

Data Warehouse Copo de Nieve

El Data Warehouse Copo de Nieve, también conocido como Snowflake, es un tipo de diseño de base de datos utilizado en la construcción de data warehouses. Es una arquitectura de modelado de datos que se utiliza para organizar y almacenar grandes cantidades de datos de manera eficiente, especialmente para análisis y consultas de datos.

El nombre "copo de nieve" se refiere a la forma en que se organiza la estructura de la base de datos en este diseño. En una base de datos Copo de Nieve, los datos se organizan en tablas que están conectadas mediante relaciones de uno a muchos. Las tablas están normalizadas, lo que significa que se dividen en varias tablas más pequeñas para evitar la redundancia de datos y mejorar la eficiencia en el almacenamiento.

En un Data Warehouse Copo de Nieve, la tabla de hechos (que contiene datos numéricos) está conectada a varias tablas de dimensiones (que contienen datos descriptivos) a través de claves primarias y claves foráneas. Sin embargo, a diferencia de otros diseños de bases de datos las tablas de dimensiones también pueden estar normalizadas, lo que significa que pueden dividirse en más tablas más pequeñas. Esto da lugar a una estructura en forma de copo de nieve, con la tabla de hechos en el centro y las tablas de dimensiones ramificándose hacia afuera en niveles múltiples.

Las limitaciones que tiene este esquema son bajas, sin embargo, una de las más importantes es lo costoso que es mantener este tipo de esquema, debido a la amplia variación de tablas que tiene (dimensionales y sub-dimensionales), que le suman complejidad al esquema. También, a pesar de que el resultado de las consultas realizadas sea fiable, el proceso que lleva suele dedicarle una mayor cantidad de tiempo.

Algunas de las ventajas de usar este modelo, son las siguientes:

- Optimiza algunas herramientas de modelado multidimensional OLAP.
- No hay redundancia de datos.
- Esquema de datos simple y necesita poco almacenamiento de datos.

Por otro lado, las desventajas son las siguientes:

- Consultas complejas debido a la normalización de datos.
- Bajo rendimiento debido a la normalización.
- Más de una tabla de dimensión por cada dimensión.

Preguntas:

1. ¿Qué es un modelo de tipo Copo de Nieve?
2. ¿Cuál es la característica principal que diferencia este modelo con el de estrella?
3. ¿Qué tipo de rendimiento tiene el modelo de Copo de nieve (alto/bajo)?
4. ¿Por qué se le dice Copo de Nieve a este modelo?
5. ¿Cuáles son las 2 desventajas de este modelo debido a la normalización de datos?

Integrantes:

- Alexander Romero Mitiaeva (0248744)
- Adrián Reyes Mercado (0231620)
- Camila Gómez Salas (0247507)