

万里分布式数据库管理系统 技术白皮书



北京万里开源软件有限公司

©版权 2021 北京万里开源软件有限公司

万里分布式数据库管理系统 是北京万里开源软件有限公司的注册商标。所有其它商标归其相应的所有者所有。

目录

1. 万里分布式数据库管理系统产品简介.....	1
2. 万里分布式数据库管理系统产品特性.....	3
简洁的架构.....	3
动态横向扩展.....	3
灵活的数据分布策略.....	3
分布式事务.....	4
完整的 SQL 标准支持.....	5
丰富的数据类型.....	6
高可靠与数据冗余.....	7
高性能数据服务.....	7
良好的上下游生态.....	8
完整的国产化支持.....	8
3. 数据库管理系统产品平台与指标.....	9
硬件环境支持.....	9
操作系统支持.....	9
技术指标.....	9
4. 开发接口支持.....	11
5. 竞争优势.....	12
6. 联系我们.....	13

1. 万里分布式数据库管理系统产品简介

万里分布式数据库管理系统是一款国产自主可控的分布式关系型数据库管理系统，能够应用于金融、电信、政府、互联网等对数据库事务处理性能要求较高的行业核心系统。

它采用 Shared-Nothing 的计算存储分离分布式架构的设计，具有极强的可扩展性，满足用户按需动态扩缩容的需求；基于经典的 paxos 一致性协议实现数据库多副本的强一致性，满足金融级数据高可用；数据自动均衡分布，兼容主流 SQL 标准，对应用开发完全透明，应用平滑迁移；采用了线程池、并行计算、计划下推等多种优化策略，深度适配多种平台，满足了高并发，低响应延迟的场景要求。

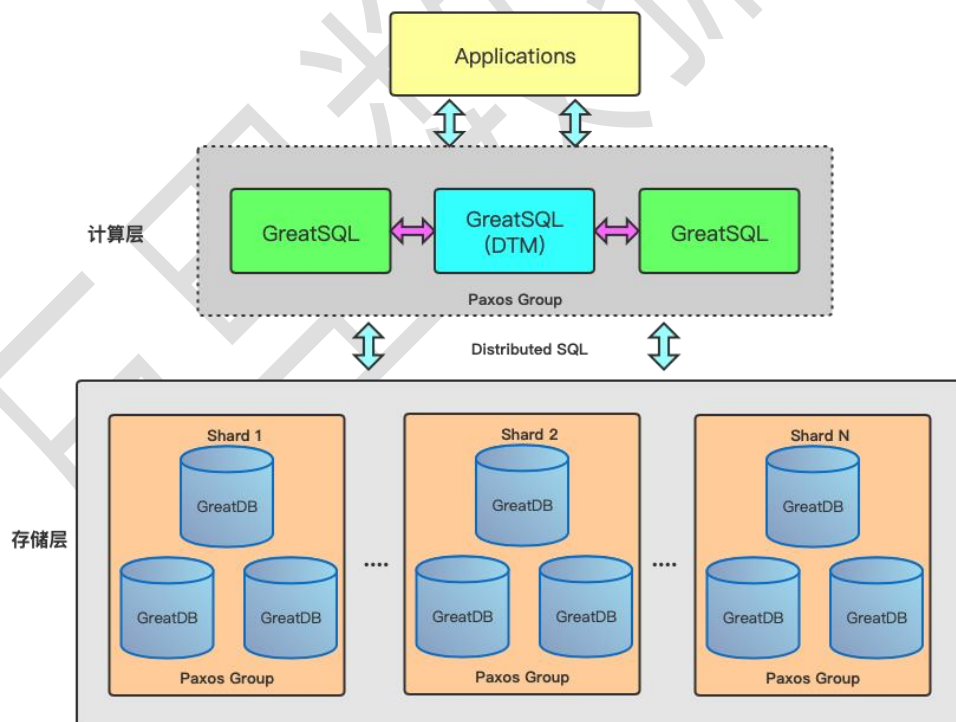


图 1 万里分布式数据库管理系统架构图

如上图 1 所示，万里开源万里分布式数据库管理系统数据库架构包括两部分：

SQL 节点 (GreatSQL)：所有的 GreatSQL 节点组成了可横向扩展的计算层。SQL 节点用于接收应用请求，解析 SQL 语法，生成执行计划以及执行 SQL 计划。SQL 节点只存储集群的元数据，SQL 节点之间通过 paxos 进行元数据同步，保证整个集群元数据的强一致性。

DTM (Distributed Transaction Manager) 节点：分布式事务管理器，负责维护全局事务 ID 的申请，全局事务活跃列表的获取等相关任务。

数据节点 (GreatDB)：存储真实的业务数据。数据节点接收 SQL 节点的 SQL 请求，给计算层提供数据存储和抽取接口。数据节点以 shard 模式进行组织，一个 shard 包含多个 GreatDB 节点，shard 内部的 GreatDB 节点保存同一份数据，通过 paxos 协议进行数据同步，从而保证 Shard 数据的副本强一致。所有的 Shard 组成了整个数据存储。

2. 万里分布式数据库管理系统产品特性

简洁的架构

万里分布式数据库管理系统架构简洁，采用计算与存储分离架构，SQL 节点负责执行计划构建与执行，数据节点负责业务数据存储和部分计算功能。

动态横向扩展

万里分布式数据库管理系统支持在线分层按需横向扩展。

SQL 节点不存储业务数据，是轻量节点，可快速进行横向扩展以增强集群计算能力。新扩增的 SQL 节点会通过 paxos 自动同步集群元数据，待完全同步完成后，可对外提供服务。

数据节点主要负责存储业务数据，在线扩容数据节点会触发数据单调重分布，实现集群存储能力的扩容。数据节点扩容以 shard 为单位，并保证足够的高可用。

灵活的数据分布策略

万里分布式数据库管理系统支持如下多种类型的表分布策略：

- Normal：普通表，随机的分布在某个 shard 上。
- Global：全局表，分布在所有的 shard 上，每个 shard 上拥有全量的表数据。
- Partition

□ Hash: 数据按照分区字段的 hash 值, 分布在不同的 shard 上, 每个 shard 上存储部分表数据, 不同分片数据之间没有交集。

Range: 数据按照分区字段的范围, 分布在不同的 shard 上, 每个 shard 上存储部分分片数据, 不同分片数据之间没有交集。

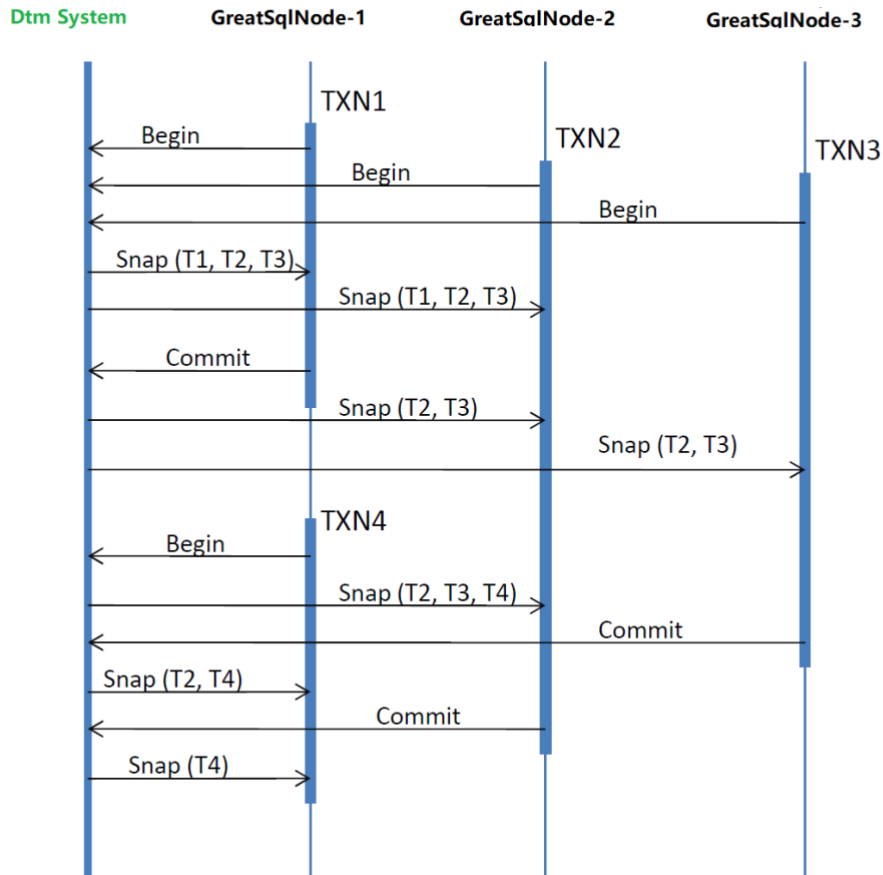
分布式事务

万里分布式数据库管理系统提供分布式事务支持, 通过全局事务管理器 (DTM)、两阶段提交 (XA)、MVCC 多版本控制、行级锁等技术实现完整的分布式事务的能力。确保集群事务操作 ACID 完整性。支持读已提交、可重复读两种隔离级别, 实现高效的读写访问性能。同时 DTM 具备高可用特性, 可以在 GreatSQL 层中进行快速故障转移。

DTM 的基本功能包括:

1. 给事务提供全局可见的事务 ID (DTID)。
2. 给事务提供全局一致性的视图。
3. 同时也会管理全局事务的启动、提交、回滚等操作。

GreatSQL 计算节点将会和 DTM 交互来处理各种事务请求, 从下图可以看出, GreatSqlNode 会通过 DTM 给全局事务指派全局事务 ID (DTID), 以及获取全局一致性的视图信息 (图中的 Snapshot)。



完整的 SQL 标准支持

万里分布式数据库管理系统 支持完整的 SQL 语法标准，完美兼容 MySQL 通信协议、SQL 语法及各种数据库对象。万里分布式数据库管理系统 SQL 标准支持如下

1. DML 语句

例如 INSERT、UPDATE、DELETE 以及各类复杂查询 SQL。

2. DDL 语句

例如 ALTER TABLE、CREATE DATABASE、DROP FUNCTION 等。

3. DCL 语句

例如 GRANT、SHOW、SET、KILL 等管理命令。同时提供了集群管理命令扩展。

4. 视图

提供视图对象的创建、修改、删除，允许 SQL 查询使用视图来简化业务逻辑的编写。

5. 存储过程

提供存储过程的创建、修改、删除，允许将业务逻辑以 SQL 语言的方式编写到存储过程中，简化应用开发维护，同时通过减少网络交互实现高效的执行性能。

6. 用户自定义函数

提供用户自定义函数的创建、修改、删除，除了支持数据库系统函数，用户还可以通过 SQL 创建业务逻辑函数，来简化 SQL 的开发过程。

7. 触发器

提供触发器的创建、修改、删除，允许 DBA 配置行级触发器进行数据维护操作。支持针对插入、修改和删除的事前触发器和事后触发器。

8. 定时任务

提供定时任务的创建、修改和删除，允许 DBA 配置定时任务来简化日常运维调度过程。

9. 临时表

提供会话级临时表的创建、修改和删除，允许业务程序在互不干扰的情况下，通过临时表进行业务逻辑执行过程中临时性数据的存储。

丰富的数据类型

万里分布式数据库管理系统提供丰富的数据类型支持。

数值类型	整数: integer, int, smallint, tinyint, mediumint, bigint
------	--

	浮点数: decimal, numeric, float, double
	Bit 数: bit
时间类型	Date, datetime, timestamp, time, year
字符类型	文本: char, varchar
	二进制: binary, binary
	大对象: blob, text
	枚举与 set: enum, set

高可靠与数据冗余

万里分布式数据库管理系统采用全组件冗余架构，任意组件故障不会影响集群的可用性，高可用达到 99.99%。

集群所有的组件都基于 Paxos 保证元数据与业务数据的强一致，任一组件都至少部署 3 个节点，在某台机器出现故障的时候，自动进行故障转移，任意故障场景 RPO=0，确保数据零丢失。

万里分布式数据库管理系统支持同城双机房、两地三中心等部署架构，可以构建去中心化的跨机房对等部署集群，实现数据库级别的容灾能力。

高性能数据服务

万里分布式数据库管理系统采用计算存储分离的原生分布式架构，通过分布式并行计算、SQL 负载均衡技术实现集群高性能吞吐。

同时基于数据库集群优异的线性扩展能力，万里分布式数据库管理系统吞吐性能随集群节点规模准线性提升。

良好的上下游生态

万里分布式数据库管理系统集群具备良好的生态，完，可与产业链上下游产品无缝衔接，比如 ogg、informatica、kettle。所有的第三方工具、编程语言、客户程序可以像使用单机数据库一样与万里分布式数据库管理系统集群进行完美的交互。

完整的国产化支持

万里分布式数据库管理系统集群提供丰富的国产化支持，包括龙芯、飞腾、鲲鹏等国产主流芯片；麒麟、统一等国产操作系统和云平台等国产生态。

3. 数据库管理系统产品平台与指标

硬件环境支持

支持本地存储（SATA、SAS、SSD/NVMe SSD 等）

支持千兆、万兆 Ethernet 网络

支持通用的 X86 和 ARM 架构芯片，对国产主流芯片进行了兼容适配，适配的国产芯片包括龙芯、飞腾、鲲鹏、海光、申威、兆芯等。

操作系统支持

支持的国产操作系统包括：

1. 中标麒麟服务器操作系统
2. 统信服务器操作系统
3. 深度操作系统服务器版
4. 银河麒麟服务器操作系统
5. Open Euler
6. Turbo Linux
7. 其他操作系统：RedHat Linux、SUSE Linux 等 Linux 操作系统。

技术指标

下表列出了适用于数据库系统的约束和最大值：

每个表中列的个数	1000
每个表中索引的个数	60
数据节点数大小	1024
数据库、表、视图、列、索引名长度	64

单个数据库存储大小	无限制
最大单表数据空间	64TB
数据库最大个数	无限制
主键最大长度	3072
索引最大长度	3072
服务器最大连接数	无限制
单表最大行数	无限制
结果集最大列个数	1017
单个 BLOB(CLOB)大对象的大小	4GB
用户和角色个数	无限制

4. 开发接口支持

万里分布式数据库管理系统提供标准的数据库接口 ODBC、JDBC、OLEDB、ADO.NET、PDO 等常用接口；

万里分布式数据库管理系统支持各种开发语言和工具，如 JAVA、C/C++、C#、PHP、Python、GO 等。

5. 竞争优势

万里开源万里分布式数据库管理系统产品具备透明可扩展、金融级高可用、多种心部署、高并发低延迟、完备的 SQL 能力等特性。

主要优势包括：

相比传统国产数据库厂家，万里开源数据库在集群扩展性、性能、以及数据库生态方面有较大优势；

相比开源数据库 MySQL/PostgreSQL，万里开源数据库在可靠性、扩展性和性能方面具备碾压优势；

相比新兴分布式数据库厂家，万里开源数据库具备更加完备的 SQL 语法支持，具备更加强大的事务吞吐性能，对应用适配更加友好。

6. 联系我们

网址：www.greatdb.com

地址：北京市朝阳区东三环建外 SOHO 东区 B 座 32 层

服务热线：400-032-7868

服务传真：(010) 64087839

电子邮件地址：info@greatdb.com

官微：

