



fit@hcmus

VNUHCM - UNIVERSITY OF SCIENCE
FACULTY OF INFORMATION TECHNOLOGY

BÁO CÁO ĐỒ ÁN
MẠNG MÁY TÍNH - 22CTT5

ĐỒ ÁN THỰC HÀNH 1
LẬP TRÌNH SOCKET

Giảng viên hướng dẫn

Ths. LÊ HÀ MINH

Thành viên nhóm

Võ Quốc Quang - 22120299

Hồ Quang Sang - 22120309

Lê Hoàng Sơn - 22120311

Mục lục

1	Thông tin của nhóm	2
2	Đánh giá mức độ hoàn thành	2
3	Kịch bản giao tiếp của chương trình	3
3.1	Giao thức trao đổi giữa client và server	3
3.1.1	Giao thức SMTP	3
3.1.2	Giao thức POP3	3
3.2	Cấu trúc thông điệp gửi lên máy chủ SMTP	4
3.2.1	Mail không đính kèm file	4
3.2.2	Mail đính kèm file	5
3.3	Cấu trúc thông điệp gửi lên máy chủ POP3	6
3.4	Kiểu dữ liệu của thông điệp	6
4	Cách tổ chức cơ sở dữ liệu	6
5	Môi trường lập trình và framework	7
6	Hướng dẫn sử dụng các tính năng chương trình	7
6.1	Khởi động Mail Server và Mail Client	7
6.2	Gửi Email	8
6.3	Đọc Email và tải tệp đính kèm	8
6.4	Những tính năng khác	8
6.4.1	Tải email tự động	8
6.4.2	Lọc email và chuyển vào thư mục	9
6.4.3	Cập nhật trạng thái đã đọc/chưa đọc của email	9
7	Tài liệu tham khảo	9

1 Thông tin của nhóm

STT	Họ và tên	MSSV	Phân công công việc
1	Võ Quốc Quang	22120299	Gửi file đính kèm email Tổ chức cơ sở dữ liệu Tải email về máy từ mail server Xử lý tự động tải Email theo cấu hình file config
2	Hồ Quang Sang	22120309	Gửi Email với TO, CC và BCC Bài báo cáo
3	Lê Hoàng Sơn	22120311	Quản lý trạng thái Email Xử lý lọc Email

2 Đánh giá mức độ hoàn thành

Về tổng quan, đề án đã được hoàn thành ở mức **95%** các yêu cầu được đưa ra. Mức độ hoàn thành đề án như sau:

- Gửi được email với To, CC, BCC có thể đính kèm tệp (dung lượng dưới 3MB)
- Tổ chức file config để lưu trữ thông tin về người dùng, trạng thái các email và thời gian cập nhật dữ liệu về máy chủ
- Quản lý trạng thái email.
- Xử lý lọc email vào các folder cụ thể (phân loại theo địa chỉ người gửi, theo subject, hoặc từ khóa trong nội dung email).
- Tổ chức file config có cấu trúc (.json)
- Tự động tải đủ và đúng các email từ mailbox của Server theo cấu hình file config.

Có một vấn đề khi gửi email đính kèm file, việc giới hạn dung lượng file gửi dưới **3 MB** sẽ làm những email **chỉ gửi 1 file** vượt mức đó không được gửi đi và người dùng phải quay lại gửi 1 email khác chứ không thể thay đổi file gửi trong tiến trình hiện tại (nghĩa là khi phát hiện file quá dung lượng thì dừng tiến trình gửi email hiện tại lại).

Chương trình mail client này chỉ tải được những email có dữ liệu người nhận trong file config. Nếu có 1 chương trình mail client khác (ví dụ như Thunderbird) gửi lên mail server thì chương trình sẽ không có dữ liệu người nhận (trong trường hợp dữ liệu người nhận chưa có sẵn) và không thể sử dụng thực hiện các chức năng vốn có.

3 Kịch bản giao tiếp của chương trình

3.1 Giao thức trao đổi giữa client và server

- Kịch bản giao tiếp giữa máy chủ mail client và mail server là sử dụng giao thức SMTP (Simple Mail Transfer Protocol) để gửi email và giao thức POP3 (Post Office Protocol version 3) để lấy email từ mailbox.
- SMTP, viết tắt của Simple Mail Transfer Protocol, là một giao thức chuẩn được sử dụng để truyền tải thư điện tử (email) trên mạng Internet. SMTP được dùng chủ yếu để gửi tin nhắn từ máy khách (client) đến máy chủ thư (mail server) hoặc từ máy chủ này đến máy chủ khác. Đây là một phần của bộ giao thức TCP/IP, nằm ở tầng ứng dụng.
- POP3, viết tắt của Post Office Protocol version 3, là một giao thức chuẩn được sử dụng để nhận email từ một máy chủ thư qua Internet. Nó hoạt động dựa trên mô hình client-server và cho phép các email clients như Outlook, Thunderbird, hoặc các ứng dụng di động tải email xuống từ máy chủ của họ để xử lý offline.

3.1.1 Giao thức SMTP

- Mail client kết nối đến server SMTP sử dụng cổng mặc định 465 hoặc một cổng khác đã được cấu hình đặc biệt cho mục đích này.
- Mail client gửi lệnh "EHLO" (hoặc "HELO") đến server để bắt đầu phiên giao tiếp và thông báo cho server về danh sách các lệnh mà client hỗ trợ.
- Lệnh "MAIL FROM" được gửi bởi mail client để thông báo cho server về địa chỉ email của người gửi.
- Lệnh "RCPT TO" được sử dụng để xác định người nhận của email.
- Lệnh "DATA" được mail client gửi đến server để bắt đầu quá trình gửi nội dung của email.
- Mail client sau đó gửi tiêu đề, nội dung của thư, và thông tin của các tệp đính kèm (nếu có) tuân theo định dạng quy định bởi SMTP (thường được mã hóa dưới dạng Base64 hoặc một dạng mã hóa tương tự để chúng có thể được truyền qua giao thức SMTP dưới dạng văn bản).
- Cuối cùng, lệnh "QUIT" được gửi bởi mail client để kết thúc phiên làm việc và đóng kết nối với server.

3.1.2 Giao thức POP3

- Mail client kết nối đến server SMTP sử dụng cổng mặc định 1100 hoặc một cổng khác đã được cấu hình đặc biệt cho mục đích này.
- Lệnh "USER" được gửi từ mail client đến server với tên người dùng (thường là địa chỉ email) để bắt đầu quá trình xác thực.

- Mail client tiếp tục quá trình xác thực bằng cách gửi mật khẩu tương ứng với lệnh "PASS". Khi xác thực thành công, client có thể tiếp tục với các lệnh khác.
- Lệnh "UIDL" được gửi để yêu cầu danh sách mã UIDL (Unique Identifier) của từng email trong mailbox. Mỗi email sẽ có một mã UIDL riêng biệt.
- Mail client lưu mã UIDL của từng email vào máy chủ client, giúp quản lý việc tải email từ server, tránh việc tải về các email đã được tải về trước đó.
- Lệnh "STAT" được gửi để lấy thông tin về số lượng và kích thước tổng cộng của các email có trong mailbox (ở đây ta chỉ lấy số lượng email).
- Lệnh "LIST" được gửi để lấy danh sách chi tiết về các email, bao gồm số thứ tự và kích thước của từng email (ở đây ta chỉ lấy số lượng email).
- Lệnh "RETR" được sử dụng cùng số thứ tự của email để truy xuất và tải về nội dung của email đó từ server.
- Mail client nhận và lưu trữ nội dung của email đã được truy xuất vào mailbox của client.
- Lệnh "QUIT" được gửi để kết thúc phiên làm việc và đóng kết nối với server.

3.2 Cấu trúc thông điệp gửi lên máy chủ SMTP

3.2.1 Mail không đính kèm file

```
EHLO mydomain.com\r\n
MAIL FROM:<sender>\r\n
RCPT TO:<receiver>\r\n
DATA\r\n
From: sender\r\n
To: receiver\r\n
CC: receiver\r\n
Subject: subject\r\n
Content-Type: text/plain\r\n
    mail body\r\n
.\r\n
QUIT\r\n
```

3.2.2 Mail đính kèm file

```
EHLO mydomain.com\r\n
MAIL FROM:<sender>\r\n
RCPT TO:<receiver>\r\n
DATA\r\n
From: sender\r\n
To: receiver\r\n
CC: receiver\r\n
Subject: subject\r\n
Content-Type: multipart/mixed; boundary="----=_Part_Boundary_YuSato"\r\n
----=_Part_Boundary_YuSato\r\n
Content-Type: text/plain\r\n
    mail body\r\n
----=_Part_Boundary_YuSato\r\n
Content-Type: application/octet-stream; name="file name 1"\r\n
Content-Transfer-Encoding: base64\r\n
Content-Disposition: attachment; filename="file name 1"\r\n\r\n
    file content 1/r/n
----=_Part_Boundary_YuSato\r\n
Content-Type: application/octet-stream; name="file name 2"\r\n
Content-Transfer-Encoding: base64\r\n
Content-Disposition: attachment; filename="file name 2"\r\n\r\n
    file content 2/r/n
    ...

----=_Part_Boundary_YuSato\r\n
.\r\n
QUIT\r\n
```

3.3 Cấu trúc thông điệp gửi lên máy chủ POP3

```
<Chú thích: i là số thứ tự từ 0>
USER username\r\n
PASS password\r\n
UIDL\r\n
STAT\r\n
LIST\r\n
RETR i\r\n
QUIT\r\n
```

3.4 Kiểu dữ liệu của thông điệp

- Khi làm việc với máy chủ SMTP, thông điệp được trao đổi giữa client và server là chuỗi văn bản. Chuỗi này bao gồm các lệnh SMTP như HELO, EHLO, MAIL FROM, RCPT TO, DATA, và QUIT, cũng như nội dung email nếu đang ở trong quá trình gửi email. Các lệnh này được phân cách bởi kí tự nguồn cấp dòng và kí tự xuống dòng, đây là ký tự xuống dòng chuẩn cho giao thức mạng.
- Khi làm việc với máy chủ POP3 để tải email, các lệnh cũng là chuỗi văn bản và bao gồm lệnh như USER, PASS, UIDL, STAT, LIST, RETR, và QUIT. Các thông điệp từ server POP3 trả về cũng là văn bản, chứa thông tin về email và nội dung của email được yêu cầu.

4 Cách tổ chức cơ sở dữ liệu

config.json

```
"Filter": [
  "D:\\socket project\\ClientProject.py\\Mail Filter\\Inbox",
  "D:\\socket project\\ClientProject.py\\Mail Filter\\Work",
  "D:\\socket project\\ClientProject.py\\Mail Filter\\Project",
  "D:\\socket project\\ClientProject.py\\Mail Filter\\Important",
  "D:\\socket project\\ClientProject.py\\Mail Filter\\Spam"
],

"Sender" : [...],
"Receiver" : [...],
"Autoload" : 10,
"Sender_Project" : ["ahihi@testing.com","ahuu@testing.com"],
```

```
"Subject" : ["urgent","ASAP","instant","now","deadline"],
"Content" : ["report","meeting"],
"Spam_Keywords" : ["hack","crack","virus","free","voucher","discount","bonus"]
```

- Lưu thông tin người nhận vào **Receiver**.
- Lưu thông tin người gửi vào **Sender**.
- Di chuyển các email có chứa các địa chỉ trong **Sender,Project** vào thư mục **Project**.
- Di chuyển các email subject có chứa các key trong **Subject** vào thư mục **Important**.
- Di chuyển các email nội dung có chứa các key trong **Content** vào thư mục **Work**.
- Di chuyển các email mà subject hoặc nội dung chứa các key trong **SpamKeywords** vào thư mục **Spam**.
- Di chuyển các email không chứa bất kì key nào trong **SenderProject, Subject, Content, SpamKeywords** vào thư mục **Inbox**.

5 Môi trường lập trình và framework

- Ngôn ngữ lập trình được sử dụng: Python 3.10+
- Môi trường lập trình: Visual Studio Community 2022 tích hợp IDE dành cho Python
- Không sử dụng framework. Nhóm hoàn thành dự án bằng cách tự xây dựng chức năng và logic cho chương trình.

6 Hướng dẫn sử dụng các tính năng chương trình

Yêu cầu: Python phiên bản 3.10+

6.1 Khởi động Mail Server và Mail Client

- Khởi động mail server bằng cách chạy **test_mail_server.jar**.
- Khởi động mail client bằng cách chạy **Main.py**
- Trong mail server, nhập thông tin cấu hình bao gồm cổng SMTP(465) và cổng POP3(1100).

6.2 Gửi Email

- Trong giao diện mail client, chọn tùy chọn **1. Để gửi email** để bắt đầu viết email mới.
- Điền các thông tin cần thiết như địa chỉ người nhận, chủ đề, nội dung email.
- * Lưu ý: Nếu có nhiều địa chỉ người nhận thì mỗi địa chỉ phải được ngăn cách bởi dấu ','.
- Nếu có tệp tin đính kèm, chọn tùy chọn **1. Có** và nhập số lượng tệp, sau đó nhập đường dẫn tệp tin trên máy tính của bạn.
- Sau khi hoàn thành điền thông tin, màn hình sẽ hiển thị **Gửi email thành công**.
- * Lưu ý: Những tệp tin vượt quá 3 MB sẽ không được gửi nếu gửi nhiều tệp tin cùng lúc và nếu chỉ gửi 1 tệp tin mà dung lượng vượt quá 3 MB thì tiến trình gửi mail sẽ tự động bị ngắt. Người dùng phải gửi mail lại từ đầu

6.3 Đọc Email và tải tệp đính kèm

- Trong giao diện mail client, chọn tùy chọn **2. Để xem danh sách các email đã nhận** để xem danh sách các thư mục chứa email đã nhận.
- Chọn thư mục bạn muốn đọc từ danh sách.
- Chọn email bạn muốn đọc trong thư mục.
- Màn hình hiển thị email với thông tin chi tiết như người gửi, tiêu đề, nội dung, và danh sách tệp tin đính kèm (nếu có).
- Bạn có thể lựa chọn tải về toàn bộ tệp tin đính kèm, sau đó nhập đường dẫn lưu trên máy tính của bạn.

6.4 Những tính năng khác

6.4.1 Tải email tự động

- Ngay sau khi mail client được khởi động, chức năng tự động tải email mỗi 10 giây từ mailbox của Server về cơ sở dữ liệu sẽ được thực thi.
- Chương trình sẽ duyệt qua những username có thông tin trong phần **Receiver** của cơ sở dữ liệu để tự động cập nhật những email mới trong mailbox (chỉ cập nhật email từ những username có dữ liệu trong cơ sở dữ liệu).
- Lưu trữ thông tin email vào cơ sở dữ liệu: **Inbox, Project, Important, Work, Spam** dưới những dạng:
(chưa đọc)<sender><subject><attachment> cho email có đính kèm nhưng chưa được đọc
(chưa đọc)<sender><subject><no attachment> cho email không đính kèm nhưng chưa được đọc

- Khi người dùng đọc mail thì sẽ trạng thái sẽ được cập nhật như sau:
`<sender><subject><attachment>` cho email có đính kèm
`<sender><subject><no attachment>` cho email không đính kèm

6.4.2 Lọc email và chuyển vào thư mục

- Xác định các quy tắc lọc email dựa trên cấu hình file config như dựa trên địa chỉ người gửi, chủ đề, nội dung, hoặc các tiêu chí khác.
- Di chuyển các email có chứa các địa chỉ trong **Sender,Project** vào thư mục **Project**.
- Di chuyển các email subject có chứa các key trong **Subject** vào thư mục **Important**.
- Di chuyển các email nội dung có chứa các key trong **Content** vào thư mục **Work**.
- Di chuyển các email mà subject hoặc nội dung chứa các key trong **SpamKeywords** vào thư mục **Spam**.
- Di chuyển các email không chứa bất kì key nào trong **SenderProject, Subject, Content, SpamKeywords** vào thư mục **Inbox**.

6.4.3 Cập nhật trạng thái đã đọc/chưa đọc của email

- Thông tin email được lưu trữ trong cơ sở dữ liệu: **Inbox, Project,Important, Work, Spam** dưới những dạng:
`(chưa đọc)<sender><subject><attachment>` cho email có đính kèm nhưng chưa được đọc
`(chưa đọc)<sender><subject><no attachment>` cho email không đính kèm nhưng chưa được đọc
- Khi người dùng đọc mail thì sẽ trạng thái sẽ được cập nhật như sau:
`<sender><subject><attachment>` cho email có đính kèm
`<sender><subject><no attachment>` cho email không đính kèm

7 Tài liệu tham khảo

[Simple Mail Transfer Protocol \(SMTP\)](#)

[What is POP3 \(Post Office Protocol Version 3\)?](#)

[How to Utilize POP3 Protocol to Access DataMotion and Direct Accounts](#)

[The TCP/IP Guide - SMTP Commands How to Perform Threading Timer in Python](#)

[Working with config files in Python](#)

[ChatGPT](#)