VE482 — Introduction to Operating Systems

Lab 4

Manuel — UM-JI (Fall 2019)

Goals of the lab

- Learn basics on database
- Import a databse
- Run simple databse queries

1 System preparation

This is the evening, you are exhausted after a long day of work, but you want to ensure you fully master database systems before you go any further in your LemonDB project. This Mr. Frown is quite scary, you do not want to get into any trouble with him.

1.1 Database creation

As a first step you need to find a database, you start your web-browser and try to connect to a proper search engine. Unfortunately you VPN takes too long to load so you decide to come up with another idea. After a bit of thinking you realise that you still have a git version of the Linux kernel, and as you know everything about git you can easily generate logs from the git commits. To ensure a proper formatting you refer to git pretty format documentation page.

So you open a terminal running `mumsh` and type a simple command line to test database generation.

```
mumsh $ git log --pretty="%H,%aN,%aI,%s" > db.csv
```

The goal being to fully master the type of queries needed by LemonDB, you do not need a very complicated database and in the end only generate two csv files[1] containing the following fields.

Fields for `timestamp.csv`:

- Hash of the commit
- Author name
- Author date, strict ISO 8601 format
- Author date, UNIX timestamp

Fields for `db.csv`:

- Hash of the commit
- Author name
- Subject

1.2 Database system installation

Your memories on database systems being a bit rusty you decide to refresh them. Luckily your VPN is now working,so you can use a proper search engine and ensure the correctness of the information found.

- What are the most common database systems?
- Briefly list the pros and cons of the three most common ones.

After completing your reading you decide to install `SQLite` on your Linux system. The next step is now to import your git database into two tables.

[1] The `db.csv` and `timestamp.csv` can be found on the server in the directory `ve482/l4`.

- Create an empty SQLite database.
- Use the SQLite shell to prepare two empty tables for each of your .csv file.
- Import each .csv file in its corresponding SQLite table.

2 Database queries

At this stage it is important to run basic queries to verify that the database has been imported correctly. This is also a great opportunity to check whether you fully understand what Mr. Frown wants you do implement. Therefore you spend the rest of the evening playing around the database and running queries.

- Who are the top five contributors to the Linux kernel since the beginning?
- Who are the top five contributors to the Linux kernel for each year over the past five years?
- What is the most common “commit subject”?
- On which day is the number of commits the highest?
- Determine the average time between two commits for the five main contributor.