

# DTCC

## 数 / 造 / 未 / 来

### 第十二届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA 2021

2021 年 10 月 18 日 - 20 日 | 北京国际会议中心



数 / 造 / 未 / 来  
第十二届中国数据库技术大会  
DATABASE TECHNOLOGY CONFERENCE CHINA 2021

# SQL代码自动审核方案及平台建设

中国工商银行软件开发中心 林镇熙

DTCC  
2021



北京国际会议中心

2021/10/18-10/20



ChinaUnix.net

ITPUB

# 目录

---

## 01. 痛点和难点

---

## 02. 一种代码自动审核实现方案

## 03. SQL代码审核平台建设



# 痛点

- SQL问题未在研发阶段暴露，生产出了问题，为时已晚
- 性能问题尤为棘手！
- 出事后的灵魂拷问：
  - 1、SQL没用上索引，为什么没发现？
  - 2、过滤条件建了索引，但SQL没用上索引，为什么没发现？
  - 3、SQL用了索引，但索引效率很低，为什么没发现？
  - 4、规范不允许SQL这样写，为什么没发现？



# 难点

- 传统处理方案存在短板

方案	短板
压力测试	一般选取个别、重点交易，无法全量覆盖
代码扫描	没有结合表结构、生产统计数据等信息，无法定位性能问题，多用于代码规范性检查
基于测试环境数据库系统视图检查SQL运行情况	依赖测试环境测试覆盖情况，且测试环境的数据量一般偏小，数据分布可能与生产也有偏差，效果有限



# 目录

---

01. 痛点和难点

02. 一种代码自动审核实现方案

---

03. SQL代码审核平台建设



# 典型性能问题1 - 没建索引

- 分析方法1：结合表结构，才能分析全表扫/全索引扫的性能问题！

```
<select id="getInfo"
parameterType="TestObject">
SELECT * FROM t1 WHERE b = #{b}
</select>
```

+

```
create table t1
(
    id int primary key,
    a varchar(10),
    b varchar(10),
    c int,
    key idx_a (a)
);
```

+

```
class TestObject
{
    private String a;
    private String b;
    ....
}
```

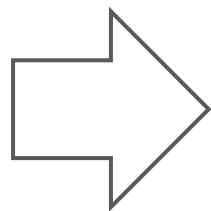


table	type	possible_keys	key	rows	生产记录数
t1	ALL	null	null	1000	1000000





# 典型性能问题2 - 隐式数据转换

- 分析方法2：结合应用JAVA程序，才能分析隐式数据转换导致无法匹配索引的性能问题！

```
<select id="getInfo"
parameterType="TestObject">
SELECT * FROM t1 WHERE a = #{a}
</select>
```

+

```
create table t1
(
    id int primary key,
    a varchar(10),
    b varchar(10),
    c int,
    key idx_a (a)
);
```

+

```
class TestObject
{
    private int a;
    private String b;
    ....
}
```

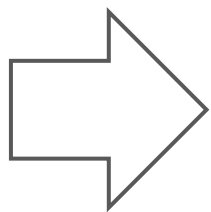


table	type	possible_keys	key	rows	生产记录数
t1	ALL	idx_a	null	1000	1000000





# 典型性能问题3 - 使用低效索引

- 分析方法3：结合生产统计信息，才能定位使用低效索引的性能问题！

```
<select id="getInfo"
parameterType="TestObject">
SELECT * FROM t1 WHERE a = #{a}
</select>
```

+

```
create table t1
(
    id int primary key,
    a varchar(10),
    b varchar(10),
    c int,
    key idx_a (a)
);
```

+

```
class TestObject
{
    private String a;
    private String b;
    ....
}
```

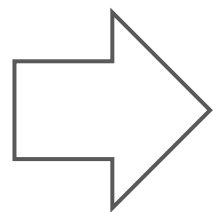


table	type	possible_keys	key	rows	生产记录数	生产索引区分度
t1	ref	idx_a	idx_a	1000	1000000	10





# SQL分析5要素

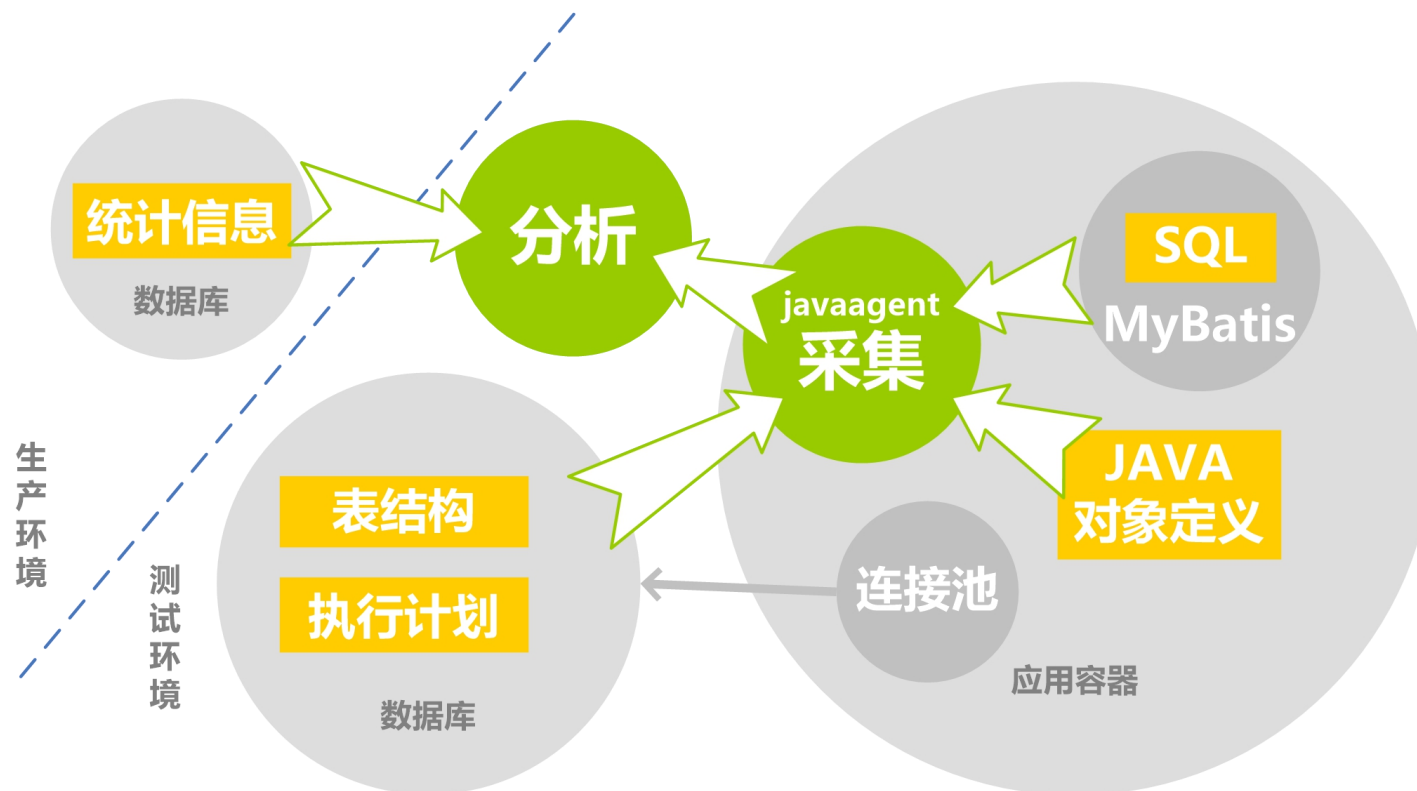
---

- SQL
- 表结构
- 输入参数/JAVA对象定义
- 统计信息
- 执行计划





# 基于javaagent+MyBatis的采集分析



# 目录

---

01. 痛点和难点

02. 一种代码自动审核实现方案

03. SQL代码审核平台建设



# 审核平台建设

## ■ 1、数据采集

- 范围：研发 + 生产
- 实施条件：应用容器设置环境变量，应用名、镜像名、模板名、版本号
- 难点
  - ✓ javaagent
  - ✓ MyBatis对象处理
  - ✓ 连接池多样化
  - ✓ 获取执行计划



# 审核平台建设

## ■ 2、数据存储

- 数据结构设计
- SQL、表结构、JAVA对象、执行计划随时都可能变化，都需要保存下来
- 难点：分析对象界定，例如：应用+镜像+SQL ID



# 审核平台建设

## ■ 3、数据分析

- 自动化 vs 人工介入
- 白名单管理。例如：部分场景无法检查、问题暂无法优化
- 难点：问题界定，不是所有的全表扫都需要解决





# 审核平台建设

## ■ 4、问题管理

- 对接企业研发管理系统
- 团队分工，人员职责



# 优点

---

- 采集分析对应用透明，应用无需改造
- 无需执行交易即可检查，不受测试覆盖率影响
- 只在研发环境实施，不影响生产运行





# 企业提升

- 分析->解决->模型优化，形成良性循环，不搞一次性运动
- 持续提升自动化能力，减少资源投入
- 专注业务系统研发
- 保障业务系统稳定高效运行





# THANKS