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

SUS205

Architecting sustainably and reducing your AWS carbon footprint

Margaret O'Toole

Worldwide Tech Leader, Sustainability
AWS

Elliot Nash

Sr. Manager, Software Development
Amazon Prime Video

aws

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

We’re on a journey together



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

What does sustainable even mean?

Meet the **needs of the present** without compromising the ability of **future generations to meet their needs**

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

Amazon’s Sustainability Journey

FARTHER 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 emissions by 2040; Amazon is a co-founder and the first signatory of The Climate Pledge

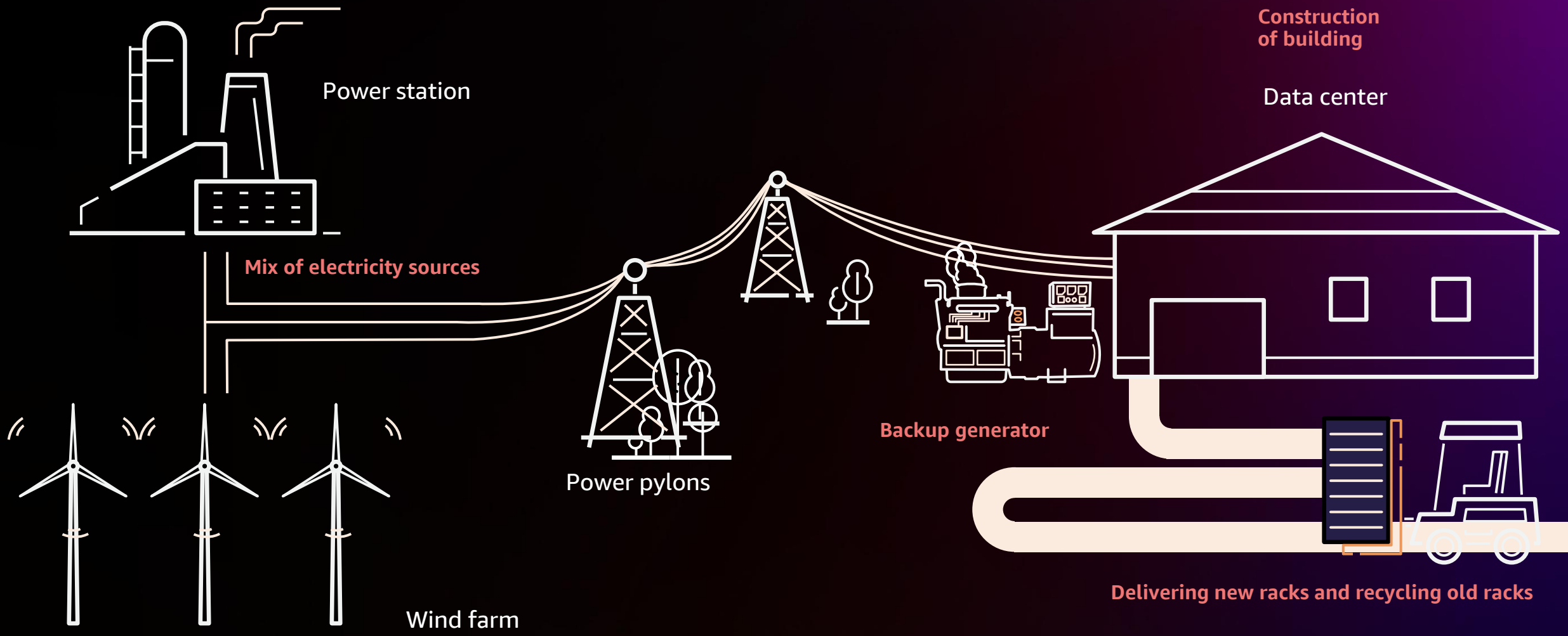


aws

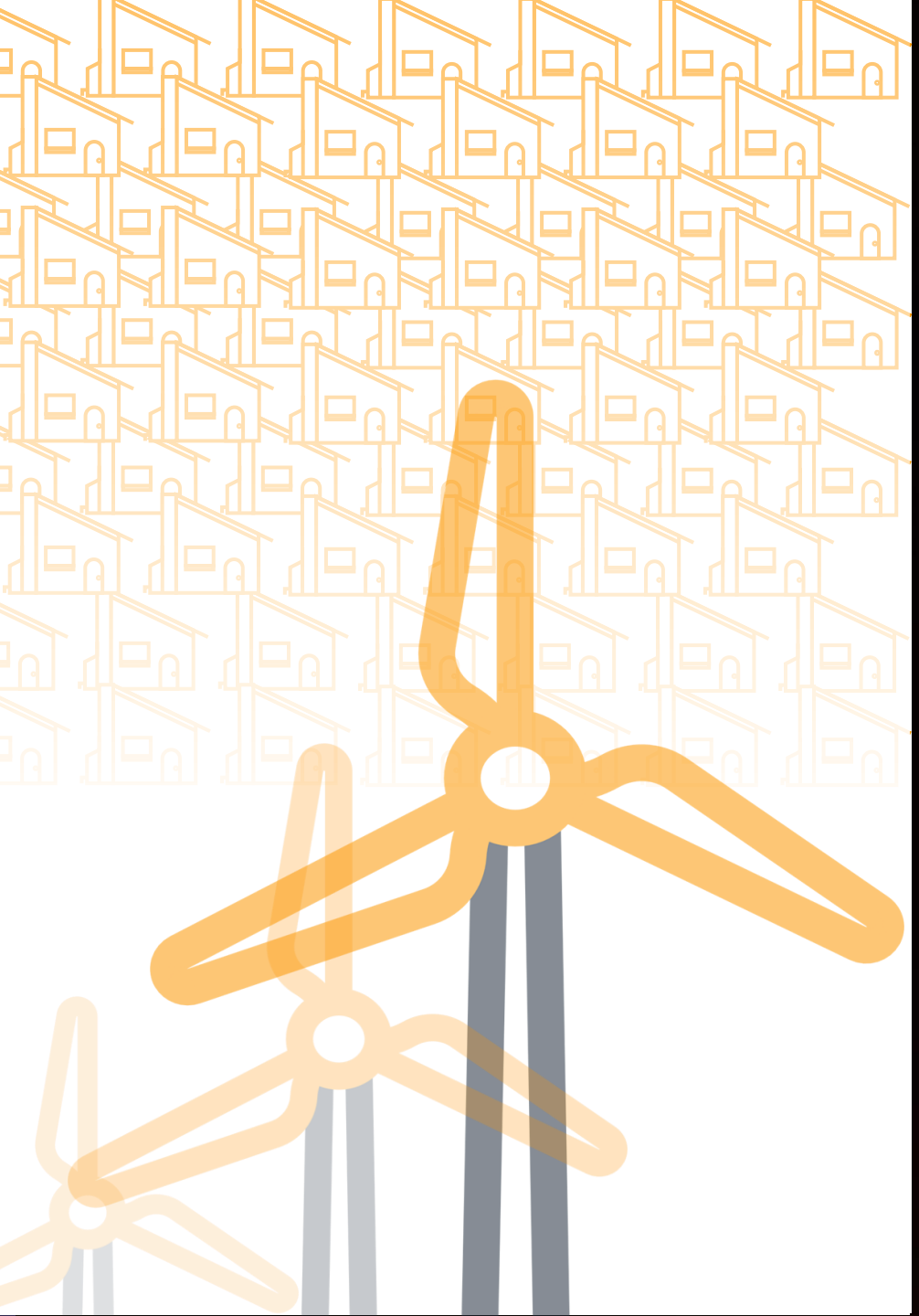
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

The carbon footprint of data centers

TOP-DOWN CONTEXT



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



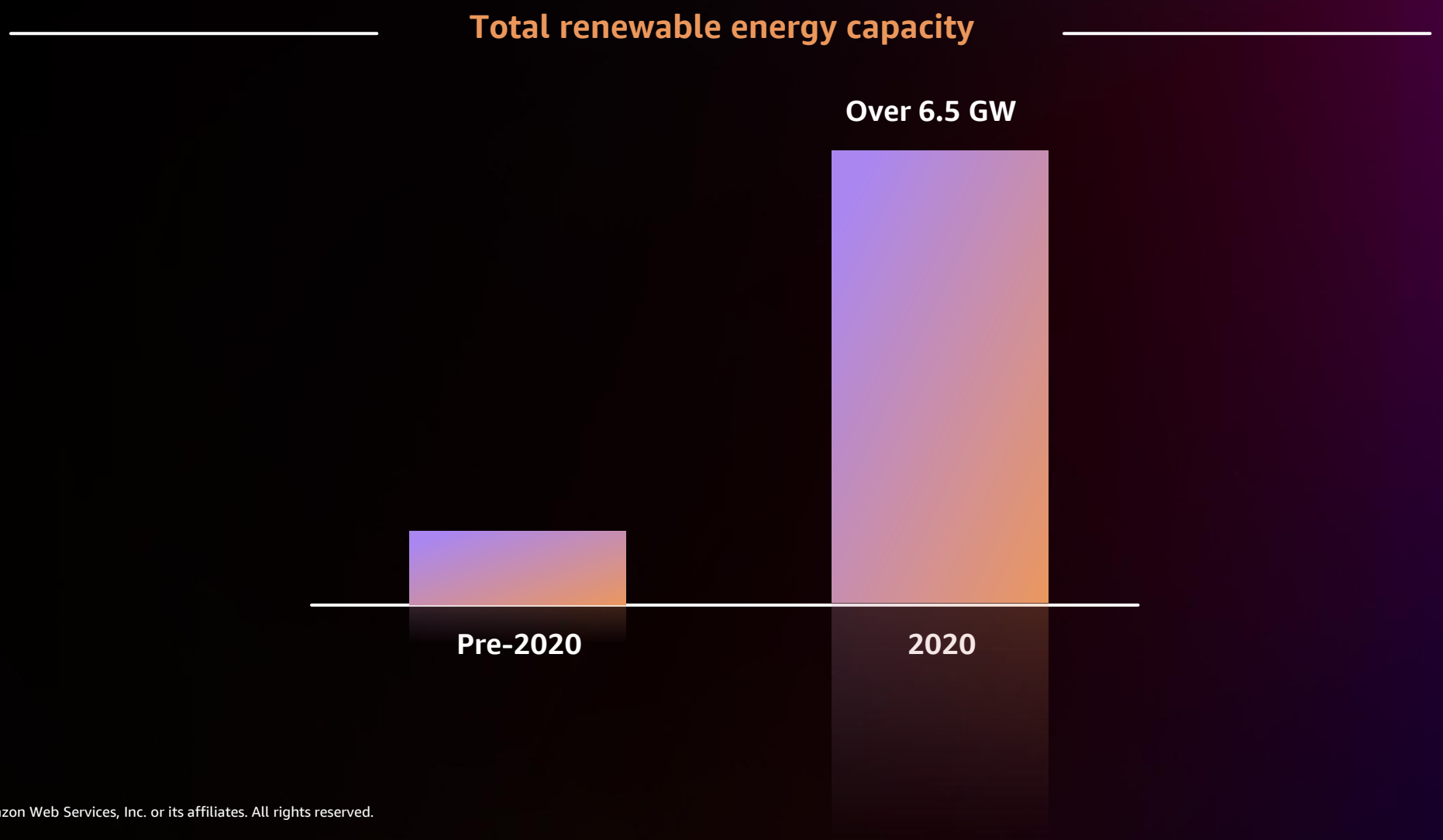
Amazon is the **largest corporate buyer of renewable energy** globally

---

More than 380 projects around the world, on the path to power operations with 100% renewable energy by 2025

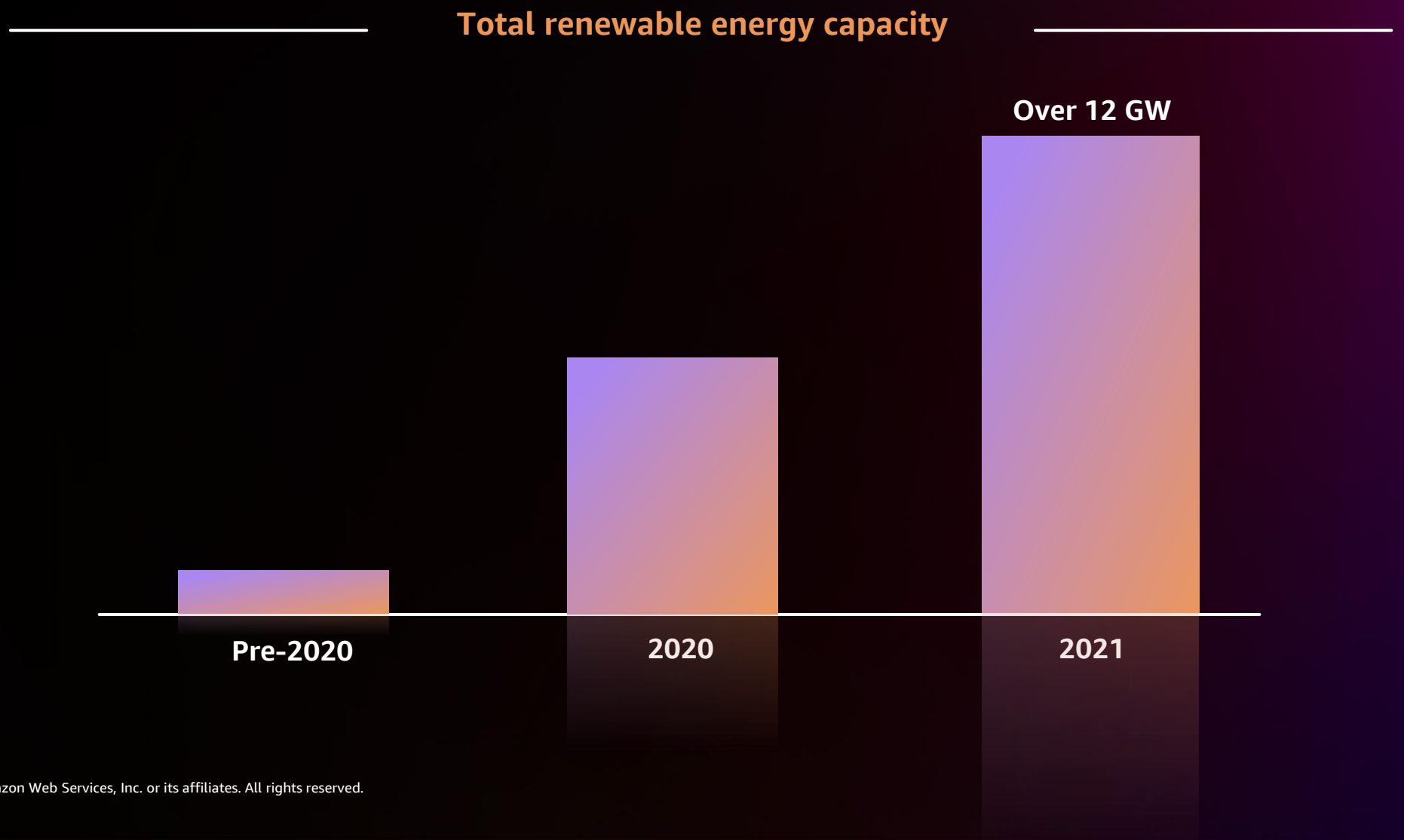
Once online, enough power for more than **4.6 million US homes/year** or **13.4 million European homes/year**

Path to 100% renewable energy by 2025



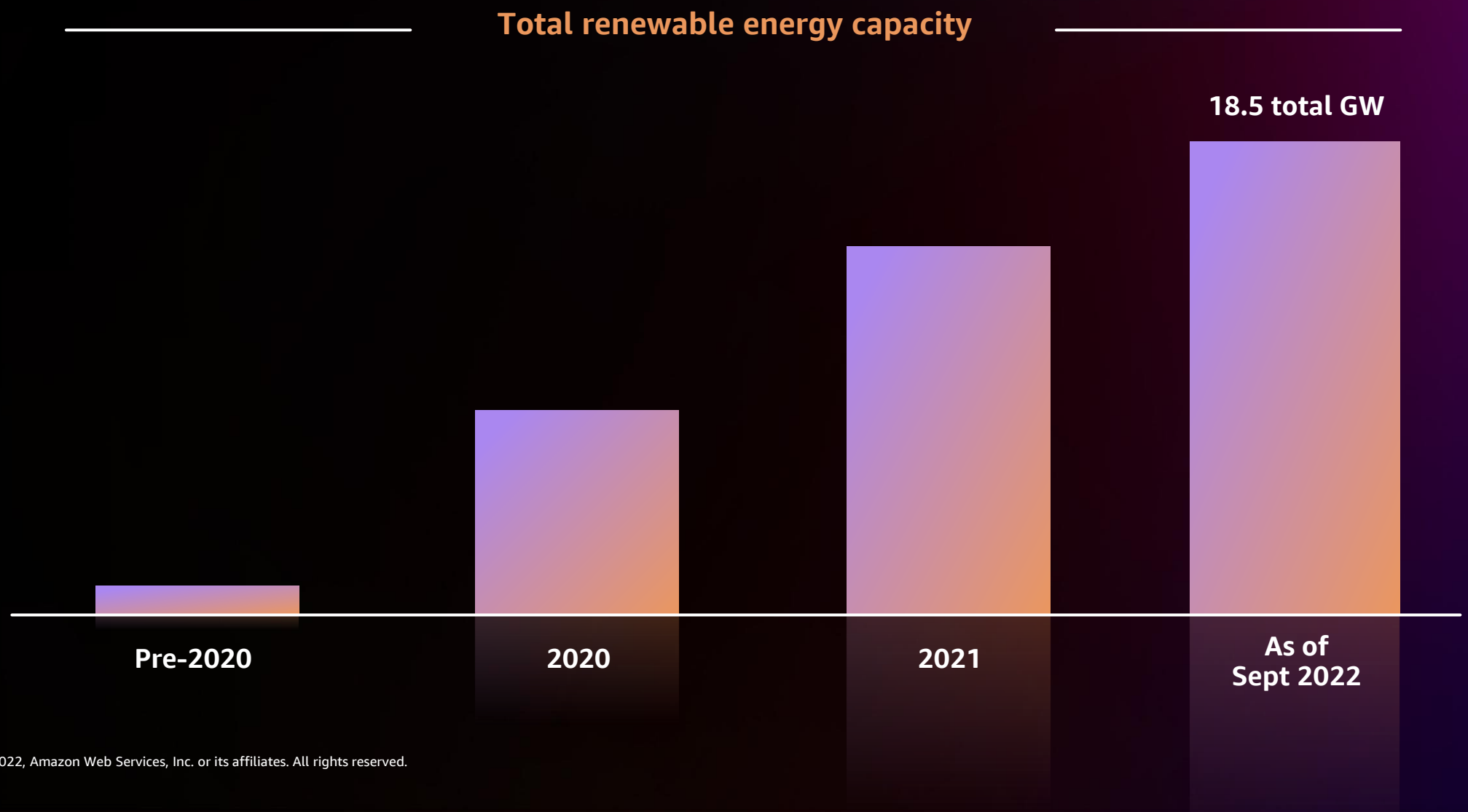
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Path to 100% renewable energy by 2025



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

Path to 100% renewable energy by 2025



© 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 5 times more energy efficient

US
3.6 times more energy efficient



Asia Pacific
5 times more energy efficient

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

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

Water positive

Recycled water

20

data centers using recycled water

Water efficiency

0.25

liters per kilowatt-hour in 2021

Water reuse

96%

of AWS's cooling water is reused in irrigation in US West (Oregon) Region

Water replenishment

2.4 billion

liters returned to communities through replenishment projects

aws

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

Let’s take a look back in time

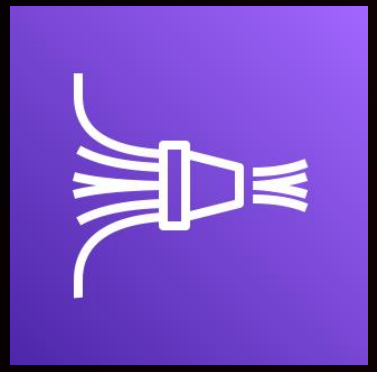


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

What happened in 2015?



Amazon ECR



Amazon Kinesis
Data Firehose



Amazon EC2
t2.nano and
X1 instances

aws

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

Are you well-architected?

AWS News Blog

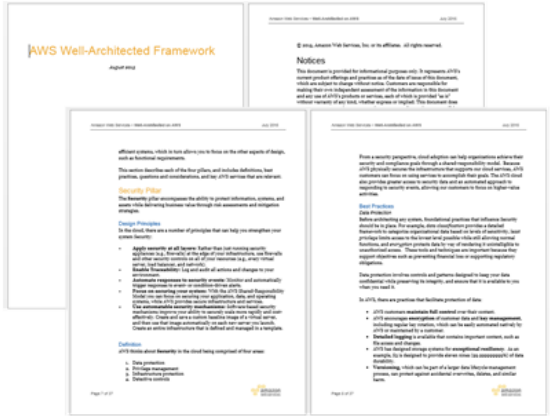
Are You Well-Architected?

by Jeff Barr | on 02 OCT 2015 | Permalink | Share

Seattle-born musical legend Jimi Hendrix started out his career with a landmark album titled Are You Experienced?

I've got a similar question for you: **Are You Well-Architected?** In other words, have you chosen a cloud architecture that is in alignment with the best practices for the use of AWS?

We want to make sure that your applications are well-architected. After working with thousands of customers, the AWS Solutions Architects have identified a set of core strategies and best practices for architecting systems in the cloud and have codified them in our new AWS Well-Architected Framework. This document contains a set of foundational questions that will allow you to measure your architecture against these best practices and to learn how to address any shortcomings.



The AWS Well-Architected Framework is based around four pillars:

- **Security** – The ability to protect information systems and assets while delivering business value through risk assessments and mitigation strategies.
- **Reliability** – The ability to recover from infrastructure or service failures, dynamically acquire computing resources to meet demand, and mitigate disruptions such as misconfigurations or transient network issues.
- **Performance Efficiency** -The efficient use of computing resources to meet system requirements, and maintaining that efficiency as demand changes and technologies evolve.
- **Cost Optimization** – The ability to avoid or eliminate unneeded cost or suboptimal resources.

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

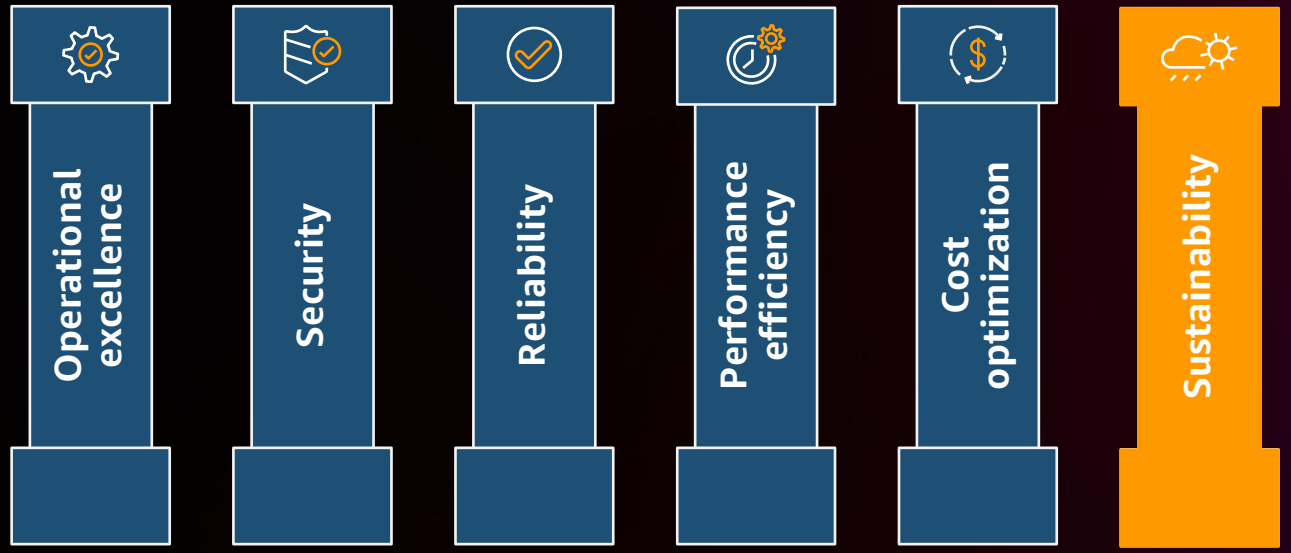
What “good” looks like has changed over the years

AND WE NEED TO CHANGE OUR ARCHITECTURAL PRACTICES WITH THAT



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

AWS Well-Architected Framework



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

The right resource, in the right amount, at the right time to deliver on your needs

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

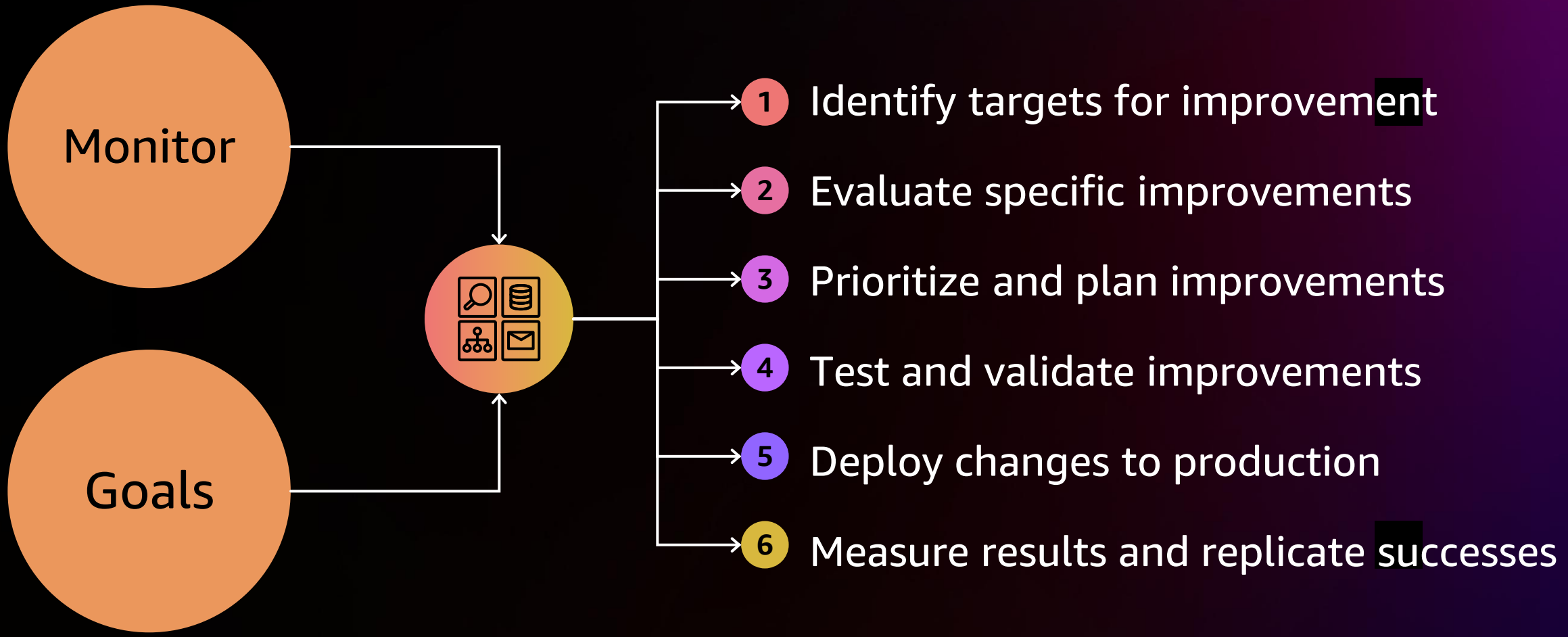
“Good intentions never work. You need good mechanisms to make anything happen.”

Jeff Bezos

Founder of Amazon

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

Improvement process



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

AWS customer carbon footprint tool



Track, measure, review, and forecast the carbon emissions generated from your AWS usage

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

Sustainability proxy metrics

- Help us understand how efficient we're being with our resources
- Understand what value we get from the things we provision
- Multiple metrics drive success



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

We’re on a journey together



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

Select a Region based on sustainability goals

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

Select a Region based on sustainability goals

SUS01-BP01

1. Assess and shortlist potential Regions for your workload based on your business requirements.

2. Choose Region(s) near Amazon renewable energy projects or Regions where the grid has a published carbon intensity that is lower than others.

aws

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

Shortlist potential Regions based on your business requirements



Latency



Cost



Services & features



Compliance

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

Choose Region(s) near Amazon renewable energy projects or Regions where the grid has a published carbon intensity that is lower than others



Market-based



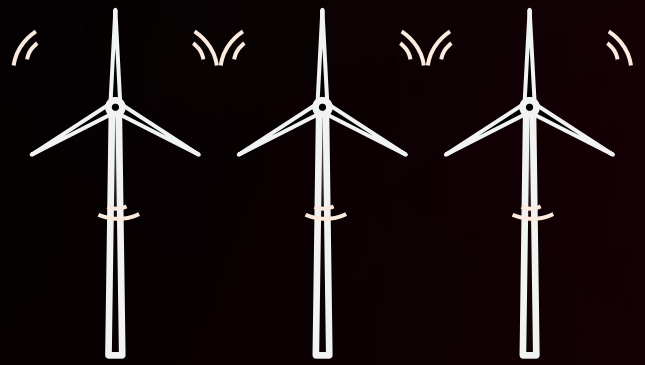
Location-based

aws

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

AWS Regions powered by over 95% renewable energy in 2021

US East (Northern Virginia)
GovCloud (US-East)
US East (Ohio)
US West (Oregon)
GovCloud (US-West)
US West (Northern California)



Canada (Central)
Europe (Ireland)
Europe (Frankfurt)
Europe (London)
Europe (Milan)
Europe (Paris)
Europe (Stockholm)

Renewable Energy Methodology

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

Well-Architected Sustainability Pillar Updates

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

Updated format

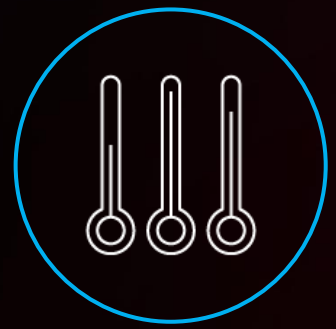
PRESCRIPTIVE BEST PRACTICES



Understand common anti-patterns



Learn benefits of establishing this best practice



Understand level of risk



See implementation guidance



Review relevant resources

aws

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

Common anti-patterns

SUS05-BP02 USE INSTANCE TYPES WITH THE LEAST IMPACT



Using one family of instances

Only using x86 instances

Using instances in a manner they were not designed for

Not evaluating new instance types

Not checking recommendations from AWS Compute Optimizer

One instance type in AWS Auto Scaling configuration

aws

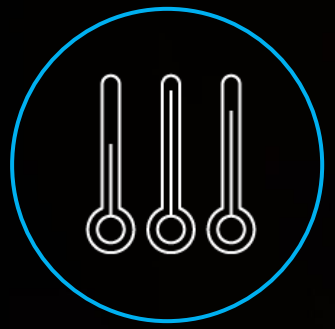
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Benefits and risks

SUS05-BP02 USE INSTANCE TYPES WITH THE LEAST IMPACT



By using energy-efficient and right-sized instances, you can greatly reduce the environmental impact and cost of your workload



Level of risk exposed if this best practice is not established: Low

aws

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

Implementation guidance

SUS05-BP02 USE INSTANCE TYPES WITH THE LEAST IMPACT



Learn and explore instance types that can lower your workload environmental impact

Plan and transition your workload to instance types with the least impact

Operate and optimize your workload instance

aws

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

Let’s take a real-world optimization example

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

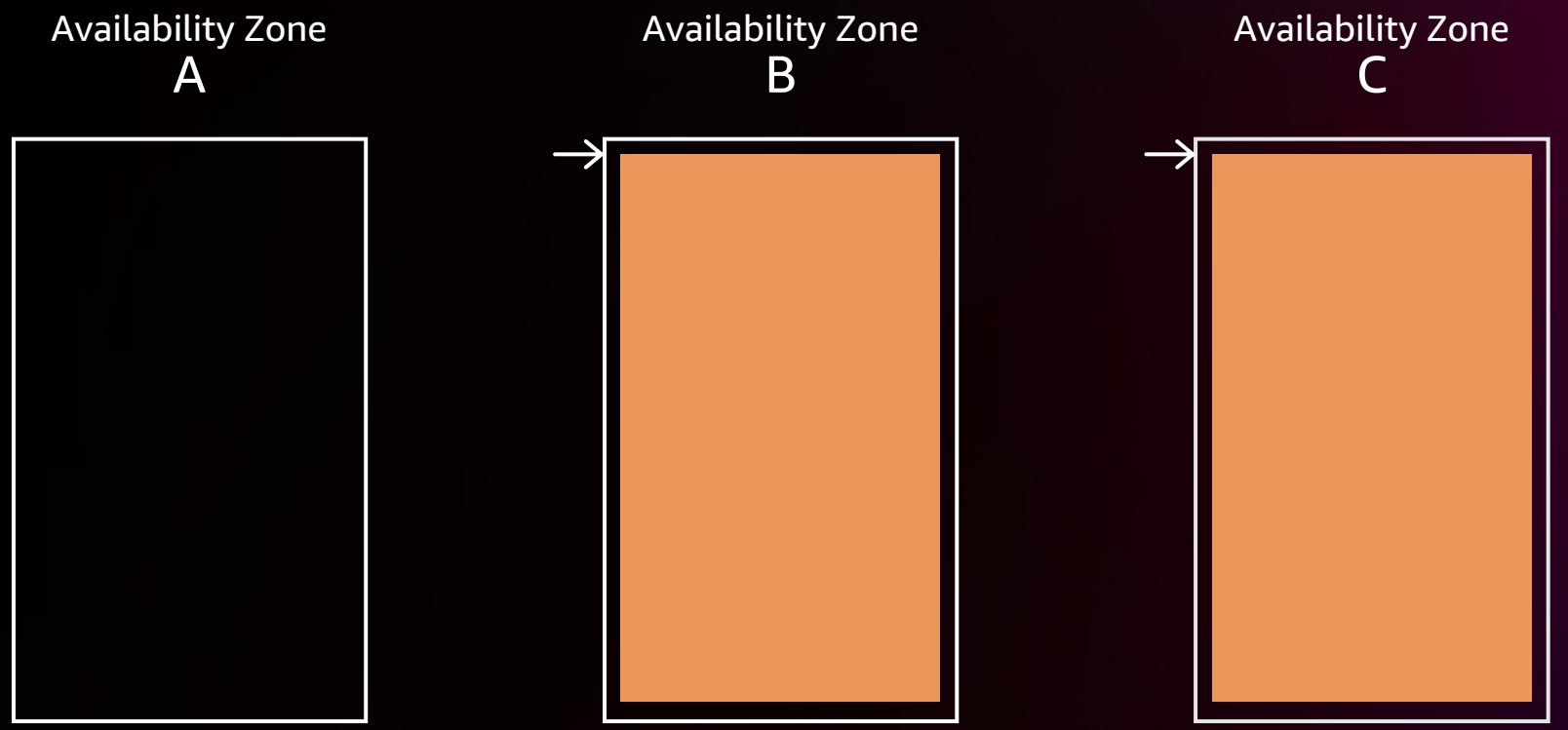
Optimize capacity for sustainability



AZ downscaling multiplier of 1.5x for scaling to handle an AZ outage

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

Optimize capacity for sustainability



Workload requirement: Immediate failover

AZ downscaling multiplier of 1.5x for scaling to handle an AZ outage

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

Optimize capacity for sustainability



AZ downscaling multiplier of 1.33x for scaling to handle an AZ outage

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

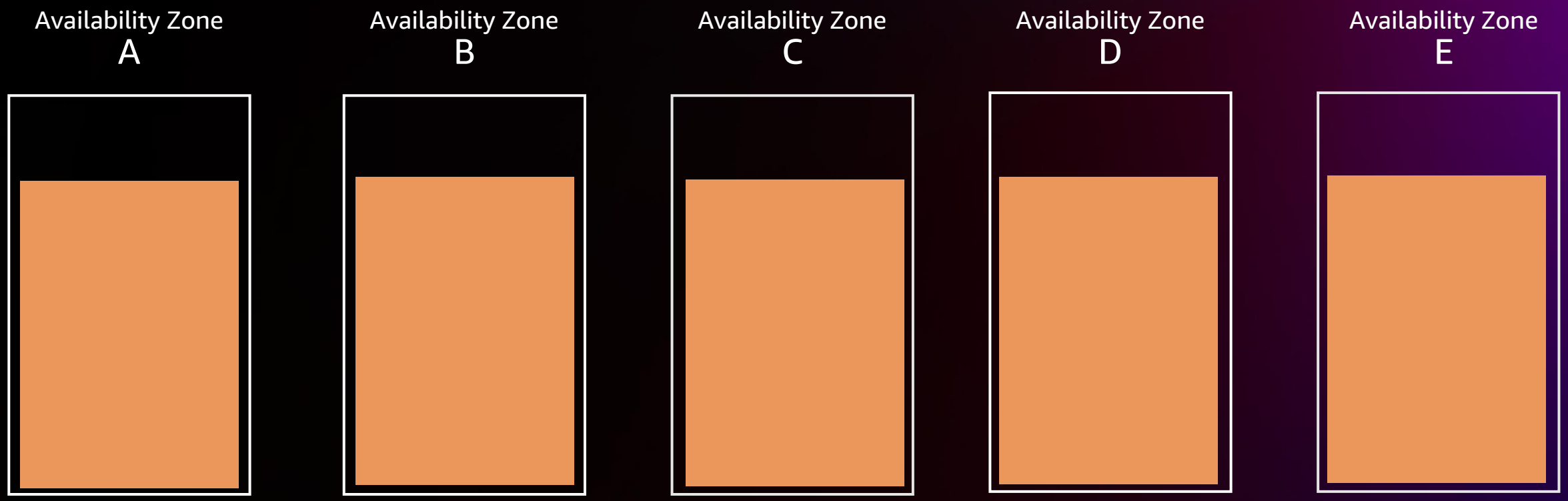
Optimize capacity for sustainability



AZ downscaling multiplier of 1.33x for scaling to handle an AZ outage

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

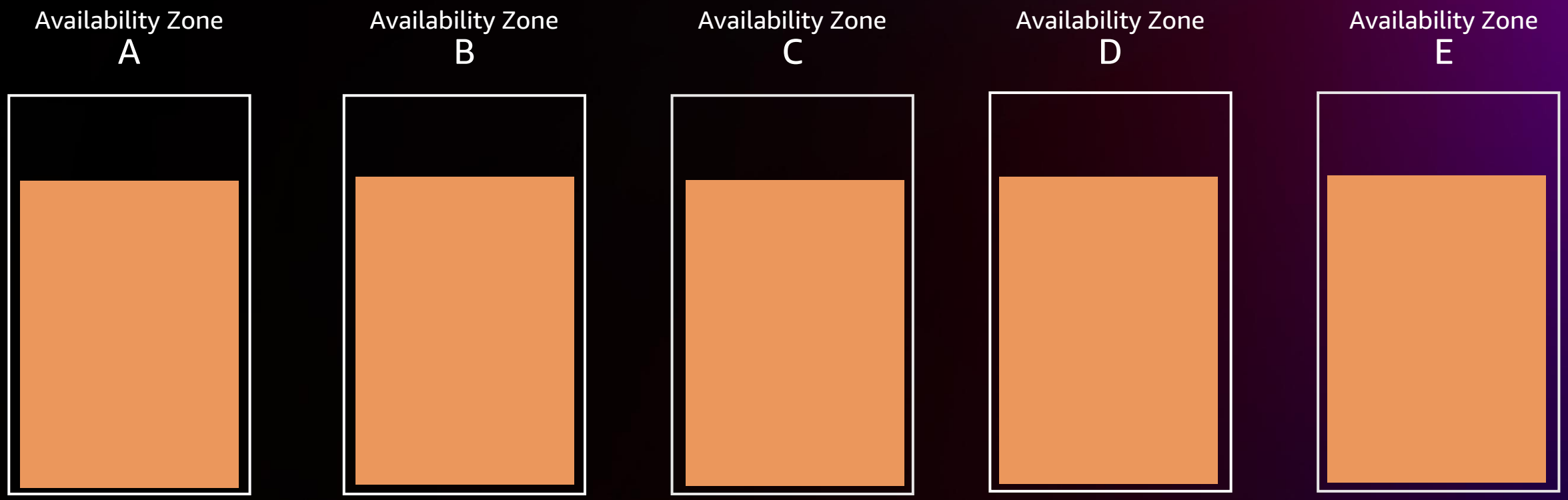
Optimize capacity for sustainability



AZ downscaling multiplier of 1.25x for scaling to handle an AZ outage

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

Optimize capacity for sustainability



AZ downscaling multiplier of 1.25x for scaling to handle an AZ outage

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

Distributing the workload

17.24% reduction in instance count



Faster failover

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

Get specific best practices for your workload

AWS Architecture blog channel

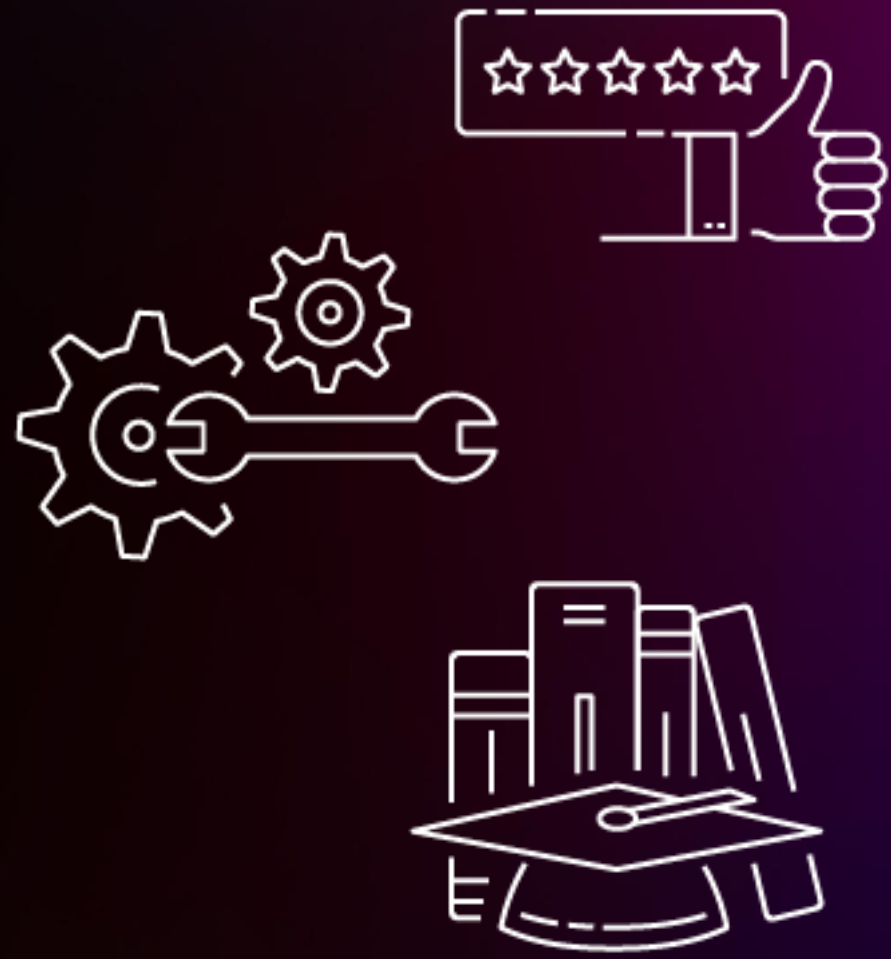
Compute – https://go.aws/3U1nLYc

Networking – https://go.aws/3h2c1pS

Storage – https://go.aws/3FDeukN

AI/ML – https://go.aws/3Uidoiv

Databases – https://go.aws/3UnBOax

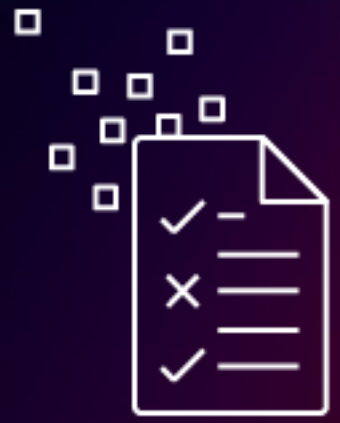


Get hands on at
https://wellarchitectedlabs.com/

aws

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

Reality: How do you prioritize these kinds of optimizations?



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

Introducing Amazon Prime Video



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

“Sustainability, social responsibility and ethics are key core elements of the Griffith Business School curriculum.”

Griffith Centre for Sustainable Enterprise

https://bit.ly/3DwQTQ2



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

Prime Video

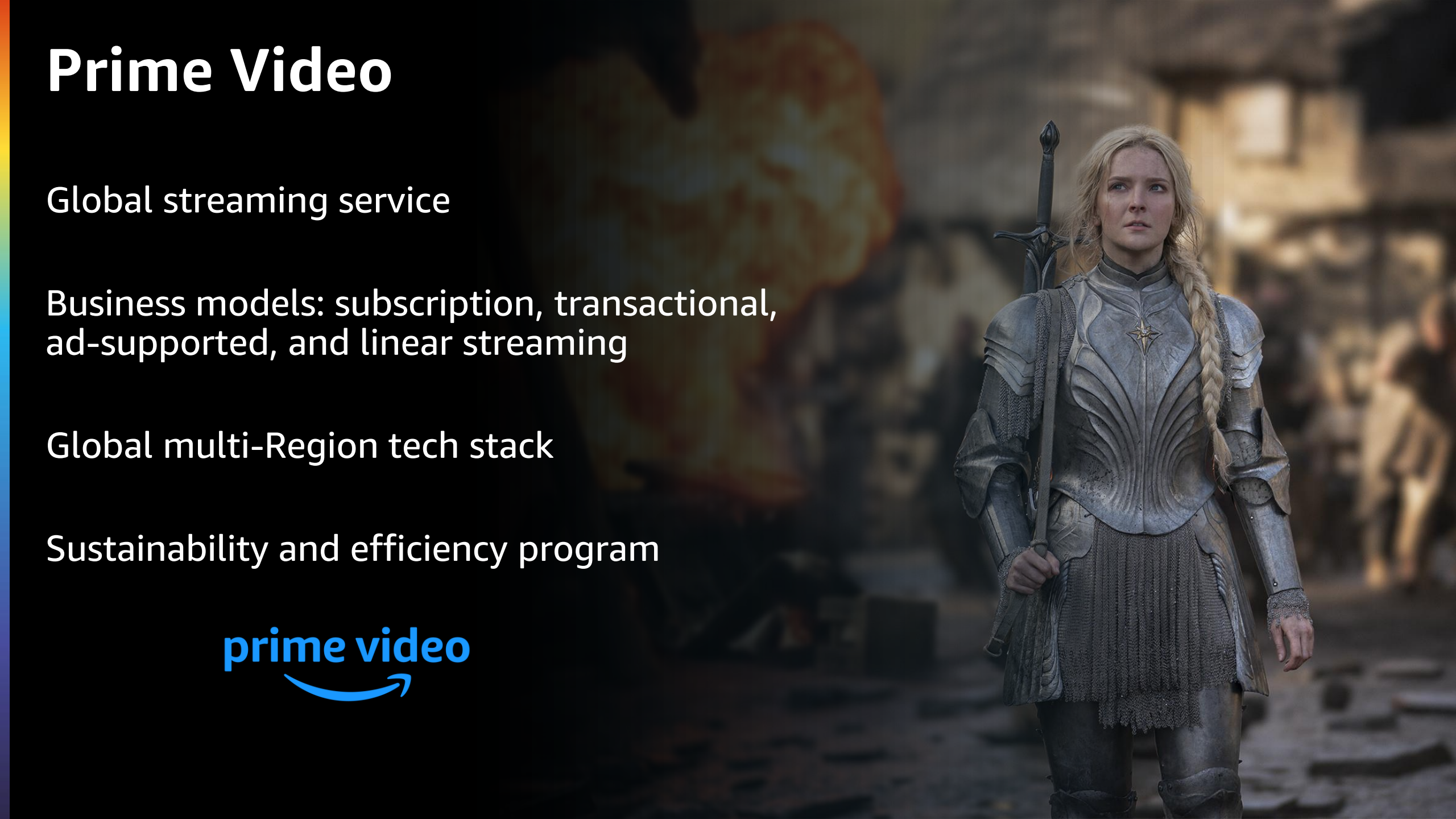
Global streaming service

Business models: subscription, transactional, ad-supported, and linear streaming

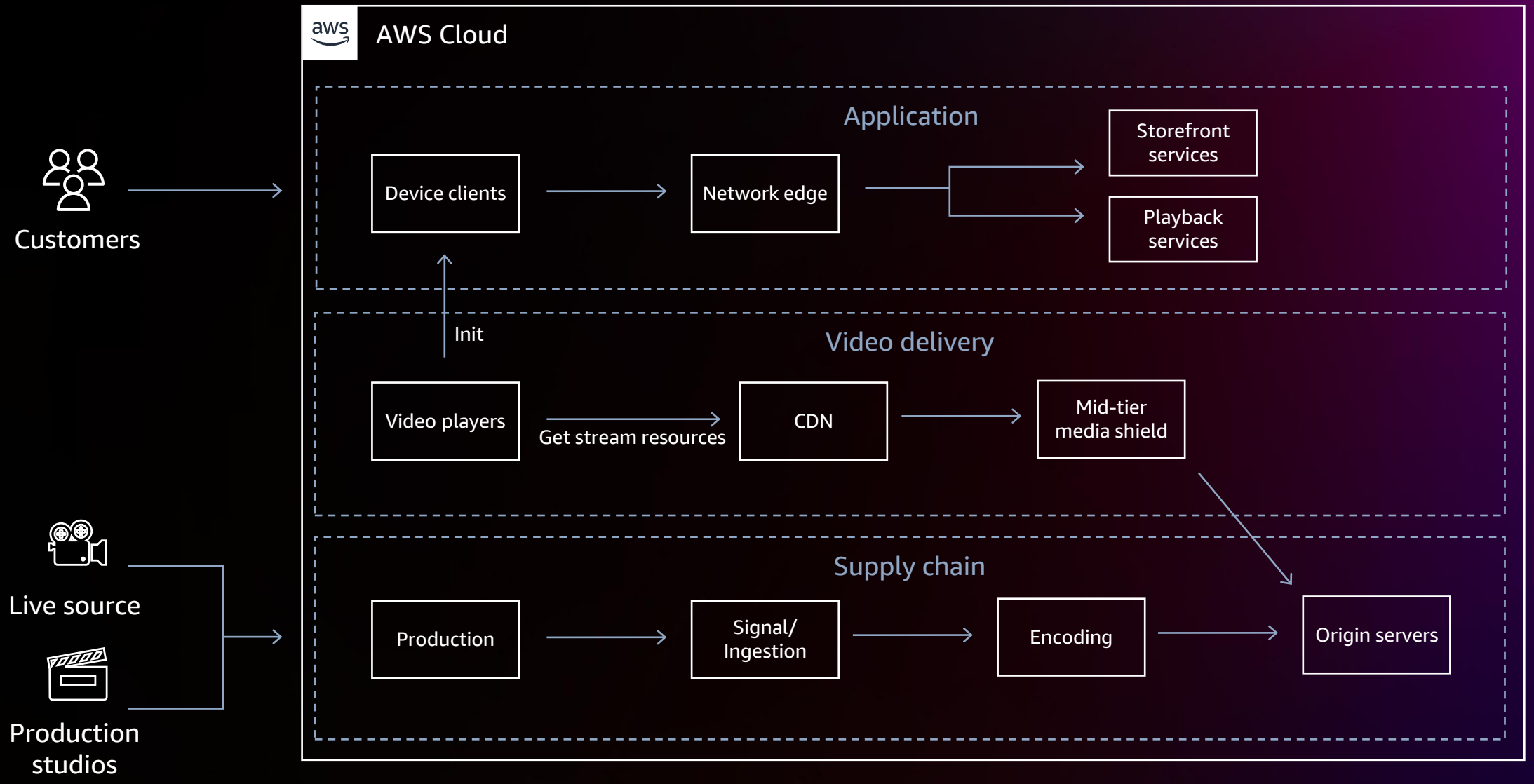
Global multi-Region tech stack

Sustainability and efficiency program

prime video



10,000-foot view



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

Building a carbon reduction program

Start with goal(s)

Baseline your metrics

Iterative improvement process



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

Start with goal(s)

Sustainability in the cloud

Encode optimizations

- Reduce bits on the wire by X% YoY

Infrastructure optimization

- Improve infrastructure efficiency by Y% YoY

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

Baseline your metrics

Desired output metric: AWS carbon footprint tool – metric tons of carbon dioxide equivalents per year for our AWS infrastructure

Proxy metrics (controllable inputs)

- Proxy metrics are closer to real time, and often more controllable
- Examples of proxy metrics:
  - Playback experience: infrastructure cost ($) per 1000 concurrent streams
  - Content delivery: delivery bandwidth (Gbps) per stream
  - Content discovery experience: Normalized Instance Hour (NIH) per 1000 page impressions
  - Customer acquisition: infrastructure cost ($) per subscription

aws

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

1 Identify targets for improvement

Pilot 1: Edge request termination

- Ingress point for all requests from client devices

Pilot 2: Video quality analysis

- Detect block corruption on a video feed
- Models became increasingly large and complex
- Goal: scale sustainably, by improving cost per analysis hour

2 Evaluate specific improvements

Pilot 1: Edge request termination

- AWS Compute Optimizer
- Migrate from c5.2xl to C6g (AWS Graviton2)



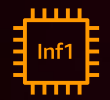
C6g instance



AWS Compute Optimizer

Pilot 2: Video quality analysis

- Benchmarking
- Migrate from G4dn to Inf1 (AWS Inferentia)



Inf1 instance

aws

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

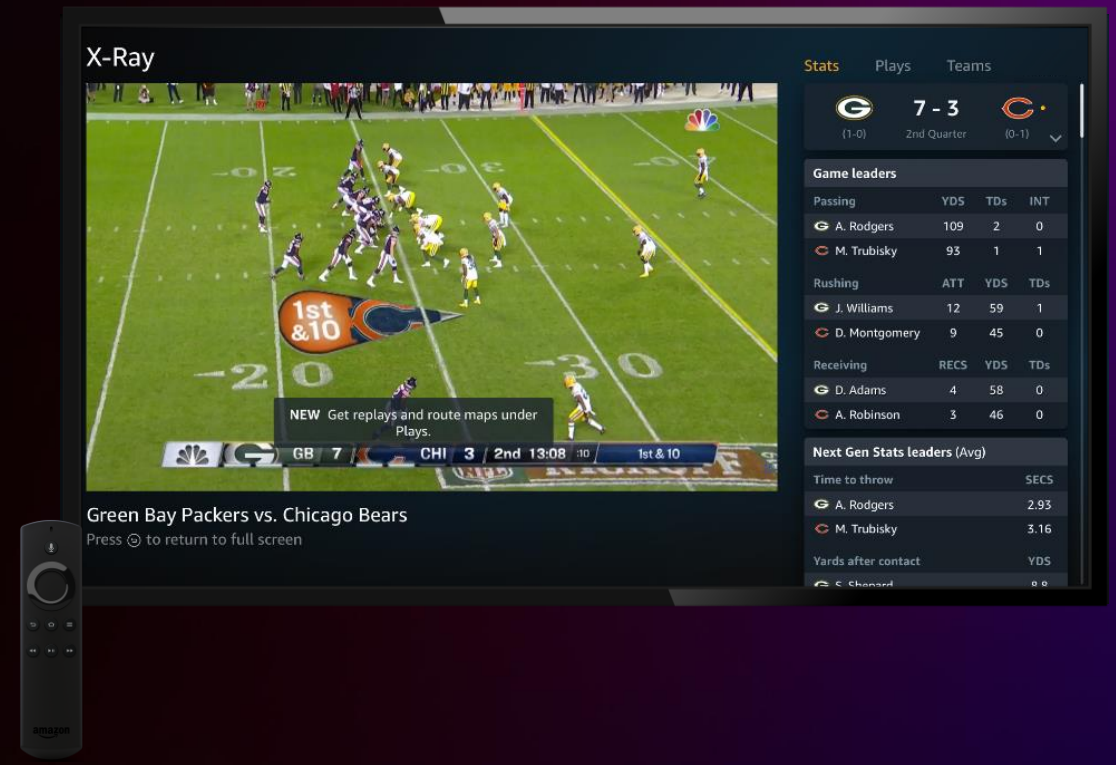
3 Prioritize and plan improvements

How do you get this work prioritized?

- Use your proxy metric
- Tie the outcome back to the business KPI

Define success criteria and plan sprint work

- Latency reduction
- Cost reduction
- Physical serving capacity
- Effort needed

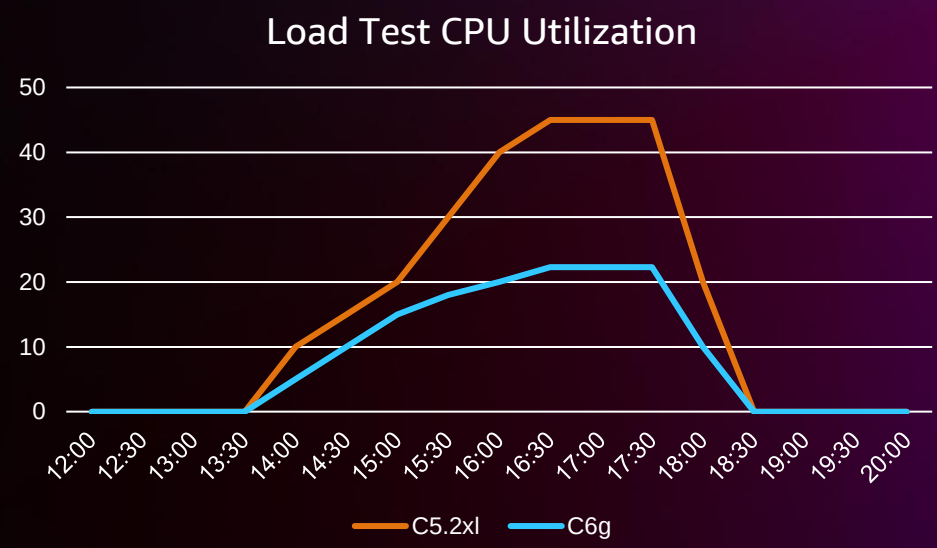


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

4 Test and validate improvements

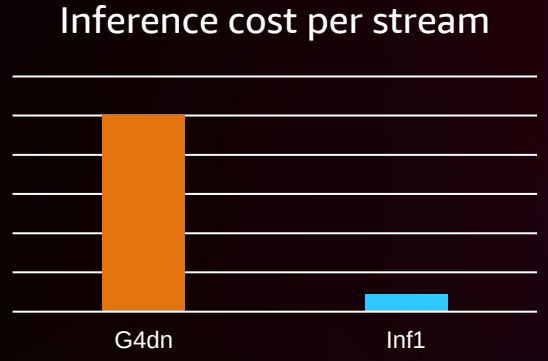
Pilot 1: Edge request termination

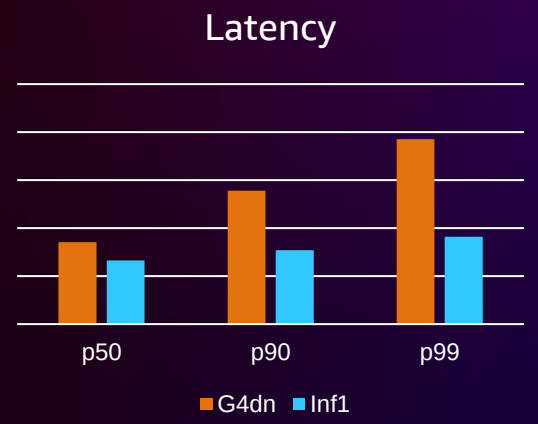
- Performance test under load
- 50% CPU performance improvement



Pilot 2: Video quality analysis

- Inference cost per concurrent streams
- Latency – p50, p90, p99





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

5 Deploy changes to production

Pilot 1: Edge request termination

- Temporary parallel pipeline
- Prod test in first Region before extending to others

Pilot 2: Video quality analysis

- Compile model
- Publish model artifact

aws

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

6 Measure results and replicate successes

Pilot 1: Edge request termination

- Compared to C5.2xl
  - Total cost reduction of 40% to run the global fleet

Pilot 2: Video quality analysis

- Compared to the previous GPU instance
  - Scaled to 10x more concurrent streams with 10% less cost

aws

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

NFL Thursday Night Football (TNF)

This year Prime Video launched an industry first – exclusive live streaming of the American Football, the NFL's TNF

Consider coupling infrastructure efficiency efforts with your project launches

TNF Top 20 and efficiency bar raisers

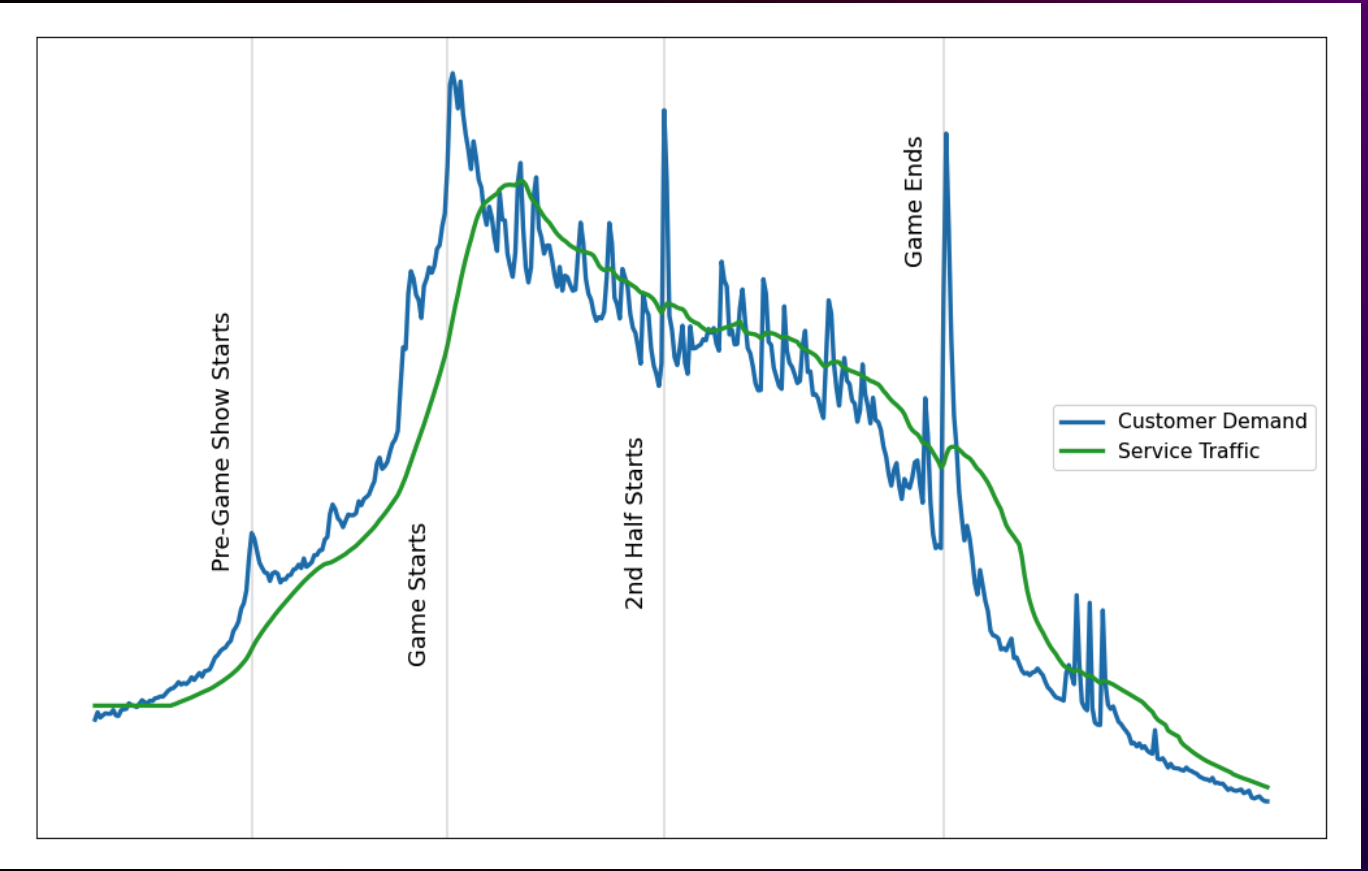


Scale with user load

Predictable, yet spikey demand patterns

Optimize your scaling profile around your demand constraints

Replicate success/learnings



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

“We started in a garage, but we’re not there anymore. We are big, we impact the world, and we are far from perfect. Our local communities, planet, and future generations need us to be better every day.”

Success and Scale Bring Broad Responsibility

Amazon Leadership Principle

aws

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

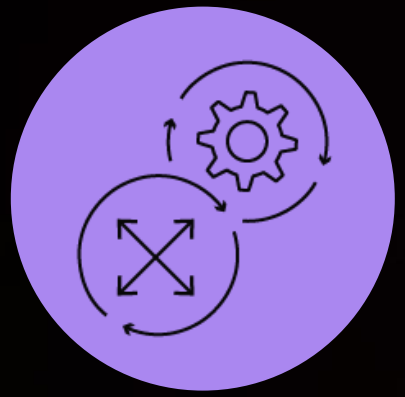
We’re on a journey together

NORMALIZE SUSTAINABLE PRACTICES

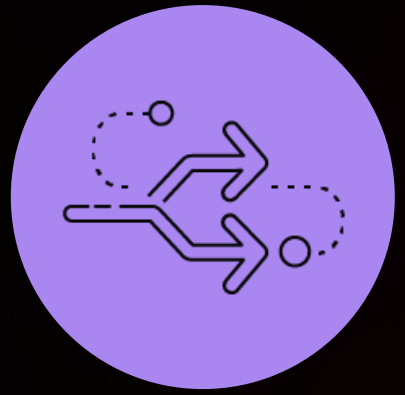


© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Call to action



How can my organization use proxy metrics as part of our overall strategy?



How can I create a culture of experimentation that supports our sustainability objectives?

© 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 chalk talks and workshops

SUS201-R: 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-R: Optimizing architectures for sustainability

aws

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

Thank you!



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

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