INSA INSTITUT NATIONAL DES SCIENCES APPLIQUÉES TOULOUSE

Docker

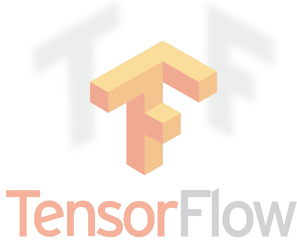
IA Frameworks

Brendan Guillouet - 30/11/2020

Institut National des Sciences Appliquées de Toulouse

Google cloud platform

ML Python Libraries



Python Environment

Viz' Python Libraries

Framework & Tool

Table of contents

Introduction

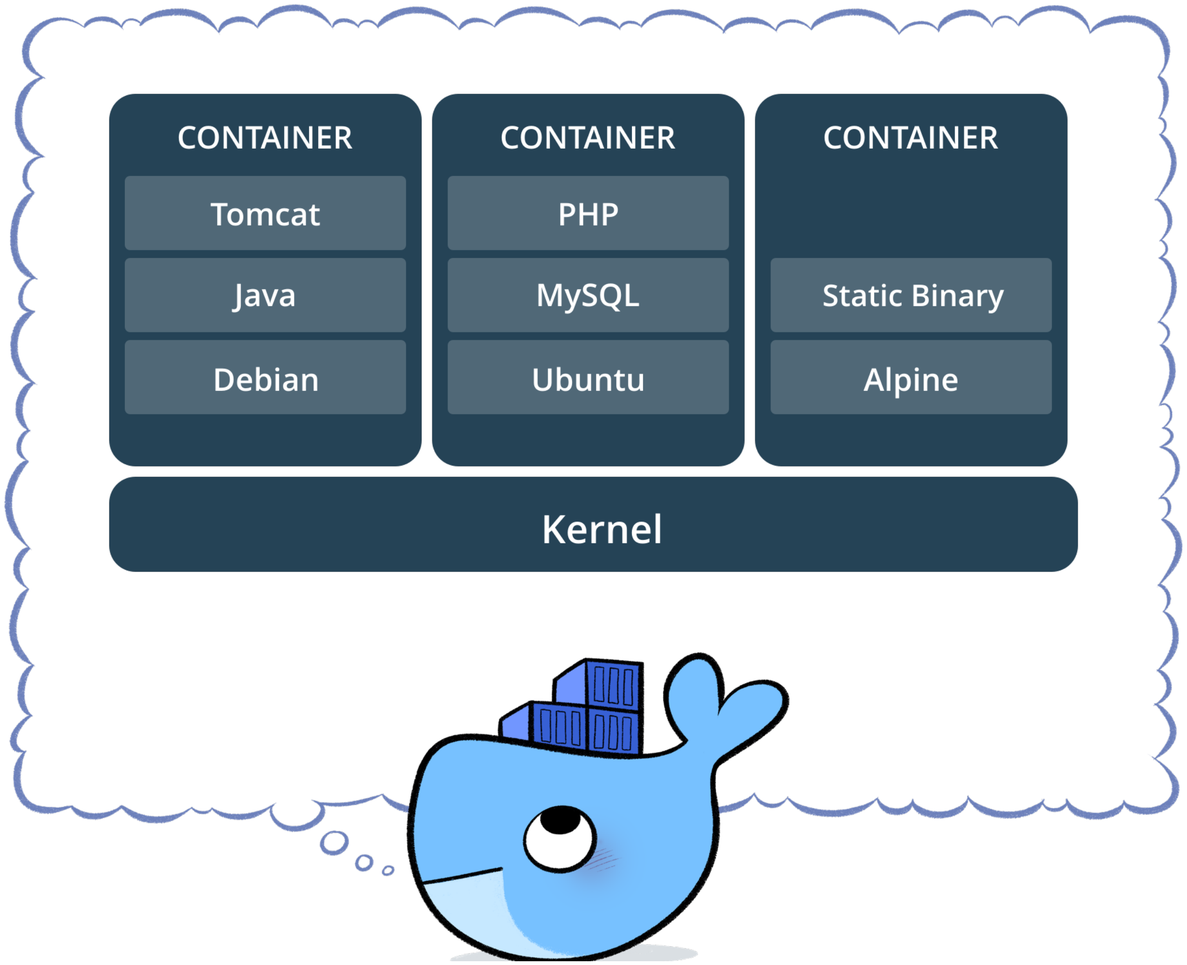
What is docker ?

As for python environment, docker enables to create environment isolated from each other called container.

This environment are built at an OS-level, which means you make run a windows program on unix and *vice versa*.

It enables to easily recreate environment for application whatever the dependencies are.

Facilitate development and test.



Why should I use docker as a data scientist ?

- **Sharing your work** within a docker container ensure that anyone can recreate it whatever the devices he is using. More people can have access to your work.

- **And use the work of others.** Docker is widely used today to share their works. Being comfortable with it ensure you easy access to it.

- **Ease use of computational power.** Developing within a docker container ensure you to easily move your code to devices with more computational power.

- **Ease use of deployment.** Developing within a docker container ensure you to easily move your model to be use in production environment.

Docker Terminology

- **Image**: It's like a turned-off VM which contains the tools you want. Ex: Ubuntu + Tensorflow with Nvidia drives and a Jupyter server.

- **Container**: is an instantiation of an image. You can have multiple copies of the same image running.

- **Dockerfile**: Recipe for creating an image

- **DockerHub / Image registry**: Place where Organisation or individual can post public or private docker images fo ease collaboration and sharing.
  https://hub.docker.com

Application

LAB - OBJECTIVES

- Write a Dockerfile based on the official *Tensorflow Dockerfile* available on *DockerHub.*
- Use the Dockerfile to build and Docker image.
- Launch containers with different options from the built image.
- Run training and prediction on cats vs dogs (with code developed on previous lab) on a google cloud instance.

DockerFile

- The Dockerfile we use on this lab is based on the official Tensorflow gpu image.
- Below is the dockerfile we'll use.

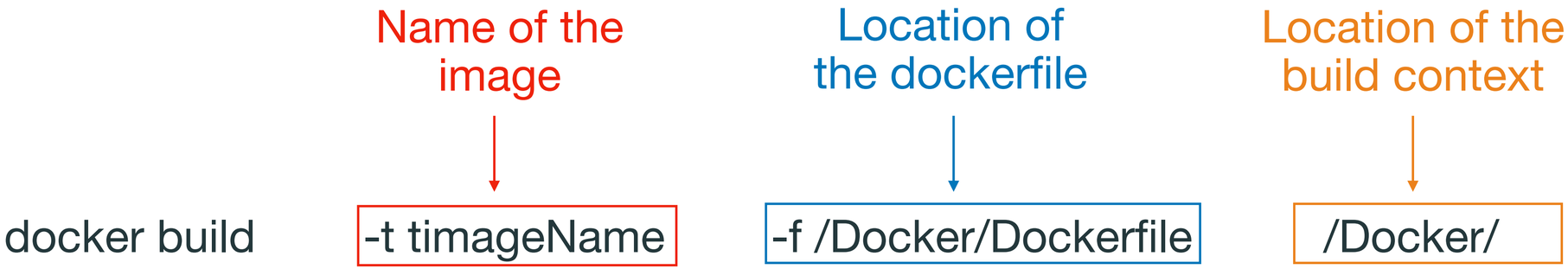
```
FROM tensorflow/tensorflow:2.1.0-gpu-py3
ENV TZ=Europe/Paris
RUN ln -snf /usr/share/zoneinfo/$TZ /etc/localtime && echo $TZ > /etc/timezone
RUN apt-get update && apt-get install -y python-opencv python-tk unzip pv
RUN pip install h5py scikit-image tqdm
```

- The **FROM** instruction sets the based image. A valid docker file starts with a from instruction.
- The **ENV** instruction sets environment variable (PYTHONPATH, etc.)
- The **RUN** instruction will execute any commands.

Official documentation of **dockerfile.**

Build image

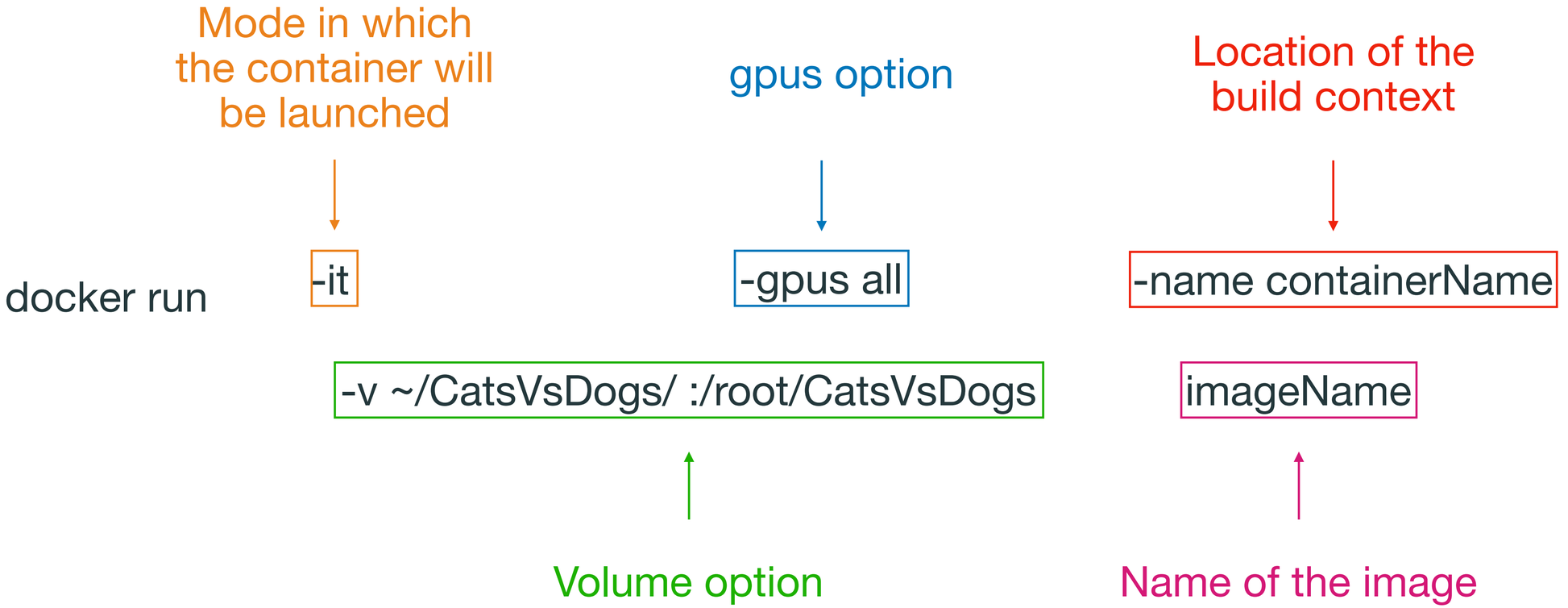
Run the build command in in order to build a docker image



The build context is the location of the folder to which the ADD statement will reference.
All the files required by the dockerfile to build the image are located here.

Run container

Run the run command in order to run your container.



Mode of the container could be:
-it : interactive mode
-dt : detached mode

Mounted Volume

The -v option allows you to use some data you have in your machine within a container.

Other utils command

- *sudo docker images ls -a:* list all existing images.
- *sudo docker containers ls -a:* list all existing containers.
- *sudo docker start/stop/rm container_name:* start, stop or remove a container.
- *sudo docker exec container_name `command`:* Execute the `command` within a container.