# Amazon SageMaker Bring Your Own Script/Container

Michael Lin

Sr. Solutions Architect
Amazon Web Services

Immersion Day



aws machine learning

# Model Options

## In this session: Bring Your Own Script/Container

AWS Marketplace for Machine Learning

Amazon SageMaker AutoPilot

Training code

- XGBoost - Gradient Boosted Trees
- Matrix Factorization
- Regression
- Principal Component Analysis
- K-Means Clustering
- And More!

Built-in Algorithms (17)
No ML coding required

TensorFlow, Keras, mxnet, PyTorch, Horovod, Gluon, dmlc XGBoost, Apache Spark, Chainer, scikit learn

Bring Your Own Script
Amazon SageMaker builds the container
Open source containers

Bring Your Own Container
Full control, you build the container
R, C++, etc

# Fully Managed, Distributed, Auto-Scaled, Secured



aws

# Containers for Machine Learning

# AWS Deep Learning Containers

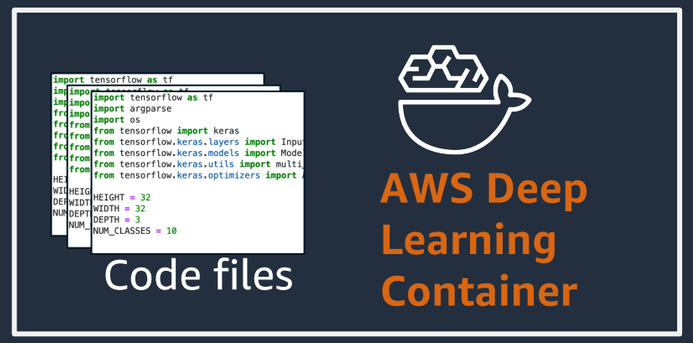
- Prepackaged Docker container images fully configured and validated
- Optimized for performance with latest NVIDIA driver, CUDA libraries, and Intel libraries
- Consistent and reproducible deployment and lightweight
- Optimized for distributed machine learning
- Runs on Amazon ECS, Amazon EKS and Amazon SageMaker



aws

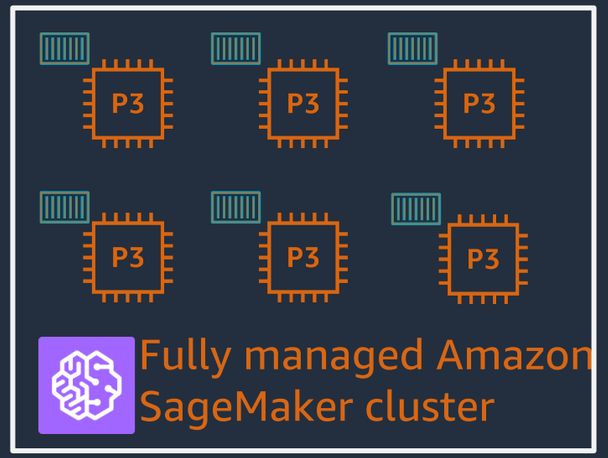
# Amazon SageMaker
# Bring your own Script

# High Level Workflow



```
from sagemaker.tensorflow import TensorFlow

cifar10_estimator = TensorFlow(entry_point='source/cifar10.py',
                               role=role,
                               framework_version='1.14',
                               train_instance_count=1,
                               train_instance_type='ml.p3.xlarge')
```







aws

# Bring Your Own Script

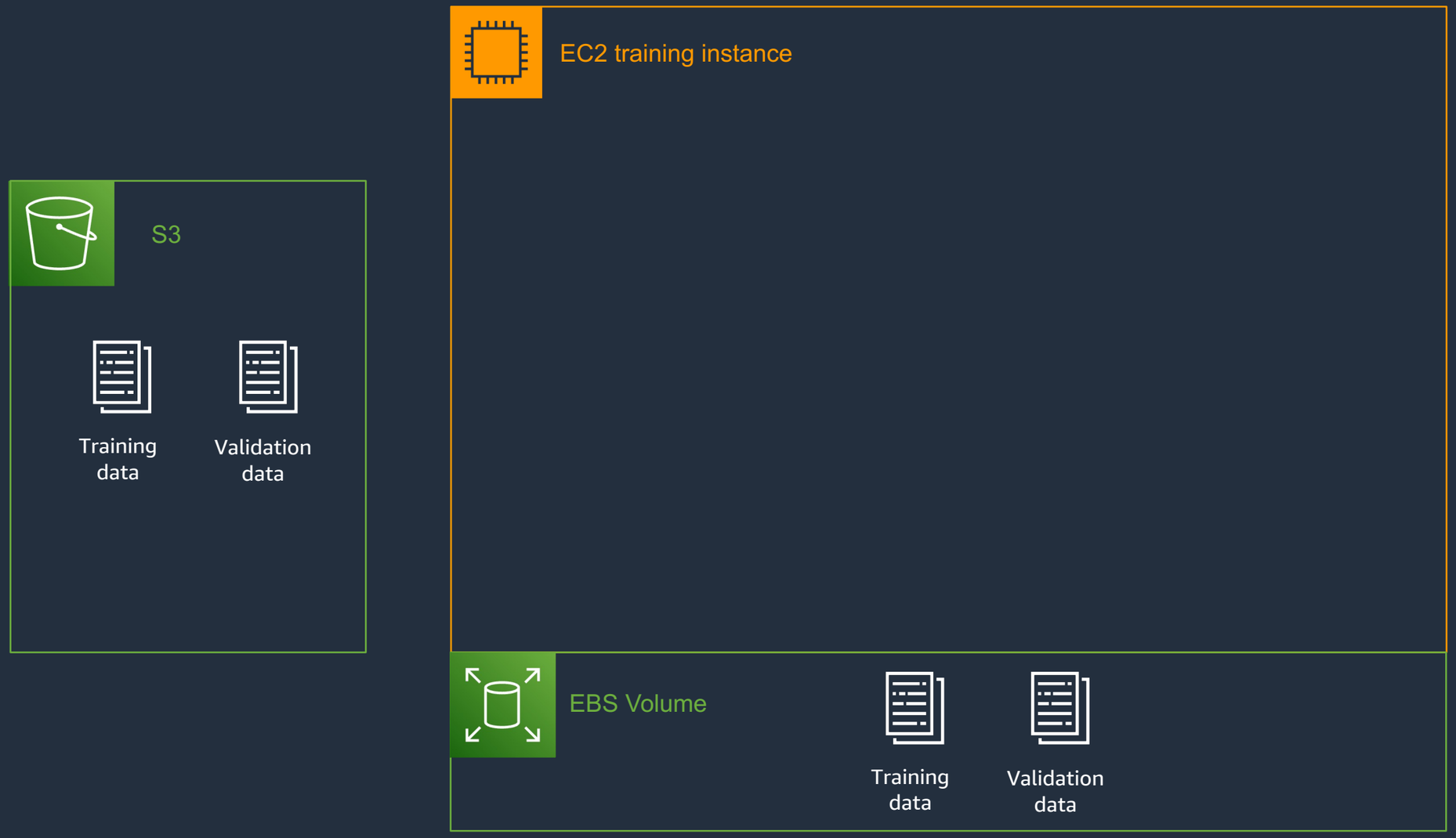
AWS
Managed

1. Point to the AWS-managed container of your choice
2. Write your model as a bundle of files
3. Specify the entry-point (script) in the SageMaker Estimator
4. Include any extra libraries with requirements.txt
5. Use our web-server for inference

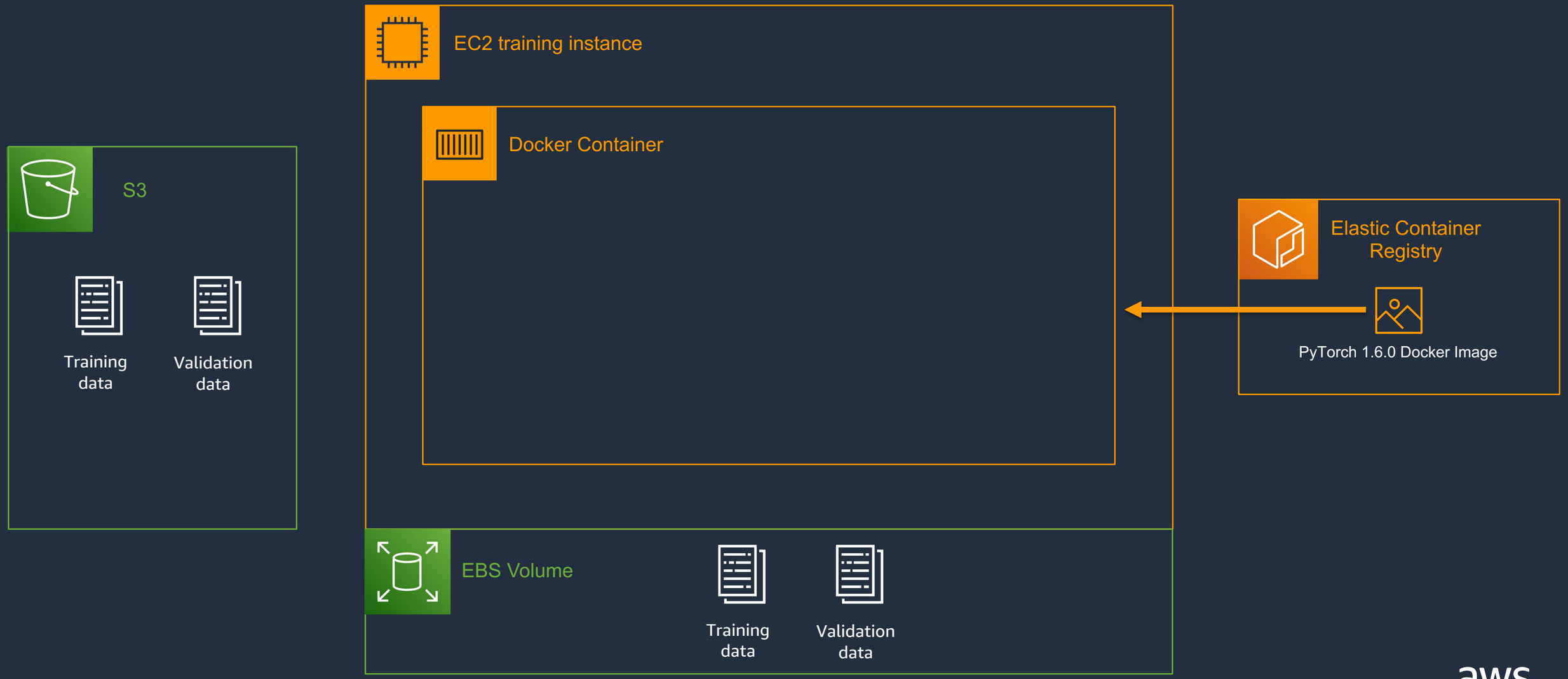


aws

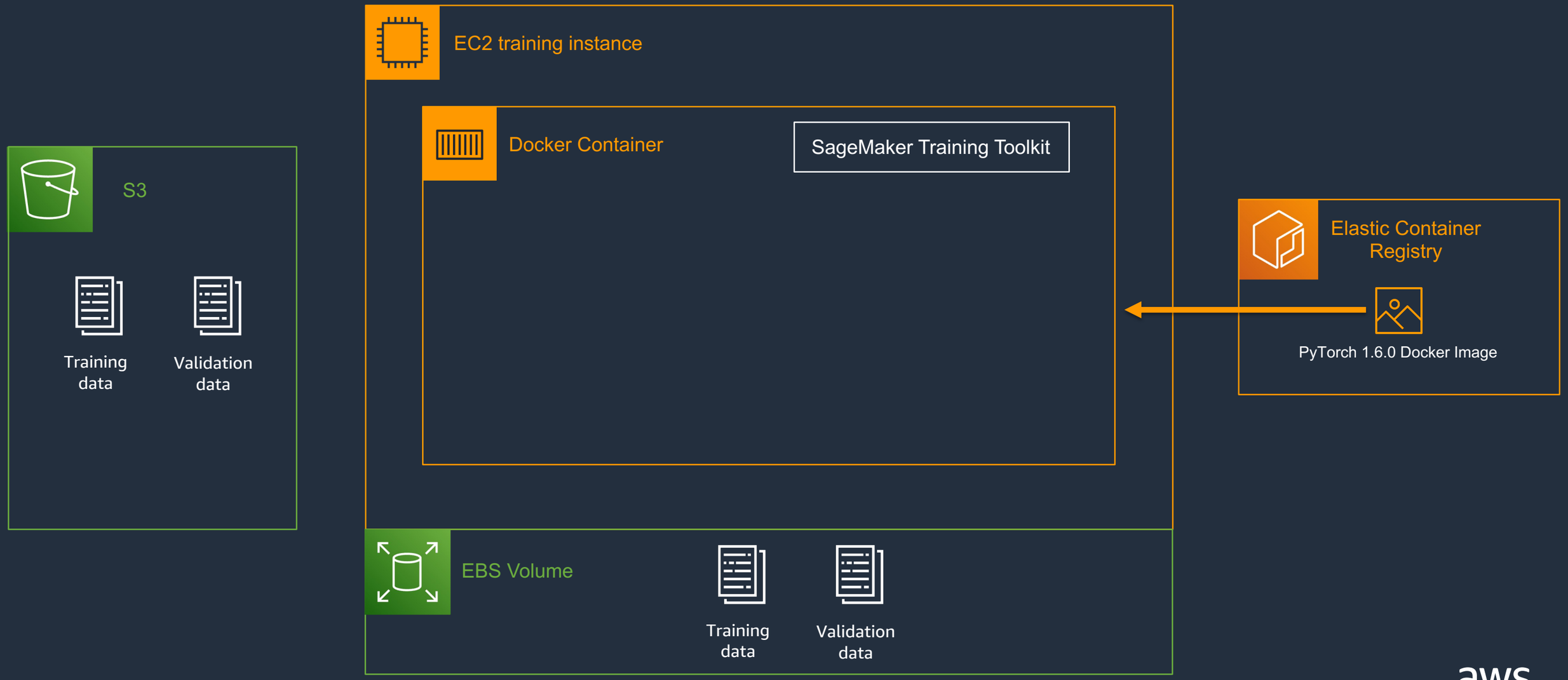
# Copy the data from S3 to EBS volume of the EC2 instance



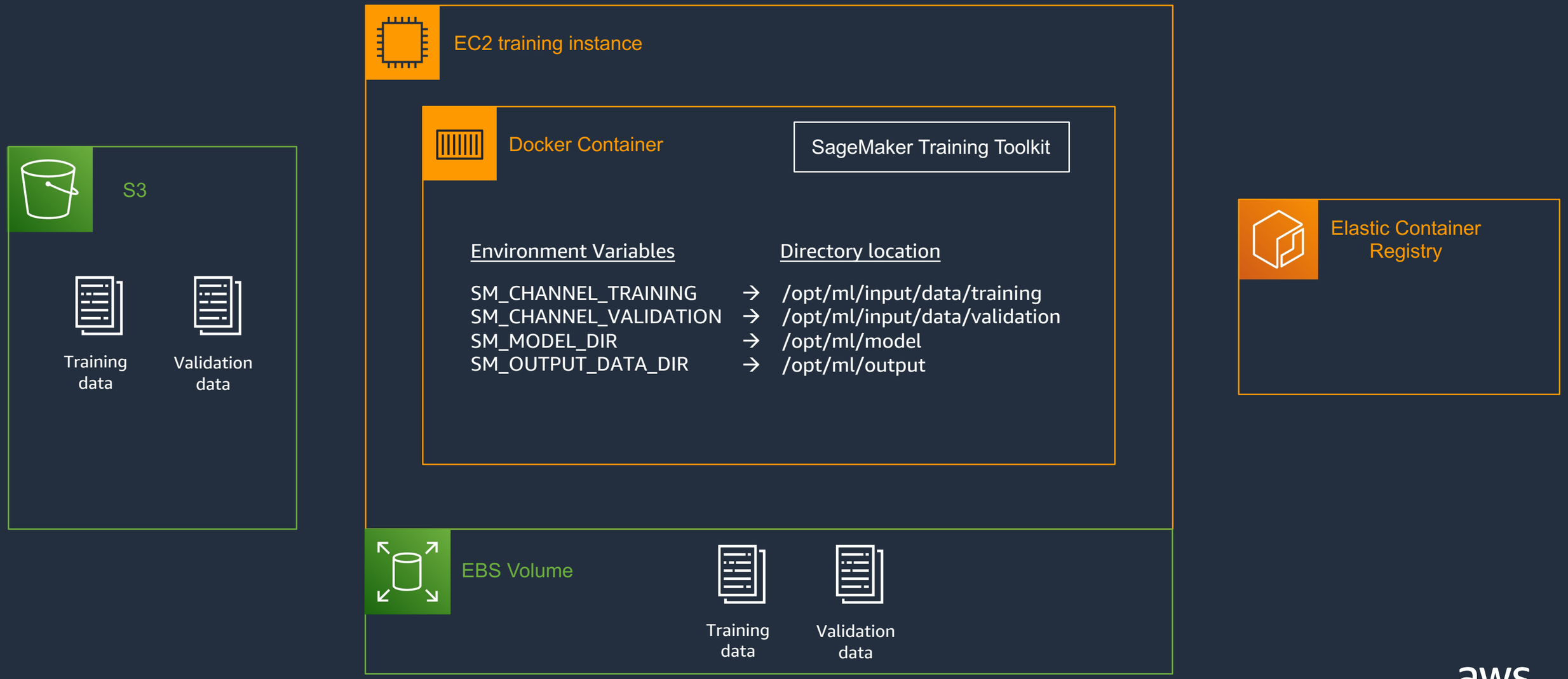
# Docker Container



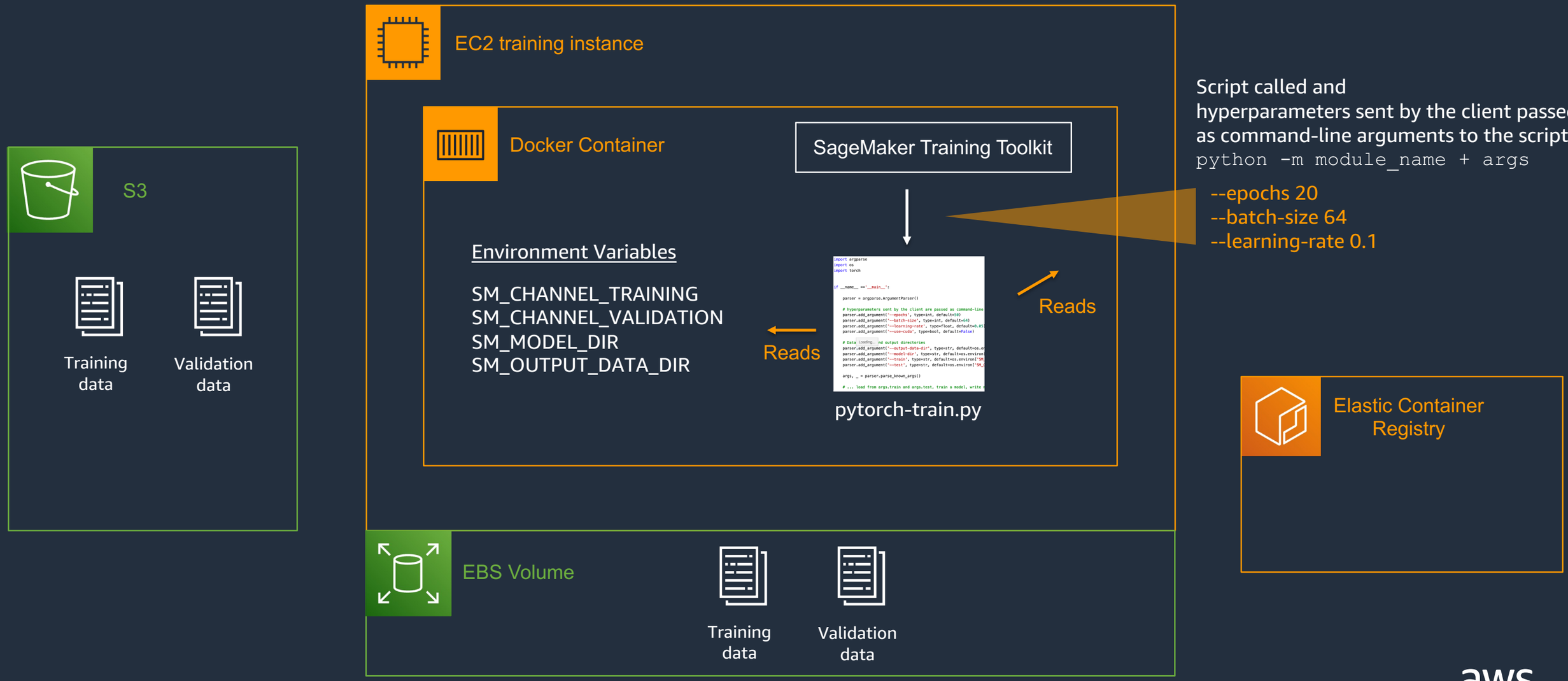
# SageMaker Training Toolkit



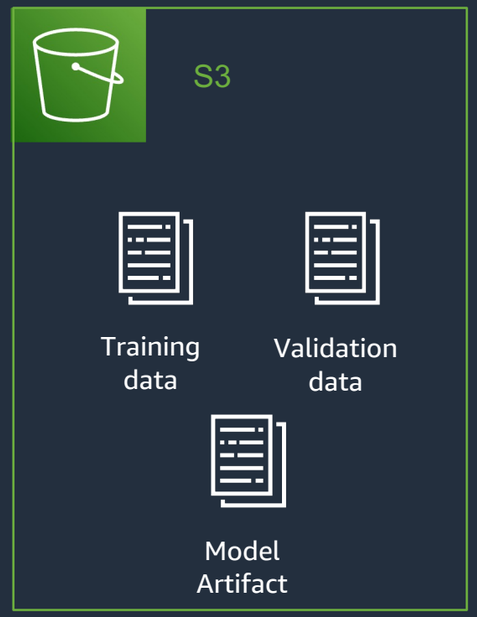
# Environment variables and directory location

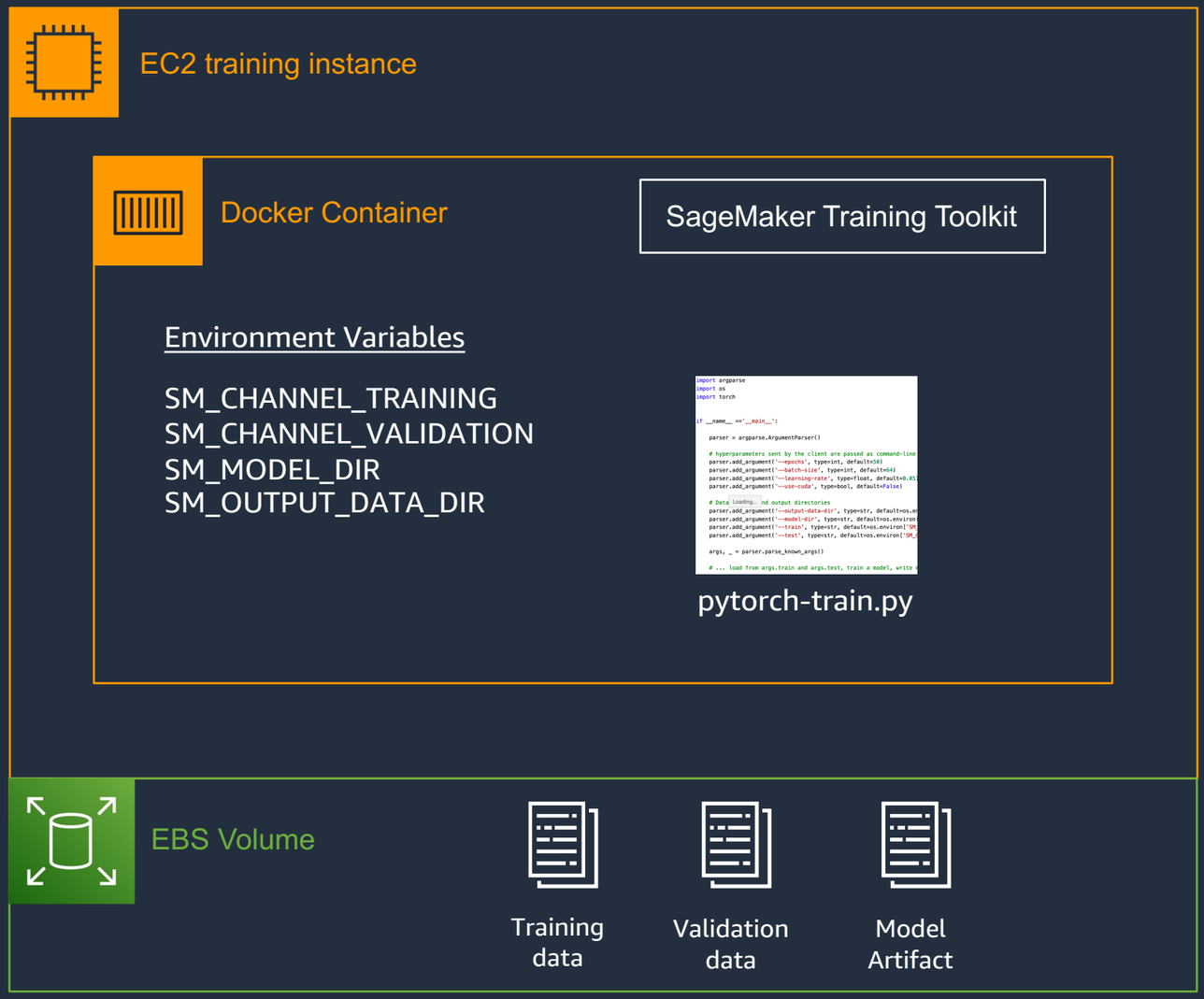


# Your script is called with arguments present in the Estimator



# Copy of the Model Artifact to S3 after training







aws

# Amazon SageMaker
# Bring your own Container

# Bring Your Own Docker File - R Example

```
FROM ubuntu:16.04

MAINTAINER Amazon SageMaker Examples <amazon-sagemaker-examples@amazon.com>

RUN apt-get -y update && apt-get install -y --no-install-recommends \
    wget \
    r-base \
    r-base-dev \
    ca-certificates

RUN R -e "install.packages(c('mda', 'plumber'), repos='https://cloud.r-project.org')"

COPY mars.R /opt/ml/mars.R
COPY plumber.R /opt/ml/plumber.R

ENTRYPOINT ["/usr/bin/Rscript", "/opt/ml/mars.R", "--no-save"]
```



aws

# Bring Your Own Docker File - Python Example

```
FROM ubuntu:18.04

MAINTAINER Amazon AI <sage-learner@amazon.com>


RUN apt-get -y update && apt-get install -y --no-install-recommends \
         wget \
         python3-pip \
         python3-setuptools \
         nginx \
         ca-certificates \
    && rm -rf /var/lib/apt/lists/*

RUN ln -s /usr/bin/python3 /usr/bin/python
RUN ln -s /usr/bin/pip3 /usr/bin/pip
```

```
RUN pip --no-cache-dir install numpy==1.16.2 scipy==1.2.1 scikit-learn==0.2
0.2 pandas flask gunicorn
```

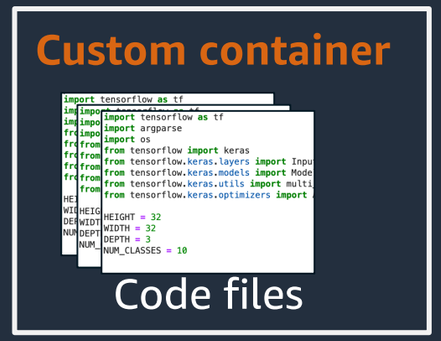
```
ENV PYTHONUNBUFFERED=TRUE
ENV PYTHONDONTWRITEBYTECODE=TRUE
ENV PATH="/opt/program:${PATH}"

# Set up the program in the image
COPY decision_trees /opt/program
WORKDIR /opt/program
```



aws

# High Level Workflow

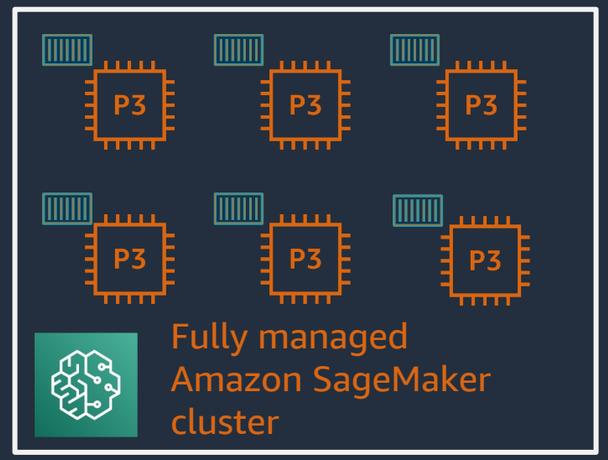

Docker build





Container registry

```
ic = sagemaker.estimator.Estimator(training_image,
                                   role,
                                   train_instance_count=1,
                                   train_instance_type='ml.p3.xlarge',
                                   train_volume_size = 50,
                                   train_max_run = 360000,
                                   input_mode= 'File',
                                   output_path=s3_output_location,
                                   sagemaker_session=sess)
```




Amazon Simple Storage Service



aws

# Bring Your Own Docker File

Customer Managed

1. Write your model however you please
2. Point to your model within your Docker file
3. Register your container on ECR
4. Point to your container’s address in ECR
5. Don’t forget to implement a serve() function!



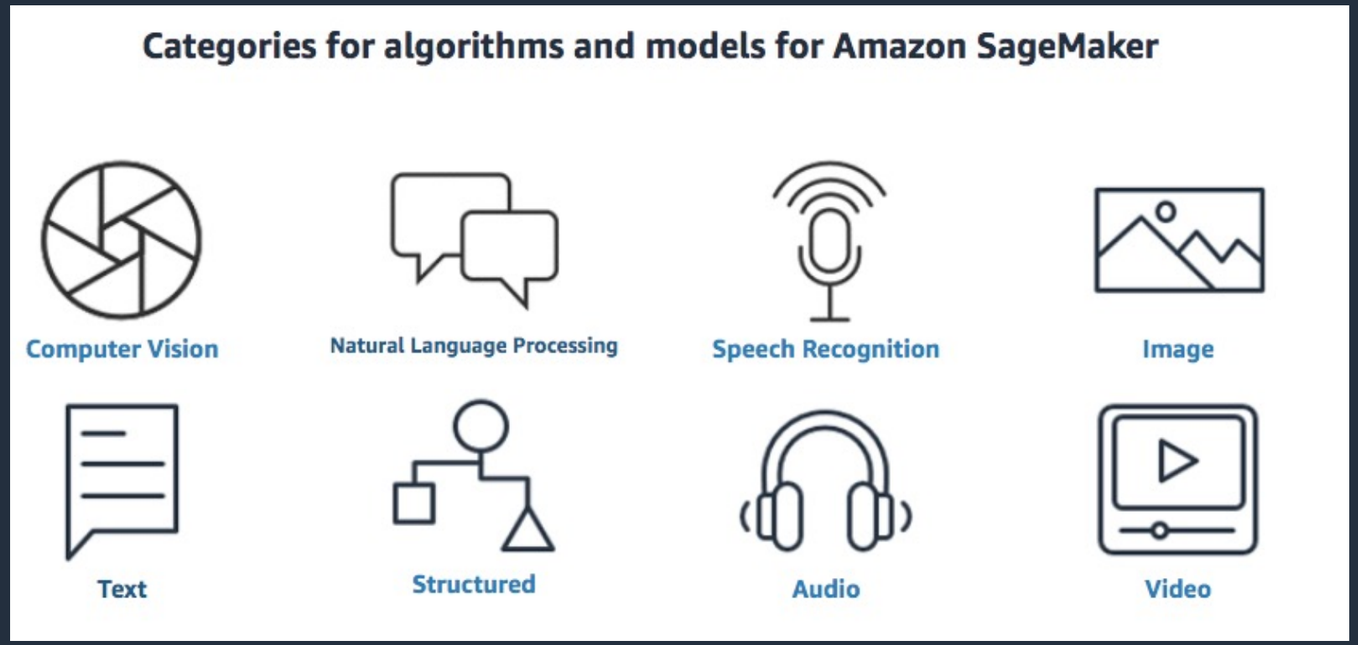
aws

# ML Marketplace

| Algorithms | Models |
| --- | --- |
| • You can train on your data | • Use a pretrained model artifact |

- Subscription model
- Free tier!

230+ solutions!






aws

# Pro Tips

| | *Script Mode* | *BYO Docker* | *ML Marketplace* |
|---|---|---|---|
| Pros | Quickly train your own model | Most flexible | Huge variety: easy to add value quickly |
| Cons | Limited to managed options | More time consuming | Less insight into solution |

SageMaker reads from /opt/ml

Find an example to modify



aws

# Thank you!

Michael Lin

linmicht@amazon.com