Universitat de les Illes Balears

11756 Storage and Data Recovery

Master's Degree in Intelligent Systems

Exercise 5

A **Reservations.JSON** file is available containing data on completed hotel reservations. For each reservation, the room number, the dates when the reservation was made, check-in and check-out dates, the client's name and nationality, and the number of people included in the reservation are provided.

- You are asked to create a database with the tables **ROOM** (with the room number and a price, which you may generate randomly), **RESERVATION** (with the reservation dates), **CLIENT** (with the client's personal information), and **COUNTRY** (with the country code and full name). In addition to the attributes mentioned above, you must include any appropriate **FOREIGN KEYS**.
- In addition to the tables above, you must create a table **ALL** that contains the attributes included in the JSON file. Then, you should create a program that fills this **ALL** table with the reservation data.
- Using an external program (in any programming language you prefer), read all the records from the **ALL** table and populate the **ROOM**, **CLIENT**, **RESERVATION**, and **COUNTRY** tables. Measure the execution time of the entire process.
- Create new tables equivalent to the original ones — **ROOM2**, **RESERVATION2**, **CLIENT2**, **COUNTRY2** — and, using an internal program (a **PROCEDURE** or **FUNCTION** of the DBMS), read all the records from the **ALL** table and populate these new tables. Measure the execution time of the full process.
- Create a program that splits the original JSON file into 10 smaller JSON files and, using the **MAP REDUCE** technique, calculate the occupancy of each room. Finally, a JSON file must be generated containing the room numbers and the number of days each room has been occupied.
- It has been proposed to avoid JOINs when determining the room occupancy dates. For this reason, it has been decided to include a text attribute in the **ROOM** and **ROOM2** tables that contains a JSON structure with the reservation dates for each room. You are asked to fulfill this requirement using an external program (for table **ROOM**) and a DBMS **PROCEDURE** or **FUNCTION** (for **ROOM2**). Measure the execution time in each case.
- Include a **DISCUSSION** section evaluating the previous procedures, the difficulties of each process, the execution times, and any other aspects you consider relevant.

The final report consists in a PPT slides presentation (about 10-15 slides), the programs and the exportation of the generated database.