于覆盖网络,VXLAN实施方式在此在本说明书中致力于讨论系统10的一些能力。[迫切要注意的是,虽然本公开内容的实施例可以被应用于任何种类的覆盖网络(例如,NV-GRE、MAC-in-MAC等),但是这样的讨论不应该被以任何方式解释成限制本公开内容及其潜在应用的广泛范围。]VXLAN是可以被用来将第2层网络覆盖在第3层网络上的隧道方案。每个覆盖都是VXLAN段。在相同的VXLAN段内的VM能够与彼此进行通信。每个VXLAN段都可以由其VNI来识别,所述VNI可以是24位段ID。VXLAN隧道的端点(被称作虚拟隧道端点或VTEP)可以包括在包括VM的服务器上的VEM。因此,VNI和VXLAN相关的隧道仅为VTEP(而不为诸如VM之类的其它实体)所知。

[0026] VTEP一般地旨在处于网络的边缘,典型地将接入交换机(虚拟的或物理的)连接到IP传输网络。预期VTEP功能性将被内置到接入交换机中,但它逻辑上与接入交换机分离。VTEP可以表现为用于它所被连接到的IP网络的IP主机。IP网络可以包括组播骨干。每个VTEP都可以基于其IP接口所被连接到的子网而配置有IP地址。VTEP能够使用这种IP接口来与其它VTEP交换携带封装的以太网帧的数据分组。VTEP还可以通过使用IGMP来加入IP组播组而作为IP主机。

[0027] VXLAN典型地通过一个或多个服务器而被部署在虚拟化主机(例如,VM)上的数据中心中。服务器可以是相同的(例如,第2层)网络或不同的(例如,第3层)网络的一部分。VXLAN段被重叠在第2层或第3层网络之上。VM不知道覆盖网络和传输方法,因为封装和解封装在相应服务器上的VTEP处透明地发生。VXLAN可以被部署在其中所有主机理解VXLAN协议的网络上。替换地,VXLAN可以被部署在其中仅主机中的一些理解VXLAN协议的网络上。例如,在VXLAN覆盖网络上的节点可能需要与可能是基于VLAN的旧网络上的节点进行通信。为了使得能实现通信,网络能够包括VXLAN网关,所述VXLAN网关转发VXLAN和非VXLAN环境之间的业务。配置网关从传入帧剥离VXLAN报头并且基于内以太网帧的目的地MAC地址将它们转发到物理端口。在相反方向上,针对非VXLAN接口的传入帧基于帧中的VLAN ID被映射到特定VXLAN覆盖网络。

[0028] 采用VXLAN,当VM发送帧时,它连同VNI一起被封装在用户数据报协议(UDP)数据分组中,所述VNI唯一地识别VXLAN段。传统的IEEE 802.1Q交换在该段内发生。代替像在未知单播的情况下那样广播帧,UDP数据分组被组播到在相同的VXLAN段上具有VM的VEM组。当在现有VXLAN上提出新的VM时,从VM的相应VEM发送针对VXLAN组播组的IGMP加入。组播组地址被用来封装来自VXLAN VM的广播、组播以及未知单播业务。包括VXLAN VNI的报头被插入在组播数据分组开头。组播数据分组被发出(即,被组播)到VXLAN在其上被实现的组播组。接收者端点使用IP单播来发送标准地址解析协议(ARP)响应。在VEM已学习到虚拟机针对其所被定位的IP地址之后,数据分组能够在没有组播支持的情况下被直接地发送到目的地VEM。

[0029] 在云环境中,VXLAN能够横跨多个端点并且每个VEM能够属于多个VXLAN。VXLAN能够依赖于物理基础设施的组播骨干。因为VXLAN能够在集群内并且横跨集群,所以这种环境中的任何故障可以影响VXLAN的网络基础设施,从而使调试/故障诊断变得困难。例如,转向图1的覆盖网络12,假定VM 18、20以及22是在具有VNI为5000的相同VXLAN上并且属于具有

地址225.1.1.1的组播组。可能存在其中VM 18和VM 20正在与彼此很好地进行通信、但是VM 18和VM 22不能够交换通信的情形。网络管理员没有关于问题是在相应的VM、交换机14还是上游设备(例如物理基础设施的物理网络交换机和路由器)中的线索。在另一情况下,单播业务可能在VM 17a-c之间良好,但组播业务可能被中断。这样的场景费力地难以调试,其中这样的调试活动性质上是乏味的。附加地,在组播查验中,仅传送路径是组播数据分组,然而返回路径是来自端点的单播。因此,不能够保证双向组播路径验证。

[0030] 系统10被配置成在提供用于在网络环境中发现多点端点的系统和方法时解决这些问题(和其它问题)。根据各种实施例,调试引擎26能够针对在覆盖网络12中有VXLAN能力的虚拟接口(例如,虚拟机内核网络接口卡(VMKNIC))上的每个VXLAN启用发现分组。如本文所用的那样,术语“发现分组”包含任何适合的链路层(例如,在第2层中的数据链路层)网络协议数据分组。能够可选地以类型长度值(TLV)帧格式提供发现分组,所述类型长度值帧格式可以被用来共享关于网络中的节点的信息(例如,身份、IP地址、能力等)。因此,发现分组可以与诸如思科发现协议(CDP)、链路层发现协议(LLDP)、极端发现协议(EDP)、网捷发现协议(FDP)、北电网络发现协议(NDP)、SynOptics网络管理协议(SONMP)等协议相关联。在不背离本公开内容的广泛范围的情况下,可以在系统10中使用任何链路层发现协议。为了便于说明,并且不作为限制,CDP协议将在本文中被描述来说明本公开内容的特定能力和特征。

[0031] CDP是被典型地用来共享与其它直连节点相关联的信息(例如,操作系统版本和IP地址)的数据链路层网络协议。支持CDP的网络设备将CDP通告发送到组播目的地址01:00:0c:cc:cc:cc。这些组播数据分组可以由其它直连节点(其支持CDP)在它们连接的网络接口处接收。支持CDP的每个节点都将从其它节点接收到的信息存储在表中,所述表能够使用诸如show CDP neighbors命令之类的适当的命令来查看。包含在CDP通告中的信息可以包括操作系统版本、主机名、来自配置于CDP帧被发送的端口上的(一个或多个)所有协议的IP地址、通告从其被发送的端口标识符、设备类型和型号、双工设定、VTP域、本机VALN、功耗(针对以太网供电设备)以及其它设备特定信息。

[0032] 典型地,CDP分组能够被用来获得一般网络(不必实施VXLAN技术或等效覆盖方案)中的邻近设备的协议地址。CDP不能够被用来提供关于未被直接地连接的设备的信息,因为CDP分组是第2层数据分组。路由器在默认情况下一般地不转发第2层业务;交换机将由CDP所用的MAC组播地址识别为特殊地址并且将不从其它接口转发分组。因此,仅直连邻居接收CDP分组。然而,系统10的实施例能够被配置成使用CDP分组(或具有其它链路层格式的发现分组)以便获得关于无法通过诸如VXLAN之类的覆盖网络直接地连接的端点的信息。

[0033] 本公开内容的详细资料提供在使用适当地配置的发现分组来实现覆盖方案的网络环境中跨越组播骨干发现端点。因为VXLAN(和其它类似的覆盖方案)在每个VXLAN段中呈现第2层网络(而不管端点是否被直接地连接),所以可以在VXLAN段中的节点之间交换发现分组(诸如CDP分组)以排除连通性故障。

[0034] 根据本公开内容的实施例,在接收到发现分组之后,每个VEM 17a-c都可以验证发现分组上的TLV参数。每个VEM 17a-c都可以将发送者识别为邻近设备,并且适当地响应。发现分组中的应用TLV能够被用来至少交换远程VEM的VNI和IP地址(例如,VMKNIC IP)以便验证点对点连通性。命令行接口(CLI)能够为调试提供过滤器(例如,每VXLAN)。由系统10所提供的操作和基础设施能够适用于诸如VXLAN、NV-GRE以及MAC-in-MAC之类的封装解决方案,

以及支持这样的封装解决方案的虚拟或物理设备。系统10的实施例能够使得网络管理员能够清楚地使上游(例如,组播骨干)问题与虚拟交换机或VM问题隔离。系统10的实施例能够使得易于跨越多个分布式虚拟交换机(DVS)在大数据中心部署中故障诊断并且提供用来发现VXLAN段的主机的机制。

[0035] 系统10的实施例还能够在隧道环境中例如在任何组播覆盖网络中发现端点。在数据中心环境(诸如正由覆盖网络12所提供的数据中心环境)中,网络管理员能够提供CLI以从VSM 24触发发现。VSM 24可以引导VEM 17a-c发送发现分组。端点(例如,VEM 17a)能够将发现分组(其可以使用组播覆盖来封装)发送到覆盖网络12中的所有其它端点(例如,VEM 17b和17c)。其它端点(例如,VEM 17b和17c)从而能够学习到源端点(例如,VEM 17a)的存在。这种信息能够被推送给VSM 24,其能够聚合来自所有端点的信息并且将所发现的端点提供(例如,显示)给网络管理员。

[0036] 如果在VXLAN段中存在连通性问题,则该问题可能位于在下列中的一个或多个处:(1)一个或多个VM;(2)在VM与VEM之间;以及(3)在VEM之间。使用发现分组可以使得网络管理员能够高效地(例如,使用CLI上的一个或多个命令)确定连通性问题是否是在VEM之间,并且如果是这样的话,则确定哪些VEM受该连通性问题影响。

[0037] 而且,使用发现分组的优点(在TLV被启用情况下)可以是应用特定TLV可以被插入到发现分组中。例如,端点能够被每应用隔离。应用TLV能够携带特定GRE密钥/VNI(或任何其它标识符),并且仅可以发现携带这种特定识别信息的那些端点。进一步地,各种实施例可以将组播MAC地址用作目的地MAC地址,从而验证双向组播路径。能够在VSM 24处使用适合的显示机制查看所发现的多点信息。

[0038] 应用特定TLV还能够被用来故障诊断覆盖网络12中的问题,其中上游连通性(例如,配置)中的问题可以由这种方法容易地识别。例如,服务器管理员可以识别两个VM(例如,一个为VM 18,一个为VM 20)不在与彼此对话,并且可以相应地通知网络管理员。网络管理员能够识别VM 18和20被托管在其上的VEM(例如,VEM 17a和17b),并且然后验证上游组播连通性的工作条件。两个端点之间的路由能够通过将端点标识符(例如,在VXLAN的情况下为VMKNIC IP地址)使得仅该特定端点应答(例如,VEM 17b或17a)来验证。网络管理员因此能够容易地推断出问题是否是因为上游连通性。

[0039] 转向系统10的基础设施,网络拓扑能够包括任何数目的服务器、VM、DVS、虚拟路由器、VSM以及被互连以形成大且复杂的虚拟覆盖网络12的其它节点。更具体地,调试引擎26、服务器16a-c、VEM 17a-c、交换机14、VSM 24和/或服务器28可以被配置成执行本文中所述讨论的特定活动(或者与彼此协作)。此外,这些设备是能够促进本文中所讨论的许多处理、通信和/或发现活动的网络元件。如本说明书中所用的那样,术语“网络元件”旨在包含路由器、服务器、交换机、电缆盒、网关、桥接器、负载均衡器、防火墙、联机服务节点、代理、处理器、模块或任何其它适合的设备、组件、元件、专有装置、用户设备,或可操作来在网络环境中交换信息的对象。这些网络元件可以包括任何适合的硬件、软件、组件、模块、接口,或促进其操作的对象。这可以包括允许数据或信息的有效交换实现的适当的算法和通信协议。

[0040] 图1的元件可以通过采用任何适合的连接(有线或无线)的一个或多个接口被耦合到彼此,所述连接为电子通信提供可行的路径。附加地,这些元件中的任何一个或多个可以基于特定配置需要被组合或者从架构去除。系统10可以包括有能够用于数据分组在网络中

的电子传输或接收的传输控制协议/网际协议 (TCP/IP) 通信的配置。系统10还可以在适当情况下并且基于特定需要与用户数据报协议/IP (UDP/IP) 或任何其它适合的协议相结合地操作。此外,网关、路由器、交换机以及任何其它适合的节点(物理的或虚拟的)可以被用来促进网络中的各种节点之间的电子通信。

[0041] 示例网络环境可以被配置在可以包括一个或多个网络的物理基础设施之上,并且进一步地,可以以任何形式加以配置,所述任何形式包括但不限于局域网(LAN)、无线局域网(WLAN)、VLAN、城域网(MAN)、广域网(WAN)、VPN、内部网、外联网、任何其他适当的架构或系统,或促进网络中的通信的其任何组合。在一些实施例中,通信链路可以表示支持LAN环境的任何电子链路,诸如例如电缆、以太网、无线技术(例如,IEEE 802.11x)、ATM、光纤等或其任何适合的组合。在其它实施例中,通信链路可以表示通过任何适当的介质(例如,数字订户线路(DSL)、电话线路、T1线路、T3线路、无线、卫星、光纤、电缆、以太网等或其任何组合)和/或通过诸如广域网(例如,因特网)之类的任何附加的网络的远程连接。

[0042] 转向图2,图2是图示根据本公开的实施例的系统10的附加细节的简化框图。调试引擎26可以包括消息生成器40、发现模块42、输出生成器44、处理器46以及存储器48。调试引擎26可以通过覆盖网络12被连接到组播骨干50。组播骨干50互连托管‘n’个VM 70-82的‘m’个VEM(还被称为端点)60-66。组播骨干50可以包括交换机、路由器、链路以及促进覆盖网络12中的各种节点之间的通信的其它网络设备。每个VEM都可以包括一个或多个VM。例如,VEM 60包括VM 70和72。VEM 62包括VM 74和76。VEM 64包括VM 78和80。VEM 66包括VM 82。一个或多个VM可以是能够由对应的VNI来识别的VXLAN段的一部分。例如,VM 70、72以及74是VXLAN段90的一部分,并且VM 76、78、80以及82是VXLAN段92的一部分。调试引擎26可以被连接到管理机94,所述管理机可以通过诸如显示器96上的CLI之类的适合的用户接口来控制。在一个实施例中,管理机94可以是VSM 24的一部分。

[0043] 出于讨论的目的,做出了VEM 60中的VM 70不正在与VEM 62中的VM 74交互的假设。缺少连通性的原因可以包括组播骨干50(例如,在VEM 60与VEM 74之间的通信链路)的问题,或个别地VM 70或VM 74的问题,或在VM与VEM之间(例如,在VM 60与VEM 60之间或在VM 74与VEM 62之间)的问题。根据本公开的实施例,调试引擎26可以使发现分组被发送到适当的端点,所述适当的端点在这个示例场景中包括VEM 60和VEM 62使得连通性问题可以被解析(或者至少变窄)。

[0044] 根据各种实施例,调试引擎26可以在覆盖网络12中组播发现分组。发现分组可以被配置成发现跨越组播骨干50连接的并且属于覆盖网络12的一段(例如,具有对应VNI的VXLAN段)的端点。消息生成器40可以生成包括适当的TLV信息和适当的标识符的发现分组,所述适当的标识符诸如与VM 70和74所属于的VXLAN段90相对应的VNI。在一些实施例中,可以以周期性时间间隔规则地(例如,例如每五分钟)生成发现分组。在其它实施例中,发现分组可以由用户视需要或情况而定来生成。自动轮询可以使网络泛洪并且对于VEM(例如,VEM 60-66)和管理机94引入不必要的处理。在用户触发的选项中,用户出于诸如调试网络之类的特定目的而触发发现分组。发现模块42可以将发现分组组播到覆盖网络12中属于适当的VXLAN段90的端点。在本文中所描述的示例中这样的端点是VEM 60和62。在一个实施例中,覆盖网络12是跨越第3层网络的第2层覆盖方案,发现分组是第2层数据分组,并且分发的方法是组播。因此,端点可以对第2层发现分组做出响应,而不管它们被直接地还是间接地连

接到调试引擎26。

[0045] 发现模块42可以等待来自各种VEM (例如,VEM 60和62) 的响应或超时并且聚合所接收到的信息 (还在此被称为发现的多点信息)。如果连通性问题是组播骨干50处,则取决于哪一个链路具有连通性问题VEM 60或62 (或两者) 可能未能对发现分组做出响应。因此,响应可以超时,并且发现模块42可以记录这样的超时。视情况而定,输出生成器44可以引起所发现的多点信息在显示器96上的显示。在其中连通性问题是组播骨干50 (例如,将VEM 64与组播骨干50连接的链路) 处的情况下,输出可以指示VEM 62没有正在响应。

[0046] 在另一示例场景中,网络管理员可能想要发现VXLAN段 (例如,VXLAN段92) 中的所有VEM。网络管理员可以使调试引擎26将发现分组组播到VXLAN段92中的端点 (因为仍然未知)。消息生成器40生成包括适当的TLV信息和适当的标识符 (例如,与VXLAN段92相对应的VNI或GRE密钥) 的发现分组。发现模块42可以将发现分组组播到覆盖网络12中属于VXLAN段92的所有端点。VEM 64-66可以适当地对发现分组做出响应。发现模块42可以聚合所发现的多点信息。所发现的多点信息可以包括做出响应的所有VEM的VEM识别 (诸如IP地址)。输出生成器44可以引起所发现的多点信息在显示器96上的显示。因此,可以基于对来自相应端点的发现分组的响应来识别有关端点。

[0047] 在各种实施例中,调试引擎26可以位于管理机94上,所述管理机和与VEM 60-66分离的专用服务器 (例如,服务器28) 和虚拟交换机相关联。在其它实施例中,调试引擎26可以位于与覆盖网络12中的一个或多个服务器 (例如,收容VEM 60-66的服务器) 相关联的虚拟交换机上。例如,调试引擎26可以位于收容VEM 60的相同服务器上。在又一个示例实施例中,调试引擎26可以位于与VSM 24相关联的虚拟服务装置上。调试引擎26的各种其它配置在覆盖网络12的虚拟化环境中是可能的。注意在各种实施例中,处理器46和存储器48可以是物理上与调试引擎26的其它组件分离的网络元件 (例如,消息生成器40、发现模块42等) 的一部分。在各种实施例中,调试引擎26的组件可以被一起耦合在一个物理设备上,或者它们可以基于特定需要位于单独的物理设备上并且以任何适合的组合逻辑上互连到调试引擎26中。

[0048] 转向图3,图3是图示可以与本公开的实施例相关联的操作的简化流程图100。当调试引擎26被激活时,操作可以在102处开始。在104处,多点端点消息由消息生成器40来生成。在106处,消息例如通过插入适合的TLV信息使得发现分组被创建而被适当地封装。在108处,可以通过发现模块42在组播覆盖之上发送具有信息密钥 (例如,VNI等) 的已封装消息。在110处,发现模块42可以等待响应 (或替换地,来自接收者的超时)。在112处,发现模块42可以聚合所接收到的信息并且输出生成器40可以引起所发现的多点信息在显示器96上的显示。当端点已被发现时,过程在116处结束。

[0049] 转向图4,图4是图示根据本公开的实施例的示例输出屏幕200。输出屏幕200可以显示由发现模块42所接收到的多点信息。多点信息可以包括VEM (例如,由VEM名称或适合的ID来识别)、VEM IP地址等的列表。能够获得 (和显示) 的多点信息能够包括任何设备、系统或用来识别特定网络中的VEM的网络信息。多点信息可以由输出生成器44基于CLI上的适合的命令自动地显示,或者由事件触发器 (诸如与VXLAN段的VEM断开) 造成。

[0050] 例如,诸如“`cinquedia (config) #show cdp neighbors`”之类的适合的命令可以提出属于覆盖网络12的VEM的列表。输出屏幕200能够指示四个VEM (例如,VEM1、VEM2、VEM3以

及VEM4)在VXLAN中被连接。诸如“cinquedia(config)#show cdp neighbors vxlan 5000”之类的另一适合的命令可以提出属于相同VXLAN段(例如,由所对应的VNI 5000来识别)的VEM的列表。输出屏幕200指示VEM 1和VEM 3属于VXLAN段5000。诸如“cinquedia(config)#show cdp neighbors verify 172.23.231.100 172.23.231.200”之类的又一个适合的命令可以验证两个VEM(例如,在IP地址172.23.231.100和172.23.231.200处)之间的点对点组播连通性。输出屏幕200能够指示虽然从发送者到接收者的连接不具有问题,但是返回连接已失败。注意,本文中所列举的多点信息的命令和集合仅用于示例目的。可以基于特定主机设备、管理机软件或功能性、软件开发套件(SDK)、版本等来配置命令。本文中所列举的多点信息的命令和集合不旨在以任何方式限于本公开内容的广泛范围。

[0051] 转向图5,图5是图示可以与本公开的实施例相关联的示例操作步骤的简化流程图300。当调试引擎26被激活时,操作可以在步骤302处开始。在304处,可以识别具有连通性问题的VM。在一个实施例中,网络管理员可以识别VM。在另一实施例中,VM可以由相应的VEM来自动地识别。在306处,可以确定受影响的VM的VXLAN段。在各种实施例中,每个VM都可以属于多个VXLAN段。在308处,受影响的VXLAN段可以被选择成被测试。在一个实施例中,选择可以是随机的。在另一实施例中,选择可以是基于网络策略、企业策略或可能是基于特定需要的其它适合的准则。

[0052] 在310处,发现分组可以被调试引擎26发送到所选VXLAN段中的VEM。在312处,发现模块26可以等待响应或超时。在314处,调试引擎26可以聚合来自VEM的响应。在316处,调试引擎26可以视需要而定显示所发现的多点信息。在一个实施例中,多点信息可以被显示在屏幕上。在另一实施例中,多点信息可以被插入到报告中。在又一个实施例中,可以打印多点信息。在318处,可以做出VXLAN段中的所有VEM是否已经响应并且它们是否已被包括在多点信息中的确定。如果VEM中的至少一些具有连通性问题,则这样的VEM将未被包括在多点信息中。在320处,可以识别这样断开的VEM并且过程在322处结束,其中可以执行后续的集中式故障诊断。

[0053] 替换地,如果VXLAN段中的所有VEM都被包括在多点信息中,如在318处所确定的那样,则在324处做出所有有关的VXLAN段是否已被测试的后续确定。如果所有VXLAN段已被测试,则它可以指示在组播骨干中没有连通性问题,并且连通性问题可以与各个VM相关联,或者在受影响的VM与相应的VEM之间。可以在322处结束过程。注意,如果所有有关的VXLAN段尚未被测试,则操作可以进行到步骤326,其中下一个VXLAN段可以被选取用于测试。随后,可以针对在306处所识别的每个有关的VXLAN段重复步骤310至320。

[0054] 在示例实施方式中,本文中所概述的活动中的至少一些部分可以用在例如调试引擎26、服务器16a-c、VEM 17a-c、交换机14、VSM 24和/或服务器28中所提供的软件加以实现。在一些实施例中,所讨论的特征中的一个或多个可以用在这些元件外部提供的硬件加以实现,或者以任何适当的方式合并来实现所预定的功能性。另外地,调试引擎26、服务器16a-c、VEM 17a-c、交换机14、VSM 24和/或服务器28可以包括能够协调以便实现如本文中所概述的操作的软件(或往复软件)。在仍然其它的实施例中,这些元件可以包括任何适合的算法、硬件、软件、组件、模块、接口或促进其操作的对象。停止

[0055] 此外,本文中所描述并示出的调试引擎26、服务器16a-c、VEM 17a-c、交换机14、VSM 24和/或服务器28(和/或它们关联的结构)还可以包括用于在网络环境中接收、传送

和/或以其它方式传递数据或信息的适合的接口。另外地,与各种节点相关联的处理器和存储器中的一些可以被去除或者以其它方式合并,使得单个处理器和单个存储器位置负责特定活动。在一般意义上,图中所描绘的布置在它们的表示方面可能是更有逻辑的,然而物理架构可以包括这些元件的各种排列、组合和/或混合。必须注意的是,无数可能的设计配置能够被用来实现在这里所概述的操作目标。因此,所关联的基础设施具有无数的组成布置、设计选择、设备潜力、硬件配置、软件实施方式、设备选项等。

[0056] 在示例实施例中的一些中,一个或多个存储器(例如,存储器48)能够存储用于本文中所描述的操作的数据。这包括能够存储被执行来执行本说明书中所描述的活动的指令(例如,软件、逻辑、代码等)的存储器。处理器能够执行与数据相关联的任何类型的指令以实现在本说明书中此处所详述的操作。在一个示例中,处理器46能够将元件或物品(例如,数据)从一个状态或事物变换为另一状态或事物。在另一示例中,本文中所概述的活动可以用固定逻辑或可编程逻辑(例如,由处理器所执行的软件/计算机指令)来实现,并且本文中所识别的元件可以是某种类型的可编程处理器、可编程数字逻辑(例如,现场可编程门阵列(FPGA)、可擦可编程只读存储器(EPROM)、电可擦可编程只读存储器(EEPROM)、包括数字逻辑、软件、代码、电子指令的ASIC、闪存存储器、光盘、CD-ROM、DVD ROM、磁或光学卡、适合于存储电子指令的其它类型的机器可读介质,或其任何适合的组合。

[0057] 系统10中的组件中的任一个都能够包括一个或多个存储器(例如,存储器48)以用于存储在实现如本文中所概述的操作时要被使用的信息。这些设备可以进一步将信息保持在任何适合类型的非暂时性存储介质(例如,随机存取存储器(RAM)、只读存储器(ROM)、现场可编程门阵列(FPGA)、可擦可编程只读存储器(EPROM)、电可擦可编程ROM(EEPROM)等)、软件、固件中,或者在适当情况下并且基于特定需要保持在任何其它适合的组件、设备、元件或对象中。能够基于特定需要和实施方式在任何数据库、寄存器、表、高速缓存、队列、控件列表或存储结构中提供在系统10中被跟踪、发送、接收或存储的信息,其全部都能够任何适合的时间系中被参考。本文中所讨论的存储器术语中的任一个应该被解释为被包含在广义术语“存储器”内。类似地,本说明书中所描述的潜在的处理元件、模块以及机器中的任一个都应该被解释为被包含在广义术语“处理器”内。

[0058] 同样重要的是要注意,参考前面的图所描述的操作和步骤仅图示可以被或者在系统内执行的可能场景中的一些。可以在适当情况下删除或者去除这些操作中的一些,或者在不背离所讨论的构思的范围的情况下,可以大量地修改或者改变这些步骤。此外,这些操作的定时可以被大量地变更并且仍然实现本公开内容中所教导的结果。已经出于示例和讨论的目的提供了前面的操作流程。必要灵活性由系统以该任何适合的布置、年表、配置来提供,并且在不背离所讨论的构思的教导的情况下可以提供定时机制。

[0059] 尽管已经参考特定布置和配置详细地描述了本公开内容,但是在不背离本公开内容的范围的情况下,可以显著地改变这些示例配置和布置。例如,尽管已经参考牵涉特定网络接入和协议的特定通信交换对本公开内容进行了描述,但是系统10可以适用于其它交换或路由协议。而且,尽管已经参考促进通信过程的特定元件或操作图示了系统10,但是这些元件和操作可以用实现系统10的预定功能性的任何适合的架构或过程代替。

[0060] 注意在本说明书中,对包括在“一个实施例”、“示例实施例”、“实施例”、“另一实施例”、“一些实施例”、“各种实施例”、“其它实施例”、“替代实施例”等等中的各种特征(例如,

元件、结构、模块、组件、步骤、操作、特性等)的参考旨在意味着任何这样的特征被包括在本公开内容的一个或多个实施例中,但是可以或者可能未必被组合在相同的实施例中。此外,单词“优化”、“最佳化”以及相关术语是指的是规定的结果的速度和/或效率方面的改进的术语,并且不意味着指示用于实现所规定的结果的过程已实现了,或者能够实现“最佳的”或完全快速的/完全高效的状态。

[0061] 许多其它改变、替换、变化、变更以及修改可以被本领域的技术人员想到,并且意图是,本公开包含如落入所附权利要求的范围内的所有这样的改变、替换、变型、变更以及修改。为了帮助美国专利商标局 (USPTO) 并且另外地,帮助对本申请颁布的任何专利的任何读者解释所附权利要求,申请人希望注意本申请人:(a) 除非单词“用于...的装置”或“用于...的步骤”被具体地用在特定权利要求中,否则不打算所附权利要求中的任何一个援引美国法典第35条112节的第六(6)段因为它于其申请日存在;以及(b) 不通过本说明书中的任何声明旨在以在所附权利要求中未被以其它方式反映的任何方式限制本公开内容。

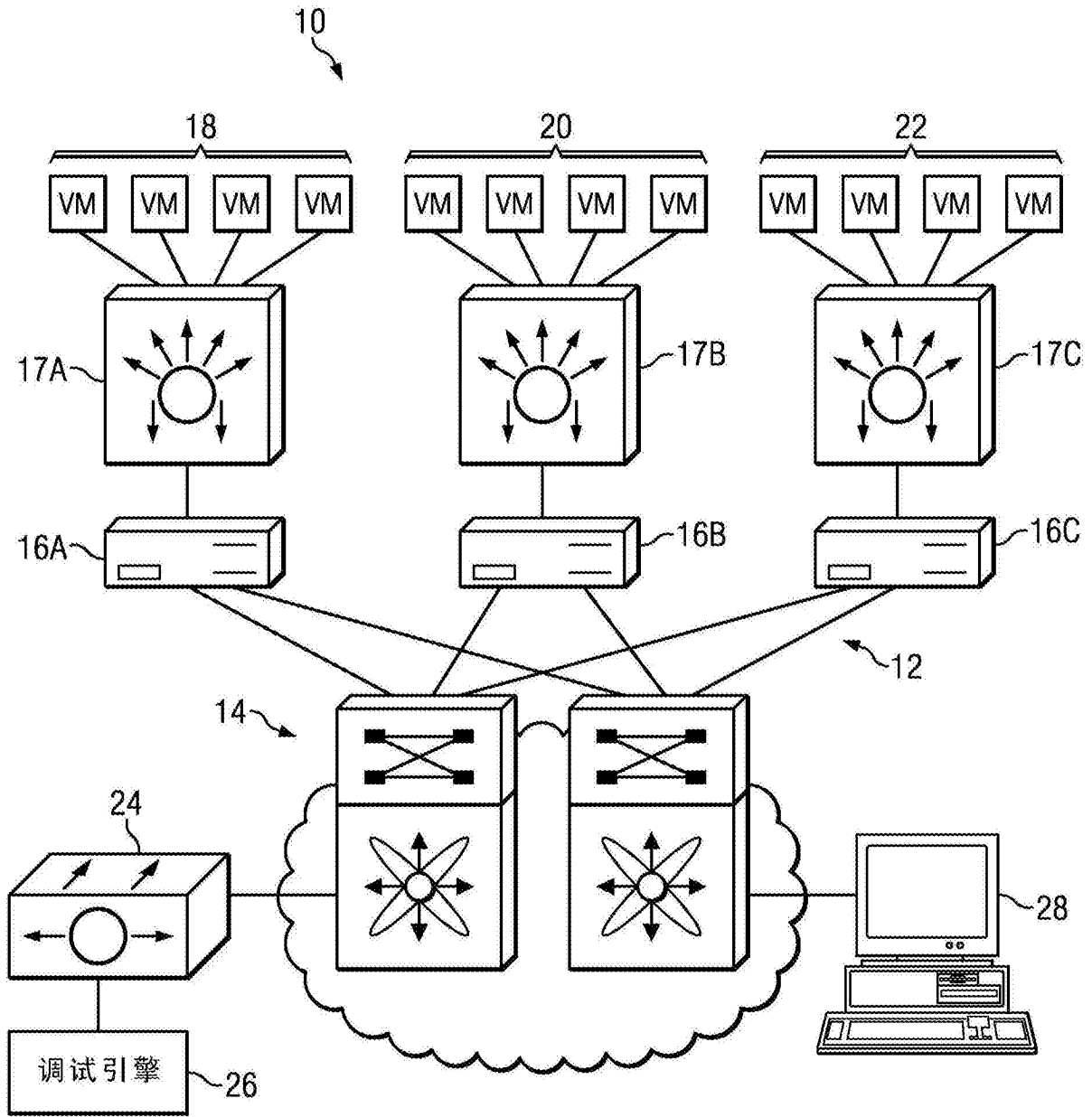


图1

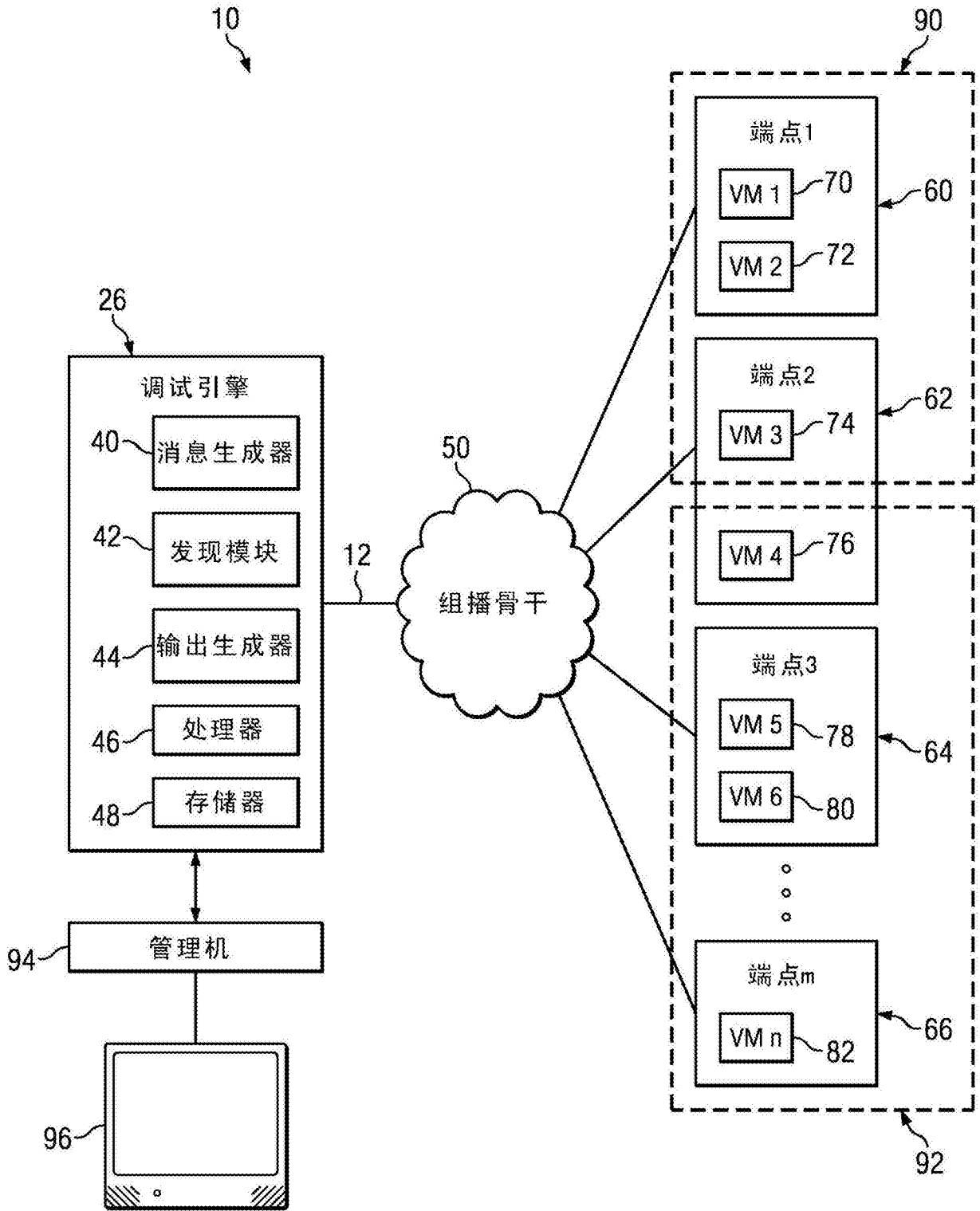


图2

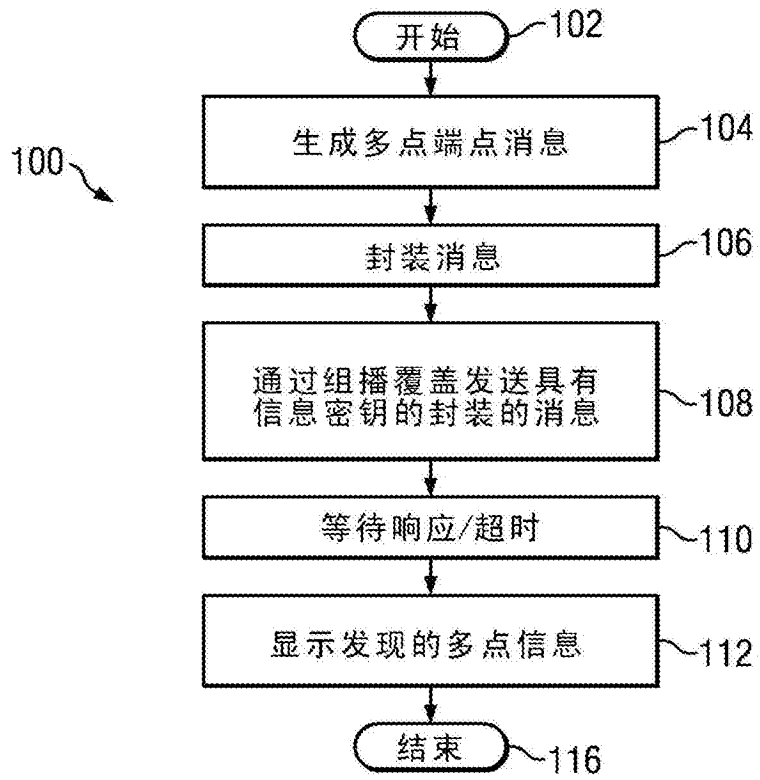


图3

200

```
cinquedia(config)# show cdp neighbors
```

Device-ID	Hldtme	Platform	Port ID
VEM-1 vmknic ip	123	N1k	Eth3/5
VEM-2 vmknic ip	123	N1k	Eth4/5
VEM-3 vmknic ip	123	N1k	Eth3/5
VEM-4 vmknic ip	123	N1k	Eth4/5

```
cinquedia(config)# show cdp neighbors vxlan 5000
```

Device-ID	Hldtme	Platform	Port ID
VEM-1 vmknic ip	123	N1k	Eth3/5
VEM-3 vmknic ip	123	N1k	Eth4/5

```
cinquedia(config)# show cdp neighbors verify 172.23.231.100 172.23.231.200
```

Device-ID (from)	Device-ID (to)	Reachable
172.23.231.100	172.23.231.200	Success
172.23.231.200	172.23.231.100	Failure

图4

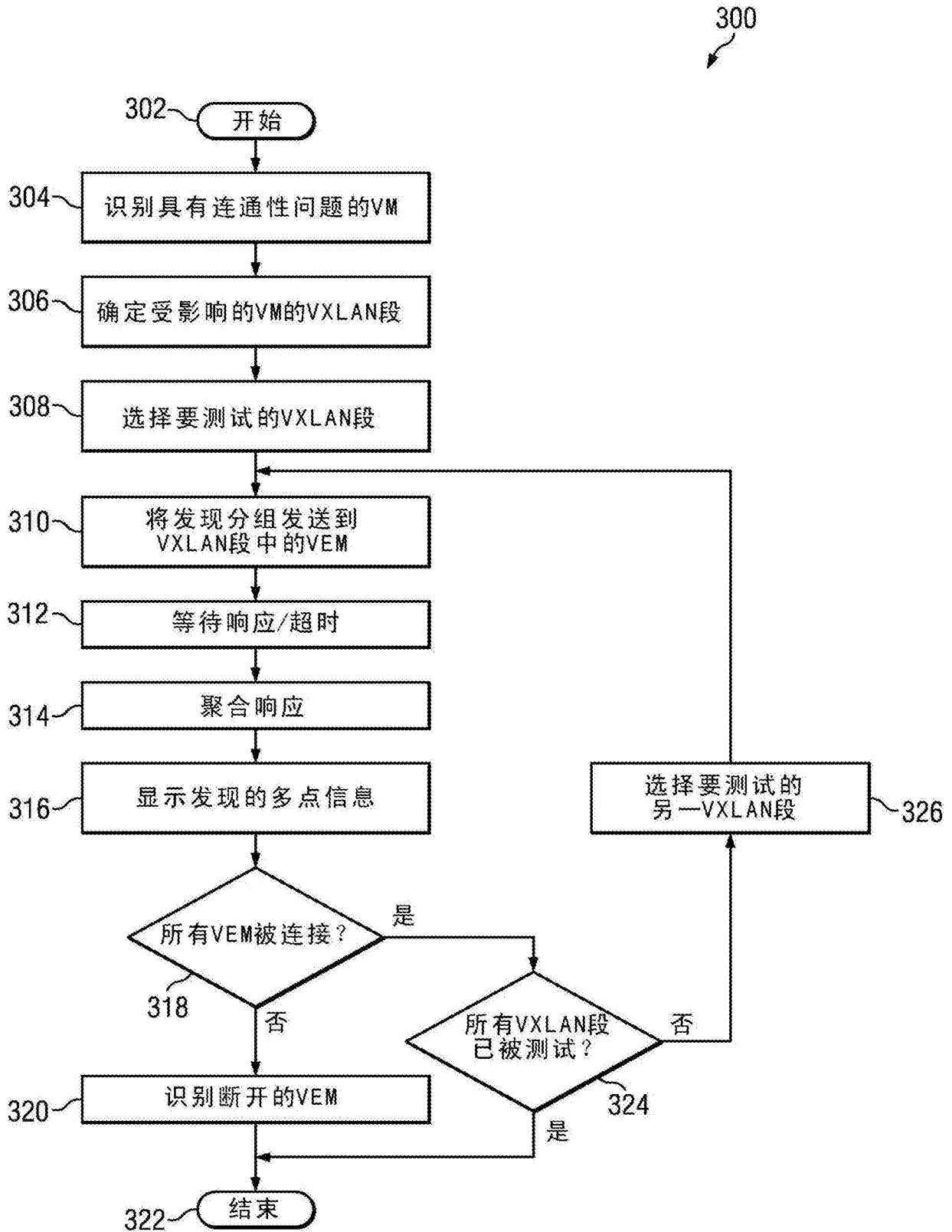


图5