

BASES DE DATOS

Introducción

C. J. Date Cap. 1 y 2



Agenda



- ❑ Definición.
- ❑ Historia
- ❑ Componentes
- ❑ Usuarios. Tipos.
- ❑ Tipos de BD
- ❑ Ventajas
- ❑ Arquitectura ANSI/SPARC
- ❑ Preguntas, Ejercicios

Autor Hernán Puelman



Definición

BASE DE DATOS (BD): *Sistema de software cuyo objetivo es almacenar datos y permitirle a los usuarios realizar operaciones sobre dichos datos.*

En el próximo *slide* veremos una BASE DE DATOS (BD) muy pequeña, de una sola tabla y algunas operaciones que se pueden realizar sobre ella.



Definición

NOTEBOOKS

#ID	MARCA	MODELO	RAM	FECHA FABRICACIÓN	SSD
1	Lenovo	Thinkpad	16	01/04/2022	S
2	Dell	Inspiron	12	04/04/2020	N
5	Lenovo	Yoga	24	10/01/2021	S
33	Samsung	Galaxy	16	01/12/2022	S
11	HP	Pavilion	8	11/05/2022	N
47	Apple	Mac Pro	32	20/12/2022	S

Operaciones

NOTEBOOKS

#ID	MARCA	MODELO	RAM	FECHA FABRICACIÓN	SSD
1	Lenovo	Thinkpad	16	01/04/2022	S
2	Dell	Inspiron	12	04/04/2020	N
5	Lenovo	Yoga	24	10/01/2021	S
11	HP	Pavilion	8	11/05/2022	N
33	Samsung	Galaxy	16	01/12/2022	S
47	Apple	Mac Pro	32	20/12/2022	S
53	Bangho	Bacis	8	06/01/2021	

**INSERT INTO Notebooks (#id, marca, modelo, RAM, FechaFabricacion)
VALUES (53, 'Bangho', 'Bacis', 8, '2021/01/06')**

Operaciones

NOTEBOOKS

#ID	MARCA	MODELO	RAM	FECHA FABRICACIÓN	SSD
1	Lenovo	Thinkpad	16	01/04/2022	S
2	Dell	Inspiron	12	04/04/2020	N
5	Lenovo	Yoga	24	10/01/2021	S
11	HP	Pavilion	8	11/05/2022	N
33	Samsung	Galaxy	16	01/12/2022	S
47	Apple	Mac Pro	32	20/12/2022	S
53	Bangho	Ban Pro	8	06/01/2021	null

UPDATE Notebooks

SET modelo = 'Ban Pro'

WHERE #id = 53;

Operaciones

NOTEBOOKS

#ID	MARCA	MODELO	RAM	FECHA FABRICACIÓN	SSD
1	Lenovo	Thinkpad	16	01/04/2022	S
2	Dell	Inspiron	12	04/04/2020	N
5	Lenovo	Yoga	24	10/01/2021	S
11	HP	Pavilion	8	11/05/2022	N
33	Samsung	Galaxy	16	01/12/2022	S
53	Bangho	Ban Pro	8	06/01/2021	null

DELETE Notebooks
WHERE #ID = 47;

Historia



Figura 4.8 Base de datos jerárquica. Estructura lógica y ejemplo



```

{
  "id": "ObjectID%795d55479ba749924a3e",
  "cliente": {
    "tipo": "Regular",
    "apellido": "Carvajal",
    "nombre": "Juan Alexander",
    "edad": "24",
    "sexo": "M",
    "telefono": "132-981234"
  },
  "tipoPago": "Contado",
  "fechaDePago": "15/04/2016 01:27:16:31:49Z",
  "item": {
    "album": "Peace Sells...But Who's Buying?",
    "anio": 1986,
    "cantidad": 1,
    "precio": 35000,
    "artista": "Megadeth"
  },
  "album": "Black Sabbath",
  "anio": 1979,
  "cantidad": 1,
  "precio": 45000,
  "artista": "Black Sabbath"
},
  "factura": {
    "numero": "0001",
    "total": "Heavy Metal",
    "trazo": "Metal"
  }
}
  
```

BD Jerárquicas (IMS IBM)

60's

BD Red (CODASYL) IDS

1969

Codd - Modelo Relacional

1970

System R

1974

Oracle

1979

Informix

1980

IBM-DB2

1981

ANSI

1986

ISO

1987

Sql Server

1989

MySQL, MariaDB, PostgreSQL

1990's

Neo4J

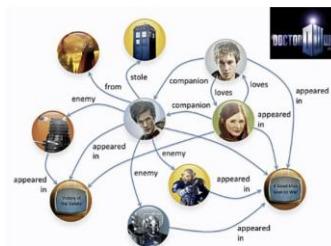
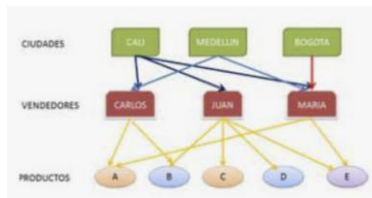
2007

Cassandra

2008

Redis, MongoDB

2009



Key	Value
K1	AAA,BBB,CCC
K2	AAA,BBB
K3	AAA,DDD
K4	AAA,2,01/01/2015
K5	3,ZZZ,5623

Componentes

DATOS

SOFTWARE

HARDWARE

USUARIOS



DATOS

“

- Información consistente (nula o mínima redundancia).
- Información compartida.
- Se percibe a la BD de diferentes formas (vistas).



SOFTWARE



- DBMS: Database Management System
- Interfaz entre la base de datos y los programas o usuarios finales
- Almacenamiento, lectura, actualización y borrado de información
- Accesos y control de la información
- Integridad
- Copias de seguridad



HARDWARE



- Servidores, clusters.
- Procesadores.
- Memoria
- Unidades de almacenamiento.



USUARIOS

- ❑ **Usuarios finales:** No saben de BD.
- ❑ **Programadores.**
- ❑ **Usuarios especializados:** Análisis de datos.
- ❑ **Analistas, Diseñadores de sistemas.** Requerimientos. Modelos.
- ❑ **Administrador de BD (DBA).** “Super usuario”



Usuarios - DBA



- ☐ Crear el nivel conceptual y físico.
- ☐ Privilegios y accesos
- ☐ Asistencia a desarrolladores
- ☐ Almacenamiento
- ☐ Instalaciones
- ☐ Monitoreo
- ☐ Mantenimiento
- ☐ Backup/Recovery



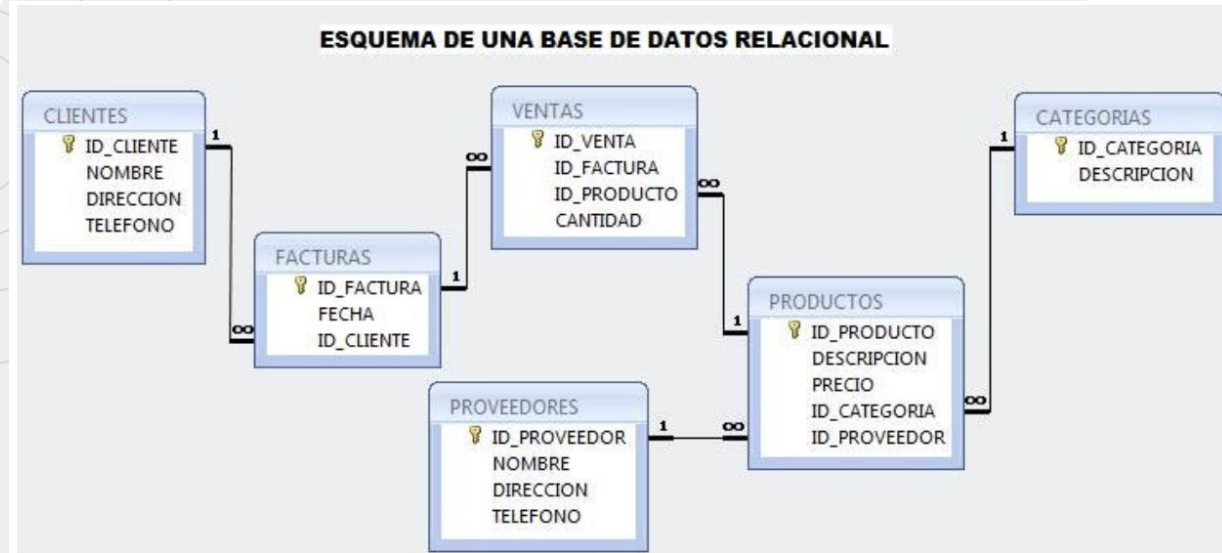
TIPOS DE BASES DE DATOS

- Relacionales
- No Sql
 - Clave/Valor
 - Documentales
 - Columnares
 - Grafos
- Otros (Time Series, OO, etc.)



Relacional

“



Oracle, MySQL, Sql Server, Postgres, etc.



Relacionales

“

Se utilizan para un sin número de casos.
Especialmente cuando se requiere manejo de transacciones y consistencia.
Ej. Aplicaciones que manejan dinero, stock en tiempo real, etc.



No Sql

“

- Datos almacenados que no requieren estructuras fijas
- Normalmente no garantizan consistencia [ACID](#)
- Escalan bien [horizontalmente](#).
- Mayor volumen de datos.
- También se las denomina "no solo SQL"



Clave/Valor

Key	Value
K1	AAA,BBB,CCC
K2	AAA,BBB
K3	AAA,DDD
K4	AAA,2,01/01/2015
K5	3,ZZZ,5623

KEY	VALUE
User:1	Jorge, Vieytes, 18-05-2013, \$100000, Ingeniero
User:222	Ana, Ferreyra, 03-10-2001, \$900000, Independiente
Prod:43	BAC5673 , A, 56, 19-10-2024, Azul, 467
Cat:A	A, móvil
Prod:97	PET89 , B2, 220, , , 3092
User:34	Vera, Torres, 22-08-1999, \$100000

Productos: Redis, MemCache, Riak KV

Clave/Valor

“

- Buenos tiempos de respuesta
- No implementan integridad.
- Tamaño pequeño
- Utilizadas para Cache y
- Manejo de sesiones.



Documentales

Colección

```
{  
  "id_": "123",  
  "nombre": "Juan",  
  "apellido": "Perez",  
  "edad": 54,  
  "estado civil": "S",  
  "domicilio": {  
    "calle": "Olavarria",  
    "nro": "238" },  
  "hobbies": [futbol, golf]  
}
```

```
{  
  "id_": "927",  
  "nombre": "Ana",  
  "apellido": "Saenz",  
  "edad": 25,  
  "domicilio": {  
    "calle": "Guemes",  
    "nro": "4321" },  
  "hobbies": [lectura, tiro]  
}
```

```
{  
  "id_": "333",  
  "nombre": "Vera",  
  "apellido": "Torres",  
  "edad": 55,  
}
```

Productos: MongoDB, CouchBase, etc.

Documentales

“

- Almacenan datos en forma de árbol.
- Son “*Schemaless*”
- Se utilizan para almacenar datos no estructurados (logs, entidades con propiedades diferentes, diferentes categorías).
- No suelen manejar transacciones.



Wide column

ID	Nombre	Edad	Sueldo
1	Juan	45	
10	Ana		2.000
23	Cecilia	54	3.500

ID	Columna	Valor
1	Nombre	Juan
1	Edad	45
10	Nombre	Ana
10	Sueldo	2.000
23	Nombre	Cecilia
23	Edad	54
23	Sueldo	3.500

Productos: Cassandra, Hbase.

Wide column

“

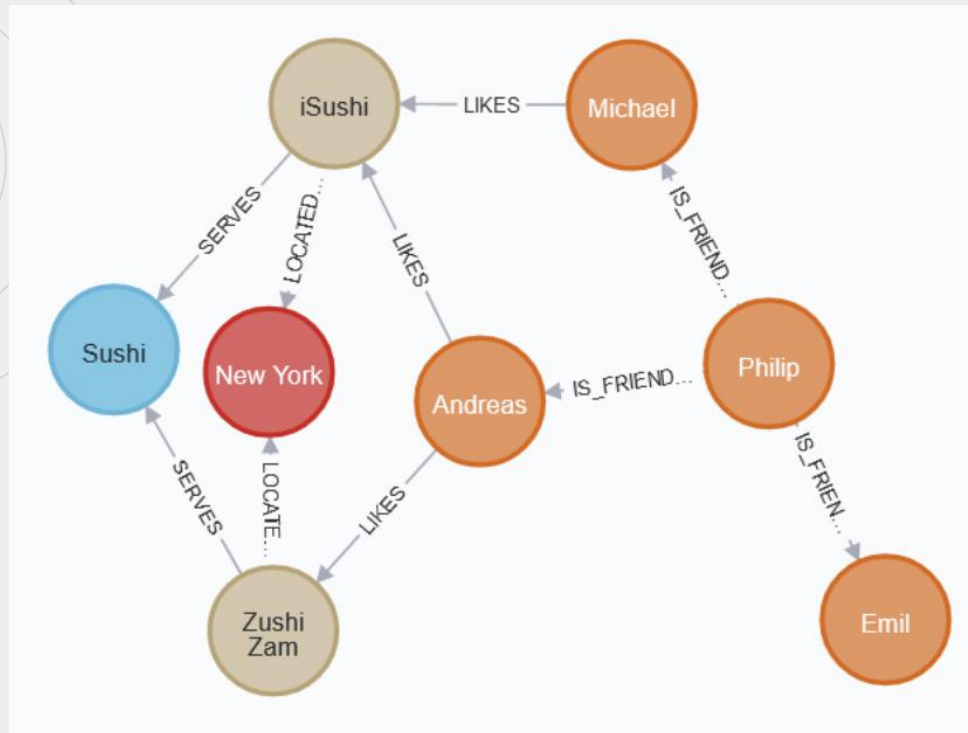
- Son ***schemaless***
- Solo almacenan las **columnas con datos reales.**
- Escalan horizontalmente.
- No existe integridad.
- No garantizan la consistencia
- No manejan Transacciones.

Productos: Cassandra, Hbase.



Grafos

“



Productos: Neo4j ...

Autor Hernán Puelman



Grafos



- Están formadas por nodos y arcos
- Redes sociales, logística, Detección de fraudes, recomendaciones, etc.

Productos: Neo4j ...

Autor Hernán Puelman



RDBMS

“

Relational Data Base Management System:

“BD que almacena datos relacionados entre si, representados por tablas con mínimo o nulo nivel de redundancia.”



Entidades, relaciones y Propiedades.

Entidades: Conjunto de objetos distinguibles y agrupados de los cuales se desea registrar información.

Relaciones: Vínculos entre las entidades.

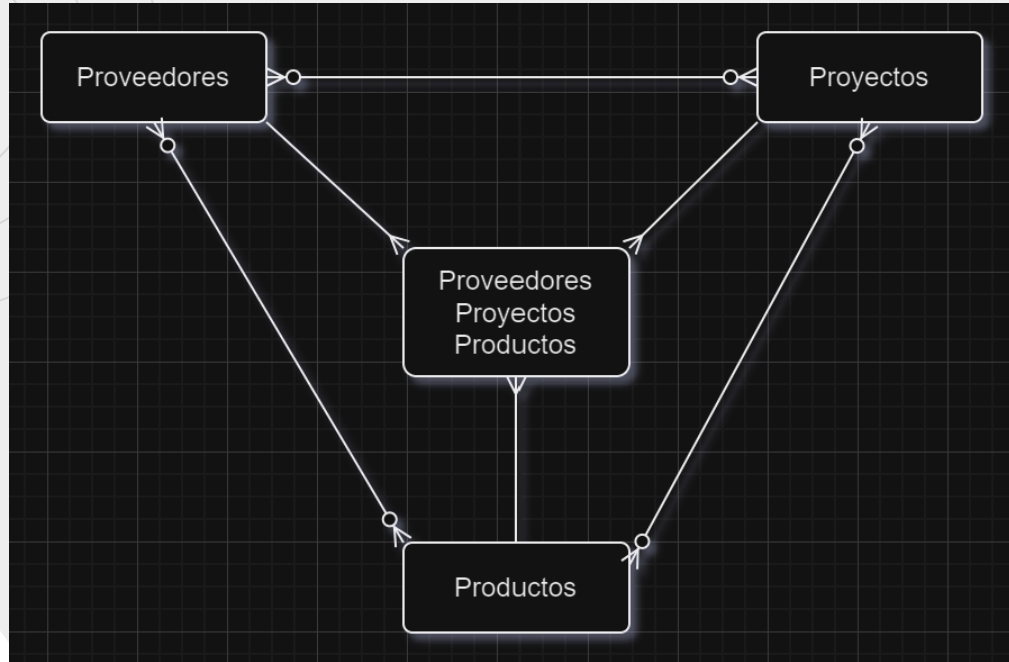
Propiedades: Datos relativos a las entidades. Simples y de un tipo específico.



Entidades, relaciones y Propiedades.

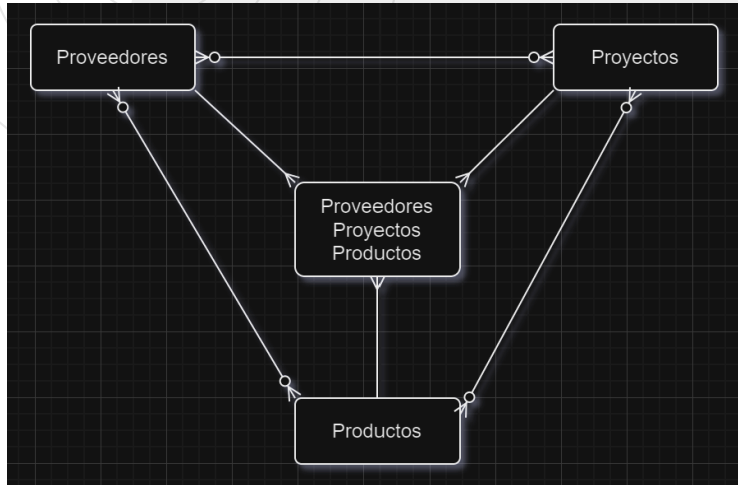
DER

“



Modelo de datos.

“ Lógico



Implementación

Físico



RDBMS – Ventajas



- Los datos pueden compartirse
- Control de la redundancia
- Minimizar la inconsistencia de datos
- Manejo de transacciones
- Integridad de datos
- Seguridad
- Independencia de Datos *(slide siguiente)*
- Lenguaje estándar
- Diferentes “visiones” para diferentes usuarios.



Independencia de Datos

“ **Independencia Lógica:** Inmunidad de las aplicaciones frente a cambios en las tablas de la aplicación.

Independencia Física: Inmunidad de las aplicaciones frente a cambios en la representación física y acceso a los datos.

En los sistemas de almacenamiento que no son Bases de Datos suelen coincidir los registros lógicos y los físicos.

RDBMS – Transacciones

“

*Conjunto de operaciones que **deben** (o no) realizarse en su totalidad.”*



RDBMS – Transacciones

“ *Propiedades*

A

Atomicidad

C

Consistencia

I

Aislamiento (Isolation)

D

Durabilidad



RDBMS – Transacciones



Atomicidad: Las transacciones son atómicas

Consistencia: La BD pasa de un estado consistente a otro consistente.

Aislamiento: Concurrencia. Las operaciones de una transacción están ocultas a las demás.

Durabilidad: Las actualizaciones perduran en la BD aunque haya una caída posterior.



Finalmente ... Qué es un sistema relacional?

“ *Es aquel que:*

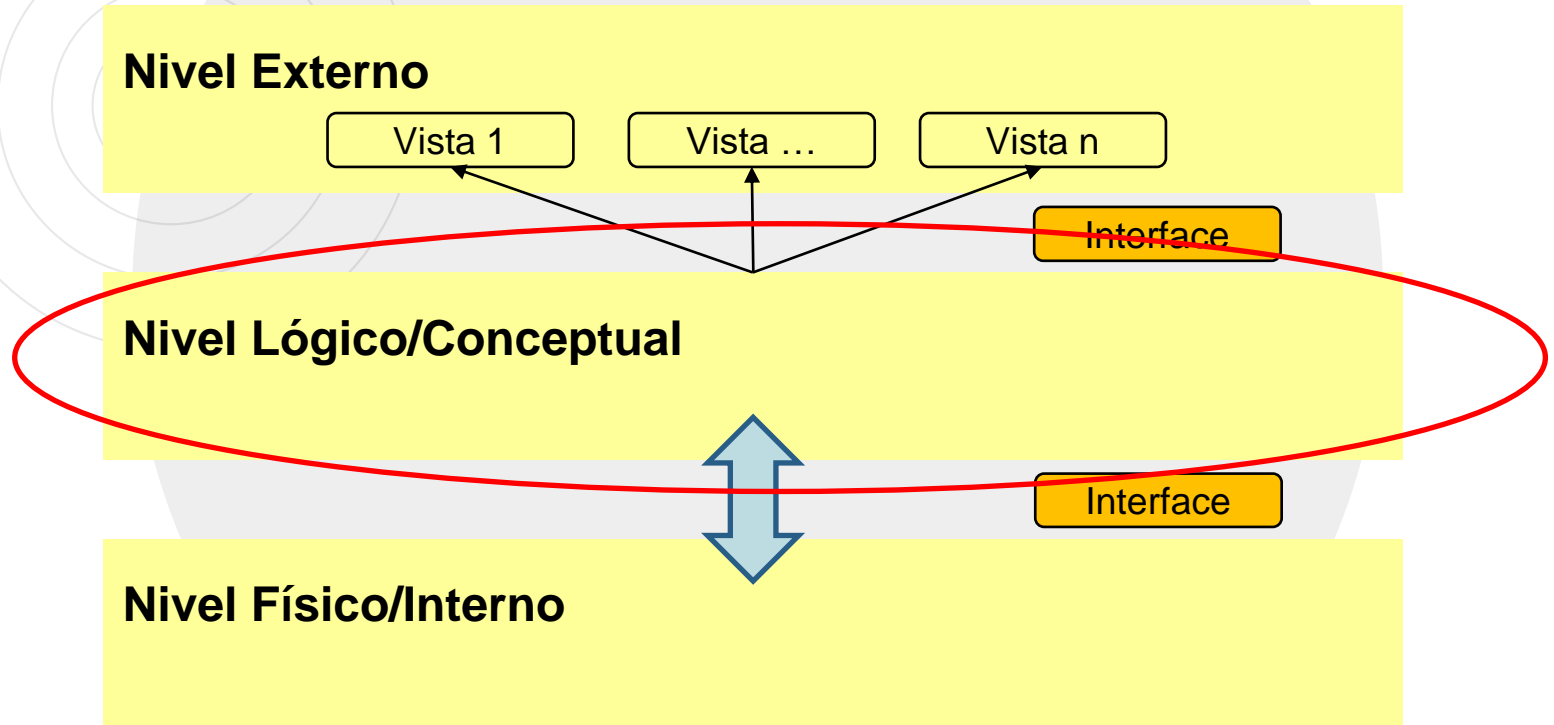
- Los datos son percibidos como tablas.
- Las operaciones disponibles generan nuevas tablas a partir de las anteriores.

Definición. *Esquema de Base de datos:*

Es la visión que tienen los diferentes usuarios de la Base de Datos.



Arquitectura - ANSI-SPARC

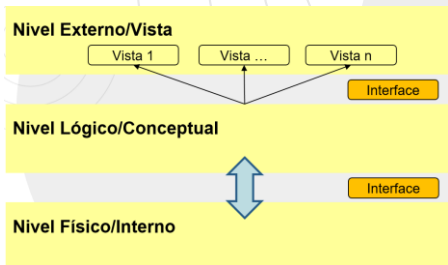


RDBMS – Arq. ANSI-SPARC



Externo

Nivel de abstracción más alto. Provee diferentes vistas de la misma BD a diferentes usuarios brindando seguridad escondiendo partes de la DB a usuarios particulares.



Lógico

Describe la estructura de la BD. Explica que datos se almacenan en las tablas de la DB, tipos y relaciones entre las tablas. Hay un solo esquema conceptual por DB.

Nivel Físico

Nivel mas bajo de abstracción. Describe como se almacenan los datos en la BD y provee los métodos para accederlos.

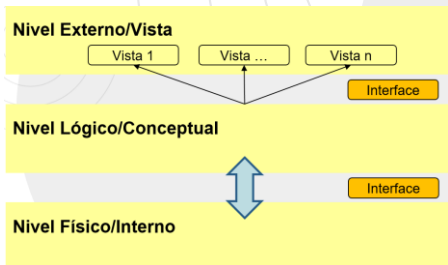


RDBMS – Arq. ANSI-SPARC



Interface entre el nivel Externo y Conceptual

Define la correspondencia entre los datos de la capa Externa y la capa conceptual.



Interface entre el nivel Conceptual e Interno

La interface entre el nivel conceptual e interno identifica como se almacenan y acceden los registros lógicos en el nivel interno.



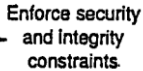
RDBMS - Funciones



- Definición de datos.
- Consulta y recuperación de datos.
- Manipulación de datos.
- Ejecución.
- Optimización.
- Integridad de datos
- Seguridad
- Catálogo



“



Autor Hernán Puelman

Herramientas y librerías



- **Generadores de reportes.**
- **Herramientas de administración y monitoreo.**
- **Lenguajes de aplicaciones.**
- **Reorganización**
- **Estadísticas**
- **Carga/descarga (Import/Export)**
- **Backup/recovery**



File System vs RDBMS



- No conoce la estructura interna de los registros.
- En los sistemas No DBMS coincide la estructura interna de los registros.
- No implementan restricciones de seguridad e integridad.
- No hay soporte de concurrencia.
- No poseen un catálogo.
- No proporcionan independencia de datos.
- No hay integración entre los diferentes archivos.





Preguntas ?





Ejercicios

Autor Hernán Puelman



Ejercicios



1. *Qué tipos de componentes tiene una BD?*
2. *El DBMS a qué tipo de componente de la BD pertenece?*
3. *Diferencia entre el modelo de datos y su implementación.*
4. *De un ejemplo de Independencia de Datos física*
5. *De un ejemplo de Independencia de Datos lógica*
6. *Qué es la A en ACID y qué significa?*
7. *Las tablas a qué nivel de la arquitectura ANSI-SPARC pertenecen?*
8. *Qué niveles del modelo ANSI SPARC son relacionales?*
9. *Cuántos esquemas tiene cada nivel en el modelo ANSI SPARC?*
10. *Qué es y cuál es la función del Catálogo en un RDBMS?*
11. *Por qué un archivo implementado por File System no proporciona independencia de datos?*
12. *Qué diferencias hay en cómo implementa la seguridad un File System y un DBMS respecto al acceso a los datos?*
13. *De dos ejemplos de utilitarios que comúnmente acompañan a un DBMS*

