**poseidon平台使用手册**

|  |  |  |
| --- | --- | --- |
| 文件状态：  [] 草稿  [√] 正式发布  [ ]正在修改 | 文件标识： | poseidon平台使用手册 |
| 当前版本： | 1.0 |
| 作者： | 安培 |
| 完成日期： |  |

修订历史

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 序号 | 版本 | 编写/修订说明 | 修订人 | 修订日期 | 备注 |
| 1 |  | 创建文档 | 安培 | 2018/06/19 |  |

目录

[一、 简介 4](#_Toc517356741)

[二、 用户使用指南 5](#_Toc517356742)

[1、配置管理 5](#_Toc517356743)

[2、用户管理 8](#_Toc517356744)

[3、StreamCql 10](#_Toc517356745)

[4、任务管理 11](#_Toc517356746)

[5、监控管理 22](#_Toc517356747)

[6、历史记录管理 24](#_Toc517356748)

[7、任务告警与恢复 25](#_Toc517356749)

[三、 附录 26](#_Toc517356750)

[1. 通用配置项解释 26](#_Toc517356751)

[2. Jstorm任务开发插件 29](#_Toc517356752)

# 简介

本文档隶属于poseidon实时计算平台，poseidon平台为管理和监控jstorm和 Flink任务提供部署、管理与监控服务。

**github地址：**

<http://poseidontest.10101111.com/>

[http://poseidon.10101111.com/](http://poseidontest.10101111.com/)

**本文档用于**

* 方便用户快速了解poseidon项目；
* 方便用户快速熟悉poseidon项目的相关功能；
* 方便用户及时解决操作中遇到的相关问题。

# 用户使用指南

## 1、配置管理

1. **通用配置信息**

用户通过此功能来对系统变量进行配置。可以添加和修改配置项。超级管理员可以进行“添加配置项”和“修改配置项”，而普通用户只能查看配置项列表。

配置项的详细说明见此手册的“通用配置项解释”部分

****

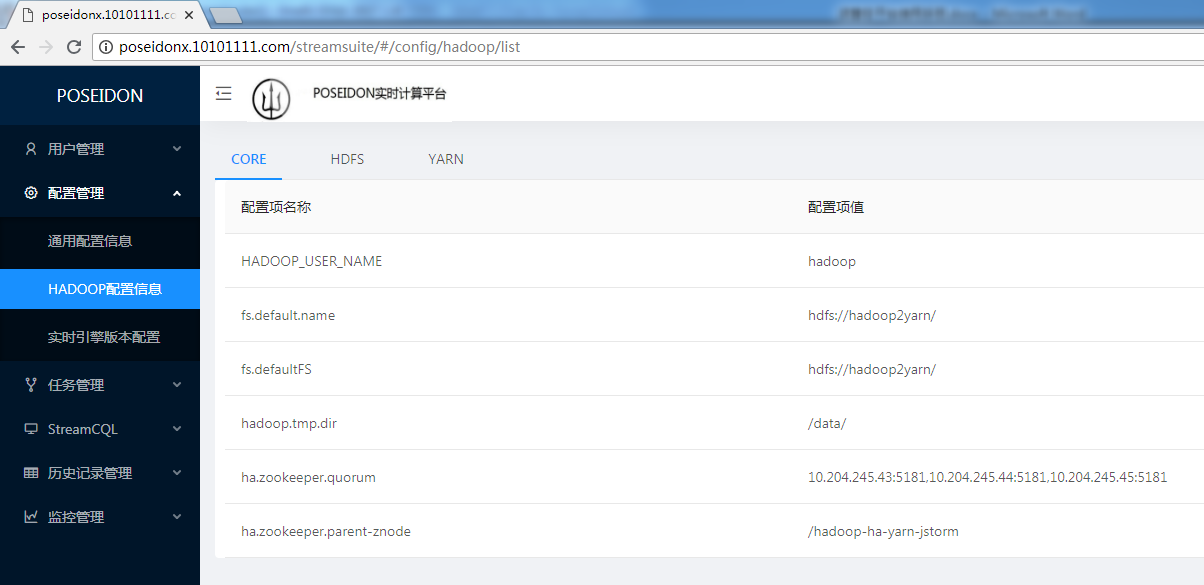
添加和修改配置项，需要录入配置项名称，配置项值，配置项备注。

****

一般情况下，用户无需添加系统配置项，都是系统内置好的，用户可以根据实际需要修改配置项的值。比如修改 FLINK\_AUTO\_RECOVERY\_RETRY 来设置flink任务失败时自动恢复的重试次数。

1. **Hadoop配置信息**

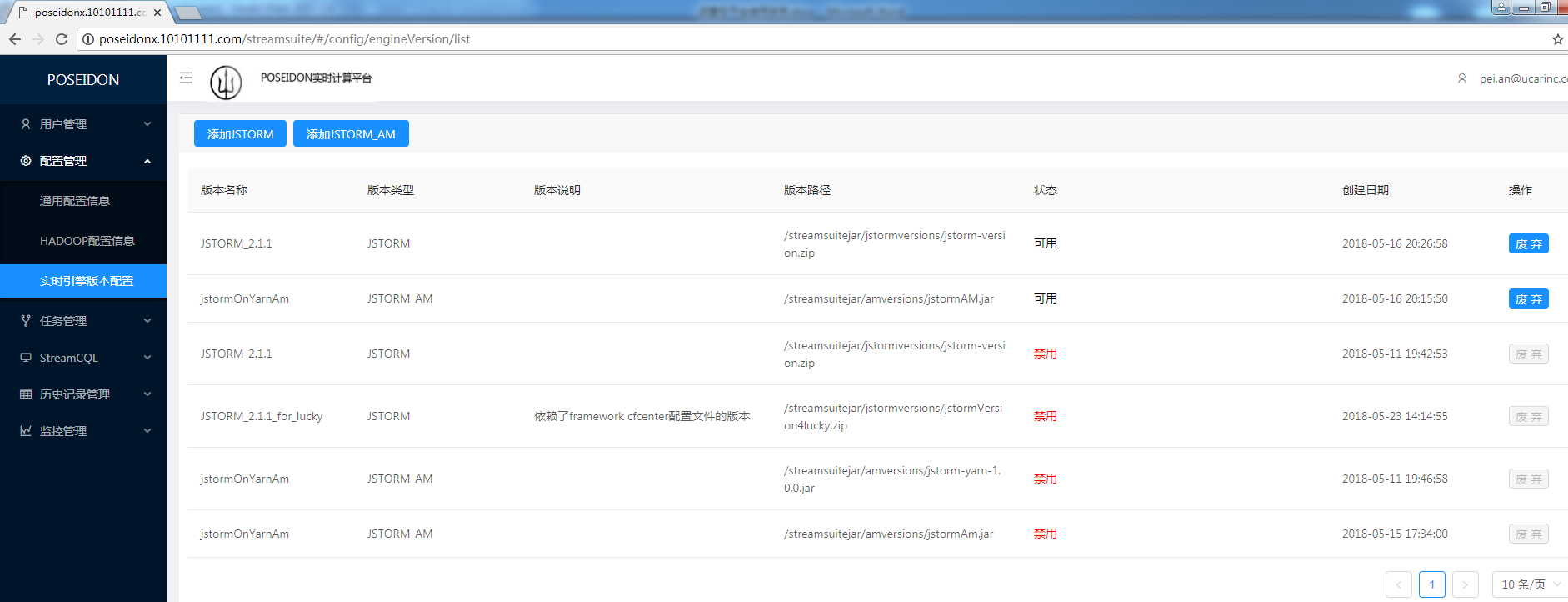
此功能展示了hadoop配置文件 core-site.xml，hdfs-site.xml，yarn-site.xml 的配置值。

****

1. **实时引擎版本配置**

此功能专门用于上传jstormAm引擎包，和jstorm引擎包。平台提交jstorm任务时会先在yarn环境启动 Am，再通过Am来进一步启动jstorm集群。

超级管理员可以进行上传引擎包和废弃引擎包操作，而普通用户只能查看引擎包信息列表。

****

点击添加jstorm引擎包上传jstorm引擎包

****

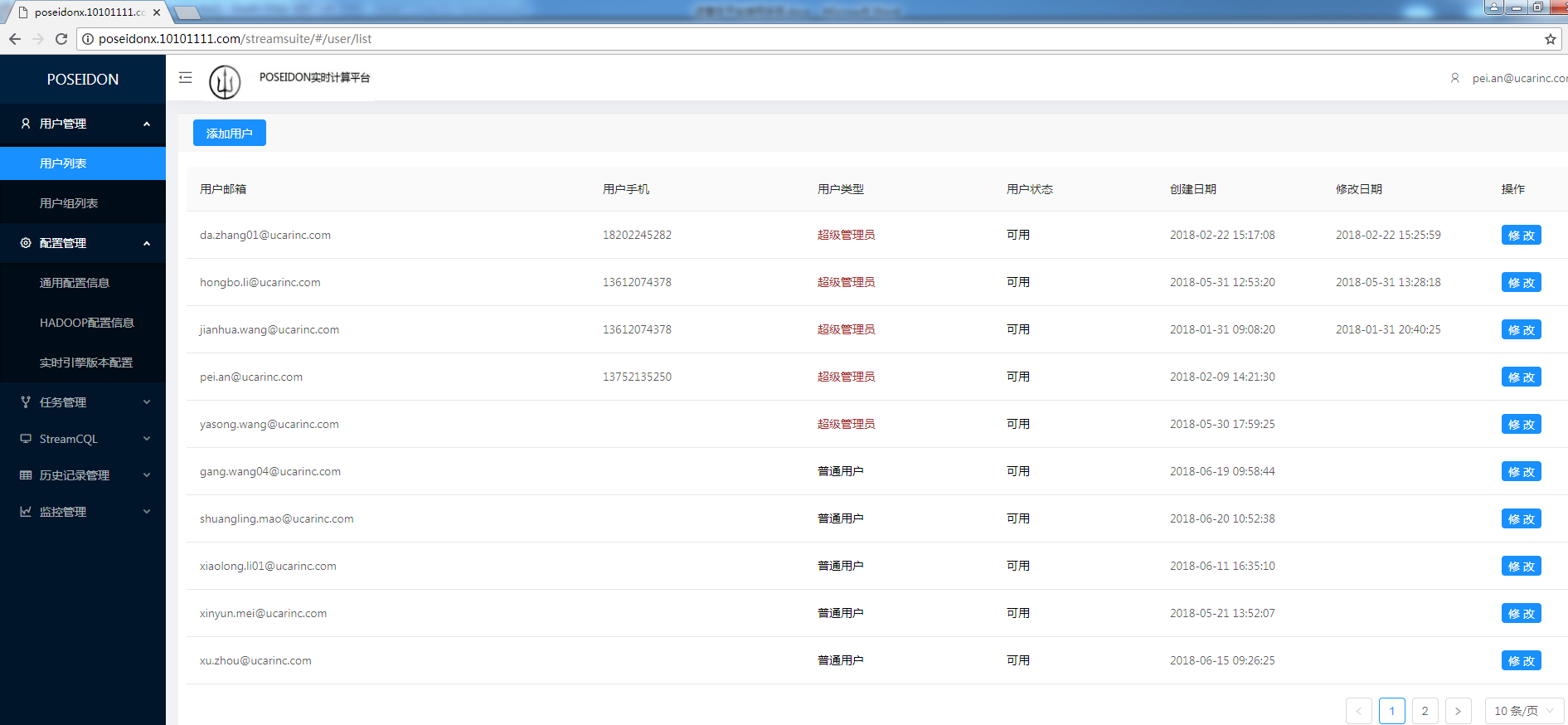
点击添加jstormAM引擎包上传jstormAM引擎包

****

## 2、用户管理

1. **用户列表**

此功能用于用户信息的维护，用户类型分为超级管理员和普通用户，超级管理员可以进行“添加用户”和对所有人员进行“修改”操作，而普通用户只能对自己的信息进行“修改”操作。

****

点击添加进行添加用户。用户名必须是邮箱格式，此邮箱也用于接收告警信息。可以同时录入用户的手机号，此手机号也会用于接收告警短信使用。

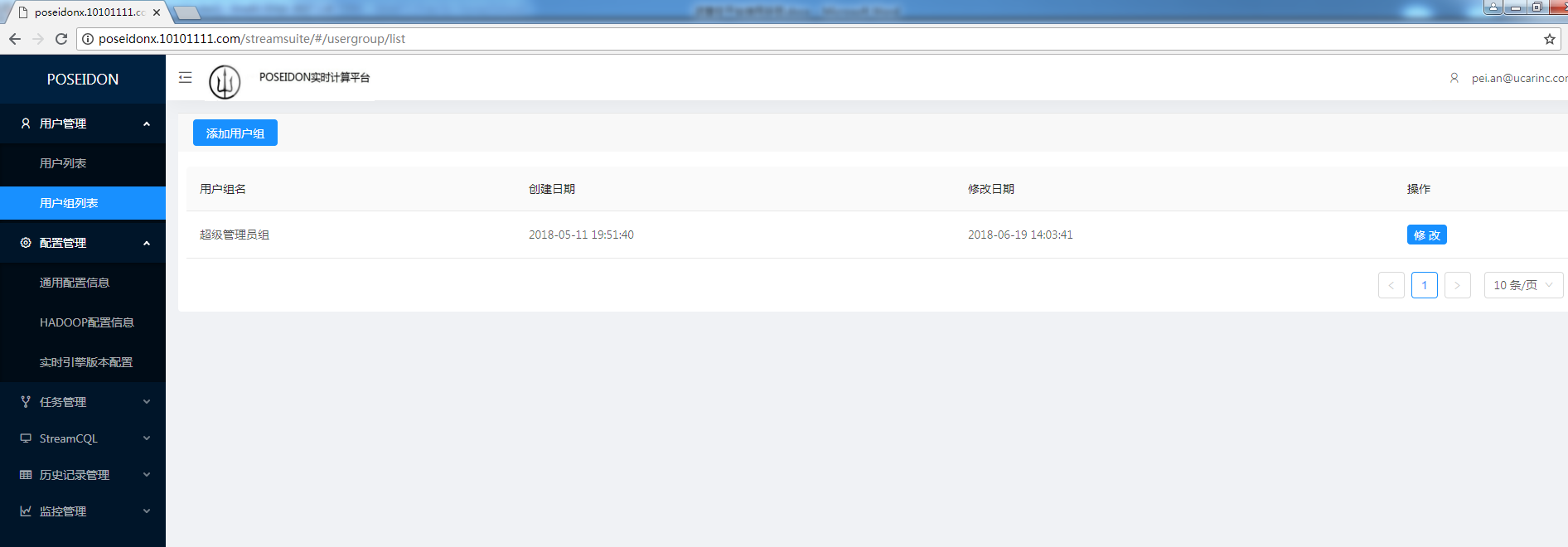


1. **用户组列表**

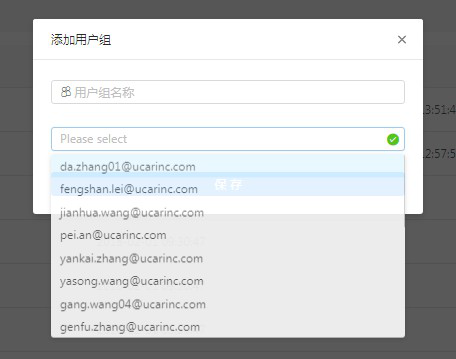
此功能用于用户组的维护。超级管理员可以进行“添加用户组操作”，普通用户只能查看列表信息。

使用用户组的概念是便于对 cql脚本和任务文件进行管理。用户在创建cql脚本和任务文件时需要将其指定到一个用户组。这样在创建任务时对于普通用户只可以选择自己所属用户组的cql脚本和任务文件。

并且同样对于普通用户只能看到自己创建的和自己所属用户组的任务。



添加用户组可以通过可多选的下拉框进行成员选择。



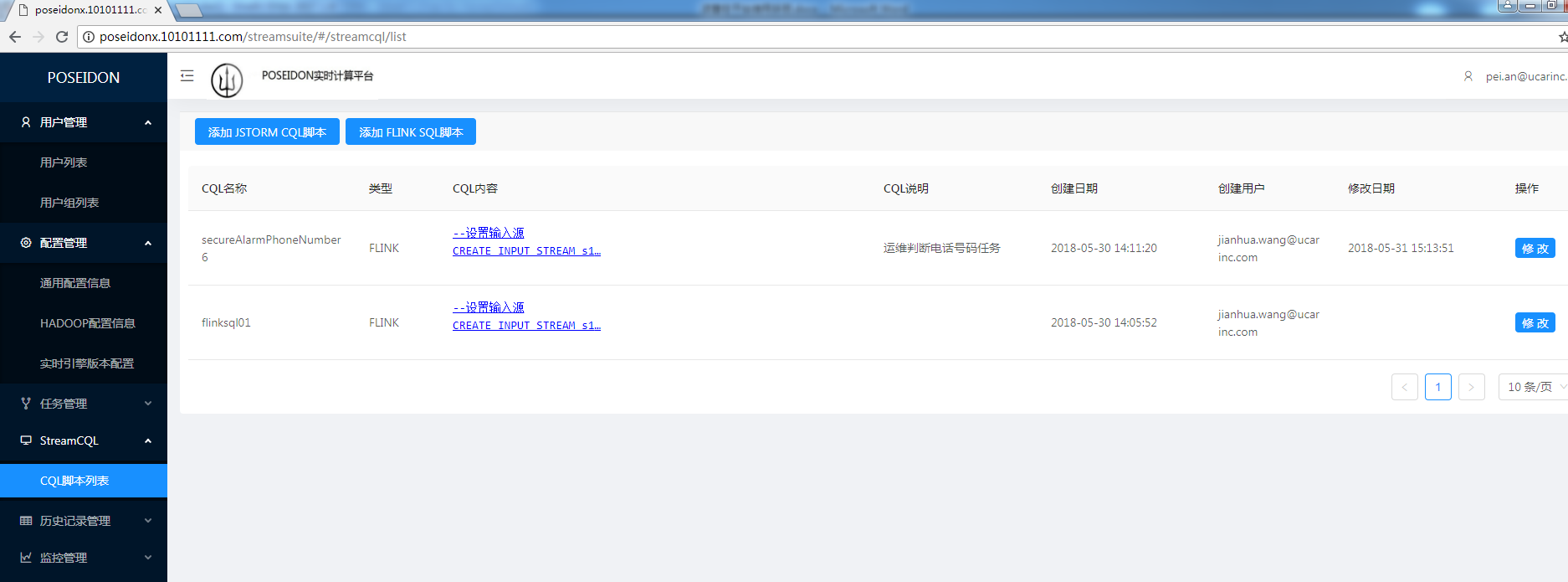
## 3、StreamCql

1. **CQL脚本列表**

此功能用于对CQL脚本的维护。

平台中的任务提交方式有两种，任务文件方式提交和CQL方式提交。此处就是对CQL脚本的维护功能。

具体的CQL语法参考《CQL语法用户指南》

****

添加脚本时可以对语法进行检查。填写cql名称、备注、选择用户组后进行保存。

****

## 4、任务管理

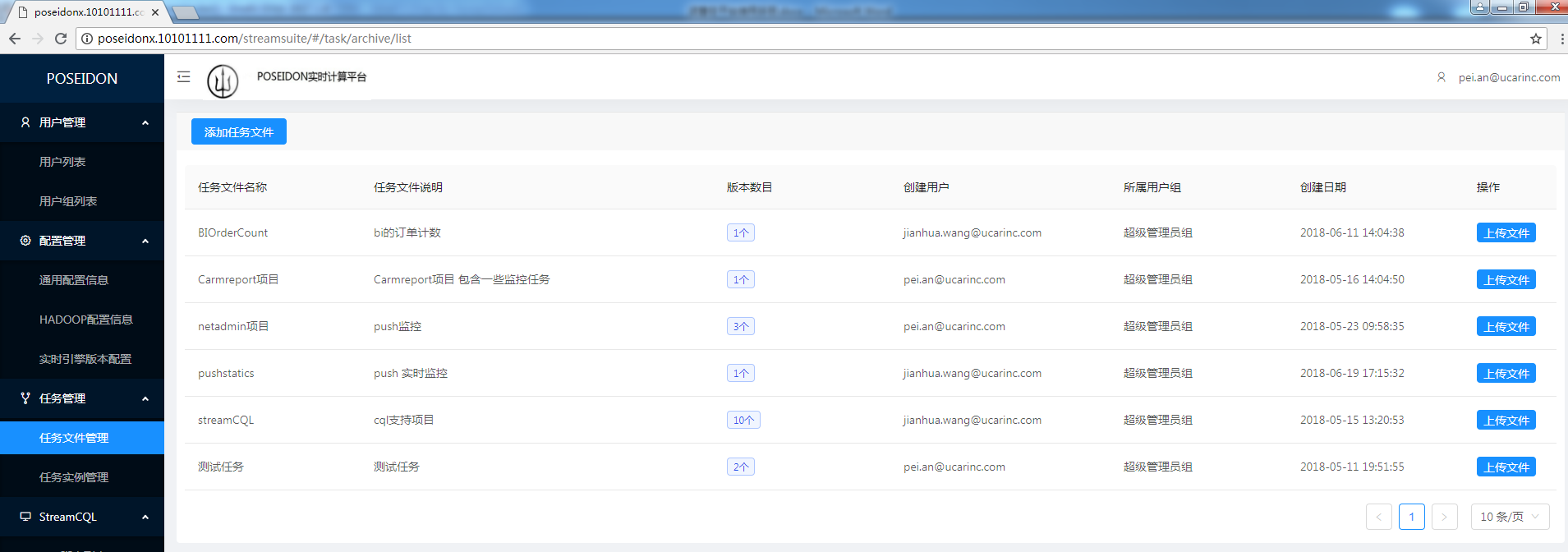
1. **任务文件管理**

此功能用于对任务文件进行维护。

平台中的任务提交方式有两种，任务文件方式提交和CQL方式提交。此处就是对任务文件的维护功能。

任务文件是个虚拟的概念，可以理解为一个任务的项目名称。一个任务文件记录，可以上传多个版本的war包，称为任务文件版本，任务文件版本被存储于hdfs中。

创建任务时通过选择任务文件，再选择任务文件的版本，并设置main类的路径对任务进行配置。

****

添加任务文件。填写任务文件名称、备注、选择用户组后进行保存。任务文件名称不能重复。

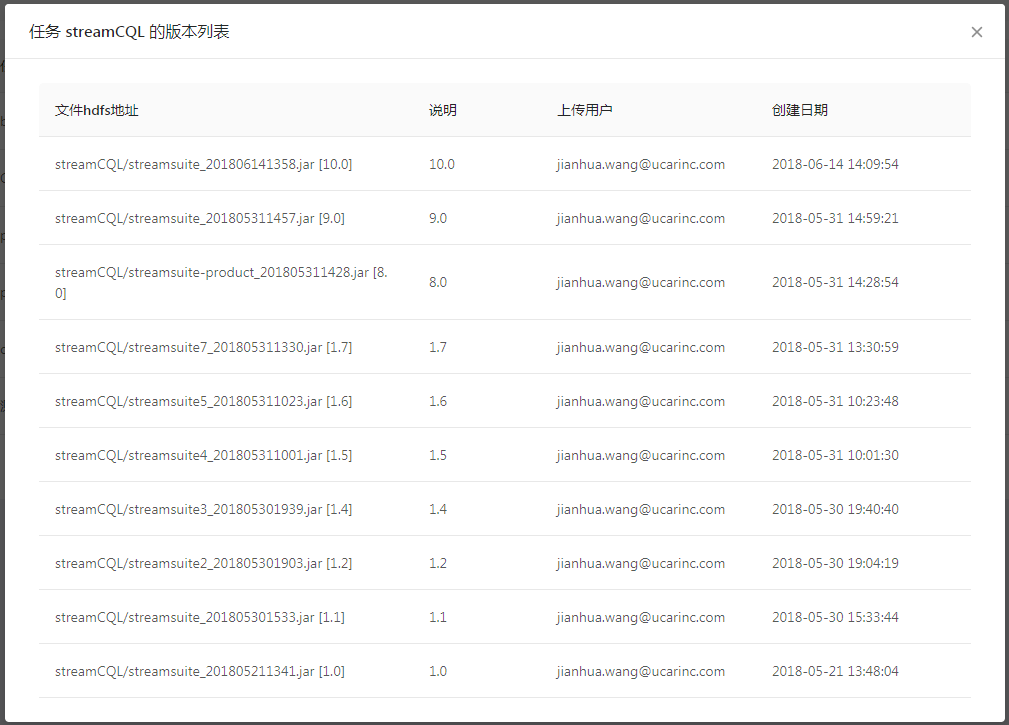
****

点击上传文件，可以上传任务文件版本。上传的任务文件必须是.war文件。

war包上传之后，会被重新整合为一个jar包文件提交到hdfs上。

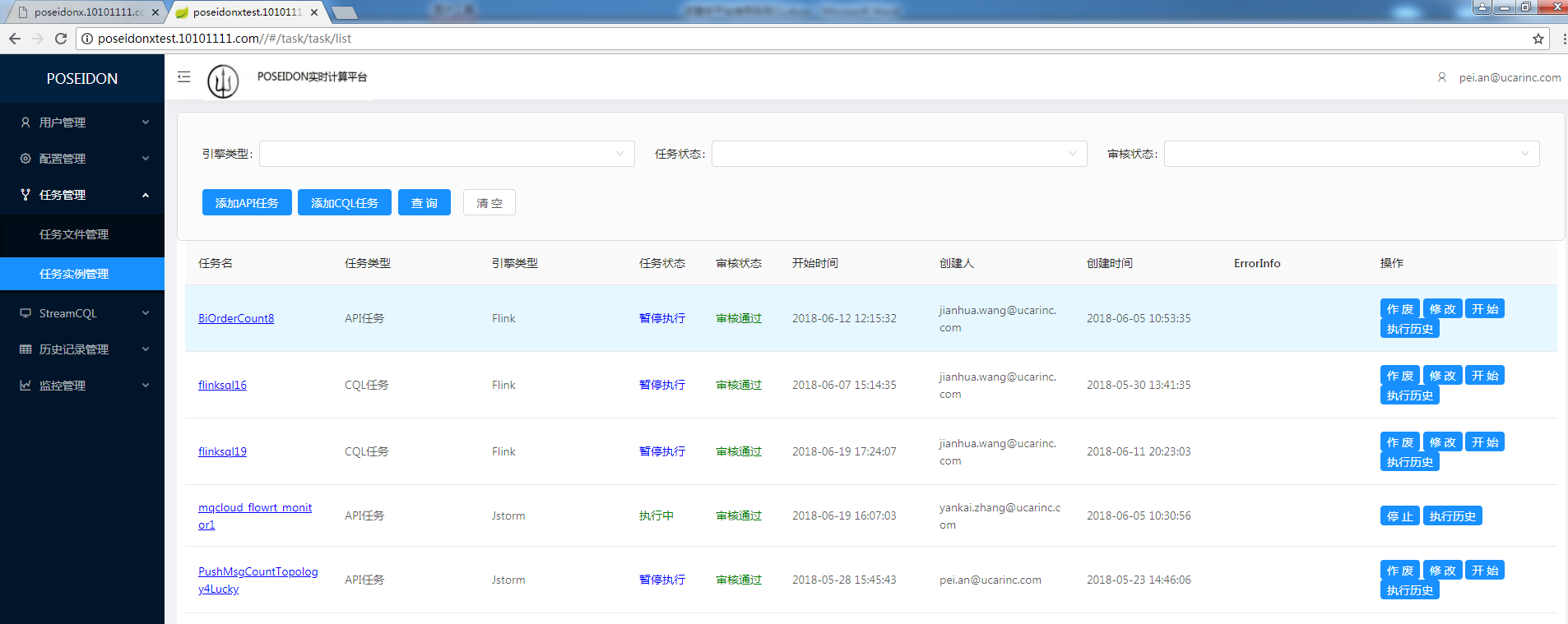
****

点击版本数目会显示所有的任务文件版本。

****

为了可以正常提交CQL任务。用户需要在任务文件管理中上传一个特殊的任务文件 CQL引擎任务文件。CQL引擎任务文件的名称在通用配置信息中，由配置项CQL\_ENGINE\_PROJECT\_NAME进行配置。CQL引擎的war包如果打包参考 《安装文档.doc》准备任务执行引擎部分内容。

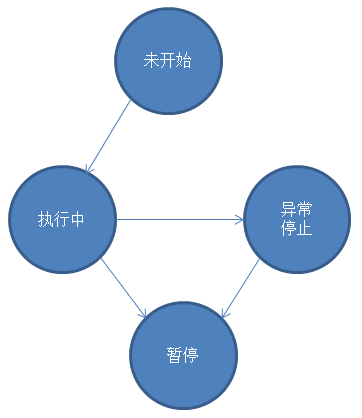
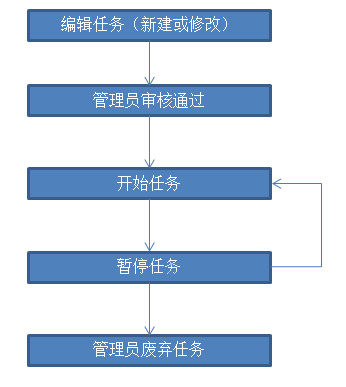
1. **任务实例管理**



任务管理是平台的核心功能。普通用户可以创建、修改、开始、停止任务，查看任务提交历史等操作。管理员除此之外还可以对任务进行作废和审核操作。

任务类型：平台支持执行两种任务，API任务和CQL任务，执行API任务需要先上传任务文件war包。执行CQL任务需要先编辑StreamCQL语句。

引擎类型：平台支持两种引擎，jstorm和flink。



任务生命周期 任务状态转换

点击添加API任务，选择jstorm引擎新建Jstorm的API任务。



点击添加API任务，选择Flink引擎新建Flink的API任务。



点击添加CQL任务，选择Jstorm引擎新建Jstorm的CQL任务。



点击添加CQL任务，选择Flink引擎新建Flink的CQL任务。



管理员审核完毕后，点击开始任务。此时会有时间轴出现，请耐心等待提交。

如果任务提交成功，则每个时间轴都是绿色。提示任务开始成功。

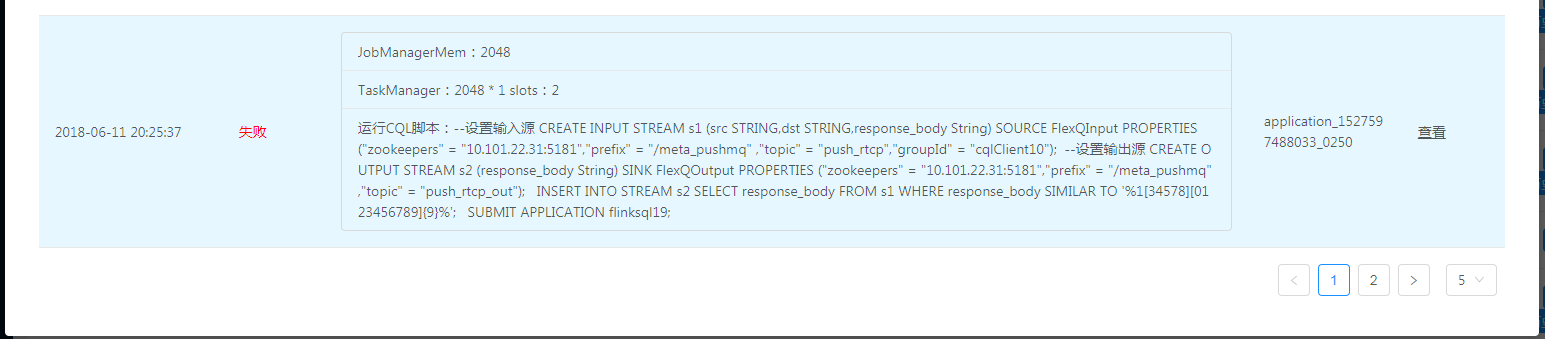


如果任务开始失败，会回滚所分配的所有资源。



无论任务是否提交成功，都会记录提交历史，并记录执行结果。



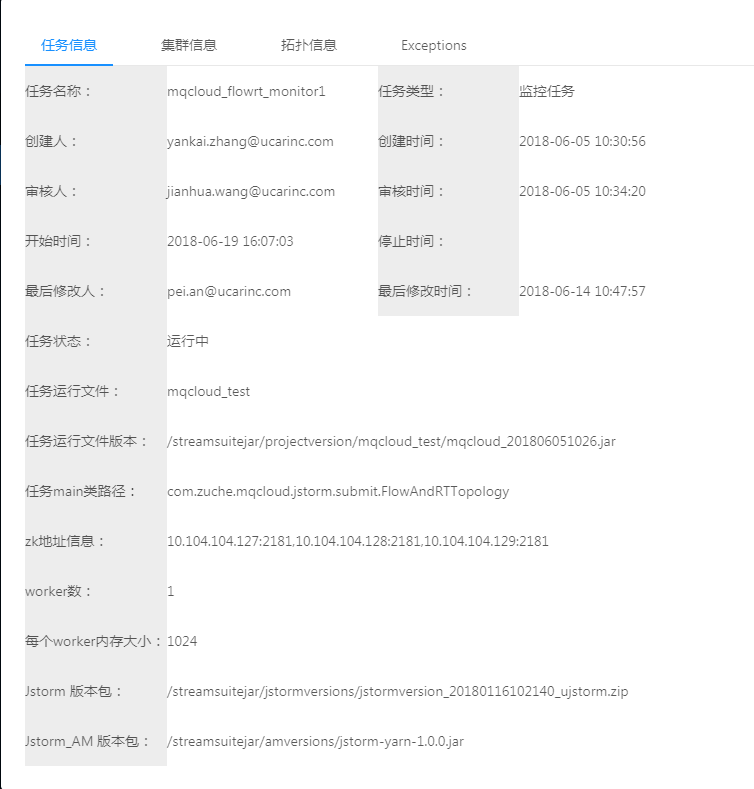


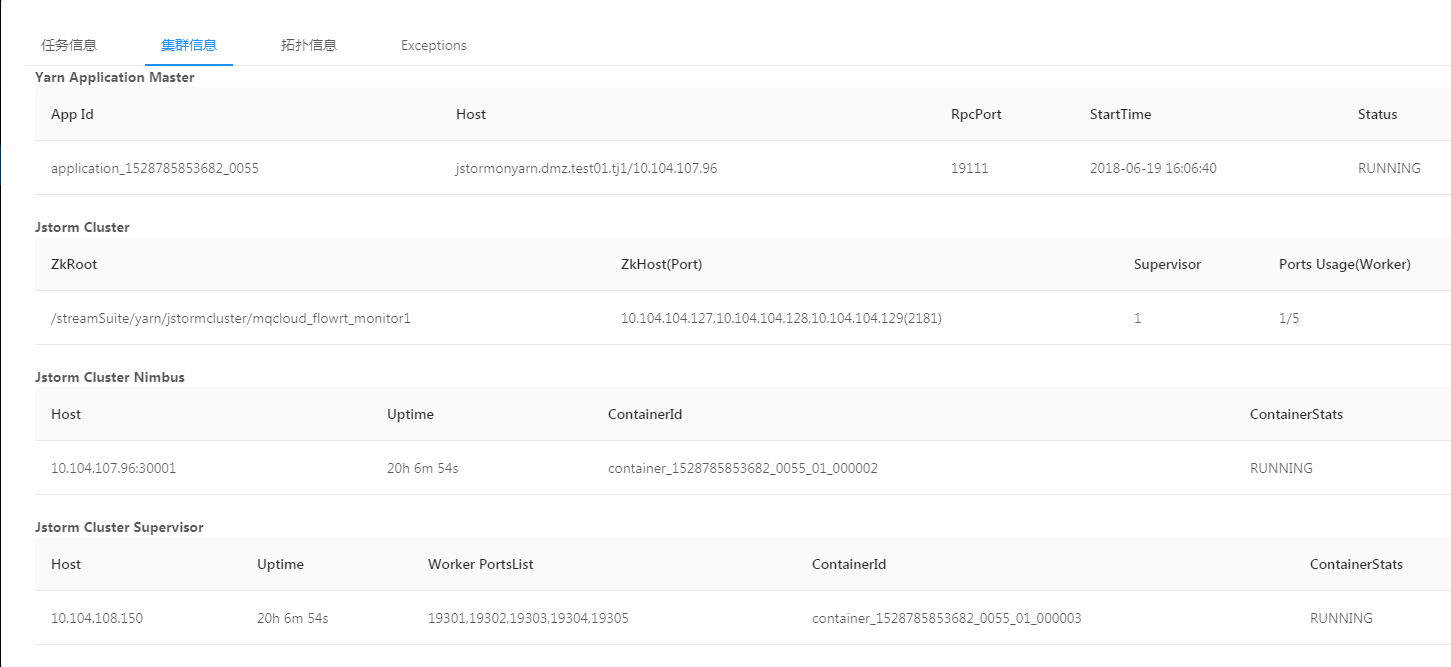
运行中和异常停止的任务可以点击停止暂停任务执行。

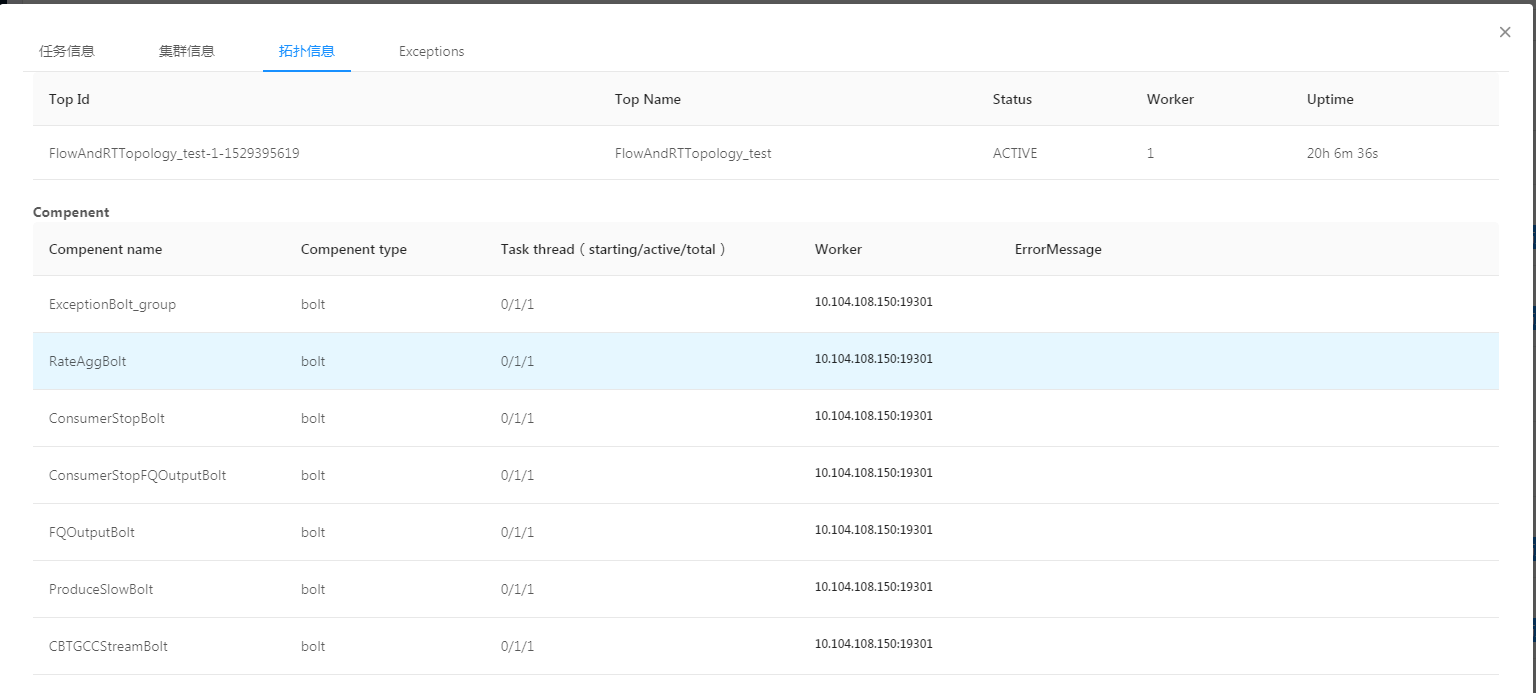


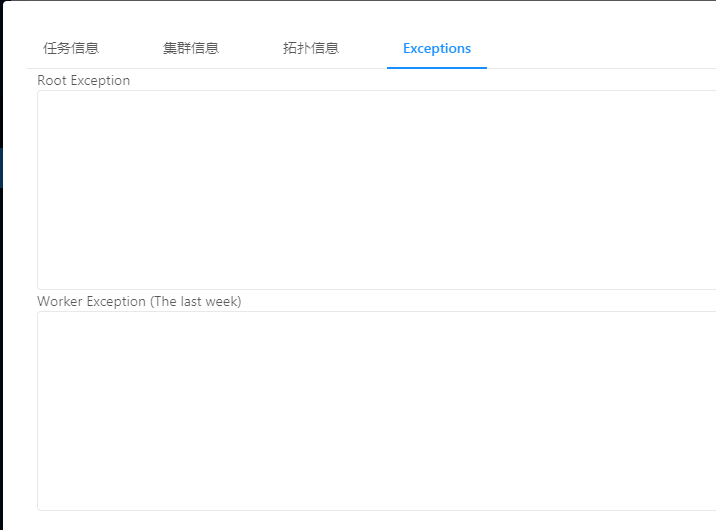
Jstorm任务的详情信息分为4个标签。

任务基本信息，集群信息，拓扑信息和异常信息。







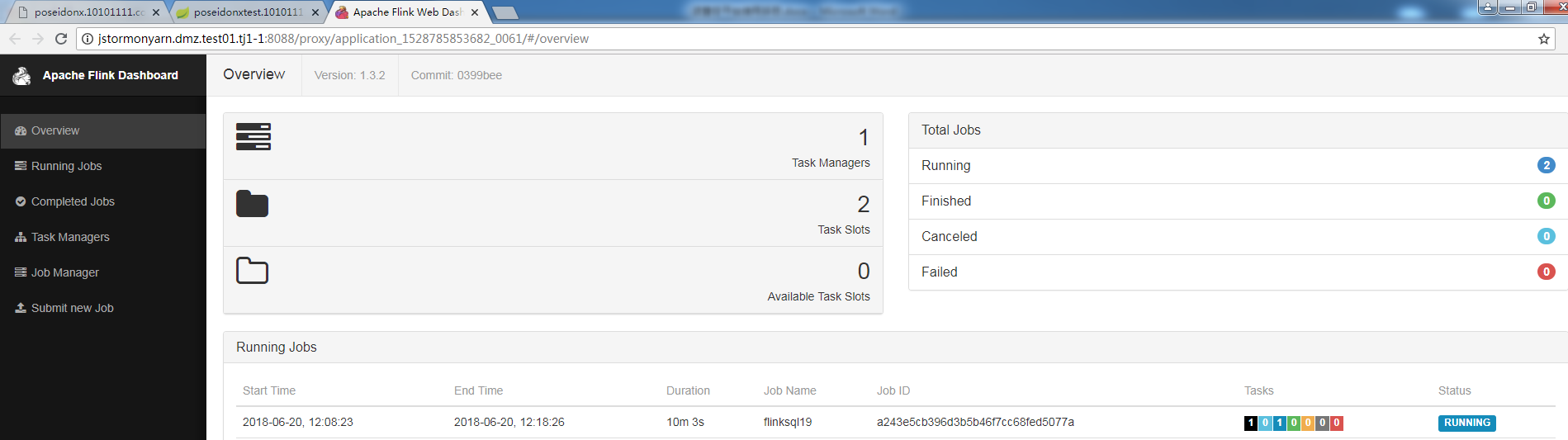


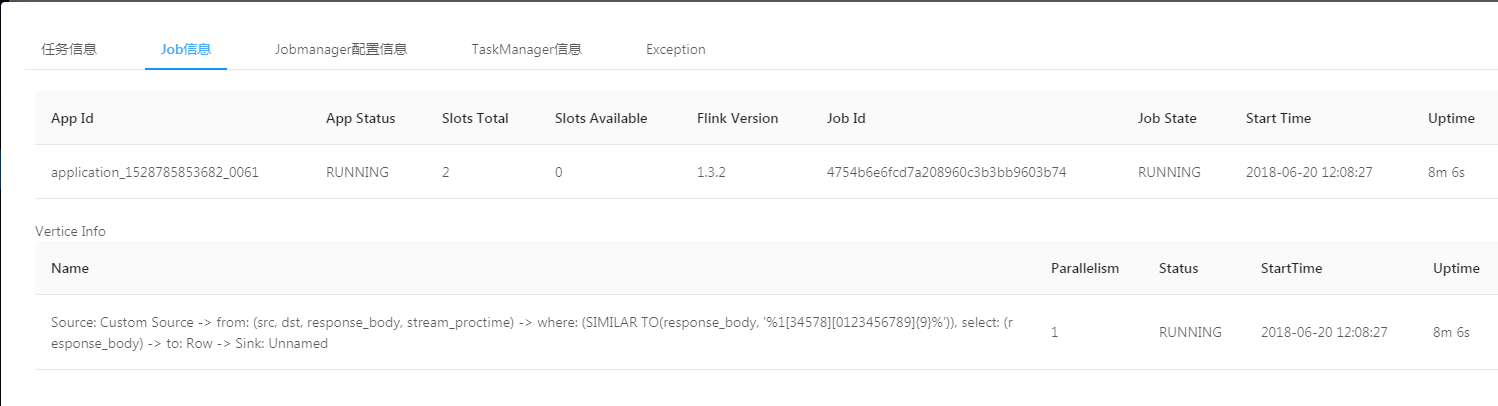
flink任务的详情信息分为5个标签。

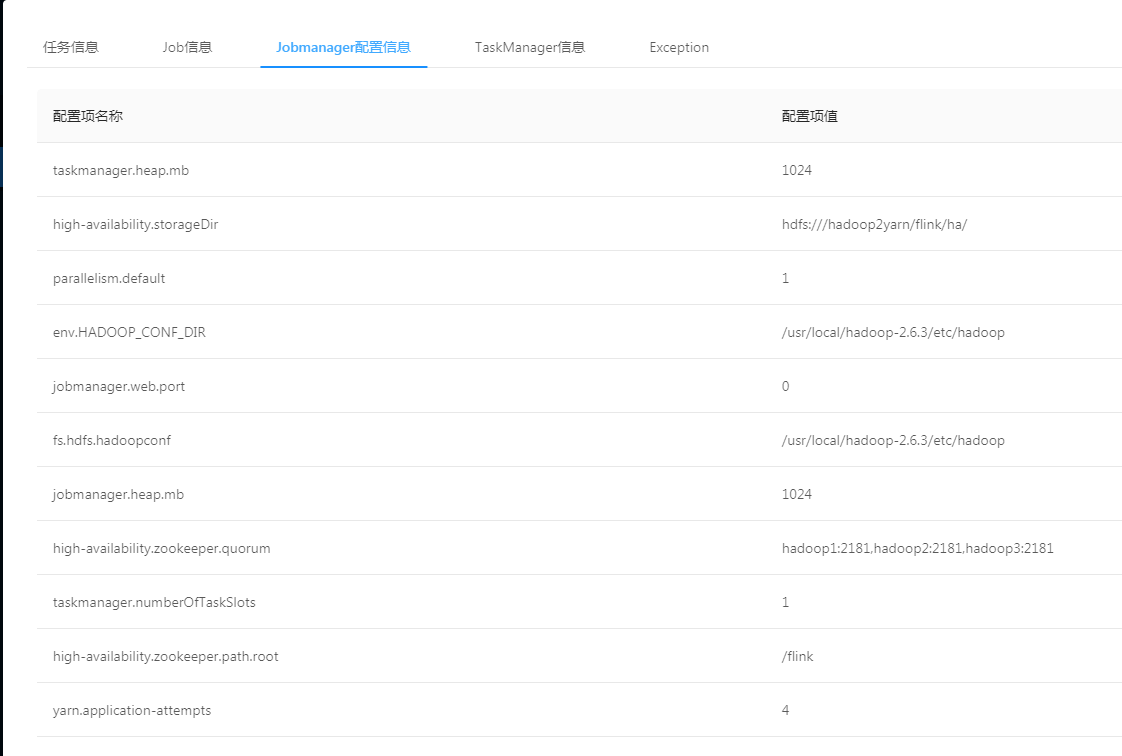
任务基本信息，job信息，jobManager的配置信息，TaskManager信息和异常信息。

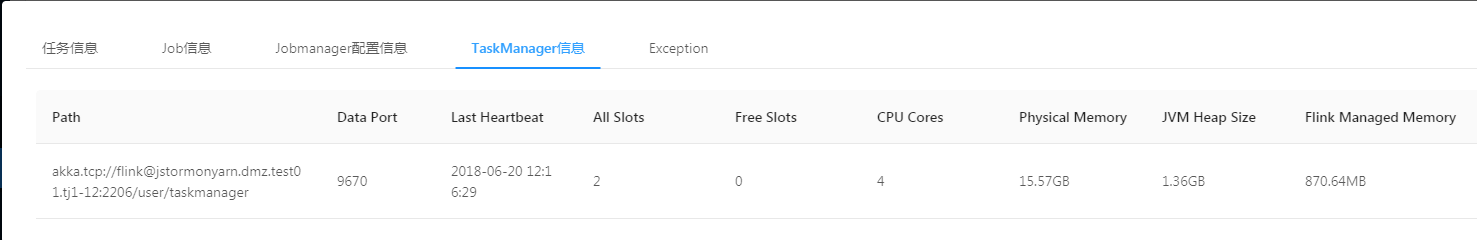


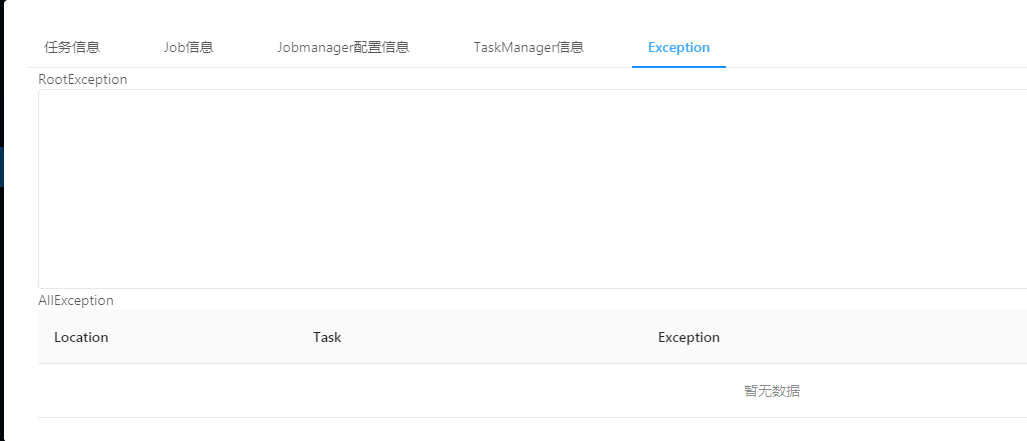
对于运行中的flink任务可以点击查看app控制台，进入Flink控制台页面。











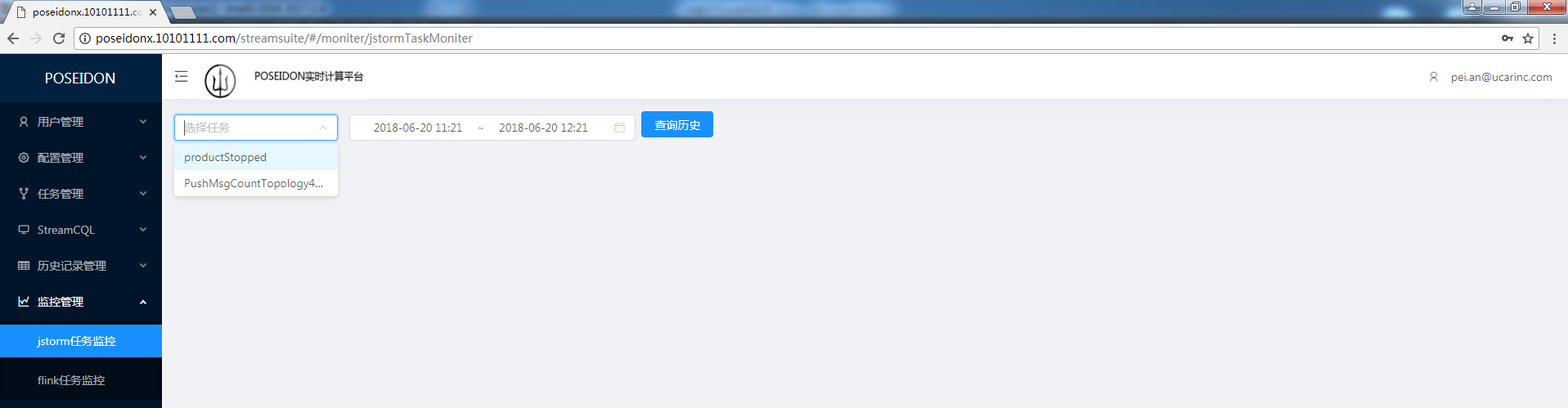
## 5、监控管理

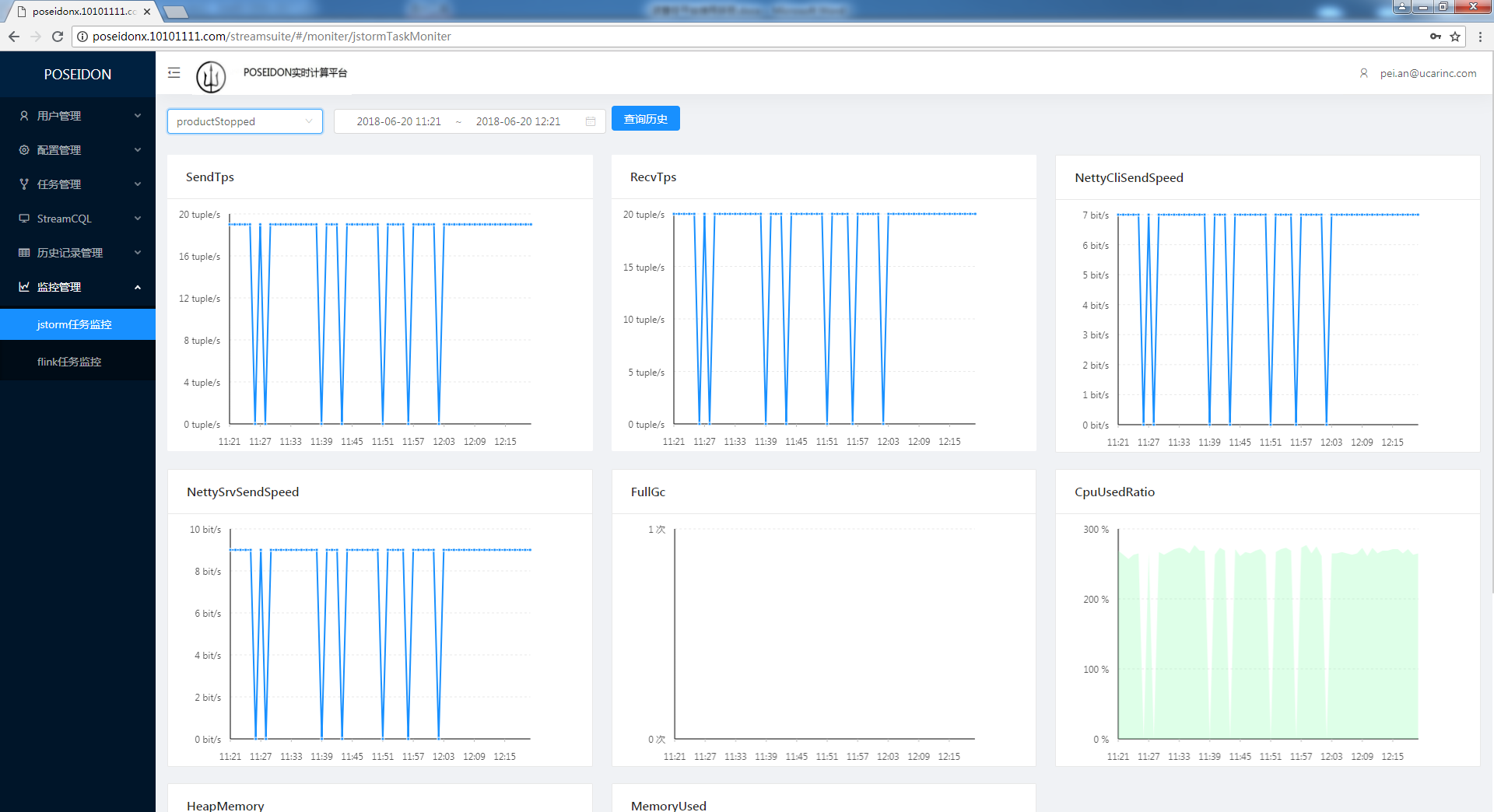
1. **Jstorm任务监控**

监控管理jstorm任务监控功能对jstorm任务的指标进行监控，展示监控曲线。

选择想要查看监控的任务会默认展示，最近一小时的监控曲线。并且曲线会以分钟为单位，每分钟刷新一次。

如果想要查看历史曲线，选择时间段点击查询历史进行查询。

****

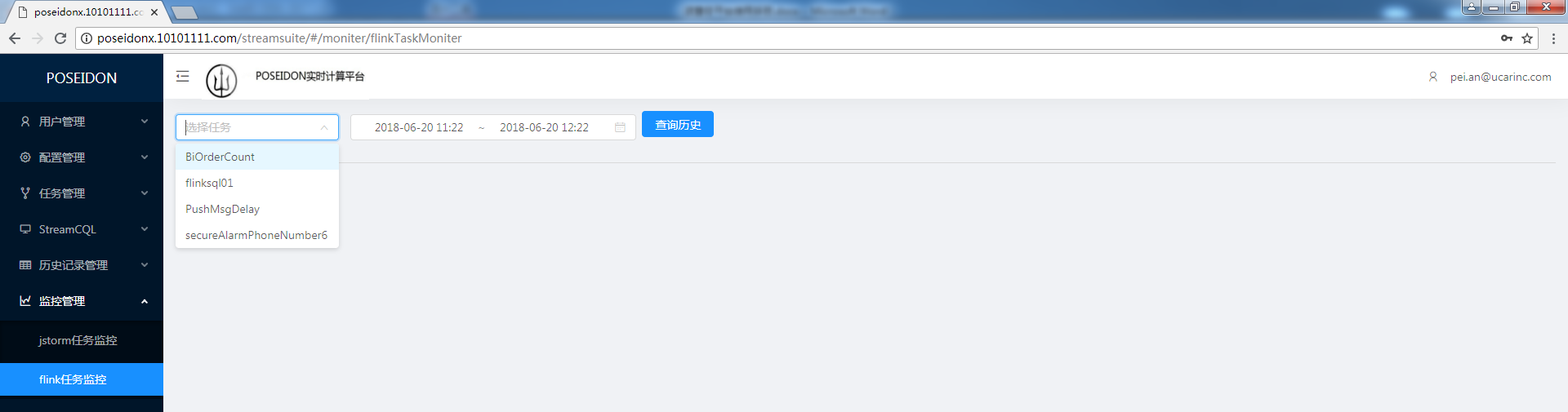


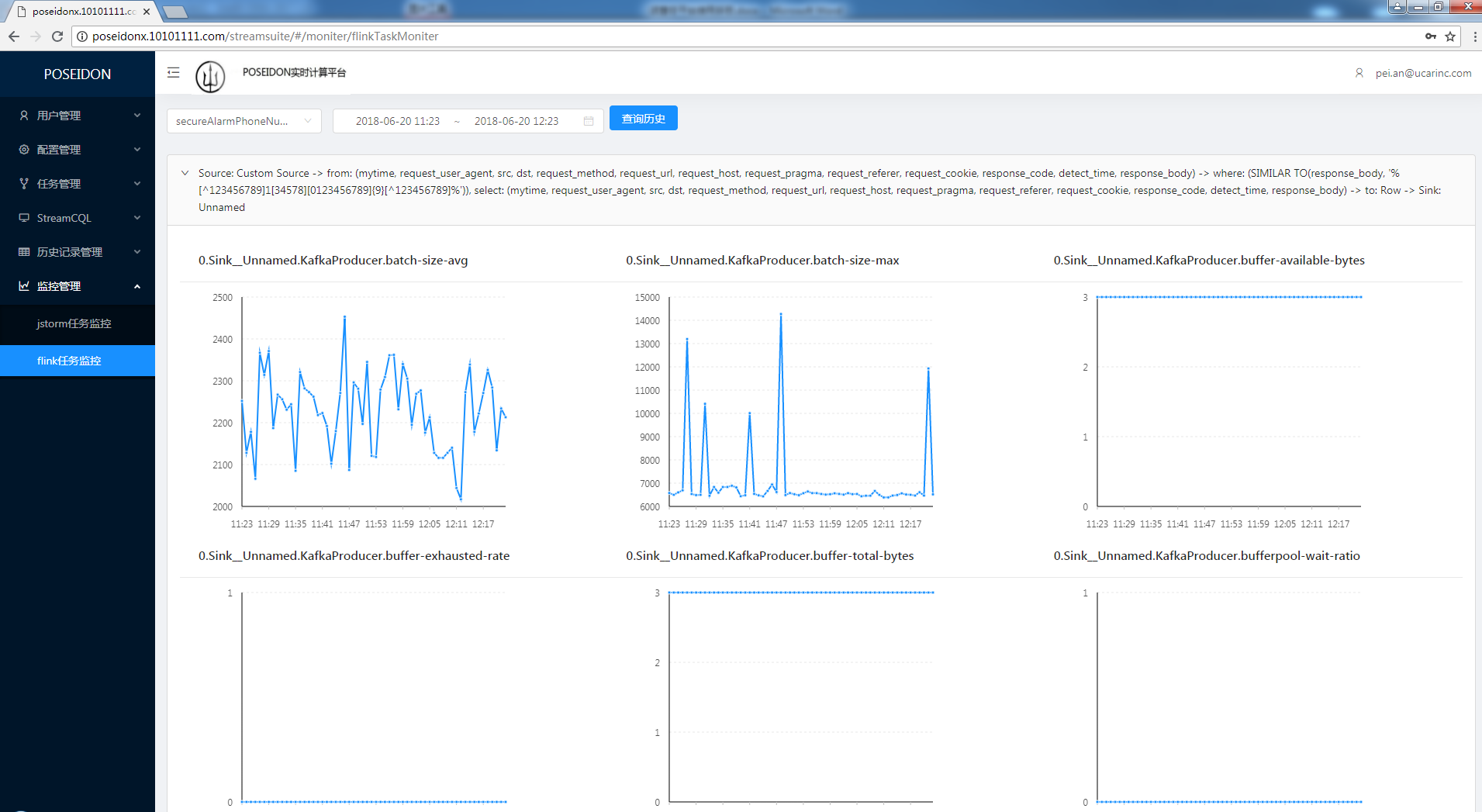
1. **Flink任务监控**

监控管理Flink任务监控功能对flink任务的指标进行监控，展示监控曲线。

选择想要查看监控的任务会默认展示，最近一小时的监控曲线。

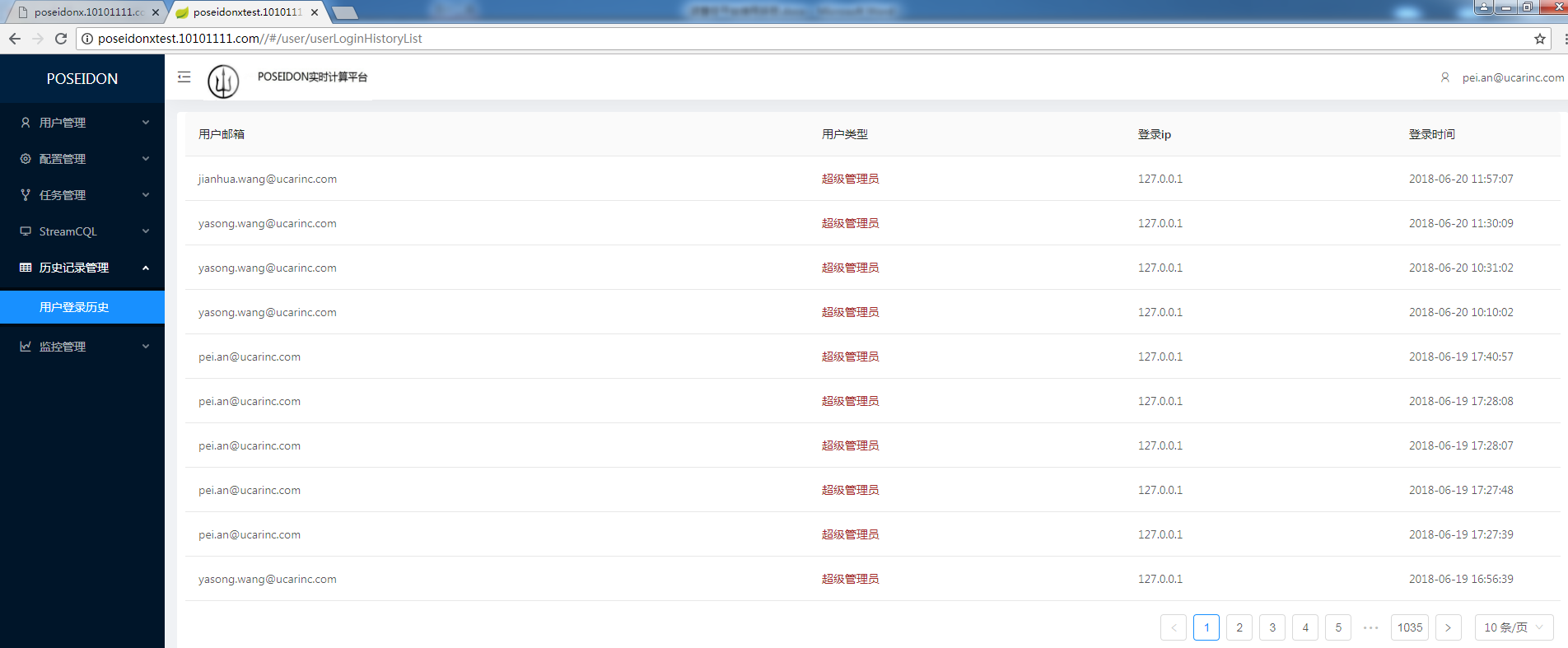
如果想要查看历史曲线，选择时间段点击查询历史进行查询。



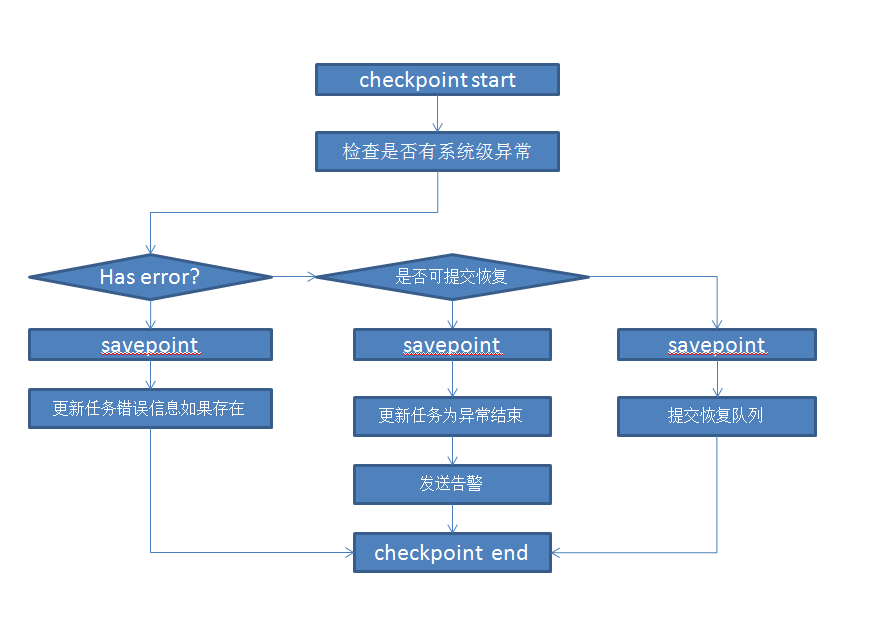


## 6、历史记录管理

1. **用户登录历史**



## 7、任务告警与恢复



每分钟checkpoint流程如图所示

一旦发现任务出现系统级别异常，例如yarn app无法连接等情况。

如果无法提交恢复程序，进行任务保存点存储任务状态，并发送告警。

如果可以提交任务恢复队列，则进入恢复流程。

如果未发现系统系统，则任务保存点存储任务状态，并根据是否存在exception更新任务错误信息。

恢复流程为

1. 恢复队列中拿到待恢复任务。
2. 强制结束所有老任务相关资源。
3. 尝试复活一次。
4. 如果复活失败,则3分钟后再复活，重试N次。重试次数可由“通用配置项”配置。
5. 上述复活全没成功，放弃复活，并发送告警复活失败，更新任务为异常结束。
6. 复活成功，也发送告警提醒任务失败后复活成功。

对于告警信息的发送方式，用户可以自行实现AlarmBusiness接口的实现类，并配置于

streamsuite.properties 文件中。

# 附录

## 通用配置项解释

**配置项名称：**JSTORM\_HOME

**初始值：**/usr/local/jstorm/jstorm-2.2.1/

**是否必填：**是

poseidon服务器安装jstorm软件包的路径。此值可以根据软件包安装实际位置来进行修改

**配置项名称：**FLINK\_HOME

**初始值：**/usr/local/flink-1.3.2

**是否必填：**是

poseidon服务器安装flink软件包的路径。此值可以根据软件包安装实际位置来进行修改

**配置项名称：**LOCAL\_PROJECT\_ITEM\_DIR

**初始值：**/usr/local/tmpjar/

**是否必填：**是

poseidon服务器，本地任务文件版本文件的暂存路径。此值可以根据实际情况来进行修改

**配置项名称：**ZK\_HOST

**初始值：**10.104.104.127,10.104.104.128,10.104.104.129

**是否必填：**是

实时引擎用到的zk地址，此值可以根据实际情况来进行修改。

**配置项名称：**ZK\_PORT

**初始值：**2181

**是否必填：**是

实时引擎用到的zk端口，此值可以根据实际情况来进行修改。

**配置项名称：**HBASE\_ZK\_ROOT

**初始值：**/hbase\_0.98

**是否必填：**是

hbase的zk root，监控曲线等功能都会依赖于hbase，此值可以根据实情况来进行修改。

**配置项名称：**HBASE\_ZK\_PORT

**初始值：**5181

**是否必填：**是

hbase的zk port，监控曲线等功能都会依赖于hbase，此值可以根据实情况来进行修改。

**配置项名称：**HBASE\_ZK\_HOST

**初始值：**10.101.22.31

**是否必填：**是

hbase的zk host，监控曲线等功能都会依赖于hbase，此值可以根据实情况来进行修改。

**配置项名称：**EMAIL\_ALARM\_OPEN

**初始值：**1

**是否必填：**不填写时为关闭

邮箱告警总开关，1为打开0为关闭

**配置项名称：**PHONE\_ALARM\_OPEN

**初始值：**1

**是否必填：**不填写时为关闭

手机告警总开关，1为打开0为关闭

**配置项名称：**CQL\_ENGINE\_PROJECT\_NAME

**初始值：**streamCQL

**是否必填：**是

CQL引擎的任务文件的名称。此名称的任务文件会被识别为是CQL引擎处理。

**配置项名称：**JSTORM\_TASK\_LOG\_PREFIX

**初始值：**/data/log/hadoop/

**是否必填：**请务必使用初始值的目录位置，否则可能存在兼容性问题

服务器jstorm任务日志前缀。

**配置项名称：**JSTORM\_TOP\_START\_TIMEOUT

**初始值：**150

**是否必填：**默认150秒

JSTORM TOP任务启动的最大等待时间。超过此时间认为启动失败 (秒)。

注意这个时间不是整个任务启动的时间，还包括45秒的yarn检查时间，和nimbus检查时间(最大30秒，重试10次)

**配置项名称：**JSTORM\_AUTO\_RECOVERY\_RETRY

**初始值：**3

**是否必填：**默认重试3次

JSTORM 任务失败的时候，自动恢复的重试次数

这里配置初始值的3指的是如果恢复失败的时候，进行的重试次数，加上默认的恢复1次就是尝试恢复4次。

**配置项名称：**JSTORM\_CLUSTER\_RECOVERY\_FLAG

**初始值：**0

**是否必填：**默认为0

当Nimbus无心跳超时，是否强制恢复。1打开，0关闭

这个值指的是每分钟checkpoint执行时，经检测多次无法获取到nimbusClient，是否进行强制恢复。

**配置项名称：**JSTORM\_NIMBUS\_MEM

**初始值：**2048

**是否必填：**默认1024M

Jstorm 任务，启动时为Nimbus Container分配的内存数

**配置项名称：**FLINK\_JOB\_MANGER\_MEM

**初始值：**2048

**是否必填：**默认2048M

Flink任务启动时，为JOB\_MANGER 分配的内存数

**配置项名称：**FLINK\_JOB\_START\_TIMEOUT

**初始值：**60

**是否必填：**默认60秒

FLINK\_JOB 启动的最大等待时间。超过此时间认为启动失败 (秒) 默认为1分钟

注意这个时间不是yarn-session的最大启动等待时间，yarn-session的最大启动等待时间为90秒。

**配置项名称：**FLINK\_AUTO\_RECOVERY\_RETRY

**配置值：**3

**是否必填：**不填写时为默认重试3次

FLINK任务失败的时候，FLINK 自动恢复的重试次数(默认为3次)。

这里配置初始值的3指的是如果恢复失败的时候，进行的重试次数，加上默认的恢复1次就是尝试恢复4次。

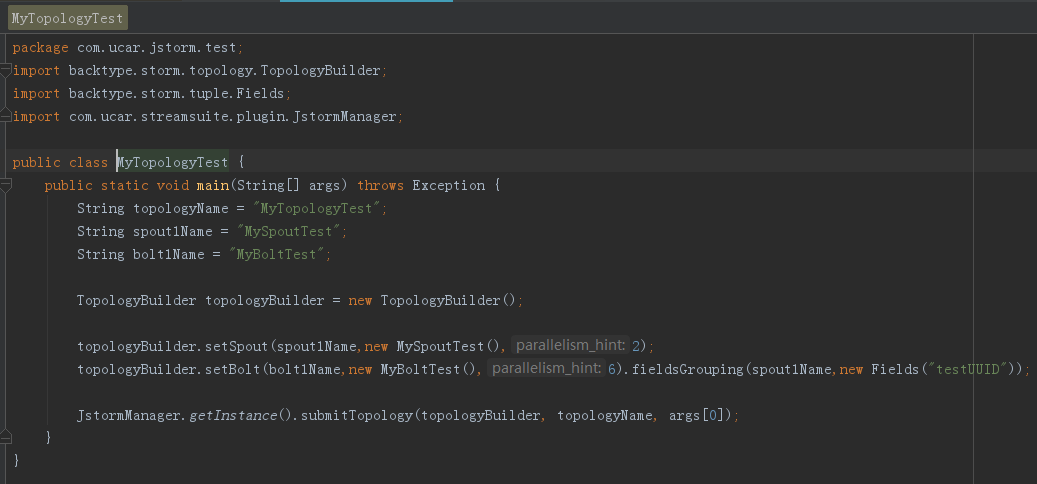
## Jstorm任务开发插件

编写 jstorm任务需要使用平台提供的插件来提交任务。否则平台无法对任务的稳定性进行保证，也无法对任务进行监控。

pom引入

<dependency>  
 <groupId>com.ucar.streamsuite</groupId>  
 <artifactId>jstorm-plugin</artifactId>  
 <version>1.0.0</version>  
</dependency>

任务开发案例如图



并发度可以从任务创建页面进行设置，如果代码中设计了则以代码为准，覆盖页面设置的值。