(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 101346681 B
(45)授权公告日 2011.07.06

(21)申请号 200480004000.3
(22)申请日 2004.01.30
(30)优先权数据
10/367,387 2003.02.14 US
(85)PCT申请进入国家阶段日
2005.08.11
(86)PCT申请的申请数据
PCT/US2004/002840 2004.01.30
(87)PCT申请的公布数据
WO2004/074983 EN 2004.09.02
(73)专利权人 英特尔公司
地址 美国加利福尼亚州
(72)发明人 D·博达斯
(74)专利代理机构 上海专利商标事务所有限公司 31100
代理人 钱慰民
(51)Int.Cl.
G06F 1/26(2006.01)
(56)对比文件
CN 1214130 A,1999.04.14, 全文.
CN 1344389 A,2002.04.10, 全文.
CN 1282911 A,2001.02.07, 全文.
US 2002/0087897 A1,2002.07.04, 说明书第0011-0040段.
审查员 王亮

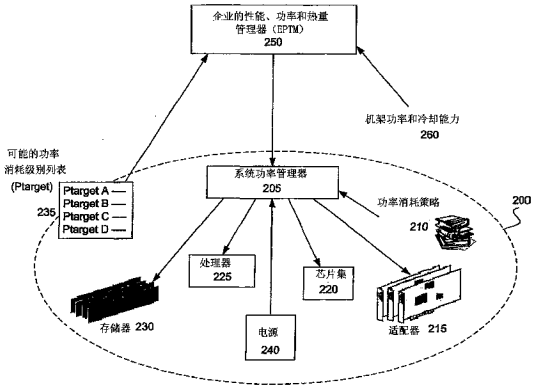
权利要求书 5 页 说明书 11 页 附图 9 页

(54)发明名称

企业的功率和热量管理

(57)摘要

一个计算机系统的功率消耗会基于软件和工作负载而一直变化。像数据中心这样的设备管理很多的计算机系统。随着不断增长的对于计算机系统的功率和冷却的要求,数据中心面临它们提供这些功率和冷却能力的限制。这些限制偶尔会由于功率或者冷却系统的问题而加剧。计算机系统可以用一种方法来维持总体功率消耗低于设定的目标级别。企业的功率和热量管理器,EPTM,可以动态地改变这一设定从而改进支持功率和冷却基础结构的效率。此外,EPTM可使用这一能力来改进性能、可用性,并且改进执行各种管理策略的能力。



CN 101346681 B

1. 一种用于管理计算机系统的功率的方法，包括：

通过控制分配给第一计算机系统的功率来管理所述第一计算机系统的功率消耗，

其中当所述第一计算机系统的功率消耗将被降低时，分配给所述第一计算机系统的功率从第一功率目标降到第二功率目标，

其中当所述第一计算机系统的功率消耗将从第一功率目标增加到第三功率目标时，产生对所述第一计算机系统将分配更多功率的请求，其中当没有足够功率来满足所述请求时，分配给所述第一计算机系统的一个或多个元件的功率被节制。

2. 如权利要求 1 所述的方法，其特征在于，当有足够的可用功率时，分配给所述第一计算机系统的功率被增加。

3. 如权利要求 2 所述的方法，其特征在于，当没有足够的可用功率时，通过降低分配给第二计算机系统的功率来增加分配给所述第一计算机系统的功率。

4. 如权利要求 3 所述的方法，其特征在于，当所述第二计算机系统的功率消耗低于分配给所述第二计算机系统的功率时，分配给所述第二计算机系统的功率被降低。

5. 如权利要求 1 所述的方法，其特征在于，所述第一计算机系统的功率消耗将被降低以响应与所述第一计算机系统相关的警告。

6. 如权利要求 5 所述的方法，其特征在于，所述警告涉及不充足的冷却空气。

7. 如权利要求 5 所述的方法，其特征在于，所述警告涉及高的热能。

8. 如权利要求 5 所述的方法，其特征在于，所述警告涉及可用功率的变化。

9. 如权利要求 1 所述的方法，其特征在于，所述第一计算机系统的功率消耗会基于一个或多个策略而被降低或者升高。

10. 如权利要求 9 所述的方法，其特征在于，所述一个或多个策略包括基于成本的策略。

11. 如权利要求 9 所述的方法，其特征在于，所述一个或多个策略包括基于时间的策略。

12. 如权利要求 9 所述的方法，其特征在于，所述一个或多个策略包括基于收益的策略。

13. 如权利要求 9 所述的方法，其特征在于，所述一个或多个策略包括基于费率的策略。

14. 如权利要求 9 所述的方法，其特征在于，所述一个或多个策略包括基于成本的策略、基于时间的策略、基于费用的策略、基于费率的策略。

15. 一种用于管理计算机系统的功率的方法，包括：

为第一计算机系统定义一个或多个目标功率，

其中所述一个或多个目标功率分别与分配给所述第一计算机系统的一定量的功率相关联；

将所述第一计算机系统设置于第一目标功率；以及

当所述第一计算机系统的功率消耗接近所述第一目标功率时，产生对所述第一计算机系统将分配更多功率的请求，其中当没有足够功率来满足所述请求时，分配给所述第一计算机系统的一个或多个元件的功率被节制。

16. 如权利要求 15 所述的方法，还包括：

对所述第一计算机系统发出的对更多功率的请求作出响应，确定是否有足够的功率用来满足所述请求。

17. 如权利要求 16 所述的方法，其特征在于，其中所述确定是否有足够功率包括：

确定分配给第二计算机系统的功率是否可以被降低。

18. 如权利要求 17 所述的方法，其特征在于，当所述第二计算机系统的性能受最低程度的影响时，分配给所述第二计算机系统的功率可以被降低。

19. 如权利要求 16 所述的方法，还包括：

当有足够的功率以满足所述请求时，将所述第一计算机系统设置于第二目标功率。

20. 如权利要求 19 所述的方法，其特征在于，所述第二目标功率比所述第一目标功率高。

21. 一种用于管理计算机系统的功率的方法，包括：

接收来自第一计算机系统的增加被分配功率的请求，所述第一计算机系统位于由两个或更多个计算机系统构成的组中；以及

响应所述请求，再评估分配给所述组中两个或多个计算机系统的功率以确定所述请求是否可被满足

当不能满足所述请求时，节制分配给所述第一计算机系统的元件的功率。

22. 如权利要求 21 所述的方法，其特征在于，所述再评估分配给所述两个或更多个计算机系统的功率不包括所述第一计算机系统。

23. 如权利要求 22 所述的方法，其特征在于，所述再评估功率包括：

将第二计算机系统的功率消耗和分配给所述第二计算机系统的功率作比较，从而确定分配给所述第二计算机系统的功率是否可被降低。

24. 如权利要求 23 所述的方法，其特征在于，分配给所述第二计算机系统的功率可被降低到高于所述第二计算机系统的功率消耗的级别上。

25. 如权利要求 21 所述的方法，其特征在于，当所述第一计算机系统的功率消耗接近被分配功率时，所述请求被接收。

26. 一种用于管理计算机系统的功率的装置，包括：

用于接收来自第一计算机系统的请求以将分配的功率从第一数额增加至第二数额的装置，所述第一计算机系统处于由两个或更多个计算机系统构成的组中；以及

用于响应所述请求，再评估分配给所述组中两个或更多个计算机系统的功率来决定所述请求是否可以被满足的装置；以及

当不能满足所述请求时，用于节制分配给所述第一计算机系统的元件的功率的装置。

27. 如权利要求 26 所述的装置，其特征在于，所述再评估分配给所述两个或更多计算机系统的功率不包括所述第一计算机系统。

28. 如权利要求 27 所述的装置，其特征在于，所述用于再评估功率的装置包括：

用于将第二计算机系统的功率消耗和分配给所述第二计算机系统的功率的数额进行比较，以确定分配给所述第二计算机系统的功率的数额是否可被减少的装置。

29. 如权利要求 28 所述的装置，其特征在于，分配给所述第二计算机系统的功率的数额可被减少到高于所述第二计算机系统的当前功率消耗级别的数额。

30. 如权利要求 26 所述的装置，其特征在于，所述请求在所述第一计算机系统的功率

消耗接近被分配功率的第一数额时被接收。

31. 如权利要求 26 所述的装置，其特征在于，所述第一数额和所述第二数额是预定的。

32. 一种用于管理计算机系统的功率的装置，包括：

用于通过控制分配给所述第一计算机系统的功率来管理第一计算机系统的功率消耗的装置，其中当所述第一计算机系统的功率消耗要被降低时，分配给所述第一计算机系统的功率从第一目标功率降低到第二目标功率，

当所述第一计算机系统的功率消耗将从第一功率目标增加到第三功率目标时，用于产生对所述第一计算机系统将分配更多功率的请求的装置，其中当没有足够功率来满足所述请求时，分配给所述第一计算机系统的一个或多个元件的功率被节制。

33. 如权利要求 32 所述的装置，其特征在于，当有足够的可用功率时，分配给所述第一计算机系统的功率增加。

34. 如权利要求 33 所述的装置，其特征在于，当没有足够的可用功率时，通过降低分配给第二计算机系统的功率来增加分配给所述第一计算机系统的功率。

35. 如权利要求 34 所述的装置，其特征在于，当所述第二计算机系统的功率消耗比分配给所述第二计算机系统的功率低时，分配给所述第二计算机系统的功率被降低。

36. 如权利要求 32 所述的装置，其特征在于，所述第一计算机系统的功率消耗将被降低以响应与所述第一计算机系统相关联的警告。

37. 如权利要求 36 所述的装置，其特征在于，所述警告涉及不充足的冷却空气。

38. 如权利要求 36 所述的装置，其特征在于，所述警告涉及高的热能。

39. 如权利要求 36 所述的装置，其特征在于，所述警告涉及可用功率的变化。

40. 如权利要求 36 所述的装置，其特征在于，所述警告涉及轻的工作负载。

41. 如权利要求 41 所述的装置，其特征在于，所述第一计算机系统的功率消耗基于一个或多个策略将会被减少。

42. 如权利要求 41 所述的装置，其特征在于，所述一个或多个策略包括基于成本的策略。

43. 如权利要求 41 所述的装置，其特征在于，所述一个或多个策略包括基于时间的策略。

44. 如权利要求 41 所述的装置，其特征在于，所述一个或多个策略包括基于费用的策略。

45. 如权利要求 41 所述的装置，其特征在于，所述一个或多个策略包括基于费率的策略。

46. 一种用于管理计算机系统的功率的装置，包括：

用于为第一计算机系统定义一个或多个目标功率消耗级别的装置，

其中所述一个或多个目标功率消耗级别中每一个与分配给所述第一计算机系统的功率的数额相关；

用于设置所述第一计算机系统于第一目标功率的装置；

当所述第一计算机系统的功率消耗接近第一目标功率时，用于产生给第一计算机系统分配更多功率的请求的装置，其中当没有足够功率来满足所述请求时，分配给所述第一计算机系统的一个或多个元件的功率被节制。

47. 如权利要求 46 所述的装置，还包括：

用于对所述第一计算机系统对更多功率的请求作出响应，确定是否有足够的功率用于满足所述请求的装置。

48. 如权利要求 47 所述的装置，其特征在于，所述用于确定是否有足够的功率的装置包括：

用于确定分配给第二计算机系统的功率可否被降低的装置。

49. 如权利要求 48 所述的装置，其特征在于，当所述第二计算机系统的性能不受影响时，分配给所述第二计算机系统的功率可以被降低。

50. 如权利要求 47 所述的装置，还包括：当有足够的功率来满足所述请求时，用于将所述第一计算机系统设置于第二目标功率上的装置。

51. 如权利要求 50 所述的装置，其特征在于，所述第二目标功率高于所述第一目标功率。

52. 一种用于管理计算机系统的功率的系统，包括：

控制器，所述控制器通过控制分配给所述第一计算机系统的功率来管理第一计算机系统的功率消耗，其中当所述第一计算机系统的功率消耗将被减少时，所述控制器将减少分配给所述第一计算机系统系统的功率，

耦合到所述控制器的性能监视器，所述性能监视器监视所述第一计算机系统和所述第二计算机系统的性能，其中当所述第一计算机系统的性能要提高时，所述性能监视器请求所述控制器为所述第一计算机系统分配额外的功率，其中当没有足够的可用功率时，所述控制器将降低分配给第一计算机系统的一个或多个组件的功率。

53. 如权利要求 52 所述的系统，其特征在于，所述控制器将把分配给所述第一计算机系统的功率由第一级别降低到第二级别。

54. 如权利要求 53 所述的系统，其特征在于，所述第一级别和所述第二级别是预设定的。

55. 如权利要求 52 所述的系统，其特征在于，当所述第一计算机系统的功率消耗将被增加时，所述控制器将增加分配给所述第一计算机系统的功率。

56. 如权利要求 55 所述的系统，其特征在于，当有足够可用功率时，控制器将把分配给所述第一计算机系统的功率由第一级别增加到第三级别。

57. 如权利要求 56 所述的系统，其特征在于，所述第一级别和所述第三级别是预设定的。

58. 如权利要求 52 所述的系统，其特征在于，当所述第二计算机系统的功率消耗低于分配给所述第二计算机系统的功率时，分配给所述第二计算机系统的功率被降低。

59. 一种用于管理计算机系统的功率的系统，包括：

控制器，所述控制器用于控制分配给第一计算机系统和第二计算机系统的功率；以及

耦合到所述控制器的性能监视器，所述性能监视器监视所述第一计算机系统和所述第二计算机系统的性能，其中当所述第一计算机系统的性能要提高时，所述性能监视器请求所述控制器为所述第一计算机系统分配额外的功率，当没有足够的可用功率时，所述控制器将降低分配给第一计算机系统的一个或多个组件的功率。。

60. 如权利要求 59 所述的系统，其特征在于，所述控制器通过降低分配给所述第二计

算机系统的功率来分配额外的功率给所述第一计算机系统。

61. 如权利要求 59 所述的系统，还包括耦合到所述控制器上的温度监视器，所述温度监视器监视与所述第一计算机系统相关的温度，其中当所述第一计算机的温度将被减低时，所述温度监视器请求所述控制器减少分配给所述第一计算机系统的功率。

62. 如权利要求 59 所述的系统，还包括耦合到控制器上的策略处理机，所述策略处理机将为所述控制器提供一种在分配功率给所述第一计算机系统和所述第二计算机系统时应遵循的功率分配策略。

63. 如权利要求 62 所述的系统，其特征在于，所述功率分配策略包括基于成本的策略、基于费用的策略、基于费率的策略、和基于时间的策略。

企业的功率和热量管理

[0001] 相关申请

[0002] 本申请涉及标题为“管理在给特定规格内的系统功率使用的方法和设备”，提交于2002年1月2日的，序列号为10/037,391的待审申请，和标题为“评估系统功率和冷却要求的方法”，提交于2001年12月13日的，序列号为10/022,448的待审申请。

发明领域

[0003] 本发明主要涉及系统管理领域，尤其涉及管理计算机系统的性能、功率、热量和其他属性的方法和装置。

[0004] 背景

[0005] 通常，要衡量计算机系统在各种可能情况下的功率使用是困难的。许多新的使用模型、应用和数据模式定期地被发现。当一种新的使用模型使得计算机系统的元件需要的功率比电源所能提供的多时，电源、从而计算机系统就会发生故障。

[0006] 对于许多计算机系统来说，功率规格是基于分析模型的。这样的分析模型针对计算机系统各元件对功率使用的分配作出某些假设。功率规格可能是所有元件的最大估计功率消耗的总和。电源应被设计为可支持所估计的功率消耗。

[0007] 典型地，计算机系统制造商会为他们的每种计算机系统提供额定功率或者功率规格。该功率规格是基于计算机系统中元件的最大估计功率消耗的。所以功率规格指示出计算机系统的电源应该能够应付的最大功率，这可以被称作P_{MAX}。在确定P_{MAX}的值的过程中，计算机系统设计人员们通常考虑最坏情况配置，在这一典型情况下，P_{MAX}的赋值是基于一个装配了所有元件的计算机系统。

[0008] 此外，P_{MAX}的值也可基于计算机系统装配了最消耗功率的元件，包括硬件元件和软件元件的假设。例如，一个服务器计算机系统被设计为支持四个运行在1.5GHz到2.2GHz之间的处理器，12个内存插槽，8个输入输出(I/O)适配器插槽以及5个硬盘驱动器架。这样一个计算机系统的P_{MAX}值假设它装配了4个2.2GHz(最大功率)的处理器，内存插槽和I/O插槽被完全使用，且装有5个硬盘驱动器。为了让情况更槽糕一点，计算机系统设计者们会加入一个警戒线来降低系统失败的可能性。例如，一个电源可能有一个比估计功率消耗高某一百分比(例如，20％)的最大额定值，结果导致估计功率消耗变大。

[0009] 超裕度设计的电源可以驱动相关联的基础结构的更高要求。这在通常把计算机系统安装在机架上的数据中心很明显。每个机架有限定的功率和冷却能力。数据中心的管理者们经常使用夸大的估计功率消耗(例如，基于铭牌的规格)来决定可以上机架的计算机系统的数量。随着每一代计算机功率规格的增长，机架可以支持的计算机系统的数量减少了。结果，机架上的空位越来越多。此外，超裕度设计的电源过于昂贵了。

附图说明

[0010] 本发明通过例子而非限制的方式，在附图中进行说明，图中相似的标号表示相同的元件，其中：

[0011] 图 1A 是一个计算机机架的例子。

[0012] 图 1B 是对计算机系统的不同的目标功率消耗级别以及功率规格的例子进行说明的图。

[0013] 图 2 是对根据一个实施例，用于改进支持基础结构的效率的控制器的例子进行说明的方块图。

[0014] 图 3 是对根据一个实施例，用来管理一组位于计算机机架上的计算机系统控制器的例子进行说明的方块图。

[0015] 图 4A 和 4B 是一个根据一个实施例，由 EPTM 进行功率分配的例子。

[0016] 图 5A 和 5B 是另一个根据一个实施例，由 EPTM 进行功率分配的例子。

[0017] 图 6 是对根据一个实施例由计算机系统运行的用于请求更多功率的进程进行说明的方块图。

[0018] 图 7 是对根据一个实施例使用 EPTM 的功率分配进程的例子进行说明的流程图。

[0019] 图 8 是对根据一个实施例的，用于为一组计算机系统中的每一计算机系统评估目标功率消耗级别的进程的例子进行说明的方块图。

[0020] 详细描述

[0021] 对于一个实施例，已经揭示了改进支持一个或多个计算机系统基础结构的效率的方法和装置。基于计算机系统中的元件在不同时刻的功率消耗，计算机系统中的功率消耗可能会在不同级别间改变。可以为计算机系统设定一个目标功率消耗级别。当计算机系统的当前功率消耗级别接近于目标功率消耗级别时，计算机系统将被分配额外的功率。

[0022] 在以下描述中，为了解释的需要，将会阐述大量具体的细节，以便透彻了解本发明。不过显然地，本领域技术人员可以不采用这些具体细节来实践本发明。在其他实施例中，众所周知的结构、进程和设备在以方块图的形式被显示，或者以概括的方式被提及，以免提供过于详细的解释说明。

[0023] 目标功率消耗级别

[0024] 图 1A 是一个计算机机架的例子。计算机机架 100 可包括一组含计算机系统 105 在内的计算机系统。这组计算机系统中的每一个计算机系统可被设置成执行不同的计算角色，例如，网络服务器、路由器等。计算机机架 100 被放置在一个密闭的房间内。该房间配有充足的空调装置以便为包含在计算机机架 100 中的计算机系统保持合适的操作温度。该房间可能有其他类似于计算机机架 100 的计算机机架（未示出），这些其他计算机机架可能包括其他计算机系统。该房间可能位于一个数据中心，且它的尺寸可变化。

[0025] 计算机机架 100 与特定的功率和冷却能力相关联。典型地，数据中心管理员们使用由计算机系统制造商们指定的功率规格来决定可以包括于计算机机架上的计算机系统的数量。图 1B 是对一个计算机系统的不同目标功率消耗级别和功率规格的例子进行说明的图。计算机系统的功率规格 125 可被称作 P_{MAX}。在很多情况下，包括在计算机机架 100 中的所有计算机系统的 P_{MAX} 的合计值被用来决定计算机机架 100 如何被装配。每一代计算机系统都有更高的功率规格。结果就是，可以包括在计算机机架上，例如计算机机架 100 上的计算机系统的数量可能会减少。

[0026] 由计算机系统制造商提供的功率规格 125(P_{MAX}) 可能会被高估。就是说，功率规格 125(P_{MAX}) 可能会比计算机系统的实际功耗级别高一点。这可能是因为大多数计算机系统没

有安装所有可安装的元件。即使一个计算机系统被充分装配了，这些元件也未必是计算机系统制造商估计时采用的最消耗功率的元件。对于存放许多计算机系统的数据中心，基于功率规格 125(P_{MAX}) 来决定必需的功率会导致在支持基础结构方面不必要的需求，例如，不必要的计算机机架和冷却能力。

[0027] 在一个实施例中，每个包括在计算机机架 100 中的计算机系统与一个或者更多个目标功率消耗级别相关联。目标功率消耗级别可被称作 P_{TARGET}。P_{TARGET} 可以设置得低于或等于 P_{MAX}。例如，参照图 1B，可以有三个不同的目标功率消耗级别 110、115、120（或者 $P_{TARGET\ A}$、$P_{TARGET\ B}$、$P_{TARGET\ C}$）。在一个实施例中，P_{TARGET} 可用来分配功率给计算机系统。使用 P_{TARGET}110、115、120 可允许包括例如可用功率和冷却能量的支持基础结构得到更好的利用和提高的效率。例如，在 t1 时刻，可以给计算机系统 105 设置目标功率消耗级别 110($P_{TARGET\ A}$)。在 t2 时刻，当计算机系统 105 的功率消耗级别接近目标功率消耗级别 110($P_{TARGET\ A}$) 时，额外的功率可以被分配给计算机系统 105。为了实现这一点，可以为计算机系统 105 设置目标功率消耗级别 115($P_{TARGET\ B}$)。类似的，在 t3 时刻，当计算机系统 105 的功率消耗级别接近目标功率消耗级别 115($P_{TARGET\ B}$) 时，通过采用目标功率消耗级别 120($P_{TARGET\ C}$)，额外的功率可以被分配给计算机系统 105。需要指出的是分配给计算机系统 105 的功率的数额要比使用功率规格 125(P_{MAX}) 时所分配的数额少。同时也需要指出的是当有可用的功率时，额外的功率可以被用于分配。

[0028] P_{TARGET} 的设置取决于几个因素。这些因素可包括：例如计算机系统中元件的数量、工作负载、可用的功率容量、可用的冷却能力、性能要求、环境要求、管理要求、等等。可以用手动方法或者自动的方法来设置 P_{TARGET}。例如，手动方法可涉及使用典型的工作负载来确定计算机系统 105 的最高功率消耗级别。一个百分数（例如，5%）会被用来作为到达 P_{TARGET} 的安全警戒线。在另一个例子中，可以通过在一段使用高峰时期自动地监测计算机系统 105 的功率消耗级别然后计算功率消耗级别的平均值来设置 P_{TARGET}。在一个实施例中，可以通过使用在标题为“评估系统功率和冷却要求的方法”，提交于 2001 年 11 月 13 日的，序列号为 10/022,448 的申请中所描述的技术来决定 P_{TARGET}。其它技术也可以用于建立 P_{TARGET}。

[0029] 功率节制

[0030] 计算机系统 105 可以用 P_{TARGET} 来确定何时开始节制计算机系统 105 中的一个或更多元件的功率。例如，给计算机系统 105 中的硬盘驱动器（未示出）的功率可被降低，这样计算机系统 105 的功率消耗会维持在低于或等于目标功率 P_{TARGET} 的水平。在一个实施例中，通过使用标题为“管理在给特定规格内的系统功率使用的方法和设备”，提交于 2002 年 1 月 2 日的，序列号为 10/037,391 的申请中所描述的技术，给一个计算机系统中一个或更多元件的功率可被节制。其它功率节制的技术也可以被使用。

[0031] 图 2 是方块图，说明根据一个实施例，用于改进支持基础结构的效率的控制器的例子。控制器可以被称作企业的功率和热量管理器 (EPTM)。在本例中，计算机系统 200 包括多种元件（例如，处理器 225、芯片集 220、内存 230、等等）。在一个实施例中，系统功率管理器 205 可从电源 240 处接收关于计算机系统 200 的当前功率消耗级别的信息。例如，电源 240 可提供一端口（未示出）以允许系统功率管理器 205 提取关于计算机系统 200 的当前功率消耗级别的信息。其它技术也可以用于提供计算机系统 200 的当前功率消耗级别。例如，元件可直接向系统功率管理器 205 报告他们的功率消耗。

[0032] 系统功率管理器 205 负责监视和管理计算机系统 200 中的一个或多个元件的功率消耗。系统功率管理器 205 使用功率消耗策略 210 来管理元件的功率消耗以便在必要时进行适当的功率增加或降低操作。例如，当计算机系统 200 的当前功率消耗级别正接近 P_{TARGET} 时，系统功率管理器 205 就会使用由功率消耗策略 210 提供的信息来确定如何降低一个或多个元件的功率。降低一个或多个元件的功率会在使当前功率消耗级别被保持在接近或低于 P_{TARGET} 的水平的同时影响元件和计算机系统 200 的性能。

[0033] 在一个实施例中，因为工作负载会经常改变，所以能够动态地设定 P_{TARGET} 以适应不同的功率要求将是有利的。例如，当目标功率消耗级别 110($P_{TARGET\ A}$) 被使用，且当前功率消耗级别接近目标功率消耗级别 110($P_{TARGET\ A}$) 时，下一目标功率消耗级别 115($P_{TARGET\ B}$) 被设置，而不是节制这些元件的功率。当然，设置目标功率消耗级别 115($P_{TARGET\ B}$) 包括验证确实有充足的可用功率来分配给计算机系统以支持目标功率消耗级别 115($P_{TARGET\ B}$)。在一个实施例中，功率节制和设置下一目标功率消耗级别会被一起使用，如果当前功率消耗所处的级别需要进行上述操作的话。需要指出的是当将 P_{TARGET} 设置得过高时，分配给计算机系统 200 的功率不会被充分地利用。

[0034] 功率与热量管理系统

[0035] 图 2 中所示的功率与热量管理器 (EPTM)250 基于功率和冷却能力 260 对分配给计算机系统 200 的功率进行管理。其它因素（例如性能等等）也会被 EPTM250 用来管理分配给计算机系统 200 的功率。计算机系统 200 可以是位于像图 1 中所示的机架 100 那样的机架中许多计算机系统中的一个。数据中心管理员们可确定功率和冷却能力 260。

[0036] EPTM 250 通过计算机系统 200 中的系统功率管理器 205 与计算机系统 200 进行通信。这可以通过带内信道，例如，局域网 (LAN)、或者交互应用通信，或者带外信道，例如总线来实现。其它技术也可以用于使 EPTM 250 与计算机系统 200 通信。EPTM 250 把 EPTM 250 要分配给计算机系统 200 的功率的数额指示给系统功率管理器 205。该被分配的功率数额将通过电源 240 提供。该被分配的功率数额是基于已设定的目标功率消耗级别 P_{TARGET} 的。EPTM 250 还可询问计算机系统 200 其当前功率消耗级别。EPTM 250 可询问功率消耗级别的历史或者在特定时间的功率消耗级别。

[0037] 在一个实施例中，计算机系统 200 可与一个或多个目标功率消耗级别 (P_{TARGET}) 相关联。每一个 P_{TARGET} 可被 EPTM 250 使用以分配适当的功率数额给计算机系统 200。在一个实施例中，不同目标功率消耗级别 235 的列表将被存储在计算机系统 200 上，并提供给 EPTM 250。例如，当计算机系统 200 是空闲的并且可以在一个低的目标功率消耗级别上运行时，就会被设置在这个低的目标功率消耗级别，而 EPTM 250 依此会分配小量数额的功率给计算机系统 200。

[0038] 图 3 是方块图，对根据一个实施例，用于管理一组位于一个计算机机架内的计算机系统的控制器的例子进行说明。EPTM 250 可管理除计算机机架 100 外的一个或多个计算机机架（未示出）。对分配给计算机机架 100 中的计算机系统 300-306 的功率的管理取决于计算机系统 300-306 中各自的角色。例如，计算机系统 300 可能担当一个高优先权的网络路由器，那么 EPTM 250 就要给计算机系统 300 分配尽可能多的，满足其需要的功率。这一分配可能会影响分配给其余的与计算机系统 300 处于同一机架上的计算机系统 301-306 的功率。除了接收涉及功率和冷却能力 260 的信息外，EPTM 250 还接收关于可选功率源例

如，不间断电源 (UPS)285 或者其它备用发电机（未示出）的信息。EPTM 250 会接收提供功率的电力公司 275 的状态信息。EPTM 250 也会接收例如像为计算机机架 100 提供冷却空气的空气调节单元 270 那样的冷却系统的元件的状态信息。其它冷却系统的元件的例子包括但不局限于热交换机（未示出）。

[0039] EPTM 250 可使用管理策略 255 来使得 EPTM 250 可以，举例来说，调节计算机系统 300-306 中各自的 P_{TARGET} 以完成特定的任务。管理策略 255 可以通过一个将适当策略通知给 EPTM 250 的策略处理机（未示出）来进行。例如，有些计算机系统会由管理策略 255 设计成高优先权或者高性能，因而 EPTM250 有必要为针对这些计算机系统的功率分配区分优先权次序。所以，EPTM 250 可能在牺牲位于计算机机架 100 中其它计算机系统的前提下分配功率给这些计算机系统。管理策略 255 将指引 EPTM 250 降低计算机机架 100 的功率或者冷却（例如，10％），从而降低开支。管理策略 255 将指引 EPTM 250 调节计算机系统 300-306 中各自的 P_{TARGET} 来改变功率消耗的模式，从而利用电力公司 275 的可能的“日常分时计量”。

[0040] 在一个实施例中，EPTM 250 也接收来自性能监视器 280 信息。性能监视器 280 可以是一个监视位于计算机机架 100 中的计算机系统 300-306 中的一个或多个的性能和 / 或工作负载的系统。性能监视器 280 的例子是网络负载平衡器。性能监视器 280 会请求 EPTM 250 基于当前工作负载调节计算机系统 300-306 中各自的 P_{TARGET}。这包括使用更高 P_{TARGET} 的来获得更高的性能。在另一个例子里，性能监视器 280 可充当负载平衡器，并移动计算机系统 300-306 间的工作负载以使得在一给定的工作负载下，功率的节约最大化。性能监视器 280 可以是一个与 EPTM 250 分开的计算机系统，或者可以是 EPTM 250 的一部分，或者是计算机系统 300-306 中的一个或多个的一部分。

[0041] 在一个实施例中，EPTM 250 也可以接收来自热电偶 265 的信息。热电偶 265 可接收来自计算机系统、计算机机架 100，和 / 或计算机机架 100 所处的房间或者数据中心中的所有或者部分的成员的热输入。接收来自热电偶 265 的信息使得 EPTM 250 可以管理用于保持计算机系统 300-306 在一个可接受的操作条件所需要的冷却空气。根据接收自热电偶的信息，EPTM 250 可以降低供给一个或多个计算机系统 300-306 的 P_{TARGET} 以冷却那些计算机系统。

[0042] 在一个实施例中，EPTM 250 接收指示电能来源的状态的信息，其中电能包括了例如来自电力公司、不间断电源 (UPS) 或者备用发电机，例如用柴油机工作的发电机的电力。当备用发电机和 / 或 UPS 系统由于数量短缺（或者一些其它原因）而发生故障的时候，数据中心将会经历电力危机，因为电力公司不能满足所有需求。常规的应对这种危机的方法是关掉一个或多个计算机系统。在一个实施例中，EPTM 250 可降低一个或多个计算机系统的 P_{TARGET} 来降低总体功率消耗。这使得计算机机架 100 中的所有计算机系统在经过最小停机时间后可以继续运行。取决于工作负载、当前功率消耗级别、以及计算机系统的 P_{TARGET} 的综合情况，这可能会也可能不会对计算机系统的性能产生影响。

[0043] 可用功率的分配

[0044] 图 4A 和 4B 是根据一个实施例的，由 EPTM 进行的功率分配的一个例子。图 4A 和 4B 包括相同的 EPTM 250，和装配相同计算机系统的相同计算机机架。本例中的机架有能力提供和冷却 6000 瓦。计算机系统包括计算机系统 405 和 410。机架中每个计算机系统的 P_{TARGET} 可被设为 600 瓦。每个计算机系统的 P_{TARGET} 比其功率规格 P_{MAX} 要低。图 4A 中所示的

每个计算机系统都与一个 P_{TARGET} 相关联。

[0045] 每个计算机系统分别在某个当前功率消耗级别上运行。当前功率消耗级别可能低于或者接近 P_{TARGET}。在本例中，与计算机系统 405 和 410 相关联的 P_{TARGET} 可被设为 600 瓦。需要指出的是，图 4A 中所示的计算机系统 405 和 410 的当前功率消耗级别也是 600 瓦，而图 4A 中机架内所有计算机系统的总体功率消耗大约为 4700 瓦。在本例中，由于机架的功率能力是 6000 瓦，所以有 1300 瓦 (6000-4700) 的可用功率可供 EPTM 250 在必要时分配。EPTM 250 可使用这些可调用功率来满足一个或多个计算机系统对额外功率的请求。

[0046] 如图 4A 所述，计算机系统 405 和 410 的当前功率消耗级别已经达到了他们 600 瓦的 P_{TARGET}，将向 EPTM 250 请求额外的功率。这在本例中，被表示为从计算机系统 405 和 410 指向 EPTM 250 的方向箭头。EPTM 250 通过动态地分配可用功率给计算机系统 405 和 410 来响应该请求。例如，EPTM 250 将计算机系统 405 和 410 各自的目标功率消耗级别 P_{TARGET} 设置到下一更高级别，然后根据这些更高的 P_{TARGET} 以及可用的 1300 瓦给计算机系统 405 和 410 分配功率。在本例中，分配给计算机系统 405 的功率从 600 瓦提高到 700 瓦，而分配给计算机系统 410 的功率从 600 瓦提高到 650 瓦，如图 4B 所示。这由从 EPTM 250 指向计算机系统 405 和 410 的方向箭头表示。更高的功率分配使得计算机系统 405 和 410 可以提供更高的性能。在该例中，由于有可用功率，EPTM 250 可以在对机架中其它计算机系统的性能造成的影响最小的情况下，满足计算机系统 405 和 410 的请求。需要指出的是，当警戒线被加入 P_{TARGET} 时，计算机系统 405 和 410 会在它们的当前功率消耗级别到达它们适当的 P_{TARGET} 之前就产生请求。

[0047] 通过对分配功率评估进行的功率分配

[0048] 图 5A 和 5B 是另一个根据本发明的，由 EPTM 进行的功率分配的例子。在该例中，计算机系统 405 和 410 要求 EPTM 分配更多的功率。在某种情况下，可能没有足够多的可调用功率来满足额外功率的请求，除非采取进一步的操作。例如，可能没有任何可调用功率。在一个实施例中，EPTM 250 可对照机架中计算机系统各自的当前功率消耗级别，对它们各自的 P_{TARGET} 进行评估。EPTM250 不需要对计算机系统 405 和 410 的当前功率消耗级别进行评估，因为是它们发起对更多功率的请求的，所以它们应该运行于它们的 P_{TARGET}。

[0049] 如上文所描述的，每个计算机系统可以与一个以上的 P_{TARGET} 相关联。例如，EPTM 250 可检查每个计算机系统可能的目标功率消耗级别 235（在图 2 中所描述的）的列表，以此确定 EPTM 250 是否可以令计算机系统运行在一个较低的目标功率消耗级别。参照图 5A 和 5B 中所示的例子，EPTM 250 可决定让计算机系统 415 运行在更低的 P_{TARGET} 上，从而可以把分配给计算机系统 415 的功率从 550 瓦降低到 400 瓦。从计算机系统 415 那里得到的降低的功率 (150 瓦）可以用来重新分配给计算机系统 410(50 瓦）和计算机系统 405(100 瓦）。需要指出的是，EPTM 250 可能需要降低分配到一个以上的计算机系统的功率来满足该请求。同时需要指出的是，被 EPTM 250 降低所得功率的计算机系统可能是空闲的或者不是很忙，以使得对它们的性能造成的影响最小。

[0050] 在一个实施例中，发出对更多功率的请求的计算机系统要等待一段预设定的时间，使得 EPTM 250 来确定可用功率并分配必要的功率以满足请求。在一个实施例中，当 EPTM 250 不能满足来自计算机系统的对更多功率的请求时，与该计算机系统相关联的功率管理器 205（在图 2 中所描述的）将被激活以节制该计算机系统的一个或多个元件的功率

消耗。

[0051] 需要指出的是，一个计算机系统可能有也可能没有与一个以上的目标功率消耗级别相关联。而且，当计算机系统与一个以上的目标功率消耗级别相关联时，EPTM 250 就没有必要使用不同的功率消耗级别来管理分配给该计算机系统的功率。

[0052] 功率管理

[0053] 通常情况下，数据中心（例如包括计算机机架 100 中的计算机系统）的电是由电力公司提供的。当由电力公司提供的功率级别发生波动时可能会发生状况。在这种情况下，可调用功率量有时候会比用于维持数据中心里的计算机系统运行的期望值低。当没有其它功率源来补偿这一功率波动时，数据中心管理员们会采用关闭一个或者更多计算机系统来减少对功率的需求。关闭计算机系统会导致，例如，服务缺失、产生收益的机会减少、可用性降低等等。在一个实施例中，如果有功率波动，会产生一个功率波动警告，EPTM 250 将被用于降低一个或更多计算机系统的可用功率的级别，并且 / 或者用于使计算机系统的总体功率需求符合可调用功率。例如，当 EPTM 250 确认可用功率降低了 50％，那么 EPTM 250 会以此为根据降低每个计算机系统的目标功率消耗级别。由于这不需要通过关闭计算机系统来实现，可用性得到了保持。需要指出的是，计算机系统可能需要对它们的一个或多个元件进行功率节制操作来应付更低的功率消耗级别。

[0054] 当接收来自电力公司的电力发生问题时，数据中心管理员们可使用不间断电源 (UPS) 系统。当来自电力公司的电力长期中断时，许多数据中心管理员们也使用各种其它的备用发电机。在一个实施例中，当一个或者更多个 UPS 系统故障时，EPTM 250 可被用于降低一个或多个计算机系统可用的功率的级别。例如，当接收来自电力公司的电力发生故障时，且当 UPS 系统有故障时，数据中心管理员们会切换至备用发电机来弥补功率的波动。在这些情况下，EPTM 250 会相应地降低每个计算机系统的目标功率消耗级别以便与备用发电机产生的功率数额相符合。类似地，当备用发电机故障时，EPTM 250 会基于可调用功率采取适当的操作来保持计算机系统的运行。在该例中，需要指出的是，取决于各个计算机系统的当前工作负载，它们的性能可能会也可能不会受 EPTM 250 的动作的影响。

[0055] 热量管理

[0056] 在运行时，数据中心的每一个计算机系统会产生热能或热量。该热量可以被数据中心的空气调节系统产生的循环冷却空气带走。大多数空气调节系统具有可动元件，例如，风扇、压缩机、等等，而它们可能会发生故障。当一个或者更多这样的元件发生故障时，数据中心的冷却能力就会降低。除非该事件期间计算机系统的功率消耗级别也被降低，否则数据中心的温度会升高，且计算机系统中的组件可能发生故障。通常地，为了避免这种不愿看到的温度升高，数据中心的管理员们需要通过关闭一个或多个计算机系统来降低计算机系统的功率消耗级别。这会导致服务丢失、可用性减低、等等。在一个实施例中，在没有足够的冷却空气时，可能会产生温度警告，然后 EPTM 250 会降低计算机系统的目标功率消耗级别。这会有助于在不必关机的情况下维持计算机系统的运行。

[0057] 冷却空气在数据中心里的分布可能不是均匀的。这意味着数据中心的有些区域会比其它区域凉。冷却空气较少的区域可被称作热区。位于热区的计算机系统由于其相对于那些不位于热区的计算机系统更加高的温度而可能更容易发生故障。更糟的情况是热区可能会由于诸如人的行为而发生变化。在一个实施例中，每个计算机机架，例如计算机机架

100，都与一个用来测量其邻近的温度的热电偶相关联。在另一个实施例中，每个计算机系统也都与一个热电偶相关联。无论计算机系统是否位于热区都如此。在一个实施例中，由热电偶测得的温度被送给 EPTM 250。EPTM 250 然后用这一信息来降低计算机系统的功率消耗。例如，EPTM 250 可降低其热电偶显示高温的计算机系统的目标功率消耗级别。这会有助于降低热输出，从而也有助于降低位于计算机系统内或周围的空气温度。

[0058] 管理策略的执行

[0059] 在一个实施例中，EPTM 250 也可用于执行不同的管理策略。这可能在可以为多个计算机系统设置目标功率消耗级别以应对可用功率和冷却能力的 EPTM250 以外。管理策略会创造成本节约、增加收益产生、增加可用性、改进性能的机会等等。为了节约成本，EPTM 250 可用于降低计算机系统的功率消耗以便减少例如 10％的功率和冷却成本。这可以被称作基于成本的管理策略。

[0060] 为了增加收益，EPTM 250 可以用于为不同的计算机机架（例如，2000 瓦机架、3000 瓦机架、等等）维持不同的功率和冷却能力。基于客户愿意支付的容量或者服务级别对他们进行收费。这可以被称作基于费用的管理策略。

[0061] 一些电力公司会在每天不同的使用时间改变他们的电费率。在一个实施例中，EPTM 250 利用这一费率信息并且执行基于时间的策略来设定目标功率消耗级别以利用不同的功率费率。这可以被称作基于费率的管理策略。

[0062] 在一个实施例中，EPTM 250 保存涉及一个或更多计算机系统的功率消耗级别的数据记录。数据将被分析以使得 EPTM 250 确认一个或更多计算机系统的进一步的功率需求从而优先占有对更多功率的请求。例如，EPTM 250 注意到特定的网络服务计算机系统在一年的第 1 周至第 50 周的工作日的 8AM 到 10AM 期间有相当大的工作负载（因此有更高的功率消耗级别）。结果，EPTM 250 在第 1 周至第 50 周的工作日的 8AM 之前把网络服务计算机系统的目标功率消耗级别 P_{TARGET} 改变成较高的目标功率消耗级别。然后，EPTM 250 在 10AM 之后把网络服务计算机系统的目标功率消耗级别（P_{TARGET}）改变成较低的功率消耗级别。这可以被称作基于时间的管理策略。这一策略可以减少任何涉及等待计算机去请求额外功率以及等待被 EPTM 250 分配的实际功率导致的延时。在一个实施例中，EPTM 250 会采用基于成本的策略、基于费用的策略、基于费率的策略、基于时间的策略中的一个或多个策略。

[0063] 请求更多功率的进程

[0064] 图 6 是说明根据一个实施例，由一个计算机系统执行的请求更多功率的进程的方块图。请求会被如上文描述那样送给 EPTM 250。在方块 605 处，进行一个测试以确定当前功率消耗级别是否超过目标功率消耗级别。从方块 605 开始，在当前功率消耗级别没有超过目标功率消耗级别时，目标功率消耗级别（相应的，当前已分配功率的数额）仍然是可接受的。在这种情况下，进程仍然留在方块 605 处，直到当前功率消耗级别超过目标功率消耗级别。当目标功率消耗级别被超过时，将产生一个对更多功率的请求，如方块 610 中所示。

[0065] 当对更多功率的请求在方块 610 中产生后，进程会等待来自 EPTM 250 的响应。换句话说，进程会等待一段预设定时间，而且不管是否从 EPTM 250 收到了响应，进程流到方块 615 以确定请求是否被满足。当请求不能被满足时，或者在预设定时间内没有收到响应，进程由方块 615 流到方块 620，在方块 620 处计算机系统需要节制其一个或多个元件的功率消耗以确保其当前功率消耗级别符合其被设定的目标功率消耗级别。从方块 615 处，当请

求可被满足时，新的目标功率消耗级别被设定，如方块625中所示。进程然后由方块625流到方块605，在方块605处计算机系统继续监视其当前功率消耗级别并与新的目标功率消耗级别进行比较。需要指出的是，将会与目标功率消耗级别一起使用一个阈值，这样在当前功率消耗级别接近目标功率消耗级别时，可产生对更多功率的请求。

[0066] 处理要求额外功率的请求的进程

[0067] 图7是根据一个实施例，对使用EPTM进行功率分配的例子进行说明的流程图。进程可以由上述的EPTM 250来执行。在方块705处，对更多功率的请求被EPTM 250接收。请求可由被EPTM 250管理的一组计算机系统中的一个计算机系统产生的。例如，这组计算机系统可安装在一个计算机机架上，例如图1中所示的计算机机架100。在一个实施例中，EPTM 250可监视一个或更多计算机系统的当前功率消耗级别，并且可自行判断何时被监视的计算机系统需要额外的功率，并且动态决定是否可分配额外功率。在这种情况下，被监视的计算机系统可以不发出请求。

[0068] 在方块710处，进行一个测试以判断是否有足够的可用功率来满足请求。如果EPTM 250没有把所有可用功率分配给该组计算机系统，那么功率就是可用的，如图4A中的例子所示。当还有可用的功率用于满足请求时，进程由方块710流到方块715，以满足对更多功率的请求。这可包括EPTM 250把更多功率分配给产生请求的计算机系统。

[0069] 当没有足够的可用功率来满足请求时，进程从方块710流到方块720，在方块720处位于该组计算机系统中的其它计算机系统的各自的当前功率消耗级别被确定。在方块725处，进行一个测试以确定为该组内其他计算机系统各自所设的目标功率消耗级别(P_{TARGET})是否可被降低。如上文所描述的，设定的目标功率消耗级别(P_{TARGET})可用于确定分配给计算机系统的功率量。这可能低于计算机系统制造商建议的功率规格(P_{max})。例如，测试会基于每个计算机系统的当前行为和工作负载确定它们是否可以在较低的功率消耗级别上运行。

[0070] 当分配给其它计算机系统的功率不能被降低时，进程由方块725流到方块740。在方块740处，会产生一个响应以说明对更多功率的请求不能被满足。当分配给一个或更多其它计算机系统的功率可被降低时，进程由方块725流到方块730。

[0071] 在方块730，进行一个测试以决定由降低分配给一个或更多其它计算机系统产生的功率是否能够满足对更多功率的请求。如果不是，进程由方块730流到方块740，请求不能被满足。然而，如果来自方块730的结果是肯定的，进程流到方块735，此处分配给这些其它计算机系统的功率可以被降低。这可包括为这些其他计算机系统各自设置降低的目标功率消耗级别。进程然后从方块735流到方块715，此处可将功率分配给产生对更多功率的请求的计算机系统。

[0072] 再评估已分配功率的进程

[0073] 图8是根据一个实施例，对用于为一组计算机系统中的每个计算机系统重新评估目标功率消耗级别的进程的例子进行说明的方块图。在方块805处，该组计算机系统中的每个计算机系统（可不含产生对更多功率请求的计算机系统）的目标功率消耗级别被确定。如上文所描述的，分配给计算机系统的功率是基于计算机系统被设定的目标功率消耗级别的。

[0074] 在方块810处，计算机系统的当前功率消耗级别被确定。当前功率消耗级别可能

会低于或接近目标功率消耗级别。在方块 815 处，进行一次比较来确定当前功率消耗级别是否比目标功率消耗级别要低，从而有可能降低目标功率消耗级别而不对计算机系统的性能造成太多影响。例如，计算机系统可以对应于一个较低的目标功率消耗级别来消耗功率。在这种情况下，该计算机系统的目标功率消耗级别将被改到较低的目标功率消耗级别。当计算机系统的目标功率消耗级别不能被更改（例如，计算机系统只能在一个目标功率消耗级别上运行）时，进程由方块 815 流到方块 830。在方块 830，进行一次测试以确定组内是否有其它的计算机系统需要为了可能的被分配功率的减少而被检查。

[0075] 在方块 815，当计算机系统的目标功率消耗级别可被减少时，进程流到方块 820，此处新的较低的目标功率消耗级别被设定。新的目标功率消耗级别可对应于较低的分配给计算机系统的功率量。这意味着更多的功率变得可用。在方块 825 处，累积的可调用功率量被更新。该进程然后流到方块 830 处以确定组内是否有其它计算机系统需要为了可能的被分配功率的减少而被检查。当所有计算机系统都被检查后，进程由方块 830 流到方块 835，此处累积的可调用的功率量被用于确定是否足以满足对更多功率的需求。

[0076] 图 8 中的进程试图在确定请求是否可被满足之前降低组内所有其它计算机系统的功率。或者，图 8 中的进程不需要试图减少所有计算机系统的功率，而是代之以仅仅降低组内某些计算机系统的功率，直到请求被满足。有可能即使进程试图降低所有计算机系统的功率，可调用功率的数额还不足以满足请求。例如，每个计算机系统可能已经运行在或者接近其相应的目标功率消耗级别。

[0077] 需要指出的是，EPTM 250 可能会也可能不会再评估已分配给计算机系统的功率。例如，当对更多功率的请求被接收时，EPTM 250 可能由于没有足够的可调用功率而无法满足该请求，即使其它计算机系统可能没有充分利用它们被分配的功率。或者，EPTM 250 可如图 8 所示进行再评估以满足请求。

[0078] 计算机系统

[0079] 本发明的各种实施例的操作可以由一个计算机系统内的执行计算机程序指令序列的处理单元来实现。EPTM 250 可以通过软件、硬件或者软硬件结合的方式实现。例如，EPTM 250 可被实现为包含硬件电路的芯片或芯片集，该芯片或芯片集包括了专用于执行功率、性能和热量管理功能的辅助处理器。该芯片或芯片集还进一步包括内部存储器、和至计算机系统内各元件（例如，系统 CPU、系统内存、等等）的总线连接。该芯片或芯片集也包括用于接收例如来自诸如那些安装于计算机机架 100 内的其它计算机系统的功率请求的总线连接。EPTM250 然后被耦合到性能监视器 280、热电偶 265 和提供一个或多个管理策略的策略处理机 255 上。

[0080] 计算机可读媒介

[0081] 在一个实施例中，EPTM 250（在图 7-8 中的例子中所描述的）可以通过软件来实现，其中包括了被存储在可被视为机器可读存储媒介的存储器中的指令。存储器可以是随机存取存储器、只读存储器、或者持久保存存储器，例如海量存储设备或者任何这些设备的结合物。根据本发明的实施例，指令序列的执行会使得处理单元执行操作。指令可以从存储设备或者通过网络连接从一个或多个计算机系统（例如，服务器计算机系统）加载到计算机系统的存储器中。指令可被同时存储在几个存储设备中。类似的，由被 EPTM 250 管理的计算机系统产生的对更多功率的请求（在图 6 中的例子中所描述的）可以通过硬件、软

件或软硬件结合实现。所以，本发明的实施例不局限于任何明确的硬件和软件的结合，也不局限于任何特定的被计算机系统执行的指令来源。

[0082] 已经在此披露了使用目标功率消耗级别 P_{TARGET} 技术以帮助一个或更多计算机系统改进支持基础结构的效率的方法。该技术可用于调节分配给一个或更多计算机系统的功率。功率分配的调整可基于一个或更多的因素，例如，温度、性能、工作负载、功率容量、冷却能力、成本、可用性、等等。该技术可用于实现管理计算机系统中的管理策略。

[0083] 尽管本发明已经参照特定的示例性实施例描述，但显然可以在不背离如权利要求中阐述的更宽的精神和本发明的范围的前提下对这些实施例作出各种修正和改变。因此，这些说明和附图应该被看做是一种解释而不是一种限制。

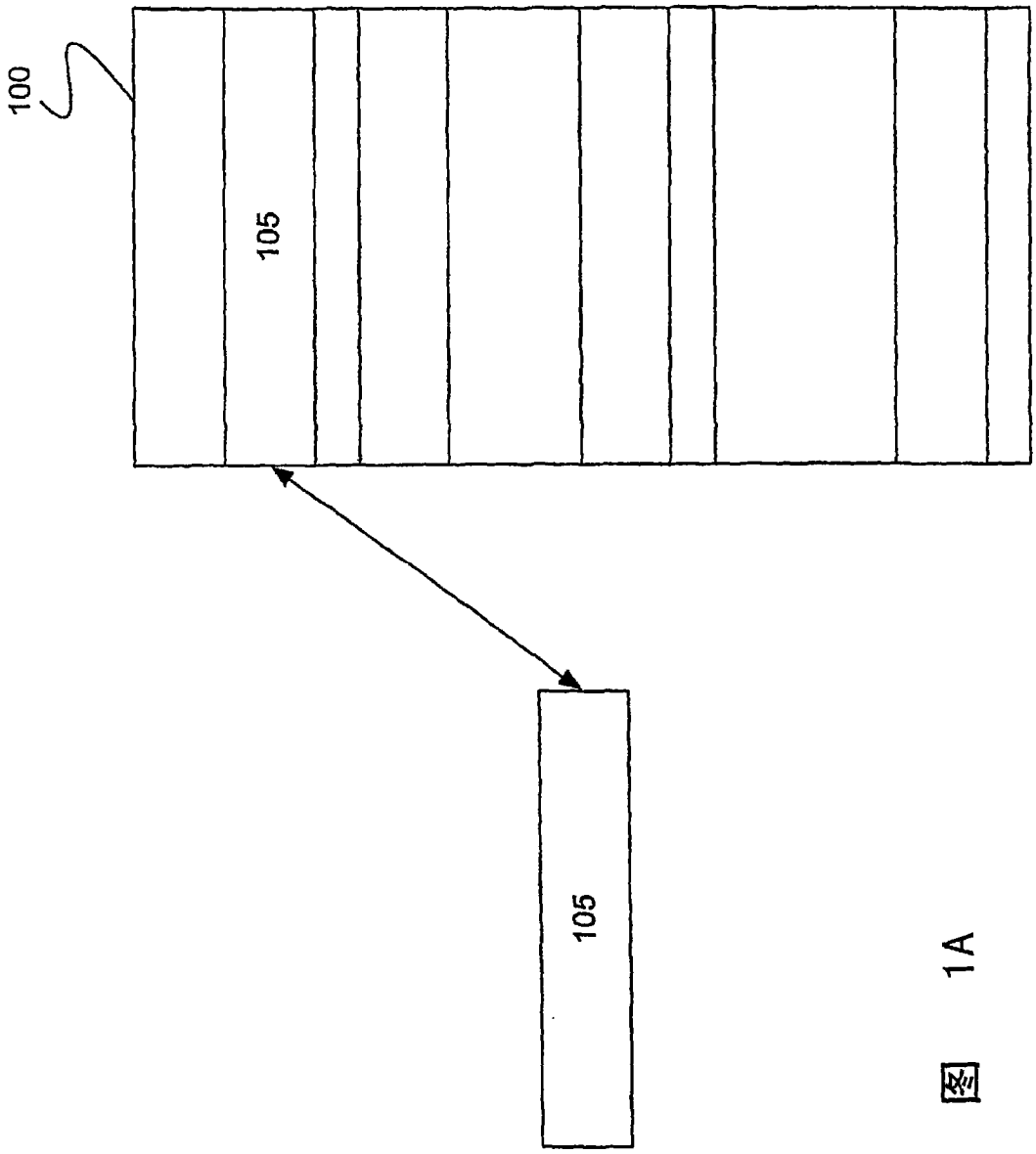


图 1A

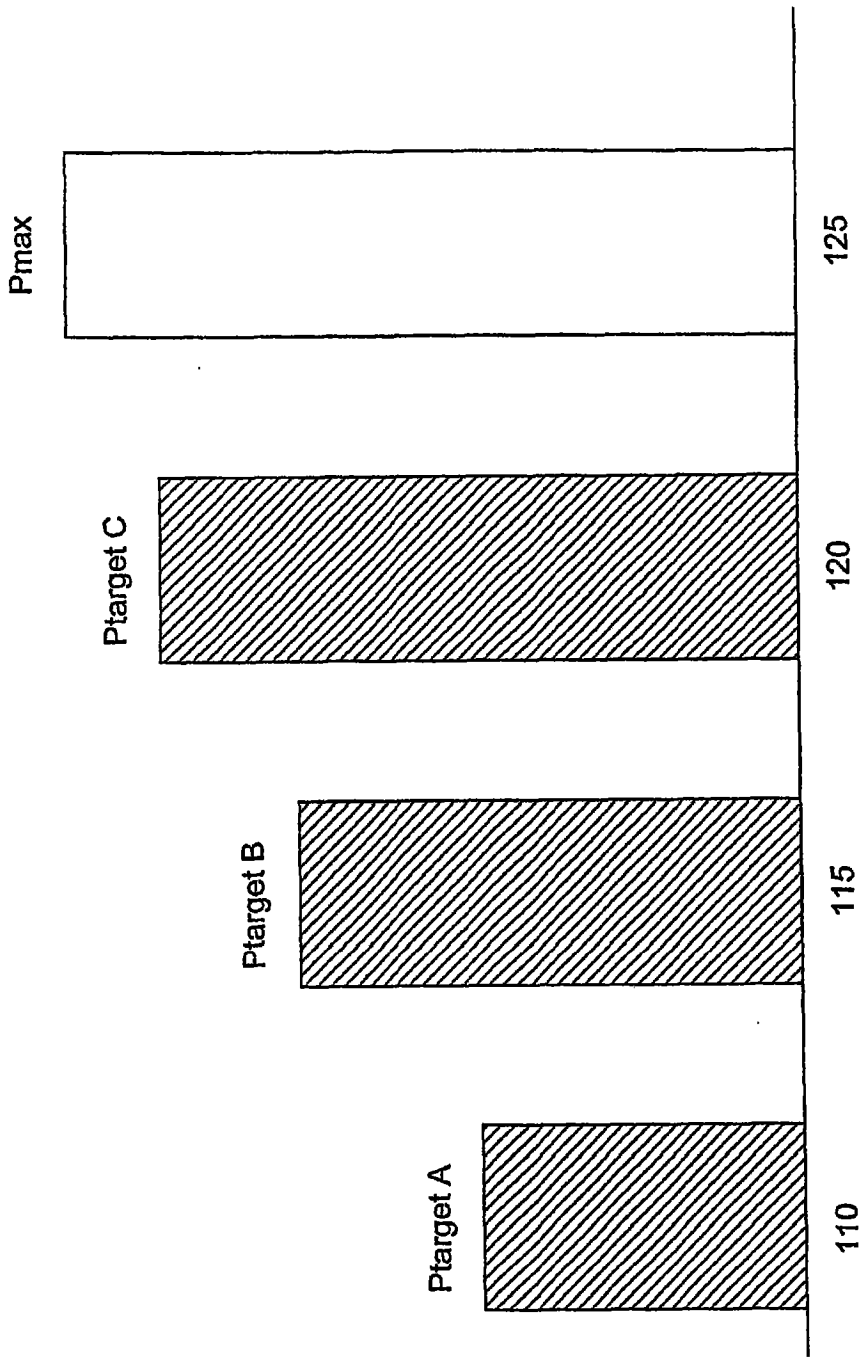


图 1B

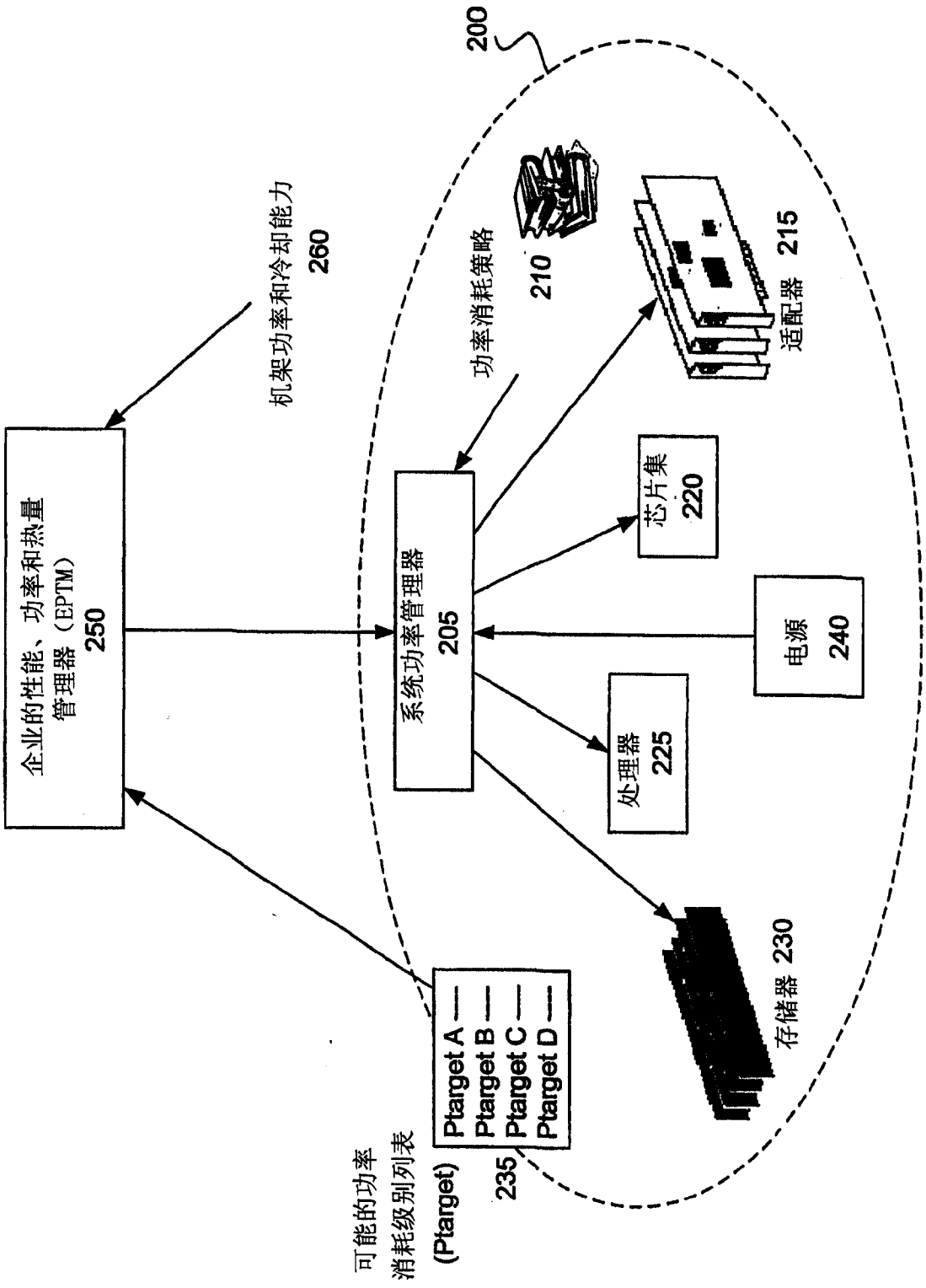

200
机架功率和冷却能力
260
企业的性能、功率和热量
管理器（EPTM）
250
功率消耗策略
210
系统功率管理器
205
芯片集
220
适配器
215
电源
240
处理器
225
存储器
230
可能的功率
消耗级别列表
(Ptarget)
Ptarget A
Ptarget B
Ptarget C
Ptarget D
235

图 2

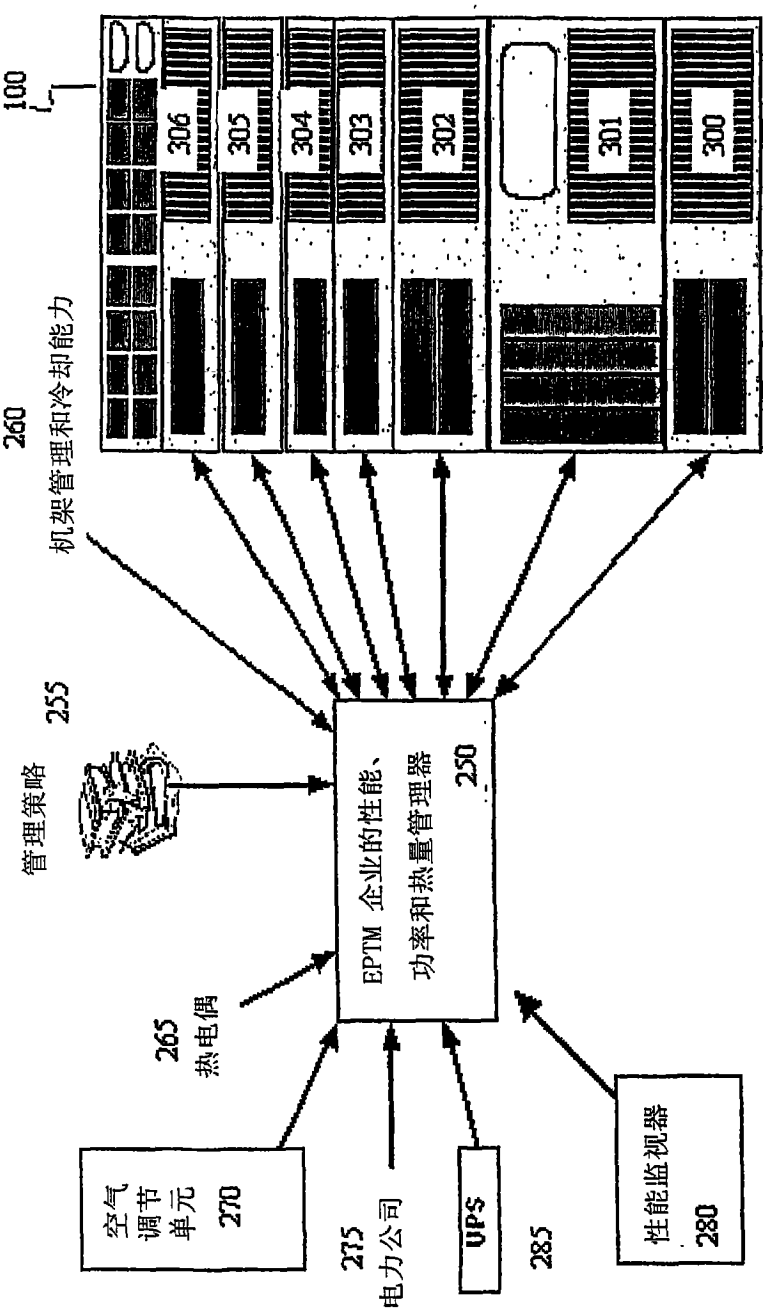

100
306
305
304
303
302
301
300
260
机架管理和冷却能力
255
管理策略
EPTM 企业的性能、
功率和热量管理器
250
265
热电偶
空气
调节
单元
270
275
电力公司
UPS
285
性能监视器
280

图 3

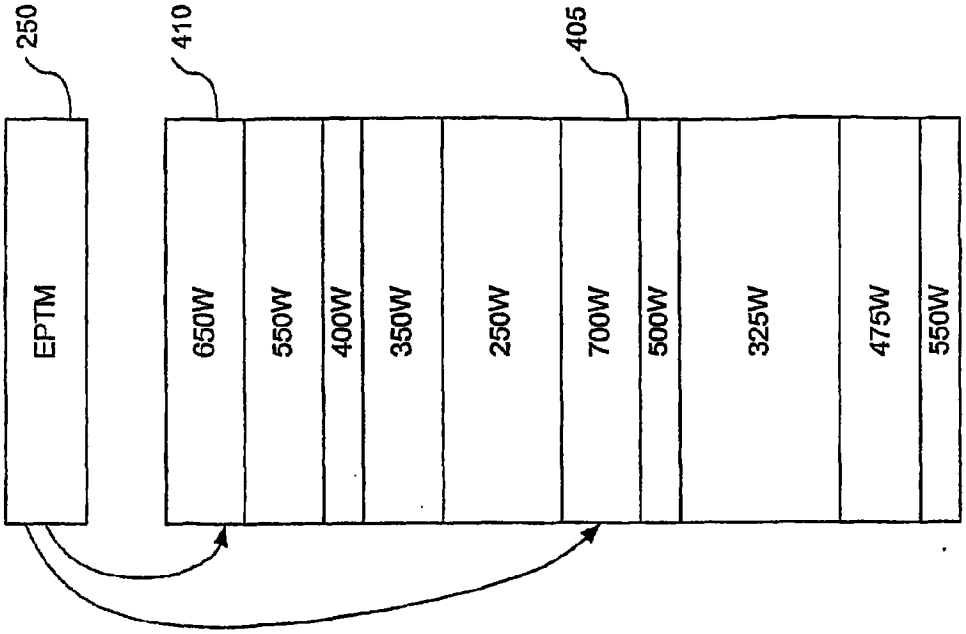

250
410
405
EPTM
650W
550W
400W
350W
250W
700W
500W
325W
475W
550W

图 4B

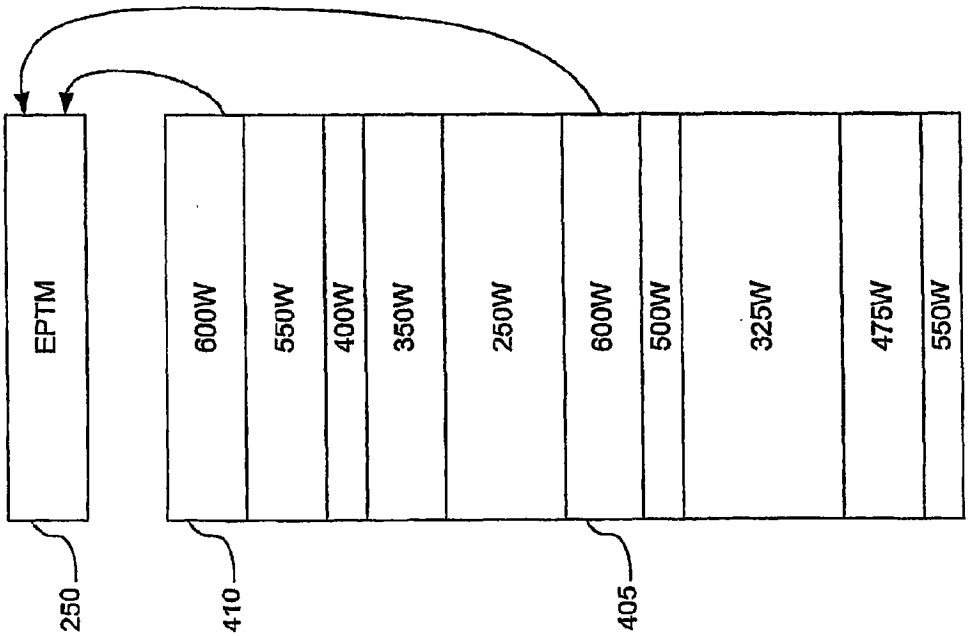

EPTM
600W
550W
400W
350W
250W
600W
500W
325W
475W
550W
250
410
405

图 4A

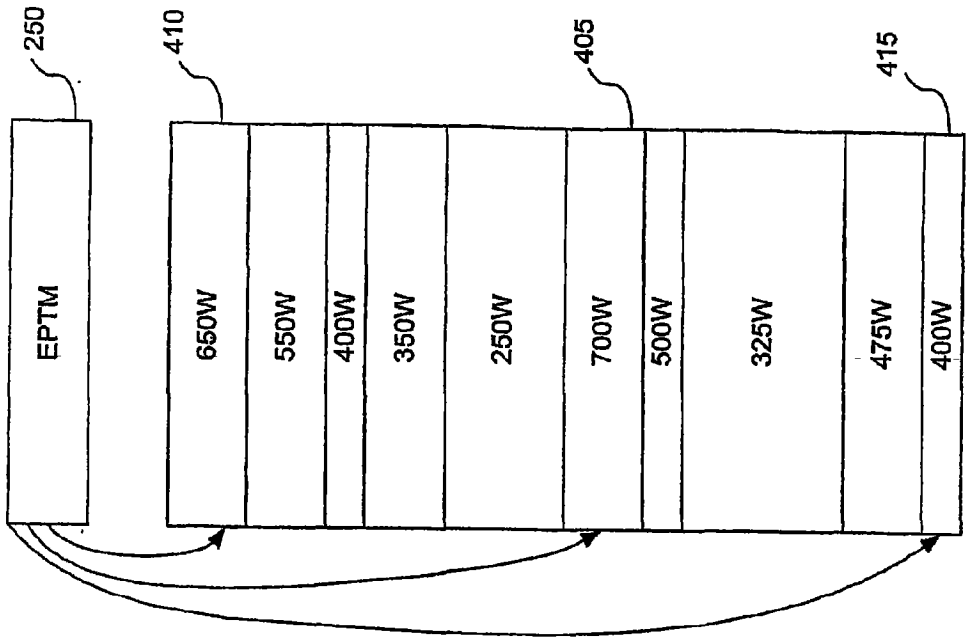

250
410
405
415
EPTM
650W
550W
400W
350W
250W
700W
500W
325W
475W
400W

图 5B

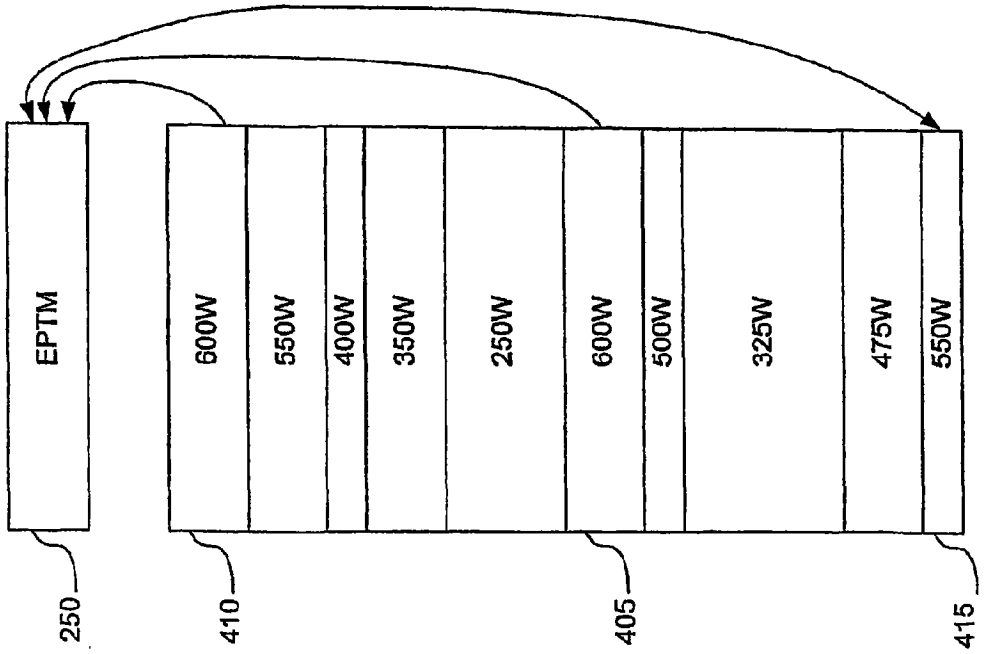

EPTM
600W
550W
400W
350W
250W
600W
500W
325W
475W
550W
250
410
405
415

图 5A

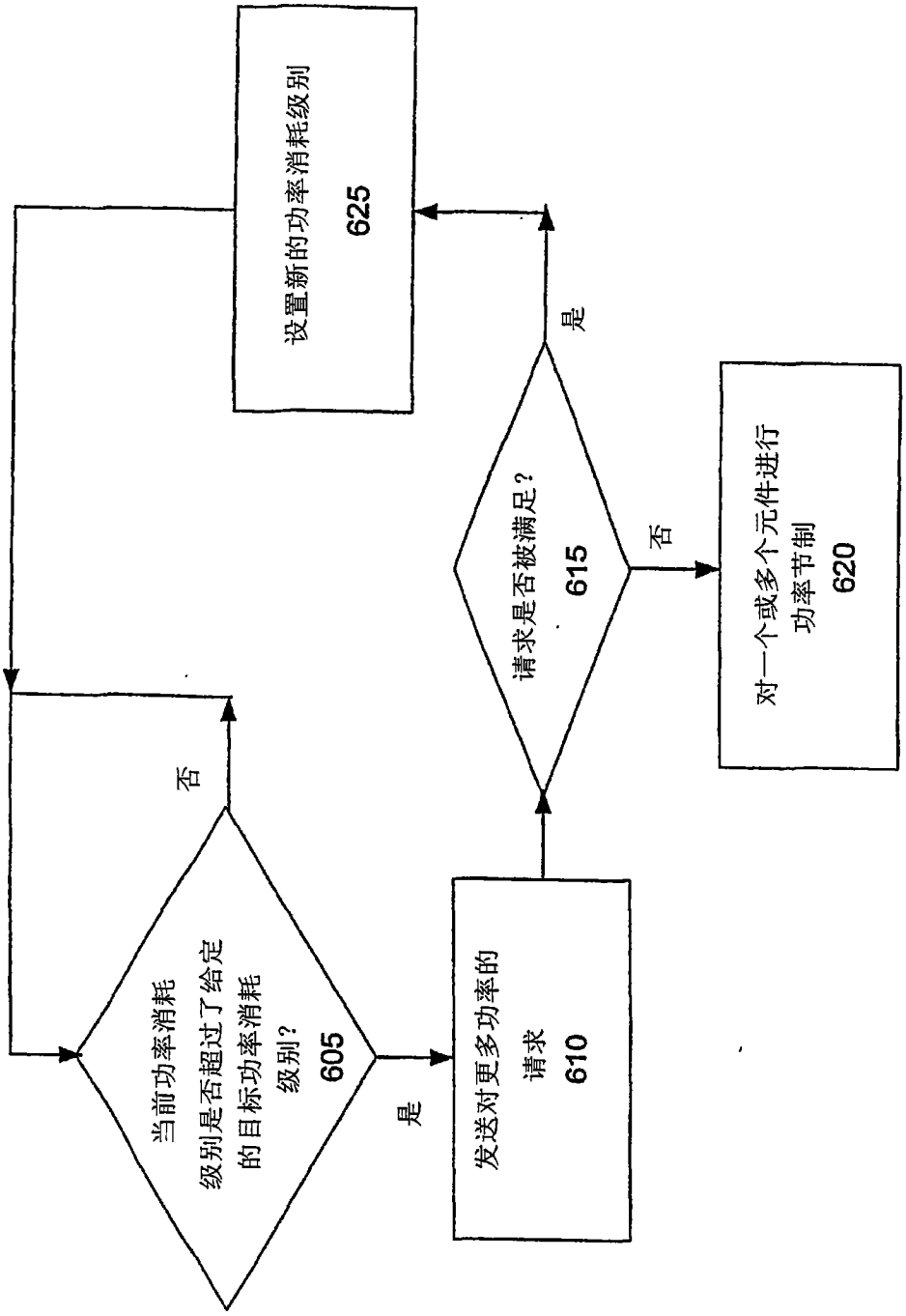

当前功率消耗级别是否超过了给定的目标功率消耗级别？
605
否
是
发送对更多功率的请求
610
请求是否被满足？
615
否
是
对一个或多个元件进行功率节制
620
设置新的功率消耗级别
625

图 6

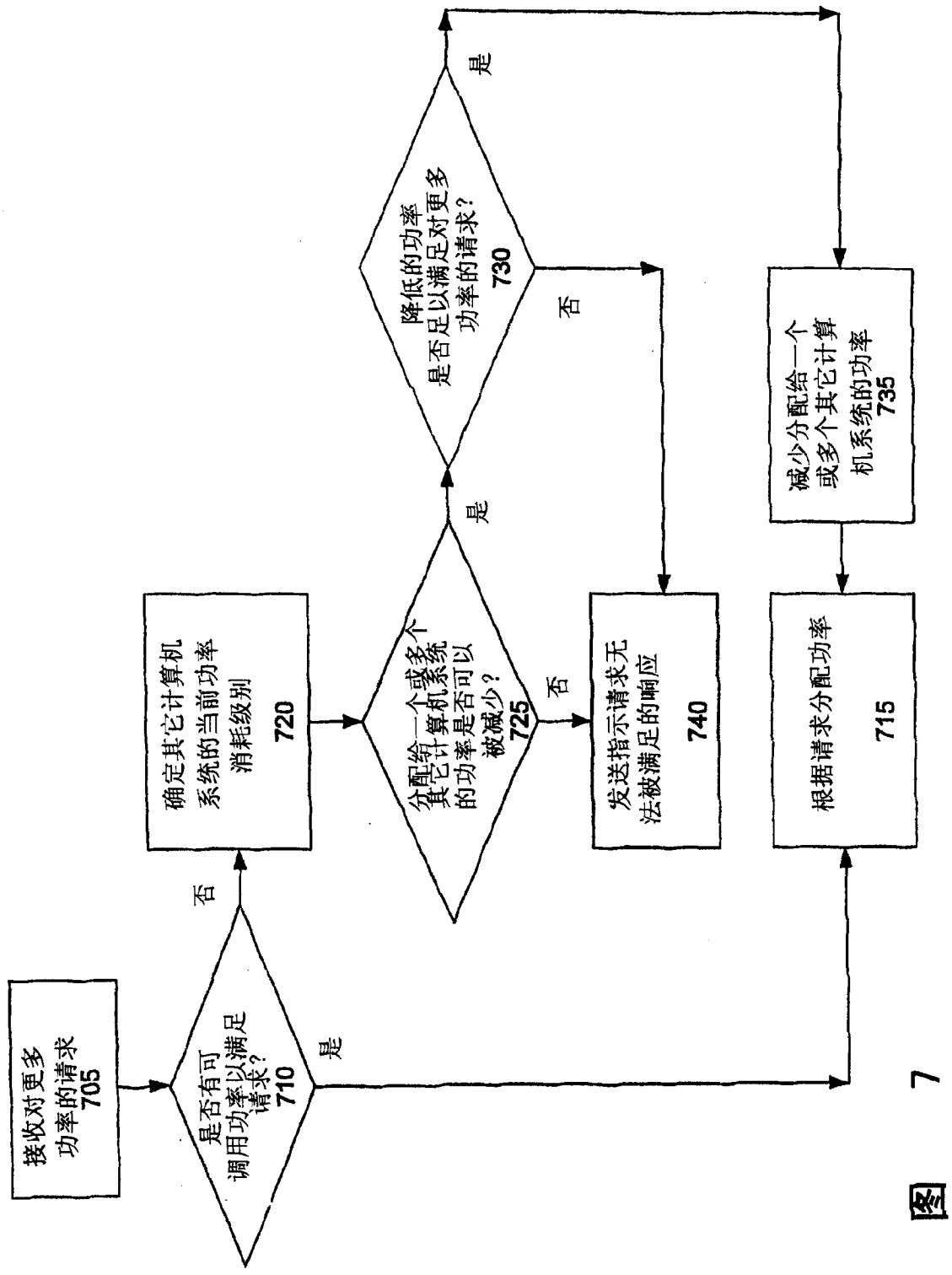

接收对更多功率的请求 705
是否有可调用功率以满足请求? 710
否
是
确定其它计算机系统的当前功率消耗级别 720
分配给一个或多个其它计算机系统的功率是否可以被减少? 725
否
是
降低的功率是否足以满足对更多功率的请求? 730
否
是
发送指示请求无法被满足的响应 740
根据请求分配功率 715
减少分配给一个或多个其它计算机系统的功率 735

图 7

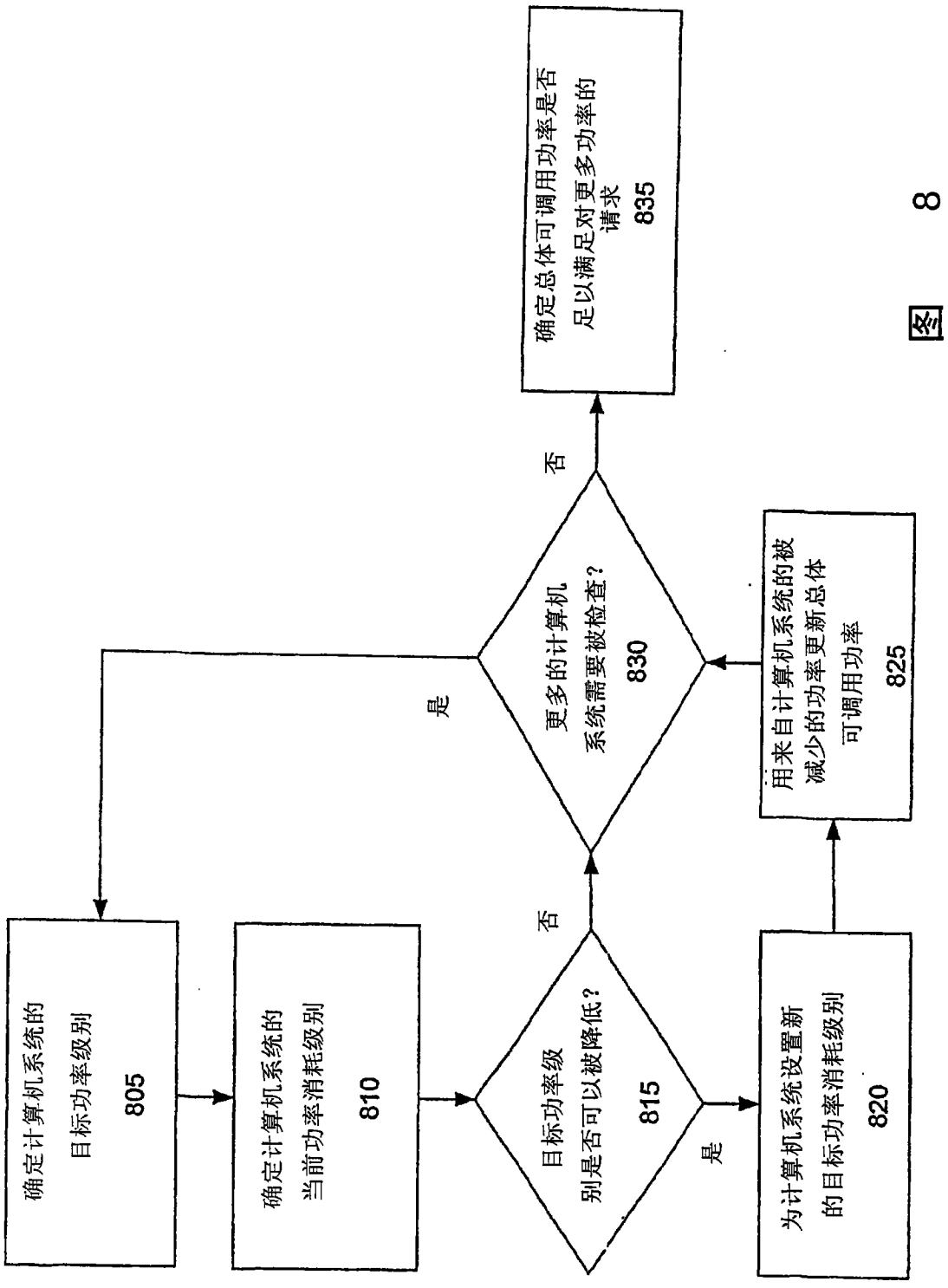

确定计算机系统的
目标功率级别
805
确定计算机系统的
当前功率消耗级别
810
目标功率级
别是否可以被降低?
815
否
是
为计算机系统设置新
的目标功率消耗级别
820
用来自计算机系统的被
减少的功率更新总体
可调用功率
825
更多的计算机
系统需要被检查?
830
是
否
确定总体可调用功率是否
足以满足对更多功率的
请求
835

图 8