



# TAREA #1

## Bases de Datos Distribuidas

### Descripción breve

Reporte escrito sobre distintos conceptos de BDD

- Urrutia González Brenda
  - Olea García Alan
- Juárez Anguiano Mario Alexis

Equipo 03  
14/03/2025

Las Bases de Datos Distribuidas son un conjunto de múltiples Bases de Datos que nos permiten almacenar y procesar datos; se encuentran lógicamente relacionadas a través de computadoras interconectadas entre sí.

Existen diversos conceptos, diferencias, estrategias, etc. de las BDD (Bases de Datos Distribuidas), las cuales se mencionarán en el presente reporte.

### Conceptos de bases de datos distribuidos

Algunos conceptos clave para poder entender las Bases de Datos Distribuidas son:

- Fragmentación: Literalmente se fragmentan o dividen los datos en partes distintas para poder distribuirse de forma más fácil.
- Replicación: Es una copia de los datos en cada nodo para evitar fallos.
- Transparencia de distribución: No le permite al usuario visualizar la distribución de los datos.
- Control de concurrencia: Se encarga de coordinar el acceso a los datos de manera simultánea, evitando de esta manera los conflictos entre intercambios que puedan existir.
- Esquema híbrido: Combinación del esquema de fragmentación y replicación.
- Esquema cliente-servidor: Es una arquitectura donde los clientes envían solicitudes a un servidor y este responde con lo solicitado.
  - Tradicional: Un servidor (central) maneja todas las peticiones.
  - Distribuido: Los datos se encuentran en varios servidores interconectados y se accede a ellos desde distintos sitios.
  - P2P: Cada nodo actúa como cliente y servidor a la vez. *(Colbert, 2014)*

### Diferencias entre BD y BDD

Entre las múltiples diferencias, mencionaremos las más importantes:

- En BDD, los datos se encuentran en diversos servidores de la red, mientras que en BD todos los datos están en un mismo servidor o ubicación.
- En BDD, si un nodo<sup>1</sup> falla, otros pueden responder; en BD, si el servidor que almacena la información falla, todo el sistema falla.
- En cuanto a escalabilidad, en BDD se tiene la posibilidad de añadir más servidores si es necesario, mientras en BD la escalabilidad es limitada, pues depende de los recursos del servidor principal.
- En BDD se puede encontrar mayor redundancia, lo cual implica que se puede tener replicación de datos en diversos nodos para evitar pérdida de información.
- BD es mucho más sencillo de administrar ya que únicamente se utiliza un servidor.

---

<sup>1</sup> Un nodo es cualquier máquina o servidor que forma parte del sistema distribuido y almacena, procesa o gestiona datos.

La principal diferencia entre ambas es que BD funciona con un solo archivo de base de datos, mientras que BDD utiliza diversos archivos almacenados en varias computadoras que están conectadas en la red. *(M.I.D. Salazar, s.f.)*

### Arquitectura de BDD

La arquitectura de bases de datos distribuidas define la organización y el diseño de sistemas en los que los datos no se encuentran en un único lugar, sino que están distribuidos en múltiples nodos interconectados. Este tipo de arquitectura permite mejorar la disponibilidad, el rendimiento y la tolerancia a fallos en comparación con una base de datos centralizada *(We Learn Data, s.f.)*.

Existen tres enfoques principales en la arquitectura de bases de datos distribuidas:

1. **Arquitectura Cliente-Servidor:** En este modelo, la base de datos se encuentra en un servidor que gestiona las peticiones de múltiples clientes. Los clientes envían consultas y reciben respuestas sin necesidad de almacenar datos localmente. Esta arquitectura permite una mejor administración de los recursos y facilita la seguridad y el mantenimiento del sistema *(Elmasri & Navathe, 2015)*.
2. **Arquitectura de Múltiples Servidores:** También conocida como arquitectura de bases de datos distribuidas puras, en este modelo los datos están distribuidos entre varios servidores que pueden trabajar de manera independiente o coordinada. Cada servidor puede gestionar una parte de la base de datos y responder a consultas de los clientes de forma eficiente. Este diseño mejora la escalabilidad y la tolerancia a fallos, ya que la carga se distribuye entre varios nodos *(Weiss & Anderl, 2020)*.
3. **Arquitectura Centralizada:** En este enfoque, aunque los usuarios pueden acceder a los datos desde distintos puntos de la red, toda la información se almacena en un único servidor central. Aunque este modelo simplifica la administración y la seguridad, puede convertirse en un cuello de botella en términos de rendimiento y disponibilidad, ya que la dependencia de un único servidor puede generar problemas si esta falla *(Basesdedatos's Weblog, 2007)*.

Cada uno de estos modelos tiene ventajas y desafíos dependiendo de las necesidades de la organización y de los recursos disponibles para la gestión de la base de datos distribuida.

### Estrategias para distribuir una BD

La distribución de una base de datos puede realizarse mediante diversas estrategias que buscan mejorar el rendimiento, la disponibilidad y la escalabilidad del sistema. A continuación, se detallan algunas de las principales:

1. **Fragmentación (Particionamiento):** Consiste en dividir la base de datos en fragmentos más pequeños que pueden almacenarse en diferentes nodos. Existen varios tipos: fragmentación horizontal primaria, fragmentación horizontal derivada, fragmentación vertical y fragmentación mixta *(Silberschatz, Korth & Sudarshan, 2020)*.

2. Replicación: Implica mantener copias idénticas de la base de datos en múltiples nodos. Esto mejora la disponibilidad y la tolerancia a fallos. Puede implementarse de forma completa (cada nodo contiene toda la base) o de manera selectiva (**Özsu & Valduriez, 2020**).

3. Middleware: Se refiere a una capa de software que facilita la comunicación entre aplicaciones distribuidas y bases de datos. Coordina transacciones distribuidas, maneja la transparencia de ubicación y optimiza consultas (**Elmasri & Navathe, 2015**).

Además, se deben considerar criterios importantes como la localidad de los datos, la fiabilidad, los costos de almacenamiento, la distribución de carga, el costo de comunicación y el tipo de uso del sistema (**Özsu & Valduriez, 2020**).

### Gestores y funciones para BDD: Relacionales y NoSQL

Las bases de datos distribuidas pueden clasificarse como Relacionales o NoSQL, dependiendo del modelo de datos que utilizan. A continuación, se presentan los gestores más comunes y sus funciones principales.

*Bases de datos relacionales:*

- MySQL: Multiplataforma, rápido y de código abierto. Comúnmente usado en aplicaciones web.
- PostgreSQL: Orientado a objetos, extensible y conforme al estándar SQL.
- SQL Server: Propiedad de Microsoft, optimizado para empresas con grandes volúmenes de datos.
- Oracle Database: Utilizado en entornos empresariales críticos por su rendimiento y seguridad (**INESIEM, s.f.**).

*Bases de datos NoSQL:*

- MongoDB: Almacena documentos en formato JSON. Flexible y escalable.
- Cassandra: Optimizada para grandes volúmenes de datos y alta disponibilidad.
- DynamoDB: Diseñada por Amazon. Escalabilidad automática y baja latencia.
- Couchbase Server: Combina modelo documental y clave-valor.
- RavenDB: Ofrece soporte para transacciones distribuidas y replicación (**E-Dea, s.f.**).

Funciones clave de los gestores distribuidos:

- Replicación de datos
- Particionamiento (sharding)
- Garantía de consistencia y disponibilidad
- Escalabilidad horizontal
- Gestión de transacciones (ACID en relacionales y modelos flexibles en NoSQL) (**Silberschatz, Korth & Sudarshan, 2020**).

#### Bibliografías:

- [1] conceptos de bases de datos distribuidos - Search Videos. (s. f.). <https://www.bing.com/videos/riverview/relatedvideo?&q=conceptos+de+bases+de+datos+distribuidos&&mid=569293741EEA4419EE03569293741EEA4419EE03&&FORM=VRDGAR>
- [2] 5.1.3 Diferencia entre una Base de Datos Centralizada y una Distribuida. (s. f.). [http://cidcame.uaeh.edu.mx/lcc/mapa/PROYECTO/libro21/513\\_diferencia\\_entre\\_una\\_base\\_de\\_datos\\_centralizada\\_y\\_una\\_distribuida.html](http://cidcame.uaeh.edu.mx/lcc/mapa/PROYECTO/libro21/513_diferencia_entre_una_base_de_datos_centralizada_y_una_distribuida.html)
- [3] Elmasri, R., & Navathe, S. B. (2015). Fundamentals of Database Systems. Pearson.
- E-Dea. (s.f.). Gestores de bases de datos NoSQL. Recuperado de <https://www.e-dea.co/blog/gestores-bases-de-datos-nosql/>
- [4] INESIEM. (s.f.). Los gestores de bases de datos más usados. Recuperado de <https://www.inesem.es/revistadigital/informatica-y-tics/los-gestores-de-bases-de-datos-mas-usados/>
- [5] Özsu, M. T., & Valduriez, P. (2020). Principles of Distributed Database Systems. Springer.
- [6] Silberschatz, A., Korth, H. F., & Sudarshan, S. (2020). Database System Concepts. McGraw-Hill Education.
- [7] Elmasri, R., & Navathe, S. (2015). Fundamentals of Database Systems. Pearson.
- [8] Weiss, S., & Anderl, R. (2020). Distributed Database Systems: Principles and Applications. Springer.
- [9] We Learn Data. (s.f.). Database Architectures.
- Basesdedatos's Weblog. (2007). Conceptos de bases de datos.