# APUNTES DE CLASE – CLASE 3 – PRESENCIAL

## Bases de Datos Transaccionales

Almacenan transacciones

Su fin es enviar y recepcionar datos a **gran velocidad.**

Se orientan generalmente al análisis de calidad, datos de producción e industrial.

Transacción 🡪 composición de datos 🡪 evento

**Objetivo principal:** Asegurar la ***consistencia*** de las transacciones dentro de una DB.

En caso de que no se puedan confirmar, se deben poder **revertir.**

Las transacciones se realizan **completas o no se realizan**.

En resumen 🡪 integridad de datos 🡪 ACID

### Características:

* Permiten llevar a cabo un **gran número de trx**.
* Manejan datos operativos que provienen de sistemas OLTP (on-line transactional processing)
* Capturan datos sobre el **contexto histórico de la trx** para su posterior análisis
* Gran velocidad de procesamiento

Servicio 🡪 App q se ejecuta en segundo plano 🡪 Sin interacción con el usuario

### Ventajas

* Aseguran la **integridad** de los datos
* Rápidas y muy baja latencia
* …

### Desventajas

* Limitación para generar informes
* El historial de datos que facilitan a través de su consulta…

### Bases de datos relacional vs transaccional

Comprenden dos aspectos de la gestión de datos que van de la mano.

La bd **transaccional** funciona de manera asociada a una bd **relacional**

La función de la bd transaccional es asegurar que las trx dentro de la bd relacional se cumplen de manera completa.

## Arquitectura de Bases de Datos Distribuidas

Un sistema distribuido se puede ver como un *clúster* de servidores

Es un conjunto de protocolos que permite que varios sistemas de bd funcionen conjuntamente.

Tiene capas de seguridad y demás servicios que aseguran la alta disponibilidad

Vienen a solucionar los problemas de:

Escalabilidad 🡪 Capacidad de estar preparado ante el crecimiento de la base de datos

Alta disponibilidad 🡪 Si se cae un nodo, se redirigen las peticiones por otro camino, pero de igual manera se llegará a la información solicitada

**Clúster**

Conjunto de HW y SW

* Clúster en la nube: Configuración y despliegue de clústeres de servidores virtuales o físicos dentro de un entorno de computación en la nube.
* Clúster de servidores dedicados: Se refiere a un entorno de hosting en el cual un servidor físico es asignado exclusivamente a un solo cliente, organización o para un propósito específico.

Pueden desplegarse en nubes públicas, privadas o entornos híbridos.

## Bases de Datos – Data Warehouse – Data Lake

Conceptos muy relacionados entre sí, pero con propósitos distintos.

**Data Lake** 🡪 “Gran salón que guarda todo lo que le mandes”. Está preparada para almacenar datos no estructurados. No se le asigna un tipo de dato. A partir de la información se le aplica la funcionalidad ETL

**ETL** 🡪Extracción, Transformación, Carga (Load)

**Data Warehouse (DWH) 🡪** Se obtienen a partir de la información lo que se está buscando. Se aplica “inteligencia” o análisis. Tiene los algoritmos de procesamiento y extracción de datos. Transforma los datos en una estructura OLAP (Online Analytical Processing). Crea una capa de datos optimizada sobre la cual se puede analizar.

El DWH es una bd con estructura OLAP, que se encuentra por encima de las bd trx y que está diseñada para el análisis de los datos que hay en el sistema.

*DATALAKE 🡪 ETL 🡪 DATA WAREHOUSE 🡪 ANALISIS DE DATOS 🡪 TOMA DE DECISIONES*

CRM (Customer Relationship Manager) 🡪 sistemas de sw que permiten almacenar la relación con el cliente.