



(12) 发明专利

(10) 授权公告号 CN 112558889 B

(45) 授权公告日 2021.05.28

(21) 申请号 202110214197.9

审查员 杨佳玉

(22) 申请日 2021.02.26

(65) 同一申请的已公布的文献号

申请公布号 CN 112558889 A

(43) 申请公布日 2021.03.26

(73) 专利权人 北京微核芯科技有限公司

地址 100082 北京市海淀区中关村东路66

号1号楼10层1105-2

(72) 发明人 赵继业 郇丹丹

(74) 专利代理机构 北京超凡宏宇专利代理事务

所(特殊普通合伙) 11463

代理人 曹瑞敏

(51) Int. Cl.

G06F 3/06 (2006.01)

G06F 9/50 (2006.01)

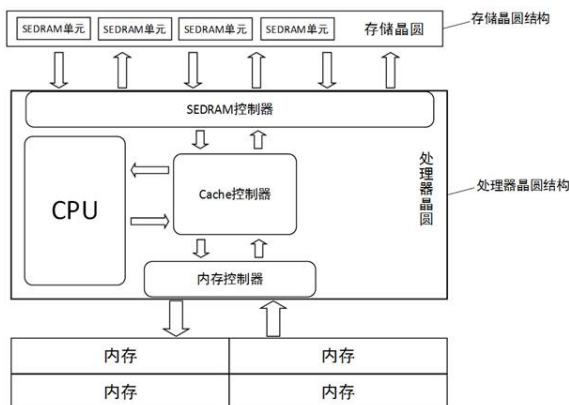
权利要求书3页 说明书13页 附图6页

(54) 发明名称

一种基于SED-DRAM的堆叠式Cache系统、控制方法和Cache装置

(57) 摘要

本发明涉及一种基于SED-DRAM的堆叠式Cache系统、控制方法和Cache装置,该堆叠式Cache系统集成于多层键合晶圆,并包括高速缓冲存储器、Cache控制器及SED-DRAM控制器;多层键合晶圆包括存储晶圆结构和处理器晶圆结构;SED-DRAM单元集成于存储晶圆结构中的每一层存储晶圆中,并用作高速缓冲存储器的存储空间;处理器晶圆结构集成有CPU、Cache控制器、SED-DRAM控制器及内存控制器。上述Cache系统采用集成于存储晶圆结构每一层存储晶圆的SED-DRAM单元作为高速缓冲存储器的存储空间,可以极大地提高高速缓冲存储器的容量和带宽,提高高速缓冲存储器的命中率,加快热点数据的读取速度,提升CPU内部读取数据的命中率,同时还可以节省处理器芯片的存储资源。



1. 一种基于SEDAM的堆叠式Cache系统,其特征在于,所述堆叠式Cache系统集成于多层键合晶圆,并包括:高速缓冲存储器、Cache控制器以及SEDAM控制器;

所述多层键合晶圆包括存储晶圆结构和处理器晶圆结构;

所述存储晶圆结构包括至少一层存储晶圆;

所述处理器晶圆结构包括至少一层处理器晶圆;

SEDAM单元集成于所述存储晶圆结构的每一层存储晶圆中,并用作所述高速缓冲存储器的存储空间;

所述处理器晶圆结构,集成有CPU、Cache控制器、SEDAM控制器及内存控制器;

晶圆之间通过键合结构进行管脚对管脚的数据直接传输;

所述晶圆之间通过键合结构进行管脚对管脚的数据直接传输,包括:

所述存储晶圆结构中的存储晶圆与所述处理器晶圆结构中的处理器晶圆之间通过键合结构进行管脚对管脚的数据直接传输;

所述存储晶圆结构中的存储晶圆之间通过键合结构进行管脚对管脚的数据直接传输;

所述处理器晶圆结构中的处理器晶圆之间通过键合结构进行管脚对管脚的数据直接传输;

所述Cache控制器用于判断所述高速缓冲存储器是否命中,向所述CPU、所述SEDAM控制器以及所述内存控制器发出相应的命令操作及数据交互请求;

所述SEDAM控制器控制读出或写入所述SEDAM单元的内容。

2. 根据权利要求1所述的堆叠式Cache系统,其特征在于,所述SEDAM单元存储的内容,包括控制域和Cache数据域,所述控制域包括Cache状态域、Cache标签域。

3. 根据权利要求2所述的堆叠式Cache系统,其特征在于,所述控制域还包括目录域信息。

4. 根据权利要求1所述的堆叠式Cache系统,其特征在于,所述处理器晶圆结构还包括集成于所述处理器晶圆结构的Cache SRAM,所述SEDAM单元和Cache SRAM用作所述高速缓冲存储器的存储空间。

5. 根据权利要求4所述的堆叠式Cache系统,其特征在于,所述Cache控制器用于判断所述高速缓冲存储器是否命中,向所述CPU、所述SEDAM控制器以及所述内存控制器发出相应的命令操作及数据交互请求;

所述SEDAM控制器控制读出或写入所述SEDAM单元的内容;

所述Cache控制器还用于读出或写入所述Cache SRAM中存储的内容。

6. 根据权利要求5所述的堆叠式Cache系统,其特征在于,所述SEDAM单元存储的内容为Cache数据域;

所述Cache SRAM存储的内容为控制域;

所述控制域包括Cache状态域和Cache标签域。

7. 根据权利要求6所述的堆叠式Cache系统,其特征在于,所述控制域还包括有目录域信息。

8. 根据权利要求1-7任一项所述的堆叠式Cache系统,其特征在于,采用所述SEDAM单元作为所述高速缓冲存储器的其中任一级的高速缓冲存储器的存储空间。

9. 根据权利要求8所述的堆叠式Cache系统,其特征在于,所述处理器晶圆结构设置一

层处理器晶圆,所述存储晶圆结构设置一层或两层存储晶圆。

10. 根据权利要求1-7任一项所述的堆叠式Cache系统,其特征在于,所述处理器晶圆结构设置一层处理器晶圆,所述存储晶圆结构设置一层或两层存储晶圆。

11. 一种用于如权利要求1-3中任一项所述的堆叠式Cache系统的Cache控制方法,其特征在于,所述Cache控制方法将集成于存储晶圆结构每一层的存储晶圆的SEDRAM单元,作为高速缓冲存储器的整体存储空间进行控制,控制域和Cache数据域都存储在所述SEDRAM单元中,具体步骤如下:

第一步骤:集成于处理器晶圆结构的CPU输出访存地址;

第二步骤:根据访存地址,集成于所述处理器晶圆结构的Cache控制器发送读请求给集成于所述处理器晶圆结构的SEDRAM控制器,所述SEDRAM控制器从所述SEDRAM单元中读出所需要的Cache Line,所述Cache Line包括所述控制域和所述Cache数据域,所述控制域包括Cache状态域和Cache标签域;

第三步骤:对访存地址中对应于Tag的位与所述Cache Line中的所述Cache标签域进行地址比较;

第四步骤:如果访存地址对应于Tag的位与其中一路的Cache标签域完全相同、且所述Cache Line的状态位处于有效状态,则判断所述高速缓冲存储器访问命中,否则判断所述高速缓冲存储器访问未命中;

第五步骤:当所述高速缓冲存储器访问命中后,所述Cache控制器直接把所述Cache Line中的所述Cache数据域内容读回到CPU中,作为本次访存操作的结果,并直接执行第十步骤;

第六步骤:当所述高速缓冲存储器访问未命中时,如果存在下一级高速缓冲存储器,所述Cache控制器向下一级高速缓冲存储器发送访存操作请求,并等待所述访存操作请求所对应的数据返回;当所述访存操作请求所对应的数据返回所述Cache控制器后,执行第八步骤;

第七步骤:如果所述高速缓冲存储器访问未命中,且本级高速缓冲存储器是最后一级高速缓冲存储器时,所述Cache控制器直接通过集成于所述处理器晶圆结构的内存控制器向内存发送访存操作请求,并等待访问内存的访存操作请求所对应的数据返回;当访问内存的访存操作请求所对应的数据返回所述Cache控制器后,执行第八步骤;

第八步骤:所述Cache控制器将返回的访存数据发送给CPU,同时通过高速缓冲存储器替换算法找出一个Cache路中的Cache Line;被找出的Cache Line作为被替换的Cache Line,根据被替换的Cache Line是未被使用的Cache Line、被使用但非脏的Cache Line、或脏的Cache Line,以及Cache写策略进行写回块的处理;

第九步骤:所述Cache控制器将第八步骤中返回的访存数据填充到被替换的Cache Line所对应的SEDRAM单元存储空间中,由所述SEDRAM控制器向所述存储晶圆结构上与写地址对应的SEDRAM单元发出写操作,写操作的写地址由所述Cache控制器根据第八步骤中找出的被替换的Cache Line给出,写数据是需要被写入的Cache Line的完整内容,包括控制域和Cache数据域,其中控制域信息根据访问Cache Line的请求确定,Cache数据域内容由第八步骤中返回的访存数据确定;

第十步骤:本次高速缓冲存储器访问操作结束。

12. 一种用于如权利要求4-7中任一项所述的堆叠式Cache系统的Cache控制方法,其特征在于,所述Cache控制方法将集成于存储晶圆结构每一层的存储晶圆的SEDRAM单元,作为高速缓冲存储器的一部分存储空间进行控制,Cache数据域存储在所述SEDRAM单元中,控制域存储在集成于处理器晶圆结构的Cache SRAM中,具体步骤如下:

第一步骤:集成于所述处理器晶圆结构的CPU给出访存地址;

第二步骤:根据访存地址,集成于所述处理器晶圆结构的Cache控制器先访问所述Cache SRAM,从所述Cache SRAM中读出对应的控制域,所述控制域包括Cache状态域和Cache标签域;

第三步骤:对访存地址与Tag相对应的位与从所述Cache SRAM读出的Cache标签域进行地址比较;

第四步骤:如果访存地址对应于Tag的位与其中一路的Cache标签域完全相同、且Cache Line的状态位处于有效状态,则判断所述高速缓冲存储器访问命中,否则判断所述高速缓冲存储器访问未命中;

第五步骤:当所述高速缓冲存储器访问命中时,根据访存地址,通过集成于所述处理器晶圆结构的SEDRAM控制器从所述SEDRAM单元中读出访存地址对应的Cache Line中的Cache数据域的内容,作为本次访存操作的结果,直接执行第十步骤;

第六步骤:当所述高速缓冲存储器访问未命中时,如果存在下一级高速缓冲存储器,所述Cache控制器向下一级高速缓冲存储器发送访存操作请求,并等待所述访存操作请求所对应的数据返回;当所述访存操作请求所对应的数据返回所述Cache控制器后,执行第八步骤;

第七步骤:如果所述高速缓冲存储器访问未命中,且本级高速缓冲存储器是最后一级高速缓冲存储器时,所述Cache控制器直接通过内存控制器向内存发送访存操作请求,并等待访问内存的访存操作请求所对应的数据返回;当访问内存的访存操作请求所对应的数据返回所述Cache控制器后,执行第八步骤;

第八步骤:所述Cache控制器将返回的访存数据发送给CPU,同时通过高速缓冲存储替换算法找出一个Cache路中的Cache Line;被找出的Cache Line作为被替换的Cache Line,根据被替换的Cache Line是未被使用的Cache Line、被使用但非脏的Cache Line、或脏的Cache Line,以及Cache写策略进行写回块的处理;

第九步骤:所述Cache控制器在第八步骤中返回的访存数据返回后对第八步骤中找出的被替换Cache Line进行控制域和Cache数据域的填充,其中控制域内容根据第六、七步骤中发出的访存请求确定,通过所述Cache控制器填充到第八步骤中找出的被替换Cache Line所对应的Cache SRAM存储空间中;Cache数据域内容根据返回的访存数据确定,通过所述SEDRAM控制器向所述存储晶圆结构上与写地址对应的SEDRAM单元发出写操作,填充到被替换的Cache Line所对应的SEDRAM单元存储空间中,写地址由所述Cache控制器根据被替换的Cache Line给出;

第十步骤:本次高速缓冲存储器访问操作结束。

13. 一种基于SEDRAM的堆叠式Cache装置,其特征在于,包括权利要求1-10任一项所述的堆叠式Cache系统。

一种基于SEDRAM的堆叠式Cache系统、控制方法和Cache装置

技术领域

[0001] 本发明涉及计算机存储器技术领域,具体涉及一种基于SEDRAM的堆叠式Cache系统、控制方法和Cache装置。

背景技术

[0002] 现有技术中,高速缓冲存储器(Cache)是一种小容量的高速存储器,由快速SRAM(Static Random-Access Memory)存储元件组成,可以直接集成在CPU(Central Processing Unit,中央处理器)芯片内。在CPU和内存之间设置高速缓存Cache,把内存中被频繁访问的活跃程序块和数据块复制到Cache中,以提高CPU读写指令和数据的速度。由于程序访问的局部性,在大多数情况下,CPU能直接从Cache中取得指令和数据,而不必访问内存。

[0003] SRAM的集成度较低,相同容量的DRAM(Dynamic Random Access Memory,动态随机存取存储器)内存可以设计为较小的体积,但SRAM却需要很大的体积,而且价格较高,这也是不能将缓存容量做大的重要原因,容量增大必然导致CPU内部晶体管数量的增加,要在有限的CPU面积上集成更大的缓存,对制造工艺要求更高。例如IBM最新的Power9处理器,其缓存容量只有120MB;而Intel的服务器级芯片至强Platinum的缓存容量也只有38.5MB。

发明内容

[0004] 本发明提供了一种基于SEDRAM的堆叠式Cache系统、控制方法和Cache装置,该堆叠式Cache系统采用SEDRAM(Stacking Embedded Dynamic Random Access Memory,三维堆叠嵌入式动态随机存储器)单元作为高速缓冲存储器(Cache)的存储空间,处理器晶圆和集成SEDRAM单元的存储晶圆直接键合连接,大幅度增加Cache容量,可以将Cache容量提高到GB(Gigabyte, 2^{20} 字节)量级,与传统Cache容量相比至少可以提高一个数量级。

[0005] 第一方面,本发明提供一种基于SEDRAM的堆叠式Cache系统,所述堆叠式Cache系统集成于多层键合晶圆,并包括:高速缓冲存储器(Cache)、Cache控制器以及SEDRAM控制器;

[0006] 所述多层键合晶圆包括存储晶圆结构和处理器晶圆结构;

[0007] 所述存储晶圆结构包括至少一层存储晶圆;

[0008] 所述处理器晶圆结构包括至少一层处理器晶圆;

[0009] SEDRAM单元集成于所述存储晶圆结构的每一层存储晶圆中,并用作所述高速缓冲存储器的存储空间;

[0010] 所述处理器晶圆结构,集成有CPU、Cache控制器、SEDRAM控制器及内存控制器;

[0011] 晶圆之间通过键合结构进行管脚对管脚(Pin to Pin)的数据直接传输;

[0012] 所述晶圆之间通过键合结构进行管脚对管脚的数据直接传输,包括:

[0013] 所述存储晶圆结构中的存储晶圆与所述处理器晶圆结构中的处理器晶圆之间通过键合结构进行管脚对管脚的数据直接传输;

[0014] 所述存储晶圆结构中的存储晶圆之间通过键合结构进行管脚对管脚的数据直接传输；

[0015] 所述处理器晶圆结构中的处理器晶圆之间通过键合结构进行管脚对管脚的数据直接传输。

[0016] 结合上述第一方面,在第二种可能的实现方式中,所述Cache控制器用于判断所述高速缓冲存储器是否命中,向所述CPU、所述SEDRAM控制器以及所述内存控制器发出相应的命令操作及数据交互请求；

[0017] 所述SEDRAM控制器控制读出或写入所述SEDRAM单元的内容。

[0018] 结合上述第一方面的第二种可能的实现方式,在第三种可能的实现方式中,所述SEDRAM单元存储的内容,包括控制域(Cache Control)和Cache数据域(Cache Data),所述控制域包括Cache状态域(Cache State)、Cache标签域(Cache Tag)。

[0019] 结合上述第一方面的第三种可能的实现方式,在第四种可能的实现方式中,所述控制域还包括目录域(Directory)信息。

[0020] 结合上述第一方面,在第五种可能的实现方式中,所述处理器晶圆结构所述堆叠式Cache系统,还包括集成于所述处理器晶圆结构的有Cache SRAM,所述SEDRAM单元和Cache SRAM用作所述高速缓冲存储器的存储空间。

[0021] 结合上述第一方面的第五种可能的实现方式,在第六种可能的实现方式中,所述Cache控制器用于判断所述高速缓冲存储器是否命中,向所述CPU、所述SEDRAM控制器以及所述内存控制器发出相应的命令操作及数据交互请求；

[0022] 所述SEDRAM控制器控制读出或写入所述SEDRAM单元的内容；

[0023] 所述Cache控制器还用于读出或写入所述Cache SRAM中存储的内容。

[0024] 结合上述第一方面的第六种可能的实现方式,在第七种可能的实现方式中,所述SEDRAM单元存储的内容为Cache数据域；

[0025] 所述Cache SRAM存储的内容为控制域；

[0026] 所述控制域包括Cache状态域和Cache标签域。

[0027] 结合上述第一方面的第七种可能的实现方式,在第八种可能的实现方式中,所述控制域还包括有目录域信息。

[0028] 结合第一方面、第一种可能的实现方式、第二种可能的实现方式、第三种可能的实现方式、第四种可能的实现方式、第五种可能的实现方式、第六种可能的实现方式、第七种可能的实现方式、第八种可能的实现方式,在第九种可能的实现方式中,采用所述SEDRAM单元作为所述高速缓冲存储器其中任一级的高速缓冲存储器的存储空间。

[0029] 结合上述第一方面的第九种可能的实现方式,在第十种可能的实现方式中,所述处理器晶圆结构设置一层处理器晶圆,所述存储晶圆结构设置一层或两层存储晶圆。

[0030] 结合第一方面、第一种可能的实现方式、第二种可能的实现方式、第三种可能的实现方式、第四种可能的实现方式、第五种可能的实现方式、第六种可能的实现方式、第七种可能的实现方式、第八种可能的实现方式,在第十一种可能的实现方式中,所述处理器晶圆结构设置一层处理器晶圆,所述存储晶圆结构设置一层或两层存储晶圆。

[0031] 第二方面,提供一种用于上述技术方案提供的任意一种堆叠式Cache系统的Cache控制方法,所述Cache控制方法将集成于存储晶圆结构每一层的存储晶圆的SEDRAM单元,作

为高速缓冲存储器的整体存储空间进行控制,控制域和Cache数据域都存储在所述SEDRAM单元中,具体步骤如下:

[0032] 第一步骤:集成于处理器晶圆结构的CPU输出访存地址;

[0033] 第二步骤:根据访存地址,集成于所述处理器晶圆结构的Cache控制器发送读请求给集成于所述处理器晶圆结构的SEDRAM控制器,所述SEDRAM控制器从所述SEDRAM单元中读出所需要的Cache Line,所述Cache Line包括所述控制域和所述Cache数据域(Cache Data),所述控制域包括Cache状态域(Cache State)和Cache标签域(Cache Tag);

[0034] 第三步骤:对访存地址中对应于Tag的位与所述Cache Line中的所述Cache标签域进行地址比较;

[0035] 第四步骤:如果访存地址对应于Tag的位与其中一路的Cache 标签域完全相同、且所述Cache Line的状态位处于有效状态,则判断所述高速缓冲存储器访问命中(Cache Hit),否则判断所述高速缓冲存储器访问未命中(Cache Miss);

[0036] 第五步骤:当所述高速缓冲存储器访问命中后,所述Cache控制器直接把所述Cache Line中的所述Cache数据域内容读回到CPU中,作为本次访存操作的结果,并直接执行第十步骤;

[0037] 第六步骤:当所述高速缓冲存储器访问未命中时,如果存在下一级所述高速缓冲存储器,所述Cache控制器向下一级所述高速缓冲存储器发送访存操作请求,并等待所述访存操作请求所对应的数据返回;当所述访存操作请求所对应的数据返回所述Cache控制器后,执行第八步骤;

[0038] 第七步骤:如果所述高速缓冲存储器访问未命中,且本级高速缓冲存储器是最后一级高速缓冲存储器时,所述Cache控制器直接通过集成于所述处理器晶圆结构的内存控制器向内存发送访存操作请求,并等待该访问内存的访存操作请求所对应的数据返回;当该访问内存的访存操作请求所对应的数据返回所述Cache控制器后,执行第八步骤;

[0039] 第八步骤:所述Cache控制器将返回的访存数据发送给CPU,同时通过高速缓冲存储器替换算法找出一个Cache路中的Cache Line;被找出的Cache Line作为被替换(Replace)的Cache Line,根据被替换的Cache Line是未被使用的Cache Line、被使用但非脏(dirty)的Cache Line、或脏(dirty)的Cache Line,以及Cache写策略进行写回块的处理;

[0040] 第九步骤:所述Cache控制器将第八步骤中返回的访存数据填充到被替换的Cache Line所对应的SEDRAM单元存储空间中,由所述SEDRAM控制器向所述存储晶圆结构上与写地址对应的SEDRAM单元发出写操作,写操作的写地址由所述Cache控制器根据第八步骤中找出的被替换Cache Line给出,写数据是需要被写入的Cache Line完整内容,包括控制域和Cache数据域,其中控制域信息根据访问Cache Line的请求确定,Cache数据域内容由第八步骤中返回的访存数据确定;

[0041] 第十步骤:本次高速缓冲存储器访问操作结束。

[0042] 第三方面,提供一种用于上述技术方案提供的任意一种堆叠式Cache系统的Cache控制方法,所述Cache控制方法将集成于存储晶圆结构每一层的存储晶圆的SEDRAM单元作为高速缓冲存储器(Cache)的部分存储空间进行控制,Cache数据域(Cache Data)存储在所述SEDRAM单元中,控制域(Cache Control)存储在集成于处理器晶圆结构的Cache SRAM中,

具体步骤如下：

[0043] 第一步骤：集成于所述处理器晶圆结构的CPU给出访存地址；

[0044] 第二步骤：根据访存地址，集成于所述处理器晶圆结构的Cache控制器先访问所述Cache SRAM，从所述Cache SRAM中读出对应的控制域，所述控制域包括Cache状态域和Cache标签域；

[0045] 第三步骤：对访存地址与Tag相对应的位与从所述Cache SRAM读出的Cache标签域进行地址比较；

[0046] 第四步骤：如果访存地址对应于Tag的位与其中一路的Cache标签域完全相同、且Cache Line的状态位处于有效状态，则判断所述高速缓冲存储器访问命中(Cache Hit)，否则判断所述高速缓冲存储器访问未命中(Cache Miss)；

[0047] 第五步骤：当所述高速缓冲存储器访问命中时，根据访存地址，通过集成于所述处理器晶圆结构的SEDRAM控制器从所述SEDRAM单元中读出访存地址对应的Cache Line中的Cache数据域(Cache Data)的内容，作为本次访存操作的结果，直接执行第十步骤；

[0048] 第六步骤：当所述高速缓冲存储器访问未命中时，如果存在下一级高速缓冲存储器，所述Cache控制器向下一级高速缓冲存储器发送访存操作请求，并等待该访存操作请求所对应的数据返回；当该访存操作请求所对应的数据返回所述Cache控制器后，执行第八步骤；

[0049] 第七步骤：如果所述高速缓冲存储器访问未命中，且本级高速缓冲存储器是最后一级高速缓冲存储器时，所述Cache控制器直接通过集成于处理器晶圆的内存控制器向内存发送访存操作请求，并等待该访问内存的访存操作请求所对应的数据返回；当该访问内存的访存操作请求所对应的数据返回所述Cache控制器后，执行第八步骤；

[0050] 第八步骤：Cache控制器将返回的访存数据发送给CPU，同时通过高速缓冲存储器替换算法找出一个Cache路中的Cache Line；被找出的Cache Line作为被替换(Replace)的Cache Line，根据被替换的Cache Line是未被使用的Cache Line、被使用但非脏(dirty)的Cache Line、或脏(dirty)的Cache Line，以及Cache写策略进行写回块的处理；

[0051] 第九步骤：所述Cache控制器在第八步骤中返回的访存数据返回后对第八步骤中找出的被替换Cache Line进行控制域和Cache数据域的填充，其中控制域内容根据第六、七步骤中发出的访存请求确定，通过所述Cache控制器填充到第八步骤中找出的被替换Cache Line所对应的Cache SRAM存储空间中；Cache数据域内容根据返回的访存数据确定，通过所述SEDRAM控制器向所述存储晶圆结构上与写地址对应的SEDRAM单元发出写操作，填充到第八步骤中找出的被替换Cache Line所对应的SEDRAM单元存储空间中，写地址由所述Cache控制器根据被替换的Cache Line给出；

[0052] 第十步骤：本次高速缓冲存储器访问操作结束。

[0053] 第四方面，提供一种基于SEDRAM的堆叠式Cache装置，包括上述第一方面的十一种可能的实现方式中的任意一种堆叠式Cache系统。

[0054] 根据第一、二、三、四方面提供的基于SEDRAM的堆叠式Cache系统、Cache控制方法和Cache装置，采用SEDRAM单元作为该Cache系统的高速缓冲存储器(Cache)的存储空间，可以大幅度增加Cache容量，实现不同功能、不同工艺的晶圆之间的数据快速、短距离传输并有效突破“存储墙”的限制，打造真正的存算一体。

[0055] 根据第一、二、三、四方面提供的基于SEDRAM的堆叠式Cache系统、Cache控制方法和Cache装置,采用SEDRAM单元作为该Cache系统的Cache的整体存储空间或者其中一级Cache的整体存储空间,处理器晶圆和集成有SEDRAM单元的存储晶圆直接键合连接,大幅度增加Cache容量,可以将Cache的存储容量提高到GB(Gigabyte, 2^{20} 字节)量级,与传统Cache容量相比至少可以提高一个数量级;同时,还能提高Cache的命中率,加快热点数据的读取速度,提升CPU内部读取数据的命中率,综合提高系统性能,同时降低成本、功耗和延时;除此之外,还可以大幅度增加大容量存储器的数据传输带宽,节省处理器芯片的存储资源,实现在同样大小面积的芯片上集成更多的处理逻辑,如集成更多人工智能(Artificial Intelligence, AI)核或其它协处理器(Coprocessor)核,实现更多的功能和更强的计算能力。

[0056] 根据第一、二、三、四方面提供的基于SEDRAM的堆叠式Cache系统、Cache控制方法和Cache装置,采用SEDRAM单元作为该Cache系统的Cache的部分存储空间或者其中一级Cache的部分存储空间,片内Cache SRAM中可以存储更多的控制域,大幅度增加Cache容量,可以将Cache的存储容量提高到GB(Gigabyte, 2^{20} 字节)量级,与传统Cache容量相比至少可以提高一个数量级,还可提高Cache的命中率,加快热点数据的读取速度,提升CPU内部读取数据的命中率,综合提高系统性能,同时降低成本、功耗和延时,并大幅度增加大容量存储器的数据传输带宽。

附图说明

[0057] 为了更清楚地说明本申请实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本申请的某些实施例,因此不应被看作是对范围的限定,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他相关的附图。

[0058] 图1为本发明一种实施例提供的SEDRAM作为Cache整体存储空间的堆叠式Cache系统示意图;

[0059] 图2为本发明一种实施例提供的SEDRAM作为Cache部分存储空间的堆叠式Cache系统示意图;

[0060] 图3为本发明一种实施例提供的SEDRAM和CPU的堆叠式Cache系统结构示意图;

[0061] 图4为本发明另一种实施例提供的SEDRAM和CPU的堆叠式Cache系统结构示意图;

[0062] 图5为本发明一种实施例提供的SEDRAM作为Cache整体存储空间的Cache控制方法流程图;

[0063] 图6为本发明一种实施例提供的SEDRAM作为Cache部分存储空间的Cache控制方法流程图。

具体实施方式

[0064] 下面将结合本发明实施例中的附图,对本发明实施例中的技术方案进行清楚地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本发明保护的范围。

[0065] 在本发明实施例中,SEDRAM(Stacking Embedded Dynamic Random Access Memory,三维堆叠嵌入式动态随机存储器)是一种DRAM结构,其利用纳米级互连技术将多片晶圆或晶圆与芯片在垂直方向互相连接在一起,两片晶圆堆叠技术S-stacking或多片晶圆堆叠技术M-stacking可以实现两片或多片不同工艺、不同功能的晶圆直接金属互连,达到更高的互连密度及对准精度。S-stacking技术和M-stacking技术可直接将存储(Memory)晶圆和逻辑(Logic)晶圆堆叠在一起,并行高效处理架构,在不同功能的晶圆之间实现数据快速传输,实现更短的连接距离,提升连接速度和通道数目。

[0066] 第63届国际电子器件大会(IEDM 2020)上发表的题为《采用3D混合键合技术具有34GB/s/1Gb带宽和0.88pJ/b能效接口的异质集成嵌入式LPDDR4/LPDDR4X DRAM》(A Stacked Embedded DRAM Array for LPDDR4/4X using Hybrid Bonding 3D Integration with 34GB/s/1Gb 0.88pJ/b Logic-to-Memory Interface)的技术论文中公开了SEDRAM。

[0067] 高速缓冲存储器(Cache)通常分为多级Cache,每一级Cache内部又可以分成多个路(way),每个路有多个Cache Line(Cache行),Cache Line作为一个基本存储块由控制域(简称Cache Control)和Cache数据域(简称Cache Data)两个部分构成;控制域再划分为Cache状态域(Cache State)和Cache标签域(Cache Tag),多核处理器系统还包括目录域(简称Directory)信息等。

[0068] Cache状态域代表本Cache Line的状态,包括标记这个Cache Line是否保存着有效的数据(有效位,valid位),只有在之前被访问过的存储地址,它的数据才会存在于对应的Cache Line中,相应的有效位也会被置为1;是否有被写过的最新数据(脏位,dirty位),即数据是否与下一级存储系统(下一级Cache或内存)保持一致等等。

[0069] Cache标签域存储Cache Line数据对应的地址,用于判断本Cache Line是否是当前访问地址所对应的内容。

[0070] 目录域(Directory)记录多核一致性相关信息,记录多核处理器中哪个处理器核拥有该Cache Line的数据备份,数据是独占使用(Exclusive)还是共享使用(Shared)等信息。

[0071] Cache数据域存放访存数据,供CPU使用。

[0072] Cache Line中的控制域除Cache状态域(Cache State)、Cache标签域(Cache Tag)、目录域(Directory)外,还可包含年龄位、指令标志位和预取算法的标志位等,年龄位是指Cache替换算法用到的Cache Line的年龄,指令标志位用来标志Cache Line存储的是指令还是数据,预取算法的标志位用来标志是否是预取回来的数据等。

[0073] 本发明实施例提供了一种基于SEDRAM的堆叠式Cache系统、控制方法和Cache装置,该堆叠式Cache系统集成于多层键合晶圆,并包括:高速缓冲存储器(Cache)、Cache控制器以及SEDRAM控制器,该堆叠式Cache系统采用SEDRAM单元作为Cache的存储空间,处理器晶圆和集成SEDRAM单元的存储晶圆直接键合连接,大幅度增加Cache容量,可以将Cache容量提高到GB(Gigabyte, 2^{20} 字节)量级,与传统Cache容量相比至少可以提高一个数量级。

[0074] 实施例一

[0075] 参考图1、图3,本实施例提供了一种基于SEDRAM的堆叠式Cache系统,堆叠式Cache系统集成于多层键合晶圆,并包括高速缓冲存储器(Cache)、Cache控制器以及SEDRAM控制器;SEDRAM控制器可以为单通道也可以为多通道并行访问;根据SEDRAM的分体(bank)数量,

SEDRAM控制器可以是一个也可以是多个；

[0076] 多层键合晶圆包括存储晶圆结构和处理器晶圆结构,存储晶圆结构包括至少一层存储晶圆,处理器晶圆结构包括至少一层处理器晶圆;即:多层键合晶圆可以包括一层处理器晶圆和一层存储晶圆,多层键合晶圆也可以包括一层处理器晶圆和两层存储晶圆,多层键合晶圆还可以包括两层处理器晶圆和一层存储晶圆,处理器晶圆结构中的处理器晶圆的具体层数和存储晶圆结构中的存储晶圆的具体层数,可以根据实际情况进行确定;如图3所示,该Cache系统设置有一层存储晶圆和一层处理器晶圆,即存储晶圆0和处理器晶圆0;

[0077] SEDRAM单元集成于存储晶圆结构的每一层存储晶圆中,并用作该Cache系统中的Cache的整体存储空间;也可以采用SEDRAM单元作为其中一级Cache的整体存储空间;如图3所示,SEDRAM单元集成在存储晶圆0上,SEDRAM单元作为该Cache系统中的Cache的整体存储空间,存储Cache Line的Cache数据域和控制域,并且处理器晶圆0集成有CPU、Cache控制器、SEDRAM控制器及内存控制器;晶圆之间通过键合结构进行管脚对管脚(Pin to Pin)的数据直接传输,这里的晶圆之间是指:存储晶圆和处理器晶圆之间、存储晶圆和存储晶圆之间以及处理器晶圆和处理器晶圆之间;

[0078] 多层键合晶圆上集成的全部逻辑模块构成完整的处理器系统;内存一般集成在芯片外的主板上;存储晶圆和处理器晶圆可以先进行整晶圆键合,再进行划片,并完成封装;

[0079] Cache控制器是该堆叠式Cache系统的控制核心,Cache控制器用于判断该Cache系统中的Cache是否命中,向CPU、SEDRAM控制器以及内存控制器发出相应的命令操作及数据交互请求;SEDRAM控制器控制读出或写入SEDRAM单元的内容。

[0080] 更进一步地,SEDRAM单元存储的内容包括控制域(Cache Control)和Cache数据域(Cache Data);控制域包括Cache状态域(Cache State)和Cache标签域(Cache Tag),控制域还可以包括目录域(Directory)信息。

[0081] 采用SEDRAM单元作为该Cache系统中的Cache的整体存储空间或者其中一级Cache的整体存储空间,处理器晶圆和集成SEDRAM单元的存储晶圆直接键合连接,大幅度增加Cache容量,可以将Cache的存储容量提高到GB(Gigabyte, 2^{20} 字节)量级,与传统Cache容量相比至少可以提高一个数量级;同时,还能提高Cache的命中率,加快热点数据的读取速度,提升CPU内部读取数据的命中率,综合提高系统性能,同时降低成本、功耗和延时;除此之外,还可以大幅度增加大容量存储器的数据传输带宽,节省处理器芯片的存储资源,实现在同样大小面积的芯片上集成更多的处理逻辑,如集成更多人工智能(Artificial Intelligence, AI)核或其它协处理器(Coprocessor)核,实现更多的功能和更强的计算能力。

[0082] 实施例二

[0083] 参考图2、图4,本实施例提供了一种基于SEDRAM的堆叠式Cache系统,堆叠式Cache系统集成于多层键合晶圆,并包括高速缓冲存储器(Cache)、Cache控制器以及SEDRAM控制器;SEDRAM控制器可以为单通道也可以为多通道并行访问;根据SEDRAM的分体(bank)数量,SEDRAM控制器可以是一个也可以是多个;

[0084] 多层键合晶圆包括存储晶圆结构和处理器晶圆结构,存储晶圆结构包括至少一层存储晶圆,处理器晶圆结构包括至少一层处理器晶圆;即:多层键合晶圆可以包括一层处理器晶圆和一层存储晶圆,多层键合晶圆也可以包括一层处理器晶圆和两层存储晶圆,多层

键合晶圆还可以包括两层处理器晶圆和一层存储晶圆,处理器晶圆结构中的处理器晶圆的层数和存储晶圆结构中的存储晶圆的层数,可以根据实际情况进行确定;如图4所示,该Cache系统设置有两层存储晶圆和一层处理器晶圆,即存储晶圆0、存储晶圆1以及处理器晶圆0;

[0085] SEDRAM单元集成于存储晶圆结构的每一层存储晶圆中,并用作该Cache系统中的Cache的部分存储空间;也可以采用SEDRAM单元作为其中一级Cache的部分存储空间;如图4所示,SEDRAM单元集成在存储晶圆0和存储晶圆1上,处理器晶圆0上集成有CPU、SEDRAM控制器、Cache控制器、Cache SRAM及内存控制器,SEDRAM单元作为该Cache系统中的Cache的部分存储空间存储Cache Line中的Cache数据域,Cache Line的控制域信息存储在Cache SRAM中;晶圆之间通过键合结构进行管脚对管脚(Pin to Pin)的数据直接传输,这里的晶圆之间是指:存储晶圆和处理器晶圆之间、存储晶圆和存储晶圆之间以及处理器晶圆和处理器晶圆之间;

[0086] 多层键合晶圆上集成的全部逻辑模块构成完整的处理器系统;内存一般集成在芯片外的主板上;存储晶圆和处理器晶圆可以先进行整晶圆键合,再进行划片,并完成封装;

[0087] 更进一步地,Cache控制器是该堆叠式Cache系统的控制核心,Cache控制器用于判断该Cache系统中的Cache是否命中,向CPU、SEDRAM控制器以及内存控制器发出相应的命令操作及数据交互请求;SEDRAM控制器控制读出或写入SEDRAM单元的内容;Cache控制器还用于读出或写入Cache SRAM中存储的内容。

[0088] 具体地,SEDRAM单元作为该Cache系统中的Cache的部分存储空间,SEDRAM单元存储的内容为Cache数据域;Cache SRAM存储的内容为控制域;控制域包括Cache状态域和Cache标签域,控制域还可以包括目录域信息。

[0089] 采用SEDRAM单元作为该Cache系统中的Cache的部分存储空间或者其中一级Cache的部分存储空间,Cache SRAM中可以存储更多的控制域,大幅度增加Cache容量,可以将Cache的存储容量提高到GB(Gigabyte, 2^{20} 字节)量级,与传统Cache容量相比至少可以提高一个数量级,还可提高Cache的命中率,加快热点数据的读取速度,提升CPU内部读取数据的命中率,综合提高系统性能,同时降低成本、功耗和延时,并大幅度增加大容量存储器的数据传输带宽。

[0090] 基于SEDRAM的堆叠式Cache系统的组织方式可以是组相联(set-associative)、直接相联(direct-mapped)和全相联(fully-associative)等。堆叠式Cache系统既可以作为处理器系统的其中一级Cache(其余各级Cache通常实现在CPU内部),也可以作为处理器系统的唯一一级Cache存在。

[0091] 在上述堆叠式Cache系统中,SEDRAM单元并不集成在处理器芯片内部,而是集成在存储晶圆结构的每一层存储晶圆上;采用SEDRAM单元作为该Cache系统中的Cache的存储空间,可以提高访问带宽,实现不同功能、不同工艺的晶圆之间的数据快速、短距离传输并有效突破“存储墙”的限制,打造真正的存算一体;除此之外,还可以极大地提高Cache容量,提高Cache的命中率,加快热点数据的读取速度,提升CPU内部读取数据的命中率,综合提高系统性能,同时降低成本、功耗和延时。

[0092] 实施例三

[0093] 参考图5,本实施例提供了一种用于上述实施例中任意一种堆叠式Cache系统的

Cache控制方法,将集成于存储晶圆结构的每一层存储晶圆中的SEDRAM单元作为该Cache系统中的Cache的整体存储空间进行控制,控制域和Cache数据域都存储在SEDRAM单元中;

[0094] 该Cache控制方法的具体步骤如下:

[0095] 第一步骤S10:集成于处理器晶圆结构的CPU输出访存地址;

[0096] 第二步骤S11:根据访存地址,Cache控制器发送读请求给集成于处理器晶圆结构的SEDRAM控制器,SEDRAM控制器从SEDRAM单元中读出所需要的Cache Line,Cache Line包括控制域和Cache数据域(Cache Data),控制域包括Cache状态域(Cache State)和Cache标签域(Cache Tag);

[0097] 第三步骤S12:对访存地址中对应于Tag的位与Cache Line中的Cache Tag进行地址比较;

[0098] 第四步骤S13:如果访存地址对应于Tag的位与其中一路的Cache Tag完全相同、且Cache Line的状态位处于有效状态,则判断Cache访问命中(Cache Hit),否则判断Cache访问未命中(Cache Miss);

[0099] 第五步骤S14:当Cache访问命中后,直接把Cache Line中的Cache Data内容读回到CPU中,作为本次访存操作的结果,并直接执行第十步骤S19;

[0100] 第六步骤S15:当Cache访问未命中时,如果存在下一级Cache,Cache控制器向下一级Cache发送访存操作请求,并等待该访存操作请求所对应的数据返回;当该访存操作请求所对应的数据返回Cache控制器后,执行第八步骤S17;

[0101] 第七步骤S16:如果Cache访问未命中,且本级Cache是最后一级Cache时,Cache控制器直接通过集成于处理器晶圆结构的内存控制器向内存发送访存操作请求,并等待该访问内存的访存操作请求所对应的数据返回;当该访问内存的访存操作请求所对应的数据返回Cache控制器后,执行第八步骤S17;

[0102] 第八步骤S17:Cache控制器将返回的访存数据发送给CPU,同时通过Cache替换算法找出一个Cache路中的Cache Line;被找出的Cache Line作为被替换(Replace)的Cache Line,根据被替换的Cache Line是未被使用的Cache Line、被使用但非脏(dirty)的Cache Line、或脏(dirty)的Cache Line,以及Cache写策略进行写回块的处理;Cache替换算法有随机替换(Random)、最近最少使用(LRU)、最不经常使用(LFU)和先进先出(FIFO)等;Cache写策略有写回(Write Back)或写穿透(Write Through)等;

[0103] 第九步骤S18:所述Cache控制器将第八步骤S17中返回的访存数据填充(Refill)到被替换的Cache Line所对应的SEDRAM单元存储空间中,由所述SEDRAM控制器向所述存储晶圆结构上与写地址对应的SEDRAM单元发出写操作,写操作的写地址由所述Cache控制器根据第八步骤S17中找出的被替换Cache Line给出,写数据是需要被写入的Cache Line完整内容,包括控制域和Cache数据域,其中控制域信息根据访问Cache Line的请求确定,Cache数据域内容由第八步骤S17中返回的访存数据确定;

[0104] 第十步骤S19:本次Cache访问操作结束。

[0105] 下面以单核处理器为例进行说明,因为处理器晶圆和集成SEDRAM单元的存储晶圆采用堆叠式结构,可以提供大容量高带宽,单核处理器的堆叠式Cache为4路组相联1GB Cache,每路的容量为 2^{18} 字节,每个Cache Line对应1024位宽的Cache数据域,即每个Cache Line对应 2^7 字节数据,每路 2^{11} 个Cache Line;Cache替换策略为随机(Random)替换,Cache替

换块写策略为写回策略。Cache状态位为2位,1位为valid位(即,有效位),即Cache数据是否有效,1位为dirty位(即,脏位),表示Cache中的数据是否是被写过,即是否与下一级存储系统一致;CPU的访存地址为40位。

[0106] 对于上述单核处理器,采用上述方法时可以包括以下具体步骤:

[0107] 第一步骤S10,CPU给出访存物理地址为40位1fc0000000(采用16进制表示),采用二进制表示为:0001 1111 1100 0000 0000 0000 0000 0000 0000 0000;

[0108] 第二步骤S11,根据访存地址,Cache控制器发送请求访问地址1fc0000000给SEDRAM控制器,SEDRAM控制器从SEDRAM单元中读出所需要的4路Cache Line;访存地址低7位([6:0])(0000000)为Cache Line的行内地址,称为Offset,即找到Cache Line内对应的字节的地址;访存地址接下来的11位([17:7])(00000000000)为Cache Line地址,称为Index,即定位具体的Cache Line的地址,即从四路的第0行读取Cache Line;

[0109] 第三步骤S12:访存地址对应于Tag的位为地址位的高22位([39:18])(0001 1111 1100 0000 0000 00),与从SEDRAM单元读出来的Cache Line中的Cache Tag域进行地址比较;

[0110] 第四步骤S13:地址比较结果为访存地址对应于Tag的位([39:18])(0001 1111 1100 0000 0000 00)与其中一路(第01路)的Tag一致,且该路对应的Cache Line的状态位处于有效状态,即valid为1,则认为Cache访问命中(Cache Hit);

[0111] 第五步骤S14:当Cache访问命中后,直接把Cache Line中的Cache Data内容读回到CPU中,作为本次访存操作的结果,直接执行第十步骤S19;

[0112] 第十步骤S19:CPU为单核处理器,本次Cache访问操作结束。

[0113] 采用SEDRAM单元作为该Cache系统中的Cache的整体存储空间或者该Cache系统中的Cache的其中一级Cache的整体存储空间,处理器晶圆和集成SEDRAM单元的存储晶圆直接键合连接,大幅度增加Cache容量,可以将Cache的存储容量提高到GB(Gigabyte, 2^{20} 字节)量级,与传统Cache容量相比至少可以提高一个数量级;同时,还能提高Cache的命中率,加快热点数据的读取速度,提升CPU内部读取数据的命中率,综合提高系统性能,同时降低成本、功耗和延时;除此之外,还可以大幅度增加大容量存储器的数据传输带宽,节省处理器芯片的存储资源,实现在同样大小面积的芯片上集成更多的处理逻辑,如集成更多人工智能(Artificial Intelligence, AI)核或其它协处理器(Coprocessor)核,实现更多的功能和更强的计算能力。

[0114] 实施例四

[0115] 参考图6,本实施例提供了一种用于上述实施例中任意一种堆叠式Cache系统的Cache控制方法,Cache控制方法将集成于存储晶圆结构的每一层存储晶圆中的SEDRAM单元作为该Cache系统中的Cache的部分存储空间进行控制,Cache数据域(Cache Data)存储在SEDRAM单元中,控制域(Cache Control)存储在与CPU一起集成于处理器晶圆结构的Cache SRAM中;

[0116] 该控制方法具体步骤如下:

[0117] 第一步骤S20:集成于处理器晶圆结构的CPU给出访存地址;

[0118] 第二步骤S21:根据访存地址,集成于处理器晶圆结构的Cache控制器先访问Cache SRAM,从Cache SRAM中读出对应的控制域,控制域包括Cache状态域和Cache标签域;

[0119] 第三步骤S22:对访存地址与Tag相对应的位与从Cache SRAM读出的Cache Tag进行地址比较;

[0120] 第四步骤S23:如果访存地址对应于Tag的位与其中一路的Cache Tag完全相同、且Cache Line的状态位处于有效状态,则判断Cache访问命中(Cache Hit),否则判断Cache访问未命中(Cache Miss);

[0121] 第五步骤S24:当Cache访问命中时,根据访存地址,通过集成于处理器晶圆结构的SEDRAM控制器从SEDRAM单元中读出访存地址对应的Cache Line中的Cache数据域(Cache Data)的内容,作为本次访存操作的结果,直接执行第十步骤S29;

[0122] 第六步骤S25:当Cache访问未命中时,如果存在下一级Cache,Cache控制器向下一级Cache发送访存操作请求,并等待该访存操作请求所对应的数据返回;当该访存操作请求所对应的数据返回Cache控制器后,执行第八步骤S27;

[0123] 第七步骤S26:如果Cache访问未命中,且本级Cache是最后一级Cache时,Cache控制器直接通过内存控制器向内存发送访存操作请求,并等待该访问内存的访存操作请求所对应的数据返回;当该访问内存的访存操作请求所对应的数据返回Cache控制器后,执行第八步骤S27;

[0124] 第八步骤S27:Cache控制器将返回的访存数据发送给CPU,同时通过Cache替换算法找出一个Cache路中的Cache Line;被找出的Cache Line作为被替换(Replace)的Cache Line,根据被替换的Cache Line是未被使用的Cache Line、被使用但非脏(dirty)的Cache Line、或脏(dirty)的Cache Line,以及Cache写策略进行写回块的处理;

[0125] 第九步骤S28:所述Cache控制器在第八步骤S27中返回的访存数据返回后对第八步骤S27中找出的被替换Cache Line进行控制域和Cache数据域的填充,其中控制域内容根据第六步骤S25或第七步骤S26中发出的访存请求确定,通过所述Cache控制器填充(Refill)到第八步骤S27中找出的被替换Cache Line所对应的Cache SRAM存储空间中;Cache数据域内容根据返回的访存数据确定,通过所述SEDRAM控制器向所述存储晶圆结构上与写地址对应的SEDRAM单元发出写操作,填充(Refill)到第八步骤S27中找出的被替换Cache Line所对应的SEDRAM单元存储空间中,写地址由所述Cache控制器根据被替换的Cache Line给出;

[0126] 第十步骤S29:本次Cache访问操作结束。

[0127] 下面以单核处理器为例进行说明,因为处理器晶圆和集成SEDRAM单元的存储晶圆采用堆叠式结构,可以提供大容量高带宽,单核处理器的堆叠式Cache为8路组相联2GB Cache,每路的容量为 2^{18} 字节,每个Cache Line对应512位宽的数据域,即每个Cache Line包括 2^6 字节,每路 2^{12} 个Cache Line;Cache替换策略为最近最少使用(LRU)替换,Cache替换块写策略为写回策略。Cache状态位为2位,1位为valid位(即,有效位),即Cache数据是否有效,1位为dirty位(即,脏位),表示Cache中的数据是否是被写过;CPU的访存物理地址为48位。

[0128] 对于上述单核处理器,采用上述控制方法时Cache访问步骤可以包括以下内容:

[0129] 第一步骤S20:CPU给出访存地址,CPU给出访存物理地址为48位8e000580000(采用十六进制表示),访存物理地址采用二进制表示为1000 1110 0000 0000 0000 0101 1000 0000 0000 0000 0000 0000;

[0130] 第二步骤S21:根据访存地址,Cache控制器先访问存放控制域(包括Cache状态域(Cache State)和Cache标签域(Cache Tag))等的Cache SRAM,从中读出对应8路的Cache状态域和Cache标签域;访存地址低6位([5:0]) (000000)为Cache Line的行内地址,称为Offset,即找到Cache Line内对应的字节的地址;访存地址接下来的12位([17:6]) (000000000000)为Cache Line地址,称为Index,即定位具体的Cache Line的地址,从8路Cache SRAM每一路的第0行读出控制域信息;

[0131] 第三步骤S22:访存地址对应于Tag的位为地址位的高30位([47:18]) (1000111000000000000000101100000),与从SEDRAM单元读出来的8路Cache Line中的Cache Tag域进行地址比较;

[0132] 第四步骤S23:通过Tag比较发现从Cache SRAM读出的8路Tag与访存地址对应的Tag位(1000111000000000000000101100000)都不一致,则认为Cache访问未命中(Cache Miss),执行第六步骤S25;

[0133] 第六步骤S25:Cache访问未命中,而且不存在下一级Cache;

[0134] 第七步骤S26:如果Cache访问未命中,且本级Cache是最后一级Cache时,Cache控制器直接通过内存控制器向内存发送访存操作请求,访存地址为8e0005800000,并等待该访问内存的访存操作请求所对应的数据返回;访问内存的访存操作请求所对应的数据返回Cache控制器后,执行第八步骤S27;

[0135] 第八步骤S27:Cache控制器将返回的数据发送给CPU内部,同时通过Cache替换算法LRU算法从8路中找出一个最近最少使用Cache路中的Cache Line,被找出的Cache Line作为被替换(Replace)的Cache Line,该被替换的Cache Line的dirty位为1,表示是脏的Cache Line,写策略为写回,进行写回块的处理,采用十六进制表示的地址为1f0005800000,将SEDRAM单元中该Cache Line的脏数据写回内存对应的地址;

[0136] 第九步骤S28:内存访问的地址为8e0005800000的访存数据返回后,Cache控制器根据第八步骤S27中找出的被替换(Replace)的Cache Line向Cache SRAM中填充控制域信息,同时发送请求给SEDRAM控制器进行Cache数据域的填充,其中Cache SRAM内被替换的Cache Line所对应的存储空间中的Cache状态域(Cache State)的valid位置为1和dirty位置为0,Cache标签域(Cache Tag)填充为访存地址的[47:18]位的内容(1000111000000000000000101100000),Cache数据域(Cache Data)的内容则通过SEDRAM控制器填充(Refill)到第八步骤S27中找出的被替换Cache Line所对应的SEDRAM单元存储空间中;

[0137] 第十步骤S29:作为单核处理器,本次Cache访问操作结束。

[0138] 采用SEDRAM单元作为该Cache系统中的Cache的部分存储空间或者其中一级Cache的部分存储空间,Cache SRAM中可以存储更多的控制域,大幅度增加Cache容量,可以将Cache的存储容量提高到GB(Gigabyte, 2^{20} 字节)量级,与传统Cache容量相比至少可以提高一个数量级,还可提高Cache的命中率,加快热点数据的读取速度,提升CPU内部读取数据的命中率,综合提高系统性能,同时降低成本、功耗和延时,并大幅度增加大容量存储器的数据传输带宽。

[0139] 实施例五

[0140] 本实施例提供了一种堆叠式Cache装置,包括上述实施例提供的任意一种堆叠式

Cache系统。

[0141] 需要说明的是：本发明实施例中提到的Cache容量并非本发明能达到的最高Cache容量，Cache容量大小和处理器芯片的面积、集成SEDRAM单元的密度、SEDRAM存储晶圆堆叠层数、SEDRAM的制作工艺密切相关，处理器芯片的面积越大、存储晶圆密度越大、存储晶圆堆叠层数越多且能堆叠的SEDRAM制作工艺越先进，Cache容量也就越大，Cache容量可以达到4GB、8GB甚至更高。

[0142] 本发明采用SEDRAM作为Cache的存储空间。这里提到的Cache可以是处理器系统内所有的Cache作为一个整体设置在存储晶圆上，也可以作为其中一级Cache设置在存储晶圆上，处理器系统中的其它级Cache可以集成在CPU内，也可以设置在CPU外围，其它级Cache可以是SRAM，也可以是DRAM、HBM(High Bandwidth Memory)等，这些情况均不影响本发明的保护范围。

[0143] 需要说明的是，在本文中，诸如“第一”和“第二”等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来，而不一定要求或者暗示这些实体或操作之间存在任何这种实际的关系或者顺序。而且，术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含，从而使得包括一系列要素的过程、方法、物品或者设备不仅包括那些要素，而且还包括没有明确列出的其他要素，或者是还包括为这种过程、方法、物品或者设备所固有的要素。在没有更多限制的情况下，由语句“包括一个……”限定的要素，并不排除在包括所述要素的过程、方法、物品或者设备中还存在另外的相同要素。

[0144] 以上所述仅为本申请的优选实施例而已，并不用于限制本申请，对于本领域的技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。应注意到：相似的标号和字母在下面的附图中表示类似项，因此，一旦某一项在一个附图中被定义，则在随后的附图中不需要对其进行进一步定义和解释。以上所述仅为本申请的优选实施例而已，并不用于限制本申请，对于本领域的技术人员来说，本申请可以有各种更改和变化。凡在本申请的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本申请的保护范围之内。

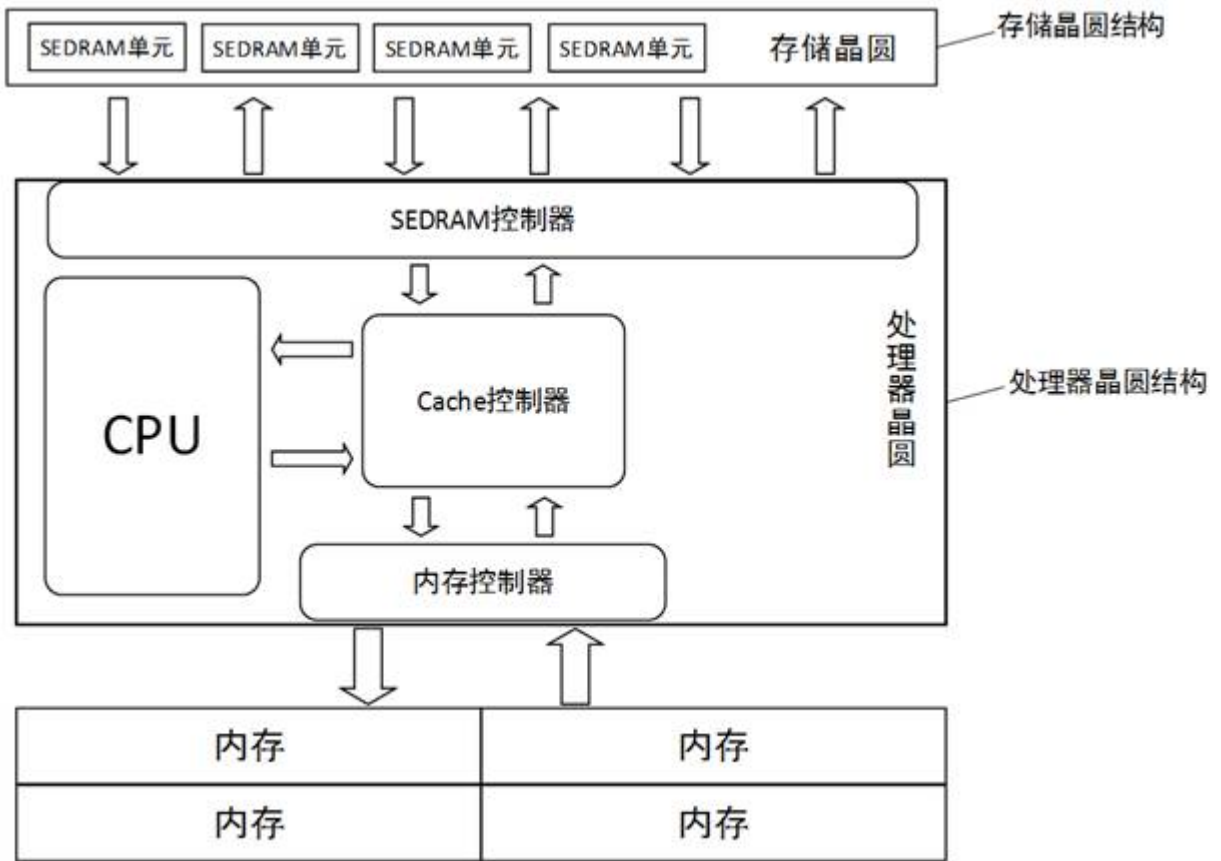


图1

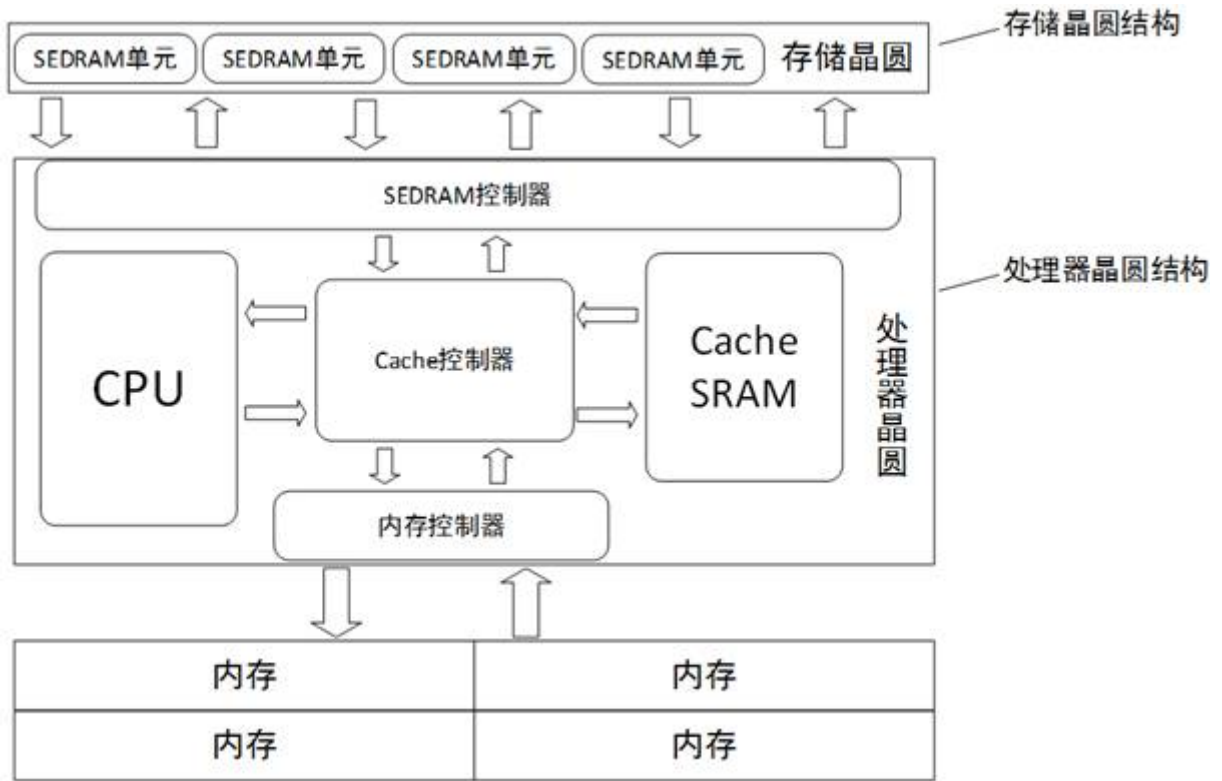


图2

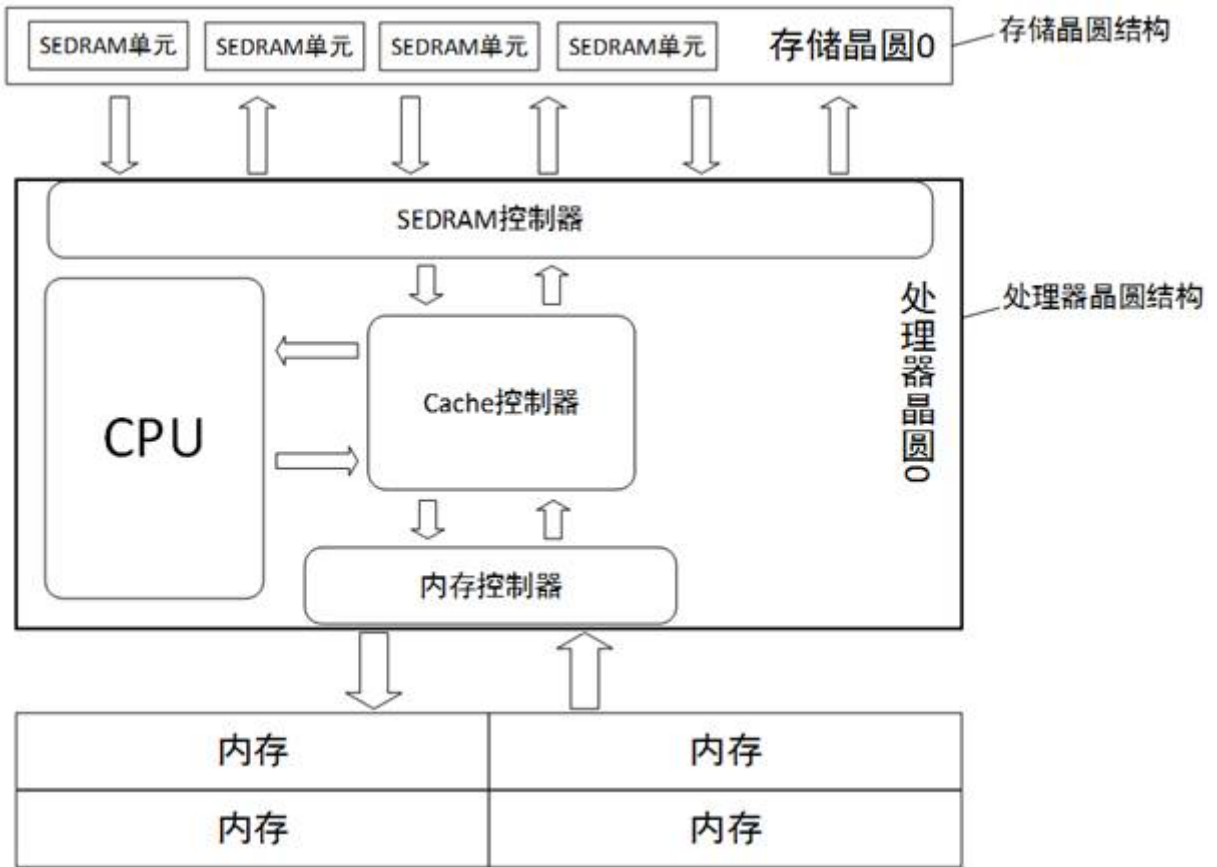


图3

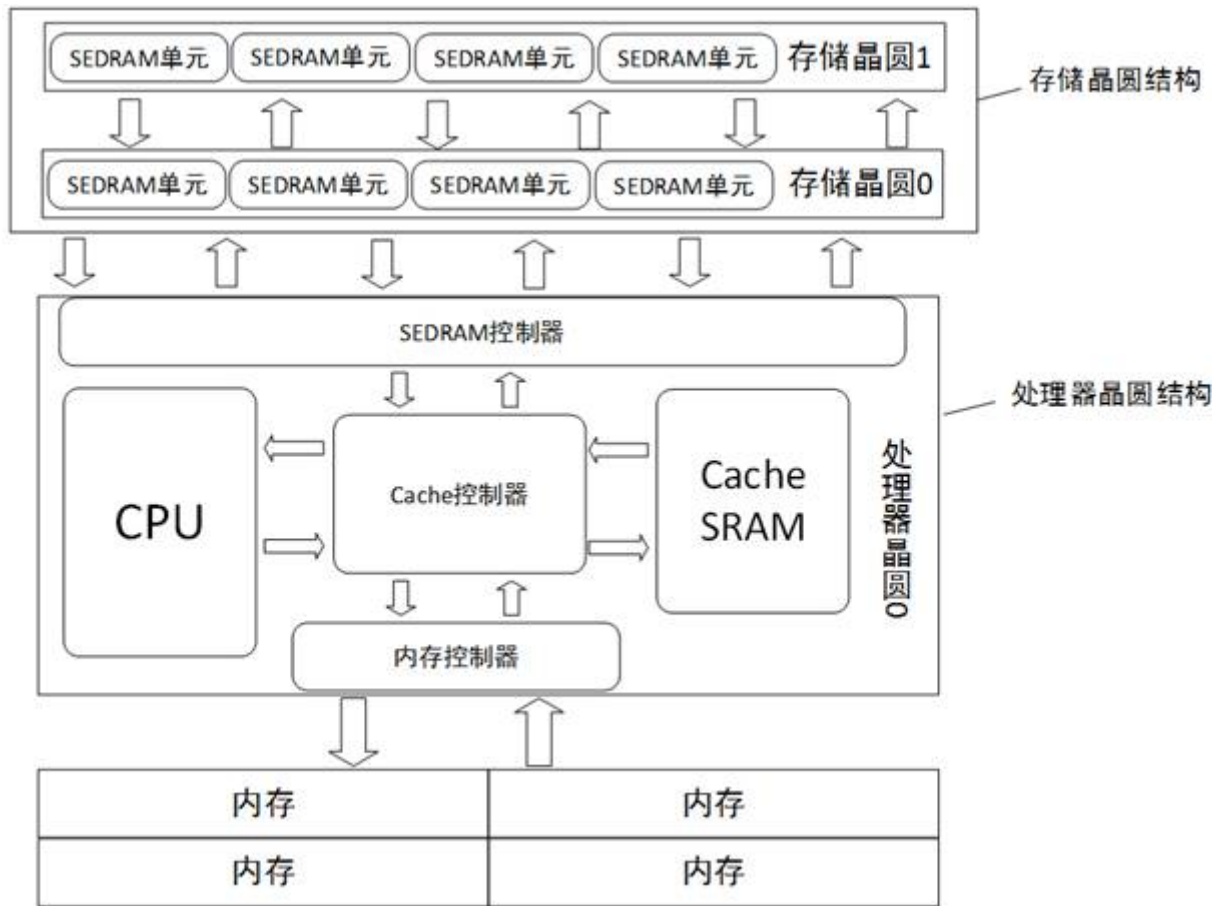


图4

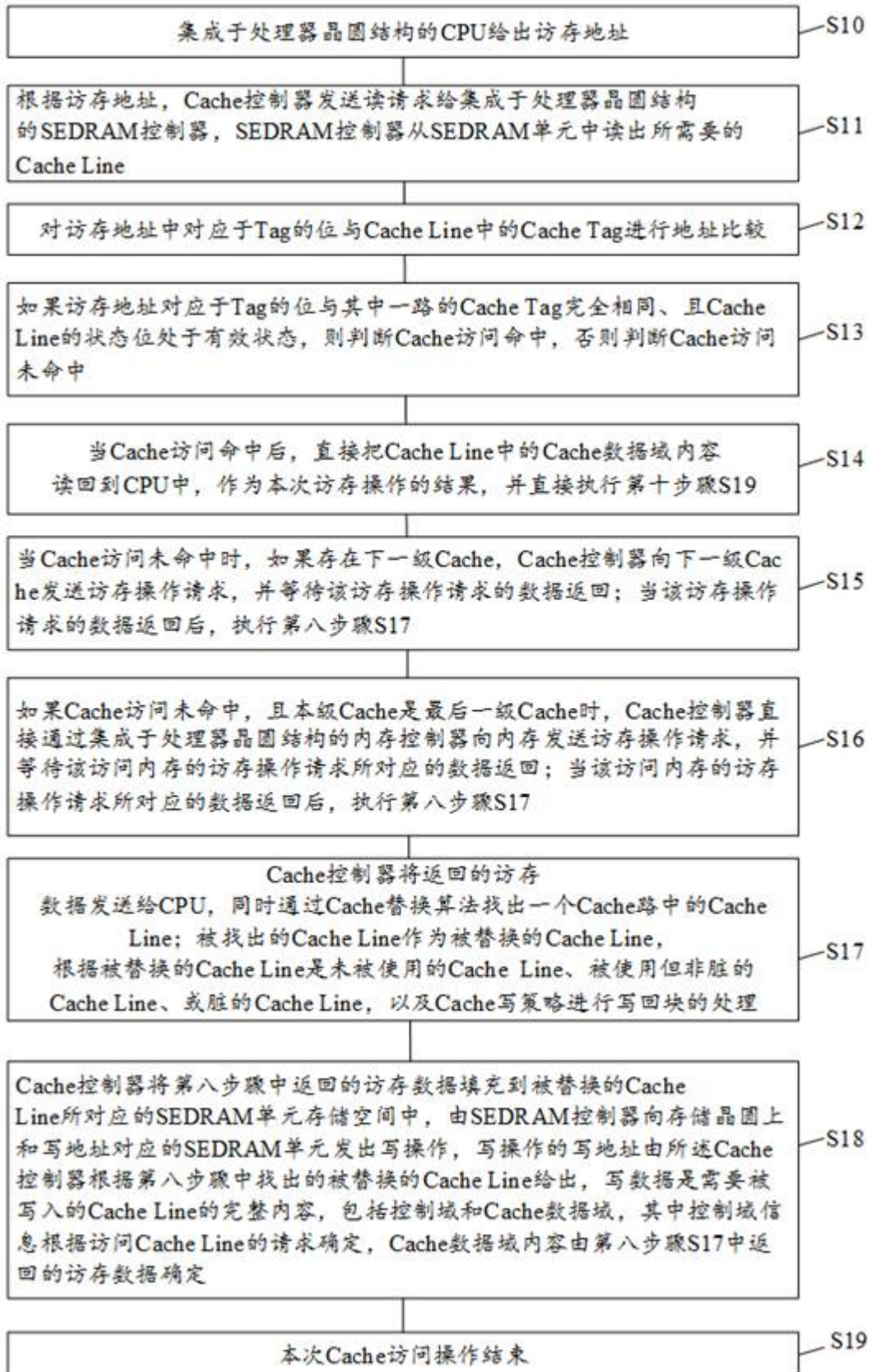


图5

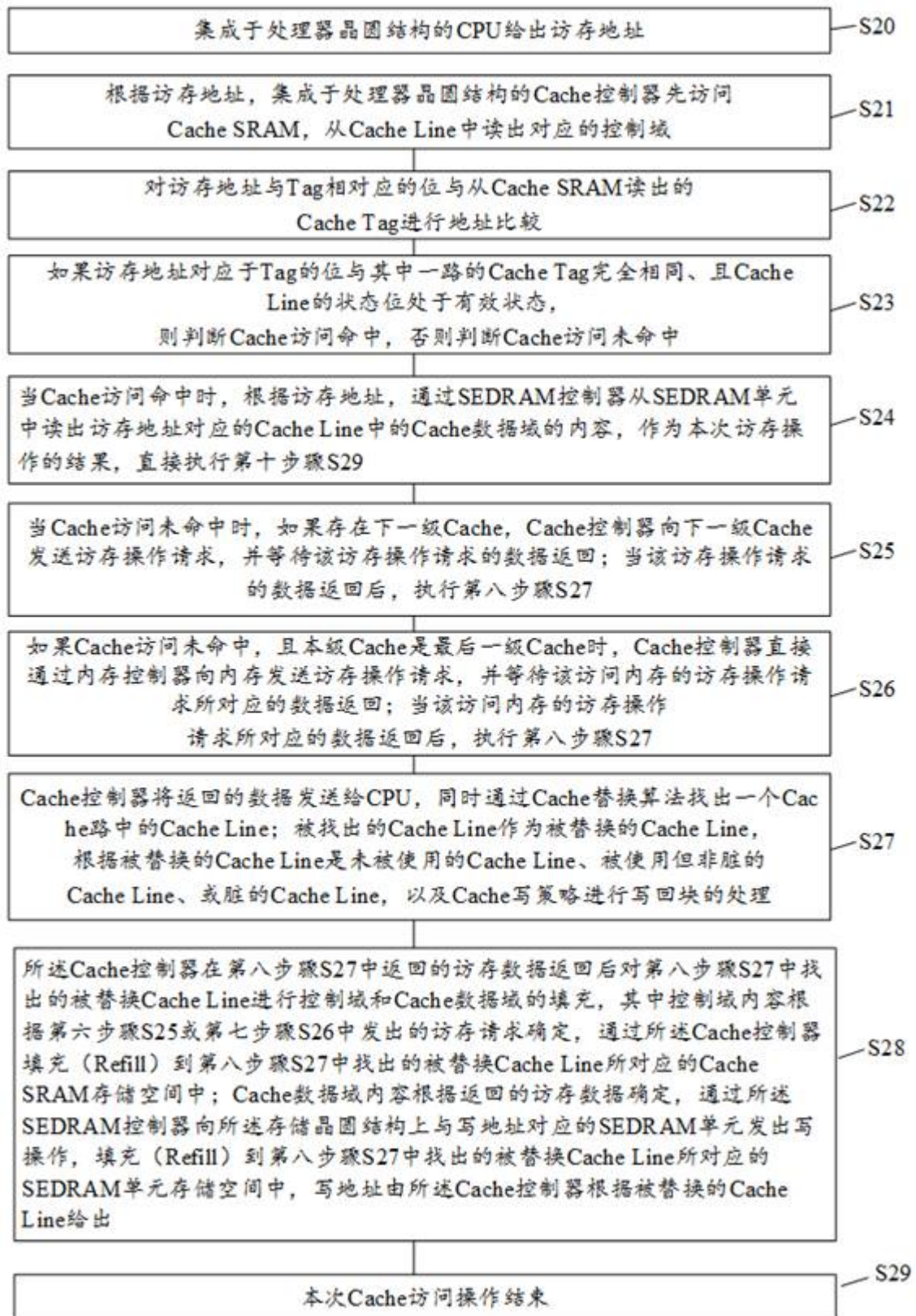


图6