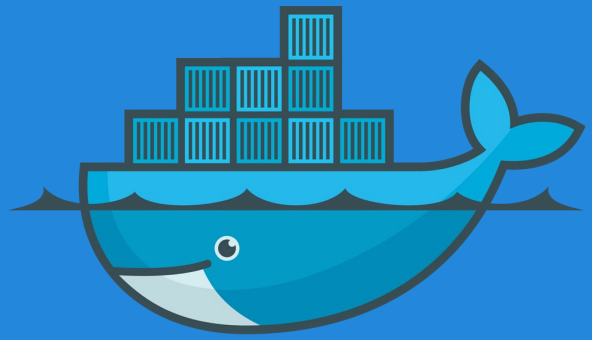


docker容器技術課程

資料科學應用

Philipz
鄭淳尹

Philipz (鄭淳尹)

Docker.Taipei 共同發起人

2014 COSCUP/iThome Summit 講者

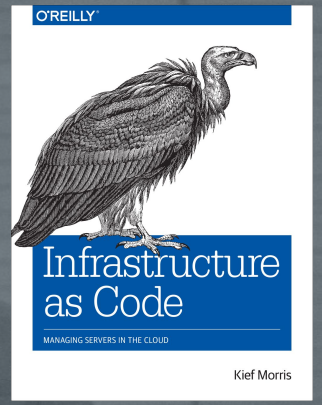
2015 Microsoft Azure 開發者大會 講者

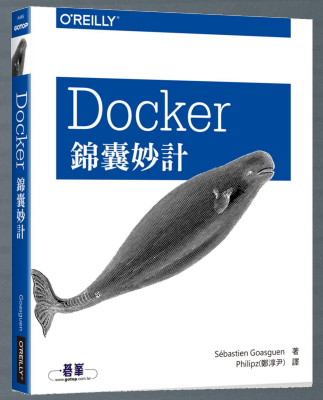
2016 COSCUP Docker 進階工作坊

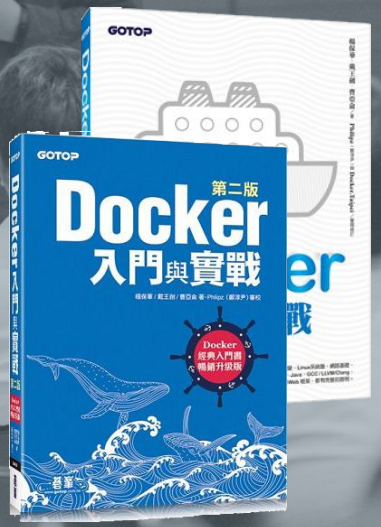
2016 元智大學資工系 Docker 專題演講

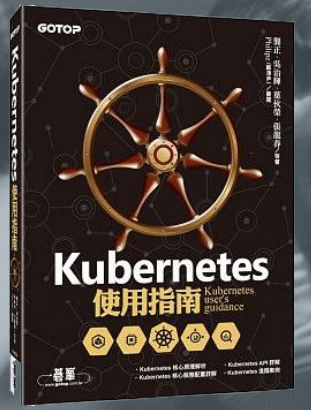
2016 義守大學資工系 Docker 研習營

2017 逢甲大學資工系 Docker 研習班









TradingBot 演算法

- 影像辨識訊號處理
- Machine Learning
  - SVM
- 自動交易系統
  - Java
  - 群益API

基於模式辨識及支援向量機之期貨當沖交易策略和系統

A Strong Futures Day-Trading Strategy and System Using Pattern Recognition and Support Vector Machine

鄭淳尹*
Philip Zheng

CERTIFICATE
OF APPRECIATION

鄭淳尹 老師

於西元 2017 年 3 月 19 日擔任 2017HackNTU
數據分析黑客松 Tech Talk 講師，內容精湛，嘉惠
特頒此狀 以茲感謝

Hackntu 數據分析 黑客松

2017HackNTU
臺大黑客松 總召

元智大學
資訊工程學系
Yuan Ze University
Department of Computer Science and Engineering

135 Yuan-Tung Road, Jung-Li,
Taoyuan, Taiwan, 32003, R.O.C.
TEL: +886-3-4638800 ext. 2372
FAX: +886-3-4638850
http://www.cse.yzu.edu.tw/

邀請函

中華民國台灣省桃園市 320 中壢區內壢遠東路 135 號
135 Yuan-Tung Road, Chungli, Taoyuan, Taiwan, 320, R.O.C.
03-463-8800*2360 or 2372(TEL) 03-463-8850 (FAX) www.cse.yzu.edu.tw

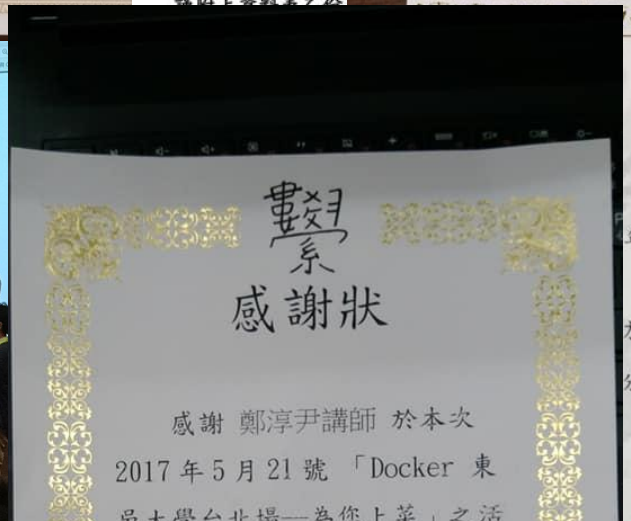
鄭淳尹先生 惠鑑:

素仰 先生學術淵博，特邀先生蒞臨本校作專題演講，承蒙慨允，謹申謝悃。

專題講座時間訂為
1401B 室，敬請 屆
金鳳教授研究室休

東吳大學 數學系
感 謝 狀

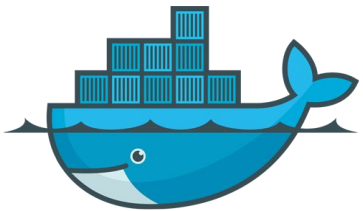
謝 鄭淳尹先生 蒞臨本系
題 演 講



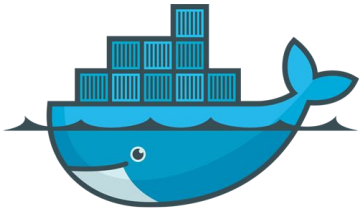
群益期貨
PITAL FUTURES
Fintech
來賓分享：
金湯尼、Philipz

課程大綱

1. 容器與資料科學
2. Rocker - RStudio
3. Python & Jupyter
4. Nvidia-docker 工具
5. Docker & TensorFlow GPU
6. Docker Compose for Data Science
7. 結語



1. 容器與資料科學



NATURE | TOOLBOX

Software simplified

Containerization technology takes the hassle out of setting up software and can boost the reproducibility of data-driven research.

Andrew Silver

29 May 2017

Researchers can access Docker images either from the platform's own hosting service, Docker Hub, or from registries of containers such as BioContainers and Dockstore, which allow the sharing of tools vetted by other scientists. Brian O'Connor at the University of California, Santa Cruz, who was the technical lead for the Dockstore registry, recommends that scientists look through container registries to find a tool that works for their project instead of trying to reinvent something that already exists.

運算膠囊整合研究計算所需的資料、研究方法、運算平台，輕易就可重現研究成果。

But actually getting the underlying Docker software to run properly can be challenging, says Simon Adar, chief executive of Code Ocean in New York, an online service that aims to simplify the process. “It's too technical, it was designed for developers to deploy complex systems.” The service, launched in February, creates what Adar calls “compute capsules”, which comprise code, data, results and the Docker container itself. Researchers upload their code and data, and then either execute it in a web browser or share it with others — no installation required. Adar likens the process to sharing a YouTube video. The company even offers a widget that enables users to embed executable code in web pages.

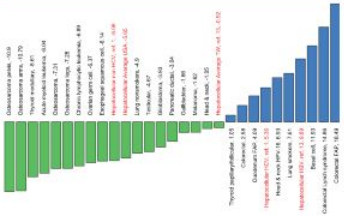
Discover & Run Scientific Code

Code Ocean is a cloud-based **executable** research platform

+ UPLOAD YOUR CODE

PHYSICS COMPUTER SCIENCE CHEMISTRY ENGINEERING BIOLOGY SOCIAL SCIENCES MATHEMATICS ECONOMICS

MEDICAL SCIENCES APR 2017



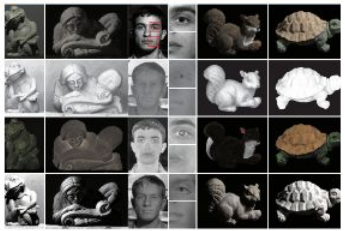
Maxime Tarabichi, Vincent Detours

R script and data for: A research note regarding "Variation in cancer risk...

Tomasetti and Vogelstein argued that 2/3 of human cancers are due to 'bad luck' and that "primary...

F1000Research, 2016

ENGINEERING JAN 2017



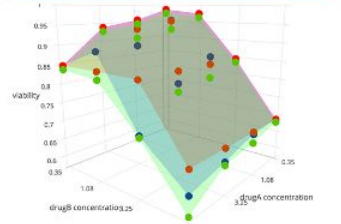
Daniel Hauagge, Scott Wehrwein, Kavita Bala, Noah..

Photometric Ambient Occlusion for Intrinsic Image Decomposition

We present a method for computing ambient occlusion (AO) for a stack of images of a...

IEEE Transactions on Pattern Analysis an..., 2015

BIOINFORMATICS JAN 2017



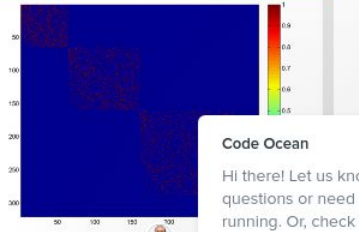
Jennifer O'Neil, Yair Benita, Igor Feldman, Melissa C...

Combo Synergy of Two Drugs

Combination drug therapy is a widely used paradigm for managing numerous...

Molecular Cancer Therapeutics, 2016

COMPUTER SCIENCE MAY 2017



Bo Yang, William K. Ch...m

The FEC Algorithm f... Community Minin... Signed Network

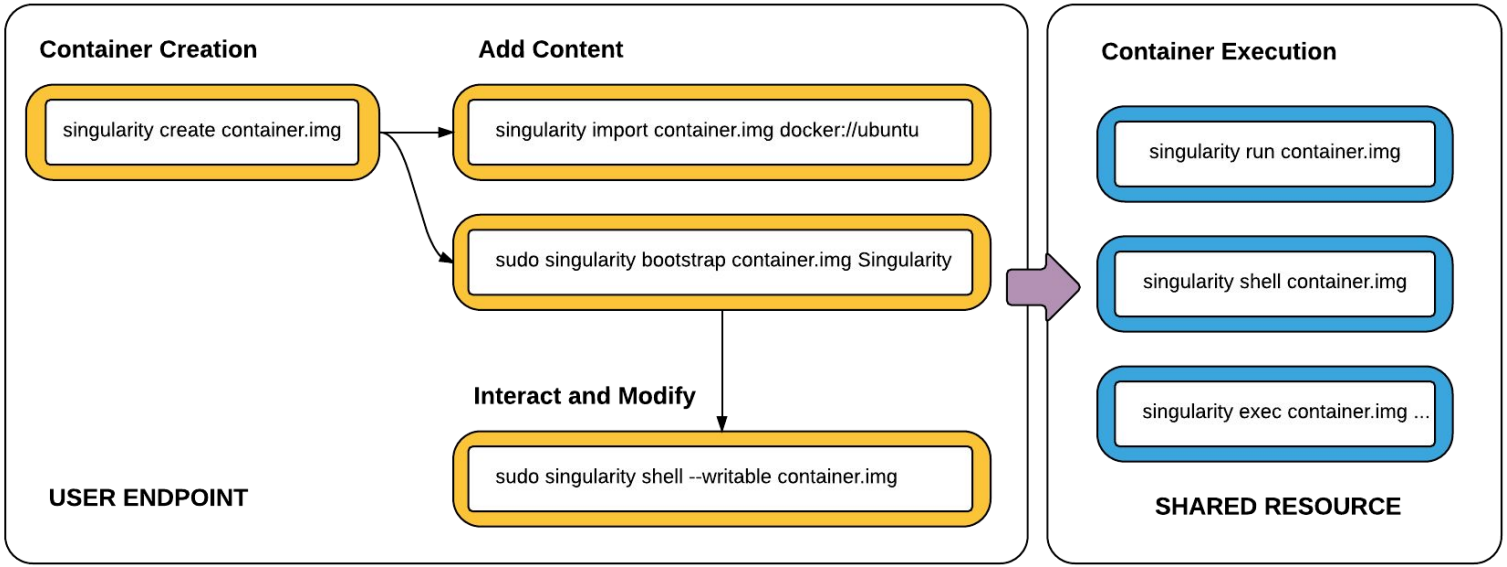
This code is used for community mining from signed networks. The algorithm contains two main...

IEEE Transactions on Knowledge and Dat.., 2007

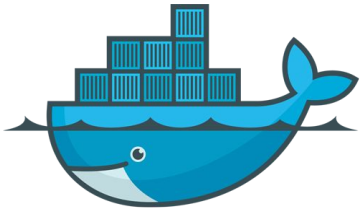
Singularity

A Container for HPC

Singularity HPC Container Technology Moves Out of the Lab



2. Rocker - RStudio



R & RStudio

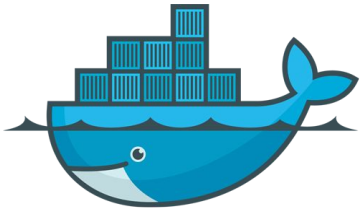
https://github.com/rocker-org/rocker/wiki/Using-the-RStudio-image

$ docker run -d -p 8787:8787 rocker/rstudio

- username: rstudio
- password: rstudio

http://ropenscilabs.github.io/r-docker-tutorial/

http://tutorials.iq.harvard.edu/R/Rgraphics/Rgraphics.html



Commit to a New Docker image

install.packages("ggplot2")

docker commit CONTAINER_ID

docker run NEW_DOCKER_IMAGE

rOpenSci

rOpenSci packages

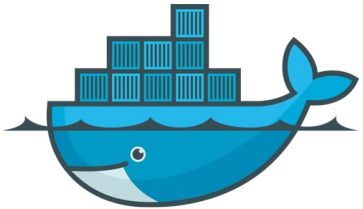
https://ropensci.org/tutorials/

This is a complete list of all available rOpenSci packages. Packages are grouped by ones that acquire data, full-text of journal articles, altmetrics, data publication, focus on scalable and reproducibile computing, data visualization, data tools, image processing, taxonomy, HTTP tools, data analysis or geospatial work. Packages with a CRAN, or BIOC sign are stable versions that you can quickly install from your nearest mirror using `install.packages("PACKAGE_NAME")`. Others are in various stages of development (bleeding edge packages are not listed here) and you can learn more by following our GitHub organization page, and our GitHub organization for bleeding edge projects. All of our software packages are open source. Please see package description files for more information on specific licenses.

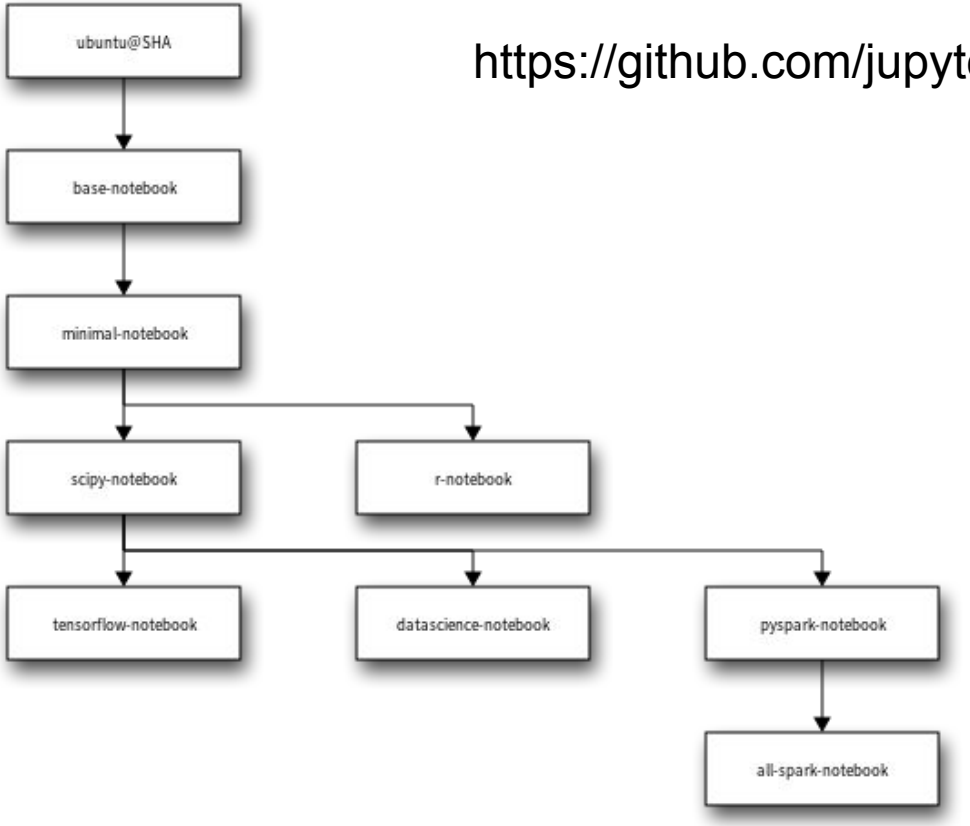
Many of our packages are community-contributed. If you are interested in contributing a package, please visit our onboarding repository for details.

Data Publication | Data Access | Literature | Altmetrics | Scalable & Reproducible Computing | Databases | Data Vizualization | Image Processing | Data Tools | Taxonomy | HTTP tools | Geospatial | Data Analysis

3. Python & Jupyter



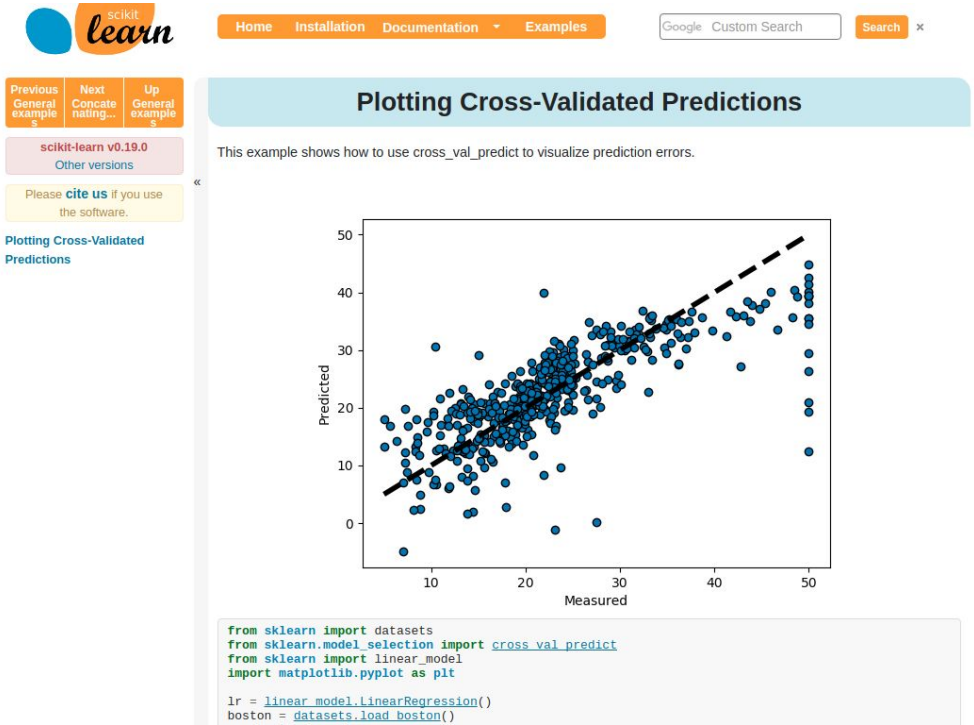
Jupyter Docker Stack



https://github.com/jupyter/docker-stacks

How to Use

$ docker run -it --rm -p 8888:8888 jupyter/scipy-notebook



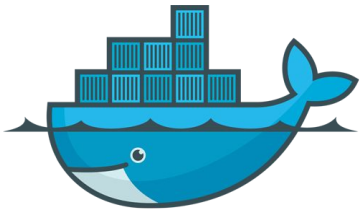
Volume for Save

$ docker run -it --rm -p 8888:8888 -v YOUR_FOLDER:/home/jovyan/work jupyter/scipy-notebook

Docker Volume

vSphere, NFS, AWS EFS, Azure File Storage

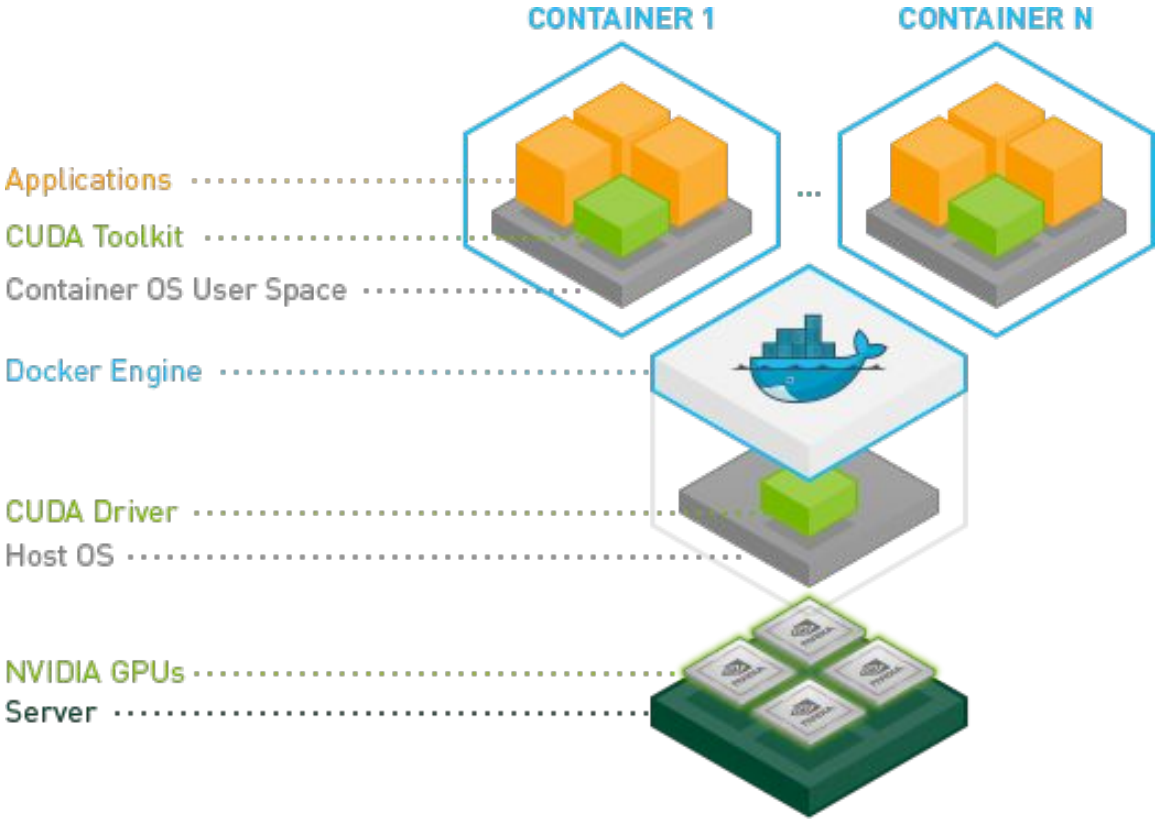
4. Nvidia-docker 工具



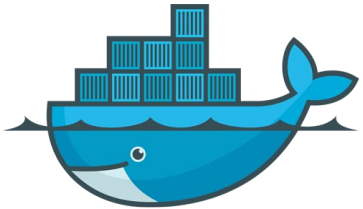
Docker + TensorFlow + GPU

- Machine Learning, Deep Learning
- TensorFlow Docker images
- nvidia-docker, All-in-one DL image Deep Learning

Docker & NVIDIA GPUs



5. Docker & TensorFlow GPU

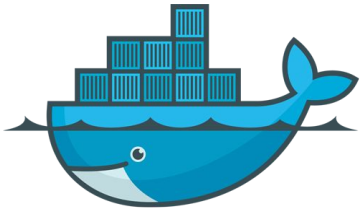


Just ONE Line

nvidia-docker run -it --rm -p 8888:8888 tensorflow/tensorflow:1.3.0-gpu

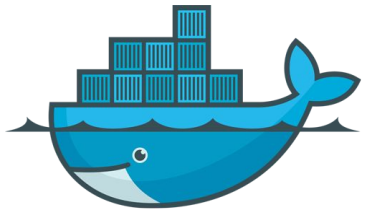
Only Support Linux

6. Docker Compose for Data Science



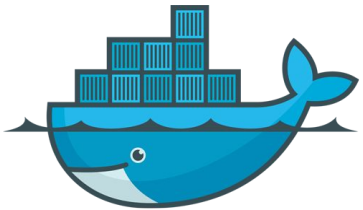
Compose File Sample (1/3)

```
version: '2'
services:
   db:
     image: mysql:5.7
     volumes:
       - db_data:/var/lib/mysql
     restart: always
     environment:
       MYSQL_ROOT_PASSWORD: wordpress
       MYSQL_DATABASE: wordpress
       MYSQL_USER: wordpress
       MYSQL_PASSWORD: wordpress
```



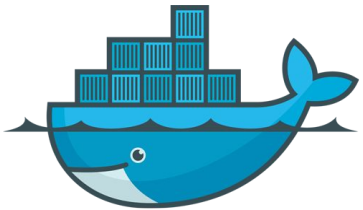
Compose File Sample (2/3)

```
wordpress:
  depends_on:
    - db
  image: wordpress:latest
  ports:
    - "80:80"
  restart: always
  environment:
    WORDPRESS_DB_HOST: db:3306
    WORDPRESS_DB_PASSWORD: wordpress
```

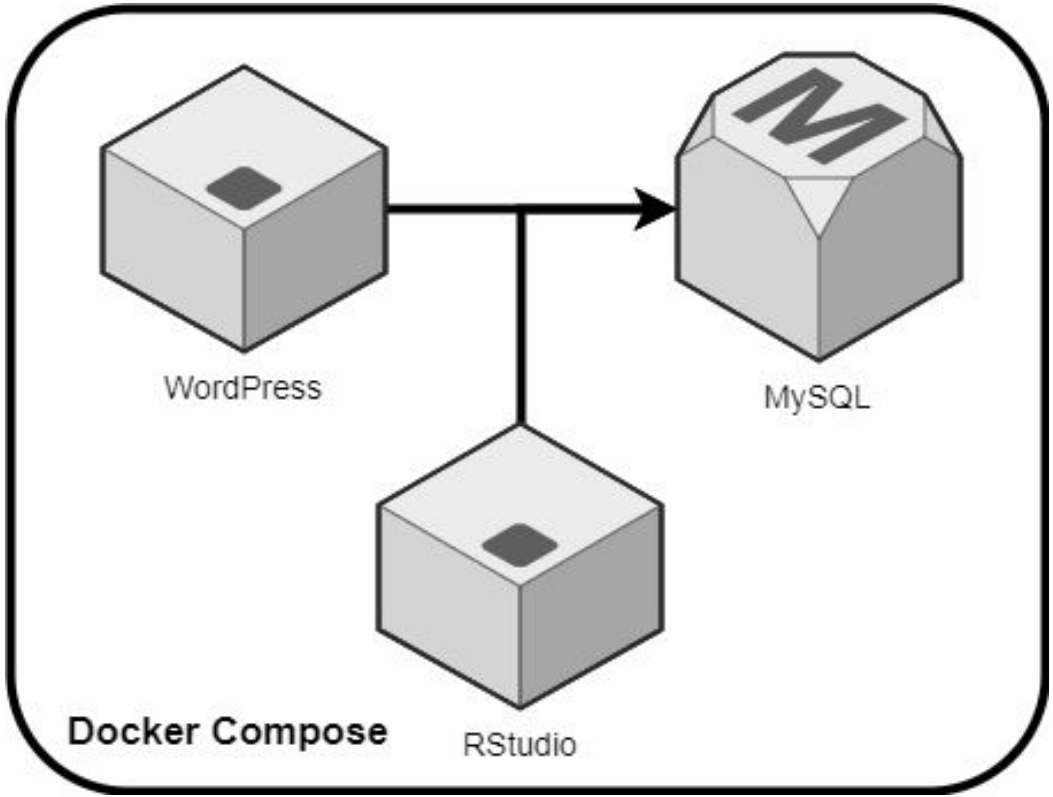


Compose File Sample (3/3)

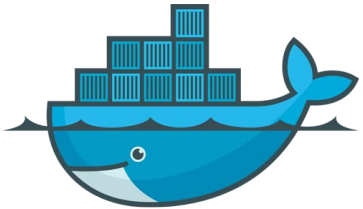
```
  ropensci:
    depends_on:
      - wordpress
    image: rocker/ropensci:latest
    ports:
      - "8787:8787"
    restart: always
volumes:
   db_data:
```



RStudio with MySQL DB



7. 結語



容器思維

1. 各階段如何使用容器
2. 直接用容器，不再重頭安裝
3. 官方映像檔或依需要客製
4. 營運環境亦使用容器
5. 在易用性與尺寸取得平衡
6. 使用標籤(tag)區分版本
7. 擺脫程式語言限制，善用各語言優點

TensorFlow & Singularity

README.md

Tensorflow-singularity-container-with-GPU-support

This repository provides a bootstrap definition file to build Tensorflow (1.1.0) singularity container with Nvidia GPU support based on singularity 2.3 release.

How to build

1. Install singularity 2.3 release. You can see the installation instructions on singularity homepage (section: Build an RPM from the source).
2. Download cuda 8.0 (cuda_8.0.61_375.26_linux-run) and cudnn5.1 (cudnn-8.0-linux-x64-v5.1.tgz) (Here I assume that the nvidia driver has been installed in your host machine) and store the downloaded files and above scripts under the same folder.
3. Run "sh build.sh" (assume that you have sudo access)
4. copy tensorflow_gpu-1.1.0-cp27-linux_x86_64.img into your own local folder and change its owner and group (sudo chown your_user_id:your_group_id tensorflow_gpu-1.1.0-cp27-linux_x86_64.img) so that you can run it with local user.
5. Run "singularity exec --nv tensorflow_gpu-1.1.0-cp27-linux_x86_64.img python hello_world.py" to check whether it works (where flag '--nv' is used by singularity to automatically detect nvidia driver in the host machine since release 2.3).

FaaS & Machine Learning

FaaS Gateway

CREATE NEW FUNCTION

func_tensorflow

func_tensorflow

Replicas: 1

Invocation count: 3

Image: jmkhael/faas-tensorflow:1.0@sha256:fdae576e95e18c6cb2db0d94217242c3da221139

Invoke function

INVOKE

Text

JSON

Request body

https://upload.wikimedia.org/wikipedia/zh/3/34/Lenna.jpg

Response status

200

Response body

Downloading https://upload.wikimedia.org/wikipedia/zh/3/34/Lenna.jpg under /root/images

Successfully downloaded Lenna.jpg 105880 bytes.
bonnet, poke bonnet (score = 0.17255)
sombrero (score = 0.05238)
cloak (score = 0.04602)

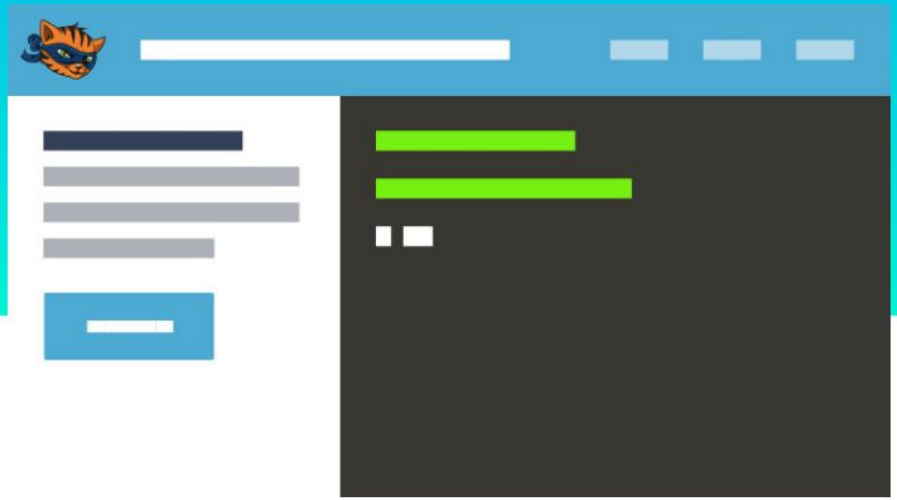
https://github.com/jmkhael/faas-tensorflow

Katacoda

Learn Teach Embed Log In Sign Up

Learn new technologies right in your browser

Interactive Technical Learning Platform for Software Engineers



Learn these technologies (with more to come)

ker_p

Play with Docker Classroom

The Play with Docker classroom brings you labs and tutorials that help you get hands-on experience using Docker. In this classroom you will find a mix labs and tutorials that will help Docker users, including SysAdmins, IT Pros, and Developers. There is a mix of hands-on tutorials right in the browser, instructions on setting up and using Docker in your own environment, and resources about best practices for developing and deploying your own applications.

We recommend you start with one of our Getting Started Guides, and then explore the individual labs that explore many advanced features of Docker

Getting Started Guides

For a comprehensive approach to understanding Docker, choose your preferred journey.

Getting Started Walk-through for IT Pros and System Administrators
Learn more about Docker, how it works and how it can help you deploy secure, scalable applications and save money along the way.

Getting Started Walk-through for Developers
Learn the core concepts of Docker and how it can make building apps faster, easier, and more secure.

Or for a full list of individual labs on this site, check out our labs page

Full list of individual labs


Thank you
Docker可省下比金錢更寶貴的時間！