(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 106797328 A
(43)申请公布日 2017.05.31

(21)申请号 201580047773.8

(22)申请日 2015.08.31

(30)优先权数据
14/475,927 2014.09.03 US

(85)PCT国际申请进入国家阶段日
2017.03.03

(86)PCT国际申请的申请数据
PCT/US2015/047633 2015.08.31

(87)PCT国际申请的公布数据
WO2016/036627 EN 2016.03.10

(71)申请人 微软技术许可有限责任公司
地址 美国华盛顿州

(72)发明人 张铭 吕国晗 袁利华

(74)专利代理机构 北京市金杜律师事务所 11256
代理人 王茂华

(51)Int.Cl.
H04L 12/26(2006.01)

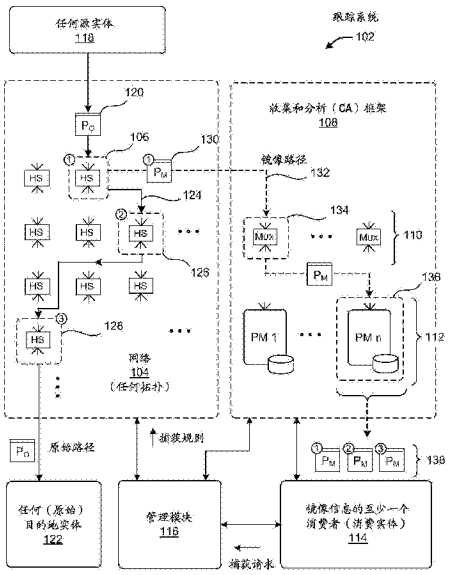
权利要求书2页 说明书16页 附图15页

(54)发明名称

收集和分析所选择的网络流量

(57)摘要

本文描述了用于调查网络的行为的跟踪系统。在操作中，网络中的每个交换机(或交换机的某一子集中的每个交换机)可以确定其处理的每个原始分组是否满足一个或多个分组检测规则。如果满足，则交换机生成镜像分组并将该分组发送给负载平衡器复用器，负载平衡器复用器转而将镜像分组转发给处理模块用于进一步分析。由交换机托管的分组检测规则可以被设计为基于任何环境特定目标来选择最感兴趣的分组的子集。作为这种行为的结果，在不被太多的信息压倒的情况下，跟踪系统可以有效地并且快速地指出网络的不期望的(和潜在期望的)行为。



CN 106797328 A

1.一种用于从网络收集分组的方法,包括:

在网络内的交换机处接收原始分组;

确定是否将所述原始分组镜像;

在做出将所述原始分组镜像的决定的情况下,基于所述原始分组生成镜像分组,所述镜像分组至少包括在所述原始分组中提供的信息的子集;

向负载平衡复用器发送所述镜像分组;以及

向由所述原始分组指定的目标目的地发送所述原始分组。

2.根据权利要求1所述的方法,其中所述确定是否将所述原始分组镜像包括:

关于分组检测规则来分析所述原始分组;

确定所述原始分组是否满足所述分组检测规则;以及

如果所述原始分组满足所述分组检测规则,则生成将所述原始分组镜像的指令。

3.根据权利要求2所述的方法,其中所述分组检测规则指定表示指定协议相关特性的每个原始分组要被镜像。

4.根据权利要求2所述的方法,其中所述分组检测规则指定来源于指定应用的每个原始分组要被镜像。

5.根据权利要求2所述的方法,其中所述分组检测规则对应于用户创建的分组检测规则,并且其中所述用户创建的分组检测规则指定满足用户指定匹配条件的每个原始分组要被镜像。

6.根据权利要求2所述的方法,其中所述分组检测规则指定以下每个原始分组要被镜像,所述每个原始分组表示在处理所述分组时所述交换机遇到指定条件。

7.根据权利要求1所述的方法,还包括:基于至少一个负载平衡考虑,从复用器候选的集合中选择所述复用器。

8.根据权利要求1所述的方法,其中所述复用器是硬件实现的复用器。

9.一个或多个用于分析从网络收集的分组的计算设备,包括:

接口模块,用于从至少一个处理模块接收多个镜像分组,

在原始分组满足分组检测规则集合中的至少一个分组检测规则的情况下,响应于处理所述原始分组,每个镜像分组由所述网络中的交换机产生并且被转发给所述至少一个处理模块,并且

每个镜像分组至少包括在所述原始分组中提供的信息的子集;

至少一个处理引擎,被配置为处理所述镜像分组以得到关于在所述网络中已发生或正发生的事件的至少一个结论;以及

动作采取模块,被配置为基于所述至少一个结论来采取动作。

10.一种用于在网络中使用的交换机,所述交换机对应于物理设备,所述交换机包括:

接收模块,被配置为接收原始分组;

匹配模块,被配置为通过确定所述原始分组是否满足分组检测规则集合中的至少一个分组检测规则,来确定是否将所述原始分组镜像;

镜像模块,被配置为在做出将所述原始分组镜像的决定的情况下,基于所述原始分组生成镜像分组,所述镜像分组至少包括在所述原始分组中提供的信息的子集;

镜像分组发送模块,被配置为向负载平衡复用器发送所述镜像分组;以及

原始分组发送模块，被配置为向由所述原始分组指定的目标目的地发送所述原始分组。

11.根据权利要求3所述的方法，其中所述协议相关特性由传输层协议产生的至少一个信息项表示。

12.根据权利要求3所述的方法，其中所述协议相关特性由路由协议产生的至少一个信息项表示。

13.根据权利要求8所述的方法，其中所述指定条件指示所述原始分组要被所述交换机丢弃。

14.根据权利要求1所述的方法，还包括：

在所述复用器处接收所述镜像分组；

基于至少一个负载平衡考虑，从处理模块候选的集合中选择处理模块；以及

向选择的所述处理模块发送所述镜像分组。

15.根据权利要求10所述的交换机，还包括：目标复用器选择模块，被配置为基于至少一个负载平衡考虑来从复用器候选的集合中选择所述复用器。

收集和分析所选择的网络流量

背景技术

[0001] 通常难以确定发生在网络内的故障和其它异常事件的原因。这种困难产生于现代网络的复杂性，加上这样的网络在任何给定时间处理的大量信息。有经验的分析者可以通过调查网络的那些假定为最有可能故障的组件(例如，通过检查由那些组件记录的控制信息)的行为来解决这个问题。然而，分析者不能确保被审查的信息将揭示问题的根源。分析者可以扩大分析范围以解决这个问题，但是这样的策略可以导致太多的信息压倒分析者。

发明内容

[0002] 本文描述了用于调查网络的行为的跟踪系统。在操作中，网络中的每个交换机(或网络中的至少一些交换机中的每个交换机)可以确定其处理的每个原始分组是否满足一个或多个分组检测规则。如果满足，则交换机可以生成镜像分组。镜像分组至少包括原始分组中的信息的子集。然后，交换机可以将镜像分组转发给负载平衡复用器。交换机还将原始分组以未改变的形式发送给由原始分组指定的目标目的地。

[0003] 在接收镜像分组时，基于至少一个负载平衡考虑，复用器可以从候选处理模块集合中选择处理模块。然后，复用器向所选择的处理模块发送镜像分组，其中使用一个或多个处理引擎对其分析。

[0004] 考虑任何应用特定的目标，由交换机托管的分组检测规则可以被设计为选择被认为是高兴趣值的分组的子集。作为该行为的结果，在分析者不被太多的信息压倒的情况下，跟踪系统可以有效地并且快速地指出网络的不期望的(和潜在期望的)行为。

[0005] 上述方法可以表现在各种类型的系统、设备、组件、方法、计算机可读存储介质、数据结构、图形用户界面呈现、制品等中。

[0006] 提供本发明内容来以简化的形式介绍概念的选择；这些概念在下面的具体实施方式中进一步描述。本发明内容不旨在标识所要求保护的主题的关键特征或必要特征，也不旨在用于限制所要求保护的主题的范围。

附图说明

[0007] 图1示出了跟踪系统的一个示例的概览。跟踪系统从网络中提取所选择的信息以用于分析。

[0008] 图2示出了图1的跟踪系统的一个非限制性实施方式。

[0009] 图3示出了被配置为执行镜像功能的网络中的交换机的一个实施方式。该配置的交换机是由图1的跟踪系统使用的镜像功能的一个组件。

[0010] 图4示出了与图1的跟踪系统的另一组件对应的复用器的一个实施方式。

[0011] 图5示出了图3的交换机的复用行为。

[0012] 图6示出了图4的复用器的复用行为。

[0013] 图7示出了根据一个实施方式的图4的复用器可以利用来执行其复用功能的说明性表数据结构。

[0014] 图8示出了由图3的交换机输出的信息的一个示例。

[0015] 图9示出了由图4的复用器输出的信息的一个示例。

[0016] 图10示出了处理模块的一个实施方式，该处理模块是图1的跟踪系统的另一组件。

[0017] 图11示出了消费实体的一个实施方式，该消费实体是与图1的跟踪系统交互的组件。

[0018] 图12示出了管理模块的一个实施方式，该管理模块是图1的跟踪系统的另一组件。

[0019] 图13示出了解释图3的交换机的一种操作方式的处理。

[0020] 图14示出了解释匹配模块的一种操作方式的处理，该匹配模块是图3的交换机的组件。

[0021] 图15示出了解释图4的复用器的一种操作方式的处理。

[0022] 图16示出了解释图10的处理模块的一种操作方式的处理。

[0023] 图17示出了解释图11的消费实体的一种操作方式的处理。

[0024] 图18示出了解释图12的管理模块的一种操作方式的处理。

[0025] 图19示出了可用于实现前述附图中所示的特征的任何方面的说明性计算功能。

[0026] 贯穿公开内容和附图，使用相同的附图标记来指代相同的组件和特征。系列100的附图标记指代最初在图1中找到的特征，系列200的附图标记指代最初在图2中找到的特征，系列300的附图标记指代最初在图3中找到的特征，依此类推。

具体实施方式

[0027] 本公开组织如下。部分A描述了用于例如通过选择性地提取的流经网络的特定类型的分组来选择性地收集和分析网络流量的说明性跟踪系统。部分B阐述了解释部分A的跟踪系统的操作的说明性方法。部分C描述了可用于实施部分A和部分B中描述的特征的任何方面的说明性计算功能。

[0028] 作为前序事项，一些附图描述了在一个或多个结构组件(不同地被称为功能、模块、特征、元件等)的上下文中的概念。图中所示的各种组件可以以任何方式通过任何物理和有形机构(例如，通过在计算机设备上运行的软件、硬件(例如，芯片实现的逻辑功能)等、和/或其任何组合)来实现。在一种情况下，图中所示的各种组件分离成不同的单元，可以反映在实际实施方式中对应的不同物理和有形组件的使用。备选地或附加地，图中所示的任何单个组件可以由多个实际物理组件来实现。备选地或附加地，图中的任何两个或更多个分离组件的描绘可以反映由单个实际物理组件执行的不同功能。将依次描述的图19提供了关于图中所示的功能的一个说明性物理实施方式的附加细节。

[0029] 其它附图以流程图形式描述了概念。在该形式中，特定操作被描述为构成以特定顺序执行的不同块。这样的实施方式是说明性的而非限制性的。本文描述的特定块可以被分组在一起并且在单个操作中执行，特定块可以被分解为多个组件块，并且可以以与本文所示的顺序不同的顺序(包括执行块的并行方式)来执行特定块。流程图中所示的块可以以任何方式通过任何物理和有形的机构(例如，通过在计算机设备上运行的软件、硬件(例如，芯片实现的逻辑功能)等、和/或其任何组合)来实现。

[0030] 关于术语，短语“配置为”包括可以构造任何种类的物理和有形的功能以执行标识的操作的任何方式。功能可以被配置为使用例如在计算机设备上运行的软件、硬件(例如，

芯片实现的逻辑功能)等、和/或其任何组合来执行操作。

[0031] 术语“逻辑”包括用于执行任务的任何物理和有形功能。例如,流程图中所示的每个操作对应于用于执行该操作的逻辑组件。可以使用例如在计算机设备上运行的软件、硬件(例如,芯片实现的逻辑功能)等、和/或其任何组合来执行操作。当由计算设备实现时,逻辑组件表示作为无论以何种方式实现的计算系统的物理部分的电组件。

[0032] 以下的解释可以将一个或多个特征标识为“可选的”。这种类型的陈述不应被解释为可以被认为是可选的特征的详尽指示;即,虽然在文本中没有明确地标识,其它特征可以被认为是可选的。此外,单个实体的任何描述不旨在排除使用多个这样的实体;类似地,多个实体的描述不旨在排除使用单个实体。此外,虽然描述可以将特定特征解释为执行所标识的功能或实施所标识的机构的备选方式,但是特征也可以以任何组合来组合在一起。最后,术语“示例性”或“说明性”是指潜在地许多实施方式之中的一个实施方式。

[0033] A.说明性跟踪系统

[0034] A.1.概览

[0035] 图1示出了跟踪系统102的一个示例的概览。跟踪系统102提取关于通过网络104传输的所选择的分组的信息,然后分析那些分组。在一个使用场景中,分析者可以使用由跟踪系统102提供的信息,来调查异常事件或不期望的事件。在其它情况下,分析者可以使用跟踪系统102提供的信息,来调查网络104中的期望行为。总体上,由跟踪系统102提供的信息可以提供关于正在研究的无论什么事件的原因的洞察。

[0036] 在其它潜在的益处中,跟踪系统102从网络104挑选信息的选择性减少了呈现给人类分析者或其它消费者的“噪声”的量,从而便于他或她的调查。其还有助于跟踪系统的可伸缩性和整体效率。以下描述的跟踪系统102的其它方面进一步有助于由跟踪系统102提供的分组收集功能的可伸缩性和效率。

[0037] 网络104由多个硬件交换机(诸如,代表性交换机106)组成。例如,每个交换机可以由专用集成电路(ASIC)等提供的逻辑功能来实现。虽然未示出,但是网络104附加地或备选地可以包括一个或多个软件实现的交换机。基于一个或多个路由考虑,每个交换机以其构造的无论什么方式来执行向目的地路由从源接收的输入分组的主要功能。源可以对应于沿多跳路径的另一“上游”交换机、或分组的最终起始点。类似地,目的地可以对应于沿路径的另一交换机、或分组的最后目的地。

[0038] 网络104在图1中仅以高级形式描绘。实际上,网络104可以具有任何拓扑。拓扑确定网络104中的交换机的选择以及那些交换机的布置(和互连)。此外,网络104可以在任何环境中被使用。在一种情况下,例如,网络104可以用于在数据中心内路由分组,并且在外部实体和数据中心之间路由分组。在另一种情况下,网络104可以在企业环境中被使用。在另一种情况下,网络104可以在中间上下文中例如通过在两个或更多个环境之间(例如,在两个或更多个数据中心之间等)路由信息来操作。还可能是其它应用。

[0039] 跟踪系统102具有两个主要组件:镜像功能以及收集和分析(CA)框架108。镜像功能共同表示由网络104中的所有相应交换机提供的镜像机构。在其它实施方式中,交换机的子集(但不是所有交换机)包括镜像机构。当其托管的交换机接收与一个或多个分组检测规则匹配的原始分组时,每个镜像机构生成镜像分组。镜像分组包含从原始分组提取的信息(诸如,原始分组的报头信息)的子集。镜像分组还包含指定新目的地地址(与原始分组的原

始目的地地址相比)的新报头。然后,根据由镜像机构分配的地址,交换机向CA框架108传递镜像分组。CA框架108然后以各种实施方式特定的方式来处理镜像分组。

[0040] 更具体地,交换机可以将镜像分组发送给从一个或多个复用器(Mux)110的集合中选择的复用器。基于至少一个负载平衡考虑,所选择的复用器然后可以将镜像分组发送给处理模块(PM)112集合中的一个处理模块。然后,所选择的处理模块可以使用一个或多个处理引擎来处理镜像分组(以及其它先前接收的镜像分组)。

[0041] 至少一个消费实体114可以与处理模块112交互,以获得镜像分组。然后,消费实体114可以使用一个或多个处理引擎对镜像分组执行任何专用分析。在一种情况下,消费实体114可以对应于在计算设备上运行的以自动方式操作的分析程序。在另一种情况下,消费实体114可以对应于在人类分析者的指导下在计算设备上运行的分析程序。在一些情况下,消费实体114还隶属于特定应用。考虑到这种关联,消费实体可以对网络中影响其自己的应用的事件特别感兴趣。

[0042] 管理模块116可以控制跟踪系统102的任何方面。例如,管理模块116可以指示网络104中的交换机加载特定的分组检测规则,以用于捕获流经网络104的特定类型的分组。管理模块116还可以与任何消费实体交互。例如,消费实体114可以标识网络中的问题,并且作为响应,请求管理模块116将分组检测规则传播给交换机;作为这些规则的结果而产生的镜像分组将帮助消费实体114标识问题的原因。

[0043] 为了澄清上述解释,图1描绘了通过网络104的一个原始分组连同其镜像对应物的流。稍后的子部分(下面)提供了关于在描述图1的代表性流中引入的每个操作的附加说明性细节。

[0044] 如图所示,任何源实体118将原始分组(P_0)120发送到网络104中,其中最终旨在将其发送给任何目的地实体122。例如但不限于,源实体118可以对应于第一计算设备以及目的地实体122可以对应于第二计算设备。更具体地,例如,目的地实体122可以对应于位于数据中心中的、托管特定应用的服务器计算设备。源实体118可以对应于出于任何目的希望与应用交互的任何计算设备。

[0045] 如本文所使用的术语,分组是指任何信息单元。在一个特定实施方式中,原始分组120对应于具有由IP协议指定的报头和有效载荷的因特网协议(IP)分组。更具体地,原始分组可以提供标识目的地实体的虚拟IP(VIP)地址。反过来,目的地实体122可以与直接IP(DIP)地址相关联。除了其它功能,网络104中的至少一个组件将VIP地址映射到目的地实体122的适当DIP地址。

[0046] 网络104可以使用任何路由协议,通过其交换结构,来将原始分组120从源实体118向目的地实体122路由。可以在建立路由中起作用的一个这样的协议是如在RFC 4271中定义的边界网关协议(BGP)。另外注意,当原始分组120遍历其路由时,在原始分组120上操作的网络104中的不同组件可以向(或从)原始分组120附加(或移除)各种封装报头。

[0047] 更具体地,图1描绘了其中原始分组120遍历具有多个段或跳的路径124的仅说明性情况。在第一段中,原始分组120被路由给交换机106。在第二段中,原始分组120被路由给另一交换机126。在第三段中,原始分组120被路由给另一交换机128。在第四段中,原始分组120被路由给目的地实体122。在实际实践中,路径124可以具有任何数量的跳(包括单跳),并且可以遍历由交换机定义的交换结构中的任何交换机。此外,如上所述,网络104可以使

用一个或多个隧道协议，来将原始分组封装在其它封闭分组中；这样的规定本质上是环境特定的，并且从图1中省略以便于解释。

[0048] 每个交换机(或交换机的子集中的每一个)上的镜像机构分析原始分组，以首先确定其是否满足一个或多个分组检测规则。如果满足，则镜像机构将生成与原始分组对应的镜像分组，同时保持原始分组本身完整，并且不干扰沿路径124的原始分组的路由。

[0049] 例如，考虑交换机106的操作。(当其它交换机处理原始分组120时，其它交换机将表现出相同的行为。)假设交换机106首先确定原始分组120匹配至少一个分组检测规则。然后，其生成镜像分组130。然后，交换机106可以沿路径132将镜像分组130转发给指定的目的地(对应于复用器110中的一个)。更具体地，沿路径132的不同的传播实体可以将封装报头附加到镜像分组130(或去除封装报头)。但是，为了便于图示和解释，图1将镜像信息简单地称为镜像分组130。

[0050] 更具体地，在一个实施方式中，交换机106可以应用至少一个承载考虑，来选择复用器110集合中的复用器。例如，假设交换机106选择复用器134。在其它实施方式中，CA框架108可以提供单个复用器；在那种情况下，交换机106将镜像分组130发送给该复用器，而不在多个可用的复用器之间进行选择。

[0051] 复用器134基于至少一个承载考虑，来执行将镜像分组130进一步路由给处理模块112中的一个的功能。复用器134还将选择目标处理模块，使得关于通过网络104的流的镜像分组被发送给同一处理模块。复用器134本身可以以任何方式实现。在一种情况下，复用器134可以对应于硬件实现的复用器，诸如由专用集成电路(ASIC)提供的逻辑功能。在另一种情况下，复用器134对应于软件实现的复用器，诸如在服务器计算设备上运行的复用程序。在其它情况下，复用器110的集合可以包括硬件复用器和软件复用器的组合。

[0052] 假设复用器134将镜像分组130路由给特定处理模块136。在一个实施方式中，处理模块136可对应于服务器计算设备。在接收时，处理模块136可以对镜像分组130执行各种操作。在一个这样的功能中，处理模块136可以将镜像分组与关于相同路径124(如果存在)的其它分组相关联，并且然后将镜像分组以其由交换机创建的顺序进行排序。例如，在原始分组遍历其路径124完成时，处理模块136可以生成分组序列138，分组序列138对应于由交换机106、126和128创建的镜像分组的序列。

[0053] 消费实体114可以提取由处理模块136存储的任何分组相关的信息，并且然后以任何方式分析该信息。以下描述提供可以由消费实体114执行的分析的示例。图1具体示出了消费实体114至少提取或以其它方式访问与原始分组120通过网络104的路径124相关联的序列138。在其它情况下，消费实体114可以请求并接收特定的镜像分组，而不是分组序列。

[0054] A.2.特定网络环境的示例

[0055] 图2示出了环境202，环境202包括图1的跟踪系统102的一个非限制性实施方式。环境202对应于包括多个计算设备204(诸如多个服务器)的数据中心。网络206允许数据中心内的计算设备204与数据中心内的其它计算设备通信。网络206还允许外部实体208与计算设备204交互。诸如因特网的广域网210可以将数据中心的网络206与实体208耦合。

[0056] 网络206可以具有任何拓扑。如图2的特定和非限制性示例所示，网络206包括胖树(fat-tree)分级拓扑中的多个交换机。非限制性地，交换机可以包括核心交换机212、聚集交换机214、机架顶(TOR)交换机216等。此外，网络206可以将计算设备204组织成诸如容器

218和220的容器。实际的数据中心可以包括许多更多的交换机和计算单元;图2仅示出了数据中心环境的功能的代表性和简化示例。

[0057] 网络206中的所有交换机或其某一子集包括镜像机构。在它们处理原始分组时(假定原始分组满足一个或多个分组检测规则),镜像机构生成镜像分组。然后,镜像机构将镜像分组转发给收集和分析(CA)框架222。

[0058] 更具体地,CA框架222可以提供用于处理镜像分组的收集和分析的专用设备。换言之,CA框架222可以在通过网络206的原始分组的路由中不执行任何角色。(但是在其它实施方式中,CA框架222可以执行路由原始分组和处理镜像分组的双重角色)。在一种情况下,CA框架222包括一个或多个复用器224。复用器可以对应于硬件复用器,并且更具体地,可以对应于已经被重新配置为执行复用角色的硬件交换机。备选地或附加地,至少复用器224的子集可以对应于软件实现的复用器(例如,对应于一个或多个服务器计算设备)。

[0059] 复用器224可以被耦合到网络206的顶层交换机212和/或被耦合到其它交换机。此外,复用器224可以被直接耦合到一个或多个处理模块226。备选地,如图2所示,复用器224可以使用任何连接拓扑经由交换机228被连接到处理模块。

[0060] A.3.具有镜像能力的说明性交换机

[0061] 图3示出了具有镜像能力的说明性交换机302,意味着其具有生成和转发作为原始分组的镜像对应物的分组的能力。如上所述,交换机302可以被实现为硬件单元(例如,作为ASIC)。

[0062] 从高级的角度来看,交换机302可以包括用于执行三个主要功能的功能。功能304允许交换机302执行其将所接收的原始分组转发给目标目的地的传统角色。功能306执行交换机操作的镜像方面。并且功能308执行各种管理功能。更具体地,为了便于解释,图3将这三个功能(304、306、308)示为三个单独的域。然而,在一些实施方式中,单个物理模块可以执行归于图3所示的不同域的两个或更多个功能。

[0063] 从功能304开始,接收模块310从任何源接收原始分组120。源可以对应于图1的源实体118或另一个“上游”交换机。路由选择模块312选择对应于下一跳314的原始分组的下一目的地。转而,下一跳314可以对应于原始分组的最终目标目的地、或沿多跳路径的另一个“下游”交换机。在选择下一跳314时,路由选择模块312可以查询在数据存储器316中提供的路由信息。在选择下一跳314时,路由选择模块312还可以使用任何协议(诸如BGP)。发送模块318将原始分组发送给下一跳314。尽管图3中未明确示出,在将原始分组发送给下一跳314之前,发送模块318可以可选地使用任何封装协议来将原始分组封装在另一分组中。

[0064] 关于镜像功能306,匹配模块320确定已接收的原始分组120是否与存储在数据存储器322中的任何分组检测规则匹配。以下将阐述说明性规则。如果原始分组120满足任何一个或多个分组检测规则,则镜像模块324生成镜像分组326。如上所述,镜像模块324可以通过从原始分组120提取信息的子集(诸如,原始分组的报头)来产生镜像分组326。镜像模块324还可以添加原始分组120中不存在的信息,诸如在处理原始分组120的过程中由交换机302本身产生的元数据。在一些实施方式中,镜像模块324可以使用可用的分组-复制技术(诸如,由加利福尼亚州圣何塞的思科系统公司提供的封装远程交换端口分析器(ERSPAN)技术),来创建镜像分组326。

[0065] 复用器选择模块328从(图1的)复用器110集合中选择要向其发送镜像分组326的

复用器。在图3的上下文中，假设复用器选择模块328选择了复用器332。例如，复用器选择模块328可使用哈希算法来哈希由镜像分组传达的信息项(诸如，在原始分组的IP报头(被复制到镜像分组中的信息)的报头中提供的不同信息项)的任何元组。哈希操作产生哈希结果，哈希结果转而可以被映射到特定的复用器。具有镜像机构的所有交换机采用相同的哈希功能。总体来说，哈希操作具有在可用的复用器110集合上扩散镜像分组的效果。数据存储器330可以提供复用器选择模块328在执行其操作时可以参考的信息；例如，数据存储器330可以例如通过提供其相应地址来标识可用的复用器110。

[0066] 发送模块334将镜像分组发送给复用器332。在一种情况下，发送模块334可以使用任何隧道协议(诸如通用路由封装(GRE))来将镜像分组封装在隧道分组中，然后在隧道协议报头的“顶部”附加复用IP报头。GRE例如在RFC 2784中被描述。发送模块318产生封装的镜像分组336。

[0067] 关于管理功能308，交换机302可以包括用于处理其它相应任务的其它控制模块338。例如，路由管理模块可以执行诸如以下的任务：向网络中的其它交换机广播交换机302的存在、确定其它交换机的存在、更新数据存储器(316、330)中的路由信息等。接口模块340可以从管理模块116接收管理信息和其它指令。

[0068] 现在更详细地参考匹配模块320，该组件可以将原始分组120与不同类型的分组检测规则进行比较。以下解释提供分组检测规则的代表性示例。这样的列表是以说明性的而不是限制性的精神提供的；其它实施方式可以依赖于下面未提及的附加类型的分组检测规则。

[0069] 第一类型的分组检测规则可以指定：如果原始分组120(诸如，通过例如在原始分组120的报头和/或主体中，包含指定的协议相关的一个或多个信息项)表示协议相关的特性，则原始分组120要被镜像。例如，该信息可以对应于由传输级错误检查协议(诸如传输控制协议(TCP))产生的标记。在另一种情况下，触发条件可以对应于由诸如BGP的路由协议产生的一个或多个信息项。

[0070] 第二类型的分组检测规则可以指定：如果原始分组120例如通过包含应用相关的一个或多个信息项来表示其起源于特定应用，则原始分组120要被镜像。应用相关的信息项可以对应于标记、代码、地址等。应用可以将其在正常执行的过程中产生的信息项添加到分组。

[0071] 第三类型的分组检测规则对应于用户创建的分组检测规则。这类型的规则指定：如果原始分组满足用户指定的匹配条件，则原始分组要被镜像。用户可以对应于网络管理员、测试工程师、应用或系统开发者、网络104的终端用户等。例如，用户可以创建规则，该规则指定包含所标识的报头信息的任何分组要被镜像。

[0072] 第四类型的分组检测规则可以指定：如果原始分组120表示当交换机302处理原始分组120时遇到特定条件或情况，则原始分组120要被镜像。例如，可以是在检测到由交换机302添加的原始分组中的信息项时触发规则；该信息项指示交换机302在处理原始分组120时遇到错误情况或其它事件。

[0073] 更具体地，例如，由交换机302用于转发原始分组120的功能304可以被实现为处理管线，其中对原始分组120串行地执行一系列操作。在一个或多个阶段，错误检测功能342可以检测其对原始分组120的处理中的错误情况。例如，在分析的接收或路由选择阶段期间，

错误检测功能342可以确定原始分组120已被损坏，并且因此不能被有意义地解译，并且因此不能被转发给下一跳314。作为响应，错误检测功能342可以向原始分组120附加标记或其它信息项，指示其将被丢弃。然后，功能304的处理管线的稍后阶段可以执行丢弃原始分组120的明确步骤。

[0074] 然而，在发生丢弃之前，匹配模块320可以检测已添加的信息项的存在，并且作为响应，镜像模块324可以将具有向其添加的信息的原始分组120(即使，如所述的，该分组将最终被丢弃)镜像。这样的镜像分组在分析期间提供有用的信息，以标识分组丢弃的原因。

[0075] 匹配模块320包括输入344，以通常指示匹配模块320可以在由交换机302执行的处理中的任何阶段、而不必仅在接收阶段将原始分组120与分组检测规则进行比较。因此，在一些情况下，原始分组120在初始接收时可以不包含触发分组检测规则的特定信息字段；但是交换机302本身可以在其处理的稍后阶段添加触发信息项，提示匹配模块320稍后成功地匹配修改的分组与规则之一。

[0076] 此外，跟踪系统102可以提供用于检测分组丢弃的附加技术。例如，处理模块或消费实体可以通过分析沿原始分组遍历网络的路径产生的镜像分组的序列来检测分组丢弃的存在。如原始分组未到达其旨在的最后目的地的事实所证明的，分组丢弃可以本身表现为序列的过早截断。或者，序列可以在序列中揭示“洞”，该“洞”指示期望跳目的地来接收分组，但是该跳目的地没有接收(尽管在那种情况下，分组可以最终仍然到达其最后目的地)。

[0077] 在其它情况下，交换机302可以向原始分组120添加元数据信息，以指示在处理原始分组120时，交换机302遇到一些其它条件，其中该条件不一定与错误相关联。

[0078] 第五类型的分组检测规则可以指定：如果原始分组120指定要被镜像的所标识的服务类型，则原始分组120要被镜像。例如，该类型的分组检测规则可以基于由原始分组120指定的差分服务代码点(DSCP)值等来决定将原始分组120镜像。

[0079] 第六类型的分组检测规则可以指定：如果原始分组120由ping相关应用产生，则原始分组120要被镜像。更具体地，ping相关应用通过将原始分组发送给目标实体来操作，此时请求目标实体发送对原始分组的响应。

[0080] 为了重复，其它环境可以应用附加类型的分组检测规则。例如，在检测到特定IP源和/或目的地地址、或TCP或UDP源和/或目的地端口等时，可以触发其它规则。此外，在一些情况下，在检测到原始分组120中的单个信息项(诸如，原始分组120中的单个标记)时，可以触发分组检测规则。但是在其它情况下，分组检测规则可以是在检测到原始分组120中的两个或更多个信息项的组合(诸如，原始分组120中的两个标记的组合)时触发。此外，在上述任何情况下，信息项可以出现在原始分组120的报头和/或主体中。备选地或附加地，分组检测规则可以由原始分组120的其它特性(即，除了特定信息项在原始分组120的报头或主体中的存在或不存在之外的一些特性)触发。例如，在检测到原始分组120被损坏、或具有某一其它错误、或满足某一其它匹配条件时，可以触发规则。

[0081] 在附图的序列中暂时向前跳跃，图5示出了由图3的复用器选择模块328执行的复用功能。如那里所示，复用器选择模块328使用某一扩散算法506(诸如，对原始分组IP报头的某一元组操作的哈希算法)将原始分组502映射到复用器集合504中的一个复用器。

[0082] 更具体地，在一种情况下，复用器中的每一个可以由其自己的唯一VIP地址表示。因此，复用器选择模块328具有在不同的VIP地址中进行选择的作用。在另一种情况下，复用

器的集合可以具有不同的直接DIP地址，但是具有相同的VIP地址。可以使用任何负载平衡协议(诸如，等价多路径路由(ECMP))来在复用器之间扩散镜像分组。ECMP在RFC 2991中定义。

[0083] 图8示出了在能够镜像的交换机302的输出处生成的经封装的镜像分组336的说明性结构。经封装的镜像分组336包括由镜像模块324产生的上述镜像分组326，例如，对应于原始分组120中的信息的子集(例如通过至少提供原始分组120的报头)。封装外部字段包括镜像隧道报头802(诸如GRE隧道报头)。下一个封装外部字段包括镜像IP报头804。其它实施方式可以采用封装镜像分组326的其它方式。

[0084] A.4.说明性复用器

[0085] 图4示出了复用器402的一个实施方式。复用器402可以对应于图1所示的复用器110集合中的一个。或者，复用器402可以对应于由跟踪系统102提供的唯一的复用器。复用器402可以对应于硬件实现的设备或软件实现的设备或其某种组合。在前一种情况下，硬件复用器可以对应于已经被重新编程和改变用途以执行复用功能的商品交换机。或者，硬件复用器可以对应于被构造为执行下面描述的功能的定制设计的组件。

[0086] 复用器402包括用于执行实际复用功能的功能404以及用于管理复用功能的功能406。例如，功能404可以包括用于接收镜像分组412的接收模块410。(更精确地，镜像分组412对应于在交换机302的输出处产生的封装镜像分组336的类型，但是为了简洁，其在下面被简称为“镜像分组”412。)功能404还可以包括用于从候选处理模块112集合中选择处理模块的PM选择模块414。在执行其操作时，PM选择模块414查询数据存储器416中的路由信息。假设PM选择模块414选择将镜像分组412发送给PM 418。发送模块420将镜像分组412发送给PM 418。在这样做时，发送模块420可以将镜像分组412封装在隧道协议报头(例如GRE报头)中，然后将该信息封装在又一个外部IP报头中，以产生封装的镜像分组422。控制相关模块424可以管理复用器的操作的任何方面。例如，控制相关模块424可以提供地址信息，用于存储在数据存储器416中，该地址信息标识PM的地址。接口模块426例如通过从管理模块116接收用于配置复用器402的操作的控制指令，与(图1的)管理模块116交互。

[0087] PM选择模块414可以基于任何负载平衡考虑从PM 112集合中选择PM。在一种方法中，PM选择模块414使用哈希算法来将使用原始分组的报头而包含的信息项哈希，该信息项是也在镜像分组中捕获的信息。得到的哈希映射到处理模块112中的一个。哈希算法还确保属于相同分组流的分组被映射到相同的处理模块。跟踪系统102可以通过从原始分组中选择输入信息项(其用作哈希算法的输入密钥)来实现该结果，输入信息项将在原始分组遍历穿过网络104的路径时保持相同，或者当被哈希算法作用时，输入信息项将以其它方式产生相同的输出哈希值。此外，跟踪系统102在所有复用器110上部署相同的哈希算法。

[0088] 图6描绘了由图4的PM选择模块414执行的复用功能。如那里所指示的，PM选择模块414使用某一扩散算法606(诸如上述哈希算法)将所接收的镜像分组602映射到PM集合604中的一个。

[0089] 在一种情况下，处理模块112中的每一个可以由其自己的唯一VIP地址表示。PM选择模块414因此具有在不同VIP地址之中进行选择的作用。在另一种情况下，处理模块112的集合可以具有不同的直接地址(DIP)，但是具有相同的VIP地址。任何负载平衡协议(诸如ECMP)可以被用于在处理模块112之间扩散镜像分组。

[0090] 图7示出了PM选择模块414可以用于执行其复用功能的说明性表数据结构702。数据存储器416可以存储表数据结构702。更具体地，图7对应于其中通过对硬件交换机重新编程和改变用途来产生复用器402的一个实施方式。在那种情况下，交换机可以具有表集合，表集合可以被重新编程和改变用途以支持复用功能，复用功能不是这些表的本机功能。

[0091] 更具体地，在一个实施方式中，表数据结构702包括四个链接表的集合，该集合包括表T_1、表T_2、表T_3和表T_4。图7示出了以高级方式表示的表中的几个代表性条目。在实践中，条目可以采取任何形式。假设复用器402从例如对应于镜像分组412的任何源接收分组。分组具有报头，报头指定与分组所指向的目的地相关联的特定地址。PM选择模块414首先使用输入地址作为索引，来定位第一表T_1中的条目($entry_w$)。该条目转而指向第二表T_2中的另一个条目($entry_x$)。该条目转而指向第三表T_3中的条目的连续块704。PM选择模块414基于任何选择逻辑来选择块704中的条目中的一个。例如，如上所述，PM选择模块414可以对从原始分组的IP报头提取的一个或多个信息项进行哈希，以产生哈希结果；该哈希结果转而落入与块704中的条目相关联的二进数(bin)中的一个，从而选择与该二进数相关联的条目。第三表T_3中的所选条目(例如$entry_{y2}$)指向第四表T_4中的条目($entry_z$)。

[0092] 在这一阶段，PM选择模块414可以使用由第四表中的$entry_z$赋予的信息，来生成与特定PM模块相关联的地址。发送模块420然后将分组封装到例如对应于封装的镜像分组422的新分组中。然后，发送模块420将封装的镜像分组422发送给所选择的PM。

[0093] 在一个实施方式中，表T_1可以对应于L3表，表T_2可以对应于组表，表T_3可以对应于ECMP表，并且表T_4可以对应于隧道表。这些是商品硬件交换机可以本机提供的表，但是它们不以图7中指定的方式链接在一起。它们也不填充上面指定的类型的映射信息。更具体地，在一些实施方式中，这些表包括具有在网络内执行本机分组转发功能所使用的条目的槽(slot)以及空闲(未使用)槽。跟踪系统102可以以上面阐述的特定方式将表链接，并且然后可以将条目加载到未使用的槽中，以共同提供用于复用目的的映射信息的实例。

[0094] 图9示出了在复用器402的输出处生成的封装的镜像分组422的说明性结构。封装的镜像分组422包括(作为其第一部分的)在交换机302的输出处产生的封装的镜像分组336。更具体地，封装的镜像分组422包括镜像分组326、镜像隧道报头802和镜像IP报头804。另外，封装的镜像分组422包括新的封装负载平衡器隧道报头902(诸如GRE隧道报头)。下一个封装外部字段包括负载平衡器IP报头904。其它实施方式可以采用封装在复用器402的输出处的镜像分组信息的其它方式。

[0095] 作为最后的评论，特别是在其中复用器110对应于改变用途的硬件交换机或其它硬件设备的情况下，复用器110具有高吞吐量。该特性是允许跟踪系统104处理高流量的一个特征；该特性还促进跟踪系统104的可伸缩性。

[0096] A.5.说明性处理模块

[0097] 图10示出了处理模块1002的一个实施方式，处理模块1002是图1的跟踪系统102的另一组件。处理模块1002从复用器110接收镜像分组流。如上所述，复用器110将与通过网络104的相同路径有关的镜像分组转发给同一处理模块。因此，在一个实施方式中，由处理模块1002接收的镜像分组流将不包含与其它处理模块处理的流有关的镜像分组。

[0098] 解封装模块1004从所接收的镜像分组中移除外部报头。例如，关于图9的封装的镜像分组422，解封装模块1004移除报头(802、804、902、904)，以留下由(图3的)镜像模块324

产生的原始镜像分组326。然而，为了简化以下说明，此后将由处理模块1002处理的镜像信息简称为镜像分组。在其它实施方式中，处理模块1002可以保留在外部报头中提供的至少一些信息，只要该信息提供有用的诊断信息。

[0099] 处理模块1002可以包括在镜像分组流上操作的一个或多个处理引擎1006的集合。例如，至少一个踪迹组装模块1008可以将属于通过网络104的同一流或路径的镜像分组集合分组在一起。在图1的示例中，例如，踪迹组装模块1008可以将由交换机106、126和128产生的镜像分组组装成单个组，以产生镜像分组序列138。踪迹组装模块1008还可以根据其创建的顺序在组中将镜像分组排序。踪迹组装模块1008可以通过查询由镜像分组捕获的时间戳、序列号和/或其它信息来执行其功能。

[0100] 至少一个过滤和选择(FS)模块1010可以从接收的镜像分组流中挑选出一个或多个类型的分组。例如，FS模块1010可以挑选出与特定TCP标记、或特定错误条件、或特定应用等有关的分组。例如通过使用正则表达式功能等，FS模块1010可以通过将所接收的镜像分组中提供的信息与匹配规则进行匹配来执行其功能。

[0101] 归档模块1012存储接收的原始镜像分组和/或由其它处理引擎1006生成的任何更高级的信息。归档模块1012可以在数据存储器1014中存储任何这样的信息，数据存储器1014可以对应于一个或多个物理存储机构、提供在单个站点处或分布在多个站点上。例如，在一种情况下，归档模块1004可以存储由处理模块1002接收的所有原始镜像分组。附加地或备选地，归档模块1012可以存储由踪迹组装模块1008产生的踪迹。附加地或备选地，归档模块1012可以存储由FS模块1010标识的镜像分组的所选择的子集等。

[0102] 更具体地，取决于将消费镜像分组的消费实体的计划需要，归档模块1012可以以针对不同类型的镜像分组的不同方式来存储镜像分组。在一些情况下，归档模块1012可以记录镜像分组的完整踪迹。在其它情况下，归档模块1012可以存储在路径中产生的特定镜像分组，而不必存储这些路径的完整踪迹。例如，如果捕获了指示在特定交换机处发生分组丢弃的明确信息，则归档模块1012可以避免捕获直到分组丢弃点的整个跳序列。

[0103] 接口模块1016允许任何消费实体(诸如图1的消费实体114)来检索由处理模块1002收集和处理的任何信息。在一种情况下，消费实体114可以对应于正在使用任何性质的计算设备来接收和分析所收集的信息的人类分析者。备选地或附加地，消费实体114可以对应于自动化分析程序。

[0104] 在一种情况下，消费实体114可以接收已经归档在数据存储器1014中的信息。备选地或附加地，消费实体114可以在它们被处理模块1002接收时接收镜像分组(例如，作为这种信息的实时流)。在一种情况下，接口模块1016允许任何消费实体经由一个或多个应用编程接口(API)与其资源交互。例如，接口模块1016可以提供用于不同的信息提取模式的不同API。API还可以允许消费实体指定用于在提取的期望镜像分组中使用的过滤标准等。

[0105] 接口模块1016还可以从消费实体接收指令。例如，自动化分析程序(例如，由消费实体实现)可以指示归档模块1012基于分析程序的信息需求，自动地和动态地改变其记录的信息的类型和性质。

[0106] 另一个接口模块1018提供用于执行处理模块1002和(图1的)管理模块116之间的通信的机构。例如，基于其分析，处理模块1002可以自动地向管理模块116发送指令，指示管理模块116转而向网络104中的交换机发送更新的分组检测规则。新的分组检测规则将改变

去往处理模块1002的镜像分组的流。例如，处理模块1002可以要求管理模块116提供新的规则集合，以增加或减少(例如，通过使选择标准更少或更多限制)其接收的镜像分组的量。备选地或附加地，处理模块1002可以动态地对正在接收的信息的类型做出反应。即，对于任何应用特定的原因，它可以影响分组检测规则中的改变，以捕获特定类型的附加类型的分组或者特定类型的较少分组。例如，处理模块1002可以收集特定量的证据，以暗示洪泛攻击当前正在发生；此后，其可以请求管理模块116降低其接收的进一步确认洪泛攻击的存在的镜像分组的量。

[0107] 出于任何应用特定的原因，管理模块116同样可以使用接口模块1018向处理模块1002发送指令。例如，管理模块116可以主动地向处理模块1002请求性能数据。管理模块116可以使用性能数据来以上述任何方式改变镜像功能的行为。可以执行管理模块116和处理模块1002之间的再一些其它环境特定的交互。

[0108] A.6.说明性消费实体

[0109] 图11示出了在图1的上下文中引入的消费实体114的一个实施方式。如上所述，消费实体114可以对应于计算设备，人类分析者通过该计算设备对镜像分组执行分析。备选地或附加地，消费实体可以对应于运行一个任何类型的计算设备的一个或多个分析程序。

[0110] 消费实体114包括接口模块1102，用于例如通过由处理模块112提供的一个或多个API与处理模块112交互。消费实体114可以获得由处理模块112捕获和处理的任何信息。在一种情况下，消费实体114可以向处理模块112的整个集合做出信息请求；然后，保持期望信息的一个或多个特定处理模块将通过提供期望信息来做出响应。备选地或附加地，处理模块112可以自动向消费实体114提供镜像分组信息。例如，消费实体114可以注册一个或多个事件处理器，以用于接收期望的分组相关信息的目的。处理模块112可以通过在遇到它时提供期望信息来对这些事件处理器做出响应。消费实体114可以将其收集的信息存储在数据存储器1104中。如上所述，消费实体114还可以向处理模块112发送指令和其它反馈。

[0111] 消费实体114可以提供用于分析所接收的镜像分组信息的一个或多个专用处理引擎1106。在一种情况下，例如，处理引擎可以检查所收集的镜像分组的报头中的TCP报头信息。该信息揭示了在通信实体之间建立的连接的数量。处理引擎可以将连接的数量与阈值进行比较，以确定是否发生了洪泛攻击或其它异常条件。

[0112] 另一个处理引擎可以针对断开的链路或行为不良的组件来检查网络104，断开的链路或行为不良的组件可以贡献于丢失或损坏的信息流。这样的处理引擎可以基于各种证据(诸如，通过标识分组的过早截断的序列(例如，其中分组未到达其旨在目的地))和/或基于包含缺失的跳、异常路由等的分组的序列来确定故障的存在。附加地或备选地，处理引擎可以检查以下证据中的任何证据：BGP或其它路由信息、由交换机添加的错误条件元数据、ping相关的分组信息等。即，BGP信息可以直接揭示网络中的路由问题(诸如，链路的故障或行为不良等)。错误条件信息可以揭示特定交换机已经由于其损坏或其它因素而丢弃分组。ping相关分组信息可以揭示网络中的两个实体之间的连接性问题。如上所述，ping应用对应于如下应用：该应用通过向远程实体发送测试消息并且监听远程实体对ping消息的响应，来测试到远程实体的连接的质量。

[0113] 消费实体114可以使用再一些其它类型的处理引擎1106；上述示例是在说明的而不是限制的精神上描述的。

[0114] 处理引擎可以以任何方式实现，诸如通过基于规则的引擎、人工智能引擎、机器训练的模型等。例如，一个基于规则的处理引擎可以采用反映诊断规则集合的映射表或分支算法。每个规则可以以IF-THEN格式来构造。即，规则可以指定：如果证据集合{X_1、X_2、...、X_n}存在于所捕获的镜像分组中，则网络可能遭受异常Y。这些规则的特定性质本质上将是环境特定的，取决于正被监控的网络104的性质、分析的目标和/或任何其它因素。

[0115] 在一些情况下，处理引擎还可以动态地执行一系列测试，其中后续测试可以由一个或多个先前测试的结果触发，并且可以依赖于在先前测试中生成的结论。

[0116] 至少一个动作采取模块1108可以基于由任何处理引擎1106提供的分析结果来采取动作。例如，一个动作采取模块可以以任何形式(例如，通过提供警报信号、检测到的故障的原因的文本解释等)向人类分析者通知分析结果。在另一种情况下，动作采取模块可以主动地禁用或以其它方式修改已被确定为行为不良的网络104的一部分的性能。例如，这种动作采取模块可以禁用到正被攻击的特定服务器或其它资源的通信路由、阻断源自可疑恶意实体的流量等。

[0117] 接口模块1110允许消费实体114与管理模块116交互。例如，至少为了处理模块112可以这样做的相同原因，消费实体114可以向管理模块116发送请求。例如，处理引擎可以希望改变正在接收的分组的类型，或者改变其正在接收的分组的量。为此，处理引擎可以向管理模块116做出请求，指示其向网络104中的交换机发送更新的分组检测规则。更新的规则在实际中由交换机放置时，将实现处理引擎的目标。

[0118] 作为关于图1和图11的最后注释，这些附图示出了处理模块112作为与消费实体分离的代理。在其它实施方式中，上面描述为由处理模块112执行的一个或多个功能可以替代地由消费实体执行。实际上，在一些实施方式中，可以完全消除处理模块112，并且消费实体可以直接从复用器110接收镜像分组。

[0119] A.7.说明性管理模块

[0120] 最后，图12示出了管理模块116的一个实施方式。管理模块116可以使用至少一个控制模块1202来控制网络交换机、复用器110、处理模块112等中的各种操作。例如，控制模块1202可以向交换机提供分组检测规则集合，其管理交换机的后续镜像行为。控制模块1202可以基于一个或多个因素(诸如，来自管理员的明确指令、与消费实体相关联的人类分析者的明确请求、任何处理模块或消费实体的自动化请求等)生成新规则。

[0121] 在一种情况下，管理模块116指示所有交换机加载相同的分组检测规则集合。在其它情况下，管理模块116可以指示交换机的不同子集加载分组检测规则的不同的相应集合。管理模块116可以针对任何环境特定原因采取后面的方法，例如以降低由具有高流量的交换机产生的镜像分组的量等。

[0122] 管理模块116还可以包括至少一个性能监控模块1204。该组件接收关于网络104和跟踪系统102的各种组件的行为的反馈信息。基于该信息，性能监控模块1204可以生成反映网络104和跟踪系统102的性能水平的一个或多个性能相关测量。例如，性能监控模块1204可以确定由跟踪系统102创建的镜像分组的量。可以以各种方式将镜像分组与原始分组区分开。例如，交换机上提供的每个镜像机构可以向其创建的镜像分组添加服务类型(TOS)标记，该标记可以将分组标识为镜像分组。

[0123] 控制模块1202还可以基于由性能监控模块1204提供的性能数据来更新其传播给

交换机的规则。例如，控制模块1202可以降低镜像分组的数量，以减少在峰值流量负载的时段期间网络104中的拥塞，使得跟踪系统102的镜像行为将不会不利地影响原始分组的流。

[0124] 管理模块116还可以包括执行其它管理操作的任何其它功能1206。例如，尽管在图12中未明确陈述，但是功能1206可以向交换机编译和发送路由信息。该路由信息确定交换机通过网络104路由原始分组和镜像分组的方式。

[0125] 最后，管理模块116可以包括用于与跟踪系统102的各种参与者交互的多个接口，包括用于与网络104中的交换机交互的接口模块1208、用于与复用器110交互的接口模块1210、用于与处理模块112交互的接口模块1212、以及用于与消费实体交互的接口模块1214。

[0126] B.说明性处理

[0127] 图13-图18以流程图形式示出了解释部分A的跟踪系统102的操作的处理。由于跟踪系统102的操作底层的原理已经在部分A中描述，因此在本部分中将以概括的方式解决特定操作。

[0128] 从图13开始，该图示出了解释图3的交换机302的一种操作方式的处理1302。在块1302中，交换机302接收通过网络104传输的原始分组。在块1306中，交换机302确定是否将原始分组镜像。在块1308中，假设做出镜像原始分组的决定，交换机基于原始分组生成镜像分组。镜像分组至少包括在原始分组中提供的信息的子集。在块1310中，交换机302可选地基于至少一个负载平衡考虑从候选复用器集合110中选择复用器。在以下意义上，该操作是可选的：在一些实施方式中，跟踪系统102可以仅提供单个复用器，因此在这种情况下在复用器之间不需要复用。在块1312中，交换机302将镜像分组发送给所选择(或默认)的负载平衡复用器。在块1314中，交换机302将原始分组发送给由原始分组指定的目标目的地。上述操作被串行地描述以简化说明；但是这些操作中的任何操作也可以并行执行(诸如操作1312和1314)。

[0129] 图14示出了解释匹配模块320的一种操作方式的处理1402，匹配模块320是图3的交换机302的组件。在块1404中，匹配模块320关于至少一个分组检测规则来分析原始分组。在块1406中，匹配模块320确定原始分组是否满足分组检测规则。在块1408中，如果原始分组满足分组检测规则，则匹配模块320生成镜像原始分组的指令。在实际实践中，匹配模块320可以关于分组检测规则的集合串行地或并行地执行图14的操作。

[0130] 图15示出了解释图4的复用器402的一种操作方式的处理1502。在块1504中，复用器402接收镜像分组。在块1506中，复用器402基于至少一个负载平衡选择考虑从处理模块候选集合中选择处理模块。例如，复用器402可以使用上述哈希技术来在处理模块候选之间进行选择，同时还确保属于相同流的分组被发送给同一处理模块。在块1508中，复用器402将镜像分组发送给已经选择的处理模块。

[0131] 图16示出了解释图10的处理模块1002的一种操作方式的处理1602。在块1604中，处理模块1002从复用器110接收镜像分组。在块1606中，处理模块1002对镜像分组执行任何类型的处理，诸如但不限于：组装相关镜像分组(例如，其属于相同的流)的序列；过滤和选择特定镜像分组；将镜像分组和/或由处理模块1002执行的分析结果归档等。

[0132] 图17示出了解释图11的消费实体114的一种非限制性和代表性操作方式的处理1702。在块1704中，消费实体114确定是否开始其对镜像分组的分析。例如，假设消费实体

114与和网络104交互或在网络104中发挥某一作用的特定应用(诸如,TCP相关应用或BGP相关应用)相关联。在一种操作模式中,这样的应用可以独立于跟踪系统102确定在网络104中发生故障或其它不期望的事件。作为响应,应用可以请求交换机开始收集特定类型的镜像分组。即,应用可以向管理模块116做出这样的请求,管理模块116转而向交换机发送一个或多个分组检测规则,分组检测规则当由交换机应用时将具有捕获期望分组的最终效果。在另一种操作模式中,在不首先遇到异常条件的情况下,应用可以请求交换机在正常操作过程中收集特定分组。再一些其它操作模式是可能的。

[0133] 在块1706中,消费实体114接收由处理模块112提供的镜像分组和/或分析结果。消费实体114可以使用推送(push)技术、拉收(pull)技术或其组合,来获得块1706中的信息。在块1708中,消费实体114分析镜像分组以得到关于网络104中已发生的或者当前正在网络104中发生的事件的第一结论。此后,基于该第一结论,消费实体114可以采取一个或多个动作(其示例在图17中概括)。

[0134] 例如,在块1710中,消费实体114可以向人类分析者、管理员或任何其他实体通知网络104内的异常条件。消费实体114可以使用任何用户接口呈现来传送这些结果。备选地或附加地,在块1712中,消费实体114可以记录其分析的结果。备选地或附加地,在块1714中,消费实体114可以采取任何其它动作,诸如,通过禁用或以其它方式改变网络104的任何部分的行为。

[0135] 备选地或附加地,在块1716中,消费实体114可以使用第一结论来触发另一轮分析。该第二轮分析可以使用第一结论作为输入数据。这样的迭代调查可以重复任何次数,直到人类分析者或自动程序得到期望的最后结论。注意,块1716的分析相对于消费实体114已从处理模块112接收的镜像分组信息进行。

[0136] 备选地或附加地,在块1718中,消费实体114可以与处理模块112交互,以从处理模块112获得附加的分组相关信息。备选地或附加地,消费实体114可以与管理模块116交互,来请求管理模块116改变加载在交换机上的分组检测规则。这种改变转而将改变消费实体114从处理模块112接收的分组的类型和/或量。然后,当附加的分组相关信息已被接收时,消费实体114可以重复上述任何操作。

[0137] 最后,图18示出了解释图12的管理模块116的一种操作方式的处理1802。在块1804中,管理模块116可以向跟踪系统102的组件(诸如,网络104中的交换机、复用器110、处理模块112等)发送各种指令。例如,管理模块116可以向交换机发送更新的分组检测规则集合,交换机此后将以特定方式管理其分组镜像行为。在块1806中,管理模块116从诸如交换机、复用器110、处理模块112、消费实体等的各种实体接收反馈。以上述方式,即在块1804的后续执行中,管理模块116可以随后使用反馈来更新其发送给各种代理的指令。管理模块116还可以执行在图18中未表示的其它管理功能。

[0138] C.代表性计算功能

[0139] 图19示出了可以用于实现在上述附图中阐述的跟踪功能的任何方面的计算功能1902。例如,图19中所示的计算功能1902的类型可以用于实现以下中的任何项:软件实现的复用器(如果在图1的跟踪系统102中使用)、任何分组处理模块、管理模块116、任何消费实体(诸如消费实体114)等。在所有情况下,计算功能1902表示一个或多个物理的和有形的处理机构。

[0140] 计算功能1902可以包括一个或多个处理设备1904，诸如一个或多个中央处理单元(CPU)、和/或一个或多个图形处理单元(GPU)等。

[0141] 计算功能1902还可以包括用于存储诸如代码、设置、数据等的任何类型的信息的任何存储资源1906。在非限制情况下，例如，存储资源1906可以包括以下中的任何项：任何类型的RAM、任何类型的ROM、闪存设备、硬盘、光盘等。更一般地，任何存储资源可以使用任何用于存储信息的技术。此外，任何存储资源可以提供易失性或非易失性信息的保留。此外，任何存储资源可以表示计算功能1902的固定或可移除组件。当处理设备1904执行存储在任何存储资源或任何存储资源组合中的指令时，计算功能1902可以执行上述任何功能。

[0142] 关于术语，任何存储资源1906或存储资源1906的任何组合可以被视为计算机可读介质。在许多情况下，计算机可读介质表示某种形式的物理和有形实体。术语计算机可读介质还包括例如经由物理管道和/或空气或其它无线介质等发射或接收的传播信号。然而，特定术语“计算机可读存储介质”和“计算机可读介质设备”明确地排除传播信号本身，同时包括所有其它形式的计算机可读介质。

[0143] 计算功能1902还包括用于与任何存储资源交互的一个或多个驱动机构1908(诸如，硬盘驱动机构、光盘驱动机构等)。

[0144] 计算功能1902还包括用于接收各种输入(经由输入设备1912)和用于提供各种输出(经由输出设备1914)的输入/输出模块1910。说明性输入设备包括键盘设备、鼠标输入设备、触摸屏输入设备、数字化板、一个或多个摄像机、一个或多个深度相机、自由空间姿势识别机构、一个或多个麦克风、语音识别机构、任何运动检测机构(例如加速度计、陀螺仪等)等。一个特定输出机构可以包括呈现设备1916和相关联的图形用户界面(GUI)1918。其它输出设备包括打印机、模型生成机构、触觉输出机构、归档机构(用于存储输出信息)等。计算功能1902还可以包括用于经由一个或多个通信管道1922与其它设备交换数据的一个或多个网络接口1920。一个或多个通信总线1924将上述组件通信地耦合在一起。

[0145] 通信管道1922可以以任何方式实现，例如通过局域网、广域网(例如因特网)、点对点连接等、或其任何组合。通信管道1922可以包括由任何协议或任何协议组合管理的硬连线链路、无线链路、路由器、网关功能、名称服务器等的任何组合。

[0146] 备选地或附加地，可以至少部分地通过一个或多个硬件逻辑组件来执行前面部分中描述的任何功能。例如但不限于，计算功能1902可以使用以下中的一个或多个项来实现：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。

[0147] 最后，尽管以特定于结构特征和/或方法动作的语言描述了主题，但是应当理解，所附权利要求中限定的主题不一定限于以上所描述的具体特征或动作。相反，以上描述的具体特征和动作被公开为实现权利要求的示例形式。

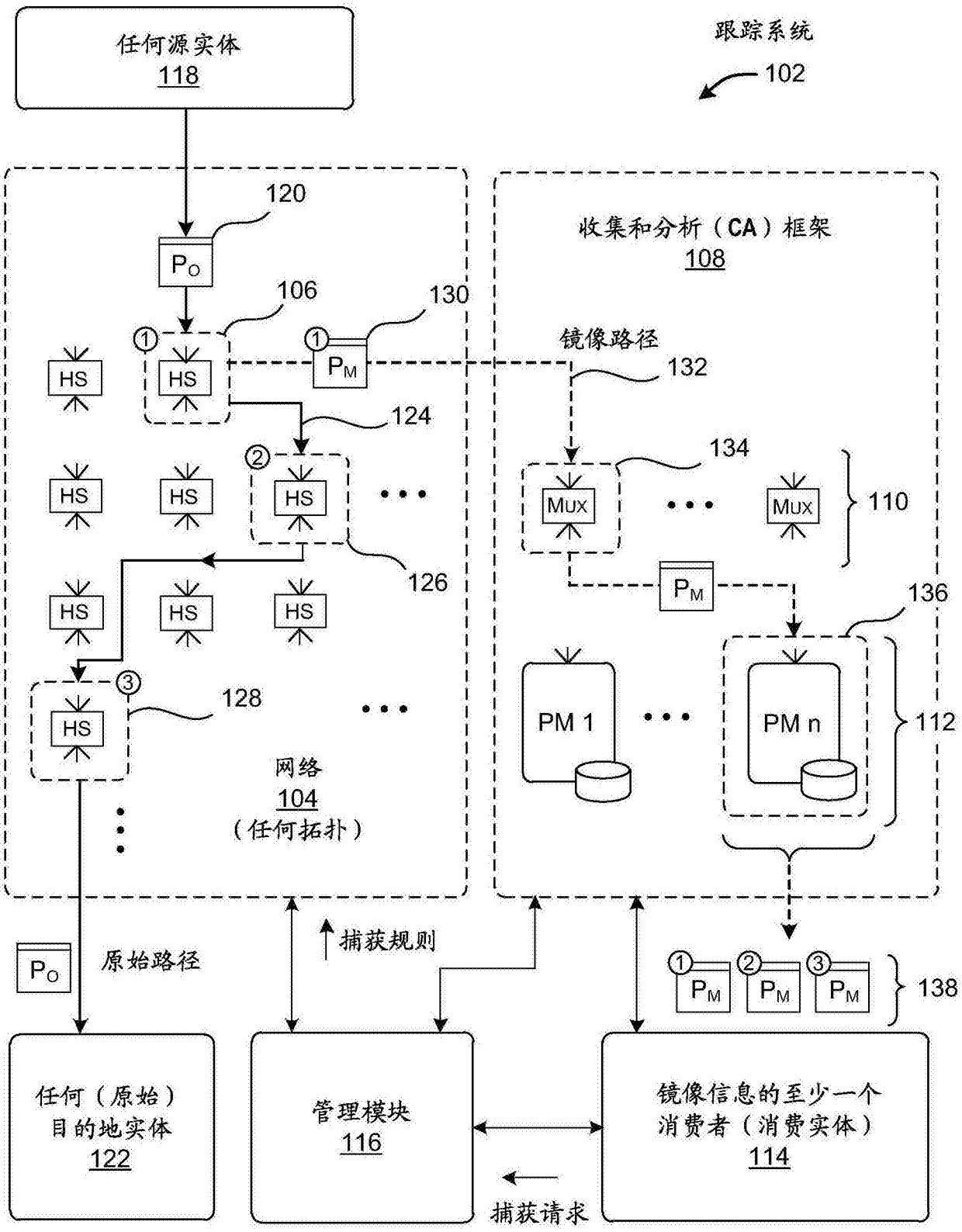

任何源实体
118
跟踪系统
102
120
P_O
106
130
收集和分析（CA）框架
108
镜像路径
132
HS
P_M
124
134
Mux
110
126
136
PM 1
PM n
112
128
网络
104
（任何拓扑）
捕获规则
原始路径
138
任何（原始）
目的地实体
122
管理模块
116
镜像信息的至少一个
消费者（消费实体）
114
捕获请求

图1

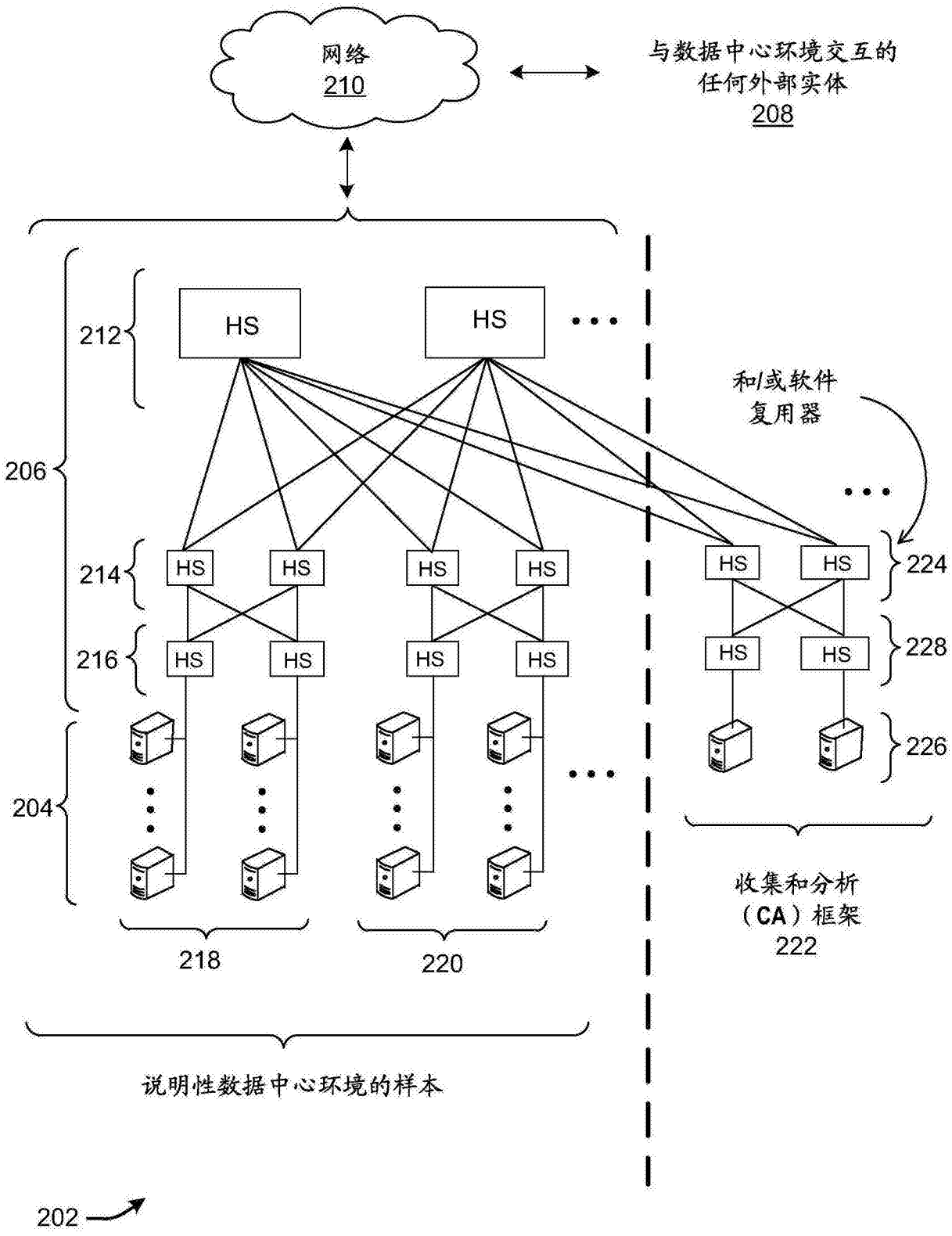

网络
210
与数据中心环境交互的
任何外部实体
208
212
HS
HS
和/或软件
复用器
206
214
HS
HS
HS
HS
HS
HS
224
216
HS
HS
HS
HS
HS
HS
228
204
226
收集和分析
（CA）框架
222
218
220
说明性数据中心环境的样本
202

图2

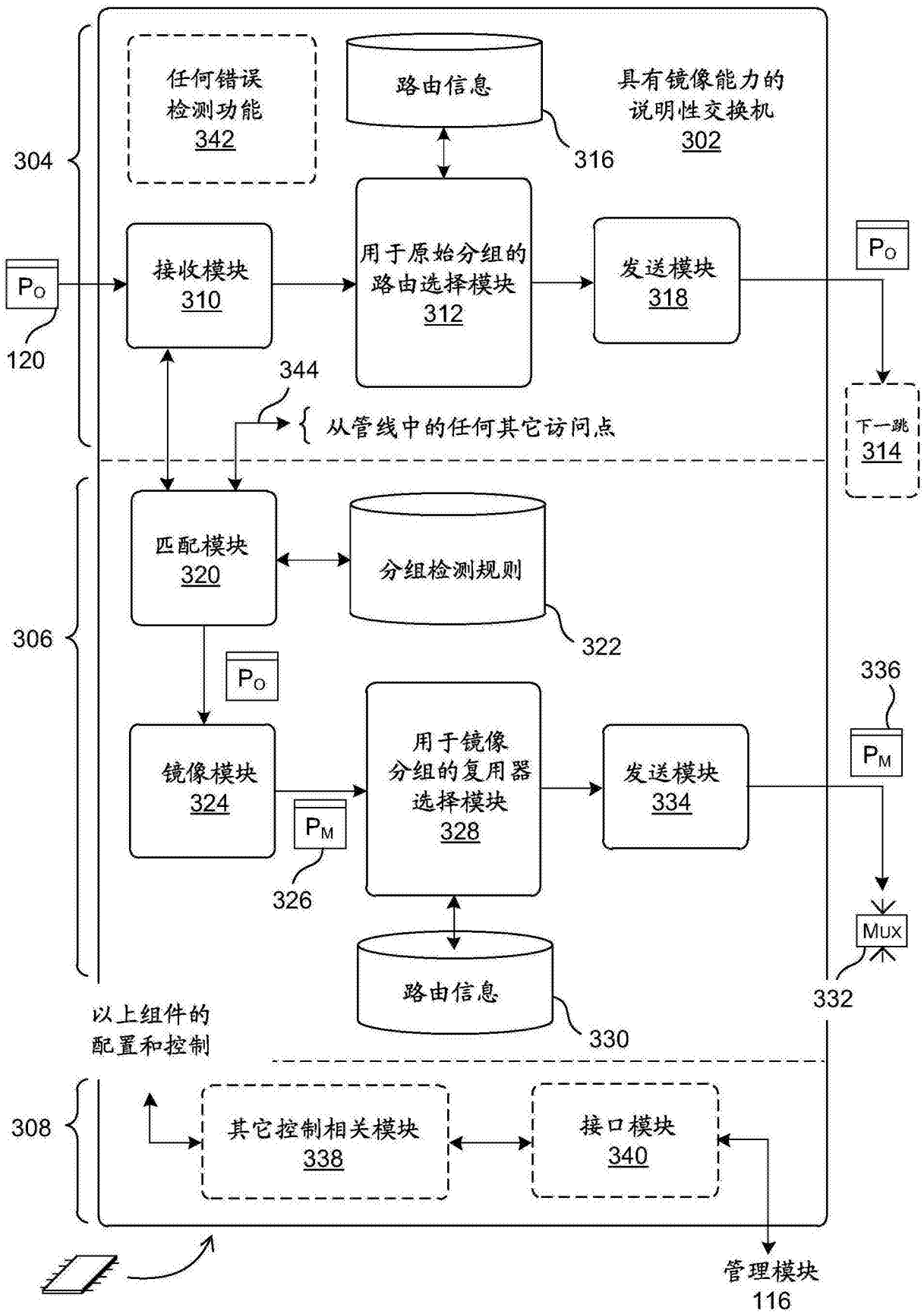

任何错误
检测功能
342
路由信息
316
具有镜像能力的
说明性交换机
302
304
接收模块
310
用于原始分组的
路由选择模块
312
发送模块
318
PO
120
344
从管线中的任何其它访问点
下一跳
314
匹配模块
320
分组检测规则
322
306
PO
336
镜像模块
324
用于镜像
分组的复用器
选择模块
328
发送模块
334
PM
PM
326
MUX
路由信息
330
332
以上组件的
配置和控制
308
其它控制相关模块
338
接口模块
340
管理模块
116

图3

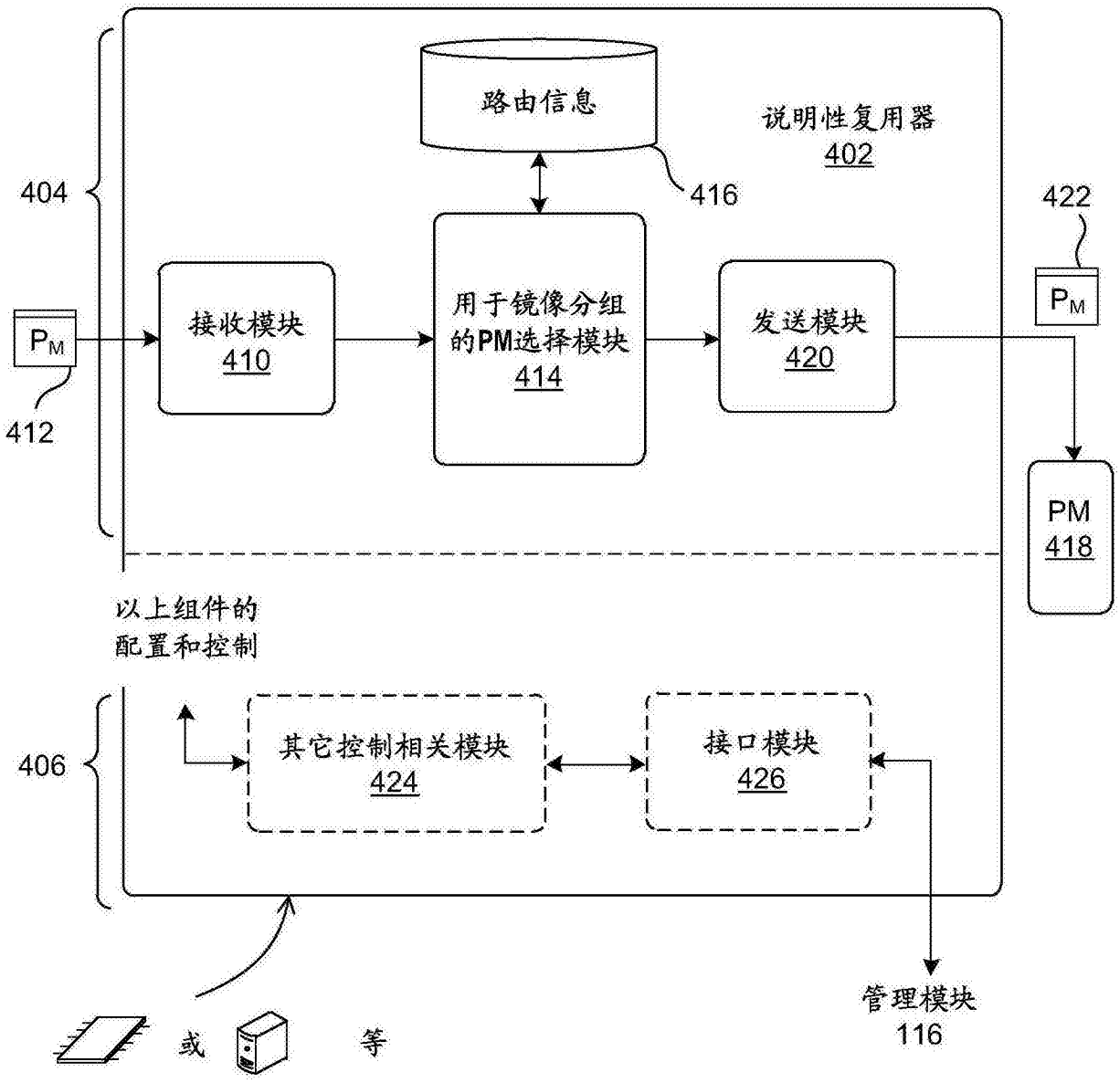

路由信息
416
说明性复用器
402
404
422
P_M
接收模块
410
用于镜像分组
的PM选择模块
414
发送模块
420
P_M
412
PM
418
以上组件的
配置和控制
406
其它控制相关模块
424
接口模块
426
或
等
管理模块
116

图4

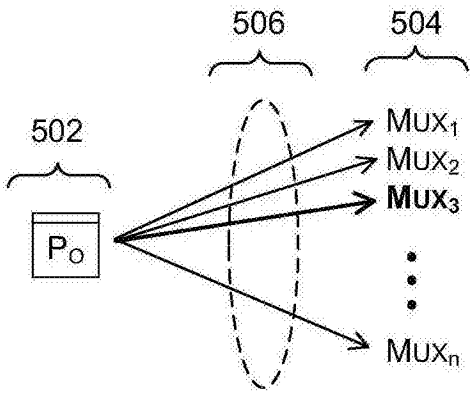

506
504
502
MUX_1
MUX_2
MUX_3
P_O
MUX_n

图5

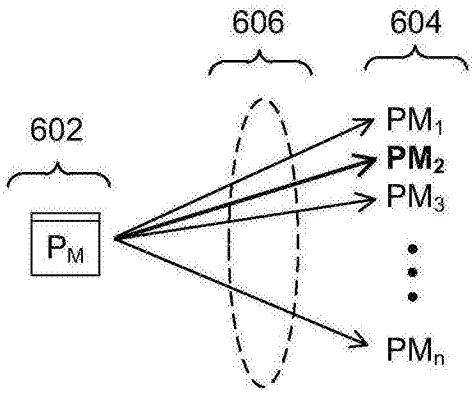

606
604
602
PM_1
PM_2
PM_3
P_M
PM_n

图6

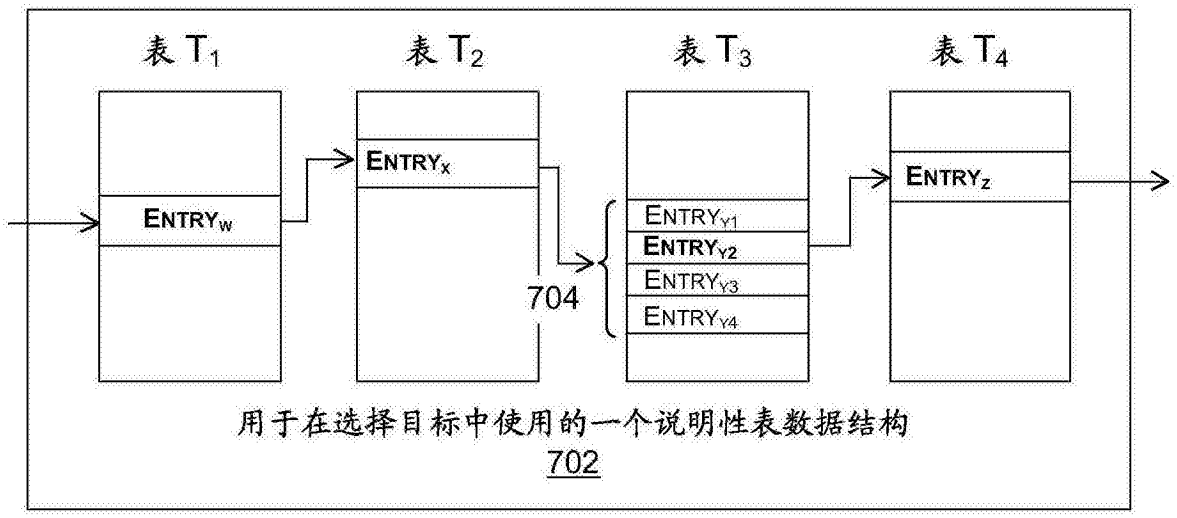


表 T1
表 T2
表 T3
表 T4
ENTRYW
ENTRYX
ENTRYY1
ENTRYY2
ENTRYY3
ENTRYY4
ENTRYZ
704
用于在选择目标中使用的一个说明性表数据结构
702

图7


镜像IP报头
804
镜像隧道
（例如GRE）
报头
802
创建的
镜像分组
326
336

图8


负载平衡器
IP报头
904
负载平衡器隧道
（例如GRE）报头
902
镜像IP报头
804
镜像隧道
（例如GRE）
报头
802
创建的
镜像分组
326
422

图9

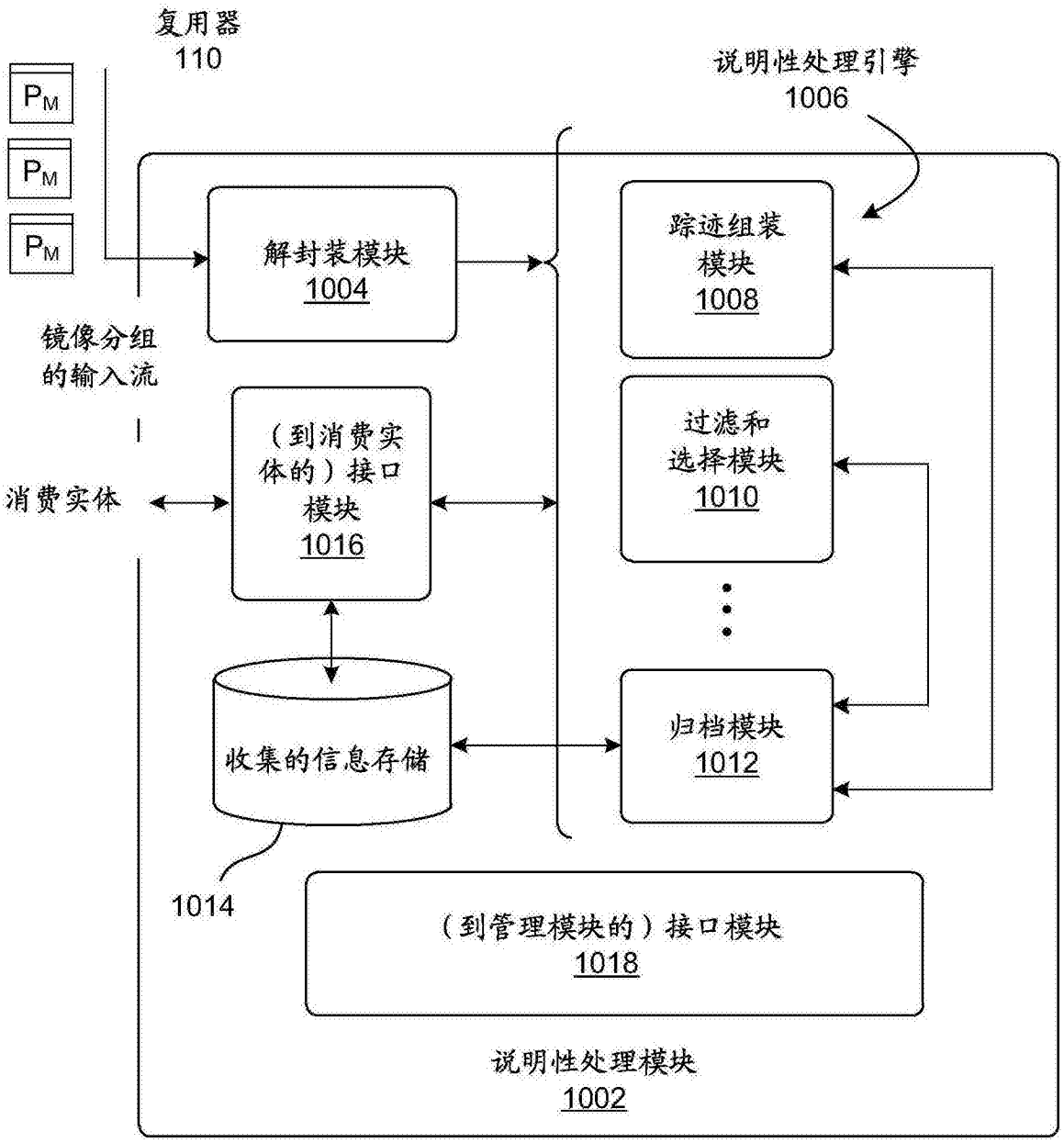

复用器
110
PM
PM
PM
镜像分组
的输入流
消费实体
解封装模块
1004
（到消费实
体的）接口
模块
1016
收集的信息存储
1014
说明性处理引擎
1006
踪迹组装
模块
1008
过滤和
选择模块
1010
归档模块
1012
（到管理模块的）接口模块
1018
说明性处理模块
1002

图10

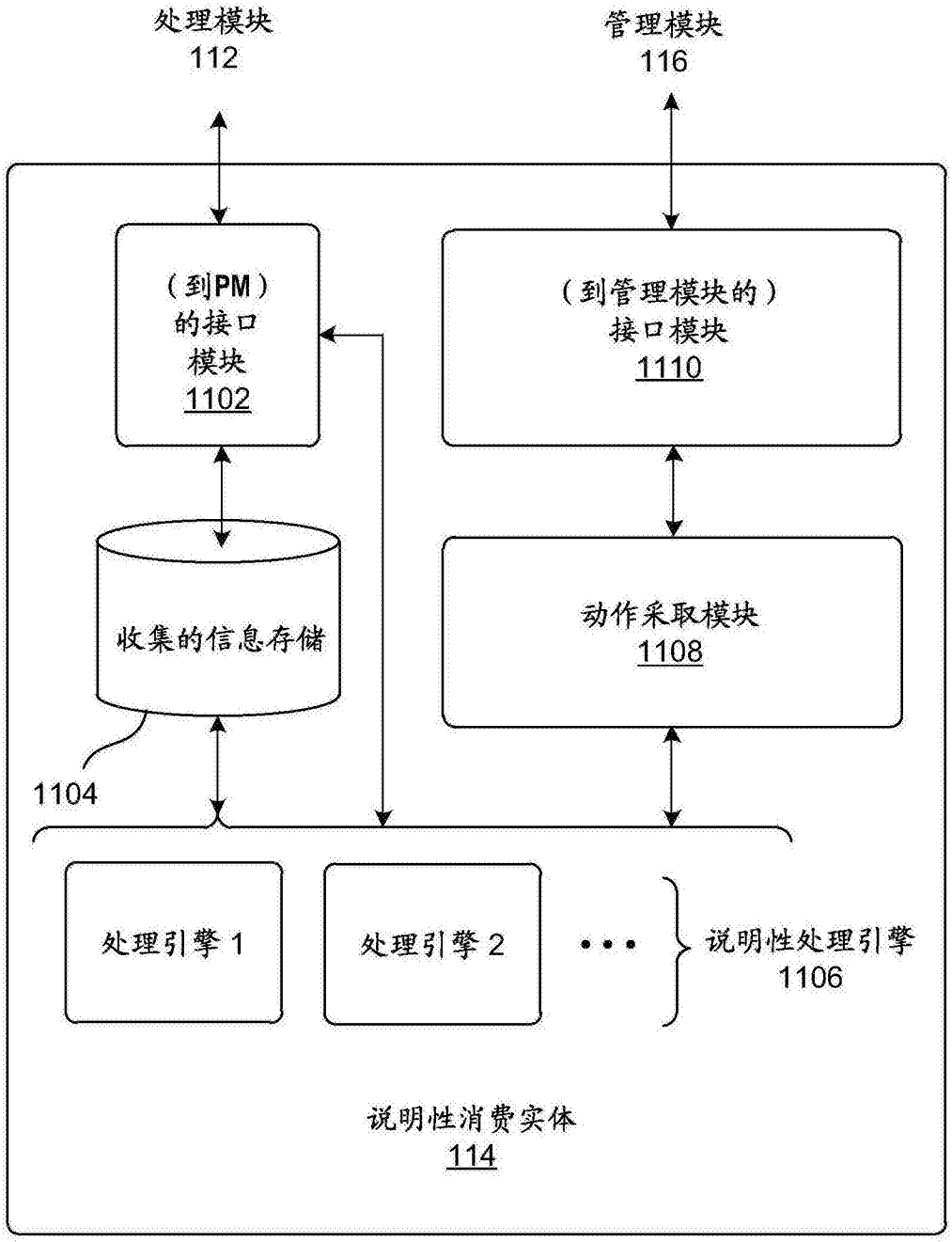

处理模块
112
管理模块
116
（到PM）
的接口
模块
1102
（到管理模块的）
接口模块
1110
收集的信息存储
动作采取模块
1108
1104
处理引擎 1
处理引擎 2
• • •
说明性处理引擎
1106
说明性消费实体
114

图11

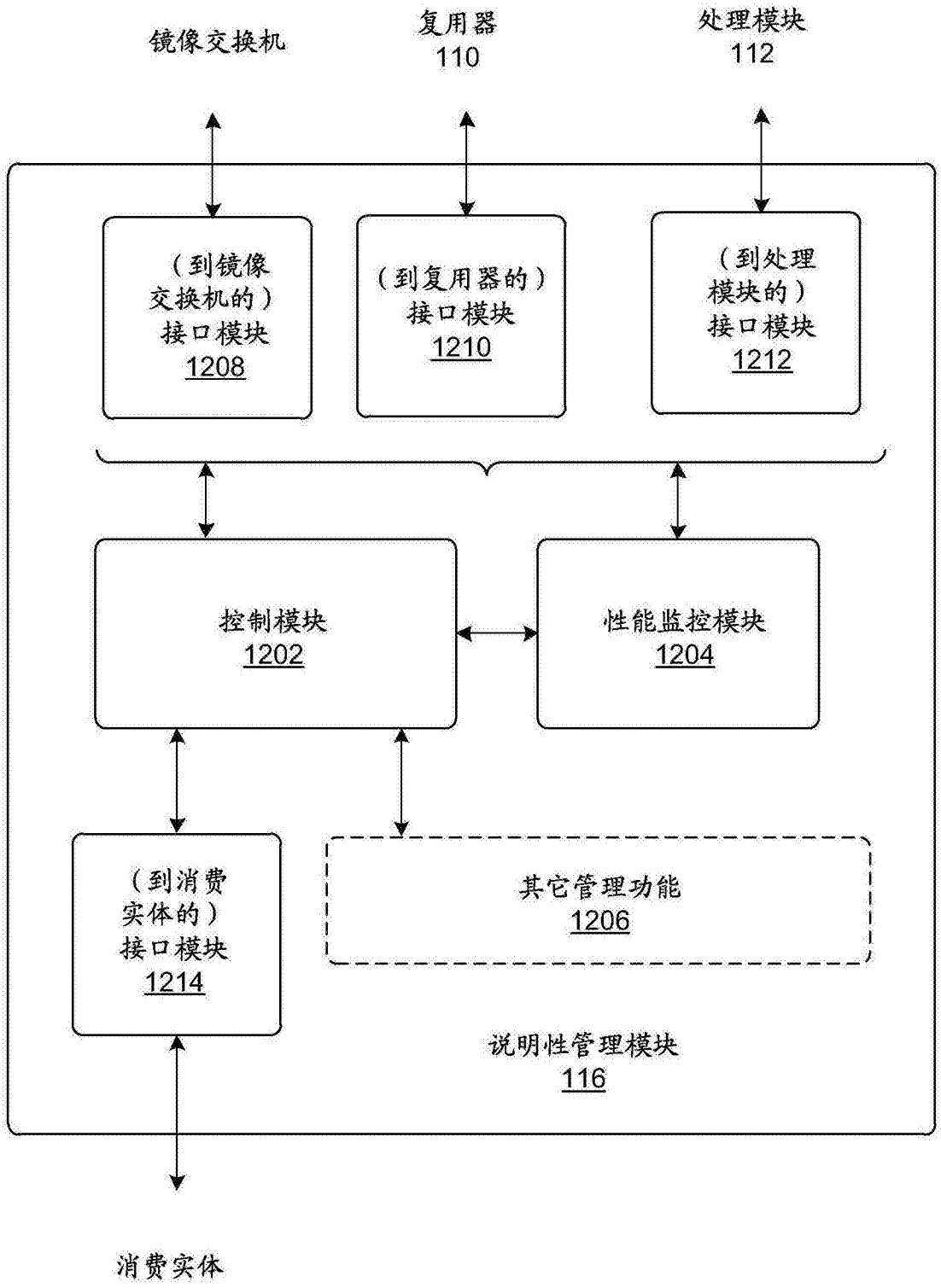

镜像交换机
复用器
110
处理模块
112
（到镜像交换机的）接口模块
1208
（到复用器的）接口模块
1210
（到处理模块的）接口模块
1212
控制模块
1202
性能监控模块
1204
（到消费实体的）接口模块
1214
其它管理功能
1206
说明性管理模块
116
消费实体

图12

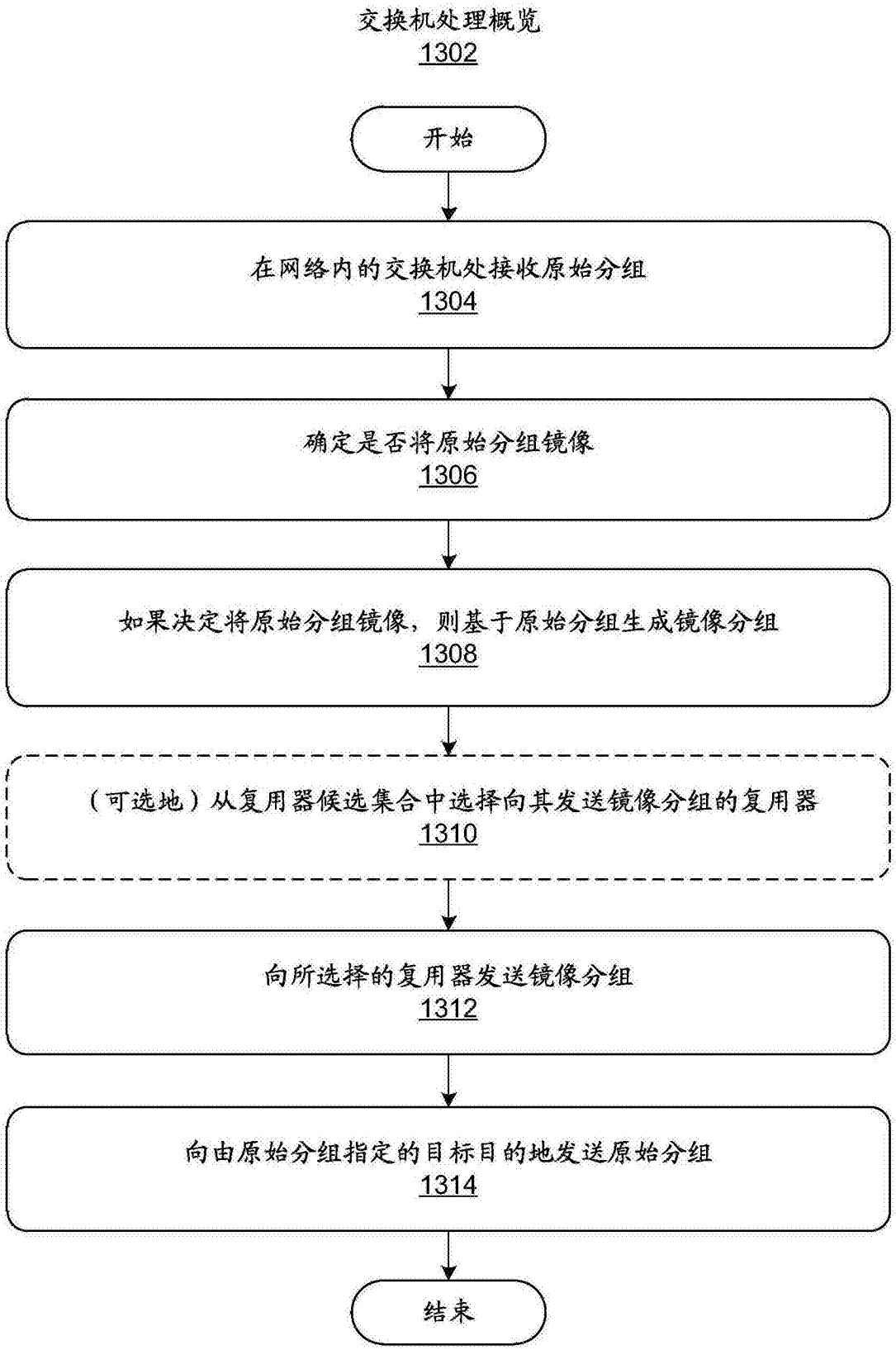

交换机处理概览
1302
开始
在网络内的交换机处接收原始分组
1304
确定是否将原始分组镜像
1306
如果决定将原始分组镜像，则基于原始分组生成镜像分组
1308
（可选地）从复用器候选集合中选择向其发送镜像分组的复用器
1310
向所选择的复用器发送镜像分组
1312
向由原始分组指定的目标目的地发送原始分组
1314
结束

图13

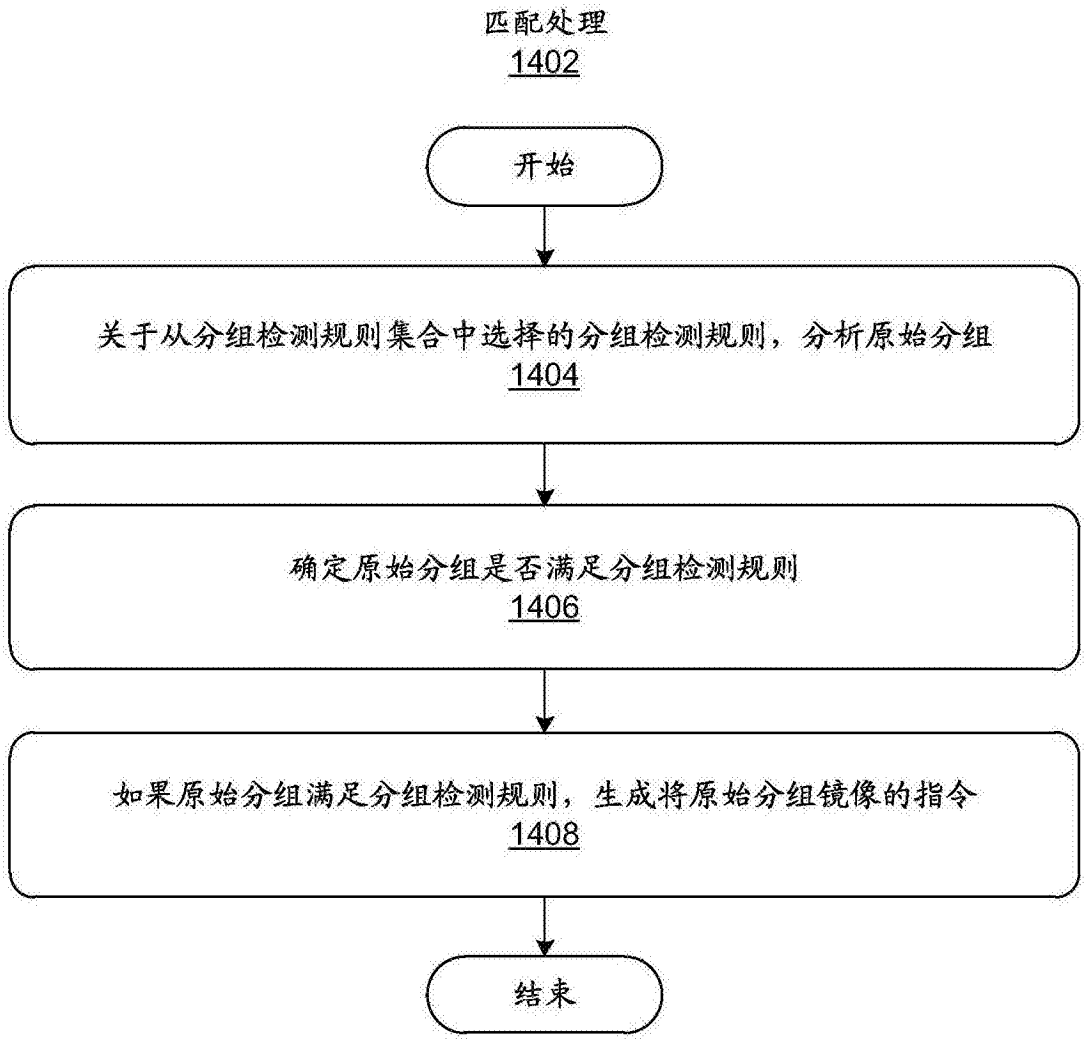

匹配处理
1402
开始
关于从分组检测规则集合中选择的分组检测规则，分析原始分组
1404
确定原始分组是否满足分组检测规则
1406
如果原始分组满足分组检测规则，生成将原始分组镜像的指令
1408
结束

图14

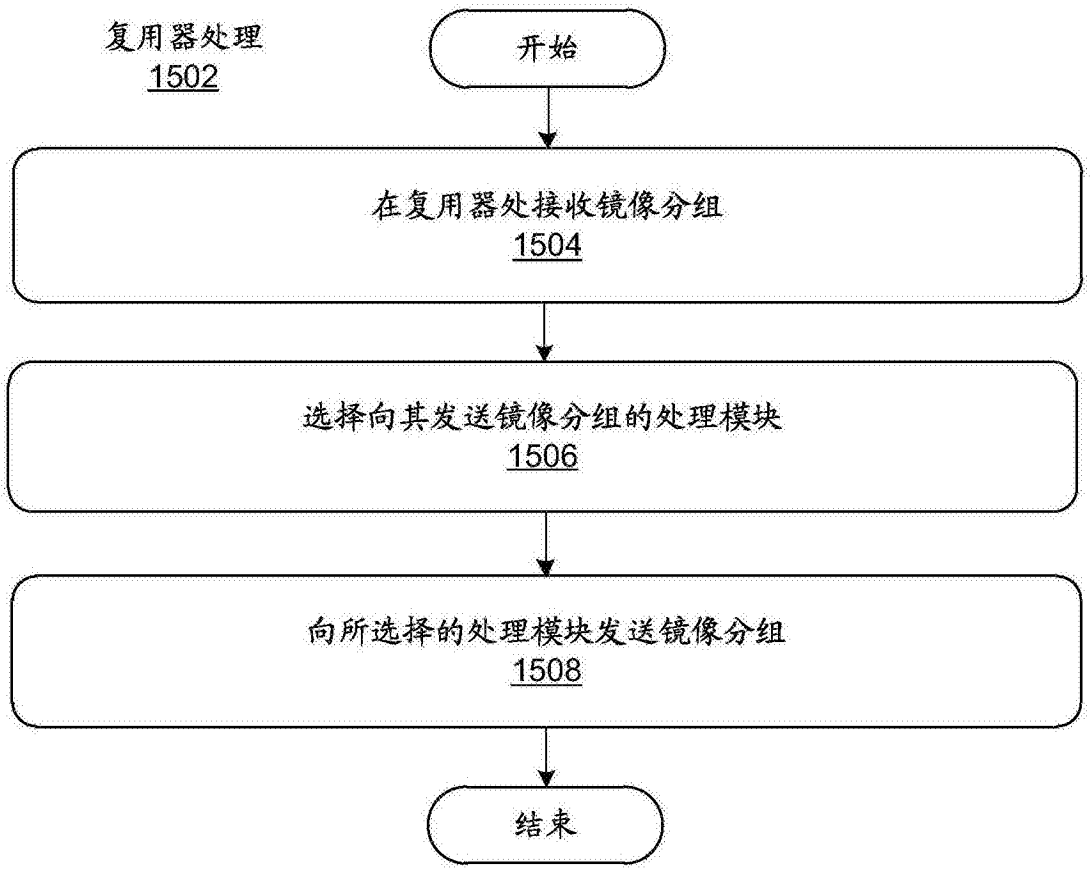

复用器处理
1502
开始
在复用器处接收镜像分组
1504
选择向其发送镜像分组的处理模块
1506
向所选择的处理模块发送镜像分组
1508
结束

图15

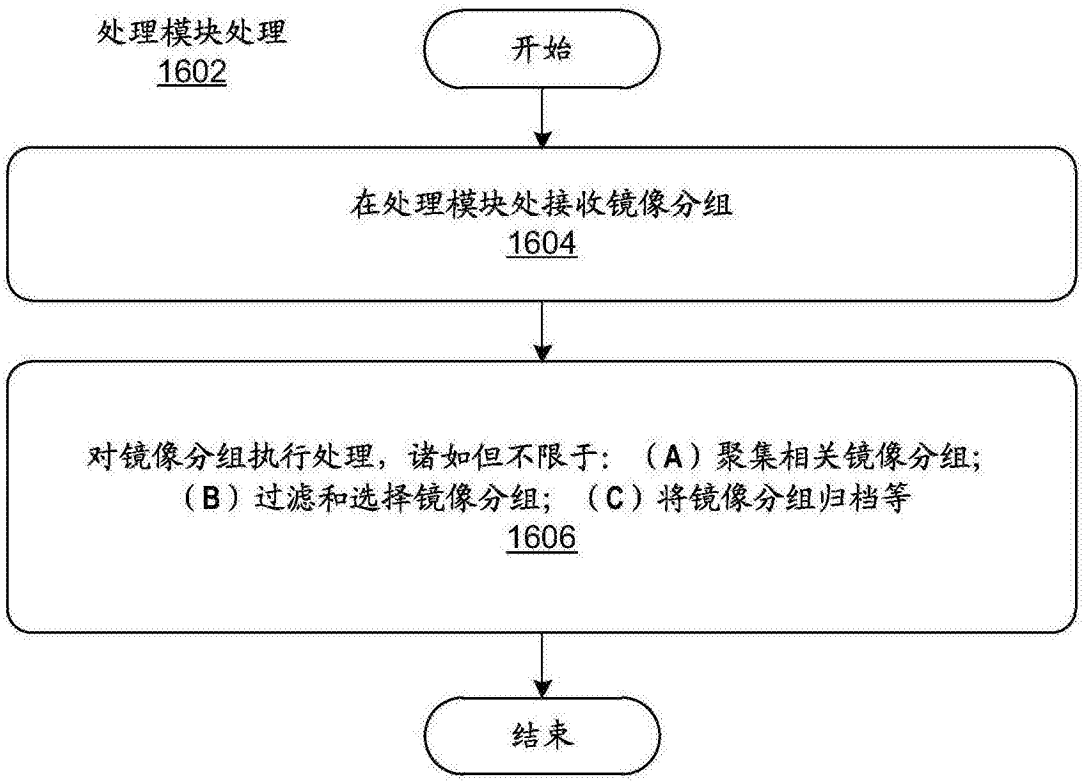

处理模块处理
1602
开始
在处理模块处接收镜像分组
1604
对镜像分组执行处理，诸如但不限于：（A）聚集相关镜像分组；
（B）过滤和选择镜像分组；（C）将镜像分组归档等
1606
结束

图16

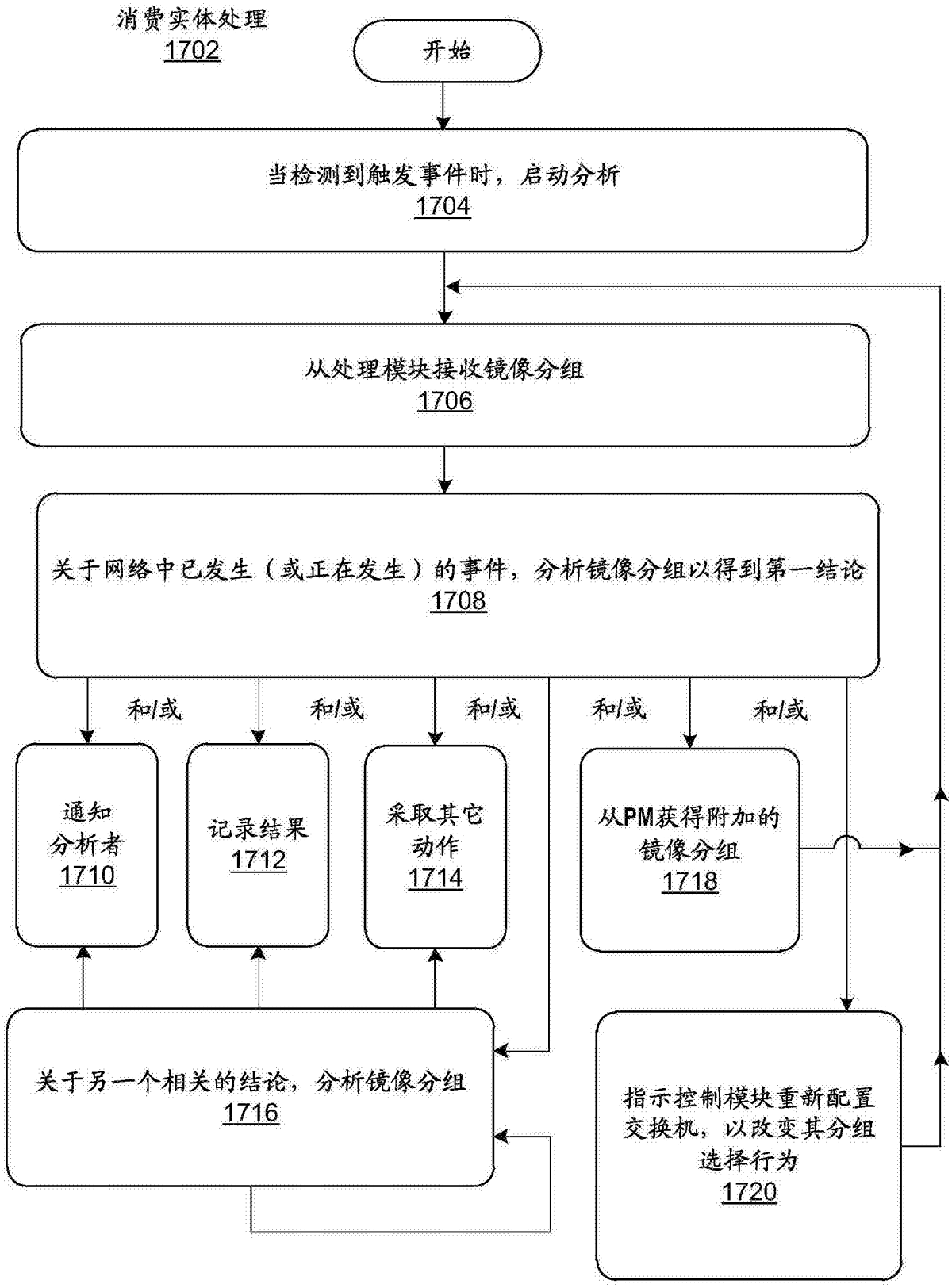

消费实体处理
1702
开始
当检测到触发事件时，启动分析
1704
从处理模块接收镜像分组
1706
关于网络中已发生（或正在发生）的事件，分析镜像分组以得到第一结论
1708
和/或
和/或
和/或
和/或
和/或
通知
分析者
1710
记录结果
1712
采取其它
动作
1714
从PM获得附加的
镜像分组
1718
关于另一个相关的结论，分析镜像分组
1716
指示控制模块重新配置
交换机，以改变其分组
选择行为
1720

图17

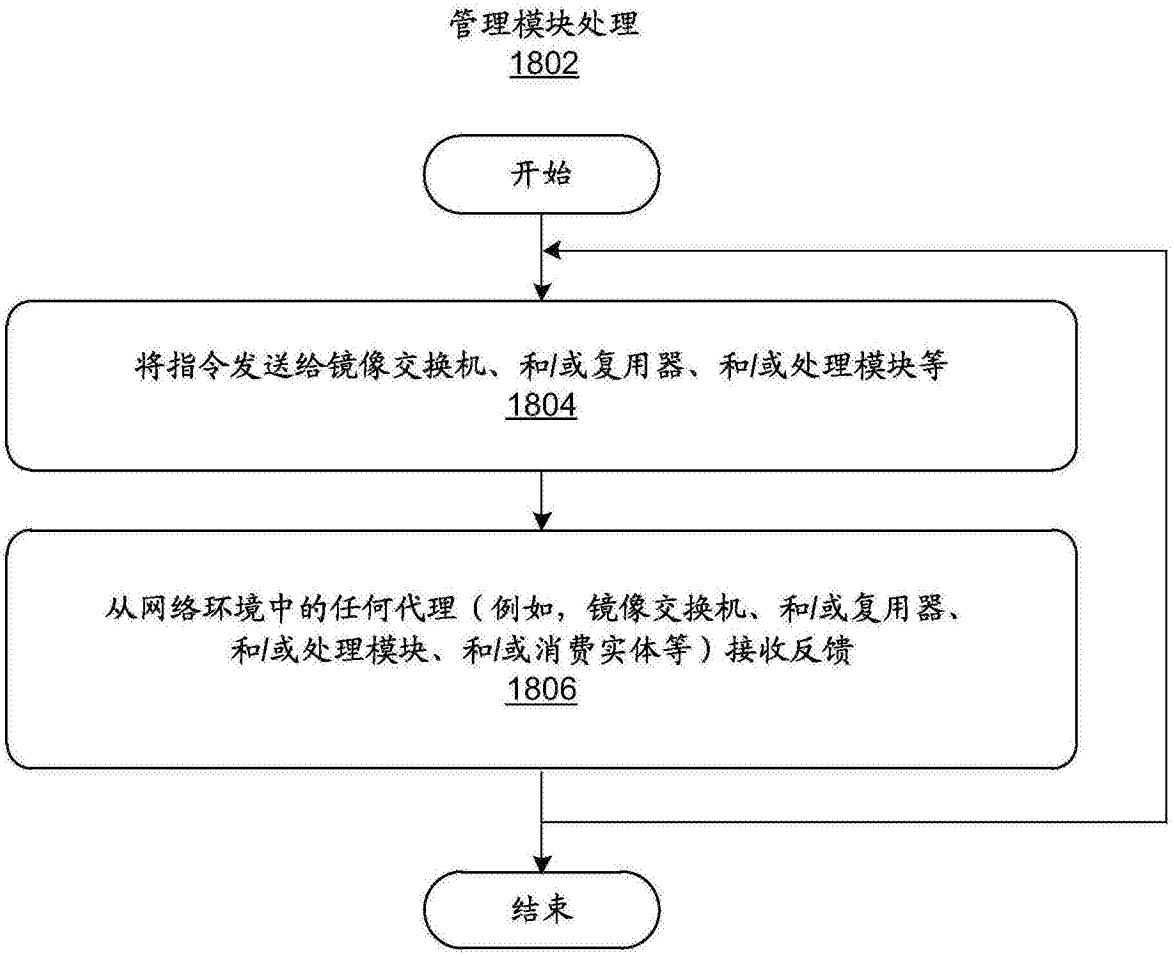

管理模块处理
1802
开始
将指令发送给镜像交换机、和/或复用器、和/或处理模块等
1804
从网络环境中的任何代理（例如，镜像交换机、和/或复用器、
和/或处理模块、和/或消费实体等）接收反馈
1806
结束

图18

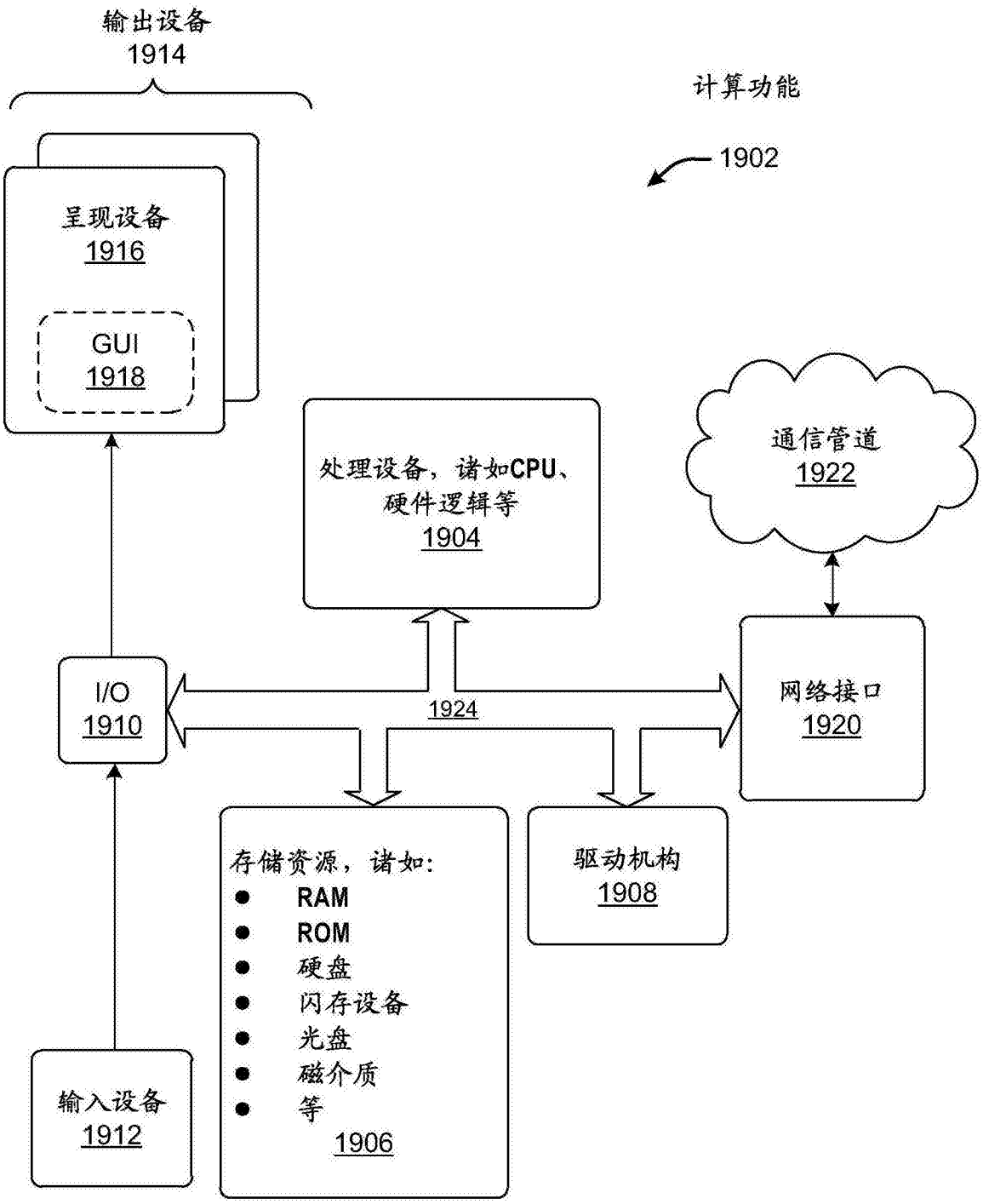

输出设备
1914
计算功能
1902
呈现设备
1916
GUI
1918
处理设备，诸如CPU、
硬件逻辑等
1904
通信管道
1922
I/O
1910
1924
网络接口
1920
存储资源，诸如:
RAM
ROM
硬盘
闪存设备
光盘
磁介质
等
1906
驱动机构
1908
输入设备
1912

图19