

# myAb压力测试

## 概述

myAb为投递搜狐Python实习的一个在线题目，题目内容为：  
编写一个 ab(Apache Benchmark)的替代程序，并尽可能兼容 ab 的参数。

ab程序的主要功能和参数可参考 <http://mo2g.com/view/38/>，下面简单的说明一下 myAb 程序的运行结果以及兼容的主要参数。

在本机命令行输入 `./myAb.py -h -V -c 10 -n 20 -e www.baidu.com`  
能够得到如下输出：

```
xiyuanBuptdeMacBook-Pro:ab xiyuanbupt$ ./myAb.py -h -V -c 10 -n 20 -e www.baidu.com
-A 采用base64编码验证身份
-c 并发访问的用户数
-n 一个用户访问次数
-H 设置代理浏览器
-t 在一段时间内访问
more 更多信息参考ab的使用
VERSION: 0.0
200 requests      : 20
Time per request  : 85.411 [ms]
Server Software   : BWS/1.1
Time taken for tests : 1.708212 seconds
Document Path     : /
Requests per second : 11.708 [#/sec]
400 requests      : 0
Document Length   : 1951966 bytes
Server Hostname    : www.baidu.com
HTML transferred  : 1951966.000 bytes
300 requests      : 0
Sever Port         : 443
100 requests      : 0
Transfer rate      : 1138.066 [Kbytes/sec]
600 requests      : 0
500 requests      : 0
total requests    : 20
Total transferred程序的: 1990715.000 bytes

Connection Times (ms)
item    min      mean     median    max
Waiting 30       162      92        1006
Total    84       506      245       1708
Connect  6        10       10        33
Create file 10_user_20_visit_host_www.baidu.com_2015-09-26_18:32:27.415020.csv
xiyuanBuptdeMacBook-Pro:ab xiyuanbupt$
```

## myAb 兼容的一些功能

- n 设置请求的总次数
- c 设置请求的并发数
- H 设置浏览器代理
- A 采用base64 编码向服务器提供身份验证
- C cookie 信息
- e 保存为csv格式，文件名 由 myAb 自己产生，会在执行结果中显示
- g 保存为TSV
- h 显示ab可选参数列表
- i 发送HEAD请求，默认发送GET请求
- p 通过POST发送的数据（urlencode后的数据）
- s 使用https发送请求
- V 显示myAb 的版本号
- w 采用HTML表格打印结果
- X 设置代理服务器

## myAb 的实现思路

1、用于发送http请求的库有很多，包括urllib urllib2 httplib httpplib2 requests pycurl，pycurl也是在用了urllib2 httplib2 很久之后发现才选择了，原因如下：pycurl接口更丰富，pycurl是 curl 的 py 版，发送请求的速度更快。后续需要考虑在传输层是否需要完成一些工作。

2、我将 myAb 的实现分为三部分：

一、部分是发送请求获得各种请求参数的部分，这一部分的输入接口是例如用户数量，请求次数，请求方式，浏览器代理，cookie，post数据（如果为post方法）等内容，输出的接口为通过发送http请求后获得一系列参数有结构的字典。对应aRequest.py pressurer.py 两个文件

二、数据展示部分，比较简单的一个部分，输入为一模块输出的字典dict，并且数据展示部分设计的比较通用，只要`type(dict)==type({)}`，以及 `type(dict[key])==type({})`，以及 `type(dict[key][key])==type({})` 的形式都能以合理的格式输出，包括csv，html，tsv，标准输出等。对应show.py 文件

三、逻辑控制模块，用于控制运行逻辑。对应myAb.py 文件

3、模块一具体实现是，首先实现一个用户发送n次请求，获得一个参数的列表，列表元素为每一次请求的情况，对于多个用户同时发送请求，控制逻辑为执行多个线程，并将每一个线程的结果输出至一个全局变量（对应于这些线程为全局）保存，通过分析逻辑获得运行的数据。

4、后续回顾程序的设计，觉得不好的一点是在模块一就过滤了很多数据，只根据现在的需求把需要的数据给提取出来了，这对以后程序的扩展不是方便，好的做法是在数据展示阶段过滤不需要的数据，而且数据展示部分我也将接口设计为不需要展示的数据就不会展示。以后再写程序应当注意这点。

5、myAb 并未涉及到python 传输层的模块，另外模拟多用户是用多线程模拟的，不知是否会有问题。

6、代码中也有一些亮点，有一些程序员“偷懒”又有效的做法，欢迎有兴趣的人阅读并指正错误。

## myAb 的环境依赖

编写环境

系统 OS X 10.10.5

python -V : Python 2.7.10

依赖包 pycurl tablib