

## Migrando desde Oracle a BigQuery para analítica en tiempo real

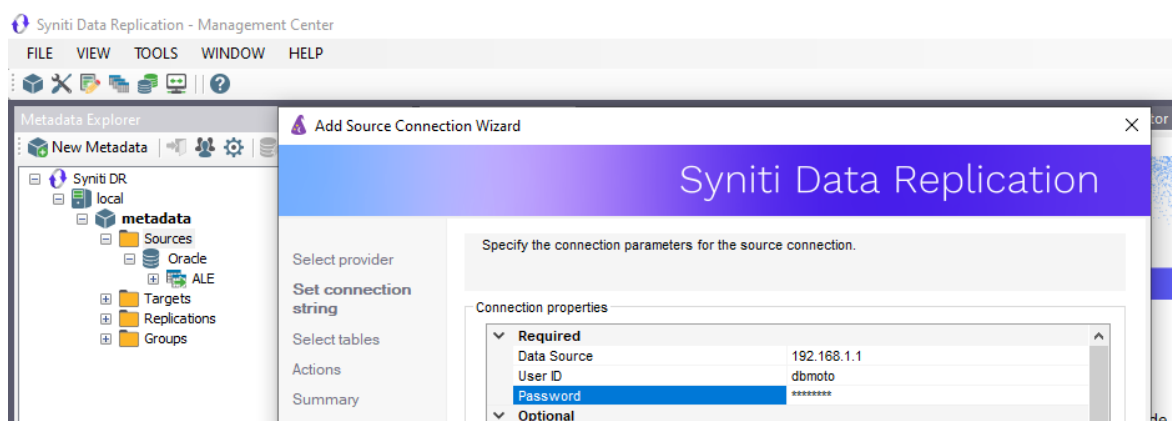
En este artículo vamos a revisar una replicación de datos en tiempo real entre una base de datos Oracle on-premise a BigQuery utilizando tecnología Change Data Capture.

Comúnmente el método para cargar información hacia un Data Warehouse es a través de cargas masivas. Hoy en día, ingerir, almacenar y manipular datos con servicios cloud como Google BigQuery hace todo el proceso fácil y con bajo impacto financiero.

Por otro lado, Syniti Data Replication le permite mover datos selectivos en tiempo real, permitiendo una replicación incremental sin impacto en los sistemas productivos, utilizando Change Data Capture para mayor precisión.

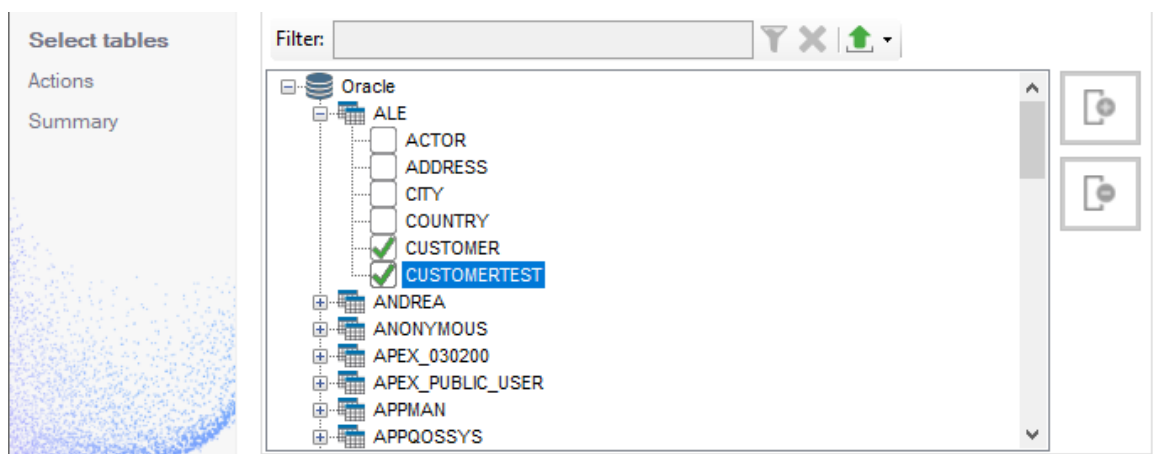
### Creando la conexión en Oracle

El primer paso que debemos hacer al momento de crear una replicación de datos en tiempo real desde Oracle, es crear la conexión de origen en el Management Center de Syniti Data Replication, ingresando la dirección IP/instancia, usuario y contraseña.



Conexión a Oracle Syniti Data Replication

A continuación, en el mismo asistente, seleccionar las tablas de Oracle que serán replicadas a BigQuery.



Selección de tablas a replicar en Syniti Data Replication

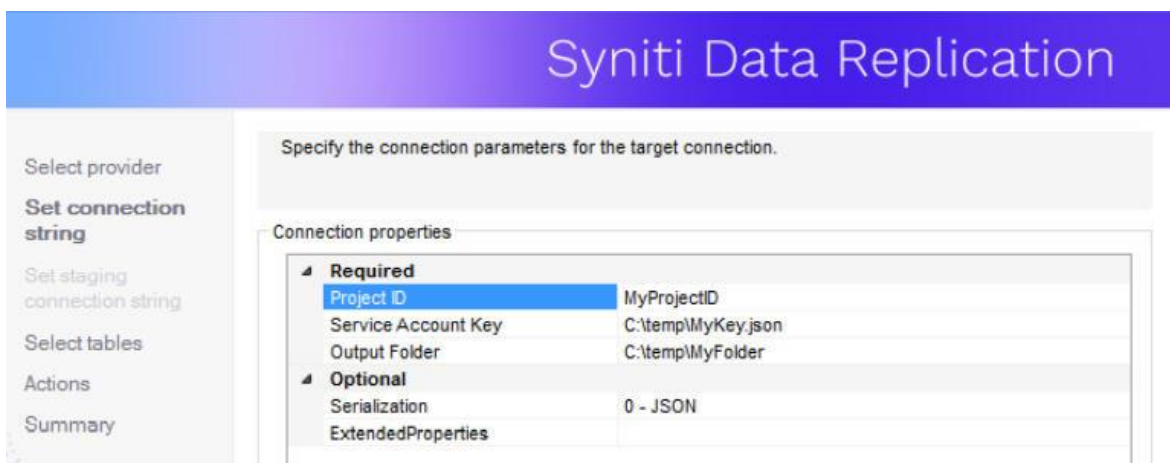
Por último, se debe habilitar la opción de replicación transaccional de Oracle, definiendo el método de lectura del Log Transaccional. Syniti Data Replication ofrece una alternativa de replicación transaccional para bases de datos que no tengan redo logs, a través de triggers, por ejemplo, para el caso de bases de datos Oracle v8.

[El presente video en Youtube](#) puede asistirlo en la creación de una conexión completa hacia una base de datos Oracle paso a paso, a través de logs transaccionales.

### Creación de la conexión a BigQuery

Antes de realizar la conexión a BigQuery usted necesita instalar el conector de datos apropiado, para ello solo debe contactar a nuestro equipo de soporte técnico, y ellos le ayudaran a obtener los componentes necesarios para realizar la conexión, incluso si usted desea utilizar serialización AVRO™.

Para crear la conexión solo debemos seguir el intuitivo asistente gráfico de Syniti Data Replication, ingresando la ubicación del proyecto, la ubicación de los servicios locales y el tipo de serialización, entre las que se encuentran disponibles CSV, JSON y AVRO



## Crear conexión en Google BigQuery

### Creando las tablas de destino en BigQuery

Si bien tablas físicas no existen en BigQuery, dentro de Syniti Data Replication debe visualizar hacia donde estamos enviando los datos para un mayor orden y rápida administración.

Para completar este paso, solo debemos tomar la tabla de origen en Oracle y soltarla sobre la conexión en BigQuery, esto abrirá un asistente que permitirá definir la ubicación que contendrán nuestros datos replicados.

### Creando las replicaciones en Syniti Data Replication

El siguiente paso es la creación de las replicaciones, en este paso se debe definir el tipo de replicación a utilizar, con BigQuery como destino, están disponibles dos poderosas alternativas de replicación:

**Refresh:** Un proceso de carga de datos a alta velocidad, el cual puede programarse para ser ejecutado a horas determinadas por el usuario.

**Change Data Capture:** El proceso de replicación transaccional, utilizando logs de Oracle para capturar las actualizaciones sin impacto significantes en los sistemas productivos.

Source log info  
Target connection  
Target log info  
Mapping info  
Scheduling  
Actions  
Summary

Replication Name

Replication Name: CUSTOMER

Description:

Use Group: <undefined> Create

Replication Type

Refresh

Continuous Mirroring

Synchronization

## Tipos de replicación para BigQuery

Una vez definido el tipo de replicación para nuestros datos y el intervalo de tiempo en el cual vamos a replicar nuestra información, debemos definir que tipos de datos vamos a replicar, para ello, Syniti Data Replication ofrece la posibilidad de descartar campos que no se requieran replicar, permitiendo dejar datos específicos en BigQuery.

Una vez definidas nuestras reglas de mapeo podemos continuar con el proceso de replicación.

Fields Mapping

[Oracle] ALE.CUSTOMER

Field name	Ordinal	Type
CUSTOMER_ID	01	NUMBER
STORE_ID	02	NUMBER
FIRST_NAME	03	VARCHAR2
LAST_NAME	04	VARCHAR2
EMAIL	05	VARCHAR2
ADDRESS_ID	06	NUMBER
ACTIVE	07	NUMBER
CREATE_DATE	08	DATE
LAST_UPDATE	09	TIMESTAMP

[PostgreSQL] Syniti.customer

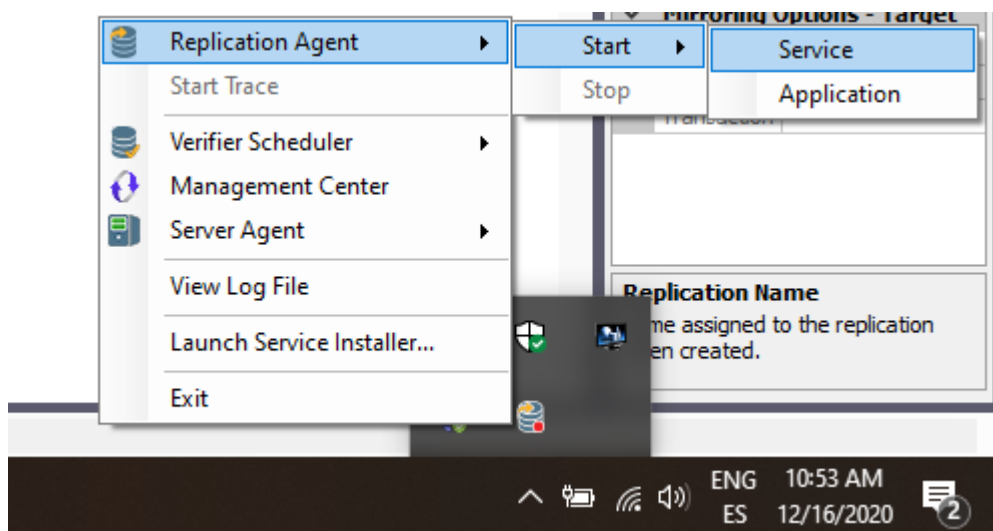
Field name	Ordinal	Type
customer_id	01	numeric
store_id	02	numeric
first_name	03	character ...
last_name	04	character ...
email	05	character ...
address_id	06	numeric
active	07	numeric
create_date	08	timestamp ...
last_update	09	timestamp ...

## Mapeo de datos en BigQuery

## Ejecutando el motor de replicaciones

Por último, solo debemos activar el servicio de replicación, desde el menú de tareas de Windows, el cual puede ser activado como un servicio en Windows o como una aplicación.

En ambientes productivos es altamente recomendado utilizar la opción de servicio de Windows, ya que esto permitirá que el servicio se levante en forma automática en caso de algún reinicio de sistema.



Ejecutar el motor de replicaciones para BigQuery

### Cuáles son los siguientes pasos

Como puede ver, el proceso de mover datos desde Oracle a BigQuery es bastante simple, rápido y no requiere de programación o conocimiento de sintaxis propietaria. Hasta este punto su sistema BigQuery contendrá las últimas actualizaciones, por ende, ya podrá avanzar en sus modelos analíticos y de machine learning, minimizando el impacto en sus sistemas productivos.

Syniti Data Replication, no solo soporta integraciones de datos entre Oracle y BigQuery, usted puede revisar que plataformas soportamos como origen y destino de datos.