Journal of Advanced Smart Convergence Vol.1 No.2 26-29 (2012)
DOI 10.7236/JASC2012.1.2.5

JASC 12-2-5

Recovery Methods in Main Memory DBMS

Jeong-Joon Kim[*], Jeong-Jin Kang[+], Ki-Young Lee[†]

[*]Department of Computer Science and Information Engineering, KonKuk University, Korea
[+]Department of Information and Communication, Dong Seoul University, Korea
[†]Department of Medical IT and Marketing, Eulji University, Korea

Abstract

Recently, to efficiently support the real-time requirements of RTLS(Real Time Location System) services, interest in the main memory DBMS is rising. In the main memory DBMS, because all data can be lost when the system failure happens, the recovery method is very important for the stability of the database. Especially, disk I/O in executing the log and the checkpoint becomes the bottleneck of letting down the total system performance. Therefore, it is urgently necessary to research about the recovery method to reduce disk I/O in the main memory DBMS. Therefore, In this paper, we analyzed existing log techniques and check point techniques and existing main memory DBMSs' recovery techniques for recovery techniques research for main memory DBMS.

Key words: RTLS, MMDBMS, Recovery, Log, Checkpoint

1. INTRODUCTION

The fast and efficient processing about complex spatial data has been urgently required in GIS applications such as RTLS(Real Time Location System) depending on development of geographic information system technology. To achieve this, A Study on Main memory DBMS about that entire database resides in main memory are currently underway. these main memory DBMS can lose the whole database by an unexpected glitch such as system error, transaction error, computer malfunction, etc[1-3].

Therefore, it is necessary to maintain log information to disk for backup of the database files and transaction processing for recovery. However, log records for recovery and incurred disk IO during the course of performing checkpoint are important factor to degrade the performance of the system. Therefore, we need the efficient recovery system which can reduce disk IO in order to maximize the system performance.

Current main memory DBMS occurs a lot of disk operations and disk space become less efficient. Because it is a way to save a copy of the data space to perform transactions at the checkpoint that has been updated by the dual disk. Thus, Recovery Techniques are very important to maintaining the consistency of main memory DBMS through managing transaction log while minimizing the overhead for disk IO in main memory DBMS. And Also, Changed data by transaction perform an effective checkpoint to maintain consistency.

In this paper, we analyze the existing log technique, checkpoint technique and main memory DBMS recovery techniques for research of main memory DBMS recovery technique.

2. LOG TECHNIQUE

Record of the log is essential for recovery in main memory DBMS. But At this point, overhead about occurring disk IO is important factor to degrade performance of the system due to occupy the much part of entire transaction execution time. Existing research to solve that problem can be divided into two techniques; non-volatile memory is used as log storage device and the way to improve the performance of recording the log techniques.

The former technique is the way of electronic hardware solutions, non-volatile memory is used as log storage device to improve performance of system such as battery backup RAM, flash memory. However these techniques have to use

Manuscript received : Sept. 7, 2012 / revised : Nov. 1, 2012
[†]Corresponding Author: kylee@eulji.ac.kr
Tel: +82-10-7313-4689
Department of Medical IT and Marketing, Eulji University, Korea

extra hardware. And it has a problem that is dependent on specific hardware.

The latter technique is the way to improve the performance of recording the log techniques. There are three representative techniques; immediate commit, group commit, pre-commit.

Immediate commit stores changed data during transaction execution in main memory log buffer, log file is recorded all logs of the transaction on the disk before changing the partially committed to committed, and then transaction will complete. This technique has the advantage of simple structure and a short response time of individual transactions in comparison with group completion or pre-completion technique. But because the all log is recorded each transaction log on log file upon completion of transaction, this technique has a disadvantage that frequent disk IO.

Group completion technique is not recorded each transaction log information in the log file. Transaction will commit after all partially commit transaction log is recorded log file in batches saving the each transaction log information in log buffer. It can reduce the number of disk IO. This technique provides better performance than immediately commit technique because of reduction of the number of disk reflection about log buffer information. But, It has a disadvantage that degrade concurrency to maintain lock until partial commit of transaction will commit. One way of log techniques, it solve the degrade concurrency retaining lock of corresponding transaction's resource through transaction's commit is delayed up to specific point at group commit technique [4].

It released corresponding transaction's resource when transaction is partially commit status, not commit. First of all, it unlock without waiting for disk output the log buffer's contents and allow access to subsequent transaction's resource. And all transaction log about partially commit status in log buffer is recorded in log file in transaction order in batches. Then corresponding transaction commit.

The operations of transaction are processed very quickly in main memory DBMS, But it is required relatively long time to write the log on disk. Thus, if you release the lock before perform this task, transactions waiting for subsequent transaction's resource is able to escape from wait state and It should increase the concurrency.

As a prerequisite to use this technique, concurrency control technique should be used to ensure log's disk output that is completed in order after the start of commit task. For example, if you are using the strict 2PL(2-phase locking), dependence between transaction is maintained and even the resiliency about failure of the system can be guaranteed. Because it does not occur transaction commit than one's depending transaction in advance even if it apply pre-commit.

This technique is no change of the number of log buffer's disk output and the response time of individual transaction by comparing the group-commit. But, it reduce other transaction's lock waiting time waiting for a possessed lock by corresponding transaction and it has the advantage of improving concurrency of transaction performance than group commit[8].

3. CHECK POINT TECHNIQUE

There are representative three check point technique in main memory DBMS; transaction consistent check point, action consistent check point, fuzzy check point. Transaction consistent checkpoint perform check point when performing transaction all commit[5].

This technique, when check point occurs, first all preceding transaction will commit and creation of subsequent transaction is stopped. if preceding transaction is committed, all the actual check point transaction performs and if perform of check point is completed, creation and perform of subsequent transaction is allowed. This technique is maintained transactional consistency on database using check point. In recovery status, there is no need to undo about performed transaction after current check point, it can be restored due to redo about commit transaction's log. However, it has a disadvantage of performance degradation because create, perform of subsequent transaction is stopped during performing check point.

Action consistent check point technique performs check point when action is completed, because transaction consider as continuous act of data base. While performing a check point, all the update operations on the database need to stop and UNDO log and REDO log should be recorded in this technique.

This technique is synchronized with performing of transaction such as Transaction Consistent Check Point, it has a disadvantage of performance degradation because create, perform of subsequent transaction is stopped during performing check point. Fuzzy Check point technique is performed check point asynchronously with other transaction by creating process for perform of check point when check point occurs. Fuzzy check point has superior performance than other checkpoint techniques. Because it proceeds regardless of transaction perform[6].

Fuzzy Checkpoint technique is difficult to data base consistency during performing of check point. It can be performed simultaneously renew operation transaction and check point about the same main memory data base's page. Thus, when system error occurs during performing of check point, database file on disk will not be able to use recovery.
Many systems is used ping-pong update that have two database files on disk to perform the checkpoint alternately[5,6]. Looking at you through the process of performing this technique, as shown in Figure 1, when the

checkpoint occurs, create a process for check point. After then, check the value of check point position recorded in anchor file, check point performs in corresponding data base file. At that time, check point is performed with other transaction asynchronously.

The performing of check point is completed when it is all reflected in the database file about changed data in main memory data base. At that time, the value of check point position in anchor file is changed information about other database file to be used in performing the next check point.

This technique can be performed simultaneously update operation transaction and check point about page of same main memory data base. Thus, If a system error occurs during performing of check point, recovery is performed a different data base files.

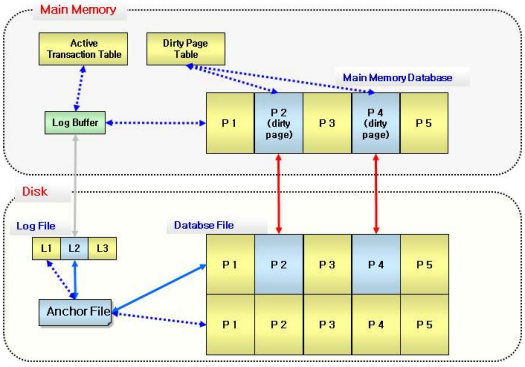

Fig. 1. The Structure of Fuzzy Check Point

Update page of main memory database must be all recorded two database file in fuzzy check point based on ping-pong update. This technique has a disadvantage of performance degradation to recovery system because twice disk IO on the same page. Also, it need additional space to maintain two database file, occurs management maintenance cost for them.

4. Existing Main Memory Recovery Technique

This chapter discusses the existing recovery techniques in main memory DBMS.

A. Dali

Dali is main memory DBMS created at AT&T Bell Laboratories[7]. REDO log and UNDO log is managed by each transaction in order to reduce performance degradation caused by contention between transactions. In transaction is partially commit status, it used pre-commit that transaction REDO log in system global log buffer is moved and it is unlocked before record log on dick. In Dali, it has a disadvantage of performance degradation using act consistent checkpoint technique[8].

B. Mr.RT

Mr.RT is main memory DBMS created by ETRI in priority[9]. This system uses ECBH and T-tree as the indexing system for efficient access to data that resides in the main memory. It is applied bach configuration to configure the new index at system restart. And it used 2PL-PI based static 2PL for concurrency control. Recovery system is used fuzzy check point, tuple logging technique that is record log related tuple in order to reduce amount of log.

In addition, log buffer is used NYSIMM of stable memory in order to improve the real-time processing. UNDO logs and REDO logs are managed separately to each transaction unit, it reflects only REDO log to disk when transaction is committed like Dali's recovery technique.

C. System M

System Mis main memory testbed created at university of Princeton. This system consists message server to simulate the application, transaction server to transaction management, log server and check point server to recovery management, lock server to concurrency control. Database is being loaded in shared virtual memory, log server and check point server provides a diversity of options to seek appropriate recovery algorithm on main memory structure.

D. Altibase

Altibase is main memory DBMS created at domestic Altibase company, is shown the latest version of 4.3.1. And Altibase currently has No.1 domestic market share in main memory DBMS. This system solved the problem that is limited space on existing main memory through integrating main memory DBMS to disk DBMS by Hybrid method. Also, it provides log technique based on WAL, support fuzzy check point based on update ping-pong for recovery.

E. TimesTen

TimesTen is main memory DBMS created at TimesTen company, is shown the latest version of 5.1. And TimesTen currently has No.1 global market share in main memory DBMS. For the recovery, it provides immediately commit and group commit and supports transaction consistent check point and update ping-pong technique[10,11].

5. Conclusion

Efficient recovery technique is important to stability of database in main memory DBMS Because all stored data in main memory may lose when unexpected situation occurs such as computer failure, system crash, system error, etc.

In this paper, we analyzed existing log technique, check point and recovery technique of existing main memory DBMS for research of recovery technique on main memory DBMS. Results of the analysis of these previous studies, in the future, we should be studied effective recovery technique to solve the problem that is duplicated check point processing

on the same page and waste of space in fuzzy check point based update ping-pong which is using the majority of main memory DBMS.

References

[1] K. Y. Lee, M. J. Lim, J. J. Kim, K. H. Kim, J. L. Kim, "Design and Implementation of a Data Management System for Mobile Spatio-Temporal Query," *Journal of the Institute of Webcasting, Internet and Telecommunication(IWIT)*, Vol. 11, No. 1, pp.109-113, 2011.

[2] J. Park, G. Choi, "A Embedded System Technology for Web based monitoring and control system," *Journal of the Institute of Webcasting, Internet and Telecommunication(IWIT)*, Vol. 9, No. 1, pp.61-69, 2009.

[3] Y. G. Jung, J. K. Choi, Y. J. Choi, "Design and Implementation of Medical Education System using Mesh-Up Meta-Search Program," *Journal of the Institute of Webcasting, Internet and Telecommunication(IWIT)*, Vol. 11, No. 4, pp.119-123, 2011.

[4] Lehman, T. J., and Carey, M. J., "A Recovery Algorithm for A High Performance Memory-Resident Database System," *Proc. ACM SIGMOD Conf.*, Vol.16, No.3, pp.104-117, 1987.

[5] Young-Sik Kwon, Yang-Sae Moon, Ki-Woong Cang, and Kyu-Young Whang, "Design and Implementation of a Recovery Technique based on Nonstop Transaction - Consistent Check pointing in a Main Memory Storage System," *SIGDB-KDBC, Data-Base Task Group,* pp.232-239, 2003.

[6] Salem, K., and Garcia-Molina, H., "System M : A Transaction Processing Testbed for Memory Resident Data," *IEEE Transactions on Knowledge and Data Engineering,* Vol.2, No.1, pp.161-172, 1990.

[7] Bohannon, P., Daniel, F., Rajeev, R., Abraham, S., Seshadri, S., and Sudarshan, S., "The Architecture of the Dali Main-Memory Storage Manager," *The Journal of Multimedia Tools and Applications*, Vol.4, No.2, pp.115-151, 1997.

[8] Eich, M., "MARS : The Design of a Main Memory Database Machine," *Proc. Int. Workshop on Database Machines*, pp.251-268, 1987.

[9] Kyung-Mo Lee, Jung-Ouk Lim, Gyoung-Bae Kim, Kwang-Chul Jung, Soon-jo Lee, Jin-Go Kim,Hae-Young Bae, "Mr .RT 3.0 : Main - Memory Resident Real - Time Database System for High Performance of Real - Time Transactions," *Fall Congress of Korean Institute of Information Scientists and Engineers*, Vol.25, No.2, pp.208-210, 1998.

[10] Margaret, H.D., Jun-Lin, L., and Xi, L., Fuzzy Checkpointing Alternatives for Main Memory Databases, *Recovery Mechanisms in Database Systems, Prentice-Hall, Inc.,* 1997.

[11] The TimesTen Team, "In-Memory Data Management for Consumer Transactions The TimesTen Approach," *Proc. of the ACM SIGMOD Int. Conf. on Management of data,* 1999.

Jeong-Joon Kim received his B.S. and M.S. degrees in Computer Science at Konkuk University in 2003 and 2005, respectively. In 2010, he received his Ph.D. degree in Databases at Konkuk University. He is currently a professor at the department of Computer Science at Konkuk University. His research interests include database systems, geographic information systems(GIS), location-based systems(LBS), main memory databases, embedded databases, and wireless sensor networks (WSN) etc.

Jeong-Jin Kang is currently the faculty of the Department of Information and Communication at Dong Seoul University in SeongNam, Korea since 1991, and currently the President of the Institute of Webcasting, Internet and Telecommunication(IWIT). During 3 years from Feb. 2007 to Feb. 2010, he worked as a Visiting Professor at the Department of Electrical and Computer Engineering, The Michigan State University. He was a lecturer of the Department of Electronic Engineering at (Under)Graduate School(1991-2005), The Konkuk University. Dr. Kang is a member of the IEEE Antennas and Propagation Society(IEEE AP-S), the IEEE Microwave Theory and Techniques Society(IEEE MTT-S), and a member of the Institute of Webcasting, Internet and Telecommunication(IWIT), Korea. His research interests involve Smart Mobile Electronics, RF Mobile Communication, Smart Convergence of Science and Technology, RFID/USN, u-Healthcare and ultrafast microwave photonics, as well as GIS, LBS, moving objects databases, and telematics etc.

Ki-Young Lee received his B.S. degree in Computer Science at Soongsil University in 1984. In 1988 and 2005, he received M.S. and Ph.D. degrees in Databases at Konkuk University, respectively. From 1984 until 1991, he worked for Korea Institute of Ocean Science & Technology(KIOST) as a researcher in Data Information & Processing department. He is currently a professor at the department of Medical IT and Marketing at Eulji University. He was also the head of department of S/W development in Bio-Meditech Regional Innovation Center at Eulji University. He is the director of the Korea Webcasting, Internet & Telecommunication(IWIT). His research interests include spatial databases, geographic information systems(GIS), location-based services(LBS), u-Healthcare, ubiquitous sensor network(USN), moving objects databases, and telematics etc.