

2018年PostgreSQL中国技术大会



基于Odyssey连接池实现企业级 PostgreSQL数据分布中间件

权宗亮

zongliang.quan@postgresdata.com

神州飞象(北京)数据科技有限公司

PgBouncer

- 单进程
 - 性能瓶颈
- 新功能
 - 负载均衡
 - 读写分离

负载均衡

- 分组
 - 多个主节点
 - 权重
 - 随机

依赖客户端的读写分离

- JDBC setReadOnly方法
 - 客户端调用此方法发出命令
 - SET SESSION CHARACTERISTICS AS TRANSACTION READ ONLY
 - PgBouncer收到后将后续的事务作为只读处理
 - 需要应用程序修改设置或者代码
- 简单方法
 - JDBC设置只读发出固定语句
 - 应用框架
- 缺点
 - 方式僵化
 - 必须修改应用

基于语法解析的读写分离

- 类似pgpool的实现方式
 - 单进程性能瓶颈，语法解析引擎开销过大
- 单个语句事务
- BEGIN语句
 - SERIALIZABLE
 - READ WRITE
 - READ ONLY

基于语法解析的读写分离

- 复合语句事务
 - 主备库双事务
 - 遇到写语句关闭备库事务
 - 或者直到结束所有事务
 - 无法保证两个事务快照完全一致
 - 两个事务管理复杂
 - 主库有额外开销

Odyssey

- Yandex开源项目
- <https://github.com/yandex/odyssey>
- 多线程
- 更好的连接池控制



区分业务会话

- 业务系统的事务
 - SQL语句比较固定
 - 偶尔会有些管理员或者监控SQL
 - 单语句事务比较少
 - 第一条语句一般是 BEGIN
 - 第二条语句基本不同
- 区分
 - BEGIN之后的第二条语句基本可以区分不同的事务
 - 如果不能，使用hint
 - BEGIN ISOLATION LEVEL READ ONLY/READ WRITE/SERIALIZABLE

跟踪SQL语句序列的读写分离

- 会话库
 - 跟踪所有事务，记录SQL语句
 - 统计使用频度
 - 数据源可选SQLite/PG

采集阶段

- 采集时间
- 采集次数
- 25006错误: `ERRCODE_READ_ONLY_SQL_TRANSACTION`
 - 备库执行写操作
- 语句同时发往主备库
 - 识别 25006, 标记为写

错误校正

- 写事务错标为只读
 - 采集周期过短，样本不全
 - 应用修改
- 自动校正
 - 识别到 25006
- 无法校正错标为写的事务
 - 手工修改

语法分析与语句跟踪的比较

- 语法分析
 - 发现未识别语句需要修改代码
- 语句跟踪
 - 开销更小，flex+bison 不是效率很高的引擎
 - 人工干预简单

Thanks

