

数字转型 架构演进

SACC

2019 中国系统架构师大会

SYSTEM ARCHITECT CONFERENCE CHINA 2019



2019年10月31-11月2日



北京海淀永泰福朋喜来登酒店



全新IT技术私域交流平台

ApacheShardingSphere数据平台 演进之路

京东数科-潘娟
panjuan@apache.org



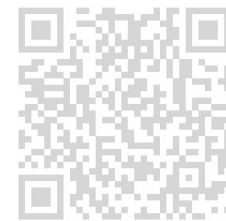
全新IT技术私域交流平台



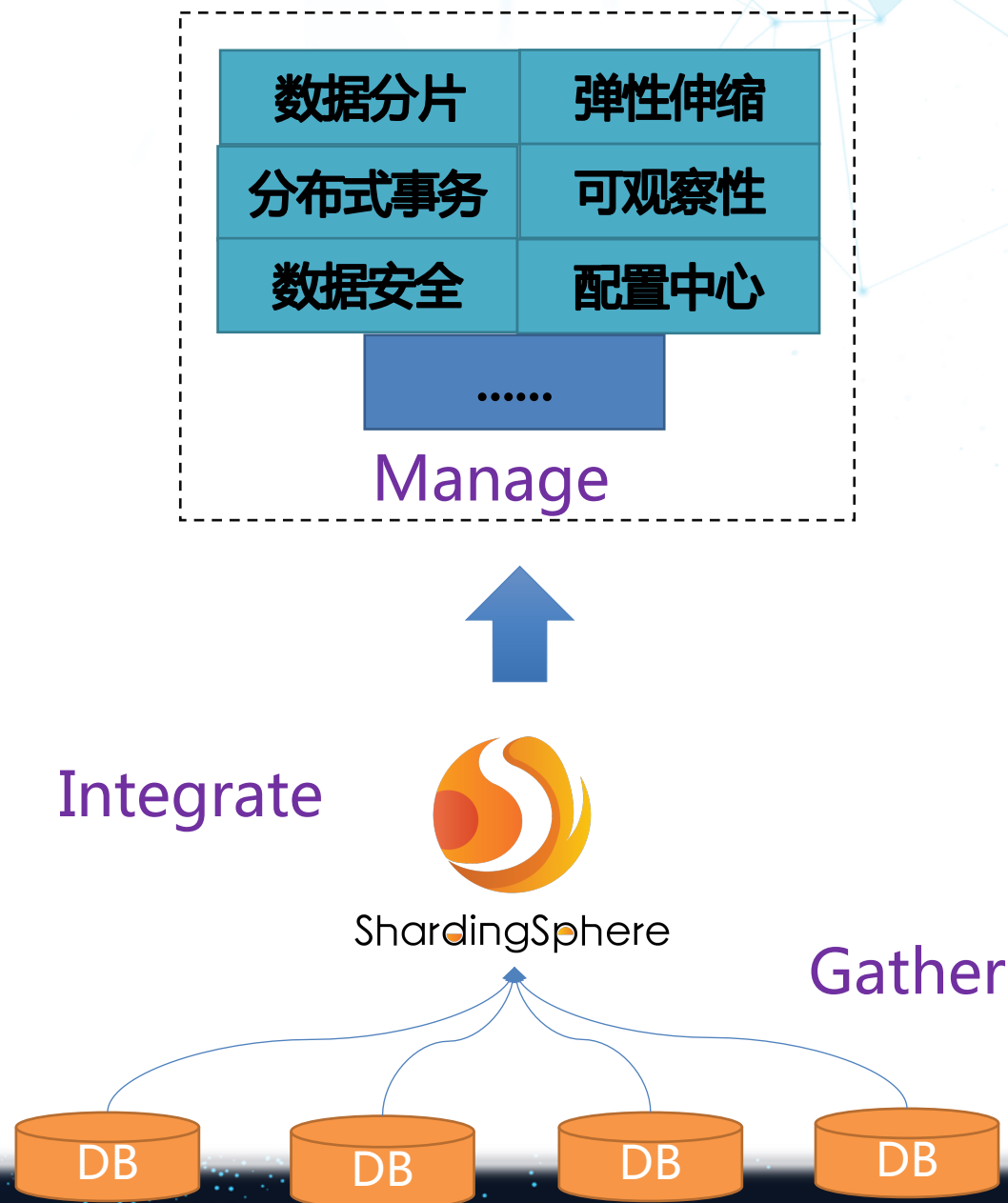
潘娟

京东数科高级DBA&Apache ShardingSphere PPMC

主要负责京东数科数据库中间件开发、数据库运维自动化平台开发、生产数据库运维工作。多次参与京东6.18、11.11等大促活动的护航工作。曾负责京东数科数据库自动化平台设计与开发项目，现专注于Apache ShardingSphere分布式数据库中间件开发和发展。乐于在数据库、分布式、中间件、开源社区等相关领域进行学习和探索。多次受邀参加数据库&架构领域的相关会议并进行分享交流。也多次进行在线技术和经验交流分享、公众号文章分享。



全新IT技术私域交流平台



架构



ShardingJDBC

接入端

ShardingProxy

分库分表

读写分离

数据分片

两阶段事务

柔性事务

分布式事务

弹性伸缩

配置中心

数据安全

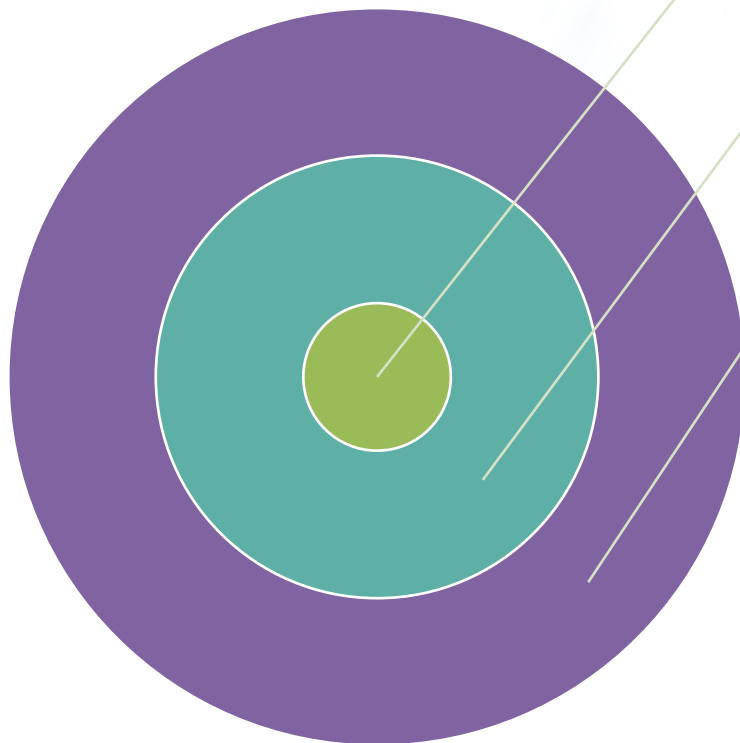
可观察性

分布式治理



数据库集群

弹性扩展



数据库

- 计算
- 存储

ShardingSphere

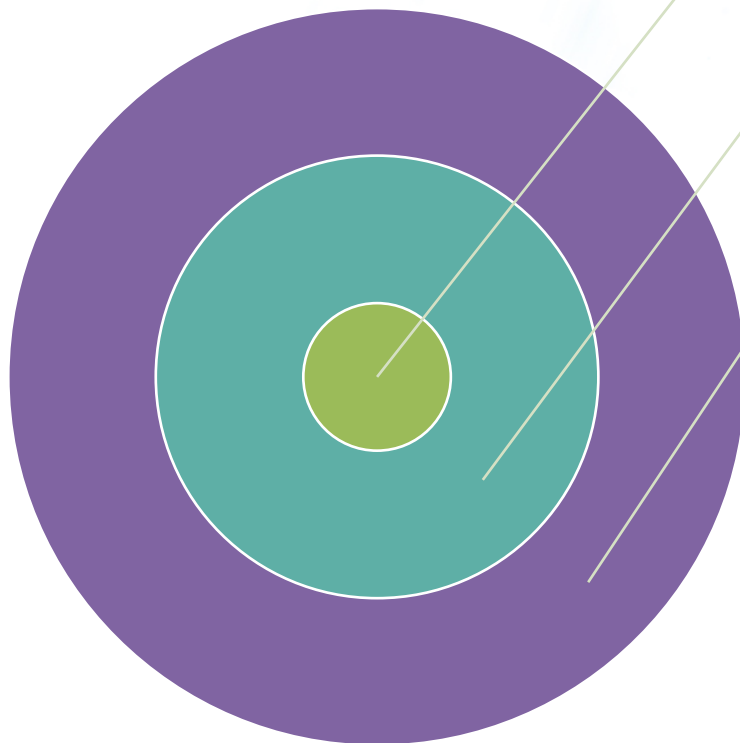
- SQL解析&改写&路由
- 分布式事务拓展
- 高可用
- 数据脱敏
- 可观察性
- 弹性伸缩
-

- ◆ 充分利用数据库的存储能力
- ◆ 在保证查询结果正确的前提下，充分利用数据库的计算能力。

例如：SELECT AVG(score) FROM student;

$$AVG(score) = avg_1 + avg_2 + avg_3$$

$$AVG(score) = \frac{SUM(score)}{COUNT(score)}$$



数据库

- 计算
- 存储

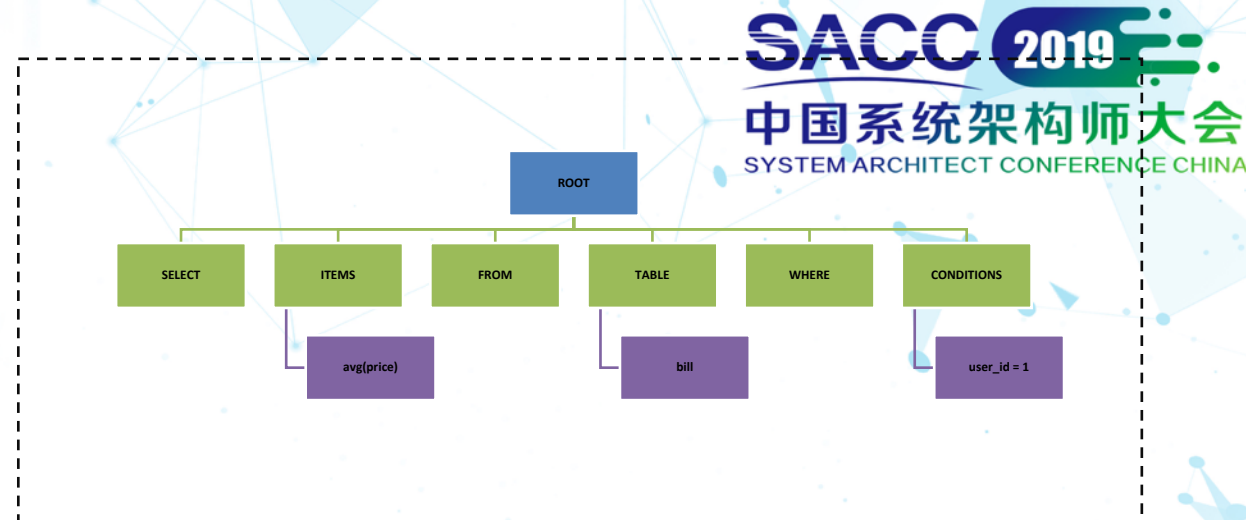
ShardingSphere

- SQL解析&改写&路由
- 分布式事务拓展
- 高可用
- 数据脱敏
- 可观察性
- 弹性伸缩
-

SQL理解

```
SELECT AVG(price) FROM bill
WHERE user_id = 1
```

解析



规则

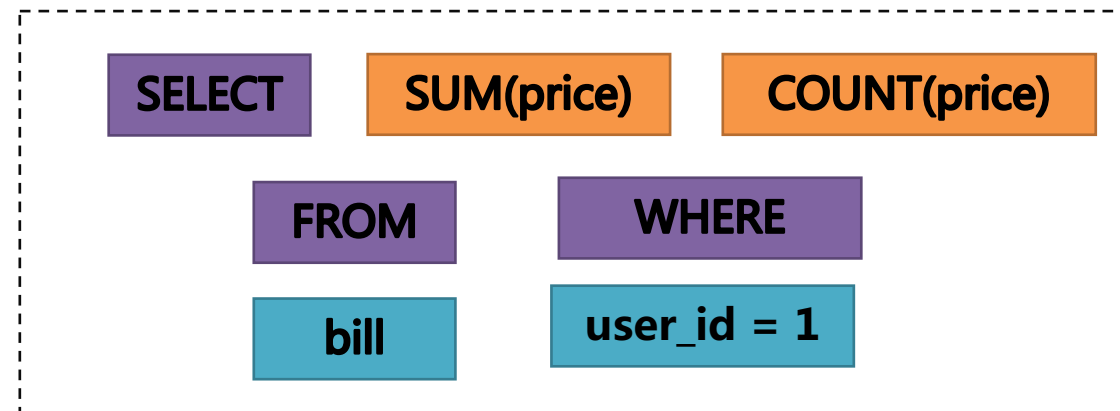
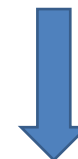


```
SELECT SUM(price)/COUNT(price)
FROM bill_0
WHERE user_id = 1
```

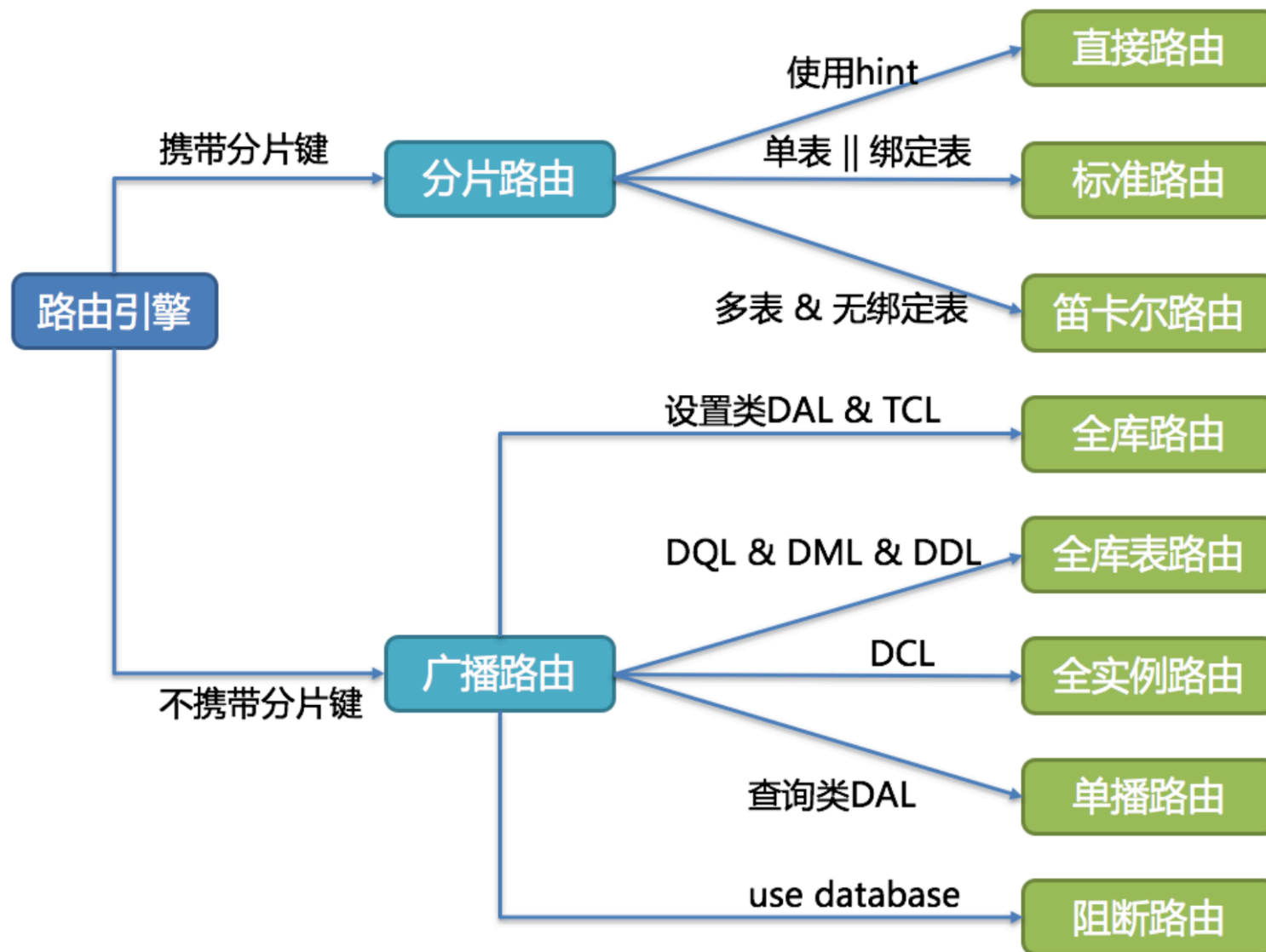
改写



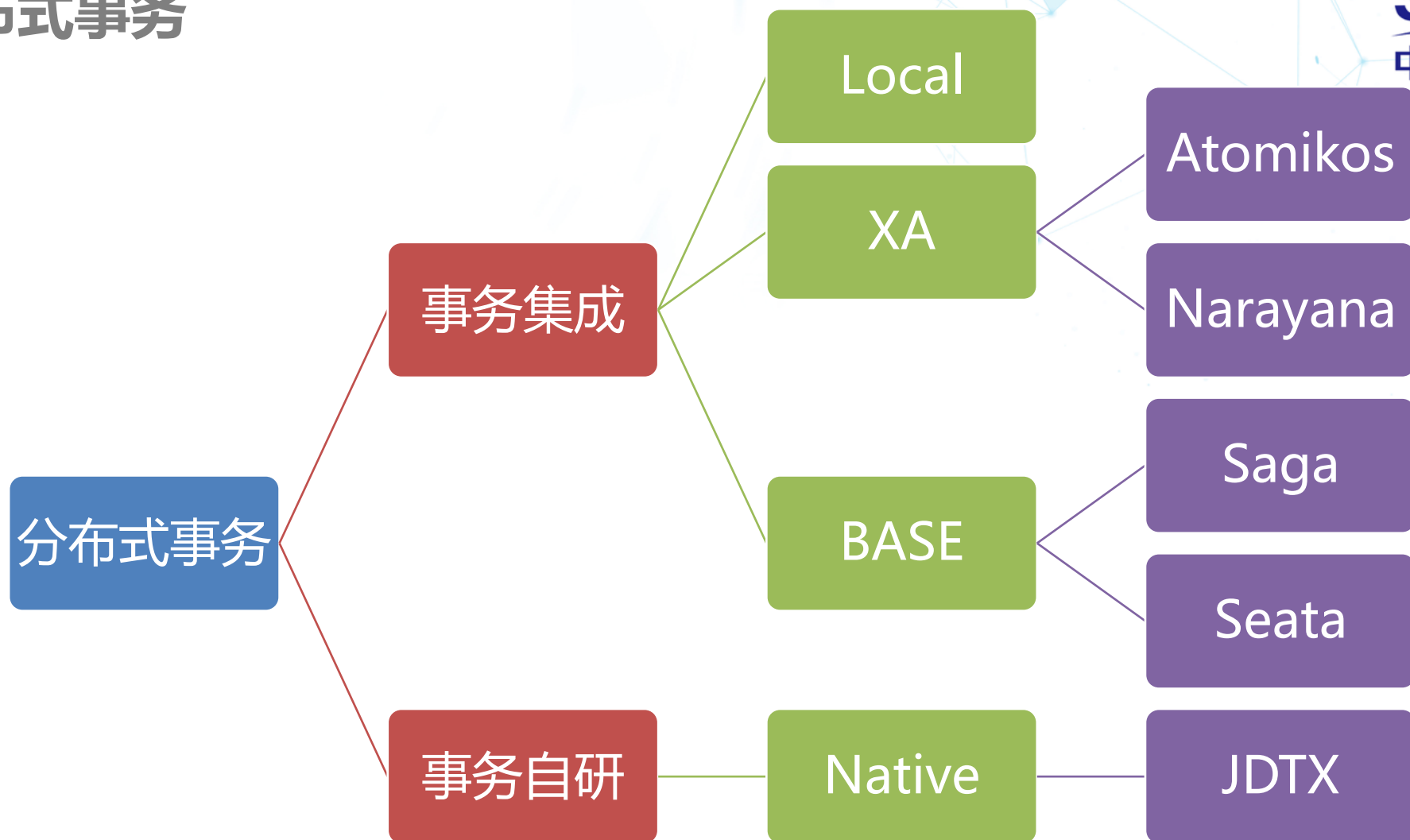
处理



	方案	优点	缺点
第一代解析引擎	Druid SQL Parser	支持多数据库 解析效率较高	新旧版本不兼容 SQL解析引擎无法 自主可控
第二代解析引擎	自研的“半理解” 方式解析器	支持常见的关系型 数据库 解析效率最高	无法实现 100%SQL兼容 维护成本高
第三代解析引擎	ANTLR语言解析 器	自定义解析规则 实现100%SQL兼 容	穿透缓存的情况下 解析效率略低



	本地事务	两阶段事务	柔性事务
业务改造	无	无	实现相关接口
一致性	不支持	强一致	最终一致
隔离性	不支持	原生支持	服务保证
并发性能	无影响	严重衰退	略微衰退
适合场景	业务方处理不一致	短事务/低并发	长事务/高并发
事务层级	数据库层	数据库层	服务层

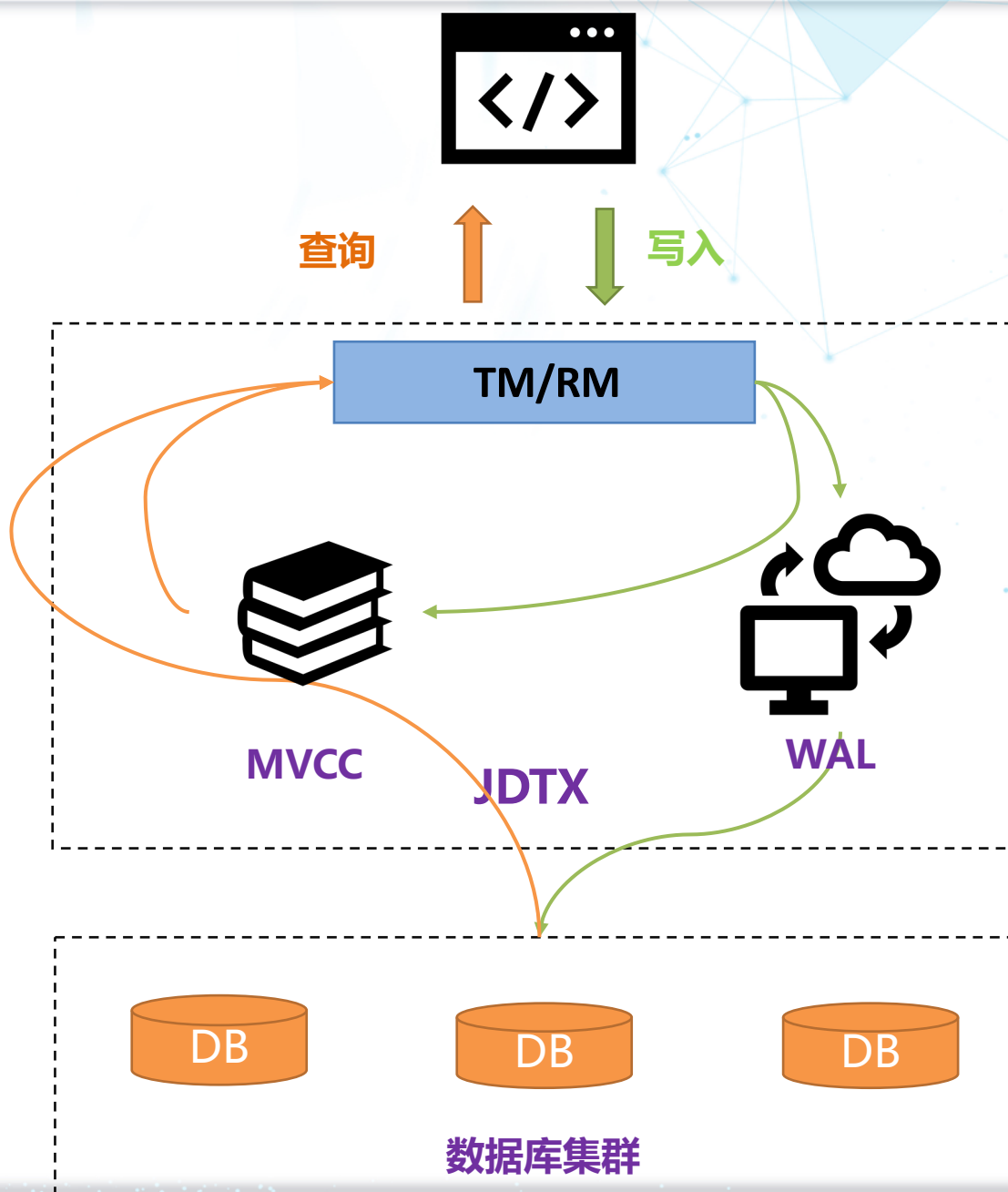


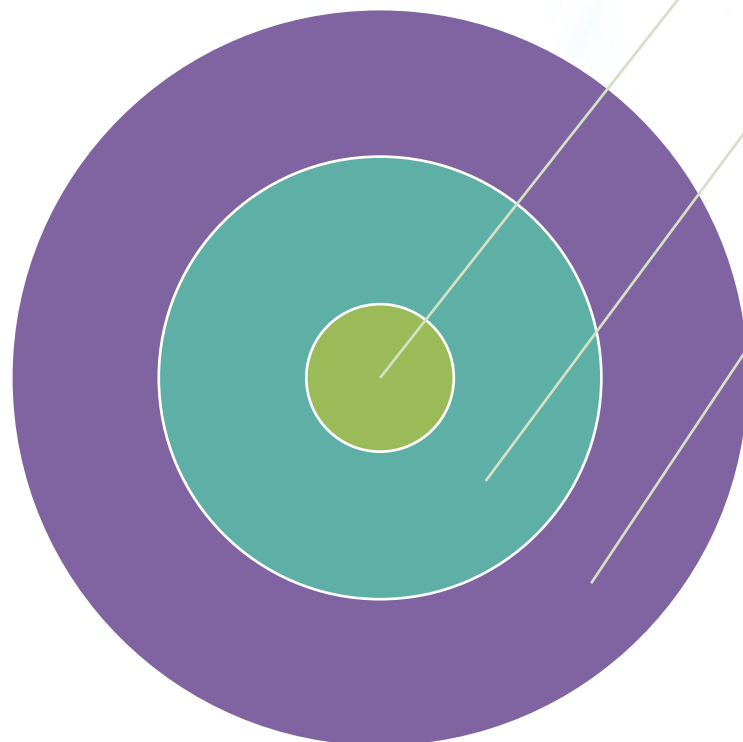
Saga : <https://github.com/apache/servicecomb-saga-actuator>

Seata : <https://github.com/seata/seata>

优势：

- ◆ 透明化底层分布式事务实现细节，用户可使用标准API开启分布式事务。
- ◆ 支持同时使用数据分片，分布式事务等。
- ◆ 依据实际场景，用户可选择使用不同类型的分布式事务。





数据库

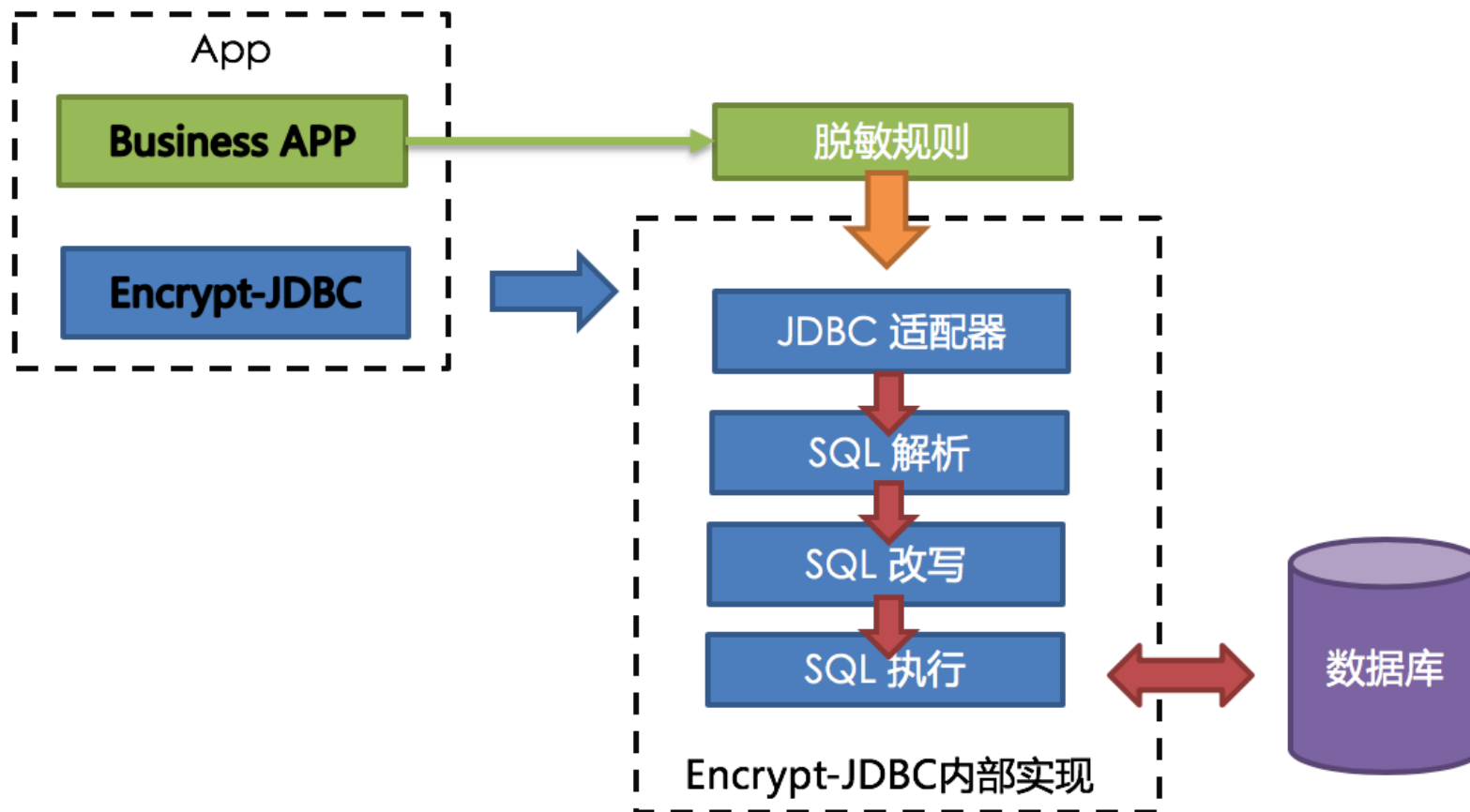
- 计算
- 存储

ShardingSphere

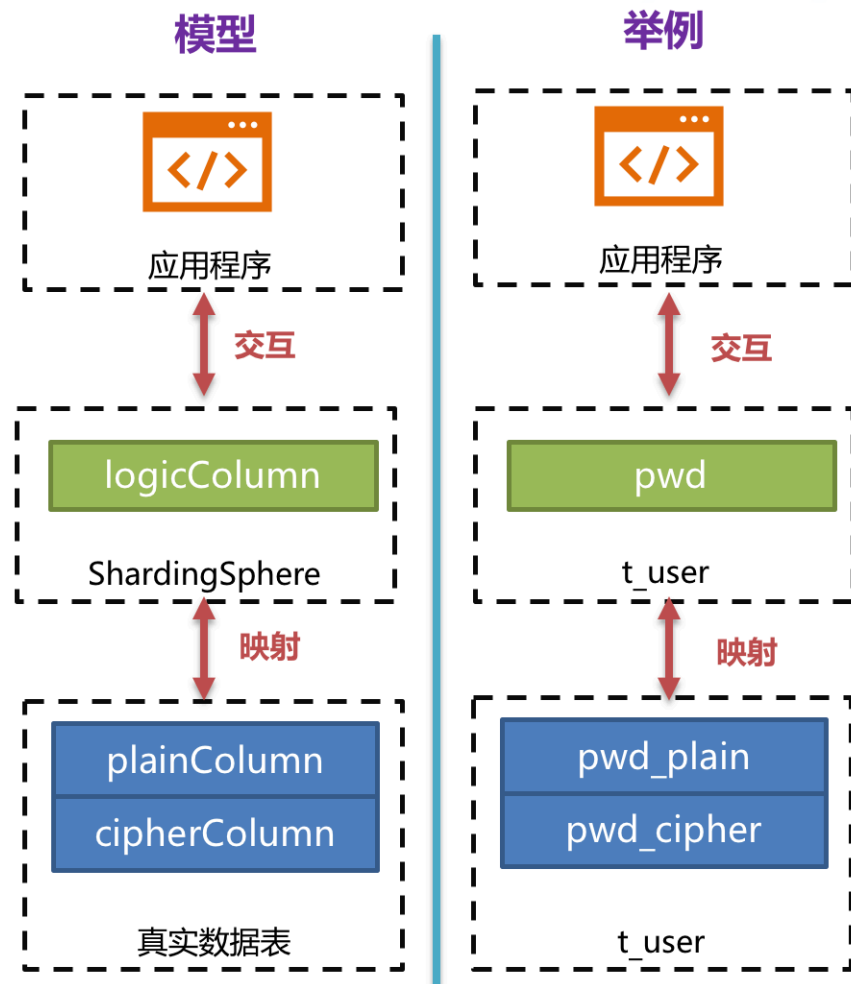
- SQL解析&改写&路由
- 分布式事务

拓展

- 高可用
- 数据脱敏
- 可观察性
- 弹性伸缩
-



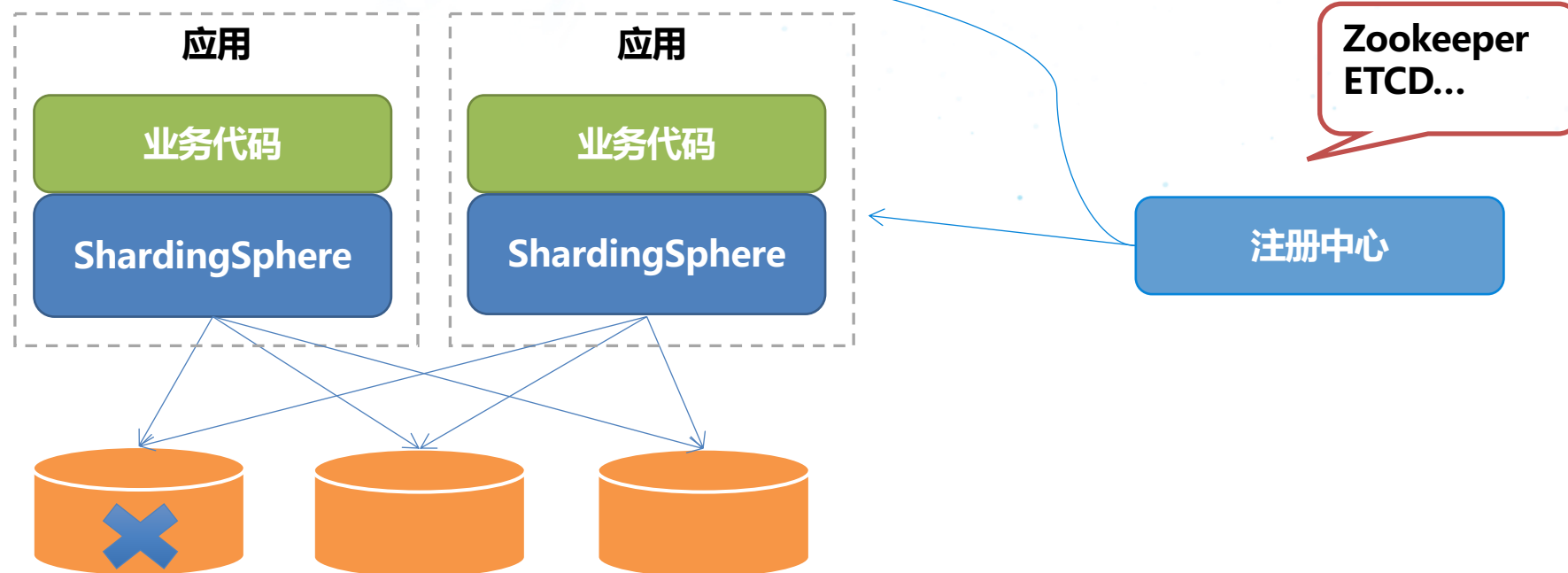
参考文章：https://mp.weixin.qq.com/s/_wfG_gQFwpvaMVwwCQVpfQ

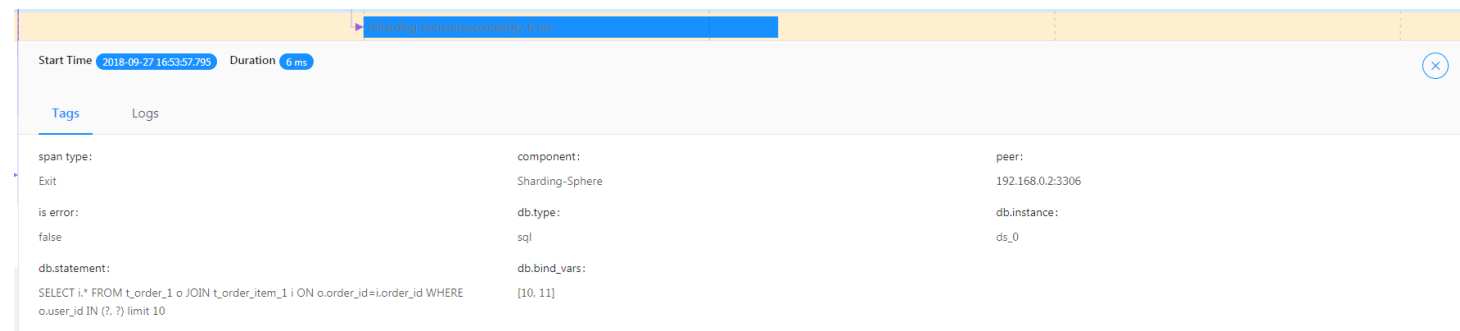
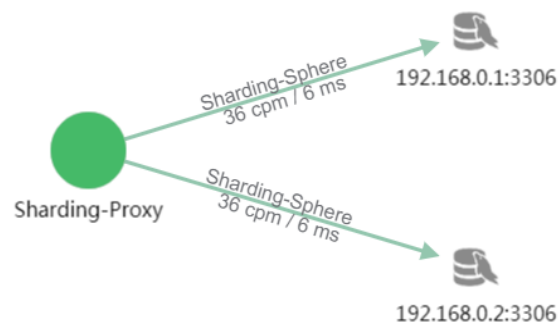
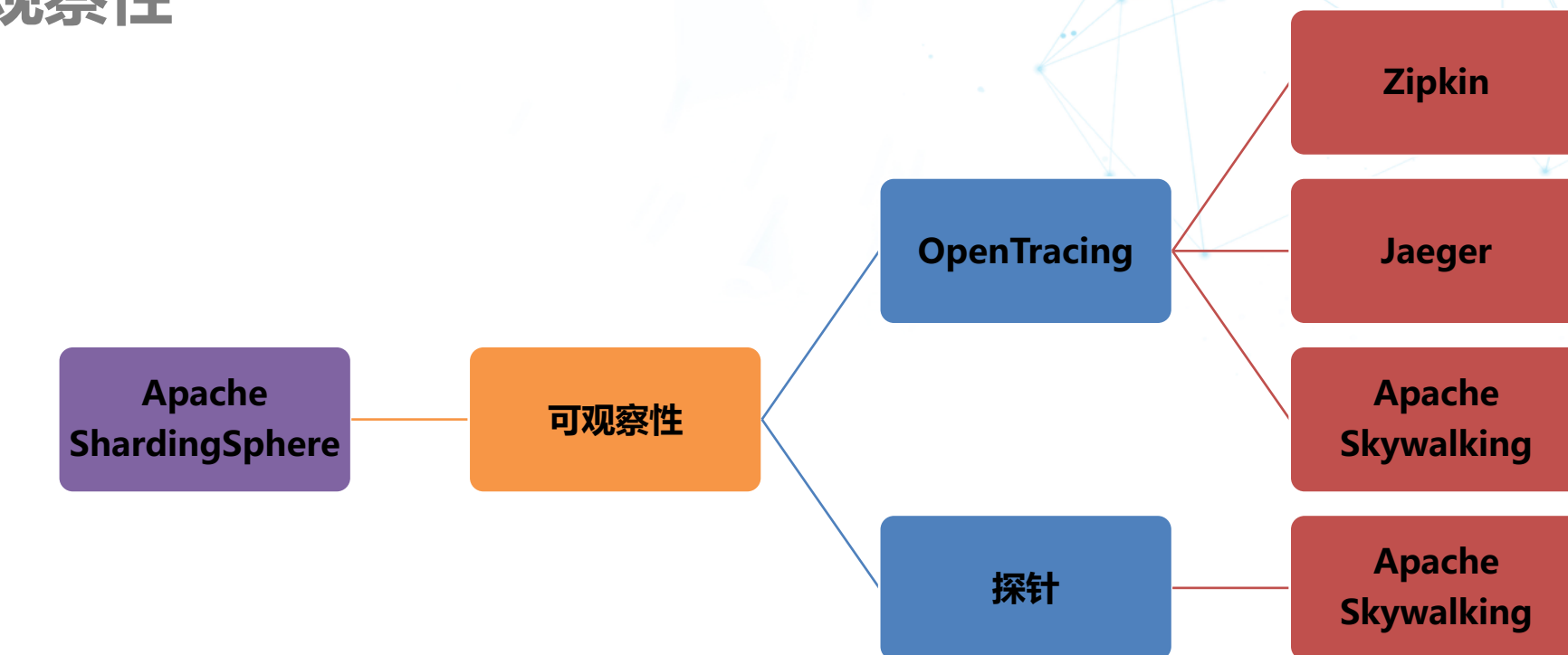


中间件脱敏服务优势

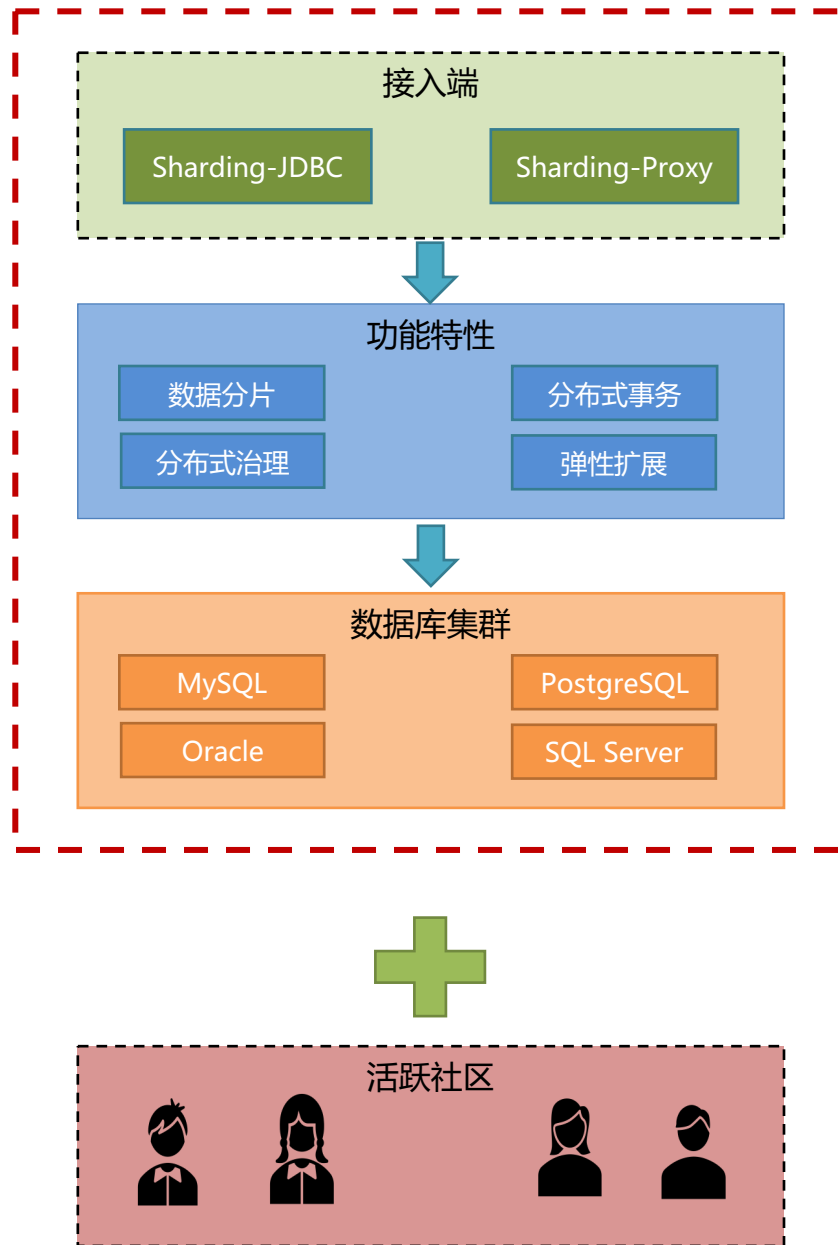
1. 自动化&透明化数据脱敏过程，用户无需关注脱敏中间实现细节。
2. 提供多种内置、第三方脱敏策略，用户仅需简单配置即可使用。
3. 针对已上线业务，可实现明文数据与密文数据同步存储，并通过配置决定使用明文列还是密文列进行查询。可实现在不改变业务查询SQL前提下，已上线系统对加密前后数据进行安全、透明化迁移。

- ✓ 配置动态化推动
- ✓ 配置统一管理
- ✓ 数据库从库禁用
- ✓ 应用实例熔断





组成



Apache
ShardingSphere
(开源分布式数据库中间件平台)

贡献社区

来源

- 云公司开源战略
- 业务系统采用
- 个人兴趣提升

工作与开源

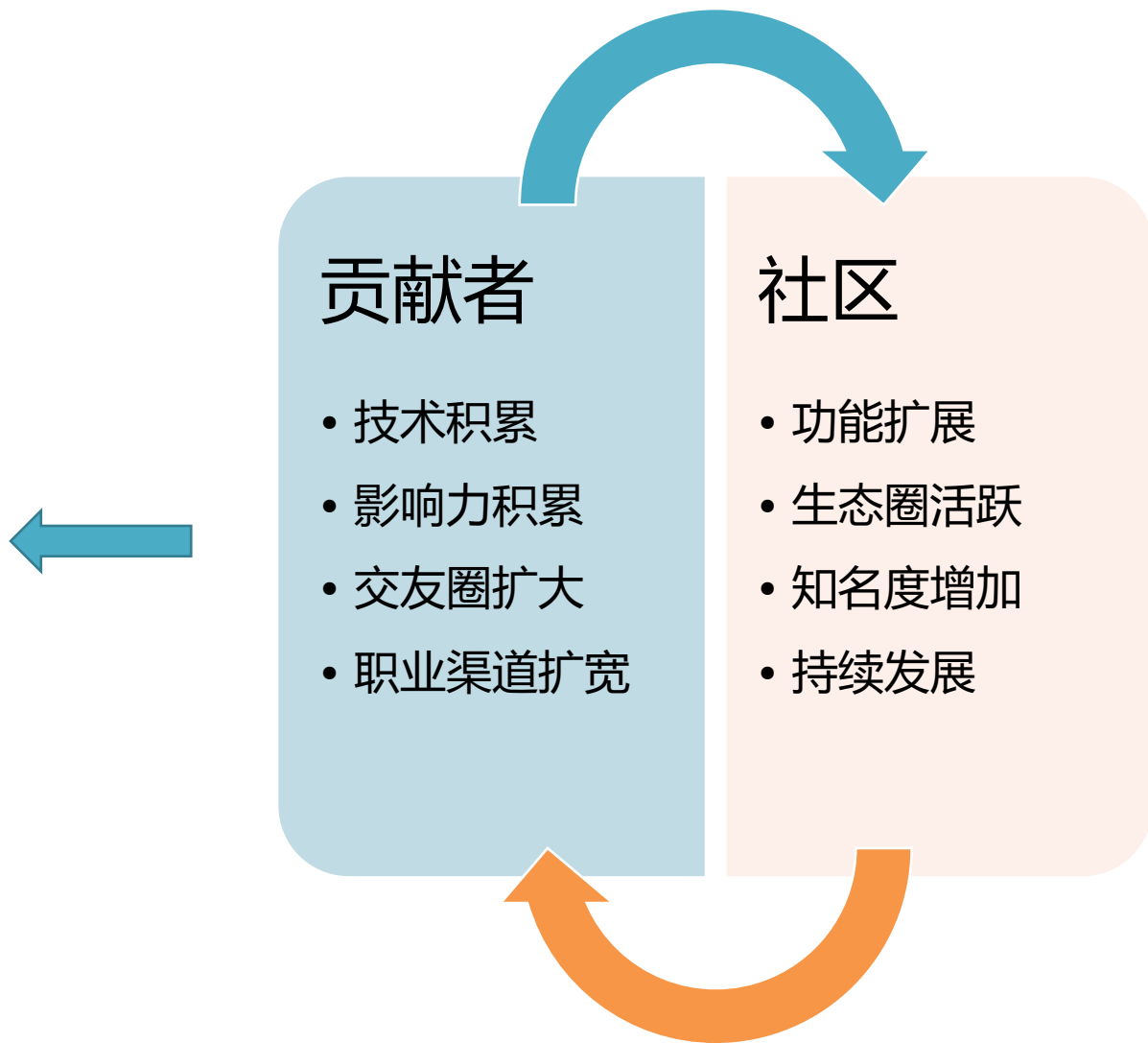
- 相互促进
- 弹性控制
- 兴趣与目标驱动

贡献者

- 技术积累
- 影响力积累
- 交友圈扩大
- 职业渠道扩宽

社区

- 功能扩展
- 生态圈活跃
- 知名度增加
- 持续发展



成为Committer

SACC 2019
中国系统架构师大会
SYSTEM ARCHITECT CONFERENCE CHINA



◆ GitHub

<https://github.com/apache/incubator-shardingsphere>

◆ HomePage

<https://shardingsphere.apache.org>





Thanks