Docker Instruction

Lulu Zhu 17 Fall BIA

This instruction aims at helping you set up the Docker environment and share documents between local host and Docker. If you successfully set up Docker during the Hadoop/Spark Bootcamp, jump to step 3 directly.

1. **Download and Install Docker**

    1.1   https://runnable.com/docker/install-docker-on-macos

    1.2   https://runnable.com/docker/install-docker-on-windows-10

2. **Pull Jupyter Notebook Image**

    2.1   **Command Line:** *sudo docker pull jupyter/all-spark-notebook*

    2.2   Check out all images and containers:

    Right now, we have one image – jupyter/all-spark-notebook- and no container.

```
Last login: Sat Apr 13 12:40:32 on ttys000
[➜  ~ docker image ls                                                          ]
REPOSITORY                    TAG              IMAGE ID           CREATED            SIZE
jupyter/all-spark-notebook    latest           43618b207424       7 weeks ago        6.39GB
[➜  ~ docker container ls                                                       ]
CONTAINER ID       IMAGE              COMMAND              CREATED            STATUS
[    PORTS                NAMES                                                  ]
```

3. **Share documents between host and Docker**

    3.1   Create a new folder locally which is the host_directory

    Command Line: *mkdir path_of_ host_director*

```
[➜  ~ ls /Users/lulu/Desktop/678_19_Spring/case_study
[➜  ~ mkdir /Users/lulu/Desktop/678_19_Spring/case_study/container-data
[➜  ~ ls /Users/lulu/Desktop/678_19_Spring/case_study
container-data
```

    3.2   Connect local folder to Docker when start a new container

    Command Line: *docker run -p 8888:8888 -v*

    host_directory:container_directory jupyter/all-spark-notebook

```
→ ~ docker run -p 8888:8888 -v /Users/lulu/Desktop/678_19_Spring/case_study/container-data:/home/jovyan jupyter/all-spark-notebook
Executing the command: jupyter notebook
[I 22:09:15.696 NotebookApp] Writing notebook server cookie secret to /home/jovyan/.local/share/jupyter/runtime/notebook_cookie_secret
[I 22:09:15.924 NotebookApp] JupyterLab extension loaded from /opt/conda/lib/python3.6/site-packages/jupyterlab
[I 22:09:15.925 NotebookApp] JupyterLab application directory is /opt/conda/share/jupyter/lab
[I 22:09:15.927 NotebookApp] Serving notebooks from local directory: /home/jovyan
[I 22:09:15.927 NotebookApp] The Jupyter Notebook is running at:
[I 22:09:15.927 NotebookApp] http://(45621cc5949b or 127.0.0.1):8888/?token=9b8b34e564929bc3a17fd94551c855547d0548021d0f61fd
[I 22:09:15.928 NotebookApp] Use Control-C to stop this server and shut down all kernels (twice to skip confirmation).
[C 22:09:15.933 NotebookApp]

    Copy/paste this URL into your browser when you connect for the first time,
    to login with a token:
        http://(45621cc5949b or 127.0.0.1):8888/?token=9b8b34e564929bc3a17fd94551c855547d0548021d0f61fd
```

Note: (1) container_directory will always be /home/jovyan

(2) Make sure no space in your path or file name

(3) Make sure no Jupyter Notebook running locally before starting Notebook on Docker.

3.3   Copy/paste the URL shown in your terminal into browser to open Jupyter notebook on Docker.

3.4   Move the documents or data to the host directory, then you should be able to see the documents from Docker page. If you create a new Notebook directly on Docker, this Notebook can also be seen in the host directory.

Reference:

Learning Spark.ppt Hao Han, JIMMY ZHENG