

nsd_1908_devops_day01

多进程

- 编写的程序，默认是单进程、单线程的。程序只能按编写的序列一条条地扫行。
- 执行指令时，父进程将自身资拷贝一份，生成子进程，指令在子进程中运行，称作fork
- 在python中，os.fork()实现多进程
- 注意：windows不支持多进程编程
- os.fork()的返回值是数字。在父进程中这个数字是非0值（子进程的PID号），在子进程中，这个数字是0。

多进程编程思路

- 父进程只负责产生子进程
- 子进程负责做具体的工作
- 子进程工作结束后，需要彻底退出结束

多线程编程

- 程序是在磁盘上存储的一些可执行文件
- 程序运行时，将会在内存中出现一到多个进程。所以说，进程可以认为是加载到内存中的一系列指令。
- 一个进程中，可以包含多个线程。
- 每个进程都有自己独立的运行空间，但是线程共享进程的运行空间。
- windows支持线程

多线程编程思路

- 主线程只负责产生工作线程
- 工作线程负责做具体的工作

urllib模块

- 包含了4个子模块：request / error / parse / robotparse
- 常用于http / ftp客户端编程

```
>>> from urllib import request
>>> html = request.urlopen('http://www.163.com')
>>> html.read(10)
b' <!DOCTYPE'
>>> html.readline()
b' HTML>\n'
>>> html.readlines()
```

模拟客户端访问

```
>>> url = 'http://www.jianshu.com'
>>> request.urlopen(url)
... ..
urllib.error.HTTPError: HTTP Error 403: Forbidden

# 设定请求头的User-Agent为Firefox
>>> headers = {"User-Agent": "Mozilla/5.0 (X11; Linux x86_64; rv:52.0) Gecko/20100101 Firefox/52.0"}
# 构建一个请求对象
>>> r = request.Request(url, headers=headers)
>>> html = request.urlopen(r)
>>> html.read() # OK
```

数据编码：

- url只允许一部分字符，其他字符必须进行编码

```
>>> url = 'https://www.sogou.com/web?query=春节'
>>> html = request.urlopen(url)
... ..
UnicodeEncodeError: 'ascii' codec can't encode characters in position 15-16: ordinal not in range(128)

>>> request.quote('元旦')
'%E5%85%83%E6%97%A6'
>>> url = 'https://www.sogou.com/web?query=' + request.quote('元旦')
>>> url
'https://www.sogou.com/web?query=%E5%85%83%E6%97%A6'
>>> html = request.urlopen(url)
```

wget模块

- 底层采用urllib，实现下载功能

```
(nsd1908) [root@room8pc16 day01]# pip install wget
>>> import wget
>>>
wget.download('https://img01.sogoucdn.com/app/a/100520021/3d14b6bc9f25f642eb896d30a3d8e413', '/tmp/aa.jpg')
```

paramiko模块

- 实现ssh客户端功能

```
(nsd1908) [root@room8pc16 day01]# pip install /var/ftp/pub/zzg_pypkgs/paramiko_pkgs/*

>>> import paramiko
>>> ssh = paramiko.SSHClient() # 创建sshclient客户端
```

```
# 自动接受服务器发来的证书
>>> ssh.set_missing_host_key_policy(paramiko.AutoAddPolicy())
# 登陆
>>> ssh.connect('192.168.4.4', username='root', password='123456')
# 执行命令后，返回值是由3个类文件对象构成的元组：（输入，输出，错误）
>>> result = ssh.exec_command('id root; id zhangsan')
>>> len(result)
3
# 也可以把执行任务后的返回数据分别赋值
>>> stdin, stdout, stderr = ssh.exec_command('id root; id zhangsan')
>>> out = stdout.read()
>>> err = stderr.read()
>>> out
b'uid=0(root) gid=0(root) \xe7\xbb\x84=0(root)\n'
>>> err
b'id: zhangsan: no such user\n'
>>> out.decode() # bytes类型转str类型
'uid=0(root) gid=0(root) 组=0(root)\n'
# 关闭连接
>>> ssh.close()
```