(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 105103142 A
(43)申请公布日 2015.11.25

(21)申请号 201380075221.9

(22)申请日 2013.03.29

(85)PCT国际申请进入国家阶段日
2015.09.29

(86)PCT国际申请的申请数据
PCT/US2013/034532 2013.03.29

(87)PCT国际申请的公布数据
WO2014/158181 EN 2014.10.02

(71)申请人 惠普发展公司,有限责任合伙企业
地址 美国德克萨斯

(72)发明人 B·S·巴齐尔 A·布朗
J·K·弗朗科姆 M·斯特恩斯
C·V·华 D·J·赛普利斯
P·汉森

(74)专利代理机构 永新专利商标代理有限公司
72002
代理人 刘瑜 王英

(51)Int.Cl.
G06F 13/14(2006.01)
G06F 1/26(2006.01)
G06F 9/44(2006.01)

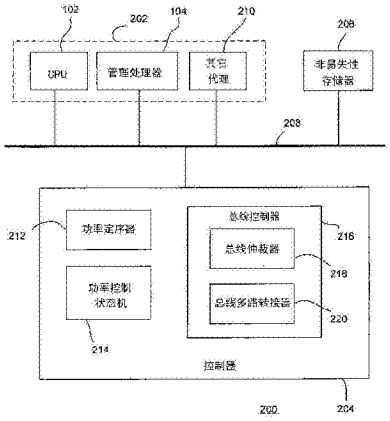
权利要求书2页 说明书4页 附图5页

(54)发明名称

在计算节点中的代理之间共享固件

(57)摘要

描述了在节点上包括多个中央处理单元(CPU)的多个代理之间共享固件。在示例中,计算节点包括:总线;耦合到总线的非易失性存储器,其用于存储针对多个代理的固件;功率定序器,其用于实现针对于多个CPU的加电顺序;多个功率控制状态机,其基于功率定序器的输出来分别控制多个CPU的状态;以及总线控制器,其用于基于多个功率控制状态机的状态将多个代理选择性地耦合到非易失性存储器。



CN 105103142 A

1. 一种用于在节点上包括多个中央处理单元 (CPU) 的多个代理之间共享固件的装置，包括：

总线；

耦合到所述总线的非易失性存储器，其用于存储针对所述多个代理的固件；

功率定序器，其用于实现针对所述多个 CPU 的加电顺序；

多个功率控制状态机，其基于所述功率定序器的输出来分别控制所述多个 CPU 的状态；

总线控制器，其用于基于所述多个功率控制状态机的状态来将所述多个代理选择性地耦合到所述非易失性存储器。

2. 如权利要求 1 所述的装置，其中，所述总线控制器包括：

总线仲裁器，其用于选择所述多个代理中的一个以用于与所述非易失性存储器进行通信；以及

总线多路复用器，其用于在所述非易失性存储器和由所述总线仲裁器选择的所述多个代理中的一个之间建立通信链路。

3. 如权利要求 1 所述的装置，其中，所述总线是串行数据总线。

4. 如权利要求 1 所述的装置，其中，所述多个代理进一步包括用于将所述固件的映象加载到所述非易失性存储器的管理代理。

5. 如权利要求 1 所述的装置，其中，所述多个功率控制状态机中的每一个将所述多个 CPU 中的相应的一个保持在重置状态，直到被所述功率定序器选择用于加电。

6. 一种在节点上包括被连接到总线的多个中央处理单元 (CPU) 的多个代理之间共享固件的方法，包括：

在被耦合到所述总线的非易失性存储器中存储针对所述多个代理的固件；

实现针对所述多个 CPU 的加电顺序；

基于所述加电顺序来控制所述多个 CPU 的状态；以及

基于所述多个 CPU 的状态来将所述多个代理选择性地耦合到所述非易失性存储器。

7. 如权利要求 6 所述的方法，其中，控制所述状态的步骤包括：

基于所述加电顺序来选择所述多个 CPU 中被允许加电的 CPU；

授权所述 CPU 访问所述非易失性存储器；

将所述多个 CPU 中的每个而不是所选择的 CPU 保持在重置状态；以及

针对所述多个 CPU 中的至少一个另外的 CPU 重复选择、授权和保持的步骤。

8. 如权利要求 6 所述的方法，进一步包括：

授权管理进程访问所述非易失性存储器以更新存储于其中的所述固件。

9. 如权利要求 6 所述的方法，进一步包括：

从所述多个代理中提出请求的代理接收要求访问所述非易失性存储器的请求；以及

基于所述请求来继而向所述提出请求的代理授权独占访问。

10. 如权利要求 6 所述的方法，其中，所述总线是串行数据总线。

11. 一种计算机系统，包括：

至少一个节点，包括：

包括多个中央处理单元 (CPU) 的多个代理；

总线；

耦合到所述总线的非易失性存储器，其用于存储针对所述多个代理的固件；以及

耦合到所述总线的集成电路，包括：

功率定序器，其用于实现针对所述多个 CPU 的加电顺序；

多个功率控制状态机，其基于所述功率定序器电路的输出来相应地控制所述多个 CPU 的状态；

总线控制器，其用于基于多个功率控制状态机电路的状态来将所述多个代理选择性地耦合到所述非易失性存储器。

12. 如权利要求 11 所述的计算机系统，其中，所述总线控制器包括：

总线仲裁器，其用于选择所述多个代理中的一个以用于与所述非易失性存储器进行通信；以及

总线多路复用器，其用于在所述非易失性存储器和由所述总线仲裁器选择的所述多个代理中的一个之间建立通信链路。

13. 如权利要求 11 所述的计算机系统，其中，所述总线是串行数据总线。

14. 如权利要求 11 所述的计算机系统，其中，所述多个代理进一步包括用于将所述固件的映象加载到所述非易失性存储器的管理代理。

15. 如权利要求 11 所述的计算机系统，其中，所述多个功率控制状态机中的每一个将所述多个 CPU 中的相应的一个保持在重置状态，直到被所述功率定序器选择用于加电。

在计算节点中的代理之间共享固件

背景技术

[0001] 计算机系统包括非易失性存储器以存储当被供电或“启动”时执行的第一代码。该种非易失性存储器可以被称为“固件”。该种固件的代码可以提供固件接口，例如，基本输入/输出系统(BIOS)、标准可扩展固件接口(UEFI)等。该固件的代码的至少一部分可以是可更新的。固件中可更新代码的当前状态被称为“映象”。因此，固件的当前映象可以用新的映象替换。固件更新过程可能涉及对固件的非易失性存储器的擦除和再编程。

[0002] 现代计算机通常具有多个处理器，所述多个处理器提供了相比于单个处理器系统改善的处理速度和性能。通常，系统中的每个处理器都具有专用固件，该专用固件使处理器能够加载操作系统(OS)。该专用固件被存储在针对每个处理器的单独的非易失性存储器中。为了升级固件，更新的固件需要被加载到每个处理器的每个存储器中。

附图说明

[0003] 参照下面的附图来描述本发明的一些实施例：

[0004] 图1是根据示例实现的计算节点的框图。

[0005] 图2是根据本发明的示例的图1的计算节点的固件子系统的框图。

[0006] 图3是描绘了根据本发明的示例的计算机系统的框图。

[0007] 图4是描绘了一种根据示例实现的在节点上多个代理之间共享固件的方法的流程图，所述多个代理包括被连接到总线的多个CPU。

[0008] 图5是描绘了根据本发明的示例的控制CPU状态的方法的流程图。

具体实施方式

[0009] 本文描述了在计算节点中的代理之间共享固件。在示例中，非易失性存储器耦合到总线以存储多个代理的固件，所述固件包括多个中央处理单元(CPU)。功率定序器实现针对多个CPU的加电顺序。多个控制状态机基于功率定序器的输出分别控制CPU的状态。总线控制器基于功率控制状态机的状态选择性地将代理耦合到非易失性存储器。通过这种方式，单个非易失性存储器可以在多个代理之间被共享以存储固件。此外，总线控制器基于功率定序器的输出来在CPU之间对到非易失性存储器的访问进行仲裁。固件访问仲裁和功率定序之间的这种耦合允许CPU当需要基于任意特定加电顺序时获得并且执行固件。

[0010] 在示例中，硬件和软件的组合可以用于管理对单个非易失性存储设备的共享访问，所述非易失性存储设备包含用于启动多个中央处理单元(CPU)的固件。管理代理可以用于在非易失性存储器没有由CPU中的任何一个使用时更新固件，使得全部CPU可以同时看到更新。非易失性存储器可以用于存储针对计算节点中其它代理的固件。在多个代理之间共享具有固件的单个非易失性存储器降低了节点成本，并且要求较少的占用面积(real estate)。由于仅存在具有固件的单个非易失性存储器，所以对于所有代理的固件存在单个更新点。这可以节约更新时间。在示例中，管理代理可以具有向非易失性存储器写入的独占权利，以提供针对在CPU上运行的恶意软件的破坏的更高级别的安全性。

[0011] 图 1 是根据示例实现的计算节点 100 的框图。计算节点 100 可以是单个计算机系统，或者是包括多个这样的计算节点的较大计算机系统的一部分。计算节点 100 包括多个中央处理单元 (CPU)102、管理处理器 104、各种支持电路 106、存储器 108、各种输入 / 输出 (IO) 电路 120、固件 114 以及互连电路 101。互连电路 101 可以提供总线、桥等来促进计算机系统 100 的部件之间的通信。CPU 102 可以包括在本技术领域中公知的任意类型的微处理器。支持电路 106 可以包括高速缓存、电源、时钟电路、数据寄存器等。存储器 108 可以包括随机存取存储器、只读存储器、高速缓冲存储器、磁读 / 写存储器等或诸如此种存储设备的任意组合。

[0012] 管理处理器 104 可以包括任意类型的微处理器、微控制器、微计算器等。管理处理器 104 提供系统管理环境与计算节点 100 的硬件部件之间的接口，所述计算节点 100 的硬件部件包括 CPU 102、支持电路 106、存储器 108、IO 电路 120 和 / 或固件 114。在一些实现中，管理处理器 104 可以被称为基板管理控制器 (BMC)。管理处理器 104 及其功能独立于 CPU 102 的功能。

[0013] 固件 114 可以包括存储由包括 CPU 102 的节点 100 中的各种设备使用的代码的非易失性存储器。固件可以包括 BIOS、UEFI 等。固件 114 也可以包括一旦启动或重置由 CPU 102 首先执行的代码，其被称为“启动代码”。在本文中所使用的术语“非易失性存储器”可以指代任意类型的非易失性存储装置。示例包括只读存储器 (ROM)、电可擦除可编程 ROM(EEPROM)、闪速存储器、铁电随机存取存储器 (F-RAM) 等，以及这样的设备的组合。

[0014] 图 2 是根据本发明的示例的用于计算节点 100 的固件子系统 200 的框图。固件子系统 200 包括多个代理 202、控制器 204、非易失性存储器 206 及总线 208。代理 202 可以包括 CPU 102 和管理处理器 104。在示例中，代理 202 可以包括至少一个其它代理 (“其它一个或多个代理 210”)。非易失性存储器 206 存储固件 114。固件 114 可以包括用于由每一个代理 202 执行的代码。总线 208 可以是串行数据总线，例如，串行外围接口 (SPI) 总线等。在另一示例中，总线可以是包括并行总线的任意类型的总线。代理 202、控制器 204 以及非易失性存储器 206 耦合到总线 208 以用于通信。

[0015] 控制器 204 可以包括功率定序器 212、多个功率控制状态机 214、以及总线控制器 216。在示例中，控制器 204 可以是集成电路，如专用集成电路 (ASIC)、可编程逻辑器件 (PLD)(如复杂可编程逻辑器件 (CPLD) 或现场可编程门阵列 (FPGA)) 等。在示例中，功率定序器 212、多个功率控制状态机 214 以及总线控制器 216 中的一个或多个可以是在集成电路中实现的电路。在示例中，功率定序器 212、控制状态机 214 以及总线控制器 216 中的一个或多个可以被实现为在集成电路中由处理器执行的软件。在另一示例中，控制器 204 的元件可以利用硬件电路和软件的组合来被实现。

[0016] 功率定序器 212 实现针对 CPU 102 的加电顺序。在示例中，功率定序器 212 每次选择一个 CPU 用于加电。在给定的 CPU 已经完成它的加电之后，功率定序器 212 选择另一 CPU。通过这种方式，CPU 102 被顺序地并且不是全部同时地被加电。“供电”和“加电”这两个术语在本文中作为同义词使用。通常，CPU 通过依靠执行在特定的预定义位置处开始的指令 (如重置向量) 来“供电”。

[0017] 功率控制状态机 214 基于功率定序器 212 的输出来控制 CPU 102 的状态。在示例中，每个 CPU 可以处于各种状态，如断电、重置、供电以及任意各种部分供电状态 (如各种睡

眠状态）。CPU 102 中的每一个包括专用功率控制状态机 214。在示例中，功率控制状态机 214 将未被供电的 CPU 102 中的每一个保持在重置状态。

[0018] 总线控制器 216 基于功率控制状态机 214 的状态来将代理 202 选择性地耦合到非易失性存储器 206。当功率控制状态机 214 指示 CPU 102 中的一个将被供电时，总线控制器 216 将所选择的 CPU 102 耦合到非易失性存储器 206。在示例中，总线控制器 216 包括总线仲裁器 218 和总线多路复用器 220。总线仲裁器 218 选择任意的代理 202 以用于通过总线 208 与非易失性存储器 206 进行通信。换言之，总线仲裁器 218 每次向一个代理授权总线访问。总线仲裁器 218 可以在 CPU 102 中的每一个基于功率控制状态机 214 的输出（以及间接地基于功率定序器 212 的输出）而被供电时，向该 CPU 授权总线访问。总线多路复用器 220 在非易失性存储器和由总线仲裁器 218 选择的代理 202 之间建立通信链路。可以理解，基于可以与本发明一起使用的不同类型的公知的总线，总线控制器 216 可以具有不同的配置。通常，总线控制器 216 促进了在多个代理 202 之间对非易失性存储器 206 的共享访问。一旦 CPU 102 已经访问了非易失性存储器 206，CPU 102 就可以检索其固件并且执行加电。

[0019] 总线控制器 216 可以接收将总线授权赋予代理 202 而不是 CPU102 的另外的输入。例如，总线控制器 216 可以服务来自其它代理 202 的针对访问非易失性存储器 206 的总线访问请求。在示例中，管理处理器 104 可以将这种请求发送给总线控制器 216。管理处理器 104 可以请求访问非易失性存储器 206，以便对固件写和 / 或读。例如，管理处理器 104 可以将固件的各种（多个）映像写入非易失性存储器（如针对任意代理 202 的升级后的固件）。任意其它代理 210 可以类似地请求访问非易失性存储器以用于写和 / 或读存储其中的固件。

[0020] 图 3 是描绘了根据本发明的示例的计算机系统 300 的框图。计算机系统 300 包括多个计算节点 302。每个计算节点 302 可以类似与计算节点 100 地进行配置。计算节点 302 的每一个可以包括类似于图 2 中所示固件子系统的固件子系统 200。也即，每个计算节点 302 可以包括多个已经对非易失性存储器中的固件共享访问的代理。代理包括多个 CPU，所述多个 CPU 获得对非易失性存储器的共享访问以检索它们的固件用于供电和启动。

[0021] 图 4 是描绘了一种根据示例实现的在节点上多个代理之间共享固件的方法的流程图，所述多个代理包括被连接到总线的多个 CPU。方法 400 从步骤 402 开始，其中固件被存储于针对多个代理的被连接到总线的非易失性存储器中。在步骤 404，实现针对多个 CPU 的加电顺序。在步骤 406，基于加电顺序来控制多个 CPU 的状态。在步骤 408，基于 CPU 的状态来选择性地将代理耦合到非易失性存储器。

[0022] 在步骤 410，可以做出针对访问非易失性存储器和授予请求代理独占访问的一个（多个）另外的请求。特别地，在步骤 412，管理处理器可以被授予访问非易失性存储器以用于更新存储于其中的固件。

[0023] 图 5 是描绘了根据本发明的示例的控制 CPU 状态的方法 500 的流程图。方法 500 可以在方法 400 中的步骤 406 处执行。在步骤 502，基于加电顺序来选择允许被供电的 CPU。在步骤 504，CPU 被授权对非易失性存储器的总线访问。在步骤 506，其它 CPU 中的每一个被保持在重置状态中。接下来可以针对各 CPU 来重复方法 500。

[0024] 在上述描述中，阐述了许多细节来提供对本发明的理解。然而，本领域中的技术人员将理解，可以在不具有这些细节的情况下实施本发明。尽管参照限定数量的实施例公开

了本发明，但本领域中的技术人员将意识到由此的多种修改和变形。意图是附加的权利要求覆盖落入本发明的真正精神和范围的这些修改和变形。

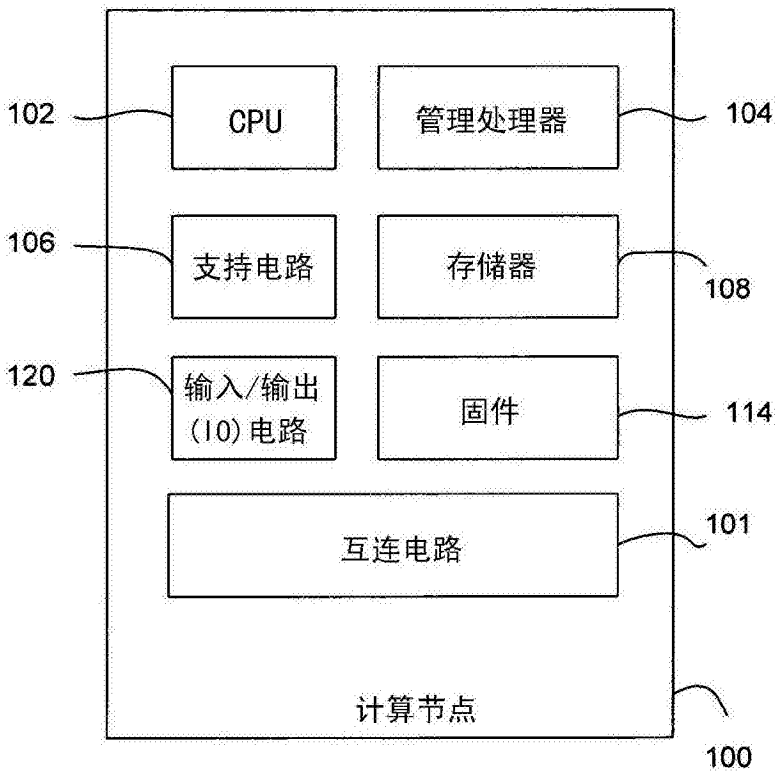

102
CPU
管理处理器
104
106
支持电路
存储器
108
120
输入/输出
(IO)电路
固件
114
互连电路
101
计算节点
100

图 1

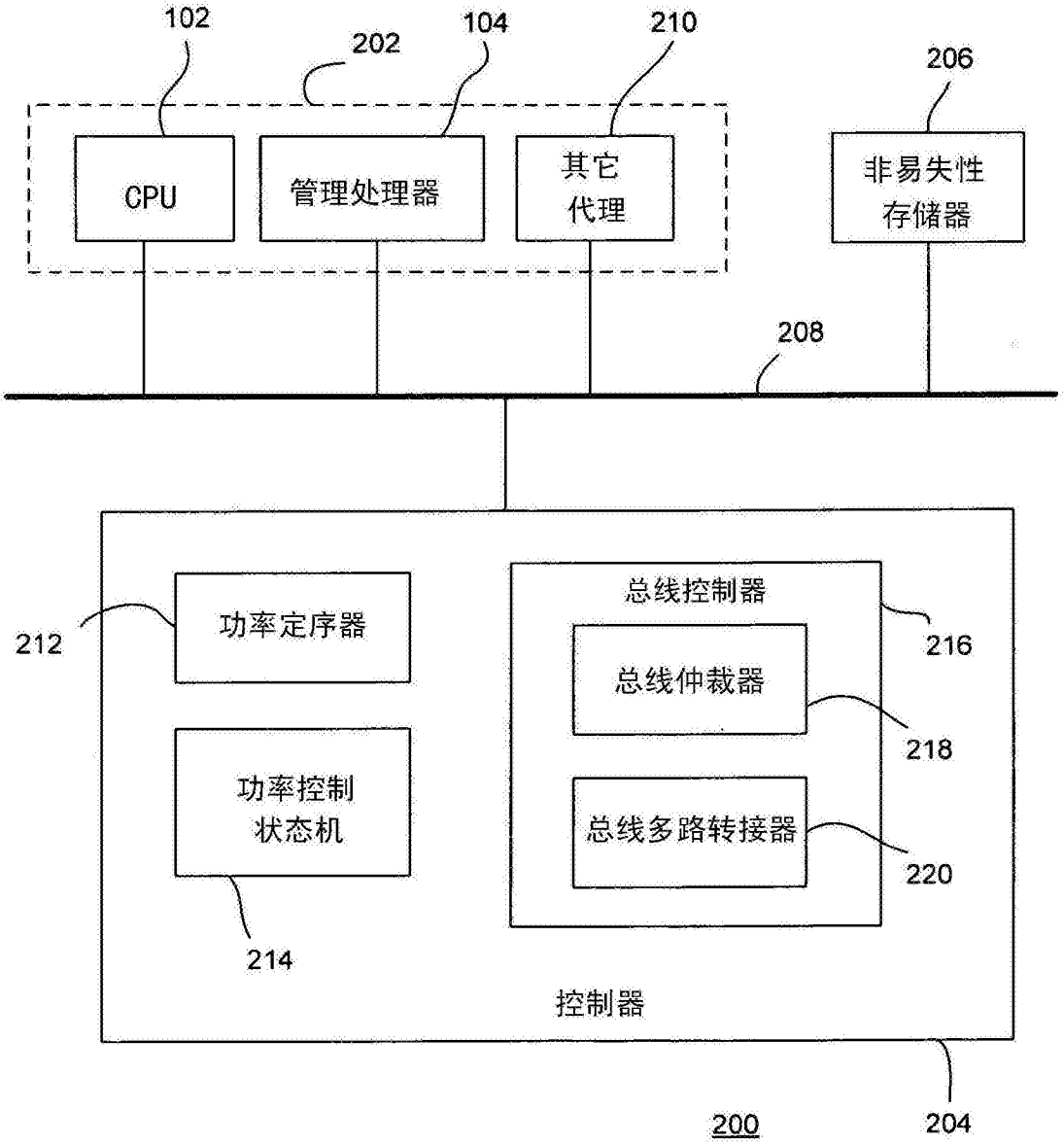

102
202
104
210
206
CPU
管理处理器
其它
代理
非易失性
存储器
208
212
功率定序器
总线控制器
216
总线仲裁器
218
功率控制
状态机
总线多路转接器
220
214
控制器
200
204

图 2

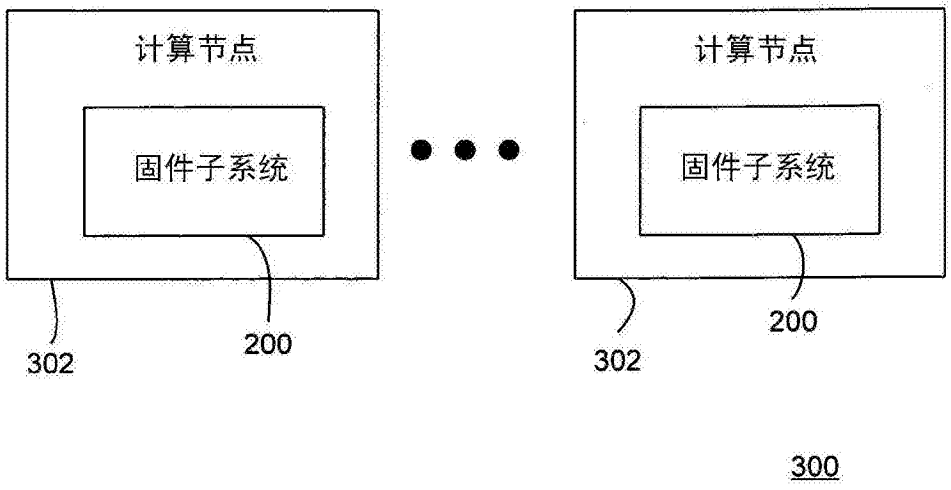


图 3

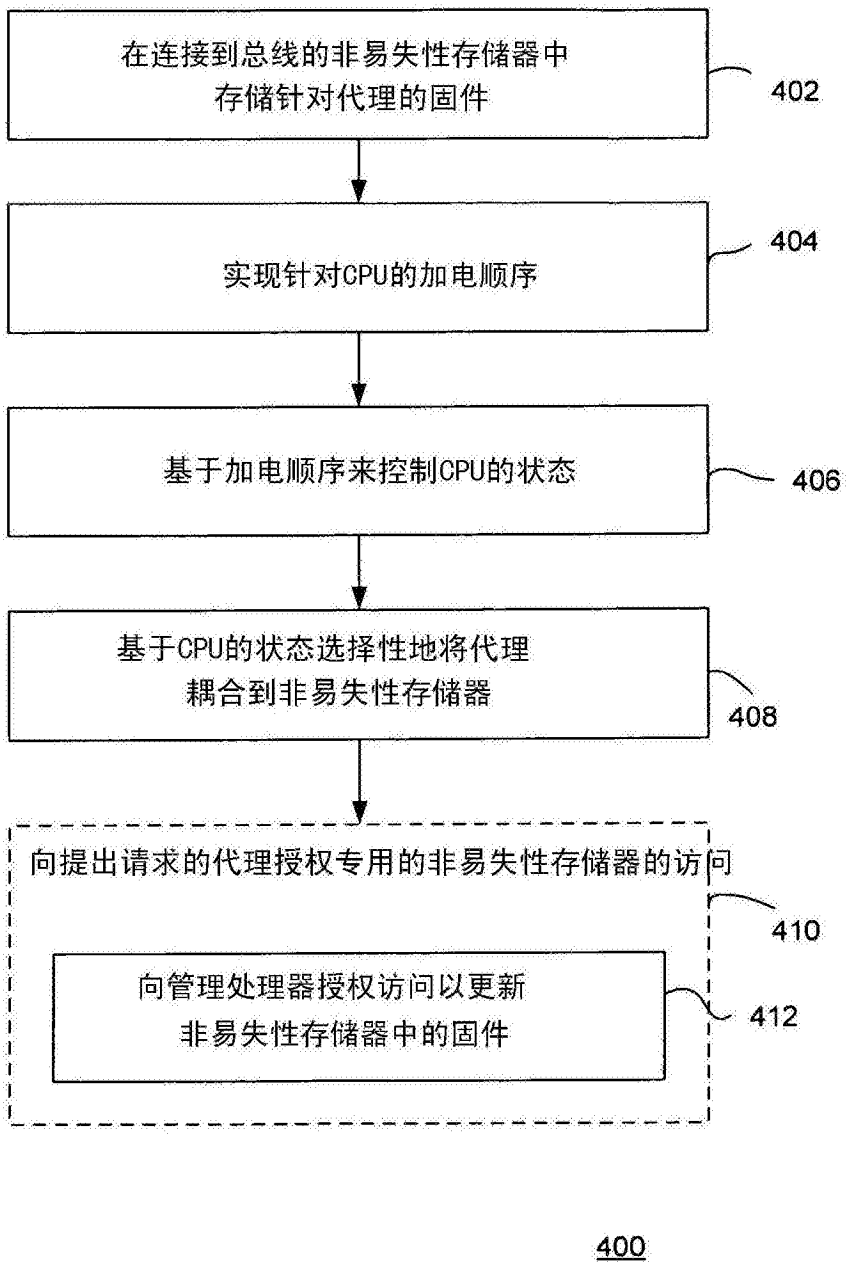

在连接到总线的非易失性存储器中
存储针对代理的固件
402
实现针对CPU的加电顺序
404
基于加电顺序来控制CPU的状态
406
基于CPU的状态选择性地将代理
耦合到非易失性存储器
408
向提出请求的代理授权专用的非易失性存储器的访问
410
向管理处理器授权访问以更新
非易失性存储器中的固件
412
400

图 4

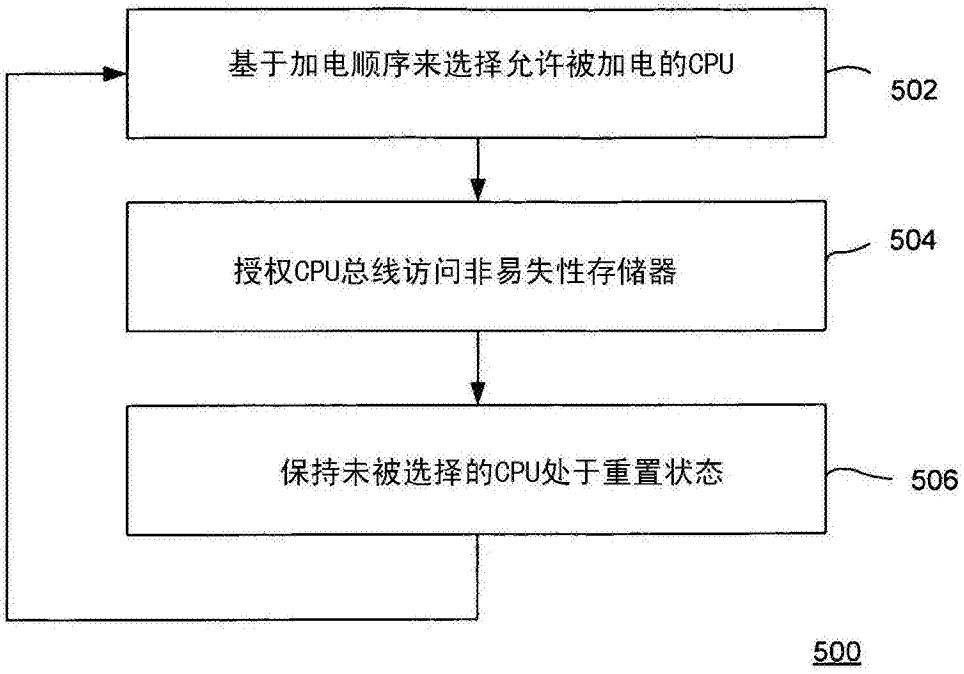

基于加电顺序来选择允许被加电的CPU
502
授权CPU总线访问非易失性存储器
504
保持未被选择的CPU处于重置状态
506
500

图 5