Journal of Web Systems and Applications (2017) Vol. 1, Numuber 1
Clausius Scientific Press, Canada

Distributed Database Integrated Transaction Processing Technology Research

Zhi-yong Liu[1], Qiao SUN[1], Shao-wei Zhang[1], Xu-bin Pei[2], Lan-mei FU[1], Jia-song SUN[3, a*]

[1] Beijing GuoDianTong Network Technology Co., Ltd., Beijing, China

[2] State Grid Zhejiang Electric Power Company, Hangzhou, China

[3] E. E. Department, Tsinghua University, Beijing, China

[a]email: littlesmart@yeah.net

*Corresponding author

Keywords: Distributed Database, Integrated Transaction Processing, Comprehensive Distributed Database Management, Optimistic concurrency control.

Abstract. In this paper, we propose one comprehensive distributed database transaction method for InfiniBand transaction integrity in the distributed heterogeneous environment caused by data and application expansion. The method first integrates the functions of the standard online transaction engine and then uses the optimistic concurrency control technique to deal with the online memory allocation problem. On the one hand, it reduces the delay time of the online transaction and on the other hand, allocates the computing cost of the transaction to the server and the client. In the existing enterprise network environment using YCSB as a benchmark test shows that the different keys under the method of response speed are the fastest. The result shows that this kind of hybrid processing scheme can make full use of the advantages of remote direct memory access and fast network, which can further improve the consistency and reliability of the distributed database system.

1. Introduction

Distributed database system refers to the data is physically dispersed and logically centralized database system. Distributed database system uses the computer network to manage the geographical location and decentralized need to control different levels of concentration of multiple logical units (usually centralized database system) constitute a unified database system. Distributed database with the distribution and logical coordination. A transaction symbolizes a unit of work performed within a database management system (or similar system) against a database, and treated in a coherent and reliable way independent of other transactions. A transaction generally represents any change in database. Transactions in a database environment have two main purposes [1,2].

Big data technology was first introduced into the telecom industry for marketing and decisions-making purposes. From a database perspective, such workload is characterized as OLAP (On-Line Analytical Processing) tasks [3,4]. However, today's big data technology has

greatly evolved to a level where general staff are beginning to use big data to increase productivity in daily routine work. These new applications not only contain high analysis workload, but also involve in heavy OLTP (On-Line Transactional Processing) tasks. Examples are batch inserts of newly constructed road records into a geographic information table, or updates of a certain customer's personal data record. The InfiniBand Architecture is an industry standard interconnection technology which aims to provide low-latency and high-bandwidth communication [5]. Current generation InfiniBand products, from Mellanox provide low latency and high bandwidth.

In this paper, an integrated network distributed database transaction processing method based on InfiniBand is proposed. The classic OLTP engine is combined with InfiniBand's RDMA technology for OLTP design to take full advantage of the low latency of RDMA and fast network. This kind of comprehensive processing can effectively control the delay in the database transaction, and further enhance the robustness and efficient response of the distributed database system. In our internal database system test our program, the experiment shows that our framework than the traditional OLTP advantage is that it can provide more efficient and stable distributed database transaction processing.

The structure of this paper is as follows: Firstly, we introduce the architecture of InfiniBand RDMA, and then give the basic principle of OLTP. The fourth part presents the integrated distributed database network transaction processing. Finally, we give the summary and prospect.

2. The Architecture of InfiniBand RDMA

The InfiniBand Architecture Specification describes a first order interconnect technology for interconnecting processor nodes and I/O nodes to form a system area network. The architecture shown in Fig.1and Fig.2 is independent of the host operating system (OS) and processor platform [6,7]. However, due to their complexity in hardware implementation and non-transparency to the remote side, send/receive operations do not perform as well as RDMA operations in current InfiniBand platforms. Thus these designs have not achieved the best performance for small data messages and control messages [8, 9].

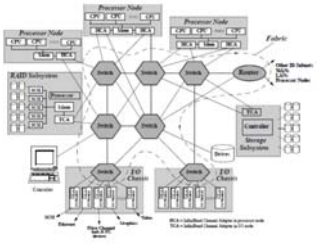

Figure 1 InfiniBand System Area Network Structure

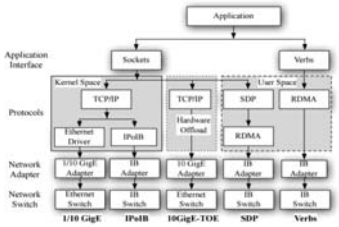

Figure 2 Networking layers of InfiniBand RDMA

3. Basic principle of OLTP

Relational database systems have been the backbone of business applications for more than 20 years. A typical OLTP environment consists of a number of users managing their transactions via a terminal or a desktop computer connected to a database management system (DBMS) via a local area network or through the Web. An OLTP system is thus a typically client-server system or a multi-tier system. In a simplified approach, the server is composed by three main components: the hardware platform (including the disk subsystem), the operating system, and the transactional engine. Most of the transactional systems available today use a DBMS as transactional engine, which is in practice the main component of any OLTP systems, assuring not only the transactional properties but also the

recovery mechanisms. Dependability is an integrative concept that includes the following attributes: Availability, Reliability, Safety, Confidentiality, Integrity and Maintainability. The importance of generating real-time business intelligence is that it is a building block to achieve better business process management and true business process optimization [10,11].

4. Integrated distributed database network transactions processing

In this paper, we propose one OLTP distributed database transaction scheme based on InfiniBand with optimistic concurrency controlling and make use of RDMA and fast network's low latency as far as possible. RDMA's architecture is neither shared memory nor shared. This new architecture requires a fundamental rethinking of the design of the database. By integrating the functionality provided by the OLTP engine and using optimistic concurrency controlling for OLTP design, it will take full advantage of the low latency advantages of RDMA and fast networks. The advantages of this method are: Reduced transaction latency; transaction burden sharing lead to a more balanced system. The detailed method is as follows: Our optimistic concurrency control algorithm differs from the traditional method in that it extends the traditional OCC to compute transactions with distributed delays in its commit time and recover the timestamp of the transactions before the transaction is executed, and solves the problem of highly competitive workloads difficult. On the other hand, in the workload by using more parallel computing, reduce the dependence on the transaction to improve the ability of the transaction. To overcome the scalability bottleneck caused by redundant serial execution points, the OLTP analysis engine in this article uses multivariate concurrency control to improve the transaction recovery performance in order to make the OCC available to a wider range of OLTP workloads with the high cost of aborting and restarting the contention. The transaction recovery in this article attempts to reuse the execution results without restarting invalid transactions from the beginning and with low latency transactions from multiple concurrent connections that have high data insertion rates that cannot be addressed by disk processing. This kind of OLTP helps to improve the efficiency of transaction of whole cycle. In this paper, the precompiled collections are compiled by optimizing compilation and OLTP in memory to have the advantage that they can be invoked by different applications to speed up transactional efficiency.

5. Experiments and test results

In order to test the performance of the above methods, we have in their own enterprise network environment for the test platform to build and verify. The specific configuration is: 10 IBM-x3650M4-2U server (2CPU, 6-core 12-thread Xeon E5-2620, 64G memory, PCIe-2 bus, Mellanox HCA, 2T hard disk), one for the central data node, the other server as a data node. Each server has 32K of first-order instruction and data cache and 1238K of L2 cache. The operating system is Windows8 and InfiniBand configured.

The Yahoo! Cloud Serving Benchmark(YSTB) is used to benchmark multiple systems and compare them. YCSB has done a lot of optimization to improve client performance, such as the use of data types in the most primitive array of bits to reduce the data object itself to create the necessary time to convert and so on. YCSB several major features: Support for common database read and write operations, such as insert, modify, delete and read; Multi-thread support. YCSB with Java, a very good multi-threaded support; Flexible definition of scene files. You can specify test scenarios flexibly, such as 100% insert, 50% read 50% write, etc.; Data request distribution: support for random (only a small part of the data to access most of the request) and the latest data several request distributions; Scalability: You can extend the way Workload to modify or extend the functionality of YCSB.

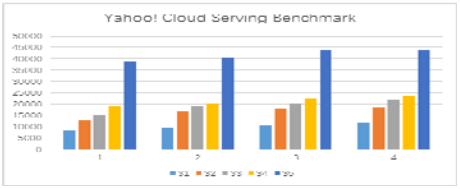

Figure 3(a) S1-S5 test results of YCSB benchmark on database I

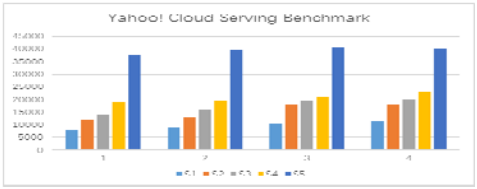

Figure 3(b) S1-S5 test results of YCSB benchmark on database II

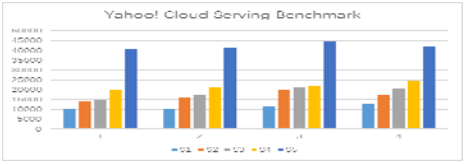

Figure 3(c) S1-S5 test results of YCSB benchmark on database III

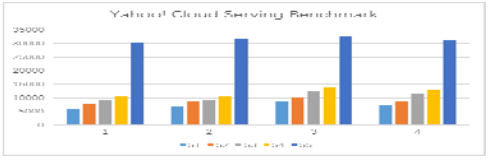

Figure 3(d) S1-S5 test results of YCSB benchmark on database IV

As YCSB itself will take a lot of work, so the deployment of YCSB on a separate machine. Both the YCSB and the database server guarantee Gigabit bandwidth. We set up the separate directories of our original database to the database I, II and III to complete the five tests of YCSB S1-S5. Figure 3 (a)-(d) show all the test results. As can be seen from the table, this paper presents a comprehensive distributed database network transaction processing methods in the four database test results are better. The best results are obtained in database 3, the lowest results are obtained in database 4, and it is clear that the influence of different data sets is obvious.

6. Conclusions

Under the poor network conditions, RDMA-based OLTP method had a high computational cost. We propose one comprehensive distributed database transaction method, which adopted InfiniBand transaction integrity in the distributed heterogeneous environment. The experimental tests show that our methods had the fastest response speed. It means our hybrid processing scheme can be promoted with using the mix of remote direct memory access and fast network. Next, we will further improve the consistency and reliability of the distributed database system.

Acknowledgements

This research was financially supported by Science and Technology Project of the State Grid Corporation of China (SGZJ0000BGJS1500433) and the State Grid Information & Telecommunication Group CO., LTD. (SGITG-KJ-JSKF [2015]0003).

References

[1] Research on the Synchronization Technology of Distributed Database System[D]. Master's thesis, Changchun University of Science and Technology, Apr. 2008.

[2] Database transaction on https://en.wikipedia.org/wiki/Database_transaction

[3] Lu X, Su F, Liu H, et al. A unified OLAP/OLTP big data processing framework in telecom industry[C]. Communications and Information Technologies (ISCIT), 2016 16th International Symposium on. IEEE, 2016: 290-295.

[4] NAGASURENDRAN P. A Novel Perspective of Data Warehousing and Data Mining in Education[J]. Global Journal for Research Analysis, 2017, 5(9).

[5] Sur S, Koop M J, Chai L, et al. Performance analysis and evaluation of Mellanox ConnectX InfiniBand architecture with multi-core platforms[C].15th Annual IEEE Symposium on High-Performance Interconnects (HOTI 2007). IEEE, 2007: 125-134.

[6] Shanley T, Winkles J. InfiniBand Network Architecture[M]. Addison-Wesley Professional, 2003.

[7] InfiniBand Trade Association. InfiniBand Architecture Specification: Release 1.0[M]. InfiniBand Trade Association, 2000.

[8] Liu J, Wu J, Panda D K. High performance RDMA-based MPI implementation over InfiniBand[J]. International Journal of Parallel Programming, 2004, 32(3): 167-198.

[9] Islam N S, Rahman M W, Jose J, et al. High performance RDMA-based design of HDFS over InfiniBand[C]. Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis. IEEE Computer Society Press, 2012: 35.

[10] Plattner H. A common database approach for OLTP and OLAP using an in-memory column database[C]. Proceedings of the 2009 ACM SIGMOD International Conference on Management of data. ACM, 2009: 1-2.

[11] Vieira M, Madeira H. A dependability benchmark for OLTP application environments[C]. Proceedings of the 29th international conference on Very large data bases. Volume 29. VLDB Endowment, 2003: 742-753.