# 编程作业2设计文档

## 作者：1853931-郑涵文

### 01 开发环境

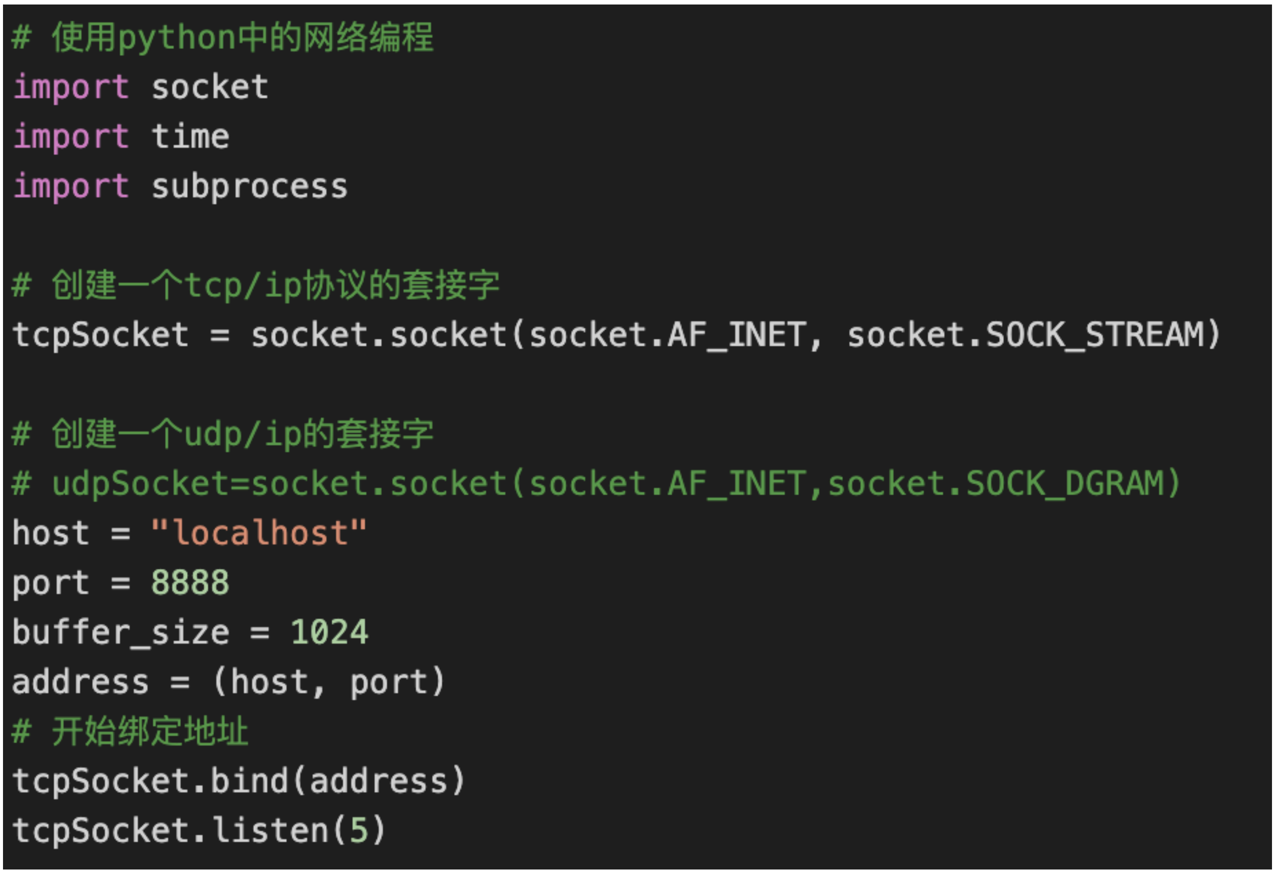
本地环境：VMware Fusion 专业版 11.5.6 + Ubuntu 64位 20.04 + Python 3.7

### 02 设计细节

**网络通信设计：**

python socket编程

client:



```python
# 使用python中的网络编程
import socket
import time
import subprocess

# 创建一个tcp/ip协议的套接字
tcpSocket = socket.socket(socket.AF_INET, socket.SOCK_STREAM)

# 创建一个udp/ip的套接字
# udpSocket=socket.socket(socket.AF_INET,socket.SOCK_DGRAM)
host = "localhost"
port = 8888
buffer_size = 1024
address = (host, port)
# 开始绑定地址
tcpSocket.bind(address)
tcpSocket.listen(5)
```

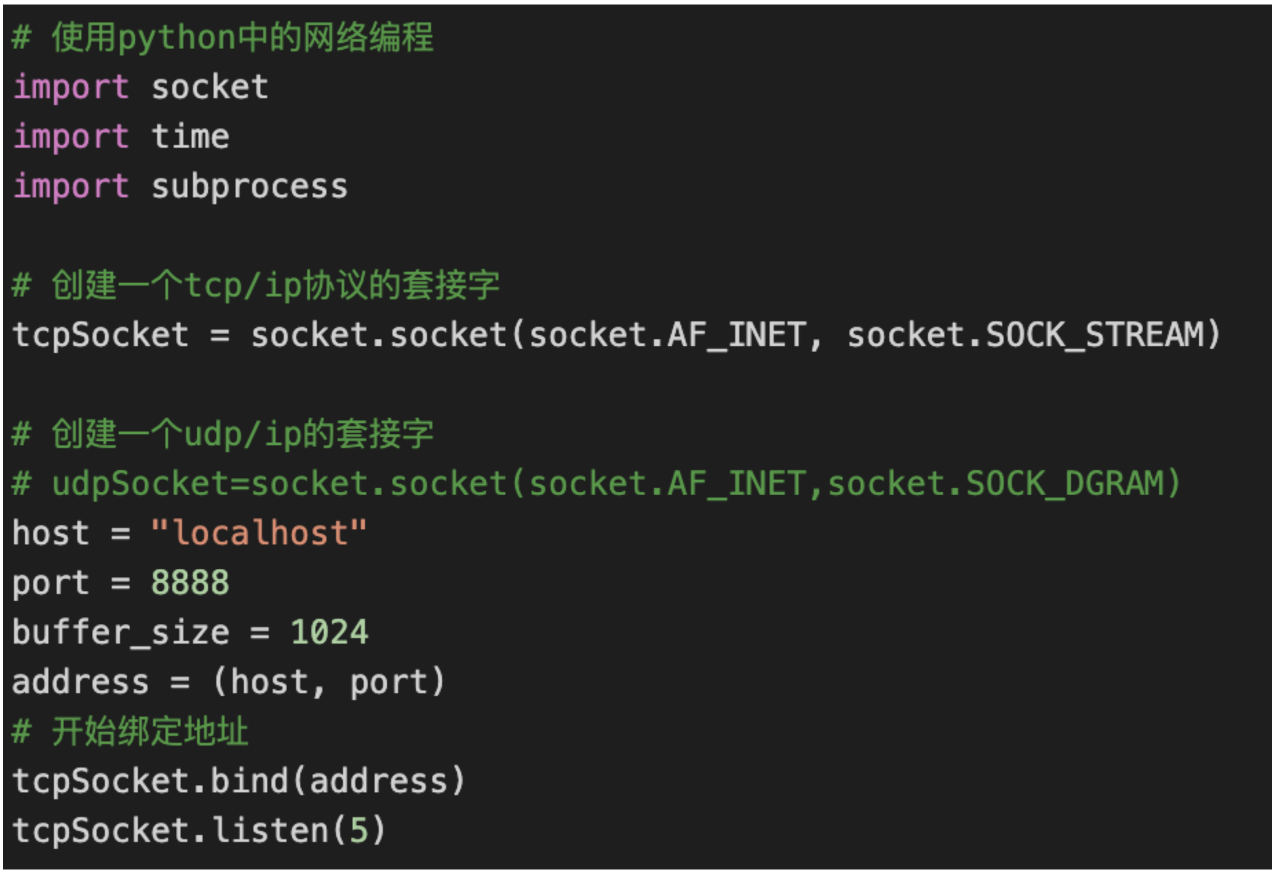
# 发送数据到服务器

clientSocket1.send(bytes(data, encoding="utf8"))

# 接收并处理服务器发送的数据

```
        data = clientSocket.recv(buffer_size)
```

server:



```python
# 使用python中的网络编程
import socket
import time
import subprocess

# 创建一个tcp/ip协议的套接字
tcpSocket = socket.socket(socket.AF_INET, socket.SOCK_STREAM)

# 创建一个udp/ip的套接字
# udpSocket=socket.socket(socket.AF_INET,socket.SOCK_DGRAM)
host = "localhost"
port = 8888
buffer_size = 1024
address = (host, port)
# 开始绑定地址
tcpSocket.bind(address)
tcpSocket.listen(5)
```

```
        # 服务器连接

        clientSock, client_addr = tcpSocket.accept()

        # 接受客户端数据

        data = clientSock.recv(buffer_size)

        #关闭套接字

        tcpSocket.close()
```

**存储负载平衡设计：**

```
# 统计dblp.xml行数

wc -l dblp.xml

# 按行数分割dblp.xml->dblp_0.xml & dblp_1.xml

split -l "xx" dblp.xml -d -a 1 dblp_
```

**查询容错：**

正常情况下server1处理dblp_0.xml server2处理dblp_1.xml

```
    # server1异常 把"wrong"消息返回client
    clientSock.send("wrong")
    # client若接收到"wrong" 把"Socket1 wrong"消息发送给server2
    clientSocket2.send(bytes("Socket1 wrong", encoding="utf8"))
    # server2若接收到"Socket1 wrong" 则开始处理dblp_0.xml
    output = subprocess.check_output('grep -o %s dblp_1.xml | wc -l' % newdata, shell=True)
    # server2将dblp_0.xml处理结果返回
    clientSock.send(output)
    # server2处理server1异常后 继续处理dblp_1.xml
    output = subprocess.check_output('grep -o %s dblp_0.xml | wc -l' % newdata, shell=True)
    # server2返回dblp_1.xml的处理结果
    clientSock.send(output)

    # server2异常同理
```