



(12)发明专利申请

(10)申请公布号 CN 106878106 A

(43)申请公布日 2017.06.20

(21)申请号 201710104661.2

(22)申请日 2017.02.24

(71)申请人 新华三技术有限公司

地址 310052 浙江省杭州市滨江区长河路
466号

(72)发明人 宋小恒

(74)专利代理机构 北京博思佳知识产权代理有
限公司 11415

代理人 陈蕾

(51) Int. Cl.

H04L 12/26(2006.01)

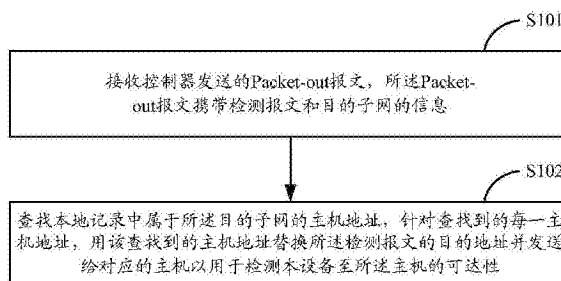
权利要求书2页 说明书6页 附图2页

(54)发明名称

一种可达性检测方法及装置

(57)摘要

本发明提供一种可达性检测方法及装置,应用于SDN中的源检测设备,该方法包括:接收控制器发送的Packet-out报文,所述Packet-out报文携带检测报文和目的子网的信息;查找本地记录中属于所述目的子网的主机地址,针对查找到的每一主机地址,用该查找到的主机地址替换所述检测报文的目的地并发送给对应的主机以用于检测本设备至所述主机的可达性。通过本发明可减少SDN控制器与源检测设备之间用于可达性检测的Packet-out报文数量。



1. 一种可达性检测方法,应用于SDN中的源检测设备,其特征在于,所述方法包括:
接收控制器发送的Packet-out报文,所述Packet-out报文携带检测报文和目的子网的信息;
查找本地记录中属于所述目的子网的主机地址,针对查找到的每一主机地址,用该查找到的主机地址替换所述检测报文的地址并发送给对应的主机以用于检测本设备至所述主机的可达性。
2. 如权利要求1所述的方法,其特征在于,所述Packet-out报文还携带VXLAN的标识,所述查找本地记录中属于所述目的子网的主机地址,包括:
在所述VXLAN内查找属于所述目的子网的主机地址。
3. 如权利要求1所述的方法,其特征在于,所述接收控制器发送的Packet-out报文,包括:
接收所述控制器发送的Bundle创建消息,创建Bundle缓存;
接收所述控制器发送的携带所述Packet-out报文的Bundle添加消息,将所述Packet-out报文添加到已创建的Bundle缓存中;
所述查找本地记录中属于所述目的子网的主机地址,包括:
接收所述控制器发送的Bundle执行消息;
根据所述Bundle缓存中Packet-out报文携带的目的子网的信息查找本地记录中属于所述目的子网的主机地址;
所述用该查找到的主机地址替换所述检测报文的地址并发送给对应的主机之后,还包括:
禁止关闭和丢弃所述Bundle缓存。
4. 如权利要求3所述的方法,其特征在于,所述方法还包括:
接收所述控制器在可达性检测时间达到预设的检测时间阈值时发送的Bundle丢弃消息,丢弃所述Bundle缓存。
5. 如权利要求3所述的方法,其特征在于,所述Packet-out报文还携带序列号修改步长,所述用该查找到的主机地址替换所述检测报文的地址并发送给对应的主机之后,还包括:
根据所述Bundle缓存中Packet-out报文携带的序列号修改步长修改检测报文的报文序号。
6. 一种可达性检测装置,应用于SDN中的源检测设备,其特征在于,所述装置包括:
接收单元,用于接收控制器发送的Packet-out报文,所述Packet-out报文携带检测报文和目的子网的信息;
查找单元,用于查找本地记录中属于所述目的子网的主机地址;
发送单元,用于针对所述查找单元查找到的每一主机地址,用该查找到的主机地址替换所述检测报文的地址并发送给对应的主机以用于检测本设备至所述主机的可达性。
7. 如权利要求6所述的装置,其特征在于,所述Packet-out报文还携带VXLAN的标识,所述查找单元,具体用于在所述VXLAN内查找属于所述目的子网的主机地址。
8. 如权利要求6所述的装置,其特征在于,所述装置还包括:
所述接收单元,具体用于接收所述控制器发送的Bundle创建消息,创建Bundle缓存;接

收所述控制器发送的携带所述Packet-out报文的Bundle添加消息,将所述Packet-out报文添加到已创建的Bundle缓存中;

所述查找单元,具体用于接收所述控制器发送的Bundle执行消息;根据所述Bundle缓存中Packet-out报文携带的目的子网的信息查找本地记录中属于所述目的子网的主机地址;

禁止单元,用于在所述发送单元用查找到的主机地址替换所述检测报文的目的地并发送给对应的主机之后,禁止关闭和丢弃所述Bundle缓存。

9. 如权利要求8所述的装置,其特征在于,所述装置还包括:

丢弃单元,用于接收所述控制器在可达性检测时间达到预设的检测时间阈值时发送的Bundle丢弃消息,丢弃所述Bundle缓存。

10. 如权利要求8所述的装置,其特征在于,所述Packet-out报文还携带序列号修改步长,所述装置还包括:

修改单元,用于在所述发送单元用查找到的主机地址替换所述检测报文的目的地并发送给对应的主机之后,根据所述Bundle缓存中Packet-out报文携带的序列号修改步长修改检测报文的报文序号。

一种可达性检测方法及装置

技术领域

[0001] 本发明涉及网络通信技术领域,尤其涉及一种可达性检测方法及装置。

背景技术

[0002] SDN(Software Defined Networks,软件定义网络)是一种新型的网络创新架构,通过将控制平面和数据平面分离,实现网络流量的灵活控制。在应用SDN技术的数据中心中,SDN控制器负责网络的智能化管理和监控,其中,可达性检测就是运维监控中的一项重要内容。

[0003] 以SDN网络中租户的网关设备和虚拟机之间的可达性检测为例,SDN控制器针对每一个虚拟机构造一个检测报文,检测报文的地址为目的虚拟机的地址,源地址为网关设备的地址,然后通过Packet-out报文将检测报文发送给网关设备,再由网关设备向目的虚拟机转发检测报文,沿途所经设备通过Packet-in报文将检测报文上报控制器,控制器根据上报的Packet-in报文确定从网关设备到目的虚拟机之间的路径是否正常,完成可达性检测。

[0004] 当租户内虚拟机的数量很大时,控制器需要发送大量的检测报文,每个检测报文需要一个Packet-out报文封装,这将占用大量的控制器资源,同时,占用控制器与网关设备之间的连接资源。倘若控制器对全网内任意虚拟机之间的可达性进行探测,其构造报文和发送报文的工作量可能是控制器无法承担的。

发明内容

[0005] 本发明的目的在于提供一种可达性检测方法及装置,用以减少控制器与源检测设备之间用于可达性检测的Packet-out报文数量。

[0006] 为实现上述发明目的,本发明提供了如下技术方案:

[0007] 本发明提供一种可达性检测方法,应用于SDN中的源检测设备,所述方法包括:

[0008] 接收控制器发送的Packet-out报文,所述Packet-out报文携带检测报文和目的子网的信息;

[0009] 查找本地记录中属于所述目的子网的主机地址,针对查找到的每一主机地址,用该查找到的主机地址替换所述检测报文的地址并发送给对应的主机以用于检测本设备至所述主机的可达性。

[0010] 本发明还提供一种可达性检测装置,应用于SDN中的源检测设备,所述装置包括:

[0011] 接收单元,用于接收控制器发送的Packet-out报文,所述Packet-out报文携带检测报文和目的子网的信息;

[0012] 查找单元,用于查找本地记录中属于所述目的子网的主机地址;

[0013] 发送单元,用于针对所述查找单元查找到的每一主机地址,用该查找到的主机地址替换所述检测报文的地址并发送给对应的主机以用于检测本设备至所述主机的可达性。

[0014] 由以上描述可以看出,本发明控制器在发送给源检测设备的Packet-out报文中携带检测报文和目的子网的信息,源检测设备根据目的子网的信息查找本地记录中属于该目的子网的主机地址,将检测报文的地址替换为查找到的主机地址发送给对应主机,以检测源检测设备至主机的可达性。在该检测过程中,控制器发送的单个Packet-out报文可触发对目的子网中多个主机的可达性检测,大大降低了控制器与源检测设备之间Packet-out报文的数量,进而减少了对控制器以及控制器与源检测设备之间网络资源的占用。

附图说明

- [0015] 图1是本发明实施例示出的一种可达性检测方法流程图;
- [0016] 图2是本发明实施例示出的一种SDN组网示意图;
- [0017] 图3是本发明实施例示出的源检测设备的结构示意图;
- [0018] 图4是本发明实施例示出的一种可达性检测装置的结构示意图。

具体实施方式

[0019] 这里将详细地对示例性实施例进行说明,其示例表示在附图中。下面的描述涉及附图时,除非另有表示,不同附图中的相同数字表示相同或相似的要素。以下示例性实施例中所描述的实施方式并不代表与本发明相一致的所有实施方式。相反,它们仅是与如所附权利要求书中所详述的、本发明的一些方面相一致的装置和方法的例子。

[0020] 在本发明使用的术语是仅仅出于描述特定实施例的目的,而非旨在限制本发明。在本发明和所附权利要求书中所使用的单数形式的“一种”、“所述”和“该”也旨在包括多数形式,除非上下文清楚地表示其他含义。还应当理解,本文中使用的术语“和/或”是指并包含一个或多个相关联的列出项目的任何或所有可能组合。

[0021] 应当理解,尽管在本发明可能采用术语第一、第二、第三等来描述各种信息,但这些信息不应限于这些术语。这些术语仅用来将同一类型的信息彼此区分开。例如,在不脱离本发明范围的情况下,第一信息也可以被称为第二信息,类似地,第二信息也可以被称为第一信息。取决于语境,如在此所使用的词语“如果”可以被解释成为“在……时”或“当……时”或“响应于确定”。

[0022] 本发明实施例提出一种可达性检测方法,该方法由控制器在发送给源检测设备的Packet-out报文中携带检测报文及目的子网的信息,源检测设备根据目的子网的信息查找本地记录的属于该目的子网的主机地址,并将检测报文的地址替换为查找到的主机地址发送给对应主机,以完成可达性检测。

[0023] 参见图1,为本发明可达性检测方法的一个实施例流程图,该实施例从源检测设备侧对可达性检测过程进行描述。本发明的源检测设备为可达性检测的起点设备,例如,当检测网关设备到主机的可达性时,网关设备为此次可达性检测的源检测设备;当检测源主机到目的主机的可达性时,源主机的接入交换设备为可达性检测的源检测设备。

[0024] 步骤101,源检测设备接收控制器发送的Packet-out报文。

[0025] 控制器构造用于可达性检测的检测报文,该检测报文的地址不作限定,源地址为源检测设备地址。控制器将检测报文携带在Packet-out报文中,并将要检测的目的子网的信息也携带在Packet-out报文中,作为一个实施例,可在Packet-out报文的Action

list(动作列表)中携带目的子网的信息,该Action list包含若干动作项,具体可利用Set-Field(设置域)动作项携带目的子网的信息(例如,10.1.10.0/24)。控制器将该Packet-out报文发送给源检测设备。

[0026] 源检测设备接收到该Packet-out报文后,获取该Packet-out报文中携带的检测报文和目的子网的信息。

[0027] 步骤102,查找本地记录中属于所述目的子网的主机地址,针对查找到的每一主机地址,用该查找到的主机地址替换所述检测报文的地址并发送给对应的主机以用于检测本设备至所述主机的可达性。

[0028] 源检测设备根据Packet-out报文携带的目的子网的信息查询本地记录中属于该目的子网的主机地址(主机上线时,会主动通告自身地址信息,源检测设备接收并记录主机通告的地址信息及出端口等信息),例如,目的子网的信息为10.1.10.0/24,假设本地记录中存在IP地址为10.1.10.11和IP地址为10.1.10.22的两台主机,则这两台主机均属于该目的子网。进一步地,在Packet-out报文中还可以携带VXLAN标识,源检测设备可基于VXLAN标识,在VXLAN内查找属于目的子网的主机地址。

[0029] 源检测设备针对查找到的每一主机地址,用查找到的主机地址替换检测报文的地址,并通过该主机地址对应的出端口将该替换地址后的检测报文发送给对应主机,以检测本设备至主机的可达性。

[0030] 在一个例子中,检测报文在从源检测设备发往主机的过程中会经过一些中间网络设备,这些中间网络设备将检测报文镜像一份通过Packet-in报文发送给控制器,控制器根据从不同网络设备接收到的Packet-in报文,确定检测报文的转发路径,从而确定源检测设备至主机的可达性。

[0031] 从上述描述可以看出,控制器与源检测设备之间通过一个Packet-out报文可触发从源检测设备到多个主机的可达性检测,减少了控制器与源检测设备之间的Packet-out报文数量,节约设备资源和带宽资源。

[0032] 当控制器需要对源检测设备到多个目的子网的主机进行可达性检测,且需要重复检测时,为了进一步减少控制器与源检测设备之间的Packet-out报文数量,本发明充分利用Bundle(捆绑)机制,并对Bundle机制进行了部分改进。

[0033] 具体为,控制器向源检测设备发送Bundle创建消息,该Bundle创建消息可通过设置Bundle控制消息(OFPT_BUNDLE_CONTROL)的消息类型来实现,此处消息类型为OFPBCT_OPEN_REQUEST。源检测设备接收到Bundle创建消息后,创建Bundle缓存。

[0034] 控制器将Packet-out报文携带在Bundle添加消息(OFPT_BUNDLE_ADD_MESSAGE)中发送给源检测设备。源检测设备接收到Bundle添加消息后,将Packet-out报文添加到已创建的Bundle缓存中。控制器通过发送多次Bundle添加消息,使源检测设备的Bundle缓存中包含多个Packet-out报文。

[0035] 在完成Packet-out报文添加后,控制器向源检测设备发送Bundle执行消息,该Bundle执行消息可通过设置Bundle控制消息(OFPT_BUNDLE_CONTROL)的消息类型为OFPBCT_EXECUTE_REQUEST来实现,该OFPBCT_EXECUTE_REQUEST为本发明新增消息类型,源检测设备在接收到Bundle执行消息后,根据Bundle缓存中Packet-out报文携带的目的子网的信息,查找本地记录中属于目的子网的主机地址,用查找到的主机地址替换检测报文的

目的地址发送给对应主机。

[0036] 需要特殊说明的是,本发明中,源检测设备在完成对Bundle缓存中所有Packet-out报文的处理后,会禁止关闭和丢弃已执行的Bundle缓存,以使控制器可通过重复发送Bundle执行消息,触发Bundle缓存中多个Packet-out报文的多次执行,进一步减少控制器与源检测设备之间的Packet-out报文数量。

[0037] 当然,控制器还可以对可达性检测时间进行统计,当统计的可达性检测时间达到预设的检测时间阈值时,向源检测设备下发Bundle丢弃消息。源检测设备根据该Bundle丢弃消息,丢弃本设备中的Bundle缓存,以节约源检测设备的内存空间。例如,在网络组建初期,设置检测时间阈值为3天,在3天内对网络进行频繁检测,3天后网络趋于稳定,停止可达性检测,源检测设备丢弃用于可达性检测的Bundle缓存。

[0038] 此外,为了使源检测设备在每次接收到控制器发送的Bundle执行消息时,可以自动修改Packet-out报文中检测报文的报文序号,控制器在发送的Packet-out报文中还携带了序列号修改步长,作为一个实施例,该序列号修改步长可通过Packet-out报文中的Action list携带,具体可利用EXPERIMENTER(实验)动作项携带,例如,定义序列号修改步长为1。源检测设备在执行Bundle缓存中的Packet-out报文后,根据Packet-out报文中携带的序列号修改步长修改Packet-out报文中检测报文的报文序号,当源检测设备再次执行Packet-out报文时,基于上次修改的报文序号发送检测报文给主机。

[0039] 检测报文在从源检测设备发往主机的过程中会经过一些中间网络设备,这些中间网络设备将检测报文镜像一份通过Packet-in报文发送给控制器,控制器根据从不同网络设备接收到的Packet-in报文,确定检测报文的转发路径。同时,根据上送的检测报文的报文序号可以进一步判断转发路径的传输性能是否稳定,例如,针对某一主机发送100次检测报文(报文序号1~100),控制器未接收到中间网络设备上送的报文序号为5和10的检测报文,或者根据报文序号为5和10的检测报文进行检测时未得到完整的转发路径,说明从源检测设备到该主机的转发路径存在一定的不稳定因素。此外,对于从同一网络设备上送的具有相同报文序号的检测报文,控制器只会处理一次,以过滤掉网络不稳定因素的影响。

[0040] 综上所述可知,本发明通过在Packet-out报文中携带目的子网的信息,可一次触发从源检测设备到目的子网中多个主机的可达性检测,从而减少控制器与源检测设备之间Packet-out报文的数量,此外,通过改进现有Bundle机制,使Bundle具有容器功能,可以被重复使用,从而进一步减少控制器与源检测设备之间Packet-out报文的数量。

[0041] 现以图2所示SDN网络为例,介绍从网关设备到虚拟机的可达性检测过程。

[0042] 图2所示SDN网络包括控制器Controller、网关设备GW、路由器R1和R2、接入交换设备SW1和SW2。为了简化描述,假设图中仅示出一个租户的一个VXLAN网络(记为VXLAN10)内的虚拟机VM10~VM12以及VM20~VM22,其中,VM10和VM20属于子网1(10.1.1.0/24);其它VM属于子网2(10.1.2.0/24)。

[0043] Controller向GW发送消息类型为OFPBCT_OPEN_REQUEST的OFPT_BUNDLE_CONTROL消息,GW根据该消息创建Bundle缓存,记为Bundle1。

[0044] Controller针对每一个子网构造Packet-out报文,以子网1为例:生成检测报文,记为Packet1,Packet1的源IP地址为GW的IP地址,目的IP地址不限定。将Packet1携带在Packet-out报文(记为Packet-out1)中,在Packet-out1的Action List的Set-Field动作项

中携带子网1的网段信息(10.1.1.0/24)、VXLAN ID (VXLAN10);在EXPERIMENTER动作项中携带序列号修改步长为1。将Packet-out1携带在OFPT_BUNDLE_ADD_MESSAGE中发送给GW,GW在接收到OFPT_BUNDLE_ADD_MESSAGE后,将携带的Packet-out1添加到Bundle1中。

[0045] 同理,Controller构建子网2的Packet-out报文(记为Packet-out2),通过OFPT_BUNDLE_ADD_MESSAGE消息发送给GW,GW在接收到OFPT_BUNDLE_ADD_MESSAGE后,GW将Packet-out2添加到Bundle1中。以此类推,当SDN网络有多个租户的多个VXLAN的多个子网需要进行可达性检测时,可构建若干Packet-out报文添加到Bundle1中。

[0046] 在完成Packet-out报文添加后,Controller向GW发送消息类型为OFPBCT_EXECUTE_REQUEST的OFPT_BUNDLE_CONTROL消息,触发GW执行Bundle1中的所有Packet-out报文。

[0047] 以Packet-out1为例,GW根据Packet-out1的Action List中的Set-Field动作项,获取携带的VXLAN ID (VXLAN10)、子网信息(10.1.1.0/24),GW在VXLAN10内遍历属于子网10.1.1.0/24的虚拟机,确认VM10和VM20属于子网10.1.1.0/24。

[0048] GW根据本地记录的VM10的IP地址和MAC地址替换Packet1的目的IP地址和目的MAC地址,生成新的检测报文,记为Packet1_1,将Packet1_1发送给VM10;同理,根据本地记录的VM20的IP地址和MAC地址替换Packet1的目的IP地址和目的MAC地址,生成新的检测报文,记为Packet1_2,将Packet1_2发送给VM20。

[0049] Packet1_1沿R1、SW1转发,R1和SW1均将Packet1_1镜像一份通过Packet-in报文发送给Controller,因此,Controller可获知Packet1_1的转发路径为GW→R1→SW1;同理,获知Packet1_2的转发路径为GW→R2→SW2,完成可达性检测。

[0050] 需要补充说明的是,GW在生成Packet1_1和Packet1_2后,根据Packet-out1的Action List中EXPERIMENTER动作项的序列号修改步长,修改Bundle1中Packet-out1消息携带的Packet1的报文序号,报文序号加1,该报文序号的变化会导致下一次检测时以该新的报文序号生成Packet1_1和Packet1_2,以使Controller可区分出是哪一次的检测。

[0051] Bundle1中所有Packet-out报文的处理过程与上述Packet-out1的处理过程相同。Controller向GW发送一次消息类型为OFPBCT_EXECUTE_REQUEST的OFPT_BUNDLE_CONTROL消息可触发Bundle1中所有Packet-out报文的执行。

[0052] 由于Bundle1执行后不会自动关闭和丢弃,因此,Controller可通过多次发送消息类型为OFPBCT_EXECUTE_REQUEST的OFPT_BUNDLE_CONTROL消息触发GW执行一次对Bundle1所有Packet-out报文的处理,从而避免Controller针对每一个子网重新下发Packet-out报文,最大程度地减少Controller和GW之间Packet-out报文的数量,提升Controller对整个SDN网络的检测效率。

[0053] 与前述可达性检测方法的实施例相对应,本发明还提供了可达性检测装置的实施例。

[0054] 本发明可达性检测装置的实施例可以应用在源检测设备上。装置实施例可以通过软件实现,也可以通过硬件或者软硬件结合的方式实现。以软件实现为例,作为一个逻辑意义上的装置,是通过其所在设备的处理器运行存储器中对应的计算机程序指令形成的。从硬件层面而言,如图3所示,为本发明可达性检测装置所在设备的一种硬件结构图,除了图3所示的处理器以及非易失性存储器之外,实施例中装置所在的设备通常根据该设备的实际

功能,还可以包括其他硬件,对此不再赘述。

[0055] 请参考图4,为本发明一个实施例中的可达性检测装置的结构示意图。该可达性检测装置包括接收单元401、查找单元402以及发送单元403,其中:

[0056] 接收单元401,用于接收控制器发送的Packet-out报文,所述Packet-out报文携带检测报文和目的子网的信息;

[0057] 查找单元402,用于查找本地记录中属于所述目的子网的主机地址;

[0058] 发送单元403,用于针对所述查找单元402查找到的每一主机地址,用该查找到的主机地址替换所述检测报文的地址并发送给对应的主机以用于检测本设备至所述主机的可达性。

[0059] 进一步地,所述Packet-out报文还携带VXLAN的标识,

[0060] 所述查找单元402,具体用于在所述VXLAN内查找属于所述目的子网的主机地址。

[0061] 进一步地,所述装置还包括:

[0062] 所述接收单元401,具体用于接收所述控制器发送的Bundle创建消息,创建Bundle缓存;接收所述控制器发送的携带所述Packet-out报文的Bundle添加消息,将所述Packet-out报文添加到已创建的Bundle缓存中;

[0063] 所述查找单元402,具体用于接收所述控制器发送的Bundle执行消息;根据所述Bundle缓存中Packet-out报文携带的目的子网的信息查找本地记录中属于所述目的子网的主机地址;

[0064] 禁止单元,用于在所述发送单元403用查找到的主机地址替换所述检测报文的地址并发送给对应的主机之后,禁止关闭和丢弃所述Bundle缓存。

[0065] 进一步地,所述装置还包括:

[0066] 丢弃单元,用于接收所述控制器在可达性检测时间达到预设的检测时间阈值时发送的Bundle丢弃消息,丢弃所述Bundle缓存。

[0067] 进一步地,所述Packet-out报文还携带序列号修改步长,所述装置还包括:

[0068] 修改单元,用于在所述发送单元403用查找到的主机地址替换所述检测报文的地址并发送给对应的主机之后,根据所述Bundle缓存中Packet-out报文携带的序列号修改步长修改检测报文的报文序号。

[0069] 上述装置中各个单元的功能和作用的实现过程具体详见上述方法中对应步骤的实现过程,在此不再赘述。

[0070] 对于装置实施例而言,由于其基本对应于方法实施例,所以相关之处参见方法实施例的部分说明即可。以上所描述的装置实施例仅仅是示意性的,其中所述作为分离部件说明的单元可以是或者也可以不是物理上分开的,作为单元显示的部件可以是或者也可以不是物理单元,即可以位于一个地方,或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本发明方案的目的。本领域普通技术人员在不付出创造性劳动的情况下,即可以理解并实施。

[0071] 以上所述仅为本发明的较佳实施例而已,并不用以限制本发明,凡在本发明的精神和原则之内,所做的任何修改、等同替换、改进等,均应包含在本发明保护的范围之内。

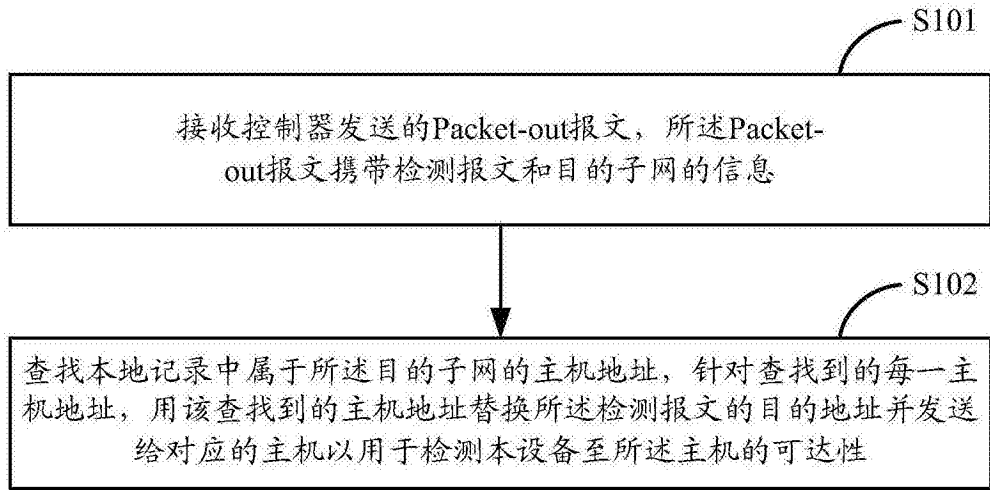


图1

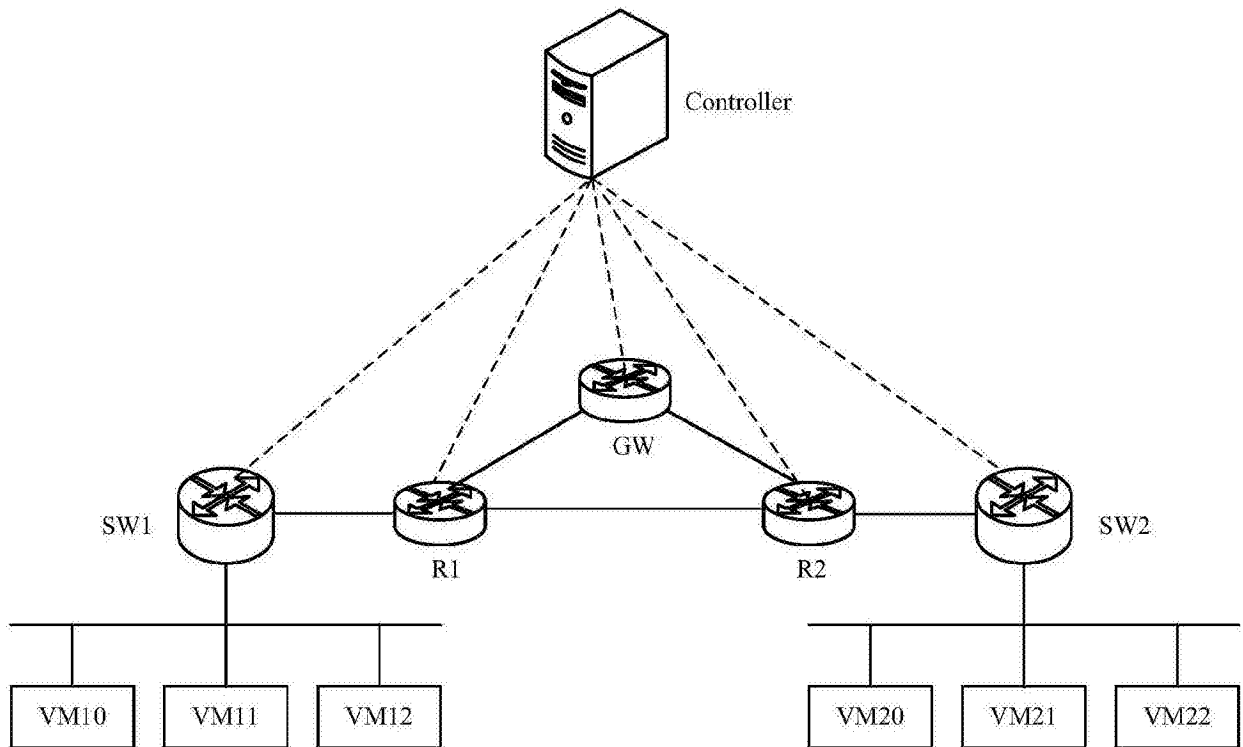


图2

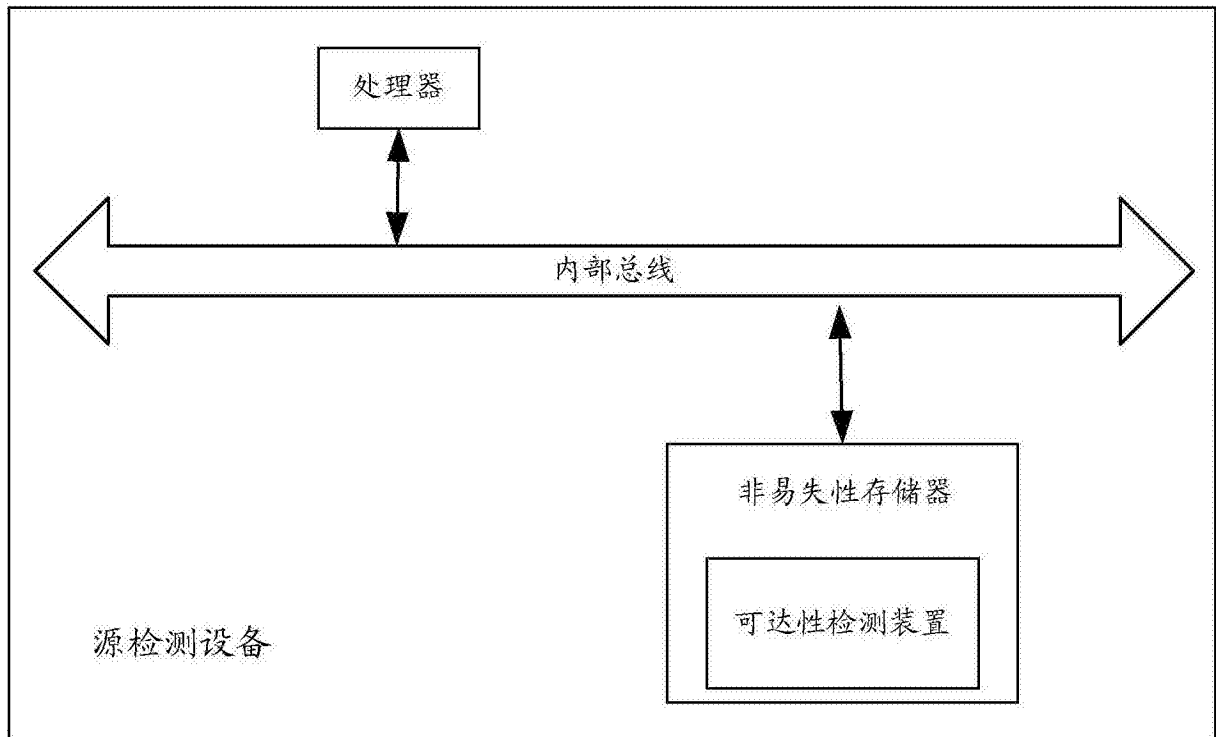


图3

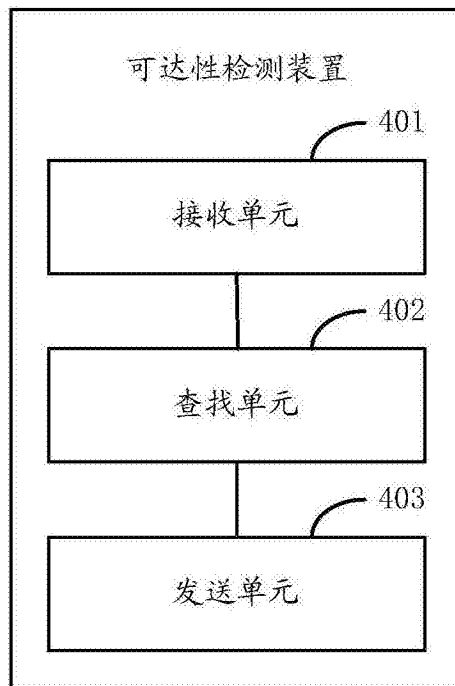


图4