# 1 安装python包管理器

下载文件

wget https://bootstrap.pypa.io/get-pip.py --no-check-certificate2

执行安装

python get-pip.py

# 2 安装shadowsocks(连接vps的socks5代理)

sudo pip install shadowsocks

然后执行：

防火墙开启端口：firewall-cmd --permanent --add-port=1080/tcp

重启防火墙：firewall-cmd --reload

sslocal -s 192.126.123.195 -p 18000 -k 123abc123@mxx -m aes-256-cfb -l 1080 -b 127.0.0.1

其中 -s 是vps的ip -p socsk服务端的监听端口 -k 密钥 -m 加密方式

-l 客户端监听端口 -b 客户端监听IP地址

# 3 安装privoxy (将socks5代理转换成http代理)

yum install privoxy

如果找不到安装包先执行：

rpm -Uvh <http://dl.fedoraproject.org/pub/epel/epel-release-latest-7.noarch.rpm>

后在进行安装。

修改配置文件： vim /etc/privoxy/config

在文件中添加：

listen-address 0.0.0.0:8118

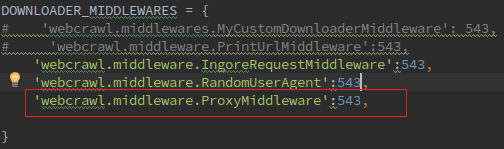
forward-socks5 / 127.0.0.1:1080 .

保存文件后重启privoxy : privoxy /etc/privoxy/config

其中listen-address表示监听的ip地址和端口号，可以将ip改成127.0.0.1表示只监听本地ip

# 4 爬虫配置http代理

爬虫源码的seething.py中添加http代理



注：webcrawl只是表示当前爬虫的项目名称，变更爬虫项目需要更改。