

腾讯音乐 数据资源管理实践

腾讯音乐 高级大数据开发工程师
卢宇鹏



目录 CONTENT

01 背景

03 效果

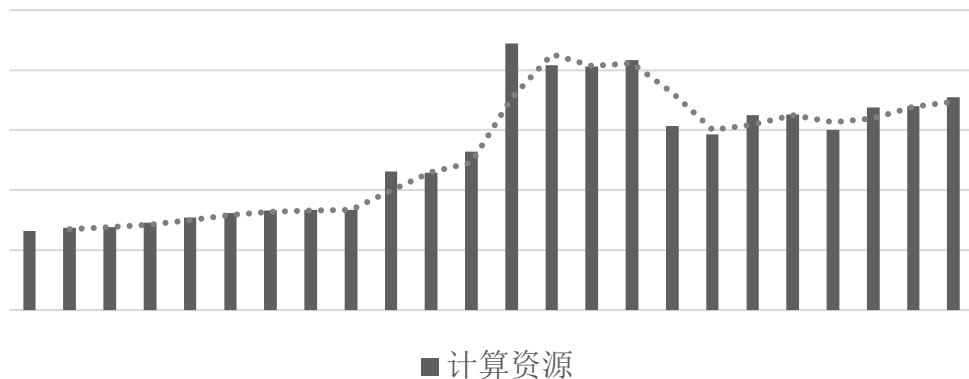
02 数据治理
方案及实践

01 背景

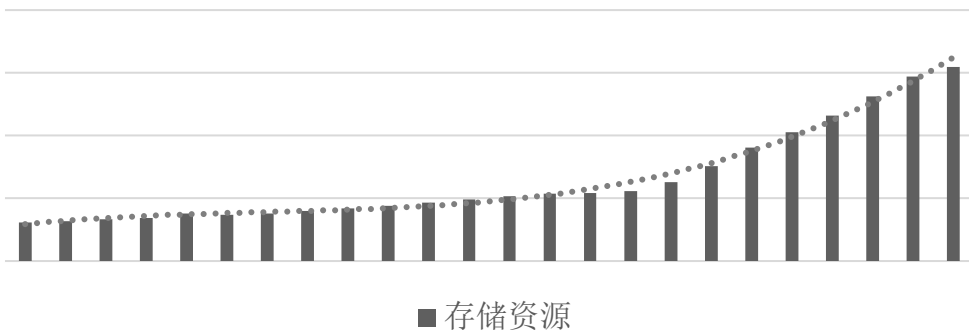


资源增长趋势

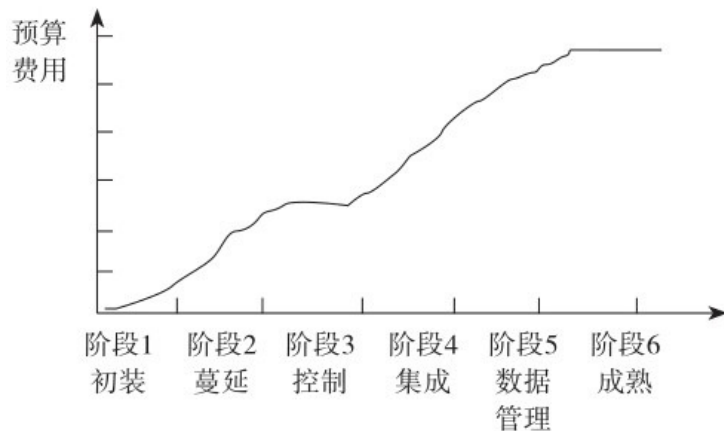
计算资源增长趋势



存储资源增长趋势



诺兰阶段模型



资源成本管理难点



数据量大、涉及人员多



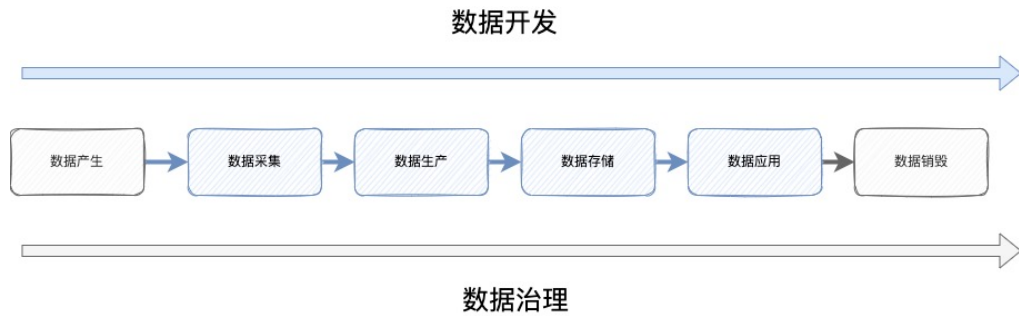
缺乏数仓建设规范，多为烟囱开发模式



数据开发过程中缺乏成本意识，造成资源浪费



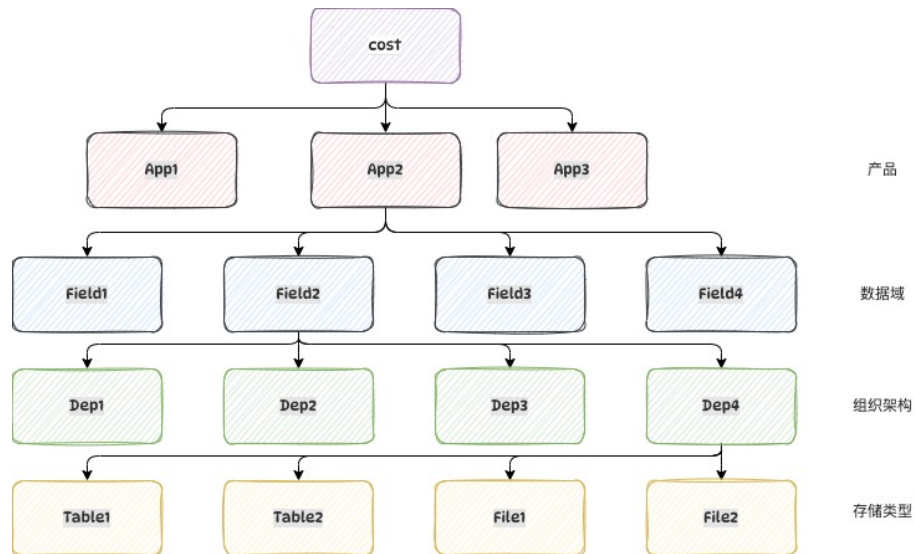
缺乏数据配套优化工具



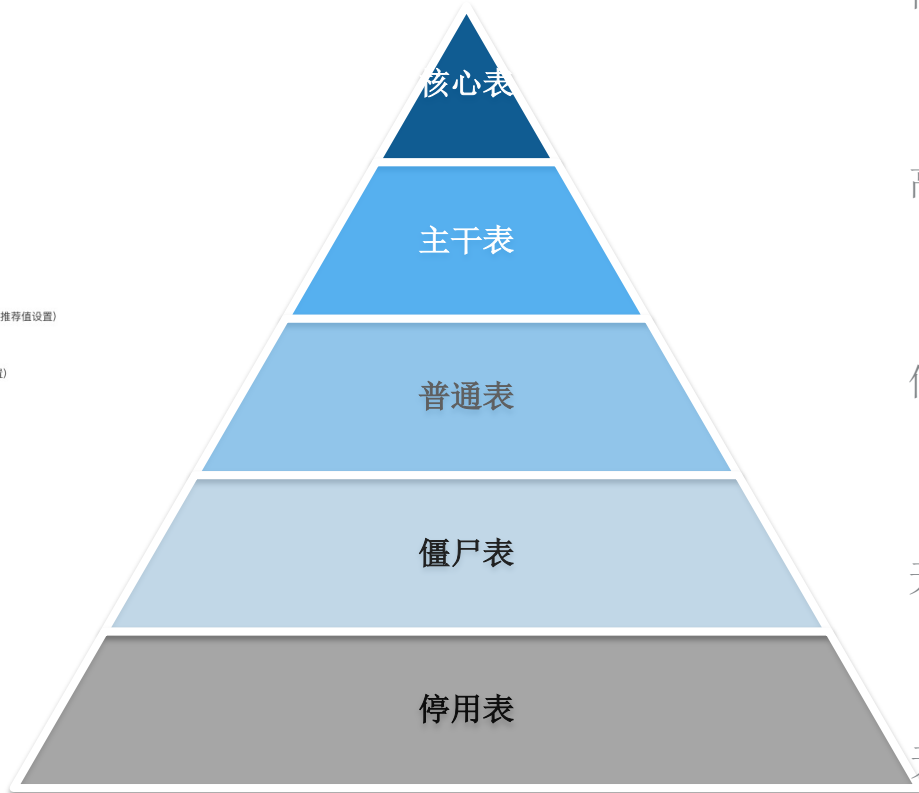
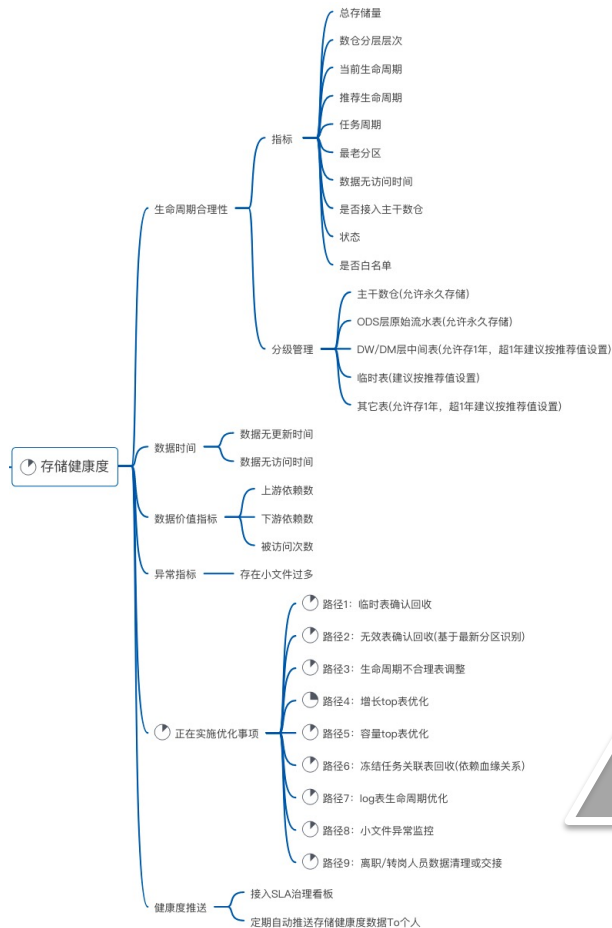
02 数据治理 方案及实践



宏观治理-方法论



宏观治理-方法论



高访问且高扇出

高访问或高扇出

低访问、低扇出

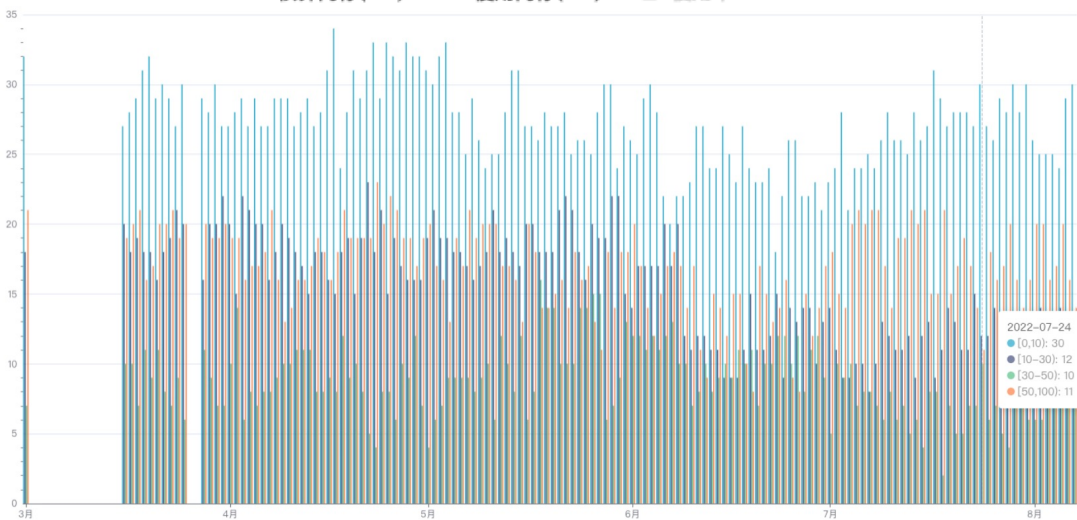
无访问、无扇出

无更新

宏观治理-提升资源利用效率



核算内存(MB) 使用内存(MB) 使用率

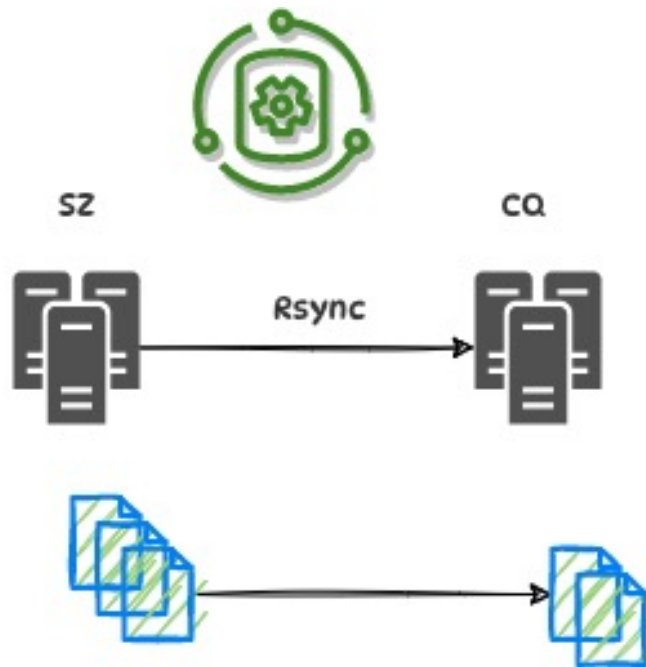


资源时租系统

大规模计算任务普遍具有明显的时间聚集特性，全天资源使用出现波峰波谷现象。凌晨高峰时段资源满负载，任务出现堆积，下午到次日凌晨资源富裕。

时租系统将计算资源使用精确到小时级别，属于一种基于规则的动态扩缩容策略系统。

宏观治理-分级存储



数据治理覆盖范围



元数据治理

数据常见的问题

业务指标口径不统一、业务理解难度大

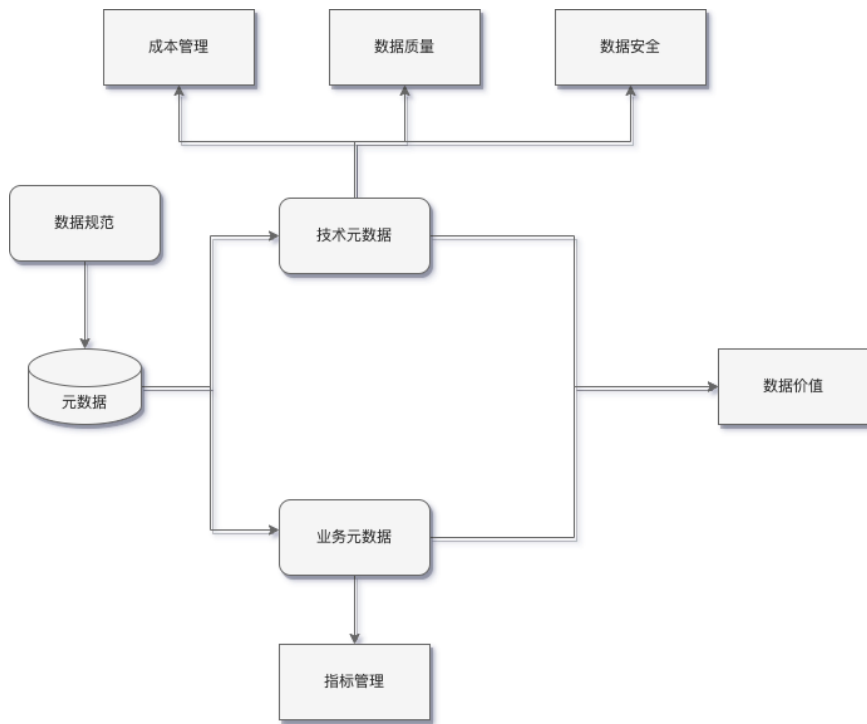
数据质量较差、缺乏有效的监督手段

数据开发丧失资源成本意识，产生冗余存储造成资源浪费

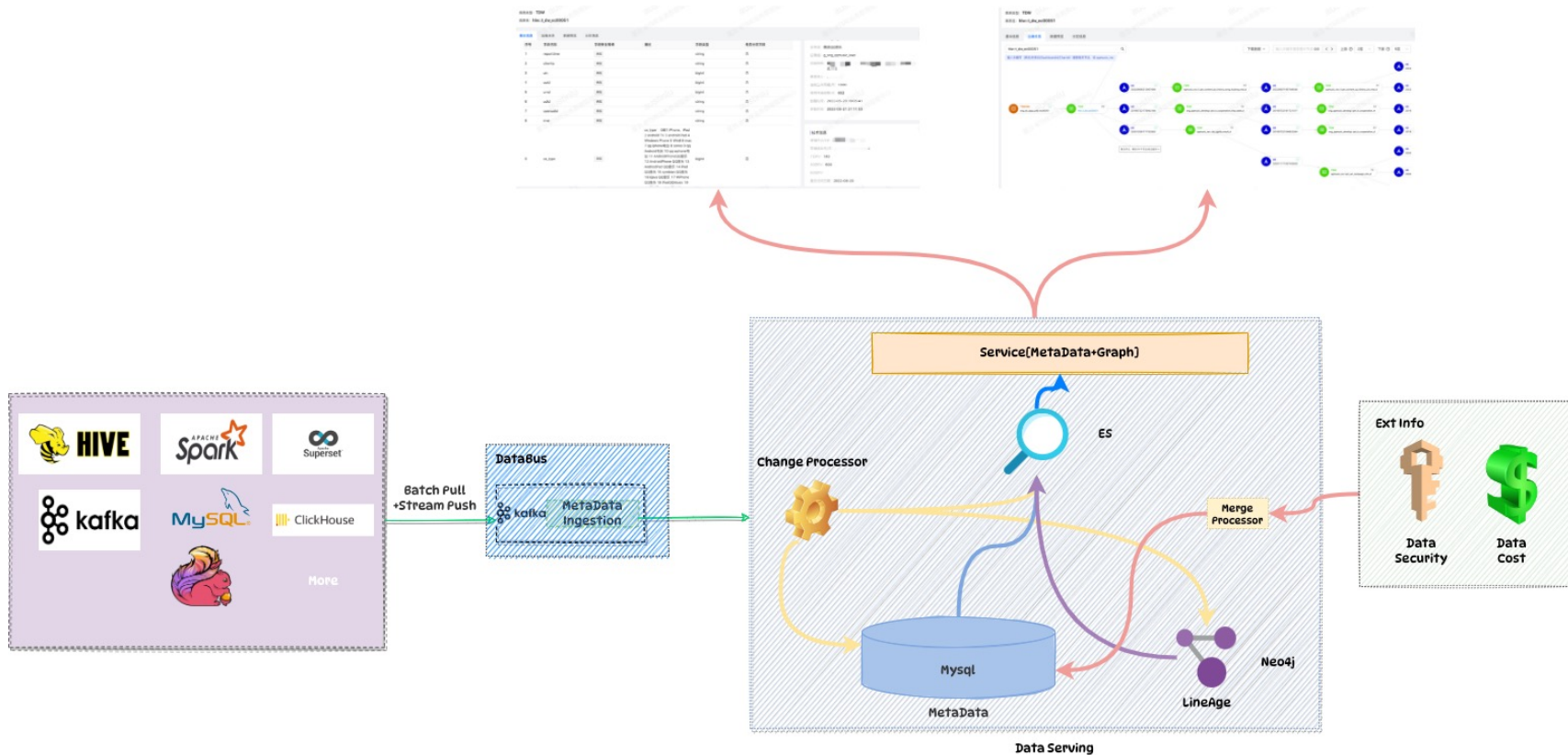
数据缺乏安全级别标识，具有敏感数据信息泄漏风险

缺失数据规范监控，产生大量垃圾元数据，导致元数据管理难度增高

缺乏数据价值量化手段



Lego元数据



是什么

库表类型: TDW

库表名:

基本信息

血缘关系

数据预览

分区信息

复制DDL语句

序号	字段名称	字段安全等级	描述	字段类型	是否分区字段
1		普通		string	否
2		普通		string	否
3		机密		bigint	否
4		机密		bigint	否
5		普通		bigint	否
6		普通		string	否
7		机密		string	否
8		机密		string	否
9		普通		bigint	否

基本属性

唯一id

业务线:

应用组:

组织架构:

表责任人:

当前生命周期/天: 1800

推荐生命周期/天: 602

创建时间: 2022-05-20 19:05:41

更新时间: 2022-08-23 11:37:17

技术信息

存储大小/TB:

存储成本元/月:

7日PV: 140

30日PV: 600

60日PV: 1130

最近访问日期: 2022-08-23

有什么

库表类型: TDW

库表名

基本信息

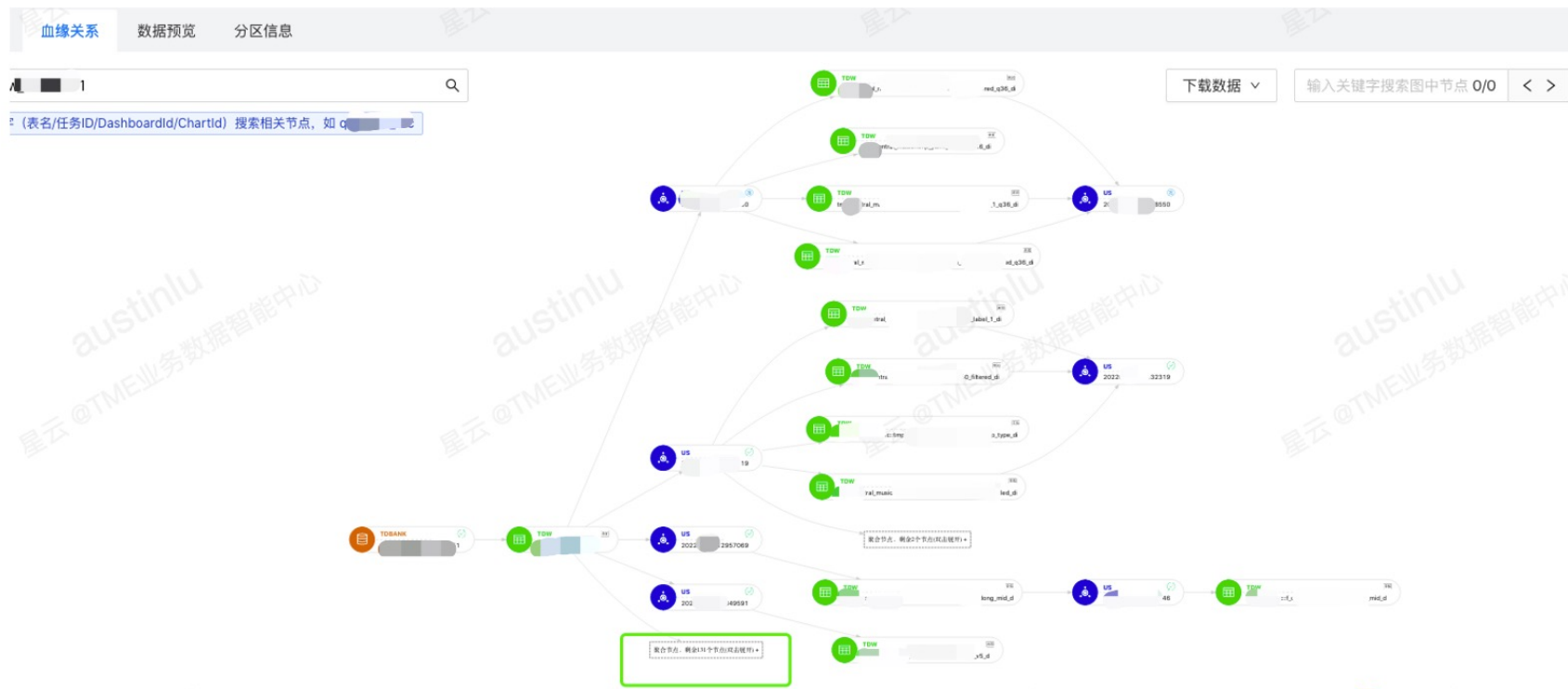
血緣關係

数据预览

分区信息

姓名	工号	性别	出生日期	身份证号	手机号	入职日期	离职日期
王	630154073	0	20201222	20201213			
李	630263527	0	20201222	20201213			
张	1608568589	0	20201222	20201213			11
王	1608570728	0	20201222	20201213			1
王	630264081	0	20201222	20201213			
王	1608571107	0	20201222	20201213			
王	1608570703	0	20201222	20201213			
王	630264092	0	20201222	20201213			1

表的价值



案例



看板ID: 1373

看板名称: W

预估成本/元

健康分: 40

60访问次数: 2

看板创建人:

看板修改人:

创建人组织:

看板创建时间: 2020-04-28 19:05:31

看板修改时间: 2021-08-12 16:28:07

创建时间: 2022-08-24 04:05:19

更新时间: 2022-08-24 04:05:19

资源预警+大盘分析

Warning 应用组:g_

austinlu(卢宇鹏)

昨天 下午 5:35

DI小助手邀请

和你加入了群聊

austinlu(卢宇鹏)

DI小助手 (DI小助手)

⚠ 应用组计算资源利用率较低提示 ⚠

应用组: 集群 (sql默认集群)

集群: g_

在过去15天中, 有10天平均内存利用率低于 32.2 % 为提升资源利用率, 您可以:

👉 启用时租系统, 了解时租: <https://wiki.wondershare.com/pages/viewpage.action?pagelId=4006714080>

👉 前往安全中心, 回收闲置资源: <https://security.wondershare.com/auth/group>
详情咨询:

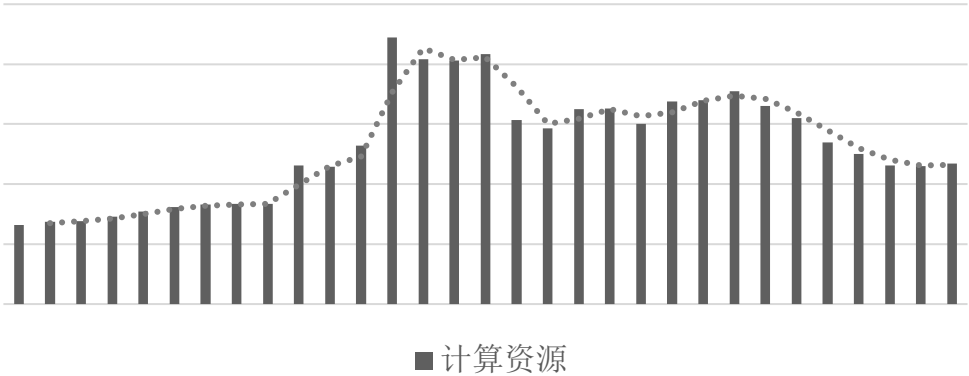
austinlu(卢宇鹏)

03 效果

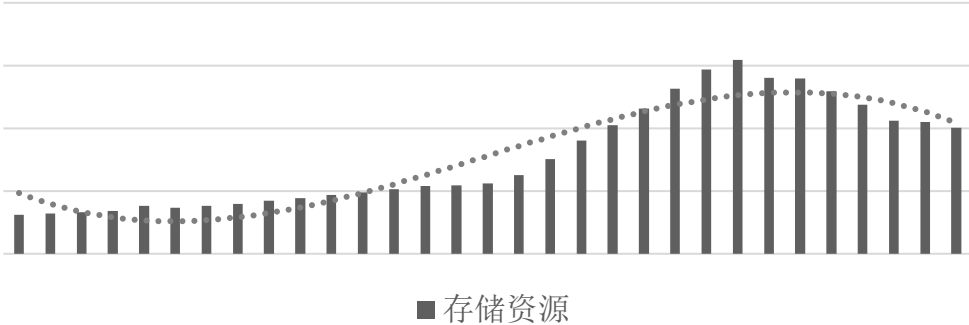


效果

治理后-计算资源增长趋势



治理后-存储资源增长趋势



表总量

■
-7.6% 月环比



非常感谢您的观看

