



(12)发明专利

(10)授权公告号 CN 106464545 B

(45)授权公告日 2020.02.07

(21)申请号 201580023950.9

袁利华 张永光

(22)申请日 2015.04.10

(74)专利代理机构 北京市金杜律师事务所
11256

(65)同一申请的已公布的文献号

申请公布号 CN 106464545 A

代理人 王茂华

(43)申请公布日 2017.02.22

(51)Int.Cl.

H04L 12/26(2006.01)

(30)优先权数据

14/272,728 2014.05.08 US

(85)PCT国际申请进入国家阶段日

2016.11.07

(86)PCT国际申请的申请数据

PCT/US2015/025381 2015.04.10

(87)PCT国际申请的公布数据

W02015/171260 EN 2015.11.12

(73)专利权人 微软技术许可有限责任公司

地址 美国华盛顿州

(56)对比文件

US 2008/0080507 A1, 2008.04.03,

US 7937492 B1, 2011.05.03,

US 7937492 B1, 2011.05.03,

US 2010138885 A1, 2010.06.03,

US 7356031 B1, 2008.04.08,

US 2008080507 A1, 2008.04.03,

EP 2398188 A1, 2011.12.21,

US 2006203819 A1, 2006.09.14,

US 5488608 A, 1996.01.30,

审查员 高修会

(72)发明人 邬海涛 郭传雄 D·A·马尔兹

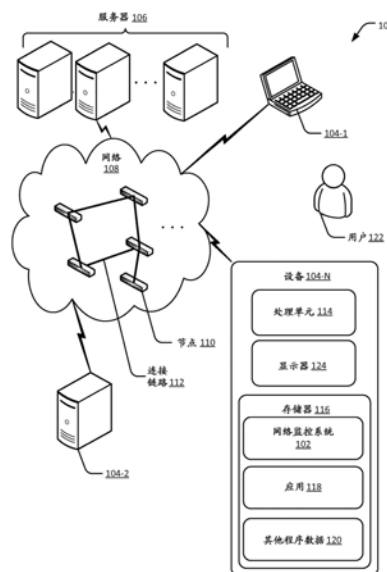
权利要求书2页 说明书12页 附图6页

(54)发明名称

一种细粒度网络监控方法及系统

(57)摘要

在一些情况下,一种网络监控系统可以通过根据隧道协议准备封装的数据分组来确定网络(例如,数据中心网络)中的节点或连接链路的操作状况或健康状况。取决于对封装的数据分组进行路由的结果,网络监控系统确定节点或连接链路是正在正常运行还是正在经历诸如过载或失灵的问题。



1. 一种用于网络管理的方法,包括:

在利用可执行指令配置的一个或多个处理单元的控制之下:

将测试数据分组传送到目的地节点,所述测试数据分组包括所述测试数据分组要被路由通过的多于一个特定节点的信息;以及

至少部分基于所述测试数据分组的所述传送的结果来确定所述多于一个特定节点的健康状况;

所述方法还包括至少将第一分组迭代地封装在第二分组中,并且至少将所述第二分组迭代地封装在第三分组中以形成所述测试数据分组。

2. 根据权利要求1所述的方法,其中所述封装还包括:将所述多于一个特定节点中的相应特定节点的地址信息包含在所述第一分组和所述第二分组的分组头部中,并且其中所述地址信息包括所述相应特定节点的全局地址或局部地址,所述全局地址对应于在包含所述多于一个特定节点和所述目的地节点的网络中可路由的地址,并且所述局部地址对应于由所述相应特定节点的一个或多个邻居可到达的地址。

3. 根据权利要求2所述的方法,其中传送所述测试数据分组包括:根据被包含在所述第三分组的分组头部中的目的地地址信息来将所述测试数据分组传送到第一中间节点,所述第一中间节点对所述第三分组进行解封装以获得所述第二分组,并且所述第一中间节点根据被包含在所述第二分组的分组头部中的目的地地址信息来将所述第二分组转发到第二中间节点,所述第二中间节点对所述第二分组进行解封装以获得所述第一分组,并且所述第二中间节点根据被包含在所述第一分组的分组头部中的目的地地址信息来将所述第一分组转发到另一节点。

4. 根据权利要求1所述的方法,其中所述特定节点中的至少一个特定节点的所述信息包括所述测试数据分组从发送所述测试数据分组的节点到所述特定节点中的所述至少一个特定节点遍历的跳的数目。

5. 根据权利要求1所述的方法,还包括:当所述测试数据分组的所述传送的所述结果对应于在预定时间阈值内在所述目的地节点处对所述测试数据分组的接收成功时,确定所述特定节点正在运行。

6. 一个或多个计算机可读介质,其存储可执行指令,所述可执行指令当由一个或多个处理器执行时,使所述一个或多个处理器执行包括以下动作:

将测试数据分组传送到目的地节点,所述测试数据分组包括所述测试数据分组要被路由通过的多于一个特定节点的信息;以及

至少部分基于所述测试数据分组的所述传送的结果来确定所述多于一个特定节点的健康状况;以及

至少将第一分组迭代地封装在第二分组中,并且至少将所述第二分组迭代地封装在第三分组中以形成所述测试数据分组。

7. 根据权利要求6所述的一个或多个计算机可读介质,其中迭代地封装多于一个相应的数据分组包括:迭代地在所述多于一个相应的数据分组的分组头部中设置所述多于一个特定节点的目的地地址。

8. 根据权利要求7所述的一个或多个计算机可读介质,还包括:确定是全局地址、局部地址还是其组合被用作特定节点的相应的数据分组的分组头部中的特定节点的目的地地

址。

9. 根据权利要求6所述的一个或多个计算机可读介质,所述动作还包括添加所述测试数据分组从发送所述测试数据分组的节点到所述多于一个特定节点中的至少一个特定节点遍历的跳的数目的信息。

10. 根据权利要求6所述的一个或多个计算机可读介质,其中所述测试数据分组作为目的地的接收节点与所述测试数据分组源自于的发送节点相同,并且所述动作还包括:

发送出所述测试数据分组;以及

其中确定所述多于一个特定节点的所述健康状况至少部分基于所述测试数据分组是否被接收。

11. 一种用于网络管理的系统,包括:

一个或多个处理单元;

存储器,其存储可执行指令,所述可执行指令当由所述一个或多个处理单元执行时,使所述一个或多个处理单元执行包括以下动作:

至少将第一数据分组迭代地包含在第二数据分组内,并且至少将所述第二数据分组迭代地包含在第三数据分组内,所述第二数据分组的分组头部包括与多于一个特定节点相关联的信息,所述多于一个特定节点的健康状况要被确定以用于网络监控,所述第三数据分组的分组头部包括与另一特定节点相关联的信息,所述另一特定节点的健康状况要被确定,并且所述第一数据分组的分组头部包括与目的地节点相关联的信息,所述目的地节点确定所述特定节点的操作状况;以及

发送出所述第三数据分组;

其中所述多于一个特定节点是所述第三数据分组要被路由通过的节点。

12. 根据权利要求11所述的系统,其中与所述特定节点相关联的所述信息包括所述特定节点的全局地址或局部地址。

一种细粒度网络监控方法及系统

背景技术

[0001] 数据中心网络将巨大数量的设备相互连接并且实现从网络中的一个设备到另一设备的数据传输。为了提供可靠数据传输,数据中心网络的拓扑结构通常被设计为允许在网络中的两个设备之间的多条路径用于数据传输。尽管多条路径设计可以在故障和流量拥塞的时间提供性能上的平滑下降,但是该设计也可能增大标识网络中的任何发生故障的或不健康的设备或在网络中的设备之间的连接链路的难度。

发明内容

[0002] 本发明内容介绍了细粒度网络监控的简化构思,下面在具体实施方式中进一步描述细粒度网络监控。本发明内容不旨在确定要求保护的主题的必要特征,也不旨在用于限制要求保护的主题的范围。

[0003] 本申请描述细粒度网络监控的示例实施例。在一个实施例中,发送节点确定或选择一个或多个特定节点或连接链路,该一个或多个特定节点或连接链路的操作状况或健康状况将被分析。在确定一个或多个特定节点或连接链路时,发送节点可以迭代地对对应于一个或多个特定节点或连接链路的数据分组进行封装或包裹以形成测试数据分组。在一个实施例中,发送节点可以将使得测试数据分组能够路由通过一个或多个特定节点或连接链路的信息插入或包含在对应于一个或多个特定节点或连接链路的数据分组的分组头部中。在一些实施例中,在形成测试数据分组之后,发送节点可以将测试数据分组发送出到网络(例如,数据中心网络)。接收测试数据分组的发送节点或目的地节点可以至少部分地基于测试数据分组是否是根据预定准则接收的来确定一个或多个特定节点或连接链路的操作状况或健康状况。

附图说明

[0004] 参考附图阐述具体实施方式。在附图中,附图标记的最左边的(多个)数字标识该附图标记首次出现的附图。在不同的附图中对相同的附图标记的使用指示相似或相同的项。

[0005] 图1图示了细粒度网络监控系统的示例环境。

[0006] 图2图示了如图1所示的示例细粒度网络监控系统的设备的示例。

[0007] 图3A和图3B图示了用于在网络中对测试数据分组进行路由的示例框架。

[0008] 图4图示了细粒度网络监控的示例方法。

[0009] 图5图示了如图1所示的示例细粒度网络监控系统的设备的示例用户界面。

具体实施方式

[0010] 已经提出了用于网络监控和对发生故障的设备和连接链路的标识的许多算法。然而,这样的算法可以不仅在网络监控期间将大量额外流量引入在网络中,而且还可能不能有效地且准确地标识发生故障的设备或连接链路(例如,不正常地工作和/或已经过载的发

生故障的设备或连接链路)。

[0011] 本公开内容描述了一种可在诸如数据中心网络的数据网络或通信网络中使用的网络监控系统。网络监控系统确定或选择节点或链路,该节点或链路的操作状况或健康状况将被分析或探查,并且创建被配置为路由或遍历通过该节点或链路的数据分组(例如,测试数据分组)以便确定该节点或链路的操作状况或健康状况。在一个实施例中,网络监控系统可以使用隧道封装来创建或生成包括数据分组的多个层或级的测试数据分组,其中数据分组的至少一个层或级被配置为遍历要分析的特定节点或链路。基于测试数据分组的路由结果,网络监控系统可以确定要分析的节点或链路的操作状况或健康状况,并且可以将分析结果报告给相关的人,例如网络管理员或网络的操作员,以用于稍后分析和维护。

[0012] 在一个实施例中,该节点或链路的操作状况或健康状况可以包括但不限于该节点或链路是否正常地或如所预期的工作,该节点或链路是否过载,该节点或链路是否被断开,等等。在一些实施例中,网络监控系统可以随机地或策略性地(例如,基于特定选择算法、等等)选择节点或路径。额外地或备选地,网络监控系统可以确定或选择多于一个节点和/或链路以用于分析单个监控探查或多个监控探查中的相应的操作状况或健康状况。

[0013] 在确定或选择要分析或探查其操作状况或健康状况的节点或链路之后,网络监控系统可以确定或选择用于将数据分组路由到该节点或链路的策略。在一个实施例中,网络监控系统可以创建第一数据分组和第二数据分组,并且将第二数据分组包裹或包含在第一数据分组内以生成或创建测试数据分组。在一个实例中,网络监控系统可以使用隧道封装来生成或创建测试数据分组。网络监控系统可以设置第一数据分组的分组头部中的一个或多个属性以指引第一数据分组遍历或路由通过所选择的节点或路径。通过举例而非限制的方式,网络监控系统可以将第一数据分组的分组头部中的目的地地址设置为与所选择的节点或链路相关联的地址。在一个实施例中,与所选择的节点或链路相关联的地址可以包括例如对应于可在网络中路由的地址的全局地址、对应于可仅仅通过例如与所选择的节点或链路相邻的一个或多个节点路由或到达的地址的局部地址。

[0014] 额外地或备选地,在一些实施例中,网络监控系统可以设置第一数据分组的分组头部中的可以用于反映或定义到所选择的节点或链路的路径的一个或多个其他属性。例如,网络监控系统可以设置针对跳数(即跳的数目)或跳限制(即,在丢弃或丢掉数据分组之前允许数据分组行进的跳的最大数目)的属性值以反映或定义到第一数据分组的分组头部中的所选择的节点或链路的路径。

[0015] 如果多于一个节点或链路要在该监控探查中被分析,则网络监控系统可以迭代地包裹或封装数据分组以生成或创建测试数据分组,其中相应的分组头部中的一个或多个属性要被设置为定义要被分析的对应的节点或链路。

[0016] 在生成测试数据分组时,网络监控系统可以将包含一个或多个其他数据分组的测试数据分组发送到网络。在一个实施例中,网络监控系统可以将测试数据分组发送到目的地节点,该目的地节点可以与测试数据分组从其被发送或来源于其的节点相同或不同。在一个实例中,将测试数据分组发送回到网络监控系统驻存在其中的节点使网络监控系统免于请求和等待来自在其处接收到测试数据分组的另一节点的测试数据分组的路由结果。

[0017] 在一个实施例中,网络监控系统可以基于一个或多个预定准则来确定所选择的节点或链路的操作状况或健康状况。一个或多个预定准则可以包括但不限于是否在目的地节

点处成功接收到数据分组,是否在预定时间段内在目的地节点处接收到数据分组,等等。

[0018] 如果数据分组满足一个或多个预定准则,则网络监控系统可以确定所选择的节点或链路是否正常地或如预期的工作。如果数据分组未能满足一个或多个预定准则,则网络监控系统可以确定所选择的节点或链路易受失灵和/或过载的影响。在一些实施例中,网络监控系统可以执行进一步的分析来确定所选择的节点或链路是否失灵或过载。

[0019] 所描述的系统选择或确定特定节点或链路以用于确定该特定节点或链路是否正常工作,并且因此允许诸如数据中心网络的网络中的一个或多个节点的细粒度监控。所描述的系统可以将该节点或链路的操作状况报告给网络管理员或操作员以用于进一步分析和/或后续维护。

[0020] 在本文中描述的示例中,网络监控系统确定用于分析的节点或链路,将第一数据分组封装在第二数据分组内,发送出第二数据分组,以及基于第一数据分组是否是根据一个或多个预定准则接收到的来确定该节点或链路的操作。然而,在其他实施例中,这些功能可以由位于相同位置或不同位置处的一个或多个服务执行。例如,在至少一个实施例中,选择服务可以选择要探查哪个节点或链路,而准备服务可以准备要被发送的包括封装的分组的数据分组。发送服务可以将数据分组发送出到目的地节点,并且确定服务可以基于一个或多个预定准则来确定所选择的节点或链路的操作状况或健康状况。

[0021] 另外,尽管在本文中描述的示例中,网络监控系统可以被实施为被安装在单个设备中的软件和/或硬件或被实施为服务,但是在其他实施例中,网络监控系统可以被实施在通过网络被提供在一个或多个服务器中的和/或被分布在分布式计算体系结构或云计算体系结构中的多个设备和/或服务中。

[0022] 本申请描述多个和不同的实施方式和实施例。下面的部分描述可以被用于实践各种实施方式的框架的说明性示例。接下来,本申请描述用于实施网络监控系统的示例系统、设备和过程。

[0023] 示例框架

[0024] 图1图示了可用于实施网络监控系统102的示例框架100。在该示例中,网络监控系统102被描述为被包含于多个设备104-1、104-2、...、104-N(其统称为设备104)中的一个中。然而,在其他实例中,网络监控系统102可以是与设备104独立或分离的实体。例如,网络监控系统102可以被包含于一个或多个服务器106中和/或被分布在一个或多个服务器106之间,一个或多个服务器106可以经由网络108与彼此和/或与设备104通信数据。额外地或备选地,在一些实例中,网络监控系统102的功能可以被包含于一个或多个设备104和一个或多个服务器106中和/或被分布在一个或多个设备104和一个或多个服务器106之间。例如,一个或多个服务器106可以包含网络监控系统102的功能的部分,而网络监控系统102的其他功能可以被包含于一个或多个设备104中。另外,在一些实施例中,例如,网络监控系统102的一些或所有功能可以被包含于由服务器106和/或设备104形成的云计算系统或体系结构中。在其他实例中,一个或多个服务器106可以是网络108的部分。

[0025] 设备104中的一个或多个可以被实施为各种计算设备中的任何,包括但不限于台式计算机、笔记本计算机或便携式计算机、手持设备、上网本、互联网设备、平板或平板式计算机、移动设备(例如,移动电话、个人数字助理、智能电话、等等)、等等或其组合。

[0026] 网络108可以是无线网络或有线网络或其组合。网络108可以是与彼此相互连接的

并且用作单个大型网络(例如,互联网或内联网)的个体网络的汇集。这样的个体网络的示例包括但不限于电话网络、线缆网络、局域网(LAN)、广域网(WAN)和城域网(MAN)。另外,个体网络可以是无线网络或有线网络或其组合。有线网络可以包括电载体连接(例如通信线缆,等等)和/或光学载体或连接(例如光纤连接,等等)。无线网络可以包括例如WiFi网络、其他射频网络(例如,蓝牙®、Zigbee、等等)、等等。在一个实施例中,网络108可以包括数据中心网络。

[0027] 额外地,在一个实例中,网络108可以包括多个节点110和多个连接链路112。多个节点110可以包括交换或路由组件,例如交换机(例如,商品交换机、等等)、路由器、集线器、等等。在一些实施例中,多个节点110还可以包括具有处理和/或存储能力的一个或多个设备,例如设备104。与多个连接链路112一起,多个节点110可以将多个设备104和一个或多个服务器106与彼此进行相互连接。在一个实施例中,多个节点110可以经由多个连接链路112被组织以形成针对整个网络108的特定拓扑结构或网络108的不同部分中的不同拓扑结构。示例拓扑结构可以包括但不限于胖形拓扑结构、环形拓扑结构、星形拓扑结构、总线拓扑结构、混合拓扑结构或其各种组合。在一些实例中,多个节点110可以被组织为多个层的交换机,包括机架顶(ToR)交换机、聚合交换机、核心交换机、等等。

[0028] 在一个实施例中,特定设备(例如,设备104-N)可以包括耦合到存储器116的一个或多个处理单元114。一个或多个处理单元114可以被实施为一个或多个硬件处理器,包括例如微处理器、专用指令集处理器、图形处理单元、物理处理单元(PPU)、中央处理单元(CPU)、图形处理单元(GPU)、数字信号处理器、等等。额外地或备选地,本文描述的功能能够至少部分地由一个或多个硬件逻辑组件来执行。例如但不限于,可以被使用的说明性类型的硬件逻辑组件包括现场可编程门阵列(FPGA)、专用集成电路(ASIC)、专用标准产品(ASSP)、片上系统(SOC)、复杂可编程逻辑器件(CPLD)、等等。

[0029] 存储器116可以包括或存储由一个或多个处理单元114运行的一个或多个应用118(例如,网络监控应用、等等)以及其他程序数据120。存储器116可以耦合到其他设备、与其他设备相关联和/或对其他设备可访问,其他设备例如网络服务器、路由器和/或服务器106。

[0030] 存储器116可以包括易失性存储器,例如随机访问存储器(RAM)和/或非易失性存储器,例如只读存储器(ROM)或闪速RAM。存储器116是计算机可读介质的示例。计算机可读介质至少包括两种类型的计算机可读介质,即计算机存储介质和通信介质。

[0031] 计算机存储介质包括以任何方法或技术实现的用于存储诸如计算机可读指令、数据结构、程序模块或其他数据的信息的易失性介质和非易失性介质、可移除介质和不可移除介质。计算机存储介质包括但不限于相变存储器(PRAM)、静态随机访问存储器(SRAM)、动态随机访问存储器(DRAM)、其他类型的随机访问存储器(RAM)、只读存储器(ROM)、电可擦可编程只读存储器(EEPROM)、闪存或其他存储器技术、紧凑盘只读存储器(CD-ROM)、数字多用盘(DVD)或其他光学存储装置、磁带盒、磁带、磁盘存储装置或其他磁性存储设备、或者能够被用于存储用于由计算设备访问的信息的任何其他非传输介质。

[0032] 对比之下,通信介质可以将计算机可读指令、数据结构、程序模块、或其他数据实现在诸如载波的经调制的数据信号或其他传输机制中。如本文中所限定的,计算机存储介质不包括通信介质。

[0033] 用户122可以使用设备104的应用118(例如网络监控应用、浏览器应用、等等)以监控网络108中的节点110和/或连接链路112的操作状况或健康状况。在一个实施例中,网络监控应用可以是由网络监控系统102提供的应用。在一些实施例中,网络监控应用可以是能够与网络监控系统102进行通信和与网络监控系统102合作以执行网络监控的独立应用。网络监控应用可以提供与网络108的拓扑结构相关联的形式为映射和/或列表的信息,并且可以允许用户122选择特定节点110和/或特定连接链路112以用于分析。例如,在接收到节点110和/或连接链路112的选择的指示之后,网络监控系统102创建探查(例如,测试数据分组、等等)以确定所选择的节点110和/或连接链路112的操作状况或健康状况,并且经由设备104的显示器124将分析的结果返回给用户122。

[0034] 示例网络监控系统

[0035] 图2更详细地图示了示例网络监控系统102。在该示例中,示例网络监控系统102被描述为被包含于设备104中或设备104的部分中。如在前面所描述的,设备104可以包括但不限于一个或多个处理单元114和存储器116。额外地,设备104还可以包括一个或多个应用118。在一些实施例中,设备104还可以包括网络接口202和输入/输出接口204。一个或多个处理单元114被配置为运行从网络接口202接收到的、从输入/输出接口204接收到的和/或被存储在存储器116中的指令。在一个实施例中,设备104还包括显示器124。显示器124可以包括触摸屏、普通屏幕(即,没有触敏能力的屏幕)、等等。

[0036] 网络监控系统102可以包括程序模块206和程序数据208。在一个实施例中,网络监控系统102可以包括输入模块210。输入模块210可以从用户122接收与节点110或连接链路112相关联的信息,该节点110或连接链路112的操作状况或健康状况将被分析。例如,网络监控系统102的输出模块212可以提供多个节点110和/或连接链路112中的全部或子集的拓扑映射(或列表)以用于在设备104的显示器124中呈现给用户122。在一个实施例中,例如,输出模块212可以提供拓扑映射或列表以用于经由设备104的应用118(例如网络监控应用、浏览器应用)呈现。在用户122可以从拓扑映射或列表选择节点110或连接链路112之后,输入模块210从设备104或应用118接收所选择的节点110或连接链路112的信息。

[0037] 额外地或备选地,网络监控系统102可以包括选择模块214。选择模块214可以自动地或半自动地基于一个或多个选择算法或策略来选择节点110或连接链路112。例如,选择模块214可以从多个节点110和/或连接链路112中随机地选择节点110和/或连接链路112。在一些实例中,用户122可以已经向网络监控系统102指示了要分析的网络108的特定部分。在经由输入模块210接收到网络108的特定部分的信息之后,选择模块214可以从网络108的该特定部分随机地选择节点110和/或连接链路112。在一个实施例中,选择模块214可以例如通过逐步地从最接近设备104的节点110或连接链路112开始直到如由用户122指示的节点110或连接链路112选择节点110和/或连接链路112来策略性地选择节点110或连接链路112。在一些实例中,选择模块214可以随机地选择网络108的特定部分内的一个或多个节点110和/或连接链路112。在一个实例中,网络监控系统102可以确定或选择要被分析用于单个监控探查或多个监控探查的一个或多个节点110和/或连接链路112。额外地或备选地,网络监控系统102可以基本上在相同的时间或在不同的时间探查一个或多个节点110和/或连接链路112。

[0038] 在选择要被分析的节点110或连接链路112时,网络监控系统102可以采用准备模

块216来准备要被发送用于探查所选择的节点110 或连接链路112的操作状况或健康状况的数据分组。在一个实施例中,准备模块216可以获得可用于将数据分组路由到所选择的节点 110或连接链路112的信息。通过举例而非限制的方式,可用于将数据分组路由到所选择的节点110或连接链路112的信息可以包括全局地址、局部地址、到达所选择的节点110的跳的(最小)数目、等等。准备模块216可以从包括与多个节点110和多个连接链路112 相关联的地址和/或标识信息的数据库218获得可用于将数据分组路由到所选择的节点110或连接链路112的信息。额外地,在一些实施例中,数据库218还可以包括网络108的拓扑信息和/或在一个或多个先前探查处确定的多个节点110和多个连接链路112的操作状况或健康状况、等等。

[0039] 在一个实施例中,节点的局部地址可以对应于旨在用于在节点的邻域内通信的并且(仅仅)可由在节点的邻域内的一个或多个邻域节点和/或位于网络108的部分或全部内的节点的子集到达的网络地址。额外地,节点的全局地址可以对应于可由在网络108内部和/或外部的另一节点(例如,设备104、服务器106、等等)路由和/或访问的网络地址。与节点相关联的跳的数目或跳数对应于将数据分组从发送出数据分组的节点或设备(例如在该示例中设备104)路由到节点110需要的跳的数目。在一些实施例中,与所选择的连接链路112相关联的信息可以包括但不限于与对应于所选择的连接链路112 的两个终端的节点相关联的地址或路由信息(例如,全局地址、局部地址、或其组合、等等)。

[0040] 在获得可用于将数据分组路由到所选择的节点110或连接链路 112的信息时,准备模块216可以通过基于所获得的信息准备数据分组来控制或指定数据分组要被遍历或路由通过的路径或路径的部分。通过举例而非限制的方式,准备模块216可以根据隧道协议来准备或生成测试数据分组。隧道协议的示例可以包括但不限于 IP-in-IP、GRE(通用路由封装)、MPLS(多协议标记交换)、等等。为便于描述,IP-in-IP隧道协议在下文中被用于图示。然而,本公开内容不被理解为该IP-in-IP隧道协议,但是还适用于如以上所描述的其他隧道协议。

[0041] 在一个实施例中,如果选择了单个节点110,则准备模块216可以将第一数据分组(例如,“内”数据分组)包裹或包含于第二数据分组(例如,“外”数据分组)的数据主体或有效载荷中以形成测试数据分组。准备模块216还可以生成两个分组头部,针对内数据分组的内分组头部和针对外数据分组的外分组头部。准备模块216 可以在外数据分组的外分组头部中设置或包含与所选择的节点110 相关联的路由信息(例如,全局地址、局部地址、跳数、等等),并且在内数据分组的内分组头部中设置或包含另一节点110(例如,目的地节点110或设备104)的路由信息。取决于网络108采用和/ 或设备104或节点110利用的协议的类型或寻址方案,分组头部可以包括IPv6分组头部、IPv4头部、等等。

[0042] 额外地或备选地,如果选择了连接链路112,则准备模块216可以将第一数据分组包裹或包含于第二数据分组中以形成测试数据分组。准备模块216还生成两个分组头部,针对第一数据分组的第一分组头部和针对第二数据分组的第二分组头部。准备模块216可以在第一数据分组的第一分组头部中设置或包含与连接链路112的第一终端相关联的路由信息(例如,全局地址、局部地址、跳数、等等),并且在第二数据分组的第二分组头部中设置或包含连接链路 112的第二终端的路由信息。额外地,如果目的地节点110与连接链路112的第一终端不同,则准备模块216可以将第三数据分组包裹或包含于具有包含与目的地节点

110或设备104相关联的路由信息的第三数据分组的分组头部的第一数据分组中。

[0043] 在一些实施例中,如果选择了多于一个节点110和/或连接链路 112,则准备模块216可以如以上所描述的一个接着一个地迭代地包裹或包含对应于所选择的节点110和/或连接链路112的相应的数据分组以形成测试数据分组。另外,准备模块216可以因此生成相应的数据分组的分组头部以在其中分别设置或包含所选择的节点110 和/或连接链路112的路由信息。在一些实例中,包含于测试数据分组的最里面的数据分组的分组头部中的路由信息可以对应于针对目的地节点110或设备104的路由信息。

[0044] 在一些实施例中,准备模块216可以设置在与测试数据分组的所选择的节点110和/或连接链路112相关联的一个或多个数据分组的相应的分组头部中的一个或多个其他属性或参数。例如,准备模块 216可以设置在与所选择的节点110或连接链路112相关联的数据分组的分组头部上的DSCP(差分服务代码点)的值以控制在数据分组遍历的路径上的数据分组所属的优先级组(PG)。额外地或备选地,准备模块216可以设置在与所选择的节点110或连接链路112相关联的数据分组的分组头部上的ECN(显式拥塞通知)的值以控制数据分组可以在数据分组遍历的路径上具有的拥塞经验。

[0045] 在创建或生成测试数据分组时,网络监控系统102的发送模块 220可以经由网络108将测试数据分组发送出到目的地节点110或设备104。额外地,网络监控系统102可以包括接收模块222,其被配置为等待或监听测试数据分组的路由结果。在一个实施例中,路由结果可以包括是否在目的地节点110或设备104处接收到测试数据分组的封装的数据分组(例如,封装在测试分组中的最里面的数据分组)的结果、等等。取决于目的地节点110或设备104是否包括网络监控系统102或其部分,接收模块222可以接收在(网络监控系统102的至少部分被包含于其中的)目的地节点110或设备104 处的路由结果,或者(在网络监控系统102是独立于目的地节点110 或设备104的实体的情况下)接收作为从目的地节点110或设备104 发送的通知消息的路由结果。

[0046] 在一个实施例中,网络监控系统102的确定模块224可以根据一个或多个预定准则来确定所选择的节点110或连接链路112的操作状况或健康状况。一个或多个预定准则可以包括例如是否由接收模块222接收到路由结果,是否在预定时间段内由接收模块222接收到路由结果,等等。如果未接收到路由结果或者如果在预定时间段外接收到路由结果,则确定模块224可以确定要分析其操作状况或健康状况的所选择的节点110或连接链路112易受有问题的状况影响,例如所选择的节点110或连接链路112过载或遭受流量拥塞、被断开或失灵、等等。例如,如果路由结果在预定时间段外由接收模块222接收到,则确定模块224可以确定所选择的节点110或连接链路112易受过载或流量拥塞影响。例如,如果没有由接收模块 22接收到路由结果,则确定模块224可以确定所选择的节点110或连接链路112可以易受失灵或被断开影响。

[0047] 在一些实施例中,确定模块224可以确定在网络108中存在问题,但是可能不能够区分问题是与所选择的节点110或连接链路112相关,还是与网络108中的其他节点或连接链路相关。确定模块224 可以确定进一步的分析是值得期望的。确定模块224可以指令选择模块214选择与所选择的节点110或连接链路112相邻的一个或多个其他节点110和/或连接链路112以分析或确定它们的操作状况或健康状况。在一个实施例中,如果第一节点或连接链路处在远离第二节点或连接链路的预定跳数(例如,一个、两个、三个、等等) 处,则第一

节点或连接链路和第二节点或连接链路相邻。例如,预定跳数可以由网络108的网络管理员或操作员(例如,用户122)定义。额外地或备选地,确定模块224可以将提示提供给用户122并请求用户122提供要分析一个或多个其他节点110和/或连接链路 112的指令和/或指示。在一些实施例中,网络监控系统102还可以包括其他程序数据120,例如已经被分析的一个或多个节点110和/或连接链路112的操作状况和/或对应的问题的记录。

[0048] 示例情景

[0049] 图3A图示了根据前面的实施例的遍历测试数据分组的第一示例框架或情景300。在该示例中,网络监控系统102准备测试数据分组 302,测试数据分组302包括外分组头部304、内分组头部306和数据主体(或有效载荷) 308,如在前面的实施例中所描述的。例如,外分组头部304包括作为目的地地址的与要分析其操作状况的特定节点312(例如,诸如核心交换机的交换机) 相关联的路由信息(例如,全局地址、局部地址或跳数、等等)。此外,在该示例中,内分组头部306具有作为内分组的对应的目的地地址的源310的路由信息。网络监控系统102可以之后经由网络108将测试数据分组302 从源310(例如,设备104或服务器106)发送。

[0050] 在一个实施例中,测试数据分组302可以被路由通过一个或多个中间节点314并到达特定节点312。测试数据分组302可以在特定节点312的数据平面或层处解封装或解包裹而不消耗特定节点312的处理资源。在解封装或解包裹时,经解封装或解包裹的数据分组316 可以基于包含于内分组头部306中的目的地地址经由一个或多个中间节点314(当测试数据分组302从源310被路由到特定节点312 时与这些节点可以相同或可以不相同)被路由或转发到源310。在一个实施例中,经解封装或解包裹的数据分组316可以使用用于转发或路由普通或常见数据分组的相同机制经由一个或多个中间节点 314被路由或转发到源310。如在前面的实施例中描述的,在到达源 310之后,网络监控系统102可以分析与经解封装或解包裹的数据分组316相关联的信息以确定特定节点312的操作状况。

[0051] 在一些实施例中,如果针对特定节点312存在问题,例如特定节点312被断开或已经过载,则测试数据分组302可能不能够到达特定节点312并且因此在源310处接收不到与测试数据分组302相关联的数据分组。在这种情况下,网络监控系统102可以确定或检测特定节点312当前正在经历问题。网络监控系统102可以之后提供分析结果以用于经由设备104的显示器124呈现给用户122,并且等待来自用户122的另外的指令。额外地或备选地,网络监控系统102 可以在具有或没有来自用户122的介入或指令的情况下逐步地选择与特定节点312相邻的一个或多个节点110和/或连接链路112以标识或定位问题的源。

[0052] 图3B图示了根据前面的实施例的遍历测试数据分组的第二示例框架或情景318。在该示例中,网络监控系统102可以以单个监控探查或尝试确定或分析多个节点110和/或连接链路112的操作状况或健康状况。在一个实施例中,如在前面的实施例中描述的,测试数据分组320可以包括外分组头部322、多个内分组头部324-1、……、324-K和数据主体(有效载荷) 326,其中K是大于一的整数。外分组头部322可以包括作为目的地地址的与要分析其操作状况的第一节点328(例如,诸如ToR交换机的交换机) 相关联的路由信息(例如,全局地址、局部地址或跳数、等等)。内分组头部324-1可以包括作为目的地地址的与要分析其操作状况的第二节点330(例如,诸如聚合交换机的交换机) 相关联的路由信息(例如,全局地址、局部地址或跳数、等等)。此外,在该示例中,最里面的分组头部324-K 可以包括作为最

里面的数据分组334的对应的目的地地址的目的地 322的路由信息。网络监控系统102可以之后经由网络108将测试数据分组320从源336 (例如,设备104或服务器106) 发送。

[0053] 在一个实施例中,当测试数据分组320成功地被路由和到达在第一节点328处时,则测试数据分组320在消耗或不消耗第一节点328 的处理资源的情况下在第一节点328的数据平面或层处被解封装或解包裹以暴露内分组头部324-1。在解封装或解包裹时,第一未包裹的数据分组338可以随后基于包含于内分组头部324-1中的目的地地址而经由M数目的中间节点340被路由到第二节点330,其中M是等于或大于零的整数。在一个实施例中,第一未包裹的数据分组338 可以使用用于转发或路由普通或常见数据分组的相同的机制而被路由或转发到第二节点330。在一个实例中,如果成功地到达第二节点 330处,则第一未包裹的数据分组338可以在第二节点330的对应的数据平面或层处被解封装或被解包裹以生成第二未包裹的数据分组 342。第二未包裹的数据分组342可以之后经由零或一些中间节点被路由到要分析其操作状况的另一节点或目的地332。在一个实施例中,第二未包裹的数据分组342可以使用用于转发或路由普通或常见数据分组的相同的机制而被路由或转发到目的地332。如在前面的实施例中描述的,取决于测试数据分组320的最里面的数据分组334是否成功地到达目的地332,网络监控系统102可以确定多个节点110和/或连接链路112中的一个或多个是正常工作还是过载、等等。

[0054] 备选实施方式

[0055] 尽管网络监控系统102被描述为从其发送和/或接收测试数据分组的设备104的部分或被包含于从其发送和/或接收测试数据分组的设备104中,但是在一些实施例中,网络监控系统102可以被包含于一个或多个设备104和/或与从其发送测试数据分组的设备104和/ 或测试数据分组被送往其的设备104不同的一个或多个服务器106 中。在该实例中,参考图1作为示例,网络监控系统102可以将请求第一设备104以准备和发送测试数据分组的请求或指令发送到第一设备104 (例如,设备104-1)。网络监控系统102还可以将另一请求或指令发送到可以与第一设备104相同或不同的第二设备104,另一请求或指令请求第二设备104将与测试数据分组的路由结果有关的通知消息发送回到网络监控系统102。在一些实例中,如果第一设备104和第二设备104相同,则网络监控系统102可以将单个请求或指令发送到该相同的设备104以实现对测试数据分组的发送和对测试数据分组的路由结果的通知两者。

[0056] 另外,尽管前面的实施例描述了网络监控系统102生成或创建用于探查所选择的节点的操作状况或健康状况的测试数据分组,但是在其他实例中,网络监控系统102可以将该探查并入到初始或最初与网络监控不相关的原始或普通数据分组中。例如,原始或普通数据分组可以包括由于与网络监控不相关的另一应用118、第一设备或第二设备的请求或操作要从第一设备被发送到第二设备的数据分组。在这种情况下,网络监控系统102可以将原始数据分组包裹或封装在对应于要分析其操作状况的节点的数据分组内以形成新的数据分组,并经由要被分析的节点将新的数据分组发送到第二设备。在一些实例中,网络监控系统102可以已经与第二设备关于指示探查与原始数据分组的传输的该并入的数据格式进行了协商或同意,并且可以修改或可以不修改原始数据分组的分组头部和/或数据主体 (例如,有效负荷) 以因此指示该并入。在接收到 (具有或没有取决于协商或同意的格式的修改的) 原始数据分组之后,第二设备可以基于同意的数据格式来识别修改并向网络监控系统

统102通知原始数据分组的路由结果(并且因此通知对应于要被分析的节点的数据分组的路由结果)。

[0057] 示例方法

[0058] 图4是描绘了网络监控的示例方法400的流程图。在一些情况下,图4的方法可以使用图2的网络监控系统 and/或在对应于图3的相似情景下被实施在图1的框架中。为了便于解释,参考图1-3描述方法 400。然而,方法400可以备选地被实施在其他环境中 and/或使用其他系统来实施。

[0059] 在计算机可执行指令的总体背景下描述图4中图示的方法400。总体上,计算机可执行指令可以包括执行特定功能或实现特定抽象数据类型的例程、程序、对象、组件、数据结构、流程、模块、功能、等等。方法还可以被实践在分布式计算环境中,其中功能由通过通信网络链接的远程处理设备执行。在分布式计算环境中,计算机可执行指令可以被定位在包括存储器存储设备的本地计算机存储介质和/或远程计算机存储介质中。

[0060] 示例方法被图示为逻辑流程图中的表示可以采用硬件、软件、固件或其组合实施的操作的序列的框的汇集。方法被描述的顺序不旨在被理解为限制,并且任何数量的所描述的方法框能够以任何顺序来组合以实施该方法或备选方法。额外地,个体框可以在不脱离本文中描述的主题的精神和范围的情况下从方法中被省略。在软件的背景下,框表当由一个或多个处理器运行时执行所记载的操作的计算机指令。在硬件的背景下,框中的一些或全部可以表示执行所记载的操作的专用集成电路 (ASIC) 或其他物理组件。

[0061] 参考图4,在框402处,方法400包括确定要被分析的节点或连接链路。例如,参考图2,输入模块210或选择模块214可以确定或选择一个或多个节点110和/或一个或多个连接链路112,该一个或多个节点110和/或一个或多个连接链路112的操作状况或健康状况将被分析。

[0062] 在框404处,方法400包括获得要被分析的节点或连接链路的路由信息。例如,参考图2,准备模块216可以获得与一个或多个所选择的节点110和/或一个或多个所选择的连接链路112相关联的相应的路由信息。

[0063] 在框406处,方法400包括对一个或多个数据分组进行封装或包裹以形成测试数据分组。例如,参考图2,准备模块216可以根据隧道协议来将第一数据分组封装或包裹到第二数据分组中以形成测试数据分组。在一个实施例中,第一数据分组和第二数据分组中的至少一个对应于所选择的节点110或所选择的连接链路112的终端。

[0064] 在框408处,方法400包括确定是否要分析一个或多个节点和/或连接链路。例如,参考图2,准备模块216可以确定是否要对另一所选择的节点110或另一所选择的连接链路112的终端执行更多的封装或包裹。如果要执行更多的封装或包裹,则准备模块216迭代地或重复地将测试数据分组封装或包裹在对应于另一所选择的节点 110或另一所选择的连接链路112的终端的另一数据分组中。

[0065] 在框410处,方法400包括发送测试数据分组。例如,参考图2,如果不再需要封装或包裹,则发送模块220可以发送出测试数据分组。

[0066] 在框412处,方法400包括接收测试数据分组的路由结果。例如,参考图2,接收模块222可以在目的地节点或最终节点110或设备 104处或从目的地节点或最终节点110或设备104接收测试数据分组的路由结果。

[0067] 在框414处,方法400包括确定节点或连接链路的操作状况。例如,参考图2,确定模块224可以基于路由结果并且根据一个或多个预定准则来确定一个或多个所选择的节点110和/或一个或多个所选择的连接链路112的相应的操作状况。

[0068] 在框416处,方法400包括确定是否要分析另一节点或连接链路。例如,参考图2,输入模块210或选择模块214可以确定是否要分析另一节点110或连接链路112。如果不需要分析额外的节点110或连接链路112,则输入模块210或选择模块214可以停止并等待下一请求或指令。如果需要分析额外的节点110或连接链路112,则准备模块216可以获得与额外的节点110或连接链路112相关联的路由信息,并且准备新的测试数据分组。

[0069] 本文中描述的方法中的任何的动作的任何可以至少部分地由处理器或其他电子设备基于存储在一个或多个计算机可读介质上的指令而被实施。通过举例而非限制的方式,本文中描述的方法中的任何的动作的任何可以在使用可执行指令被配置的一个或多个处理器的控制下来实施,可执行指令可以被存储在诸如一个或多个计算机存储介质的一个或多个计算机可读介质上。另外,如在前文中描述的各种实施例的组件和操作可以在不脱离本公开内容的情况下进行组合、重新布置、替代和/或省略。

[0070] 示例用户界面

[0071] 图5图示了可以由示例网络监控系统102用于与用户122交互或通信的示例用户界面500。在该示例中,用户界面500被描述为通过应用118提供的用户界面,应用118例如设备104的网络监控应用。在其他实例中,用户界面500可以是由网络监控系统102远程地提供的并且经由设备104的显示器124通过应用118(例如,浏览器应用)呈现给用户122的用户界面。在该示例中,显示器124被描述为触摸屏。在其他实例中,显示器124可以包括没有触敏能力的普通屏幕。

[0072] 在一个实施例中,用户界面500可以包括网络108中的节点110 和/或连接链路112的部分或全部的拓扑映射502和/或列表504。拓扑映射502和/或列表504可以提供与网络108中的一个或多个节点 110和/或连接链路112相关联的信息。提供的信息的示例可以包括但不限于一个或多个节点110和/或连接链路112的标识信息、一个或多个节点110和/或连接链路112的操作状况、一个或多个节点110 和/或连接链路112的上次更新的时间、等等。在一个实例中,与节点110和/或连接链路112相关联的信息可以当用户122将指向工具(例如手指或触摸屏的触笔、普通屏幕的鼠标、等等)放置在表示该节点110或连接链路112的图形对象上时在拓扑映射502上被呈现给用户112。

[0073] 在一个实施例中,用户122可以选择506特定节点110或连接链路112以用于指令网络监控系统102确定该特定节点110或连接链路112的操作状况或健康状况。在一些实施例中,用户122可以选择508网络108的包括要分析其操作状况的一个或多个节点110和/ 或一个或多个连接链路112的部分或子集。

[0074] 在经由输入模块210接收到特定节点110或连接链路112(或网络108的子集)的选择的指示时,网络监控系统102可以执行如在前面的实施例中所描述的所选择的节点110或连接链路112(或网络 108的所选择的子集)的网络监控。响应于确定所选择的节点110 或连接链路112(或网络108的所选择的子集)的操作状况,网络监控系统102可以提供用于在用户界面500的结果部分510中呈现给用户122的分析结果。

[0075] 额外地或备选地,在一些实施例中,应用118可以在用户界面500 的不同的区域

(即,结果部分)中呈现分析结果。通过举例而非限制的方式,用户界面500或应用118可以更新对应于所选择的节点 110或连接链路112 (或网络108的所选择的子集) 的拓扑映射502和/或列表504的部分以示出分析结果。在一个实施例中,用户界面 500或应用118可以突出拓扑映射502和/或列表504的更新的部分以允许用户122容易地检测所选择的节点110或连接链路112 (或网络108的所选择的子集) 的分析结果。用户界面500或应用118可以通过使用不同的颜色、样式 (例如,文本大小、字体、样式、等等) 突出更新的部分。额外地或备选地,例如,用户界面500或应用118可以通过使更新的部分闪光或闪烁突出更新的部分。

[0076] 另外,在一些实施例中,网络监控系统102可以允许用户122确定是否要执行对所选择的节点110或连接链路112 (或网络108的所选择的子集的一个或多个节点110和/或连接链路112) 的进一步分析以确定哪个问题 (过载、失灵、断开、等等) 最有可能由所选择的节点110或连接链路112 (或网络108的所选择的子集的一个或多个节点110和/或连接链路112) 经历。

[0077] 结论

[0078] 尽管已经以对结构特征和/或方法动作特定的语言描述了实施例,但是应理解权利要求不必限于所描述的特定特征或动作。相反,特定特征和动作被公开为实施要求保护的主题的示例性形式。

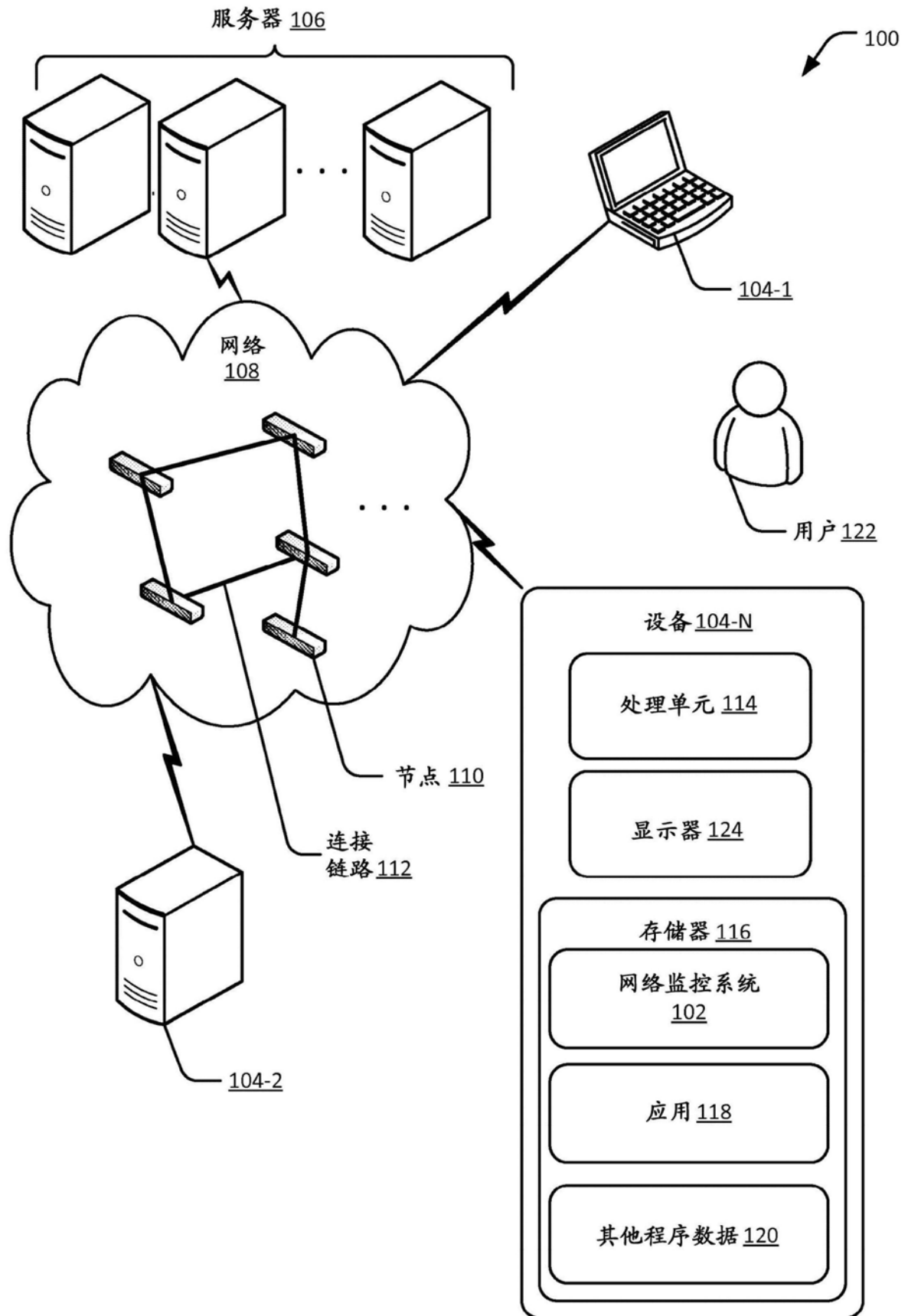


图1

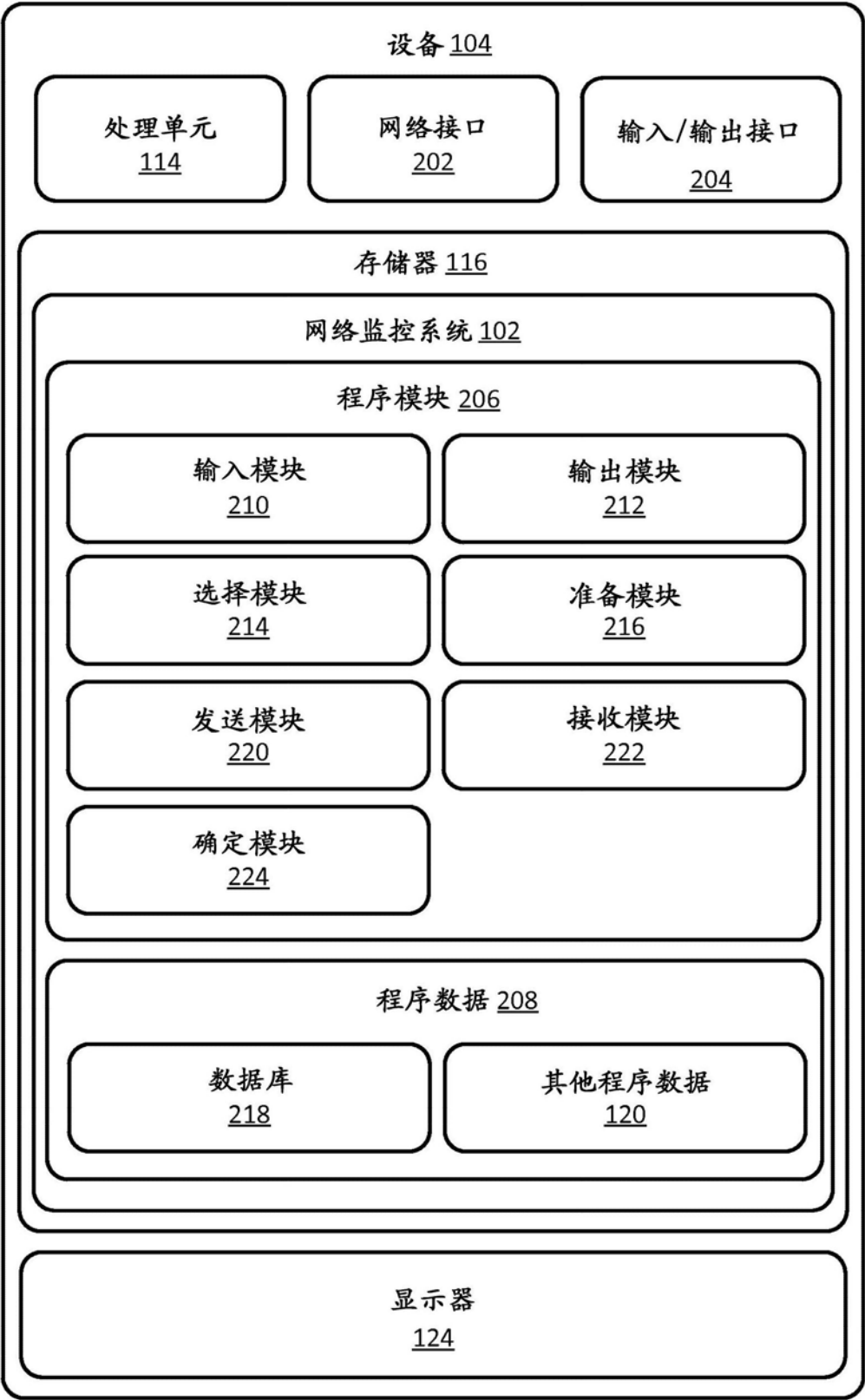


图2

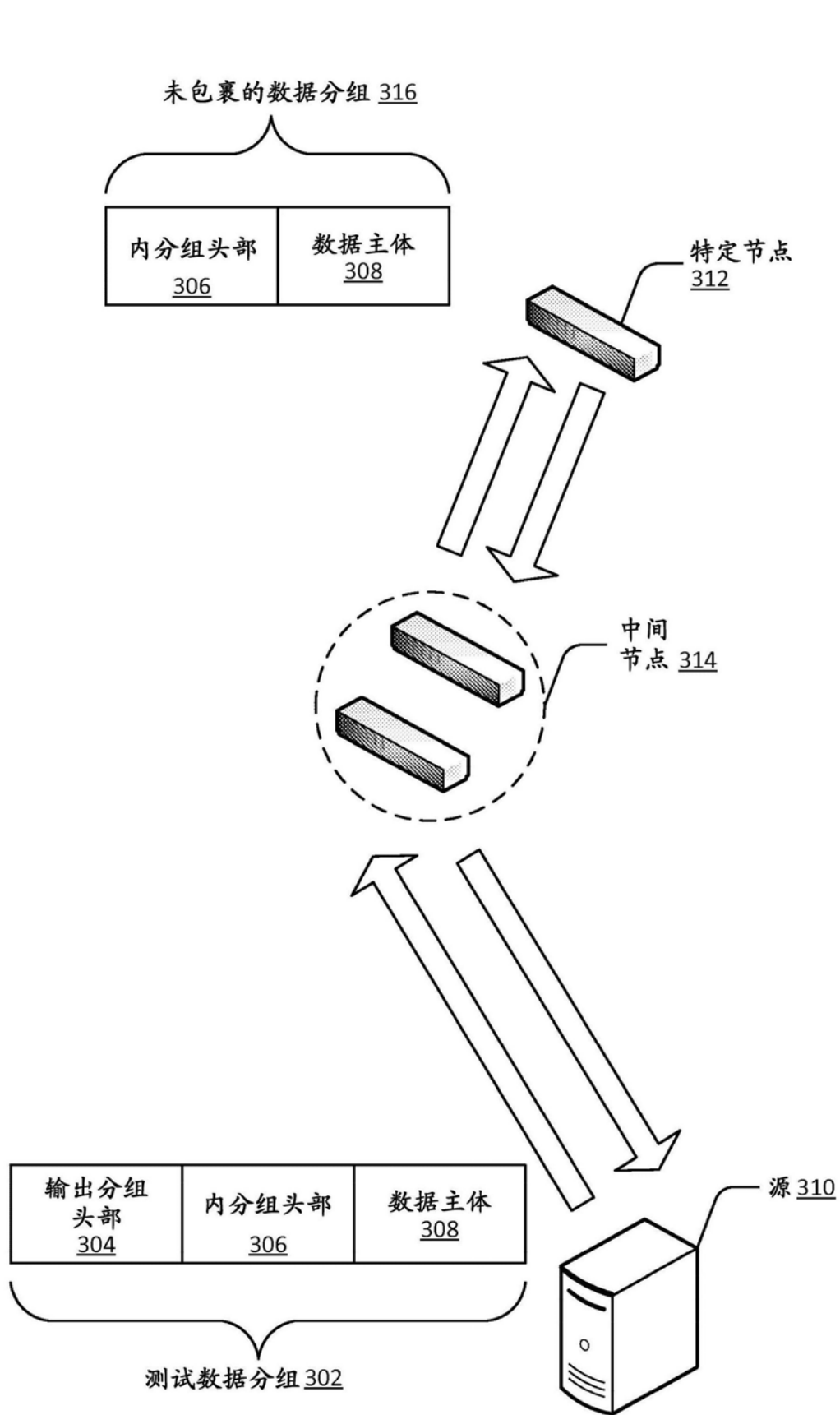


图3A

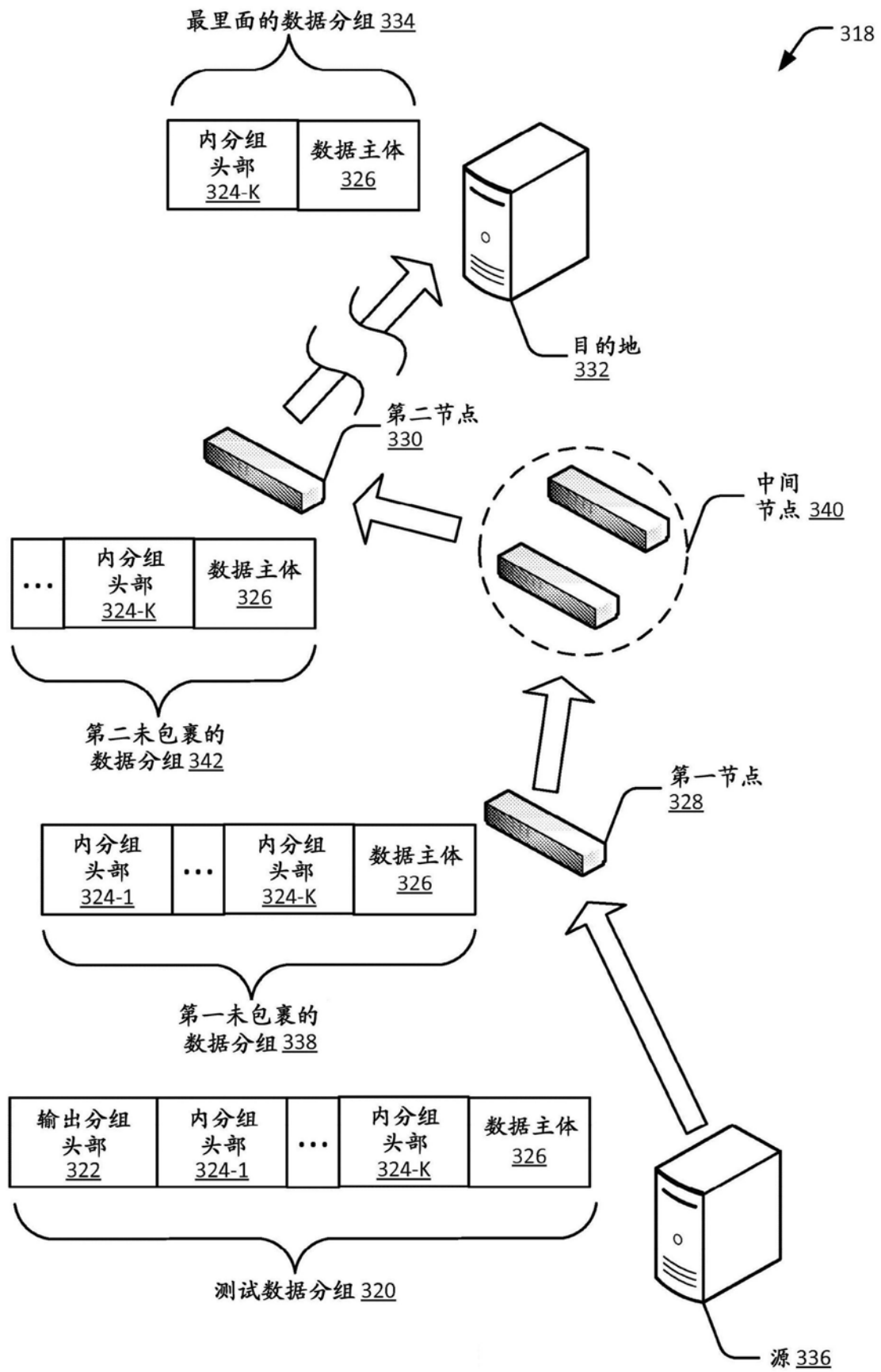


图3B

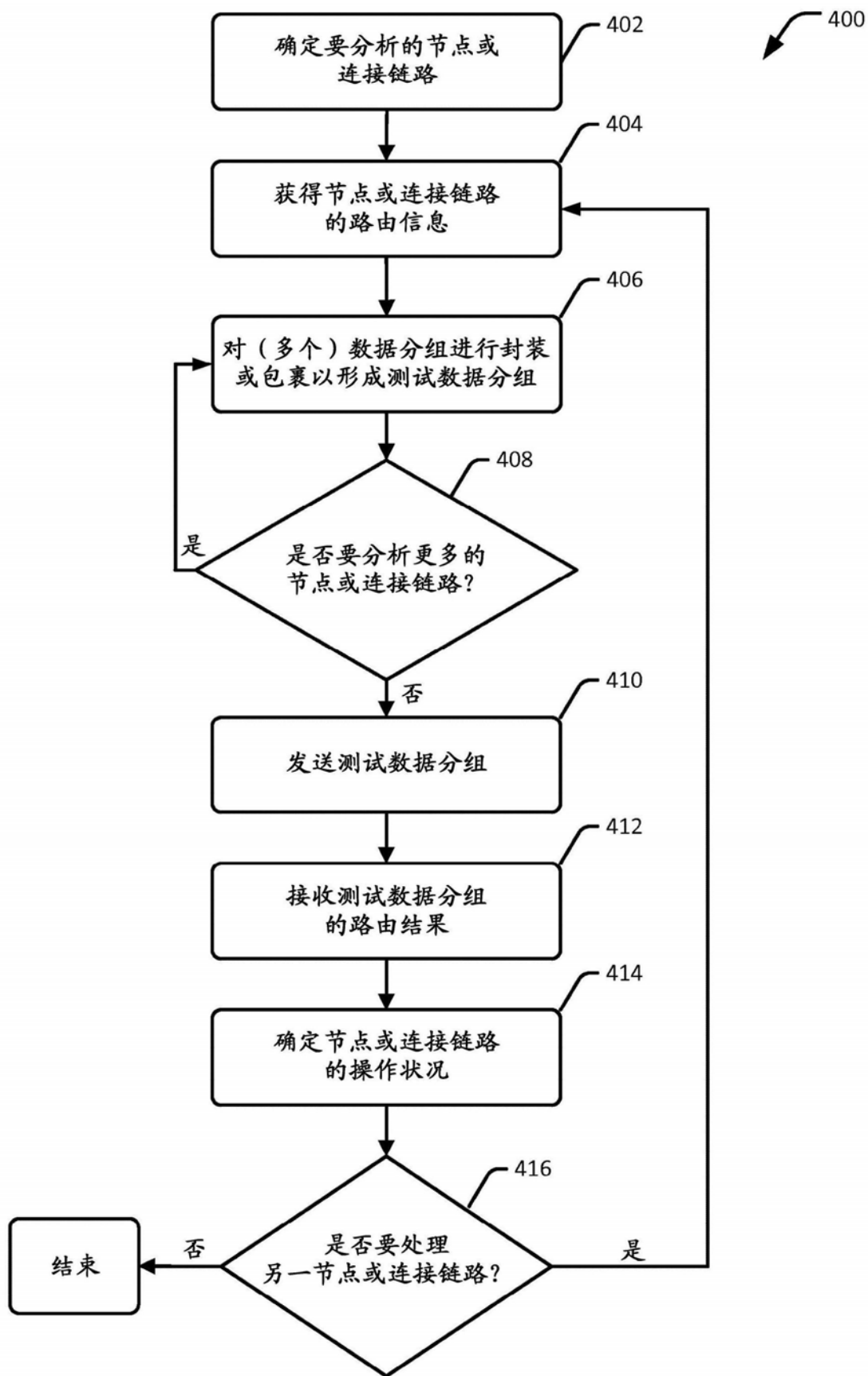


图4

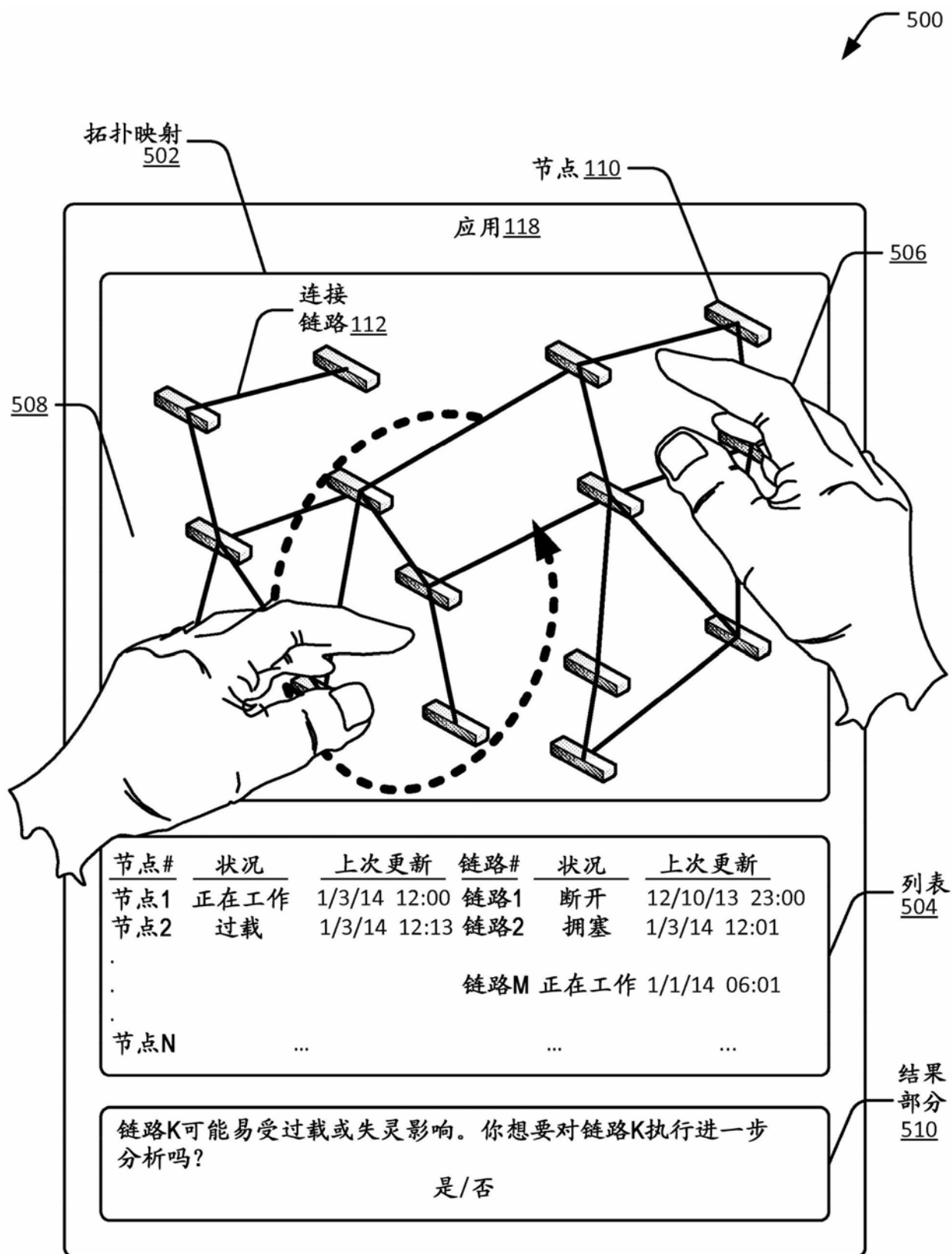


图5