

(19) 中华人民共和国国家知识产权局



(12) 发明专利说明书

(10) 申请公布号 CN 102770826 A

(43) 申请公布日 2012.11.07

(21) 申请号 CN201180011019.0

(22) 申请日 2011.02.25

(71) 申请人 微软公司

地址 美国华盛顿州

(72) 发明人 A·坎萨尔 J·D·伯格 A·A·巴塔查里亚

(74) 专利代理机构 上海专利商标事务所有限公司

代理人 杨洁

(51) Int. CI

G06F1/28

G06F9/44

权利要求说明书 说明书 幅图

(54) 发明名称

虚拟机功耗测量和管理

(57) 摘要

虚拟机功率计量系统和方法的各实施例测量各个虚拟机的功耗。物理主机服务器的功率表测量被转换成测量驻留在主机服务器上的每个单独虚拟机的功耗的各个虚拟机功率表。通过使用主机服务器的总功耗以及虚拟机的资源利用率生成功率模型，来计算虚拟机功耗。使用该功率模

型来计算最优功率模型系数。使用两个实施例之一来计算虚拟机使用的能量。该系统和方法的各实施例还可用于获得特定活动（诸如，服务、请求或搜索查询）的功耗。另外，虚拟机功率计量可用于虚拟机功率设限，以允许虚拟化环境中的功率过度预订。

法律状态

法律状态公告日

法律状态信息

法律状态

权利要求说明书

1.一种在计算设备上实现的方法，用于计算跨驻留于物理主机服务器上的至少一个虚拟机而发生的活动的功耗，包括：

将活动时间段定义为在起始事件与结束事件之间流逝的时间量；

在所述活动时间段期间跟踪用于所述活动的每个虚拟机上的每个资源；

计算在所述活动上工作的每个虚拟机在所述活动时间段期间的功耗，以获得功耗测量结果；

将在所述活动上工作的每个虚拟机在所述活动时间段期间的功耗测量结果相加，以获得总的活动功耗；以及

输出表示所述活动消耗的总功耗的所述总的活动功耗。

2.如权利要求 1 所述的方法，其特征在于，还包括：

计算将虚拟机的总的资源利用率与所述主机服务器的总功耗相关的功率模型；以及

使用所述功率模型以及最优功率模型系数来计算所述虚拟机的功耗。

3.如权利要求 1 所述的方法，其特征在于，还包括：

计算在所述活动时间段期间被使用但未作为资源被跟踪的设备的总功耗；

以及

确定未被跟踪设备的所述总功耗的一部分，所述部分在所述活动时间段期间被用于所述活动中。

4.如权利要求 2 所述的方法，其特征在于，还包括将在所述活动上工作

的每个虚拟机的功耗测量结果、以及在所述活动时间段期间所使用的所述未被跟踪设备的功耗的所述部分相加，以获得总的活动功耗。

5.如权利要求 4 所述的方法，其特征在于，还包括：

当在所述活动在所述虚拟机上被初始化时，生成起始事件；以及

当所述活动被终止并被呈现给用户时，确定并标记结束事件。

6.如权利要求 2 所述的方法，其特征在于，还包括：

确定每个虚拟机所使用的每个资源的百分比；

将每个资源百分比乘以其对应的最优功率模型系数，以获得每个虚拟机的中间资源结果；以及

将每个虚拟机的所述中间资源结果的每一个相加，以获得每个虚拟机的单个功耗。

7.如权利要求 2 所述的计算机实现的方法，其特征在于，还包括：

确定为每个虚拟机工作的中央处理单元（CPU）的百分比；以及

使用所述最优功率模型系数来计算每个虚拟机的每个资源的虚拟机专用的系数。

8.一种在计算设备上实现的方法，用于单独地控制驻留于物理主机服务器上的多个虚拟机的功耗，包括：

使用所述计算设备执行以下动作：

监视所述主机服务器上的每个虚拟机的功耗；

测量所述主机服务器上的每个虚拟机所提取的峰值功耗；

确定所述多个虚拟机中的任一个是否已经超出其功率预算；

若是，则确定所述主机服务器的功率预算是否已被超出；以及

若是，则将所述虚拟机中至少一个的功耗降至所述主机服务器的总功耗。

9.如权利要求 10 所述的方法，其特征在于，还包括基于每个虚拟机的已测量的峰值功耗来定位所述主机服务器上的虚拟机。

10.如权利要求 10 所述的方法，其特征在于，还包括：

生成功率模型，所述功率模型将所述多个虚拟机中每一个的总的资源利用率与所述主机服务器的总功耗相关；

使用所述功率模型来获得最优功率模型系数；以及

使用所述功率模型以及所述最优功率模型系数来计算所述多个虚拟机中每一个的功耗。

11.如权利要求10所述的方法，其特征在于，还包括降低超出其功率预算的每个虚拟机的功耗，以便不超出所述主机服务器的功率预算。

12.如权利要求10所述的方法，其特征在于，还包括降低在所述主机服

服务器上驻留的每个较低优先级的虚拟机的功耗，以便不超出所述主机服务器的功率预算，其中与已经超出其功率预算的虚拟机相比，所述较低优先级的虚拟机具有较低的优先级。

13.如权利要求 10 所述的方法，其特征在于，还包括关闭已经超出其功率预算的所述虚拟机中的一个或多个，以便不超出所述主机服务器的功率预算。

14.如权利要求 10 所述的方法，其特征在于，还包括将一个或多个虚拟机迁移到另一个服务器，以便不超出所述主机服务器的功率预算。

说明书

<p>背景技术

数据中心、云计算基础设施、计算集群、企业计算工具和高性能计算工具

在将工作负载运行在共享物理服务器上的同时，越来越多地使用虚拟机来隔离工作负载。虚拟机是对计算设备的仿真，其严格意义上是软件但像物理计算设备那样运作。这意味着一个或多个虚拟机可驻留在单个物理主机服务器上。然而，在运行在虚拟机上的客户端和全部软件看来，每个虚拟机是独立的服务器。

虚拟化（或使用虚拟机）允许将多个应用安全地置于单个物理主机服务器

上，因为每个应用在其“虚拟”计算机内部运行，并且不影响其他应用。这允许服务器资源被更好地利用，并且降低所需的服务器的数量以及托管它们（诸如电力和冷却）的操作成本。因此，虚拟化是管理计算资源的高效技术，并正在数据中心和大规模计算工具中获得广泛的接受。

数据中心和其他计算工具经常监视服务器的功率使用，以便正确地管理功

率容量和分布。功率测量结果数据还用于制定若干手动和自动的功率管理决策。事实上，功率计量在数据中心的广泛，以至于功率分配单元（PDU）、电源板、服务器电源都供应测量功率所需的功率计量仪器，并且在新服务器中，甚至是服务器主板已经包含了测量功率所需的功率计量仪器。测量物理服务器中的功耗的另一种方法是将功率计量仪器连接到服务器。

物理服务器的功耗可通过计量被供应给服务器的电量来测量。然而，虚拟

机是在软件中被创建的，并且不存在将功率表连接到虚拟机的简单方法。这使得测量虚拟机的功耗非常困难。用于物理服务器的技术仅能报告服务器的总的功率提取。但是，它们无法用于测量驻留在单个物理主机服务

器上的多个虚拟机中每一个的单独功耗。事实上，单纯地使用硬件方案无法测量虚拟机的功耗。

测量虚拟机的功耗的能力是重要的。用户或服务器管理器不能测量每台虚拟机所使用的功率。出于若干理由，测量该功率是有用的。例如，如果虚拟机正在使用过量的功率，则它可能必须被关闭或被迁移到具有额外能力的不同的电源电路。另外，软件开发者可测量出它们的软件在被运行于物理服务器上时消耗了多少能量，但无法在软件运行在虚拟机上时测量该能量。无法测量功率使用削弱了开发者优化软件以得到更低功耗的能力。测量虚拟机功耗的另一个用途是基于功率使用向用户返还费用。诸如在云计算中使用的共享物理服务器允许多个用户在共享的物理基础设施上运行虚拟机。物理服务器的所有者可能希望基于功率使用向虚拟机用户收费。

而虚拟机功率计量的另一个用途、以及下面更详细地描述的一个用途产生于虚拟机功率设限。功率设限允许电力基础设施的过度预订，并带来功率容量的更高效的利用率。然而，当前的功率设限技术被设计用于物理服务器，而无法在被应用到虚拟机时很好地工作。例如，对物理服务器的功率进行设限导致了在该服务器上托管的全部虚拟机的设限。这破坏了虚拟机所提供的工作负载的隔离，并可使得使用虚拟化的益处受挫。

因此，大规模计算工具中功率管理的主要组件之一是对功率的可视性。然而，越来越多地在数据中心的虚拟化机器正缺少这种能力。使用虚拟机允许安全地隔离多个处于相同位置的工作负载，使得多个工作负载能在更少的服务器上被合并，并且带来被改善的资源利用率。然而，缺少对虚拟机功率使用的可见性带走了在无虚拟化的情况下可用的服务器功率计量的许多好处。

发明内容

提供本发明内容以便以简化形式介绍将在以下详细描述中进一步描述的

一些概念。本发明内容并不旨在标识所要求保护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主题的范围。

虚拟机功率计量系统和方法的各实施例测量各个虚拟机的功耗。功耗计量

通常用在物理服务器上，以制定若干功率管理和操作决策。虚拟机功率计量系统和方法的各实施例允许在虚拟机上也能作出这样的决策。

一般的想法是测量整个物理主机服务器的功耗的功率表测量被转换成各

个虚拟机功率表，该虚拟机功率表测量驻留在主机服务器上的每个单独虚拟机的功耗。虚拟机功率计量系统和方法的各实施例将与任意数量的虚拟机一同工作，并且跨多个物理主机服务器工作。

了解单个虚拟机的功耗对于在数据中心或任一虚拟化服务器中制定的若

干功率管理决策来说是有用的。例如，虚拟机功率计量可用于虚拟机功率设限，以允许虚拟化环境中的功率过度预订。可以执行该虚拟机管理设限，而不破坏隔离属性。各个虚拟机功耗测量结果还可用于跟踪跨越多个虚拟机的活动（诸如，搜索查询）的功耗。

虚拟机功率计量系统和方法的各实施例不需要供应功率模型的硬件。在进

行中得知所需要的任何信息。该系统和方法的各实施例自动地学习将硬件（或资源）使用（诸如，中央处理单元（CPU）使用、磁盘使用、存储器使用、网络使用等）与功耗相关的功率模型。随后跟踪每个虚拟机的硬件使用以推断该虚拟机的功耗。

该系统和方法的各实施例在两大步骤中计算该功耗。第一，使用物理主机服务器的总功耗以及特定虚拟机的资源利用率来生成并训练功率模型。资源利用率是对虚拟机正在使用多少资源的判定。将这些资源利用率相加以发现虚拟机的总的资源利用率。生成将总的资源利用率与总功耗相关的功率模型。

该系统和方法的各实施例还在该步骤期间确定最优功率模型系数。这通过确定由虚拟机正在使用的资源的数量 N 来实现。生成并求解具有 N 个未知数的一系列 N 个等式，以获得最优功率模型系数。

该系统和方法的各实施例还包括功率模型学习技术，其中所学到的模型对于在每个虚拟机内部托管的工作负载是唯一的。这说明了在功率使用中工作负载之间可被发现的差距，即使工作负载具有相同或看起来相似的资源使用。

第二，该系统和方法的各实施例计算虚拟机所使用的能量。可使用若干实施例来计算该能量。一个实施例使用虚拟机所使用的资源的百分比，并随后将这些百分比相加以获得该虚拟机的功耗。在其他实施例中，按照单位而不是百分比来测量所使用的资源。作为示例，仅举数例，所测量的资源可包括 CPU、存储器、磁盘、网卡和风扇。在这些实施例中，可通过考虑主机服务器的 CPU 频率来扩充功率模型。

该系统和方法的各实施例还可用于获得特定活动的功耗。活动可以是服务、请求（诸如，web 搜索查询）、电子邮件递送、网页检索、数据库操作、或文件更新。即使该活动是跨多个虚拟机并且跨多个主机服务器，也可以计算该活动的功耗。确定活动时间段，该活动时间段是在计算基础设施上该活动为活跃的时间段。该系统和方法的各实施例随后通过将各个活动时间段期间在该活动上工作的每个虚

拟机的功耗相加以获得总的活动功耗，来计算功耗。

应当注意，替换实施例也是可能的，并且此处所讨论的步骤和元素可取决于特定实施例而改变、添加或消除。这些替换实施例包括可使用的替换步骤和替换元素，以及可做出的结构上的改变，而不脱离本发明的范围。

附图说明

现在参考附图，在全部附图中，相同的参考标号表示相应的部分：

图 1 是示出在主机服务器计算设备上实现的虚拟机功率计量系统和方法的各实施例的综合概述的框图。

图 2 是示出在图 1 所示的虚拟机功率计量系统和方法的各实施例中使用的模块和数据的框图。

图 3 是示出图 1 和 2 所示的虚拟机功率计量系统的各实施例的总体操作的流程图。

图 4 是示出图 2 所示的功率模型训练模块的各实施例的操作细节的流程图。

图 5 是示出图 2 所示的能量计算模块的第一实施例的操作细节的流程图。

图 6 是示出图 2 所示的能量计算模块的第二实施例的操作细节的流程图。

图 7 是示出图 2 所示的功率设限模块的各实施例的操作细节的流程图。

图 8 示出了其中可实现图 1-7 所示的虚拟机功率计量系统和方法的各实施
例的合适的计算系统环境的一个示例。

具体实施方式

在以下对虚拟机功率计量系统和方法的各实施例的描述中，对附图进行了

参考，附图形成了该描述的一部分，并且其中作为说明示出了可实施虚拟机
功率计量系统和方法的各实施例的一个具体示例。可以理解，可以利用
其他实施 例，并且可以作出结构上的改变而不背离所要求保护的主题的
范围。

I. <u>功率计量概念</u>

由于没有硬件功率测量设备可被连接到单个虚拟机，因此，虚拟机功率计
量系统和方法从资源使用来推断能量。原则上，虚拟机
功率计量系统和方法跟 踪软件中每台虚拟机所使用的资源，并且随后通
过利用各个资源的功率模型将 资源使用转换成能量。

对于该方法，系统和方法跟踪对全部硬件资源的使用，诸如 CPU 核、磁
盘阵列、存储体（memory bank）、网卡和图形卡。跟踪资源使用包括确定
资源处于什么功率状态，以及它为什么样的虚拟机工作。如果系统和方法根
据 哪一个虚拟机正在使用资源来正确地标记每个资源，并还在该期间
（通常是几 微秒或毫秒）测量功率提取，则该系统和方法可计算每台虚拟
机的活动能量使 用。活动能量可被定义为资源在代表某个虚拟机或系统
工作时所消耗的能量。

除活动能量以外，系统还使用某些空闲能量和共享能量。作为示例，对于

CPU，空闲能量可包括对于活动状态之间的小的持续时间，在睡眠状态中消

耗的能量。共享能量包括当所执行的工作对多个虚拟机有益时资源所使用的能量，诸如在来自多个虚拟机的输入/输出操作之前在磁盘加速自旋上消耗的能量。无论任一虚拟机是否在运行，空闲能量的某些都被消耗，而空闲能量的一部分受到虚拟机是否存在的影响。

例如，虚拟机的空闲时间段之间的持续时间可能影响在空闲时间期间更深的睡眠状态是否被处理器采用。空闲和共享能量可被分开地报告，或者依赖于将如何使用计量信息，与虚拟机的活动能量使用相等地或成比例地在虚拟机之间被划分。对于下面详细地讨论的功率设限应用，分开地报告最有效。

因此，跟踪虚拟机能量归结于两个挑战。第一，微秒或毫秒粒度的功率测量；应当每隔几微秒确定每个资源的瞬时功率，因为使用资源的虚拟机可能（依赖于用于上下文切换的时间段）每隔几微秒就改变。这可能导致组件使用和功率状态改变。

第二，标记每个单独的虚拟机的资源使用。必须确定哪个虚拟机负责使用每个资源。理想地，这个过程不需要操纵在虚拟机内部运行的应用源代码，因为平台开发者可能无法对其进行访问，并且在实践中，要求应用开发者操纵其源代码不可能衡量。

为了解决第一个挑战，系统和方法的各实施例利用将资源的软件可观察状态与其功率使用相关的功率模型。如果能以微秒粒度来观察状态改变，则能推断在该精细的时间刻度的功率使用。观察功率状态是不同寻常的，并且必须作出若干近似。例如，在方法和系统的某些实施例中可能无法完全知道资源状态，诸如处理器内子组件的时钟选通、或磁盘阵列内机械运动的确切特性。另外，可能不存在对平台内某些硬件资源的功率状态（诸如，硬件控制的设备功率状态、或图形处

理活动) 的可视性。

在观察资源状态之后, 下一个步骤是使用其功率模型来确定该状态中的功

率使用。然而, 功率模型不容易获得。系统和方法的各实施例就地构建功率模型。为了解决第二个挑战, 系统和方法的各实施例利用关于调度虚拟机的资源的知识。同样, 缺少对全部资源的完整的可视性, 并且必须进行操纵开销和准确性之间的折衷。

II. 系统概览

图 1 是示出在主机服务器计算设备 110 上实现的虚拟机功率计量系统 100

和方法的各实施例的综合概述的框图。一般而言, 虚拟机功率计量系统 100 和方法的各实施例确定在主机服务器计算设备 110 上运行的各个虚拟机的功耗, 而不需要专用硬件。

虚拟机功率计量系统 100 的各实施例驻留在主机服务器计算设备 110 上。

主机服务器计算设备 110 表示物理服务器。一个或多个虚拟机 120 由主机服务器计算设备 110 来托管。这些一个或多个虚拟机 120 可被表示为 VM_1 、 VM_2 ... 到 VM_N , 其中 N 是虚拟机的数量。功率跟踪模块 130 测量全部的系统功率, 或者由主机服务器计算设备 110 消耗的总功率。在某些实施例中, 功率跟踪模块 130 代表软件驱动器, 从主机服务器计算设备 110 主板上的可用的功率传感器读取全部的系统功耗。

资源跟踪模块 140 代表跟踪虚拟机的资源使用的系统管理程序性能计数

器。系统管理程序(也被称为虚拟机管理器)是允许多个操作系统在主机计算机上并发地运行的管理软件。资源跟踪模块 140 跟踪主机服务器计算设备 110 的资源。为了保持虚拟机功率计量系统 100 的各实施例可用于

现有的系统管理程序，系统 100 的各实施例使用现有的系统管理程序性能计数器。然而，在系统和方法的某些实施例中，与现有系统管理程序的可用性不是必需时，可添加额外的性能计数器和工具以监视资源使用。系统 100 的各实施例处理该信息以计算并输出每台虚拟机 150 的功耗。

图 2 是示出在图 1 所示的虚拟机功率计量系统 100 和方法的各实施例中使用的模块和数据的框图。一般而言，虚拟机功率计量系统 100 的各实施例包括用于虚拟机中功率测量和管理的模块。更具体地，虚拟机功率计量系统 100 的各实施例包括功率模型训练模块 200，该功率模型训练模块 200 基于来自功率跟踪模块 130 和资源跟踪模块 140 的信息来生成功率模型。在模块 200 的某些实施例中，该功率模型还通过考虑中央处理单元（CPU）的频率来细化。

系统 100 的各实施例还包括能量计算模块 210，该能量计算模块 210 使用该功率模型来计算一个或多个虚拟机 120 中每一个的功耗。如下面所解释的，该模块的各实施例使用至少两个实施例来获得功耗。结果是每个单独的虚拟机 150 的功耗。功率设限模块 220 使用该信息来按需提供每台虚拟机的功率限制。

功率模型训练模块 200 使用主机服务器 230 的总的功率使用和资源利用率测量结果 240 来生成功率模型 250。如下面详细地解释的，采用多个测量结果来获得一系列功率模型等式。对该系列等式求解以发现最优功率模型系数 260。

能量计算模块 210 使用功率跟踪模块 130 信息、资源跟踪模块 140 信息、和最优功率模型系数 260 来计算一个或多个虚拟机 120 中每一个的功耗。存在能量计算模块 210 的两个实施例，如图 2 中虚线所示。

能量计算模块 210 的第一个实施例使用虚拟机资源百分比来进行计算 270。

在该实施例中，虚拟机中每一个所使用的系数是相同的。例如，如果两

个虚拟机都使用 10%CPU（并且没有其他资源），则它们的能量使用将是相同的。能量计算模块 210 的第二个实施例使用虚拟机资源百分比和虚拟机专用的功率模型来进行计算 280。在该实施例中，作为示例，即使两个虚拟机中每一个都使用 10%CPU，但能量使用可能不同，因为一个虚拟机可能在执行浮点运算，而另一个虚拟机可能在执行定点运算（这两个运算使用 CPU 的不同的子单元，并因此可能使用不同量的能量）。

下面讨论这两个实施例的数学细节。这两个实施例都通过使用功率模型扩

充模块 290 稍微增加功率模型的复杂度来细化。功率模型扩充模块 290 通过考虑主机服务器计算设备 110 的中央处理单元（CPU）的频率来扩充功率模型。

结果是每个单独的虚拟机 150 的功耗。该结果随后由功率设限模块 220

用来对每个虚拟机所使用的功率进行设限。下面详细地讨论功率设限过程的细节。

III. 操作概览

计算基础设施（诸如数据中心）中的计算活动通常包括由在不同的物理机

器上托管的多个虚拟机或服务器角色所执行的多个任务。例如，搜索查询可首先被 web 前端服务（也被称为 web 前端服务器角色）接收，并可被主存在虚拟机内部。随后，该搜索查询可由多个搜索数据检索服务器、广告服务器、和内容呈现服务器来处理。处理的结果被收集并被发送给用户。

作为另一个示例，活动可包括接收特定用户的电子邮件。首先，（可能在虚拟机内部托管的）传入电子邮件服务器角色将接收该消息，并确定该消息正确的用户账户。病毒扫描服务器可以检查该电子邮件消息的内容。垃圾邮件过滤服务器也可以检查该消息的内容。随后，存储服务器可将该消息置于其存储中。每一情形中的活动包括多个分布式服务器角色上的能量支出，服务器角色中的某些可在同一机器上被托管，而某些可在不同的机器上。期望确定活动作为整体所使用的能量。

图 3 是示出图 1 和 2 所示的虚拟机功率计量系统 100 的各实施例的总体操作的流程图。一般而言，虚拟机功率计量系统 100 和方法的各实施例确定单独活动的能量使用，该单独活动可被分布在若干虚拟机上。活动被定义为服务或请求。例如，搜索查询是具有可验证的起始和结束的活动。

参考图 3，该方法通过选择要确定其功耗的活动（框 300）来开始。该活动跨越多个物理主机服务器中每一个上的至少一个虚拟机。接下来，当首先在虚拟机上初始化该活动时，生成一起始事件（框 310）。该起始事件标记开始跟踪该活动所使用的每个资源（框 320）。当活动使用特定的虚拟机时，确定在活动使用期间相关虚拟机的能量使用（或功耗）（框 330）。

应当注意，活动可在不同的时间、在不同的虚拟机上开始和结束。例如，在刚给出的电子邮件的示例中，虚拟机可在时间 T_{1} 开始对电子邮件进行病毒扫描，并在时间 T_{2} 结束扫描。另一个虚拟机随后可在时间 T_{3} 开始对同一电子邮件进行垃圾邮件的分析，并在时间 T_{4} 结束该分析。

可使用下面详细地描述的虚拟机功率计量方法来确定每个虚拟机的功耗。

这生成在所选择的活动中工作的每个虚拟机的功耗测量结果。随着活动进行，将跨多个服务器角色使用的能量相加以生成该活动所使用的总能量。

系统 100 的各实施例随后确定并标记该活动在其中被终止的结束事件（框 340）。该结束事件可以是履行搜索请求，或者标记该活动被完成或被终止并呈现给用户的时间的某些其它事件。活动时间段被定义为起始事件和结束事件之间的时间量（框 350）。

另外，不是虚拟机功率计量方法所跟踪的计算基础设施的一部分的设备可能消耗一些能量。这些可包括冷却计算工具和这些外围设备的仪器。如下估计在这些设备中使用的能量。系统 100 的各实施例确定在活动时间段期间被使用、但未被跟踪的设备的总功耗（框 360）。

使用硬件功率计量或使用（基于规范和数据表或就地学到的）已知的功率模型来测量这些设备所使用的总能量。随后确定由所选择的活动在活动时间段期间所使用的总功耗的部分（框 370）。总能量可在全部活动间被均等地分配、按其执行时间的比率或按先前计算的它们在虚拟机内消耗的能量比率被划分。

系统 100 的各实施例随后对所选择的活动在活动时间段期间的功耗求和，以获得总的活动功耗（框 380）。具体而言，对在活动时间段期间为该活动工作的每个虚拟机的功耗测量结果、以及在活动时间段期间所使用的未被跟踪设备的功耗的部分求和。这产生总的活动功耗，或活动中所使用的全部虚拟机上的活动执行中使用的功耗，无论虚拟机驻留在哪个物理主机服务器上。随后输出总的活动功耗（框 390）。

上面的讨论使用术语“能量使用”和“功耗”。这些可以能量为单位（诸如，焦耳或千瓦时）来报告。还可使用功率单位（诸如瓦特），就已提取的功率来报告功耗。在报告是就功率而言的情形中，可以提供表示在单位时间间隔期间提取的功率的时间序列。总的活动功耗可被报告为每个虚拟机活动的以能量为单位的单个数字，或者在活动时间段期间每个虚拟机的以能量为单位的单个数字。或者，总的活动功耗可被报告为活动时间段期间以功率为单位的具有多个值的时间序列。

IV. 系统和操作细节

现在将讨论虚拟机功率计量系统 100 和方法的各实施例的系统和操作细节。这些实施例包括功率模型训练模块 200、能量计算模块 210、功率模型扩充模块 290 和功率设限模块 220 的各个实施例。现在将详细讨论该系统和这些模块中的每一个的操作细节。

IV. A. 功率模型训练模块

功率模型训练模块 200 生成用于确定虚拟机功耗的功率模型。图 4 是示出图 2 所示的功率模型训练模块 200 的各实施例的操作细节的流程图。方法通过输入总功耗来开始（框 400）。这由功率跟踪模块 130 来报告。换言之，主机服务器计算设备 110 向功率跟踪模块 120 报告所使用的总功率。

另一个输入是对虚拟机正在使用多少给定资源进行分解（框 410）。该信息由资源跟踪模块 140 来报告。“资源”可以是主机服务器计算设备 110 的计算机基础设施的一部分的任一组件。例如，资源可以是处理器（诸如，中央处理单元（CPU））、磁盘、存储器等。资源跟踪模块 140 确定虚拟机正在利用每个资源中的多少（被称为“资源利用率”）。随后对每个资源

的资源利用率 求和，以获得总的资源利用率（框 420）。

主机服务器计算设备 110 包括将这些资源分配到驻留在其上的任意虚拟机的操作系统（未示出）。例如，主机操作系统可认为虚拟机能得到多达 CPU 储蓄的 30%。当主机操作系统保留了这 30% 时，它还跟踪到该虚拟机当前正在使用它被分配的 CPU 存储的 30% 中的 17%。因此，主机操作系统既要强调它绝不超出 30%，又要跟踪该虚拟机当前正在使用什么。为驻留在主机服务器计算设备 110 上的每个虚拟机完成以上操作。以此方式，主机服务器计算设备 110 向资源跟踪模块给出关于每个虚拟机正在使用一个资源中的多少的信息。

令 M 表示资源的数量，并令 N 表示观察的数量。对于物理服务器中 M 数量个资源中的每一个，模块 200 的各实施例确定将全部虚拟机相加的总的资源率（框 430）。

一般而言，基于每个虚拟机的资源利用率来生成功率模型。这通过取得每个资源的利用率并将虚拟机正使用的全部资源相加以获得总的资源利用率来完成。在数学上，假设存在由虚拟机正在使用的三个资源，并且 a_1 、 a_2 和 a_3 是最优功率模型系统，它们均是常数。随后功率模型将资源利用率与总功耗（TPC）如下相关：

在理论上，如果对资源利用率的三个不同的值进行总功耗的三个不同的测量，则我们具有三个等式和三个未知数。因此，模块 200 的各实施例确定由虚拟机正在使用的 M 数量个资源。可对一系列 M 个等式求解以确定 M 个未知数。然而，在实践中，系统的大多数实施例收集 N 个测量结果，

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/458077073106006103>