Movert动态采集器

scraj-\*+sfe4j

# 第1章 简介

movert是一个使用浏览器进行采集的项目，使用selenium的webdriver操作浏览器（firefox），完成对动态渲染（js/ajax）网页的数据采集。由于使用的浏览器来采集，所以采集机器上需要先安装firefox浏览器，然后才能使用，另外资源占用也会比较大，并且采集速度也会较慢。

使用的三个rmi服务和uyint共用一套，不需要重新启动rmi服务。

# 第2章 安装部署

## 2.1 准备工作

参考uyint，两个是一样的。

需要额外注意的是，movert依赖firefox浏览器，所以前提是先安装firefox。

## 2.2 启动movert启动线程

对应的启动类：com.hiekn.scraj.movert.fix.MovertFixApp

# 第3章 movert项目编译

相关项目：sfe4j、scraj-common、scraj-uyint-common、scraj-movert-common、scraj-movert

# sfe4j是一个单独项目，需要首先编译它，因为后面有项目依赖它。

# scraj-movert-common依赖 scraj-uyint-common和scraj-common。

# scraj-common、和scraj-uyint-common这个两个项目之间没有相互依赖，先编译谁都可以。

# scraj-movert-common 是movert的核心。

# scraj-movert这个项目包含爬虫任务，以及爬虫线程等。