

SQL*Loader - (Parte 1): Introducción

by admin - lunes, octubre 26, 2020

<https://dbandtech.com/sqlloader-parte-1-introduccion/>

SQL*Loader

- **Que es Sql Loader.**
- **Características.**
- **Funcionamiento de la Carga de Datos.**
- **Métodos de Carga: Ruta Convencional y directa.**
- **Ficheros erróneos y descartados.**

Introducción y Objetivos:

En este documento trabajaremos con la carga y descarga de datos desde la **Version 11g** de Oracle, mediante línea de comando y herramientas tales como **Enterprise Manager** y **Keep Tool**

Se deben cumplir los siguientes objetivos:

- Comprender la estructura del fichero de control de SQL*Loader.
- Conocer los ficheros de log de SQL*Loader
- Realizar cargas masivas de datos con formatos de longitud fija y variable desde la línea de comando.
- Rellenar columnas con datos derivados durante una carga masiva de datos.
- Realizar cargas de datos de un fichero a varias tablas y a una tabla desde varios ficheros.
- Realizar cargas de registros que cumplan determinadas condiciones.
- Realizar cargas de datos que incluyan campos BLOB.
- Volcar en un fichero de texto datos provenientes de las tablas de ORACLE usando

SQL*Plus y alguna herramienta grafica (TOAD, KeepTool, IxUnload o similares).

- Conocer las posibilidades que ofrece Enterprise Manager para carga y descarga de datos.

Que es SQL*Loader?

Es una utilidad que proporciona Oracle para cargar datos desde ficheros externos a una Base de Datos ORACLE. Las tablas en las que se quiere cargar información deben estar creadas ya que Sql*Loader no crea tablas, y estas tablas pueden estar vacías o con datos ya incorporados.

Características:

- Permite cargas desde ficheros de texto aunque también pueden ser binarios.
- Soporta varios Formatos de carga, carga selectiva, cargas multitaslas.

- Puede utilizarse en distintas plataformas.
- Puede leer desde múltiple ficheros de datos en una misma sesión de carga.
- Puede cargar datos en diferentes tablas en una misma sesión de carga.
- Puede manipular ficheros de longitud fija y variable
- Soporta dos tipos de ruta de carga: ruta convencional y directa
- Permite transformación de datos durante la carga

Funcionamiento de la Carga de Datos:

Para ejecutar Sql*Loader debemos ir a la línea de comando de Windows, el fichero ejecutable, en mi caso se encuentra ubicado en C:/app/oracle/product/11.2.0/dbhome_1/BIN/sqlldr.exe

En sistemas UNIX se utiliza el mismo comando.

Al escribir en la línea de comando sqlldr se nos desplegara una lista de los parámetros a utilizar.

- USERID={username[/password][@net_service_name]/} ? usuario
- CONTROL=control_file_name ? ruta y archivo de control utilizado
- LOG=path_file_name ? ruta y archivo de log
- BAD=path_file_name ?ruta y archivo generado por datos erróneos
- DATA=path_file_name ? ruta y archivo con datos de entrada
- DISCARD=path_file_name ? ruta y archivo generado con datos descartados
- DISCARDMAX=logical_record_count ? Max cantidad de archivos descartados
- SKIP=logical_record_count ? salto de registros
- LOAD=logical_record_count ? registros a cargar por defecto(all)
- ERRORS=insert_error_count ? cantidad de errores aceptados
- ROWS=rows_in_bind_array ? numero de filas que se cargan con cada INSERT
- BINDSIZE=bytes_in_bind_array ? Maximo tamano 65536b, 64k
- SILENT=[(keyword[,keyword...][)] --> à permite suprimir cabeceras, mensajes
- DIRECT={TRUE | FALSE} ? Tipo de path usado para la carga
- PARFILE=path_file_name ? Lee parámetros desde fichero de texto
- READSIZE=bytes_in_read_buffer ? cant. buffer utilizado en la carga
- FILE=database_datafile_name ? ruta y nombre de fichero de entrada

SQL Loader conecta con un usuario de la base de datos, se le dice cuál es el fichero de texto que contiene la información a importar a la Base de datos y cuál es el que contiene las especificaciones sobre lo que se va a hacer con esa información (archivo de control .CTL).

De los parámetros vistos anteriormente utilizaremos con más frecuencia los siguientes para trabajar mediante línea de comando:

- **userid=usuario/contrasena@orcl**
- **Control**
- **Data**
- **log**
- **Discard**

- bad

Métodos de Carga:

Carga de Ruta Convencional:

Por defecto es el utilizado y utiliza el parámetro **INSERT** y hace uso de **bind array** para cargar datos en las tablas de la Base de Datos. Cuando se utiliza este método Sql Loader compite con el resto de los procesos por la utilización de los recursos del buffer y a veces puede ralentizar las cargas.

- Pocos los datos a cargar.
- Acceso a tablas que se encuentran en un Cluster
- Permite aplicar funciones SQL a los campos de datos.
- Comprueba constraints de la base de datos.

Carga de Ruta Directa:

La carga de ruta directa analiza los datos de entrada con la descripción dada en el fichero de control y convierte los campos del fichero de entrada a su correspondiente formato en los campos de las tablas de Oracle. Disminuye tiempo de carga y uso del CPU

- Mucha cantidad de datos a cargar.
- La carga convencional da error en la conversión de datos.
- Debe ejecutarse un Script para que la base de datos esté preparada para este tipo de carga

\$ORACLE_HOME/rdbms/admin/catldr.sql

- Deshabilita constraints y triggers antes de la carga de datos. (check y foreign key)
- Rehabilita las constraints y triggers al terminar la carga de datos.

```
sqlldr userid=system/manager@oracl  
e DIRECT=TRUE CONTROL=ruta_directa.ctl
```

SQL*Loader

Ejemplo 1 desde la línea de comando "CMD":

```
sqlldr userid=usuario/contrasena@NombreBasedeDatos control=c:/carga.ct  
l data=c:/datos.dat
```

Ejemplo 2 desde la línea de comando "CMD":

```
sqlldr userid=usuario/contrasena@orcl  
control=/RUTA/DEL/FICHEROde CONTROL.CTL  
data=/RUTA/DEL/FICHERO/QUE/CONTIENE/LOS/DATOS  
log=/RUTA/GUARDAR/LOGS  
bad=/A/DONDE/VAN/DATOS/RECHAZADOS.bad  
discard=/A/DONDE/VAN/DATOS/DESCARTADOS.dsc
```

Bad Files.

Los datos erróneos son aquellos que no han sido insertados en las tablas de la base de datos porque no han podido ser leídos correctamente desde el fichero de datos de entrada o bien su inserción causa errores de incumplimiento de restricciones definidas en las tablas.

- Violación de constraint existentes.
- Falta de espacio en las tablespaces.

Discard Files:

Mientras que los datos descartados no se insertan en la base de datos porque no verifican una determinada condición que puede imponerse en el fichero de control, de manera que tan solo se inserten los datos que satisfacen dicha condición.

Si estos dos últimos archivos no son especificados, se crearan automáticamente al encontrarse datos erróneos y descartados.

Estos parámetros pueden pasarse bien desde la línea de comando o pueden ir incluidos dentro del fichero de control, también pueden ir incluidos dentro del fichero de control los datos a ser cargados cuando es poca la cantidad a cargar sin especificar ruta y nombre de archivo.