



BAZE PODATAKA I

2017/18

○ KURSU

- Model 70+30
 - redovno pohadjanje nastave = 4 poena
 - 3 kolokvijuma = 66 poena
 - završni deo ispita = 30 poena
- Uslov za izlazak na završni ispit – više od polovine poena na predispitnim obavezama
- BONUS poene je moguće osvojiti!

LITERATURA

Preporučena, ne i obavezna.

- [1] G. Pavlović-Lažetić, *Osnove relacionih baza podataka*, drugo izdanje, Matematički fakultet, 1999.
- [2] D. Stefanović, *SQL i programiranje u relacionim bazama podataka*, PMF, Kragujevac, 2009.
- [3] B. Lazarević, Z. Marjanović, N. Aničić, S. Babrogić, *Baze podataka*, FON, Beograd, 2003.
- [4] H. Garcia-Molina, J.D.Ulman, J.Widom, *Database Systems: The Complete Book*, Prentice Hall, 2002.
- [5] Hoffer, Jeffrey A., *Modern database management*, Prentice Hall, 2011
- [6] J. Ullman, J. Widom, *A First Course in Database Systems*, Prentice Hall, 2008
- [7] V. Blagojević, *Relacione baze podataka I*, ICNT, Beograd, 2005.

NAJVAŽNIJE TEME

- Ovaj kurs se bavi temama **čuvanja, održavanja i upotrebe kolekcija podataka**.
- Kao uvodni kurs, pre svega, se bavi maistream-om, tj. opšteprihvaćenim konceptom organizacije i održavanja podataka – **relacionim bazama podataka**.
- Ciljevi kursa su da:
 - ovladate osnovnim pojmovima vezanim za:
 - baze podataka uopšte
 - relacione modele
 - sisteme za upravljanje relacionim bazama podataka
 - naučite neke od komandi SQL*-a
 - dobijete sliku o tome šta se SQL*-om može, a šta ne može uraditi
 - naučite da što uspješnije koristite komande SQL-a i time osbosobite za praktičan rad

*SQL - jezik za manipulisanje podacima



INFORMACIJA, PODATAK, METAPODATAK



PODATAK

- Podatak ↔ Informacija

Baker, Kenneth D. 324917628
Doyle, Joan E. 476193248
Finkle, Clive R. 548429344
Lewis, John C. 551742186
McFerran, Debra R. 409723145

Podatak

Zabeležena reprezentacija objekata i događaja koji su od značaja za korisnika.

Class Roster			
Course:	MGT 500	Semester: Spring 2010	
	Business Policy		
Section:	2		
Name	ID	Major	GPA
Baker, Kenneth D.	324917628	MGT	2.9
Doyle, Joan E.	476193248	MKT	3.4
Finkle, Clive R.	548429344	PRM	2.8
Lewis, John C.	551742186	MGT	3.7
McFerran, Debra R.	409723145	IS	2.9
Sisneros, Michael	392416582	ACCT	3.3

Informacija

Podatak obrađen na način da donosi novo znanje osobi koja ga koristi.

Prevođenje podataka u informaciju

INFORMACIJA

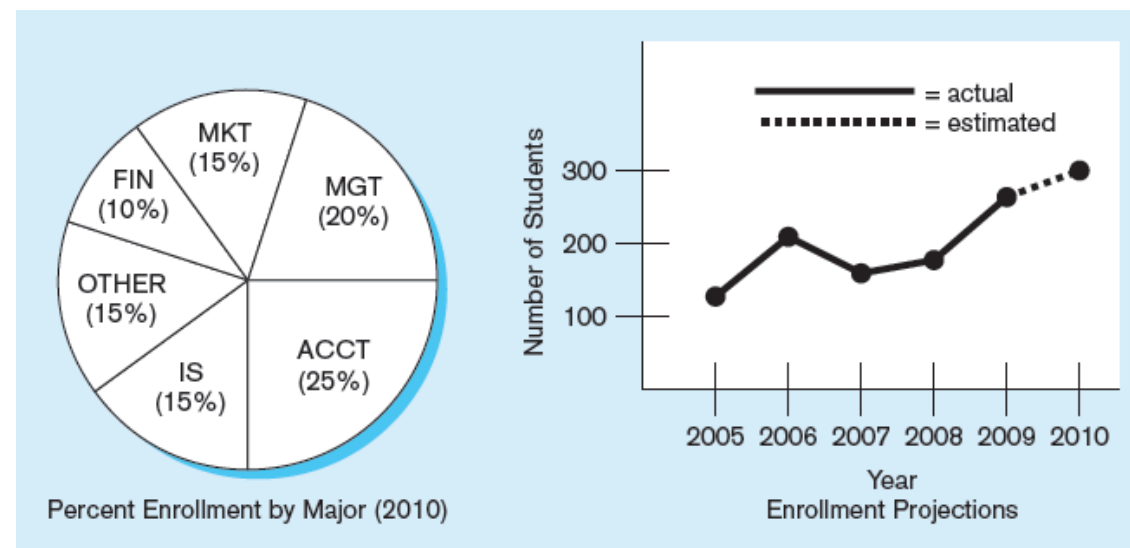
Drugi način prevođenja podataka u informaciju – sumiranje ili druge vrste procesiranja koje omogućavaju dodatnu interpretaciju.

Podatak

- je kodirana predstava o nekoj činjenici iz realnog sveta;
- služi za tehničko uobličavanje informacije kako bi se ona mogla sačuvati i preneti.

Informacija

- je protumačen podatak o pojavi koju podatak pokazuje.



METAPODACI

Osnovni mehanizam davanja konteksta podacima su METAPODACI.

Metapodatak – podatak koji opisuje karakteristike i kontekst podataka razumljivih korisniku.

“data about data”

Data Item			Metadata			
Name	Type	Length	Min	Max	Description	Source
Course	Alphanumeric	30			Course ID and name	Academic Unit
Section	Integer	1	1	9	Section number	Registrar
Semester	Alphanumeric	10			Semester and year	Registrar
Name	Alphanumeric	30			Student name	Student IS
ID	Integer	9			Student ID (SSN)	Student IS
Major	Alphanumeric	4			Student major	Student IS
GPA	Decimal	3	0.0	4.0	Student grade point average	Academic Unit



UVOD



GDE SU?

Hi, Ana | Sign Out | Help

Make Y! My Homepage

YAHOO! SPORTS

Search

Search Web

Home NFL MLB NBA NHL NCAAF NCAAB NASCAR Golf UFC Boxing Soccer Tennis Action Sports More ThePostGame Shop Fantasy

Tennis Home Rankings Matches Schedule Players Video Photos Blog Tickets

Sports Search

TRENDING NOW: Reggie Wayne Kobe Russell Martin Evgeni Malkin Bronson Arroyo

2012 ATP Points Standings

[Men's Singles](#) | [Women's Singles](#) | [Men's Doubles](#) | [Women's Doubles](#)

Top 100					Oct 8
Rank	Last	Name	Country	Ranking Pts	
1	-	Roger Federer	Switzerland	11805	
2	-	Novak Djokovic	Serbia	10970	
3	-	Andy Murray	Scotland	8090	
4	-	Rafael Nadal	Spain	7085	
5	-	David Ferrer	Spain	5960	
6	-	Jo-Wilfried Tsonga	France	4640	
7	-	Tomas Berdych	Czech Republic	4570	
8	-	Juan Martin del Potro	Argentina	3670	
9	-	Janko Tipsarevic	Serbia	3185	
10	-	Juan Monaco	Argentina	2775	
11	-	Nicolas Pietrangeli	Spain	2515	
12	-	John Isner	United States	2475	
13	-	Richard Gasquet	France	2370	
14	-	Milos Raonic	Canada	2335	
15	-	Kei Nishikori	Japan	2315	
16	-	Marin Cilic	Croatia	2200	
17	-	Stanislas Wawrinka	Switzerland	1955	
18	-	Gilles Simon	France	1860	
19	-	Philipp Kohlschreiber	Germany	1855	
20	-	Alexandr Dolgopolov	Ukraine	1770	
21	-	Tommy Haas	Germany	1553	
22	-	Sam Querrey	United States	1530	
23	-	Fernando Verdasco	Spain	1480	
24	-	Marcel Granollers	Spain	1465	
25	-	Florian Mayer	Germany	1430	
26	-	Andreas Seppi	Italy	1415	
27	-	Andy Roddick	United States	1375	

ADVERTISEMENT

Replay

Personally Yours

YAHOO! MAIL

Get It Now

GOOGLE, SKYPE, ANDROID KORISTE

Google koristi sopstveni backend sistem nazvan BigTable.
Open source implementacija BigTable je Apache-ov **Hadoop**.

Skype - PostgresSQL

C:\Users\Username\AppData\Roaming\Skype\

Android aplikacije - **SQLite**

MOTIVACIJA

- Skoro svakog sata direktno ili indirektno koristite neku vrstu baze podataka, npr. kada
 - bilo kom pretraživaču zadate ključne reči,
 - sa bankomata podižete novac sa računa,
 - koristite bilo koju društvenu mrežu,
 - kupujete u bilo kojoj prodavnici koja koristi fiskalni račun...
- Baza podataka se koristi za smeštanje, manipulisanje i upotrebu podataka u svim tipovima organizacija, uključujući poslovne, vladine, javne itd.

ŠTA JE BAZA PODATAKA?

Uopštena definicija

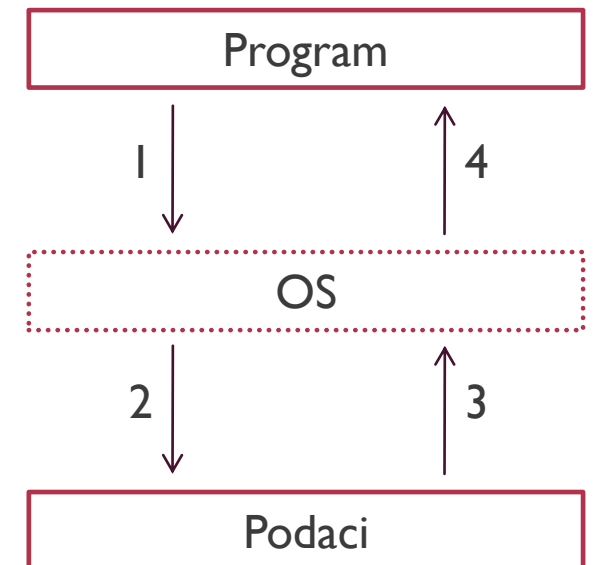
Baza podataka je organizovana kolekcija logički povezanih podataka;

Savremena definicija

Savremena definicija se vezuje za računarski podržano čuvanje i obradu podataka.

RAČUNARSKI PODRŽANO ČUVANJE I OBRADA

- Da bi računarski podržano čuvanje i obrada podataka bili mogući prvi korak bio je razvoj **fajl sistema**.
- Prve aplikacije za obradu i čuvanje podataka su bile izgrađene **direktno nad fajl sistemom**, podaci su bili čuvani u datotekama.



FAJL SISTEM – NEDOSTACI

- Ponavljanje istih podataka uz različite aplikacije.
- Nekonzistentnost podataka.
- Programi za obradu podataka zavise od načina struktuiranja podataka (sekvencijalna, direktna, indeksirana datoteka itd).
- Obrada podataka je skupa, s obzirom na nekonzistentnost i zavisnost programa od organizacije podataka.
- Korišćenje istih podataka od strane većeg broja korisnika je otežano.
 - Na primer, istovremeni pokušaj dva ili više korisnika da promene sadržaj jedne iste datoteke završiće se, u mnogim slučajevima, pamćenjem promena samo onog korisnika koji je poslednji završio rad sa datotekom.
- Neadekvatna realizacija oporavka od pada sistema.
 - U slučaju pada sistema aktivni poslovi nemaju mogućnost poništavanja svojih delimičnih izvršenja (ako su ona deo jedinstvene logičke celine), a često, po uspostavljanju sistema, ni evidenciju o svom delimičnom izvršenju.

BAZE PODATKA

- Pojam baza podataka pojavio se krajem 60-tih godina i označavao je:
 - skup međusobno povezanih podataka koji se čuvaju zajedno, i
 - među kojima ima samo onoliko ponavljanja koliko je neophodno za njihovo optimalno korišćenje pri višekorisničkom radu.
- Podaci se:
 - pamte tako da budu nezavisni od programa koji ih koriste, i
 - strukturiraju se tako da je omogućen porast baze.
- Za efikasan rad sa podacima i održavanje konzistentnog stanja baze koristi se **sistem za upravljanje bazama podataka (SUBP – DBMS, Data Base Management System)**.

Baze podataka + SUBP
=
SISTEM BAZA PODATAKA

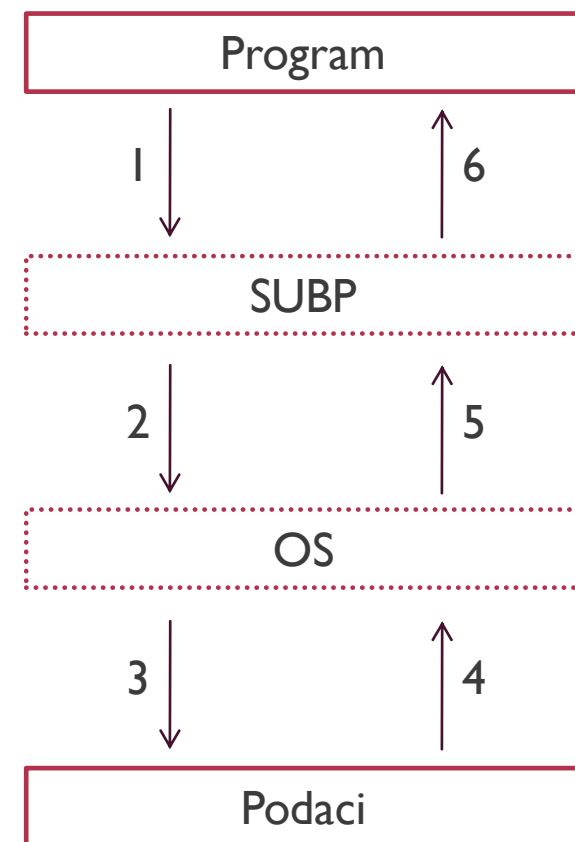
ČEMU SLUŽI SUBP?

SUBP se koristi za manipulisanje velikim skupovima podataka.

- Održava i manipuliše i podacima i metapodacima.

Dodatno, SUBP obezbeđuje:

- Oporavak od pada sistema,
- Konkurentni pristup,
- Brzi razvoj aplikacija,
- Integritet i sigurnost podataka.

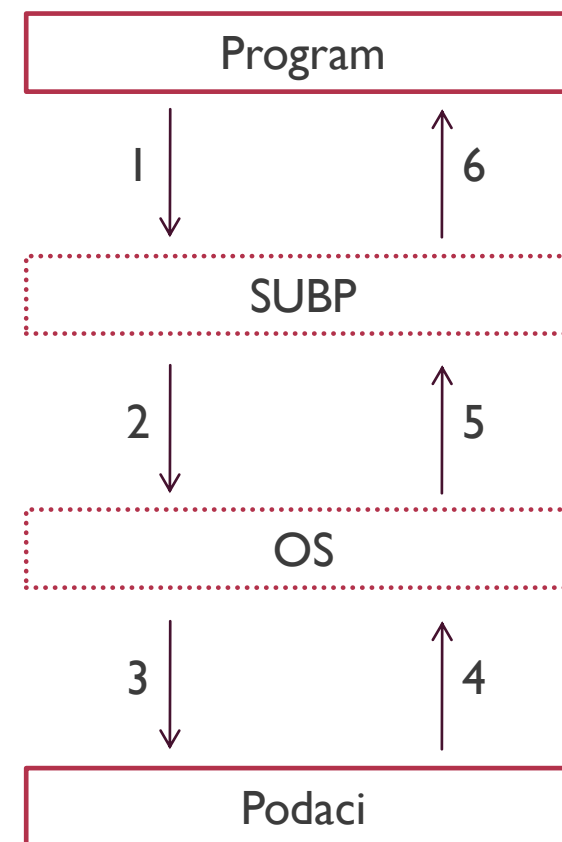


PRIMERI

- MS SQL server
- Oracle
- MySQL
- Postgres
- NO SQL baze
 - MongoDB, CouchDB
 - Cassandra, Hbase

BAZE PODATKA

- Kod ovakvog pristupa podacima, slog koji je definisan u programu uopšte ne mora odgovarati strukturi sloga u nekoj datoteci, nego može biti izveden iz stvarnog sloga mehanizmom preslikavanja podataka.

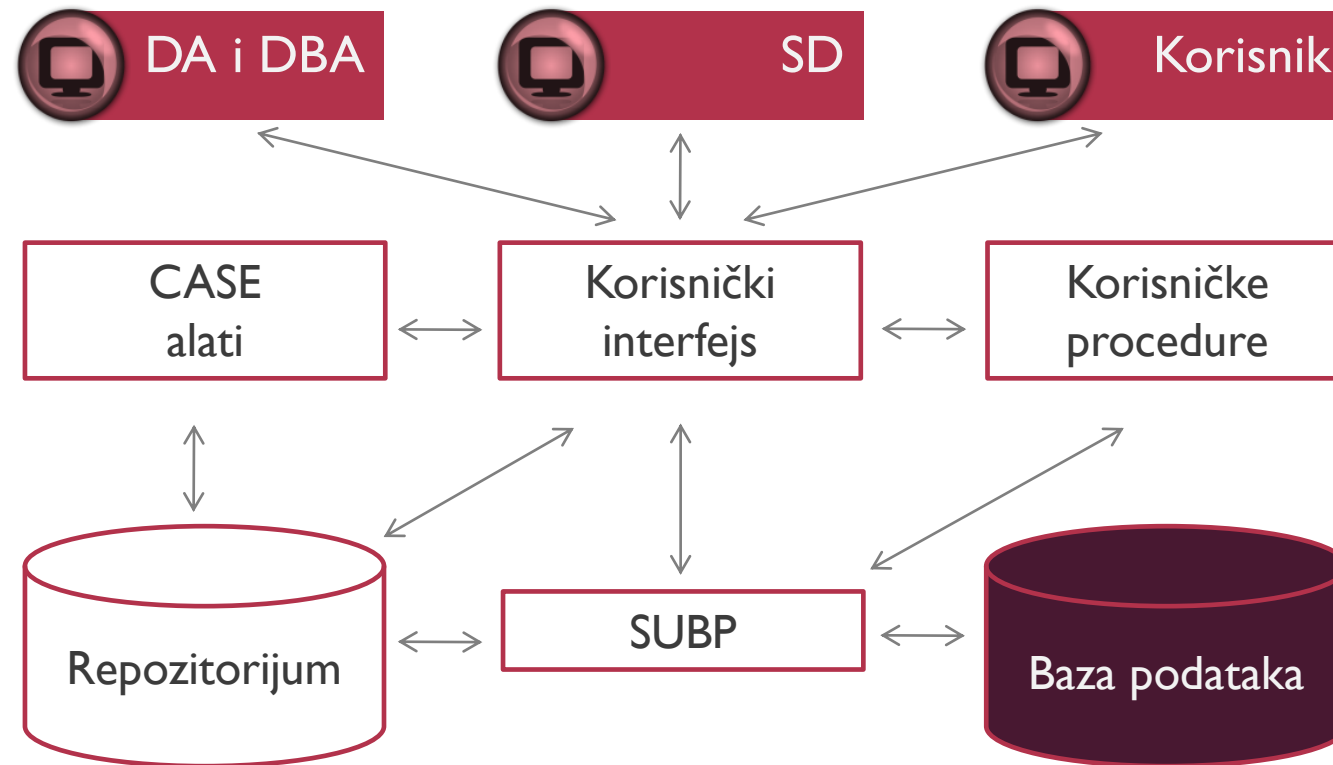




OKRUŽENJE BAZE PODATAKA I SUBP ARHITEKTURA



OKRUŽENJE



OKRUŽENJE

- **DA i DBA**
 - **Data administrators** – administratori podataka su osobe zadužene za upravljanje podacima (informacionim resursima).
 - **Database administrators** – administratori baze podataka su osobe zadužene za fizički dizajn baze i za regulisanje tehničkih pitanja vezanih za okruženje baze podataka.
- **SD**
 - **System developers (sistem analitičari i programeri)** su osobe koje se bave dizajnom i programiranjem aplikacija i baza podataka koje te aplikacije koriste.
- **Korisnici**
 - **‘Naivni’ korisnici (End/naive users)** su osobe koje su direktni korisnici podataka koji se čuvaju u bazi podataka.

OKRUŽENJE

- U opštem slučaju **CASE** alati su automatizovani alati namenjeni dizajnu modela podataka, baze podataka, različitih softvera, ... U trenutnom kontekstu, pre svega, se misli na alate za definisanje logičkog modela podataka.

... ovi alati često imaju i mogućnost automatskog generisanja SQL koda kojim se kreiraju objekti baze podataka.

- **Korisničke procedure (application programs)** su programi/procedure pisane za konkretnu korisničku bazu podataka i služe kreiranju i održavanju njenih objekata.
- **Korisnički interfejs** je interfejs preko kojeg sve vrste korisnika (DA, DBA, SD i krajnji korisnik) komuniciraju sa ostalim komponentama sistema.

OKRUŽENJE

- **Repozitorijum** je centralizovana kolekcija meta podataka.
- **Baza podataka** je organizovana kolekcija logički povezanih podataka, obično, namenjena različitim grupama korisnika.



ZADACI SUBPA

- **Integrisanost i kontrolisana redundansa**

krajnji cilj integrisanosti je minimalna redundansa (višestruko ponavljanje) podataka;

- **Organizacija prema potrebama korisnika**

podrazumeva mogućnost definisanja izvedenih slogova sa podacima;

- **Sigurnost**

podrazumeva efikasnu kontrolu pristupa podacima, u smislu ko može da pristupi bazi podataka, kojim podacima i šta može da radi sa tim podacima;

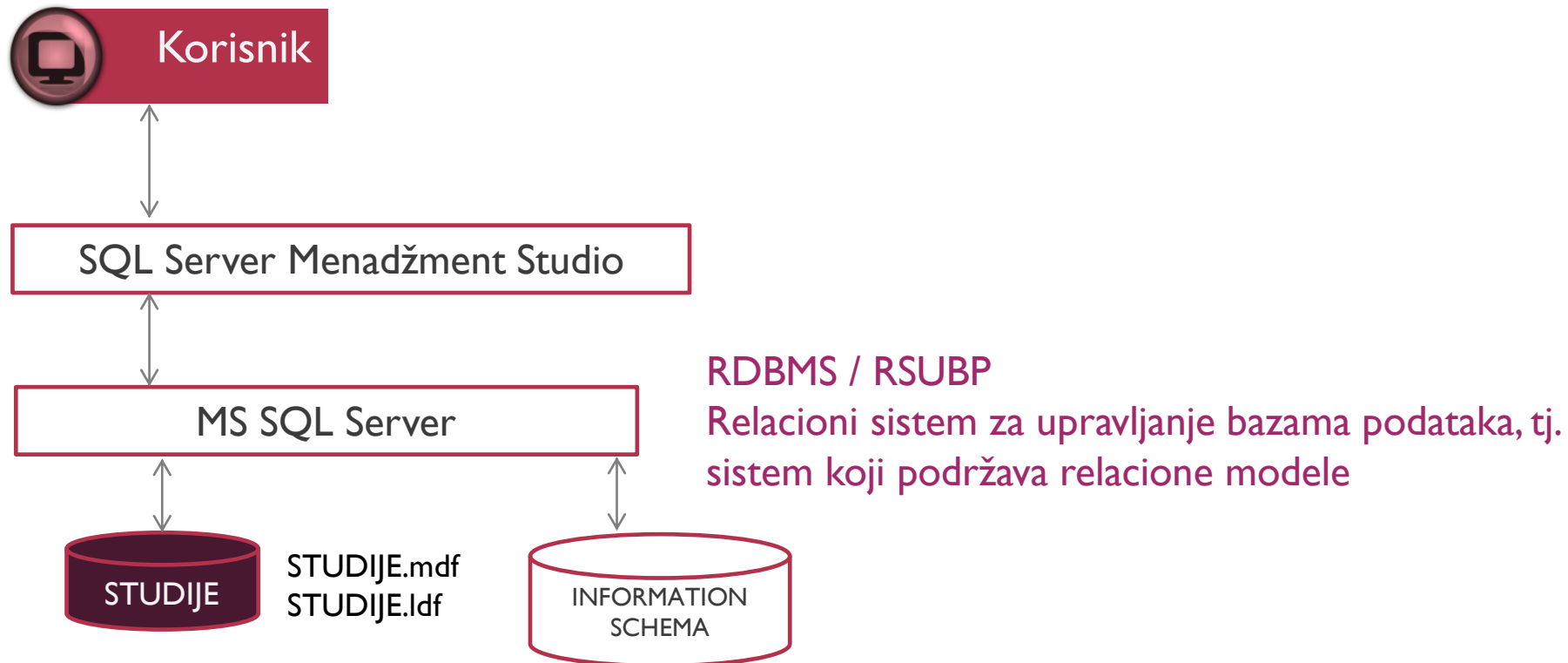
- **Konkurentnost**

podrazumeva mogućnost sinhronizovanog rada više korisnika istovremeno;

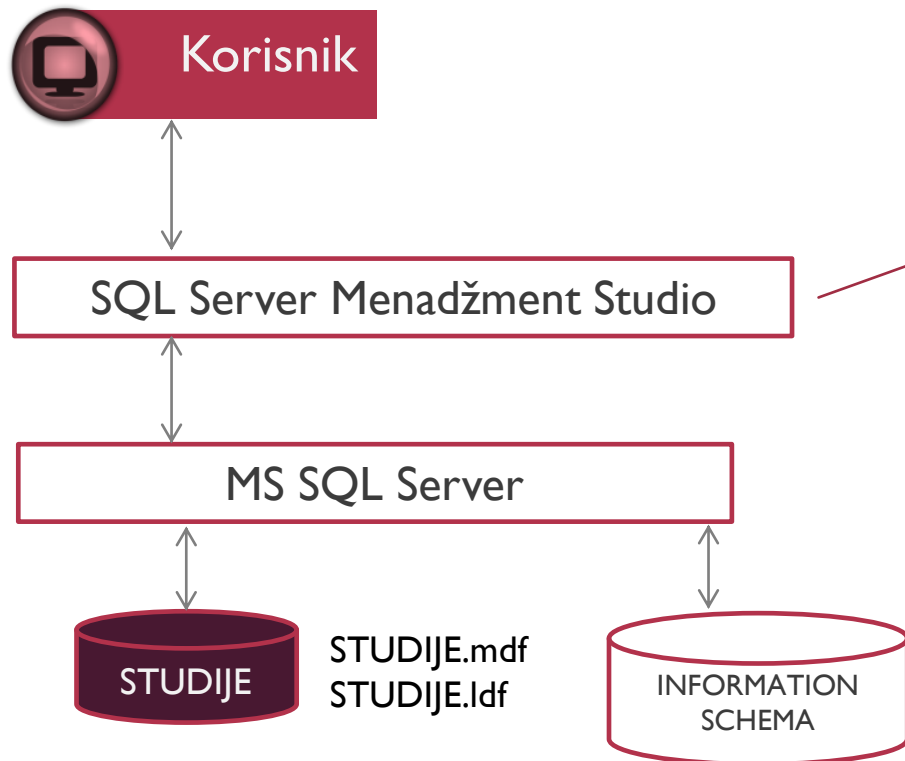
- **Integritet**

podrazumeva automatski oporavak od nasilnih prekida u radu koji dovode do tzv. nekonzistentnih stanja usled delimično izvršenih ažuriranja (unos, izmene ili brisanja) podataka;

OKRUŽENJE U KOM ĆEMO MI RADITI



OKRUŽENJE U KOM ĆEMO MI RADITI



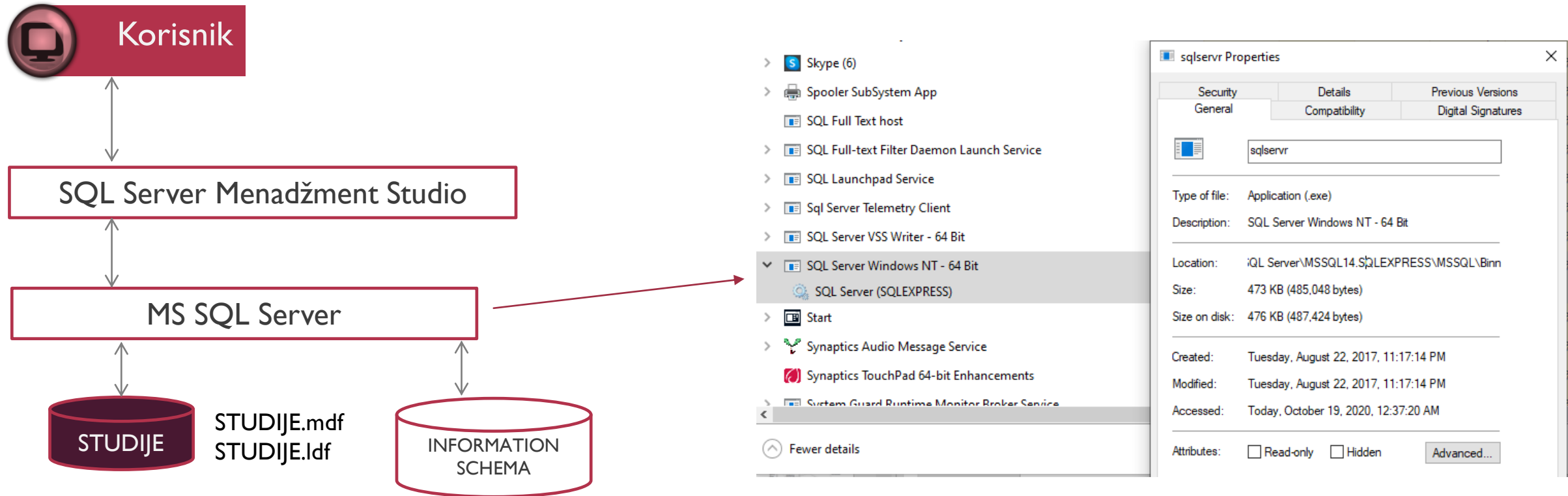
The screenshot shows the Microsoft SQL Server Management Studio interface. The **Object Explorer** on the left displays the server structure for **(localdb)\test2**, including a database named **STUDIJE** with various tables. The **Query Editor** on the right contains the following SQL query:

```
SELECT *  
FROM [dbo].[Studenti]
```

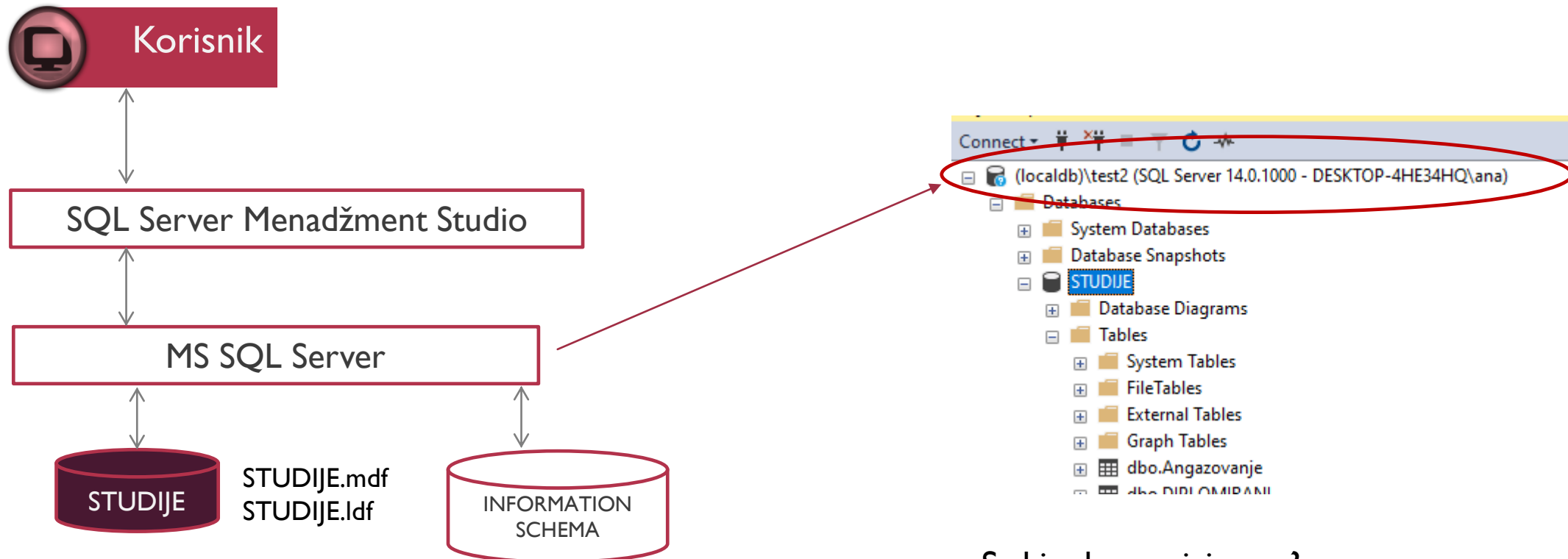
The **Results** pane at the bottom shows the output of the query:

	Indeks	Upisan	Ime	Mesto	Datr	Smer
1	1	2000	Ana	Kragujevac	1982-05-21 00:00:00.000	2
2	1	2002	Nenad	Kragujevac	1984-06-15 00:00:00.000	1
3	2	2001	Sanja	Jagodina	1983-04-20 00:00:00.000	3
4	2	2000	Vica	Kragujevac	1984-06-15 00:00:00.000	4

OKRUŽENJE U KOM ĆEMO MI RADITI



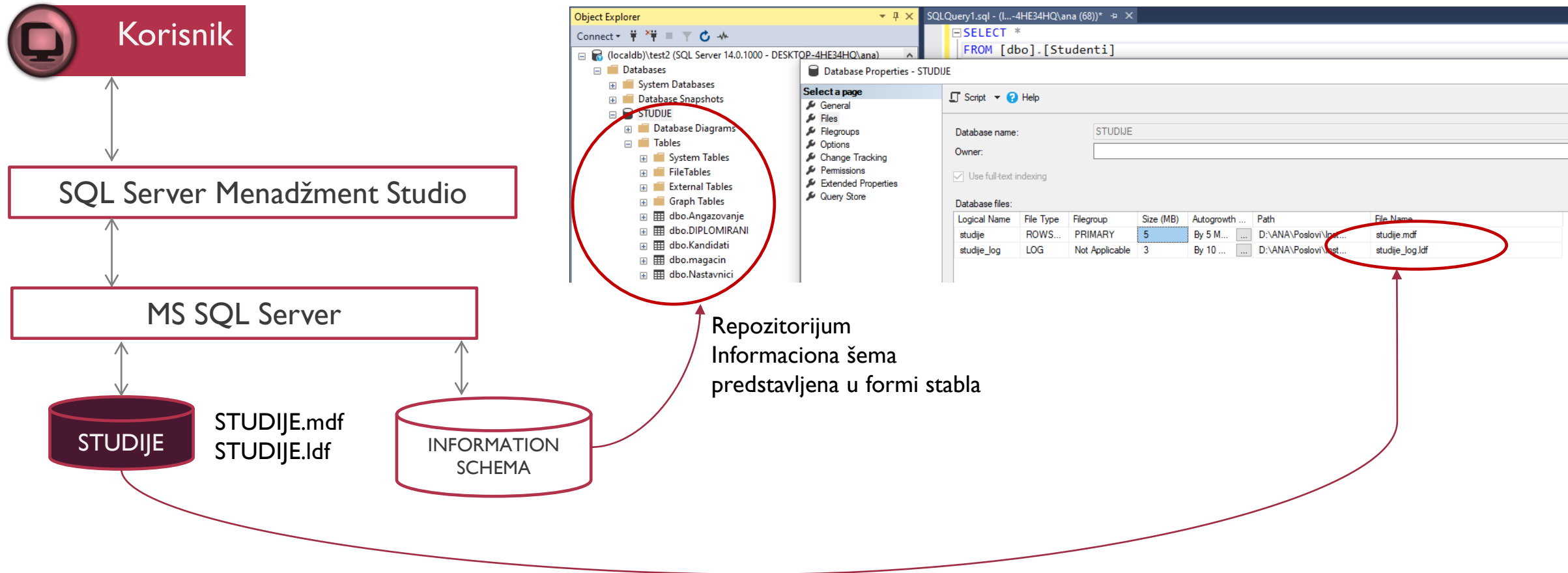
OKRUŽENJE U KOM ĆEMO MI RADITI



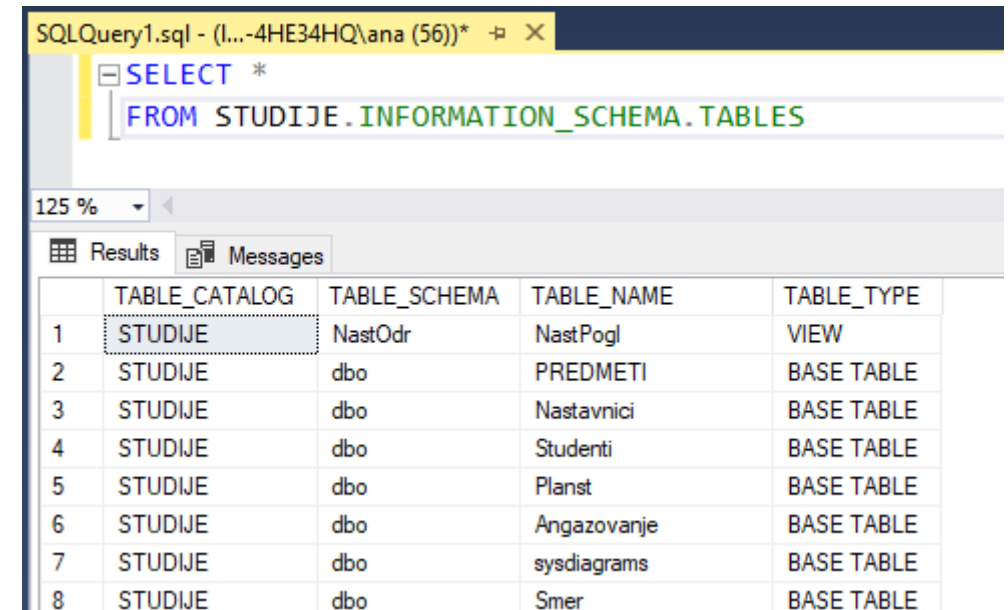
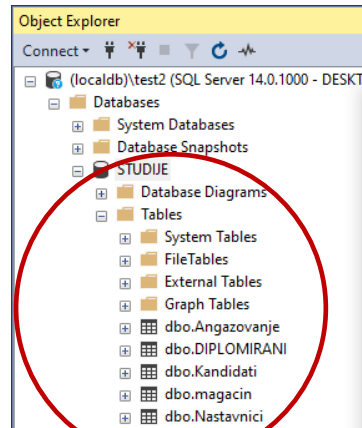
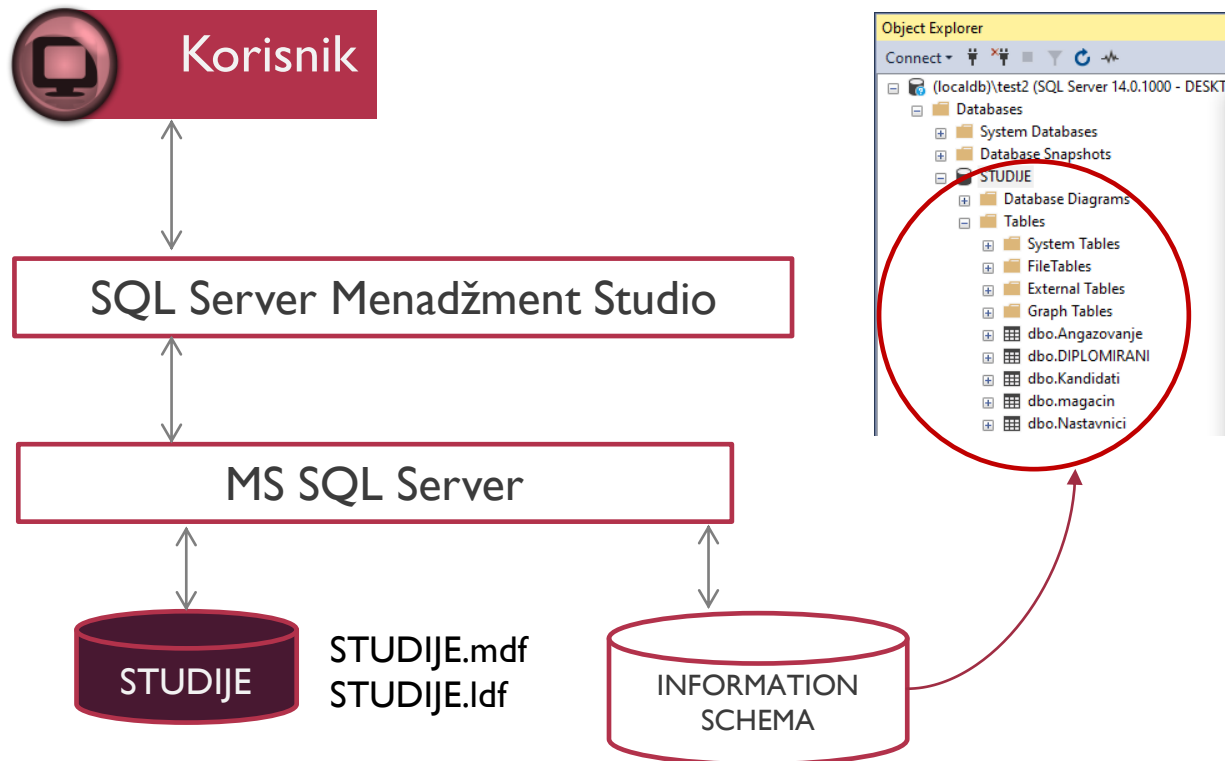
Sa kim komuniciramo?

Instanca SQL servera čiji su servisi pokrenuti i dostupni

OKRUŽENJE U KOM ĆEMO MI RADITI



OKRUŽENJE U KOM ĆEMO MI RADITI

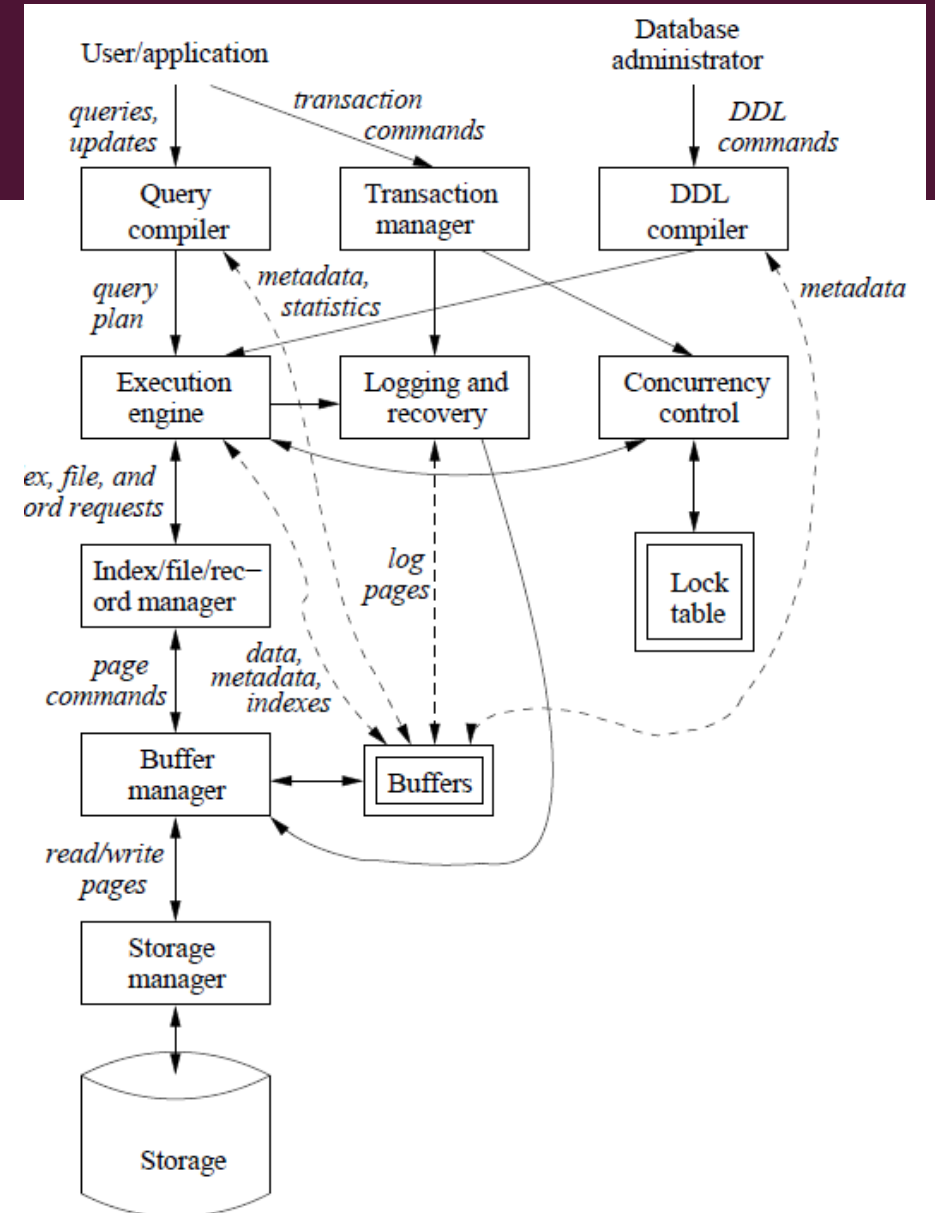


Informaciona šema (baza podataka o bazi podataka) se čuva u formi u kojoj se čuvaju i podaci korisnika (u tabelama).

Uz odgovarajuću dozvolu, pretraživa je na isti način na koji se pretražuju i korisnikovi podaci u bazi podataka.

SUBP ARHITEKTURA

- Prevođenje i izvršavanje upita
- Upravljanje transakcijama i kontrola konkurentnosti
- Upravljanje memorijom
- Vođenje logova i obazbeđivanje oporavka
- Upravljanje internom memorijom
- Organizacija podataka unutar fajlova



PROCESIRANJE UPITA

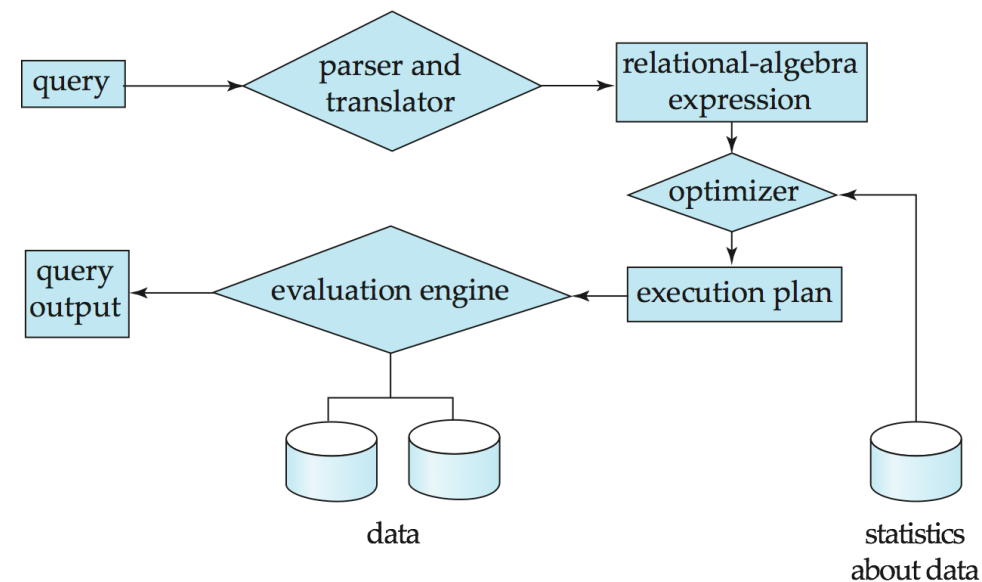
■ Parsiranje

- Parsira upit
- Izvršava provjere na osnovu podataka u katalogu
 - korektnost (da li postoje objekti navedeni u upitu)
 - autorizovanost
 - integritetna ograničenja

■ Optimizacija

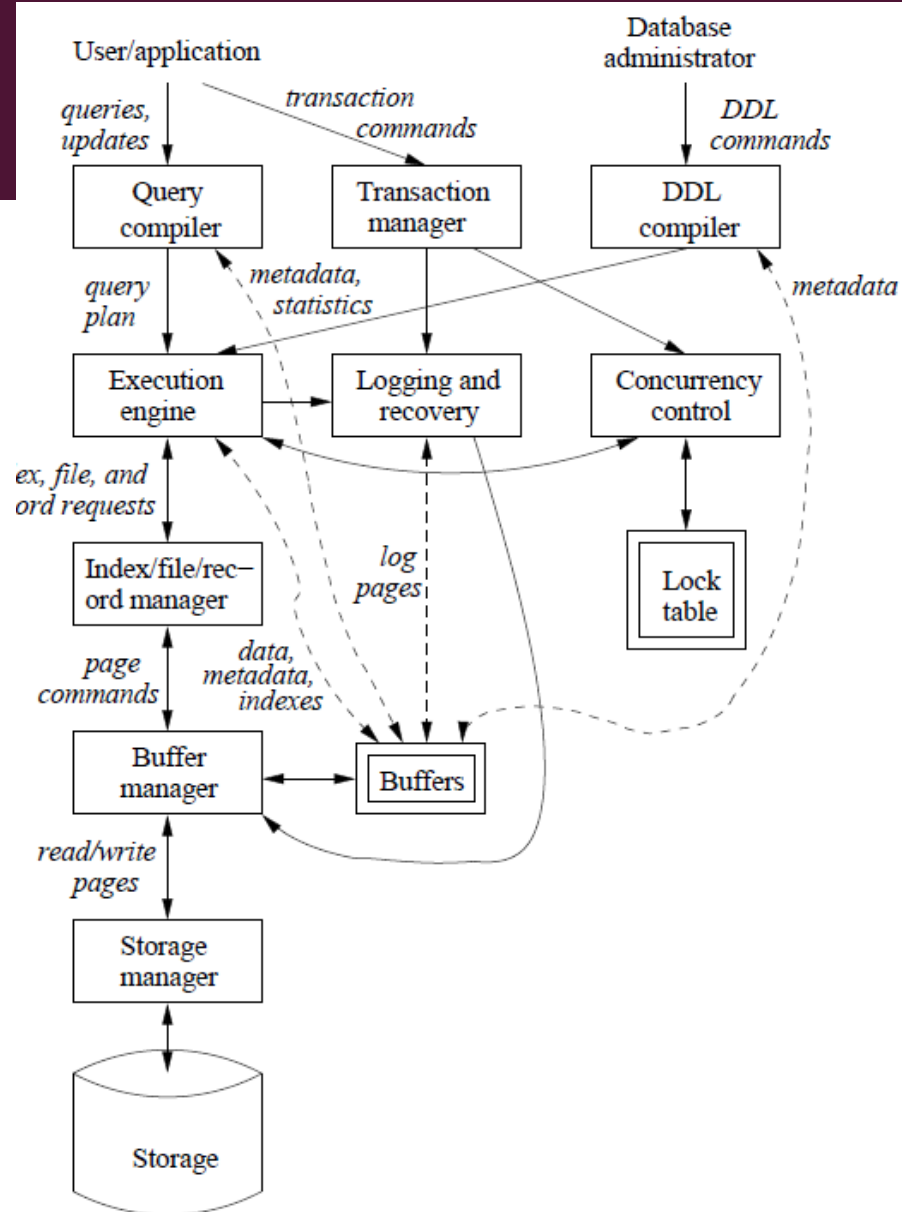
- Kreiranje logičkih i fizičkih planova
- Pronalaženje najefikasnijeg plana za izvršavanje upita

■ Izvršavanje



UPRAVLJANJE MEMORIOM

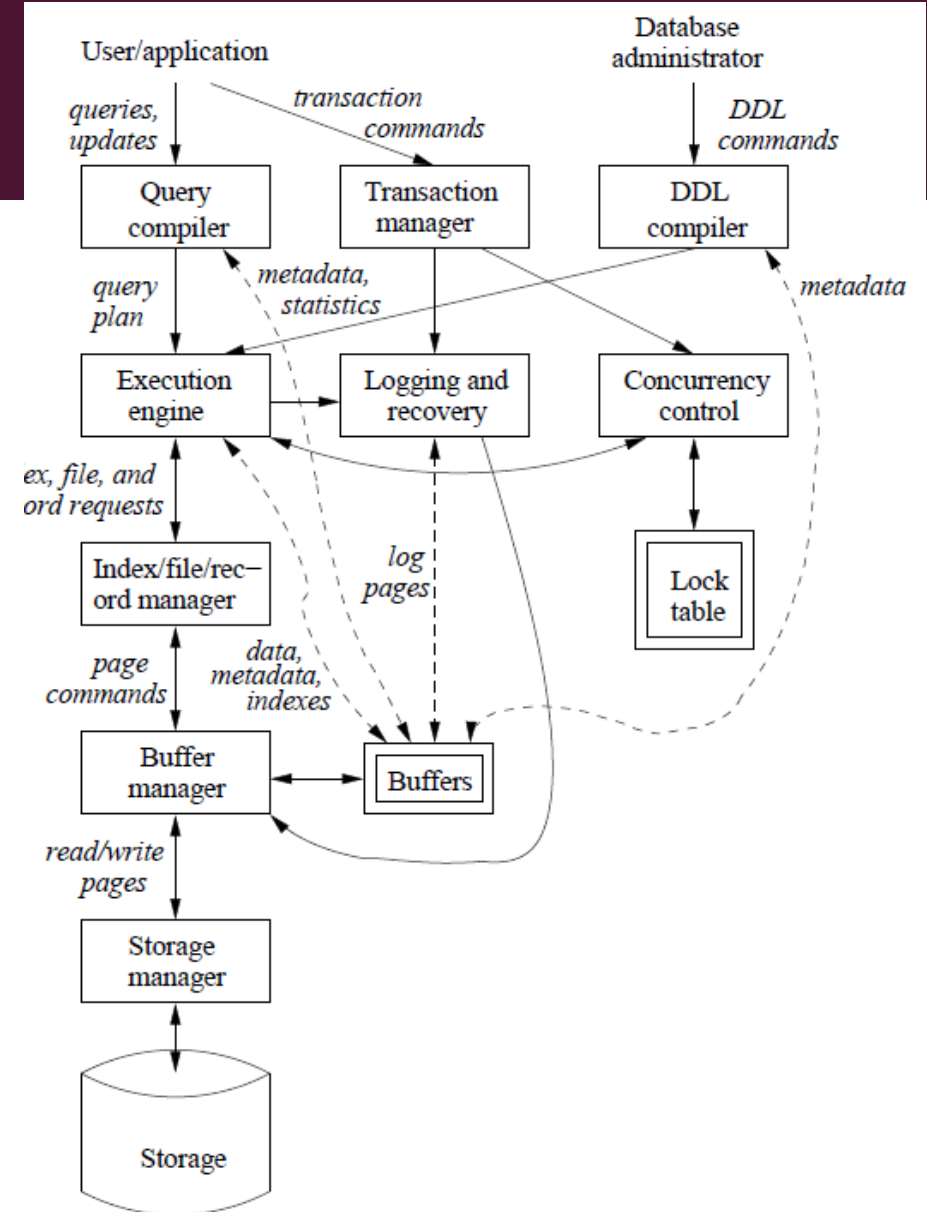
- **Bafer menadžer**
 - Odgovoran za upravljanje memorijom dodeljenom SUBP-u. Podatke pročitane sa diska smešta u interne bafere.
 - Sve komponente koje zahtevaju podatke iste dobijaju obraćajući se bafer menadžeru.



UPRAVLJANJE SKLADIŠTEM

■ Menadžer skladišta

- Vodi evidenciju o fajlovima na disku.
- Sa njim komunicira bafer menadžer, a on komunicira sa OS-om.



PREDNOSTI SUBP-A

- Jednostavni modeli podataka – korisnik nije svestan fizičke organizacije podataka, podatke organizuje prema modelu koji SUBP podržava (ove godine relacioni model)
- Jednostavan jezik za komunikaciju / upravljanje podacima
- Optimizovano izvršavanje upita
- Obezbeđen višekorisnički rad i kontrolu pristupa
- Obezbeđena podrška za paralelno izvršavanje više poslova istovremeno
- Obezbeđen oporavak od pada sistema