[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
G06F 12/08 (2006.01)

[12] 发明专利说明书

专利号 ZL 200310123566.5

[45] 授权公告日 2008 年 12 月 24 日

[11] 授权公告号 CN 100445964C

[22] 申请日 2003.12.26
[21] 申请号 200310123566.5
[30] 优先权
[32] 2002.12.27 [33] US [31] 10/330986
[73] 专利权人 英特尔公司
地址 美国加利福尼亚州
[72] 发明人 C·D·哈尔 R·L·坎贝尔
[56] 参考文献
CN85106711A 1987.2.4
US5560013A 1996.9.24
CN1359496A 2002.7.17
CN1382277A 2002.11.27
US4742447 1988.5.3
审查员 俞立文

[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 程天正 陈 霁

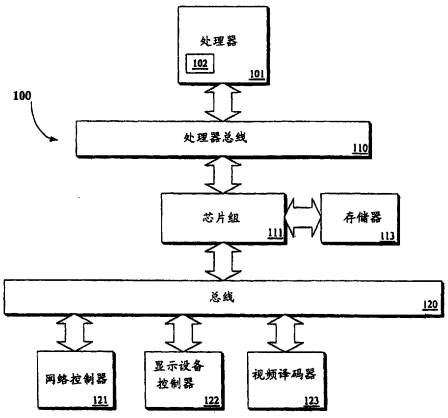
权利要求书 2 页 说明书 7 页 附图 5 页

[54] 发明名称

用于后期重映射虚拟机存储器页面的机制

[57] 摘要

依照一个实施例，公开了一种计算机系统(100)。所述计算机系统包括处理器(101)、与该处理器(101)相耦合的芯片组(111)以及与该芯片组相耦合的存储器(113)。所述芯片组(111)将从该处理器(101)接收的分区虚拟机存储器地址转换为页面级地址。



1. 一种计算机系统，包括：

处理器，包括分区逻辑，用于产生分区的虚拟机存储器地址；

与所述处理器相耦合的芯片组，用于将所述分区的虚拟机存储器地址转换为页面级地址；以及

与所述芯片组相耦合的存储器设备，所述存储器设备能够根据所述页面级地址而被访问。

2. 如权利要求 1 所述的计算机系统，其中，所述芯片组包括转换后援缓冲器 TLB，转换后援缓冲器 TLB 用于将所述分区的虚拟机存储器地址转换为所述页面级地址。

3. 如权利要求 2 所述的计算机系统，其中所述 TLB 包括：

用于接收分区的地址的输入；

与所述输入相耦合的重映射目录；

与所述目录和输入相耦合的重映射表；以及

与所述重映射表相耦合的输出，用于传送物理地址。

4. 如权利要求 3 所述的计算机系统，其中所述存储器设备包括：

重映射表，用于将所述分区的虚拟机存储器地址转换为物理存储器地址；以及

页表，用于规定所述分区的虚拟机存储器地址到页面级地址的重映射。

5. 如权利要求 2 所述的计算机系统，还包括：

总线主控器设备；以及

与总线主控器设备和 TLB 相耦合的偏移寄存器。

6. 如权利要求 5 所述的计算机系统，其中所述 TLB 根据 TLB 基址内的标志位来控制总线主控器设备访问存储器的权限。

7. 如权利要求 6 所述的计算机系统，其中所述总线主控器设备是图形控制器。

8. 如权利要求 6 所述的计算机系统，其中所述总线主控器设备是磁盘控制器。

9. 一种芯片组，包括：

输入，用于接收来自处理器的分区的虚拟机存储器地址；

重映射逻辑，耦合到所述输入，用于把所述分区的虚拟机存储器地址

转换为页面级地址；以及

输出，耦合到所述重映射逻辑，用于把所述页面级地址传送到存储器设备。

10. 如权利要求 9 所述的芯片组，其中所述重映射逻辑包括：

与所述输入相耦合的重映射目录；以及

与所述重映射目录和所述输入相耦合的重映射表。

11. 如权利要求 10 所述的芯片组，还包括转换后援缓冲器 TLB，所述 TLB 根据 TLB 基址内的标志位来控制总线主控器设备访问存储器的权限，其中所述总线主控器设备与所述芯片组相耦合。

12. 如权利要求 11 所述的芯片组，其中所述总线主控器设备是图形控制器。

13. 一种用于重映射分区的虚拟存储器的方法，包括：

产生分区的虚拟机存储器地址；

在芯片组接收所述分区的虚拟机存储器地址；

把所述分区的虚拟机存储器地址转换为页面级地址；以及

根据所述页面级地址访问存储器设备。

14. 如权利要求 13 所述的方法，其中产生分区的虚拟机存储器地址进一步包括：

在处理器将逻辑地址转换为虚拟机存储器物理地址；以及

将所述虚拟机存储器物理地址转换为所述分区的虚拟机存储器地址。

15. 如权利要求 14 所述的方法，进一步包括：将所述分区的虚拟机存储器地址传送到所述芯片组。

用于后期重映射虚拟机存储器页面的机制

技术领域

此发明涉及诸如微处理器的计算机处理器的虚拟机。具体来讲，本发明涉及对虚拟机的存储器管理。

背景技术

操作系统（OS）是用于控制物理计算机硬件（例如，处理器、存储器、磁盘和 CD－ROM 驱动器）的一种软件程序，并且给应用程序提供一组统一的抽象服务（例如，文件系统）。虚拟机管理器（VMM）也是一种控制物理计算机硬件的软件程序，所述物理计算机硬件例如像处理器、存储器以及磁盘驱动器。与 OS 不同，VMM 给在虚拟机（VM）内部执行的程序提供一种错觉，好像所述程序是在真的物理计算机硬件上执行，所述物理计算机硬件例如是处理器、存储器和磁盘驱动器。

每个 VM 通常作为自包含实体来运行，使得在 VM 中执行的软件好像正单独运行在“裸”机上，而非在虚拟机内部那样来执行，其中所述虚拟机与其它 VM 共享处理器和其它物理硬件。这就是 VMM，其仿效“裸”机的确定功能，以便使在 VM 内部执行的软件执行起来就好像它是一个在计算机上执行的单独实体一样。

已经开发了各种技术来为运行在系统上的虚拟机分配物理系统存储器。一种这样的技术是利用分区存储器。所述分区存储器是将物理系统存储器分为多个连续区域的情形。虽然分区存储器在 CPU 中实现起来不算昂贵，但是存储器在虚拟机之间却不能够被容易地动态重新配置。

还开发了多种技术以用于通过给定的存储容量来获得更多的性能。一种这样的技术是利用虚拟存储器。虚拟存储器是基于这样一种概念，即当运行程序时，不必将整个程序一次载入主存储器。相反，计算机的操作系统将根据执行的需要将程序段从二级存储器设备（例如硬盘驱动器）载入主存储器。

为使该方案可行，操作系统维护了如下的表，所述表用于持续跟踪每个程序段驻留在主存储器和二级存储器中的位置。作为以该方式执行程序的结果，程序的逻辑地址不再对应于主存储器中的物理地址。为处理这种情况，中央处理单元（CPU）将程序的有效地址或虚拟地址映射成它们对应的物理地址。

然而，在执行分区存储器技术的计算机系统中，常常希望为每个虚拟机动态地重新分配存储器。还希望以单个页面为基础，而不是以大容量的区域为基础来管理存储器。采用当前分区存储器系统，不可以动态地重新分配存储器。当前分区存储器系统还对页面级的灵活性强加了许多限制，而许多操作系统都需要所述页面级的灵活性来支持虚拟存储器。

发明内容

本发明提供一种计算机系统，包括：

处理器，包括分区逻辑，用于产生分区虚拟机存储器地址；

与所述处理器相耦合的芯片组，用于将所述分区虚拟机存储器地址转换为页面级地址；以及

与所述芯片组相耦合的存储器设备，该存储器设备可以根据所述页面级地址加以访问。

本发明还提供了一种芯片组，包括：

输入，用于接收来自处理器的分区虚拟机存储器地址；

重映射逻辑，耦合到所述输入，用于把所述分区虚拟机存储器地址转换为页面级地址；以及

输出，耦合到所述重映射逻辑，用于把所述页面级地址传送到存储器设备。

本发明又提供了一种用于重映射分区虚拟存储器的方法，包括：

产生分区虚拟机存储器地址；

在芯片组接收所述分区虚拟机存储器地址；

把所述分区虚拟机存储器地址转换为页面级地址；以及

根据所述页面级地址访问存储器设备。

附图说明

通过下面所给出的详细描述以及通过本发明不同实施例的附图，将会更加全面地理解本发明。然而，不应该将所述附图理解为将本发

明限定为具体的实施例，而仅仅是为了说明和理解。

图 1 是计算机系统的一个实施例的框图；

图 2 是耦合到芯片组和存储器设备的处理器的一个实施例的框图；

图 3 是芯片组重映射机制的一个实施例的框图；

图 4 是将虚拟地址转换为页面地址的一个实施例的流程图；以及

图 5 是耦合到芯片组和存储器设备的总线主控器的一个实施例的框图。

具体实施方式

下面描述用于将分区虚拟机存储器重映射到页面粒度存储器（page granular memory）的机制。在说明书中提及的“一个实施例”或“一实施例”意指就该实施例所描述的特定特征、结构或特性，包括在本发明的至少一个实施例中。在说明书不同地方出现的短语“在一个实施例中”未必都指的是同一实施例。

在随后的描述中，提出了很多细节。然而，对于本领域技术人员而言，显而易见的是，在没有这些具体细节的情况下也可以实施本发明。在其他情况下，为了避免模糊本发明，以框图形状示出了公知的结构和设备，而没有以细节的方式示出。

图 1 是计算机系统 100 的一个实施例的框图。计算机系统 100 包括用于处理数据信号的处理器 101。处理器 101 可以是复杂指令集计算机（CISC）微处理器、精简指令集计算（RISC）微处理器、超长指令字（VLIW）微处理器、实现指令集组合的处理器或其它处理器设备。

在一个实施例中，处理器 101 是包括奔腾®IV 系列和移动奔腾®的奔腾®系列处理器中的一种处理器，并且奔腾®IV 可以从 California 的 Santa Clara 的 Intel 公司购买到。作为选择，还可以使用其它处理器。图 1 示出了采用单一处理器计算机的计算机系统 100 的例子。然而，本领域普通技术人员将会意识到，计算机系统 100 可以使用多个处理器来实现。

处理器 101 与处理器总线 110 相耦合。处理器总线 110 在计算机系统 100 中的处理器 101 和其它部件之间传送数据信号。计算机系统 100 还包括存储器 113。在一个实施例中，存储器 113 是动态随机存取存储器（DRAM）设备。然而，在其它实施例中，存储器 113 可以是静

态随机存取存储器（SRAM）设备，或其它存储器设备。

存储器 113 可以存储由数据信号表示的指令和代码，所述指令和代码可以由处理器 101 执行。依照一个实施例，高速缓冲存储器 102 驻留在处理器 101 内，并存储如下的数据信号，这些数据信号还被存储在存储器 113 中。高速缓冲存储器 102 由处理器 101 通过利用其访问的局部性来加速存储器访问。在另一个实施例中，高速缓冲存储器 102 驻留在处理器 101 的外部。

计算机系统 100 进一步包括芯片组 111，所述芯片组 111 与处理器总线 110 和存储器 113 相耦合。芯片组 111 控制计算机系统 100 中的处理器 101、存储器 113 以及其它部件之间的数据信号，并且桥接在处理器总线 110、存储器 113 以及第一输入/输出（I/O）总线 120 之间的数据信号。

在一个实施例中，I/O 总线 120 可以是单条总线或多条总线的组合。在其它的实施例中，I/O 总线 120 可以是遵守 2.1 版本规范的外围部件互连（Peripheral Component Interconnect）总线，其是由设立于 Oregon 的 Portland 的 PCI 专业组开发的。在另一个实施例中，I/O 总线 120 还可以是由设立于 California 的 San Jose 的 PCMCIA 开发的个人计算机存储卡国际协会(PCMCIA)总线。作为选择，还可以使用其它总线来实现 I/O 总线。I/O 总线 120 提供了计算机系统 100 中的部件之间的通信链路。

网络控制器 121 与 I/O 总线 120 相耦合。网络控制器 121 将计算机系统 100 链接到计算机网络（图 1 中未示出）中，并且支持机器之中的通信。在一个实施例中，计算机系统 100 经由网络控制器 121 从计算机 110 接收流式视频数据。

显示设备控制器 122 也与 I/O 总线 120 相耦合。显示设备控制器 122 允许将显示设备耦合到计算机系统 100，并且充当显示设备和计算机系统 100 之间的接口。在一个实施例中，显示设备控制器 122 是单色显示适配器（MDA）卡。

在其它实施例中，显示设备控制器 122 还可以是彩色图形适配器（CGA）卡、增强型图形适配器（EGA）卡、扩展型图形阵列（XGA）卡，或其它显示设备控制器。该显示设备可以是电视机、计算机监视器、平板显示器或其它显示设备。该显示设备经由显示设备控制器 122 从

处理器 101 来接收数据信号，并向计算机系统 100 的用户显示信息和数据信号。

视频译码器 123 也与 I/O 总线 120 相耦合。视频译码器 123 是用于将所接收到的编码数据转换为其原始格式的硬件设备。依照一个实施例，视频译码器 123 是运动图像专家组 4（MPEG-4）译码器。然而，本领域普通技术人员将会意识到，视频译码器 123 还可以采用其它类型的 MPEG 译码器来实现。

依照一个实施例，计算机系统 100 支持使用分区存储器的虚拟机。在其它实施例中，计算机系统 100 包括用于后期将虚拟机存储器重映射到页面粒度存储器的机制。图 2 是与芯片组 111 和存储器设备 113 相耦合的处理器 101 的一个实施例的框图。处理器 101 包括分区逻辑 210 以及转换后援缓冲器（TLB）218。

分区逻辑 210 通过存储器 113 中的地址分区来支持虚拟机系统。分区逻辑 210 存储映射信息，该映射信息表示存储器 113 中的页在 TLB218 的位置。具体来讲，分区逻辑 210 为每个虚拟机生成一地址范围。例如，可以将 0-1Gb 的范围分配给第一虚拟机，同时将 1-2Gb、以及 2-3Gb 的范围分别分配给第二和第三虚拟机。

TLB 218 与分区逻辑 210 相耦合。TLB 218 是在存储器 113 中最频繁使用的页表项（PTE）的高速缓存器。具体来讲，TLB 218 包括正在存储器 113 中使用的当前有效地址。因此，每当执行地址转换时，未必要访问存储器 113 中的 PTE。

在常规的虚拟机系统中，分区逻辑 210 以及 TLB 218 执行所有必要的地址转换。然而，常常希望以单个页面为基础来管理存储器 113，而不是以大容量的区域为基础。采用当前分区存储器系统，不可以指定如下页面级灵活性的存储器，所述页级灵活性是大多数操作系统所要求的。

芯片组 111 包括 TLB 220。TLB 220 是一高速缓冲存储器，其包括在存储器 113 内的重映射表中正在使用的当前有效地址。存储器 113 包括页表 236 以及重映射表 238。页表 236 将分区虚拟机地址转换为物理存储器地址。页表 236 包括 PTE 的集合。页表 236 为每个映射的虚拟页保持一个 PTE。为了访问物理存储器中的页，查找适当的 PTE 以便找到该页驻留的位置。

与页表 236 相似，重映射表 238 包括一表项的列表，该表项用于规定分区虚拟地址到页面级地址的重映射。在一个实施例中，实现 TLB 220 以及重映射表 238，以便相反地转换存储器地址从而支持页面级操作，由此，规避了由处理器 101 强加的分区范围。

图 5 是与芯片组 111 及存储器设备 113 相耦合的处理器 101 和总线主控器 530 的另一个实施例的框图。总线主控器 530 是 I/O 设备（例如，磁盘控制器、图形控制器等），其有权访问存储器 113。偏移寄存器 550 耦合在总线主控器 530 和 TLB 220 之间。

偏移寄存器 550 提供到虚拟机地址空间的映射，所述虚拟机控制总线主控器 530。通常，计算机系统 100 的操作系统期望分区虚拟机物理地址是实际地址。然而，由于所述地址因为芯片组进行的重映射而不是真的物理地址，所以由总线主控器使用的地址是不正确的。由此，偏移管理器 550 以及 TLB 220 纠正该地址，以便将该地址转换为与在 VM 中运行的操作系统的地址相同的地址。依照一个实施例，TLB 220 可以根据 TLB 220 的基址内的标志位阻塞总线主控器 530 访问存储器 113。该标志可以依据由具体特定实现所提供的定义来阻止读访问或写访问或者它们两者。

图 3 是芯片组 111 的一个实施例的框图。如上所述，芯片组 111 包括 TLB 220。TLB 220 包括 TLB 输入 305、TLB 输出 310、重映射目录 320 以及重映射表 330。TLB 输入 305 接收来自于处理器 101 内部的分区逻辑 210 的分区地址。TLB 输出 310 将已转换的物理地址传送到存储器 113。

目录项 320 接收分区地址的低 10 位，以便在重映射表 330 定义起始位置。重映射表 330 接收下面的 10 个位，所述下面的 10 个位表示重映射表 330 内的、包括对应的物理地址的实际表项。在选择了物理地址之后，将该地址连同分区地址的最先 12 位一起从 TLB 输出 310 被传送。

图 4 是将虚拟地址转换为页面地址的一个实施例的流程图。在处理块 410 处，处理器 101 中的页表（例如 TLB 218）将逻辑地址转换为虚拟存储器物理地址。在处理块 420 处，分区逻辑 210 将虚拟存储器物理地址转换到分区地址空间。在处理块 430 处，芯片组重映射机制将分区地址空间转换到物理地址空间。在处理块 440 处，将物理地

址传送到存储器 113。

芯片组重映射机制结合了存储器分区和页面粒度重映射的优点。由此，芯片组重映射机制能够使处理器支持具有分区存储器空间的虚拟机充分利用由页面粒度存储器空间提供的灵活性。例如，现在可以实现为虚拟机动态调整存储器大小，而这在常规系统中几乎是不可能的。

然而，对于本领域普通技术人员来说，在已阅读前面的描述之后，本发明的多种替换和修改无疑是显而易见的，应该理解的是，为了举例说明而所示出以及所描述的具体实施例不能理解为是对本发明的限制。因此，提及各种实施例的细节不意味限制权利要求书的范围，该权利要求书本身陈述了被视为本发明的那些特征。

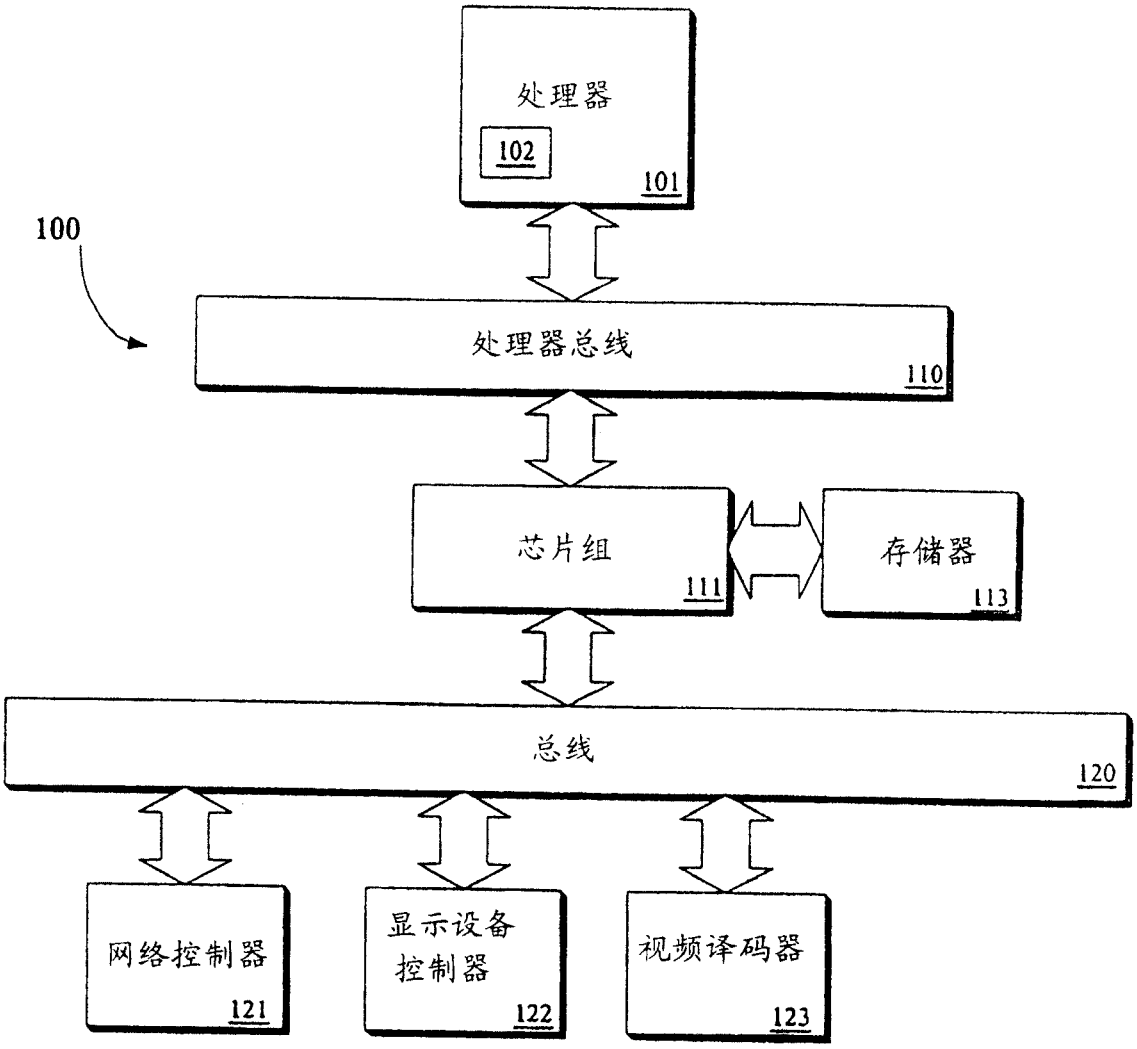

处理器
102
101
100
处理器总线
110
芯片组
111
存储器
113
总线
120
网络控制器
121
显示设备
控制器
122
视频译码器
123

图 1

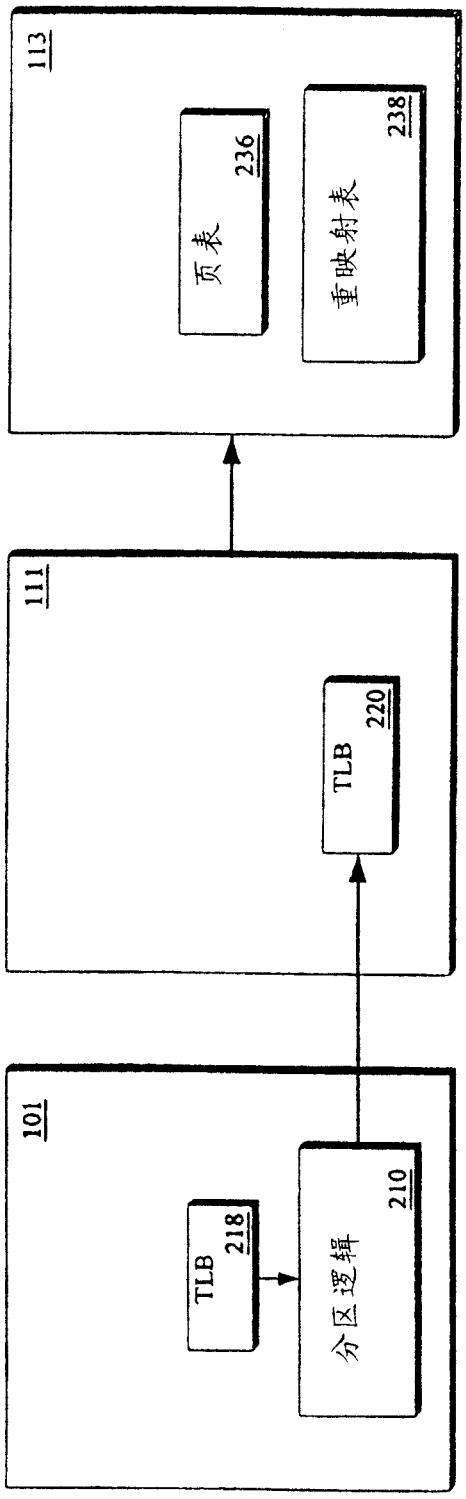


图 2

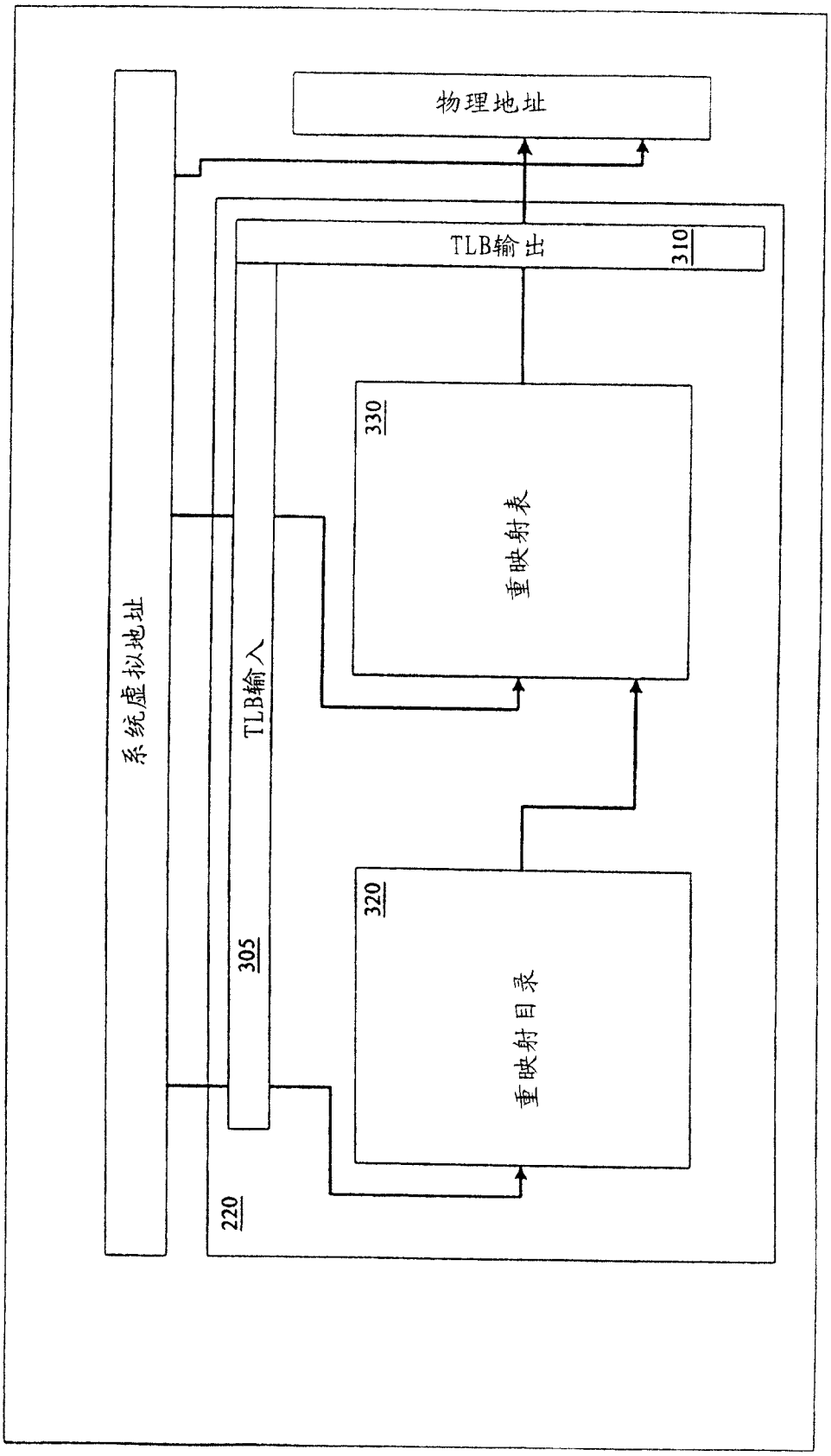


图 3

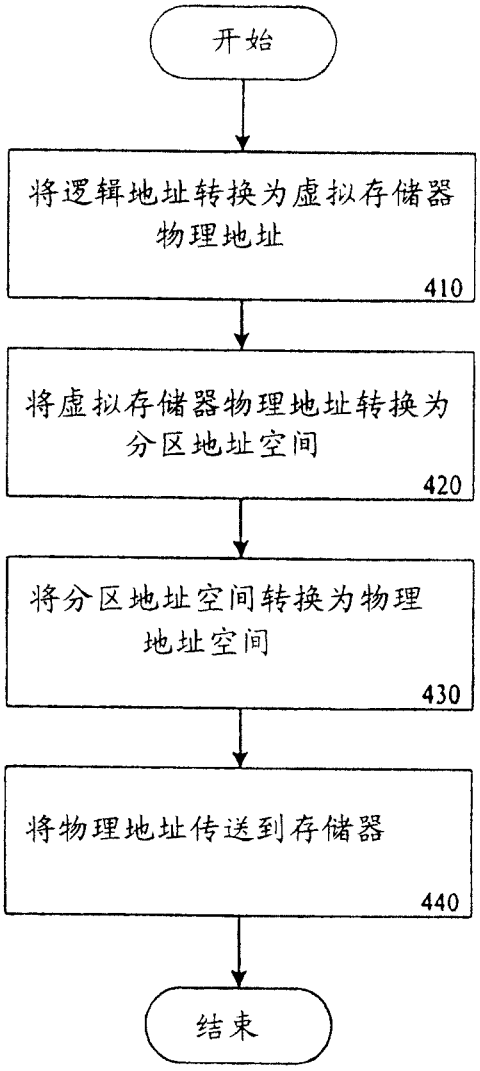

开始
将逻辑地址转换为虚拟存储器物理地址
410
将虚拟存储器物理地址转换为分区地址空间
420
将分区地址空间转换为物理地址空间
430
将物理地址传送到存储器
440
结束

图 4

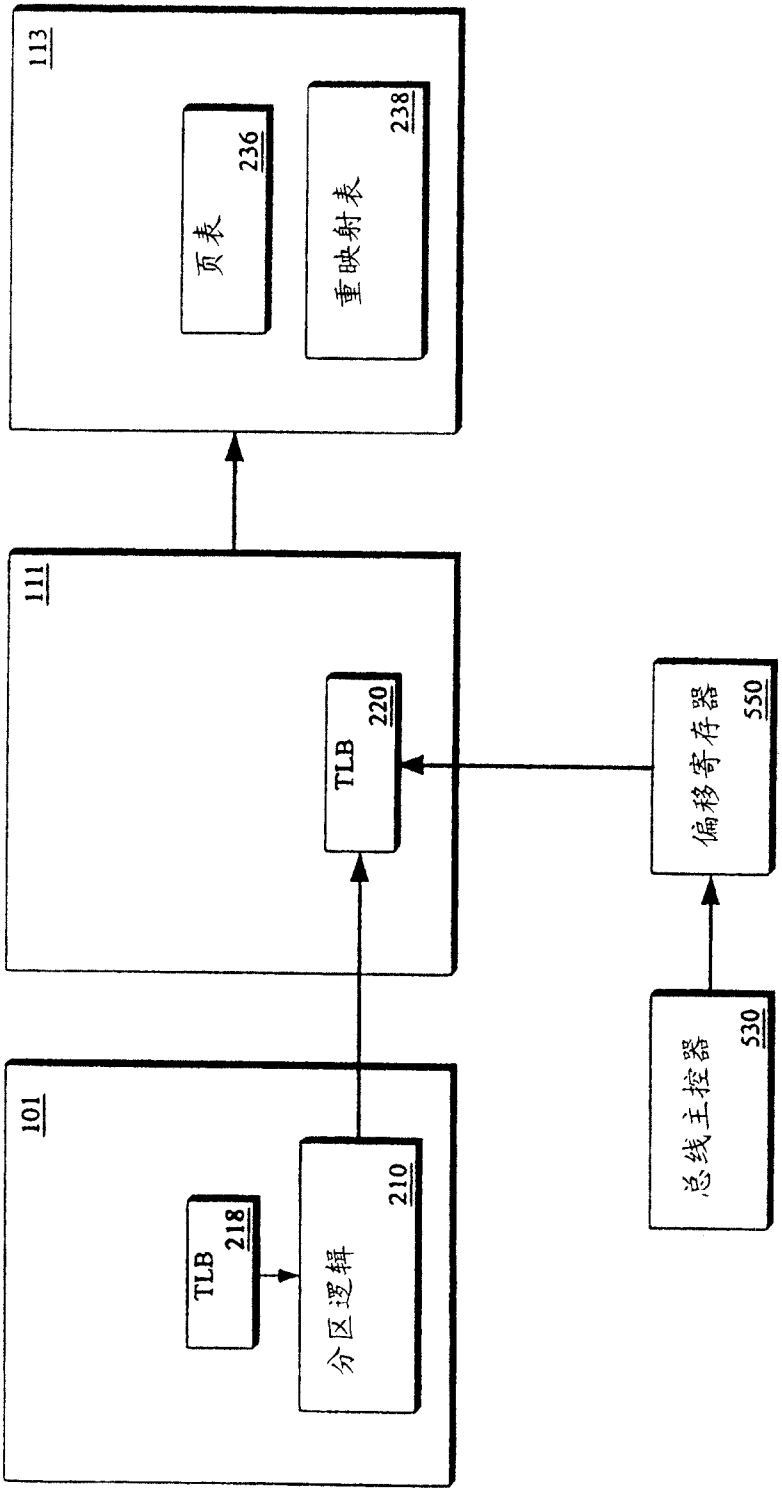

113
页表
236
重映射表
238
111
TLB
220
偏移寄存器
550
101
TLB
218
分区逻辑
210
总线主控器
530

图 5