



数据技术嘉年华

Data Technology Carnival

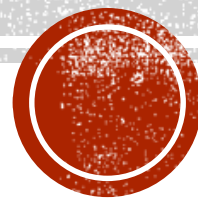
云·数据·智能 - 数聚价值智胜未来

关注公众号回复help,
可获取更多经典学习
资料和文档, 电子书



基于工具的 SQL审核实践

福建农信 谢方恺



第七届



数据技术嘉年华
Data Technology Carnival



背景介绍

- 上百套数据库
- 应用数量持续增加
- 更新迭代速度不断加快
- 大部分生产问题由新上线的代码引起
- 人工代码审核效率无法满足



痛点

01

审核工作量大

02

规范缺乏及时监督约束

03

性能隐患带入生产环境

04

审查结果无法可视化

第七届



数据技术嘉年华

Data Technology Carnival



自动化工具

主要目标

- 提高审核效率
- 保证数据库规范落地
- 量化分析数据库SQL问题
- 产出SQL审核报告



第七届



数据技术嘉年华
Data Technology Carnival



自动化工具



功能需求

- ◆ 与行内资源管理联动
- ◆ 支持审核规则定制扩展
- ◆ 问题自动分析汇总
- ◆ 实现审核版本管理



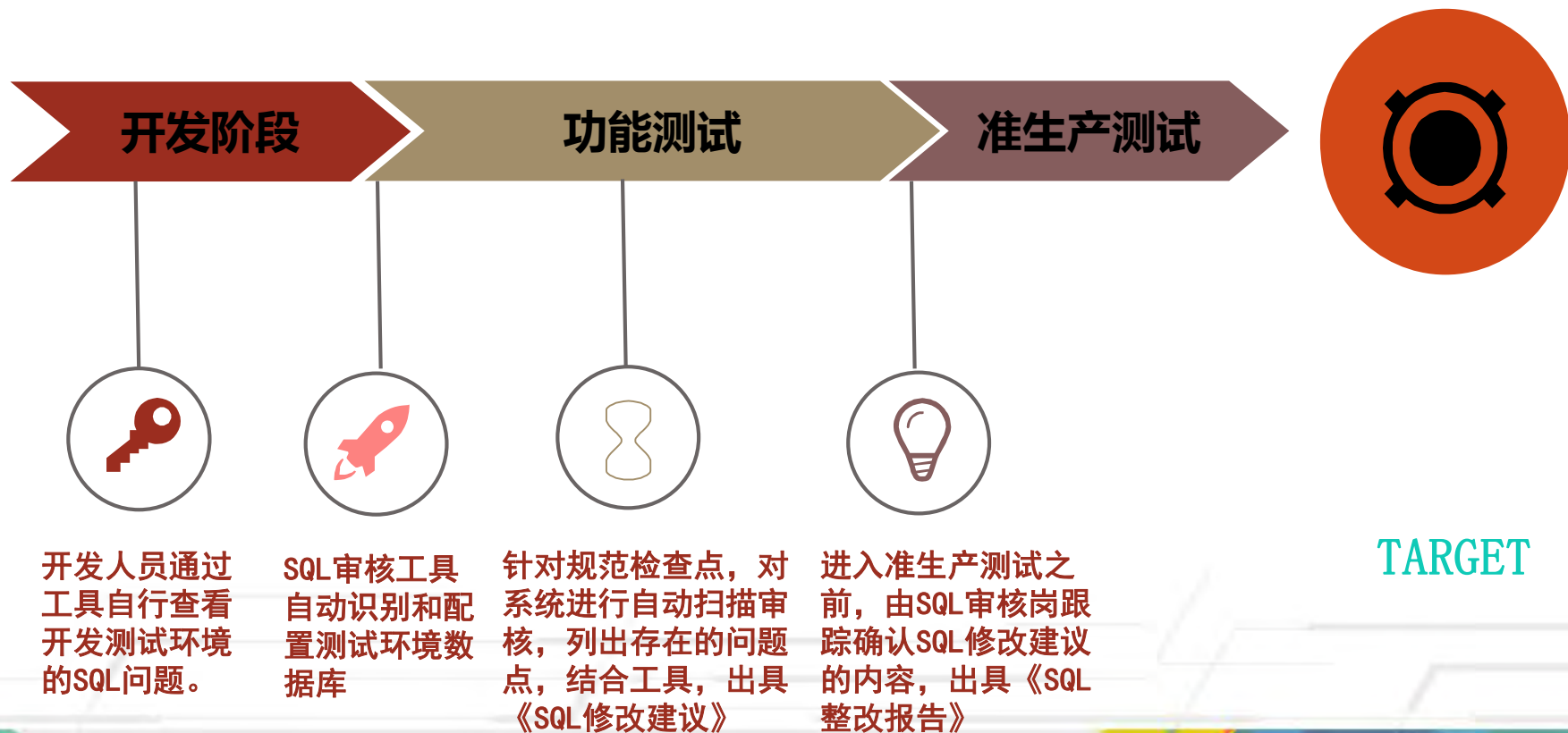
第七届



数据技术嘉年华
Data Technology Carnival



流程整合



开发规范

Oracle开发设计规范

- 命名规范
- SQL/PLSQL编写规范
- 数据库配置规范
- 表设计规范
- 索引设计规范
- 其他对象设计规范

The screenshot shows a web-based Oracle database management interface. The top navigation bar includes '首页' (Home), '数据库' (Database), 'SQL分析' (SQL Analysis), '工作台' (Workbench), and '工具箱' (Toolbox). The main content area displays a '风险分类' (Risk Classification) report for a specific schema. The report lists various SQL issues with their counts and progress bars. The left sidebar shows a list of databases under the 'Oracle' category.

所有数据库

搜索数据库...

Oracle

- 电子渠道-修改-银联前
- 核心&柜面-新增-集团
- 地区特色业务统一接入
- 生源地助学贷款系统二
- 省财政国库集中支付电
- 资金统一管理平台(三
- 黑名单检索系统
- 2017年移动自助柜员机
- 统一收单平台-新增-银
- ATM-新增-跨行转账-身
- 集团客户服务平台
- 中间业务系统-新增-苗
- 农信银-修改-通存来账
- 银联无卡快捷支付
- 核心系统-修改-应付利
- 客户关系管理系统

数据库 > 192.111.6.180@zjglcs > Schema: ZJGL_FJNX

快捷选择: 天 周 月 选择日期: 2017年11月10日 导出报告

风险分类

57个	执行计划中存在全表扫描
19个	使用select *
18个	执行计划中存在索引跳扫
12个	使用反向查询
3个	缺少连接条件
3个	查询语句缺少where条件
3个	在查询列上使用表达式
1个	使用左模糊查询
1个	未使用绑定变量



第七届



数据技术嘉年华
Data Technology Carnival



审核报告

报告内容

- 以项目为维度
- 区分各关联应用和模块
- 内容区分不同的等级

2.2 数据库 192.111.31.10@ecifdb

审核类型	审核项	问题个数
SQL	查询语句缺少 where 条件,执行计划中存在全表扫描	4
	多个条件 or 查询,执行计划中存在全表扫描	1
	存在多表连接,执行计划中存在全表扫描	4
	执行计划中存在全表扫描	1
	查询语句缺少 where 条件,执行计划中存在全表扫描,执行计划中存在全分区扫描	1
	查询语句缺少 where 条件,使用 select *,执行计划中存在全表扫描	2
	查询语句缺少 where 条件	2
对象	所有表/视图必须有 comment	8
	所有字段必须有 comment	10

5.1.1.3.SQL 语句 3

```
SELECT * FROM busservices WHERE SERVICEID = :1 AND 1 = 1
```

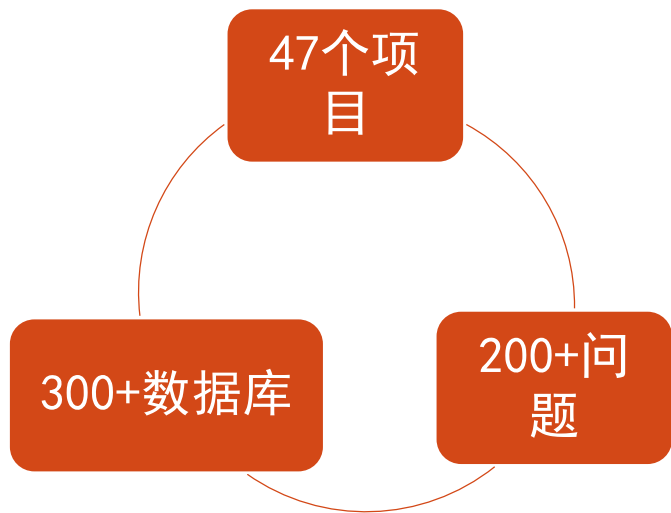
现象: 语句测试期间执行了 41 次, 使用 select *,执行计划中存在全表扫描, 占比较高

DBA 建议: 需要核对查询逻辑是否正确, 是否需要在 SERVICEID 上创建组合索引。

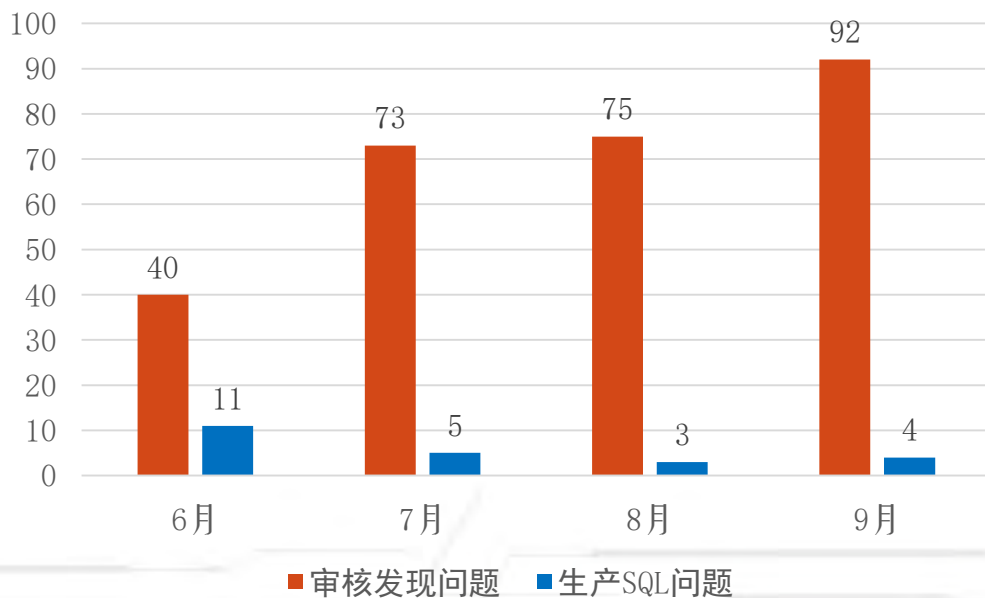
*/*由用户回复及填写修改后内容, 由 SQL 审核员复核确认是否修改完成*/*



成效



审核发现问题与生产问题



意义

用工具化自动化的手段

让DBA从急诊医生转身为系统的保健医生



第七届



数据技术嘉年华

Data Technology Carnival



审核工具MZ3 V2的主要功能

01

数据库动态SQL审核

02

SQL脚本审核

03

结构变更脚本审核



第七届



数据技术嘉年华

Data Technology Carnival





数据库
内存

自动抓取
收集数据

1

2
存储+分析



资料库

3

返回分析后数据

MZ3



SQL文件

人工上传
收集数据

1

展现给用户

4



第七届



数据技术嘉年华

Data Technology Carnival



数据库动态SQL审核

1. 对数据库上运行的SQL进行风险评估
2. 通过采集模块自动抓取数据库上运行的SQL，根据既定规则进行分析，找出其中存在风险的SQL语句和执行计划
3. 适用场景：应用测试阶段、应用上线初期阶段



第七届



数据技术嘉年华
Data Technology Carnival



典型场景-应用测试阶段

现状

- 大多只做功能测试，不关注SQL质量
- 测试环境与生产环境在硬件配置、数据规模等存在差异，性能测试无法全面发现低效SQL

使用工具

- 全面扫描应用使用的SQL，根据规则发现隐患，提前解决



第七届



数据技术嘉年华
Data Technology Carnival



SQL质量



用户

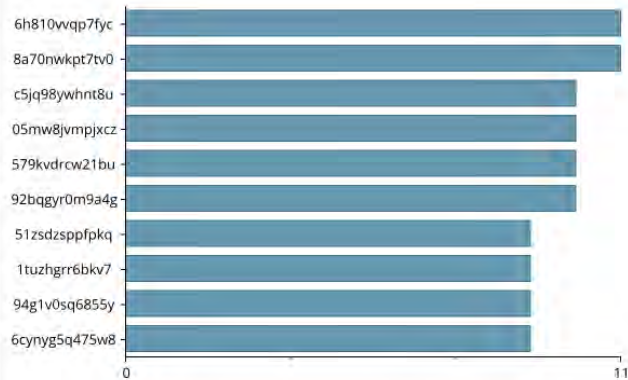


● CUST_U_ENMOTECH

风险分类



SQL语句



类型: SQL语句 执行计划

数据库: cloudcc

选择用户

时间范围: 出现时间 从: 2017/11/01 至: 2017/11/13

高级搜索: 不限

审核规则: [选择审核规则](#) 已选规则: 执行计划中存
在全表扫描

审核时间: 2017年11月13日 -- 2017年11月13日

搜索

生成问题单

导出报告

忽略

取消选择

<input type="checkbox"/>	SQL ID	执行用户	SQL文本	出现时间	最近活动	相似SQL(条)	风险点	操作
<input type="checkbox"/>	ga0snhdr6mspx	CUST_U_ENMO...	select t0.id,t0.beizhu as beizhu,t0.reco	最近一天	最近一天	2	5	
<input type="checkbox"/>	d2qytfj3f7qdg	CUST_U_ENMO...	select t0.id,t0.khmc,(select tt00.name f	最近一天	最近一天	2	8	
<input type="checkbox"/>	7pyz7gazftavc	CUST_U_ENMO...	select t0.id,t0.khmc,(select tt00.name f	最近一天	最近一天	2	8	
<input type="checkbox"/>	6841q2sqfnk44	CUST_U_ENMO...	select t0.id,t0.name,to_char(t0.created	最近一天	最近一天	2	5	
<input type="checkbox"/>	dmc8n2cb1bn7n	CUST_U_ENMO...	select /*+index_desc(t0_idx_cloudcccase	最近一月	最近一天	3	5	
<input type="checkbox"/>	1q38j6qyqdzsy	CUST_U_ENMO...	select t0.id,(case when (select case whe	最近一月	最近一天	26	7	
<input type="checkbox"/>	7q0678gwn4ghh	CUST_U_ENMO...	SELECT count(1) C from (select t0.id fr	最近一天	最近一天	2	1	
<input type="checkbox"/>	g72x3kzn671aw	CUST_U_ENMO...	select t0.id,t0.khmc,(select tt00.name f	最近一天	最近一天	2	8	
<input type="checkbox"/>	7kjf66nh7vj4f	CUST_U_ENMO...	select count(t0.id) as zjrecord from (se	最近一天	最近一天	2	6	
<input type="checkbox"/>	66fczuv3p1m9x	CUST_U_ENMO...	select t2.id,t2.name,t2.obj_id,t2.url,t2	最近一天	最近一天	2	2	

基本信息

首次出现时间: 2017年11月13日 12点10分37秒

最近活动时间: 2017年11月13日 12点10分37秒

执行用户: CUST_U_ENMOTECH

SQL文本

```
SELECT t0.id,
        t0.khmc,
        (SELECT tt00.name
         FROM Account tt00
         WHERE tt00.is_deleted='0'
          AND tt00.id=t0.khmc) AS khmcccname,t0.name,t0.ownerid,(case
        WHEN
```

复制

风险点

风险	影响程度	风险说明	违反数量	查看建议
存在多层次嵌套	警告	SQL中使用in、exists做嵌套条件，嵌套深度超过5层(层次可自定义)	10	
执行计划中出现笛卡尔积	警告	可能缺少连接条件，导致SQL使用笛卡尔积的方式进行连接。	2	
执行计划中存在全表扫描	警告	执行计划中存在全表扫描。阈值参照cost大小，默认值：开发测试...	5	
查询语句使用union	提示	查看语句使用union而不是union all	2	
在查询列上使用表达式	严重	在查询列上使用表达式，可能会导致无法使用列上的索引，影响SQ...	6	
使用反向查询	提示	where子句使用!=、<>、not in、not exists、not like、not等	3	
使用嵌套子查询	提示	使用嵌套子查询	28	
使用标量子查询	提示	在SQL使用标量子查询，当子查询返回的数据行大于1时将导致SQL...	12	

SQL 片段 - 查询列上使用表达式

SQL 片段 【1】

```
concat((t0.knx),")
```

SQL 片段 【5】

```
to_number(case when (t0.knx) is null then '0' else concat((t0.knx),") end)
```

SQL 片段 【3】

```
round(to_number(case when (t0.knx) is null then '0' else concat((t0.knx),") end),0)
```



第七届



数据技术嘉年华

Data Technology Carnival



SQL 片段 - 标量子查询

SQL 片段 【2】

```
select case when c0.codevalue is null then c1.codevalue else c0.codevalue end from tp_sys_code c1 left join tp_sys_code c0 on c0.lang='zh' and c0.codetype='ffe20160263C12BCHj6e' and c0.render=c1.id wh
```

SQL 片段 【3】

```
select q.name from tp_sys_queue q where q.id=t0.ownerid
```

SQL 片段 【4】

```
select tt00.name from Account tt00 where tt00.is_deleted='0' and tt00.id=t0.khmc
```

SQL 片段 【7】

```
select yyc.rate from tp_sys_currency_adv yyc where yyc.currency_code='CNY' and yyc.is_default='1'
```

SQL 片段 【9】

```
select yyc.rate from tp_sys_currency_adv yyc where yyc.currency_code='CNY' and ( t0.jsrq >=yyc.begin_date and yyc.begin_date =(select Max(yyc2.begin_date) from tp_sys_currency_adv yyc2 where t0.js
```



【d2qytfj3f7qdg】详情

基本信息

SQL文本

违反规则

执行计划

1. 4006637172

相关应用

相关表

1. TP_STD_DATATABLE15

1.1 基础信息

1.2 索引

1.3 约束

1.4 分区

1.5 列

返回顶部 ↑

SQL文本

```
...
ELSE
(SELECT u.name
FROM ccuser u
WHERE u.id = t0.ownerid
END ) AS owneridccname,(rtrim(ltrim(to_char(case
WHEN t0.knx is NULL THEN
'0'
ELSE t0.knx end,'999,999,999,999,999,999,990'))))||'%' AS knx,case
WHEN t0.jine is NULL THEN
'0'when t0.jine='0.00' THEN
'0'when t0.jine='0' THEN
'0' else(case
WHEN (t0.currency is null) THEN
'CNY' || ' ' ||rtrim(ltrim(to_char(t0.jine,'999,999,999,999,999,999,990.99'))))
WHEN (t0.currency='CNY') THEN
t0.currency || ' ' ||rtrim(ltrim(to_char(t0.jine,'999,999,999,999,999,999,990.99'))))
ELSE t0.currency || ' ' ||rtrim(ltrim(to_char(t0.jine,'999,999,999,999,999,999,990.99')))) || ' (CNY ' ||rtrim(ltrim(to_char(round((t0.jine/ (case
```

违反规则

审核规则	规则说明	风险SQL (个)
存在多层次嵌套	SQL中使用in、exists做嵌套条件，嵌套深度超过5层(层次可自定义)	10
执行计划中出现笛卡尔积	可能缺少连接条件，导致SQL使用笛卡尔积的方式进行连接。	2
执行计划中存在全表扫描	执行计划中存在全表扫描。阈值参照cost大小，默认值：开发测试环境-10、生产环境-20.	5
查询语句使用union	查看语句使用union而不是union all	2
在查询列上使用表达式	在查询列上使用表达式，可能会导致无法使用列上的索引，影响SQL性能	6
使用反向查询	where子句使用!=、<>、not in、not exists、not like、not等	3

1. 4006637172

1. TP_STD_DATATABLE1

1.1 基础信息

1.2 索引

1.3 约束

1.4 分区

1.5 列

2. TP_SYS_AREA

2.1 基础信息

执行计划

4006637172

基本信息

Hash值: 4006637172
成本: 810

首次出现: 2017-11-13 12:10:37
最近执行: 2017-11-13 12:10:37

详细步骤

Id	Operation	Name	Rows	Bytes	Cost (%CPU)
0	SELECT STATEMENT				810 (100)
1	TABLE ACCESS BY INDEX ROWID	TP_STD_DATATABLE1	1	57	2 (0)
2	INDEX UNIQUE SCAN	SYS_C0016086	1		1 (0)
3	TABLE ACCESS BY INDEX ROWID	TP_SYS_USER	1	31	2 (0)
4	INDEX UNIQUE SCAN	SYS_C0015862	1		1 (0)
5	TABLE ACCESS BY INDEX ROWID	TP_SYS_QUEUE	1	406	0 (0)
6	INDEX UNIQUE SCAN	PK_TP_SYS_QUEUE_ID	1		0 (0)
7	TABLE ACCESS BY INDEX ROWID	TP_SYS_USER	1	31	2 (0)
8	INDEX UNIQUE SCAN	SYS_C0015862	1		1 (0)
9	FILTER				
10	TABLE ACCESS FULL	TP_SYS_CURRENCY_ADV	1	10	3 (0)
11	SORT AGGREGATE		1	1	

相关应用

相关表

1. TP_STD_DATATABLE15

1.1 基础信息

1.2 索引

1.3 约束

1.4 分区

1.5 列

2. TP_SYS_AREA

2.1 基础信息

2.2 索引

2.3 约束

2.4 分区

2.5 列

3. TP_SYS_ROLE

返回顶部 ↑

TP_STD_DATATABLE15

基础信息

创建时间	分析时间	分区键	子分区键	采样大小 (行)	数据 (行)	数据块 (个)	空块 (个)	平均行长 (Bbyte)	描述
2017-08-13 22:24:16	2017-08-14 22:08:14		—	0	0	0	0	0	

索引

索引名	索引列	分析时间	采样大小 (行)	数据 (行)	集群因子	层高	唯一值 (个)	页块 (个)
SYS_C0016296	ID	2017-08-14 22:08:16	0	0	0	0	0	0

约束

约束名	约束列	类型	状态	依赖索引
SYS_C0016296	ID	P	ENABLED	SYS_C0016296

分区

无数据

SQL 脚本审核

1. 用户可以直接将SQL语句提交到平台进行风险评估
2. 平台对SQL语句进行分析，标记其中不符合开发规范、存在风险的部分
3. 适用场景：应用开发阶段



第七届



数据技术嘉年华
Data Technology Carnival



典型场景-代码开发过程

现状

- 缺少SQL代码检查工具，无法保证开发人员严格遵守SQL开发规范

使用工具

- 开发人员自助式审核SQL，并根据规则建议主动修复SQL代码中的问题



第七届



数据技术嘉年华
Data Technology Carnival



分析汇总



风险分类

- 8个 过多表连接
- 3个 where条件中禁止使用嵌套子查询
- 2个 select/update/delete SQL中必须包含where子句
- 1个 like只允许使用右模糊, 不允许使用全模糊
- 1个 表join时必须有join条件

SQL列表

状态: 不限 分析完毕 分析失败

审核规则:

编号	SQL文本	状态	风险点	备注	操作
1	select distinct device_name,sn from rms_device_info where device_name in (?, ?)	分析完毕	0	【StatementId】: getDeviceSnByDeviceName 【文件名】: DeviceSearchDao.xml	
2	select s.id,s.description,s.mip,s.server_status as physical_status,ds.devStatus as service_status,d.a...	分析失败	1	【StatementId】: deviceGroupAccessSearchByTag 【文件名】: DeviceSearchDao.xml	
3	select distinct device_name FROM device_extra_prop WHERE cluster = (select cluster FROM device...	分析完毕	5	【StatementId】: getCacheDeviceByLvsDevice 【文件名】: DeviceSearchDao.xml	
4	SELECT dev.device_name, dev.node_name as secondtier_node, node.province_code, node.node_e...	分析完毕	2	【StatementId】: selectDeviceProAndNode 【文件名】: DeviceSearchDao.xml	
5	select distinct d.device_name from rms_device_info d join hope_node n on d.toptire_id = n.node_...	分析失败	1	【StatementId】: selectDetectDevice 【文件名】: DeviceSearchDao.xml	
6	select s.id,s.description,s.mip,s.server_status as physical_status,ds.devStatus as service_status,d.a...	分析失败	1	【StatementId】: deviceAccessSearchByTag 【文件名】: DeviceSearchDao.xml	
7	select i.device_id,i.device_name description,i.mip," as physical_status,ds.devStatus as service...	分析失败	1	【StatementId】: deviceSearchByTag 【文件名】: DeviceSearchDao.xml	
8	SELECT DISTINCT server_status FROM 'server' where server_status in ('正常','关闭','挂起','未知')	分析完毕	0	【StatementId】: selectAllPhysicalStatus 【文件名】: DeviceSearchDao.xml	
9	select s.id,s.description,s.mip,s.server_status as physical_status,ds.devStatus as service_status,d.a...	分析失败	0	检查\${}符以及其他MyBatis语法, 【StatementId】: deviceAccessSearchByCondition 【文件名】: D...	
10	SELECT DISTINCT t.device_name, d.device_name AS refresh_device FROM (SELECT cluster, device_...	分析完毕	2	【StatementId】: selectRefreshDevices 【文件名】: DeviceSearchDao.xml	

```

SELECT DISTINCT DEVICE_NAME
  FROM DEVICE_EXTRA_PROP
 WHERE CLUSTER =
      (SELECT CLUSTER FROM DEVICE_EXTRA_PROP WHERE DEVICE_NAME = ?)
 AND DEVICESTATE = 'Online'
 AND DEVICE_NAME IN
      (SELECT DEVICE_MRTG.DEVICE_NAME
       FROM DEVICE_MRTG
       LEFT JOIN MRTG_PROPERTY_CONTENT
         ON DEVICE_MRTG.DEVICE_GROUP = MRTG_PROPERTY_CONTENT.ID
       WHERE MRTG_PROPERTY_CONTENT.CONTENT LIKE '%_CACHE%')

```

SQL文本

风险点

编号	违反规则	级别	规则描述
1	where条件中禁止使用嵌套子查询	提示	where条件中禁止使用嵌套子查询
2	where条件中禁止使用嵌套子查询	提示	where条件中禁止使用嵌套子查询
3	like只允许使用右模糊，不允许使用全模糊	提示	like使用只允许使用右模糊，不允许使用全模糊
4	表join时必须有join条件	提示	表join时必须有join条件
5	过多表连接	警告	禁止出现3个以上的表join，可配置表数量

结构变更脚本审核

1. 用户可以直接将DDL语句提交到平台进行风险评估
2. 平台对DDL语句进行分析，标记其中不符合开发规范、存在风险的部分，并且分析变更影响范围：相关依赖对象、相关SQL等
3. 适用场景：表设计阶段



典型场景-表设计阶段

现状

- 缺少结构设计检查工具，无法保证开发人员严格遵守结构设计规范
- 开发人员不了解线上结构使用情况，无法全面评估结构变更的影响范围

使用工具

- 开发人员自助式审核DDL，并根据规则建议主动修复结构设计中存在的问题
- 根据生产库上对象-对象、对象-SQL等依赖关系，自动评估变更影响范围



基础信息

名称: ddl_test

创建人: zdemo

创建时间: 2017年08月22日

状态: 分析完毕

脚本类型: DDL脚本

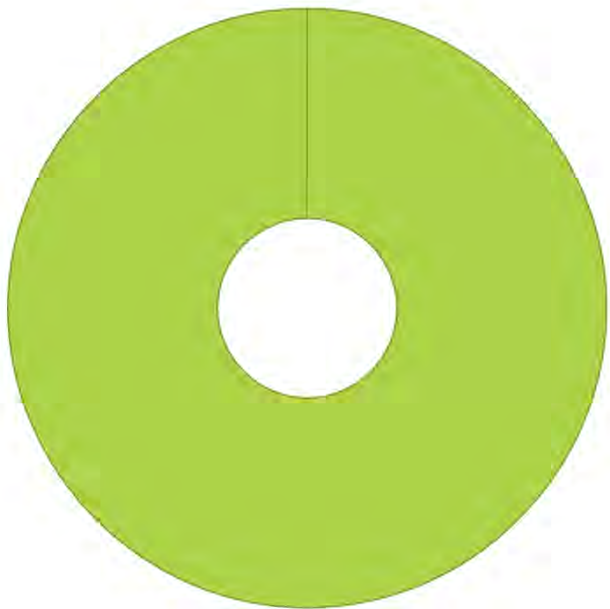
数据库类型: Oracle

数据库: MZ3_PRI

schema: MZ3

分析结果

分析汇总



● 分析完成

风险分类

1个 未指定表所在表空间

1个 对象缺少属主

1个 表名称含有数字等非标准字符

1个 缺少列注释

1个 索引/表使用默认Initrans值

1个 缺少表注释

影响范围

依赖对象		
属主	对象名称	对象类型
MZ3	TRG_BASE_INFO	TRIGGER

« < | 第 1 页,共 1 页 | > » |  显示 1-1 条, 共 1 条

相关SQL			
SQL ID	执行用户	SQL文本	上线时间
没有找到相关的信息			

« < | 第 0 页,共 0 页 | > » |  没有数据

DDL脚本

状态: 不限 分析完毕 分析失败 审核规则:

编号	DDL脚本	状态	风险点	备注	操作
1	DROP PACKAGE MZ3.PACKAGE_DB_STAT	分析完毕	1		
2	drop table DB_BASE_INFO	分析完毕	2		

« < | 第 1 页,共 1 页 | > » |  显示 1-2 条, 共 2 条

一个分享交流的地方



微信号: eyygle



Long Press QR Code To
Identify The Concern

长按二维码识别关注



扫一扫，加入我们，分享更多知识



第七屆



数据技术嘉年华

Data Technology Carnival





THANKS

