Aprendizaje automático en la nube

María Camila Durango
Software Engineer@Mercado Libre
mdurango@udemedellin.edu.co

2025

Clases

Marzo 19 - Introducción al MLOps, AWS (User root, user IAM, políticas, servicios, costos), prácticas de desarrollo, requerimientos del proyecto.

Marzo 25 - Introducción a S3, MLFlow tracking con ec2 y buckets

Marzo 26 - Introducción a dockers, ECR, feature engineering

Marzo 31 - Introducción a Sagemakers, basic deploys and jobs training

Abril 1 - Customs models y monitoring

Abril 2- Pipelines en SageMakers y Pipelines de MLOps

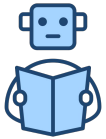
Abril 7 - LLM y RAGS (pendiente por validar) - kubernetes

Abril 8 - Entrega de proyecto final

Requerimientos para este curso

Cuenta de AWS con usuario IAM. Recomendación: Añadir doble auth a la cuenta para una mayor seguridad. Crear una cuenta con tarjeta de crédito.

Cuenta de Github y creación de repositorio público/privado con el material del curso. Se requiere previa configuración. Recomendación: Configuración con key SSH. Darle amor a github un fin de semana para conocer comandos básicos (git status, git add, git commit, git stash, git checkout, git switch) Aprende github en 15m



Conocimientos en Python, ML, Deep Learning.

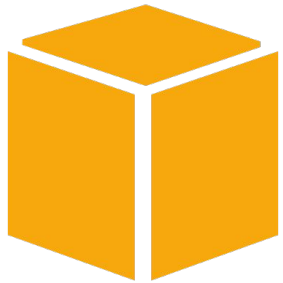
https://github.com/CamilaCortex/MLOps-AWS/tree/main (repo del curso)

AWS

Amazon Web Services (AWS) es un proveedor líder de servicios en la nube, ofreciendo una amplia gama de servicios, desde almacenamiento hasta potencia computacional.

Características Clave:

- Escalabilidad: Permite ajustar recursos según la demanda.
- Flexibilidad: Adaptación a diversos casos de uso.
- Pago por uso: Solo pagas por lo que utilizas.



¿Cómo lo hace?

Instancia

Definición: En el contexto de AWS, una instancia generalmente se refiere a una instancia de Amazon EC2 (Elastic Compute Cloud). Es un servidor virtual que ejecuta aplicaciones en la nube. Las instancias son configurables y tienen varios tipos (p. ej., t2, m5, c5) según las necesidades de computación, memoria y almacenamiento.

Uso: Cada instancia en AWS es independiente y tiene su propio sistema operativo. Las instancias se lanzan y gestionan a través de la consola de EC2 o mediante APIs de AWS.

Ejemplo: Si inicias una instancia de EC2 con un tipo de máquina específica (como una instancia m5.large), estás creando un servidor virtual que puedes usar para ejecutar aplicaciones, almacenar datos, etc.

clúster

Un clúster es un grupo de nodos que trabajan juntos para completar tareas complejas de procesamiento de datos o realizar operaciones que requieren recursos compartidos. Los clústeres están diseñados para proporcionar alta disponibilidad y escalabilidad.

Función en AWS:

En AWS, un clúster permite agrupar instancias de EC2 (o nodos en AWS EMR, por ejemplo) para gestionar aplicaciones distribuidas. Proporciona un ambiente donde varios nodos pueden comunicarse y trabajar en conjunto.

Ejemplos:

Un clúster de Amazon ECS (Elastic Container Service) que gestiona contenedores de aplicaciones distribuidas.

Un clúster de Amazon RDS (Relational Database Service) que agrupa instancias para equilibrar la carga y aumentar la disponibilidad.

Conceptos importantes

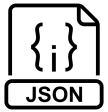
Usuario:

- Persona o aplicación
- Cada usuario tiene credenciales de acceso (user, password)
- Los usuarios siempre tienen políticas de permiso (están pueden darse por usuario o por grupos)

Roles:

- Lo asumen las entidades
- Es un conjunto de permisos que se pueden asumir por una entidad (usuario, servicio o aplicación) y que permite realizar acciones específicas en AWS.
- Son útiles para delegar permisos sin necesidad de compartir credenciales.
- Un rol que permite a un servicio de EC2 acceder a un bucket S3 para leer y escribir datos.

Política:



- Es un documento en formato JSON que define qué acciones están permitidas (o denegadas) sobre recursos en AWS.
- Las políticas determinan los permisos específicos que los roles o usuarios tienen sobre los recursos.

Tipos de usuarios:

- Root (super-admin): Tiene permisos ilimitados y puede gestionar todos los aspectos de la cuenta, incluidos la gestión de usuarios IAM, la configuración de facturación, y las acciones relacionadas con la seguridad (como la activación de MFA code).

- Usuario IAM: Se crea mediante el servicio de IAM y se le asignan permisos específicos para acceder a recursos de AWS. Cada usuario IAM es independiente y puede tener su propio conjunto de privilegios basados en las políticas asignadas.

Diferencias entre Roles y permisos a usuarios IAM

Permanencia vs Temporalidad:

Rol: Permisos temporales, asumidos por una entidad.

Usuario: Permisos permanentes hasta que sean modificados.

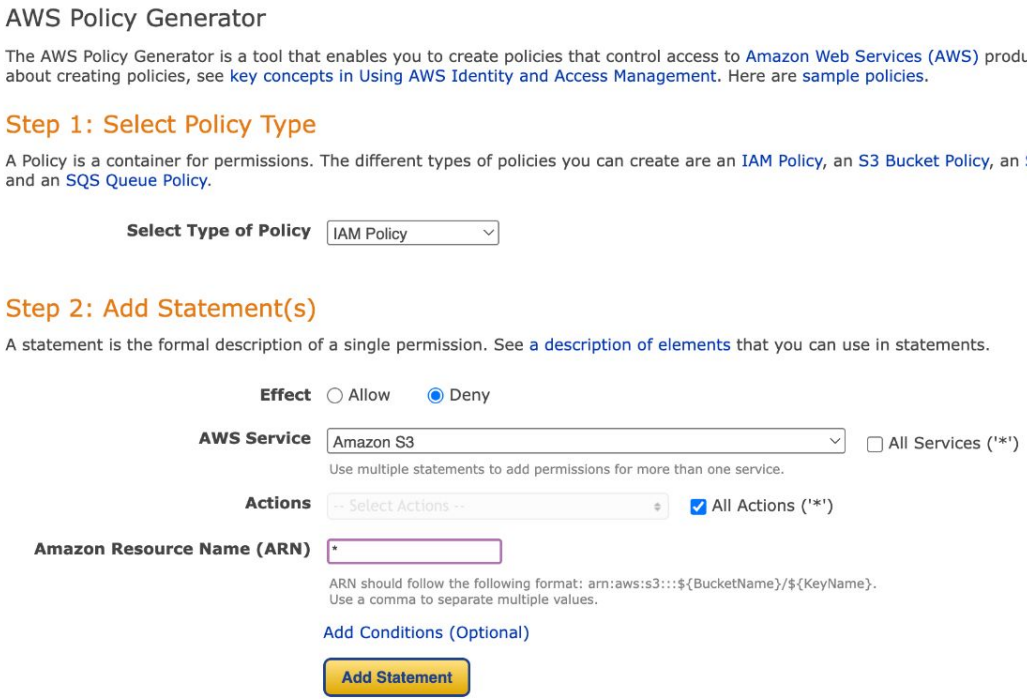
Ámbito de Uso:

- Rol: Se utiliza principalmente para servicios o situaciones donde no se desea definir usuarios específicos (por ejemplo, instancias de EC2).
- Usuario: Es más personal; cada usuario tiene su propio conjunto de permisos.

Práctica:

- Crear cuenta de AWS
- Usuario IAM
- Crear un alias
- Añadir MAF
- Ejemplo de política usando https://awspolicygen.s3.amazonaws.com/policygen.html

Ejemplo de política para no permitir creación de buckets en s3



EC2

Antes de hablar de EC2…

Puertos: Es una puerta en la MV que nos permite ejecutar diferentes aplicaciones y servicios que se comuniquen entre sí a través de una red, como Internet. Cada puerto tiene un id único que lo identifica, y con este número nos ayuda a saber qué información recibe/envía a través de esa puerta.

Amazon Elastic Compute Cloud (EC2)

Permite a los usuarios alquilar máquinas virtuales en las que podemos ejecutar aplicaciones. EC2 proporciona una capacidad informática escalable en la nube, lo que significa que podemos aumentar o disminuir la cantidad de recursos computacional a nuestras necesidades.

- Pagamos por el uso.
- Variedad de instancias
- Facilidad de administración
- Seguridad (Security groups, IAM)

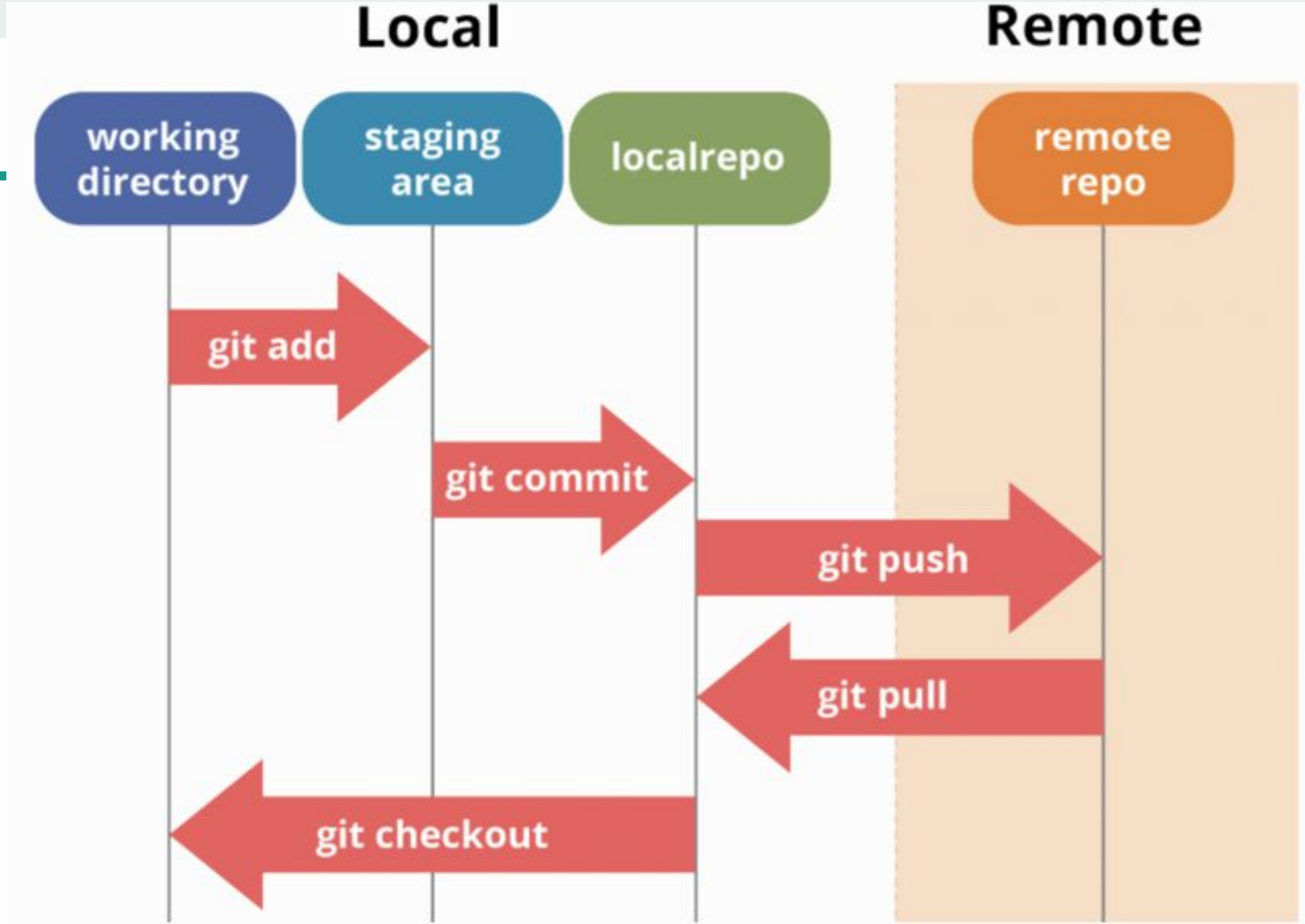
Usos:

- Alojamiento de aplicaciones web
- Análisis de datos
- Desarrollo de pruebas
- Almacenamiento de datos y backup.

Buckets s3

Es un servicio para almacenar cualquier tipo de datos, AWS te cobra por la cantidad de objetos almacenados.

- El nombre del bucket debe de ser único.
- Dependiendo del servicio por e.g SageMaker tendrás un bucket creado por default para almacenar todo lo de tu sesión.
- Usa la estrategia key-value.
- El key se refiere a el prefix y el object name.
-


Local
Remote
working directory
staging area
localrepo
remote repo
git add
git commit
git push
git pull
git checkout

Branches

1.

Stable Branches

1. `master`
2. `develop`

2.

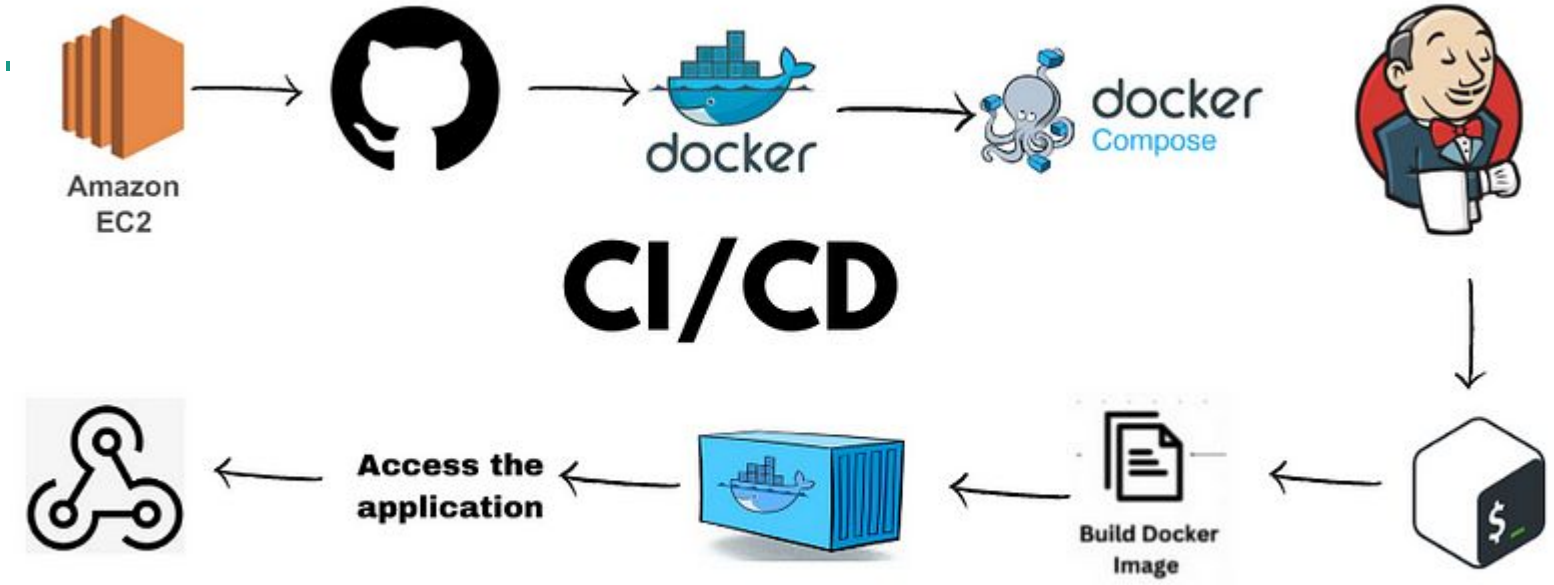
Releasable Branches

1. `master`
2. `hotfix/.*`

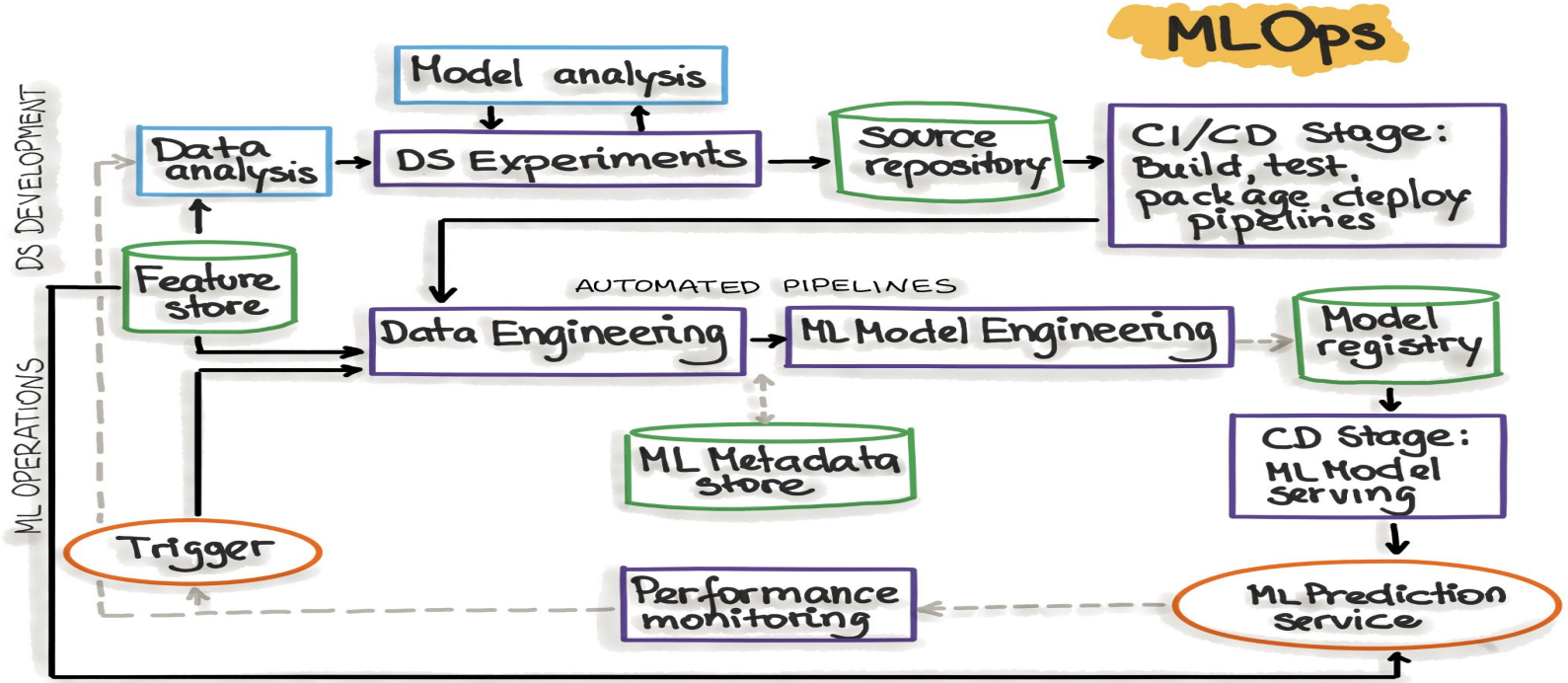
3.

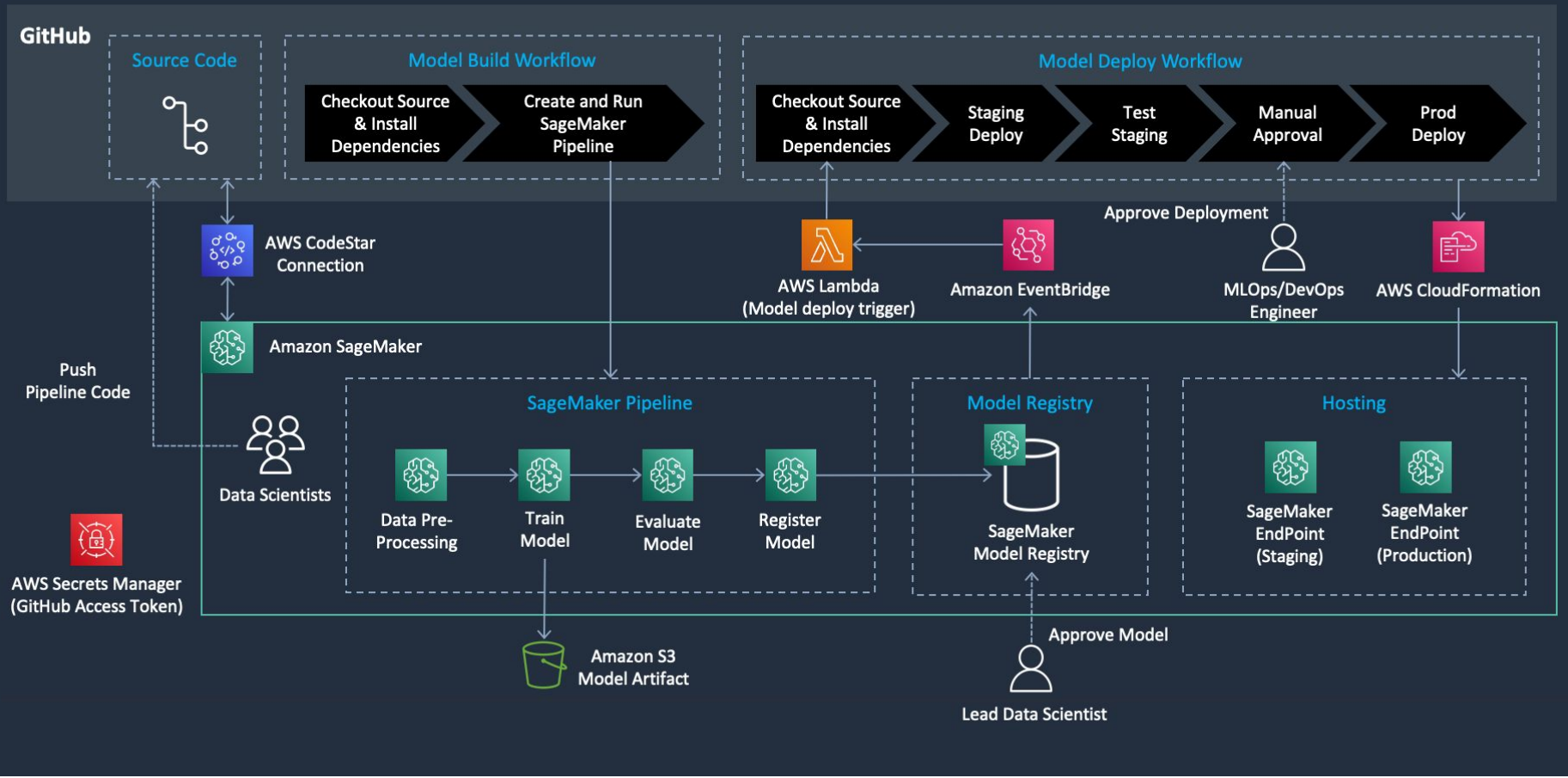
Feature Branches

1. `feature/.*`
2. `enhancement/.*`
3. `fix/.*`
4. `release/.*`
5. `backport/.*`
6. …


Amazon
EC2
docker
docker
Compose
CI/CD
Build Docker
Image
Access the
application

Introducción al MLOps



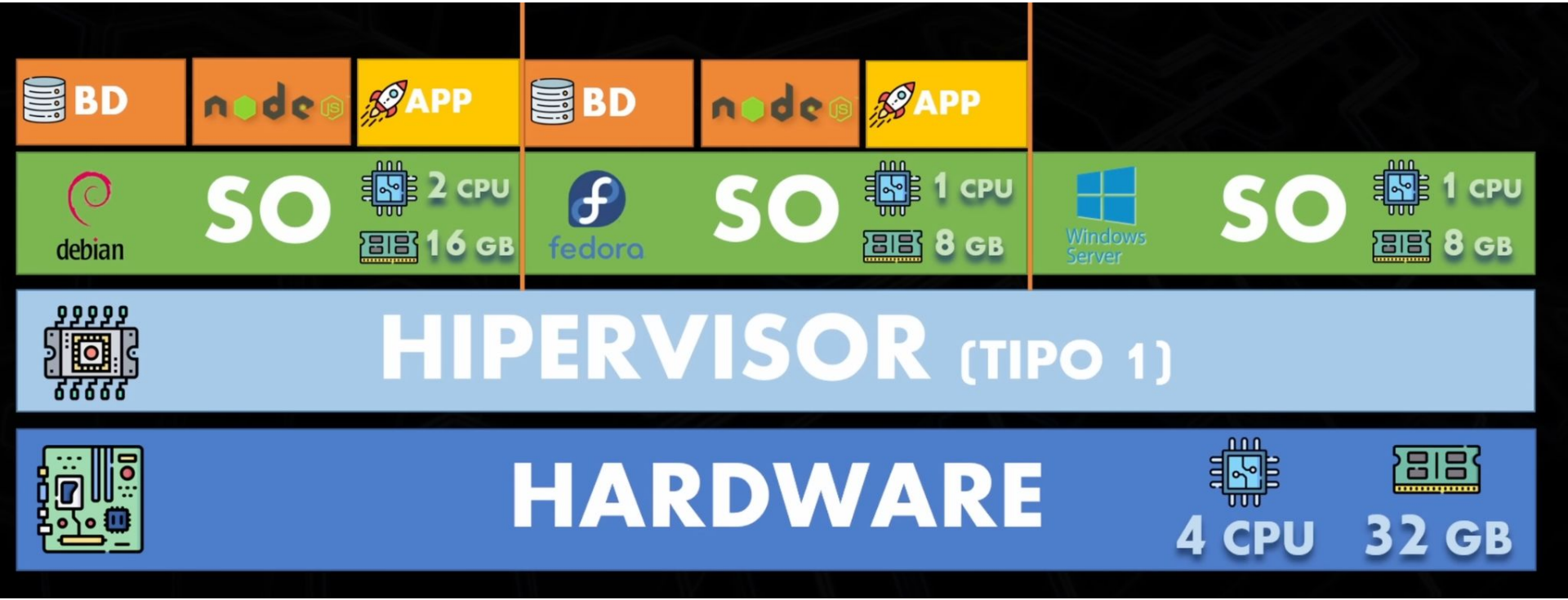

GitHub
Source Code
Model Build Workflow
Checkout Source & Install Dependencies
Create and Run SageMaker Pipeline
Model Deploy Workflow
Checkout Source & Install Dependencies
Staging Deploy
Test Staging
Manual Approval
Prod Deploy
AWS CodeStar Connection
AWS Lambda (Model deploy trigger)
Amazon EventBridge
Approve Deployment
MLOps/DevOps Engineer
AWS CloudFormation
Amazon SageMaker
Push Pipeline Code
Data Scientists
SageMaker Pipeline
Data Pre-Processing
Train Model
Evaluate Model
Register Model
Model Registry
SageMaker Model Registry
Hosting
SageMaker EndPoint (Staging)
SageMaker EndPoint (Production)
AWS Secrets Manager (GitHub Access Token)
Amazon S3 Model Artifact
Approve Model
Lead Data Scientist

Introducción a Dockers

https://www.youtube.com/watch?v=9eTVZwMZJsA&ab_channel=RingaTech

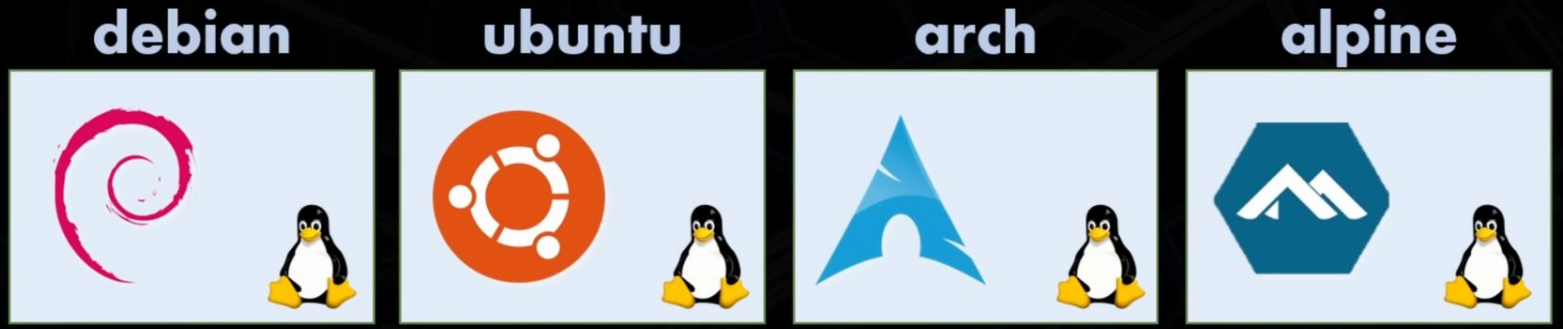
Aplicación 1

No comparten sistema de archivos ni microprocesos, están aislados entre sí.



Hipervisor virtualiza el hardware, haciendo que las máquinas virtuales sean independientes entre sí. Los docker virtualizan a nivel operativo.

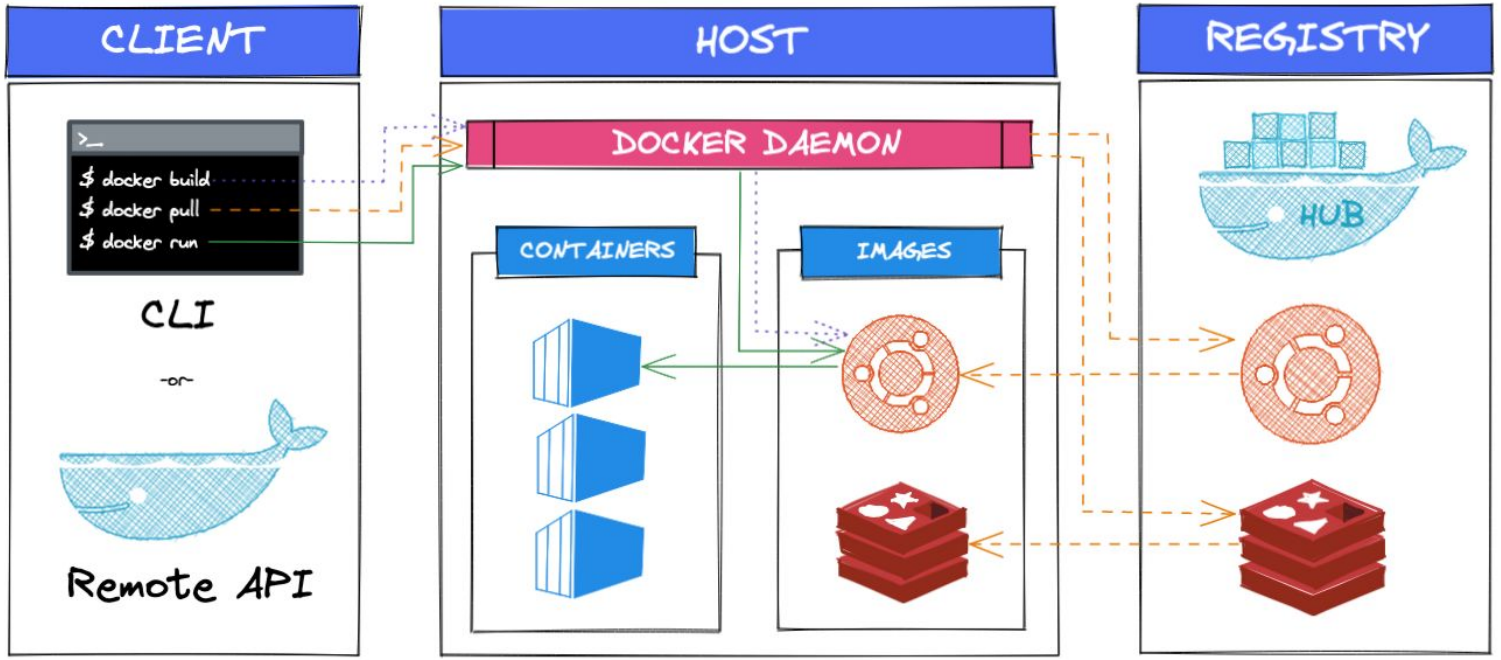
Distribuciones de Linux



Un driver o controlador es un software que permite que un dispositivo de hardware se comunique con el sistema operativo de una computadora.

ARDUINO™

Docker



Docker Daemon es un servicio que se encarga de crear, administrar y ejecutar contenedores de Docker.

SageMaker

Amazon SageMaker AI es un servicio de aprendizaje automático (ML) completamente gestionado. Con SageMaker AI, los científicos de datos y desarrolladores pueden crear, entrenar e implementar modelos de ML con rapidez y seguridad en un entorno alojado listo para producción. Ofrece una interfaz de usuario para ejecutar flujos de trabajo de ML que permite que las herramientas de ML de SageMaker AI estén disponibles en múltiples entornos de desarrollo integrados (IDE).