

可配置的拓展2MB事务功能设计文档

修订历史

版本	修订日期	修订描述	作者	备注
Cedar0.3	2017-09-28	可配置的拓展2MB事务功能设计文档	屈兴 朱涛	

1 需求分析

Cedar是华东师范大学计算机科学与软件工程学院基于OceanBase研发的可扩展的关系数据库，实现了巨大数据量上的跨行跨表事务。OceanBase0.4.2中，一个事务写操作(IUD)的操作日志数据量的限制，即不能超过2MB大小。而这样的限制导致了OB数据库不支持长事务的执行和大批量更新操作。因此，对于大数据量的写事务执行，OB只能通过多个类似的短，小数据量的写事务组合完成。这种方式，显然具有明显的缺点：

- 长事务的原子性等性质被破坏
- 执行时间长、资源开销大
- 业务逻辑复杂

针对这一问题，在Cedar0.3版本中，实现了基于存储过程的可配置的拓展2MB事务的功能。该功能支持手动配置参数设置事务最大操作日志数据量限制，使以存储过程执行的事务可以突破OB原有的2MB大小限制。

2 功能简述

可配置的拓展2MB事务的功能，支持事务级的写数据量限制，可以自定义配置，修改配置项后，重启集群后生效。单条sql级的写数据量不超过2MB大小。

3 设计思路

3.1 子功能模块划分

可配置的拓展2MB事务的子模块，主要有：

- MergeServer长事务词法语法解析模块
解析长事务hint，设置相关事务属性。
- UpdateServer长事务执行管理模块
新增长事务执行流程，与短事务资源隔离。
- 日志处理流程缓冲区大小配置模块
新增UpdateServer系统配置项，重启UpdateServer后根据配置项设置日志相关缓冲区大小。

3.2 总体设计思路

考虑到一般的begin transaction开始事务一般不会超过2MB，为了支持超过2MB事务的执行，Cedar0.3在存储过程的基础上来支持此功能。即存储过程中的for，while循环语句可以产生写数据量大，执行时间长的事务，并将其称之为长事务。

1.MergeServer长事务词法语法解析模块

长事务的定义与普通存储过程的定义完全一致，当传入的存储过程参数不同，循环次数不同，导致存储过程可以是长事务也可能是短事务。为了使长事务可以执行，而不会因为写数据量过大、执行时间过长和其他高并发的小事务抢用资源（主要是内存），导致系统资源耗尽，最终长事务和众多小事务执行失败而回滚。在存储过程调用的时候增加一个hint（LONG_TRANS）表明该事务是长事务，进入长事务的执行流程。

2.UpdateServer长事务执行管理模块

长事务的执行流程与普通的短事务的执行流程实际上相同，只是长事务执行过程的所用到的内存由长事务的内存分配器分配。

3.日志处理流程缓冲区大小设置模块

为了保证事务的持久化，修改事务日志处理流程中相关缓冲区大小的常量为读取配置的变量。使得UpdateServer重启后，读取用户重新配置的UpdateServer系统配置项，完成日志等缓冲区大小重置。