

BASES DE DATOS AVANZADAS

Clase anterior

- Tema 2. Bases de Datos Distribuidas (BDD).
- **Características de Transparencia en los Sistemas Manejadores de Bases de Datos Distribuidos (SMBDD)**
 - ❖ Transparencia de replicación.
 - ❖ Transparencia de falla.
 - ❖ Transparencia de desempeño.
 - ❖ Transparencia de heterogeneidad.

Clase 8

- Tema 2. Bases de Datos Distribuidas (BDD).
- **Arquitectura de los sistemas de administración de base de datos distribuida**

Bases de Datos Distribuidas

Bibliografía

Bell, David (1992). Distributed database systems. Wokingham, Eng. : Addison-Wesley.

Connolly, Thomas M. (2005). Sistemas de bases de datos: un enfoque práctico para diseño, implementación y gestión. (4ta ed.). Madrid : Pearson Educación Limited.

Date, C. J. (2001). Introducción a los sistemas de bases de datos. (7ma ed.). México: Pearson Educación: Addison Wesley.

Marqués, M. (2001). Apuntes de ficheros y bases de datos. Universitat Jaume I, Campus de Riu Sec. España. consultado el 13 de noviembre de 2007, en:
<http://www3.uji.es/~mmarques/f47/apun/apun.html> Si no encuentras el documento, pulsa aquí.

Rob, Peter (2004). Sistemas de bases de datos: diseño, implementación y administración. (5ta ed.). México, D.F.: Thomson.

Clase 8

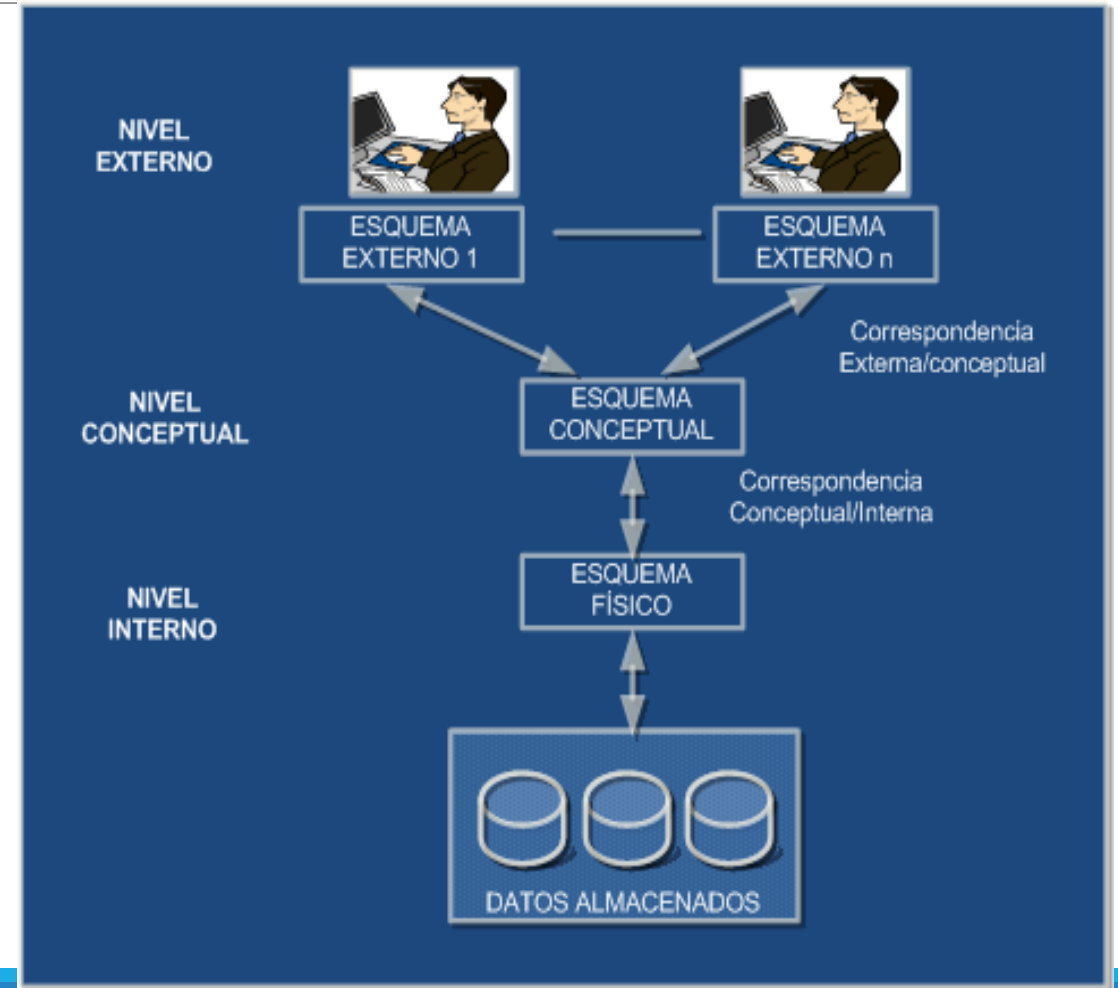
Tema 2. Bases de Datos Distribuidas (BDD).

- **Arquitectura de los sistemas de administración de base de datos distribuida**
- Arquitectura de los **Sistemas Manejadores de Bases de Datos** ANSI-SPARC divide a un sistema en tres niveles: interno, conceptual y externo.
- Surgió en 1975, el comité **ANSI-SPARC** (American National Standard Institute - Standards Planning and Requirements Committee) propuso una arquitectura de tres niveles para los sistemas de bases de datos.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

- **Arquitectura ANSI-SPARC** de los sistemas de administración de base de datos (SMBD)



Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI-SPARC** de los sistemas de administración de base de datos

- El objetivo de la arquitectura de tres niveles es separar los programas de aplicación de la base de datos física.
- En esta arquitectura, el esquema de una base de datos se define en tres niveles de abstracción distintos (Marqués, 2001)

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI-SPARC** de los sistemas de administración de base de datos

- En el nivel interno:

Describe la estructura física de la base de datos mediante un esquema interno.

Se especifica mediante un modelo físico y describe todos los detalles para el almacenamiento de la base de datos, así como los métodos de acceso.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI-SPARC** de los sistemas de administración de base de datos

Nivel conceptual:

- Describe la estructura de toda la base de datos para el usuario, mediante un esquema conceptual.
- Se concentra en describir entidades, atributos, relaciones, operaciones de los usuarios y restricciones.
- Se puede utilizar un modelo conceptual o un modelo lógico para especificar el esquema.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI-SPARC** de los sistemas de administración de base de datos

Nivel externo:

- Describe varios esquemas externos o vistas de usuario.
- Cada esquema externo describe la parte de la base de datos que interesa a un grupo de usuarios determinado y oculta a ese grupo el resto de la base de datos.
- En este nivel se puede utilizar un modelo conceptual o un modelo lógico para especificar los esquemas.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI-SPARC** de los sistemas de administración de base de datos

- La mayoría de los SMBD no distinguen del todo los tres niveles. Algunos incluyen detalles del nivel físico en el esquema conceptual por ejemplo.
- Los tres esquemas son descripciones de los mismos datos pero con distintos niveles de abstracción.

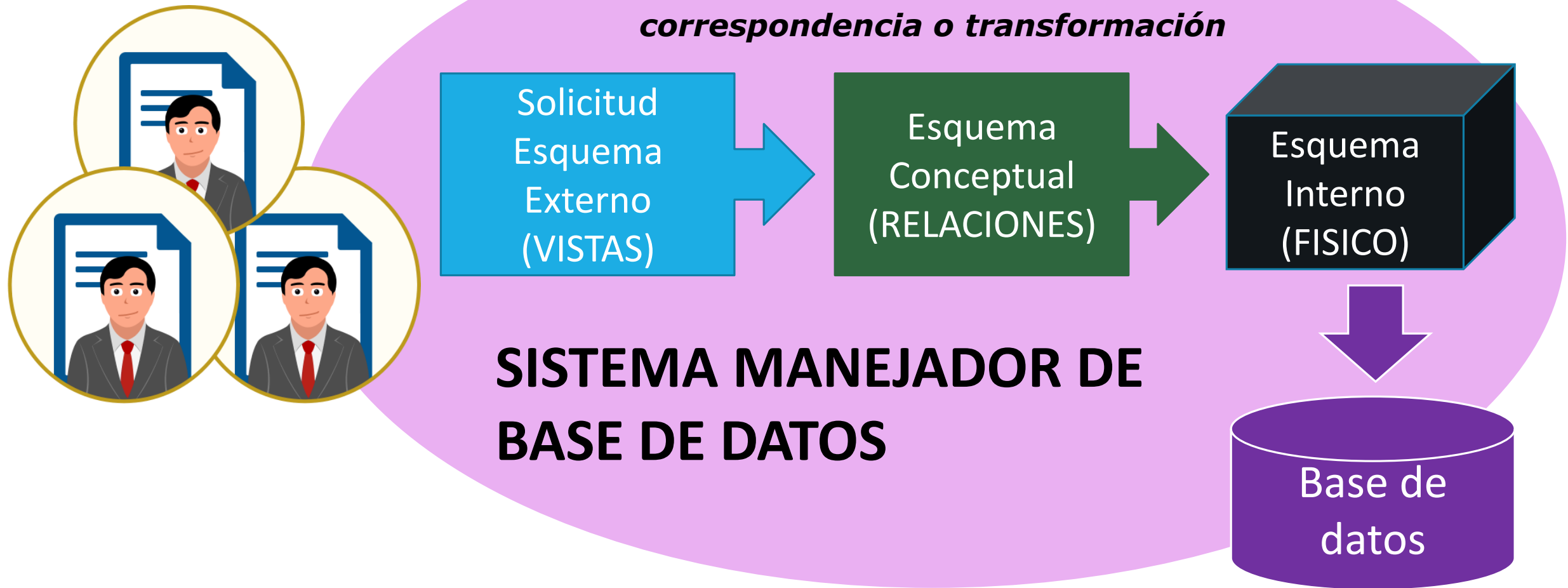
Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI-SPARC** de los sistemas de administración de base de datos

- Los únicos datos que existen realmente están a nivel físico, almacenados en un dispositivo como puede ser un disco.

Clase 8. Tema 2. Bases de Datos Distribuidas (BDD). Arquitectura **ANSI-SPARC** de los sistemas de administración de base de datos



Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- Según David Bell (Bell, 1992) para definir un esquema de estandarización en bases de datos distribuidas se debe definir un modelo de referencia.
- Para definir ese modelo de referencia se puede seguir uno de los siguientes tres enfoques.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- Basado en componentes.
- Basado en funciones.
- Basado en datos.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- **Basado en componentes.** Un SMBD consiste de un número de componentes y sus relaciones, cada uno de los cuales proporciona alguna funcionalidad.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- **Basado en funciones.** Se identifican las diferentes clases de usuarios junto con la funcionalidad que el sistema ofrecerá para cada clase.
- Estructura jerárquica de clases.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- **Basado en datos.** Se identifican los diferentes tipos de descripción de datos.
- Define las unidades funcionales que realizarán y/o usarán los datos de acuerdo con las diferentes vistas.

**Este es el enfoque seguido por el modelo
ANSI/SPARC.**

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- **Basado en datos.**
- Para definir la arquitectura ANSI/SPARC de un sistema de administración de base de datos distribuida utilizando el modelo de referencia basado en datos, recordemos, los sistemas de datos distribuidos están divididos en dos clases separadas:
- **Sistemas de administración de bases de datos distribuidas homogéneos.**
- **Sistemas de administración de bases de datos distribuidas heterogéneas.**

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- **Sistemas de administración de bases de datos distribuidas homogéneos.**
 - Tiene múltiples colecciones de datos.
 - Integra múltiples recursos de datos
 - Se parecen a un sistema centralizado. Datos en varios sitios comunicados por la red.
 - No existen usuarios locales y todos ellos acceden la base de datos a través de una interfaz global.
 - El esquema global es la unión de toda las descripciones de datos locales.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

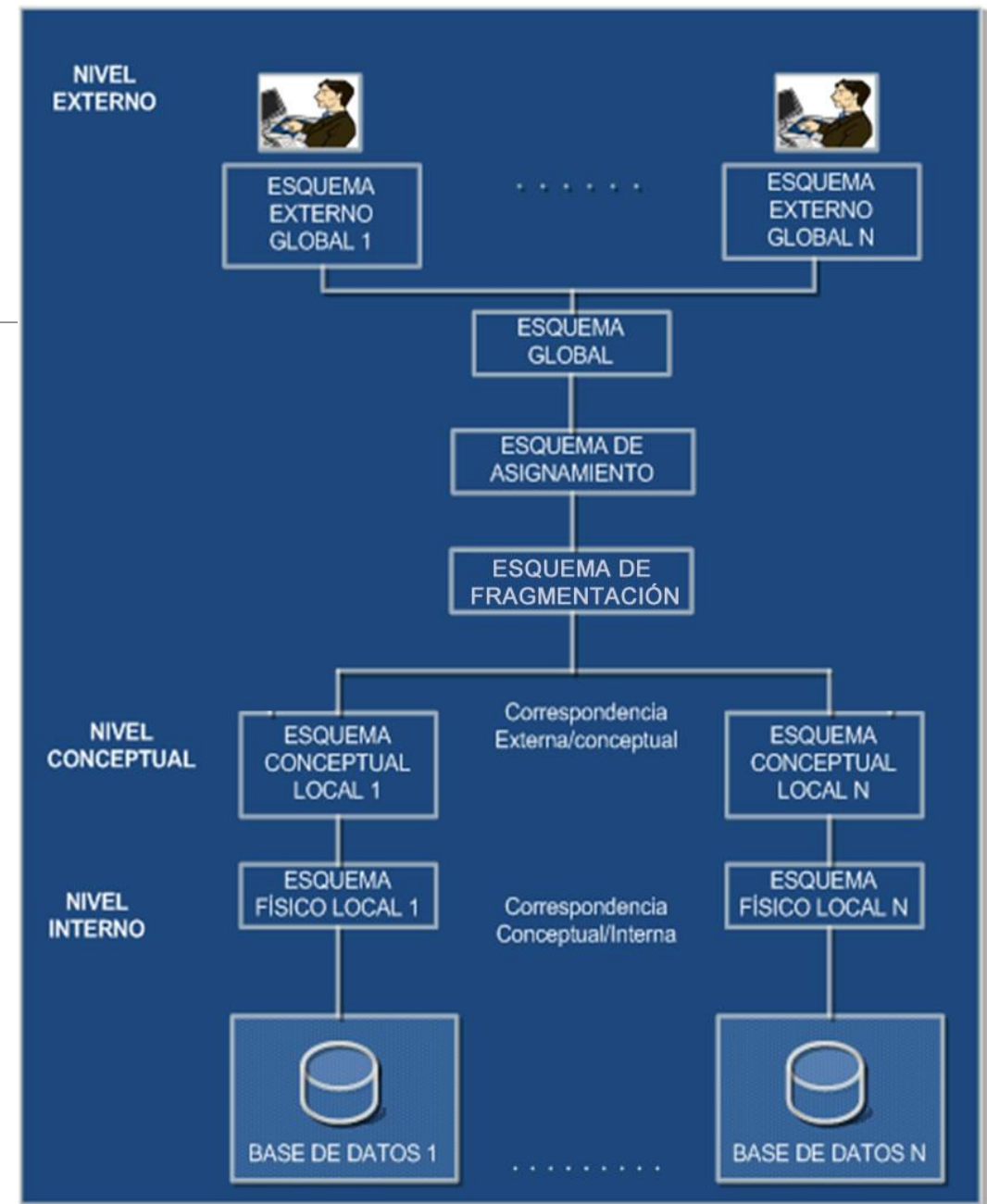
- **Sistemas de administración de bases de datos distribuidas homogéneos.**
 - Para manejar los aspectos de la distribución, se deben agregar dos niveles a la arquitectura estándar ANSI-SPARC:
 - El esquema de fragmentación describe la forma en que las relaciones globales se dividen entre las bases de datos locales
 - El esquema de asignamiento especifica el lugar en el cual cada fragmento es almacenado.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- **Sistemas de administración de bases de datos distribuidas homogéneos.**



Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

- **Sistemas de administración de bases de datos distribuidas heterogéneas.**
 - Según David Bell (Bell, 1992), se caracterizan por manejar diferentes sistemas de administración de base de datos (DBMS) en los nodos locales.
 - Una subclase son los **sistemas de administración de multi-bases de datos.**

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Arquitectura **ANSI/SPARC** de un sistema distribuido

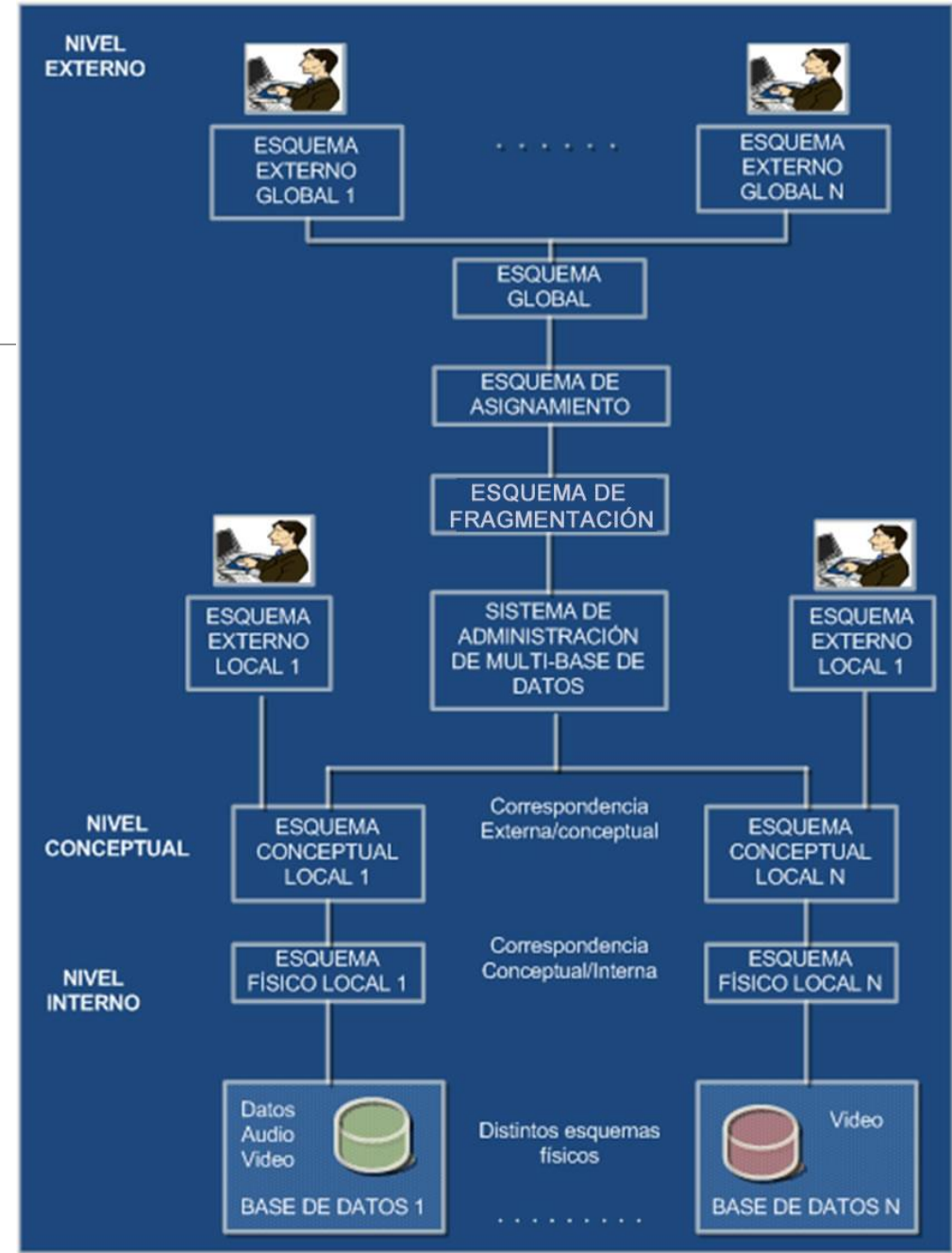
- **Sistemas de administración de bases de datos distribuidas heterogéneos.**
 - Un sistema multi-bases de datos (Smulti-BD) **tiene múltiples sistemas manejadores de base de datos**, que pueden ser de tipos diferentes y múltiples bases de datos existentes. La integración de todos ellos se realiza mediante subsistemas de software.
 - Existen **usuarios locales y globales**. Los usuarios locales acceden sus bases de datos locales sin verse afectados por la presencia del sistema administrador de multi-base de datos.

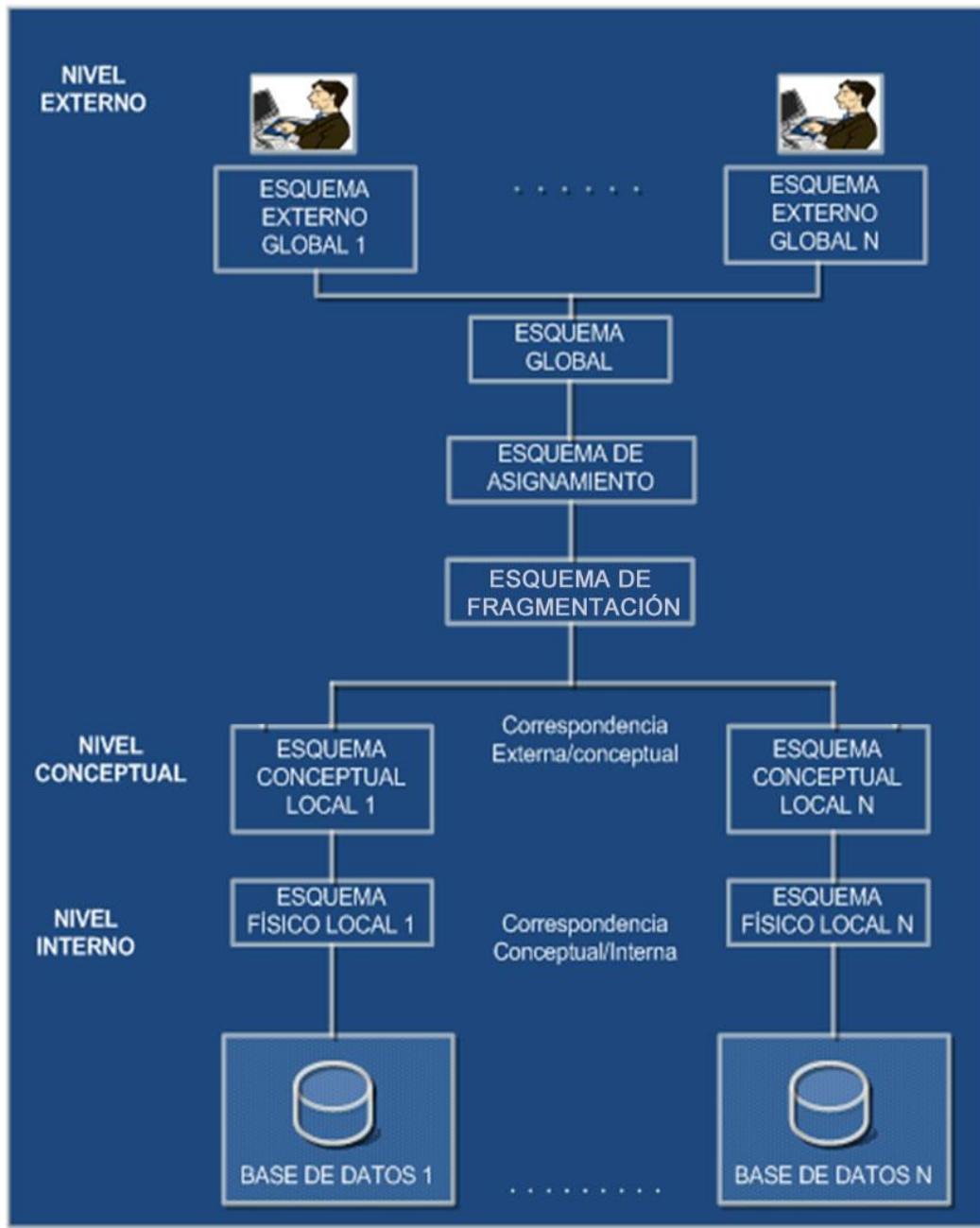
Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

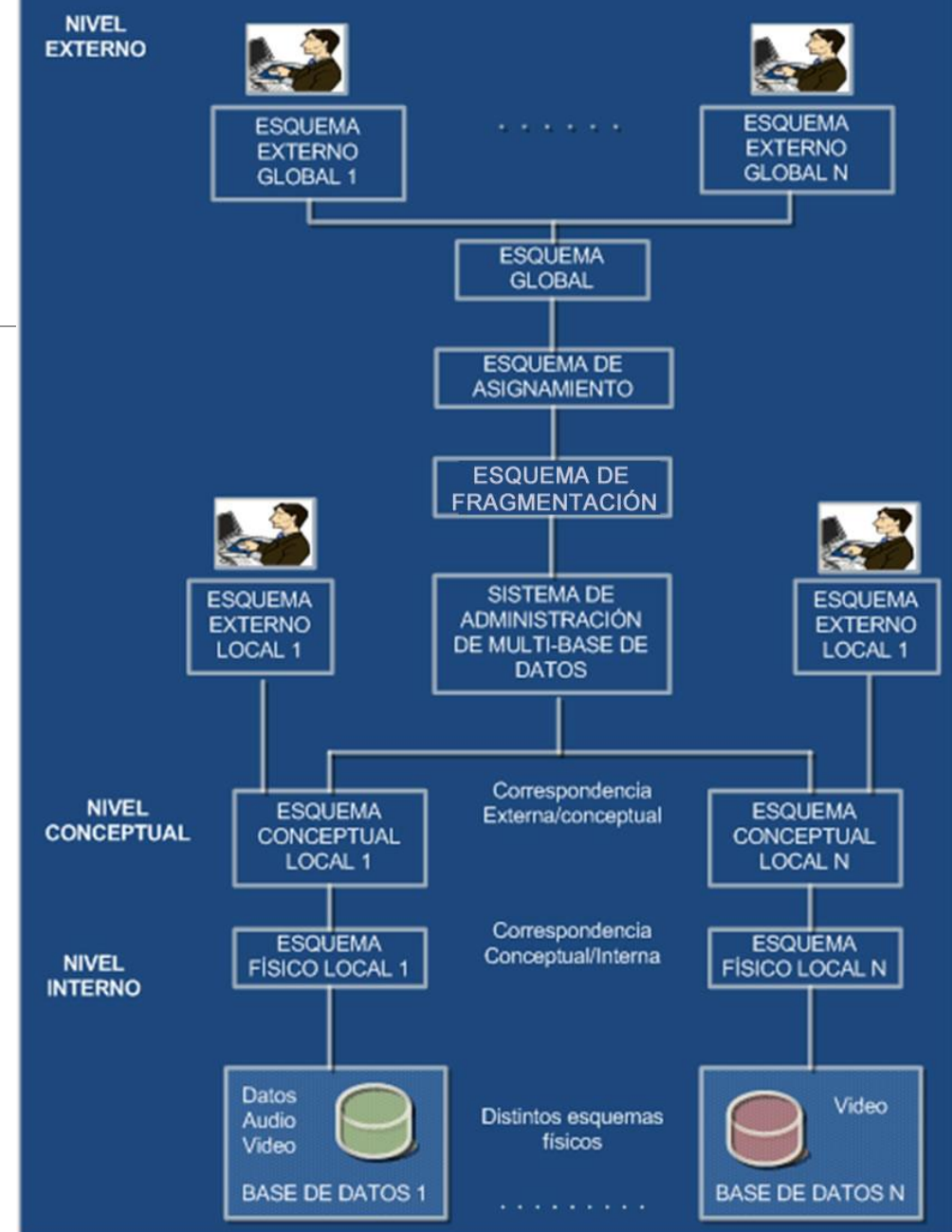
Arquitectura **ANSI/SPARC** de un sistema distribuido

- **Sistemas de administración de bases de datos distribuidas heterogéneos.**





SMBDD homogéneo



SMBDD heterogéneo

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

- Se basa en los principios del diseño de una base de datos centralizada y contempla los tres puntos siguientes (Marqués, 2001):
 - i. Diseño conceptual de la base de datos.
 - ii. Diseño lógico de la base de datos.
 - iii. Diseño físico de la base de datos
- * La ubicación de datos y programas a través de los diferentes sitios de una red de computadoras.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

i. Diseño conceptual de la base de datos.

- Se construye el esquema de la información, independientemente de cualquier consideración física.
- A este esquema se le denomina **esquema conceptual**.
- Al construir el esquema, se expone la semántica (significado) de los datos encontrando las entidades, atributos y relaciones.
- Su objetivo es comprender:
 - ✓ La perspectiva que cada usuario tiene de los datos.
 - ✓ La naturaleza de los datos, independientemente de su representación física.
 - ✓ El uso de los datos a través de las áreas de aplicación.
- Comunicación Diseñador ↔ Organización
- La notación más popular es el modelo entidad-relación.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

i. Diseño lógico de la base de datos.

- Se construye el esquema de la información, **basándose en un modelo de base de datos específico**, independientemente de cualquier consideración física y del SMBD.
- Se transforma **el esquema conceptual en un esquema lógico** que utilizará las estructuras de datos del modelo de base de datos en el que se basa el DBMS que se vaya a utilizar.
- Puede ser el **modelo relacional**, el modelo de red, el modelo jerárquico o el modelo orientado a objetos.

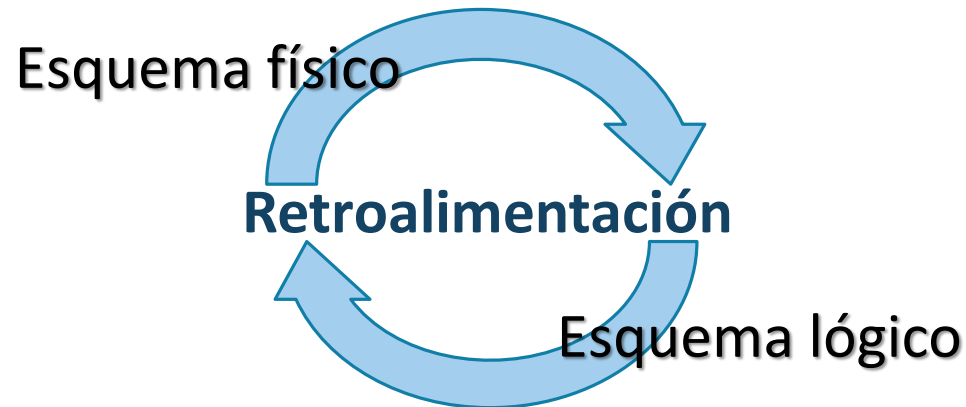
Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

i. Diseño físico de la base de datos.

- Descripción de la implementación de la base de datos en memoria secundaria: estructuras de almacenamiento y métodos de acceso que garanticen un acceso eficiente a los datos.
- Cuál es el DBMS que se va a utilizar, ya que el esquema físico se adapta a él.



Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

i. Diseño físico de la base de datos.

- En el caso de las **bases de datos distribuidas** se tienen que considerar adicionalmente los tres problemas siguientes (Rob, 2004):
 - I. Cómo dividir la base de datos en fragmentos.
 - II. Qué fragmentos replicar.
 - III. Dónde localizar estos fragmentos y réplicas.
- La fragmentación y replicación de los datos se ocupa de los dos primeros problemas, mientras que la colocación de los datos con el tercero.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

**¿Cuáles son los objetivos
del diseño de la
distribución de los datos?**



Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Objetivos del diseño de la distribución de los datos(Bell, 1992):

- ✓ Procesamiento local.
- ✓ Distribución de la carga de trabajo.
- ✓ Costo de almacenamiento y disponibilidad.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Objetivos del diseño de la distribución de los datos(Bell, 1992):

✓ Procesamiento local.

Para maximizar el procesamiento local, colocar los datos tan cerca como sea posible de las aplicaciones que los utilizan.

En el diseño de la distribución de los datos se puede agregar el **número de referencias locales y remotas** que le corresponden a cada **fragmentación candidata** y la **localización del fragmento**, y de esta forma se seleccione la **mejor solución** de ellas.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Objetivos del diseño de la distribución de los datos(Bell, 1992):

- ✓ Distribución de la carga de trabajo.

Se realiza para tomar ventaja de las diferentes características (potenciales) o utilizaciones de las computadoras de cada sitio, y maximizar el grado de ejecución de paralelismo de las aplicaciones.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Objetivos del diseño de la distribución de los datos(Bell, 1992):

- ✓ Costo de almacenamiento

Es posible tener sitios especializados en la red para el almacenamiento de datos.

El costo de almacenamiento de datos no es tan relevante si éste se compara con el del **CPU, I/O** y costos de transmisión de las aplicaciones y disponibilidad.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

El problema de la fragmentación



Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

El problema de la fragmentación.

Se refiere al particionamiento de la información para distribuir cada parte en los diferentes sitios de la red.

El diseño de una base de datos distribuida, cualquiera que sea el enfoque que se siga, debe responder satisfactoriamente las siguientes preguntas:

- a. ¿Por qué hacer una fragmentación de datos?
- b. ¿Cómo realizar la fragmentación?
- c. ¿Qué tanto se debe fragmentar?
- d. ¿Cómo probar la validez de una fragmentación?
- e. ¿Cómo realizar el asignamiento de fragmentos?
- f. ¿Cómo considerar los requerimientos de la información?

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

El problema de la fragmentación.

Nombre del fragmento: CLIENTE_H1 Ubicación: DF Nodo: DF

Numero	Nombre	Direccion	Estado	Edad
1	Rodolfo Pérez	Las palmas 34	DF	34
3	Jorge Ortega	Priv. Jarras 12	DF	35
*	(Nuevo)			

Nombre del fragmento: CLIENTE_H2 Ubicación: VERACRUZ Nodo: XAL

Numero	Nombre	Direccion	Estado	Edad
2	Arturo Carmor	Carlos A. Carril	VERACRUZ	40
4	Saúl Jiménez S	Díaz Mirón 45	VERACRUZ	60
5	Samuel Cruz	Ruiz Cortínez	VERACRUZ	53

Nombre del fragmento: CLIENTE_H3 Ubicación: NL Nodo: MON

Numero	Nombre	Direccion	Estado	Edad
6	Erika Ochoa	20 de Nov. 24	NL	28

Fragmentos de una tabla en tres lugares

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

Peter Rob (Rob, 2004), afirma que **la fragmentación de los datos** permite dividir un objeto en dos o más segmentos o fragmentos.

Objeto:

- Base de datos de usuario
- Base de datos de sistema
- Una tabla.

Cada fragmento puede guardarse en cualquier sitio en una red de computadoras.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

La información de la fragmentación de los datos se guarda en un **Catálogo de Datos Distribuidos** (DDC, Distributed Data Catalog por sus siglas en inglés), desde donde es accedida por el procesador de transacciones para procesar las solicitudes de los usuarios.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

- Las estrategias de fragmentación de los datos a estudiar, están basadas a nivel de tabla.
- La tabla se divide en fragmentos lógicos.
 - **Fragmentación horizontal**
 - **Fragmentación vertical**
 - **Fragmentación mezclada**

Una tabla fragmentada siempre puede recrearse con sus partes fragmentadas mediante una combinación de uniones y articulaciones.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

La **fragmentación horizontal** se refiere a la división de una relación en subconjunto (fragmentos) de tuplas (filas).

Cada fragmento se guarda en un nodo diferente, y cada uno de ellos tiene filas únicas.

Todas las filas únicas tienen los mismos atributos (columnas). En suma, cada fragmento equivale a una sentencia SELECT, con la cláusula WHERE en un solo atributo.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

La **fragmentación vertical** se refiere a la división de una relación en subconjuntos de atributos (columna).

Cada subconjunto (fragmento) se guarda en un nodo diferente, y cada fragmento tiene columnas únicas, con la excepción de la columna clave, la cuál es común a todos los fragmentos.

Esto es el equivalente de la sentencia `SELECT columna1, columna2 INTO Nueva_Tabla FROM Tabla.`

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

La **fragmentación mezclada** se refiere a una combinación de estrategias horizontales y verticales.

En otras palabras, una tabla puede dividirse en varios subconjuntos horizontales (filas), y cada una tiene un subconjunto de los atributos (columnas).

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

Tabla a fragmentar

CLI_NUM	CLI_NOM	CLI_DIR	CLI_EST	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA
1	Selene Aguirre	Las palmas 34	DF	\$3,500.00	\$2,700.00	3	\$1,245.00
2	Martin Porres	Blvd. Lopez Ma	VE	\$6,000.00	\$12,000.00	1	\$0.00
3	Miriam Gutierri	Águila 34	DF	\$4,000.00	\$3,500.00	3	\$3,400.00
4	Benito López	Rueda 23	VE	\$6,000.00	\$5,890.00	3	\$1,090.00
5	Victor Pérez	Carlos Carrillo	VE	\$1,200.00	\$550.00	1	\$0.00
6	Nicolás Rosas	20 de nov. 123	NL	\$2,000.00	\$350.00	2	\$50.00

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **horizontal** de una base de datos distribuida.

Tabla CLIENTE de la compañía XYZ, ilustrada en la figura. La tabla contiene los atributos CLI_NUM, CLI_NOM, CLI_DIR, CLI_EST, CLI_LIM, CLI_BAL, CLI_NIVEL, CLI_DEUDA.

CLI_NUM	CLI_NOM	CLI_DIR	CLI_EST	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA
1	Selene Aguirre	Las palmas 34	DF	\$3,500.00	\$2,700.00	3	\$1,245.00
2	Martin Porres	Blvd. Lopez Ma	VE	\$6,000.00	\$12,000.00	1	\$0.00
3	Miriam Gutieri	Águila 34	DF	\$4,000.00	\$3,500.00	3	\$3,400.00
4	Benito López	Rueda 23	VE	\$6,000.00	\$5,890.00	3	\$1,090.00
5	Victor Pérez	Carlos Carrillo	VE	\$1,200.00	\$550.00	1	\$0.00
6	Nicolás Rosas	20 de nov. 123	NL	\$2,000.00	\$350.00	2	\$50.00

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

La **fragmentación horizontal** se refiere a la división de una relación en subconjunto (fragmentos) de tuplas (filas).

Cada fragmento se guarda en un nodo diferente, y cada uno de ellos tiene filas únicas.

Todas las filas únicas tienen los mismos atributos (columnas). En suma, cada fragmento equivale a una sentencia SELECT, con la cláusula WHERE en un solo atributo.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **horizontal** de una base de datos distribuida.

La Compañía **XYZ** requiere información sobre sus **clientes** en los tres estados, pero las **ubicaciones** de la compañía en cada estado (DF, VE y MY) solamente requieren datos con respecto a **clientes locales**.

Con base en esos requerimientos, se decide **distribuir los datos por estado**. Por consiguiente, se definen los fragmentos horizontales de acuerdo con la estructura mostrada en la tabla 1.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **horizontal** de una base de datos distribuida.

NOMBRE DEL FRAGMENTO	UBICACIÓN	CONDICIÓN	NOMBRE DEL NODO	NÚMEROS DE CLIENTE	NÚMERO DE REGISTROS
CLI_H1	Distrito Federal	CLI_EST = 'DF'	DF	1, 3	2
CLI_H2	Nuevo León	CLI_EST = 'NL'	MY	6	1
CLI_H2	Veracruz	CLI_EST = 'VE'	XAL	2, 4, 5	3

Cada fragmento horizontal puede tener un número diferente de filas, pero cada uno de ellos **DEBE** tener los mismos atributos. Los fragmentos resultantes producen las tres tablas siguientes.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **horizontal** de una base de datos distribuida.

Nombre del fragmento: CLI_H1			Ubicación: Distrito Federal			Nodo: DF	
CLI_NUM	CLI_NOM	CLI_DIR	CLI_EST	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA
1	Selene Aguirre	Las palmas 34	DF	\$3,500.00	\$2,700.00	3	\$1,245.00
3	Miriam Gutierrez	Águila 34	DF	\$4,000.00	\$3,500.00	3	\$3,400.00

Nombre del fragmento: CLI_H2			Ubicación: Nuevo León			Nodo: MY	
CLI_NUM	CLI_NOM	CLI_DIR	CLI_EST	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA
6	Nicolás Rosas	20 de nov. 123	NL	\$2,000.00	\$350.00	2	\$50.00

Nombre del fragmento: CLI_H3			Ubicación: Veracruz			Nodo: XAL	
CLI_NUM	CLI_NOM	CLI_DIR	CLI_EST	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA
2	Martin Porres	Blvd. Lopez Mateo	VE	\$6,000.00	\$12,000.00	1	\$0.00
4	Benito López	Rueda 23	VE	\$6,000.00	\$5,890.00	3	\$1,090.00
5	Victor Pérez	Carlos Carrillo	VE	\$1,200.00	\$550.00	1	\$0.00

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **vertical** de una base de datos distribuida.

Tabla CLIENTE de la compañía XYZ, ilustrada en la figura 1. La tabla contiene los atributos CLI_NUM, CLI_NOM, CLI_DIR, CLI_EST, CLI_LIM, CLI_BAL, CLI_NIVEL, CLI_DEUDA.

CLI_NUM	CLI_NOM	CLI_DIR	CLI_EST	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA
1	Selene Aguirre	Las palmas 34	DF	\$3,500.00	\$2,700.00	3	\$1,245.00
2	Martin Porres	Blvd. Lopez Ma	VE	\$6,000.00	\$12,000.00	1	\$0.00
3	Miriam Gutieri	Águila 34	DF	\$4,000.00	\$3,500.00	3	\$3,400.00
4	Benito López	Rueda 23	VE	\$6,000.00	\$5,890.00	3	\$1,090.00
5	Victor Pérez	Carlos Carrillo	VE	\$1,200.00	\$550.00	1	\$0.00
6	Nicolás Rosas	20 de nov. 123	NL	\$2,000.00	\$350.00	2	\$50.00

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

La **fragmentación vertical** se refiere a la división de una relación en subconjuntos de atributos (columna).

Cada subconjunto (fragmento) se guarda en un nodo diferente, y cada fragmento tiene columnas únicas, con la excepción de la columna clave, la cuál es común a todos los fragmentos.

Esto es el equivalente de la sentencia `SELECT columna1, columna2 INTO Nueva_Tabla FROM Tabla.`

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **vertical** de una base de datos distribuida.

Supongamos que la compañía está dividida en dos fragmentos, el de **servicio** y el de **colecciones**.

Cada departamento está ubicado en un edificio distinto, y cada uno tiene interés solamente en unos cuantos de los **atributos** de la tabla CLIENTE.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **vertical** de una base de datos distribuida.

Cada fragmento vertical debe tener el **mismo número de filas**, pero la inclusión de los diferentes atributos depende de la columna **clave**.

NOMBRE DEL FRAGMENTO	UBICACIÓN	NOMBRE DEL NODO	NOMBRES DE ATRIBUTO
CLI_V1	Edif. de servicio	ES	CLI_NUM, CLI_NOM, CLI_DIR, CLI_EST
CLI_V2	Edif. de colección	EC	CLI_NUM, CLI_LIM, CLI_BAL, CLI_NIVEL, CLI_DEUDA

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **vertical** de una base de datos distribuida.

Observemos que el atributo clave (CLI_NUM) es común en ambos fragmentos CLI_V1 y CLI_V2.

NOMBRE DEL FRAGMENTO	UBICACIÓN	NOMBRE DEL NODO	NOMBRES DE ATRIBUTO
CLI_V1	Edif. de servicio	ES	CLI_NUM, CLI_NOM, CLI_DIR, CLI_EST
CLI_V2	Edif. de colección	EC	CLI_NUM, CLI_LIM, CLI_BAL, CLI_NIVEL, CLI_DEUDA

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación
vertical de una
base de datos
distribuida.

Nombre del fragmento: CLI_V1				Ubicación: Edif. de servicio		Nodo: ES	
CLI_NUM	CLI_NOM	CLI_DIR	CLI_EST				
1	Selene Aguirre	Las palmas 34	DF				
2	Martin Porres	Bld. Lopez Ma	VE				
3	Miriam Gutierr	Águila 34	DF				
4	Benito López	Rueda 23	VE				
5	Victor Pérez	Carlos Carrillo	VE				
6	Nicolás Rosas	20 de nov. 123	NL				

Nombre del fragmento: CLI_V2					Ubicación: Edif. De colección		Nodo: EC	
CLI_NUM	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA				
1	\$3,500.00	\$2,700.00	3	\$1,245.00				
2	\$6,000.00	\$12,000.00	1	\$0.00				
3	\$4,000.00	\$3,500.00	3	\$3,400.00				
4	\$6,000.00	\$5,890.00	3	\$1,090.00				
5	\$1,200.00	\$550.00	1	\$0.00				
6	\$2,000.00	\$350.00	2	\$50.00				

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **mezclada** de una base de datos distribuida.

Tabla CLIENTE de la compañía XYZ, ilustrada en la figura. La tabla contiene los atributos CLI_NUM, CLI_NOM, CLI_DIR, CLI_EST, CLI_LIM, CLI_BAL, CLI_NIVEL, CLI_DEUDA.

CLI_NUM	CLI_NOM	CLI_DIR	CLI_EST	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA
1	Selene Aguirre	Las palmas 34	DF	\$3,500.00	\$2,700.00	3	\$1,245.00
2	Martin Porres	Blvd. Lopez Ma	VE	\$6,000.00	\$12,000.00	1	\$0.00
3	Miriam Gutieri	Águila 34	DF	\$4,000.00	\$3,500.00	3	\$3,400.00
4	Benito López	Rueda 23	VE	\$6,000.00	\$5,890.00	3	\$1,090.00
5	Victor Pérez	Carlos Carrillo	VE	\$1,200.00	\$550.00	1	\$0.00
6	Nicolás Rosas	20 de nov. 123	NL	\$2,000.00	\$350.00	2	\$50.00

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación de una base de datos distribuida.

La **fragmentación mezclada** se refiere a una combinación de estrategias horizontales y verticales.

En otras palabras, una tabla puede dividirse en varios subconjuntos horizontales (filas), y cada una tiene un subconjunto de los atributos (columnas).

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **mezclada** de una base de datos distribuida.

La compañía XYZ requiere que los datos CLIENTE se fragmenten **horizontalmente** para acomodar las diferentes ubicaciones de la compañía; dentro de las ubicaciones, los datos deben ser fragmentados **verticalmente** para acomodar los diferentes departamentos (servicio y colección).

En suma, la tabla CLIENTE requiere una fragmentación mezclada.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **mezclada** de una base de datos distribuida.

La fragmentación mezclada requiere un procedimiento de dos pasos.

Primero, se introduce la **fragmentación horizontal** por cada sitio, con base en la ubicación dentro de un estado (CLI_EST).

La fragmentación horizontal produce los subconjuntos de tuplas cliente (fragmentos horizontales) localizados en cada sitio.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **mezclada** de una base de datos distribuida.

Como los departamentos están localizados en edificios diferentes; se utiliza una **fragmentación vertical** dentro de cada **fragmento horizontal** para dividir los atributos, con lo cual se satisfacen las necesidades de información de cada departamento en cada sub sitio.

Clase 8

Tema 2. Bases de Datos Distribuidas (BDD).

Diseño de una base de datos distribuidas

Fragmentación **mezclada** de una base de datos distribuida.

NOMBRE DEL FRAGMENTO	UBICACIÓN	CRITERIOS HORIZONTALES	NOMBRE DEL NODO	FILAS RESULTANTES EN EL SITIO	NOMBRES DE ATRIBUTOS EN EL SITIO
CLI_M1	DF-Servicio	CLI_EST = 'DF'	DF-S	1, 3	CLI_NUM, CLI_NOM, CLI_DIR, CLI_EST
CLI_M2	DF-Colección	CLI_EST = 'DF'	DF-C	1, 3	CLI_NUM, CLI_LIM, CLI_BAL, CLI_NIVEL, CLI_DEUDA
CLI_M3	NL-Servicio	CLI_EST = 'NL'	MY-S	6	CLI_NUM, CLI_NOM, CLI_DIR, CLI_EST
CLI_M4	NL-Colección	CLI_EST = 'NL'	MY-C	6	CLI_NUM, CLI_LIM, CLI_BAL, CLI_NIVEL, CLI_DEUDA
CLI_M5	'VE'-Servicio	CLI_EST = 'VE'	XAL-S	2, 4, 5	CLI_NUM, CLI_NOM, CLI_DIR, CLI_EST
CLI_M6	'VE'-Colección	CLI_EST = 'VE'	XAL-C	2, 4, 5	CLI_NUM, CLI_LIM, CLI_BAL, CLI_NIVEL, CLI_DEUDA

Cada fragmento mostrado contiene datos de clientes por estado y, dentro de cada estado, por ubicación de departamento, para adaptarse a los requerimientos de datos de cada departamento

Fragmentación **mezclada** de una base de datos distribuida.

Nombre del fragmento: CLI_M1 Ubicación: DF-Servicio Nodo: DF-S

CLI_NUM	CLI_NOM	CLI_DIR	Expr1003	CLI_EST
1	Selene Aguirre	Las palmas 34	DF	DF
3	Miriam Gutierr	Águila 34	DF	DF

Nombre del fragmento: CLI_M2 Ubicación: DF-colección Nodo: DF-C

CLI_NUM	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA	CLI_EST
1	\$3,500.00	\$2,700.00	3	\$1,245.00	DF
3	\$4,000.00	\$3,500.00	3	\$3,400.00	DF

Nombre del fragmento: CLI_M3 Ubicación: NL-Servicio Nodo: MY-S

CLI_NUM	CLI_NOM	CLI_DIR	Expr1003	CLI_EST
6	Nicolás Rosas	20 de nov. 123	NL	NL

Nombre del fragmento: CLI_M4 Ubicación: NL-Colección Nodo: MY-C

CLI_NUM	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA	CLI_EST
6	\$2,000.00	\$350.00	2	\$50.00	NL

Nombre del fragmento: CLI_M5 Ubicación: 'VE'-Servicio Nodo: XAL-S

CLI_NUM	CLI_NOM	CLI_DIR	Expr1003	CLI_EST
2	Martin Porres	Blvd. Lopez Mz	VE	VE
4	Benito López	Rueda 23	VE	VE
5	Victor Pérez	Carlos Carrillo	VE	VE

Nombre del fragmento: CLI_M6 Ubicación: 'VE'-Colección Nodo: XAL-C

CLI_NUM	CLI_LIM	CLI_BAL	CLI_NIVEL	CLI_DEUDA	CLI_EST
2	\$6,000.00	\$12,000.00	1	\$0.00	VE
4	\$6,000.00	\$5,890.00	3	\$1,090.00	VE
5	\$1,200.00	\$550.00	1	\$0.00	VE