

NAND Flash 固态存储卡的设计与实现

尹 蕾

(中国工程物理研究院电子工程研究所, 四川 绵阳 621900)

摘要:描述了一种基于 CPCI 的高速固态存储板卡的设计和实现. 该存储卡用于某项目中频采样存储设备上, 提供中频信号的采集存储、回放及分析等功能. 该存储板卡采用 NAND Flash 作为固态存储体, 利用 FPGA 作为固态盘的存储控制器, 同时采用文件的方式对数据进行管理. 本设计采用了一种建立有效块表的方式对坏块进行管理, 并通过 DDR2 作为坏块数据备份, 进一步提高了存储时的写入速度. 该存储卡数据写入速度理论上可达 2GB/s, 可以兼容多个应用场合, 实际中频存储设备只用到了 500 MB/s, 存储容量高达 768 GB.

关键词: NAND; FLASH; CPCI; 流水线; 坏块

中图分类号: TP 303

文献标志码: A

文章编号: 1006-7930(2012)02-0300-05

随着信息科学的飞速发展, 数据的实时存储技术已经成为通信领域中非常重要的环节, 系统对存储容量和存储带宽的要求在不断提高.

固态存储卡由控制单元和存储单元组成, 它相较于传统的硬盘而言有以下的优点: 由于没有电机加速旋转的过程, 所以启动快; 由于不用磁头, 快速随机读取, 所以读延迟极小; 由于没有机械马达和风扇, 工作时几乎没有噪音^[1-2]; 由于内部不存在任何的机械活动部件, 不会发生机械故障, 所以不怕碰撞、冲击和振动^[3-4]; 写入速度快, 容量大等. 由于固态盘具有的这些优点, 使得它在高速大容量存储领域得到了广泛的使用. 本文设计并实现了一种基于 CPCI 的高速大容量的固态存储卡, 它的最终工作的实时数据传输速率为 500 MB/s(理论上可以达到 2 GB/s), 存储容量为 768 GB.

1 存储卡的系统设计

本存储卡为标准 6U 的 CPCI 板卡^[5], 其基本结构框图如图 1 所示, 板上使用两片 Virtex5 系列的 FPGA 作为盘阵控制器^[6], 使用 96 片 NAND Flash 组成固态盘阵, 一片 FPGA 管理 48 片 Flash. 两片 FPGA 作为 PCI9656 的局部端, 通过不同的地址空间来进行访问. 板间高速数据的传输通过 J4 上自定义的高速差分数据线进行. 由于存储的数据是按记录的文件进行管理, 所以选择了一片 DDR2 来存储各种文件管理表格. 另外 NAND Flash 都会存在坏块的问题, 设计中选择了一片 DDR2 来做为坏块数据的备份存储器.

另外, 根据芯片手册^[7], 每片最多有 640 个坏块, 在最坏的情况下, 每个 FLASH 模块有 $640 \times 4 \times 4$ 个坏块, 共 24 组, 总的坏块数为 245 760, 容量为 60 GB. 而实际应用中, 芯片的坏块数一般一片里面只有几十个坏块存在, 因此坏块的容量一般只有几个 GB, 满足项目的需求.

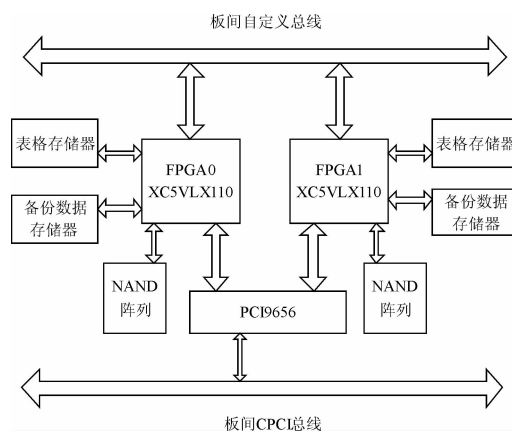


图1 存储板卡系统架构图

Fig. 1 System frame of storage card

该存储卡用于高速采样存储系统,要求写入速度快、存储容量大.系统选用的 NAND Flash 芯片为三星的 K9NCG08U5M,该 Flash 芯片单片的容量为 8 GB.该 Flash 的基本结构为:4 个 die,每个 die 包含 8 192 块,每块包含 64 页,每页包含 $(4\ 096 + 128)$ Byte.它的每个 die 的基本结构如图 2 所示.

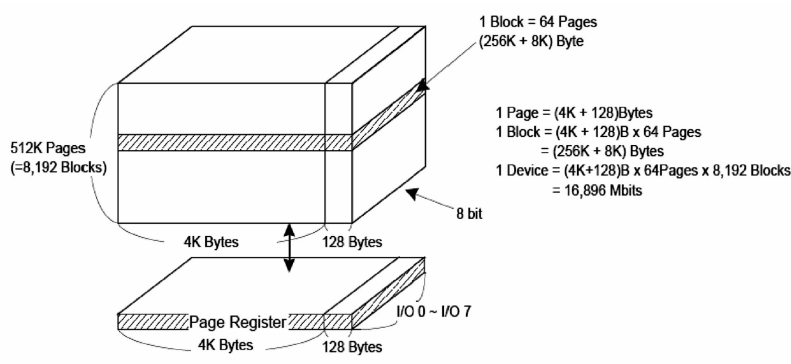


图 2 Flash Die 基本结构

Fig. 2 Basic structure of Flash Die

该 Flash 的写入时间为最小 45ns/Byte,当一页数据写入完成后就是芯片内部的编程操作了,

根据 Flash 手册上提供的数据,编程操作的最大时间为 700 us.由此可见,单片 Flash 的存储速度是很慢的(小于 5 MB/s).所以为了满足高带宽的要求,需要将多片 Flash 并行操作.另外 NAND Flash 的读写都是按页操作的,所以进出 Flash 的数据都需要进行缓存,单片 Flash 的缓存空间为至少一页即 4 KB.在考虑并行片数时,如果并行片数越多则需要的缓存空间就越大;并行片数太少,处理速度又跟不上数据率.在本设计中将 96 片 Flash 分为两个阵,分别由两片 FPGA 来控制其操作.每个阵的 48 片 Flash 又分为 12 组,每组 4 片并行操作,12 组 Flash 流水缓存.从整体看,相当于两个阵 96 片 Flash 并行操作,提高了操作速度的同时也节省了存储空间.即便这样设计后,系统的写入速度也只能提高到 480 MB/s,对于某些高速系统仍不能满足要求,在本设计中通过组内采用流水线操作来提高了系统的写入速度.该 Flash 芯片内有 4 个 die,4 个 die 共用数据线和读写控制线,但它们有各自的片选信号线和 R/B 信号线.鉴于此 Flash 结构,将同一组 Flash 的 4 个 die 设计为流水线操作.这样当 die0 的写操作完成后,不需要等待 die0 的编程完成就可以直接进行 die1 的写操作,当 die3 的写操作完成时,已经距离 die0 写操作完成有 700 多 us 的时间,大于 die0 的最大编程时间 700 us,这时又可以继续进行 die0 的第 2 页写操作,依次类推.如图 3 所示.经过这样的流水加并行设计后,系统的写入速度最快理论上可以达到 2 GB/s.

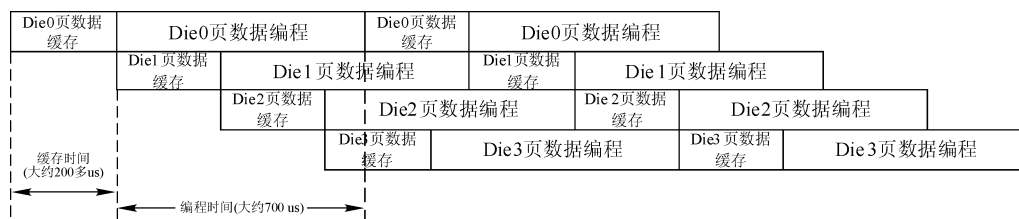


图 3 Flash 流水线操作示意图

Fig. 3 Pipeline diagram

通过前面系统描述,本设计的理论写入速度可以达到 2 GB/s,但是由于 NAND Flash 固有的缺点(在使用过程中会不时地出现新的坏块),一旦在采集写入过程中出现了新的坏块,该块数据需要重新写入,这样一来就会打断原来的写入过程.从整个采集过程来看,不仅减慢了采集存储的速度,还会增加额外的缓存空间.所以,如果在采集存储的过程中去实时处理坏块数据会加大系统的存储开销和降低系统的存储带宽.在本设计中通过文件管理、坏块管理和存储备份三个模块协调工作来解决该问题.将采集过程中输入 Flash 盘阵的数据同时备份到 DDR2,同时将遇到的坏块打上标记,然后在采集结束以后再将 DDR2 中的数据写入 Flash 中,修改 FAT 表的内容和链接关系.

2 固态盘存储控制器的功能设计

由于 NAND 型 Flash 具有写入带宽低和有坏块的缺点,所以需要设计合理有效的控制逻辑来克服

这些固有缺点.对于写入带宽低的缺点,如前所述,设计中采用并行操作和流水操作相结合的方法来提高带宽;对于有坏块的缺点,本设计则采用预先建立有效块表的方法,避免数据写入到坏块中去.在写入和擦除的过程中,如果遇到新的坏块,则将该坏块从有效块表中去除;对于数据的流入流出,为了数据读写的连贯性,为每个 Flash 控制器配置了一个大小为 16 KB 的缓存器,以用来匹配 J4 口和 CPCI 上的数据速率.数据写入时,当单个缓存器写满后,才能进行 Flash 的写入操作;而当数据读出时,只能当单个缓存器被读空以后才能将数据从 Flash 中读出.整个盘阵控制器的逻辑结构如图 4 所示.

下面分别简要介绍下各模块的功能:

2.1 盘阵控制器

盘阵控制器 FAC 接收来自主机端的命令,进行解析后分别给各模块发送操作指令.如,接收到的是采集命令,则 FAC 同时给 12 组 Flash 发送写入指令,并启动相应的文件管理;接收到的是擦除命令,则具体指定到某组的某个 die 进行擦除操作,等等.

2.2 Flash 控制器

Flash 控制器主要由缓存模块、状态机控制模块和寄存器堆组成^[8].其中缓存模块用来接收采集板送来的数据,容量为 16 KB,即一个 FLASH 组的一页.状态机控制模块用来控制对 FLASH 模块的各种操作,包括读、写、擦除、状态判断等.寄存器堆用来产生 FLASH 模块的各种控制信号,包括 RE、WE、ALE、CLE 以及 CE.

2.3 文件系统管理

由于 NAND Flash 存储单元中的数据不能简单的实现直接覆写,使得存储器不能像普通的硬盘那样随意的写入和更新.所以有必要设计一套适用于 Flash 存储器的简单文件系统.

文件管理系统用于管理文件的创建、删除、回放及导入\导出等工作.系统对每组 Flash 建立相应的 FREE 表和 FAT 表,FREE 表用于记录该组 Flash 中的空闲块,FAT 表用于记录每个文件的起始地址、结束地址以及链接关系等.每个文件创建时,系统从 FREE 表中选择一个空闲块作为文件起始,并记录到 FAT 表中,在文件录入的过程中,系统不停的从 FREE 表中取出空闲块直到文件采集完成,再将文件的结束地址和各地址间的链接关系记录到 FAT 表中.在文件回放或导出时,系统根据该文件的 FAT 表信息依次读出 Flash 中的数据.

2.4 坏块管理

NAND Flash 存储器存在一些初始无效块,并且在使用过程中也会出现存储单元的损坏,因此有必要记录每片 Flash 中坏块的位置.目前在 Flash 操作中处理坏块有很多种方法,在本设计中采用了建立有效块表的方法,即本文中涉及到的 FREE 表.在进行写操作时,从 FREE 表中取出数据即为有效块地址.

在本存储板卡中,由于对 Flash 进行分组并行操作,将会产生新的坏块问题.由于组内每片 Flash 的坏块地址不可能完全相同,而读写操作是统一的,因此一旦组内某片 Flash 的第 n 块是坏块,那么就认为组内所有 Flash 的第 n 块都是坏块.

2.5 数据备份

由于采集过程中数据在不停的输入,如果 Flash 在写入操作时产生了新的坏块,就需要对出错页的数据进行重新写入.当然,在实时存储的过程中,这个出错页重新写入的过程只能在采集完成后进行.另外新的坏块的出现是随机的,不固定的.基于上述原因,就需要在采集的过程中将输入 Flash 盘阵的数据同时进行备份.

在本设计中,将一片大小为 2 Gbit 的 DDR2 存储器作为专门的备份存储器使用.对应于 48 个 FLASH 模块,将 DDR2 分成 48 个部分,这样每个 FLASH 模块在每次采集的时候最多可以允许有 256

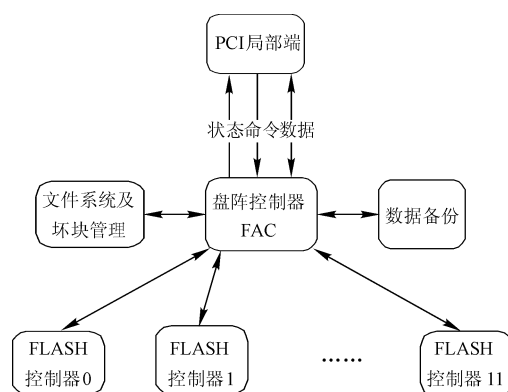


图4 存储控制器整体结构框图

Fig. 4 Holistic frame of Flash Array controller

个出错页的备份.

3 基于文件的存储卡工作流程举例

该存储卡的操作是基于文件方式的,存储卡的工作:初始化、采集、回放、数据导入和导出都是以文件的方式进行管理的.系统需要采集存储时,首先需要建立新文件,并为新文件分配新的 FREE 表和 FAT 表.然后存储卡接收主机发送的采集存储命令,缓存采集板过来的数据并进行固态存储.在数据采集存储的过程中需要不停地读取文件 FREE 表和更新文件 FAT 表.当存储卡收到主机发送的采集停止命令后,结束实时存储过程,并将之前存储过程中的坏块数据重新进行写入,最后更新 FAT 表及链接,整个存储过程结束.采集存储过程的流程如图 5 所示.

4 存储卡控制器的仿真与实现

该控制器在 FPGA 实现时采用了 100 M 的输入时钟,而 Flash 控制器的操作时钟为 16.67M.控制器最后实现占用的资源情况如图 5 所示.从图中可以看出,该控制器对 FPGA 逻辑资源的占用率不高,存储器的使用也还有冗余,所以接下来的工作可以通过增加数据的缓存级数以及提高 Flash 的操作时钟来提高存储卡的写入带宽.

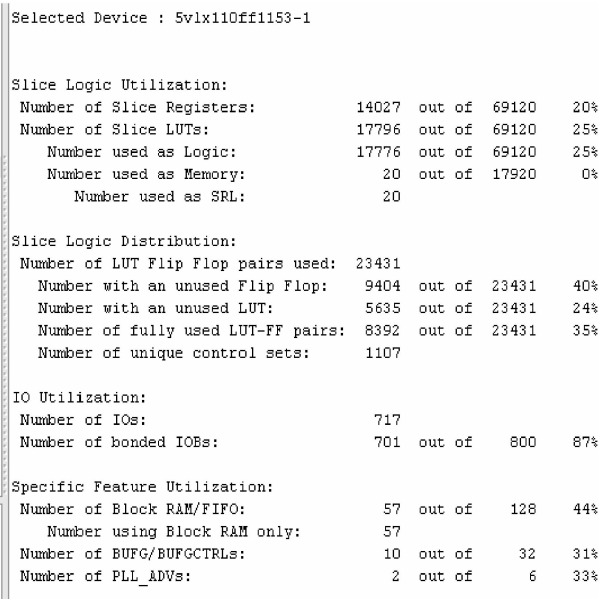


图 6 FPGA 资源占用情况图

Fig. 6 FPGA resource usage

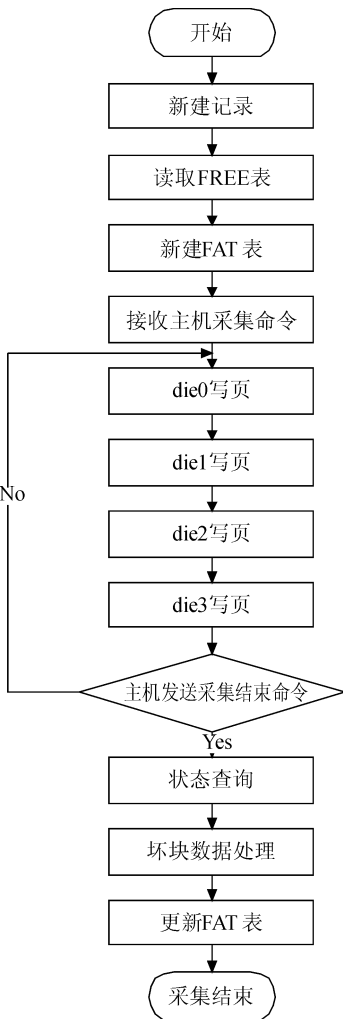


图 5 存储卡采集流程图

Fig. 5 Flow chart of storage

5 性能测试分析

本存储卡是中频采样存储设备中的一张板卡,到目前为止已经完成了系统级联测试、高温\低温工作试验、高温\低温储存试验,试验结果良好.对输入的-60 dBm~0 dBm 中频输入信号能够实时存储,回放时误码性能恶化≤1 dB.本存储卡的主要性能参数见表 1 所示.

表 1 存储卡性能参数	
Tab. 1 Performance parameter of the storage card	
参数	说明
板卡尺寸	标准 6U 尺寸
功耗	<5 W
板卡容量	768 GB
板卡存储速度	理论上为 2 GB/s,实际工作在 500 MB/s
板卡回放速度	理论上为 2 GB/s,实际工作在 500 MB/s
板卡数据导出速度	CPCI 32 bit 数据宽度,80 MB/s

6 结 语

本文设计并实现了基于 FPGA 为盘阵控制器的固态存储板卡,解决了海量高速数据的实时存储、多途径回放等问题.该存储卡在现有设备中的存储速度为 500 MB/s,理论上可以达到 2 GB/s 的输入速度,不仅能很好的用于该中频采样存储设备上,同时还具有很广泛的应用前景.

参考文献 References

- [1] 周 骞,陈金树.超高速数据采集存储系统的设计与实现[J].计算机应用研究,2003(8):139-141.
ZHOU Qian, CHEN Jin-shu. Design and Implementation of an Ultra-high Speed Data Acquisition and Storage System[J]. Computer Application and Research, 2003(8):139-141.
- [2] 李 超,王虹现,邢孟道.高速大容量 FLASH 存储系统设计[J].火控雷达技术,2007,36:83-88.
LI Chao, WANG Hong-xian, XING Meng-dao. Design of a FLASH Storage System with High Speed and Large Capacity[J]. Fire Control Radar Technology, 2007, 36:83-88.
- [3] 宋 杰,何 友,唐小明.基于 FPGA 的超高速雷达信号实时采集存储系统[J].电子技术应用,2005(11):18-20.
SONG Jie, HE You, TANG Xiao-ming. The Ultra-high Speed Radar Signal Real-time Acquisition and Storage System Based on FPGA[J]. Application of Electronic Technique, 2005(11):18-20.
- [4] 朱知博.基于 NAND FLASH 的高速大容量存储系统设计[J].现代电子技术,2011,34(8):170-173.
ZHU Zhi-bo. Design of High-speed and Mass Storage System Based on NAND Flash[J]. Modern Electronics Technique, 2011, 34(8):170-173.
- [5] TOM Shanley, DON Aderson. PCI 系统结构[M]. 4 版. 刘 晖,译.北京:电子工业出版社,2000:5-26.
- [6] XILINX. Virtex-5 User Guide[Z]. <http://www.xilinx.com>, 2010.
- [7] SAMSUNG. 2G x 8Bit / 4G x 8Bit / 8G x 8Bit NAND Flash Memory Datasheet <http://www.samsung.com>, 2007.
- [8] Micron. Micron NAND Flash Controller via Xilinx Spartan-3 FPGA. <http://www.micron.com>, 2005.

Design and realization of solid state storage card based on NAND Flash

YIN Lei

(China Academy of Engineering Physics, Institute of Electronic Engineering, Mianyang 621900, China)

Abstract: This paper describes the design and implementation of a high speed storage system based on CPCI, which provides data storage, process and replay analysis service during the intermediate frequency sampling procedure. The storage system is composed of two main parts, i. e. NAND Flash as its static memorizer and FPGA as its storage controller, and it already functions on XX intermediate frequency sampling project. The writing speed for data storage is further increased to 2 GB/s, while 500 MB/s is good enough to meet the project's need. The total storage volume reached 768 GB and it is capable to work in multiple applications as well. The paper provided several methods to meet such improvements as in managing the data in file format, managing the bad block by effective block in the mean time using DDR2 for data backup.

Key words: NAND; FLASH; CPCI; pipeline; bad block