

阿里巴巴云时代的数据库管理

目录

阿里数据库技术团队概述&&个人简介

业务发展与规模化

挑战与机遇：性能、系统性问题、多样化

DBA的挑战与进化

云时代：让研发具备DBA的能力

个人简介

- 阿里巴巴数据库团队高级专家
- 《高性能MySQL 第三版》译者
- 09年研究生毕业加入阿里数据库团队，负责淘宝核心数据库运维、管理、优化、配置，构建了淘宝数据库第一代MySQL运维监控系统。
- 2013年开始支持阿里云数据库RDS运维与线上故障处理，并开始探索并实践数据库服务产品化之路，向用户提供数据库管理、数据传输、数据库优化等服务产品。
- 个人博客：<http://www.orczhou.com>
- Google: **orczhou**



阿里巴巴数据库技术团队概述



负责阿里旗下所有子公司的数据库管理。

负责历年双11、双12、春节红包等大型活动的数据库容量规划、架构升级、运行稳定性等。

主导了阿里去IOE、异地多活等架构落地，负责阿里巴巴MySQL内核研发。

研发了iDB(企业数据库研发流程平台)、OceanBase云平台、数据管理DMS、数据传输DTS等云产品。

数据库团队的进化

运维工具时代

- Oracle、小型机、存储、脚本

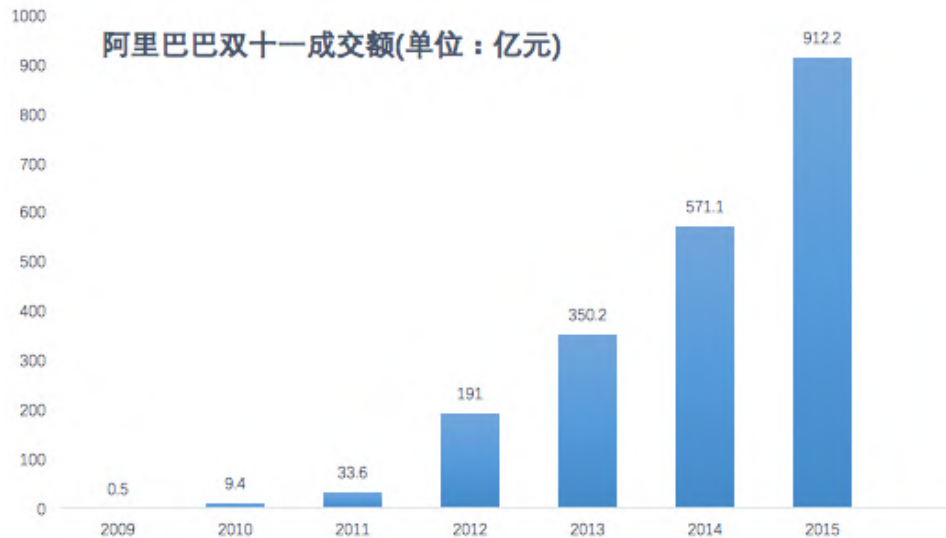
平台与服务时代

- 双11、去IOE、异地多活

云服务时代

- 云计算、OceanBase、Docker、DevOps、自助

来自业务的压力



2012年：『某核心集群总共执行293亿次SQL/天，集群总QPS达86万/秒，集群TPS11万/秒，单机QPS高达6.5万/秒』

2015年：约500万QPS

规模化与性能优化

- 简单的数学：5%的性能提升意味着什么？

10000台主机*5万*5% = 2500万

10000台主机*5万*10% = 5000万

技术进化1：对性能极致的追求

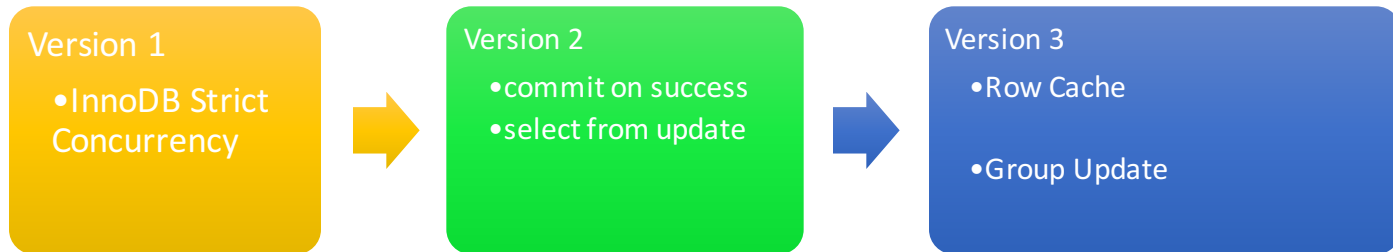
对性能极致的追求：电商库存场景

- 原生MySQL版本事务数：550
- 业务：每秒卖出多少商品

成本巨大、架构复杂



对性能极致的追求：电商库存场景



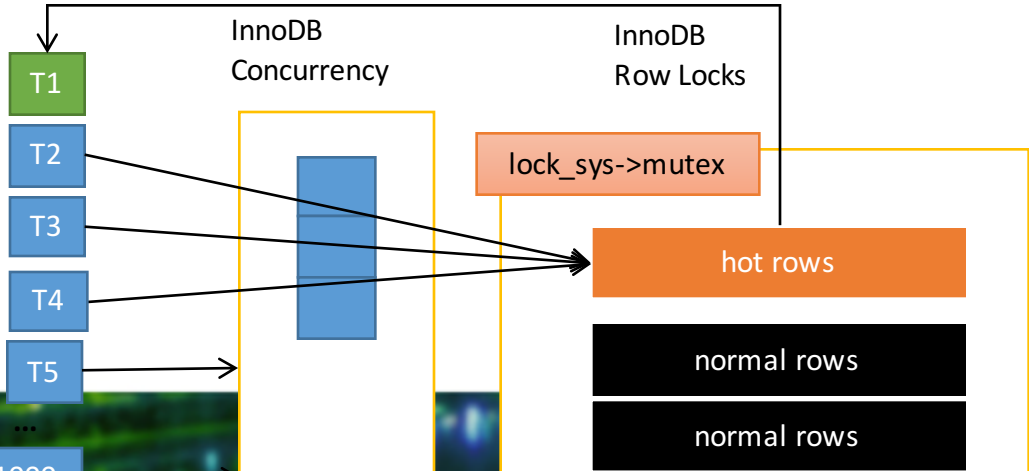
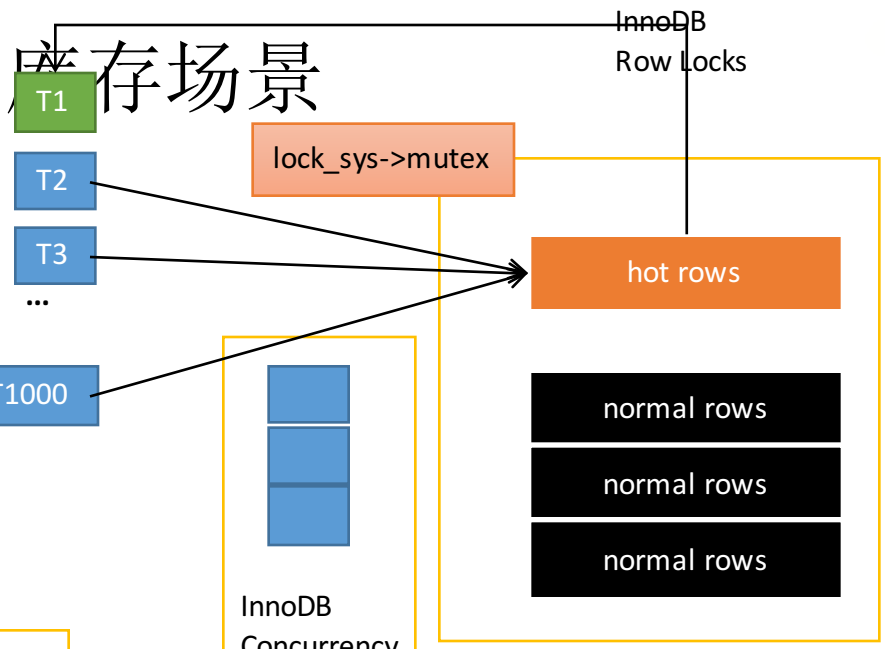
对性能极致的追求：电商库存场景



对性能极致的追求：电商库存场景

Transaction model:
① begin;
② insert normal row;
③ update hot row;
④ select hot row;
⑤ commit;

Deadlock Searching

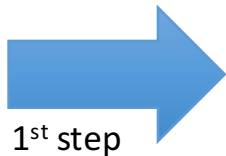


Limit The Waiters

对性能极致的追求： 电商库存场景

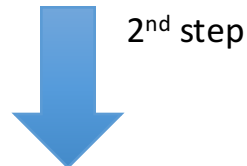
Transaction model:

- ① begin;
- ② insert normal row;
- ③ update hot row;
- ④ select hot row;
- ⑤ commit;



Transaction model:

- ① begin;
- ② insert normal row;
- ③ select * from update hot row;
- ④ commit;

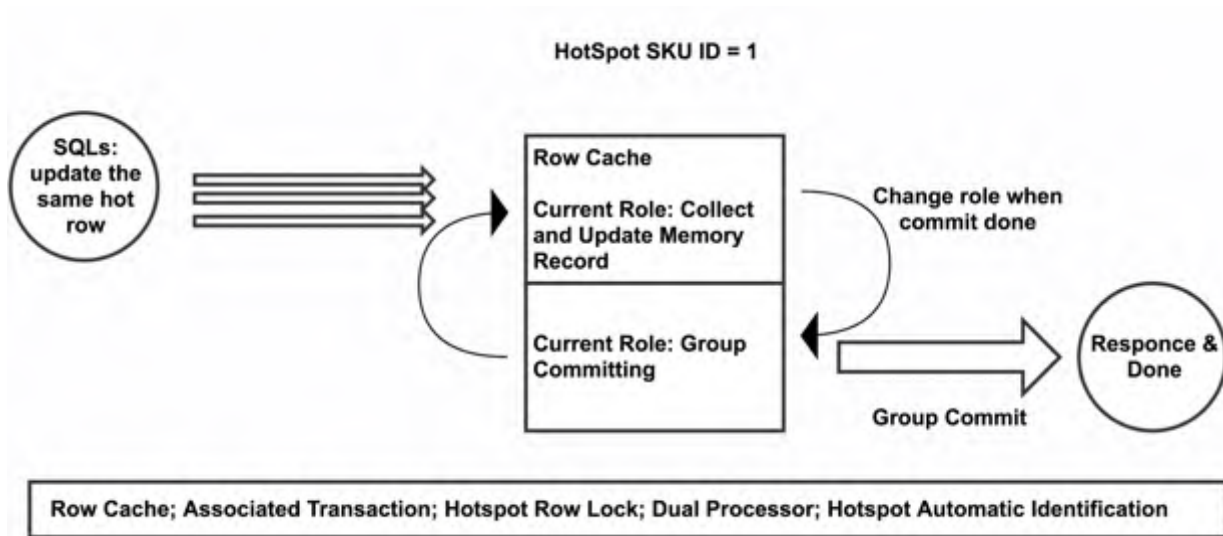


Transaction model:

- ① begin;
- ② insert normal row;
- ③ select * from update commit_on_success rollback_on_fail

对性能极致的追求：电商库存场景

- 自动热点识别



- 批量提交

规模化与系统问题

- 简单的数学：如果某个系统正常运行十年出一次故障，那么运行一万个该系统的实例，每个月会有多少次故障？

$$(1/10/12)*10000 = 83.33$$

规模化与系统问题

- 标准化：让系统更稳定
 - 上线前压测提升稳定性
 - 让自动化成为可能
 - 每次解决一个问题，其实是解决了整个平台的问题
- 解决每一个暴露的系统问题

系统问题案例： Bug#77572

- 现象： 5.6在online DDL的时候， 如果其他客户端写入数据出现了UK冲突后， 那么会导致online DDL失败。
 - DDL表有UK键
 - DDL的表比较大时会出现
 - DDL时其他客户端写入了冲突数据会出现

系统问题案例： Bug#77572

- 背后的原因分析：
 - 在Online DDL过程中，会记录所有对目标表的PK的修改行为，即Online Log
 - 在Insert的时候先插入PK，发现成功，记录一条Insert log
 - 再插入UK，发现冲突，回滚Insert操作，记录一条Delete log
 - 全量数据拷贝完成后，在应用增量Online Log的时候，在应用第一条Insert log时，发现UK冲突，报错退出，DDL失败

系统问题案例： Bug#77572

- 官方态度：
 - 是一个明确的Limitation;
 - 只修改报错内容;
- 阿里数据库：
 - 修复该bug
 - 通过平台自动化处理

业务扩张与多样化



数据库技术与产品服务



目录

阿里数据库技术团队概述&&个人简介

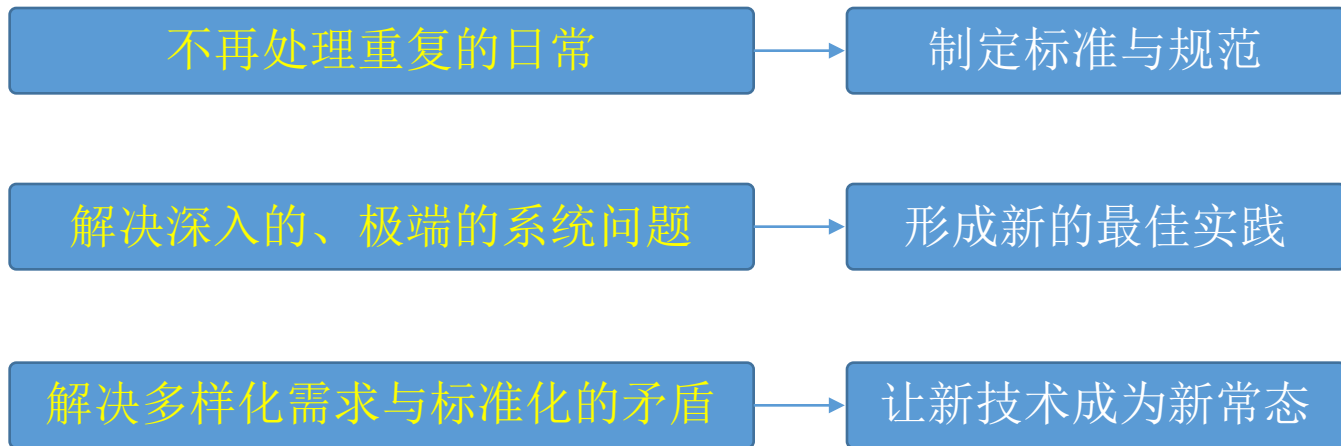
业务发展与规模化

挑战与机遇：性能与系统稳定性

DBA的挑战与进化

云时代：让研发具备DBA的能力

DBA的挑战与进化



专职的DBA会越来越少：平台会取代DBA的所具备的基础能力
专业的DBA会越来越贵：DBA的专业能力会被平台所放大

云时代：让研发具备DBA的能力

iDB：企业数据库研发流程平台

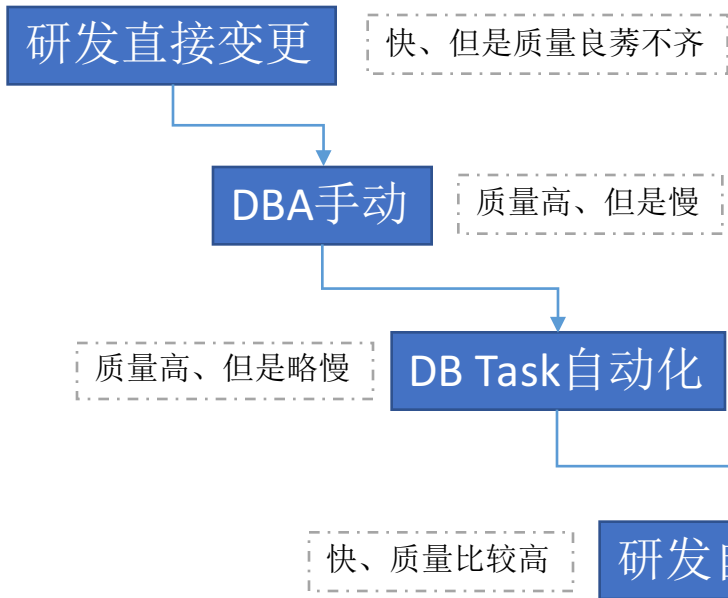


流程与规范

性能优化与诊断



让研发具备DBA的能力：流程与规范

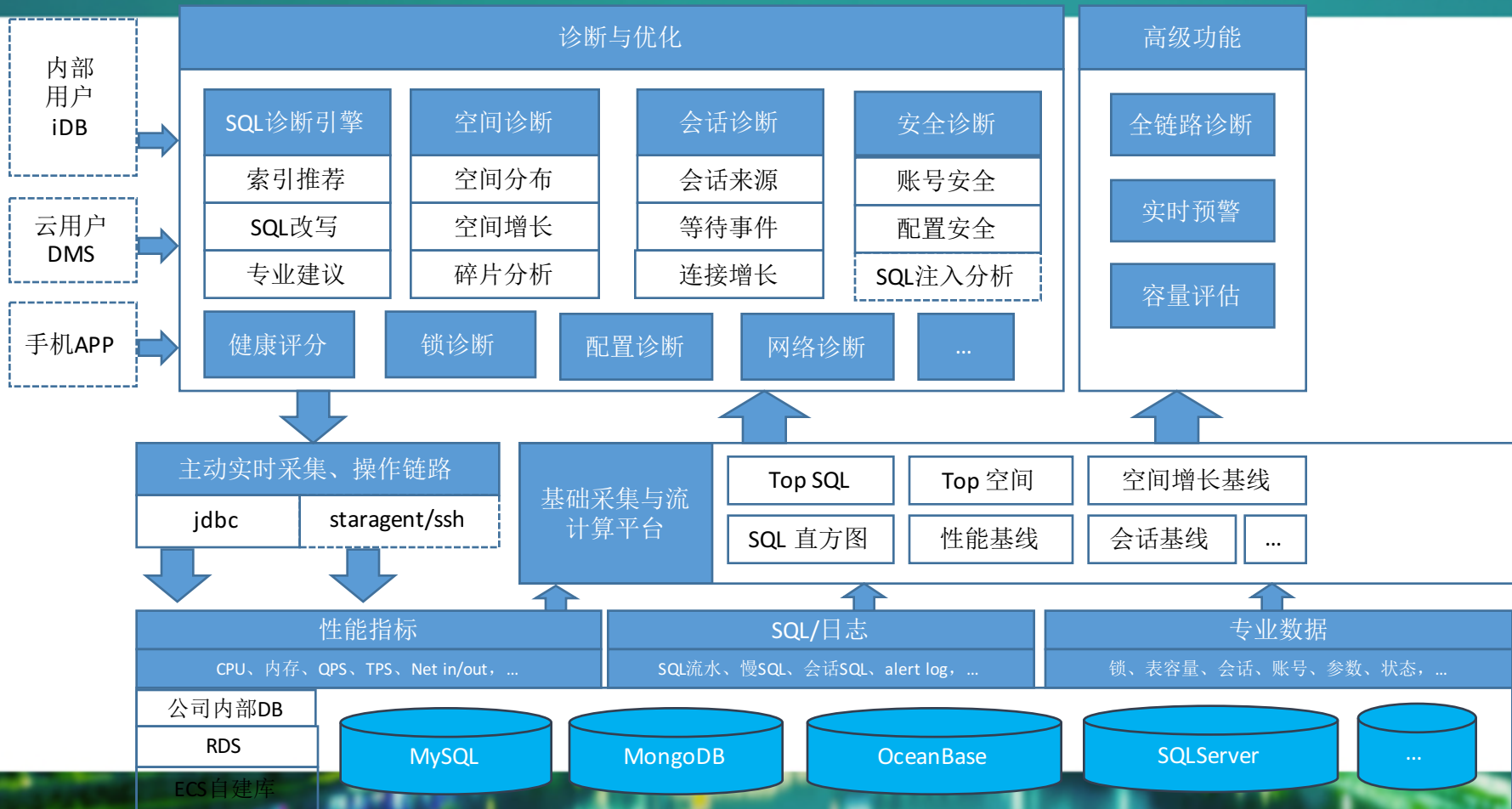


结构规范
索引合理
SQL高效

- 引擎使用InnoDB
- 不使用外键
- 表必须加上注释
- 尽量使用UNSIGNED
- 字符集使用UTF8
- 默认定义为NOT NULL
-
- 表必须有数值主键(InnoDB)
- 索引命名规范
- 不要过多使用单列索引
- 字符串尽量定义前缀
-
- 尽量使用索引
- 尽量使用索引覆盖
- 核心SQL避免复杂逻辑：计算、
- 避免WHERE中的隐式转换
- 建议使用合理的分页实现
-

让研发具备DBA的能力：诊断与优化

运维技术与实战峰会



挑战还在继续，欢迎加入

