### ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN KHOA CÔNG NGHỆ THÔNG TIN

  



**BÁO CÁO ĐỒ ÁN**

# SOCKET

Môn học: Mạng máy tính

### GIÁO VIÊN HƯỚNG DẪN

**Giáo viên lý thuyết**:

Huỳnh Thụy Bảo Trân

**Giáo viên thực hành**:

Chung Thùy Linh

### ĐẠI HỌC QUỐC GIA THÀNH PHỐ HỒ CHÍ MINH TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN KHOA CÔNG NGHỆ THÔNG TIN

  



**BÁO CÁO ĐỒ ÁN**

# SOCKET

Môn học: Mạng máy tính

Đề 2: Web Client

### THÀNH VIÊN NHÓM

21120313 – Trần Nam Phương

21120317 – Nguyễn Phan Anh Quốc

21120321 – Nguyễn Văn Siêu

## MỤC LỤC

[MỤC LỤC 3](#_bookmark0)

[THÔNG TIN THÀNH VIÊN 4](#_bookmark2)

[ĐÁNH GIÁ MỨC ĐỘ HOÀN THÀNH 4](#_bookmark3)

[KỊCH BẢN GIAO TIẾP CỦA CHƯƠNG TRÌNH 5](#_bookmark4)

1. [Giao thức trao đổi: 5](#_bookmark5)
2. [Cấu trúc thông điệp: 5](#_bookmark6)
3. [Kiểu dữ liệu thông điệp: 5](#_bookmark7)

[MÔI TRƯỜNG LẬP TRÌNH 6](#_bookmark10)

[HƯỚNG DẪN SỬ DỤNG CHƯƠNG TRÌNH 7](#_bookmark11)

[BẢNG PHÂN CÔNG CÔNG VIỆC 9](#_bookmark26)

[NGUỒN THAM KHẢO 10](#_bookmark27)

## THÔNG TIN THÀNH VIÊN

|  |  |  |
| --- | --- | --- |
| **Mã số sinh viên** | **Họ và tên** | **Chú thích** |
| 21120313 | Trần Nam Phương | [21120313@student.hcmus.edu.vn](mailto:21120313@student.hcmus.edu.vn) |
| 21120317 | Nguyễn Phan Anh Quốc | [21120317@student.hcmus.edu.vn](mailto:21120317@student.hcmus.edu.vn) |
| 21120321 | Nguyễn Văn Siêu | [21120321@student.hcmus.edu.vn](mailto:21120321@student.hcmus.edu.vn) |

## ĐÁNH GIÁ MỨC ĐỘ HOÀN THÀNH

|  |  |  |  |
| --- | --- | --- | --- |
| STT | Chức năng | Đánh giá | Chú thích |
| 1 | Kết Nối | 100% |  |
| 2 | Quản lý kết nối | 100% |  |
| 3 | Tải và lưu thành file dạng “Content-length” | 100% |  |
| 4 | Tải và lưu thành file dạng “Transfer-encoding: chunked” | 50% | Chưa xử lý được việc web google trả về chunk size lớn nên lưu dữ liệu còn thiếu |
| 5 | Tải các file trong folder | 100% |  |
| 6 | Multiple request | 0% | Chưa xử lý được việc server đóng kết nối sau khi trả về response thứ nhất |
| 7 | Multiple connection | 100% |  |

## KỊCH BẢN GIAO TIẾP CỦA CHƯƠNG TRÌNH

#### Giao thức trao đổi:

Giao thức trao đổi giữa Client và Server được sử dụng trong đồ án là TCP.

#### Cấu trúc thông điệp:

Khi Client muốn truy suất thông tin của Server thì Client sẽ gửi cho Sever thông điệp có cấu trúc như sau:

GET **resource** HTTP/1.1\r\n

Host: **host**\r\n

Connection: Keep-Alive\r\n

\r\n

Mỗi dòng sẽ kết thúc với ‘\r\n’ và toàn bộ thông điệp cũng sẽ kết thúc với ‘\r\n’. Trong đó, **resource** là tài nguyên mà client cần truy cập tới trong server, **host** là trang chủ của server đó. Sau khi nhận được thông điệp từ client, server sẽ phản hồi dữ liệu lại cho client.

#### Kiểu dữ liệu thông điệp:

Dữ liệu được gửi đến và dữ liệu được nhận vào đều là kiểu chuỗi ký tự (string).

## MÔI TRƯỜNG LẬP TRÌNH

* Ứng dụng viết trên hệ điều hành **Windows 11.**
* Ngôn ngữ lập trình **Python** (Phiên bản 3.9.7 64-bit).
* Các thư viện hỗ trợ: **socket, threading, re, os, time.**

## HƯỚNG DẪN SỬ DỤNG CHƯƠNG TRÌNH

Khi bắt đầu chạy file client.py thì chương trình sẽ yêu cầu nhập vào link url của các trang web cần tải.

Ảnh có chứa văn bản, màn hình, ảnh chụp màn hình, máy tính

Mô tả được tạo tự động

Sau khi nhập thì Client tiến hành kết nối tới server và gửi request.

Ảnh có chứa văn bản, màn hình, ảnh chụp màn hình, máy tính

Mô tả được tạo tự động

Sau khi tải xong sẽ in địa chỉ của file đã tải .

Ảnh có chứa văn bản, ảnh chụp màn hình, màn hình, máy tính

Mô tả được tạo tự động

Nếu nhập nhiều url thì các url cách nhau bằng khoảng trắng

Ảnh có chứa văn bản, ảnh chụp màn hình, màn hình, máy tính

Mô tả được tạo tự động

**BẢNG PHÂN CÔNG CÔNG VIỆC**

|  |  |
| --- | --- |
| Họ và tên – MSSV | Công việc |
| Trần Nam Phương  21120313 | Gửi request tới server |
| Xử lý header response |
| Xử lý body response và lưu về dạng “Content-length” |
| Báo cáo đồ án |
| Nguyễn Phan Anh Quốc  21120317 | Tạo và quản lý kết nối đến server |
| Phân tích url để lấy địa chỉ host và tên file tải về |
| Thực hiện multiple request |
| Xử lí đa luồng để gửi nhiều kết nối đến server(multiple connection) |
| Nguyễn Văn Siêu  21120321 | Xử lý body response và lưu về dạng “Transfer-encoding: chunked” |
| Xử lý dữ liệu html để lấy url của các file cần download trong folder |
| Xử lý đa luồng để tải tất cả file trong folder |

## NGUỒN THAM KHẢO

* + Đa luồng – Thread

<https://www.geeksforgeeks.org/multithreading-python-set-1/amp/?fbclid=IwAR0iYlwmMhxbbM7gA-LAoMKMygwQOES6m9PlJh9UYU5nee_vv0nbvuDlRHM>

* + HTTP Header Response

<https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers>

* + Write file in Python and save in computer

<https://www.w3schools.com/python/python_file_write.asp>

<https://www.geeksforgeeks.org/python-os-path-join-method/>

* + Transfer-encoding: chunked

<https://bunny.net/academy/http/what-is-chunked-encoding/>

* + Thư viện re

<https://www.w3schools.com/python/python_regex.asp?fbclid=IwAR2RgTZ8mOm77SIUAIay6EmXaHvzpWv0ImwIQr8oeWHPPBU8ZVnzD0qCgXk>

<https://docs.python.org/3/library/re.html?fbclid=IwAR2q9A9bZzE4W5XHAbzrjvtc8ewHITqzDqt_Pre9o-2Sym495odn24Sdu64>

* + Và các tài liệu hướng dẫn thực hành môn học đã cung cấp.