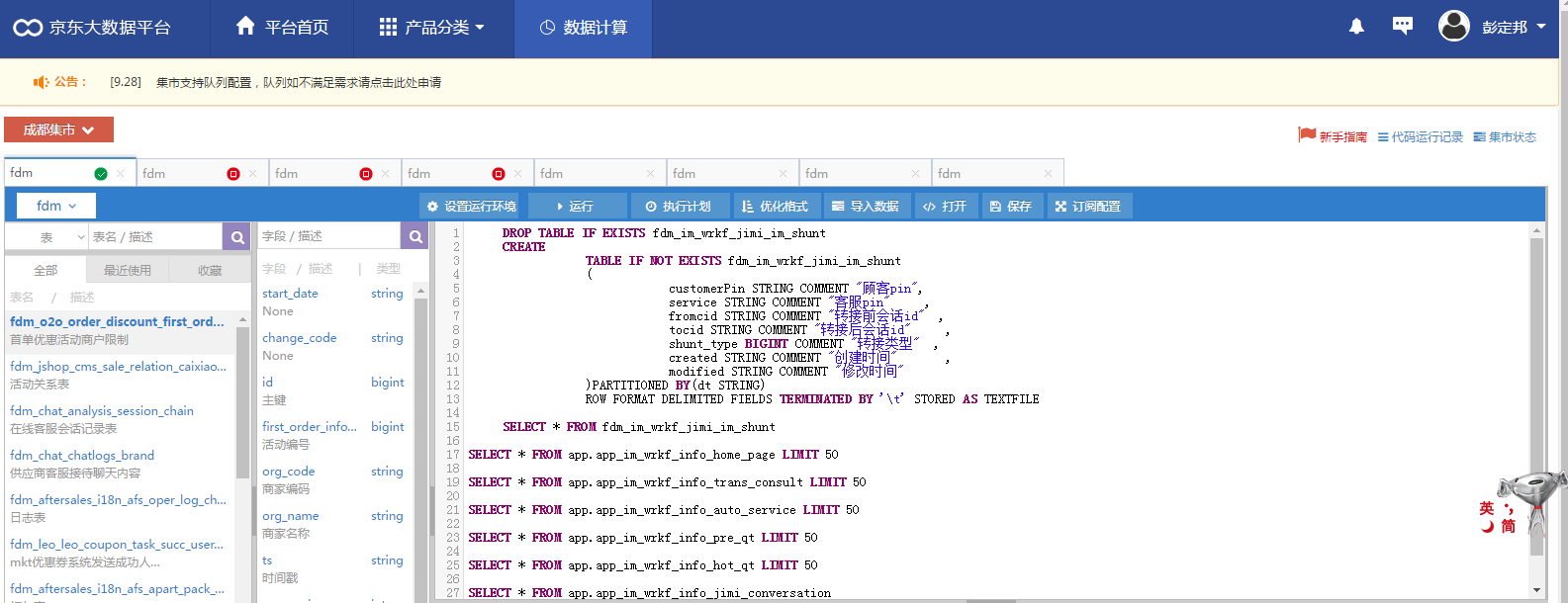
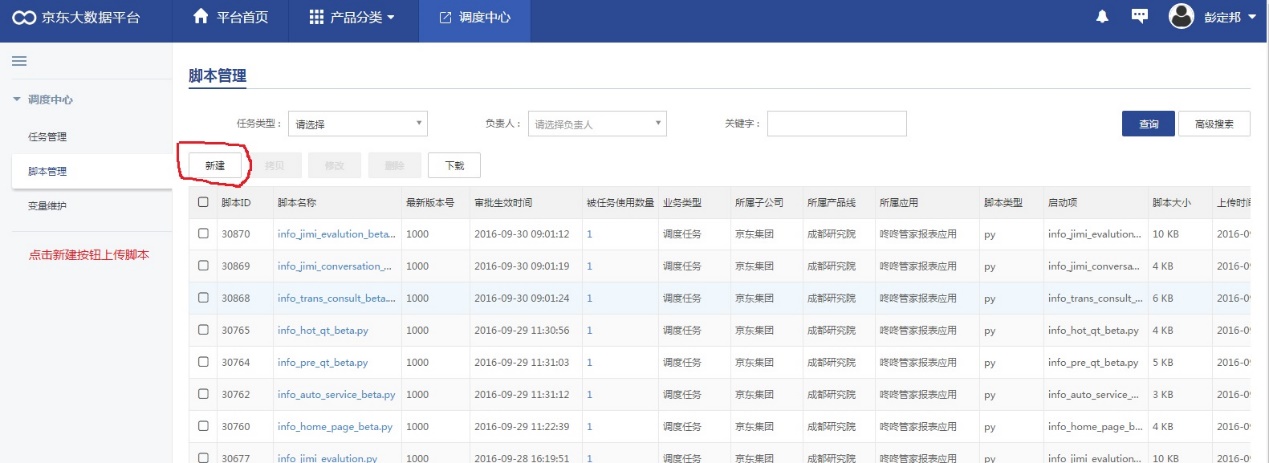
大数据平台数据外推流程

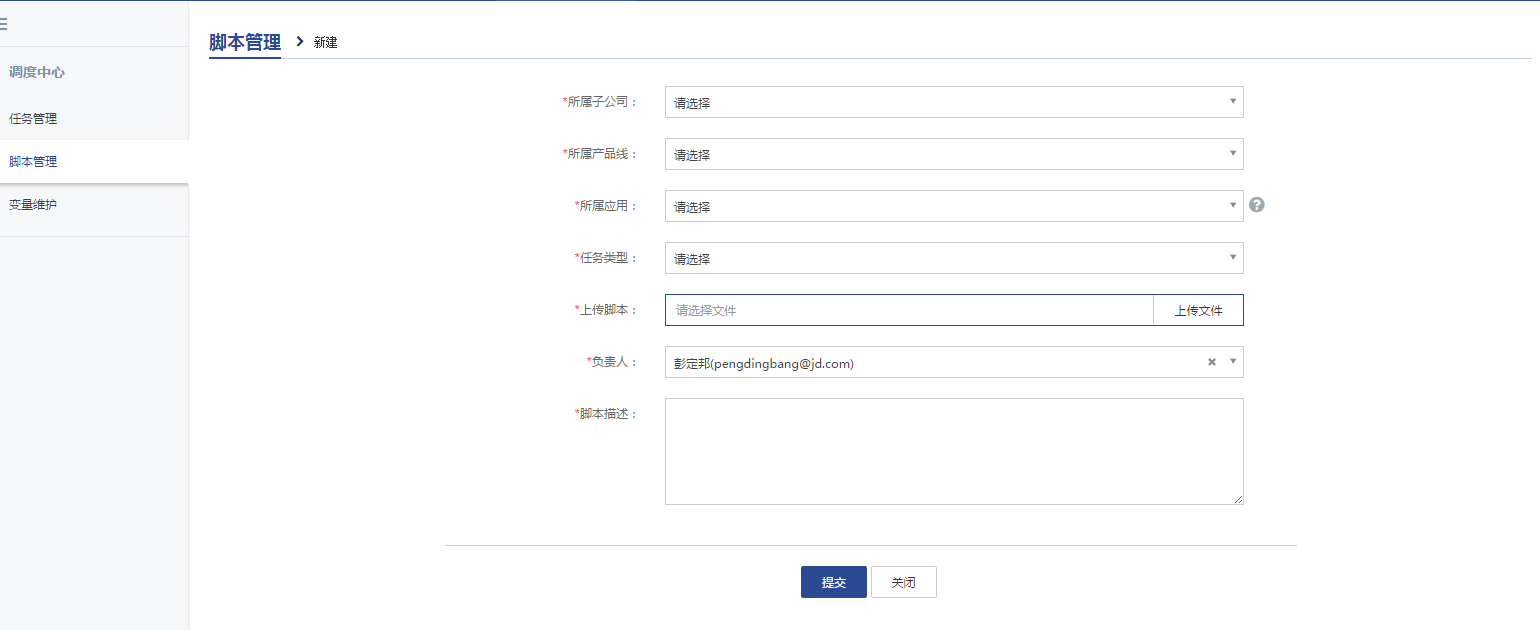
1. 步骤说明
2. HQL编写，验证无误。（使用单引号，避免奇怪错误）
3. 将HQL改写为python脚本。（UTF-8编码，可使用平台上运行脚本改写）
4. 脚本上传，并在流程中心审批通过。
5. 建立数据计算任务，选择对应的python脚本。
6. 数据外推流程申请。
7. 建立抽取数据任务。
8. 详细说明
9. 在大数据平台HQL编写，如下图所示



1. 将运行无误的HQL改写为python脚本，调度任务脚本可参考线上正常运行的任务脚本。
2. 在调度中心，脚本管理模块上传脚本，填写对应申请单，如下图所示

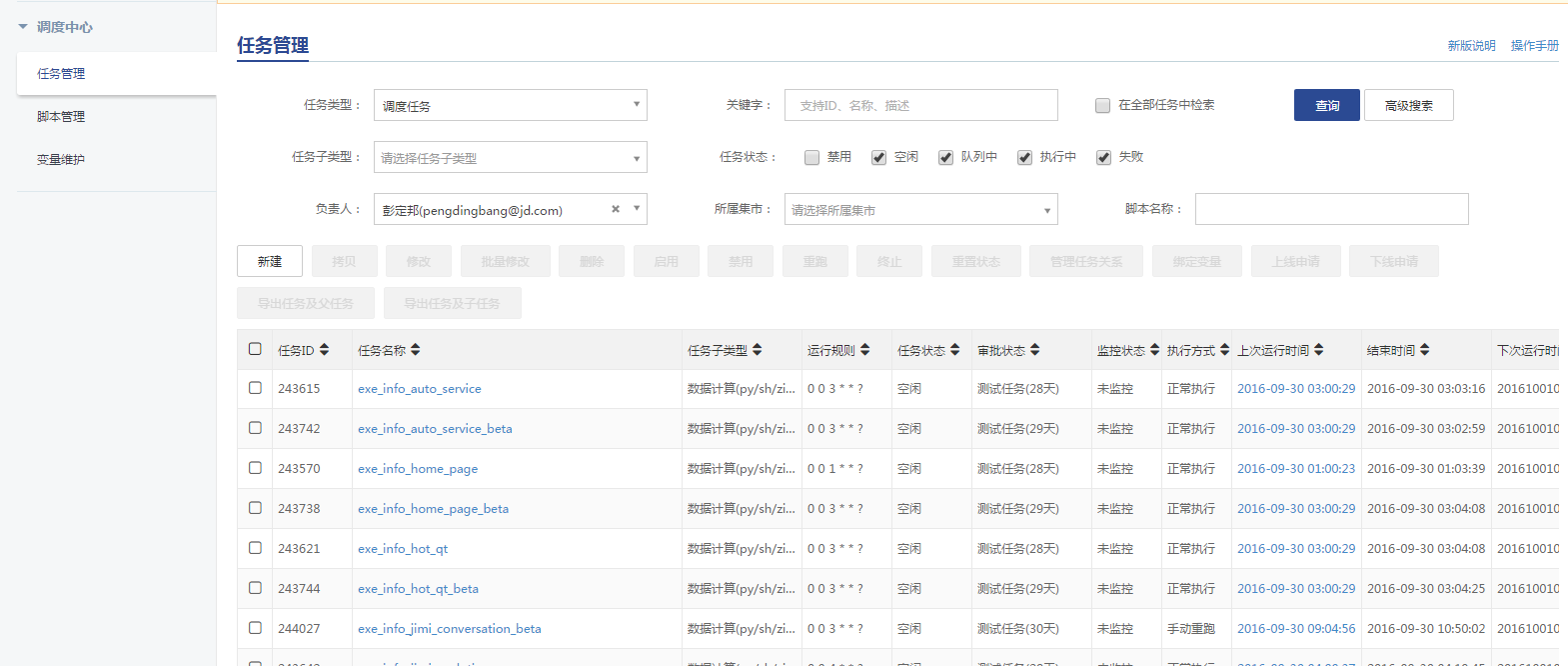


点击图中新建按钮，填写新建脚本申请



点击提交按钮后，应用负责人需要在流程中心完成审批。

1. 脚本审批完成后，在调度中心任务管理模块建立对应的数据计算任务，如下图所示。



点击新建按钮，填写任务表单



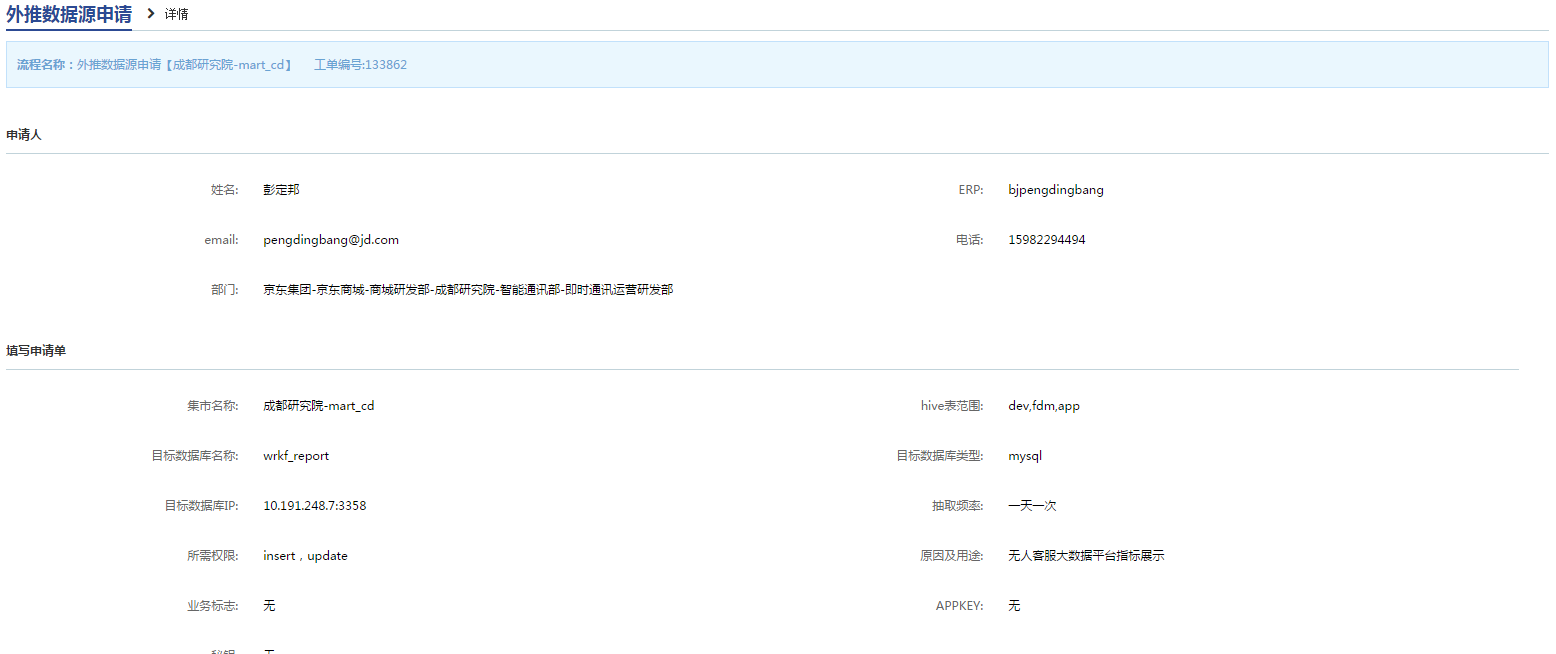
在任务子类型中，选择数据计算任务；脚本名称中选择已上传的对应脚本；节点编号选择支持所属集市的节点。

数据计算任务依赖大数据平台的业务表，需要等业务表抽数完成后，才能执行计算任务，需要设置依赖关系。

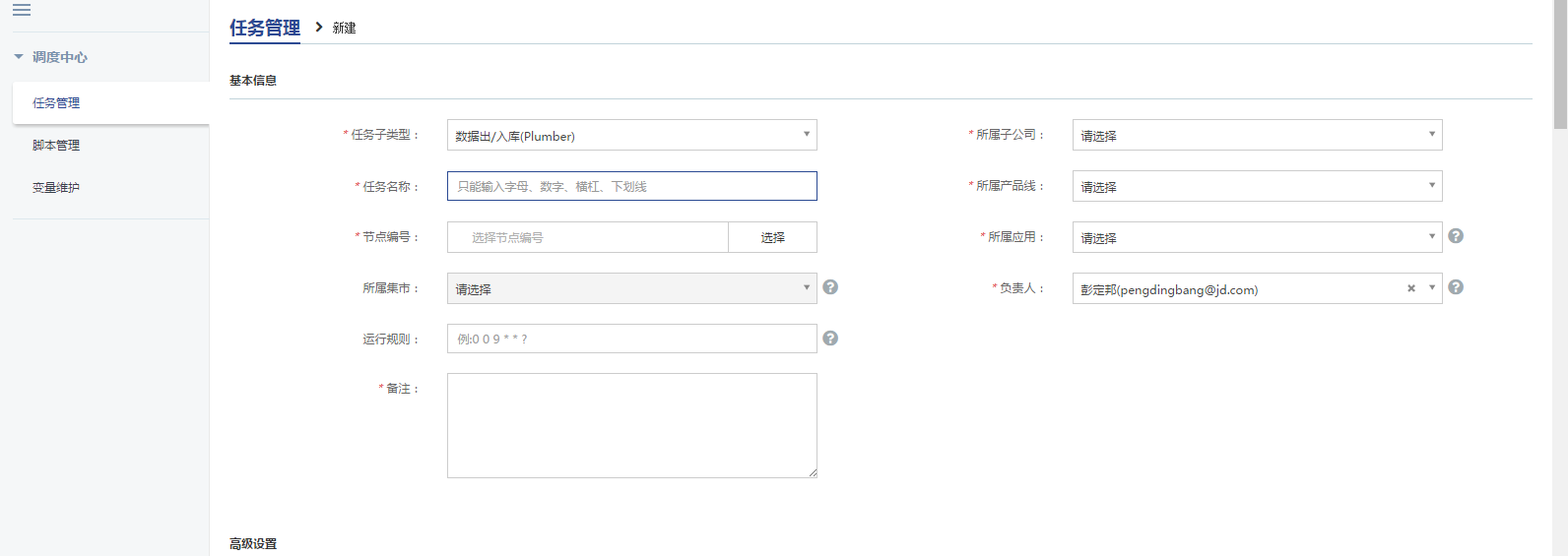
1. 数据计算任务建立完成之后，可建立数据外推任务，在流程中心提外推数据源申请，如图所示。



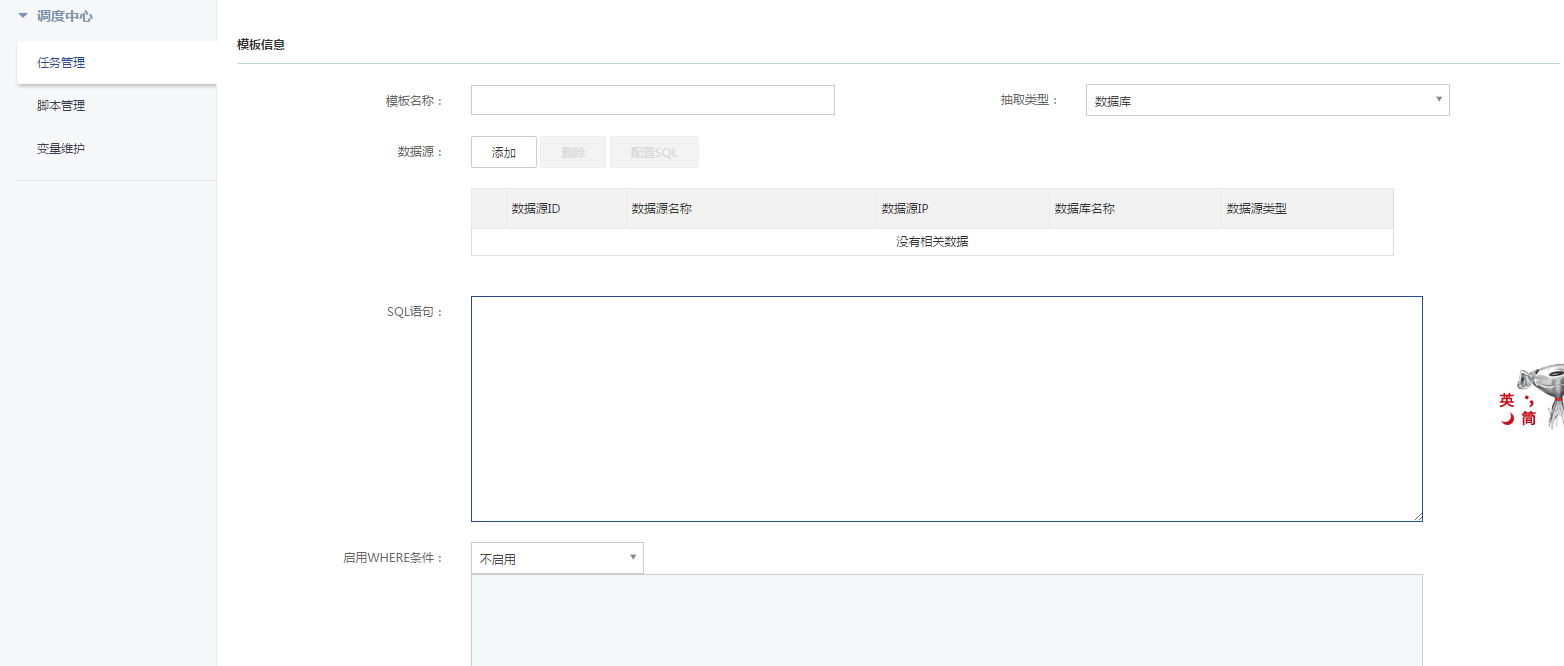
外推数据源申请填写完成后，申请进度及各个流程审批负责人可在“我的申请”中查看。



1. 外推数据源申请流程通过之后，需要建立plumber任务。在调度中心-任务管理新建任务，如下图所示



任务子类型选择为数据出/入库，在模板信息中，添加数据源；数据源为业务数据来源，例如fdm\_chat\_im\_pop\_upv数据源库为imsl.mysql.jddb.com。（这个不需要严格对应）SQL语句用于筛选推送到目标数据库的数据。可参考线上推数任务。



目标数据源选择为线上数据，目标表名为线上数据库中表名

