Klient IMAP s podporou TLS

Dokumentace k projektu

Autor: Slabik Yaroslav

Login: xslabi01

Obsah

---

1. Úvod

Tato dokumentace popisuje IMAP klienta s názvem **imapcl**, vytvořeného v rámci projektu předmětu ISA na Fakultě informačních technologií VUT v Brně. Cílem projektu bylo implementovat klienta pro protokol IMAP4rev1 (RFC 3501), který umožňuje stahování elektronické pošty ze zadaného serveru a ukládání zpráv do specifikovaného adresáře. Klient podporuje bezpečné spojení pomocí SSL/TLS, autentizaci uživatele, výběr různých schránek a možnost stahování pouze nových zpráv či pouze jejich hlaviček.

2. Popis zadání

Program **imapcl** má následující funkce:

- **Připojení k IMAP serveru:** Připojuje se k zadanému IMAP serveru pomocí protokolu IMAP4rev1.
- **Autentizace uživatele:** Provádí přihlášení uživatele pomocí poskytnutých přihlašovacích údajů.

- **Stahování zpráv:** Stahuje zprávy ze zadané schránky (výchozí je "INBOX") a ukládá je do specifikovaného adresáře.
- **Výpis počtu zpráv:** Vypisuje informaci o počtu stažených zpráv na standardní výstup.
- **Podpora parametrů:** Umožňuje nastavení dodatečných parametrů pro změnu funkcionality, jako je použití TLS, výběr schránky, výběr portu, stahování pouze nových zpráv či pouze hlaviček zpráv.

3. Návrh aplikace

3.1 Architektura programu

Program byl navržen modulárně s důrazem na přehlednost a snadnou údržbu. Hlavní komponenty programu:

- **Parsers argumentů příkazové řádky**: Zpracovává vstupní parametry a nastavuje chování programu.
- **Síťový modul**: Zajišťuje navázání spojení se serverem a komunikaci pomocí IMAP protokolu.
- **Autentizační modul**: Provádí přihlášení uživatele k serveru.
- **Modul zpracování zpráv**: Stará se o stahování, zpracování a ukládání zpráv.
- **Modul pro práci s lokálním indexem zpráv:** Zodpovědný za čtení, aktualizaci a správu indexu UID zpráv.
- **Modul Dekódování Zpráv:** Zpracovává enkódování zpráv, jako je Base64 a Quoted-Printable.
- **SSL/TLS modul**: Zajišťuje bezpečné spojení se serverem.

3.2 Struktura kódu

Kód je rozdělen do několika souborů a funkcí, z nichž každá plní specifickou úlohu:

- **main.cpp:** Obsahuje hlavní funkci programu, zpracování argumentů a orchestraci hlavních kroků (připojení, autentizace, stahování zpráv).
- **imap_client.cpp/imap_client.h:** Implementuje funkce pro komunikaci s IMAP serverem, zpracování zpráv a dekódování.
    - **generate_tag:** Generuje unikátní tag pro IMAP příkazy.
    - **ssl_read:** Funkce pro čtení dat.
    - **ssl_write:** Funkce pro zápis dat.
    - **connect_to_server**: Navazuje spojení se serverem, nastavuje SSL/TLS pokud je vyžadováno.
    - **send_command:** Odesílá IMAP příkaz na server a přijímá odpověď.
    - **read_line:** Čte jednu řádku z odpovědi serveru.
    - **read_literal:** Čte blok dat specifikované velikosti z serveru. Používá se k načítání obsahu zpráv nebo jiných velkých datových bloků.

    - **login:** Provádí autentizaci uživatele.
    - **select_mailbox:** Vybere specifikovanou schránku na serveru.
    - **read_local_index:** Čte lokální index UID ze souboru **index.txt**.
    - **update_local_index:** Aktualizuje lokální index UID ve souboru **index.txt**.
    - **search_unseen_messages:** Vyhledává nepřečtené zprávy ve vybrané schránce.
    - **save_message:** Ukládá jednotlivé zprávy do souborů.
    - **fetch_messages:** Stahuje zprávy a ukládá je do adresáře.
    - **read_credentials, directory_exists:** Pomocné funkce pro práci se soubory a adresáři.
    - Dekodovací funkce: **base64_decode, decode_quoted_printable, decode_encoded_word.**

- **ssl_utils.cpp/ssl_utils.h**: Implementuje funkce pro práci s SSL/TLS.
    - **initialize_ssl**: Inicializuje OpenSSL knihovnu.
    - **create_context**: Vytváří SSL kontext.
    - **configure_ssl_context:** Konfiguruje SSL kontext s certifikáty.
    - **cleanup_ssl:** Čistí SSL kontext a uvolňuje zdroje.

4. Popis implementace

4.1 Zpracování argumentů příkazové řádky

Program využívá knihovnu **getopt** pro zpracování parametrů příkazové řádky.

- **server**: Povinný argument určuje IP adresu nebo doménové jméno IMAP serveru.
- **-p port**: Specifikuje číslo portu na serveru. Výchozí hodnota je 143 pro nešifrované spojení a 993 při použití TLS.
- **-T**: Zapíná šifrování pomocí SSL/TLS.
- **-c certfile**: Určuje soubor s certifikáty pro ověření serveru.
- **-C certaddr**: Určuje adresář s certifikáty (výchozí: **/etc/ssl/certs**).
- **-n**: Stahování pouze nových zpráv.
- **-h**: Stahování pouze hlaviček zpráv.
- **-a auth_file**: Povinný parametr určující cestu k souboru s autentizačními údaji.
- **-b MAILBOX**: Určuje název schránky na serveru (výchozí: **INBOX**)
- **-o out_dir**: Povinný parametr určující výstupní adresář pro ukládání stažených zpráv.

4.2 Navázání spojení se serverem

Funkce **connect_to_server** zajišťuje navázání spojení s IMAP serverem. Postup zahrnuje:

1. **Získání Informací o Serveru:** Pomocí funkce **getaddrinfo** získá informace o serveru na základě zadané IP adresy nebo doménového jména.
2. **Vytvoření Socketu:** Vytvoří TCP socket pro komunikaci se serverem.
3. **Navázání TCP Spojení:** Připojí se k serveru na specifikovaném portu.
4. **Inicializace SSL/TLS (pokud je požadováno):**
    4.1. Vytvoří nový SSL objekt a připojí ho k socketu.
    4.2. Naváže SSL spojení pomocí **SSL_connect**.
    4.3. Ověří platnost certifikátu serveru.
    4.4. Pokud ověření selže, program vypíše chybovou zprávu a ukončí se.

4.3 Autentizace uživatele

Funkce **login** provádí autentizaci uživatele pomocí příkazu **LOGIN**. Postup zahrnuje:

1. **Přijetí Uvítací Zprávy:** Přijme uvítací zprávu od serveru po navázání spojení.
2. **Generování Unikátního Tagu:** Vytvoří unikátní tag pro IMAP příkaz.
3. **Odeslání Přihlašovacího Příkazu:** Odešle příkaz **LOGIN** spolu s uživatelským jménem a heslem.
4. **Zpracování Odpovědi:** Kontroluje, zda byla autentizace úspěšná na základě odpovědi serveru.

4.4 Výběr schránky

Funkce **select_mailbox** umožňuje výběr specifikované schránky na serveru pomocí příkazu **SELECT**. Postup zahrnuje:

1. **Odeslání Příkazu SELECT:** Vybere požadovanou schránku (např. **INBOX**, **Sent**, **Trash**).
2. **Získání Seznamu UID Zpráv:** Po úspěšném výběru schránky získá unikátní ID všech zpráv ve schránce pomocí příkazu **UID SEARCH ALL**.
3. **Uložení UID do Vektoru:** UID jsou uložena do vektoru pro následné zpracování.

4.5 Stahování a ukládání zpráv

Funkce **fetch_messages** zajišťuje stahování zpráv ze serveru a jejich ukládání do specifikovaného adresáře. Postup zahrnuje:

1. **Čtení Lokálního Indexu:** Načte lokální index stažených zpráv z **index.txt** v výstupním adresáři.
2. **Určení Zpráv k Stažení:** Na základě parametrů **-n** a **-h** určí, které zprávy je třeba stáhnout (všechny nebo pouze nové, hlavičky nebo celé zprávy).
3. **Odeslání Příkazů FETCH:** Pro každou zprávu odešle příkaz **UID FETCH** pro stažení požadovaných částí zprávy.
4. **Zpracování Odpovědí:** Přečte odpovědi serveru, dekóduje zprávy (např. Base64, Quoted-Printable) a ukládá je do souborů ve formátu RFC 5322.

5. **Aktualizace Lokálního Indexu:** Po úspěšném stažení zpráv aktualizuje **index.txt** s novými UID.

4.6 Práce se SSL/TLS certifikáty

Funkce **configure_ssl_context** a další v **ssl_utils** modulu zajišťují správnou práci s SSL/TLS certifikáty:

1. **Načtení Certifikátů:** Program načte certifikáty ze souboru (**-c certfile**) nebo adresáře (**-C certaddr**). Pokud nejsou specifikovány, použije výchozí systémové certifikáty.
2. **Ověření Certifikátu Serveru:** Po navázání SSL spojení ověří platnost certifikátu serveru pomocí OpenSSL funkcí.
3. **Nastavení Ověřovacího Režimu:** SSL kontext je nastaven na ověřování certifikátu serveru (**SSL_VERIFY_PEER**).

5. Popis řešení

5.1 Synchronizační Logika Obsahu

Program synchronizuje obsah výstupní složky s obsahem serveru následujícím způsobem:

- **Přepnutí na jinou schránku:** Při výběru nové schránky jsou zprávy z předchozí schránky odstraněny z výstupní složky. Toto zajišťuje, že obsah složky vždy odpovídá aktuálně vybrané schránce.
- **Opětovné použití programu ve stejné schránce:** Pokud je program spuštěn znovu ve stejné schránce, stáhne všechny zprávy znovu, bez ohledu na to, zda již byly dříve staženy. Tato logika byla zvolena, aby uživatel mohl nejprve stáhnout všechny své zprávy a poté používat program k stahování pouze jejich hlaviček a naopak.

5.2 Správa Příznaku "unseen"

Po stažení nepřečtené zprávy klient resetuje příznak "unseen" (`\UNSEEN`) na serveru. Tímto krokem se snaží synchronizovat stav zprávy, jelikož stažení zprávy může indikovat její přečtení uživatelem. I když není zcela přesné, zda uživatel zprávu skutečně přečetl, resetování příznaku umožňuje vyhýbat se opakovanému stahování již zpracovaných zpráv.

6. Návod na použití

6.1 Formát souboru s autentizačními údaji

Soubor s autentizačními údaji (**auth_file**) musí být ve formátu jednoduchého textu a obsahovat následující položky, každou na samostatném řádku:

username = uživatelské_jméno
password = heslo

6.2 Příklady spuštění programu

1. Stažení všech zpráv bez TLS:

   ./imapcl server -a auth.txt -o zpravy

2. Stažení nových zpráv s TLS a certifikátem:

   ./imapcl server -T -c cert.pem -n -a auth.txt -o zpravy

3. Stažení hlaviček zpráv:

   ./imapcl server -h -a auth.txt -o -h

7. Popis testování aplikace

7.1 Popis provedených testů a výsledky testování

1. **Test připojení s a bez TLS**:
   1.1. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt*
   Výstup: ***Staženo 9 zpráv ze schránky INBOX.***

   1.2. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -T*
   Výstup: ***TSL connection established. Staženo 9 zpráv ze schránky INBOX.***

2. **Test připojení na správnou a nesprávnou adresu:**
   2.1. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt*
   Výstup: ***Staženo 9 zpráv ze schránky INBOX.***

   2.2. *Vstup: ./imapcl fit.vutbr.cz -o maildir -a credentials.txt*
   Výstup: ***Invalid address/Domain name not supported***

3. **Test parametrů portu:**
   3.1. *Vstup: ./imapcl eva.fit.vutbr.cz -p 143 -o maildir -a credentials.txt*
   Výstup: ***Staženo 9 zpráv ze schránky INBOX.***

   3.2. *Vstup: ./imapcl eva.fit.vutbr.cz -p 14 -o maildir -a credentials.txt*
   Výstup: ***Connection Failed***

   3.3. *Vstup: ./imapcl eva.fit.vutbr.cz -p asdf -o maildir -a credentials.txt*
   Výstup: ***Error: port must contain only numbers.***

4. **Test autentizace s platnými a neplatnými údaji**:
   4.1. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt*
   Výstup: ***Staženo 9 zpráv ze schránky INBOX.***

4.2. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a wrong_credentials.txt*
Výstup: ***Server error: a001 NO [AUTHENTICATIONFAILED] Authentication failed.***

5. **Stažení různých typů zpráv**:

5.1. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -n*
Výstup: ***Žádné nové zprávy ve schránce INBOX.*** - pokud není k dispozici

Stažena 1 nová zpráva ze schránky INBOX. - pokud je k dispozici

5.2. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -h*
Výstup: ***Staženy hlavičky 6 zpráv ze schránky INBOX.***

5.3. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -h -n*
Výstup: ***Žádné nové zprávy ve schránce INBOX.*** - pokud není k dispozici

Stažena hlavička 1 nové zprávy ze schránky INBOX. - pokud je k dispozici

6. **Test práce s certifikáty**:

6.1. *Vstup:* ./imapcl eva.fit.vutbr.cz -T -C Sertificate/ -o maildir -a credentials.txt
Výstup: ***TSL connection established. Staženo 9 zpráv ze schránky INBOX***

6.2. *Vstup:* ./imapcl eva.fit.vutbr.cz -T -C Wrong_Sertificate/ -o maildir -a credentials.txt
Výstup: ***Certificate directory is empty or contains no valid certificates***

6.3. *Vstup:* ./imapcl eva.fit.vutbr.cz -T -c cacert.pem -o maildir -a credentials.txt
Výstup: ***TSL connection established. Staženo 9 zpráv ze schránky INBOX.***

6.4. *Vstup:* ./imapcl eva.fit.vutbr.cz -T -c wrong.pem -o maildir -a credentials.txt
Výstup: ***Error loading certificate***

7. **Test ukládání zpráv:**

7.1. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt*

Výstup: ***Staženo 9 zpráv ze schránky INBOX.***

7.2. *Vstup: ./imapcl eva.fit.vutbr.cz -o wrong_maildir -a credentials.txt*

Výstup: ***Error: output directory does not exist: wrong_maildir***

8. **Test parametrů schránky:**

8.1. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -b Sent*

Výstup: ***Staženo 67 zpráv ze schránky Sent.***

8.2. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -b Trash*

Výstup: ***Staženo 1568 zpráv ze schránky Trash.***

8.3. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -b Error*

Výstup: ***Server error: a002 NO Mailbox doesn't exist: Error (0.001 + 0.000 secs).***

9. **Testování neplatného formátu příkazového řádku:**

9.1. *Vstup: ./imapcl -o maildir -a credentials.txt*

Výstup: ***Error: server IP or domain name is required***
Usage: ./imapcl server [-p port] [-T [-c certfile] [-C certaddr]] [-n] [-h] -a auth_file [-b MAILBOX] -o out_dir

9.2. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir*

Výstup: ***Error: credentials file is required (-a auth_file)***
Usage: ./imapcl server [-p port] [-T [-c certfile] [-C certaddr]] [-n] [-h] -a auth_file [-b MAILBOX] -o out_dir

9.3. *Vstup: ./imapcl eva.fit.vutbr.cz -a credentials.txt*

Výstup: ***Error: output directory is required (-o out_dir)***
Usage: ./imapcl server [-p port] [-T [-c certfile] [-C certaddr]] [-n] [-h] -a auth_file [-b MAILBOX] -o out_dir

9.4. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -z*

Výstup: ***./imapcl: invalid option -- 'z'***
Usage: ./imapcl server [-p port] [-T [-c certfile] [-C certaddr]] [-n] [-h] -a auth_file [-b MAILBOX] -o out_dir

9.5. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -b*

Výstup: ***./imapcl: option requires an argument -- 'b'***
Usage: ./imapcl server [-p port] [-T [-c certfile] [-C certaddr]] [-n] [-h] -a auth_file [-b MAILBOX] -o out_dir

9.6. *Vstup: ./imapcl eva.fit.vutbr.cz extra_argument -o maildir -a credentials.txt*

Výstup: ***Error: unexpected argument(s): extra_argument***
Usage: ./imapcl server [-p port] [-T [-c certfile] [-C certaddr]] [-n] [-h] -a auth_file [-b MAILBOX] -o out_dir

9.7. *Vstup: ./imapcl eva.fit.vutbr.cz -o maildir -a credentials.txt -C* Sertificate/

Výstup: ***Error: -C and -c options require -T to be specified.***
Usage: ./imapcl server [-p port] [-T [-c certfile] [-C certaddr]] [-n] [-h] -a auth_file [-b MAILBOX] -o out_dir

8. Použitá literatura

1. RFC 3501 - IMAP4rev1
2. RFC 5322 - Internet Message Format
3. Dokumentace OpenSSL
4. Standardní knihovny C/C++