

Projekt na získanie zápočtu z predmetu Počítačové siete LS 2024/25

27. marca 2025

1 Získanie zápočtu

1.1 Formálne aspekty

Vaše repozitáre musia byť súkromné (private), nie verejné (public).
Predídeme tým nepríjemným nedorozumeniam ohľadom autorstva.

- Odovzdávať sa bude prostredníctvom githubu, sprístupnením repozitára menom **reader** užívateľovi **gjenca**
- Repozitár bude obsahovať iba jeden súbor so zdrojovým kódom, nič iné.
- Termín odovzdania projektu je 12.4.2025 23:59:59.
- Po tomto termíne budú projekty vyhodnotené, bude oznámený počet bodov.
- V prípade, že niekto bude mať záujem o zvýšenie počtu bodov, bude vypísaný ďalší termín dokedy budú môcť záujemcovia odovzdať druhú verziu programu.

1.2 Kritériá hodnotenia

- Posudzovať sa bude korektnosť implementácie v zmysle špecifikácie protokolu.

- Posudzovať sa bude aj zdrojový kód: pokiaľ niekto odovzdá príliš škaredý kód, dostane menej bodov.
- Používajte, (adekvátne účelu) prostriedky Pythonu: funkcie, slovníky, triedy, zabudované dátové typy.

1.3 Testovač zadania

Na URL https://github.com/gjenca/test_reader.git je malý testovací systém pre testovanie vášho riešenia. Pokyny pre použitie vid' tam. Určite ho použijete predtým, ako mi pošlete riešenie.

Ak nájdete nejakú chybu v tých testoch (snažil som sa ich písať robustne, ale je možné, že tom v nich spravil *ja* nejakú chybu), nahláste mi to cez issue tracker na URL https://github.com/gjenca/test_reader/issues, alebo mi pošlite email.

POZOR! Testy sa budú časom meniť a pribúdať, tak ako budem Vaše zadania opravovať.

2 Zadanie

Napíšte forkujúci sa TCP server, ktorý bude vykonávať službu zapisovania a čítania riadkov z textových súborov z podadresára `data` aktuálneho adresára.

3 Protokol

3.1 Základné pravidlá

- Protokol bude pre texty používať kódovanie UTF-8, koniec riadku je `<LF>` (ASCII kód 10), ako v Unixe.
- Bude postavený nad TCP.
- Klient bude serveru posilať *požiadavky* (anglicky *request*) a server po každej požiadavke pošle naspäť *odpoveď* (anglicky *response*).

3.2 Požiadavka

- Požiadavka bude vždy obsahovať v prvom riadku *metódu* (anglicky *method*), t.j. jedno slovo z množiny slov `READ`, `LS`, `LENGTH`.

- Nasledovať bude bližšia špecifikácia požiadavky, pomocou niekoľkých *hlavičiek* (headers). Hlavičky nemusia byť uvedené (môže ich byť 0). Každá hlavička je na jednom riadku. Hlavička sa skladá z *identifikátora* a *hodnoty*, sú oddelené dvojbodkou. Identifikátor musí byť ASCII reťazec bez bielych znakov a nesmie obsahovať znak : (dvojbodka). ¹ Príklad hlavičky:

```
File:list_of_students.txt
```

- Po hlavičkách bude nasledovať vždy práve jeden prázdny riadok, ktorý značí koniec hlavičiek.

3.3 Odpoveď

- Odpoveď bude mať v prvom riadku *stav* (anglicky status).
- Stav obsahuje ako prvé slovo trojciferné číslo, nasledované práve jednou medzerou a popisom. Príklad stavu:

```
200 Bad request
```

- Nasledujú hlavičky, v rovnakom tvare ako pri požiadavke.
- Potom bude vždy práve jeden prázdny riadok.
- Nasledovať obsah odpovede (alebo nič), dĺžka vyplýva z metódy a hlavičiek (*Lines*). Obsah sú vždy nula alebo viac riadkov.
- Spojenie za normálnych okolností ukončí klient, zavretím socketu.

3.4 Metódy

Teraz ideme špecifikovať po jednotlivých metódach typy rôznych požiadaviek, možných odpovedí a tomu zodpovedajúcich efektov a prípadne obsahov odpovedí. Ak nie je v tabuľke v poslednom stĺpci uvedený žiaden obsah odpovede, obsah odpovede je vtedy prázdny.

¹Vid' metódy `isascii` a `isspace` v dokumentácii Pythonu

3.4.1 READ

Hlavičky:

- **Názov hlavičky** `File`
Popis hlavičky Názov textového súboru odkiaľ treba prečítať obsah. Nesmie obsahovať znak `/`. Súbor sa má nachádzať v podadresári `data`.
- **Názov hlavičky** `From`
Popis hlavičky Nezáporné prirodzené číslo. Prvý riadok, ktorý chceme prečítať, číslovanie riadkov začína od 0.
- **Názov hlavičky** `To`
Popis hlavičky Nezáporné prirodzené číslo. Riadok po poslednom riadku, ktorý chceme prečítať. `From` a `To` majú teda rovnakú sémantiku ako funkcia `range` v Pythone, alebo vystrihovanie zo zoznamov v Pythone.

Správanie sa v rôznych situáciách:

- **Situácia** V hlavičkách sú neprípustné hodnoty, alebo niektorá hlavička chýba. Toto zahŕňa aj situáciu, keď je hodnota `From` väčšia ako hodnota `To`.
Odpoveď (status) `200 Bad request`
Odpoveď (hlavičky) Žiadne hlavičky odpovede.
Odpoveď (obsah) Prázdny obsah odpovede.
- **Situácia** Požaduje sa viac riadkov, ako je v súbore.
Odpoveď (status) `201 Bad line number`.
Odpoveď (hlavičky) Žiadne hlavičky odpovede.
Odpoveď (obsah) Prázdny obsah odpovede.
- **Situácia** Súbor neexistuje (`FileNotFoundError`).
Odpoveď (status) `202 No such file`
Odpoveď (hlavičky) Žiadne hlavičky odpovede.
Odpoveď (obsah) Prázdny obsah odpovede.
- **Situácia** Súbor sa nedá prečítať z iných dôvodov (`OSError`).
Odpoveď (status) `203 Read error`

Odpoveď (hlavičky) Žiadne hlavičky odpovede.

Odpoveď (obsah) Prázdny obsah odpovede.

- **Situácia** Normálny stav

Odpoveď (status) 100 OK

Odpoveď (hlavičky) Lines: *dĺžka odpovede v riadkoch*

Odpoveď (obsah) Riadky zo súboru.

3.4.2 Pomôcky

Pre otváranie súboru pre čítanie použite `with/open` idióm, aby ste nemuseli myslieť na to, že ho treba zavrieť.

```
>>> with open('data/a.txt') as f:  
...     lines=f.readlines()
```

3.4.3 LS

Hlavičky: Žiadne

Správanie sa v rôznych situáciách:

- **Situácia** Normálny stav

Odpoveď (status) 100 OK

Odpoveď (hlavičky) **Lines:** n , kde n je počet správ/súborov (riadkov odpovede).

Odpoveď (obsah) Názvy všetkých správ/súborov adresári **data**.

3.4.4 Pomôcky

Zoznam súborov adresári zistíte pomocou `os.listdir`.

```
>>> import os
>>> os.listdir('data')
['b.txt', 'a.txt']
```


3.4.5 LENGTH

Hlavičky:

- **Názov hlavičky** File

Popis hlavičky Názov textového súboru ktorého počet riadkov chceme zistiť. Nesmie obsahovať znak /. Súbor sa má nachádzať v podadresári data.

Správanie sa v rôznych situáciách:

- **Situácia** V hlavičkách sú neprípustné hodnoty, alebo niektorá hlavička chýba.
Odpoveď (status) 200 Bad request
Odpoveď (hlavičky) Žiadne hlavičky odpovede.
Odpoveď (obsah) Prázdny obsah odpovede.
- **Situácia** Súbor neexistuje (`FileNotFoundException`).
Odpoveď (status) 202 No such file
Odpoveď (hlavičky) Žiadne hlavičky odpovede.
Odpoveď (obsah) Prázdny obsah odpovede.
- **Situácia** Súbor sa nedá prečítať z iných dôvodov (`OSError`).
Odpoveď (status) 203 Read error
Odpoveď (hlavičky) Žiadne hlavičky odpovede.
Odpoveď (obsah) Prázdny obsah odpovede.
- **Situácia** Normálny stav
Odpoveď (status) 100 OK
Odpoveď (hlavičky) Lines: 1
Odpoveď (obsah) Jeden riadok, obsahujúci počet riadkov v tom súbore.

3.4.6 Pomôcky

3.4.7 Neznáma metóda

Ak klient pošle na začiatku požiadavky čokoľvek iné, ako je riadok obsahujúci známu metódu, server vráti status `204 Unknown method` s prázdnyimi hlavičkami a obsahom a *uzavrie spojenie*.

3.5 Príklad komunikácie klient-server

`C->S` sú označené dáta zasielané klientom serveru, `S->C` naopak. Všimnite si aj prázdne riadky označujúce koniec hlavičiek, sú súčasťou správnej implementácie protokolu.

```
C->S LS
C->S
S->C 100 OK
S->C Lines:2
S->C
S->C b.txt
S->C a.txt
C->S LENGTH
C->S File:a.txt
C->S
S->C 100 OK
S->C Lines:1
S->C
S->C 3
C->S READ
C->S File:a.txt
C->S From:0
C->S To:2
C->S
S->C 100 OK
S->C Lines:2
S->C
S->C This is line 1 of a.txt
S->C This is line 2 of a.txt
C closes connection
```