

Docker: o que é e para que serve

Luiz Sol e Marcos Vinicius

2019-03-11

O problema

A pessoa A desenvolveu um *pipeline* de *Machile Leearning* em Python no seu computador pessoal.

O *setup* da pessoa A é:

- Majaro 17.1 (Arch Linux)
- Python 3.5
- PostgreSQL 10.3 (Banco de Dados)
- PyTorch 0.9
- Pandas 0.14

Agora a pessoa B precisa revisar o código (ver se foi bem escrito e se funciona corretamente).

O *setup* da pessoa *B* é:

- Windows 10
- Python 3.7
- MySQL 5.7 (Banco de Dados)
- TensorFlow 0.9
- Pandas 0.21

E uma vez que o código for aprovado ele deverá ser colocado no servidor de produção (máquina que irá rodar de fato a aplicação) C.

O *setup* do servidor *C* é:

- RHEL 7.6 (Red Hat Linux)
- Python 3.6
- Cassandra 3.11 (Banco de Dados)
- TensorFlow 0.7
- Pandas 0.26

E agora José?

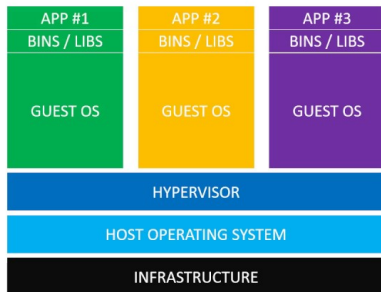
Possíveis soluções:

- **Solução 1:** Obrigar todo mundo a usar os mesmos softwares que estão no servidor de produção
- **Problemas da solução 1:**
 - Servidores tendem a utilizar versões antigas e estáveis, o que pode atrapalhar os pesquisadores
 - Nem sempre é possível usar para desenvolvimento o que se usa em produção (licenças, interface com o usuário, demanda computacional etc)
 - Restringir pesquisadores e desenvolvedores pode implicar a fuga de capital humano qualificado

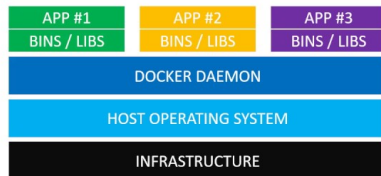
- **Solução 2:** Utilizar *máquinas* virtuais que espelhem o setup de produção
- **Problemas da solução 2:**
 - Máquinas virtuais são grandes (~10GB) e consomem bastante memória (~6GB) por si só
 - Versionamento (controle de versões) de máquinas virtuais não é uma tarefa simples (arquivos binários)
 - A interação entre a máquina hópedeira e a máquina hóspede (máquina virtualizada) nem sempre é simples (arquivos, rede, *clipboard* etc)

Docker ao resgate

O *Docker* se propõe a utilizar o próprio kernel do sistema operacional para executar as aplicações das máquinas virtualizadas.



Virtual Machines



Docker Containers

Os principais conceitos do docker são:

- **Imagens:** instruções de como construir os containers
- **Containers:** as “máquinas virtuais” já construídas a partir dos containers
- **Volumes:** pastas no computador hospedeiro que o container irá utilizar para armazenar arquivos permanentemente

Imagens

São as “plantas” que serão utilizadas para construir os *containers*

São contruídas a partir de outras imagens que estão disponíveis no repositório público do Docker (entitulado Docker Hub)

Imagens de sistemas operacionais “puros”

☐ Docker Certified

Images

☐ Verified Publisher
Docker Certified And Verified Publisher Content

☒ Official Images
Official Images Published By Docker

Categories

☐ Analytics

☐ Application Frameworks

☐ Application Infrastructure

☐ Application Services

☐ Base Images

☐ Databases

☐ DevOps Tools

☐ Featured Images

☐ Messaging Services

☐ Monitoring

☒ Operating Systems

☐ Programming Languages

☐ Security

☐ Storage

Operating Systems

☐ Linux

☐ Windows

Architectures

☐ ARM

☐ ARM 64

☐ IBM POWER

☐ IBM Z

☐ PowerPC 64 LE

☐ x86

☐ x86-64

alpine
Updated an hour ago
A minimal Docker image based on Alpine Linux with a complete package index and only 5 MB in size!

Container Linux x86-64 IBM Z PowerPC 64 LE 386 ARM 64 ARM Featured Images Base Images Operating Systems

ubuntu
Updated an hour ago
Ubuntu is a Debian-based Linux operating system based on free software.

Container Linux IBM Z PowerPC 64 LE 386 ARM 64 ARM x86-64 Base Images Operating Systems

centos
Updated 44 minutes ago
The official build of CentOS.

Container Linux ARM 64 ARM x86-64 PowerPC 64 LE 386 Base Images Operating Systems

debian
Updated 44 minutes ago
Debian is a Linux distribution that's composed entirely of free and open-source software.

Container Linux PowerPC 64 LE ARM 64 386 ARM x86-64 IBM Z Base Images Operating Systems

fedora
Updated 44 minutes ago
Official Docker builds of Fedora

Container Linux IBM Z PowerPC 64 LE ARM 64 ARM x86-64 Base Images Operating Systems

amazonlinux

Imagens de sistemas operacionais com linguagens pré-instaladas

The screenshot displays the Docker Hub interface with a sidebar on the left and a main content area on the right. The sidebar includes filters for 'Images' (Official Images is selected), 'Categories' (Programming Languages is selected), 'Operating Systems' (Linux and Windows are selected), and 'Architectures' (ARM, ARM 64, IBM POWER, IBM Z, PowerPC 64 LE, x86, and x86-64 are selected). The main content area lists several programming language images:

- golang**: Updated 41 minutes ago. Description: Go (golang) is a general purpose, higher-level, imperative programming language. Tags: Container, Linux, Windows, PowerPC 64 LE, 386, ARM 64, ARM, x86-64, IBM Z, Programming Languages. Stats: 10M+ Downloads, 2.3K Stars.
- python**: Updated 41 minutes ago. Description: Python is an interpreted, interactive, object-oriented, open-source programming language. Tags: Container, Windows, Linux, PowerPC 64 LE, 386, ARM 64, ARM, x86-64, IBM Z, Programming Languages. Stats: 10M+ Downloads, 3.9K Stars.
- openjdk**: Updated 41 minutes ago. Description: OpenJDK is an open-source implementation of the Java Platform, Standard Edition. Tags: Container, Linux, Windows, ARM 64, x86-64, 386, ARM, IBM Z, PowerPC 64 LE, Programming Languages. Stats: 10M+ Downloads, 1.3K Stars.
- php**: Updated 41 minutes ago. Description: While designed for web development, the PHP scripting language also provides general-purpose use. Tags: Container, Linux, ARM 64, ARM, x86-64, IBM Z, PowerPC 64 LE, 386, Programming Languages. Stats: 10M+ Downloads, 4.3K Stars.
- ruby**: Updated 41 minutes ago. Description: Ruby is a dynamic, reflective, object-oriented, general-purpose, open-source programming language. Tags: Container, Linux, IBM Z, PowerPC 64 LE, 386, ARM 64, ARM, x86-64, Programming Languages. Stats: 10M+ Downloads, 1.6K Stars.
- java**: Updated 40 minutes ago. Stats: 10M+ Downloads, 2.0K Stars.

Imagens de sistemas operacionais com bancos de dados pré-instalados

The screenshot displays the Docker Hub interface for database images. On the left, there is a sidebar with filters for 'Images' and 'Categories'. The 'Images' section includes 'Verified Publisher' and 'Official Images' (checked). The 'Categories' section lists various categories, with 'Databases' checked. Below these are sections for 'Operating Systems' (Linux, Windows) and 'Architectures' (ARM, ARM 64, IBM POWER, IBM Z, PowerPC 64 LE, x86, x86-64). The main content area shows a list of database images: mongo, postgres, redis, mariadb, mysql, and crate. Each image entry includes its icon, name, update status, description, and a list of supported architectures. The 'mongo' image is highlighted as the 'OFFICIAL IMAGE' with 10M+ downloads and 5.6K stars. The 'postgres' image is also an 'OFFICIAL IMAGE' with 10M+ downloads and 6.3K stars. The 'redis' image is an 'OFFICIAL IMAGE' with 10M+ downloads and 6.6K stars. The 'mariadb' image is an 'OFFICIAL IMAGE' with 10M+ downloads and 2.6K stars. The 'mysql' image is an 'OFFICIAL IMAGE' with 10M+ downloads and 7.8K stars. The 'crate' image is an 'OFFICIAL IMAGE' with 10M+ downloads and 115 stars.

☐ Docker Certified

Images

☐ Verified Publisher
Docker Certified And Verified Publisher Content

☒ Official Images
Official Images Published By Docker

Categories

☐ Analytics

☐ Application Frameworks

☐ Application Infrastructure

☐ Application Services

☐ Base Images

☒ Databases

☐ DevOps Tools

☐ Featured Images

☐ Messaging Services

☐ Monitoring

☐ Operating Systems

☐ Programming Languages

☐ Security

☐ Storage

Operating Systems

☐ Linux

☐ Windows

Architectures

☐ ARM

☐ ARM 64

☐ IBM POWER

☐ IBM Z

☐ PowerPC 64 LE

☐ x86

☐ x86-64

mongo
Updated an hour ago
MongoDB document databases provide high availability and easy scalability.
Container Windows Linux x86-64 ARM 64 Databases

postgres
Updated an hour ago
The PostgreSQL object-relational database system provides reliability and data integrity.
Container Linux ARM 64 ARM x86-64 IBM Z PowerPC 64 LE 386 Databases

redis
Updated an hour ago
Redis is an open source key-value store that functions as a data structure server.
Container Linux Windows IBM Z PowerPC 64 LE 386 ARM 64 ARM x86-64 Databases

mariadb
Updated an hour ago
MariaDB is a community-developed fork of MySQL intended to remain free under the GNU GPL.
Container Linux 386 PowerPC 64 LE ARM 64 x86-64 Databases

mysql
Updated an hour ago
MySQL is a widely used, open-source relational database management system (RDBMS).
Container Linux x86-64 Databases

crate
Updated an hour ago

OFFICIAL IMAGE
10M+ 5.6K
Downloads Stars

OFFICIAL IMAGE
10M+ 6.3K
Downloads Stars

OFFICIAL IMAGE
10M+ 6.6K
Downloads Stars

OFFICIAL IMAGE
10M+ 2.6K
Downloads Stars

OFFICIAL IMAGE
10M+ 7.8K
Downloads Stars

OFFICIAL IMAGE
10M+ 115
Downloads Stars

Ex: Se eu quiser construir uma imagem para um container Linux que irá fazer *webscrapping* utilizando Celery eu posso criar a imagem:

- A partir da imagem de um Linux puro (Alpine, Ubuntu etc)
- A partir da imagem oficial do Python (que por sua vez foi criada a partir da imagem de um Linux puro)
- A partir de uma imagem que já tenha Celery instalado e configurado

Trechos do Dockerfile do *MySQL*

```
# Usando outra imagem como ponto de partida
FROM debian:stretch-slim # ...
# Executando comandos para instalar pacotes
RUN apt-get update && # ...
# Definindo variáveis de ambiente
ENV GOSU_VERSION 1.7 # ...
# Expondo uma das pastas para o hospedeiro
VOLUME /var/lib/mysql # ...
# Copiando arquivos para a imagem
COPY config/ /etc/mysql/ # ...
# Expondo a porta 3306 para o hospedeiro
EXPOSE 3306 33060 # ...
# Executando o daemon do MySQL
CMD ["mysqld"]
```

Fonte: MySQL Docker File

Esses comandos serão executados toda vez que um novo container de MySQL for construído.

Observações importante: **o sistema de arquivo dos *containers* é volátil!**
(Isso ficará mais claro quando falarmos sobre volumes)

Containers

Containers são instâncias de imagens.

- Um *container* é gerado a partir de uma imagem
- Um hospedeiro pode executar vários *containers* simultaneamente

Analogia: a imagem é o molde e os *containers* são os objetos criados a partir desse molde

TODO: colocar print de vários containers rodando (meu blog)