

学校编码: 10384
学号: 23020141153142

分类号 _____ 密级 _____
UDC _____

廈門大學

硕士学位论文

基于应用感知的重复数据删除存储系统的
性能优化研究

**Research on Application-aware Performance Optimizations
for Deduplication-based Storage Systems**

陈 晓

指导教师姓名: 吴素贞副教授
专业名称: 计算机科学与技术
论文提交日期: 2017年4月
论文答辩时间: 2017年5月
学位授予日期: 2017年 月

答辩委员会主席:
评阅人:

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

厦门大学博硕士学位论文摘要库

摘要

近几年，数据量急剧膨胀，主存储系统和小型移动设备都面临着数据存储能力不足的挑战。重复数据删除技术可以有效地减少需要存储的数据量，缓解存储设备的空间压力。但是在节省存储空间的同时，重复数据删除技术会带来一定的额外开销，主要包括元数据的查找和管理、数据碎片化引起的读放大问题等，在一定程度上制约了重复数据删除技术在大数据环境下的性能和系统的可扩展性。

对主存储系统和移动设备中的读写数据内容的分析结果表明，在主存储系统和移动设备上，应用所产生的数据存在着大量的冗余，但是这些冗余数据均存在着局部性现象，即数据冗余基本上存在于同种应用产生的数据中，不同应用的数据之间的冗余率很低。

在主存储系统中，提出一种应用感知的重删性能优化方法 AA-Plus，利用数据冗余的局部性现象将重复数据删除后的非冗余数据块按照应用类型分开存储和管理，使得同种应用类型的数据块在存储设备上的存放位置更加集中，减轻了数据的碎片化问题，缓解了主存储系统中重复数据删除造成的读放大问题。性能测试结果表明，AA-Plus 显著提升了单个应用的数据块的集中程度，减少了数据的碎片化程度，提高了重复数据删除存储系统的吞吐量。

在处理器和内存资源紧缺的移动设备中，提出一种应用感知的智能手机存储性能优化方法 APP-Dedupe，按照应用分别管理重复数据删除的元数据缓存，根据系统需求对重复数据删除的元数据缓存进行换入和换出，减小了单元数据操作涉及的元数据规模，减少了基于闪存的移动设备上重复数据删除的元数据缓存开销及元数据查询和修改时的系统资源使用量。性能测试结果表明，APP-Dedupe 减少了手机存储系统的数据写入量，降低了应用的平均响应时间，减少了闪存发生垃圾回收操作的次数，提升了智能手机存储系统的性能和可靠性。

关键词：重复数据删除；应用感知；元数据管理；

Abstract

In recent years, with rapidly increasing of data, primary storage systems and small mobile devices encounter the challenges of insufficient data storage capacity. Data deduplication can effectively reduce the amount of data that should be stored, thus reducing the space pressure of storage systems. However, data deduplication will bring some additional overheads, such as the search and management of metadata, and the read amplification problem caused by data fragmentation, which restricts the performance and extendability of data deduplication in the big data environment to a certain extent.

The analysis of the read and write data shows that there is a great deal of data redundancy generated by applications in the primary storage system and mobile device. However, there is strong locality in the redundant data. That is, data redundancy mainly exists in the data that is generated by the same application. There is low redundancy rate among the data of different applications.

In primary storage systems, we propose an application-aware performance optimization method for data deduplication, called AA-Plus. By exploiting the locality of data redundancy, the non-redundant data is stored and managed according to the application types after data deduplication. The data belonging to the same type of applications is stored more centrally on the storage device, alleviating the data fragmentation problem and the read amplification problem caused by data deduplication in the primary storage system. The performance evaluation results show that, AA-Plus significantly increases the centralization of data blocks belonging to an individual application, reduces the degree of data fragmentation and improves the throughput of primary deduplication-based storage systems.

In mobile devices with insufficient processor and memory resources, we propose an application-aware deduplication optimization method, called APP-Dedupe, to improve the performance of smartphone-based storage systems. APP-Dedupe

manages the metadata cache of data deduplication respectively according to the type of applications. The metadata of data deduplication is dynamically swapped in to and out from the cache according to the system requirement, which reduces the amount of metadata involved by a single metadata operation, and reduces the overhead of metadata cache of data deduplication on flash-based mobile devices as well as the amount of system resources used for metadata queries and modifications. Performance evaluation results show that, APP-Dedupe reduces the amount of data written in the mobile storage system, the average response time of applications and the number of garbage collection operations for flash, thus improving the performance and reliability of smartphone-based storage systems.

Key words: Data Deduplication; Application-aware; Metadata Management;

厦门大学博硕士学位论文摘要库

目录

摘要.....	I
Abstract	II
第一章 绪论.....	1
1.1 课题背景和研究意义.....	1
1.2 重复数据删除技术的研究现状	3
1.3 论文主要研究内容	6
1.4 章节安排.....	7
第二章 主存储系统中应用感知的重删性能优化	9
2.1 问题分析与研究动机.....	9
2.1.1 重复数据删除基本流程.....	10
2.1.2 数据碎片化现象及其对性能的影响.....	14
2.1.3 重复数据的局部性现象.....	17
2.1.4 研究动机.....	18
2.2 系统设计与实现	18
2.2.1 系统整体架构.....	19
2.2.2 应用感知的数据块存储设计.....	20
2.2.3 数据一致性维护	21
2.3 性能评估与分析	22
2.3.1 测试环境与数据集.....	22
2.3.2 测试方法与测试参数.....	23
2.3.3 测试结果与分析.....	25
2.4 本章小结.....	29
第三章 应用感知的智能手机存储性能优化.....	31
3.1 问题分析与研究动机.....	31

3.1.1	移动设备和主存储系统的区别.....	33
3.1.2	内容感知的应用负载收集.....	35
3.1.3	智能手机的负载特性分析.....	38
3.1.4	研究动机.....	39
3.2	系统设计与实现	41
3.2.1	系统整体架构.....	41
3.2.2	应用感知的元数据缓存管理.....	42
3.2.3	应用感知的数据块布局.....	44
3.3	性能评估与分析	44
3.3.1	测试环境与方法.....	44
3.3.2	性能结果与分析.....	45
3.4	本章小结.....	50
第四章	全文总结	51
参考文献	53
攻读硕士学位期间完成的论文及申请的专利	61
致谢	62

Content

Abstract Chinese	I
Abstract English	II
Chapter 1 Introduction	1
1.1 Research Background	1
1.2 Relate Workon Deduplication	3
1.3 Main Content of the Paper	6
1.4 Paper Structure	7
Chapter 2 Application-aware Optimizations for Primary Deduplication Storage	9
2.1 Background and Motivation	9
2.1.1 Basic Processes of Deduplication	10
2.1.2 Data Fragmentation and Its Impact on Performance	14
2.1.3 Locality Phenomenon of Deduplication	17
2.1.4 Motivation.....	18
2.2 Design and Implementation	18
2.2.1 System Architecture	19
2.2.2 Design of Application-aware Block Storage.....	20
2.2.3 Data Consistency	21
2.3 Performance Evaluation	22
2.3.1 Experimental Setup and Data Sets	22
2.3.2 Evaluation Methodology and Parameters	23
2.3.3 Performance Results and Analysis.....	25
2.4 Conclusion	29

Chapter 3 Application-aware Optimizations for Smartphones	31
3.1 Background and Motivation	31
3.1.1 Differences between Primary Storage System and Mobile Devices	33
3.1.2 Content-Aware Application Trace Collection	35
3.1.3 Analysis of Workload Characteristics of Mobile Devices	38
3.1.4 Motivation.....	39
3.2 Design and Implementation	41
3.2.1 System Architecture	41
3.2.2 Application-Aware Metadata Cache Management	42
3.2.3 Application-Aware Block Management.....	44
3.3 Performance Evaluation	44
3.3.1 Experimental Setup and Methodology.....	44
3.3.2 Performance Results and Analysis.....	45
3.4 Conclusion.....	50
Chapter 4 Summary	51
References.....	53
Publications	61
Acknowledgments.....	62

第一章 绪论

1.1 课题背景和研究意义

根据 IDC 给出的数据，全球的数据量在 2010 年正式突破 ZB，预计 10 年后将达到 44ZB，这相当于 30 倍的数据量增长，大约平均每两年翻一番^[1]。

数据量的增加给企业带来了许多挑战。存储空间不能满足数据增长的需求，只能通过不断地增加存储设备来解决。与此同时，几乎所有的数据中心都通过网络提供服务。数据量的膨胀导致数据在网络中的传输时间变长，网络变的更加拥堵，同时也更加耗费服务器的处理资源。

重复数据删除是一种减少数据所需存储空间的技术^[2]。在大多数的存储系统包含许多数据片段的重复副本。例如，相同的文件可以由不同的用户保存在几个不同的位置，或者不相同的两个或更多个文件可能仍然包括许多相同的数据。重复数据删除通过仅保存一个数据副本并用指向其他副本的指针替换原始副本来消除这些额外的副本。重复数据删除可用于灾备系统和归档系统，也可以用于主存储系统。

在 NetApp 的报告中提到，在某些情况下，重复数据删除可以将存储所需要的空间要求降低高达 95%^[3]，这与存储的文件类型的共享的数量相关。消除额外的数据副本不仅节省了直接磁盘硬件成本，而且节省了相关成本，例如电力，冷却，维护，占地面积等，还可以降低数据传输过程对于网络带宽需求^[4]。

随着存储系统上的数据量的不断增大，重复数据删除带来的对于存储空间和网络带宽的节省也愈加明显，但同时也存在一些可扩展性方面的挑战^[5]。如何提高重删率以节省更多的存储空间，减少带来的处理器和内存等额外开销，成为当前对于重复数据删除研究关注的重点。

重复数据删除技术给存储系统带来的好处主要有以下三个方面：

- (1) 可以有效控制数据的急剧增长。
- (2) 增加有效存储空间，提高存储效率，从而节省空间、电力供应、冷却等

运维成本，节省存储总成本和管理成本^[6]。

(3) 节省数据传输的网络带宽，提高备份的速度和系统可靠性等。

重复数据删除的基本原理如图 1 所示，图中左边有 4 个去重前的文件。将这几个文件分别切分成多个块，并用同一个字母标识相同的数据块，可以看到在文件中存在相同的数据块。图中右边为存储设备上存放的去重后的数据。要写入文件的某个块时，首先会判断这个块是否已经在存储设备上存在。如果不存在，则写入存储设备，并将引用率置 1。如果要写入的块已经在存储设备上存在，则将已存在的块的引用率增 1，并将文件的块指针指向该块，而不是将数据块再写入存储设备一次。这样做使得文件中相同的数据块只在存储设备上保存一份，从而达到节省存储空间的目的。

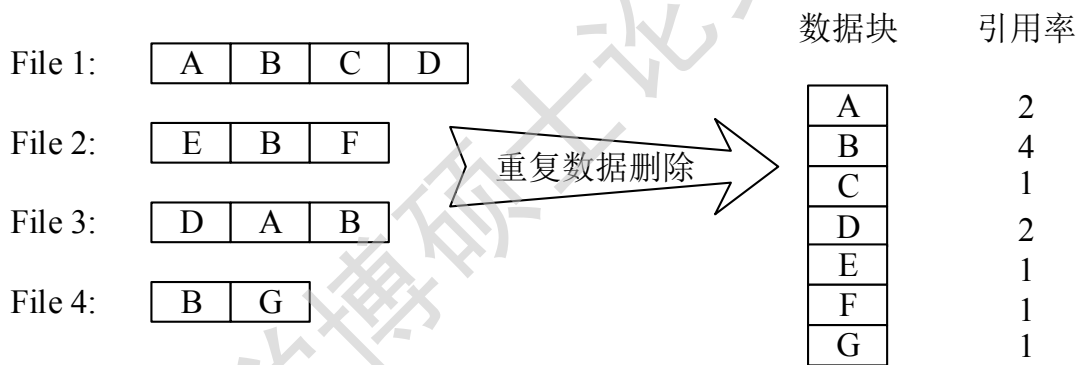


图 1. 重复数据删除技术的基本原理

随着重复数据删除技术的推广，研究人员发现重删技术在减少数据占用的存储空间，节省了成本的同时带来了一定的额外开销。其中最主要的开销包括两个方面：

(1) 元数据的管理开销。数据判重是重删的关键步骤之一，它主要是在当前系统上已有的指纹库中寻找是否存在要写入数据块的指纹信息。随着重删系统中的数据量的增大，元数据的量也不断增大。当元数据的大小超过内存大小时，频繁的元数据访问会造成大量的 I/O 请求，从而使系统吞吐量急剧下降。

(2) 数据块的碎片化问题。由于存储设备上非顺序 I/O 的性能低下，各种存储系统都在设法尽量减少数据的碎片化。由于重删系统要利用系统上已经存在的

非冗余数据来减少数据的写入量,这样就导致了连续的存储介质访问变为随机访问,并且数据块的分散程度随着数据量的增大愈发严重,有可能使一个读 I/O 请求分散在存储介质的多个区域造成写放大。

以上两个开销在一定程度上限制了主存储系统上的重删系统的数据规模和可扩展性。如果重删技术在带来数据定入量减少的同时尽可能地减少系统资源的使用,那么该技术可运用的范围也将更加广泛。

智能手机近年来发展迅速,手机的存储子系统性能得到了大量的关注。由于手机存储一般基于 Flash 的 eMMC 芯片,存储写入的数据量和垃圾回收的触发次数成为影响存储性能的关键因素。重删技术可以减少写入数据量以达到性能提升的效果,写数据的减少导致垃圾回收的次数减少,进一步提升了智能手机上的存储系统的性能。

然而,由于手机的处理器和内存资源有限,对于重删带来的额外开销的限制更加严格。如果能在不占用过多系统资源的情况下应用重删系统,那重删技术对手机存储系统带来的性能提升将更加显著。

重删技术是一把双刃剑。不管是在主存储系统还是手机存储上,重删均能带来写入数据量的减少以提升性能,但是该技术需要额外的系统开销用于去重的过程。为了使其带来的存储资源节省和性能提升最大化,有必要对重删的过程进行性能优化,减少其对系统资源的使用,提高重删系统的吞吐量。

1.2 重复数据删除技术的研究现状

目前,重复数据删除技术已被大量应用于灾备系统和归档系统中,并逐渐开始应用于虚拟机、邮件服务器、云存储等环境中^[7]。很多著名的计算机存储厂商,如 EMC、NetApp、IBM、HP、华赛、华为、中兴等,都投入了大量的人力和物力研发并推广重复数据删除技术的应用,许多相关的产品也已经得到终端用户的认可和使用^{[8][9][10][11]}。除了在工业界得到了大量关注,在学术界中也有大量的科研机构对重删进行研究,如美国的普林斯顿大学^{[12][13]}、明尼苏达大学^[14]和加州大学圣克鲁兹分校^[15]等,国内的华中科技大学^{[16][17][18][19]}、清华大学^{[6][20]}、中国科学院^[21]和南开大学^[22]等,也已经在重复数据删除技术的研究中取得了相当丰

富的研究成果。

对重删的研究主要集中在如何更快地执行重删过程,如何更快地重构或读取数据,以及如何探测和发现更多重复数据。

关于如何更快地执行重复数据删除过程,研究人员主要致力于解决哈希索引的磁盘瓶颈问题。例如,Data Domain公司提出的DDFS^[12]利用摘要向量等多项技术减轻哈希索引引起的性能问题。加州大学圣克鲁兹分校和HP实验室共同提出的稀疏索引^[23]利用样本和分层技术进一步减少了索引表在磁盘中的查找开销。SiLo^[24]采用另一种思路,它充分挖掘了备份数据流中的相似性和位置语义,在更少的内存下实现了更高的吞吐量。SiLo减少了主节点的内存空间索引,然后通过挖掘局部语义增强了重复检测过程中的概率相似性检测。AA-Dedupe^[25]是一个应用感知的源端重删方案,它将个人计算机上的所有数据基于不同的分块方式和哈希计算方法分成多个子集,每个子集在重删过程中采用与其包含数据最合适的分块方式和哈希计算方法。加州大学圣克鲁兹分校和IBM共同研究的HANDS系统^[26]利用启发式的预取策略,根据负载特征变化从磁盘中预取指纹信息到内存中,以减少指纹查找延迟带来的系统开销,使得重删技术在主存储系统中更好地被使用。

以上方法都将索引表保存在磁盘上,但是明尼苏达大学和微软研究院共同提出的ChunkStash^[14]以及帕德博恩并行计算中心提出的Dedupv1^[27]改变了这种存放方式,利用闪存固态盘的读性能优势将哈希索引表和元数据存放在固态盘上,显著提高了哈希索引表的查找效率,进而提高了重复数据删除的吞吐率。然而,利用闪存作为辅助的方案会增加重删的硬件成本。此外,南开大学的研究人员利用低能耗的协处理器加快指纹计算的过程来提高重删的性能^[22]。

关于如何更快地重构或读取数据,目前的解决方案大都是通过保持数据在物理设备上的顺序性,尽量降低数据在物理设备上的分散程度来保证数据的重构性能不会降低太大^{[28][29][30]}。iDedup^[28]是一种面向容量的重删方案,主要面向主存储系统。它利用数据写入的空间局部性,对于非连续的冗余数据块不进行重删。Lillibridge等人提出了同时使用增加缓存大小、容器封装、使用正向的装配区域^[31],以减轻备份应用的数据块的碎片化问题。HAR(History-Aware Rewriting)^{[32][33]}

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士论文摘要库