首页 / 谈谈分布式数据库的分片键选择准则和数据重分布的思考



谈谈分布式数据库的分片键选择准则和数据重分布的思考





原创 🛭 范达宏 2024–08–24

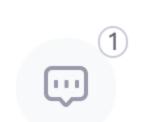


• 396

270浏览

192浏览

100浏览



()

一、分片键选择的重要性

熟悉分布式数据库的伙伴们应该都明白,选择正确的分片【布】键【shard key】的是分布式数据库 表现出卓越性能的基础,一旦选择了不合适的分片键,有可能的跨节点的分布式事务、繁重的数据重分布 将会拖累整个系统,其性能表现不说优于集中式数据库,能否正常运行都未可知。

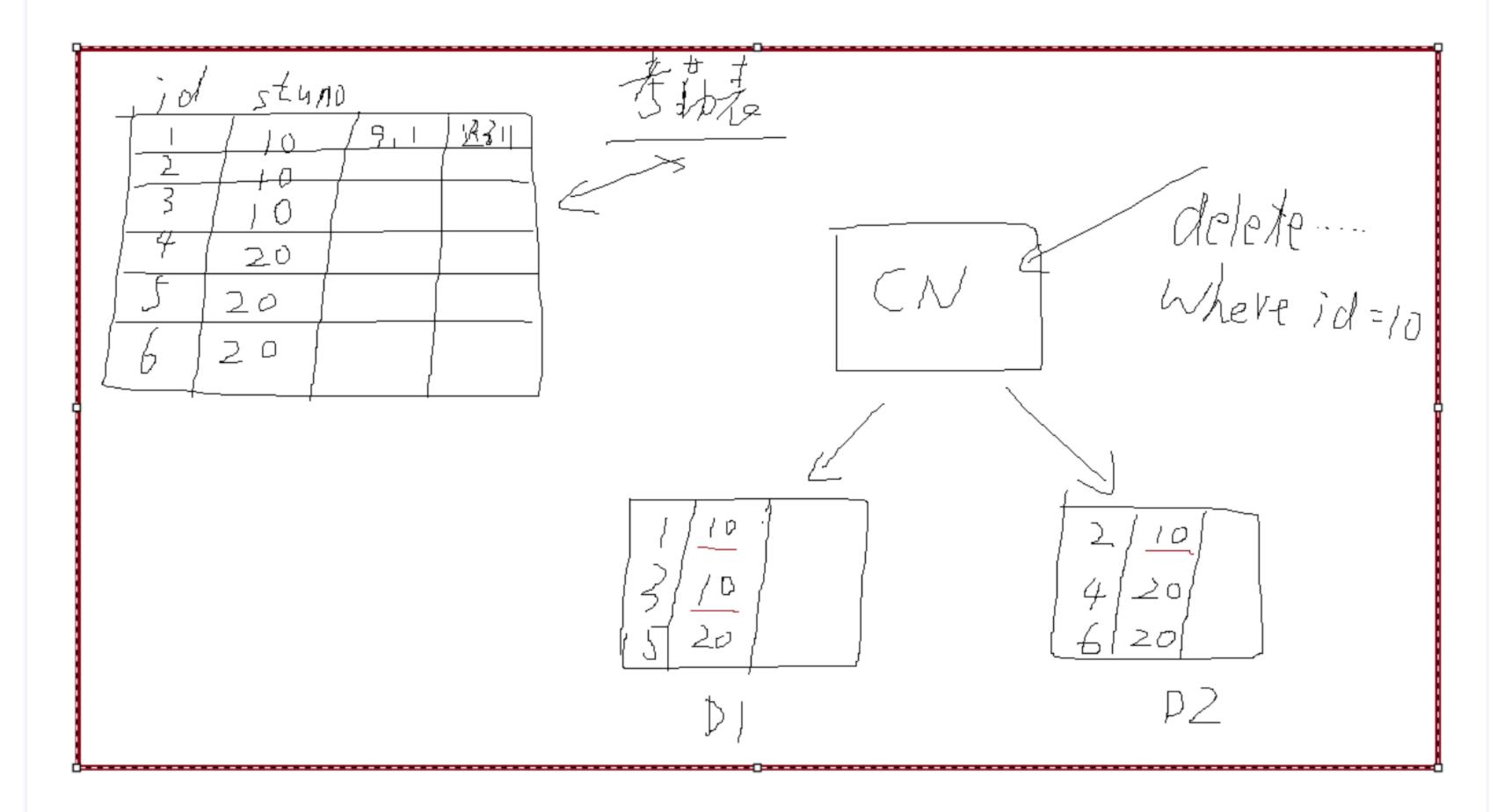
多数分布式数据库都是HATP【即混合事务/分析处理】类数据库,即OLTP【事务性】和OLAP【分析 型】,因此本文中讨论的分片键的选择准则也是基于HATP类型数据库的特征展开的。

二、分片键选择的准则

第一部分讨论了分片键选择的重要性,作者现在阐述作者导师总结的三条分片准则,并进行详细解 析。

准则1.尽量使用业务键作为分片键,避免产生分布式事务,避免产生跨节点的增删改。

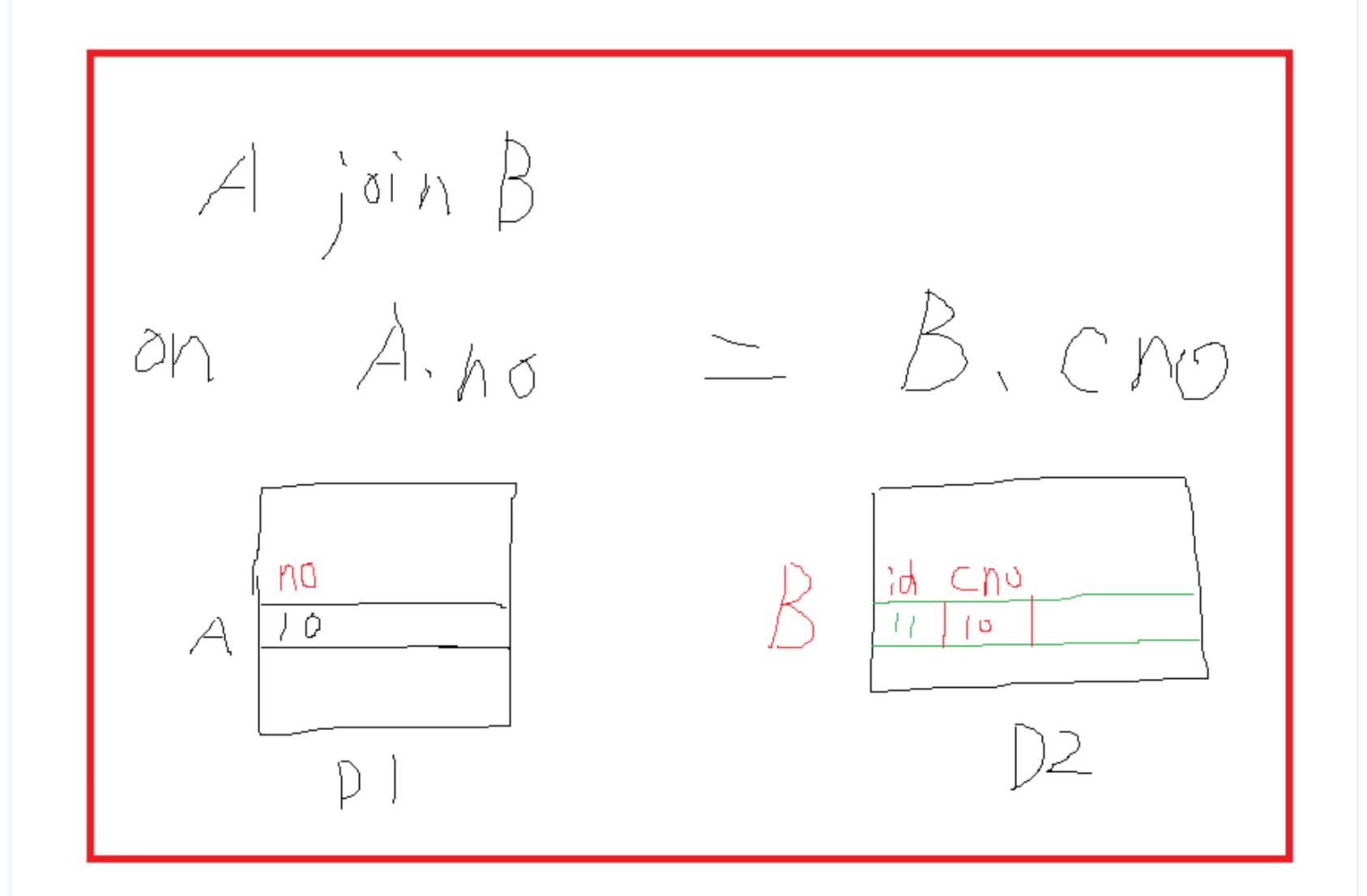
所谓业务键指的是在表中该键对应的数据表示了元组的某个实际存在的属性。比喻说很多表的主键喜 欢用序列生成,这种键的键值就没有表示实际含义,就不是业务键,因此在实际业务中,该键基本不会出 现在SQL语句的条件列和join列中。那么为什么使用了这类非业务键作为分片键,容易产生分布式事务、 跨节点的增删改操作呢?如下图所示【图很丑,意思能表达清楚就好】



上图假设了一张学生考勤记录表,其中id列为序列生成的非业务键,如果按照id列作为分片键的话, 数据节点D1和D2各分布了3条数据【假设只有6条数据,不影响举例】。如果我们要删除学号=10【学号为 业务键,更容易在SQL语句中出现】的学生的考勤记录,则需要将D1中的2条数据删除和D2中的1条数 据,即产生了分布式事务,跨节点进行了删除操作,影响了系统的高并发。

准则2.尽量避免数据重分布,大表关联条件可作为分片键

数据重分布是分布式数据库无法避开也无法完全避免的一个话题,数据重分布的大致含义就是某个DN 节点在收到CN节点下发的任务时,任务存在表关联的情况,由于另外一个表的等值条件的元组不在本DN 节点,因此需要从其它DN读取数据,如下图所示:





一网打尽linux 7系列安装11g RAC上所有

三种连接方式的通俗解释-应该都能看懂

2024-10-28

2024-07-11

2024-10-30

的坑

还原程序员两次的懵逼事件带来的小知识 59浏览 2024-07-15 最新文章 三种连接方式的通俗解释-应该都能看懂 100浏览 2024-10-30 关于Oracle和Tbase中的varchar2和varch ar的探讨 270浏览 2024-10-28 还原程序员两次的懵逼事件带来的小知识 59浏览 2024-07-15 一网打尽linux 7系列安装11g RAC上所有 的坑 192浏览 2024-07-11 小白可以这样走上DBA之路 475浏览 2024-07-09

目录

• 二、分片键选择的准则

- 准则2.尽量避免数据重分布,大...
- 准则1、2、3出现冲突怎么办,...

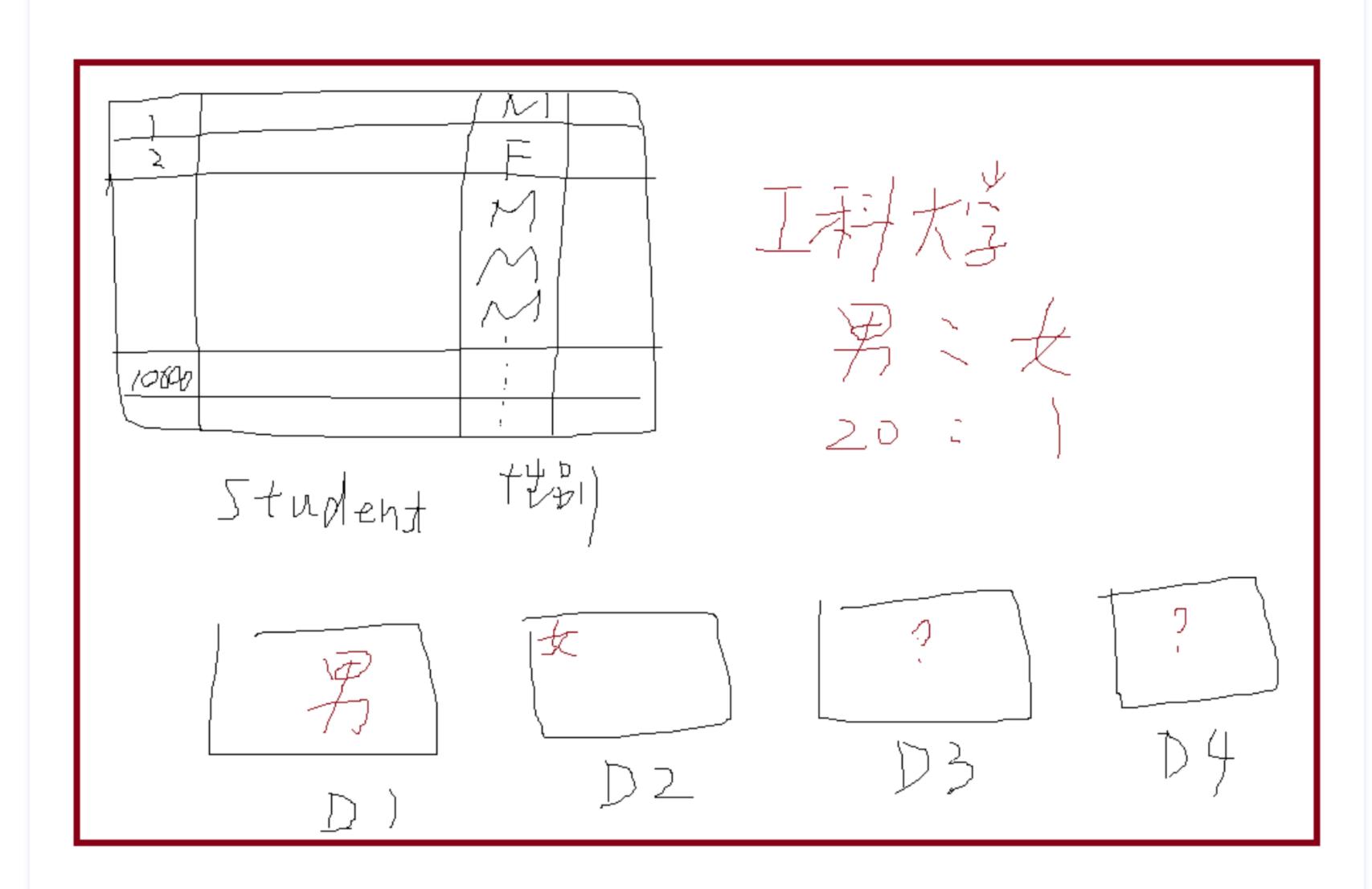
• 三、关于数据重分布的思考

- 思考4:星形模式中事实表怎么选...
- 思考5:消除很小的表的数据重分...

A表和B表join时,A表派出了分片键no列,B表派出了自己非分片键cno,因此A表no=10的元组和B表中cno=10的元组并不在一个DN上,因此需要在DN之间建立一个连接【因此数据重分布也会增加DN之间的连接数】,将B表中的数据移动到D1上进行操作,即发生了数据重分布行为。而该案例中B表如果是选择了cno作为分片键,则B表中cno=10的元组也将出现在D1中,则不会发生数据重分布行为。

准则3.尽量避免数据倾斜

分布式的优势在于可以通过MR让很多台机器同时干一件事情,映射到分布式数据库就是可以在多个DN上同时运算某个查询任务,然后在CN上汇总结果。但如果出现数据倾斜,则数据大量出现在某一个节点【或某几个少数节点】,对于该表的查询将全部集中在这个节点上,因此分布式数据库的优势将荡然无存。



如上图,某工科大学10000名学生,男女学生比例为20:1,学生表选择了性别列作为分片键,则会出现D1节点存放约9532名男生信息,D2节点存放468名女生信息,D3、D4节点不存放该表任何行数据。

准则1、2、3出现冲突怎么办,比如说某表按照准则1选择A列作为分片键,按照准则2则选择B列作为分片键,此时需要根据业务的实际规律进行分析,即从客户的角度综合考虑,选择对整个系统最优的分片键。一般是情况是准则1>准则2、3的。

三、关于数据重分布的思考

说明:本部分基于作者自己最近学习分布式数据库的思考,可能存在理解不深刻的地方,望大家一起思考、指正。

思考1: A join B on A.id = B.id and A.type='X' A表选择id还是type作为分片键?

我的思考是两种选择是两种维度,选择id虽然在多个DN上进行了运算,但是没有数据重分布,选择ty pe列虽然在一个DN上运算,但是出现了数据重分布。由于数据重分布消耗资源较为严重,因此倾向于使用id列作为分片键。

思考2:A join B on A.id = B.id and A.name=B.name 这种情况下id列[或name列]作为双方的分片键,运算时会出现数据重分布吗?

我的答案是不会出现数据库重分布。因为双方均按id列进行了分片,id相同的数据元组已经分在了同一个DN上,第2个条件只不过是在同一个DN上继续筛选数据。

思考3:两个表join的列是否为分片键的3种组合的数据重分布情况。

我的思考是:

1)双方均有诚意,拿出了自己的分片键进行join,则无数据重分布。

2)A表拿分片键,B表没拿分片键,则B表进行数据重分布。

3)双方均没有拿出自己的分片键join,系统会将每个表的完整数据发送到所有节点,这可能是最糟糕的一种情况了。

思考4:星形模式中事实表怎么选择分片键?

我的思考是由于事实表的每一个业务列都有可能和另外一个表进行join,在同一个SQL中用不同列跟不同的表join亦属常见情况,因此该SQL出现数据库重分布已经无法避免,分片键只能选择其中一个了,一般可按出现最多的频率去选。

思考5:消除很小的表的数据重分布之后,SQL性能出现几倍的提升是为什么?

我的答案是对于这种场景我也不明白,下周我再问问导师,目前只能是理解到系统是不是到了一个瓶颈,哪怕多出少数负担,性能也会急剧下降。

墨力计划 分布式数据库 国产分布式数据库 分片表 分片键

最后修改时间: 2024-08-26 09:55:35

「喜欢这篇文章,您的关注和赞赏是给作者最好的鼓励」

关注作者

赞赏

【版权声明】本文为墨天轮用户原创内容,转载时必须标注文章的来源(墨天轮),文章链接,文章作者等基本信息,否则作者和墨天轮有权追究责任。如果您发现墨天轮中有涉嫌抄袭或者侵权的内容,欢迎发送邮件至:contact@modb.pro进行举报,并提供相关证据,一经查实,墨天轮将立刻删除相关内容。

评论

分享你的看法,一起交流吧~



范达宏 ◎ LV.3

其实这种偏理论的东西含金量挺高的

5月前 △ 点赞 □ 评论

相关阅读

OceanBase数据库一体机获创新产品奖!

OceanBase数据库 162次阅读 2025-01-20 11:14:13

「订单来了」携手OB Cloud,实现订单处理能力翻倍,存储成本削减75%

OceanBase数据库 156次阅读 2025-02-09 16:11:53

【技术干货】磐维数据库_分布式数据库运维-正确使用物理备份特性

磐维数据库 122次阅读 2025-01-24 17:56:24

盛京银行:基于OceanBase构建全栈国产分布式数据库平台的实践探索与应用

OceanBase数据库 99次阅读 2025-01-23 18:09:31

208万!深圳证券交易所、深圳证券通信有限公司采购20节点GoldenDB

天下观查 97次阅读 2025-02-02 07:04:37

OB Cloud 两大重要更新发布!跨云容灾能力再升级、AP 场景迎来首个LTS版本

OceanBase数据库 64次阅读 2025-01-21 13:55:59

基于 OceanBase 列存与物化视图的数仓探索 | 优秀征文分享

OceanBase数据库 52次阅读 2025-01-21 13:19:30

突破单一资源限制:OceanBase 混闪机型部署验证及生产实战应用探索 | 优秀征文分享

OceanBase数据库 50次阅读 2025-01-23 17:35:46

常见问题场景日志解读 | OceanBase最佳实践 12

OceanBase数据库 46次阅读 2025-01-21 12:41:22

无锡农商行信贷核算系统升级实践:基于OceanBase的新一代系统实施

OceanBase数据库 44次阅读 2025-01-27 22:35:39