

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.

G06F 12/06 (2006.01)

G06F 13/00 (2006.01)



[12] 发明专利说明书

专利号 ZL 02145738.7

[45] 授权公告日 2006 年 1 月 18 日

[11] 授权公告号 CN 1237452C

[22] 申请日 2002.10.8 [21] 申请号 02145738.7

[71] 专利权人 凌阳科技股份有限公司

地址 台湾省新竹县科学园区

[72] 发明人 梁伯嵩

审查员 赵 强

[74] 专利代理机构 中科专利商标代理有限责任公司

代理人 汤保平

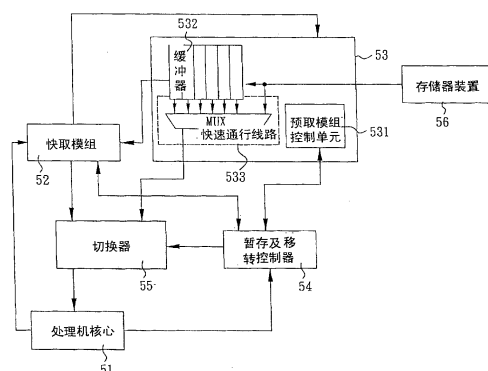
权利要求书 2 页 说明书 9 页 附图 7 页

[54] 发明名称

可暂存及移转指令与资料的存取结构及方法

[57] 摘要

本发明是有关一种可暂存及移转指令与资料的存取结构及方法，其可在预取漏失而由外部存储器读取指令或资料时，先行将所需的指令或资料移转至处理机核心，以达到指令与资料存取效能的最佳化，该结构包括有一处理机核心、一快取模组、一预取模组、一切换器、及一暂存及移转控制器，该切换器用以切换该处理机核心连通至该快取模组或该预取模组，该暂存及移转控制器侦测该预取模组与外部存储器装置介面的资料存取行为，以在该预取模组利用突发传输来传输资料时，驱动该切换器使该预取模组连通至该处理机核心，使得由该外部存储器装置传输至该预取模组的资料亦一并传输至该处理机核心。



1. 一种可暂存及移转指令与资料的存取结构，用以存取一外部存储器装置的指令与资料，其特征在于，该结构主要包括：

一处理机核心，是以单一指令或资料为传输单位，其具有突发传输的能力，可以连续送出指定的突发长度的资料串来进行存取资料；

一快取模组，连接至该处理机核心，并储存有来自该外部存储器装置的部份指令与资料，而提供一可供该处理机核心快速存取指令与资料的存储空间，其中，该快取模组是以数个指令或资料为一次连续传输单位；

一预取模组，连接于该快取模组和该外部存储器装置之间，用以预测该处理机核心将来可能需要使用到的指令与资料，而预先将此指令与资料由该外部存储器装置读入，其中，该预取模组是以数个指令或资料为一次连续传输单位；

一切换器，连接于该处理机核心以及该快取模组与该预取模组之间，用以切换该处理机核心连通至该快取模组或该预取模组，其中，该切换器的预设状态是切换该处理机核心连通至该快取模组；以及

一暂存及移转控制器，与该处理机核心、该快取模组、该预取模组、该切换器相连接以侦测该预取模组与外部存储器装置介面的资料存取行为，以在该预取模组利用突发传输来传输资料时，驱动该切换器使该预取模组连通至该处理机核心，以使得由该外部存储器装置传输至该预取模组的资料亦一并传输至该处理机核心。

2. 如权利要求1所述的可暂存及移转指令与资料的存取结构，其特征在于，其中，该外部存储器装置是为 DRAM。

3. 如权利要求1所述的可暂存及移转指令与资料的存取结构，其特征在于，其中，该外部存储器装置是为 SRAM。

4. 一种可暂存及移转指令与资料的存取方法，用以在具有处理机核心、快取模组、暂存及移转控制器与预取模组的处理机结构中，存取一外部存储器装置的指令与资料，该处理机核心具有突发传输的能力，

可以一突发单位来进行存取资料，该快取模组储存有来自该外部存储器装置的部份指令与资料，而提供一可供该处理机核心快速存取指令与资料的存储空间，该预取模组可预测该处理机核心将来可能需要使用到的指令与资料，而预先将此指令与资料由该外部存储器装置读入，其特征在于，该方法包括步骤：

5 (A) 该处理机核心发出一读取要求时，将读取地址送往该暂存及移转控制器，用以查询是否在快取模组或预取模组内存在有该读取地址所指示的有效指令或资料；若是查询结果为所指示的有效指令或资料不存在于快取模组而是存在于预取模组中时，其造成一快取漏失，但预取
10 命中，因此以一突发单位来读取资料或指令；

(B) 该暂存及移转控制器判断该突发单位的所有资料是否已全部载入该预取模组，如该突发单位的所有资料已全部载入该预取模组，则将该预取模组的该指令或资料送入该处理机核心；

15 (C) 如该突发单位的所有资料尚未全部载入该预取模组，且所要读取的指令或资料已经读入该预取模组，则将该预取模组的该指令或资料送入处理机核心；以及

(D) 如该突发单位的所有资料尚未全部载入该预取模组，且所要读取的指令或资料尚未读入该预取模组，则等待该所需的指令或资料读入后，再将该预取模组的该指令或资料送入处理机核心。

可暂存及移转指令与资料的存取结构及方法

5

技术领域

本发明是关于一种存取指令与资料的结构及方法，尤指一种可暂存及移转指令与资料的存取结构及方法。

10 背景技术

在现今的电脑系统的处理机的结构中，由于处理机核心的处理速度大幅提高，而使存储器的存取成为效能的瓶颈，因此，快取（Cache）结构便成为一重要的改善方式。快取是将以固定长度的数个指令或数笔资料的长度为一单位来存取存储器，此一单位称为一快取丛（Cache
15 Line）。此快取丛的大小则与存储器的存取单位有关。例如，在现今许多具有突发传输（Burst Transfer）的能力的存储器中，其资料存取的进行可以只需给定一次地址与各项相关设定后，连续送出指定的突发长度（Burst Length）的资料串，因而减少在传送资料的前的设定所花的初始化延迟（Initial Delay）时间。而在此种存储器中，快取丛的大小即相关
20 于突发长度（Burst Length）的大小。

图1即显示一具有前述的快取功能的处理机结构，其中，在处理机运作时，若所需的指令或资料所属的快取丛正好在快取模组11中时，处理机核心14可以快速取得该指令，而无时间延迟，或是时间延迟极低。然而，若是所需的指令或是资料不在快取模组11中，则发生快
25 取漏失（Cache Miss），此时必须从外界存储器13读入所需的资料，此动作称作快取填回（Cache Refill）。由于整个快取丛皆需读入置于快取模组11中，会造成的相当高的系统延迟，称为快取漏失惩罚（Cache Miss Penalty）。

快取漏失惩罚常在处理机第一次执行存取到某一程序码段或资料段
30 时连续出现，因而严重影响系统效能。为解决此一问题，预取（Prefetch）

便成提高效率的重要技巧。如图 2 所示，其是以预取模组 1 2 来预测处理机核心 1 4 接下来可能需要使用到的程序码区或资料区域，而预先将此区域读入预取模组 1 2 内。一旦处理机核心 1 4 发现无法从快取模组 1 1 中取得所需的指令或资料而发生快取漏失时，即可以在预取模组 1 2 搜寻。若所需的指令或资料已经被预先取入预取模组 1 2，即可以从预取模组 1 2 中，将所需的快取丛读入快取模组 1 1 内，因而大幅减少快取漏失惩罚时所需的代价。但若所需的指令或资料亦不在预取模组 1 2，则会发生预取漏失（Prefetch Miss），而仍然必需至外界存储器 1 3 中取得所需的快取丛。亦有相当高的系统延迟，称为预取漏失惩罚（Prefetch Miss Penalty）。

前述的预取模组 1 2 在计算机组织中介于外部记忆储存装置 1 3、快取模组 1 1 与处理机核心 1 4 之间，传统的结构均直接延用快取模组 1 1 的组态，而使用同于快取丛的资料长度，也就是动态存储器突发传输长度作为资料传输单位。然而，由于预取模组 1 2 与快取模组 1 1 以及处理机核心 1 4 之间并非动态存储器的介面，因此并不需要以快取丛作为资料传输单位。以快取丛作为资料传输单位，反而会使得资料传输时迟滞而不灵活。

具体而言，在具有快取模组 1 1 与预取模组 1 2 的处理机结构中，模组间具有三个主要介面。第一介面 1 5 为外部介面，其衔接外部记忆储存装置 1 3。第二介面 1 6 为预取模组 1 2 与快取模组 1 1 介面。第三介面 1 7 为快取模组 1 1 与处理器核心 1 4 介面，此介面 1 7 用以将快取模组中的指令与资料送入处理器。传统上，在第一、二介面 1 5 及 1 6 中，是使用相同于快取丛的资料长度作为资料传输单位，而第三介面 1 7 的资料存取，若与第一、二介面的存取相关时，也多需要等待快取丛的资料长度存取完成后，才能进行动作。然而，快取丛的资料长度并不一定是预取模组 1 6 与外部记忆储存装置 1 3、快取模组 1 1（快取存储器）与处理机核心 1 4 之间最佳的传输资料单位。因为快取丛的大小，跟快取模组 1 1 的存储器组织结构相关，在处理机执行过程中，快取丛的大小固定。然而，在处理机实际执行程序，在撷取指令和存取资料时，其行为模式不断的动态改变。以固定的快取丛的资料长度作为

资料传输单位，难以对各种状况作最佳化，会在不同状况下发生不经济的资源浪费。

如图3的时序图所示，为说明方便起见，假设处理机核心12需求指令或资料（Instruction / Data）的顺序是由#0指令或资料起循序执行，初始化延迟在图中标示为 Init，由预取模组12所取得的指令或资料先送进快取模组11，再由快取模组11送进处理机核心14（CPU core）。因#0指令或资料的读取需求引发快取漏失，需等待预取模组12完成对外界存储器装置13的读取动作，将资料传至快取模组11，才能让处理机核心14取得所需的指令或资料。在#4指令或资料的读取需求引发快取漏失，此时由于预取模组的作用发挥，所以不再需要等待对外界存储器装置13的读取动作，直接由预取模组12将资料传至快取模组11，再由处理机核心14取得所需的指令或资料。

图4显示另一范例时序图，其假设现今有一读取指令或资料的序列，在#2指令或资料的读取需求后，即跳到#80指令或资料的读取，因读取#80指令或资料的需求引发快取漏失，然而因为#80指令或资料并非预取模组12中已读到的#4—#7的指令或资料，因而进一步发生预取漏失，原预取模组12中的资料需舍弃。且预取模组12须再启动一次完成对外界存储器装置13的读取动作，将资料传至快取模组11，才能让处理机核心14取得所需的指令或资料。此一延宕，让处理机核心14停顿而严重影响效能。由此可知，前述现有的指令与资料存取的结构仍有予以改进的必要。

发明人爰因于此，本于积极发明的精神，亟思一种可以解决上述问题的“可暂存及移转指令与资料的存取结构及方法”，几经研究实验终至完成此项嘉惠世人的发明。

发明内容

本发明的主要目的是在提供一种可暂存及移转指令与资料的存取结构及方法，其可在预取漏失而由外部存储器读取指令或资料时，先行将所需的指令或资料移转至处理机核心，以达到指令与资料存取效能的最佳化。

依据本发明的一特色，所提出的一种可暂存及移转指令与资料的存取结构，用以存取一外部存储器装置的指令与资料，其特征在于，该结构主要包括：

一处理机核心，是以单一指令或资料为传输单位，其具有突发传输的能力，可以连续送出指定的突发长度的资料串来进行存取资料；

一快取模组，连接至该处理机核心，并储存有来自该外部存储器装置的部份指令与资料，而提供一可供该处理机核心快速存取指令与资料的存储空间，其中，该快取模组是以数个指令或资料为一次连续传输单位；

一预取模组，连接于该快取模组和该外部存储器装置之间，用以预测该处理机核心将来可能需要使用到的指令与资料，而预先将此指令与资料由该外部存储器装置读入，其中，该预取模组是以数个指令或资料为一次连续传输单位；

一切换器，连接于该处理机核心以及该快取模组与该预取模组之间，用以切换该处理机核心连通至该快取模组或该预取模组，其中，该切换器的预设状态是切换该处理机核心连通至该快取模组；以及

一暂存及移转控制器，与该处理机核心、该快取模组、该预取模组、该切换器相连接以侦测该预取模组与外部存储器装置介面的资料存取行为，以在该预取模组利用突发传输来传输资料时，驱动该切换器使该预取模组连通至该处理机核心，以使得由该外部存储器装置传输至该预取模组的资料亦一并传输至该处理机核心。

其中，该外部存储器装置是为 DRAM。

其中，该外部存储器装置是为 SRAM。

本发明一种可暂存及移转指令与资料的存取方法，用以在具有处理机核心、快取模组、暂存及移转控制器与预取模组的处理机结构中，存取一外部存储器装置的指令与资料，该处理机核心具有突发传输的能力，可以一突发单位来进行存取资料，该快取模组储存有来自该外部存储器装置的部份指令与资料，而提供一可供该处理机核心快速存取指令与资料的存储空间，该预取模组可预测该处理机核心将来可能需要使用到的指令与资料，而预先将此指令与资料由该外部存储器装置读入，其

特征在于，该方法包括步骤：

（A）该处理机核心发出一读取要求时，将读取地址送往该暂存及移转控制器，用以查询是否在快取模组或预取模组内存在有该读取地址所指示的有效指令或资料；若是查询结果为所指示的有效指令或资料不存在于快取模组而是存在于预取模组中时，其造成一快取漏失，但预取命中，因此以一突发单位来读取资料或指令；

（B）该暂存及移转控制器判断该突发单位的所有资料是否已全部载入该预取模组，如该突发单位的所有资料已全部载入该预取模组，则将该预取模组的该指令或资料送入该处理机核心；

10 （C）如该突发单位的所有资料尚未全部载入该预取模组，且所要读取的指令或资料已经读入该预取模组，则将该预取模组的该指令或资料送入处理机核心；以及

（D）如该突发单位的所有资料尚未全部载入该预取模组，且所要读取的指令或资料尚未读入该预取模组，则等待该所需的指令或资料读入后，再将该预取模组的该指令或资料送入处理机核心。

15 由于本发明构造新颖，能提供产业上利用，且确有增进功效，故依法申请发明专利。

附图说明

20 为能审查员能更了解本发明的技术内容，特举一较佳具体实施例及附图说明如下，其中：

图 1 是显示一具有快取功能的现有处理机结构。

图 2 是显示一具有快取及预取功能的现有处理机结构。

图 3 是显示图 2 的处理机结构的一工作时序图。

25 图 4 是显示图 2 的处理机结构的另一工作时序图。

图 5 是显示本发明的可暂存及移转指令与资料的存取结构。

图 6 是显示以本发明的可暂存及移转指令与资料的存取方法的控制流程。

图 7 是显示依据本发明的可暂存及移转指令与资料的存取结构的一工作时序图。

图 8 是显示依据本发明的可暂存及移转指令与资料的存取结构的另一工作时序图。

具体实施方式

5 图 5 显示本发明的可暂存及移转指令与资料的存取结构，其包括一处理机核心 5 1、一快取模组 5 2、一预取模组 5 3、一暂存及移转控制器 5 4、一切换器 5 5 及一外部存储器装置 5 6，以一微处理器结构为例，该外部存储器装置 5 6 是为 DRAM 或 SRAM 的存储器，该处理机核心 5 1 则具有突发传输（Burst Transfer）的能力，可以连续送出指定的突发长度（Burst Length）的资料串来进行存取资料。

前述快取模组 5 2 用以储存有来自外部存储器装置 5 6 的部份指令与资料，而提供一可供处理机核心 5 1 快速存取指令与资料的存储空间，前述预取模组 5 3 是用以预测处理机核心 1 4 将来可能需要使用到的指令与资料，而预先将此指令与资料读入预取模组 5 3，如图所示，
15 该。预取模组 5 3 主要包括一预取控制单元 5 3 1、一缓冲器 5 3 2 及一快速通行线路 5 3 3，前述缓冲器 5 3 2 是用以预存可能需要使用到的指令与资料，前述快速通行线路 5 3 3 是例如为一多工器，其可在该预取控制单元 5 3 1 的控制下，将预存的指令与资料送出至切换器 5 5。

20 前述切换器 5 5 是由该暂存及移转控制器 5 4 所控制，以切换该处理机核心 5 1 连通至该快取模组 5 2 或该预取模组 5 3，在一般的状态，该切换器 5 5 是切换该处理机核心 5 1 连通至该快取模组 5 2。该暂存及移转控制器 5 4 是侦测处理器核心 5 1 的存取要求（Request）所需指令或资料的地址，当处理器核心 5 1 发出存取要求时，暂存及移转
25 控制器 5 4 会检查储存于快取模组 5 2 与预取模组 5 3 内，是否存有该地址所指示的有效指令或资料，参照图 6 所示，其控制流程包含如下的步骤：

步骤 S 6 0 1：对于处理器核心 5 1 的存取要求，其首先检查是否为读取要求；

30 步骤 S 6 0 2：若处理器核心 5 1 是要写出资料，则直接送往快取

模组 5 2 或存储器装置 5 6 处理；

步骤 S 6 0 3：若处理器核心 5 1 是要读取资料，则将读取地址送往暂存及移转控制器 5 4；

步骤 S 6 0 4：该暂存及转移控制器 5 4 查询此地址在快取模组 5 2 与预取模组 5 3 内，是否存有该位该地址所指示的有效指令或资料，如是，表示为“快取命中”状态，执行步骤 S 6 0 5；

步骤 S 6 0 5：快取模组 5 2 送出该指令或资料，切换器 5 5 将来自快取模组 5 2 的指令或资料送入处理机核心 5 1；

若步骤 S 6 0 4 判断该资料或指令不在快取模组，或者离然存在，但非有效状态，则为“快取漏失”状态，则执行以下的步骤：

步骤 S 6 0 6：检查该资料或指令所属的突发单位（Burst）是否在预取模组中，若无，则表示同时亦为“预取漏失”状态，此时启动预取漏失程序；若有，即为“预取命中”状态。

在预取命中的状况下，依据步骤 S 6 0 7、S 6 0 7 及 S 6 0 8 的判断，预取模组 5 3 可能有下列状况：

（a）资料或指令所属的 Burst 的所有资料，已全部载入预取模组 5 3（步骤 S 6 0 7 为是）；而且预取模组 5 3 并未处于传输资料给快取模组 5 2 的状态（步骤 S 6 0 8 为否）。于此状况，因为此时预取模组 5 3 内的存储器并未有任何其他的读写动作，故预取模组 5 3 可以送出该指令或资料；而切换器 5 5 可将来自预取模组 5 3 的指令或资料送入处理机核心 5 1。

（b）资料或指令所属的 Burst 的所有资料，尚未全部载入预取模组 5 3（步骤 S 6 0 7 为否）。其又可分两种状况：（b 1）所需要的该笔指令或资料已经读入预取模组 5 3（步骤 S 6 0 9 为是），此时切换器 5 5 可将来自预取模组 5 3 的指令或资料送入处理机核心 5 1；（b 2）所需要的该笔指令或资料尚未读入预取模组 5 3（步骤 S 6 0 9 为否），此时尚需等待，直到该笔所需的指令或资料读入，的后，切换器 5 5 将来自预取模组 5 3 的该指令或资料送入处理机核心 5 1。举例来说，假设突发长度（Burst Length）为 8，现在所需的资料是此 Burst 的第 4 笔。然而当处理器核心提出读取资料需求时，预取模组才刚读取到此 Burst 的

第2笔。此时需等待两个时脉周期，等到此 Burst 的第4笔由存储器装置56读入预取模组53的同时，也一并透过快速通行线路533，同时馈送往切换器55以进入处理机核心51。对一般传统结构而言，需要等待到全部的 Burst 读取完成后才能馈送，亦即至少要等七个时脉周期后5 才能进行馈送，因此，本发明所揭示的方法明显较有效率。

(c) 预取模组53正处于传输资料给快取模组52的状态（步骤S608为是）。因为本发明的快速通行线路533为一独立的读出埠，其不会影响到预取模组53传输资料给快取模组52的资料读出动作。故此时仍可送出该指令或资料，切换器55将来自预取模组53的指令10 或资料送入处理机核心51。在传统结构中，则需将整个快取丛长度的资料从预取模组传输资料给快取模组，处理器核心才能从快取模组取得资料或指令，如此耗时甚久，浪费时间。因此可得知本发明的方法明显较有效率。

由以上的方法，处理机核心51可无需等待资料由外部存储器装置15 56传输至预取模组53、再由预取模组53传输至快取模组52、最后由快取模组52传输至处理机核心51的冗长过程。

以前述的结构，因处理机核心51是以单一指令或资料为单位，而快取模组52及预取模组53是以数个指令资料为一次连续传输单位，因此，此暂存及移转控制器54可以在快取模组52及预取模组53单20 一传输单位完成前，将已由外界取得的指令或资料送入处理机核心51。例如，假设由外部存储器装置56传输至预取模组53的突发传输长度为八个指令长度，然而若此时处理机核心51需要该次突发传输中前二个指令，在如图2所示的现有结构中，预取模组1的必须在第一介面15中等待传完八个指令长度的突发传输，处理机核心14才能透过25 第二介面16来取得前四个指令。然而在图5所示的本发明的结构中，在预取模组53接收到突发传输的八个指令的前四个指令后，暂存及移转控制器54即可控制切换器55而传送资料给处理机核心51。

图7显示一范例时序图，其假设处理机核心51需求指令或资料的顺序是由#0指令或资料起循序执行，初始化延迟在图中标示为 Init。其中，当#0指令或资料的读取需求引发快取漏失，而预取模组53进行30

对外界存储器装置 5 6 读取# 0、# 1、# 2、# 3 指令或资料时，该暂存及移转控制器 5 4 驱动切换器 5 5 使该预取模组 5 3 连通至该处理机模组 5 1，以将# 0、# 1、# 2、# 3 指令或资料传至处理机核心 5 1。而在# 4 指令或资料的读取需求引发快取漏失时，此时由于预取模组 5 3 的作用，而无需要等待对外界存储器装置 5 6 的读取动作，直接由预取模组 5 3 将资料传至快取模组 5 2 及处理机核心 5 1。

图 8 显示一不连续的指令与资料存取的时序图，其中，有一读取指令或资料的序列，在# 2 指令或资料的读取需求后，即跳到# 8 0 指令或资料的读取。由于从外部存储器装置 5 6 送至预取模组 5 3 的指令与资料可以快速送入处理机核心 5 1 利用，因而当发生读取# 8 0 指令或资料引发快取漏失时，尚可及时修正预取模组 5 3 存取资料的地址，因而不致造成预取漏失。预取模组 5 3 在下一次从外界存储器读入指令或资料时，已经是正确的# 8 0—# 8 3 指令或资料，因而删去预取漏失所造成的延误，而可提高效率。

由以上的说明可知，本发明由暂存及移转控制器 5 4 的调整，可以在指令或资料尚未完全传输到预取模组 5 3 前，只要所需的部分资料已经可用，即启动预取模组 5 3 与处理机核心 5 1 间的传输，因而大幅所短整体所需时间。

综上所述，本发明无论就目的、手段及功效，均不同于现有的技术特征，实为一极具实用价值的发明。惟应注意的是，上述诸多实施例仅是为了便于说明而举例而已，本发明所主张的权利范围自应以申请专利范围所述为准，而非仅限于上述实施例。

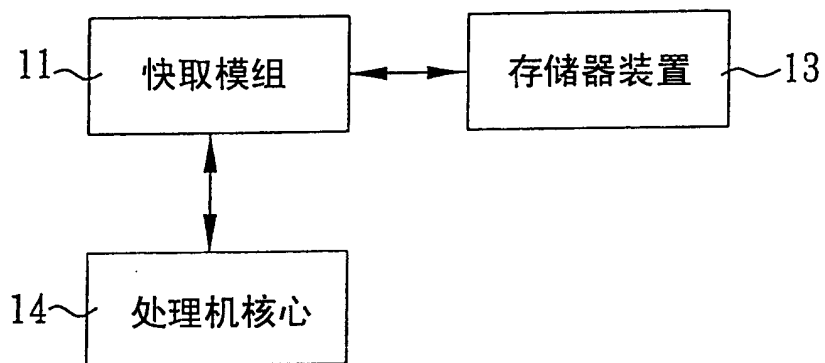


图 1

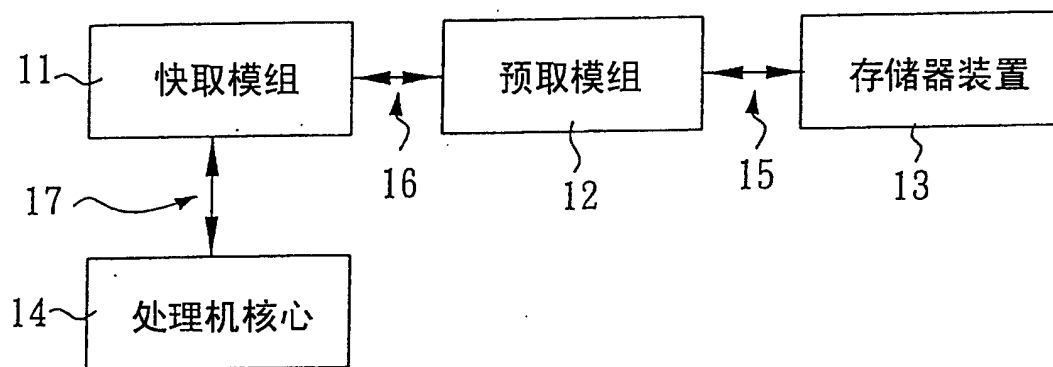


图 2

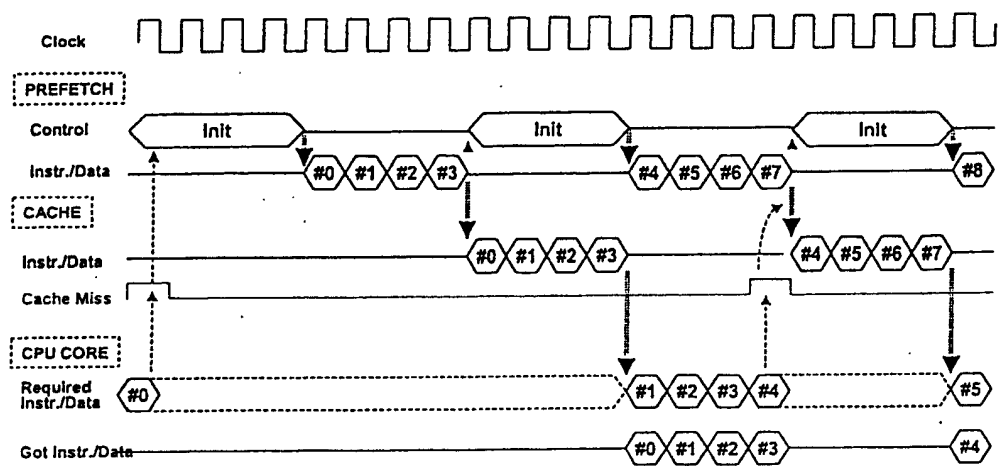


图 3

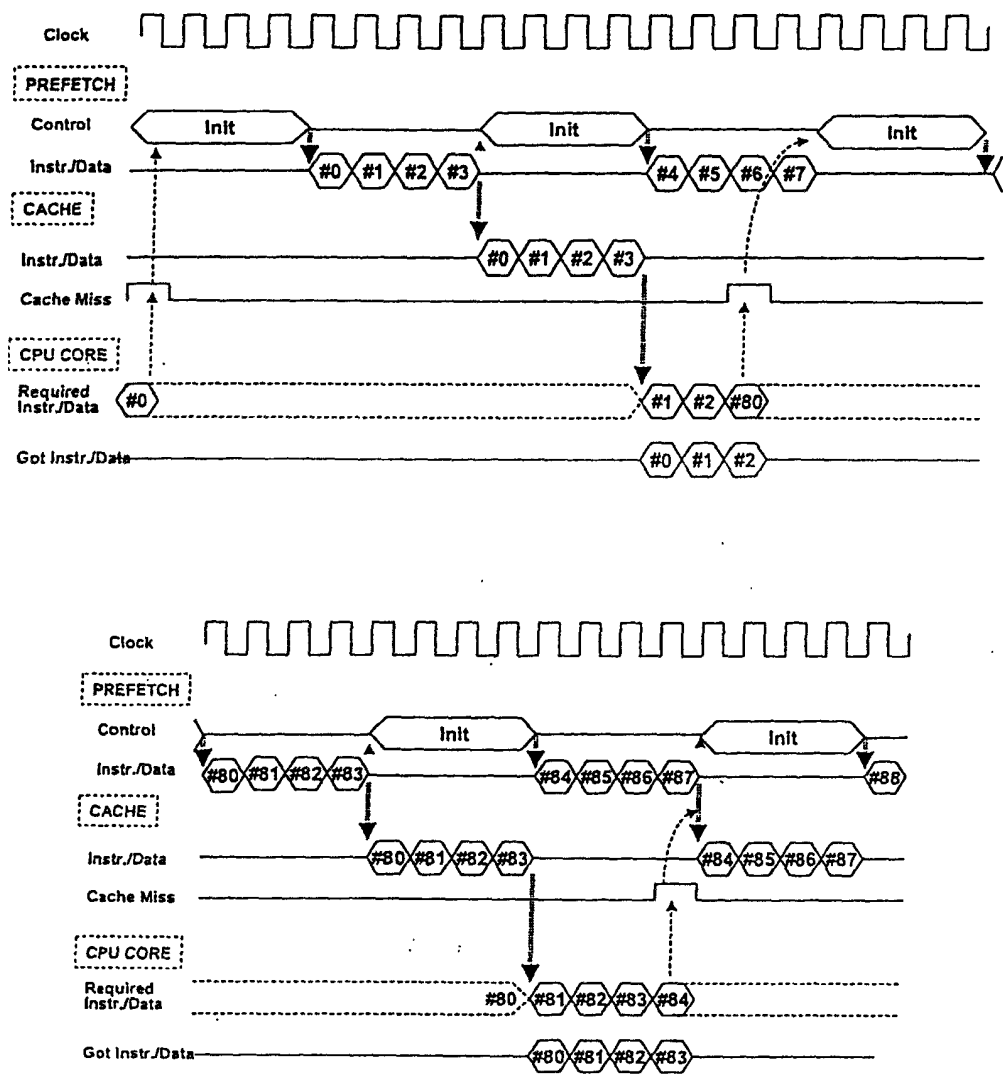


图 4

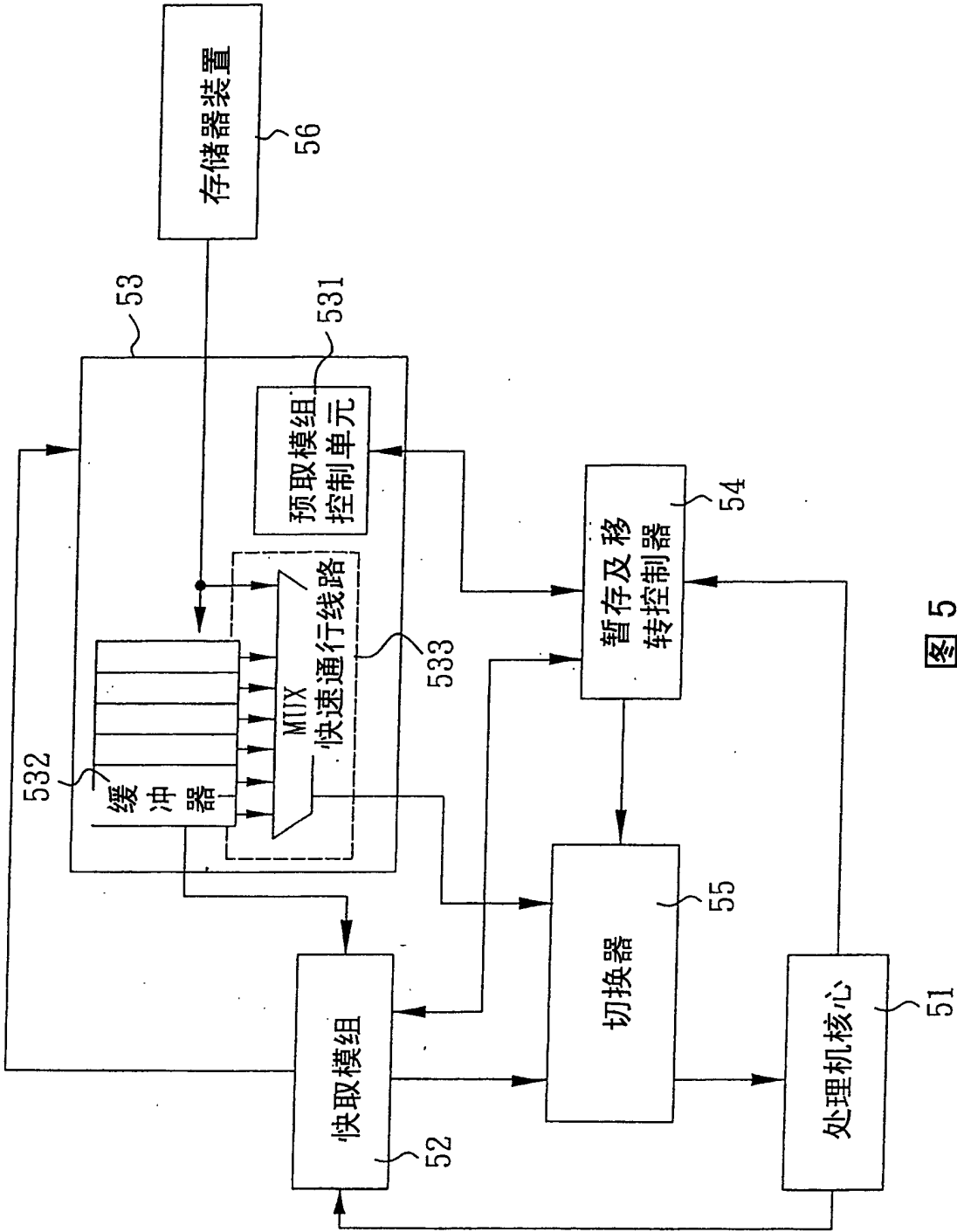


图 5

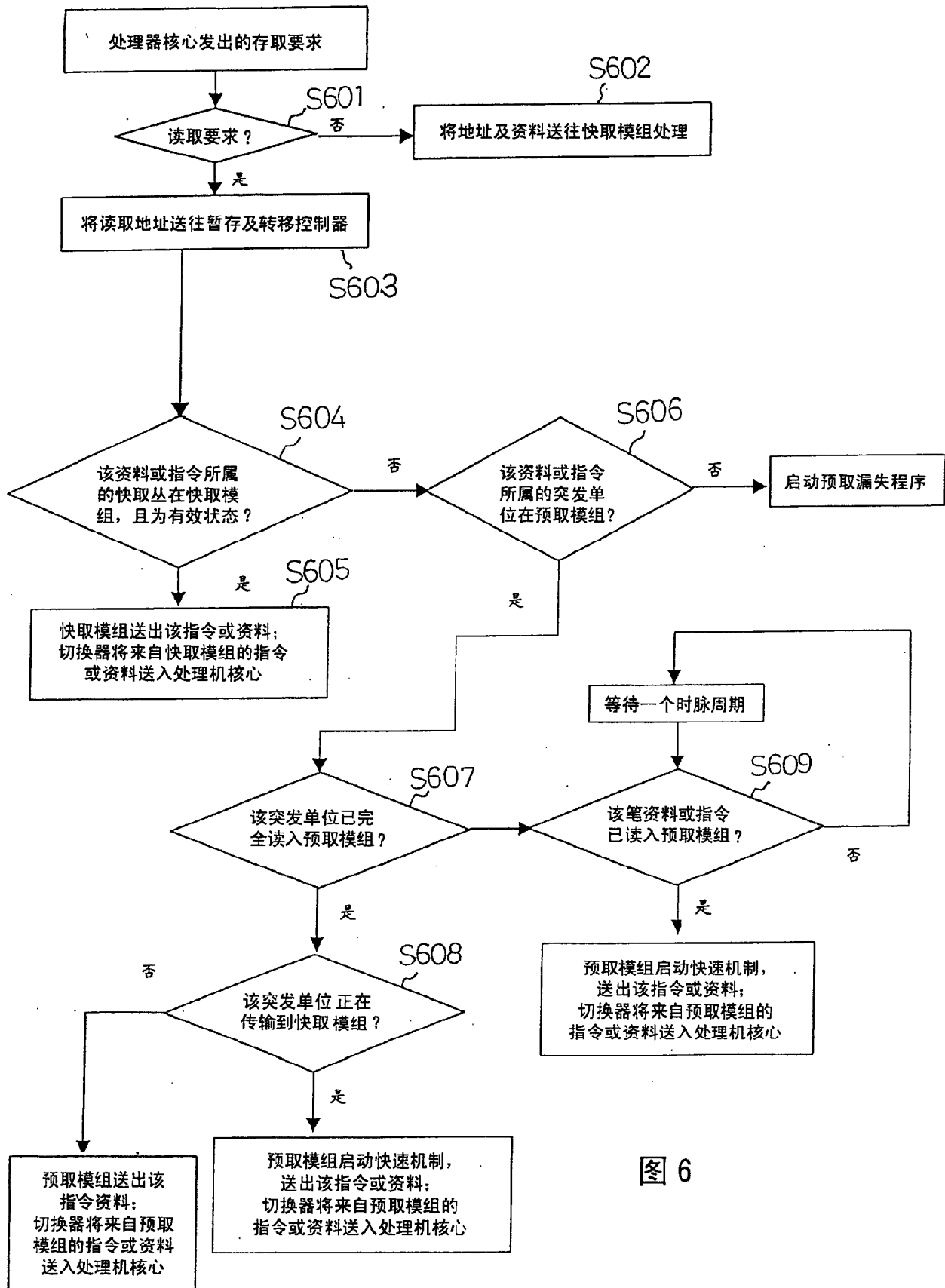


图 6

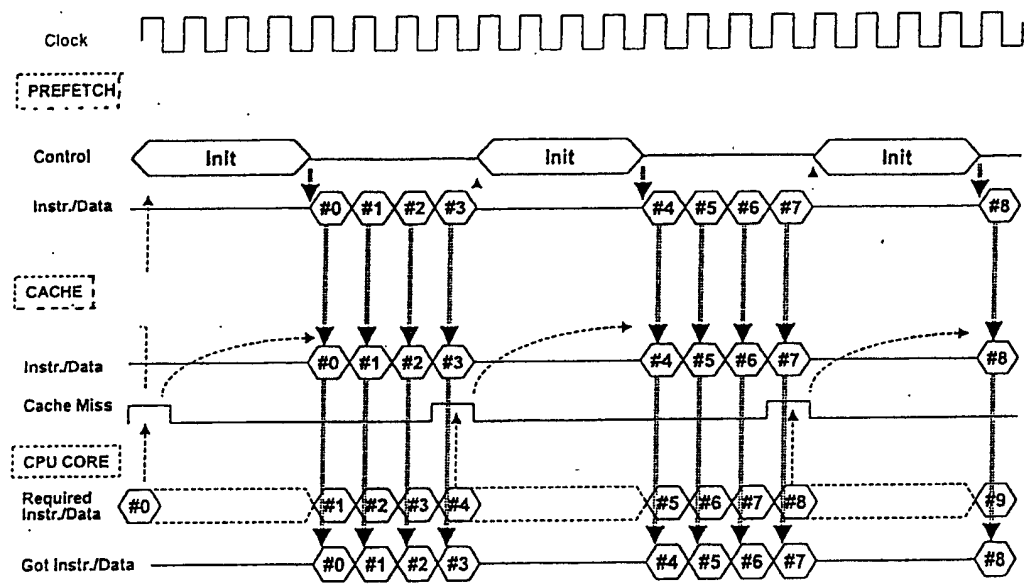


图 7

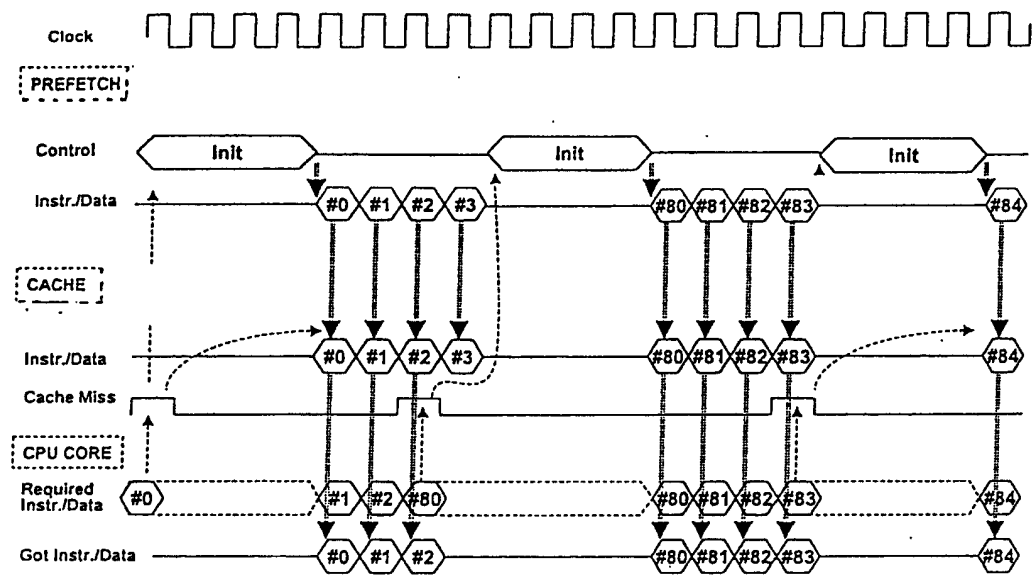


图 8