



(12) 发明专利申请

(10) 申请公布号 CN 102326132 A

(43) 申请公布日 2012. 01. 18

(21) 申请号 201080008247. 8

代理人 于静 杨晓光

(22) 申请日 2010. 02. 15

(51) Int. Cl.

(30) 优先权数据

G06F 1/32 (2006. 01)

12/393, 475 2009. 02. 26 US

(85) PCT申请进入国家阶段日

2011. 08. 18

(86) PCT申请的申请数据

PCT/EP2010/051856 2010. 02. 15

(87) PCT申请的公布数据

W02010/097306 EN 2010. 09. 02

(71) 申请人 国际商业机器公司

地址 美国纽约

(72) 发明人 J·博策克 E·苏费尔

J·伍尔德里奇

(74) 专利代理机构 北京市中咨律师事务所

11247

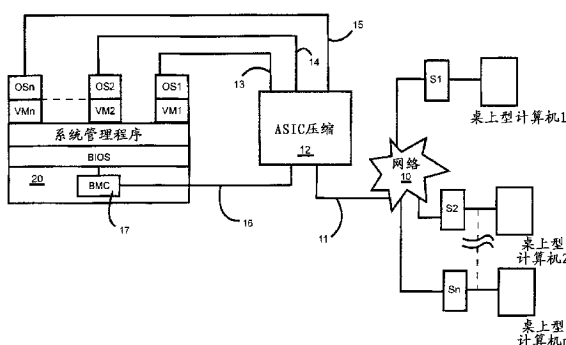
权利要求书 2 页 说明书 6 页 附图 4 页

(54) 发明名称

最大化虚拟机平台的降低功率状态的电源管理

(57) 摘要

本发明揭示虚拟机环境中的电源管理, 所述虚拟机环境包括提供多个虚拟机的至少一个虚拟机平台以及多个单独 (用户) 计算机, 每个单独计算机连接到典型虚拟机分布环境内的虚拟机中的一个相应虚拟机。本发明还提供: 用于独立于用户计算机至虚拟机的连接, 判定连接到所述虚拟机的计算机中的每个计算机是否处于活动状态的功能; 以及用于在连接到所述虚拟机的所有计算机均处于非活动状态时, 将所述虚拟机平台切换至平台中的降低功耗状态的功能。



1. 一种在虚拟机环境中的电源管理系统,包含:
虚拟机平台,所述虚拟机平台提供多个虚拟机;
多个单独客户端设备,每个客户端设备都连接到所述虚拟机中的一个相应虚拟机;
判定装置,用于独立于所述客户端设备至所述虚拟机的连接而判定连接到所述虚拟机的所述客户端设备中的每个客户端设备是否处于活动状态;及
切换装置,用于在连接到所述虚拟机的所述客户端设备中的所有客户端设备均处于非活动状态时,将所述虚拟机平台切换至降低功耗状态。
2. 如权利要求 1 所述的电源管理系统,其中:
所述多个客户端设备远离所述虚拟机平台;及
还包括用于将所述多个客户端设备分别连接到所述虚拟机平台的网络装置。
3. 如权利要求 2 所述的电源管理系统,其中
所述网络装置是万维网网络装置。
4. 如权利要求 2 所述的电源管理系统,还包括:
多个所述虚拟机平台,每个平台分别提供多个虚拟机;及
多个集合,每个集合包括分别连接到所述多个虚拟机平台之一中的虚拟机的多个客户端设备;其中
当连接到所述虚拟机平台中的一个相应平台中的虚拟机的所有所述客户端设备均处于非活动状态时,所述切换装置将所述一个相应平台切换至降低功耗状态。
5. 如权利要求 4 所述的电源管理系统,其中:
所述多个虚拟机平台是布置为虚拟服务器场的服务器;及
所述多个客户端设备集合中的每个客户端设备集合是连接到相应服务器中的虚拟机的客户端计算机。
6. 如权利要求 2 所述的电源管理系统,其中:
用于将所述虚拟机平台切换至所述降低功率状态的所述切换装置在所述虚拟机平台中;及
还包括通信装置,用于通过独立于所述虚拟机平台的操作系统的路径,告知所述切换装置所述客户端设备中的每个客户端设备是否处于非活动状态。
7. 如权利要求 6 所述的电源管理系统,其中所述虚拟机平台包括系统管理程序和 BIOS;及
用于告知非活动状态的所述路径独立于所述系统管理程序而连接到所述 BIOS。
8. 如权利要求 7 所述的电源管理系统,其中:所述虚拟机平台包括底板管理控制器(BMC),所述 BMC 具有用于追踪所述客户端设备的非活动状态的装置;及
用于告知非活动状态的所述路径通过所述 BMC 而连接到所述 BIOS。
9. 一种在虚拟机平台中的电源管理方法,所述虚拟机平台包括多个虚拟机,每个虚拟机分别连接到多个单独客户端设备中的每个客户端设备,所述方法包含:
独立于所述客户端设备至所述虚拟机的连接,判定连接到所述虚拟机的所述客户端设备中的每个客户端设备是否处于活动状态;及
在连接到所述虚拟机的所述客户端设备中的所有客户端设备均处于非活动状态时,将所述虚拟机平台切换至降低功耗状态。

10. 如权利要求 9 所述的方法,其中:

所述多个客户端设备远离所述虚拟机平台;及
通过网络将所述多个客户端设备分别连接到所述虚拟机平台。

11. 如权利要求 10 所述的方法,其中所述网络是万维网。

12. 如权利要求 10 所述的方法,其中所述电源管理用于多个所述虚拟机平台,每个平台分别提供多个虚拟机,并且所述方法包括:

将每个集合都包括多个客户端设备的多个集合分别连接到所述多个虚拟机平台之一的虚拟机;

当连接到所述虚拟机平台中的一个相应平台中的虚拟机的所有所述客户端设备均处于非活动状态时,将所述一个相应平台切换至降低功耗状态。

13. 如权利要求 10 所述的方法,其中:

在所述虚拟机平台中控制所述虚拟机平台至所述降低功耗状态的切换;及

还包括用于通过独立于所述虚拟机平台的操作系统的路径,告知所述虚拟机平台所述客户端设备中的每个客户端设备是否处于非活动状态的步骤。

14. 如权利要求 13 所述的方法,其中用于告知非活动状态的所述路径独立于所述虚拟机平台的系统管理程序而连接到所述虚拟机平台的 BIOS。

15. 一种包括计算机程序代码的计算机程序,当所述计算机程序代码被加载到计算机中并在其上执行时,将导致所述计算机执行如权利要求 9 至 14 中的任一权利要求所述的方法的所有步骤。

最大化虚拟机平台的降低功率状态的电源管理

技术领域

[0001] 本发明涉及包括多个虚拟机的虚拟机平台,各虚拟机分别连接到多个单独计算机中的每个计算机,本发明尤其涉及此类虚拟机平台的电源管理。

背景技术

[0002] 过去十年中,随着万维网(网站)系统大幅增加,组织所需的计算机处理功率每年按指数方式增长,因此如今需要使用上百个及甚至上千个服务器。这导致越来越大的主机计算机的复苏。尤其是在虚拟机(VM)模式中工作的主机及类似的大型计算机,在VM模式中,操作系统的多个实例及相关的应用程序驻留在相同的物理硬件中。此类虚拟机一直在满足通常配置为虚拟机服务器场的大量服务器的需求。有关更多背景,请参考论文:“z/VM下的虚拟Linux服务器:安全性、性能、及管理问题(Virtual Linux servers under z/VM: security, performance, and administrative issues)”,作者D. Turk,发表于IBM系统期刊(IBM Systems Journal),2005年7月;及请参考论文:“更多功率,扩展虚拟化管理器能力帮助客户成长及管理虚拟化环境(More POWER to Ya, Expanded Virtualization Manager capabilities help customers grow and manage virtualized environments)”,作者Jim Fall,发表于IBM系统杂志(IBM Systems Magazine),2007年9月。

[0003] 在此类虚拟机环境中(其中多个用户计算机连接到每个虚拟机平台,每个平台提供分别连接到这些多个用户的多个虚拟机),很难控制电源管理。在启动虚拟机会话之前,每个虚拟机平台均需要使用全功率。由于虚拟机平台对于需要访问适当虚拟机的远程用户计算机必须始终可供使用,以全功率模式持续操作任何在线平台已司空见惯。在虚拟机平台被配置为虚拟机服务器场时,浪费功耗变得特别严重。

发明内容

[0004] 本发明在第一方面中相应地提供了一种在虚拟机环境中的电源管理系统,包含:虚拟机平台,所述虚拟机平台提供多个虚拟机;多个单独客户端设备,每个客户端设备都连接到所述虚拟机中的一个相应虚拟机;判定装置,用于独立于所述客户端设备至所述虚拟机的连接而判定连接到所述虚拟机的所述客户端设备中的每个客户端设备是否处于活动状态;及切换装置,用于在连接到所述虚拟机的所述客户端设备中的所有客户端设备均处于非活动状态时,将所述虚拟机平台切换至降低功耗状态。

[0005] 优选地,所述多个客户端设备远离所述虚拟机平台;及还包括用于将所述多个客户端设备分别连接到所述虚拟机平台的网络装置。优选地,所述网络装置是万维网网络装置。所述系统优选地还包括多个所述虚拟机平台,每个平台分别提供多个虚拟机;及多个集合,每个集合包括分别连接到所述多个虚拟机平台之一中的虚拟机的多个客户端设备;其中当连接到所述虚拟机平台中的一个相应平台中的虚拟机的所有所述客户端设备均处于非活动状态时,所述切换装置将所述一个相应平台切换至降低功耗状态。

[0006] 优选地,所述多个虚拟机平台是布置为虚拟服务器场的服务器;及所述多个客户

端设备集合中的每个客户端设备集合是连接到相应服务器中的虚拟机的客户端计算机。优选地,用于将所述虚拟机平台切换至所述降低功率状态的所述切换装置在所述虚拟机平台中;及还包括通信装置,用于通过独立于所述虚拟机平台的操作系统的路径,告知所述切换装置所述客户端设备中的每个客户端设备是否处于非活动状态。优选地,所述虚拟机平台包括系统管理程序和 BIOS;及用于告知非活动状态的所述路径独立于所述系统管理程序而连接到所述 BIOS。优选地,所述虚拟机平台包括底板管理控制器(BMC),所述 BMC 具有用于追踪所述客户端设备的非活动状态的装置;及用于告知非活动状态的所述路径通过所述 BMC 而连接到所述 BIOS。

[0007] 在第二方面中,提供了一种在虚拟机平台中的电源管理方法,所述虚拟机平台包括多个虚拟机,每个虚拟机分别连接到多个单独客户端设备中的每个客户端设备,所述方法包含:独立于所述客户端设备至所述虚拟机的连接,判定连接到所述虚拟机的所述客户端设备中的每个客户端设备是否处于活动状态;及在连接到所述虚拟机的所述客户端设备中的所有客户端设备均处于非活动状态时,将所述虚拟机平台切换至降低功耗状态。

[0008] 优选地,所述多个客户端设备远离所述虚拟机平台;及通过网络将所述多个客户端设备分别连接到所述虚拟机平台。优选地,所述网络是万维网。优选地,所述电源管理用于多个所述虚拟机平台,每个平台分别提供多个虚拟机,并且所述方法包括:将每个集合都包括多个客户端设备的多个集合分别连接到所述多个虚拟机平台之一中的虚拟机;其中当连接到所述虚拟机平台中的一个相应平台中的虚拟机的所有所述客户端设备均处于非活动状态时,将所述一个相应平台切换至降低功耗状态。优选地,在所述虚拟机平台中控制所述虚拟机平台至所述降低功耗状态的切换;及还包括用于通过独立于所述虚拟机平台的操作系统的路径,告知所述虚拟机平台所述客户端设备中的每个客户端设备是否处于非活动状态的步骤。优选地,用于告知非活动状态的所述路径独立于所述虚拟机平台的系统管理程序而连接到所述虚拟机平台的 BIOS。

[0009] 在第三方面中,提供了一种包括计算机程序代码的计算机程序,当所述计算机程序代码被加载到计算机中并在其上执行时,将导致所述计算机执行第二方面所述的方法的所有步骤。

[0010] 所述第三方面的计算机程序的形式可以体现为其上存储有计算机可读程序的计算机可用介质,所述计算机可读程序用于虚拟机平台中的电源管理,所述虚拟机平台包括多个虚拟机,每个虚拟机分别连接到多个单独客户端设备中的每个客户端设备,其中所述计算机可读程序,当在计算机上执行时,导致所述计算机执行以下操作:独立于所述客户端设备至所述虚拟机的连接,判定连接到所述虚拟机的所述客户端设备中的每个客户端设备是否处于活动状态;及在连接到所述虚拟机的所述客户端设备中的所有客户端设备均处于非活动状态时,将所述虚拟机平台切换至降低功耗状态。

[0011] 优选地,所述多个客户端设备远离所述虚拟机平台;及通过网络将所述多个客户端设备分别连接到所述虚拟机平台。优选地,所述网络是万维网。优选地,所述电源管理用于多个所述虚拟机平台,每个平台分别提供多个虚拟机,并且所述计算机程序在执行时:将每个集合都包括多个客户端设备的多个集合分别连接到所述多个虚拟机平台之一中的虚拟机;并且当连接到所述虚拟机平台中的一个相应平台中的虚拟机的所有所述客户端设备均处于非活动状态时,将所述一个相应平台切换至降低功耗状态。

[0012] 优选地,所述计算机程序在执行时将导致所述计算机:在所述虚拟机平台中控制所述虚拟机平台至所述降低功耗状态的切换;及通过独立于所述虚拟机平台的操作系统的路径,告知所述虚拟机平台所述客户端设备中的每个客户端设备是否处于非活动状态。优选地,所述计算机程序在执行时将导致所述计算机:独立于所述虚拟机平台的系统管理程序而向所述虚拟机平台的 BIOS 告知非活动状态。

[0013] 本发明解决即使在没有任何连接到虚拟机平台的用户计算机处于活动模式时仍维持每个虚拟机平台处于全功率模式的功耗问题。

[0014] 为此目的,本发明的各实施例提供一种在虚拟机环境中进行电源管理的系统、方法、及计算机程序,所述虚拟机环境包括:提供多个虚拟机的至少一个物理虚拟机平台;及多个单独(用户)客户端设备,每个客户端设备连接到典型虚拟机分布环境中的虚拟机中的一个相应虚拟机。已知客户端设备包括用户计算机及计算机子系统,其包括打印机、盘驱动器及串行端口等。在本发明的以下描述中,在使用术语“用户计算机”时,旨在包括所有此类客户端设备。

[0015] 本发明还提供:用于独立于用户计算机或客户端设备至虚拟机的连接,判定连接到所述虚拟机的客户端设备中的每个客户端设备是否处于活动状态的功能;以及用于在连接到所述虚拟机的所有用户计算机均处于非活动状态时,将所述虚拟机平台切换至平台中的降低功耗状态的功能。

[0016] 在优选操作中,本发明实施于多个用户计算机远离虚拟机平台时,及其中存在网络(通常为因特网或网站,这两个术语在本说明中可交换使用)用于分别将多个用户计算机连接到虚拟机平台时。

[0017] 如上文所提,本发明的优选实施例在例如虚拟服务器场的系统中尤其有利,所述系统中具有:多个虚拟机平台,每个平台分别提供多个虚拟机;及多个用户集合,每个用户集合包括多个用户计算机或设备,其分别连接到多个虚拟机平台之一中的虚拟机;使得切换装置在连接到所述虚拟机平台中的一个相应平台中的虚拟机的所有用户计算机或客户端设备处于非活动状态时,将所述一个相应平台切换至降低功耗状态。此类实施可有效用于多个虚拟机平台提供布置为虚拟服务器场的服务器时,及多个用户计算机集合中的每个集合为连接到相应服务器中的虚拟机的客户端设备或用户计算机时。

[0018] 如将于下文所详细说明的,在本发明的一个优选实施例中,将虚拟机平台切换至降低功率状态的实施是在虚拟机平台中,及用于传达所述计算机中的每个计算机是否处于非活动状态的装置通过独立于虚拟机平台的操作系统的路径而连接到用于切换的装置。此用于传达非活动状态的路径通过独立于平台的系统管理程序的路线而连接到虚拟机平台的基本输入输出系统(BIOS)固件。在此类布置中,虚拟机平台包括具有追踪所连接用户计算机的非活动状态的功能的底板管理控制器(BMC),及用于传达非活动状态的路径通过所述 BMC 而连接到所述 BIOS。

附图说明

[0019] 现在将仅通过实例的方式参考附图描述本发明的优选实施例,这些附图是:

[0020] 图 1 是网络部分的总体示意图,其包括单个代表性虚拟机平台及连接到平台中的虚拟机的远程用户计算机集合或客户端设备集合,用以图示用户计算机的非活动状态监

视；

[0021] 图 2 是与图 1 的示意图相似的示意图，其显示分别连接到多个用户计算机集合的多个虚拟机平台；

[0022] 图 3 是设定为可实施本发明的一个优选实施例以在虚拟机环境中进行电源管理的程序的总体流程图；及

[0023] 图 4 是图 3 中设定的程序的示例性运行的流程图。

具体实施方式

[0024] 参考图 1，其中显示网络部分的总体示意图，包括：单个代表性虚拟机平台 20，及远程用户计算机（即，客户端设备）集合：桌上型计算机 1、2 及 n，它们经由网络 10（如网站）和相应服务器 S1、S2 及 Sn 而连接到平台中使用操作系统 OS1、OS2 及 OSn 的虚拟机 VM1、VM2 及 VMn。

[0025] 特定虚拟机 VM 到特定用户计算机（桌上型计算机 1 至 n）的指派由系统管理程序 21（虚拟机平台管理程序）分配。系统管理程序管理多个操作系统 OS1-OSn，以及视需要管理平台的处理器、BIOS 22、存储器及其它资源。由于这是虚拟机平台，因此可将系统管理程序视为虚拟化管理器。当前常规 VM 连接代理（未显示）可操作以将当前在平台 20 上运行的虚拟机指派给特定客户端设备（如用户桌上型计算机 1 至 n 所示）。然而，此类连接代理仅在登录及中断连接过程中使用。因此，如果特定的桌上型计算机 1、2 或 n 登录并且然后闲置，则虚拟机平台 20 中特定的被指派 VM 及其相关资源仍维持全功率，而与指派的用户桌上型计算机 1 至 n 计算机的非活动状态无关。

[0026] 据此，本发明提供一种独立于平台 20 的操作系统及系统管理程序 21 而检测分配的用户桌上型计算机 1-n 等计算机的非活动状态的实施方式。因此，在图 1 的环境中，系统管理程序（可购自 VMware、Citrix 或 Calista 环境）具有配合硬件压缩引擎工作的个体软件组件，以允许多个用户（桌上型计算机 1-n）访问同一平台 20。因此，当平台硬件接通电源时，系统管理程序分配用户专用的存储器资源。当用户不再需要使用资源时，如注销时，资源被释放回平台池中。然而，尽管不再需要使用平台，平台 20 仍维持完全电源开启及完全工作。

[0027] 如以下所详细描述的，本发明提供用于当用户计算机或客户端设备不再需要使用虚拟机资源（处理器、存储器、操作系统及应用软件）时，即，所有用户计算机或客户端设备处于非活动状态时，将平台置于较低功耗状态；但当用户计算机或客户端设备确实需要使用平台所支持的 VM 时，即，用户计算机变成活动时，平台 20 将退出低功率状态并变成完全工作。

[0028] 本发明的实施方式利用现有设备来改变虚拟机平台 20 的功率状态。BIOS 22 提供能够从平台 20 的不同部分移除电源的工业标准高级配置与电源接口（ACPI）状态。

[0029] 在本发明的实施方式中，存在通过网络 10（可为因特网或网站）经由连接 11 且通过 ASIC（专用集成电路）压缩核心 12 的路径，其中用户桌上型计算机 1-n 可分别连接到它们被分配的 VM 1-n 及操作系统 OS1-n。这由系统管理程序 21 控制。ASIC 压缩核心使得来自用户计算机（桌上型计算机 1-n）的数据能够被适当地压缩，从而以最有效率且与虚拟机平台相关联的方式存储数据以供虚拟机使用；图中所示经由连接 13、14 及 15 而施加数

据。标准 ASIC 核心在以下刊物中描述：“ASIC 核心中的数据压缩技术 (Data compression technology in ASIC cores)”，作者 S. H. Burroughs 等人，IBM 研发期刊 (IBM Journal of Research and Development)，第 42 卷，第 6 号，1998 年。

[0030] 定制 ASIC 核心具有充分灵活性，致使可形成从 ASIC 压缩核心 12 通过平台 20 中的底板管理控制器 (BMC) 17 直接进入 BIOS 22 的连接 16。BMC 是嵌入常规服务器计算机底板的常规专用微型控制器，其可用作使用常规智能平台管理接口 (IPMI) 的智能单元。因此，BMC 可用于管理在系统管理软件与平台硬件之间的接口。

[0031] 如以下关于本发明的程序 (可存储在系统管理程序 21 中) 所详细描述的，BMC 17 追踪在虚拟机平台 20 的环境中的用户数。当追踪到在客户端设备 (用户桌上型计算机) 及其被分配的 VM 之间的会话已在一段预定时间内处于非活动状态时，BMC 将用户计算机自活动池中移除。每个客户端设备或用户计算机每次进入或离开非活动状态时，均被记录下来。在参考图 4 所说明的一个实施例中，利用标准 ASIC 核心中所提供的嵌入活动定时器。BMC 在特定用户计算机的活动停止时设定定时器。然后，在时间期满后，从位于 BMC 中的活动追踪表，通过对 BIOS 22 的输入来移除不活动的用户计算机。以此实施方式，用户计算机或客户端设备的活动状态及非活动状态的追踪及传达独立于虚拟机平台的操作系统和系统管理程序而实行，因而允许独立控制虚拟机平台的低功率状态。

[0032] 现在参考图 2，其中显示电源管理的实施方式，所述实施方式应用于多个虚拟机平台 (A-N) 分别连接到多个用户计算机 (客户端设备) 集合的系统。此类布置用在虚拟服务器场所使用的虚拟机环境中。在此类布置中，每个平台 A-N 均可按照针对图 1 的虚拟机平台所描述的方式独立工作。为了便于图示，平台 N 中的组件 111 至 117 和 120 至 122 对应于平台 A 中的组件 11 至 17 和 20 至 22，且与以上针对平台 A 中的组件 11 至 17 和 20 至 22 所述的相同方式工作。

[0033] 图 3 是显示根据本发明的在虚拟机环境中进行电源管理的过程发展的流程图。在多虚拟机平台环境中，提供其中每个虚拟机平台具有多个虚拟机的实施方式，虚拟机中的每个虚拟机均适于连接到多个用户计算机 (客户端设备) 之一，步骤 71。提供以经由网络将步骤 71 的用户计算机连接到虚拟机，步骤 72。提供以监视 (独立于步骤 71 的任何连接) 多个计算机中的每个计算机是否处于活动状态，步骤 73。提供以判定多个连接的计算机中的所有计算机是否均处于非活动状态，步骤 74。提供以响应于步骤 74 中所有连接的计算机均处于非活动状态的判定，将虚拟机平台切换至降低功耗状态，步骤 75。提供以在虚拟机平台内控制步骤 75 的切换，其中连接的计算机是否处于非活动状态的传达是通过独立于虚拟机平台的操作系统的路径，步骤 76。提供以将在步骤 76 中判定的非活动状态直接传达至虚拟机平台的 BIOS，步骤 77。

[0034] 参考图 4 的流程图，其说明图 3 中设定的过程的运行。初始地，步骤 81，系统已接通电源。相应地，存在连接的远程用户计算机 (客户端设备)，步骤 82。设定计数器 (步骤 83)：

[0035] 活动用户 = 远程用户 + 1

[0036] 然后在步骤 84 判定远程用户计算机活动中是否存在改变，步骤 84。如果是，则在步骤 85 判定远程用户活动是否中断 (或变更)，步骤 85。在此步骤中，用户可连接、中断连接或变成不活动 (如果用户的活动定时器已经逾时)。如果是，则用户活动中断，远程用户

已从活动状态中断连接,步骤 86。然后变更计数器(步骤 87):

[0037] 活动用户 = 活动用户 - 1

[0038] 然后判定(步骤 88):

[0039] 活动用户 = 0, 步骤 88

[0040] 如果否,则流程分支返回步骤 84,然后过程继续。如果是,则步骤 89,虚拟机平台中存在 BIOS/ACIP 中断,以允许系统进入低功率的“S3 休眠”状态。BIOS 接着建议平台的系统管理程序保存上下文,步骤 90。在此低功耗休眠状态中,继续监视远程用户计算机的活动以查看是否变更回远程用户活动状态,步骤 90。如果是,则在步骤 91,虚拟机平台中存在 BIOS/ACIP 变更,从而进入高活动电源接通“S5”全功率状态。BIOS 接着建议平台的系统管理程序恢复上下文,步骤 92。

[0041] 虽已显示及说明特定优选实施例,但是本领域的技术人员将理解,可以在其中做出许多变更和修改。

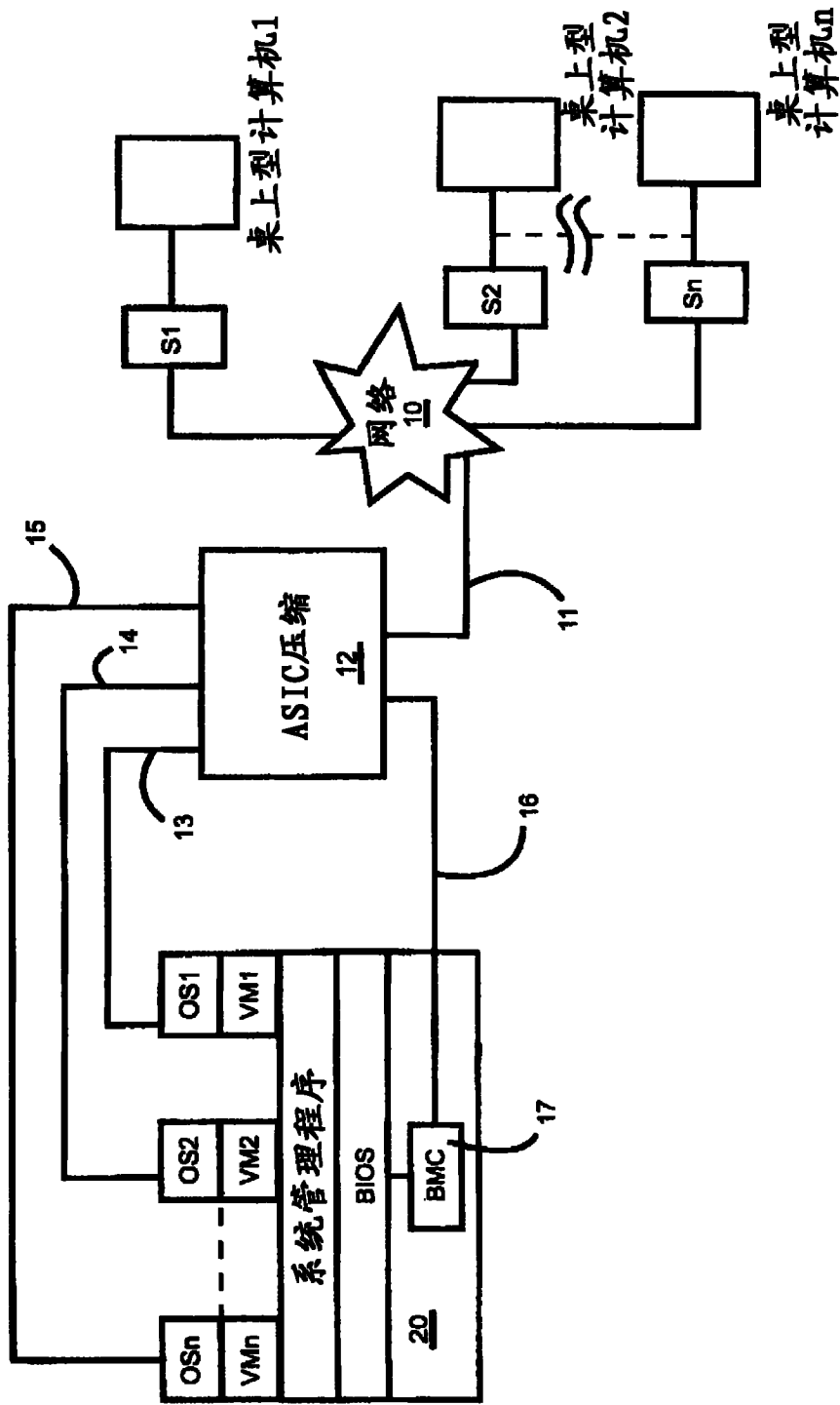


图 1

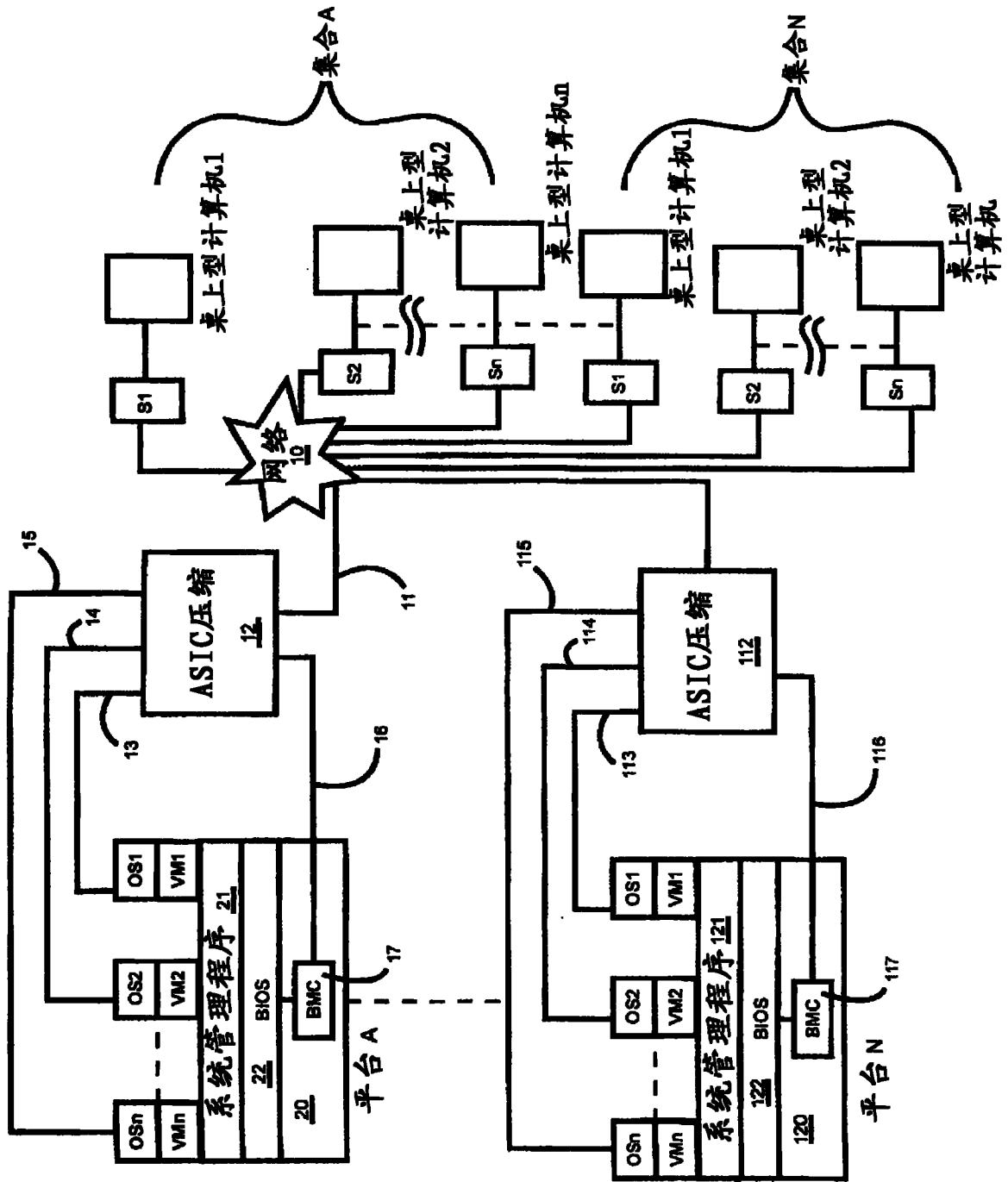


图 2

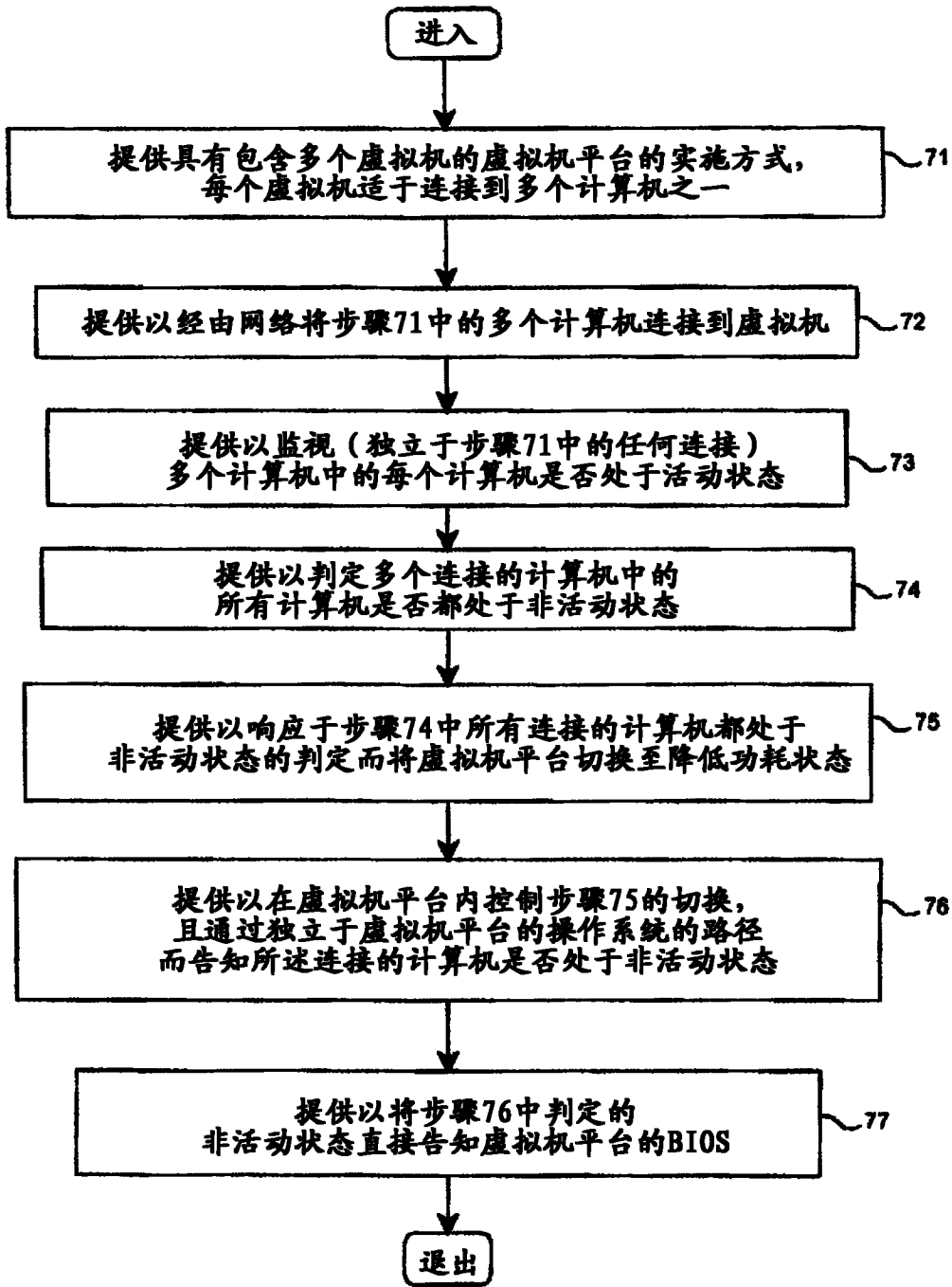


图 3

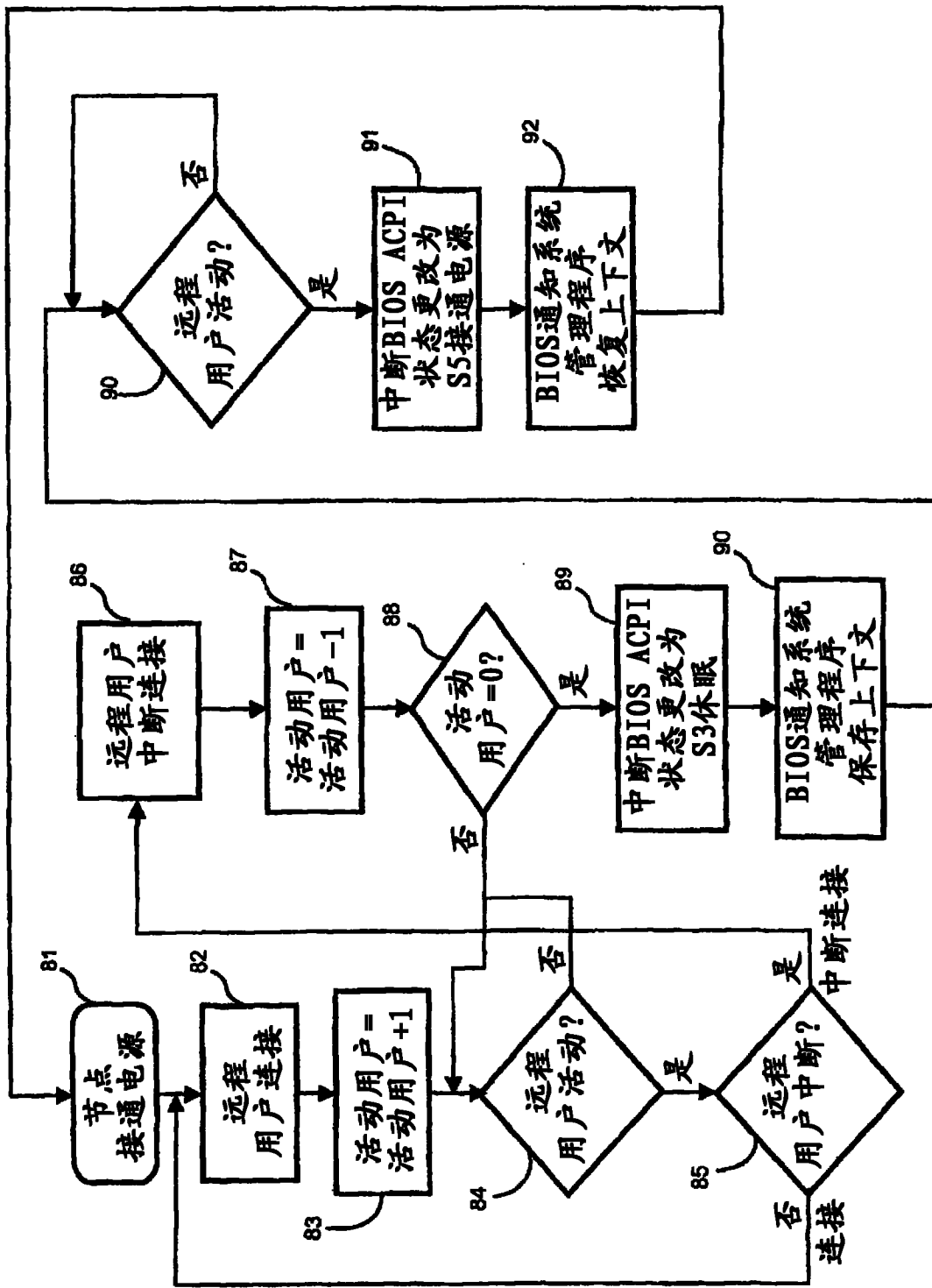


图 4