



# CS443 Cloud Computing Week 1

Eren Akbaba,

[eakbaba@amazon.com](mailto:eakbaba@amazon.com),  
[eren1akbaba@gmail.com](mailto:eren1akbaba@gmail.com)

Bilkent CS

Solutions Architect, AWS Turkey



# Eren Akbaba

- Solutions Architect, AWS Turkey
- Bilkent University, CS, 2014 Alumni ☺
- Hacettepe University, IS, Master's, 2018
- Istanbul Technical University, MBA, 2022
- ~10 years of experience
- Pre-sales, Cloud Architect, Solutions Architect,



@erenakbaba

# CS 443 – Cloud Computing

## Syllabus

- Week 1 - Introduction to Cloud Computing
- Week 2 - Service Models: IaaS, PaaS, SaaS and examples from enterprise for each
- Week 3 - Virtualization techniques (VMs vs Containers) **HW 1**
- Week 4 - Cloud compute services (VM, Containers, Kubernetes, Lambda, Microservices)
- Week 5 - Cloud Networking (VPC) **HW 2**
- Week 6 - **Data Centers and Datacenter Networks + Midterm**
- Week 7 - Cloud Storage - Replication; Distributed File Systems; Structured and unstructured storage
- Week 8 - Cloud Databases - SQL/NoSQL databases; **HW 3**
- Week 9 - Cloud Security
- Week 10 - Cloud Endpoints – APIs, API Gateway, Load Balancing **HW 4**
- Week 11 - Classical 3 Tier Web Applications
- Week 12 - DevOPS **HW 5**
- Week 13 - Machine learning and Generative AI
- Week 14 - Project Presentations – Demos – Make Up Week

# Grading

## Assessment Methods:



	Type	Label	Count	Total Contribution
1	Midterm:Essay/written	Midterm 1	1	20
2	Final:Essay/written	Final	1	30
3	Project	Project	1	20
4	Homework	Homeworks	5	25
5	Attendance	Attendance	42	5

# Agenda

**Cloud overview**

---

**AWS's cloud and locations**

---

**AWS's cloud design**

---

**Climate pledge and AWS**



# Why people choose the cloud?

Most functionality

Largest community of customers and partners

Most secure

Fastest pace of innovation

Most proven operational expertise

# How do customers benefit from the cloud?

## Agility

Teams can experiment and innovate quickly and frequently

---

## Cost Savings

Only pay for what you use, leveraging economies of scale

---

## Elasticity

Easily scale up or down with the needs of the business

---

## Innovation

Focus on business differentiators, not infrastructure

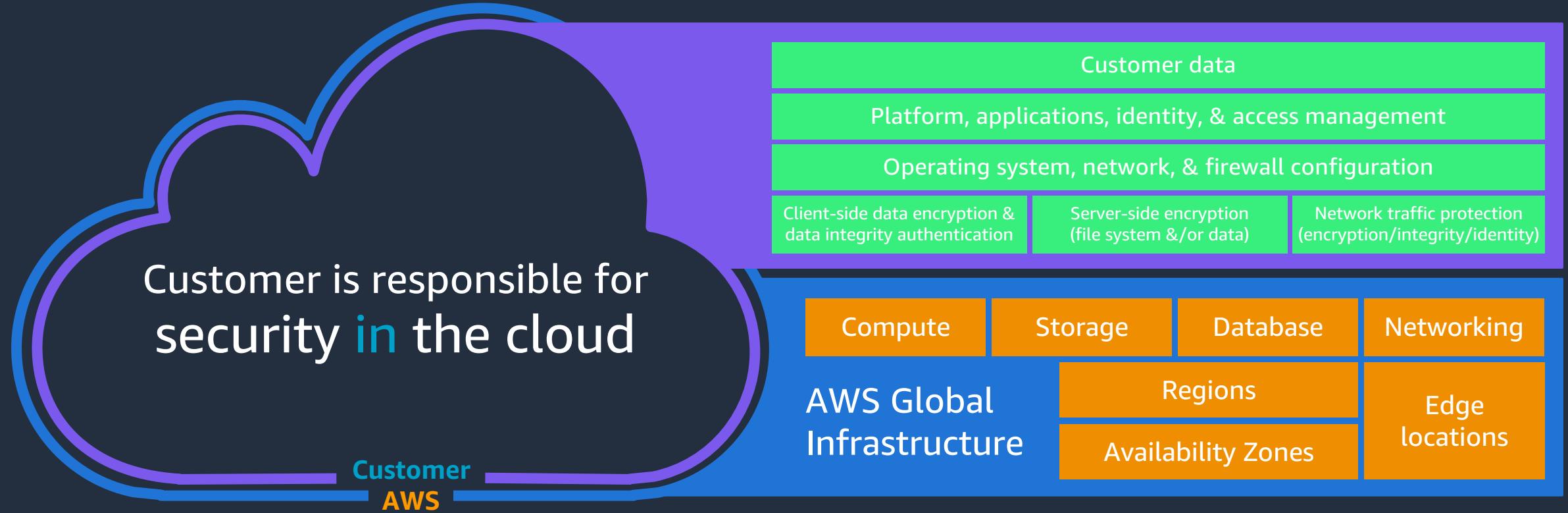
---

## Global Footprint

Extensive, reliable, and secure global cloud infrastructure



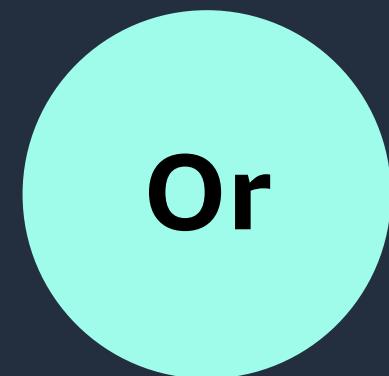
# Share your security responsibility with AWS



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



Develop  
Fast



Be  
Secure

Before...

Develop  
Fast



Be  
Secure

Now...

# Questions to consider when evaluating cloud vendors



## GLOBAL INFRASTRUCTURE

What is the underlying architecture and design of this to enable High Availability and synchronous replication?



## SECURITY

How many third-party attestations and security features do the vendors have?

Do they have international customers that can attest to this?



## INNOVATION

Is the vendor investing for the long term and have demonstrated capability with reference customers taking advantage of innovative services?



## COST

What is the philosophy of the company when it comes to cost? Is it economies of scale driven for long term or is it undercutting the market or credits for the short term?

Have you evaluated price to features across vendors?

# Why customers choose AWS

## EXPERIENCE

**17 years**

helping millions  
of customers

## GLOBAL REACH

**32 Regions**

spanning 102  
Availability Zones and over  
130 Direct Connect  
locations

## SECURITY

**300+**

security features

## INNOVATION

**200+**

service offerings

## AWS INFRASTRUCTURE

**3.6x+**

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

## TCO

**111**

price reductions  
since 2006

## ECOSYSTEM

**4500**

software listings from  
1,400 ISVs

# The **most secure,** **extensive, and reliable** Global Cloud Infrastructure

Infrastructure allows you to run workload

You have the same access and  
capabilities no matter where you are

200+ fully featured services from  
data centers globally



# Understanding the AWS Global Infrastructure

## REGION

A physical location where we cluster data centers. Each Region consists of a minimum of three isolated, and physically separate AZ's.

## AVAILABILITY ZONE (AZ)

One or more discrete data centers with redundant power, networking, and connectivity in an AWS Region.

## LOCAL ZONES

AWS Local Zones place compute, storage, database, and other AWS services closer to end-users so you can run latency-sensitive applications and meet data residency requirements in more locations

## POINTS OF PRESENCE

Reduce latency by delivering data through 550+ globally dispersed Points of Presence (PoPs) with automated network mapping and intelligent routing using Amazon CloudFront.

## WAVELENGTH ZONES

AWS infrastructure deployments that embed AWS services within telecommunications providers' data centers at the edge of the 5G network.

## DIRECT CONNECT LOCATIONS

The Direct Connect service establishes a private, physical network connection between AWS and your data center or office environment – bypassing the public internet.

## OUTPOSTS

AWS Outposts are a family of fully managed solutions delivering AWS infrastructure and services to virtually any on-premises or edge location for a truly consistent hybrid experience.

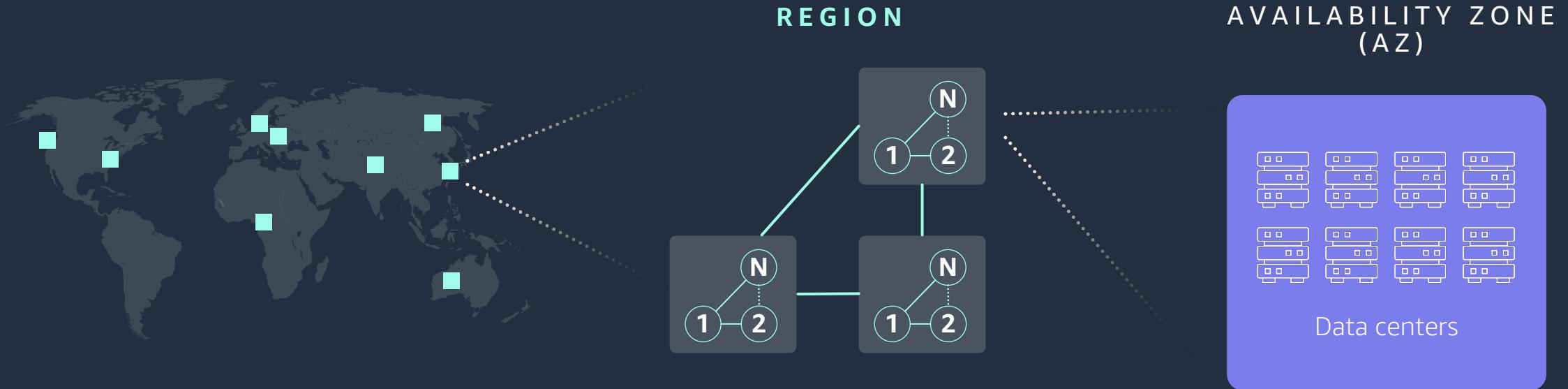
# Global Presence

99+ Availability Zones  
32 Regions ●

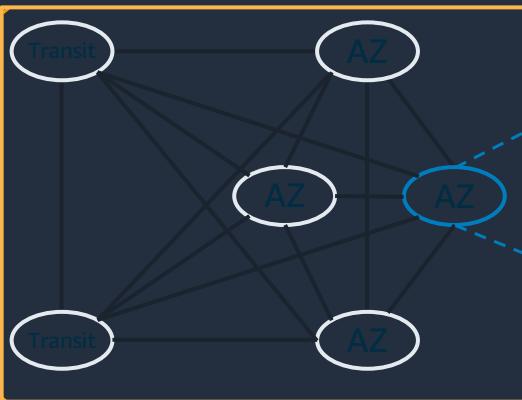


# AWS Region design

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.



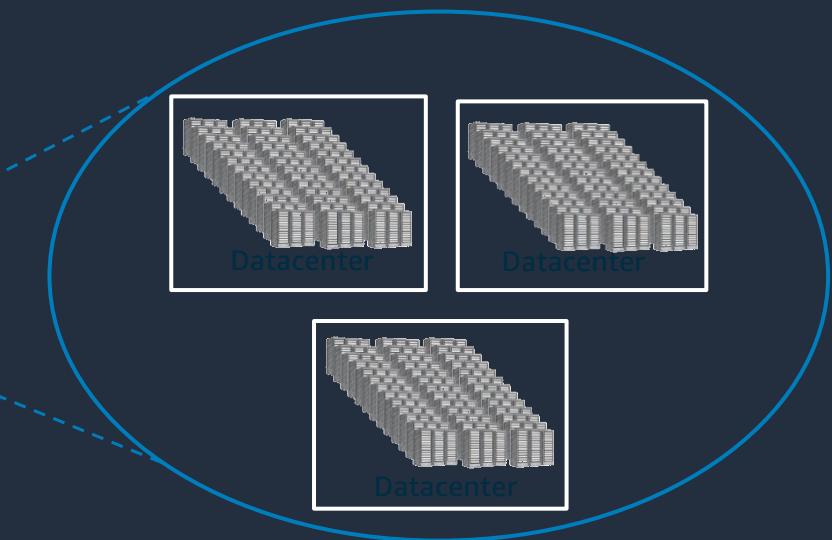
# AWS Region design



## AWS Region

**A Region** is a physical location around the world where we cluster data centers. We call each group of logical data centers an **Availability Zone**.

Transit – Add some info on slide notes



## AWS Availability Zone (AZ)

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

# Regional Expansion

**32 total regions** with 4 coming soon

2006 - 2011

**8 New Regions**

N. AMERICA  
**GovCloud US-West**  
Northern Virginia

Northern California  
Oregon

S. AMERICA  
São Paulo

EUROPE  
**Ireland**

ASIA PACIFIC  
**Singapore**  
Tokyo

2012 - 2017

**10 New Regions**

N. AMERICA  
**Canada Central**  
Ohio

EUROPE  
**Frankfurt**  
London  
Paris

ASIA PACIFIC  
**Mumbai**  
Seoul

\*Beijing  
\*Ningxia

AUSTRALIA &  
NEW ZEALAND  
**Sydney**

2018 - 2023

**14 New Regions**

N. AMERICA  
**GovCloud US-East**

EUROPE  
**Milan**  
Spain  
Stockholm  
Zurich

MIDDLE EAST  
**Tel Aviv**  
Bahrain  
UAE

ASIA PACIFIC  
**Hong Kong**  
Hyderabad  
Jakarta  
Osaka

AFRICA  
**Cape Town**

AUSTRALIA &  
NEW ZEALAND  
**Melbourne**

COMING SOON

**4 New Regions**

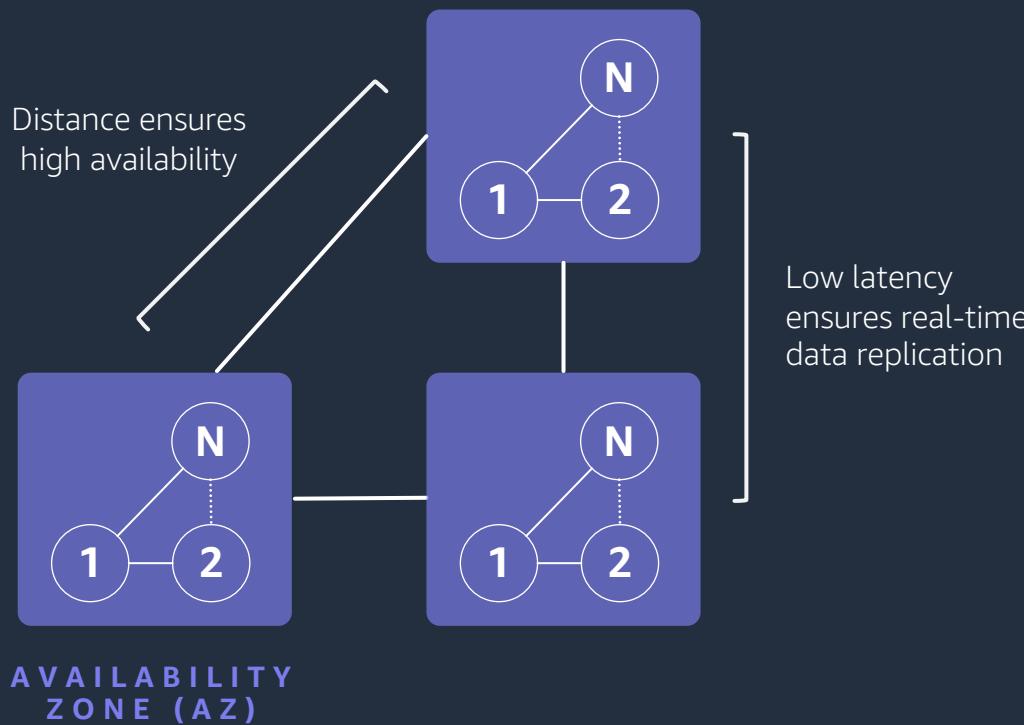
N. AMERICA  
**Canada West**

ASIA PACIFIC  
**Malaysia**  
Thailand

AUSTRALIA &  
NEW ZEALAND  
**Auckland**

# Availability zones design for resiliency

**REGION**



**100K+** servers at scale

## ISOLATED PARTITION

Fully isolated with one or more datacenters

## POWER

Highly available, fault tolerant, and scalable

## DISTANCE

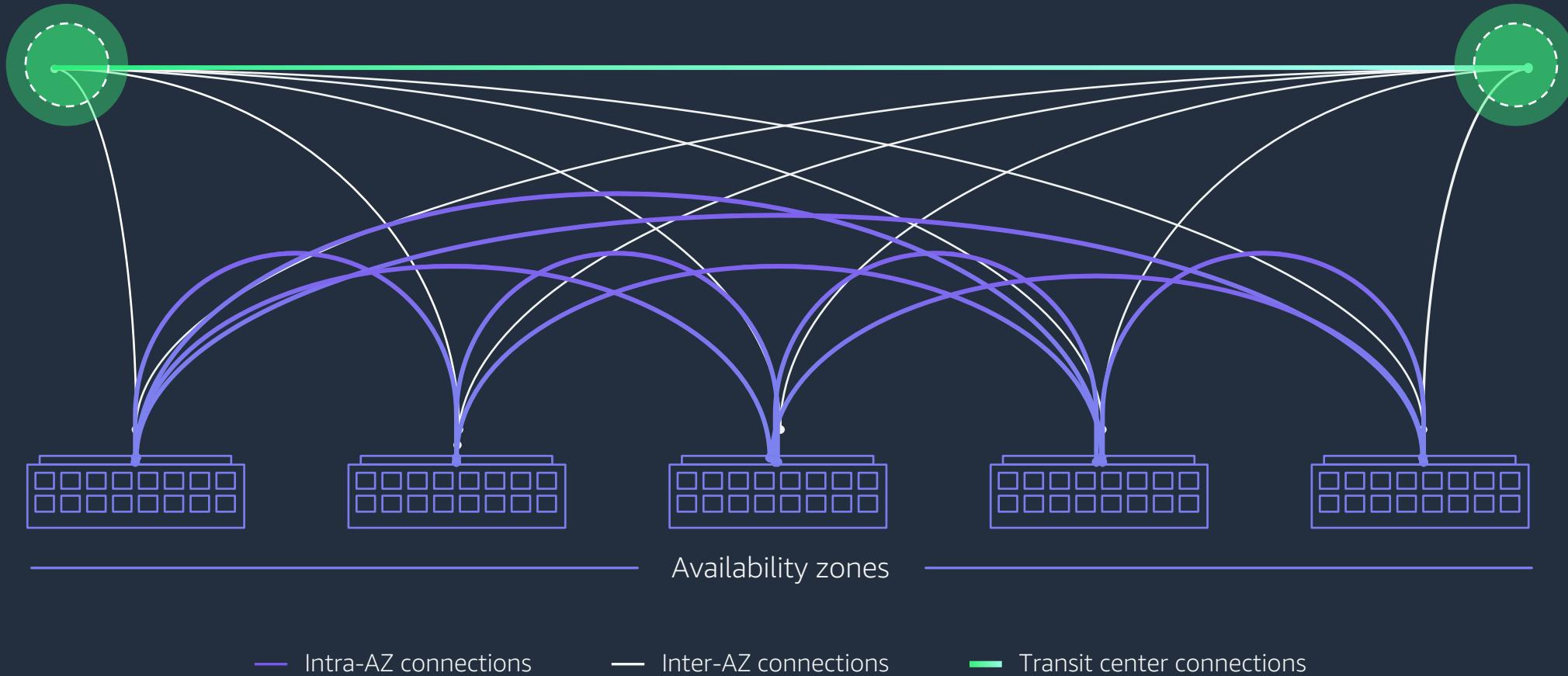
Physically separated by a meaningful distance – all within 60 miles (100km) of each other

## INTERCONNECTION

Datacenters connected via fully redundant and isolated metro fiber

# AWS network design

At least 2 redundant transit centers      Highly peered & connected



# THE Paris... CLIMATE 10 years PLEDGE Early

## PRINCIPLES

Regular reporting

Carbon elimination

Credible offsets

## PROGRESS

**200+**  
signatories

**26**  
industries

**21**  
countries



A large field of solar panels under a cloudy sky. In the foreground, a person wearing a high-visibility vest and a hard hat is kneeling in the grass, looking up at one of the solar panels while holding a laptop.

It's more  
sustainable  
in the cloud

**3.6x**

More energy efficient

**88%**

lower carbon footprint

# AWS Global Infrastructure Regions & AZs



## N AMERICA

- Canada Central (3)
- GovCloud US-East (3)
- GovCloud US-West (3)
- Northern California (3)
- Northern Virginia (6)
- Ohio (3)
- Oregon (4)
- Canada West



## EUROPE

- Frankfurt (3)
- Ireland (3)
- London (3)
- Milan (3)
- Paris (3)
- Spain (3)
- Stockholm (3)
- Zurich (3)



## ASIA PACIFIC

- \*Beijing (3)
- \*Ningxia (3)
- Hong Kong (3)
- Hyderabad (3)
- Jakarta (3)
- Mumbai (3)
- Osaka (3)
- Seoul (4)
- Singapore (3)
- Tokyo (4)
- Malaysia
- Thailand



## MIDDLE EAST

- Bahrain (3)
- Tel Aviv (3)
- UAE (3)



## AUSTRALIA & NEW ZEALAND

- Melbourne (3)
- Sydney (3)
- Auckland

## S AMERICA

- São Paulo (3)



## AFRICA

- Cape Town (3)

● Available Region

● Announced

# Availability Zone

# AWS Global Infrastructure Local Zones



## N AMERICA

Atlanta      Los Angeles (2)      Seattle  
Boston      Miami      Toronto  
Chicago      Minneapolis      Vancouver  
Dallas      New York City  
Denver      Philadelphia  
Houston      Phoenix  
Kansas City      Portland  
Las Vegas      Querétaro



## EUROPE

Copenhagen      Hamburg      Helsinki  
Warsaw      Amsterdam      Lisbon  
Athens      Berlin      Prague  
Brussels      Vienna



## ASIA PACIFIC

Bangkok      Delhi      Kolkata  
Manila      Taipei      Bengaluru  
Chennai      Hanoi



## S AMERICA

Buenos Aires  
Lima  
Santiago  
Bogotá  
Rio de Janeiro



## AFRICA

Lagos  
Johannesburg  
Nairobi



## MIDDLE EAST

Muscat



## AUSTRALIA & NEW ZEALAND

Auckland  
Perth  
Brisbane



Available Local Zone



Announced Local Zone

# AWS Direct Connect



## N AMERICA

Ashburn (2)	Houston	New York	Seattle (3)
Atlanta (2)	Kansas City	Newark	Secaucus
Boston	Las Vegas (2)	Philadelphia	Toronto (2)
Chicago (2)	Los Angeles (2)	Phoenix (2)	Vancouver
Columbus	Miami	Portland (2)	
Dallas	Milpitas	Reston	
Denver	Minneapolis	San Jose (2)	
El Segundo	Montréal (2)	Santa Clara	



## E U R O P E

Amsterdam (2)	London (3)	Prague
Berlin	Madrid (2)	Slough (2)
Copenhagen	Manchester	Stockholm (2)
Dubai	Marseille	Vienna
Dublin (3)	Milano (2)	Warsaw
Frankfurt (2)	Munich	Zurich (2)
Hamburg	Oslo	
Helsinki	Paris (4)	



## A S I A P A C I F I C

Auckland	Inzai	Perth
Bangalore	Jakarta (2)	Seoul (2)
Bangkok	Kolkata	Singapore (2)
Canberra	Kuala Lumpur	Sydney (3)
Chennai	Makati City	Taipei (2)
Delhi	Melbourne (2)	Tokyo (2)
Hong Kong SAR (2)	Mumbai (4)	
Hyderabad	Osaka	



## S AMERICA

Bogota
Buenos Aires
Lima
Rio de Janeiro
Santiago
Sao Paulo (2)



## AFRICA

Cape Town
Johannesburg
Lagos



## MIDDLE EAST

Fujairah
Haifa
Herzliya
Manama (2)
Muscat



## A U S T R A L I A & N E W Z E A L A N D

Auckland
Canberra
Melbourne (2)
Perth
Sydney (3)

● Direct Connect location

# AWS Global Infrastructure

## NORTH AMERICA

Canada Central	3
Toronto	
Vancouver	
GovCloud US-East	3
GovCloud US-West	3
Northern California	3
Northern Virginia	6
Atlanta	
Boston	
Chicago	
Dallas	
Houston	
Kansas City	
Miami	
Minneapolis	
New York City	
Philadelphia	
Querétaro	

Canada West

Ohio

Oregon

Denver

Las Vegas

Los Angeles (2)

Phoenix

Portland

Seattle

GovCloud US-West

Oregon

Northern California

### KEY

Region

Local Zone

# Availability Zones

Region, announced

Local Zone, announced

Direct Connect

# AWS Global Infrastructure

## SOUTH AMERICA

São Paulo 3

Northern Virginia 6

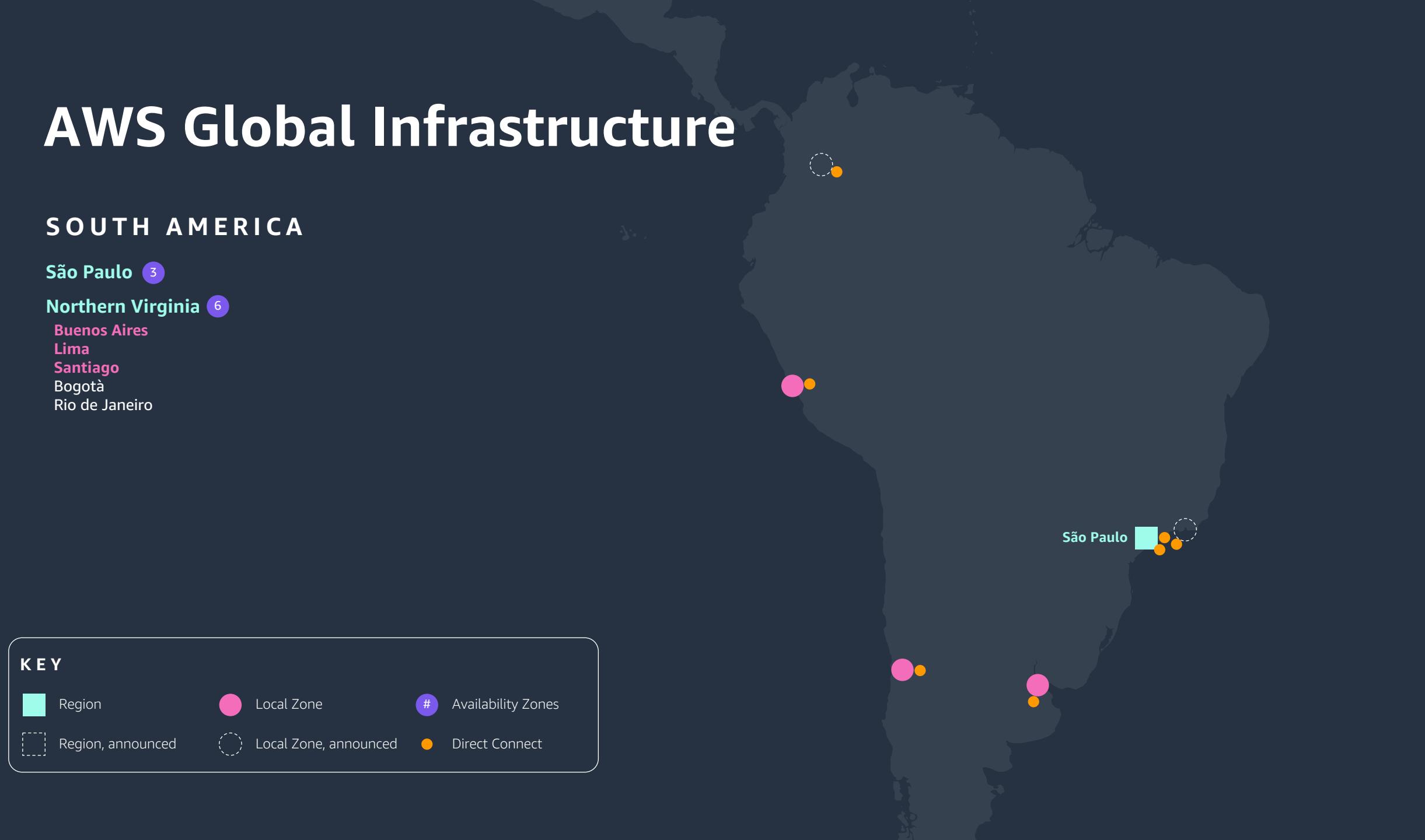
Buenos Aires

Lima

Santiago

Bogotá

Rio de Janeiro

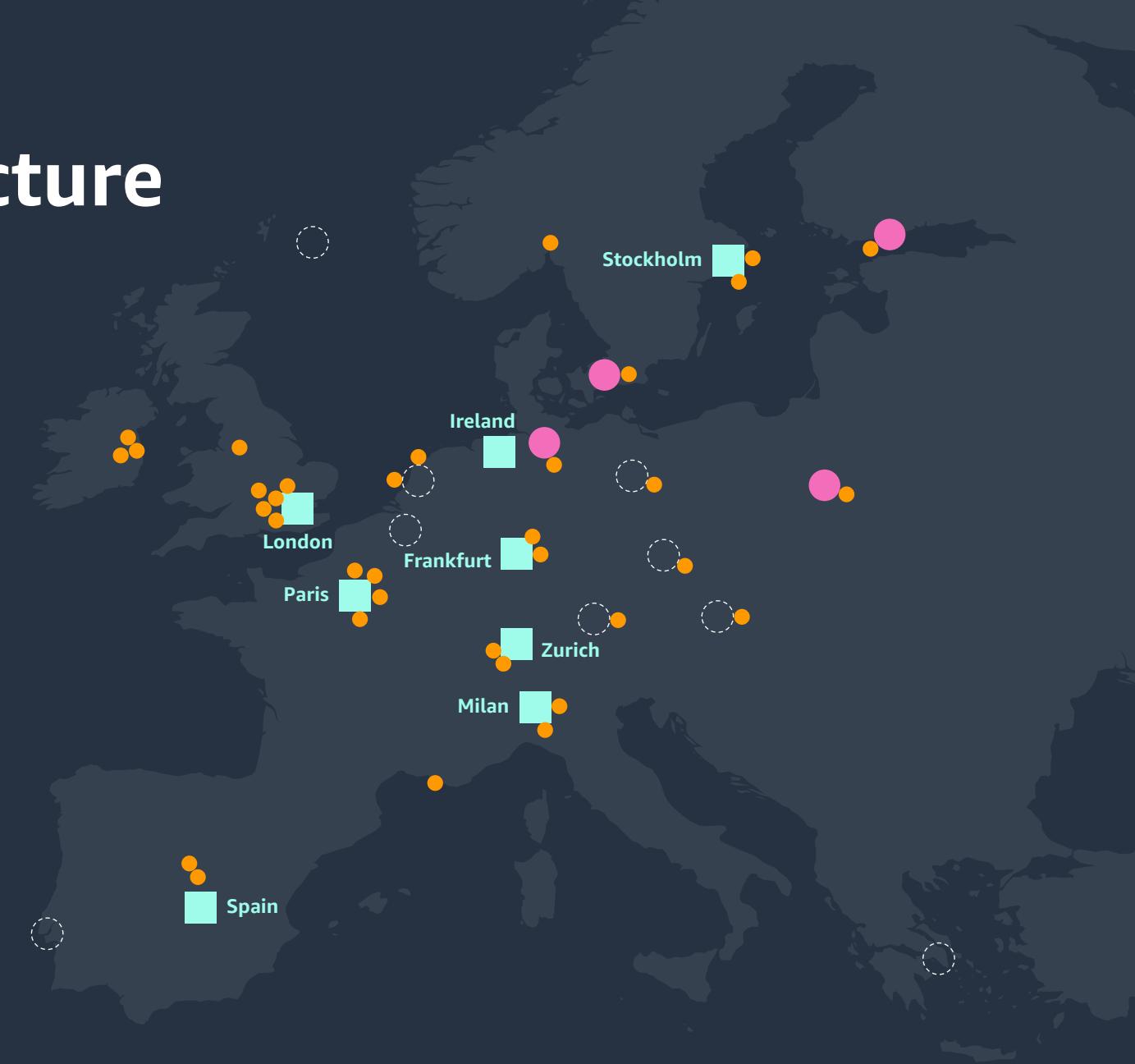


# AWS Global Infrastructure

## EUROPE

Frankfurt (3)  
Hamburg  
Warsaw  
Amsterdam  
Berlin  
Munich  
Prague  
Vienna  
**Ireland** (3)  
**London** (3)  
Milan (3)  
Athens

Paris (3)  
Brussels  
**Spain** (3)  
Lisbon  
**Stockholm** (3)  
Copenhagen  
Helsinki  
Oslo  
**Zurich** (3)



## KEY

- |                   |                       |                      |
|-------------------|-----------------------|----------------------|
| Region            | Local Zone            | # Availability Zones |
| Region, announced | Local Zone, announced | Direct Connect       |

# AWS Global Infrastructure

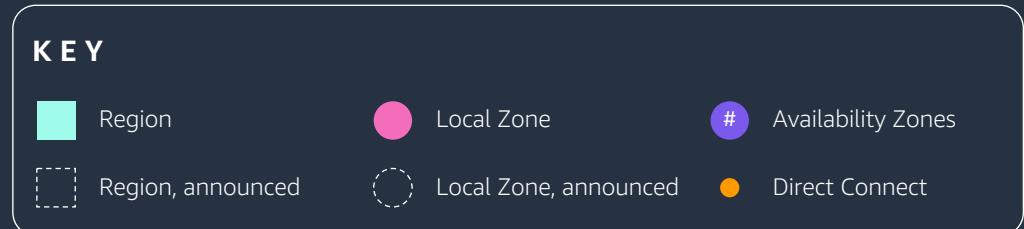
## AFRICA

Cape Town 3

Lagos

Johannesburg

Nairobi



■ Cape Town



# AWS Global Infrastructure

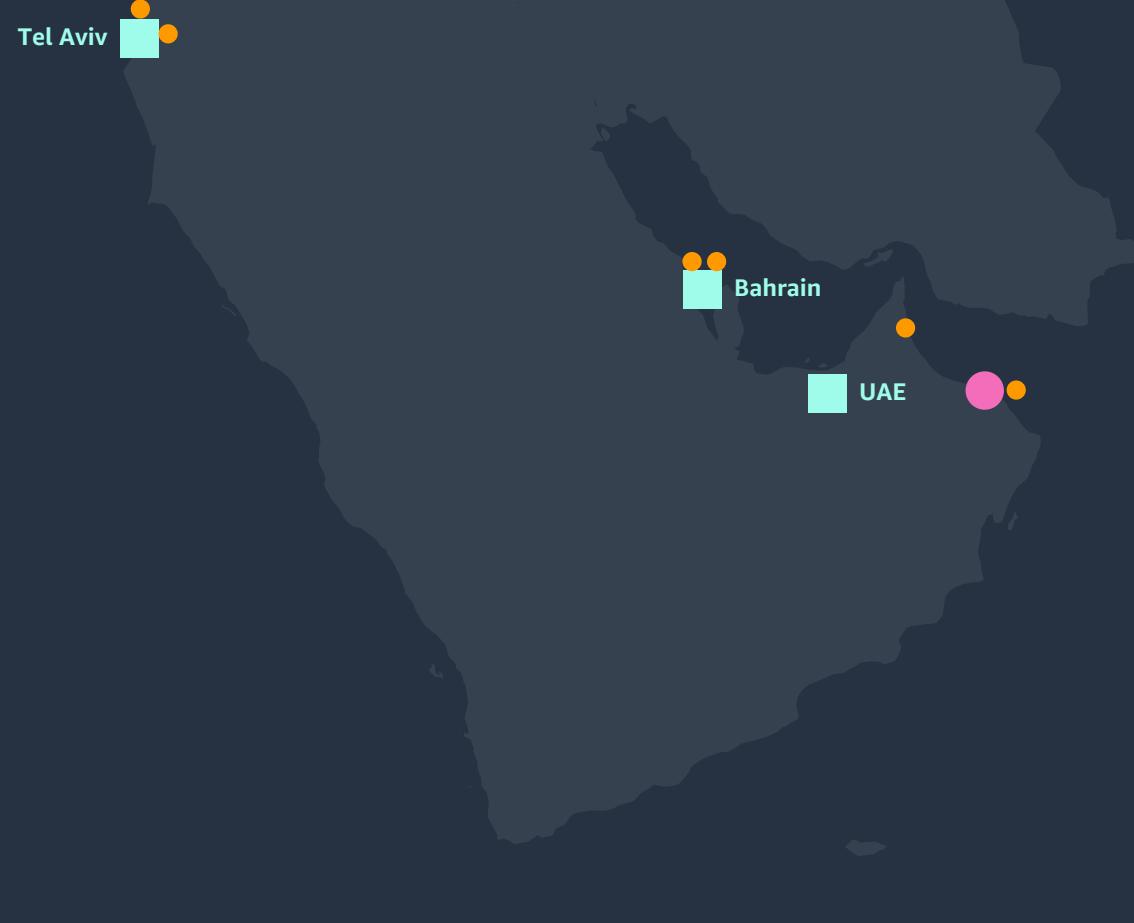
## MIDDLE EAST

Bahrain 3

Muscat

Tel Aviv 3

UAE 3



### KEY

Region

Local Zone

# Availability Zones

Region, announced

Local Zone, announced

Direct Connect

# AWS Global Infrastructure

## ASIA PACIFIC

\*Beijing 3  
\*Ningxia 3  
Hong Kong 3  
Hyderabad 3  
Bengaluru  
Chennai  
Jakarta 3  
Mumbai 3  
Delhi  
Kolkata  
Osaka 3  
Seoul 4

Singapore 3  
Bangkok  
Manila  
Hanoi  
Tokyo 4  
Taipei  
Malaysia  
Thailand

\*Beijing  
\*Ningxia  
Seoul  
Tokyo  
Osaka

### KEY

Region

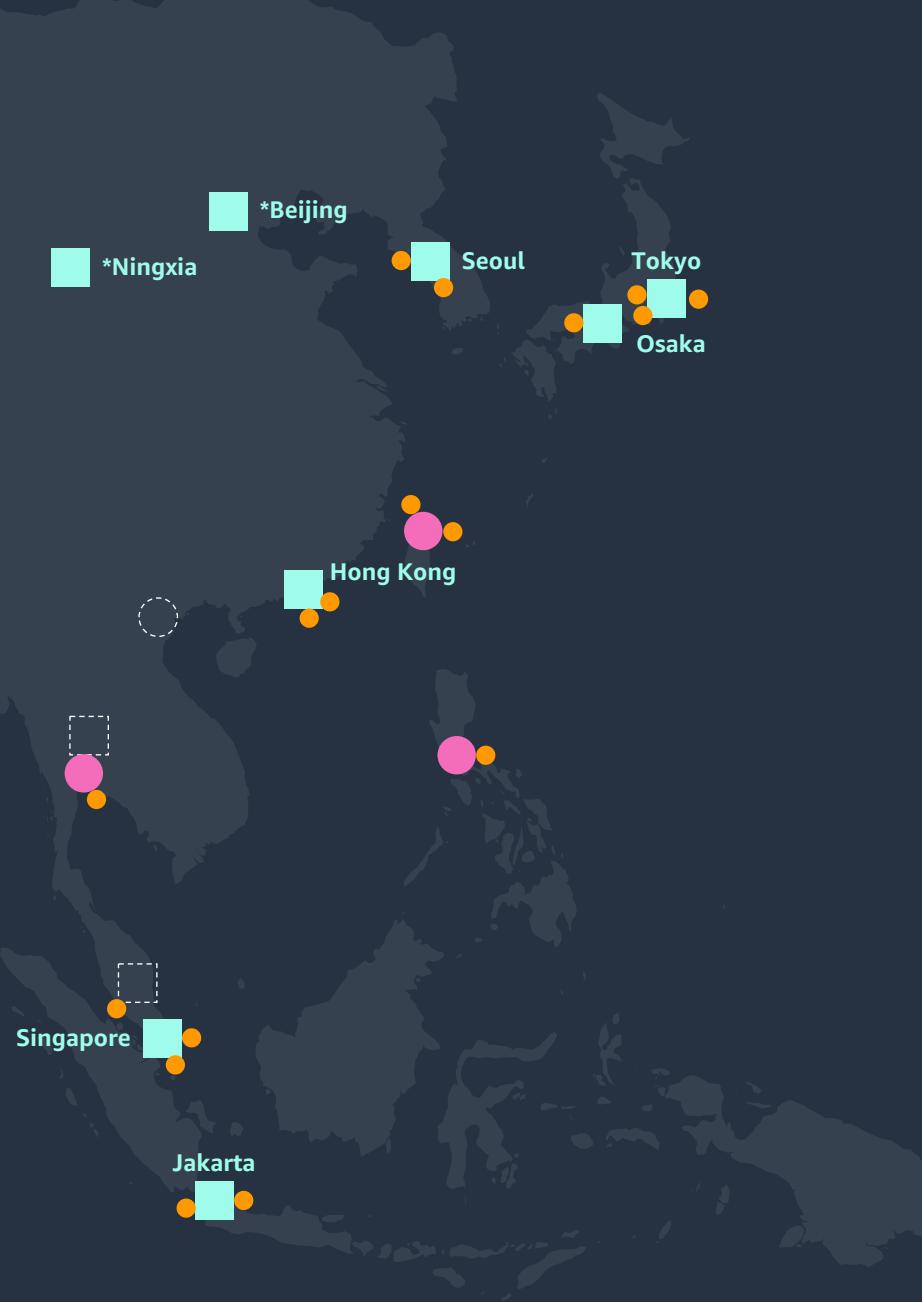
Local Zone

# Availability Zones

Region, announced

Local Zone, announced

Direct Connect



# AWS Global Infrastructure

## AUSTRALIA & NEW ZEALAND

Melbourne 3

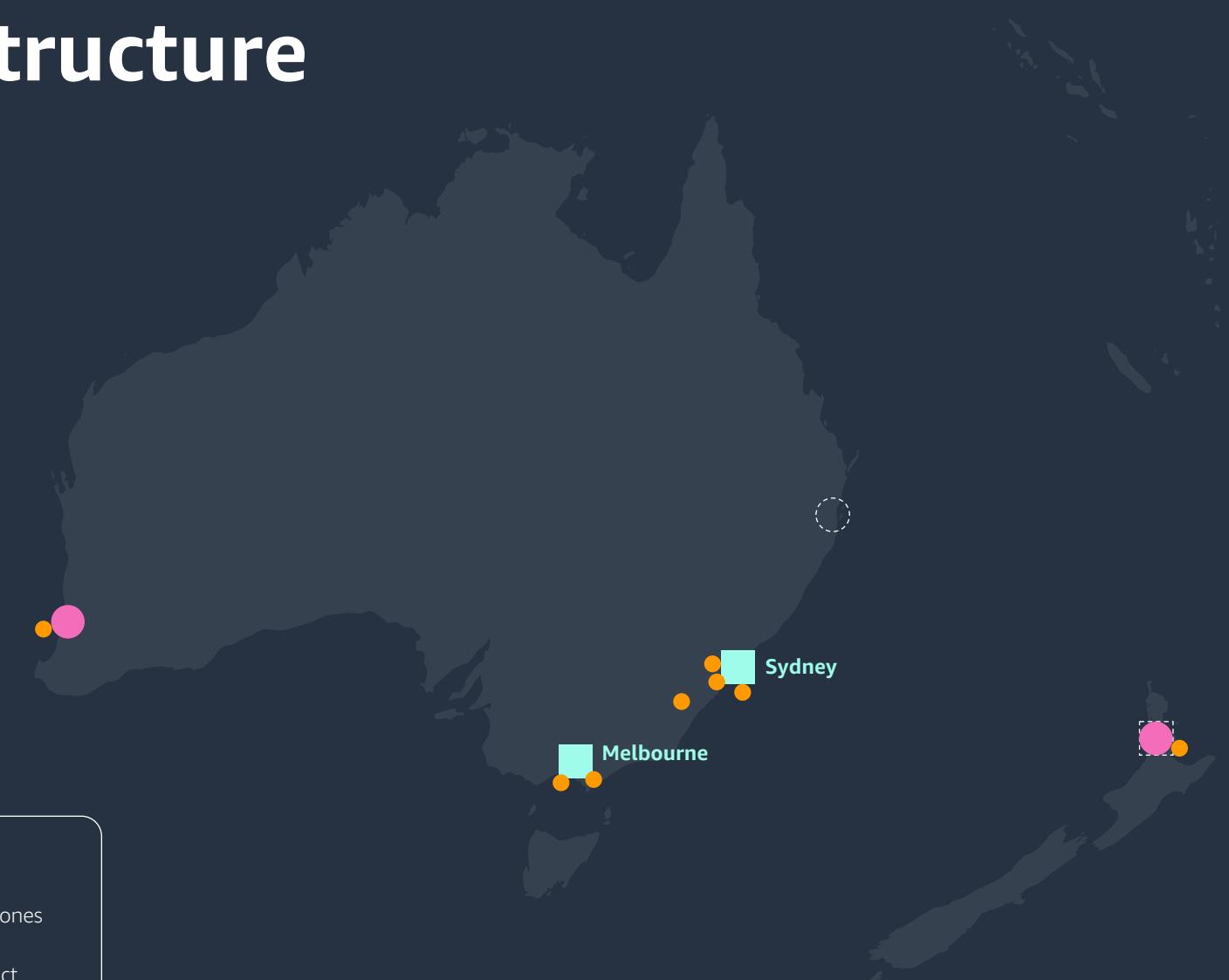
Sydney 3

Auckland

Perth

Brisbane

Auckland



### KEY

Region

Local Zone

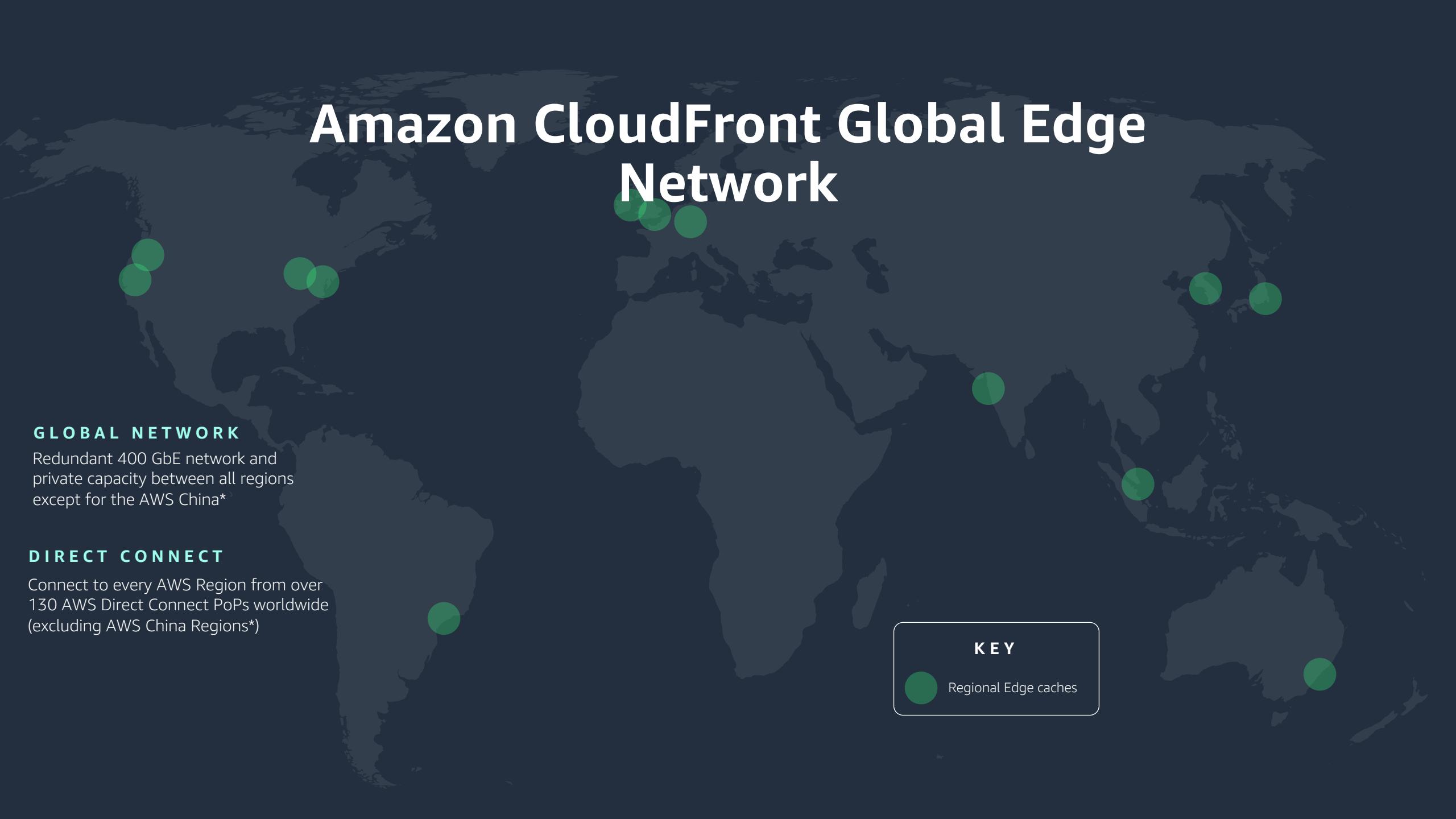
# Availability Zones

Region, announced

Local Zone, announced

Direct Connect

# Amazon CloudFront Global Edge Network



## GLOBAL NETWORK

Redundant 400 GbE network and private capacity between all regions except for the AWS China\*

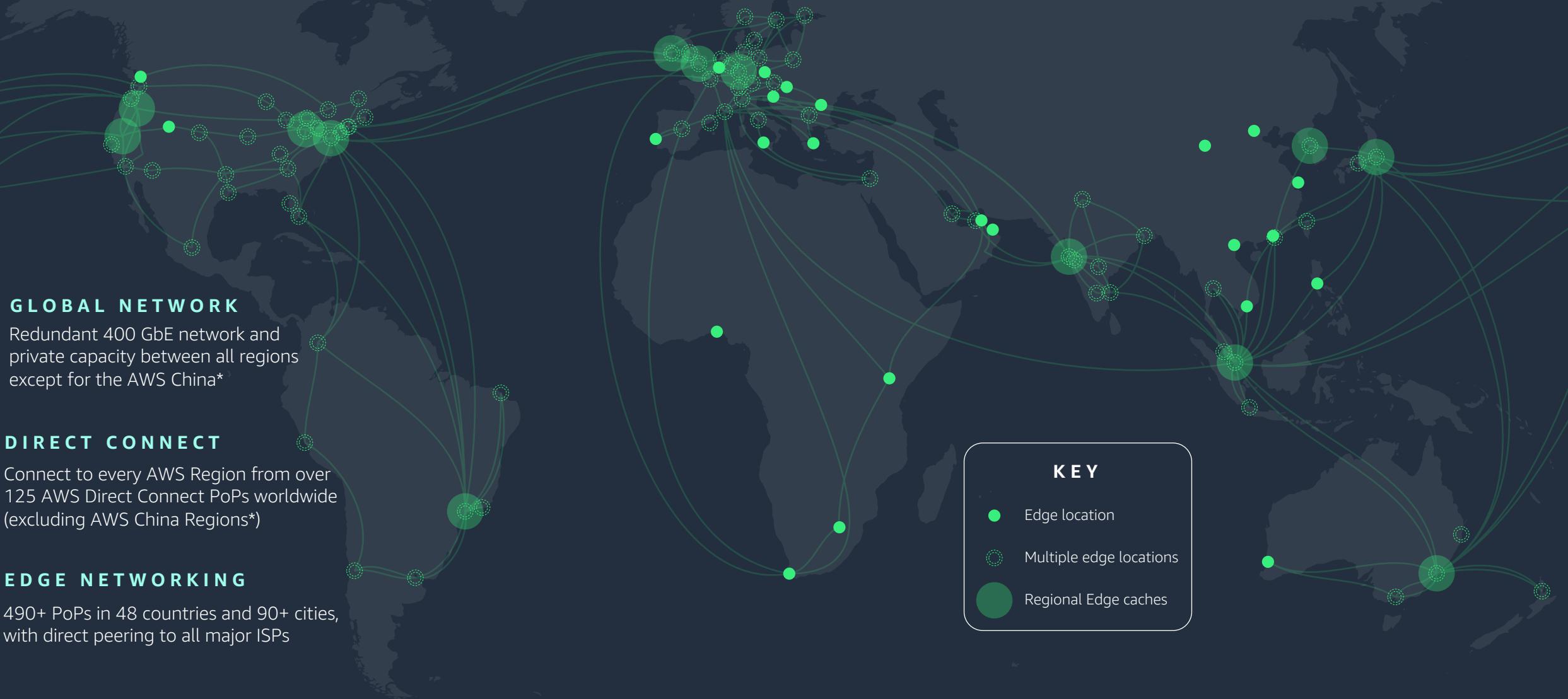
## DIRECT CONNECT

Connect to every AWS Region from over 130 AWS Direct Connect PoPs worldwide (excluding AWS China Regions\*)

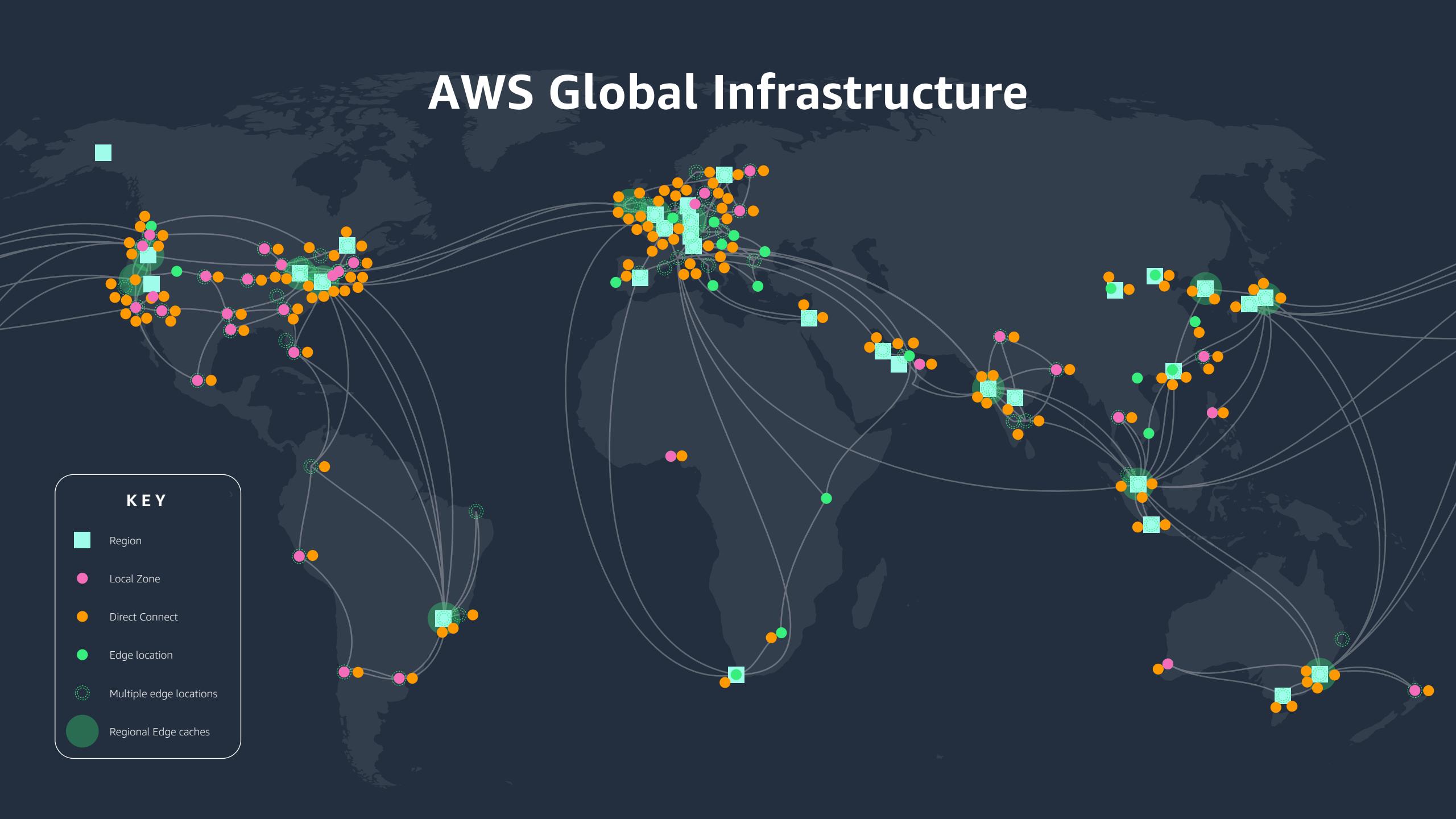
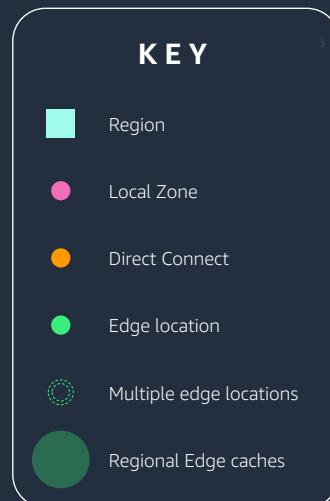
### KEY

Regional Edge caches

# AWS CloudFront Global Edge Network



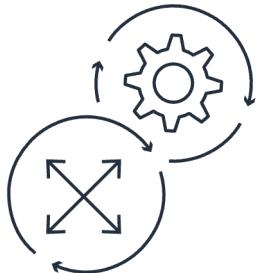
# AWS Global Infrastructure



# Support virtually any cloud workload

Application Development	End User Computing	Application Integration	200+ Fully feature services
Analytics	Machine Learning	Internet of Things	
Media	Mobile	Databases	
Core Services Compute – Storage - Networking	Security and Compliance	400	
Infrastructure Regions – Availability Zones	Management Tools	types of compute instances	

# Pillars of AWS Well-Architected



Operational Excellence



Security



Reliability



Performance Efficiency

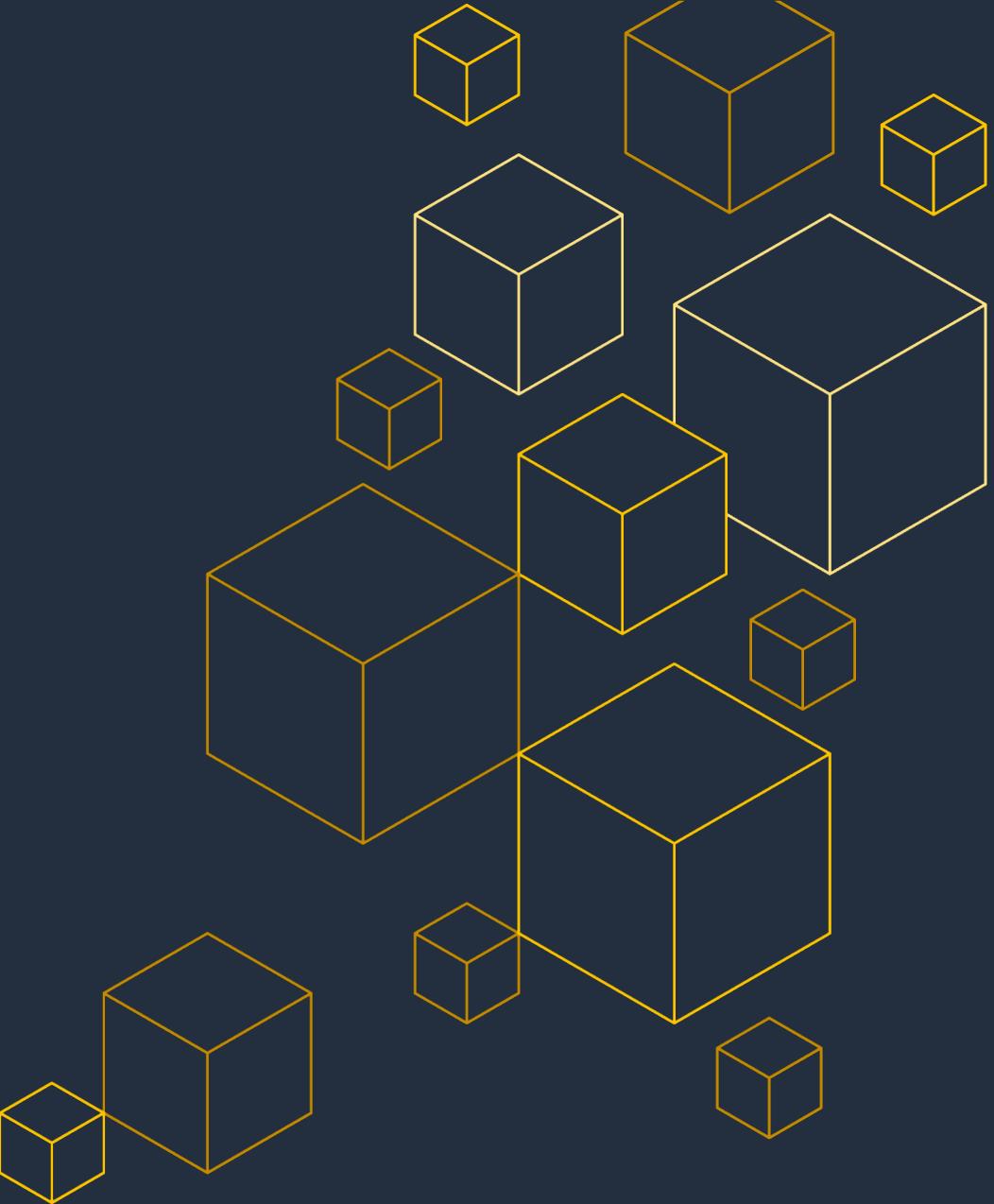


Cost Optimization



Sustainability

# Compute



# Compute in AWS



Amazon Elastic  
Container Registry  
(Amazon ECR)



VMware Cloud  
on AWS



Amazon Lightsail



AWS Batch



AWS Outposts



AWS Serverless  
Application Repository

# The widest selection of processors and architectures



INTEL® Xeon Scalable  
(Skylake and Cascade  
Lake) processors



AMD EPYC processor



AWS Graviton Processor

# Storage



# Storage

- Data is stored as objects within buckets
- Virtually unlimited storage
  - A single object is limited to 5 TB
- Designed for durability of 99.9999999%
- Detailed access to buckets and objects



Amazon Simple Storage Service  
(Amazon S3)

# Storage



Amazon Elastic Block Store  
(Amazon EBS)

# Volume types in Amazon EBS

	Solid State Disk (SSD)		Hard Disk Drive (HDD)	
	General Purpose	Provisioned IOPS	Throughput Optimized	Cold
Max volume size	16 TiB	16 TiB	16 TiB	16 TiB
Max IOPS per volume	16 000	64 000	500	250
Max throughput per volume	250 MiB/s	1000 MiB/s	500 MiB/s	250 MiB/s

# Databases



# Purpose-built databases

Relational	Key-value	Document	In-memory	Graph	Time-series	Ledger
Referential integrity, ACID transactions, schema-on-write	High throughput, Low latency reads and writes, endless scale	Store documents and quickly access querying on any attribute	Query by key with microsecond latency	Quickly and easily create and navigate relationships between data	Collect, store, and process data sequenced by time	Complete, immutable, and verifiable history of all changes to application data
<i>AWS Service(s)</i> Aurora      RDS	DynamoDB	DocumentDB	ElastiCache	Neptune	Timestream	QLDB
<i>Common Use Cases</i> Lift and shift, ERP, CRM, finance	Real-time bidding, shopping cart, social, product catalog, customer preferences	Content management, personalization, mobile	Leaderboards, real-time analytics, caching	Fraud detection, social networking, recommendation engine	IoT applications, event tracking	Systems of record, supply chain, health care, registrations, financial



# Security

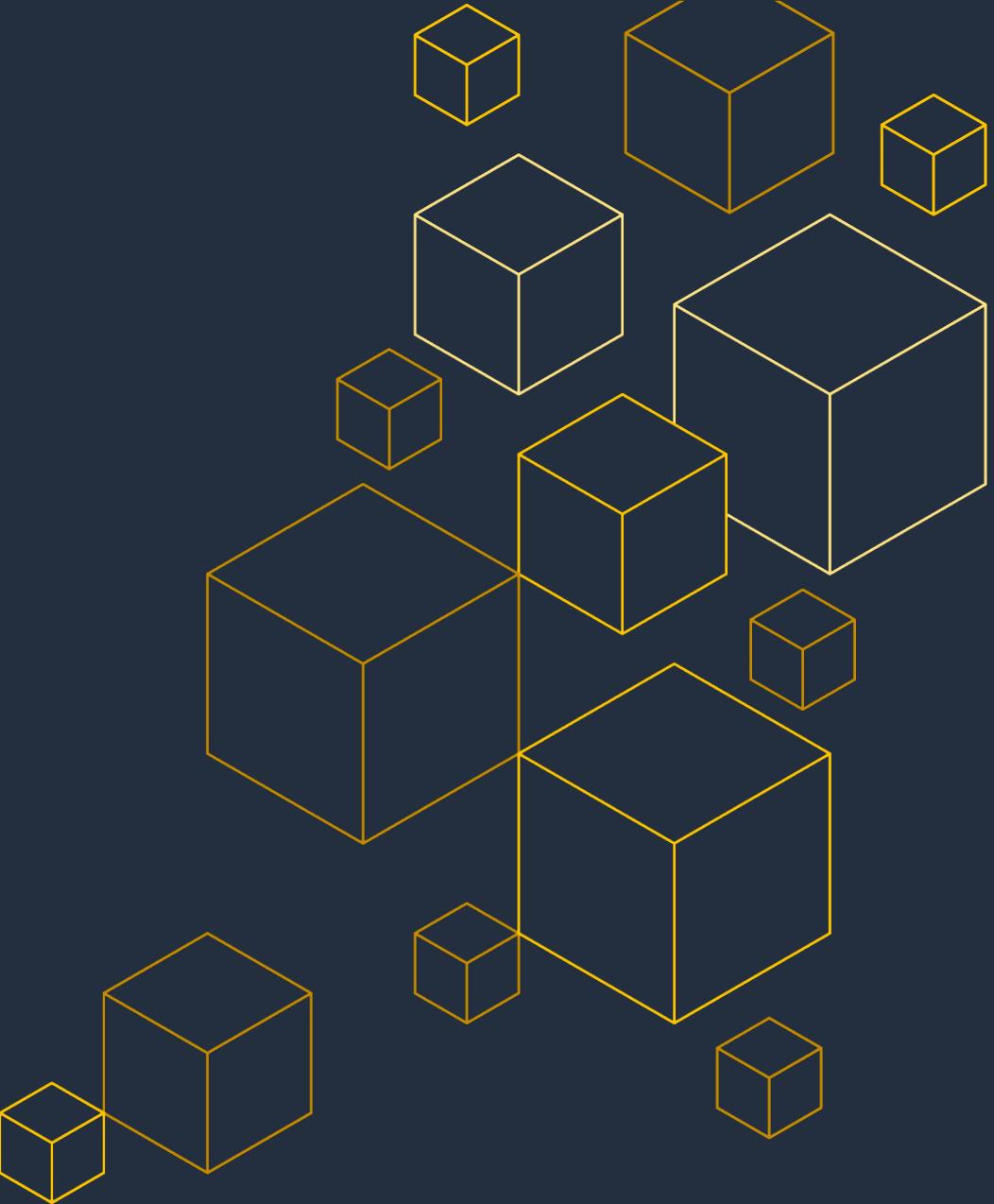


# AWS Security, Identity, and Regulatory Compliance Solutions

 Identity and access management	 Detective controls	 Infrastructure protection	 Data protection	 Incident response	 Compliance
AWS Identity and Access Management (IAM)	AWS Security Hub	AWS Firewall Manager	Amazon Macie	Amazon Detective	AWS Artifact
AWS Single Sign-On	Amazon GuardDuty	AWS Network Firewall	AWS Key Management Service (KMS)	Amazon EventBridge	AWS Audit Manager
AWS Organizations	Amazon Inspector	AWS Shield	AWS CloudHSM	AWS Backup	
AWS Directory Service	Amazon CloudWatch	AWS WAF	AWS Certificate Manager	AWS Security Hub	
Amazon Cognito	AWS Config	Amazon VPC	AWS Secrets Manager	AWS Elastic Disaster Recovery	
AWS Resource Access Manager	AWS CloudTrail	AWS PrivateLink	AWS VPN		
	VPC Flow Logs	AWS Systems Manager	Server-Side Encryption		
	AWS IoT Device Defender				



# Networking



# Amazon VPC



It allows you to provision a **logically isolated** section of the AWS cloud where you can launch AWS resources on a virtual network that you define.

Lets you **control your virtual network** resources, including the following:

- Selecting the IP Address Range
- Creating Subnets
- Configuring Routing Tables and Network Gateways

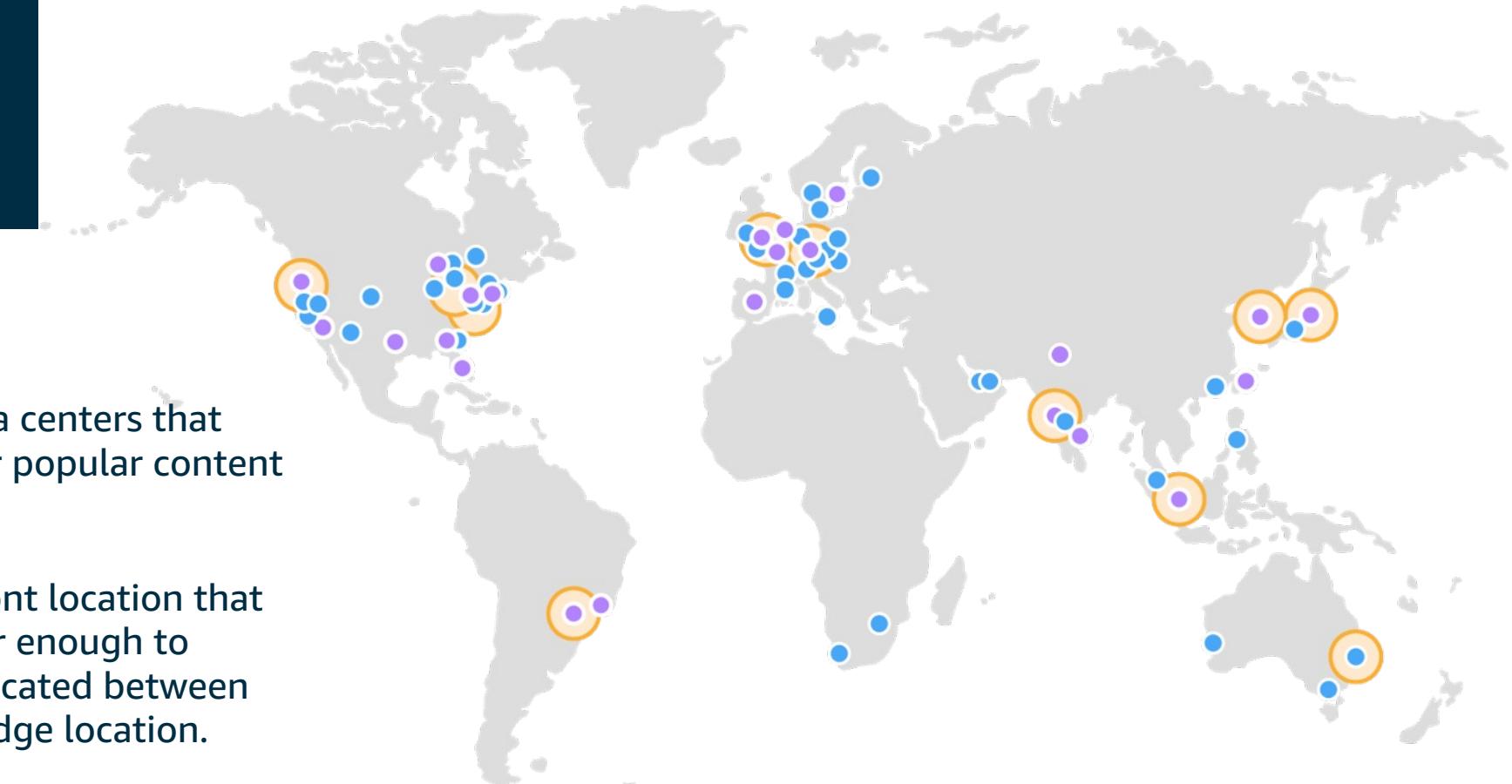
Allows you to **customize the network configuration** of your VPC.

It allows you to use **multiple layers of security**.

# Amazon CloudFront Infrastructure

- Edge Location
- Several Edge Location
- Regional Edge Cache

- **Edge locations:** A network of data centers that CloudFront uses to quickly deliver popular content to customers.
- **Regional Edge Cache:** A CloudFront location that caches content that is not popular enough to remain in an edge location. It is located between the origin server and the global edge location.



# Amazon Route 53

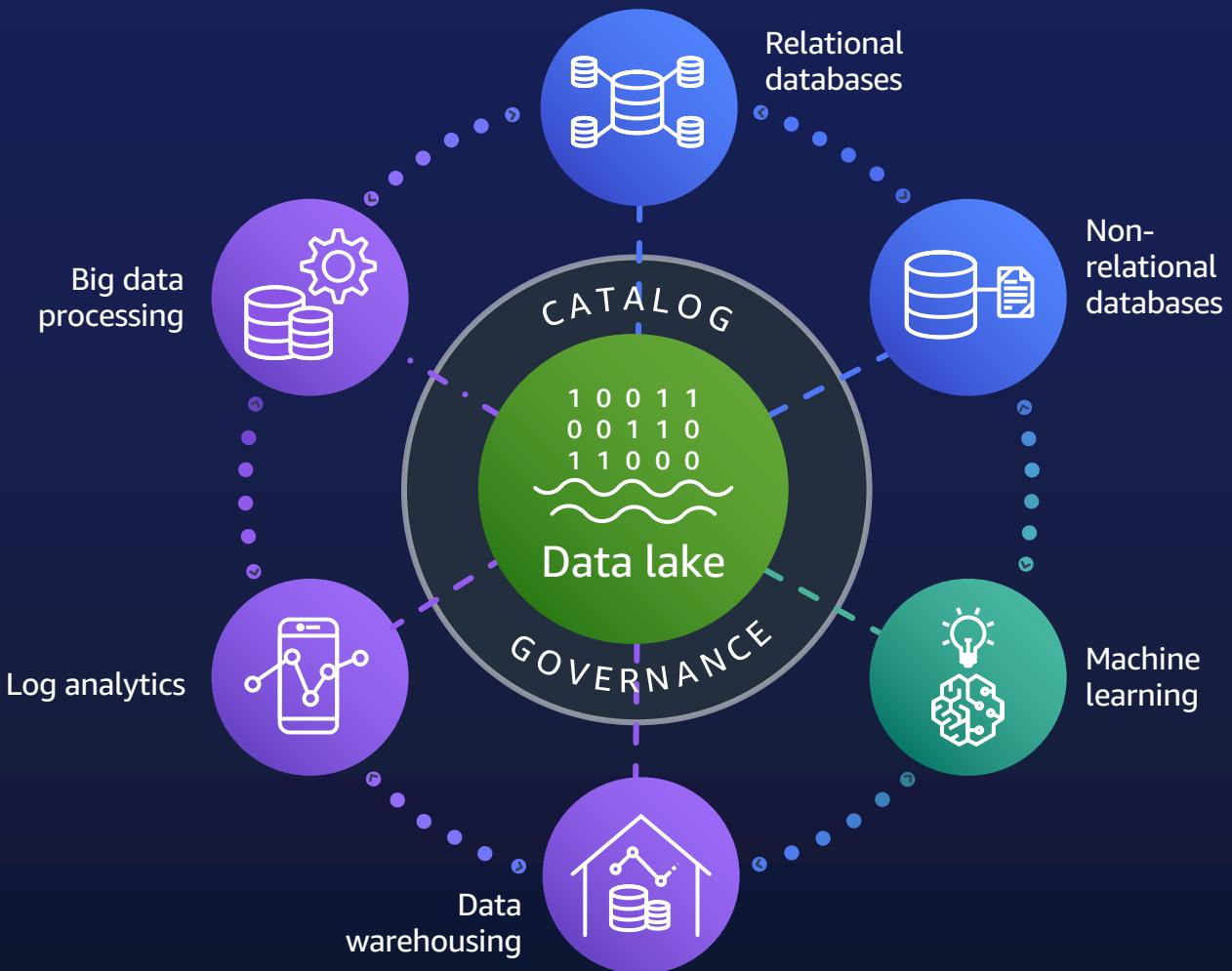


- It is a highly available and scalable DNS (Domain Name System) web service.
- It is used to redirect end users to applications on the Internet by translating names (such as [www.example.com](http://www.example.com)) into numeric IP addresses (such as 192.0.2.1) that computers use to connect to each other.
- It is fully compatible with IPv4 and IPv6.

# Analytics / Machine Learning



# Modern Data architecture on AWS



**Scalable data lakes**

**Purpose-build data services**

**Seamless data movement**

**Unified governance**

**Performant and cost-effective**

# Machine Learning Frameworks

## ML frameworks & infrastructure

PyTorch, Apache MXNet, TensorFlow

AMIs & containers

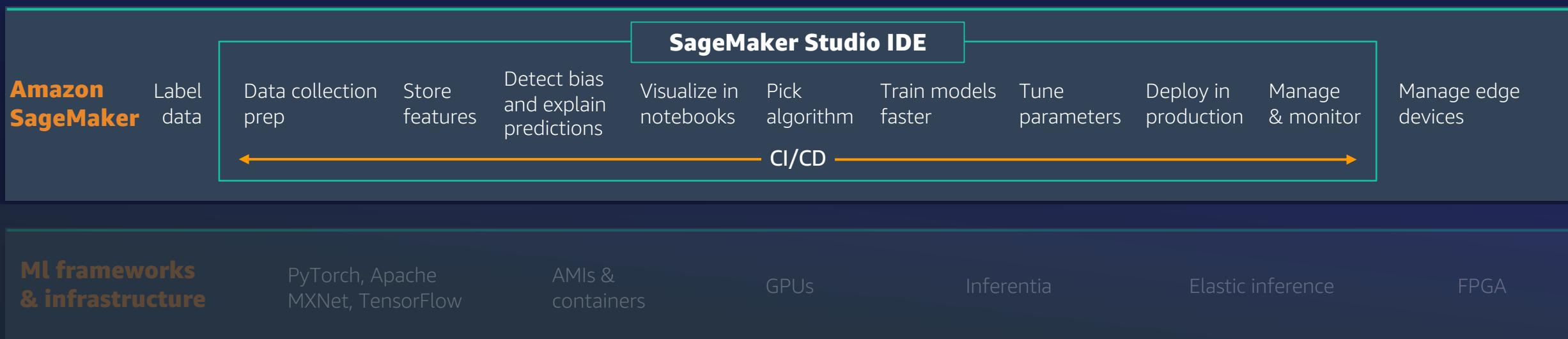
GPUs

Inferentia

Elastic inference

FPGA

# Machine Learning Services



# Artificial Intelligence Services

<b>Vision</b> Amazon Rekognition	<b>Chatbots</b> Amazon Lex	<b>Business tools</b> Amazon Personalize, Amazon Forecast, Amazon Fraud Detector, Amazon Lookout for Metrics	<b>Search</b> Amazon Kendra	<b>Healthcare</b> Amazon HealthLake, Amazon Comprehend Medical, Amazon Transcribe Medical
<b>AI services</b>				
<b>Speech</b> Amazon Polly Amazon Transcribe	<b>Text</b> Amazon Comprehend Amazon Translate Amazon Textract	<b>Contact centers</b> Contact Lens for Amazon Connect Amazon Connect Voice ID	<b>Code &amp; DevOps</b> Amazon CodeGuru DevOps Guru	<b>Industrial</b> AWS Panorama Appliance and SDK, Amazon Monitron, Amazon Lookout for Equipment, Amazon Lookout for Vision

