实施过程遇到的问题与对策

跨节点的数据网络传输带来的IO成本

问题: IO成为性能瓶颈

对策:客户号存在并作为分片键,并且在表关联

操作中将客户号相等的计算条件作为必要条件;

元数据表尽可能设计为复制表;

避免在分片键上出现空值,导致数据分布过度倾

斜



避免复杂的SQL编写

问题:编译器难以判断嵌套SQL在每个节点上的初始结果集是否可以驻留直接使用,结果往往需要汇聚后再广播给每个节点,增加大量的网络IO成本

对策:尽可能把过滤条件放到嵌套SQL内部执行,减少中间结果集大小,降低数据广播带来的处理延迟

分布式计算任务带来的调度协调成本

问题:看似较小成本的SQL实际执行成本(时间)

比传统数据库要高很多

对策:包括规避存储过程内游标使用,尽可能用聚

合成复合SQL语句或拆分成若干批量数据更新处理

步骤





技术支持工具

<mark>问题</mark>:跨集群数据自动复制,系统监控、

SQL性能分析诊断等技术工具尚不完善。

对策: 与平台厂商合作, 自开发技术工具

作为补充



基础数据平台技术优化工作



> 数据权限管理

1.多分支机构行级权限管控 2.列级权限,用户级数据脱 敏定义

> 服务质量管控

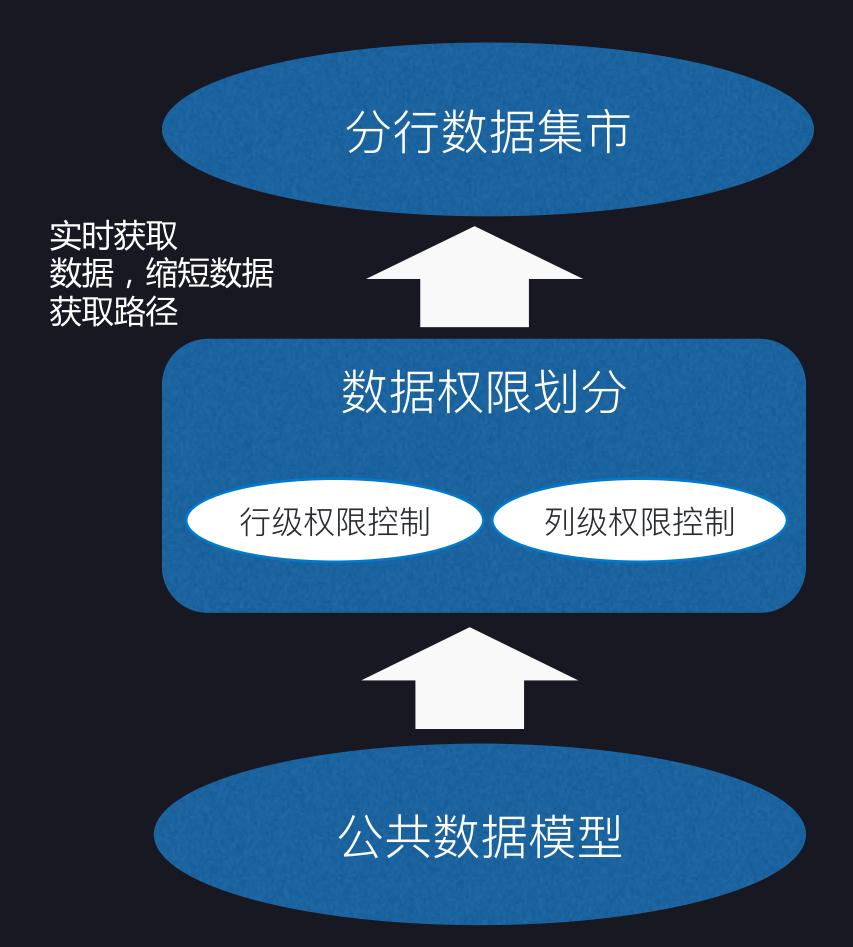
1.任务级资源管控避免不良设计 影响整体性能 2.SQL执行成本、执行时间、排 队时间等多维质量管控策略

>实时监控预警

- 1.针对实例和组件不同层级的实时监控 搭建
- 2.实时采集组件可用性、资源占用情况、任务排队数、平台事件等信息 3.配置智能告警规则



数据权限管控



行级权限

- ●行级权限实现数据的行级授权
- •分行用户只能查询该分行数据
- ●基础模型统一口径加工

列级权限

- •列级权限实现数据脱敏
- •针对不同用户设定不同的查询结 果



集群灾备以及数据同步



在线数据平台

- 1. 在线跑批集群
- 2. 数据集市
- 3. 对接业务系统
- 4. 提供数据服务

同步



同步

灾备在线数据平台

- 1. 灾备在线集群
- 2. 可做灾备切换



灾备历史数据平台

- 1. 灾备历史集群
- 2. 可做历史集群切换

集群间日常增量同步

历史数据平台

1. 历史数据备份

2. 提供历史数据服务

- 1. 针对ORC类型表,从在线数据平台每日获取增量数据,同步到其他三套数据平台
- 2. 针对Hbase表类型,使用Hbase表同步功能进行准实时同步

数据 同步

集群间全表同步

- 1. 针对ORC类型表,从在线数据平台获取表结构及全表数据,同步到其他三套数据平台,并保留历史平台的授权情况
- 2. 针对Hbase表类型,使用Hbase表同步功能进行整表数据同步

集群间区间同步

1. 从在线数据同步获取需同步区间的增量数据,同步到其他三套数据平台

技术实施关键点

性能优化设计

- 1.建立性能设计规范 2.始于开发阶段的性能监控 3.分布式数据平台与传统数据
 - 平台性能问题的异同点

应用项目管理

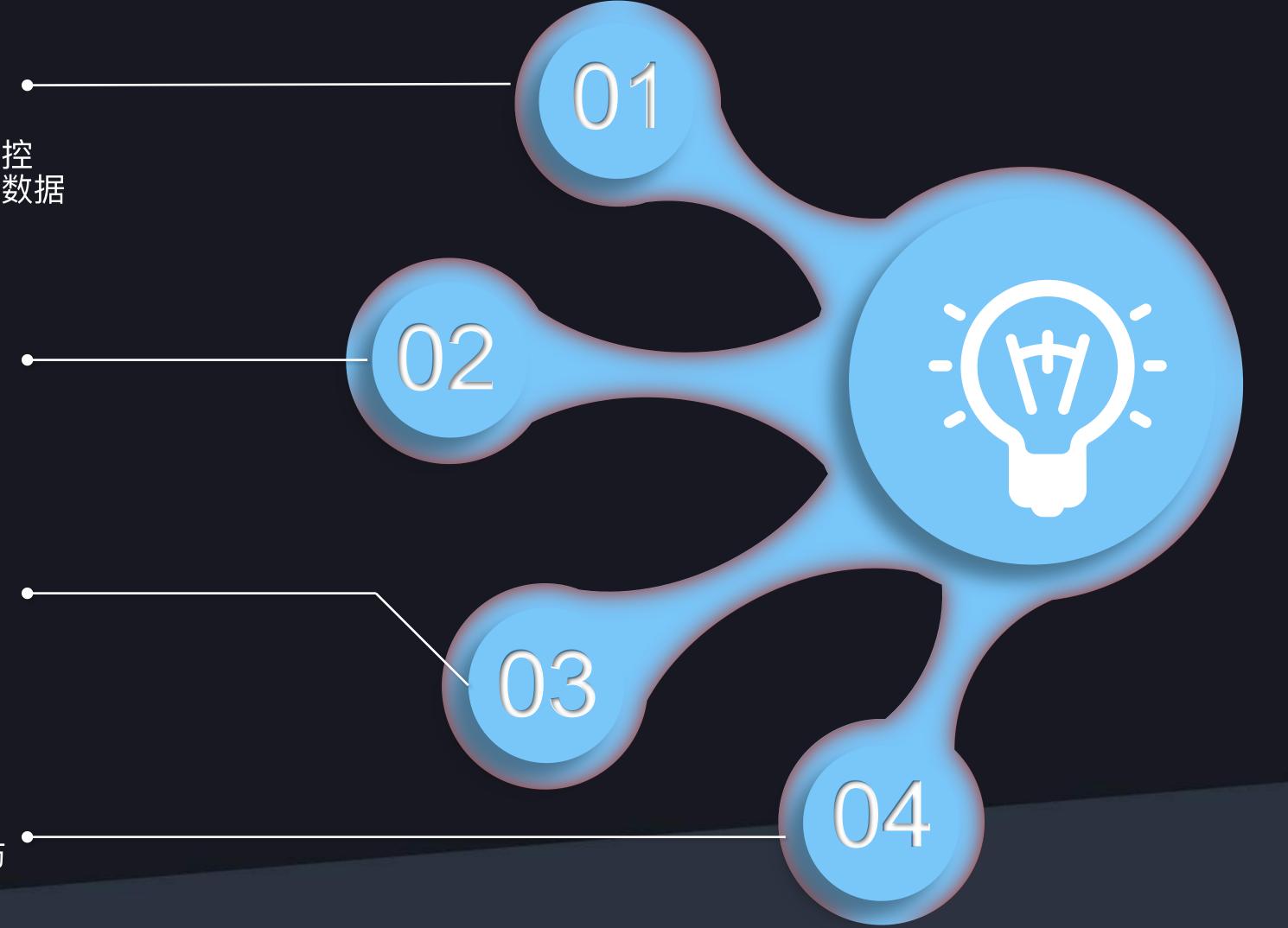
- 1.数据模型设计管控 2.资源按需弹性分配
- 3.批处理任务集中调度

> 知识传递

- 1.细分主题多波次培训
- 2.成立技术兴趣小组
- 3.技术反模式案例总结

> 数据质量管控

1.自动化数据质量检核 2.构建加工路径上多个系统相对独立的数据质量防火墙

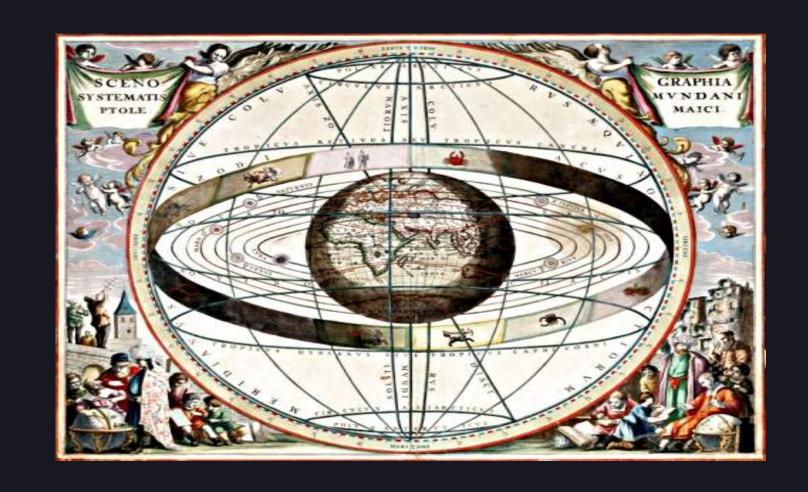




目录

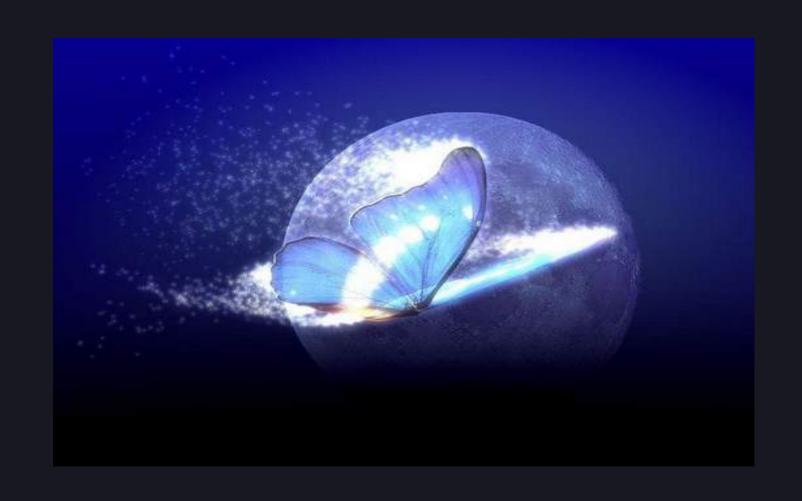
- 1 平台建设背景
- 2 数仓应用体系建设
- 3 风控领域创新应用
- 4 取得成果与未来展望

思路-风险管理的两种思考体系



传统思维方法

- 1.需要准确财报数据才能判断企业经营状况
- 2.选取合适的变量和模型,通过对历史数据的拟合验证,可以预测客户是否将出现信用违约
- 3.评审与贷后环节对人的判断能力非常依赖
- 4.外部环境复杂多变,系统性风险难以防范



大数据思维方法

- 1.大部分客户财报数据不准确,是否能从公开信息变化辅助判断客户经营是否正常
- 2.信用违约的原因很复杂,没有足够多数据支撑模型精准预测,但 多个关联信息可以推测违约概率是否在变大
- 3.多个环节引入并更多的客观数据可减少人为判断风险
- 4.可以建立数字化监控体系来对系统性风险感知预警



风险管理应用规划

行业数字地图:

- 1、自动遴选、组织数据,提升规划效能
- 2、量化分析提升授信决策效率

3、感知行业趋势变化,敏捷应对系统性风险

客户交易风险管理:

- 1、流处理技术实现实时欺诈风险识别
- 2、自主学习和优化的反欺诈模型
- 3、识别和阻断事中风险

柜员操作风险管理:

- 1、智能决策引擎实时产生预警工单
- 2、自动化风险审查作业模式
- 3、提升内审作业效率,防范道德风险

业务规划平台

作系 统 大数据风控

风险

监测系统

反欺

信风险额系统

移动

信贷

应用

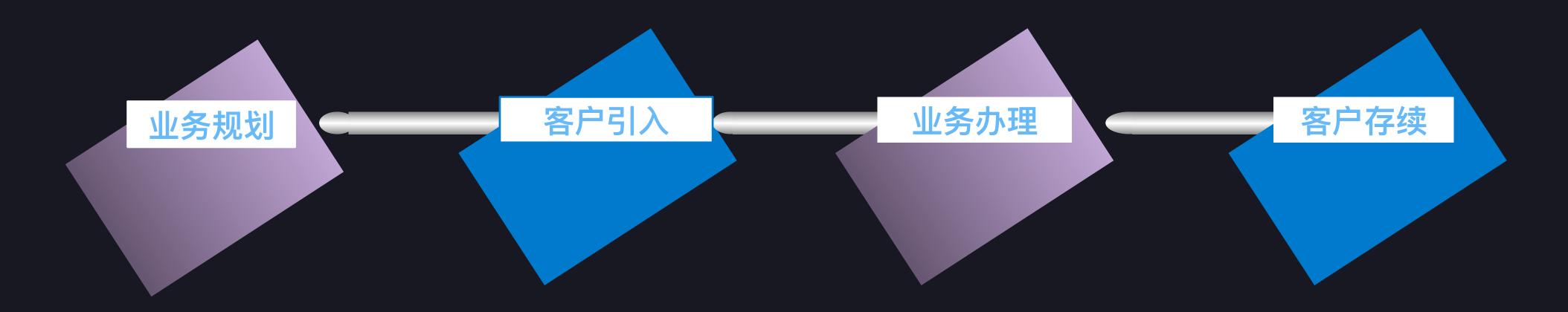
业务协作、风控前置:

- 1、便捷采集申请信息,提升流程效率
- 2、自动交叉验证外部数据
- 3、减少人为判断,规避人员道德风险

全面的信用风险预警服务:

- 1、多维风险视图提升决策效率
- 2、知识图谱技术分析风险传导,提升 风险感知能力
- 3、智能决策引擎汇聚专家智慧,降低人员要求
- 4、信用欺诈检测有效防范道德风险

风险管理全生命期系统支持



业务规划平台:

- 1.区域行业发展 趋势分析
- 2.行业龙头企业动态跟踪
- 3.潜在客户智能检索
- 4.授信政策制定

移动信贷应用:

- 1.目标客户信息 采集
- 2.贷前风险排查
- 3.额度测算

信用风险预警:

- 1.企业风险视图
- 2.简易评分卡
- 3.风险预警信号

信用风险预警:

- 1.客户风险预警
- 2.评分卡
- 3.风险决策树
- 4.模型与策略
- 5.信用欺诈检测

运营风险监测:

1.柜员操作风险

交易反欺诈:

1.客户交易风险

信用风险预警:

- 1.客户风险预警
- 2.组合风险预警
- 3.信用违约预测
- 4.风险影响分析
- 5.风险缓释

移动信贷应用:

贷后检查

业务规划平台

授信政策重评估



风险预警系统-业务能力规划





基本数据



衍生数据



联机接口



风险防控规则



风险分析



决策支持

基础数据支撑服务

统一联机接口服务

统一风险防控规则配置服务

综合风险分析、决策支持

分区域行业宏观分析

客户风险综合分析

风险事件跟踪分析

历史信息趋势分析

风险演变分析

风险预测及其跟踪验证分析

