

Puesta en Marcha

Este documento detalla los pasos para crear las bases de datos, migrar la información y correr las pruebas con los programas creados para automatizar dichas tareas. El proyecto fue construido sobre Ubuntu y es necesario correrlas bajo un sistema operativo basado en Debian o bien realizar las acciones manualmente.

IMPORTANTE

Todos los scripts se hicieron en base al bash de Linux. Si por algún motivo el docente no dispone de Ubuntu, los scripts no funcionarán y deberá importar manualmente los archivos en PostgreSQL y en MongoDB -ver último paso *carga manual*- y ejecutar todas las consultas manualmente. **Recomiendo fuertemente correr el proyecto en Ubuntu ya que se construyó y testeó en este entorno.**

Instalación de Software

Para ejecutar correctamente las consultas es necesario instalar los siguientes programas:

- *MongoDB (versión 3.4.4+)*: <https://www.mongodb.com/download-center#community>
- *PostgreSQL (versión 9.5.8+)*: <https://www.postgresql.org/download/>
- *Java (versión 1.8+)*: <https://www.java.com/es/download/>

Además, se recomienda instalar los siguientes para correr las pruebas:

- *PgAdmin3*: <https://www.pgadmin.org/download/>
- *RoboMongo*: <https://robomongo.org/download>

Preparar PostgreSQL

Paso 1: Posicionarse en la carpeta.

Una vez descargado el .zip, descomprimirlo y dirigirse a dicha carpeta utilizando el comando `cd` o bien abriendo la carpeta con la interfaz gráfica, dar click secundario y clickear en “abrir en terminal”.

Paso 2: Crear la base de datos, el esquema e índices en postgresQL.

Primero debe crearse una base de datos en PostgreSQL llamada “*scada*” y un esquema público. Puede realizarse mediante *pgadmin3* o bien manualmente. Es necesario que la base de datos se llama *scada* ya que así será referenciada a lo largo del proyecto. A continuación, seguir los siguientes pasos:

- Dirigirse a la carpeta del proyecto contenedora de datos de PostgreSQL: tipear en la consola `cd postgresql`.
- Cambiar la contraseña de la base de datos en el script `prepare.sh`. Para ello, tipear `sudo gedit prepare.sh`. Se abrirá un editor de texto y se debe modificar la siguiente línea **export PGPASSWORD="PONER ACÁ LA CONTRASEÑA"**. Una vez puesta la contraseña se debe guardar el script.
- Dar permiso de ejecución al script `prepare.sh` tipeando en consola `sudo chmod +x prepare.sh`. En este paso el sistema solicitará la contraseña del usuario root.
- Correr el script tipeando `./prepare.sh`.

Paso 3: Insertar los datos en las tablas de PostgreSQL.

- Dirigirse a la carpeta `data` tipeando `cd data`.
- Cambiar la contraseña de la base de datos en el script `load_data.sh`. Para ello, tipear `sudo gedit load_data.sh`. Se abrirá un editor de texto y se debe modificar la siguiente línea **export PGPASSWORD="PONER ACÁ LA CONTRASEÑA"**. Una vez puesta la contraseña se debe guardar el script.
- Dar permisos al script `load_data.sh` tipeando `sudo chmod +x load_data.sh`.
- Correr el script tipeando `./load_data.sh` y esperar a que finalice la ejecución (~15 minutos). Luego, los datos estarán insertados en las tablas del esquema. El script arroja el tiempo que tardó la inserción de datos en el fichero `time.txt` dentro de la carpeta `postgresql/data`.

Preparar MongoDB

Las pruebas sobre MongoDB se realizaron sin contraseña ni usuario particular; se utilizó la configuración por defecto.

Paso 1: Iniciar servidor.

Iniciar el servidor de MongoDB en una nueva terminal tipeando `mongod`. Se la debe dejar corriendo durante el resto de la instalación.

Paso 1: Posicionarse en la carpeta.

Ir a la carpeta raíz del proyecto utilizando el comando `cd` o bien abriendo la carpeta con la interfaz gráfica, dar click secundario y clickear en "abrir en terminal".

Paso 2: Crear la base de datos, el esquema e índices en postgresQL.

- Dirigirse a la carpeta del proyecto contenedora de datos de MongoDB: tipear en la consola `cd mongodb`.
- Correr el script de creación de índices tipeando `mongo < indexes/indexes_create.js`.

Migrar Datos de PostgreSQL a MongoDB

Ir a la carpeta raíz del proyecto utilizando el comando `cd` o bien abriendo la carpeta con la interfaz gráfica, dar click secundario y clickear en “abrir en terminal”.

- a. Típear en la consola “`sudo java -jar "java/target/migrator-1.0-SNAPSHOT.jar"`”. Esperar a que finalice (~30min).

Exportar/Importar Colecciones en MongoDB

La migración de PostgreSQL a MongoDB fue solo para obtener esa misma información replicada en las distintas bases de datos. De todas formas, es necesario guardar la información en formato *json* realizando un *dump* y luego *importarlo* para tomar métricas de inserción.

Paso 1: Realizar Dump de las Colecciones (Exportar).

- a. Dirigirse a la carpeta del proyecto contenedora de datos de MongoDB: típear en la consola “`cd mongodb`”.
- b. Dar permisos al script `save_data.sh` típeando “`sudo chmod +x save_data.sh`”.
- c. Correr el script típeando “`./save_data.sh`” y esperar a que finalice la ejecución. Luego, los documentos de cada colección estarán exportados y guardados en formato *json* en la carpeta `mongodb/data`.

Paso 1: Realizar Carga de las Colecciones (Importar).

- a. Dar permisos al script `load_data.sh` típeando “`sudo chmod +x load_data.sh`”.
- b. Correr el script típeando “`./load_data.sh`” y esperar a que finalice la ejecución. Este script arroja el tiempo que tardó la ejecución de la carga de datos en el fichero `time.txt` dentro de la carpeta `mongodb/data`.

Realizar Pruebas en Consultas de Selección

Una vez que ambas bases de datos están preparadas es necesario correr las pruebas de cada consulta. Para ello, se crearon carpetas para cada grupo de consultas (con agrupadores y funciones de agregación, sin repeticiones y con ensambles). Dependiendo del motor, pararse en la carpeta raíz y dirigirse hacia la carpeta `/postgres` o `/mongodb` y luego a la carpeta `selects`.

- Si se quiere ejecutar las pruebas de PostgreSQL: posicionarse en la carpeta raíz del proyecto y típear `cd postgres/selects`.

- Si se quiere ejecutar las pruebas de MongoDB: posicionarse en la carpeta raíz del proyecto y tipear `cd mongodb/selects`.

Aclaraciones

Dentro de la carpeta `/selects` (tanto de mongodb como de postgresql) existen tres carpetas: `/1`, `/2` y `/3`, las cuales se corresponden con las consultas con agrupadores y funciones de agregación, consultas sin repeticiones y consultas con ensambles, respectivamente.

Para correr las pruebas se debe dirigir a la carpeta con las consultas deseadas (`/1`, `/2` o `/3`), una vez allí se deben realizar los siguientes pasos:

- a. Dar permisos al script `run.sh` tipeando `"sudo chmod +x run.sh"`.
- b. Correr el script tipeando `"./run.sh"` y esperar a que finalice la ejecución. Este script ejecuta 30 veces cada consulta y arroja el tiempo que tardó cada una de las ejecuciones por consulta y el promedio por cada consulta en un fichero llamado `time.txt` dentro de la misma carpeta.

Este script existe en cada una de las carpetas de consulta y para ambos motores.

Consultas de Modificación/Eliminación

Estas consultas se encuentran en las carpetas `/updates` o `/deletes` dentro de la carpeta de PostgreSQL o MongoDB.

- Si se quieren ejecutar las consultas de eliminación de PostgreSQL: posicionarse en la carpeta raíz del proyecto y tipear `cd postgres/deletes`.
- Si se quieren ejecutar las consultas de modificación de PostgreSQL: posicionarse en la carpeta raíz del proyecto y tipear `cd postgres/updates`.
- Si se quieren ejecutar las consultas de eliminación de MongoDB: posicionarse en la carpeta raíz del proyecto y tipear `cd mongodb/deletes`.
- Si se quieren ejecutar las consultas de modificación de MongoDB: posicionarse en la carpeta raíz del proyecto y tipear `cd mongodb/updates`.

Allí se encuentran las consultas en código SQL/JS depende del motor. Se deben copiar las consultas y ejecutarlas manualmente, ya que no hay script para correrlas porque modifican el estado de la base de datos.

Carga Manual

Ignorar esta sección si se está en Linux y **ver sólo si se opera desde Windows**. Se debe ingresar a la carpeta raíz y luego dirigirse a `/cargamanual`. Dentro de esta carpeta se encuentra el dump de ambas bases de datos. Importar a PostgreSQL la información de la carpeta `/cargamanual/postgresql` e importar a Mongo la información de `/cargamanual/mongodb`. **Recomiendo fuertemente correr el proyecto en Ubuntu ya que se construyó y testeó en este entorno.**