



# Intro to Amazon Cloud

## EC2 overview

Abhay Gwaliorkar  
Solutions Architect

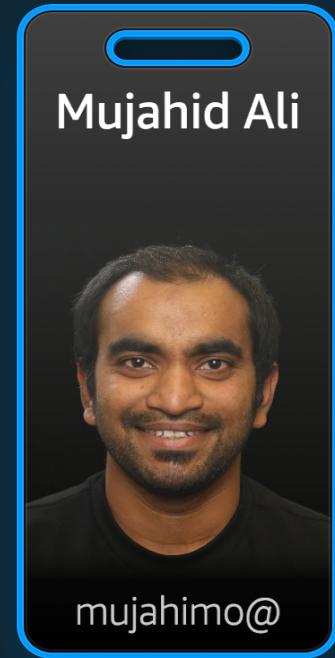
April 2022



# AWS Team



**Abhay Gwaliorkar**  
Solutions Architect  
[abhaygwa@amazon.com](mailto:abhaygwa@amazon.com)



**Mujahid Ali Mohammed**  
Solutions Architect  
[mujahimo@amazon.com](mailto:mujahimo@amazon.com)

# Agenda

- Introduction to AWS Cloud
- Global Reach
- EC2 Overview
- EC2 Details

# What is AWS?

AWS provides a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers millions of businesses in over 190 countries around the world.

## Benefits

- Low Cost
- Elasticity
- Agility
- Secure
- Global Reach



# What sets AWS apart?



Security

Fine-grained control



Service Breadth & Depth; pace of innovation

200 + services to support any cloud workload; rapid customer driven releases



Experience: 1M+ customers

Building and managing cloud since 2006



Global Footprint

84 Availability Zones within 26 geographic Regions



Machine Learning

More machine learning happens on AWS than anywhere else.  
Machine learning in the hands of every developer and data scientist.



Ecosystem

Tens of thousands of APN partners. The AWS Marketplace offers 50 categories, and more than 8,000 software listings

# Pricing Philosophy

High volume / low margin businesses are in our core DNA

Trade CapEX for  
variable expense

Pay for what  
you use

Our economies of  
scale provide us  
with lower costs

85 price  
reductions  
since 2006

Pricing model  
choice to support  
variable and stable  
workloads

On-demand  
Reserved Instances  
Spot

Save more money as  
you grow bigger

Tiered pricing  
Volume discounts  
Custom pricing

# Customer obsessed



90%  
of roadmap originates with customer requests  
and are designed to meet specific needs



“Performance, reliability, and responsiveness are fundamental to our customer experience, and T3 instances help us to deliver on that customer promise while also controlling our costs.”

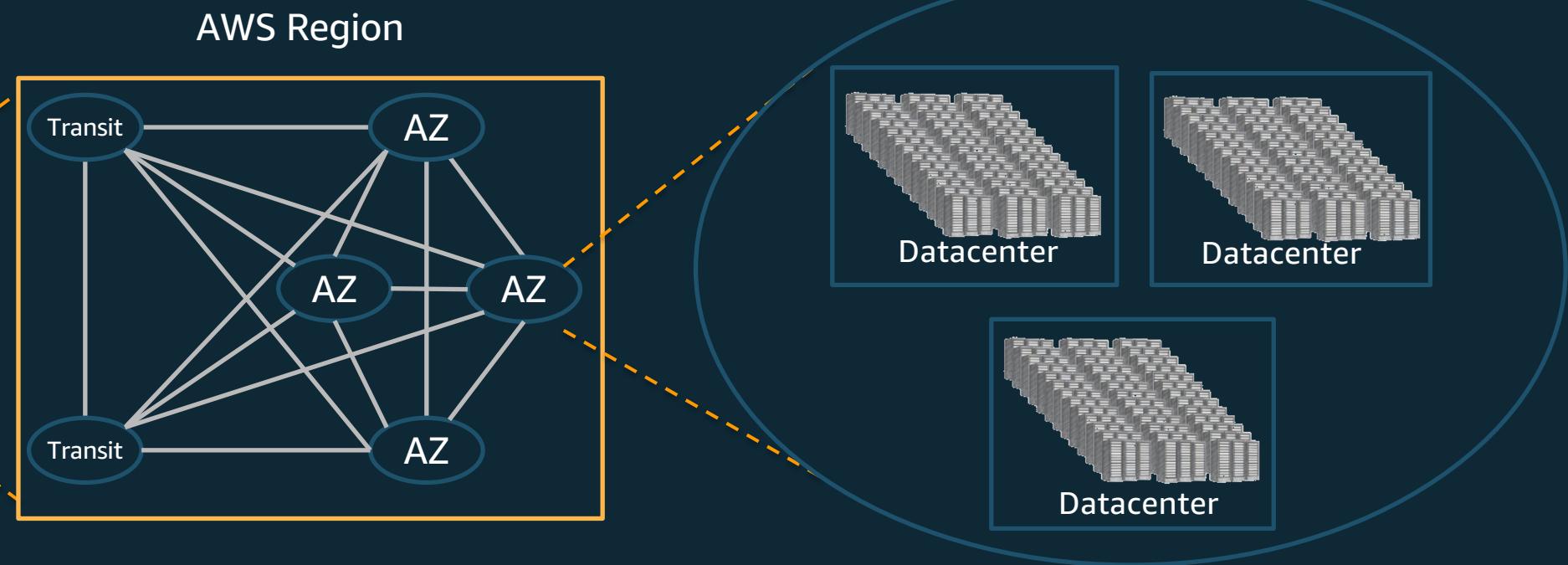
—Heroku

1

AWS Global Reach

# 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.

# 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



# AWS Global Infrastructure

26 geographical regions, 84 availability zones, 310+ POPs



## Region & Number of Availability Zones (AZs)

### GovCloud (US)

US-East (3), US-West (3)

### US West

Oregon (4)

Northern California (3)

### US East

N. Virginia (6), Ohio (3)

### Canada

Central (3)

### South America

São Paulo (3)

### Africa

Cape Town (3)

## Announced Regions

8 Regions and 24 AZs in Australia, Canada, India, Israel, New Zealand, Spain, Switzerland, and United Arab Emirates (UAE).

### Europe

Frankfurt (3), Paris (3),  
Ireland (3), Stockholm (3),  
London (3), Milan (3)

### Middle East

Bahrain (3)

### Asia Pacific

Singapore (3), Sydney (3)  
Tokyo (4), Osaka (3)  
Seoul (3), Mumbai (3),  
Hong Kong (3), Indonesia (3)

### China

Beijing (2), Ningxia (3)



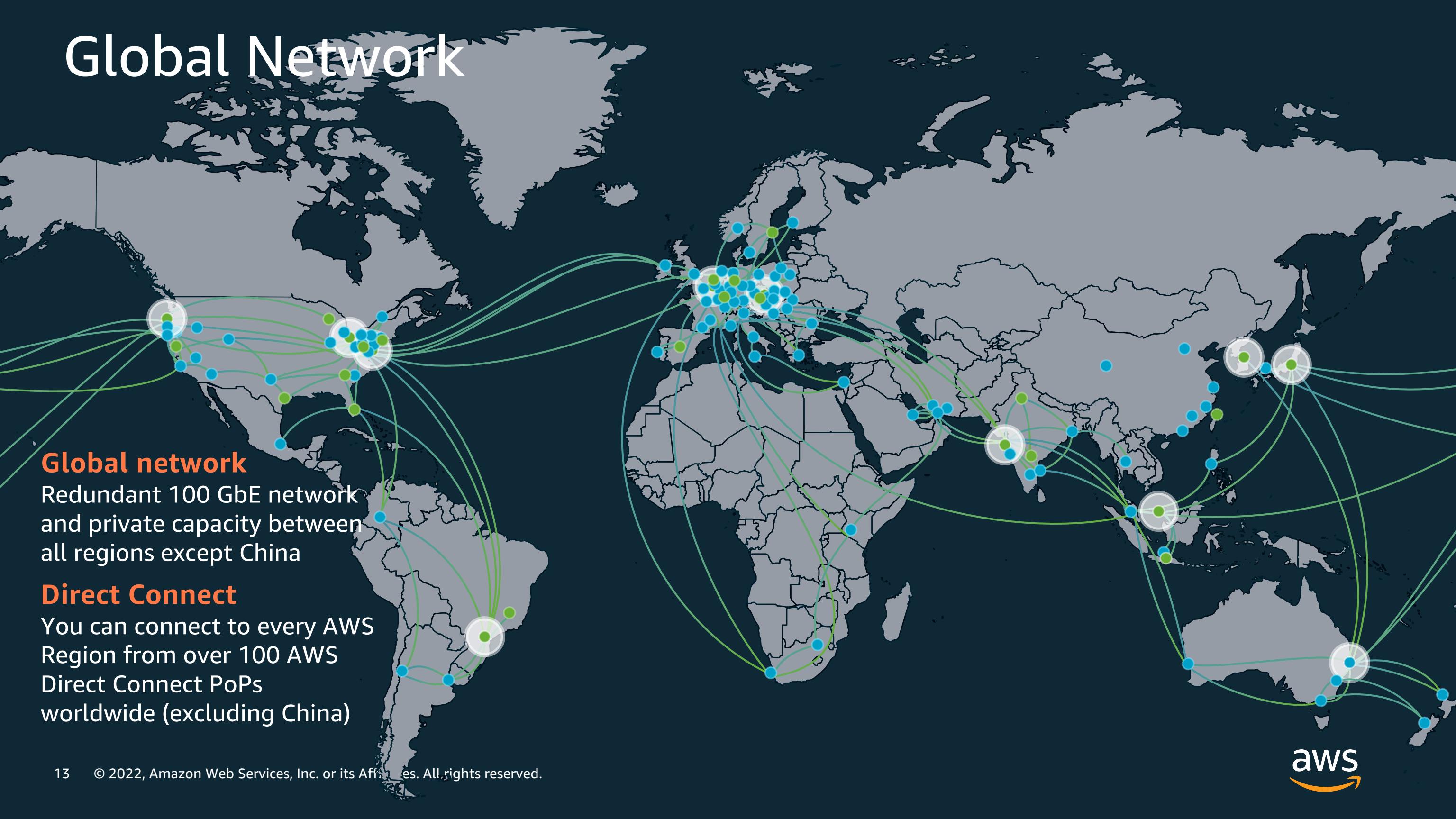
# Global Network

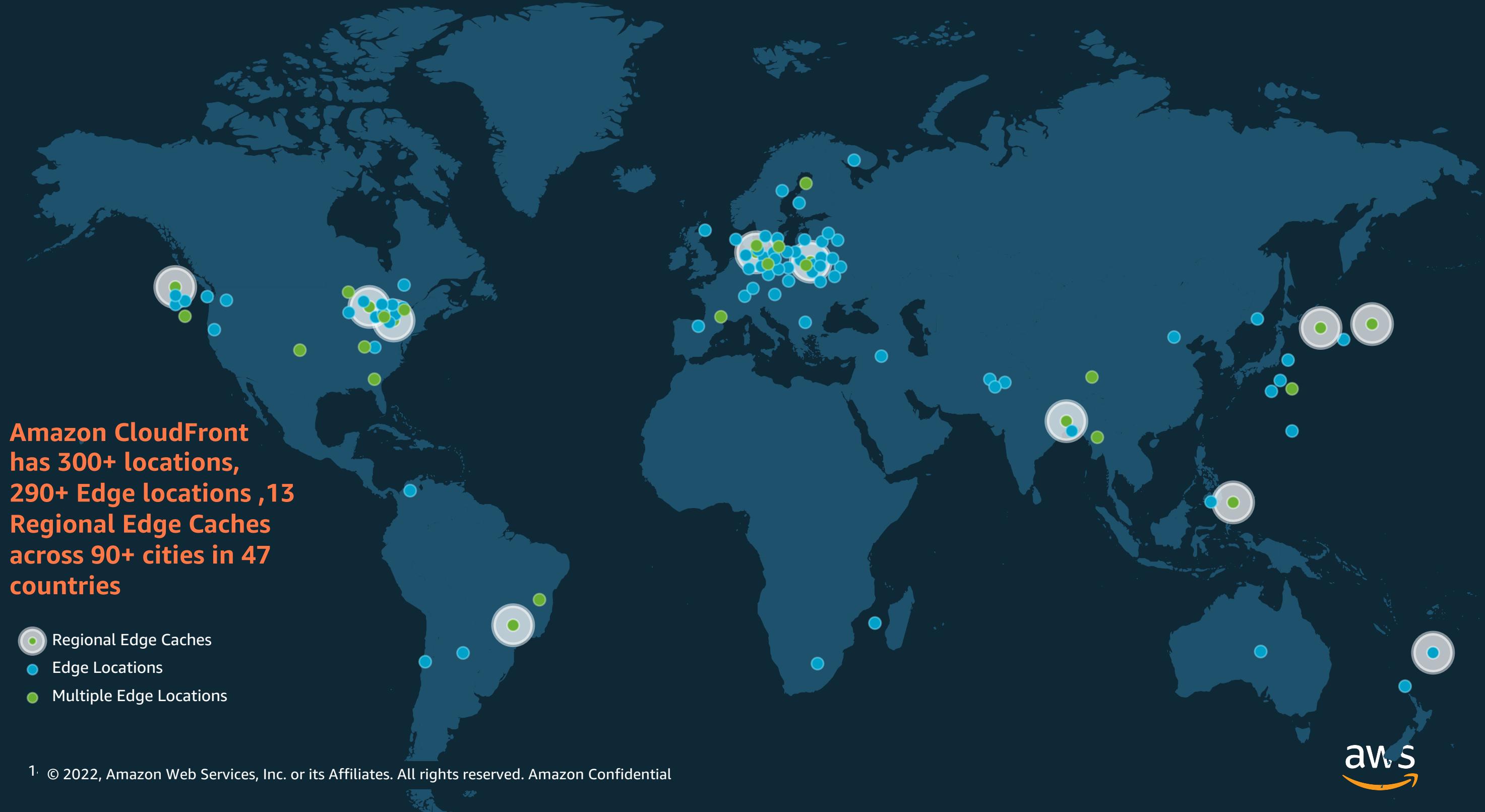
## Global network

Redundant 100 GbE 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)



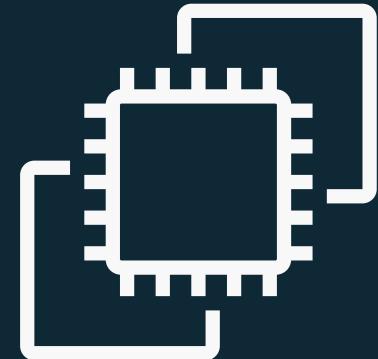


2

# EC2 Overview



# 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

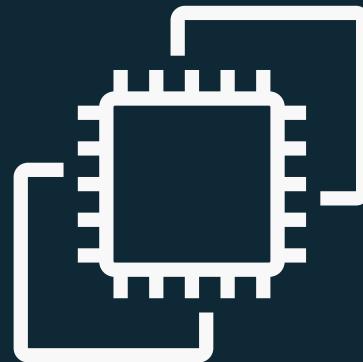


## AWS Lambda

Serverless compute  
for stateless code execution in  
response to triggers

# Virtual Machines

*Secure and scalable compute capacity in the cloud. Launch servers when needed without up-front commitments.*



**Amazon EC2**

- OS** Linux | Windows | Mac
- CPU** x86 | Arm (*Intel, AMD, NVIDIA, AWS Graviton*)
- Storage** Local | Network
- Purchase** On-demand | Reserved Instance | Spot
- Images** AWS Provided | Community | Custom
- Use Cases** Web | Application | Database | Container Host | ...

# EC2 Concepts

AMI

Reusable template for launching EC2 instances.

Security Groups

Instance-specific virtual stateful firewall.

Instance Type

Ratio of resources (vCPU, RAM, I/O, ...)

Instance Size

Amount of resources (vCPU, RAM, I/O, ...)

	vCPU / RAM (GB)		
	General (M5)	Compute (C5)	Memory (R5)
large	2 / 8	2 / 4	2 / 16
x-large	4 / 16	4 / 8	4 / 32
2x-large	8 / 32	8 / 16	8 / 64

*Examples*



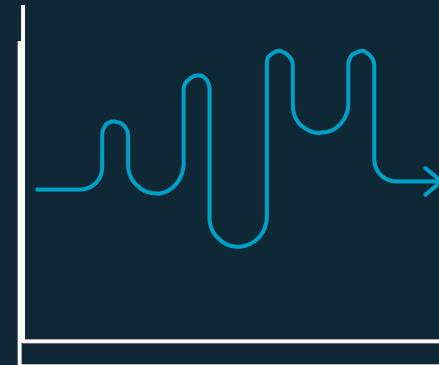
# Instance Types

	General Purpose	Compute Optimized	Memory Optimized	Accelerated Computing	Storage Optimized
	Burstable performance General Purpose	Compute Intensive Compute +memory up to 100 Gbps	Memory Optimized In-memory Memory Intensive Compute and Memory Intensive	Graphics Intensive General Purpose GPU FPGA	High I/O Dense Storage Big Data Optimized
<b>intel</b>	T3 M5	C5 C5n	R5 X1 X1e	G3 P3 F1	D2 H1
	M5d	C5d	R5d Z1d		I3
<b>AMD</b>	T3a M5a		R5a		
<b>metal</b>	M5m	c5m	R5m u-12tb1 Z1dm		I3m
<b>others</b>	A1 M6g	C6g	R6g	P3dn	I3en

# 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

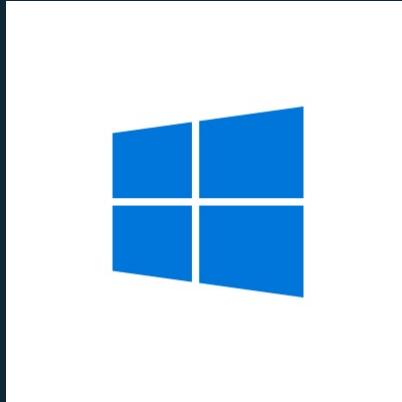
Spare Amazon EC2 capacity at **savings of up to 90%** off On-Demand prices



Fault-tolerant, flexible,  
stateless workloads

# EC2 Supported Operating Systems

- Windows 2003R2\*/2008\*/2008R2\*/2012/2012R2/2016/2019
- Amazon Linux
- Debian
- Suse
- CentOS
- Red Hat Enterprise Linux
- Ubuntu
- Mac



for more OSes see:

<https://aws.amazon.com/ec2/faqs/>

# Docker Containers

*Build applications, not infrastructure. Scale quickly and seamlessly. Security and isolation by design.*



**Amazon ECS, EKS  
and Fargate**

**Elastic  
Container  
Service (ECS)**

**Elastic  
Kubernetes  
Service (EKS)**

**Fargate**

Native managed container orchestration service to deploy, manage and scale containerized apps.

Managed certified Kubernetes service.

Automatically manages EC2 instances for ECS or EKS containers.

# Serverless Code Execution

*Run code without provisioning or managing servers or containers.*



AWS Lambda

- Runtimes** Node.js, Java, C#, Ruby, Go, .NET Core, Python
- Scaling** Continuous, automatic parallelization
- Triggers** HTTP Request, Object Event, Table Update, API, ...
- Pricing** By the millisecond

# Any Questions?

Up Next: EC2 Hands on Lab