



(12) 发明专利

(10) 授权公告号 CN 102804164 B

(45) 授权公告日 2015.12.02

(21) 申请号 201180013986.0

第 59 行至第 2 栏第 22 行、第 3 栏第 28 行至第 7
栏第 8 行，以及权利要求 1-2，说明书附图 4、6。

(22) 申请日 2011.03.11

US 7149187 B1, 2006.12.12, 说明书第 1 栏
第 59 行至第 2 样第 22 行、第 3 样第 28 行至第 7
栏第 8 行，以及权利要求 1-2，说明书附图 4、6。

(30) 优先权数据

12/724,507 2010.03.16 US

US 2006/0221990 A1, 2006.10.05, 说明书第
[0159]-[0164] 段，说明书附图 20.

(85) PCT 国际申请进入国家阶段日

2012.09.14

US 7283470 B1, 2007.10.16, 说明书第 7 样
第 18 行至第 8 样第 7 行，以及第 11 样第 52 行至
第 12 样第 31 行。

(86) PCT 国际申请的申请数据

PCT/US2011/028159 2011.03.11

CN 101640639 A, 2010.02.03, 全文。

(87) PCT 国际申请的公布数据

W02011/115844 EN 2011.09.22

审查员 刘洋

(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

(72) 发明人 D·班塞尔 P·R·沃克赫德

(74) 专利代理机构 上海专利商标事务所有限公

司 31100

代理人 杨洁

(51) Int. Cl.

G06F 15/16(2006.01)

G06F 9/44(2006.01)

(56) 对比文件

US 7149187 B1, 2006.12.12, 说明书第 1 样

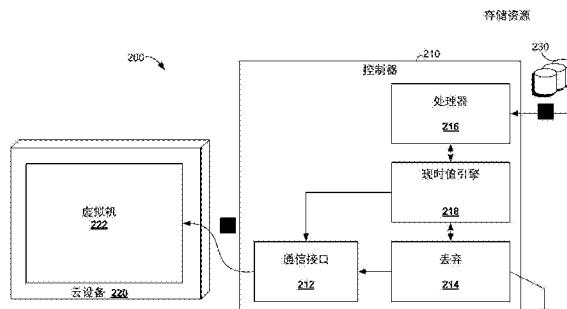
权利要求书 2 页 说明书 6 页 附图 4 页

(54) 发明名称

修整虚拟机通信流量

(57) 摘要

云计算平台具有计算机可读介质，其执行修
整虚拟机通信流量的方法。云计算平台包括虚
拟机和控制器。控制器限制与虚拟机相关联的流
量，尤其是服务该虚拟机的网络拥塞的时候，以允
许虚拟机获得所需的通信速率。控制器可基于为虚
拟机求值所得的丢弃概率来丢弃与虚拟机相关联
的通信消息。



1. 一种用于利用无队列和无状态机制管理在具有多个虚拟机的云计算平台中的虚拟机通信流量的计算机实现方法,所述方法包括 :

维护所述云计算平台中的虚拟机的平均通信速率,所述平均通信速率是由所述云计算平台为所述虚拟机求值的平均发送速率 ASR,其中 $ASR = \alpha * previous_ASR + (1 - \alpha) * current_SR$,所述“ α ”由所述云计算平台设置,“previous_ASР”是先前为所述虚拟机求值的 ASR,而“current_SR”是由所述云计算平台为所述虚拟机检测的当前发送速率;

根据所述平均通信速率和所述虚拟机的所需通信速率 DSR 来计算所述虚拟机的丢弃概率,所述丢弃概率 DP 由所述云计算平台为所述虚拟机求值,其中 $DP = \frac{\beta * (ASR - DSR)}{ASR}$,

“ β ”由所述云计算平台设置;以及

当所述虚拟机的丢弃概率大于由所述云计算平台生成的现时值时,随机丢弃所述虚拟机的通信消息,其中现时值 η 是由所述云计算平台为虚拟机随机生成的,其中 $\eta = \frac{Random_Number}{Maximum_Random_Number}$,“Random_Number”是所述云计算平台生成的,而“Maximum_Random_Number”是所述云计算平台设置的。

2. 如权利要求 1 所述的计算机实现的方法,其特征在于,还包括:限制在所述虚拟机上的传输通信消息或接收通信消息,以获得所述云计算平台中的每个虚拟机的所需通信速率。

3. 如权利要求 2 所述的计算机实现的方法,其特征在于,对传输通信消息或接收通信消息的所述限制,在所述云计算平台检测到服务所述虚拟机的网络上的拥塞时被触发。

4. 如权利要求 3 所述的计算机实现的方法,其特征在于,在无需队列的情况下执行所述虚拟机的通信速率的限制。

5. 如权利要求 1 所述的计算机实现的方法,其特征在于,所需通信速率是设置在所述虚拟机的通信策略中的所需发送速率 DSR。

6. 一种用于利用无队列和无状态机制执行丢弃虚拟机通信流量的方法,所述方法包括 :

计算虚拟机的发送速率的指数加权平均,通过云计算平台为所述虚拟机计算指数加权平均发送速率 ASR,其中 $ASR = \alpha * previous_ASR + (1 - \alpha) * current_SR$,“ α ”由所述云计算平台设置,“previous_ASР”是先前为所述虚拟机求值的 ASR,而“current_SR”是由所述云计算平台为所述虚拟机检测的当前发送速率;

标识所述虚拟机的所需发送速率 DSR;

对发送自所述虚拟机的通信消息的丢弃概率求值,所述丢弃概率 DP 由所述云计算平台为所述虚拟机求值,其中 $DP = \frac{\beta * (ASR - DSR)}{ASR}$,“ β ”由所述云计算平台设置;以及

当为所述虚拟机生成的现时值大于所述丢弃概率时,丢弃通信消息,所述现时值 η 是由所述云计算平台为所述虚拟机随机生成的,其中 $\eta = \frac{Random_Number}{Maximum_Random_Number}$,“Random_Number”是所述云计算平台生成的,而“Maximum_Random_Number”是所述云计算平

台设置的。

7. 如权利要求 6 所述的方法,其特征在于,所述通信消息是包,并且所述虚拟机通信流量被传输控制协议 TCP 所管控。

8. 如权利要求 6 所述的方法,其特征在于,所需发送速率 DSR 被设置在所述虚拟机的服务级协定中。

9. 一种云计算平台系统,被配置来利用无队列和无状态机制修整虚拟机的通信流量,所述云计算平台系统包括:

通过网络通信的虚拟机;以及

控制器,配置成用于当为所述虚拟机生成的现时值大于丢弃概率时,丢弃与所述虚拟机相关联的通信消息,其中所述控制器对所述虚拟机的平均发送速率 ASR 求值,其中 $ASR = \alpha * previous_ASR + (1 - \alpha) * current_SR$, “ α ”由所述控制器设置,“previous_ASR”是先前为所述虚拟机求值的 ASR,而“current_SR”是由所述控制器为所述虚拟机检测的当前发送速率,并且所述控制器对所述虚拟机的丢弃概率 DP 求值,其中 $DP = \frac{\beta * (ASR - DSR)}{ASR}$, “ β ”

由所述控制器设置,而所需发送速率 DSR 被设置在所述虚拟机的服务级协定中并且如果服务所述虚拟机的网络被拥塞则由所述控制器使用,并且所述控制器随机生成现时值 η ,其中 $\eta = \frac{Random_Number}{Maximum_Random_Number}$, “Random_Number”是所述控制器生成的,而“Maximum_Random_Number”是所述控制器设置的。

修整虚拟机通信流量

背景技术

[0001] 通常，云计算平台在因特网可访问的虚拟环境中托管软件应用。云计算平台允许一组织使用由第三方设计和维护的数据中心。常规的虚拟环境向小型或大型组织提供所请求的硬件资源、软件应用资源、网络资源和存储资源。虚拟环境还提供应用安全性、应用可靠性、应用伸缩性和应用可用性。

[0002] 常规的数据中心提供物理计算资源、物理存储资源和物理网络资源。数据中心中的物理资源被虚拟化并且作为一组应用编程接口被展示给所述组织。该组织不需要维护它们自己的硬件资源或软件资源、或者维护可靠和可缩放的数据中心。

[0003] 该组织可以通过虚拟环境有效地访问这些物理资源，而无需知道软件或底层物理硬件的细节。在常规的云计算平台中，硬件资源和软件资源可由互不信任的组织共享。为防止对任一组织拒绝服务，常规云平台提供维护对共享资源的公平访问的过程。大多数过程需要大量的状态信息，这增加了云计算平台中的计算和成本开销。

[0004] 例如，共享过程可包括漏桶、令牌桶和公平排队。漏桶过程将通信消息储存在队列中，并以恒定速率传输通信消息。当队列满时，丢弃通信消息。令牌桶过程将通信消息储存在队列中，并以基于与队列相关联的令牌的数量的速率传输通信。如果通信消息在桶中等待了预定时间，可丢弃通信消息。公平排队过程将通信消息储存在队列中，并且以与由其它通信会话的其它队列经历的速率相同或成比例的速率传输通信消息。这些过程提供了在云中共享资源的机制，但因维护队列及相关联的队列状态的需求而具有高 CPU 开销。

发明内容

[0005] 本发明的实施例，在一方面，涉及在云计算平台中修整(shape)虚拟机流量的云计算平台、计算机可读介质、以及计算机实现的方法。云计算平台包括连接到通信网络的控制器和虚拟机。控制器提供用于以最小开销修整通信流量的无状态机制。

[0006] 控制器监视通信网络上的流量。控制器计算连接到通信网络的虚拟机的平均通信速率。进而，由控制器获得虚拟机的所需的发送速率。控制器基于该平均通信速率来确定虚拟机的丢弃概率。网络中的通信消息基于虚拟机的丢弃概率被丢弃。

[0007] 提供本发明内容以便以简化的形式介绍将在以下详细描述中进一步描述的一些概念。该发明内容不旨在标识所要求保护的主题的关键特征或本质特征，也不旨在独立地用于帮助确定所要求保护的主题的范围。

附图说明

[0008] 图 1 是示出根据本发明的各实施例的示例性云计算平台的网络图；

[0009] 图 2 是示出了根据本发明的实施例的通信系统的框图，该通信系统具有示例性云计算平台中的示例性控制器和虚拟机；

[0010] 图 3 是示出了根据本发明的实施例的管理虚拟机通信流量的示例性方法的逻辑图；以及

[0011] 图 4 是示出了根据本发明的实施例的丢弃虚拟机通信流量的示例性方法的逻辑图。

具体实施方式

[0012] 本专利用具体细节来描述申请专利的主题以满足法定要求。然而，该描述本身并非旨在限制本专利的范围。相反，发明人设想所要求保护的主题还可结合其他当前或未来技术按照其他方式来具体化，以包括不同的步骤或类似于本专利中所描述的步骤的步骤组合。此外，尽管术语“步骤”和“框”可在此处用于指示所采用的方法的不同元素，但除非而且仅当明确描述了各个步骤的顺序时，该术语不应被解释为意味着此处公开的各个步骤之中或之间的任何特定顺序。此外，以下参考附图详细描述了各实施例，各附图整体通过引用结合于此。

[0013] 如此处所使用的，术语“组件”指的是硬件、软件、和固件的任何组合。

[0014] 云计算平台允许在各虚拟机之间共享硬件和软件资源。在一些实施例中，限制虚拟机流量以防止云计算平台中的资源缺乏。云计算平台执行限制由虚拟机利用的带宽的流量修整过程。因此，云计算平台保护共享资源，避免恶意虚拟机为了减少其它虚拟机的可访问性而企图访问过量资源。云计算平台还保护共享资源，避免非恶意虚拟机无意地使用过量网络资源。

[0015] 在一些实施例中，云计算平台利用无队列和无状态机制来实现虚拟机的带宽限制。这机制使每个虚拟机保持低开销。在一个实施例中，虚拟机的通信消息是传输控制协议(TCP)通信消息，即，包。

[0016] 带宽限制允许云计算平台的管理员基于与给定虚拟机的关系，来设置对于该给定的虚拟机可用的带宽。例如，小虚拟机可被限制在 100Mbps 而大虚拟机可被限制在 1Gbps。因此，所需的通信速率可以设置在与在云计算平台中的每个虚拟机协商的服务级协定中。

[0017] 本领域技术人员能够理解，该云计算平台可以包括硬件、软件、或硬件和软件的组合。硬件包括配置成执行存储在存储器中的指令的处理器和存储器。在一个实施例中，存储器包括存储具有用于计算机实现的方法的计算机可使用的指令的计算机程序产品的计算机可读介质。计算机可读介质包括易失性和非易失性介质、可移动和不可移动介质、以及可由数据库、交换机和各种其它网络设备读取的介质。网络交换机、路由器以及相关组件本质上是常规的，与这些组件通信的手段也是如此。作为示例而非限制，计算机可读介质包括计算机存储介质和通信介质。计算机存储介质，即机器可读介质，包括以用于存储信息的任何方法或技术来实现的介质。存储的信息的示例包括计算机可使用指令、数据结构、程序模块以及其他数据表示。计算机存储介质包括但不限于，随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、闪存或其他存储器技术、紧致盘只读存储器(CD-ROM)、数字多功能盘(DVD)、全息介质或其他光盘存储、磁带盒、磁带、磁盘存储、以及其他磁存储设备。这些存储器技术可瞬间、临时或永久地存储数据。

[0018] 在一个实施例中，云计算平台包括对客户端设备可用的云应用。客户端设备访问云计算平台来在一个或多个虚拟机上执行云应用。云应用是使用在云计算平台中可用的存储和处理资源来实现的。

[0019] 图 1 是示出根据本发明的各实施例的示例性计算系统 100 的网络图。在一实施例

中,图1所示的计算系统100仅仅是示例性的并且不旨在暗示对范围或功能的任何限制。本发明的实施例可以以大量其他配置来运行。参考图1,计算系统100包括云计算平台110、云应用120和客户端设备130。

[0020] 云计算平台110被配置为执行由客户端设备130所请求的云应用120。云计算平台110维护提供执行云应用120的虚拟机的计算设备。云计算平台还包括储存应用和系统信息的存储资源。云计算平台110通过诸如无线网络、局域网、有线网络或因特网之类的通信网络连接到客户端设备130。

[0021] 云应用120对客户端设备130可用。在云计算平台110上执行的软件实现云应用120。在一个实施例中,由云计算平台110提供的虚拟机执行云应用120。云应用120可以包括,但不限于,编辑应用、网络管理应用、财务应用、或由客户端设备130所请求或开发的任何应用。在某些实施例中,云应用120的一些功能可以在客户端设备130上执行。

[0022] 客户端设备130被用户用于与云计算平台110所提供的云应用120交互。在一些实施例中,客户端设备130必须向云计算平台110注册才能访问云应用120。具有来自云计算平台110的账户的任何客户端设备130都可以访问云应用120和云计算平台110中提供的其他资源。客户端设备130包括但不限于个人数字助理、智能电话、膝上型计算机、个人计算机、游戏系统、机顶盒、或任何其他合适的客户端计算设备。客户端设备130包括用户和系统信息存储以在客户端设备130上存储用户和系统信息。用户信息可包括搜索历史、cookie以及口令。系统信息可以包括因特网协议地址、经高速缓存的网页以及系统使用率。客户端设备130与云计算平台110通信以从云应用120接收结果。

[0023] 因此,计算系统100是用云计算平台110来配置的,该云计算平台110向客户端设备130提供云应用120。云应用120去除了在客户端设备130上更新和管理多个本地客户端应用这一负担。

[0024] 在特定实施例中,云计算平台提供修整虚拟机通信流量的控制器。控制器为每个虚拟机提供计数器。计数器可跟踪流量的平均发送速率。云计算平台可对虚拟机的指数加权平均发送速率求值。基于该平均发送速率,由控制器为每个虚拟机计算丢弃速率。基于所计算的平均发送速率,通信消息被控制器随机丢弃。在一实施例中,丢弃概率可通过管控虚拟机的服务级协定所包括的策略来设置。另选地,丢弃概率可由中央策略基于云计算平台检测的网络拥塞根据需要触发。

[0025] 图2是示出了具有示例性云计算平台中的示例性控制器210和虚拟机222的通信系统200。控制器210连接到通信网络。通信网络连接到云设备220和存储资源230。

[0026] 控制器210监视通信网络上与执行云应用的虚拟机222相关的通信流量。控制器210包括通信接口212、丢弃循环214、处理器216和现时值引擎218。在特定实施例中,控制器210被配置来通过基于为虚拟机检测的发送速率丢弃出站流量的一小部分来修整虚拟机通信。控制器能够修整流量而无需维护虚拟机或通信会话的队列。控制器210测量每个虚拟机的发送速率,并选择一个或多个通信消息用于以所确定的概率丢弃。

[0027] 通信接口212从执行虚拟机222的云设备220接收通信消息。通信接口还向执行虚拟机222的云设备220发送通信消息。通信接口212可包括,但不限于,以太网或异步传输模式(ATM)接口。控制器210可通过通信接口212向云设备220发送通信消息。

[0028] 控制器210使用丢弃循环214来丢弃通信消息。可通过向接地的接口发送消息来

丢弃通信消息。当控制器 210 确定虚拟机 222 具有超过指定阈值的丢弃概率时,通信消息被发送到丢弃循环。

[0029] 处理器 216 为控制器 210 执行计算。计算被控制器使用,来确定是否丢弃虚拟机 222 的通信消息。计算可包括对平均发送速率、丢弃概率和现时值求值。基于这些计算,处理器 216 可设置标志来丢弃虚拟机 222 的一个或多个通信消息。

[0030] 处理器 216 获取云设备 220 执行的虚拟机 222 的当前发送速率。在另一实施例中,在每次从虚拟机 222 接收到通信消息时,处理器 216 可通过将通信消息的大小除以自控制器 210 接收到来自虚拟机 222 的最后一次通信消息的间隔,以计算当前发送速率。当虚拟机 222 的服务级协定中没有设置丢弃概率时,当前发送速率被用来确定虚拟机 222 的丢弃概率。服务级协定被储存在存储资源 230 中,其由处理器 216 访问来确定丢弃概率是否被设置。

[0031] 进而,处理器 216 计算平均发送速率。在一个实施例中,平均发送速率是指数平均的发送速率(ASR)。处理器可求值 $ASR = \alpha * previous_ASR + (1 - \alpha) * current_SR$, 其中“ α ”由云计算平台设置,“previous_ASR (先前 ASR)”是先前为虚拟机求值的 ASR,而“current_SR (当前 SR)”是为虚拟机检测的当前发送速率。“ α ”是有理数。指数平均使用比其它平均函数更少的存储器。在特定实施例中,可使用其它平均函数,例如,中间发送速率。

[0032] 访问虚拟机 220 的所需发送速率以求值丢弃概率。可以在服务级协定中设置所需的发送速率。在其它实施例中,管理员可基于网络拥塞实时设置所需的发送速率。在又一实施例中,所需的发送速率可基于网络上检测的拥塞自动设置。所需的发送速率可以随着通信网络上的拥塞的增加而自动降低。

[0033] 处理器 216 计算丢弃概率(DP)。处理器求值 $DP = \frac{\beta * (ASR - DSR)}{ASR}$, “ β ”由云计算平台设置。 β 可由管理员设置来减少或增加丢弃的通信消息的数量。 β 的范围从 0 到 1。

[0034] 现时值引擎 218 是随机数生成器。它为处理器 216 生成随机数。现时值引擎 218 还可向处理器提供最大随机数。进而,处理器 216 可生成被用作阈值的现时值(η)。处理器 216 求值 $\eta = \frac{Random_Number}{Maximum_Random_Number}$, “Random Number(随机数)”是现时值引擎 218 生成的随机数,而“Maximum Random Number (最大随机数)”是现时值引擎 218 提供的最大随机数。“Random Number”和“Maximum Random Number”两者都是有理数。

[0035] 处理器 216 将现时值和丢弃概率进行对比。当丢弃概率大于现时值时,处理器 216 可丢弃通信消息。

[0036] 在一些实施例中,虚拟机 222 的丢弃概率可由服务级协定基于网络拥塞来设置。例如,当控制器 210 没有检测到网络拥塞时,服务级协定可指示应该将平均发送速率设置到所需发送速率。因此,丢弃概率是零。虚拟机 222 被允许无限制资源利用。但是当控制器 210 检测到网络拥塞时,平均发送速率被设置到特定值,例如 50Mbps、20Mbps,这取决于拥塞级别和访问云计算平台中的资源的虚拟机 222 的数量。

[0037] 存储资源 230 储存每个虚拟机 222 的云应用和服务级协定。在一些实施例中,虚拟机 222 可具有多个服务级协定。每个服务级协定可对应于虚拟机执行的不同云应用。服

务级协定可设置虚拟机 222 的所需发送速率、云应用的所需发送速率，和云设备 220 的所需发送速率。在一些实施例中，服务级协定也可存储虚拟机 222 的丢弃概率、云应用的丢弃概率，和云设备 220 的丢弃概率。响应于对所需的发送速率或丢弃概率的请求，存储资源 230 可将服务级协定传输到控制器 210。

[0038] 在一实施例中，云计算平台在运行于云设备上的虚拟机上执行云应用。云计算平台修整虚拟机的通信流量。云计算平台监视虚拟机的通信速率并基于通信速率丢弃包。通信速率可包括传输速率和接收速率。

[0039] 图 3 是示出了管理虚拟机通信流量的示例性方法的逻辑图。该方法在步骤 310 初始化。在步骤 320，云计算平台维护云计算平台中的虚拟机的平均通信速率。云计算平台可使用虚拟机的指数平均或中间平均。在步骤 330，云计算平台根据平均通信速率以及虚拟机的所需通信速率来计算虚拟机的丢弃概率。所需通信速率可从储存了服务级协定的存储资源获得，该服务级协定指定虚拟机的所需发送速率。另选地，所需发送速率可被云计算平台的管理员设置。

[0040] 云计算平台限制与虚拟机相关联的传输通信消息或接收通信消息，以获得云计算平台中的虚拟机的所需通信速率。在一个实施例中，传输通信消息或接收通信消息的限制，在云计算平台检测到服务虚拟机的网络上的拥塞时被触发。进而，在步骤 340，云计算平台在虚拟机的丢弃概率大于云计算平台生成的现时值时随机丢弃虚拟机的通信消息。该方法在步骤 350 终止。

[0041] 在一些实施例中，云计算平台基于在云设备上执行的虚拟机的发送速率来修整通信流量。云计算平台可丢弃通信消息以降低整体的网络拥塞。在决定是否丢弃通信消息时，云计算平台试图维护虚拟机的所同意的服务级别。

[0042] 图 4 是示出了丢弃虚拟机通信流量的示例性方法的逻辑图。虚拟机通信流量可通过传输控制协议(TCP)管控。该方法在步骤 410 初始化。在步骤 420，云计算平台计算虚拟机的发送速率的指数加权平均。通过云计算平台为虚拟机计算指数加权平均发送速率(ASR)，其中 $ASR = \alpha * previous_ASR + (1 - \alpha) * current_SR$ ，“ α ”由云计算平台设置，“ $previous_ASR$ ”是先前为虚拟机求值的 ASR，而“ $current_SR$ ”是为虚拟机检测的当前发送速率。

[0043] 在步骤 430，云计算平台标识虚拟机所需的发送速率。所需发送速率(DSR)可在虚拟机的服务级协定中设置。另选地，所需发送速率可被云计算平台的管理员设置。

[0044] 在步骤 440，云计算平台还为虚拟机发送的通信消息的丢弃概率进行求值。丢弃概率(DP)由云计算平台为虚拟机求值，其中 $DP = \frac{\beta * (ASR - DSR)}{ASR}$ ，“ β ”由云计算平台设置。

[0045] 在步骤 450，在云计算平台所生成的虚拟机的现时值时大于丢弃概率时，云计算平台丢弃通信消息。现时值(η)是由云计算平台为虚拟机随机生成的，其中

$$\eta = \frac{Random_Number}{Maximum_Random_Number} \text{, "Random Number" is 由云计算平台生成的, 而 "Maximum}$$

Random Number" 是云计算平台设置的。在一实施例中，通信消息可以是包。该方法在步骤 460 终止。

[0046] 总之，云计算平台修整与执行在云设备上的虚拟机相关联的通信流量。云计算平台监视通信速率并试图为云计算平台中的每个虚拟机维护合适的通信速率。与虚拟机相关

联的服务级协定由云计算平台处理,以确保为虚拟机维护合适的服务级别。

[0047] 本发明的各实施例的上述描述是说明性的,且配置和实现中的各修改都处于当前描述的范围内。例如,尽管本发明的实施例是参照图 1 — 4 来概括性地描述的,但是那些描述是示例性的。尽管用对结构特征或方法动作专用的语言描述了本主题,但可以理解,所附权利要求书中定义的主题不必限于上述具体特征或动作。更确切而言,上述具体特征和动作是作为实现权利要求的示例形式公开的。因此,本发明的实施例的范围只在仅由所附权利要求书来限定。

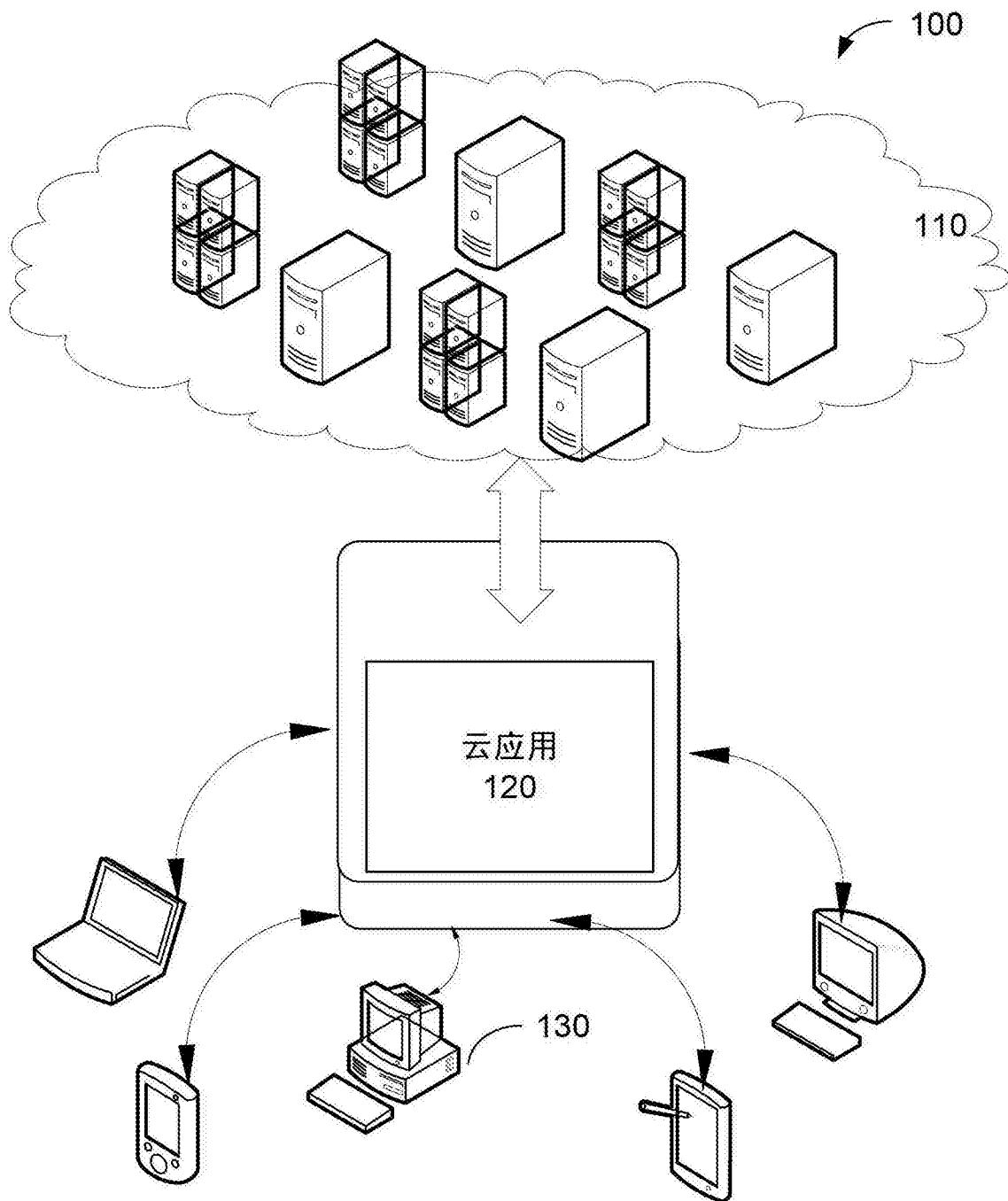


图 1

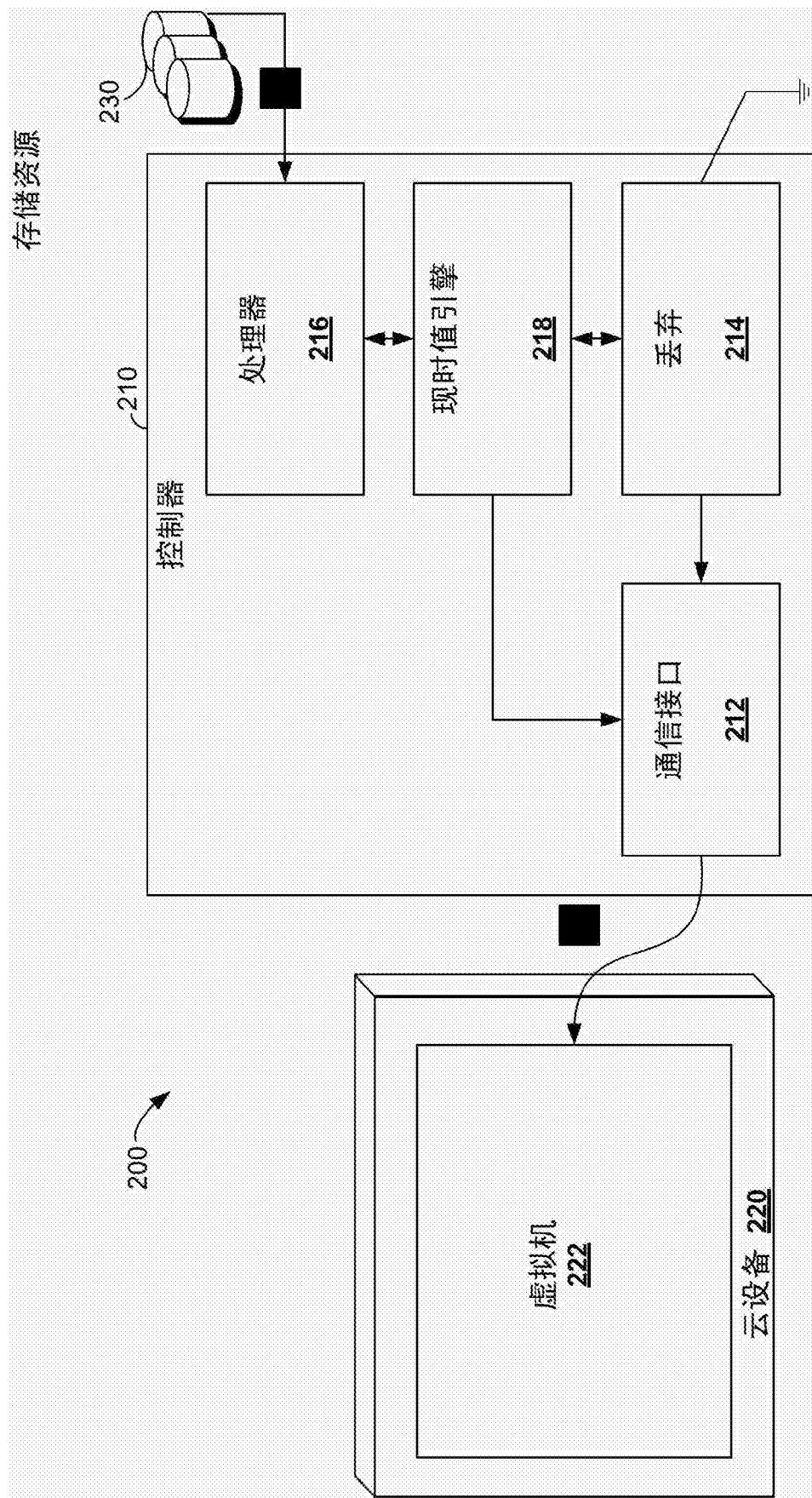


图 2

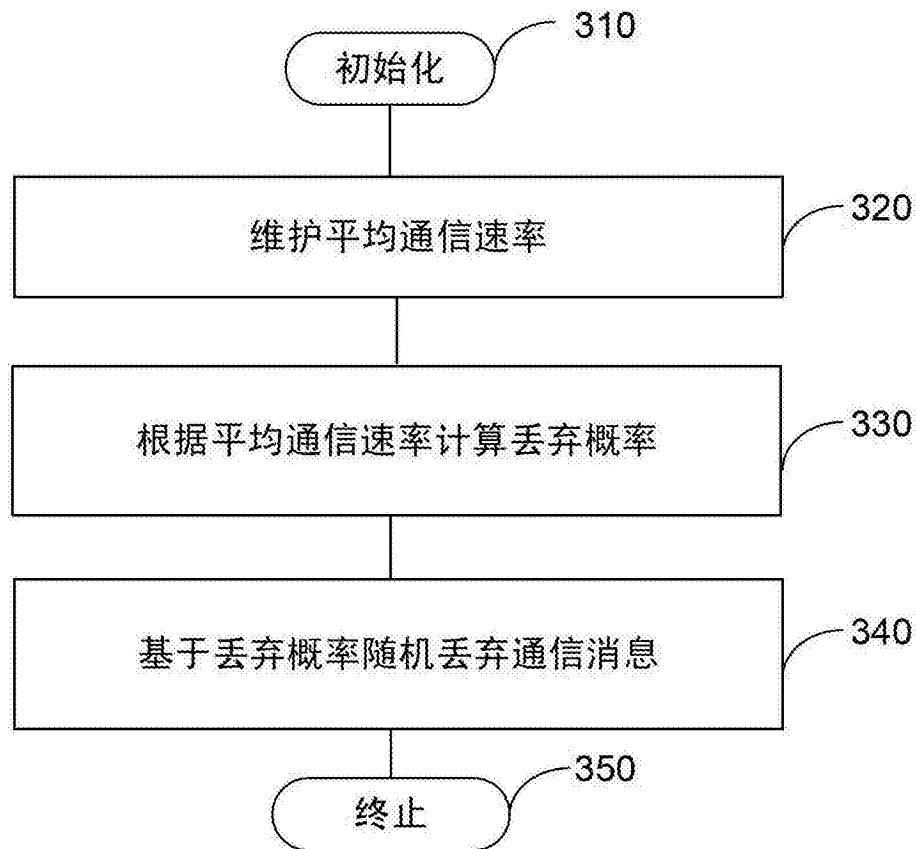


图 3

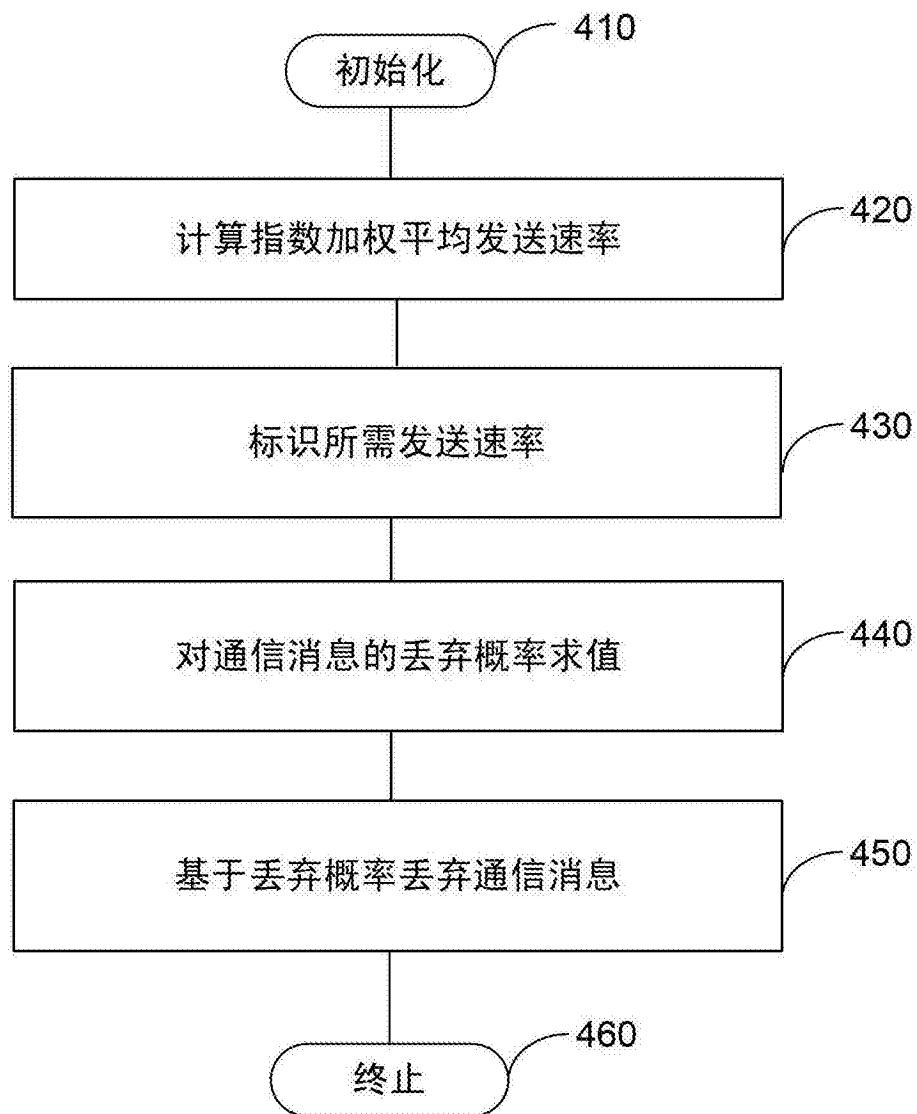


图 4