
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

SUS301

Sustainable machine learning for protecting natural resources

Nelson Gonzalez (he/him)
Head of Global Impact Computing
AWS

Benoit de Chateauvieux (he/him)
Solutions Architect
AWS

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Agenda

1. AWS’s commitment to sustainable technology and technology for sustainability
2. Impact computing for natural resource business intelligence: oceans, forests
3. Innovation for ML at the extreme edge and supply chain resilience
4. Reducing the carbon footprint of ML workloads
5. Optimizing along the ML lifecycle
6. Call to action

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Success and scale bring broad responsibility

We started in a garage, but we're not there anymore. We are big, we impact the world, and we are far from perfect. **We must be humble and thoughtful about even the secondary effects of our actions**. Our local communities, planet, and future generations need us to be better every day. We must begin each day with a determination to make things better, do better, and be better for our customers, our employees, our partners, **and the world at large**. We must end every day knowing we can do even more tomorrow. **Leaders create more than they consume and always leave things better than how they found them**.

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Further and faster, together

September 19, 2019

Amazon and Global Optimism announced The Climate Pledge, a commitment to meet the goals of the Paris Agreement 10 years early – and achieve net-zero carbon by 2040

Amazon is a co-founder and first signatory of The Climate Pledge







aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.


MiiR
NESPRESSO.
CARBON CURE
asos
cabify
gsk
BEST BUY
Snap Inc.
itv
CLIF
Mercedes-Benz
reckitt
Weyerhaeuser
Microsoft
Unilever
jetBlue
CANARY WHARF GROUP PLC
BROOKS
VAUDE
hotelbeds
PROSEGUR
JLL
CBRE
MAERSK
Infosys
verizon
SIEMENS
Ingenuity for life
With 375+ signatories from around the world
HRS
amazon
P&G
SAP
Uber
acciona
Klarna.
hp
Iceland
Johnson Controls
VISA
IBM
Coca-Cola EUROPEAN PARTNERS
Mckinstry
For The Life Of Your Building
CRANSWICK plc
Great British Taste
Telefónica
Lime
IPG
MACQUARIE
PEPSICO
BOOM
PHILIPS
Sainsbury's

The Climate Pledge Fund

A corporate venture capital fund that invests in companies that can **accelerate** the path to meeting The Climate Pledge

The Climate Pledge Fund invests in companies across multiple industry sectors with an initial focus on:



Circular economy



Energy generation, storage, and utilization



Food and agriculture



Manufacturing and materials



Renewable energy technology



Transportation and logistics

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Amazon is the largest corporate buyer of renewable energy globally

---

More than 379 projects around the world, on the path to power operations with 100% renewable energy by 2025

Once online, enough power for more than 4.6 million US homes/year or 13.4 million European homes/year

Impact computing

Help customers **improve lives** of humans, other species, and natural ecosystems through modeling, simulation, and data-rich compute solutions













© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.







The **Amazon Sustainability Data Initiative (ASDI)** significantly reduces the **cost**, **time**, and **technical barriers** associated with analyzing large datasets to generate sustainability insights

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ASDI: Making access to data faster, cheaper, and easier











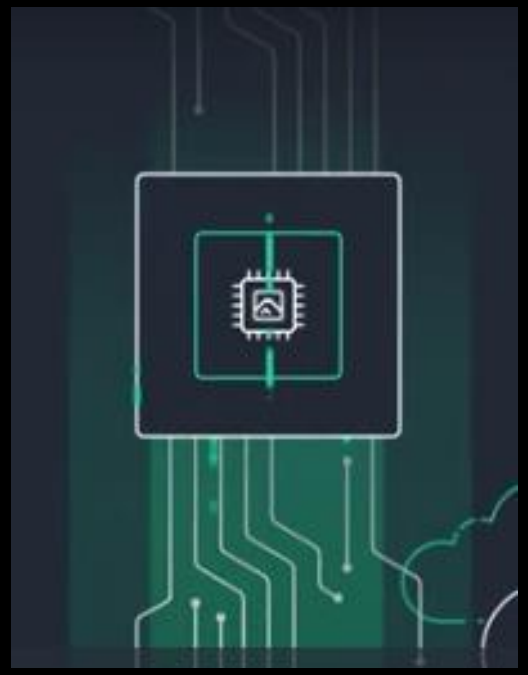




- Climate data projections
- Weather forecasts
- Historical weather records
- Satellite imagery
- Air quality
- Water
- Energy
- Ocean forecasts
- Sea surface temperature
- Environmental indicators

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



AWS Graviton and Inferentia

The latest generation of AWS-designed processors built for the cloud

---

Graviton3-based Amazon EC2 instances use up to 60% less energy for the same performance than comparable instances

---

Inferentia is one of our most power-efficient inference processor

Innovation for social impact

Automating large-scale HPC runs

HPC

AWS Lake Formation	Life-in-the-loop simulation	Systems' digital twins
Amazon Sustainability Data Initiative	Energy/water/climate simulation for policy design	National/regional economy for UN SDGs
AWS social determinants of Amazon HealthLake	Financial health simulation	Social determinants of health IoT for building design/engineering
AWS ESG measurement	Nuclear fusion simulation	Circular economy modeling

ML

Applying extreme edge AI/ML – responsible AI

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine learning for sustainability use cases

Electricity systems

Enabling low-carbon electricity | Reducing current system impacts | Reducing current impact

Transportation

Reducing transport activity | Optimizing air traffic | Alternative fuels and electrification

Buildings/cities

Optimizing buildings | Urban planning | Smart cities | Digital twins

Industry

Optimizing supply chain | Improving materials | Production and energy

Farms and forests

Precision agriculture | Managing forests | Estimating carbon stock

Responsible AI

CO2 emissions | Bias | Explainability | Privacy | Model accuracy | Data drift | Service health

Finance

Designing carbon markets | Modernize risk management | Microloan | NLP for ESG

Circular economy

Design circular products, components, and materials | Operate circular business models | Optimize infrastructure

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine learning for sustainability approaches

Computer vision	Reinforcement learning	Time series	Transfer learning
Uncertainty quantification	Unsupervised learning	Responsible AI	Natural language

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The biodiversity/supply chain crisis



Half of global GDP, or $44 trillion, is threatened by nature loss

Biodiversity and enterprise resilience both win by investing upstream in long-term, place-based strategies to improve supply chain sustainability through adaptive management of the first mile of our supply chains



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Natural resource business intelligence



Databases

+



Data warehouses
+ data lakes

+



Business
intelligence tools

AMAZON AURORA ML



AMAZON NEPTUNE ML



AMAZON REDSHIFT ML



AMAZON ATHENA ML



AMAZON QUICKSIGHT ML



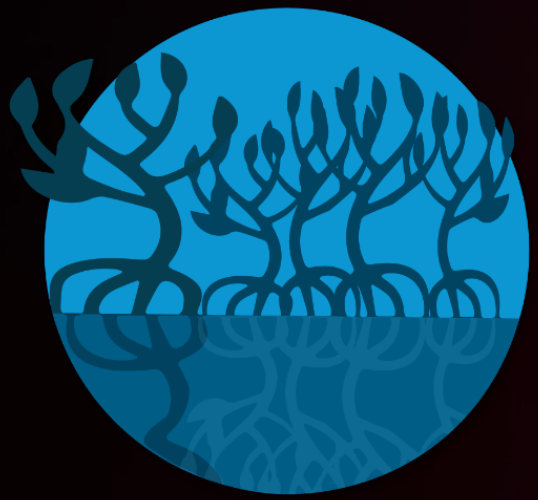
aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

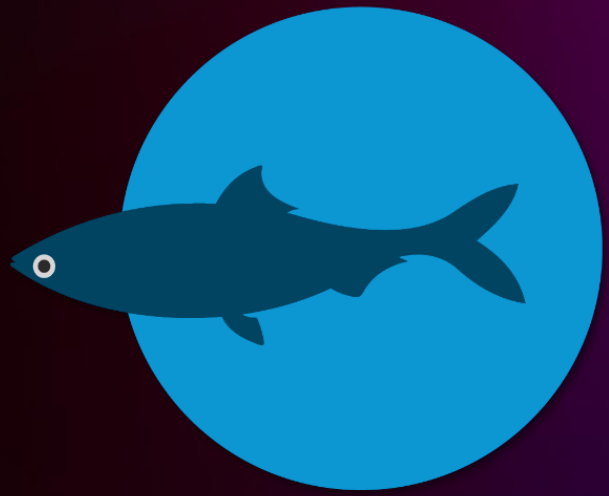
Our oceans are critical



90% of habitable space on the planet



50% of our oxygen



Food and income to **>3 billion people**

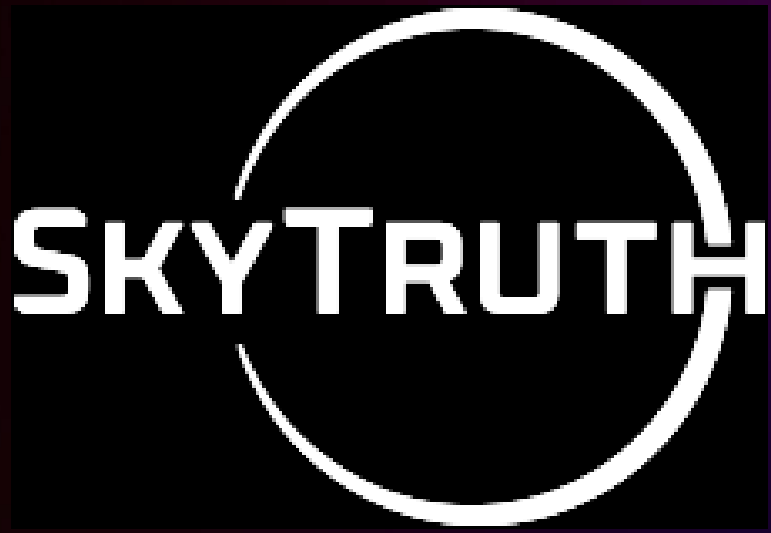
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Oceans





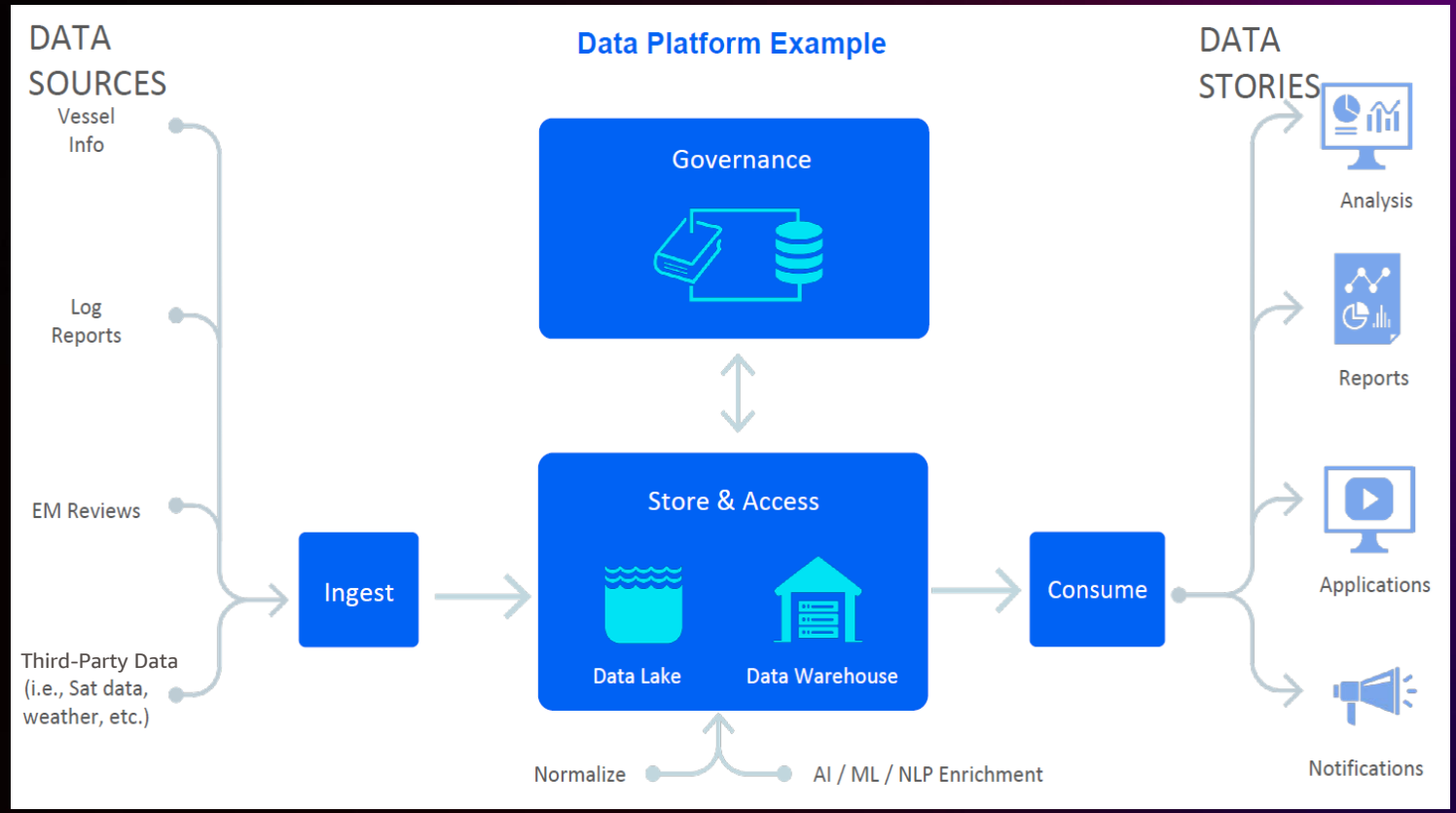




© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Electronic monitoring data architecture

- Architecture, including a data model, that outlines governance, security, data warehouse, ingest, APIs, etc.
- Open-source stack that can be used by companies, RFMOs, etc.
- Platform as a Service (PaaS) that members could join to both add and retrieve data

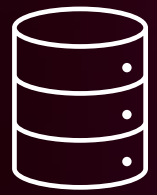


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

AWS is helping to address major pain points and drive change on the water



Employing AI/ML at review centers and on the edge



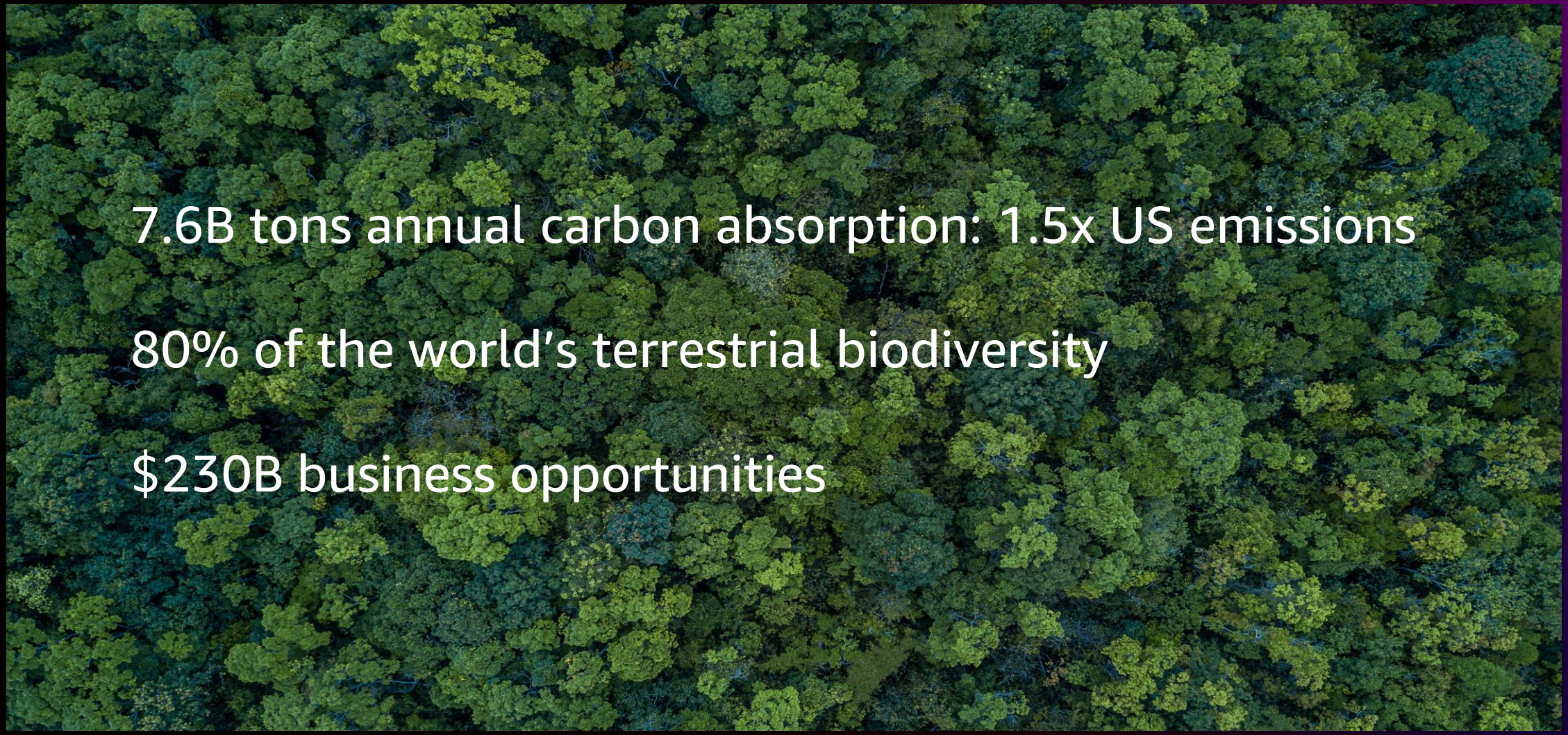
Building a standardized (but flexible) data architecture

Together, these two strategies pave the way for faster, more effective monitoring of conditions on vessels

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Forests



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Forests

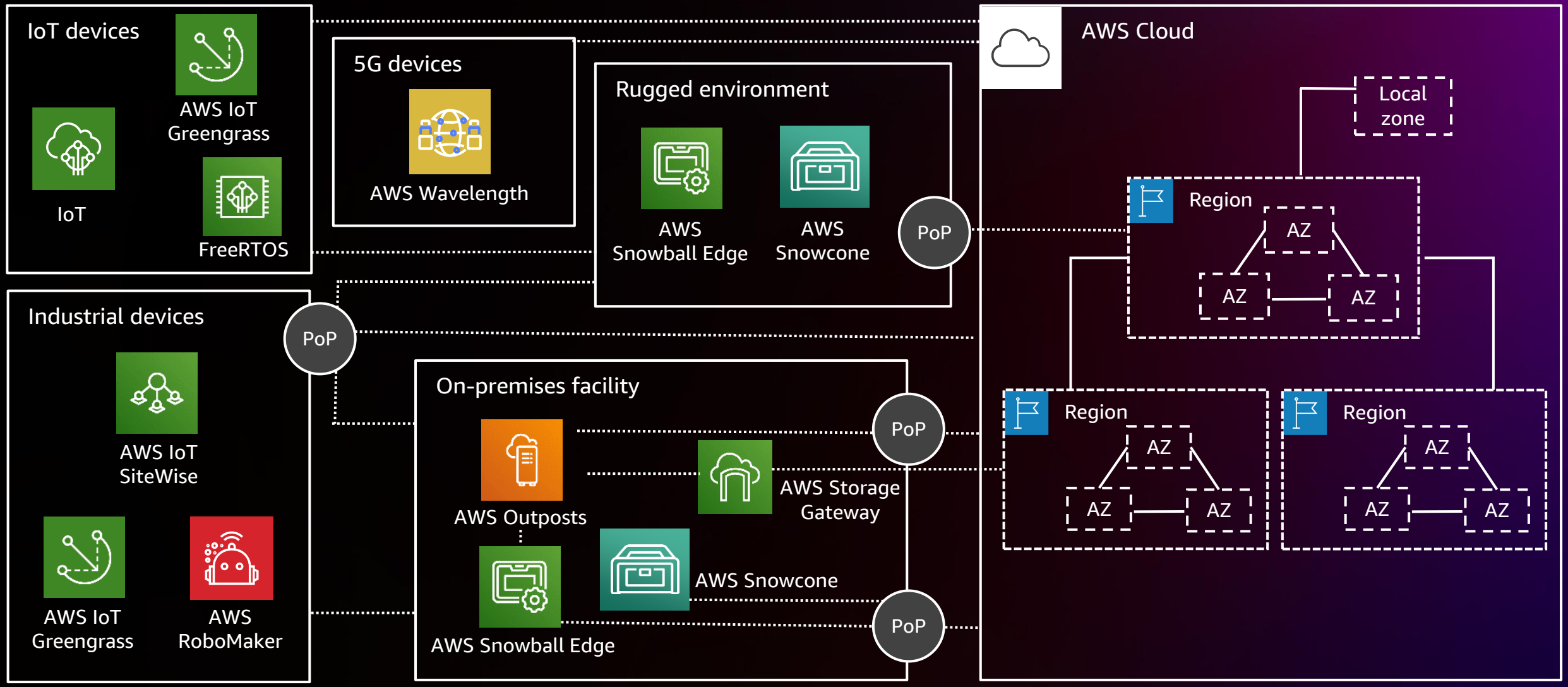






aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Implication: ML for extreme edge

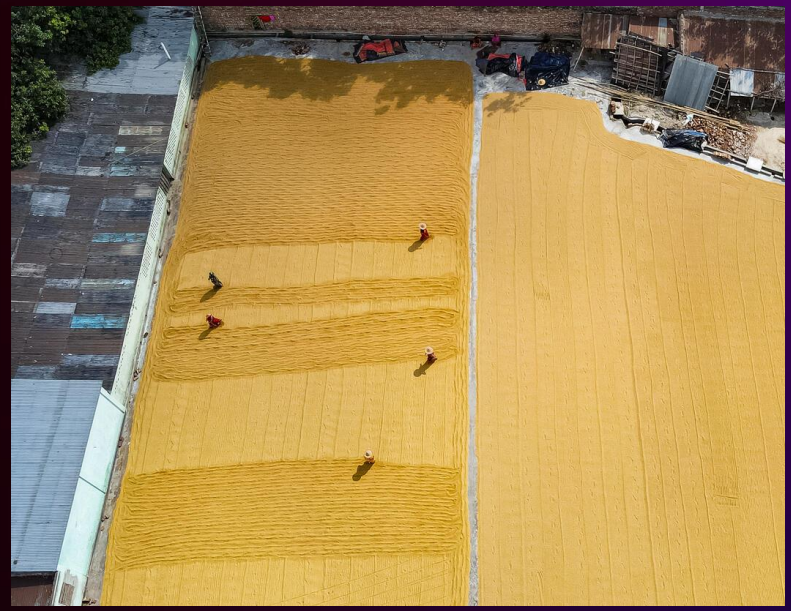


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Implication: “First mile” issues in other opaque and intermediated commodity supply chains







© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

We are just beginning to discover what ML can do for conservation

Adaptive management of the first mile of our supply chains will require more than just the ability to collect data

We will need to nimbly translate data points into actionable information that can drive adaptive management



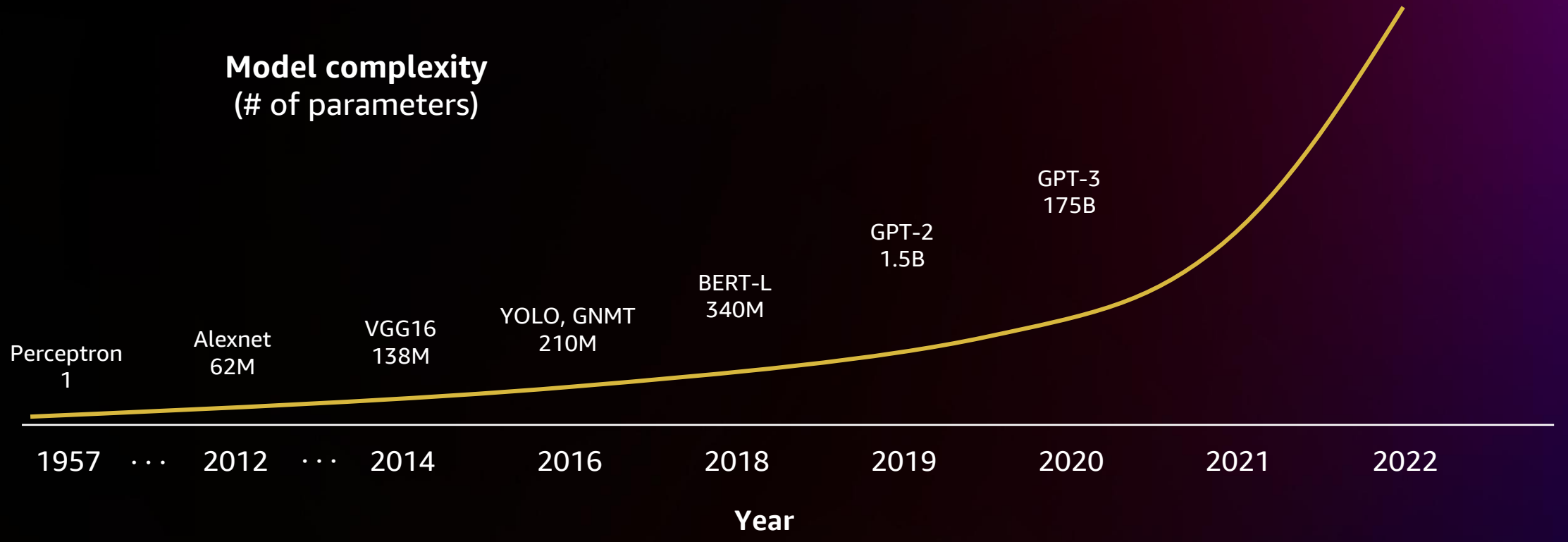
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Machine learning is a powerful tool to **decarbonize the planet** and **protect natural resources** . . .

It also questions how to build AI systems that are **sustainable** and benefit everyone

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

ML models are becoming bigger and more complex



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

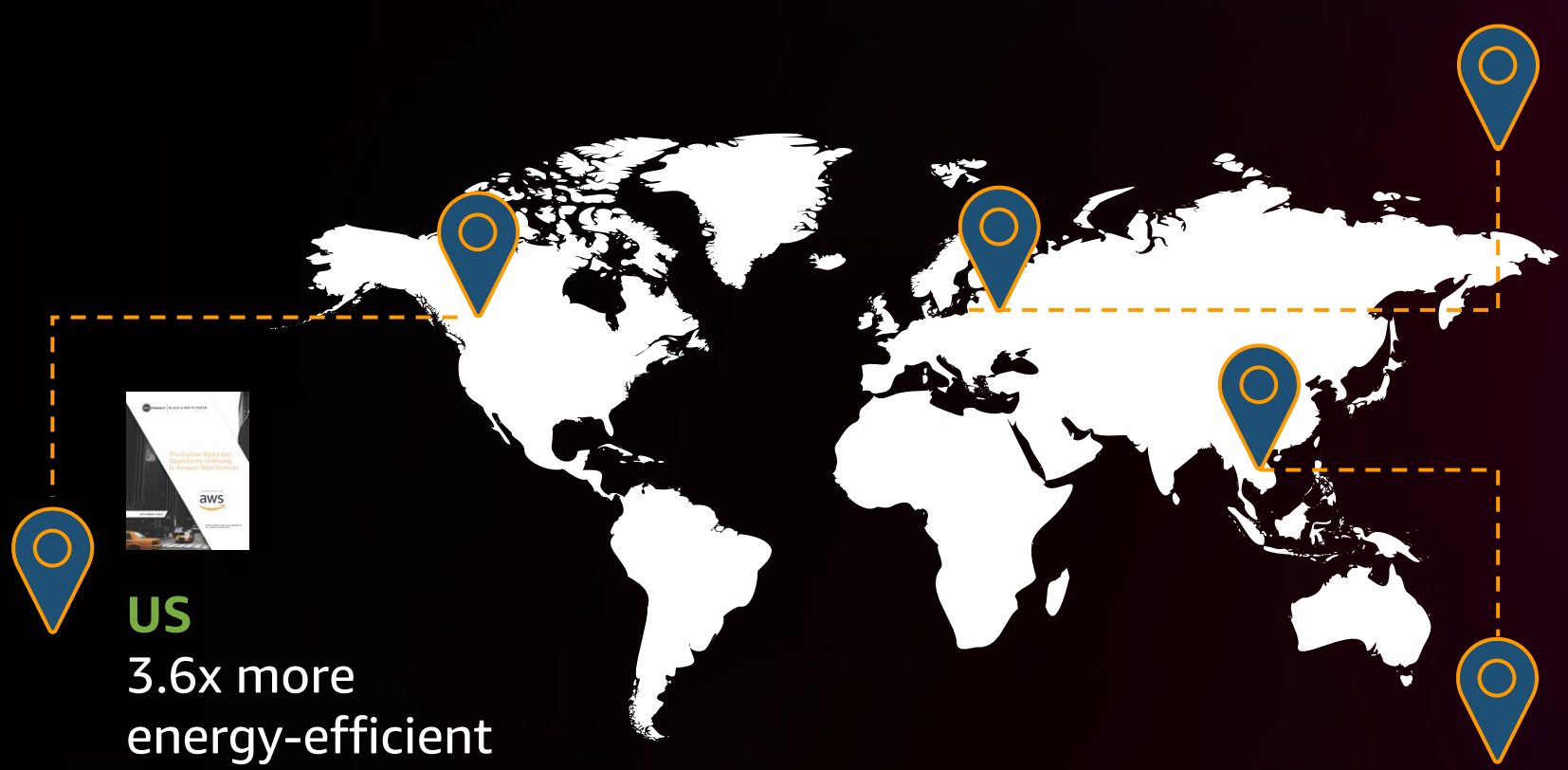
Source: Parameter, Compute and Data Trends in Machine Learning

Reducing the carbon footprint of ML workloads

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Carbon reduction opportunity

AWS CAN LOWER THE CARBON FOOTPRINT OF AVERAGE ON-PREMISES DATA CENTER WORKLOADS BY NEARLY 80% TODAY AND UP TO 96% ONCE AWS IS POWERED WITH 100% RENEWABLE ENERGY





Europe
Up to 5x more energy-efficient

US
3.6x more energy-efficient

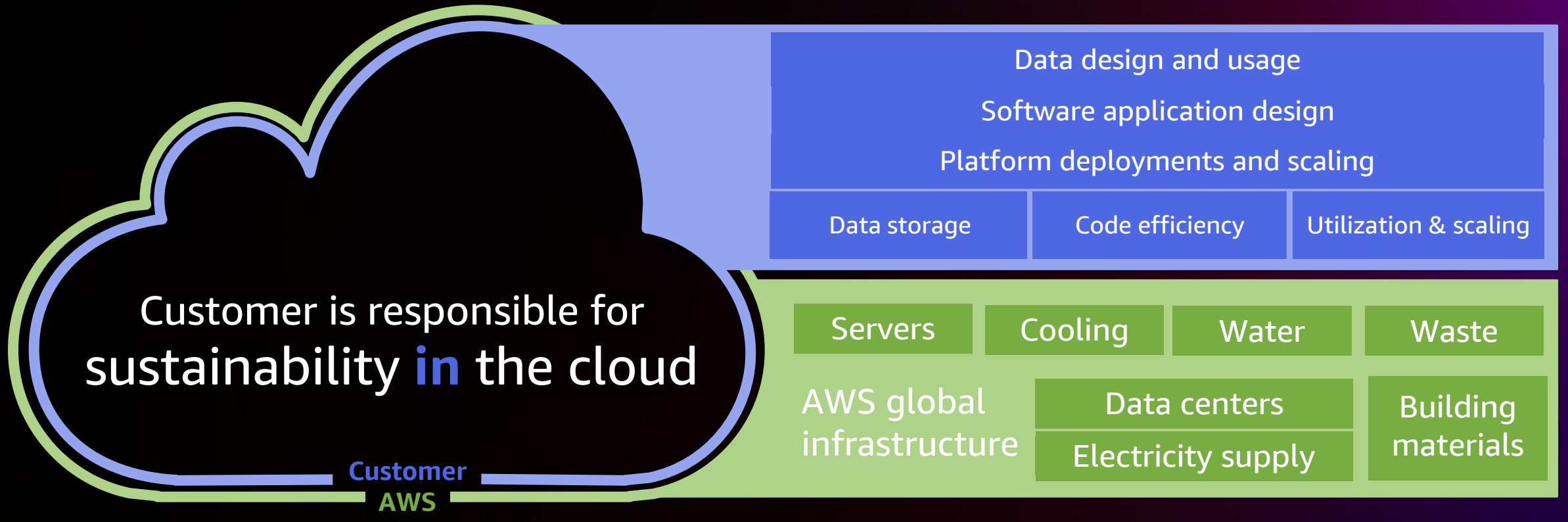


Asia Pacific
5x more energy-efficient

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Find all the reports on aws.amazon.com/sustainability/resources/

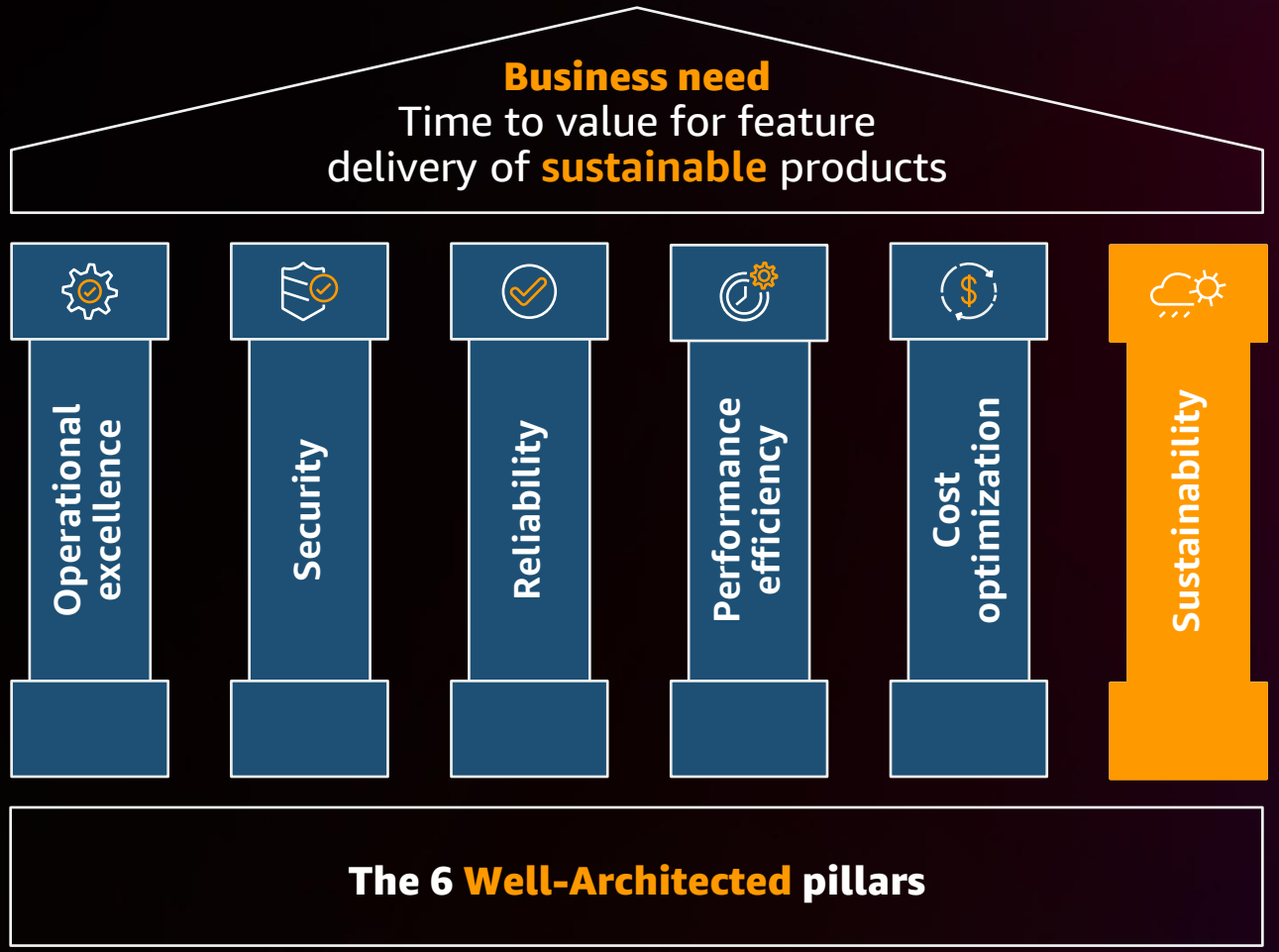
In the cloud, sustainability is a shared responsibility



AWS is responsible for
sustainability **of** the cloud

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

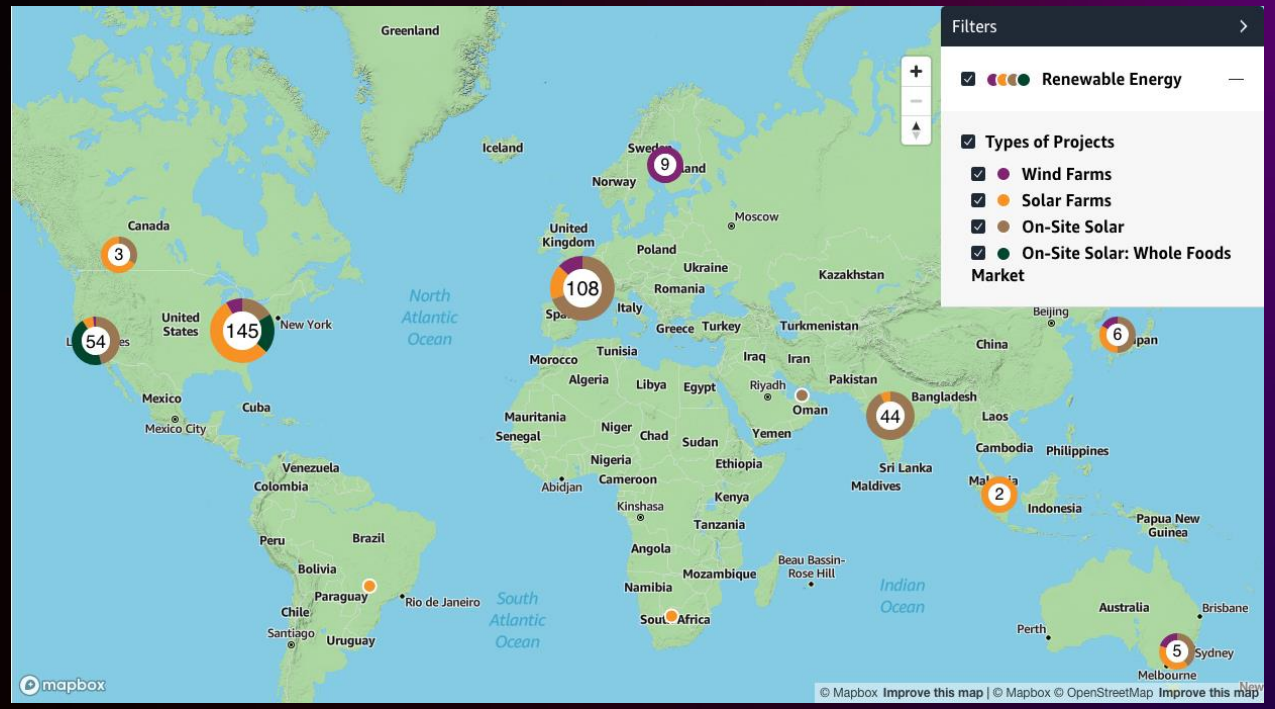
Optimize: AWS Well-Architected Framework for sustainability



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Region selection

- AWS is on a path to powering operations with **100% renewable energy** by 2025
- Choose Regions
  - Near Amazon renewable energy projects
  - Where the grid has a published carbon intensity that is lower than other locations (or Regions)



https://sustainability.aboutamazon.com/environment/renewable-energy

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Reducing resource usage



Reducing
energy usage

Reducing usage of
cloud resources

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Goals for optimization

Maximize

Utilization

CPU efficiency
via software design

Idle resources

Data processing

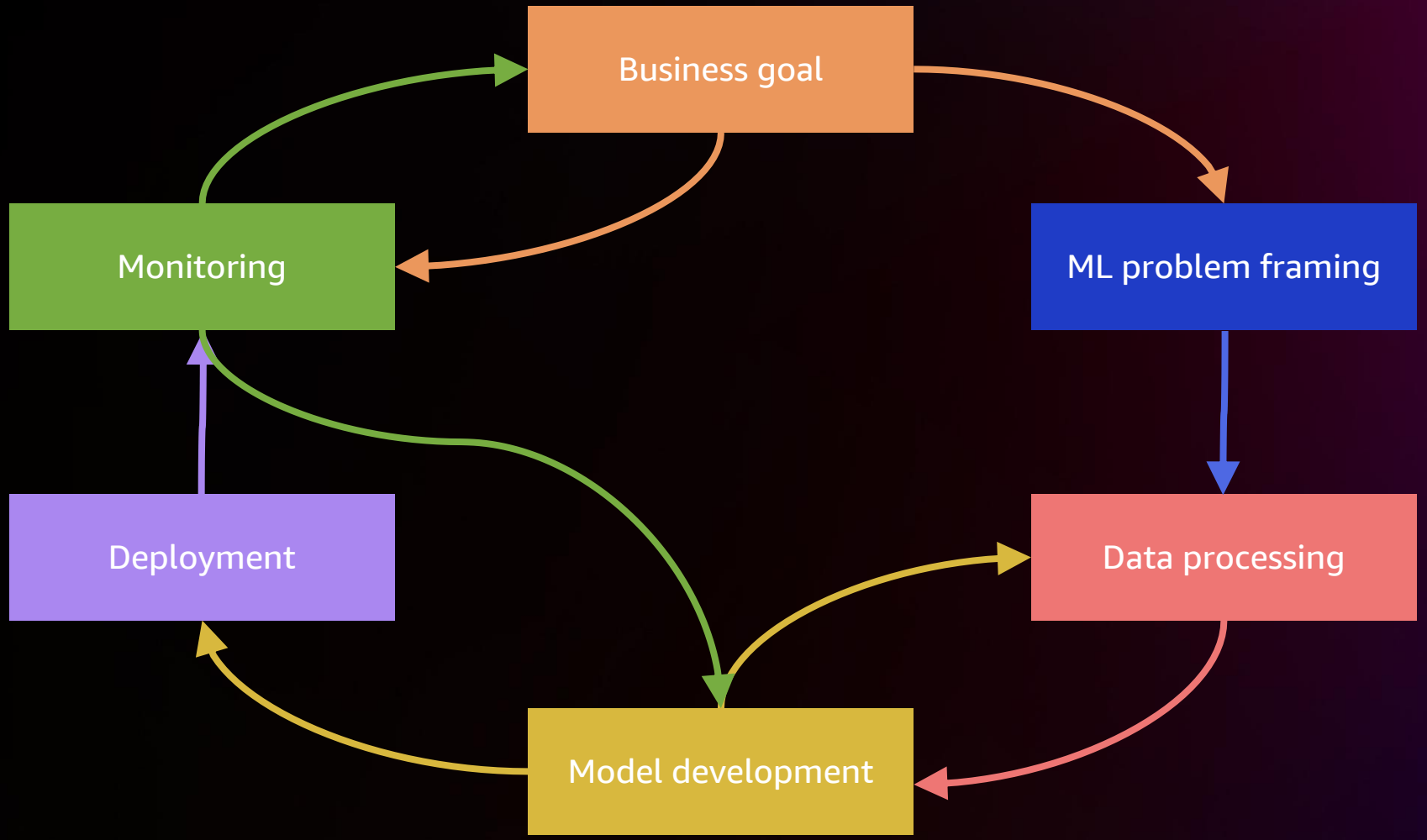
Data storage

Eliminate or minimize

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Optimizing along the ML lifecycle



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

https://docs.aws.amazon.com/wellarchitected/latest/machine-learning-lens/well-architected-machine-learning-lifecycle.html

ML problem framing phase

Managed AI services



Amazon Polly



Amazon Rekognition

Pre-existing datasets, algorithms, models



Amazon Marketplace



Amazon SageMaker JumpStart



Third-parties (Hugging Face . . .)

Custom models



Custom model

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data processing phase

Processing

- Avoid processing duplication
- Minimize idle resources with serverless data pipelines
- Minimize data movement across networks

Storage

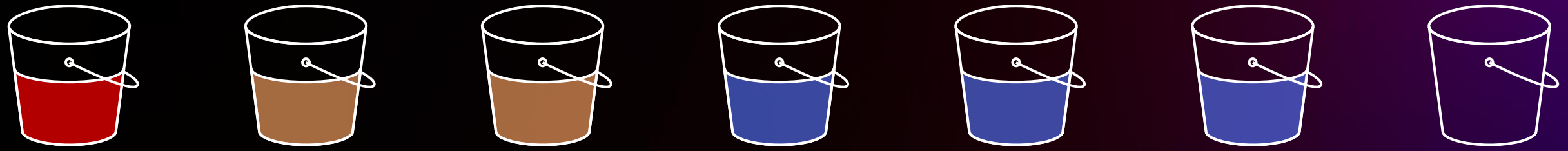
- Avoid dataset duplication
- Select efficient file formats and compression algorithms

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Sustainable storage options

Frequent Access Infrequent

Hot Warm Cold Deleted



Amazon S3 Standard

Active data
Millisecond access
≥3 AZ

S3 Standard-IA

Inactive data
Millisecond access
≥3 AZ

S3 One Zone-IA

Inactive data
Millisecond access
1 AZ

S3 Glacier Instant Retrieval

Archive data
Millisecond access
≥3 AZ

S3 Glacier Flexible Retrieval

Archive data
Minutes to hours
≥3 AZ

S3 Glacier Deep Archive

Archive data
Hours access
≥3 AZ

Useless data
0 AZ

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Model development phase

Model building

- Define acceptable performance criteria
- Select energy-efficient algorithms
- Use pretrained or partially pretrained models

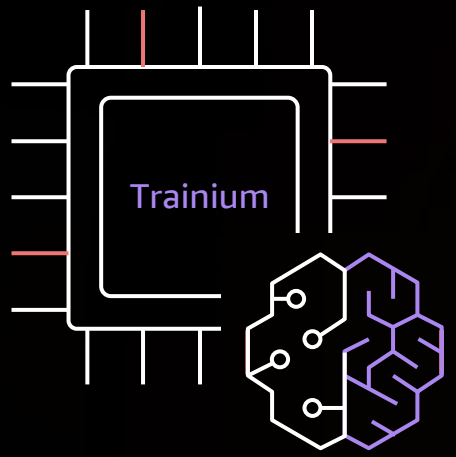
Model training and tuning

- Optimize the resources of your training environment
- Use efficient cross-validation techniques for hyperparameter optimization
- Use a debugger

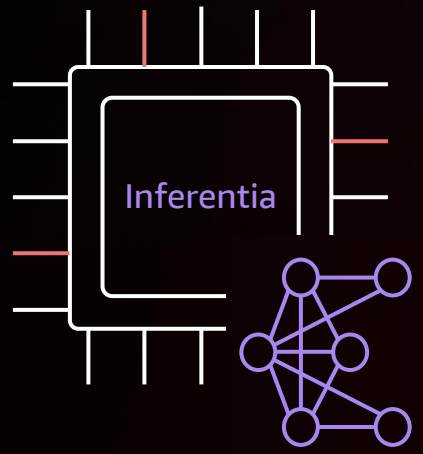
aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data processing phase

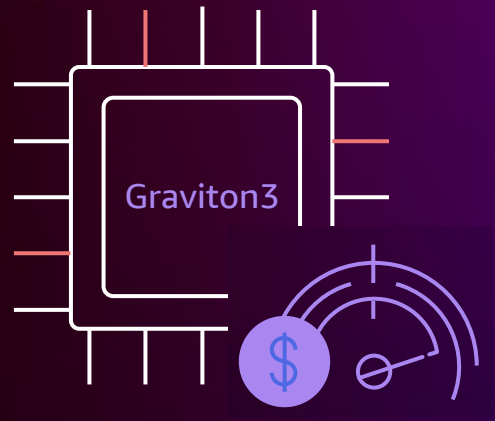
USE EFFICIENT SILICON



Expected to be one of the most **energy-efficient** AWS processors for training



Inf1 has **50% higher performance per watt** than g4dn, which makes it one of the **most sustainable** ML accelerators that Amazon EC2 offers



Some of the best performances per watt in Amazon EC2; up to **60% less energy** than comparable EC2 instances

Training

Inference

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

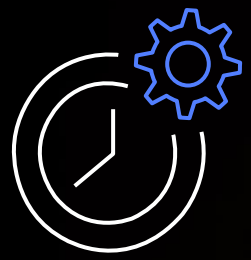
Model deployment and monitoring phases

Model deployment

- Make efficient use of GPU
- Optimize models for inference
- Deploy multiple models behind a single endpoint
- Right-size your inference environment
- Consider inference at the edge

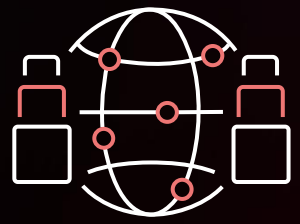
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Negotiate impact-friendly SLAs



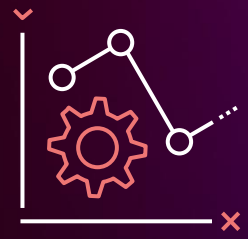
If your users can tolerate **some latency**

Deploy your model on asynchronous endpoints



If your users can tolerate **some latency** in the event of a failover

Don't provision extra-capacity



When you don’t need **real-time inference**

Use Amazon SageMaker batch transform

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Model deployment and monitoring phases

Model deployment

- Make efficient use of GPU
- Optimize models for inference
- Deploy multiple models behind a single endpoint
- Right-size your inference environment
- Consider inference at the edge
- Negotiate impact-friendly SLAs

Model monitoring

Retrain only when necessary

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Call to action!

Build



Amazon Sustainability Data Initiative
Browse ASDI datasets and propose an impact-computing POC



AWS Promotional Credit Program
Apply for an ASDI grant to develop sustainability-related projects

Optimize



Optimize AI/ML workloads for sustainability
Learn best practices from the blog post series



AWS Well-Architected Framework
Optimize with the Machine Learning Lens and the new sustainability pillar

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Sustainability breakout sessions

SUS204: Sustainability in AWS global infrastructure

SUS205: Architecting sustainably and reducing your AWS carbon footprint

SUS206: Sustainability and AWS silicon

SUS208: Utilizing sustainability data at scale

SUS210: Modeling climate change impacts and risks at scale

SUS211: Water stewardship and renewable energy at scale

SUS212: Accelerating decarbonization and sustainability transformation

SUS301: Sustainable machine learning for protecting natural resources

SUS303: Delivering sustainable, high-performing architectures

SUS312: How innovators are driving more sustainable manufacturing

STP213: Scaling global carbon footprint management

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Sustainability workshops and chalk talks

SUS201: Detecting deforestation with geospatial images and Amazon SageMaker

SUS202: Using Amazon SageMaker, AutoGluon, and ASDI to predict air quality

SUS203: Carbon emissions reporting using modern data architecture on AWS

SUS207: Patterns for obtaining insights from sustainability data

SUS302: Optimizing architectures for sustainability

aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Nelson Gonzalez
nrg@amazon.com

Benoit de Chateauvieux
chateauv@amazon.com



Please complete the session survey in the **mobile app**

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.