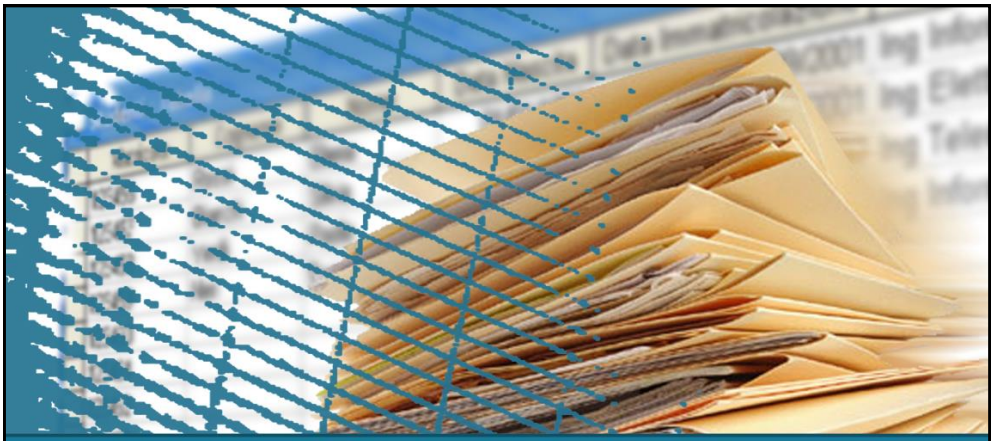


Database Management Systems

Introduction to the course

DBMG 1

Transaction processing

- On Line Transaction Processing (OLTP)
  - Traditional DBMS usage
- Characterized by
  - snapshot of current data values
  - detailed data, relational representation
  - structured, repetitive operations
  - read/write access to few records
  - short transactions
  - isolation, reliability, and integrity are critical (ACID)
  - database size ≈ 100MB-GB

DBMG 2

Analytical processing

- On Line Analytical Processing (OLAP)
  - Decision support applications
- Characterized by
  - "historical" data
  - consolidated, integrated data
  - ad hoc applications
  - read access to millions of records
  - complex queries
  - consistency before and after periodical loads
  - database size ≈ 100GB-TB

DBMG 3

Course content

- First part (weeks 1-7)
  - DBMS server technology
  - SQL Triggers
  - Distributed databases
- Second part (weeks 8-14)
  - Data warehouse design
  - OLAP analysis
  - Data mining

DBMG 4

Books

- Course books
  - Atzeni, Ceri, Fraternali, Paraboschi, Torlone, *Basi di dati – Architetture e linee di evoluzione*, McGraw Hill, 2007
  - Atzeni, Ceri, Paraboschi, Torlone, *Database systems*, McGraw Hill, 1999
  - Golfarelli, Rizzi, *Data warehouse: teoria e pratica della progettazione*, McGraw-Hill, 2006
  - Tan, Steinbach, Kumar, *Introduction to data mining*, Pearson, 2006

DBMG 5

Books

- Other books
  - Ramakrishnan, Gehrke, *Database Management Systems*, McGraw-Hill, 2004
  - Kimball e altri, *several books and white papers on data warehouse design methodologies and case studies*, Wiley
  - Han, Kamber, *Data mining: concepts and techniques*, Morgan Kaufmann, 2006

DBMG 6

Database Management Systems

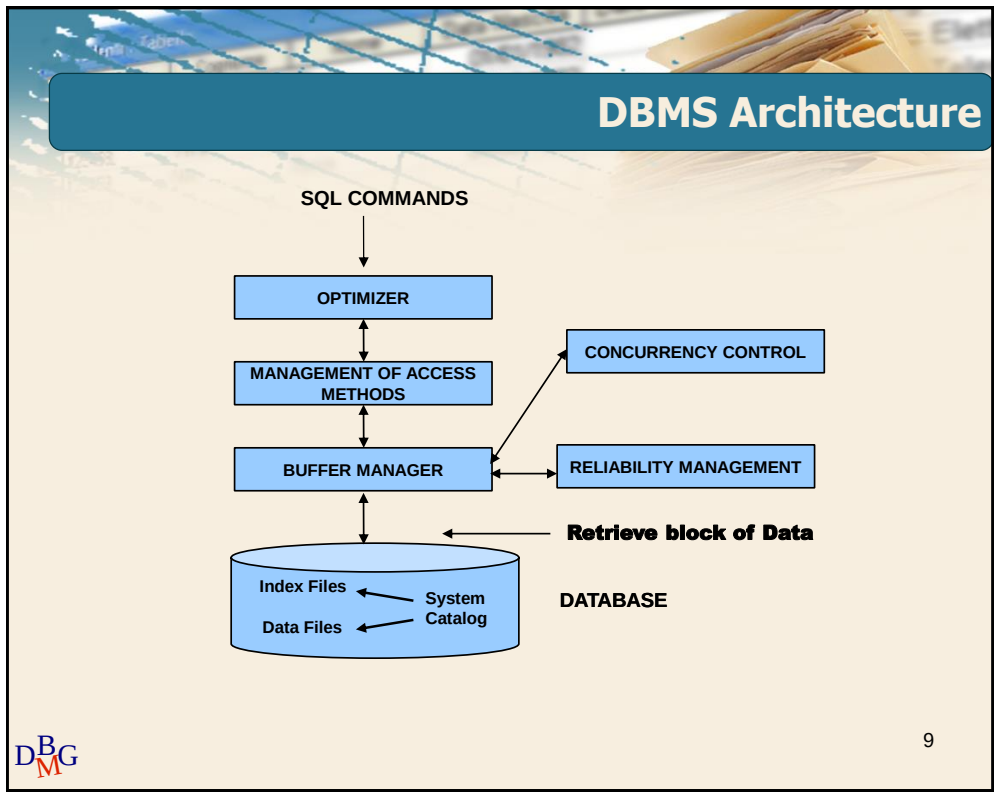
Introduction to DBMS

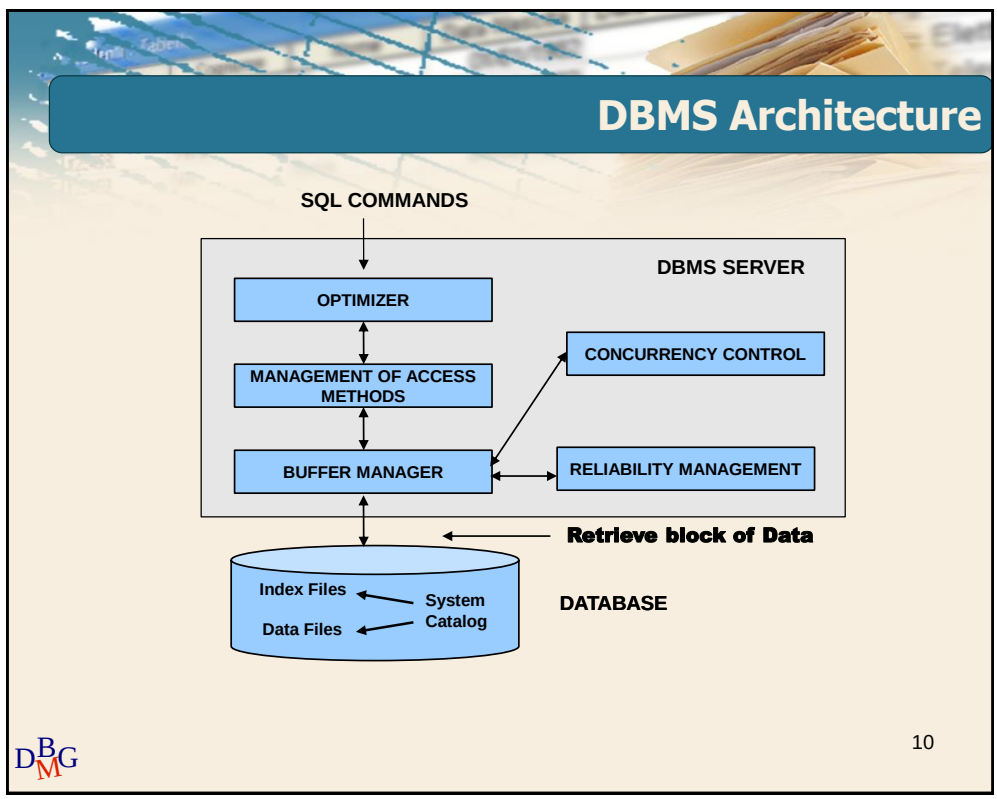
DBMG 7

Introduction to DBMS

- Data Base Management System (DBMS)
  - A software package designed to store and manage databases
- We are interested in internal mechanisms of a DBMS providing services to applications
  - Useful for making the right design choices
    - System configuration
    - Physical design of applications
  - Some services are becoming available also in operating systems

DBMG 8


DBMS Architecture
SQL COMMANDS
OPTIMIZER
MANAGEMENT OF ACCESS METHODS
BUFFER MANAGER
CONCURRENCY CONTROL
RELIABILITY MANAGEMENT
Retrieve block of Data
Index Files
Data Files
System Catalog
DATABASE
9


DBMS Architecture
SQL COMMANDS
DBMS SERVER
OPTIMIZER
MANAGEMENT OF ACCESS METHODS
BUFFER MANAGER
CONCURRENCY CONTROL
RELIABILITY MANAGEMENT
Retrieve block of Data
Index Files
Data Files
System Catalog
DATABASE
10

DBMS Components

- Optimizer
  - It selects the appropriate execution strategy for accessing data to answer queries
  - It receives in input a SQL instruction (DML)
  - It executes lexical, syntactic, and semantic parsing and detects (some) errors
  - It transforms the query in an internal representation (based on relational algebra)
  - It selects the "right" strategy for accessing data
- This component guarantees the *data independence* property in the relational model

DBMG 11

DBMS Components

- Access Method Manager
  - It performs physical access to data
  - It implements the strategy selected by the optimizer

DBMG 12

DBMS Components

- Buffer Manager
  - It manages page transfer from disk to main memory and vice versa
  - It manages the main memory portion that is pre-allocated to the DBMS
    - e.g., Oracle SGA
- The memory block pre-allocated to the DBMS is *shared* among many applications

DBMG 13

DBMS Components

- Concurrency Control
  - It manages concurrent access to data
    - Important for write operations
  - It guarantees that applications do not interfere with each other, thus yielding consistency problems

DBMG 14

DBMS Components

- Reliability Manager
  - It guarantees correctness of the database content when the system crashes
  - It guarantees atomic execution of a transaction (sequence of operations)
  - It exploits auxiliary structures (log files) to recover the correct database state after a failure

DBMG 15

Transaction

- A *transaction* is a logical unit of work performed by an application
  - It is a sequence of one or more SQL instructions, performing read and write operations on the database
- It is characterized by
  - Correctness
  - Reliability
  - Isolation

DBMG 16

Transaction example: Bank Tranfer

- The following transaction moves 100 euro from account xxx to account yyy

```
UPDATE ACCOUNTS
SET Balance = Balance - 100
WHERE Account_Number = xxx

UPDATE ACCOUNTS
SET Balance = Balance + 100
WHERE Account_Number = yyy
```

DBMG 17

Transaction delimiters

- Transaction start
  - Typically implicit
  - First SQL instruction
    - At the beginning of a program
    - After the end of the former transaction
- Transaction end
  - COMMIT: correct end of a transaction
  - ROLLBACK: end with error
    - The database state goes back to the state at the beginning of the transaction

DBMG 18

Transaction end

- 99.9% of transactions commit
- Remaining transactions rollback
  - Rollback is required by the transaction (suicide)
  - Rollback is required by the system (murder)

DBMG 19

Transaction properties

- ACID properties of transactions
  - **A**tomicity
  - **C**onsistency
  - **I**solation
  - **D**urability

DBMG 20

Atomicity

- A transaction cannot be divided in smaller units
  - It is *not* possible to leave the database in a intermediate state of execution
- Guaranteed by
  - *Undo.* The system undoes all the work performed by the transaction up to the current point
    - It is used for rollback
  - *Redo.* The system redoes all work performed by committed transactions
    - It is used to guarantee transaction commit in presence of failure

DBMG 21

Consistency

- A transaction execution should not violate integrity constraints on a database
  - Enforced by defining integrity constraints in the database schema (Create table, ….)
    - Primary key
    - Referential Integrity (Foreign key)
    - Domain Constraints
    - …
  - When a violation is detected, the system may
    - Rollback the transaction
    - Automatically correct the violation

DBMG 22

Isolation

- The execution of a transaction is *independent* of the concurrent execution of other transactions
  - Enforced by the Concurrency Control block of the DBMS

DBMG 23

Durability

- The effect of a committed transaction *is not lost* in presence of failures
  - It guarantees the reliability of the DBMS
  - Enforced by the Reliability Manager block of the DBMS

DBMG 24