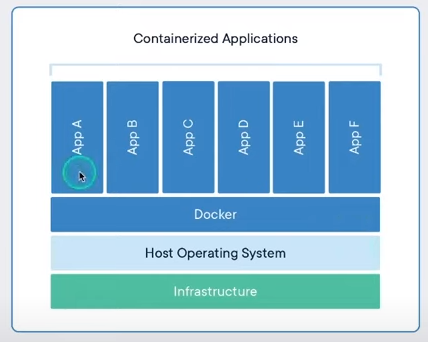
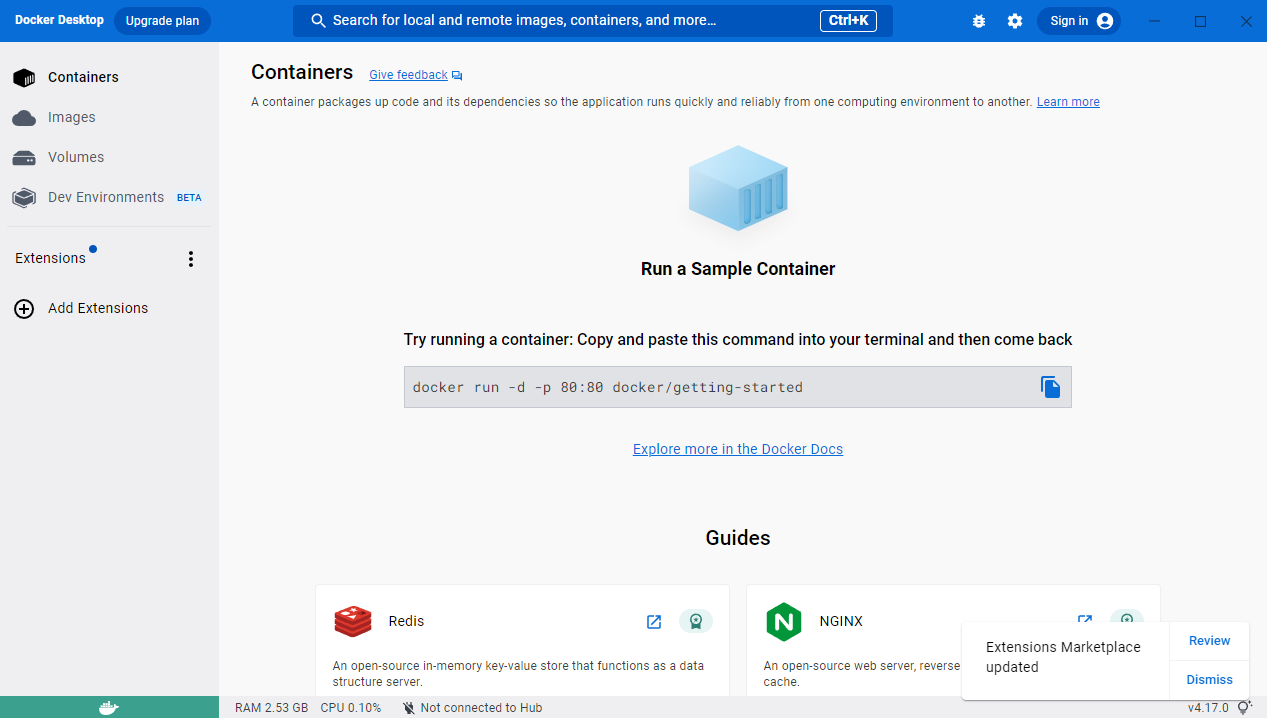
**MANUAL INSTALACIÓN AIRBYTE Y DBT**

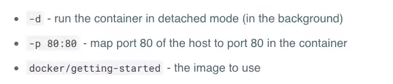
**Docker Desktop**

**REQUISITOS**: Vamos a realizar todo el proceso en Windows, por lo tanto hemos descargado Docker Desktop.

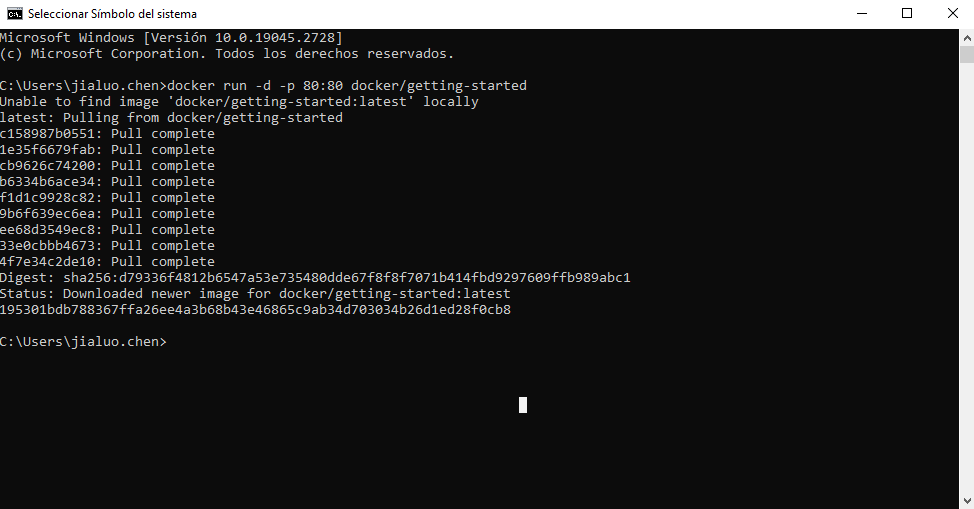
Docker tiene la siguiente arquitectura:

Una vez descargado e instalado Docker, deberíamos tener algo así:

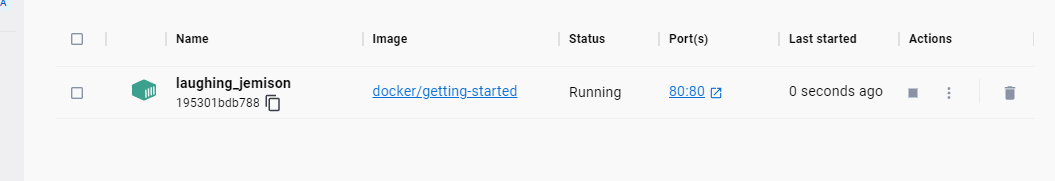


Copiamos “docker run -d -p 80:80 docker/getting-started” y lo ejecutamos en un CMD.

Y quedará algo así:



Volvemos a Docker, y observaremos que se ha desplegado un contenedor:



**Configurar Airbyte vía Docker**

Accedemos a la pagina web de Airbyte para instalar la versión open source.

Enlace: <https://airbyte.com/airbyte-open-source>

En el CMD:

1. “git clone <https://github.com/airbytehq/airbyte.git>”
2. “cd airbyte”
3. “docker-compose up”

Una vez realizado los pasos, accedemos a través de un navegador:

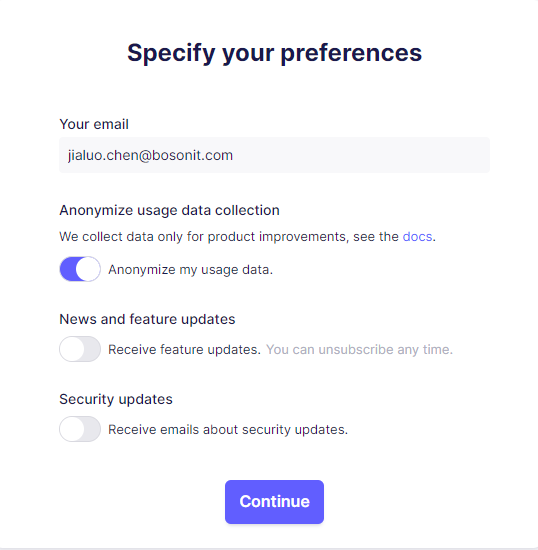
“http://:localhost:8000”

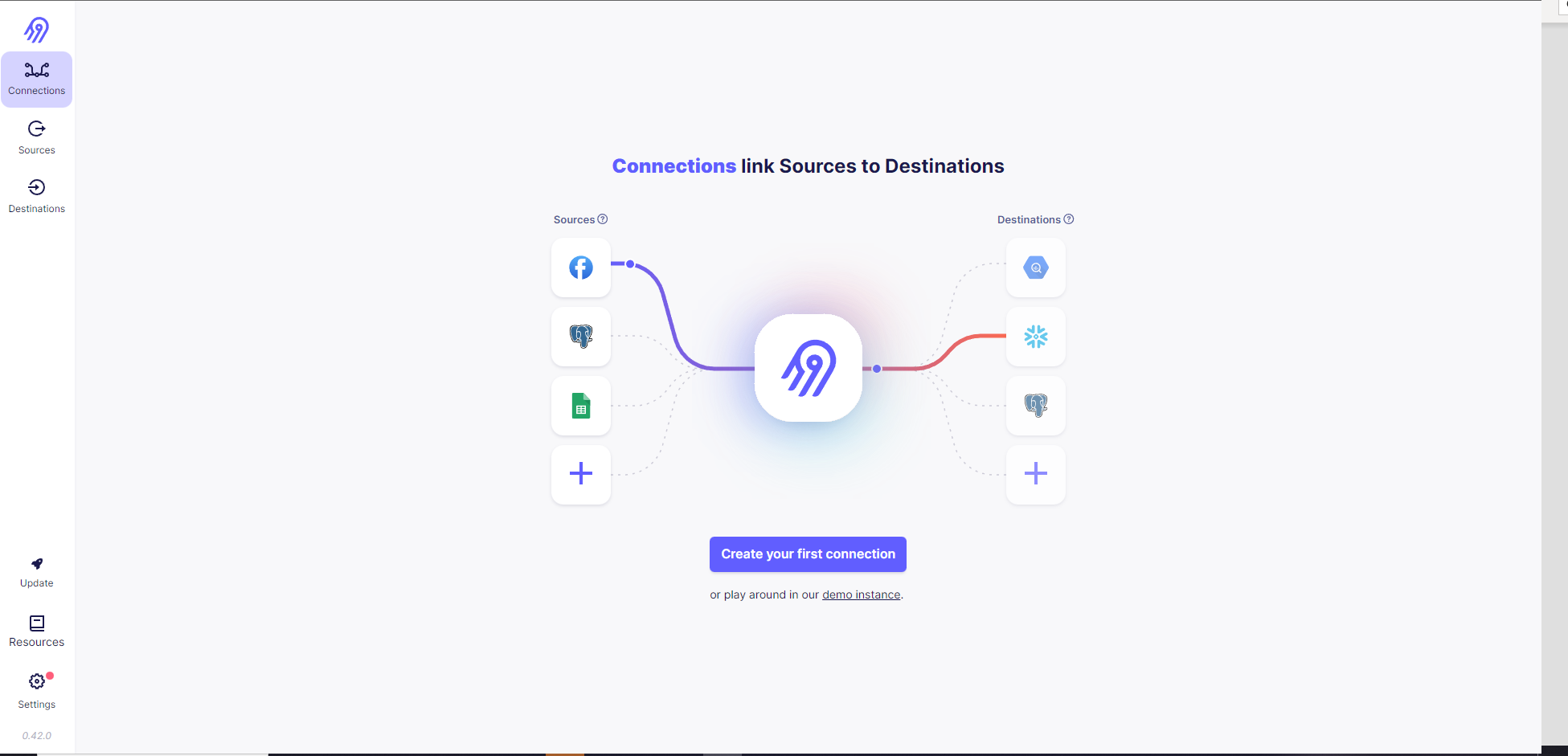
Pedirá un usuario y contraseña:

* Usuario: airbyte
* Contraseña: password

Para cambiar la contraseña, se accedería al fichero .env

Después de acceder con el usuario, configurar las preferencias:



Una vez configurado todas las preferencias nos mostrara una pantalla similar así:

**Cargar distintos tipos de datos como fuente y usar Postgre como destino a través de Airbyte**

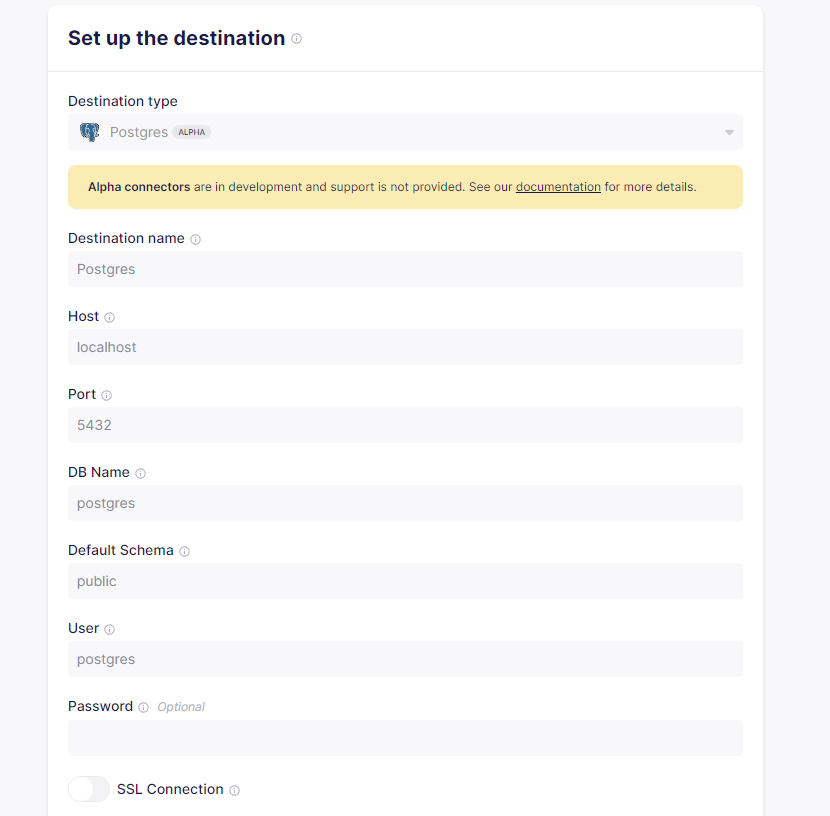
En nuestro caso configurare Postgre como destino, ya que como fuente en este caso configurare un CSV.

A través de CMD: **“docker run --name airbyte\_dst -e POSTGRES\_PASSWORD=admin -p 3000:5432 -d debezium/postgres:latest”**

(En mi caso el puerto 3000, pero se podría usar cualquier otro que este libre)

Una vez realizado, tener en cuenta que hay datos por defecto como nombre de DB y user.

El siguiente paso seria ir a Airbyte y elegir Postgre como destino y rellenar la configuración de la siguiente forma si habéis seguido los mismos ajustes: (el password seria **admin**)



En caso de que quieras acceder a psql para lanzar queries, a través de CMD: “docker exec -it airbyte-source psql --username=postgres”

Una vez configurado el destino, procedemos a configurar la fuente, que en este caso usaremos CSV, JSON.

Buscamos el directorio donde hemos instalado Airbyte, en mi caso **C:\Users\jialuo.chen\airbyte,** buscamos el fichero .env y cambiamos la configuración de:

LOCAL\_ROOT=C:\Users\jialuo.chen\OneDrive - Bosonit\Escritorio\airbyte\_dbt\db\_src

LOCAL\_DOCKER\_MOUNT=C:\Users\jialuo.chen\OneDrive - Bosonit\Escritorio\airbyte\_dbt\db\_src

HACK\_LOCAL\_ROOT\_PARENT=C:\Users\jialuo.chen\OneDrive - Bosonit\Escritorio\airbyte\_dbt

 Cargar un CSV en Postgre usando Airbyte

 Cargar un JSON en Postgre usando Airbyte

 Cargar datos en Kafka desde Postgre -> <https://airbyte.com/tutorials/postgres-to-kafka>

 Cargar datos desde MySQL a kafka y desde allí a Pinot -> <https://airbyte.com/tutorials/real-time-data-analytics-pipeline>

 Sync desde Postgre aplicando CD (debezium) -> <https://airbyte.com/tutorials/postgresql-change-data-capture-elt>

 <https://airbyte.com/tutorials/incremental-change-data-capture-cdc-replication>

 <https://airbyte.com/tutorials/incremental-data-synchronization>

 Leer datos en Postgre que contenga datos en JSON, extraer datos JSON y utilizarlos como nuevas columnas y volcar de nuevo a otra tabla Postgre utilizando dbt para la transformación