

Multithreaded Book Scraper

Описание

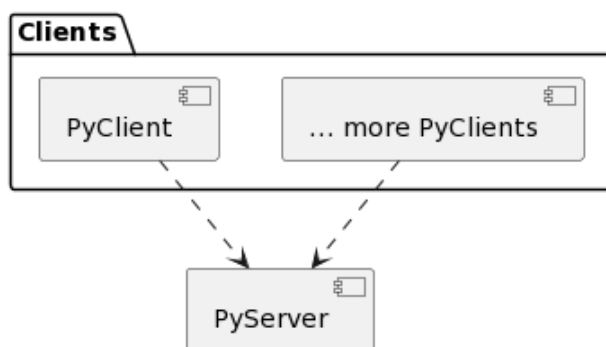
Апликация, която е написана на Python 3. Съдържа 2 файла - PyServer.py и PyClient.py. Използва две библиотеки - "BeautifulSoup" и "Requests".

Web scraping се извършва на сайта <http://books.toscrape.com>

Структура на приложението

Приложението е структурирано с клиент и сървър, написани като два отделни файла - PyClient.py и PyServer.py съответно.

Структурата може да бъде видяна на графиката приложена долу.



Класове

1. PyServer.py

Отговаря за слушането за заявки на даден адрес и порт (**localhost** и **8000**) и за тяхното обработване. За всеки нов клиент създава нова нишка, като по този начин може да работи с повече от един клиент едновременно.

- При стартиране на конзолата излиза съобщението **"OPENED A SERVER ON {HOST} WITH PORT {PORT}."**, където *HOST* е адреса, а *PORT* е портът.
- Функцията `handle_client()` отговаря за разпореждането на сървъра с даден клиент.
 - За всеки нов клиент, който се свърже към сървъра се изписва на конзолата съобщението **"New client connected: {address}"**.
- Когато клиент подаде команда **"exit"** сървърът затваря нишката на дадения клиент и продължава работа с останалите или стои в очакване за нови клиенти.
 - За всеки клиент, който се отвърже от сървъра се изписва на конзолата съобщението **"Client disconnected: {address}"**.

2. PyClient.py

Отговаря за свързването към вече съществуващ сървър на даден адрес и порт и за подаването на заявки към него.

- При стартиране на конзолата излиза съобщението **"Enter a number of titles you want to receive (or 'exit' to quit):"**, след което се чакат входни данни от потребителя във формат "<цяло число>", които обозначават колко заглавия на книги искаме да ни бъдат върнати от сървъра.

- При правилно подадени данни получаваме string със заглавия разделени от запетайки, който след това обработваме обратно в списък от string-ове и изписваме на конзолата.

Заявки

Както вече бе обяснено заявките са във вида “<цяло число>”, където това цяло число обозначава колко на брой заглавия искаме да ни бъдат върнати от сървъра, който прави web scraping на предварително зададен уебсайт.

Как да стартираме

1. Инсталираме **Python 3**.
2. Подаваме `pip3 install -r requirements.txt` в папката.
3. Подаваме `python3 PyServer.py` в папката.
4. Подаваме `python3 PyClient.py` в папката.

Изготвил: Емил Йорданов Йорданов, ФН: 82193, група 3, поток 1, КН.