

Modelo de Datos

Niveles de abstracción de una BD

- Percepción de una BD según el tipo de usuario (final, especialista o administrador)
- Arquitectura ANSI/SPARC*
- Base para la *independencia de datos*

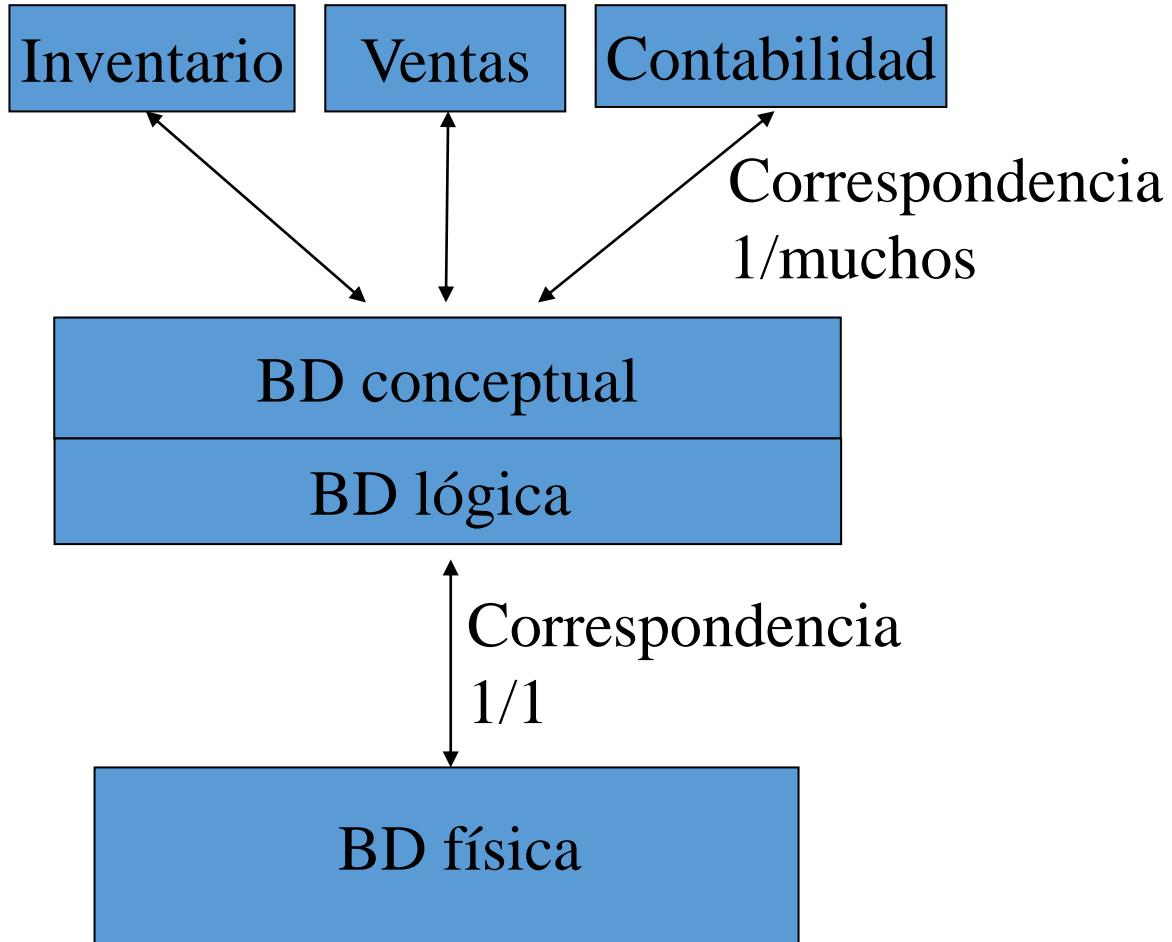
* *American National Standards Institute/Systems Planification
and Requirements Committee*

Arquitectura ANSI/SPARC

Nivel de visión o externo
(vistas parciales
de la BD)

Nivel conceptual/lógico
(vista global
de la BD)

Nivel físico
(almacenamiento
de la BD)

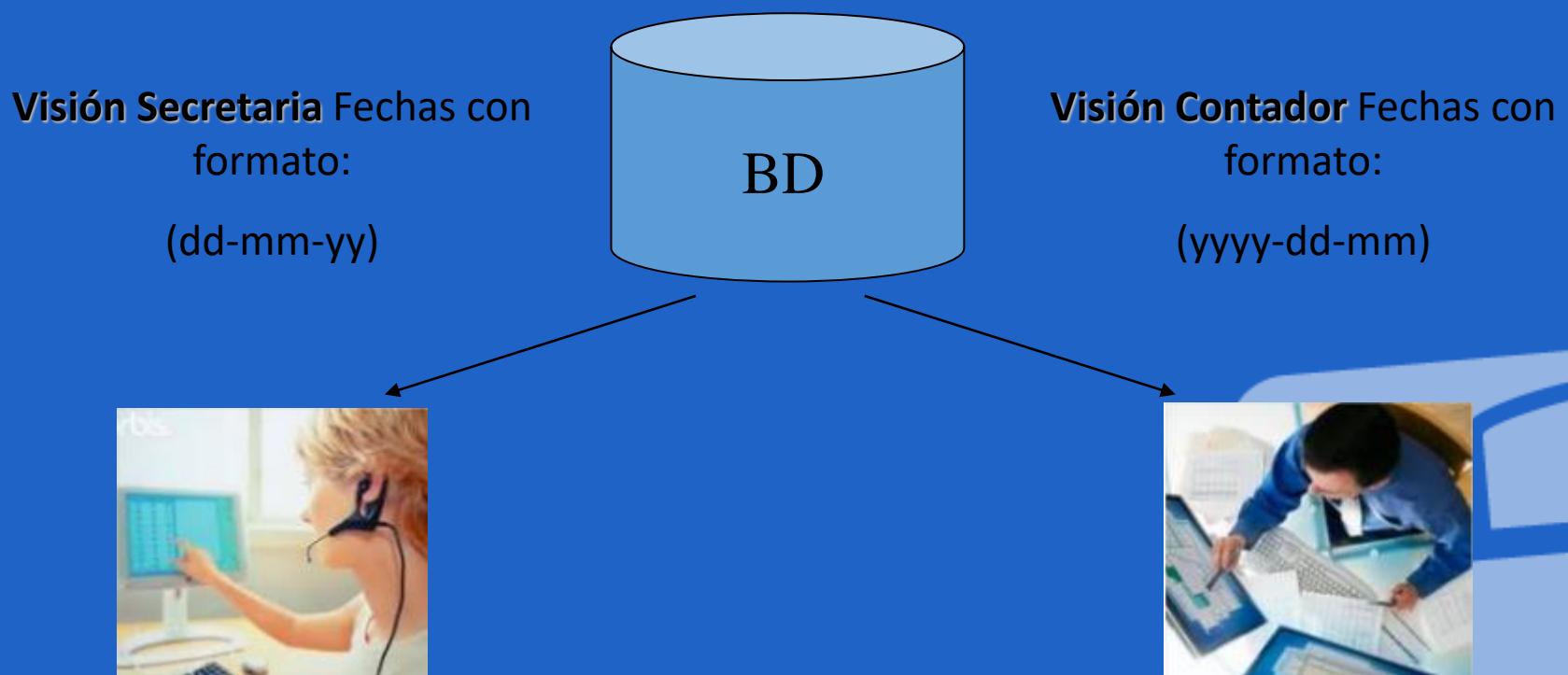


Nivel de visión/externo

- El más cercano a los **usuarios finales**
- **Percepción** de la BD por parte de los usuarios finales
- Tantas visiones como tipos de usuarios haya
- Cada visión de usuario final se puede considerar como un *subesquema*

Nivel de visión

Cada visión ofrece su representación de los datos de la BD



Nivel de visión

Una *visión* de usuario podría incluir:

- Datos Agregados: *Totales por Dpto*
- Datos Derivados:

$$\text{Sueldo total} = \text{básico} + \text{comisión}$$

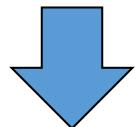
- Datos Calculados:

$$\text{Edad de una persona (inferida a partir de su fecha de nacimiento)}$$

Note que en una BD *convencional* este tipo de datos *posiblemente* no están almacenados explícitamente en la BD

BD

Cédula	Nombre	Salario	Comisión	Dpto
712123	Alesha Dixon	100	5	Ciencias
435422	Madonna	50	10	Ciencias
342311	Björk	100	5	Medicina
898765	Lorde	200	10	Medicina
231452	Cheryl Cole	50	5	Ciencias
546897	Bruno Mars	100	5	Artes
717234	Boy George	50	20	Artes



Total	Dpto
220	Ciencias
315	Medicina
175	Artes

Total sueldo
por Dpto

Visión de
un usuario

Salario + Comisión

Nivel de visión

- Los conceptos de dato agregado, derivado y calculado significan, en general, que un dato se genera a partir de otro(s)
- Puede haber a su vez:
 - datos agregados derivados
 - datos agregados calculados
 - etc.

Nivel conceptual

- Nivel mediador entre los otros dos niveles
- Interesante para el **usuario especialista**
- Se ocupa de los datos almacenados en la BD física y las *relaciones* entre ellos
- Descripción semántica de los datos que conforman la BD
- Soporta a cada visión de usuario externa

Nivel conceptual

- Es una visión completa de todos los *requisitos* y elementos de interés para la organización
- Incluye *restricciones* sobre los datos
- La descripción del nivel conceptual **no debe tener detalles dependientes del almacenamiento**
- Es un lenguaje de muy alto nivel
- El nivel lógico es un refinamiento (ofrece más detalles) que el modelo conceptual

Nivel físico

- Interno
- Más cercano a la máquina
- Interesa al **usuario administrador** y al **especialista**
- Esquema físico: Descripción y *tipos de datos* (tamaño y precisión), tipos de índices y de estructuras de almacenamiento usadas, de acuerdo con un *SGBD particular*

Nivel físico

- Describe **cómo** se almacenan los datos en términos de estructuras de datos específicas
- Se encarga de:
 - Reservar espacio para los datos e índices
 - Comprimir los datos
 - Encriptar los datos

Independencia de los datos

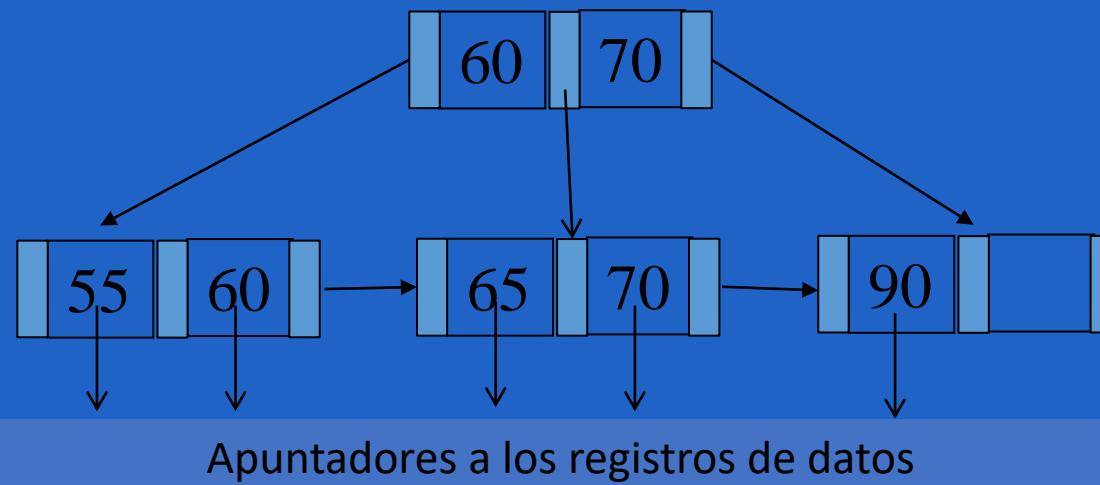
- Es uno de los objetivos de la arquitectura ANSI/SPARC
- Permite modificar la definición de un nivel sin afectar (en lo posible) el nivel inmediatamente superior
- Sin independencia de datos se requeriría mucho esfuerzo para cambiar las aplicaciones de tal forma que se adapten al nuevo esquema de la BD.
- Hay dos tipos: física y lógica

Independencia física

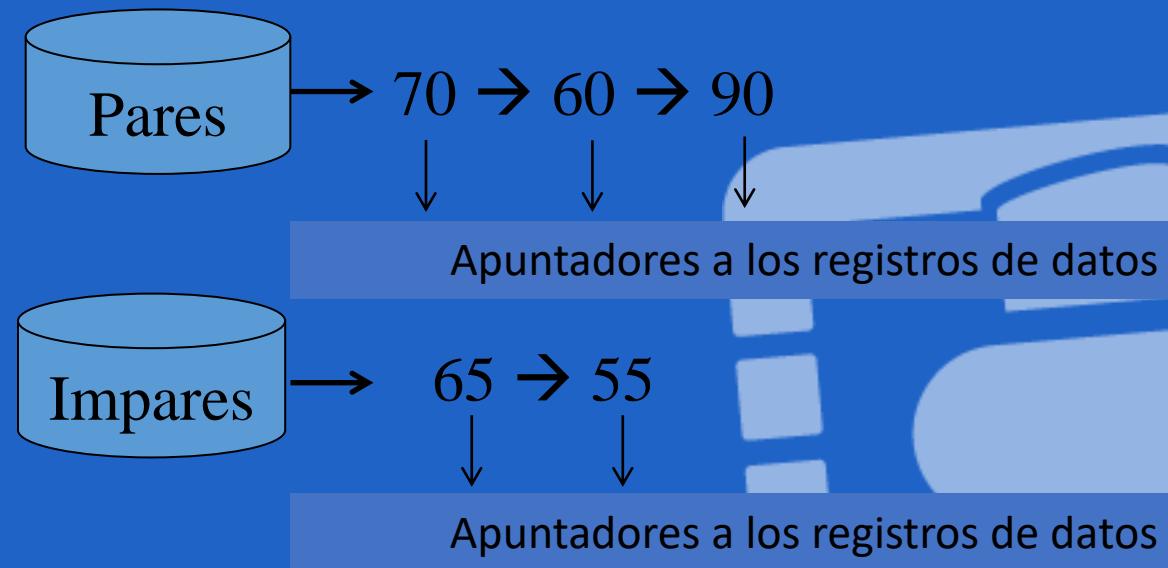
- Se presenta entre el nivel conceptual y el nivel físico
- Un cambio en el esquema físico (e.g., usar otras estructuras de almacenamiento) no conduce a cambios en el esquema conceptual

Ej. Cambiar el tipo de índice de un atributo: *B+* por *hash*, ¿con qué propósito se hace este tipo de cambios?
- **Inmunidad** del esquema conceptual ante cambios del esquema físico

Árbol B+



Función *Hash*:
por simplicidad
supóngase:
Pares e Impares



Independencia lógica

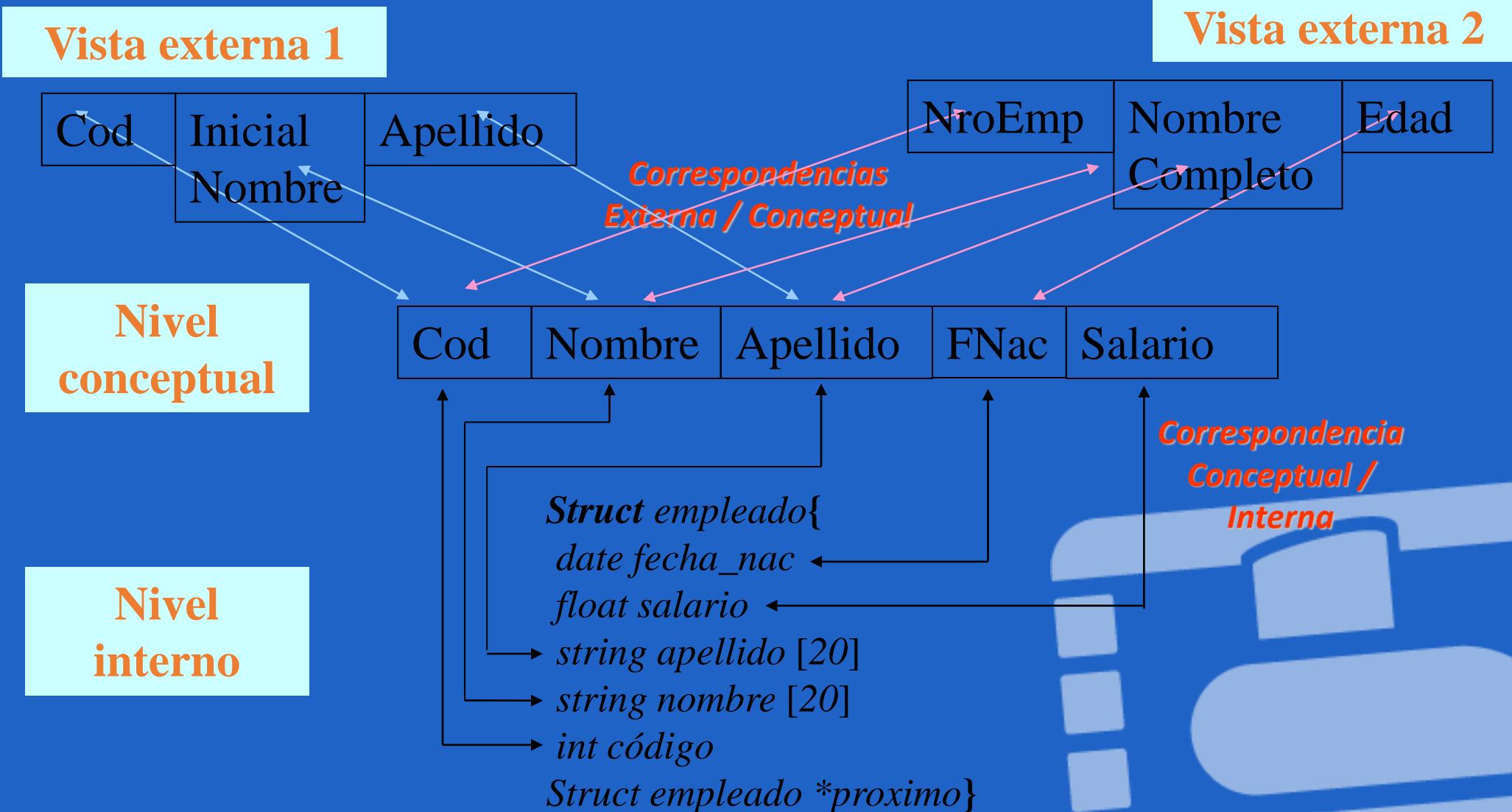
- Se presenta entre el nivel de visión y el nivel conceptual
- Significa que un cambio en el nivel conceptual no debe implicar un cambio en el nivel de visión
- Es más difícil de lograr. ¿Por qué?

Independencia lógica

Algunos de los posibles cambios en el nivel conceptual:

- Adición de elementos (atributos, entidades, etc.) →
Ej. La adición de un atributo **obligatorio** puede afectar a los subesquemas externos
- Eliminación de elementos → La eliminación de un atributo puede afectar a los subesquemas externos

Diferencias entre los niveles



- La descripción completa de una BD se denomina *esquema*
- Cada visión de usuario tiene su esquema (subesquema), existe **un** esquema conceptual/lógico y **uno** interno
- Existen correspondencias (*mappings*):
 - entre cada subesquema externo y el conceptual
 - entre el esquema conceptual y el interno
- Estas correspondencias le permiten al SGBD saber, e.g., un elemento conceptual con cual elemento del nivel físico se corresponde

- En *teoría* debería existir un lenguaje de descripción para cada esquema (nivel)*
 - Sin embargo, en la práctica en muchos SGBD el esquema lógico queda fusionado con la especificación del esquema interno
- * No siempre sucede así en los SGBD actuales...

Concepto de modelo

- La realidad concreta, con todos sus detalles no se puede captar y hay detalles que NO interesan:


Captar lo que le interesa a la organización
- N observadores de un fenómeno tendrán al menos N percepciones, posiblemente distintas, del mismo fenómeno
- Modelo: Herramienta para comunicar y plasmar la representación de algún fenómeno del mundo real
- Esa representación ayuda a la comprensión del fenómeno