



大家都在搜...



下载APP

开源软件

问答

动弹

打赏 ¥

评论

收藏 ☆

点赞

分享文章

微博

QQ

微信






春哥大魔王的博客的个人空间 > 工作日志 > 正文

jQuery MiniUI

快速开发WebUI界面, 支持Java、.Net、PHP www.miniui.com

决定分库分表的原因有哪些? 原



春哥大魔王的博客 发布于 2019/06/20 13:42 字数 965 阅读 404 收藏 0 点赞 0  评论 0

[B+Tree](#) [Java](#) [MySQL](#) [InnoDB](#)

写在前面

前几天梁大发表了mysql单表500w数据分表的铁律, 也参与了回复, 结果同一天隔壁组面试, 正好问了下面试者这个问题, 面试官想在多方面考察面试者技术扎实程度, 结果面试者回答的不好, 所以整理一篇, 决定分库分表的原因有哪些?



春哥大魔王 置顶



服务器当然是一个很大的因素。但是跟业务场景qps也很有关系，单虚拟机，单物理机mysql实例的写qps也是有极限的，为应对未来业务发展的需求，也需要按照qps规划做分库 😁

删除

| 作者



太赞了

单表数据量角度

这是梁大文章中阐述的主要原因之一：

曾经在中国互联网技术圈广为流传着这么一个说法：MySQL 单表数据量大于 2000 万行，性能会明显下降。事实上，这个传闻据说最早起源于百度。具体情况大概是这样的，当年的 DBA 测试 MySQL 性能时发现，当单表的量在 2000 万行量级的时候，SQL 操作的性能急剧下降，因此，结论由此而来。然后又据说百度的工程师流动到业界的其它公司，也带去了这个信息，所以，就在业界流传开这么一个说法。

阿里巴巴《Java 开发手册》提出单表行数超过 500 万行或者单表容量超过 2GB，才推荐进行分库分表。对此，有阿里的黄金铁律支撑，所以，很多人设计大数据存储时，多会以此为标准，进行分表操作。

事实上，这个数值和实际记录的条数无关，而与 MySQL 的配置以及机器的硬件有关。因为，MySQL 为了提高性能，会将表的索引装载到内存中。InnoDB buffer size 足够的情况下，其能完成全加载进内存，查询不会有问题。但是，当单表数据库到达某个量级的上限时，导致内存无法存储其索引，使得之后的 SQL 查询会产生磁盘 IO，从而导致性能下降。当然，这个还有具体的表结构的设计有关，最终导致的问题都是内存限制。这里，增加硬件配置，可能会带来立竿见影的

性能提升哈。

单行数据特点角度

这是飞哥的观点，分库分表原因和每行记录大小有关系，因为b+ tree索引叶子节点保存真实记录数据，每条记录大小决定了每个叶子节点能保存多少记录，影响了b+tree索引树的高度，从而影响了性能。

业务特点角度

春哥大魔王的角度是从业务角度切入，比如虚拟机mysql实例的写qps瓶颈大概在3k左右，物理机mysql实例的写qps瓶颈大概可以摸高到1w，但是随着业务场景发展，随便搞个大促，或者推广，整个系统的瓶颈就会暴露出来，所以需要提早做分库分表了。

连接数角度

另一个角度可以是连接数，因为随着单个mysql实例连接的下游服务越来越多，导致mysql和服务之间建立大量的长连接，连接数量是完全有上限的，可以通过读写分离，分库分表等多个角度降低连接数造成的性能问题。

最后

所以分库分表，数据量只是其中的因素之一，可以从机器配置，mysql原理，连接池，业务特点等多个角度考虑。



香港服务器,双向CN2极速直连

恒创科技 免备案服务器, ping低至10ms, 7x24h服务, 3天无忧退款保障
技i

© 著作权归作者所有

¥ 打赏

👍 点赞 (0)

☆ 收藏 (0)

➦ 分享

🖨 打印 🚩 举报

◀ 上一篇: 一直再说高并发, 多少QPS才算高并发?

下一篇: 大型系统演进之路-负载均衡演进 ▶



春哥大魔王的博客

粉丝 44 博文 327 码字总数 295117

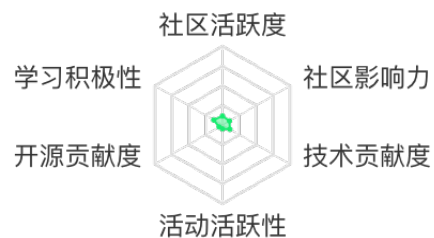
作品 0

📍 海淀 🏢 程序员

♡ 关注

✉ 私信

💬 提问



评论(0)

在这里发表你对此文的观点

😊 插入表情 # 插入软件

0/1000



发表评论

相关文章

最新文章

备战一线互联网公司Java工程师面试题 (2)

JVM 1、请介绍一下JVM内存模型??用过什么垃圾回收器都说说呗 2、线上发送频繁full gc如何处理? CPU 使用率过高怎么办? 如何定位问题?如何解决说一下解决思路和处理方法 3、知道字节码吗?字节...

j4love 2018/04/14 👁 395 💬 5

【死磕Sharding-jdbc】——基于ssm

原文作者: 阿飞Javaer 原文链接: <https://www.jianshu.com/p/602e24845ed3> 本篇文章讲解如何在ssm (spring、springmvc、mybatis) 结构的程序上集成sharding-jdbc (版本为1.5.4.1) 进行分库分...

飞哥-Javaer 2018/05/05 👁 0 💬 0

【死磕Sharding-jdbc】——基于 SSM 集成sharding-jdbc2.0.3

原文作者: 阿飞Javaer 原文链接: <https://www.jianshu.com/p/7b6997c3586d> 本篇文章讲解如何在ssm (spring、springmvc、mybatis) 结构的程序上集成sharding-jdbc (版本为2.0.3) 进行分库分表...

飞哥-Javaer 2018/05/27 👁 0 💬 0

通过四道常问面试题，带你了解什么是数据库分库分表

编者语: 为了避免被误解为:「手里有把锤子, 看什么都是钉子!」, 说明一下不是什么业务都适合分布式数据库, 更不是用了分布式数据库性能就一定得到扩展。其次: 本文为纯干货, 建议先转发...

老道士 2019/06/16 👁 55 💬 0

shardingSphere 分库分表后, 多表联查缺少数据

使用shardingSphere对mysql数据库order 和orderItem表分库分表后, 使用inner join 连接两表 数据查不出来。分库分表策略

孙亮 2019/11/21 👁 79 💬 0

加载更多

OSCHINA 社区

[关于我们](#)
[联系我们](#)
[合作伙伴](#)
[Open API](#)

在线工具

[码云 Gitee.com](#)
[企业研发管理](#)
[CopyCat-代码克隆检测](#)
[实用在线工具](#)

微信公众号



OSCHINA APP

聚合全网技术文章，根据你的阅读喜好进行个性推荐

[下载 APP](#)

©OSCHINA(OSChina.NET) 工信部 开源软件推进联盟 指定官方社区

深圳市奥思网络科技有限公司版权所有 粤ICP备12009483号