# 概述

多源同步与MySQL的主从复制不同，多源同步是在不同数据库之间的数据迁移，可以是基于消息队列实现生产者-消费者模型实现（此时主从复制的binlog方式不一定可行）。

# 原理

## 基于GTID多源复制

## 基于消息队列多源同步

# 工具

## OGG

Oracle GoldenGate是一款实时访问、基于日志变化捕捉数据，并且在异构平台之间迚行数据传输的产品。GoldenGate TDM是一种基于软件的数据复制方式，它从数据库的日志解析数据的变化（数据量只有日志的四分之一左右）。GoldenGate TDM将数据变化转化为自己的格式，直接通过TCP/IP网络传输，无需依赖于数据库自身的传递方式，而且可以通过高达10:1的压缩率对数据迚行压缩，可以大大降低带宽需求。在目标端，GoldenGate TDM可以通过交易重组，分批加载等技术手段大大加快数据投递的速度和效率，降低目标系统的资源占用，可以在亚秒级实现大量数据的复制，并且目标端数据库是活动的。

优点：性能高，不依赖数据库，高压缩率。

缺点：不开源，收费。

## Canal

Canal是由Alibaba开源的一个基于binlog的增量日志组件，其核心原理是canal伪装成Mysql的slave，发送dump协议获取binlog，解析并存储起来给客户端消费。

优点：可以同步任何非查询类操作。DDL和DML语句（除了数据查询语句select）。增量+全量都可以。

缺点：数据源只支持MySQL，因为利用了MySQL的binlog特性。

项目地址：https://github.com/alibaba/canal

## otter

otter是一款基于Java且免费、开源基于数据库增量日志解析，准实时同步到本机房或异地机房的mysql/oracle数据库的解决方案

优点：

1. 单向同步， mysql–>oracle互相同步
2. 双向同步，无冲突变更
3. 文件同步，本地/aranda文件
4. 双A同步，冲突检测&冲突补救
5. 数据迁移，中间表/行记录同步

缺点：同canal，因其是利用canal来进行同步的，canal作为其数据源组件。

项目地址：https://github.com/alibaba/otter

## Kettle

Kettle是一款国外开源的ETL工具，用java编写，可以在视窗、Linux和Unix上运行，数据抽取高效稳定，中文称水壶。这个项目的主要程序员MATT想把各种数据放进一个水壶，然后以指定的格式流出。它是一个ETL工具集，允许你管理来自不同数据库的数据，并通过提供一个图形用户环境来描述你想做什么，而不是你想怎么做。Kettle中有两个脚本文件，转换和作业。转换完成了数据的基本转换，而作业完成了对整个工作流的控制。

优点：功能强大，支持几乎所有数据库。

缺点: 需要用户自己一步步配置，学习成本高；通过查询语句select同步的。

项目地址：https://github.com/pentaho/pentaho-kettle

# 应用

## OceanBase

## PolarDB

## TiDB

TiDB Data Migration (DM)是针对MySQL或MariaDB到TiDB的一体化数据迁移任务管理平台，DM通过解析binlog event将数据以SQL语句的形式迁移到TiDB中，可以完成全量数据迁移和增量数据复制。

参考：<https://mp.weixin.qq.com/s/plMdhalL4n1cNe5pUnQ1lg>

## TDSQL

采用Kafka实现多源同步。

## GoldenDB