

## Unidad 7

# Bases de datos distribuidas

**UTPL**  
UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

# Agenda

- Introducción
- Transparencia en un SGBDD
- Clases de implementaciones distribuidas
- Diseño de una base de datos distribuida:  
Fragmentación y Replicación
- Aplicaciones de bases de datos distribuidas



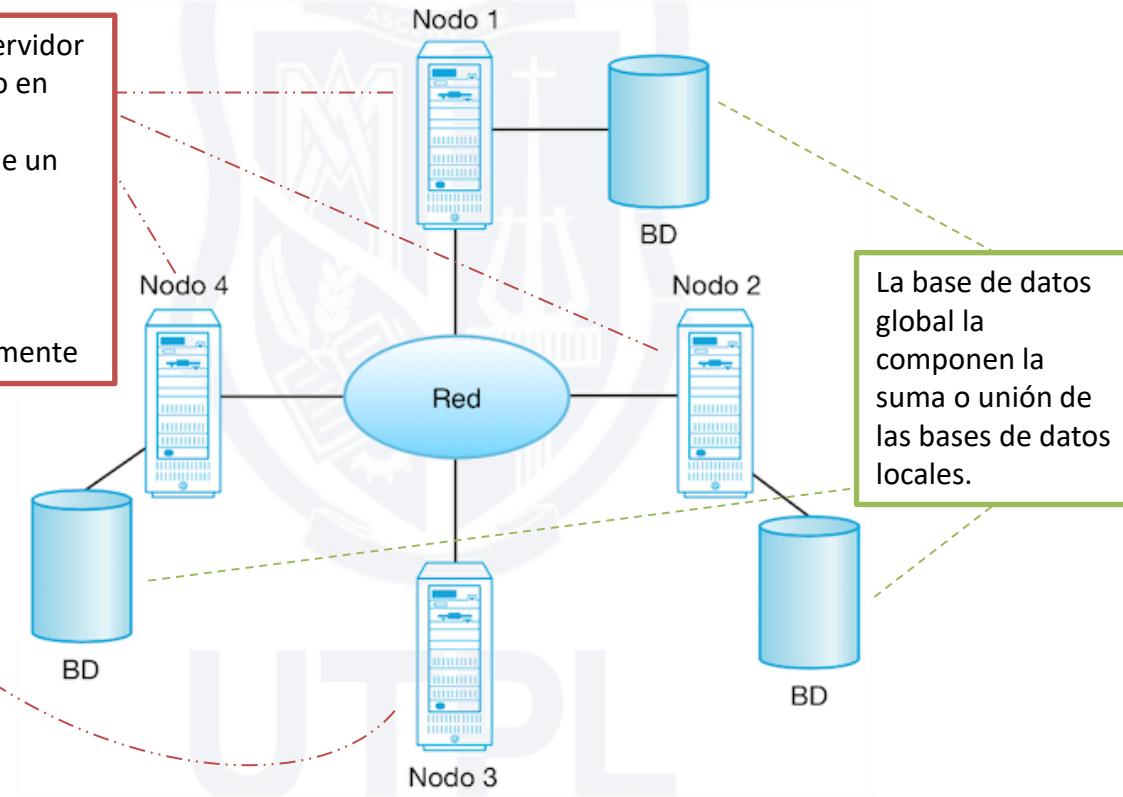
# Introducción

- Bases de Datos Distribuidas → mezcla de dos tecnologías:
  - La tecnología de base de datos.
  - La tecnología de redes y comunicación de datos.
- Las Bases de Datos Distribuidas aportan al dominio de la gestión de bases de datos las ventajas de la computación distribuida, esto es:
  - Un conjunto de elementos de procesamiento interconectados por una red cooperan en forma coordinada en la ejecución de tareas.

# Introducción

## Topología

Cada nodo es un servidor físicamente alojado en una determinada ubicación, que tiene un SGBD corriendo localmente y opcionalmente un conjunto de datos almacenados localmente



# Introducción

## Topología

- Los datos deben estar físicamente en más de una sede. Cada sede o computadora es un **nodo de la red**.
- Cada nodo proporciona un **entorno de ejecución de transacciones**, tanto **local**, como **global**:
  - **Transacciones locales**: cuando se accede a los datos de la sede que inició la transacción.
  - **Transacciones globales**: se accede a datos de sedes distintas al nodo que inició la transacción.
- **Esquema lógico global único**: Los **datos** están **lógicamente integrados**, para realizar recuperación y actualización tanto en forma local como remota.

# Introducción

## Definición

**Base de Datos Distribuida (BDD):** colección de múltiples bases de datos interrelacionadas lógicamente, distribuidas sobre una red de computadores y manejadas por un SGBDD.

**Sistema de Gestión de Bases de Datos Distribuido (SGBDD):** Sistema de software que maneja BDDs y que hace de la distribución algo **transparente** al usuario.

# Transparencia

El principio fundamental de las bases de datos distribuidas es el siguiente:

*Desde el punto de vista del usuario, un sistema distribuido deberá ser idéntico a un sistema no distribuido.*



# Transparencia

## Niveles de transparencia:

- Transparencia de distribución
- Transparencia de transacción
- Transparencia de fallas
- Transparencia de desempeño
- Transparencia de heterogeneidad

# Transparencia

- Los usuarios de un sistema distribuido deberán comportarse exactamente como si el sistema fuera local.
- Todos los problemas de los sistemas distribuidos deberían ser internos (a nivel de realización) y no externos (a nivel de usuario).
- En términos de SQL, la lógica de las operaciones SELECT, INSERT, UPDATE y DELETE no deberá sufrir cambios
- Las operaciones de definición de datos, en cambio, requerirán cierta ampliación en un sistema distribuido.



# Diseño de una base de datos distribuida

- El diseño de una base de distribuida se realiza en dos fases:
  - **FASE 1:** Desarrollar un modelo de datos de la misma forma que se haría para una base de datos centralizada.
  - **FASE 2:** Decidir cómo se distribuirán los datos en los distintos nodos del sistema distribuido.
- La premisa es que los datos se distribuyan de tal que manera que la mayoría de las transacciones se ejecuten de manera local.

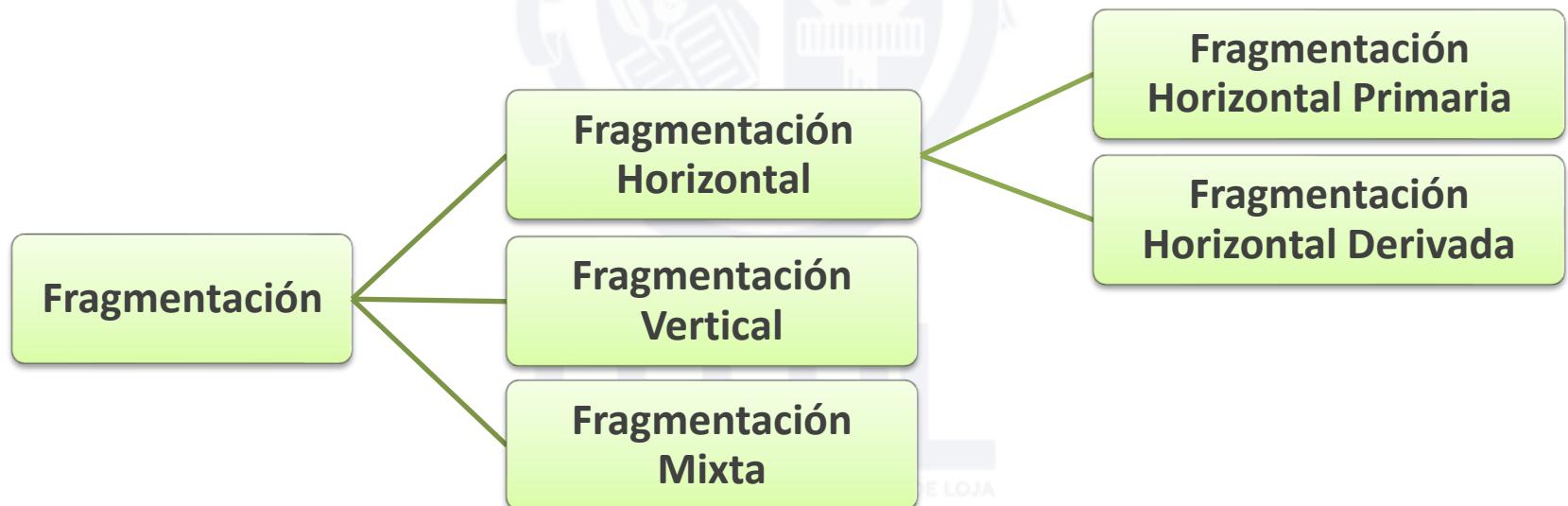
# Diseño de una base de datos distribuida

- Existen **dos patrones o esquemas de distribución:** Fragmentación y Replicación.
- La **fragmentación** divide una tabla en varias partes llamadas “fragmentos” y cada fragmento se almacena en un nodo.
- La **replicación** en cambio guarda una copia íntegra de la tabla en varios nodos del sistema distribuido.
- **Ambos patrones se utilizan** durante el proceso de diseño de una Base de Datos Distribuida.

# Fragmentación y Replicación

## Fragmentación de Datos

- La fragmentación implica dividir una unidad lógica (relación o tabla) en **unidades lógicas más pequeñas llamadas fragmentos.**
- Hay varios tipos de fragmentación:



# Fragmentación y Replicación

## Fragmentación de Datos

- Para asegurar que la base de datos no sufrirá cambios semánticos durante la fragmentación de los datos, se definen tres normas que determinan la **calidad de la fragmentación** de una relación:
  - **Completitud**.- Todos los datos han de encontrarse en al menos, un fragmento.
  - **Disyunción**.- Los fragmentos deberán ser disjuntos.
  - **Reconstrucción**.- Siempre se ha de poder reconstruir la base de datos global a partir de los fragmentos.

# Fragmentación y Replicación

## Fragmentación Horizontal

- **Fragmentación horizontal primaria:** Es un **subconjunto de las tuplas** de una relación formado por aquellas tuplas **que satisfagan una condición** sobre uno o más de sus atributos. Los fragmentos se asignan a diferentes sitios del sistema distribuido.
- **Fragmentación horizontal derivada:** la fragmentación horizontal de una relación primaria se aplica a otras relaciones secundarias, que referencian a la primaria a través de una **clave externa o foránea**. De esta forma, los datos relacionados en ambos tipos de relación se fragmentan de la misma forma.
- Para reconstruir la relación desde una fragmentación horizontal se aplica la operación **UNION**.

# Fragmentación y Replicación

## Fragmentación horizontal primaria (Ejemplo)

Tabla inicial de alumnos (T)

<u>DNI</u>	Escuela	Nombre	Nota ingreso	Beca
87633483	EUI	Concha Queta	5.6	No
99855743	EUI	Josechu Letón	7.2	Si
33887293	EUIT	Oscar Romato	6.1	Si
05399075	EUI	Bill Gates	5.0	No
44343234	EUIT	Pepe Pótamo	8.0	No
44543324	EUI	Maite Clado	7.5	Si
66553234	EUIT	Ernesto Mate	6.6	No

# Fragmentación y Replicación

## Fragmentación horizontal primaria (Ejemplo)

Fragmento T\_EUI:  $\sigma_{\text{Escuela}=\text{"EUI"}}(T)$

DNI	Escuela	Nombre	Nota ingreso	Beca
87633483	EUI	Concha Queta	5.6	No
99855743	EUI	JosechuLetón	7.2	Si
05399075	EUI	Bill Gates	5.0	No
44543324	EUI	Maite Clado	7.5	Si

Fragmento T\_EUIT:  $\sigma_{\text{Escuela}=\text{"EUIT"}}(T)$

DNI	Escuela	Nombre	Nota ingreso	Beca
33887293	EUIT	Oscar Romato	6.1	Si
44343234	EUIT	Pepe Pótamo	8.0	No
66553234	EUIT	Ernesto Mate	6.6	No

# Fragmentación y Replicación

## Fragmentación horizontal derivada (Ejemplo)

```
almacen(cod_almacen, dirección, teléfono, área)
producto(ref, nombre, descripción, cod_almacen)
```

Esta es la fragmentación horizontal primaria para almacén:

```
almacensur = select * from almacen where área= 'Sur'
almacenorte = select * from almacen where área= 'Norte'
```

Esta es la fragmentación horizontal derivada para producto

```
productosur = select p.*
from producto p, almacensur as
where p.cod_almacen = as.cod_almacen
```

```
productonorte = select p.*
from producto p, almacenorte as
where p.cod_almacen = an.cod_almacen
```

# Fragmentación y Replicación

## Replicación de datos

- **Replicación:** El sistema conserva varias copias o réplicas idénticas de una tabla. Cada réplica se almacena en un nodo diferente.
- Ventajas:
  - Disponibilidad
  - Aumento del paralelismo
  - Minimiza el tráfico de datos entre nodos
- Desventajas:
  - Aumento de la sobrecarga en las actualizaciones

# Fragmentación y Replicación

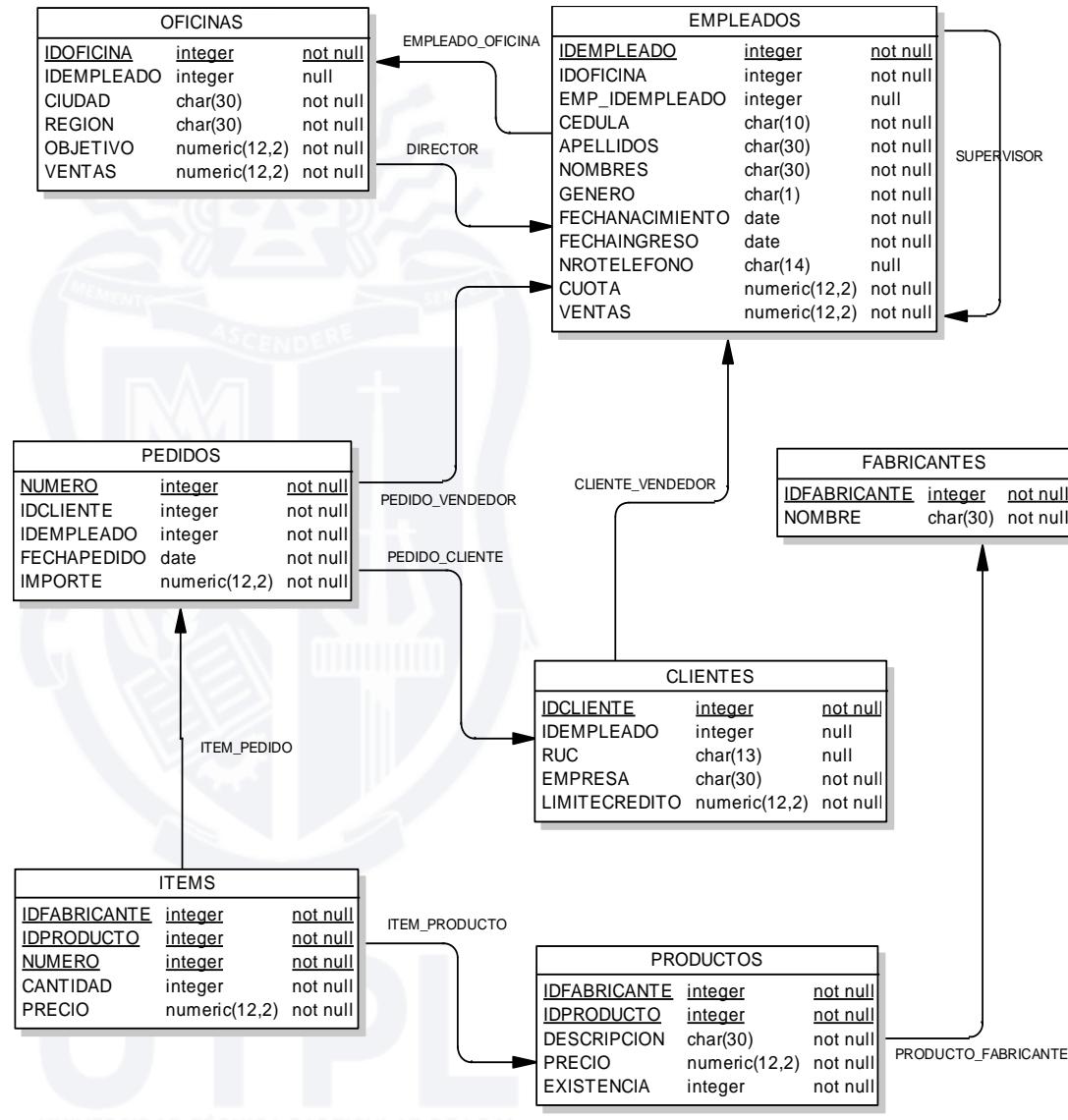
## Replicación de datos

- **Replicación completa:** toda la base de datos es replicada en todos los sitios.
  - Esto **mejora la disponibilidad** notablemente.
  - También **mejora el rendimiento de las consultas globales**.
  - La **desventaja** es que **reduce la rapidez de las operaciones de actualización**
- **Replicación parcial:** algunas partes seleccionadas de la base de datos son replicadas.
- En una base de datos distribuida se debe combinar apropiadamente **fragmentación y replicación parcial**

# Ejemplo 1

## Caso BD PEDIDOS

- Distribución con base en la ciudad
- Asumiendo dos ciudades (Quito y Guayaquil)

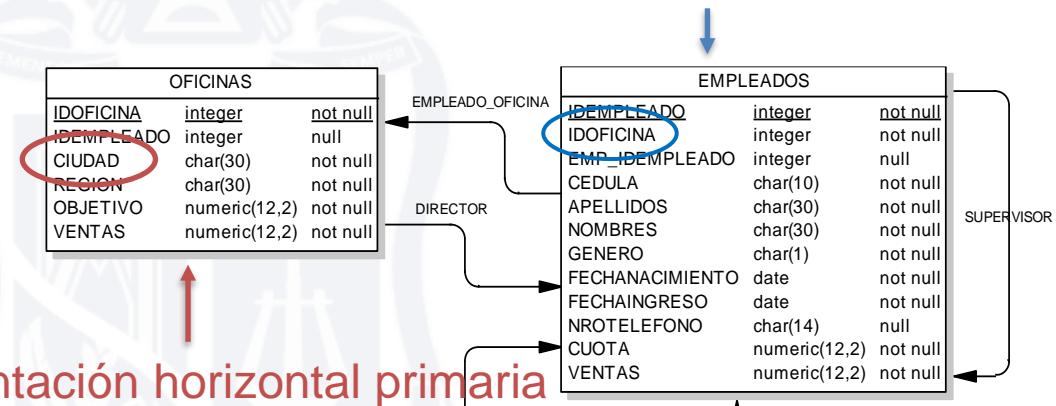


# Ejemplo 1

¿Cómo se distribuyen los datos de **OFICINAS** y **EMPLEADOS**?



Fragmentación horizontal derivada



Fragmentación horizontal primaria

$$\text{oficinas\_UIO} = \sigma_{ciudad = 'UIO'}(\text{oficinas})$$

$$\text{oficinas\_GYE} = \sigma_{ciudad = 'GYE'}(\text{oficinas})$$

$$\text{empleados\_UIO} = (\rho_e \text{empleados}) \bowtie_{e.idoficina = o.idoficina} (\rho_o \text{oficinas\_UIO})$$

$$\text{empleados\_GYE} = (\rho_e \text{empleados}) \bowtie_{e.idoficina = o.idoficina} (\rho_o \text{oficinas\_GYE})$$

# Ejemplo 1

Fragmentación horizontal primaria

Fragmentación horizontal derivada

Replicación

