**数据迁移操作手册**

发布日期： 2019.10

|  |  |
| --- | --- |
| 撰 写： | 王 振 华 |
| 文档编号： |  |

文档修改记录

|  |  |  |  |
| --- | --- | --- | --- |
| 版本号 | 主要作者 | 修改记录 | 完成日期 |
| 0.1 | 王振华 | 创建 | 2019-10-15 |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

**目 录**

[1 程序部署 1-1](#_Toc22057305)

[1.1 部署环境 1-1](#_Toc22057306)

[1.2 程序安装 1-1](#_Toc22057307)

[1.3 配置参数 1-1](#_Toc22057308)

[1.4 程序启停 1-2](#_Toc22057309)

[1.4.1 程序启动 1-2](#_Toc22057310)

[1.4.2 程序停止 1-2](#_Toc22057311)

[2 程序配置 2-3](#_Toc22057312)

[2.1 数据源管理 2-3](#_Toc22057313)

[2.2 配置管理 2-4](#_Toc22057314)

[2.3 调度管理 2-6](#_Toc22057315)

[2.4 结果跟踪 2-6](#_Toc22057316)

[2.5 迁移结果 2-7](#_Toc22057317)

# 程序部署

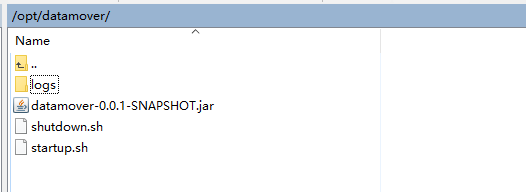
## 部署环境

* Linux服务器
* JDK1.8或以上
* MySQL5

其中JDK1.8需要安装在Linux服务器上，MySQL5可以安装在准备的Linux服务器上，也可以安装在其他服务器上。

## 程序安装

这里将程序所有文件都放到/opt/datamover目录。该目录结构如下



|  |  |
| --- | --- |
| 文件/目录 | 说明 |
| logs | 生成的日志文件夹，无需创建，程序自动生成。 |
| datamover-0.0.1-SNAPSHOT.jar | 程序主文件 |
| shutdown.sh | 程序停止脚本，需赋予执行权限 |
| startup.sh | 程序启动脚本，需赋予执行权限 |

## 配置参数

所有的程序参数可通过startup.sh脚本更改，内容如下

|  |  |
| --- | --- |
| 参数名 | 说明 |
| server.port | 端口号 |
| server.servlet.context-path | 项目访问名称 |
| spring.datasource.primary.url | Mysql配置数据存储库JDBC连接字符串 |
| spring.datasource.primary.username | Mysql配置数据存储库JDBC用户名 |
| spring.datasource.primary.password | Mysql配置数据存储库JDBC用户密码 |
| spring.datasource.primary.driver-class-name | Mysql配置数据存储库JDBC驱动类名 |
| spring.datasource.primary.initialization-mode | always: 始终执行初始化；embedded: 只初始化内存数据库（默认值）；never: 不执行初始化 |
| spring.datasource.primary.initial-size | 连接池初始化连接个数 |
| spring.datasource.primary.min-idle | 回收空闲连接时，保证至少有几个连接 |
| spring.datasource.primary.max-active | 连接池中最多支持多少个活动会话 |
| spring.datasource.primary.max-wait | 程序向连接池中请求连接时，最大等待的毫秒数，超过这个时间则认为失败 |
| spring.datasource.primary.time-between-eviction-runs-millis | 空闲连接检测间隔（毫秒） |
| spring.datasource.primary.min-evictable-idle-time-millis | 连接在池中最小生存的毫秒数 |
| spring.datasource.primary.validation-query | 检测连接有效性的SQL |
| spring.datasource.primary.test-while-idle | 在分配连接时，是否先检查该连接是否有效 |
| spring.datasource.primary.test-on-borrow | 申请连接时，是否进行连接有效性检查 |
| spring.datasource.primary.test-on-return | 返还连接时，是否进行连接有效性检查 |
| spring.datasource.primary.break-after-acquire-failure | 是否失败重连 |
| spring.datasource.common.max-wait | 页面创建的数据源最大等待的毫秒数 |
| datamover.dest.group-count.delete | 数据迁移批量删除一次操作的记录上限 |
| datamover.dest.group-count.insert | 数据迁移批量插入一次操作的记录上限 |
| datamover.thread.scheduler.pool-size | 迁移配置执行线程池大小 |

startup.sh并未列出所有参数，这些参数基本不会被更改。

## 程序启停

### 程序启动

|  |
| --- |
| # cd 到程序（datamover\*\*\*.jar）所在目录  ./startup.sh |

### 程序停止

|  |
| --- |
| # cd 到程序（datamover\*\*\*.jar）所在目录  ./shutdown.sh |

# 程序配置

为介绍方便，这里以一个示例介绍程序的使用。

源数据库为MySQL，表名称为t\_enterprise\_info（规模：40+列，16000+条数据）。

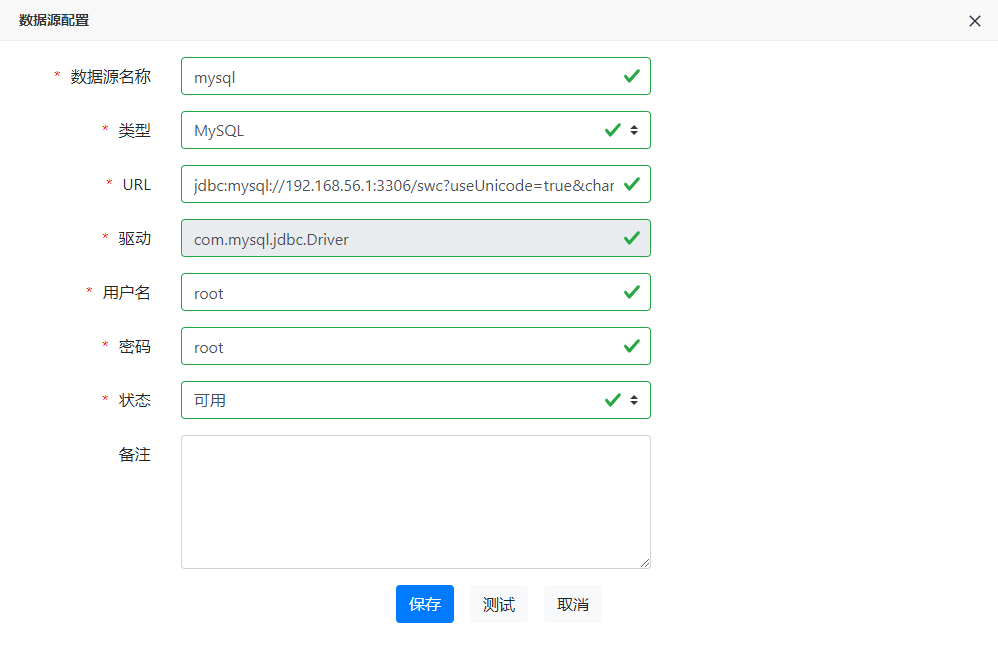
目标数据库为PostgreSQL，表名称同样为t\_enterprise\_info（实际使用中，表名及列名也可不同）。

通过该工具，将源库表数据迁移到目标表数据，每1分钟同步一次。

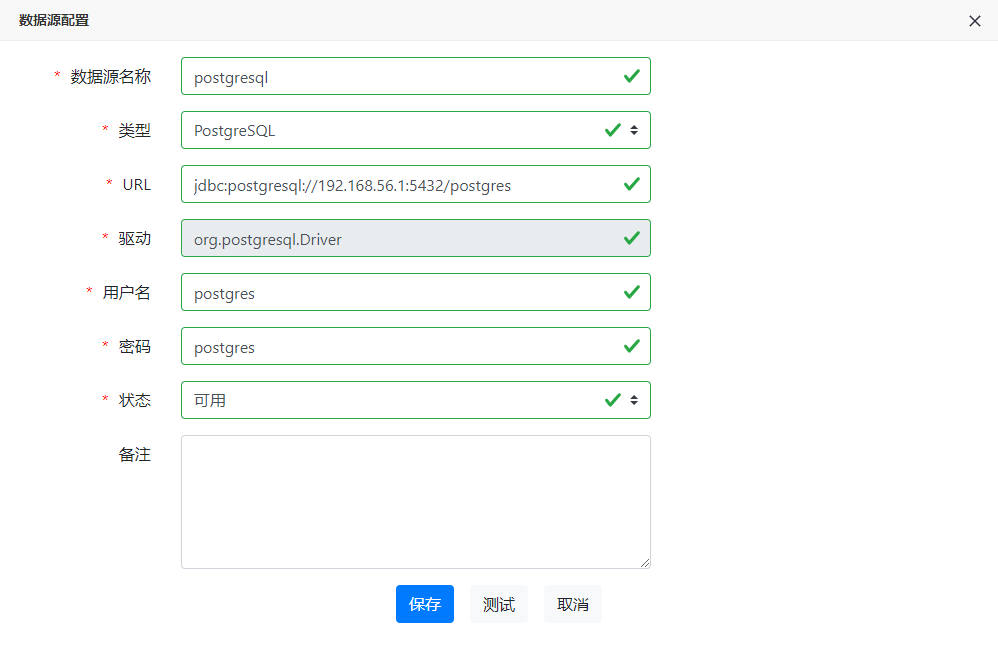
为完成上述工作，需要进行数据源配置、迁移策略配置、调度配置这些项的操作。

## 数据源管理

.新增源数据库，表单如下



新增目标数据库，表单如下



新增数据库后，无需点击“重新加载”按钮（页面触发的变更，系统会自动重新加载，只有运维人员改库方式的变更，才需要“重新加载”操作）。

## 配置管理

由于是跨库进行数据迁移，所以需要选择查询SQL进行表单填写。

查询SQL中，可以以“${xxx}”方式设置变量。这里设置一个自定义变量。

查询SQL后，根据查询数据和表主键（或联合主键），删除目标库表的数据。主键为“[“id”]”，即主键只有一列。若为多列可这样设置：“[“name”, “dept”]”。

超时时间可以为空，也可以设置为秒数。若为空则认为永不超时。

是否单例一般设置为是，这样若已经有任务在执行，可避免再次启动任务。

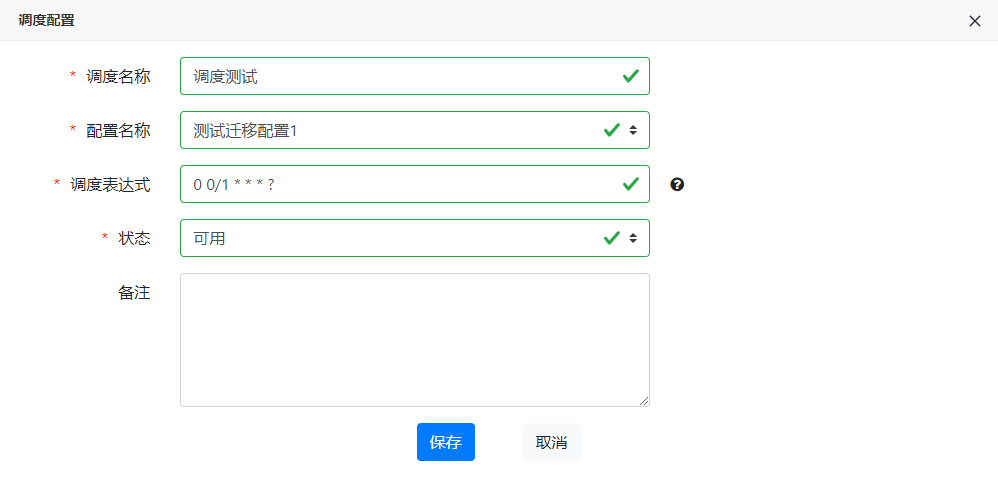
状态可控制配置是否能被执行，只有可用状态的配置才能执行。

任务完成后，可以执行一段SHELL脚本。脚本可以调用其他系统的URL来进行自动化的步骤编排。下图中的这部分配置，会获取一下本次执行结果。



## 调度管理

每个调度配置，可以关联一个迁移配置。调度的配置页面如下。



其中调度表达式与Quartz表达式一致，上述表达式代表一分钟执行一次，更多实例请参照帮助页面。保存后，若状态为可用，则配置立即生效（页面触发的变更，系统会自动重新加载，只有运维人员改库方式的变更，才需要“重新加载”操作）。

## 结果跟踪

手动下发任务后，来到结果跟踪页面，页面显示如下



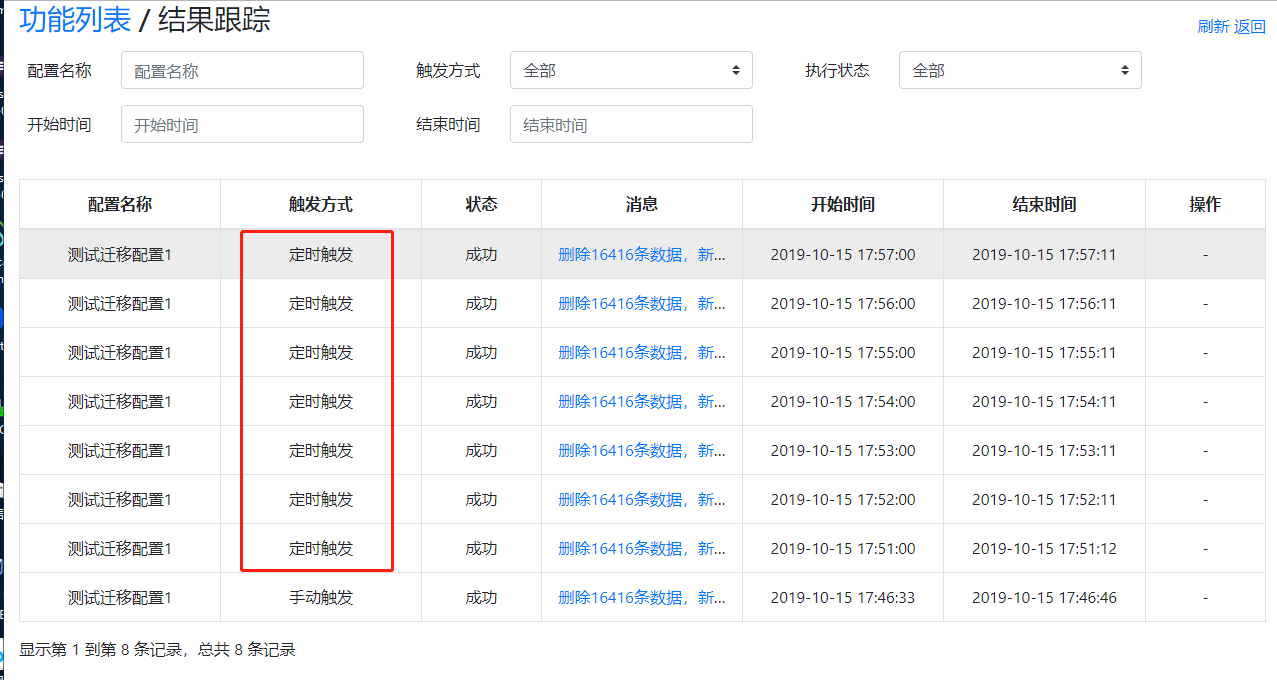
此时任务的状态为“运行中”。页面默认30秒刷新一次，也可以通过右上角“刷新”按钮手动刷新。任务完成后如下



点击消息内容，可以查看任务的消息内容，如下图。

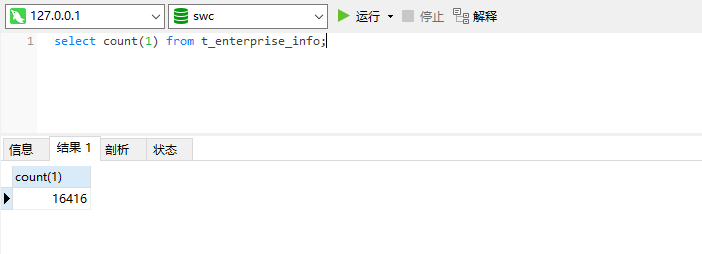


这个页面也可看到调度触发的任务结果，如下图。



## 迁移结果

源库数据条数



目标库数据条数

