Atitit 数据排重检测 与排重清理流程总结

目录

[1. 数据排重监测防止入库 1](#_Toc10405)

[1.1. 联合主键法（最简单 1](#_Toc14241)

[1.2. Unique建（简单 1](#_Toc6952)

[1.3. 触发器检测（较为复杂情况下 1](#_Toc2702)

[1.4. 程序内检测 1](#_Toc19480)

[2. 已有数据排重法（定时 1](#_Toc32564)

[2.1. Mysql timer (event机制排重 2](#_Toc18917)

[2.2. 程序timer排重法 2](#_Toc7221)

[3. 已有数据排重法（一次性 2](#_Toc13008)

[3.1. 导入导出sql法（推荐 2](#_Toc6219)

[3.2. 大数据量下，复制表法，表改名法 2](#_Toc13861)

[3.3. 导出csv等格式再导入（不推荐 2](#_Toc21289)

[3.4. 其他程序处理类 2](#_Toc13331)

# 数据排重监测防止入库

## 联合主键法（最简单

## Unique建（简单

## 触发器检测（较为复杂情况下

## 程序内检测

适合于没有数据库权限情况下。。以及提前预检测

# 已有数据排重法（定时

## Mysql timer (event机制排重

需要建立临时表。作为对比表set，然后双表inner运算，得到差异表 即是需要排除的数据。

## 程序timer排重法

比较啰嗦，只适合于没有数据库权限情况下 不得已选择

任然是临时表与原表作差异运算。。。。数据量特别大情况下，需要做循环运算，不能直接差异对比。。

# 已有数据排重法（一次性

## 导入导出sql法（推荐

，清空表，增加唯一建，导入sql（忽略错误模式

## 大数据量下，复制表法，表改名法

先复制结构，增加建，导入sql，然后原表改名，新表改为原表名。。

## 导出csv等格式再导入（不推荐

## 其他程序处理类