(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 103649938 A
(43)申请公布日 2014.03.19

(21)申请号 201280034965.1

(22)申请日 2012.06.13

(30)优先权数据
11175337.2 2011.07.26 EP

(85)PCT国际申请进入国家阶段日
2014.01.14

(86)PCT国际申请的申请数据
PCT/IB2012/052984 2012.06.13

(87)PCT国际申请的公布数据
WO2013/014545 EN 2013.01.31

(71)申请人 国际商业机器公司
地址 美国纽约阿芒克

(72)发明人 G·M·N·丹科尔 A·B·布朗

(74)专利代理机构 北京市金杜律师事务所 11256
代理人 王茂华 李峥宇

(51)Int.Cl.
*G06F 15/16*(2006.01)

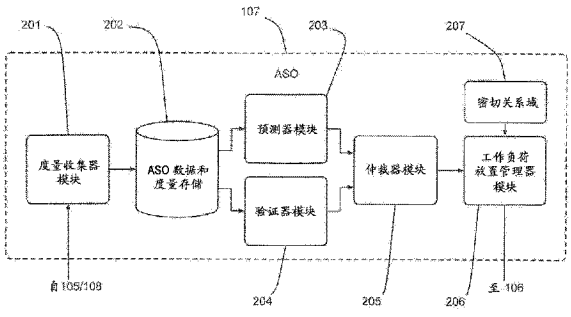
权利要求书2页 说明书8页 附图6页

(54)发明名称

在多处理计算机系统中管理工作负荷

(57)摘要

公开了一种自动系统优化器，所述自动系统优化器配置为在多处理计算机系统中优化工作负荷的处理。




CN 103649938 A

1. 一种用于在多处理计算机系统中管理工作负荷的方法，所述方法包括步骤：

针对多个处理器核限定密切关系域的集合，每一个所述密切关系域包括所述处理器核的集合；

限定密切关系测量，所述密切关系测量配置为指示给定工作负荷应当在何时被移动到包括较少处理器核的较小的密切关系域；

限定性能测量，所述性能测量配置为指示给定工作负荷的性能；

根据所述密切关系测量和所述性能测量来测量给定工作负荷；

如果所述密切关系测量指示给定工作负荷应当被移动到较小的密切关系域，则将所述工作负荷移动到较小的密切关系域；或者

如果所述性能测量指示性能降低，则将所述工作负荷移动到较大的密切关系域。

2. 根据权利要求 1 所述的方法，包括进一步的步骤：

如果来自所述密切关系测量和所述性能测量的所述指示冲突，则关于预定的协议来解决所述冲突。

3. 根据权利要求 2 所述的方法，其中，所述预定的协议包括遵循来自所述性能测量的所述指示以将所述工作负荷移动到较大的密切关系域。

4. 根据前述权利要求中的任意一项所述的方法，其中，所述密切关系域的集合包括在给定的密切关系域中根据处理器核的数量而排序的密切关系域的等级。

5. 根据前述权利要求中的任意一项所述的方法，包括进一步的步骤：

根据来自所述密切关系测量或者所述性能测量的所述指示来指示工作负荷放置方式。

6. 根据前述权利要求中的任意一项所述的方法，其中，所述密切关系测量和所述性能测量包括互相排斥的测量。

7. 根据前述权利要求中的任意一项所述的方法，其中，所述密切关系测量包括对给定工作负荷等待数据所花费的处理时间占总处理时间的比例的测量。

8. 根据前述权利要求中的任意一项所述的方法，其中，所述性能测量包括存储指令对于给定工作负荷的执行速率的测量。

9. 根据前述权利要求中的任意一项所述的方法，其中，周期性地执行根据所述密切关系测量和所述性能测量的给定工作负荷的所述测量。

10. 根据前述权利要求中的任意一项所述的方法，其中，给定工作负荷在域之间的移动被减幅。

11. 根据前述权利要求中的任意一项所述的方法，通过操作系统守护程序来执行。

12. 一种用于在多处理计算机系统中管理工作负荷的装置，所述装置操作为：

针对多个处理器核限定密切关系域的集合，每一个所述密切关系域包括所述处理器核的集合；

限定密切关系测量，所述密切关系测量配置为指示给定工作负荷应当在何时被移动到包括较少处理器核的较小的密切关系域；

限定性能测量，所述性能测量配置为指示给定工作负荷的性能；

根据所述密切关系测量和所述性能测量来测量给定工作负荷；

如果所述密切关系测量指示给定工作负荷应当被移动到较小的密切关系域，则将所述工作负荷移动到较小的密切关系域；或者

如果所述性能测量指示性能降低，则将所述工作负荷移动到较大的密切关系域。

13. 根据权利要求 12 所述的装置，进一步操作为：

如果来自所述密切关系测量和所述性能测量的所述指示冲突，则关于预定的协议来解决所述冲突。

14. 根据权利要求 13 所述的装置，其中，所述预定的协议包括遵循来自所述性能测量的所述指示以将所述工作负荷移动到较大的密切关系域。

15. 根据权利要求 12 到 14 中的任意一项所述的装置，其中，所述密切关系域的集合包括在给定的密切关系域中根据处理器核的数量而排序的密切关系域的等级。

16. 根据权利要求 12 到 15 中的任意一项所述的装置，进一步操作为：

根据来自所述密切关系测量或者所述性能测量的所述指示来指示工作负荷放置方式。

17. 根据权利要求 12 到 16 中的任意一项所述的装置，其中，所述密切关系测量和所述性能测量包括互相排斥的测量。

18. 根据权利要求 12 到 17 中的任意一项所述的装置，其中，所述密切关系测量包括对给定工作负荷等待数据所花费的处理时间占总处理时间的比例的测量。

19. 根据权利要求 12 到 18 中的任意一项所述的装置，其中，所述性能测量包括存储指令对于给定工作负荷的执行速率的测量。

20. 根据权利要求 12 到 19 中的任意一项所述的装置，其中，周期性地执行根据所述密切关系测量和所述性能测量的给定工作负荷的所述测量。

21. 根据权利要求 12 到 20 中的任意一项所述的装置，其中，给定工作负荷在域之间的移动被减幅。

22. 根据权利要求 12 到 21 中的任意一项所述的装置，通过操作系统守护程序来提供。

23. 一种存储在计算机可读介质上并且可加载到计算机的内部存储器中的计算机程序，包括当在计算机上运行所述程序时配置用于执行根据权利要求 1 到 11 中的任意一项所述的方法的软件代码部分。

# 在多处理计算机系统中管理工作负荷

## 技术领域

[0001] 本发明涉及在多处理计算机系统中管理工作负荷。

## 背景技术

[0002] 在多处理器计算机系统中，访问存储器中的数据的处理成本取决于与访问该存储器的处理器相关的存储器的物理单元。与其他线程共享的频繁访问和修改存储器区域的多个线程能够得益于全部在物理上靠近在一起。这降低了获得存在于远程处理器的高速缓存中的高速缓存线路要求的交叉节点业务的量。然而，放置工作负荷以便增加这样的高速缓存密切关系与在整个系统的所有可用资源上均衡工作的更加通常的期望相冲突。而且，将所有工作放置在单个节点上并且本地分配所有存储器将增加高速缓存密切关系，但是，由于对于该节点上的资源的增强的竞争，通常将不会对于所有工作负荷提供良好的性能。

[0003] 配置一些自动优化系统以便尝试基于诸如密切关系措施、工作负荷大小或者资源竞争来预测工作负荷的最佳放置。这样的预测优化器支持工作负荷的快速放置并且相应地支持快速性能增益。不能够受益的工作负荷不受影响，除非由于对相关联的或者链接的工作负荷的改变的结果而间接地受影响。

[0004] 然而，由于存在关于域的不同类型、复杂性能特性的工作负荷的许多可能放置，因此由预测优化器执行的分析是复杂的。而且，当进行大的改变时难于预测工作负荷放置的结果或者难于识别其他工作负荷的效果。预测优化器的测试是复杂的并且难于确定系统是鲁棒的。因而预测优化器的充分测试需要是宽泛的以便产生可接受等级的置信度。而且，在每次改变直观推断时，会需要运行大量测试。预测优化器在其优化方面通常是保守的，以便避免使性能降级的可能性。

[0005] 其他自动优化系统尝试经过搜索或者实验来优化工作负荷的放置。假定具有充分时间和稳定的性能工作负荷，这样的基于搜索的优化器能够获得接近优化的工作负荷放置。基于搜索的优化器是相对简单的系统。然而，基于搜索的优化器由于会要求分析并且随着工作负荷的数量呈指数增加的高数量的可能放置组合而具有高运行时间成本。基于搜索的优化器还会由于如下而是破坏性的：只要在实验中，则其他工作负荷以及进行实验的工作负荷就会受到不利影响。

## 发明内容

[0006] 本发明的实施方式提供一种用于在多处理计算机系统中管理工作负荷的方法，所述方法包括步骤：

[0007] 针对多个处理器核限定密切关系域的集合，每一个所述密切关系域包括所述处理器核的集合；

[0008] 限定密切关系测量，所述密切关系测量配置为指示给定工作负荷应当在何时被移动到包括较少处理器核的较小的密切关系域；

[0009] 限定性能测量，所述性能测量配置为指示给定工作负荷的性能；

[0010] 根据所述密切关系测量和所述性能测量来测量给定工作负荷；

[0011] 如果所述密切关系测量指示给定工作负荷应当被移动到较小的密切关系域，则将所述工作负荷移动到较小的密切关系域；或者

[0012] 如果所述性能测量指示性能降低，则将所述工作负荷移动到较大的密切关系域。

[0013] 所述方法可以包括进一步的步骤：如果来自所述密切关系测量和所述性能测量的所述指示相互冲突，则关于预定的协议来解决所述冲突。所述预定的协议可以包括遵循来自所述性能测量的指示以便将所述工作负荷移动到较大的密切关系域。所述密切关系域的集合可以包括在给定的密切关系域中根据处理器核的数量而排序的密切关系域的等级。所述方法可以包括进一步的步骤：根据来自所述密切关系测量或者性能测量的所述指示来指示工作负荷放置方式。所述密切关系测量和所述性能测量可以包括互相排斥的测量。所述密切关系测量可以包括对给定工作负荷等待数据所花费的处理时间占总处理时间的比例的测量。所述性能测量可以包括存储指令对于给定工作负荷的执行速率的测量。可以周期性地执行根据所述密切关系测量和所述性能测量的给定工作负荷的所述测量。给定工作负荷在域之间的移动可以被减幅。所述方法可以通过操作系统守护程序来执行。

[0014] 另一实施方式提供一种用于在多处理计算机系统中管理工作负荷的装置，所述装置操作为：

[0015] 针对多个处理器核限定密切关系域的集合，每一个所述密切关系域包括所述处理器核的集合；

[0016] 限定密切关系测量，所述密切关系测量配置为指示给定工作负荷应当在何时被移动到包括较少处理器核的较小的密切关系域；

[0017] 限定性能测量，所述性能测量配置为指示给定工作负荷的性能；

[0018] 根据所述密切关系测量和所述性能测量来测量给定工作负荷；

[0019] 如果所述密切关系测量指示给定工作负荷应当被移动到较小的密切关系域，则将所述工作负荷移动到较小的密切关系域；或者

[0020] 如果所述性能测量指示性能降低，则将所述工作负荷移动到较大的密切关系域。

[0021] 进一步实施方式提供一种存储在计算机可读介质上并且可加载到计算机的内部存储器中的计算机程序，包括当在计算机上运行所述程序时配置用于执行用于管理多处理计算机系统中的工作负荷的方法的软件代码部分，所述方法包括步骤：

[0022] 针对多个处理器核限定密切关系域的集合，每一个所述密切关系域包括所述处理器核的集合；

[0023] 限定密切关系测量，所述密切关系测量配置为指示给定工作负荷应当在何时被移动到包括较少处理器核的较小的密切关系域；

[0024] 限定性能测量，所述性能测量配置为指示给定工作负荷的性能；

[0025] 根据所述密切关系测量和所述性能测量来测量给定工作负荷；

[0026] 如果所述密切关系测量指示给定工作负荷应当被移动到较小的密切关系域，则将所述工作负荷移动到较小的密切关系域；或者

[0027] 如果所述性能测量指示性能降低，则将所述工作负荷移动到较大的密切关系域。

## 附图说明

[0028] 现在将参照附图仅通过示例的方式来描述本发明实施方式，在附图中：

[0029] 图 1 是包括自动系统优化计划(ASO)的计算机的示意性示出；

[0030] 图 2 是在图 1 的计算机中的 ASO 的部件的示意性示出；

[0031] 图 3a 是示出对于图 2 的 ASO 的度量数据规范的表；

[0032] 图 3b 是示出对于图 2 的 ASO 的度量限定的表；以及

[0033] 图 4 到图 7 是示出由图 2 的 ASO 执行的处理的流程图。

## 具体实施方式

[0034] 参照图 1，计算机 101 包括多处理处理器 102 并且装载有操作系统 103，该操作系统 103 配置为对于以工作负荷 104 形式的一个或者多个应用程序提供处理平台。处理器 102 提供有配置为监控处理器 102 的参数的预定集合的性能监视单元(PMU)105。操作系统(OS)103 进一步包括提供活跃系统优化器(ASO)107 和 OS103 的核心功能的内核 106。在当前实施方式中，ASO107 是用户空间守护程序，并且配置为从 PMU105 输入选择的参数，并且从内核 106 输入内核统计 108。ASO107 配置基于以输入的 PMU105 参数和内核统计 108 来优化每一个被监控的工作负荷 104 的性能，如下面进一步详细描述的。将对于 ASO107 的输入 PMU105 参数和内核统计 108 初始设置为缺省值，在 OS103 启动时可以由系统管理员修改该缺省值。

[0035] 参照图 2，在当前实施方式中，ASO107 包括度量收集器模块 201、ASO 数据和度量存储 202、预测器模块 203、验证器模块 204、仲裁器模块 205、工作负荷放置管理器模块 206 和一组密切关系域 207。密切关系域 207 是对于处理器 102 限定的处理器核的预定的集合或者分组。密切关系域可以包括从单个处理器核到可用的所有处理器核的任何数量的核。在当前实施方式中，每一个密切关系域 207 包括基于处理器核的各自数量的等级。在当前实施方式中，处理器 102 包括两个芯片，每个芯片分别提供八个核产生总共十六个处理器核。按照下面将处理器 102 划分为五个密切关系域：

[0036] 十六个第五级域，每一个域包括各自的单个核；

[0037] 八个第四级域，每一个域包括两个核；

[0038] 四个第三级域，每一个域包括四个核；

[0039] 两个插槽级域，每一个域包括八个核；以及

[0040] 一个书籍级(book-level)域，包括十六个核。

[0041] 度量收集器模块 201 配置为以性能数据的形式从 PMU105 收集数据的预定集合，并且从内核 106 收集内核统计 108。

[0042] 将收集的数据存储在 ASO 数据和度量存储 202 中。对于正在被监控的每一个工作负荷 104 收集数据。在当前实施方式中，参照图 3a，收集下面的内核统计 108：

[0043] 等待每指令数据花费的时钟周期(CCWD)；以及

[0044] 总的每指令时钟周期(CPI)。

[0045] 从 PMU105 收集下面的性能数据：

[0046] 执行的存储指令的速率(ESI/s)。

[0047] 将每一个数据集合与相关工作负荷 104 相关地存储在 ASO 数据记录 301 中，该相关工作负荷 104 由唯一的工作负荷标识符(WLID)识别。

[0048] 参照图 3b，ASO 进一步包括一组度量 302。ASO 度量 302 包括密切关系测量（AM）303，该密切关系测量（AM）303 的形式为等待数据以便完成指令花费的处理器时钟周期与时钟周期的总数量之间的比值。AM303 与密切关系测量阈值相关联，在当前实施方式中，该密切关系测量阈值为 30%。ASO 度量进一步包括执行的存储指令的速率（ESI/s）的形式的性能测量（PM）304。PM304 与性能阈值 305 相关联，在当前实施方式中，该性能阈值 305 包括对于给定工作负荷的由 ASO107 计算的峰值 ESI/s。

[0049] 预测器模块 203 配置为对于每一个相关工作负荷周期性地监控密切关系测量（AM）303，并且将 AM303 与 30% 的 AM 阈值进行比较。在当前实施方式中，测量时段为 5 秒钟。如果对于给定测量时段的 AM303 落在 AM 阈值之下，则预测器 203 配置为不采取进一步动作。然而，如果 AM303 对于给定测量时段满足或者超出 AM 阈值，则预测器 203 配置为向仲裁器模块 205 发出指令，如果可能，推荐将相关工作负荷分配到密切关系域 207 中的较小一个。

[0050] 验证器模块 204 配置为与预测器模块 203 并行工作，以便对于每一个相关工作负荷周期性地监控性能测量（PM）304，并且将 PM304 与 PM 阈值 305 进行比较。只要对于给定的测量时段 PM304 满足或者超出 PM 阈值 305，则验证器 204 配置为不采取进一步的动作。然而，如果对于给定测量的 PM304 小于 PM 阈值 305，则验证器 204 配置为向仲裁器模块 205 发出指令，推荐将相关工作负荷分配到密切关系域 207 中的较大一个。对于给定的工作负荷，验证器 204 配置为将 PM 阈值计算为峰值 ESI/s，并且使用其来用于测量相关的工作负荷 104。

[0051] 仲裁器模块 205 配置为在每一个数据时段中对从预测器 203 和验证器 204 接收的指令进行仲裁。如果仅从预测器 203 和验证器 204 之一接收到用于将给定工作负荷分配到较大或者较小密切关系域 207 的指令，则不存在冲突并且将该指令传送到工作负荷放置管理器模块 206。如果从预测器 203 和验证器 204 中的每一个接收到将给定工作负荷分配到较大和较小密切关系域 207 的冲突的指令，则仲裁器 205 配置为选择用于向工作负荷放置管理器模块 206 进行转发的合适指令。在当前实施方式中，响应于这样的冲突，仲裁器 205 配置为总是从验证器 204 选择指令，将给定工作负荷分配到较大密切关系域 207，用于向工作负荷放置管理器模块 206 转发。

[0052] 在当前实施方式中，工作负荷放置管理器模块 206 配置为管理测量时段，并且指示预测器模块 203 和验证器模块 204 在每一个测量时段结束时执行它们各自的度量测量。响应于经由仲裁器模块 205 接收的来自预测器模块 203 和验证器模块 204 的推荐，工作负荷放置管理器模块 206 进一步配置为确定是否可以实现该推荐，并且如果为是，则相应地指示内核 106。在一些情况下，工作负荷放置管理器模块 206 可能不能够实现推荐以便增加或者降低对于给定工作负荷的密切关系域。工作负荷放置管理器模块 206 进一步配置为对于正在被监控的每一个工作负荷提供减幅因数。提供该减幅因数以便减小给定工作负荷可以改变密切关系域的频率。通常，减幅因数配置为避免给定工作负荷返回到该工作负荷仅在近期从其移动的给定的密切关系域。

[0053] 在当前实施方式中，当工作负荷放置管理器 206 将工作负荷放置在较大密切关系域上时，对于给定工作负荷，将减幅因数设置到预定值。在当前实施方式中，该预定值为 20。如果对于每一个工作负荷的减幅因数大于零，则在每一个测量时段该减幅因数减小一。只

要减幅因数大于零，就不允许工作负荷放置管理器 206 将工作负荷放置在较小域上，而是可以与减幅因数值无关地放置在较大域上。

[0054] 因而 ASO107 配置为与各自的预定阈值相关地、以密切关系测量（AM）和性能测量（PM）的形式，监控选择的工作负荷 104 相对于预定的度量集合 302 的性能。对于给定的工作负荷，如果密切关系测量超出密切关系测量阈值，则 ASO107 指示内核 106 降低工作负荷在其上运行的密切关系域的尺寸。如果对于给定的工作负荷性能测量不满足性能测量阈值，则 ASO107 指示内核 106 增加工作负荷在其上运行的密切关系域的尺寸。

[0055] 现在将参照图 4 的流程图进一步描述由度量收集器模块 201 执行的处理。在步骤 401 处，响应于 ASO107 的初始化而发起处理，并且处理移动到步骤 402。在步骤 402 处，根据 ASO 数据 301 来识别要被监控的工作负荷 104，并且处理移动到步骤 403。在步骤 403 处，根据 ASO 数据 301 识别要在其上收集数据的数据源的步骤，并且处理移动到步骤 404。在步骤 404 处，对于每一个相关的工作负荷，在 ASO 数据 301 中将 CCWD、CCPI 和 ESI/s 记录日志，并且在 ASO107 的处理时段内，继续这一记录处理。

[0056] 现在将参照图 5 的流程图来进一步描述由预测器模块 203 执行的处理。在步骤 501 处，响应于 ASO107 的启动而发起处理，并且然后处理移动到步骤 502。在步骤 502 处，根据 ASO 度量 302 来识别要被监控的对于工作负荷的密切关系测量（AM），并且处理移动到步骤 503。在步骤 503 处，对于收集到的 ASO 数据 301，对于每一个被监控的工作负荷 104 来计算 AM，并且处理移动到步骤 504。在步骤 504 处，将计算的 AM 与 AM 阈值进行比较，并且处理移动到步骤 505。在步骤 505 处，如果当前计算的 AM 值超出 AM 阈值，则处理移动到步骤 506。在步骤 506 处，将向相关工作负荷的处理分配到较小密切关系域的推荐通信到仲裁器模块 205，并且处理移动到步骤 507。在步骤 507 处，处理等待来自工作负荷放置管理器模块 206 的指示下一个数据周期的信号。当接收到这样的信号时，处理返回到步骤 503 并且如上所述地进行。如果在步骤 505 处当前计算的 AM 值没有超出 AM 阈值，则不向仲裁器模块 205 通信响应，并且处理移动到步骤 507，并且如上所述地进行。

[0057] 现在将参照图 6 的流程图来进一步描述由验证器模块 204 执行的处理。在步骤 601 处，响应于 ASO107 的启动而发起处理，并且然后处理移动到步骤 602。在步骤 602 处，根据 ASO 度量 302 识别要被监控的对于工作负荷的性能测量（PM），并且处理移动到步骤 603。在步骤 603 处，对于收集到的 ASO 数据 301，对于每一个被监控的工作负荷 104 来计算 PM，并且处理移动到步骤 604。在步骤 604 处，将计算的 PM 与 PM 阈值进行比较，并且处理移动到步骤 605。在步骤 605 处，如果当前计算的 PM 值不匹配或者没有超出 PM 阈值，则处理移动到步骤 606。在步骤 606 处，将相关的工作负荷的处理分配到较大的密切关系域的推荐通信到仲裁器模块 205，并且处理移动到步骤 607。在步骤 607 处，处理等待来自工作负荷放置管理器模块 206 的指示下一个数据周期的信号。当接收到这样的信号时，处理移动到步骤 608 以便在返回到步骤 603 来如上所述进行处理之前，更新 PM 阈值。如果在步骤 605 处当前计算的 PM 值匹配或者超出 PM 阈值，则不向仲裁器模块 205 通信响应，并且处理移动到步骤 707 且如上所述进行处理。

[0058] 现在将参照图 7 的流程图来进一步描述由仲裁器模块 205 和工作负荷放置管理器模块 206 的组合执行的处理，其中步骤 706 和 711 包括仲裁器模块 205 的功能。在步骤 701 处，在 ASO107 启动时发起处理，并且处理移动到步骤 702。在步骤 702 处，识别要被监控的

工作负荷104，并且对于每一个被监控的工作负荷来执行随后的处理。然后处理移动到步骤703并且等待逝去预定的数据周期时段，并且然后移动到步骤704。在步骤704处，关于数据周期而向预测器和验证器模块203、204发送信令，并且处理移动到步骤705。在步骤705处，如果从预测器和验证器模块203、204接收到结果，则处理移动到步骤706。在步骤706处，如果在结果之间不存在冲突，则处理移动到步骤707。在步骤707处，检查对于相关工作负荷的减幅因数，以便确定是否允许密切关系域中的任何改变，并且如果允许，则处理移动到步骤708。在步骤708处，通过向内核106发出合适的指令来实现相关指令，并且处理移动到步骤709。在步骤709处，更新相关工作负荷的减幅因数以便反映密切关系域中的改变，并且处理返回到步骤703，并且如上所述进行处理。如果在步骤705处没有从预测器或者验证器模块203、204接收到响应，则处理移动到步骤710，并且在返回到步骤703之前使对于相关工作负荷的减幅因数减量，并且如上所述进行处理。如果在步骤706处接收到冲突的结果，则处理移动到步骤711。在步骤711处，选择来自验证器模块203的结果，并且处理移动到步骤707，并且如上所述进行处理。如果在步骤707处由于给定工作负荷的当前减幅因数而不允许移动该给定工作负荷，则处理移动到步骤710，其中在返回到步骤703以便如上所述进行处理之前使减幅因数减量。

[0059] 在另一实施方式中，仲裁器模块配置为当解决预测器和验证器模块之间的冲突时考虑其他输入。换句话说，由仲裁器模块提供的解决方案可以基于多个因数。在进一步的实施方式中，可以由单独的模块来提供仲裁器模块和工作负荷放置管理器模块的功能。

[0060] 在进一步的实施方式中，如果ASO发现由于不适合的环境而不能够优化性能，则该ASO可以配置为进入休眠模式，并且仅当环境指示ASO能够提供性能改善时才唤醒。

[0061] 在另一实施方式中，为了使系统对于小的性能变换不太敏感，在预定数量的测量时段上执行AM或者PM的平均。如果在预定数量的连续时段内平均的AM满足或者超出AM阈值，则由预测器推荐较小的密切关系域。相反，如果在该预定数量的连续时段内平均的AM落入PM阈值以下，则由验证器推荐较大的密切关系域。

[0062] 在另一实施方式中，利用乘数因数来扩展减幅因数。当工作负荷放置管理器将工作负荷放置在较大域上时，将减幅因数设置为X+旧的减幅因数×乘数，并且然后使该乘数增加1（例如，X=20，初始乘数=0）。只要减幅因数大于0，就不允许工作负荷放置管理器将工作负荷放置在较小的域上，但是能够与减幅因数值无关地放置在较大域上。如果减幅因数大于0，则在每一个测量时段，减幅因数减1。当减幅因数到达0时，重新设置乘数。对于ASO尝试频繁移动的工作负荷，这将增加减幅因数，但是将仍然允许各处移动更加逐渐变化的工作负荷。

[0063] 在进一步的实施方式中，当减幅因数移动到较小域并且不应用乘数时，将减幅因数设置到较小值。这一配置降低了朝向较小域的移动速度。例如，在从书本密切关系域到单核密切关系域的每一个移动之间，可以需要五个测量时段。

[0064] 在进一步的实施方式中，通过延迟预测器和验证器的启动，并且还通过限制对于预测器和验证器的操作时段来限制密切关系域之间的改变频率。

[0065] 本领域普通技术人员将理解，工作负荷放置管理器可能不能够实现来自预测器或者验证器的一些推荐。例如，假定两个密切关系域包括一核域和两核域。两个工作负荷A和B正在运行。工作负荷A已经被压缩到两核域并且工作负荷B当前正在所有三个核上运

行。预测器可以推荐将工作负荷 B 放置在较小域上，但是工作负荷放置管理器可以确定将两个工作负荷放置在两核域上是不可接受的，并且这两个工作负荷太大而不能被放置在仅一个核上。因此工作负荷放置管理器可以配置为使工作负荷 A 压缩在两个核上，并且使工作负荷 B 运行在所有三个核上。

[0066] 在进一步的实施方式中，预测器配置为当确定对于给定工作负荷是否降低密切关系域的尺寸时考虑进一步的度量。这样的进一步度量可以配置为通过验证器降低密切关系域改变反转的数量。例如，CPI 堆栈击穿可以用于权衡密切关系域改变的潜在优点和潜在缺点。

[0067] 在另一实施方式中，预测器提供有识别处理器拓扑的非一致元件的数据。例如，可以在一个插槽上提供六个核并且在另一插槽上提供两个核。可以使预测器知晓该拓扑并且考虑当前域的子域的尺寸。然后，预测器配置为向工作负荷放置管理器通信最小域尺寸，以使得工作负荷放置管理器能够避免将其放置在太小的域上。

[0068] 在进一步的实施方式中，验证器配置为在每一个域上保持工作负荷的性能历史。验证器进一步配置为在不推荐移动回到先前的较大域的情况下，考虑在给定域上性能的诸如 5% 的稍微下降。容忍性能的小幅下降支持工作负荷移动经过中间域以便到达其中可以显著改善性能的其他域。在密切关系域的范围上有效地采样性能之后，验证器可以配置为使用收集到的历史性能数据，以便进行推荐来移动到工作负荷在其上最佳地执行的域。例如，在向下的方向上具有减幅因数 3 对于每一个密切关系域提供性能测量的三个样本。一旦采样了每一级，则验证器配置为推荐移动到具有最佳平均性能的密切关系域。

[0069] 在另一实施方式中，验证器和预测器配置为确保与工作负荷的属性相关的相关性能在尝试优化工作负荷之前的时间段内是稳定的。在应用优化之后，然后按照相同的方式在新的域中监控这些度量。

[0070] 在进一步的实施方式中，ASO 配置为检测诸如添加处理器或存储器或者划分迁移事件的大系统事件，并且可以作为响应而配置为推荐重新设置所有优化。

[0071] 本领域普通技术人员将理解，可以使用任何适合的度量、统计和测量来提供适合于 ASO 的给定应用的密切关系测量和性能测量。而且，密切关系域的限定可以改变并且不需要是分级的，而是配置为适合对于 ASO 的任何给定应用。

[0072] 本发明的实施方式可以应用于单个插槽系统。例如，将具有强相互作用的线程放置在相同的核上会是有利的。锁定竞争提供另一机会，其中 ASO 在单个芯片环境中会是有利的。限制竞争锁定的软件线程到可用硬件线程子集的放置能够产生性能改善。

[0073] 与基于搜索的优化器相比较，本发明的实施方式配置为支持较快且不太破坏的放置或者工作负荷，并且与预测优化器相比较不太保守。预测器防止工作负荷的移动(其不太可能得益于密切关系域的改变)，并且还通常在合适的点处停止朝向较小域的移动。另一方面，由于验证器将使被证明是错误的来自预测器的任何推荐恢复，因此，预测器会是不太保守的。

[0074] 作为由预测器模块缓解的较少实验的结果，本发明的实施方式配置为与基于搜索的优化器或者预测优化器相比较更加有效地支持多个放置的并行处理。与纯实验方案相比较，极大地降低了所需的实验的量，这在具有许多工作负荷的繁忙系统上尤其有利。当一次处理许多工作负荷时，预测优化器面对高度复杂的预测问题。

[0075] 与预测优化器相比较，本发明的实施方式更加容易进行测试并且更加鲁棒。使具有通常相反目标的两个简单部件使用不同的度量降低了测试付出，并且支持并行执行两个部件的测试。在域移动之后，预测器模块配置为从新的域收集度量。因而，预测器具有预测在下一个域中是否存在改善性能的可能性的简化工作。这与预测哪一个域将产生最佳性能的重要工作形成对比。

[0076] 本领域普通技术人员将理解，体现本发明的部分或者全部的装置可以是具有配置为提供本发明的部分或者全部实施方式的软件的通用设备。该设备可以是单个设备或者一组设备，并且该软件可以是单个程序或者一组程序。而且，用于实现本发明的任何或者所有软件可以经由任何合适的传输或者存储单元进行通信，以使得能够将该软件加载到一个或者多个设备上。

[0077] 尽管通过对本发明的实施方式的描述说明了本发明，并且尽管相当详细地描述了实施方式，但是申请人并不意在限制或者以任何方式将所附权利要求的范围局限于这样的细节。附加的优点和变型对于本领域普通技术人员将容易实现。因此，本发明在其更宽泛的方面并不局限于代表性装置和方法的具体细节，以及所示出和描述的说明性示例。因此，在不脱离申请人的一般创造性概念的范围的情况下，可以根据这样的细节进行改变。

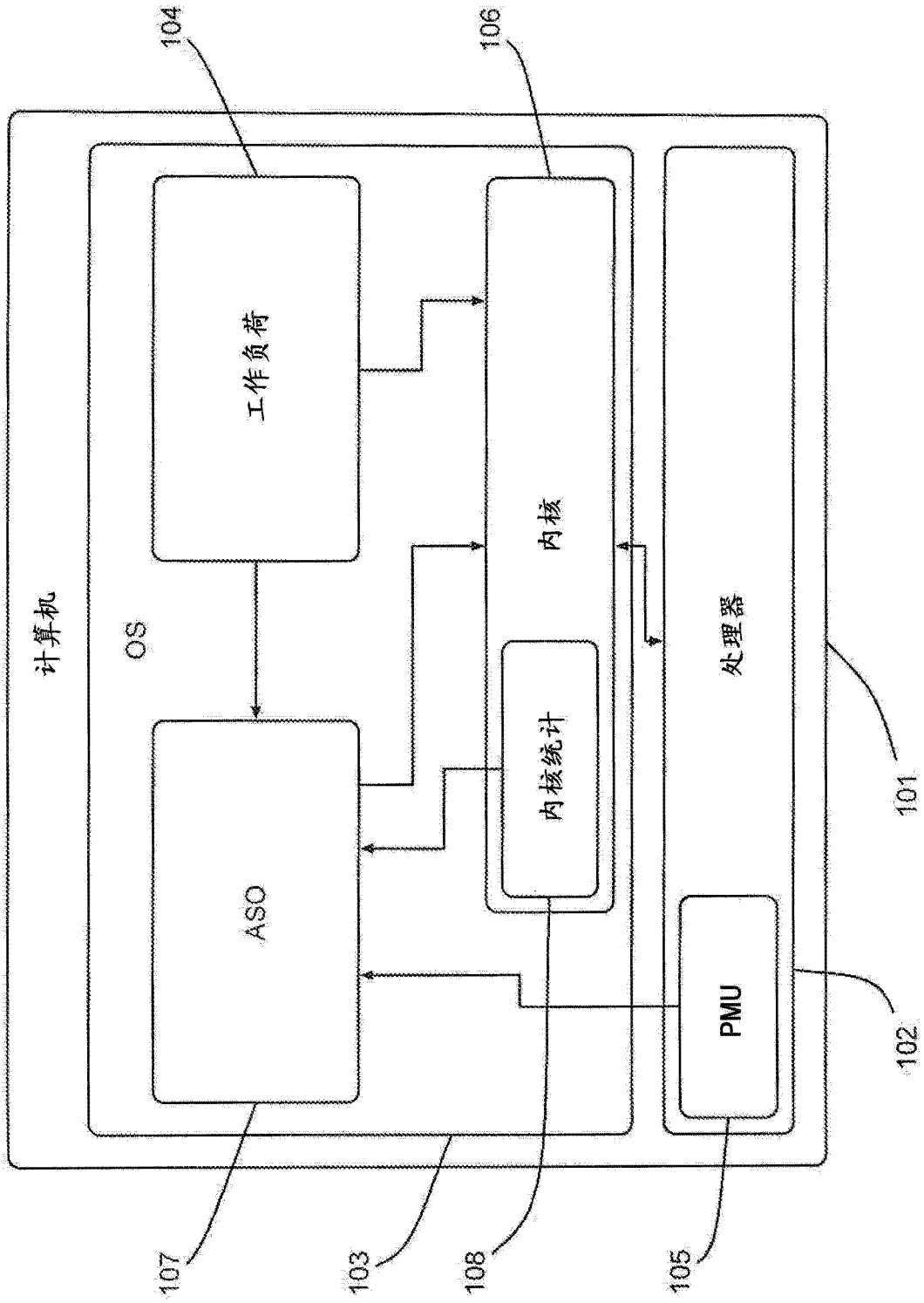



图 1

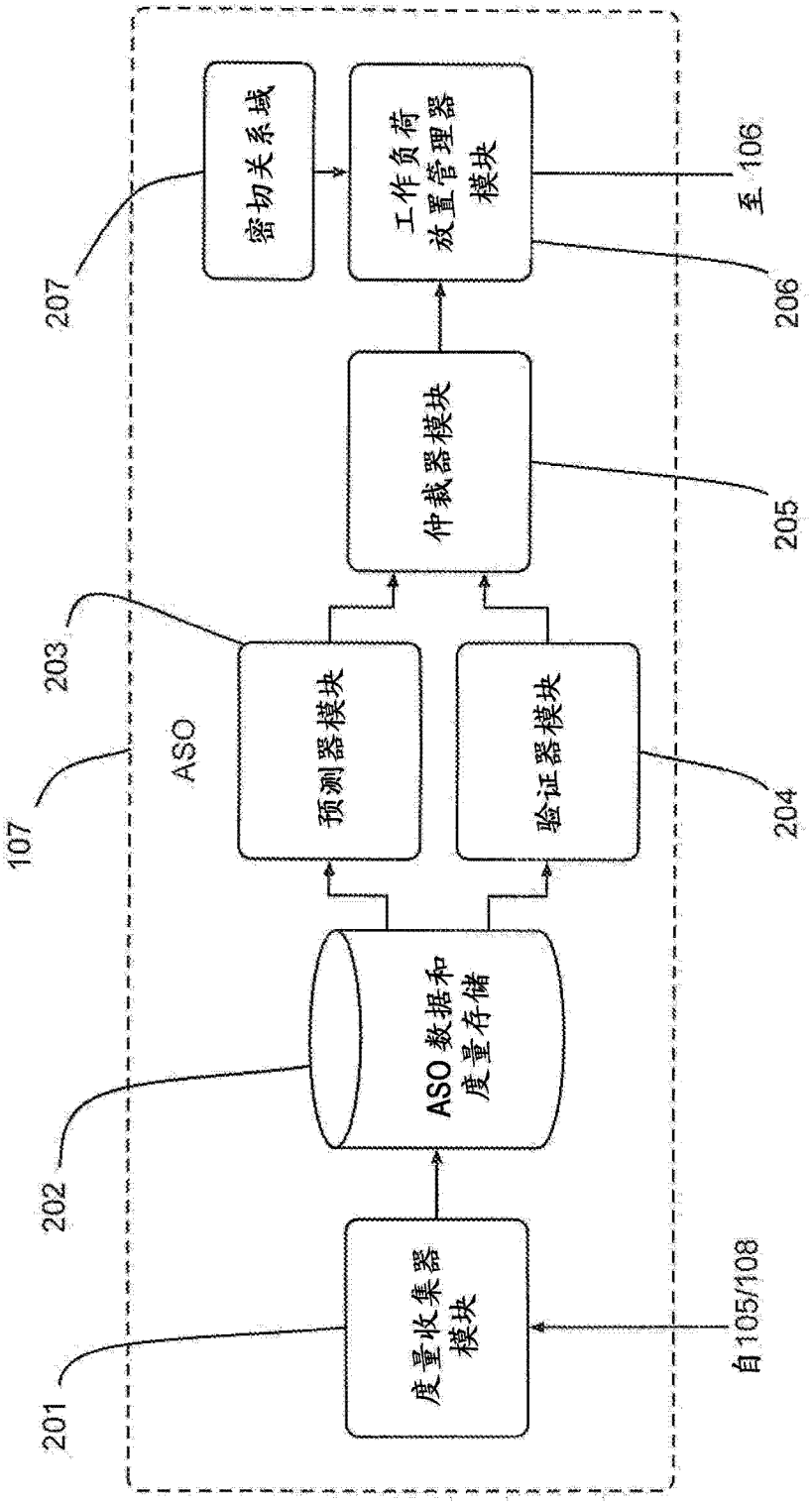

密切关系域
工作负荷放置管理器模块
至 106
207
206
仲裁器模块
205
203
ASO
107
预测器模块
验证器模块
204
ASO 数据和度量存储
202
度量收集器模块
自 105/108
201

图 2

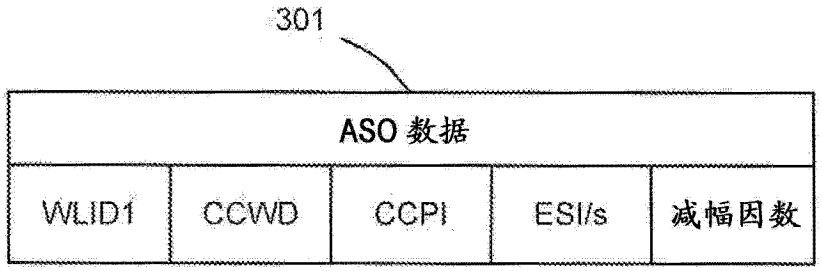

301
ASO 数据
WLID1
CCWD
CCPI
ESI/s
减幅因数

图 3a

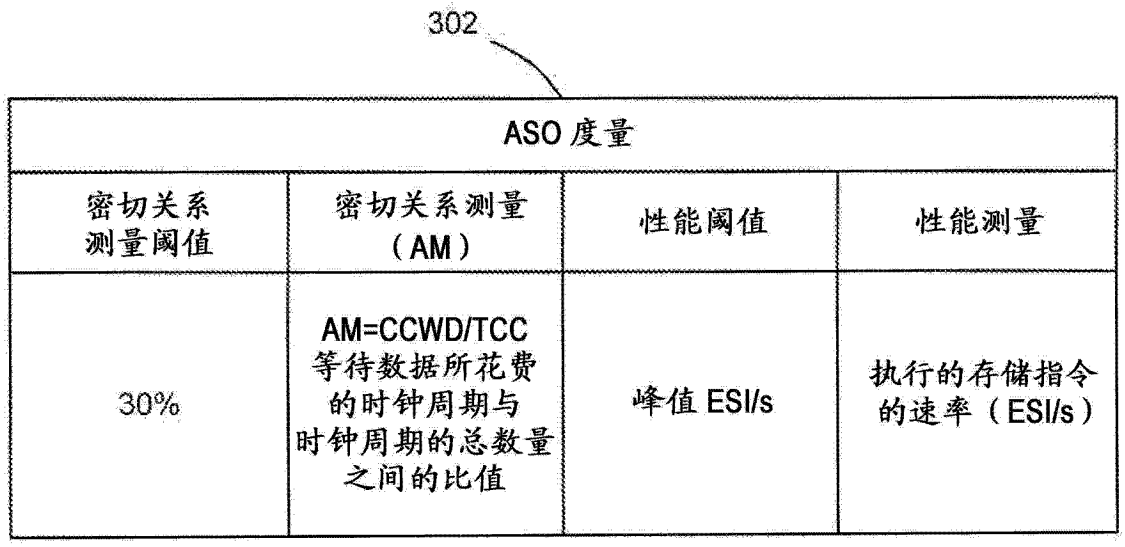

302
ASO 度量
密切关系测量阈值
密切关系测量（AM）
性能阈值
性能测量
30%
AM=CCWD/TCC 等待数据所花费的时钟周期与时钟周期的总数量之间的比值
峰值 ESI/s
执行的存储指令的速率（ESI/s）

图 3b

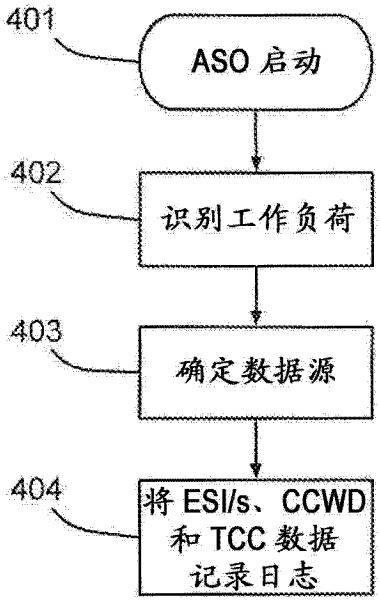

401
ASO 启动
402
识别工作负荷
403
确定数据源
404
将 ESI/s、CCWD 和 TCC 数据记录日志

图 4

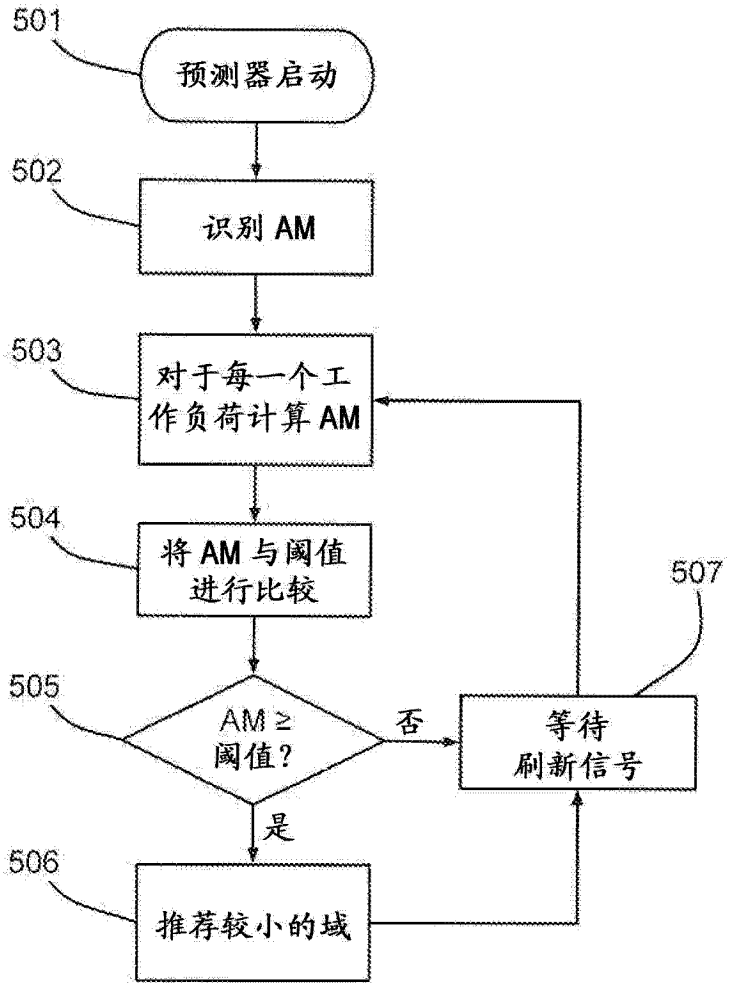

501
预测器启动
502
识别 AM
503
对于每一个工作负荷计算 AM
504
将 AM 与阈值进行比较
505
AM ≥ 阈值?
否
507
等待刷新信号
是
506
推荐较小的域

图 5

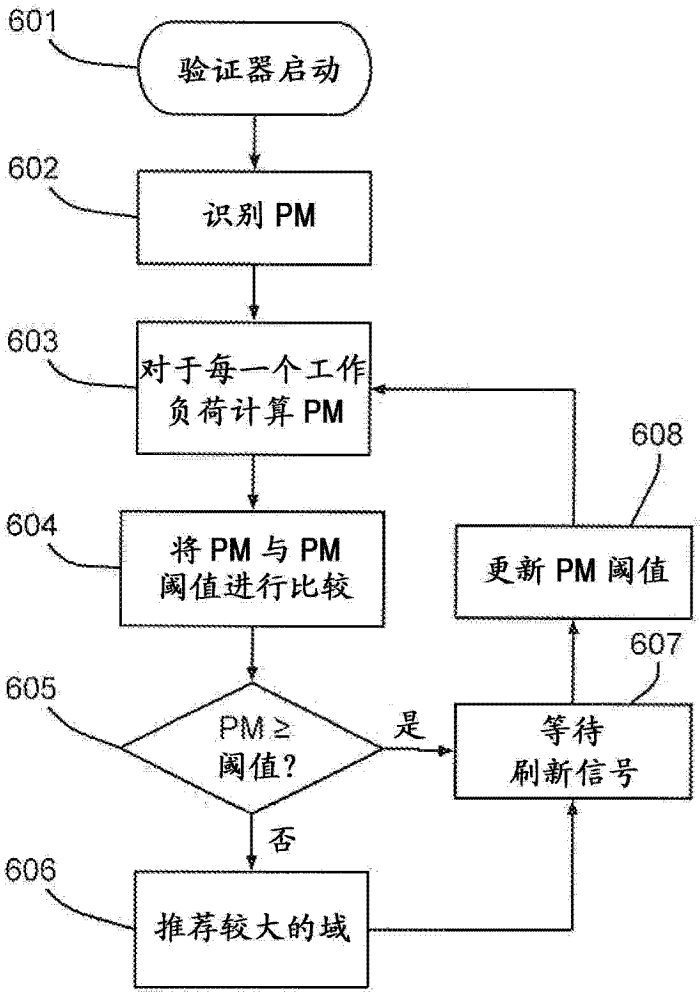

601
验证器启动
602
识别 PM
603
对于每一个工作负荷计算 PM
604
将 PM 与 PM 阈值进行比较
605
PM ≥ 阈值?
是
否
606
推荐较大的域
607
等待刷新信号
608
更新 PM 阈值

图 6

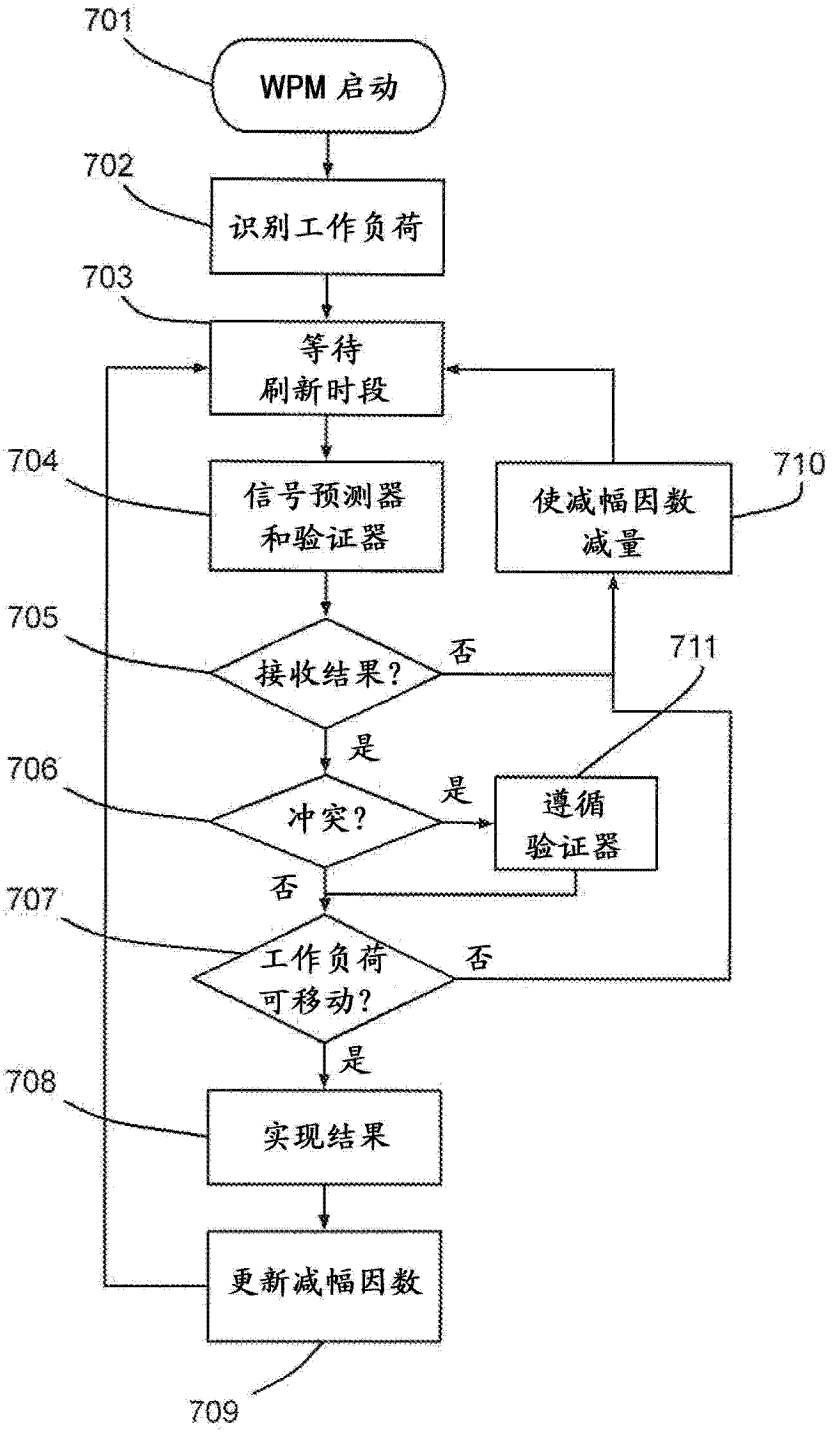

701
WPM 启动
702
识别工作负荷
703
等待
刷新时段
704
信号预测器
和验证器
710
使减幅因数
减量
705
接收结果?
否
711
是
706
冲突?
是
遵循
验证器
否
707
工作负荷
可移动?
否
是
708
实现结果
更新减幅因数
709

图 7