

Introducció a les bases de dades

Índex

1. Introducció
2. Arquitectura dels SGBD
3. Models, llenguatges i administració de BD

1. Introducció

BD i SGBD

Anomenem **base de dades** (BD) als conjunts de fitxers interrelacionats, amb estructures complexes i compartits per diversos processos simultàniament

Una **base de dades d'un SI** és la representació integrada dels conjunts d'entitats instància corresponents a les diferents entitats tipus del SI i de les seves interrelacions. Aquesta representació informàtica ha de poder ser utilitzada de manera compartida per molts usuaris de tipus diversos.

El software especialitzat en la gestió de les bases de dades es denomina **sistema de gestió de bases de dades** (SGBD).

Objectius i serveis dels SGBD

- Permetre la realització de consultes **no predefinides** i complexes.
- Oferir **flexibilitat als canvis** (independència).
- Facilitar l'eliminació i la gestió de la **redundància**.
- Assegurar el manteniment de la **qualitat** de les dades (integritat).
- Gestionar la **concurrència** d'accés a les dades.
- Proveir serveis de **seguretat**.

2. Arquitectura dels SGBD

L'esquema de la DB

Els SGBD necessiten que els donem una descripció o definició de la BD, descripció que rep el nom d'**esquema** de la BD, i que els SGBD tindran contínuament a l'abast.

Anteriorment, ja hem parlat de la distinció entre dos nivells de representació informàtica:

- El **nivell lògic** ens oculta els detalls de com s'emmagatzemen les dades, com es mantenen i com s'accedeix físicament a ells. En aquest nivell només es parla d'entitats, atributs i regles d'integritat.
- Per qüestions de rendiment, ens podrà interessar descriure elements de **nivell físic** com quins índexs tindrem, com i on volem que s'agrupin físicament els registres, de quina mida han de ser les pàgines, etc.

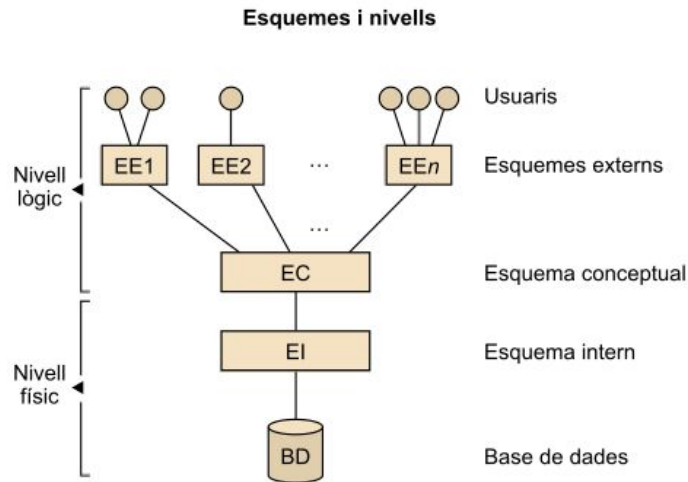
Arquitectura ANSI/SPARC: equivalències

D'acord amb l'arquitectura **ANSI/SPARC**, hi havia d'haver **tres nivells** d'esquemes, tres nivells d'abstracció. La idea bàsica d'ANSI/SPARC consistia a descompondre el nivell lògic en dos: el nivell **extern** i el nivell **conceptual**. Anomenaven nivell **intern** el que aquí hem anomenat nivell físic.

Anteriorment	ANSI/SPARC
Lògic	Extern
	Conceptual
Físic	Intern

Arquitectura ANSI/SPARC: els tres nivells I

- Al nivell extern se situen les diferents visions lògiques que els processos usuaris (programes d'aplicació i usuaris directes) tindran de les parts de la BD que utilitzaran. Aquestes visions s'anomenen **esquemes externs**.
- Al nivell conceptual hi ha una sola descripció lògica bàsica, única i global, que anomenem **esquema conceptual**, i que serveix de referència per a la resta d'esquemes.
- Al nivell físic hi ha una única descripció física, que anomenem **esquema intern**.



Arquitectura ANSI/SPARC: els tres nivells II

A l'**esquema conceptual**, s'hi descriuran les entitats tipus, els seus atributs, les interrelacions i també les restriccions o regles d'integritat.

En definir un **esquema extern**, se citaran només aquells atributs i entitats que interessin, els podrem reanomenar, podrem definir dades derivades, podrem redefinir una entitat perquè les aplicacions que utilitzen aquest esquema extern creguin que en són dos, o podrem definir combinacions d'entitats perquè en semblin una de sola, etc.

L'**esquema intern** o físic contindrà la descripció de l'organització física de la BD: camins d'accés (índexs, hashing, apuntadors...), codificació de les dades, gestió de l'espai, mida de la pàgina, etc.

Independència de les dades

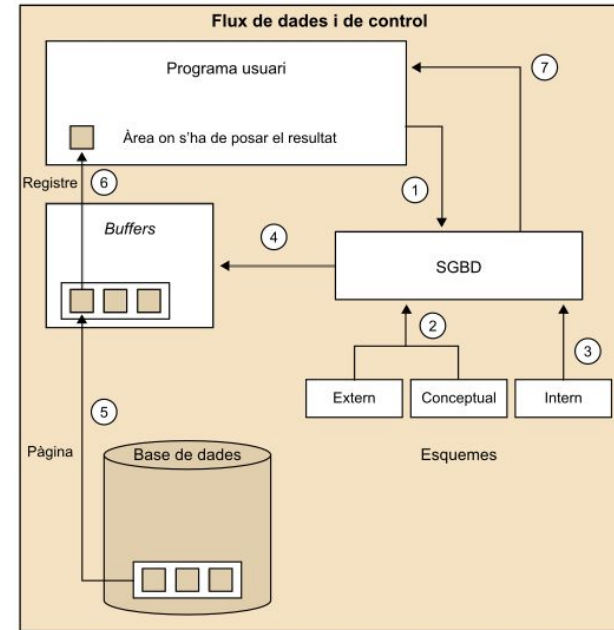
Hi ha **independència física** quan els canvis en l'organització física de la BD no afecten el món exterior (és a dir, els programes usuaris o els usuaris directes).

Hi ha **independència lògica** quan els usuaris no es veuen afectats pels canvis en el nivell lògic.

Flux de dades i de control

Els principals passos de l'execució d'una consulta enviada al SGBD per un programa d'aplicació són:

- A. El programa envia la consulta a l'SGBD (1) que consulta els esquemes conceptual i extern per validar-la (2).
- B. El SGBD determina com executar la consulta consultant l'esquema intern (3).
- C. El SGBD comprova si la pàgina d'interès està en els buffers i, si no ho està, la càrrega des del disc (5).
- D. El SGBD còpia les dades a l'àrea de treball del programa (6) i retorna el control a aquest (7).



3. Models, llenguatges i administració de BD

Models de BD

El conjunt de components o eines conceptuais que un SGBD proporciona per modelar rep el nom de **model de BD**.

Alguns exemples de models de BD són el model **relacional**, el model **jeràrquic**, el model en **xarxa** i el model **relacional amb objectes**.

Els models de BD ens proporcionen tres tipus d'eines:

- **Estructures de dades** amb les quals es pot construir la BD: taules, arbres, etc.
- Diferents tipus de **restriccions** (o regles) d'integritat que l'SGBD haurà de fer complir a les dades: dominis, claus, etc.
- Una sèrie d'**operacions** per treballar amb les dades.

Llenguatges

Per comunicar-se amb l'SGBD, l'usuari, ja sigui un programa d'aplicació o un usuari directe, es val d'un **llenguatge**. Hi ha molts llenguatges diferents, segons el tipus d'usuaris per als que estan pensats (experts o no informàtics) i el tipus de coses que els usuaris han de poder expressar amb ells (expressions complexes o simples).

- Hi ha llenguatges especialitzats en l'escriptura d'esquemes; és a dir, en la descripció de la BD. Es coneixen genèricament com **DDL o Data Definition Language**.
- Altres llenguatges estan especialitzats en la utilització de la BD (consultes i manteniment). Es coneixen com **DML o Data Management Language**.
- No obstant això, el més freqüent és que el mateix llenguatge disposi de construccions per a les dues funcions, DDL i DML.
 - Per exemple, SQL té instruccions DDL (com CREATE TABLE), instruccions DML com (SELECT) i addicionalment instruccions de control d'entorn (p.e. COMMIT).

Administració de bases de dades

Una empresa o institució que tingui SI construïts al voltant de BD necessita que algú porti a terme una sèrie de funcions centralitzades de gestió i administració, per assegurar que l'explotació de la BD és la correcta. Aquest conjunt de funcions es coneix amb el nom d'administració de BD, i els usuaris que fan aquest tipus especial de treball es denominen **administradors de BD** (ABD).

Un administrador de base de dades s'encarrega de tasques com:

- Manteniment, administració i control dels esquemes.
- Assegurar la màxima disponibilitat de les dades (backups, logs, etc.)
- Resolució d'emergències.
- Vigilància de la integritat i de la qualitat de les dades.
- Disseny físic, estratègia de camins d'accés i reestructuracions.
- Control del rendiment.
- Normativa i assessorament als programadors i als usuaris finals.
- Control i administració de la seguretat: autoritzacions, restriccions, etc.

