

Proyecto La Palma

Descripción

Creación de una infraestructura que permita almacenar, monitorizar y visualizar el estado de los embalses situados en la palma a través de unos sensores que miden diferentes parámetros.

Estructura del proyecto

El repositorio del proyecto está formado por los siguientes directorios y archivos:

- **bi:** cuadro de mandos en PowerBI para visualizar los agregados diarios
- **data:** Resultados del proceso ETL que da lugar al Data Warehouse. La carga de archivos se realiza de tanto en ficheros .csv como en una base de datos MySQL, de la cual hay un dump.
- **deployment:** docker compose necesario para levantar la infraestructura del proyecto
- **documentation:** archivos de documentación
- **grafana:** exportación de dashboard y alarmas de Grafana
- **instructions:** instrucciones, tareas y rúbrica del proyecto
- **resources:** archivos de prueba para hacer peticiones simples. Tanto en versión Python como en versión .http (requiere un cliente http)
- **src:** directorio con los archivos fuente del proyecto
 - **emulator:** emulador de una boya que reporta datos de forma periódica. El comportamiento del script está sujeto a la parametrización definida en el fichero config.py
 - **etl:** scripts para realizar el proceso de ETL de los datos concurrentes y convertirlos en datos agregados por días. El proceso carga los datos tanto en una base de datos como ficheros .csv (carpeta data). Existen 2 ficheros:
 - **aggregation.py:** script principal que hace la agregación por días
 - **buoys.py:** obtiene un listado de las diferentes boyas y los carga en una tabla aparte.
 - **utils:** contiene un script 'Setup.py' que facilita la puesta en marcha del proyecto (crea las entidades, la suscripción, etc...)
- **.env:** archivo con las variables de entorno. Útil por si no se quieren definir directamente en el sistema.
- **README.md:** documento de presentación del proyecto donde se pueden encontrar comandos útiles
- **requirements.txt:** librerías necesarias para ejecutar los scripts de Python.

Instalación y configuración

Para poner en marcha en proyecto hay que seguir los siguientes pasos:

1. Desplegar la infraestructura de Docker, utilizando el comando ``docker-compose -f docker-compose_SDD.yml up -d``
2. Dar de alta las variables de entorno definidas en el fichero README.md. (También se puede hacer uso del fichero .env)
3. Instalar las librerías necesarias con el comando ``docker-compose -f docker-compose_SDD.yml up -d``
4. Ejecutar el fichero Setup.py, el cual creará las entidades, suscripción y las tablas del Data Warehouse de los valores agregados.

Emulador

Para crear datos de manera ficticia, se puede hacer uso del script BuoyEmulator.py. Este script creará valores pseudo-aleatorios de cada medida para cada boya con un intervalo de tiempo determinado. El rango de los valores aleatorios, la frecuencia de los mensajes y el período que estos abarcan se pueden **configurar en el fichero config.py**

ETL

Para agregar los datos diariamente y cargarlos en el Data Warehouse se debe ejecutar el script aggregation.py. Este script hace todo el proceso de transformación y carga los datos en el Data Warehouse especificado en las variables de entorno.

Si también es necesario crear (o refrescar) las diferentes entidades existentes se puede ejecutar el script buoys.py. Este script creará una tabla con las diferentes boyas que están reportando datos.