

ASSIGNMENT REPORT 2: MULTIPROCESS IMPLEMENTATION

CENG2034, OPERATING SYSTEMS

Zeki Sağlam
zzekisaglam@gmail.com

Tuesday 9th June, 2020

Abstract

The purpose of this assignment is to understand child and parent relationships and use multiprocessing. method.I downloaded some files with the function using child process from a given Array list.

1 Introduction

This Assessment I compared many types of processes.I learned how to use it.

2 Assignments

I started this task by taking the necessary libraries.

```
import os
import requests
import sys
from multiprocessing import Pool
from multiprocessing import Process
import glob
import shutil
import urllib.request
import hashlib
```

2.1. For Print parent id

Print parent id to screen.Create child process using fork.If pid of process is greater than 0 that means it is parent process.os.wait() method is used by a process to wait for completion of a child process. If pid of process is equal to 0 that means it is child process.Print child id to screen.

```
print("The PID of the main process: ", os.getpid())
pid = os.fork() #

if (pid > 0): #
    os.wait() #

if (pid == 0): #
    print("The PID of the child process: ", os.getpid()) |
```

2.2 With the child process, downloading the files via the given Array list

Requests.get method makes a request to a url, and return the status code. If status code equal to 200 that means url is working.

Create a file named file_name. Print information and file name after that Print if url was not downloaded, and check for duplicates. The elements registered in the a list were checked with the multiprocessing technique and the duplicate

```
array = [
    "http://wiki.netsec.lab.mu.edu.tr/images/thumb/f/f7/MSKU-BlockchainResearchGroup.jpeg/300px-MSKU-BlockchainResearchGroup.jpeg",
    "https://upload.wikimedia.org/wikipedia/tr/9/98/Mu%C4%9Fla_S%C4%B1tk%C4%B1_Ko%C3%A7man_%C3%9Cniversitesi_Logo.png",
    "https://upload.wikimedia.org/wikipedia/commons/thumb/c/c3/Hawaii%27i.jpg/1024px-Hawaii%27i.jpg",
    "http://wiki.netsec.lab.mu.edu.tr/images/thumb/f/f7/MSKU-BlockchainResearchGroup.jpeg/300px-MSKU-BlockchainResearchGroup.jpeg",
    "https://upload.wikimedia.org/wikipedia/commons/thumb/c/c3/Hawaii%27i.jpg/1024px-Hawaii%27i.jpg"]

file_name = ["1.url", "2.url", "3.url", "4.url", "5.url"]

def download_file(url, file_name):

    r = requests.get(url, stream=True)
    if r.status_code == 200:
        r.raw.decode_content = True

        with open(file_name, 'wb') as f:
            shutil.copyfileobj(r.raw,
                               f)

        print("File Downloaded and checked", file_name)
    else:
        print("File can not Downloaded Error !")
        p = multiprocessing.Pool(processes=6)
        file_hashes_list = p.map(hashFile, files)

    for file_hash in file_hashes_list:
        pushDictAndCheckDuplicates(file_hash)
```

2.3 Avoiding the orphan process situation

This time the problem was avoiding the orphan process situation. I added wait() for this, after the child but before the parent because I wanted my parent process to wait child to complete. We created child processes and with "os.wait()" methods , avoid from orphan processes.

3 Results

I got this output showing parent and child pids. It shows a list of downloaded files, the pids of

```
The PID of the main process: 6042
The PID of the child process: 6043
File Downloaded and checked 1.url
File Downloaded and checked 2.url
File Downloaded and checked 3.url
File Downloaded and checked 4.url
File Downloaded and checked 5.url
```

the processes in the functions, duplicate files list and taken time.

4 Conclusion

I learned a lot with this task. One of the first things I learned; in multiprocessing, multiple processes or jobs can be run and managed by the CPU or a single program. This Assignment was very instructive and made me understand how processes work.

5 My github

github.com/zekisaglam/