
aws
INNOVATE
AI/ML EDITION

aws INNOVATE
AI/ML EDITION

Build enterprise scale ML workflows on Kubernetes and Amazon SageMaker with Kubeflow

KJ Pittl
ISV Solution Architect
Amazon Web Services

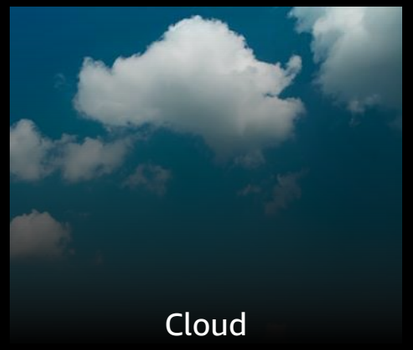
Agenda

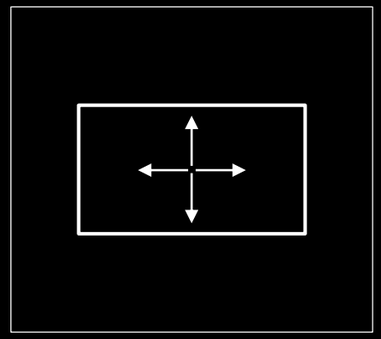
- Why machine learning with containers on Kubernetes
- Scaling ML on Kubernetes with Amazon SageMaker
- Kubeflow and Kubeflow pipelines
- Common integration patterns with AWS services
- Demo

aws INNOVATE
AI/ML EDITION

Why machine learning with containers / on Kubernetes

Why machine learning on Kubernetes



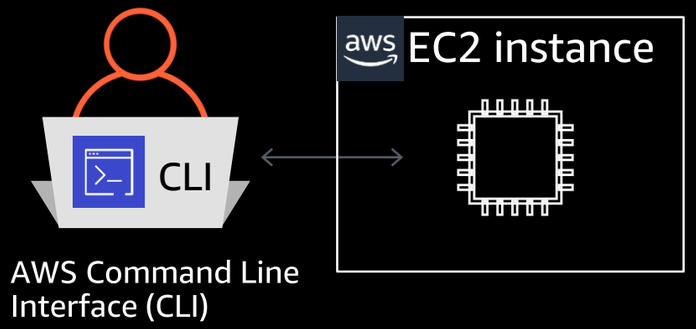


Composability

Portability

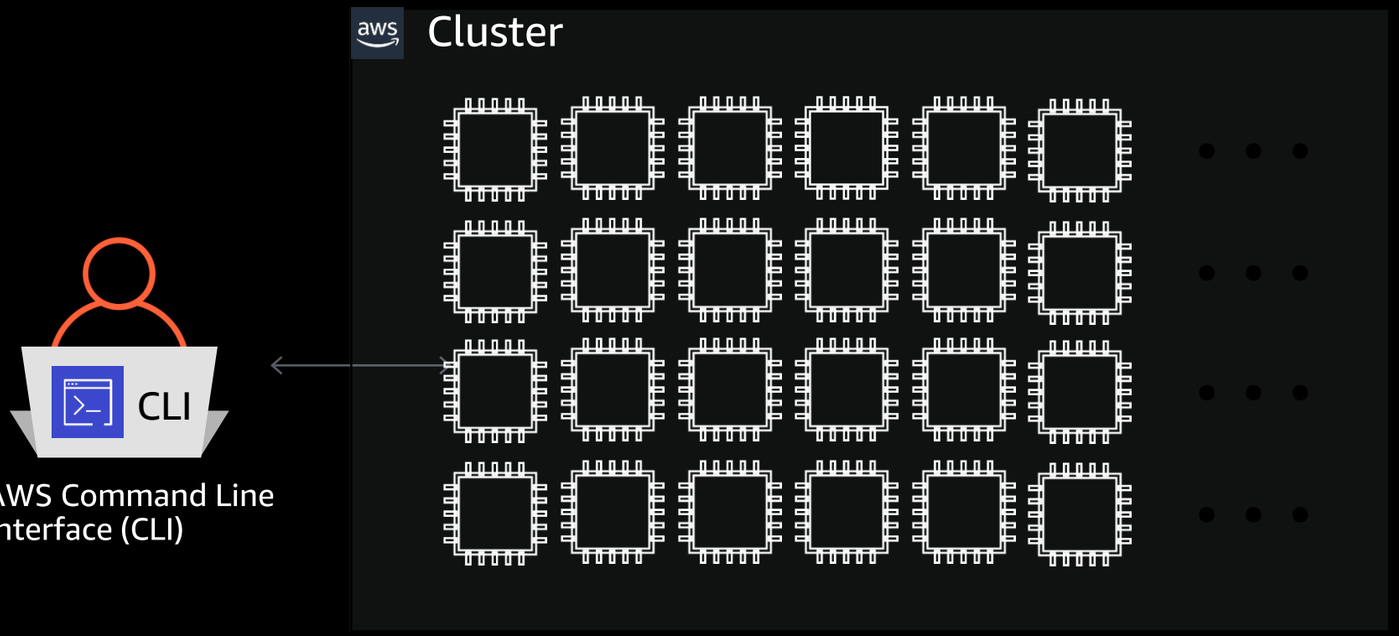
Scalability

Challenges with scaling machine learning



AWS Command Line Interface (CLI)

- Compute (CPUs, GPUs)
- Storage
- Source control
- ML Frameworks



AWS Command Line Interface (CLI)

- Scaling
- Managing infrastructure
- Managing pipelines

Why machine learning with containers

Why machine learning with containers

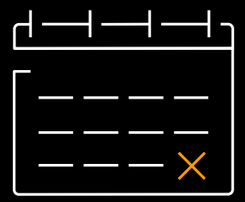
Packages

- Training code
- Dependencies
- Configurations

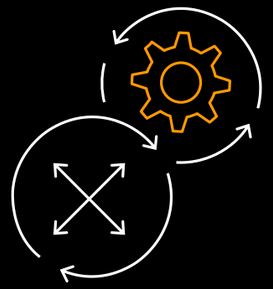
ML environments that are

- Lightweight
- Portable
- Scalable
- Consistent

Challenges of building container images for ML



Takes days to test and configure

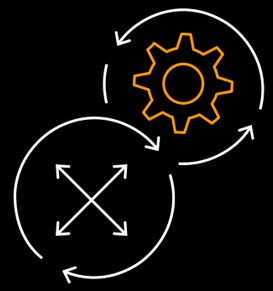


Must optimize for performance & scale

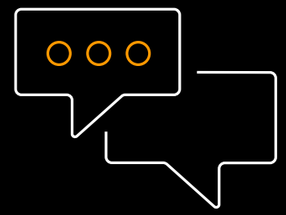
Rebuild and re-optimize new framework versions

AWS deep learning containers

Pre-packaged Docker container images fully configured and validated



Best performance and scalability without tuning



Works with Amazon EKS, Amazon ECS, and Amazon EC2

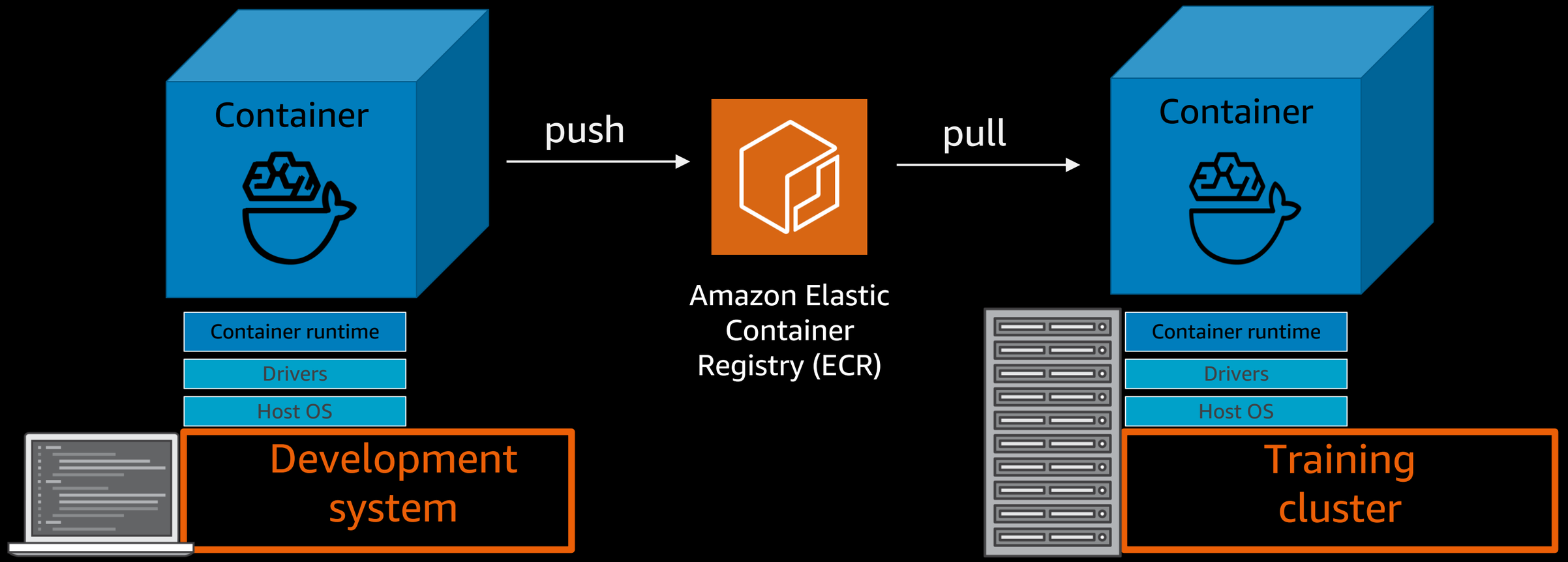
KEY FEATURES

Customizable container images

Support for TensorFlow, Apache MXNet, PyTorch

Single and multi-node training and inference

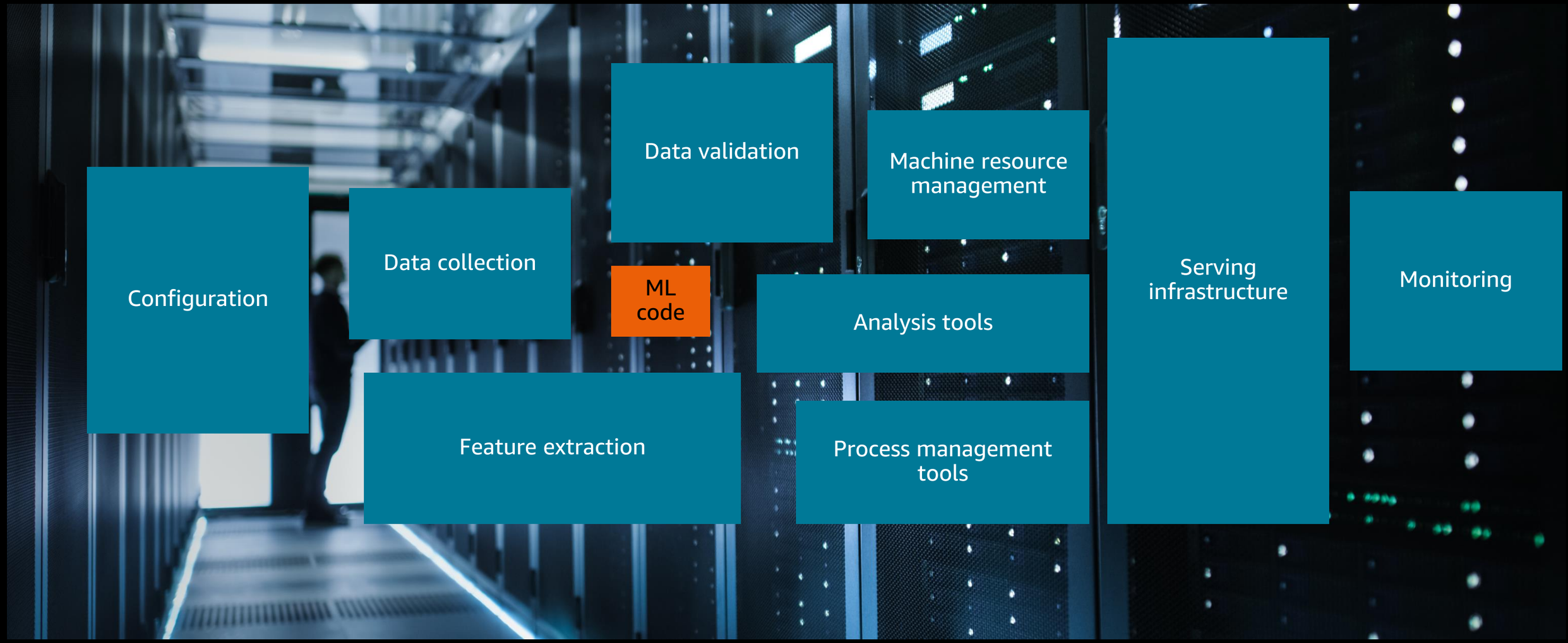
Why ML with containers?



aws INNOVATE
AI/ML EDITION

Scaling machine learning on Kubernetes with Amazon SageMaker

ML is hard



Source: Sculley, et al. "Hidden Technical Debt in Machine Learning Systems," NIPS'15: *Proceedings of the 28th International Conference on Neural Information Processing Systems.* Volume 2 December 2015, 2503–2511

Pain points of self-managed ML Platforms

The following are all obstacles to the core goal of building best-in-class models that solve business problems

- Setting up k8s without prior experience is challenging
- Configuring proper scaling of compute has a learning curve
- Right sizing instances for cost-efficiency is hard
- ML tasks need additional configuration to use GPU or CPU nodes optimally
- Libraries and toolkits need to be regularly updated, which increases technical debt that later needs to be paid off
- Additional management burden for the ops team

The AWS ML Stack

Broadest and most complete set of machine learning capabilities

AI SERVICES

HEALTH AI

- NEW Amazon HealthLake
- Amazon Transcribe Medical
- Amazon Comprehend Medical

INDUSTRIAL AI

- NEW AWS Panorama + Appliance
- NEW Amazon Monitron
- NEW Amazon Lookout for Equipment
- NEW Amazon Lookout for Vision

ANOMALY DETECTION

- NEW Amazon Lookout for Metrics

CODE AND DEVOPS

- NEW Amazon DevOps Guru
- Amazon CodeGuru

VISION

- Amazon Rekognition

SPEECH

- Amazon Polly
- Amazon Transcribe *+Medical*

TEXT

- Amazon Comprehend *+Medical*
- Amazon Translate
- Amazon Textract

SEARCH

- Amazon Kendra

CHATBOTS

- Amazon Lex

PERSONALIZATION

- Amazon Personalize

FORECASTING

- Amazon Forecast

FRAUD

- Amazon Fraud Detector

CONTACT CENTERS

- Contact Lens
- Voice ID *For Amazon Connect*

ML SERVICES

Amazon SageMaker

Label data

SAGEMAKER STUDIO IDE

- NEW Aggregate & prepare data
- NEW Store & share features
- Auto ML
- Spark/R
- NEW Detect bias
- Visualize in notebooks
- Pick algorithm
- Train models
- Tune parameters
- NEW Debug & profile
- Deploy in production
- Manage & monitor
- NEW CI/CD

Human review

NEW: SageMaker JumpStart

NEW: Model management for edge devices

FRAMEWORKS & INFRASTRUCTURE

TensorFlow, mxnet, PyTorch, GLUON, Keras, scikit-learn, Horovod, DeepGraphLibrary

Deep Learning AMIs & Containers

GPUs & CPUs

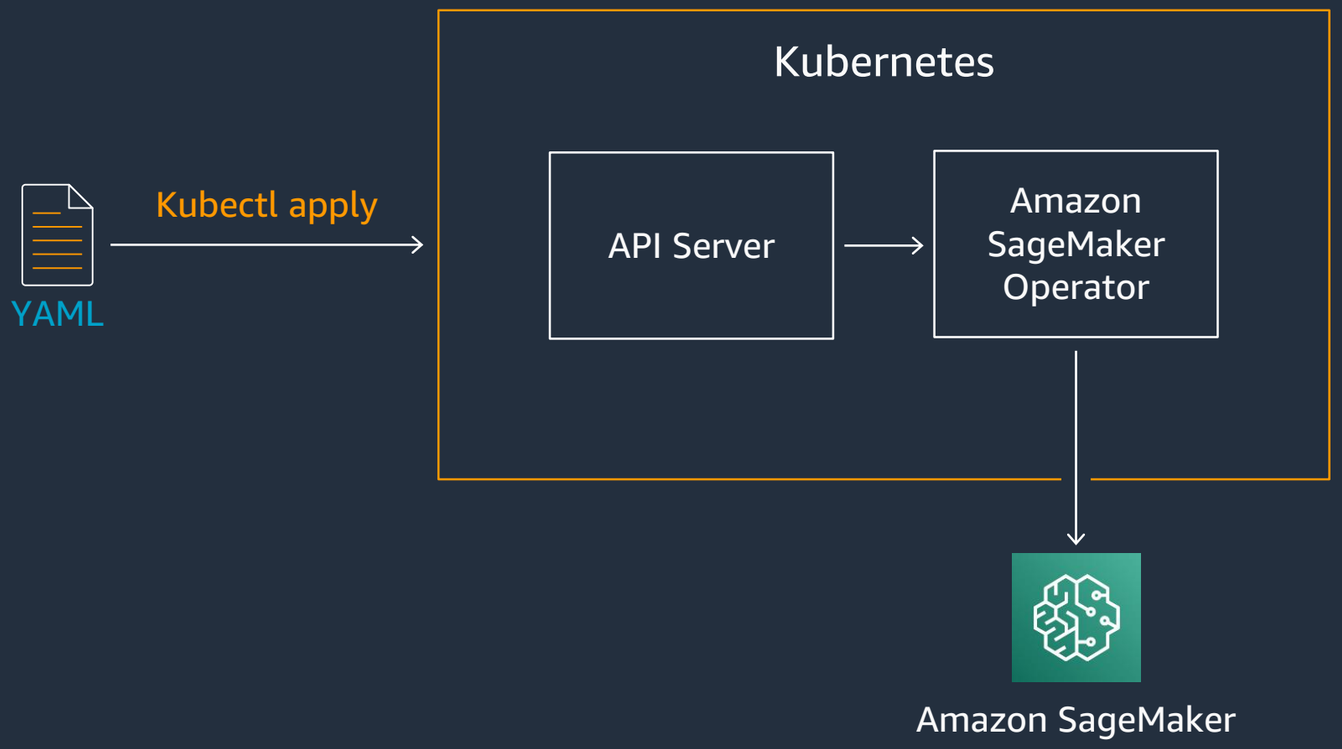
Elastic Inference

Trainium

Inferentia

FPGA

SageMaker Operators



KEY FEATURES

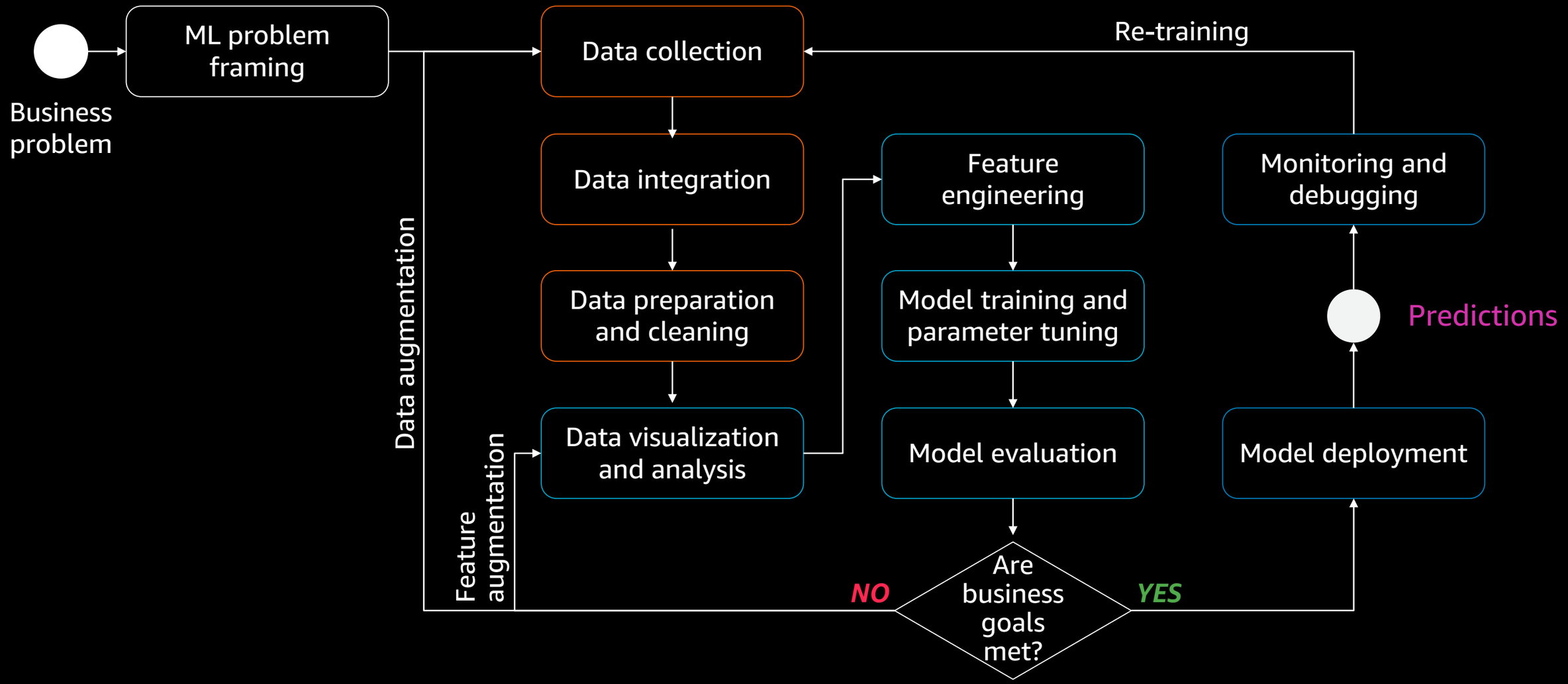
Amazon SageMaker Operators for training, tuning, inference, ground truth data labeling

Natively interact with Amazon SageMaker jobs using Kubernetes tools (e.g., get pods, describe)

Stream and view logs from Amazon SageMaker in Kubernetes

Helm Charts to assist with setup and spec creation

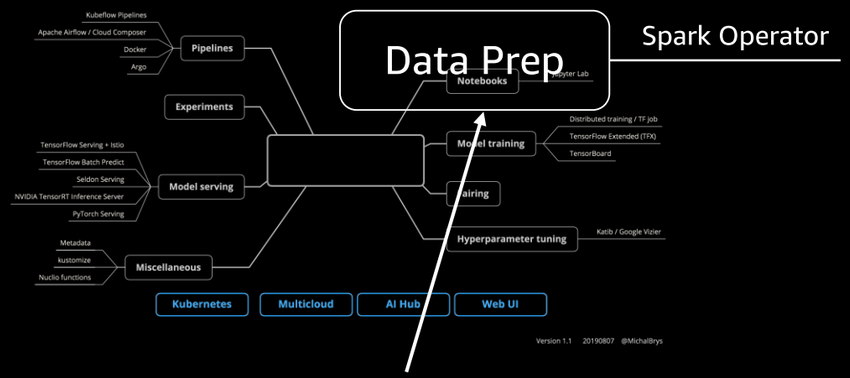
ML cycle



aws INNOVATE
AI/ML EDITION

Kubeflow and Kubeflow Pipelines

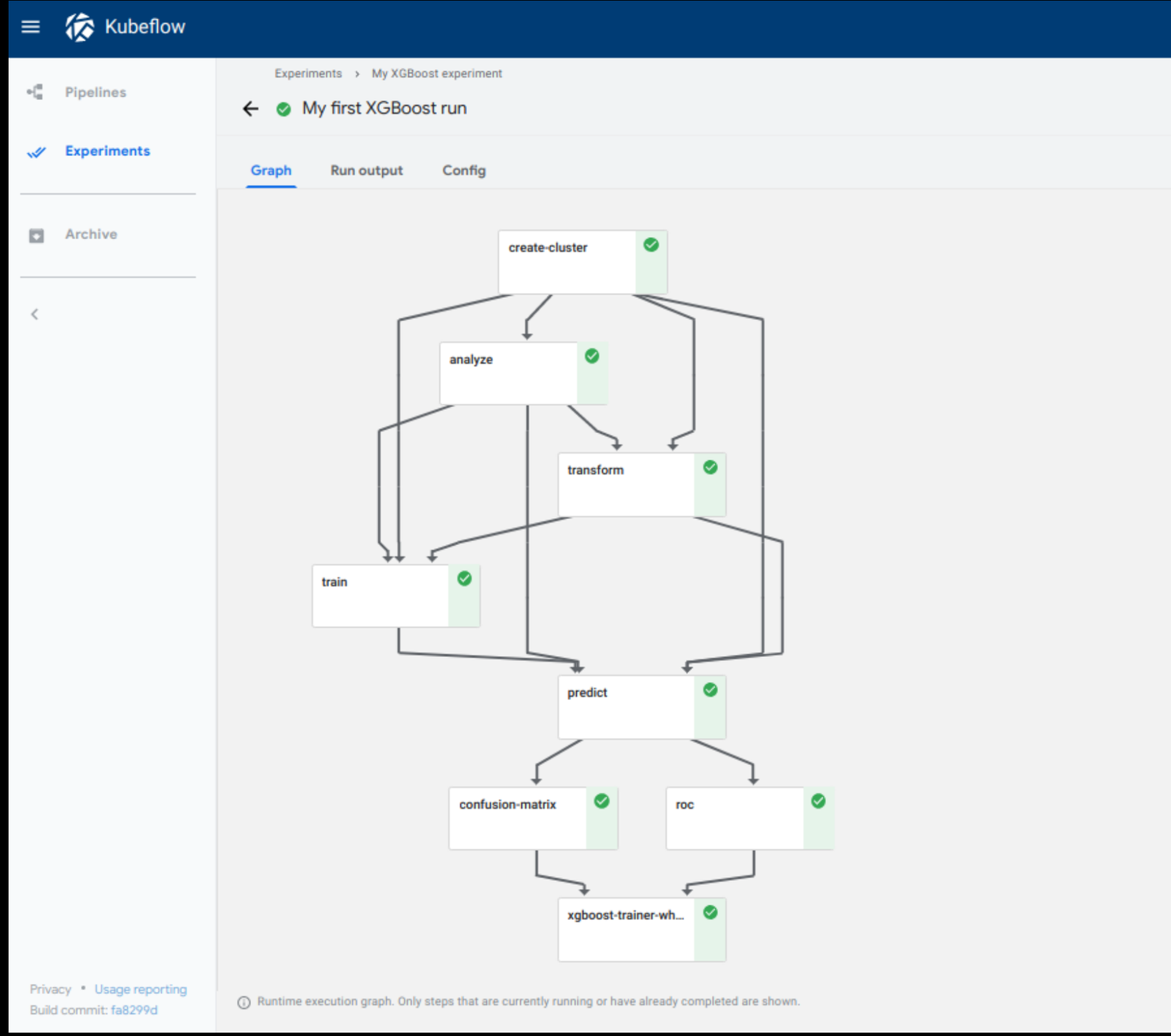
Kubeflow Components



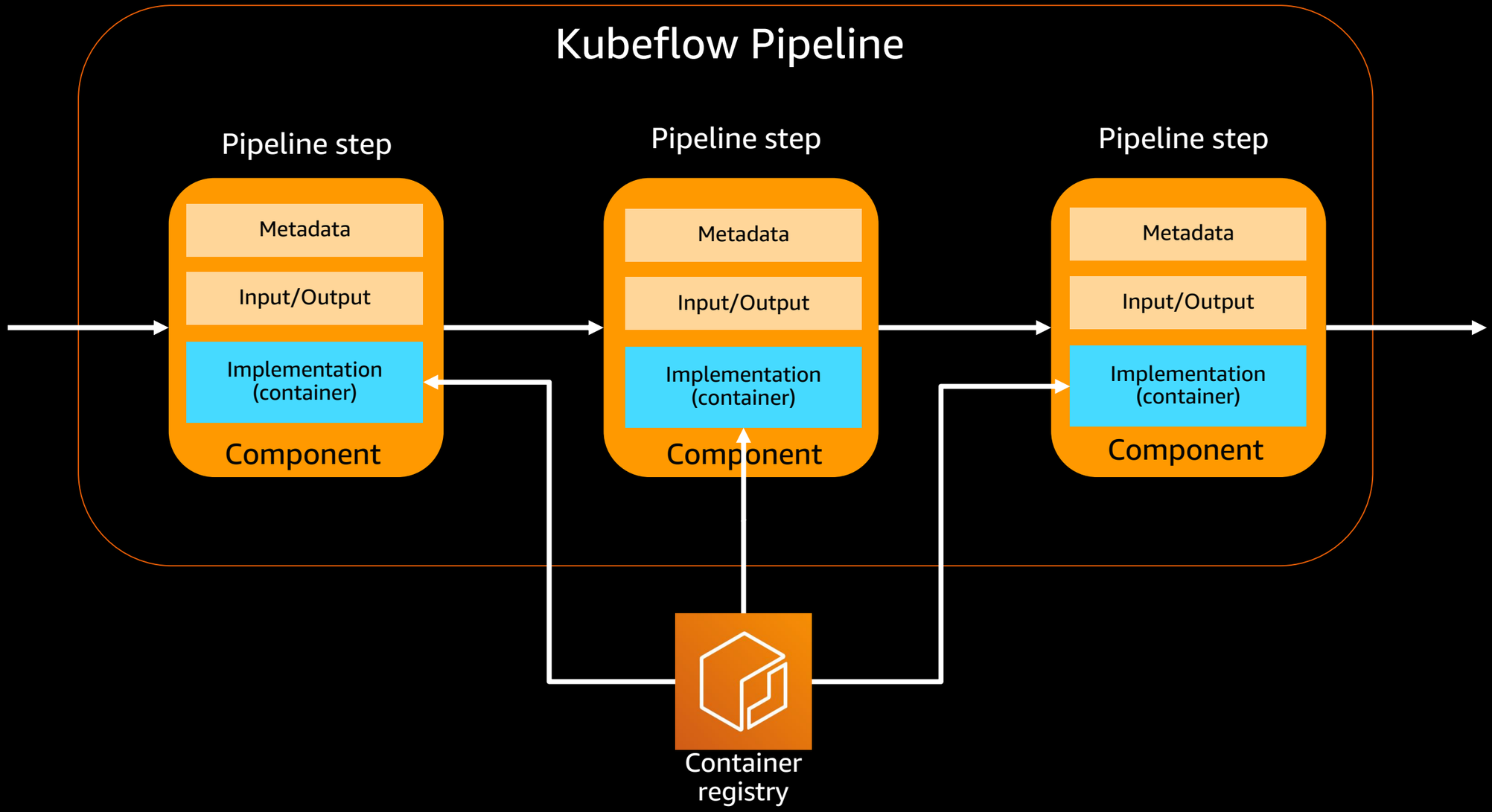
Kubeflow

Kubeflow Pipelines

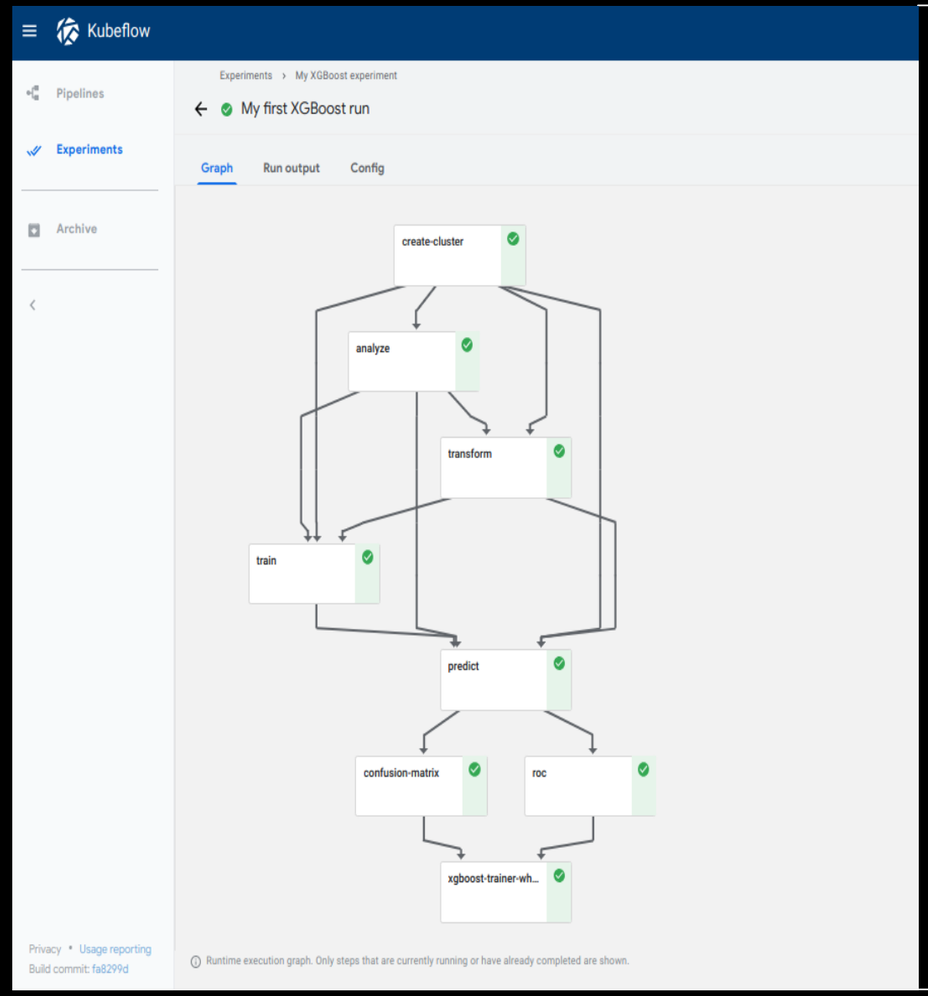
- A user interface (UI) for managing and tracking experiments, jobs, and runs
- An engine for scheduling multi-step ML workflows
- A software development kit (SDK) for defining and manipulating pipelines and components



Kubeflow pipelines



Creating a pipeline (1)



Pipeline decorator

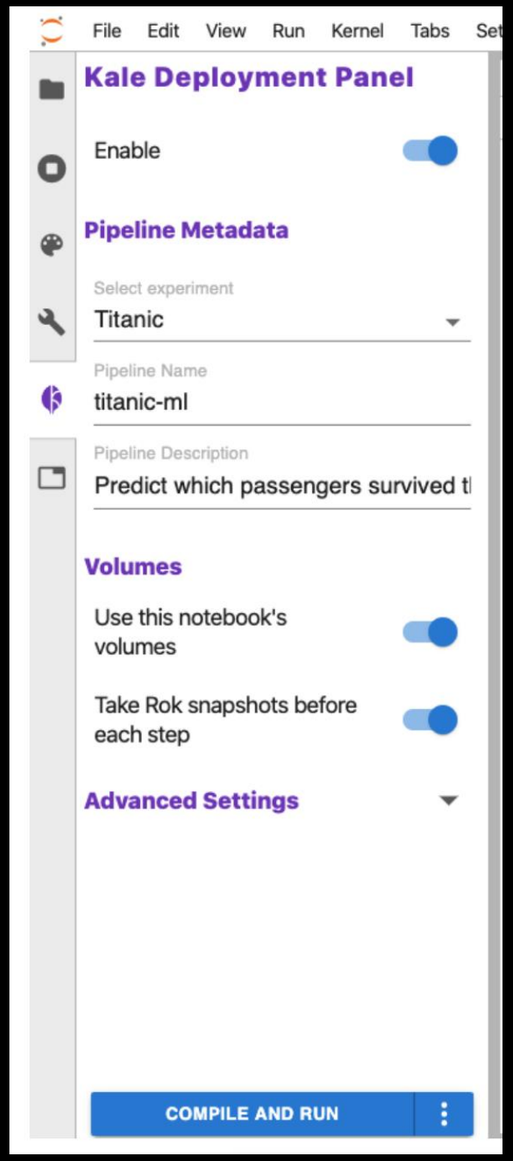
Pipeline function

Pipeline component

```
@dsl.pipeline(
  name='Sample Trainer',
  description=''
)

def sample_train_pipeline(... ):

    create_cluster_op = CreateClusterOp('create-cluster', ...)

    analyze_op = AnalyzeOp('analyze', ...)

    transform_op = TransformOp('transform', ...)

    train_op = TrainerOp('train', training_container,...)

    predict_op = PredictOp('predict', ...)

    confusion_matrix_op = ConfusionMatrixOp('confusion-matrix', ...)

    roc_op = RocOp('roc', ...)

kfp.compiler.Compiler().compile(sample_train_pipeline , 'my-pipeline.zip')
```

Compile pipeline

Creating a Kubeflow pipeline with Kale



- simple UI via Kale JupyterLab extension to define Kubeflow Pipelines

pipeline-parameters

```
[3]: nodes_number = 256
     learning_rate = 0.0001
```

step: loaddata

Cell type: Pipeline Step | Step name: loaddata | Depends on

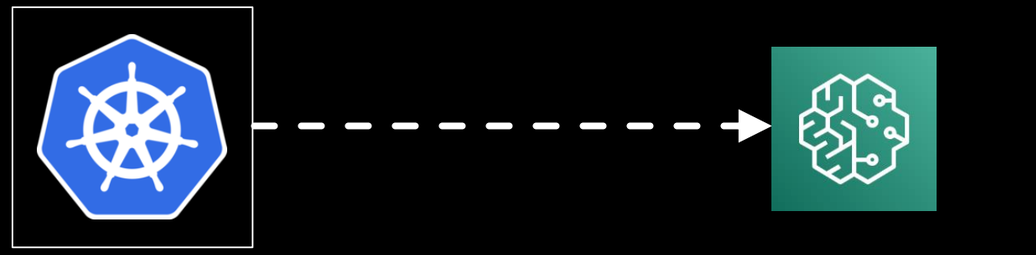
```
path = "data/"

PREDICTION_LABEL = 'Survived'

test_df = pd.read_csv(path + "test.csv")
train_df = pd.read_csv(path + "train.csv")
```

pipeline-metrics

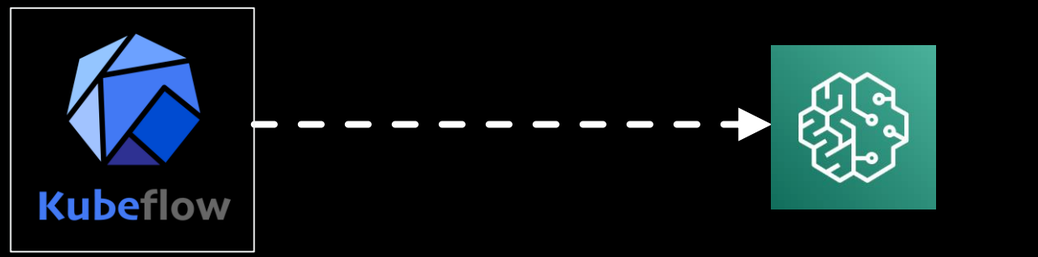
```
[ ]: print(test_accuracy_resnet)
```

Kubeflow + Amazon SageMaker

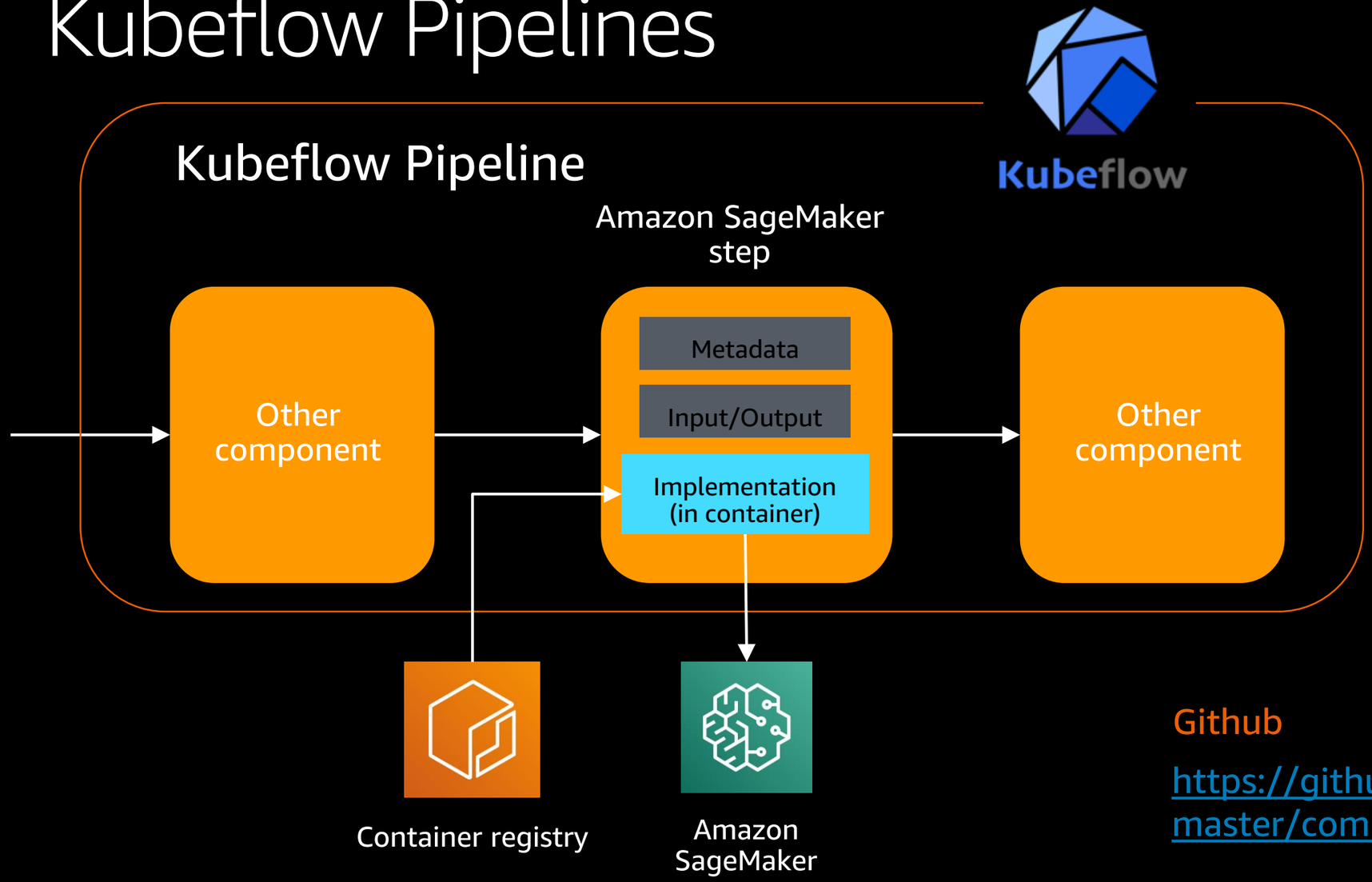
- Amazon SageMaker operators for Kubernetes



- Amazon SageMaker components for Kubeflow pipelines



Amazon SageMaker components for Kubeflow Pipelines



Supported components

- Model training
- Hyperparameter tuning
- Processing
- Model deployment
- Batch transform
- Amazon SageMaker Ground Truth

Github

https://github.com/kubeflow/pipelines/tree/master/components/aws/sagemaker

aws INNOVATE
AI/ML EDITION

Common integration patterns with AWS services

Leverage AWS innovations through Kubeflow

Do-it-yourself

Managed service

Blog Post

Enterprise-ready Kubeflow: Securing and scaling AI and ML

https://aws.amazon.com/blogs/opensource/enterprise-ready-kubeflow-securing-and-scaling-ai-and-machine-learning-pipelines-with-aws/

aws INNOVATE
AI/ML EDITION

Demo

Kubeflow + Amazon SageMaker delivers value

- Portability
- Composability
- Scalability
- Shared infrastructure
- Repeatable pipelines
- Automation
- CI / CD
- Open source

Amazon SageMaker

- Fully-managed infrastructure
- Ground truth labeling
- Automatic model tuning
- Built-in optimized algorithms
- Managed Spot Training
- Scalable inference endpoints
- Model monitoring

Get started

Demo code

https://github.com/data-science-on-aws/workshop

Blog post – Enterprise-ready Kubeflow

https://aws.amazon.com/blogs/opensource/enterprise-ready-kubeflow-securing-and-scaling-ai-and-machine-learning-pipelines-with-aws/

Kubeflow on AWS

https://www.kubeflow.org/docs/aws/

SageMaker components for Kubeflow Pipelines

http://github.com/kubeflow/pipelines/tree/master/components/aws/sagemaker

Online workshop

https://www.eksworkshop.com/advanced/420_kubeflow/pipelines/

Sample Jupyter notebooks: https://github.com/aws-samples/eks-kubeflow-workshop

Visit the AI and Machine Learning Resource Hub for more resources

Dive deeper with these resources, get inspired and learn how you can use machine learning to accelerate business outcomes.

- The machine learning journey e-book
- Machine learning enterprise guide
- 7 leading machine learning use cases e-book
- A strategic playbook for data, analytics, and machine learning
- Accelerating ML innovation through security e-book
- … and more!

Visit resource hub »

https://tinyurl.com/aiml-aws

AWS Machine Learning (ML) Training and Certification

Learn like an Amazonian, based on the curriculum we've used to train our own developers and data scientists

AWS is how you build machine learning skills

Courses built on the curriculum leveraged by Amazon's own teams. Learn from the experts at AWS.

Flexibility to learn your way

Learn online with 65+ on-demand digital courses or live with virtual instructor-led training, plus hands-on labs and opportunities for practical application.

Validate your expertise

Demonstrate expertise in building, training, tuning, and deploying machine learning models with an industry-recognized credential.

aws.training/machinelearning

Thank You for Attending AWS Innovate

We hope you found it interesting! A kind reminder to **complete the survey.**
Let us know what you thought of today's event and how we can improve the event experience for you in the future.


aws-apac-marketing@amazon.com


twitter.com/AWSCloud


facebook.com/AmazonWebServices


youtube.com/user/AmazonWebServices


slideshare.net/AmazonWebServices


twitch.tv/aws

aws INNOVATE
AI/ML EDITION

Thank you!