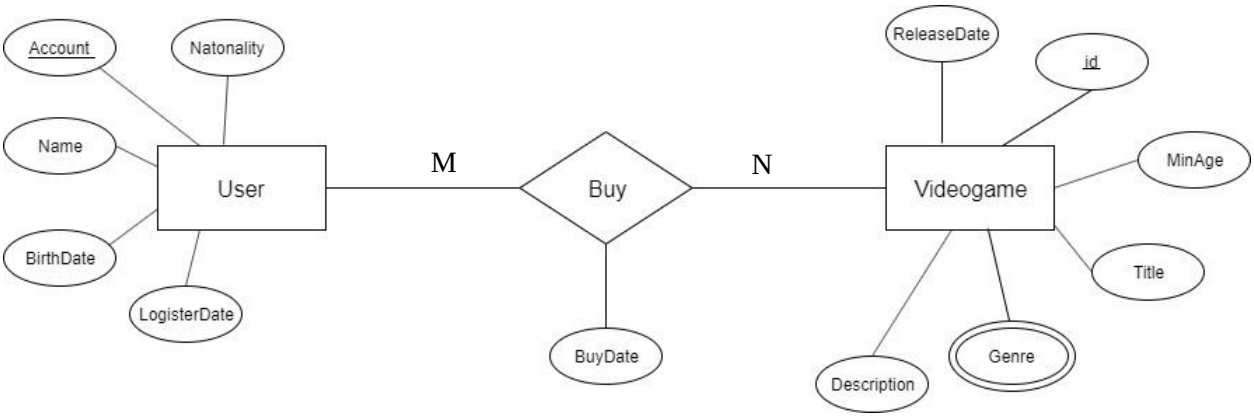
Designing a Cassandra Database

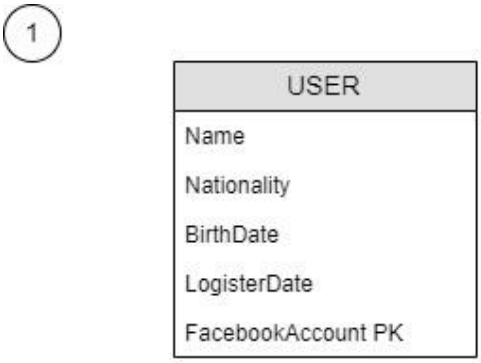
We want to design a database to store the videogames sales of the platform MAETS. The first is to translate the information to an ER Model. This model consists in two entities, one relation between them and his own attributes.



Once we have this, we need to convert the data to tables that achieve our necessities:

- Q1: User data based on his Facebook account.
- Q2: Videogame's titles bought by one user and sorted by descending buy date.
- Q3: Videogames released every month sorted by release date and the minimum age (descending).

The first query is simple: we need to create a table with all the user's info and select the Facebook (unique) as the primary key.

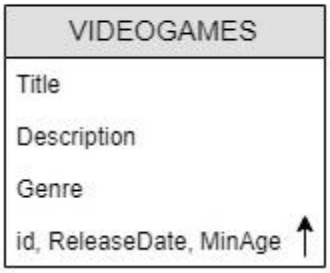


For the second one, we created a table with the videogames titles, the user (Facebook account) that bought it and the buy date. To achive uniqueness, we need a partition key, Facebook account, and a clustering column, buy date. This two build the primary key.

2

The last one request the videogames info. We need one partition, id, and two clustering columns, released date and minimun age. Again, this three build the primary key.

3



The data created for the tables can be found in the annexed PDF document.

The next step is to create the tables in the Cassandra Shell.

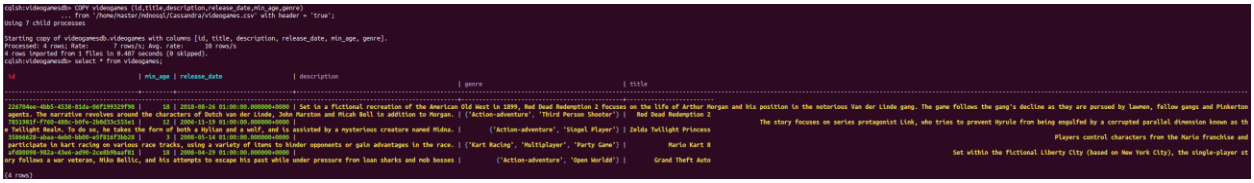
```
(base) master@master-BigData:~$ cqlsh
Connected to Test Cluster at 127.0.0.1:9042.
[cqlsh 5.0.1 | Cassandra 3.11.3.5116 | CQL spec 3.4.4 | Native protocol v4]
Use HELP for help.
cqlsh> CREATE KEYSPACE videogamesdb
   ... WITH replication = {'class': 'SimpleStrategy', 'replication_factor' : 1};
cqlsh> describe keyspaces

musicdb  masterpydb     system_auth  system_distributed  system_traces
userdb   system_schema  system       videogamesdb        killrvideo
```

```
cqlsh> use videogamesdb ;
cqlsh:videogamesdb> CREATE TABLE users (
                ... name VARCHAR,
                ... facebook_account VARCHAR,
                ... nationality VARCHAR,
                ... birthdate TIMESTAMP,
                ... logister_date TIMESTAMP,
                ... PRIMARY KEY (facebook_account)
                ... );
cqlsh:videogamesdb> CREATE TABLE videogames_by_user_date (
                ... facebook_account VARCHAR,
                ... buy_date TIMESTAMP,
                ... title VARCHAR,
                ... PRIMARY KEY ((facebook_account),buy_date)
                ... )WITH CLUSTERING ORDER BY (buy_date DESC);
cqlsh:videogamesdb> CREATE TABLE videogames (
                ... id UUID,
                ... title VARCHAR,
                ... description VARCHAR,
                ... genre SET<VARCHAR>,
                ... min_age INT,
                ... release_date TIMESTAMP,
                ... PRIMARY KEY ((id), min_age, release_date)
                ... )WITH CLUSTERING ORDER BY (min_age ASC,release_date ASC);
cqlsh:videogamesdb> describe tables

videogames  videogames_by_user_date  users
```

And copy the data from a CSV file.

```
cqlsh:videogamesdb> COPY users (facebook_account,name,birthdate,nationality,logister_date)
                ... from '/home/master/mdnosql/Cassandra/usr.csv' with header = 'true';
Using 7 child processes

Starting copy of videogamesdb.users with columns [facebook_account, name, birthdate, nationality, logister_date].
Processed: 2 rows; Rate:       3 rows/s; Avg. rate:       5 rows/s
2 rows imported from 1 files in 0.406 seconds (0 skipped).
cqlsh:videogamesdb> select * from users
                ... ;

 facebook_account | birthdate                       | logister_date                   | name         | nationality
------------------+---------------------------------+---------------------------------+--------------+-------------
    Frank_E_Graham | 1994-08-17 03:12:30.000000+0000 | 2004-08-31 22:25:03.000000+0000 |  frankgraham |     Spaniard
      Crawford_Tew | 2005-01-05 17:48:02.000000+0000 | 2017-09-15 04:02:11.000000+0000 | tewcracwford |      Italian

(2 rows)
```

```
cqlsh:videogamesdb> COPY videogames_by_user_date (title, facebook_account, buy_date)
                ... from '/home/master/mdnosql/Cassandra/videogames_by_user_date.csv' with header = 'true';
Using 7 child processes

Starting copy of videogamesdb.videogames_by_user_date with columns [title, facebook_account, buy_date].
Processed: 4 rows; Rate:       8 rows/s; Avg. rate:      11 rows/s
4 rows imported from 1 files in 0.372 seconds (0 skipped).
cqlsh:videogamesdb> select * from videogames_by_user_date ;

 facebook_account | buy_date                        | title
------------------+---------------------------------+------------------------
    Frank_E_Graham | 2019-02-15 19:12:03.000000+0000 |   Red Dead Redemption 2
    Frank_E_Graham | 2014-08-31 21:25:03.000000+0000 |        Grand Theft Auto
    Frank_E_Graham | 2007-08-31 22:25:03.000000+0000 | Zelda Twilight Princess
      Crawford_Tew | 2015-04-09 07:00:03.000000+0000 |            Mario Kart 8

(4 rows)
```



Finally, the two proposals queries.

```
46 SELECT buy_date, title
47 FROM videogames_by_user_date
48 WHERE facebook_account = 'Frank_E_Graham';
```

Results | Query Trace

buy_date	title
2019-02-15 19:12:03+0000	Red Dead Redemption 2
2014-08-31 21:25:03+0000	Grand Theft Auto
2007-08-31 22:25:03+0000	Zelda Twilight Princess

```
51 SELECT *
52 FROM videogames
53 WHERE min_age >= 18
   ALLOW FILTERING;
```

Results | Query Trace

id	min_age	release_date	description	genre	title
226704ee-4bb5-4538-81da-06f	18	2018-08-26 01:00:00+0000	Set in a fictional recreation of t	{Action-adventure, Third Perso	Red Dead Redemption 2
afd80098-982a-43e6-ad90-2ce8	18	2008-04-29 01:00:00+0000	Set within the fictional Liberty	{Action-adventure, Open World	Grand Theft Auto

María Fernández Hijano

Martín Blázquez Moreno

Juan Rafael Caro Romero