

(12) 发明专利申请

(10) 申请公布号 CN 102110025 A

(43) 申请公布日 2011. 06. 29

(21) 申请号 201010625174. 9

G06F 1/32(2006. 01)

(22) 申请日 2010. 12. 24

(30) 优先权数据

12/647671 2009. 12. 28 US

(71) 申请人 英特尔公司

地址 美国加利福尼亚州

(72) 发明人 J · J · 宋 J · H · 克拉福德

(74) 专利代理机构 中国专利代理(香港)有限公司 72001

代理人 曲卫涛 王洪斌

(51) Int. Cl.

G06F 9/50(2006. 01)

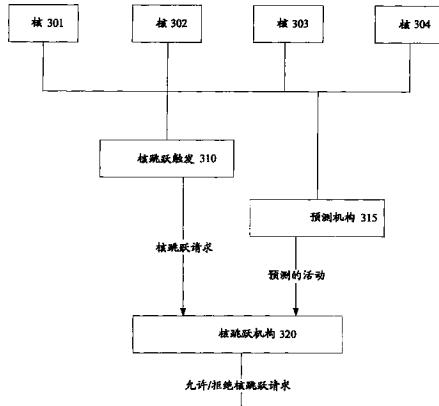
权利要求书 5 页 说明书 17 页 附图 9 页

(54) 发明名称

避免无效核跳跃和提供硬件辅助的低功率状态选择的机构

(57) 摘要

本文描述用于避免无效核跳跃以及提供硬件辅助的功率状态选择的设备和方法。预测核的未来闲置 - 活动。如果预测到有效核跳跃场景的活动模式的驻留期足够大，那么确定核有效并且得到允许。但是，如果预测到有效活动模式驻留的时间不够长，即预测到无效模式驻留的时间较长，那么拒绝核跳跃请求。因此，设计者可实现避免核跳跃的策略，该策略针对执行核跳跃的代价（例如，核跳跃的时间代价）权衡核跳跃的潜在增益（例如，减轻核跳跃状况）。单独地，可在硬件中预测与核的硬件功率状态相关联的闲置持续时间。此外，确定闲置持续时间预测的精度。一旦接收到使核进入某个功率状态的请求，功率管理单元便可选择硬件预测的功率状态（如果精度足够高的话）或利用请求的功率状态（如果硬件预测的精度不够高的话）。



1. 一种设备,包括 :

多个处理器核 ;

预测机构,用于预测在未来间隔所述多个处理器核的未来活动 ;以及

核跳跃机构,用于基于在所述未来间隔所述多个处理器核的所述未来活动确定核跳跃是否有效,并响应于基于所述多个处理器核的所述未来活动确定所述核跳跃不有效而禁止核跳跃事件。

2. 如权利要求 1 所述的设备,其中所述预测机构预测在所述未来间隔所述多个处理器核的未来活动包括所述预测机构预测在所述未来间隔第一闲置 - 繁忙模式的第一驻留期。

3. 如权利要求 2 所述的设备,其中在所述未来间隔的所述第一闲置 - 繁忙模式包括第一聚合闲置 - 繁忙模式,它是多个有效核跳跃闲置 - 繁忙模式的聚合。

4. 如权利要求 3 所述的设备,其中所述核跳跃机构基于在所述未来间隔所述多个处理器核的所述未来活动确定核跳跃是否有效包括所述核跳跃机构执行以下步骤 :响应于所述第一聚合闲置 - 繁忙模式的所述第一驻留期大于阈值而确定所述核跳跃有效 ;以及响应于所述第一聚合闲置 - 繁忙模式的所述第一驻留期不大于所述阈值而确定所述跳跃不有效。

5. 如权利要求 1 所述的设备,其中所述预测机构预测在所述未来间隔所述多个处理器核的未来活动包括所述预测机构利用在前一个间隔期间所述多个处理器核的过去活动作为在所述未来间隔所述多个处理器核的所述未来活动。

6. 如权利要求 1 所述的设备,其中所述预测机构预测在所述未来间隔所述多个处理器核的未来活动包括所述预测机构执行以下步骤 :从软件实体接收活动提示 ;以及基于来自所述软件实体的所述活动提示预测在所述未来间隔的所述未来活动。

7. 如权利要求 1 所述的设备,还包括用于至少基于热密度状况生成所述核跳跃事件的核跳跃触发逻辑,其中所述核跳跃事件包括执行核跳跃的请求,并且其中所述核跳跃机构禁止所述核跳跃事件包括所述核跳跃机构拒绝执行所述核跳跃的所述请求。

8. 如权利要求 7 所述的设备,其中所述核跳跃机构基于在所述未来间隔所述多个处理器核的所述未来活动确定核跳跃是否有效包括确定所述多个处理器核的所述未来活动是否指示用于减轻所述热密度状况的核跳跃是可行的。

9. 如权利要求 1 所述的设备,其中所述多个核、所述预测机构和所述核跳跃机构包含在微处理器中,所述微处理器耦合到存储器,其中所述存储器选自由动态随机存取存储器 (DRAM)、双倍数据速率 (DDR) RAM 和静态随机存取存储器 (SRAM) 组成的组。

10. 一种设备,包括处理器,所述处理器包括 :

多个核 ;

耦合到所述多个核的核跳跃逻辑,用于触发核跳跃请求 ;

耦合到所述多个核的预测逻辑,用于预测在下一个间隔所述多个核的活动 ;以及

耦合到所述核跳跃逻辑和所述预测逻辑的核跳跃管理器逻辑,所述核跳跃管理器逻辑响应于所述核跳跃管理器确定在所述下一个间隔所述多个核的所述活动指示核跳跃是无效的而拒绝所述核跳跃请求。

11. 如权利要求 10 所述的设备,其中用于触发核跳跃请求的所述核跳跃逻辑基于热密度状况。

12. 如权利要求 11 所述的设备,其中所述多个核的所述活动包括在所述下一个间隔期

间聚合活动模式的预测的驻留期。

13. 如权利要求 12 所述的设备,其中所述多个核包括 N 个核,并且多种可能的活动模式包括 2^N 种模式,并且其中所述聚合活动模式包括对于减轻所述热密度状况可行的所述 $2N$ 种活动模式的子集。

14. 如权利要求 13 所述的设备,其中所述核跳跃管理器确定在所述下一个间隔所述多个核的所述活动指示核跳跃是无效的包括所述核跳跃管理器确定对于减轻所述热密度状况可行的所述 $2N$ 种活动模式的所述子集的所述预测的驻留期不超过阈值驻留期。

15. 如权利要求 14 所述的设备,其中响应于所述核跳跃请求,所述核跳跃管理器响应于所述核跳跃管理器确定在所述下一个间隔所述多个核的所述活动指示核跳跃是有效的而启动核跳跃以减轻所述热密度状况,并且其中所述核跳跃管理器确定在所述下一个间隔所述多个核的所述活动指示核跳跃是有效的包括所述核跳跃管理器确定对于减轻所述热密度状况可行的所述 2^N 种活动模式的所述子集的所述预测的驻留期超过所述阈值驻留期。

16. 如权利要求 14 所述的设备,其中所述核跳跃管理器响应于所述核跳跃请求启动核跳跃以减轻所述热密度状况包括 :所述核跳跃管理器确定最佳活动模式以减轻所述热密度状况 ;以及至少启动所述核跳跃以获得所述最佳活动模式。

17. 如权利要求 10 所述的设备,其中所述多个核、所述核跳跃逻辑、所述预测逻辑和所述核跳跃管理器逻辑包含在微处理器中,所述微处理器耦合到存储器,其中所述存储器选自由动态随机存取存储器 (DRAM)、双倍数据速率 (DDR) RAM 和静态随机存取存储器 (SRAM) 组成的组。

18. 一种方法,包括 :

预测在未来间隔期间处理器上的多个核的闲置 - 活动表示的预测的驻留期 ;

接收核跳跃请求 ;

确定在所述未来间隔期间所述多个核的所述闲置 - 活动表示的所述预测的驻留期是否大于驻留期阈值 ;

响应于所述核跳跃请求,响应于确定在所述未来间隔期间所述闲置 - 活动表示的所述预测的驻留期大于所述驻留期阈值而执行核跳跃 ;以及

响应于所述核跳跃请求,响应于确定在所述未来间隔期间所述闲置 - 活动表示的所述预测的驻留期不大于所述驻留期阈值而不执行所述核跳跃。

19. 如权利要求 18 所述的方法,其中所述闲置 - 活动表示包括全部数量的可能的闲置 - 活动模式中的第一数量的闲置 - 活动模式的第一聚合,其中确定所述第一数量的闲置 - 活动模式对于核跳跃是有效的。

20. 如权利要求 19 所述的方法,其中确定所述第一数量的闲置 - 活动模式对于核跳跃是有效的是基于权衡与执行所述核跳跃相关联的时间代价以及与执行所述核跳跃相关联的热密度增益的策略。

21. 如权利要求 20 所述的方法,其中权衡与执行所述核跳跃相关联的所述时间代价以及与执行所述核跳跃相关联的所述热密度增益的所述策略包括用于将所述全部数量的可能的闲置 - 活动模式分为以下三组的策略 :所述第一数量的闲置 - 活动模式,这响应于所述第一数量的闲置 - 活动模式与核跳跃可行性相关联而进行 ;所述全部数量的闲置活动模式中的第二数量的闲置活动模式,这响应于所述第二数量的闲置活动模式与最小热密度增益

相关联而进行；以及所述全部数量的闲置活动模式中的第三数量的闲置活动模式，这响应于所述第三数量的闲置活动模式与核跳跃不可行性相关联而进行。

22. 如权利要求 18 所述的方法，还包括响应于热密度状况生成所述核跳跃请求。

23. 如权利要求 18 所述的方法，其中所述驻留期阈值可由特权级实体动态地调整。

24. 一种包括程序代码的计算机可读介质，所述程序代码在由机器执行时将执行如权利要求 18 所述的方法。

25. 一种设备，包括用于执行如权利要求 18 所述的方法的部件。

26. 一种设备，包括：

多个核；以及

耦合到所述多个核的功率机构，所述功率机构用于接收使所述多个核中的核进入请求的功率状态的请求，预测预测的功率状态，并基于所述功率机构的预测精度为所述核选择所述预测的功率状态。

27. 如权利要求 26 所述的设备，其中所述功率机构基于所述功率机构的预测精度为所述核选择所述预测的状态包括所述功率机构响应于所述功率机构的预测精度超过精度阈值而为所述核选择所述预测的功率状态，并且其中所述功率机构还响应于所述功率机构的预测精度没有超过精度阈值而为所述核选择所述请求的功率状态。

28. 如权利要求 27 所述的设备，其中所述功率机构预测所述预测的功率状态包括所述功率机构基于在一定间隔内测量的闲置持续时间预测所述预测的功率状态。

29. 如权利要求 28 所述的设备，其中所述功率机构基于在所述间隔期间测量的闲置持续时间预测所述预测的功率状态包括：所述功率机构响应于在所述间隔内所述测量的闲置持续时间超过均衡持续时间而增加预测的闲置持续时间，并响应于在所述间隔内所述测量的闲置持续时间没有超过所述均衡持续时间而减小所述预测的闲置持续时间。

30. 如权利要求 29 所述的设备，其中所述功率机构还确定所述功率机构的预测精度。

31. 如权利要求 30 所述的设备，其中所述功率机构还确定所述功率机构的预测精度包括：所述功率机构响应于在所述间隔内所述测量的闲置持续时间超过所述均衡持续时间并且所述预测的闲置持续时间对应于长闲置持续时间而增加所述预测精度；响应于在所述间隔内所述测量的闲置持续时间超过所述均衡持续时间并且所述预测的闲置持续时间对应于短闲置持续时间而减小所述预测精度；响应于在所述间隔内所述测量的闲置持续时间没有超过所述均衡持续时间并且所述预测的闲置持续时间对应于所述短闲置持续时间而增加所述预测精度；响应于在所述间隔内所述测量的闲置持续时间没有超过所述均衡持续时间并且所述预测的闲置持续时间对应于所述长闲置持续时间而减小所述预测精度。

32. 如权利要求 26 所述的设备，其中所述多个核和所述功率机构包含在微处理器中，所述微处理器耦合到存储器，其中所述存储器选自由动态随机存取存储器 (DRAM)、双倍数据速率 (DDR) RAM 和静态随机存取存储器 (SRAM) 组成的组。

33. 一种设备，包括：

多个核；

耦合到所述多个核中的核的预测逻辑，所述预测逻辑用于至少基于在一定间隔内所述多个核中的核的闲置持续时间预测预测的低功率状态；

预测精度逻辑，用于确定所述预测逻辑的预测精度；以及

耦合到所述预测逻辑的功率控制逻辑,用于接收来自软件实体的请求的低功率状态以及来自所述预测逻辑的预测的低功率状态,其中所述功率控制逻辑响应于所述预测精度大于阈值精度而为所述核选择所述预测的低功率状态,并响应于所述预测精度不大于所述阈值精度而为所述核选择所述请求的低功率状态。

34. 如权利要求 33 所述的设备,其中用于至少基于核的闲置持续时间预测预测的低功率状态的所述预测逻辑包括 :

闲置持续时间逻辑,用于确定在所述间隔内所述核的闲置持续时间 ;

耦合到所述闲置持续时间逻辑的均衡逻辑,所述均衡逻辑响应于所述闲置持续时间超过均衡持续时间而确定所述闲置持续时间是长持续时间,并响应于所述闲置持续时间没有超过所述均衡持续时间而确定所述闲置持续时间是短持续时间 ; 以及

耦合到所述均衡逻辑的预测状态逻辑,基于所述均衡逻辑响应于所述闲置持续时间超过所述均衡持续时间确定所述闲置持续时间是长持续时间并响应于所述闲置持续时间没有超过所述均衡持续时间确定所述闲置持续时间是短持续时间,所述预测状态逻辑确定所述预测的低功率状态。

35. 如权利要求 34 所述的设备,其中所述闲置持续时间逻辑包括用于对在所述间隔内所述核的闲置持续时间计数的计数逻辑,并且所述均衡逻辑包括用于比较所述闲置持续时间和所述均衡持续时间的比较逻辑。

36. 如权利要求 35 所述的设备,其中所述预测状态逻辑包括 :

状态机,用于响应于所述均衡逻辑确定所述闲置持续时间是长持续时间而移到较长的预测闲置状态,并响应于所述均衡逻辑确定所述闲置持续时间是短持续时间而移到较短的预测闲置状态 ; 以及

存储元件,用于存储基于所述较长的预测闲置状态或所述较短的预测闲置状态的所述预测的低功率状态的表示。

37. 如权利要求 34 所述的设备,其中用于确定所述预测逻辑的预测精度的所述预测精度逻辑包括 :

间隔逻辑,用于确定包括所述间隔在内的间隔总数 ;

耦合到所述均衡逻辑和所述预测状态逻辑的精度确定逻辑,所述精度确定逻辑基于由所述预测状态逻辑之前在所述间隔总数内预测的之前预测的低功率状态结合均衡逻辑在所述间隔总数内确定长持续时间以及确定短持续时间而确定所述间隔总数中的准确间隔的数量 ; 以及

精度逻辑,用于基于所述间隔总数中的所述准确间隔的数量确定所述预测精度。

38. 如权利要求 37 所述的设备,其中所述间隔逻辑包括用于对所述间隔总数计数的计数逻辑,并且所述精度逻辑包括用于根据所述准确间隔的数量除以所述间隔总数确定所述预测精度的算术逻辑。

39. 如权利要求 38 所述的设备,其中所述精度确定逻辑包括 : 比较逻辑,用于将在所述间隔总数内的之前预测的低功率状态的表示与在所述间隔总数内的长持续时间或短持续时间的对应表示进行比较 ; 以及计数逻辑,用于对所述间隔总数中的所述准确间隔的数量进行计数。

40. 如权利要求 33 所述的设备,其中所述多个核和所述功率机构包含在微处理器中,

所述微处理器耦合到存储器，其中所述存储器选自由动态随机存取存储器 (DRAM)、双倍数据速率 (DDR) RAM 和静态随机存取存储器 (SRAM) 组成的组。

41. 一种方法，包括：

在处理器的预测硬件中确定核的预测的功率状态；

确定所述预测硬件的预测精度；

从软件实体接收为核请求的功率状态；

响应于所述预测精度小于阈值精度而为所述核选择所述请求的功率状态；以及

响应于所述预测精度不小于所述阈值精度而为所述核选择所述预测的功率状态。

42. 如权利要求 41 所述的方法，还包括：

确定在一定间隔内所述核的闲置持续时间；以及

确定在所述间隔内所述核的所述闲置持续时间是否超过均衡阈值。

43. 如权利要求 42 所述的方法，其中在所述处理器的预测硬件中确定所述核的预测的功率状态包括：

在所述预测硬件中，响应于确定在所述间隔内所述核的闲置持续时间超过所述均衡阈值而确定较长的预测闲置状态；

在所述预测硬件中，响应于确定在所述间隔内所述核的闲置持续时间超过所述均衡阈值而确定比之前预测的闲置状态短的预测闲置状态；以及

基于响应于确定在所述间隔内所述核的闲置持续时间超过所述均衡阈值的比之前预测的闲置状态长的预测闲置状态以及响应于确定在所述间隔内所述核的闲置持续时间超过所述均衡阈值的所述较短的预测闲置状态，为所述核确定所述预测的功率状态。

44. 如权利要求 43 所述的方法，还包括响应于确定在所述间隔内所述核的闲置持续时间而确定总间隔计数，其中确定所述预测硬件的预测精度包括：

响应于在所述预测硬件中响应于确定在所述间隔内所述核的闲置持续时间超过所述均衡阈值而确定较长的预测闲置状态，

响应于所述之前预测的闲置状态指示长预测闲置状态而将准确间隔计数递增，并且

响应于所述之前预测的闲置状态指示短预测闲置状态而将所述准确间隔计数递减；

响应于在所述预测硬件中响应于确定在所述间隔内所述核的闲置持续时间超过所述均衡阈值而确定较短的预测闲置状态，

响应于所述之前预测的闲置状态指示短预测闲置状态而将准确间隔计数递增，并且

响应于所述之前预测的闲置状态指示长预测闲置状态而将所述准确间隔计数递减；以及

基于所述准确间隔计数和所述总间隔计数确定所述预测精度。

45. 一种包括程序代码的计算机可读介质，所述程序代码在由机器执行时将执行如权利要求 41 所述的方法。

46. 一种设备，包括用于执行如权利要求 41 所述的方法的部件。

避免无效核跳跃和提供硬件辅助的低功率状态选择的机构

技术领域

[0001] 本发明涉及处理器执行领域,具体来说,涉及处理器操作的优化。

背景技术

[0002] 半导体处理和逻辑设计的发展允许在集成电路装置上存在越来越多的逻辑。因此,计算机系统配置已经从系统中的单个或多个集成电路演变成在各个集成电路上存在多个处理器管芯、多个核、多个硬件线程和多个逻辑处理器。处理器或集成电路通常包括单个物理处理器管芯,其中该处理器管芯可包括任何数量的核、硬件线程或逻辑处理器。

[0003] 集成电路上越来越多数量的处理元件(即,核、硬件线程和逻辑处理器)使得能够并行完成更多任务。但是,作为增加的处理功率的必然结果,热密度和泄漏功率的问题也随之扩大。因此,具有多个核的处理器可以采用称为核跳跃(即,至少一个核的体系结构状态/上下文移到另一个核)的热耗散技术。但是,将整个上下文从一个核移到另一个核所固有的是成本,即,浪费的执行周期、花费在移动上的能量以及冷高速缓存。然而,当前还没有关于在基于热密度的初始核跳跃决定之外决定何时核跳跃的明智决定。因此,可能会在不需要核跳跃(即,触发的热密度状况可自我减轻)或者核跳跃不可行(即,跳跃会导致相同或更差的热状况)时启动核跳跃。结果,存在这样一些情形,其中由于热状况而触发核跳跃,但有利的却是避免核跳跃。

[0004] 随着处理器的发热和功率问题不断升级,低功率状态的智能使用变得更加重要。目前,现今的特权级软件(即,操作系统)在请求转变到低功率状态时并不非常准确。因此,之前的软件可能会请求核从过深(即,消耗较低的功率,但与核在未来成为闲置的时间量相比,没有足够的苏醒时间)或过浅(即,当闲置时间量大于苏醒时间时,消耗较多的功率)进入到无效的特定低功率状态。

发明内容

[0005] 本发明涉及一种设备,包括:

[0006] 多个处理器核;

[0007] 预测机构,用于预测在未来间隔所述多个处理器核的未来活动;以及

[0008] 核跳跃机构,用于基于在所述未来间隔所述多个处理器核的所述未来活动确定核跳跃是否有效,并响应于基于所述多个处理器核的所述未来活动确定所述核跳跃不有效而禁止核跳跃事件。

[0009] 本发明涉及一种设备,包括处理器,所述处理器包括:

[0010] 多个核;

[0011] 耦合到所述多个核的核跳跃逻辑,用于触发核跳跃请求;

[0012] 耦合到所述多个核的预测逻辑,用于预测在下一个间隔所述多个核的活动;以及

[0013] 耦合到所述核跳跃逻辑和所述预测逻辑的核跳跃管理器逻辑,所述核跳跃管理器逻辑响应于所述核跳跃管理器确定在所述下一个间隔所述多个核的所述活动指示核跳跃

是无效的而拒绝所述核跳跃请求。

[0014] 本发明涉及一种方法，包括：

[0015] 预测在未来间隔期间处理器上的多个核的闲置 - 活动表示的预测的驻留期；

[0016] 接收核跳跃请求；

[0017] 确定在所述未来间隔期间所述多个核的所述闲置 - 活动表示的所述预测的驻留期是否大于驻留期阈值；

[0018] 响应于所述核跳跃请求，响应于确定在所述未来间隔期间所述闲置 - 活动表示的所述预测的驻留期大于所述驻留期阈值而执行核跳跃；以及

[0019] 响应于所述核跳跃请求，响应于确定在所述未来间隔期间所述闲置 - 活动表示的所述预测的驻留期不大于所述驻留期阈值而不执行所述核跳跃。

[0020] 本发明涉及一种包括程序代码的计算机可读介质，所述程序代码在由机器执行时将执行如上所述的方法。

[0021] 本发明涉及一种设备，包括用于执行如上所述的方法的部件。

[0022] 本发明涉及一种设备，包括：

[0023] 多个核；以及

[0024] 耦合到所述多个核的功率机构，所述功率机构用于接收使所述多个核中的核进入请求的功率状态的请求，预测预测的功率状态，并基于所述功率机构的预测精度为所述核选择所述预测的功率状态。

[0025] 本发明涉及一种设备，包括：

[0026] 多个核；

[0027] 耦合到所述多个核中的核的预测逻辑，所述预测逻辑用于至少基于在一定间隔内所述多个核中的核的闲置持续时间预测预测的低功率状态；

[0028] 预测精度逻辑，用于确定所述预测逻辑的预测精度；以及

[0029] 耦合到所述预测逻辑的功率控制逻辑，用于接收来自软件实体的请求的低功率状态以及来自所述预测逻辑的预测的低功率状态，其中所述功率控制逻辑响应于所述预测精度大于阈值精度而为所述核选择所述预测的低功率状态，并响应于所述预测精度不大于所述阈值精度而为所述核选择所述请求的低功率状态。

[0030] 本发明涉及一种方法，包括：

[0031] 在处理器的预测硬件中确定核的预测的功率状态；

[0032] 确定所述预测硬件的预测精度；

[0033] 从软件实体接收为核请求的功率状态；

[0034] 响应于所述预测精度小于阈值精度而为所述核选择所述请求的功率状态；以及

[0035] 响应于所述预测精度不小于所述阈值精度而为所述核选择所述预测的功率状态。

[0036] 本发明涉及一种包括程序代码的计算机可读介质，所述程序代码在由机器执行时将执行如上所述的方法。

[0037] 本发明涉及一种设备，包括用于执行如上所述的方法的部件。

附图说明

[0038] 举例说明本发明，而不是要本发明受限于附图中的各图。

- [0039] 图 1 示出包括多个处理元件的处理器的一个实施例。
- [0040] 图 2 示出包括多个处理元件的处理器的另一个实施例。
- [0041] 图 3 示出用于避免无效核跳跃 (inefficient core hopping) 的机构 (mechanism) 的一个实施例。
- [0042] 图 4 示出用于预测未来活动的图 3 中的预测机构的一个实施例。
- [0043] 图 5 示出利用闲置 - 活动模式的驻留期来预测未来活动的一个实施例。
- [0044] 图 6 示出预取用于避免无效核跳跃的方法的流程图的一个实施例。
- [0045] 图 7 示出提供硬件辅助的低功率状态选择的一个实施例。
- [0046] 图 8 示出提供硬件辅助的低功率状态选择的另一个实施例。
- [0047] 图 9a 示出用于提供硬件辅助的低功率状态选择的方法的流程图的实施例。
- [0048] 图 9b 示出用于图 8 中的预测闲置状态机的状态机的实施例。

具体实施方式

[0049] 在以下详细描述中, 阐述了众多具体细节, 例如用于闲置 - 活动预测的特定硬件结构 / 机构的实例、闲置持续时间测量、精度确定、特定处理器配置、特定核跳跃状况、特定低功率状态、特定处理器单元 / 逻辑、处理元件的特定实例等, 以便充分理解本发明。但是, 本领域技术人员将明白, 无需采用这些具体细节也可实现本发明。在其它情况下, 没有详细描述公知的组件或方法, 例如特定和备选的多核和多线程化处理器体系结构、用于所示模块 / 块的特定逻辑电路以及微处理器的特定操作细节, 以免不必要地使本发明晦涩难懂。

[0050] 本文描述的方法和设备是为了避免无效核跳跃并提供处理器中的硬件辅助的低功率状态选择。具体来说, 这些优化主要是基于处理器中的核的预测的未来活动或不活动而参考跳跃和功率状态选择进行论述的。实际上, 下文将参考图 2 简短地描述用于处理器核之间的通信的说明性环组织, 以便说明潜在的处理器上核配置的实施例。然而, 本文描述的设备和方法不限于此, 它们可通过对于集成电路的单独部分利用上下文跳跃或功率状态选择而在任何集成电路中实现。

多核处理器的实施例

[0052] 参考图 1, 示出包括多个核的实施例。在一个实施例中, 处理器 100 包括在核之间移动上下文的核跳跃硬件 (未示出)。在另一个实施例中, 处理器 100 包括用于将核单独地置于低功率状态的功率硬件 (未示出)。处理器 100 包括任何处理器, 例如微处理器、嵌入式处理器、数字信号处理器 (DSP)、网络处理器或用于执行代码的其它装置。如图所示, 处理器 100 包括多个处理元件。

[0053] 在一个实施例中, 处理元件是指能够保持处理器的状态 (例如, 执行状态或体系结构状态) 的线程单元、线程槽、处理单元、上下文、逻辑处理器、硬件线程、核和 / 或任何其它元件。换句话说, 在一个实施例中, 处理元件是指能够与诸如软件线程、操作系统、应用或其它代码的代码独立相关联的任何硬件。物理处理器通常是指潜在地包括任何数量的其它处理元件 (例如, 核或硬件线程) 的集成电路。

[0054] 核通常是指设置在集成电路上的能够维持独立体系结构状态的逻辑, 其中每个独立维持的体系结构状态与至少一些专用执行资源相关联。与核相比, 硬件线程通常是指设置在集成电路上的能够维持独立体系结构状态的任何逻辑, 其中这些独立维持的体系结构

状态共享对执行资源的访问。可见,当某些资源共享、而其它资源专用于某个体系结构状态时,硬件线程与核的命名之间的线重叠。然而,操作系统通常将核与硬件线程视为是单独的逻辑处理器,其中操作系统能够在每个逻辑处理器上单独地调度操作。

[0055] 如图 1 所示,物理处理器 100 包括两个核,即核 101 和核 102。这里,可以利用核跳跃来减轻处理器的一部分上的热状况。但是,从核 101 跳到 102 可潜在地在核 102 上造成在核 101 上存在的相同的热状况,同时带来核跳跃的成本。因此,在一个实施例中,处理器 100 包括可利用核跳跃的任何数量的核。此外,处理器 100 中所包含的功率管理硬件能够使各个单元和 / 或核处于低功率状态以节省功率。这里,在一个实施例中,处理器 100 提供硬件以帮助为处理器 100 的处理元件以及潜在地为下文中将更详细地示出和论述的各个模块 / 单元进行低功率状态选择。

[0056] 尽管处理器 100 可包括不对称核,即具有不同配置、功能单元和 / 或逻辑的核,但图中示出对称核。因此,下文将不再详细论述与核 101 等同示出的核 102,以免赘述。另外,核 101 包括两个硬件线程 101a 和 101b,而核 102 包括两个硬件线程 102a 和 102b。因此,诸如操作系统的软件实体潜在地将处理器 100 视为是 4 个独立的处理器,即能够同时执行 4 个软件线程的 4 个逻辑处理器或处理元件。在一个实施例中,上下文的跳跃是指核跳跃;但是,在其它实施例中,可独立于核跳跃或与核跳跃结合执行线程跳跃。

[0057] 这里,第一线程与体系结构状态寄存器 101a 相关联,第二线程与体系结构状态寄存器 101b 相关联,第三线程与体系结构状态寄存器 102a 相关联,而第四线程与体系结构状态寄存器 102b 相关联。如图所示,体系结构状态寄存器 101a 复制在体系结构状态寄存器 101b 中,因此能够为逻辑处理器 101a 和逻辑处理器 101b 存储各个体系结构状态 / 上下文。也可为线程 101a 和 101b 复制其它较小资源,例如指令指针和再命名分配器逻辑 130 中的再命名逻辑。一些资源可通过分区来共享,例如重排序 / 引退单元 135 中的重排序缓冲器、ILTB 120、加载 / 存储缓冲器和队列。其它资源可潜在地完全共享,例如通用内部寄存器、页表基址寄存器、低级数据高速缓存和数据 -TLB 115、执行单元 140、以及无序单元 135 的部分。

[0058] 处理器 100 通常包括可完全共享、通过分区共享、或供处理元件专用 / 专用于处理元件的其它资源。在图 1 中,示出具有处理器的说明性逻辑单元 / 资源的纯粹示范性的处理器的实施例。注意,处理器可包括或省略这些功能单元中的任何一个,并且可包括没有描绘的任何其它已知的功能单元、逻辑或固件。如图所示,处理器 100 包括用于预测将要执行 / 采用的分支的分支目标缓冲器 120 以及用于存储指令的地址翻译条目的指令 - 翻译缓冲器 (I-TLB) 120。

[0059] 处理器 100 还包括耦合到提取单元 120 以解码所提取的元素的解码模块 125。在一个实施例中,处理器 100 与定义 / 指定可在处理器 100 上执行的指令的指令集体系结构 (ISA) 相关联。这里,由 ISA 辨识的机器代码指令通常包括称为操作码的指令的一部分,其引用 / 指定将要执行的指令或操作。

[0060] 在一个实例中,分配器和再命名器块 130 包括分配器以用于预约资源,例如用于存储指令处理结果的寄存器文件。然而,线程 101a 和 101b 潜在地能够无序执行,其中分配器和再命名器块 130 还预约其它资源,例如用于跟踪指令结果的重排序缓冲器。单元 130 还可包括寄存器再命名器以用于将程序 / 指令引用寄存器再命名为位于处理器 100 内部的

其它寄存器。重排序 / 引退单元 135 包括用于支持无序执行以及稍后对无序执行的指令的有序引退的组件，例如上文提到的重排序缓冲器、加载缓冲器和存储缓冲器。

[0061] 在一个实施例中，调度器和执行单元块 140 包括用于在执行单元上调度指令 / 操作的调度器单元。例如，在具有可用的浮点执行单元的执行单元的端口上调度浮点指令。还包括与执行单元相关联的寄存器文件以存储信息指令处理结果。示范性执行单元包括浮点执行单元、整数执行单元、跳转执行单元、加载执行单元、存储执行单元和其它已知的执行单元。

[0062] 较低级数据高速缓存和数据翻译缓冲器 (D-TLB) 150 耦合到执行单元 140。该数据高速缓存用于存储最近使用 / 操作的元素，例如数据操作数，它们潜在地保持为存储器一致性状态。D-TLB 用于存储最近的虚拟 / 线性 - 物理地址翻译。作为一个特定实例，处理器可包括用于将物理存储器分解为多个虚拟页面的页表结构。

[0063] 如图所描绘，核 101 和 102 共享对较高级或更远离的高速缓存 110 的访问，高速缓存 110 用于缓存最近提取的元素。注意，较高级或更远离 (further-out) 是指进一步远离 (increasing or getting further way from) 执行单元的高速缓存级。在一个实施例中，较高级的高速缓存 110 是最后一级数据高速缓存（即，处理器 100 上的存储器层级中的最后一个高速缓存），例如第二或第三级数据高速缓存。但是，较高级的高速缓存 110 并不限于此，它可包括指令高速缓存或与其相关联。踪迹高速缓存（它是指令高速缓存的一种）可以改为耦合在解码器 125 之后以便存储最近解码的踪迹。

[0064] 注意，在所描绘的配置中，处理器 100 还包括用于与位于处理器 100 外部的装置通信的总线接口模块 105，这些装置可以是例如系统存储器 175、芯片组、北桥或其它集成电路。存储器 175 可专用于处理器 100 或与系统中的其它装置共享。存储器 175 的常见类型实例包括动态随机存取存储器 (DRAM)、静态 RAM (SRAM)、非易失性存储器 (NV 存储器) 和其它已知的存储装置。

[0065] 图 1 示出具有不同模块、单元和 / 或逻辑的表示的示范性处理器的抽象逻辑视图。但是，注意，利用本文描述的方法和设备的处理器无需包括所示的单元。并且，处理器可省略所示的一些或所有单元。为了说明不同配置的可能性，论述现在转到图 2，图 2 描绘包括具有用于互连多个核的环配置的处理器上存储器接口模块（即，核外模块）的处理器 200 的实施例。处理器 200 示为包括：物理分布式高速缓存；环互连；以及核、高速缓存和存储器控制器组件。但是，本描绘纯粹是说明性的，实现所描述的方法和设备的处理器可包括用于与外部装置通信的任何处理元件、任何高速缓存类型或级和 / 或存储器、前端总线或其它接口。

[0066] 在一个实施例中，缓存代理 221-224 均用于管理物理分布式高速缓存的片段。作为一个实例，诸如组件 221 的每个高速缓存组件用于为并置核管理高速缓存的片段 (slice)，其中并置核是指为了管理高速缓存的分布式片段而与高速缓存代理相关联的核。如图所描绘，高速缓存代理 221-224 称为高速缓存片段接口逻辑 (CSIL)；它们又可称为高速缓存组件、代理或与高速缓存或其片段接口的其它已知的逻辑、单元或模块。注意，高速缓存可以是任何高速缓存级；然而，对该示范性实施例，论述集中在由核 201-204 共享的最后一级高速缓存 (LLC) 上。

[0067] 与高速缓存代理在环互连 250 上处理业务并与高速缓存片段接口非常相似，核代

理 / 组件 211/214 也处理业务并分别与核 201–204 接口。如图所描绘,核代理 221–224 称为处理器核接口逻辑 (PCIL) ;它们又可称为核组件、代理或与处理元件接口的其它已知的逻辑、单元或模块。另外,环 250 示为包括与诸如存储器控制器 (IMC) 231 和图形处理器 (未示出) 的其它模块接口的存储器控制器接口逻辑 (MCIL) 230 和图形集线器 (GFX) 240。但是,环 250 可包括或省略上述模块中的任何一个,并且可包括未示出的其它已知的处理器模块。另外,可通过诸如点对点互连或多跳互连的其它已知的互连来连接类似的模块。

[0068] 避免无效核跳跃的实施例

[0069] 在一个实施例中,诸如图 1、图 2 中示出的处理器或未示出的其它处理器的处理器将避免无效核跳跃。转到图 3,示出能够避免这种无效核跳跃的处理器的实施例。在该图示中,核跳跃是指上下文或体系结构状态从诸如核 302 的核移到诸如核 303 的另一个核。核跳跃触发机构 310 基于任何已知的核跳跃状况触发核跳跃或核跳跃事件。例如,能够基于热密度状况触发核跳跃。换句话说,当处理器的一部分、或更具体来说是处理器的核过热时,那么在一些情况下,核跳跃能够减轻热密度状况,即将工作负荷并且相应地将热量分散到处理器的其它物理位置。

[0070] 作为一个特定的说明性实例,假设核 301 和 302 并置在处理器上,并且以满容量操作,从而导致在处理器的一部分上生成过多热量,即造成热密度状况。基于指示这种状况的任何已知的设备或方法,核跳跃触发逻辑 310 生成 / 触发核跳跃请求。响应该请求,核跳跃机构 / 管理器可执行核跳跃。在这种情况下,核跳跃机构 320 可启动将核 301 的体系结构状态迁移到诸如核 304 的另一个核,其中核 301 的体系结构状态可包括核 301 上的一个或多个或潜在地所有硬件线程上下文。因此,处理器的工作负荷分散在处理器上,由此也潜在地散布热量。

[0071] 但是,如上所述,只依据当前的热量信息执行核跳跃而没有关于未来活动预测的一些额外信息会潜在地导致允许无效核跳跃。作为一个实例,在以上情形中,在下一个间隔期间,核 301 打算闲置,这会导致热密度状况的自我减轻;闲置不会生成太多热量。然而,在没有预测的情况下,对于核跳跃会带来功率和浪费的执行时间方面的成本,这本来可以通过预测未来闲置来避免。在另一个实例中,所示 4 个核中的 3 个核 301、302 和 304 具有造成引起核跳跃的热密度状况的大小相当的工作负荷。但是,繁忙核的上下文跳到核 303 不会减轻热密度状况,因为核 303 靠近仍然繁忙的其它核。

[0072] 因此,在一个实施例中,预测机构 315 用于预测在未来间隔多个处理器核的未来活动。本文所指的间隔可指时间量、执行周期数或任何其它时间度量。例如,间隔可具有几个周期到数千微妙的范围。类似地,可利用任何已知的用于预测处理器、核、处理元件或其它执行引擎的活动的方法来预测在未来 / 下一个间隔期间处理器核 301–304 的活动。这种预测的广义的说明性实例包括:利用在前一间隔期间多个处理器核的过去活动作为在未来间隔多个处理器核的未来活动;从软件实体接收活动提示并基于活动提示预测未来活动;执行未来活动的算术或统计预测;利用诸如卡尔曼滤波算法的预测算法来预测未来活动;以及利用过去的闲置、低功率状态驻留期 (residency) 来预测未来活动。

[0073] 参考图 4,示出预测未来活动和闲置状态驻留期的实施例。在一个实施例中,包含在预测机构 315 中的预测器 435 可利用卡尔曼滤波算法。基于来自核 301–30N 的激活 (有时称为工作或 C0) 状态和低功率 (有时称为非工作或 C2-CN) 状态驻留期的数据,计算预

测。注意，在本论述中，特别参考特定的低功率状态，例如 C- 状态，以便进一步论述。但是，利用闲置和 / 或激活驻留期可与任何形式的功率状态、执行状态、全局状态或其它形式的活动或非活动指示符相关联。

[0074] C- 状态最一般是指由高级配置和电源接口 (ACPI) 定义的功率状态。根据 ACPI 的处理器或核功率状态的类型实例包括：C0 是操作 / 激活状态类型；C1 通常称为暂停 (Halt)，它是其中处理器不执行指令但可基本上瞬间地返回到执行状态的状态类型；C2 通常称为停止时钟 (Stop-Clock)，它是其中处理器保持所有软件可见的状态但可能要花费较长时间苏醒的状态类型；以及 C3 通常称为睡眠 (Sleep)，它是其中处理器无需使它的高速缓存保持一致、而是维持其它状态的状态类型。一些处理器对 C3 具有变型，即深睡 (C4)、更深睡 (C5)、最深睡 (C6)，它们的不同之处在于唤醒处理器的时间和功率节省的量不同。注意，本文描述的 C 状态类型中的一些是针对示范性处理器的，例如可从 CA, Santa Clara 的 Intel 公司获得的高级 Intel® Architecture 32 (IA-32) 处理器；但实施例同样可与其它处理器一起使用。这里，Intel® 处理器的 C 状态可能与 ACPI 规范不直接相符；但是，对于本论述，这两个激活 / 失活状态集合同样适用。换句话说，处理器可将任何数量的它们自己的定制 C 状态映射到上文描述的 ACPI 类型的 C 状态。

[0075] 继续以上论述，一旦预测器 435 基于来自活动监视器 430 的数据利用卡尔曼滤波算法计算了预测，核跳跃管理器 440 便能够基于预测决定对于是否要避免核跳跃做出决定。因此，活动监视器 430 可监视数据以确定激活或闲置信息，如上所述。可通过监视器 430 监视的与这些功率状态相关联的数据的实例包括：功率状态进入 / 退出事件，例如在监视器时期或间隔期间核 301-30N 的进入 / 退出事件；以及基于进入 / 退出事件或其它计数计算在这些状态的驻留期。由此，计算相同封装中的激活 - 闲置核的重叠。在一个实施例中，活动监视和预测的间隔可以是每隔 500 μ s；但是，这可容易地在几个执行周期到数千微秒的范围内变动。

[0076] 因此，如上所述，活动监视器 430 可从各个核 301-30N 接收关于它们的当前或过去的活动级的输入数据。在一个实施例中，活动监视器 430 包括可以按各种方式布置的缓冲器。例如，缓冲器可适于为每个核 301-30N 存储与每个功率状态改变事件相关联的时间标记的指示。这里，活动监视器 430 因而拦截 CPU 核 301-30N 进入和退出功率状态的事件并对这些事件加时间标记。在一个实施例中，将记录存储在内核缓冲器中。然后，按预定间隔，例如如上所述的间隔，活动监视器 430 将监视数据提供给预测器 435。因而，这种监视数据可包括时间标记数据以及活动状态以指示在存储间隔期间每个核处于给定状态的时间。

[0077] 作为响应，预测器 435 可利用该信息生成在下一个间隔所预测的核状态的模式分布。在一个实施例中，预测器 435 可执行诸如卡尔曼滤波算法的给定预测算法以生成该模式分布，但本发明在这方面不受限制。此外，应了解，取决于所支持的低功率状态的数量以及给定的核数、预测周期的长度等，模式分布可广泛地改变。为了易于论述，本文将描述包括三种不同模式（例如，下文将参考图 5 更详细论述的三种聚合模式）的模式分布。但是，应了解，本发明的范围在这方面不受限制，并且在不同实施例中，可以提供更多或更少的这样的模式，例如在给定活动级关于核数可以有变化的粒度。

[0078] 因此，将这种模式分布信息从预测器 435 提供给核跳跃管理器 440，核跳跃管理器 440 可基于模式分布允许或禁止核跳跃请求。作为一个实例，下一个间隔的特定聚合闲

置 - 活动模式可指示核跳跃将是有效的。这里, 响应于聚合模式分布和 / 或驻留期高于阈值驻留期, 核跳跃管理器 440 允许核跳跃。换句话说, 如果预测到核 301-30N 在某个激活 - 闲置模式驻留足够长的时间以便可以将核跳跃视为是可行且有效的, 则允许核跳跃。相反, 如果如有效模式的驻留期低于阈值所指示, 预测到驻留在不有效或不可行的其它闲置 - 活动模式而不是有效模式, 则核跳跃管理器 440 禁止或拒绝核跳跃请求。因此, 在这个实例中, 基于核 301-30N 的活动的未来预测, 可避免不必要的、无效的或不可行的核跳跃。

[0079] 在一个实施例中, 可计算以下三种模式以便预测活动:(1) 模式 A : 封装闲置 (内部的所有核都闲置); (2) 模式 B : 封装繁忙 (内部的所有核都繁忙); 以及 (3) 模式 C : 封装部分闲置 (其余情况是, 至少一个核繁忙, 同时至少一个核闲置)。这第三种模式描绘了闲置 / 繁忙重叠的情况。从以上实例可见, 通过活动监视器 430 可使核 301-30N 的进入 / 退出功率状态时间标记可用。因此, 可计算这三种模式分布预测。表 1 中示出根据本发明一个实施例的预测的示范输出, 其中假设间隔周期为 500 μ s。

[0080] 表 1

[0081]

下一个 T1 中模式 A%	25% (封装在模式 A : 25% *500 = 125 μ s)
下一个 T1 中模式 B%	15% (封装在模式 B : 15% *500 = 75 μ s)
下一个 T1 中模式 C%	60% (封装在模式 C : 60% *500 = 300 μ s)

[0082] 因此, 如表 1 所示, 对于间隔周期 T1, 预测闲置封装模式为 25% 的时间 (即, 125 μ s), 而预测所有核在 15% 的时间 (即, 75 μ s) 为激活, 并且在下一个预测周期的其余 60% 期间, 至少一个核激活并且至少一个核闲置 (即, 300 μ s)。在不同实施例中, 生成这些模式预测的方式可以改变。在一个实施例中, 预测可利用卡尔曼滤波器, 这将在下文马上进一步论述; 但是, 其它实现也是可能的。

[0083] 卡尔曼滤波器模型 (KFM) 利用经受高斯噪声的线性动态和线性观测来对部分观测的随机过程建模。它是根据一系列不完全且有噪声的测量估计动态系统的状态的有效递归滤波器。基于 KFM, 可以用多个预定模式 (例如, 模式 A、B 和 C 以及诸如下文参考图 5 描述的模式的其它模式的百分比 / 驻留期) 来阐明核 301-30N 的活动, 这些预定模式被视为是在时域离散化的实数随机过程的观测, 用 $y_{1:t} = (y_1 \dots y_t)$ 指示。过程的隐藏状态 $x_{1:t} = (x_1 \dots x_t)$ 又表示为实数的向量。KFM 中的线性随机差分方程为:

[0084] $x(t) = Ax(t-1) + w(t-1) p(w) \sim N(0, Q) x(0) \sim N(x_{1|0}, V_{1|0})$ [方程 1]

[0085] 并且, 测量方程为:

[0086] $y(t) = Cx(t) + v(t) p(v) \sim N(0, R)$ [方程 2]

[0087] 差分方程 1 中的 $n \times n$ 转换矩阵 A 使之前的 t-1 时间步长的状态与当前步长 t 的状态相关, 其中缺少驱动函数或过程噪声。这里, n 是隐藏状态的数量。在我们的任务中, m = n 是可能的 CPU 活动状态的数量。 $x_{1|0}, V_{1|0}$ 是状态的初始均值和方差, Q 是转换动态噪声的系统协方差, 而 R 是观测噪声的观测协方差。观测函数的转换一直相同, 并且该模型据称是非时变性的或是齐次的。

[0088] 利用 KFM, 给定直到当前时间的所有观测, 可预测未来时间的值。但是, 我们一般对

于未来不确定，并且因而计算最佳猜测以及置信度。因此，计算在可能的未来观测上的概率分布，用 $P(Y_{t+h} = y | y_{1:t})$ 表示，其中 $k > 0$ 是时域 (horizon)，即，距离未来多远来预测。

[0089] 给定观测值的序列 $(y_1 \dots y_t)$ ，为了预测新的观测值，对于到未来的一定时域 $k > 0$ ，计算 $P(Y_{t+h} = y | y_{1:t})$ 。方程 3 是通过忽略未来隐藏状态的预测而对关于未来观测的预测的计算。

$$[0090] P(Y_{t+h} = y | y_{1:t}) = \sum_x P(Y_{t+h} = y | X_{t+h} = x) P(X_{t+h} = x | y_{1:t}) \quad [\text{方程 3}]$$

[0091] 在方程的右边，通过固定滞后平滑算法（即， $P(X_{t-L} = x | y_{1:t})$ ， $L > 0$ ， L 是滞后）来计算 $P(X_{t+h} = x | y_{1:t})$ 。因此，在研究算法的细节之前，首先介绍 KFM 中的固定滞后平滑。

[0092] 固定滞后卡尔曼平滑器 (FLKS) 是执行回溯性数据同化的方法。给定直到当前时间的所有证据，它估计过去的状态，即， $P(X_{t-L} = x | y_{1:t})$ ， $L > 0$ ，其中 L 是滞后，例如，给定当前的传感器读数，我们可能想要算出管道是否在 L 分钟之前中断。这在传统上称为“固定滞后平滑”，但术语“后知 (hindsight)”可能更合适。在离线情形中，这称为（固定间隔）平滑；这对应于计算 $P(X_{T-L} = x | y_{1:T})$ ， $T \geq L \geq 1$ 。

[0093] 在预测算法中，要多进行正向和反向 h 遍。这些遍的计算类似于平滑过程中的计算。唯一的差别在于，在预测步骤中，新观测的初始值为空，这表示 $y_{1:T+h} = [y_{1:T} \ y_{null}^1 \dots y_{null}^h]$ 。预测算法通过对直到当前时间的所有证据加上 $y_{1:T+h} = [y_{1:T} \ y_{null}^1 \dots y_{null}^h]$ 执行回溯性数据同化来估计 $y_{1:T+h} = [y_{1:T} \ y_{T+1} \dots y_{T+h}]$ 的值。实际上，我们考虑利用之前的步长作为先前数据，例如，如果 $h = 1$ ，则 $y_{T+1} = (y_{T-1} + y_T)/2$ ，而不是 $y_{T+1} = null$ 。

[0094] 表 2 示出预测算法的伪代码。

[0095] 表 2

<pre> function Predicting ($y_{1:T}, x_{1:0}, V_{1:0}, A, C, Q, R$) $x_{0:0} = x_{1:0}$ $V_{0:0} = V_{1:0}$ $y_{T+1} = (y_{T-1} + y_T)/2;$ for $t = 1:T+1$ $(x_{t:t}, V_{t:t}, L_t) = \text{Fwd}(y_t, x_{t-1:t-1}, V_{t-1:t-1}, A, C, Q, R)$ end for $t = T:-1:1$ $(x_{t:t}, V_{t:t}, V_{t-1:t}) = \text{Back}(x_{t+1:T}, V_{t+1:T}, x_{t:t}, V_{t:t}, A, Q)$ end </pre>

[0097] 在表 2 中，Fwd 和 Back 是抽象的运算符。对于第一循环 ($\text{for } t = 1:T$) 的每个 Fwd (正向遍) 操作，首先通过 $x_{t|t-1} = Ax_{t-1|t-1}$ 和 $V_{t|t-1} = AV_{t-1|t-1}A' + Q$ 计算推理均值和方差；然后分别通过 $\text{err}_t = y_t - Cx_{t|t-1}$ 、 $S_t = CV_{t|t-1}C' + R$ 、 $K_t = V_{t|t-1}C'S_t^{-1}$ 和 $L_t = \log(N(\text{err}_t; 0, S_t))$ 计算推理 (创新) 误差、误差的方差、卡尔曼增益矩阵和这个观测的条件对数似然性；最

后通过 $x_{t|t} = x_{t|t-1} + K_t \text{err}_t$ 和 $V_{t|t} = V_{t|t-1} - K_t S_t K_t'$ 更新均值和方差的估计值。

[0098] 对于第二循环 (for $t = T-1 : -1 : 1$) 的每个 Back (反向遍) 操作, 首先, 通过 $x_{t+1|t} = Ax_{t|t}$ 和 $V_{t+1|t} = AV_{t|t}A' + Q$ 计算推理量; 然后, 通过 $J_t = V_{t|t} A' V_{t+1|t}^{-1}$ 计算平滑器增益矩阵; 最后, 分别通过称为 Rauch-Tung-Striebel (RTS) 方程的 $x_{t|T} = x_{t|t} + J_t (x_{t+1|T} - x_{t+1|t})$ 、 $V_{t|T} = V_{t|t} + J_t (V_{t+1|T} - V_{t+1|t}) J_t'$ 和 $V_{t-1,t|T} = J_{t-1} V_{t|T}$ 计算均值、方差和互方差的估计值。

[0099] 表 2 中所述的计算会比较复杂, 例如, 当计算 Fwd 运算符中的卡尔曼增益矩阵和 Back 运算符中的平滑器增益矩阵时, $T+1$ 步长循环中存在矩阵求逆。并且, 计算复杂度将为 $O(TN^3)$, 其中 T 是历史观测的数量, N 是活动状态的数量, 这是因为对于一般的 $N \times N$ 矩阵, 用于求解矩阵逆的高斯消元法会导致 $O(N^3)$ 复杂度。但是, 在各个实施例中, 算法实现可简化。在题为“Saving Power in A Computer System”且申请号为 11/955326 的同时待审的申请中描述了利用 KFM 来预测未来活动以进行功率状态选择的更具体的论述。

[0100] 现在参考图 5, 示出潜在活动模式的实施例, 这些潜在活动模式又可称为闲置 - 活动模式、预测活动、未来活动、下一个间隔活动、活动表示、闲置 - 活动表示等。这里, 将预测器 435 描绘为预测三种聚合闲置 - 活动模式 501、502 和 503。在该图示中, 逻辑 1 表示核是激活的, 即处于工作或 C0 状态; 而逻辑 0 表示核是闲置的, 即处于非工作或非 C0 功率状态。但是, “激活”和“闲置”的定义内可包含任何功率状态分组, 例如 C0 和 C1 视为是激活的, 而 C2-C6 视为是闲置 / 失活的。因此, 在所描绘的实例中, 可能的活动模式 505 的字段 505a-d 分别对应于核 301-304 的活动。这里, 每个字段可表示在下一个或未来间隔期间预测的活动或不活动, 这导致对于 4 个核有 16 种可能的闲置 - 活动模式。从这个简单的实例推测, 对于 N 个核潜在地存在 2^N 种模式。

[0101] 包括预测器 435 的预测机构可预测在未来 / 下一个间隔所有这 2^N 种可能模式的任何形式的驻留期, 然后将它们聚合为聚合模式, 如图所示。但是, 也可将这 2^N 种模式的任何子集预测为聚合模式。另外, 即使当预测聚合模式时, 也可预测每种可能模式的驻留期。此外, 模式的聚合可以按任何方式来执行, 以使得分组可不同于图 5 中的图示。而且, 模式的驻留期的表达式可以按任何形式进行, 例如驻留期的分布、一定间隔内驻留期的百分比、下一个间隔内驻留期的时间量、或预测一种模式在未来间隔内出现的时间或量的其它表达式。

[0102] 这里, 将核 301-304 的预测活动模式聚合为三个组: 501、502、503。尽管图中没有描绘, 但确定核跳跃是否有效 (是否允许) 的最基本的方法是将模式聚合为两个组: 对于核跳跃有效的活动模式; 以及对于核跳跃无效的活动模式。基本上, 如图所示, 在这个实例中, 聚合模式 502 包括对于核跳跃确定为“有效”的聚合模式, 而模式 503 和 501 包括确定为“无效”的模式。但是, 这个实例对无效进行了区分。模式 501 包括这样的模式, 其中核跳跃是可行的, 即可容易地执行, 但是却无需核跳跃。具体来说, 当对照时间代价、即履行核跳跃而停止执行的时间进行权衡时, 认为与模式 501 中所聚合的模式 (例如, 0000 或 0101) 相关联的热密度状况的增益没有益处。

[0103] 例如, 假设核 301-304 驻留在环拓扑中, 因此使得核 301 相邻于核 302 和 304。这里, 如果在当前活动包括 0101 时请求核 304 的核跳跃, 则上下文从核 304 跳到核 303 可临时缓解核 304 上的发热问题。但是, 从总体角度看, 热问题可能变得更糟, 因为活动模式现在是 0110, 其中核 302 和 303 设置在彼此靠近的位置而不是像 0101 中那样散布。在这种情

况下,不仅处理器封装的整体热状况会变更差,而且每个核基础上的发热状况刚好会转移到核 303。然而,在一些实现中,设计者可将这种模式视为是有效的,并且因此将它与其它有效模式聚合在一起。

[0104] 另一方面,模式 503 包括可得益于核跳跃的模式。但是,由于不可行或者由于与由核跳跃带来的时间、执行代价相比缺少实质的热状况减轻,所以核跳跃潜在地是无效的。例如,模式 1111 指示核 301-304 都是繁忙的。因此,从一个核跳到另一个核只是从一个激活核切换到另一个激活核。另外,在这个子集组中可包含诸如 1011 的模式,因为上下文从核 301 切换 / 跳到核 302 以获得模式 0111 可使启动热状况变得甚至更糟,如上所述。

[0105] 注意,可能的 2^N 种模式的这些分组、聚合或子集不限于基于潜在热状况减轻相对于执行的时间代价或多个时间代价的权衡分组。实际上,在一些情况中,核跳跃策略可纯粹基于可行性进行分组。在这种实施例中,如果核跳跃可行,则确定有效,然而如果不可行,则确定核跳跃无效。因此,预测机构可实现将闲置 - 活动模式分组为聚合模式的任何策略。此外,聚合可包括任何模式分组或关联。在一个实施例中,聚合包括将模式的驻留期聚合。例如,假设将驻留期(例如,聚合模式 502 的驻留期)表示为某个间隔的百分比,例如 70%,这可容易地转换为对于某个间隔的时间量($70\% * 500 \mu s = 350 \mu s$)。这实质上指示,预测器 435 预测,在未来 $500 \mu s$ 间隔期间,构成聚合模式 502 的某个版本的模式将驻留 $350 \mu s$ 或该间隔的 70%。

[0106] 为了进一步说明该实例,假设预测器 435 预测,模式 1100 将驻留下一个间隔的 10%,模式 0011 将驻留下一个间隔的 30%,而模式 1001 将驻留下一个间隔的 10%。在这个实例中,聚合简单地包括各个模式的总和,即 $10\% + 20\% + 30\% + 10\%$,从而获得 70% 的聚合驻留期。但是,如上所述,聚合可包括任何模式分组,或因此包括任何驻留期分组。注意,与预测的驻留期相比,各个聚合模式的实际驻留期可改变。但是,在模拟期间发现,诸如预测器 435 的预测器极其准确。实际上,一个模拟集合指示,在 4 核处理器上的预测器只有 4.34% 的相对误差。换句话说,如果预测器预测对于 $500 \mu s$ 间隔诸如 0101 的模式的驻留期为 $300 \mu s$,则实际的测量结果为该模式驻留介于 $287 \mu s$ 与 $313 \mu s$ 之间的时间。

[0107] 在一个实施例中,在间隔期间,核跳跃机构 320 或核跳跃管理器 440 基于在该间隔多个处理器核的未来 / 预测活动来确定核跳跃是否有效。注意,此确定的一部分或全部可在预测机构 315 并且潜在地可在其它逻辑中进行。继续以上实例,一旦预测器 435 预测了聚合模式 501、502 和 503 的驻留期,核跳跃管理器 440 便确定核跳跃是否有效。在一个实施例中,确定核跳跃有效包括确定是否预测到有效核跳跃模式的聚合组在下一个间隔驻留足够长的时间以视为是核跳跃有效。

[0108] 例如,确定有效核跳跃模式的聚合模式 502 的驻留期是否大于阈值,例如驻留期阈值。注意,阈值可包括对应于模式驻留期的表达式的任何百分比、时间、分布或其它值。为了便于说明,假设阈值是 60%。在以上实例中,聚合模式 502 的驻留期包括 70%,即,预测的有效闲置 - 活动模式超过阈值,在该实例中,预测该模式驻留的时间长于阈值。这里,因为确定有效,所以核跳跃管理器 440 允许核跳跃。但是,如果阈值为 75%,则由于预测到有效模式不会在间隔内驻留足够长的时间,所以核跳跃管理器 440 拒绝核跳跃。注意,在一些实施例中,驻留期阈值可预先确定,而在其它实施例中,它可通过特权级实体、非特权级实体或其组合动态地调整。

[0109] 在一个实施例中，核跳跃管理器 440 还可基于核跳跃请求确定核跳跃的最佳配置。例如，响应于接收到来自核 304 的跳跃请求，核跳跃管理器可轮询活动监视器 430 以确定当前的激活 - 闲置模式或利用间隔内的模式驻留期的预测来选择跳跃以实现最佳配置。这里，假设预测单个模式 0011 在下一个间隔期间具有最长驻留期。然后，响应于核跳跃请求，核管理器可确定 304 的上下文将跳到核 301 而不是 302，因为在间隔期间在最长驻留的模式 0011 中，设置在彼此靠近的位置的核 302 和 303 将同时激活。确定最佳配置不限于基于单个模式。而是，可使用组合或聚合。例如，可将指示特定核激活的预测模式加以组合，并且评估该组合的驻留期以确定最佳核跳跃。

[0110] 还应注意，核跳跃机构 320 的确定可以不是在活动模式上进行的，而是在一些实施例中，可在任何未来活动预测上进行，未来活动预测可用于指示核跳跃是否有效。实际上，上文在图 4 的论述之前已经描述了未来活动预测的众多实例。因此，可以利用任何已知的方法或设备来确定核跳跃是否有效。此外，重要的是要再次注意，所描绘的聚合模式的子集分组的数量和内容纯粹是说明性的。

[0111] 论述转到图 6，描绘用于避免无效核跳跃的方法的实施例。尽管以下图 6 和图 9a 的流程按照基本上串行的方式示出，但每个流程可至少部分地并行或按不同的顺序执行。此外，所示的一些流程可省略，而其它流程可包含在不同实施例中。例如，可在前一个间隔结束时预测下一个或未来间隔的驻留期。在模拟期间，预测平均值为约 6.5 微秒。因此，间隔可结束，预测可开始，并且核跳跃请求可全部部分地并行进行。因此，图 6 和图 9a 中示出的流程的顺序纯粹是说明性的。

[0112] 在流程 605 中，预测在未来间隔期间处理器上的多个核的闲置 - 活动表示的驻留期。如上所述，闲置 - 活动表示可包括处理元件活动的任何已知的表示。在一个实例中，闲置 - 活动表示包括多个有效闲置 - 活动模式的第一聚合闲置 - 活动模式。这里，N 个核的活动可由 2^N 种模式来表示，其中通过平台策略来确定有效闲置 - 活动模式是否有效，平台策略可以是例如权衡减轻热密度状态的核跳跃的增益和执行核跳跃的时间执行成本。图 5 中将这种有效闲置 - 活动模式的特定说明性实例描绘为是聚合模式 502。注意，如图 5 所示，可预测和计算多于一种聚合模式。实际上，在一个实施例中，可预测和聚合所有可能的 2^N 种模式的驻留期，而在其它实施例中，可只预测和计算 2^N 种模式和聚合模式的子集。

[0113] 在间隔期间，接收核跳跃请求。注意，本文不会详细论述核跳跃请求及其生成以免混淆对是否避免核跳跃请求的论述。但是，核跳跃请求可基于任何已知的状况生成，例如热密度状况、工作负荷状况、故障状况等。另外，任何实体都可生成核跳跃请求。例如，核跳跃请求可由诸如图 3 中的核跳跃触发逻辑 310 的硬件触发。核跳跃请求也可由诸如内核或 OS 的特权级软件实体生成。实际上，一些处理器设计者可选择对用户级代码开放请求机构，以便核跳跃触发逻辑 310 接收用户请求并生成核跳跃请求。

[0114] 在判定流程 615 中，确定有效闲置 - 活动表示的预测的驻留期是否大于阈值。例如，确定聚合模式 502 的驻留期是否大于驻留期阈值。注意，达到相同目的的此确定的备选方法包括执行对诸如模式 501 和 503 的无效模式的聚合的驻留期的预测。然后，确定该聚合是否低于阈值。基本上，不是确定有效核跳跃模式将驻留足够长的时间以使得核跳跃有效，而是后一个实例是确定无效核跳跃模式是否将不驻留足够长的时间以使核有效。

[0115] 响应于可包括预测的驻留期与驻留期阈值的某种形式的比较的确定，如果预测的

驻留期大于阈值,即核跳跃有效,则在流程 625 中允许该请求;或者如果预测的驻留期没有超过阈值,即核跳跃无效,则在流程 620 中拒绝 / 避免该请求。允许核跳跃可简单地包括让正常的核跳跃逻辑(例如,将一个核的状态迁移到另一个核的逻辑)执行它的任务,而拒绝可包括阻止核跳跃发生的任何形式的防碍、禁止、拒绝或其它操作。

[0116] 硬件辅助的低功率状态选择的实施例

[0117] 接下来转到图 7,描绘包括基于预测精度在请求的功率状态与预测的功率状态之间选择的功率机构的处理器的实施例。如图所示,功率机构 710 接收使某个核(例如,核 701-704 中的核 701)进入请求的功率状态的请求。从以上论述注意到,请求的功率状态可包括任何功率状态,例如 C- 状态、ACPI 功率状态或其它已知的功率状态。之前,功率控制单元仅仅基于软件请求 741 便利用请求的功率状态。因此,诸如操作系统(OS)、内核、管理程序或虚拟机监视器(VMM)的软件实体 740 为核 701 请求某个功率状态,并且相应地进入该功率状态。但是,与以上的核跳跃非常相似,在没有预测核 701 的闲置持续时间的情况下,核 701 可产生进入低功率状态的成本 / 支出,并且接着会在短时间内被再次唤醒。实际上,在这种情况下,核 701 可进入该低功率状态,并且驻留在此的时间少于将核 701 从这个功率状态唤醒所花费的时间量。作为额外的构成要素,题为“Saving Power in a Computer System”、申请号为 11/955326 的同时待审的申请论述了核在一定时间内驻留在一种功率状态时的功率状态的硬件升级 / 降级。注意,硬件升级和降级可结合本文描述的设备和方法一起使用。例如,硬件可预测所预测的闲置持续时间或所预测的功率状态,而软件可请求所请求的功率状态。如果硬件是准确的,那么选择预测的功率状态。但是,如果核驻留在所预测的功率状态的时间过长或者不够长,那么如同时待审的申请中所论述,可随后将该功率状态降级 / 升级。

[0118] 在一个实施例中,功率机构 710 提供硬件辅助的功率状态选择。实际上,作为潜在的优化,功率状态选择可基于预测精度,以使得当硬件预测精度较高时,选择硬件预测的功率状态。并且,当硬件预测精度较低时,选择软件请求的功率状态。因此,当硬件准确地预测了正确的、即对于未来间隔最有效的功率状态时,那么硬件的选择具有优先权,而当预测机构不是很准确时,硬件将优先权让给软件。实际上,在模拟期间发现,操作系统功率管理功率状态选择在 34.8% 的时间是正确的;纯硬件预测的实施例在 92.7% 的时间是正确的;并且基于硬件精度的硬件辅助的功率状态选择的混合的实施例在 94.3% 的时间实现了正确的功率状态。

[0119] 在一个实施例中,功率机构 710 包括用于预测闲置持续时间或所预测的功率状态的预测机构 720。注意,预测机构可预测闲置持续时间,例如非常长、长、短或非常短的持续时间,并且通过硬件、固件、软件或其组合使功率状态与其相关联。为了说明,可使非常长和长持续时间与最深睡眠状态 C6 相关联,而使短和非常短持续时间与睡眠状态 C3 相关联。在这个图示中,术语“预测诸如 C3 和 C6 的低功率状态”的使用与预测闲置持续时间的陈述同义,或反之亦然。

[0120] 在一个实施例中,预测机构 720 响应于所测量的闲置持续时间而增加或减少所预测的闲置持续时间。例如,在间隔期间,诸如图 4 中的监视器 430 的活动监视器利用计数器逻辑和 / 或时间标记来确定 / 测量核(例如,核 701)的闲置持续时间。基于该闲置持续时间,相应地更新未来 / 预测的闲置持续时间。作为说明性实例,响应于所测量的闲置持续时

间与均衡 (break even) 阈值的比较, 增加或减少所预测的闲置持续时间。如果软件实体 740 请求功率, 那么在一个实施例中, 均衡阈值包括最深 (例如, C6) 功率状态的均衡时间。注意, C6 的均衡时间可以预先确定。均衡时间也可包括从 C6 功率状态唤醒到激活状态的时间量以及与进入 / 退出功率状态相关联的任何其它时间度量。

[0121] 继续该实例, 如果在前一个间隔期间测量的闲置持续时间长于 C6 均衡时间, 则通过预测机构 720 增加所预测的闲置持续时间。或者, 如果在前一个间隔期间测量的闲置持续时间短于 C6 均衡时间, 那么通过预测机构 720 减少所预测的闲置持续时间。重要的是要注意, 功率状态的均衡时间是纯粹说明性的阈值。并且, 可以利用任何时间量或其它值来作为阈值以将它与来自间隔的测量 / 实际闲置持续时间进行比较。如上所述, 所预测的闲置持续时间可对应于功率状态。

[0122] 但是, 在一些实施例中, 所预测的闲置持续时间的更新不一定表示所预测的功率状态的改变。为了说明, 在以上实例中, 使长和非常长的闲置持续时间与 C6 相关联, 而使短和非常短的闲置持续时间与 C3 相关联。因此, 所预测的闲置持续时间从长增加到非常长仍表示所预测的功率状态为 C6。但是, 如果所测量的闲置持续时间低于均衡阈值, 那么所预测的闲置持续时间从长减小到短, 这表示所预测的功率状态从 C6 变为 C3。在后一种情形, 假设从软件实体 740 接收到使核 701 进入到诸如 C6 的低功率状态的请求。预测机构 720 已经完成测量在间隔期间核 701 的闲置持续时间。这里, 所测量的闲置持续时间低于 C6 均衡时间, 由此使预测机构 720 预测短的闲置持续时间。假设此时认为硬件预测是准确的, 那么功率管理器 715 便可基于硬件预测的短的闲置持续时间越过软件请求的 C6 状态而为核 701 选择硬件预测的状态 C3。

[0123] 然而, 在一个实施例中, 根据硬件预测机构 720 的预测精度来调节硬件预测的功率状态的选择。实质上, 在这个实施例中, 存在防护措施, 由此当硬件没有准确地预测时, 不会越过软件请求的功率状态而选择硬件预测的功率状态。因此, 在一个实施例中, 预测精度机构 725 用于预测预测机构 720 的预测精度。作为一个实例, 预测精度机构 725 响应于在间隔内测量的闲置持续时间超过均衡阈值 / 持续时间并且所预测的闲置持续时间对应于长闲置持续时间而增加预测精度。换句话说, 如果所预测的闲置持续时间是长闲置、即长于阈值持续时间并且实际测量的闲置持续时间也超过均衡阈值, 那么闲置预测是正确的。因此, 预测精度机构 725 基于所确定的准确预测增加精度。但是, 预测到长闲置持续时间并且测量到类似的长闲置、即超过均衡持续时间不是可确定为准确的唯一预测。作为另一个实例, 当预测到短闲置持续时间时测量到短闲置持续时间、即小于均衡时间也可增加预测精度。

[0124] 以类似方式, 预测精度机构 725 响应于在间隔内测量的闲置持续时间超过均衡持续时间并且所预测的闲置持续时间对应于短闲置持续时间而减小预测精度。此外, 预测精度机构 725 响应于在间隔内测量的闲置持续时间没有超过均衡持续时间、即短的测量闲置持续时间并且所预测的闲置持续时间对应于长闲置持续时间而减小预测精度。

[0125] 注意, 以上实例论及预测闲置持续时间, 并且论述了这些预测的精度的确定。但是, 预测不限于此。作为一个实例, 预测可包括预测实际功率状态或间隔内的功率状态驻留期。这里, 活动监视器可收集关于进入 / 退出这些功率状态的数据。并且, 将预测与数据进行比较以确定直接功率状态预测的精度。

[0126] 然而, 不管预测方法如何, 一旦进行了硬件预测, 功率管理器 715 便基于功率机构

720 的预测精度为核（例如，核 701）选择预测的状态或为核 701 选择请求的状态。在一个实施例中，提供精度阈值。这里，如果预测精度超过精度阈值，那么选择 / 利用硬件预测的功率状态。但是，如果预测精度没有超过阈值，那么选择请求的功率状态。

[0127] 接下来参考图 8，描绘用于预测闲置持续时间 / 功率状态的预测逻辑和用于确定预测精度的精度逻辑的特定说明性实施例。实际上，这里论述利用示范性数字和状态的简化的说明性实例以提供对图 8 中的逻辑的操作的看法。预测逻辑 720 至少基于在间隔内多个核中的一个核（例如，核 701）的闲置持续时间来预测所预测的低功率状态。这里，闲置持续时间逻辑 805 用于确定在间隔内该核的闲置持续时间。作为一个实例，间隔可以是定时间隔，例如 500 μ s 间隔。或者，间隔可包括闲置持续时间本身。在这个实施例中，当核（例如，核 701）进入低功率状态、即闲置时，诸如计数器的计数逻辑对核 701 的闲置持续时间计数。注意，也可利用如参考图 4 论述的利用时间标记的活动监视器来确定闲置持续时间。一旦核 701 苏醒，预测逻辑 720 便更新它的预测，并且预测精度逻辑 725 便相应地更新它的预测逻辑 720 的精度。响应于苏醒，闲置持续时间逻辑 805 将测量的闲置持续时间 806 提供给均衡逻辑 810。

[0128] 均衡逻辑 810 确定测量的闲置持续时间 806 是长持续时间还是短持续时间。如上所述，在一个实施例中，利用诸如 C6 状态的功率状态的均衡时间作为均衡阈值。在这个实例中，假设均衡阈值为 100 μ s，但也可利用任何均衡值。因此，如果闲置持续时间 806 超过 100 μ s，则确定核 701 的闲置持续时间为长持续时间。反之，如果闲置持续时间 806 为或低于 100 μ s，则确定闲置持续时间为短持续时间。作为一个简单的实例，均衡逻辑 810 可包括用于将闲置持续时间 806 与均衡持续时间 100 μ s 进行比较的比较逻辑。

[0129] 预测状态逻辑 815、820 基于均衡逻辑、即来自均衡逻辑 810 的长或短闲置确定而确定所预测的低功率状态。预测的闲置 / 功率状态元件 820 保存之前预测的功率状态。注意，元件 820 中可表示任何数量的功率状态和 / 或闲置持续时间。作为一个说明性实施例，存储元件 820 保存对应于功率状态的闲置状态的二位表示。这里，00 表示非常短的闲置持续时间，它对应于 C3 功率状态；01 表示短闲置持续时间，它也对应于 C3 功率状态；10 表示长闲置持续时间，它对应于 C6 功率状态；而 11 表示非常长的闲置持续时间，它也对应于 C6 功率状态。因此，当从软件实体 740 接收到对于核 701 的请求的功率状态时，将所预测的闲置持续时间 / 功率状态 822 作为硬件预测的功率状态提供给功率管理器 715。

[0130] 然而，继续以上实例，所预测的闲置功率状态 820 可由预测的闲置状态机构 815 基于长 / 短闲置持续时间信息 811 进行更新。快速参考图 9b，示出预测的闲置状态机 815 的流程图的实施例。可见，基于来自均衡逻辑 810 的长 / 短闲置确定，相应地更新状态机 815 的状态。例如，如果保存在元件 820 中的之前状态包括 01、即预测短闲置持续时间 816b 并且确定长闲置持续时间，那么长闲置持续时间将状态机移到状态 816c，即预测长闲置状态。这个结果可保存在状态机 815 中并直接提供给功率管理器 715。或者，如图所示，将结果存储在元件 820 中。一旦提供了经过更新的闲置持续时间预测 816，预测循环便基本上结束了。因此，响应于来自软件实体 740 的进入诸如 C3 的请求功率状态的请求，将新更新的预测闲置持续时间 / 功率状态、即对应于 C6 功率状态的长闲置持续时间提供给功率管理器 715。功率管理器 715 在请求的功率状态 C3 与预测的功率状态 C6 之间进行选择。

[0131] 在一个实施例中，这个选择可基于预测逻辑 720 的精度来进行。回到均衡逻辑

810, 将长 / 短闲置持续时间确定也提供给间隔精度确定逻辑。另外, 将在通过状态机 815 更新之前的保存在元件 820 中的之前状态也提供给逻辑 825。实质上, 采用比较测量的闲置和预测的闲置是否都是长闲置或短闲置的形式将之前预测的闲置持续时间与新测量的持续时间进行比较。换句话说, 逻辑 825 确定核 701 的闲置的预测是否正确。

[0132] 间隔精度确定逻辑提供精度确定 826。在一个实施例中, 如果预测是准确的, 那么信号 826 是将精度逻辑 830 递增的递增信号。类似地, 如果预测不正确, 那么信号 826 包括不递增信号。因此, 精度逻辑 830 实质上是对预测逻辑正确的闲置间隔的数量进行计数。此外, 对于每个闲置持续时间测量, 将总间隔跟踪逻辑 830 递增。因此, 核 701 的准确的闲置间隔的数量可在精度逻辑 830 中获得, 而总间隔数可在间隔跟踪逻辑 830 中获得。因此, 精度逻辑 830 可将准确间隔数除以总间隔数以获得诸如精度百分比的精度值 831。

[0133] 因此, 功率管理器 715 或精度逻辑 830 可将这个精度值与诸如 90% 的阈值进行比较。如果精度超过阈值 90%, 那么对于核 701 利用硬件预测的状态 C6。相反, 如果精度没有超过阈值 90%, 那么对于核 701 利用请求的功率状态。注意, 在另一个实例中, 精度逻辑可容易地颠倒, 其中对于不准确的闲置间隔将精度逻辑 830 中的计数器递增并除以总间隔数以获得不准确度值。然后, 可将这个值与诸如 10% 的不准确度阈值进行比较。在操作中, 使用基本上相同。注意, 在参考图 8 的实例中所提供的精度阈值以及其余的数量和状态纯粹是说明性的。另外, 诸如均衡阈值或精度阈值的阈值可通过硬件、特权软件、用户软件或其组合来动态地调整。

[0134] 转到图 9a-9b, 示出用于预测闲置持续时间和确定该预测的精度的方法的实施例。在流程 901 中, 核从闲置或非激活状态苏醒 / 退出。注意, 该核可以是苏醒的整个处理器或一组核的一部分。在闲置期间, 在流程 905 中测量 / 确定闲置持续时间。例如, 计数器对闲置的持续时间计数。

[0135] 在流程 910 中, 确定总的闲置计数。作为一个实例, 对于每个闲置间隔, 将总的闲置计数递增。在流程 915 中, 确定闲置持续时间是否大于均衡阈值, 例如某个功率状态的均衡时间。如果闲置持续时间不大于均衡阈值, 那么在流程 920 中确定该闲置持续时间为短。类似地, 如果闲置持续时间大于均衡阈值, 那么在流程 925 中确定闲置持续时间为长。再次快速地参考图 9b, 示出基于流程 920、925 中的确定而在状态之间移动的状态机。例如, 如果如流程 925 所示, 闲置持续时间为长, 那么状态机在状态 816a-816d 中朝向更长的闲置状态移动。反之, 短持续时间确定 920 将状态朝向更短的闲置状态、例如朝向状态 816a 移动。

[0136] 此外, 在流程 930-945 中, 确定预测硬件的预测精度。这里, 如果在流程 920 中确定闲置持续时间为短, 那么从流程 930 有两种可能的精度结果:(1) 之前预测的闲置状态是正确的, 即预测的闲置持续时间为状态 816a 或 816b, 它们表示短闲置持续时间; 或 (2) 之前预测的闲置状态是不正确的, 即预测的闲置持续时间为状态 816c 或 816d, 它们表示长闲置持续时间。类似地, 在流程 935 中, 有两种类似的结果:(1) 之前预测的闲置状态是正确的, 即预测的闲置持续时间为状态 816c 或 816d, 它们表示长闲置持续时间; 或 (2) 之前预测的闲置状态是不正确的, 即预测的闲置持续时间为状态 816a 或 816b, 它们表示短闲置持续时间。因此, 当来自任一路径的预测是正确时, 在流程 940 中将精度计数递增。如果预测不正确, 那么流程移到流程 945, 它通过将精度计数除以来自流程 910 的总闲置计数来确定精度值。因此, 如果精度大于阈值精度, 那么对核 701 使用硬件预测的功率状态。反之, 如

果精度不大于阈值精度,那么对核 701 使用请求的功率状态。软件请求和准确的硬件预测相结合潜在地导致非常准确的功率状态选择,例如如所模拟的超过 94% 的正确选择。

[0137] 本文所用的模块是指任何硬件、软件、固件或其组合。通常,示为独立的模块边界通常会改变并且潜在地会重叠。例如,第一和第二模块可共享硬件、软件、固件或其组合,同时潜在地保留一定独立的硬件、软件或固件。在一个实施例中,术语“逻辑”的使用包括硬件,例如晶体管、寄存器或诸如可编程逻辑器件的其它硬件。但是,在另一个实施例中,逻辑也包括与硬件集成的软件或代码,例如固件或微代码。

[0138] 本文所用的值包括数量、状态、逻辑状态或二进制逻辑状态的任何已知的表示。通常,逻辑级、逻辑值 (logic/logical value) 的使用又称为 1 和 0,它们简单地表示二进制逻辑状态。例如,1 表示高逻辑级,而 0 表示低逻辑级。在一个实施例中,诸如晶体管或闪速单元的存储单元能够保存单个逻辑值或多个逻辑值。但是,使用了计算机系统中的其它值表示。例如,十进制数字“十”也可表示为二进制值 1010 和十六进制字母 A。因此,值包括能够保存在计算机系统中的任何信息表示。

[0139] 此外,状态可由值或值的部分来表示。作为一个实例,诸如逻辑 1 的第一个值可表示默认或初始状态,而诸如逻辑 0 的第二个值可表示非默认状态。另外,在一个实施例中,术语“重设”和“设置”分别指默认和更新后的值或状态。例如,默认值潜在地包括高逻辑值,即重设,而更新后的值潜在地包括低逻辑值,即设置。注意,可利用任何值组合来表示任何数量的状态。

[0140] 以上阐述的方法、硬件、软件、固件或代码的实施例可经由存储在机器可访问或机器可读介质上的可通过处理元件执行的指令或代码来实现。机器可访问 / 可读介质包括提供 (即,存储和 / 或传送) 可由诸如计算机或电子系统的机器读取的形式的信息的任何机构。例如,机器可访问介质包括 :随机存取存储器 (RAM), 例如静态 RAM (SRAM) 或动态 RAM (DRAM); ROM; 磁或光存储介质; 闪速存储器装置; 电存储装置、光存储装置、声存储装置或其它形式的传播信号 (例如,载波、红外信号、数字信号) 存储装置等。例如,机器可通过从介质接收诸如载波的传播信号来访问存储装置,该介质能够保存待在该传播信号上传送的信息。

[0141] 整篇说明书中提到“一个实施例”或“实施例”时表示,结合这个实施例描述的特定特征、结构或特性包含在本发明的至少一个实施例中。因此,在整篇说明书中的各个地方出现短语“在一个实施例中”或“在实施例中”时不一定都指相同的实施例。此外,这些特定特征、结构或特性可以在一个或多个实施例中以任何合适的方式加以组合。

[0142] 在以上说明书中,参考特定示范性实施例给出了详细描述。但是,很明显,在不偏离如随附权利要求所述的本发明的更宽广的精神和范围的情况下,可对此作出各种修改和改变。因此,应将说明书和附图视为是说明性而不是限制性的。此外,实施例和其它示范性语言的上述使用不一定指相同的实施例或相同的实例,而是可以指截然不同的实施例,也可以潜在地指相同的实施例。

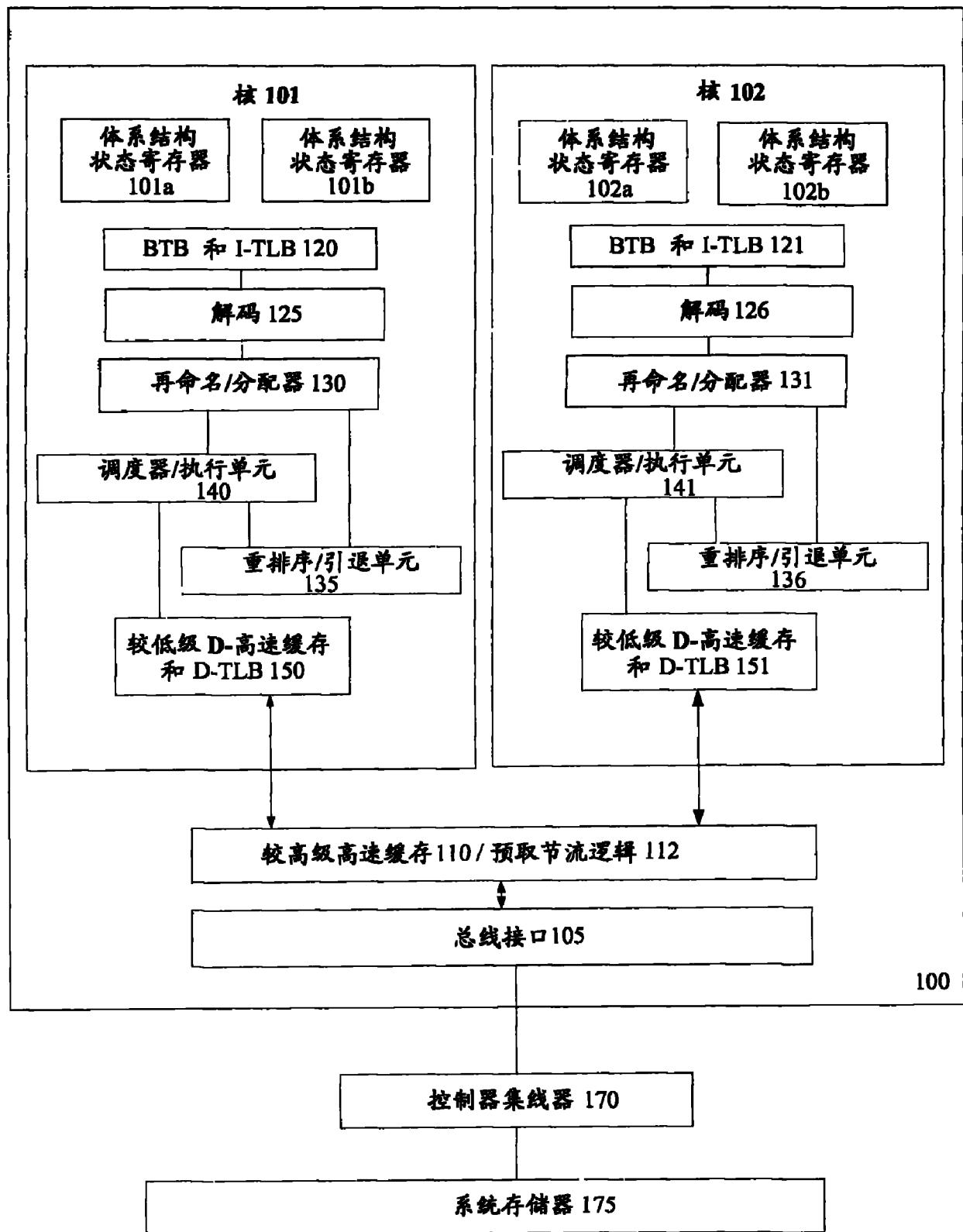


图 1

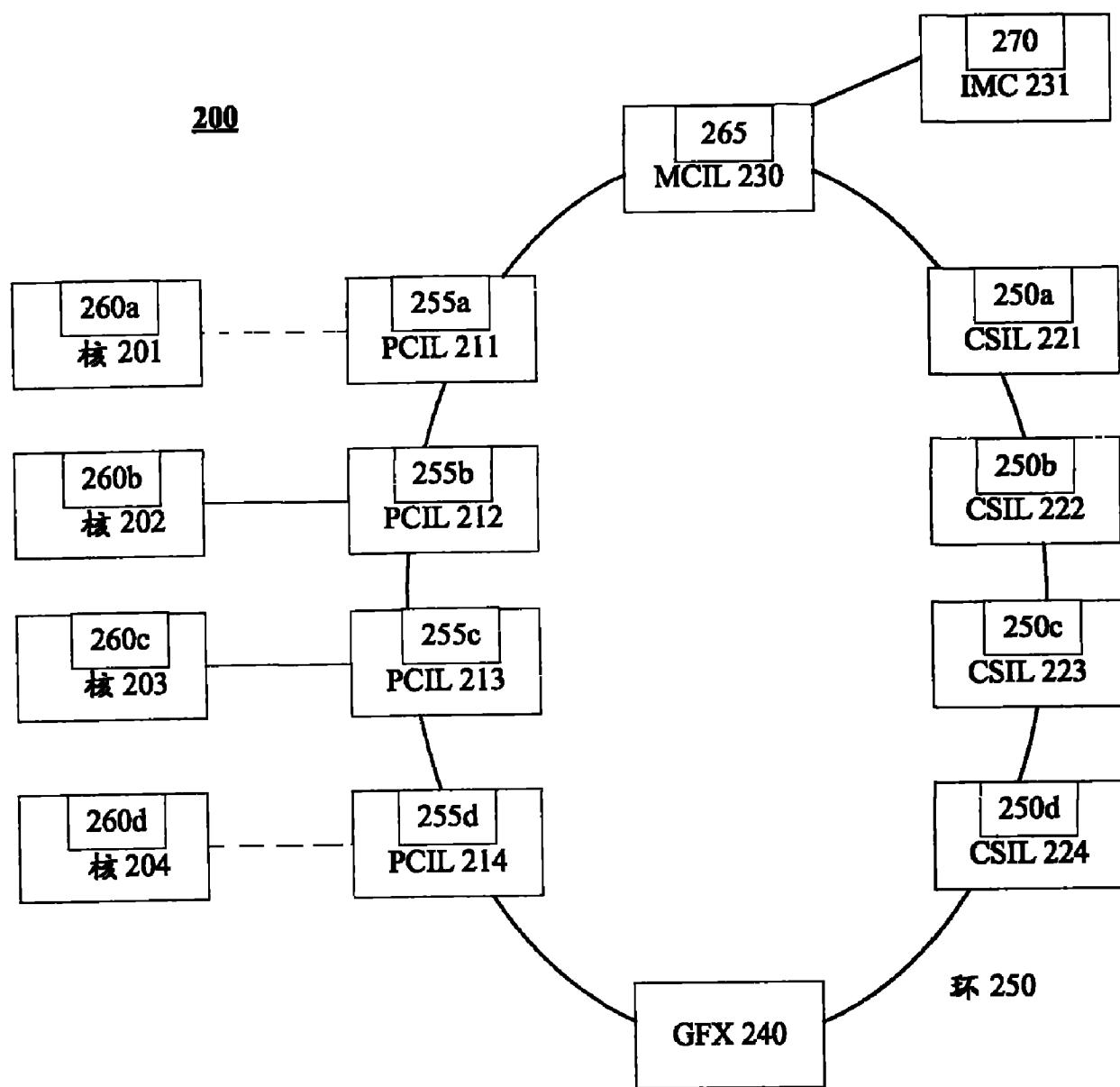


图 2

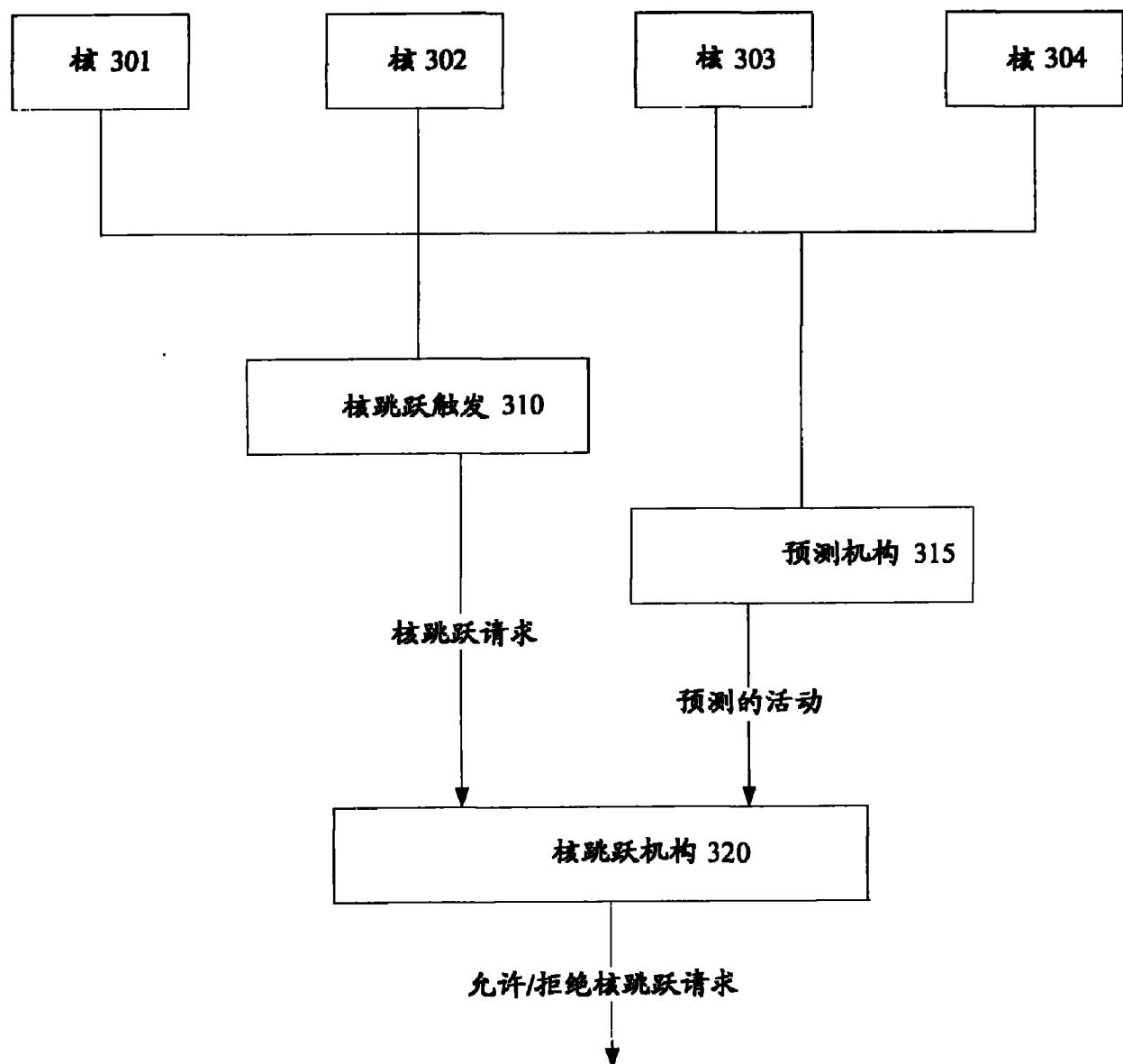


图 3

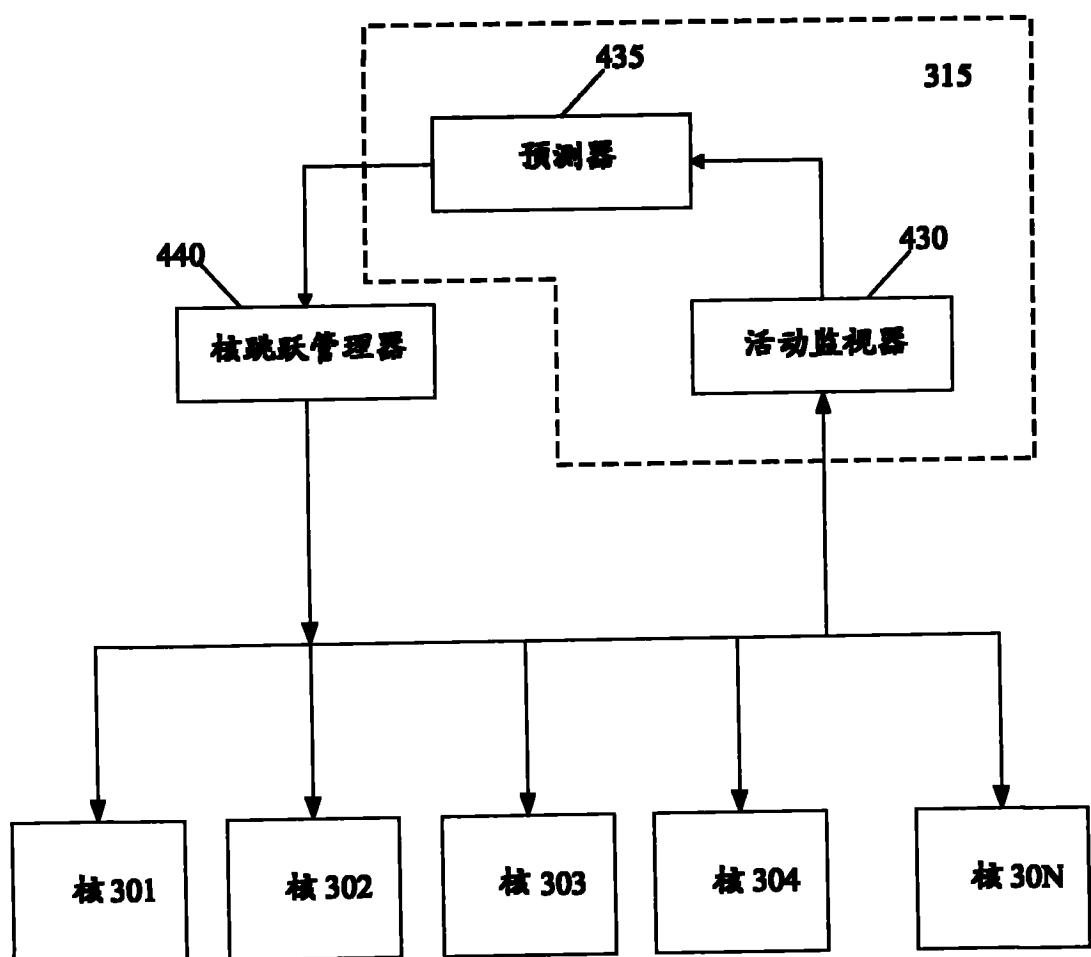
410

图 4

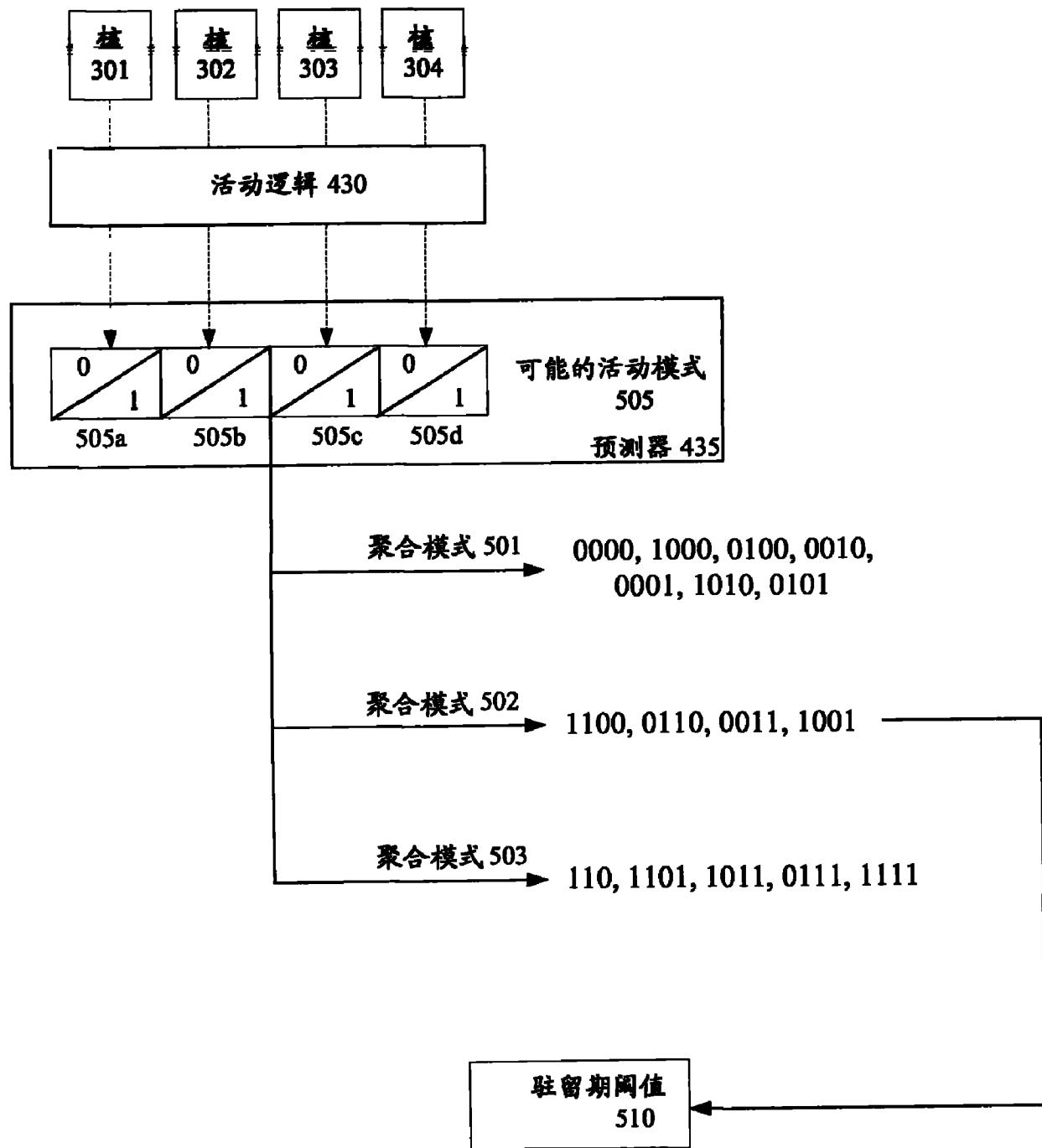


图 5

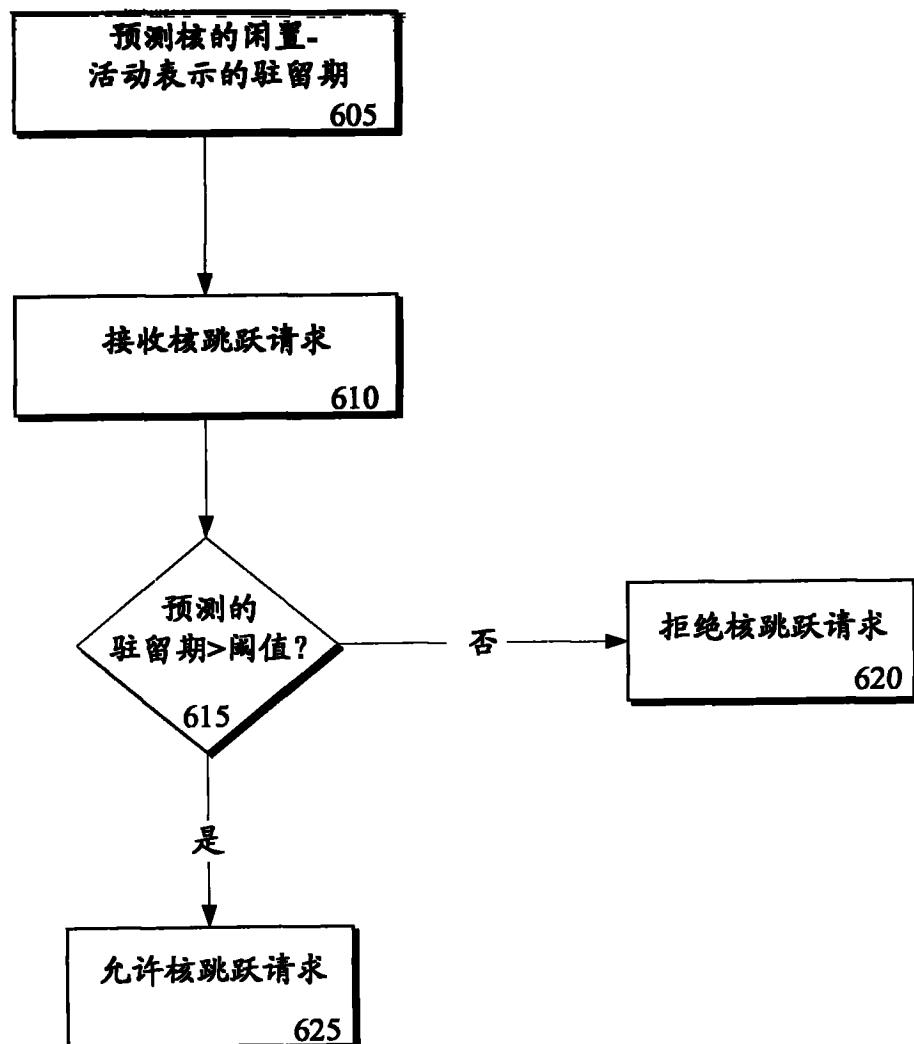


图 6

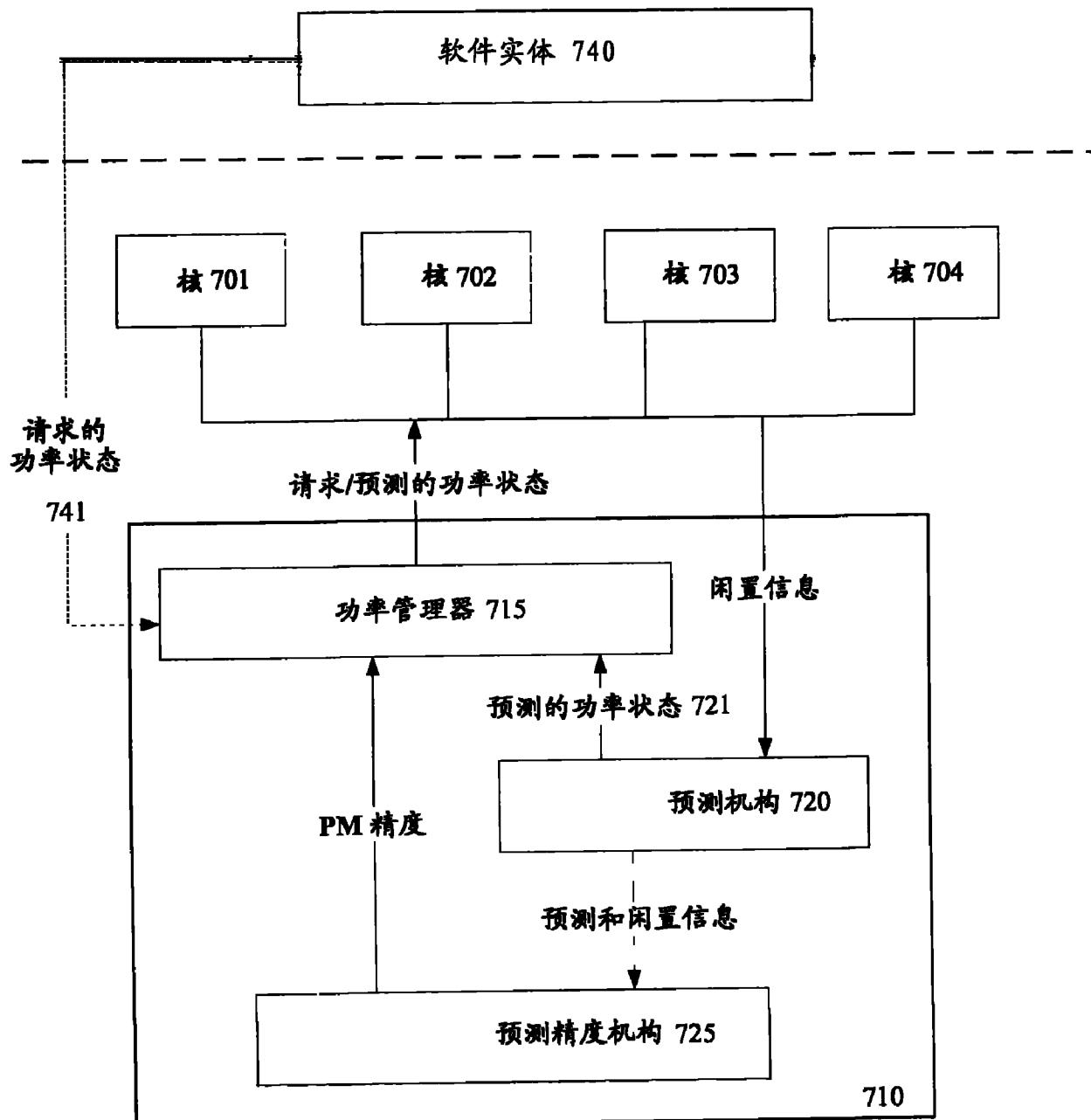


图 7

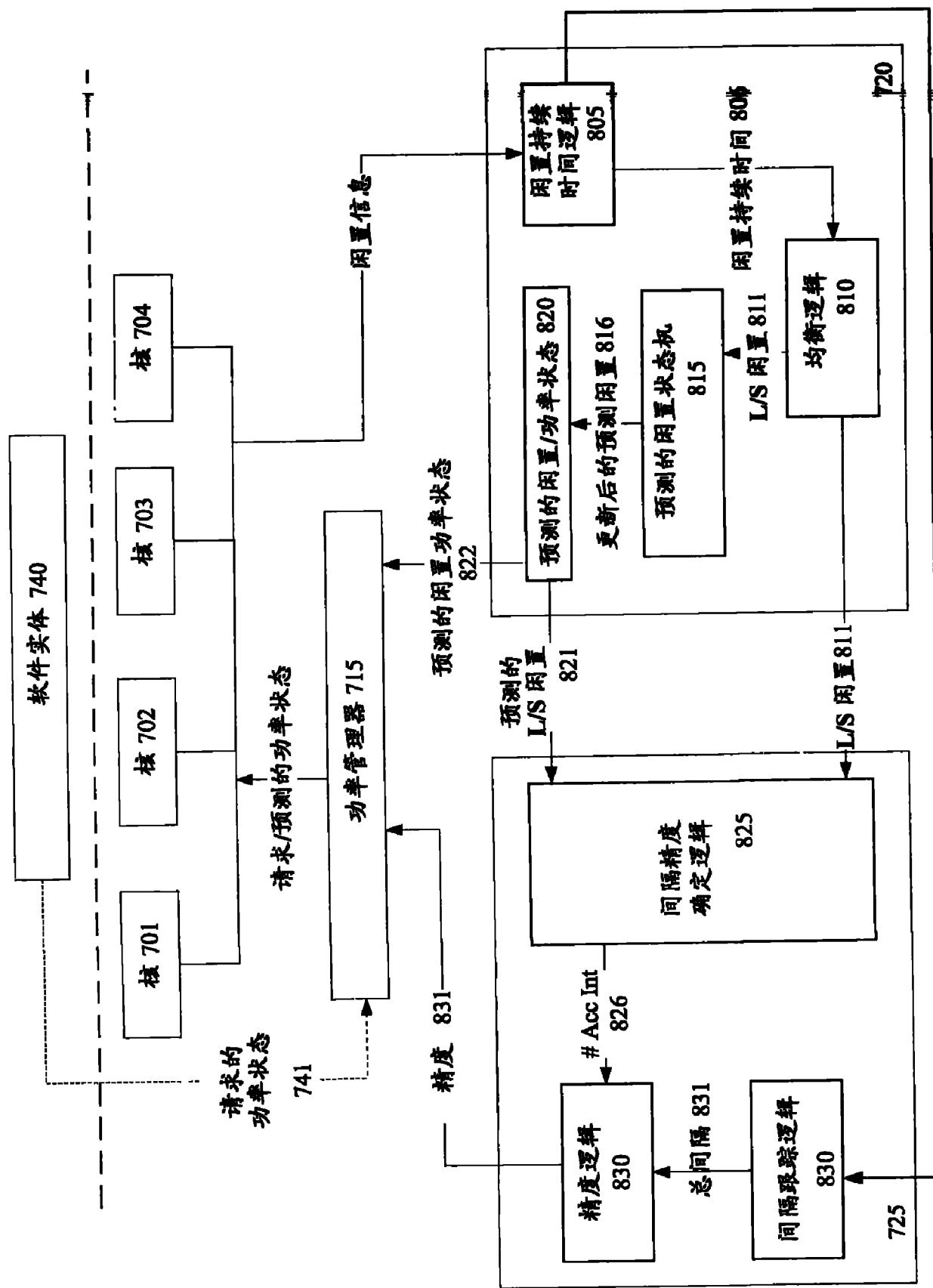


图 8

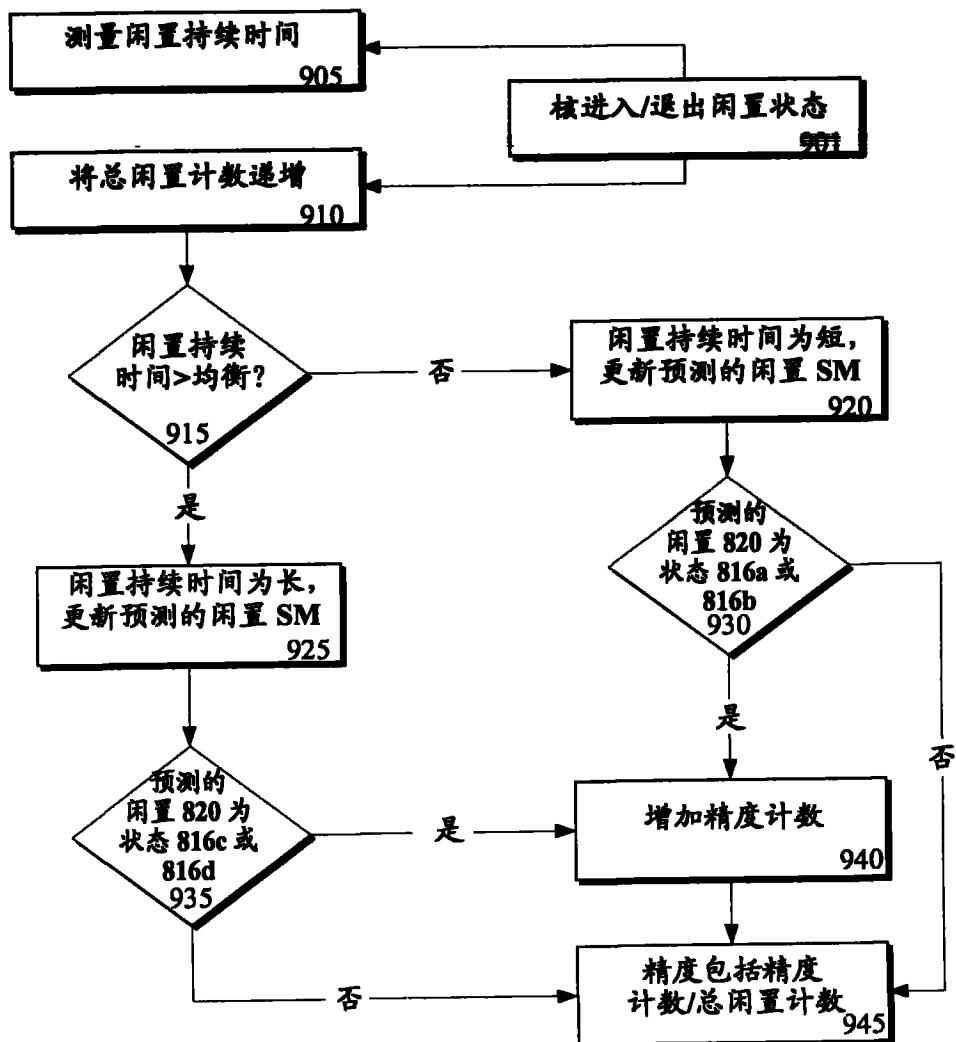
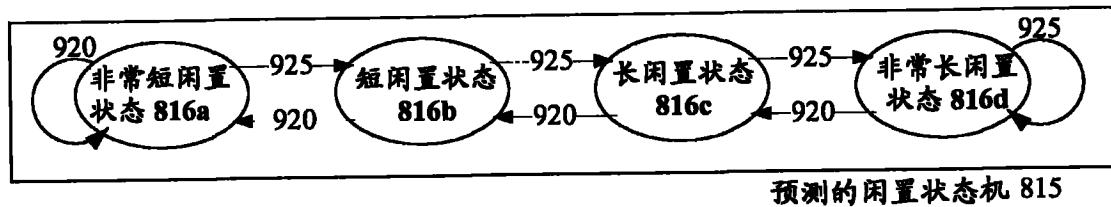


图 9a



预测的闲置状态机 815

图 9b