

*Prolećni semestar, 2023/24*

*PREDMET: CS230 Distribuirani sistemi*

Projektni zadatak

***Razviti jednostavan distribuirani sistem za deljenja datoteka koristeći peer-to-peer arhitekturu sa centralizovanim indeksiranjem.***

Ime i prezime: **Aleksa Stanojević**

Broj indeksa: **5726**

Datum izrade: **07.06.2024.**

Asistent: **Andjela Grujić**

Profesor: **Nemanja Zdravković**

Sadržaj

[1. Uvod i teoretska postavka izabrane teme 3](#_Toc169289234)

[**1.1.** **Arhitektura distribuiranog sistema** 3](#_Toc169289235)

[**1.1.1.** **Peer-to-Peer (P2P) Arhitektura** 3](#_Toc169289236)

[**1.1.2.** **Centralizovano Indeksiranje** 3](#_Toc169289237)

[1.2. Map reducer 5](#_Toc169289238)

[1.3. Replikacija 5](#_Toc169289239)

[1.4. Korišćene tehnologije 6](#_Toc169289240)

[2. Primer/studija slučaja primene izabrane teme 6](#_Toc169289241)

[**2.1.** **Kratak opis razvijenog sistema** 6](#_Toc169289242)

[**2.2.** **Funkcionalnosti sistema uključuju:** 6](#_Toc169289243)

[**2.3.** **Struktura aplikacije** 6](#_Toc169289244)

[**2.4.** **Funkcionalni tok** 9](#_Toc169289245)

[Registracija Peer-ova 9](#_Toc169289246)

[Pretraga Datoteka 9](#_Toc169289247)

[Preuzimanje Datoteka 10](#_Toc169289248)

[Replikacija Datoteka 10](#_Toc169289249)

[**2.5.** **Korisnički interfejs** 10](#_Toc169289250)

[Zaključak 10](#_Toc169289251)

[Literatura 11](#_Toc169289252)

# Uvod i teoretska postavka izabrane teme

Ovaj projekat je jednostavan distribuirani sistem za deljenje datoteka koristeći peer-to-peer arhitekturu sa centralizovanim indeksiranjem. Sistem se sastoji od centralnog servera koji indeksira datoteke koje peer-ovi (klijenti) dele, i peer-ova koji učestvuju u razmeni datoteka.

Centralni server (indeksni server) vodi evidenciju o dostupnim datotekama i peer-ovima koji ih poseduju. Peer-ovi komuniciraju sa serverom kako bi registrovali svoje datoteke i pretraživali dostupne datoteke u mreži. Kada peer pronađe željenu datoteku, direktno preuzima datoteku od drugog peer-a. Ovaj pristup kombinuje prednosti P2P arhitekture, kao što su decentralizacija i skalabilnost, sa efikasnošću centralizovanog pretraživanja.

## **Arhitektura distribuiranog sistema**

Arhitektura ovog projekta za deljenje datoteka koristi kombinaciju Peer-to-Peer (P2P) arhitekture sa centralizovanim indeksiranjem. Sistem se sastoji od centralnog servera koji služi kao indeksni server i više peer-ova koji deluju kao klijenti u mreži.

### **Peer-to-Peer (P2P) Arhitektura**

P2P arhitektura je distribuirani sistem u kojem su svi čvorovi (peer-ovi) ravnopravni i dele resurse direktno jedni s drugima bez potrebe za centralnim serverom. Svaki peer funkcioniše kao klijent i server, omogućavajući deljenje datoteka, resursa i informacija unutar mreže. Ključne karakteristike P2P sistema uključuju:

* **Decentralizacija**: Nema centralnog servera koji kontroliše mrežu; svi peer-ovi su ravnopravni.
* **Skalabilnost**: Sistem može lako da se proširi dodavanjem novih peer-ova.
* **Robustnost**: Sistem je otporan na kvarove, jer gubitak jednog peer-a ne utiče značajno na celokupnu mrežu.

### **Centralizovano Indeksiranje**

Centralizovano indeksiranje kombinuje prednosti centralizovanog servera za pretragu sa decentralizovanim skladištenjem i deljenjem datoteka. U ovom pristupu, centralni server vodi evidenciju o datotekama koje su dostupne na različitim peer-ovima i omogućava brzo pretraživanje tih datoteka. Ključne prednosti centralizovanog indeksiranja su:

* **Efikasna Pretraga**: Brz i pouzdan način za pronalaženje datoteka unutar mreže.
* **Jednostavno Upravljanje:** Lakše je upravljati i održavati indeks datoteka na jednom mestu.
* **Poboljšana Performansa:** Smanjuje se opterećenje na pojedinačnim peer-ovima tokom pretrage.

Centralizovani direktorijum je donekle sličan arhitekturi klijent-server u smislu da održava ogroman centralni server za pružanje usluga direktorijuma. Svi peers-ovi obaveštavaju ovaj centralni server o svojoj adresi i datotekama koje stavljaju na raspolaganje za deljenje. Dakle, u osnovi ovaj server održava podatake o tome koja datoteka je prisutna na kojoj adresi. Prvi sistem koji je koristio ovu metodu bio je Napster, za distribuciju Mp3.

A diagram of a computer network

Description automatically generated

A diagram of a computer system

Description automatically generated

A diagram of a computer system

Description automatically generated

## Map reducer

MapReduce je model za obradu i generisanje velikih skupova podataka koji se koristi za paralelizaciju zadataka. Sastoji se od dve glavne faze: "Map" i "Reduce". U fazi "Map", podaci se razdeljuju u manje podskupove koji se obrađuju paralelno, dok u fazi "Reduce" obrađeni podaci se kombinuju u konačni rezultat. U ovom projektu, MapReduce princip je primenjen u funkciji pretrage datoteka, gde se pretraga distribuira među više peer-ova radi paralelne obrade. Koristeći više procesorskih jezgara, funkcija pretrage koristi radnički pool da efikasno pretraži sve datoteke na različitim peer-ovima i vrati rezultate centralnom serveru.

A diagram of a map

Description automatically generated

## Replikacija

Replikacija je proces kopiranja i održavanja identičnih verzija datoteka na više računara (peer-ova) radi povećanja dostupnosti i otpornosti na greške. U ovom projektu, kada se nova datoteka doda ili izmeni na jednom peer-u, centralni server koordinira replikaciju te datoteke na druge peer-ove. Funkcija za replikaciju osigurava da svaki peer koji treba imati kopiju datoteke dobija najnoviju verziju, čime se osigurava kontinuitet i integritet podataka u celom sistemu. Ova funkcionalnost omogućava korisnicima pristup datotekama čak i ako jedan ili više peer-ova postanu nedostupni.

## Korišćene tehnologije

**Flask (Python) :** Web framework korišćen za implementaciju sistema

Centralni server je postavljen pomoću Flask-a za registraciju i ažuriranje informacija o peer-ovima, pretragu datoteka i koordinaciju replikacije. Peer-ovi su korišćenjem Flask-a postavili lokalne web servere za preuzimanje i repliciranje datoteka, kao i za interakciju sa centralnim serverom za registraciju, pretragu i replikaciju. Takođe, Flask je korišćen za generisanje HTML šablona i CSS stilova, omogućavajući jednostavan web interfejs za pregledanje i pretragu datoteka od strane korisnika. Korišćenje Flask-a kao centralne tehnologije omogućilo je efikasnu komunikaciju i koordinaciju između svih komponenti sistema za deljenje datoteka.

# Primer/studija slučaja primene izabrane teme

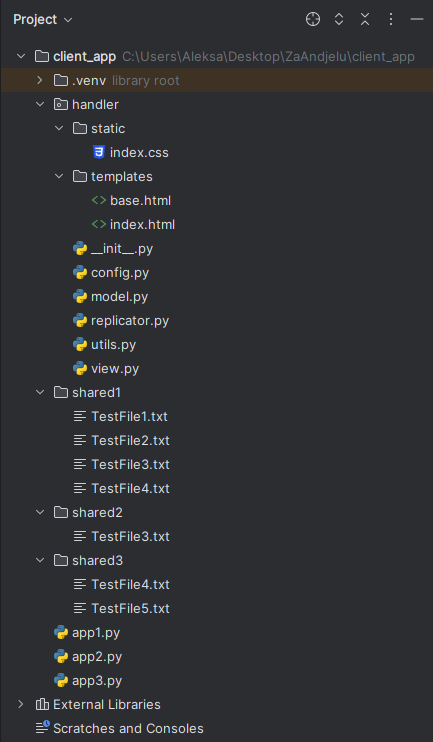
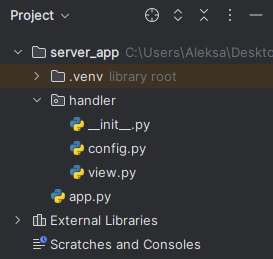
## **Kratak opis razvijenog sistema**

Razvijeni sistem za deljenje datoteka preko mreže je distribuirani sistem koji koristi peer-to-peer arhitekturu sa centralizovanim indeksiranjem. Sastoji se od centralnog servera i više peer-ova razvijenih u Flask-u. Ovaj sistem omogućava deljenje, pretragu i replikaciju datoteka među više računara, pružajući korisnicima efikasan način za razmenu podataka.

## **Funkcionalnosti sistema uključuju:**

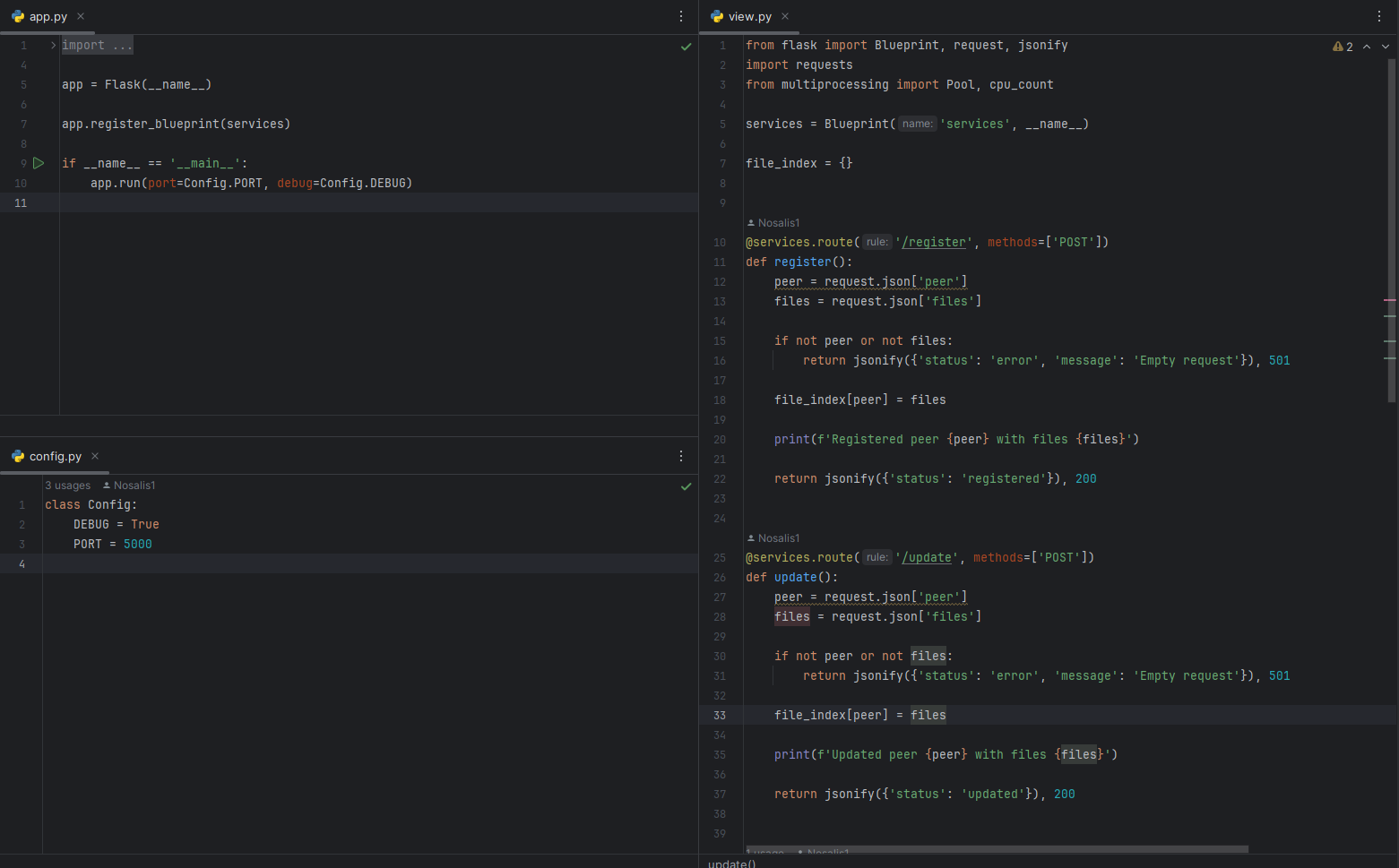
* Registracija peer-ova i njihovih datoteka na centralnom serveru.
* Ažuriranje informacija o datotekama na centralnom serveru od strane peer-ova.
* Pretraga datoteka na centralnom serveru.
* Preuzimanje i repliciranje datoteka između peer-ova.
* Prikaz i pretraga datoteka preko web interfejsa.

### **Struktura aplikacije**



**Centralni Server (handler, app.py)**

* **app.py:** Glavni fajl koji pokreće Flask server aplikaciju.
* **handler:** Sadrži rute za registraciju, ažuriranje, pretragu i replikaciju datoteka.



**Peer-ovi**

* **app.py:** Glavna datoteka za pokretanje peer-ova. Svaki peer koristi različite portove i direktorijume za deljenje datoteka.
* **config.py:** Konfiguraciona datoteka za podešavanje peer-ova.
* **model.py:** Model datoteke (DistributedFile) i funkcija za učitavanje datoteka iz direktorijuma.
* **replicator.py:** Dekorator replikator funkcije za poziv repliciranja fajlova na serveru od strane peer-a.
* **view.py:** Datoteka koja sadrži sve rute.
* **utils.py:** Pomoćne funkcije za rad sa datotekama.

**Šabloni i Stilovi**

* **base.html, index.html:** HTML šabloni za prikaz informacija o datotekama i interakciju sa korisnicima.
* **index.css:** CSS stilovi za dizajn korisničkog interfejsa.

## **Funkcionalni tok**

### Registracija Peer-ova

* Peer se pokreće i registruje kod centralnog servera, šaljući informacije o datotekama koje deli i poseduje.
* Server čuva ove informacije.

### Pretraga Datoteka

* Kada peer želi da pronađe određenu datoteku, šalje zahtev za pretragu serveru.
* Server pretražuje po imenu datoteke i vraća listu peer-ova koji poseduju traženu datoteku.

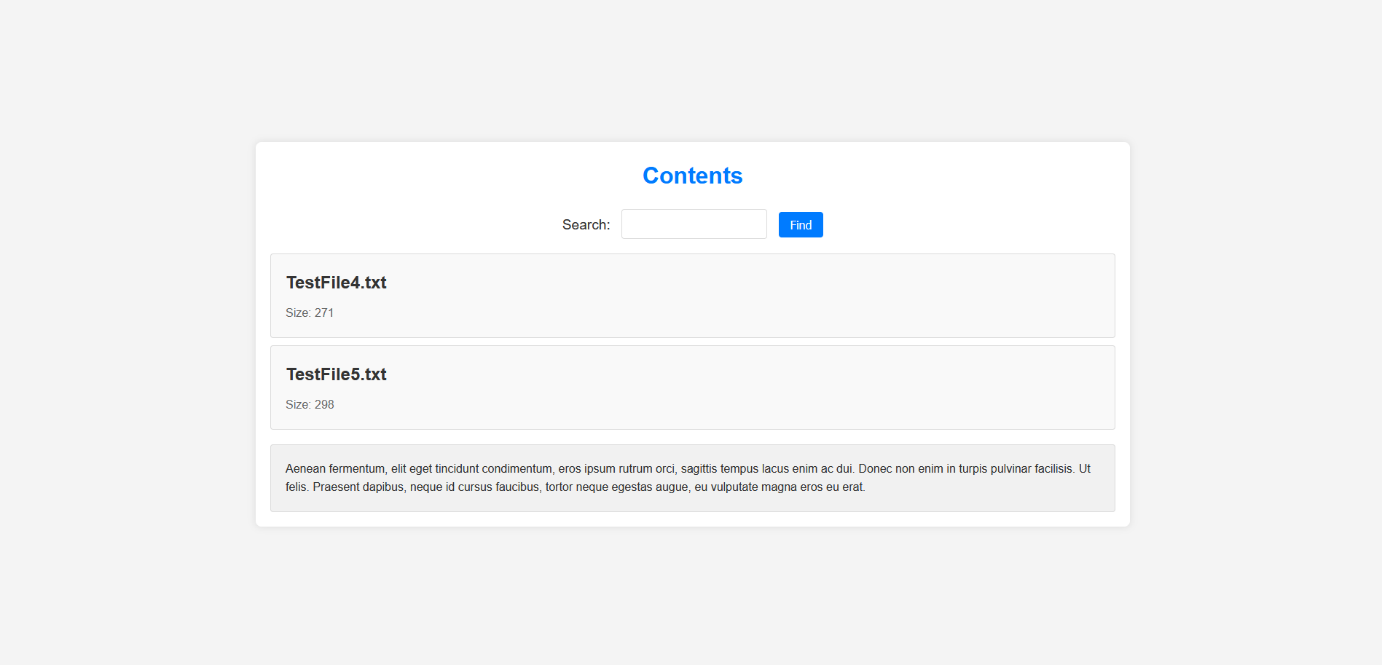
### Preuzimanje Datoteka

* Peer bira jednog od peer-ova iz rezultata pretrage tj. onog kog ima najnoviju veryiju datoteke i šalje mu zahtev za preuzimanje datoteke.
* Odabrani peer odgovara na zahtev i šalje kopiju datoteke.

### Replikacija Datoteka

* Kada peer ažurira datoteku, obaveštava centralni server o izmenama.
* Server zatim obaveštava druge relevantne peer-ove da preuzmu ažuriranu verziju datoteke.

## **Korisnički interfejs**



U korisničkom interfejsu za peer-a imamo takozvani search bar koji koristimo za pretragu distribuiranog fajla u mreži. Unosom naziva fajla i klikom na dugme „Find“ pretražiće se fajl i ukoliko se zahtev uspešno obradi fajl će se dodati u listu ispod. U listi fajlova prikazujemo ime fajla koji se čuva kao ii njegovu veličinu. Klikom na bilo koju karticu fajla otvaramo prikaz kontenta istog ispod same liste.

# Zaključak

Razvijeni sistem za deljenje datoteka putem peer-to-peer arhitekture sa centralizovanim indeksiranjem pruža efikasno i skalabilno rešenje za razmenu podataka između više računara. Korišćenjem Flask-a za razvoj centralnog servera i peer-ova, sistem omogućava jednostavnu registraciju, pretragu, preuzimanje i replikaciju datoteka. Primena Map reducera i implementacija replikacije povećava dostupnost i otpornost sistema na greške. Ovaj projekt predstavlja osnovu za naprednije distribuirane sisteme koji mogu podržati kompleksnije funkcionalnosti i veće količine podataka.

**Moguće dodatne funkcionalnosti**

* Autentifikacija i autorizacija: Implementacija sistema za autentifikaciju korisnika i autorizaciju pristupa datotekama kako bi se osigurala sigurnost i privatnost podataka.
* Šifrovanje datoteka: Dodavanje opcije za šifrovanje datoteka prilikom razmene između peer-ova radi zaštite od neovlašćenog pristupa.
* Poboljšanje pretrage: Optimizacija algoritama za pretragu i uvođenje filtera za pretragu kao što su pretraga po tipu datoteke, veličini, i datumu modifikacije.
* prilagođavanje udaljenosti peer-ova za optimizaciju preuzimanja

# Literatura

* Waterloo, Sveučilište. "Peer-to-Peer Architecture Design Activity.", <https://student.cs.uwaterloo.ca/~cs446/1171/Arch_Design_Activity/Peer2Peer.pdf>
* ResearchGate. "The centralized peer-to-peer (P2P) system: A peer E sends a message to the central server.", <https://www.researchgate.net/figure/The-centralized-peer-to-peer-P2P-system-A-peer-E-sends-a-message-to-the-central-server_fig3_356245976>
* Migara, Pramod M. "Peer-to-Peer File Sharing Protocols." , <https://migarapramodm.medium.com/peer-to-peer-file-sharing-protocols-2b6a1ba162ab>
* GeeksforGeeks. "P2P (Peer-to-Peer) File Sharing." , <https://www.geeksforgeeks.org/p2p-peer-to-peer-file-sharing/>
* Jinja. "Jinja Documentation." , <https://jinja.palletsprojects.com/en/3.1.x/>