

CUBRID RDBMS

ABHIJIT THAKRE¹

¹School of Informatics and Computing, Bloomington, IN 47408, U.S.A.

²Mechanical Engineer, Nagpur University, 2003

*Corresponding authors: abhijit.thakre@gmail.com

project-000, May 4, 2017

With advanced techniques of data mining and analysis, bigdata processing has become a key in today's world. Many of the bigdata processing uses NOSQL data for storing. However in order to avail the ACID behavior of database, the focus is again back to the RDBMS databases. This paper focuses on one the similar ORDBMS CUBRID. It also highlight the architecture of CUBRID with it key component and features provided.

© 2017 <https://creativecommons.org/licenses/>. The authors verify that the text is not plagiarized.

Keywords: Cloud, I524

<https://github.com/cloudmesh/sp17-i524/blob/master/paper1/S17-IO-3021/report.pdf>

1. INTRODUCTION

CUBRID is open source RDBMS with object support developed by Navel corporation. Developed in C language CUBRID provides key features like scalability, high availability, higher performance, online and incremental backups. CUBRID is distributed under GNU general public license for the database server engine and BCD license for API and client tool.

2. ARCHITECTURE

CUBRID has distinguished 3 layer architecture. It consists of Database server, connection broker and application layer.

2.1. Database Server

It is the core component of the CUBRID Database Management System. The main function for the database server are as below

- Saving and managing the data.
- Processing of the queries from user.
- Providing smooth functioning for multiple users.

2.2. Broker

It acts as middleware between the Database Server and GUI application to provide seamless experience. It provides functions

- Connection pooling.
- Monitoring.
- Log tracing and analysis.

2.3. CUBRID Manager

It is a GUI tool that manages database and broker. It also provides the Query Editor for executing queries on database for users.

3. ARCHITECTURE DIAGRAM

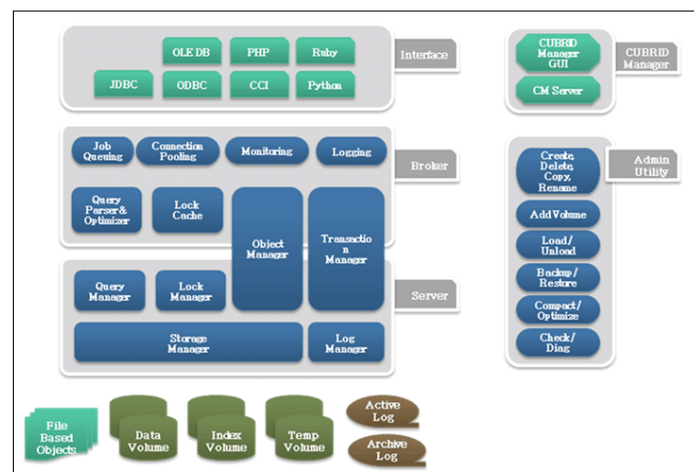


Fig. 1. [1]

4. DESCRIPTION

Database server and the broker work in co-ordination with each other as server and client respectively. The can be deployed on the different or same machine. The broker takes care of the queries from the users, it process it to the optimum level. On optimization it creates a query plan tree and sends the request to the server. The response from the server is via cursor navigation which is further returned to the user.

The client caches object instances from the database to its memory to provide fast access to data by using the query execution results or directly by users/applications. In addition, it caches locks as well as objects from the server for concurrency control. The execution of triggers or methods specified by users or applications is also performed in the client module.

References: [1]

5. MODULE CONFIGURATION

The CUBRID client and server modules consist of the following components:

- Transaction Management Component.
- Server Storage Management Component.
- Client Storage Management Component.
- Object Management Component.
- Client-Server Communications.
- Thread Management.
- Query Processing.

6. AUTHENTICATION IN CUBRID

CUBRID provides two levels of authentication.

User needs to enter credentials to login to the Host Server. On first login user need to set the admin credentials.

User needs to login to each database in the host server to access the individual database.

References: [2]

7. KEY FEATURES

7.1. High Availability and Scalability

CUBRID uses heartbeat technology to provide automated accurate fail-over and fail-back features which makes the database continuously available. The server is available during the upgrades, replacement or even during the maintenance phase.

7.2. Database Sharding

CUBRID 8.4.3 provides free sharding feature where data can be divided on multiple instance. In addition to unlimited sharding it also provides the features like connection pooling and load balancing to all the shards.

7.3. Performance

CUBRID provides high performance to the users with feature like query caching, optimized algorithm for indexing, fast object access.

CUBRID performance for hotspot read and special character for web service using logging technique have applied to disk have resulted in increase in the overall performance.

Function based indexing, filtering indexing and index skip scan provides various features to user for increasing the performance

7.4. Reliability

CUBRID is highly reliable with features like online incremental backup and restore. It provides the access restriction based on userip and databaseid. Naver is one of the top web portal in Korea rely on cubrid database.

7.5. Language Support

It provides 90 percent support to sql language support.

8. CUBRID SHARD

CUBRID shard is sharding is RDBMS specially targeted to address the problem on processing bigdata. CUBRID shard distributes the user data on multiple server to store it. So for fetching the data specific to user it needs to pass key information about the shard in the query. Parsing the query and finding the shard both things are taken care by the broker and does not needs the additional layer. This helps in increasing the performance for big data.

9. PERFORMANCE BENCHMARKING

9.1. CUBRID Vs MySQL

Test was performed by one of the social networking site on new 8.4.0 version of cubrid database to understand the time taken by cubrid database for managing the queries of 6 million active users. The queries mostly involved the IN and UNION operator.

For executing the test both Mysql and CUBRID database was made to store 3 user group table with 1520000 rows each. Each of the db had given 40 thread load for 10 min. The systems used for both the database are of same configuration. The test was conducted for five different scenarios.

9.2. Analysis Results

In most realistic scenario CUBRID IN query have performed twice as fast as Mysql UNION query and 8 times as fast as the Mysql IN query.

It also came out from the test that the effect of Key limit have considerably decreased the IO operations.

9.3. NBD BenchMark

The NBD (NHN Internet Bulletin Board Application Database) Benchmark is important benchmark to measure of performance of a DBMS used for a Bulletin Board System (BBS) type of web services.

This benchmark test was conducted to compare the performance of CUBRID R1.1 DBMS with OSS DBMS D1, Commercial DBMS D1, and Commercial DBMS D2. Factors considered for the benchmark includes database size, workload increase and cache functionality.

9.4. Conclusion of NBD Benchmark

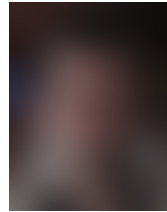
In the small and medium-sized database benchmarking, CUBRID is at second place after commercial DBMS D1. In the small database benchmarking, CUBRID performance can be improved if provided more work load. For the given workload it showed the usage of 30% CPU.

For medium size however it showed 80% of CPU usage. DBMS D2 shows the highest CPU usage of 100% in both small and medium-sized databases. References: [3]

REFERENCES

- [1] "CUBRID RDBMS," Web Page, Indiana University, Feb. 2017. [Online]. Available: <http://www.cubrid.org>
- [2] "CUBRID RDBMS," Web Page, Indiana University, Feb. 2017. [Online]. Available: <http://www.methodsandtools.com/tools/cubriddatabase.php>
- [3] "CUBRID RDBMS," Web Page, Indiana University, Feb. 2017. [Online]. Available: http://www.cubrid.org/cubrid_mysql_sns_benchmark_test

AUTHOR BIOGRAPHIES



Abhijit Thakre received his BE (Mechanical) in 2003 from The University of Nagpur.