



# Sistemas de Bases de Datos 2

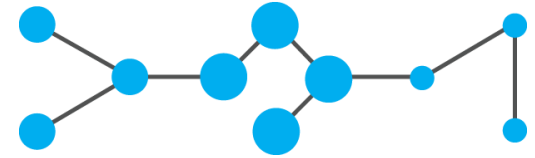
**2024**

**Ing. Luis Alberto Arias Solórzano**

**Unidad 5**

# BD Distribuidas

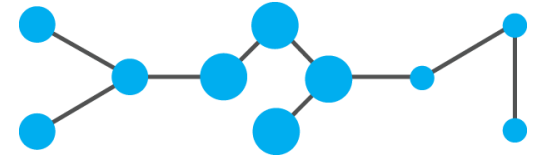
## Definiciones



- Decimos que la "base de datos distribuida" es en realidad un tipo de base de datos *virtual* cuyas partes componentes están almacenadas en varias bases de datos "reales" distintas que se encuentran en varios sitios distintos; de hecho, es la unión lógica de esas bases de datos reales.
- En un sistema distribuido de bases de datos se almacena la base de datos en varios servidores, sitios o nodos. Utilizan varios medios de comunicación para que puedan poner en contacto las distintas computadoras de un sistema distribuido. No comparten ni memoria ni discos. Las computadoras de un sistema distribuido pueden variar en tamaño y función pudiendo abarcar desde las estaciones de trabajo a los grandes sistemas.
- Cada sitio es un sitio de sistema de base de datos por derecho propio. En otras palabras, cada sitio tiene sus propias bases de datos "reales", sus propios usuarios locales, su propio DBMS local y software de administración de transacciones (incluyendo su propio software local para bloqueo, registro en bitácora, recuperación, etcétera), así como su propio administrador de comunicación de datos local.

# BD Distribuidas

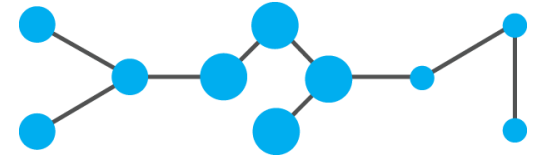
## Definiciones



- El sistema de base de datos distribuida puede ser considerado como un tipo de **sociedad** entre los DBMSs locales en cada uno de los sitios locales; un nuevo componente de software en cada sitio —de manera lógica, una extensión del DBMS local— proporciona la funcionalidad de sociedad necesaria, y es la combinación de este nuevo componente y el DBMS existente, lo que constituye lo que generalmente llamamos **sistema de administración de base de datos distribuida**.
- Las principales diferencias entre las bases de datos paralelas sin compartimientos y las bases de datos distribuidas son que las bases de datos distribuidas normalmente se encuentran en varios lugares geográficos distintos, se administran de forma separada y poseen una interconexión más lenta. Otra gran diferencia es que en un sistema distribuido se dan dos tipos de transacciones, las locales y las globales.
- Otra gran diferencia es que en un sistema distribuido se dan dos tipos de transacciones, las locales y las globales. Una **transacción local** es aquella que accede a los datos del único sitio en el cual se inició la transacción. Por otra parte, una **transacción global** es aquella que, o bien accede a los datos situados en un sitio diferente de aquel en el que se inició la transacción, o bien accede a datos de varios sitios distintos.

# BD Distribuidas

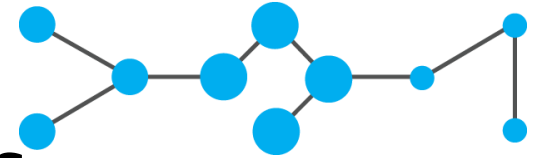
## Tipos



- **Homogéneas:** En las **bases de datos distribuidas homogéneas** todos los sitios tienen idéntico software de sistemas gestores de bases de datos, son conscientes de la existencia de los demás sitios y acuerdan cooperar en el procesamiento de las solicitudes de los usuarios. En estos sistemas los sitios locales renuncian a una parte de su autonomía en cuanto a su derecho a modificar los esquemas o el software del sistema gestor de bases de datos. Ese software también debe cooperar con los demás sitios en el intercambio de la información sobre las transacciones para hacer posible el procesamiento de las transacciones entre varios sitios.
- **Heterogéneas:** en las **bases de datos distribuidas heterogéneas** sitios diferentes puede que utilicen esquemas diferentes y diferente software de gestión de sistemas de bases de datos. Puede que unos sitios no sean conscientes de la existencia de los demás y puede que sólo proporcionen facilidades limitadas para la cooperación en el procesamiento de las transacciones. Las diferencias en los esquemas suelen constituir un problema importante para el procesamiento de las consultas, mientras que la divergencia del software supone un inconveniente para el procesamiento de transacciones que tengan acceso a varios sitios.

# BD Distribuidas

## Objetivos



Considérese una relación  $r$  que hay que almacenar en la base de datos. Hay dos enfoques del almacenamiento de esta relación en la base de datos distribuida:

- **Réplica.** El sistema conserva réplicas (copias) idénticas de la relación y guarda cada réplica en un sitio diferente. La alternativa a las réplicas es almacenar sólo una copia de la relación  $r$ .
- **Fragmentación.** El sistema divide la relación en varios fragmentos y guarda cada fragmento en un sitio diferente.

La fragmentación y la réplica pueden combinarse: Las relaciones pueden dividirse en varios fragmentos y puede haber varias réplicas de cada fragmento.



Gracias