

# TRILHA BIG DATA

28 de Agosto de 2020



**"Big Data usando recursos não gerenciados pela cloud, é possível?"**

*Erika Nagamine - Oracle*

**Início às 12h25**



# Recadinhos



- Dá um duplo clique aqui na apresentação para melhor experiência
- Ou F11 para maximizar o navegador





# Big Data in a cloud world

## Using only self-managed services

---

**Erika Nagamine**

Cloud Solutions Engineer

## Safe harbor statement

---

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

Statements in this presentation relating to Oracle's future plans, expectations, beliefs, intentions and prospects are "forward-looking statements" and are subject to material risks and uncertainties. A detailed discussion of these factors and other risks that affect our business is contained in Oracle's Securities and Exchange Commission (SEC) filings, including our most recent reports on Form 10-K and Form 10-Q under the heading "Risk Factors." These filings are available on the SEC's website or on Oracle's website at <http://www.oracle.com/investor>. All information in this presentation is current as of September 2019 and Oracle undertakes no duty to update any statement in light of new information or future events.



# Erika Nagamine

Cloud Solutions Engineer

Data Management | Data Engineer | Data Scientist | Analytics

Tech Brazil Cloud Solutions Engineer

Oracle



<https://github.com/erikanagamine>

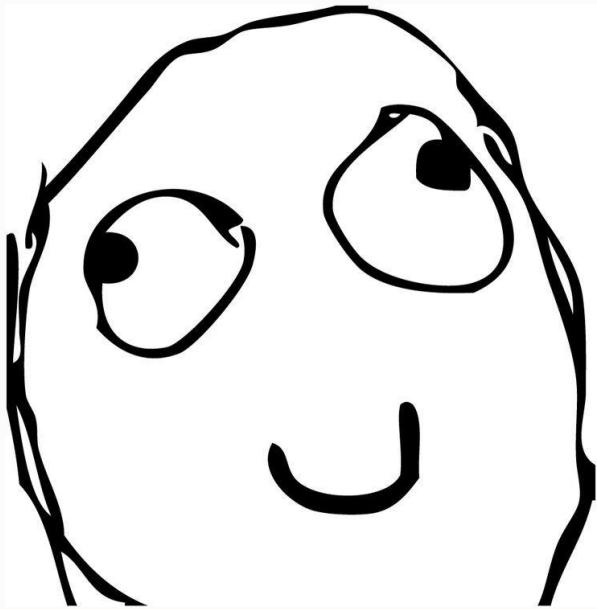


[erika.nagamine@oracle.com](mailto:erika.nagamine@oracle.com)

# Agenda

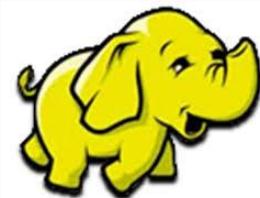
---

- Why?
- Project outcomes
- How?
- Cloud
- Docker
- terraform









# Why?

---



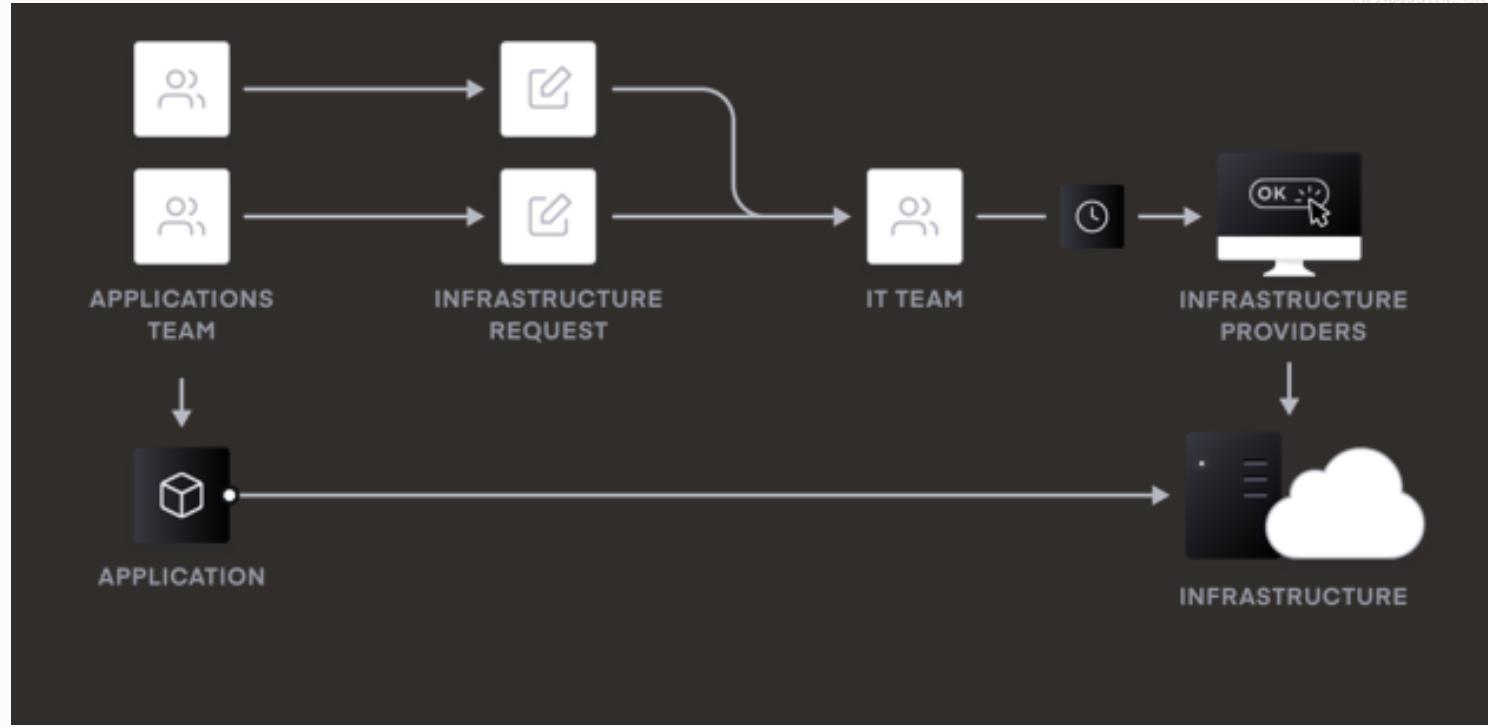
# Project outcomes....



# How?

---





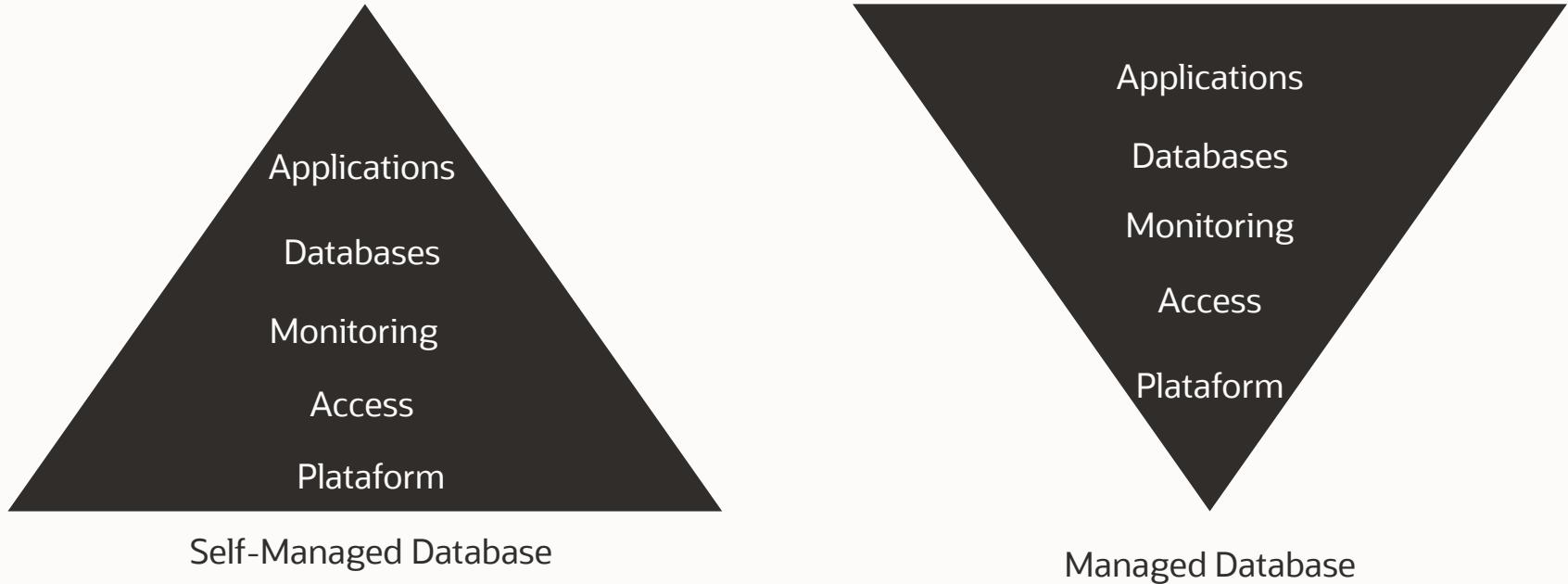
Cloud A



Cloud B



# Administration Time Spent



Source: an introduction to Cloud databases – o'reilly

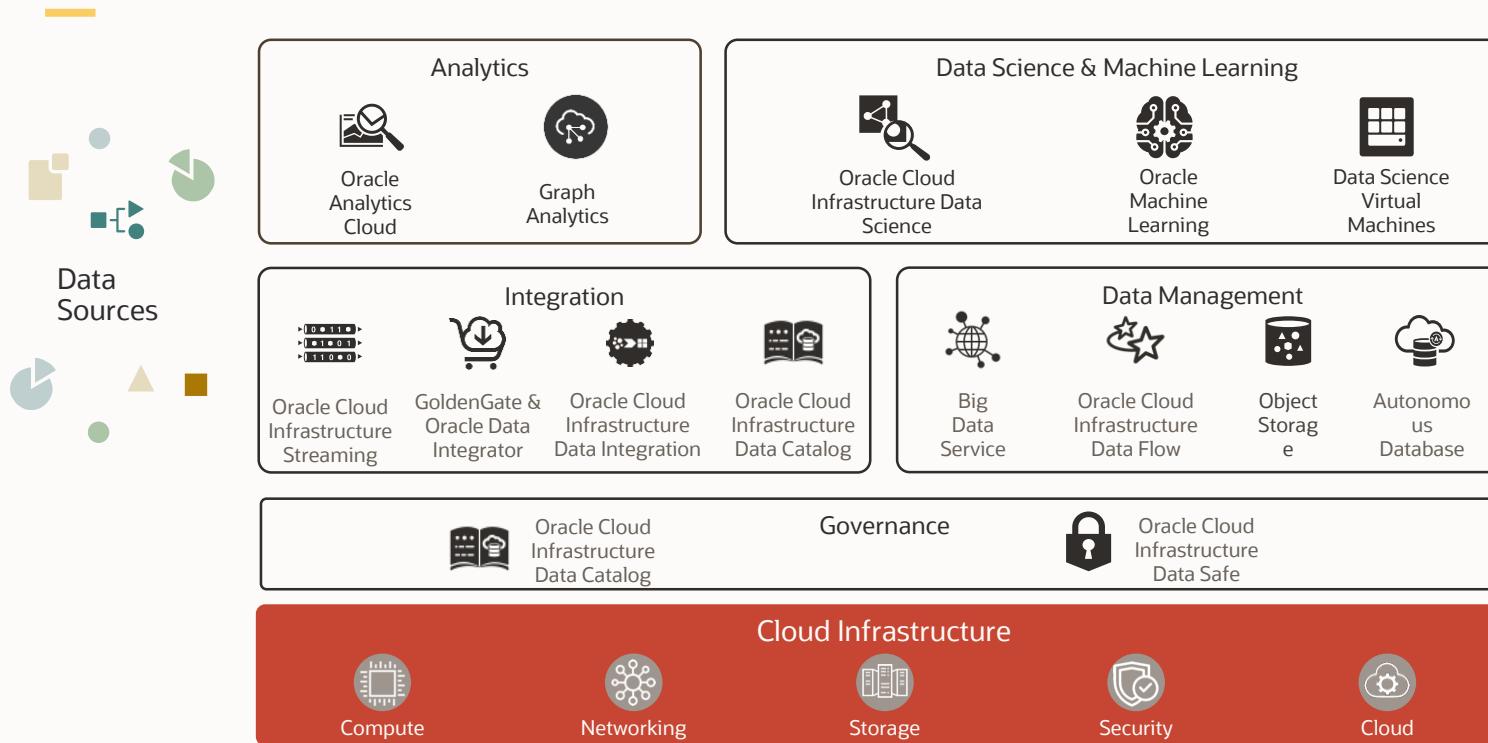
The image displays three cloud service consoles side-by-side:

- AWS (Left):** Shows the navigation bar with 'Services' and 'Resource Groups'. Below it is a sidebar with links to History, Console Home, Support, EC2, Direct Connect, Database Migration Service, and CloudWatch.
- Google Cloud Platform (Middle):** Shows the main dashboard with sections for Home, Marketplace, Billing, API & Services, Support, IAM & admin, Getting started, Security, COMPUTE, App Engine, Compute Engine, and Kubernetes Engine. A search bar at the top right says "Find a service by name or feature (for ex...)"
- Oracle Cloud (Right):** Shows the main dashboard with sections for Database, Data Safe, Big Data, Digital Assistant, Data Catalog, Data Flow, Data Integration, Data Science, Solutions and Platform, Analytics, Resource Manager, Email Delivery, Application Integration, Monitoring, Developer Services, Blockchain Platform, Marketplace, VMware Solution, More Oracle Cloud Services, Platform Services, Classic Data Management Services, Governance and Administration, Account Management, Identity, Security, Governance, and Administration. A search bar at the top right says "Search for resources, services, and documentation".

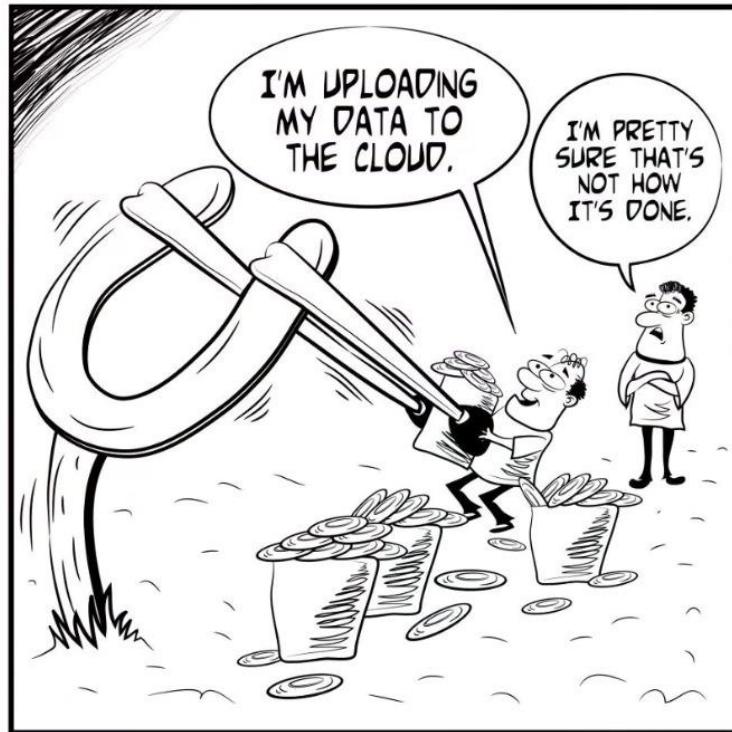
Quick Actions and Start Exploring sections are also visible on the Oracle Cloud interface.

# Oracle Data Platform

Machine learning supported by data sources, ingest, management and analytics



# Cloud



Stacks | Oracle Cloud Infrastructure X +

← → C https://console.us-ashburn-1.oraclecloud.com/resourcemanager/stacks

US East (Ashburn) ⚙️ 🔍 ⓘ 🌐

ORACLE Cloud Search for resources and services

Resource Manager

**Stacks**

Jobs

List Scope

COMPARTMENT

datascience

erikanagamine202004 (root)/datascience

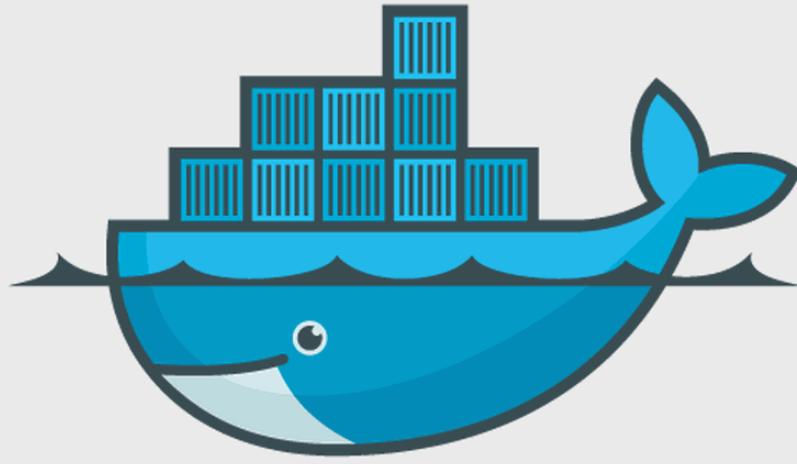
Create Stack

Name	Description	State	Created
No items			

Showing 0 Stacks < Page 1 >

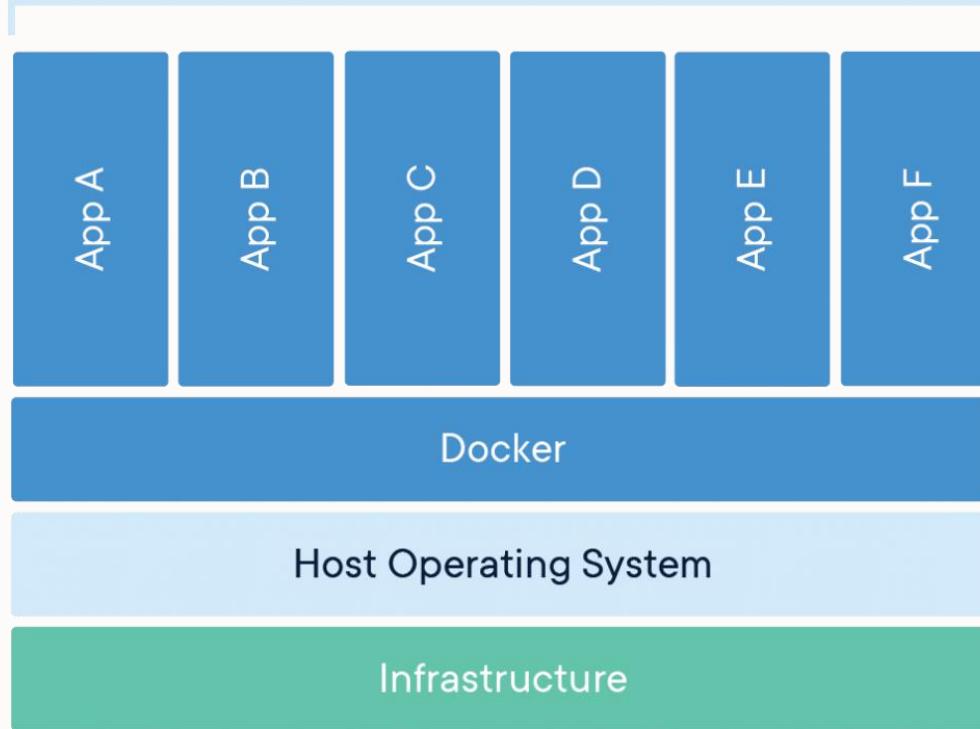
Tag Filters [add](#) | [clear](#)

no tag filters applied



# docker

## Containerized Applications



New Pricing and Packaging announced! [Learn more...](#)

Search for great content (e.g., mysql)

Explore

Repositories

Organizations

Get Help ▾

erikanagamine ▾



# Welcome to Docker Hub

## Download and Take a Tutorial

Get started by downloading Docker Desktop, and learn how you can build, tag and share a sample image on Hub.

[Get started with Docker Desktop](#)

### Create a Repository

Push container images to Docker Hub



### Create an Organization

Manage Docker Hub repositories with your team

Access the world's largest library of container images

cassandra - Docker Hub x + hub.docker.com/\_/cassandra

New Pricing and Packaging announced! [Learn more...](#)

 docker hub

Explore Repositories Organizations Get Help erikanagamine 

Explore > [cassandra](#) >

Using 0 of 1 private repositories. [Get more](#)

 **cassandra** ☆

[Docker Official Images](#)  
Apache Cassandra is an open-source distributed storage system.

↓ 100M+ 

Container Linux ARM PowerPC 64 LE ARM 64 386 x86-64 Databases Official Image

Copy and paste to pull this image

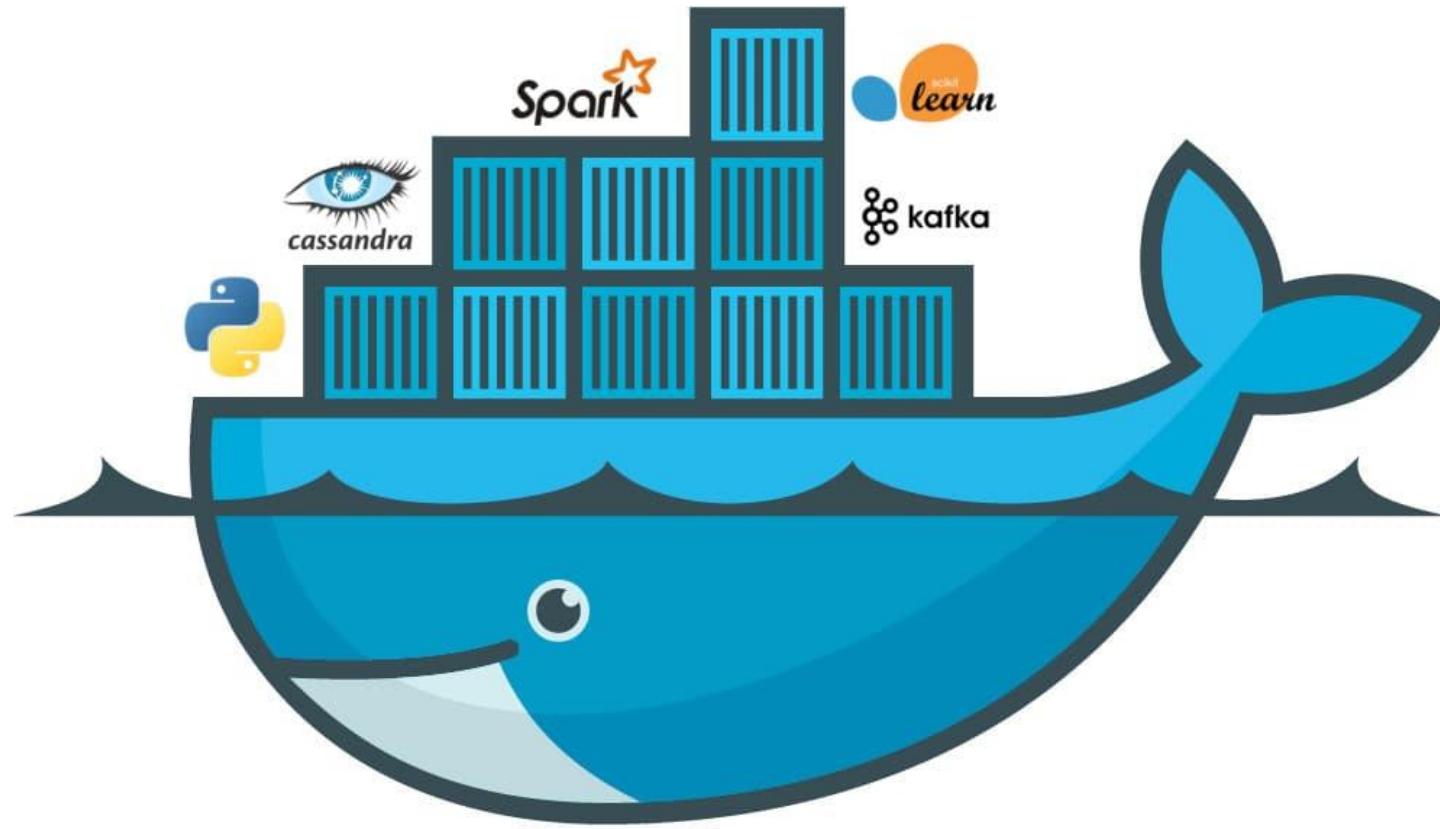
`docker pull cassandra` 

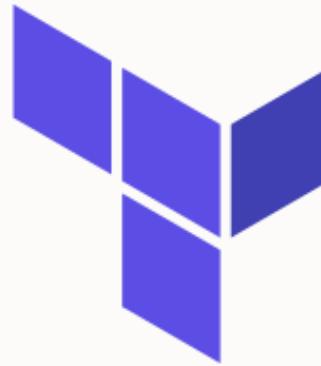
[View Available Tags](#)

Description [Reviews](#) [Tags](#)

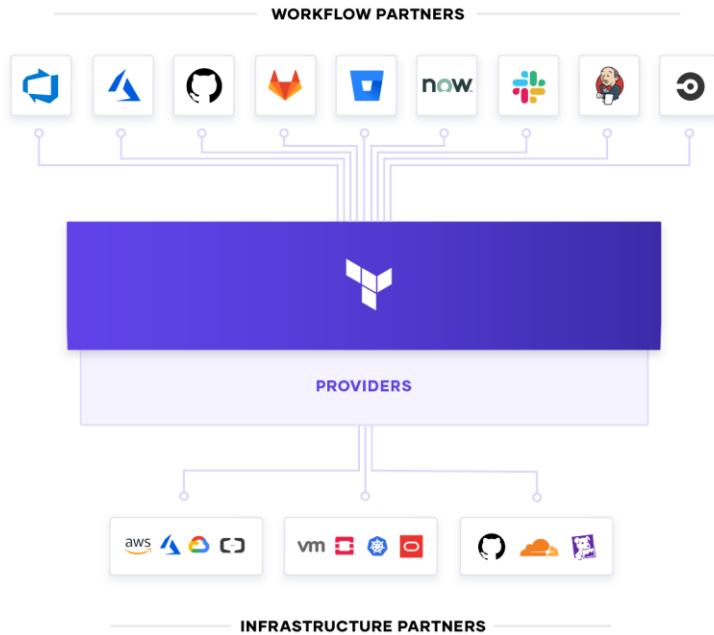
## Quick reference

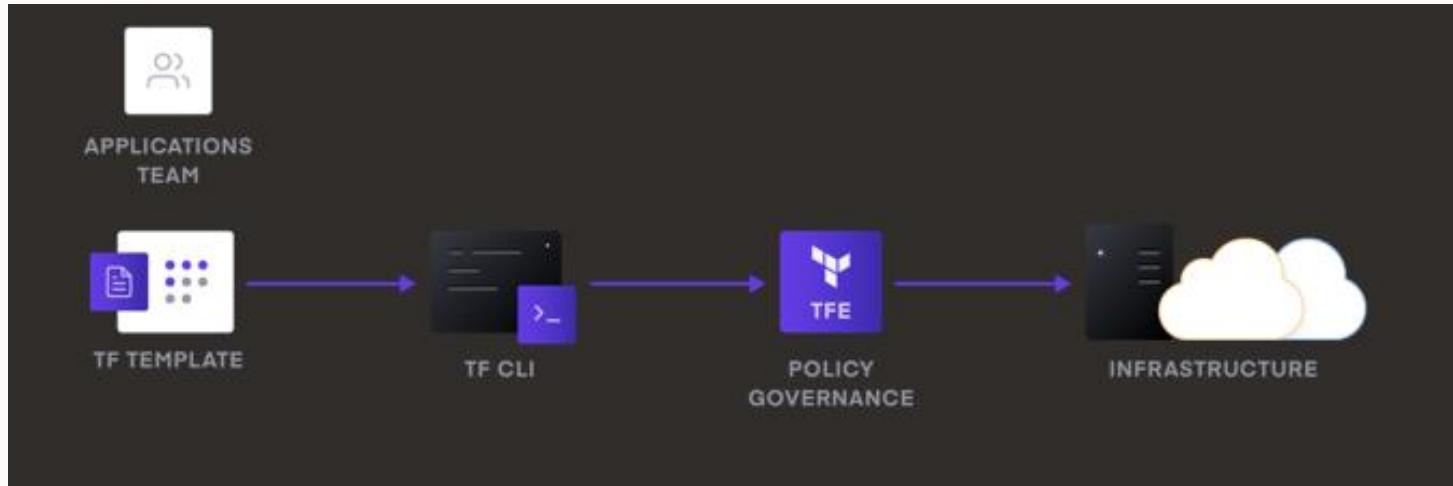
- **Maintained by:** the Docker Community
- **Where to get help:** the Docker Community Forums, the Docker Community Slack, or Stack Overflow





# Terraform





```
variables.tf / ...
1 # Configure the Oracle Cloud Infrastructure provider with an API Key
2 provider "oci" {
3     tenancy_ocid = "var.tenancy_ocid"
4     user_ocid = "var.user_ocid"
5     fingerprint = "var.fingerprint"
6     private_key_path = "var.private_key_path"
7     region = "var.region"
8 }
9
10 # Get a list of Availability Domains
11 # 0 references
12 data "oci_identity_availability_domains" "ads" {
13     compartment_id = "var.tenancy_ocid"
14 }
```

# Infrastructure as a code (IaC)

---

- Reuse
- Compose
- Configure
- Debug
- Test
- Version
- Document
- Standardizing
- Vendor independence

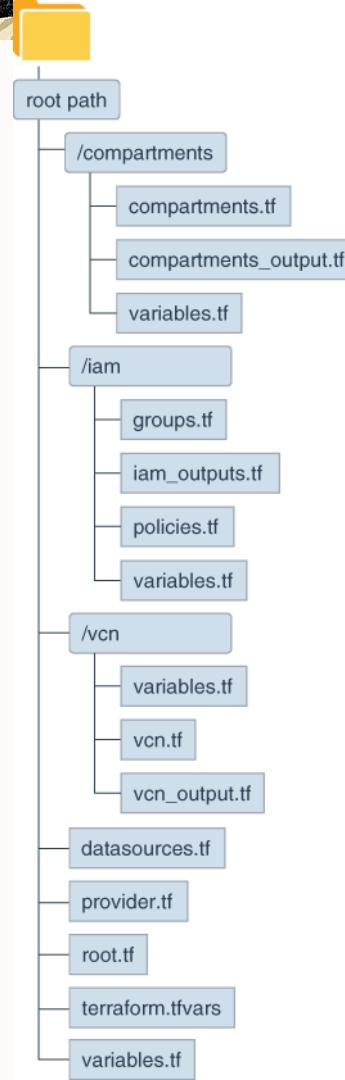
# Terraform: Before you begin

---

- Download: <https://www.terraform.io/downloads.html>
- cloud account: <https://www.oracle.com/cloud/free/>
- A computer that has the following software and access to the internet:
  - A utility to generate API signing keys.
    - Most UNIX-like systems have openssl. On Windows, you can use Git Bash.
  - A utility to generate SSH key pairs.
    - Most UNIX-like systems have ssh-keygen. On Windows, you can use PuTTY.
  - A web browser.
- Some information about terraform: <https://www.terraform.io/intro/>

# Terraform Structure

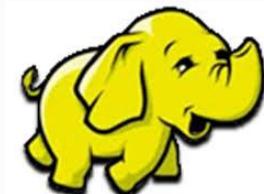
- Compartments
- IAM
- vcn
- The following files are in the root directory:
  - datasources.tf
  - provider.tf
  - root.tf
  - terraform.tfvars
  - variables.tf



# And finally

---





```
enagamine@ENAGAMINE:~/terraform
xdnig]
module.iam.oci_identity_policy.iam_admin_managers: Creation complete after 2s [id=ocid1.policy.oc1..aaaaaaaaatdiph3356z7qw7kd2xgxoarx4a2a34suvvtepp2wajq7373idla]
module.iam.oci_identity_policy.sys_admins: Creation complete after 1s [id=ocid1.policy.oc1..aaaaaaaaarhy3lmpdibaiviwin5i67d6m3kh4kqu3272gxeroqgvxp7d4ez]
module.iam.oci_identity_policy.iam_managers: Creation complete after 1s [id=ocid1.policy.oc1..aaaaaaaaascjcggnz3nqsejwx3636t1ce1zwg7fah3agjnqkrh3gscvtva]
module.a.vcn.oci_core_nat_gateway.nat_gateway: Creation complete after 2s [id=ocid1.natgateway.oc1.iad.aaaaaaaaayvipv4w7o2u2fapdk12if5mb5xu3u7h61swu6hypugd]5vmkra]
module.iam.oci_identity_policy.network_admins: Creation complete after 2s [id=ocid1.policy.oc1..aaaaaaaaasfwf7vvb43jhq2cnnzbpzmi77hh7spewilc735ospitvmecc7fohq]
module.compute.oci_core_instance.compute: Still creating... [10s elapsed]
module.compute.oci_core_instance.compute: Still creating... [20s elapsed]
module.compute.oci_core_instance.compute: Still creating... [30s elapsed]
module.compute.oci_core_instance.compute: Still creating... [40s elapsed]
module.compute.oci_core_instance.compute: Still creating... [50s elapsed]
module.compute.oci_core_instance.compute: Still creating... [1m0s elapsed]
module.compute.oci_core_instance.compute: Still creating... [1m10s elapsed]
module.compute.oci_core_instance.compute: Still creating... [1m20s elapsed]
module.compute.oci_core_instance.compute: Still creating... [1m30s elapsed]
module.compute.oci_core_instance.compute: Still creating... [1m40s elapsed]
module.compute.oci_core_instance.compute: Still creating... [1m50s elapsed]
module.compute.oci_core_instance.compute: Still creating... [2m0s elapsed]
module.compute.oci_core_instance.compute: Still creating... [2m10s elapsed]
module.compute.oci_core_instance.compute: Creation complete after 2m11s [id=ocid1.instance.oc1.iad.anuwcljthqe43hqc42intu3ywwob5qz7bprm34b2dzw6krz13w4s63fzmmea]
module.compute.data.oci_core_vnic_attachments.app_vnics: Reading...
module.compute.null_resource.remote-exec: Creating...
module.compute.null_resource.remote-exec: Provisioning with 'remote-exec'...
module.compute.null_resource.remote-exec (remote-exec): Connecting to remote host via SSH...
module.compute.null_resource.remote-exec (remote-exec): Host: 132.145.162.29
module.compute.null_resource.remote-exec (remote-exec): User: ubuntu
module.compute.null_resource.remote-exec (remote-exec): Password: false
module.compute.null_resource.remote-exec (remote-exec): Private key: true
module.compute.null_resource.remote-exec (remote-exec): Certificate: false
module.compute.null_resource.remote-exec (remote-exec): SSH Agent: false
module.compute.null_resource.remote-exec (remote-exec): Checking Host Key: false
module.compute.data.oci_core_vnic_attachments.app_vnics: Read complete after 0s [id=2020-08-27 18:19:14.8099971 +0000 UTC]
module.compute.data.oci_core_vnic.app_vnic: Reading...
module.compute.data.oci_core_vnic.app_vnic: Read complete after 0s [id=ocid1.vnic.oc1.iad.abuwcljtti4nyaxiph2btrouq4amfeuet5xir5befk3c4j4rdc4y2odvlvtq]
module.compute.null_resource.remote-exec (remote-exec): Connected!
module.compute.null_resource.remote-exec (remote-exec): --2020-08-27 18:19:23-- https://raw.githubusercontent.com/erikanagamine/terraform/master/script.sh
module.compute.null_resource.remote-exec (remote-exec): Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 151.101.200.133
module.compute.null_resource.remote-exec (remote-exec): Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|151.101.200.133|:443... connected.
module.compute.null_resource.remote-exec (remote-exec): HTTP request sent, awaiting response...
module.compute.null_resource.remote-exec (remote-exec): 200 OK
module.compute.null_resource.remote-exec (remote-exec): Length: 1384 (1.4K) [text/plain]
module.compute.null_resource.remote-exec (remote-exec): Saving to: 'script.sh'

module.compute.null_resource.remote-exec (remote-exec): script.sh      0%      0  --.-KB/s
module.compute.null_resource.remote-exec (remote-exec): script.sh    100%   1.35K  --.-KB/s      in 0s

module.compute.null_resource.remote-exec (remote-exec): 2020-08-27 18:19:23 (25.6 MB/s) - 'script.sh' saved [1384/1384]

module.compute.null_resource.remote-exec (remote-exec): nohup: appending output to 'nohup.out'
module.compute.null_resource.remote-exec: Still creating... [10s elapsed]
module.compute.null_resource.remote-exec: Still creating... [20s elapsed]
module.compute.null_resource.remote-exec: Creation complete after 20s [id=6392166431637810089]

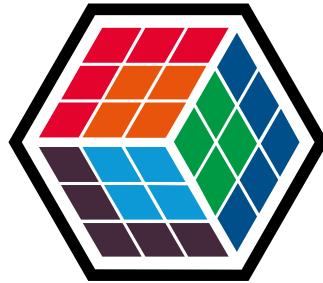
Apply complete! Resources: 30 added, 0 changed, 0 destroyed.
enagamine@ENAGAMINE:~/terraform$
```

Dúvidas?



Obrigada!





# THE DEVELOPER'S CONFERENCE