SQL LOADER

Sql Loader è un Tool di Oracle che permette di importare un intero csv su una tabella, con la possibilità di specificare anche delle where condition.

1. Installare Sql Loader

Innanzitutto bisogna scaricare l’instant client per windows e i rispettivi tool; i pacchetti dovrebbero essere i seguenti:

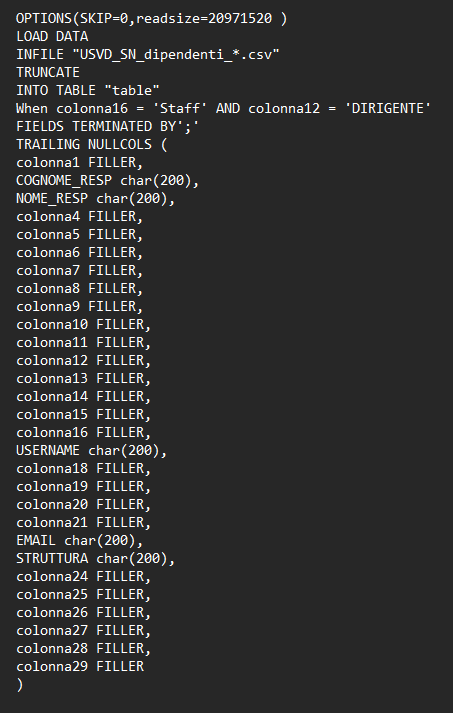
* instantclient-basic-windows.x64-21.11.0.0.0dbru
* instantclient-tools-windows

Creare su C: una cartella “oracle” dove mettere dentro, una cartella che contiene instant client e tools insieme.

Copiare il path, che più o meno dovrà essere questo “C:\oracle\instantclient\_19\_20” e inserirlo tra le variabili d’ambiente in PATH.

Siamo così pronti per utilizzare il comando sqlldr sul prompt.

1. Scrivere il Control file

Il file di controllo è un file di testo che contiene istruzioni DDL (Data Definition Language) usate per controllare i seguenti aspetti di una sessione SQL Loader:

- Dove SQL Loader troverà i dati da caricare

- Come SQL Loader prevede che i dati vengano formattati

- Come verrà configurato SQL Loader (gestione della memoria, rifiuto dei record, gestione del caricamento interrotto e così via) durante il caricamento dei dati

- Come SQL Loader manipolerà i dati caricati

Esempio di un file di controllo qui a destra :

**LOAD DATA** : inizio di un nuovo caricamento di dati;

**INFILE** : specifica il nome di un file di dati contenente i dati che si desidera caricare. Solitamente il file si indica tra apici singoli, ma nel caso in cui si metta un path si usano le virgolette (“”).

La cosa particolare, in questo caso, è L’asterisco (\*) prima del .csv.

Esso permette di generalizzare il nome del file: se ad esempio si ha l’esigenza di non tener fisso il nome, magari perché con data, in questo caso non si dovrà andare a cambiare ogni volta manualmente.

Es. USVD\_SN\_dipendenti\_YYYYMMDD.csv, in questo caso verrà preso correttamente.

**TRUNCATE**: elimina in modo rapido ed efficiente tutte le righe da una tabella o da un cluster per ottenere le migliori prestazioni possibili. Si può, insomma, specificare un metodo di caricamento per la tabella ( INSERT, APPEND, REPLACE o TRUNCATE). Questo sovrascrive il metodo di caricamento della tabella globale, che è INSERT.

* + **APPEND**: lo usiamo se la tabella è già piena e vogliamo appendere i nuovi record inseriti ai precedenti.
  + **REPLACE**: vengono eliminate tutte le righe all’interno della tabella e caricati i nuovi dati.

**INTO TABLE**: indica la tabella entro cui stiamo inserendo i dati.

**WHEN** : clausola che specifica una o più condizioni. Uno dei problemi di questo tool è che purtroppo non si possono utilizzare where clause molto complesse; non si può ad esempio, usare la OR (solo la AND) né il LIKE.

**FIELDS TERMINATED BY ‘;’** : indica come sono separati i campi all’interno del csv.

**TRAILING NULLCOLS** :

1. Lanciare i comandi

Posizioniamo la cartella con gli script su Desktop, copiamo path cartella e entriamoci dal Prompt dei comandi -> cd "path".

Una volta posizionati in quella cartella, dal prompt, possiamo lanciare il seguente comando:

sqlldr user/password@\"\(DESCRIPTION=\(ADDRESS=\(PROTOCOL=TCP\)\(Host=10.XXX.XXX.XXX\)\(Port=1521\)\)\(CONNECT\_DATA=\(SERVICE\_NAME=servicename\)\)\)\" control=FILECTL.ctl Log=Results.log ROWS=10000

- servicename ->

- FILECTL.ctl -> Control file contenente comandi su come riempire tabella

- Results.log -> File sulla quale avremo i log

- 10000 -> Max righe che usciranno come risultato

Nella sintassi DDL, utilizziamo il carattere di escape "\" quando è necessario usare dei caratteri speciali. Se ad esempio, come nel caso del comando sqlldr appena scritto, ho la necessità di utilizzare le parentesi tonde “()”, bisogna far precedere la parentesi da un escape che indicherà che la parentesi tonda deve essere presa alla lettera.