

UNIVERSIDAD TECNOLÓGICA DE CHIHUAHUA
TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN



III.1. ANÁLISIS SUPERVISADO

MATERIA: Extracción de Conocimiento en Bases de Datos

MAESTR@: Enrique Mascote

ALUMNO: Carlos Adrián Mata Nevárez

Matricula: 1123250015

GRUPO: IDGS91N

FECHA: 29/11/2025

ÍNDICE

| | |
|--|---|
| INTRODUCCIÓN | 1 |
| CARACTERÍSTICAS CLAVE | 1 |
| VENTAJAS Y DESVENTAJAS..... | 2 |
| Ventajas..... | 2 |
| Desventajas | 2 |
| Proceso de implementación..... | 3 |
| CONSIDERACIONES FINALES | 4 |
| Desafíos Comunes en la Replicación | 4 |
| Recomendaciones Esenciales para la Estabilidad | 5 |
| REFERENCIAS..... | 6 |

INTRODUCCIÓN

Oracle GoldenGate es una plataforma de replicación y captura de datos en tiempo real que permite mover, integrar y sincronizar información entre múltiples bases de datos heterogéneas.

En un proyecto de Data Warehouse, GoldenGate se utiliza para obtener datos actualizados con mínima latencia, asegurando que el DW refleje los cambios operativos casi al instante (datascientest, 2025).

CARACTERÍSTICAS CLAVE

1. Replicación en tiempo real

Permite mover cambios desde sistemas transaccionales hacia el Data Warehouse con retraso mínimo, ideal para análisis con información fresca (datascientest, 2025).

2. Captura basada en logs (CDC)

Utiliza Change Data Capture leyendo los redo/transaction logs, evitando cargar al sistema fuente y garantizando rendimiento.

3. Alta compatibilidad

Funciona con Oracle, MySQL, SQL Server, PostgreSQL, DB2, y más. Especialmente útil cuando las organizaciones usan múltiples tecnologías.

4. Replicación unidireccional y bidireccional

Soporta múltiples topologías: una sola vía, ambas vías, multi-master, broadcast, etc.

Flexible para integraciones complejas.

5. Escalabilidad y baja latencia

Procesa grandes volúmenes con eficiencia, logrando transmisión continua incluso en entornos de alto tráfico (pandorafms, 2021).

VENTAJAS Y DESVENTAJAS

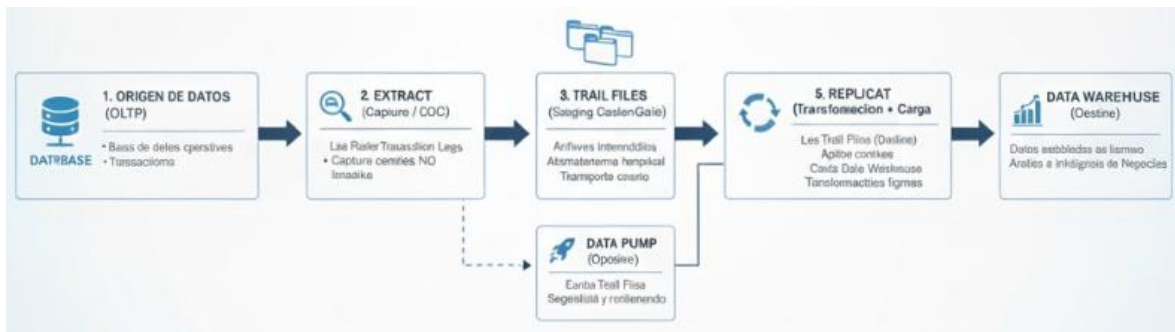
Ventajas

- Baja carga en el sistema origen gracias a la captura basada en logs.
- Replicación casi en tiempo real, ideal para DW modernos.
- Alta disponibilidad y tolerancia a fallas mediante topologías redundantes (datascientest, 2025).

Desventajas

- Licenciamiento costoso, especialmente para empresas pequeñas.
- Requiere personal capacitado, su configuración puede ser compleja.
- Dependencia de la infraestructura Oracle, aunque es multi-plataforma, su mejor rendimiento es en Oracle DB (Jesús, 2024).

Proceso de implementación



Origen de datos (OLTP)

Bases de datos operativas generan transacciones.

Extract (Capture / CDC)

GoldenGate lee los archivos de redo/transaction logs para capturar cambios de forma no invasiva.

Trail Files (Staging GoldenGate)

Los cambios capturados se almacenan en archivos intermedios (trail) para su transporte.

Data Pump (Opcional)

Envía los trail files a otro servidor GoldenGate para mayor seguridad y rendimiento.

Replicat (Transformación ligera + Carga)

Aplica los cambios en el Data Warehouse. Puede incluir conversiones de tipos, renombrado de campos o filtros simples.

Data Warehouse

Recibe los datos actualizados en tiempo real.

CONSIDERACIONES FINALES

Desafíos Comunes en la Replicación

Durante la operación de un sistema de replicación, es habitual enfrentarse a varios desafíos técnicos que deben ser gestionados proactivamente:

Evitar los Cuellos de Botella: Ajustar los parámetros de los procesos capture y replicat es un reto constante. Se debe encontrar el equilibrio justo para asegurar que la información se mueva sin sobrecargar los recursos ni ralentizar el flujo.

Gestión de Conflictos: Cuando se implementa una replicación bidireccional (datos fluyendo en ambas direcciones), el mayor obstáculo es manejar los conflictos que surgen cuando se intenta actualizar el mismo registro simultáneamente en dos bases de datos distintas.

Asegurar la Retención de Logs: Es crucial garantizar que los logs de transacciones en la base de datos de origen tengan una retención suficiente. Si los logs se borran antes de que el proceso de captura pueda leerlos, se perderá la capacidad de replicar esas transacciones.

Recomendaciones Esenciales para la Estabilidad

Para mitigar estos riesgos y mantener el sistema operativo a largo plazo, se sugieren varias prácticas probadas:

Monitoreo Continuo del Lag: Se debe vigilar constantemente el retraso (lag) de la replicación. Esto permite detectar de forma temprana si la información se está acumulando en los archivos intermedios (trail files) y tomar acciones correctivas antes de que se convierta en un problema grave.

Utilizar el Data Pump: Implementar el proceso Data Pump en la arquitectura proporciona una capa extra de resiliencia y seguridad durante el transporte de los datos, mejorando la fiabilidad general del sistema.

Arquitectura de Alta Disponibilidad (HA): Es vital mantener los servidores que alojan GoldenGate en un esquema de alta disponibilidad. Esto asegura que, si un servidor falla, el proceso de replicación pueda ser asumido inmediatamente por otro nodo, evitando interrupciones.

Documentación Rigurosa: Para facilitar el soporte futuro y la transferencia de conocimiento, es imprescindible documentar de forma clara tanto la topología completa del sistema como todos los parámetros de configuración clave.

REFERENCIAS

(19 de Oct de 2021). Obtenido de pandorafms:

<https://pandorafms.com/blog/es/monitorizar-oracle-goldengate/>

(12 de Nov de 2025). Obtenido de datascientest:

<https://datascientest.com/es/oracle-goldengate-todo-sobre>

Jesús. (31 de Mar de 2024). Obtenido de dongee:

<https://www.dongee.com/tutoriales/ventajas-y-desventajas-de-oracle/>