

SQL Replicate

La replicación es una técnica para copiar y distribuir un conjunto de datos desde una base de datos a otra. Esto permite la transferencia y sincronización de datos.

SQL Server Replication está diseñado específicamente para funcionar dentro del entorno de SQL Server, permite realizar configuraciones para adaptar la replicación con requisitos específicos.

Tipos de replicación en SQL Server

1. Replicación Instantánea

- Se crea una copia de los datos del Publicador y se distribuye a los Suscriptores. No rastrea cambios incrementales, sino que envía una instantánea completa en cada replicación.

2. Replicación Transaccional

- Comienza con una instantánea inicial de los datos, los cambios se capturan del log de transacciones del Publicador y se distribuyen en tiempo real. Es adecuado cuando se requiere que las bases de datos estén constantemente sincronizadas.

3. Replicación de Mezcla (Merge)

- En este tipo, tanto el Publicador como el Suscriptor pueden realizar cambios en los datos. Después de crear y enviar la primera instantánea de datos al suscriptor, se realiza un seguimiento de los cambios realizados en el Publicador y Suscriptor, posteriormente se fusionan los cambios.

4. Replicación punto a punto:

- Es un tipo de replicación transaccional donde se permite que varios servidores contengan copias idénticas de los datos. Cualquier cambio realizado en cualquier servidor se replica en todos los demás servidores. Lo que garantiza que cada servidor contenga los datos más recientes.

Componentes Principales

1. Publicador (Publisher)

- Es la base de datos que contiene los datos a replicar. Es la fuente de datos.
- El Publicador no solo almacena los datos, tiene la tarea de rastrear los cambios realizados en los datos y publicarlos para que los suscriptores los reciban.

2. Suscriptor (Subscriber)

- Es la base de datos que recibe los datos replicados desde el Publicador.
- Pueden optar por recibir todos los datos publicados o solo una parte de ellos, dependiendo de la configuración de la replicación.

3. Distribuidor (Distributor)

- Es un servicio que actúa como intermediario entre el Publicador y el Subscriber. Su principal tarea es almacenar las transacciones o los cambios que se publican y distribuirlos a los Suscriptores.

4. Publicación (Publication)

- Es un conjunto de datos u objetos de datos (tablas, vistas, etc.) que se van a replicar desde el Publicador hacia el Subscriber.
- Define los datos que se van a distribuir.
- Ejemplo: Si existe una tabla de vuelos que se necesita replicar, se crea una publicación que incluya esta tabla. Cualquier cambio realizado en la tabla del Publicador se replicará en el Subscriber.

5. Artículo (Article)

- Es el objeto individual dentro de una Publicación que se replica.
- Pueden ser tablas, vistas o cualquier otro objeto de base de datos.
- Una publicación puede contener uno o más artículos.
- Ejemplo: Si existe una tabla de vuelos y una tabla de pasajeros dentro de una Publicación, cada una de estas tablas es un artículo dentro de esa publicación.

6. Suscripción (Subscription)

- Es la solicitud de un Suscriptor para recibir una copia de los datos que el Publicador ha puesto a disposición a partir de una Publicación.
- Hay dos tipos de Suscripciones:
 - i. Suscripción de inserción (Pull): El publicador envía los datos al suscriptor.
 - ii. Suscripción de extracción (Push): El Suscriptor solicita los datos desde el Publicador.

7. Agentes (Agents)

- Son procesos que automatizan y gestionan las diferentes etapas de la replicación. Los agentes principales son:
 - i. Agente de Lectura del Log:
 - Es responsable de leer el log de transacciones en el Publicador y mover las transacciones que se han realizado en la base de datos hacia el Distribuidor.
 - Se hace uso de la replicación transaccional.
 - ii. Agente Distribuidor:
 - Se encarga de mover los datos del Distribuidor a los Suscriptores.
 - iii. Agente de Instantánea:
 - Genera una instantánea completa de los datos y objetos de la base de datos del Publicador y los envía a los Suscriptores.
 - iv. Agente de Mezcla:
 - Se encarga de sincronizar los datos entre el Publicador y el Suscriptor, maneja los cambios que se hagan en ambas bases de datos.

LOG de Transacciones

Registra las operaciones y transacciones que modifican datos en la base de datos. Esto incluye inserciones, actualizaciones y eliminaciones.

Algunas de las funciones principales son las siguientes:

1. Registro de transacciones

- Cuando se ejecuta una transacción que modifica los datos (UPDATE, INSERT, DELETE), el log de transacciones guarda la información sobre esos cambios.

2. Recuperación en caso de fallos

- SQL server utiliza el log de transacciones para restaurar el estado de la base de datos en caso de un fallo. Puede revertir (undo) las transacciones incompletas o replicar (redo) las transacciones confirmadas que aún no se han escrito dentro de la base de datos.

3. Rollback de transacciones

- Si una transacción falla o es cancelada, SQL server puede utilizar la información del log de transacciones para deshacer los cambios que se realizaron antes del fallo.

4. Replicación y CDC

- SQL Server utiliza el log de transacciones para capturar y distribuir los cambios realizados en la data a otros servidores, asegurando que las bases de datos se mantengan sincronizadas.

5. Copia de seguridad y recuperación a nivel

- En el modo de recuperación completa (Full Recovery Model) el log de transacciones permite realizar copias de seguridad.

Problemas comunes con el Log de Transacciones

1. Crecimiento excesivo: Si no se realizan copias de seguridad periódicas del log, el log puede crecer indefinidamente, lo que puede llenar el almacenamiento y detener el flujo.
2. Transacciones largas: Las transacciones que duran mucho tiempo pueden evitar que el log se trunque, lo que puede generar que el log crezca.
3. Log lleno: Si el log de transacciones no tiene espacio para crecer y no se puede truncar, la base de datos puede dejar de aceptar nuevas transacciones hasta que se libere espacio.

Conclusión

La replicación en SQL Server es una manera eficiente de distribuir datos entre bases de datos y mantenerlas sincronizadas. Los roles de Publisher, Suscriptor, Distribuidor y el uso del Log de Transacciones son fundamentales para que el proceso funcione de manera eficiente.

El Log de Transacciones es útil para el registro y la recuperación de transacciones, permitiendo que las bases de datos sean confiables y resilientes frente a fallos. Su manejo adecuado a través de copias de seguridad y configuración de recuperación es importante para evitar problemas de rendimiento y pérdida de datos.