

Primer parcial: Fase #1

José Fernando Pérez, 20000756

November 2020

1 Introducción

En este documento se especifican los requerimientos e implementaciones técnicas utilizadas para la realización del primer examen parcial.

2 Arquitectura de la solución

Se utilizó la tecnología de contenerización de aplicaciones utilizando *docker* para montar las imágenes y *docker-compose* para crear la infraestructura de red, automatizar la carga a la base de datos y manejar las imágenes de *rstudio* y base de datos *MySQL*.

2.1 Docker Compose

```
version: "3.8"
services:
  db:
    image: mysql:5.7
    ports:
      - 3306:3306
    command: --secure-file-priv=/var/lib/mysql-files
    environment:
      - "MYSQL_ROOT_PASSWORD=root123"
      - "MYSQL_DATABASE=parcial1"
      - "MYSQL_USER=test"
      - "MYSQL_PASSWORD=test123"
    volumes:
      - "./sql_scripts:/docker-entrypoint-initdb.d"
      - "./data:/var/lib/mysql-files"
    networks:
      parcial1_net:
        ipv4_address: 172.28.1.1
  rstudio:
```

```

    image: rocker/tidyverse
    ports:
      - 8787:8787
      - 3838:3838
    environment:
      - "DISABLE_AUTH=true"
      - "ADD=shiny"
    volumes:
      - "./app:/home/rstudio/app"
    networks:
      parcial1_net:
        ipv4_address: 172.28.1.2
networks:
  parcial1_net:
    ipam:
      driver: default
      config:
        - subnet: 172.28.0.0/16

```

Para la base de datos se utilizó la imagen de *MySQL* versión 5.7 (habían problemas de carga con versiones posteriores) y la imagen *tidyverse* de *rstudio*. Para cada imagen se montó un volumen en una carpeta específica para indicar el código y los scripts de *sql* utilizados.

2.2 Red

Se utilizó la red privada "172.28.1.0/24" y se asignaron las siguientes IPs a las imágenes:

1. Base de datos(MySQL): 172.28.1.1
2. RStudio: 172.28.1.2

2.3 Automatización de carga

En la documentación de la imagen docker de *MySQL* se especifica que se ejecutan los scripts *.sh* y *.sql* de la carpeta "/docker-entrypoint-initdb.d", por dicha razón se montó un volumen en dicha carpeta que hace referencia a los scripts de *sql* que crean los esquemas de la base de datos y cargan un archivo separado por comas en la tabla creada. Por ejemplo, para el conjunto de datos llamado "videos_stats" se generó el siguiente script:

```

USE parcial1;

CREATE TABLE videos_stats (
  id_video VARCHAR(30) NOT NULL,
  viewCount INT,

```

```
        likeCount INT,  
        dislikeCount INT,  
        favoriteCount INT,  
        commentCount INT,  
        PRIMARY KEY (id_video)  
    );  
  
LOAD DATA INFILE '/var/lib/mysql-files/academática_video_stats.csv'  
INTO TABLE videos_stats  
FIELDS TERMINATED BY ','  
IGNORE 1 ROWS;
```

3 Descripción de la solución

Se trabajo en 2 pestañas principales, visualización de los datos en crudo y manipulación de un conjunto de datos a partir de sus columnas.