



Fakultet informatike u Puli

2. Manji projekt RS

Master worker architecture

Dataset = FakeDataset (iz proslog kolokvija)

Klijent

Generira se lista od 10 000 klijenata. (client ids) Učitava dataset i uzima stupac koji sadrži python kod. Dataset se podijeli ravnomjerno klijentima. (dict klijenata i njihov python kod) Klijenti šalju zahtjeve za obradu koda. Ispisuje u konzolu za svakog klijenta prosječan broj slova koji sadrži sav njihov python kod.

Master servis (radi async..)

Pokreće na random 5-10 workera. (ovo je number of workers = N) Po pokretanju workera dodjeljuje im se id. (spremaju se u dict, potrebno da se zna koji worker je obavio sto). Servis prati broj primljenih i obavljenih zadataka. Logira se timestamp slanja i primanja taska. Salje svakom workeru po 1000 redaka.[ovo je sample size = M, task za workera] Nakon sto dobije rezultate od workera, salje sljedeće redke.

Worker servisi (rade async, svi imaju isti kod)

Posto se radi na lokalnoj masini potrebno je “simulirati” delay mreže. Kad worker dobije i vraća obavljeni task na random čeka 0.1 - 0.3 sekunde. Worker računa broj riječi u python “fileu” . (ala “word counter”)

HINTS!!!

Svaki servis pokreće se u zasebnom terminalu.

Ukoliko se radi o PC s manje RAM, onda tipa pola dataseta i manje redaka, workera (minimum 3)

Constraints!!!

- Dopustene biblioteke:
 - asynio

- pandas
- Python standard library