

*Corso di Laurea in Ingegneria Gestionale
Sapienza Università di Roma*

Corso di Basi di Dati

A.A. 2019/2020

1 – Nozioni Preliminari

Tiziana Catarci, Francesco Leotta

Ultimo aggiornamento : 25/09/2019

Risorse di una organizzazione

- ▶ Nello svolgimento di ogni attività, un'organizzazione deve disporre di determinate **risorse** per perseguire gli scopi che si è prefissa :
 - persone
 - denaro
 - materiali
 - **dati e informazioni**
- ▶ Ogni organizzazione dispone di un *sistema informativo* che organizza e gestisce un gran numero di *informazioni* rappresentate per mezzo di *dati*.
 - *I sistemi informativi esistono da molto prima dell'invenzione dei calcolatori.*
 - *Esempio : elenchi di utenze telefoniche, archivi anagrafici ...*

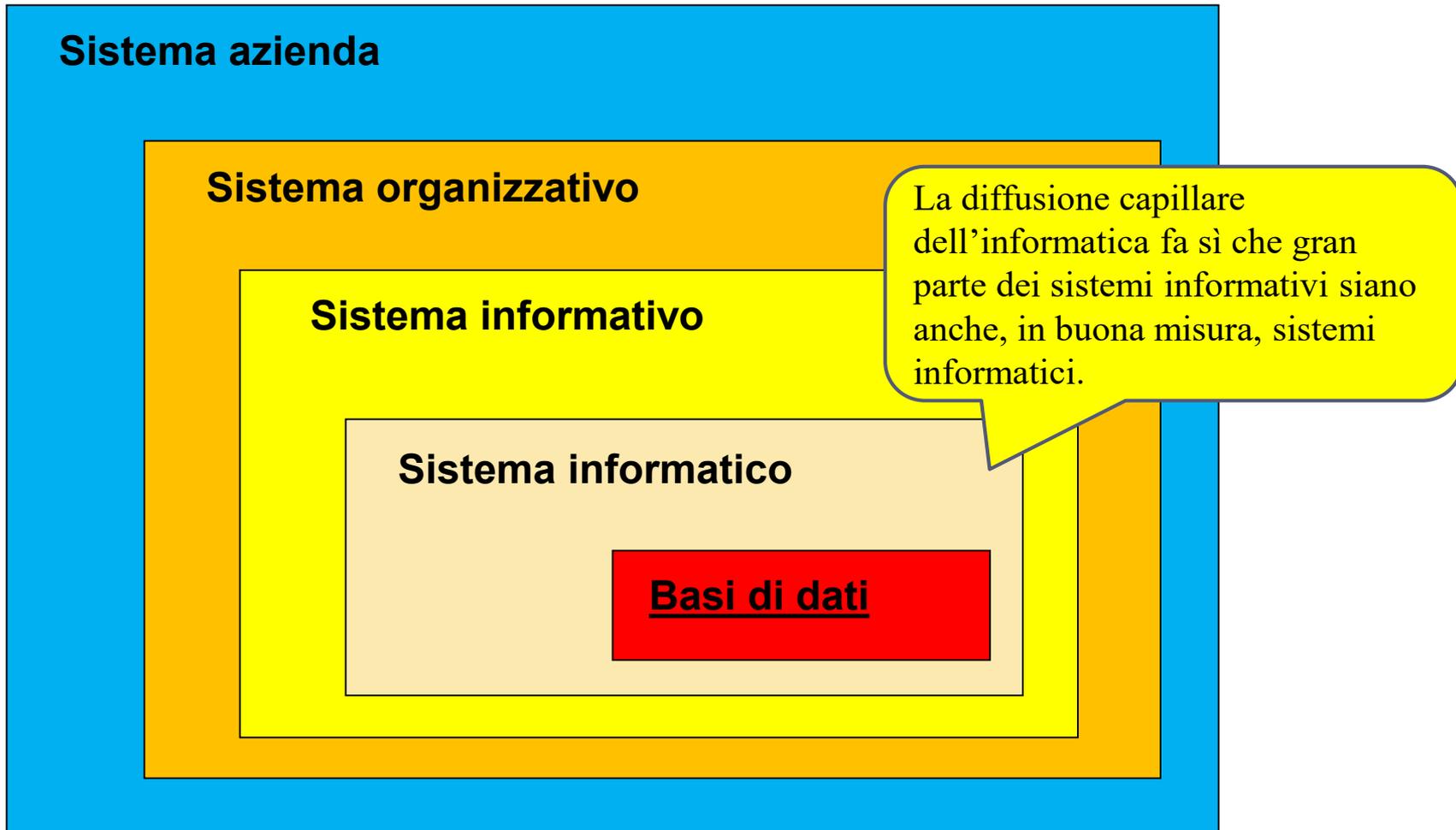
Funzioni di un Sistema Informativo

- ▶ Principali funzioni di un Sistema Informativo :
 - Raccolta e acquisizione delle informazioni;
 - archiviazione, conservazione delle informazioni;
 - elaborazione delle informazioni;
 - distribuzione, scambio di informazioni.

Il concetto di Sistema Informativo è indipendente da qualsiasi forma di automatizzazione!

- ▶ Un **Sistema Informatico** è quella porzione automatizzata del Sistema Informativo che gestisce le informazioni con tecnologia informatica.

Sistema Informatico



Dati e Informazioni

- ▶ Nelle *attività umane*, le informazioni vengono gestite in forme diverse:
 - idee informali, linguaggio naturale (scritto o parlato, formale o colloquiale), mente umana, carta, dispositivi elettronici, disegni, grafici, schemi, numeri e codici.
- ▶ Nei *sistemi informatici*, le **informazioni** vengono rappresentate per mezzo di **dati** :
 - ▶ **informazione** : notizia, dato o elemento che consente di avere conoscenza più o meno esatta di fatti, situazioni, modi di essere.
 - ▶ **dato** : ciò che è immediatamente presente alla conoscenza, prima di ogni elaborazione. In informatica, un dato è un elemento di informazione costituito da simboli che debbono essere elaborati.
 - ▶ Senza “interpretazione“, i dati hanno poca utilità.
 - ▶ Se interpretati e correlati opportunamente, essi forniscono informazioni che consentono di arricchire la nostra conoscenza del mondo.

Base di Dati

▶ Base di Dati :

(accezione generica) **Collezione di dati**, che tipicamente descrive le informazioni di interesse di una o più organizzazioni correlate.

(accezione specifica) **Collezione di dati** in memoria secondaria gestita da un apposito sistema software, chiamato **DBMS (Data Base Management System)**, o Sistema di Gestione di Basi di Dati).

- Un DBMS è in grado di gestire e interrogare grandi collezioni di dati.
- I dati costituiscono una **risorsa strategica**, perché **più stabili nel tempo** di altre componenti (processi, tecnologie, ruoli umani).
 - *Ad esempio, i dati delle banche o delle anagrafi hanno una struttura sostanzialmente invariata da decenni, mentre le procedure che agiscono su di essi variano con una certa frequenza.*

Data Base Management System (DBMS)

▶ **Sistema che gestisce collezioni di dati :**

- grandi
- persistenti
- condivise

garantendo

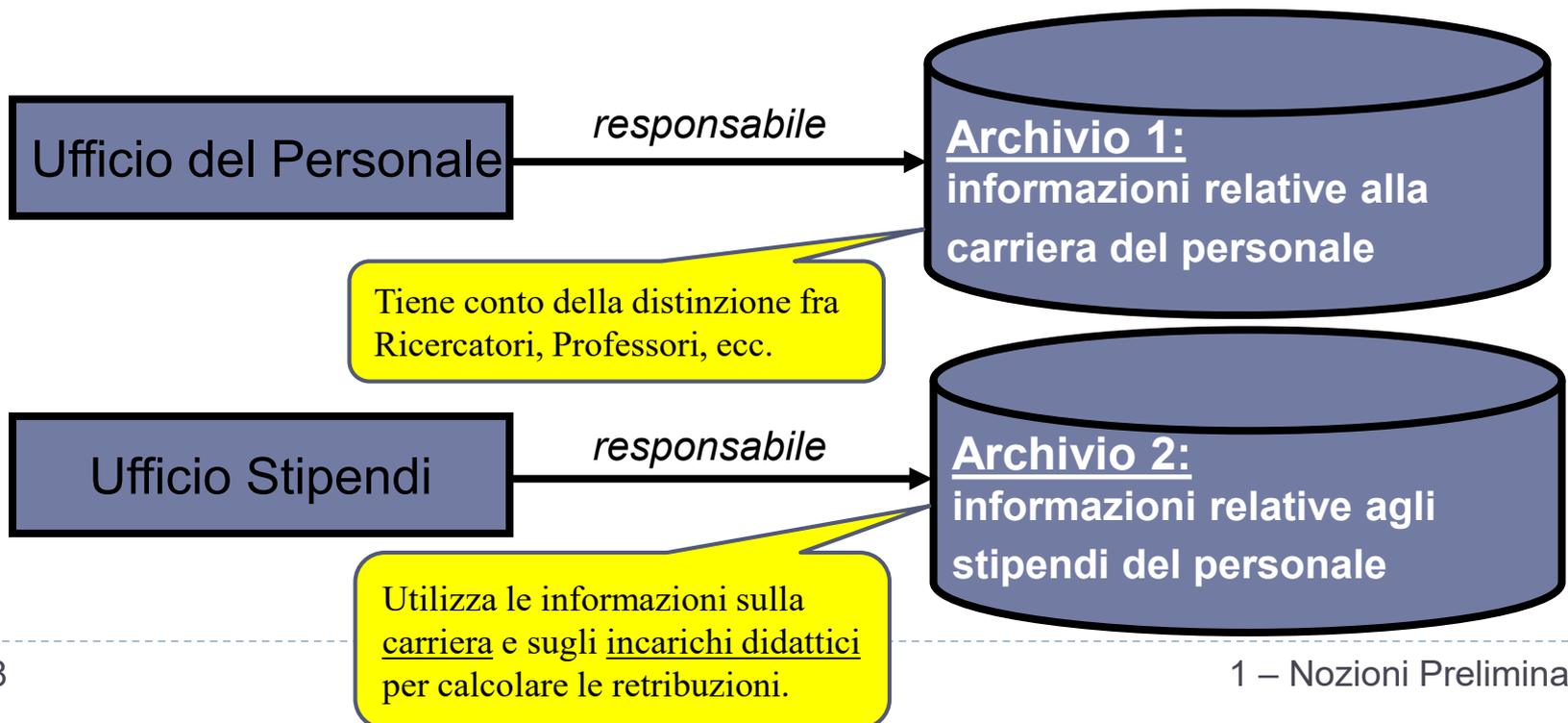
- privatezza
- affidabilità
- efficienza
- efficacia

▶ *Alcuni prodotti software disponibili sul mercato :*

- ▶ *DB2*
- ▶ *Oracle*
- ▶ *SQLServer*
- ▶ *MySQL*
- ▶ *PostgreSQL*
- ▶ *Microsoft Access*

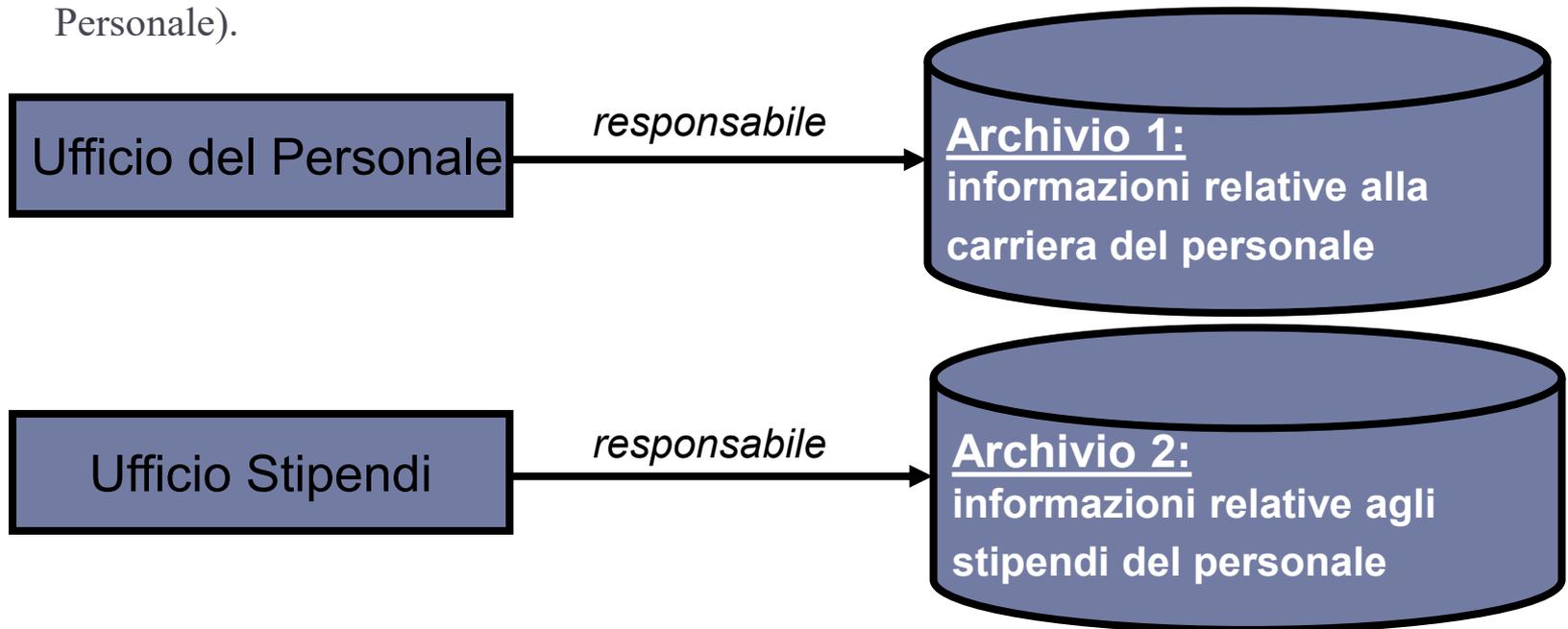
Approccio Convenzionale : *i file*

- ▶ L'approccio convenzionale alla gestione dei dati sfrutta la presenza di *archivi* o *file* per memorizzare i dati in modo persistente sulla memoria secondaria.
- ▶ Secondo quest'approccio, eventuali dati di interesse per più programmi sono **replicati** tante volte quanti sono i programmi che li utilizzano.
- ▶ Ad esempio, in un'università, le informazioni relative ai docenti possono essere utilizzate in vario modo e da diverse persone o uffici.

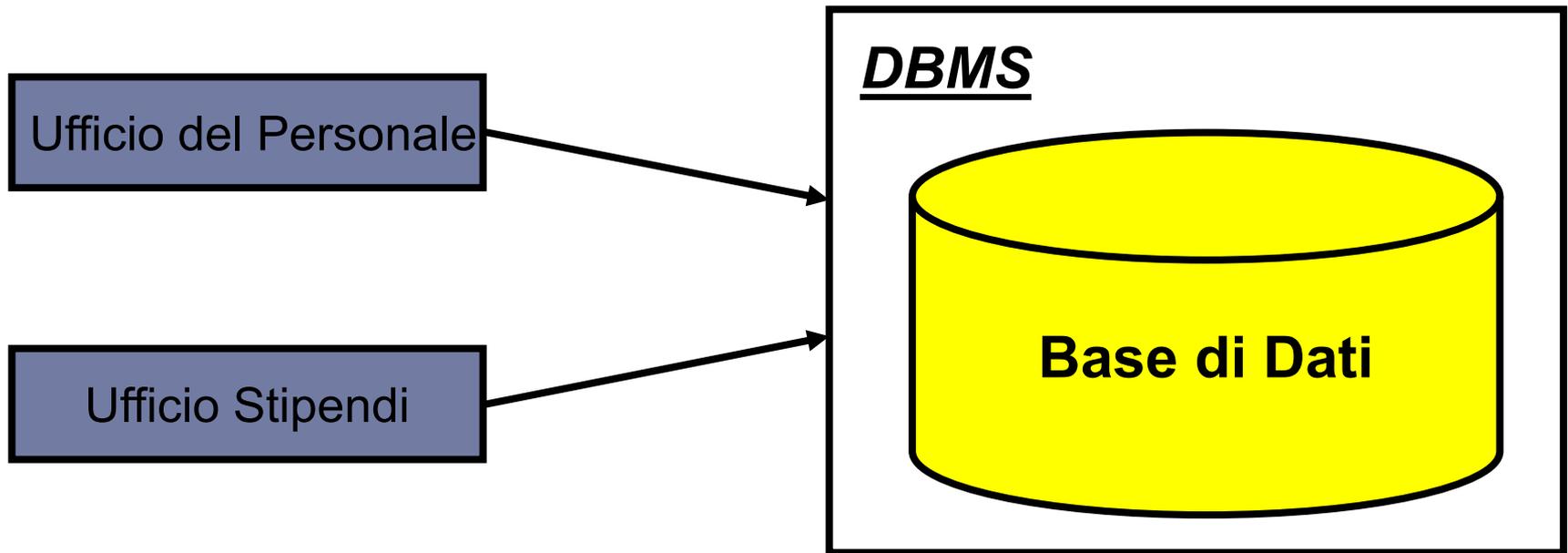


Problemi nella gestione dei *file*

- ▶ **Cosa accade se ciascun ufficio gestisce separatamente le informazioni di proprio interesse?**
 - ▶ **Presenza di duplicati** (eventuali dati di interesse per più programmi sono replicati tante volte quanti sono i programmi che li utilizzano).
 - ▶ **Difficoltà nell'aggiornamento dei dati** (a causa della presenza di duplicati, a lungo andare ci sarebbero molti dati con le copie non aggiornate allo stesso modo).
 - ▶ Possibile **incoerenza** nelle varie copie dei dati per ciascun archivio (potremmo trovare nell'Ufficio Stipendi una qualifica diversa da quella corretta presente nei dati dell'Ufficio del Personale).



Architettura del Sistema Informatico



- ▶ Le basi di dati sono state concepite principalmente per superare questo tipo di inconvenienti, gestendo in modo **integrato e flessibile** le informazioni di interesse per diversi soggetti, **limitando i rischi di ridondanza e incoerenza.**
- ▶ Il DBMS è **l'unico responsabile** della gestione della base di dati: i dati sono accessibili all'esterno **solo** tramite il DBMS.

Esempio di rappresentazione dei dati

UNIVERSITA' DEGLI STUDI DI CHISSADOVE

Corso di Studi in Ingegneria Informatica

ORARIO DELLE LEZIONI PER L'ANNO ACCADEMICO 1999-2000

INSEGNAMENTO	Docente	Aula	Orario
Analisi matematica I	Luigi Neri	N1	8:00-9:30
Basi di dati	Piero Rossi	N2	9:45-11:15
Chimica	Nicola Mori	N1	9:45-11:30
Fisica I	Mario Bruni	N1	11:45-13:00
Fisica II	Mario Bruni	N3	9:45-11:15
Sistemi informativi	Piero Rossi	N3	8:00-9:30

UNIVERSITA' DEGLI STUDI CHISSA

Corso di Studi in Ingegneria Informa

Orario di ricevimento dei docenti

DOCENTE	INSEGNAMENTI	ORARIO
Mario BRUNI	Fisica I Fisica II	Martedì' 10-12
Luigi NERI	Analisi matematica I	Lunedì' 12-13
Piero ROSSI	Basi di dati Sistemi informativi	Giovedì' 11-13
Nicola MORI	Chimica	Martedì' 16-18

Le basi di dati sono...grandi

- ▶ Dimensioni (molto) maggiori della memoria centrale dei sistemi di calcolo utilizzati.
- ▶ Il limite deve essere solo quello fisico dei dispositivi.
- ▶ I DBMS devono prevedere una gestione dei dati in memoria secondaria.

- ▶ Esempi di dimensioni molto grandi :
 - 500 Gigabyte (dati transazionali)
 - 10 Terabyte (dati decisionali)
 - 500 Terabyte (dati scientifici)
 - 100 miliardi di record

Le basi di dati sono...condivise

- ▶ Ogni organizzazione (specie se grande) è divisa in settori o comunque svolge diverse attività. Perciò utenti e applicazioni diversi devono poter accedere, secondo opportune modalità, a dati comuni.
- ▶ Ciascun settore/attività ha un (sotto)sistema informativo (non necessariamente disgiunto).
- ▶ In questo modo si riduce la *ridondanza* dei dati (poichè si evitano ripetizioni) e conseguentemente si riduce anche la possibilità di *inconsistenze*.
- ▶ Per garantire l'accesso condiviso, il DBMS dispone di un meccanismo apposito, detto *controllo di concorrenza*.

Le basi di dati sono...persistenti

- ▶ Hanno un tempo di vita indipendente dalle singole esecuzioni dei programmi che le utilizzano.
- ▶ In contrasto, i dati gestiti da un programma in memoria centrale hanno una vita che inizia e termina con l'esecuzione del programma.

I DBMS garantiscono...affidabilità

- ▶ Capacità del sistema di conservare sostanzialmente intatto il contenuto della Base di Dati (o almeno di permetterne la ricostruzione) in caso di malfunzionamenti hardware e software.
- ▶ Una base di dati è una risorsa pregiata e quindi deve essere conservata a lungo termine a fronte di qualsiasi guasto del sistema, errore umano o anche evento catastrofico.
- ▶ A questo scopo i DBMS gestiscono, in modo controllato, funzionalità di **salvataggio e ripristino** (*backup e recovery*).
 - ▶ In alcuni casi, i DBMS gestiscono in modo controllato, versioni replicate dei dati, collocate su dispositivi fisici diversi, così da garantire maggiore affidabilità complessiva.

I DBMS garantiscono...privatezza

- ▶ Ciascun utente, riconosciuto in base a un nome utente che è specificato all'atto di interagire con il DBMS, viene abilitato a svolgere solo determinate azioni sui dati, attraverso meccanismi di *autorizzazione*.
- ▶ Si possono definire meccanismi di autorizzazione :
 - l'utente A è autorizzato a leggere tutti i dati e a modificare X.
 - l'utente B è autorizzato a leggere dati X e a modificare Y.

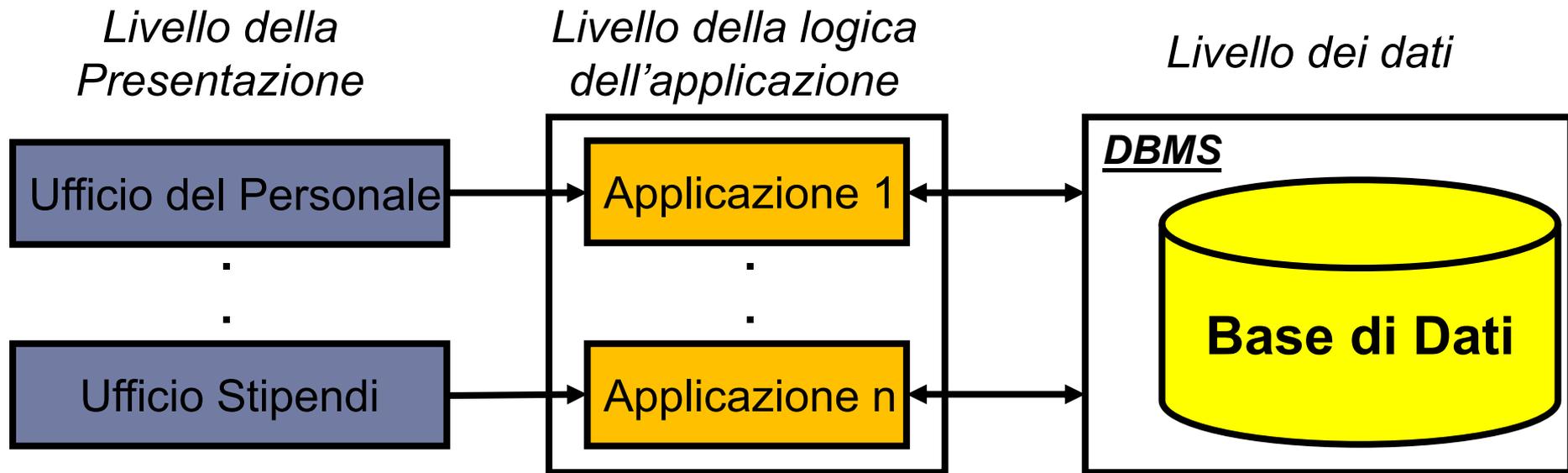
I DBMS sono efficienti

- ▶ I DBMS sono capaci di svolgere le operazioni utilizzando un insieme di risorse (tempo e spazio) che sia accettabile per gli utenti.
- ▶ *Va sottolineato che i DBMS forniscono un insieme piuttosto ampio di funzionalità che richiedono molte risorse, e quindi possono garantire efficienza solo a condizione che il sistema informatico su cui sono installati sia adeguatamente dimensionato.*

I DBMS sono efficaci

- ▶ I DBMS sono capaci di rendere produttive, in ogni senso, le attività dei loro utenti.
- ▶ *Questa definizione è chiaramente generica e non corrisponde ad un aspetto specifico. L'attività di progettazione della base di dati e delle applicazioni che la utilizzano mira essenzialmente a garantire una buona efficacia complessiva del sistema.*

Architettura a tre livelli del Sistema Informatico



- ▶ Una base di dati viene utilizzata dai vari uffici (e persone), ciascuno con le proprie competenze, attraverso **programmi diversi**.
- ▶ Il DBMS garantisce **integrazione e condivisione** coordinando i vari flussi di informazioni che giungono dalle varie sorgenti.

DBMS vs File System

- ▶ La gestione di insiemi di dati grandi e persistenti è possibile anche attraverso sistemi più semplici - i classici file system dei sistemi operativi.
- ▶ Un file consente di memorizzare e ricercare dati, ma fornisce solo semplici meccanismi di accesso e condivisione.
- ▶ I DBMS estendono le funzionalità dei file system, fornendo più servizi ed in maniera integrata.
- ▶ I file, a loro volta, sono fondamentali per gestire insiemi di dati **localmente** a una specifica procedura o applicazione.

Descrizione dei Dati

- ▶ Nei programmi tradizionali che accedono a file, ogni programma contiene una descrizione della struttura del file stesso, con i conseguenti rischi di incoerenza fra le descrizioni (ripetute in ciascun programma) e i file stessi.
- ▶ Nei DBMS, esiste una porzione della base di dati (il catalogo o dizionario) che contiene una descrizione centralizzata dei dati, utilizzabile dai vari programmi.
- ▶ La rappresentazioni dei dati avviene a livelli diversi
 - ▶ in modo da permettere l'indipendenza dei dati dalla rappresentazione fisica:
 - ▶ i programmi fanno riferimento alla struttura a livello più alto, e le rappresentazioni sottostanti possono essere modificate senza necessità di modifica dei programmi
 - ▶ E' fondamentale il concetto di **modello dei dati**.

Modello dei Dati

- ▶ Insieme di costrutti utilizzati per organizzare i dati di interesse all'interno del DBMS e descriverne la dinamica.
- ▶ Come nei linguaggi di programmazione esistono meccanismi che permettono di definire nuovi tipi, così ogni modello dei dati prevede alcuni costruttori.
- ▶ Ad esempio, il **modello relazionale** prevede il **costruttore relazione**, che **permette di definire insiemi di record omogenei**
- ▶ Una relazione viene rappresentata per mezzo di una **tabella**, le cui righe rappresentano specifici record e le cui colonne corrispondono ai campi del record.

Il Modello Relazionale

- ▶ Una relazione viene rappresentata per mezzo di una **tabella**, le cui righe rappresentano specifici record e le cui colonne corrispondono ai campi del record.
- ▶ E' il più diffuso nell'ambito delle Basi di Dati. Sarà anche il nostro modello di riferimento.

Intestazione della tabella: nome della relazione

Squadra

Attributo

Nome	Cognome	Posizione	Squadra	Età
Tommaso	Rocchi	Attaccante	S.S.Lazio	31
Alessandro	Del Piero	Attaccante	Juventus F.C.	34
Francesco	Totti	Attaccante	A.S. Roma	32

Tupla o Record

Schemi e Istanze

- ▶ In ogni base di dati si distinguono :
 - lo **schema**, sostanzialmente **invariante nel tempo**, che ne descrive la struttura (**aspetto intensionale**); *nell'esempio*, le **intestazioni** delle tabelle :

Giocatore (Nome, Cognome, Posizione, Squadra, Età)
Squadra (Nome, Anno di Fondazione)

- ▶ l'**istanza**, costituita dai valori attuali, che possono cambiare molto rapidamente (**aspetto estensionale**); *nell'esempio*, il “**corpo**” di ciascuna tabella :

Tommaso	Rocchi	Attaccante	S.S. Lazio	31
Alessandro	Del Piero	Attaccante	Juventus F.C.	34
Francesco	Totti	Attaccante	A.S. Roma	32

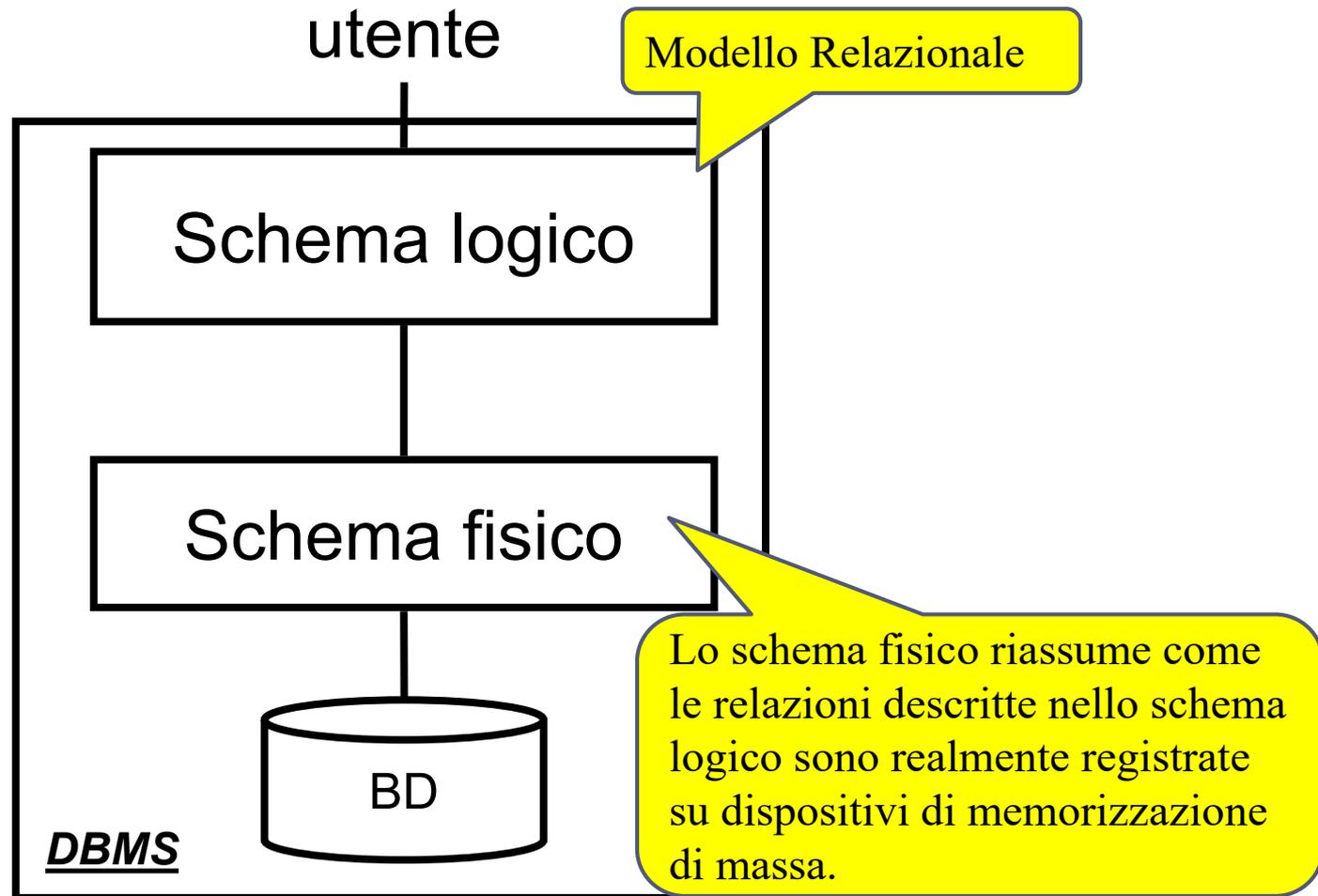
Due tipi (principali) di modelli

- ▶ **Modelli Logici:** utilizzati nei DBMS esistenti per l'organizzazione dei dati; ad essi fanno riferimento i programmi e gli utenti. I modelli logici **sono indipendenti** dalle strutture fisiche.
 - ▶ **Modello Gerarchico** (basato sull'uso di strutture ad albero) – anni 60
 - ▶ **Modello Reticolare** (basato sull'uso di grafi) – anni 70
 - ▶ **Modello Relazionale** – anni 70/80
 - ▶ **Modello ad Oggetti** (estende alle basi di dati il paradigma di programmazione ad oggetti) – anni 80
 - ▶ **Modello XML** (rivisitazione del modello gerarchico, in cui i dati devono essere presentati assieme alla loro descrizione)

Due tipi (principali) di modelli

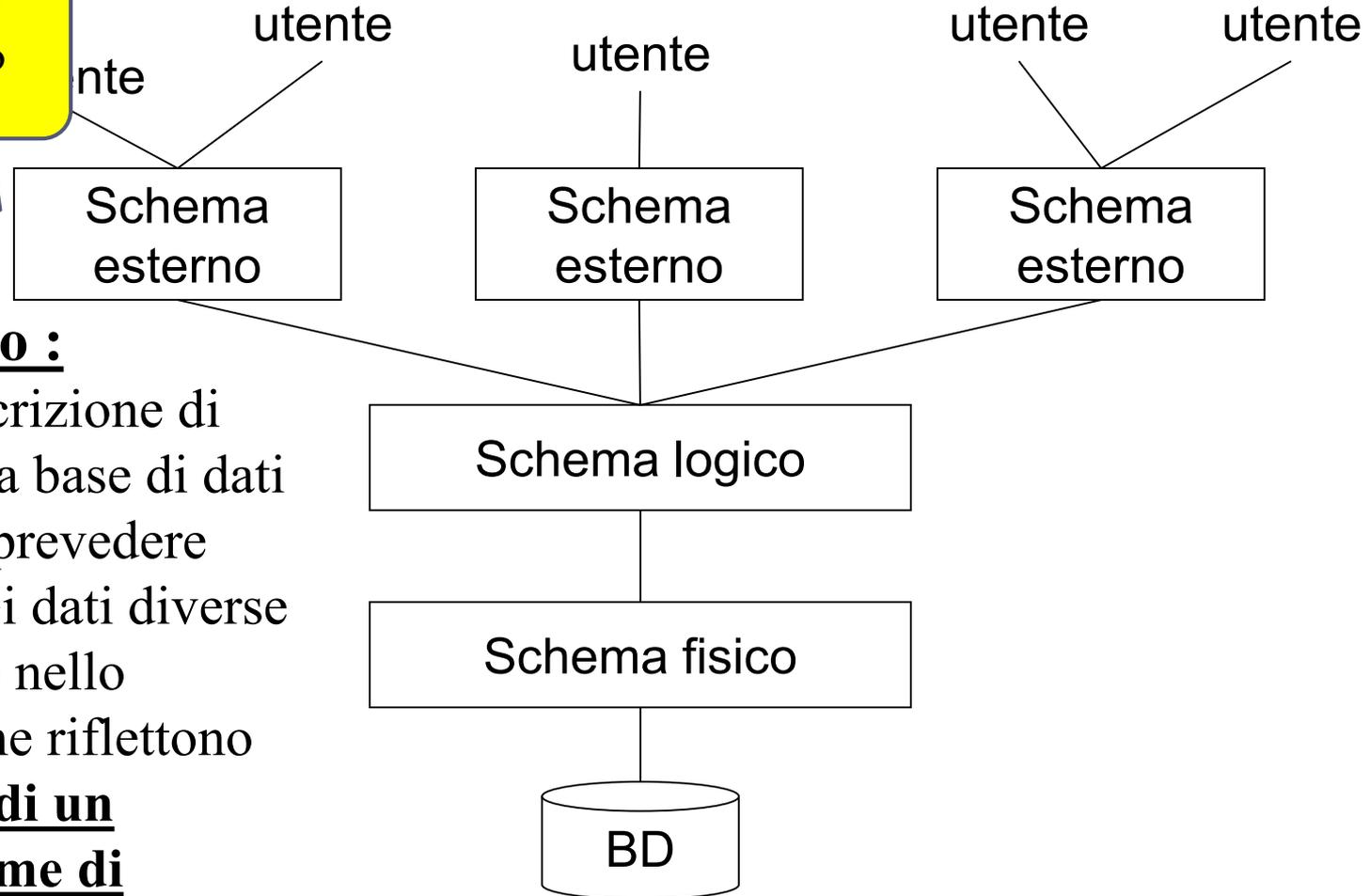
- ▶ **Modelli Concettuali:** permettono di rappresentare i dati in modo indipendente da ogni sistema, cercando di descrivere i concetti del mondo reale, piuttosto che i dati utili a rappresentarli.
 - Il più noto è il **modello Entità-Relazione**.
- ▶ Essi vengono utilizzati nelle fasi preliminari del processo di progettazione della base di dati.
 - Rappresentano il modo migliore per analizzare la realtà d'interesse, senza “contaminazioni” di tipo realizzativo.

Architettura semplificata per un DBMS



Architettura standard a tre livelli per DBMS

E' possibile associare ad uno schema logico vari schemi esterni.



Schema Esterno :

costituisce la descrizione di una porzione della base di dati di interesse. Può prevedere organizzazioni dei dati diverse a quelle utilizzate nello schema logico, che riflettono il **punto di vista di un particolare insieme di utenti.**

Indipendenza dei dati

- ▶ Conseguenza della articolazione in livelli: **l'accesso avviene solo tramite il livello esterno** (che può coincidere con il livello logico).
 - ▶ Nei sistemi moderni, lo schema esterno non è esplicitamente presente, ma è possibile definire relazioni derivate (o *viste*).
- ▶ **Due forme di indipendenza dei dati :**
 - **fisica:** il livello logico e quello esterno sono indipendenti da quello fisico; una relazione è utilizzata nello stesso modo qualunque sia la sua realizzazione fisica (che può anche cambiare nel tempo senza che debbano essere modificate le forme di utilizzo).
 - **logica:** il livello esterno è indipendente da quello logico
 - ▶ aggiunte o modifiche alle viste non richiedono modifiche al livello logico.
 - ▶ modifiche allo schema logico che lascino inalterato lo schema esterno sono trasparenti.

Esempio : Una Vista

Corsi

Corso	Docente	Aula
Basi di dati	Marrella	DS1
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

Aule

Nome	Edificio	Piano
DS1	Meschini	Primo
N3	Meschini	Terra
G	Pincherle	Primo

CorsiIstitutoMeschini

Corso	Aula	Piano
Sistemi	N3	Terra
Reti	N3	Terra
Basi di Dati	DS1	Primo

Controllo di Concorrenza

- ▶ L'esecuzione concorrente di programmi utente è essenziale per ottenere buone prestazioni dal DBMS.
 - ▶ Poichè gli accessi al disco sono frequenti, e relativamente lenti, è importante tenere occupata la CPU lavorando su diversi programmi utente concorrentemente.
 - ▶ Le azioni interallacciate di diversi programmi utente possono portare a inconsistenza: ad esempio, un assegno viene pagato mentre viene calcolato il bilancio del conto corrente.
- ▶ Il DBMS garantisce che tali problemi non si presentino: ogni utente può immaginare di essere l'unico utente del sistema.

Transazione: l'esecuzione di un programma sulla base di dati

- ▶ Il concetto chiave è la **transazione**, che è una sequenza atomica di azioni sulla base di dati (letture/scritture).
- ▶ Ciascuna transazione, eseguita completamente, deve lasciare la base di dati in uno stato consistente se esso era consistente quando la transazione ha avuto inizio.
 - ▶ Gli utenti possono specificare alcuni semplici vincoli di integrità sui dati, e il DBMS garantirà tali vincoli.
 - ▶ Al di là di questo, il DBMS non capisce realmente la semantica dei dati (ad esempio non capisce come sono calcolati gli interessi su un conto bancario).
 - ▶ Quindi, garantire che una transazione (eseguita da sola) conservi la consistenza dei dati è, in ultima analisi, responsabilità dell'utente!

Scheduling di transazioni concorrenti

- ▶ Il DBMS garantisce che l'esecuzione di $\{T_1, \dots, T_n\}$ è equivalente a una qualche esecuzione seriale $T_1' \dots T_n'$
 - ▶ Prima di leggere/scrivere un oggetto, una transazione richiede un *lock* sull'oggetto, e aspetta finché il DBMS fornisce tale *lock*. Tutti i *lock* sono rilasciati al termine della transazione (protocollo di *locking* **Strict 2PL**).
 - ▶ **Idea:** se una azione di T_i (ad esempio, scrivere X) influenza T_j (che magari legge X), una tra T_i e T_j , diciamo T_i , otterrà il blocco su X per prima, e T_j dovrà aspettare fino al termine di T_i ; ciò in pratica dà un ordinamento alle transazioni.
 - ▶ Che succede se T_j ha già un blocco su Y e T_i in seguito richiede un blocco su Y ?
 - ▶ **(Deadlock!)** T_i o T_j deve essere interrotta e fatta ripartire!

Atomicity

- ▶ Il DBMS garantisce *l'atomicità* (una transazione è una unità indivisibile - proprietà tutto-o-niente) anche se il sistema va in *crash* durante una transazione.
- ▶ **Idea:** mantenere un log (una storia) di tutte le azioni effettuate dal DBMS nell'eseguire un insieme di transazioni:
 - ▶ prima che una modifica sia fatta alla base di dati, la voce corrispondente del log viene messa al sicuro (**protocollo WAL**: spesso il supporto del sistema operativo in questo ambito è inadeguato).
 - ▶ dopo un crash, gli effetti delle transazioni eseguite parzialmente sono annullati usando il log.

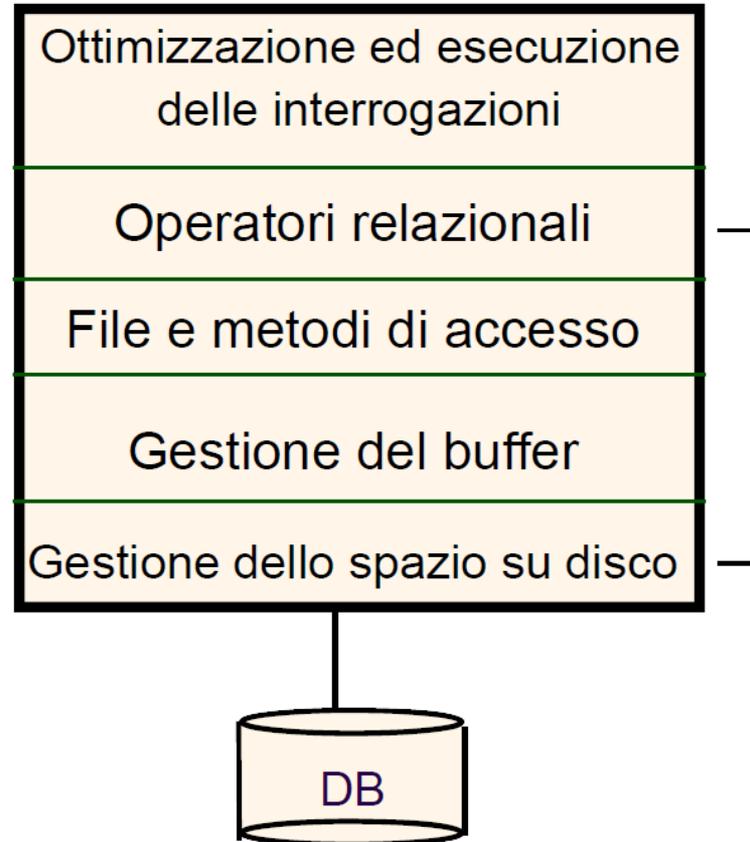
Il Log

- ▶ Le seguenti azioni sono registrate nel log:
 - ▶ Ti scrive un oggetto: il vecchio valore e il nuovo valore.
- ▶ Il record nel log deve essere salvato su disco prima della pagina modificata.
 - ▶ Ti termina/si interrompe: un record nel log registra tale azione.
- ▶ I record nel log sono collegati tramite l'ID della transazione, così che sia semplice annullare una transazione specificata (ad esempio per risolvere un deadlock).
- ▶ Il log è spesso duplicato e archiviato in una memoria “stabile”.
- ▶ Tutte le attività connesse al log (e di fatto tutte le attività connesse al controllo di concorrenza, come blocco/sblocco, gestione deadlock ecc.) sono gestite dei deadlock, dal DBMS in maniera trasparente.

Struttura di un DBMS

- ▶ Un tipico DBMS ha una struttura stratificata.
- ▶ La figura non mostra le componenti di controllo di concorrenza e di ripristino.
- ▶ Questa è una tra le possibili architetture; ciascun sistema ha le proprie varianti.

Questi strati devono tenere in considerazione il controllo di concorrenza e ripristino.



Linguaggi delle Basi di Dati

- ▶ I DBMS sono caratterizzati dalla presenza di molteplici linguaggi per la gestione dei dati.
- ▶ L'accesso ai dati può avvenire:
 - ▶ con **linguaggi testuali interattivi (ad es. SQL)**.
 - ▶ con comandi (come quelli del linguaggio interattivo) immersi in un **linguaggio ospite (Java, C, Cobol, etc.)**.
 - ▶ con comandi (come quelli del linguaggio interattivo) immersi in un **linguaggio ad hoc**, con anche altre funzionalità (ad es. per grafici o stampe strutturate), anche con l'ausilio di strumenti di sviluppo (ad es. per la gestione di maschere).
 - ▶ con **interfacce user friendly** (senza linguaggio testuale).

SQL, un linguaggio interattivo

Corsi

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

Aule

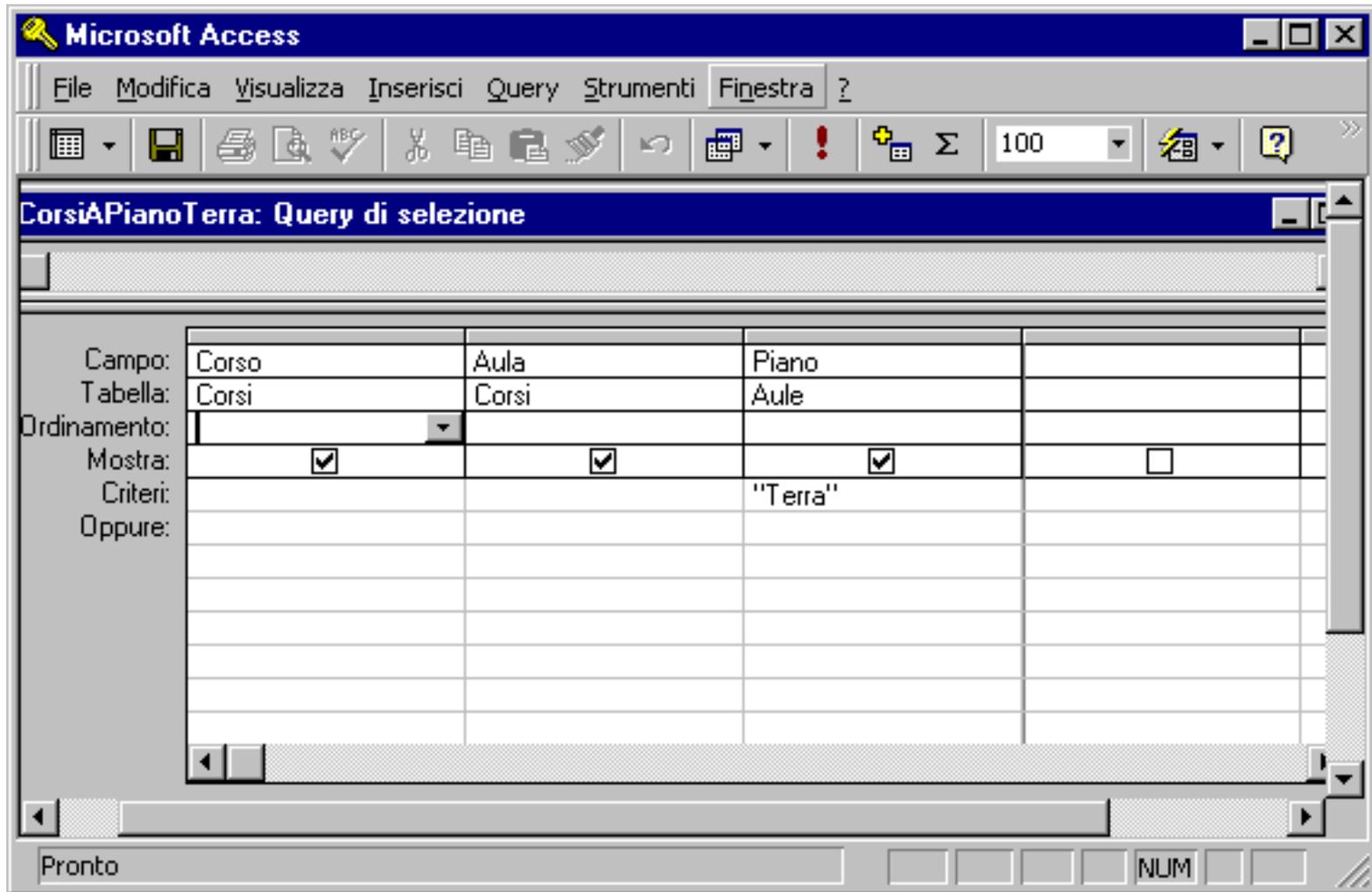
Nome	Edificio	Piano
DS1	OMI	Terra
N3	OMI	Terra
G	Pincherle	Primo

"Trovare i corsi tenuti in aule a piano terra"

```
SELECT Corso, Aula, Piano
FROM Aule, Corsi
WHERE Aule.Nome = Corsi.Aula
AND
Aule.Piano = "Terra"
```

Corso	Aula	Piano
Sistemi	N3	Terra
Reti	N3	Terra

Interazione non testuale (MS Access)



Una distinzione (separazione fra dati e programmi)

▶ data manipulation language (DML)

- per l'interrogazione e l'aggiornamento di (istanze di) basi di dati.

▶ data definition language (DDL)

- per la definizione di schemi (logici, esterni, fisici) e altre operazioni del genere.

Un'operazione DDL (sullo schema)

Corsi

```
CREATE TABLE corsi (  
  corso          CHAR(20) ,  
  docente        CHAR(20) ,  
  aula           CHAR(4))
```

Corso	Docente	Aula
Basi di dati	Rossi	DS3
Sistemi	Neri	N3
Reti	Bruni	N3
Controlli	Bruni	G

Persone che lavorano con le Basi di Dati

▶ **Progettisti e programmatori di DBMS**

- costruiscono il software dei DBMS.

▶ **Utenti**

- **utenti finali** (terminalisti): eseguono programmi che realizzano attività predefinite e di frequenza elevata, con poche eccezioni previste a priori.
- **utenti casuali**: eseguono operazioni non predefinite, non previste a priori, usando linguaggi interattivi per formulare interrogazioni (o aggiornamenti) di vario tipo.

▶ **Progettisti e programmatori di applicazioni**

- sviluppano pacchetti software che facilitano agli utenti finali l'accesso ai dati.
- Utilizzano principalmente il linguaggio di manipolazione dei dati (DML).
- i programmi applicativi dovrebbero idealmente accedere ai dati tramite lo schema esterno (si potrebbero scrivere applicazioni che accedono ad un livello più basso, ma ciò potrebbe compromettere l'indipendenza dei dati).

Persone che lavorano con le Basi di Dati

► **Amministratori della base di dati (DBA)**

- la **progettazione** e la **manutenzione** di Basi di Dati a livello di impresa viene affidata ad un professionista, responsabile di molti compiti critici :
 - **Progettazione degli schemi fisico e logico** : Il DBA è responsabile dell'interazione con gli utenti del sistema per capire quali dati devono essere memorizzati nel DBMS (progetto dello schema logico) e dove memorizzarli (progetto dello schema fisico). Il DBA può anche progettare porzioni largamente usate della schema esterno.
 - **Sicurezza e Autorizzazioni** : Il DBA è responsabile nel garantire che l'accesso non autorizzato ai dati non venga consentito. In genere, agli utenti possono essere concessi permessi per accedere solo ad alcune viste e relazioni.
 - **Disponibilità dei dati e ripristino dai guasti** : Il DBA deve garantire che se il sistema si guasta gli utenti possano continuare ad accedere ai dati non corrotti. Il DBA deve anche lavorare per riportare i dati ad uno stato consistente, attraverso un salvataggio periodico dei dati e del mantenimento di log dell'attività del sistema.
 - **Messa a punto della Base di Dati** : Dato che i bisogni degli utenti si evolvono nel tempo, il DBA è responsabile di ogni modifica alla base di dati (agli schemi logico e fisico) per garantire prestazioni adeguate al cambiamento dei requisiti richiesti.

Vantaggi e Svantaggi dei DBMS

Pro

- dati come risorsa comune di un'organizzazione, a disposizione (con opportune forme di controllo) di tutte le sue componenti.
- schema dei dati come modello unificato e preciso della realtà di interesse per l'organizzazione.
- gestione centralizzata dei dati, con riduzione di ridondanze e incoerenze.
- indipendenza dei dati (favorisce lo sviluppo di applicazioni più flessibili e modificabili).
- ▶ integrità e sicurezza dei dati, attraverso *vincoli di integrità e controlli di accesso*.
- ▶ ripristino dai crash, proteggendo gli utenti dagli effetti dei guasti del sistema.

Vantaggi e Svantaggi dei DBMS

Contro

- I DBMS sono prodotti complessi e costosi. La loro introduzione comporta notevoli investimenti diretti (acquisto del prodotto) e indiretti (acquisizione delle risorse hardware e software necessarie, conversione delle applicazioni, formazione del personale).
- I DBMS forniscono una serie di servizi, necessariamente associati ad un costo. Se alcuni servizi non sono più necessari, è difficile scorporare quelli effettivamente richiesti dagli altri, e ciò può comportare una riduzione di prestazioni.
- Applicazioni con pochi utenti e senza necessità di accessi concorrenti possono essere realizzate più proficuamente con file ordinari piuttosto che con DBMS.

Esercizi su Basi di dati e DBMS

Esercizio

Un giardiniere vuole mantenere delle informazioni (nome, descrizione, anno di acquisto, ...) sulle piante che ha nella sua serra personale.

L'elevato numero di piante che ha a disposizione lo spinge a comprare un sistema di basi di dati. Per risparmiare denaro, ne vorrebbe acquistare uno con il minor numero di funzioni, e prevede di usarlo come applicazione *stand-alone* sul suo computer. Ovviamente il giardiniere non prevede di condividere i dati con nessuno. Indicate quali delle seguenti funzioni il giardiniere dovrebbe comprare :

- ▶ *controllo di concorrenza*
- ▶ *ripristino da crash*
- ▶ *un meccanismo per le viste*
- ▶ *un linguaggio di interrogazione*

Quali delle seguenti affermazioni sono vere?

- ▶ L'indipendenza dei dati permette di scrivere programmi senza conoscere le strutture fisiche dei dati.
- ▶ L'indipendenza dei dati permette di modificare le strutture fisiche dei dati senza dover modificare i programmi che accedono alla base di dati.
- ▶ L'indipendenza dei dati permette di scrivere programmi conoscendo solo lo schema concettuale della base di dati.
- ▶ L'indipendenza dei dati permette di formulare interrogazioni senza conoscere le strutture fisiche.

Quali delle seguenti affermazioni sono vere?

- ▶ Il fatto che le basi di dati siano condivise favorisce l'efficienza dei programmi che le utilizzano.
- ▶ Il fatto che le basi di dati siano condivise permette di ridurre ridondanze e inconsistenze.
- ▶ Il fatto che le basi di dati siano persistenti ne garantisce l'affidabilità.
- ▶ Il fatto che le basi di dati siano persistenti favorisce l'efficienza dei programmi.
- ▶ Il fatto che le basi di dati siano condivise rende necessaria la gestione della privacy e delle autorizzazioni.

Quali delle seguenti affermazioni sono vere?

- ▶ La distinzione fra DDL e DML corrisponde alla distinzione fra schema e istanza.
- ▶ Le istruzioni DML permettono di interrogare la base di dati ma non di modificarla.
- ▶ Le istruzioni DDL permettono di specificare la struttura della base di dati ma non di modificarla.
- ▶ Non esistono linguaggi che includono sia istruzioni DDL sia istruzioni DML.
- ▶ SQL include istruzioni DML e DDL.
- ▶ Le istruzioni DML permettono di interrogare la base di dati e di modificarla.

Quali delle seguenti affermazioni sono vere?

- ▶ Gli utenti casuali utilizzano programmi predefiniti.
- ▶ I terminalisti utilizzano programmi predefiniti.
- ▶ Gli utenti casuali progettano la base di dati.
- ▶ I progettisti del DBMS realizzano i programmi che saranno utilizzate dai terminalisti.
- ▶ Gli amministratori della base di dati realizzano il DBMS.

Domande a risposta aperta

- ▶ Illustrare, in modo sintetico ma chiaro, supponendo di rivolgersi ad un non esperto, le caratteristiche fondamentali delle basi di dati e il ruolo che esse giocano nei sistemi informativi.
- ▶ Discutere brevemente la seguente affermazione: *“i dati sono una risorsa per una organizzazione, e come tali vanno considerati anche separatamente dalle applicazioni che li utilizzano”*.
- ▶ Illustrare brevemente (non più di mezza pagina) il concetto di indipendenza dei dati.