

本科生毕业设计

|  |
| --- |
| 面向近似应用的固态硬盘数据布局方案的设计及实现 |

|  |  |
| --- | --- |
| 院 系 | 计算机科学与技术 |
| 专业班级 | CS1610 |
| 姓 名 | 覃映超 |
| 学 号 | U201614795 |
| 指导教师 | 崔金华 |

2020年06月10日

**学位论文原创性声明**

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包括任何其他个人或集体已经发表或撰写的成果作品。本人完全意识到本声明的法律后果由本人承担。

作者签名： 年 月 日

**学位论文版权使用授权书**

本学位论文作者完全了解学校有关保障、使用学位论文的规定，同意学校保留并向有关学位论文管理部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权省级优秀学士论文评选机构将本学位论文的全部或部分内容编入有关数据进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 1、保密囗，在 年解密后适用本授权书

2、不保密囗 。

（请在以上相应方框内打“√”）

作者签名： 年 月 日

导师签名： 年 月 日

摘 要

基于近似存储的闪存系统有望消除存储系统的能耗高、I/O瓶颈等难题，其发展进步将带来计算机体系结构以及系统软件层的革新。但直接在闪存存储系统中应用现有近似存储技术时，没有充分考虑闪存系统中闪存芯片固有工艺偏差导致数据访问性能低下等现象。

针对在近似数据布局方案中，闪存系统中不可避免的闪存工艺偏差的现象，首先，提出了近似存储技术，该技术使用近似写操作来存储应用程序中的近似数据，使用近似读操作从固态盘中读取近似数据。通过增加编程电压可以提高近似写操作的速度，通过不受错误编码纠正的读取操作可以提高读取速度。然后，提出了基于工艺偏差的布局方案，该方案通过将近似数据分配到可靠性较低的块中，精确数据分配到可靠性较高的块中。实验结果表明，与传统的存储方案相比，该方案将闪存访问的平均读写响应时间分别提高了55%和39%。

**关键词**：近似存储，工艺偏差，数据放置

Abstract

Failure resilience is of pivotal importance in practical network function virtualization (NFV) systems, but has been mostly absent in the existing ones. The absence is mainly due to the challenge of patching source code of the existing NF software for extracting important NF states, a necessary step toward flow migration and replication to provide failure tolerance.

This paper proposes NFActor, a novel NFV system that uses the actor programming model to provide transparent resilience, high scalability and low overhead in network flow processing. In NFActor, a set of efficient APIs are provided for constructing NFs, with inherent support for scalability and resilience. A per-flow management principle is advocated, different from the existing practice, which provides dedicated micro service chain services for individual flows, enabling decentralized flow migration and scalable flow replication. We implement NFActor and show that it achieves good scalability, prompt flow migration and failure recovery with large numbers of concurrent flows. We also show that NFActor can enable applications such as live NF update and correct MPTCP subflow processing, which cannot be efficiently achieved in previous systems.

**Keywords:** Network Function Virtualization (NFV), Failure resilience, Actor model, Flow migration, Service Chain

撰写说明，阅后删除！！！

英文摘要在中文摘要定稿后再开始撰写，否则会多次修改，不要过于依赖百度翻译等工具。

Abstract是论文的英文摘要，一般对照中文摘要翻译，要求另起一页。Keywords对照中文关键词翻译，与英文摘要在同一页，以“Keywords：”另行顶左开始。Keywords之间用“,”隔开，最后一个Keyword后不用加任何标点符号。

目 录

[摘 要 I](#_Toc39153221)

[Abstract II](#_Toc39153222)

[1 绪 论 1](#_Toc39153223)

[1.1 课题背景 1](#_Toc39153224)

[1.2 国内外研究现状 1](#_Toc39153225)

[1.3 研究目的和主要内容 2](#_Toc39153226)

[1.4 论文结构 2](#_Toc39153227)

[1.5 课题来源 3](#_Toc39153228)

[2 固态盘背景概述 4](#_Toc39153229)

[2.1 闪存单元 4](#_Toc39153230)

[2.2 增量步进脉冲编程 4](#_Toc39153231)

[2.3 闪存工艺偏差 5](#_Toc39153232)

[2.4 固态硬盘模拟器 5](#_Toc39153233)

[2.5 本章小结 6](#_Toc39153234)

[3 近似存储设计 8](#_Toc39153235)

[3.1 近似数据识别 8](#_Toc39153236)

[3.2 近似写入 8](#_Toc39153237)

[3.3 近似读取 8](#_Toc39153238)

[3.4 本章小结 9](#_Toc39153239)

[4 XXX系统实现 10](#_Toc39153240)

[4.1 过滤器实现 10](#_Toc39153241)

[4.2 属性管理模块实现 10](#_Toc39153242)

[4.3 数据迁移模块实现 10](#_Toc39153243)

[4.4 本章小结 10](#_Toc39153244)

[5 性能测试与分析 11](#_Toc39153245)

[5.1 测试环境 11](#_Toc39153246)

[5.2 功能测试 11](#_Toc39153247)

[5.3 系统界面 11](#_Toc39153248)

[5.4 性能测试 11](#_Toc39153249)

[5.5 本章小结 13](#_Toc39153250)

[6 总结与展望 14](#_Toc39153251)

[致 谢 15](#_Toc39153252)

[参考文献 16](#_Toc39153253)

撰写说明，阅后删除！！！

目录不会自动更新，排版变动后必须更新目录，更新目录方式：鼠标点击目录，按F9键或者右键更新域，选择更新整个目录即可。

目录仅包含一级标题和二级标题和标题样式，目录是全论文的纲要。中文摘要、Abstract、论文正文的各级标题（不包括第三级）、致谢、参考文献、附录等都应编入目录，标注其页码对照关系，但目录本身不出现在其中。中文摘要、Abstract、目录等使用希腊数字“I、II、…”编连续页码；论文正文、致谢、参考文献、附录等使用“1，2，3，… ”编连续页码。

# 绪 论

本章我们首先介绍了当前存储系统面临的挑战和技术的发展趋势，然后介绍了近似存储技术的产生和发展，介绍了国内外在存储方法和闪存工艺偏差方面的相关研究工作，最后对本文的主要研究内容以及目的作了说明。

## 课题背景

如今越来越多的应用程序对存储的数据的精度要求没那么高了，因此在为了确保数据能够精确存储的操作上所花费的时间和空间对于这些应用程序来说是浪费的。例如在线观看视频时，视频文件将会被缓存起来，当缓存已满时，视频数据将被写入闪存。由于以下两个原因，缓存的视频数据可以承受较高错误率：（1）这些数据大多数是编码的多媒体数据，具有很高的容错性，因为这些数据出现少量错误时对视频的影响通常并不明显；（2）即使它们已损坏，系统也可以从Internet重新下载。还有在计算机视觉，机器学习和传感器数据分析等领域中的应用程序也可以看到对计算精度较小的妥协，以换取效率的提高。基于这种背景，这些应用的请求数据在存储层近似存储能够提高固态盘的性能，能效和容量。

## 国内外研究现状

### 近似计算

近似存储建立在三大类相关工作的基础上：近似计算，优化对存储单元的访问以及容忍固态存储器中的故障。

近似计算是一个研究领域，旨在使用硬件和软件技术来优化容错程序的执行。程序员使用语言功能，分析或程序逻辑来控制近似执行。许多有关近似计算的工作都集中在计算本身（优化算法或处理器逻辑），还有一些工作利用了非易失性固态存储技术（如PCM和Flash）的属性的技术：损耗和多级单元[2]。并针对这两种属性提出了两种机制，以此提高固态存储器的性能，寿命和密度。第一种通过减少或消除保护带，使得用于写入多级单元的编程脉冲的数量大大减少，从而加快写入速度。第二种机制通过将近似数据映射到已用尽其硬件纠错资源的块上，从而减轻了磨损故障带来的影响并延长了存储器的寿命。

### 闪存工艺偏差

随着技术的发展和位密度的提高，闪存工艺偏差（PV）的影响已经被放大。先前的工作提出了利用工艺差异来提高存储系统的写入性能[1]。在未来面对大量的近似应用时，需要同时考虑近似存储数据和工艺差异。

## 研究目的和主要内容

在本次课题的研究中，首先我们在原有的闪存模拟器的基础上，设计出符合近似存储，工艺偏差要求的闪存模拟器。然后设计出考虑了闪存工艺偏差以及近似存储的数据布局方案，以提高固态盘的性能。

## 论文结构

本文的主要内容如下：

第一章我们首先介绍了当前存储系统面临的挑战和技术发展趋势，然后分析了对象存储技术的产生及发展现状，介绍了国内外在存储策略和智能存储领域的相关研究工作，并对本文的主要研究内容及工作意义作了具体说明。

第二章首先介绍几种存储技术、特别是存储区域网（SAN）在安全性、数据共享和设备管理等方面的存在问题。然后详细论述了基于对象的存储如何解决当前的问题。最后讨论了对象存储的实现途径，OSD标准和Lustre文件系统代表了两种发展方向。本文将重点以OSD为基础进行有关研究。

第三章以磁盘阵列作为研究对象，描述了属性控制的数据放置策略。我们建立了基于闭环fork-join排队网络（CFJQN）的RAID-5读写模型，对最优分条单元大小进行了定量分析，并归纳得到可以用来准确描述一个负载的若干属性（包括存储需求和行为）。

第五章比较详细的描述了一个原型系统的设计和实现。iSCSI-OSD-RAID成功地对传统的RAID控制器进行了扩展，增加了支持iSCSI和OSD协议的访问接口。除了软件方面的工作之外，还实现了基于Intel IOP321处理器的控制器硬件，本章对硬件结构和交叉开发环境也进行了介绍。最后还给出了部分测试结果。

第六章总结了所做的工作，并计划了下一步的工作以及展望了基于对象存储的前景。

## 课题来源

本课题为湖北省自然科学基金项目“基于近似存储的闪存系统关键技术研究”中的一个基础课题。

# 固态盘背景概述

如今固态硬盘已经广泛用于移动设备，服务器等领域中。为了提高硬盘容量，闪存单元从单级单元改进为多级单元，此外，采用三维结构的闪存通过垂直堆叠闪存单元进一步提高了存储容量。但是闪存发展的不利因素是可靠性下降，以及工艺偏差带来的可靠性差异。

## 闪存单元

NAND闪存将特定数量的电荷存储在浮栅或电荷陷阱中用于表示存储的数据。如果在闪存单元中存储n位，就需要种不同的电压状态。图2‑1显示了一个多级单元（MLC）闪存在一个闪存单元中存储两位信息的例子，其中定义了四个电压状态，通过在编程写入操作中将电荷注入到浮栅中，每个闪存单元将转变为四个电压状态之一。两个相邻电压状态之间的距离称之为保护带。

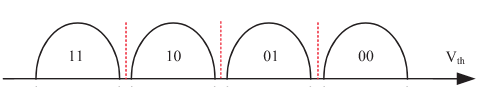


图2‑1 闪存单元的电压分布

## 增量步进脉冲编程

目前，闪存编程方式大部分采用的都是增量步进脉冲编程（ISPP）策略。典型的增量步进脉冲编程方案由多个步骤组成，每个步骤的编程电压增量通常称为ΔVpp。

研究表明，闪存的可靠性取决于每个状态的电压分布以及相邻电压状态之间的保护带。编程后单元电压可能会由于干扰和电荷泄露而产生波动，少数闪存单元的电压可能会偏离其电压状态的范围，从而导致原始位错误（RBE），因此在将数据从闪存中读取出来时，需要经过错误编码纠正（ECC）修复。闪存页面包含的位错误越多，读取期间进行纠错的过程越长，读取性能就越低。选择较小的ΔVpp有助于减小电压状态范围，使得保护带范围变大，即数据更可靠，因此读取速度也更快。但是，完成写入操作所需要的步骤增多，写入时间增加。

## 闪存工艺偏差

工艺偏差是晶体管属性自然发生的差异[8][9][10]。随着技术的发展和位密度的提高，闪存的工艺差异已经被放大，特别是对于不同块所支持的可用P/E周期数[3][4]。许多工作表明，一个块的可用P/E周期数表示该块的强度。Pan等[3]指出闪存块的误码率（BER）遵循对数高斯分布，最终反映了闪存块可用P/E周期数的分布。Woo等[4]进一步评估了块的可用P/E循环数远高于上述数字，并且块之间存在着显著差异。基于这些观察，可以得出结论，工艺偏差会导致闪存块强度明显变化。随着闪存存储容量需求的急剧增加，三维闪存（3D NAND）通过在第三个维度上堆叠闪存芯片来提升存储容量获得了广泛地研究。但是，工艺偏差在闪存块中仍然广泛存在。在3DNAND中，闪存块中的闪存单元彼此垂直堆叠，并且由于制造工艺的原因，3DNAND单元的大小在不同层之间会有很大差异，这称为层间的工艺偏差。

## 固态硬盘模拟器

本课题所采用的模拟器为3Dsim。3Dsim用来模拟固态盘的硬件结构（包括闪存颗粒，固态盘内部通道等），闪存转换层（包括数据缓存区、地址映射、垃圾回收及损耗均衡等）。3Dsim分成5个逻辑模块，分别为时间线模块、数据接口模块、数据缓存模块、闪存转换模块、高级命令模块、闪存介质模块。3Dsim系统架构如图 2‑2所示，3Dsim中始终维护有一个时间线和状态机，通过状态机的事件跳转，3Dsim处理来自上层应用发出的对3D flash执行读写负载请求，解析请求，请求的数据通过缓存调度，经过FTL层的地址映射、垃圾回收等操作后，转换成为闪存可以响应的命令，在此模块可以判断是否可以组成高级命令，通过不同高级命令对3Dflash进行操作，3Dflash响应操作，更改自己的状态表并输出统计信息到文件，最后修改时间线，反馈给状态机，进行新一条trace的读取，如此反复。

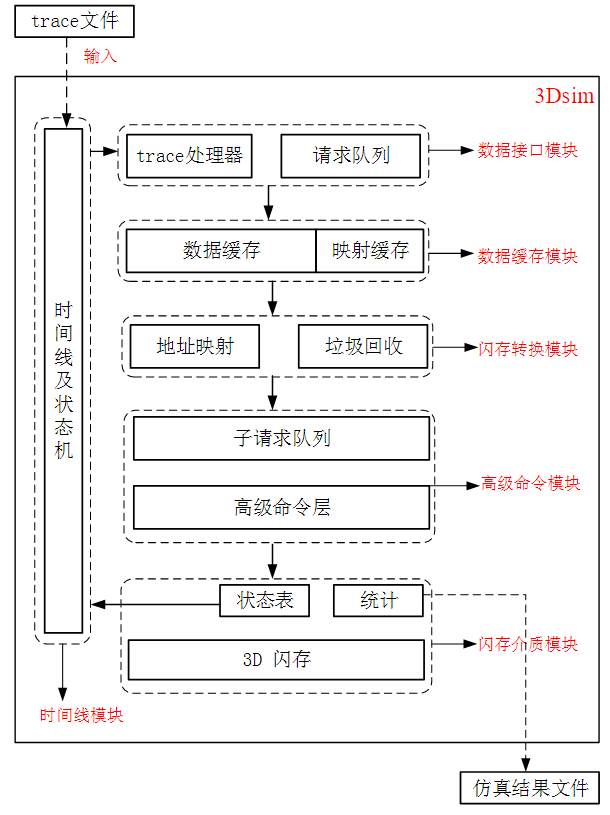


图 2‑2 3Dsim基本架构图

在具体应用时，主要需要编写的是闪存转换模块和闪存介质模块，设计出新的数据分配方案，以及对应的垃圾回收机制。

## 本章小结

本章介绍了关于闪存的一些背景知识，包括闪存单元、增量步进脉冲、闪存工艺偏差、固态盘模拟器。在介绍闪存单元时，阐述了多级单元的概念，固态盘存储数据的方法；介绍增量步进脉冲时，主要阐述了不同的ISPP方案对性能以及可靠性的影响；介绍闪存工艺偏差时，主要阐述了由于制造工艺的原因导致闪存块之间存在明显的可靠性差异；介绍固态盘模拟器时，主要阐述了模拟器的主要架构。

# 近似存储设计

近似存储技术提供了一种途径，用于缓解存储领域中面临空间开销大、能耗高、访问性能低等挑战。在本节中，我们提出了实现近似存储的一种方式。

## 近似数据识别

识别近似数据是近似存储的前提，应用程序的数据并不是都能够近似处理，所以需要标识近似数据，然后对近似数据执行近似存储操作。应用程序的近似数据通常具有容错性，因为具有容错能力的近似数据在执行近似操作后其质量降低对原应用程序的影响不大。

目前，用于区别近似数据和精确数据的方法有故障注入法和程序分析法。其中程序分析法，主要依靠程序员或近似计算感知的编译器（例如Enerj[2]）对近似数据进行注释。本文在实现基于闪存的近似存储设计时，做出了一个假设，即标记近似数据与精确数据的标签跟随每一个IO请求一起传递到数据接口模块。

在固态盘模拟器中，负载踪迹文件既通常所说的trace文件，用作记录特定的负载的所有的请求的到达时间，请求的类型，请求的逻辑地址，请求的大小。3Dsim根据这个踪迹文件，可以重现出这个特定负载或者应用环境下的请求流。在实现近似存储时，负载踪迹文件还需要添加一个参数，用于标识当前的I/O请求是否可以使用近似存储模式。也就是说，传给固态盘的I/O请求至少包括五项： （到达时间，请求的类型，请求的逻辑地址，请求的大小，近似标签）

并在闪存的页面信息中添加一位，用于表示当前页面存储的数据类型。

近似数据的识别位于数据接口模块，该模块模拟固态盘与主机逻辑通信端的数据通信，也就是接收下发的trace，从其中读取应用请求，根据请求到达的时间判断是否阻塞；根据请求的近似标签为其对应的请求结构体中的变量approxFlag进行赋值，该变量主要用于为下一阶段生成的子请求结构中的approxFlag赋值。并挂载在请求队列上，供数据缓存使用。如图 3‑1所示。

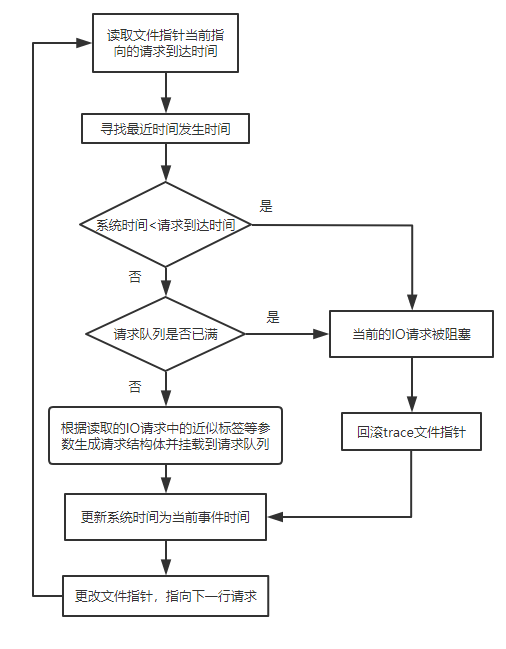


图 3‑1数据接口模块流程图

数据接口模块将读取的IO请求生成对应的请求结构体，每个结构体包含请求的所有信息：到达时间，逻辑设备号，访问起始扇区号，请求长度，操作类型，近似标签。并将请求结构体加入到请求链表中，等待进一步的处理。

## 近似写入

根据闪存的写入速度与编程步长（ΔVpp）之间的关系，近似写入的一种简单实现方式是使用比传统存储模式更大的ΔVpp，从而减少了写入步骤的数量，并提高了写入性能。近似写入操作是针对于近似数据的，当请求中访问的数据是近似数据时，就可以执行近似写入操作。

写入操作位于闪存介质模块，在执行写入操作之前需要从写子请求队列中找到符合条件的子请求，然后为该子请求分配物理地址，最后再根据子请求结构中的近似标签决定使用的写入方式。如图 3‑2所示。

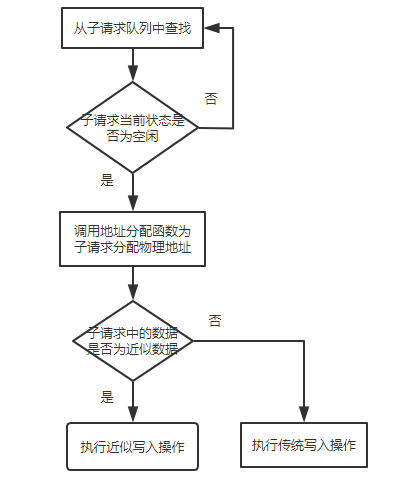


图 3‑2近似写入流程

## 近似读取

读取速度主要受到错误代码纠正（ECC）的影响，闪存页面包含的位错误越多，即RBER更高，读取期间进行纠正的过程耗时就越长，读取速度就越慢。而应用程序对近似数据具有容错性。本文采用的近似读取方式是利用近似数据的容错性来实现，即当应用程序需要读取的是近似数据时，可以在不使用ECC保护的情况下执行读取操作，进而提高读取性能。

读取操作位于闪存介质模块，但在执行读取操作前，需要从读子请求队列中查找符合条件的子请求，获取该子请求访问的物理地址信息，再根据子请求中的近似标签执行响应读取操作。如图 3‑3所示。

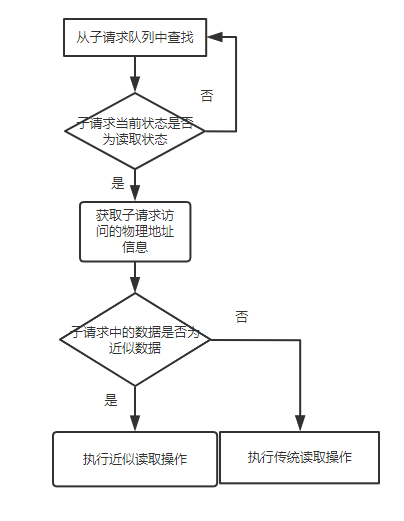


图 3‑3近似读取流程

## 本章小结

本章详细描述了本课题实现近似存储时所采用的方案，对如何识别近似数据，如何执行近似写入操作以及如何执行近似读取操作做了说明。

# 数据布局方案设计

工艺偏差已经广泛的存在于闪存中，并对闪存块之间的可靠性造成很大的差异。在本节中，我们提出了一种利用近似数据以及闪存工艺偏差来提高存储系统性能的方法。首先提出了一种块速度检测方案，以确定块的写入速度。然后，提出了一种数据分配方案进一步利用块之间的可靠性差异来提高性能。

## 速度检测

在本节中，我们首先讨论块的可靠性特征。然后提出了本文速度检测的方案。

### 块的可靠性特征

闪存块的可靠性主要体现在编程擦除（P/E）周期数和块的误码率（RBER），而块的误码率又可以体现读取速度上。在2.3节中提到了，闪存块的误码率遵循对数高斯分布。

### 写入速度

在2.2节中提到了ΔVpp与写入操作的速度成正比。多项工作指出，当闪存块的可靠性较高时，可以增加ΔVpp以提高写入速度，而不会超出闪存中错误编码纠正（ECC）的纠错能力。由于工艺偏差在闪存块之间引入了很大的可靠性差异，因此必须根据其自身的可靠性来设计单个块的写入速度。

### 读取速度

## 数据分配

在本节中，提出了一种利用近似数据和闪存工艺偏差的数据分配方案，其基本思想是使近似数据分配到可靠性较低的块，精确数据分配到可靠性较高的块。如图 4‑1所示。

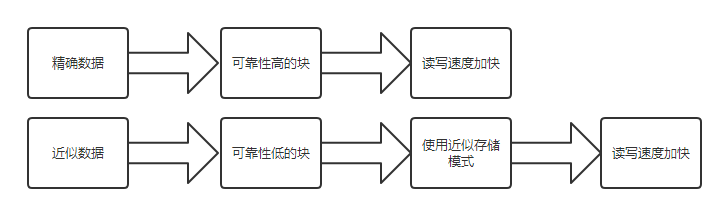


图 4‑1数据分配对应图

对于精确数据而言，一方面，存储在可靠性较高的块中，可以使用更大的编程电压ΔVpp，从而加快写入速度；另一方面由于数据存储在可靠性较高的块中，其误码率（RBER）较低，读取操作耗时较少，使得读取速度更快。对于近似数据而言，近似数据使用近似存储模式，可以利用应用程序对近似数据的容错性，使得其输出质量可以接受。这样一来闪存整体性能都得到了提升。

## 数据迁移模块实现

## 本章小结

本章描述了磁盘自愈与数据主动迁移模块的具体实现，然后对各功能模块的具体实现进行了详细分析，最后介绍了该模块的动态运行流程。

# 性能测试与分析

## 测试环境

我们首先对RAID软件进行了评估。测试平台配置为：CPU Intel Xeon 1-8GHz，512MB内存，LSI Logic 21320-R (1030) ULTRA320 SCSI HBA，连接4块Seagate Cheetah Ultra320磁盘（MODEL ST373307LC），一块Agilent 5221A FC HBA（2Gb），与主机（启动器）的Qlogic 2310F FC卡通过光纤连接。为评价磁盘阵列控制器最终性能，我们利用Iometer软件对阵列控制器进行了详细的性能评测，具体测试环境见**错误!未找到引用源。**：

表5‑1 光纤通道启动器与阵列控制器配置

|  |  |  |
| --- | --- | --- |
| 配 置 | 启动器 | 阵列控制器 |
| CPU | Xeon 2.0 | Intel 80321 |
| Memory | 512MB | 1GB |
| Fiber Channel HBA | Qlogic 23102G | Agilent DX2 |
| OS | Windows 2003 | ARM Linux |

## 功能测试

## 系统界面

不是必须

## 性能测试

其中阵列控制器挂接6个Maxtor STM3320820AS磁盘，利用6个磁盘分别构建RAID0、RAID10、RAID5三种不同的阵列级别，分条大小为128K字节，缓存大小为512MByte，LUN大小为500GByte，阵列控制器采取写回策略。



图5‑1 磁盘阵列控制器RAID5顺序读写IOPS性能

**错误!未找到引用源。**为6个磁盘构建RAID5时阵列控制器顺序读写IOPS(I/O Per Second)性能示意图，从图中可以看出，IOPS随着数据块大小变大而变小，当数据块大小为512Byte时，阵列控制器处理读请求的性能可以达到每秒15000个，写性能由于包含冗余校验信息计算，IOPS略低，约为12000，此指标高于同类美国阵列nStor 4502F，其写IOPS最大为7000左右。



图5‑2 磁盘阵列控制器RAID5读写性能

为6个磁盘构建RAID5时阵列控制器数据传输率示意图，从图中也可以看出，随着主机端I/O请求数据块大小逐渐增大，阵列控制器数据传输率逐渐增加，当数据块大小达到128Kbyte时，系统性能最佳，读性能可以达到185MByte/s，写性能约为167MByte/s。**（有坐标轴的图必须给出单位，坐标轴意义，所有测试图应该有详细分析）**

从总体性能评测结果看，本文最终设计的FC-SATA磁盘阵列控制器基本发挥了Intel IOP处理器的最佳性能，同比性能优于同类进口磁盘阵列，基本达到了设计要求，如需要更高的数据传输率，必须同时提升主机通道以及IOP本身的性能，才能有实质性的提高。

## 本章小结

本章给出了控制器完整的软硬件设计方案，提出了主机板与通道板分离的架构，可以灵活构建全系列阵列控制器产品线；提出了磁盘阵列软件分层异步体系结构，可以有效提高I/O请求处理的并行度，减少I/O处理同步等待时间；最后提出了一种针对阵列层次性结构的细粒度测试方法，可单独测试评价各模块性能，为系统性能精确评测，以及故障诊断提供决策依据，最后给出了阵列控制器性能综合评测。

# 总结与展望

基于对象的存储是为了克服当前基于块的存储存在的诸多难题，在存储接口和结构层次的重要发展。基于对象的存储设备可以具有更多智能性，可以根据应用负载选择优化的存储策略。实现属性控制的存储策略，关键需要解决两个方面的问题：第一，属性表达什么信息，怎样影响存储策略？第二，属性怎样从用户应用传递给存储设备？围绕这两个问题，作了如下几点研究和工作：

1. 提出了对象属性控制存储策略的模型。OSD协议作为SCSI的扩展集，可以在TCP/IP网络和iSCSI协议之上来传输，基于对象的存储设备接收OSD命令并进行处理。
2. 结合对应用负载特征的分析，提出根据对象属性描述的负载特征动态选择RAID级别和分条单元大小的数据放置策略。I/O请求较小的对象放在RAID-10中，同时分条单元尺寸较小；反之则放在RAID-5中。
3. 建立了属性的传递机制，一种是应用程序接口静态传递；另一种是动态机制，即根据属性当中包含的文件系统信息预测文件的分类，并根据分类做出进一步的存储策略选择。
4. 论文总结必须用这样的条目的形式给出，总结是对全文的总结，不是对毕业设计的心得，心得部分不要写。

本文为更好地认识基于对象的存储以及OSD的实现进行了有益的探索。该领域还有很多研究工作有待完成，包括：

* 以OSD为节点构建大规模（PB级）存储系统
* 属性控制的caching/prefetching策略
* 实现Device-aware（表达设备能力，按需分配资源）
* 设备的学习能力，以文件系统和历史访问信息等为提示进行预测

随着存储需求的爆炸性增长，下一代互联网络必须有新的存储结构来应对构建和管理更大规模存储系统的挑战。基于对象的存储是最有希望的突破点之一。相信存储技术的发展一定能够满足人们不断增长的存储需求。

致 谢

论文完成之际，首先要感谢我的导师崔金华教授。他站在学科发展的前沿，从论文的选题，研究工作逐步深入，到论文的撰写，都给我以细致的指导和建设性的意见，使我得以圆满而顺利地完成……。XXX严谨的治学态度、诲人不倦的师德和一丝不苟的工作作风将会给我留下不可磨灭的记忆……。

在…… 四年里，得到了许多老师和同学的大力帮助和支持，在此表示深深的谢意。感谢XXX等老师对我的关心和帮助。在课题研究和项目开发过程中，我与很多同学一起度过了一段段难忘的时光，值得怀念。我们大家共同创造的良好的学术氛围，将给我以永远而美好的回忆。

……

最后，我要深深地感谢我的父母，他们给予我无尽的关怀和无止境的爱，….

致谢属于论文的辅文部分。使用第一人称，采用散文体，对指导教师以及协助完成设计的有关人员表示谢意，并可简述自己通过本次毕业设计的体会，注意只写是查重最容易出问题的地方，请千万不要看别人写的，照搬。

参考文献

1. L.Shi, Y. Di, M. Zhao, C.-J. Xue, K. Wu, and E. Sha. Exploiting Process Variation for Write Performance Improvement on NAND Flash Memory Storage Systems. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 2016, 24(1):1-4.
2. A. Sampson, J. Nelson, K. Strauss and L. Ceze, "Approximate storage in solid-state memories," 2013 46th Annual IEEE/ACM International Symposium on Microarchitecture (MICRO), Davis, CA, 2013, pp. 25-36.
3. Y. Pan, G. Dong, and T. Zhang. Error rate-based wear-leveling for nand flash memory at highly scaled technology nodes. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, 21(7):1350–1354, July 2013.
4. Y.-J. Woo and J.-S. Kim. Diversifying wear index for mlc nand flash memory to extend the lifetime of ssds. In EMSOFT 2013, pages 1–10, Sept 2013.
5. J. Cui, Y. Zhang, L. Shi, C. Xue, W. Wu，and J. Yang. ApproxFTL: On the Performance and Lifetime Improvement of 3D NAND Flash based SSD，IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems， 2018, 37 (19) :1957-1970.
6. J. Cui，Y. Zhang，W. Wu，J. Yang, Y. Wang, and J. Huang. DLV: Exploiting Device Level Latency Variations for Performance Improvement on Flash Memory Storage Systems. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems， 2018，37(8):1546-1559.
7. Q. Li, L. Shi, J. Yang, Y. Zhang and C. J. Xue, "Leveraging Approximate Data for Robust FIash Storage," 2019 56th ACM/IEEE Design Automation Conference (DAC), Las Vegas, NV, USA, 2019, pp. 1-6.
8. M.-F. Chang and S.-J. Shen. A process variation tolerant embedded split-gate flash memory using pre-stable current sensing scheme. IEEE Journal of Solid-State Circuits, 44(3):987–994, March 2009.
9. J. Chen and L. E. Cleveland. System for providing tight program/erase speeds that are insensitive to process variations, Aug. 11 1998. US Patent 5,793,249.
10. M. Meterelliyoz, A. Goel, J. Kulkarni, and K. Roy. Accurate characterization of random process variations using a robust low-voltage highsensitivity sensor featuring replica-bias circuit. In ISSCC 2010, pages 186–187, Feb 2010.

装订毕业论文时应将任务书原件装订到最后



**本科生毕业设计任务书**

|  |  |
| --- | --- |
| 题 目 | 面向近似应用的固态硬盘数据布局方案的设计及实现 |
|  |  |

（任务起止日期：2020年1月1日～2020年6月10日）

|  |  |
| --- | --- |
| 院 系 | 计算机科学与技术 |
| 专业班级 | CS1610 |
| 姓 名 | 覃映超 |
| 学 号 | U201614795 |
| 指导教师 | 崔金华 |

教研室（系、所）负责人 2020年1月6日 审查

院（系）负责人 2020年1月9日 批准

|  |
| --- |
| 课题内容 |
| 实现一个简单的分布式网络系统，实现文件的分布式网络化存储。，实现文件的分布式网络化存储。，实现文件的分布式网络化存储。，实现文件的分布式网络化存储。，实现文件的分布式网络化存储。，实现文件的分布式网络化存储。，实现文件的分布式网络化存储。，实现文件的分布式网络化存储。 |
| 课题任务要求 |
| 研究学习文件系统的相关原理及实现研究分布式系统概念及应用研究网络文件系统的原理、架构及实现。研究学习文件系统的相关原理及实现研究分布式系统概念及应用研究网络文件系统的原理、架构及实现。研究学习文件系统的相关原理及实现研究分布式系统概念及应用研究网络文件系统的原理、架构及实现。 |
| 主要参考文献 |
| 1. 董晓明，谢长生．基于对象的进化存储系统研究．计算机科学，2005，32(11): 223~226 2. 庞丽萍编．操作系统原理（第二版）．武汉：华中理工大学出版社，1994.9．225~270 3. （美）Nils J. Nilsson著；郑扣根等译．人工智能（Artificial Intelligence）．北京：机械工业出版社，2000.9．177~194 4. （美）Tom Mitchell著；曾华军等译．机器学习（Machine Learning）．北京：机械工业出版社，2003.1．38~56 5. 谢长生，董晓明，万继光，谭志虎，刘瑞芳．磁盘阵列控制器的设计与原型实现．小型微型计算机系统，2006, 27(1): 173~176 6. （美）Nils J. Nilsson著；郑扣根等译．人工智能（Artificial Intelligence）．北京：机械工业出版社，2000.9．177~194 |
| 同组设计者 |
| 无 |

**成 绩 评 定**

**指导教师评定意见**

装订论文时应将成绩评定页放最后，导师评语放前面！

一、对毕业设计（论文）的学术评语（应具体、确切、实事求是）

|  |
| --- |
|  |

二、对毕业设计评分

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 评分项目  (分值) | 调研论证  (10分) | 外文翻译  (5分) | 设计(论文)撰写质量  (10分) | 学习态度  (10分) | 基本理论和基本技能  (50分) | 创 新  (15分) | 合 计  (100分) |
| 得分 | **8** | **4** | **8** | **8** | **40** | **12** | **80** |

指导教师签字：**崔金华** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 年 月 日

**答辩小组评定意见**

答辩小组意见和综合评定页放最后，双面！

一、评语（根据学生答辩情况及其设计（论文）质量综合评价）

|  |
| --- |
|  |

二、评分

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 评分项目  (分值) | 答 辩 情 况 | | 论 文 质 量 | | 合 计  (100分) |
| 答辩情况  (15分) | 回答问题情况  (25分) | 规范要求与文字表达  (20分) | 学术水平  (40分) |
| 得分 |  |  |  |  |  |

答辩小组长签字：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 年 月 日

**毕业答辩及成绩评定说明**

1. 毕业答辩
2. 答辩前，答辩小组应详细审阅每个学生的毕业设计（论文），为答辩做好准备。
3. 严肃认真组织答辩，开好答辩会。
4. 指导教师应参加所指导学生的答辩会，但评定其成绩时宜回避。
5. 答辩中要做好记录以供成绩评定时参考。
6. 成绩评定
7. 答辩前每个学生都要将自己的毕业设计（论文）在指定时间内交给指导教师，由指导教师审阅，写出评语并预评分。
8. 答辩工作结束后，答辩小组应举行专门会议进行讨论，在参考指导教师预评结果的基础上，结合学生毕业设计（论文）质量和学生答辩情况，综合评定每个学生的成绩。
9. 院（系）对专业答辩小组提出的优秀和不及格的毕业设计（论文），要组织院（系）级答辩，最终确定成绩，并向学生公布。
10. 各专业学生的最后成绩应符合正态分布规律。
11. 请用蓝、黑钢笔手写或五号宋体字编辑，签名须手写，A4纸双面打印。

**毕业设计（论文）成绩评定**

|  |
| --- |
| 班号：**CS1610** 学生姓名：**覃映超**  综合成绩：\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_分（折合等级\_\_\_\_\_\_\_\_\_\_\_\_\_\_）  答辩小组长（签名）：\_\_\_\_\_\_\_\_\_\_\_\_\_\_ 年 月 日 |