Efficient Transaction Recovery on Flash Disks

ON Sai Tung

A thesis submitted in partial fulfillment of the requirements
for the degree of
Master of Philosophy

Principal Supervisor: Dr. Jianliang XU

Hong Kong Baptist University

June 2010
Abstract

Due to recent advances in semiconductor technologies, flash disks have been a competitive alternative to traditional magnetic disks as secondary storage. In database management systems (DBMSs), transaction recovery is one of the most important components, which enforces both atomicity and durability of transactions. In this thesis, we study how transaction recovery can be efficiently supported in DBMSs running on single-level-cell (SLC) flash disks.

We propose a new transaction recovery scheme, called flagcommit, to exploit the unique characteristics of SLC flash disks such as out-of-place updates and partial page programming. To minimize the need of writing log records, the main idea is to embed the transaction status into flash pages through a set of chained commit flags. Based on flagcommit, we develop two specific commit protocols, namely commit-based flag commit (CFC) and abort-based flag commit (AFC), to meet different performance needs. We also extend them to support a no-force buffer management policy and a fine-grained concurrency control mechanism.

Trace-driven simulations are conducted to evaluate the performance of the proposed CFC and AFC protocols. The results show that both protocols outperform the state-of-the-art flash-aware commit protocols in terms of various performance metrics.
# Table of Contents

Declaration i

Abstract ii

Acknowledgements iii

Table of Contents iv

List of Figures vii

1 Introduction 1

1.1 Overview .................................................. 1

1.2 Characteristics of Flash Disks .......................... 3

1.3 Transaction Recovery Schemes ........................ 4

1.4 Contributions of the Thesis ............................. 8

1.5 Outline of the Thesis ..................................... 9

2 Background and Related Work 10

2.1 Flash Translation Layer (FTL) ........................... 10
2.2 Cyclic Commit Schemes (SCC and BPCC) .................................... 11
2.3 Related Work on Flash-aware Data Management ...................... 15

3 The Flag-Commit Protocols ................................................. 17
  3.1 Overview ................................................................. 17
  3.2 Commit-based Flag Commit (CFC) .................................... 19
     3.2.1 Normal Execution ............................................... 21
     3.2.2 Garbage Collection ............................................. 24
     3.2.3 Recovery ......................................................... 25
  3.3 Abort-based Flag Commit (AFC) ...................................... 28
  3.4 A Discussion of CFC and AFC ....................................... 30
  3.5 Block-based Flag Technique .......................................... 32

4 Extensions of FLAGCOMMIT Protocols .................................. 36
  4.1 Supporting Buffering and No-Force Policy .......................... 36
  4.2 Supporting a Fine-grained Concurrency Control .................. 39

5 Performance Evaluation .................................................... 41
  5.1 Experimental Setup .................................................... 41
  5.2 Performance of Flag Commit Protocols .............................. 44
     5.2.1 Effectiveness of the Block-based Flag Technique ............ 44
     5.2.2 Comparison with Cyclic Commit Protocols ................. 45
     5.2.3 Impact of Transaction Abort Ratio ............................ 49
     5.2.4 Impact of Transaction Size .................................... 51
5.2.5 Performance of No-Force Buffer Extension ................. 53
5.2.6 Performance of Record-level Concurrency Extension ....... 53

6 Conclusion and Future Work ........................................... 57

Curriculum Vitae .......................................................... 63