

ASSIGNMENT REPORT 1: PROCESS AND THREAD IMPLEMENTATION

CENG2034, OPERATING SYSTEMS

Söz Tuana KURŞUN
soztuanakursun@mu.edu.tr

Monday 4th May, 2020

Abstract

In this study, we have seen how operating systems manage processes or memory and we learned that every process is actually pid. We learned that every folder or processor is a file. Linux commands and what can be done with these commands, linux is written in c language and we use os library to use it as native in python and we learned that we can manage with the data structures of python. On the other hand, we learned multithreading. It is the ability of a program to perform more than one job at the same time. We have learned that multithreading is more useful on a time basis if there is file reading operation. We learned that all threads started in the process share the same memory. Each new process runs independently from other processes, we learned that it is not shared between memory processes and that each process uses separate memory addresses.

1 Introduction

The purpose of this laboratory is to see how operating systems manage processes and memories and in the language of python is to see the differences in multiprocessing and multithread. In this lab, we learned to learn the differences and aim to create better quality software.

2 Assignments

In this experiment, although we learned some of the linux commands, we also used them. The following commands are what we use in the experiment.

2.1 Assignment `print("PID of this process: ", pid)`

We learned how a processor's pid address works.

2.2 Assignment `print("loadavg: ", os.getloadavg())`

With this command, we learned when our computer turned on and off, how long it turned on.

2.3 Assignment `print("CPU core count: ", cpucount)`

We learned how many CPUs there are.

2.4 Assignment print ("multithreading")

Using multithreading we have seen whether it saves time.

3 Results

-As a result, we learned that every file, every folder can have a pid and all of them are different.

-We learned that if our operating system is linux, we can easily find loadavg. If not, we learned that we can do a few operations and learn.

-We learned that the values in loadavg are the output of the data in the 1st minute, 5th minute and 15th minute. We also learned how many cores our computer is.

-We learned multiprocessing and multithreading in python. We learned the differences and convenience of these two.

```
1 #!/usr/bin/python3
2 import os, threading, requests
3
4 '''q1'''
5 print("pid:", os.getpid())
6
```

Figure 1: This is a pid of run. With this code, we can learn the pid of each processor.

```
6
7 '''q2'''
8 if(os.name == "posix"):
9     print("loadavg:", os.getloadavg())
10
```

Figure 2: With this command, we learned when our computer turned on and off, how long it turned on.

```
11 '''q3'''
12 load1, load5, load15 = os.getloadavg()
13 cpu_core = os.cpu_count()
14 print("5 min loadavg:", load5)
15 print("cpu:", cpu_core)
16 if (cpu_core - load5 < 1):
```

Figure 3: We learn how many cores our computer is. We have seen that we can write with nproc. If your processor is 4-core and your incoming transactions are more than 4, we learned that the transactions made should be in line and wait.

```

17
18
19 def request(url_array):
20     response = requests.get(url_array)
21     if(response.status_code==200):
22         print("valid")
23     elif(response.status_code==404):
24         print("invalid")
25
26
27 thread1=threading.Thread(target=request, args=('https://api.github.com',))
28 thread2=threading.Thread(target=request, args=('http://bilgisayar.mu.edu.tr/',))
29 thread3=threading.Thread(target=request, args=('https://www.python.org/',))
30 thread4=threading.Thread(target=request, args=['http://akrepnalan.com/-ceng2034',])
31 thread5=threading.Thread(target=request, args=('https://github.com/caesarsalad/wow',))
32
33 thread1.start()
34 thread2.start()
35 thread3.start()
36 thread4.start()
37 thread5.start()

```

Figure 4: Case study with multithreading.

```

3 import os,sys
4 import multiprocessing
5
6 url_array= ["https://api.github.com", "http://bilgisayar.mu.edu.tr/",
7            "https://www.python.org/", "http://akrepnalan.com/ceng2034",
8            "https://github.com/caesarsalad/wow"]
9
10
11 for i in url_array:
12     checking=requests.get(i)
13     print(checking.status_code)

```

Figure 5: Case study with multiprocessing.

4 Conclusion

As a result, at the end of this experiment, as I mentioned at all stages, we learned that all processes are files and they all have a special pid. Actually, there are 3 important issues we learned in this lab. Firstly, contrary to the fact that os lesson is seen more theoretically, we have seen that we can make better software by knowing the operating system more closely with our lab. Another issue is CPU. If you have a cpu heavy task, and you want to make it faster use multiprocessing! For example, if you have 4 cores as in the examples we did, each multi-threaded core will have a capacity of approximately percent of 25 whereas with multi-processing you get percent of 100 in each core. This means that you will gain a speed of 4 in percent of 100 of 4 cores. Another topics we learned there can only be one thread running at any given time in a python process, multiprocessing is parallelism; multithreading is concurrency. Multiprocessing is for increasing speed and Multithreading is for hiding latency. Multiprocessing is best for computations. Multithreading is best for IO.

Multiprocessing vs Threading Python

▲ I am trying to understand the advantages of `multiprocessing` over `threading`. I know that **`multiprocessing`** gets around the Global Interpreter Lock, but what other advantages are there, and can **`threading`** not do the same thing?

600

▼ python multithreading multiprocessing

top answer ↓

▲ The `threading` module uses threads, the `multiprocessing` module uses processes. The difference is that threads run in the same memory space, while processes have separate memory. This makes it a bit harder to share objects between processes with multiprocessing. Since threads use the same memory, precautions have to be taken or two threads will write to the same memory at the same time. This is what the global interpreter lock is for.

524

▼
✓ Spawning processes is a bit slower than spawning threads. Once they are running, there is not much difference. 😞 Not quite true...

Figure 6: Other issues we learned multiprocessing and multithreading in python. In the picture above, we have seen a brief definition of the two terms.

```
*Adsız - Not Defteri
Dosya Düzen Biçim Görünüm Yardım
from concurrent.futures import ProcessPoolExecutor, ThreadPoolExecutor
def multithreading(func, args, workers):
    with ThreadPoolExecutor(workers) as ex:
        res = ex.map(func, args)
    return list(res)
def multiprocessing(func, args, workers):
    with ProcessPoolExecutor(workers) as ex:
        res = ex.map(func, args)
    return list(res)
```

Figure 7: Now we will see the difference between the two with a code.

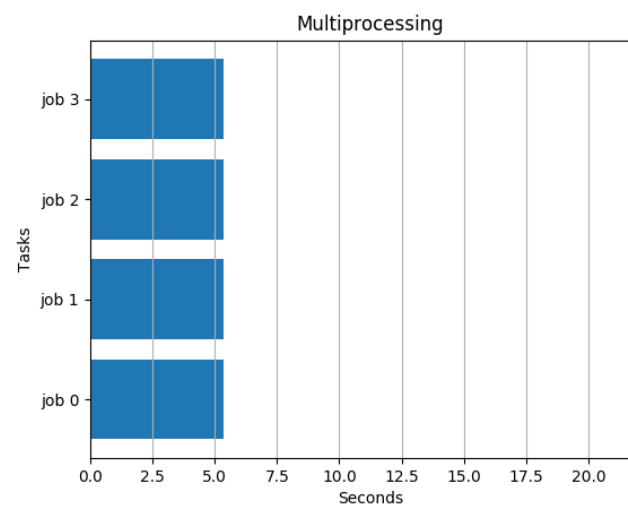
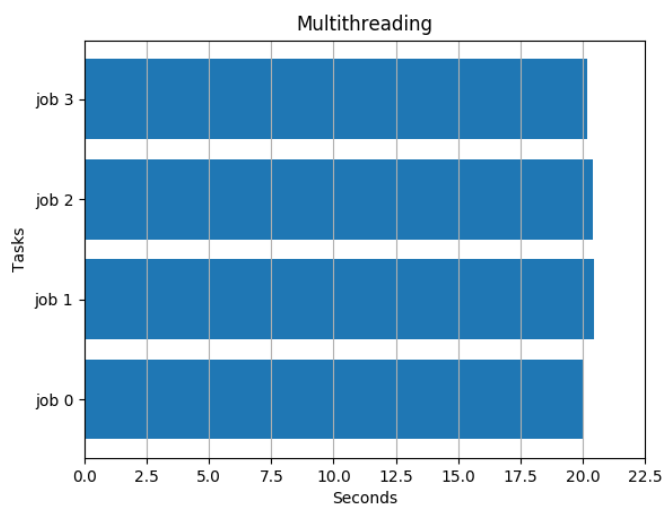


Figure 8: While Multithreading took 20 seconds, Multiprocessing took only 5 seconds. So now that we are convinced that they're not the same, we would like to know why. We have seen and understood the difference from the pictures above.