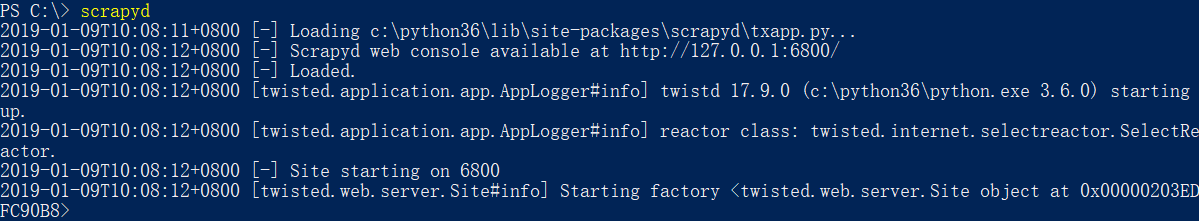
1. 准备环境

1. pip3 install scrapy
2. pip3 install scrapyd
3. pip3 install scrapyd-client
4. pip3 install scrapy-redis
5. pip3 install SpiderKeeper

2.打开cmd进入scrapy项目中(第一个项目文件夹内)

cd 到项目目录下

在打开一个cmd输入scrapyd，界面如下，浏览器<http://localhost:6800/>



出现以下界面：



在爬虫项目下新建一个scrapyd.conf文件：(与scrapy.cfg同级)

[scrapyd]

eggs\_dir = eggs

logs\_dir = logs

items\_dir =

jobs\_to\_keep = 5

dbs\_dir = dbs

max\_proc = 0

max\_proc\_per\_cpu = 4

finished\_to\_keep = 100

poll\_interval = 5.0

bind\_address = 0.0.0.0

http\_port = 6800

debug = off

runner = scrapyd.runner

application = scrapyd.app.application

launcher = scrapyd.launcher.Launcher

webroot = scrapyd.website.Root

[services]

# 启动项目

schedule.json = scrapyd.webservice.Schedule

# 0.15版本的新功能。取消蜘蛛游戏（又名作业）。如果作业处于待处理状态，则会将其删除。如果作业正在运行，它将被终止。

cancel.json = scrapyd.webservice.Cancel

# 将项目添加到项目中，如果项目不存在则创建项目。

addversion.json = scrapyd.webservice.AddVersion

# 获取上传到此Scrapy服务器的项目列表。

listprojects.json = scrapyd.webservice.ListProjects

# 获取某些项目可用的版本列表。版本按顺序返回，最后一个版本是当前使用的版本。

listversions.json = scrapyd.webservice.ListVersions

# 获取某个项目的最后一个（除非被覆盖）版本中可用的蜘蛛列表。

listspiders.json = scrapyd.webservice.ListSpiders

# 删除项目及其所有上载的版本。

delproject.json = scrapyd.webservice.DeleteProject

# 删除项目版本。如果给定项目没有更多可用版本，则该项目也将被删除。

delversion.json = scrapyd.webservice.DeleteVersion

# 0.15版本的新功能。获取某个项目的待处理，正在运行和已完成的作业列表。

listjobs.json = scrapyd.webservice.ListJobs

# 检查服务的负载状态。

daemonstatus.json = scrapyd.webservice.DaemonStatus

## 启动spiderkeeper

## cmd下运行，浏览器http://localhost:5000，必须在项目目录下运行命令界面如下：spiderkeeper

账号：admin 密码：admin

## 

4.这时需要python安装目录的Scripts下新建两个批处理文件。(如果是虚拟环境则放在虚拟环境下)

scrapy.bat和scrapyd-deploy.bat

scrapy.bat:

**@echo off**

**C:\Python36\python.exe C:\Python36\Scripts\scrapy %\***

scrapyd-deploy.bat：

@echo off

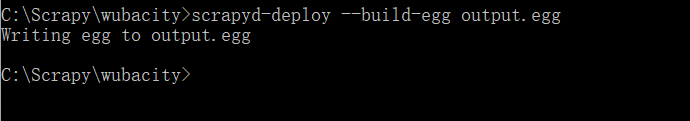
C:\Python36\python.exe C:\Python36\Scripts\scrapyd-deploy %\*

运行命令启动scrapyd：（在第一个项目文件夹下）

Scrapyd-deploy -p 项目文件名

5.项目部署到Spiderkeeper上：

cmd中使用scrapyd-deploy --build-egg output.egg



然后打开spiderkeeper的页面，点击deploy，点击create project创建新项目city

**二．spiderkeeper监控多个爬虫：**

1. 修改scrapyd安装模块内的 default\_scrapyd.conf 文件配置：

改为：

bind\_address = 0.0.0.0

1. 修改spiderkeeper模块内的confid.py文件配置：（添加）

连接设备数量：

THREADS\_PER\_PAGE = 2

连接设备的ip及端口

SERVERS = ['http://localhost:6800','http://192.168.8.95:6800']

Master端的spiderkeeper的启动：连接多台scrapyd服务器.

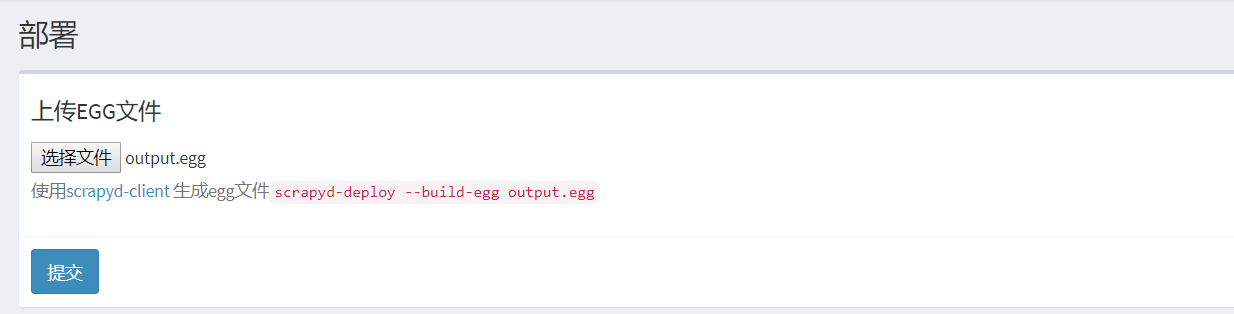
(在分布式中我们的服务器中肯定不止一台,使用spiderkeeper:)

**spiderkeeper --server=http://localhost:6800 --server=http://111.111.111.111:6800**

#启动一个spiderkeeper可以同时部署两台服务器的spider

备注：所连设备必须开启scrapyd服务和spiderkeeper服务







运行爬虫，结束。