



(12) 发明专利申请

(10) 申请公布号 CN 102662853 A

(43) 申请公布日 2012. 09. 12

(21) 申请号 201210078070. X

(22) 申请日 2012. 03. 22

(71) 申请人 北京北大众志微系统科技有限责任公司

地址 100080 北京市海淀区中关村北大街
151 号燕园资源大厦 11 层

(72) 发明人 程旭 钟祺 管雪涛 王晶

(74) 专利代理机构 常州市维益专利事务所
32211

代理人 路接洲

(51) Int. Cl.

G06F 12/06 (2006. 01)

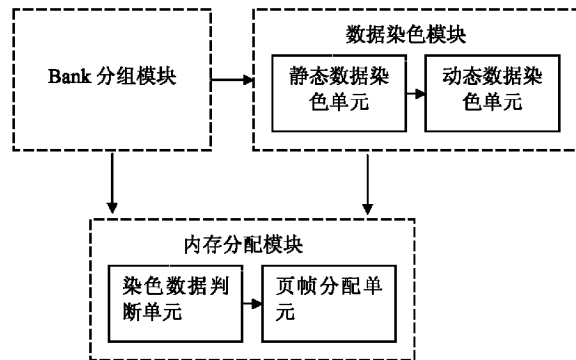
权利要求书 1 页 说明书 5 页 附图 2 页

(54) 发明名称

实现使用存储级并行的内存管理方法及装置

(57) 摘要

本发明涉及一种实现使用存储级并行的内存管理方法及装置,将 bank 的概念引入内存分配器,通过 bank 分组建立地址与 bank 间的关联,使分配器能根据地址范围标识不同的 bank。数据划分为多个大小不一的数据单元,并将数据散布到主存的所有 bank 组中,以提高访存的并行度和减少行缓存冲突。同时,本发明完全工作在操作系统层,利用编译器和操作系统提供的信息来分析数据单元间的冲突开销,并根据主存的实际配置扩展内存分配器,既不用修改应用程序,也不依赖特殊的底层硬件。



1. 一种实现使用存储级并行的内存管理方法,其特征在于,包括:

以所述的 bank 组为单位分配内存,并确定划分的 bank 组与数据间的映射关系。

2. 按照权利要求 1 所述的内存管理方法,其特征在于,还包括:

为应用的数据分配符合所述 bank 组的映射条件的物理页帧。

3. 按照权利要求 1 或 2 所述的内存管理方法,其特征在于:以所述应用的数据为单位划分使用主存的 bank,并确定划分的 bank 的映射条件,具体包括:

将与所述主存的高位片选地址对主存的 bank 进行划分,并以所述划分的 bank 组号作为所述数据的染色条件。

4. 按照权利要求 3 所述的内存管理方法,其特征在于:为冲突开销较大的数据分配来自不同 bank 组的物理页帧,具体包括:

当所述为应用申请物理内存时,根据数据间的冲突开销将数据染色,在分配物理内存时根据染色结果散布数据到不同的 bank 组;确保冲突开销较大的数据使用不同的 bank 资源;

判断新申请的内存与已申请内存之间的开销,如果开销较大,则为新申请的内存染色,保证最大的前 n 个内存区域被着以不同的颜色;如果开销较小,则将新的内存区域着以第 n 个内存区域的颜色。

5. 一种使用大容量高速缓存的内存管理装置,其特征在于:包括 bank 分组模块,数据染色模块和内存分配模块,其中:

Bank 分组模块,用于将 bank 的概念引入内存分配器。通过 bank 分组建立地址与 bank 间的关联,使分配器能根据地址范围标识不同的 bank;Bank 的分组信息输出给数据染色模块和内存分配模块。

数据染色模块,用于判断变量间的冲突开销;变量占用的存储单元越多,越难在 Cache 中驻留;当处理器同时操作多个变量时,使用变量大小的乘积表示变量间的冲突开销;冲突的开销越大,则需要将这些数据散布到不同的 bank 组;数据的染色结果输出给内存分配模块;

内存分配模块,用于为被染色的数据分配对应的物理页帧。

6. 按照权利要求 5 所述的内存管理装置,其特征在于:所述数据染色模块包括静态数据染色单元和动态数据染色单元,其中:

静态数据染色单元,在应用申请 BSS 段时,将符号表中的变量描述作为染色依据,并将染色结果输出给动态数据染色单元;

动态数据染色单元,用于将操作系统捕获的内存分配请求,将符合染色条件的动态内存插入链表,根据用户的输入动态调整染色结果。

7. 按照权利要求 5 所述的内存管理装置,其特征在于,所述内存分配模块包括依次连接的染色数据判断单元和页帧分配单元,其中:

染色数据判断单元,用于根据数据类型判断是否使用染色数据专用的内存分配接口,根据染色结果调用与 bank 组对应的内存分配接口,向页帧分配单元输出物理页帧分配指令;

页帧分配单元,用于根据物理页帧分配指令,分配映射到相应 bank 组中的物理页帧,或提示分配失败。

实现使用存储级并行的内存管理方法及装置

技术领域

[0001] 本发明涉及到计算机系统的内存管理方法,尤其涉及实现使用存储级并行的内存管理方法及装置。

背景技术

[0002] 处理器和主存之间存在较大的性能差距。Cache 能减少处理器访问主存的次数,但大容量 Cache 难以在片上集成,并且 Cache 单一的替换策略也难以满足应用访存行为的多样性。处理器仍会耗费大量的时间用于访问主存,因此需要主存更加快速地响应处理器的访存请求。

[0003] 访存地址中包含了行地址和列地址,用于定位存储阵列中的存储单元。当前主存采用高位交叉的方式来构造,使得行选和列选复用地址线和译码器,以节省成本。但单个存储阵列的数据宽度有限,为了与数据总线的宽度匹配,内存控制器会同时选中一组存储阵列,并且各个存储阵列来自不同的 DRAM 芯片。同时被选中的存储阵列成为 bank。这就是内存芯片寻址的基本原理。

[0004] RAM 被称为随机存储器,但从当前存储器的设计来看,每条指令访问存储器的延迟并不一致。从主存的组成来看,访问主存需要选中对应的存储体 (bank),使连续多条访存指令分别访问不同的 bank,是保证存储级并行的的重要手段。数据在主存中的分布决定着并行度的高低。如果分配给程序的页帧都来自不同的 bank,则前后两次访存定位到不同的 bank。在等待第一个访存请求的结果时,就可以开始响应第二个访存请求。前后两次访存的过程相叠加,进而提高访存的性能。因此,数据在主存中的布局越分散,则更容易发挥多 bank 结构提供的并行访问能力,这里的分散是指数据分布在不同的 bank 中。特别是随着主存的集成度不断提高,主存的生产成本逐步降低, bank 数与主存的容量同步增长。一旦数据聚簇在主存的部分 bank 中,应用执行时只会有部分 bank 处于工作状态,从而限制了大容量主存潜在的性能。

[0005] 因此,针对目前使用大容量主存的内存管理所存在的一系列问题,需要提供一种改进的内存管理方法,根据应用的访存行为充分利用 bank 资源,同时兼顾优化方法的灵活性和通用性,从而实现对系统性能的优化。

发明内容

[0006] 本发明要解决的技术问题是:提出一种实现使用存储级并行的内存管理方法及装置,能够减少应用数据聚簇在主存的部分 bank 中,导致主存潜在的性能被低估。

[0007] 本发明所采用的技术方案为:一种实现使用存储级并行的内存管理方法,以所述的 bank 组为单位分配内存,并确定划分的 bank 组与数据间的映射关系。并为应用的数据分配符合所述 bank 组的映射条件的物理页帧。

[0008] 本发明以所述应用的数据为单位划分使用主存的 bank,并确定划分的 bank 的映射条件,具体包括:将与所述主存的高位片选地址对主存的 bank 进行划分,并以所述划分

的 bank 组号作为所述数据的染色条件。

[0009] 为冲突开销较大的数据分配来自不同 bank 组的物理页帧,具体包括:当所述应用申请物理内存时,根据数据间的冲突开销将数据散步到不同的 bank 组。确保冲突开销较大的数据使用不同的 bank 资源;判断新申请的内存与已申请内存之间的开销,如果开销较大,则为新申请的内存染色,保证最大的前 n 个内存区域被着以不同的颜色;如果开销较小,则将新的内存区域着以第 n 个内存区域的颜色。

[0010] 同时,本发明还提供了一种使用大容量高速缓存的内存管理装置,包括 bank 分组模块,数据染色模块和内存分配模块,其中:

[0011] Bank 分组模块,用于将 bank 的概念引入内存分配器。通过 bank 分组建立地址与 bank 间的关联,使分配器能根据地址范围标识不同的 bank;Bank 的分组信息输出给数据染色模块和内存分配模块。

[0012] 数据染色模块,用于判断变量间的冲突开销;变量占用的存储单元越多,越难在 Cache 中驻留;当处理器同时操作多个变量时,使用变量大小的乘积表示变量间的冲突开销;冲突的开销越大,则需要将这些数据散布到不同的 bank 组;数据的染色结果输出给内存分配模块;

[0013] 内存分配模块,用于为被染色的数据分配对应的物理页帧。

[0014] 具体的说,本发明所述数据染色模块包括静态数据染色单元和动态数据染色单元,其中:

[0015] 静态数据染色单元,在应用申请 BSS 段时,将符号表中的变量描述作为染色依据,并将染色结果输出给动态数据染色单元;

[0016] 动态数据染色单元,用于将操作系统捕获的内存分配请求,将符合染色条件的动态内存插入链表,根据用户的输入动态调整染色结果;

[0017] 而所述内存分配模块包括依次连接的染色数据判断单元和页帧分配单元,其中:

[0018] 染色数据判断单元,用于根据数据类型判断是否使用染色数据专用的内存分配接口,根据染色结果调用与 bank 组对应的内存分配接口,向页帧分配单元输出物理页帧分配指令;

[0019] 页帧分配单元,用于根据物理页帧分配指令,分配映射到相应 bank 组中的物理页帧,或提示分配失败。

[0020] 本发明的有益效果是:将 bank 的概念引入内存分配器,通过 bank 分组建立地址与 bank 间的关联,使分配器能根据地址范围标识不同的 bank。数据划分为多个大小不一的数据单元,并将数据散布到主存的所有 bank 组中,以提高访存的并行度和减少行缓存冲突。同时,本发明完全工作在操作系统层,利用编译器和操作系统提供的信息来分析数据单元间的冲突开销,并根据主存的实际配置扩展内存分配器,既不用修改应用程序,也不依赖特殊的底层硬件。

附图说明

[0021] 下面结合附图和实施例对本发明进一步说明。

[0022] 图 1 为大容量主存中的数据分布示意图;

[0023] 图 2 为本发明对 bank 分组及其与 bank 组和物理地址映射关系示意图;

- [0024] 图 3 是本发明的标识数据染色的过程,包括对动态 / 静态数据染色 ;
- [0025] 图 4 是本发明的内存分配流程示意 ;
- [0026] 图 5 是本发明的使用存储级并行的内存管理装置实施例的结构示意图。

具体实施方式

[0027] 现在结合附图和优选实施例对本发明作进一步详细的说明。这些附图均为简化的示意图,仅以示意方式说明本发明的基本结构,因此其仅显示与本发明有关的构成。

[0028] 本发明利用大容量主存含有多个 bank 的结构特点,根据地址中的高位片选地址,将主存划分为不同的组。同时,将冲突开销较大的数据映射到不同的 bank 租,因此减少应用执行时的访存冲突。

[0029] 本发明提供的使用大容量高速缓存的内存管理方法,其流程如图 2 所示,包括如下步骤 :

[0030] 110 :根据 bank 对地址空间划分,使内存分配器根据地址标识主存中的不同 bank。

[0031] 为了使数据的布局分散,需要将 bank 的概念引入内存分配器。通过 bank 分组建立地址与 bank 间的关联,使分配器能根据地址范围标识不同的 bank。如果同时被访问的数据位于不同的地址范围,可以避免冲突,从而保证访存效率。同时,分组的结果为后续的数据染色提供依据,因为其决定了数据被染色时的颜色种类。

[0032] 图 3 描述了物理地址和 bank 间的映射关系。主存中单个阵列的数据宽度有限,为与数据总线的位宽匹配,内存控制器会同时选中一组阵列,同时被选中的阵列组成 bank。此外,DRAM 芯片中的阵列数量有限,通常为 8。为扩大容量,主存中集成了更多的 DRAM 芯片,需要联合片选地址和 bank 地址对 bank 进行选择。

[0033] 例如,利用高位片选地址将 bank 划分为多个组,组数为 $\text{bank_groups} = 2^{\text{chip_sel_bits}}$ 。当片选地址为 1 位, bank 被分为 2 组 ;片选地址为 2 位, bank 被分为 4 组。例如,2GB 的主存中包含 16 个 bank,阵列的数据宽度为 8,并且总线的位宽为 64,所有 bank 被分为两组,其中地址 0 ~ 1GB 对应组 0,1GB ~ 2GB 对应组 1。

[0034] 120 :分别对静态数据和动态数据进行染色,根据各种变量的大小来判断冲突开销。

[0035] 本发明在虚拟存储系统的基础上,改进对物理内存的管理方法,用于标识冲突开销较高的数据。

[0036] 数据染色被用于标识数据间的冲突开销。这包含两个方面的原因 :1) 由于 Cache 的作用,并非所有对数据的操作都会频繁访问主存 ;2) 数据由大小不多个变量组成,但 bank 组数有限,无法将所有变量都分布到独立的 bank 组中。为尽可能提高访存并行度,需要确定变量间的冲突开销,然后将冲突开销最大的变量散布到不同的 bank 组中。

[0037] 本发明利用变量的大小来判断冲突开销,因为变量占用的存储单元越多,越难在 Cache 中驻留。当处理器同时操作多个变量时,使用变量大小的乘积表示变量间的冲突开销。冲突的开销越大,则需要将这些数据散布到不同的 bank 组。

[0038] 如图 4 所示,表示了对图 2 所示的方法中各步骤的具体实施步骤,包括 :

[0039] 210 :根据符号表的描述和操作系统捕获的内存申请操作,挑选大于 $\text{last_level_cache_size}/2$ ($\text{last_level_cache_size}$ 为最后一级 cache 容量) 的变量 ;

[0040] 本发明通过变量的大小来判断冲突开销,因为变量占用的存储单元越多,越难在 Cache 中驻留。当处理器同时操作多个变量时,使用变量大小的乘积表示变量间的冲突开销。冲突的开销越大,则需要将这些数据散布到不同的 bank 组;

[0041] 220:对变量的地址进行取整操作,起始地址和结束地址分别以基页为大小向下、向上取整,避免同一页帧被染成不同的颜色;然后,以变量的起始地址和结束地址组成结点,将结点置为未被染色状态,插入各个进程独有的单向链表中;

[0042] 230:按结点所表示地址范围的大小对链表递减排序;

[0043] 240:链表中前 bank_groups 个结点占用的内存最多,这些结点所表示的数据之间的冲突开销也越大。因此,前 bank_groups 个结点的颜色各不相同,剩余结点与第 bank_groups-1 个结点同色,避免冲突开销较大的变量被染成相同的颜色;

[0044] 250:为每个 bank 组添加对应的分配函数。操作系统捕获到缺页异常,根据异常地址查询进程独有的链表。如果地址无法与链表中的结点匹配,说明是未染色的数据,调用默认分配函数;否则,根据链表记录的染色结果,调用对应的分配函数。

[0045] 本发明通过使用编译和操作系统提供的信息,将数据划分为多个数据单元,并依据数据单元的大小判断冲突开销。在考虑主存实际配置的情况下,将数据散布到不同的 bank 中,以提高处理器的访存效率。此外,采用该发明不需要添加额外的硬件部件,并且能根据主存的实际配置和程序的访存行为对数据进行划分,更具灵活性,实现简单。

[0046] 本发明根据上述方法实施例,相应地还提供了一种开发存储级并行的内存管理装置实施例,包括相互连接的 bank 分组模块,数据染色模块和内存分配模块,其中:

[0047] Bank 分组模块,用于将 bank 的概念引入内存分配器。通过 bank 分组建立地址与 bank 间的关联,使分配器能根据地址范围标识不同的 bank。Bank 的分组信息输出给数据染色模块和内存分配模块。

[0048] 数据染色模块,用于判断变量间的冲突开销。变量占用的存储单元越多,越难在 Cache 中驻留。当处理器同时操作多个变量时,使用变量大小的乘积表示变量间的冲突开销。冲突的开销越大,则需要将这些数据散布到不同的 bank 组。数据的染色结果输出给内存分配模块。

[0049] 内存分配模块,用于为被染色的数据分配对应的物理页帧。

[0050] 在上述装置实施例中,数据染色模块包括静态数据染色单元和动态数据染色单元,内存分配模块包括依次连接的染色数据判断单元和页帧分配单元;其中:

[0051] 静态数据染色单元,在应用申请 BSS 段时,将符号表中的变量描述作为染色依据,并将染色结果输出给动态数据染色单元;

[0052] 动态数据染色单元,用于将操作系统捕获的内存分配请求,将符合染色条件的动态内存插入链表,根据用户的输入动态调整染色结果;

[0053] 染色数据判断单元,用于根据数据类型判断是否使用染色数据专用的内存分配接口,根据染色结果调用与 bank 组对应的内存分配接口,向页帧分配单元输出物理页帧分配指令;

[0054] 页帧分配单元,用于根据物理页帧分配指令,分配映射到相应 bank 组中的物理页帧,或提示分配失败。

[0055] 以上说明书中描述的只是本发明的具体实施方式,各种举例说明不对本发明的实

质内容构成限制,所属技术领域的普通技术人员在阅读了说明书后可以对以前所述的具体实施方式做修改或变形,而不背离发明的实质和范围。

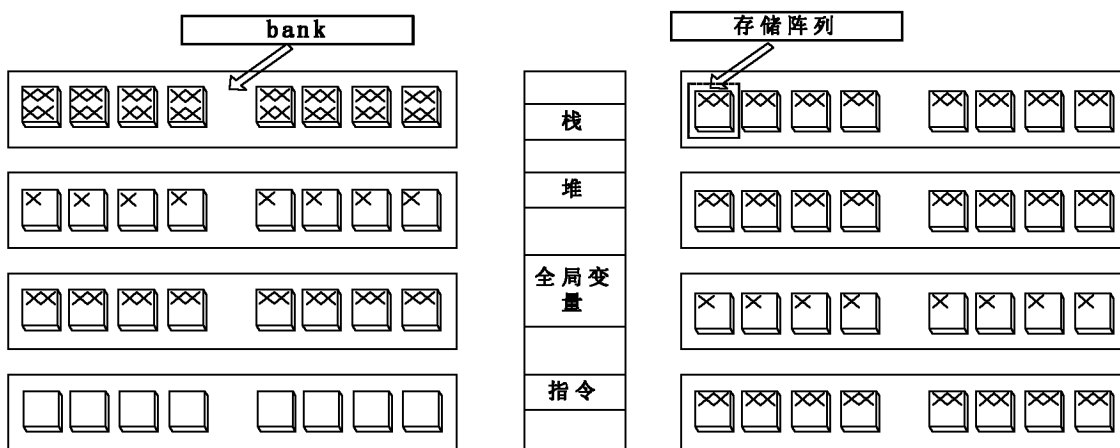


图 1

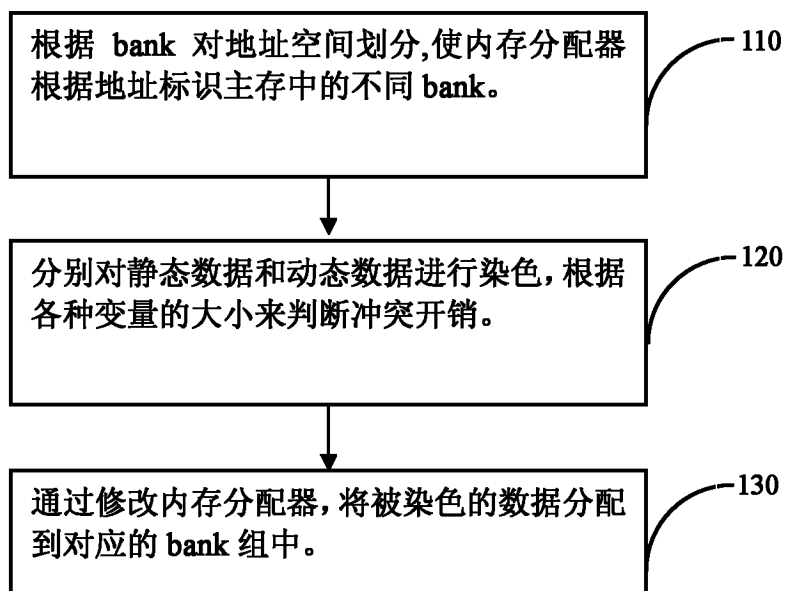


图 2

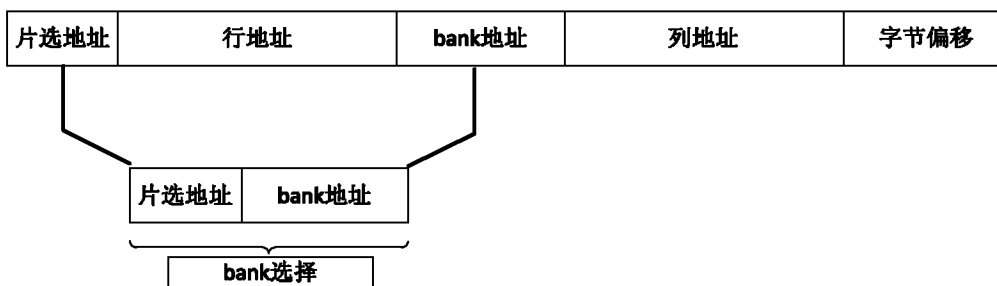


图 3

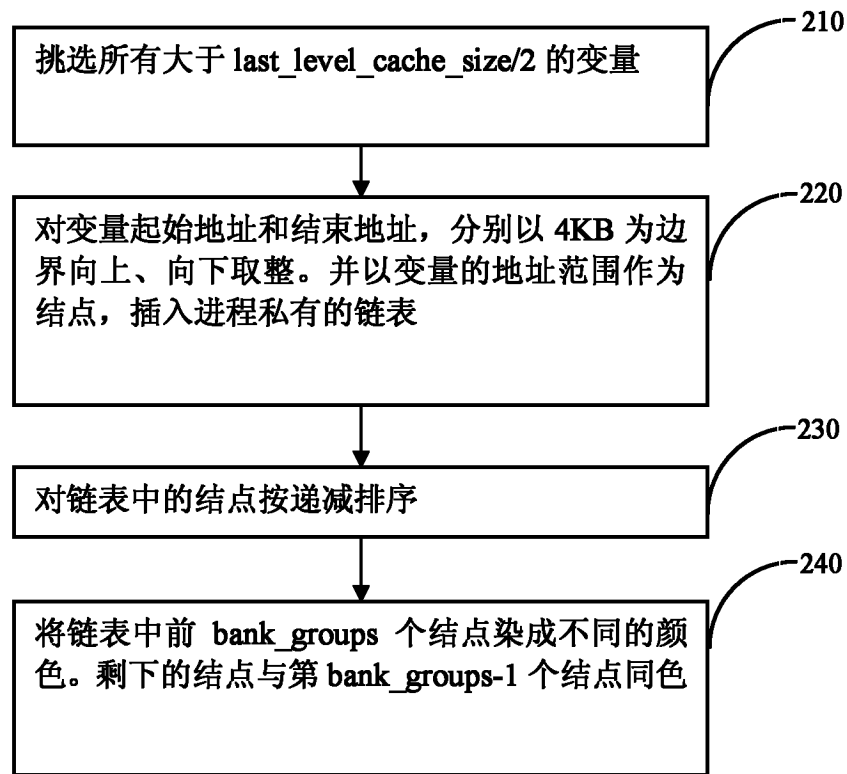


图 4

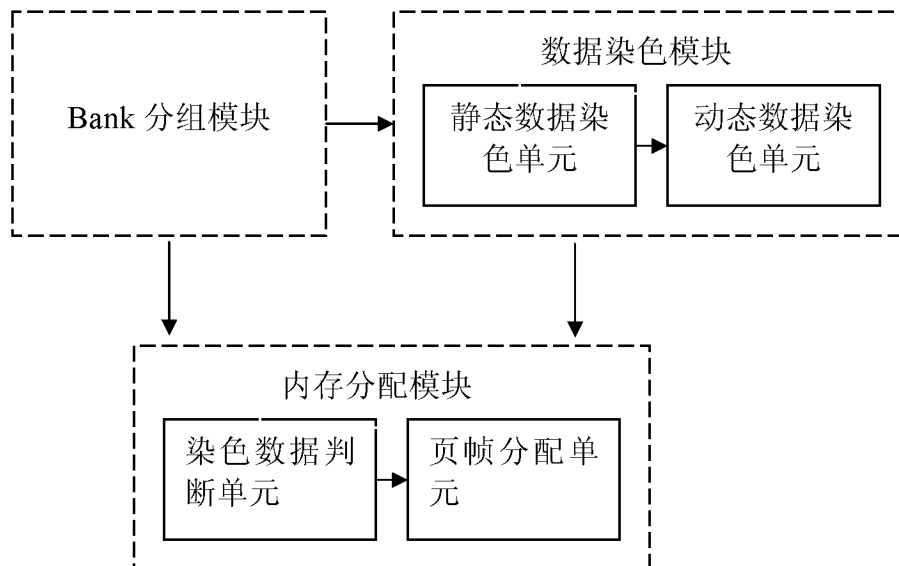


图 5