

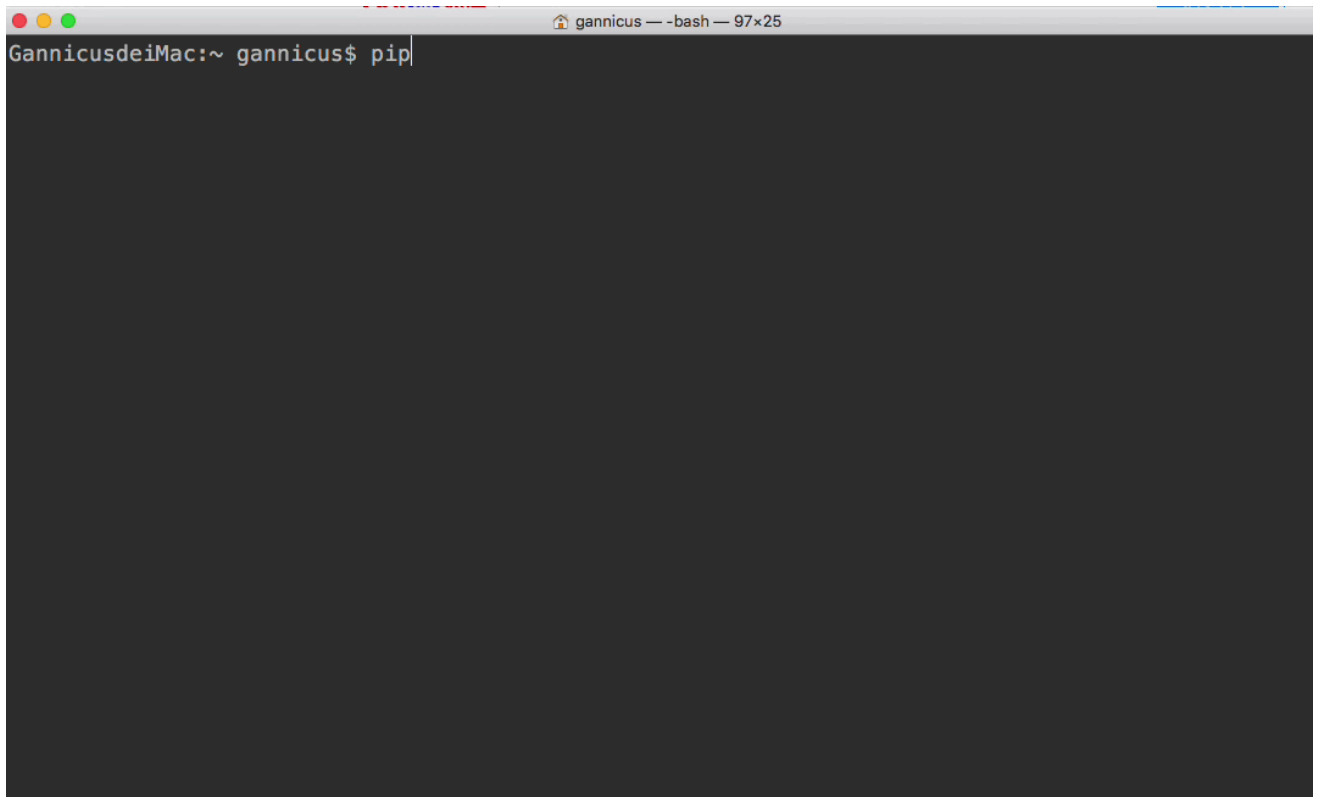
Scrapyd 的安装与启动

在第一小节中了解到 Scrapyd 的应用场景与学习的必要性之后，这一节我们就来学习如何安装以及启动 Scrapyd。

安装

Scrapyd 的安装比较简单，在[官方文档](#)的示例中有说明，我们只需要在有 Python 的环境下通过 pip 进行安装即可：

```
1 pip install scrapyd
2
```



启动

安装好之后，只需要在 Python 环境下输入命令：

```
1 scrapyd
2
```

```

Requirement already satisfied: PyHamcrest>=1.9.0 in ./anaconda3/lib/python3.6/site-packages (from Twisted>=8.0->scrapyd)
Requirement already satisfied: attrs>=17.4.0 in ./anaconda3/lib/python3.6/site-packages (from Twisted>=8.0->scrapyd)
Requirement already satisfied: pyasn1-modules in ./anaconda3/lib/python3.6/site-packages (from service-identity->Scrapy>=1.0->scrapyd)
Requirement already satisfied: pyasn1 in ./anaconda3/lib/python3.6/site-packages (from service-identity->Scrapy>=1.0->scrapyd)
Requirement already satisfied: cryptography>=2.1.4 in ./anaconda3/lib/python3.6/site-packages (from pyOpenSSL->Scrapy>=1.0->scrapyd)
Requirement already satisfied: idna>=2.5 in ./anaconda3/lib/python3.6/site-packages (from hyperlink>=17.1.1->Twisted>=8.0->scrapyd)
Requirement already satisfied: setuptools in ./anaconda3/lib/python3.6/site-packages (from zope.interface>=4.4.2->Twisted>=8.0->scrapyd)
Requirement already satisfied: asn1crypto>=0.21.0 in ./anaconda3/lib/python3.6/site-packages (from cryptography>=2.1.4->pyOpenSSL->Scrapy>=1.0->scrapyd)
Requirement already satisfied: cffi>=1.7 in ./anaconda3/lib/python3.6/site-packages (from cryptography>=2.1.4->pyOpenSSL->Scrapy>=1.0->scrapyd)
Requirement already satisfied: pycparser in ./anaconda3/lib/python3.6/site-packages (from cffi>=1.7->cryptography>=2.1.4->pyOpenSSL->Scrapy>=1.0->scrapyd)
Installing collected packages: scrapyd
Successfully installed scrapyd-1.2.0
You are using pip version 9.0.1, however version 18.1 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
GannicusdeiMac:~ gannicus$ sc

```

回车后即可看到 Scrapyd 成功启动并显示的运行信息。

```

1 2018-10-04T16:45:03+0800 [-] Scrapyd web console available at http://127.0.0.1:6800/
2 2018-10-04T16:45:03+0800 [-] Loaded.
3 2018-10-04T16:45:03+0800 [twisted.scripts._twistd_unix.UnixAppLogger#info] twistd 18.7.0
  (/Users/gannicus/anaconda3/bin/python 3.6.4) starting up.
4 2018-10-04T16:45:03+0800 [twisted.scripts._twistd_unix.UnixAppLogger#info] reactor class:
  twisted.internet.selectreactor.SelectReactor.
5 2018-10-04T16:45:03+0800 [-] Site starting on 6800
6 2018-10-04T16:45:03+0800 [twisted.web.server.Site#info] Starting factory
  <twisted.web.server.Site object at 0x107595ba8>
7 2018-10-04T16:45:03+0800 [Launcher] Scrapyd 1.2.0 started: max_proc=16,
  runner='scrapyd.runner'
8

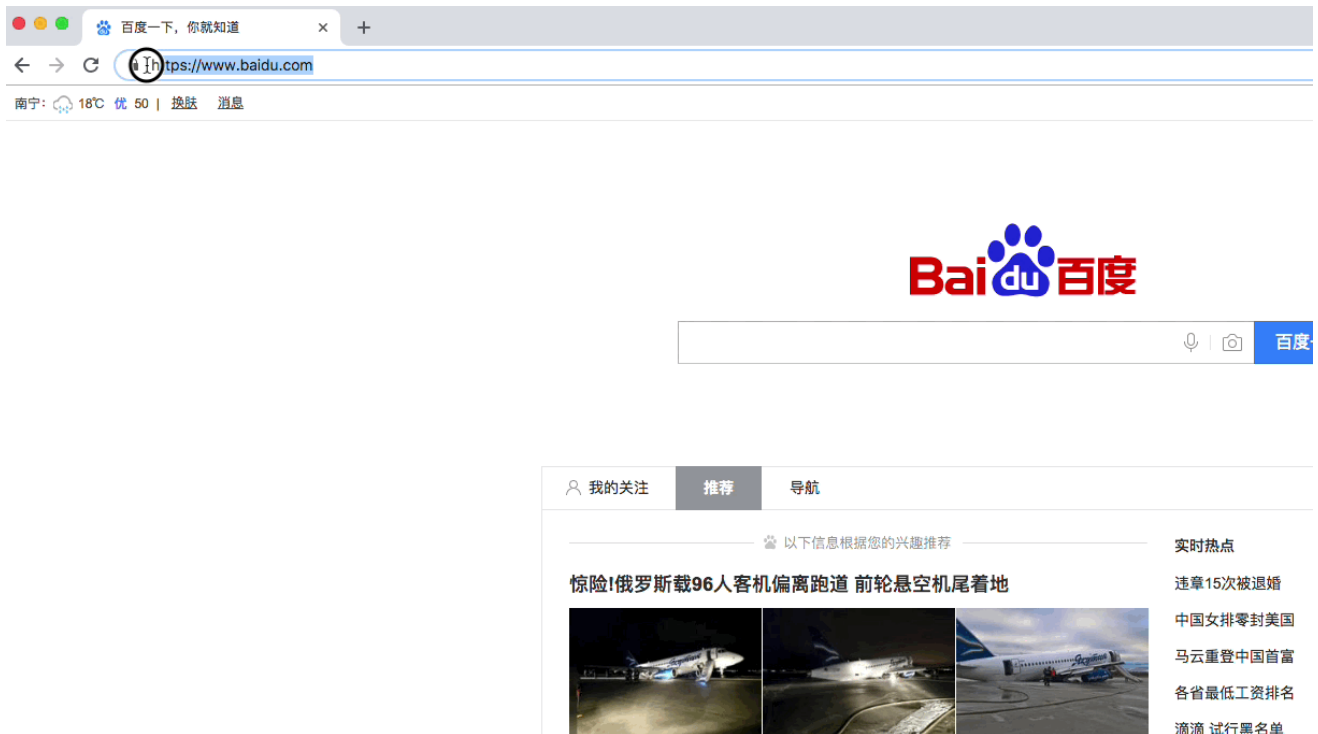
```

而后打开浏览器，输入地址：

```

1 http://localhost:6800
2

```



如果看到下图画面，即代表 Scrapyd 服务已正常运行：

Scrapyd

Available projects:

- [Jobs](#)
- [Logs](#)
- [Documentation](#)

How to schedule a spider?

To schedule a spider you need to use the API (this web UI is only for monitoring)

Example using [curl](#):

```
curl http://localhost:6800/schedule.json -d project=default -d spider=somespider
```

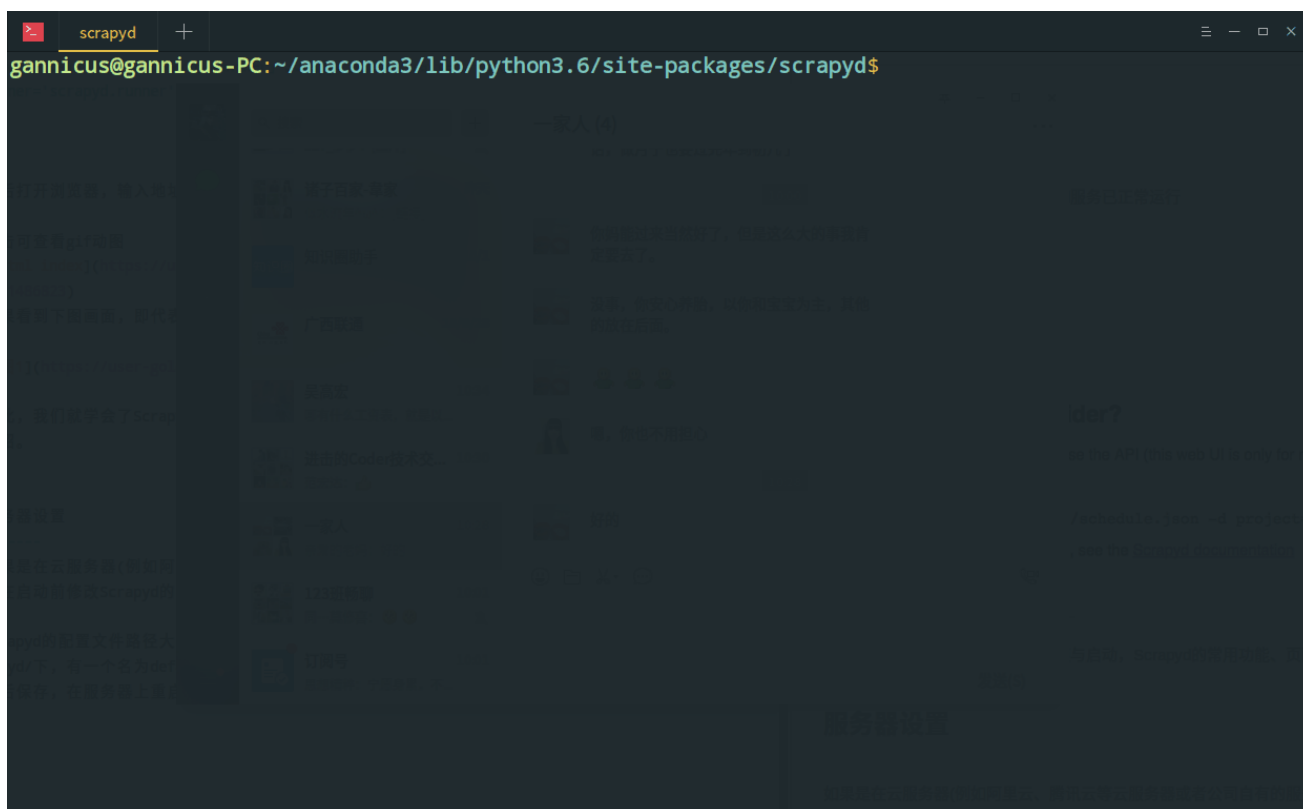
For more information about the API, see the [Scrapyd documentation](#)

至此，我们就学会了 Scrapyd 的安装与启动。关于 Scrapyd 的常用功能、页面结构等知识点笔者将在后面的小节中介绍。

用 Scrapyd 在服务器上开启远程访问

Scrapyd **默认只允许本机访问**，意味着当你在服务器上使用 Scrapyd 时，如果不开启远程访问，那你在自己的电脑上是无法访问 Scrapyd 的服务的。

如果是在云服务器(例如阿里云、腾讯云等云服务器或者公司自有的服务器)上想要启动 Scrapyd 并且你可以在任何地方通过 `ip:6800` 访问 Scrapyd 服务，则需要在启动前修改 `Scrapyd` 的配置文件，将 `bind_address` 从原来的 `127.0.0.1` 改为 `0.0.0.0`。



Scrapyd 的配置文件路径一般都会放在 Python 环境中包的安装目录内，也就是 `/site-packages/scrapy/` 下，有一个名为 `default_scrapy.conf` 的文件，其中有个设 `bind_address`，将其修改为 `0.0.0.0` 后保存，然后在服务器上重启 Scrapy 服务即可。