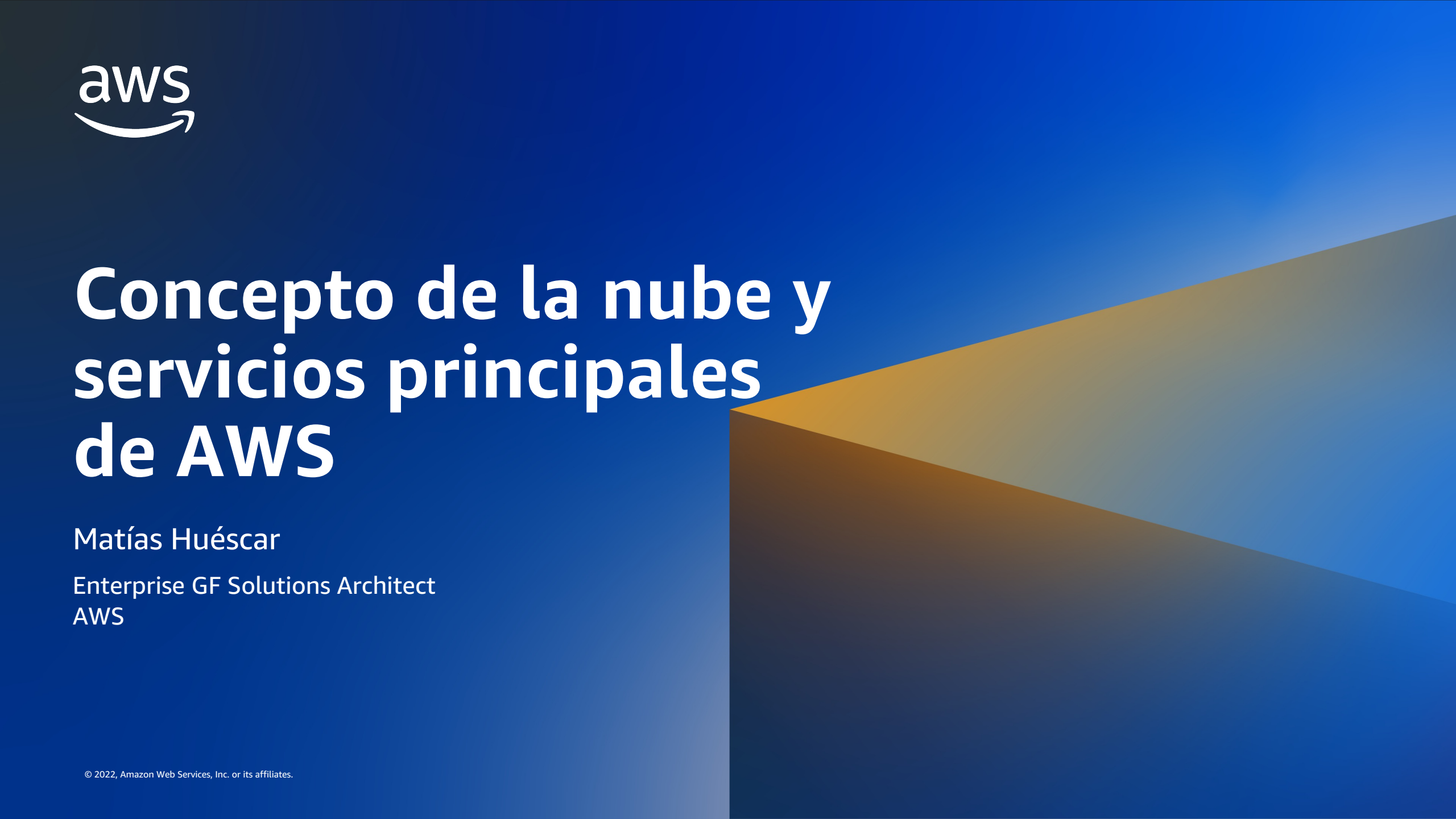


Concepto de la nube y servicios principales de AWS

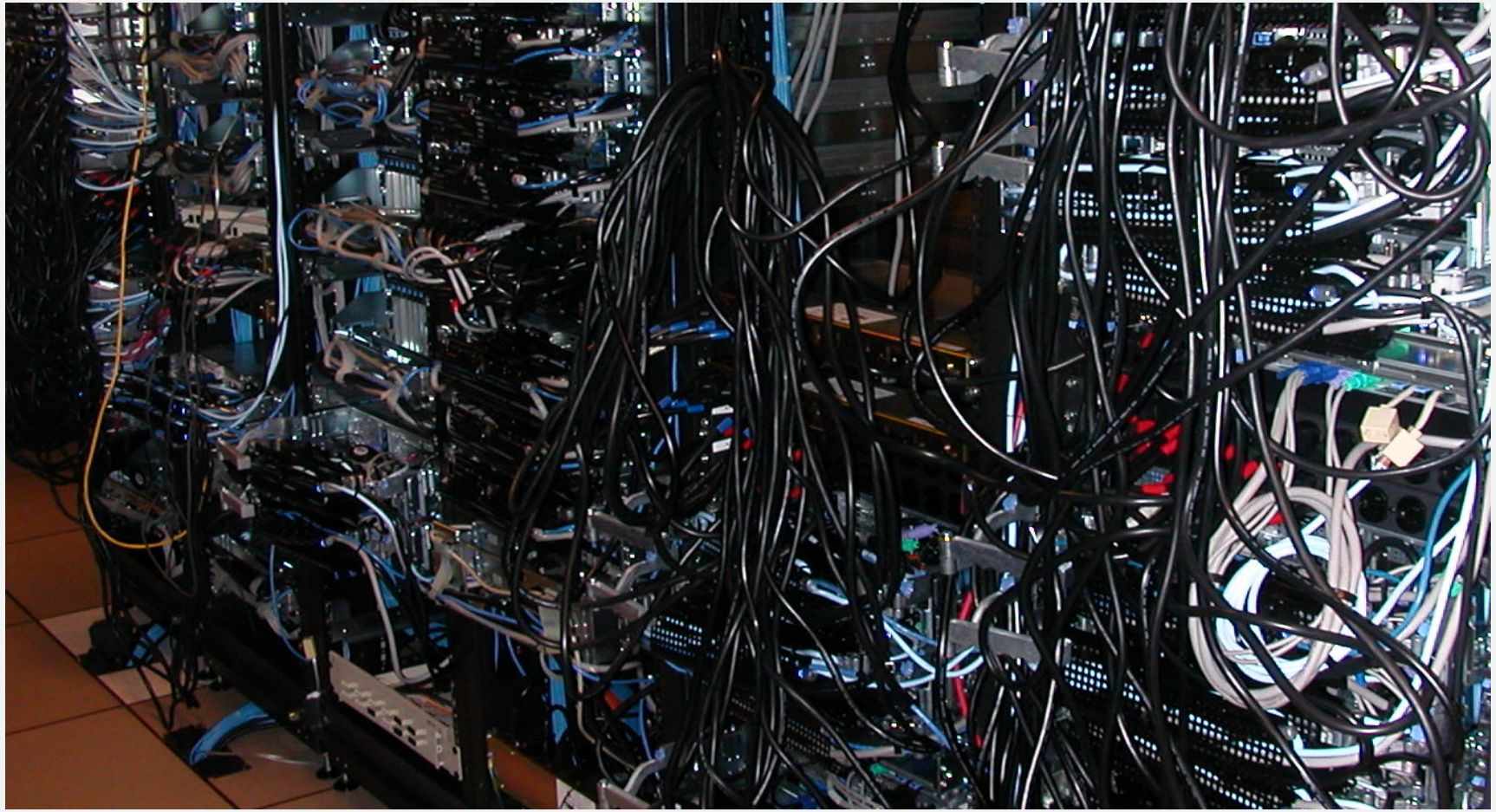
Matías Huéscar

Enterprise GF Solutions Architect
AWS

AGENDA

- What is cloud?
- Why cloud?
- AWS Cloud
- Some concepts in the cloud

What is the cloud?



What are the advantages of moving to the cloud?

Agility	Teams can experiment and innovate more quickly and frequently
Cost Savings	Only pay for what you use, lower upfront expenses
Elasticity	Easily Scale up or down with the needs of the business
Innovate Faster	Ability to focus on business differentiators, not infrastructure
Go Global in Minutes	Most extensive, reliable, and secure global cloud infrastructure

What is AWS?

A broad and deep platform that helps customers build sophisticated, scalable, secure applications

What sets AWS apart?

Security	Fine-grained control
Service Breadth & Depth	200+ fully featured services to support any cloud workload; rapid customer driven releases
Experience: 1M+ customers	Building and managing cloud since 2006
Global Footprint	96 Availability Zones within 30 geographic Regions, 21 Local Zone, 310+ Points of Presence in 245 contries and territories.
Machine Learning	More machine learning happens on AWS than anywhere else. Machine learning in the hands of every developer and data scientist.
Ecosystem	100,000 APN partners from 150 countries. The AWS Marketplace offers 50 categories, and more than 8,000 software listings
Enterprise leader	AWS positioned as a Leader in the Gartner Magic Quadrant for Cloud Infrastructure as a Service, Worldwide

AWS Well Architected Framework

Pilar de excelencia operativa

El pilar de la excelencia operativa se concentra en ejecutar y monitorear los sistemas y en mejorar constantemente los procesos y los procedimientos. Entre los temas clave se incluyen la automatización de cambios, la respuesta a eventos y la definición de estándares para administrar las operaciones diarias.

HTML | Laboratorios

Pilar de seguridad

El pilar de la seguridad se concentra en proteger la información y los sistemas. Entre los temas clave se incluyen la confidencialidad y la integridad de los datos, la administración de los permisos de usuarios y el establecimiento de controles para detectar eventos de seguridad.

HTML | Laboratorios

Pilar de fiabilidad

El pilar de fiabilidad se centra en las cargas de trabajo que realizan las funciones previstas y en cómo recuperarse rápidamente de los errores para cumplir con las demandas. Entre los temas clave se incluyen el diseño de sistemas distribuidos, la planificación de la recuperación y cómo adaptarse a los requisitos cambiantes.

HTML | Laboratorios

Pilar de eficacia del rendimiento

El pilar de eficacia del rendimiento se centra en la asignación estructurada y simplificada de TI y en los recursos informáticos. Entre los temas clave se incluyen la selección de los tipos y tamaños de recursos optimizados para los requisitos de la carga de trabajo, la supervisión del rendimiento y el mantenimiento de la eficacia a medida que evolucionan las necesidades de la empresa.

HTML | Laboratorios

Pilar de optimización de costos

El pilar de optimización de costos se centra en evitar gastos innecesarios. Entre los temas clave se incluyen la comprensión del tiempo dedicado y el control de la asignación de fondos, la selección de recursos para el tipo y la cantidad adecuados y el escalado para cumplir con las necesidades de la empresa sin gastos excesivos.

HTML | Laboratorios

Pilar de sostenibilidad

El pilar de sostenibilidad se centra en minimizar los impactos ambientales de ejecutar cargas de trabajo en la nube. Entre los temas clave se incluyen un modelo de responsabilidad compartida para la sostenibilidad, la comprensión del impacto y la maximización del uso para minimizar los recursos necesarios y reducir los impactos posteriores.

HTML | Laboratorios

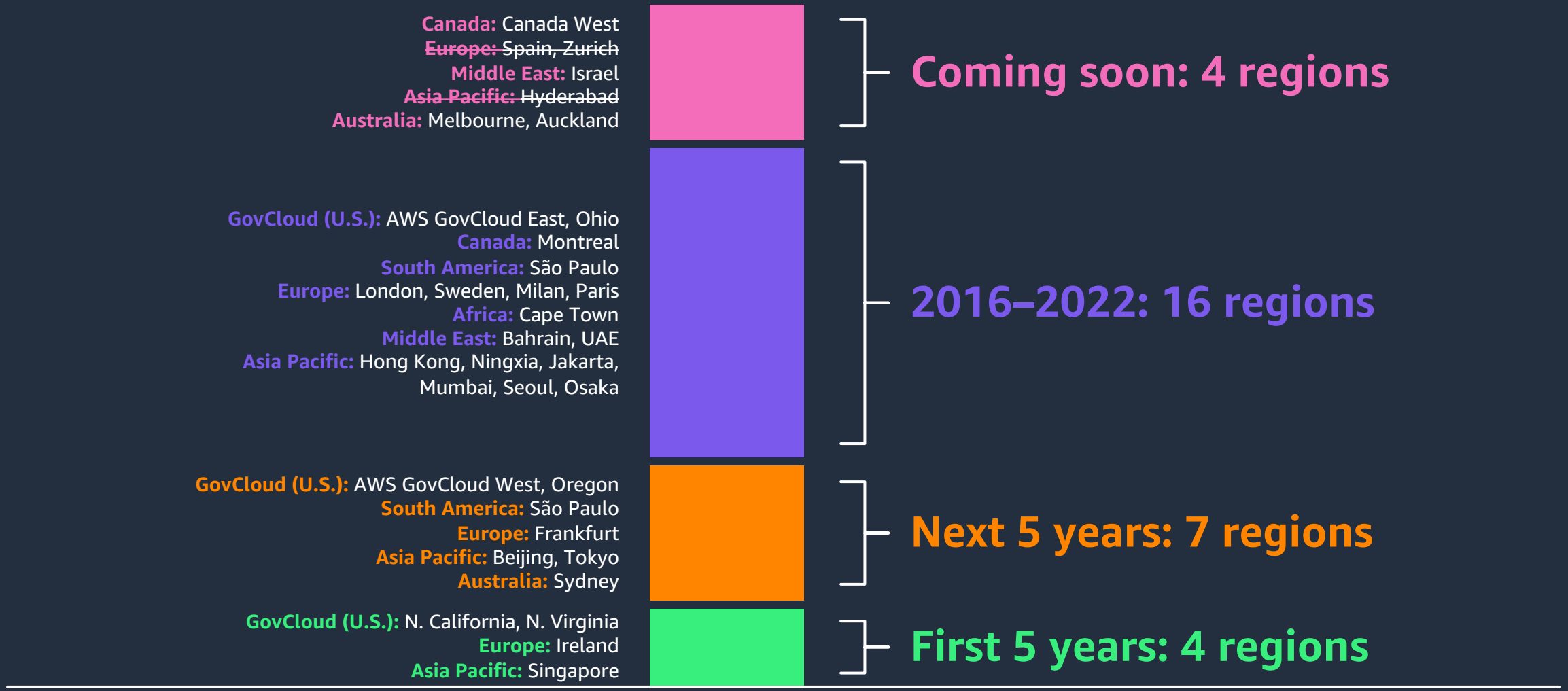
Customer Obsession

90%

of roadmap originates with customer requests
and are designed to meet specific needs

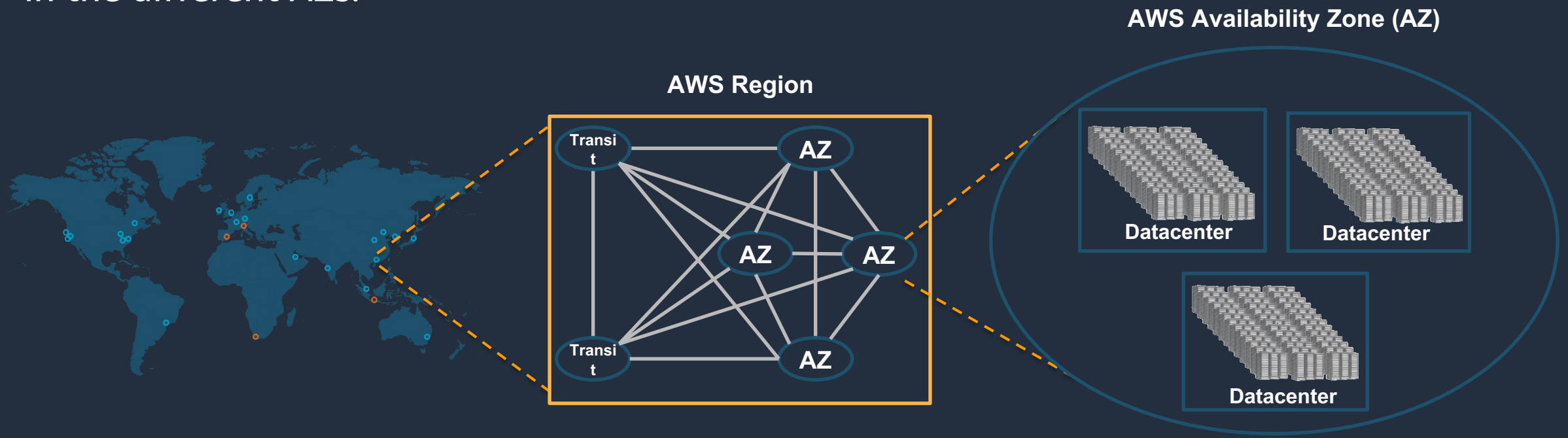
AWS Infrastructure

Regional Expansion



AWS region design

AWS Regions are comprised of multiple AZs for **high availability, high scalability,** and high **fault tolerance.** Applications and data are replicated in real time and consistent in the different AZs.

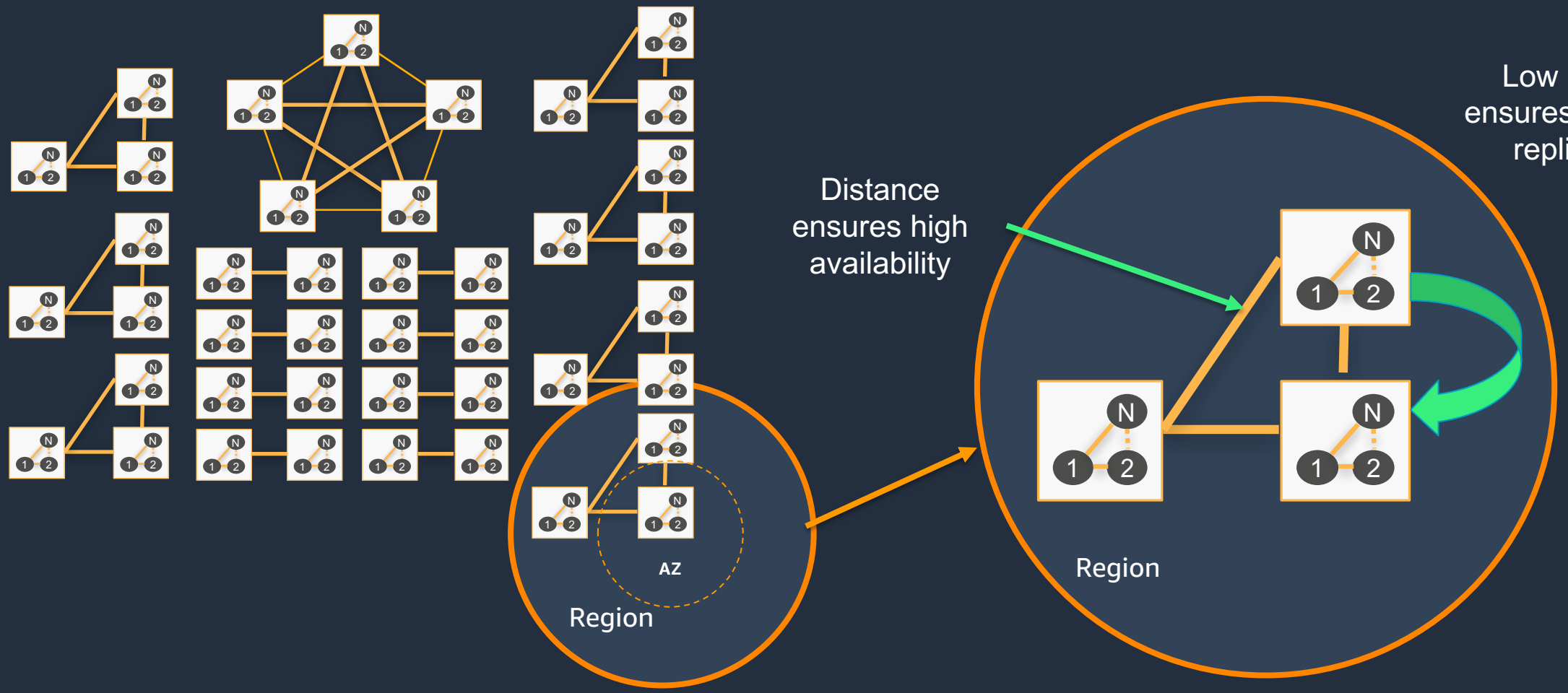


A Region is a physical location in the world where we have multiple **Availability Zones.**

Availability Zones consist of one or more discrete data centers, each with redundant power, networking, and connectivity, housed in separate facilities.

Regions and availability zones

SAME WORDS DIFFERENT MEANINGS



AWS Availability Zone (AZ) design

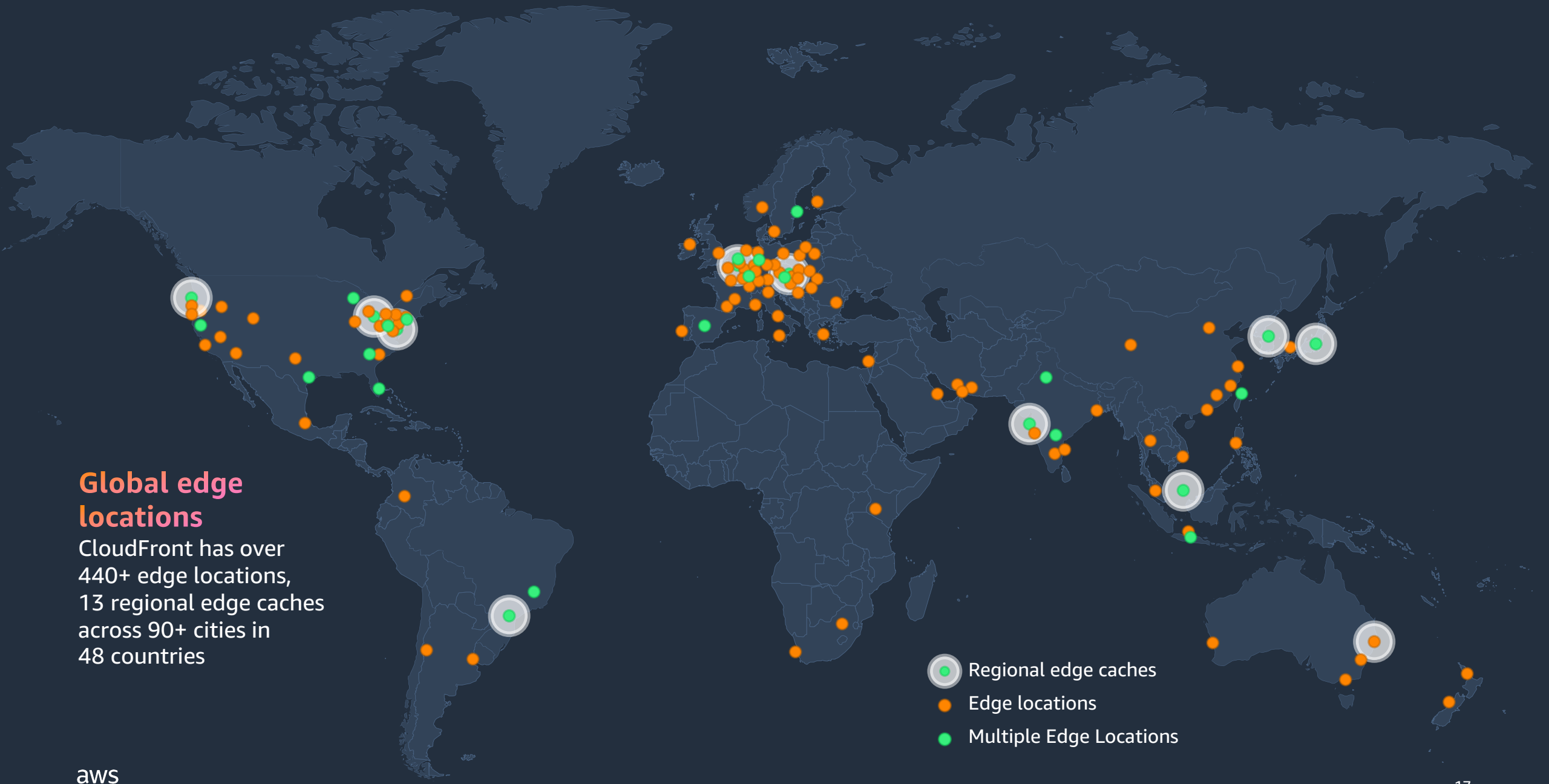
Fully isolated infrastructure with one or more datacenters

Meaningful distance of separation

Unique power infrastructure

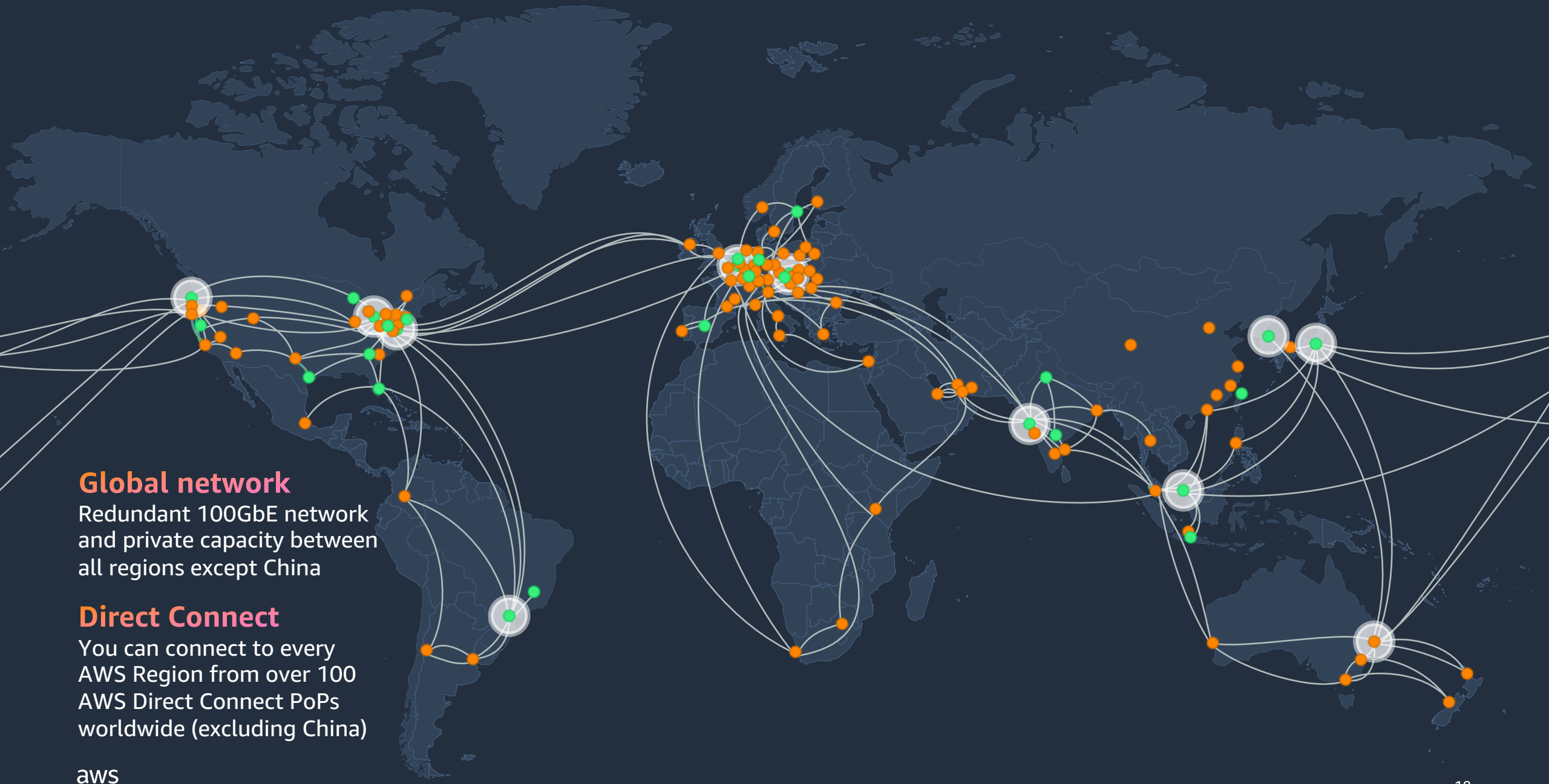
Many 100Ks of servers at scale

Datacenters connected via fully redundant and isolated metro fiber



Global edge locations

CloudFront has over 440+ edge locations, 13 regional edge caches across 90+ cities in 48 countries



Global network

Redundant 100GbE network and private capacity between all regions except China

Direct Connect

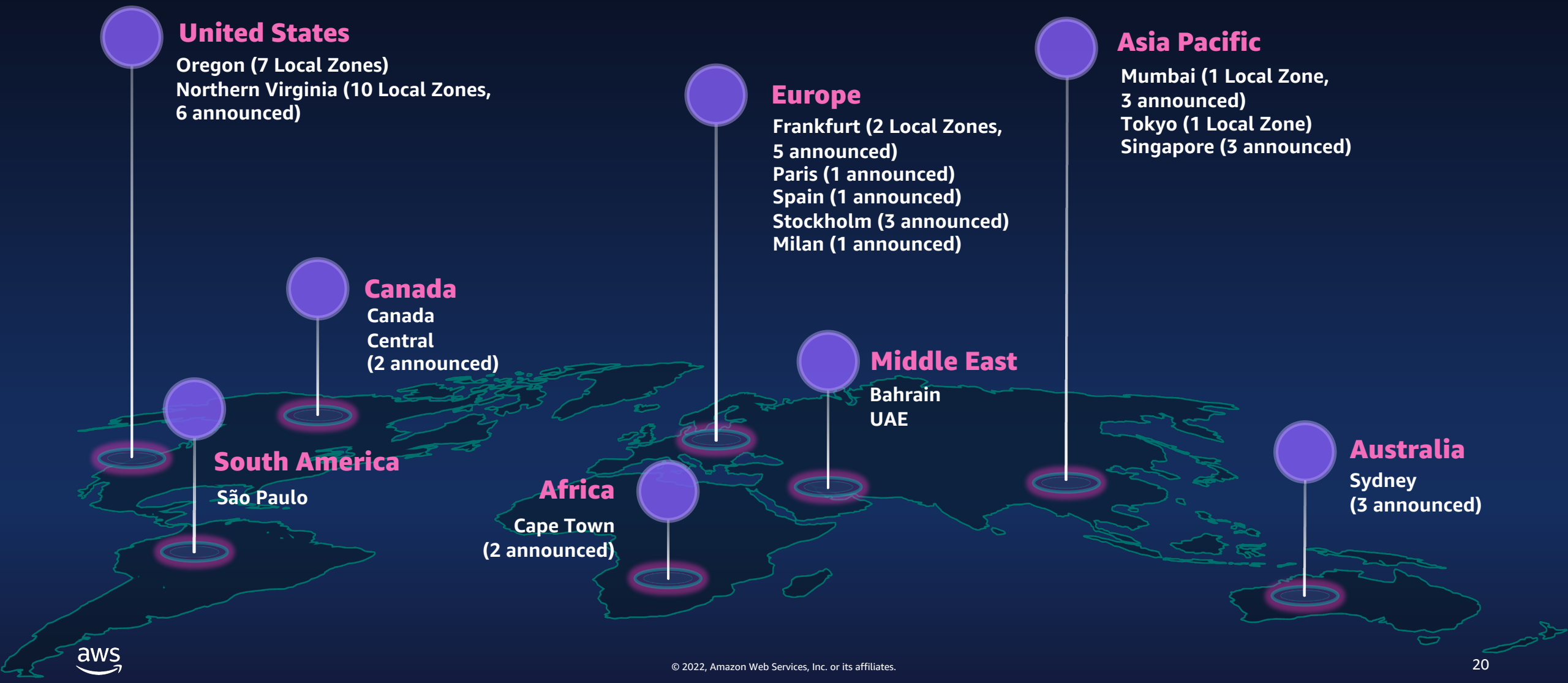
You can connect to every AWS Region from over 100 AWS Direct Connect PoPs worldwide (excluding China)

aws

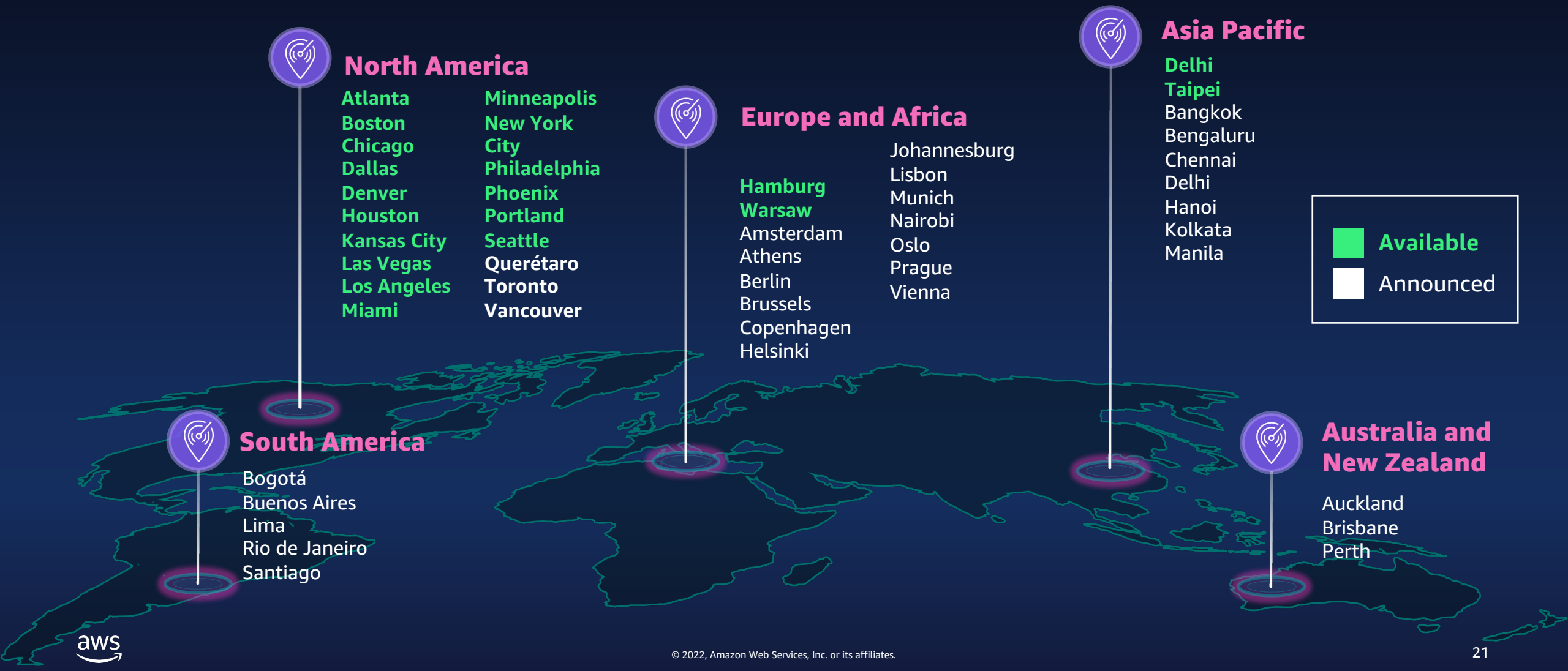
The upcoming AWS Region in Spain is an extension to our commitment to Iberia

- Opened
- 3 Availability Zones at launch
- Eighth AWS Region in Europe, joining existing regions in Dublin, Frankfurt, London, Paris, Stockholm, and Milan and upcoming one in Zurich

Available Regions and Local Zones



33 New Local Zones generally available starting 2022



Further and faster, together

Commitment to a sustainable future

On September 19, 2019, Amazon and Global Optimism announced The Climate Pledge, a commitment to meet the Paris Agreement 10 years early

Net-zero carbon by 2040

Path to 100% renewable energy by 2025

$2 billion Climate Pledge Fund

It’s more sustainable in the cloud

AWS’s infrastructure is

3.6x more energy efficient

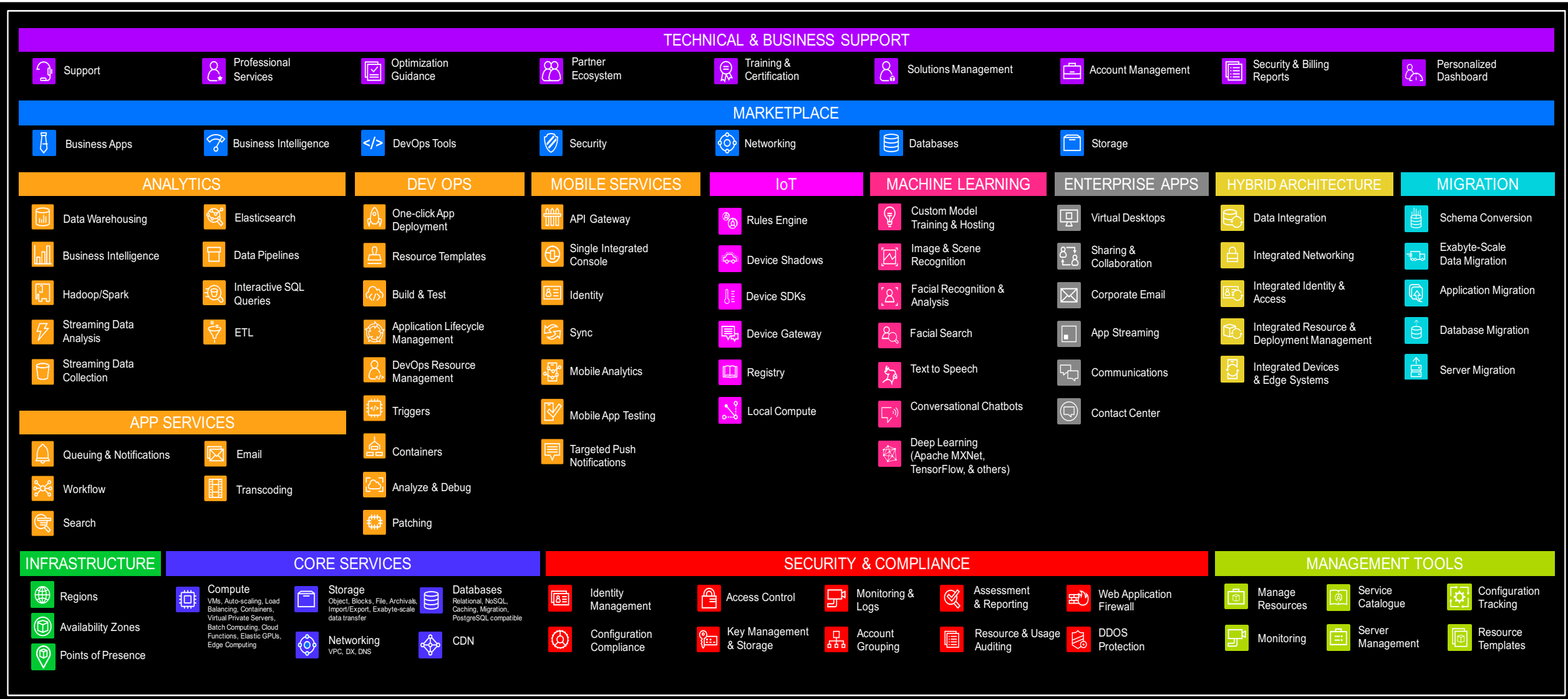
than the median of surveyed U.S. enterprise data centers

AWS performs the same task with an

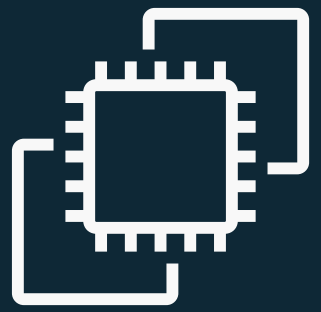
88% lower carbon footprint

Source: 451 Research, 2019, all rights reserved

Broad and Deep Functionality



Choices for Compute



Amazon EC2

Virtual server instances
in the cloud

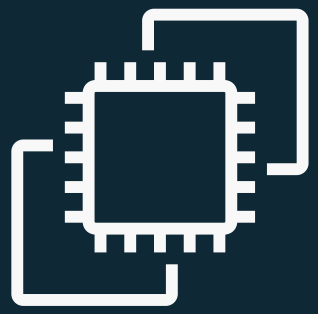
Amazon ECS, EKS, and Fargate

Container management service
for running
Docker on a managed
cluster of EC2

AWS Lambda

Serverless compute
for stateless code execution in
response to triggers

Amazon EC2



Amazon EC2

Linux | Windows | Mac

ARM and x86 architectures

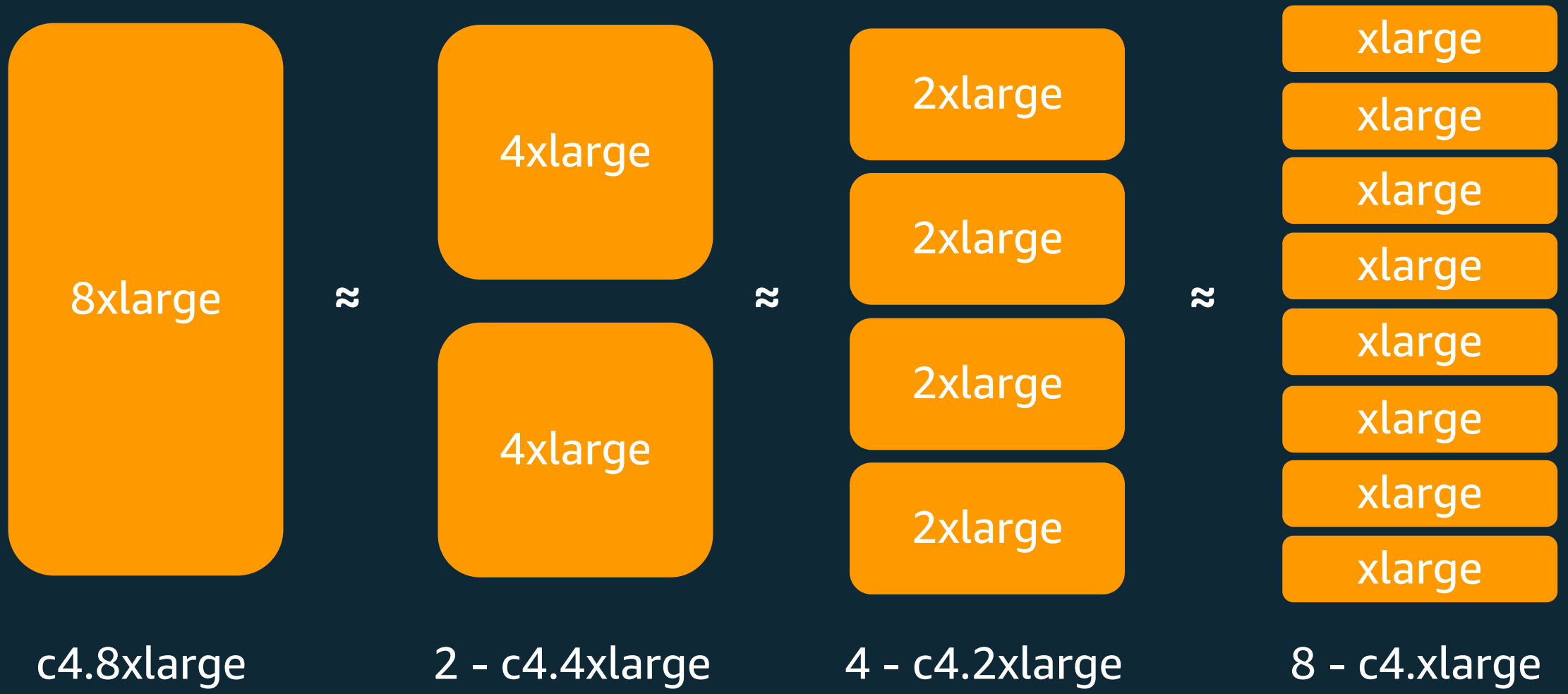
General purpose and workload optimized

Bare metal, disk, networking capabilities

Packaged | Custom | Community AMIs

Multiple purchase options: On-demand, RI, Spot

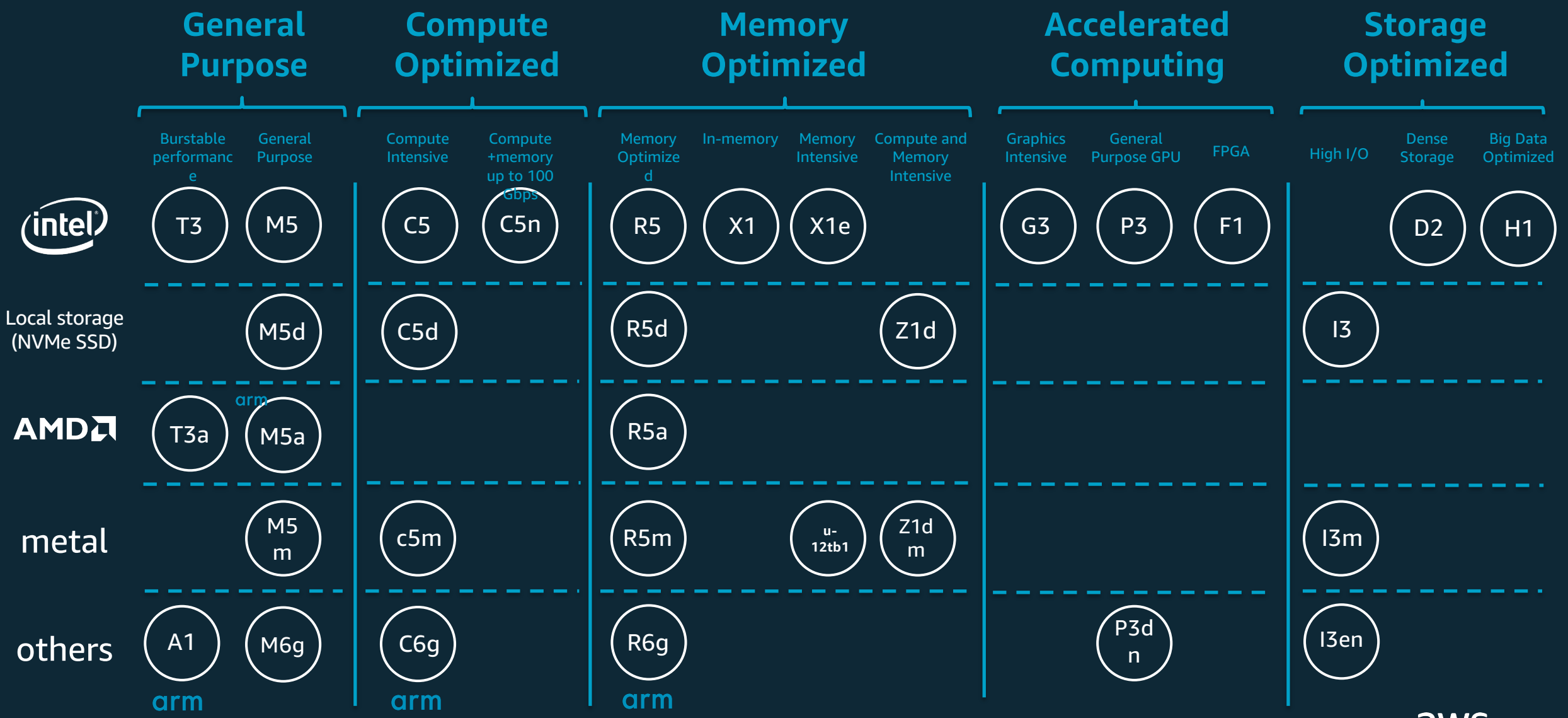
Instance sizing



c4.8xlarge | 2 - c4.4xlarge | 4 - c4.2xlarge | 8 - c4.xlarge

EC2 Naming Explained

Instance Types



Choose your processor and architecture

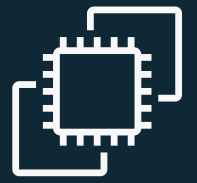
Intel® Xeon® Scalable (Skylake) processor

NVIDIA V100 Tensor Core GPUs

AMD EPYC processor

AWS Graviton Processor (arm)

AWS Inferentia



FPGAs for custom hardware acceleration

Right compute for the right application and workload

aws

AWS Graviton3 Processor

Enabling the best price/performance for your cloud workloads

Graviton2 Processor

7x performance, 4x compute cores, and 5x faster memory

Built with 64-bit Arm Neoverse cores with AWS-designed silicon using 7 nm manufacturing technology

Up to 64 vCPUs, 25 Gbps enhanced networking, 18 Gbps EBS bandwidth

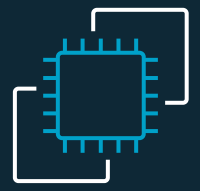
Graviton3 Processor

25% Faster than Graviton2, 2x faster floating point performance and cryptographic workloads, 3x faster for ML workloads

First in the cloud to feature DDR5, 50% higher memory bandwidth Graviton 2.

Uses up to 60% less energy. 20% higher enhanced networking bandwidth

Choice of accelerators for specialized workloads



Elastic Graphics

Easily add graphics acceleration to your EC2 instance

Configure right amount of graphics acceleration for your workload

Accelerate application for fraction of cost of standalone graphics instances

Elastic Inference

Reduce deep learning inference costs by up to 75%

Easily attach fractional sizes of a full GPU instance to EC2 or SageMaker instances

Scale inference acceleration up or down as needed with EC2 Auto Scaling

Broadest and deepest platform choice

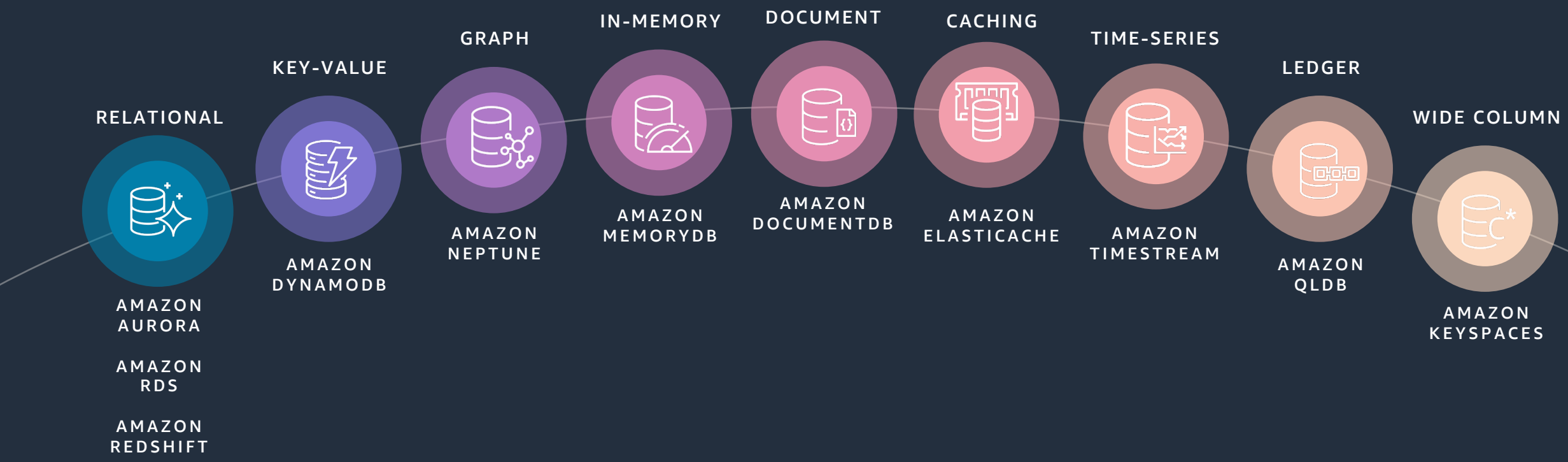
Categories	Capabilities	Options
General purpose	Choice of processor (AWS, Intel, AMD)	Elastic Block Store
Burstable	Fast processors (up to 4.5 GHz)	Elastic Inference
Compute intensive	High memory footprint (up to 24 TiB)	Elastic Graphics
Memory intensive	Instance storage (HDD and SSD)	Linux, Unix, Windows, macOS
Storage (High I/O)	Accelerated computing (GPUs ,FPGA & ASIC)	
Dense storage	Networking (up to 400 Gbps)	
GPU compute	Bare Metal	
Graphics intensive	Size (Nano to 32xlarge)	

475

instance types

for virtually every workload and business need

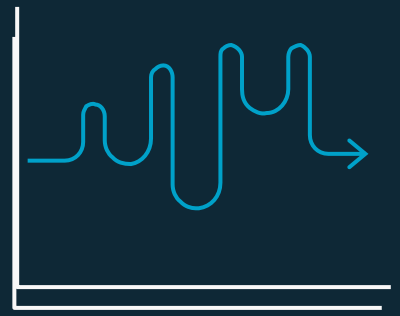
Broadest portfolio of purpose-built databases



Amazon EC2 purchase options

On-Demand

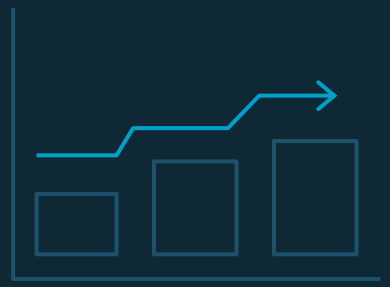
Pay for compute capacity by **the second** with no long-term commitments



Spiky workloads, to define needs

Reserved Instances

Make a 1 or 3 year commitment and receive a **significant discount** off On-Demand prices



Committed and steady-state usage

Savings Plan

Same great discounts as Amazon EC2 RIs with **more flexibility**

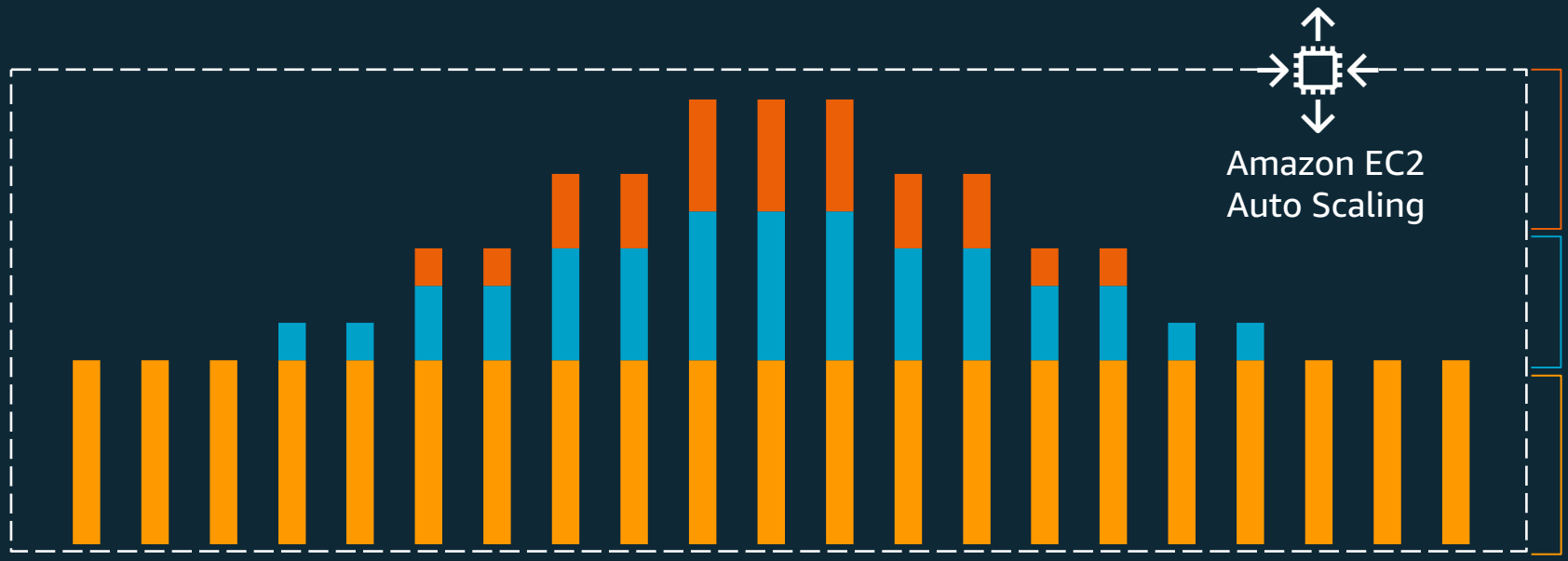
Committed flexible access to compute

Spot Instances

Spare Amazon EC2 capacity at **savings of up to 90%** off On-Demand prices

Fault-tolerant, flexible, stateless workloads

Simplifying capacity and cost optimization



Scale using Spot for short-lived, fault-tolerant, stateless workloads

Scale using On-Demand for short-lived, stateful workloads

Use Ris and Savings Plans for known/steady-state workloads

AWS SERVICES MAKE THIS EASY AND EFFICIENT

- Amazon Elastic Container Service (Amazon ECS)
- EC2 Fleet
- Amazon Elastic Kubernetes Service (Amazon EKS)
- AWS Thinkbox
- Amazon EMR
- AWS CloudFormation
- AWS Batch
- Amazon SageMaker

Motivate your teams with ***AWS Certification***

89% of IT decision makers agree that employing AWS Certified staff has shortened the time it takes to complete cloud projects[1].

Validate skills

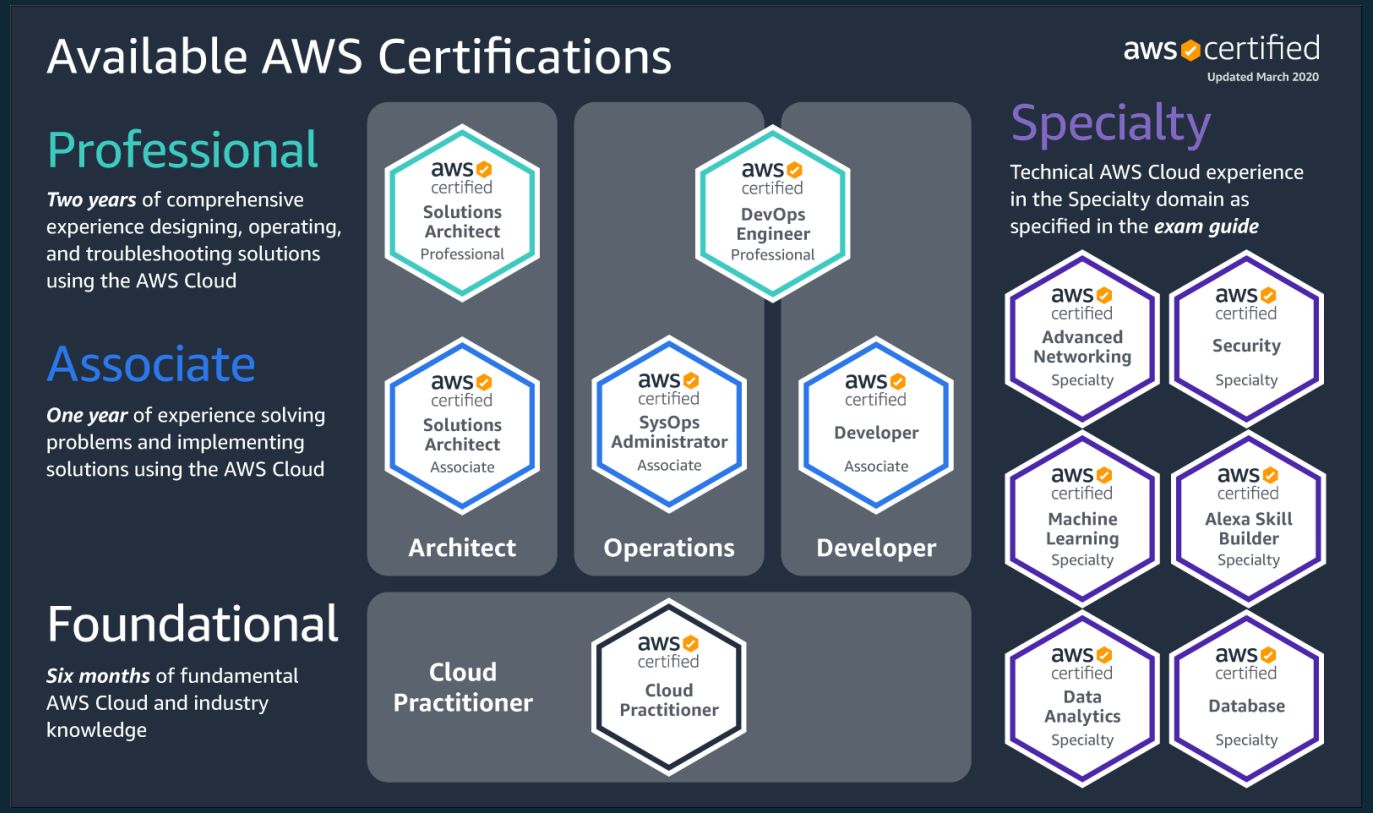
AWS Certifications help you confidently identify professionals with specific skills

Effective

AWS Certified professionals are more productive and faster troubleshooters[2]

Globally recognized

Employees value industry-recognized cloud certifications



1. 2020, ESG
2. Global Knowledge, 2019

AWS delivers broadest storage portfolio in industry

OBJECT

Amazon S3

BLOCK

Amazon EBS

FILE

Amazon EFS

Amazon FSx for NetApp ONTAP

Amazon FSx for Windows File Server

Amazon FSx for Lustre

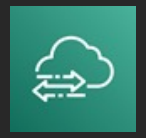
Amazon FSx for OpenZFS

BACKUP

AWS Backup

DATA TRANSFER AND MIGRATION

AWS Storage Gateway



AWS DataSync

AWS Transfer Family

AWS Snowball

AWS Snowcone

Your choice of Amazon S3 storage classes

S3 Intelligent-Tiering

S3 Standard

S3 Standard-IA

New

S3 Glacier Instant Retrieval

S3 Glacier Flexible Retrieval (formerly S3 Glacier)

S3 Glacier Deep Archive

S3 One Zone-IA

S3 Outposts

AWS Region ≥ 3 Availability Zones

Data with changing access patterns

- Milliseconds access
- No retrieval charge
- Object monitoring charge
- **Archive Instant Access tier** New
- Opt-in Async Archive tiers

Frequently accessed data

- Milliseconds access

Infrequently accessed data

- Milliseconds access
- Retrieval charge per-GB

Rarely accessed data

- **Milliseconds access**
- **Minimum storage duration**
- Retrieval charge per-GB

Archive data

- Retrieval options from minutes to hours
- **Free bulk retrievals** New
- Retrieval charge per-GB

Long-term archive data

- Retrieval in hours
- Retrieval charge per-GB

AWS AZ

Re-creatable, less accessed data

- Milliseconds access
- Retrieval charge per-GB

AWS Outposts

On-premises data

- Milliseconds access
- Retrieval charge per-GB

aws

You need the right storage for **any data**

#1
OBJECT

#1
BLOCK

#1
FILE

Amazon
S3

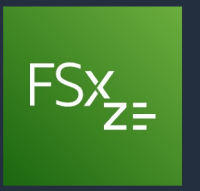
Amazon
EBS

Amazon
EFS

Amazon FSx
for NetApp
ONTAP

Amazon FSx
for Windows
File Server

Amazon FSx
for Lustre



Amazon FSx
for OpenZFS

Most complete cloud storage portfolio

Some Cloud Concepts

Shared responsibility model

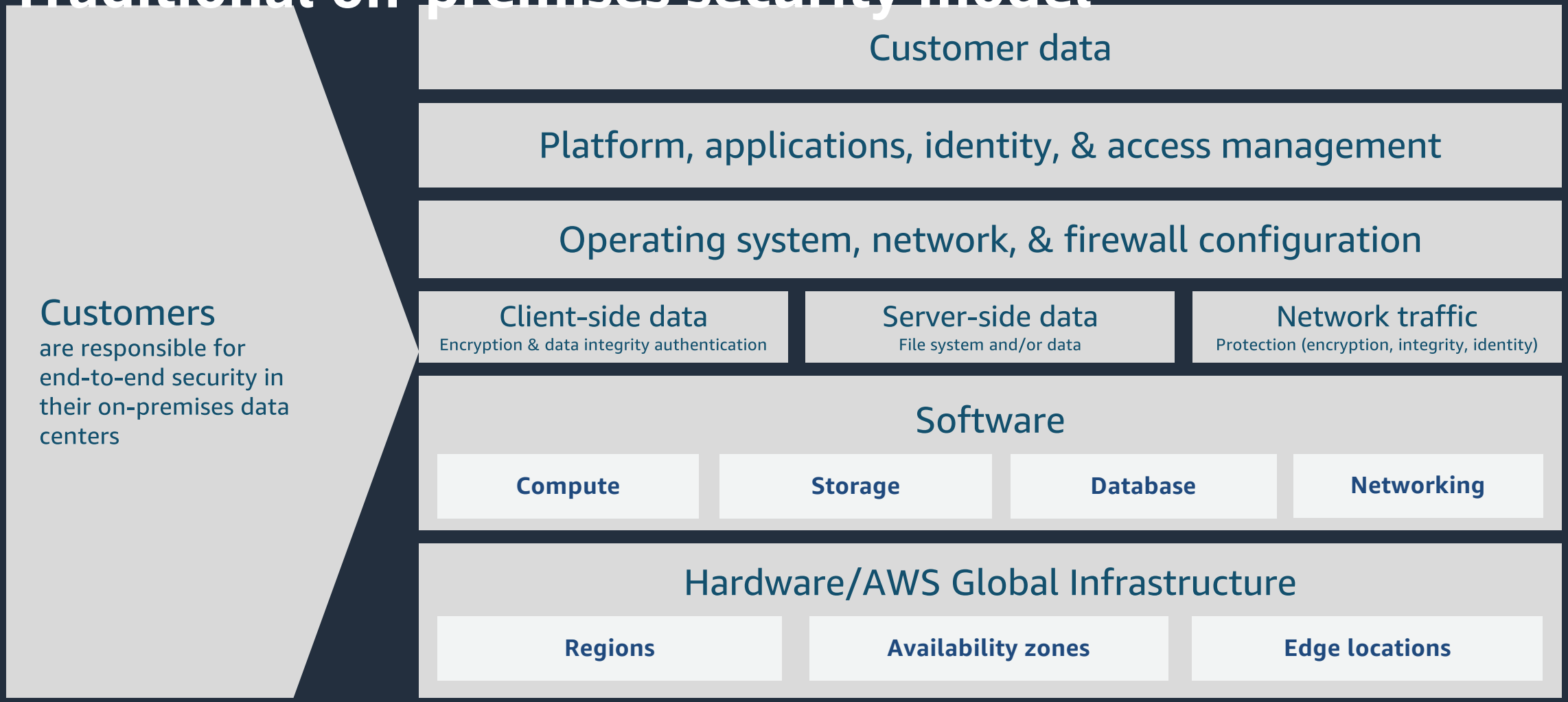
Security IN the Cloud

Customer responsibility will be determined by the AWS Cloud services that a customer selects

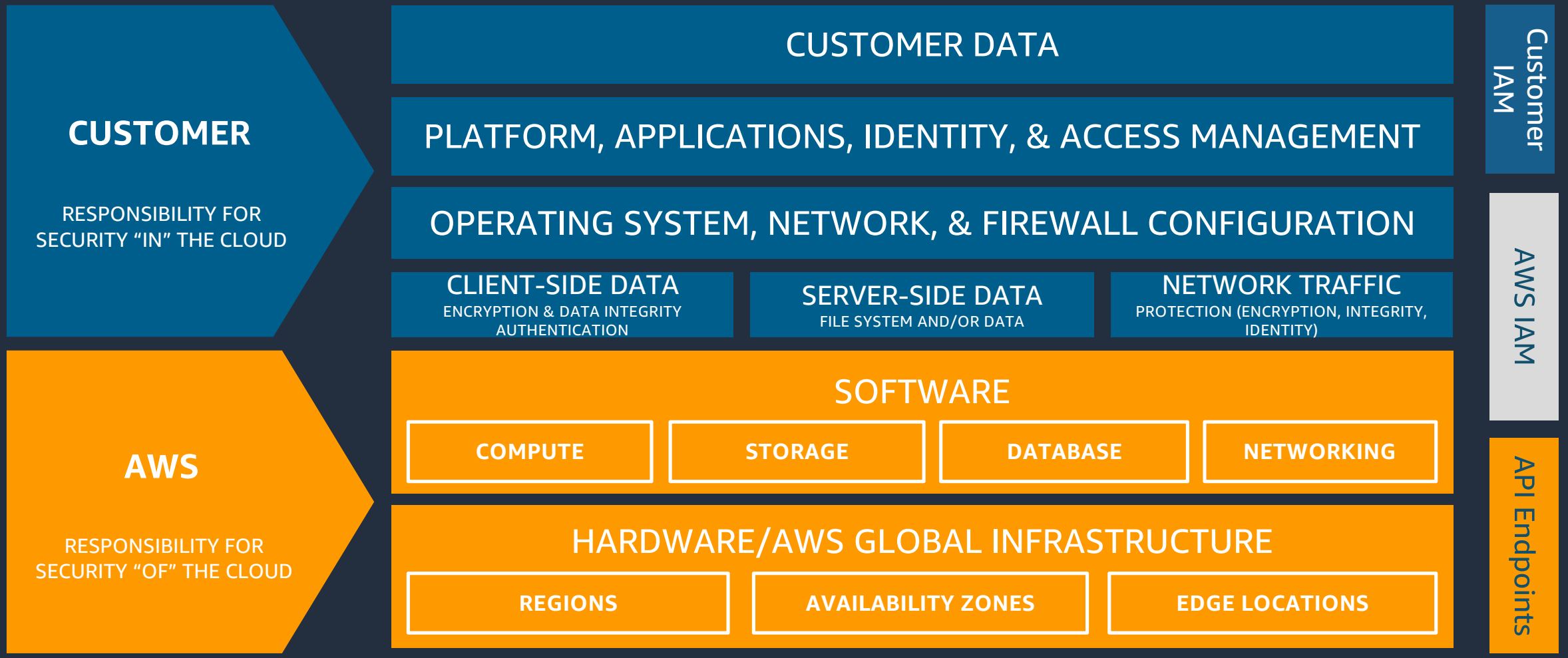
Security OF the Cloud

AWS is responsible for protecting the infrastructure that runs all of the services offered in the AWS Cloud

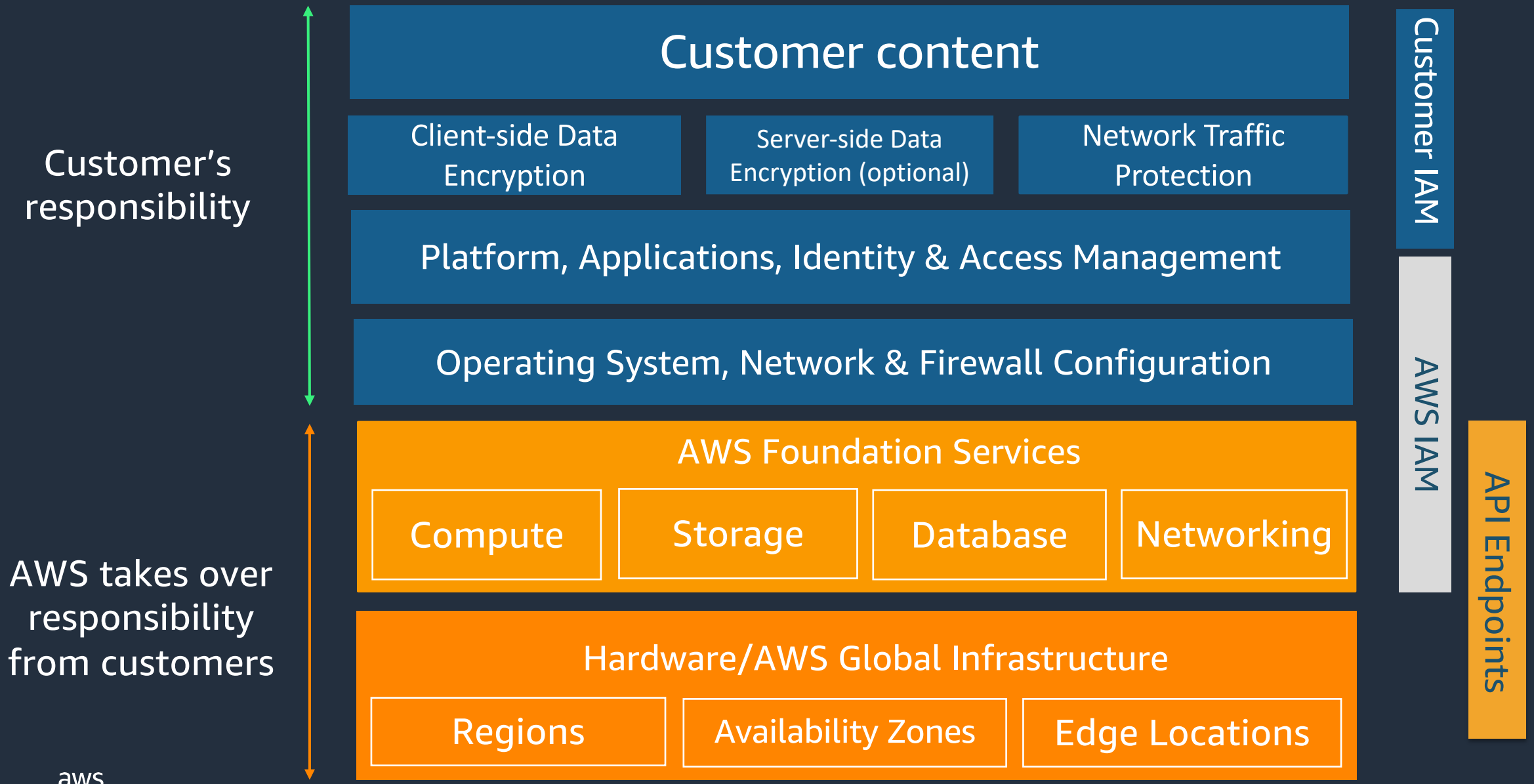
Traditional on-premises security model



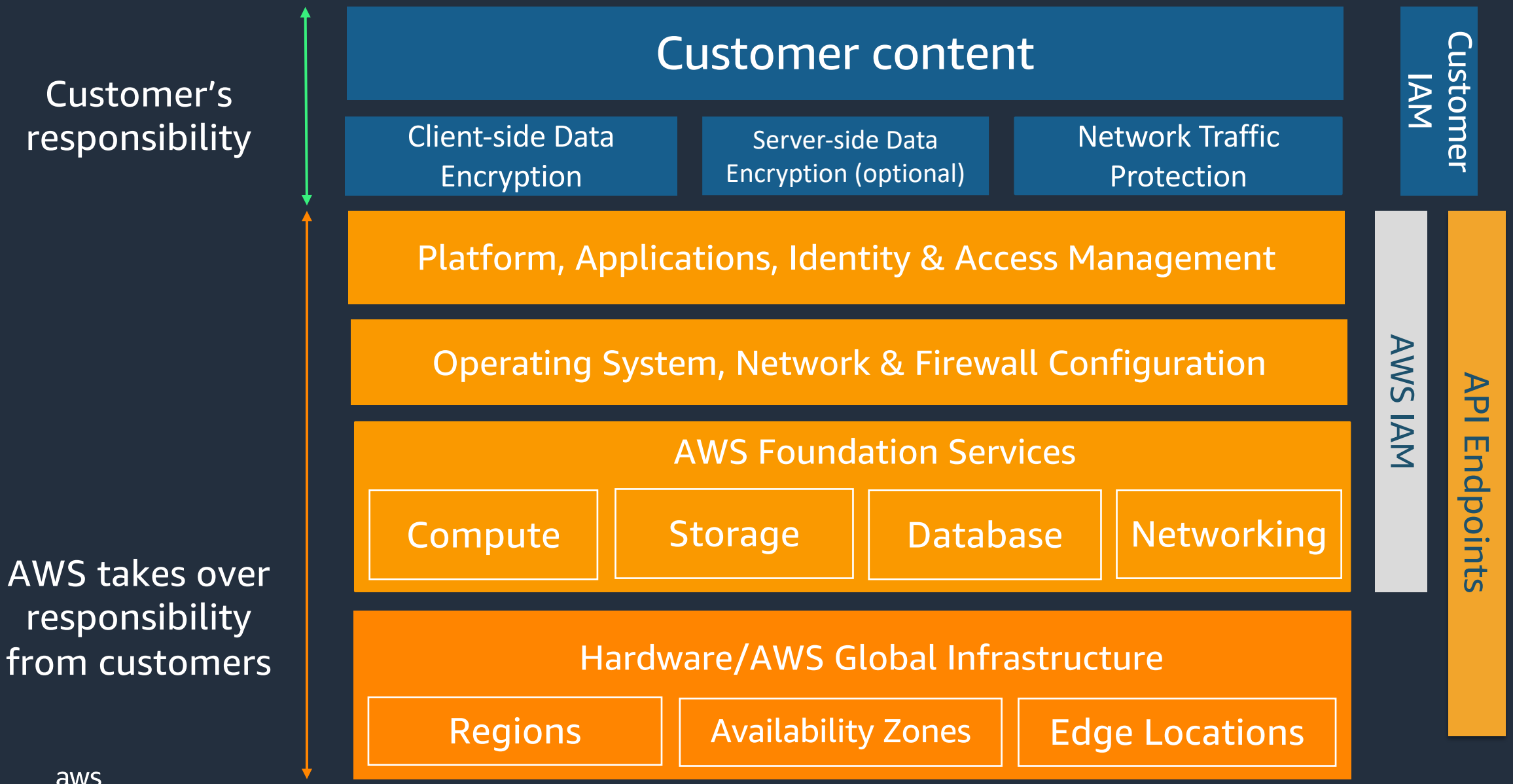
AWS Security Model when using Infrastructure Services



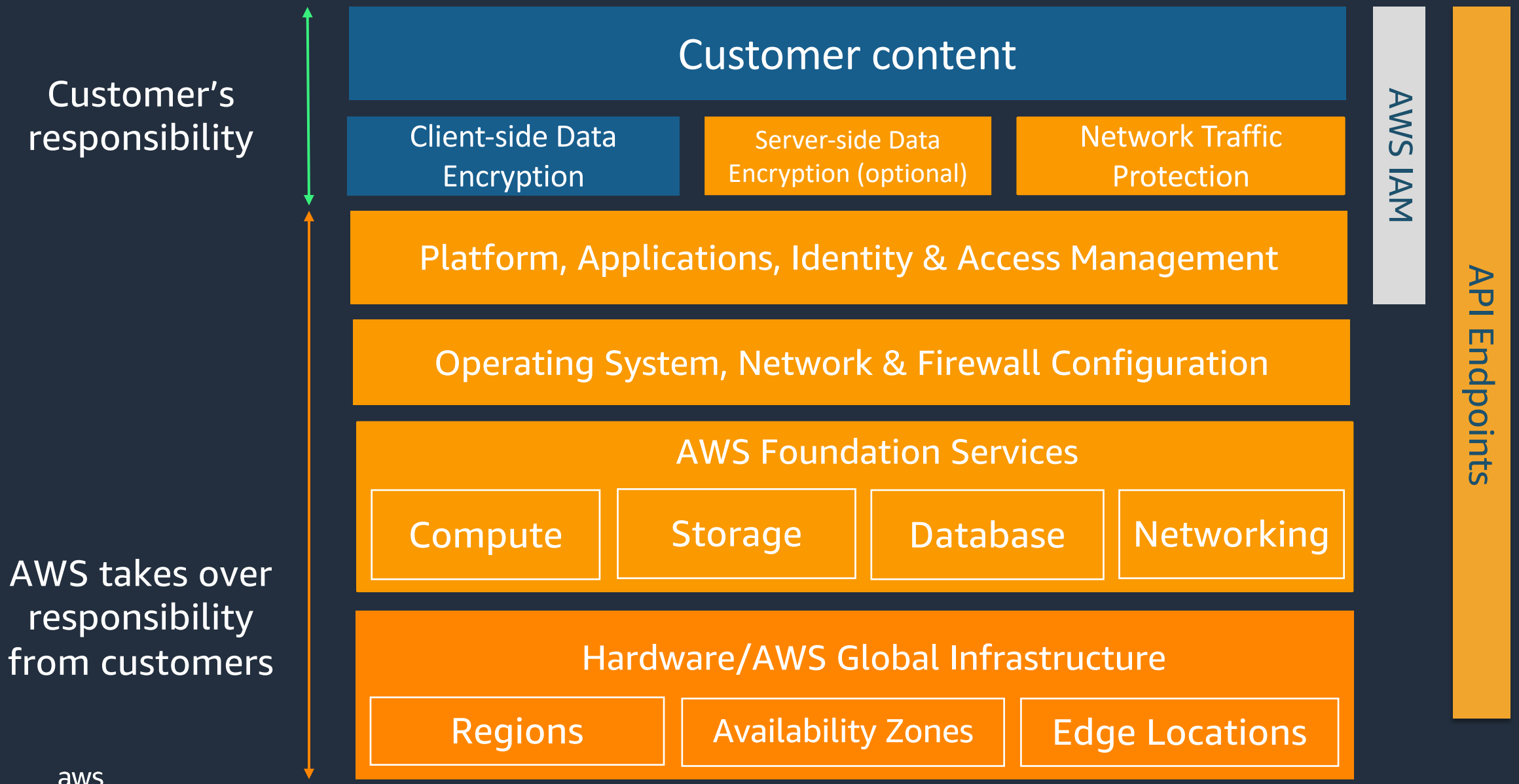
AWS Security Model when using Infrastructure Services



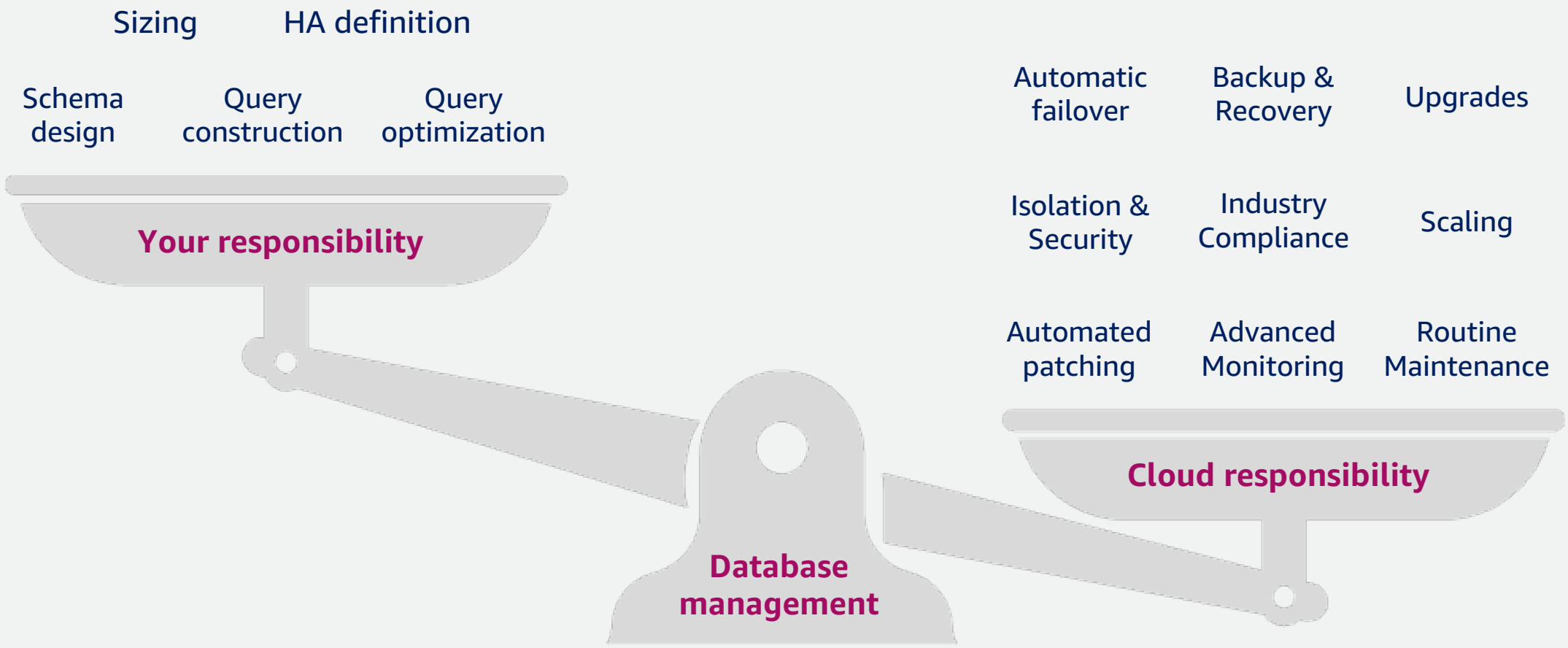
AWS Security Model when using Container Services



AWS Security Model when using Abstracted Services



What are Managed Services? What is Serverless?



Serverless Portfolio

APPLICATION PRIMITIVES – COMPUTE AND DATASTORES


Amazon S3


AWS Lambda


AWS Fargate


Amazon DynamoDB


Amazon Aurora Serverless


Amazon Kinesis

APPLICATION INTEGRATION


Amazon SNS


Amazon API Gateway


AWS Step Functions


Amazon EventBridge


Amazon SQS


AWS AppSync

Developer Tools


AWS CloudFormation


AWS Cloud9


AWS CodePipeline


AWS Config


AWS CloudTrail


Amazon CloudWatch


AWS X-Ray


AWS Serverless Application Repository

SECURITY AND ADMINISTRATION


AWS IAM


AWS SSO


Amazon GuardDuty


Amazon Inspector


Amazon VPC


AWS WAF


AWS Shield


Amazon Cognito

Compute deep dive

More

Opinionated

Less

	AWS manages	Customer manages	
AWS Lambda Serverless functions	Data source integrations Physical hardware, software, networking, and facilities Provisioning	Application code	
AWS Fargate Serverless containers	Container orchestration, provisioning Cluster scaling Physical hardware, host OS/kernel, networking, and facilities	Application code Data source integrations Security config and updates Network config	Management tasks
Amazon ECS/EKS Container-management-as-a-service	Container orchestration control plane Physical hardware software, networking, and facilities	Application code Data source integrations Work clusters	Security config and updates, network config, firewall, management tasks
Amazon EC2 Infrastructure-as-a-Service	Physical hardware software, networking, and facilities	Application code Data source integrations Scaling Security config and updates	Network config Management tasks Provisioning, managing scaling and patching of servers

Serverless Architecture

Event Source

Changes in data state

Requests to endpoints

Changes in resource state

Function

Node.js
Python
Java
C#
Go
Ruby
Bring Your Own

Services / Other

aws

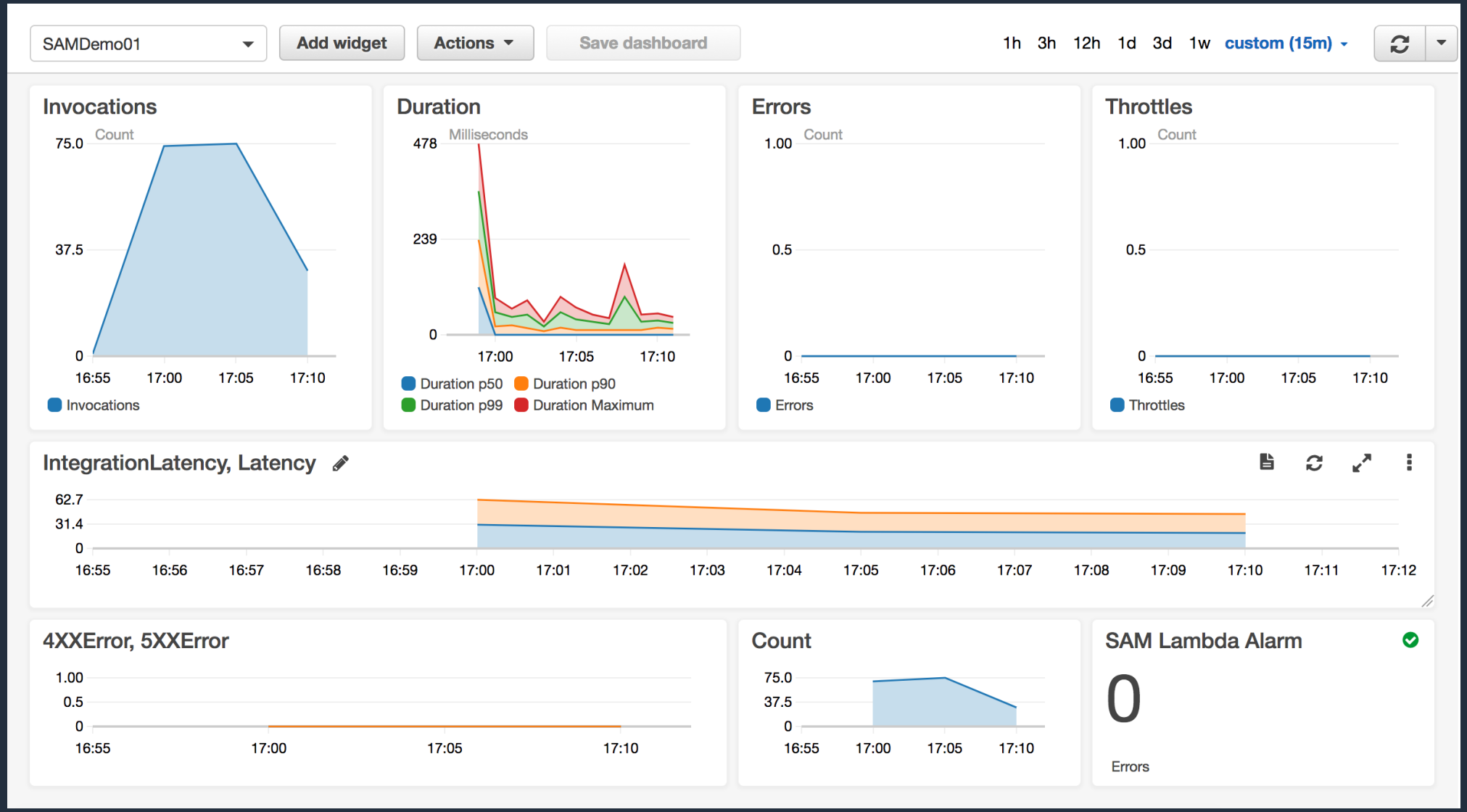
Anatomy of a Lambda Function

- Handler function
  - Function executed on invocation
  - Processes incoming event
- Event
  - Invocation data sent to function
  - Shape differs by event source
- Context
  - Additional information from Lambda service
  - Examples: request ID, time remaining

Lambda_function.py

```
def handler(event, context):
  msg = 'Hello {}'.format(
                    event['name']
                  )
  return { 'message': msg }
```

Built in monitoring

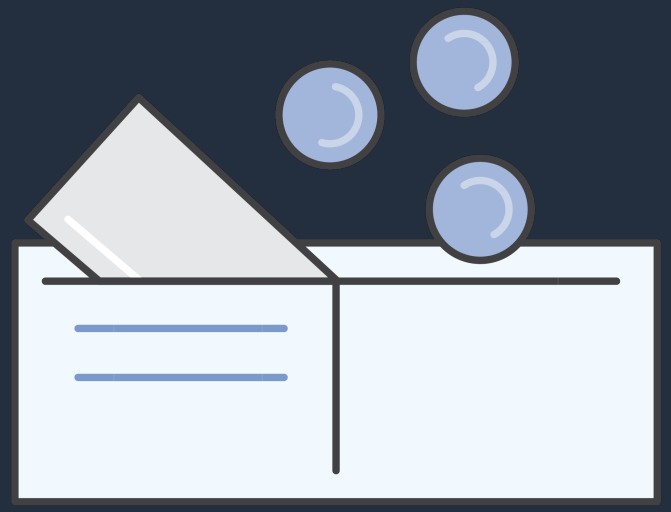


A few items to keep in mind...

- Functions are stateless, no affinity to underlying infrastructure
- Select between 128MB and 10GB
- Choose ARM or x86
- Up to 15 minutes of runtime
- Event triggers an invocation
- Lambda can handle a wide variety of event sources
    - Depending on event source, payload differs
    - Some event sources are batched (e.g. S3, SQS)
- Lambda service manages scaling, invocation
- Lambda service team manages platform security

- Build something!

Fine-grained pricing



Free Tier
1M requests and 400,000 GBs of compute.
Every month, every customer.

Pay for value

- Priced by power rating
- Charged in **1ms** increments
- Low per-request charge

- No minimum

- Never pay for idle

Container technology

Amazon EKS or Amazon ECS?

ECS

Powerful simplicity

EKS

Open flexibility

Powerful simplicity

ECS

AWS-opinionated way to run containers at scale

Reduce decisions without sacrificing scale or features

Reduce time to build, deploy, and migrate applications

https://ecsworkshop.com/

Open flexibility

EKS

Gain agility and efficiency with AWS-optimized Kubernetes, and standardize operations everywhere

Secure, highly available, with observability across all Kubernetes deployments

Build with choice of solutions from the broader community around Kubernetes

https://eksworkshop.com/

Run your containers anywhere based on your workload needs

Serverless

AWS Fargate

Spot instance

EC2 options

Amazon EC2

Spot instance

Edge and 5G

AWS Local Zones

AWS Wavelength

On-premises

AWS Outposts

EKS Anywhere
ECS Anywhere

Thank you!