

Sistemi baza podataka

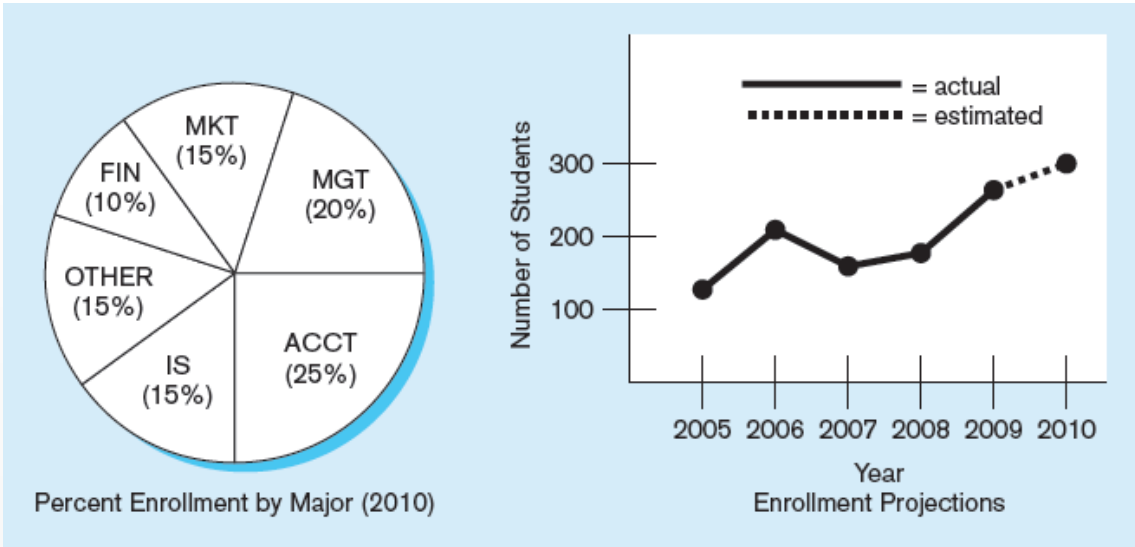
Preslišavanje

Podatak/Informacija

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

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

Informacija - protumačen podatak o pojavi koju podatak pokazuje.



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

Metapodaci - “data about data”

Osnovni mehanizam davanja konteksta podacima su METAPODACI.

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

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

Baza podataka

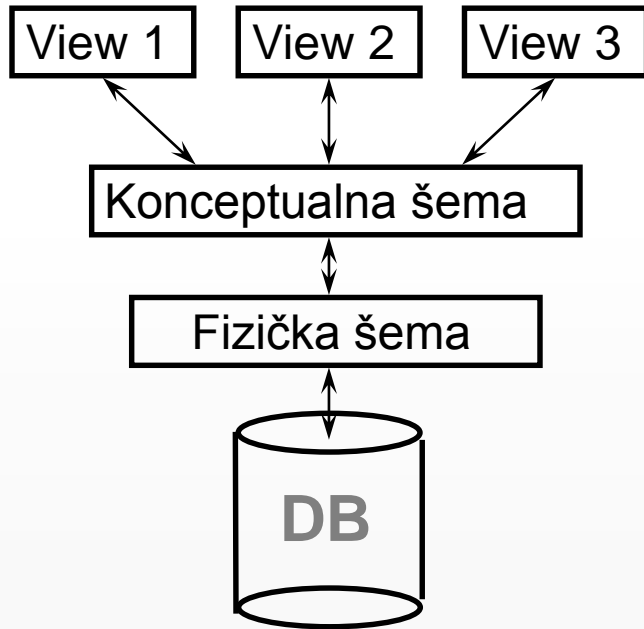
- **Baza podataka je organizovana kolekcija logički povezanih podataka.**
- Pojam baza podataka pojavio se krajem 60-tih godina i označavao je:
 - skup međusobno povezanih podataka koji se čuvaju zajedno, i
 - među kojima ima samo onoliko ponavljanja koliko je neophodno za njihovo optimalno korišćenje 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).**

Baza podataka + SUBP

=

SISTEM BAZA PODATAKA.

Baza podataka i apstrakcija



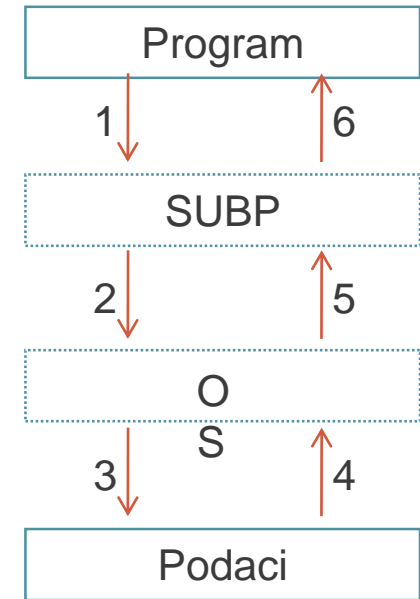
ANSI/SPARC arhitektura modela

Pogledi (Views) - sadrže podskupove konceptualnog nivoa

Konceptualni nivo (Schema) - logička struktura; sadrži opise podaka i njihovih odnosa

```
type instructor = record
  ID : string;
  name : string;
  dept_name : string;
end;
```

Fizička šema - opisuje kako su slogovi upisani na fizički medijum.



Instance i šeme

- **Database Schema** – je logička struktura baze (konceptualni nivo ANSI arhitekture)
- **Database Instance** – konkretni sadržaj baze u određenom vremenskom trenutku (analogno vrednosti promenljive)

(MS SQL Server – sistemske tabele)

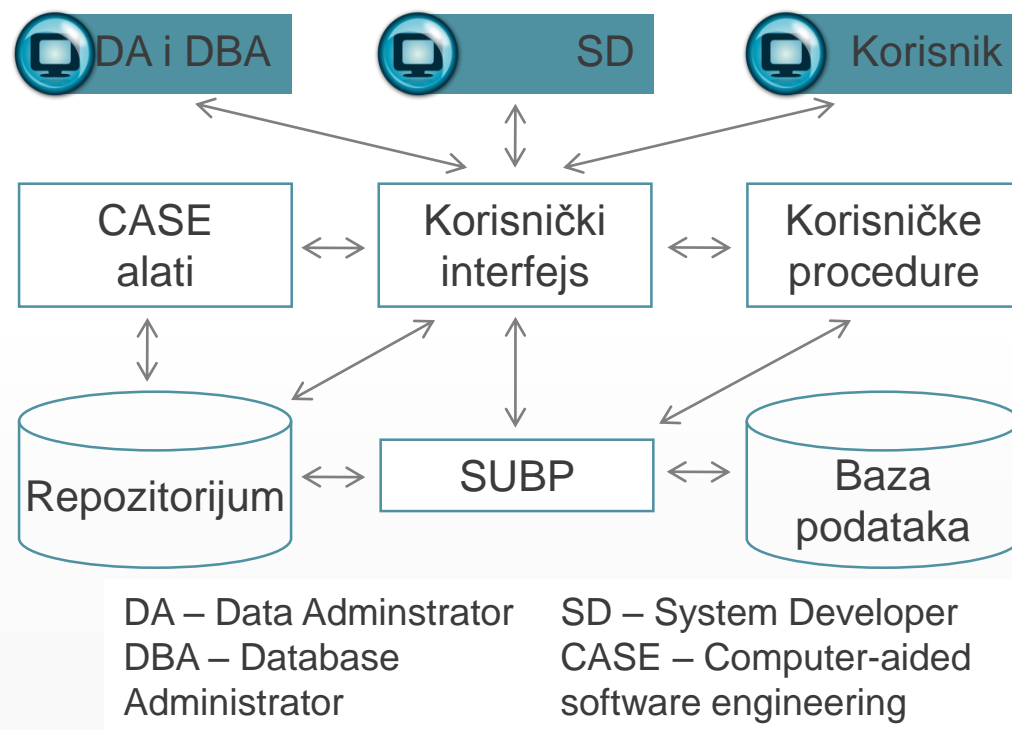
```
SELECT TABLE_NAME,*  
FROM INFORMATION_SCHEMA.TABLES  
WHERE TABLE_TYPE = 'BASE TABLE'
```

```
SELECT TABLE_NAME,*  
FROM INFORMATION_SCHEMA.COLUMNS  
WHERE TABLE_NAME = 'nastavnik'
```

Zadaci sistema baza podataka

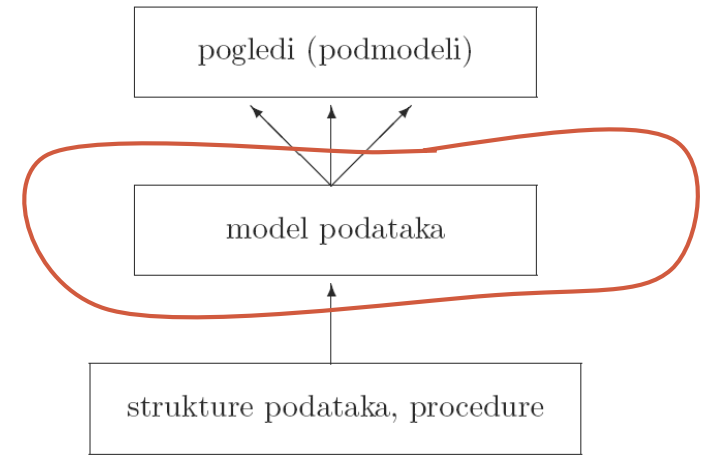
- **Nezavisnost podataka**
 - **Integrisanost i kontrolisana redundansa**
krajnji cilj integrisanosti je minimalna redundansa (višestruko ponavljanje) podataka;
 - **Organizacija prema potrebama korisnika**
podrazumeva mogućnost definisanja izvedenih slogova sa podacima;
 - **Sigurnost**
podrazumeva 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, izmena ili brisanje) podataka;
-

Okruženje baza podataka



Modeli podataka

- Kolekcija alata za opisivanje
 - Podataka
 - Relacija između podataka
 - Semantike podataka
 - Operacija
 - Ograničenja
- Najzastupljeni model podataka jeste Relacioni model
- SUBP implementira određenu vrstu modela.



- Entity-Relationship data model (mainly for database design)
- Object-based data models (Object-oriented and Object-relational)
- Semistructured data model (XML)
- Other older models:
 - Network model
 - Hierarchical model

Relacioni model i implementacija od strane SUBP-a

- Relacija -> tabela
- Ključevi
- Integritetna pravila
- Relaciona algebra/račun -> SQL

DML – pristup i manipulacija podacima

DDL - DDL compiler generiše šablone koji se smeštaju u rečnik podataka (**data dictionary**)

Data dictionary contains metadata

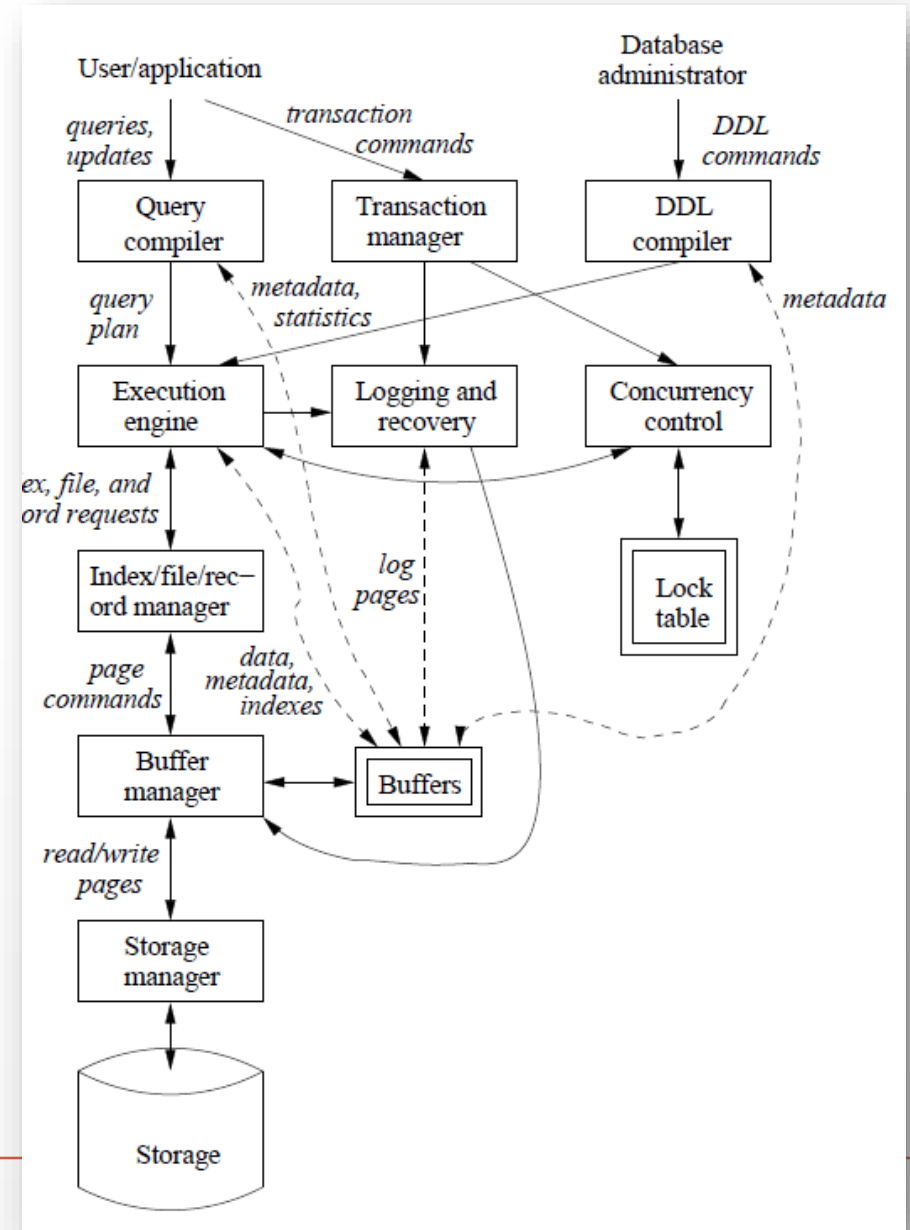
- Database schema
 - Integrity constraints
 - Primary key
 - Referential integrity
 - Authorization
-

Sistemi baza podataka

Buffer Management

Overview of DBMS

- Pravougaonici – komponente sistema
- Dupli pravougaonik – struktura podataka u memoriji
- Puna strelica - control and data flow
- Isprekidana strelica - data flow only

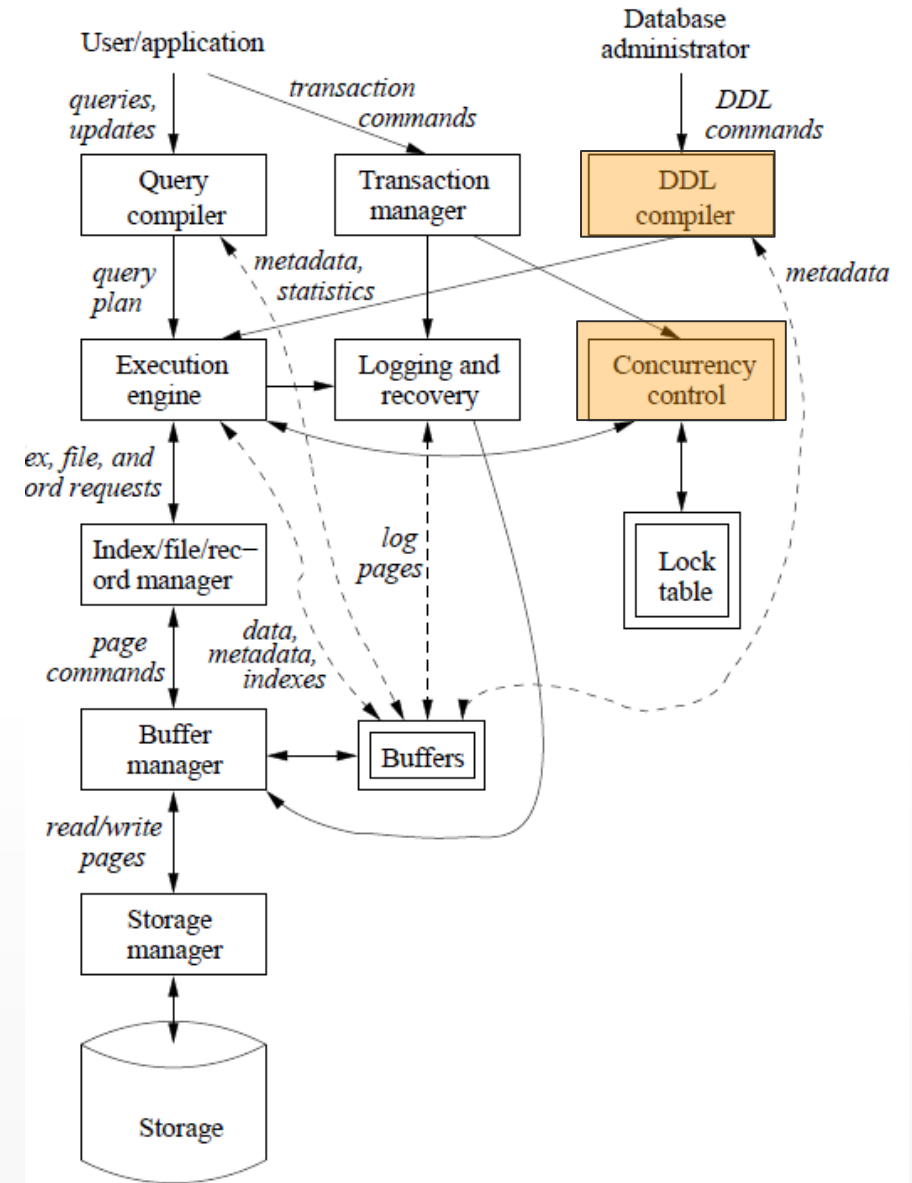


Database Definition Commands - DDL

- Database administrator (DBA) – odgovoran za strukturu/šemu baze

Example: A university DBA decides on a table with student, course, grade columns. Grade can only be (A, B, C, D)

- DBA koristi data-definition language (DDL) koji je procesiran od strane DDL kompajlera

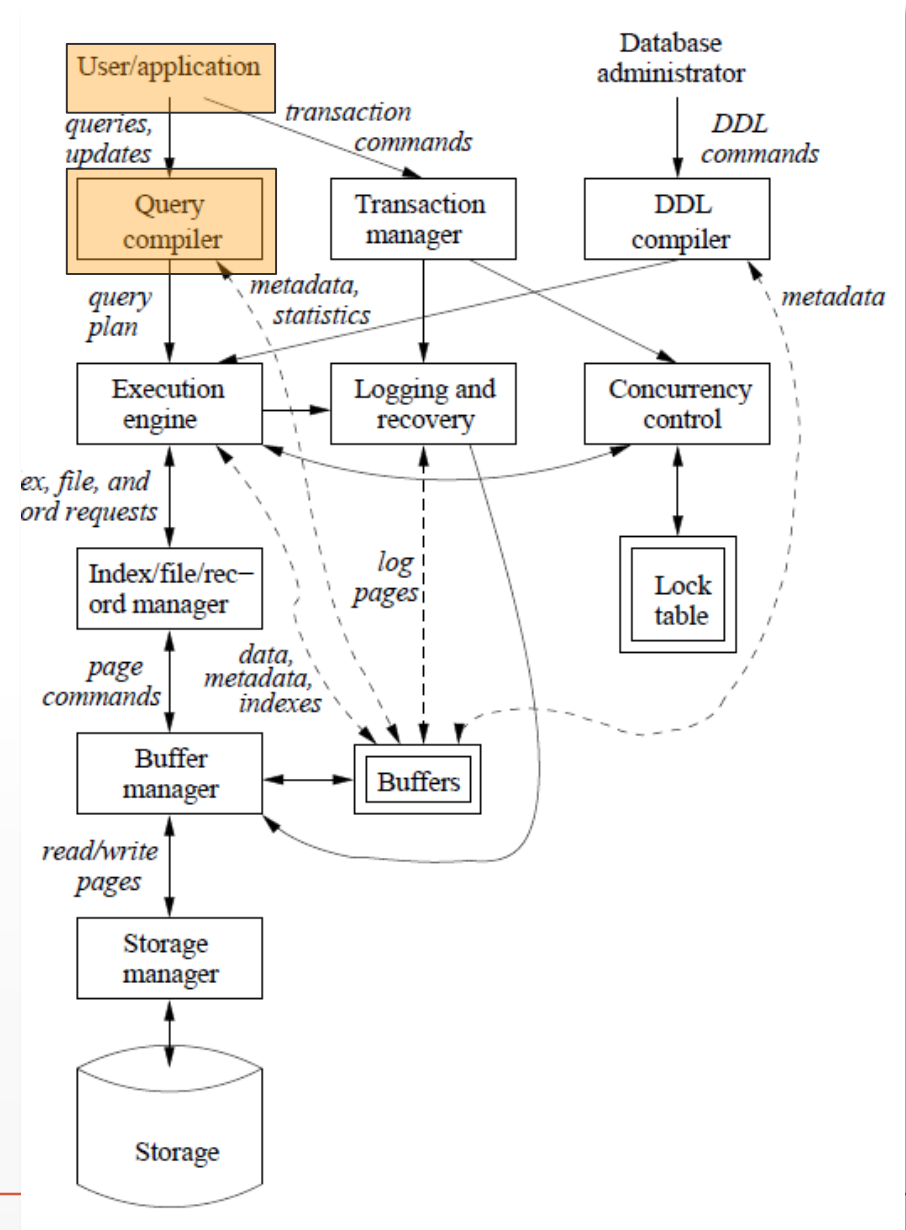


Query Processing

- Konvencionani korisnik zahteva ili menja podatke

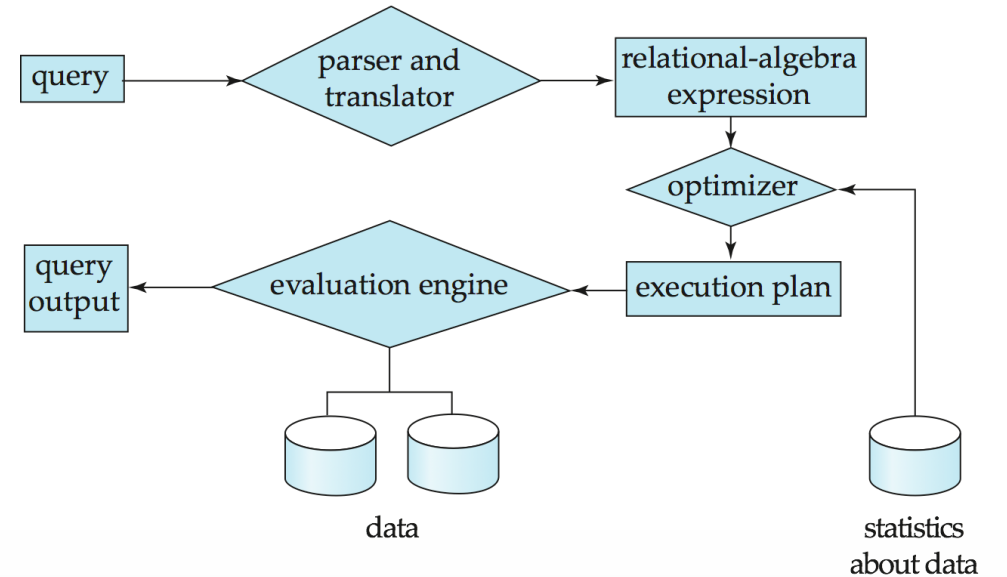
Example: a user Jack may want to take the course 101 for Spring 2012.

- Većina interakcija su upiti zabeleženi jezikom za manipulaciju podacima, data manipulation language (DML), koji je parsiran i optimizovan od strane kompajlera, Query compiler.
- Kompajler prevodi upit u interni oblik nazvan plan upita, query plan.



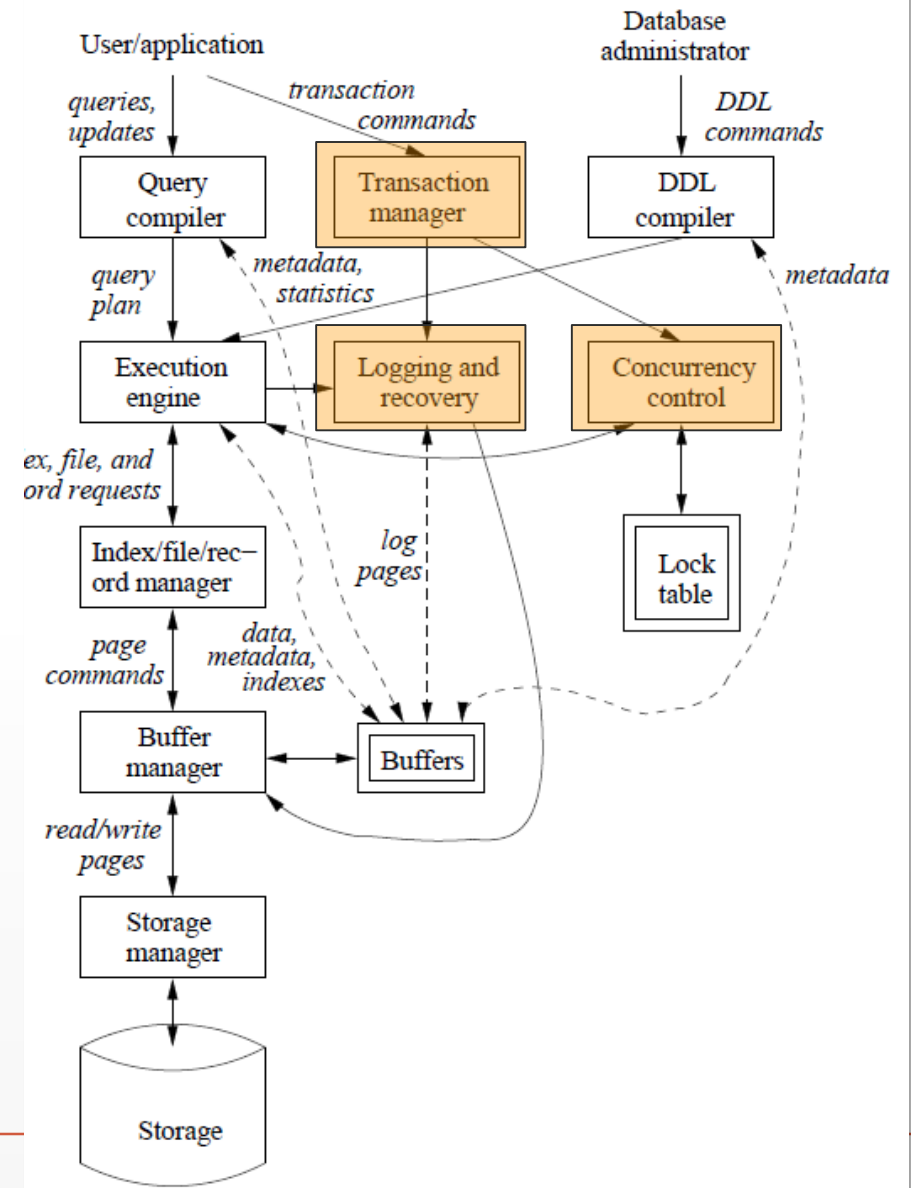
Query Processing

1. Parsiranje i prevođenje
2. Optimizacija
3. Evaluacija / izvršavanje
 - Alternative ways of evaluating a given query
 - Equivalent expressions
 - Different algorithms for each operation
 - Cost difference between a good and a bad way of evaluating a query can be enormous
 - Need to estimate the cost of operations
 - Depends critically on statistical information about relations which the database must maintain
 - Need to estimate statistics for intermediate results to compute cost of complex expressions



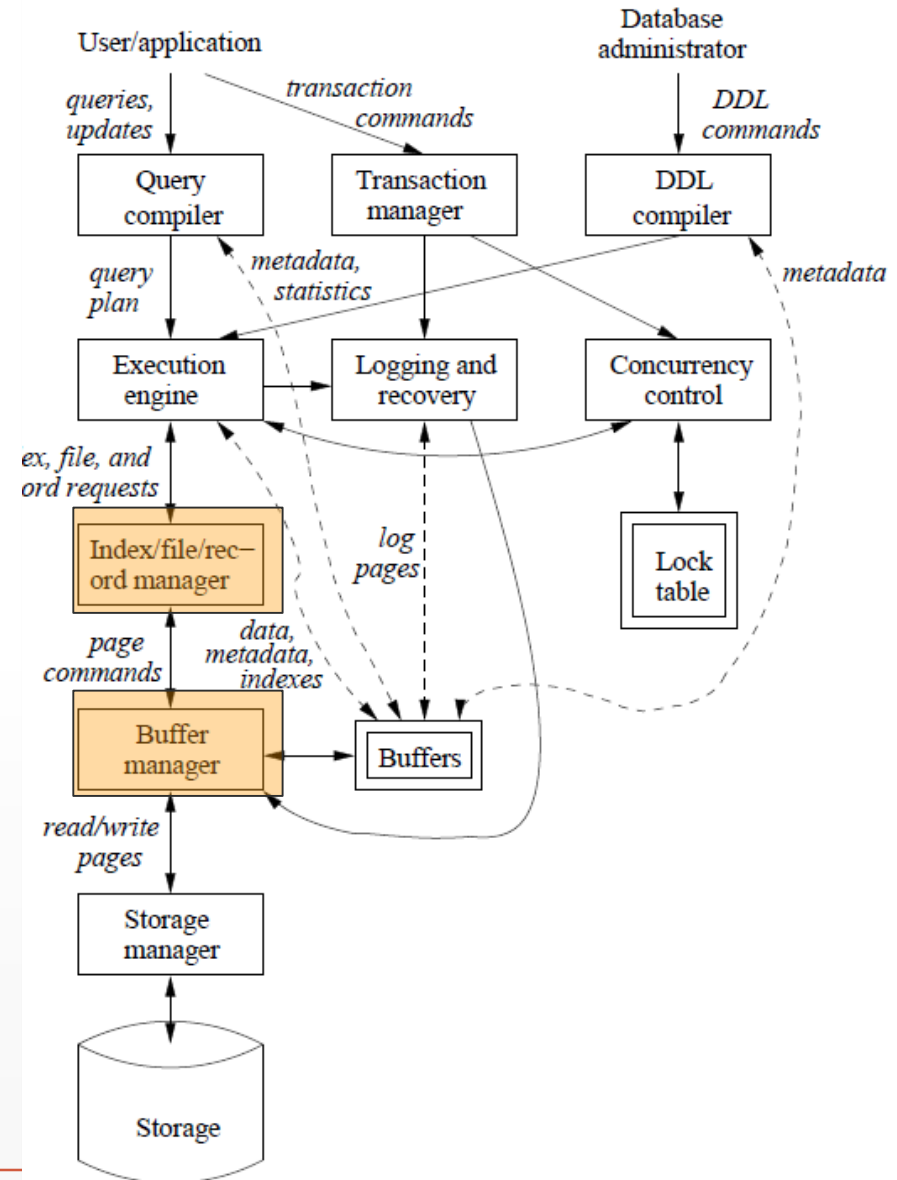
Transaction Process

- Queries and DML actions are also handled by **Transaction Manager**. They are grouped into transactions which are units that must be executed **atomically** and in **isolation** from one another. **Concurrency control** manager is responsible for this.
- The transaction must be durable – if completed must be persevered even if the system fails right after completion of transaction. **Logging and recovery** manager is responsible for this.



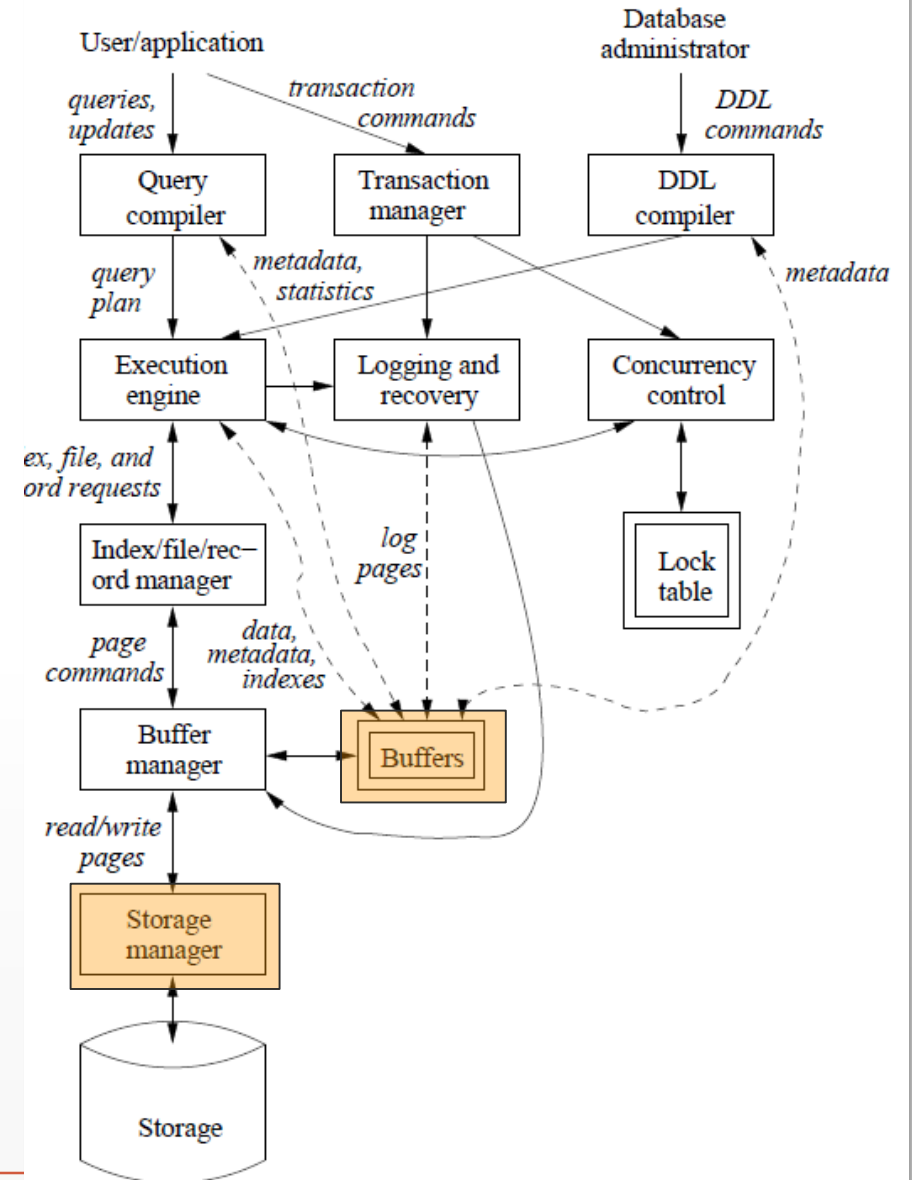
Query Execution

- The execution engine issues a sequence of request for small pieces of data (records or tuples). The request is submitted to buffer manager.
- The data is usually stored on secondary storage (hard drive). However, to perform any operation on data, it must be in the main memory. The buffer manager communicates with Storage manager to get this data.



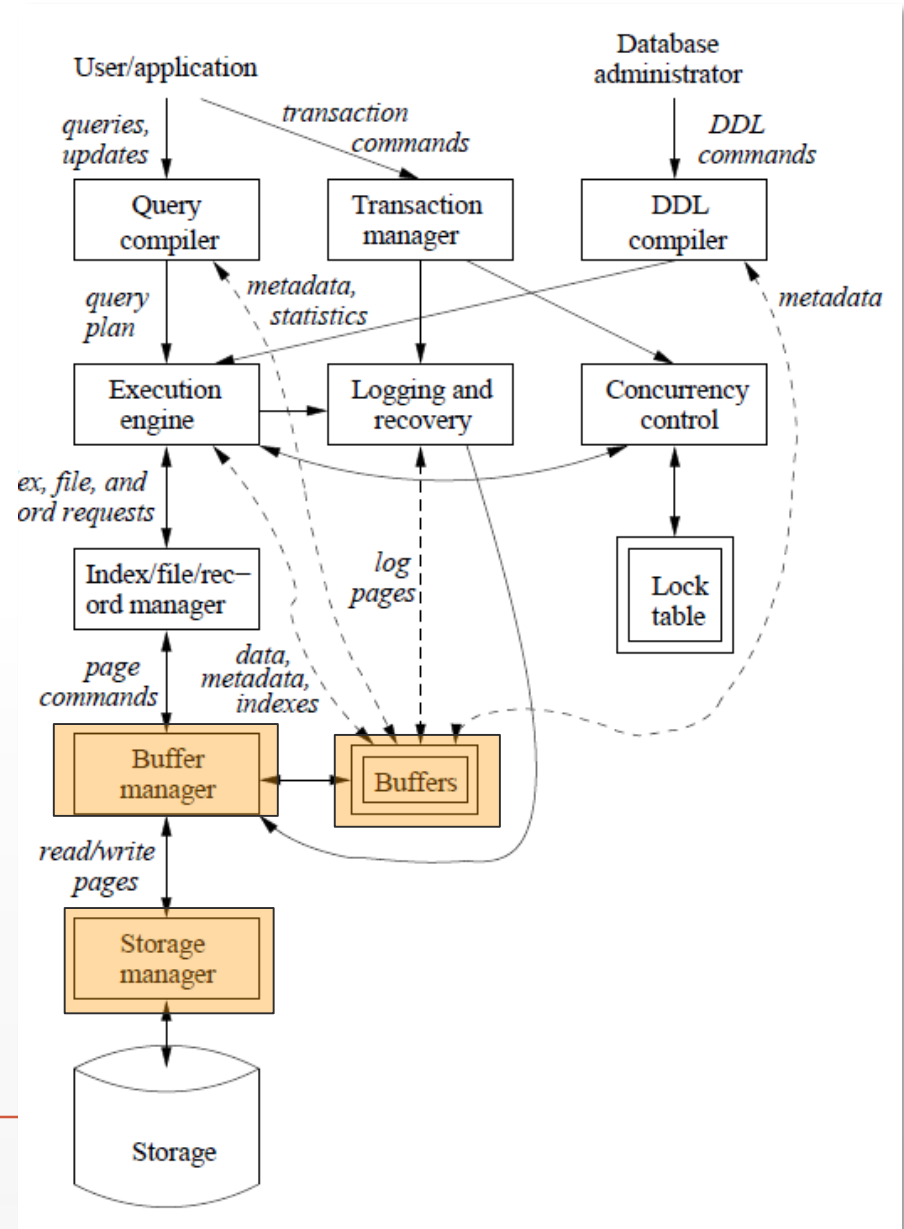
Storage and Buffer Management

- The buffer manager is responsible for partitioning the main memory into page sized buffers. All DBMS components that need information from disk interact with buffer and buffer manager directly, or through execution engine.
- This information includes
 - Data
 - Metadata
 - Log records
 - Statistics



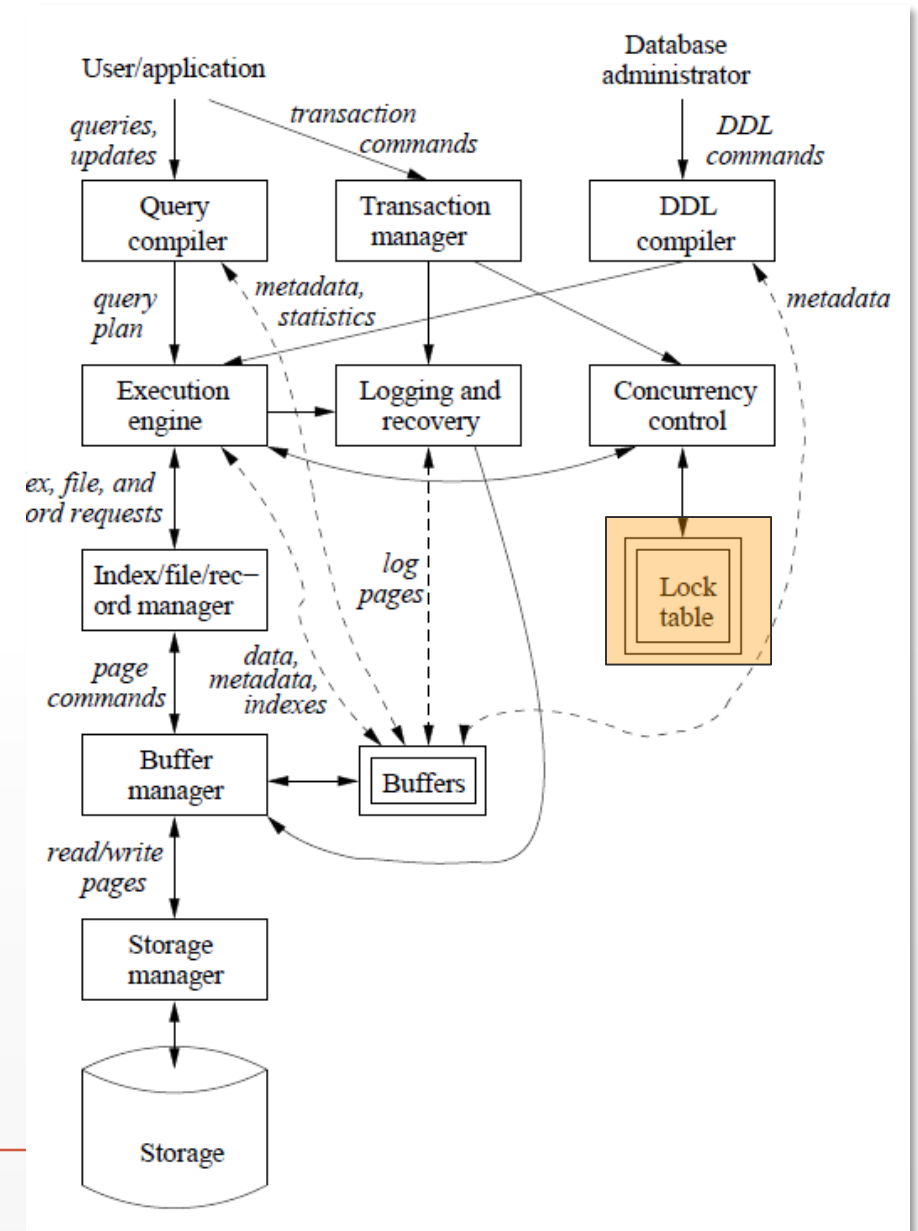
Storage and Buffer Management

- The storage manager keeps track of the location of files on the disk and provides the buffer manager with the file
- Storage manager
a program module that provides the interface between the low-level data stored in the database and the application programs and queries submitted to the system.
- The storage manager is responsible to the following tasks:
 - Interaction with the file manager
 - Efficient storing, retrieving and updating of data
- Issues:
 - Storage access
 - File organization
 - Indexing and hashing



Storage and Buffer Management

- Concurrency manager must assure that the individual actions of multiple transaction are executed such that the effect is the same as running them one at a time.
- It usually works by maintaining locks on the records. Example moving money from one account to another
- Its possible to get into a situation where all the transactions are waiting for each other due to locks. Lock table calls roll back or abort on some of the transactions to resolve deadlock.



System architecture

