

# 1. ¿Qué es una Base de Datos Distribuida?

Una **base de datos distribuida (BDD)** es un sistema en el cual los datos se almacenan en diferentes ubicaciones físicas, pero se presentan a los usuarios como una única base lógica.

Definiciones de fuentes reconocidas:

- **Gartner:** “Una base de datos cuyos objetos (tablas, vistas, columnas y archivos) residen en más de un sistema en una red, y pueden ser accedidos o actualizados desde cualquier sistema de la red”.
- **Wikipedia:** “Un sistema de base de datos en el que los datos se almacenan en distintas ubicaciones físicas a través de una red de ordenadores, pero presentándose como una única base lógica”.
- **Cockroach Labs:** “Una base de datos distribuida es aquella que se ejecuta y almacena datos en múltiples computadoras, en lugar de todo en una sola máquina”.

Características principales:

- Uso de múltiples nodos o sitios.
- Transparencia para el usuario (se percibe como una sola base).
- Soporte de replicación y fragmentación.
- Mayor tolerancia a fallos y escalabilidad.

## 2. Diferencias entre Base de Datos Centralizada y Distribuida

Aspecto	Base Centralizada	Base Distribuida
Ubicación de datos	Todos en un único servidor	Repartidos en varios nodos
Tolerancia a fallos	Baja (punto único de falla)	Alta (replicación y redundancia)
Escalabilidad	Vertical (más hardware)	Horizontal (más nodos)
Latencia	Puede ser alta en acceso remoto	Optimizada acercando datos al usuario
Complejidad	Menor, más fácil de administrar	Mayor: requiere coordinación de nodos
Costo	Inicialmente más económico	Puede ser más caro pero más flexible

### 3. Tipos de Bases de Datos Distribuidas

1. Según homogeneidad:

- **Homogéneas:** mismos SGBD y esquemas en todos los nodos.
- **Heterogéneas:** diferentes sistemas y esquemas en cada nodo.

2. Según arquitectura:

- **Cliente-servidor distribuido:** nodos clientes consultan a servidores de datos.
- **Peer-to-peer (P2P):** todos los nodos tienen roles equivalentes.
- **En la nube:** los datos se distribuyen en infraestructura cloud.

3. Según replicación:

- Totalmente replicada.
- Parcialmente replicada.
- Sin replicación.

4. Según fragmentación:

- Fragmentación horizontal (por filas).
- Fragmentación vertical (por columnas).
- Híbrida (combinación).

### 4. Enfoques para el Diseño de BDD

- **Fragmentación de datos:** división en fragmentos horizontales, verticales o mixtos.
- **Replicación:** copias completas o parciales para mejorar disponibilidad.
- **Asignación:** decidir en qué nodo se guardan fragmentos o réplicas.
- **Transparencia:** el usuario no debe notar dónde están los datos.
- **Optimización de consultas:** dividir consultas globales en subconsultas locales y combinarlas.
- **Gestión de transacciones distribuidas:** garantizar propiedades ACID mediante protocolos como *two-phase commit*.

## 5. Consulta Distribuida y Consulta Remota

- **Consulta remota:** se accede desde un nodo a una tabla que reside en otro nodo.
  - Ejemplo: un servidor en México consulta directamente una tabla almacenada en un nodo de España.
- **Consulta distribuida:** involucra datos repartidos en varios nodos.
  - El sistema divide la consulta en subconsultas, las ejecuta localmente en cada nodo y combina los resultados en uno solo.

## 6. ¿Qué es una Base de Datos en la Nube?

Una **base de datos en la nube** es aquella que funciona sobre infraestructura cloud y se ofrece como servicio (DBaaS), en lugar de instalarse en servidores propios.

### Ventajas:

- Escalabilidad bajo demanda.
- Mantenimiento delegado al proveedor.
- Alta disponibilidad y replicación geográfica.
- Pago por uso.

### Desafíos:

- Latencia de red.
- Costos variables.
- Seguridad y cumplimiento normativo.
- Dependencia del proveedor (*vendor lock-in*).

Ejemplo: **Google Spanner** (base distribuida global en la nube).