

# ***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?



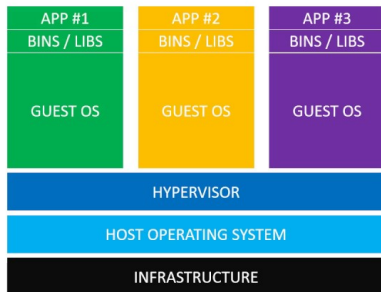
# Possíveis soluções:

- **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

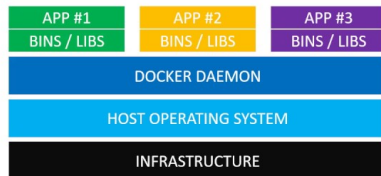
- **Solução 2:** Utilizar *máquinas* virtuais que espelhem o setup de produção
- **Problemas da solução 2:**
  - Máquinas virtuais são grandes (~10GB) e consomem bastante memória (~6GB) por si só
  - Versionamento (controle de versões) de máquinas virtuais não é uma tarefa simples (arquivos binários)
  - 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**
- ***Containers***
- **Volumes**

## Imagens

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