

Research on Technology of DRDS Distributed Database

Fang Wang^{1, a *} and Guoyun Zhong^{2, b}

¹ East China Institute of Technology, Fuzhou, JX, China, 344000

² East China Institute of Technology, Fuzhou, JX, China, 344000

^afwang@ecit.edu.cn, ^bwtwfok@126.com

Keywords: compatibility; node; mass-data; Server.

Abstract. The full name of DRDS is Distributed Relational Data Service. It uses mysql protocol to interact with external system and it compatible with the most function of mysql. It looks like a mysql database service system, but it has favorable linear expansion and mass-data processing capacity. Furthermore the cost of update and expansion is low. This article will shallowly analyze and discuss its technology characters, system frame and performance data.

Introduction

With the information going deeply into the society, application system generates more and more huge data. So the tradition relation database's processing and computing capacity faces a huge challenge. Furthermore it's cost of system expansion and device update grows so more that it can't solve the problem by updating hardware device. But the relation database is the best method to process structure and relation data and it is the most accepted and low cost database technology. So currently it is need a system like the tradition relation database to process data, compute and save data. But when it processes mass-data, it has favorable expansion and function, the cost of system update and expansion is low and it has no effect in the course of expansion and update. So DRDS is a system to realize the above purpose.

Application scene of DRDS

All of the application used mysql as relation database can use DRDS system, but the current version has some restrict such as cross database join, strong consistency guarantee and so on. These characters will get support in the follow version. This version of DRDS is especially adapt to high concurrency, mass-data read and write, and some web system with low requirement in consistency.

Technology characters of DRDS

The kernel system module is based on Java, and it has technology characters as follows:

1. Open and inter- platform is that this system can be disposed in linux, unix, windows and so on. We recommend dispose DRDS in linux system.
2. High performance is that the whole system uses event driven working mode which has two aspects. The one is high performance network components which is researched and developed independently and based on NIO model's network communication layer. It is high efficient, steady, and has high concurrency processing character. The other one is high performance parser which is completely compatible with mysql5.5 grammar and it is equal to 10 times the performance of the general antlr or javacc generating parser.
3. Ease of use is that the system shield the complicated pool table logic and make it is convenient and quick to use and define these characters for ordinary users.
4. Linear expansion is that if system processing capability is found a bottleneck then it can achieve a prospective expansion result by purchasing the single data node and joining in existing server group.
5. High availability is that the system has perfect anomaly monitoring mechanism and effective

failover mechanism, so if some node encounter problems then it can quickly change server and cut down the application effect. So that the system itself availability can reach 99.99%.

Kernel value of DRDS

- l It can be compatible with mysql. More than 95% of mysql inquiry can directly run on DRDS. In most cases DRDS can be used as a single mysql. it advisably restrict functions to ensure that users can always enjoy linear horizontal expansion capability.
- l It can automatically run and maintain data. It add machine to cluster realize horizontal expansion and auto load balancing.
- l It can make manage easier. A command can create database, create table, increase or decrease field.

System frame of DRDS

The system frame of DRDS is shown in Fig.1.

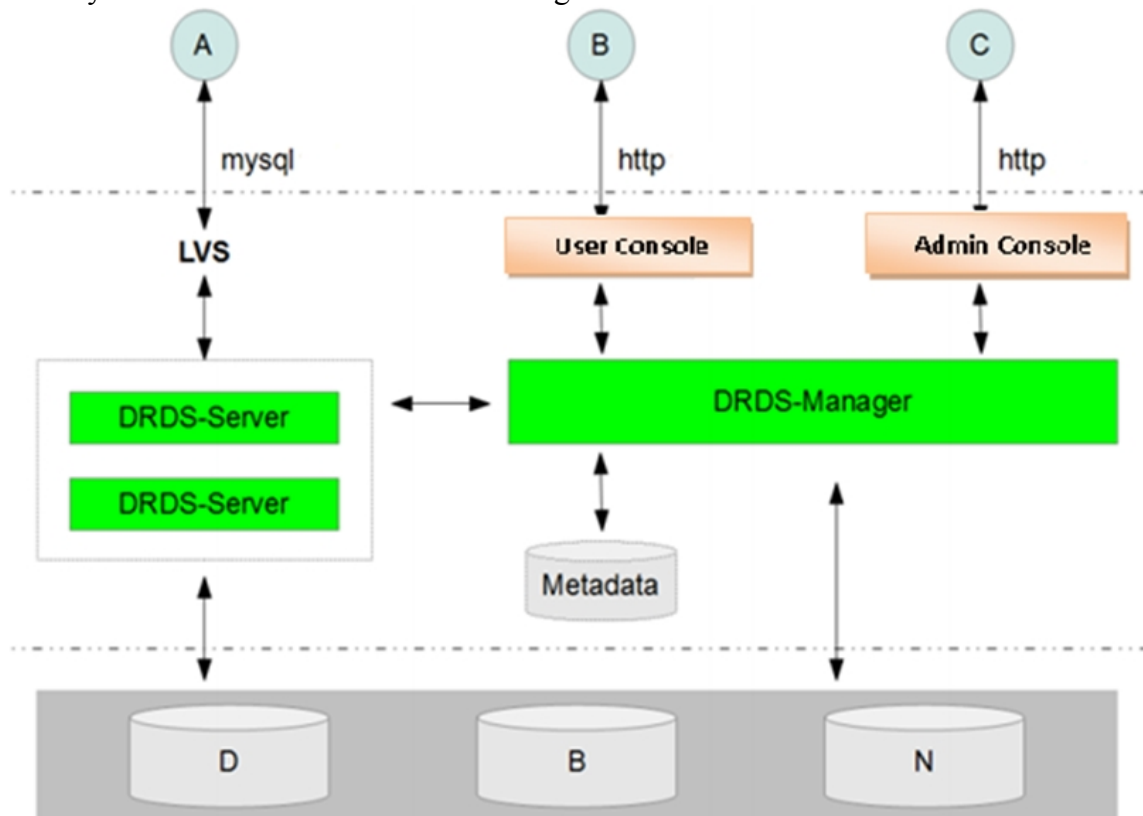


Fig.1 system frame

The current system is composed by these sub-systems as follows:

- l DRDS-Server is the kernel service of the system which provides outer service based on mysql protocol by LVS, and it is the only entrance of providing data service.
- l DRDS-Manager is the coordinating kernel of every sub-system. It makes this sub-system coordinate with each other to complete one object so that it is clear for every system in decoupling and duty.
- l User's Control is the users' managing controlling platform which is the entrance of communicating and managing by http protocol for users in the system.
- l Admin's Control is the system managing controlling platform which is the entrance of communicating and managing by http protocol for system administrators in the system.
- l Metadata is the data sets need to be long preserved such as relational data of strategy, structure and configure.
- l DBNode is the nodes to save data files and its data relation and distributes are decided by

DRDS-Server. Currently we use mysql as DBNode.

Performance data of DRDS service

Note: all the data is the average value tested in 30 minutes. For details, see the DRDS Server performance testing report shown in Table 1.

Table 1 performance data

	Scene	QPS	Latency
1	Select single library	91000	0.0018s
2	Select multi-library	63000	0.0025s
3	select multi-library limit	57000	0.0028s
4	select multi-library limit order by	47000	0.0034s
5	Select using hint	112000	0.0014s
6	Updater transaction single library	83500	0.0019s
7	update transaction multi-library	38500	0.0041s

Summary

Currently there are some restrictions and existing problems in DRDS as follows:

1. The version of Driver and client

DRDS Server is based on MySQL5.1 standard protocol. Through the above description, the users' commands can't be sent to node and run, but it is not possible to understand all commands for Server. So it may have protocol compatibility with common clients on Driver such as JDBC. Use Driver and client version tested by Ali's RDS development team.

2. Primary key and online migration

DRDS Server can automatically generate overall ID which can ensure the data whole uniqueness. When online migration turns out only the table with whole ID can migrate online currently in order to prevent primary key conflict problems after migration.

3. Database management and metadata

Manager uses meta database to save bottom database' metadata information such as table information, data distributed tactful information, field information, index information and so on. Users can update and send these metadata to Server when manage database by Manager. Then the parser and optimizer of Server select data nodes and merge data depending on these data information. So we must use UConsole to manage and maintain database. If users want to perform maintenance operation on RDS not Manager, so UConsole's RDS mode is needed.

4. SQL support issues

Because of system frame, optimizer, stream processing and so on DRDS Server can't support all function of SQL especially cross database SQL. Although the RDS team supports SQL as more as impossible basing users' need, it is a long process to complete support issues. So we must optimize the application design to prevent this issue.

Although there are so many problems and restricts in current version of DRDS, the explore team in Ali always improves and makes it perfect. So it will be better in future.

References

- [1] Jing Li, Research on DDB System Structure, [J], Intenet Fortune, (2009.5):197-198
- [2] Xun Shen, The Principle and the Practice of Ali Cloud Distributed Database (2013) on <http://wenku.baidu.com>.
- [3] Shixuan Sa, Database System Concepts, [M], Beijing: Higher Education Press(2000)

- [4] Xun Shen, Distributed Message Service Ons Principle and Practice of Ali (2015) on <http://wenku.baidu.com>
- [5] YANG Jing,LIU Tianshi,MA Gang,Data Fragment and Allocation in Distributed Database,[J], The Modern Electronic Technology (2006.9)
- [6] Qiuju Zhang, Research and Application on Data Storage Management Technology, [J], Information Technology and Standardization (2015.6)
- [7] Furong Zhong, Research on Sub-Storage System of Distributed Database, [J], Journal of Electronic Technology and Software Engineering (2014.7)