

Operativni sistemi 2 – II kolokvijum

Institut za matematiku i infomatiku

19. januar. 2015. godine

12 poena

Nad pustinjom Sahara se nalazi n satelita koji prate kretanje 4 kamiona koja učestvuju na Dakar reliju. Sateliti su dati u fajlu **sateliti.txt** kao:

ID_satelita, vreme_azuriranja_brzine

$ID_1 t_1$

$ID_2 t_2$

.

.

.

$ID_n t_n$

Brzine kretanja 4 kamiona na Dakar reliju se upisuju u fajlove za svaki od 4 kamiona posebno (**brzineKamiona_1.txt**, ..., **brzineKamiona_4.txt**). Primer izgleda fajlova **brzineKamiona_1.txt**, ..., **brzineKamiona_4.txt** je dat u nastavku:

$V_1 ID_1$

$V_2 ID_2$

$V_3 ID_3$

$V_4 ID_2$

.

.

.

Satelit S_j može upisati brzinu za kamion K u odgovarajući fajl **brzineKamiona_K.txt** u bilo kom trenutku na t_j sekundi u odnosu na vreme prethodnog upisa brzine satelita (bilo kojeg) za kamion K . Brzine su date u **m/s** i njihov broj u fajlovima nije unapred poznat. Smatrati da brzina raste linearno na vremenskom intervalu koji traje od prethodnog upisa zaključno sa novim. ID_i predstavlja satelit koji je upisao brzinu kamiona K posle t_j sekundi u odnosu na poslednji upis brzine za taj kamion. Početna brzina sva 4 kamiona je 0 m/s.

Primer: : Neka je dat fajl **sateliti.txt** kao:

2 3

3 2

5 4

...

i neka je dat fajl **brzineKamiona_K.txt** kao:

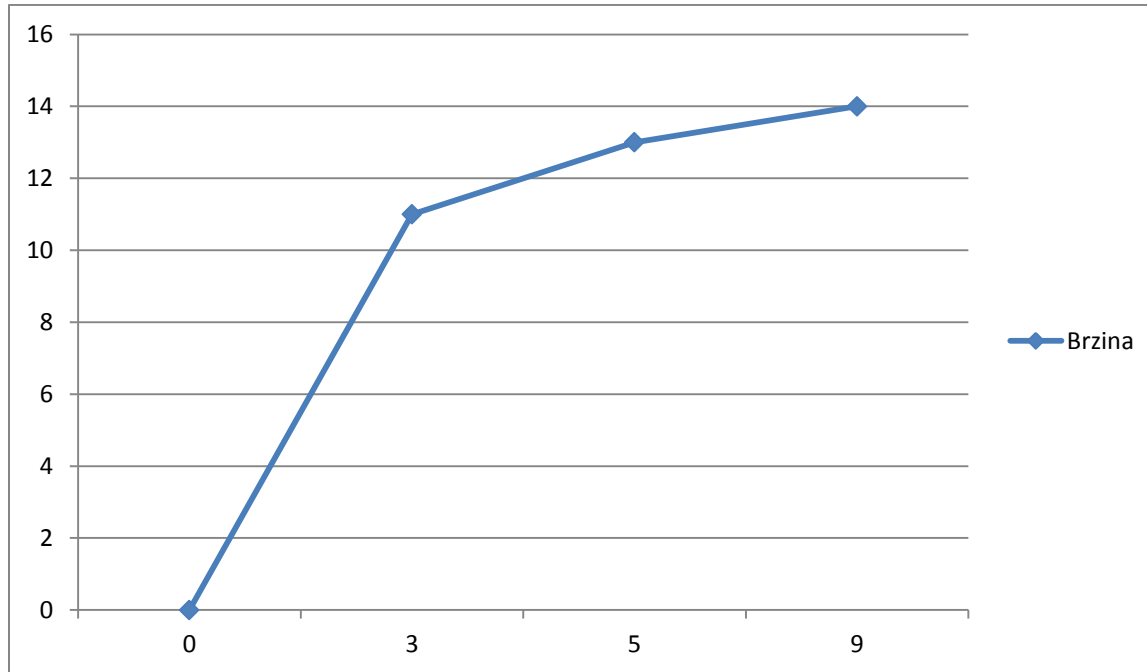
11 2

13 3

14 5

...

Linija „11 2“ znači da je satelit 2 zabeležio brzinu 11 m/s, 3 sekunde od početka kretanja kamiona. Linija „13 3“ znači da je satelit 3 zabeležio brzinu 13 m/s, 2 sekunde nakon merenja satelita 2. Linija „15 5“ znači da je satelit 5 zabeležio brzinu 14 m/s, 4 sekunde nakon merenja satelita 3.



Grafik 1 – Promena brzine kamiona

Potrebno je na klaster računaru sortirati kamione po ukupnom pređenom putu, u neopadajućem poretku. Rešenje dati u vidu parametarskog posla gde je potrebno kreirati:

- **Bash** skripta za pokretanje parametarskog posla (**pokreni.sh**), kao i integrisanje rezultata u fajl **resenje.txt**,
- Odgovarajući **posao.sub** fajl za pokretanje parametarskog posla,
- Programski kod (C programski jezik) za obradu log fajlova.

9 poena

Neka su u direktorijumu “studenti” na HDFS fajl sistemu postavljeni CSV fajlovi koji čuvaju podatke o studentima na Institutu za matematiku i informatiku u formatu: **Indeks, Ime, Prezime, Datum_rodjenja, Vrsta_studija**, dok se u direktorijumu “prijave” nalaze CSV fajlovi koji čuvaju podatke o položnim ispitima u formatu: **Datum_polaganja, Naziv_predmeta, Indeks, Ocena**.

Potrebno je napisati **map/reduce** program koji odgovara SQL upitu u nastavku:

```
select studenti.Indeks, studenti.Ime, avg(prijave.Ocena)
from studenti
join prijave on
studenti.Indeks = prijave.Indeks
where Vrsta_studija = 'informatika'
group by studenti.Indeks, s.Ime
```