

**POREĐENJE BRZINE ČITANJA PODATAKA SKLADIŠTENIH PO VRSTAMA I KOLONAMA U PostgreSQL BAZI PODATAKA****COMPARISON OF DATA SPEED READING STORAGE BY ROWS AND COLUMNS IN A PostgreSQL DATABASE**

Sanja Miljević, *Fakultet tehničkih nauka, Novi Sad*

**Oblast – ELEKTROTEHNIKA I RAČUNARSTVO**

**Kratak sadržaj** – U radu su izvršene dve grupe testova kod kojih se poredi brzina čitanja podataka skladištenih po vrstama i kolonama u PostgreSQL-u. Na osnovu dobijenih rezultata, pokazano je da je organizovanje podataka po kolonama pogodnije za agregaciju (sakupljanje, rezimiranje ili grupisanje) podataka.

**Ključne reči:** PostgreSQL, baza podataka, SQL, skladištenje podataka prema redovima, skladištenje podataka prema kolonama, izvršavanje upita

**Abstract** – Two groups of tests were performed in the paper, comparing the read speed of data stored by rows and columns in PostgreSQL. Based on the results obtained, it is shown that organizing data by columns is more suitable for aggregation (collection, summarization or grouping) of data.

**Keywords:** PostgreSQL, Database, SQL, Row-stored, Column-stored, C-store, query execution

**1. UVOD**

Baza podataka nudi organizovan mehanizam za skladištenje, upravljanje i dobijanje informacija, gde podatke možemo da skladištimo unutar tabela po redovima ili kolonama. Koje skladištenje podataka je bolje zavisi od toga koji tip upita se izvršava. U radu su kreirani eksperimenti gde je testirano izvršavanja upita. Prvi deo eksperimenta predstavljaju tabele koje se sastoje od jedne, dve i tri kolone, a drugi deo eksperimenta predstavlja realan primer u vidu ankete.

Ankete predstavljaju jednu tehniku prikupljanja podataka koje se sprovode među odabranim brojem osoba određene populacije na određenu temu. Ciljevi ankete mogu biti poslovnog ili istraživačkog karaktera [1,2]. Rezultati ankete se analiziraju i na osnovu tih rezultata se dobijaju određeni zaključci.

Podaci koji su prikupljeni putem ankete su negrupisani (sirovi) podaci. Podaci se svrstavaju u određene grupe po srodnosti na osnovu različitih kriterijuma u zavisnosti od odgovora ispitanika. Analiziranjem međusobnih odgovora ispitanika dobijaju se ključni odgovori na temu od interesa, u cilju bilo kakve vrste poboljšanja.

**NAPOMENA:**

**Ovaj rad proistekao je iz master rada čiji mentor je bio dr Slobodan Ilić, docent.**

U okviru rada analizirano je vreme koje je potrebno da se dobiju određeni podaci iz tabela i upoređena je razlika u izvršavanju upita kada se podaci skladište prema redovima i kolonama. Na osnovu dobijenih rezultata, pokazano je da je organizovanje podataka po kolonama pogodnije za anketnu analizu podataka.

**2. NAČINI SKLADIŠTENJA PODATAKA U RELACIONIM BAZAMA PODATAKA**

Baza podataka se sastoji od više tabela. Tabele se sastoje od kolona i redova. Svaka kolona odgovara jednom atributu, a svaki red odgovara pojedinačnom zapisu.

Podatke unutar tabela možemo da skladištimo atribut po atribut ili zapis po zapis. Odnosno, baza podataka može biti:

- Skladištena prema redovima (engl. *Row stored*)
- Skladištena prema kolonama (engl. *Column stored* poznato kao *columnar* ili *C-store*)

**2.1. Skladištenje podataka prema redovima**

Uobičajen, odnosno tradicionalan, način skladištenja tabela je serijalizacija svakog pojedinačnog reda. Redovi se skladište jedni pored drugih u memoriji. Ovi zapisi su lako čitljivi i lako se zapisuju. Ova vrsta skladišta je dizajnirana da vrati podatke za ceo red u što je moguće manjem broju operacija. Ovi sistemi nisu efikasni kada se traži mali broj konkrentnih podataka nad celom tabelom.

**2.2. Skladištenje podataka prema kolonama**

Za skladištenje podataka prema kolonama, podaci se za sve vrednosti određene kolone čuvaju uzastopno na disku. Operacije koje preuzimaju sve podatke za ceo red su sporije. Kod sistema koji skladište podatke prema kolonama potrebne su brojne operacije kod pristupanja disku za prikupljanje podataka iz više kolona. Ovakve vrste zahteva su uglavnom retke. U većini slučajeva, samo određeni deo podataka se preuzima. Upravo zbog ovoga skladištenje prema kolonama pokazuje odlične performanse, uprkos očiglednim nedostacima. Sistemi koji skladište podatke prema kolonama se najčešće koriste za analitičku obradu podataka kod veoma velikog broja podataka.

**3. PostgreSQL**

PostgreSQL je objektno-relacioni sistem za upravljanje bazama podataka. PostgreSQL koristi klijent server model [3,4].

### 3.1. Izvršavanje upita

Upite izvršavamo pomoću jezika SQL-a u PostgreSQL sistemu za upravljanje bazama podataka.

Pomoću upita iz tabele moguće je prikazati sirove podatke ili obraditi i prikazati jedan deo podataka. Da bi se dobio deo podataka iz tabele, potrebno je izvršiti grupisanje podataka.

Grupisanje podataka može se izvršiti prema redovima ili kolonama na osnovu zadatih kriterijuma u skupove ili skupove i podskupove, sa ciljem da podaci budu pregledni da bi mogli lakše da se analiziraju. Kriterijumi za grupisanje su specifične vrednosti za analizu od interesa. Nad podacima koji su grupisani vrednosti se mogu prebrojati, sabrati, izračunati prosek i druge operacije i time dobiti statističke serije podataka.

## 4. NAČINI SKLADIŠTENJA PODATAKA PREMA KOLONAMA U PostgreSQL-u

PostgreSQL sistem za upravljanje bazama podataka čuva podatke tako što ih skladišti redno. Za skladištenje po kolonama postoje različiti načini modifikacije u samom PostgreSQL, kao i razna literatura na ovu temu [6, 7, 8].

Modifikacija šeme, modifikacija menadžera skladištenja, modifikacija menadžera skladištenja i izvršioca upita predstavljaju izmene koje se vrše u samom PostgreSQL SUBP i oslanjaju se na skladištenje podataka orijentisanih prema redovima.

Ekstenzija `cstore_fdw` predstavlja poseban dodatak za skladištenje podataka prema kolonama koji se naknadno instalira u PostgreSQL. Ova ekstenzija koristi format za raspored podataka koji je inspirisan ORC-om (engl. *Optimized Row Columnar*). ORC čuva kolekciju redova u jednom fajlu i unutar kolekcije podaci reda se čuvaju u formi kolona. `Cstore_fdw` koristi indekse preskakanja koji omogućavaju preskakanje blokova podataka. Kod skladištenja podataka prema kolonama u ovom radu koristimo ovu ekstenziju.

## 5. EKSPERIMENTI I REZULTATI

Da bi se uporedila efikasnost izvršavanja upita nad podacima organizovanim prema redovima i kolonama osmišljena su i izvršena dva eksperimenta. Prvi eksperiment je zasnovan na veštački generisanim podacima, dok se u drugom eksperimentu koristila realna baza podataka u vidu ankete. Svaki upit je testiran jednom nad tabelom, na tek pokrenutom operativnom sistemu. Između izvršavanja dva različita upita nad istom tabelom restartovan je server. Ovim se izbegava brže izvršavanje upita zbog punjenja RAM i keš memorije.

### 5.1. Eksperimentalni deo jedan

U prvom eksperimentalnom delu obuhvaćene su tabele sa jednom kolonom, tabele sa dve kolone i tabele sa tri kolone. Tabele se popunjavaju sa numeričkim ili tekstualnim veštački generisanim podacima. Kreirane su tabele iste strukture za skladištenje podataka prema redovima i prema kolonama, sa različitim brojem  $10^1$ ,  $10^2$ ,  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$ ,  $10^7$  zapisa. Testiramo i poredimo vremena izvršavanja upita nad tabelama. Kod pisanja upita koriste se `EXPLAIN ANALYZE` naredbe kao prava naredba svakog upita. Ove naredbe pomažu da se poboljšaju

performanse baze podataka i prikazuju vreme izvršavanja (engl. *Run time*) koje se sastoji od vremena koje je potrebno za planiranje (engl. *Planning time*) kako će se upit izvršiti i vreme samog izvršavanja (engl. *Execution time*) upita. U ovom radu uzima se u obzir vreme samog izvršavanja upita nad podacima.

Testovi su izvođeni na PC sa OS ubuntu 16.4, CPU AMD A8 (2,10 GHz), 4 GB RAM i SSD 250 GB.

U nastavku je prikaz i analiza upita tabele sa tri kolone sa tekstualnim i numeričkim podacima.

#### 5.1.1. Tabele sa tri kolone sa tekstualnim i numeričkim podacima

Prva kolona *gender* sadrži podatke o polu, a moguće vrednosti su *Male*, *Female* i *NA*. Druga kolona *formaleducation* sadrži podatke o stepenu obrazovanja, a moguće vrednosti su *Bachelor's degree*, *Master's degree*, *Secondary school*, *Elementary school*. Treća kolona *age* sadrži podatke o godinama, a moguće vrednosti su 20, 25, 30, 35, 40.

Nad svim tabelama izvršavamo četiri upita koji su objašnjeni u nastavku.

**Prvi upit** koji izvršavamo je grupisanje prema vrednostima kolona *gender* i *formaleducation* i prebrojavanje sume njihovih vrednosti.

Prikaz SQL naredbi za izvršavanje prvog upita:

```
EXPLAIN ANALYZE
SELECT gender, formaleducation,
COUNT(*) FROM drugi3d4
GROUP BY gender, formaleducation;
```

**Drugi upit** koji izvršavamo je grupisanje kolone *formaleducation* i kolone *gender* za vrednosti *Female* i *Male* i prebrojavanje sume njihovih vrednosti.

Prikaz SQL naredbi za izvršavanje drugog upita:

```
EXPLAIN ANALYZE
SELECT gender, formaleducation,
COUNT(*) FROM drugi3d4
WHERE gender in ('Female', 'Male')
GROUP BY gender, formaleducation;
```

**Treći upit** koji izvršavamo je filtriranje pomoću kriterijuma *age* koji je veći ili jednak od broja 30 i grupisanje kolona *gender* i *formaleducation* po tom kriterijumu.

Prikaz SQL naredbi za izvršavanje trećeg upita:

```
EXPLAIN ANALYZE
SELECT gender, formaleducation,
COUNT(*) FROM drugi3d4
WHERE age >=30
GROUP BY gender, formaleducation;
```

**Četvrti upit** koji izvršavamo je filtriranje pomoću kriterijuma *age* koji je veći ili jednak od broja 30 i kriterijuma kolone *gender* koji ne uzima u obzir vrednost *NA* i grupisanje kolona *gender* i *formaleducation* po tim kriterijumima.

Prikaz SQL naredbi za izvršavanje četvrtog upita:

```
EXPLAIN ANALYZE
SELECT gender, formaleducation,
COUNT(*) FROM drugi3d4
WHERE age >=30 AND gender <>'NA'
GROUP BY gender, formaleducation;
```

Sledeća slika 1 prikazuje rezultat izvršavanja upita tri nad tabelom. Upit se izvršava bez EXPLAIN ANALYZE dela kod SQL naredbi.

|    | gender | formaleducation   | count  |
|----|--------|-------------------|--------|
|    | text   | text              | bigint |
| 1  | NA     | Bachler's degree  | 500    |
| 2  | NA     | Secondary school  | 501    |
| 3  | Male   | Master's degree   | 500    |
| 4  | Male   | Secondary school  | 500    |
| 5  | Female | Master's degree   | 500    |
| 6  | NA     | Elementary school | 499    |
| 7  | Male   | Bachler's degree  | 500    |
| 8  | Female | Bachler's degree  | 500    |
| 9  | Male   | Elementary school | 500    |
| 10 | Female | Elementary school | 501    |
| 11 | Female | Secondary school  | 499    |
| 12 | NA     | Master's degree   | 500    |

Slika 1. Prikaz rezultata trećeg upita kod tabela sa tri kolone

Tabela 1. Izlazna tabela vrednosti za upit tri

| Formaleducation   | Gender |        |    |
|-------------------|--------|--------|----|
|                   | Male   | Female | NA |
| Bachler's degree  | 8%     | 8%     | 9% |
| Master's degree   | 9%     | 8%     | 8% |
| Secondary school  | 8%     | 8%     | 9% |
| Elementary school | 8%     | 9%     | 8% |

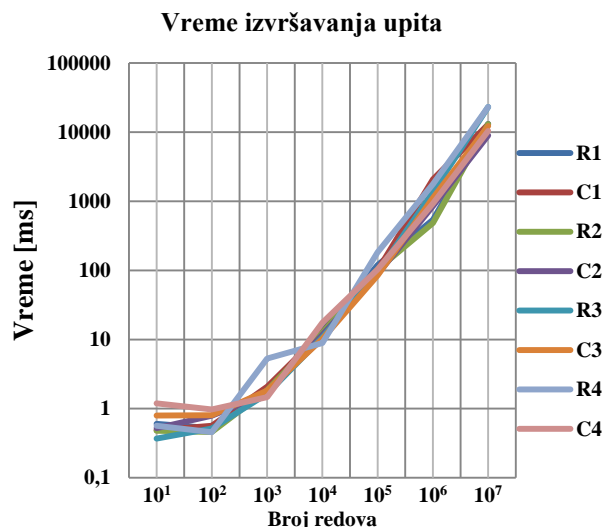
Tabela 1 prikazuje grupisanje podataka za treći upit za kriterijum godina veći ili jednak od broja 30. Grupisanje je izvršeno za grupe *Male*, *Female* i *NA* sa podgrupama *Bachler's degree*, *Master's degree*, *Secondary school* i *Elementary school*. Iz tabele 1 čitaju se izlazni podaci trećeg upita gde grupa *Male* sa podgrupom *Bachler's degree* u tabeli ima udeo 8%, za podgrupu *Master's degree* ima udeo 9%, za podgrupu *Secondary school* ima udeo 8% i za podgrupu *Elementary school* ima udeo 8%. Isto se čitaju grupe *Female* i *NA*.

Slika 2 je grafički prikaz vremena koje je potrebno da bi se prvi, drugi, treći i četvrti upit izvršio nad svakom tabelom.

Tabela 2. Tabelarni prikaz vremena izvršavanja upita nad tabelama sa tri kolone u zavisnosti od br. redova, kod skladištenje po redovima i kolonama

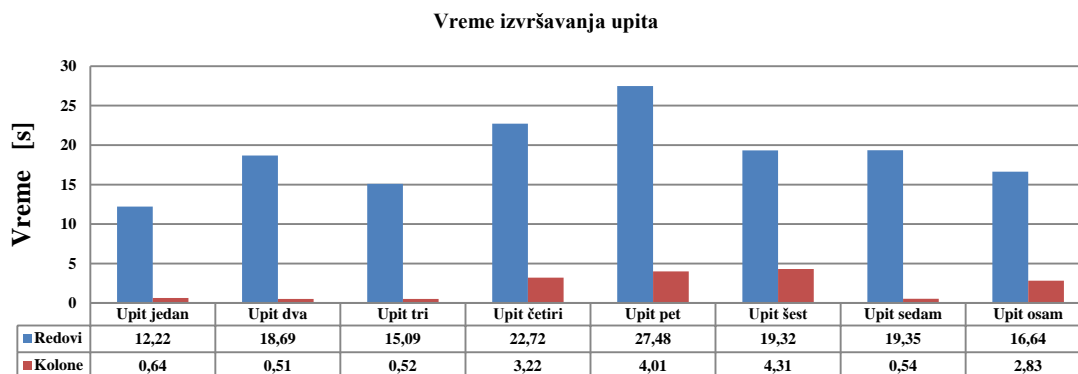
| B.r.            | Prvi upit skladištenje prema redovima | Prvi upit skladištenje prema kolonama | Drugi upit skladištenje prema redovima | Drugi upit skladištenje prema kolonama | Treći upit skladištenje prema redovima | Treći upit skladištenje prema kolonama | Četvrti upit skladištenje prema redovima | Četvrti upit skladištenje prema kolonama |
|-----------------|---------------------------------------|---------------------------------------|--|--|--|--|--|--|
| 10 <sup>1</sup> | 0,602                                 | 0,487                                 | 0,48                                   | 0,521                                  | 0,369                                  | 0,794                                  | 0,564                                    | 1,194                                    |
| 10 <sup>2</sup> | 0,523                                 | 0,564                                 | 0,458                                  | 0,789                                  | 0,515                                  | 0,805                                  | 0,457                                    | 0,976                                    |
| 10 <sup>3</sup> | 1,943                                 | 2,081                                 | 1,859                                  | 1,651                                  | 1,642                                  | 1,827                                  | 5,305                                    | 1,464                                    |
| 10 <sup>4</sup> | 13,119                                | 12,642                                | 12,434                                 | 11,678                                 | 10,79                                  | 10,245                                 | 8,915                                    | 17,518                                   |
| 10 <sup>5</sup> | 119,342                               | 100,131                               | 103,264                                | 87,977                                 | 93,006                                 | 84,313                                 | 182,941                                  | 103,646                                  |
| 10 <sup>6</sup> | 539,006                               | 2087,206                              | 486,11                                 | 853,768                                | 1368,502                               | 1122,277                               | 1770,568                                 | 929,92                                   |
| 10 <sup>7</sup> | 13195,59                              | 12950,28                              | 13139,83                               | 8960,217                               | 23188,06                               | 12313,62                               | 23187,22                                 | 10410,78                                 |

Na grafiku su predstavljene sledeće oznake: R1, R2, R3, R4 – Upit jedan, upit dva, upit tri i upit četiri koji se izvršavaju nad tabelama koje su skladištene prema redovima; C1, C2, C3, C4 – Upit jedan, upit dva, upit tri i upit četiri koji se izvršavaju nad tabelama koje su skladištene prema kolonama.



Slika 2. Grafički prikaz vremena izvršavanja upita nad tabelom sa tri kolone u zavisnosti od br. redova, kod skladištenja prema redovima i kolonama

Rezultati izvršavanja prvog upita pokazuju male razlike u vremenima izvršavanja upita između tabela koje se skladište prema redovima i kolona. Najveća razlika u vremenu izvršavanja upita ogleđa se kod 10<sup>7</sup> redova i malo je bolja kod skladištenja prema kolonama. Rezultati izvršavanja drugog, trećeg i četvrtog upita pokazuju da je razlika između tabela koje su skladištene prema redovima i kolona pokazuju male oscilacije u vremenu izvršavanja upita sa manjim brojem redova. Najveća razlika u vremenu izvršavanja upita ogleđa se kod 10<sup>7</sup> redova i značajno je bolja kod skladištenja prema kolonama. Vremena koja su u tabeli 2 predstavljaju vremena samog izvršavanja upita nad podacima i predstavljena su milisekundama.



Slika 3. Grafički prikaz vremena izvršavanja upita nad tabelama koje skladište podatke po kolonama i redovima.

## 5.2. Eksperimentalni deo dva - Izvršavanje upita na podacima iz ankete

Drugi eksperimentalni deo predstavlja konkretan primer ankete u CSV (engl. *Comma Separated Values*) fajlu. Kreirane su dve tabelle skladištene prema redovima i prema kolonama u koje su učitani podaci iz CSV fajla. Ove tabelle obuhvataju stvarne vrednosti anketa iz 1998. i 1999. godine koje su spojene u jednu tabelu.

Takođe, ovde je testirano izvršavanje osam različitih upita:

**Prvi upit** izvršava prebrojavanje redova u tabeli.

**Drugi upit** izvršava prebrojavanje samo onih redova kod kojih su vrednosti kolone za pregled glasova veće od 5.

**Treći upit** izvršava prebrojavanje redova imena proizvoda, za one proizvode koji počinju velikim slovom "W".

**Četvrti upit** prikazuje podatke imena proizvoda kod kojih je rejting između 3 i 5.

**Peti upit** izvršava prebrojavanje redova imena proizvoda gde kolona kategorija proizvoda odgovara kriterijumu "Science Fiction & Fantasy" i kolona podkategorija proizvoda odgovara kriterijumu "Fantasy".

**Šesti upit** izvršava računanje proseka kolone rang prodaje proizvoda gde je za kolonu kategorija proizvoda kriterijum "Book".

**Sedmi upit** izvršava pronalaženje svih ocena koje je korisnik napravio u "Duno" seriji 1988.

**Osmi upit** izvršava proveravanje veze između naziva knjige, dužine njegovog naslova i njegovih ocena.

Izmerena vremena koja su prikazana slikom 4 predstavljaju vremena samog izvršavanja upita nad tabelama koje skladište podatke prema kolonama i prema redovima.

Slika 3 prikazuje grafik sa kojeg se vidi da kod svih upita koji se izvršavaju nad tabelama koje se skladište prema kolonama treba manje vremena da se obrade podaci u odnosu na tabelle koje se skladište prema redovima.

## 6. ZAKLJUČAK

Kod skladištenja anketnih podataka, gde se za svrhu obrade podataka kreiraju raznovrsni upiti koji u većini slučajeva vrše delimičnu selekciju podataka, pokazalo se na osnovu eksperimenata i rezultata da skladištenje

podataka prema kolonama ubrzava rad. Odnosno, brzina izvršavanja upita je brža. Što je selekcija podataka veća i specifičnost upita komplikovanija u potraživanju podatka, razlika između skladištenja prema redovima i prema kolonama je veća i ide u prilog skladištenju prema kolonama.

Ovim radom omogućeno je dalje istraživanje na ovu temu kao i baza za formiranje istraživačko komercijalnih proizvoda upotrebom tehnologije baza podataka.

## 7. LITERATURA

- [1] Metcalfe A. V., *Statistics in Management Sciences*, Oxford University Press, New York, 2000.
- [2] Rabija Somun-Kapetanović, *Statistika u ekonomiji i menadžmentu*, Ekonomski fakultet u Sarajevu, Sarajevo, 2008
- [3] Raghu Ramakrishnan, Johannes Gehrke, *Database Management Systems*. 2nd Ed, McGraw-Hill Publishing Company, New York, 2000
- [4] Daniel J. Abadi, Samuel R. Madden, Nabil Hachem, *Column-Stores vs RowStores : How Different Are They Really?*, Vancouver, BC, Canada, June, 2008
- [5] Aditi D. Andurkar, *Column-Oriented Database Implementation in PostgreSQL for Improving Performance of Read-Only Queries*, Pune, June, 2012.
- [6] G. Graefe, *Efficient columnar storage in b-trees*, SIGMOD Rec., 2007.
- [7] Patrick O'Neil and Dallan Quass, *Improved query performance with variant indexes*, 1997.
- [8] D. J. Abadi, *Query execution in column-oriented database systems*, MIT PHD Dissertation, PhD Thesis, 2008.

## Kratka biografija:



**Sanja Miljević** rođena je u Virovitici 1989. god. Master rad na Fakultetu tehničkih nauka iz oblasti Elektrotehnike i računarstva – Automatika i upravljanje sistemima odbranila je 2020. god. kontakt: sanja\_mrmos@gmail.com