SQREAM

CLASS EXERCISE

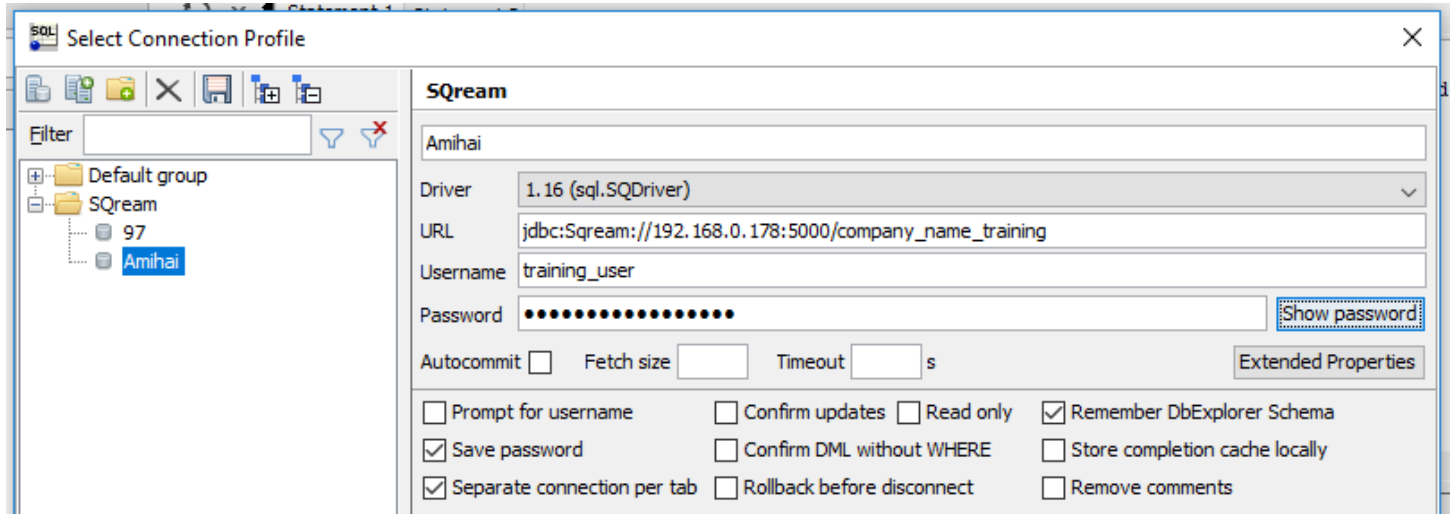
CLASS EXERCISE FOR END USERS

- SQream SQL Reference Guide: http://sqream.com/docs/latest/manual/sql_reference.html

1. Connect to a SQream DB daemon
    1. Connect to SQream Cluster using ClientCmd
        1. Direct Access:
           ```
           $ /home/sqream/sqream/bin/ClientCmd --databasename=<db_name> --username=<user_name> --password=<password> --port=<port>
           ```
        2. Load balancer:
           ```
           $ /home/sqream/sqream/bin/ClientCmd --databasename=<db_name> --username=<user_name> --password=<password> --port=3108 --clustered
           ```
    - Get parameter values from your DBA
        - Defaults are:
            - `db_name = <company_name>_training`
            - `username = training_user`
            - `password = training_password`

CLASS EXERCISE FOR END USERS

1. Connect to SQream Cluster – Cont.
    2. Connect to SQream Cluster using SQL Workbench
        1. Direct: `jdbc:Sqream//<IP>:<port>/<db>;user=<user_name>;password=<password>`
        2. Load balancer: `jdbc:Sqream//<IP>:<3108>/<db>;user=<user_name>;password=<password>;cluster=true`
        - Same parameters as above



CLASS EXERCISE FOR END USERS

2. Create Table
   - Table Name: `<student_name>_customers`
   - Fields:
     - `insert_time_stamp`
       type: `datetime not null`
     - `cust_number`
       type: `int not null`
     - `cust_name`
       type: `varchar(20) not null`
     - `cust_city`
       type: `varchar(15) null`
   - Consult "SQream SQL Reference Guide" on syntax
     - http://sqream.com/docs/latest/manual/sql_reference.html#_tables
   - Execute: "`SELECT get_ddl ('<student_name>_customers');`" to see your table definition.

   ? What is the compression type associated with your columns?

CLASS EXERCISE FOR END USERS

3. Load Data

1. Load the following 5 rows into your table using SQL `INSERT` statement
   - http://sqream.com/docs/latest/manual/sql_reference.html#_insert

insert_time_stamp	cust_number	cust_name	cust_city
'2017-07-24 11:22:33.456'	101	'AIS'	'Bangkok'
'2017-07-25 16:17:54'	102	'SQream'	'Tel Aviv'
'2017-07-26'	103	'Pizza Palace'	'Chiang Mai'
getdate()	201	'Beach Tours'	'Ko Phangan'
current_date()	202	'Diving School'	'Ko Tao'

CLASS EXERCISE FOR END USERS

3. Load Data – Cont.

2. Retrieve the loaded data using **`SELECT`** statement

   ? Note the difference in datetime representation between the first three lines.
   What do you conclude from the difference?

   - (If you are using Workbench click into the datetime value of each row to see the full value)

   ? What is the difference between **`getdate()`** and **`current_date()`**?

   ? What would be the column type for a date-only representation?

3. Insert one more row without a **`cust_city`** value:

insert_time_stamp	cust_number	cust_name
getdate()	203	'Thai Airways'

   ? What did you have to do to ensure successful load of this data as is?

CLASS EXERCISE FOR END USERS

3. Load Data – Bulk Import (`COPY FROM`)
    4. Load data into your table using the BULK Import command
        1. Create a csv file under `/home/sqream/training/<your_name>` with 3 rows of data:

```
2017-08-01,301,coca-cola,Atlanta
2017-08-01 11:22:33,302,NYSE,New York
2017-08-01 23:22:21.4,303,Disney,Los Angeles
```

            - Make sure to NOT leave an empty line at the end of the file. The last line of file should contain data.
        2. Load the csv file to your table using the `COPY` command
            - http://sqream.com/docs/latest/manual/sql_reference.html#_copy_from_bulk_import
        3. Retrieve your data using `SELECT` statement – Do you see the 3 new lines?

CLASS EXERCISE FOR END USERS

3. Load Data – Bulk Import (COPY FROM) – Cont.

    5. Error Log File

        1. Truncate your table
            - Verify table is empty by running `count(*)` query
        2. Edit your csv file by deleting the number 302 and save it
        3. Load the csv file into the table using the `COPY` command without `ERROR_LOG`
            - ? How many rows were uploaded to the table?
            - ? Why?
        4. Now load the same csv file again into the table using the `COPY` command with `ERROR_LOG <full path to error log file>`
            - ? How many rows were uploaded to the table?
            - ? What information is available in the log file?

CLASS EXERCISE FOR END USERS

3. Load Data – Bulk Import (COPY FROM) – Cont.

   6. Load the same file again into the table using the `COPY` command with `ERROR_LOG` and `ERROR_VERBOSITY 0`

      - ? How many rows were uploaded to the table?
      - ? What information is available in the log file?
      - ? In which use case could this option be useful?

CLASS EXERCISE FOR END USERS

4. Catalog Views (database, tables)
   1. Create two new tables
      - `create table t1 (a int null);`
        `create table t2 (a int not null);`
   2. Execute: `select * from sqream_catalog.tables;`
      - ? What are the `table_id` of the `t1` & `t2`?
   3. Browse the physical file location under the cluster to see the folder hierarchy according to the `table_id` you retrieved from the catalog
      - If you're unsure of cluster location you can grep for `sqreamd` process; it will point to the cluster used by it
        - » `ps aux | grep sqreamd`
      - ? How many folders does each table have?
      - ? What do these folder represent?
      - ? How many folders do you have under `Customers` table? Why?

CLASS EXERCISE FOR END USERS

4. Catalog Views (database, tables) – Cont.
    4. Execute: "`select * from sqream_catalog.columns`;"
        - Examine `column_name` & `type_name` definitions of `t1`, `t2` & `customers` tables

5. Bulk Export (COPY TO)
    1. Export data from your customers table to `/home/sqream/training/<your_name>/customers_all.csv` using "COPY TO" command with '|' delimiter
        ? Browse the exported file – how many records do you see?
    2. Export only `cust_name` and `cust_city` for `cust_numbers` greater than 302 using the tab delimiter (`\t`) to `/home/sqream/training/<your_name>/customers_gt_302.csv`
        ? Browse the exported file – how many records do you see?

CLASS EXERCISE FOR END USERS

6. Calling SQream from shell sripts
    1. create a script which would return the number of Disney records in customers table. Use **ClientCmd** with the **-c** (for command) option
        - `$ /home/sqream/sqream/bin/ClientCmd --databasename=<db_name> --username=<user_name> --password=<password> --port=<3108> --clustered -c"<your sql statement here>"`

    2. Use **ClientCmd** with **-f** option to invoke an SQL file that would:
        1. Truncate customers table
        2. Output the number of records in the table
        3. Load data from **/home/sqream/training/<your_name>/customers_all.csv**
        4. Retrieve **cust_name** and **cust_city**
        - `$ /home/sqream/sqream/bin/ClientCmd --databasename=<db_name> --username=<user_name> --password=<password> --port=<3108> --clustered -f"<full path to sql file>"`

CLASS EXERCISE FOR END USERS

7. Identity

1. Download country list from http://data.okfn.org/data/core/country-list#data and create a csv file

   `/home/sqream/training/<your_name>/country_dim.csv`

2. Create table country_dimension:
   - `country_cd int not null` identity starting at 101 skipping 2
   - `country_name nvarchar(50) not null`
   - `country_code varchar(2) not null`
3. Load the file into the table using COPY command

   ? How did you load a two column file into a 3 column table?

   ? How did you bypass loading of the column name line in the csv file?

   ? What challenge does a value containing the delimiter pose (e.g. "`Virgin Islands, U.S.`"), and how does SQream overcome it?