(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 106020926 A
(43)申请公布日 2016.10.12

(21)申请号 201610286786.7

(22)申请日 2016.04.29

(71)申请人 华为技术有限公司
地址 518129 广东省深圳市龙岗区坂田华为总部办公楼

(72)发明人 张明 陆丽娜

(74)专利代理机构 北京同达信恒知识产权代理有限公司 11291
代理人 冯艳莲

(51)Int.Cl.
G06F 9/455(2006.01)
G06F 9/54(2006.01)

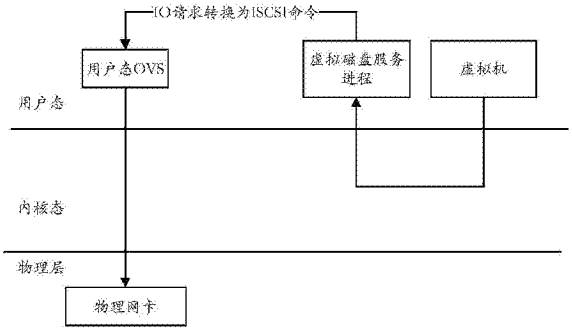
权利要求书2页 说明书10页 附图5页

(54)发明名称

一种用于虚拟交换机技术中数据传输的方法及装置

(57)摘要

本发明公开了一种用于虚拟交换机技术中数据传输的方法及装置，该方法包括：接收虚拟机VM访问文件或者磁盘的IO请求，确定是否通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机以及跨主机的虚拟机之间的网络互通；如果确定通过用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为网络小型计算机系统接口ISCSI命令后发送到所述用户态OVS；所述用户态OVS将所述ISCSI命令发送到所述物理网卡。本发明公开的方法及装置解决现有技术中用户态OVS的虚拟机处理IO请求的性能比较低的问题。



CN 106020926 A

1.一种用于虚拟交换机技术中数据传输的方法，其特征在于，所述该方法包括：

接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为网络小型计算机系统接口ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

2.如权利要求1所述的方法，其特征在于，所述确定是否通过用户态OVS将所述IO请求发送到物理网卡包括：

确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息，根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘，如果是，则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

3.如权利要求1或2所述的方法，其特征在于，所述将所述IO请求转换为网络小型计算机系统接口ISCSI命令包括：

将所述IO请求转换成小型计算机系统接口SCSI命令；

在所述SCSI命令中增加网络小型计算机系统接口ISCSI头以获得所述ISCSI命令。

4.一种用于虚拟交换机技术中数据传输的方法，其特征在于，所述方法包括：

接收物理网卡针对IO请求响应的网络小型计算机系统接口ISCSI报文；其中，所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机。

5.如权利要求4所述的方法，其特征在于，所述在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机包括：

将所述ISCSI报文转换成SCSI响应；

将所述SCSI响应转换为IO响应，并将所述IO响应发送到发起所述IO请求的虚拟机。

6.一种电子设备，其特征在于，所述电子设备包括：

确认模块，用于接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

转换模块，用于如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为网络小型计算机系统接口ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

7.如权利要求6所述的电子设备，其特征在于，所述确认模块具体用于确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息，根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘，如果是，则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

8.如权利要求6或7所述的电子设备，其特征在于，所述转换模块具体用于将所述IO请求转换成小型计算机系统接口SCSI命令；在所述SCSI命令中增加网络小型计算机系统接口ISCSI头以获得所述ISCSI命令。

9.一种电子设备，其特征在于，包括：

接收模块，用于接收物理网卡针对IO请求响应的网络小型计算机系统接口ISCSI报文；其中，所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

转换模块，用于在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机。

10.如权利要求9所述的电子设备，其特征在于，所述转换模块具体用于将所述ISCSI报文转换成SCSI响应；将所述SCSI响应转换为IO响应，并将所述IO响应发送到发起所述IO请求的虚拟机。

一种用于虚拟交换机技术中数据传输的方法及装置

技术领域

[0001] 本发明涉及通信技术领域，尤其涉及一种虚拟交换机技术中数据传输方法及装置。

背景技术

[0002] 虚拟交换机主要应用于服务器虚拟化场景，主要实现两个功能：功能1，传递虚拟机(Virtual Machine，VM)之间的网络流量；功能2，实现VM与外界网络的通信。

[0003] Linux实现中把运行空间分成两部分：用户态和内核态。其中，用户态开放虚拟交换机(Open vSwitch，OVS)是指该交换机的数据面转发功能在Linux的用户态完成，并且该用户态OVS采用数据平面开发套件(Data Plane Development Kit，DPDK)作为数据通道与物理网卡连接；另外，该用户态OVS的管理平台与内核态OVS共用，该用户态OVS设置有限速功能能够实现对网卡限速，从而为虚拟化场景提供高性能的虚拟交换机方案。

[0004] DPDK是一组快速包处理的库和驱动程序的集合。运行在Linux用户空间的Linux内核中断机制无法满足大吞吐量网卡(例如40GE)性能要求，Intel提出DPDK就是为了解决该问题，采用DPDK后可以达到网卡限速。

[0005] 现有用户态OVS方案中，虚拟机处理IO请求的路径如图1所示，虚拟磁盘服务进程接收虚拟机发送的IO请求，并将该IO请求放到存储栈；存储栈将该IO请求转换为IO报文后提交给IO协议栈；然后IO协议栈再通过socket机制将该IO报文传递给用户态OVS，最后由用户态OVS通过DPDK将报文通过物理网卡发送出去。

[0006] 通过用户态OVS实现IO报文转发的过程中，IO报文需要从用户态OVS切换到内核态，再由内核态切换到用户态；所以需要经过上下文切换；同时将IO报文从用户态转发到内核态的存储栈时，需要对IO报文进行拷贝，所以现有技术中存在用户态OVS的虚拟机处理IO请求的性能比较低的问题。

发明内容

[0007] 本发明实施例提供一种用于虚拟交换机技术中数据传输的方法及装置，本发明所提供的方法及装置解决现有技术中存在用户态OVS的虚拟机处理IO请求的性能比较低的问题。

[0008] 第一方面，提供一种用于虚拟交换机技术中数据传输的方法，该方法包括：接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0009] 如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为网络小型计算机系统接口ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

[0010] 在该实施例中，虚拟机进行IO请求的处理时，通过判断IO请求是否为通过用户态

OVS转发的，如果虚拟机则直接在用户态将IO请求转化为用户态OVS能够处理的格式，然后发送到用户态OVS进行处理。能够避免现有技术中IO请求转发时的上下文切换，从而能够有效的提高虚拟机处理IO请求的能力。

[0011] 结合第一方面，在一个可能的设计中，将所述IO请求转换为网络小型计算机系统接口ISCSI命令包括：将所述IO请求转换成小型计算机系统接口SCSI命令；在所述SCSI命令中增加网络小型计算机系统接口ISCSI头以获得所述ISCSI命令。

[0012] 可选的，为了实现对IO请求的分类，将IO请求中需要通过用户态OVS发送的IO请求分离出来，则所述确定是否通过用户态OVS将所述IO请求发送到物理网卡包括：

[0013] 确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息，根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘，如果是，则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

[0014] 第二方面，基于第一方面中IO请求的发送过程，在第二方面提供一种用于虚拟交换机技术中数据传输的方法，该方法实现IO请求的响应处理，该方法具体包括：接收物理网卡针对IO请求响应的网络小型计算机系统接口ISCSI报文；其中，所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0015] 在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机。

[0016] 结合第二方面，在一个可能的设计中，所述在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机包括：

[0017] 将所述ISCSI报文转换成SCSI响应；

[0018] 将所述SCSI响应转换为IO响应，并将所述IO响应发送到发起所述IO请求的虚拟机。

[0019] 第一方面提供的方法是IO请求的发起过程，第二方面所提供方法是IO请求的响应处理过程，所以在具体实现时第二方面所提供方法的具体效果与第一方面方法相同。

[0020] 第三方面，提供一种电子设备，该电子设备包括：

[0021] 确认模块，用于接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0022] 转换模块，用于如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为网络小型计算机系统接口ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

[0023] 结合第三方面，在一个可能的设计中，所述确认模块具体用于确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息，根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘，如果是，则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

[0024] 结合第三方面，在一个可能的设计中，所述转换模块具体用于将所述IO请求转换成小型计算机系统接口SCSI命令；在所述SCSI命令中增加网络小型计算机系统接口ISCSI头以获得所述ISCSI命令。

[0025] 第四方面，提供一种电子设备，包括：

[0026] 接收模块，用于接收物理网卡针对IO请求响应的网络小型计算机系统接口ISCSI报文；其中，所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0027] 转换模块，用于在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机。

[0028] 在一个可能的设计中，所述转换模块具体用于将所述ISCSI报文转换成SCSI响应；将所述SCSI响应转换为IO响应，并将所述IO响应发送到发起所述IO请求的虚拟机。

[0029] 第五方面，提供了一种计算机可读存储介质，所述可读存储介质上存储有实现第一方面描述的用于虚拟交换机技术中数据传输的方法的程序代码，该程序代码包含运行第一方面描述的用于虚拟交换机技术中数据传输的方法的执行指令。

[0030] 第六方面，提供了一种计算机可读存储介质，所述可读存储介质上存储有实现第二方面描述的用于虚拟交换机技术中数据传输的方法的程序代码，该程序代码包含运行第二方面描述的用于虚拟交换机技术中数据传输的方法的执行指令。

[0031] 第七方面，本发明实施例提供了一种计算机存储介质，用于储存为上述基站所用的计算机软件指令，其包含用于执行上述方面所设计的程序。

[0032] 第八方面，提供一种电子设备，该电子设备包括虚拟机、虚拟机监视器和硬件层，其中网络接口设备具体可以为物理网卡。

[0033] 该虚拟机监视器可以运行在两个状态：用户态和内核态，在该实施例中，为了提高虚拟机处理IO请求的能力，在该虚拟机监视器具体用于：

[0034] 接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为网络小型计算机系统接口ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

[0035] 第八方面的另外一种设计中，该虚拟机监视器用于确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息，根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘，如果是，则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

[0036] 第八方面的另外一种设计中，该虚拟机监视器用于将所述IO请求转换成小型计算机系统接口SCSI命令；在所述SCSI命令中增加网络小型计算机系统接口ISCSI头以获得所述ISCSI命令。

[0037] 因为IO请求的响应的具体实现是和上述实例的电子设备结构是一样的，只是虚拟机监视器320在处理物理网卡反馈的响应时的处理方式不同，该虚拟机监视器，用于接收物理网卡针对IO请求响应的网络小型计算机系统接口ISCSI报文；并在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机；其中，所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通。

[0038] 第八方面的另外一种设计中，该虚拟机监视器用于将所述ISCSI报文转换成SCSI

响应;将所述SCSI响应转换为IO响应,并将所述IO响应发送到发起所述IO请求的虚拟机。

[0039] 上述技术方案中的一个或两个,至少具有如下技术效果:

[0040] 本发明实施例提供的方法和装置,将需要通过用户态OVS发送到物理网卡的IO请求,直接在用户态将该IO请求转换为用户态OVS可以处理的格式后,发送到用户态OVS。这样避免了将IO请求从用户态切换到内核态,再由内核态切换到用户态的过程,所以能够有效地提升虚拟机处理IO请求的性能。

附图说明

[0041] 图1为现有用户态OVS方案中虚拟机处理IO请求的路径示意图;

[0042] 图2为本发明实施例所实用的服务器虚拟化场景的实现系统结构示意图;

[0043] 图3为本发明实施例提供的一种电子设备的结构示意图;

[0044] 图4为本发明实施例另外一种电子设备的结构示意图;

[0045] 图5为本发明实施例提供另外一种电子设备的结构示意图;

[0046] 图6为本发明实施例提供另外一种电子设备的结构示意图;

[0047] 图7为本发明实施例提供一种用于虚拟交换机技术中数据传输的方法的流程图;

[0048] 图8本发明实施例中用于虚拟交换机技术中数据传输的方法的流程示意图;

[0049] 图9为本发明实施例提供的另外一种用于虚拟交换机技术中数据传输的方法流程示意图;

[0050] 图10为本发明实施例提供的虚拟机处理IO请求的路径示意图。

具体实施方式

[0051] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行描述,显然,所描述的实施例是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有作出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0052] 现有用户态OVS方案中,虚拟机处理IO请求的性能比较低问题,主要的原因包括:A,IO请求从虚拟磁盘服务进程(tapdisk2)发出后,存在两次上下文切换(即用户态到内核态,再由内核态切换到用户态);B,IO请求转发过程中从存储栈到用户态OVS存在拆报文的行为;C,而且通过图1所示的流程进行虚拟机IO请求处理存在两次服务质量(Quality of Service,QoS)控制,存储栈中存在QoS控制,同时用户态OVS也存在QoS控制能力。所以在一定程度上影响了虚拟机处理IO请求的性能。

[0053] 基于上述原因,本发明实施例所提供的方案中,如果有需要通过用户态OVS发送到物理网卡的IO请求,则直接在用户态将该IO请求转换为可供用户态OVS处理的格式后发送到用户态OVS。这样避免了将IO请求从用户态切换到内核态,再由内核态切换到用户态的过程,从而也能避免上述几个问题,所以能够有效地提升虚拟机处理IO请求的性能。具体实现参照以下实例:

[0054] 实施例一

[0055] 本发明实施例中服务器虚拟化场景的实现系统的结构如图2所示。其中,用户态OVS实现虚拟机之间的网络互通,包括同主机上的虚拟机和跨主机的虚拟机之间。虚拟机的

虚拟磁盘文件存储在网络之间互连的协议存储局域网络(Internet Protocol Storage Area Network,IPSAN)设备上,虚拟机处理IO请求通过用户态OVS被发送到IPSAN设备上,实现虚拟存储访问。基于图2所示的系统结构,本发明实施例提供一种电子设备,该电子设备具体实现可以是:

[0056] 如图3所示,本发明实施例提供一种电子设备,该电子设备包括虚拟机310、虚拟机监视器320和硬件层330。该虚拟机310是基于硬件层330构建的,并且虚拟机监视器320实现虚拟机310和硬件层330之间的数据监控和传输。硬件层330包括处理器331、物理内存332、硬盘333和网络接口设备334等,其中网络接口设备334具体可以为物理网卡。

[0057] 该虚拟机监视器320可以运行在两个状态:用户态和内核态,在该实施例中,为了提高虚拟机处理IO请求的能力,在该虚拟机监视器320具体用于:

[0058] 接收虚拟机VM发送的IO请求,确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡;(在该实施例中,该物理网卡可以是图3中的网络接口设备334中的一种);其中,所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通;如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡,则在用户态将所述IO请求转换为网络小型计算机系统接口(Internet Small Computer System Interface,ISCSI)命令后发送到所述用户态OVS,以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

[0059] 可选的,为了该虚拟机监视器320确定是否通过用户态OVS将所述IO请求发送到物理网卡的具体实现为:

[0060] 确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息,根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘,如果是,则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

[0061] 可选的,为了该虚拟机监视器320将所述IO请求转换为网络小型计算机系统接口ISCSI命令包括:

[0062] 将所述IO请求转换成小型计算机系统接口SCSI命令;

[0063] 在所述SCSI命令中增加网络小型计算机系统接口ISCSI头以获得所述ISCSI命令。

[0064] 因为IO请求的响应的具体实现是和上述实例的电子设备结构是一样的,只是虚拟机监视器320在处理物理网卡反馈的响应时的处理方式不同,所以基于图3所示的结构,IO请求的响应的具体实现可以是:

[0065] 虚拟机监视器320,用于接收物理网卡针对IO请求响应的网络小型计算机系统接口ISCSI报文;并在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机;其中,所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机,所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通。

[0066] 可选的,虚拟机监视器320在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机具体实现包括:

[0067] 将所述ISCSI报文转换成SCSI响应;

[0068] 将所述SCSI响应转换为IO响应,并将所述IO响应发送到发起所述IO请求的虚拟机。

[0069] 本发明实施例提供的方案中,如果有需要通过用户态OVS发送到物理网卡的IO请

求，则直接在用户态将该IO请求转换为可供用户态OVS处理的格式后发送到用户态OVS。这样避免了将IO请求从用户态切换到内核态，再由内核态切换到用户态的过程，从而避免现有技术中存在用户态OVS的虚拟机处理IO请求的性能比较低的问题，所以能够有效地提升虚拟机处理IO请求的性能。

[0070] 实施例二

[0071] 如图4所示，本发明实施例提供一种电子设备，该电子设备具体可以包括：

[0072] 确认模块401，用于接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0073] 可选的，该确认模块401确定IO请求是否为需要通过用户态OVS发送到物理网卡时，具体实现可以是：

[0074] 该确认模块401具体用于确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息，根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘，如果是，则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

[0075] 转换模块402，用于如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为网络小型计算机系统接口ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

[0076] 可选的，该转换模块402具体用于将所述IO请求转换成小型计算机系统接口SCSI命令；在所述SCSI命令中增加网络小型计算机系统接口ISCSI头以获得所述ISCSI命令。

[0077] 实施例三

[0078] 如图5所示，本发明实施例提供一种电子设备，该电子设备具体实现可以包括：

[0079] 接收模块501，用于接收物理网卡针对IO请求响应的网络小型计算机系统接口ISCSI报文；其中，所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0080] 转换模块502，用于在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机。

[0081] 可选的，将ISCSI报文转换为IO响应的具体实现可以是：

[0082] 该转换模块502具体用于将所述ISCSI报文转换成SCSI响应；将所述SCSI响应转换为IO响应，并将所述IO响应发送到发起所述IO请求的虚拟机。

[0083] 实施例四

[0084] 如图6所示，本发明还提供另一种电子设备，用于执行前述各个实施例中的虚拟交换机技术中数据传输的方法，该电子设备包括至少一个处理器601(例如CPU)，至少一个网络接口602或者其他通信接口，存储器603，和至少一个通信总线604，用于实现这些装置之间的连接通信。处理器601用于执行存储器603中存储的可执行模块，例如计算机程序。存储器603可能包含高速随机存取存储器(RAM:Random Access Memory)，也可能还包括非不稳定的存储器(non-volatile memory)，例如至少一个磁盘存储器。通过至少一个网络接口602(可以是有线或者无线)实现该系统网关与至少一个其他网元之间的通信连接，可以使用互联网，广域网，本地网，城域网等。

[0085] 在一些实施方式中，存储器存储了程序6031，程序可以被处理器执行，这个程序包括：

[0086] 接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0087] 如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为网络小型计算机系统接口ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

[0088] 可选的，所述确定是否通过用户态OVS将所述IO请求发送到物理网卡包括：

[0089] 确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息，根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘，如果是，则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

[0090] 可选的，所述将所述IO请求转换为网络小型计算机系统接口ISCSI命令包括：

[0091] 将所述IO请求转换成小型计算机系统接口SCSI命令；

[0092] 在所述SCSI命令中增加网络小型计算机系统接口ISCSI头以获得所述ISCSI命令。

[0093] 基于图6所示的结构，为了处理物理网卡返回的IO请求的响应，对应存储器存储的程序6031还包括：

[0094] 接收物理网卡针对IO请求响应的网络小型计算机系统接口ISCSI报文；其中，所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0095] 在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机。

[0096] 可选的，所述在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机包括：

[0097] 将所述ISCSI报文转换成SCSI响应；

[0098] 将所述SCSI响应转换为IO响应，并将所述IO响应发送到发起所述IO请求的虚拟机。

[0099] 实施例五

[0100] 基于上述实施例所提供的装置结构，本发明实施例提供一种用于虚拟交换机技术中数据传输的方法，该方法具体包括以下步骤(实现流程如图7所示)：

[0101] 在本发明实施例所提供的方法可基于图3、图4和图6所示装置结构实现，其中，基于不同的装置执行本发明实施例所述方法的具体功能模块不同相同，具体实现本发明方法的功能模块可以参见上述实施例一到实施例四中不同装置的具体描述。为了方便描述以下结合实施例三对本发明实施例提供的一种用于虚拟交换机技术中数据传输的方法进行详细的描述：

[0102] 步骤701，虚拟机监视器接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡；其中，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0103] 可选的，所述确定是否通过用户态OVS将所述IO请求发送到物理网卡包括：

[0104] 确定所述IO请求所访问的磁盘或访问的文件所在磁盘的磁盘信息，根据所述磁盘信息判断所述IO请求所访问的磁盘是否为远端设备的磁盘，如果是，则确定所述IO请求需要通过所述用户态OVS发送到所述物理网卡。

[0105] 基于图2所示的系统结构，该远端设备可以是通过交换机与实现本发明实施例方法的电子设备连接的设备。

[0106] 在该实施例中，如果通过磁盘信息确定IO请求是访问设备本地磁盘，则直接通过内核态将IO请求转发到物理网卡。

[0107] 步骤702，虚拟机监视器如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡。

[0108] 为了保证IO请求能在网络上传输，IO请求跨主机的时候，需要把IO请求转换为ISCSI命令，然后到目标服务器的时候，再把报文转换为IO请求。在该实施例中，为了将tapdisk2发出的IO请求，转换为用户态OVS能够识别的ISCSI命令，在该实施例所提供的方案中，可以在用户态对该IO请求进行转换。

[0109] a1，将所述IO请求转换成小型计算机系统接口(Small Computer System Interface，SCSI)命令。

[0110] 具体的，从该IO请求中获取该IO请求对应的主次设备号(包括发起访问的存储块设备ID和IO请求需要访问的存储块设备ID)、起始扇区、请求大小、请求是读还是写；

[0111] 将该IO请求转换为SCSI命令的具体实现可以是，将IO请求需要访问的存储块设备的ID转换为对应的SCSI命令中的目的主机逻辑单元号(Logical Unit Number，LUN)，起始扇区需要转换为物理磁盘地址，读或写需要转为SCSI操作方向，实现上述转换后，将转换后的内容组装成SCSI命令。

[0112] a2，将SCSI命令转换为ISCSI命令。具体实现为：

[0113] 在SCSI命令中增加ISCSI头，ISCSI头是根据ISCSI协议组装的，ISCSI头包括基本头段(Basic Header Segment，BHS)、附加标题段(Additional Header Segment，AHS)、报文头摘要(Header-Digest)和数据摘要(Data-Digest)等。

[0114] 区别于图1所示的IO请求发送路径，本发明实施例所提供的方法中，虚拟磁盘服务发出的IO请求可以在用户态直接发送到用户态OVS，具体示意可以如图8所示。

[0115] 通过本发明实施例提供的方法，将需要通过用户态OVS发送到物理网卡的IO请求，直接在用户态将该IO请求转换为用户态OVS可以处理的格式后，发送到用户态OVS。这样避免了将IO请求从用户态切换到内核态，再由内核态切换到用户态的过程，所以能够有效地提升虚拟机处理IO请求的性能。

[0116] 实施例六

[0117] 如图9所示，本发明实施例还提供一种用于虚拟交换机技术中数据传输的方法，该方法具体包括以下步骤：

[0118] 本发明实施例所提供的方法可基于图3、图5和图6所示装置结构实现，其中，基于不同的装置执行本发明实施例所述方法的具体功能模块不相同，具体实现本发明方法的功能模块可以参见上述实施例一到实施例四中不同装置的具体描述。为了方便描述以下结合实施例三对本发明实施例提供的一种用于虚拟交换机技术中数据传输的方法进行详细的

描述：

[0119] 步骤901，虚拟机监视器接收物理网卡针对IO请求响应的ISCSI报文；其中，所述ISCSI报文需要通过用户态开放虚拟交换机OVS发送到发起所述IO请求的虚拟机，所述用户态OVS实现同主机上的虚拟机之间或跨主机的虚拟机之间的网络互通；

[0120] 步骤902，在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机。

[0121] 该实施例中，所述在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的虚拟机包括：

[0122] 将所述ISCSI报文转换成SCSI响应；

[0123] 将所述SCSI响应转换为IO响应，并将所述IO响应发送到发起所述IO请求的虚拟机。

[0124] 该实施例所提供的方法是针对IO请求的响应流程，该实施例的方案可以独立实现也可以和实施例一所提供的方法进行结合，当该实施例方法与实施例一方法结合时候可以实现IO请求和请求响应的完整过程。

[0125] 区别于图8所示的IO请求发送路径，本发明实施例所提供的方法中，电子设备在处理IO请求的响应，具体示意可以如图10所示。

[0126] 本发明实施例提供的方法中，电子设备在处理IO请求的响应时，不通过上下文切换直接在用户态对响应的报文进行处理转换，从而可以缩短响应的处理流程，有效地提升虚拟机处理IO请求的性能。

[0127] 本申请实施例中的上述一个或多个技术方案，至少具有如下的技术效果：

[0128] 本发明实施例提供的方案中，如果有需要通过用户态OVS发送到物理网卡的IO请求，则直接在用户态将该IO请求转换为可供用户态OVS处理的格式后发送到用户态OVS。这样避免了将IO请求从用户态切换到内核态，再由内核态切换到用户态的过程，从而避免现有技术中存在用户态OVS的虚拟机处理IO请求的性能比较低的问题，所以能够有效地提升虚拟机处理IO请求的性能。

[0129] 本说明书中的各个实施例均采用递进的方式描述，各个实施例之间相同相似的部分互相参见即可，每个实施例重点说明的都是与其他实施例的不同之处。尤其，对于装置实施例而言，由于其基本相似于方法实施例，所以描述得比较简单，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实施例方案的目的。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。

[0130] 本领域普通技术人员可以意识到，结合本文中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本发明的范围。

[0131] 所属领域的技术人员可以清楚地了解到，为描述的方便和简洁，上述描述的系统、

装置和单元的具体工作过程，可以参考前述方法实施例中的对应过程，在此不再赘述。

[0132] 在本申请所提供的几个实施例中，应该理解到，所揭露的系统、装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述单元的划分，仅仅为一种逻辑功能划分，实际实现时可以有另外的划分方式，例如多个单元或组件可以结合或者可以集成到另一个系统，或一些特征可以忽略，或不执行。另一点，所显示或讨论的相互之间的耦合或直接耦合或通信连接可以是通过一些接口，装置或单元的间接耦合或通信连接，可以是电性，机械或其它的形式。

[0133] 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单元来实现本实施例方案的目的。

[0134] 另外，在本发明各个实施例中的各功能单元可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。

[0135] 所述功能如果以软件功能单元的形式实现并作为独立的产品销售或使用时，可以存储在一个计算机可读取存储介质中。基于这样的理解，本发明的技术方案本质上或者说对现有技术做出贡献的部分或者该技术方案的部分可以以软件产品的形式体现出来，该计算机软件产品存储在一个存储介质中，包括若干指令用以使得一台计算机设备(可以是个人计算机，服务器，或者网络设备等)执行本发明各个实施例所述方法的全部或部分步骤。而前述的存储介质包括：U盘、移动硬盘、只读存储器(ROM，Read-Only Memory)、随机存取存储器(RAM，Random Access Memory)、磁碟或者光盘等各种可以存储程序代码的介质。

[0136] 以上所述，仅为本发明的具体实施方式，但本发明的保护范围并不局限于此，任何熟悉本技术领域的技术人员在本发明揭露的技术范围内，可轻易想到变化或替换，都应涵盖在本发明的保护范围之内。因此，本发明的保护范围应以所述权利要求的保护范围为准。

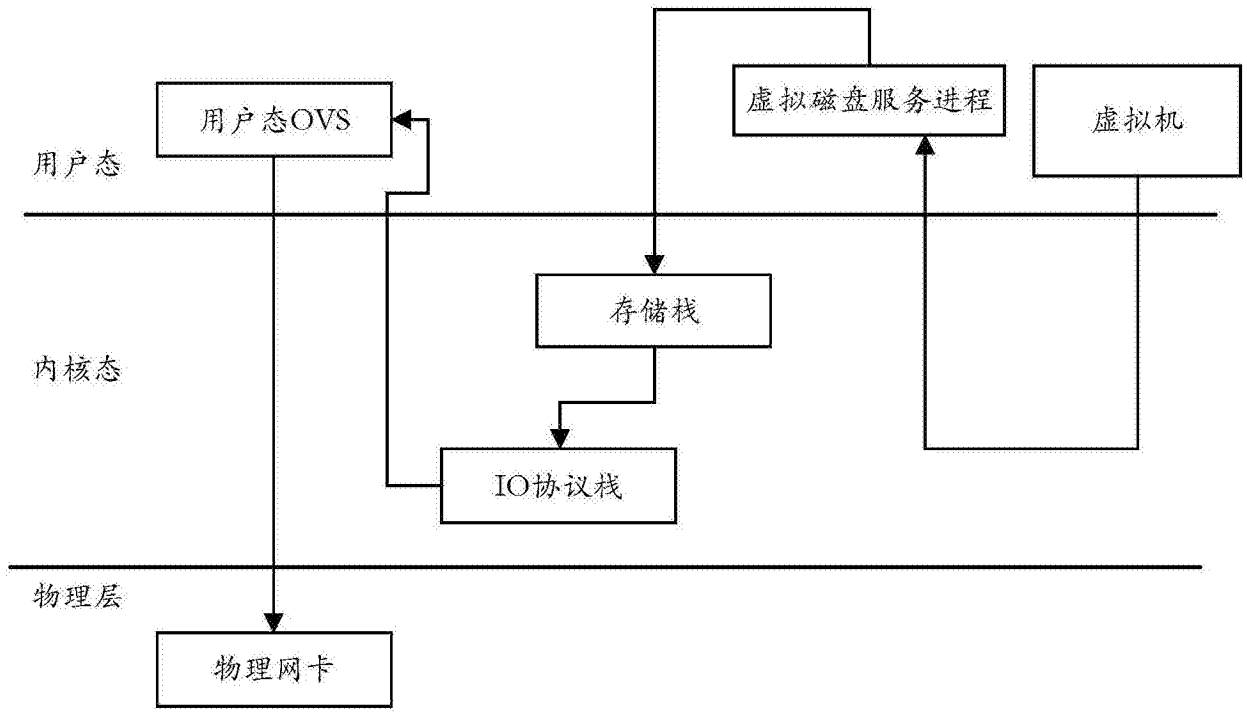

用户态OVS
虚拟磁盘服务进程
虚拟机
用户态
存储栈
内核态
IO协议栈
物理层
物理网卡

图1

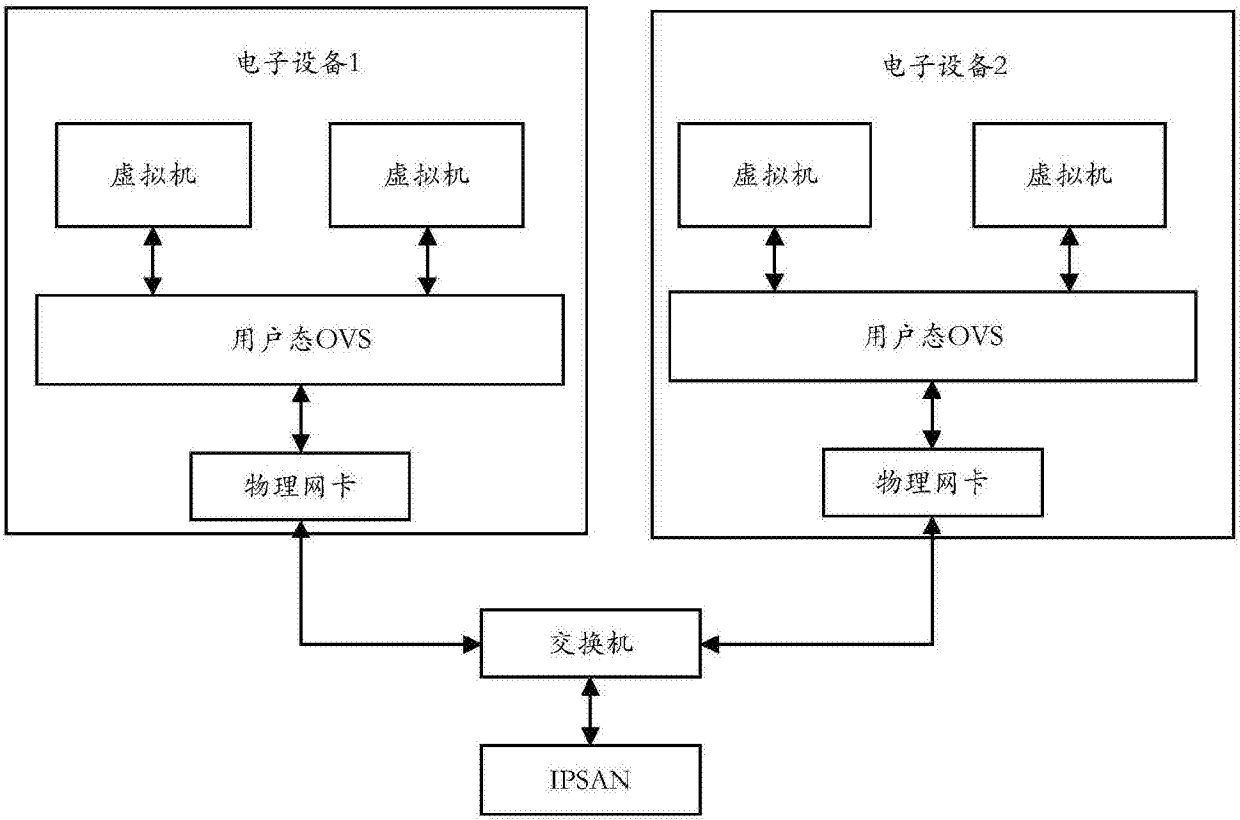

电子设备1
虚拟机
虚拟机
用户态OVS
物理网卡
电子设备2
虚拟机
虚拟机
用户态OVS
物理网卡
交换机
IPSAN

图2

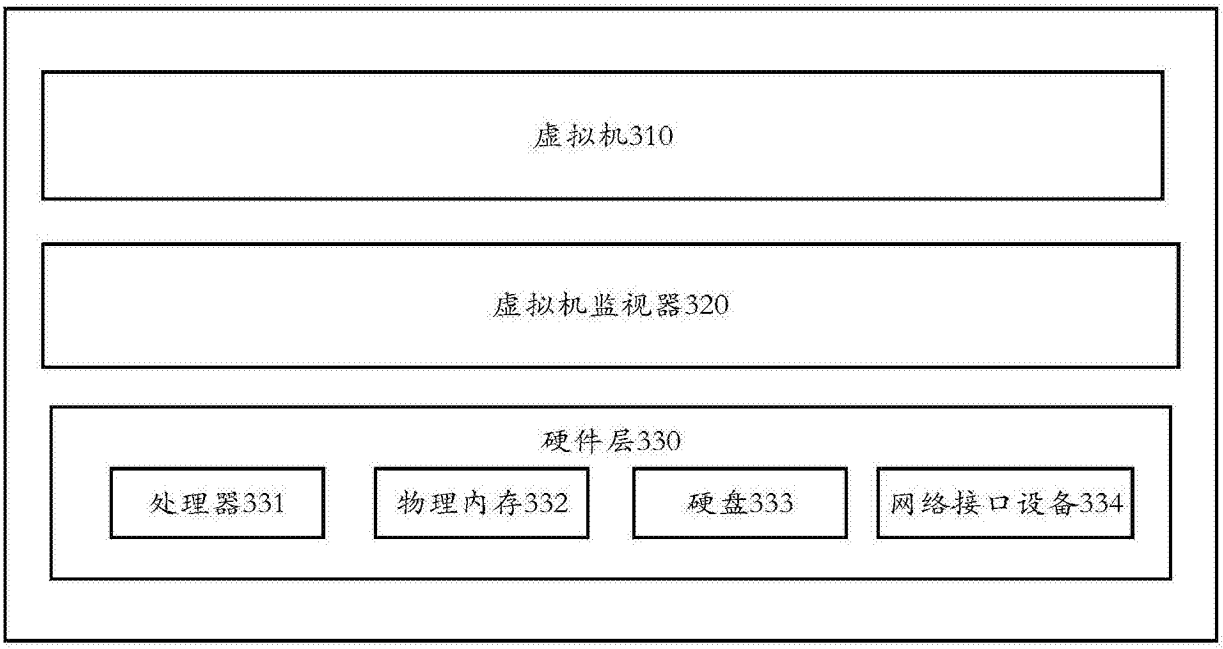

虚拟机310
虚拟机监视器320
硬件层330
处理器331
物理内存332
硬盘333
网络接口设备334

图3

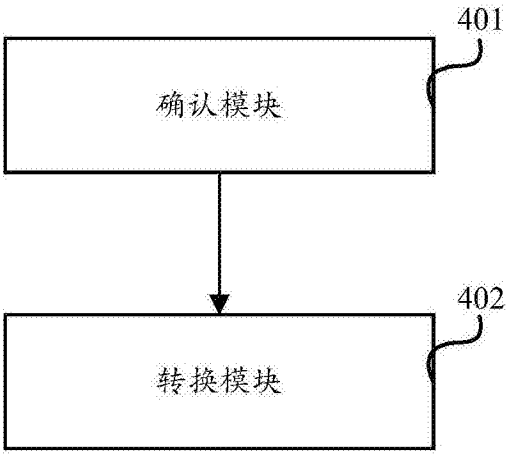

401
确认模块
402
转换模块

图4

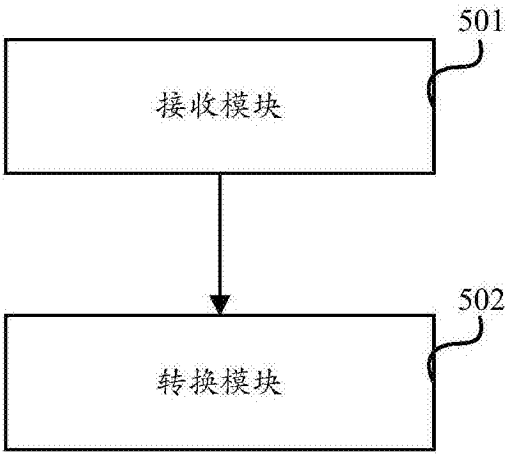

501
接收模块
502
转换模块

图5

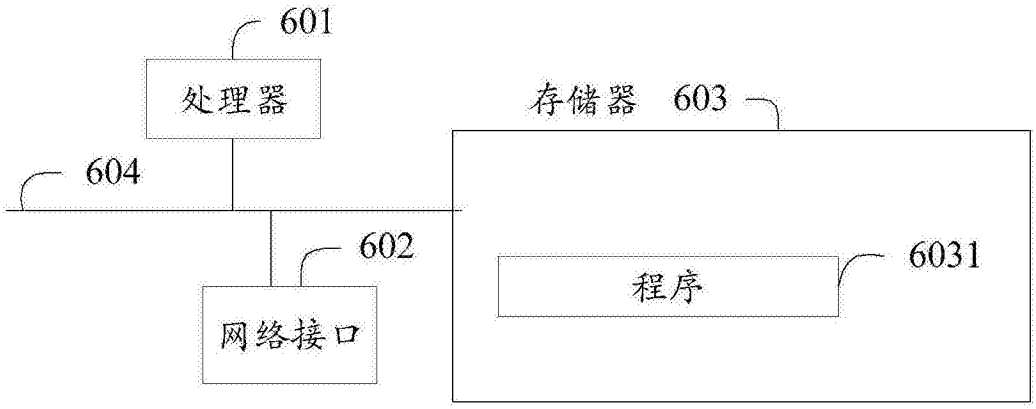

601
处理器
存储器 603
604
602
6031
程序
网络接口

图6

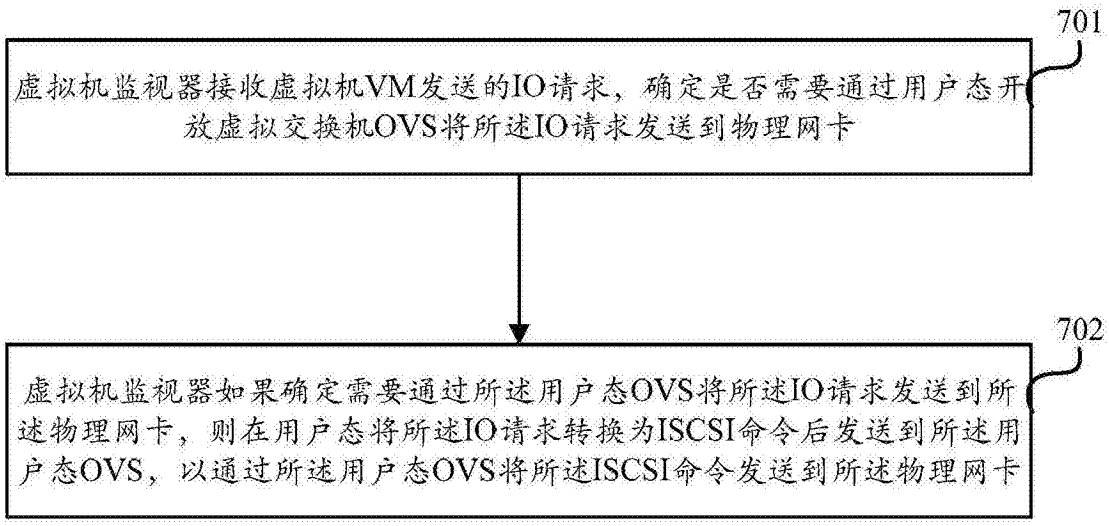

701
虚拟机监视器接收虚拟机VM发送的IO请求，确定是否需要通过用户态开放虚拟交换机OVS将所述IO请求发送到物理网卡
702
虚拟机监视器如果确定需要通过所述用户态OVS将所述IO请求发送到所述物理网卡，则在用户态将所述IO请求转换为ISCSI命令后发送到所述用户态OVS，以通过所述用户态OVS将所述ISCSI命令发送到所述物理网卡

图7

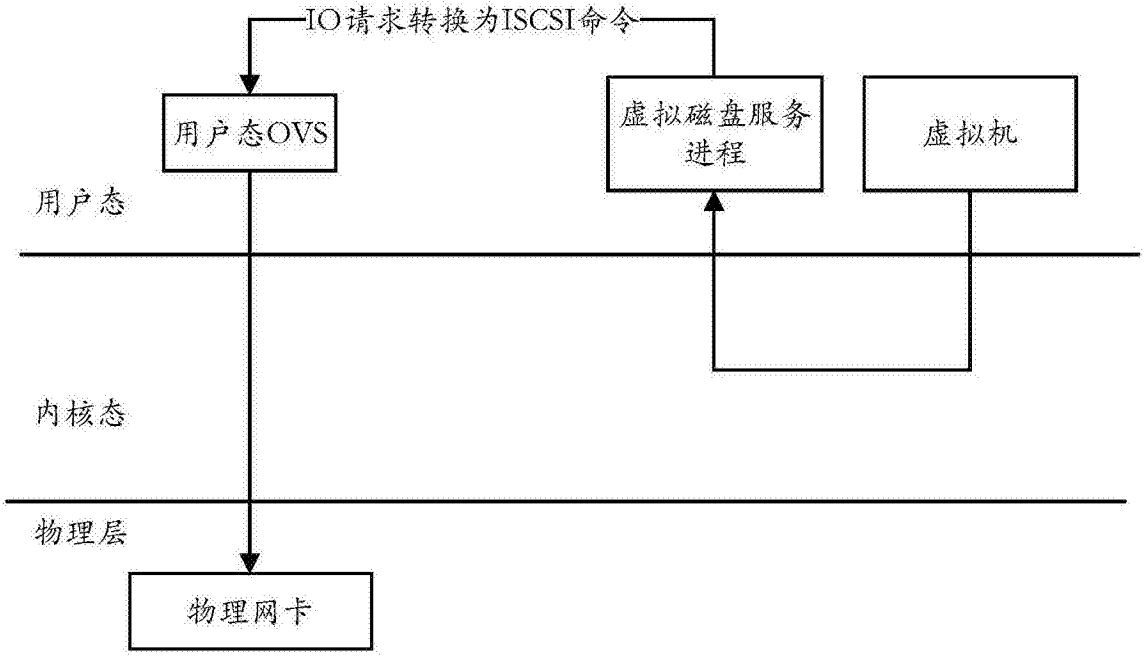

IO请求转换为ISCSI命令
用户态OVS
虚拟磁盘服务进程
虚拟机
用户态
内核态
物理层
物理网卡

图8

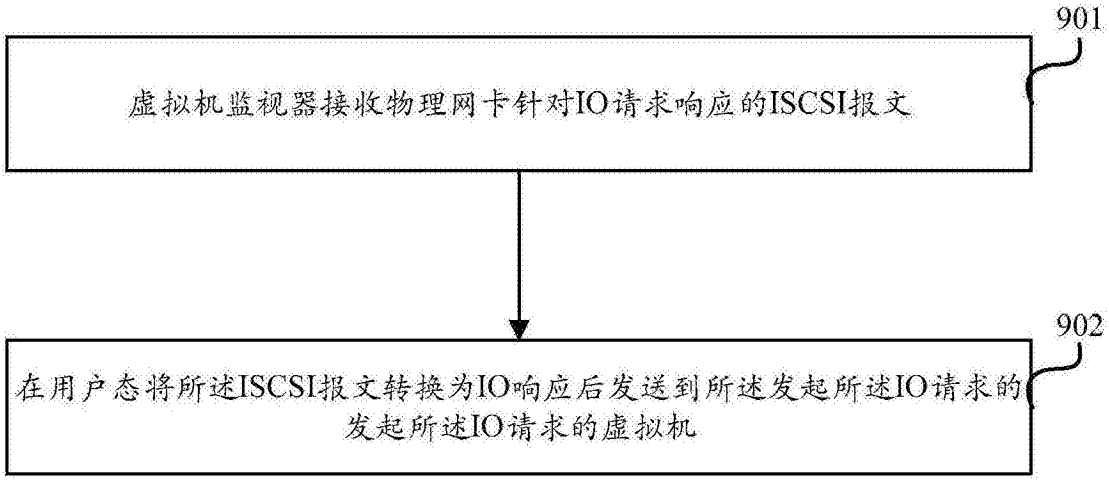

901
虚拟机监视器接收物理网卡针对IO请求响应的ISCSI报文
902
在用户态将所述ISCSI报文转换为IO响应后发送到所述发起所述IO请求的发起所述IO请求的虚拟机

图9

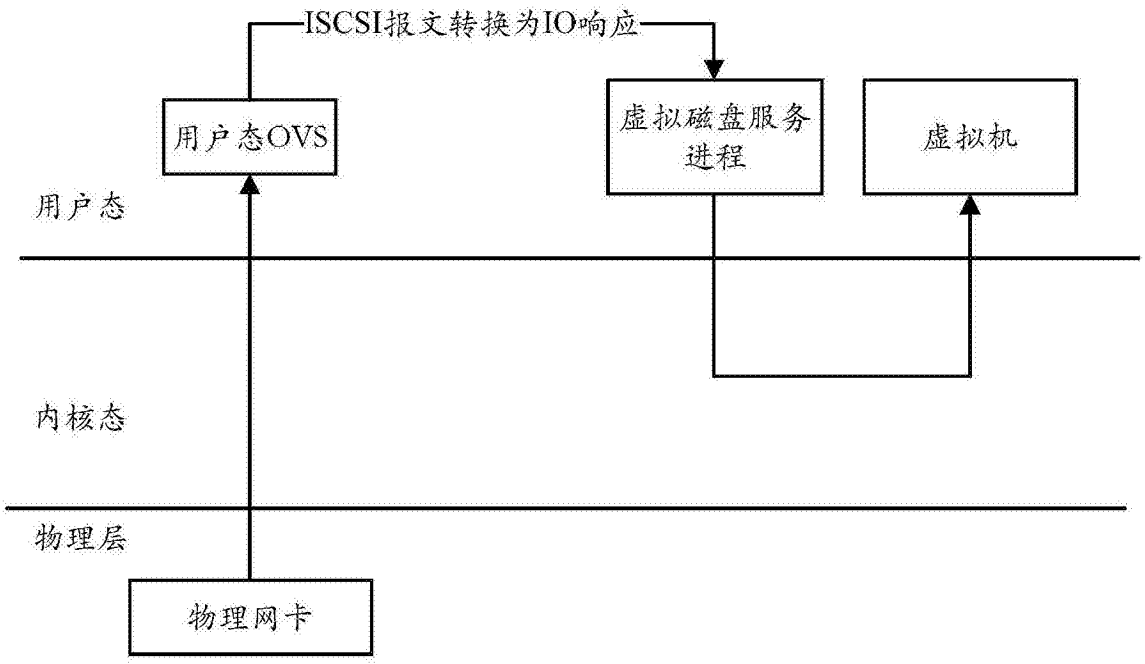


图10