

Data Warehouse y OLAP

Nacen por la necesidad de conservar datos históricos para utilizarlos en la toma de decisiones.
Sistema que agrega y almacena información de una variedad de fuentes de datos.
Para su concesión se trasladan datos periódicamente a una base de datos destinada al análisis.

• **OLTP (Procesamiento transaccional en línea)**: Sistema para el manejo de transacciones.

• **OLAP (Procesamiento analítico en línea)**

Sistema para el análisis de datos para obtener tendencias que puedan servir para la toma de decisiones.
Nos permiten procesar información de los *Data Warehouse*.

Alternativas de implementación

• **MOLAP (OLAP Multidimensionales)**: los datos se almacenan en una BD multidimensional no relacional.

• **ROLAP (OLAP Relacionales)**: Los datos se almacenan en una BD relacional.

• **HOLAP (OLAP Híbridos)**: MOLAP + ROLAP.

Características

Permiten conocer el dato histórico.
Se nutren de diferentes OLTP.
No se actualizan en tiempo real.
Los datos no se borran.
Suelen ser temáticos. Si la temática es muy específica, se suelen tener *Data Warehouse* más pequeños, llamados *Data Lakes*.

Ventajas

- **Competitividad.**
- **Productividad.**
- **Posibilidad de mayores ganancias.**

Componentes

- **Instante de recogida de datos**
 - **Arquitectura dirigida por origen**: La base de datos de origen envía la información al *Data Warehouse*.
 - **Arquitectura dirigida por destino**: El *Data Warehouse* solicita a la OLTP la información.
- **Transformación y limpieza de datos.**
- **Propagación de actualizaciones.**
- **Resúmenes de datos.**

Operativa diaria (ETL)

- **Extracción**: recuperar datos de la OLTP.
- **Transformación**: adaptar los datos a la arquitectura del *Data Warehouse*.
- **Carga (Load)**: introducción de los datos al *Data Warehouse*.

Arquitectura

- **Tabla resumen**: *Dashboard*.
- **Esquema estrella**
 - **Estrella**: combina dimensiones normalizadas y otras que no lo están.
 - **Copo de nieve**: permite que las dimensiones tengan medidas propias.

Servicios

- **Teradata.**
- **Oracle.**
- **Azure Synapse Analytics.**
- **Amazon Redshigt.**