



**VIDEO COURSE**

COURSE AUTHOR

Jon Bonso

- <https://au.linkedin.com/in/jonbonso>

COURSE LINK

- <https://portal.tutorialsdojo.com/courses/aws-certified-solutions-architect-associate-exam-video-course/>



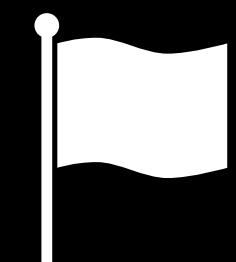
# AWS Solutions Architect Associate Exam Overview

---



Tutorials Dojo

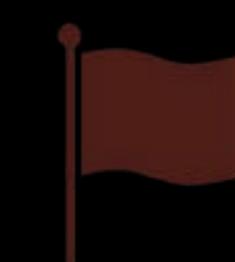
[www.tutorialsdojo.com](http://www.tutorialsdojo.com)



2013



2018



2020



2022

### FOUNDATIONAL

**Six months** of fundamental AWS Cloud and industry knowledge



### PROFESSIONAL

**Two years** of experience designing, operating, and troubleshooting solutions using the AWS Cloud



### ASSOCIATE

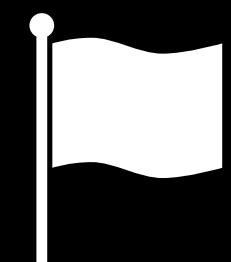
**One year** of experience solving problems and implementing solutions using the AWS Cloud



### SPECIALTY

Technical AWS Cloud experience in the Specialty domain as specified in the exam guide





2013



2018



SAA-C01



2020



SAA-C02



2022



**SAA-C03**

## Multiple Choice

Has 1 correct response and 3 incorrect responses

A company has a web application that uses Internet Information Services (IIS) for Windows Server. A file share is used to store the application data on the network-attached storage of the company's on-premises data center. To achieve a highly available system, they plan to migrate the application and file share to AWS.

Which of the following can be used to fulfill this requirement?

- Migrate the existing file share configuration to Amazon EFS.
- Migrate the existing file share configuration to Amazon EBS.
- Migrate the existing file share configuration to Amazon FSx for Windows File Server.
- Migrate the existing file share configuration to AWS Storage Gateway.

## Multiple Response

Has 2 correct responses out of 5 response options

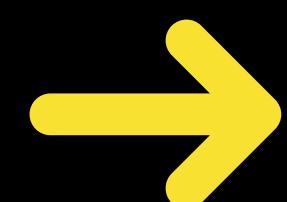
You are working as a Solutions Architect in a new startup that provides storage for high-quality photos which are infrequently accessed by the users. To make the architecture cost-effective, you designed the cloud service to use an S3 One Zone-Infrequent Access (S3 One Zone-IA) storage type for free users and an S3 Standard-Infrequent Access (S3 Standard-IA) storage type for premium users. When your manager found out about this, he asked you about the differences between using S3 One Zone-IA and S3 Standard-IA.

What will you say to your manager? (Select TWO.)

- Unlike other Amazon object storage classes, which store data in a minimum of three Availability Zones (AZs), S3 One Zone-IA stores data in a single AZ.
- Storing data in S3 One Zone-IA costs less than storing it in S3 Standard-IA.
- Storing data in S3 One Zone-IA costs more than storing it in S3 Standard-IA but provides more durability.
- Unlike other Amazon object storage classes, which store data in a minimum of three Availability Zones (AZs), S3 One Zone-IA stores data in two AZs only. Hence the name, One Zone-IA since the data replication is skipped in one Availability Zone.
- S3 One Zone-IA offers lower durability and low throughput compared with Amazon S3 Standard and S3 Standard-IA which is why it has a low per GB storage price and per GB retrieval fee.

	<b>Exam Code</b>	<b>SAA-C03</b>
	<b>Release Date</b>	<b>August 2022</b>
	<b>Prerequisites</b>	<b>None</b>
	<b>No. Of Questions</b>	<b>65</b>
	<b>Score Range</b>	<b>100 - 1000</b>
	<b>Passing Score</b>	<b>720</b>
	<b>Time Limit</b>	<b>2 hours 10 minutes</b>

Section	Score Performance		
	% of Scored Items	Needs Improvement	Meets Competencies
Section 1.0: Design Secure Architectures	30%		
Section 2.0: Design Resilient Architectures	26%		
Section 3.0: Design High-Performing Architectures	24%		
Section 4.0: Design Cost-Optimized Architectures	20%		



	<b>Exam Code</b>	<b>SAA-C03</b>
	<b>Release Date</b>	<b>August 2022</b>
	<b>Prerequisites</b>	<b>None</b>
	<b>No. Of Questions</b>	<b>65</b>
	<b>Score Range</b>	<b>100 - 1000</b>
	<b>Passing Score</b>	<b>720</b>
	<b>Time Limit</b>	<b>2 hours 10 minutes</b>



# AWS Certified Solutions Architect Associate Exam Domains

---

# Design Architectures

Domain 1:

**Design Secure  
Architectures**

Domain 2:

**Design Resilient  
Architectures**

Domain 3:

**Design High-Performing  
Architectures**

Domain 4:

**Design Cost-Optimized  
Architectures**



## AWS Certified Solutions Architect - Associate (SAA-C03) Exam Guide

### Exam results

The AWS Certified Solutions Architect - Associate exam is a pass or fail exam. The exam is scored against a minimum standard established by AWS professionals who follow certification industry best practices and guidelines.

Your results for the exam are reported as a scaled score of 100–1,000. The minimum passing score is 720. Your score shows how you performed on the exam as a whole and whether or not you passed. Scaled scoring models help equate scores across multiple exam forms that might have slightly different difficulty levels.

Your score report could contain a table of classifications of your performance at each section level. This information provides general feedback about your exam performance. The exam uses a compensatory scoring model, which means that you do not need to achieve a passing score in each section. You need to pass only the overall exam.

Each section of the exam has a specific weighting, so some sections have more questions than other sections have. The table contains general information that highlights your strengths and weaknesses. Use caution when interpreting section-level feedback. Candidates who pass the exam will not receive this additional information.

### Content outline

This exam guide includes weightings, test domains, and task statements for the exam. It is not a comprehensive listing of the content on the exam. However, additional context for each of the task statements is available to help guide your preparation for the exam. The following table lists the main content domains and their weightings. The table precedes the complete exam content outline, which includes the additional context. The percentage in each domain represents only scored content.

Domain	% of Exam
Domain 1: Design Secure Architectures	30%
Domain 2: Design Resilient Architectures	26%
Domain 3: Design High-Performing Architectures	24%
Domain 4: Design Cost-Optimized Architectures	20%
<b>TOTAL</b>	<b>100%</b>



## TASK STATEMENT

### Domain 1: Design Secure Architectures

Task Statement 1: Design secure access to AWS resources.

Knowledge of:

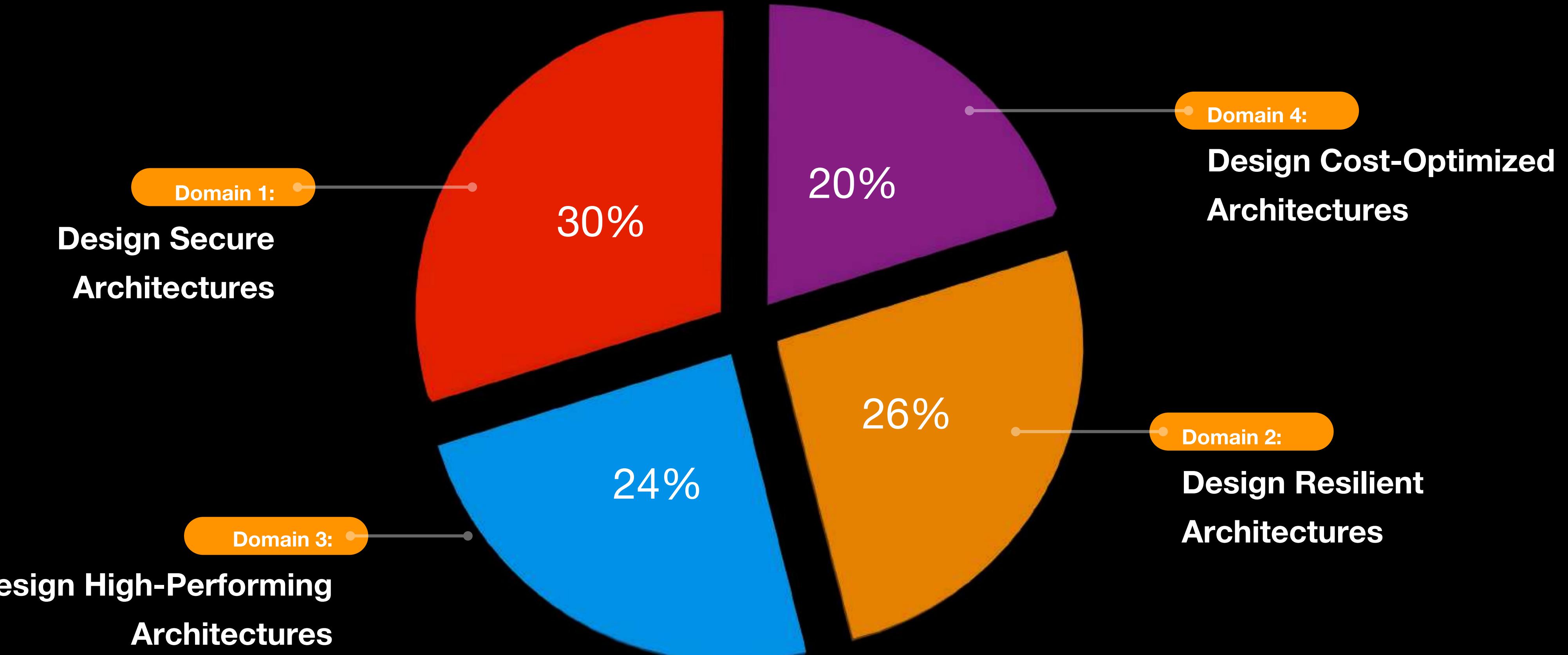
- Access controls and management across multiple accounts
- AWS federated access and identity services (for example, AWS Identity and Access Management [IAM], AWS Single Sign-On [AWS SSO])
- AWS global infrastructure (for example, Availability Zones, AWS Regions)
- AWS security best practices (for example, the principle of least privilege)
- The AWS shared responsibility model

Skills in:

- Applying AWS security best practices to IAM users and root users (for example, multi-factor authentication [MFA])
- Designing a flexible authorization model that includes IAM users, groups, roles, and policies
- Designing a role-based access control strategy (for example, AWS Security Token Service [AWS STS], role switching, cross-account access)
- Designing a security strategy for multiple AWS accounts (for example, AWS Control Tower, service control policies [SCPs])
- Determining the appropriate use of resource policies for AWS services
- Determining when to federate a directory service with IAM roles

## EXAM DOMAIN

- TASK STATEMENT #1
- TASK STATEMENT #2
- TASK STATEMENT #3



Domain 1:

**Design Secure  
Architectures**

Domain 2:

**Design Resilient  
Architectures**

Domain 3:

**Design High-Performing  
Architectures**

Domain 4:

**Design Cost-Optimized  
Architectures**

- **Design secure access to AWS resources**
- **Design secure workloads and applications**
- **Determine appropriate data security controls**

Domain 1:

**Design Secure  
Architectures**

Domain 2:

**Design Resilient  
Architectures**

Domain 3:

**Design High-Performing  
Architectures**

Domain 4:

**Design Cost-Optimized  
Architectures**

- **Design scalable and loosely coupled architecture**
- **Design highly available and/or fault-tolerant architectures**

Domain 1:

**Design Secure  
Architectures**

Domain 2:

**Design Resilient  
Architectures**

Domain 3:

**Design High-Performing  
Architectures**

Domain 4:

**Design Cost-Optimized  
Architectures**

- **Determine high-performing and/or scalable storage solutions**
- **Design high-performing and elastic compute solutions**
- **Determine high-performing database solutions**
- **Determine high-performing and/or scalable network architectures**
- **Determine high-performing data ingestion and transformation solutions**

Domain 1:

**Design Secure  
Architectures**

Domain 2:

**Design Resilient  
Architectures**

Domain 3:

**Design High-Performing  
Architectures**

Domain 4:

**Design Cost-Optimized  
Architectures**

- **Design cost-optimized storage solutions**
- **Design cost-optimized compute solutions**
- **Design cost-optimized database solutions**
- **Design cost-optimized network architectures**

## Appendix

### Which key tools, technologies, and concepts might be covered on the exam?

The following is a non-exhaustive list of the tools and technologies that could appear on the exam. This list is subject to change and is provided to help you understand the general scope of services, features, or technologies on the exam. The general tools and technologies in this list appear in no particular order. AWS services are grouped according to their primary functions. While some of these technologies will likely be covered more than others on the exam, the order and placement of them in this list is no indication of relative weight or importance:

- Compute
- Cost management
- Database
- Disaster recovery
- High performance
- Management and governance
- Microservices and component decoupling
- Migration and data transfer
- Networking, connectivity, and content delivery
- Resiliency
- Security
- Serverless and event-driven design principles
- Storage

### AWS services and features

#### Analytics:

- Amazon Athena
- AWS Data Exchange
- AWS Data Pipeline
- Amazon EMR
- AWS Glue
- Amazon Kinesis
- AWS Lake Formation
- Amazon Managed Streaming for Apache Kafka (Amazon MSK)
- Amazon OpenSearch Service (Amazon Elasticsearch Service)
- Amazon QuickSight
- Amazon Redshift

#### Application Integration:

- Amazon AppFlow
- AWS AppSync
- Amazon EventBridge (Amazon CloudWatch Events)
- Amazon MQ
- Amazon Simple Notification Service (Amazon SNS)
- Amazon Simple Queue Service (Amazon SQS)
- AWS Step Functions



# AWS Overview

**WHAT**

is AWS?

**WHEN**

did AWS start?

**WHY**

is AWS so popular?

**WHAT**  
is AWS?

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

**AWS**

**WHAT**  
is AWS?

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

# Amazon Web Services

**WHAT**  
is AWS?

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

Amazon

Web  
Services



amazon.com

**WHAT**  
is AWS?

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

Amazon

Web  
Services



**WHAT**  
is AWS?

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

**Amazon  
Web  
Services** =

**Cloud Service Provider**

- provides a cloud-based platform or cloud services
- Allows you to **rent out** virtual servers that you access remotely

**WHAT**  
is AWS?

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

# Cloud Service Provider

is *like* a

# Car Rental



**WHAT**  
is AWS?

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

# Cloud Service Provider

With different types of **CPU**, **Storage**, **Network** and other components that you can choose from!



**Virtual Machines**



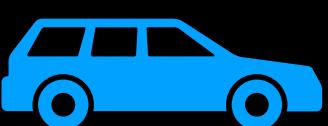
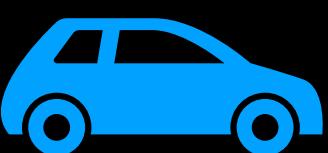
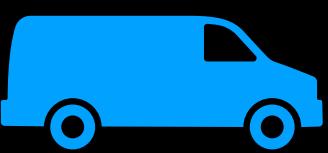
**Physical Servers**



**Storage Appliances**

**Network Devices**

# Car Rental



**WHAT**  
is AWS?



2004

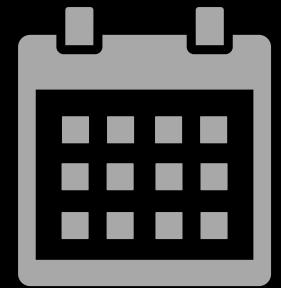
**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

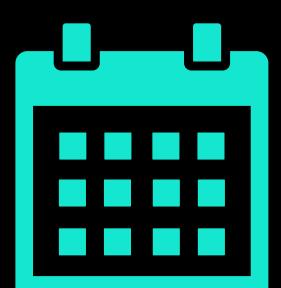
- AWS started out as a department **within** Amazon Inc.
- Used only by early Amazon customers
- Web services are not available publicly

# WHAT

is AWS?



2004



2006

# WHEN

did AWS start?

# WHY

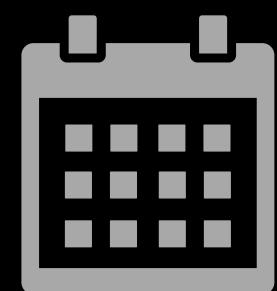
is AWS so popular?

- AWS officially started its operation as a **public cloud service provider**
- Released Amazon S3 (Simple Storage Service)
- Released Amazon SQS (Simple Queue Service)

**WHAT**  
is AWS?



2004



2006



Today

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?

- Offers hundreds of fully-featured services that are **available globally**
- Provides a highly reliable, scalable, and **low-cost** infrastructure platform in the cloud
- Boasts a broad set of cloud-based products

**WHAT**  
is AWS?

**WHEN**  
did AWS start?

**WHY**  
is AWS so popular?



Today



is the **world's leading cloud platform.**

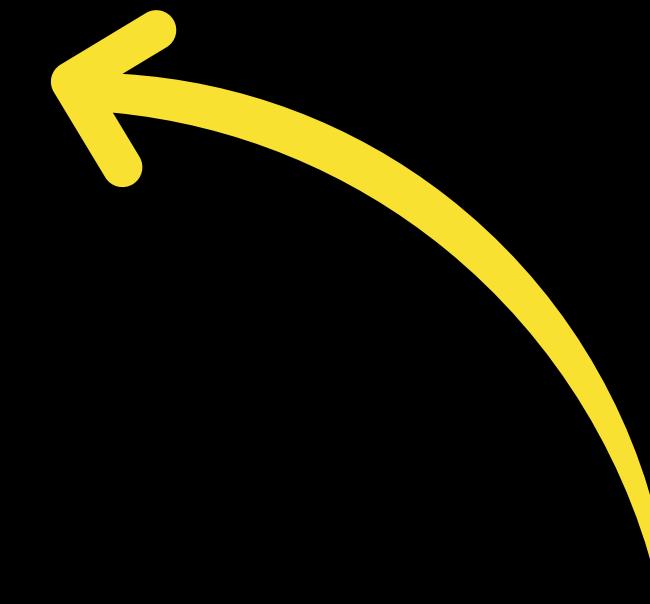
- Used by **millions** of customers
- Supports various workloads
- Significantly lower your operating costs
- Enables companies to scale globally in minutes!



# AWS Global Infrastructure

---

**Has thousands of servers!**



**These *physical* servers generate  
*virtual* machines or store your data!**

## Availability Zone

## Region

## Edge Networks



## Edge Networks

Improves the “**Availability**” of your systems

Region

Literally a Geographic “**Zone**”

Availability Zone



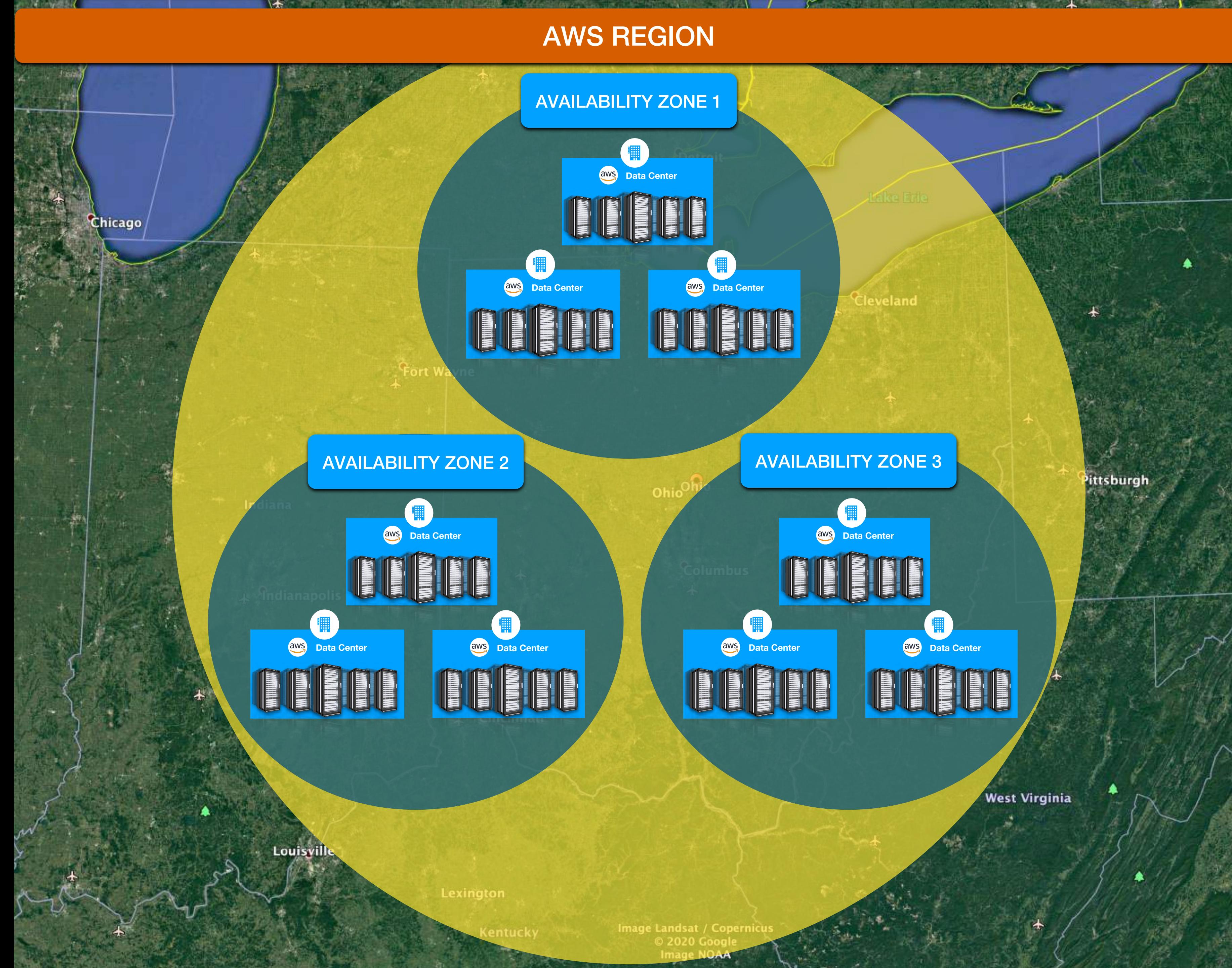
100 kilometers or 60 miles from each other

## AWS REGION

Edge Networks

Region

Availability Zone

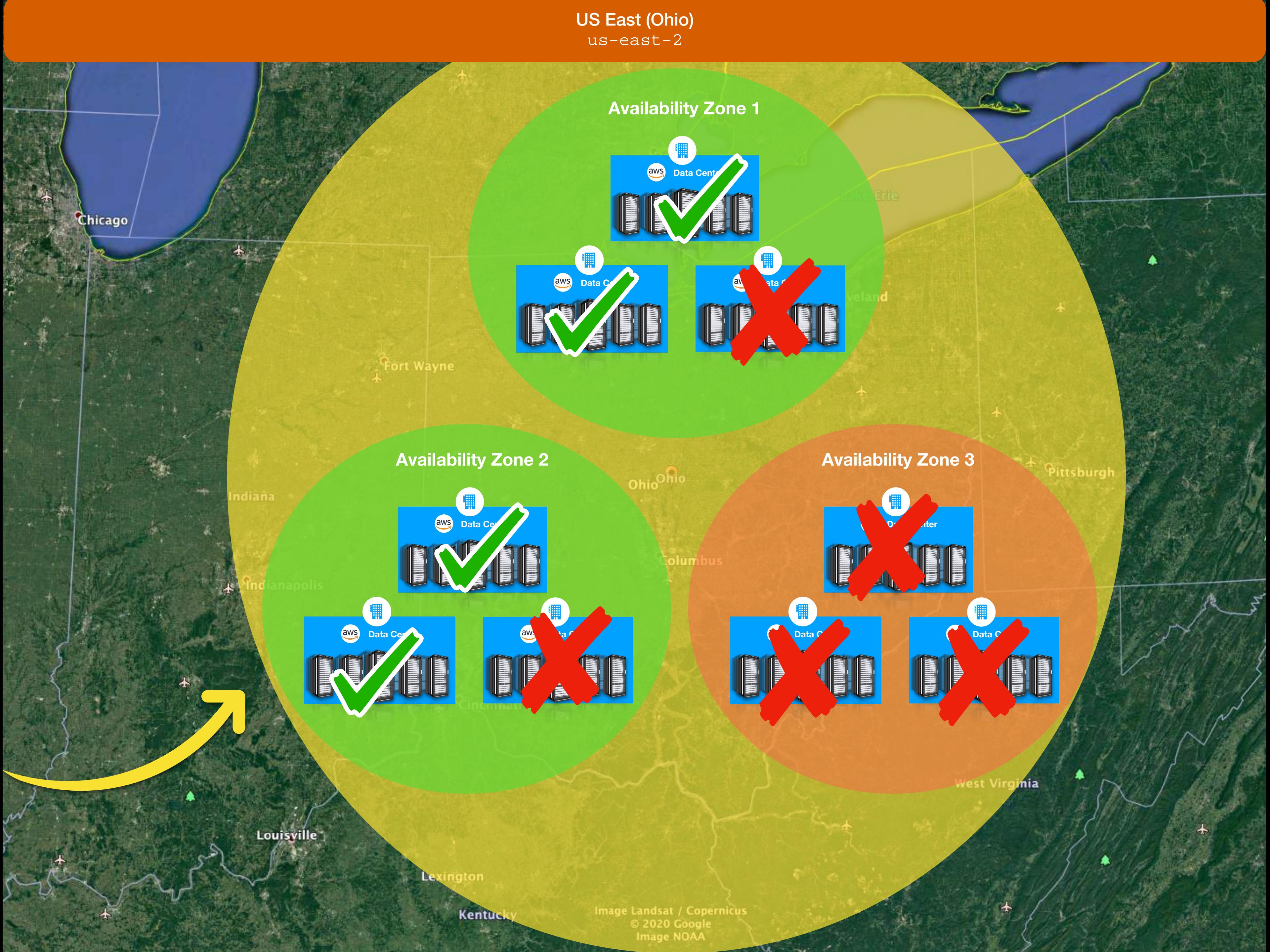


## Edge Networks

### Region

### Availability Zone

Your system will still run  
even if one or more data centers  
encountered an outage



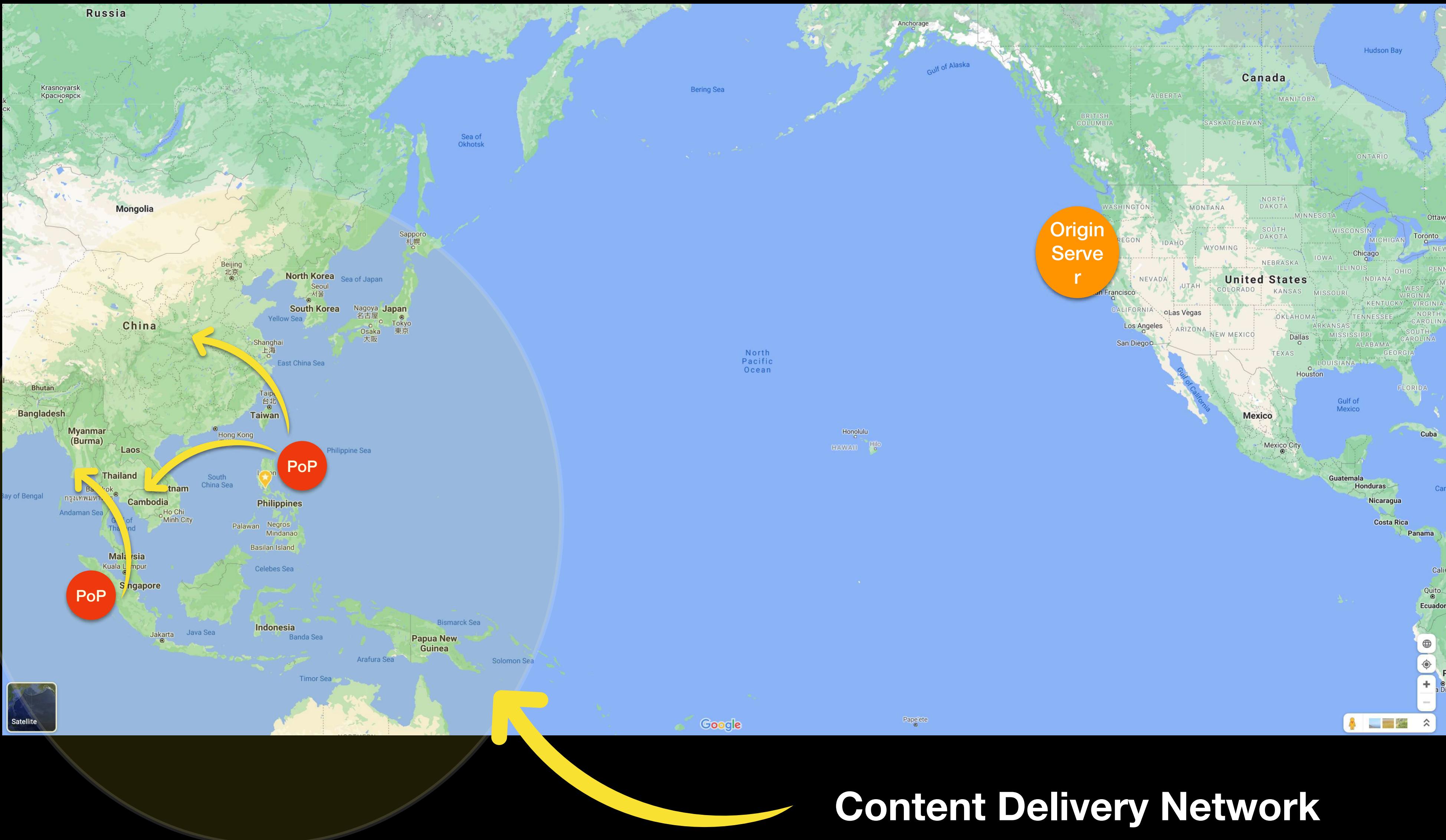
# Edge Networks

Point of Presence / Edge Location

Region

Availability Zone

Content Delivery Network





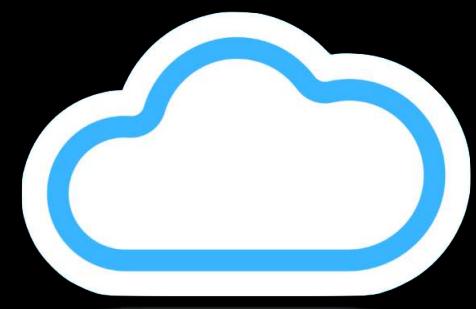
# Advantages of Cloud Computing

---

# Advantages of Cloud Computing



- Launch solutions and computing resources **in a matter of minutes**
- No need to buy & maintain costly physical servers or data centers
- On-demand access to a wide range of virtual machines, storage services, databases, and other IT resources
- Revolutionary **Cloud Economics**
- Unparalleled **Flexibility** for your enterprise IT infrastructure
- Better **Price-to-Performance Ratio**
- Lower **Total Cost of Ownership (TCO)**



# Advantages of Cloud Computing



Trade Fixed Expense for Variable Expense



Benefit from Massive Economies of Scale



Stop Guessing Capacity



Increase Speed and Agility



Stop Spending Money Running & Maintaining Data Centers



Go Global in Minutes



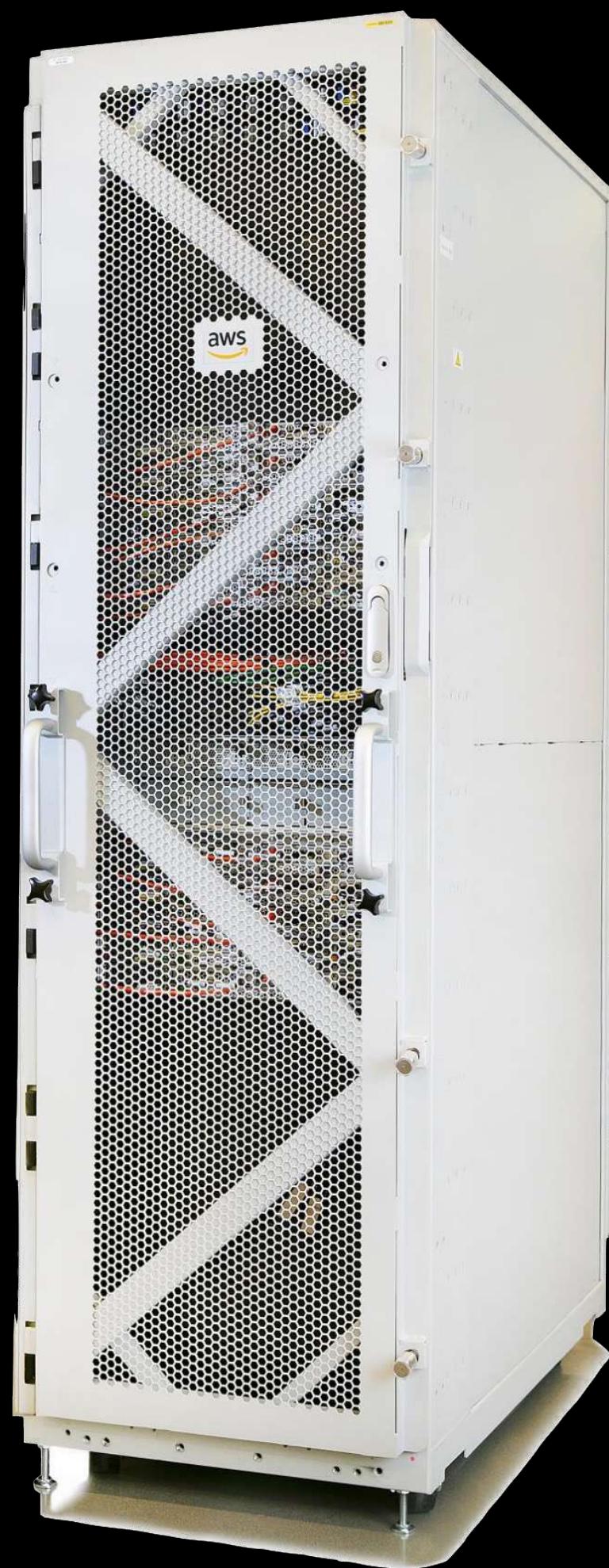
# AWS Shared Responsibility Model

---

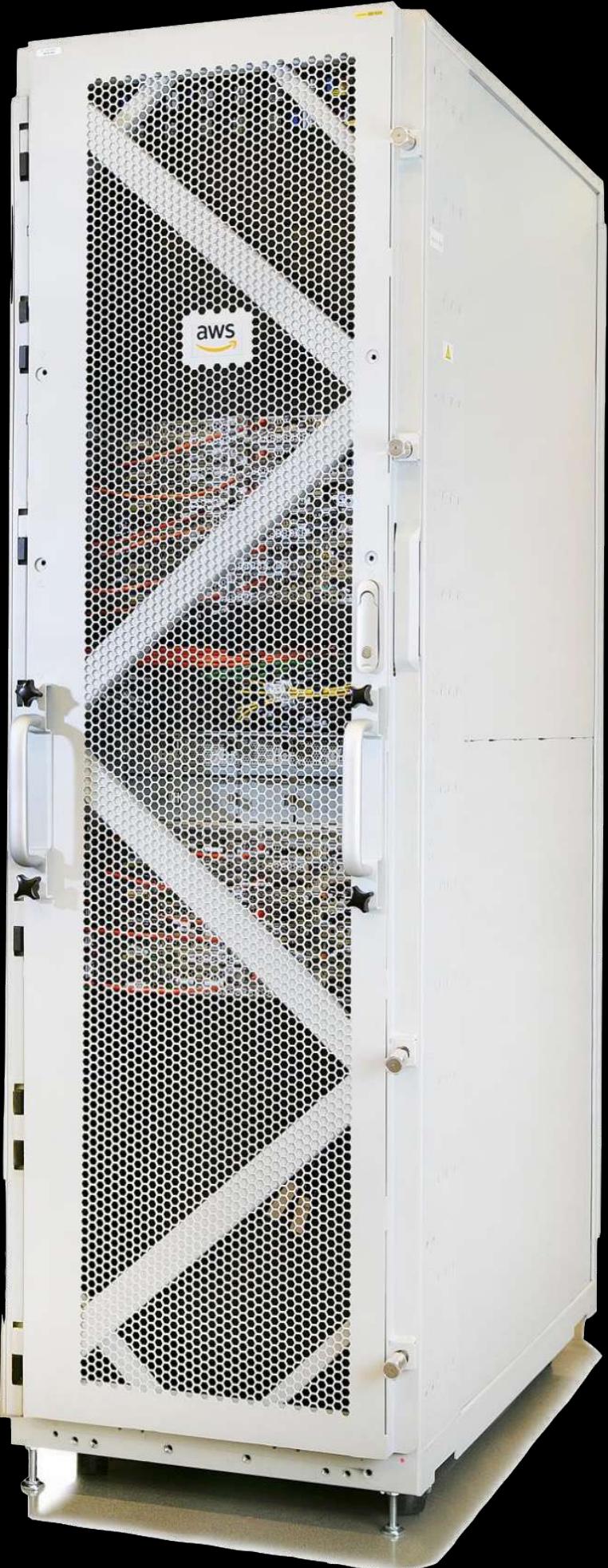
**CLOUD COMPUTING**

A model for enabling ubiquitous,  
convenient, on-demand network  
access to a shared pool of  
configurable computing  
resources  
that can be rapidly provisioned  
and released with minimal  
management effort  
or service provider interaction.

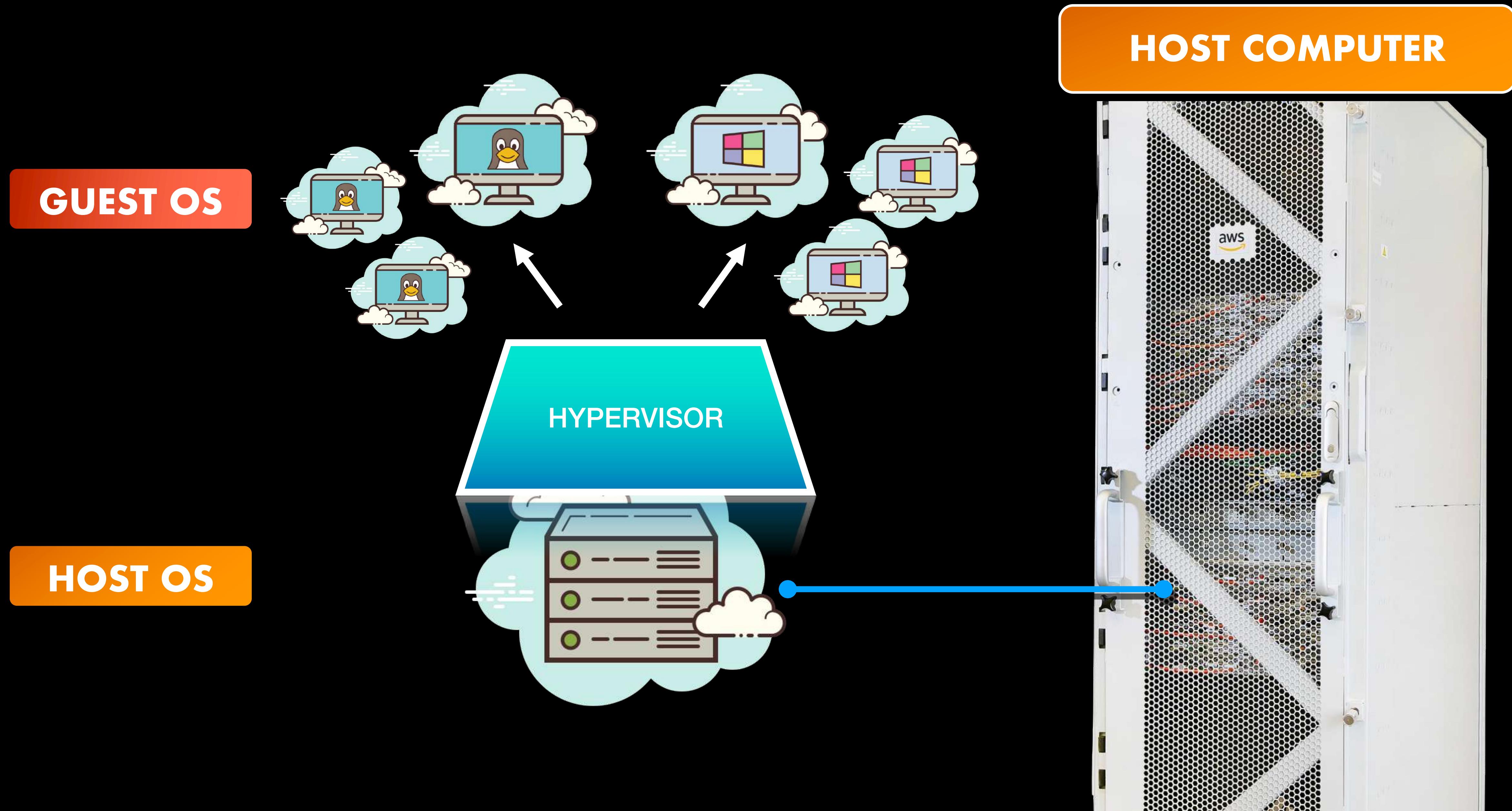
# configurable computing resources

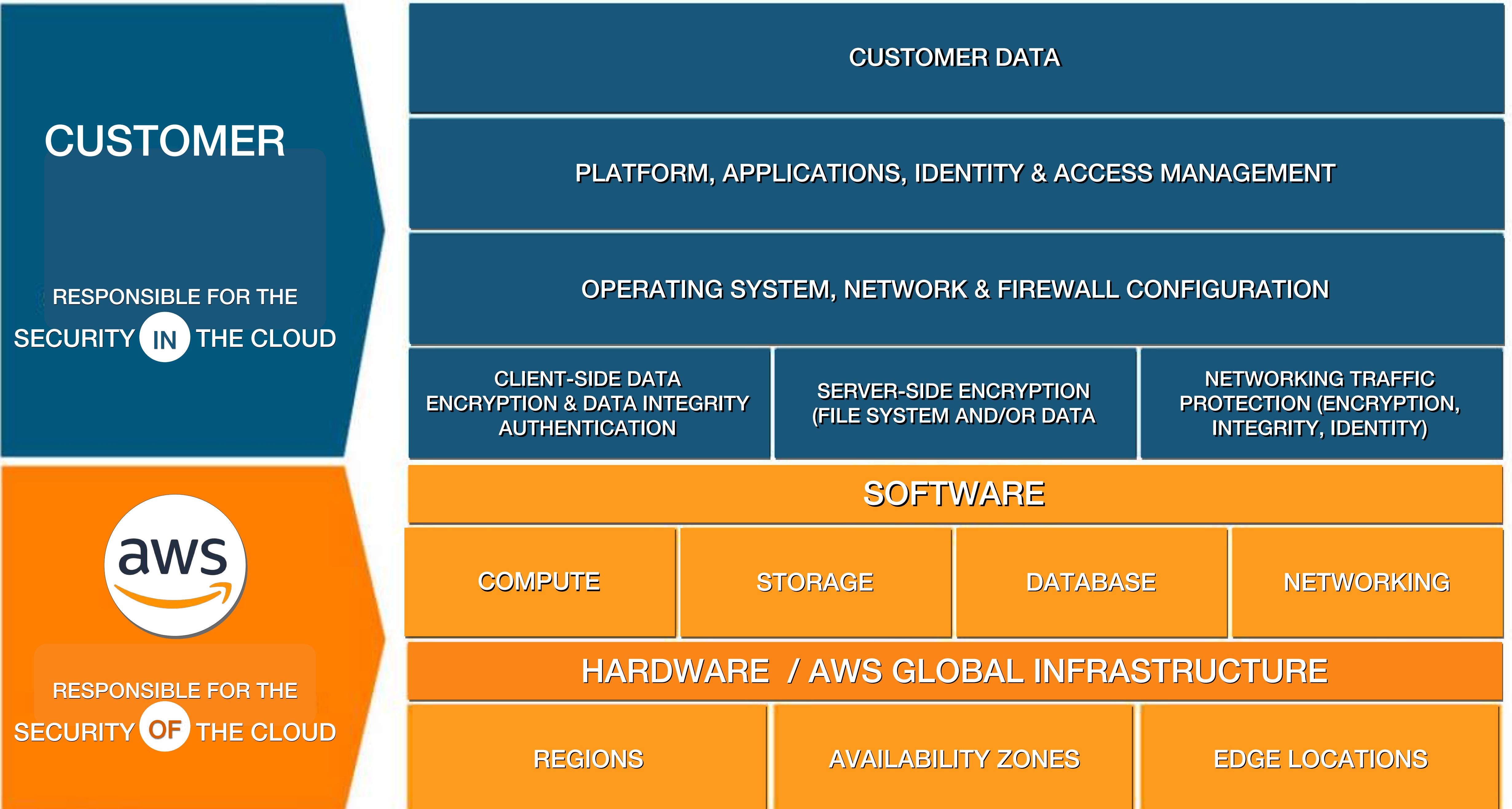


# configurable computing resources



# configurable computing resources





# WHO?

Who is responsible for **patching the operating system** of your Amazon EC2 instance?

Who is responsible for **applying the security patches of the guest operating system** that your EC2 instance is using?

Who is responsible for **running the host operating system and the virtualization layer** that powers your Amazon EC2 instances?

Who is responsible for **managing all your IAM user access** and secret keys?

Who is responsible for **maintaining the underlying server** of your AWS Lambda functions?

Who is responsible for **the Service and Communications Protection or Zone Security** of your data?

Who is responsible for **the physical security of the servers and the entire network of data centers** of the AWS Global Infrastructure?

Who is responsible for **designing encryption-at-rest strategies** and other security features in your Amazon RDS database?

Who is responsible for **the security OF the cloud** and the **security IN the cloud**?



CUSTOMER

HOST OS

GUEST OS

ABSTRACTED  
SERVICES

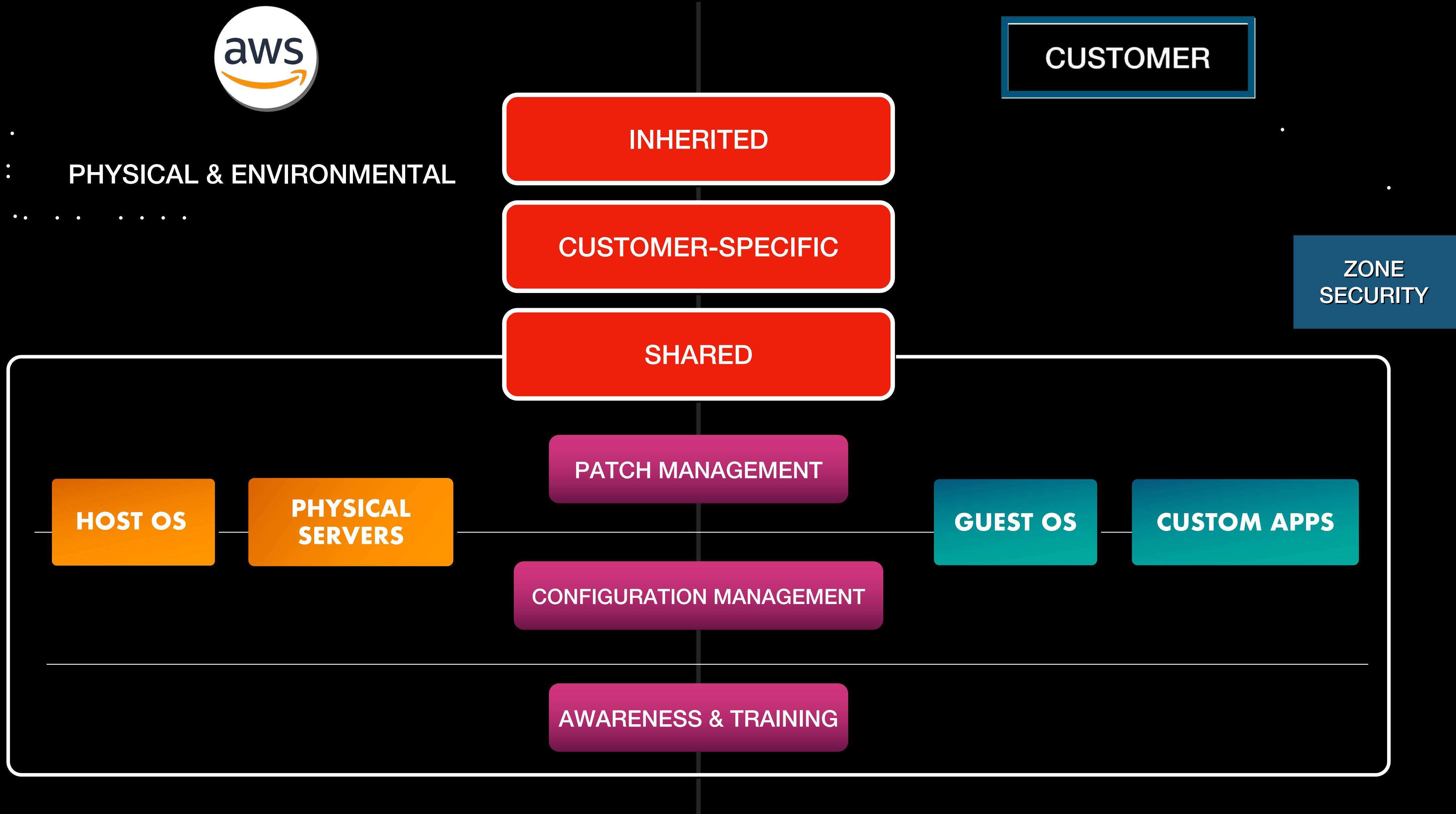
INFRASTRUCTURE  
SECURITY

CLIENT-SIDE & SERVER-SIDE  
DATA ENCRYPTION

ZONE SECURITY

CONFIGURATION  
MANAGEMENT

# IT CONTROLS





HOST OS

CUSTOMER GUEST OS

Who is responsible for **patching the operating system** of your Amazon EC2 instance?

CUSTOMER



Who is responsible for **applying the security patches of the guest operating system** that your EC2 instance is using?

Who is responsible for **running the host operating system and the virtualization layer** that powers your Amazon EC2 instances?

CUSTOMER



Who is responsible for **managing all your IAM user access** and secret keys?

Who is responsible for **maintaining the underlying server** of your AWS Lambda functions?

CUSTOMER

Who is responsible for the **Service and Communications Protection or Zone Security** of your data?



Who is responsible for the **physical security of the servers and the entire network of data centers** of the AWS Global Infrastructure?



Who is responsible for **designing encryption-at-rest strategies** and other security features in your Amazon RDS database?



CUSTOMER IN

Who is responsible for the **security OF the cloud** and the **security IN the cloud**?



# AWS Support Plans

---



TOOLS



TECHNOLOGY



PEOPLE



PROGRAMS



BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

FREE

\$

\$\$

\$\$\$

\$\$\$\$

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

RESPONSE TIME

ARCHITECTURAL  
GUIDANCE

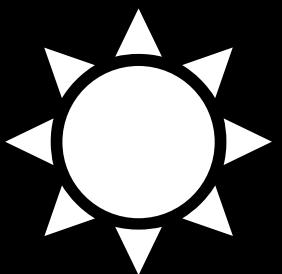
PROGRAMMATIC  
CASE MANAGEMENT

3RD-PARTY  
SOFTWARE SUPPORT

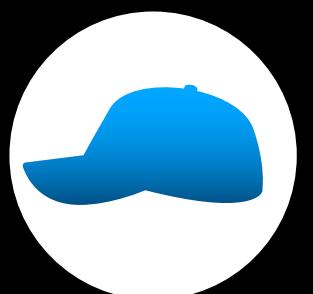
PROACTIVE SELF SERVICE  
PROGRAMS

TECHNICAL ACCOUNT  
MANAGEMENT (TAM)

ACCOUNT  
ASSISTANCE



30 DAY  
MINIMUM TERM



CLOUD SUPPORT  
ASSOCIATES



CONCIERGE SUPPORT  
TEAM



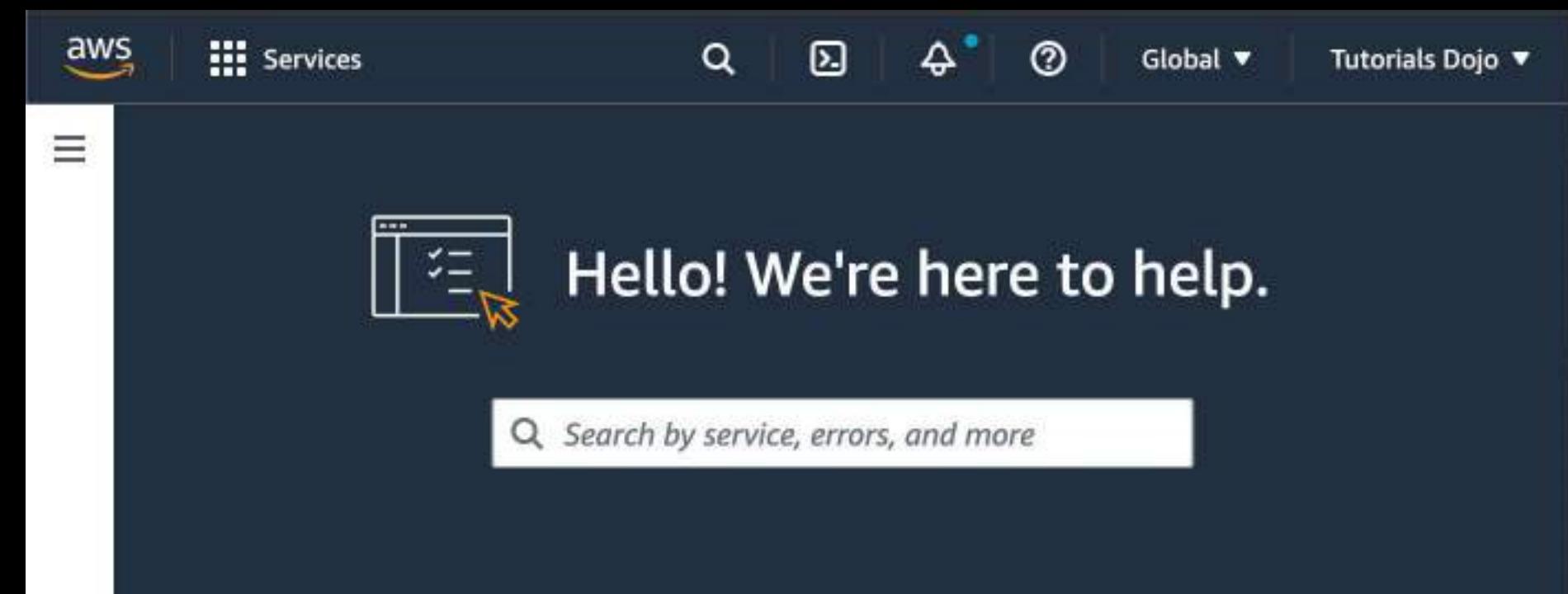
AWS MANAGED SERVICES  
TEAM



CLOUD SUPPORT  
ENGINEERS



TECHNICAL ACCOUNT  
MANAGER



BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

- Included for all AWS customers by default **FREE**
- 24/7 access to the AWS customer service, documentation, whitepapers & AWS re:Post site
- SLOW **RESPONSE TIME**
- Access to the **AWS Personal Health Dashboard**
- Access to the core security & service quota checks in **AWS Trusted Advisor** **LIMITED ACCESS**

BASIC

DEVELOPER

BUSINESS

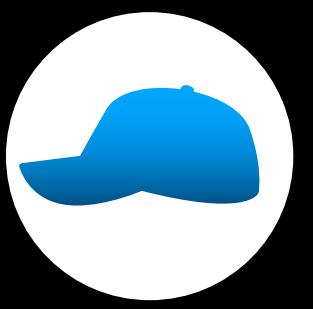
ENTERPRISE  
ON-RAMP

ENTERPRISE

- Recommended for testing or for running non-critical production workloads in AWS
- Access to the core security & service quota checks in **AWS Trusted Advisor** LIMITED ACCESS

## ENHANCED TECHNICAL SUPPORT

- Support provided by:
- Unlimited support cases with **1 primary contact**
- Prioritized responses on AWS re:Post
- Support Schedule: **Business Hours**



CLOUD SUPPORT  
ASSOCIATES

8 AM - 6 PM

MON - FRI

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

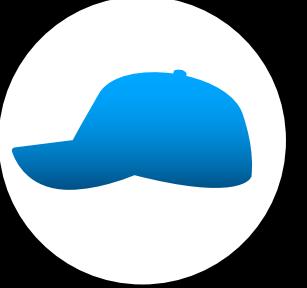
ENTERPRISE

## ENHANCED TECHNICAL SUPPORT

ARCHITECTURAL  
GUIDANCE

BASIC

## RESPONSE TIMES

- Support provided by:  CLOUD SUPPORT ASSOCIATES
- Unlimited support cases with **1 primary contact**
- Prioritized responses on AWS re:Post
- Support Schedule: **Business Hours**

8 AM - 6 PM

MON - FRI

- General guidance: < 24 hours
- System impaired: < 12 hours
- **NO** Phone or Chat Assistance



**BASIC****DEVELOPER****BUSINESS****ENTERPRISE  
ON-RAMP****ENTERPRISE****AWS Systems Manager**

Owned by Amazon   Owned by me   Shared with me   Favorites - new   All documents

**Document categories**

- AWS Documentation  
AWS user guides, tutorials
- Remediation  
Remediating common issues
- Patching  
Patching workflows
- Security  
Enforcing security best practices
- Instance management  
Tasks for EC2, EBS
- Disaster recovery and backup  
Disaster recovery, data backup, and snapshots
- AMI management  
Manage Windows and Linux AMIs
- Self service support workflows  
Troubleshooting, diagnostics and recovery
- Resource management  
Tasks for AWS resources including RDS, CloudFormation, S3, DDB, etc.
- Cost management  
Workflows to optimize costs
- Migration  
Migrate on-premises nodes to AWS
- Others

**Automation document**

Search by keyword or filter by tag or attributes

Search: AWSSupport X   Clear filters

<b>AWSSupport-ActivateWindowsWithAmazonLicense</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-AnalyzeEMRLogs</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-CalculateEBSPerformanceMetrics</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-CheckAndMountEFS</b> Owner Amazon   Platform types Windows, Linux, MacOS
<b>AWSSupport-CheckXenToNitroMigrationRequirements</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-CloneXenEC2InstanceAndMigrateToNitro</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-CollectAmazonConnectContactFlowLog</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-CollectECSInstanceLogs</b> Owner Amazon   Platform types Windows, Linux, MacOS
<b>AWSSupport-CollectEKSInstanceLogs</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-CollectElasticBeanstalkLogs</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-ConfigureDNSQueryLogging</b> Owner Amazon   Platform types Windows, Linux, MacOS	<b>AWSSupport-ConfigureEC2Metadata</b> Owner Amazon   Platform types Windows, Linux, MacOS

## SUPPORT AUTOMATION WORKFLOWS (SAW)

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

\*NOT SUPPORTED IN THE  
DEVELOPER PLAN

BASIC  
RUNBOOK

PREMIUM  
RUNBOOK



AWS Systems Manager

**SUPPORT AUTOMATION  
WORKFLOWS (SAW)**

AWSSupport-

AWSPremiumSupport -

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

BASIC  
RUNBOOK



AWS Systems Manager

## SUPPORT AUTOMATION WORKFLOWS (SAW)

- AWSSupport-CopyEC2Instance
- AWSSupport-ResetAccess
- AWSSupport-ExecuteEC2Rescue
- AWSSupport-ListEC2Resources

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

- Has all the features of the **DEVELOPER** support plan
- Recommended if you have one or more production workloads in AWS
- Access to full best practice checks in **AWS Trusted Advisor**

FULL ACCESS

## ENHANCED TECHNICAL SUPPORT

ARCHITECTURAL  
GUIDANCE

CONTEXTUAL

- Support provided by:
- Unlimited support cases by Unlimited Contacts (IAM Supported)
- Support Schedule: **24/7**
- Prioritized responses on AWS re:Post
- Access to AWS Support App in



CLOUD SUPPORT  
ENGINEERS



BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

## RESPONSE TIMES

- General guidance:
- System impaired:
- Production system impaired
- Production system outage

< 24 hours

< 12 hours

< 4 hours

< 1 hour



Tutorials Dojo

[www.tutorialsdojo.com](http://www.tutorialsdojo.com)

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE



AWS SUPPORT  
API

- A web service that provides programmatic access to AWS Support Center operations
- API endpoint: <https://support.<region>.amazonaws.com>
- Supports the following operations:
  - Support Case Management Operations
  - AWS Trusted Advisor operations

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

## 3RD-PARTY SOFTWARE SUPPORT



AWS Systems Manager

### SUPPORT AUTOMATION WORKFLOWS (SAW)

BASIC  
RUNBOOK

AWSSupport -

PREMIUM  
RUNBOOK

AWSPremiumSupport -

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

## INFRASTRUCTURE EVENT MANAGEMENT

- Available for an **additional fee**.
- Offers architecture guidance and operational support during the preparation and execution of your **planned events** (e.g. scheduled shopping holiday, product launches, system migrations et cetera)
- Prevents unnecessary system degradation or site outages by **optimizing your cloud architecture prior to your event**
- Allows you to easily assess operational readiness, mitigate risks, and execute your planned activity confidently with **assistance from AWS experts**

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE



## AWS MANAGED SERVICES TEAM

- Available for an **additional** fee.
- Helps you operate your AWS infrastructure on your behalf
- Augments your existing internal teams with advanced cloud operation skills
- Provides you with AWS experts such as a designated Cloud Service Delivery Manager, a Cloud Architect, an AMS security team, or all three.

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

- Recommended if you have business-critical production workloads with strict SLA (high RTO and RPO requirements)
- Has all the features of the **BUSINESS** support plan

## RESPONSE TIMES

- General guidance: < 24 hours
- System impaired: < 12 hours
- Production system impaired < 4 hours
- Production system outage < 1 hour
- Business-critical system outage < 30 mins

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE



CONCIERGE SUPPORT  
TEAM



TECHNICAL ACCOUNT  
MANAGER

INFRASTRUCTURE  
EVENT MANAGEMENT

ARCHITECTURAL  
GUIDANCE

- Primary contact for AWS Billing & AWS Support
- Access to a **pool of Technical Account Managers** to provide proactive guidance and assistance
- Included without any additional fees
- Use for 1 Event per year only
- Consultative review
- Architectural Guidance based on your applications (**one-per-year only**)

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

- Recommended if you have **mission-critical** production workloads with strict SLA (high RTO and RPO requirements)
- Has all the features of the **ENTERPRISE ON-RAMP** support plan
- **Most expensive** AWS Support Plan



- Access to the premium **AWS Trusted Advisor Priority** feature

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

## RESPONSE TIMES

- General guidance:
- System impaired:
- Production system impaired
- Production system outage
- Business/Mission-critical system outage

< 24 hours

< 12 hours

< 4 hours

< 1 hour

< 15 mins

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE

## INFRASTRUCTURE EVENT MANAGEMENT

- Can be used for multiple corporate events

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE



ONLINE SELF-PACED  
LABS



AWS Incident Detection  
and Response

- Provide a hands-on learning environment based on real-world scenarios.

- Available for an **additional** fee
- 24/7 proactive monitoring & incident management for your selected production workloads that are regularly conducted by AWS experts.

AWS SUPPORT  
PROACTIVE SERVICES

- Workload reviews, best practices workshops, and deep dives delivered by AWS Experts



TECHNICAL ACCOUNT  
MANAGER

- Access to a **dedicated Technical Account Manager**



# AWS Support Plans

BASIC

DEVELOPER

BUSINESS

ENTERPRISE  
ON-RAMP

ENTERPRISE



# AWS Well-Architected Framework

---



# AWS Well-Architected Framework

- Conceptualized from extensive years of cloud research, development, and real-world experience
- A knowledge base of design principles, best practices and architectural guidance
- Helps you avoid costly mistakes
- Allows you to establish key performance indicators (KPIs) to measure workload performance

# AWS Well-Architected Framework



■	AWS Well-Architected Framework
■	Table of Contents
▼	AWS Well-Architected Framework
	Introduction
	Definitions
	On architecture
	General design principles
▼	The pillars of the framework
>	Operational excellence
>	Security
>	Reliability
>	Performance efficiency
>	Cost optimization
>	Sustainability
	The review process
	Conclusion
	Contributors
	Further reading
	Document revisions
▼	Appendix: Questions and best practices
>	Operational excellence
>	Security
>	Reliability
>	Performance efficiency
>	Cost optimization
>	Sustainability
	Notices
	AWS glossary

## AWS Well-Architected Framework

Publication date: October 20, 2022 ([Document revisions \(p. 47\)](#))

The AWS Well-Architected Framework helps you understand the pros and cons of decisions you make while building systems on AWS. By using the Framework you will learn architectural best practices for designing and operating reliable, secure, efficient, cost-effective, and sustainable systems in the cloud.

### Introduction

The AWS Well-Architected Framework helps you understand the pros and cons of decisions you make while building systems on AWS. Using the Framework helps you learn architectural best practices for designing and operating secure, reliable, efficient, cost-effective, and sustainable workloads in the AWS Cloud. It provides a way for you to consistently measure your architectures against best practices and identify areas for improvement. The process for reviewing an architecture is a constructive conversation about architectural decisions, and is not an audit mechanism. We believe that having well-architected systems greatly increases the likelihood of business success.

AWS Solutions Architects have years of experience architecting solutions across a wide variety of business verticals and use cases. We have helped design and review thousands of customers' architectures on AWS. From this experience, we have identified best practices and core strategies for architecting systems in the cloud.

The AWS Well-Architected Framework documents a set of foundational questions that allow you to understand if a specific architecture aligns well with cloud best practices. The framework provides a consistent approach to evaluating systems against the qualities you expect from modern cloud-based systems, and the remediation that would be required to achieve those qualities. As AWS continues to evolve, and we continue to learn more from working with our customers, we will continue to refine the definition of well-architected.

This framework is intended for those in technology roles, such as chief technology officers (CTOs), architects, developers, and operations team members. It describes AWS best practices and strategies to use when designing and operating a cloud workload, and provides links to further implementation details and architectural patterns. For more information, see the [AWS Well-Architected homepage](#).

AWS also provides a service for reviewing your workloads at no charge. The [AWS Well-Architected Tool](#) (AWS WA Tool) is a service in the cloud that provides a consistent process for you to review and measure your architecture using the AWS Well-Architected Framework. The AWS WA Tool provides recommendations for making your workloads more reliable, secure, efficient, and cost-effective.

To help you apply best practices, we have created [AWS Well-Architected Labs](#), which provides you with a repository of code and documentation to give you hands-on experience implementing best practices. We also have teamed up with select AWS Partner Network (APN) Partners, who are members of the [AWS Well-Architected Partner program](#). These AWS Partners have deep AWS knowledge, and can help you review and improve your workloads.

### Definitions

Every day, experts at AWS assist customers in architecting systems to take advantage of best practices in the cloud. We work with you on making architectural trade-offs as your designs evolve. As you deploy

these systems into live environments, we learn how well these systems perform and the consequences of those trade-offs.

Based on what we have learned, we have created the AWS Well-Architected Framework, which provides a consistent set of best practices for customers and partners to evaluate architectures, and provides a set of questions you can use to evaluate how well an architecture is aligned to AWS best practices.

The AWS Well-Architected Framework is based on six pillars — operational excellence, security, reliability, performance efficiency, cost optimization, and sustainability.

Table 1. The pillars of the AWS Well-Architected Framework

Name	Description
------	-------------



# AWS Well-Architected Framework

## cloud architectural QUESTIONS



AWS Well-Architected Framework 138 / 440 | - 100% + ⌂ ⌃ ⌁

- Sustainability
- The review process
- Conclusion
- Contributors
- Further reading
- Document revisions
- Appendix: Questions and best practices
  - Operational excellence
  - Security
    - Security foundations
      - SEC 1 How do you securely operate your workload?
    - Identity and access management
      - SEC 2 How do you manage authentication for people and machines?
      - SEC 3 How do you manage permissions for people and machines?
    - Detection
    - Infrastructure protection
    - Data protection
    - Incident response
    - Reliability
    - Performance efficiency
    - Cost optimization
    - Sustainability
  - Notices
  - AWS glossary
- Security Best Practices the Well-Architected Way

133

AWS Well-Architected Framework  
Identity and access management

### Identity and access management

**Questions**

- [SEC 2 How do you manage authentication for people and machines? \(p. 134\)](#)
- [SEC 3 How do you manage permissions for people and machines? \(p. 140\)](#)

#### SEC 2 How do you manage authentication for people and machines?

There are two types of identities you need to manage when approaching operating secure AWS workloads. Understanding the type of identity you need to manage and grant access helps you ensure the right identities have access to the right resources under the right conditions.

**Human Identities:** Your administrators, developers, operators, and end users require an identity to access your AWS environments and applications. These are members of your organization, or external users with whom you collaborate, and who interact with your AWS resources via a web browser, client application, or interactive command line tools.

**Machine Identities:** Your service applications, operational tools, and workloads require an identity to make requests to AWS services for example, to read data. These identities include machines running in your AWS environment such as Amazon EC2 instances or AWS Lambda functions. You may also manage machine identities for external parties who need access. Additionally, you may also have machines outside of AWS that need access to your AWS environment.

**Best practices**

- [SEC02-BP01 Use strong sign-in mechanisms \(p. 134\)](#)
- [SEC02-BP02 Use temporary credentials \(p. 135\)](#)
- [SEC02-BP03 Store and use secrets securely \(p. 137\)](#)
- [SEC02-BP04 Rely on a centralized identity provider \(p. 137\)](#)
- [SEC02-BP05 Audit and rotate credentials periodically \(p. 138\)](#)
- [SEC02-BP06 Leverage user groups and attributes \(p. 139\)](#)

#### SEC02-BP01 Use strong sign-in mechanisms

Enforce minimum password length, and educate your users to avoid common or reused passwords. Enforce multi-factor authentication (MFA) with software or hardware mechanisms to provide an additional layer of verification. For example, when using IAM Identity Center as the identity source, configure the "context-aware" or "always-on" setting for MFA, and allow users to enroll their own MFA devices to accelerate adoption. When using an external identity provider (IdP), configure your IdP for MFA.



# AWS Well-Architected Framework



## Pillars



**Pillar 1**

**Pillar 2**

**Pillar 3**

**Pillar 4**

**Pillar 5**

**Pillar n...**



**Design Principles**



**Key Topics**

• **Design Patterns**

• **Anti-Patterns**



**Implementation Guide**



**Risks**



**Benefits**



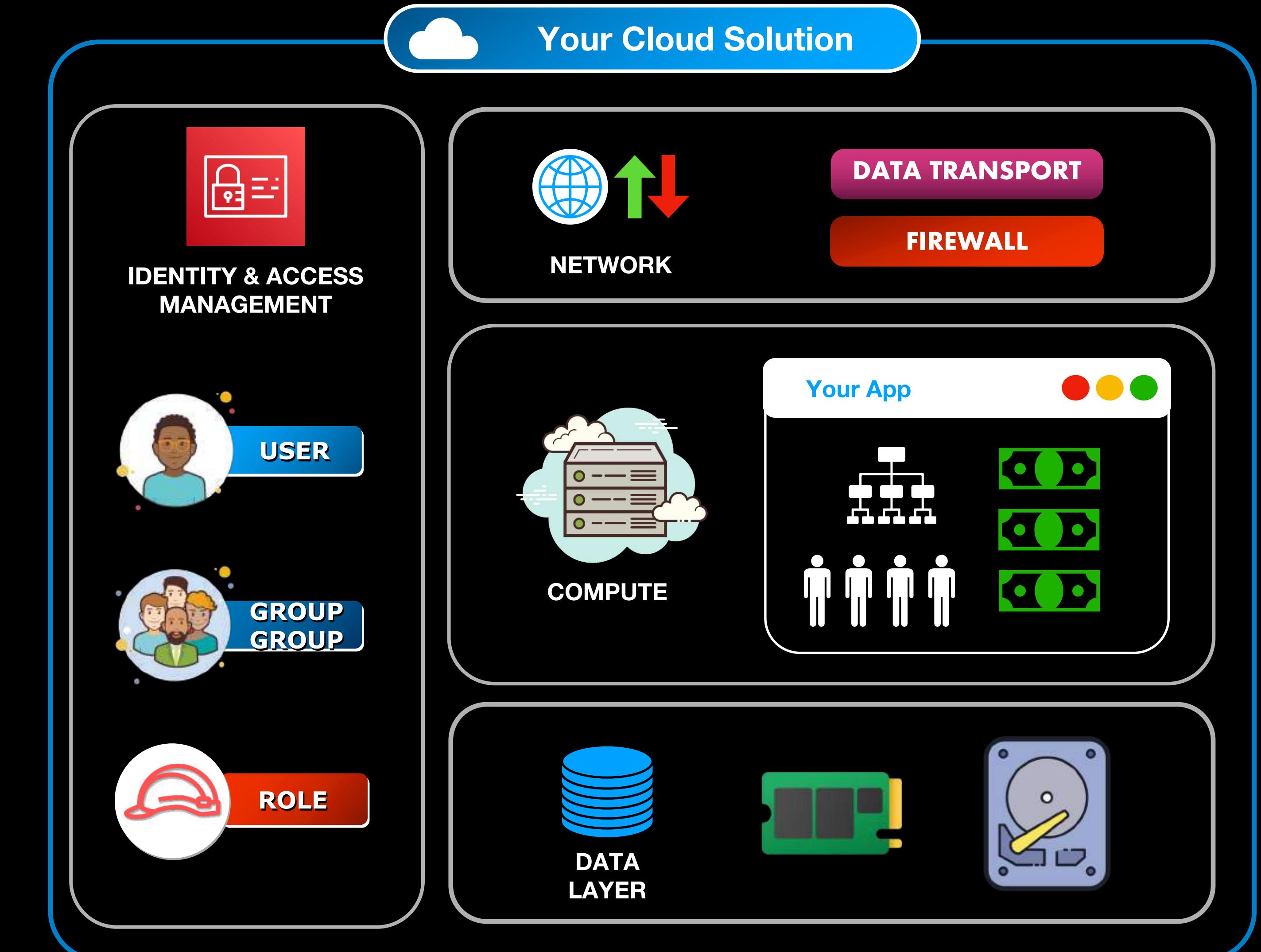
**Best Practices**



# HOW DOES IT WORK?



## Security Pillar





# HOW DOES IT WORK?

## III Security Pillar



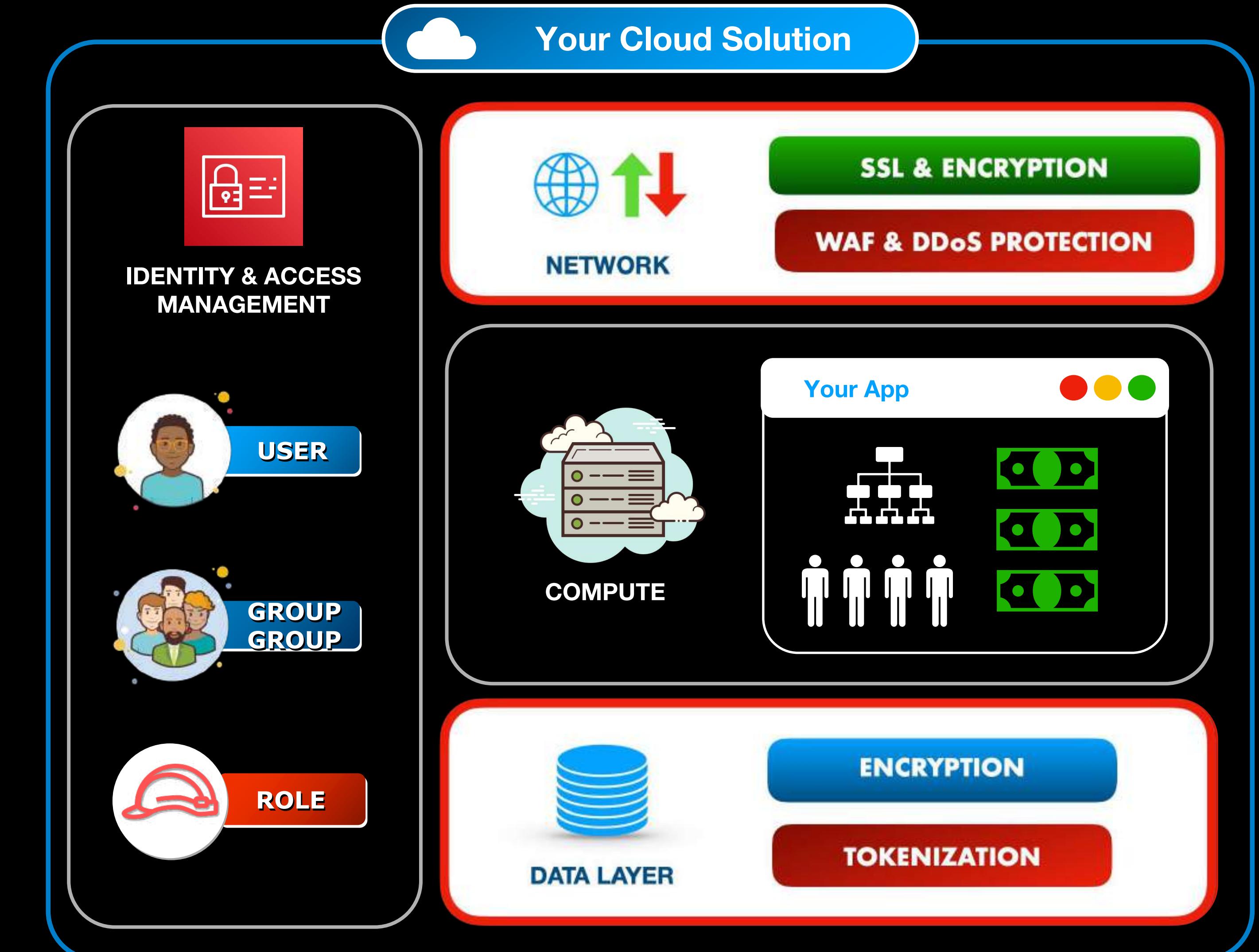
How do you protect your **data at rest**?



How do you protect your **data in transit**?



How do you manage **identities** for people and machines?





The screenshot shows a page from the AWS Well-Architected Framework document. The left sidebar contains a navigation menu with sections like Sustainability, The review process, Conclusion, Contributors, Further reading, Document revisions, Appendix: Questions and best practices (which is expanded), Operational excellence, Security (expanded), Security foundations (expanded), SEC 1 How do you securely operate your workload? (expanded), Identity and access management (expanded), SEC 2 How do you manage authentication for people and machines? (expanded), SEC 3 How do you manage permissions for people and machines? (expanded), Detection, Infrastructure protection, Data protection, Incident response, Reliability, Performance efficiency, Cost optimization, Sustainability, Notices, and AWS glossary. The main content area is titled "Identity and access management". It includes a "Questions" section with links to SEC 2 and SEC 3. Below that is a "SEC 2 How do you manage authentication for people and machines?" section with text about managing identities for AWS workloads. It then details Human Identities and Machine Identities. The "Best practices" section lists SEC02-BP01 through SEC02-BP06. Finally, there is a detailed description of SEC02-BP01.

138 / 440 | 100% | [Download](#) | [Print](#) | [Share](#)

• [Security Best Practices the Well-Architected Way](#)

133

AWS Well-Architected Framework  
Identity and access management

## Identity and access management

**Questions**

- [SEC 2 How do you manage authentication for people and machines? \(p. 134\)](#)
- [SEC 3 How do you manage permissions for people and machines? \(p. 140\)](#)

### SEC 2 How do you manage authentication for people and machines?

There are two types of identities you need to manage when approaching operating secure AWS workloads. Understanding the type of identity you need to manage and grant access helps you ensure the right identities have access to the right resources under the right conditions.

**Human Identities:** Your administrators, developers, operators, and end users require an identity to access your AWS environments and applications. These are members of your organization, or external users with whom you collaborate, and who interact with your AWS resources via a web browser, client application, or interactive command line tools.

**Machine Identities:** Your service applications, operational tools, and workloads require an identity to make requests to AWS services for example, to read data. These identities include machines running in your AWS environment such as Amazon EC2 instances or AWS Lambda functions. You may also manage machine identities for external parties who need access. Additionally, you may also have machines outside of AWS that need access to your AWS environment.

**Best practices**

- [SEC02-BP01 Use strong sign-in mechanisms \(p. 134\)](#)
- [SEC02-BP02 Use temporary credentials \(p. 135\)](#)
- [SEC02-BP03 Store and use secrets securely \(p. 137\)](#)
- [SEC02-BP04 Rely on a centralized identity provider \(p. 137\)](#)
- [SEC02-BP05 Audit and rotate credentials periodically \(p. 138\)](#)
- [SEC02-BP06 Leverage user groups and attributes \(p. 139\)](#)

### SEC02-BP01 Use strong sign-in mechanisms

Enforce minimum password length, and educate your users to avoid common or reused passwords. Enforce multi-factor authentication (MFA) with software or hardware mechanisms to provide an additional layer of verification. For example, when using IAM Identity Center as the identity source, configure the “context-aware” or “always-on” setting for MFA, and allow users to enroll their own MFA devices to accelerate adoption. When using an external identity provider (IdP), configure your IdP for MFA.

# DO YOU REALLY NEED TO FOLLOW ALL THE GUIDELINES OF THE AWS WELL-ARCHITECTED FRAMEWORK?



# TRADE-OFFS

REQUIREMENTS					AVERAGE COST
ENVIRONMENT	RELIABILITY	DATA SECURITY	SCALABILITY	COMPLIANCE	
PROD	HIGH	AT REST	MUST	HIPAA	
PRE PROD	MID	IN TRANSIT	OPTIONAL	GDPR	
DEV	LOW	NONE		PCI-DSS	



# TRADE-OFFS

ENVIRONMENT

DEV

TRADE-OFF

RELIABILITY

over

LOW COST





# TRADE-OFFS

## REQUIREMENTS

ENVIRONMENT

RELIABILITY

DATA SECURITY

SCALABILITY

COMPLIANCE

PROD

HIGH

AT REST

MUST

HIPAA

PRE PROD

MID

IN TRANSIT

OPTIONAL

GDPR

TEST

LOW

NONE

PCI-DSS



LOW COST



# TRADE-OFFS

MISSION-CRITICAL  
APPLICATIONS

## REQUIREMENTS



HIGH COST

ENVIRONMENT

PROD

PRE PROD

DEV

RELIABILITY

HIGH

MID

LOW

MORE REDUNDANT  
RESOURCES

DATA SECURITY

AT REST

IN TRANSIT

NONE

SCALABILITY

MUST

OPTIONAL

MORE COMPUTE &  
STORAGE  
RESOURCES

COMPLIANCE

HIPAA

GDPR

PCI-DSS



# TRADE-OFFS

## REQUIREMENTS

ENVIRONMENT	RELIABILITY	DATA SECURITY	SCALABILITY	COMPLIANCE
PROD	HIGH	AT REST	MUST	HIPAA
PRE PROD	MID	IN TRANSIT	OPTIONAL	GDPR
DEV	LOW	NONE		PCI-DSS

**IN PRODUCTION, SECURITY IS  
NOT USUALLY  
TRADED-OFF WITH ANY OTHER  
FACTORS**



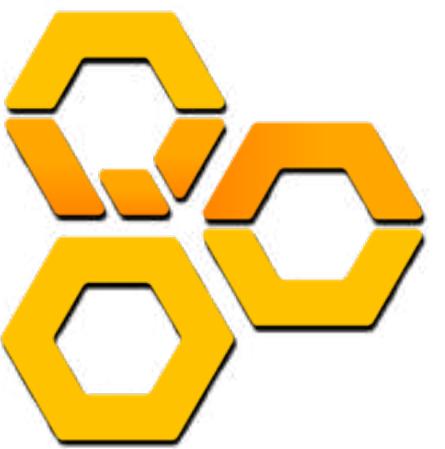
## Foreign Laws & Security Requirements

- **Covers Data Sovereignty requirements**
- **Abide by the Regional Rules that needs to be strictly followed**
- **Quickly establish a digital presence in other countries while being compliant with its data protection and privacy laws**
- **Example: General Data Protection Regulation (GDPR)**
- **Each country has its own data privacy law with a unique data residency and data sovereignty requirements**

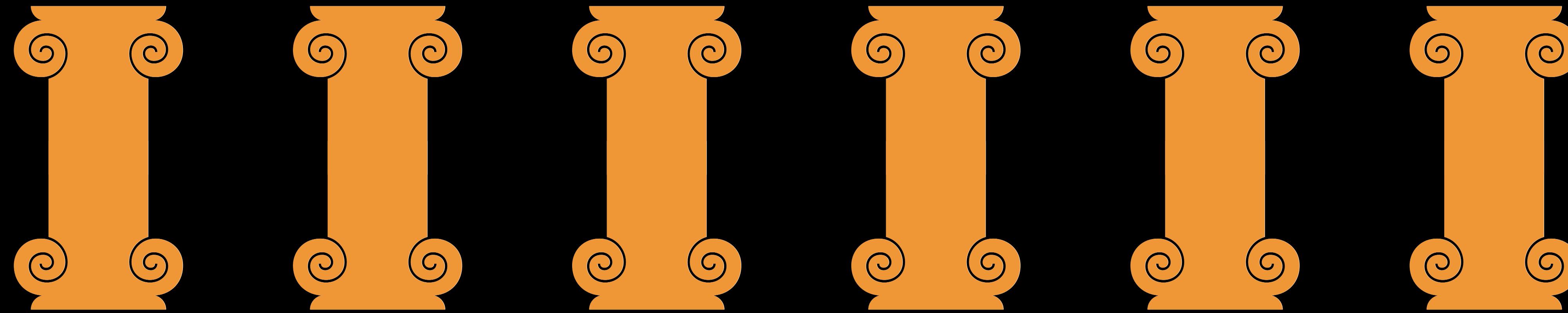


# The Pillars of the AWS Well-Architected Framework

---

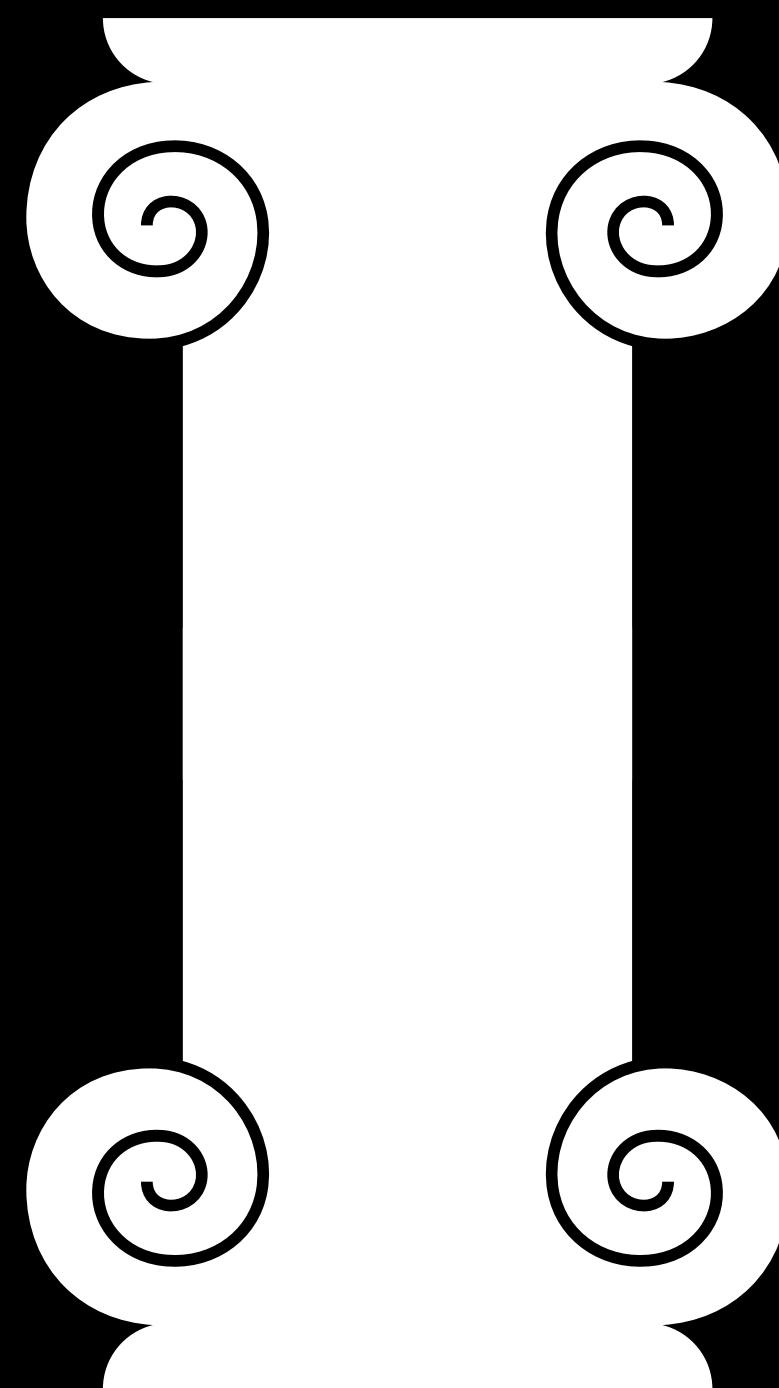


# AWS Well-Architected Framework Pillars





# AWS Well-Architected Framework Pillars



- **OPERATIONAL EXCELLENCE**
- **SECURITY**
- **RELIABILITY**
- **PERFORMANCE EFFICIENCY**
- **COST OPTIMIZATION**
- **SUSTAINABILITY**



- Revolves around **how you run your operations to deliver business value**
- Allows you to verify that your AWS workloads are operating excellently or poorly
- Provides the ability to:
  - Effectively run workloads in AWS
  - Gain helpful insight into your cloud operations
  - Continuously improve your supporting processes & procedures
- Example of an **Operationally Excellent AWS solution:**
  - An AWS workload with **loosely-coupled components which can be updated on a regular basis and where the changes can be made in small, reversible increments.**



- Can be achieved by **establishing protocols in place to continuously improve the supporting processes of your cloud operations**
- **Supporting Processes:**
  - **Continuous Improvement**
  - **Knowledge Management**
  - **Post-incident Analysis**
  - **Feedback Loops**
  - **Other protocols that support your primary processes**
- **Includes the concepts of Risk Mitigation, Disaster Recovery Exercises, Game Days or Team Drills to test your Disaster Recovery Action Plan**



## DESIGN PRINCIPLES

- **Perform Operations as Code**
- **Make Frequent, Small, Reversible Changes**
- **Refine Operations Procedures Frequently**
- **Anticipate Failure**
- **Learn from All Operational Failures**

## BEST PRACTICE AREAS

- **Organization**
- **Prepare**
- **Operate**
- **Evolve**



- **Covers the overall security of your AWS workloads**
- **Not usually traded off over other aspect of your system**
- **Checks the use of various security-related AWS services to protect the data, systems, and assets of your cloud solutions**
- **Includes the concept of Traceability (monitoring & tracking the changes made to your environment and resources)**
- **Root Cause Analysis and Remediation Automation of production incidents**
- **Aims to improve your overall Security Posture**



- **Examples of Secure AWS solutions:**
  - **Enabling Traceability via AWS Config to record, audit, and evaluate changes to AWS resources in your production environment.**
  - **Implementing data encryption, tokenization, SSL, and firewalls to protect your sensitive data in transit and data at rest**
  - **Granting the least privilege to your staff with the minimum permissions required to perform a task**



## DESIGN PRINCIPLES

- **Implement a Strong Identity Foundation**
- **Enable Traceability**
- **Apply Security at All Layers**
- **Automate Security Best Practices**
- **Protect Data in Transit and at Rest**
- **Keep People Away from Data**
- **Prepare for Security Events**

## BEST PRACTICE AREAS

- **Foundations for Security**
- **Identity and Access Management**
- **Detection**
- **Infrastructure Protection**
- **Data Protection**
- **Incident Response**



- Focused on the ability of your systems to recover and work consistently & accurately
- Ensures your applications remain reliable even if there are traffic surges, unexpected system changes, or natural disasters
- Includes the ability to operate and test your AWS workloads throughout its entire lifecycle
- Accentuates the concept of Recovery to your cloud solutions in AWS to meet your strict Recovery Time Objective (RTO) and Recovery Point Objective (RPO) requirements
- Verifies that your application has the ability to recover from service disruptions, natural disasters, application failures, and other type of outages
- Checks if your cloud architecture can dynamically acquire computing resources to meet the changing demand of your application



- Examples of Reliable AWS solutions:
  - A system that is able to recover from infrastructure or service disruptions by using redundant AWS resources such as an Amazon RDS database in Multi-AZ Deployments configuration, Amazon Aurora Global Database or an application deployed in multiple Availability Zones or AWS Regions.
  - Implementing Amazon EC2 Auto Scaling on multiple Availability Zones behind an Application Load Balancer to automatically recover from outages and dynamically acquire computing resources to avoid system degradation.
  - Using Cross-Region Replication for databases, S3 buckets, and other resources to increase the ability of your systems to recover.



## DESIGN PRINCIPLES

- **Automatically Recover from Failure**
- **Test Recovery Procedures**
- **Scale Horizontally to Increase Aggregate Workload Availability**
- **Stop Guessing Capacity**
- **Manage Change through Automation**

## BEST PRACTICE AREAS

- **Foundations for Reliability**
- **Workload Architecture**
- **Change Management**
- **Failure Management**



- **Covers the ability to improve the performance factors efficiently to meet your system requirements**
- **Focuses on achieving and maintaining a high level of efficiency even as your customer demand changes**
- **Adopting new technologies (e.g. Serverless, Containerization)**
- **Re-factoring/re-architecting the existing design of your system to improve application performance**
- **Example AWS solution that demonstrates Performance Efficiency:**
  - **Re-architecting an on-premises monolithic system to become a Serverless Application to efficiently lessen the operating cost, enhance scalability and further improve other performance factors.**



## DESIGN PRINCIPLES

- **Democratize Advanced Technologies**
- **Go Global in Minutes**
- **Use Serverless Architectures**
- **Experiment More Often**
- **Consider Mechanical Sympathy**

## BEST PRACTICE AREAS

- **Selection**
- **Review**
- **Monitoring**
- **Trade-offs**



- Focuses on **the ability to run your systems and deliver business value at the lowest price point possible**
- A **continual process of improving your AWS workloads while minimizing costs to achieve the outcomes expected of the business in a cost-effective manner**
- Aims to **increase revenue and maximize return on investment (ROI)**
- Example of a **Cost-Optimized AWS Solution:**
  - **Adopting a Consumption Model via Pay-as-you-go pricing where you only pay for the resources that you actually consume or by using AWS Serverless services.**



- Removes the reliance on elaborate forecasting to determine what would be the expected usage of your compute resources
- Less dependency on extremely inaccurate forecasting and guesswork in terms of capital expenditures (CAPEX) or operating expenses (OPEX)
- Trade Fixed Expense with Variable Expense by choosing Pay-As-You-Go Pricing and adopting a cost-effective Serverless Architecture
- Have the ability to dynamically increase or decrease resource usage to meet the ever-changing requirements of the business



## DESIGN PRINCIPLES

- **Implement Cloud Financial Management**
- **Adopt a Consumption Model**
- **Measure Overall Efficiency**
- **Stop Spending Money on Undifferentiated Heavy Lifting**
- **Analyze and Attribute Expenditure**

## BEST PRACTICE AREAS

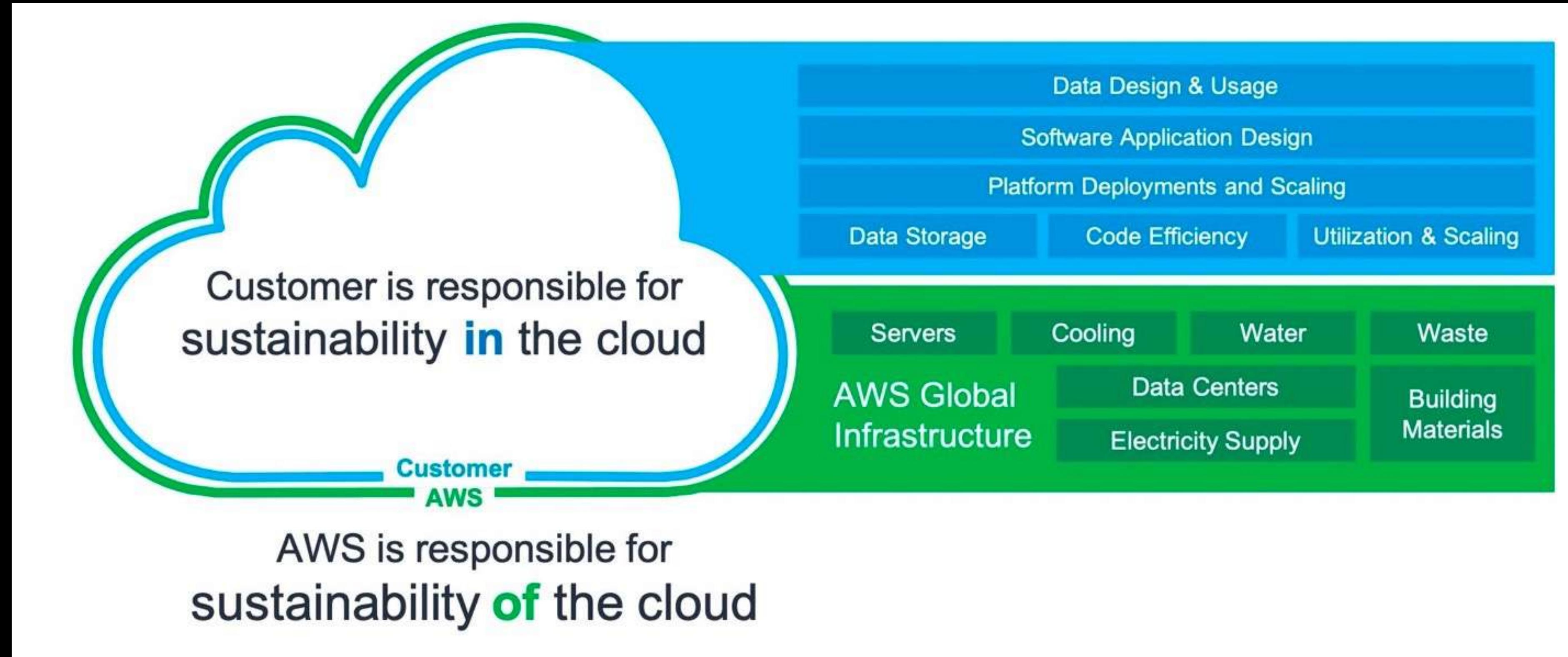
- **Practice Cloud Financial Management**
- **Expenditure & Usage Awareness**
- **Cost-effective Resources**
- **Manage Demand & Supplying Resources**
- **Optimize over Time**



- All about sustainable development, which addresses the long-term environmental, economic, and societal impact of your business operations as you use the AWS Cloud
- A Sustainable Development is:
  - "...a type of development that meets the needs of the present without compromising the ability of future generations to meet their own needs"
- Aims to lessen negative environmental impacts such as carbon emissions, unrecyclable waste, and damage to shared natural resources
- Focuses on Environmental Sustainability which is a shared responsibility between you & AWS



# SUSTAINABILITY PILLAR





## DESIGN PRINCIPLES

- **Understand your Impact**
- **Establish Sustainability Goals**
- **Maximize Utilization**
- **Anticipate and Adopt New, More Efficient Hardware & Software Offerings**
- **Use Managed Services**
- **Reduce the Downstream Impact of your Cloud Workloads**

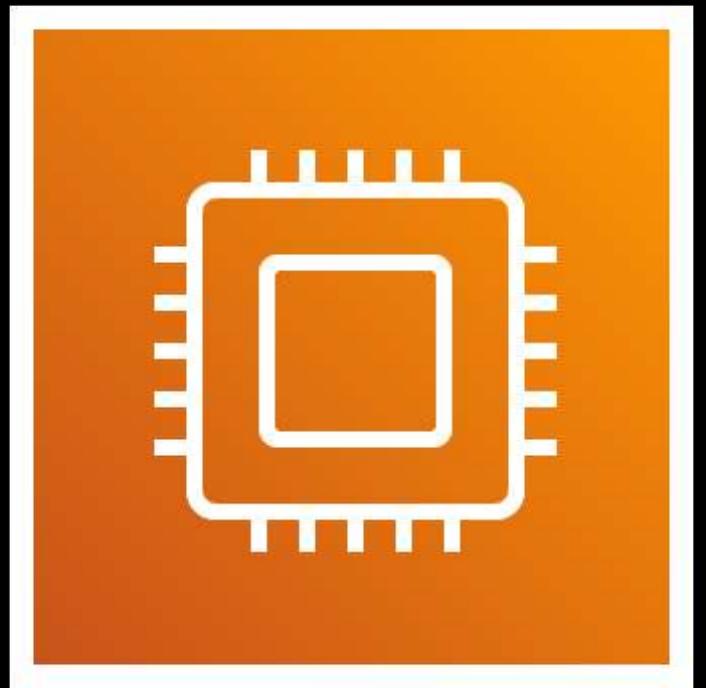
## BEST PRACTICE AREAS

- **Region Selection**
- **User Behavior Patterns**
- **Software & Architecture Patterns**
- **Data Patterns**
- **Hardware Patterns**
- **Development & Deployment Process**



# AWS Services Overview

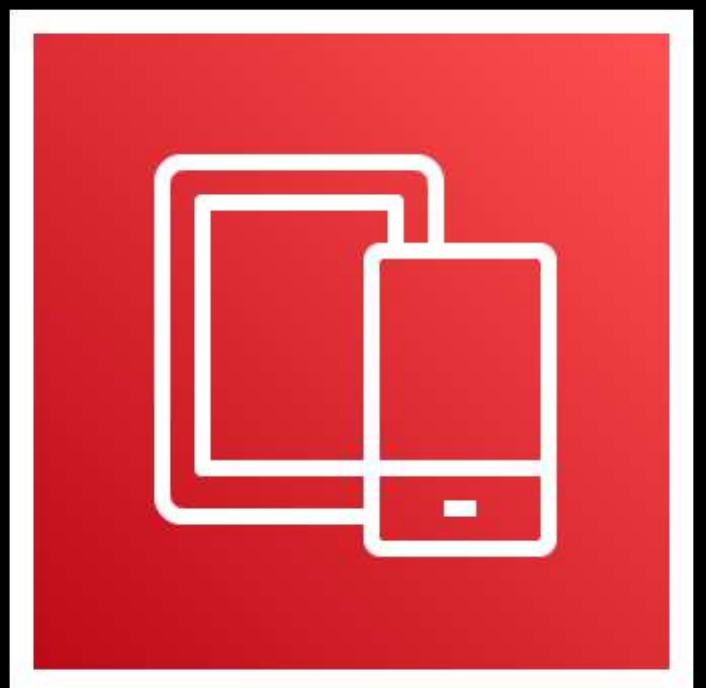
---



**Host  
Web Apps**



**Run Real-Time  
Data Analytics**



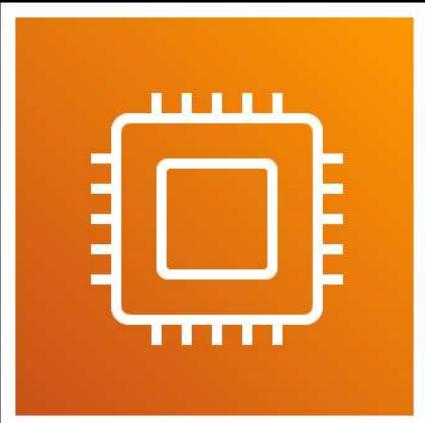
**Develop  
Mobile Apps**



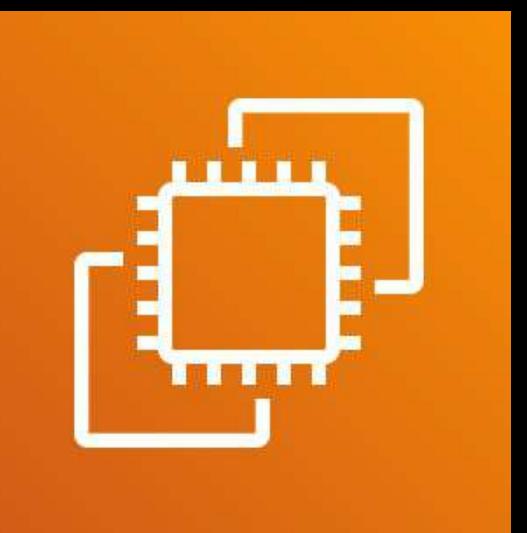
**Store Data  
for Backup**

	Analytics
	Application Integration
	AWS Cost Management
	Blockchain
	Business Applications
	Compute
	Containers
	Customer Engagement
	Database
	Developer Tools
	End User Computing
	Front-End Web & Mobile
	Game Tech
	Internet of Things
	Machine Learning
	Management & Governance
	Media Services
	Migration & Transfer
	Networking & Content Delivery
	Quantum Technologies
	Robotics
	Satellite
	Security, Identity & Compliance
	Serverless

# PER CATEGORY



## COMPUTE SERVICES



Amazon EC2



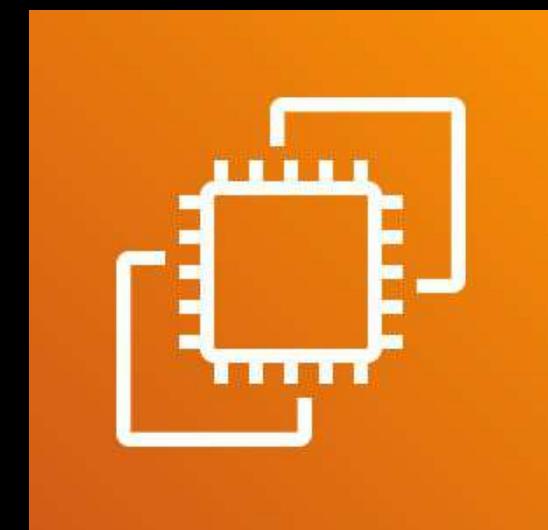
AWS Lambda



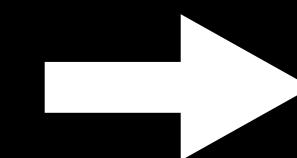
AWS Outposts



Amazon Lightsail



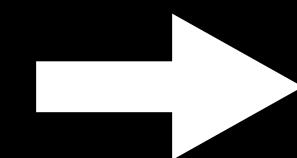
Amazon EC2



**Amazon Elastic Compute Cloud**



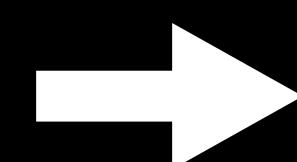
Amazon S3



**Amazon Simple Storage Service**



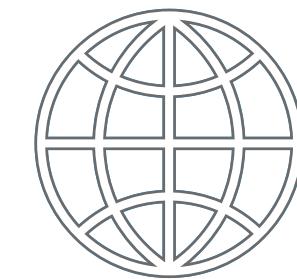
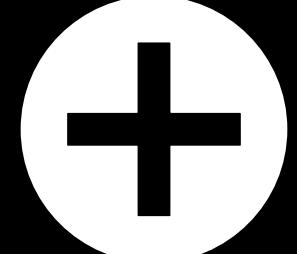
Amazon RDS



**Amazon Relational Database Service**



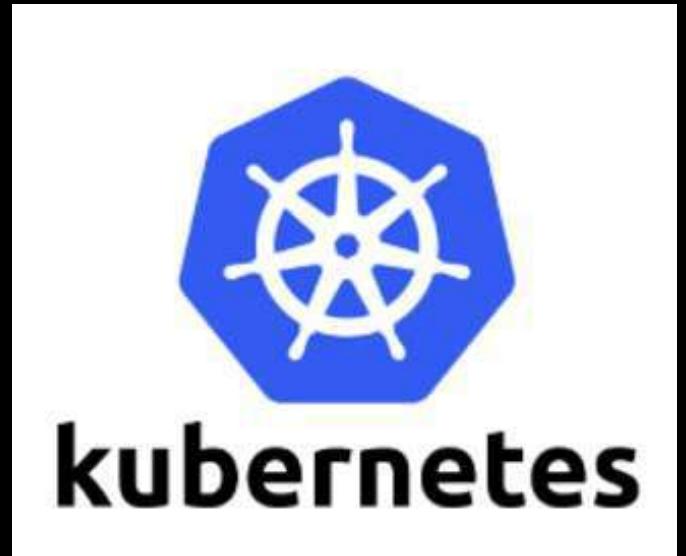
Fully Managed *By:*



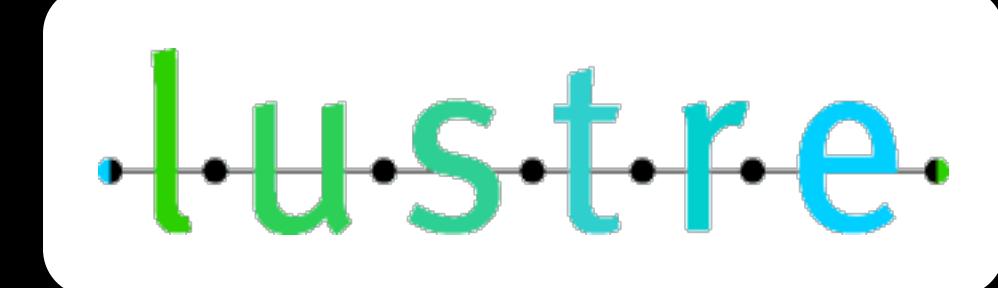
Open Source Technology



Amazon Elastic **Kubernetes** Service (EKS)

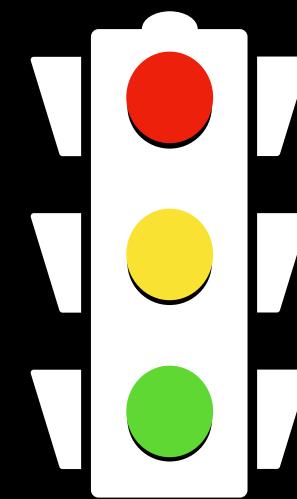


Amazon FSx for **Lustre** (FSx)



Amazon **Elasticsearch** Service





Routes Traffic



Amazon Route 53

What's the  
meaning of  
this  
number?

The number **53** is the TCP and UDP **Port Number**  
used for the Domain Name System (**DNS**) protocol transport

PORT



Featured Services



Analytics



Application Integration



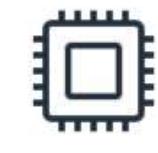
AWS Cost Management



Blockchain



Business Applications



Compute



Containers



Customer Engagement



Database



Developer Tools



End User Computing



Front-End Web & Mobile



Game Tech



Internet of Things



Machine Learning



Management & Governance



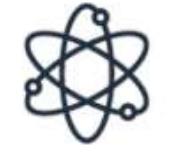
Media Services



Migration & Transfer



Networking & Content Delivery



Quantum Technologies



Robotics



Satellite



Security, Identity & Compliance



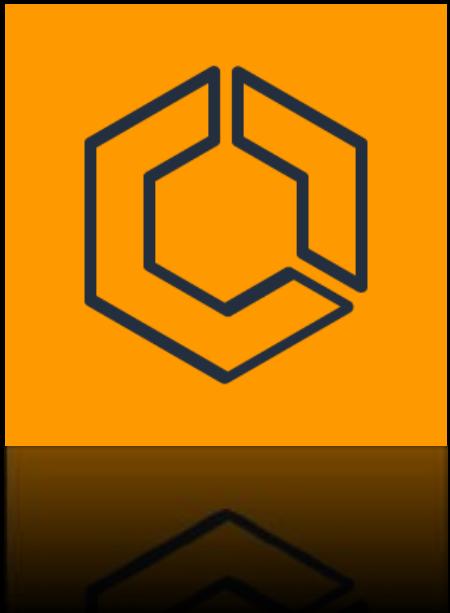
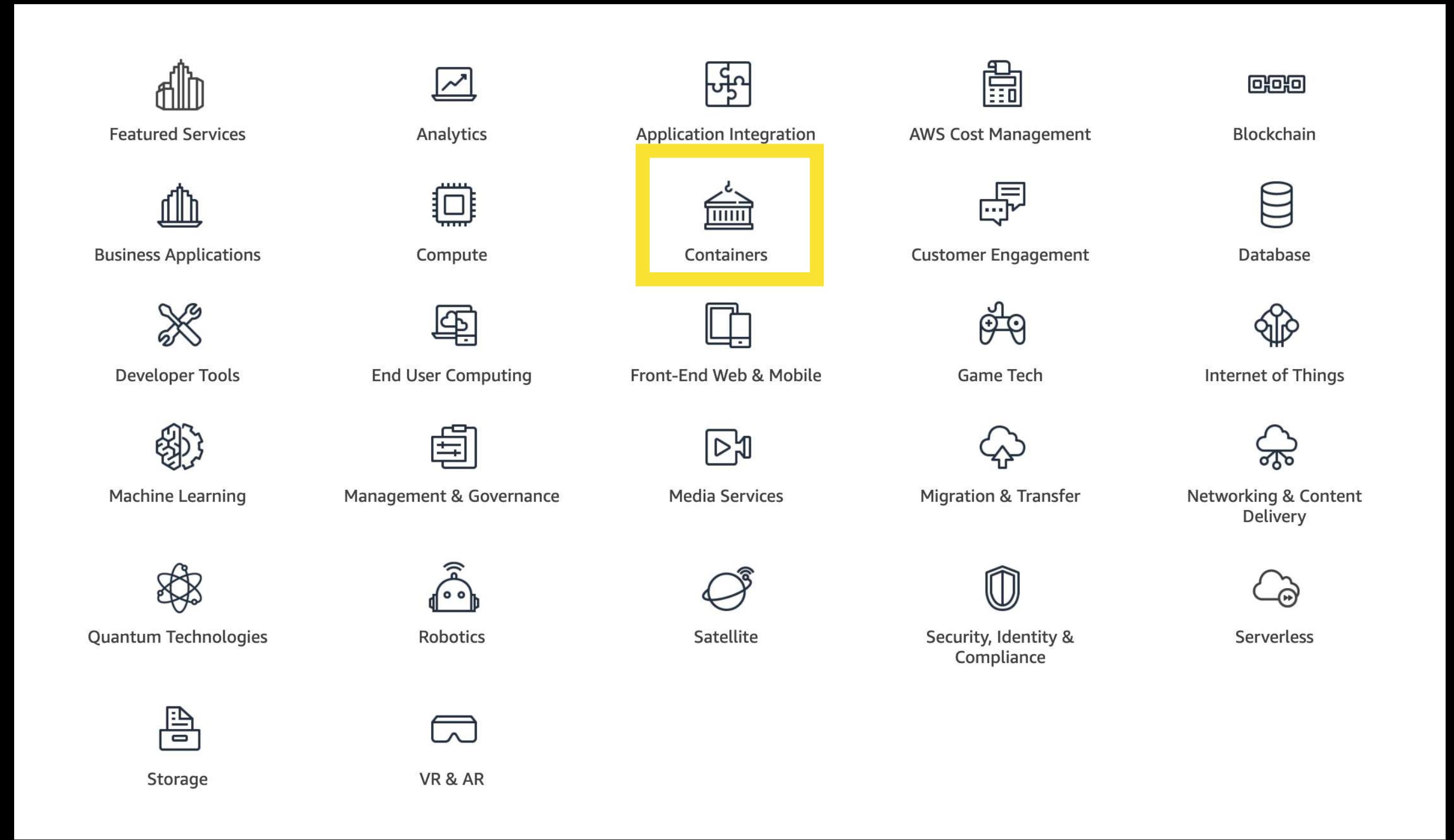
Serverless



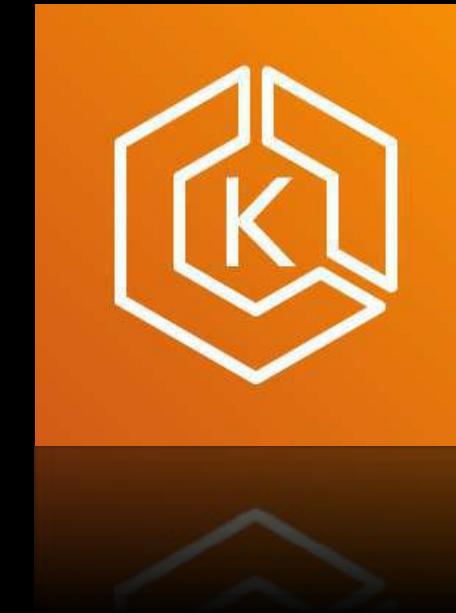
Storage



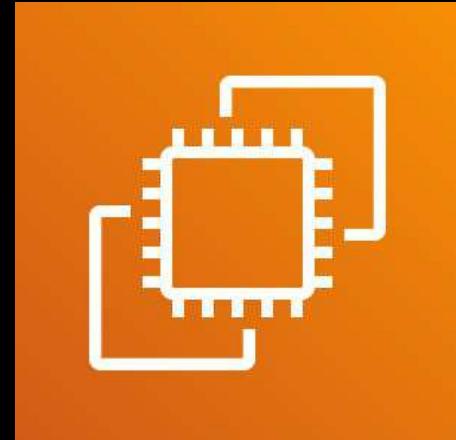
VR & AR



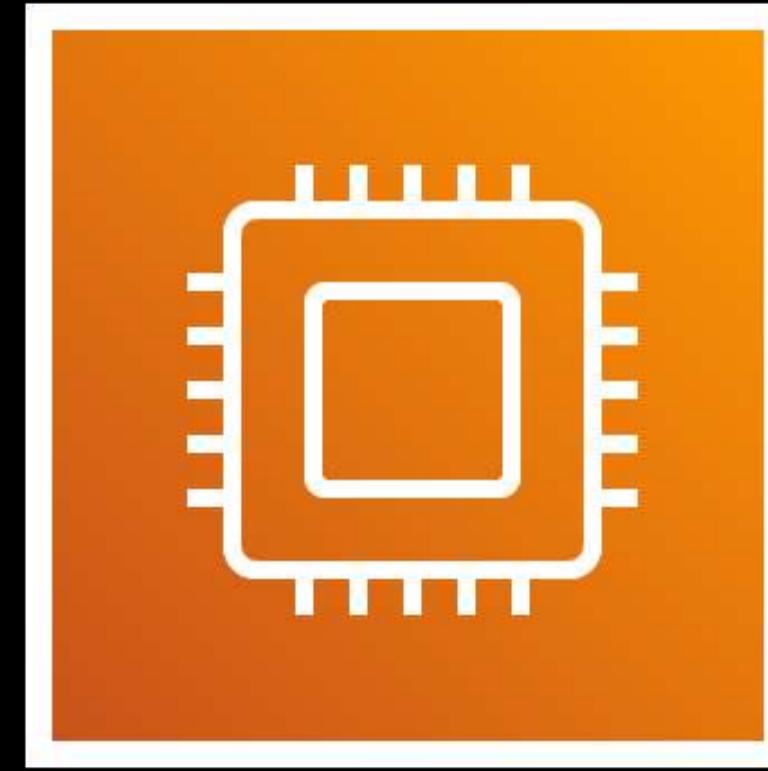
**Amazon Elastic Container Service**



**Amazon Elastic Kubernetes Service**

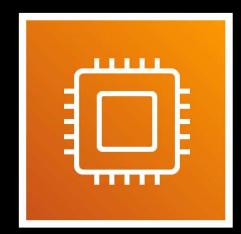


**Amazon EC2**



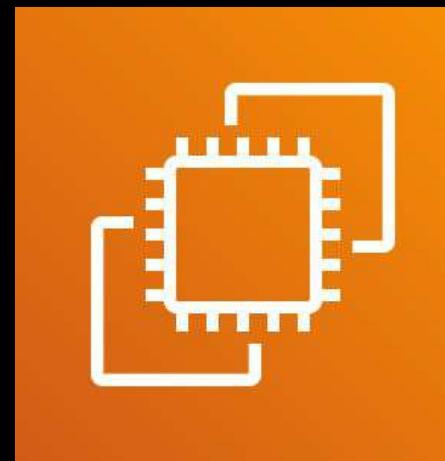
# AWS Compute Services Overview

---



# AWS Compute Services

## Virtual Machines



Amazon EC2

## Serverless



AWS Lambda

## Orchestration



AWS Elastic Beanstalk

## Container



Amazon EKS



Amazon LightSail



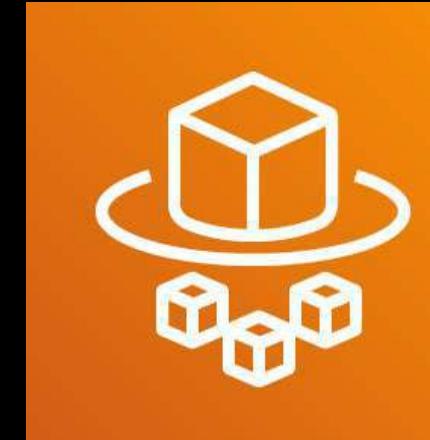
AWS Batch



Amazon ECS



AWS Outposts



AWS Fargate

## Virtual Machines

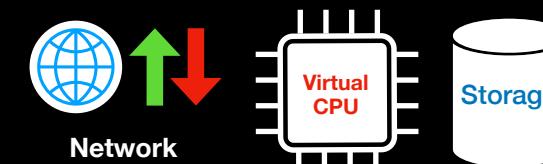




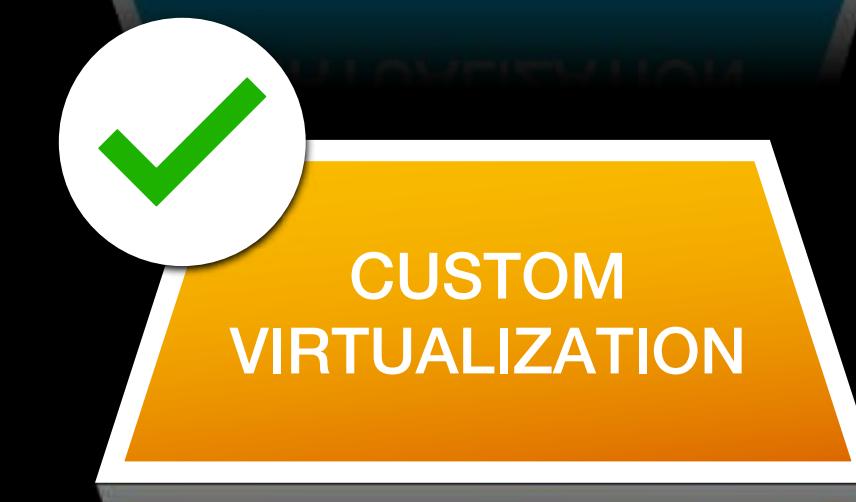
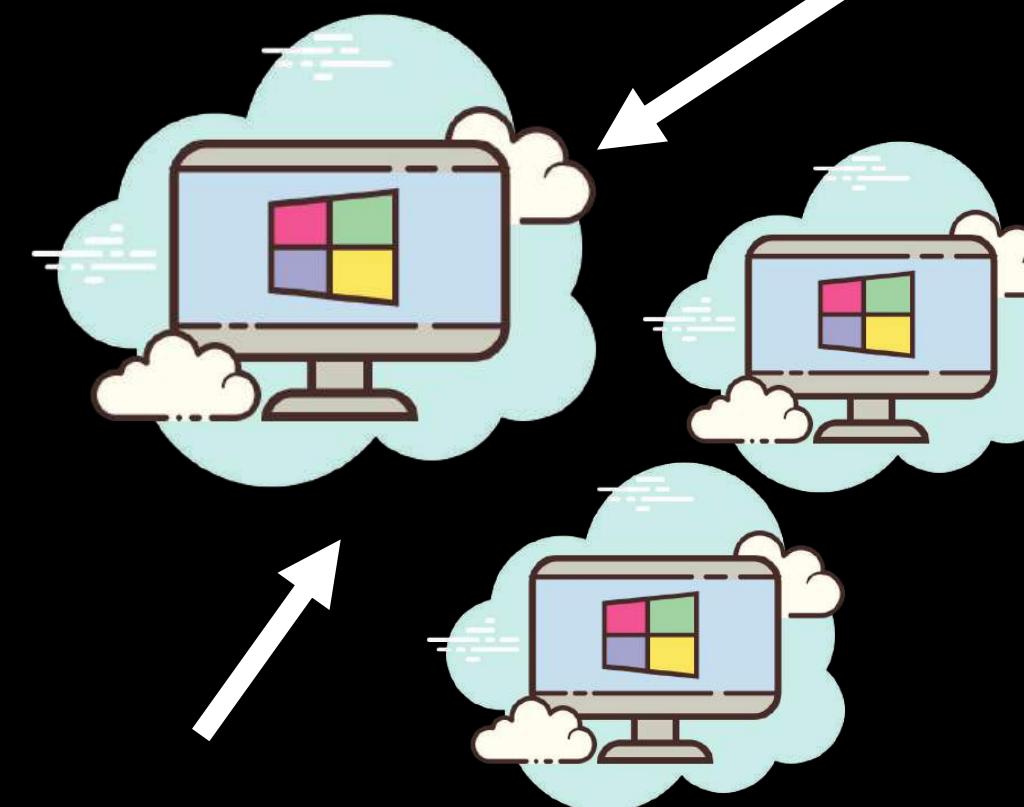
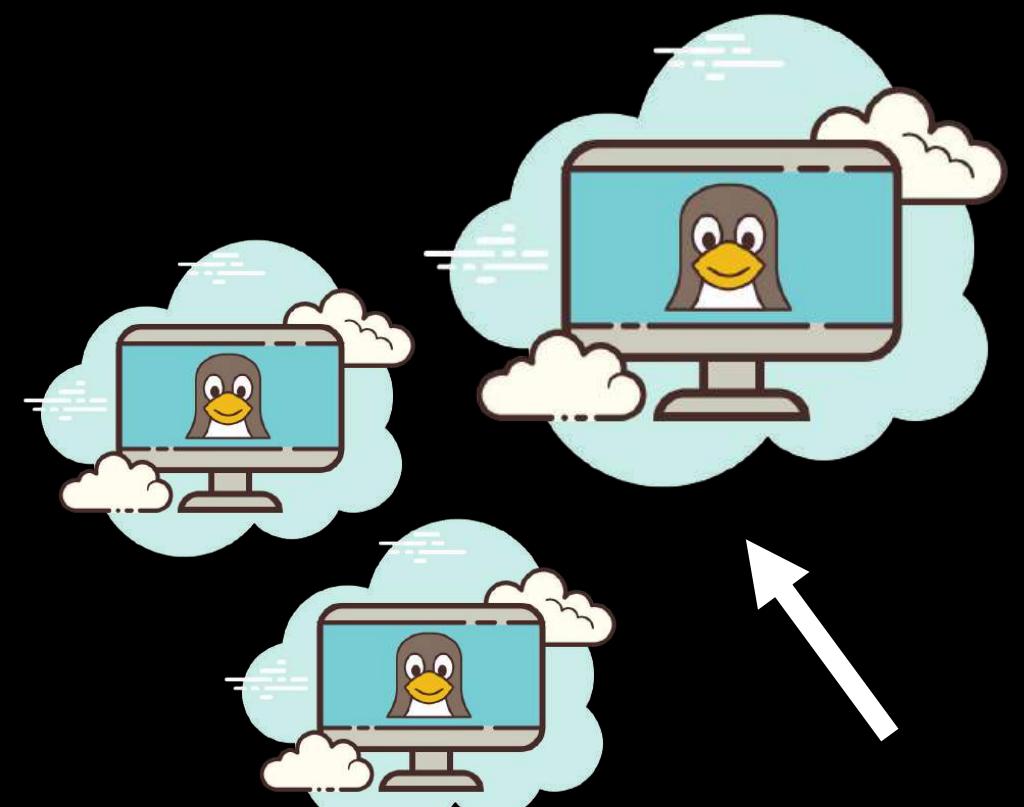
Used by **MULTIPLE** Tenants / Customers



Used by a **SINGLE** Customer



**Instance**



Also called a  
**Virtual Machine Monitor**  
or a  
**Hypervisor**



## Serverless

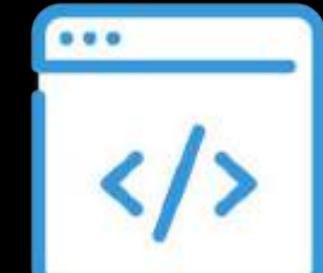
Fully Managed *By:* 



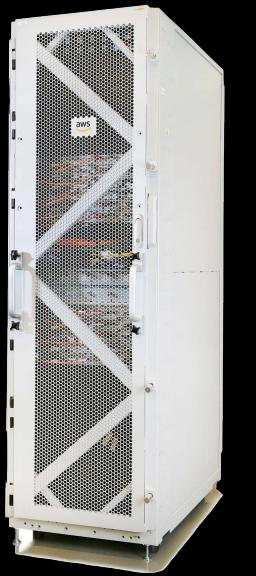
NO DIRECT  
Server access  
via:

SSH or RDP

Unlike  
Amazon EC2

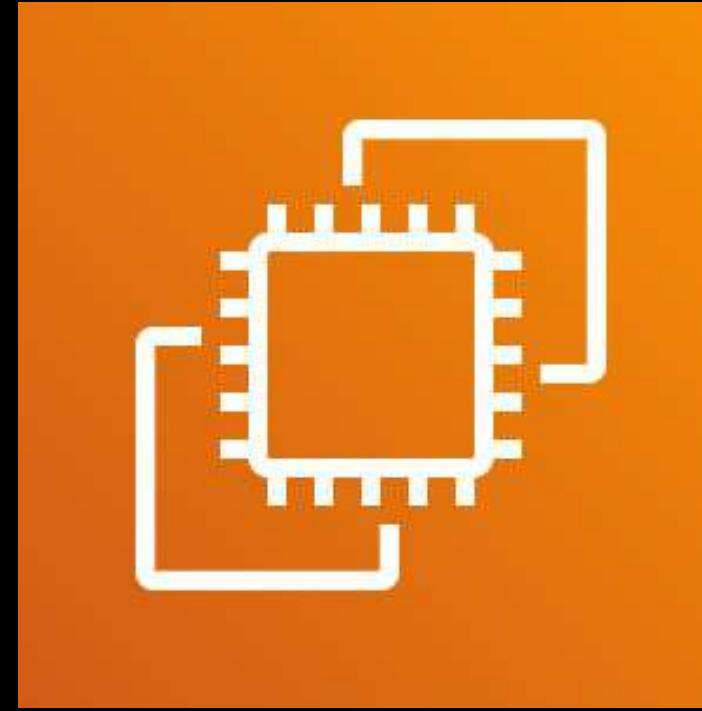


## Hybrid



On-premises data center

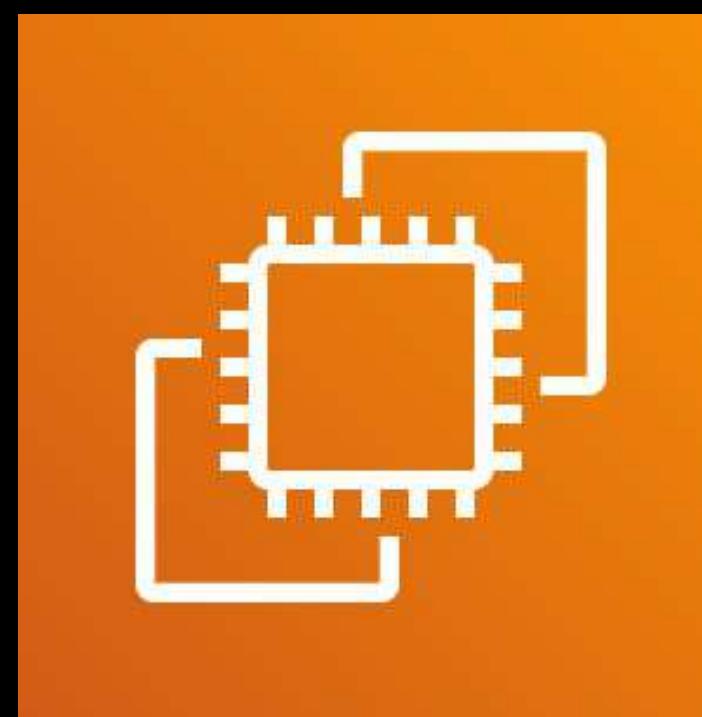




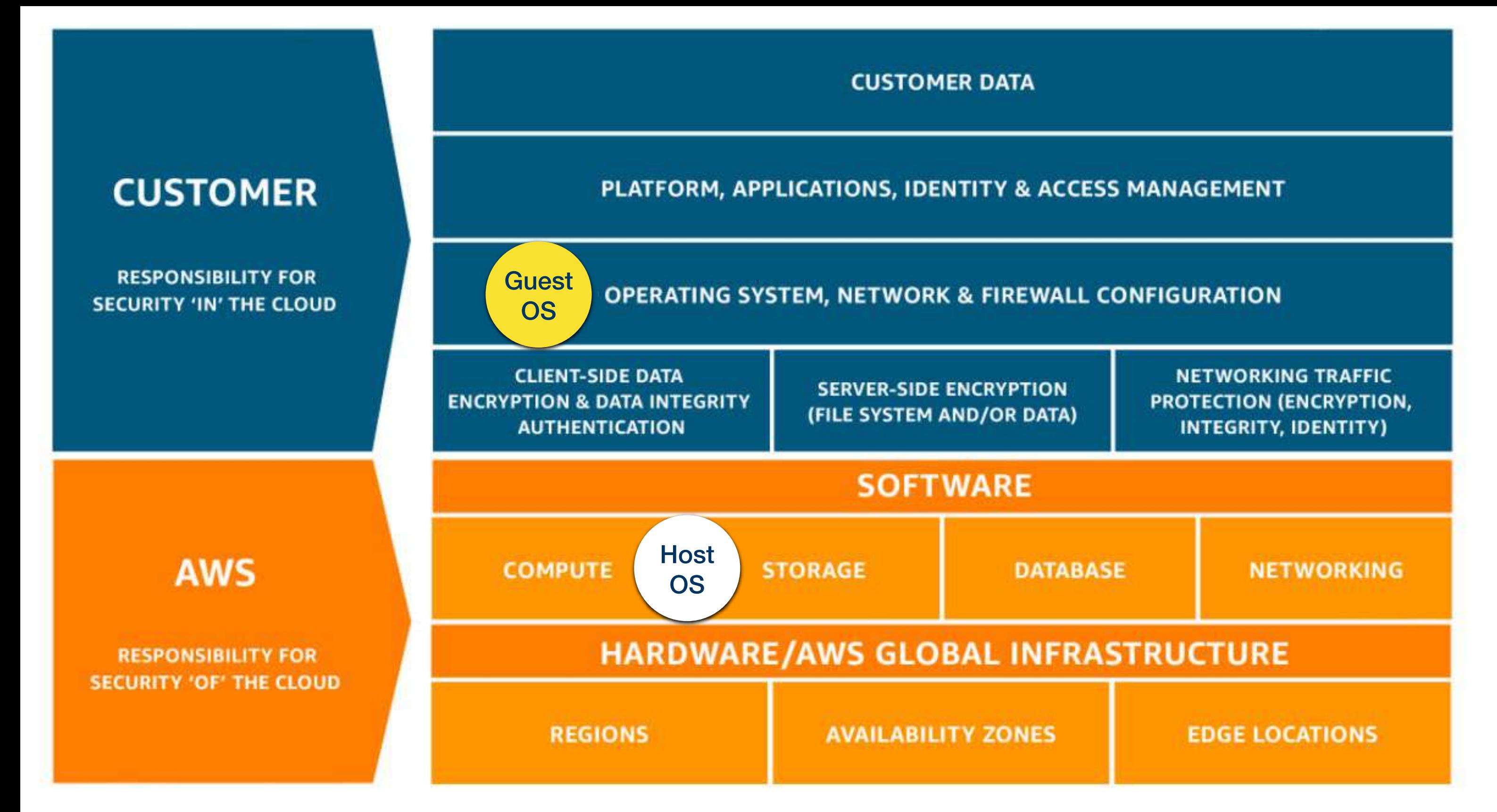
Amazon EC2

- A **computing service** that runs virtual servers in AWS
- Allows you to launch Windows, Linux or even MacOS virtual machines
- A type of an **Infrastructure as a Service (IaaS)**
- A basic building block for your cloud architecture
- Used by other AWS services as an underlying compute service

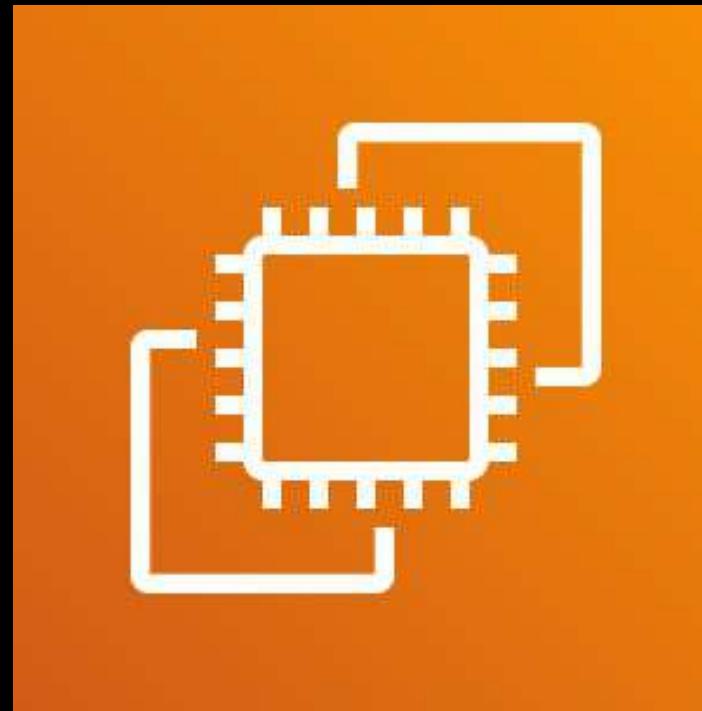
## Shared Responsibility Model



Amazon EC2

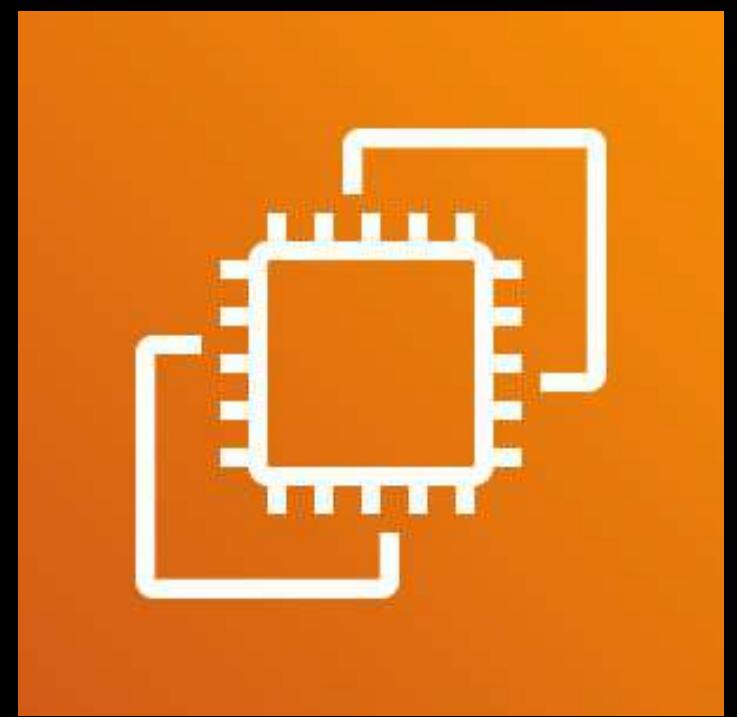


# Elastic Compute Cloud



Amazon EC2

- **Flexible**
- **Customizable**
- **Scalable**



Amazon EC2

Elastic Compute Cloud

EC2



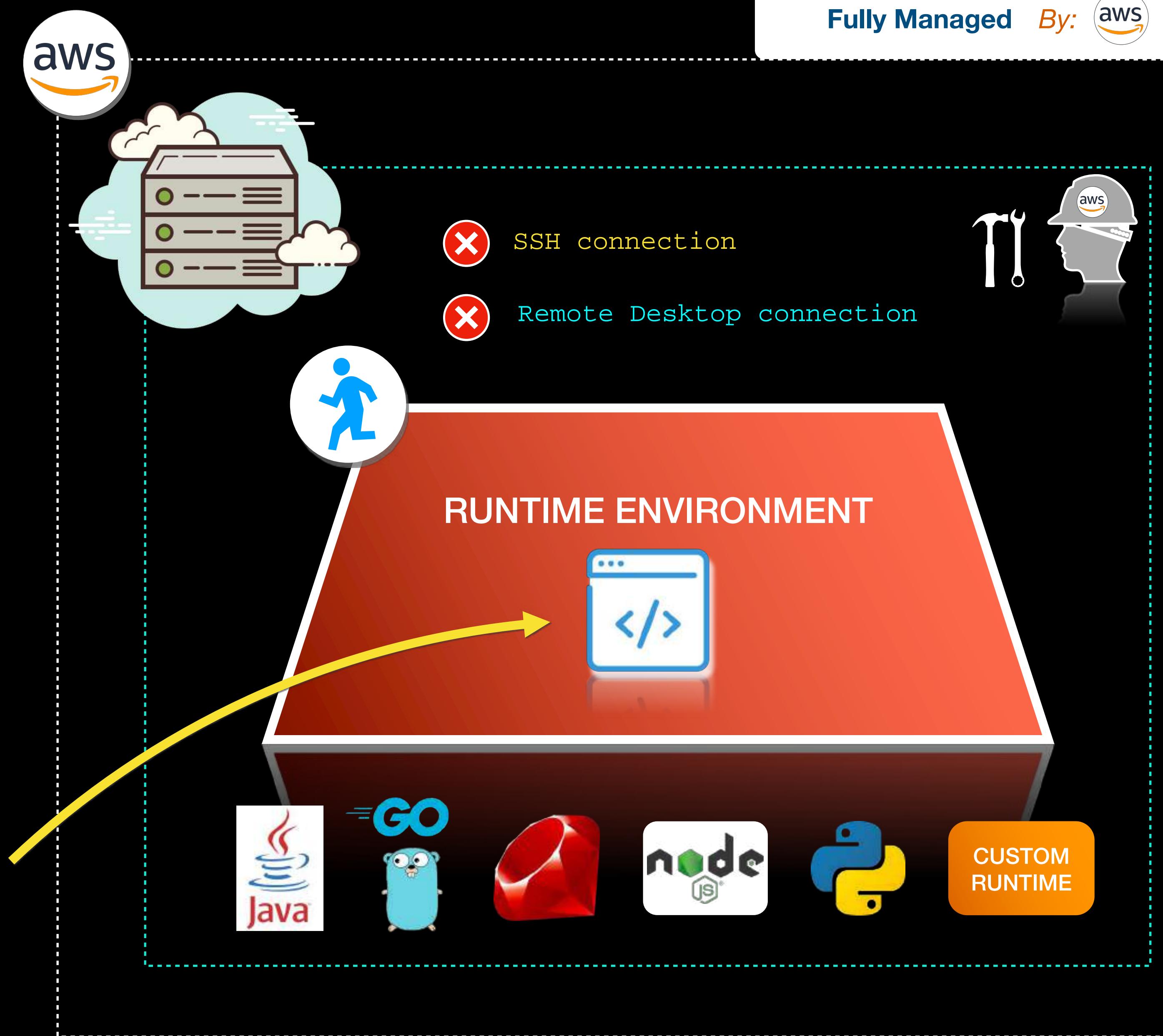
Fully Managed By: 

**Serverless**



AWS Lambda

**Lambda function**



## Orchestration



AWS Batch



AWS Elastic Beanstalk



AWS Batch

- Enables you to run **batch** computing workloads
- Dynamically provisions the optimal quantity and type of compute resources, based on the volume and specific resource requirements.
- Does the planning, scheduling, and execution of your batch computing workloads **using Amazon EC2 instances.**

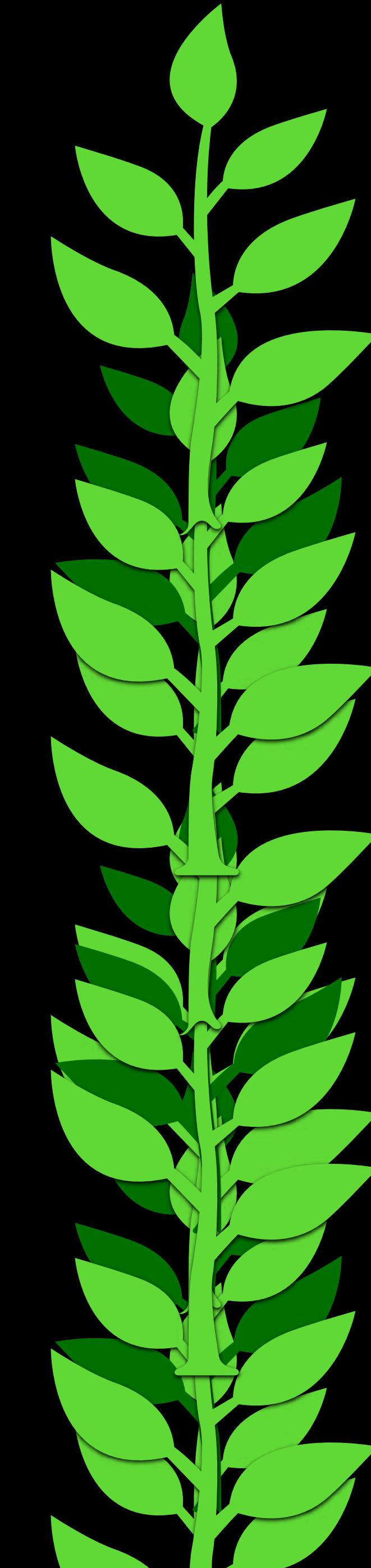


AWS Elastic  
Beanstalk

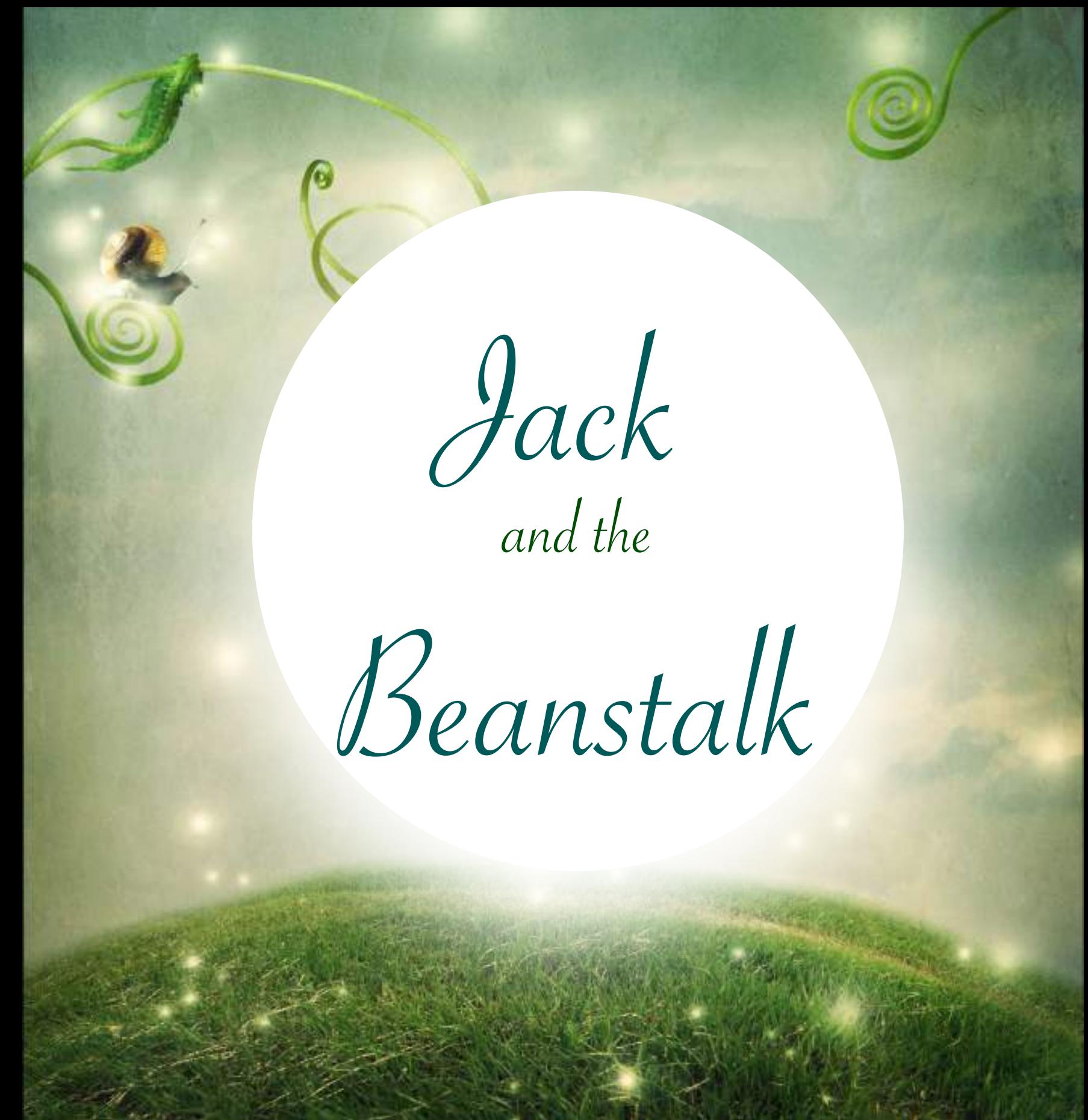
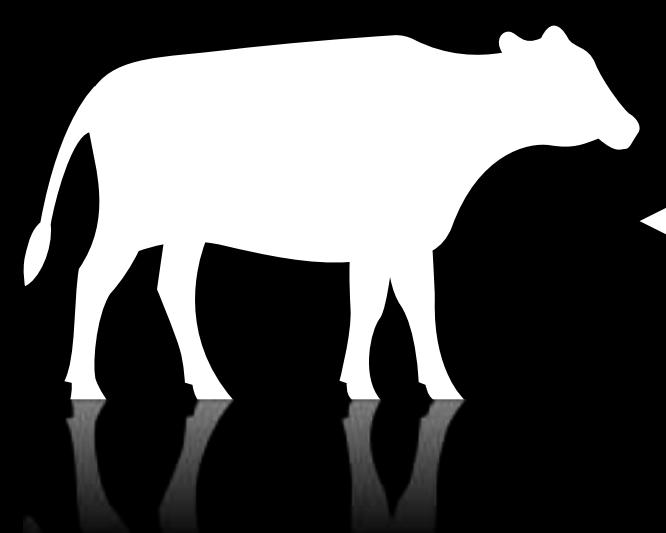
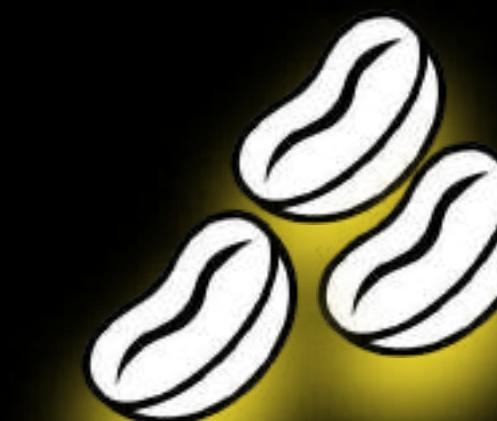
- Automates the deployment, management, scaling, and monitoring of your custom applications in AWS
- Just upload your application and it will automatically handle the common tasks to run your application.
- Handles capacity provisioning, load balancing, database management, auto-scaling, and health monitoring



AWS Elastic  
Beanstalk



*Your Applications*

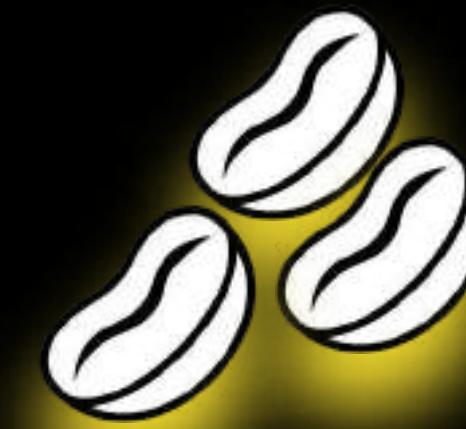




AWS Elastic  
Beanstalk



# Beanstalk



*Your Applications*



Amazon LightSail

- An easy-to-use **Virtual Private Server** (VPS)
- Has its **own web management console**
- Also provides other services like databases, load balancers, DNS records and many more.



## AWS Outposts

- A hybrid service that allows you to run AWS services, like Amazon EC2, in your on-premises data center



AWS Outposts





# AWS Container Services Overview

---



# AWS Container Services



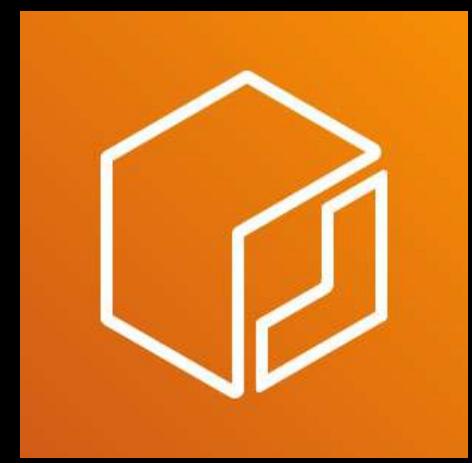
Amazon ECS



Amazon EKS

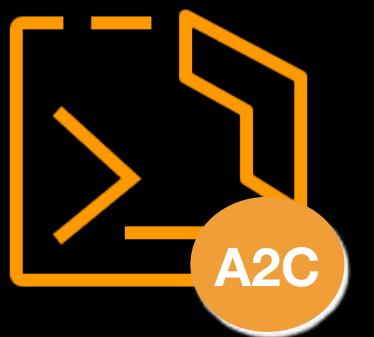


AWS Fargate



Amazon ECR

**CLI Tools**



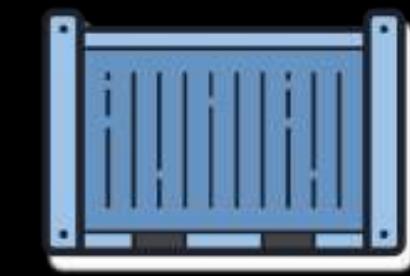
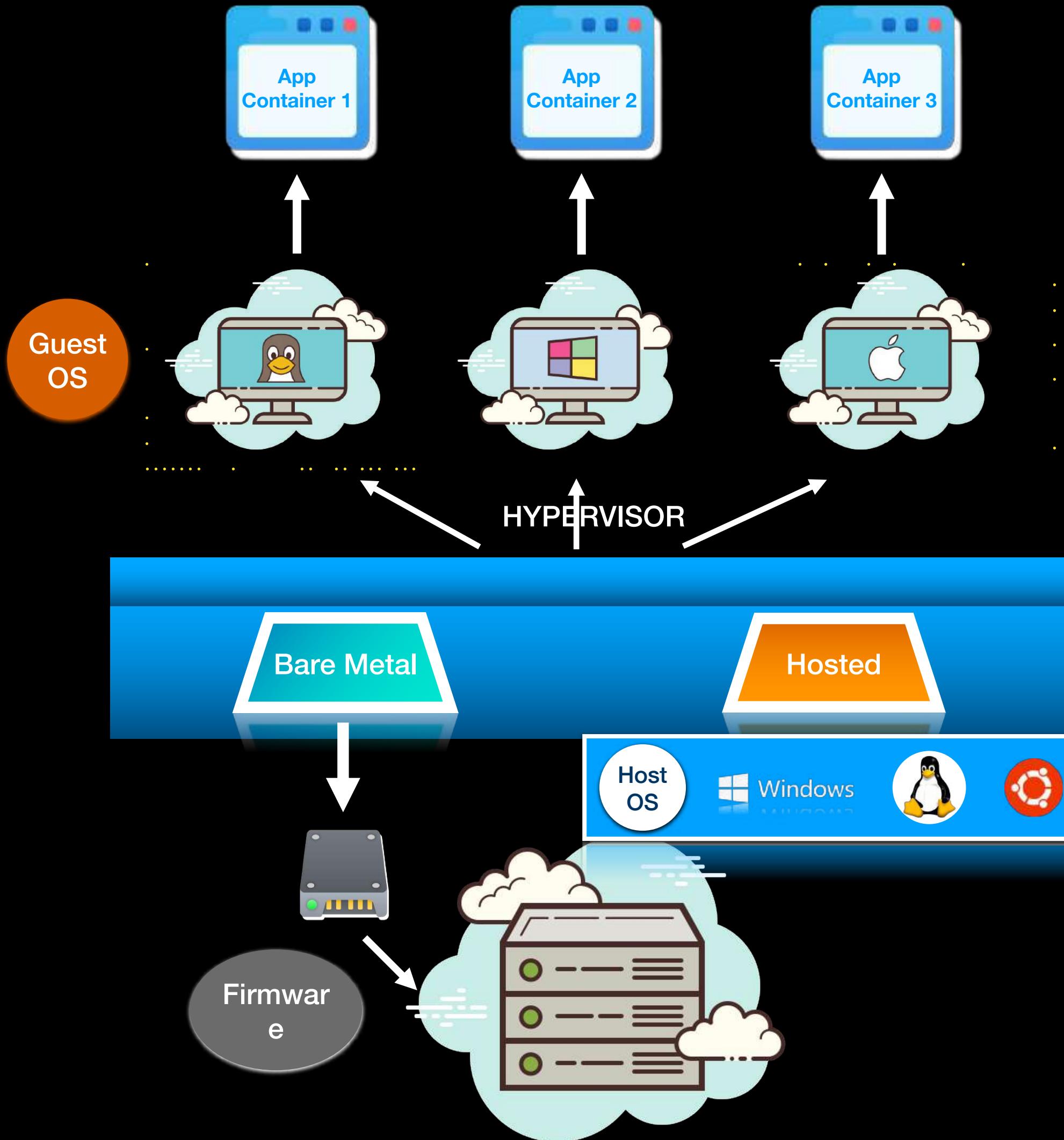
AWS App2Container  
(A2C)



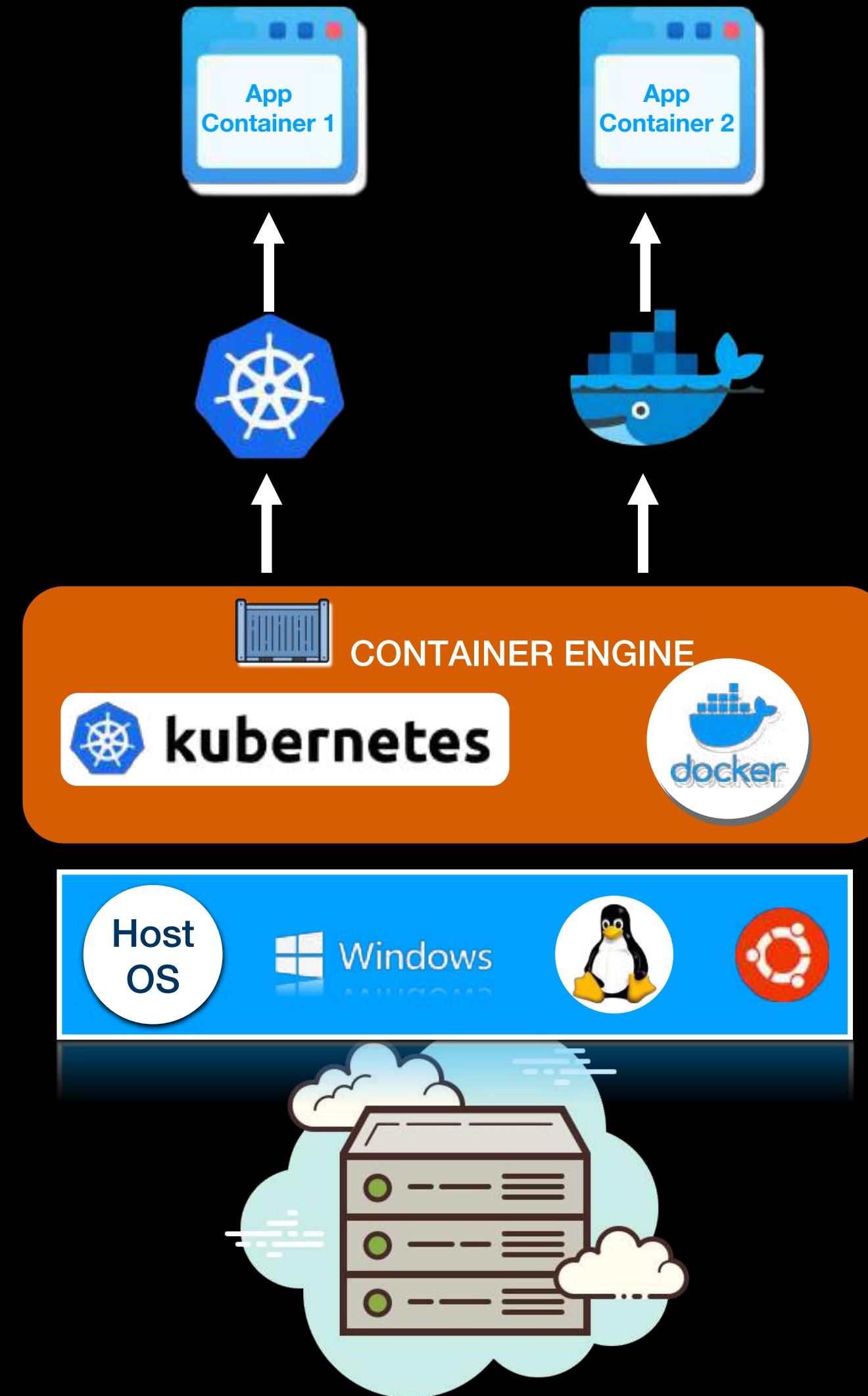
AWS Copilot



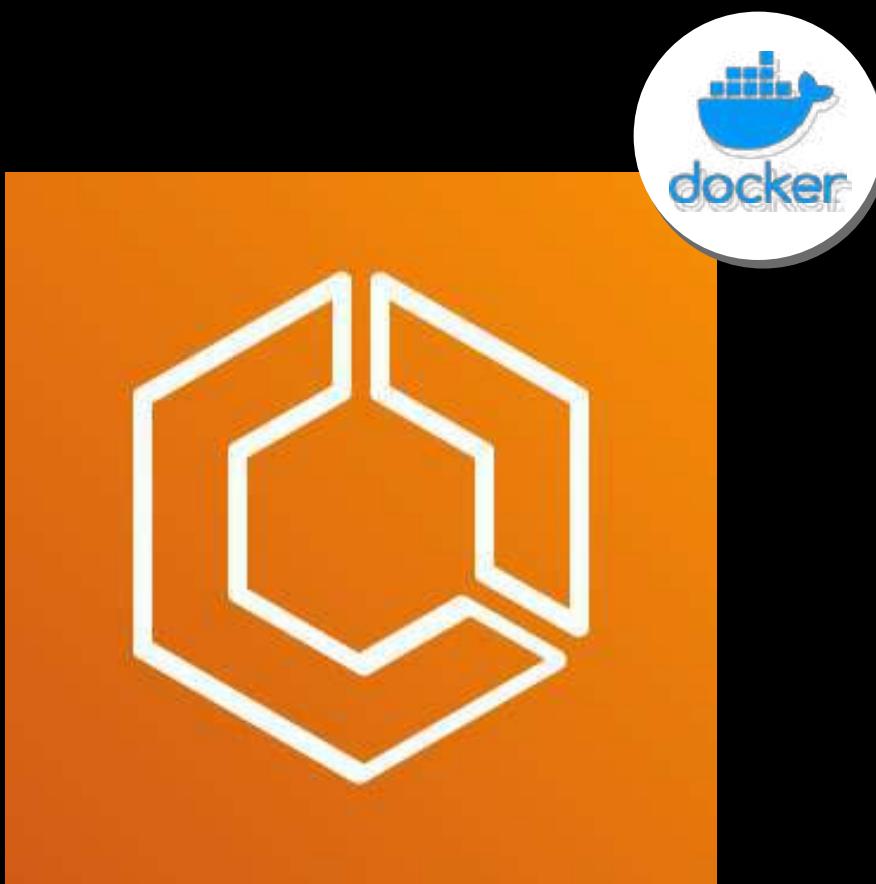
## Virtual Machine



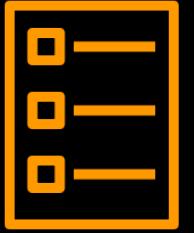
## Container



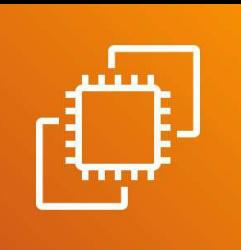
- Amazon **Elastic Container Service** (Amazon ECS)
- A **container orchestration service** that supports Docker containers.
- Allows you to easily install, operate, and scale your cluster management infrastructure in AWS
- Containers are defined in a **task definition** which you use to run an **ECS task** or are **grouped** together as an **ECS service**



Amazon ECS



- Runs your ECS tasks using:



Amazon EC2



AWS Fargate

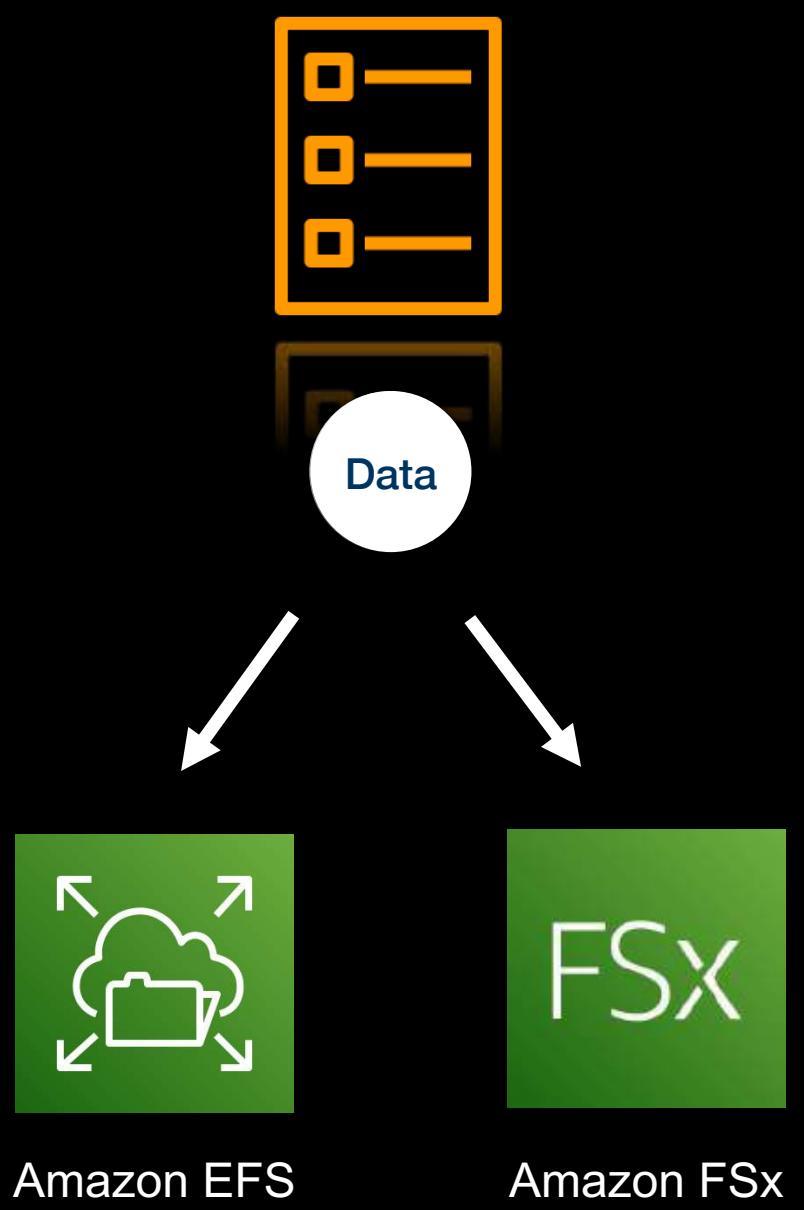
- An **IAM Role** can be attached to your ECS task in the **TaskRoleArn** property of your task definition for security control
- Store your Docker Images to:



Amazon ECR



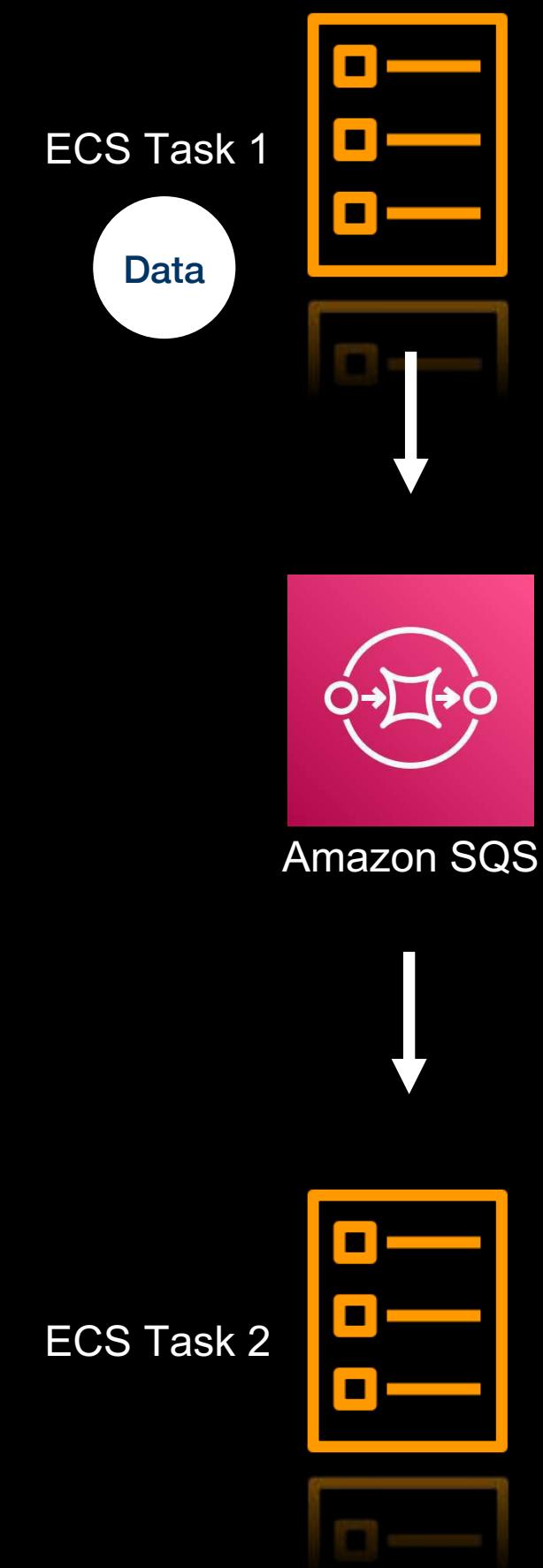
Amazon ECS



## Storage

## Integration

## Scaling



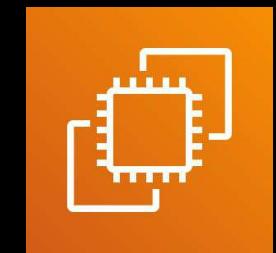
Amazon ECS  
Service Auto Scaling



- Amazon **Elastic Kubernetes Service** (Amazon EKS)
- A fully-managed **Kubernetes** service
- Portable, extensible, and open-source platform for managing containerized workloads and services
- Containers are grouped into **Pods** — the basic operational unit for Kubernetes.
- Launches and orchestrates a cluster of compute resources using:



Amazon EKS

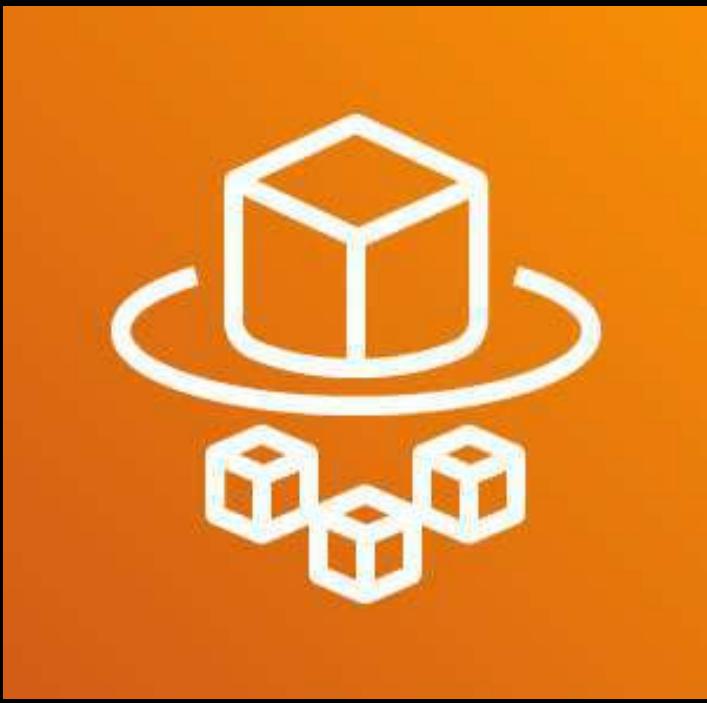


Amazon EC2



AWS Fargate

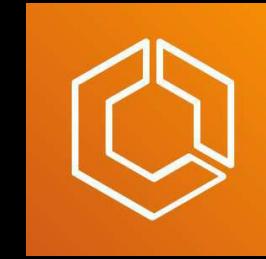
- Considered as **Cloud-agnostic** as it allows you to easily move your workloads to your on-premises network or to other cloud service providers like Microsoft Azure, Google Cloud Platform (GCP) et cetera.



AWS Fargate

- A **serverless** compute engine

- Works on:



Amazon ECS



Amazon EKS

- Allows you to focus on building your applications without worrying about server provisioning, scaling, and management
- Provides a more **cost-effective** solution than a container running on **Amazon EC2 launch type**
- Runs each ECS task or Kubernetes pod in its own kernel.
- Provides the tasks and pods in their own isolated compute environment.

- Amazon **Elastic Container Registry** (Amazon ECR)
- A fully-managed **Docker container registry**
- Allows you to store, manage, and deploy Docker container images.
- Integrated with **Amazon ECS**



Amazon ECR



- **Stores your docker images** in a highly available and scalable architecture
- You can use IAM to provide resource-level control of each repository.



AWS App2Container  
(A2C)

- A **command-line tool**
- Transforms .NET & Java applications to **containerized applications**
- Packages the application artifact and dependencies into container images.
- Configures the network ports and generates the ECS task and Kubernetes pod definitions.



AWS Copilot

- Also a **command-line tool**, just like AWS App2Container (A2C)
- Transforms .NET & Java applications to **containerized applications**
- Enables you to quickly launch and easily manage containerized applications on AWS
- Automates the deployment lifecycle of your containers



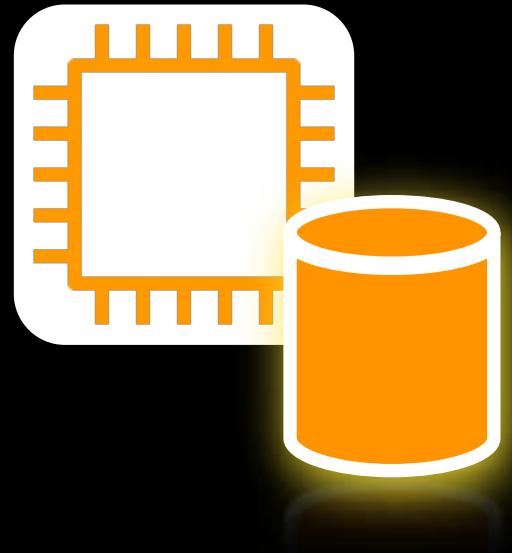
# AWS Storage Services Overview

---



# AWS Storage Services

Built-in component and **NOT**  
a full-fledged AWS Service



Amazon EC2  
Instance Store



Amazon Elastic Block  
Store  
(Amazon EBS)



Amazon Simple Storage  
Service  
(Amazon S3)



Amazon S3 Glacier



Amazon Elastic File  
System  
(Amazon EFS)



Amazon FSx for Lustre



Amazon FSx for Windows  
File Server

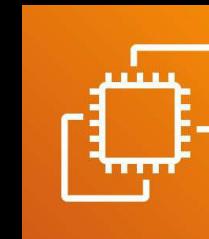


AWS Backup

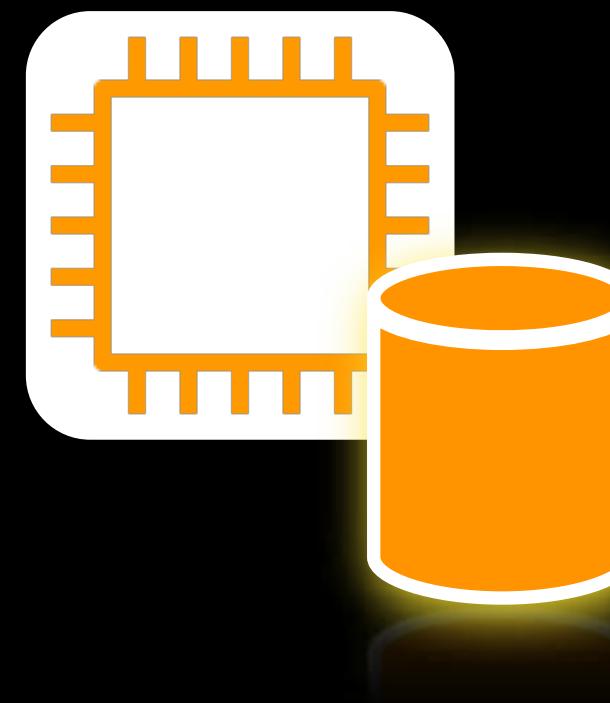


AWS Storage Gateway

Underlying Host Computer that  
powers your



Amazon EC2 Instances

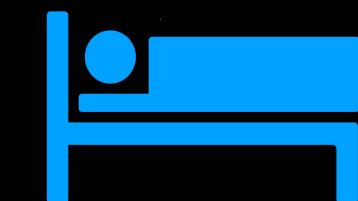


**Amazon EC2  
Instance Store**

- A temporary or **ephemeral** block-level storage
- Uses the local disks or storage volumes that are **physically attached** to the **underlying host computer** of the Amazon EC2 instance.
- Provides **low-latency** access to your data
- Loses its stored data if:
  - The underlying local storage fails
  - The Amazon EC2 Instance:



**Stops**



**Hibernates**



**Terminates**



**e·phem·er·al**

/ə'fem(ə)rel/

*adjective*

lasting for a very short time.  
"fashions are ephemeral"

Similar:

transitory

transient

fleeting

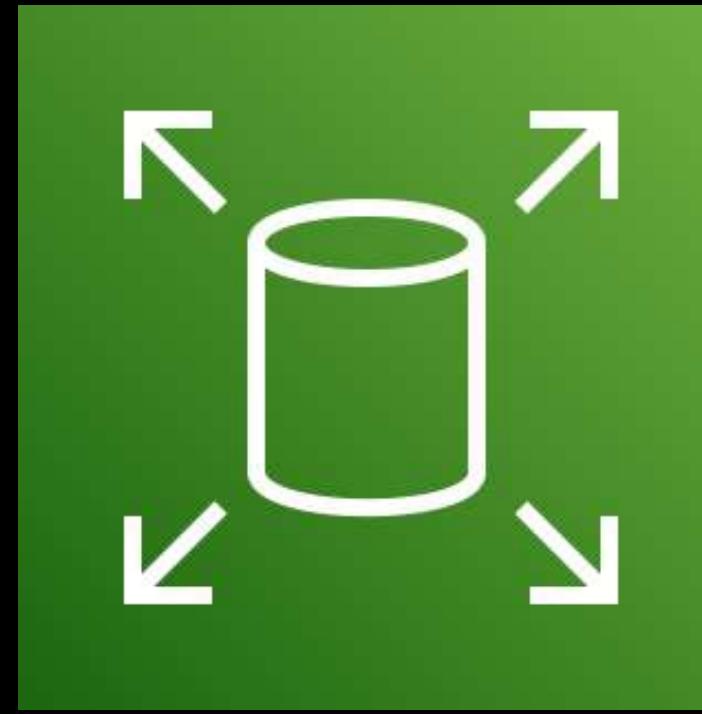
passing

short-lived

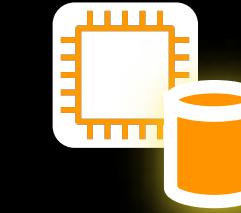
momentary

▼

## Amazon Elastic Block Store (Amazon EBS)



- A persistent block-level storage service
- Your data will still be there even if you stop, restart, or terminate your Amazon EC2 instance, unlike:



Amazon EC2  
Instance Store

- Also called **EBS Volumes**
- Mounted or attached to your Amazon EC2 instances
- **Zonal** in scope — you can only attach a volume to any EC2 instances in the **same Availability Zone**.

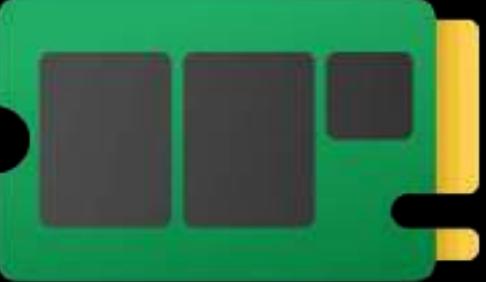
- Can be encrypted at rest using:



AWS Key Management Service  
(AWS KMS)



**Amazon Elastic Block Store**  
(Amazon EBS)

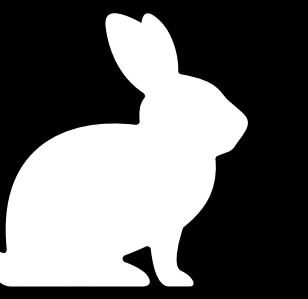


**Solid State Drive  
(SSD)**

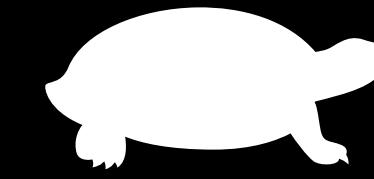


**Hard Disk Drive  
(HDD)**

### Read & Write Speeds



Fast !



Slow...

### Use Case

For workloads with frequent read/write operations

For **data archiving, backups** or throughput-oriented storage

### Dominant Performance Attribute

IOPS

Input/Out operations Per Second

Throughput

Megabit per second (Mbps)

Can be used as Boot Volume for



?

Yes

No



**Amazon Elastic Block Store**  
(Amazon EBS)



**Solid State Drive  
(SSD)**



**Hard Disk Drive  
(HDD)**

## TYPES

:



**General Purpose SSD**

.

.



**Provisioned IOPS SSD**

.



**Throughput Optimized HDD**

.



**Cold HDD**

.

**Can only be attached to a single**



**at a time**

Faster data retrieval than:

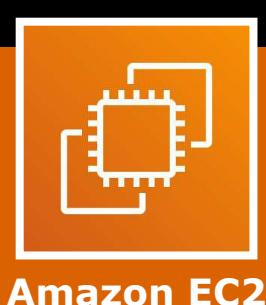


**Amazon S3**



**Amazon EFS**

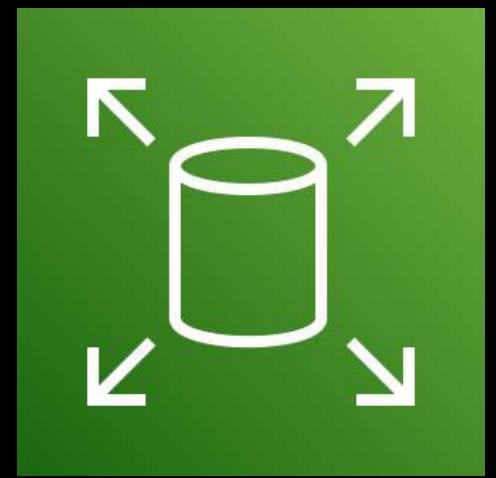
**Can be used as  
Boot Volume for**



**Amazon EC2**

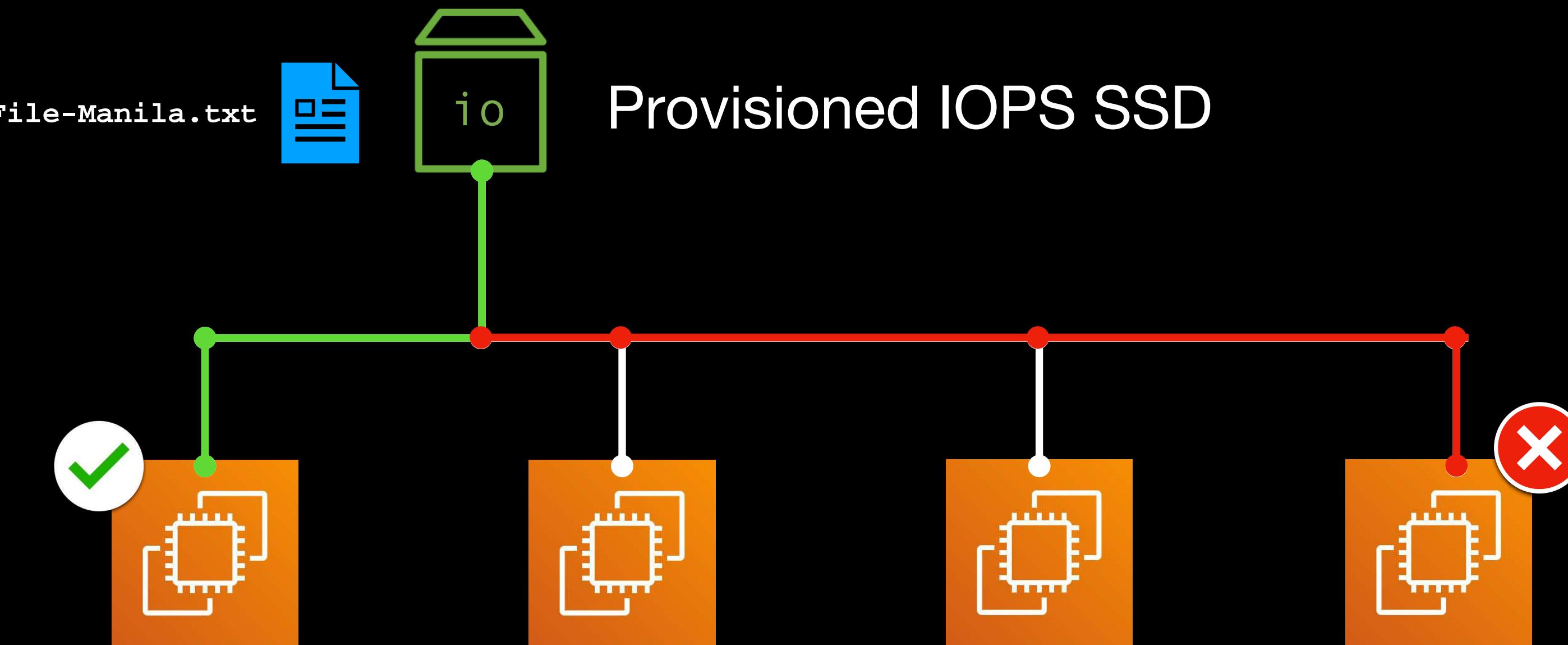


**Cannot be used  
as a Boot Volume**



**Amazon Elastic Block Store**  
(Amazon EBS)

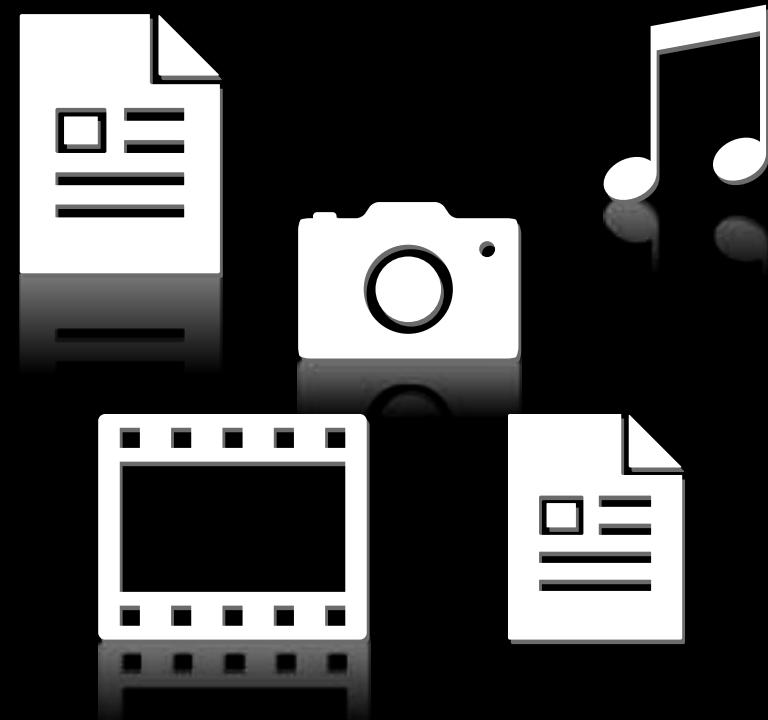
## EBS Multi-Attach



No concurrent file modification



Amazon EFS



## **Amazon Simple Storage Service** **(Amazon S3)**

- An **object storage service**
- Highly **durable** and scalable
- Can store virtually **unlimited amounts** of data
- The files are called “**objects**” that you upload to an **S3 Bucket**
- Access files via a **REST API call**



## Amazon S3 Storage Classes



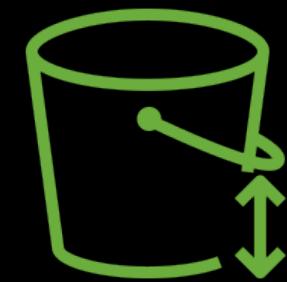
S3 Standard

For **frequently accessed** data



S3 Intelligent-Tiering

For changing or  
**unknown access patterns**



S3 Standard-IA  
(Infrequent Access)



S3 **One** Zone-IA  
(Infrequent Access)

For storing long-lived,  
yet **less frequently accessed** data



S3 Glacier



S3 Glacier **Deep** Archive

For **low-cost long-term storage**  
and data archiving



## Lifecycle Policy



## Access Control List (ACL)

- Secure access to your S3 buckets and objects



## Bucket Policy

- Control external access to your Amazon S3 bucket.



**S3 Versioning**



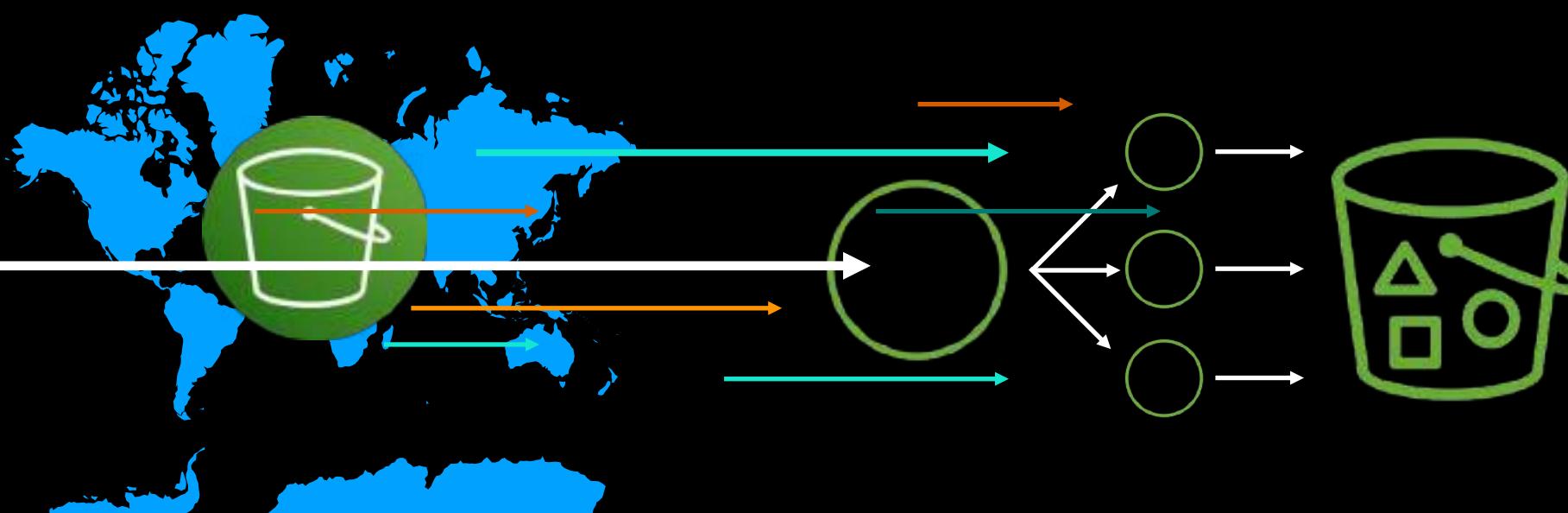
**Multi-Factor Authentication (MFA)**

- Prevent accidental data deletion in Amazon S3.



**Cross Region Replication (CRR)**

- Automatically replicate objects to a different AWS Region for backup purposes



**Transfer Acceleration**

**Multipart Upload**

- Accelerate or expedite the data transfer (upload/download) of S3 objects

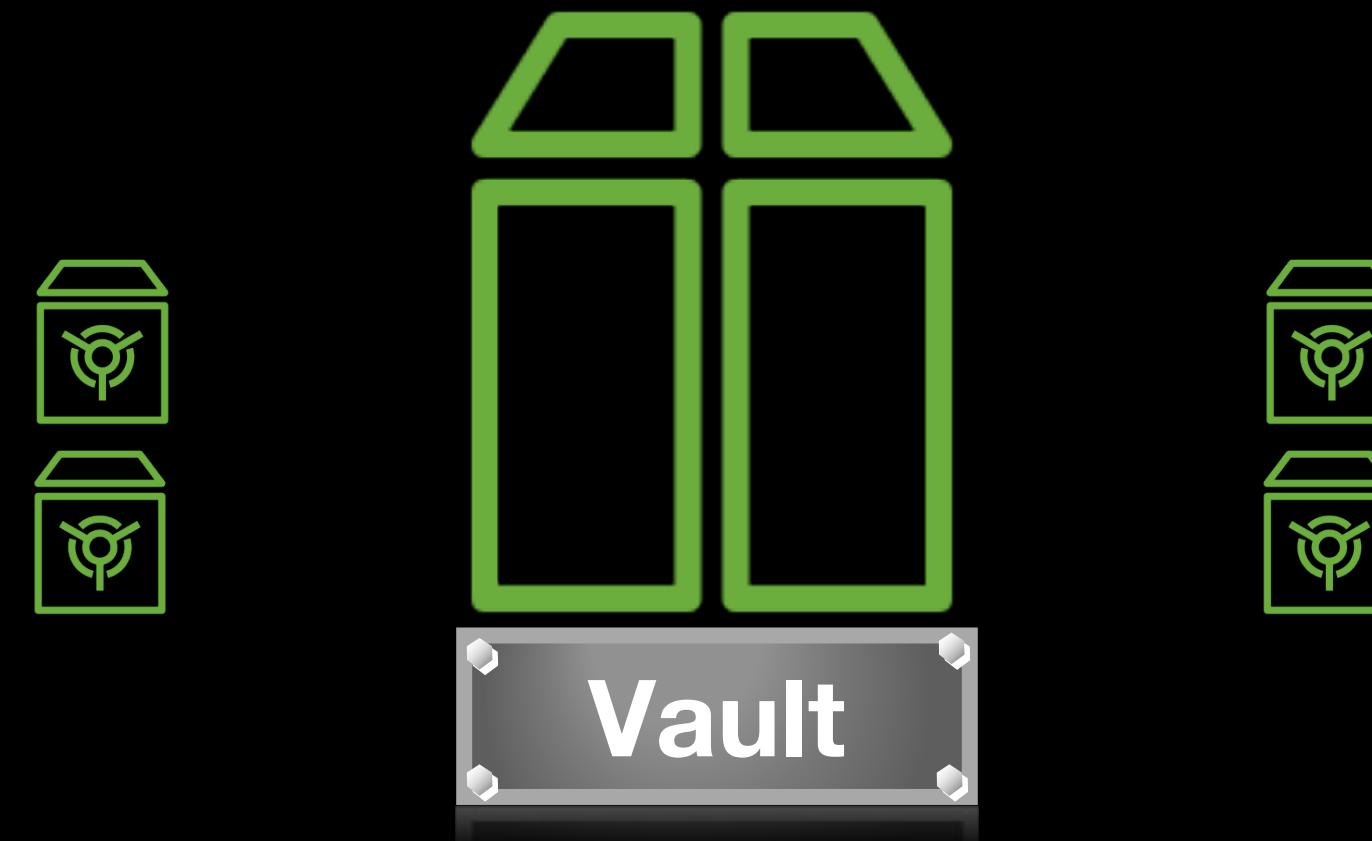
**...and many more S3 features!**



## Amazon S3 Glacier

- One of the storage classes in Amazon S3
- Has its own web management console apart from Amazon S3
- Based on the word — **Glacier**:
  - **Rarely** Accessed Data (Cold)   Cold HDD
  - **Frequently** Accessed (Hot) 
- Low-cost storage for data archiving and long-term backup.





S3 Glacier	VS	S3 Glacier Deep Archive
<b>LOW</b> <b>\$ \$</b>		<b>LOWEST</b> <b>\$</b>
<b>90 Days</b>		<b>180 days</b>
You will be billed for the <b>entire 90 Days</b>		You will be billed for the <b>entire 180 Days</b>
Normal storage usage charge	<b>COST</b>	Normal storage usage charge
Normal storage usage charge	<b>MINIMUM STORAGE DURATION</b>	Normal storage usage charge
	<b>DATA DELETED AFTER 1 DAY (24 HOURS)</b>	
	<b>DATA DELETED AFTER 90 DAY</b>	
	<b>DATA DELETED AFTER 180 DAYS</b>	



## S3 Standard

**HIGHEST \$ \$ \$**

**None**

Regular storage usage charge  
(24 hours)

Regular storage usage charge  
(30 days)

Regular storage usage charge  
(90 days)

**VS**



## S3 Glacier

**LOWEST \$**

**90 days**

You will be billed for the **entire 90 Days**

You will be billed for the **entire 90 Days**

**COST** Timed Storage - Byte Hours

**MINIMUM STORAGE DURATION**

**DATA DELETED AFTER  
1 DAY (24 HOURS)**

**DATA DELETED AFTER  
30 DAYS**

**DATA DELETED AFTER  
90 DAYS**



## Archive Retrieval Options

EXPEDITED

STANDARD

BULK



S3 Glacier

1 - 5 minutes

3 - 5 hours

5 - 12 hours



S3 Glacier  
Deep Archive

NOT AVAILABLE

Within  
12 Hours

Within  
48 hours

- A **scalable shared file storage** service
- Provides a **POSIX**-compliant (Portable Operating System Interface) shared file system
- Can be **simultaneously** accessed by multiple Amazon Linux EC2 instances in different Availability Zones.
- Uses the **Network File System (NFS)** protocol. Works as a **file share**



## Amazon Elastic File System (Amazon EFS)

- Only supports:



**Linux Servers**



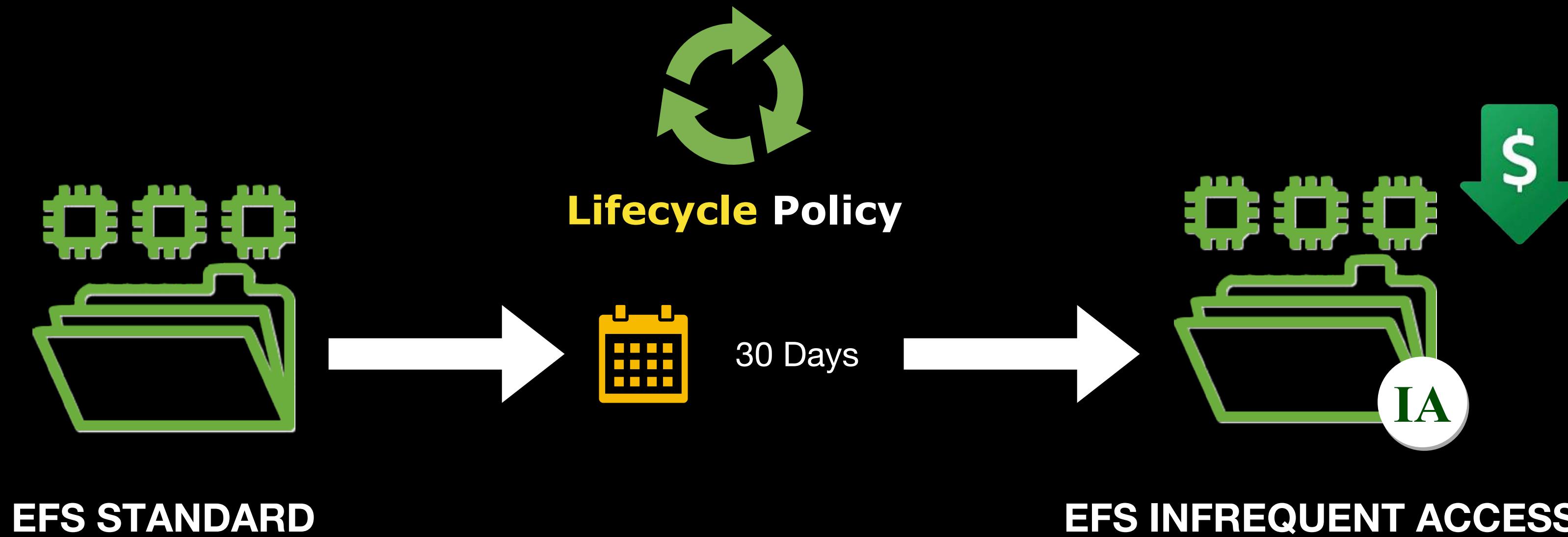
=



Amazon FSx for  
Windows File Server



**Amazon Elastic File System**  
(Amazon EFS)





**Amazon FSx**



Amazon FSx for Lustre



Amazon FSx for  
Windows File Server



## Amazon FSx for Lustre

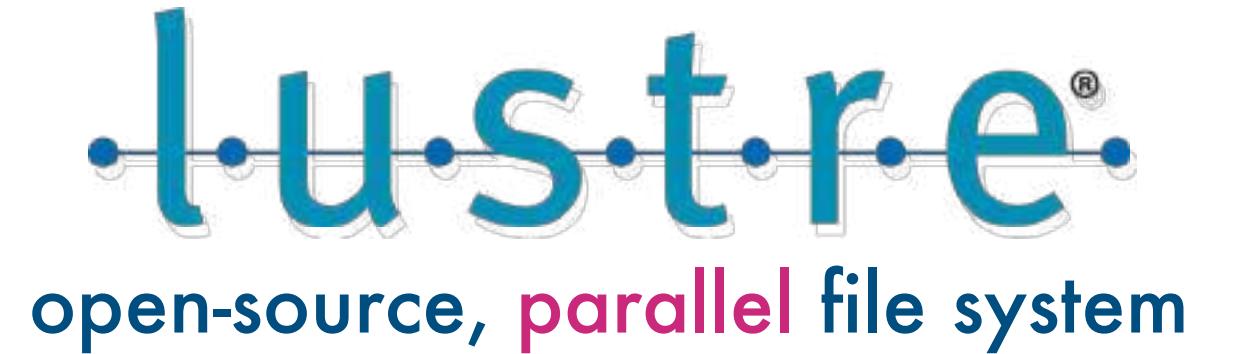


## Amazon Elastic File System (Amazon EFS)

- A **scalable shared file storage** service
- Provides a **POSIX**-compliant (Portable Operating System Interface) shared file system
- Can be simultaneously accessed by multiple Amazon Linux EC2 instances in different Availability Zones.
- Uses the **Network File System (NFS)** protocol
- Only supports:



**Linux Servers**



Linux Cluster

=

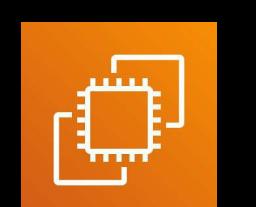


a parallel file system used for large-scale cluster computing.



## Amazon FSx for Lustre

- Primarily used for High-Performance Computing, Machine Learning, or HPC applications
- For workloads that need high-performance parallel storage for frequently accessed hot 😭 data.
- Provides a throughput of hundreds of gigabytes per second
- Offers millions of IOPS
- You can mount an Amazon FSX for Lustre file share to:



Amazon EC2



Amazon ECS



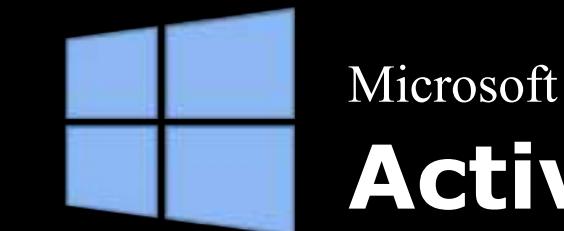
Amazon EKS

- Use the Container Storage Interface (CSI) to connect to your Amazon EKS cluster.



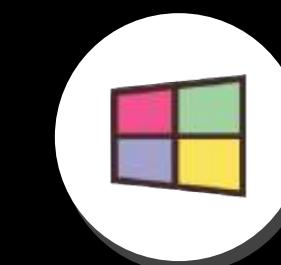
**Amazon FSx for  
Windows File Server**

- A fully managed Microsoft Windows file server service
- Uses the Server Message Block (SMB) protocol
- Can be integrated to your existing:



Microsoft

**Active Directory**

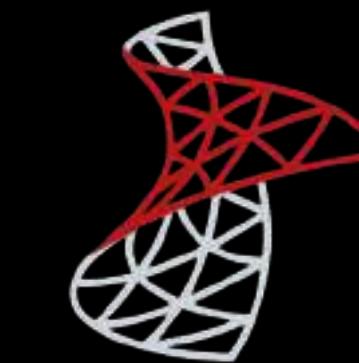


AWS Managed  
Microsoft AD



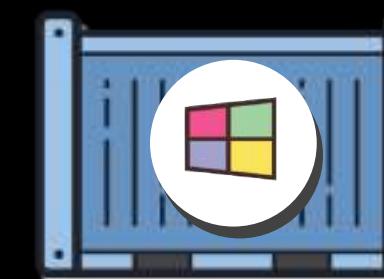
Microsoft

**SharePoint**



Microsoft

**SQL Server**



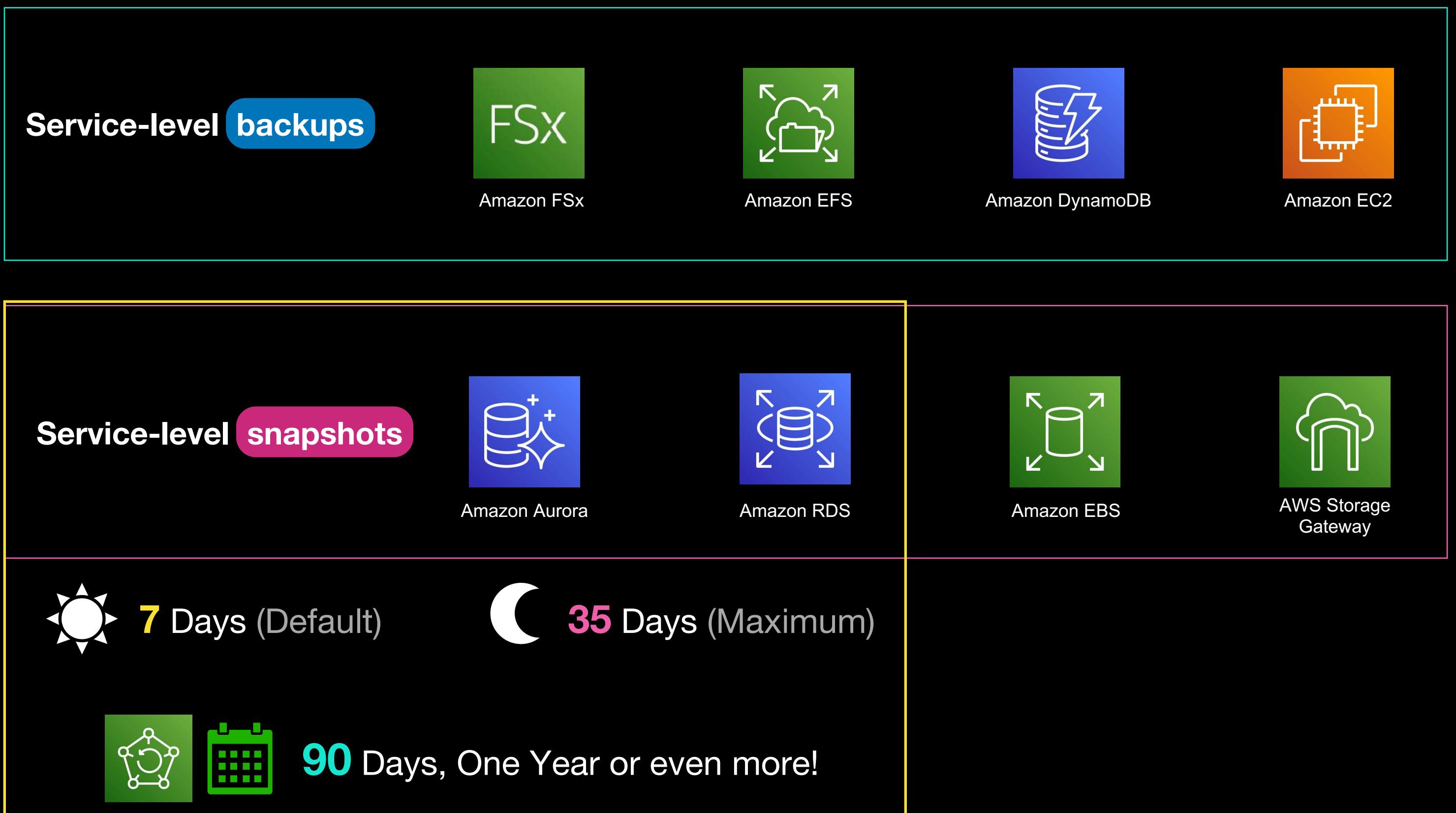
Microsoft

**Containers**

- A fully managed backup service
- Automates your server and database backup processes.



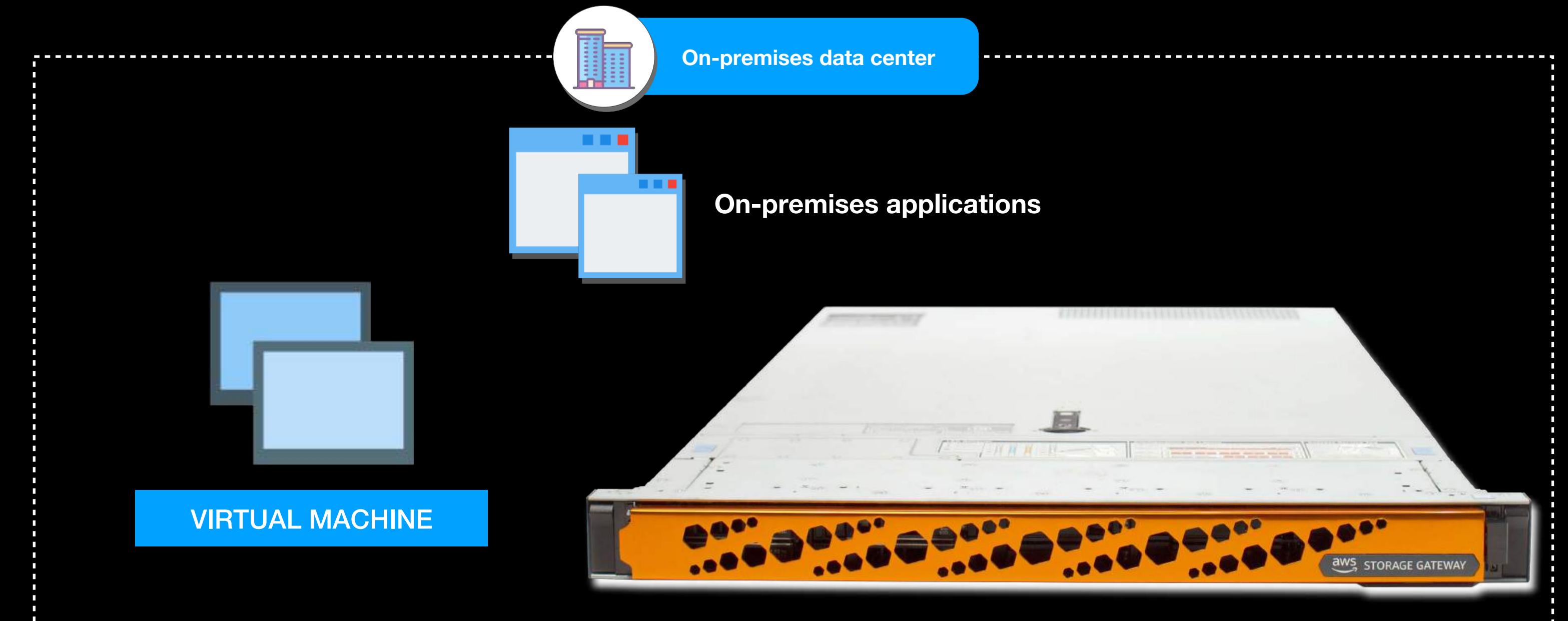
## AWS Backup



- A hybrid cloud storage service
- Connects your on-premises applications and data storage to the AWS Cloud.
- Integrate your local & cloud storage systems by using a gateway.



## AWS Storage Gateway





## File Gateway

Store and retrieve objects in  Amazon S3 using **NFS** and **SMB** protocols

Can be integrated with:  
 AWS Managed Microsoft AD    

Provides a **hardware appliance** hosted on-premises

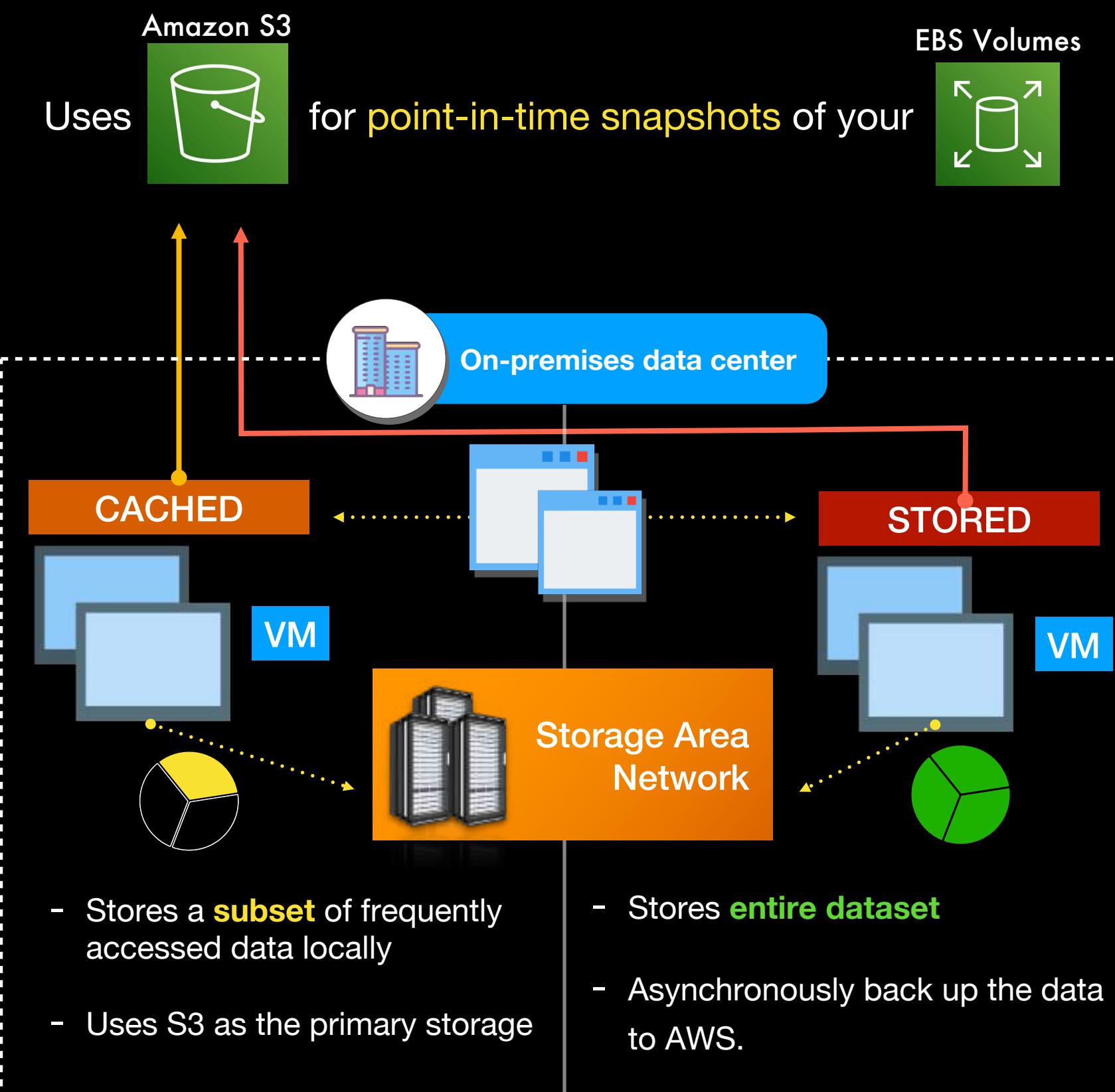


To replicate your local data to  Amazon S3



## Volume Gateway

Provides **block storage** to your **on-premises apps** with low-latency via the Internet Small Computer System Interface (**iSCSI**)



## Tape Gateway

A cloud-based Virtual Tape Library

Uses  to back up the tapes

Can store the archived tapes in:



- On-premises apps can connect to the tape gateway as iSCSI devices

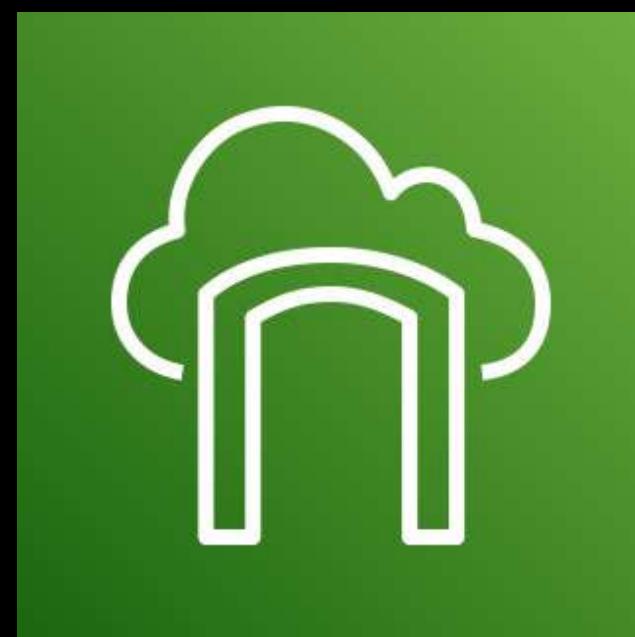
- **Reduce costs** by eliminating the use of physical backup tapes

**REPLICATE DATA**

**INTEGRATION**

**MOVE DATA**

**Migration**



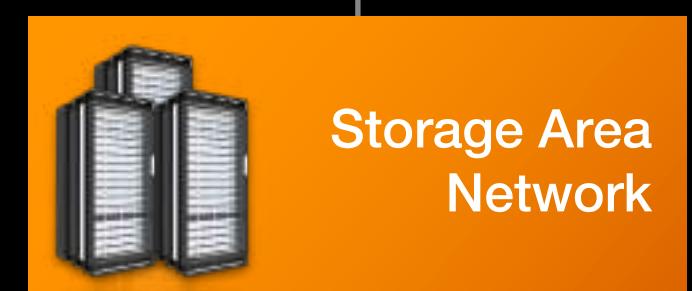
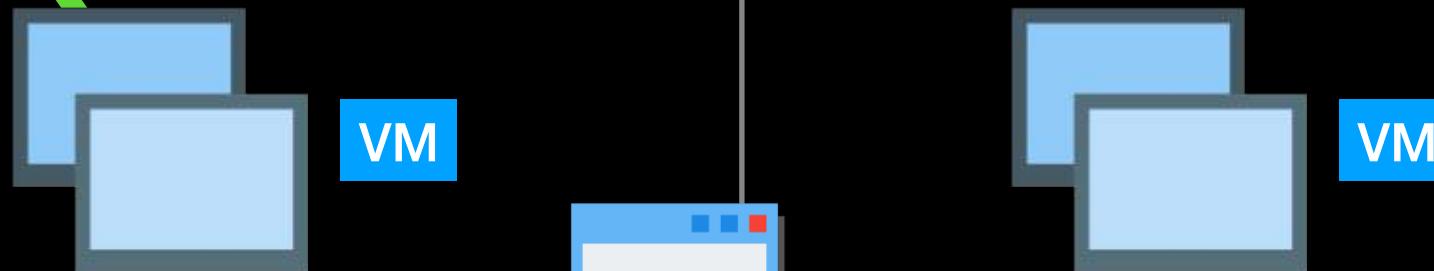
**AWS Storage Gateway**



*On-premises data will  
still be actively used*



On-premises data center



**AWS DataSync**



*On-premises data would not  
be utilized anymore/will be  
decommissioned*

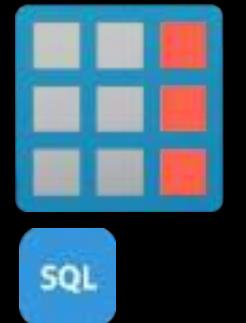


Tutorials Dojo



# AWS Database Services Overview

---



Relational

A C I D

Atomicity  
Consistency  
Isolation  
Durability



Amazon RDS



Amazon Aurora



Data warehouse



Amazon Redshift



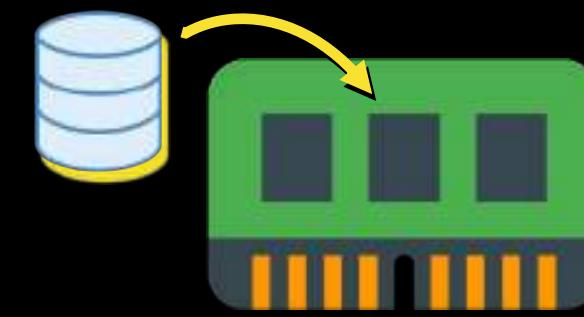
NoSQL



Amazon DynamoDB



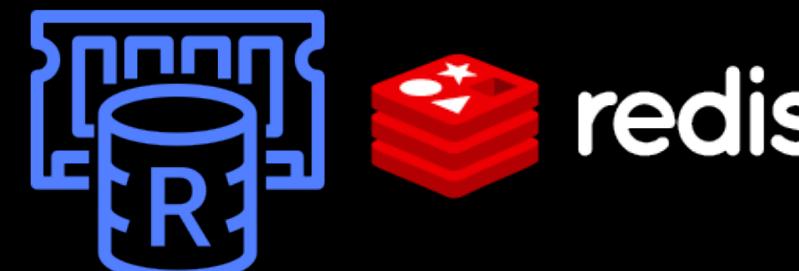
Amazon DocumentDB



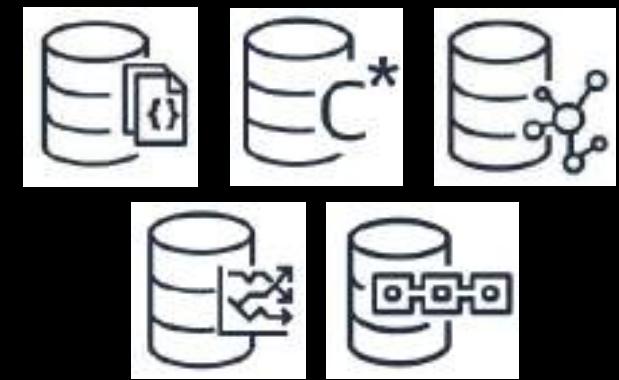
In-Memory



Amazon ElastiCache



memcached



Other Databases



Amazon Keyspaces



Amazon Neptune



Amazon Timestream

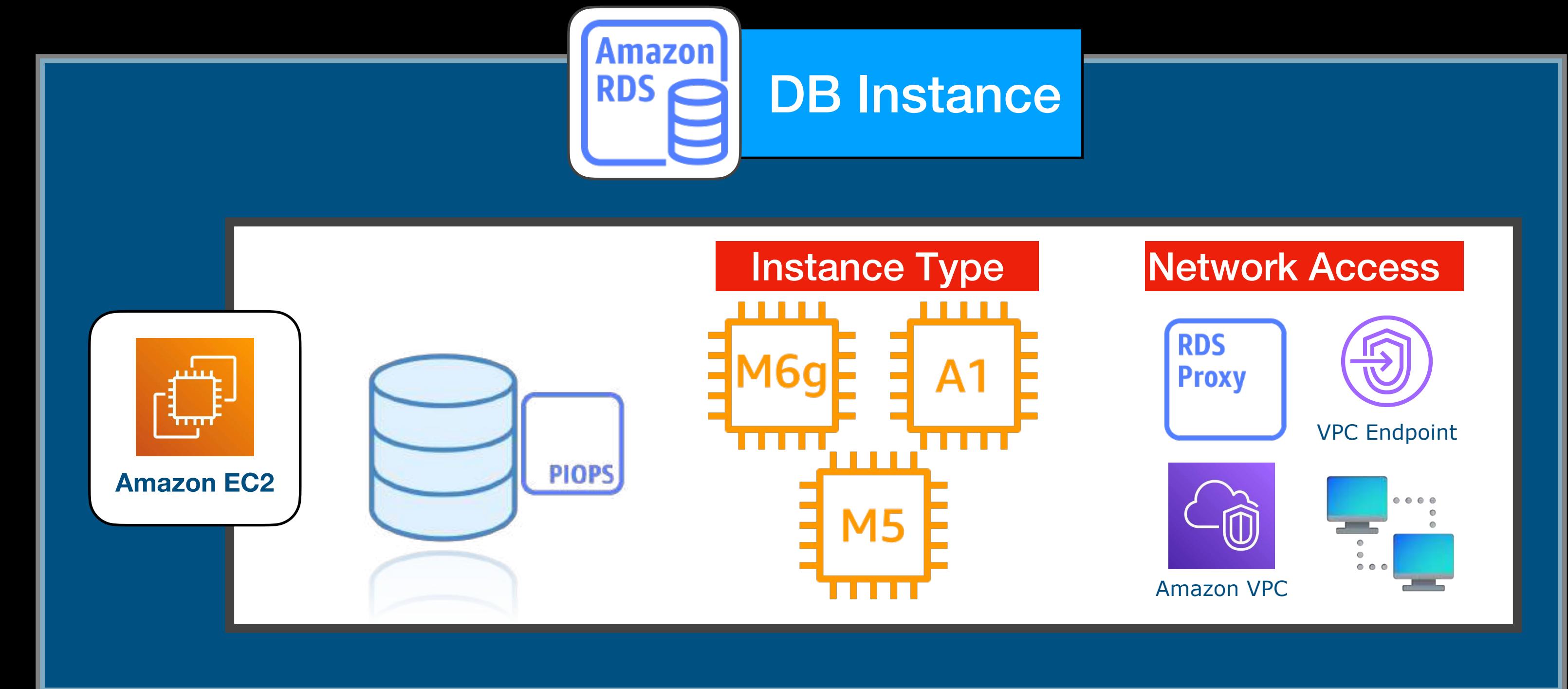


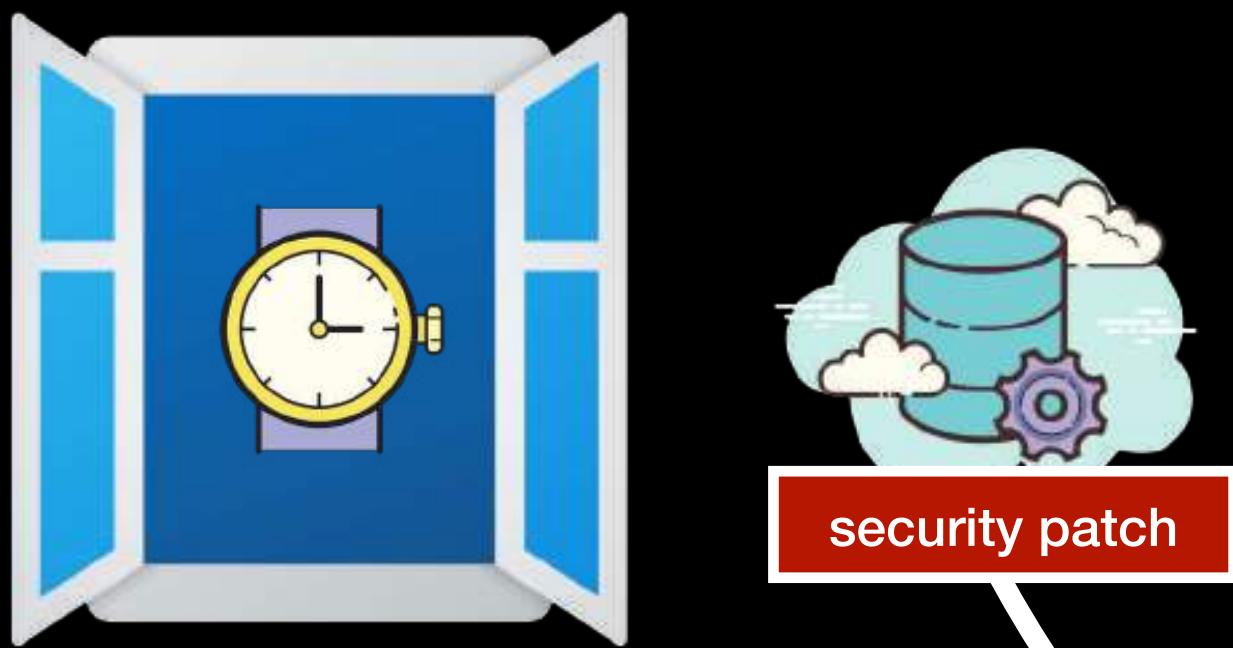
Amazon Quantum Ledger

- A **relational** database that is managed by both you (limited access) and AWS.
- **The time-consuming tasks are handled by AWS** — such as hardware provisioning, patching, backups, and maintenance.
- You can **configure the underlying EC2 instance** used by Amazon RDS



## Amazon Relational Database Service (Amazon RDS)

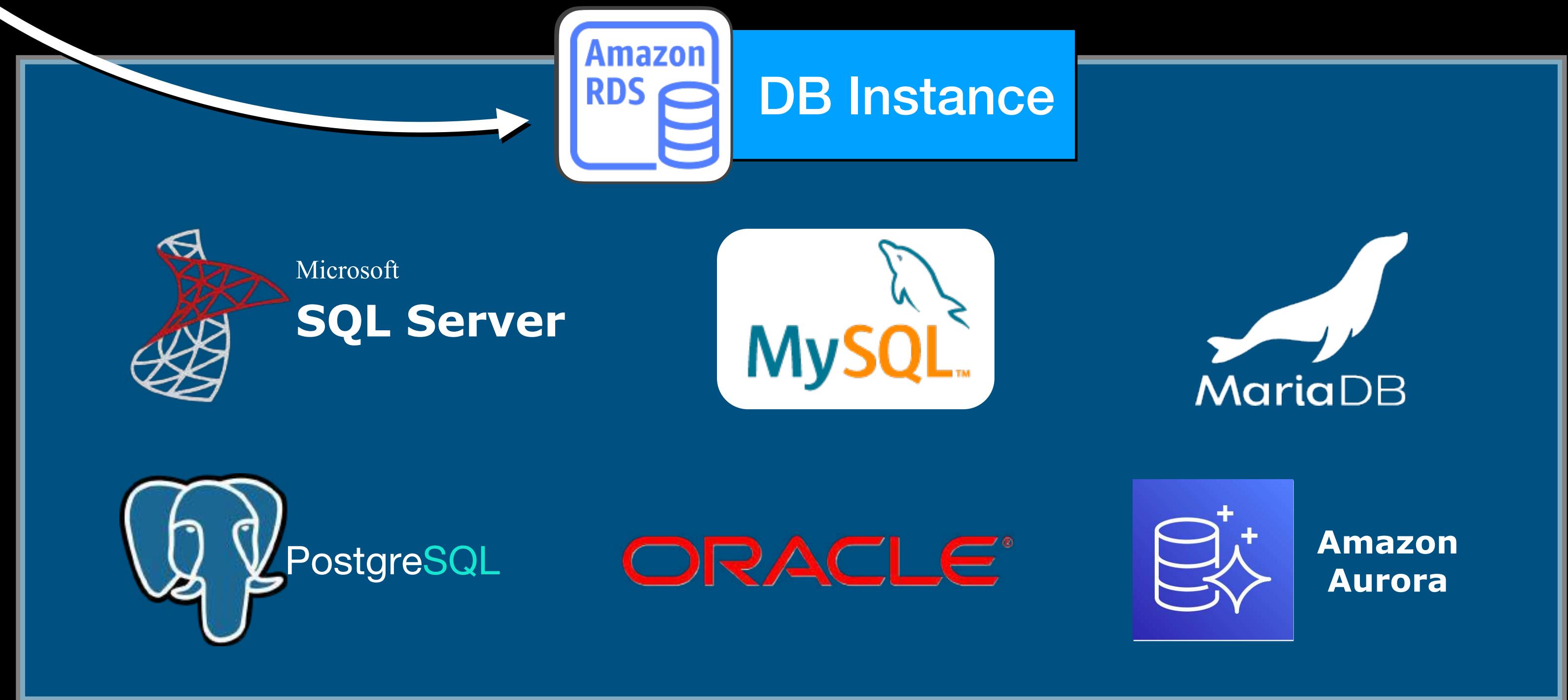




- You decide the **actual time** for the patches to be applied on its **maintenance window**
- Can run various types of **database engines**:

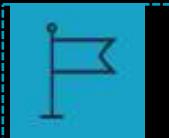


**Amazon Relational Database Service**  
(Amazon RDS)



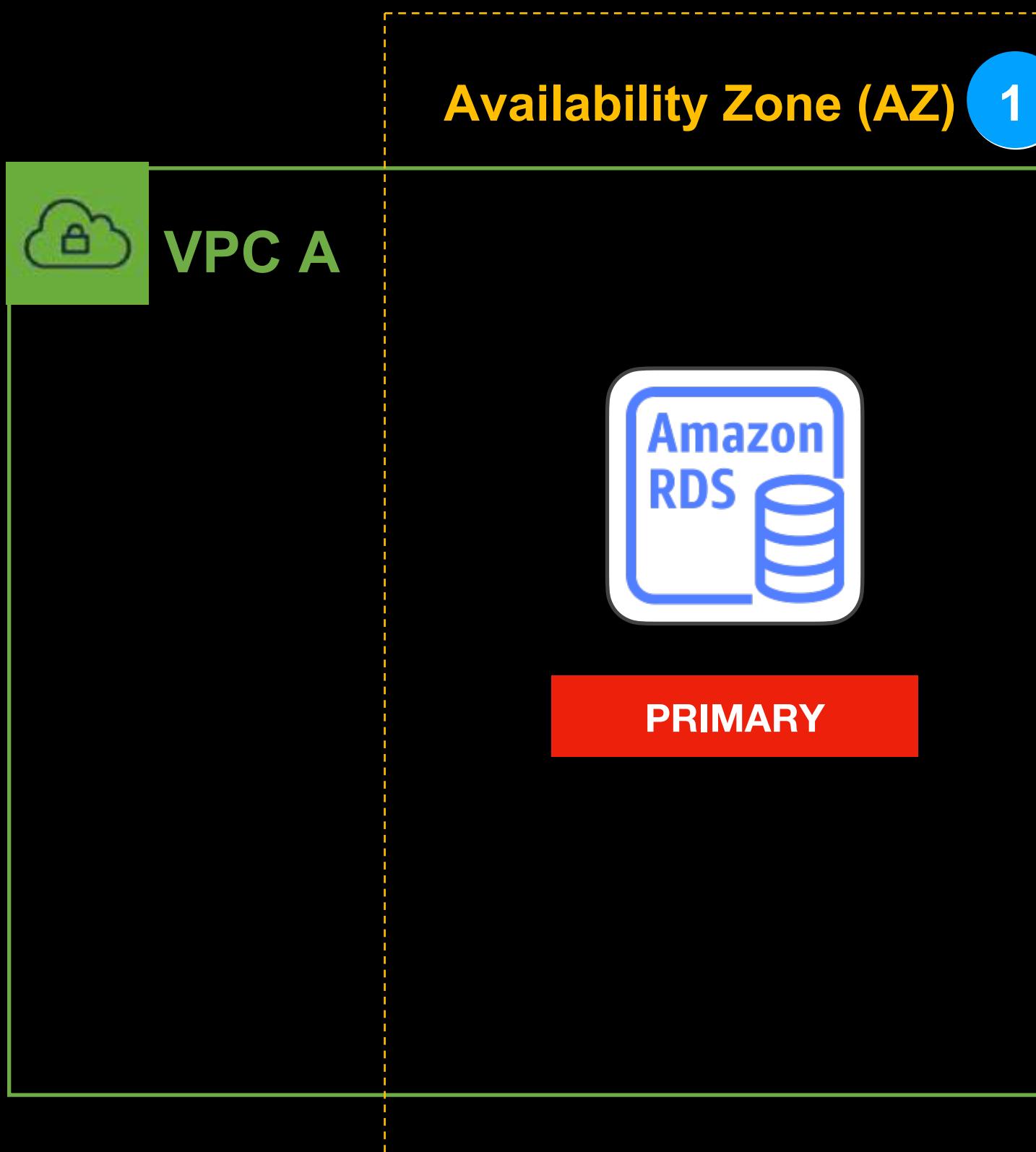


AWS Cloud

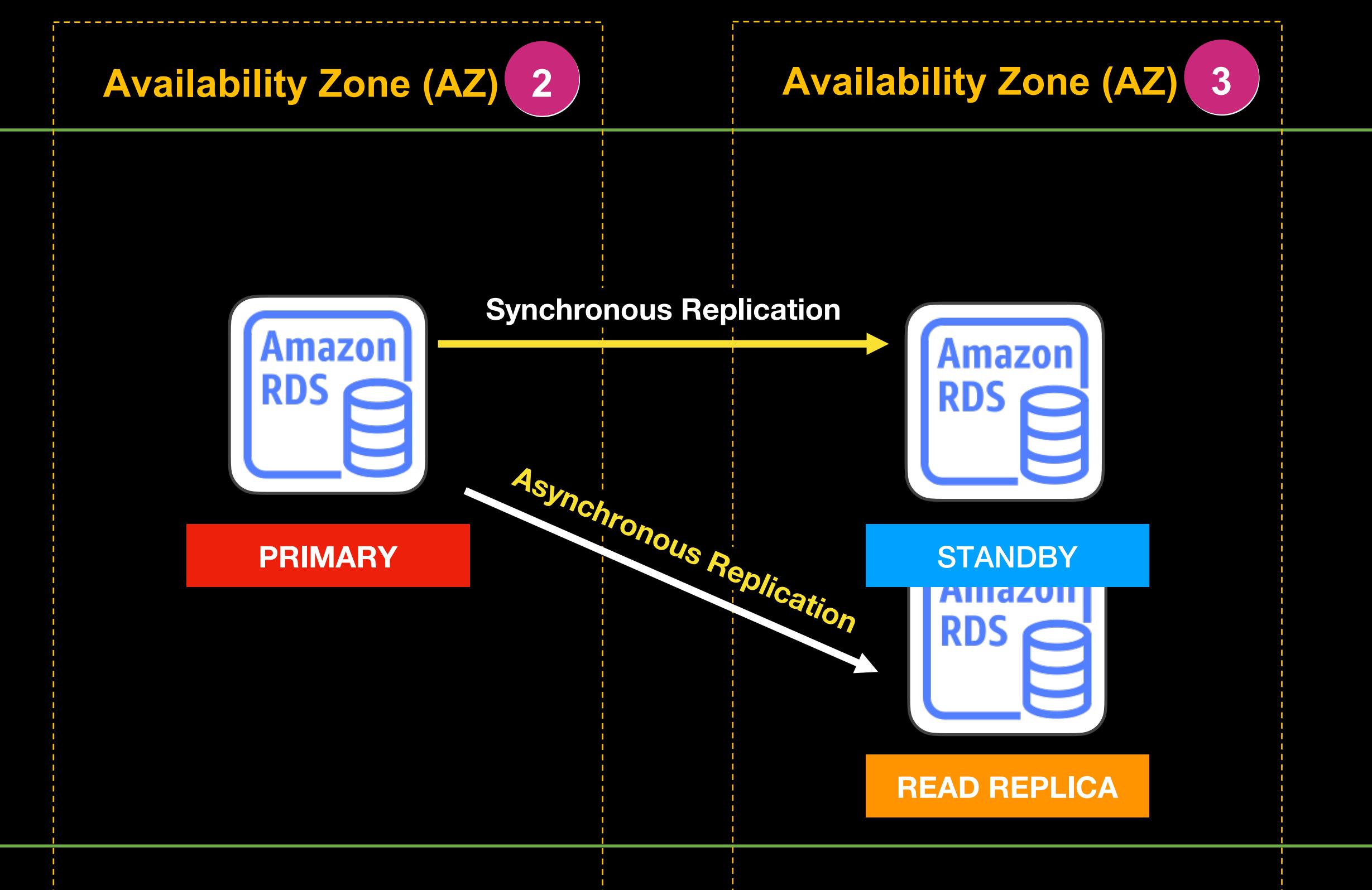


N. Virginia Region

## Single AZ

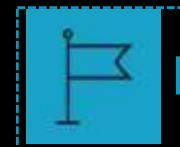


## Multi-AZ





AWS Cloud



N. Virginia Region

Single AZ

Multi-AZ



VPC A

Availability Zone (AZ) 1



PRIMARY

Availability Zone (AZ) 2



PRIMARY

Availability Zone (AZ) 3



STANDBY



READ REPLICA

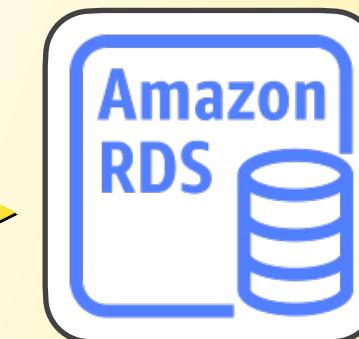
Asynchronous Replication



VPC B



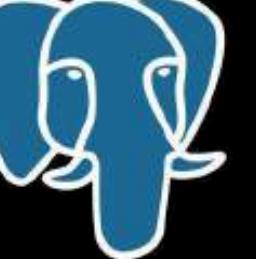
Ohio Region



READ REPLICA



Amazon Aurora

- A type of a **database engine** (that you can run on Amazon RDS) and a **fully managed database** service.
- Compatible with:  MySQL  PostgreSQL
- Scales automatically, performs faster, and costs lower than other databases
- Can **automatically grow its data storage**
- Deployed as a **database cluster** that consists of:
  - **Similar to Multi-AZ Deployments** in Amazon RDS
  - A cluster has a **single-master configuration** where applications can only write data to a single, master DB instance.
  - In a **multi-master cluster**, all DB instances have read/write capability.





**Amazon Aurora**



**Amazon Relational Database Service**  
(Amazon RDS)



- Suitable for applications that read or write **constantly changing data**, such as [Online Transaction Processing](#) applications or OLTP.



## Data warehouse

- A fully managed **data warehouse**
- Allows you to **analyze all your data using standard SQL** or through your existing Business Intelligence tools
- Optimized to **analyze relational data** coming from transactional systems, business applications, and other sources for fast SQL queries.
- Offers a **concurrency scaling** feature that supports virtually unlimited concurrent users and concurrent queries
- Has a feature called **Redshift Spectrum** that allows you to query and retrieve structured and semistructured data from files stored in:



## Amazon Redshift



## Amazon S3



**Amazon Redshift**



- Primarily used for **Online Analytical Processing or OLAP** applications like data reporting and analytics.



## NoSQL Databases

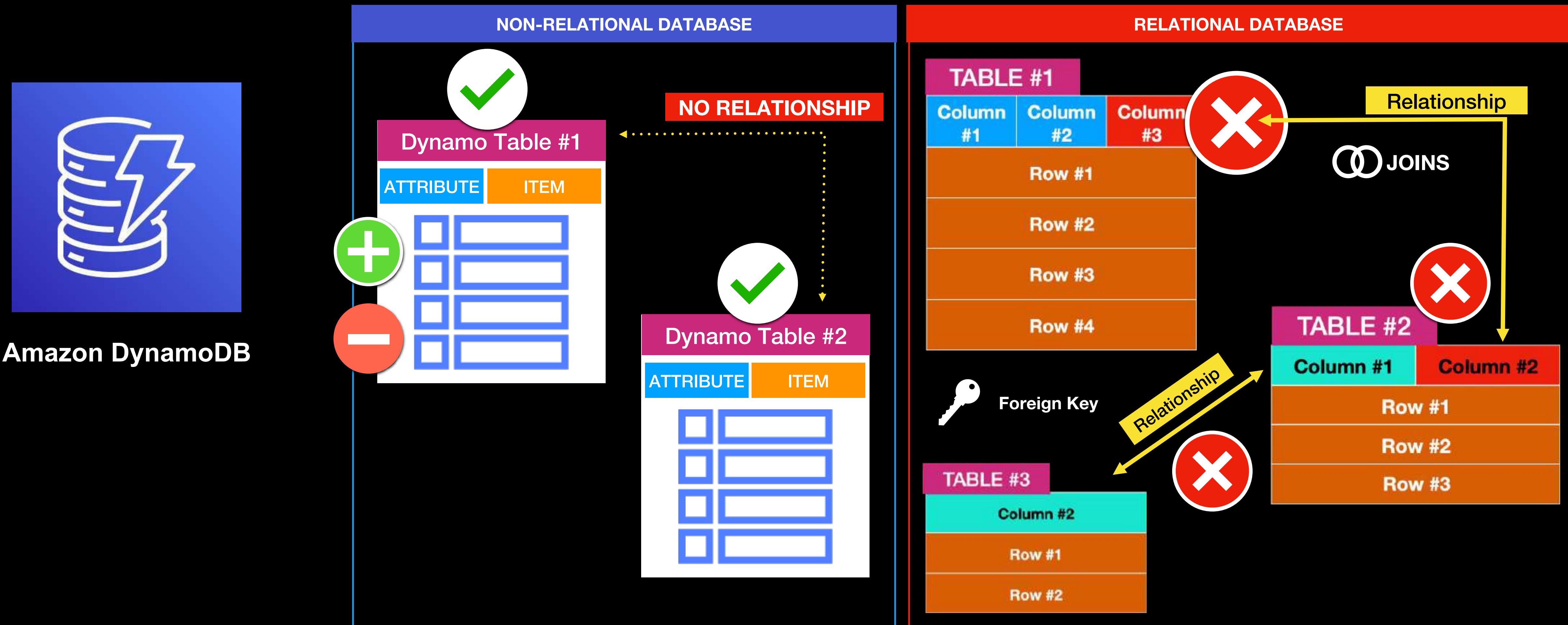


Amazon DynamoDB



Amazon DocumentDB

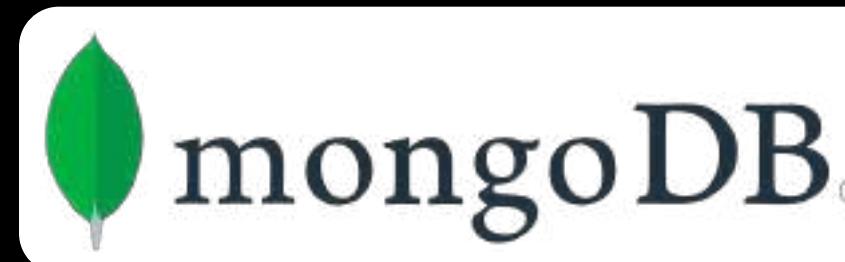
- A fully managed **NoSQL database** service
- A non-relational database that **does not have a rigid schema** or extensive table relationships.



## DOCUMENT

JSON

```
{  
  id: 1898,  
  gid: "tutorialsdojo1898",  
  firstName: "Jose",  
  lastName: "Rizal",  
  profile: {  
    nationality: "Filipino",  
    country: "Philippines",  
    birthPlace: "Laguna"  
  }  
}
```



- A fast, scalable, highly available **MongoDB-compatible database** service.
- A **document**-oriented database program
- Cross-platform, NoSQL database
- Each **document** contains fields and values in **JSON format** with no rigid schema enforced



Amazon DocumentDB

DOCUMENT DATABASE COLLECTION		
JSON Document #1	Field #1	Value #1
	Field #2	Value #2
	Field #3	Value #3
JSON Document #2	Field #1	Value #1
	Field #2	Value #2
	Field #3	Value #3
JSON Document #3	Field #1	Value #1
	Field #2	Value #2
	Field #3	Value #3

RELATIONAL DATABASE		
TABLE #1	Column #1	Column #2
Row #1	Column #1	Column #2
Row #2	Column #1	Column #2
Row #3	Column #1	Column #2
Row #4	Column #1	Column #2

- A **caching service**
- Allows you to set up, run, and scale **open-source in-memory databases** like:



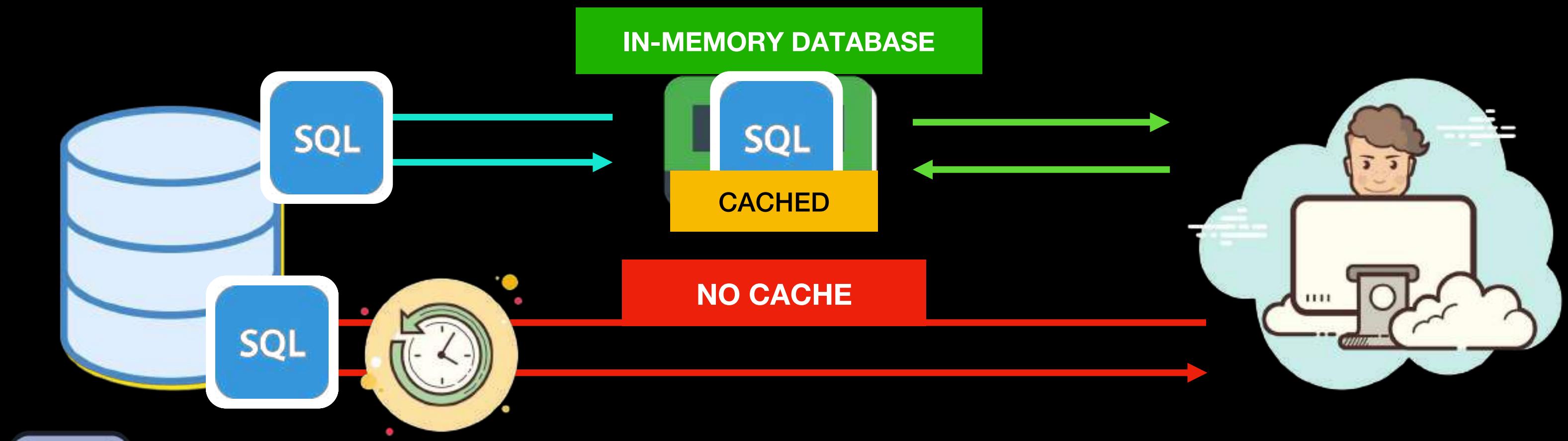
emcached



redis

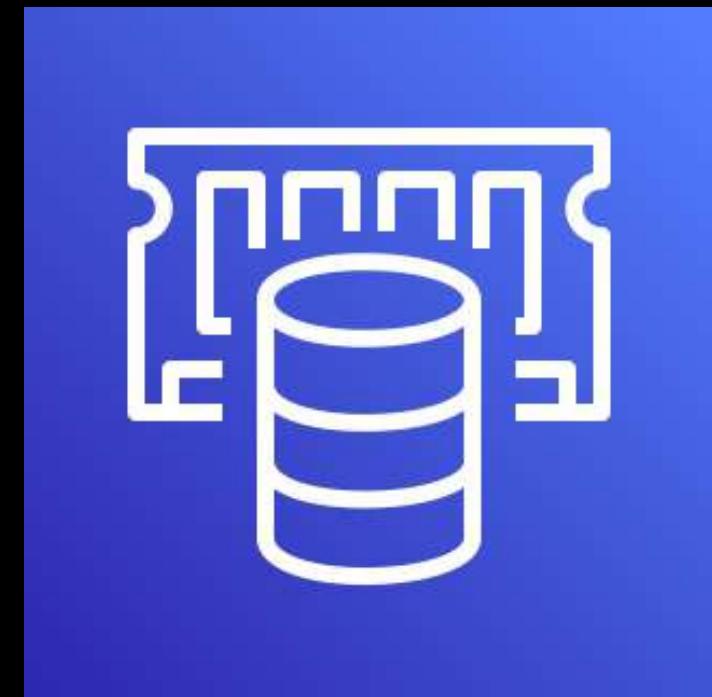


## Amazon ElastiCache



**Faster** than disk-based databases

- Useful for **database caching** that eliminates unnecessary frequent calls to the database just to **return identical datasets**
- Useful for real-time analytics, distributed session management, geospatial services, and many more



## Amazon ElastiCache



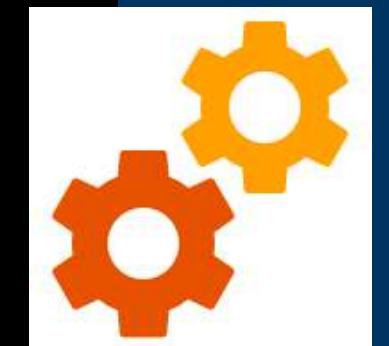
memcached



Sub-millisecond latency



Data Partitioning



Can be integrated  
to your apps with  
minimal code change



redis



emcached



Amazon ElastiCache for  
Memcached

- Based on the open-source **Memcached** in-memory data store.
- Suitable for building a simple, scalable caching layer for your data-intensive apps.
- **Multithreaded** — it can utilize multiple processing cores.
- **Lacks data replication capability**
  - **Does not:**
    - ✖ Support Advanced Data Structures
    - ✖ Provide Highly Available Caching Layer



stands for

REmote DIctionary Server



Amazon ElastiCache for  
Redis

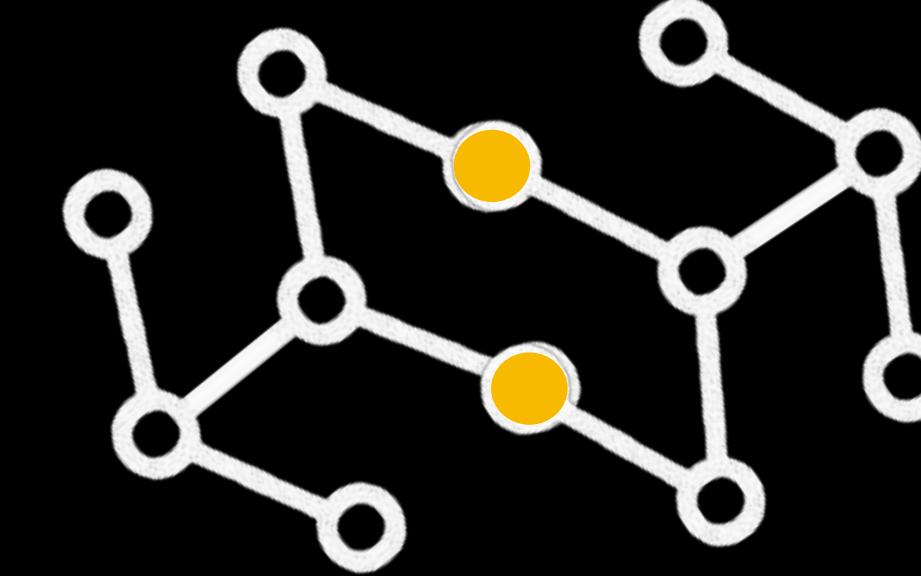
- Based on the open-source **Redis** in-memory data store.
- Provides:
  - Advanced Data Structures
  - Pub/Sub messaging
  - Geospatial support
  - Point-in-Time Snapshot support
- Has a **replication feature** that provides high availability via data replication.
- You can enable the **Cluster Mode** in Redis to have multiple primary nodes and replicas across **two or more Availability Zones**.



## Amazon KeySpaces

- A scalable, highly available, and managed **Apache Cassandra**-compatible database service
- An open-source, **wide column data store** that is designed to handle large amounts of data.
- **Run your Cassandra workloads** on AWS without having to provision, patch, or manage servers.





## Amazon Neptune

- A fast, reliable, fully-managed **graph database** service
- Makes it easy for you to build and run applications that work with **highly connected datasets**
- Allows you to **store billions of relationships** and query your data graphs with **milliseconds latency**.
- **Uses nodes to store data entities and edges** to store relationships between entities.



Time Series

9 AM



10 AM



11 AM



12 PM



Amazon Timestream

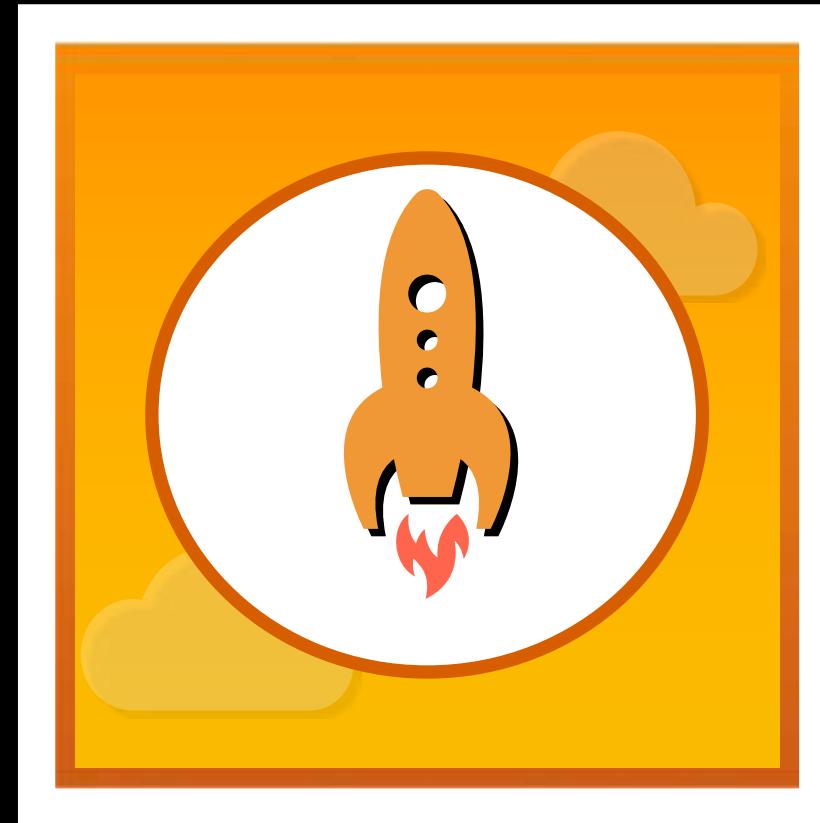
- A fast, scalable, and **serverless time series database** service
- Primarily used **for Internet-of-Things and operational applications.**
- **Track the changes of your data**
- Can be used to track stock prices, temperature measurements, and the CPU utilization of an EC2 instance over a specific amount of time.



## Amazon Quantum Ledger (Amazon QLDB)

- A fully managed **ledger database** service.
- Provides a transparent and **immutable transaction log** that is owned by a central trusted authority.
- Creates logs that are **cryptographically verifiable**
- Provide an **auditable history of all changes** made to your application data.
- Can be used to track each and every application data change.





# AWS Deployment Services Overview

---

## DEFINITION FILE

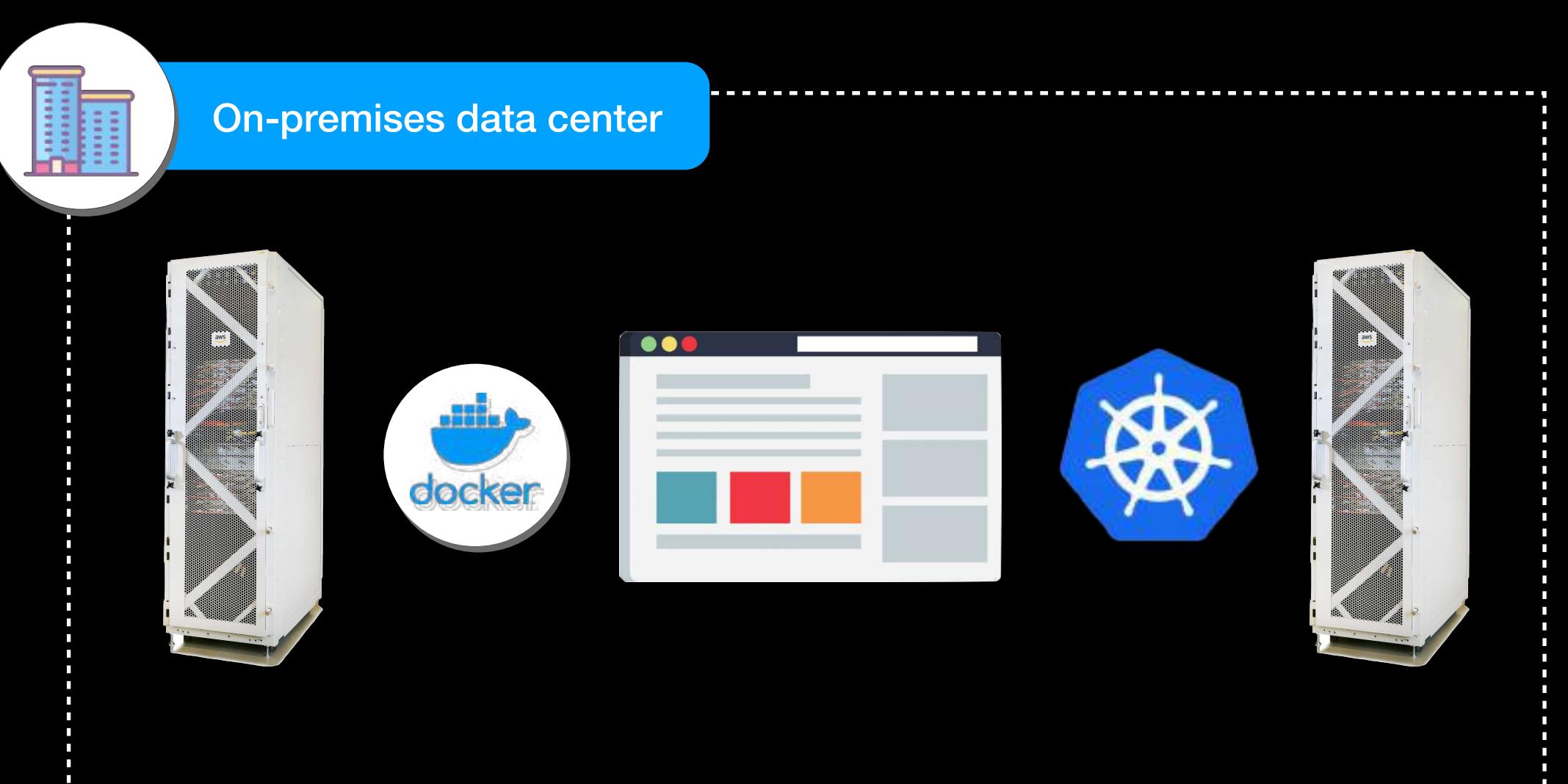
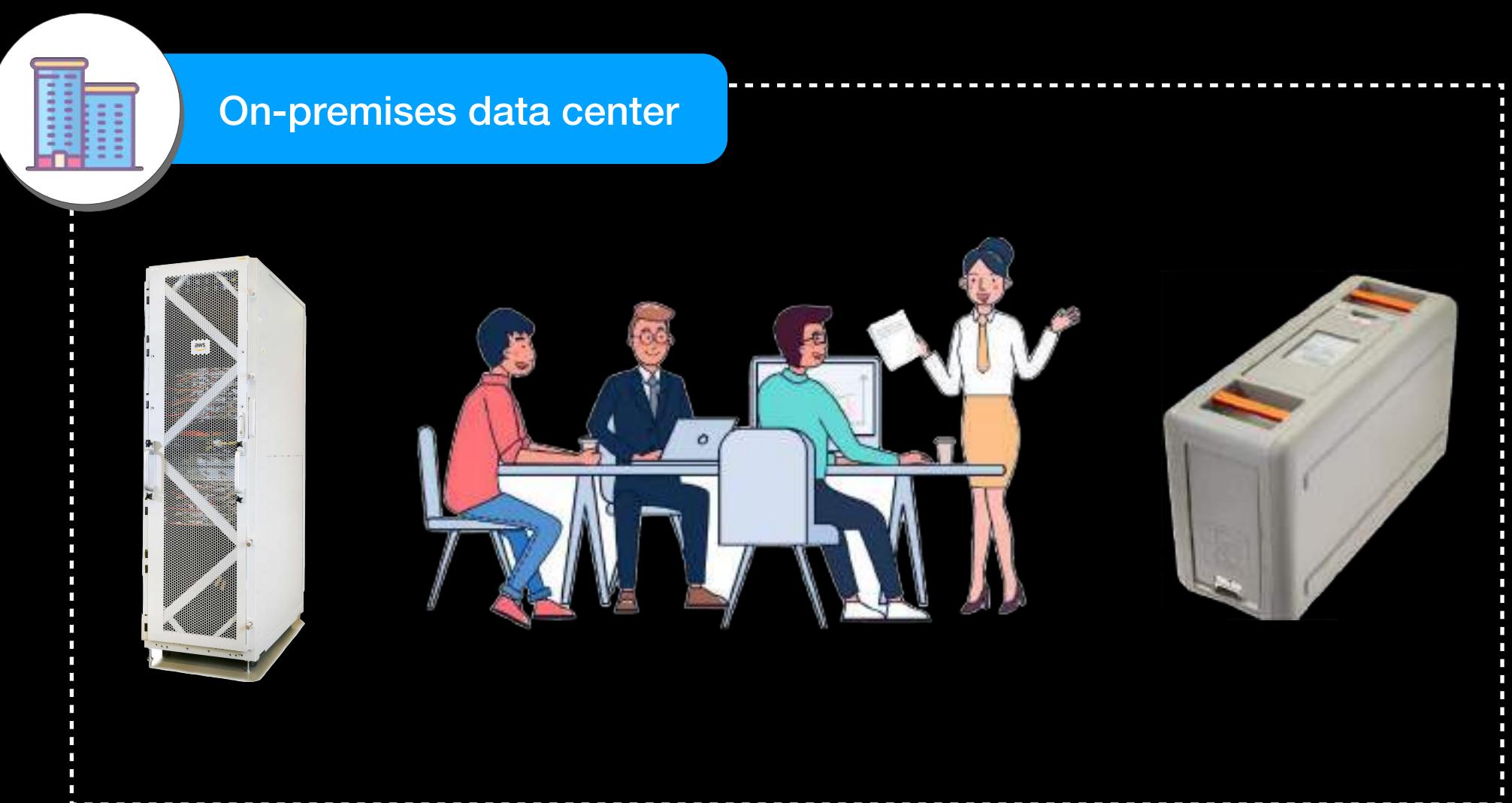
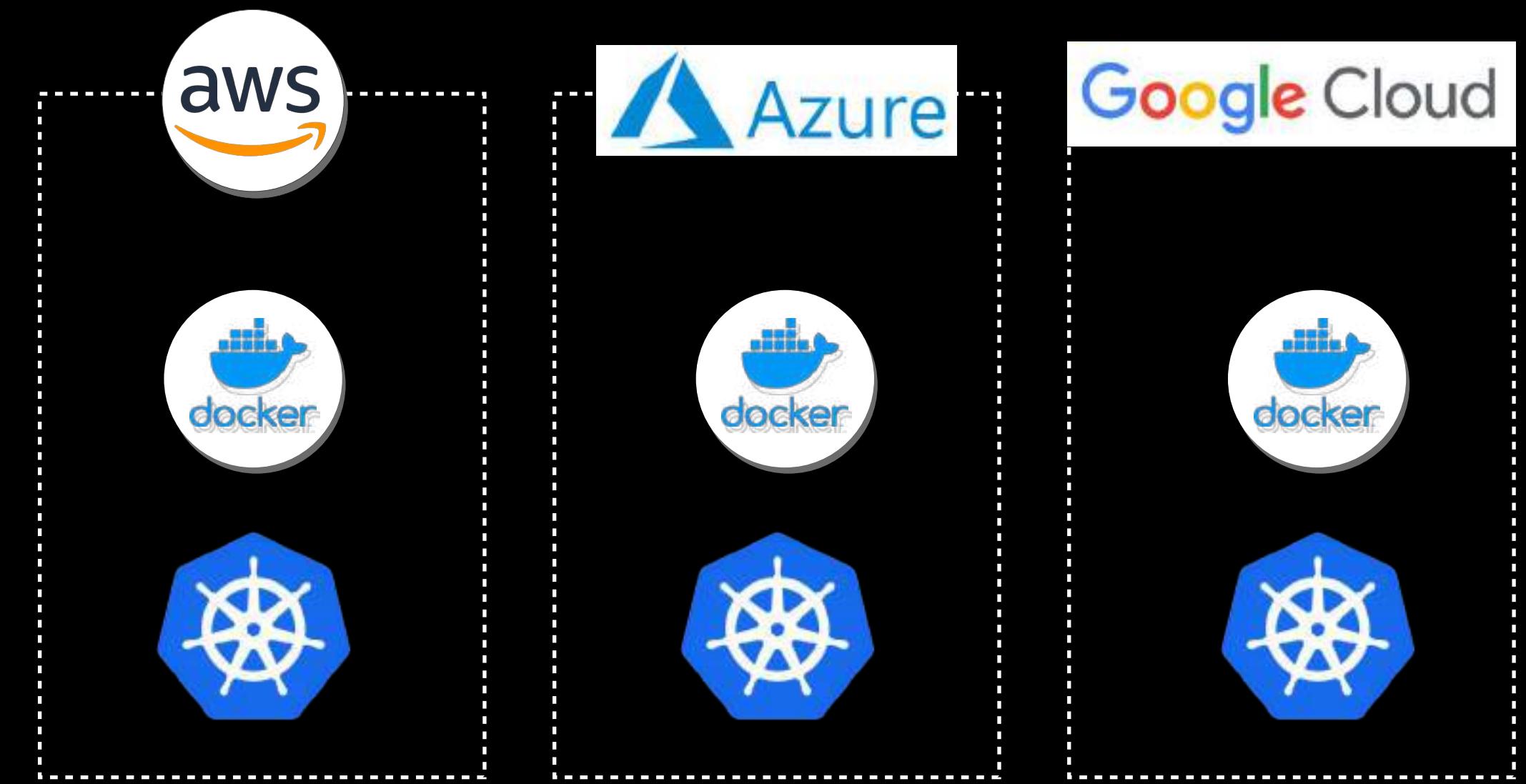
# Infrastructure as Code (IaC)

```
1 {
2     "AWSTemplateFormatVersion": "2010-09-09",
3
4     "Description": "Tutorials Dojo AWS CloudFormation Sample Template",
5
6     "Resources": {
7         "EC2Instance": {
8             "Type": "AWS::EC2::Instance",
9             "Properties": {
10                 "InstanceType": { "Ref": "InstanceType" },
11                 "SecurityGroups": [ { "Ref": "InstanceSecurityGroup" } ],
12                 "KeyName": { "Ref": "KeyName" },
13             }
14         },
15
16         "InstanceSecurityGroup": {
17             "Type": "AWS::EC2::SecurityGroup",
18             "Properties": {
19                 "GroupDescription": "Enable SSH access to Jon Bonso EC2 Instance",
20                 "SecurityGroupIngress": [
21                     { "IpProtocol": "tcp",
22                     "FromPort": "22",
23                     "ToPort": "22", "CidrIp": { "Ref": "SSHLocation" } }
24                 ]
25             }
26         },
27
28         "IPAssoc": {
29             "Type": "AWS::EC2::EIPAssociation",
30             "Properties": {
31                 "InstanceId": { "Ref": "EC2Instance" },
32                 "EIP": { "Ref": "IPAddress" }
33             }
34         }
35     }
36 }
37 }
```

## Hybrid



## Multi-Cloud





# AWS Deployment Services



AWS CloudFormation



AWS Elastic Beanstalk



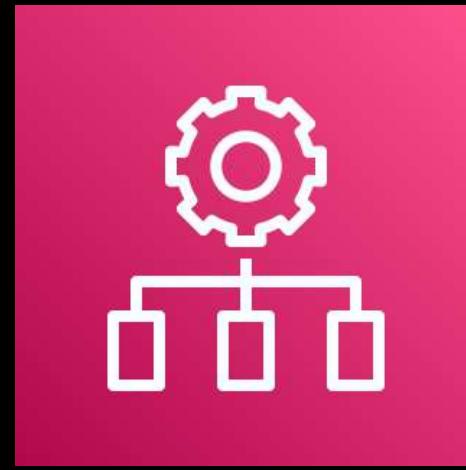
AWS CodeDeploy



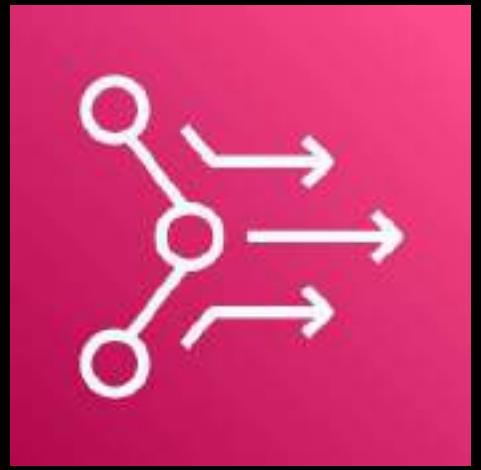
Amazon ECS  
Anywhere



Amazon EKS  
Anywhere



AWS OpsWorks



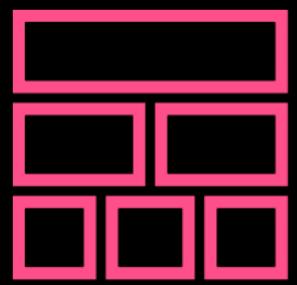
AWS Proton



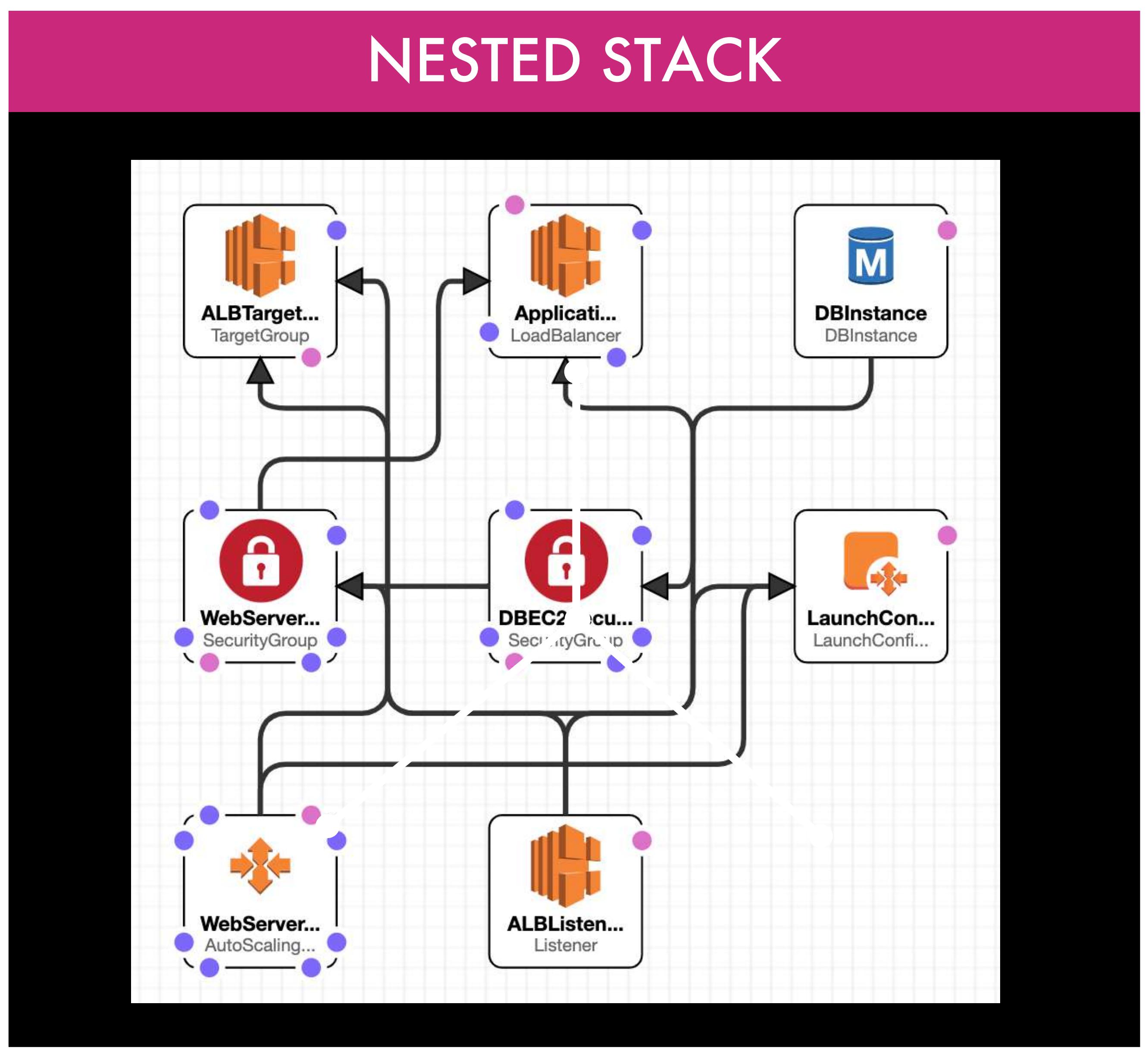
## AWS CloudFormation

- Provisions and manages your AWS resources using a custom code template in JSON or YAML format
- Has a built-in graphical drag-n-drop online tool called **CloudFormation designer**
- Primary Infrastructure as Code (IaC) service in AWS
- Provides different features such as **Nested Stacks**, **Change Sets**, **StackSets** and others

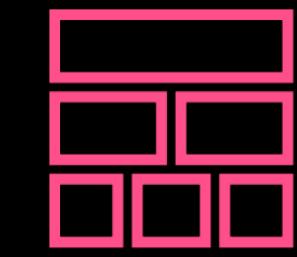
## ROOT STACK



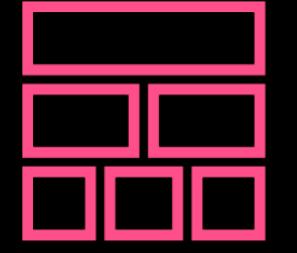
## STACK



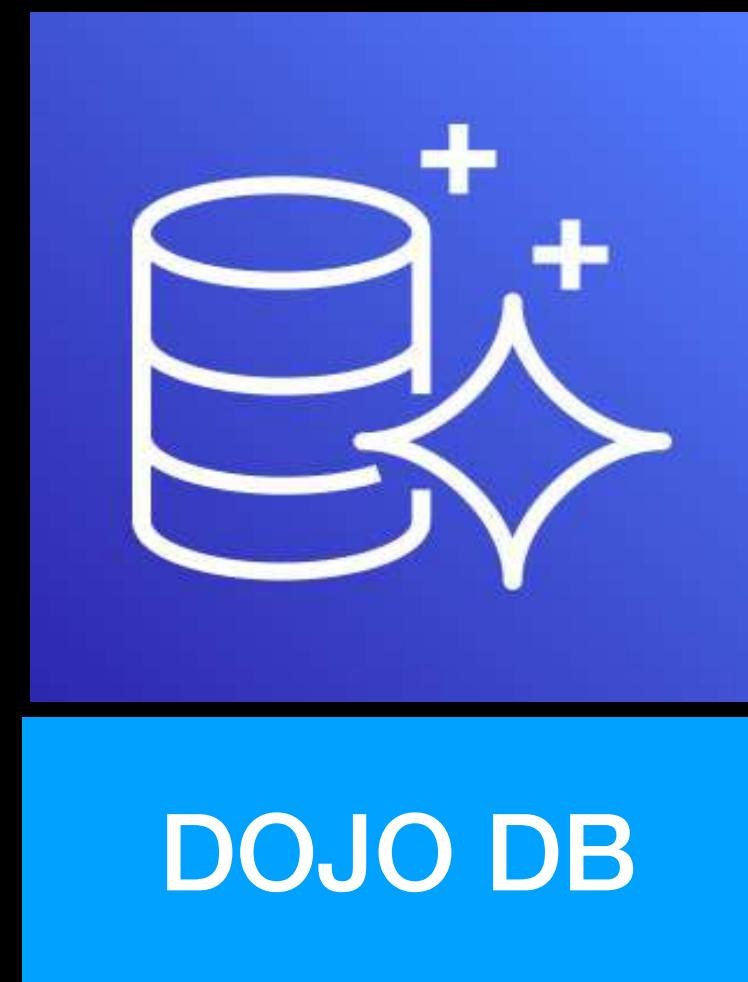
## DATABASE STACK



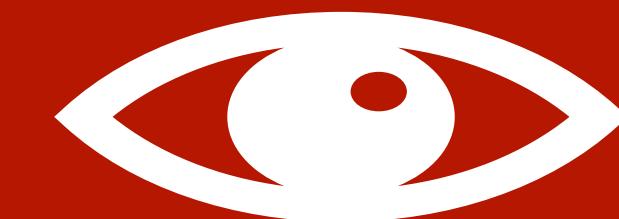
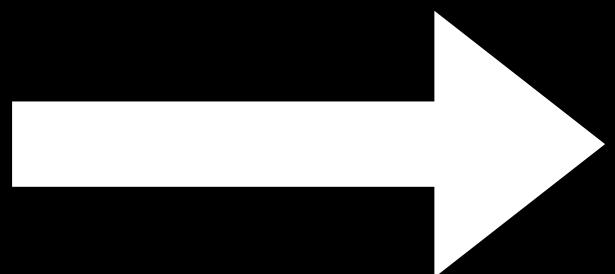
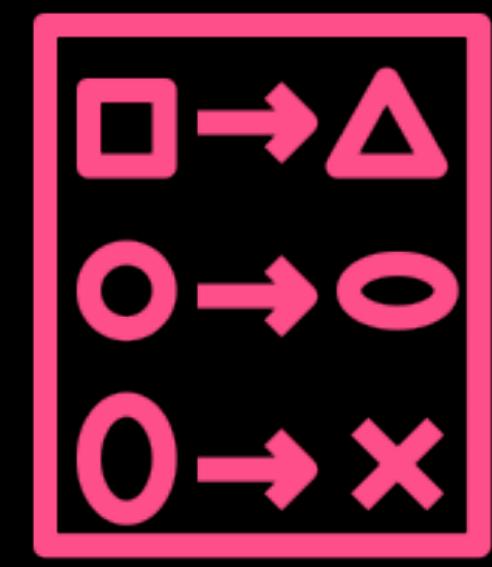
## APPLICATION STACK



**CHANGE SET**

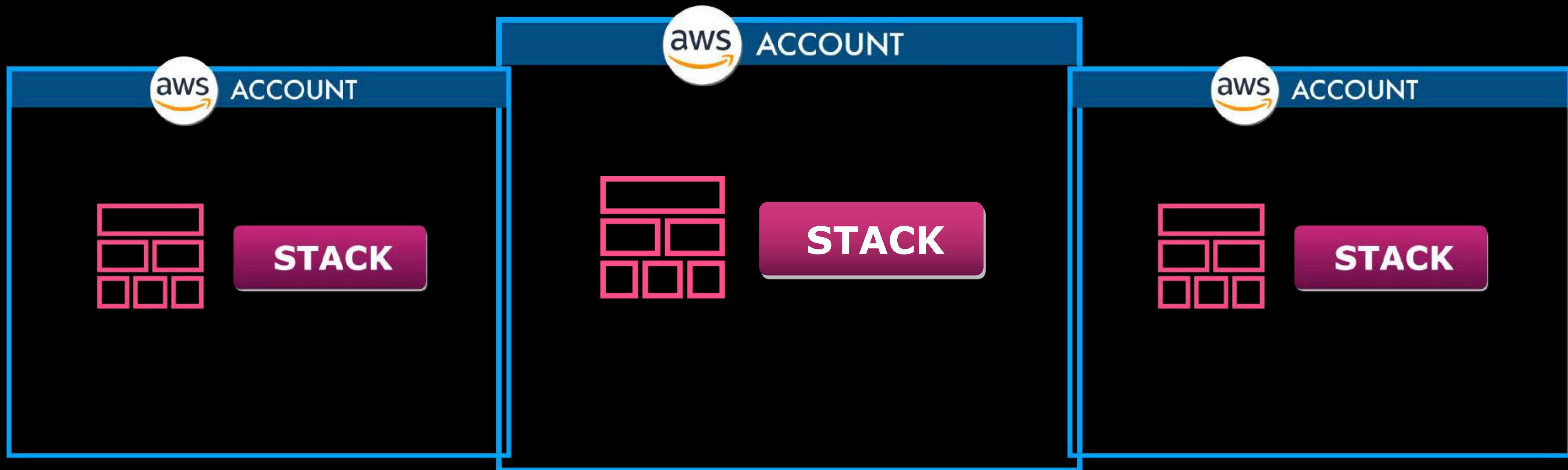


**CHANGE**



**PROVIDES A PREVIEW BEFORE  
THE ACTUAL CHANGE**

# STACKSET





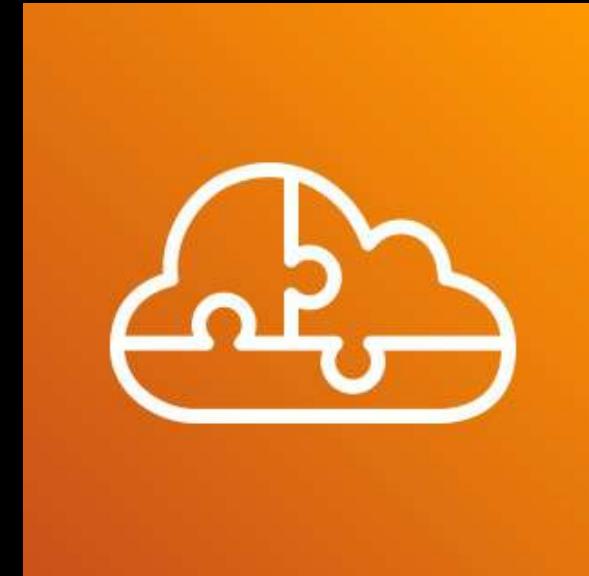
**AWS Cloud Development Kit**  
(AWS CDK)



AWS CloudFormation



**AWS Serverless Application Model**  
(AWS SAM)



**AWS Serverless Application Repository**

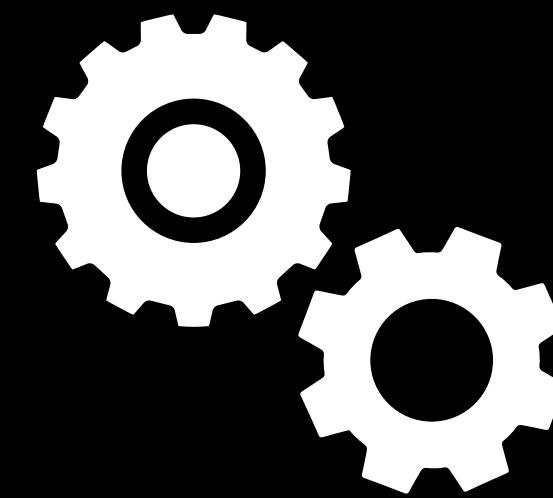
- Allows you to upload your application code in AWS and provision the required cloud environment easily
- Automatically deploys the necessary AWS resources and components to run your application
- Environment Tiers:



**AWS Elastic  
Beanstalk**



**Web Server**



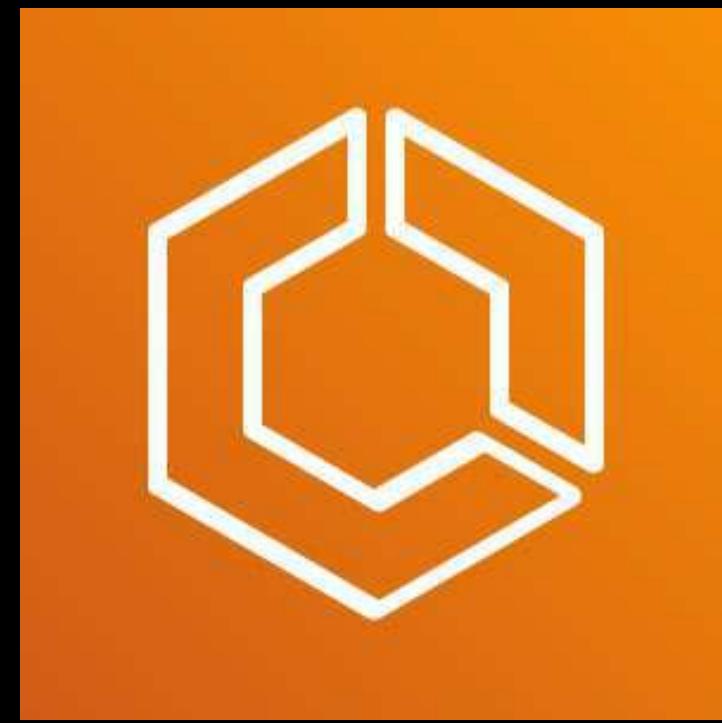
**Worker**

- Uses a configuration file to automatically deploy and configure your applications. All configuration files are stored in the **.ebextensions** folder



## AWS CodeDeploy

- A fully managed **deployment service**
- Automates your application deployments to Amazon EC2 instances, Amazon ECS clusters, AWS Lambda functions, and other computing services in AWS
- **Capable of doing hybrid deployment** of your applications to your on-premises data center and to AWS
- Does NOT create or provision AWS resources, unlike the AWS CloudFormation service



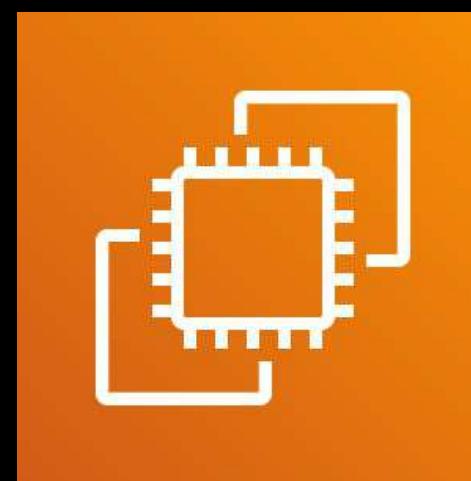
## Amazon ECS

- A **container orchestration service** that supports Docker containers
- Automates the process of installing, operating, managing, networking and scaling your cluster management infrastructure in AWS





**Amazon VPC**



Amazon EC2  
Instances



**AWS Fargate**

Serverless



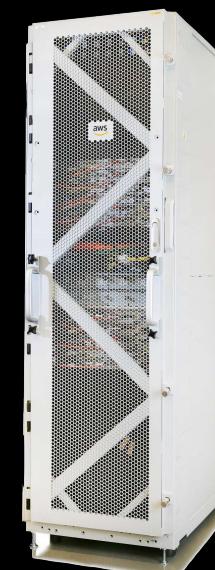
internally powered by:



**Amazon ECS Anywhere**



On-premises data center



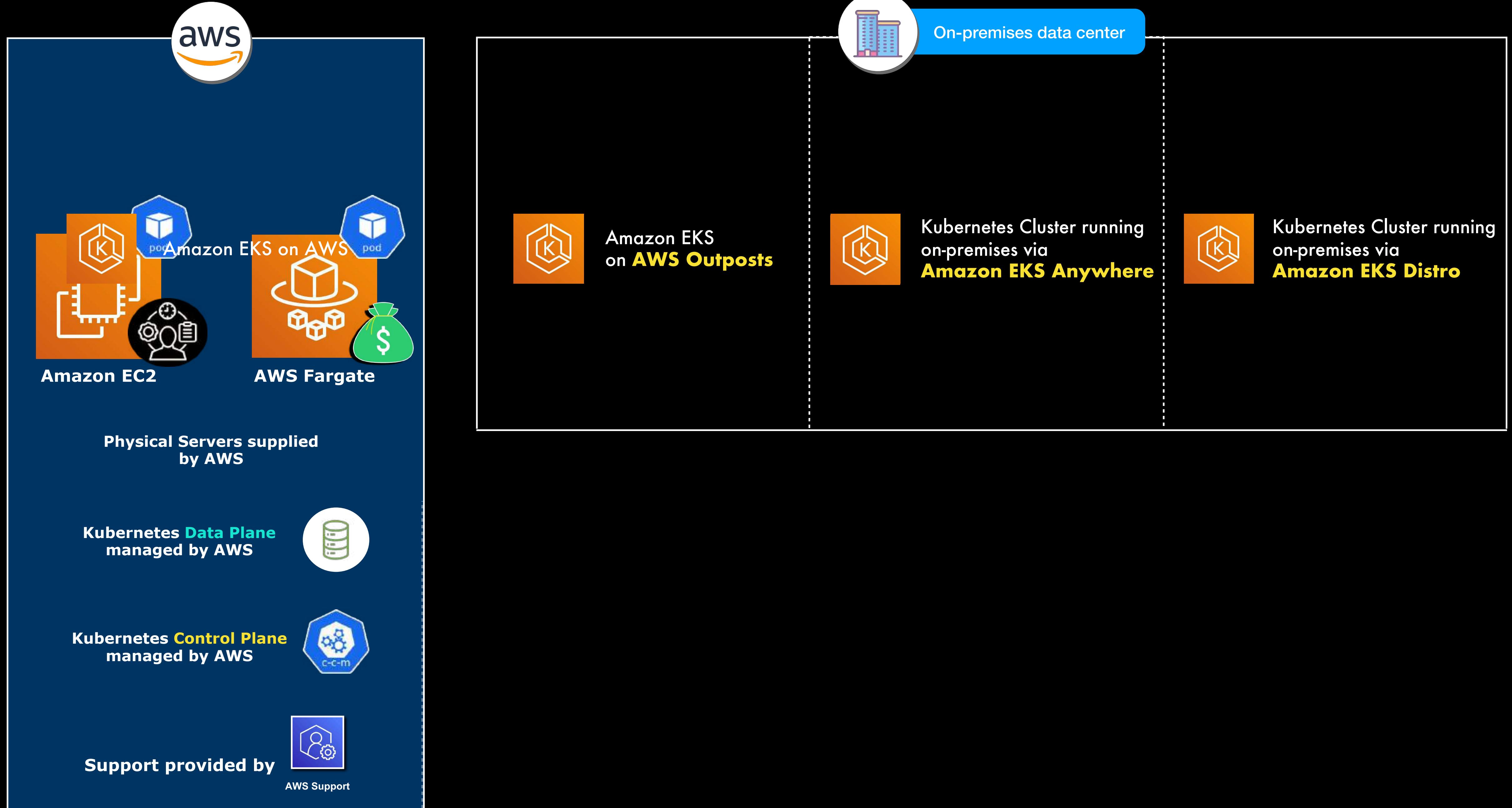
**Amazon CloudWatch Container Insights**



## Amazon EKS

- A **managed orchestration service** that supports Kubernetes containers
- Automates the process of installing, operating, managing, networking and scaling your Kubernetes control plane, pods and nodes in AWS







Amazon EKS on AWS



Amazon EC2



AWS Fargate

Physical Servers supplied  
by AWS



Kubernetes Data Plane  
managed by AWS

Kubernetes Control Plane  
managed by AWS



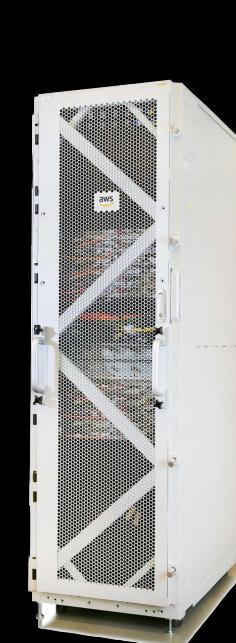
Support provided by



AWS Support



Amazon EKS  
on **AWS Outposts**



AWS Outposts

Physical Rack Server supplied by  
AWS but managed by you



Kubernetes Data Plane  
managed by you

Kubernetes Control Plane  
managed by AWS



Support provided by



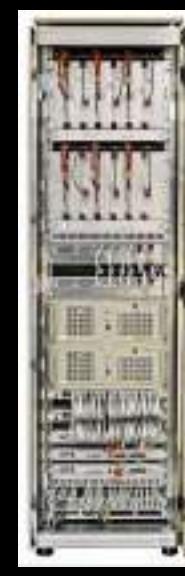
AWS Support



On-premises data center



Kubernetes Cluster running  
on-premises via  
**Amazon EKS Anywhere**



Amazon EKS Anywhere

Physical Server supplied  
and managed by you



Kubernetes Data Plane  
managed by you

Kubernetes Control Plane  
managed by you



Support provided by



AWS Support



Kubernetes Cluster running  
on-premises via  
**Amazon EKS Distro**



Amazon EKS Distro

Physical Server supplied  
and managed by you



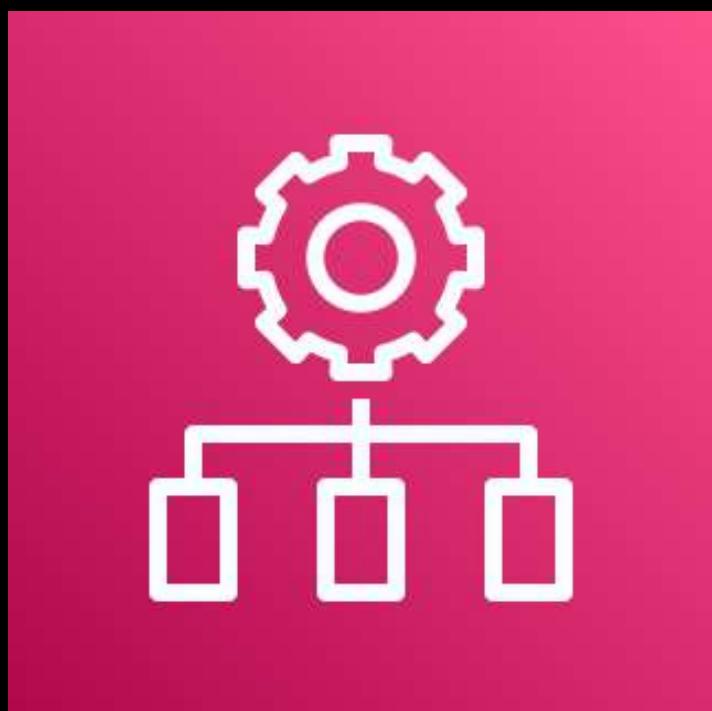
Kubernetes Data Plane  
managed by you



Kubernetes Control Plane  
managed by you

No AWS Support

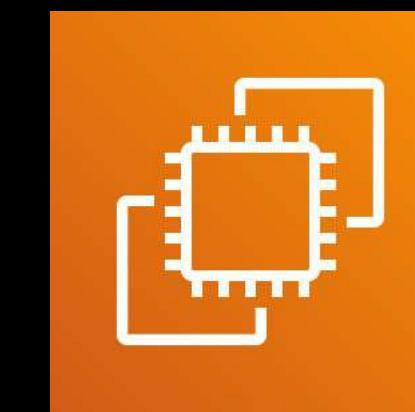
- A **configuration management service**
- Provides managed instances for your automation platforms based on:



**AWS OpsWorks**



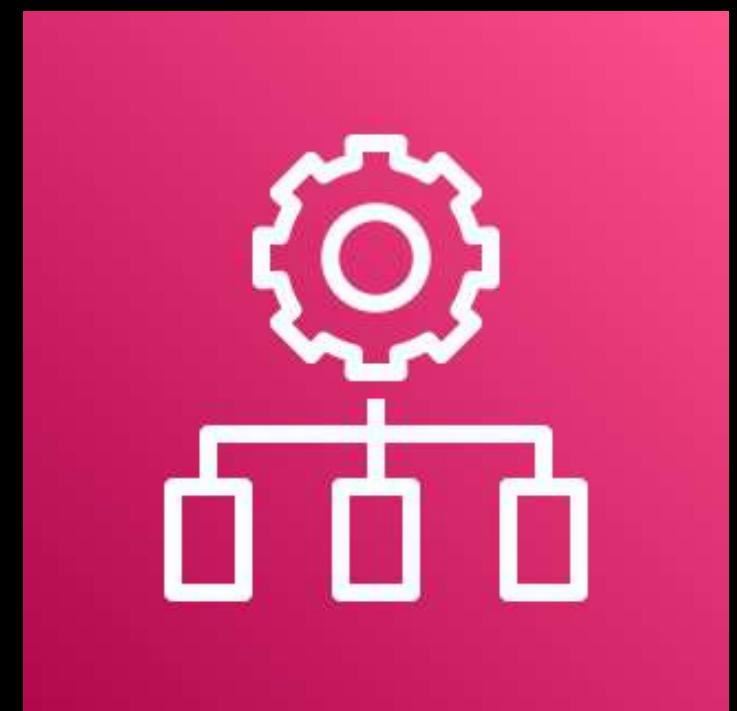
- Automates how your servers are provisioned, configured, and managed across:



**Amazon EC2  
Instances**



**On-premises  
Servers**



**AWS OpsWorks**



**AWS OpsWorks for **Chef Automate****

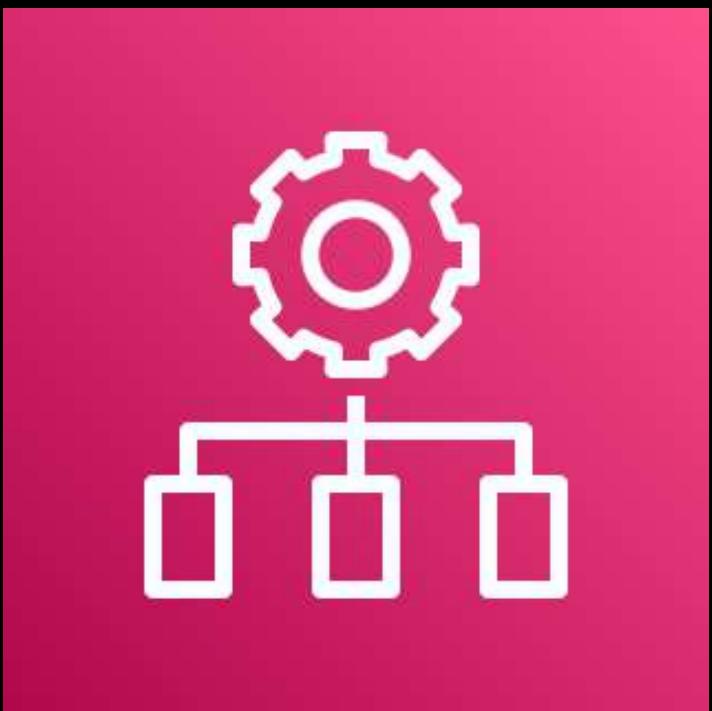


**AWS OpsWorks for **Puppet Enterprise****



**AWS OpsWorks **Stacks****

- A service that **automates container & serverless deployment**
- Ensures that you have consistent development standards and best practices across your AWS account
- **Deploys container and serverless applications using pre-approved stacks** that your platform team manages.
- Grants developers the freedom to innovate but still within the set guardrails that the security team implemented
- Offers a **self-service portal** for your developers
- Provides **AWS Proton template** which contains all the information required to deploy your custom environments and services



**AWS Proton**

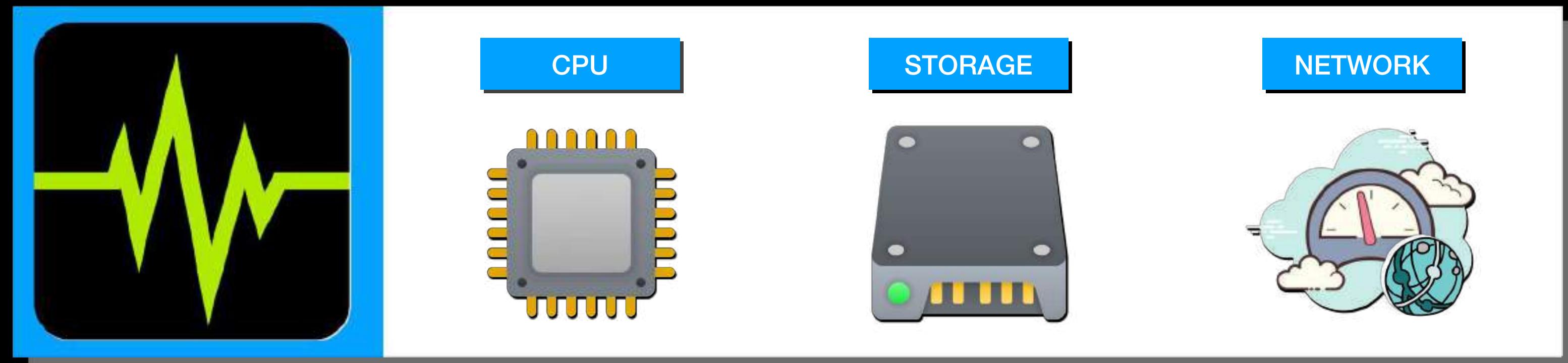
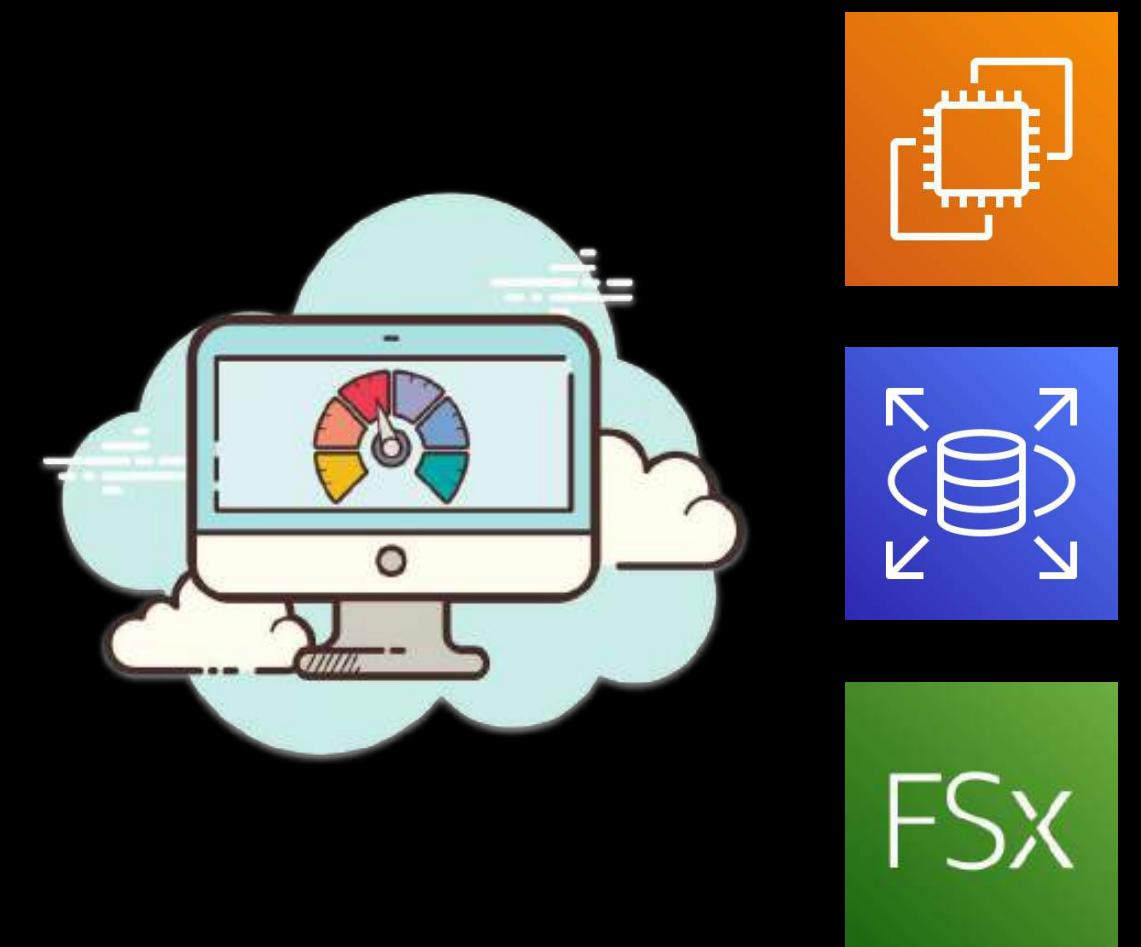
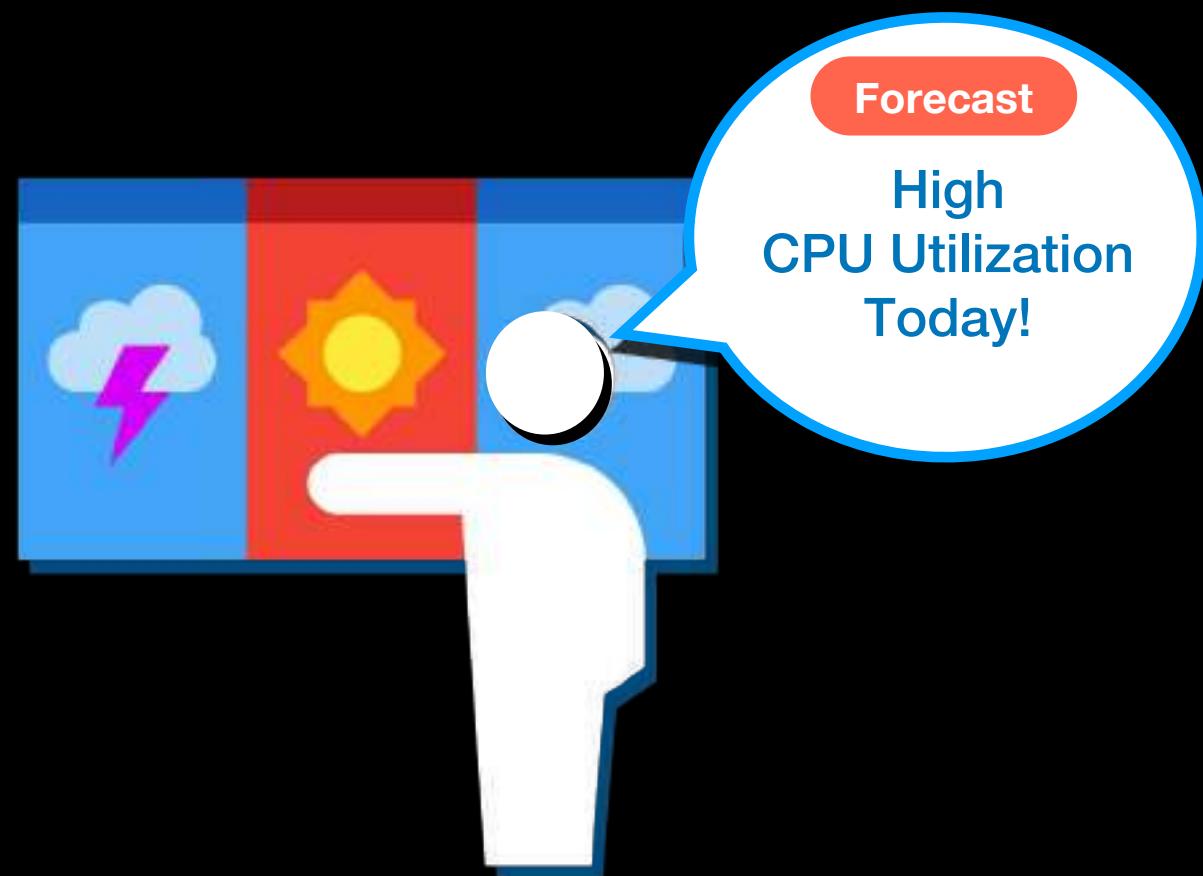


# AWS Monitoring Services Overview

---



# AWS Monitoring Services





AWS Monitoring Services



Amazon CloudWatch



AWS Service Health Dashboard



AWS Personal Health Dashboard

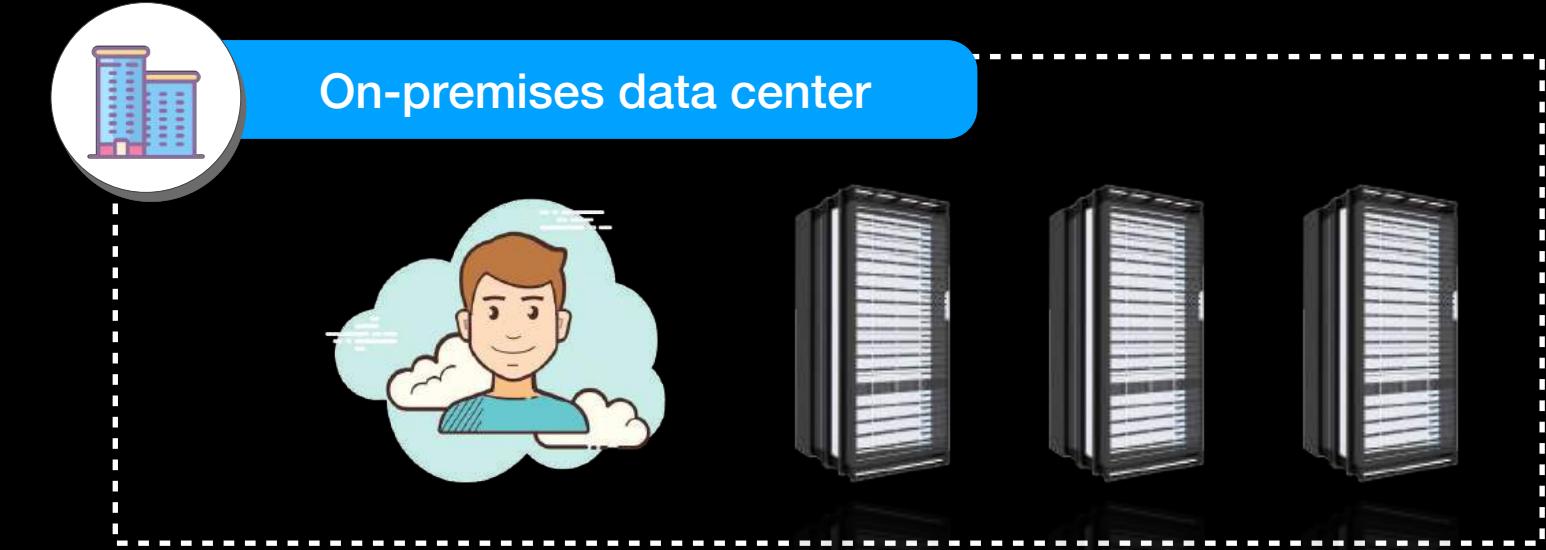


AWS Health API



## Amazon CloudWatch

- A **suite of AWS services** used in monitoring your systems on both:



- A **metrics repository** that collects system data from AWS services as well as your **custom metrics**
- **Monitors** and analyzes system metrics
- **Notifies** you if a certain threshold has been reached
- **Triggers an action** based on a **specific threshold or events** that you define



## Amazon CloudWatch



Metrics



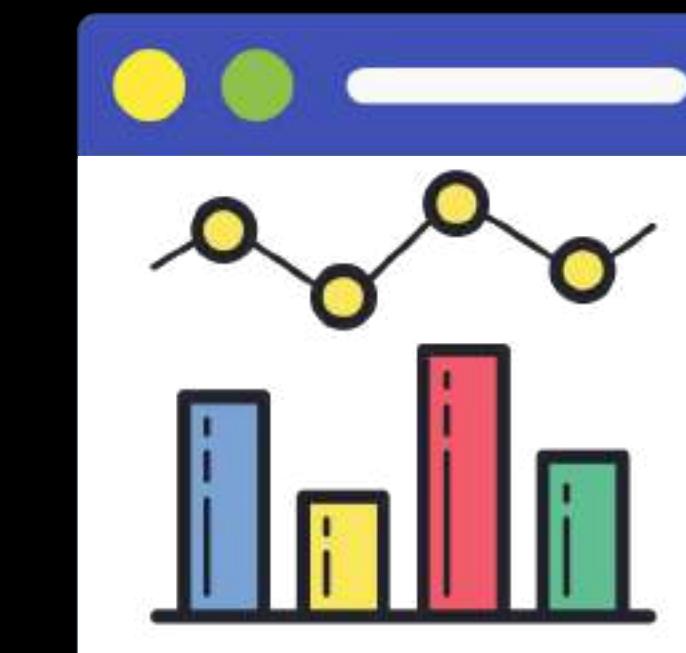
Logs



Alarms



Events



Dashboards



## Amazon CloudWatch METRICS

- **Collect metrics** from various AWS Services and your custom applications
- **Aggregate (combine) metrics** across multiple resources
- Most AWS services send metric data to CloudWatch every **1 minute** by default
  - For Amazon EC2, the default frequency is every **5 minutes**
  - **Detailed Monitoring** sends EC2 metrics data every **1 minute**

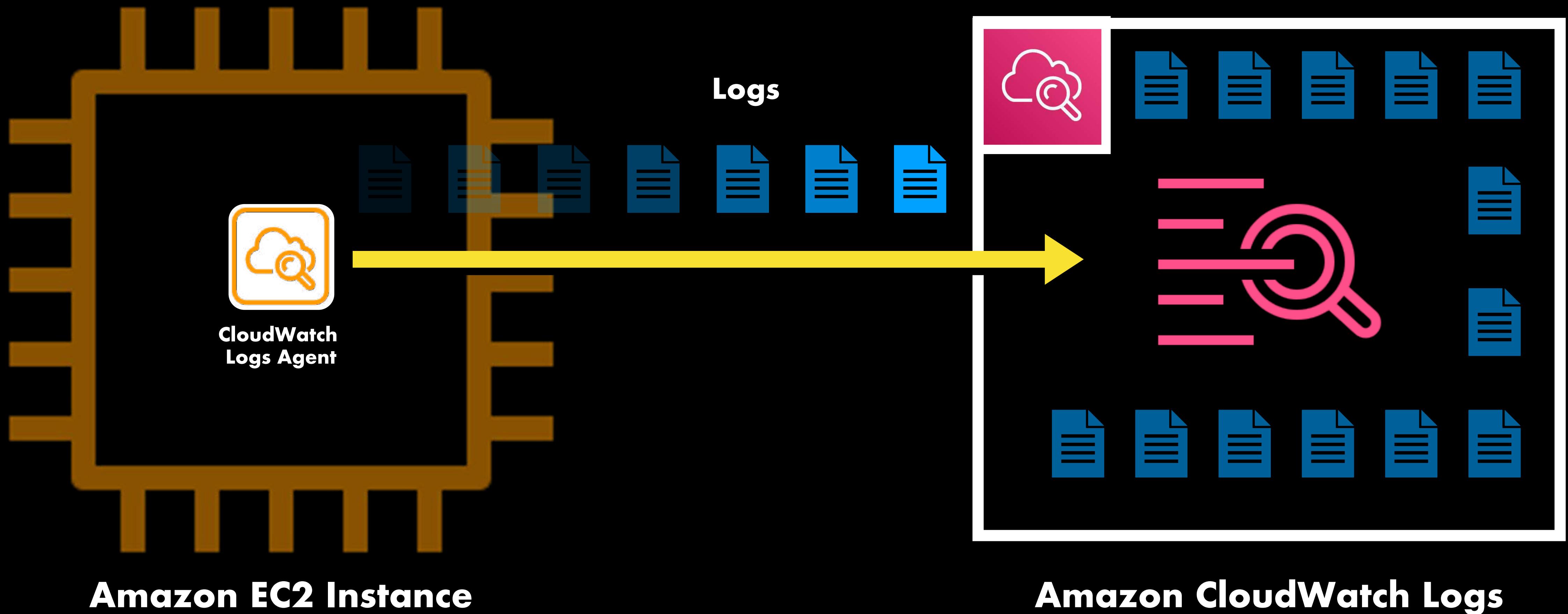




## Amazon CloudWatch Logs

- Primarily used for **logs** monitoring
- Allows you to **monitor, store, access, analyze or query the logs** from your AWS resources or from your custom applications
- Install **CloudWatch Logs agent** to your EC2 instances to automatically collect and publish your application logs to CloudWatch





- Allows you to create **alarms** for your monitoring
- Performs one or more actions based on a system metric and a **specific threshold**
- Can **notify you or other systems/services** using Amazon SNS
- Can **trigger a custom action**, such as:
  - Auto Scaling your EC2 instances
  - Sending a billing alert
  - Invoking a Lambda function
  - ... and many more!



**Amazon CloudWatch**

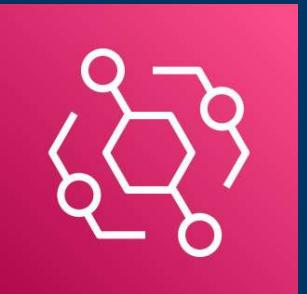
**ALARMS**



CloudWatch Events and Amazon EventBridge

have the **same underlying service and API**,

but the latter provides more features.



Amazon EventBridge

- **Monitors and responds to the system/service events** of your AWS resource in near real-time



## Amazon CloudWatch EVENTS

- Allows you to **create a CloudWatch Event rule** to track the changes or the state of your services
- **Invokes a certain action** if a specific event matched your Event rule
- Allows you to create a **scheduled job** that invokes a Lambda function on a regular basis, like every hour, every day, every week, or any schedule that you like.





## Amazon CloudWatch

### DASHBOARDS

- A customizable **dashboard** containing your AWS system metrics
- **Monitor your resources in a single view**, even if those resources are located [across different AWS Regions](#)
- Allows you to publish and view your **custom metrics**





## AWS Service Health Dashboard

### REGIONS

### SERVICE STATUS

Service Status	
Region	Status
Alexa for Business (N. Virginia)	Service is operating normally
Amazon API Gateway (Montreal)	Service is operating normally
Amazon API Gateway (N. California)	Service is operating normally
Amazon API Gateway (N. Virginia)	Service is operating normally
Amazon API Gateway (Ohio)	Service is operating normally
Amazon API Gateway (Oregon)	Service is operating normally
Amazon AppFlow (Montreal)	Service is operating normally
Amazon AppFlow (N. California)	Service is operating normally
Amazon AppFlow (N. Virginia)	Service is operating normally
Amazon AppFlow (Ohio)	Service is operating normally
Amazon AppFlow (Oregon)	Service is operating normally
Amazon AppStream 2.0 (N. Virginia)	Service is operating normally
Amazon AppStream 2.0 (Oregon)	Service is operating normally
Amazon Athena (Montreal)	Service is operating normally
Amazon Athena (N. California)	Service is operating normally
Amazon Athena (N. Virginia)	Service is operating normally
Amazon Athena (Ohio)	Service is operating normally
Amazon Athena (Oregon)	Service is operating normally
Amazon Augmented AI (Montreal)	Service is operating normally
Amazon Augmented AI (N. Virginia)	Service is operating normally
Amazon Augmented AI (Ohio)	Service is operating normally
Amazon Augmented AI (Oregon)	Service is operating normally
Amazon Braket (N. California)	Service is operating normally
Amazon Braket (N. Virginia)	Service is operating normally

Contact Us

RSS



## AWS Personal Health Dashboard

- A **personalized** dashboard that shows the **status of the AWS services that you are using**
- Does **NOT** show you the status of all the AWS services globally but only the status of the AWS services that you have in your account.
- **Shows the AWS Health events** that might affect your applications running on AWS such as **scheduled maintenance or system outages**
- **Allows you to create alerts and notifications** based on the health of your AWS resources



## AWS Health API

- Provides **programmatic access to the AWS Health** information that appears in your **AWS Personal Health Dashboard**
- A **RESTful web service** that you can access via HTTPS
- **NOT** available by default
- Only available in **Business or Enterprise support plans**

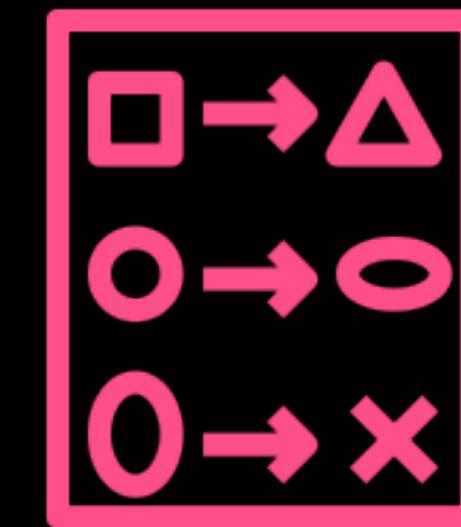


# AWS Audit & Compliance Services Overview

---



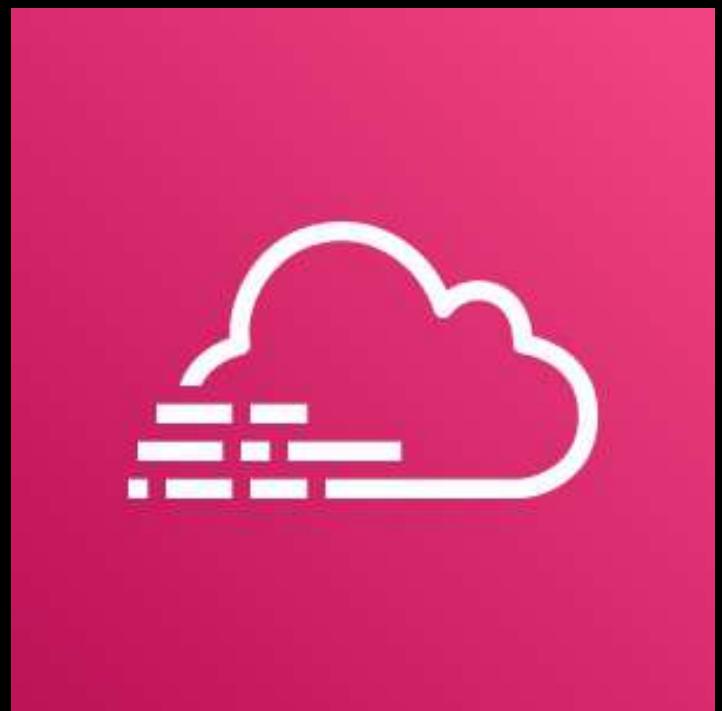
## AWS Audit & Compliance Services



RESOURCE CHANGES



## AWS Audit & Compliance Services



**AWS CloudTrail**



**AWS Artifact**



**AWS Security Hub**

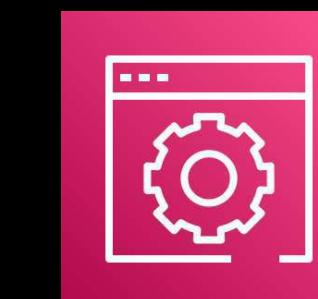


## AWS CloudTrail

- **Tracks user activity and API usage** in your AWS account
- **Stores the audit log data** in:
- **Enables risk auditing** by continuously monitoring and logging account activities, such as user actions:



Amazon S3 Bucket



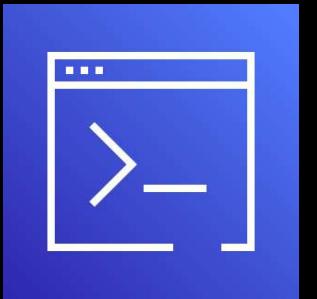
AWS Management  
Console



AWS SDK



AWS API



AWS Command Line  
Interface (CLI)



## MANAGEMENT EVENTS

### Control Plane



## AWS CloudTrail

- Attaching an IAM Role
- Creating a new VPC
- Creating a subnet



## DATA EVENTS

### Data Plane

Provide information about the **resource operations** performed **ON** (e.g. S3 **bucket**) your resources or performed **IN** (e.g. S3 **objects**) your resources

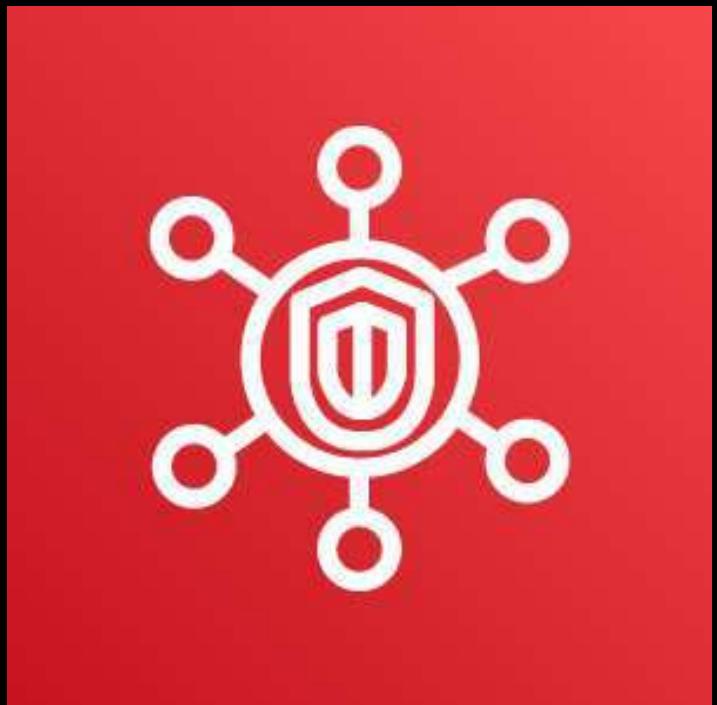
- Amazon S3 **object-level API** activities
- Invoking an AWS Lambda function

- Provides on-demand **AWS security and compliance reports**
- Acts as a self-service portal to find compliance-related information and reports for:
  - ISO Reports
  - Payment Card Industry (PCI) reports
  - Service Organization Control (SOC) reports
  - . . . and many more!
- Allows you to download AWS security and compliance documents such as **SOC 1 report, ISO certifications, and other reports**



**AWS Artifact**

- Provides a **centralized & comprehensive view of the security posture** of your cloud infrastructure across multiple AWS accounts



## AWS Security Hub



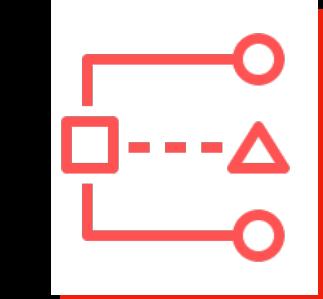
Amazon GuardDuty



Amazon Inspector



Amazon Macie



AWS IAM Access  
Analyzer



AWS Firewall  
Manager

- Helps you to comply with your company's specific security standards and best practices
- **Collects security alerts and findings** from:



# AWS Networking & Content Delivery Services

## Overview

---



# AWS Networking & Content Delivery Services



Amazon VPC



Elastic Load  
Balancing



Amazon  
Route 53



AWS  
Global Accelerator



Amazon  
CloudFront



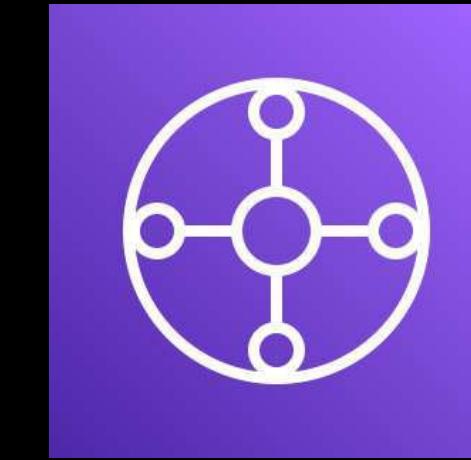
AWS PrivateLink



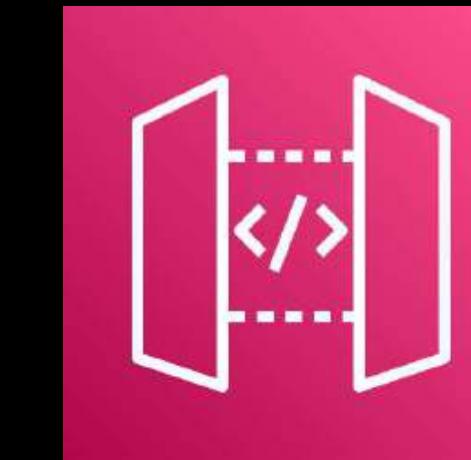
AWS VPN



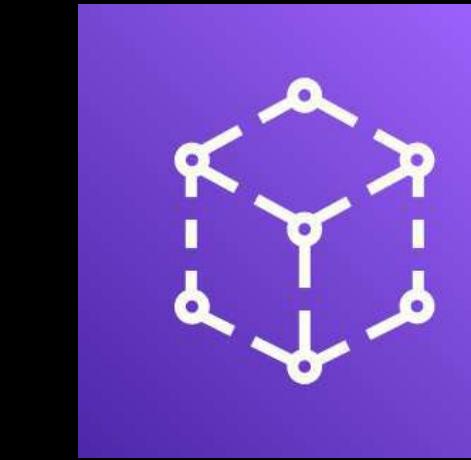
AWS Direct  
Connect



AWS  
Transit Gateway



Amazon  
API Gateway



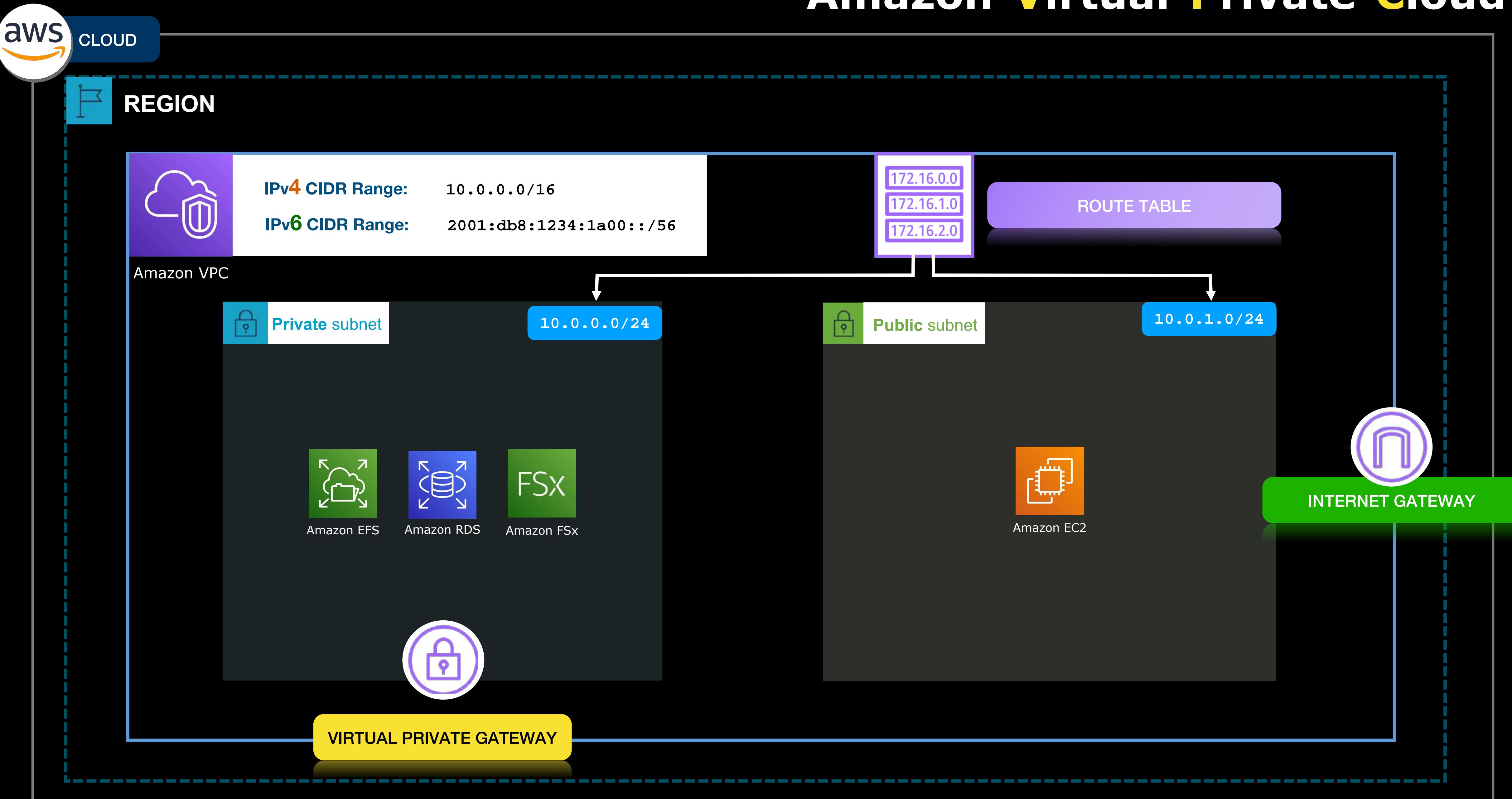
AWS App Mesh



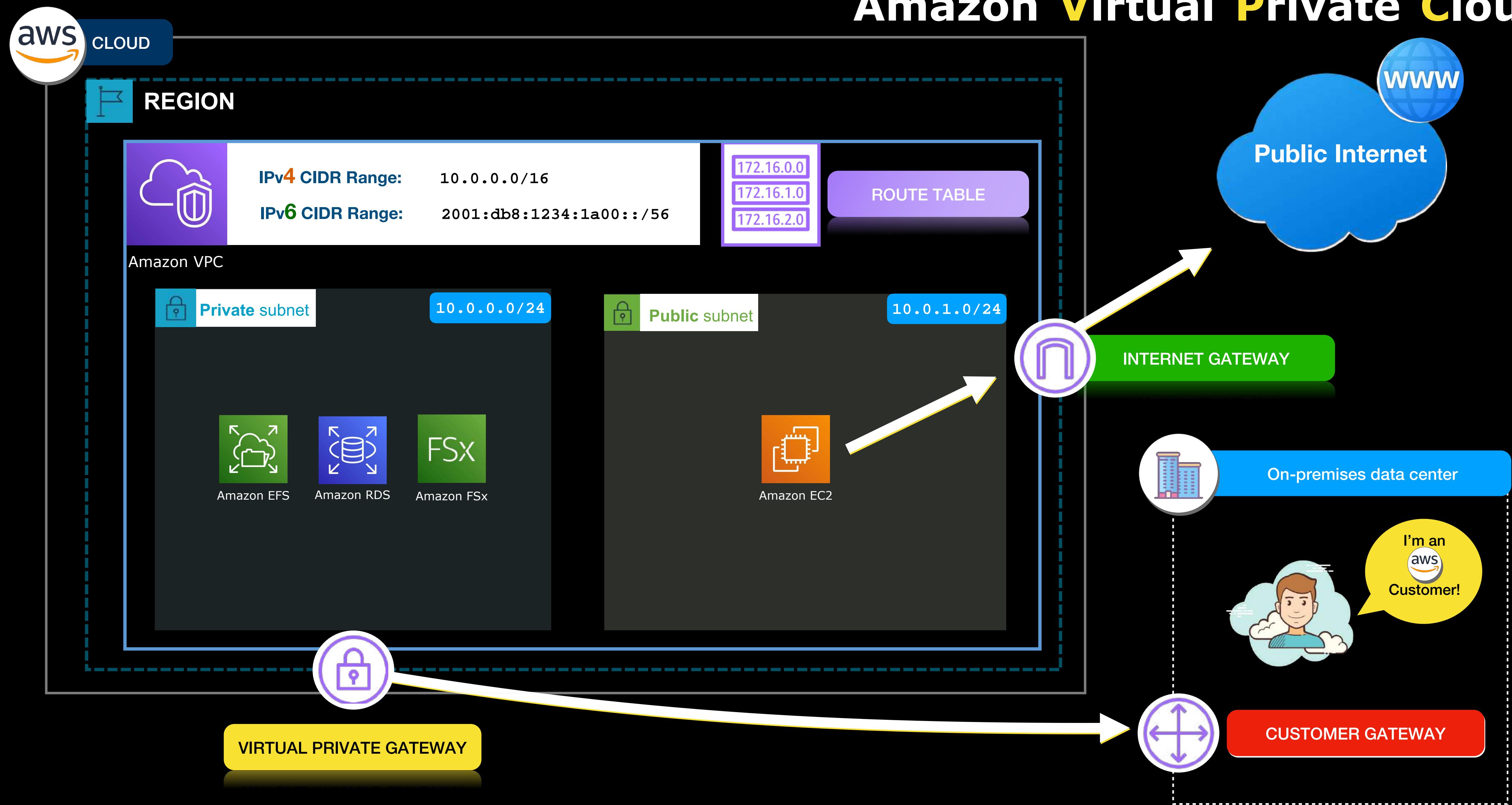
AWS Cloud Map

Also categorized as an  
**Application Integration Service**

# Amazon Virtual Private Cloud



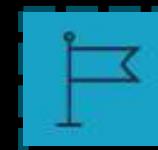
# Amazon Virtual Private Cloud



# Amazon Virtual Private Cloud

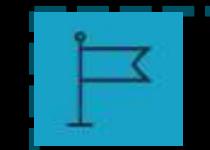
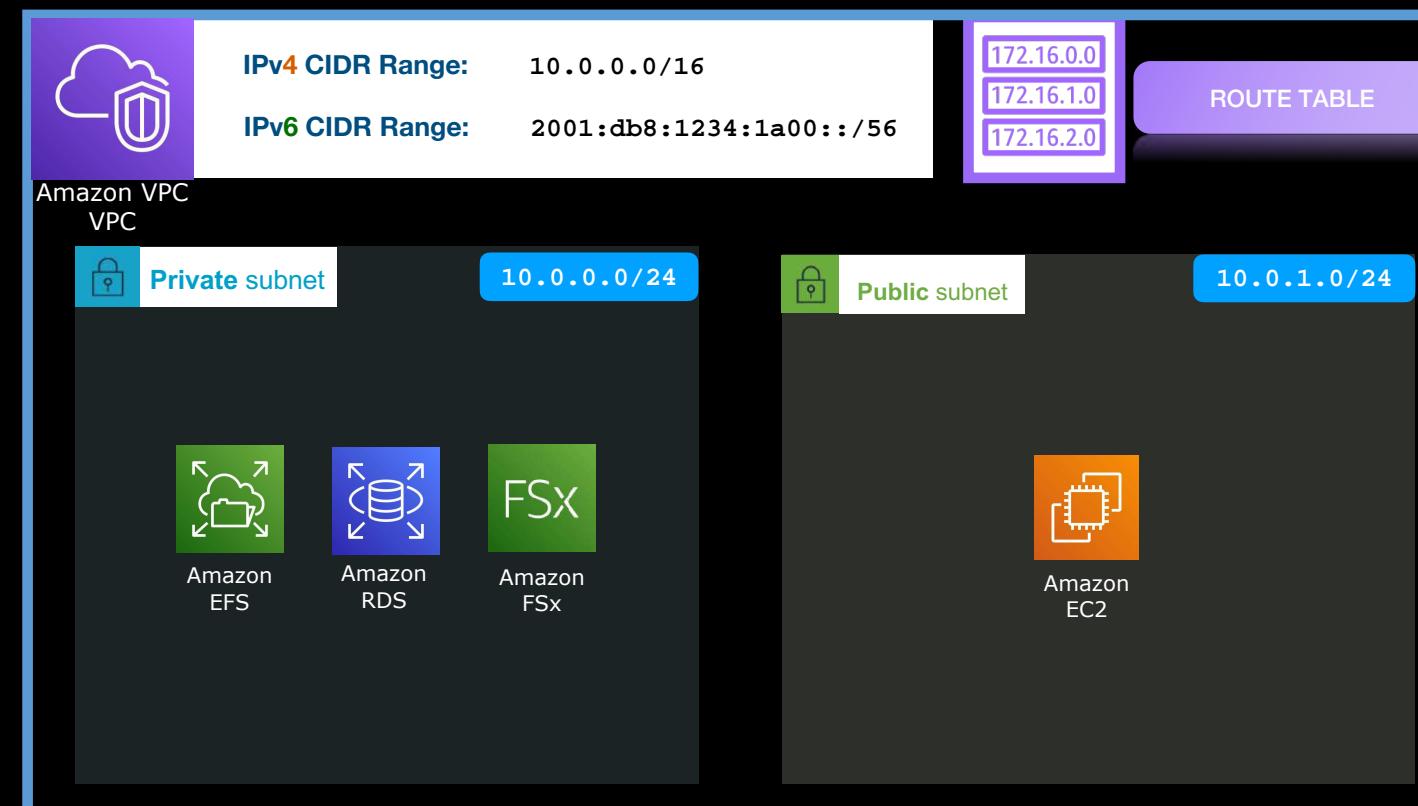


CLOUD



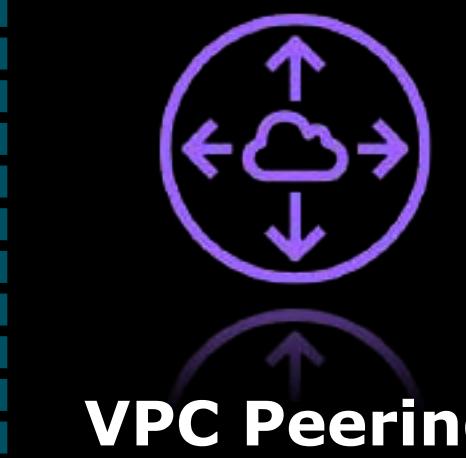
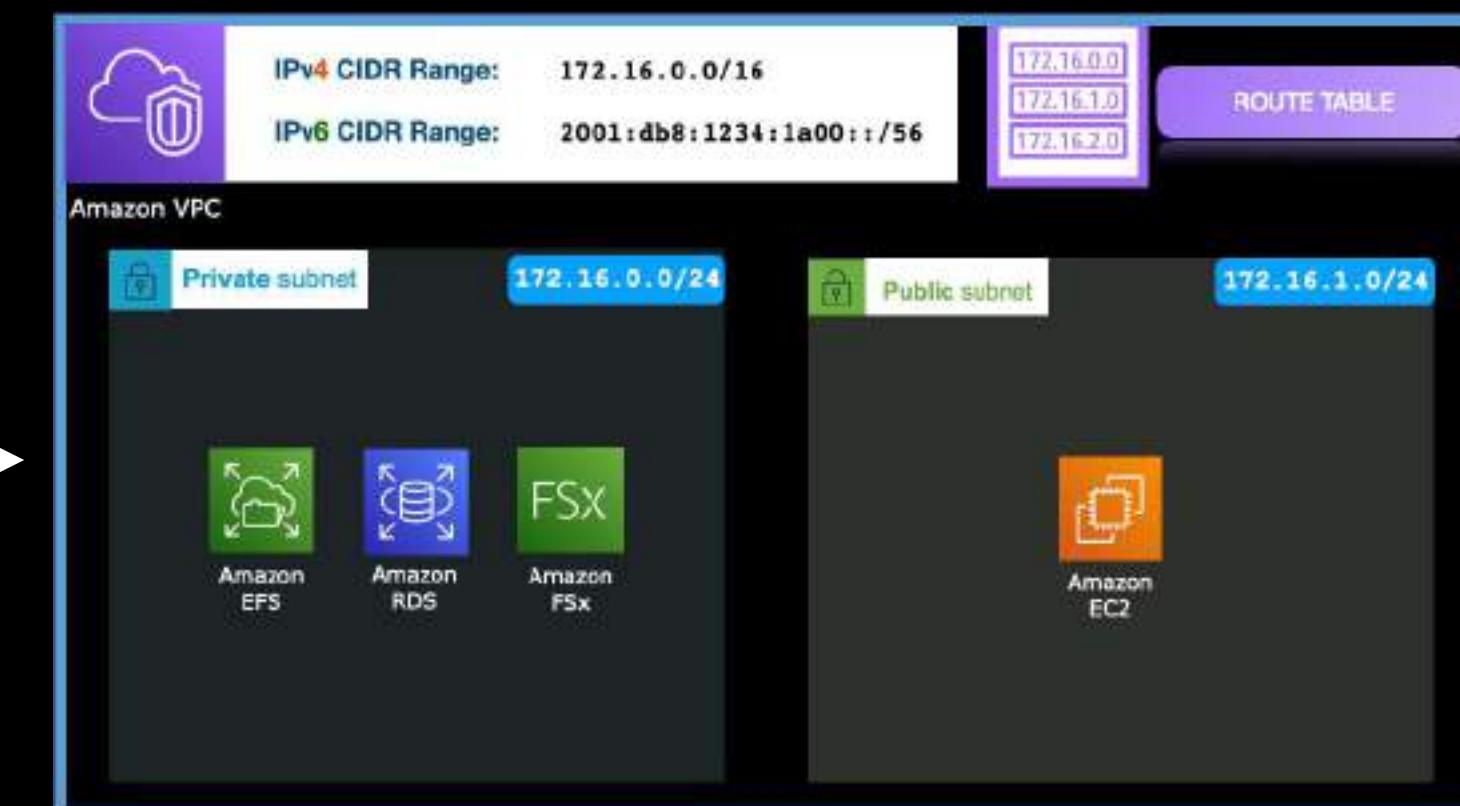
ASIA PACIFIC (Singapore)

VPC A - Manila Branch



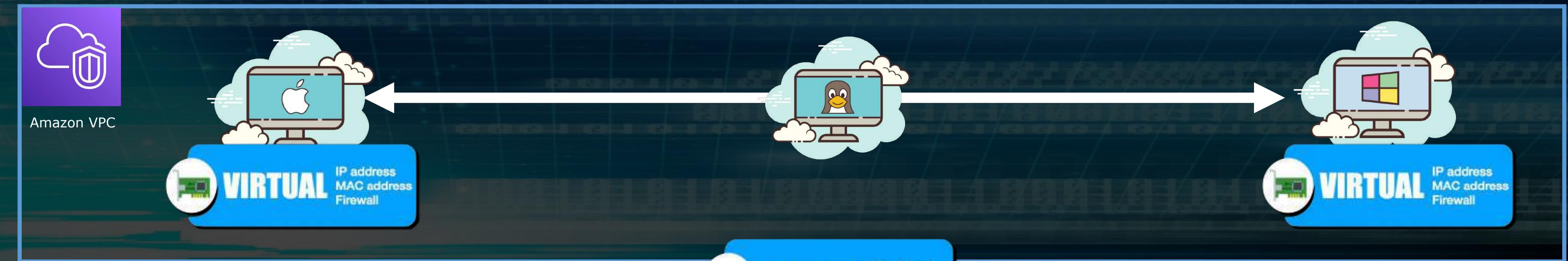
US EAST (Northern Virginia)

VPC B - New York Branch



# Virtual Private Cloud

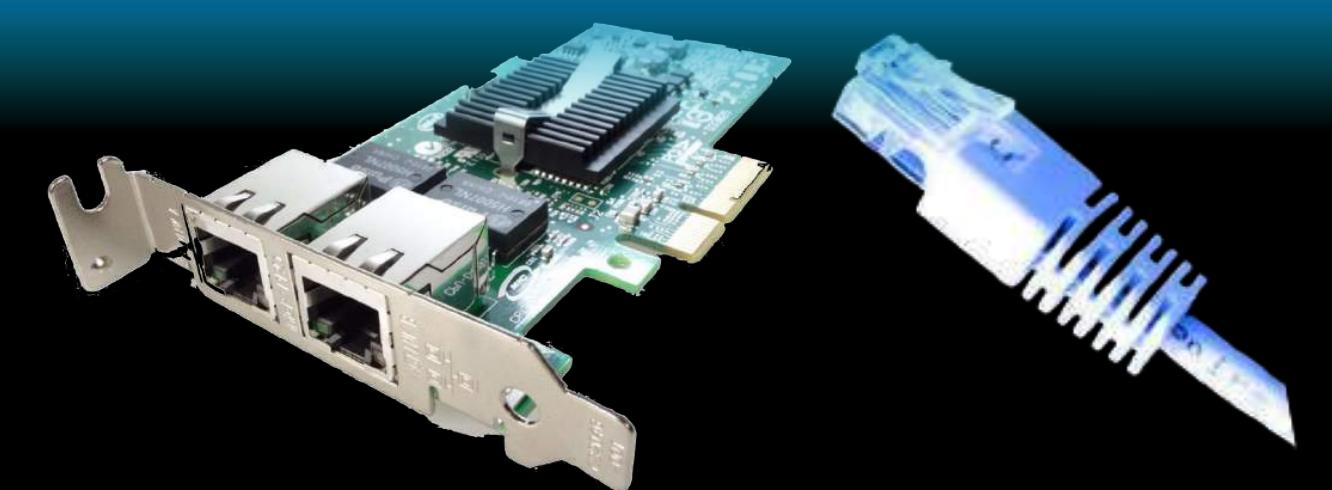
## Virtual Devices



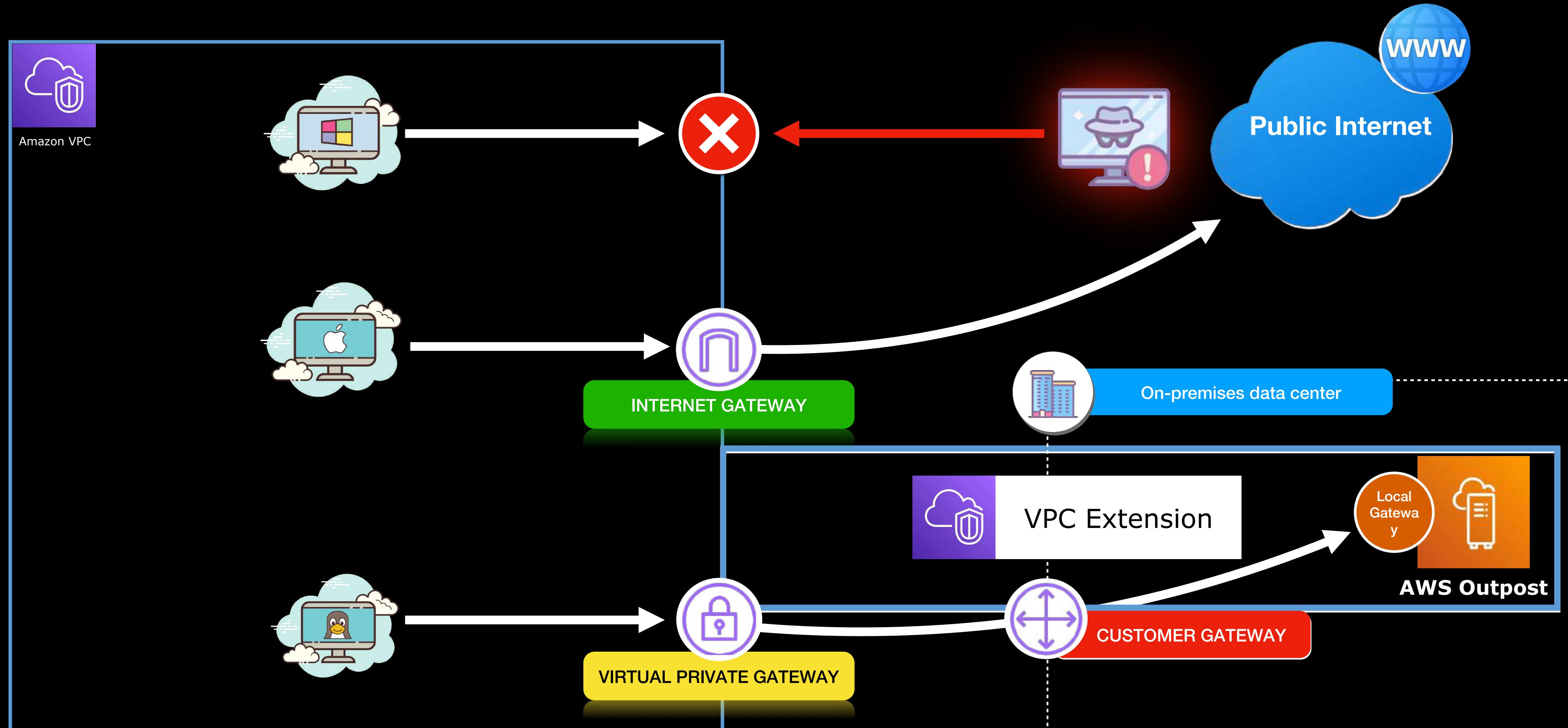
## Physical Devices

PCIe Network Interface Card

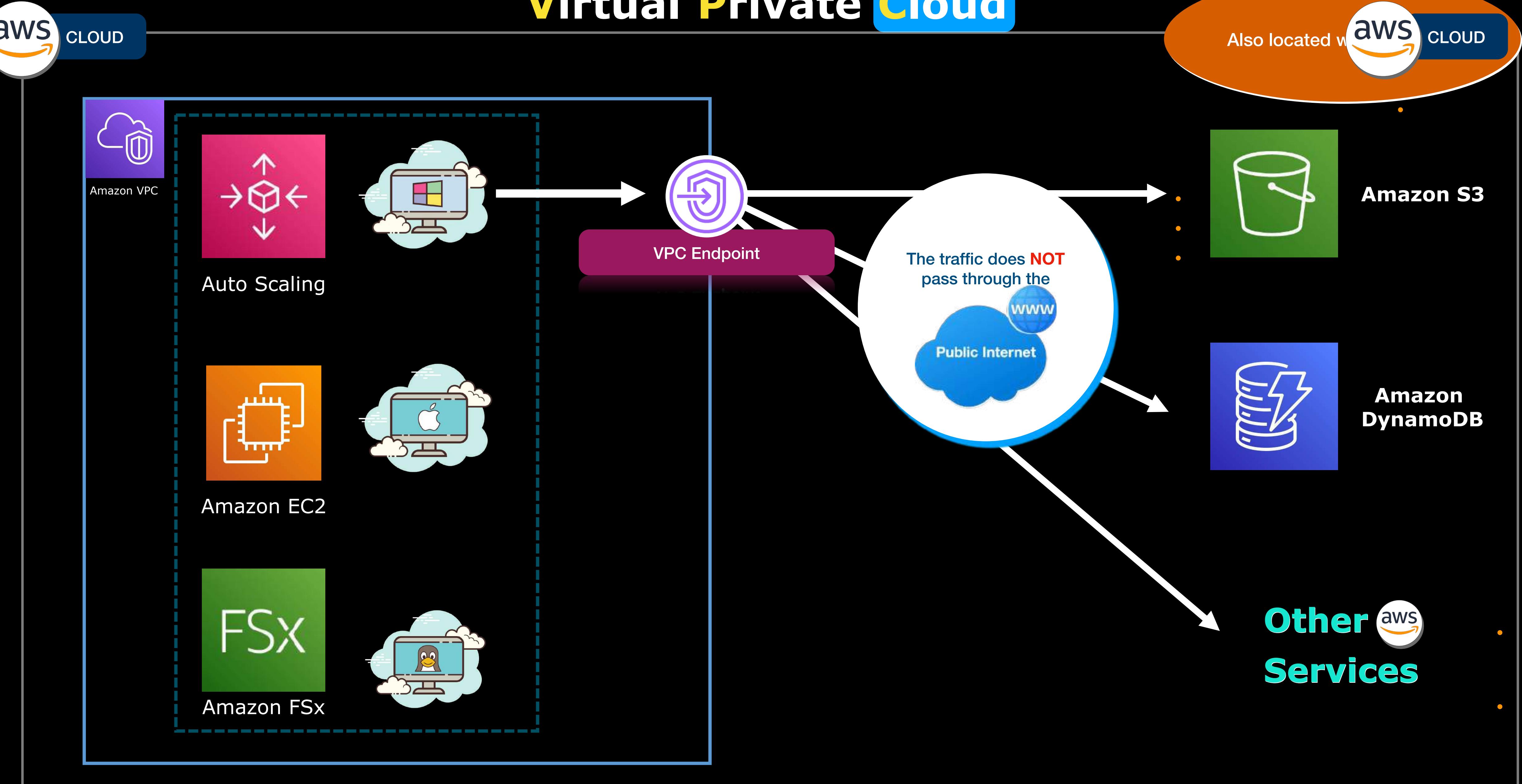
Nitro Card for VPC

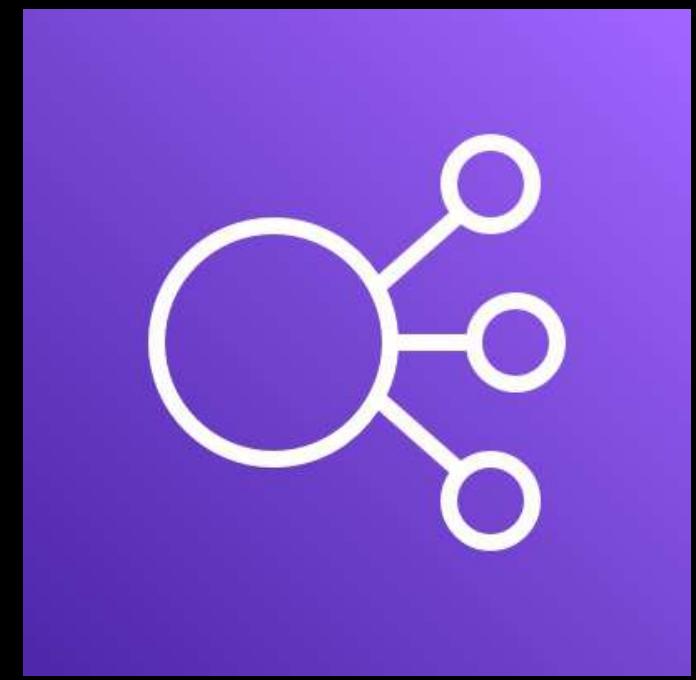


# Virtual Private Cloud



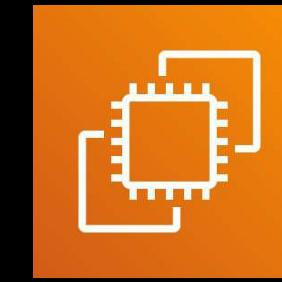
# Virtual Private Cloud





## Elastic Load Balancing

- Automatically distributes incoming traffic across **multiple targets** such as:



Amazon EC2  
Instance



Amazon ECS  
Task



AWS Fargate  
Task



AWS Lambda  
Function



IP Address

- It **distributes** (*load balances*) the incoming traffic to your underlying resources
- Provides **high-availability** to your web applications
- if one of your servers or EC2 instances fails (*unhealthy resource*), the request will be routed to another server (*healthy resource*)
- Routes incoming traffic across multiple Availability Zones, within a **single** AWS Region only.



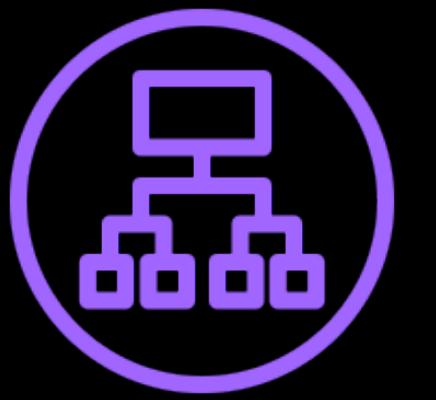
## Elastic Load Balancing TYPES

### PROTOCOL LISTENERS

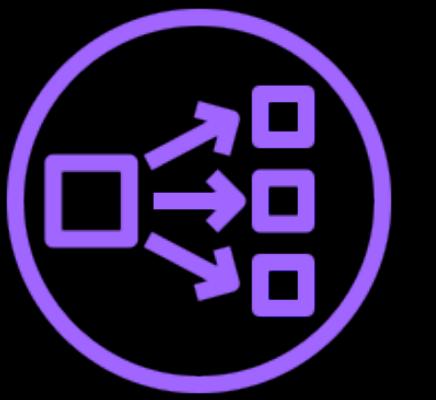
HTTP / HTTPS  
gRPC

### USE CASES

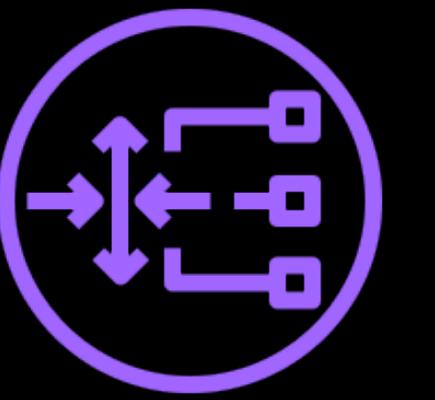
For **web apps**,  
microservices  
& containers



**Application**  
Load Balancer  
( ALB )



**Network**  
Load Balancer  
( NLB )



**Gateway**  
Load Balancer  
( GWLB )



**Classic**  
Load Balancer  
( CLB )

TCP / UDP  
TLS

IP

HTTP / HTTPS  
TCP  
SSL/TLS

Handling  
**millions of requests  
per second**  
while maintaining  
**ultra-low latencies**

Running third-party  
**virtual appliances**  
in AWS

For **legacy** applications  
in AWS  
  
For implementing  
**Custom Security Policies**  
and  
**TCP passthrough  
configuration**



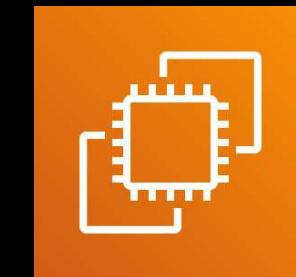
## Amazon Route 53

- A **Domain Name System** (DNS) web service
- DNS is a system that routes a **domain name** to a particular **IP address**

- **Map** domain names to:



Elastic IP  
address



Amazon EC2  
Instance



Amazon S3  
Static Website



Elastic Load  
Balancers



Amazon CloudFront  
Web Distributions



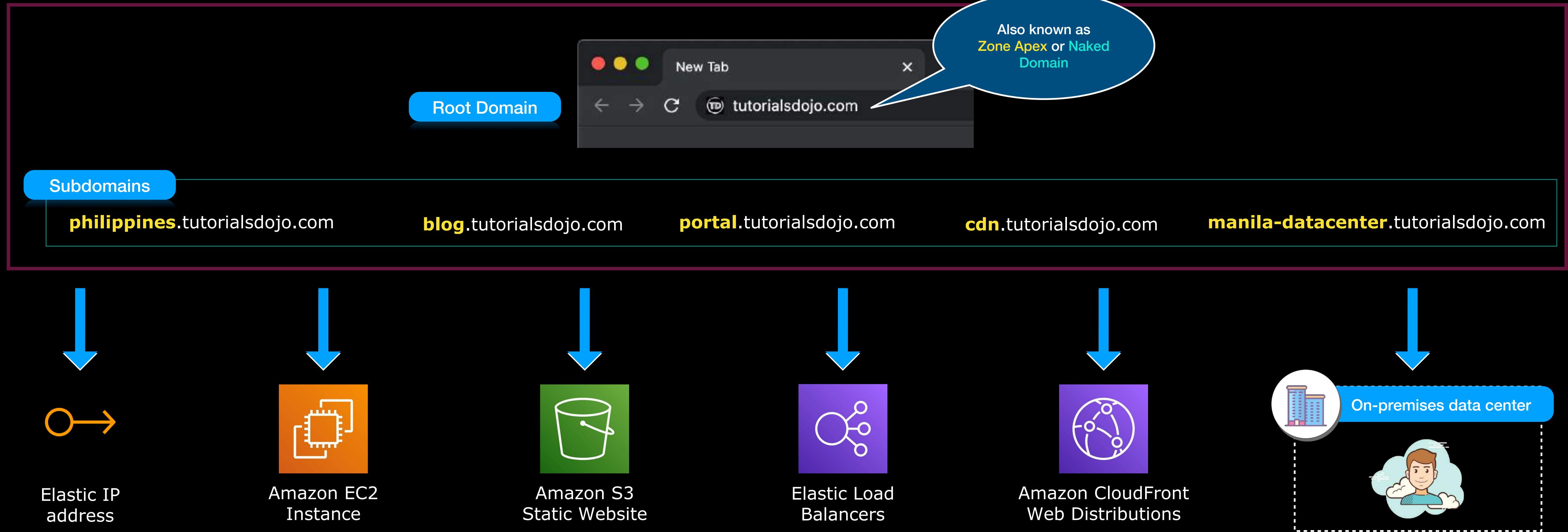
Amazon Route 53



Buy Domains



Manage Domains





## ROUTING POLICIES

Simple

Failover

Geolocation

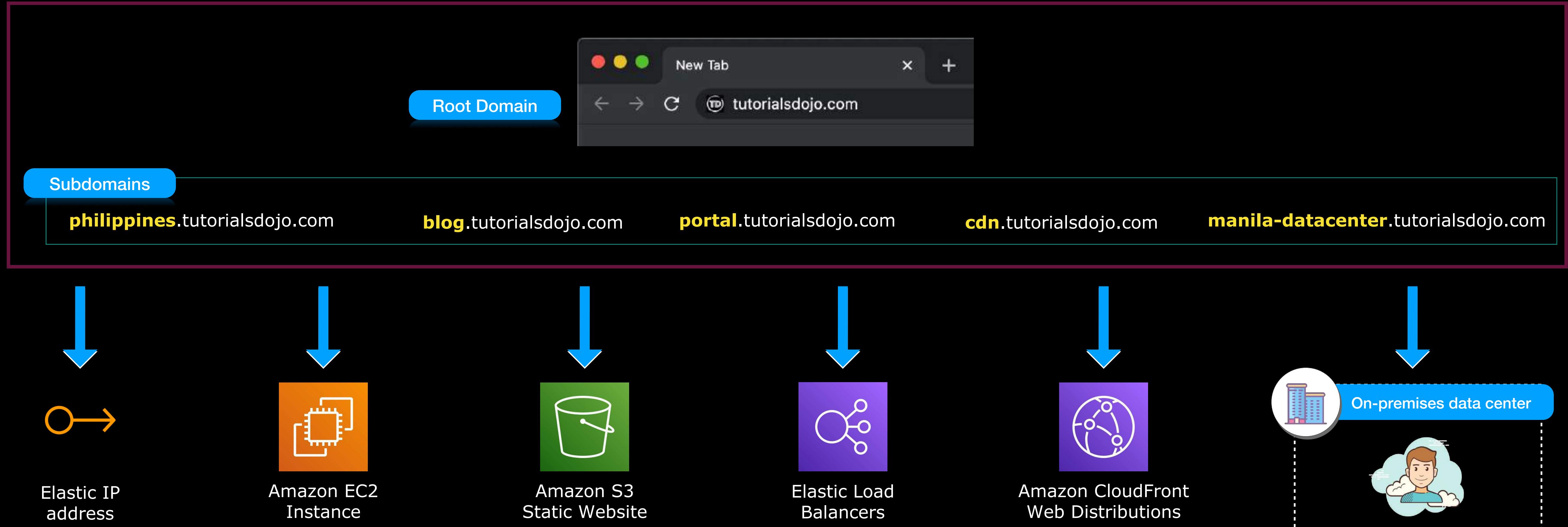
Geoproximity

Latency-Based

Multivalue Answer

Weighted

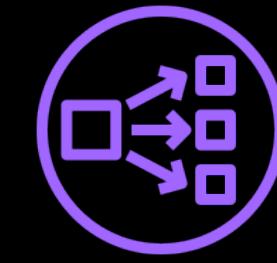
### Amazon Route 53



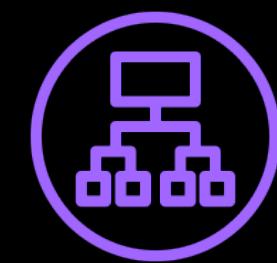


## AWS Global Accelerator

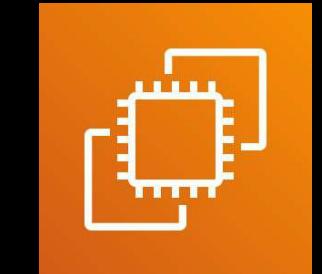
- Provides a set of **static anycast IP** addresses
- The static IP address serves as a **single fixed entry point** to:



Network  
Load Balancer



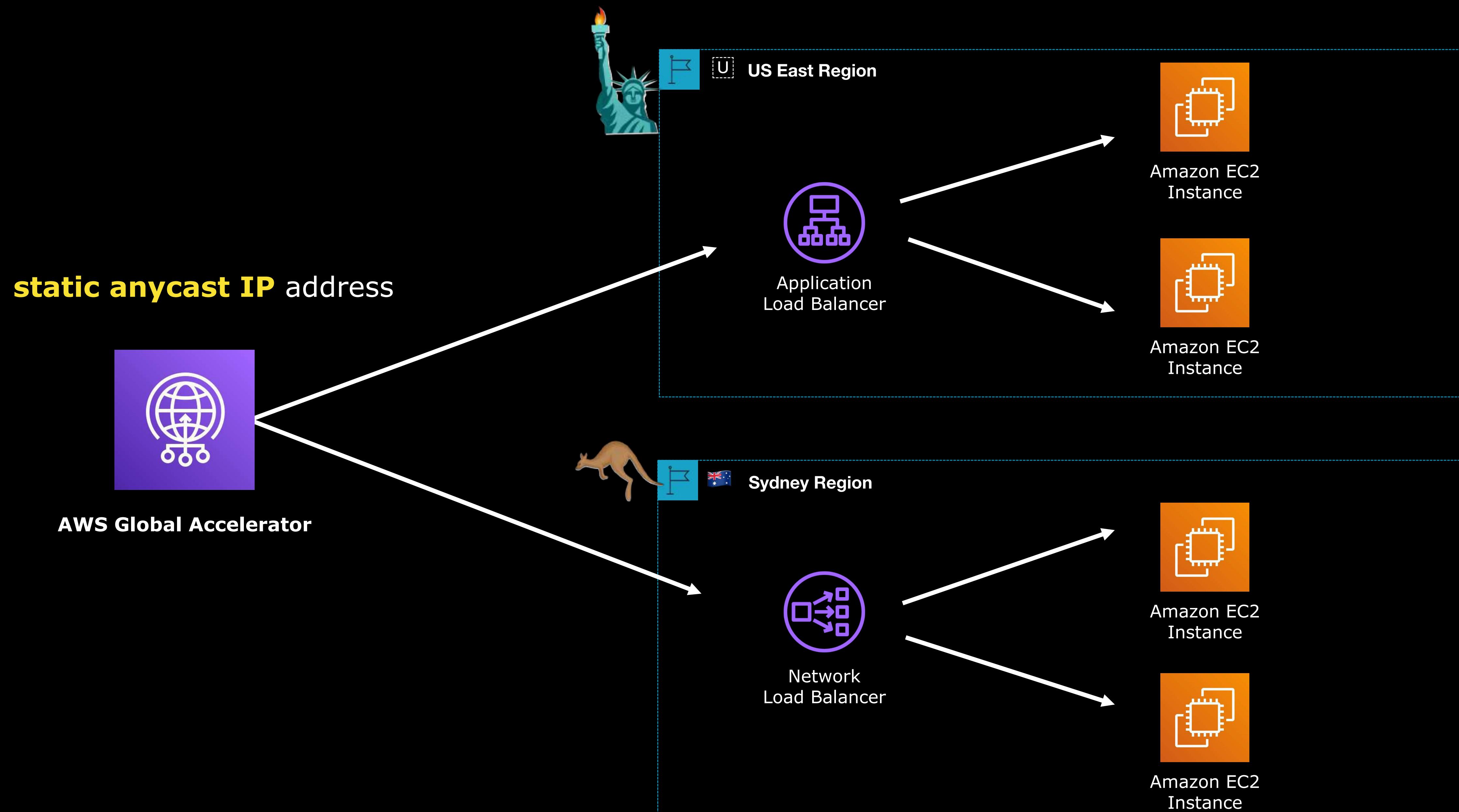
Application  
Load Balancer

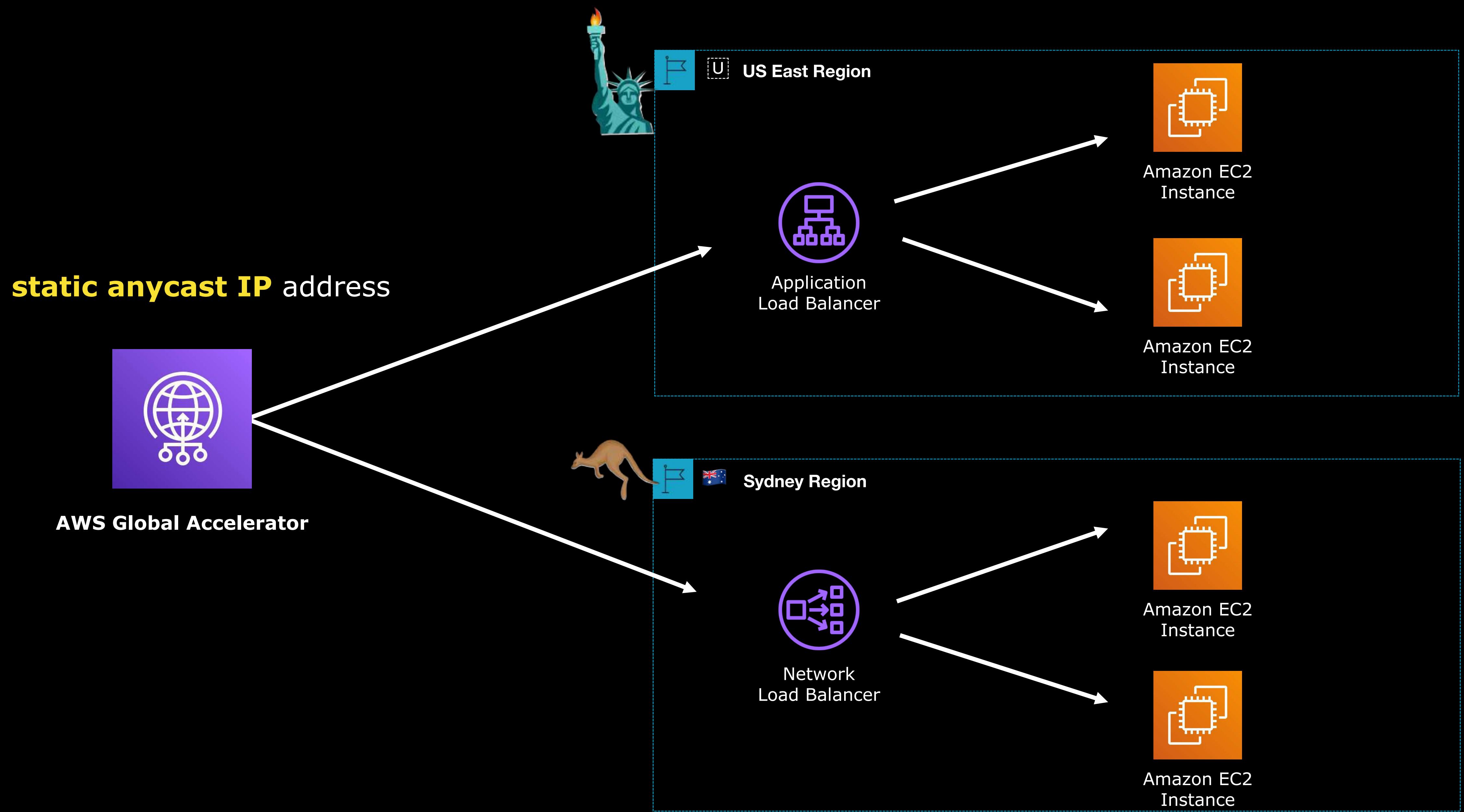


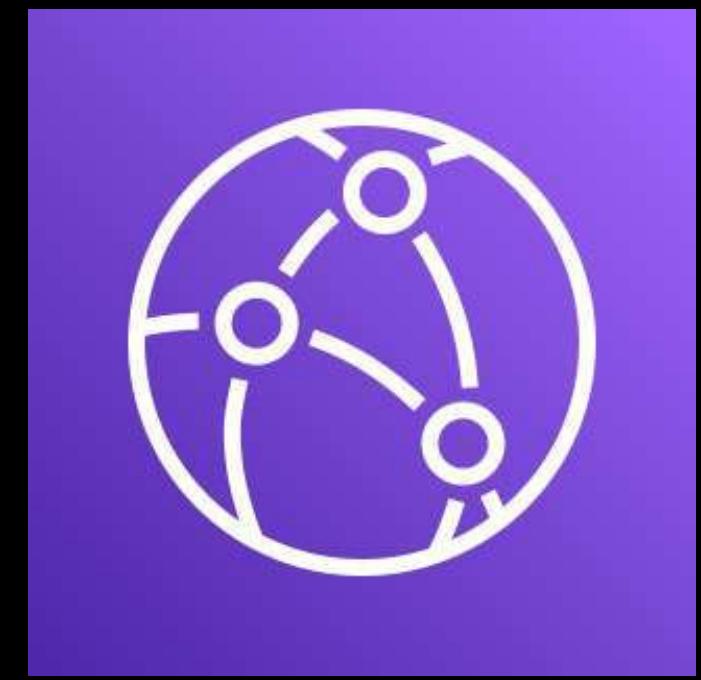
Amazon EC2  
Instance



Elastic IP  
address







## Amazon CloudFront

- A **content delivery network** (CDN) service
- Quickly delivers static content and video stream to your clients.
- A CDN is a **globally-distributed network** of services/servers spread around the globe that stores or **caches** your files.
- **Reduces latency** by shortening the time it takes to deliver your data to your users
- **Improves the response time** of your application.
- **Caches** your images, videos, media files, or software packages

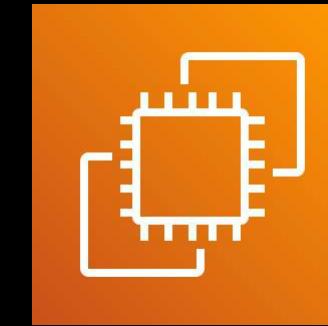


## AWS PrivateLink

- Allows **private connectivity** to various AWS services
- **Does not pass through the public Internet.**
- Provides a **private endpoint** that you can use for your:



Amazon VPC



Amazon EC2



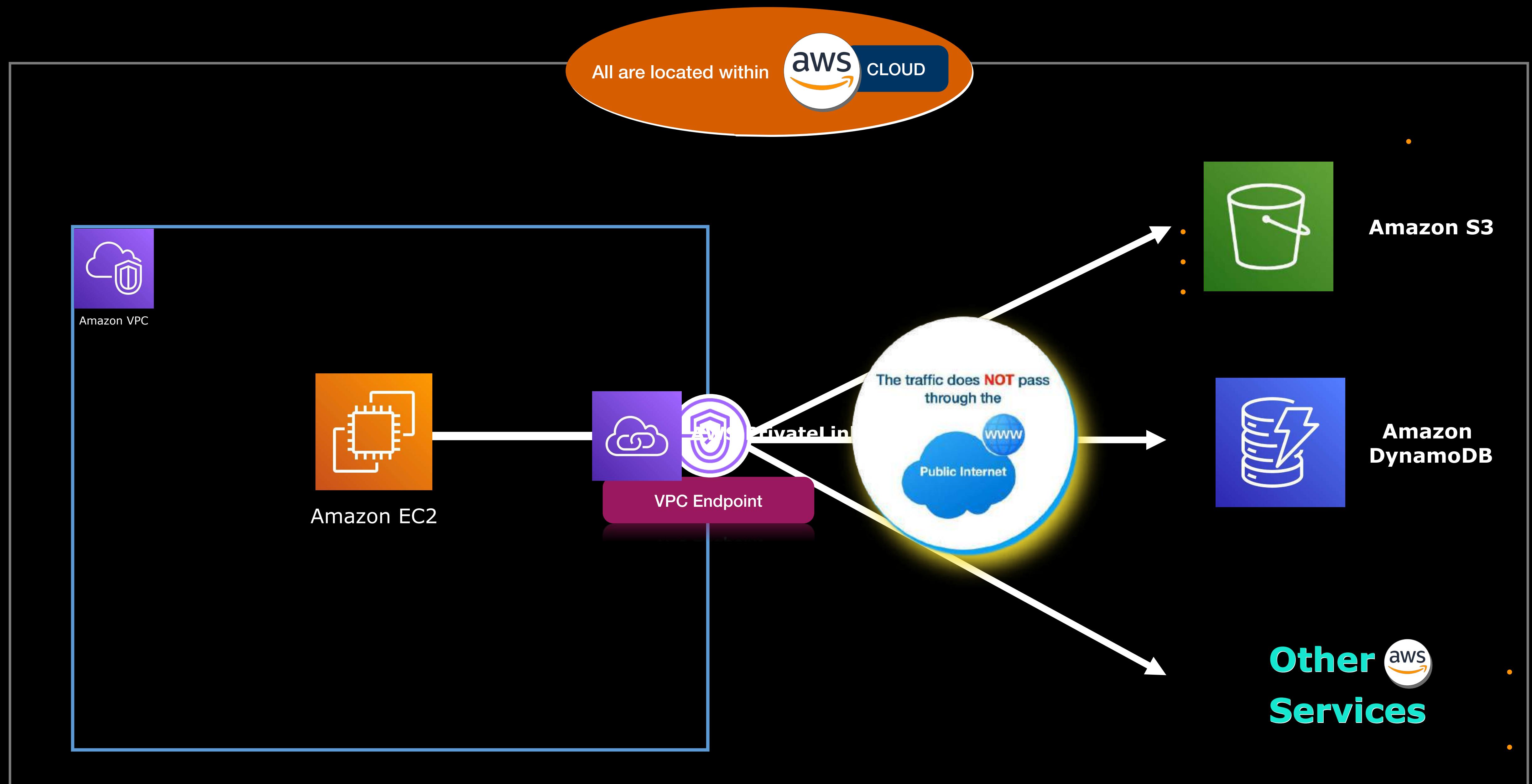
Amazon S3



Amazon  
DynamoDB

**Other  
Services**

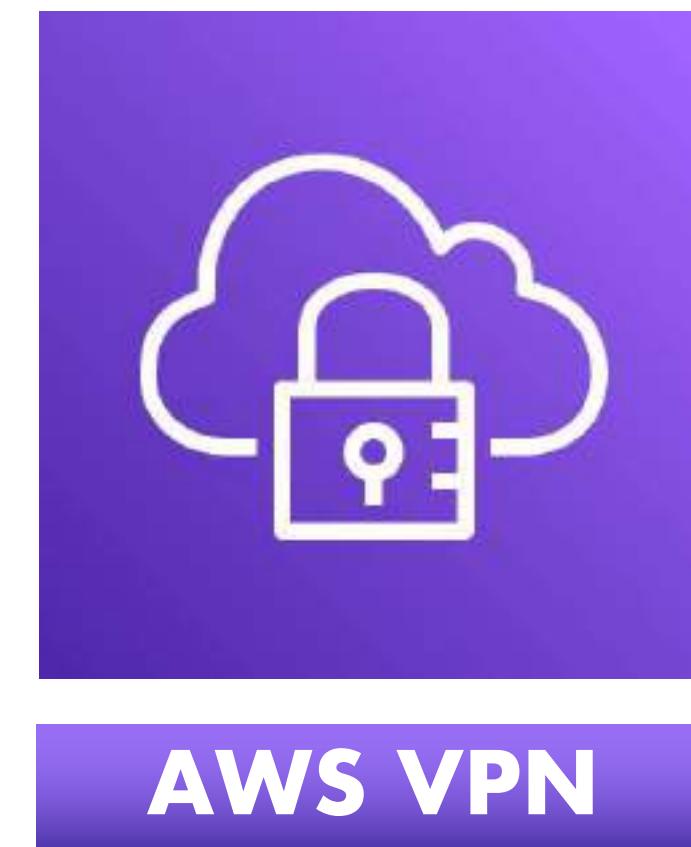




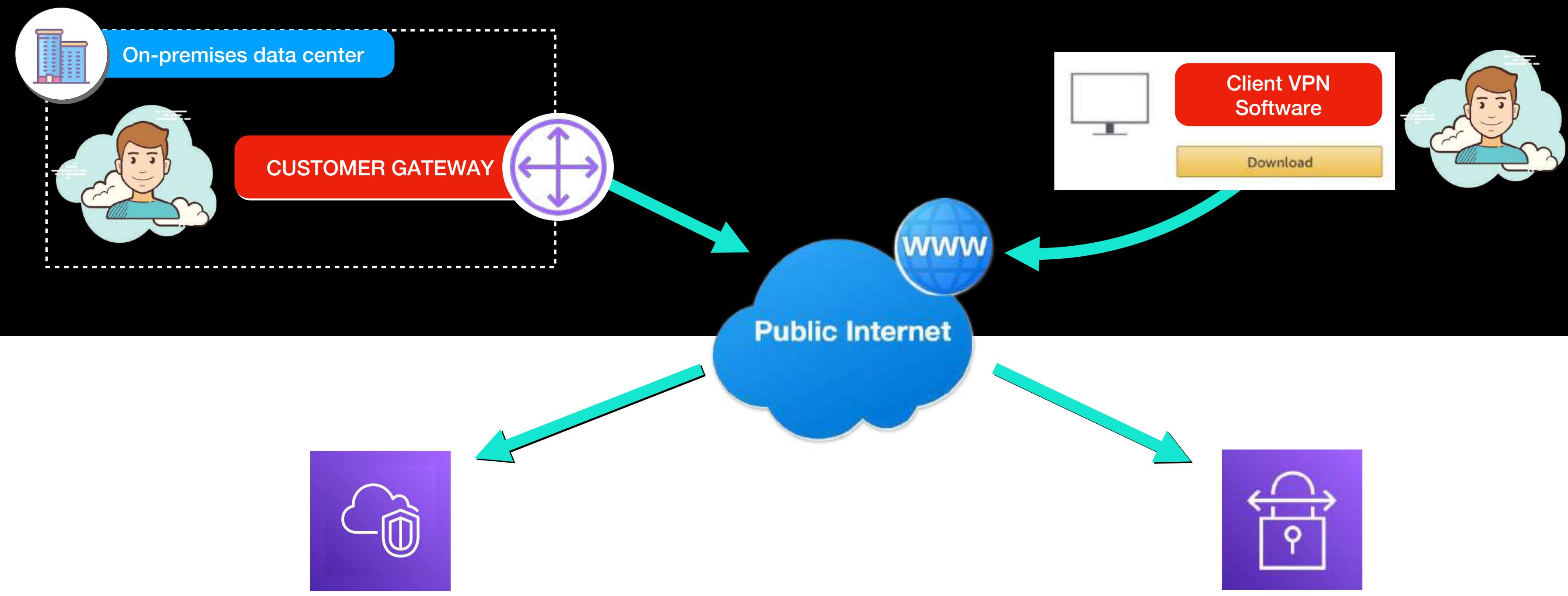


## AWS VPN

- AWS **V**irtual **P**rivate **N**etwork, or AWS VPN
- Enables you to **connect your on-premises network to AWS**.
- An encrypted connection that **passes through the public Internet**.
- Uses the **IPsec protocol** to authenticate and encrypt your data in transit.



**AWS VPN**



**AWS Site-to-Site VPN**

**AWS Client VPN**

**ENDPOINTS**

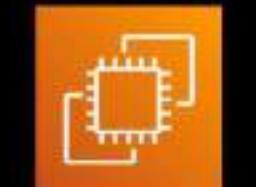
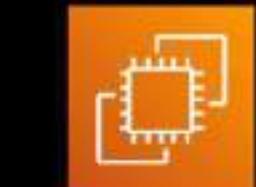


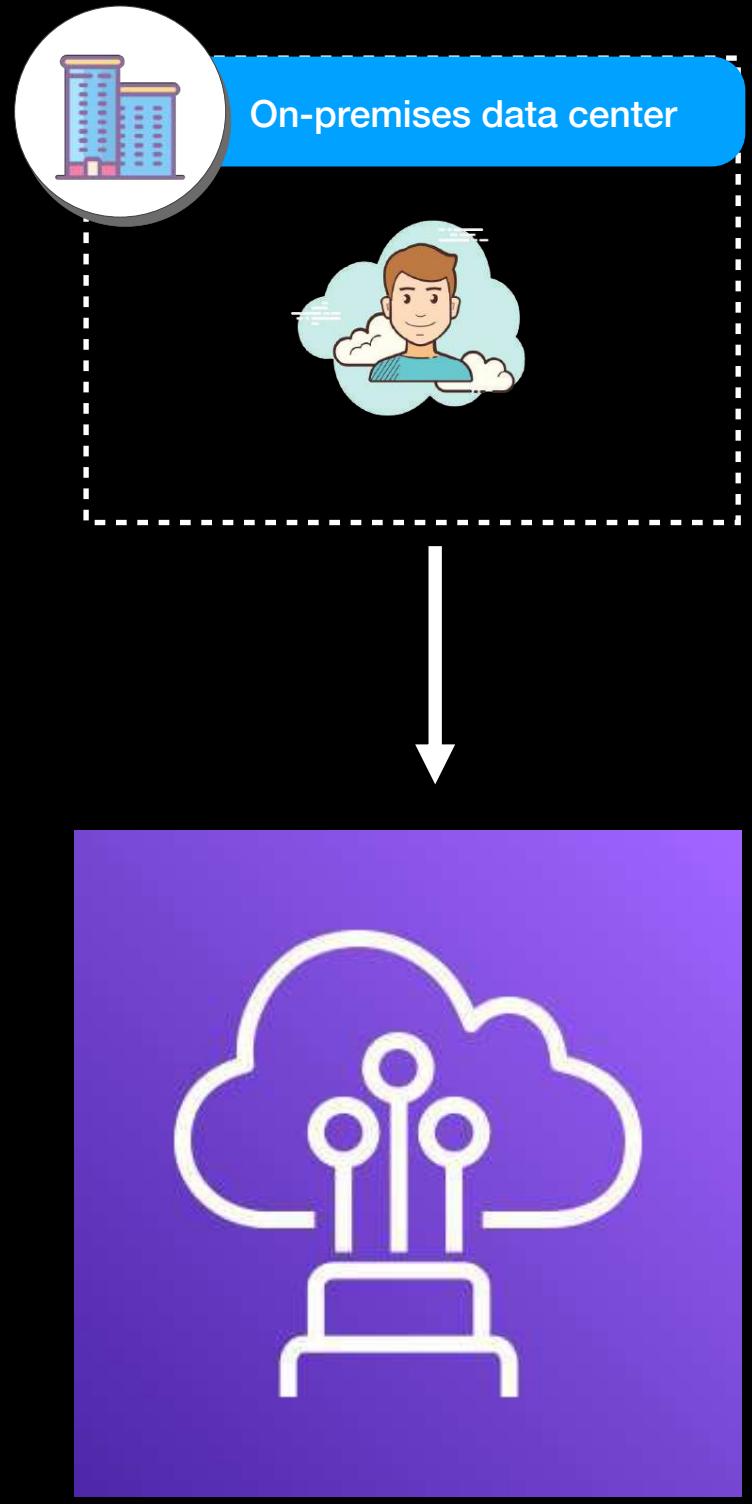
Site-to-Site VPN Endpoint



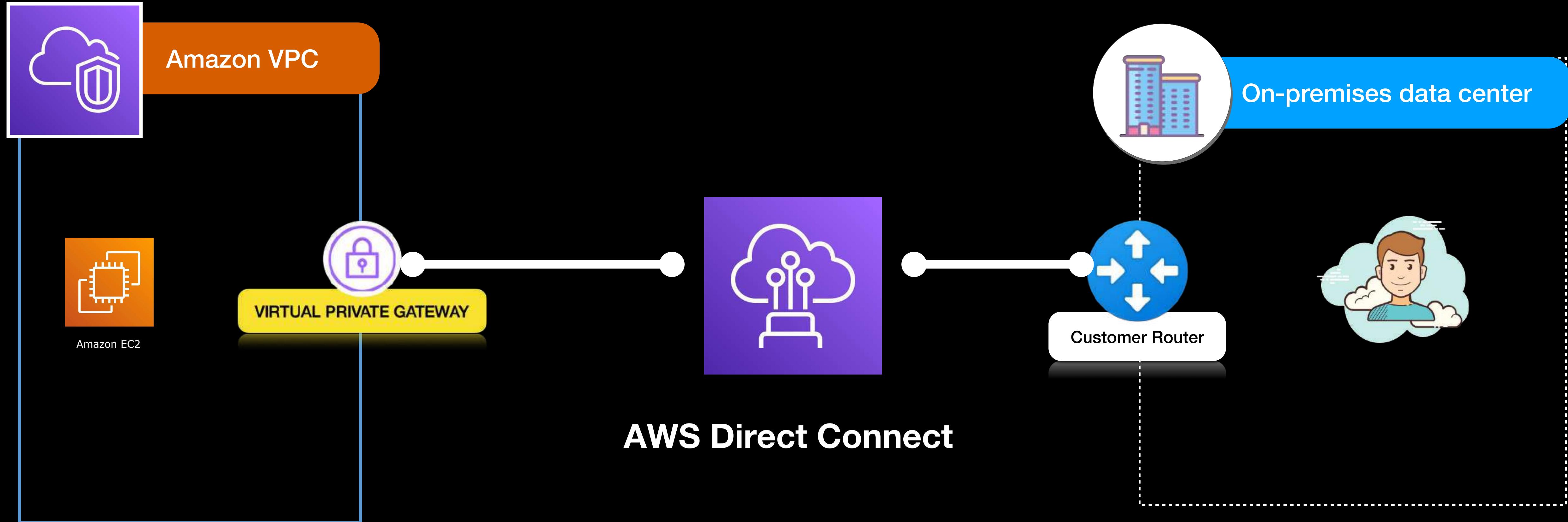
Client VPN Endpoint

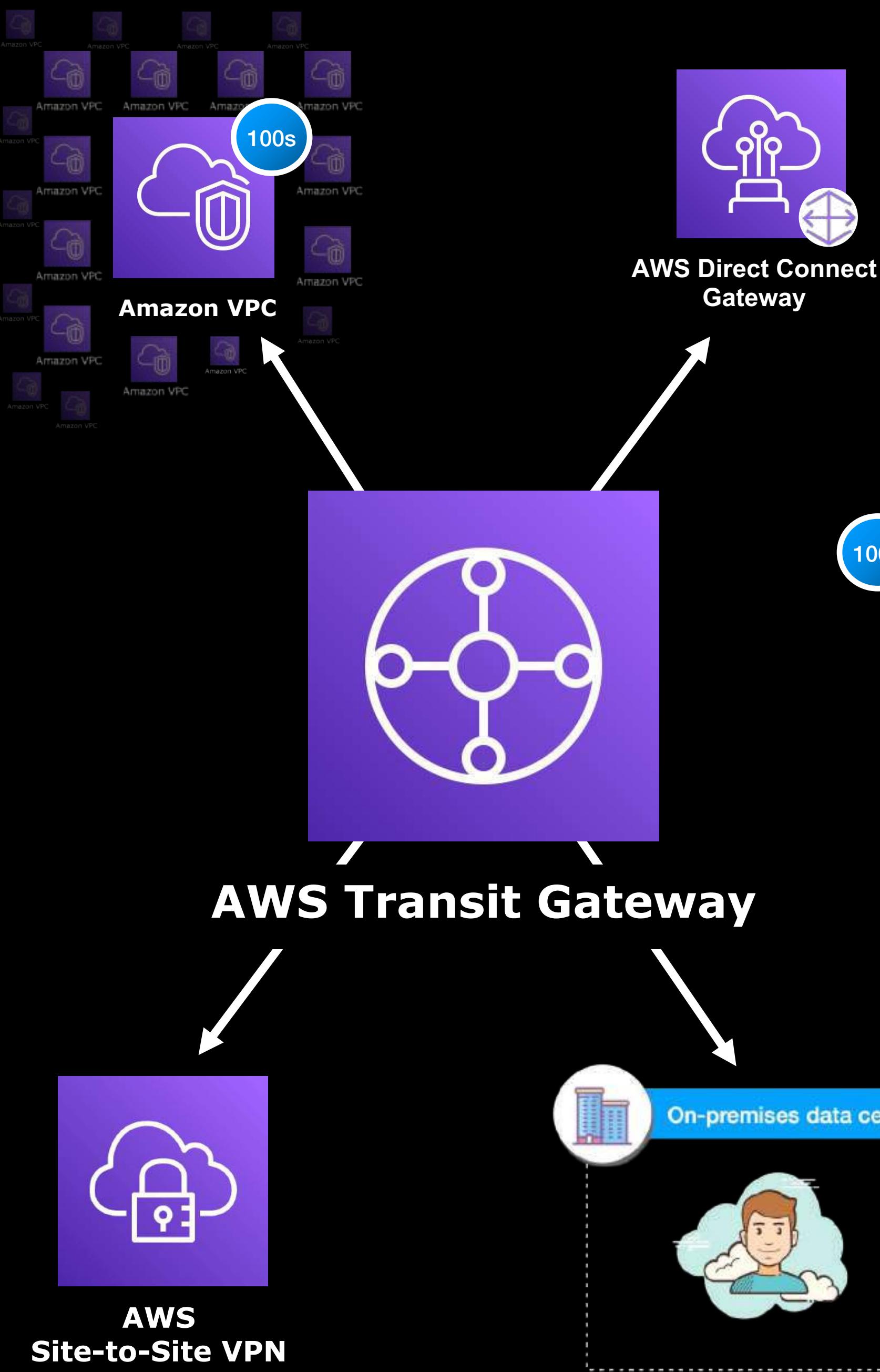
AWS Transit Gateway





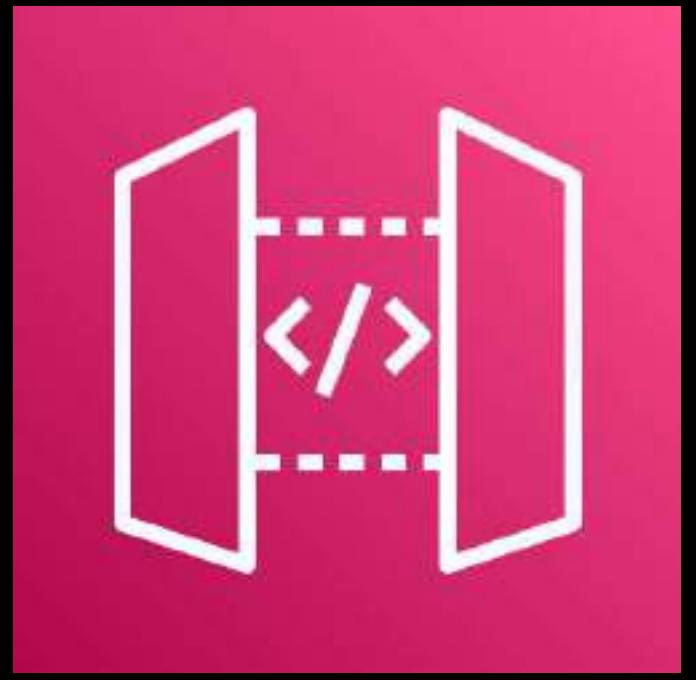
- Allows you to establish a **dedicated network connection** from your on-premises network to AWS
- Provides a **more consistent network experience** over Internet-based connections such as a VPN, and a **higher bandwidth**.
- You can create a **private virtual interface** to enable your on-premises servers to connect to the virtual private gateway of your Amazon VPC.
- You can group your virtual private gateways and private virtual interfaces using a **Direct Connect Gateway**.
- You can also use a **public virtual interface** to connect to your Amazon S3 buckets and other public resources in AWS.
- **The traffic does NOT pass through the public Internet.**



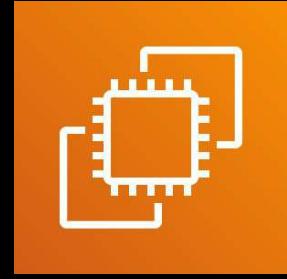


- **Connects your cloud networks** (e.g. Amazon VPCs, VPNs, Direct Connect Gateways, and on-premises networks) **to a single gateway**.
- Recommended for large organizations with **hundreds of Amazon VPCs, site-to-site VPNs, and external networks**.
- Reduces the complexity of your infrastructure and makes scaling easier

- Allows you to publish, maintain, monitor, and secure your **RESTful APIs**.
- Also supports **WebSockets** for real-time message communication
- **Acts as a front door for your back-end services** that are running on:



**Amazon API Gateway**



Amazon EC2



Amazon ECS



AWS Fargate

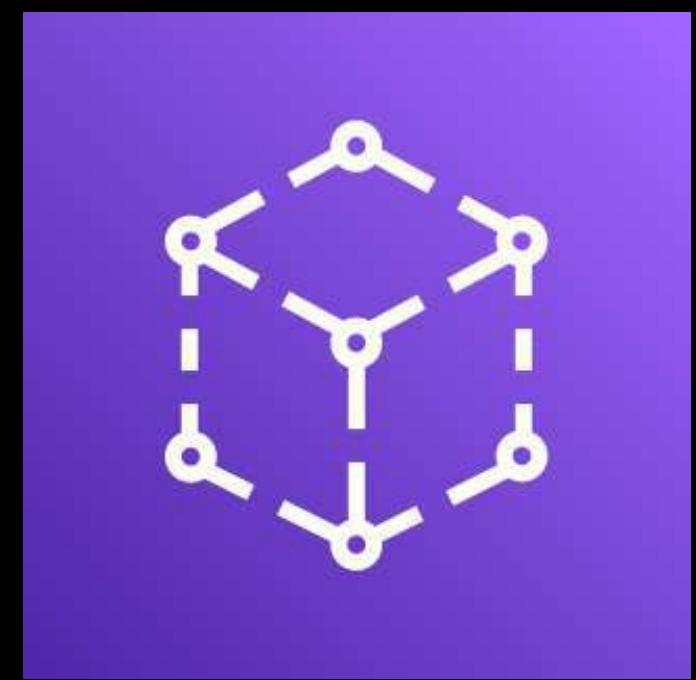


AWS Lambda



AWS Elastic  
Beanstalk

- Works as a **Proxy** – similar to APIGEE, Mulesoft and other proxies/integration platforms



## AWS App Mesh

- A **service mesh** (*an infrastructure layer that handles communication between microservices*)
- Provides application-level networking for the different types of **containerized applications** in AWS.
- **Allows your services to communicate with each other** across multiple types of computing infrastructure.
- Uses *(an open-source service mesh proxy)*
- Can be used with microservice containers managed by:



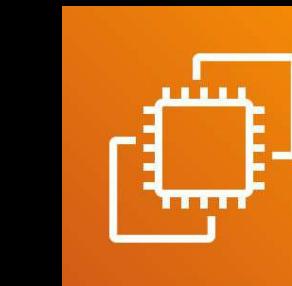
Amazon ECS



Amazon EKS



AWS Fargate

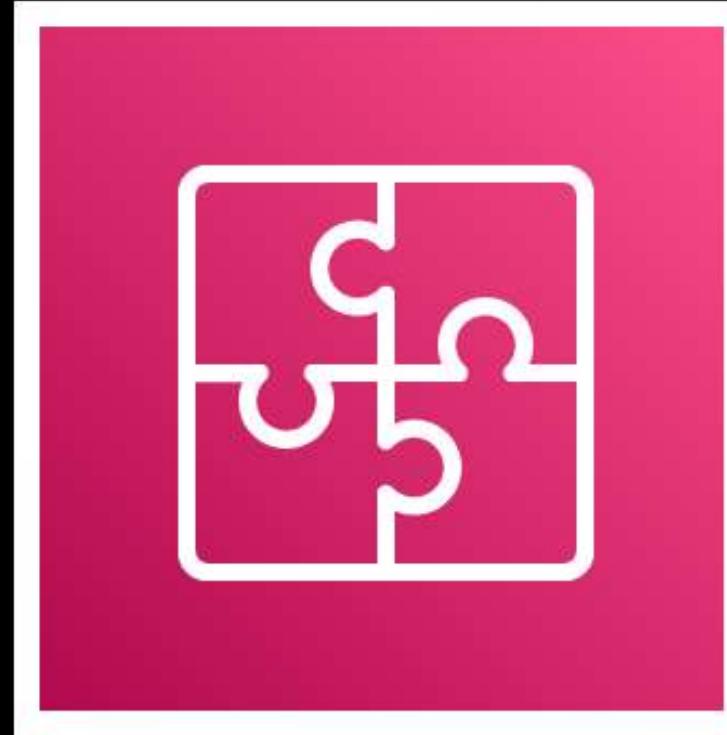


Amazon EC2



## AWS Cloud Map

- A **cloud resource discovery** service.
- Commonly used in microservices and containerized applications that have **dynamically changing resources**.
- You can **name your containerized application resources** with **custom names**.
- Improves your containerized applications in AWS by always discovering the **most up-to-date locations of your resources**
- **Improves the availability** of your system.

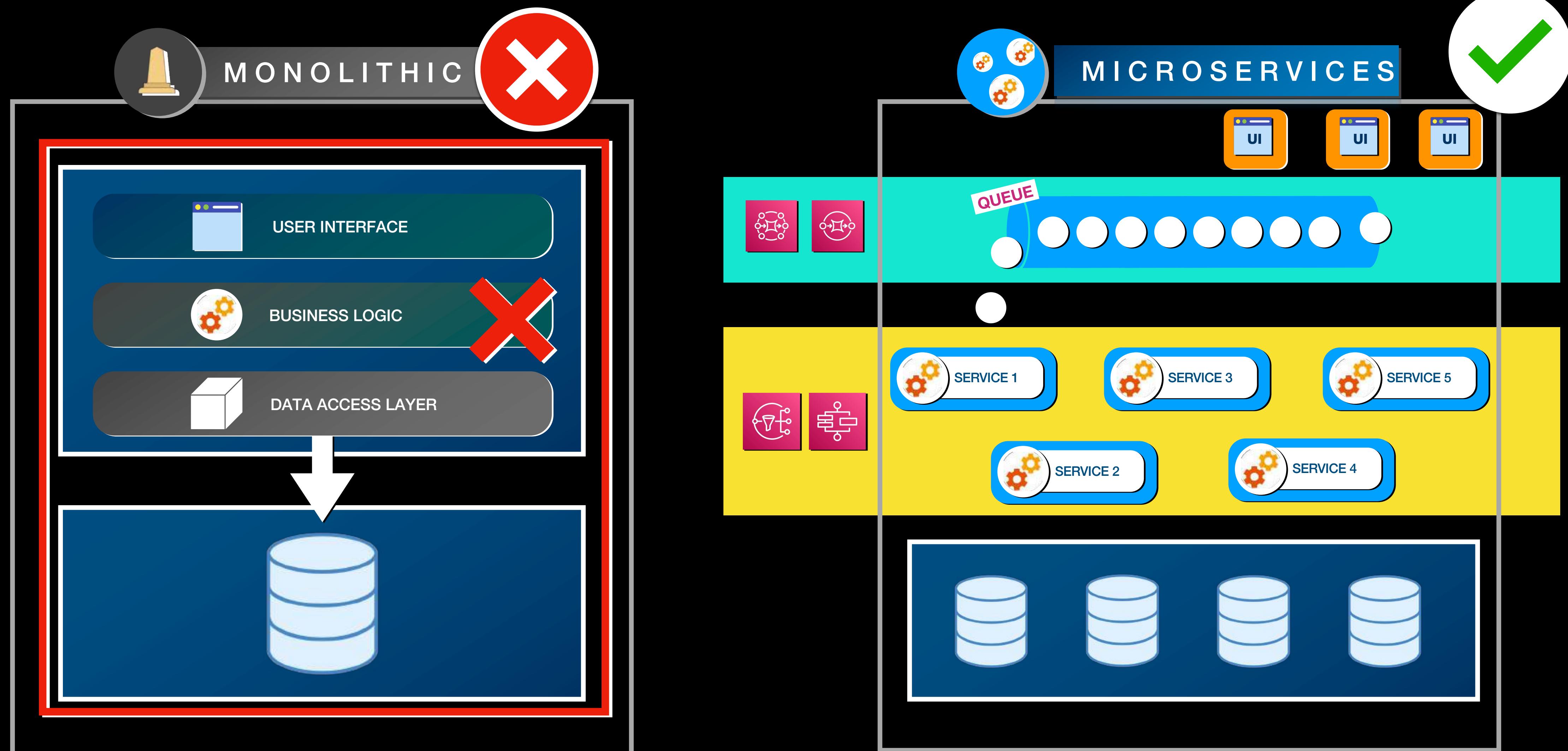


# Application Integration Services Overview

---

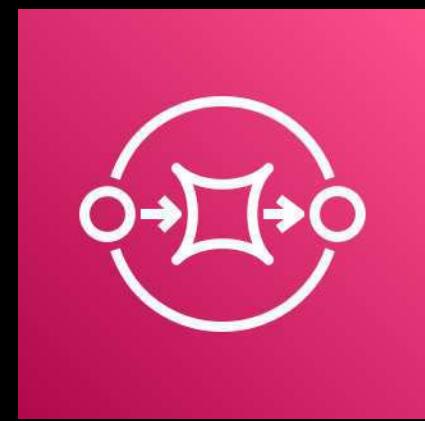


# Application Integration Services





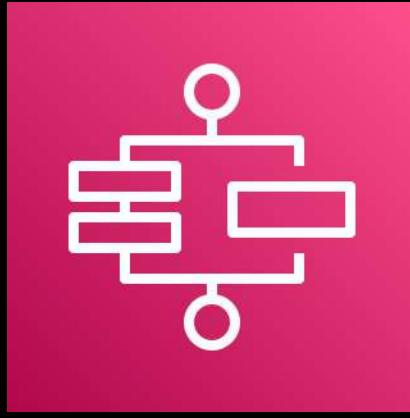
## Application Integration Services



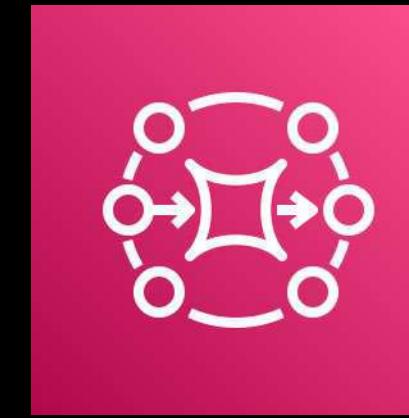
**Amazon Simple Queue Service  
(Amazon SQS)**



**Amazon Simple Notification  
Service (Amazon SNS)**



**AWS Step Functions**



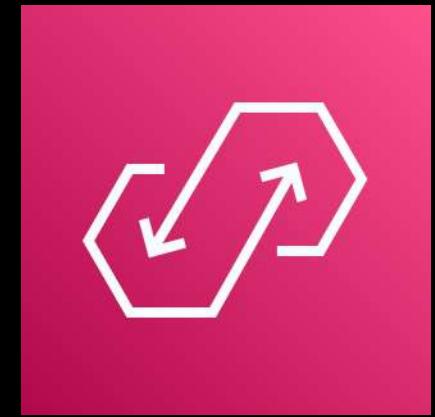
**Amazon MQ**



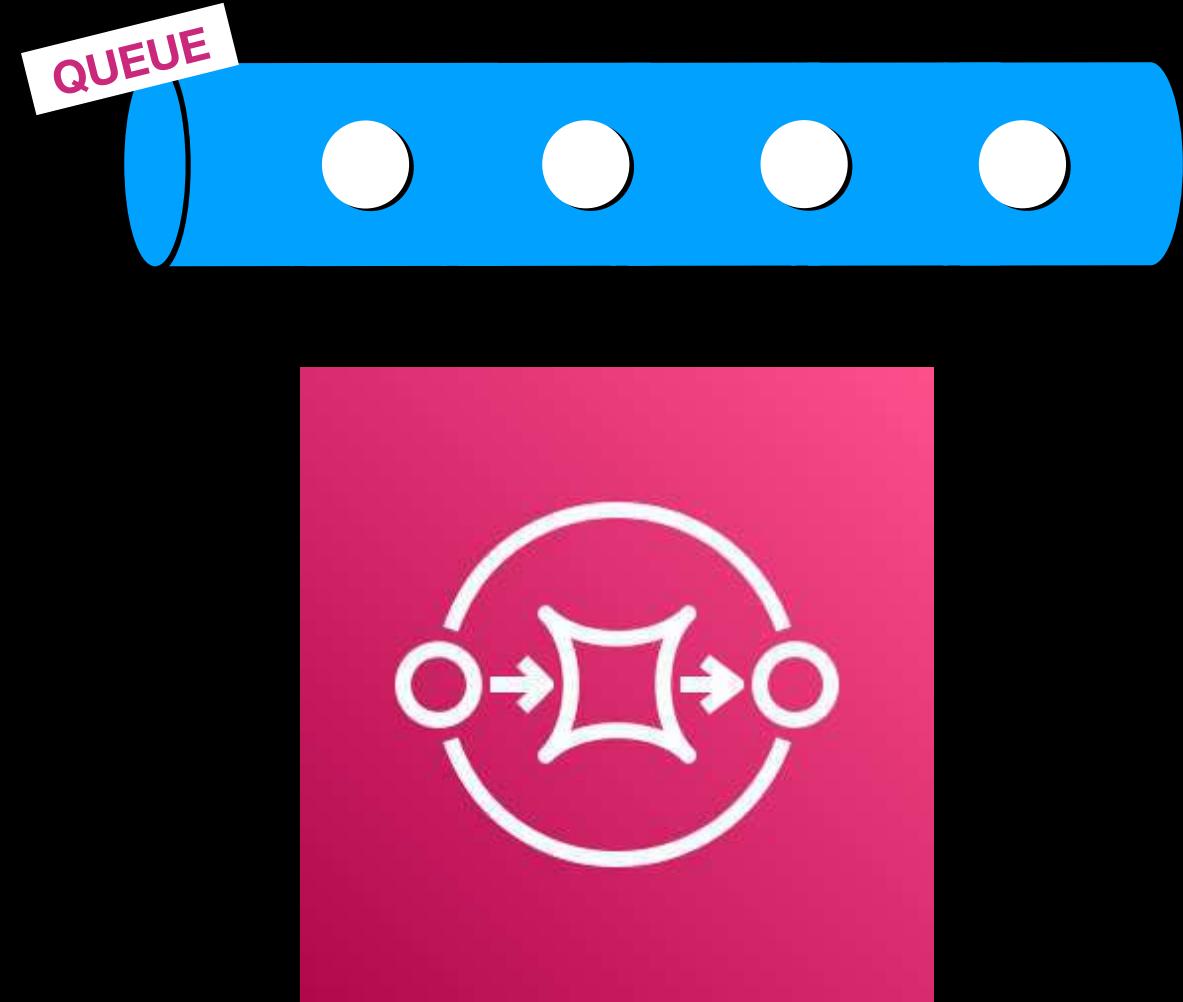
**Amazon EventBridge**



**AWS  
AppSync**

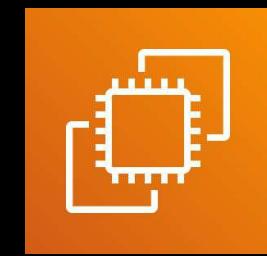


**Amazon AppFlow**



## Amazon Simple Queue Service (Amazon SQS)

- A fully managed message queuing service
- The messages can be consumed or processed by:



Amazon EC2



AWS Lambda



Amazon ECS



Other Consumers

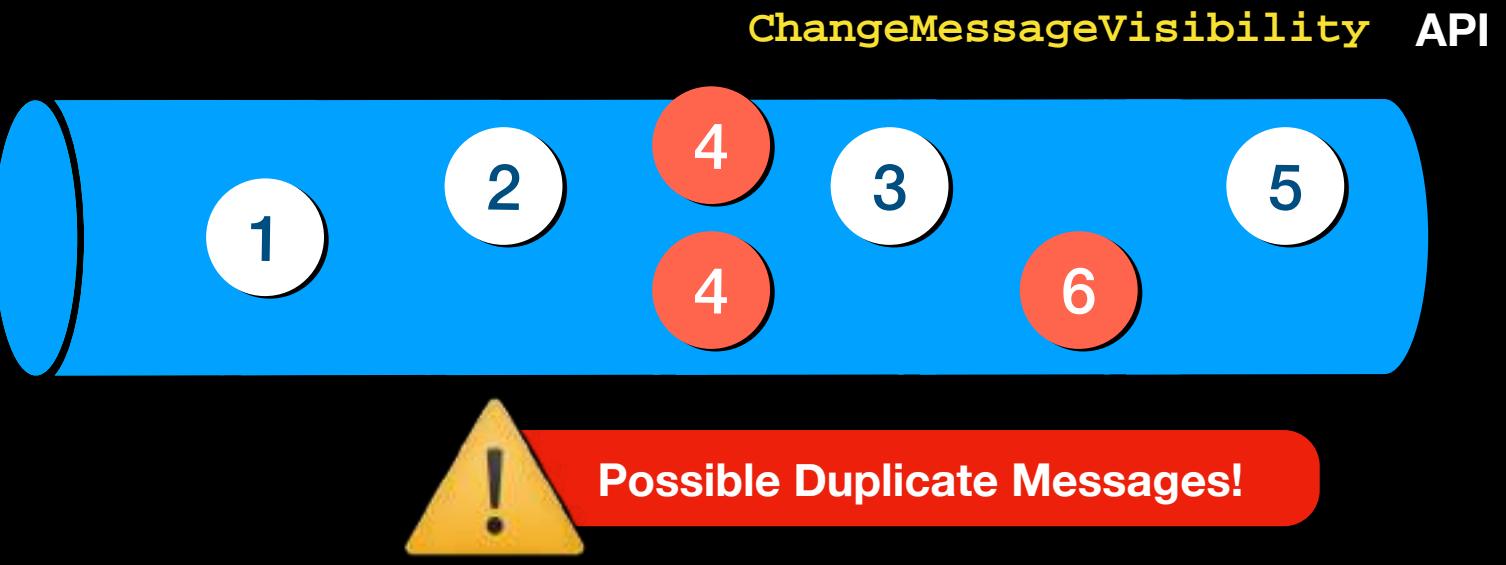
- Can replace your traditional message-oriented middleware without having to manage any servers or resources



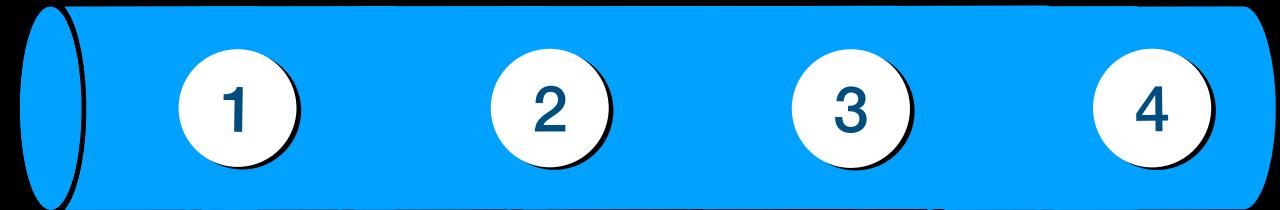


## Amazon SQS TYPES

### STANDARD



**FIFO**  
First In, First Out



### DELIVERY

At Least Once

Exactly Once

### ORDERING

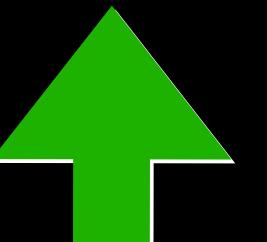
Best Effort

Messages might be delivered in a different order

Preserves the exact order  
in which the messages are received

### THROUGHPUT

HIGH

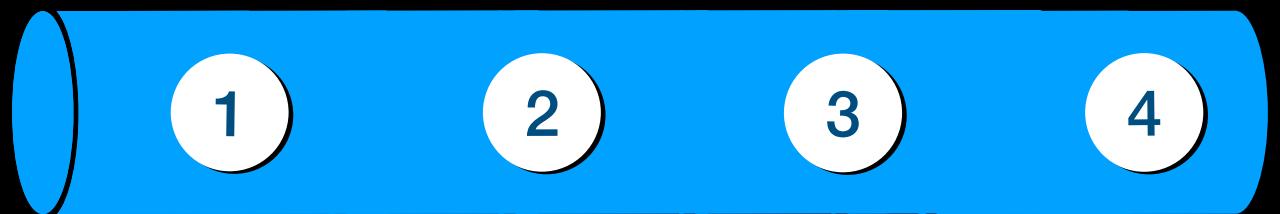
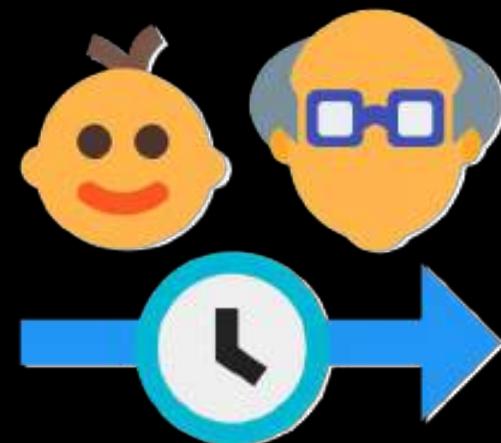


LIMITED

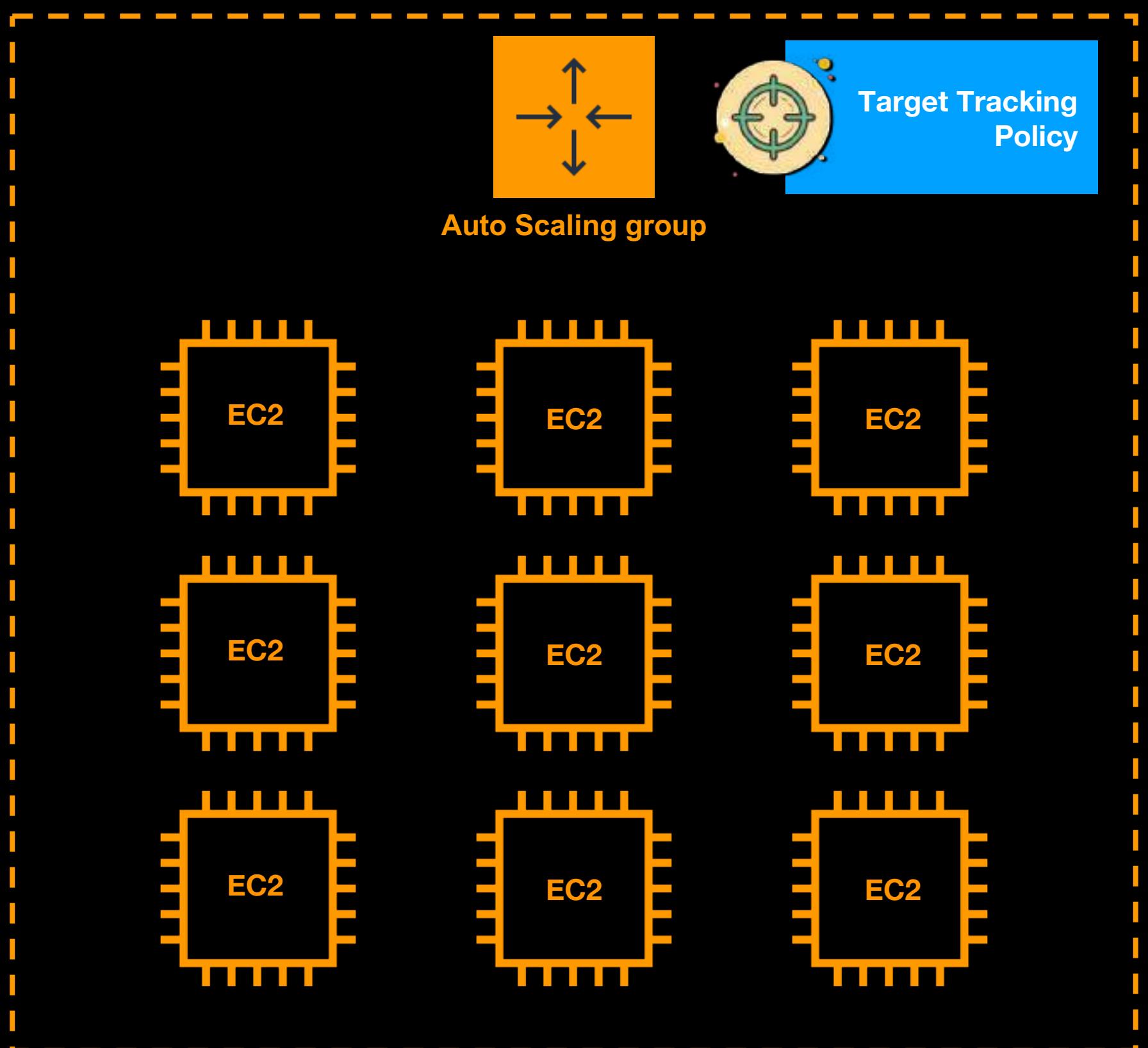




Amazon SQS



- Age of the Oldest Message
- Queue Depth
- Number of Messages





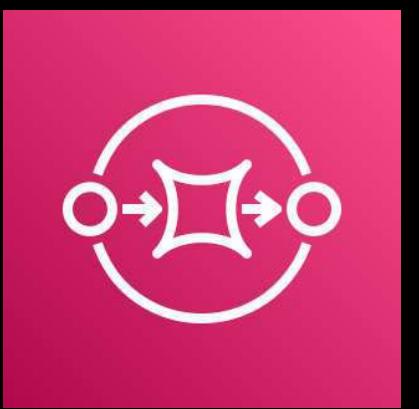
Amazon Simple Notification Service  
(Amazon SNS)



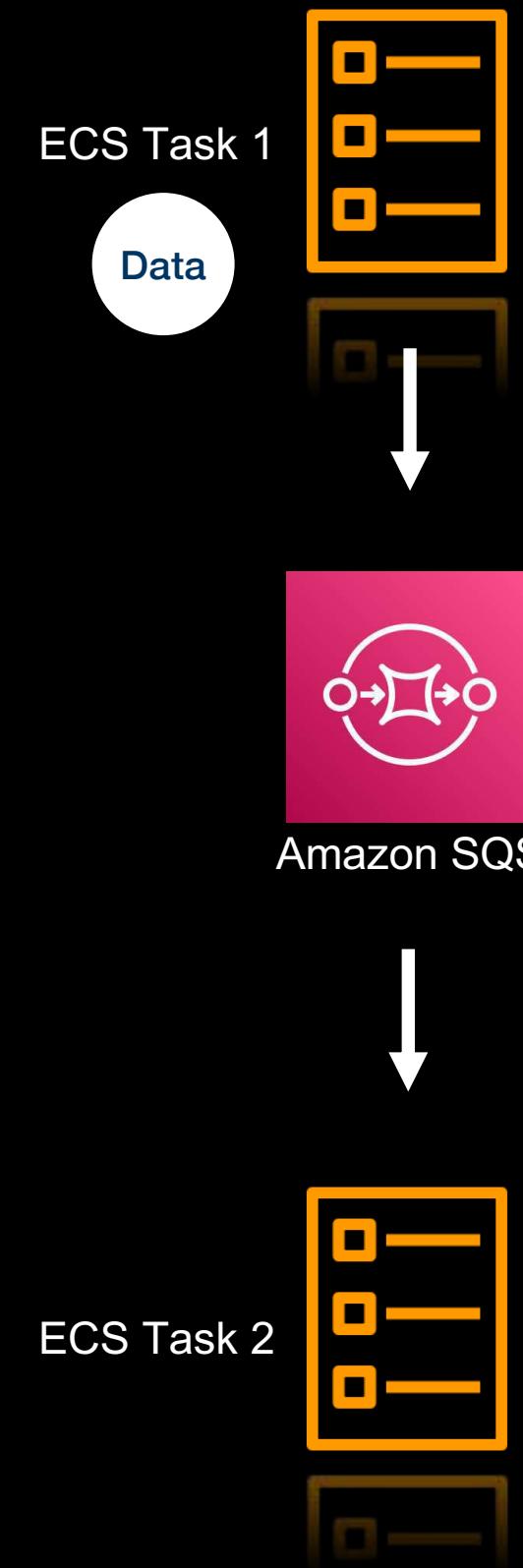
Amazon ECS



Amazon S3 Bucket



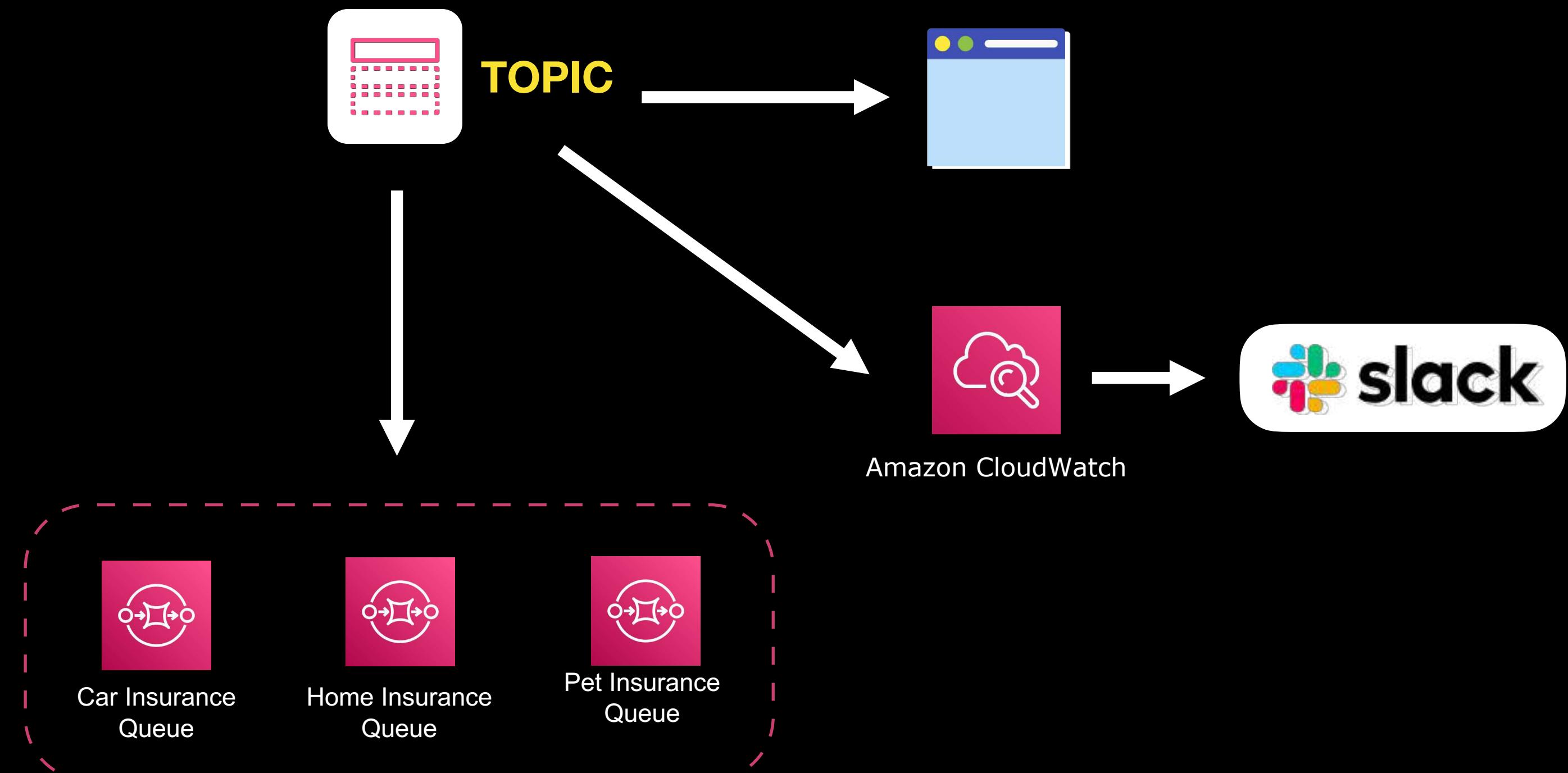
Amazon SQS

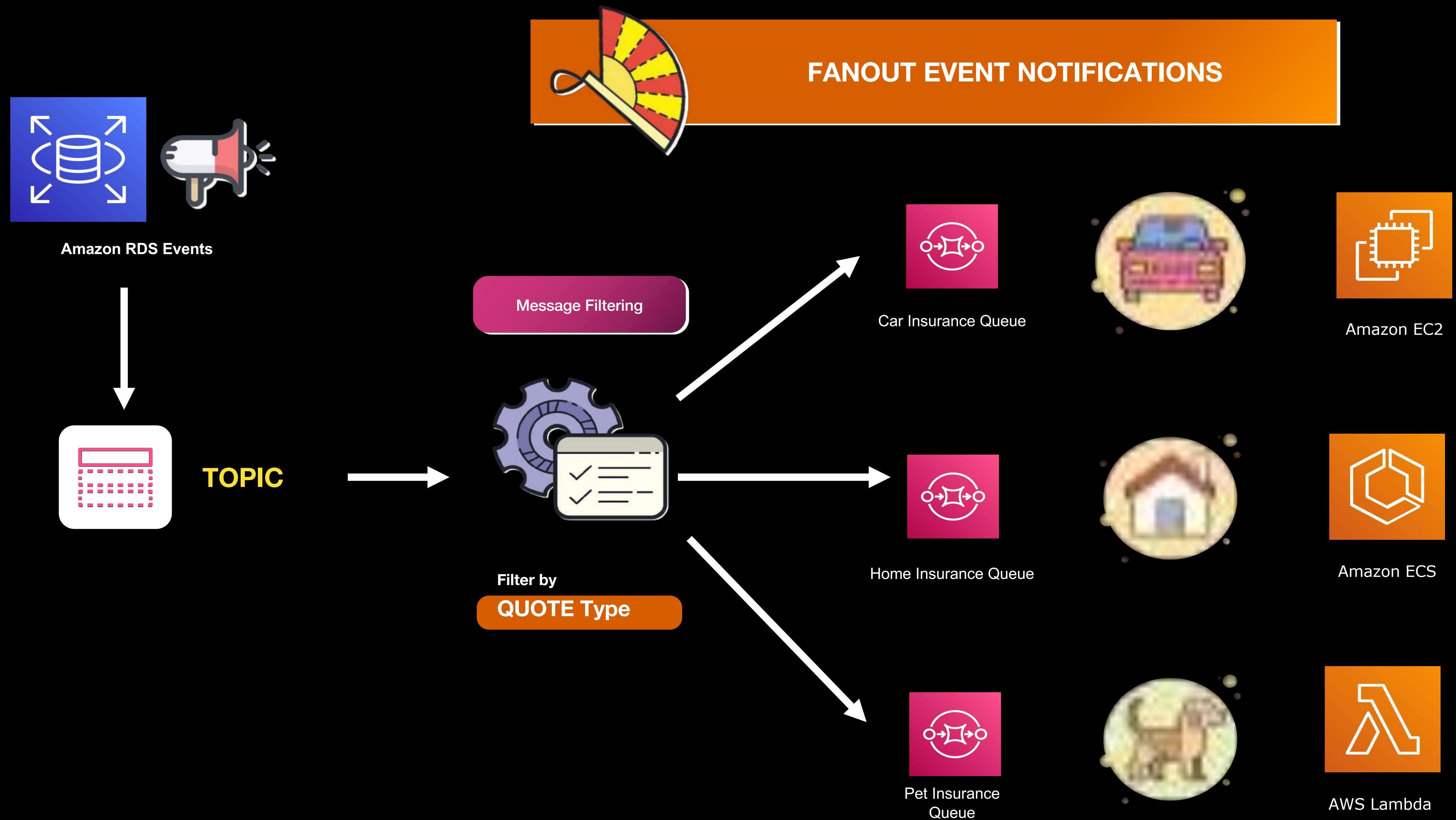


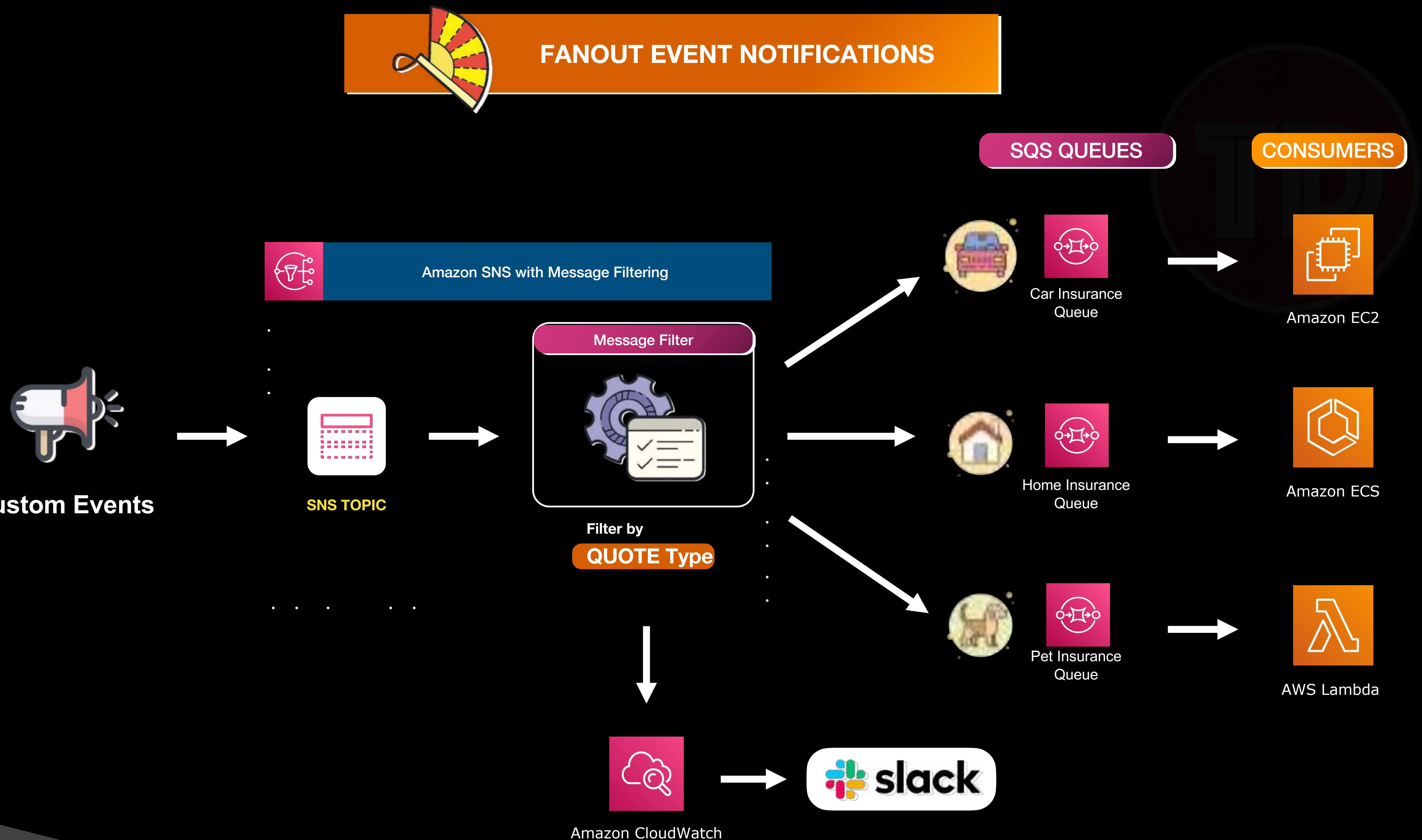
- A fully managed **messaging** and notification service
- Enables you to communicate between systems through publish/subscribe patterns or pub/sub messaging
- Messaging via mobile push, email, or SMS

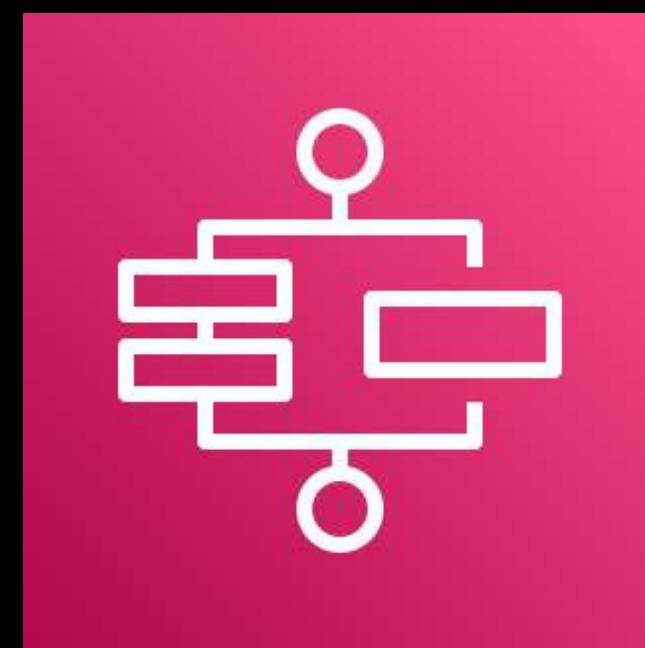


**Amazon Simple Notification Service (Amazon SNS)**







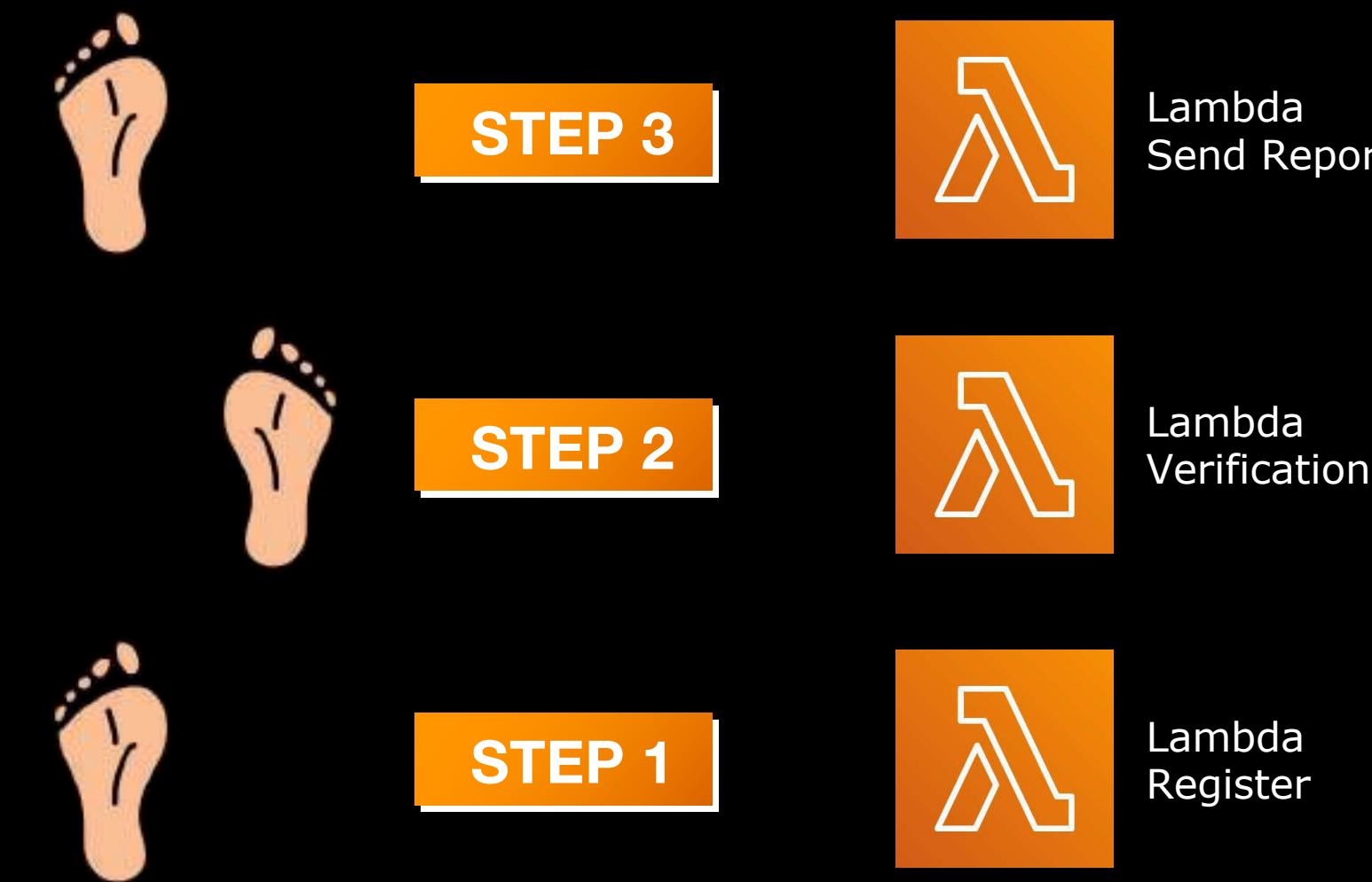


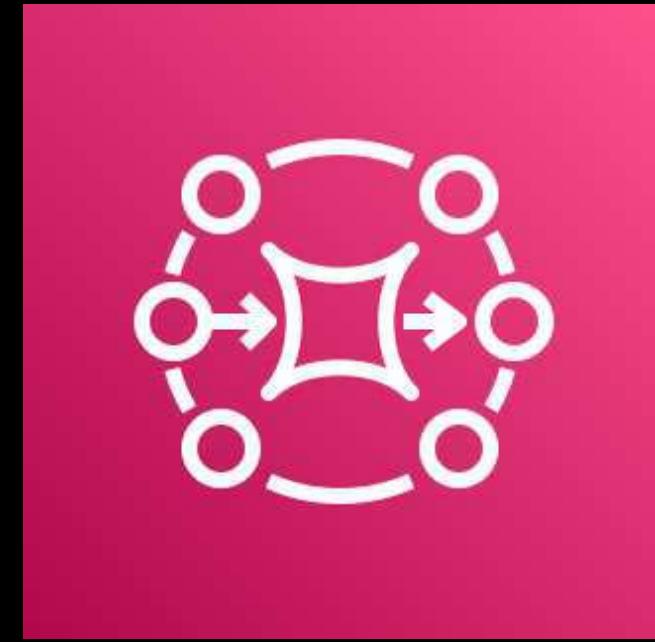
## AWS Step Functions

- A **serverless function orchestrator** for:
- Allows you to **orchestrate multiple AWS Lambda functions**, in order to achieve a specific workflow
- Enables you to create a **state machine** containing a combination of steps, activities and service tasks



AWS Lambda





## Amazon MQ

- A managed message broker service
- Uses the open-source  message broker
- The “MQ” in Amazon MQ stands for Message Queue, which is a form of asynchronous communication
- Works like  but supports more messaging protocol types
  - Supports Java Message Service (JMS), .NET Message Service (NMS), AMQP, MQTT, WebSocket and many others.

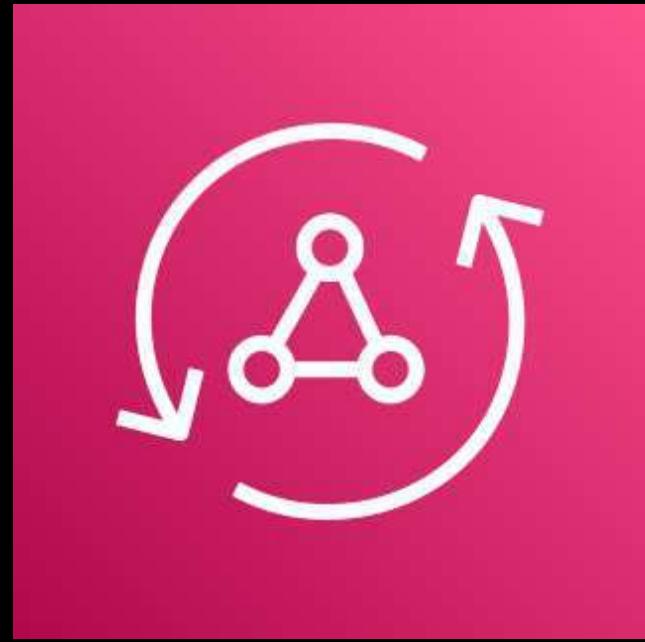
- A **serverless event bus service**
- Enables you to connect applications together using data from your own applications, Software-as-a-Service (SaaS) applications, and other AWS services.
- **Uses the same service API, endpoint, and** the underlying service infrastructure of:



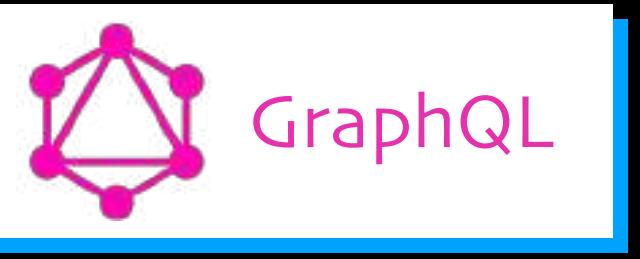
## Amazon EventBridge

- Recommended to be used for your own applications, 3rd party Software-as-a-Service apps, and other external sources
- Suitable for building **event-driven applications**



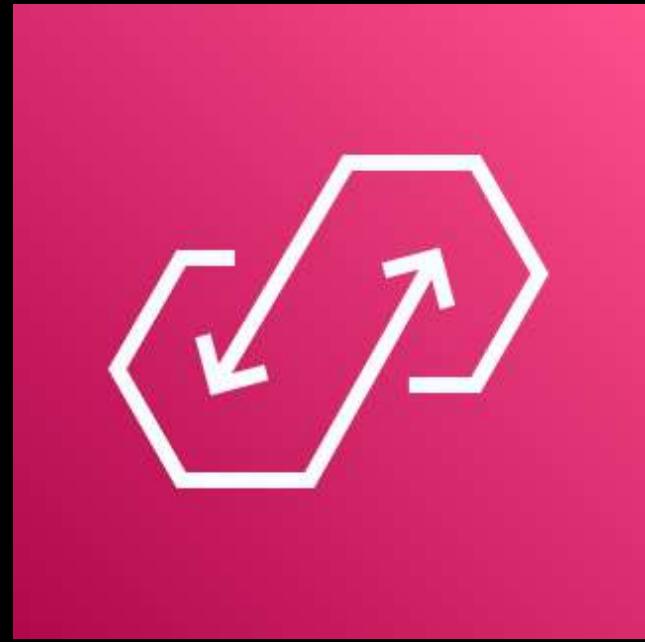


## AWS AppSync

- A managed service that uses  GraphQL
- GraphQL is a **data query language** that basically allows you to query your REST APIs
- Has different types of **schema**

<b>QUERY</b>	Read Data
<b>MUTATION</b>	Write Data
<b>SUBSCRIPTION</b>	Download/Upload Data

- Only fetches the data that you want and not the entire data set
- Unlike REST API, you can query different APIs or resources easily using a single API call
- Uses a **Resolver** which populates the data in your schema
- Simplifies application development by easily integrating GraphQL with your applications



## Amazon AppFlow

- A fully managed integration service
- Enables you to **securely transfer data between various systems** such as your Software-as-a-Service (SaaS) applications and different AWS Services
- Supports different SaaS apps such as Salesforce, Marketo, Slack, ServiceNow and many more
- Can be integrated with other AWS services
- Allows you to run your data flows on-demand, by schedule or as a response to a business event
- Provides you with powerful data transformation capabilities like filtering and validation

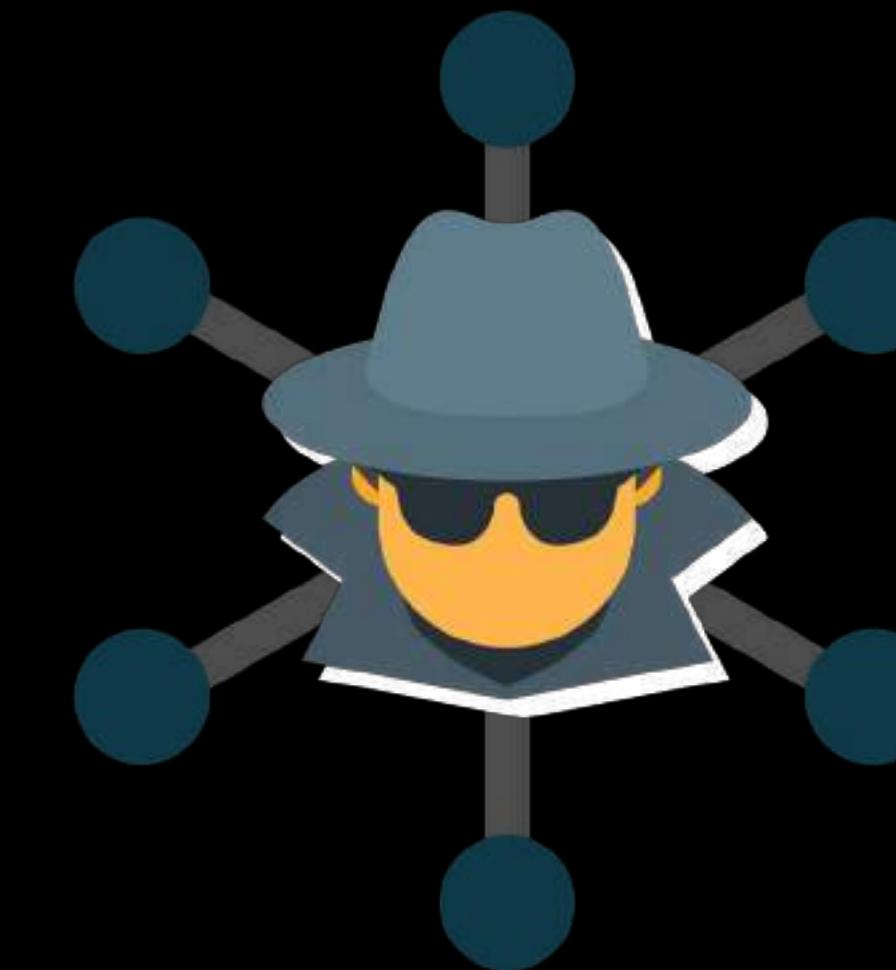
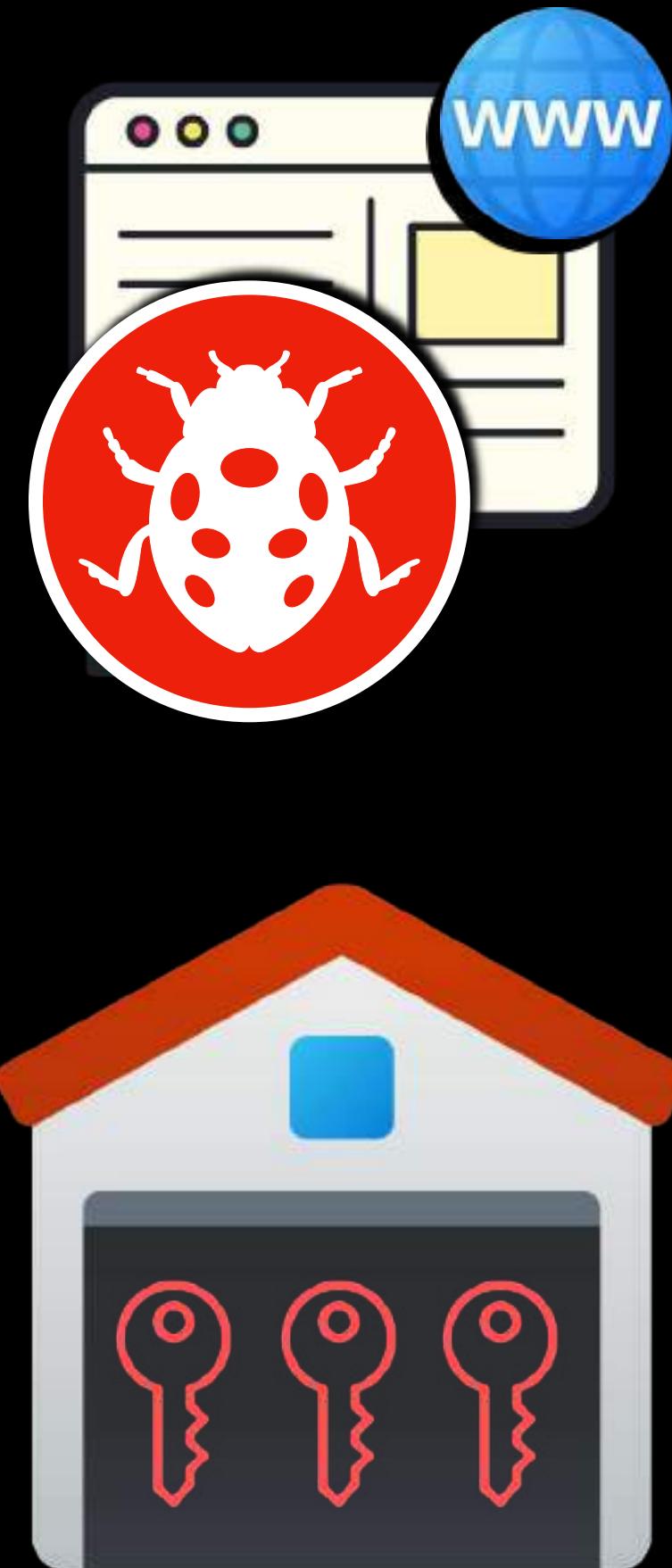


# AWS Security Services Overview

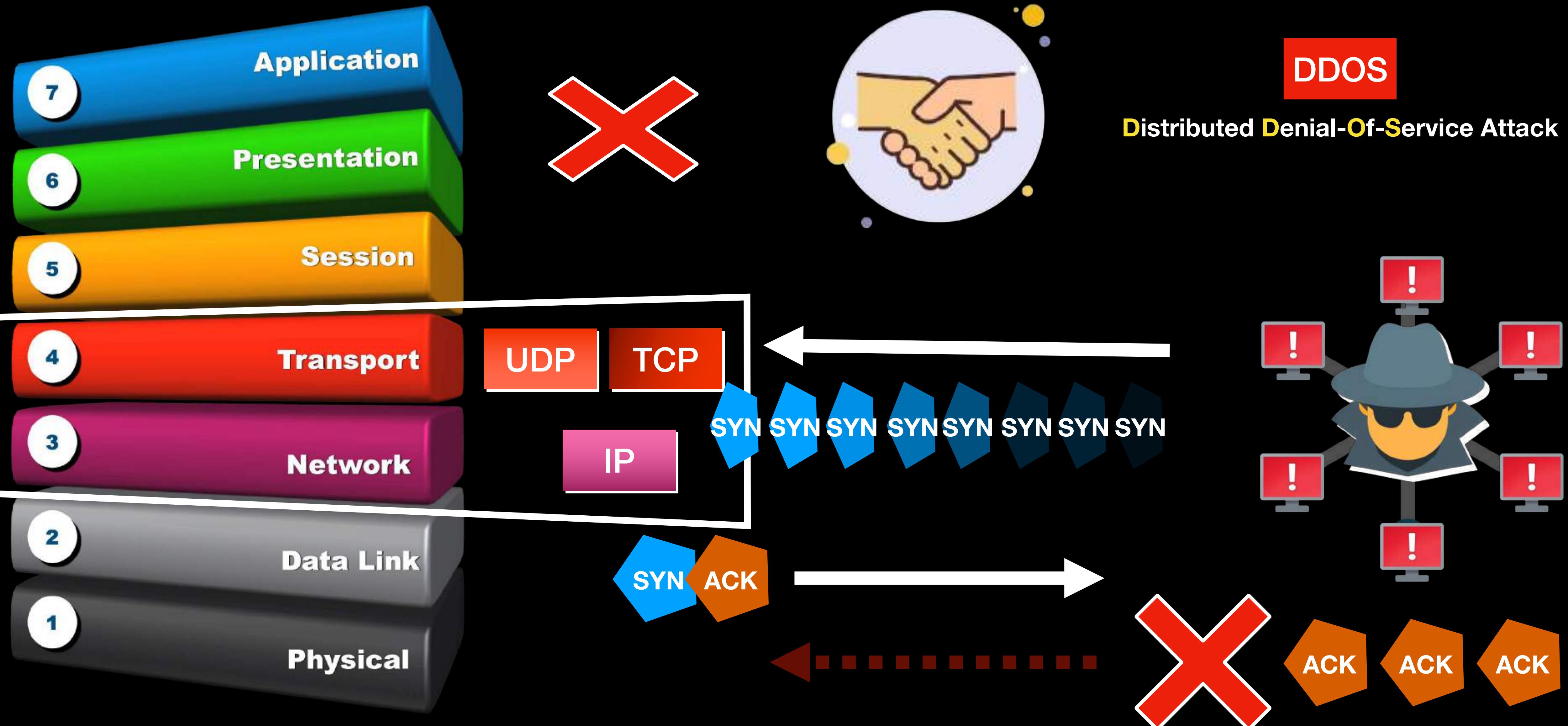
---



## AWS Security Services



# 7 Open Systems Interconnection (OSI) Model Layers





## AWS Security Services



AWS Web Application Firewall (AWS WAF)



AWS Firewall Manager



AWS Shield



Amazon GuardDuty



AWS CloudHSM



AWS Key Management Service (AWS KMS)



AWS Secrets Manager



AWS Certificate Manager (AWS ACM)



Amazon Macie



Amazon Inspector



Amazon Detective



## AWS Web Application Firewall (AWS WAF)

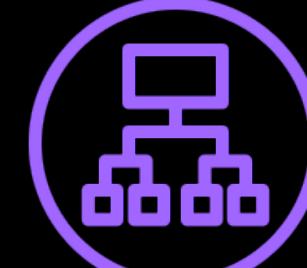
- A **web application firewall** service
- Protects your web applications from **common web exploits**
- Allows you to create **custom rules** that block **common attack patterns** such as:



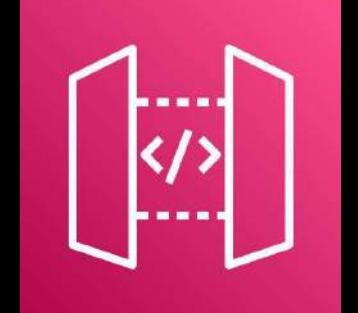
- **Can be integrated with:**



Amazon CloudFront



Application Load  
Balancer

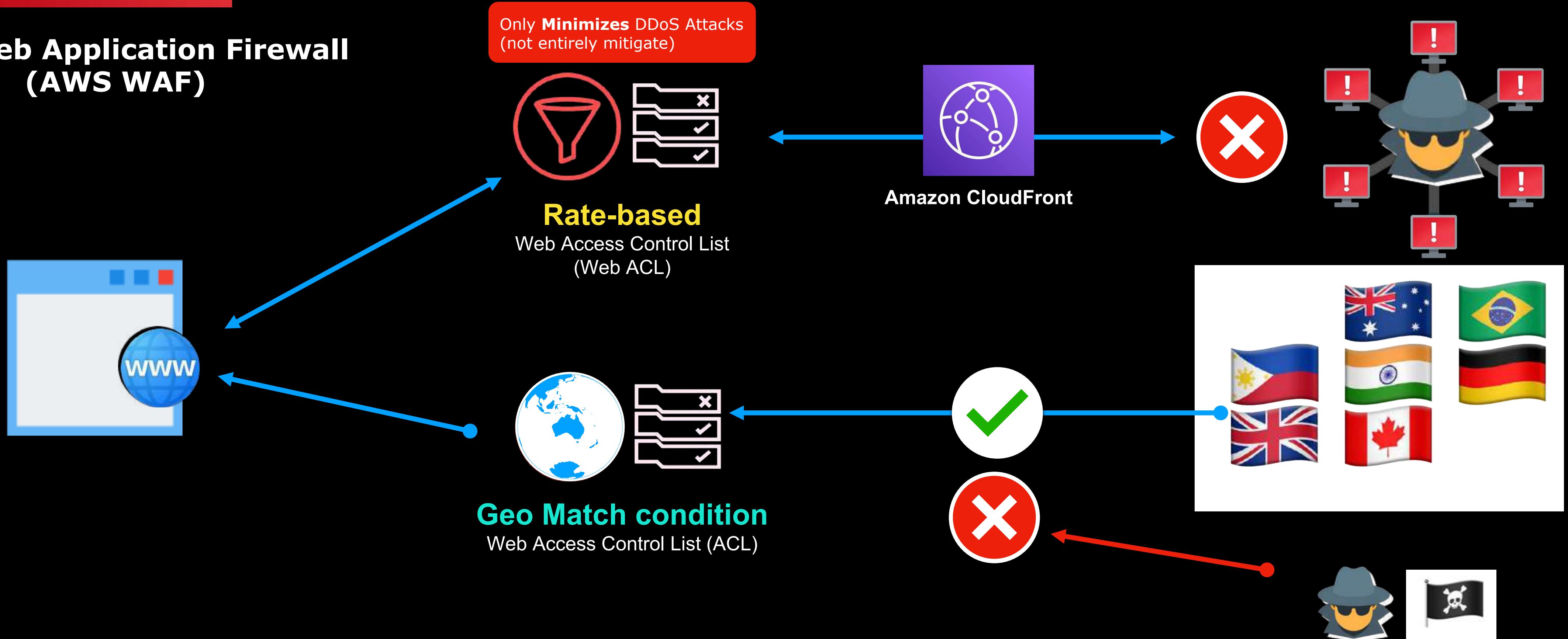


Amazon API  
Gateway



## AWS Web Application Firewall (AWS WAF)

- Has an **IP Match condition** feature, you can block malicious requests from a recurring set of IP addresses.
- Can protect your application from **illegitimate requests** sent by illegitimate external systems, through its **rate-limiting rule**.



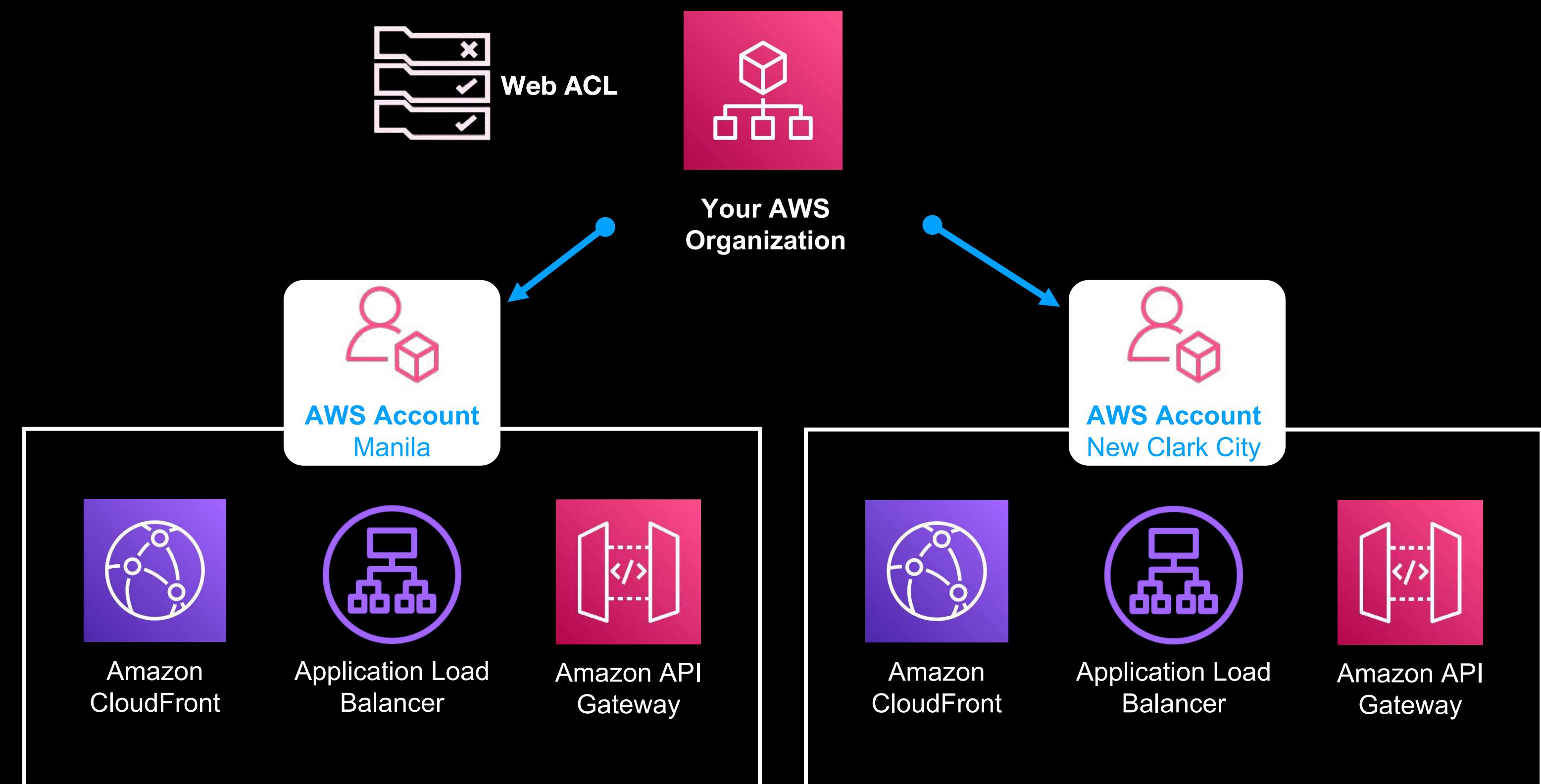


AWS WAF Rules

- A **security management service** designed for:
- Allows you to **centrally configure and manage WAF rules across multiple AWS accounts** and applications.
- Enables you to roll out your custom rules to your **AWS Organization**



## AWS Firewall Manager

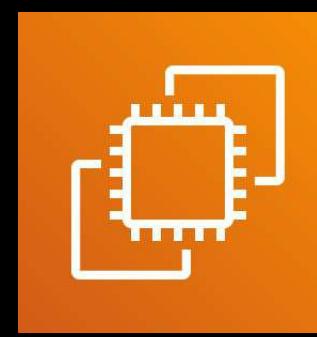


- A **managed DDoS protection** service
- Provides **detection and automatic mitigations** that minimize application downtime and latency.
- **Mitigate different types of flood attacks** such as UDP reflection, SYN flood, DNS Query flood, and HTTP flood attacks.



## AWS Shield

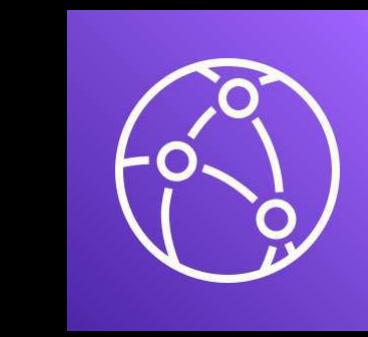
- Protects your applications that use:



Amazon EC2



Elastic Load  
Balancer



Amazon  
CloudFront



AWS Global  
Accelerator



Amazon  
Route 53

- Two Tiers:

- **Standard**

- Built-in by default
    - No extra charge

- **Advanced**

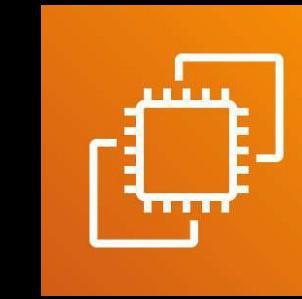
- Has an additional charge
    - Provides access to real-time DDoS attack notification
    - **DDoS Response Team (DRT)** supports you during DDoS Attack



## Amazon GuardDuty

- A **managed threat detection** service
- **Identifies malicious or unauthorized activities** in your AWS accounts and workloads.
- **Monitors activities** such as unusual API calls, cryptocurrency mining, or potentially unauthorized deployments that indicate a possible account compromise.

- Also detects potentially compromised:



Amazon EC2 Instances

- Produces security reports called:



Findings

- Able to **send notifications using CloudWatch Events** when a change was detected
- **NOT capable of doing any resource changes** by itself, like rate-limiting protection or DDoS attack mitigation.



**AWS CloudHSM**



**AWS Key Management  
Service (AWS KMS)**

- A fully managed, **cloud-based hardware security module** or HSM.
- The **HSM** in Cloud**HSM** means:

**Hardware Security Module**



**AWS CloudHSM**



- Enables you to easily **generate and use your own encryption keys**.
- Encryption keys can be in 128-bit or 256-bit

# HSM

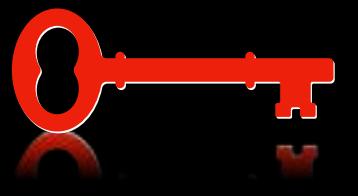
Hardware Security Module



AWS CloudHSM



- A **physical hardware device**
- Performs **cryptographic operations**
- Securely stores cryptographic **key material**



## Leading HSM Providers

THALES



yubico

utimaco®



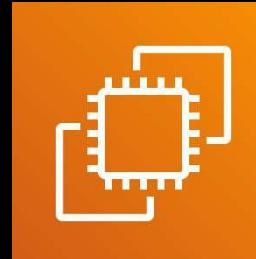
Hewlett Packard  
Enterprise

- A random, **Base64 or hexadecimal string**
- **Binary** format ( .bin )
- Used by your **encryption key**.





## AWS CloudHSM

- The **CloudHSM clients** is installed and **hosted in your:**  Amazon EC2 Instances
- The HSM cluster is **deployed in your:**  Amazon VPC
- **Single Tenant** — Only used by one tenant or user (you) 
- Can be used to:
  - Offload SSL Processing
  - Enabling Transparent Data Encryption (TDE) for Oracle databases
  - Protecting the private keys for an Issuing Certificate Authority (CA).
- Integrate CloudHSM and  AWS KMS to create a **custom key store.**

AWS KMS



## AWS Key Management Service (AWS KMS)

- A managed service that **works like**:
- Internally, it **also uses hardware security modules (HSMs)** for creating and controlling your encryption keys.
- Has **multi-tenant access**

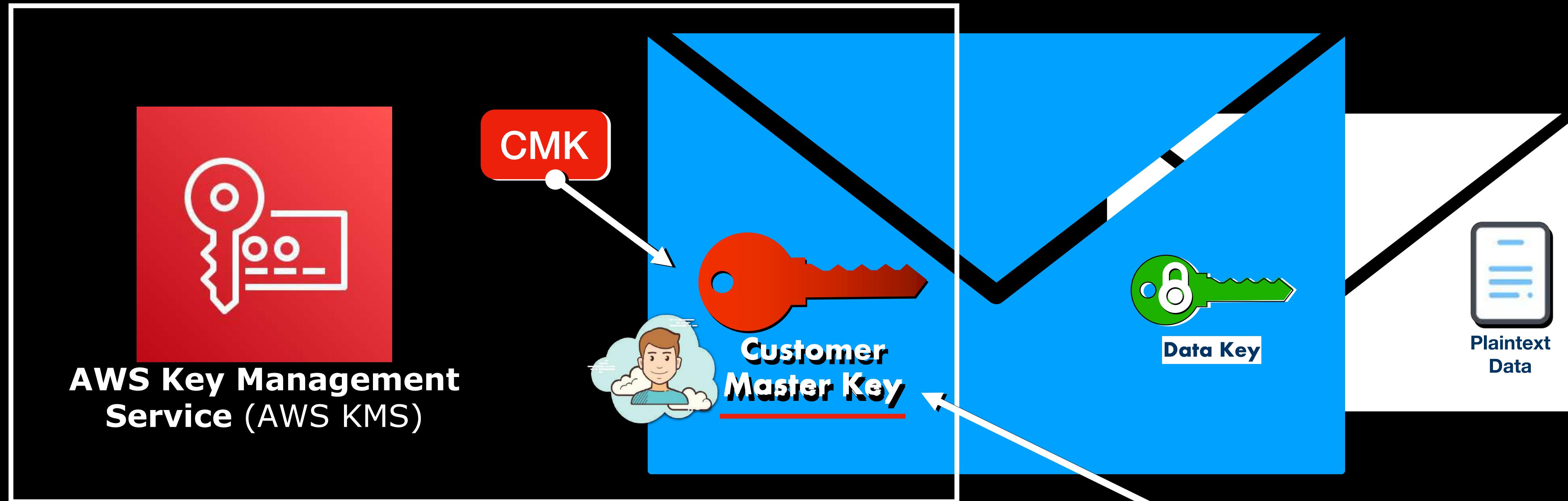
You **share** the HSM with other tenants or AWS customers


- Unlike CloudHSM, you **cannot launch the HSM to Amazon VPC or EC2 instances** (as clients with direct HSM access) that you own.
- Can be integrated with other AWS services to help you protect the data you store with these services.





## ENVELOPE ENCRYPTION



- AWS KMS **automatically rotates** your CMK



AWS CloudHSM

- You can also create a **custom key store** in AWS KMS with



**AWS Key Management Service (AWS KMS)**



- Provides **complete control** over your **encryption key lifecycle management**
- Allows you to **remove the key material** of your encryption keys.



- You can **audit key usage independently** of:



AWS CloudTrail



AWS KMS



AWS Secrets Manager

- Protect the **secrets** of your applications, services, and IT resources.
- Enables you to easily **rotate, manage, and retrieve** your secrets
- A **secret** can be:
  - A database password
  - API key
  - Authentication token
  - Other sensitive data
- Eliminates hardcoded sensitive information in plain text in:
- Offers **secret rotation** with built-in integration for:



Amazon RDS



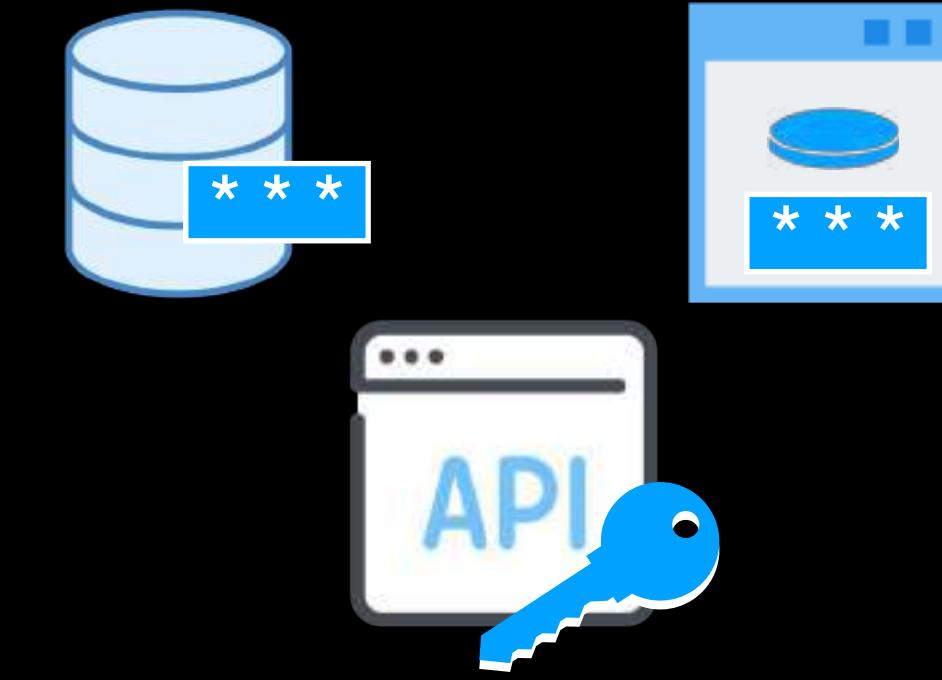
Amazon Redshift



Amazon DocumentDB



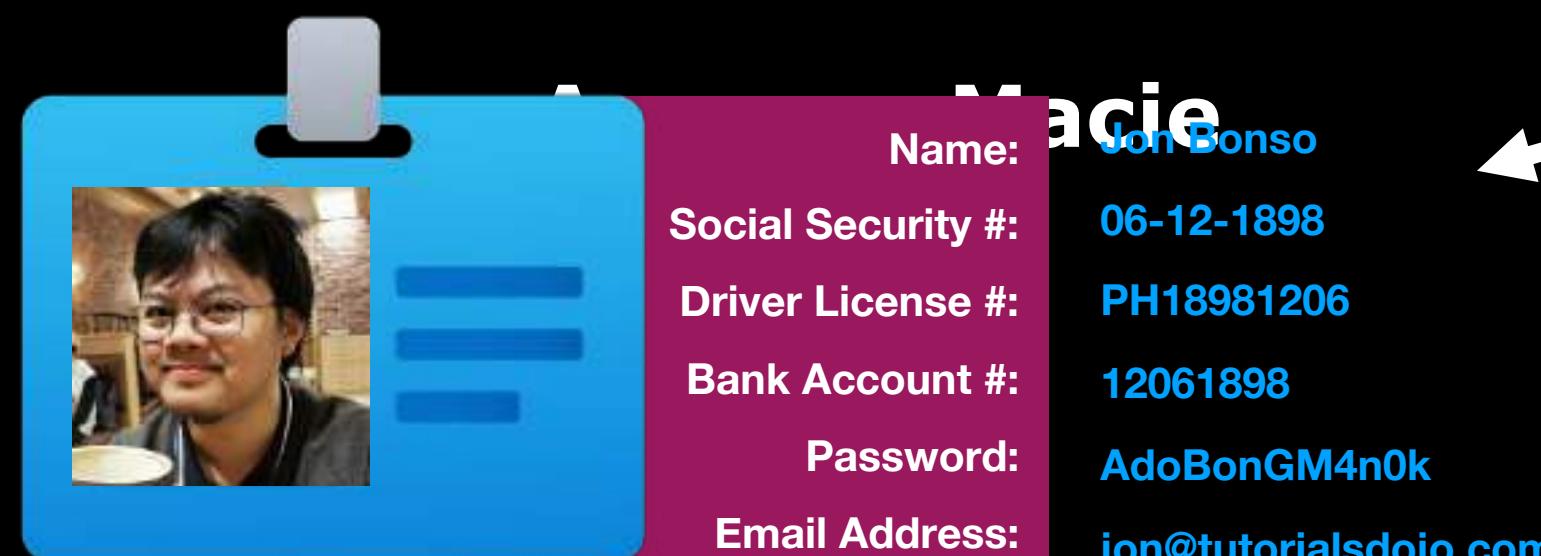
Other Services



AWS Lambda

- **Control access to secrets** using fine-grained permissions and centrally audit your secrets.
- **Not recommended** for storing encryption keys or key materials since it does not use an HSM

- A fully managed **data security and data privacy** service
- Automatically **recognizes and classifies sensitive data** or intellectual property
- **Uses machine learning** to automatically discover, classify, and protect sensitive data stored in your:



Amazon S3  
bucket



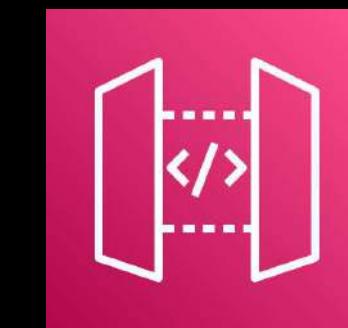
Other Services

- Recognizes sensitive data such as **personally identifiable information** or PII.
- **Provides dashboards and alerts** that give visibility into how sensitive data is being accessed or moved.

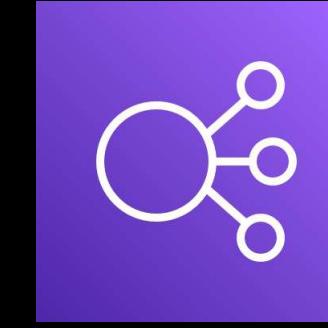


## AWS Certificate Manager (AWS ACM)

- Provisions, manages, and deploys public and private Secure Sockets Layer/Transport Layer Security **(SSL/TLS) certificates**
- Enables you to **create private certificates** for your internal resources and **manage the certificate lifecycle centrally**
- SSL Certificates are **free of charge for ACM-integrated services** such as:



Amazon API  
Gateway

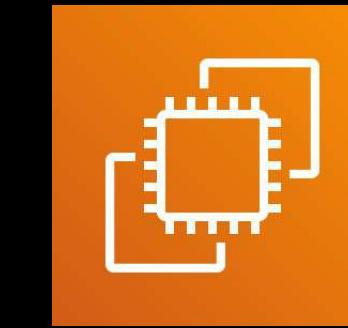


Elastic Load  
Balancing



## Amazon Inspector

- An **automated security assessment** service
- Improves the security and compliance of applications deployed on your AWS cloud infrastructure
- **Automatically assesses applications for vulnerabilities** or deviations from best practices.
- Produces a detailed list of security findings prioritized by level of security risk severity
- Provides an **automated security assessment report** that will identify unintended network access to your:



Amazon EC2 Instances

- The detailed assessment reports are available via the Amazon Inspector console or API

- Helps you **detect the root cause of your security issues** easier
- It analyzes, investigates, and quickly identifies the potential security issues or suspicious activities in your AWS infrastructure
- **Automatically collects log data from various AWS resources** such as:



**Amazon Detective**



AWS CloudTrail



VPC Flow Logs



GuardDuty Findings

- **Uses machine learning** to analyze and conduct security investigations.



# AWS Management & Governance Services Overview

---



# AWS Management & Governance Services



**S O P**

Standard Operating Procedures



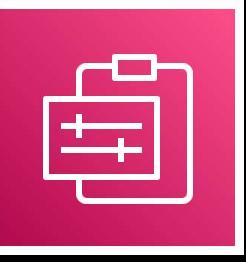
**H I P A A**

Health Insurance Portability and  
Accountability Act of 1996



**G D P R**

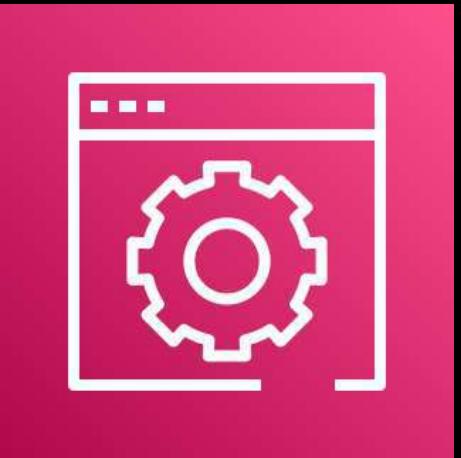
General Data Protection Regulation



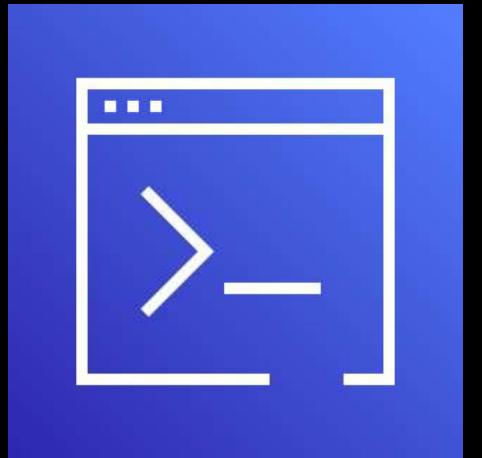
# AWS Management & Governance Services



– control resources



**AWS Management Console**



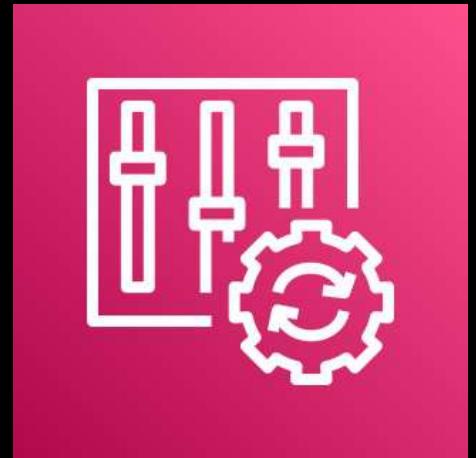
**AWS Command Line Interface (AWS CLI)**



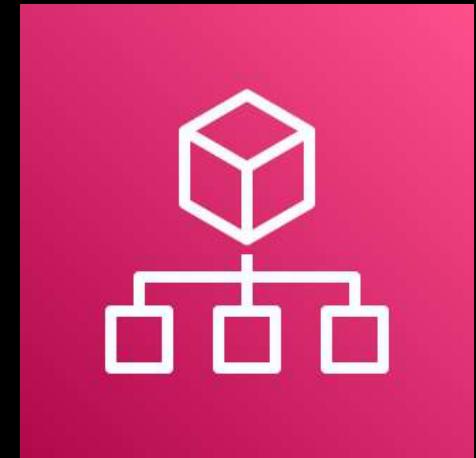
**AWS Console Mobile Application**



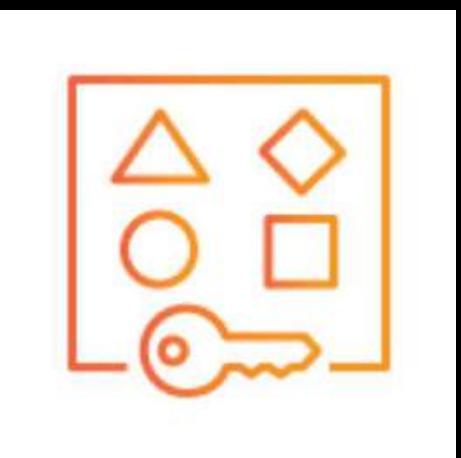
– enforce standards  
– ensure compliance



**AWS Config**



**AWS Organizations**



**AWS Resource Access Manager**



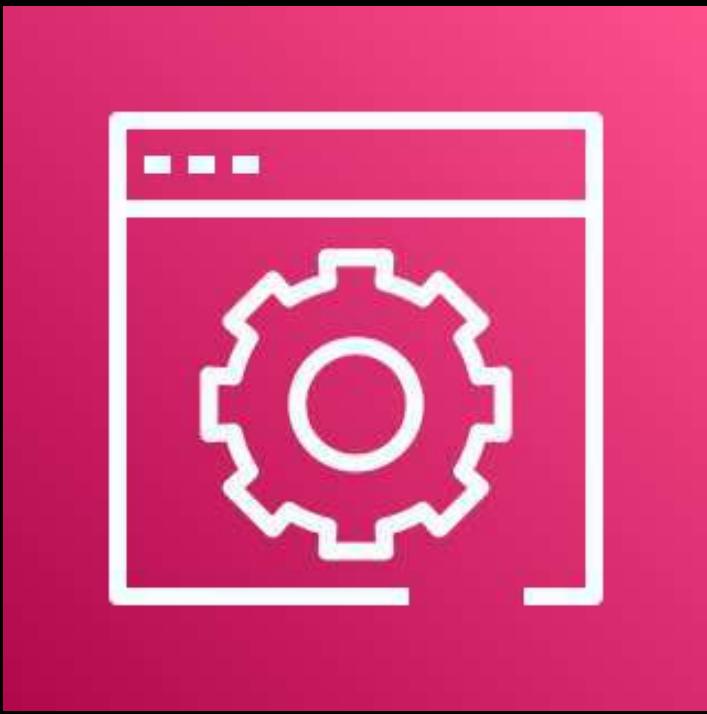
**AWS Systems Manager (SSM)**



**AWS Service Catalog**



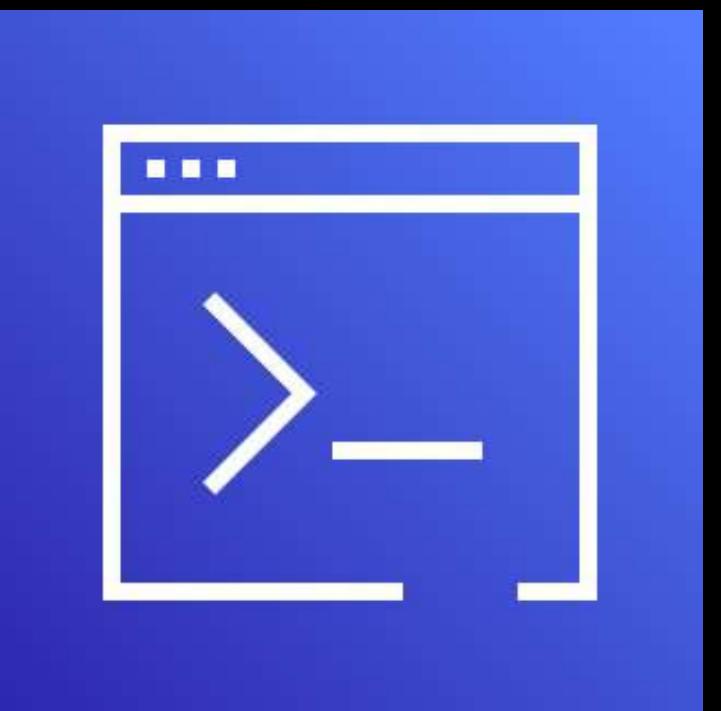
**AWS Control Tower**



## AWS Management Console

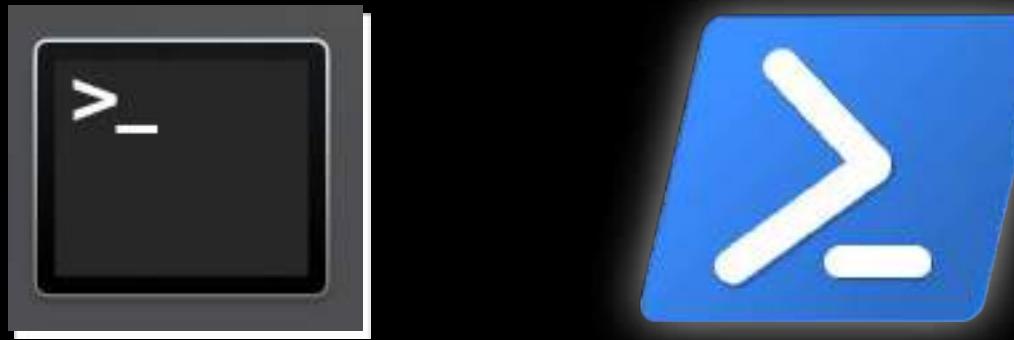
- A **web interface** to control your AWS resources
- Accessible through your **web browser**
- Log in using your IAM username and password
- Supports **Multi-Factor Authentication** (MFA)
- Accessible via this URL: <https://console.aws.amazon.com>

- A **command-line interface** to control your AWS resources

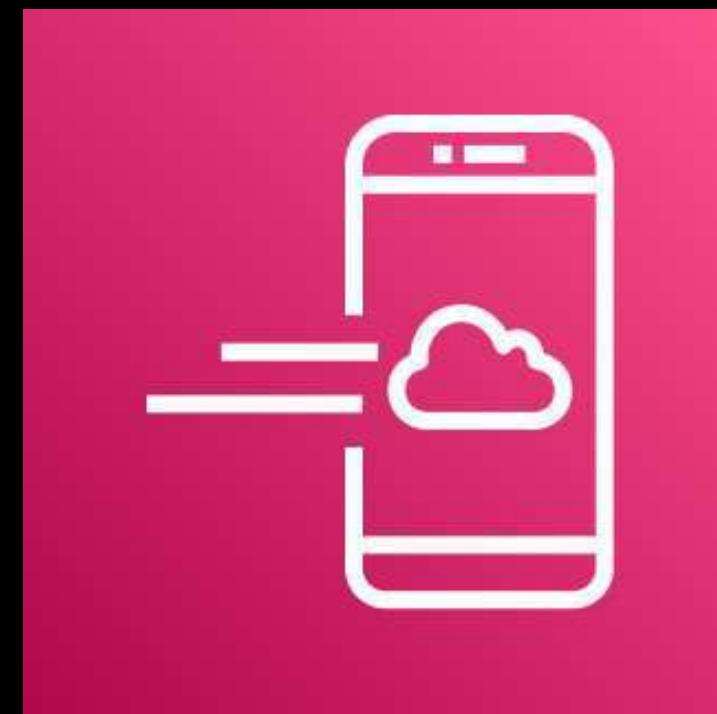


## AWS Command Line Interface (AWS CLI)

- Accessible through your terminal, command prompt or Windows PowerShell

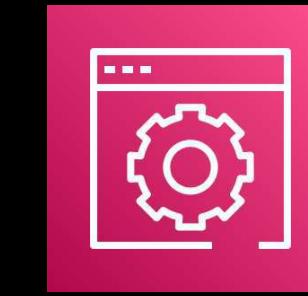


- Allows you to **develop custom shell scripts** that invoke different AWS CLI commands

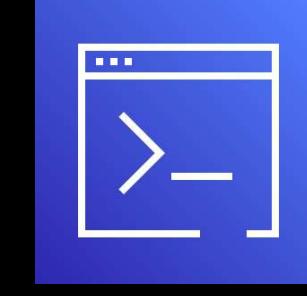


## AWS Console Mobile Application

- The **official mobile app** provided by Amazon Web Services
- Allows you to monitor your resources through a dedicated dashboard
- Enables you to view your configuration details, metrics, and alarms of **select AWS services** (not all services) on your mobile device
- Provides an overview of the account status, real-time CloudWatch metrics, Personal Health Dashboard, and AWS Billing
- Has **limited capabilities** compared with:

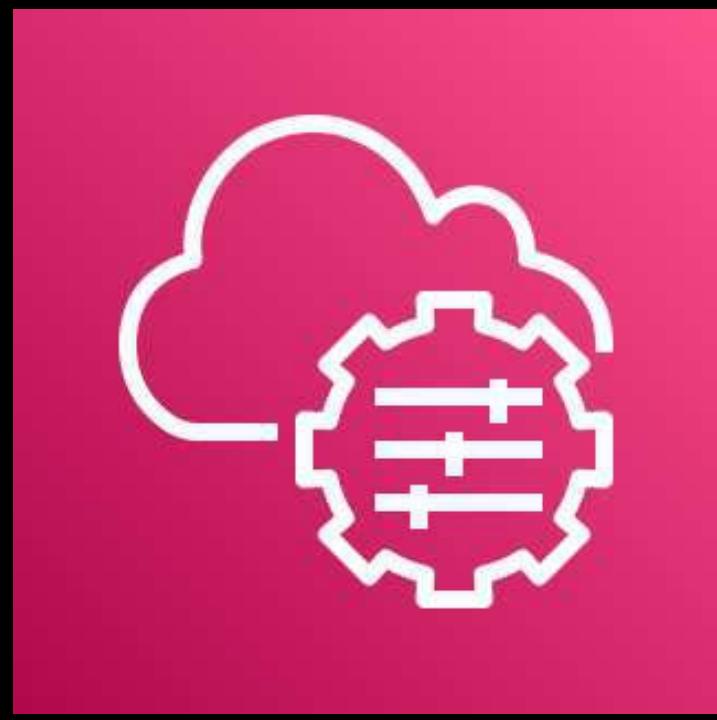


AWS Management  
Console

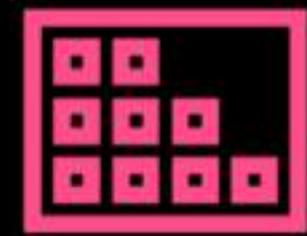


AWS CLI

- A **suite of services** that allows you to manage your resources
- Allows you to **control both of your AWS Cloud and on-premises** infrastructure
- Composed of:



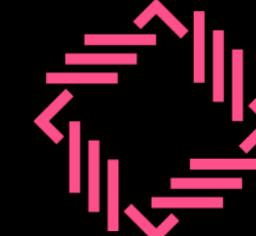
**AWS Systems Manager**  
(SSM)



Session Manager



State Manager



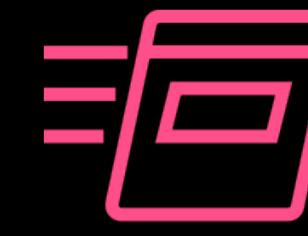
Patch Manager



Automation



Maintenance  
Windows



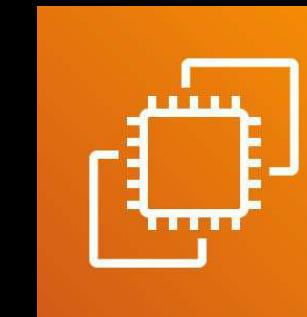
Run Command



Parameter Store



Others



Amazon EC2  
Instances

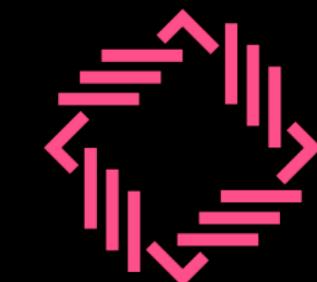


On-premises  
Servers

- Also has an **SSM agent** that you can install on your EC2 instances or on-premises servers to centrally manage your resources



## AWS Systems Manager (SSM)



Patch Manager



State Manager



Parameter Store



## STATE

- Installed softwares (e.g. startup script, antivirus etc)
- Server configurations
- Firewall settings
- Associate Ansible playbooks, Chef recipes, PowerShell modules, and other SSM Documents

## PARAMETER

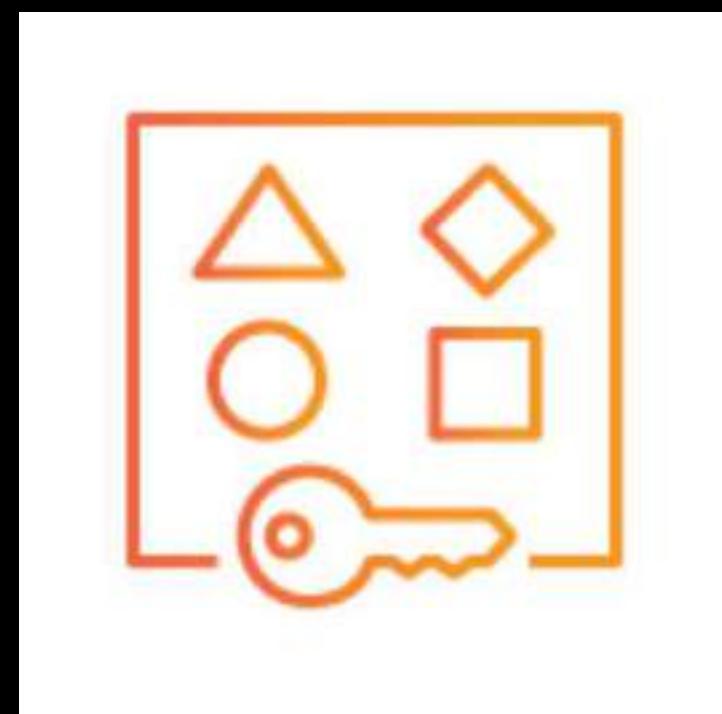
- Passwords
- Database Strings
- Amazon Machine Image (AMI) IDs
- License Codes
- Environment Variables

Secure String

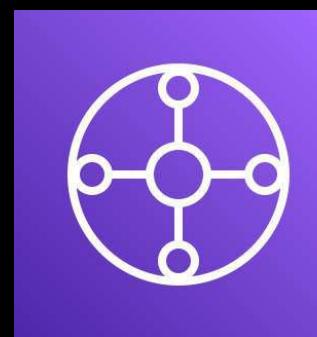


AWS KMS

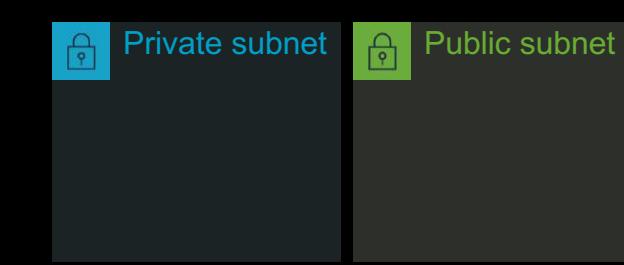
- Enables you to easily and **securely share your AWS resources** with any AWS account or within your AWS Organization
- Allows you to share:



**AWS Resource Access Manager (AWS RAM)**



AWS Transit Gateway



Subnets



AWS License Manager

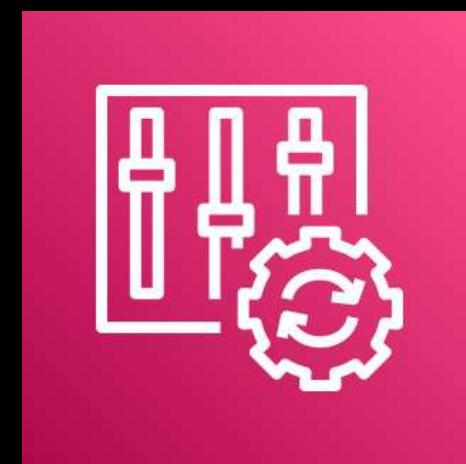
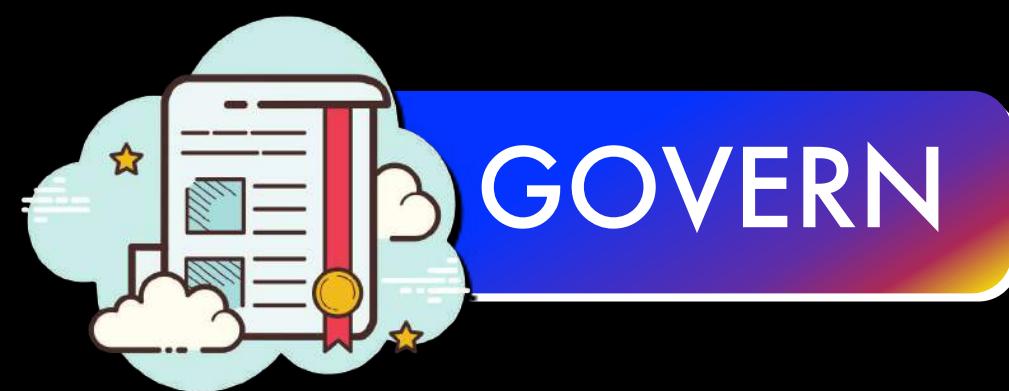


Amazon Route 53 Resolver

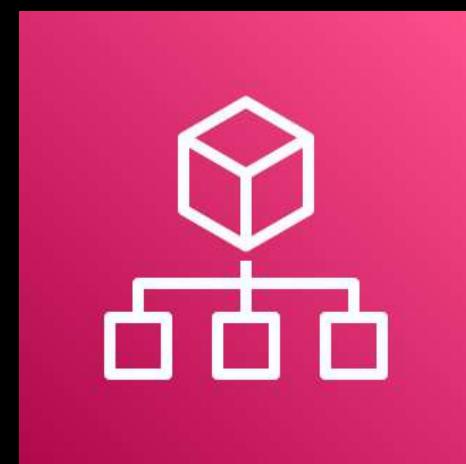


Other AWS Resources

- **Eliminates the need to create duplicate resources** in multiple accounts
- **Reduces the operational overhead** of managing multiple resources in each and every single account you own.



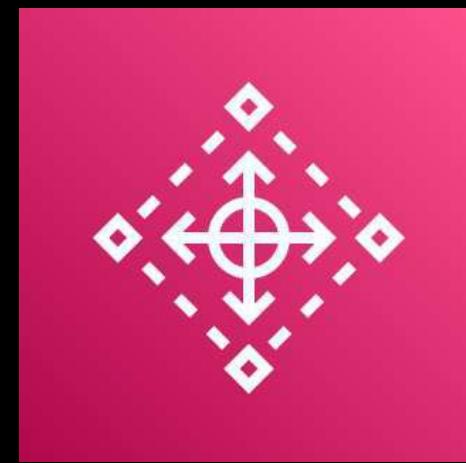
**AWS Config**



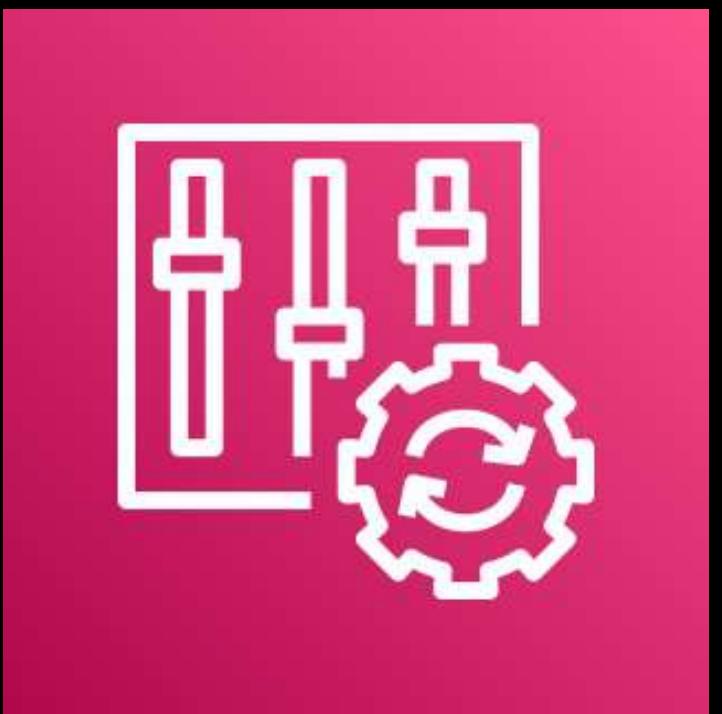
**AWS Organizations**



**AWS Service Catalog**



**AWS Control Tower**

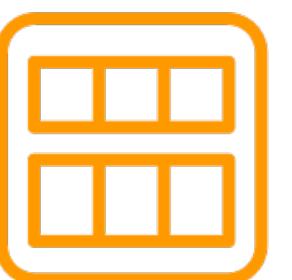


## AWS Config

- Enables you to **assess, audit, and evaluate the configurations** of your AWS resources
- Automates your **compliance assessment** process
- **Provides visibility on the existing configurations** of your various AWS services and third-party resources (such as your on-premises servers)
- Enables you to **identify the changes** made to a specific resource over time



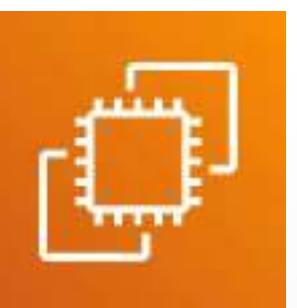
## RESOURCES



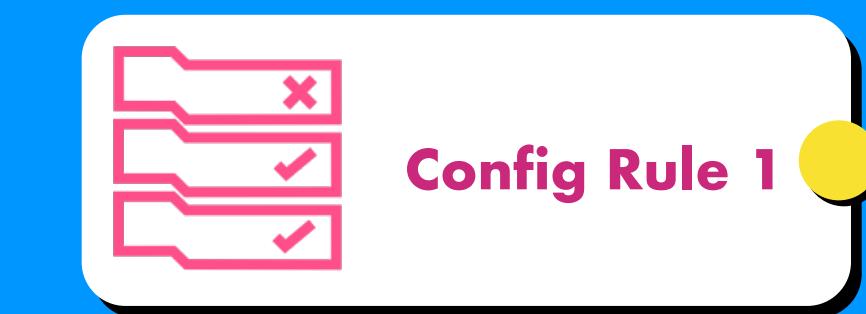
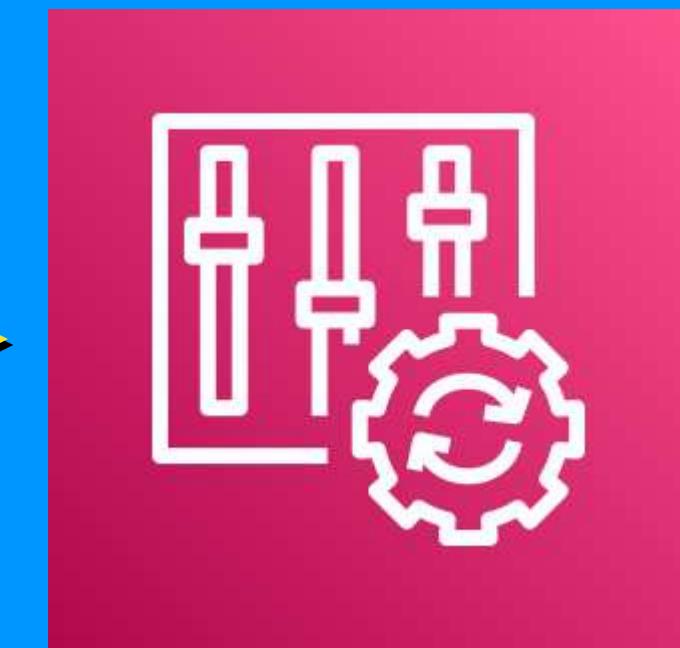
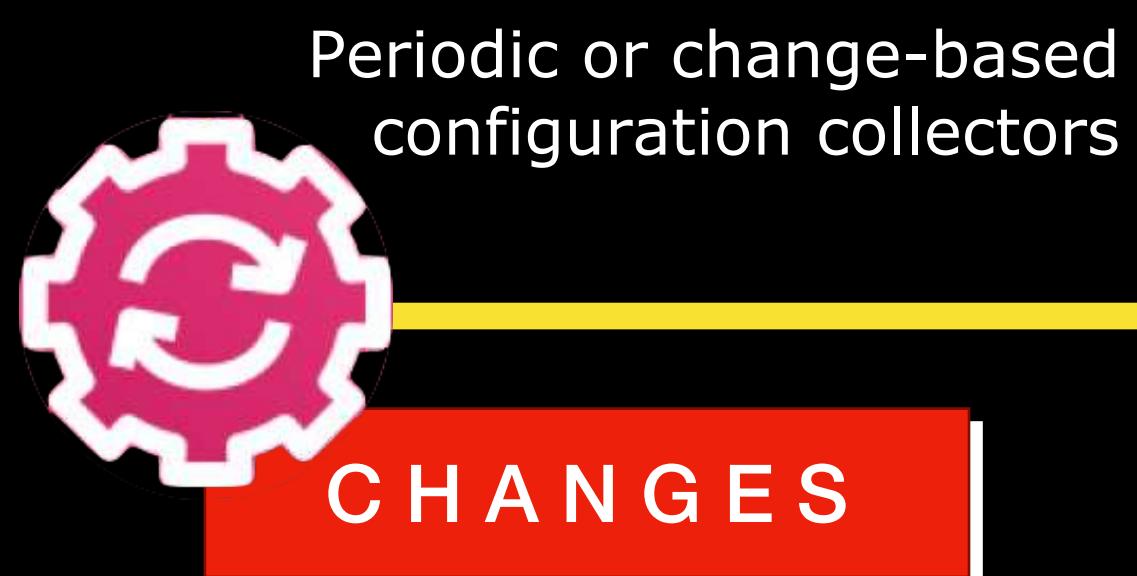
AMI



S3 Bucket

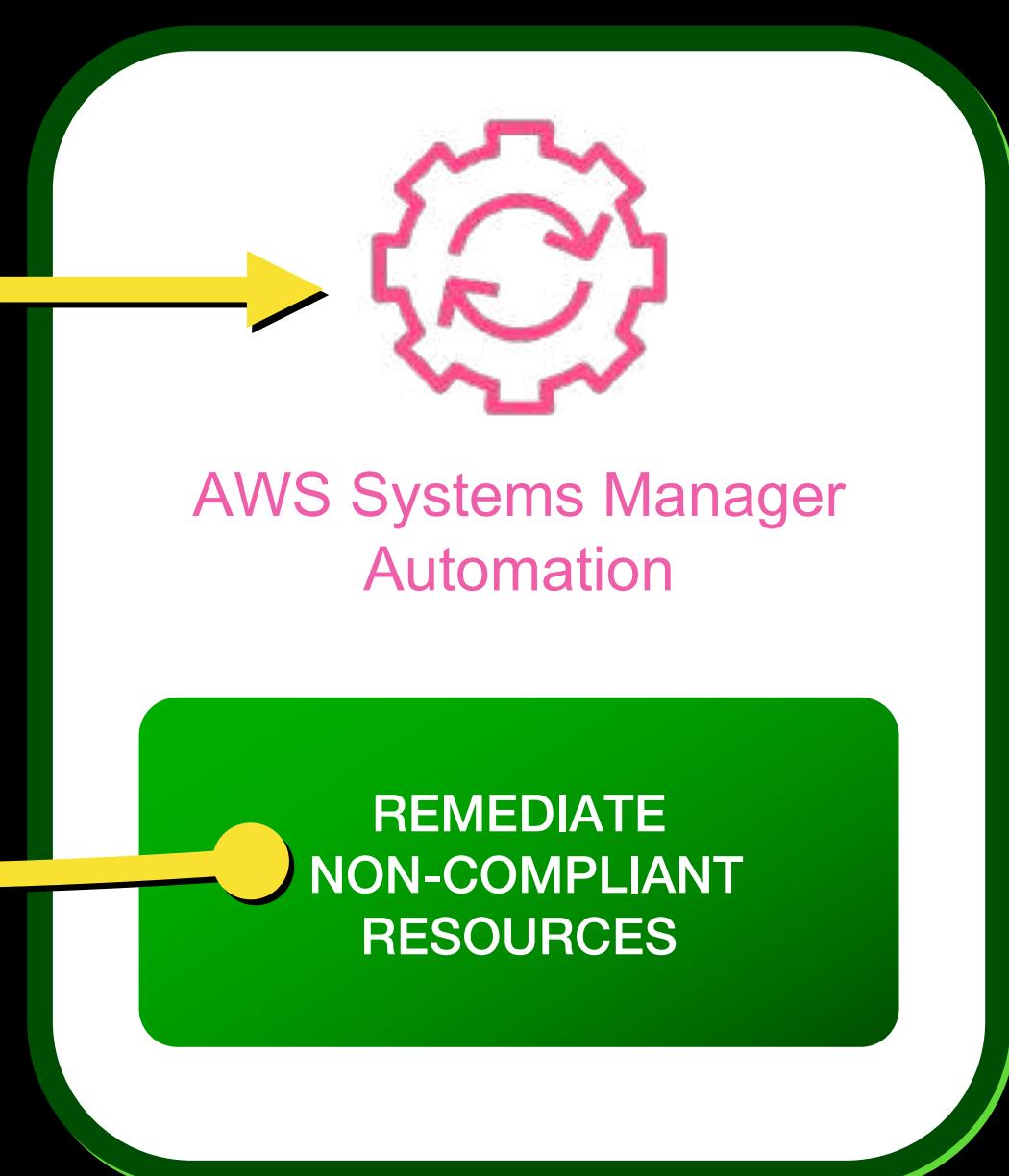
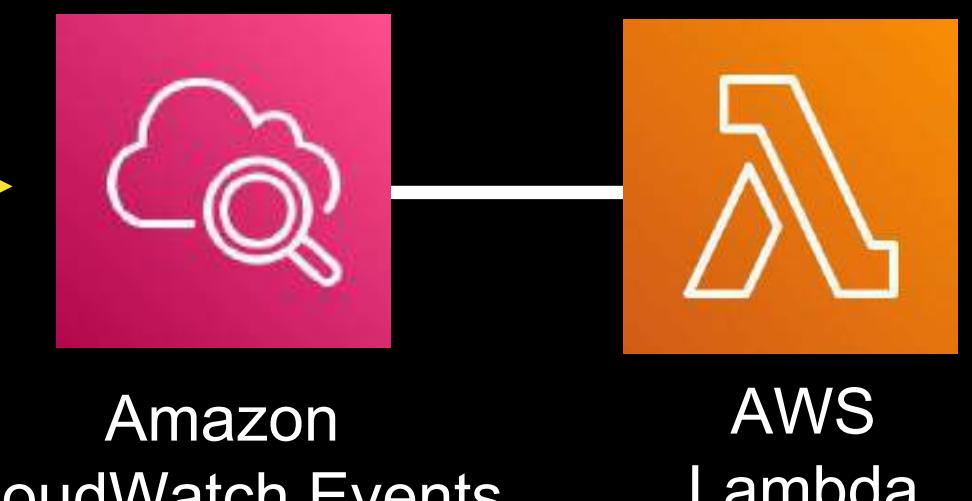


EC2 Instance



NOTIFICATION

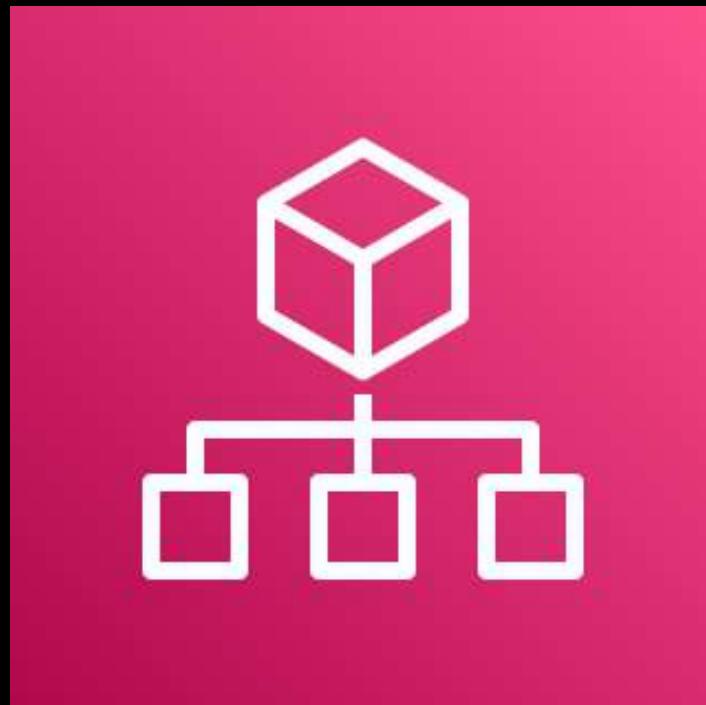
REMEDIATION



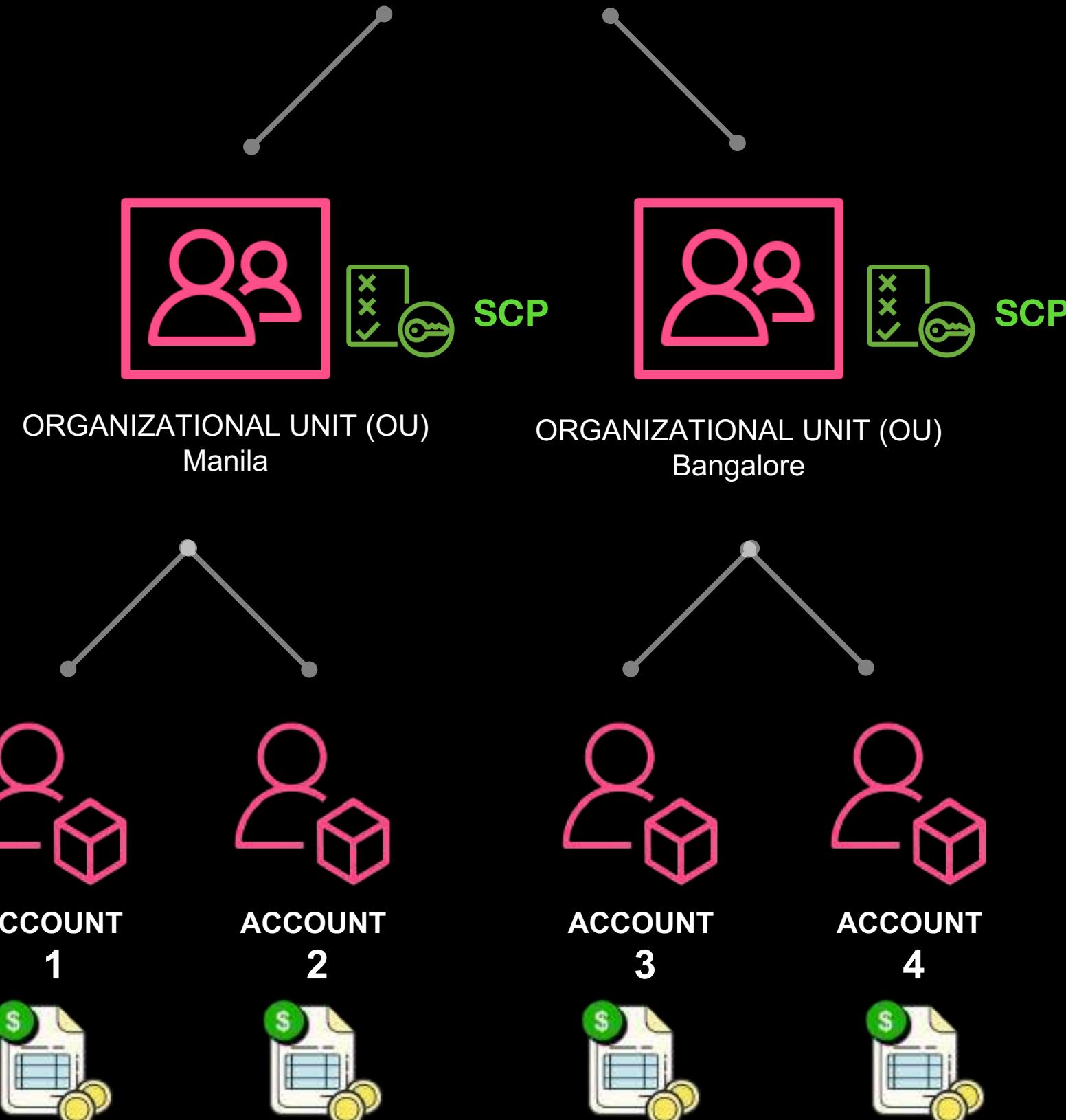
The AMI was shared to the AWS Marketplace

The bucket was set to public

The associated Elastic IP address was removed



## AWS Organizations



- **Consolidate and centrally manage** multiple AWS accounts

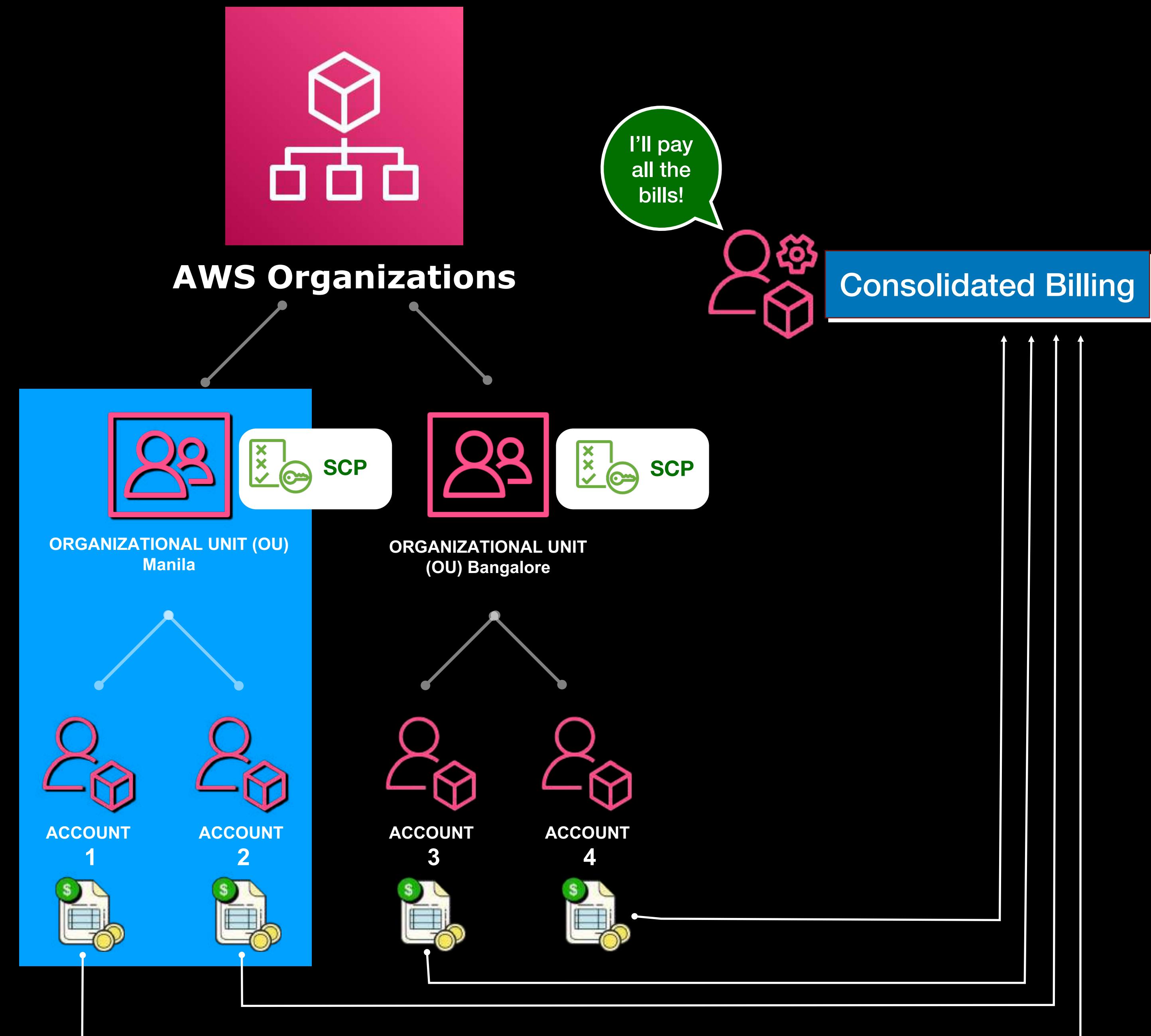


## Consolidated Billing

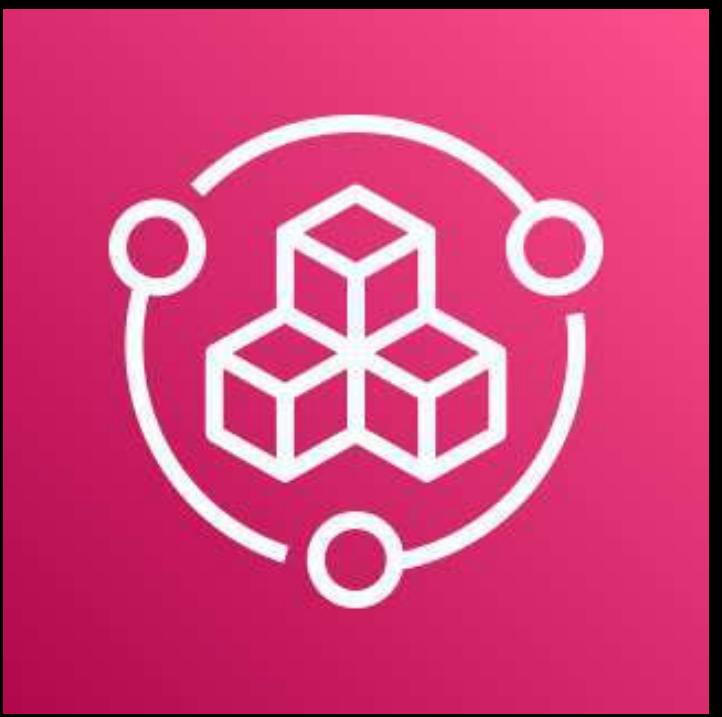
- Combines the bills of multiple AWS accounts
- Provides **volume discounts** to further lower down your costs

- Uses **Service Control Policies (SCP)** to control access and ensure organizational compliance across your AWS accounts
- Offers **Central Logging** to monitor all activities performed across your organization using AWS CloudTrail
- **Aggregate data from all your AWS Config rules** to quickly audit your environment for compliance.

A single AWS Organization can have two or more Organizational Unit (OU) and underlying AWS accounts with Service Control Policies (SCPs) attached



- Empowers you to **set up and centrally manage catalogs** of approved IT services
- Allows you to manage various IT services, referred to as "**products**" in Service Catalog then group them in a portfolio



PRODUCT

- Machine image (AMI)
- Application server
- Program
- Tool
- Database
- Other services

## AWS Service Catalog

- Assists you in meeting your compliance requirements
- **Enforce granular access control** to your resources



## AWS Control Tower

- Helps you **set up and govern a secure multi-account** AWS environment
- Automates the setup of your multi-account AWS environment
- Uses **blueprints** that follow AWS best practices for security and management
- Provides mandatory high-level rules called **guardrails**
- Help enforce your policies using service control policies (SCPs)
- **Detect policy violations** using AWS Config rules



# AWS Identity Services Overview

---

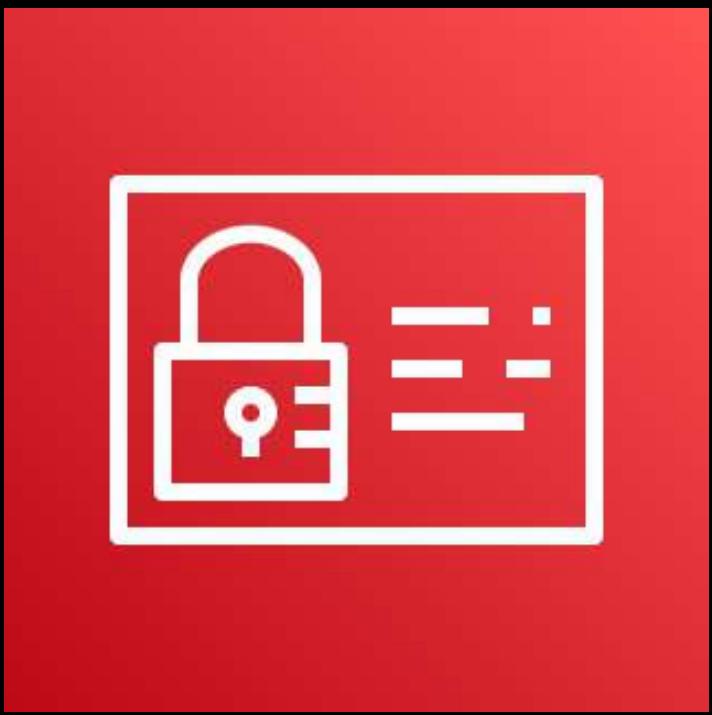


## AWS Identity Services





## AWS Identity Services



**AWS Identity & Access  
Management (IAM)**



**AWS Single Sign-On**



**Amazon Cognito**

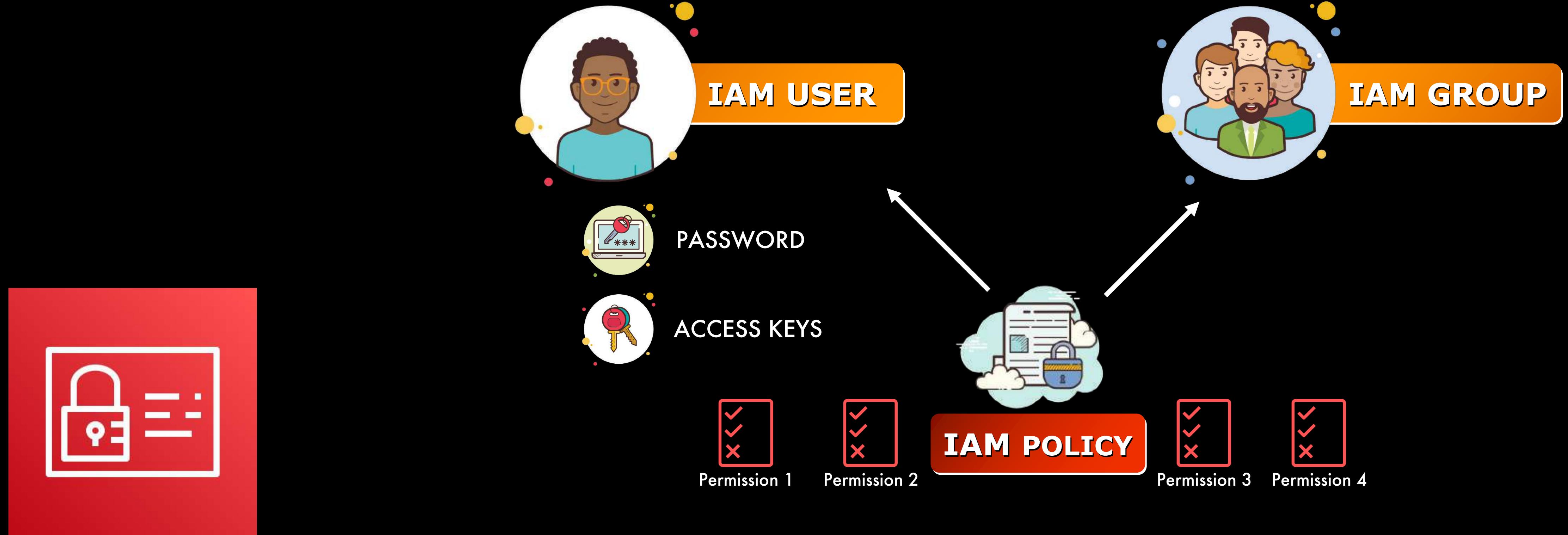


**AWS Directory  
Service**



- The primary **identity service** in AWS
- Allows you to **manage access to various AWS services** and resources

## AWS Identity & Access Management (IAM)



## AWS Identity & Access Management (IAM)





- Let you **add user sign-up, sign-in, and access control** features to your web or mobile apps
- Allows users to log in to your application with their:



**Amazon Cognito**

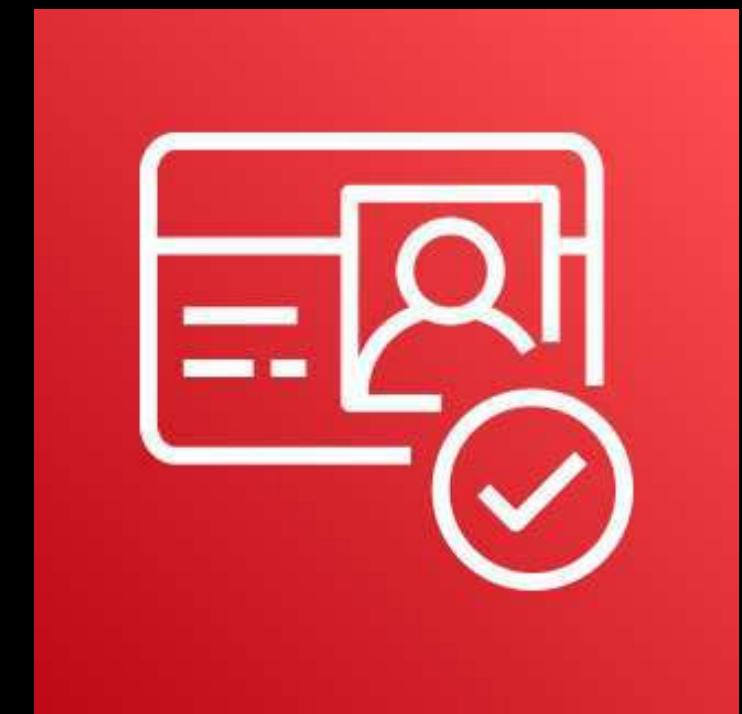


and other  
social media accounts!



**S A M L**

Security Assertion Markup Language



## Amazon Cognito



For **Authentication**

Users can sign in by authenticating through their **social identity providers**



For **Authorization**

Users can obtain **temporary and limited-privilege AWS credentials** that authorize access to other AWS services



## AWS Single Sign-On

- A **single sign-on service** in AWS
- Allows a user to **log in with a single ID and password** to access multiple and independent, software systems
- Provides a **user portal** that allows users to access the roles that they can assume
- Offers pre-configured SAML integrations to many business applications



## AWS Directory Service



- A managed Microsoft Active Directory
- Does not require you to synchronize or replicate data from your existing Active Directory to the cloud
- No need to install and manage an Active Directory domain controller
- Improves security and minimizes administrative overhead
- Allows you to assign IAM roles to your Active Directory users and groups
- Allows you to assign IAM roles to your on-premises Microsoft Active Directory using:



AD Connector

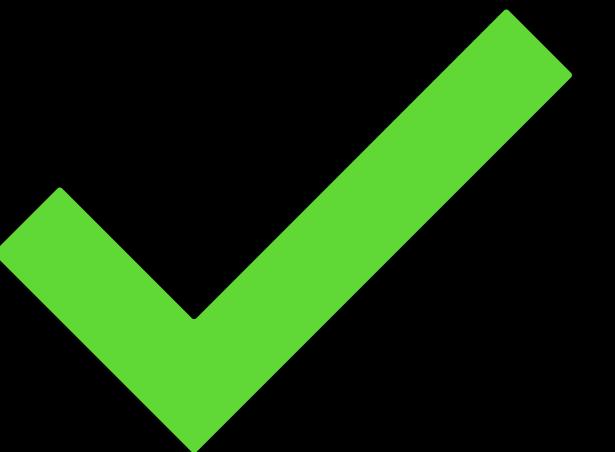
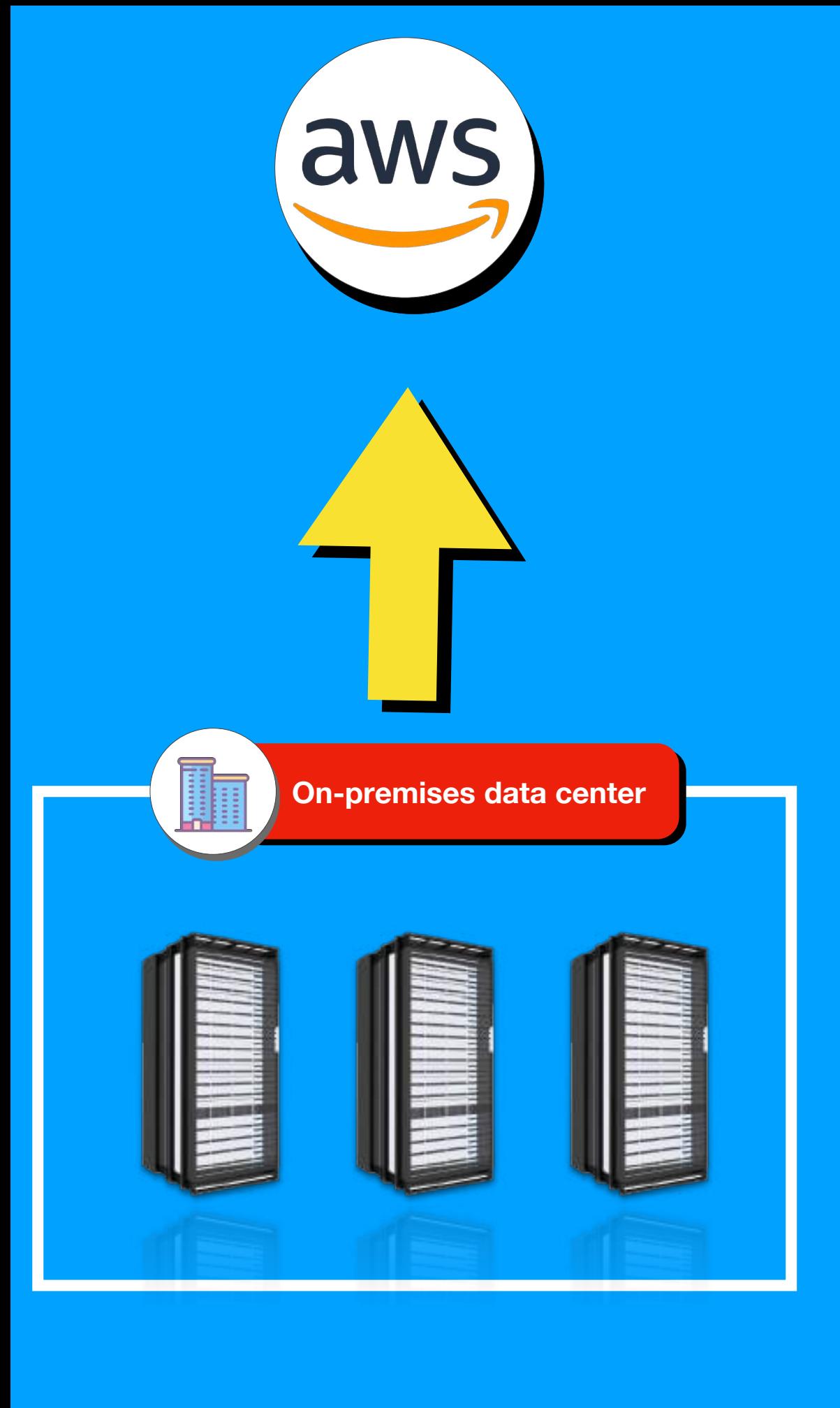


# AWS Transfer & Migration Services Overview

---



# AWS Transfer & Migration Services





# AWS Transfer & Migration Services



**AWS DataSync**



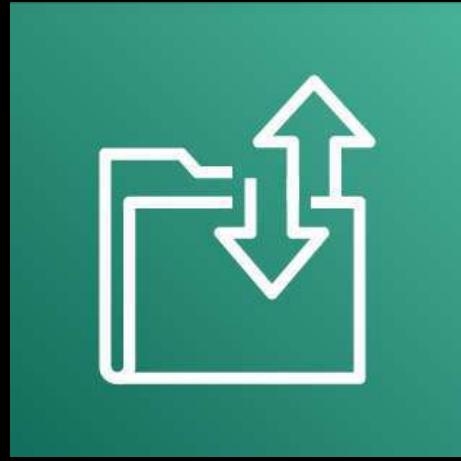
**AWS Application  
Discovery Service**



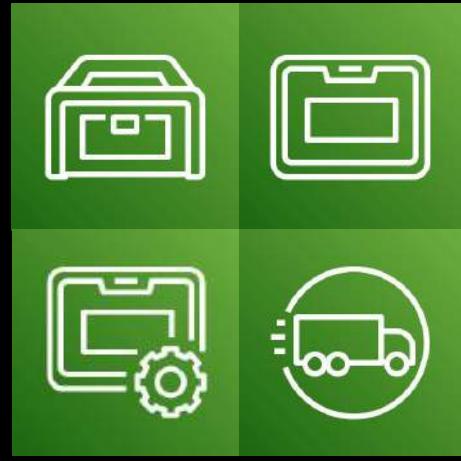
**AWS Database  
Migration Service  
(AWS DMS)**



**AWS Server  
Migration Service  
(AWS SMS)**



**AWS Transfer  
Family**



**AWS Snowball  
Family**



**Migration Hub**



**Migration Evaluator**

- An **online data transfer** service
- Automate and accelerate the replication of data between your **on-premises storage systems and AWS storage services**

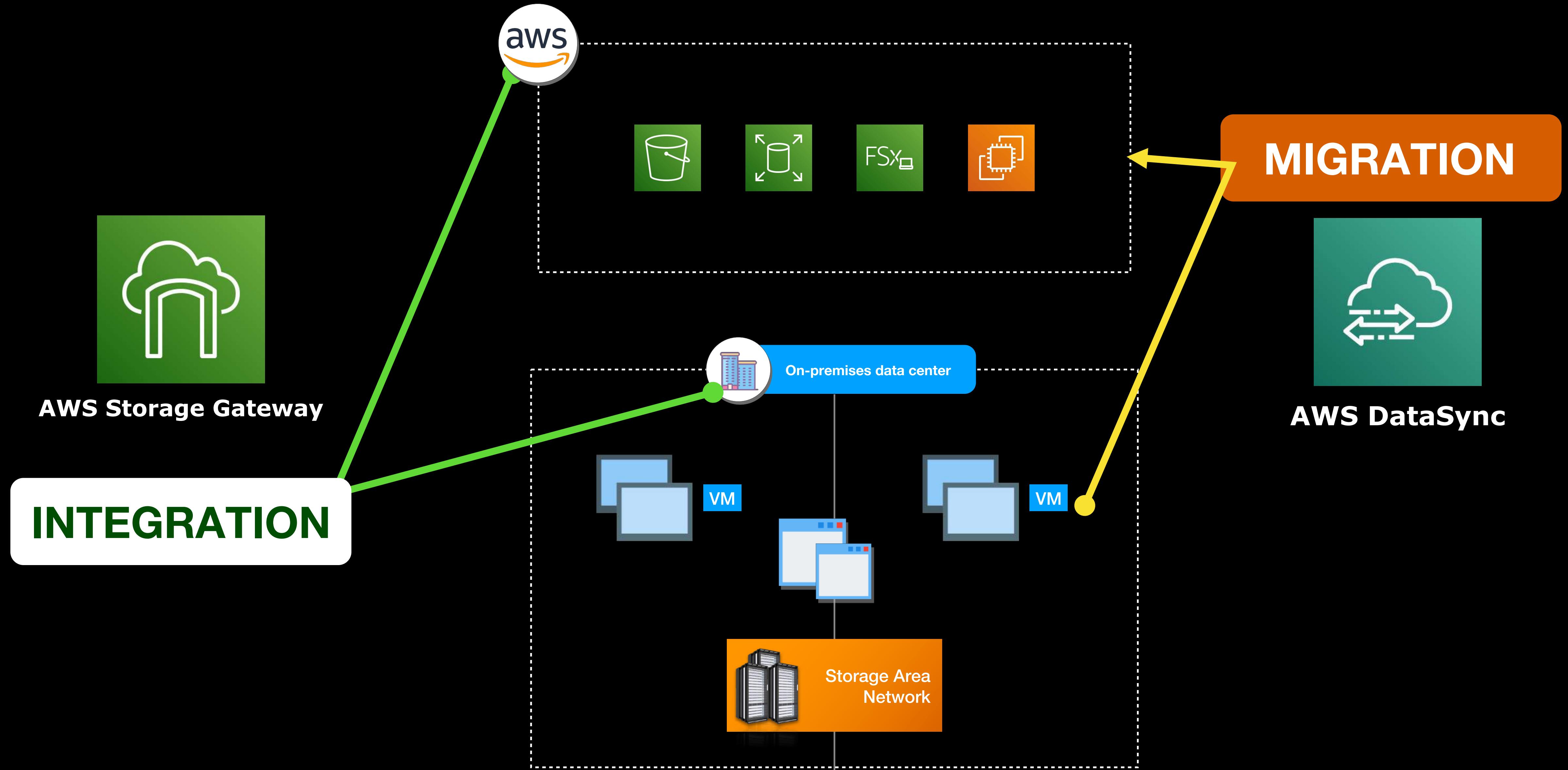


## AWS DataSync

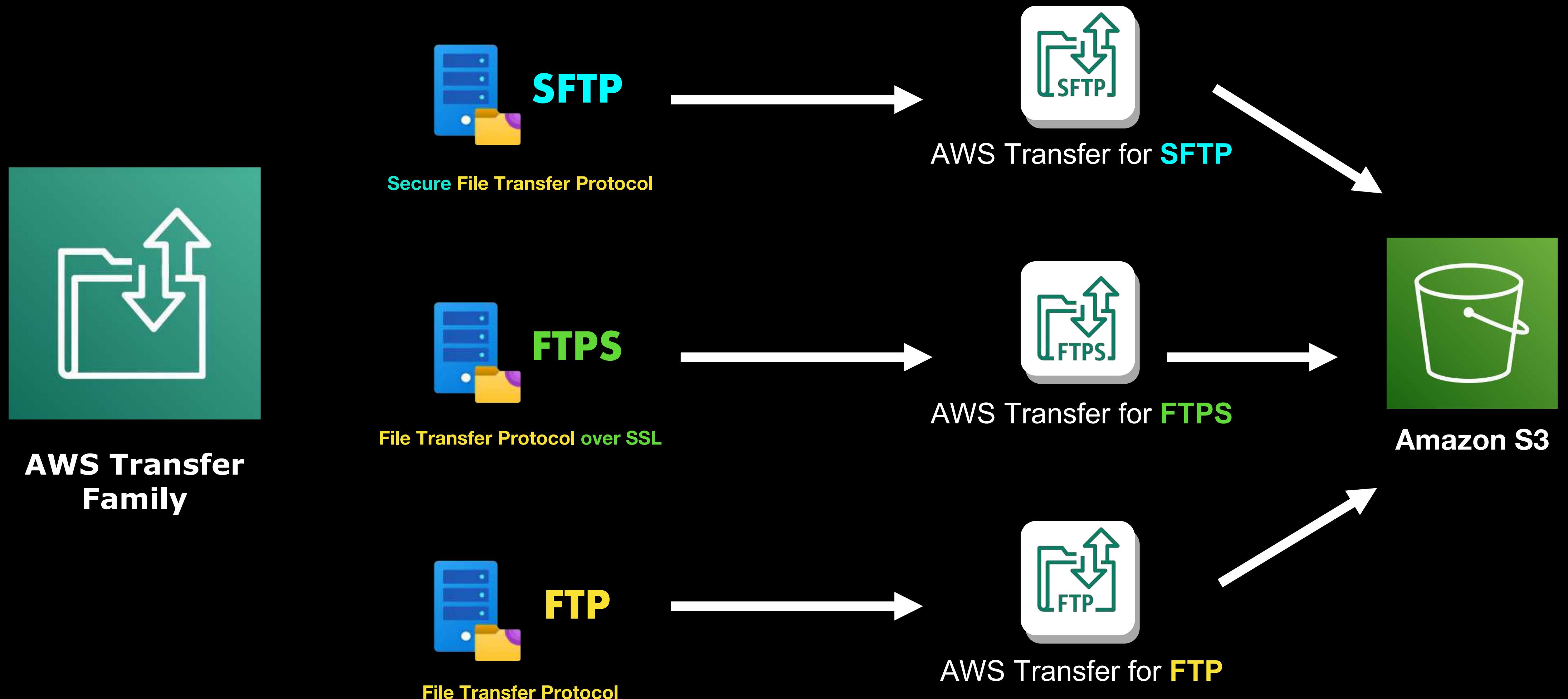
- **Copy large amounts of data** to and from AWS storage services over the Internet or via AWS Direct Connect
- Can copy data between:
  - Shared file servers
  - Self-managed object storage
  - AWS Snowcone
  - Amazon S3 buckets
  - Amazon EFS file systems
  - Amazon FSx for Windows File Server file systems
- Transfers your data from your on-premises data center to AWS through the use of:



**DataSync Agent**



A suite of services that provides a **simple and seamless file transfer**  
to **Amazon S3**



Provides **physical storage devices** and capacity points to help you move your on-premises data to AWS



**AWS Snowball  
Family**



AWS Snowcone



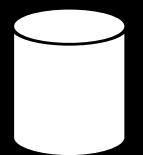
AWS Snowball



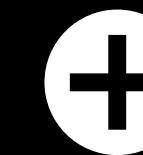
AWS Snowmobile



4.5 lbs / 2.1 kgs



8 TB of Usable Storage



Load data via **NFS mount**



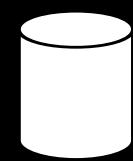
Uploads data to **Amazon S3**

## AWS Snowcone

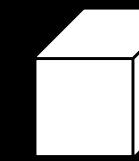




Around 50 lbs / 22.5 kgs



80 TB of Usable Storage



- Over 1 foot in height
- 11 inches wide
- 2.3 inches in length



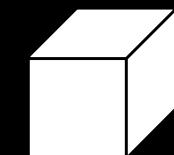
Uploads data to **Amazon S3**

## AWS Snowball

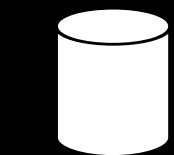




Pulled by  
a semi-trailer truck



45-foot long ruggedized  
shipping container



- Move 100 Petabytes of data
- Exabyte-scale data transfer



Uploads data to Amazon S3

## AWS Snowmobile



- Helps enterprise customers **plan migration projects**
- **Gathers information** about the customer's on-premises resources
- Enable customers to understand the configuration, usage, and behavior of servers in their IT environments
- An **AWS Discovery Agent** is required to be installed to your on-premises servers or virtual machines to capture system configuration, system performance, running processes et cetera
- Helps you **Discover** the technical details of your **Applications** running on your on-premises data center



## AWS Application Discovery Service

- Helps you **migrate your databases to AWS** quickly and securely
- Allows the source database to **remain fully operational** during the migration, which minimize the downtime



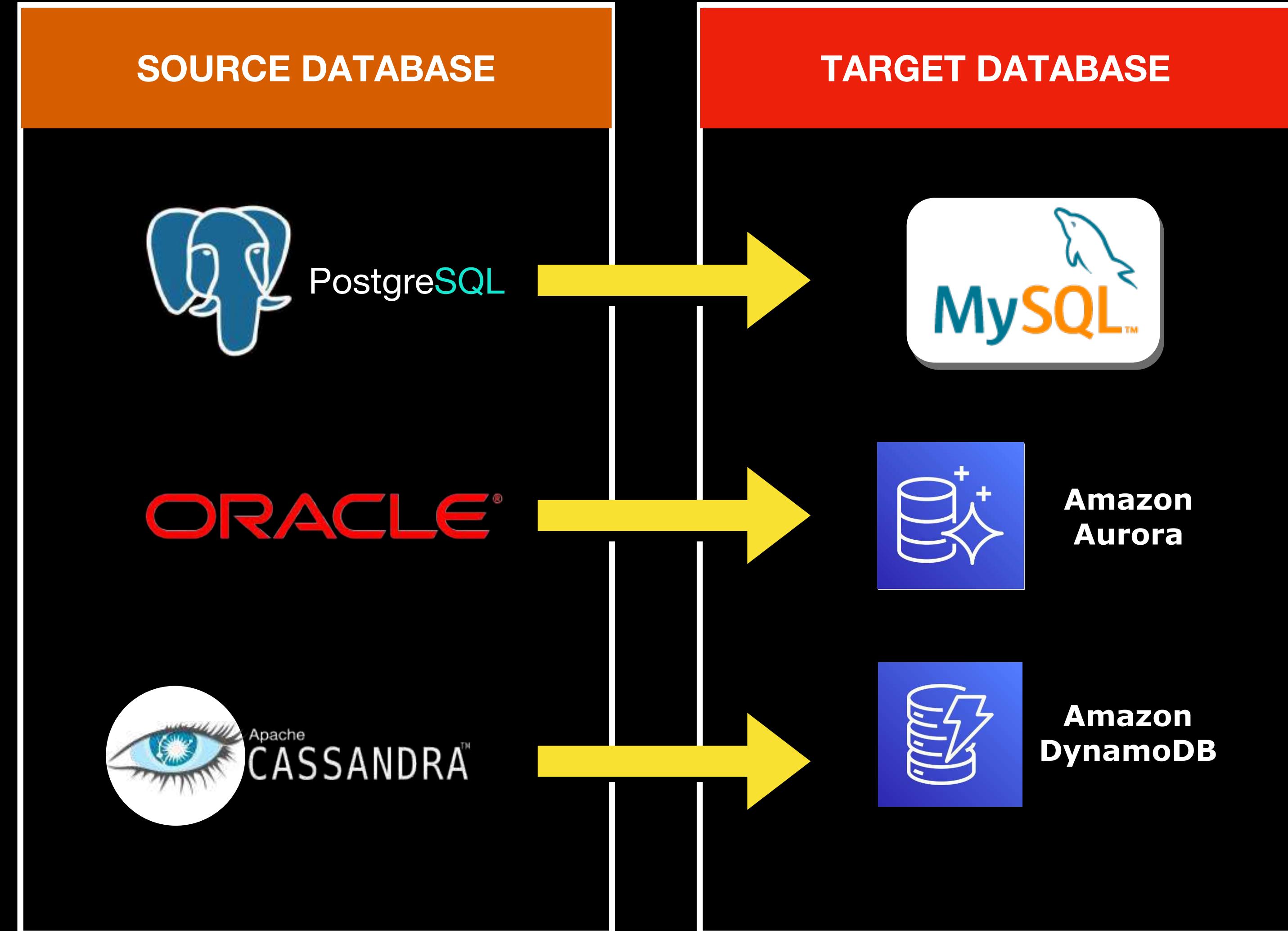
## AWS Database Migration Service (AWS DMS)

- Migrates your data to and from the most widely used commercial and open-source databases
- Allows continuous data replication via **change data capture (CDC)**
- Can be used along with **AWS Schema Conversion Tool (AWS SCT)**
- Supports both **homogeneous** (e.g. Oracle to Oracle, MySQL to MySQL) and **heterogeneous** (e.g. Oracle to MySQL, MS SQL to Amazon Aurora) database migrations

## HETEROGENEOUS DATABASE MIGRATION



**AWS Database Migration  
Service  
(AWS DMS)**





## AWS Server Migration Service (AWS SMS)

- An **agentless service that migrates on-premises workloads** and resources to AWS
- **NO NEED to install and set up an agent** like a System Manager or DataSync agent on-premises
- Uses an **SMS connector**, which can be installed on your VMware vCenter environment, to establish connection to your AWS resources
- **Automate, schedule, and track incremental replications** of your live server volumes



## Migration Hub

- A **single place to discover** your existing servers, **plan migrations**, and **track the status** of each application migration
- **DOES NOT execute actual data migration** — only track its progress
- **Provides visibility** into your application portfolio and streamlines planning and tracking
- **Shows the status of the servers and databases** that you are migrating



## Migration Evaluator

- A **migration assessment** service
- Helps customers to make the best business case for their mission-critical **AWS cloud planning and migration** activities
- Provides a clear **baseline of what workloads you're running** today
- Recommends **future-state configurations**
- Creates a **statistical model** of compute patterns for all your instances, that shows:
  - How much is being spent
  - Which AWS resources are over-provisioned
  - Specific opportunities to realize significant savings



# AWS Machine Learning Services Overview

---



# AWS Machine Learning Services

COMPUTER VISION

CUSTOMER EXPERIENCE IMPROVEMENT

AUTOMATED DATA EXTRACTION & ANALYSIS

BUSINESS METRICS

LANGUAGE AI

DEVOPS & MLOPS

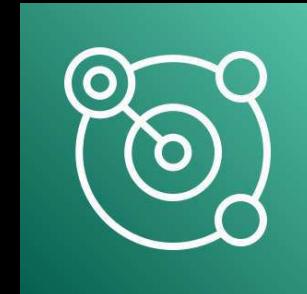


# AWS Machine Learning Services

## COMPUTER VISION



Amazon Rekognition



Amazon Lookout for  
Vision



AWS Panorama

## AWS ML Platform



Amazon SageMaker

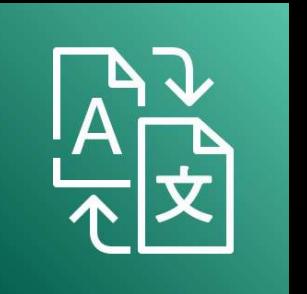
## CUSTOMER EXPERIENCE IMPROVEMENT



Amazon Kendra



Amazon Personalize



Amazon Translate

## AUTOMATED DATA EXTRACTION & ANALYSIS



Amazon Textract



Amazon Augmented  
AI (A2I)



Amazon Comprehend



Amazon Comprehend  
Medical



Amazon Forecast



Amazon Fraud Detector



Amazon Lookout for  
Metrics

## LANGUAGE AI



Amazon Lex



Amazon Transcribe



Amazon Polly

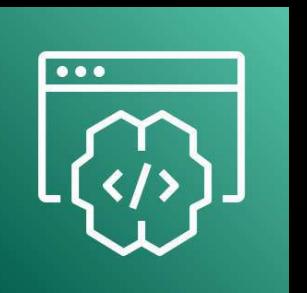
## DEVOPS & MLOPS



Amazon DevOps Guru

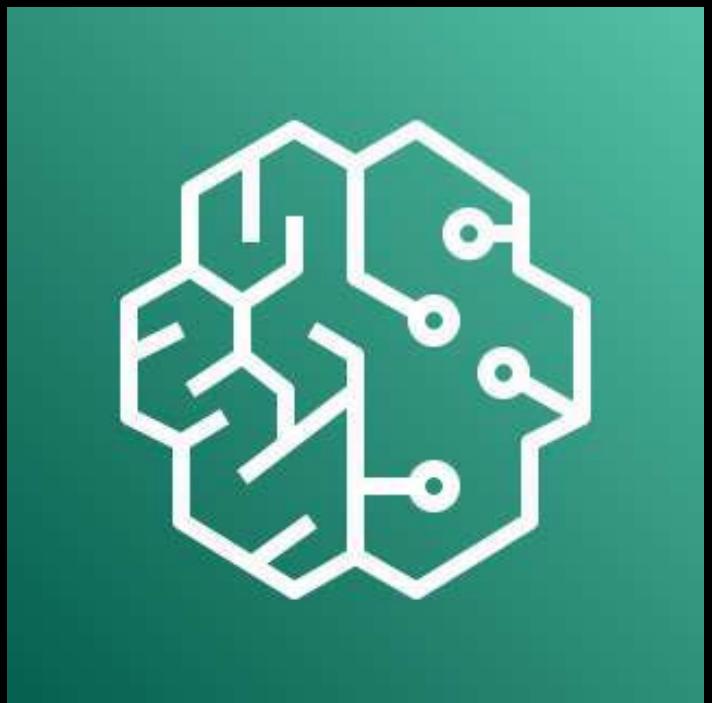


Amazon CodeGuru  
Reviewer & Profiler



Amazon  
CodeWhisperer

## AWS Machine Learning Platform



### Amazon SageMaker

- Full-fledged machine learning platform in AWS
- Allows you to build, train, and deploy machine learning (ML) models for any use case with fully managed infrastructure, tools, and workflows
- Provides a suite of features and modules, such as:



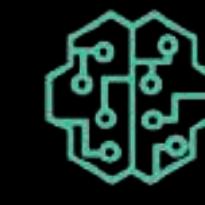
Amazon SageMaker  
Canvas



Amazon SageMaker  
Studio Lab



Amazon SageMaker  
Ground Truth



Amazon SageMaker  
Built-In Models



Amazon SageMaker  
Notebook

and many more!

## COMPUTER VISION



### Amazon Rekognition

- Extract information and insights from your images and videos using computer vision
- It can recognize:
  - Objects, texts, scenes, labels, and other attributes
  - Face of a person or a popular celebrity
  - Personal Protective Equipment (e.g. mask, helmet)
- Has a feature called Amazon Rekognition Custom Labels that allows you to classify custom components or products from your dataset

## COMPUTER VISION



### Amazon Lookout for Vision

- One of the services in the Amazon Lookout Family
- Detects **defects** on industrial products
- Used in factories and manufacturing lines to identify defects
- Actual **images of defect-free products are used as a dataset**. These images can be stored in Amazon S3 and used as **baseline images** to build a custom ML model for you
- Can automatically detect anomalies in your product like dents, cracks, scratches et cetera

## AUTOMATED DATA EXTRACTION & ANALYSIS



### Amazon Textract

- Its name is a portmanteau of the words “text” and “extract”
- Extract texts from scanned documents, PDFs, Word documents, hand-written notes, receipts, passports, IDs, and many others
- Can generate the results into a table form or a CSV file
- Has a query feature that extracts a particular field using natural language questions
- Can batch upload your documents to Amazon S3 and automate the text analysis process

## AUTOMATED DATA EXTRACTION & ANALYSIS



## Amazon Augmented AI (A2I)

- Provides human review workflows for common machine learning use cases
- The review is done by actual people and not by a computer
- Ensures the accuracy of prediction results and helps provide continuous improvements to your machine learning model
- Can be directly integrated to Amazon Rekognition, Amazon Textract and other services
- Useful for image moderation such as explicit adult or violent content
- Allows you to run a human review with a custom machine learning workflow of your choice

## AUTOMATED DATA EXTRACTION & ANALYSIS



- A natural language processing service
- Finds insights and relationships from text documents
- Can extract key phrases, sentiment, language, syntax, topics, and even Personally Identifiable Information (PII) from unstructured data
- Can implement patient data privacy solutions and identify protected health information (PHI) using:



**Amazon  
Comprehend  
Medical**

## Amazon Comprehend

- Can comprehend or understand the information written in your text documents
- Raw text data must be supplied first in order to use the Amazon Comprehend service

## LANGUAGE AI



### Amazon Lex

- Enables you to **develop conversational chatbots**
- Allows you to build Voice-based or Text-based chatbots
- Useful for developing a self-service bot or a virtual agent for your conversational Interactive Voice Response (IVR) system, corporate website, or others
- Reduces costs in maintaining a contact center

## LANGUAGE AI



### Amazon Transcribe

- A speech-to-text transcription service
- Transcribes, or makes a written record of, a speech, a phone call, or any spoken language
- Can generate call transcripts and provide conversation insights to improve customer experience and agent productivity
- Offers real-time transcription

## LANGUAGE AI



### Amazon Polly

- Converts text into speech
- Generates a lifelike speech in different voices based on a raw text file you uploaded
- If you typed: *Beautiful Philippine Islands*, the Amazon Polly service will generate an audio file saying that phrase in a male voice, a female voice, a kid's voice, or in any voice that you want your text to be spoken
- Allows you to upload custom lexicon files which can help you to customize the pronunciation of specific words and phrases

## CUSTOMER EXPERIENCE IMPROVEMENT



### Amazon Kendra

- An intelligent search service in AWS
- Can search items from multiple data sources containing both structured and unstructured data
- Supports natural language processing:
  - *"Who is the founder of the EdTech startup: Tutorials Dojo?"*
  - *"Where is the JP Rizal Hospital located?"*
  - *"How much did Mr. Jon Bonso earn a year ago?"*
- Searches all of the documents in your S3 bucket, FSx file systems, RDS databases, Github repository, Jira, Slack, Sharepoint and other data sources
- Uses machine learning to provide context to your search results for a better customer experience

## CUSTOMER EXPERIENCE IMPROVEMENT



### Amazon Personalize

- Provides **personalized recommendations** to your customers based on their past activity and behavior
- Similar to the recommendation feature in Amazon Prime, Netflix and other online streaming platforms
- Gives **recommendations based on the customer's profile, viewing history and past activities**
- Improves customer experience and sales since you can offer products that your customers wanted

## CUSTOMER EXPERIENCE IMPROVEMENT



### Amazon Translate

- A real-time language translation service
- Works like Google Translate
- Enables you to create custom terminologies based on a company-specific and domain-specific vocabulary
- For example:
  - Set the acronym "TD" as "Tutorials Dojo"
  - Enter the Tagalog phrase: "*Magandang umaga, TD*"
  - It will return: "*Good morning, Tutorials Dojo*" as an output
- Has a Formality option that controls whether the translation output uses a formal tone
- Can mask profane words or phrases

## BUSINESS METRICS



### Amazon Forecast

- Helps you **forecast** a future outcome based on your historical records and other relevant data
- You can either import or stream your time-series data to the Amazon Forecast service
- Can **provide intelligent predictions** to your sales, web traffic, inventory, revenue, cloud resource capacity, weather, future AWS bill et cetera
- Has a range of **built-in datasets** such as Weather Index, national holidays for various countries and many more
- Uses a **Predictor** machine learning model that consumes all the time-series data that you provide to make a prediction

## BUSINESS METRICS



### Amazon Fraud Detector

- Automates the fraud detection process in your applications
- Identifies potential fraudulent activity, fake reviews and spam account creation in near-real-time
- Use cases:
  - Detecting the IP addresses with a history of spamming, hacking attempts, and DDoS attacks
  - Blocking users with exactly the same IP address are posting spam and fraudulent review on your website
  - Preventing a malicious user who uses an offending IP address, an email domain, or a key attribute

## BUSINESS METRICS



### Amazon Lookout for Metrics

- One of the services of the Amazon Lookout family
- Detects anomalies in your business metrics, such as:
  - A sudden nosedive in your sales revenue
  - Unexpected drop in your customer acquisition rates
  - Causal relationships
- Identifies unusual variances in your business metrics
- Can be integrated with Amazon SNS to send alerts whenever an anomaly is detected

## DEVOPS & MLOPS



### Amazon DevOps Guru

- A machine learning service that **detects abnormal behavior in your application or AWS resources**
- Prevents unexpected downtimes or operational issues in the near future
- Monitors applications and AWS resources within your own account or on all accounts across your AWS Organization
- **Identifies operational defects** such as:
  - An unusually high DB load that is more than three times or 5 times its normal value
  - Extremely high number of invocations in your Lambda function beyond the provisioned concurrency
  - Overprovisioned write capacity on your DynamoDB tables

- A suite of development services in AWS with different tools and features such as:

### Amazon CodeGuru Reviewer

DEVOPS & MLOPS



### Amazon CodeGuru Profiler

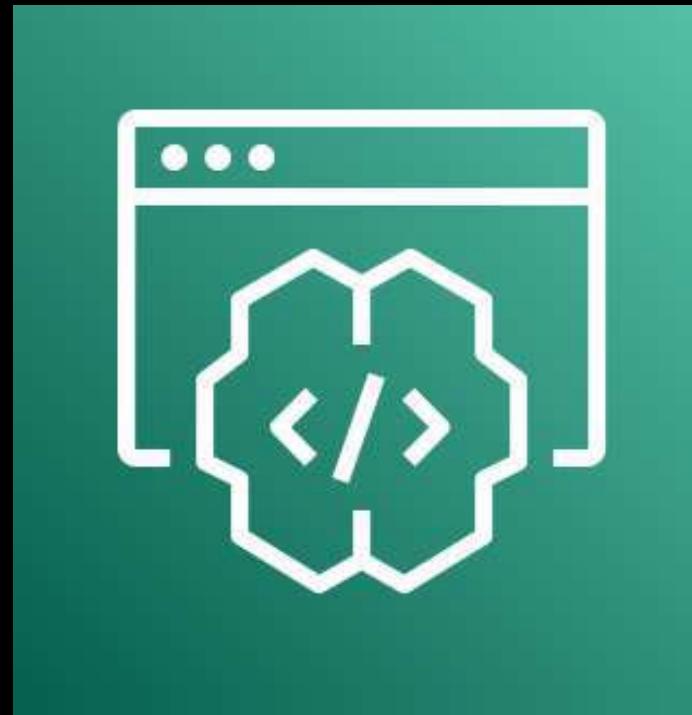
Provides recommendations for improving your efficiency, and code quality

- Scans your code and detect a range of code defects like bad exception handling, insecure CORS policy, path traversal, hardcoded credentials et cetera
- Can be integrated with your CI/CD workflow to automate the code review process

## Amazon CodeGuru

- A component that collects CPU data and analyzes the runtime performance data from your live applications
- Identifies expensive lines of codes that inefficiently use the CPU, which causes CPU bottlenecks.

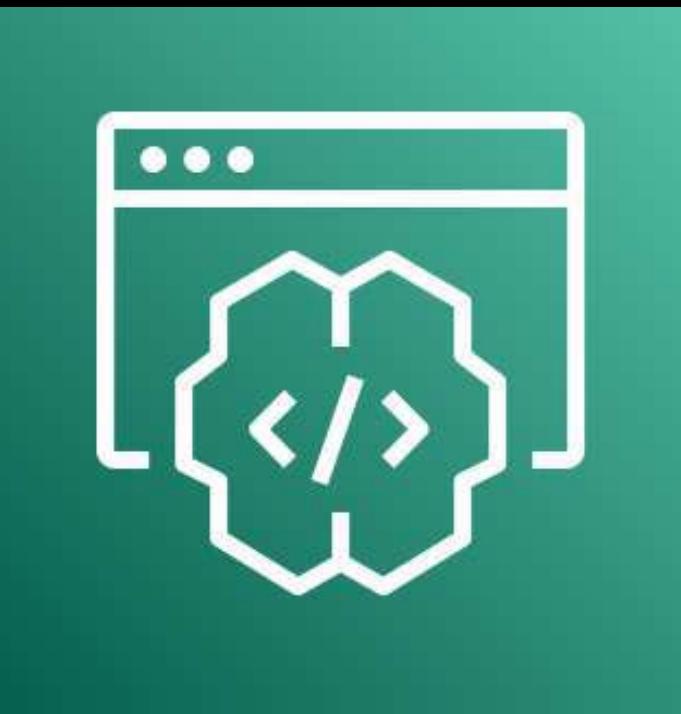
## DEVOPS & MLOPS



- Automatically generates code and functions in real-time
- Similar to Github CoPilot
- Installed in your Visual Studio IDE
- The lines of codes are generated right from your IDE editor based on the comments that you write

## Amazon CodeWhisperer

## DEVOPS & MLOPS



# Amazon CodeWhisperer

Python

```
import boto3
from botocore.exceptions import ClientError
```

```
# Function to upload a file to an S3 bucket
def upload_file(file_name, bucket, object_name=None):
```

```
    """Upload a file to an S3 bucket
```

```
:param file_name: File to upload
```

```
:param bucket: Bucket to upload to
```

```
:param object_name: S3 object name. If not specified then file_name is used
```

```
:return: True if file was uploaded, else False
```

```
"""
```

COMMENTS

```
# If S3 object_name was not specified, use file_name
```

```
if object_name is None:
```

```
    object_name = file_name
```

GENERATED LINES  
OF CODES

```
# Upload the file
```

```
s3_client = boto3.client('s3')
```

```
try:
```

```
    response = s3_client.upload_file(file_name, bucket, object_name)
```

```
except ClientError as e:
```

```
    logging.error(e)
```

```
    return False
```

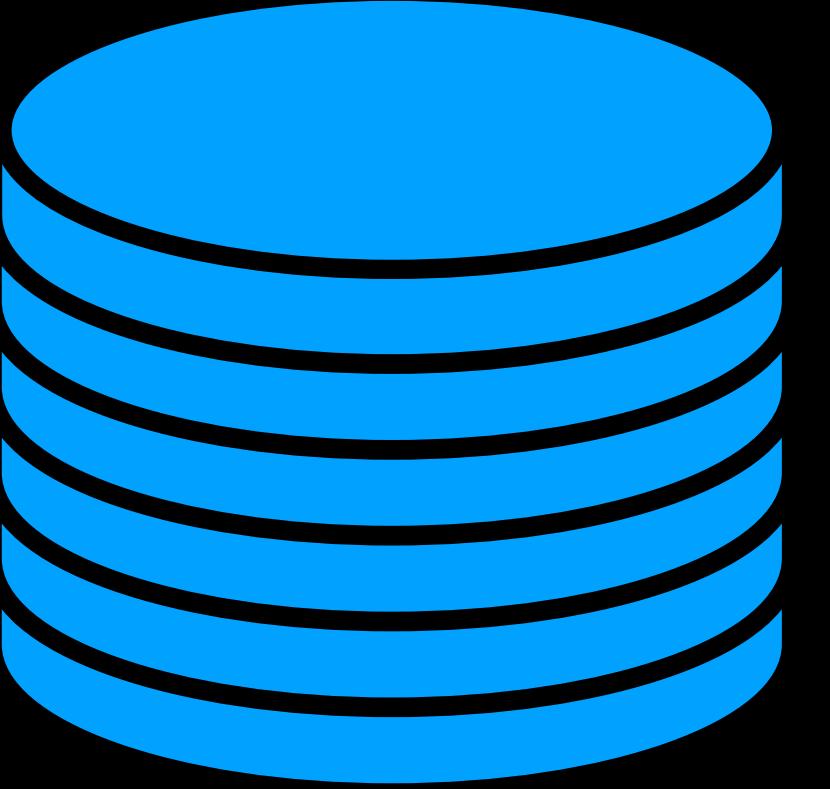
```
return True
```

Amazon CodeWhisperer



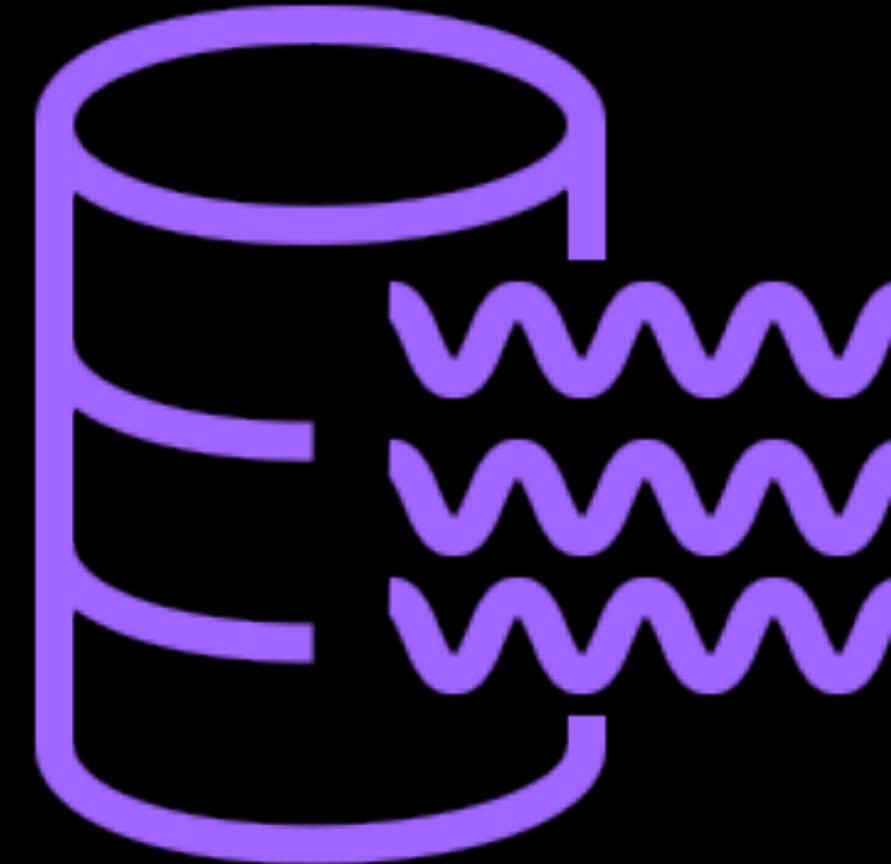
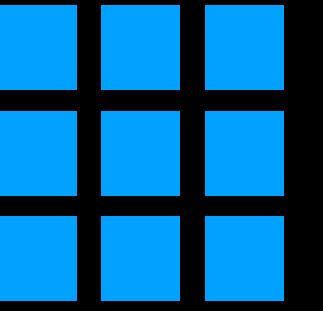
# AWS Analytics Services Overview

---



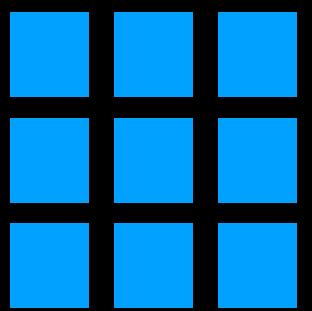
# Data Warehouse

STRUCTURED DATA

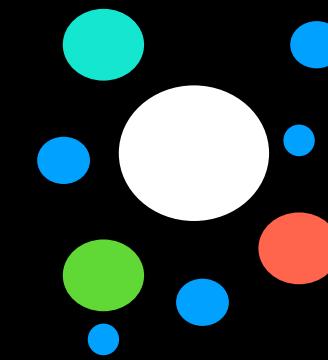


# Data Lake

STRUCTURED DATA



UNSTRUCTURED DATA





# Open Source Technologies used by AWS Analytics Services



...and many other open-source projects!

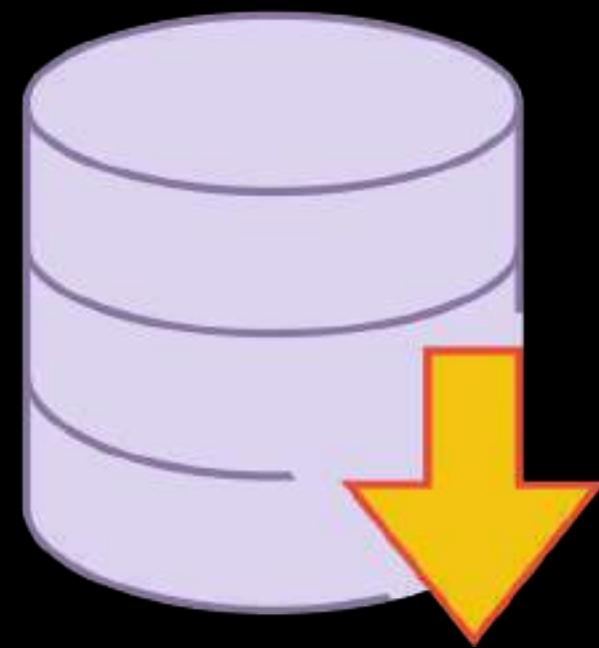


## 3rd Party Technologies used by AWS Analytics Services

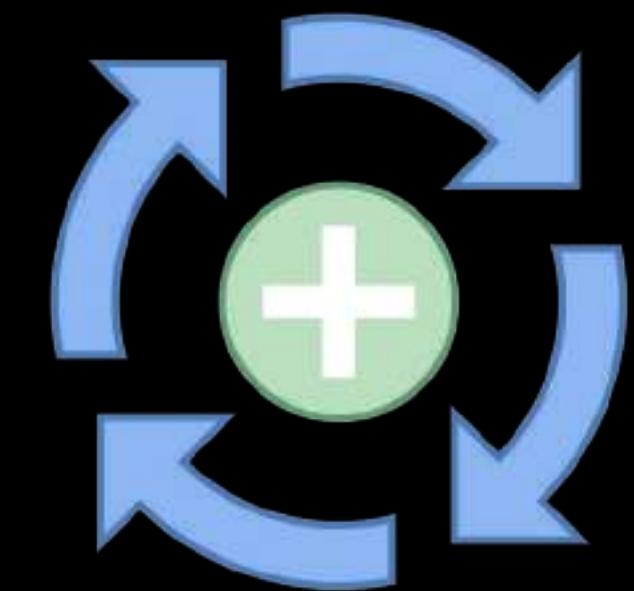


...and many more!

**Extract**

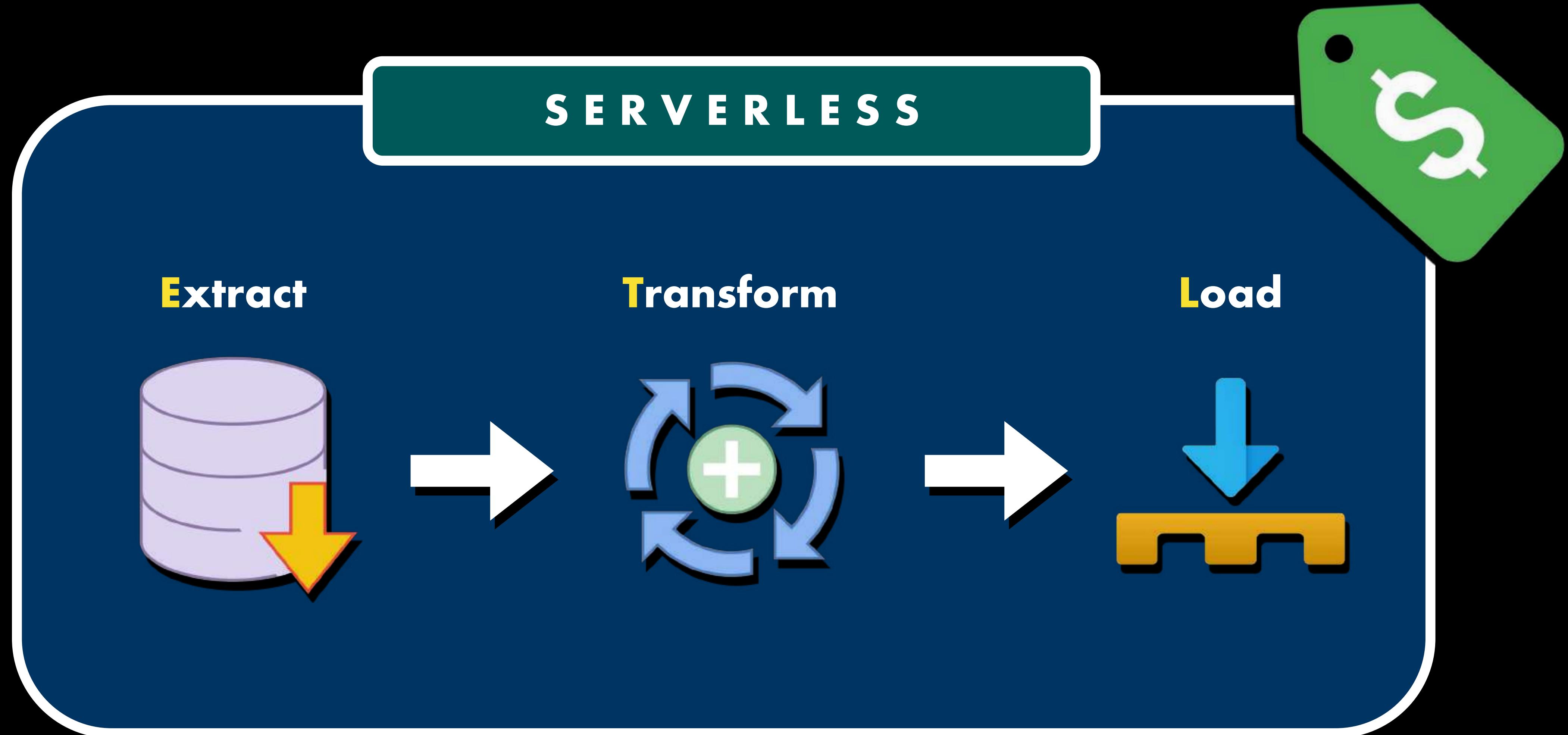


**Transform**



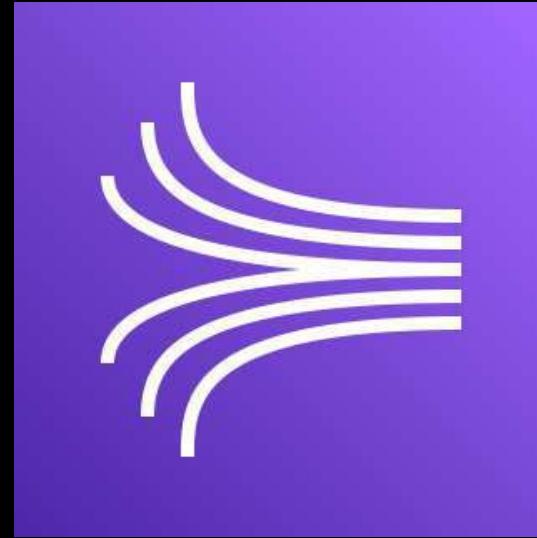
**Load**







# AWS Analytics Services



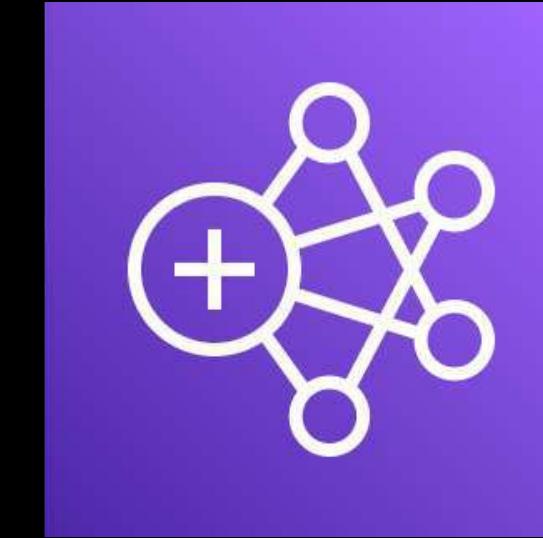
Amazon Kinesis



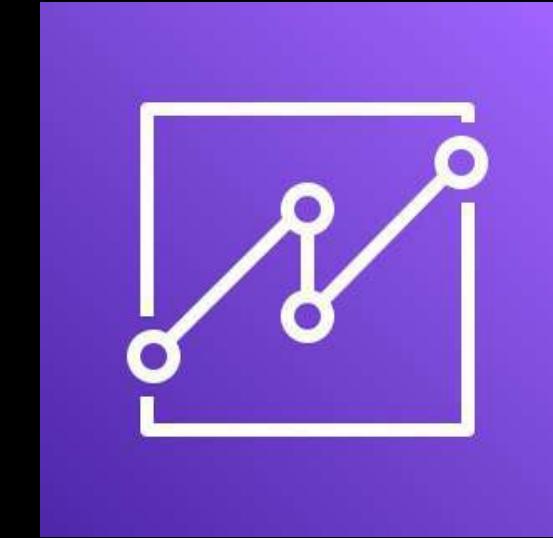
Amazon Athena



Amazon Elasticsearch  
(Amazon ES)



Amazon Elastic MapReduce  
(Amazon EMR)



Amazon QuickSight



Amazon CloudSearch



Amazon Redshift



AWS Data Pipeline



AWS Glue



Amazon Managed  
Streaming for Apache Kafka

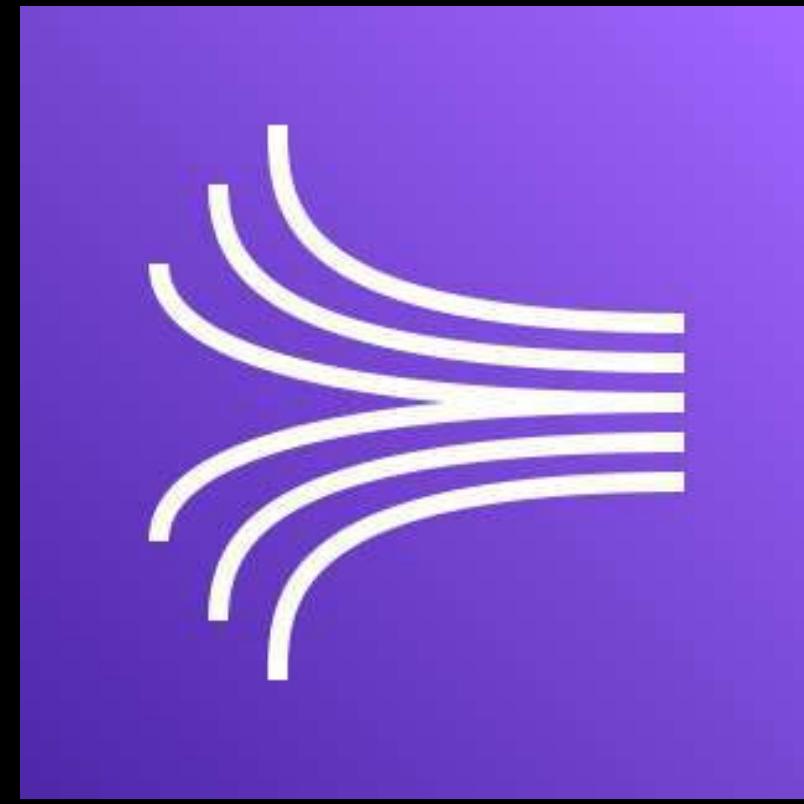


AWS Lake Formation



## Amazon Kinesis

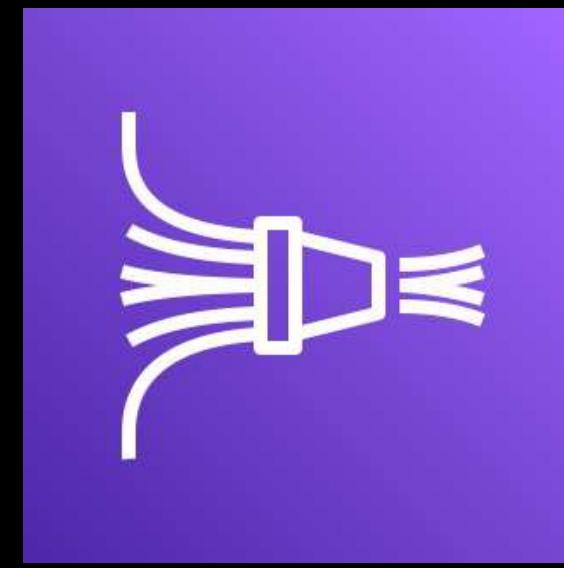
- **A suite of services** for processing your data streams
- Analyzes your **data streams** in **real-time**
- Allows you to **collect, transform, process, load, and analyze the streaming data** in real-time to help you acquire the data insights and respond to data changes



## Amazon Kinesis



Amazon Kinesis  
**Data Streams**



Amazon Kinesis  
**Data Firehose**



Amazon Kinesis  
**Data Analytics**



Amazon Kinesis  
**Video Streams**



## Amazon Kinesis Data Streams

- A massively scalable, durable, secure and low-cost **real-time data streaming service**
- Can **continuously capture gigabytes of data** per second from thousands of different sources
- Collects and sends data to your data analytics applications and consumers in real-time



## Amazon Kinesis **Data Streams**

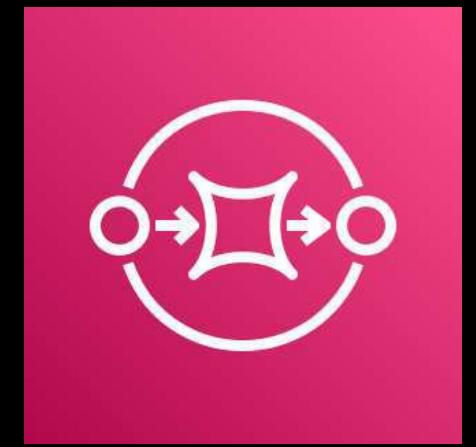
- Can be used in:
  - ▶ Real-time Applications
  - ▶ Website Clickstreams
  - ▶ Database Event Streams
  - ▶ IoT Telemetry
  - ▶ Location-tracking Events
  - ▶ Predictive Maintenance
  - ▶ Mobile Game Data Streams
  - ▶ Online Marketplaces
  - ▶ Real-time Recommendations Systems
  - ▶ ...and many more!
- Provides **ordering of records**
- Can **read & replay records** in the same order
- Suitable if you have a requirement where:
  - ▶ The data events must be received in an **ordered manner**
  - ▶ There's a need to process the data stream of your web applications, or mobile game updates, in **order of receipt**



## Amazon Kinesis Data Streams

- Can be used to decouple your cloud architecture like Amazon SQS by accepting data from your data sources and forward it to different compute resources

- Similar to **Amazon SQS** with notable differences:
  - SQS **can't process data in real-time**
  - SQS Standard queue **doesn't maintain the order of data** records by default
  - SQS FIFO queue maintains the order of data records but is **significantly slower** than SQS Standard and doesn't perform in real-time



## Amazon SQS

## USE CASES



### Amazon Kinesis Data Streams

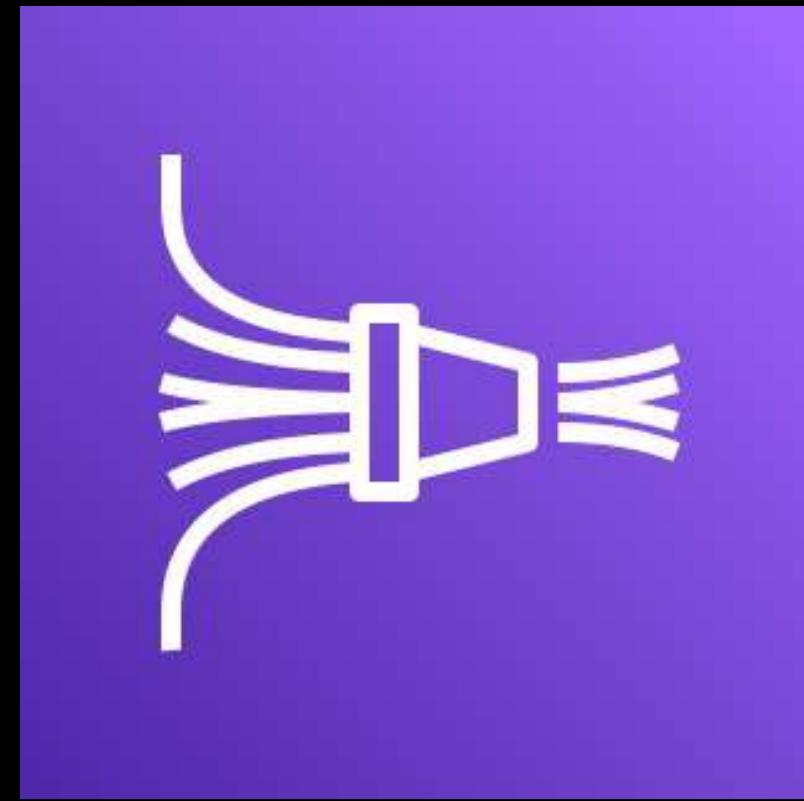
- If you need a solution that **captures the clickstream data from multiple websites in real-time** and analyzes it using batch processing
- For setting up and building a scalable, **near-real-time recommendations** for your users
- For **mobile games that stream score updates** to a backend system and post the results on a leaderboard
- For collecting the mobile game scores in **order of receipt** which can then be processed by an AWS Lambda function and stored in DynamoDB

## USE CASES



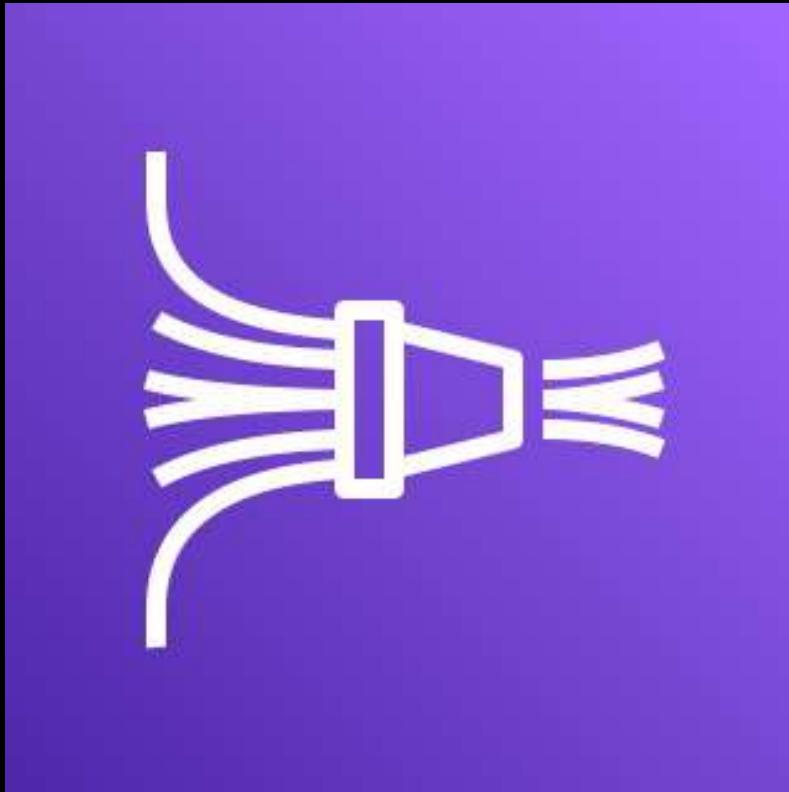
### Amazon Kinesis **Data Streams**

- For implementing **predictive maintenance** on different types of machinery equipment using IoT sensors
- For sending data to AWS in real-time wherein the data stream will **receive events in an ordered manner** for each connected device, data producer or machinery asset
- For implementing a scalable, near-real-time solution in processing millions of financial transactions
- For launching a data stream that can be consumed by Amazon Kinesis Data Analytics which can be queried using SQL queries



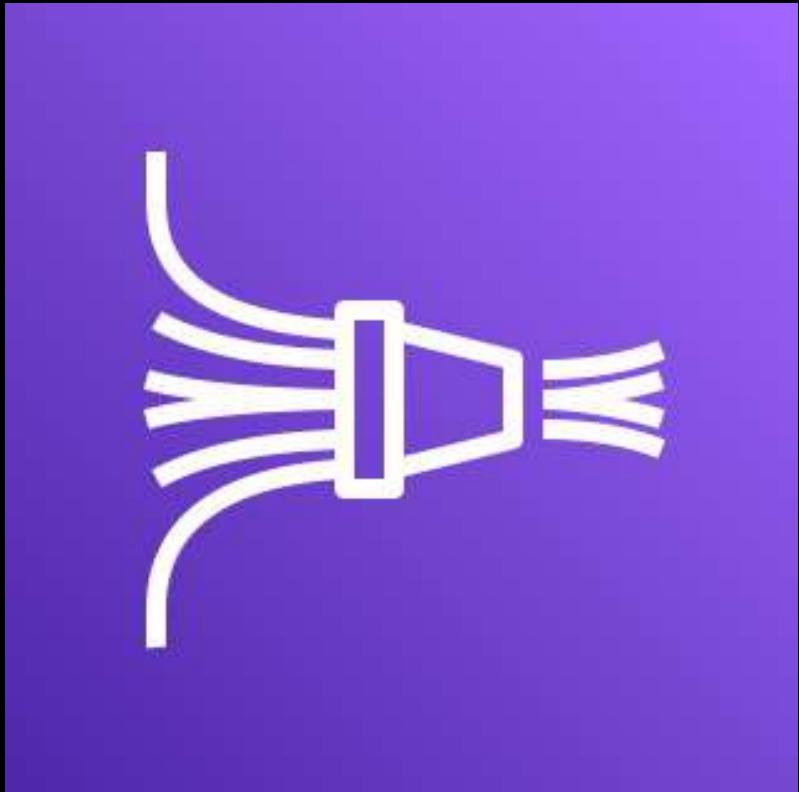
## Amazon Kinesis **Firehose**

- A fully managed service that reliably **transforms and loads your streaming data** into data stores and analytics tools
- Directly delivers data to Amazon S3, Amazon Redshift, Amazon Elasticsearch Service, and any HTTP endpoint
- Can be integrated with your third-party service providers
- Enables your data producers to directly send data to a specific destination or data store that **without any custom applications or consumers**
- Can **transform your data before sending it** to a specified destination to remove sensitive data or for data pre-processing procedures



## Amazon Kinesis **Firehose**

- Similar to Amazon Kinesis Data Stream but with certain differences:
  - ▶ Both service can accept streaming data in real-time
  - ▶ However, Kinesis Data Stream requires an external consumer to store the records while Kinesis Data Firehose does not
- Acts like a “firehose” to **immediately send the streams of data to your data store**
- Delivers your data stream directly to your Amazon S3 buckets, Redshift databases, Amazon ES clusters, and others without the need for a consumer



## Amazon Kinesis **Firehose**

- Can transform the data before it is sent to its destination
- Internally **invokes an AWS Lambda function** to transform the incoming source data and deliver the processed data to its destination
- Recommended if you need to parse the data stream to **remove any sensitive data such as personal data or protected health information (PHI)**



## Amazon Kinesis Video Streams

- A service that **securely streams video** from connected devices or sources to AWS
- Commonly used for data analytics, machine learning, video playback, and other types of media processing
- Automatically provisions and scales all the required infrastructure to ingest streaming video data from millions of devices
- **Stores, encrypts, and indexes video data** in your streams to improve performance
- Provides access to your video data through a collection of easy-to-use APIs



## Amazon Kinesis Data Analytics

- A serverless service that enables you to **analyze your streaming data**, acquire actionable insights, and respond to events in real-time
- Reduces the complexity of building, managing, and integrating streaming applications with your custom applications and other AWS services
- **Serverless**
- **Uses Apache Flink** to process and analyze streaming data
- Eliminates the manual tasks of setting up and maintaining Apache Flink



## Amazon Kinesis Data Analytics

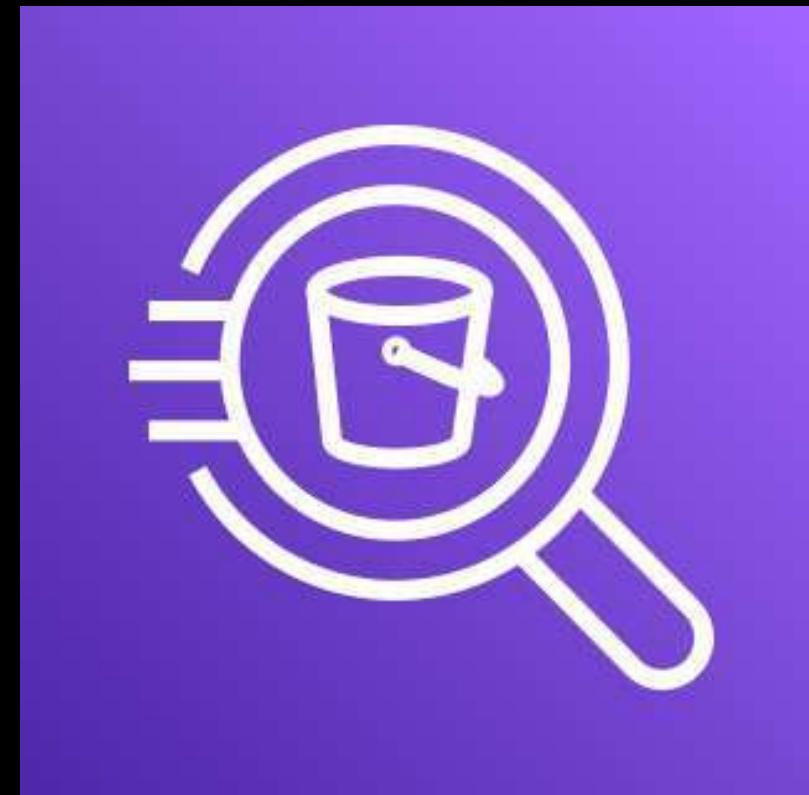
- Enables you to **author and run code** against streaming sources
- The data **can be analyzed using SQL queries** and the results can be delivered to Amazon S3, Amazon Redshift, and other data stores using Kinesis Data Firehose
- Java or Scala can be used to process and analyze your streaming data

## USE CASES



### Amazon Kinesis Data Analytics

- In **near-real-time data processing** and data querying for acquiring timely insights of your application
- For processing your streaming data **with minimal effort and operational overhead**
- For providing scalable and near-real-time data querying with minimal data loss
- For **analyzing the location data points of your GPS application** that tracks the movement of people, bikes, automobiles, or any other moving object
- You can expose a REST API using API Gateway that can be used as an **Amazon Kinesis proxy**



## Amazon Athena

- An **interactive query service** for your data that is stored in Amazon S3
- Simplifies data analysis in Amazon S3 using standard SQL queries
- Unlike S3 Select, you can **query the entire data in your Amazon S3 bucket** with Amazon Athena and not just its subset
- **Serverless**



## Amazon Athena

- Sample use case:
  - ▶ A global eCommerce website stores 250 gigabytes of transactional data each month in Amazon S3
  - ▶ You need to identify the number of items sold in each particular region for the previous month in the most cost-effective way
- Athena **costs less** than Amazon Redshift, Amazon EMR, or Amazon ES since it's serverless
- **Can use an AWS Glue Data Catalog** to store and retrieve table metadata for your Amazon S3 data and provide data visualization using Amazon QuickSight



# elasticsearch



## Amazon Elasticsearch Service (Amazon ES)

- A fully managed **Elasticsearch** service
- Elasticsearch is a distributed, multitenant-capable full-text search engine based on the Apache Lucene library
- Provides an HTTP web interface that can store data as a **schemaless JSON document**
- Provisions the necessary infrastructure and automatically manages the resources needed to run the Amazon ES cluster



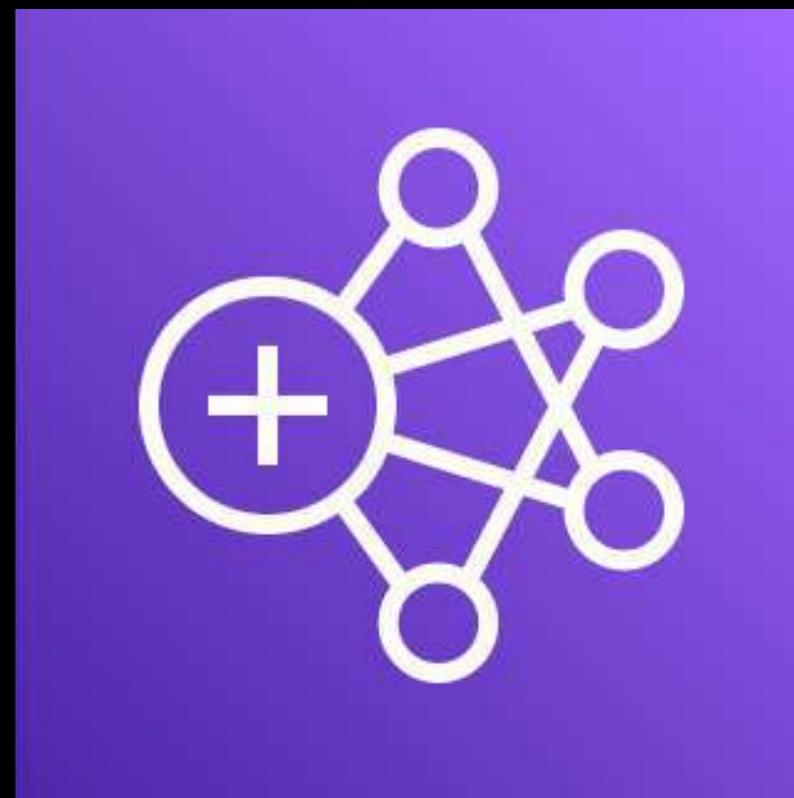
# elasticsearch



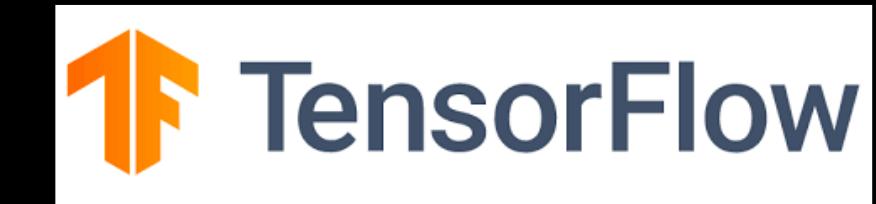
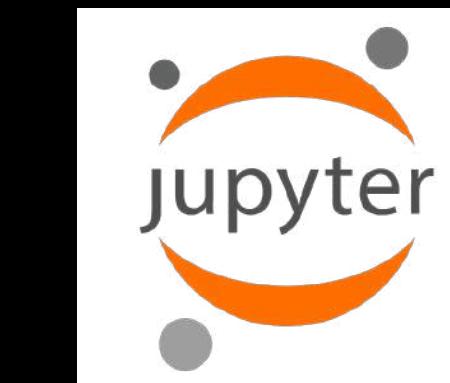
## Amazon Elasticsearch Service (Amazon ES)

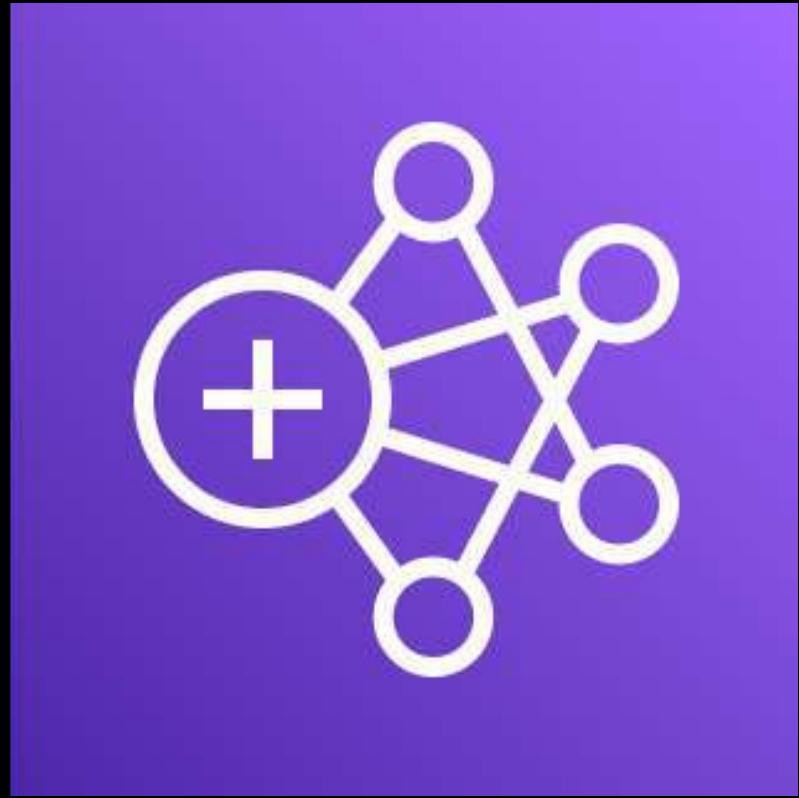
- Also allows you to launch an ELK (Elasticsearch, Logstash, and Kibana) stack in AWS
- ELK Stack:
  - **Elasticsearch** - full-text search engine
  - **Logstash** - server-side data processing pipeline
  - **Kibana** - user interface to visualize Elasticsearch data
- Provides support for open-source Elasticsearch APIs, managed Kibana, integration with Logstash and other AWS services
- Lets you pay only for what you use (*no upfront costs or usage requirements*)

- Allows you to **run different types of big data frameworks in AWS**
- A managed big data platform for processing vast amounts of data using open source tools such as:



**Amazon  
Elastic MapReduce**  
(Amazon EMR)

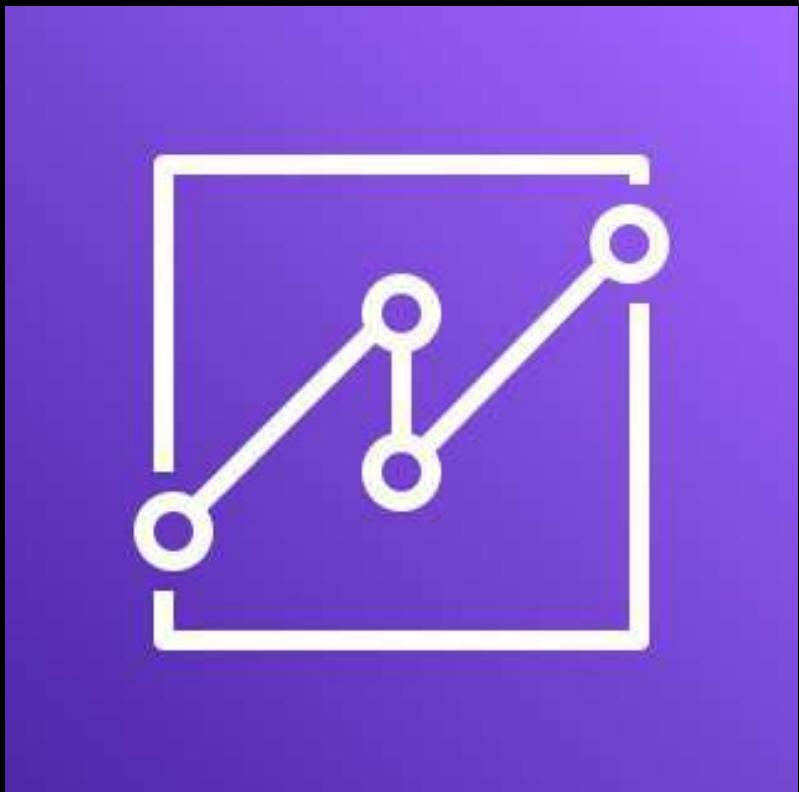




## Amazon Elastic MapReduce (Amazon EMR)

- Runs your big data framework on Amazon EC2 instances, Amazon Elastic Kubernetes Service clusters, or in your on-premises EMR cluster via AWS Outposts
- The compute resources launched by Amazon EMR are deployed in your VPC and then grouped as an Amazon EMR cluster
- You can **directly access and control the underlying EC2 instances** of your EMR cluster
- **NOT serverless**
- Automates the server provisioning and management process for you and allows your data to interact with other AWS data stores such as Amazon S3 and Amazon DynamoDB

- A scalable, serverless, embeddable, machine learning-powered **business intelligence service**



- Allows you to **create and publish interactive dashboards** that can be accessed from different browsers or mobile devices
- Allows you to **embed dashboards** into your applications

## Amazon QuickSight

- Highly scalable and can easily scale up to thousands of users globally
- **Serverless**



## Amazon CloudSearch

- A **managed search service** in AWS
- Can be used to add a search feature in your application or websites
- You can use this to:
  - ▶ Retrieve contents of selected fields
  - ▶ Provide facet information to categorize results
  - ▶ Provide statistics for numeric fields
  - ▶ Provide highlights showing search hits in the field data
  - ▶ Autocomplete suggestions
  - ▶ Geospatial search
  - ▶ and many more!



## Amazon CloudSearch

- Allows you to **create a search domain**, specify an index and upload your data as documents
- Provisions and manages all the underlying servers and resources needed to build and deploy search indexes
- Simply upload your data to any data store, create a search domain in CloudSearch, and integrate it into your applications



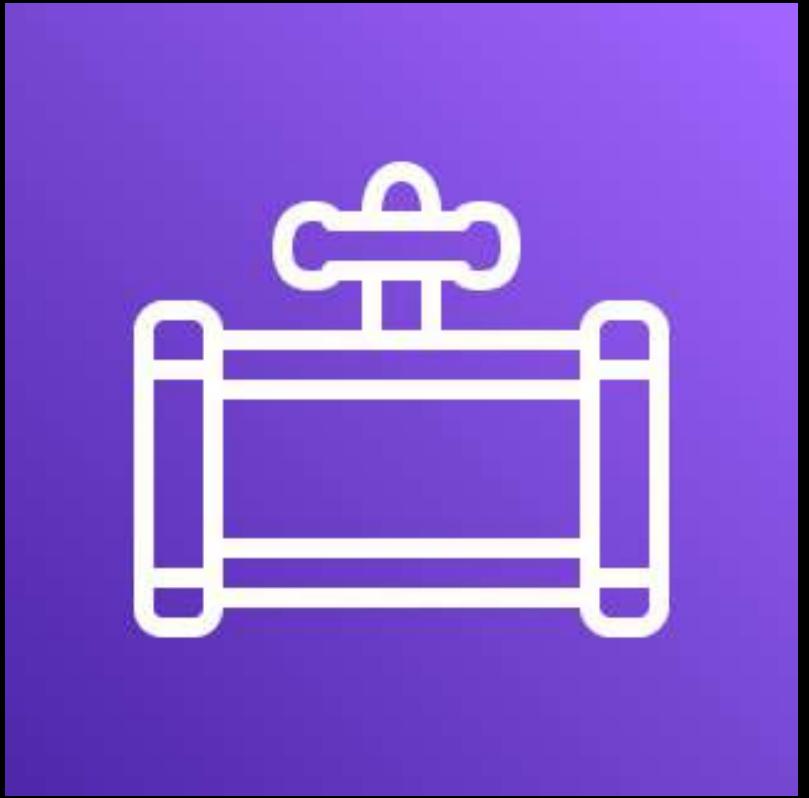
## Amazon Redshift

- A fast, scalable **data warehouse**
- Allows you to analyze all your data across your data warehouse and data lake
- Delivers faster performance than other data warehouses through the use of machine learning, massively parallel query execution and **columnar storage** on high-performance disks
- Can **run queries across petabytes of data** in your Redshift data warehouse and analyze exabytes of data in your S3 data lake
- Primarily used **for Online Analytical Processing (OLAP) applications** and reporting tools



## Amazon Redshift

- Redshift clusters run in internal Amazon EC2 instances that are configured as nodes
- You can select the particular node type and instance size that you prefer
- **Not a serverless** service
- Has a feature called **Redshift Spectrum** that allows you to query data from Amazon S3 without loading the entire data into Redshift tables
- Redshift Spectrum queries use massive parallelism to **quickly execute large datasets** at a fraction of the cost



## Amazon Data Pipeline

- A service that processes and **moves your data between different AWS compute and storage services**
- Enables you to process and move your data in specific intervals that you define to transfer your data to and from your on-premises data center
- Allows you to access, transform and process your data where it's stored at scale
- Empowers you to **transfer and store the results to various AWS services** such as Amazon S3, Amazon RDS, Amazon DynamoDB, and Amazon EMR

- A fully managed and serverless service that is primarily **used for extract, transform, and load workloads** or ETL



## AWS Glue

- Simplifies the process of preparing and loading your data before running your data analytics workload
- Creates a **Data Catalog** that allows you to specify and search your data that is stored on Amazon S3 and other AWS services
- Automatically discovers your data and store the associated metadata in the AWS Glue Data Catalog
- The data will be immediately searchable, queryable, and available for ETL once the metadata is stored



## Amazon Managed Streaming for Apache Kafka

- A **fully managed Apache Kafka service** in AWS
- Apache Kafka is an open-source platform that allows you to **build real-time streaming data pipelines** and applications
- Allows you to use Apache Kafka APIs to stream changes to and from different databases, populate your Amazon S3 data lakes, and empower machine learning and analytics applications

## AWS Lake Formation



- Makes it easy for you to **set up a secure data lake**
- Allows you to **create data catalogs** for your external data just like AWS Glue
- Collects and catalogs your data from different data sources and moves the data into a new Amazon S3 data lake
- **Classifies and processes your data** using machine learning algorithms, and secures access to your sensitive data
- Data can be queried and analyzed using Amazon Athena, Amazon Redshift, Amazon EMR, and other services



# IAM Overview



# Identity and Access Management

AUTHENTICATION

AUTHORIZATION

# Identity

## AUTHENTICATION

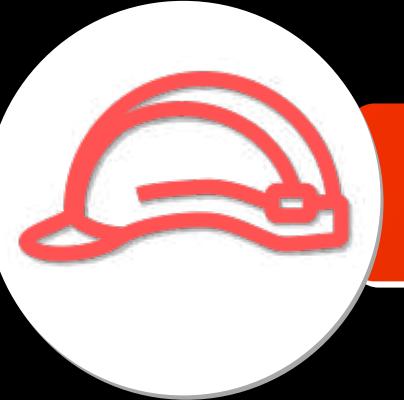
### IAM ENTITIES



**IAM USER**



**IAM GROUP**



**IAM ROLE**

### TYPES:

- Root User
- Regular IAM User

# Access Management

## AUTHORIZATION



**IAM POLICY**



**Permission 1**



**Permission 2**



**Permission 3**

AWS-managed Policy

Customer-managed Policy

Inline Policy



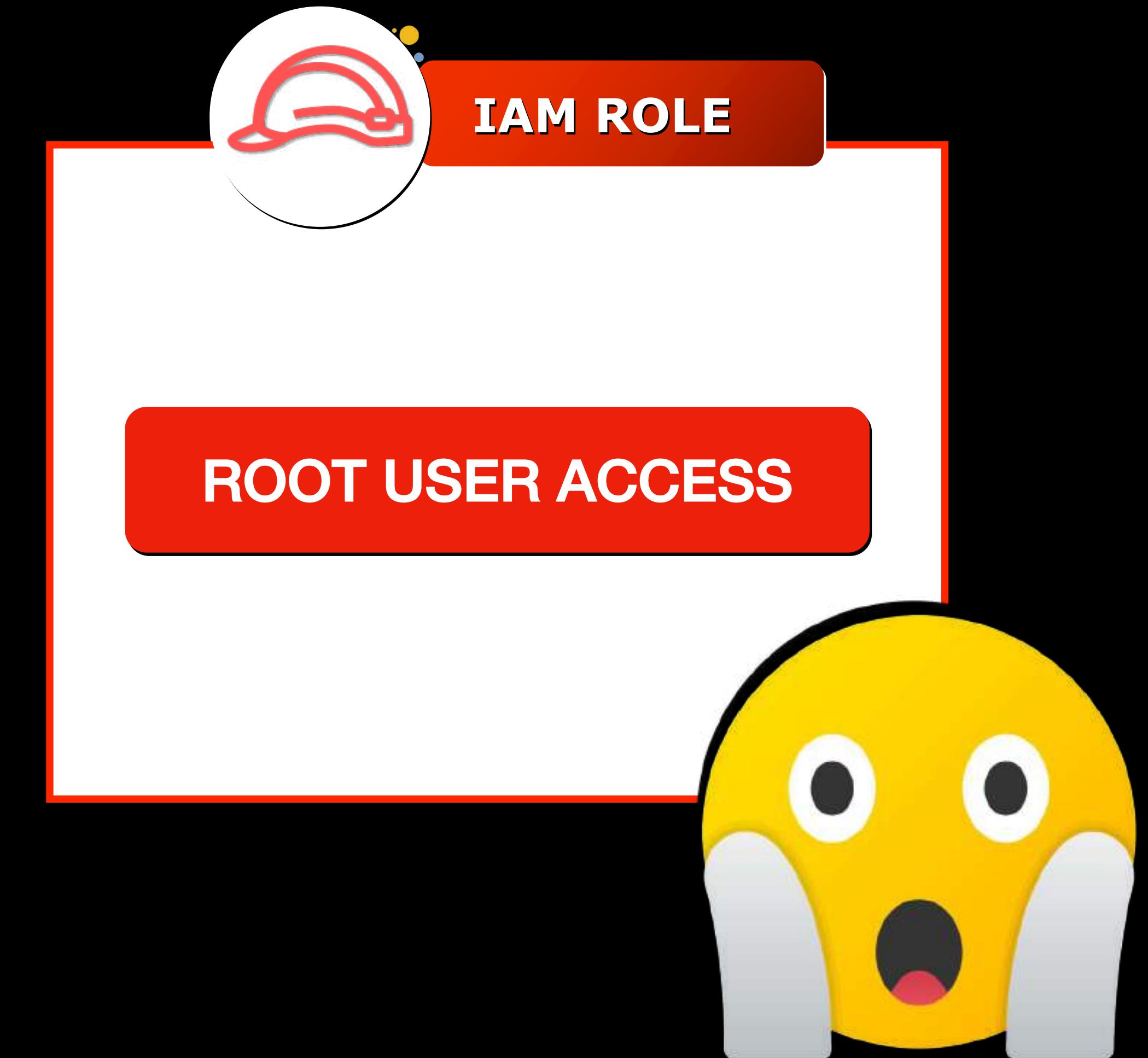
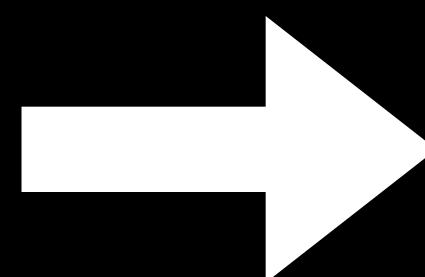
# Grant **Least** Privilege

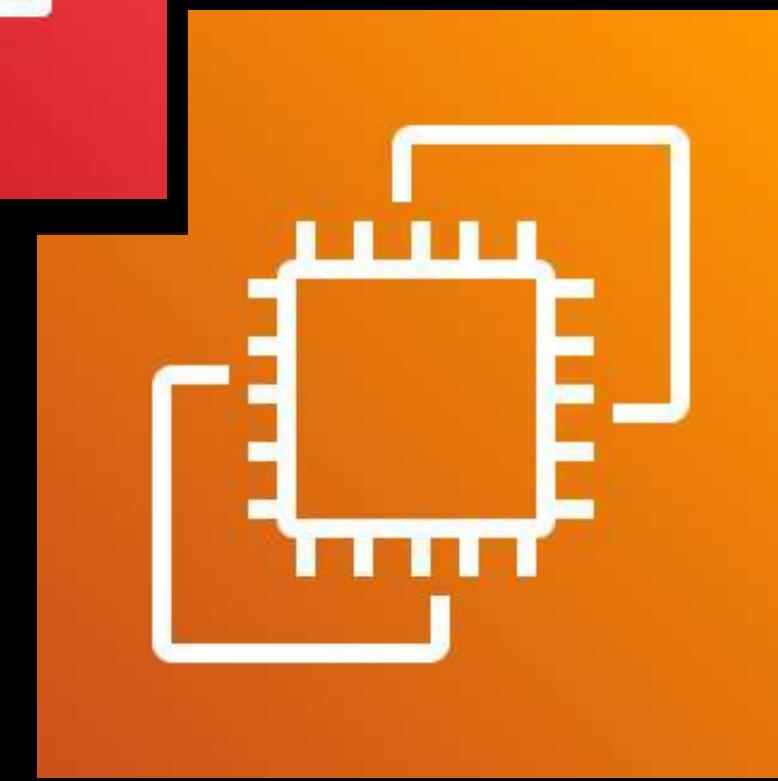
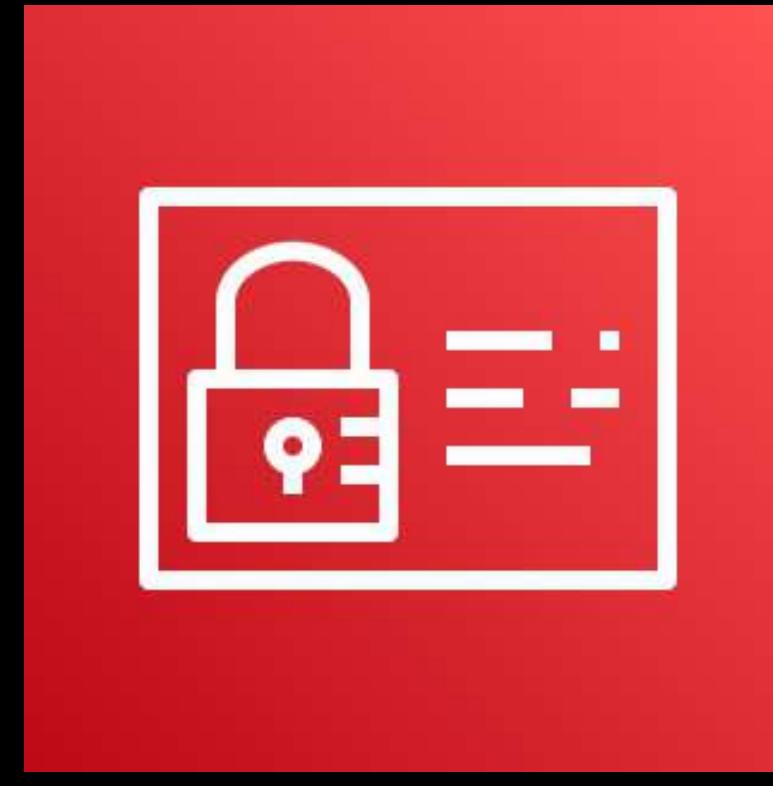


# Does not grant the least privilege



External User

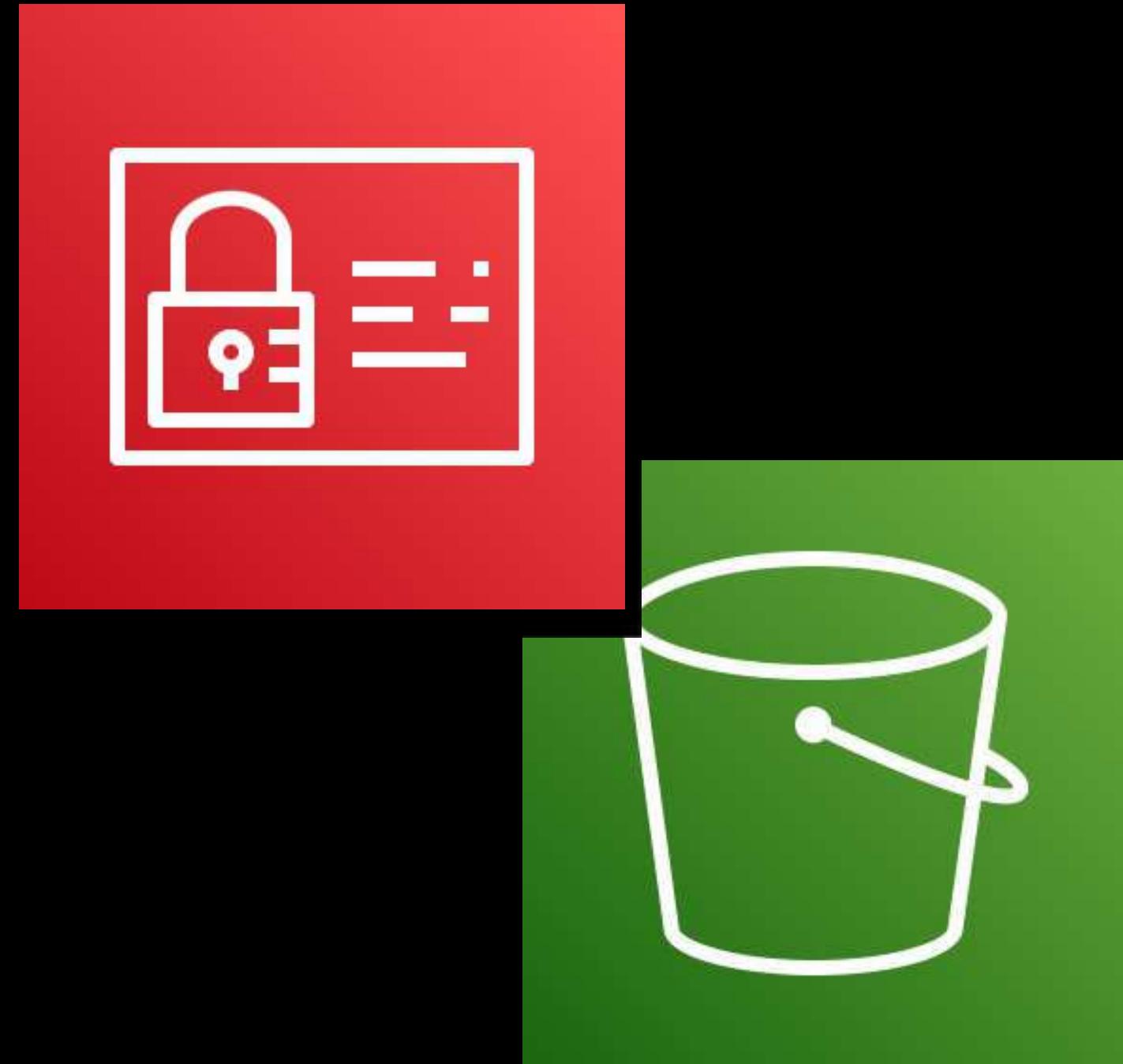




- Use the **Instance Profile** to pass a specific IAM role to your Amazon EC2 instance for it to perform certain actions
- IAM roles attached to your instance can also be viewed on your EC2 metadata.

```
curl http://169.254.169.254/latest/meta-data/iam/info
```

## Amazon EC2 and AWS IAM



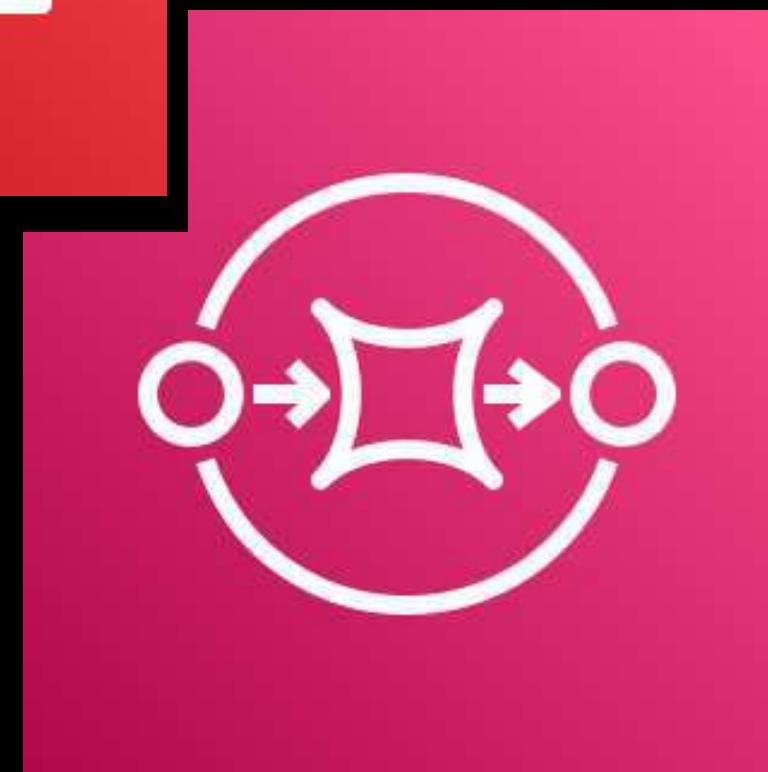
- You can set up a **bucket policy** to grant IAM users and other AWS accounts the access permissions for your bucket and its objects.
- In AWS Organization, you can set up an S3 bucket policy that allows cross-account access to other departments of your organization.

## Amazon S3 and AWS IAM



- For **DynamoDB**, you can design an **IAM policy** that allows access to put, update, and delete items in one specific table.
- **IAM DB Authentication** is a feature available for **Amazon RDS and Aurora**. This allows you to use **IAM** to centrally manage access to your database resources

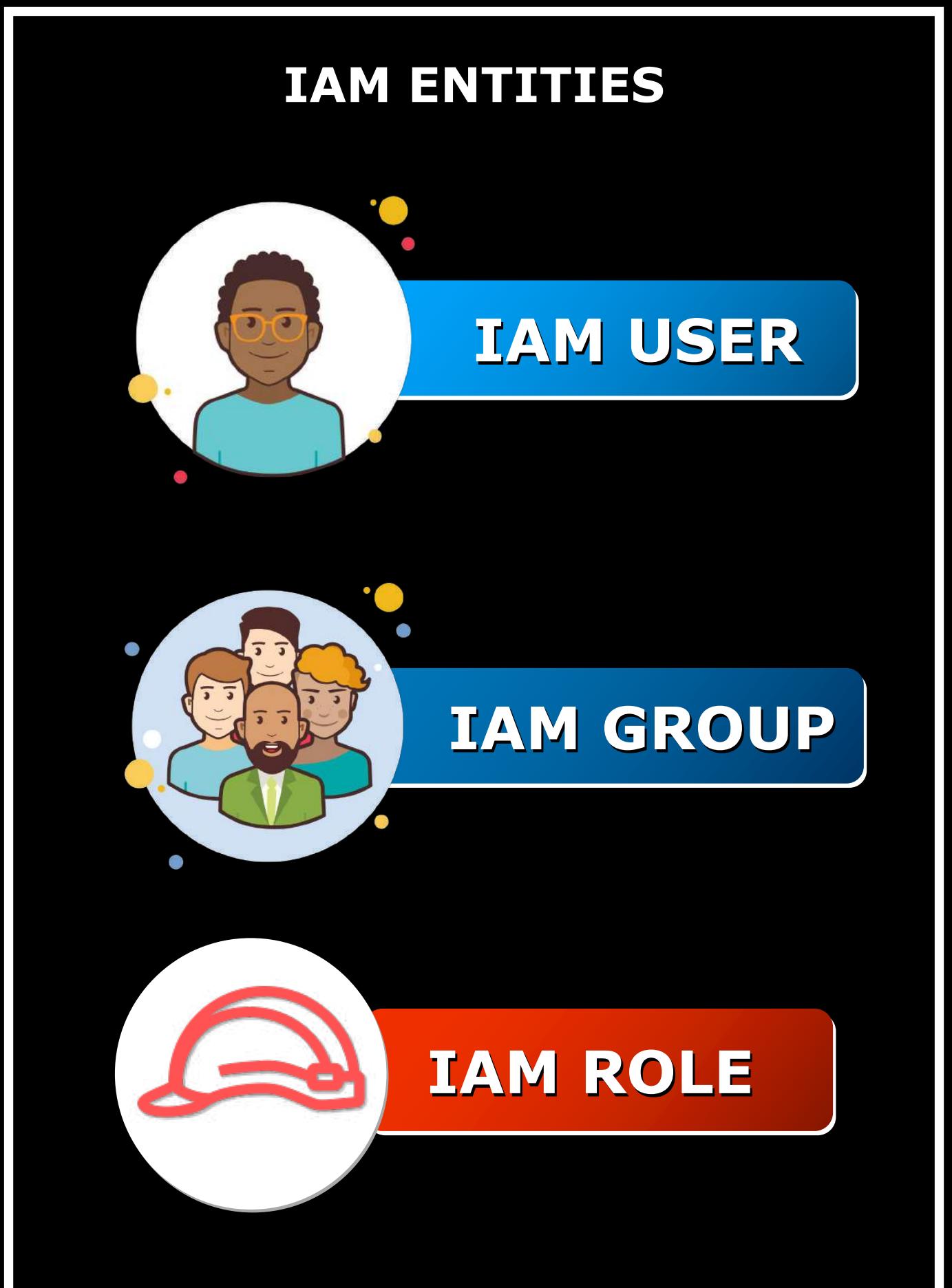
## AWS Databases and AWS IAM



- An **Access Policy** can be provisioned to control external access to your SQS queue.
- Helps you grant permissions to an external company to access your queue.
- An **SQS access policy** can allow external companies to poll the queue without giving up the permissions of your own account.

## Amazon SQS and AWS IAM

## IDENTITY-BASED POLICY



## RESOURCE-BASED POLICY



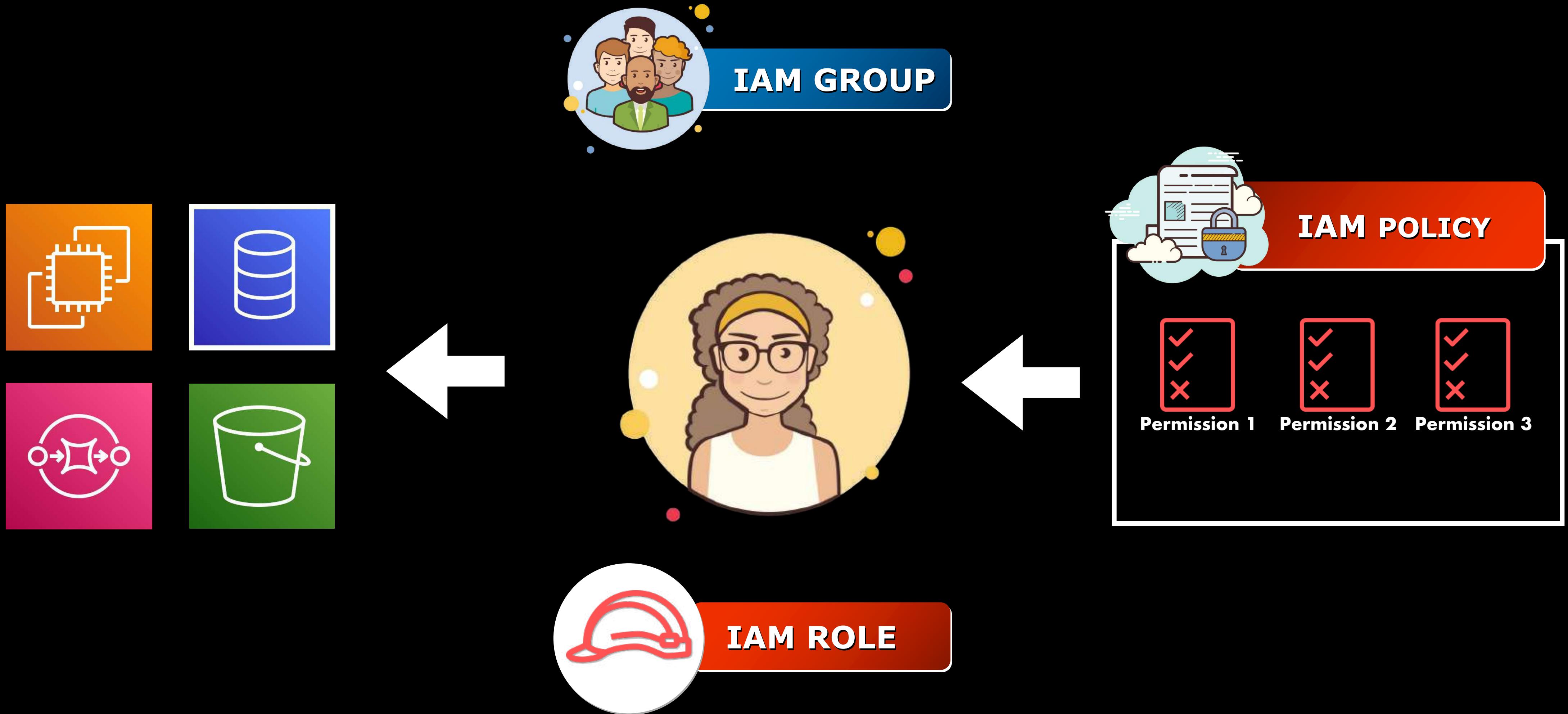
## PERMISSIONS BOUNDARY

- Allows you to set the **maximum permissions** that an identity-based policy can grant to an IAM entity.
- Ensure that the entity can only perform the actions that are allowed by both its identity-based policies and its **permissions boundaries**.



# IAM Identities

---



# IAM IDENTITIES



**IAM USER**



**IAM GROUP**



**IAM ROLE**



IAM USER

- An entity that represents an actual person or a service
- Can interact with your AWS resources using the AWS command-line interface, AWS API, or through the AWS management web console
- Provides someone the ability to sign in to the AWS Management Console and programmatic access to AWS APIs



IAM USER

Consists of:

- NAME
- PASSWORD
- ACCESS KEY PAIR



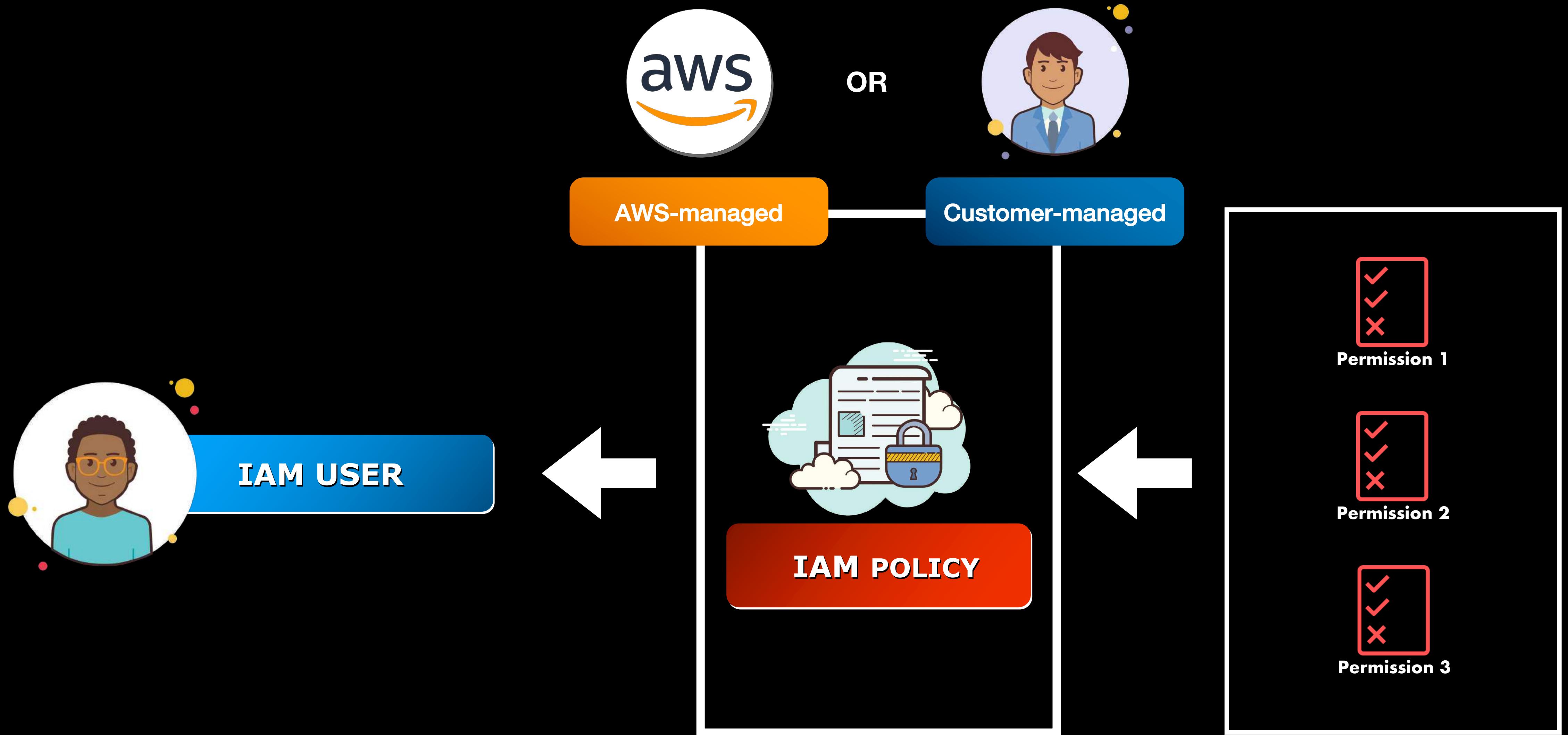
Access Key ID



Secret Access Key



- AWS CLI
- AWS APIs
- AWS SDKs
- AWS CDKs





## IAM POLICY TYPES

AWS-managed

- Managed by **AWS**
- **Cannot** be fully customized
- Has **AWS Managed-Policies** for **Job Functions** that you can readily use:
  - Administrator
  - Support User
  - Security Auditor
  - Network Administrator
  - Developer Power User
  - Billing
  - ...and others



Customer-managed

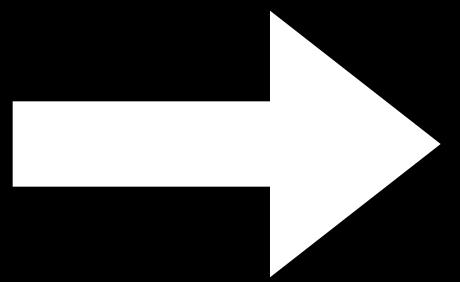
- Managed by **you (the customer)**
- **Can** be fully customized
- You have to manually create a policy for a particular job function



IAM USER



IAM USER



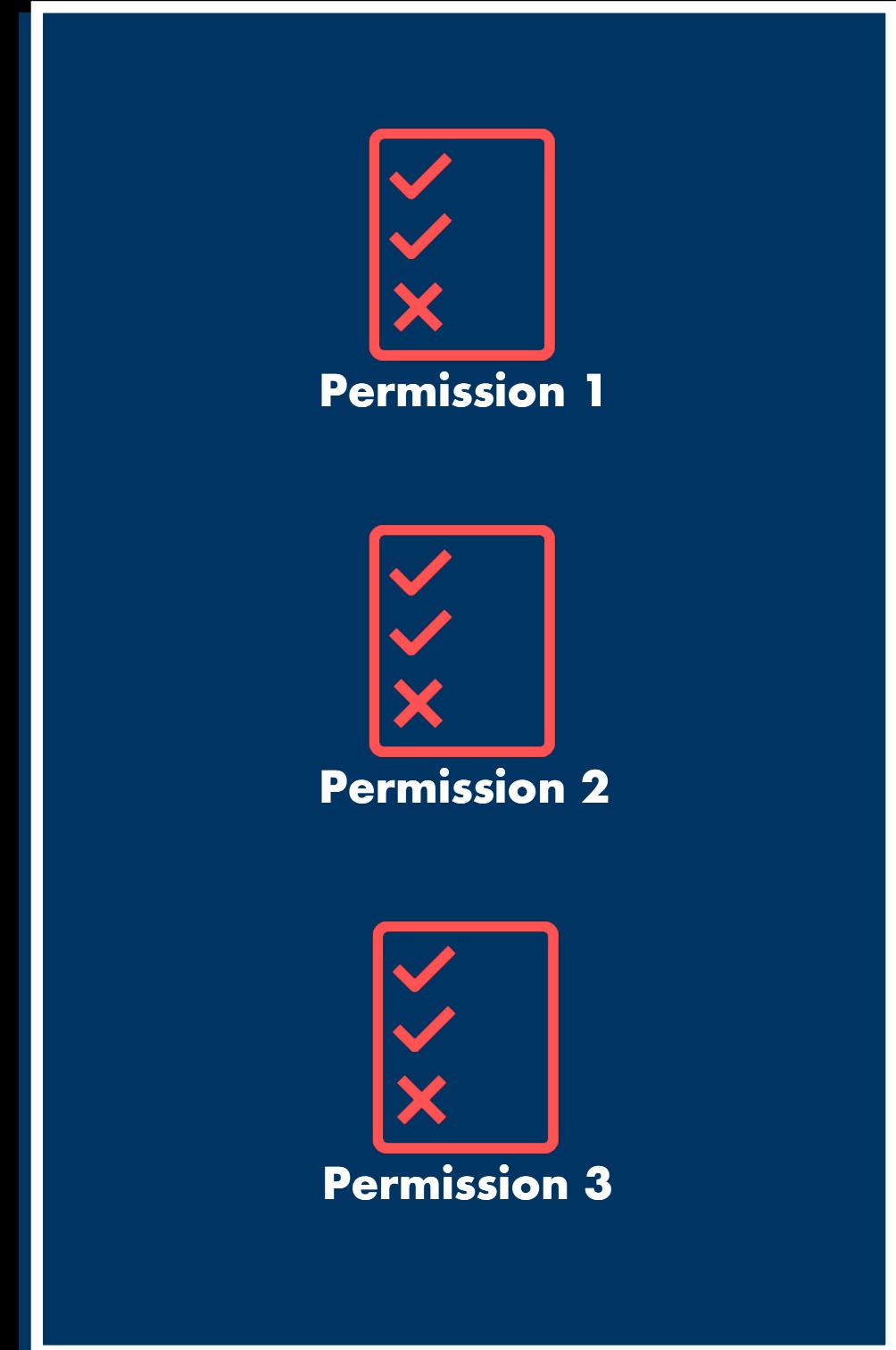
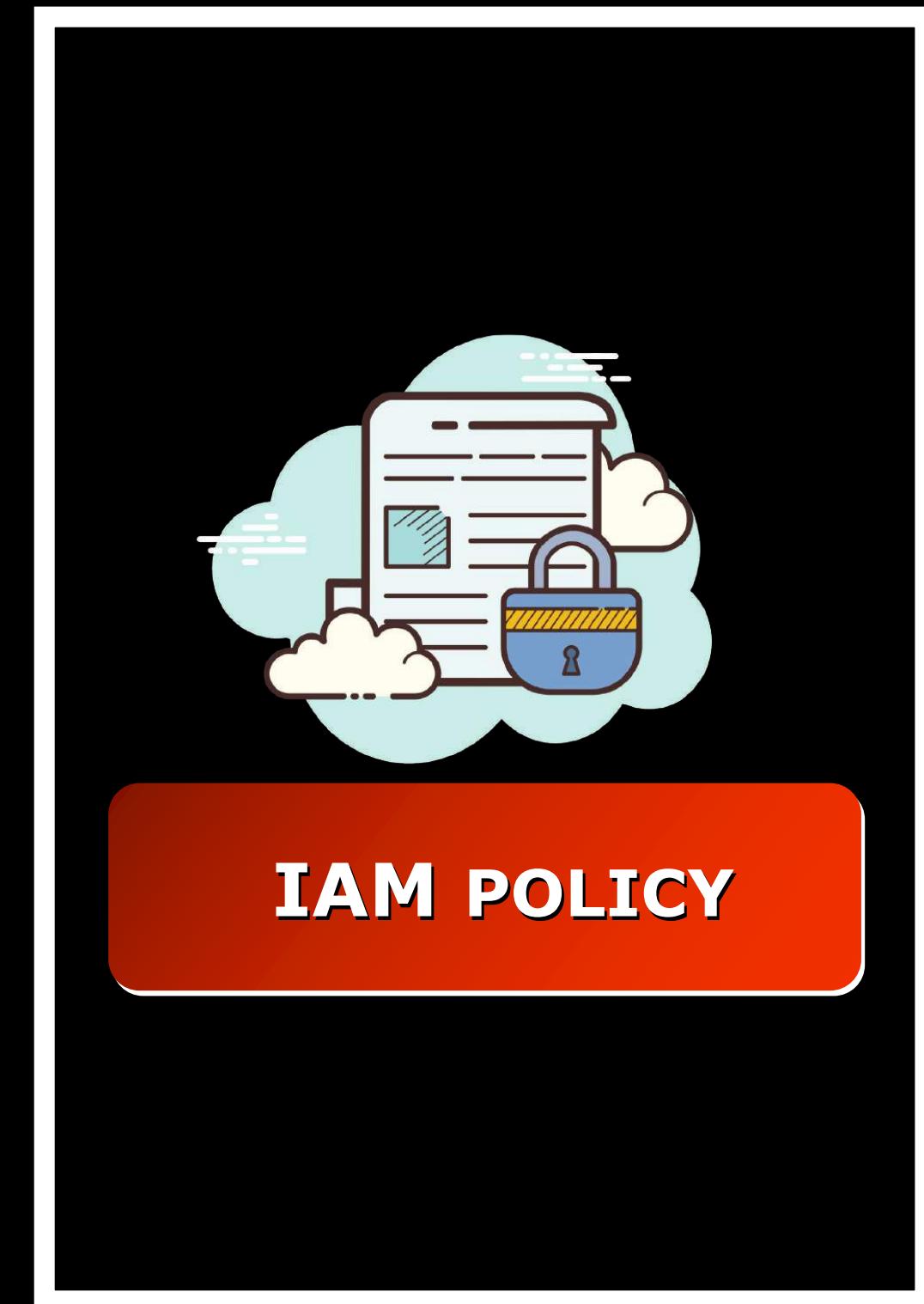
IAM GROUP

Welcome to  
the Group!



## IAM GROUP

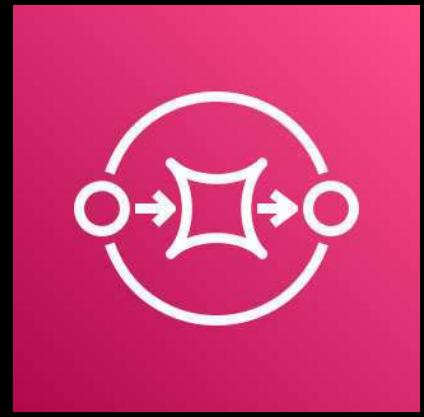
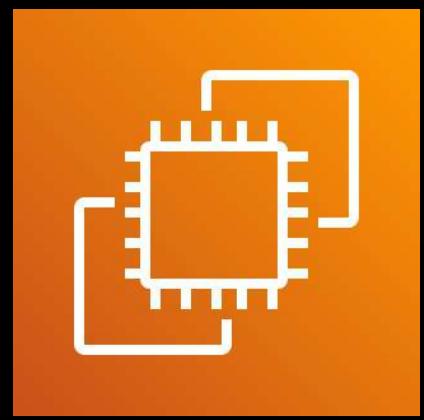
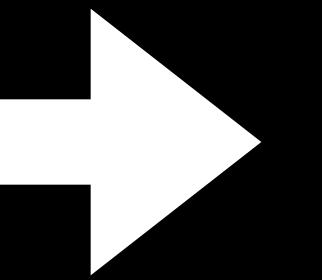
- Can contain multiple IAM Users
- A single IAM User can belong to multiple IAM Groups
  - Cannot be nested
- It can only contain IAM users and not other IAM Groups
- There is no default user group that automatically includes all of the IAM Users in your AWS account





IAM ROLE

assumed by





IAM ROLE



IAM USER

- Intended to be **assumed** by one or more AWS resources
  - No long-term credentials
- Uniquely **associated** with one single person only
  - Has long-term credentials:
    - AWS Management Console password
    - Access Keys



US - AWS ACCOUNT #1

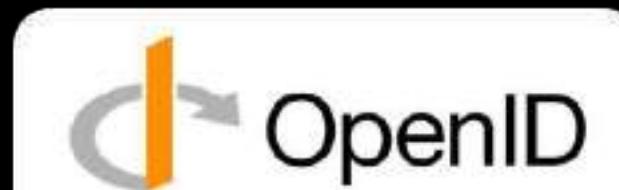
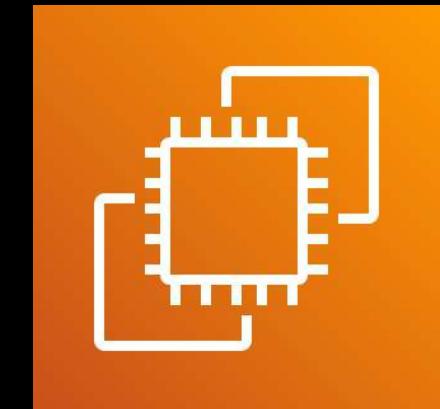


IAM ROLE

CROSS-ACCOUNT



INDIA - AWS ACCOUNT #2

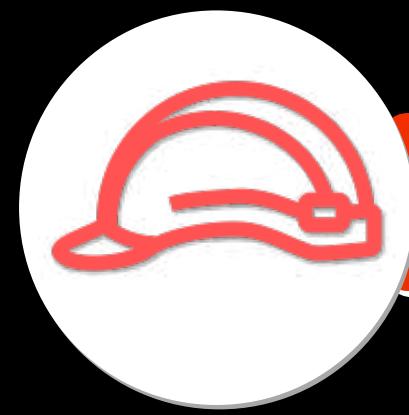


S A M L

Security Assertion Markup Language



Amazon Cognito



IAM ROLE

AWS SERVICE ROLE

AWS SERVICE-LINKED  
ROLE

- Grants access to your resources in one account to a trusted principal in a different AWS account
- Assumed by an AWS service or applications running in your EC2 instance
- Limited within your AWS account only
- The custom applications hosted in Amazon EC2 can assume an AWS service role to perform certain actions
- A predefined role that is directly linked to an AWS service



# IAM Policy Types

# IAM IDENTITIES



# RESOURCES





## IAM POLICY

- Contains permissions that explicitly **ALLOW** or **DENY** access to certain AWS services
- It provides fine-grained **access control** to specific API actions as well as the AWS resources that the policy should be applied to



## IAM POLICY

```
1 {  
2   "Version": "2012-10-17",  
3   "Statement": [  
4     {  
5       "Sid": "TutorialsDojo",  
6       "Effect": "Allow",  
7       "Action": "s3:PutObject",  
8       "Resource": "arn:aws:s3:::tutorialsdojo-manila/*"  
9     }  
10   ]  
11 }
```

**API action**      s3:PutObject

ALLOWS THE API ACTIONS  
YOU SPECIFY



## IAM POLICY

```
1 {  
2   "Version": "2012-10-17",  
3   "Statement": {  
4     "Sid": "tutorialsdojo",  
5     "Effect": "Deny", ← DENIES THE API ACTIONS  
6     "Action": [  
7       "lambda>CreateFunction",  
8       "lambda>DeleteFunction"  
9     ],  
10    "Resource": "*",  
11    "Condition": {  
12      "IpAddress": {  
13        "aws:SourceIp": "220.110.16.0/20"  
14      }  
15    }  
16  }  
17}  
18 }
```

**API actions** (highlighted in orange box): "lambda>CreateFunction", "lambda>DeleteFunction".

**IP Condition** (highlighted in blue box): "aws:SourceIp": "220.110.16.0/20".



## IAM POLICY

```
1 {  
2   "Version": "2012-10-17",  
3   "Statement": {  
4     "Sid": "tutorialsdojo",  
5     "Effect": "Deny",  
6     "Action": [  
7       "lambda>CreateFunction",  
8       "lambda>DeleteFunction"  
9     ],  
10    "Resource": "*",  
11    "Condition": {  
12      "BoolIfExists": {  
13        "aws:MultiFactorAuthPresent": "false"  
14      }  
15    }  
16  }  
17 }
```

API actions

MFA Condition

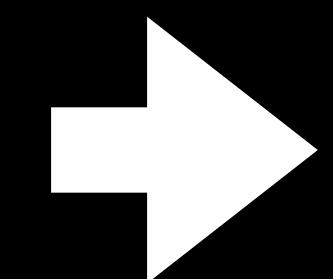


Multi-Factor Authentication  
(MFA)

## JSON EDITOR

```
1 {  
2     "Version": "2012-10-17",  
3     "Statement": [  
4         {"Sid": "tutorialsdojo",  
5          "Effect": "Deny",  
6          "Action": [  
7              "lambda>CreateFunction",  
8              "lambda>DeleteFunction"  
9          ],  
10         "Resource": "*",  
11         "Condition": {  
12             "BoolIfExists": {  
13                 "aws:MultiFactorAuthPresent": "false"  
14             }  
15         }  
16     }  
17 }
```

## VISUAL EDITOR



DENY Lambda (2 actions)

Clone | Remove

Service Lambda

Actions Write

CreateFunction  
DeleteFunction

Resources All resources

Request aws:MultiFactorAuthPresent (If exists, Bool false)  
conditions

## Standalone Policy



## Inline Policy

- **Remains unchanged even if you delete its associated IAM identity**
- **It doesn't have a strict one-to-one relationship to its associated IAM identity**
- **Will be automatically be deleted if you delete its associated identity**
- **Has a strict one-to-one relationship to its associated IAM identity**



## IAM Policy Types

- **Identity-based Policies**
- **Resource-based Policies**
- **Permissions Boundaries**
- **AWS Organizations SCPs**
- **S3 Access Control Lists (ACLs)**
- **Session Policies**

## Identity-Based Policy

- A policy that you attach to an IAM Identity
- Two Types:

### Managed Policies

- A type of a standalone policy
- Can either be AWS managed or Customer-managed

### Inline Policies

- Maintains a strict one-to-one relationship between a policy and an IAM identity.
- Tightly-coupled with its associated IAM Identity

## Resource-Based Policy

- Attaches an inline policy to a specific AWS Resource
- Types:



S3 Bucket Policy



SQS Access Policy

Permissions   Trust relationships   Tags   Access Advisor   Revoke sessions

You can view the trusted entities that can assume the role and the access conditions for the role. [Show policy document](#)

[Edit trust relationship](#)

Trusted entities

The following trusted entities can assume this role.

Trusted entities

The identity provider(s) monitoring.rds.amazonaws.com

Conditions

The following conditions define how and when trusted entities can assume the role.

There are no conditions associated with this role.



IAM ROLE

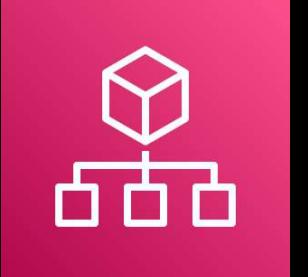
Trust Policy

## Permissions Boundaries

- Defines the **maximum permissions** that an identity-based policy can grant to an IAM entity
- Does not explicitly grant permissions
- Sets a **clear boundary** to ensure that a given IAM policy will not over-provision the permissions to your AWS resources

## Service Control Policies (SCPs)

- Primarily used in:



AWS Organizations

- Defines the **maximum permissions for account members of an organization or organizational unit.**
- Limits the **permissions that identity-based policies or resource-based policies grant to the IAM users or roles within the AWS account**
- **IAM policies can't restrict the AWS account root user. In the contrary, the specified actions from an attached SCP can affect all IAM identities, including the root user, of the member account**



- **Primarily used in:**  **Amazon S3**
- **Controls which principals in other AWS accounts can access a particular bucket**
- **These are cross-account permission policies that grant certain permissions to a specified principal that you define**
- **ACLs cannot grant permissions to entities within the same account**

## Sessions Policies

- Limits the permissions that an identity-based policy grants to a particular session
- Works like **Permissions Boundaries**
- Sets a limit of what kind of permission a session has, without granting any permissions.
- Aside from an identity-based policy, the permissions of a session policy can also come from a resource-based policy
- If there's an explicit deny in any of the policies, then it will effectively override any allowed permissions



# IAM Policy Basics

---

## Policy-wide Information

```
{  
  "Id": "TutorialsDojoPolicy1",  
  "Version": "2012-10-17",  
  
  "Statement": [  
    {  
      "Sid": "AllowAllActionsOnBooksTable",  
      "Effect": "Allow",  
      "Action": "dynamodb:*",  
      "Resource": "arn:aws:dynamodb:us-west-2:123456789012:table/Books"  
    }  
  ]  
}
```

## Statements

Logical OR

```
"Sid": "ListObjectsInBucket",  
"Effect": "Allow",  
"Action": ["s3>ListBucket", "s3>DeleteObject"],  
"Resource": ["arn:aws:s3:::tutorialsdojo-manila"]  
}  
]  
}
```

## IAM Statement Elements

```
{  
    "Sid": "AllowActionsOnBooksTable",  
    "Effect": "Allow",  
    "Principal": { "AWS": "arn:aws:iam::123456789012:root" }  
    "Action": [  
        "dynamodb:DeleteItem",  
        "dynamodb:PutItem",  
        "s3:*",  
        "dynamodb:UpdateItem"  
    ],  
    "Resource": "arn:aws:s3:::tutorialsdojo/*",  
    "Condition": {  
        "IpAddress": {  
            "aws:SourceIp": "220.110.16.0/20"  
        }  
    }  
}
```

Statement ID

ALLOW or DENY

CONDITION ELEMENT

# CONDITION ELEMENT

- **String**
- **Numeric**
- **Date**
- **Boolean**
- **Binary**
- **ARN**
- **IfExists**
- **IpAddress**
- **...and many more!**

```
"Condition" : { "{condition-operator}" : { "{condition-key}" : "{condition-value}" }}
```

StringEquals

StringNotEquals

StringEqualsIgnoreCase

StringNotEqualsIgnoreCase

StringLike

StringNotLike

DateEquals

DateNotEquals

DateLessThan

DateLessThanEquals

DateGreaterThan

DateGreaterThanOrEqual

NumericEquals

NumericNotEquals

NumericLessThan

NumericLessThanEquals

NumericGreaterThan

NumericGreaterThanOrEqual

ArnEquals, ArnLike

ArnNotEquals,

ArnNotLike

IpAddress

NotIpAddress

## CONDITION ELEMENT

IfExists

- **StringEqualsIfExists**
- **NumericEqualsIfExists**
- **BoolIfExists**
- **IpAddressIfExists**
- etc...



Shares the Amazon S3 bucket named `tutorialsdojo-manila` with an external vendor while ensuring that the **bucket owner** is still be able to access all objects

```
...  
  "Action": [  
    "s3:PutObject"  
  ],  
  "Resource": "arn:aws:s3:::tutorialsdojo-manila/*",  
  "Condition": {  
    "StringEquals": {  
      "s3:x-amz-acl": "bucket-owner-full-control"  
    }  
  }  
...  
  ...
```

Users will be denied of all API actions ( except for the **s3:PutObject** action ) if their **multi-factor authentication (MFA)** is not enabled

```
{  
  "Version": "2012-10-17",  
  "Statement": [ {  
    "Sid": "DenyAllTDojoUsersNotUsingMFA",  
    "Effect": "Deny",  
    "NotAction": "s3:PutObject",  
    "Resource": "*",  
    "Condition": {  
      "BoolIfExists": {  
        "aws:MultiFactorAuthPresent": "false"  
      }  
    } ]  
}
```



# IAM Policy Evaluation Logic



IAM POLICY

```
{  
  "Id": "TutorialsDojoPolicy1",  
  "Version": "2012-10-17",  
  
  "Statement": [  
    {  
      "Effect": "Allow",  
      "Action": "lambda:*",  
      "Resource": "*"  
    },  
    {  
      "Effect": "Deny",  
      "Action": ["lambda>CreateFunction", "lambda>DeleteFunction"],  
      "Resource": "*"  
    }  
  ]  
}
```

Allows the API Action



Denies the API Action

Logical OR



Will the API  
action be  
Allowed or  
Denied?



- 1. Authentication**
- 2. Process the request context**
- 3. Evaluate all policies within a single account**



If the IAM policies are within a **single account**...

All requests will be implicitly denied

Process the explicit **ALLOW** statements for identity-based or resource-based policy

An explicit **DENY** in any policy overrides any type of **ALLOW** actions

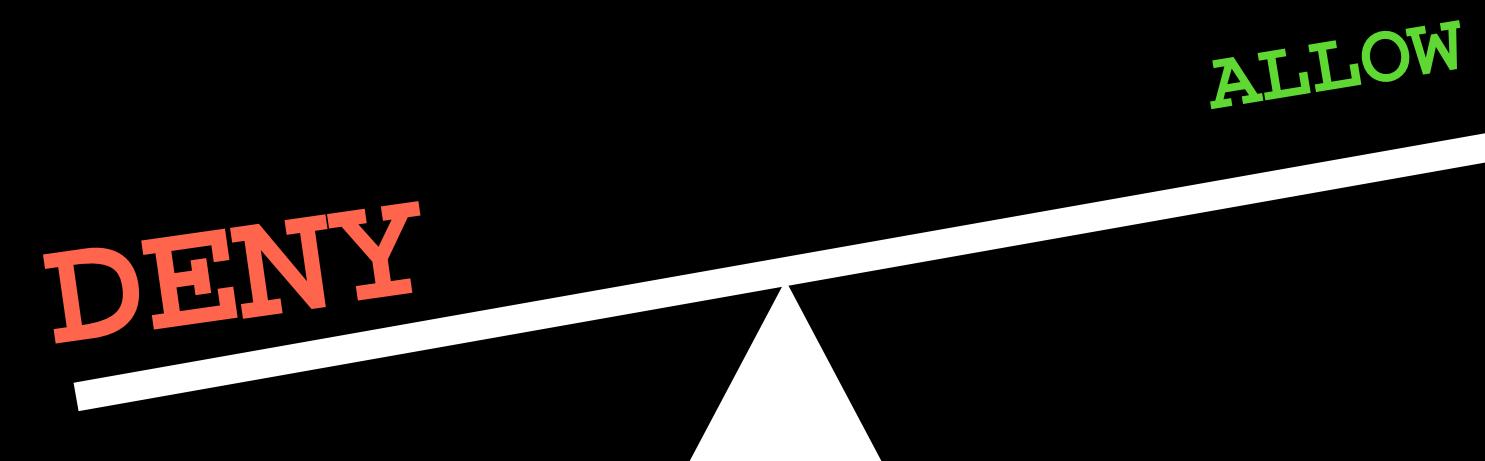


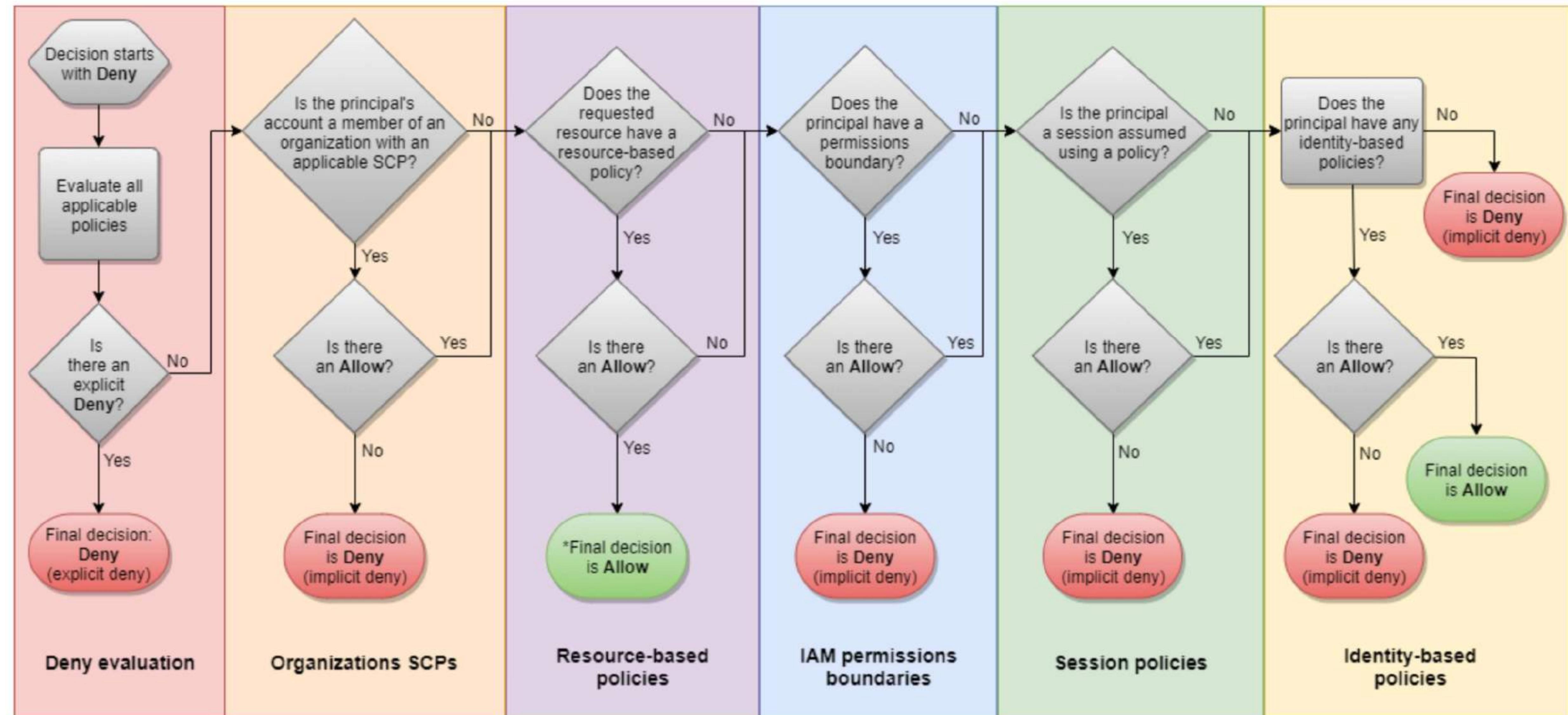
Except for the **AWS account root user**

Permissions Boundaries

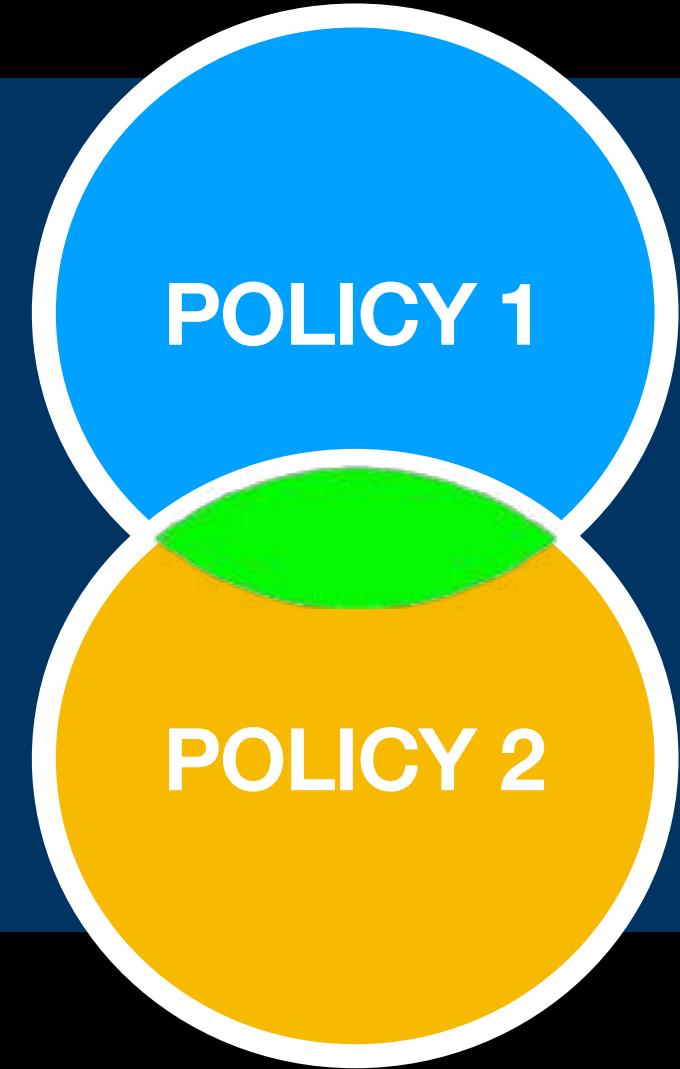
Sessions Policies

Service Control Policies (SCPs)



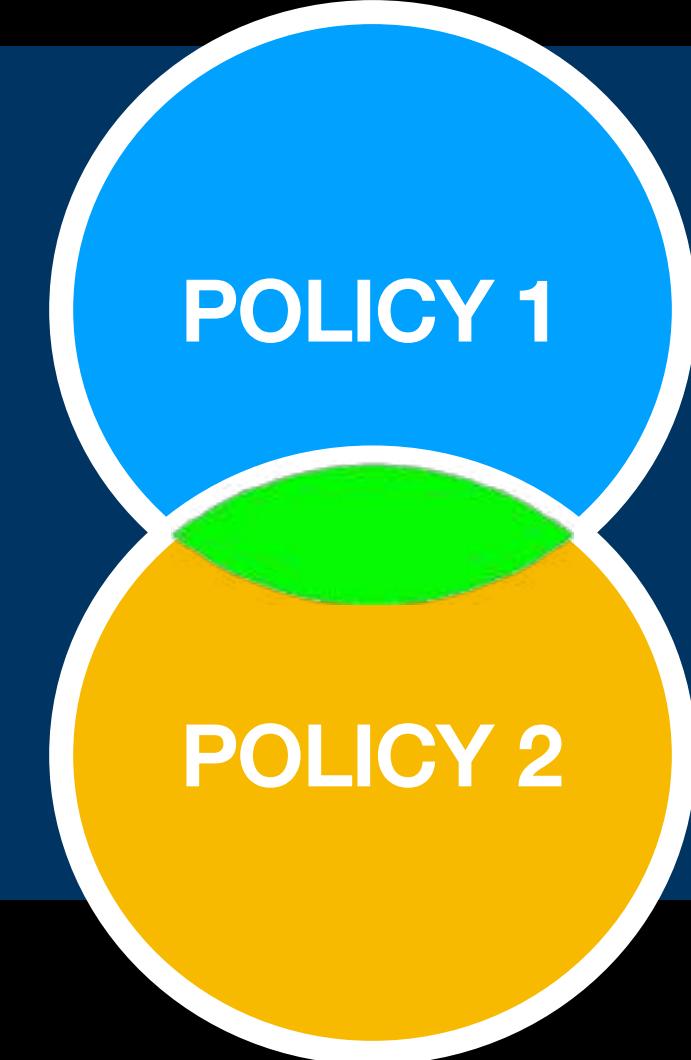


```
{  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Effect": "Allow",  
      "Action": "ec2:TerminateInstances",  
      "Resource": "*",  
      "Condition": {  
        "IpAddress": {  
          "aws:SourceIp": "49.147.194.0/24"  
        }  
      }  
    },  
    {  
      "Effect": "Deny",  
      "Action": "ec2:*",  
      "Resource": "*",  
      "Condition": {  
        "StringNotEquals": {  
          "ec2:Region": "us-west-1"  
        }  
      }  
    }  
  ]  
}
```



This policy will allow you to **terminate** an Amazon EC2 instance in the **us-west-1** region as long as your source IP is **within** the **49.147.194.0/24** CIDR block.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ec2:*",
        "ds:/*"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Deny",
      "Action": "ds:Delete*",
      "Resource": "*"
    }
  ]
}
```



This policy provides full access to Amazon EC2. It also allows creating, reading and updating the AWS Directory Service (DS) directories but not delete them.

```
{  
    "Version": "2012-10-17",  
    "Statement": [  
        {  
            "Effect": "Allow",  
            "Action": "lambda:*",  
            "Resource": "*"  
        },  
        {  
            "Effect": "Deny",  
            "Action": [  
                "lambda>CreateFunction",  
                "lambda>DeleteFunction"  
            ]  
            "Resource": "*",  
            "Condition": {  
                "IpAddress": {  
                    "aws:SourceIp": "220.200.16.0/24"  
                }  
            }  
        }  
    ]  
}
```



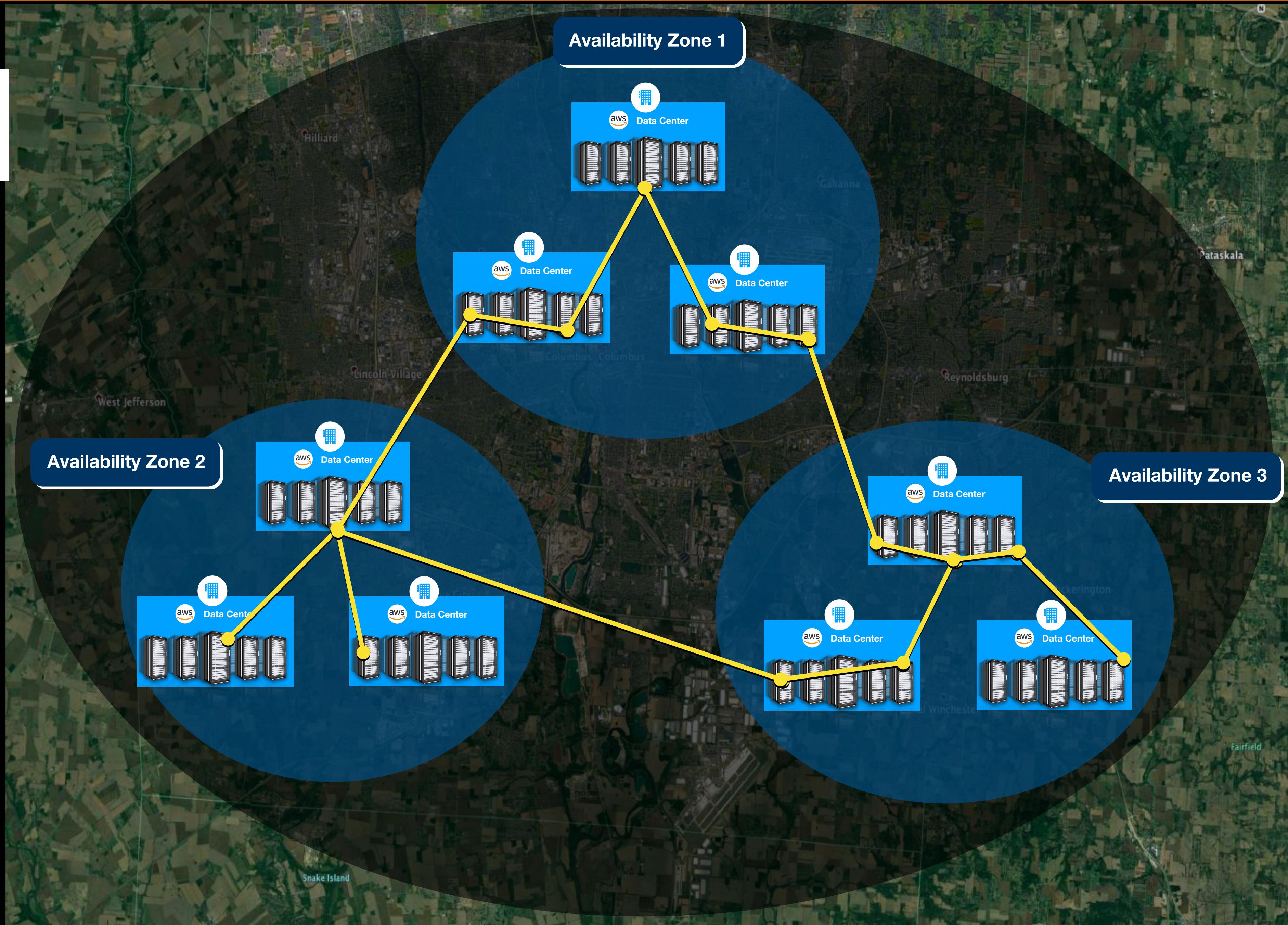
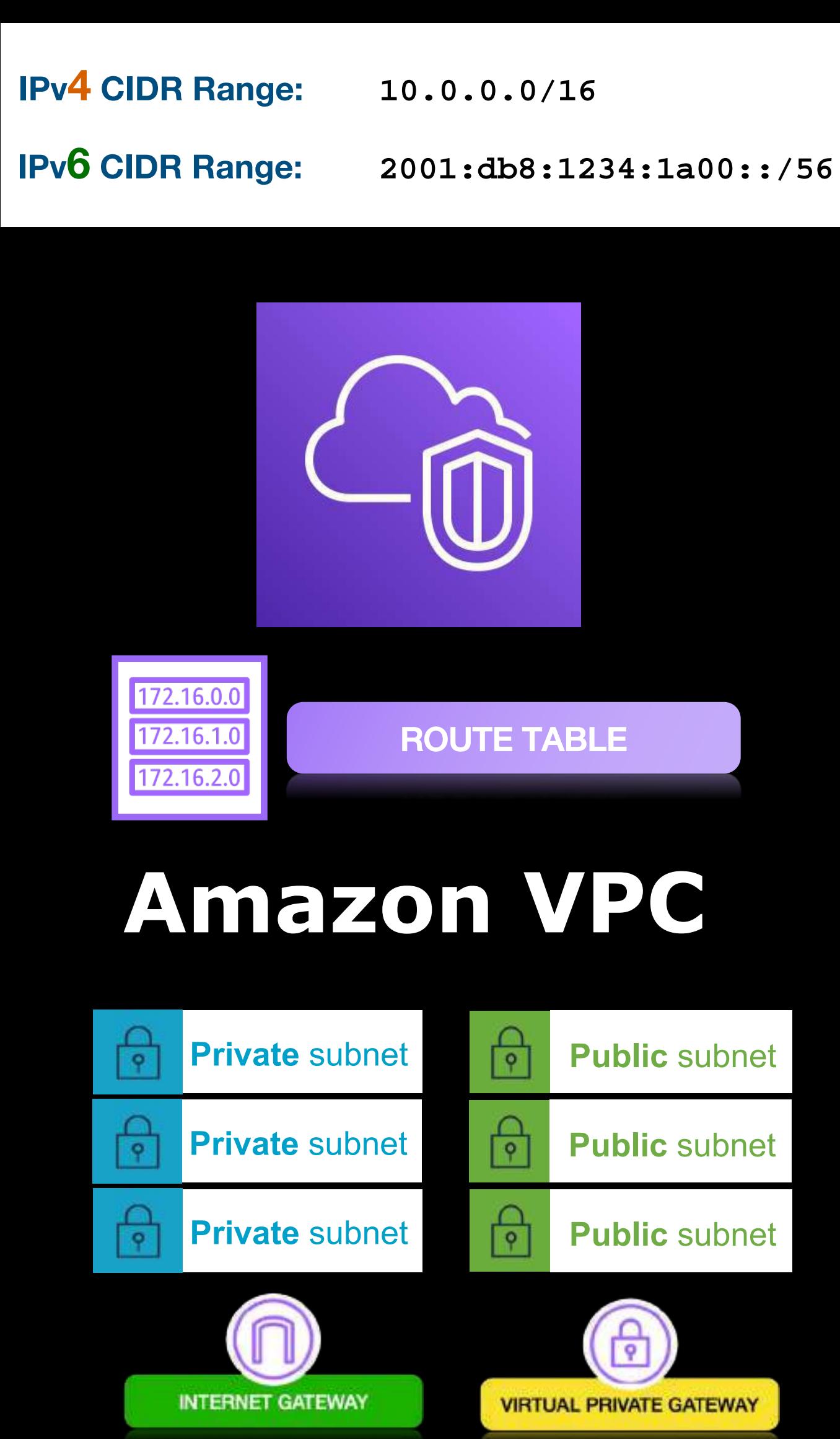
This policy will allow you to **terminate** an Amazon EC2 instance in the **us-west-1** region as long as your source IP is **within** the **49.147.194.0/24** CIDR block.



# Amazon VPC Overview

---

# US East (Ohio) us-east-2





CLOUD

## REGION

Amazon  
VPC

IPv4 CIDR Range: 10.0.0.0/16

IPv6 CIDR Range: 2001:db8:1234:1a00::/56

172.16.0.0  
172.16.1.0  
172.16.2.0

ROUTE TABLE



10.0.0.0/24

- A subnet **must reside entirely within one Availability Zone only**
- One subnet **cannot span to two or more AZs.**



10.0.1.0/24

- You can **have multiple subnets in the same Availability Zone.**





CLOUD

## REGION



Amazon VPC

IPv4 CIDR Range: 10.0.0.0/16

IPv6 CIDR Range: 2001:db8:1234:1a00::/56

172.16.0.0
172.16.1.0
172.16.2.0

ROUTE TABLE

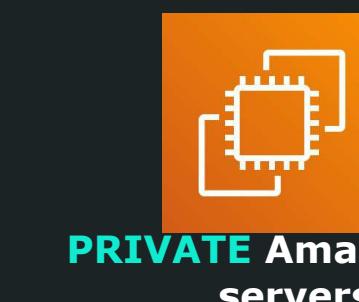
Private subnet



- For backend systems like **databases** or **application servers** that are not meant to be accessed publicly



Amazon EFS    Amazon RDS    Amazon FSx

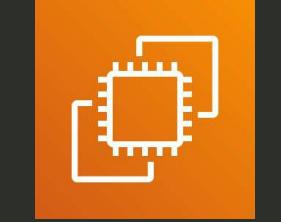
PRIVATE Amazon EC2  
servers

Security Group

Public subnet



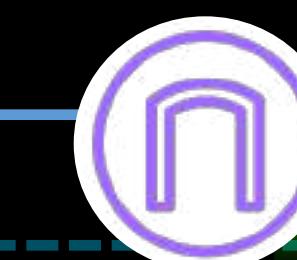
- For **publicly accessible web servers** and resources
- This subnet has a connection to the Internet Gateway of the VPC



PUBLIC Amazon EC2 web servers



Security Group



INTERNET GATEWAY

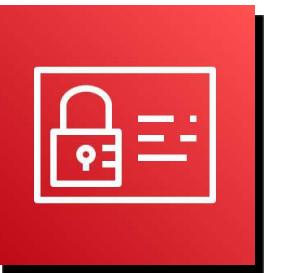


# Anatomy of an **Amazon VPC**

---



CLOUD



AWS IAM



## REGION



IPv4 CIDR Range: 10.0.0.0/16  
IPv6 CIDR Range: 2001:db8:1234:1a00::/56

Amazon VPC



Private subnet



Network ACL



Amazon EFS



Amazon RDS



Amazon FSx

PRIVATE Amazon EC2 servers



Security Group



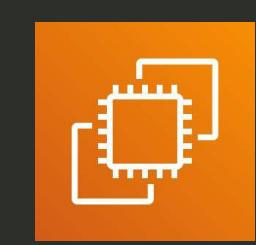
ROUTE TABLE



Public subnet



Network ACL



PUBLIC Amazon EC2 web servers



Security Group



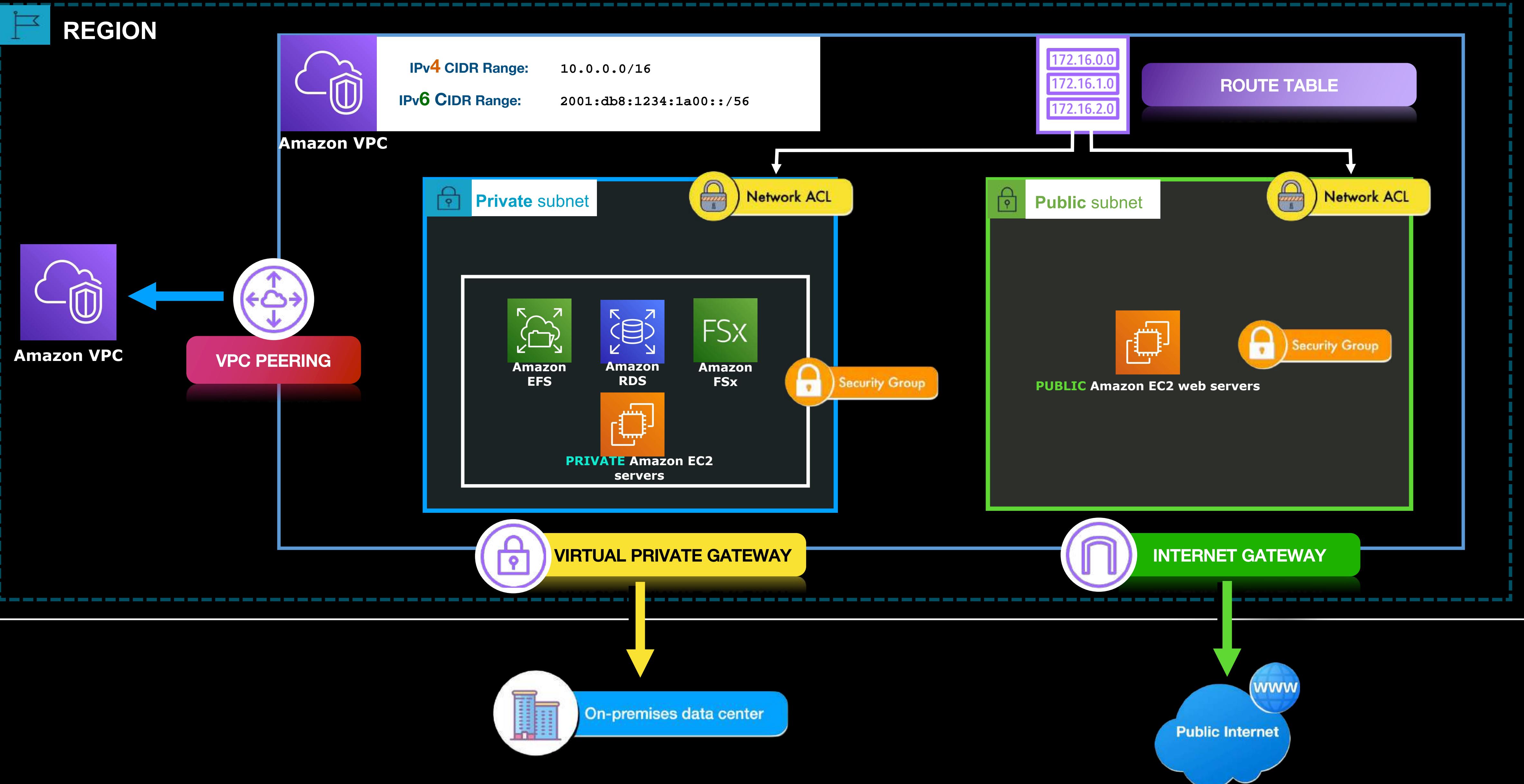
VIRTUAL PRIVATE GATEWAY

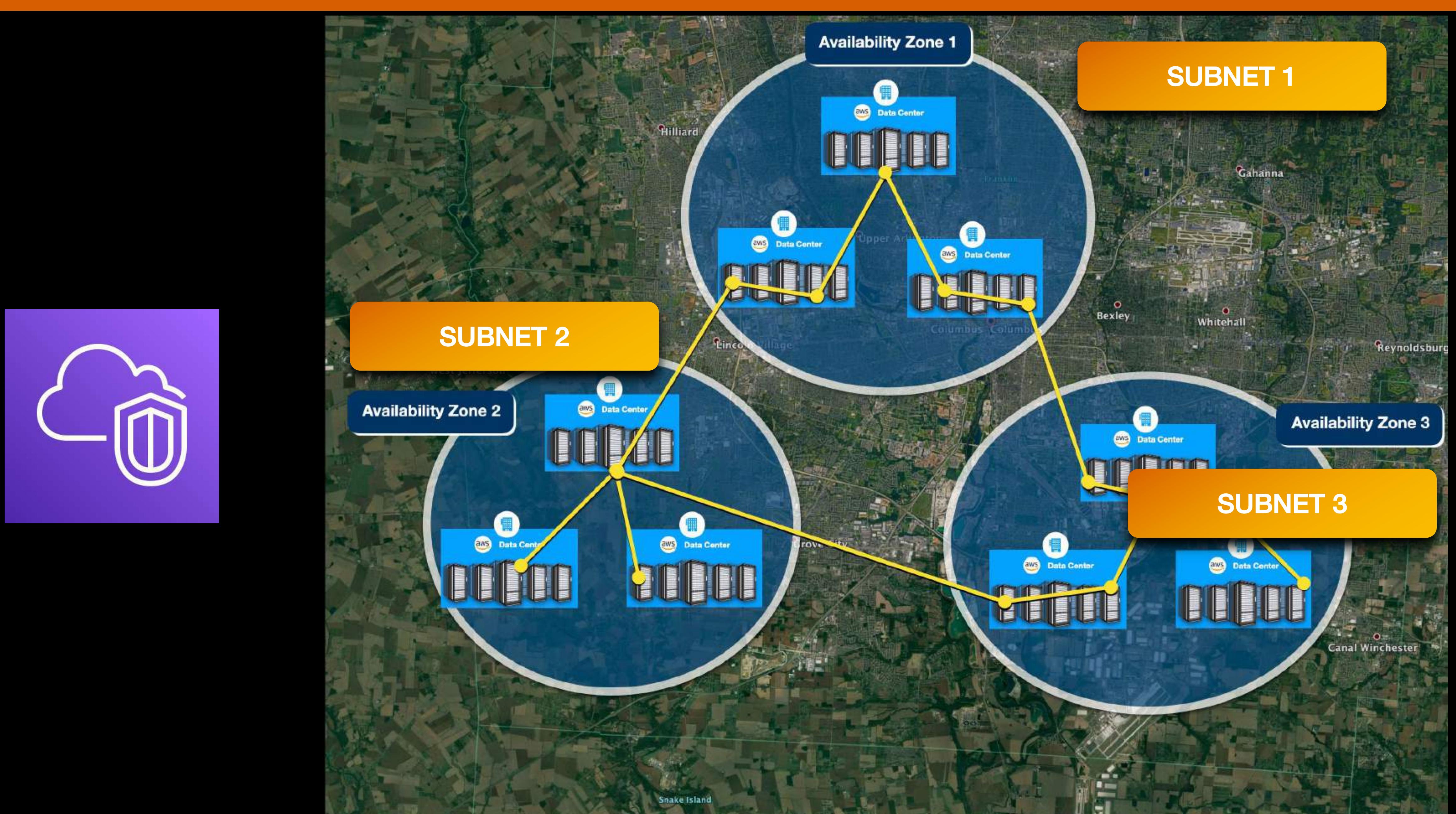


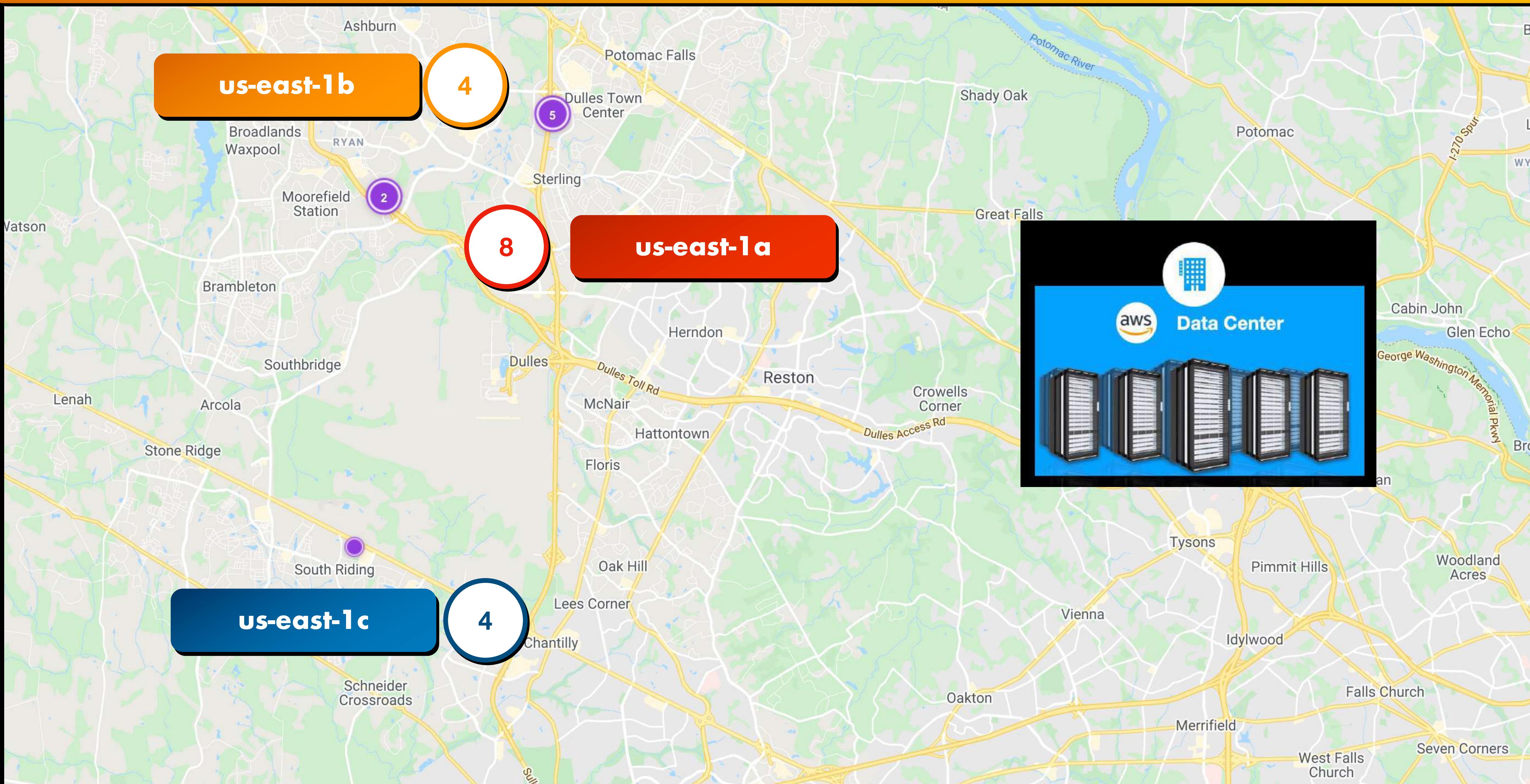
INTERNET GATEWAY



CLOUD







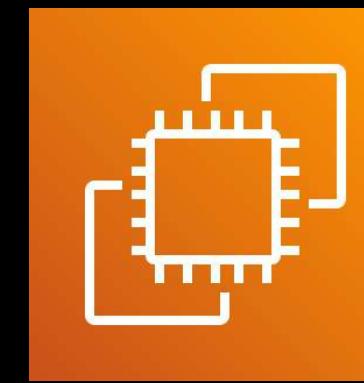


CLOUD

Fully Managed By:



Amazon VPC

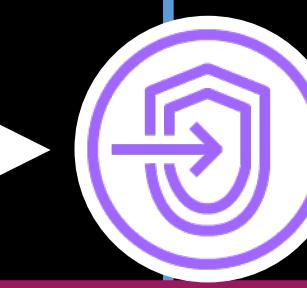
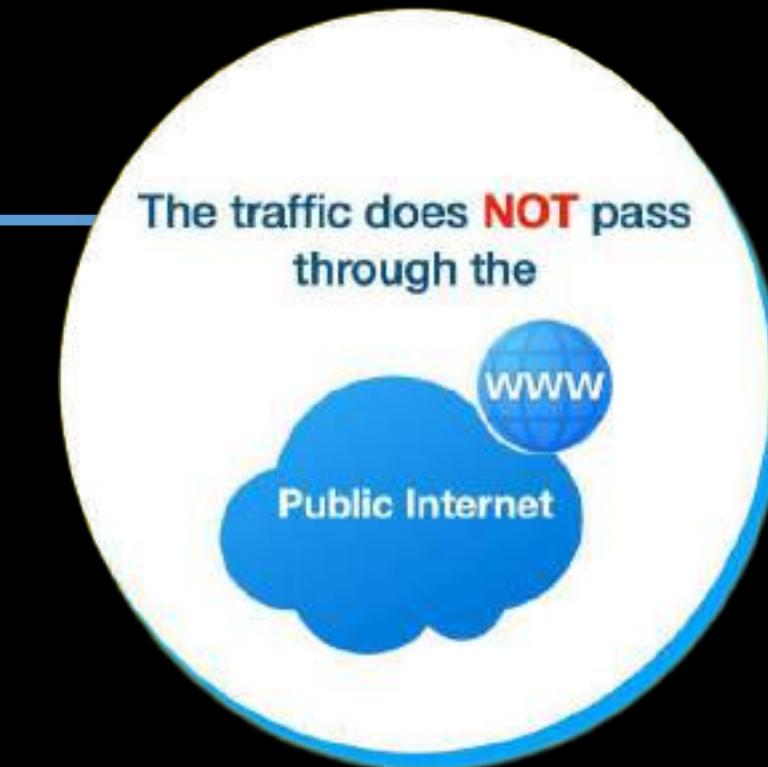


Amazon EC2



Amazon S3

Amazon S3 is not hosted in an Amazon VPC



VPC Endpoint



AWS Lambda



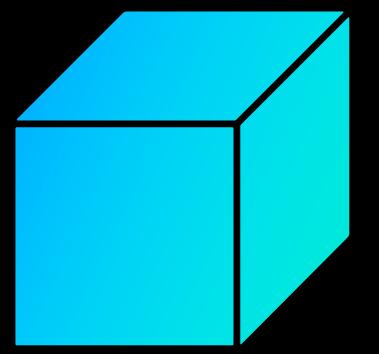
Amazon  
DynamoDB

Other Services



## Amazon VPC Components

- **CIDR Block**
- **Subnets**
- **Route Table**
- **DHCP Options Set**
- **NAT Devices**
- **Network ACLs**
- **Security Groups**
- **Different types of Gateways**

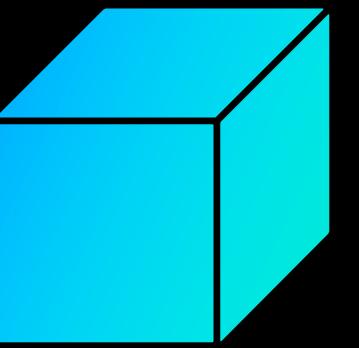


## CIDR BLOCK

- Allows you to specify the size of your network
- The allowed block size for a VPC is between /16 to /28 netmask
- A netmask (subnet mask) tells you the total number of available hosts for your network

/16	= 65,536 IP addresses
/17	= 32,768 IP addresses
/18	= 16,384 IP addresses
/28	= 16 IP addresses

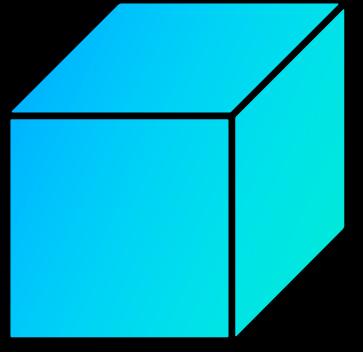
- AWS reserves a total of **5 IP addresses** from your CIDR block
- The **first four IP addresses** and the **last IP address** in each subnet CIDR block are reserved



## CIDR BLOCK

CIDR 10.0.0.0/24

10.0.0.0	- Network Address
10.0.0.1	- VPC Router
10.0.0.2	- DNS Server
10.0.0.3	- Reserved for Future Use
10.0.0.255	- Network Broadcast Address



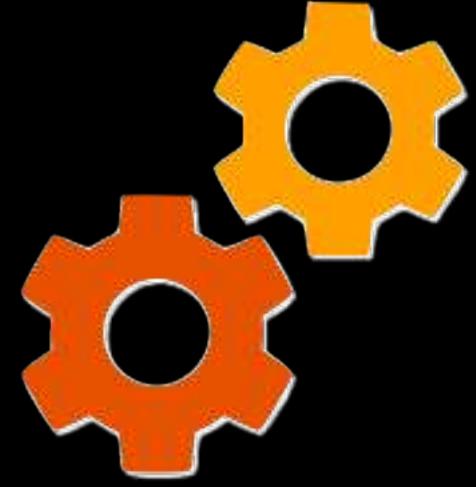
**CIDR BLOCK**

**IPv4 CIDR Range:** 10.0.0.0/16

**IPv6 CIDR Range:** 2001:db8:1234:1a00::/56

- The **implicit router** in Amazon VPC
- **Controls the network traffic** in your VPC through subnet routing
- All subnets in your VPC must be associated with a route table.
- A route table can either be the **main route table** or a **custom route table**
- A subnet in your VPC can only be associated with one route table at a time but you can associate multiple subnets with the same subnet route table.





## DHCP OPTIONS SET

- A set of options that **controls the automatic provisioning of IP addresses to your Amazon EC2 instances and other resources**
- **Uses the Dynamic Host Configuration Protocol**
- **Allocates an IP address to every host, virtual machine, EC2 instance, RDS database, load balancer, or any other AWS resources in your VPC**
- **Configures your DNS, NetBios Name Server, and Network Time Protocol (NTP)**

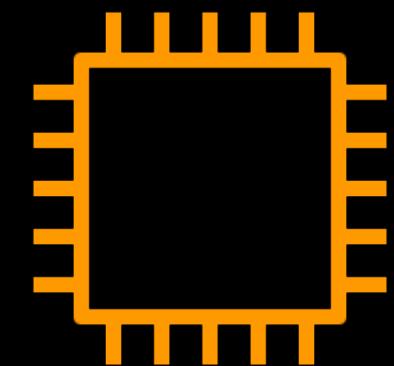


## NAT DEVICES

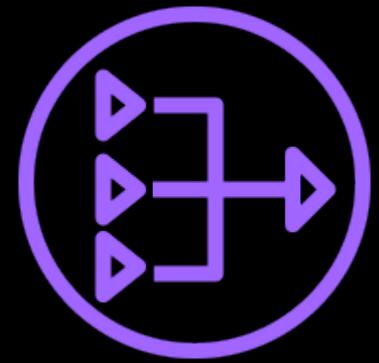
- **Uses Network Address Translation (NAT)**
- **Enable Amazon EC2 instances that are in a private subnet to connect to the public Internet or other AWS services**
- **Prevents the public Internet from initiating connections with your private EC2 instances.**
- **Works like a one-way street which means only the traffic initiated within your VPC is allowed but not vice versa**



## NAT DEVICES

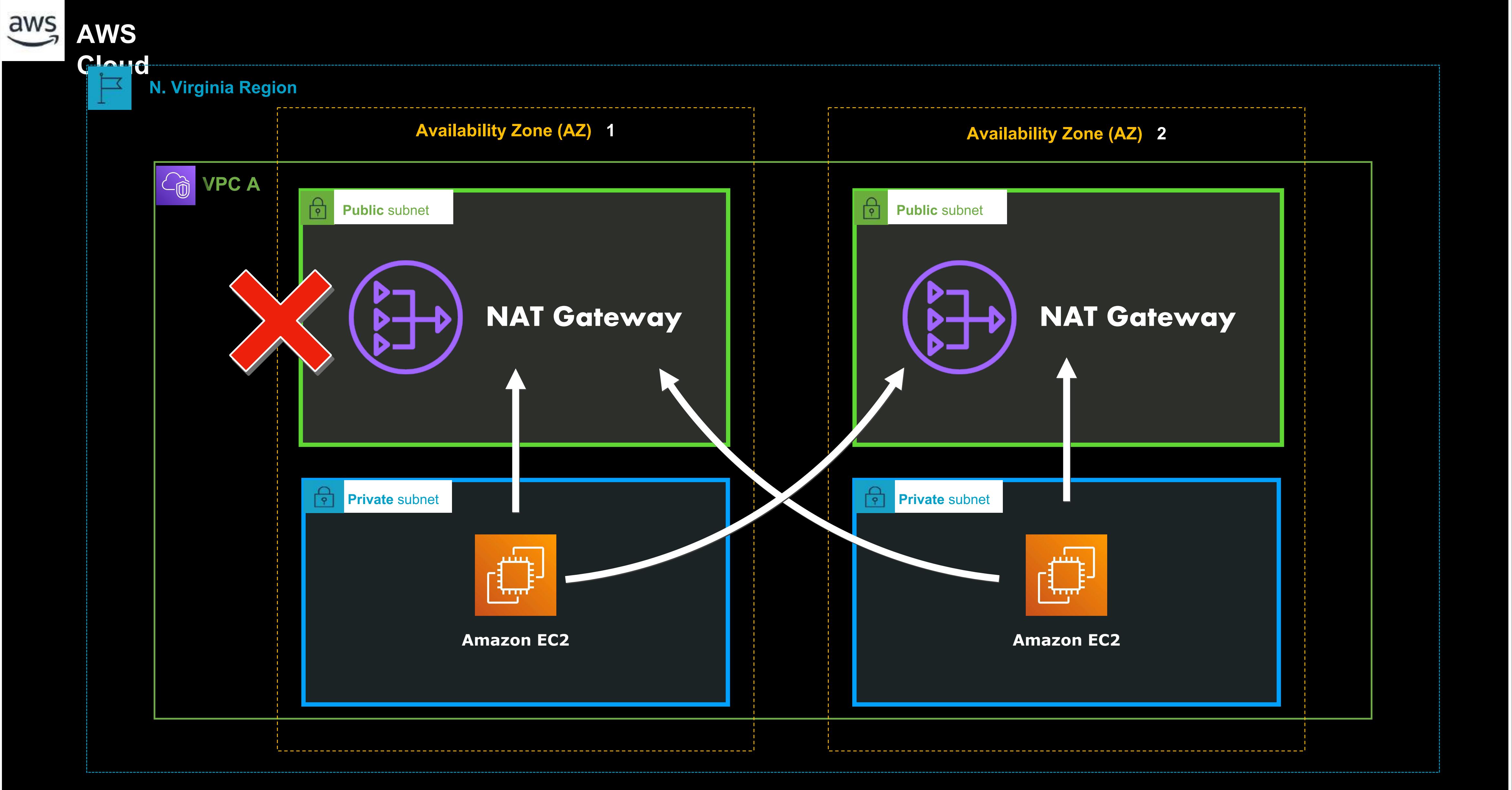


**NAT Instance**



**NAT Gateway**

- A **virtualized NAT device running in an EC2 instance within your VPC**
- **Managed by the customer (you)**
- **Not highly available nor scalable**
- An **advanced NAT device that is not running in your VPC**
- **Managed by AWS**
- **Highly available and scalable**





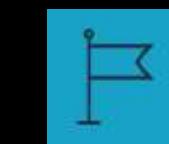
**Security Groups**



**Network Access Control List  
(Network ACL)**



## AWS Cloud



N. Virginia Region



Network ACL

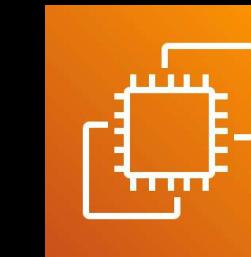
SUBNET



VPC A



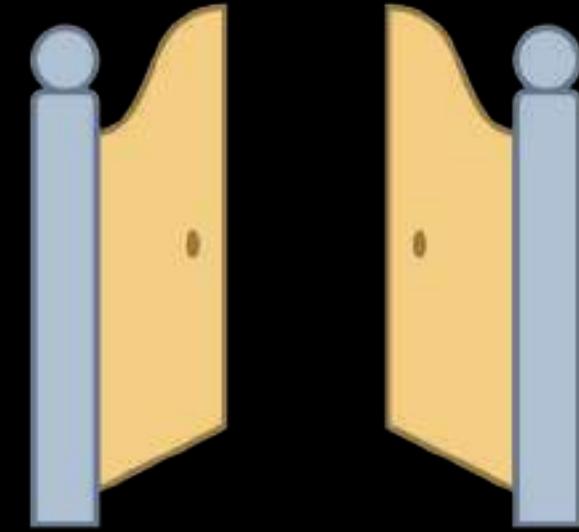
Security Group



Amazon EC2

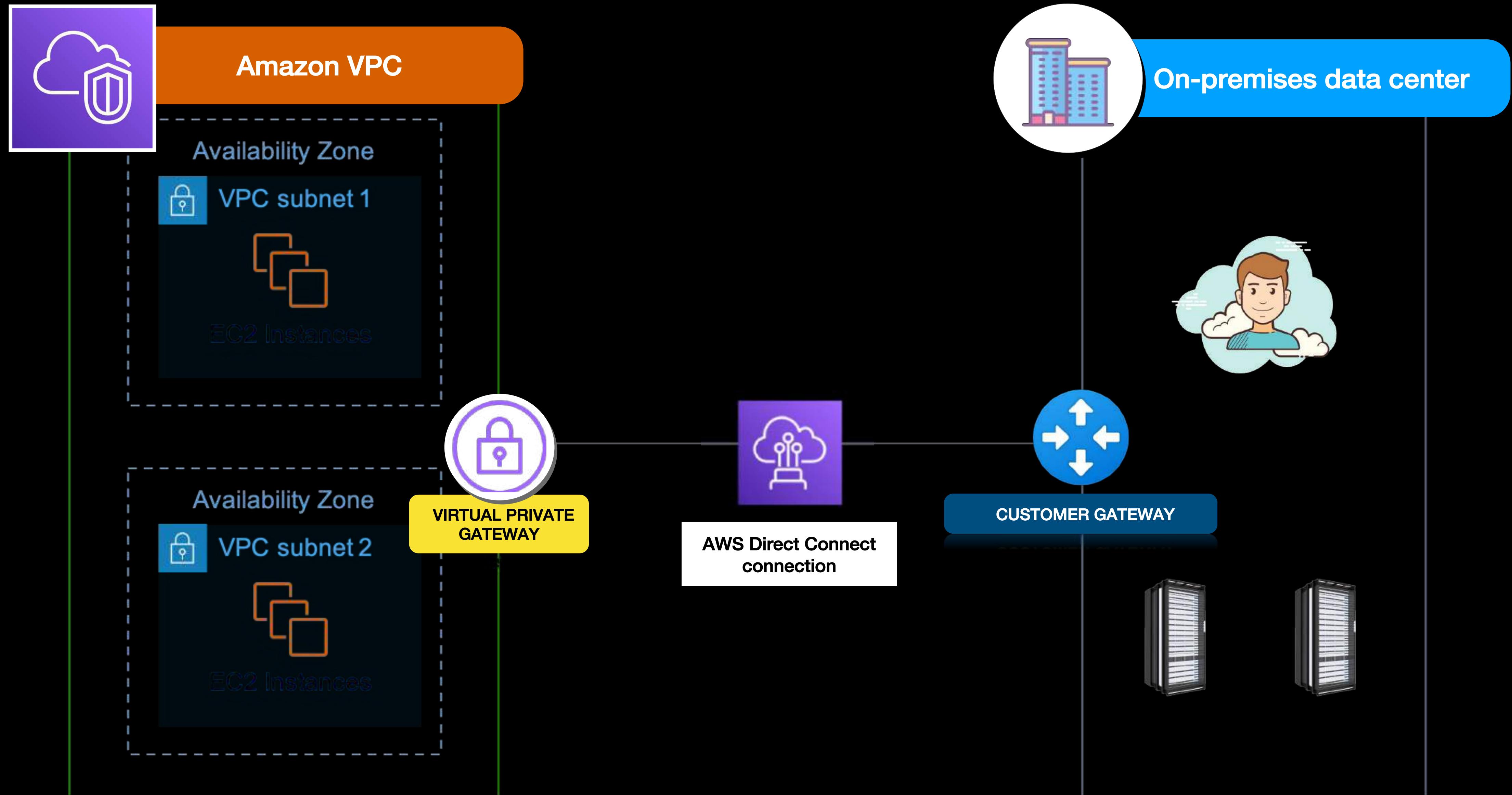
You can **create a rule that explicitly allows or denies traffic by its IP address, port, or destination**

You can only specify **ALLOW** rules in a Security group, but not **DENY** rules



## Gateways

- **Internet Gateway** 
- **Customer Gateway**
- **Virtual Private Gateway**
- **Carrier Gateways**
- **Egress-only Internet Gateway**

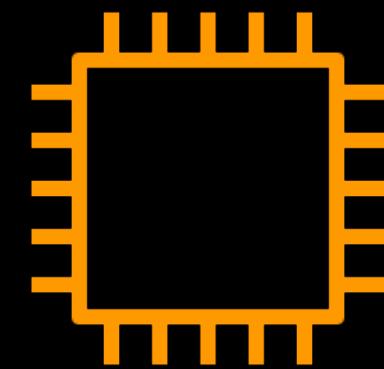


**IPv6**

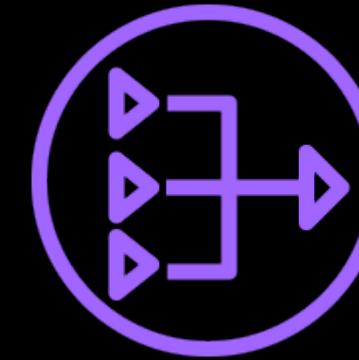


**Egress-only Gateway**

**IPv4**



**NAT Instance**

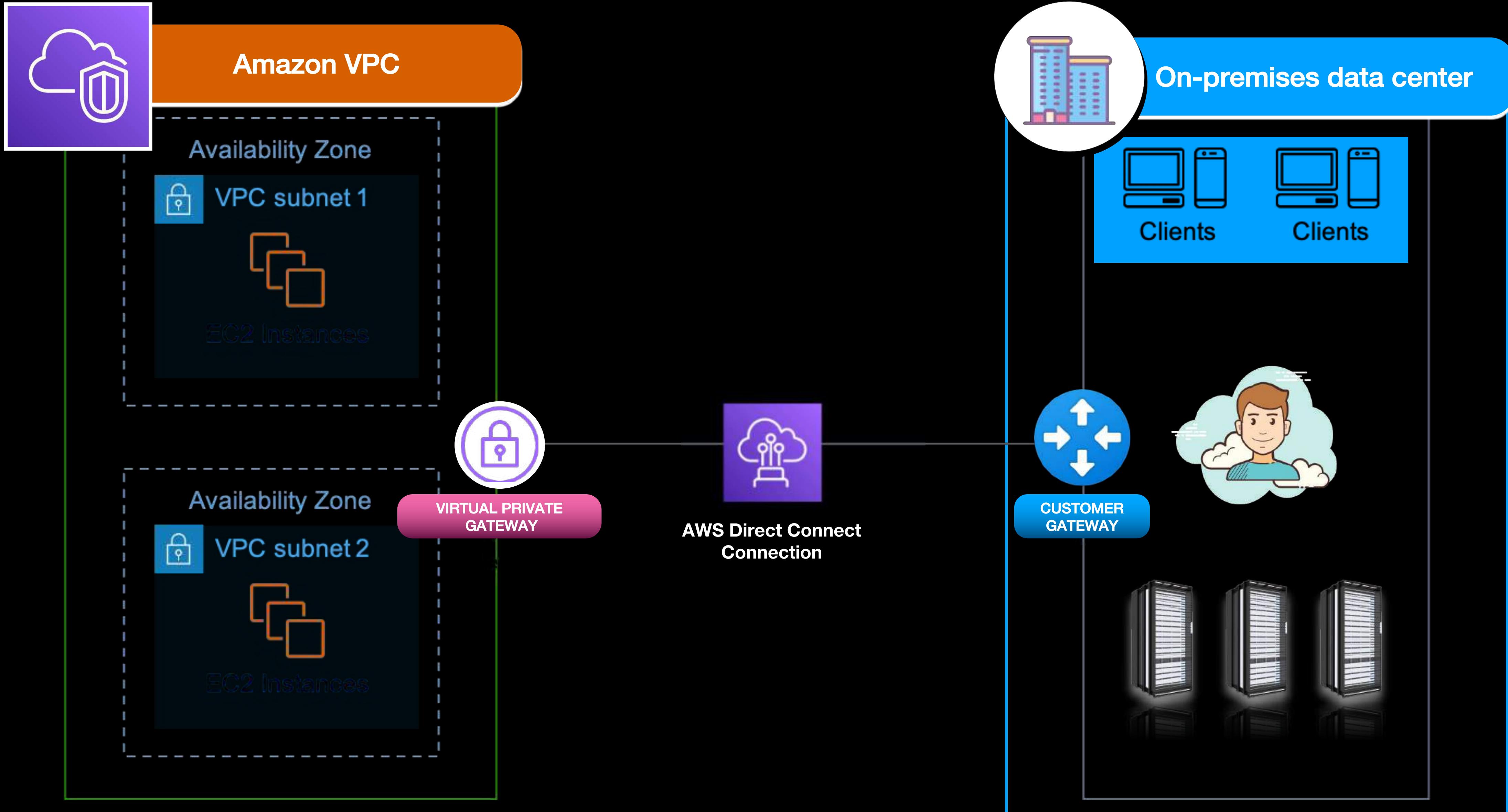


**NAT Gateway**



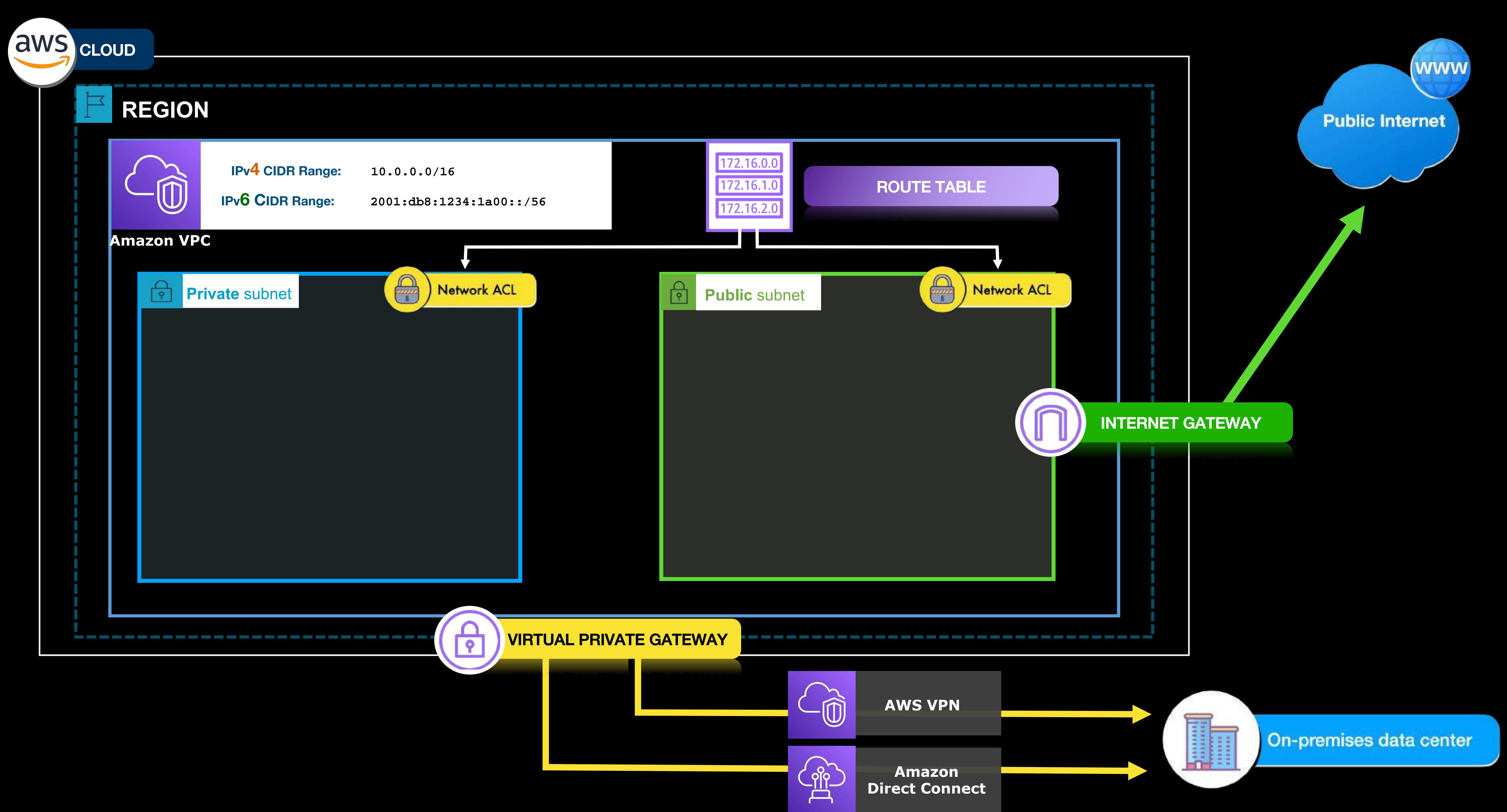
## CARRIER GATEWAY

- For VPCs that use **AWS Wavelength** to deliver ultra-low latency applications for 5G devices.
- Allows incoming traffic from a carrier network in a specific location
- Allows outgoing traffic to the carrier network and to the public Internet.
- Only available for VPCs that contain subnets in a **Wavelength Zone**





# Amazon VPC Network Architectures





## Default VPC

- There is a default VPC in each AWS Region
- A default VPC can immediately be used to launch your Amazon EC2 instances, Elastic Load Balancers, Amazon RDS databases, and other resources.
- Perfect for quickly launching simple public websites or applications
- The existing components of your default VPC can be configured
- Has an attached Internet Gateway by default



Default VPC

IPv4 CIDR Range: 172.31.0.0/16

/16

= 65,536 IP addresses

- 172.31.0.0 – Network Address
- 172.31.0.1 – VPC Router
- 172.31.0.2 – DNS Server
- 172.31.0.3 – Reserved for Future Use
- 172.31.255.255 – Network Broadcast Address

The first 4 IP addresses and the last IP address of that range are reserved.

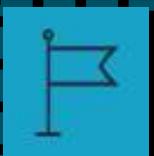
You have a total of 5 IP addresses that are not usable



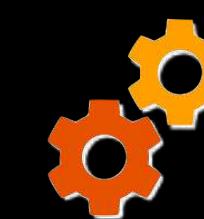
CLOUD

/20

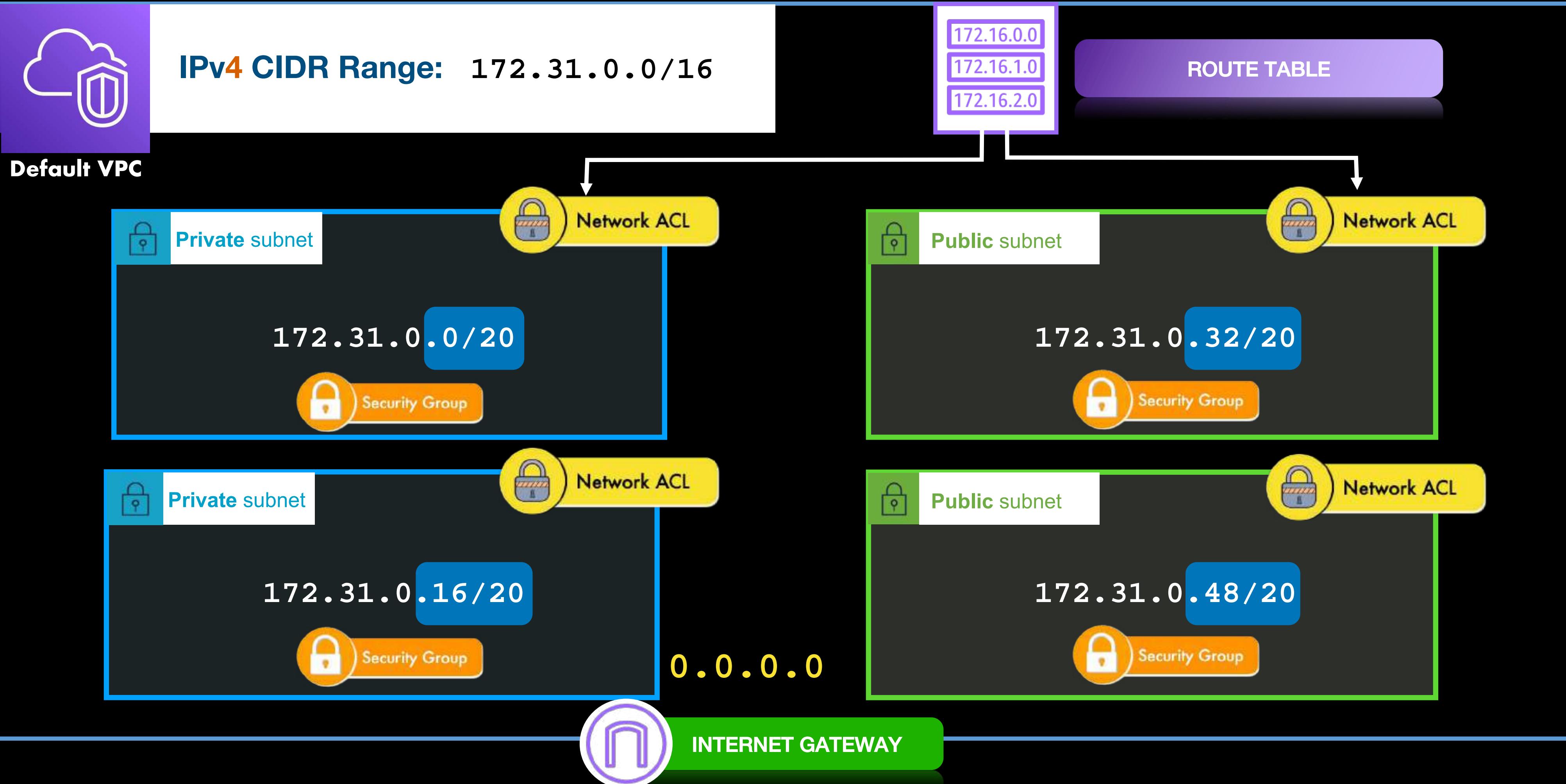
= 4,096 Total IP addresses - Reserved AWS IPs ~ 4,090 Usable IPs



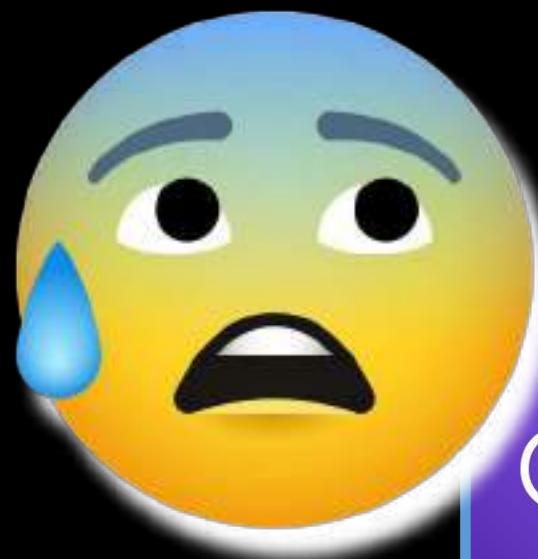
REGION



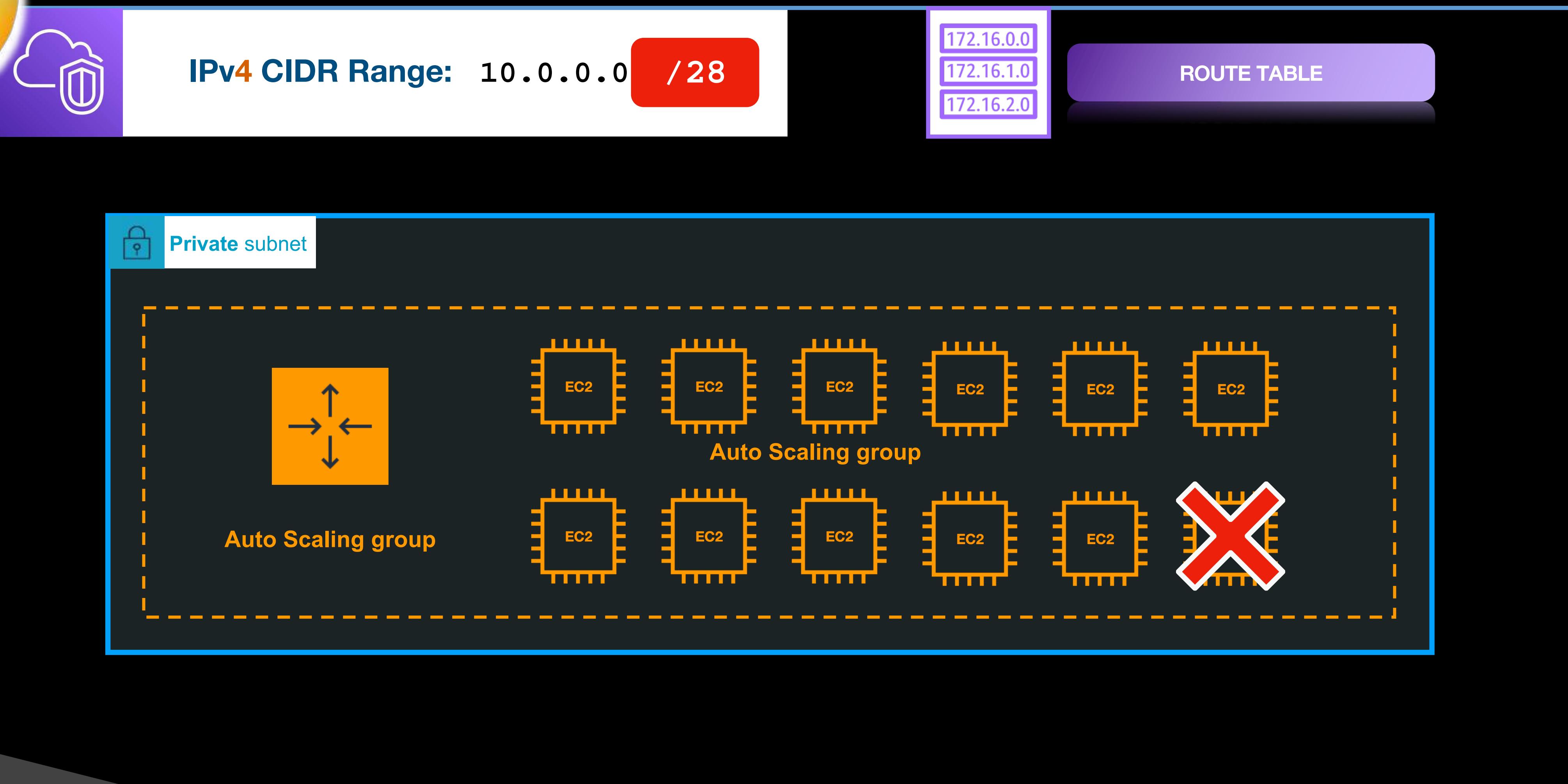
## DHCP OPTIONS SET



# CUSTOM AMAZON VPC



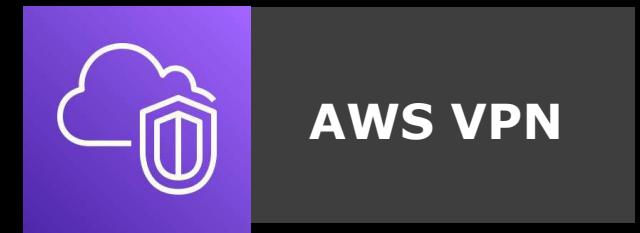
**/28 = 16 Total IP addresses - 5 Reserved AWS IPs = 11 Usable IPs**



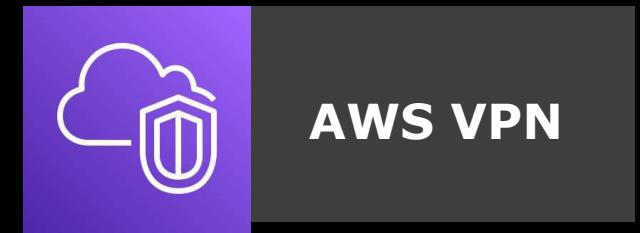
## Amazon VPC Network Architecture Types



- A VPC with a single **public** subnet
- A VPC with **public** and **private** subnets
- A VPC with **public** and **private** subnets and **Hardware VPN Access**
- A VPC with a **private** subnet only and **Hardware VPN Access**

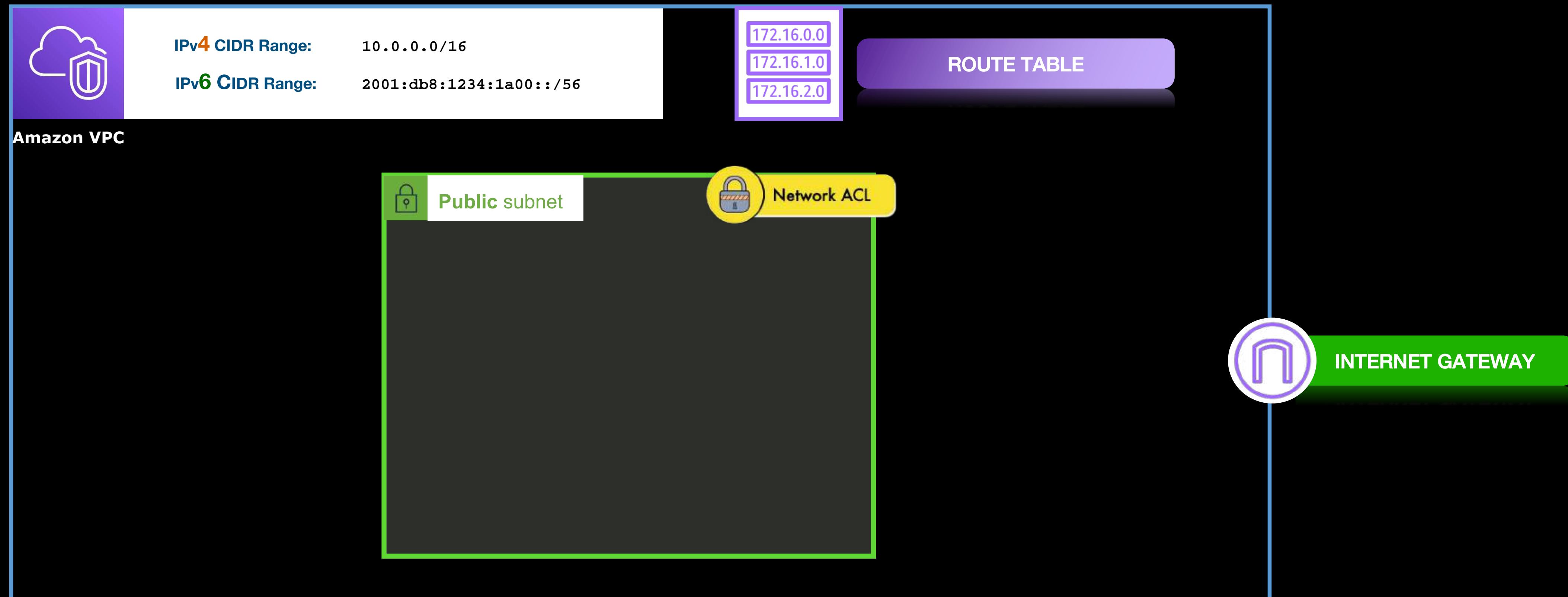


AWS VPN

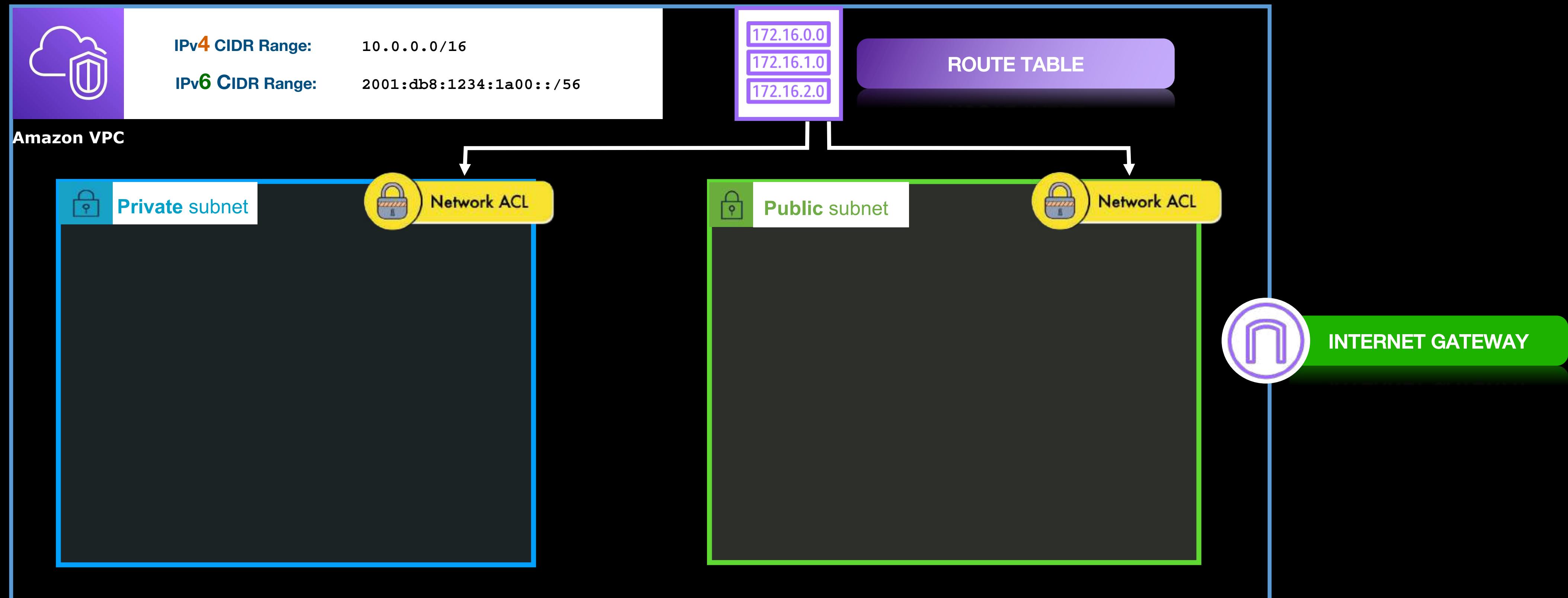


AWS VPN

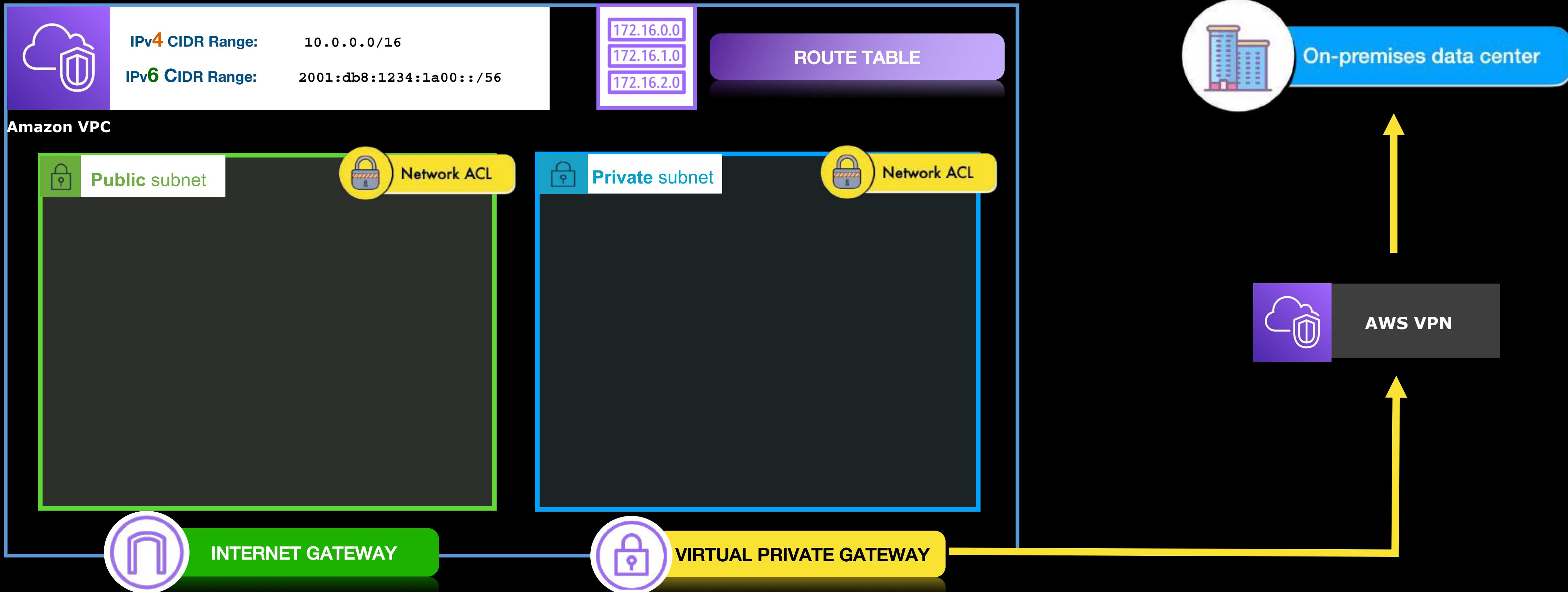
# A VPC with a single **public** subnet



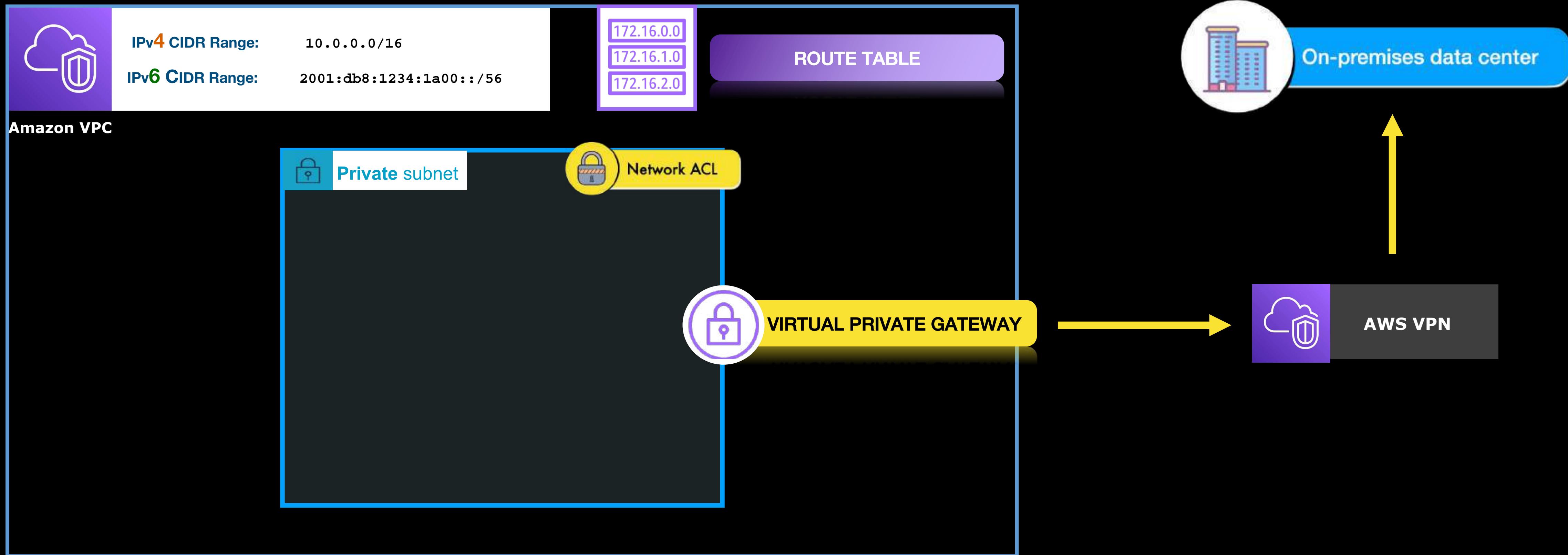
# A VPC with **public** and **private** subnets



# A VPC with **public** and **private** subnets and **Hardware VPN Access**

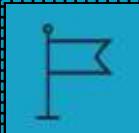


# A VPC with private subnet and Hardware VPN Access





# AWS Cloud

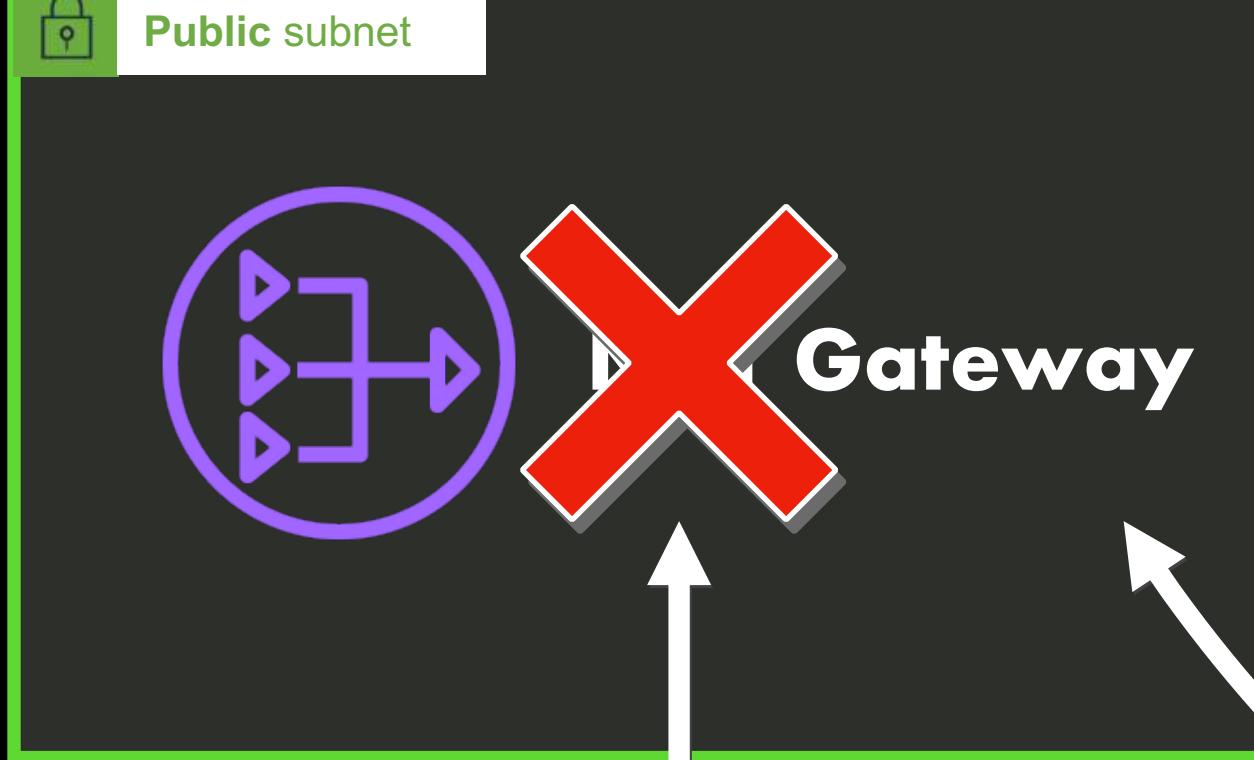


N. Virginia Region

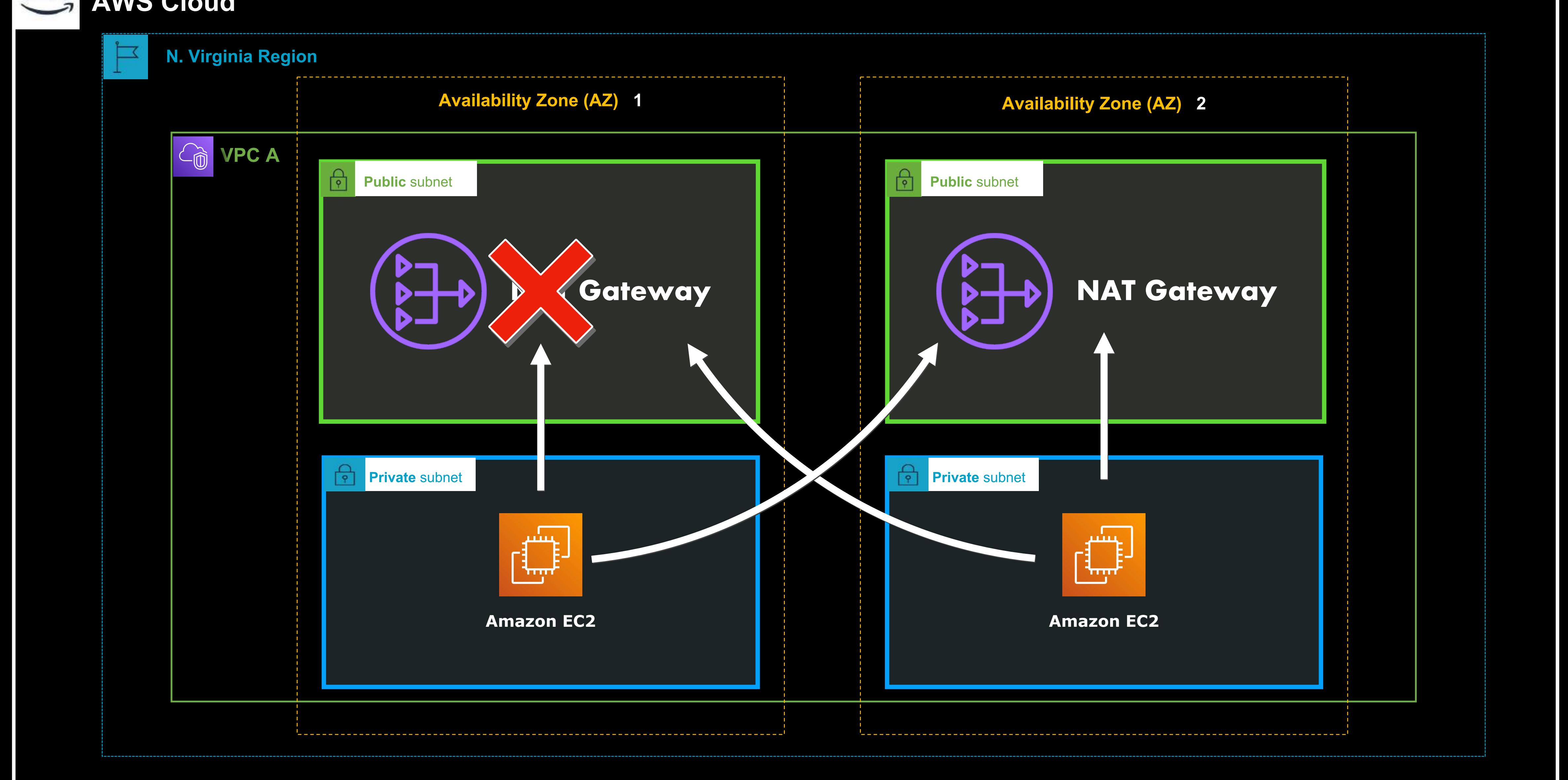
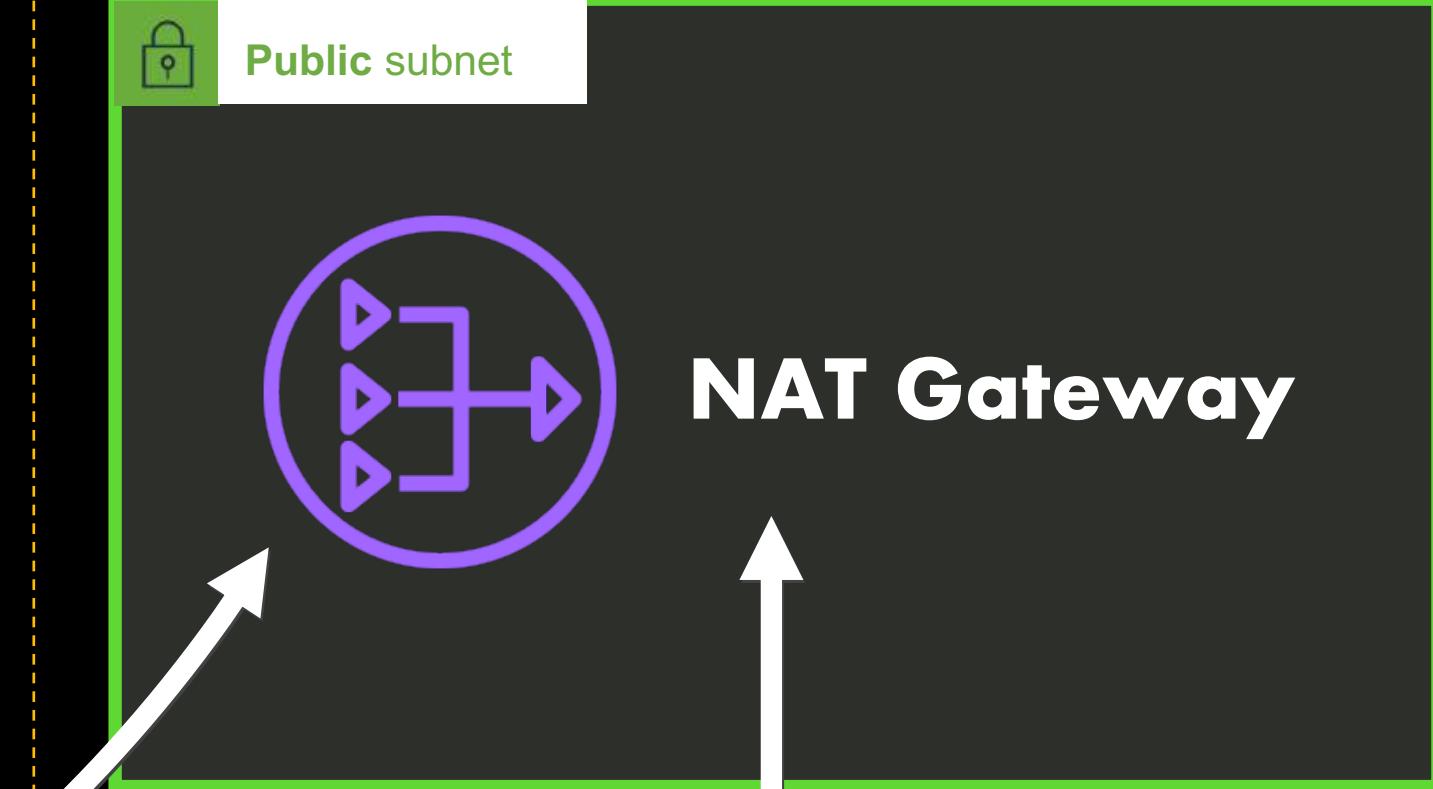
Availability Zone (AZ) 1

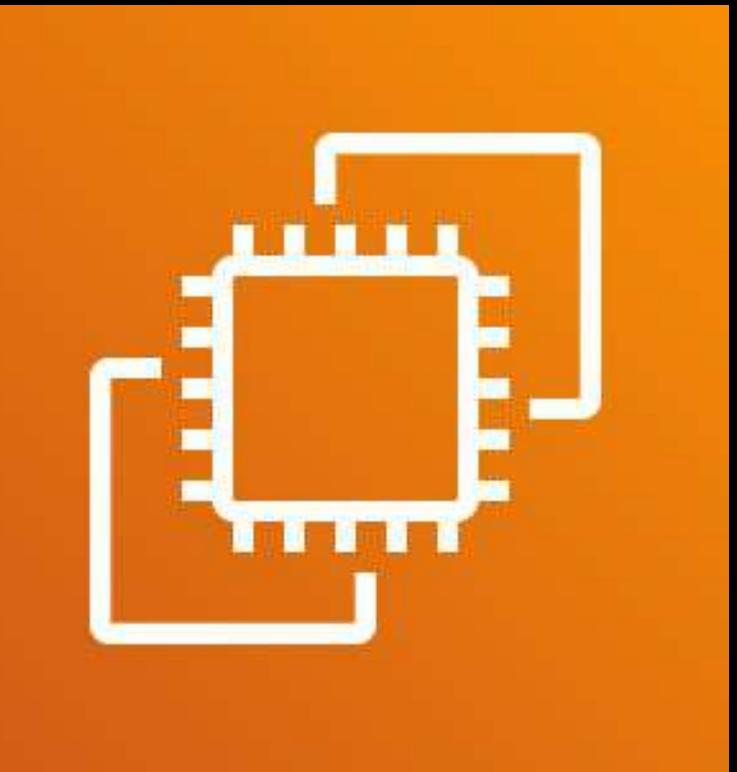


VPC A



Availability Zone (AZ) 2





# Amazon EC2 Overview

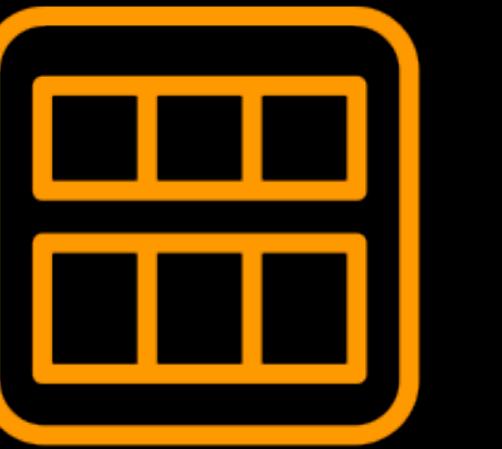
---



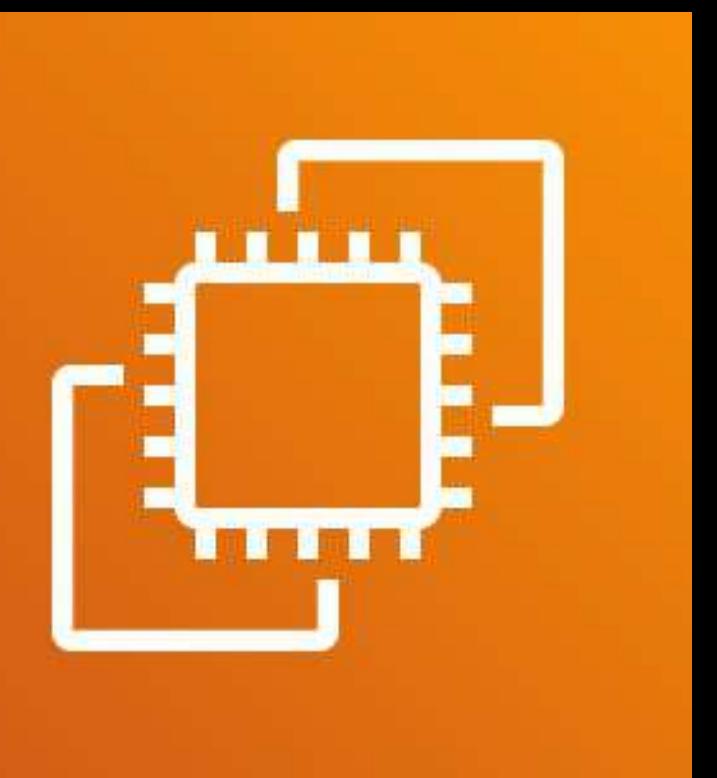


**Shared Responsibility**

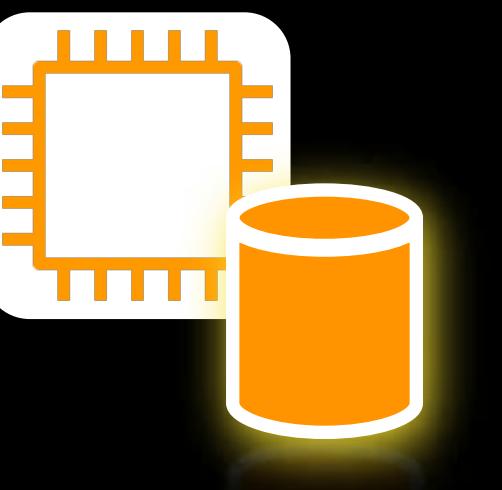
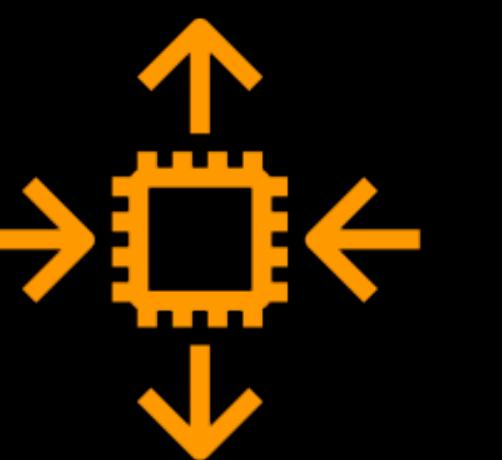




Can be **integrated** with  
a lot of AWS Services



Amazon EC2





Amazon VPC



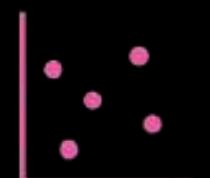
Elastic IP Address



Elastic Network Interface (ENI)



Elastic Network Adapter (ENA)



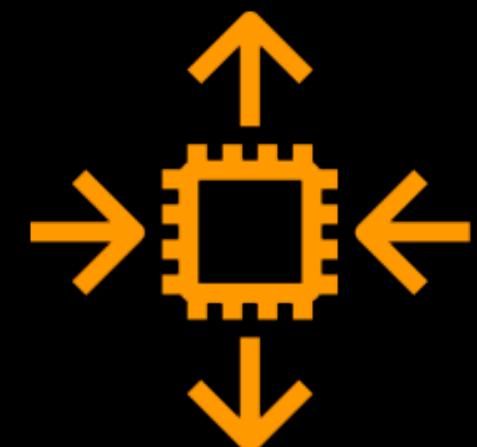
Placement Groups



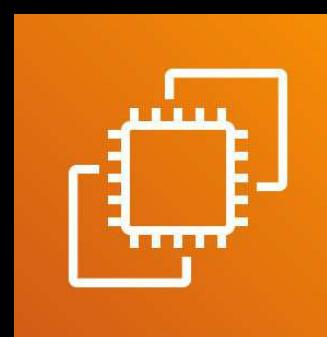
Elastic Fabric Adapter (EFA)



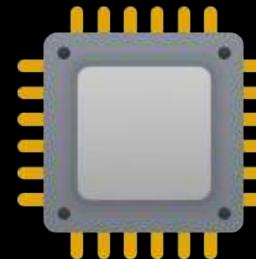
Your Computer



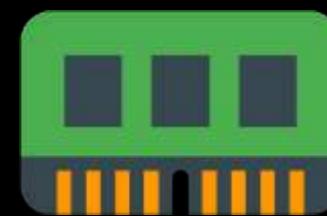
both have



Amazon EC2



CPU



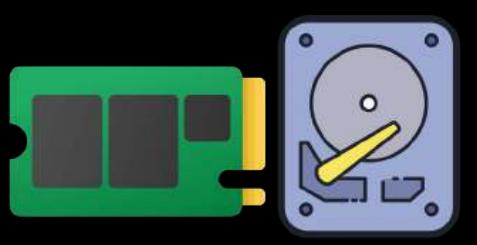
MEMORY (RAM)



NETWORK



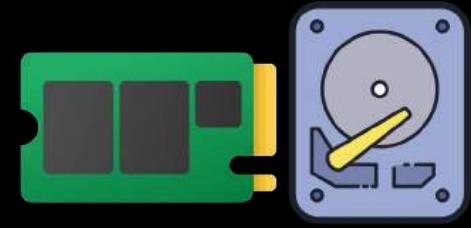
DISK IMAGE (ISO)



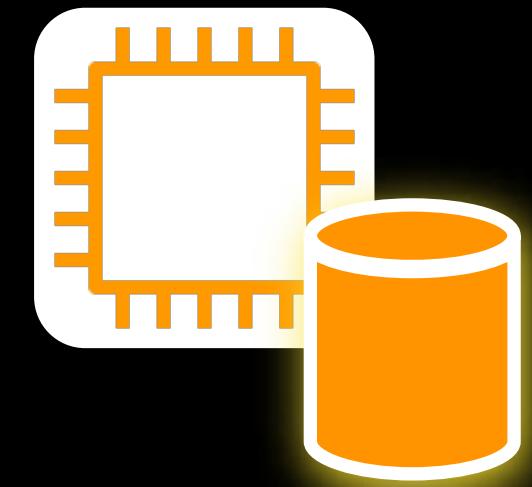
SSD/HDD STORAGE



SHARED FILE SERVER



## SSD/HDD STORAGE



Instance Store



Amazon EBS



## SHARED FILE SERVER



Amazon EFS



Amazon FSx for Lustre



Amazon FSx for Windows  
File Server



## OBJECT STORAGE



Amazon S3



## NETWORK



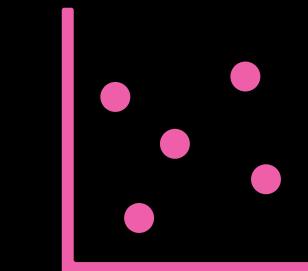
Amazon VPC



Elastic IP Address



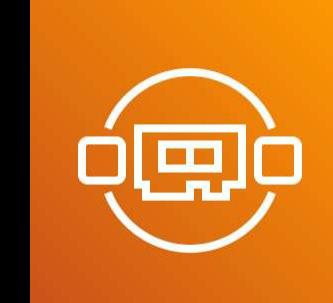
Elastic Network Interface (ENI)



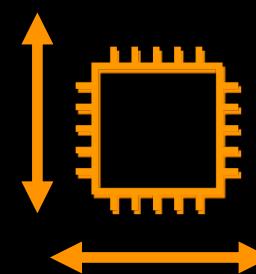
Placement Groups



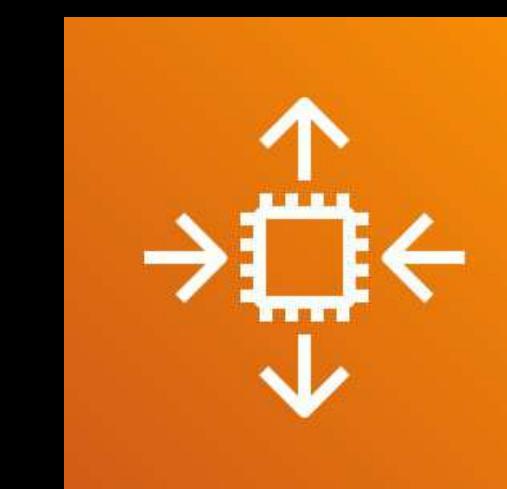
Elastic Network Adapter (ENA)



Elastic Fabric Adapter (EFA)



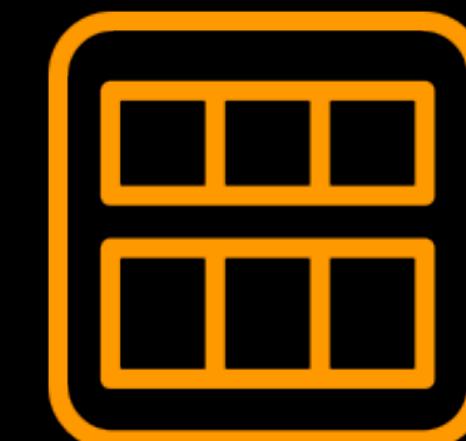
## AUTO SCALING



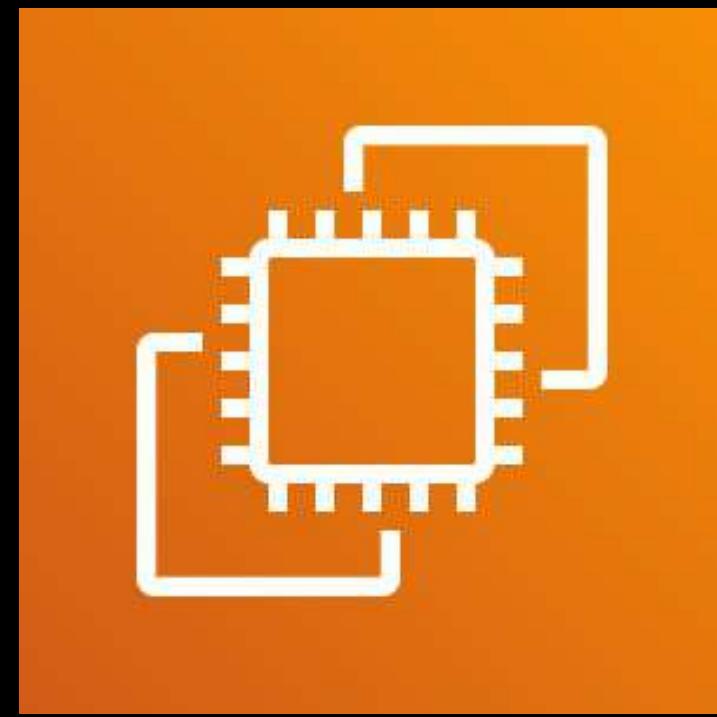
Amazon EC2 Auto Scaling



## DISK IMAGE



Amazon Machine Image (AMI)

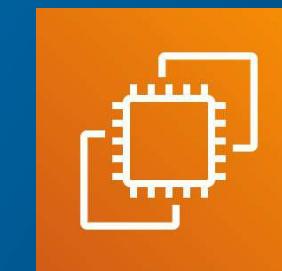


**Amazon EC2**

# Instance Purchasing Options

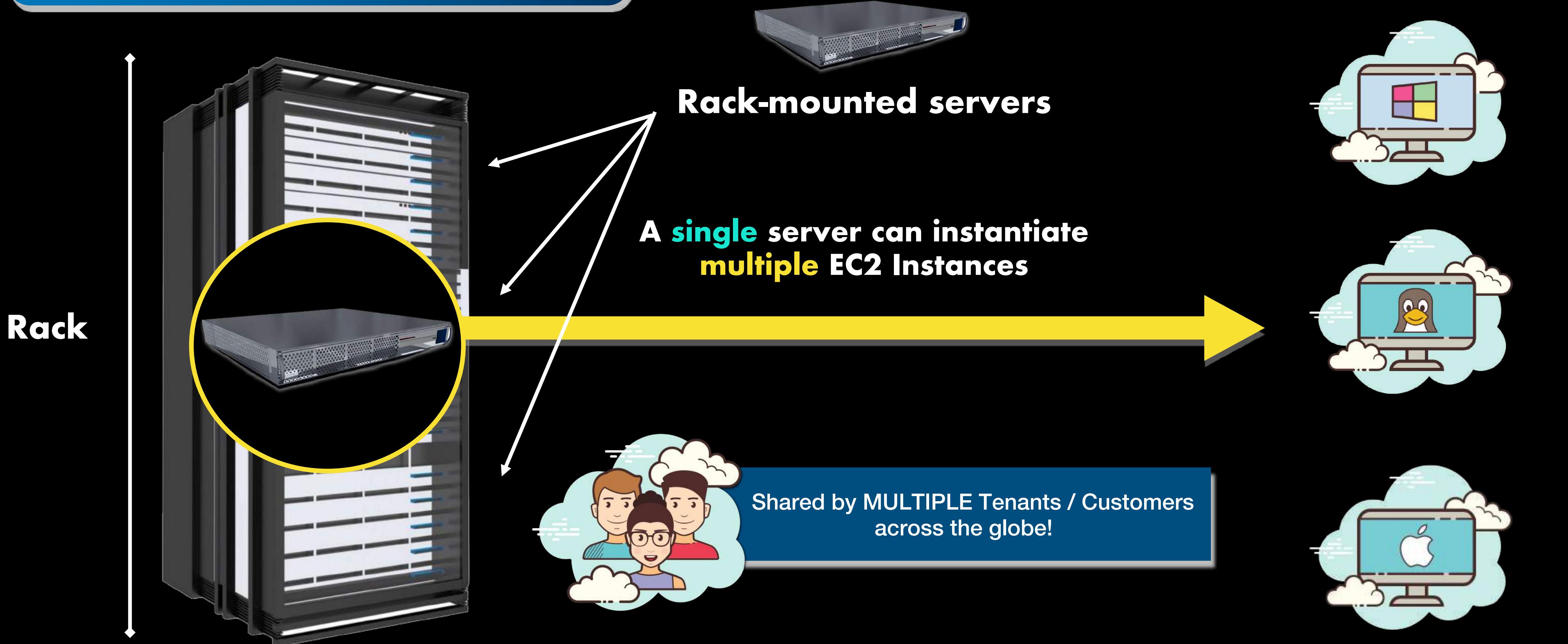
---

## Underlying Physical Servers of



Amazon EC2

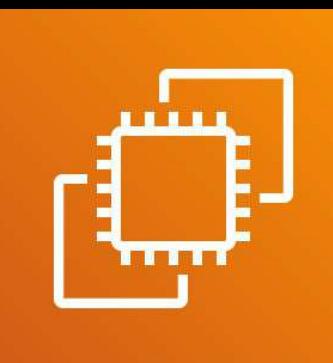
## Virtual Machines



**Spare or  
Unused Server**



## INTERRUPTION



**Amazon EC2  
Service**



## INTERRUPTS

(Automatically Terminates  
Your Spot EC2 Instance)

I would like to rent the entire server  
without any virtualization & is **dedicated**  
for my exclusive use!

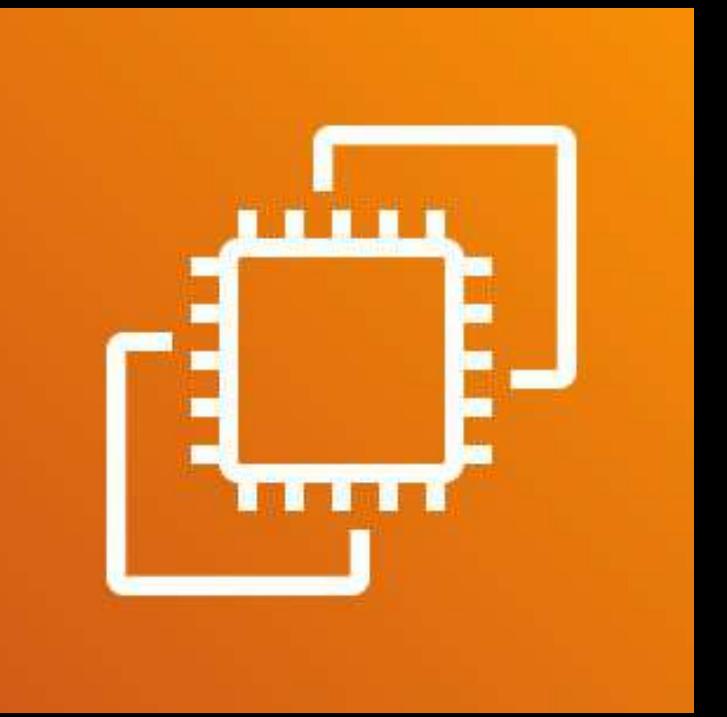
I'll pay that spare EC2  
Instance for  
**\$1** / hour



I would like to **reserve**  
this instance for  
1 year at  
**\$1.5** / hour

I want to order an  
EC2 Instance for  
**\$2** / hour





## **Amazon EC2 Instance Purchasing Options**

- **On-Demand**
- **Spot**
- **Reserved**
- **Dedicated**
- **Savings Plans**
- **Capacity Reservation**



# Spot Instances

---

**LOW** Supply  
=  
**HIGH** Price



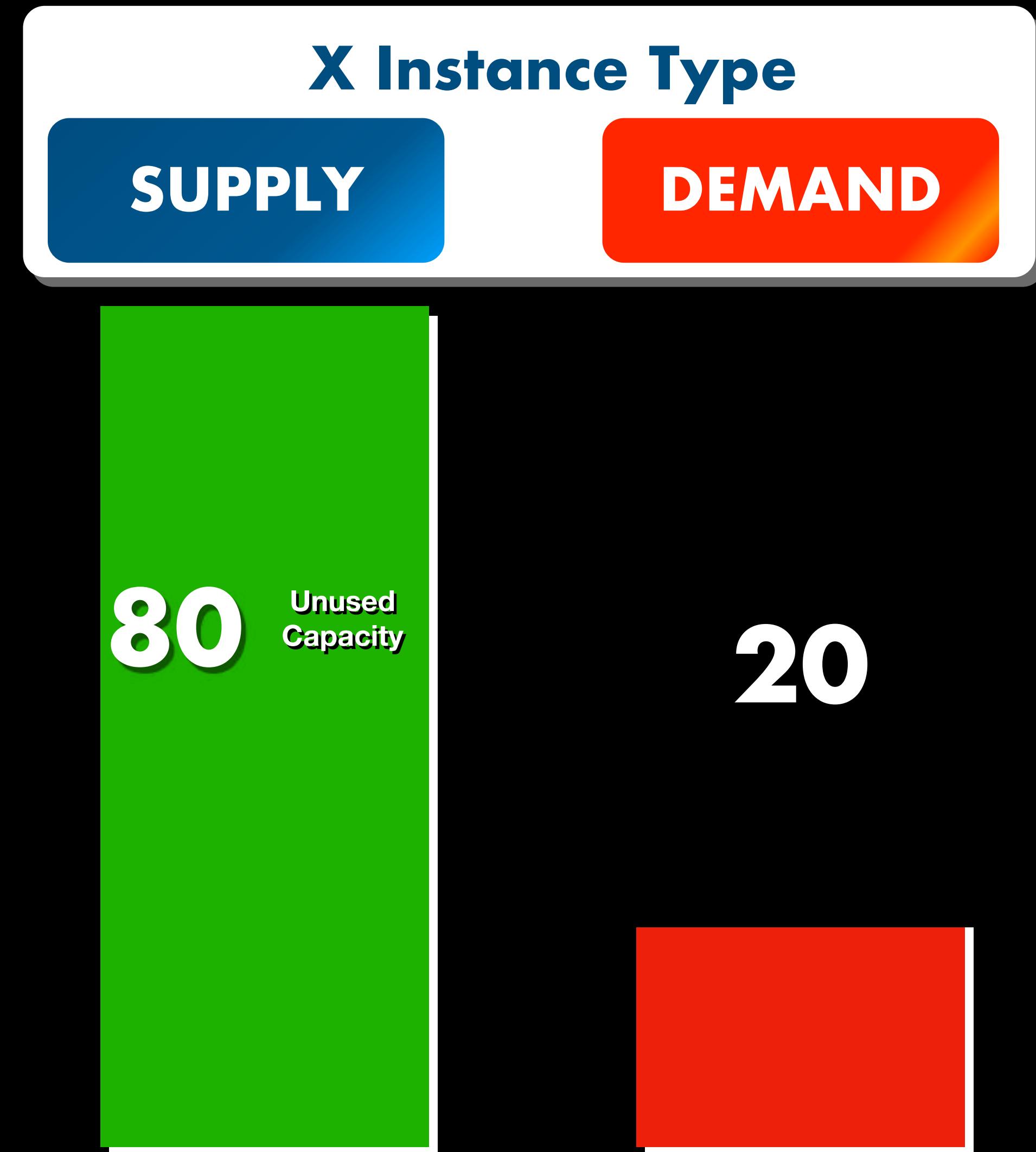
## Spot Instances

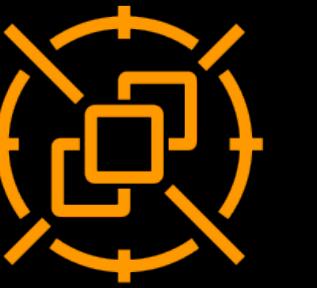
LOWEST COST

Spare or  
Unused EC2  
Capacity



SURPLUS





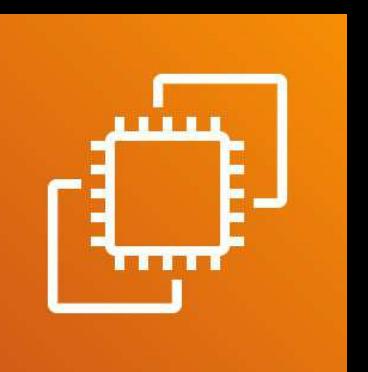
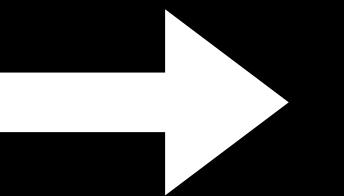
## Spot Instances

Spare or  
Unused Capacity

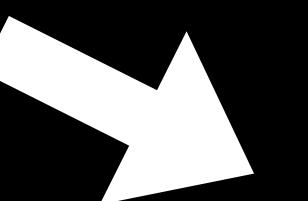


INTERRUPTS

(Automatically Terminates  
Your Spot EC2 Instance)



Amazon EC2  
Service



I want to order a Spot  
EC2 Instance for  
**\$1** / hour



I want to order an On-  
Demand  
EC2 Instance for  
**\$2** / hour



## Spot Instances

**Based on Spot Market**



**Spot Price**

**Buy “On the Spot”  
for lower prices**



## Spot Instances

### FEATURES

- Provide discounts of up to 90% compared to an On-Demand instance
- The most cost-effective type among the Instance purchasing options
- The interruption/termination is based on the Instance Type available in the AWS Global Infrastructure
- Can be interrupted, or be automatically terminated by AWS
- Suitable for non-critical and infrequent jobs that can be interrupted or processed again



## Spot Instances

### USE CASES

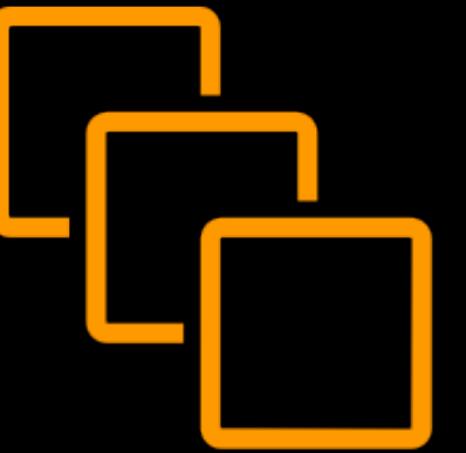
- Servers on your **development or test environments** that do not require to be 100% up all the time
- Applications with **flexible start and end times**
- Interruptible **workloads** that can handle failures gracefully
- Handling the **peak load or the additional load** of your application on top of your Reserved or On-Demand EC2 instances
- Infrequent and interruptible jobs
- Workloads that are **infrequently executed**



## Spot Instances

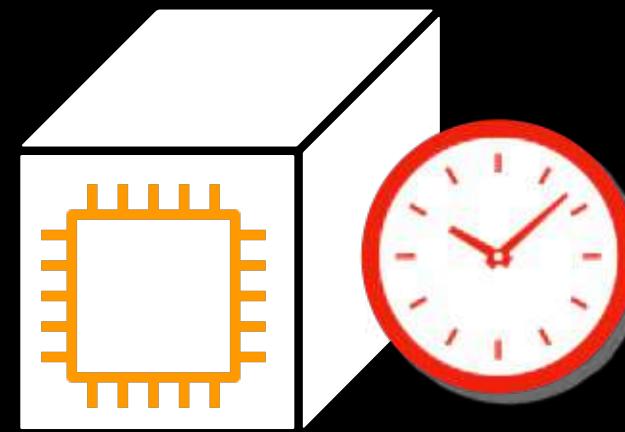
### USE CASES

- **Interruptible batch jobs or non-production applications that are currently hosted on your On-Demand Instances**
- **Running the task nodes of your Amazon Elastic MapReduce cluster**
- **Highly dynamic batch processing where each job:**
  - **Is stateless in nature**
  - **Can be started and stopped at any given time**
  - **Typically takes upwards of 60 minutes or an hour in total to complete**
- **For whenever you need the MOST cost-effective solution in running your interruptible workloads**



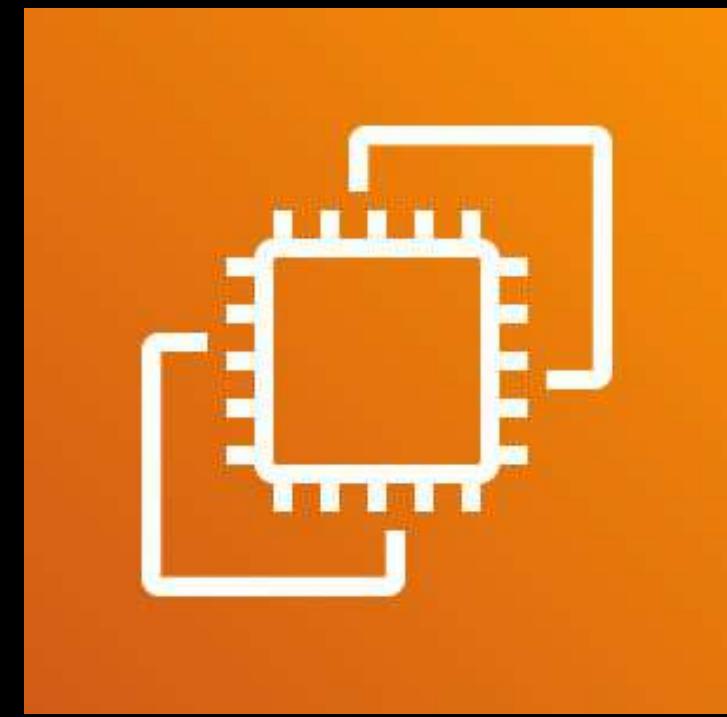
## Spot Fleet

- A **collection, or fleet, of Spot Instances**
- Can **optionally have On-Demand Instances**



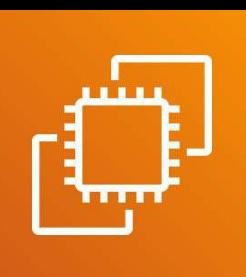
## Spot Block

- Specify a “**block of time**” or the duration in which your instance will run **continuously**
- **Rarely interrupted than your regular Spot instances.**



# On-Demand Instances

---



## On-Demand Instances



Demand #1



Right now, I want to launch an EC2 Instance for my app!



NO

**INTERRUPTIONS**

Demand #2

My batch job processing has been completed. I want to terminate my EC2 instance now

## On-Demand Instances USE CASES

- **Mission-critical workloads that must not experience any interruptions**
- **Servers of your mission-critical applications that are running on your production environment**
- **Short-term workloads that cannot be interrupted**
- **Handling the steady-state load of your applications**
- **Running the master node and the core nodes of your Amazon EMR cluster**
- **Any workloads that require uninterruptible processing**

## On-Demand Capacity Reservation

- Allows you to reserve EC2 capacity for a **specific Availability Zone for a period of time**
- Ensures that you always have access to EC2 capacity
- **No one-year or three-year term reservation or commitment**
- Suitable for scenarios where you require a guaranteed compute capacity for a week or a few months



OS Type



Linux

Pay  
by the **second**

**Minimum of**  
**1 minute**



Windows

Pay  
by the **hour**

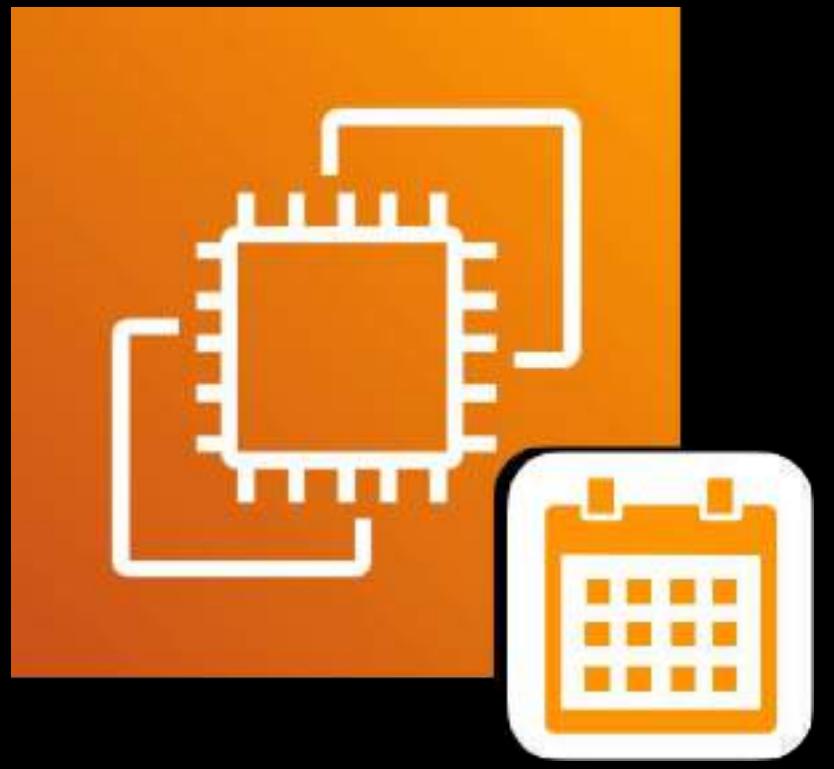
**Minimum of**  
**1 hour**



Has the **highest** cost among the other EC2 Instance Purchasing Options

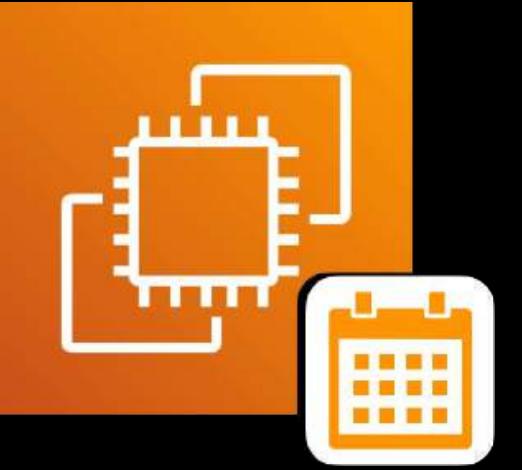


The high price you pay ensures that your EC2 Instance will NOT be interrupted

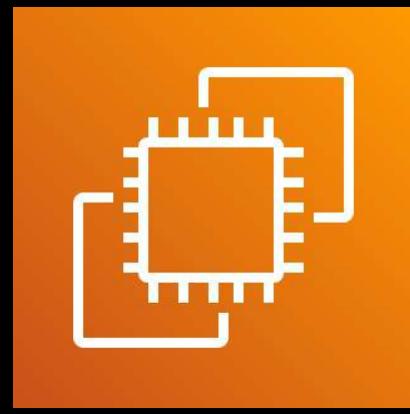


# Reserved Instances

---



## Reserved Instances



## On-Demand Instances

**FOR MISSION-CRITICAL  
APPLICATIONS**

**UNINTERRUPTIBLE**

**CHEAPER THAN  
ON-DEMAND INSTANCES**



## Spot Instances



## Reserved Instance Marketplace



# RESERVE

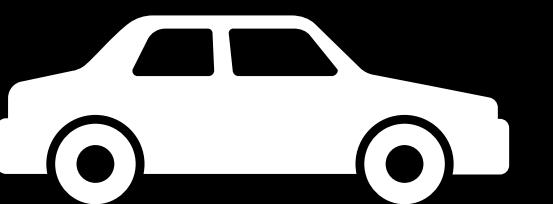


1 year  
3 years

All Upfront

Partial Upfront

No Upfront



All Upfront



Partial Upfront

No Upfront

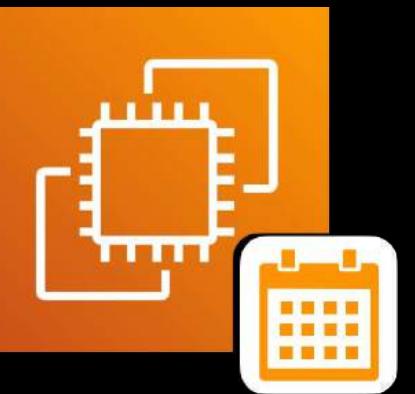
**Pay the FULL Price**

**Pay the PARTIAL Price**

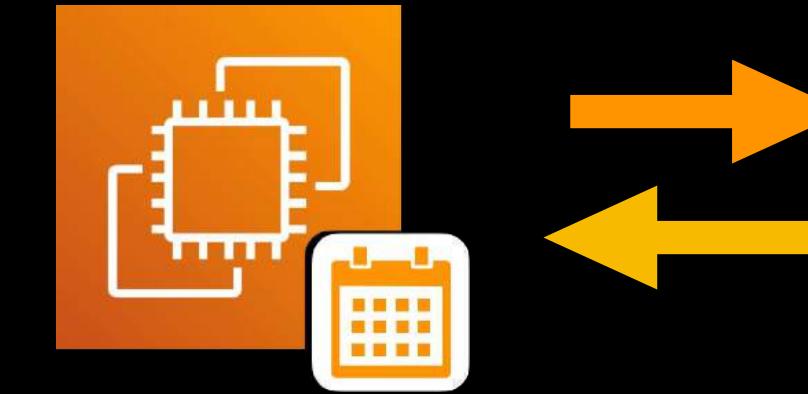
**Pay on a MONTHLY basis**

Provides the  
**highest savings!**

Provides the  
**least amount of discount**



## Standard Reserved Instance



## Convertible Reserved Instance

Both can modify the attributes such as the Availability Zone or Network

Both can modify the Instance Size using other sizes within the same instance family

Both require a fixed 1-year or 3-year commitment

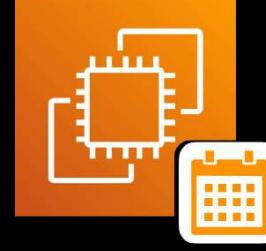
Can be sold  
in the Reserved Instance Marketplace



Cannot be sold  
in the Reserved Instance Marketplace

Cannot be exchanged for any other Reserved  
Instance

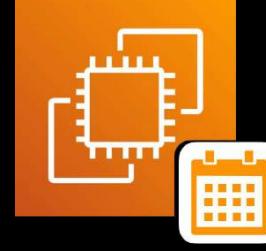
Can be exchanged for another Convertible Reserved  
Instance with a different configuration, including instance  
family, operating system, and tenancy



## Reserved Instances

### USE CASES

- Running **non-interruptible workloads** for a **one-year or three-year time frame**
- **Workloads with predictable capacity** and **uptime requirements**
- **Hosting the application servers** of your **production environment**
- For processing the **steady-state load** or the **baseline capacity** of your workloads



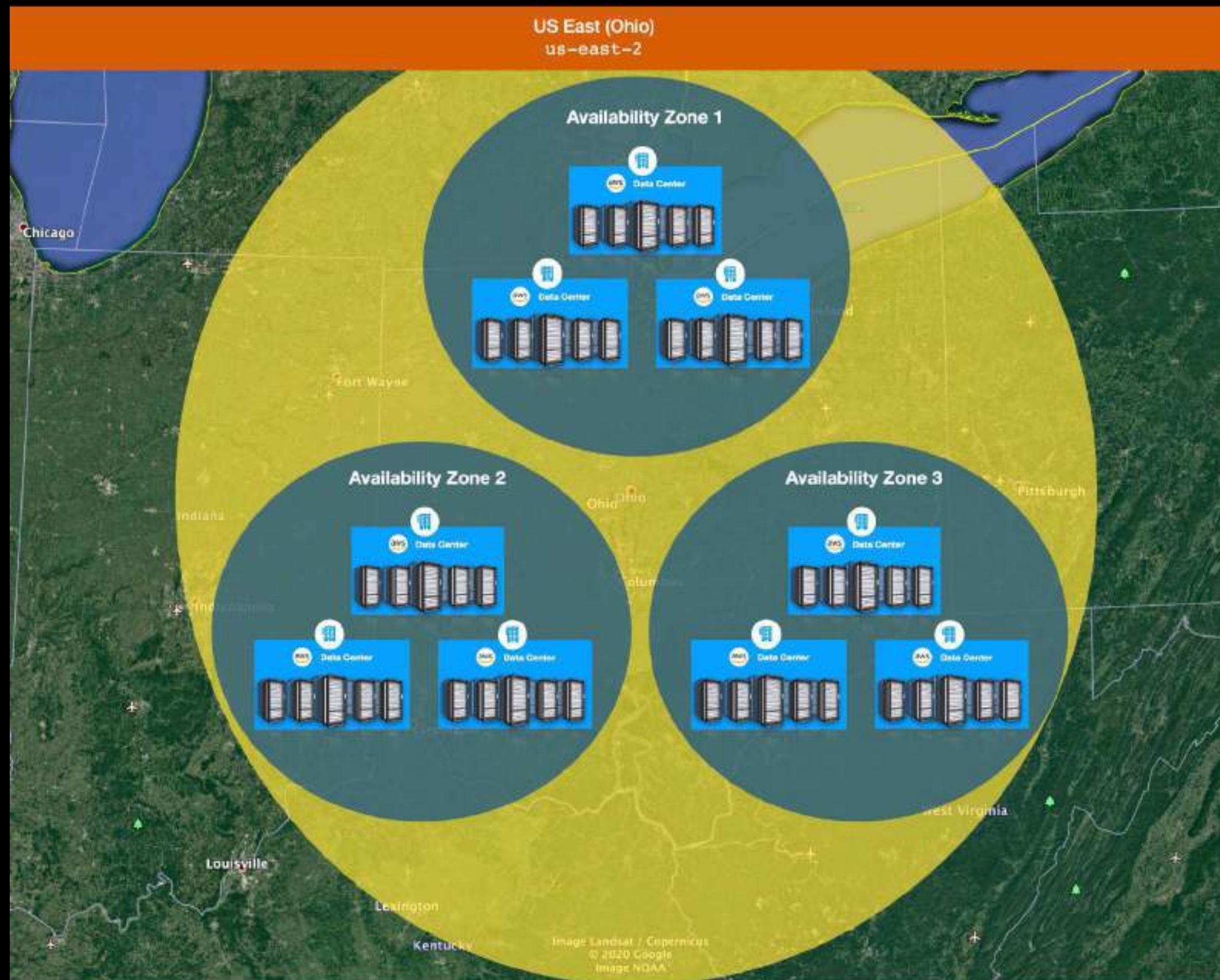
## Reserved Instances

### USE CASES

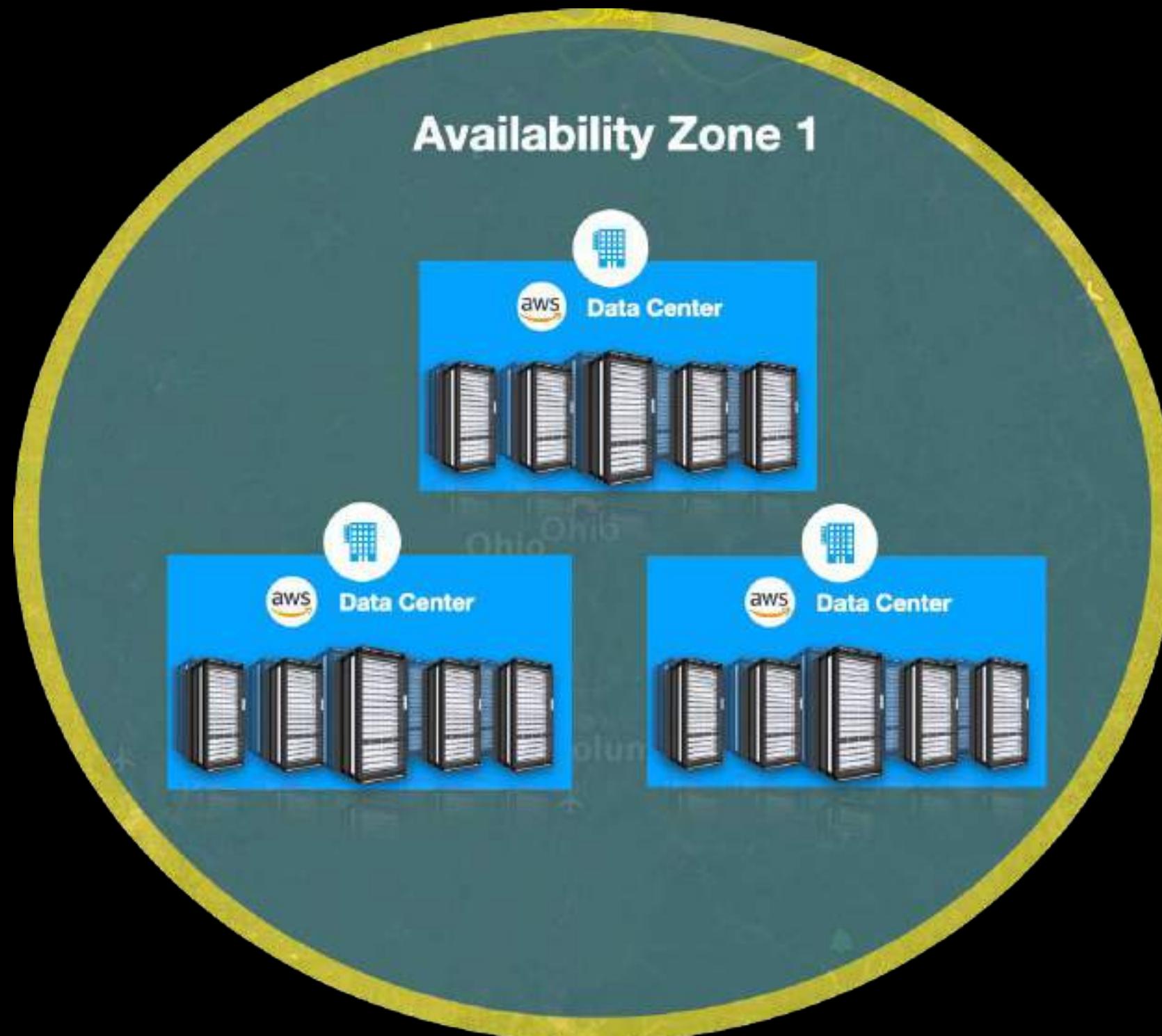
- For **Batch jobs that cannot be interrupted once started**
- For **consuming Amazon SQS queue messages in which the application should continually process messages without any downtime**
- Running the **master node or core nodes of your Amazon Elastic MapReduce cluster (*cheaper than On-Demand Instances*)**
- And many more!

# SCOPE

## Regional



## Zonal





# Dedicated Hosts & Dedicated Instances

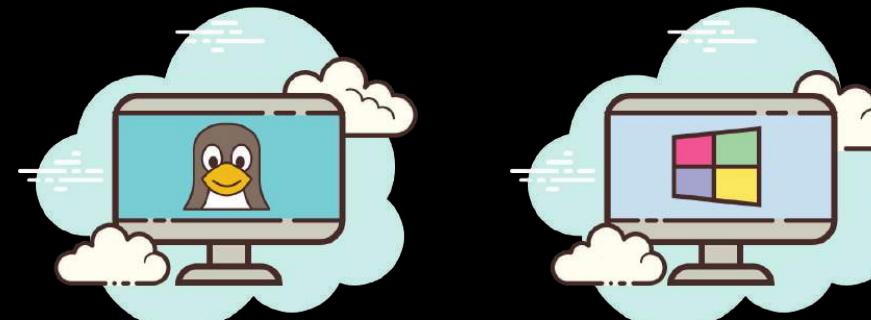
---

## TENANCY

It's like "renting" an entire house for your family, which you are the sole tenant (single-tenant).

If you share a house with your friends or co-workers, then there are multiple tenants (multi-tenant).

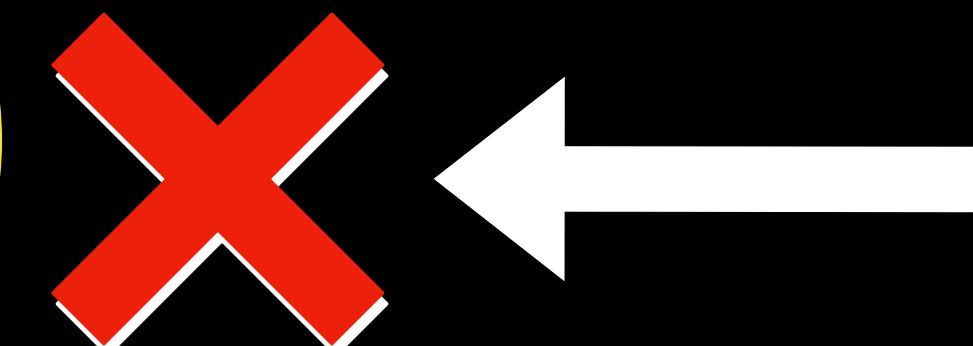
Dedicated Instance



Dedicated Host

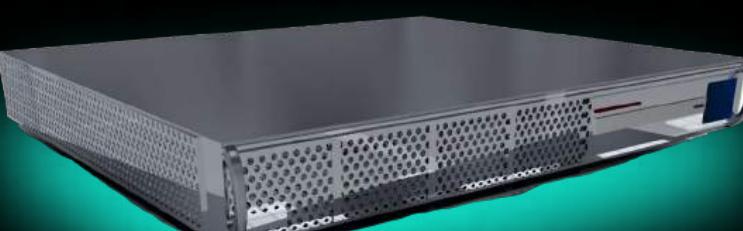


A **rack-mounted server** is also called a **HOST**



DEFAULT TENANCY

Used by a **SINGLE** Customer / Tenant



## DEDICATED HOST

A single, physical **rack-mounted server**  
or also known as a **host**



- **per-socket**
- **per-core** = CPU Core
- **per-VM**



Used by a **SINGLE** Customer / Tenant

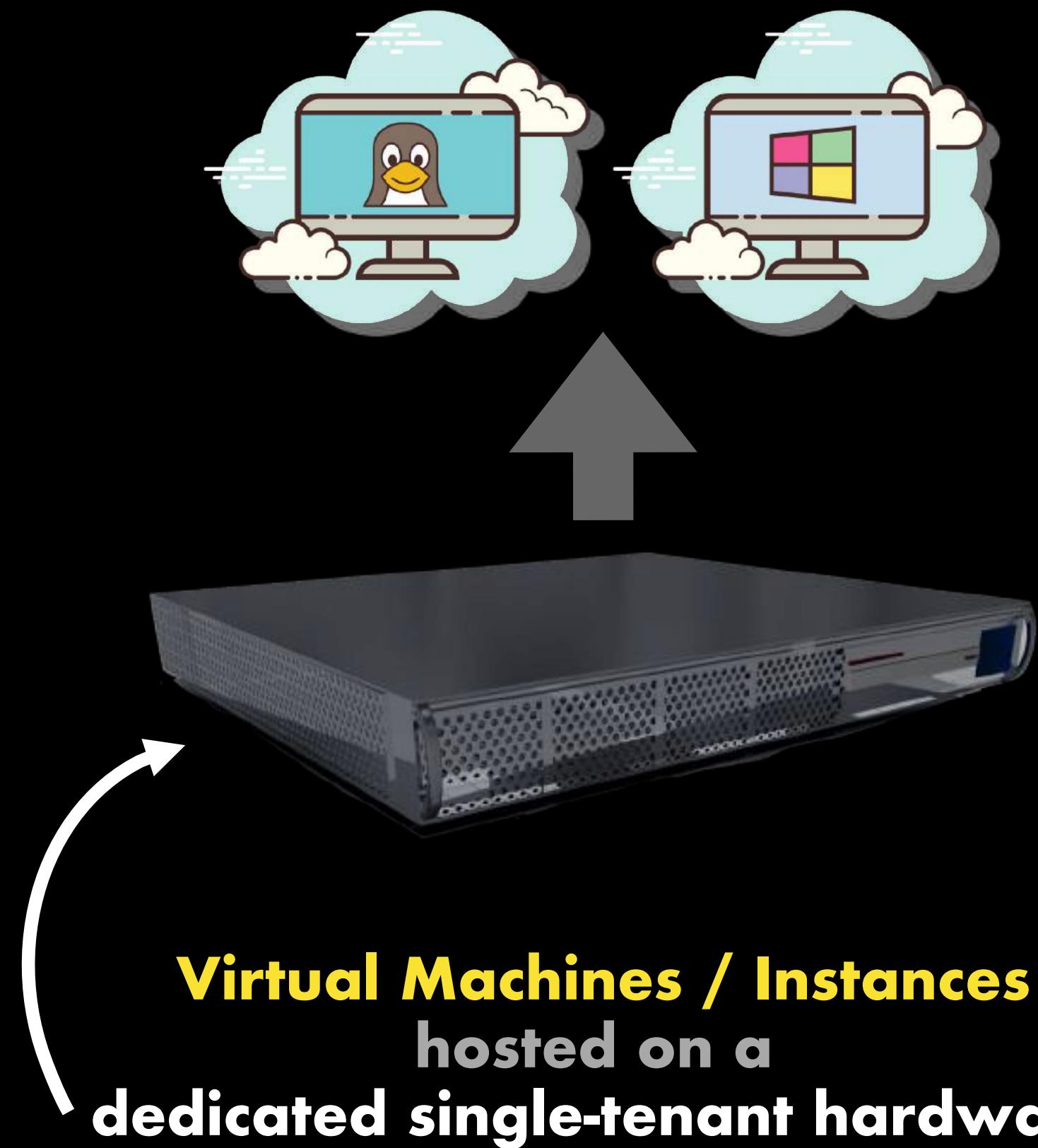
## DEDICATED HOST



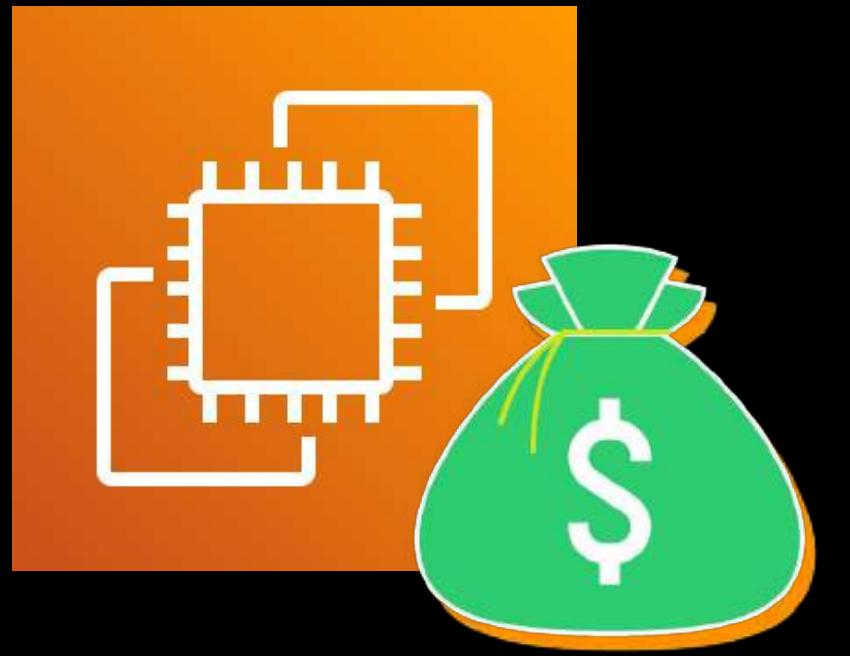
A **rack-mounted server / host**

- For cases when the **existing server-bound software licenses must be used by customers**
- To comply with your **per-core software license requirements**
- For compliance and software licensing requirements mandating that **a workload must be hosted on a physical server**
- For migrating commercial off-the-shelf applications with licenses that must still be utilized upon migration
- For performing cost analysis that supports **physical isolation of a customer workload**
- Launching Windows Server, SQL Server, SUSE Linux Enterprise Server, Red Hat Enterprise Linux, or other software licenses that are **bound to particular VMs, sockets, or physical CPU cores**

## DEDICATED INSTANCE

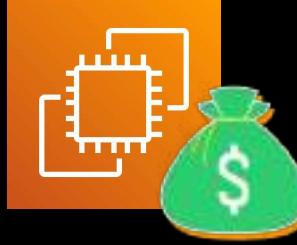


- Regular virtual machines that run in a virtual private cloud (VPC) on hardware that's dedicated to a single customer
- Dedicated Instances that belong to different AWS accounts are physically isolated at a hardware level
- Dedicated Instances may share hardware with other Amazon EC2 instances if the instances are:
  - In the same AWS account
  - Not a type of Dedicated Instance
- Allows you to launch Dedicated Spot Instances, Dedicated On-Demand Instances, or Dedicated Reserved Instances



# Savings Plans

---



## Savings Plans

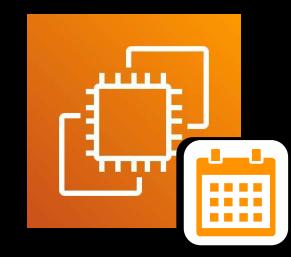
### FEATURES

- A flexible pricing model in AWS that helps you save on the usage of your:
  - Amazon EC2
  - AWS Fargate
  - AWS Lambda
- Provides discounts in exchange for a commitment to a consistent usage amount that is measured in dollars per hour for a one or three-year term
- Aside from Amazon EC2, it also cover other compute resources such as AWS Fargate and AWS Lambda
- Can be purchased from:
  - Any AWS account
  - Management account of your AWS Organization
  - Member account of your AWS Organization



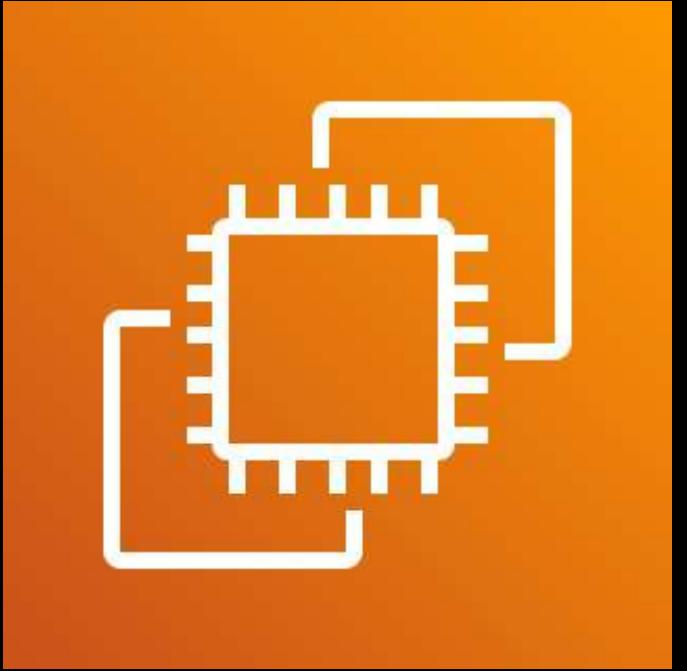
## Savings Plans

- Both require a fixed one-year or three-year commitment
- Both provides Billing Discounts



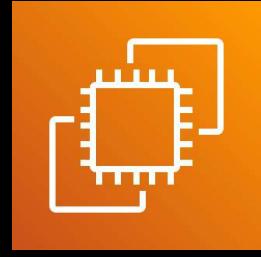
## Reserved Instances

- Based on a **consistent amount of compute usage**
- Provides **flexibility** to use a more suitable compute option at low prices **without any exchanges or modification**
- Based on a specific Instance Type or Instance Size
- Must exchange or modify the Reserved Instance to suit your current requirements



# Capacity Reservation

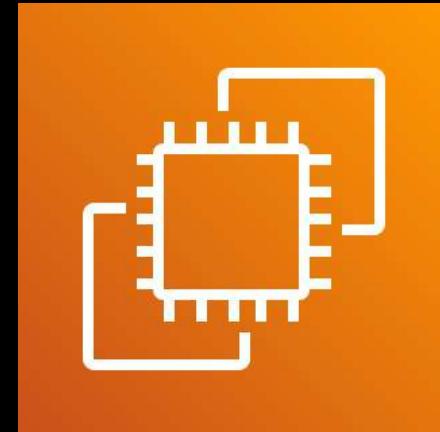
---



## Capacity Reservation

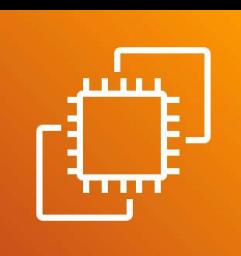
### FEATURES

- Allows you to reserve capacity for your EC2 instances in a specific **Availability Zone**
- Independent of the billing discounts offered by **Savings Plans** or **regional Reserved Instances**
- Works like a **Zonal Reserved Instance**
- No 1-year or 3-year commitment
- You can reserve a particular **Availability Zone** only (**Zonal**), no Regional reservations in scope
- Can be applied to **On-Demand EC2 Instances**



## Capacity Reservation REQUIREMENTS

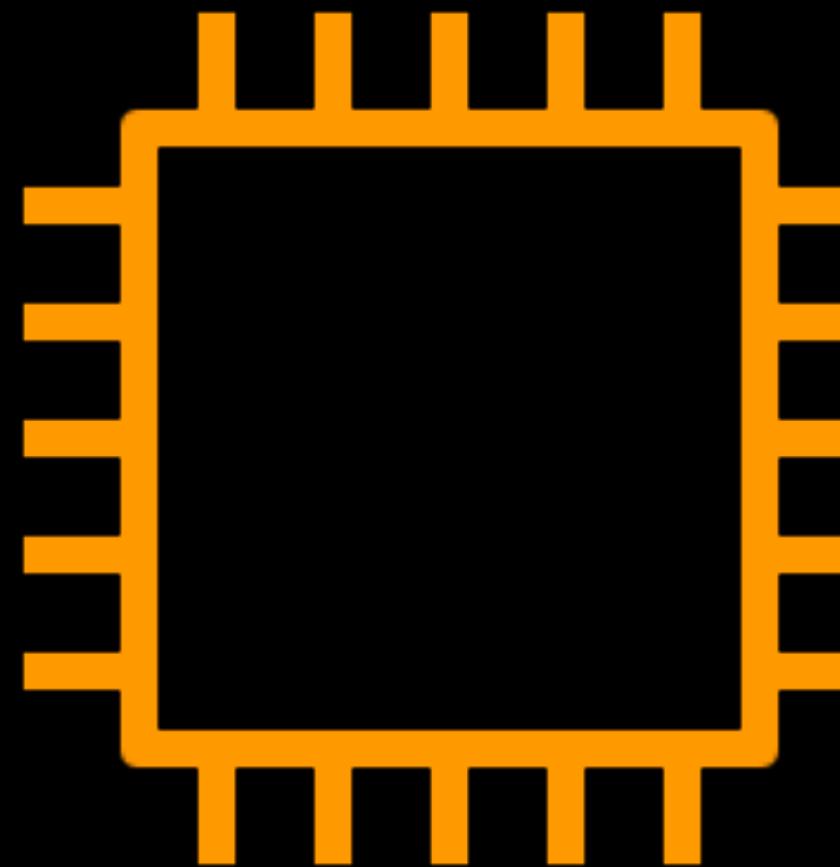
- **Availability Zone**
- **Number of Amazon EC2 Instances**
- **Instance Attributes (e.g. *instance type, OS, etc*)**



# Capacity Reservation

MATCH

Running EC2 Instances in your VPC

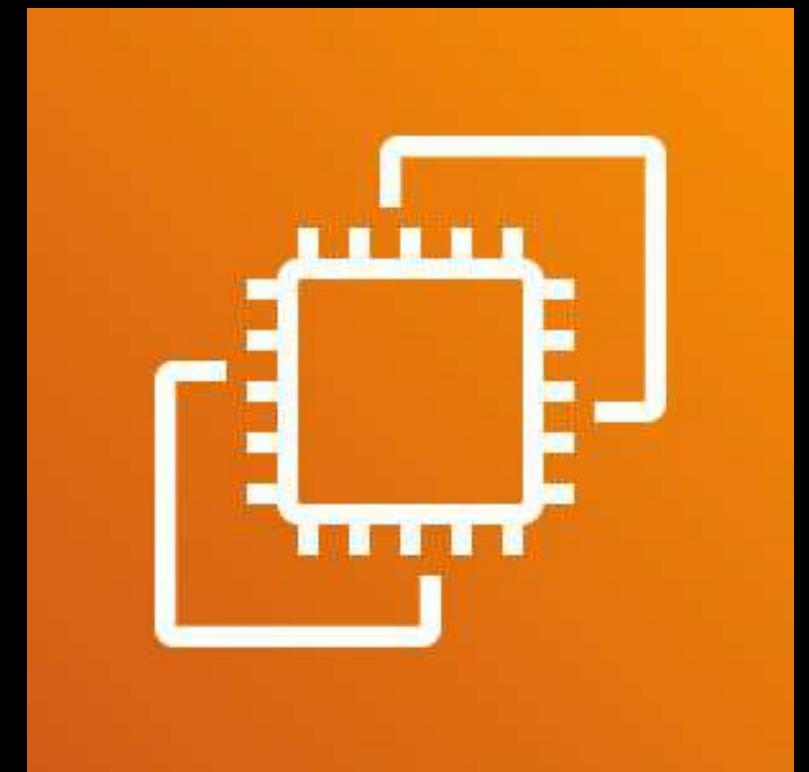


us-east-1a

2

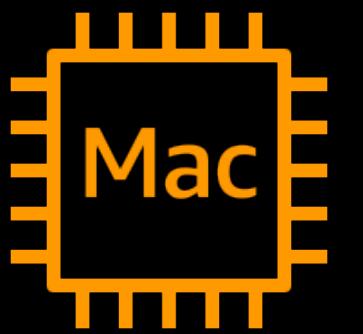
Instance Type:  
A3

- **Availability Zone**
- **Number of Amazon EC2 Instances**
- **Instance Attributes** (e.g. *instance type, OS, etc*)



# Amazon EC2 Instance Types

---



Mac Instances

\*Powered by **Mac Mini**

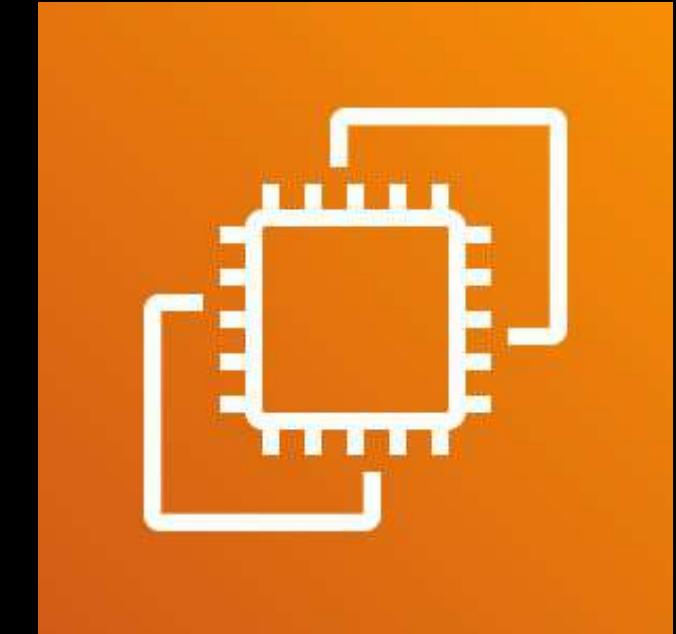


CPU

RAM

GRAPHICS

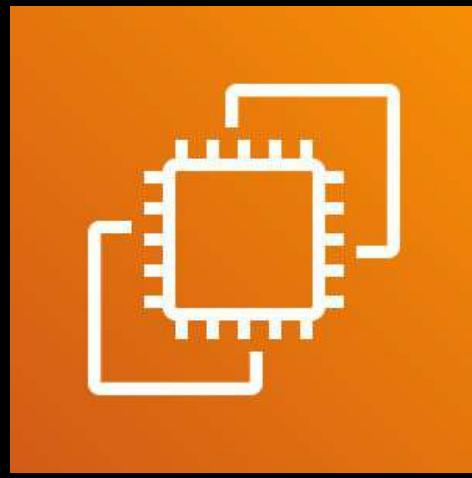
NETWORK



Amazon EC2 **Instance Type**

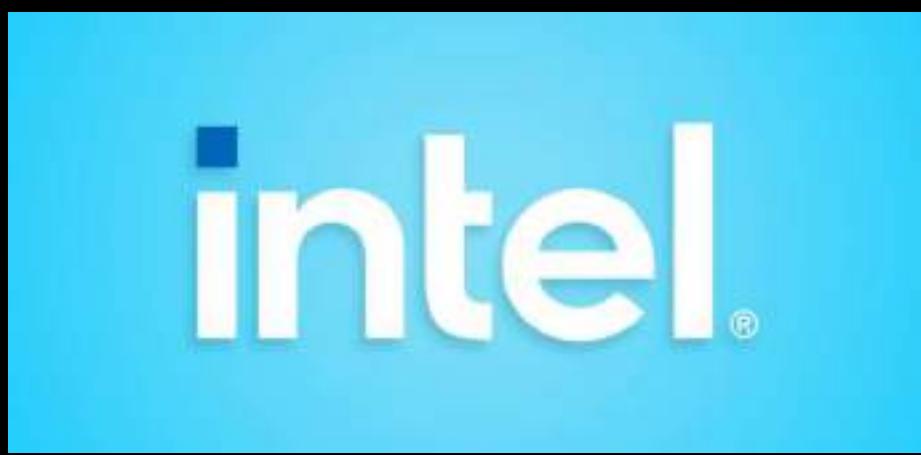
STORAGE

OTHER  
COMPONENTS...



## CPU OPTIONS

Amazon EC2 Instance Type



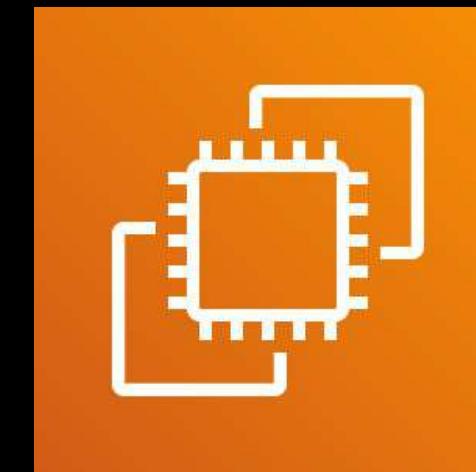
AWS Graviton

The **newer** your EC2 instance type is,  
the more **cost-efficient** and **powerful** it is.

also known as...

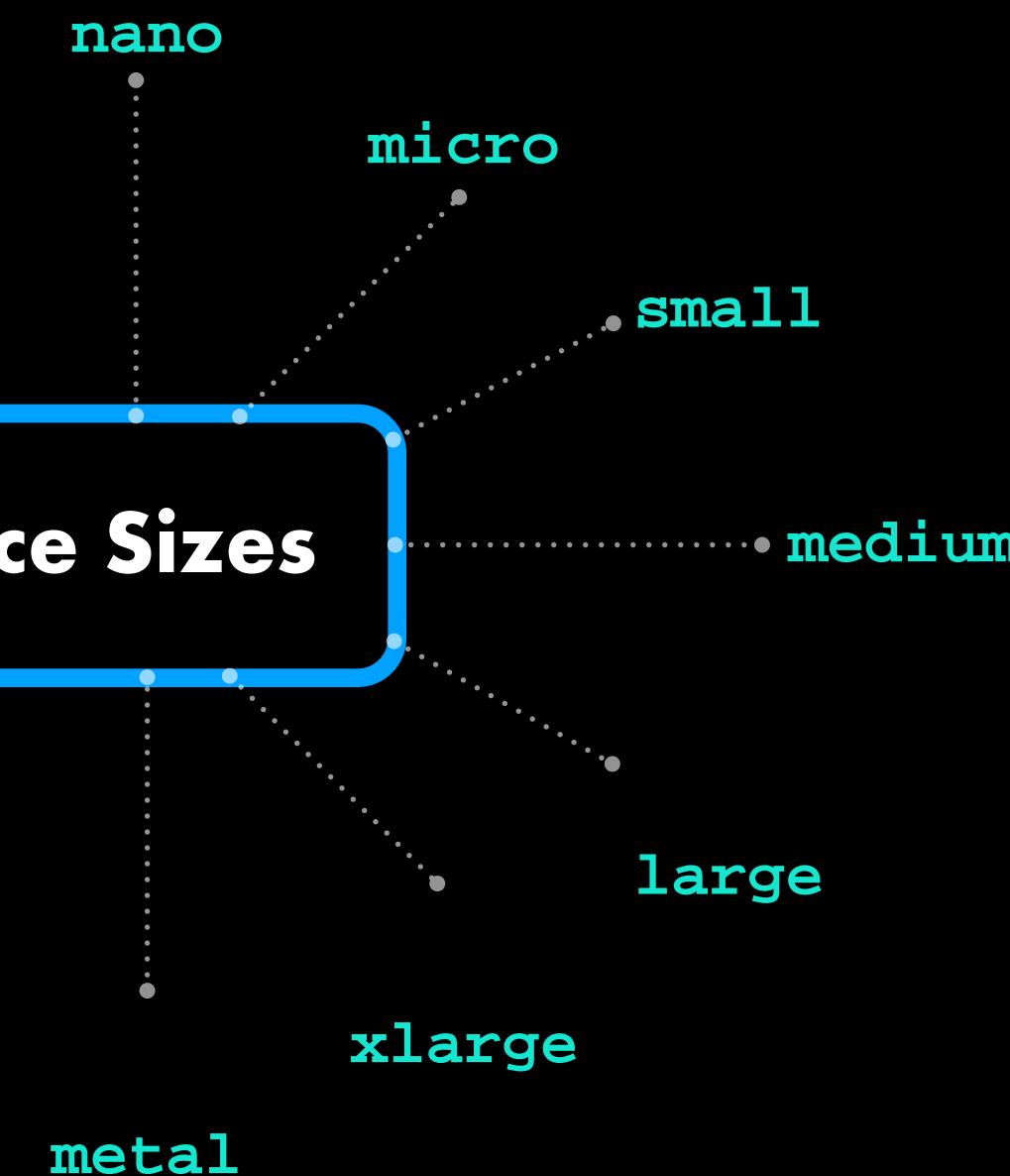


## INSTANCE FAMILY



Amazon EC2 Instance Type

Instance Sizes



## INSTANCE CATEGORIES

- General Purpose
- Compute Optimized
- Memory Optimized
- Storage Optimized
- Accelerated Computing
- Others

## INSTANCE FAMILY / TYPES

Mac, T\*, M\*, A\*

C\*

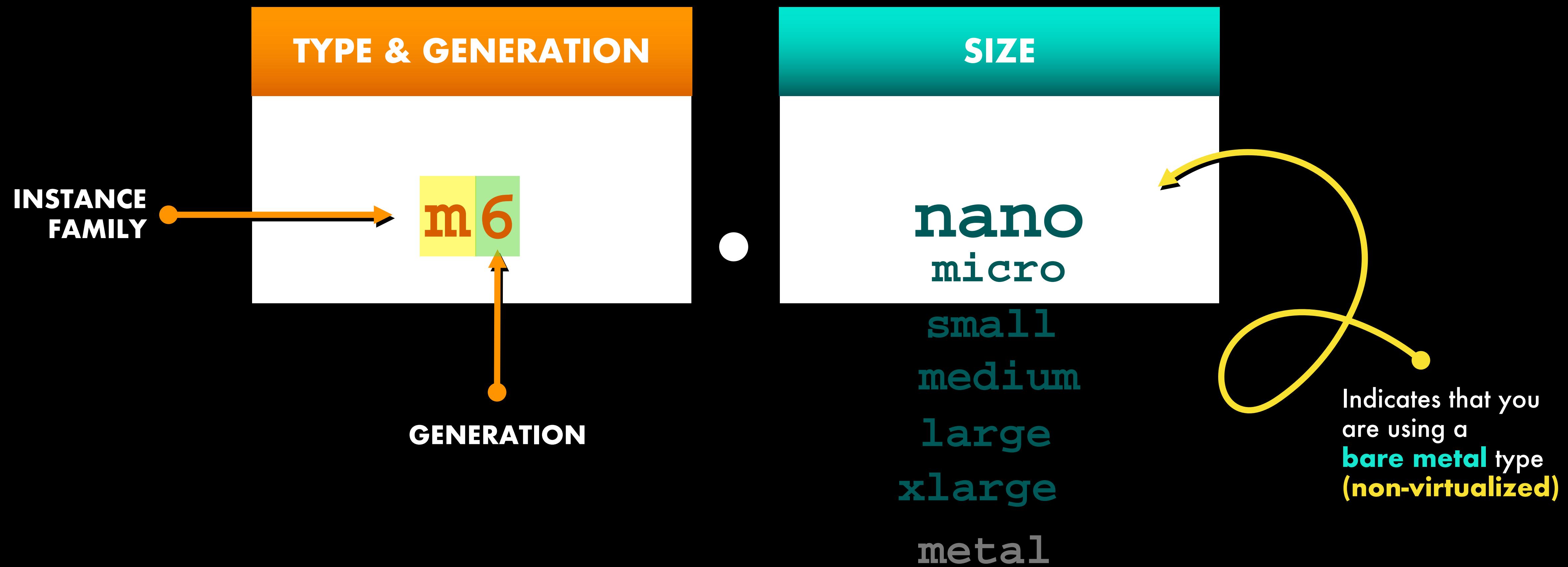
R\*, X\*, Z\*, U\*

I\*, D\*, H\*

P\*, Inf\*, G\*, F\*

More Instance Types to be launched soon!

# INSTANCE TYPE NAMING CONVENTION



# INSTANCE TYPE NAMING CONVENTION

## TYPE & GENERATION

m4 & below

m5

m6

m7 & above

PREVIOUS GENERATION

5th GENERATION

6th GENERATION

NEXT GENERATION

# INSTANCE TYPE NAMING CONVENTION

## TYPE & GENERATION

\*\*a

\*\*g

## CPU TYPE

AMD

AWS  
Graviton



# INSTANCE TYPE NAMING CONVENTION

## TYPE & GENERATION

t3a

m6g

t3, m5, r5

AMD

AWS  
Graviton

intel®

# INSTANCE TYPE NAMING CONVENTION

## TYPE & GENERATION

**\*\*\*d**

Has a local NVMe-based SSD storage

**\*\*\*n**

Has enhanced networking capabilities

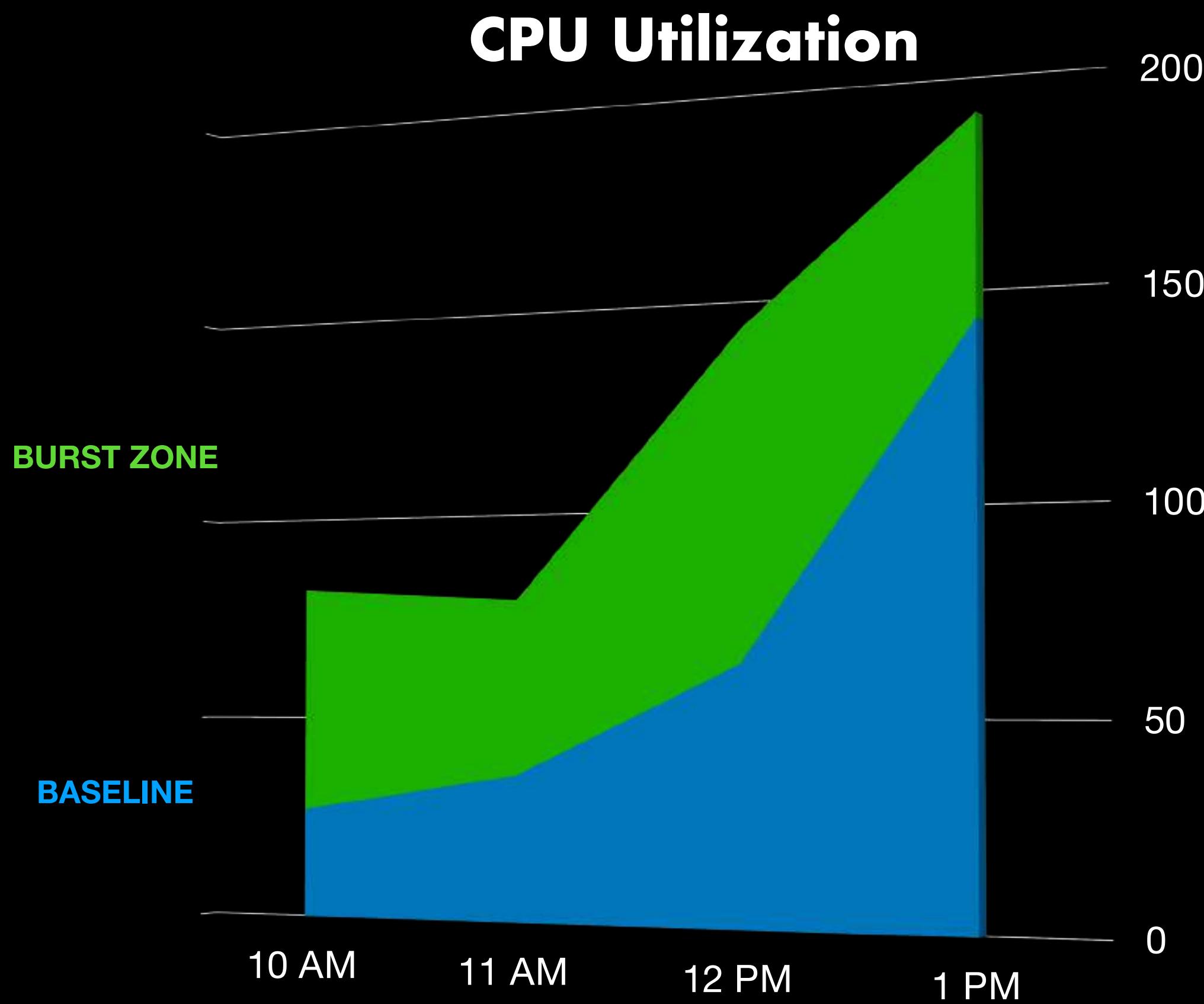
# INSTANCE TYPE NAMING CONVENTION

## TYPE & GENERATION

T

- **Burstable Performance Instances**
- **Provides a baseline level of CPU performance with the ability to burst above the baseline**
- **The ability to burst is governed by CPU Credits**

# INSTANCE TYPE NAMING CONVENTION



- A CPU Credit accrued when the instance is idle
- A sort of ‘vertical scaling’ since it temporarily provides higher CPU performance over the maximum CPU capacity of the instance
- A CPU Credit provides a full CPU core performance for one minute

# INSTANCE TYPE NAMING CONVENTION

- Bare metal instances
- Grants **direct access to the CPU and memory resources of the underlying server**
- Doesn't have a pre-installed KVM, Xen, or AWS Nitro Hypervisor that other EC2 instances use
- Allows you to fully access the CPU, Storage, and Networking bandwidth of the underlying server
- Allows customers to run their own hypervisor or virtualization secured containers such as Clear Linux Containers



# INSTANCE TYPE NAMING CONVENTION

- Meant for customers who have enterprise applications that need to run in non-virtualized environments or need to use their own hypervisor
- Can still be integrated with Amazon EBS, Elastic Load Balancers, and other resources on your Amazon VPC
- Provides the highest attributes across all other types in its Instance Family
- Have equal or more attributes than the largest instance type in the instance family

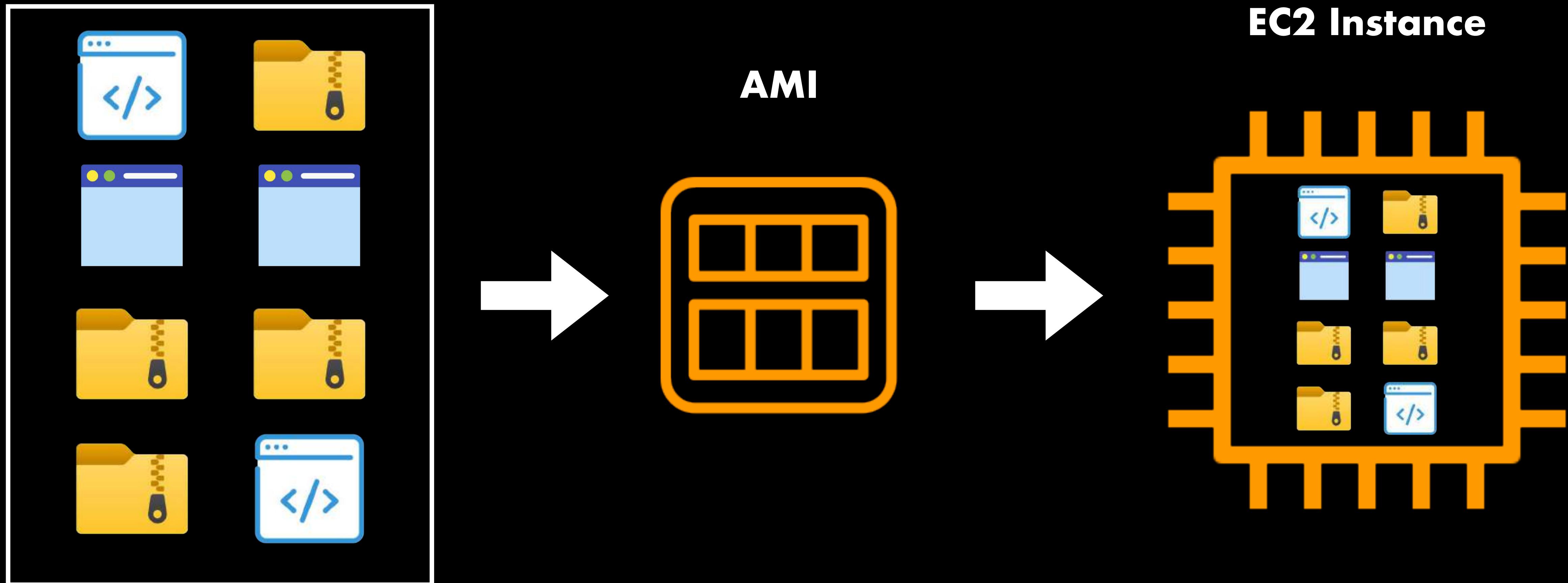
Instance Size	vCPU	Memory (GiB)	Instance Storage (GiB)	Network Bandwidth (Gbps)	EBS Bandwidth (Mbps)
c6g.medium	1	2	EBS-Only	Up to 10	Up to 4,750
c6g.large	2	4	EBS-Only	Up to 10	Up to 4,750
c6g.xlarge	4	8	EBS-Only	Up to 10	Up to 4,750
c6g.2xlarge	8	16	EBS-Only	Up to 10	Up to 4,750
c6g.4xlarge	16	32	EBS-Only	Up to 10	4750
c6g.8xlarge	32	64	EBS-Only	12	9000
c6g.12xlarge	48	96	EBS-Only	20	13500
c6g.16xlarge	64	128	EBS-Only	25	19000
c6g.metal	64	128	EBS-Only	25	19000
c6gd.medium	1	2	1 x 59 NVMe SSD	Up to 10	Up to 4,750
c6gd.large	2	4	1 x 118 NVMe SSD	Up to 10	Up to 4,750
c6gd.xlarge	4	8	1 x 237 NVMe SSD	Up to 10	Up to 4,750
c6gd.2xlarge	8	16	1 x 474 NVMe SSD	Up to 10	Up to 4,750
c6gd.4xlarge	16	32	1 x 950 NVMe SSD	Up to 10	4,750
c6gd.8xlarge	32	64	1 x 1900 NVMe SSD	12	9,000
c6gd.12xlarge	48	96	2 x 1425 NVMe SSD	20	13,500
c6gd.16xlarge	64	128	2 x 1900 NVMe SSD	25	19,000
c6gd.metal	64	128	2 x 1900 NVMe SSD	25	19,000



# Amazon Machine Image (AMI)

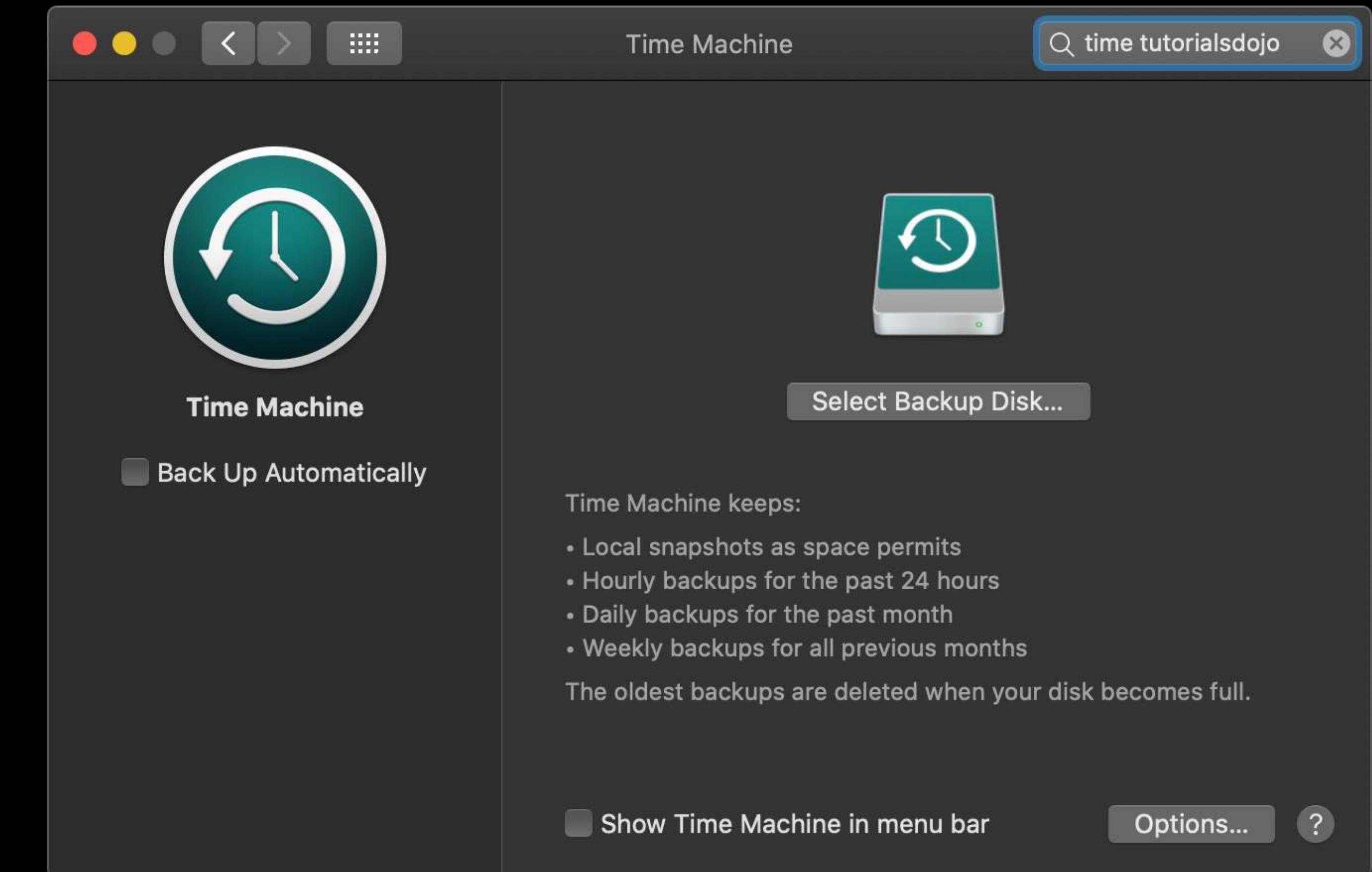
---

## apps & configurations



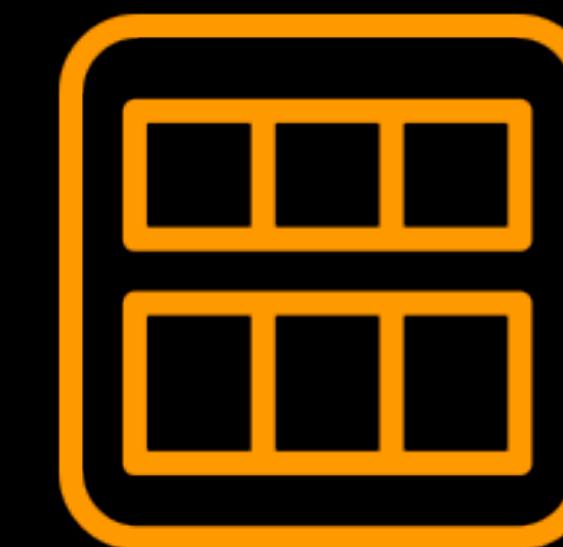
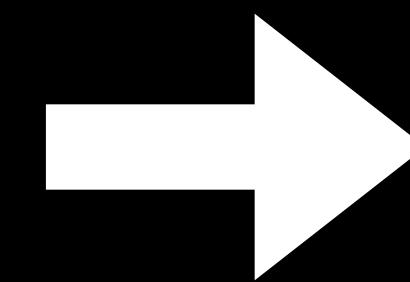


# DISK IMAGE

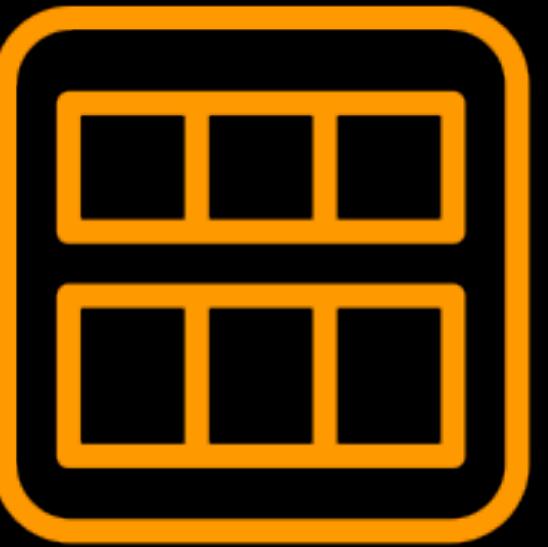




**DISK IMAGE**



**Amazon Machine Image  
(AMI)**



## Amazon Machine Image (AMI)

Volume Snapshots

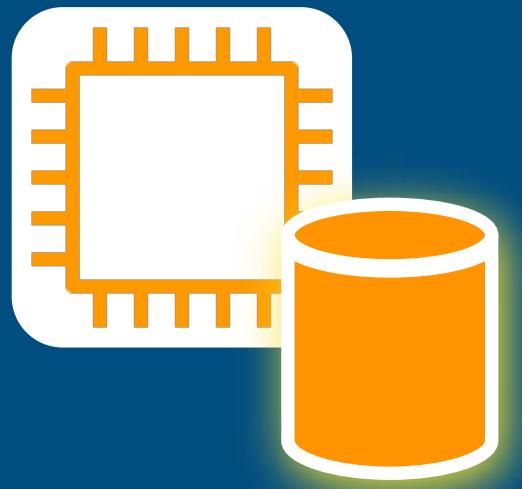
Block Device Mapping

Launch Permissions

## BLOCK STORE TYPE



**Amazon EBS**

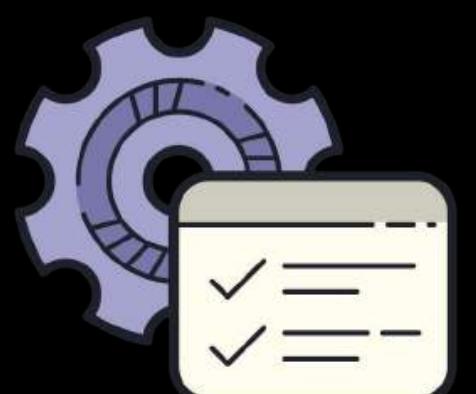


**Amazon EC2  
Instance Store**

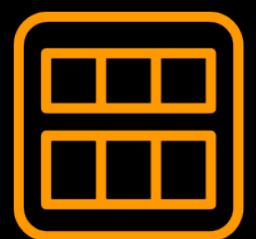
Volume Snapshots



EBS Snapshots

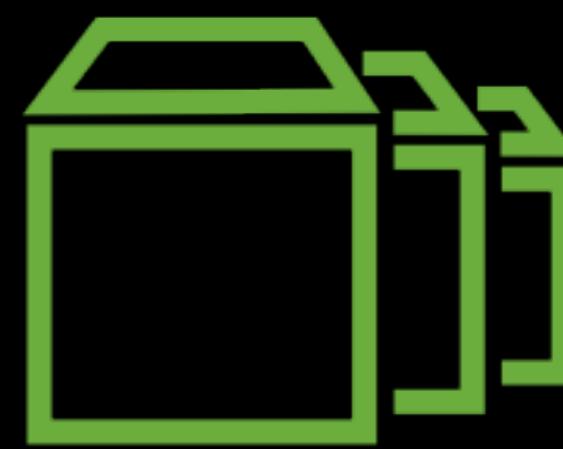


Template for the root  
volume



**Amazon Machine Image  
(AMI)**

Block Device Mapping



Amazon EBS Volumes  
mapping

N/A

Launch Permissions

- Public
- Explicit
- Implicit

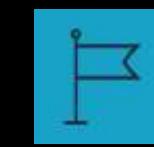


## Amazon Machine Image (AMI)

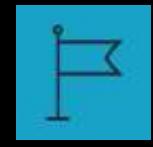
- **Regional** in scope
- You can **copy your AMI to another AWS Region**
- You can also copy your AMI to **another AWS account**



AWS Cloud



N. Virginia Region

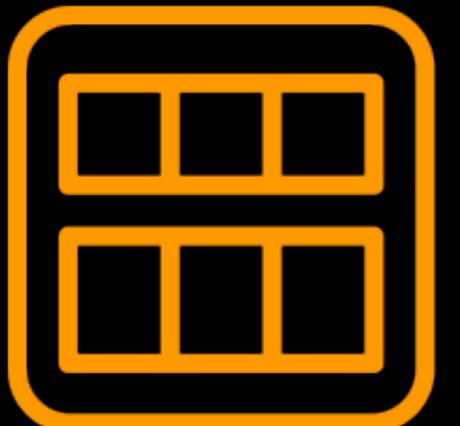
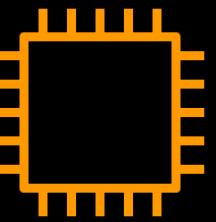


Ohio Region

Availability Zone (AZ)

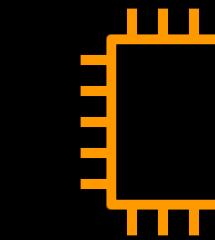
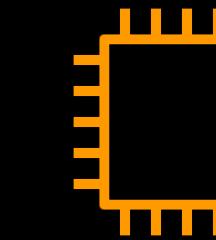
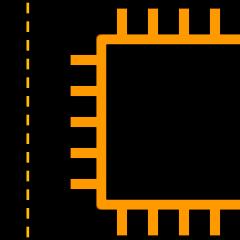
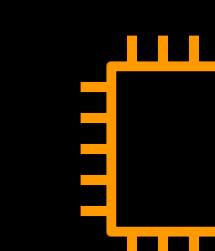
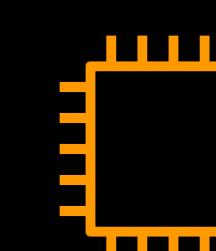
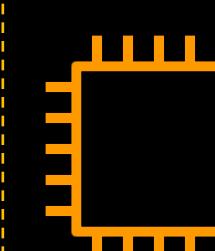


VPC A



COPY AMI

VPC A



# AWS Marketplace





## VIRTUALIZATION TYPE

# PV

Paravirtual

# HVM

Hardware  
Virtual Machine



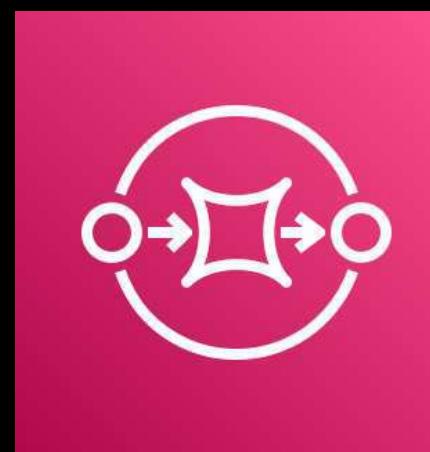
### BOOT UP PROCESS

Uses special boot  
loader called **PV-GRUB**

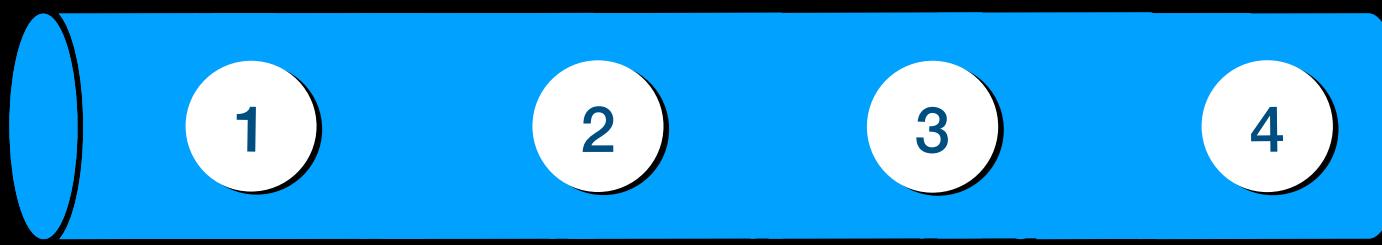
N/A

Executes the master boot  
record of the root block  
device of your image

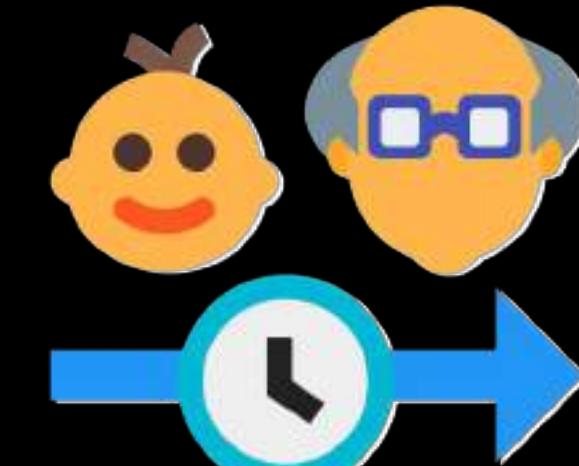
Uses several  
special hardware extensions  
such as  
**enhanced networking or**  
**GPU processing**



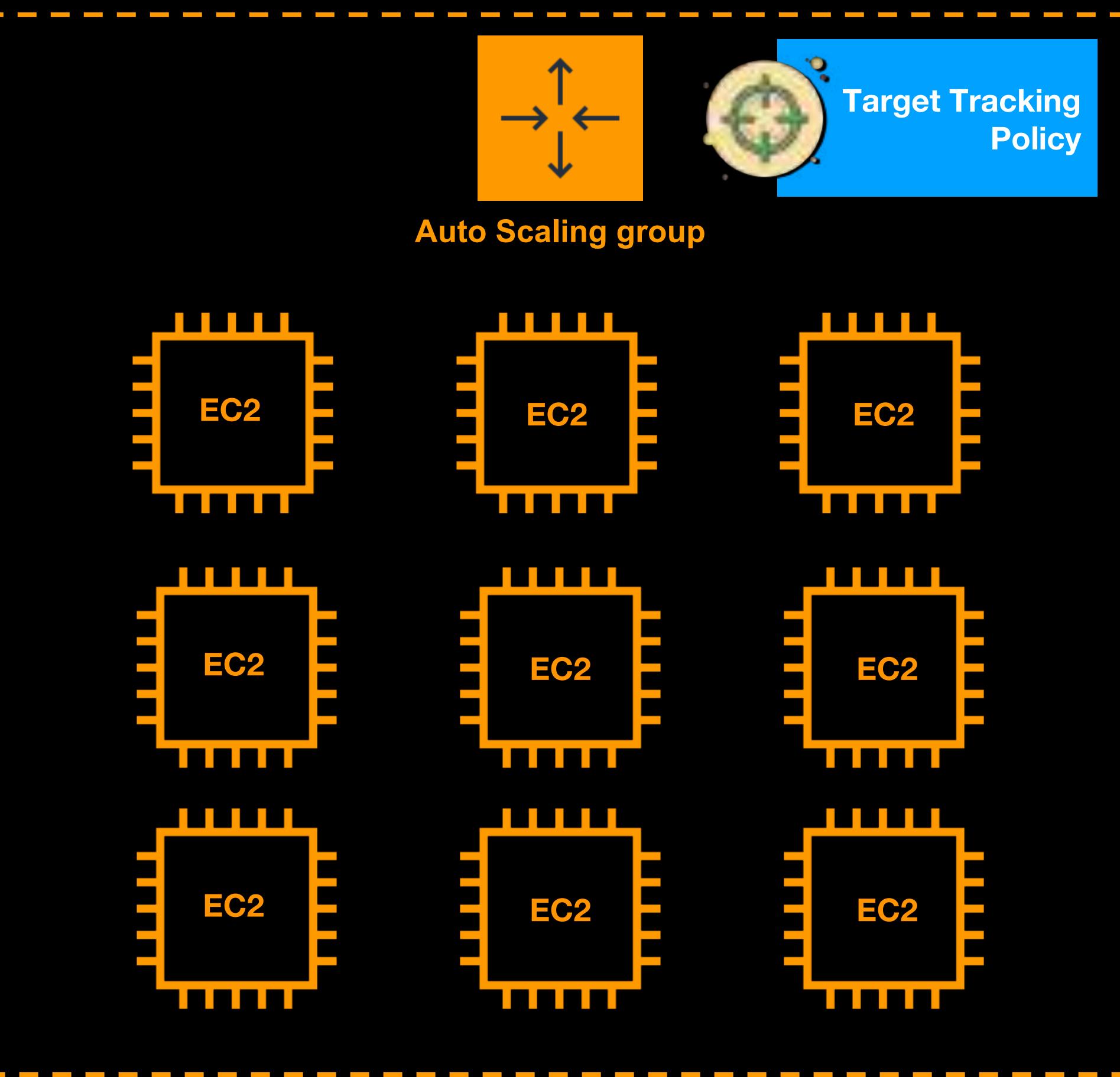
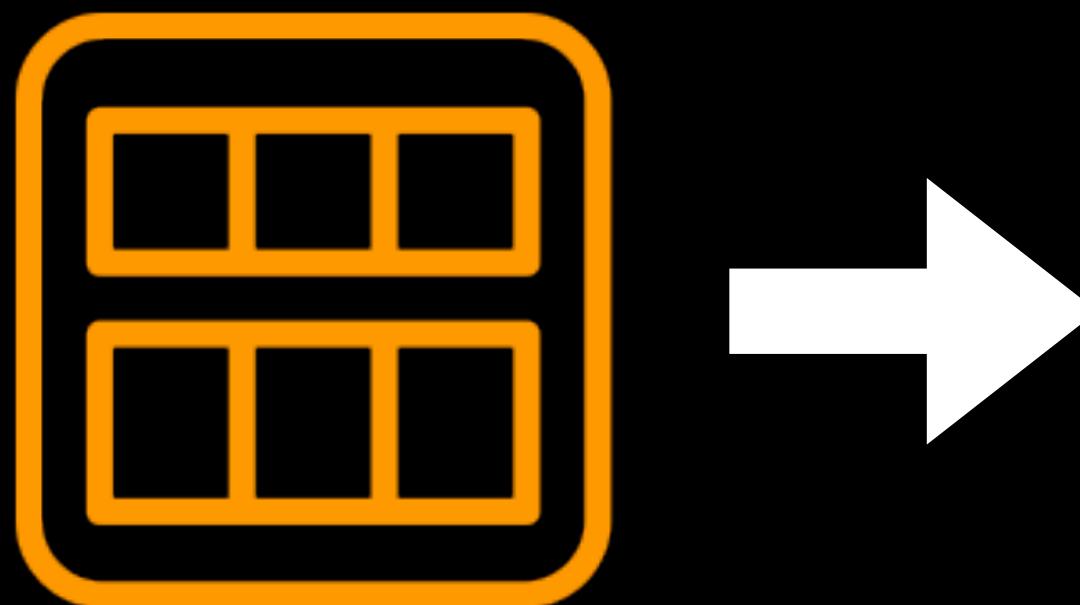
Amazon SQS

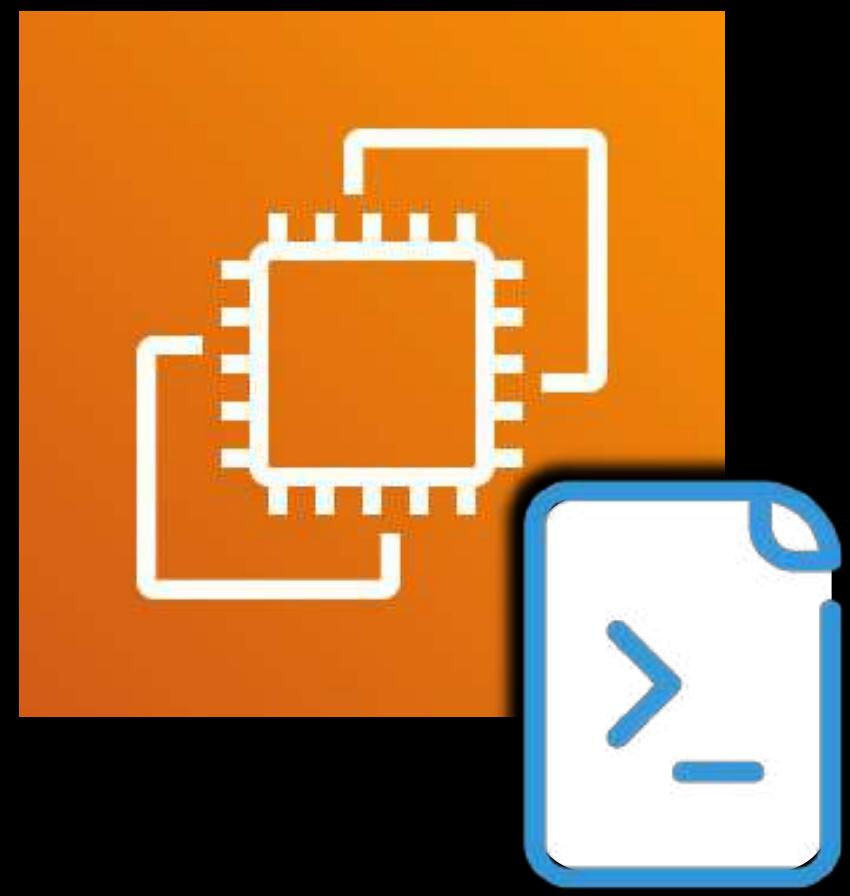


- **Age** of the Oldest Message



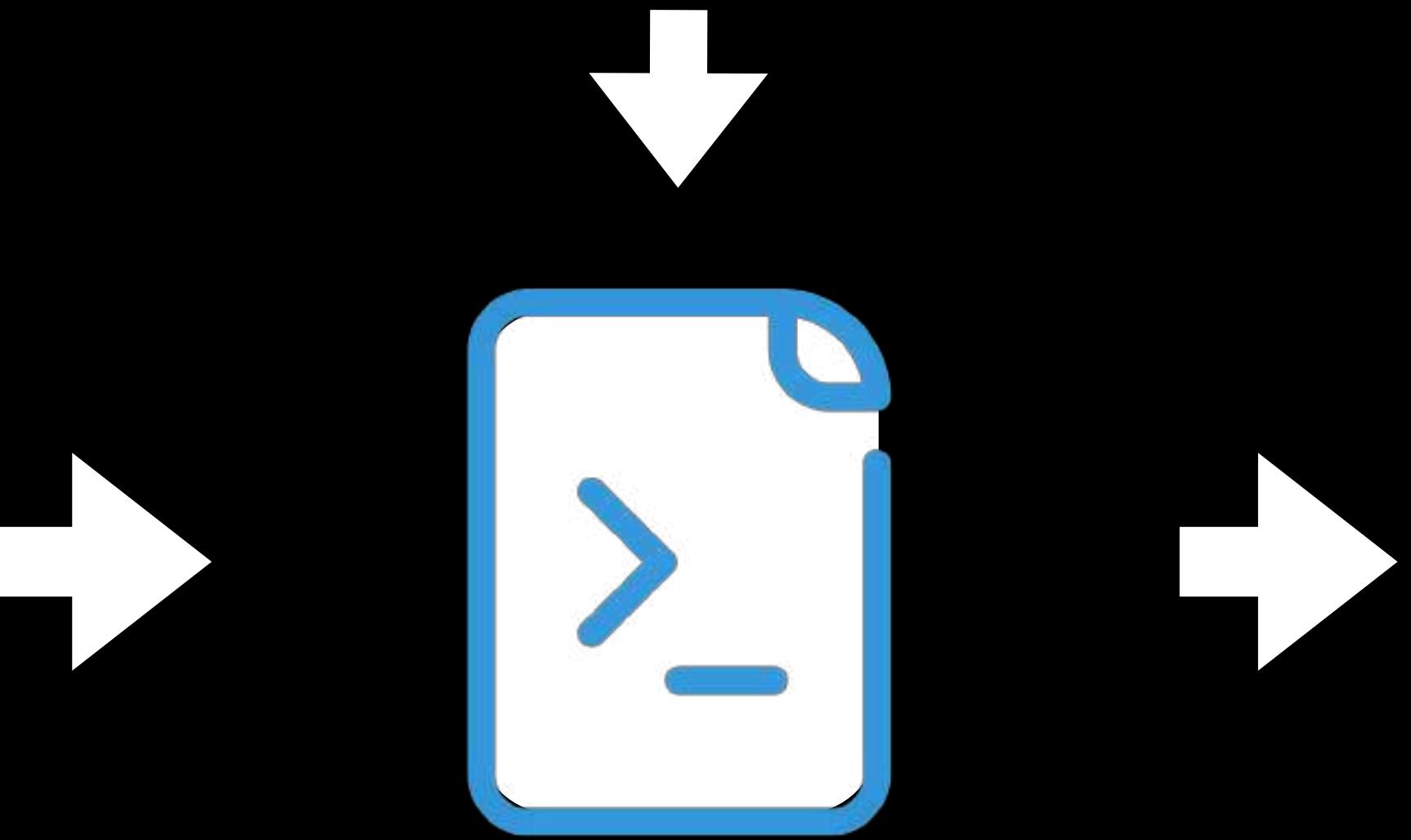
Amazon Machine Image  
(AMI)





# Instance User Data

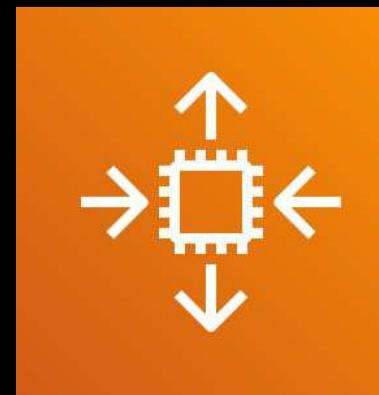
```
#!/bin/bash  
yum update -y  
mkdir tdojologs  
systemctl start httpd  
echo "tutorialsdojo OK!"
```



## User Data

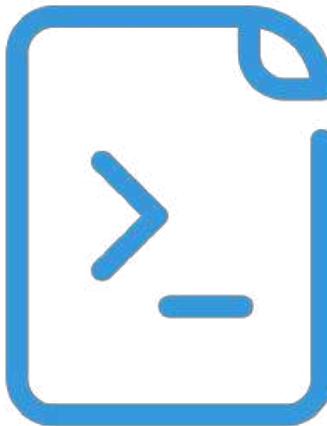


Amazon EFS



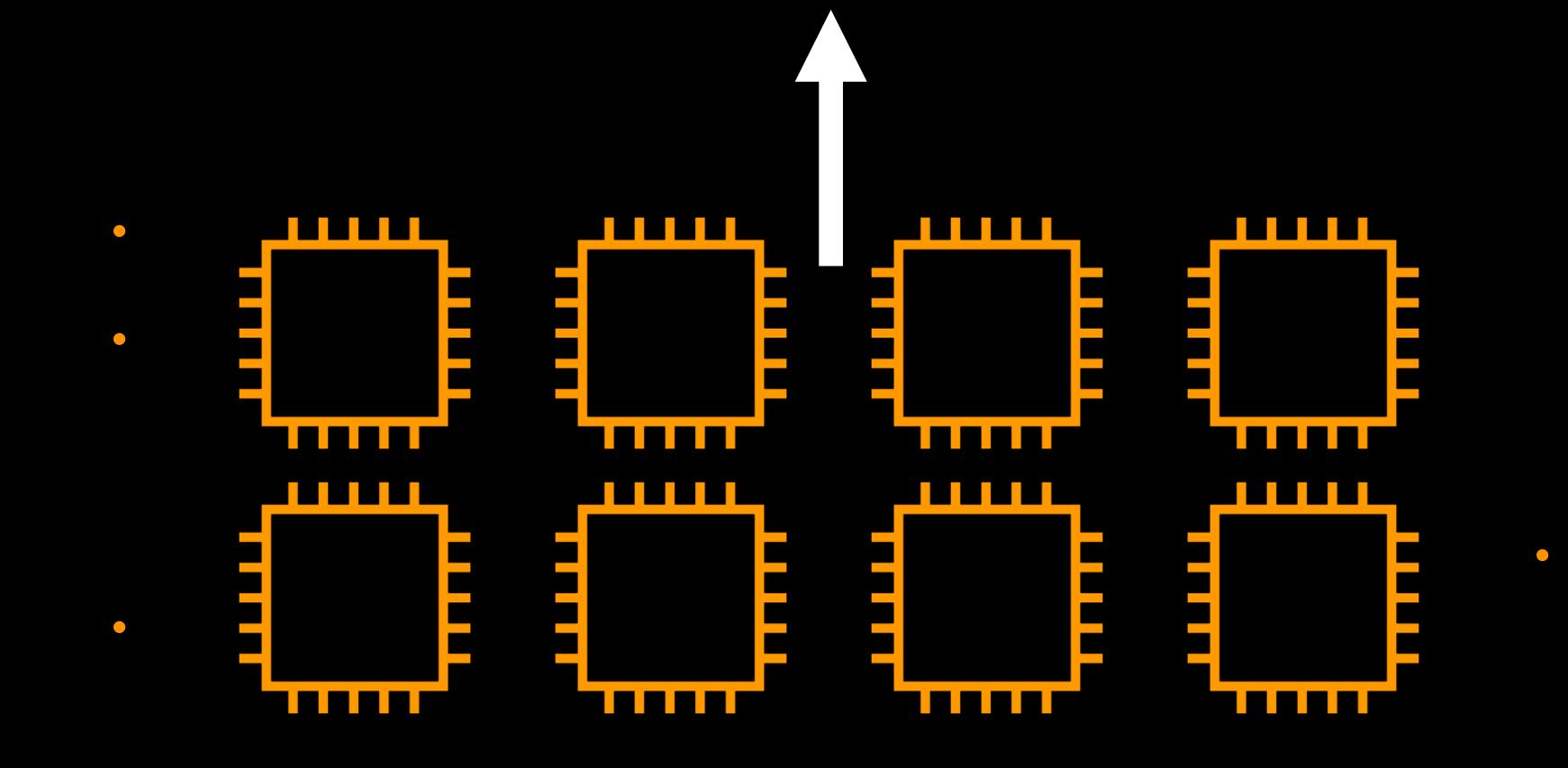
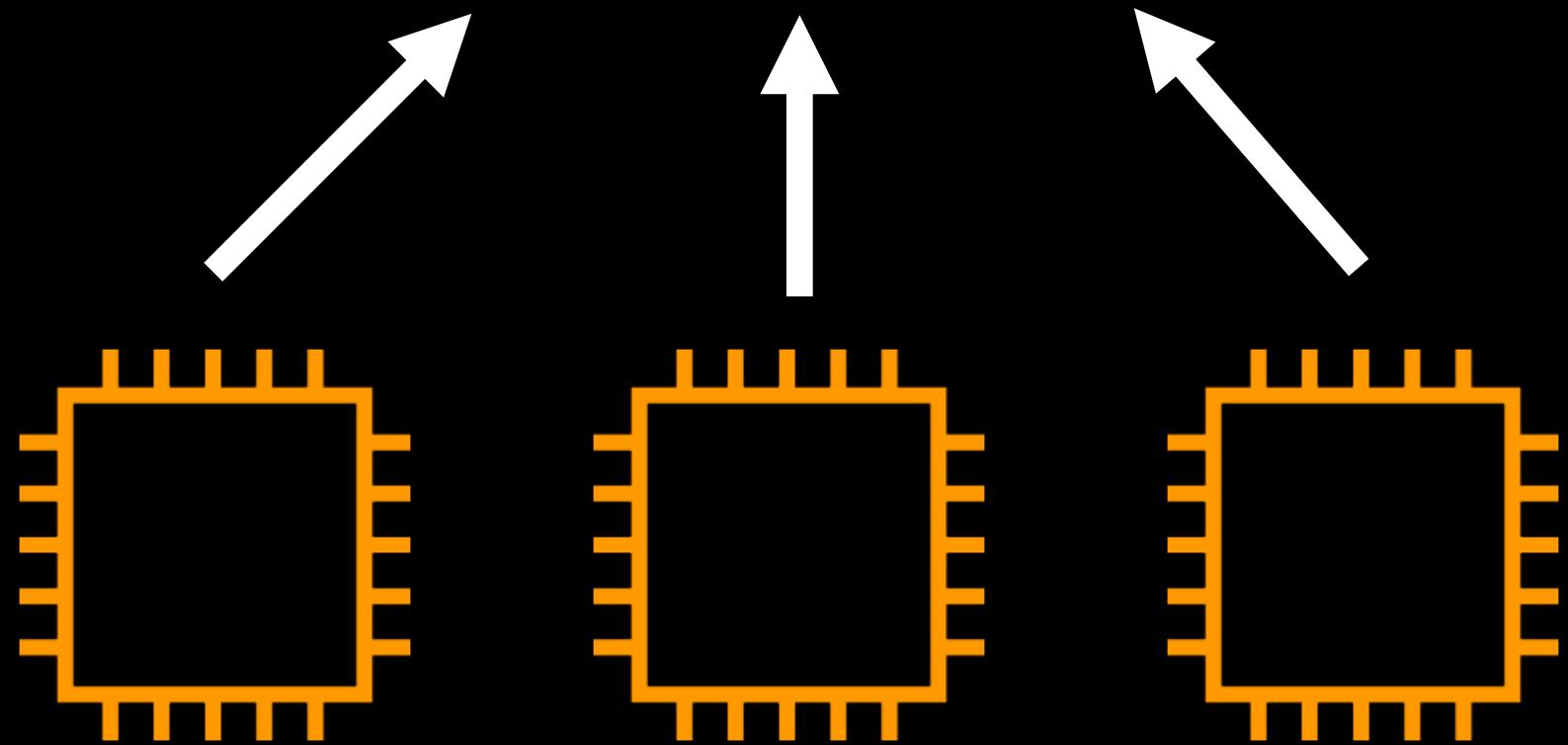
Auto Scaling Group

```
mkdir ~/tutorialsdojo-efs  
  
sudo mount -t nfs -o nfsvers=4.1,\  
rsize=1048576,wsize=1048576,hard,\  
timeo=600,retrans=2,noresvport \  
awsjonbonsoefs:/ ~/tutorialsdojo-efs
```



```
#!/bin/bash  
curl https://s3.amazonaws.com/aws-  
cloudwatch/downloads/latest/awslogs-agent-  
setup.py -O  
chmod +x ./awslogs-agent-setup.py  
./awslogs-agent-setup.py -n -r us-east-1 -c  
s3://tutorialsdojo
```

## User Data



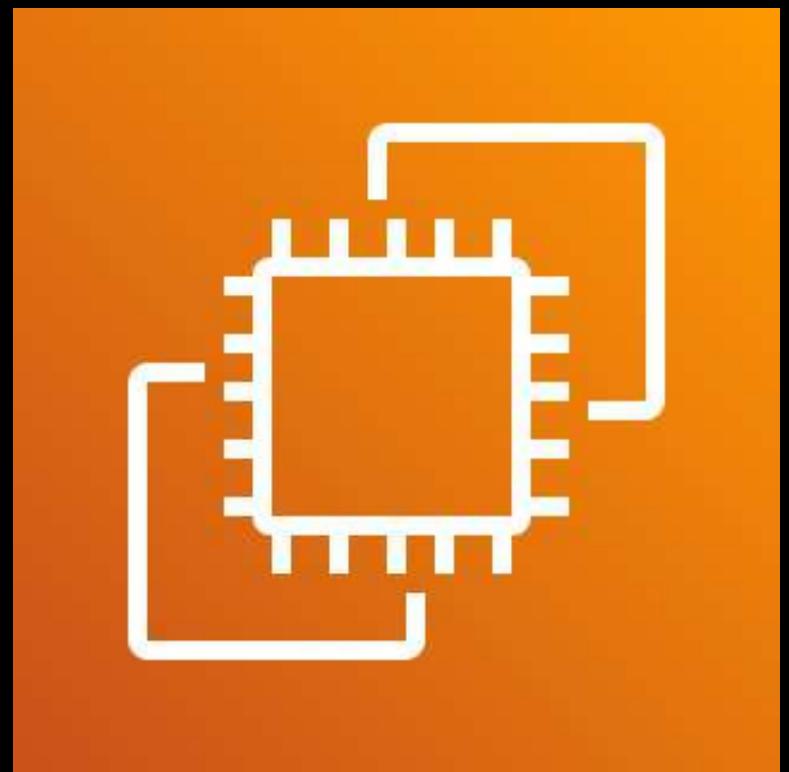
- Must be in a **base64-encoded format**
- Limited to **16 KB only when in raw form**
- Accessible from the Instance Metadata using this URI:



## User Data

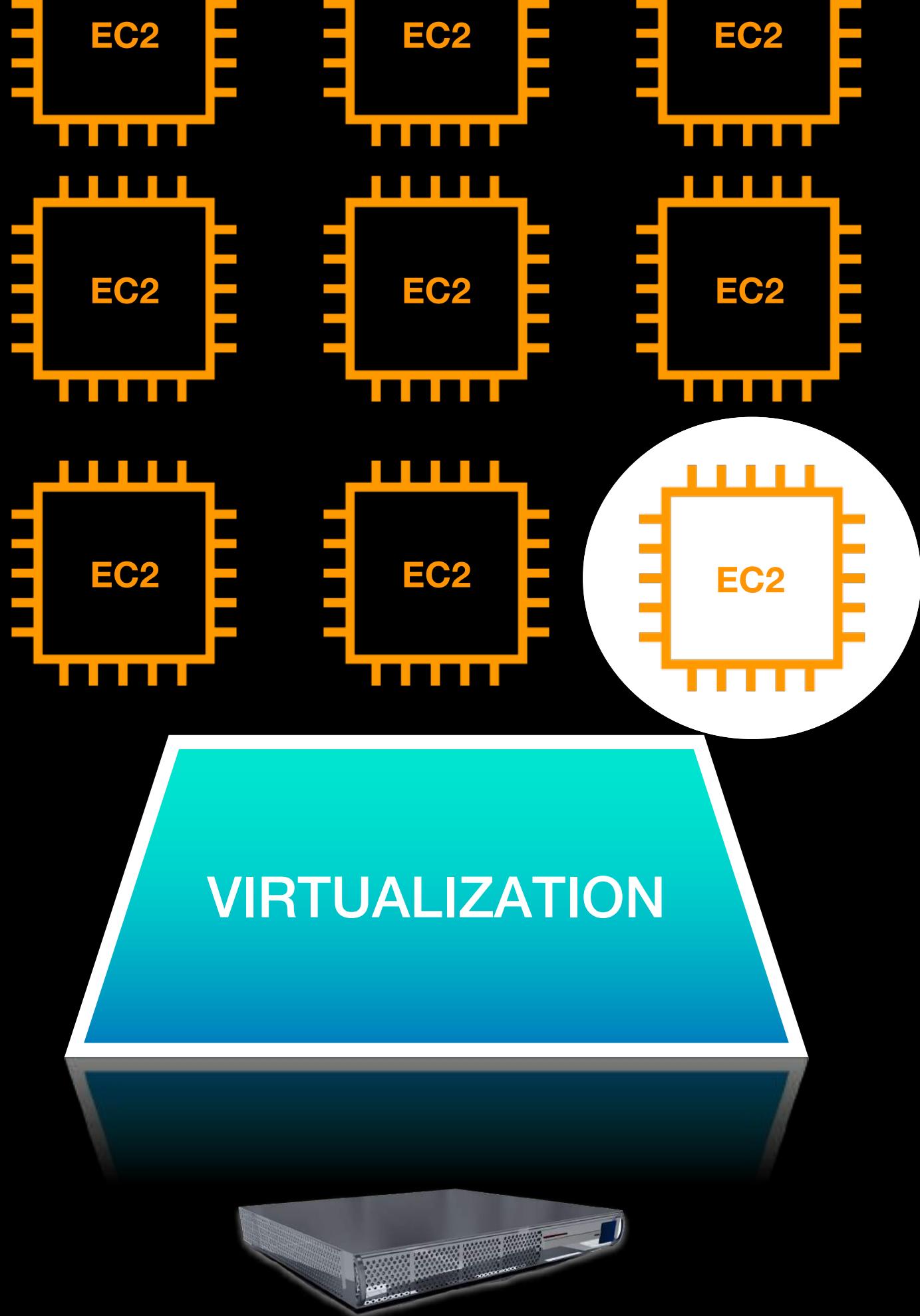
- Only run once upon the first EC2 Instance Launch
- Modifying the User Data and restarting the instance won't affect the initial User Data

<http://169.254.169.254/latest/user-data>

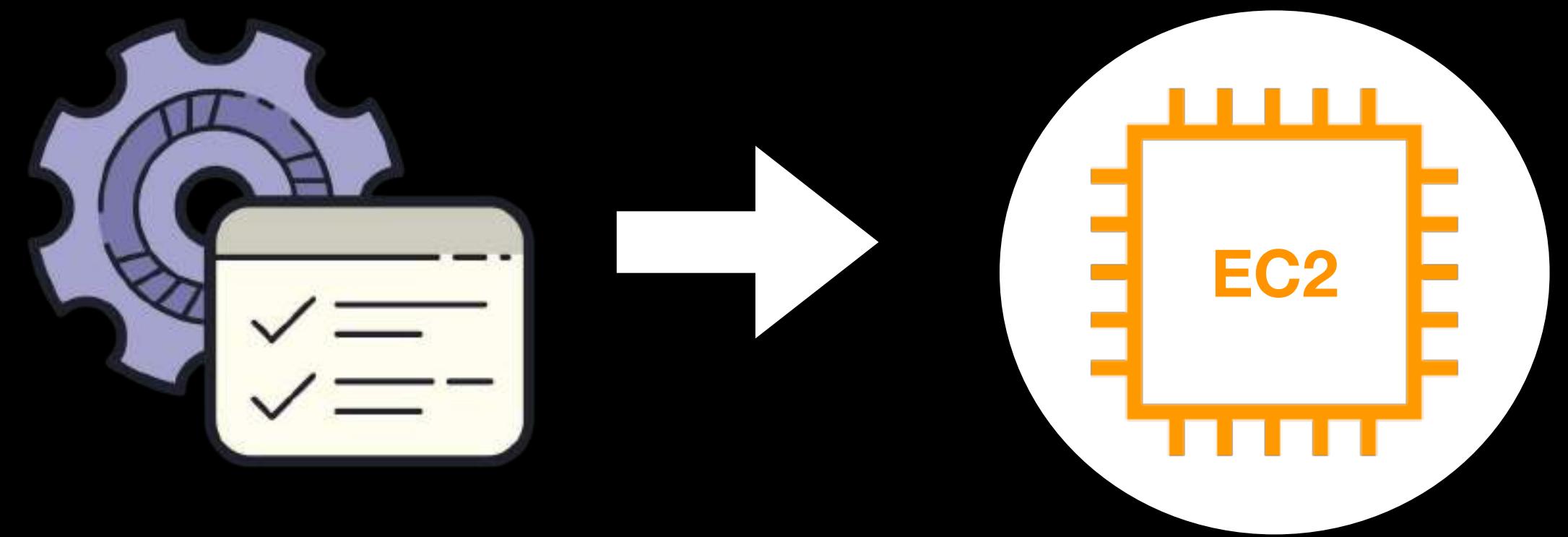


# Instance Metadata

---



**MANIFEST**

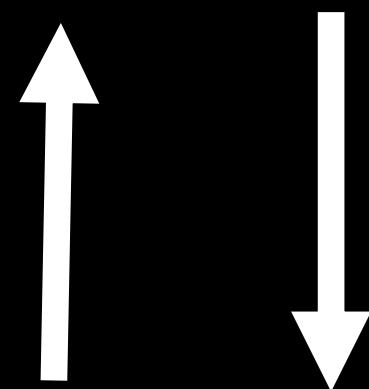




## INSTANCE METADATA

- **AMI**
- **Hostname**
- **Public IP address**
- **Private IP address**
- **Instance type**
- **MAC address**
- **Security groups**
- **Security credentials**
- **IAM Roles of your instance**
- **... and many more!**

## **Link-local Address**



**<http://169.254.169.254/latest/meta-data/>**

**INSTANCE METADATA SERVICE**



```
jonbonso@tutorialsdojo ~ curl http://169.254.169.254/latest/meta-data
ami-id
ami-launch-index
ami-manifest-path
block-device-mapping/
events/
hostname
iam/
identity-credentials/
instance-action
instance-id
instance-life-cycle
instance-type
local-hostname
local-ipv4
mac
metrics/
network/
placement/
profile
public-hostname
```



```
Administrator: Windows PowerShell
Windows PowerShell
Copyright (c) Microsoft Corporation. All rights reserved.

PS C:\Users\Administrator> Invoke-RestMethod -uri http://169.254.169.254/latest/meta-data/
ami-id
ami-launch-index
ami-manifest-path
block-device-mapping/
events/
hostname
identity-credentials/
instance-action
instance-id
instance-life-cycle
instance-type
local-hostname
local-ipv4
mac
metrics/
network/
placement/
profile
public-hostname
public-ipv4
public-keys/
reservation-id
security-groups
services/
PS C:\Users\Administrator>
```

# INSTANCE METADATA SERVICE

## version 2

Session Oriented

# CATEGORIES

```
jonbonso@tutorialsdojo >curl http://169.254.169.254/latest/meta-data
ami-id
ami-launch-index
ami-manifest-path
block-device-mapping/
events/
hostname
iam/
identity-credentials/
instance-action
instance-id
instance-life-cycle
instance-type
local-hostname
local-ipv4
mac
metrics/
network/
placement/
profile
public-hostname
```

# Private IP Address



```
portal — Tutorials Dojo Philippines - Amazon Elastic Compute Cloud Connection - to TEST Environment - Testing Instance Metadata - https://tutorialsdojo.com - All Rights Reserved -  
Cebu Server  
jonbonso@tutorialsdojo > curl http://169.254.169.254/latest/meta-data/local-ipv4  
172.31.76.5  
jonbonso@tutorialsdojo >
```

## Public IP or Elastic IP Address



```
portal — Tutorials Dojo Philippines - Amazon Elastic Compute Cloud Connection - to TEST Environment - Testing instance Metadata - https://tutorialsdojo.com - Cebu Server  
jonbonso@tutorialsdojo > curl http://169.254.169.254/latest/meta-data/public-ipv4  
12.18.98.110  
jonbonso@tutorialsdojo >
```

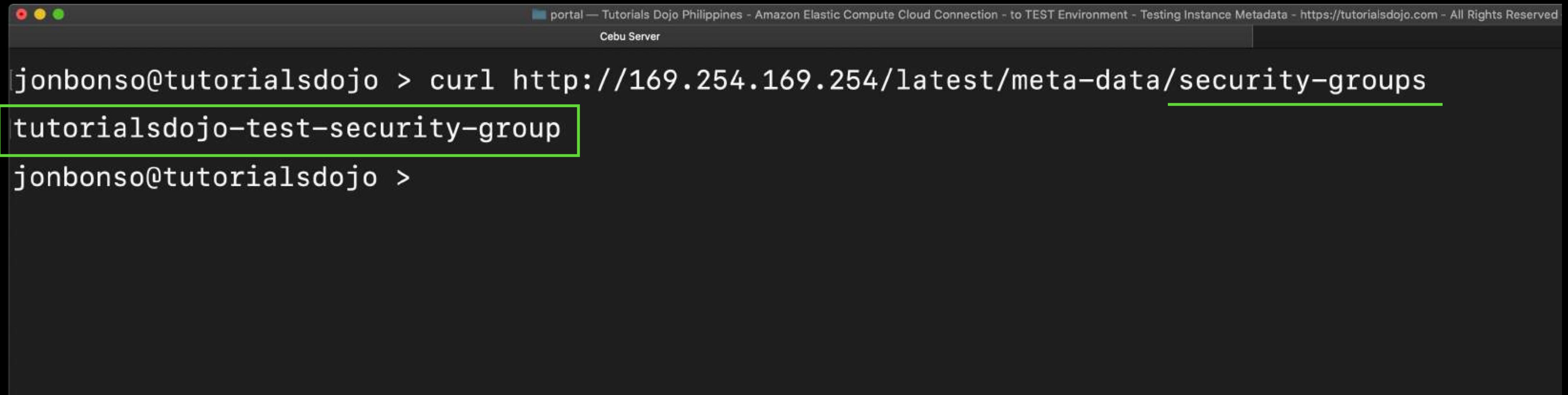
The screenshot shows a terminal window with a dark background and light-colored text. The title bar of the window reads "portal — Tutorials Dojo Philippines - Amazon Elastic Compute Cloud Connection - to TEST Environment - Testing instance Metadata - https://tutorialsdojo.com - Cebu Server". The terminal prompt is "jonbonso@tutorialsdojo >". The user runs the command "curl http://169.254.169.254/latest/meta-data/public-ipv4". The output of the command, "12.18.98.110", is highlighted with a green rectangular box. The terminal prompt appears again at the bottom.

## Media Access Control (MAC) Address

```
portal — Tutorials Dojo Philippines - Amazon Elastic Compute Cloud Connection - to TEST Environment - Testing Instance Metadata ·
Cebu Server

[jonbonso@tutorialsdojo > curl http://169.254.169.254/latest/meta-data/mac
02:4d:72:fc:21:b9
jonbonso@tutorialsdojo >
```

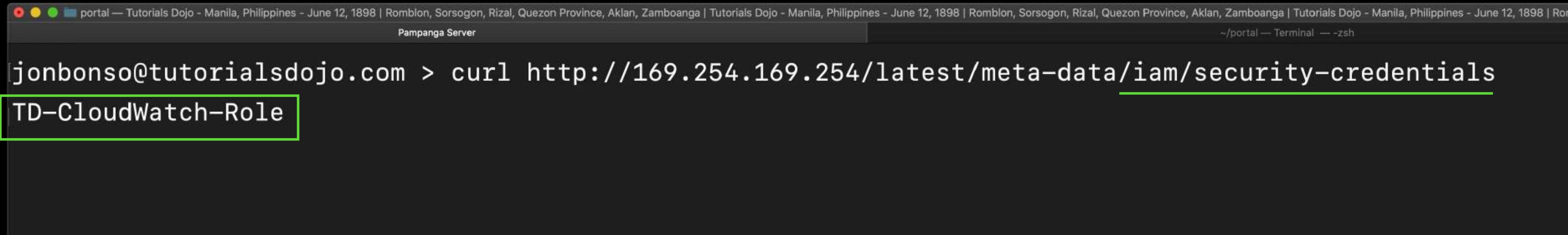
# Security Groups



A screenshot of a terminal window titled "Cebu Server". The window shows the command "curl http://169.254.169.254/latest/meta-data/security-groups" being run by user "jonbonso@tutorialsdojo". The output of the command, "tutorialsdojo-test-security-group", is highlighted with a green border. The terminal has a dark background with white text.

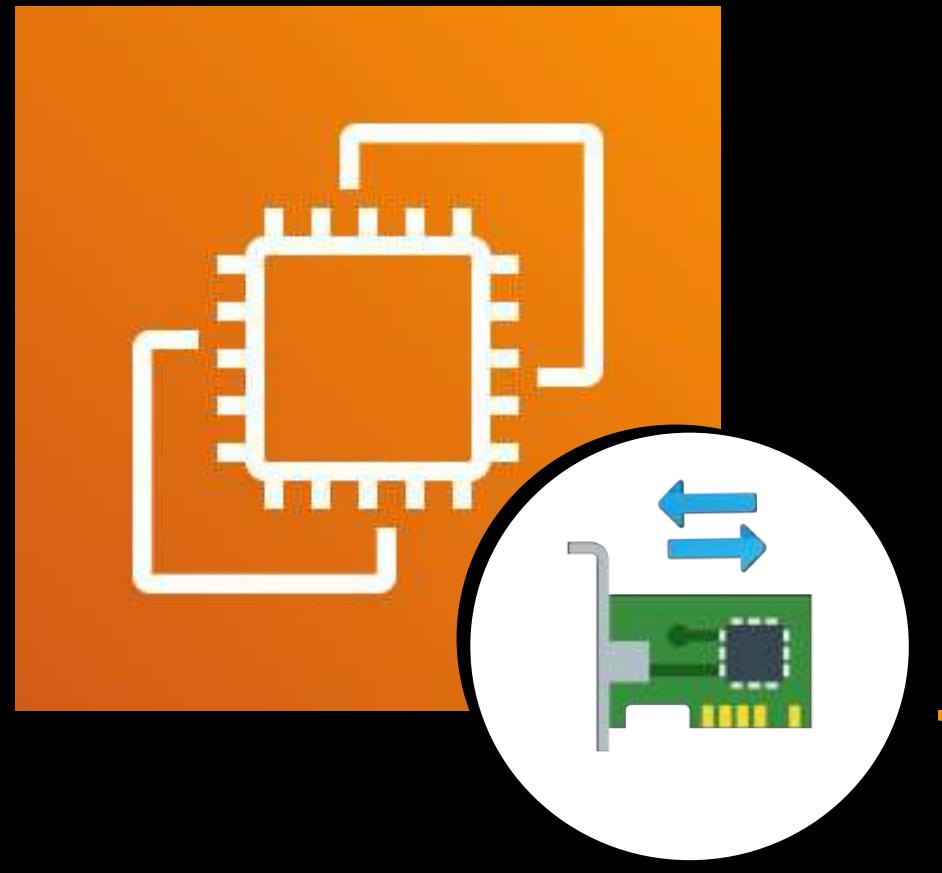
```
jonbonso@tutorialsdojo > curl http://169.254.169.254/latest/meta-data/security-groups
tutorialsdojo-test-security-group
jonbonso@tutorialsdojo >
```

# Instance Profile

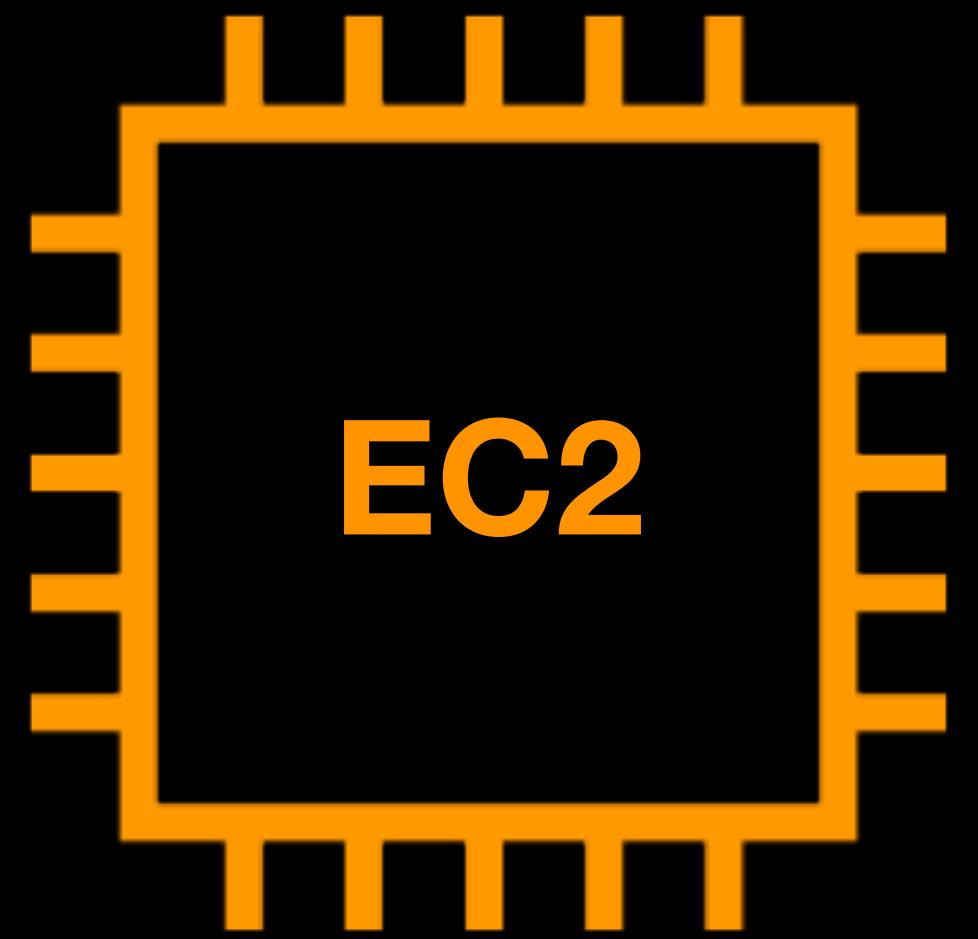


A screenshot of a terminal window titled "Pampanga Server". The window shows the command "curl http://169.254.169.254/latest/meta-data/iam/security-credentials TD-CloudWatch-Role" being run. The output of the command is highlighted with a green border and shows the role name "TD-CloudWatch-Role".

```
[jonbonso@tutorialsdojo.com > curl http://169.254.169.254/latest/meta-data/iam/security-credentials
TD-CloudWatch-Role]
```



# Amazon EC2 Networking

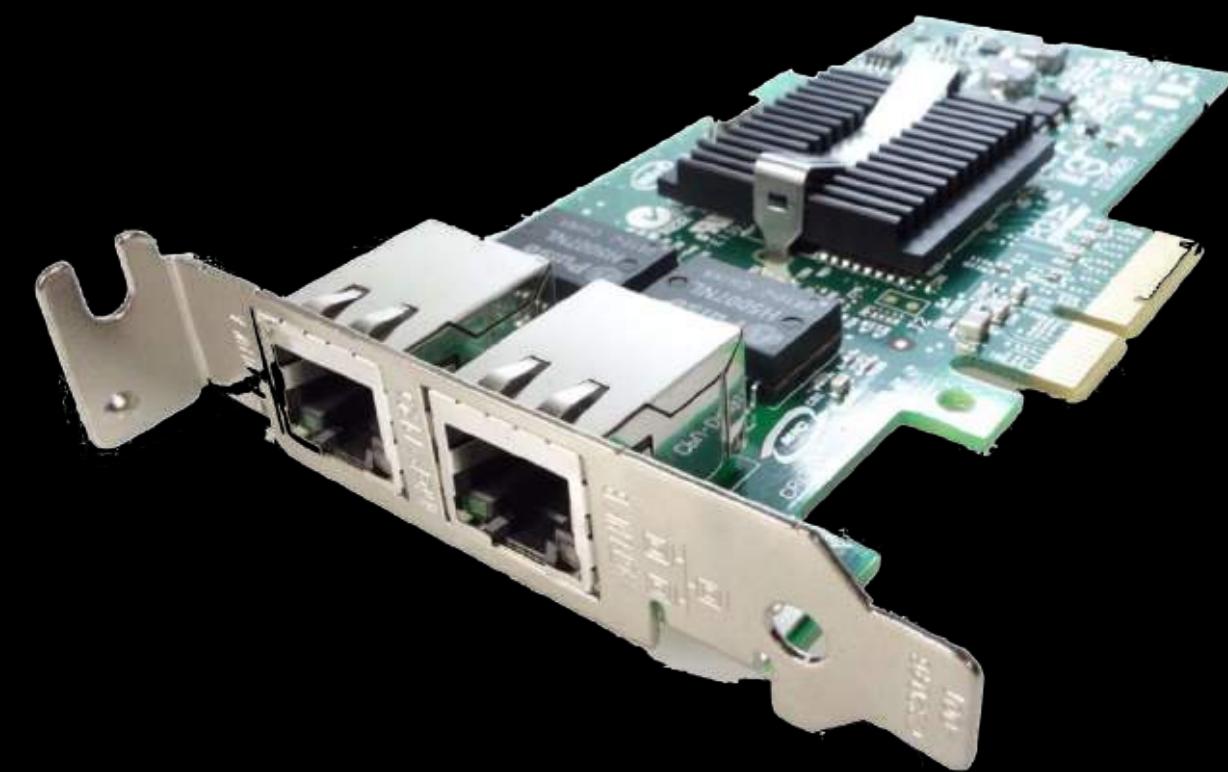
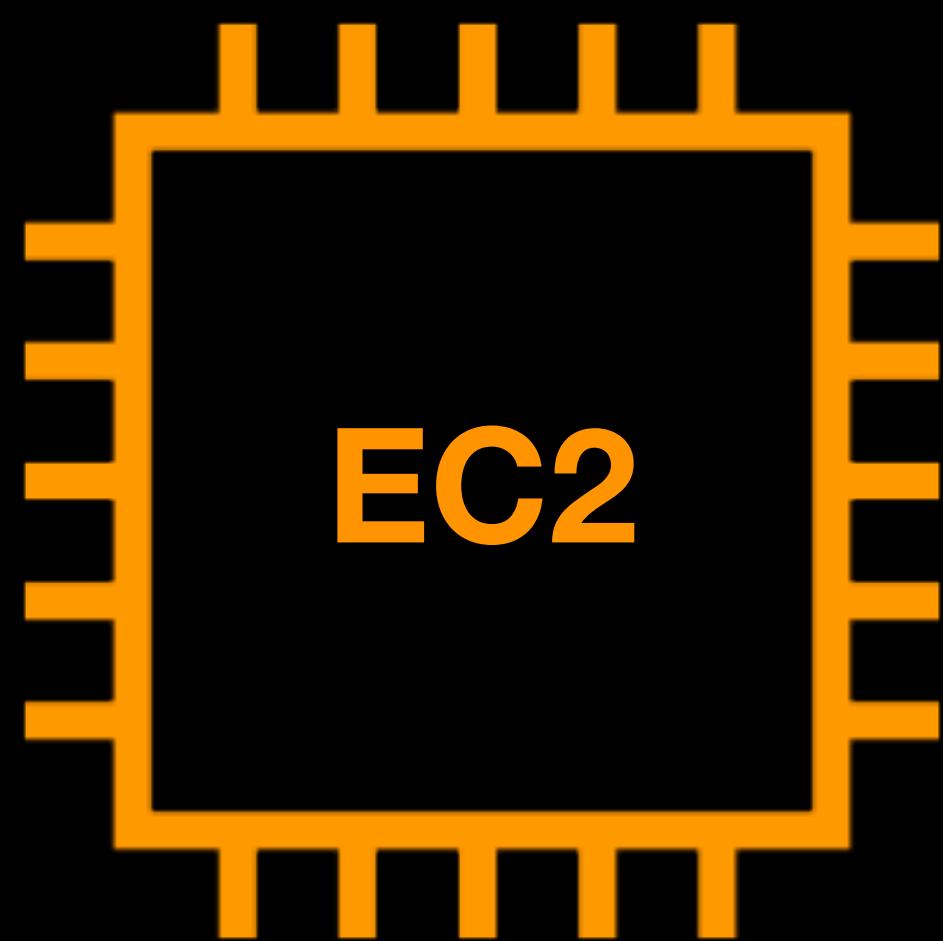


**PUBLIC INTERNET**

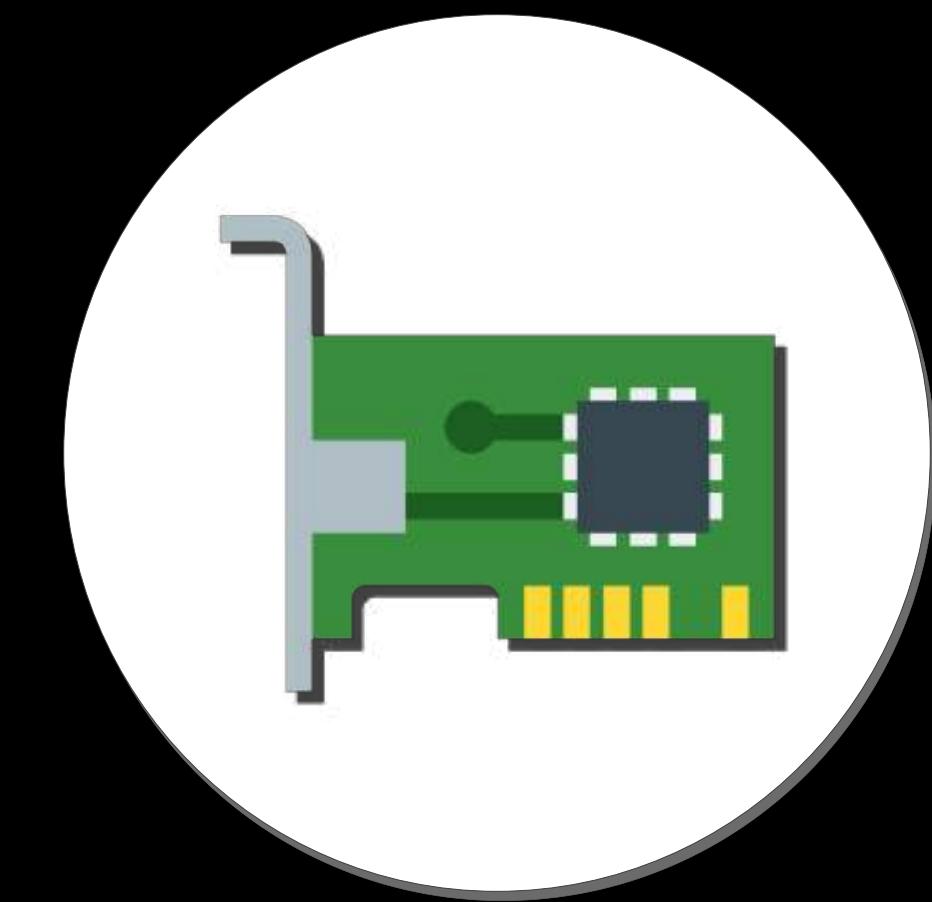
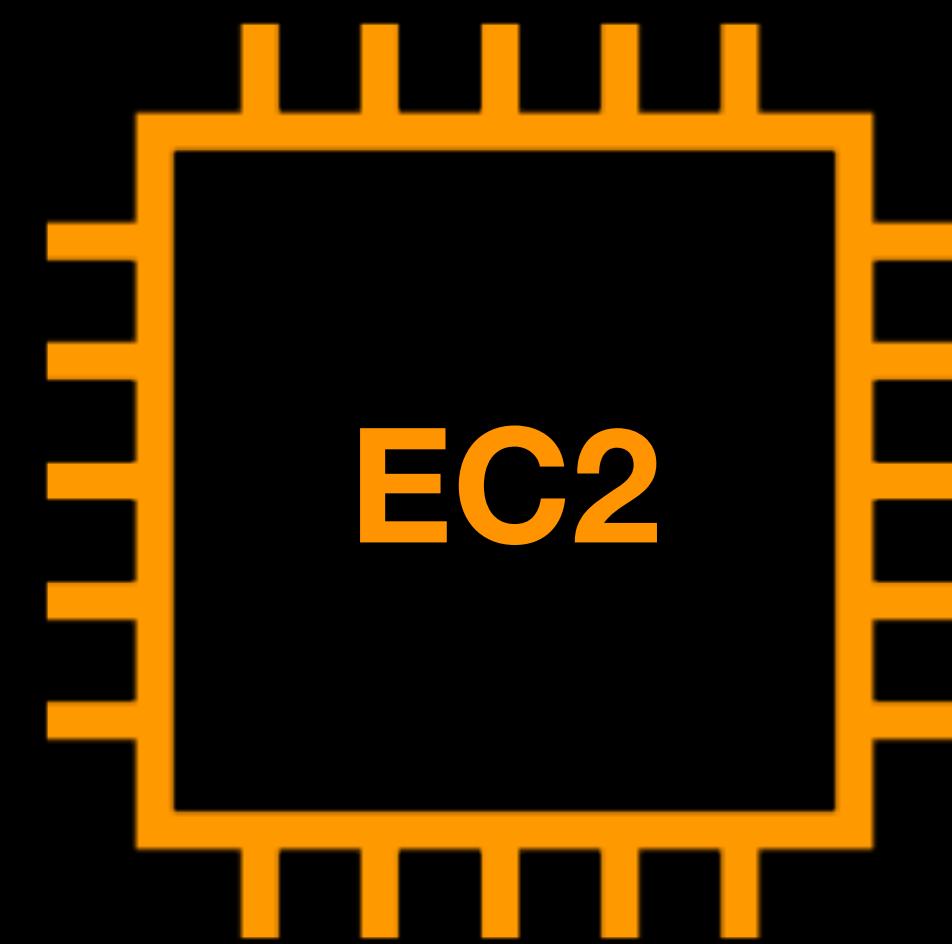


**PRIVATE NETWORK  
in AWS**

**Powered by Physical  
Networking Devices**

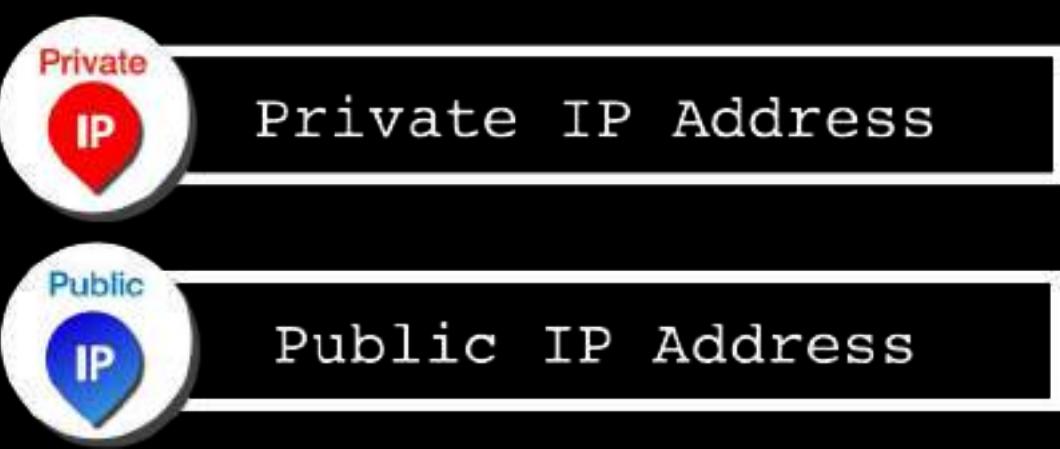


**VIRTUAL**  
Network Interface Card

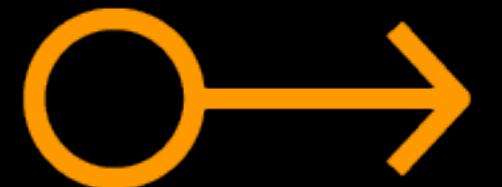


**PHYSICAL**  
Network Interface Card





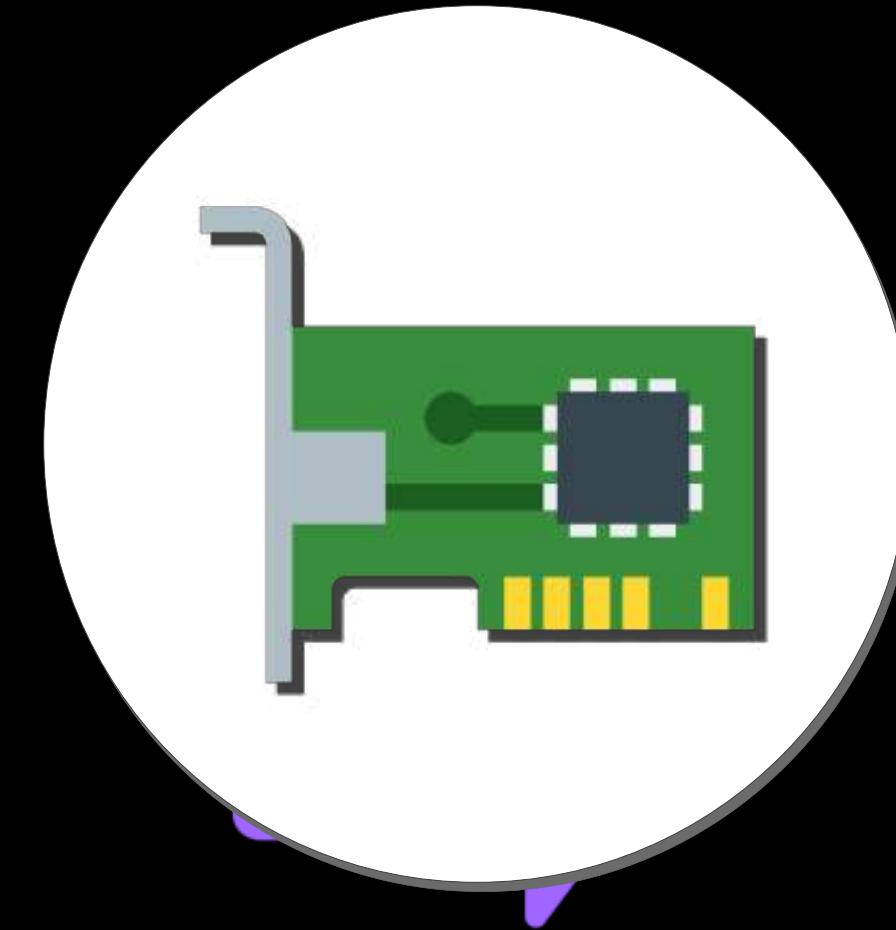
## IP Addressing



## Elastic IP Address

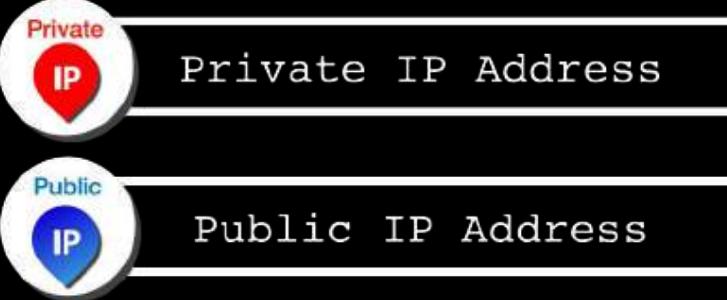


## Elastic Network Interface



## Enhanced Networking

## Elastic Fabric Adapter (EFA)

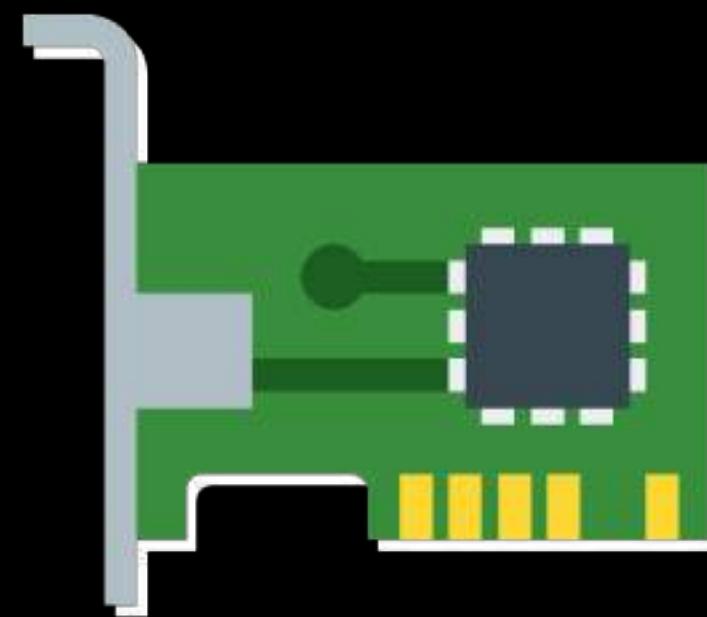


are powered by

## NETWORK INTERFACE CARD



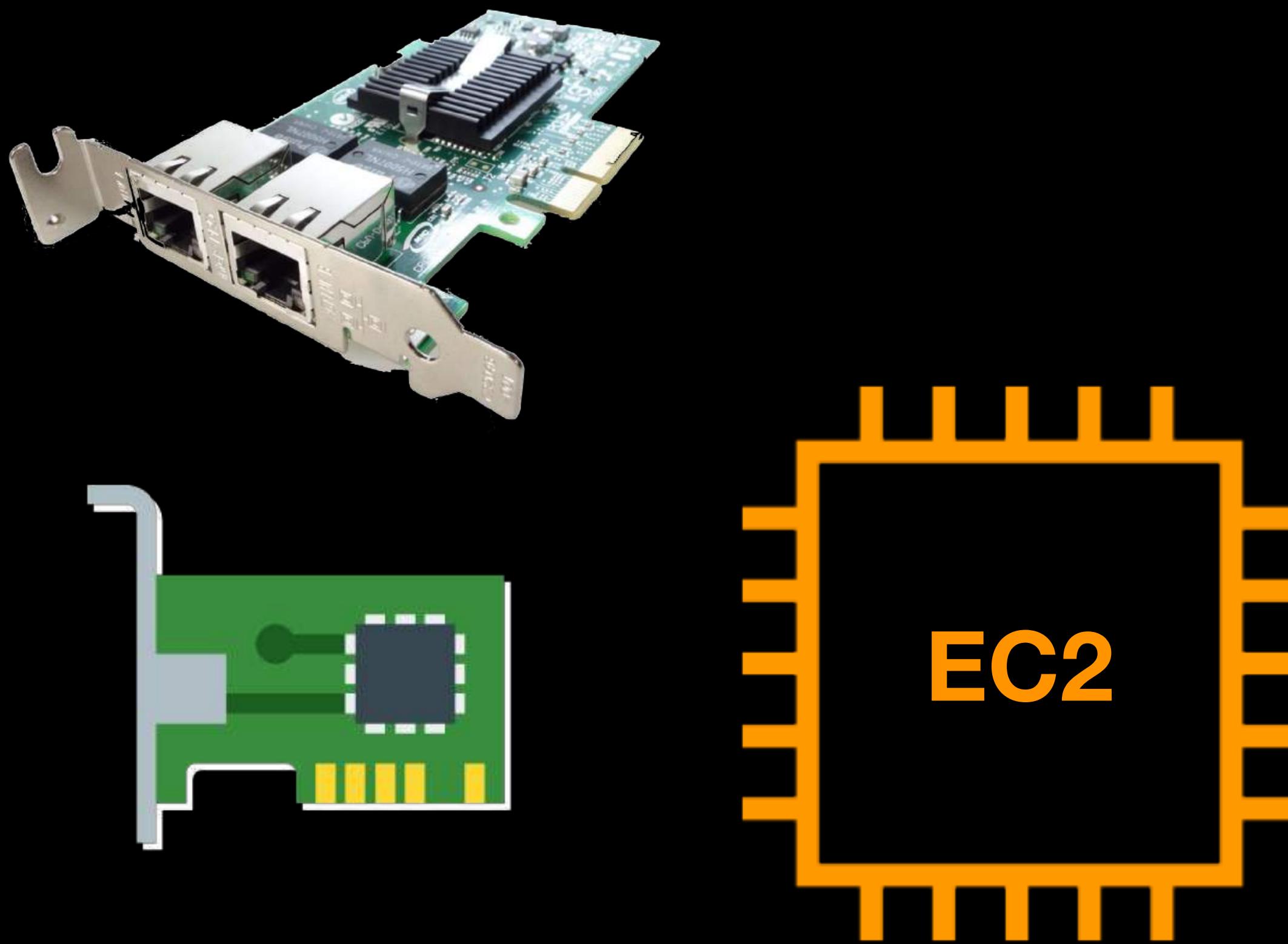
**PHYSICAL**



**VIRTUAL**



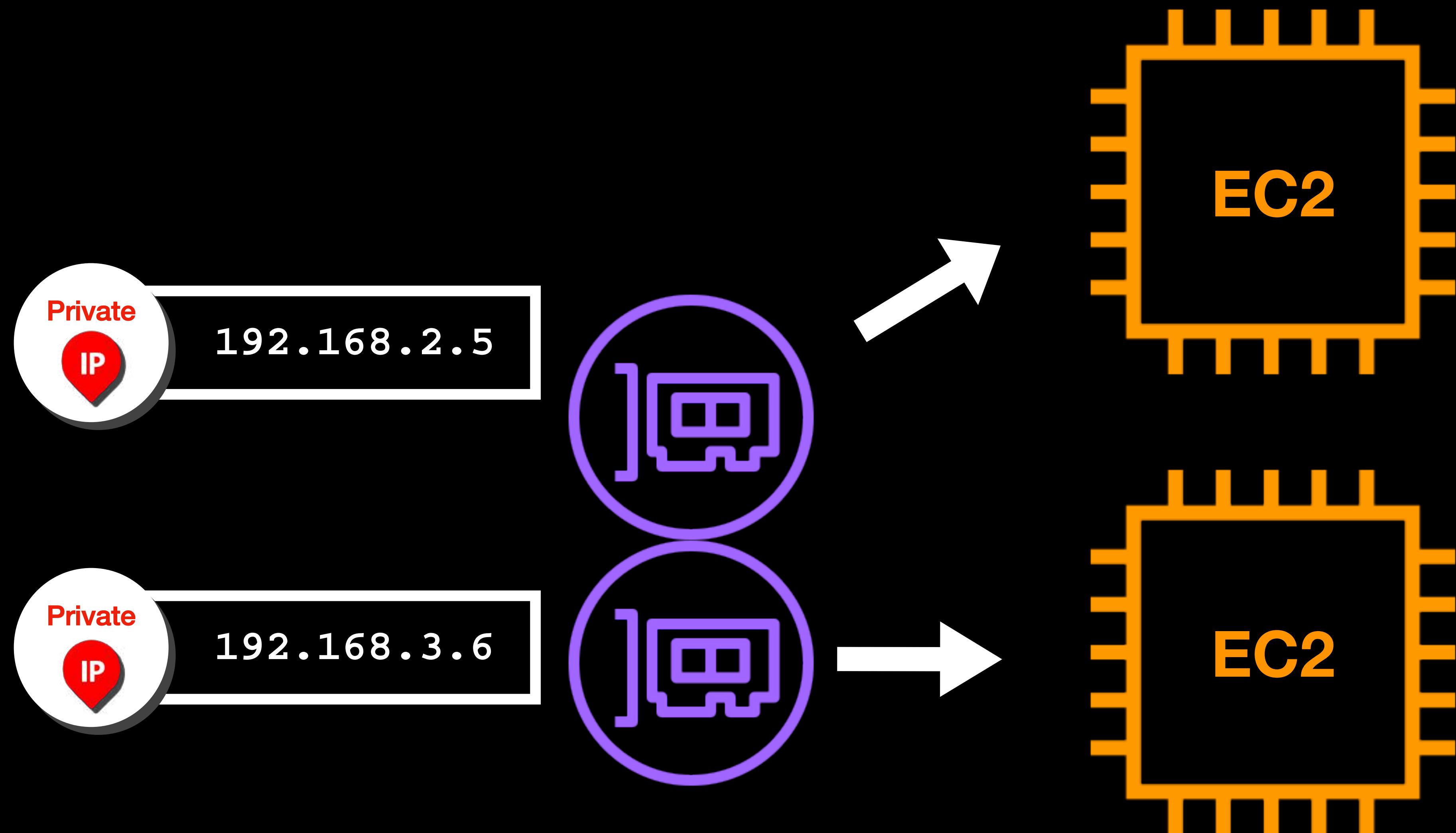
**Elastic Network Interface**





## Elastic Network Interface

- Primary private IPv4 address
- Secondary private IPv4 addresses
- One Elastic IP address per private IPv4 address
- One public IPv4 address
- One or more IPv6 addresses
- One or more security groups
- Media Access Control (MAC) address
- Source-Destination check flag
- Custom description



## CIDR



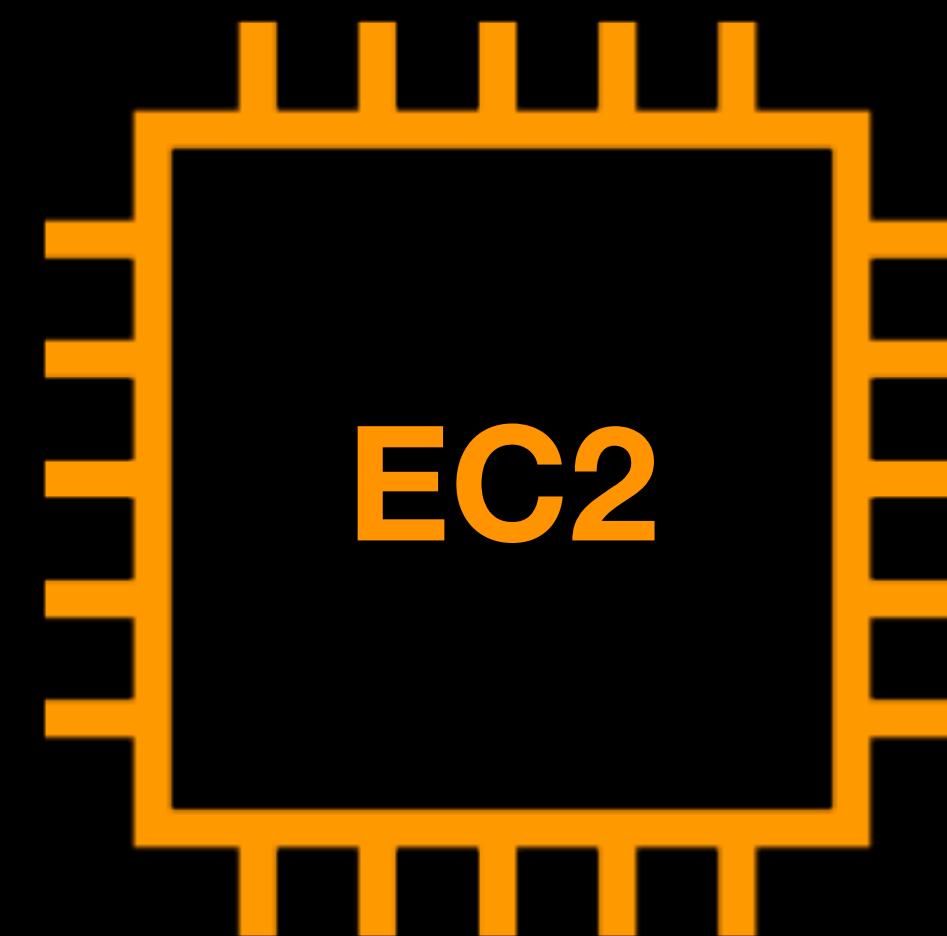
- Classless Inter-Domain Routing
- A method for allocating IP addresses
- Also used for IP Routing

IPv4 Address

IPv6 Address

Private  
IP

192.168.2.5



# Request For Comments 1918



# RFC 1918



Private IP Address



## Private IP Address

Class	IP Address Range	CIDR Block Prefix
Class A	10.0.0.0	/8
Class B	172.16.0.0	/12
Class C	192.168.0.0	/16



## Private IP Address

Class	IP Address Range	Total IP Address	CIDR Block Prefix
Class A <span>A</span>	<span>10.0.0.0</span> - <span>10.255.255.255</span>	Over 16 million	/8
Class B <span>B</span>	<span>172.16.0.0</span> - <span>172.31.255.255</span>	Over 1 million	/12
Class C <span>C</span>	<span>192.168.0.0</span> - <span>192.168.255.255</span>	Over 64,000	/16



## Private IP Address

10.0.0.0

-

10.255.255.255

172.16.0.0

-

172.31.255.255

192.168.0.0

-

192.168.255.255



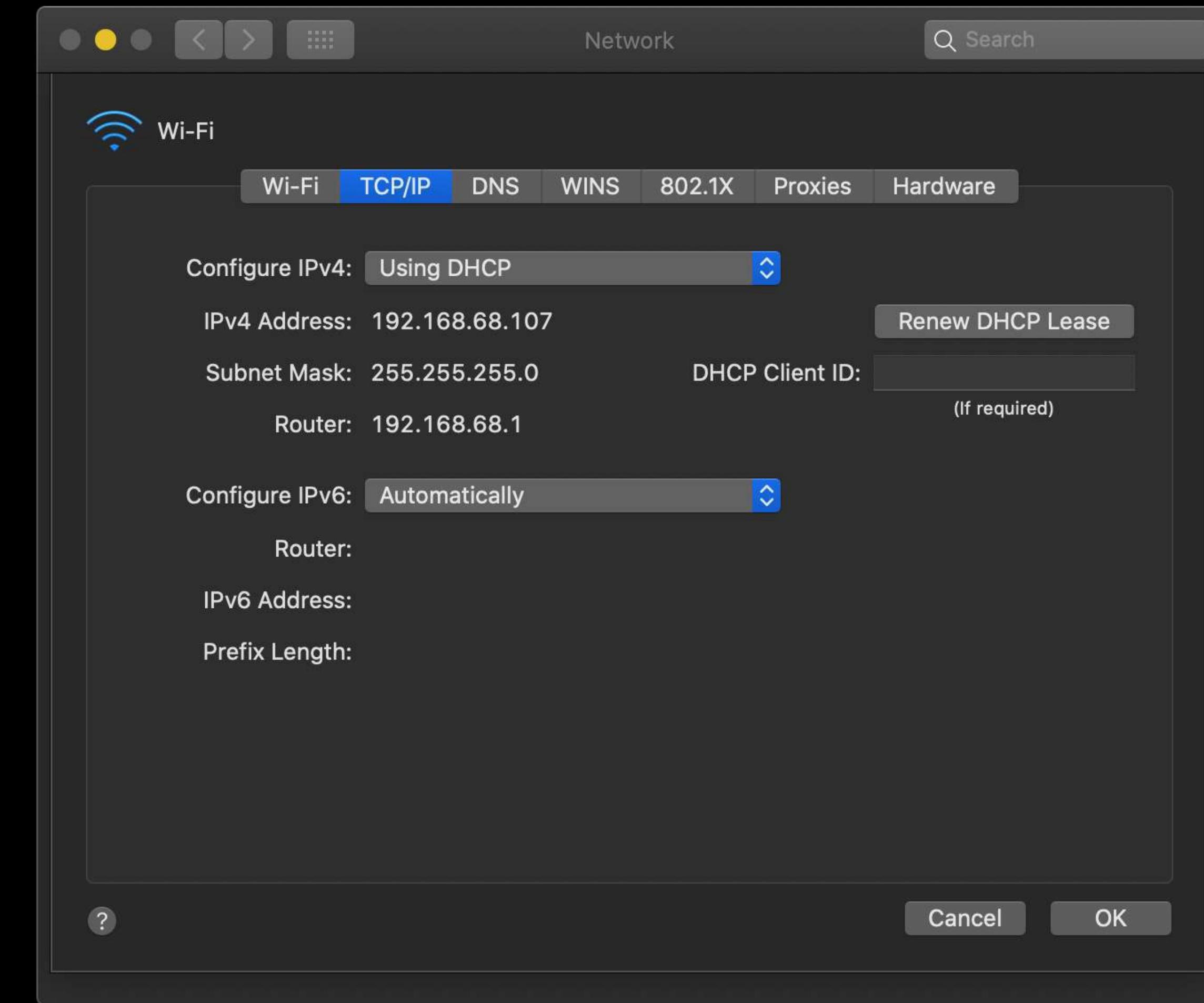


## Private IP Address

10.0.\*.\*

172.16.\*.\*

192.168.\*.\*





## Private IP Address

10.0.\*.\*

172.16.\*.\*

192.168.\*.\*

Network

Wi-Fi

Wi-Fi TCP/IP DNS WINS 802.1X

Configure IPv4: Using DHCP

IPv4 Address: 192.168.68.107

Subnet Mask: 255.255.255.0

Router: 192.168.68.1

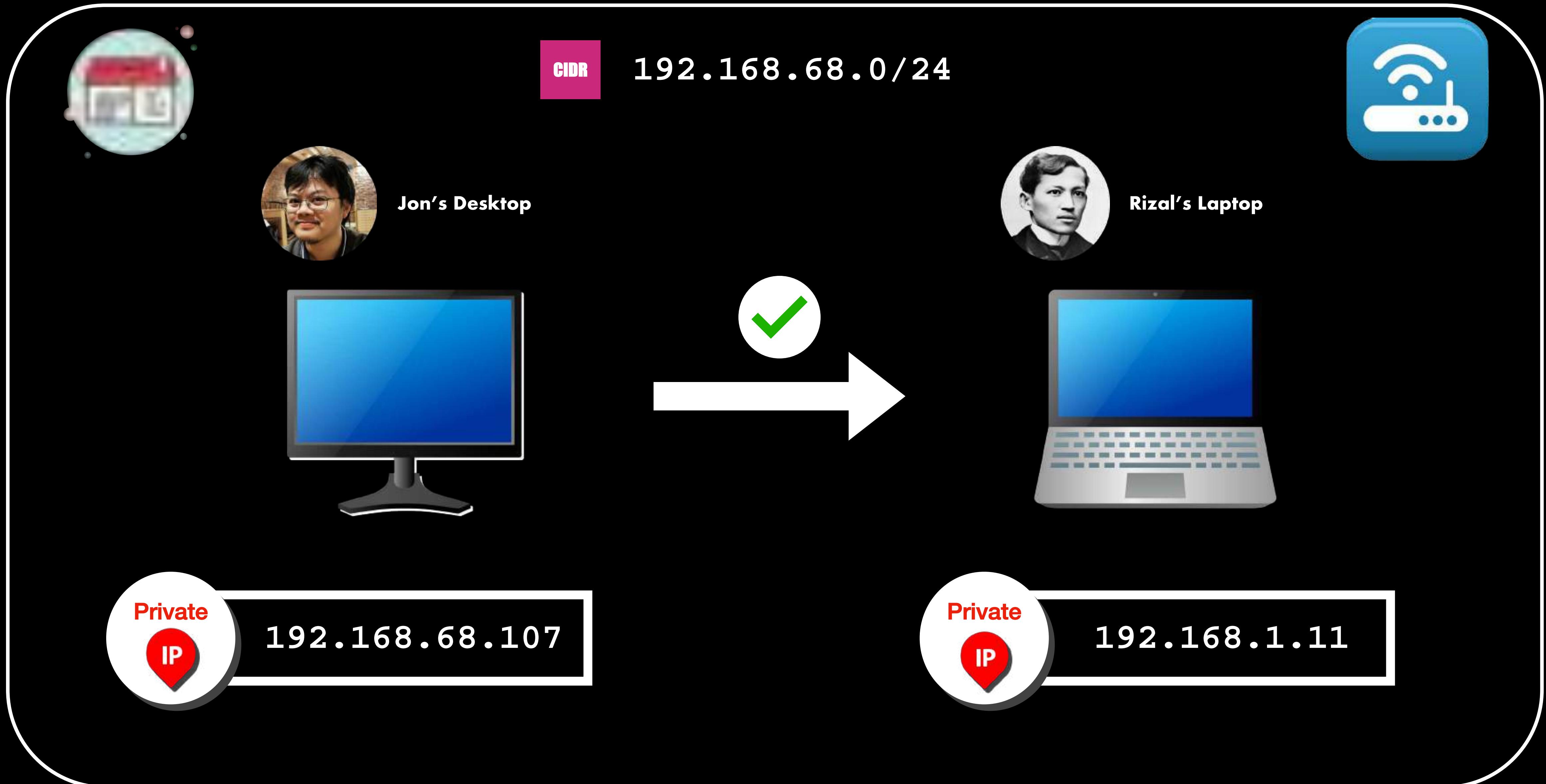
Configure IPv6: Automatically

Router:

IPv6 Address:

Prefix Length:

A screenshot of the Mac OS X Network preferences window. The window title is "Network". The "Wi-Fi" tab is selected. Under the "TCP/IP" tab, the IPv4 address is set to 192.168.68.107, which is highlighted with a green underline. The subnet mask is 255.255.255.0 and the router is 192.168.68.1. The IPv6 tab shows "Automatically" selected. The "Router:" field is empty. The "IPv6 Address:" and "Prefix Length:" fields are also empty.





**Internal DNS hostname**

**ip-10-251-50-12.ec2.internal**

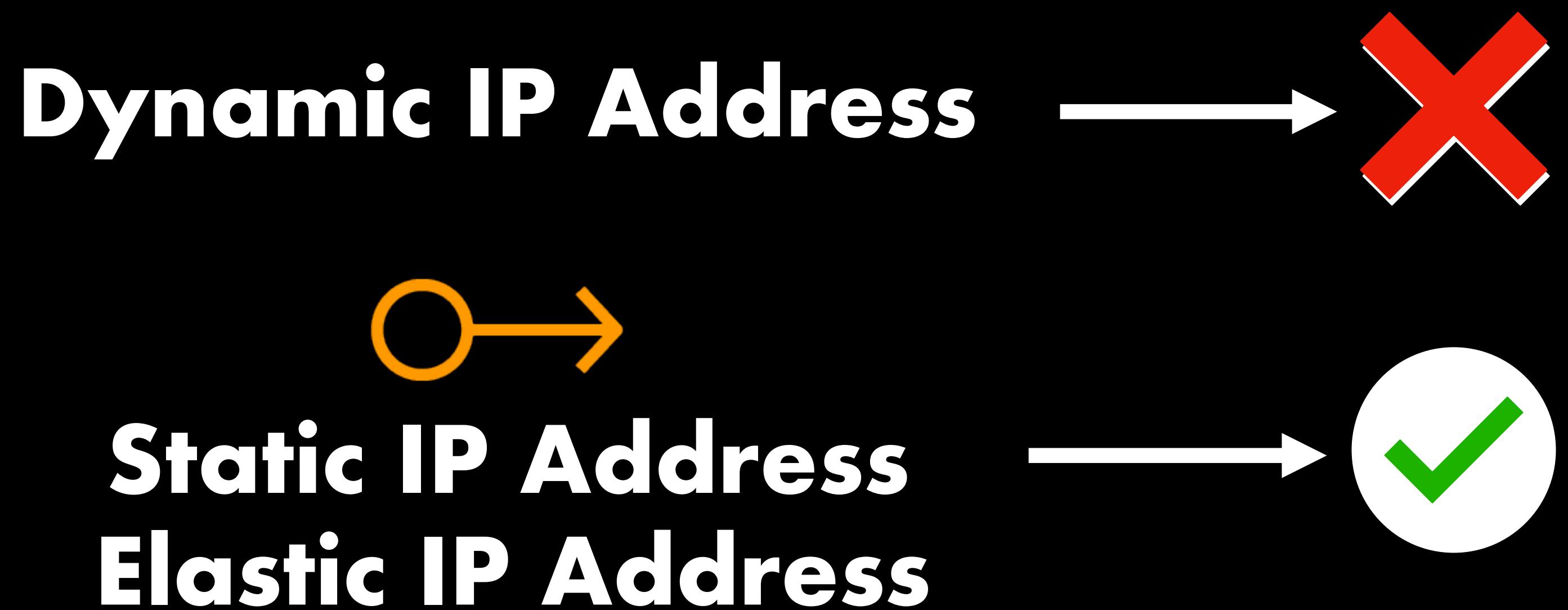


**Private IP Address**



**Public IP Address**





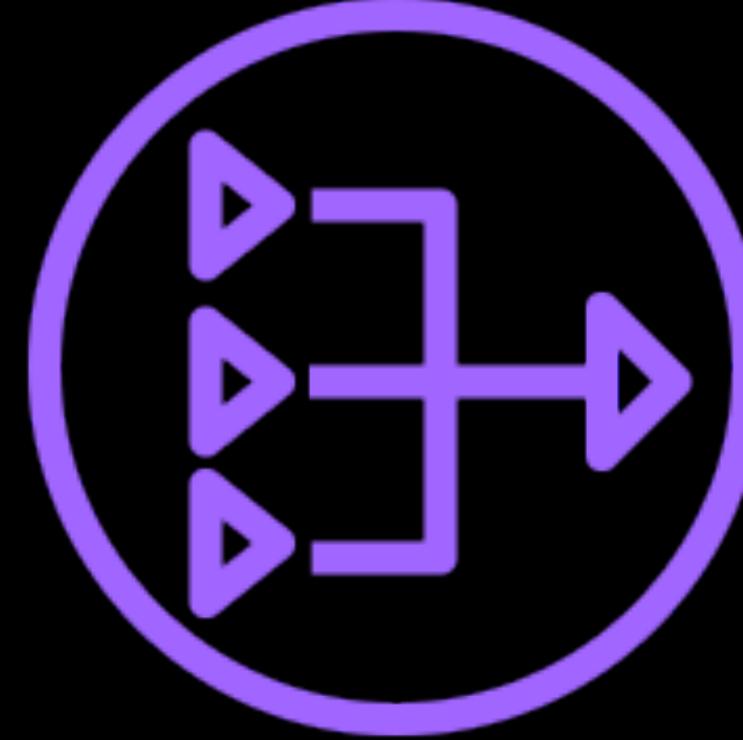


External DNS hostname

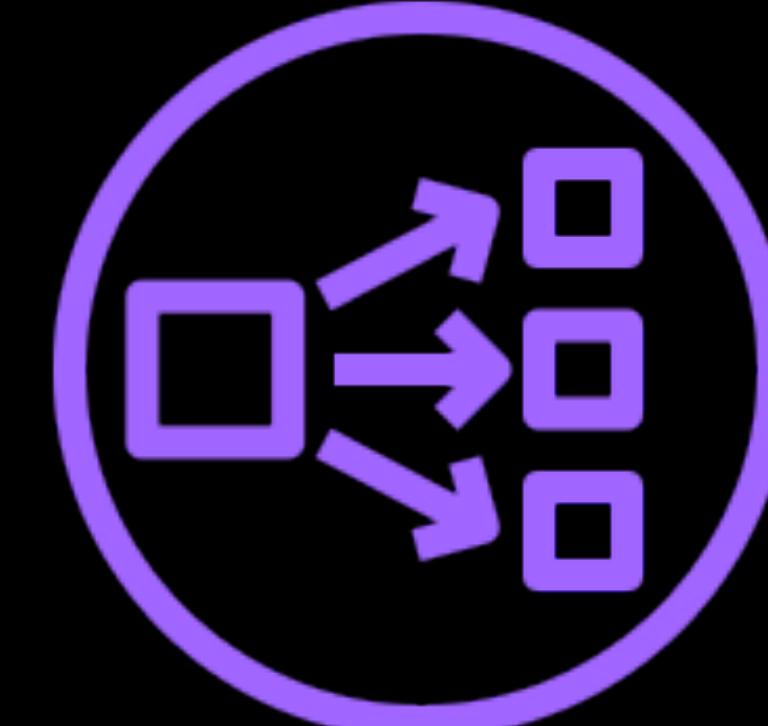
**ec2-136-158-28-50.compute-1.amazonaws.com**



**Elastic IP Address**



**NAT Gateway**

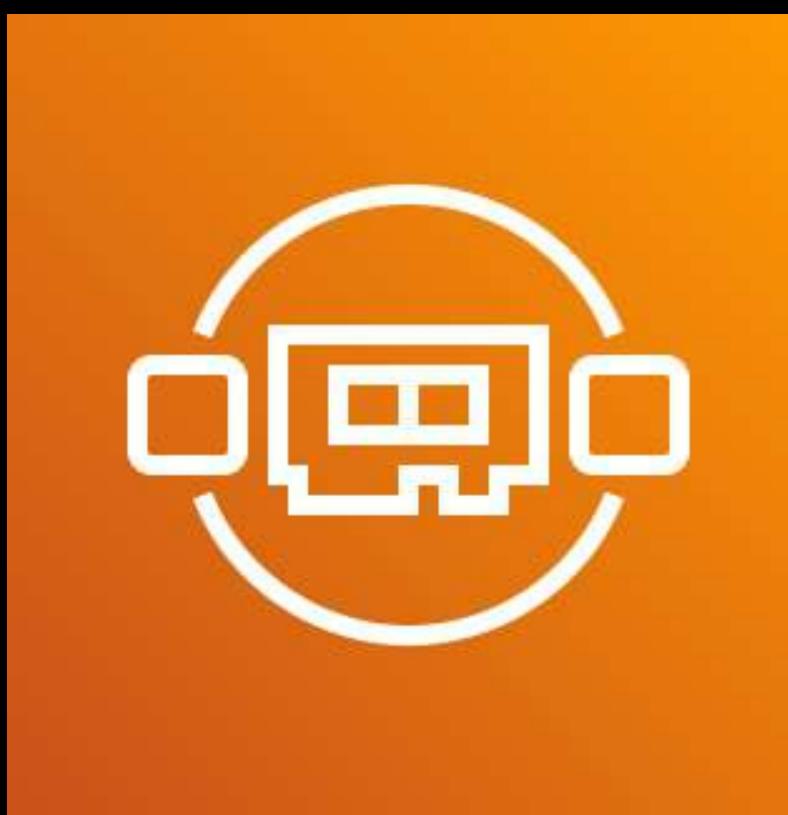


**Network Load Balancer**

**Features that **enhances** and accelerates the network capability  
of your EC2 instances:**



**Enhanced Networking**



**Elastic Fabric Adapter  
(EFA)**

## Enhanced Networking



- Based on the **network adapter drivers** of the underlying physical host
- The **network adapter drivers** can be:
  - Intel® Network Adapter Virtual Function Driver
  - AWS-built custom-based network adapter driver called **Elastic Network Adapter (ENA)**
  - Network drivers provided by AWS or other companies
  - Similar to the “**driver**” or the software package that allows your computer to access a printer or other physical computer devices



## Enhanced Networking

- **Uses single root I/O virtualization or SR-IOV**
- **Provides higher I/O performance and lower CPU utilization than the traditional virtualization techniques**
- **Controlled by network drivers (software)**
- **Provides:**
  - **Higher bandwidth**
  - **Consistent lower inter-instance latencies**
  - **Higher packet per second performance (PPS)**

# Network Drivers

intel

PRODUCTS SUPPORT SOLUTIONS DEVELOPERS PARTNERS

USA (ENGLISH)

Products Home > Drivers & Software

## Intel® Network Adapter Virtual Function Driver for Intel® 10 Gigabit Ethernet Network Connections

Version: 4.11.1 (Latest)

### Available Downloads

ixgbevf-4.11.1.tar.gz

Linux\*

Language: English

Size: 0.21 MB

MD5: ae4a7792028762c268de47728f7616bb

[Download](#)

### Other Versions

4.10.2  
4.9.3  
4.8.1  
4.7.1

### Detailed Description

#### Overview

This is the most current release of the ixgbevf driver for Linux\*, which supports kernel versions 5.11.2 through 5.11.2. It also has been tested on the following distributions:

- RHEL\* 7.9
- RHEL 8.3
- SLES\* 12sp5
- SLES 15sp2
- Ubuntu\* 19.04
- Ubuntu 20.04

#### Changes in this release:

- Added support for 5.11.2 kernel version

Note that while we attempt to keep the driver version number (4.11.1) in sync with its corresponding Linux kernel that has similar functionality this is far from authoritative. If you are using a different kernel or distro it is likely that its ixgbevf driver is at least as up to date as the out-of-tree driver.

Search or jump to... Pull requests Issues Marketplace Explore

amzn / amzn-drivers

Code Issues 13 Pull requests 1 Actions Projects Security Insights

master 4 branches 50 tags Go to file Add file Code

galpress linux/efa: Bump driver version to 1.12.0 ... b91fcfd3 8 days ago 347 commits

kernel linux/efa: Bump driver version to 1.12.0 8 days ago

userspace/dpdk Merge pull request #160 from Semihalf/master 4 months ago

README.md Revert "Update README.md" 5 months ago

About Official AWS drivers repository for Elastic Network Adapter (ENA) and Elastic Fabric Adapter (EFA)

README

Releases 50 efa\_linux\_1.12.0 Latest 8 days ago + 49 releases

Packages No packages published

Contributors 22 + 11 contributors

## AMAZON DRIVERS

Official repository of the open source drivers for devices used on AWS platforms.

The following drivers are included:

- Linux kernel driver for Elastic Network Adapter (ENA)
- FreeBSD kernel driver for Elastic Network Adapter (ENA)
- DPDK driver for ENA for critical fixes to previously released ENA DPDK drivers in DPDK official releases
- Linux kernel driver for Elastic Fabric Adapter (EFA)

For Linux driver SRPM build instruction available [here](#)



## Elastic Fabric Adapter (EFA)

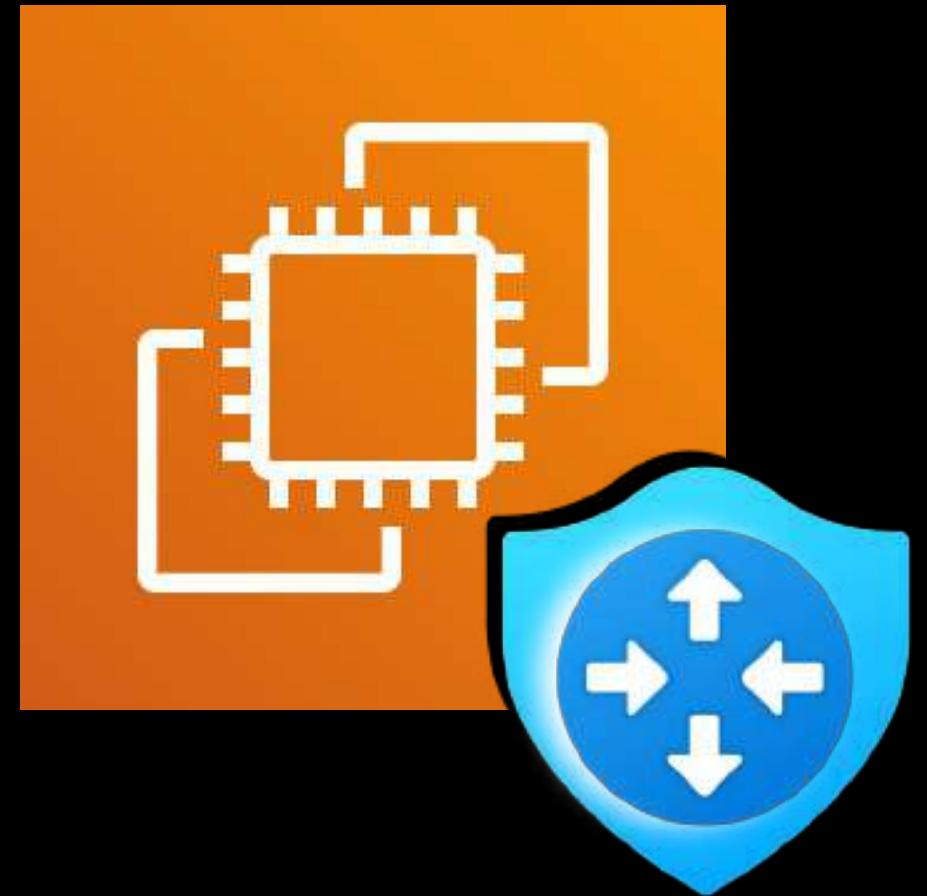
- Just like



Elastic Network Interface  
(ENI)

with additional capabilities

- Can directly communicate to the network interface hardware without passing through the Linux Kernel – also known as OS-Bypass
- Provides low-latency and reliable transport functionality to your virtual machines.
- Accelerates the networking capabilities of your High-Performance Computing or HPC workloads
- Enhances inter-instance communication



# Amazon EC2 Network Security

---



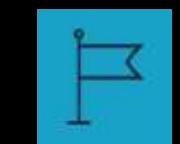
**Security Groups**



**Network Access Control List  
(Network ACL)**



AWS  
Cloud



N. Virginia Region



Network ACL

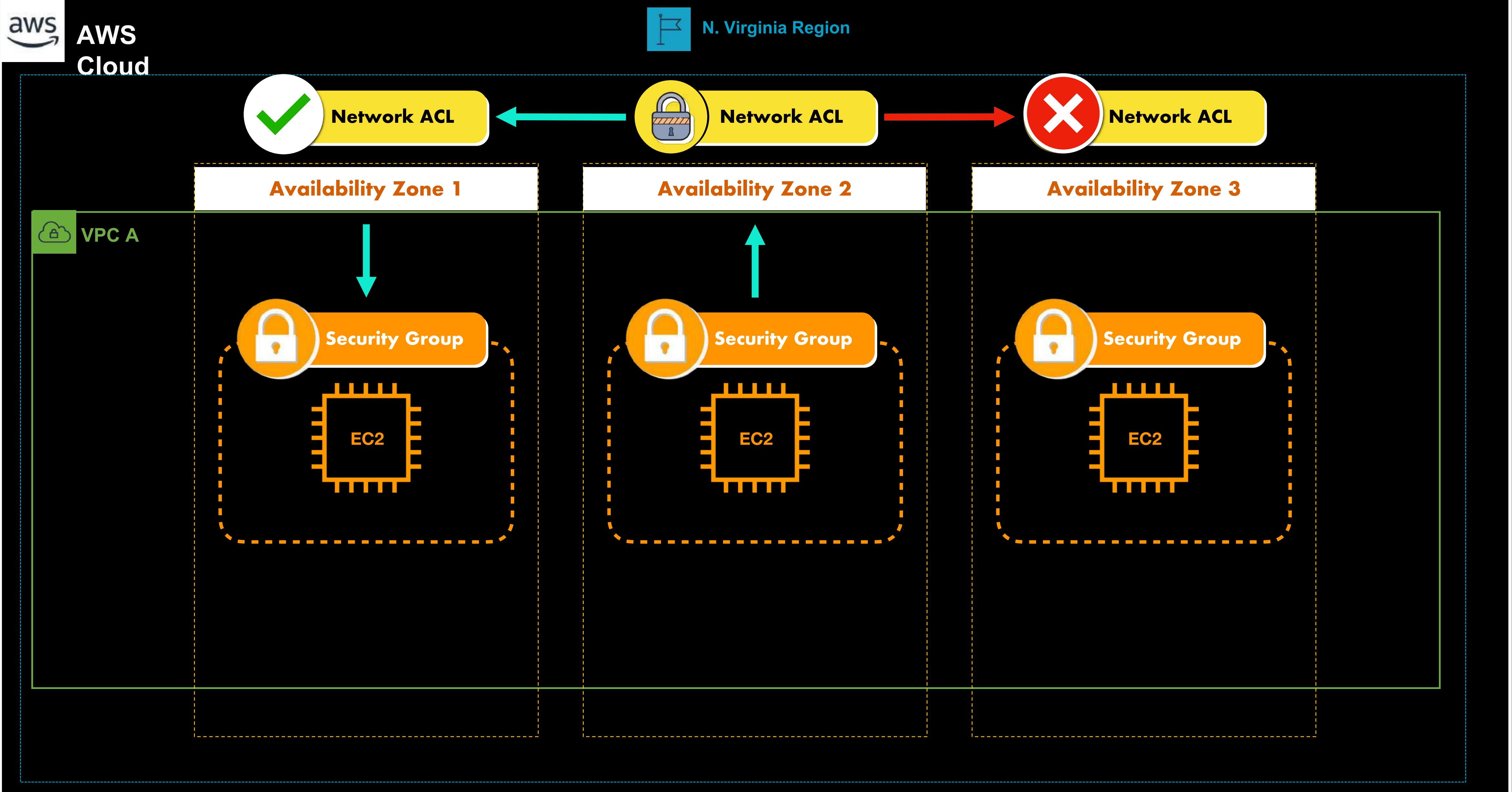
SUBNET

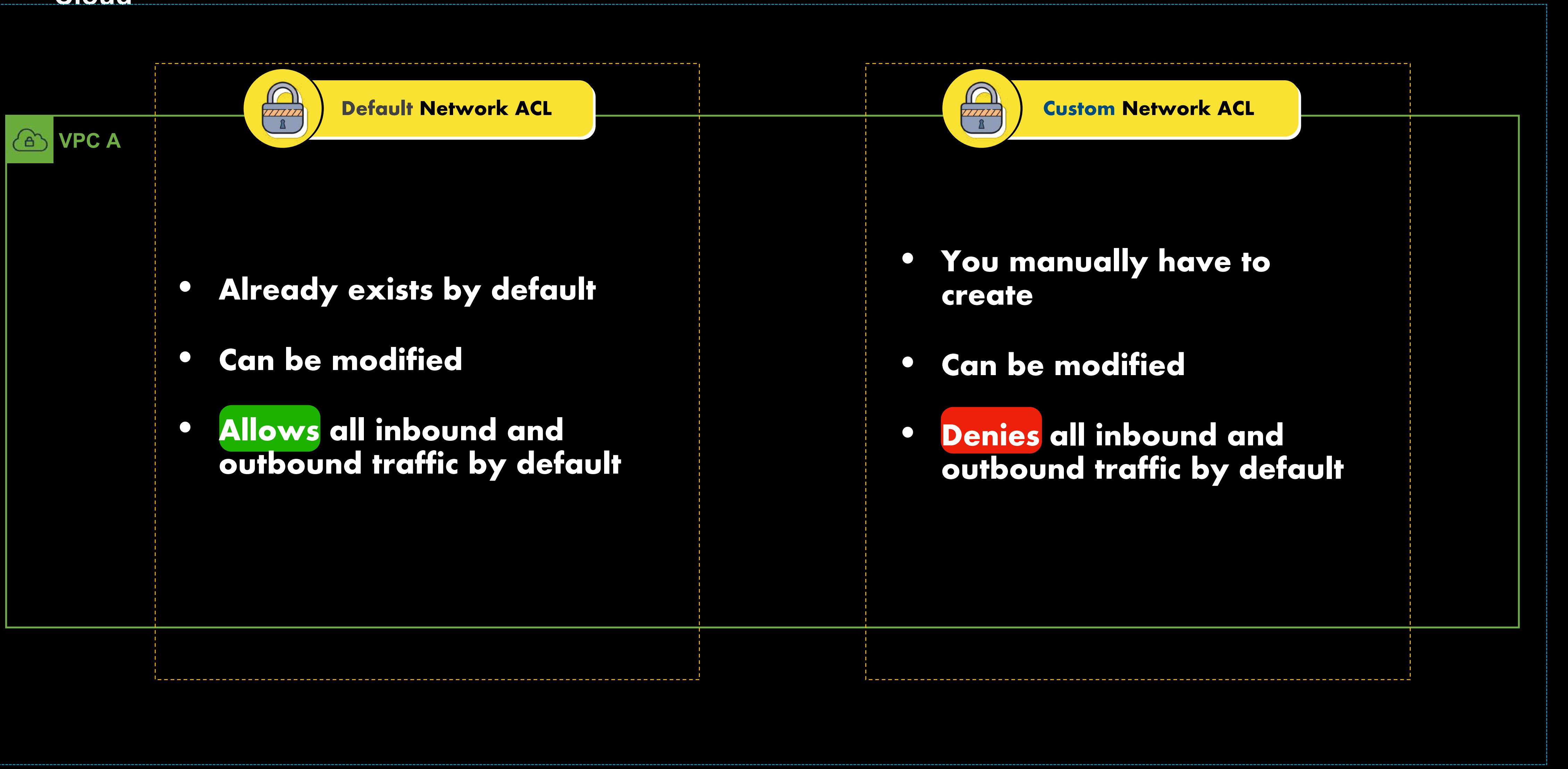
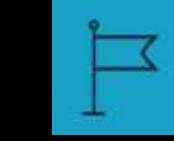


VPC A



Security Group

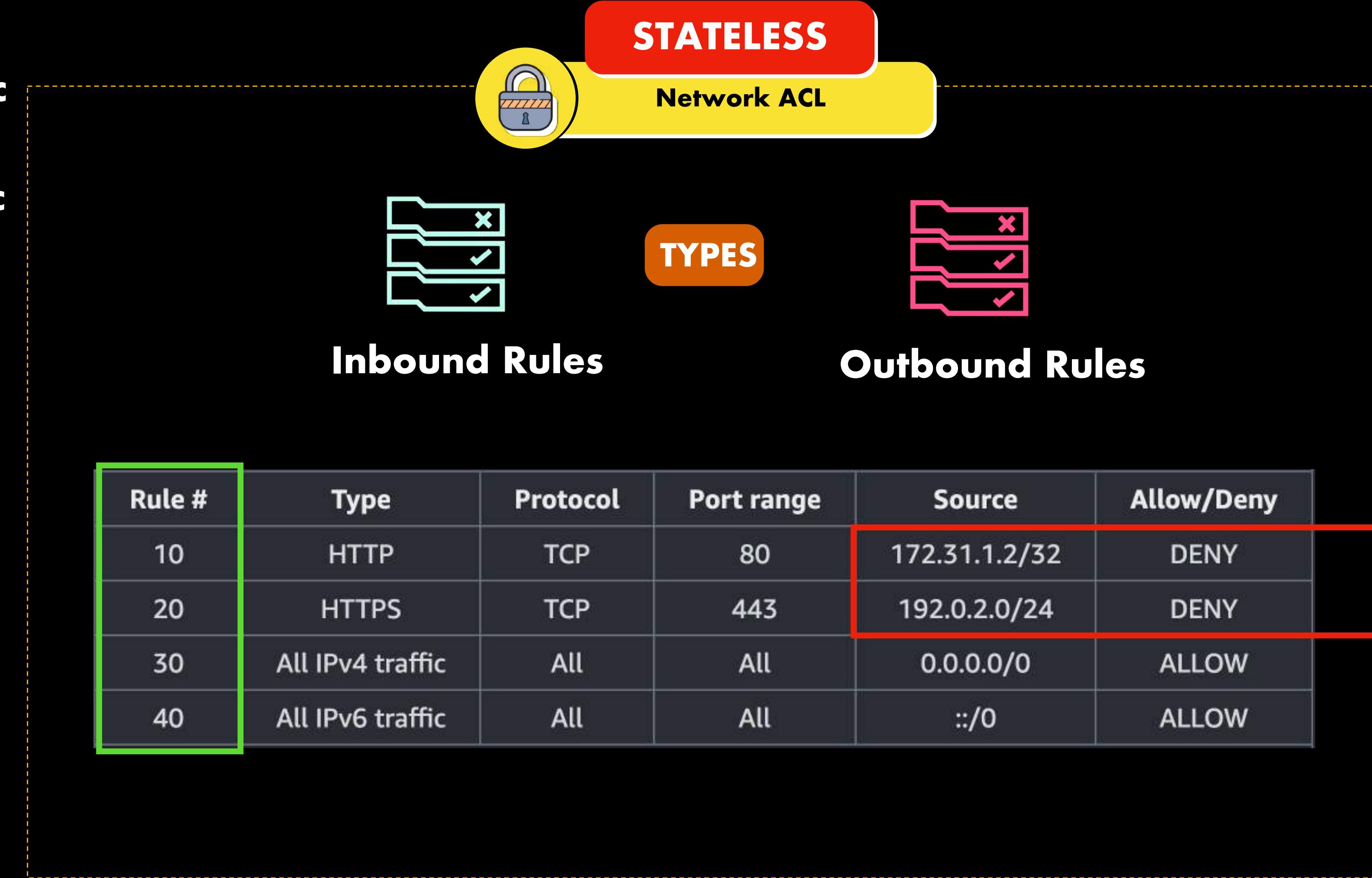




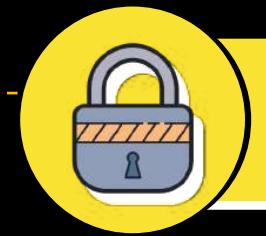
You can:

State

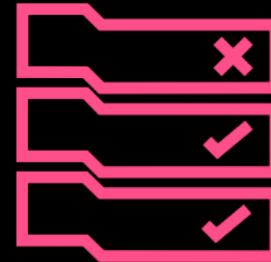
- **Allow** Traffic
- **Deny** Traffic



- An address prefix of /32 denotes a single IP address
- The /24 denotes the CIDR block which contains 256 different IP addresses



## Network ACL



## Outbound Rules

### Ephemeral Ports

- **Short-lived port numbers**
- **The range varies depending on the Operating System**

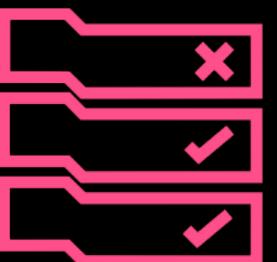


## Network ACL



### Inbound Rules

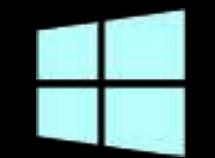
Type	Protocol	Port range
HTTP	TCP	80



### Outbound Rules



- **32768 – 61000**
- **49152 – 65535**
- **1024 – 65535**



- **Short-lived port numbers**
- **The range varies depending on the Operating System**



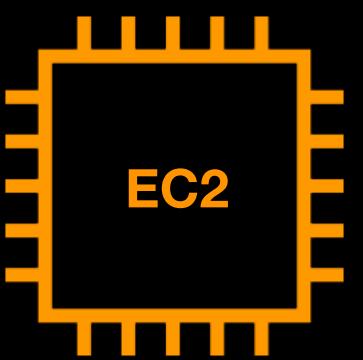
Network ACL



Inbound Rules



Outbound Rules





## Security Groups



## Security Groups

- A **virtual firewall** that controls the incoming and outgoing traffic of one or more EC2 instances
- 1 EC2 instance can have **one or more security groups**
- Cannot have an explicit DENY Rule (unlike Network ACL)
- Aside from EC2 Instances, it can also be attached to Amazon RDS, Amazon ElastiCache and other AWS resources

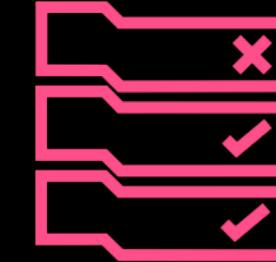


## Security Groups

- Allows incoming traffic
- Can't explicitly DENY traffic
- Not affected by Outbound Rules
- Allows outgoing traffic
- Controls traffic originated from the EC2 instance itself
- Does not affect the outgoing response traffic
- Examples:
  - EC2-initiated API call
  - Scheduled OS Patches



Inbound Rules



Outbound Rules

# 7 Open Systems Interconnection (OSI) Model Layers

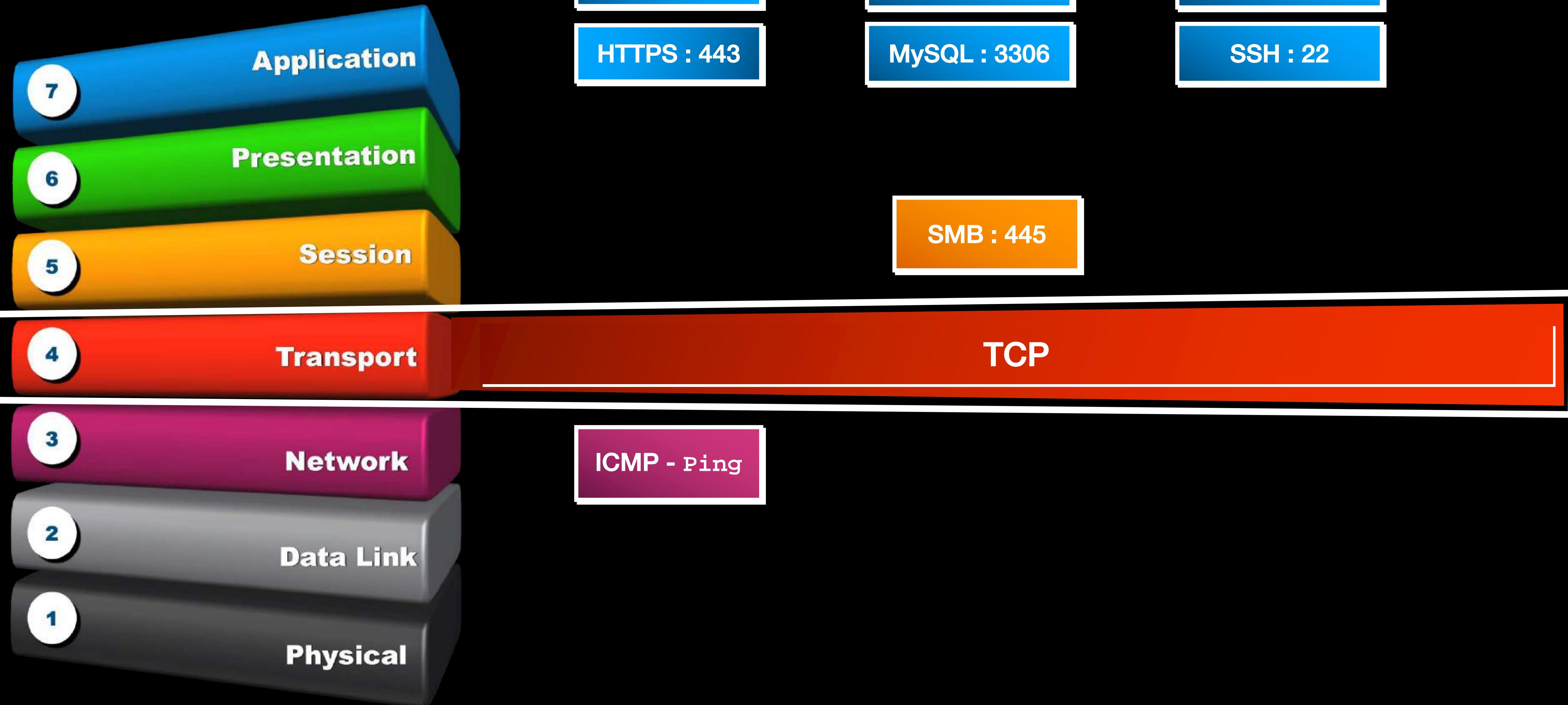


Source	Protocol	Port range	Description
0.0.0.0/0	TCP	80	Allow inbound HTTP access from all IPv4 addresses
::/0	TCP	80	Allow inbound HTTP access from all IPv6 addresses
0.0.0.0/0	TCP	443	Allow inbound HTTPS access from all IPv4 addresses
::/0	TCP	443	Allow inbound HTTPS access from all IPv6 addresses
Your network's public IPv4 address range	TCP	22	Allow inbound SSH access to Linux instances from IPv4 IP addresses in your network (over the internet gateway)
Your network's public IPv4 address range	TCP	3389	Allow inbound RDP access to Windows instances from IPv4 IP addresses in your network (over the internet gateway)

**TCP**

**UDP**

# 7 Open Systems Interconnection (OSI) Model Layers



Inbound rules <small>Info</small>																					
Type <small>Info</small>	Protocol <small>Info</small>	Port range <small>Info</small>	Source <small>Info</small>																		
All traffic	All	All	My IP ▾																		
RDP	TCP	3389	Custom ▾																		
SSH	TCP	22	Custom ▾																		
SMB	TCP	445	Custom ▾																		
<table><thead><tr><th colspan="2">Description - optional <small>Info</small></th></tr></thead><tbody><tr><td><input type="text"/> Jon Bonso's IP Address</td><td><button>Delete</button></td></tr><tr><td><input type="text"/> 6.12.18.98/32 <span>X</span></td><td></td></tr><tr><td><input type="text"/> CIDR Block</td><td><button>Delete</button></td></tr><tr><td><input type="text"/> 192.0.0.0/24 <span>X</span></td><td></td></tr><tr><td><input type="text"/> Prefix List</td><td><button>Delete</button></td></tr><tr><td><input type="text"/> pl-02cd2c6b <span>X</span></td><td></td></tr><tr><td><input type="text"/> Another Security Group</td><td><button>Delete</button></td></tr><tr><td><input type="text"/> sg-e1023de3 <span>X</span></td><td></td></tr></tbody></table>				Description - optional <small>Info</small>		<input type="text"/> Jon Bonso's IP Address	<button>Delete</button>	<input type="text"/> 6.12.18.98/32 <span>X</span>		<input type="text"/> CIDR Block	<button>Delete</button>	<input type="text"/> 192.0.0.0/24 <span>X</span>		<input type="text"/> Prefix List	<button>Delete</button>	<input type="text"/> pl-02cd2c6b <span>X</span>		<input type="text"/> Another Security Group	<button>Delete</button>	<input type="text"/> sg-e1023de3 <span>X</span>	
Description - optional <small>Info</small>																					
<input type="text"/> Jon Bonso's IP Address	<button>Delete</button>																				
<input type="text"/> 6.12.18.98/32 <span>X</span>																					
<input type="text"/> CIDR Block	<button>Delete</button>																				
<input type="text"/> 192.0.0.0/24 <span>X</span>																					
<input type="text"/> Prefix List	<button>Delete</button>																				
<input type="text"/> pl-02cd2c6b <span>X</span>																					
<input type="text"/> Another Security Group	<button>Delete</button>																				
<input type="text"/> sg-e1023de3 <span>X</span>																					

**Inbound rules** [Info](#)

Type <a href="#">Info</a>	Protocol <a href="#">Info</a>	Port range <a href="#">Info</a>	Source <a href="#">Info</a>	Description - optional <a href="#">Info</a>
All traffic ▾	All	All	<ul style="list-style-type: none"><li>Anywhere ▲</li><li>Custom</li><li><b>Anywhere</b></li><li>My IP</li></ul>	Allow Traffic from ANYWHERE <a href="#">Delete</a>

[Add rule](#)

# You can only

# Whitelisting

- Allow Traffic

Inbound rules <small>Info</small>				
Type <small>Info</small>	Protocol <small>Info</small>	Port range <small>Info</small>	Source <small>Info</small>	Description - optional <small>Info</small>
All traffic	All	All	My IP ▾	Jon Bonso's IP Address <span style="float: right;">Delete</span>
			6.12.18.98/32 <span style="float: right;">X</span>	
RDP	TCP	3389	Custom ▾	CIDR Block <span style="float: right;">Delete</span>
			192.0.0.0/24 <span style="float: right;">X</span>	
SSH	TCP	22	Custom ▾	Prefix List <span style="float: right;">Delete</span>
			pl-02cd2c6b <span style="float: right;">X</span>	
SMB	TCP	445	Custom ▾	Another Security Group <span style="float: right;">Delete</span>
			sg-e1023de3 <span style="float: right;">X</span>	



### Default Security Group

- **Already exists on your default VPC**
- **Has one inbound rule and one outbound rule by default**
- **Will be attached to your EC2 instance if you didn't specify a particular security group**
- **Automatically allows incoming traffic from any resource that also uses the default security group**
- **Allows all outgoing traffic that originated from the instance itself**



### Custom Security Group

- **You manually have to create**
- **Has a default outbound rule that allows all traffic**
- **Doesn't have a default inbound rule**
- **Denies all inbound and outbound traffic by default**

# You can only

- Allow Traffic



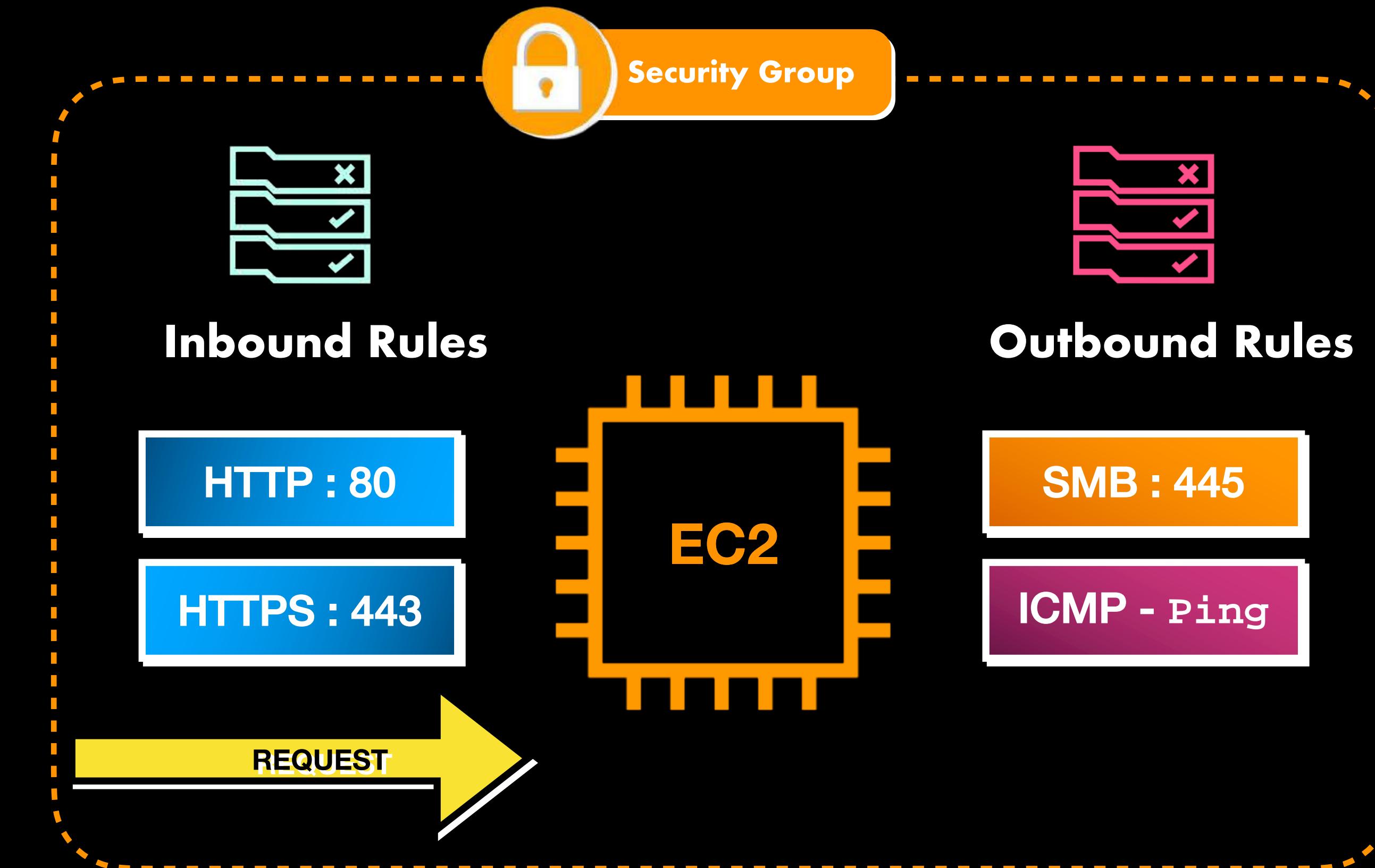
## Security Groups

```
root@ip-172-31-3-100:~# iptables -nL INPUT --line-numbers
Chain INPUT (policy ACCEPT)
num  target     prot opt source          destination
1    DROP       all  --  77.126.68.20    0.0.0.0/0
2    DROP       all  --  0.0.0.0/0      0.0.0.0/0      match-set tor src
root@ip-172-31-3-100:~#
```

## Security Groups

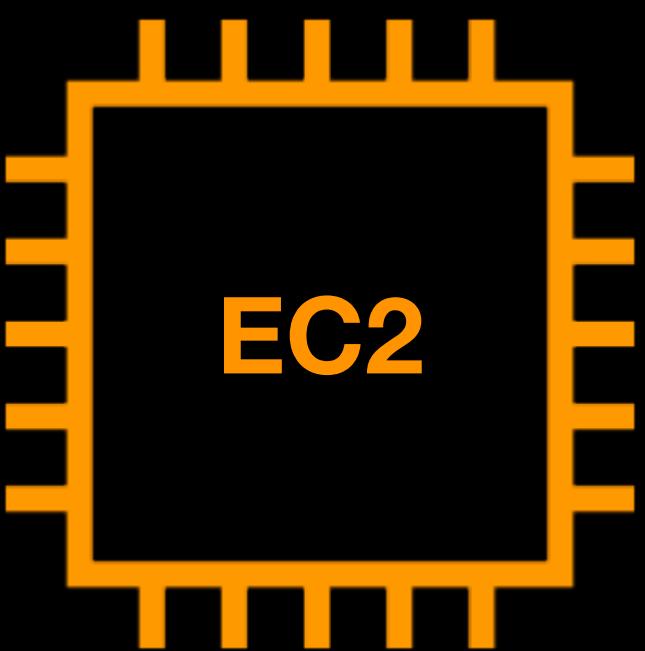


**STATEFUL**





**Security Groups**



**Amazon EC2**



**Amazon Aurora**



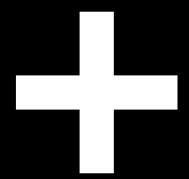
**Amazon RDS**



**Amazon ElastiCache**



## Security Groups



## Network Access Control List (Network ACL)

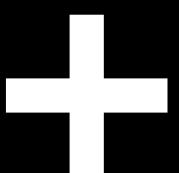


AWS Web Application Firewall (AWS WAF)

- You can't apply a security group or network ACL to your Amazon S3 buckets
- Both of these features do not provide enough protection against Cross-Site Scripting or SQL Injection attacks
- These two are also inefficient in geographic match conditions or blocking certain countries



## Security Groups



## Network Access Control List (Network ACL)

### Selected resources Info

Name	Resource ID	State
Tutorials Dojo Manila VPC	vpc-4ba22d36	Available

### Flow log settings

#### Name - optional

#### Filter

The type of traffic to capture (accepted traffic only, rejected traffic only, or all traffic).

- Accept
- Reject
- All

#### Maximum aggregation interval Info

The maximum interval of time during which a flow of packets is captured and aggregated into a flow log record.

- 10 minutes
- 1 minute

#### Destination

The destination to which to publish the flow log data.

- Send to CloudWatch Logs
- Send to an Amazon S3 bucket

#### Destination log group Info

The name of the Amazon CloudWatch log group to which the flow log is published. A new log stream is created for each monitored network interface.

▼
C

#### IAM role Info

The IAM role that has permission to publish to the Amazon CloudWatch log group.

▼
C

The IAM role must have permission to publish to the CloudWatch log group. [Set up permissions](#)

#### Log record format

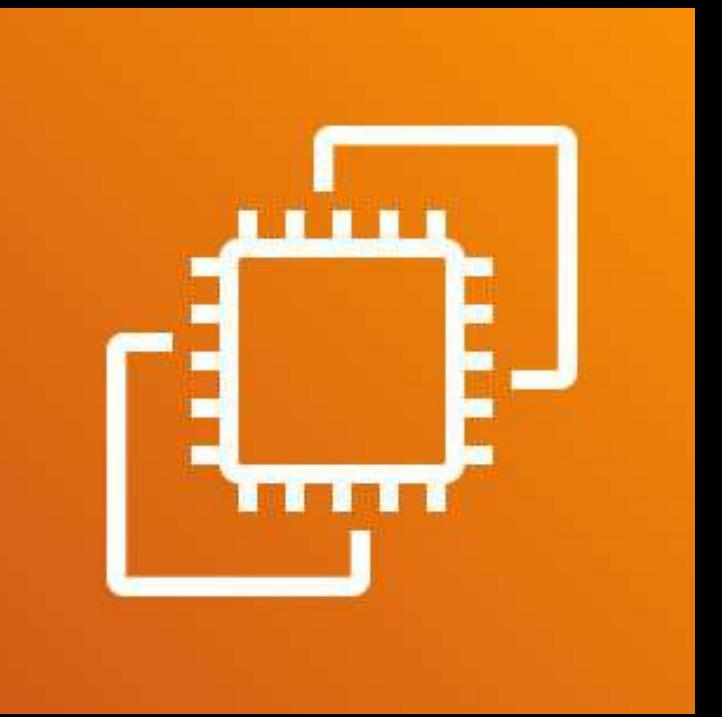
Specify the fields to include in the flow log record.

- AWS default format
- Custom format

#### Format preview

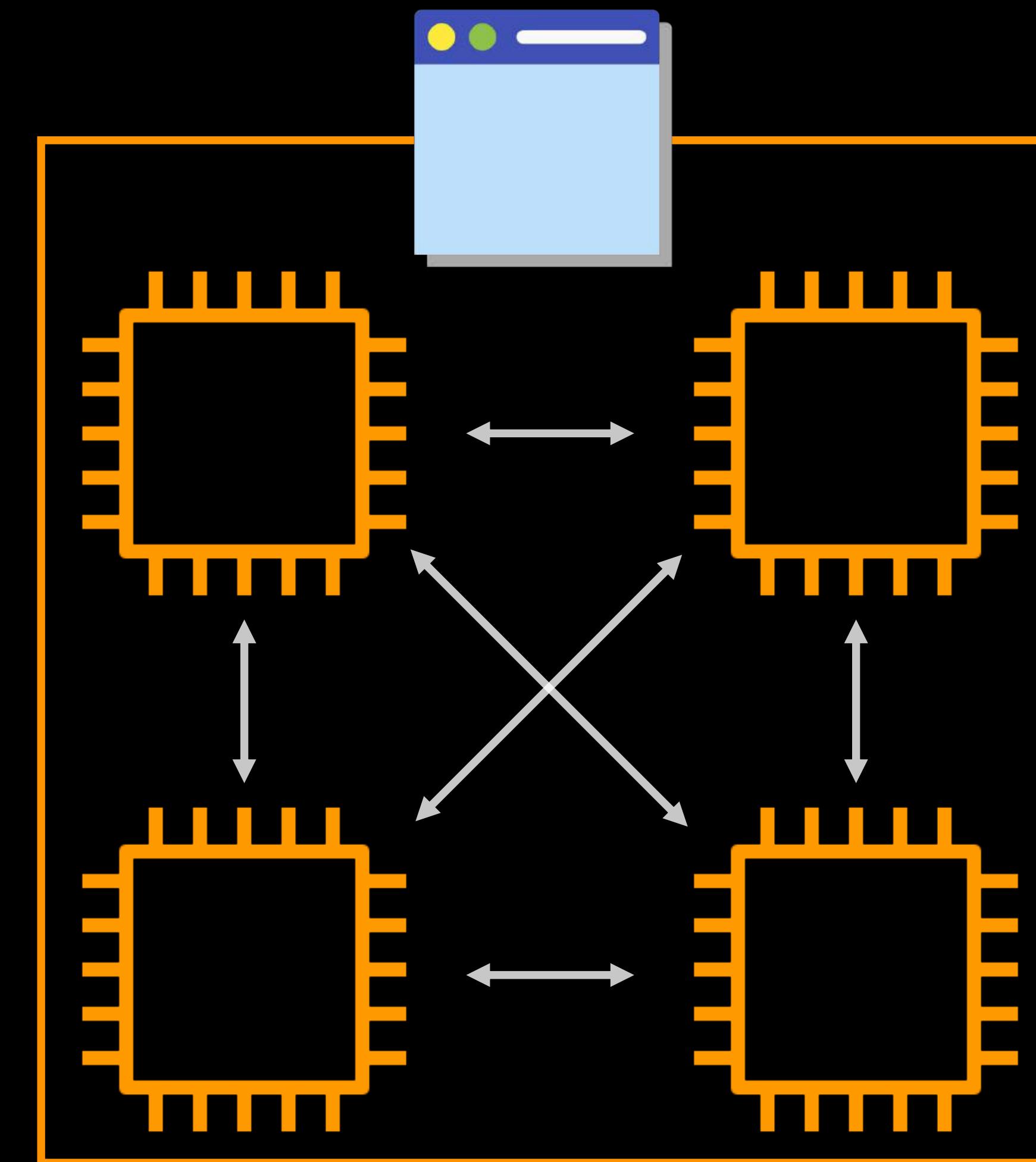
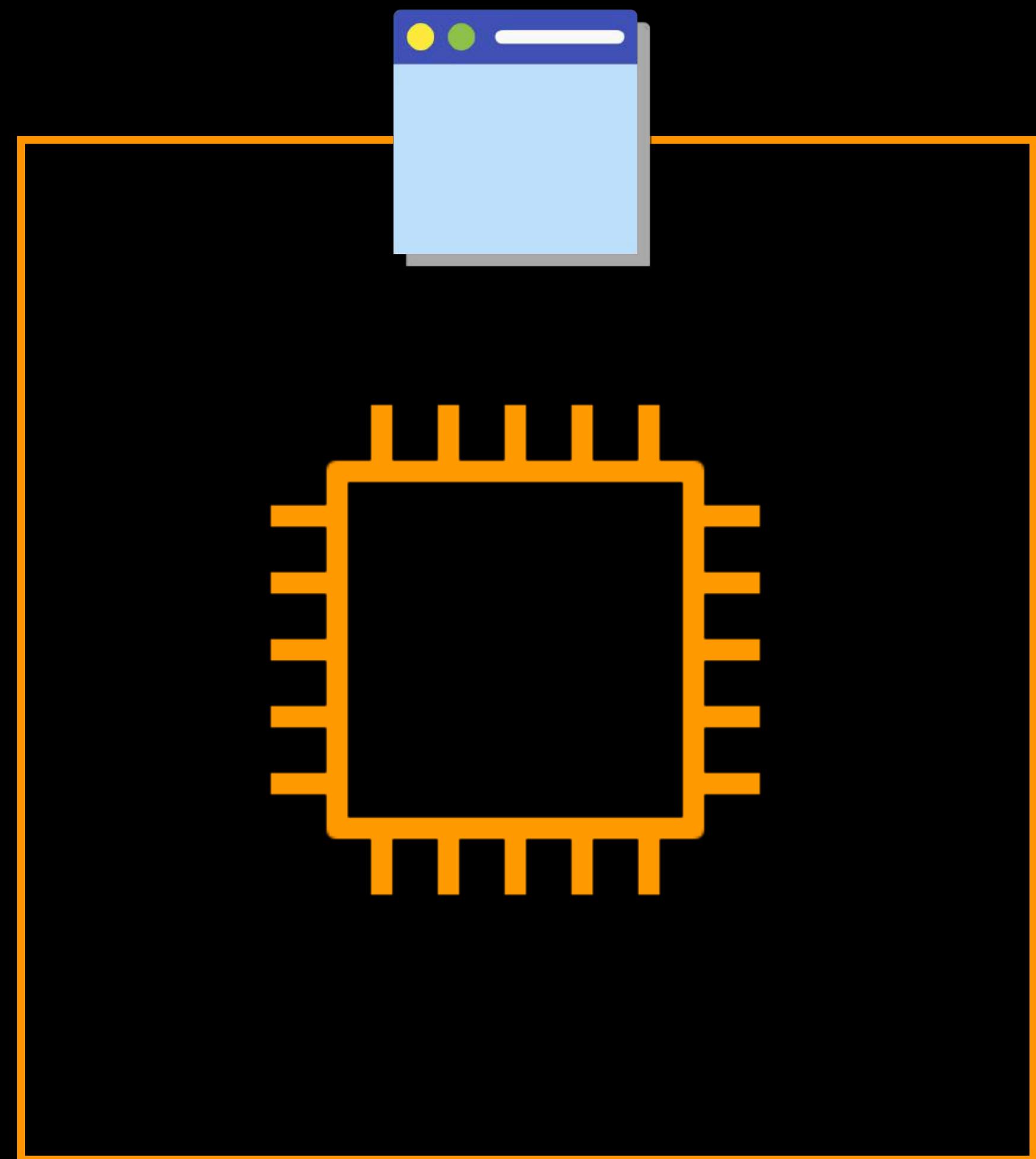
```
 ${version} ${account-id} ${interface-id} ${srcaddr} ${dstaddr} ${srcport} ${dstport}
 ${protocol} ${packets} ${bytes} ${start} ${end} ${action} ${log-status}
```

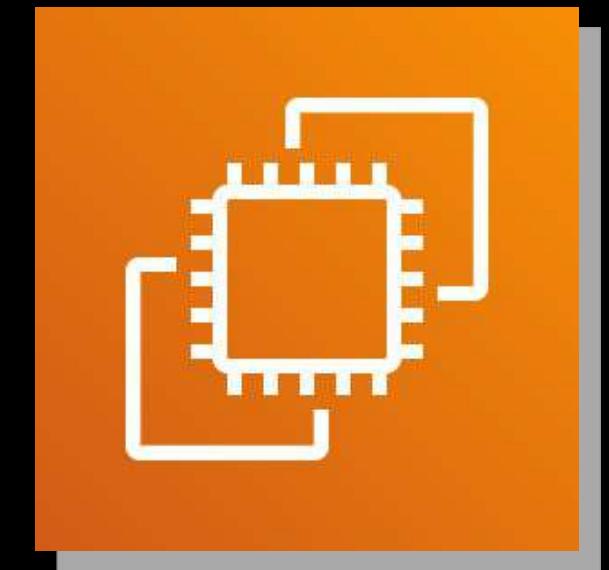
Copy



# Placement Groups

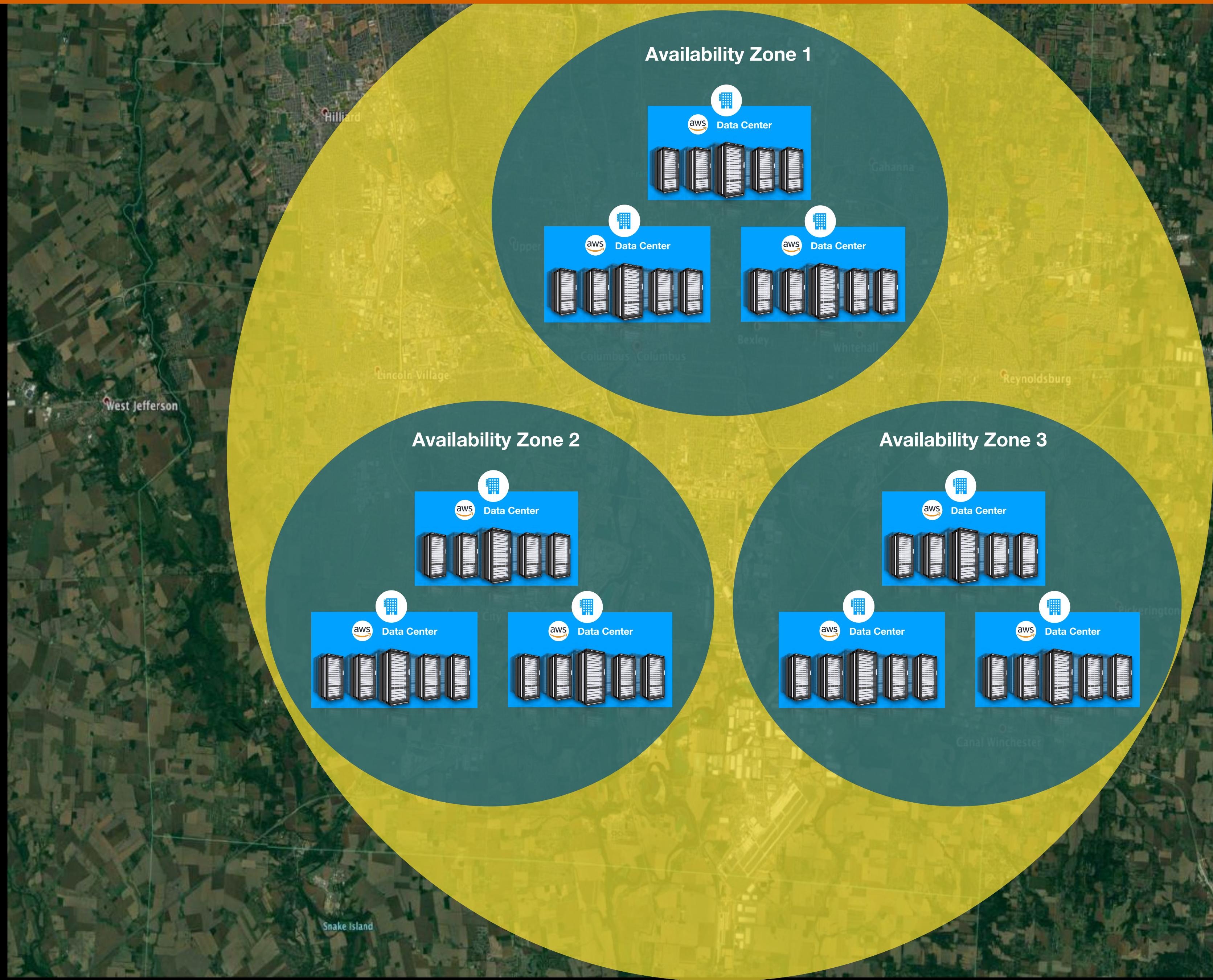
---





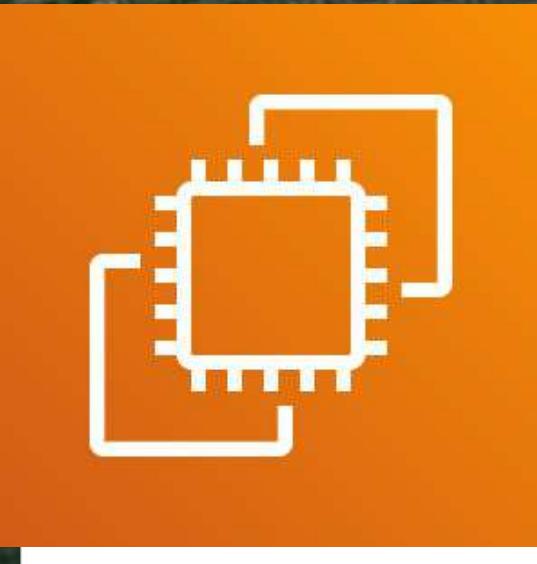
## Amazon EC2 Service

US East (Ohio)  
us-east-2

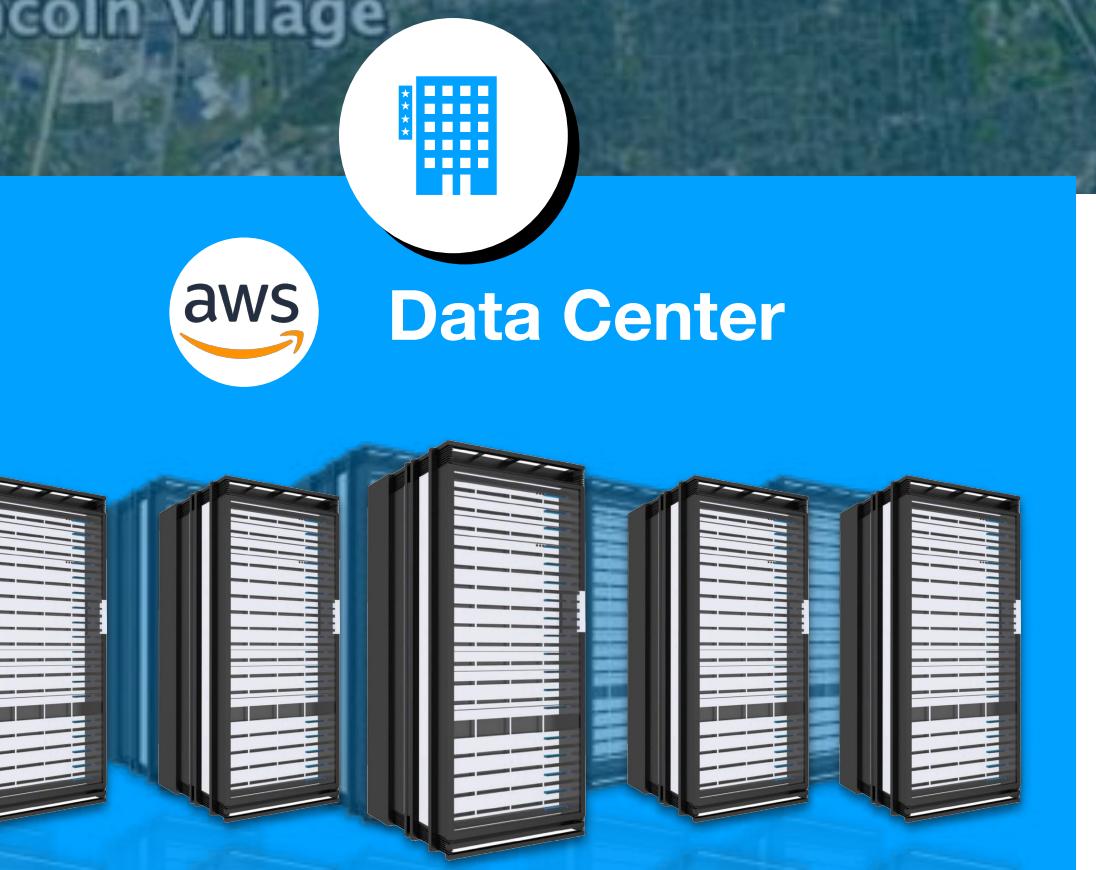


US East (Ohio)  
us-east-2

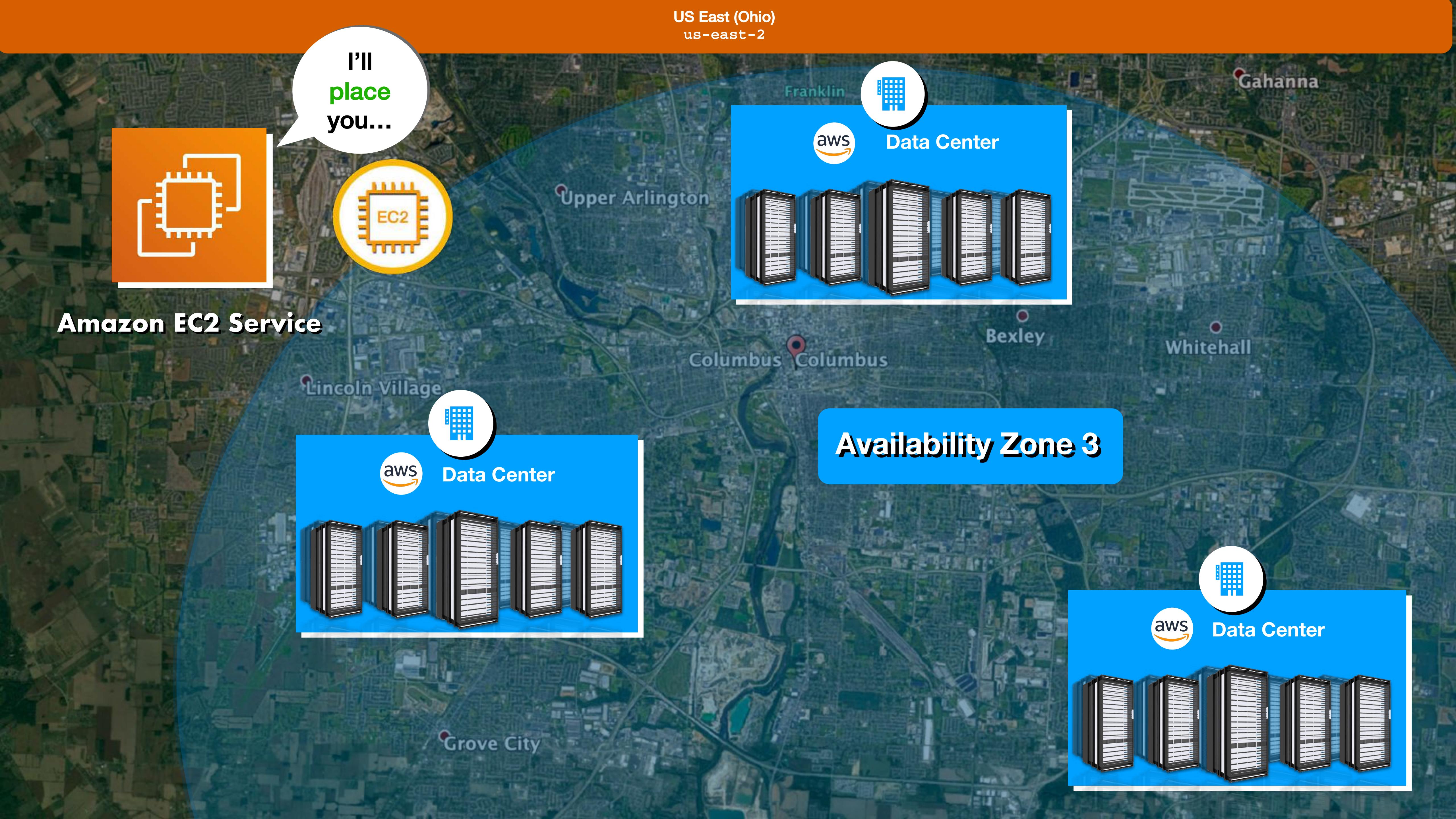
I'll  
place  
you...



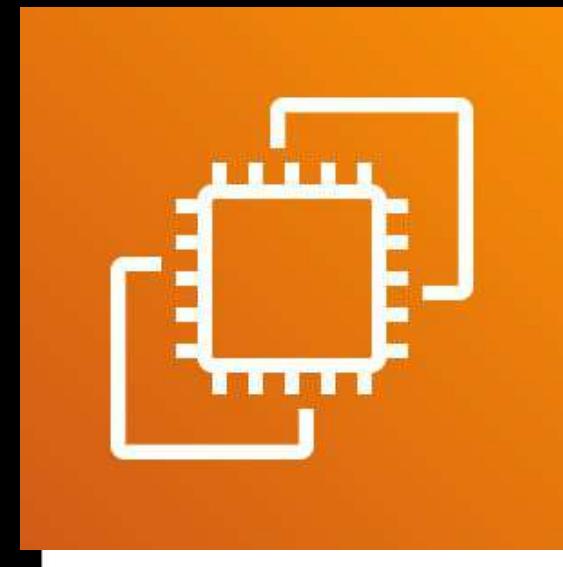
**Amazon EC2 Service**



**Availability Zone 3**



US East (Ohio)  
us-east-2

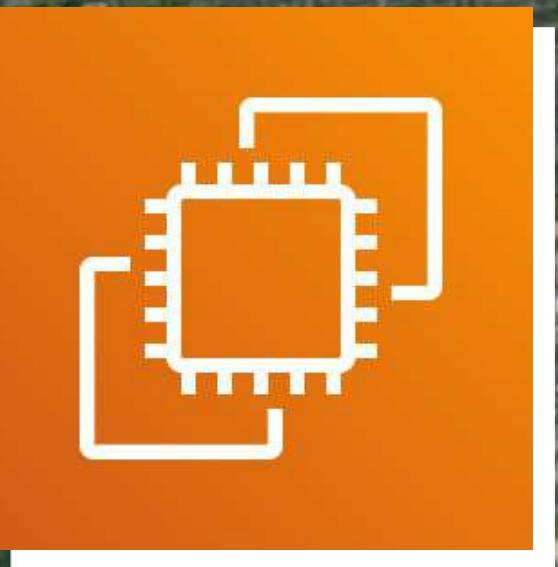


**Amazon EC2 Service**



**Availability Zone 3**





**Amazon EC2 Service**

**Availability Zone 3**



Grove City

Lincoln Village

Upper Arlington

Columbus Columbus

Franklin

Bexley

Whitehall

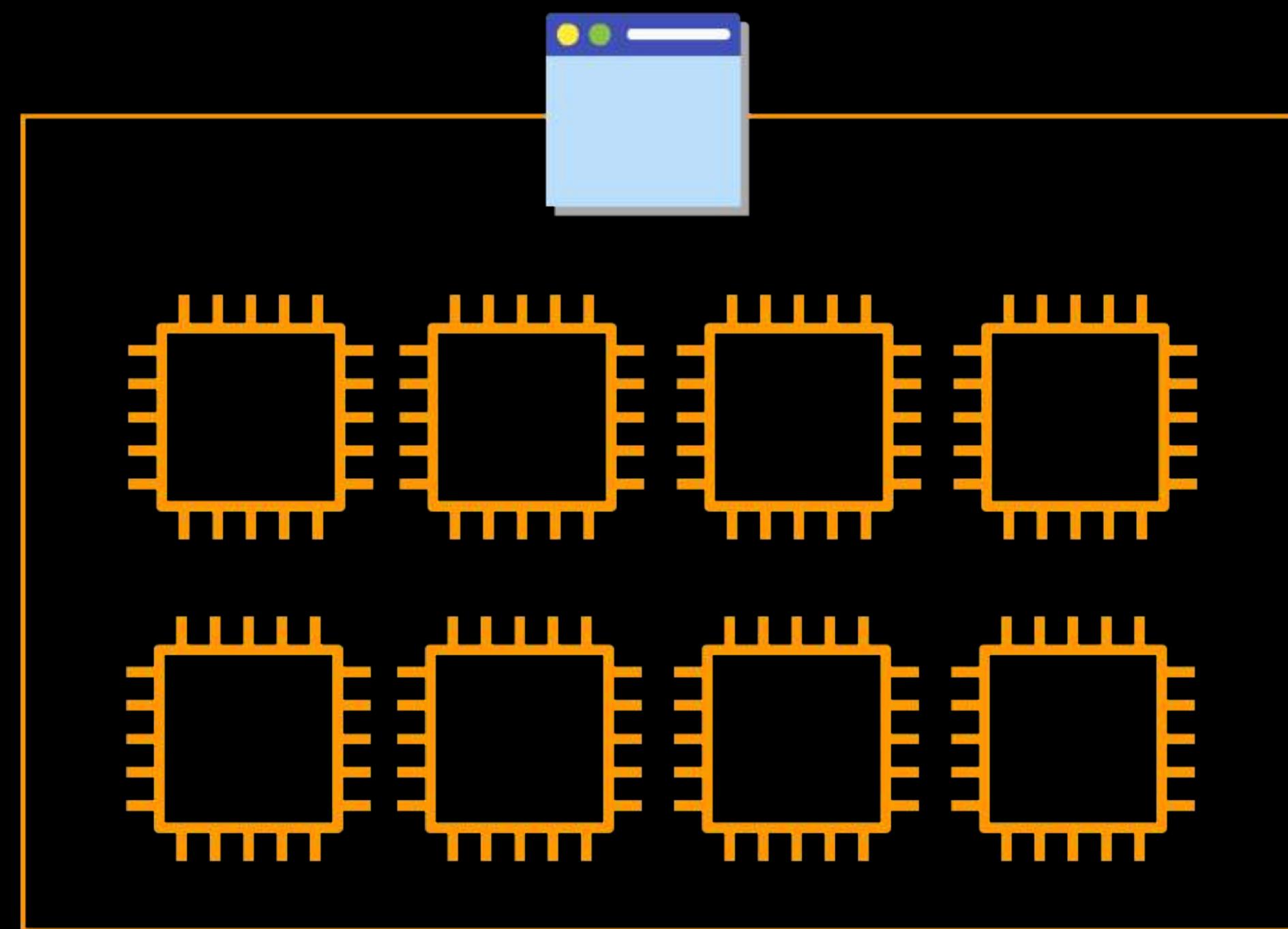
Gahanna

# Placement Groups

CLUSTER

PARTITION

SPREAD

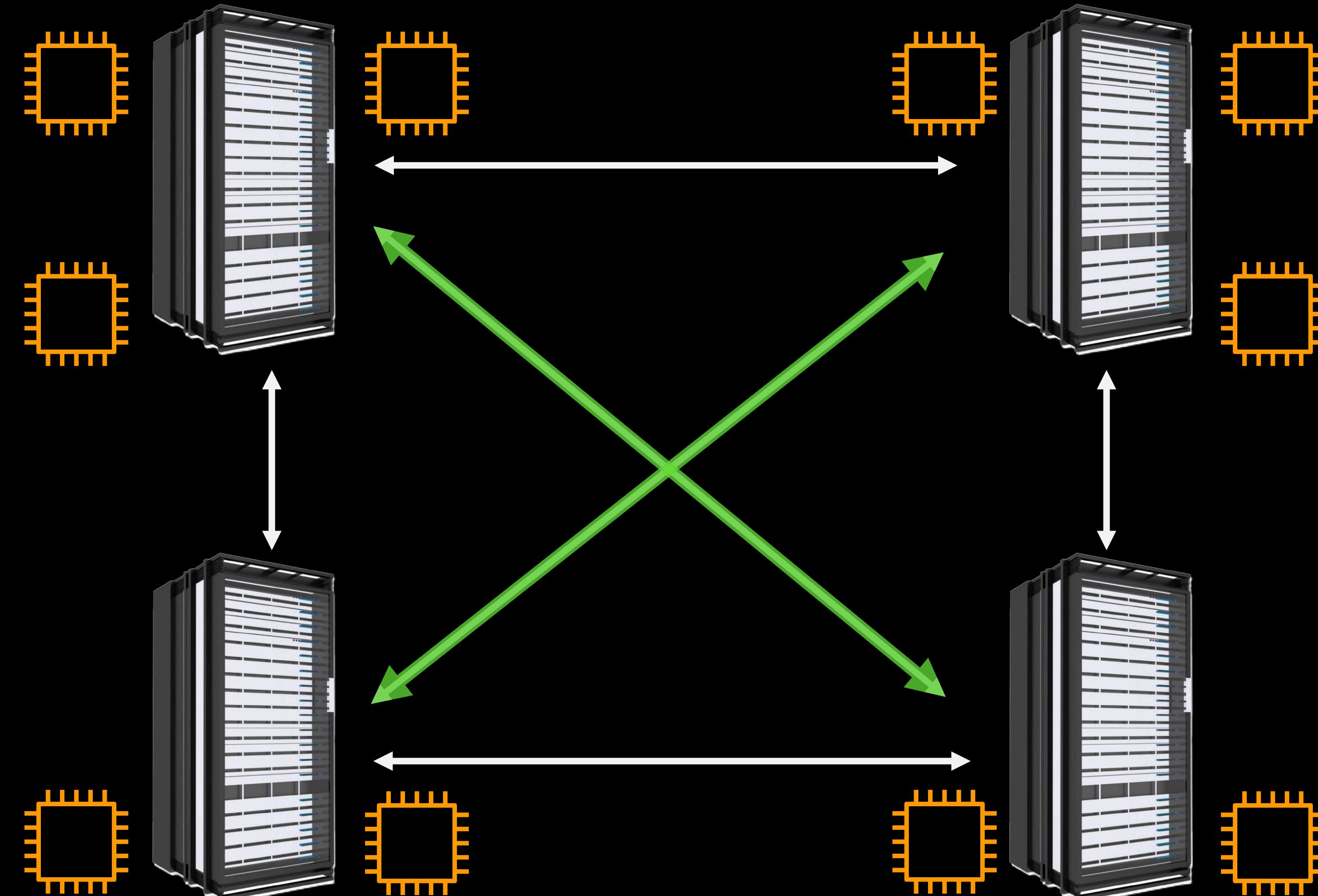


## CLUSTER

Provide low-latency network performance and high network throughput

Availability Zone

Logical Group / Host Rack Networking



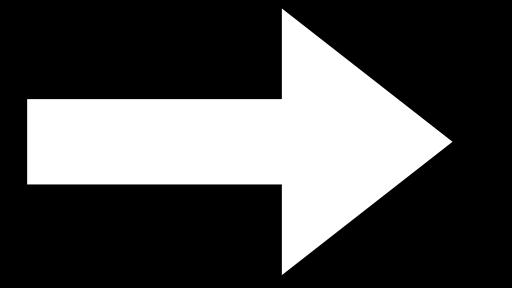
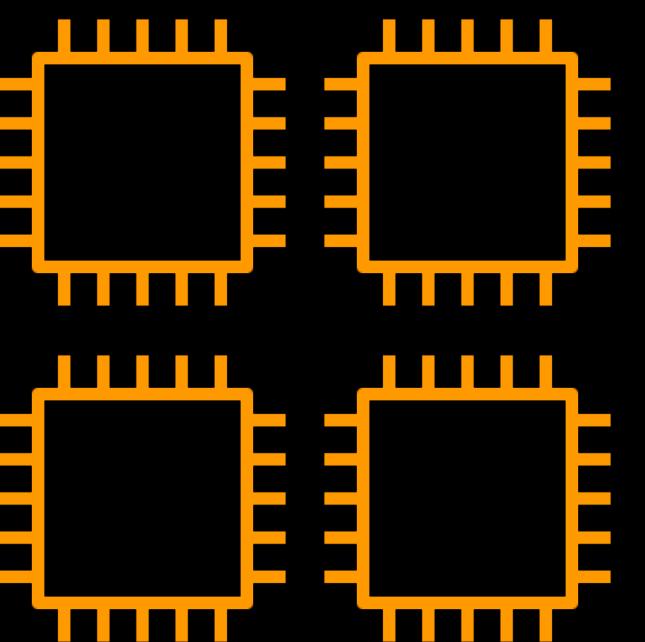
Group of rack servers on a network building block with special routing configuration

## PARTITION

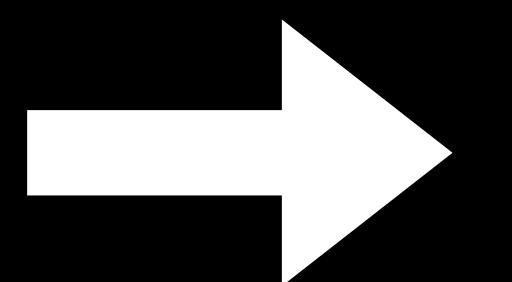
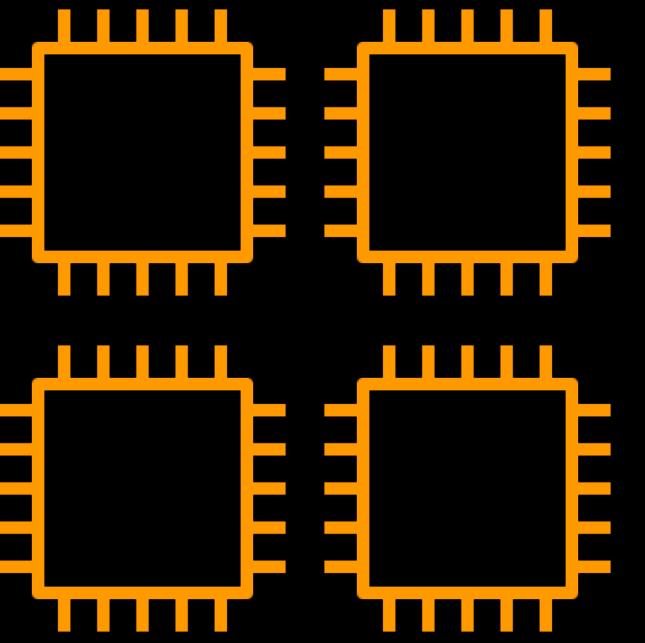
Commonly used on large distributed and replicated workloads, such as Hadoop, Cassandra, and Kafka

Availability Zone

Partition 1



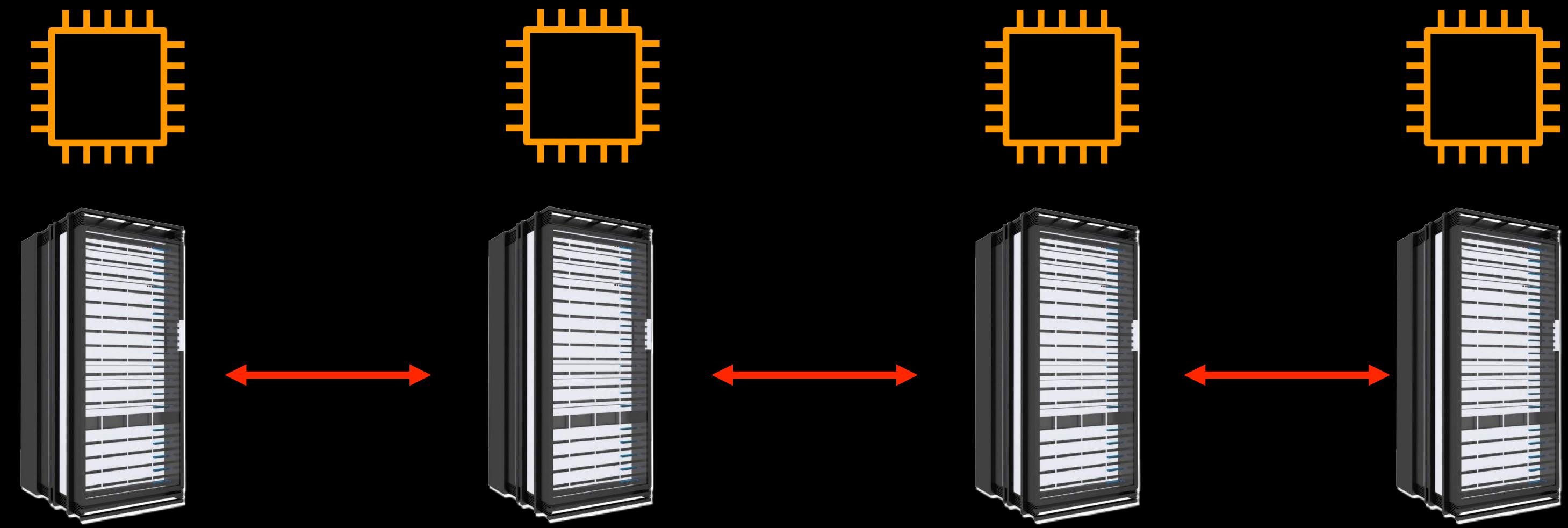
Partition 2



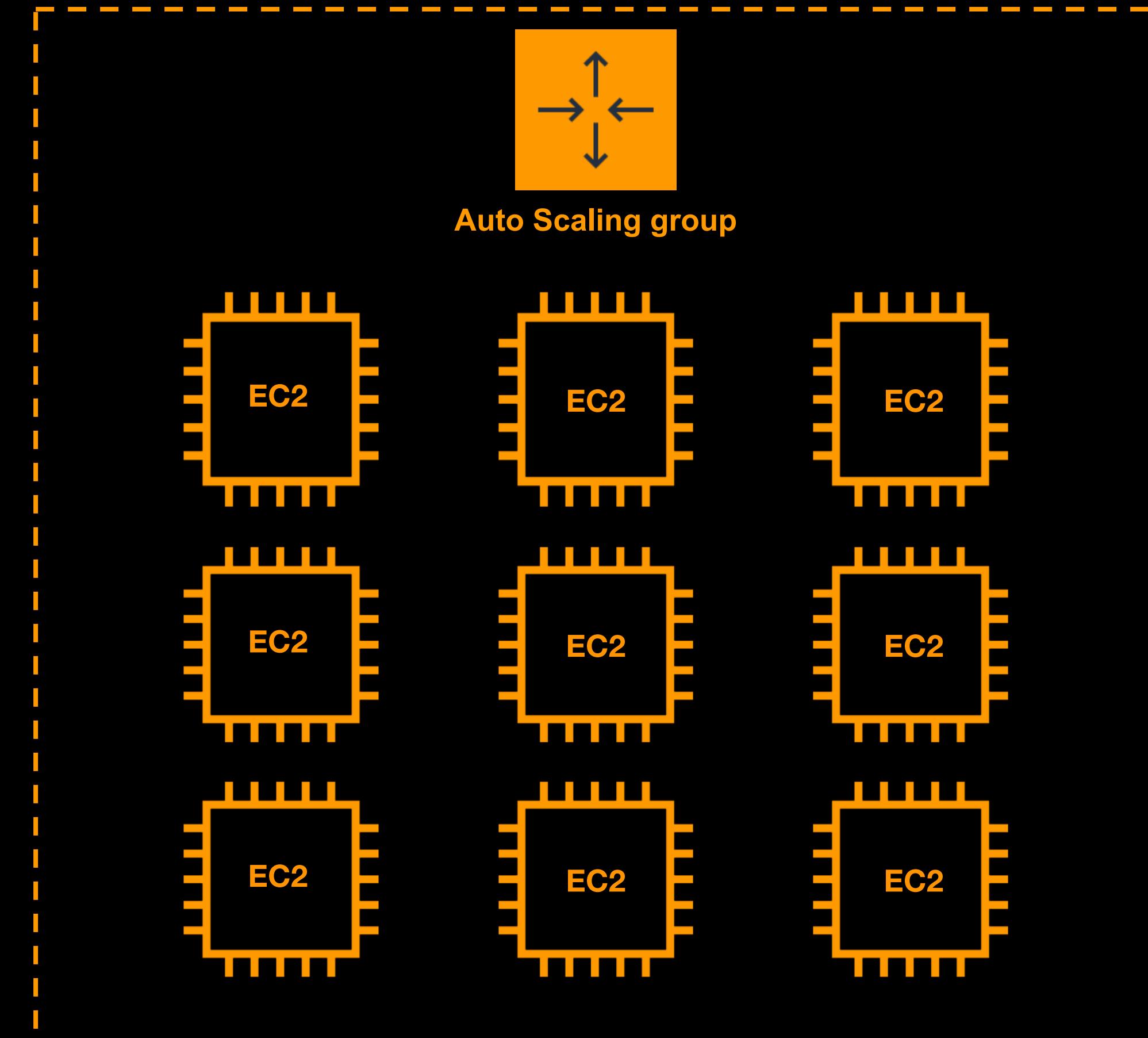
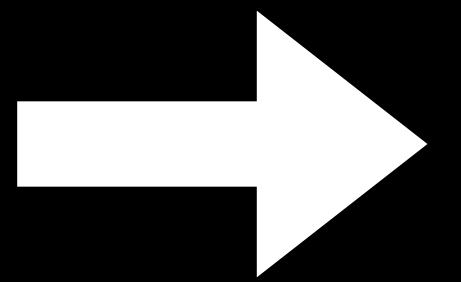
Availability Zone

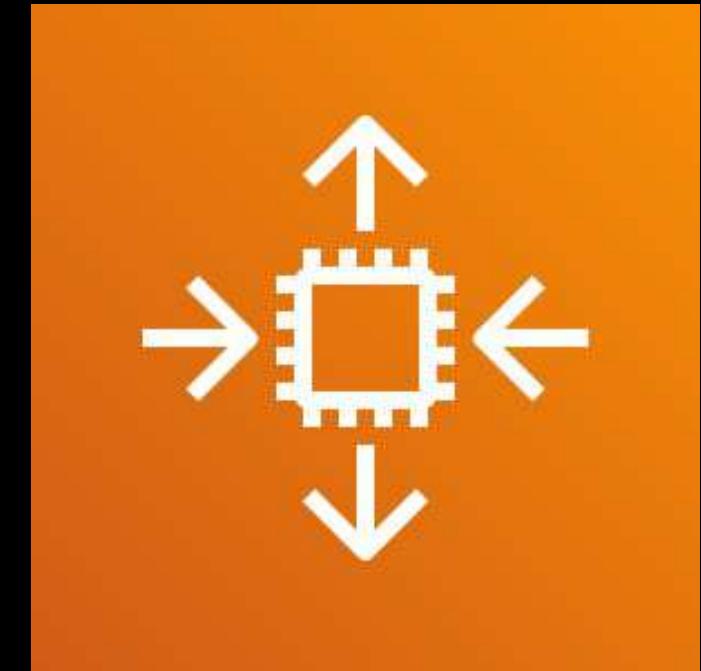
## SPREAD

Reduces correlated failures and improves availability

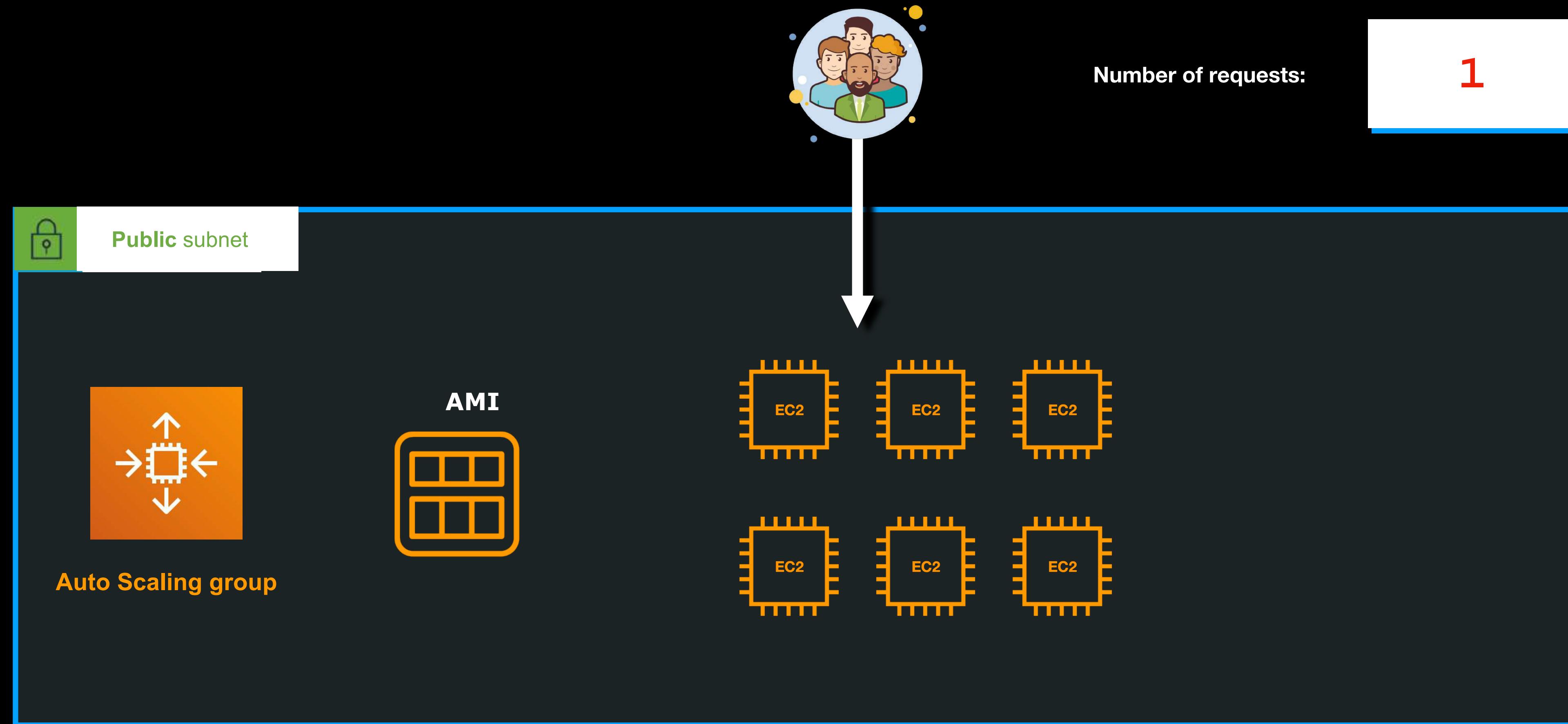


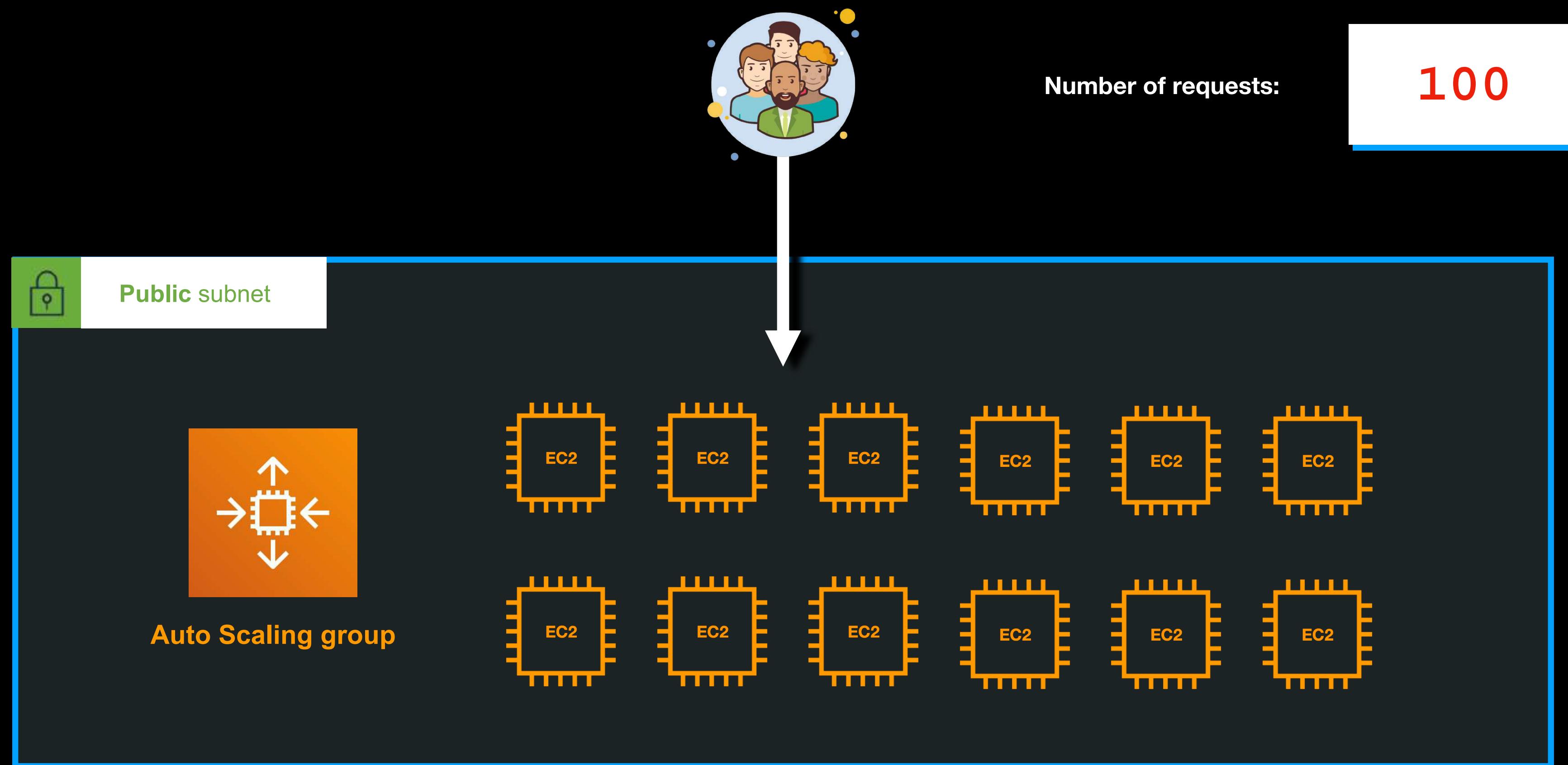
**Placement Group**

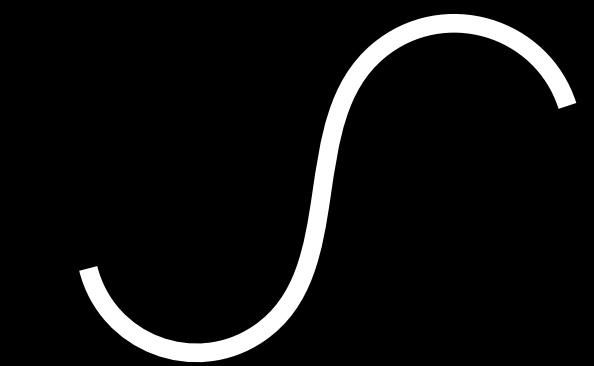




# Amazon EC2 Auto Scaling Overview

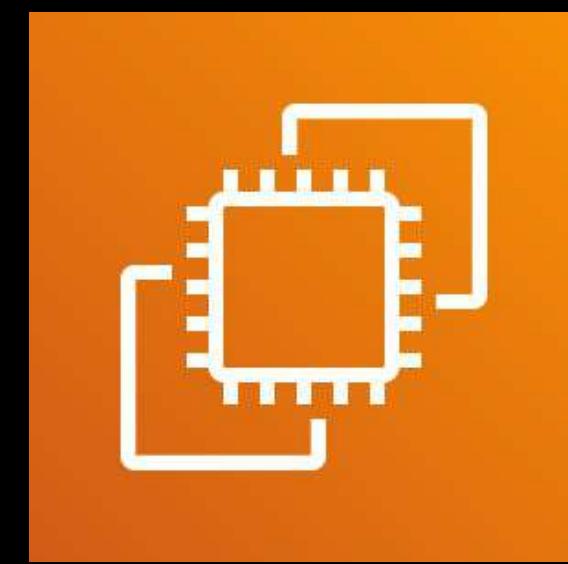




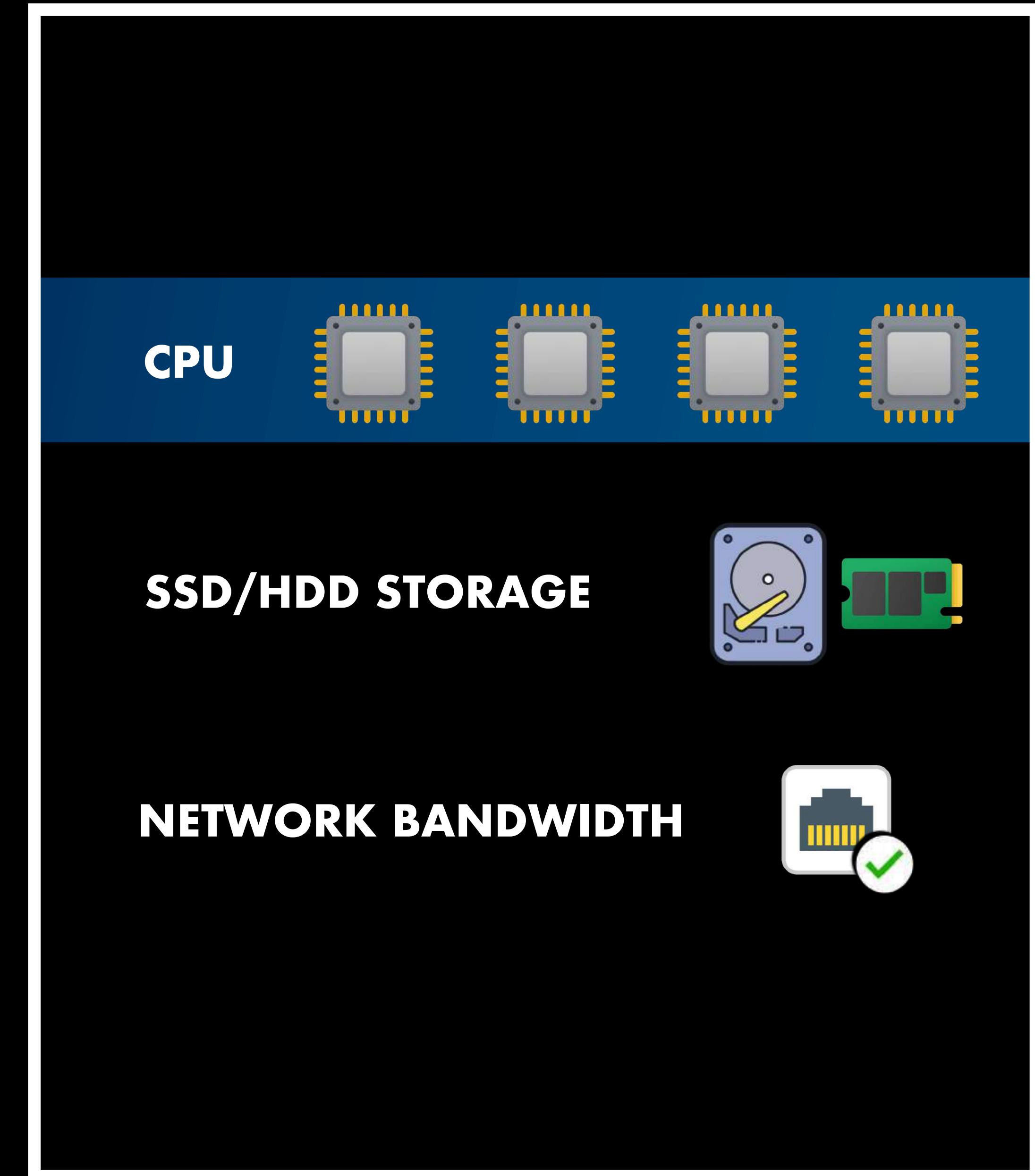


## ELASTICITY

- The ability to **dynamically acquire or release resources** when you need them
- Can be easily done in the cloud since it has hundreds of thousands of servers
- Improves the performance of your application when it is experiencing a surge of requests
- Avoids over-provisioning of your resources
- Lowers down your operating costs significantly by eliminating idle resources



Amazon EC2





On-premises data center



**RIGID** and **NOT FLEXIBLE**

# SCALING TYPES

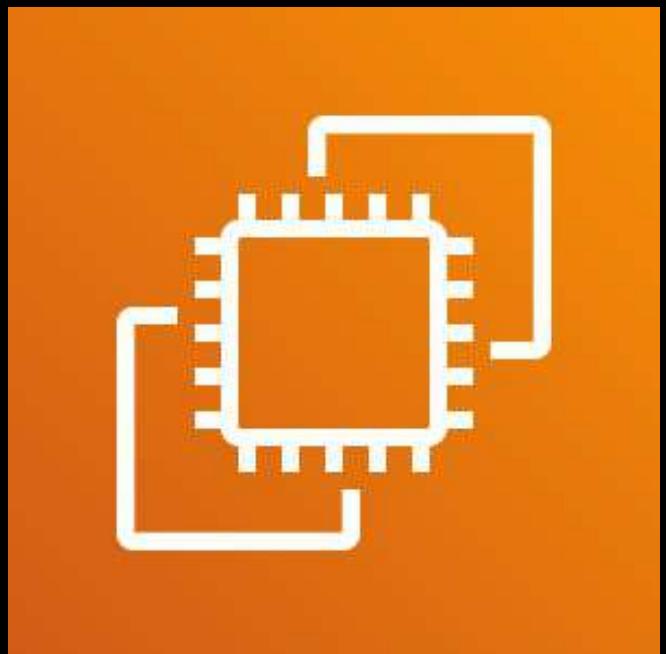
VERTICAL SCALING

HORIZONTAL SCALING

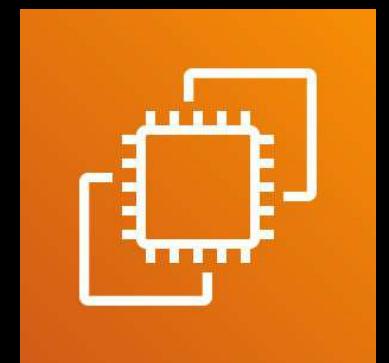
**SCALE UP**

**VERTICAL SCALING**

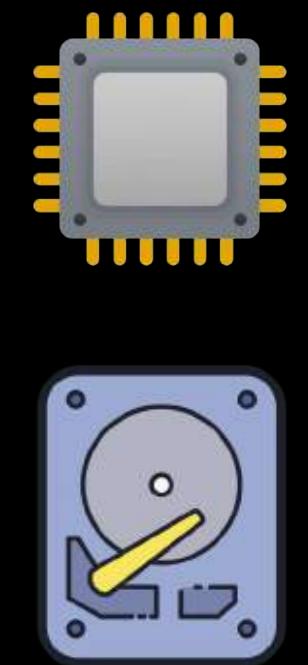
**SCALE DOWN**



Large Amazon EC2  
Instance Type

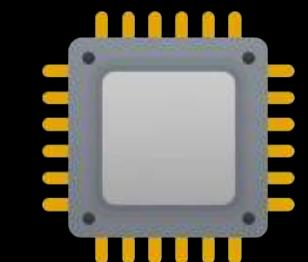


Small Amazon EC2  
Instance Type



**30 vCPU**

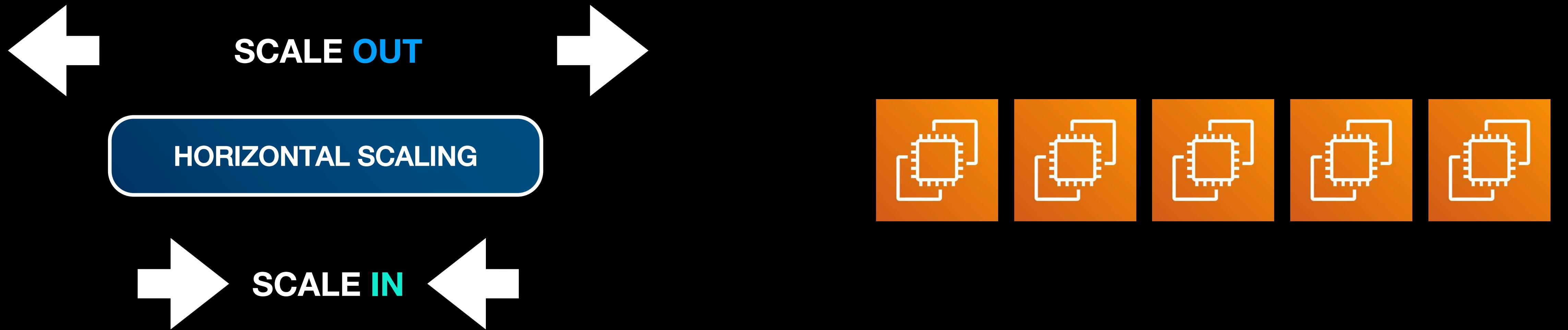
**300 GB**

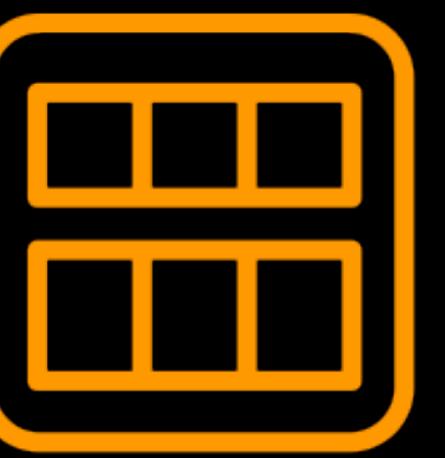


**10 vCPU**

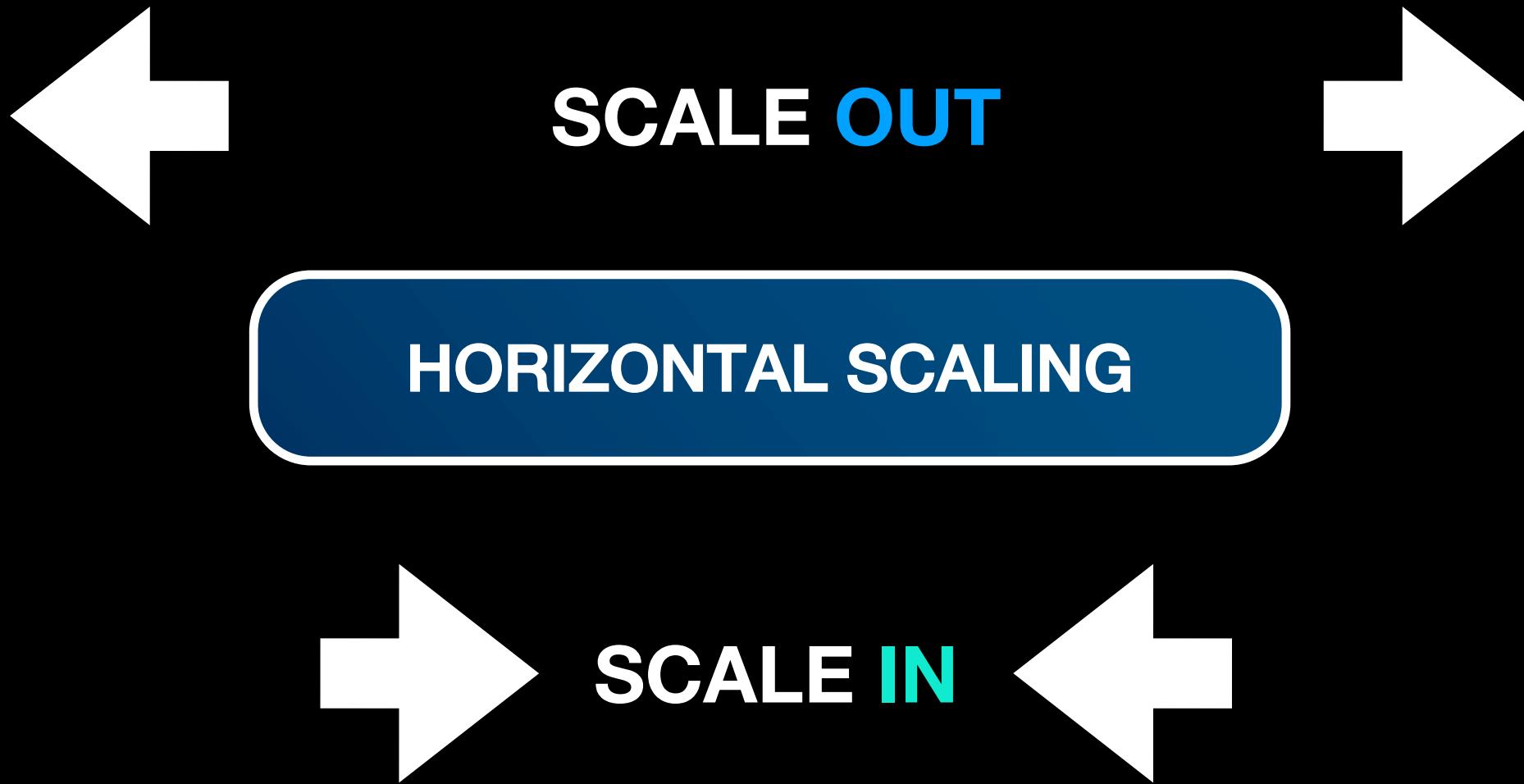


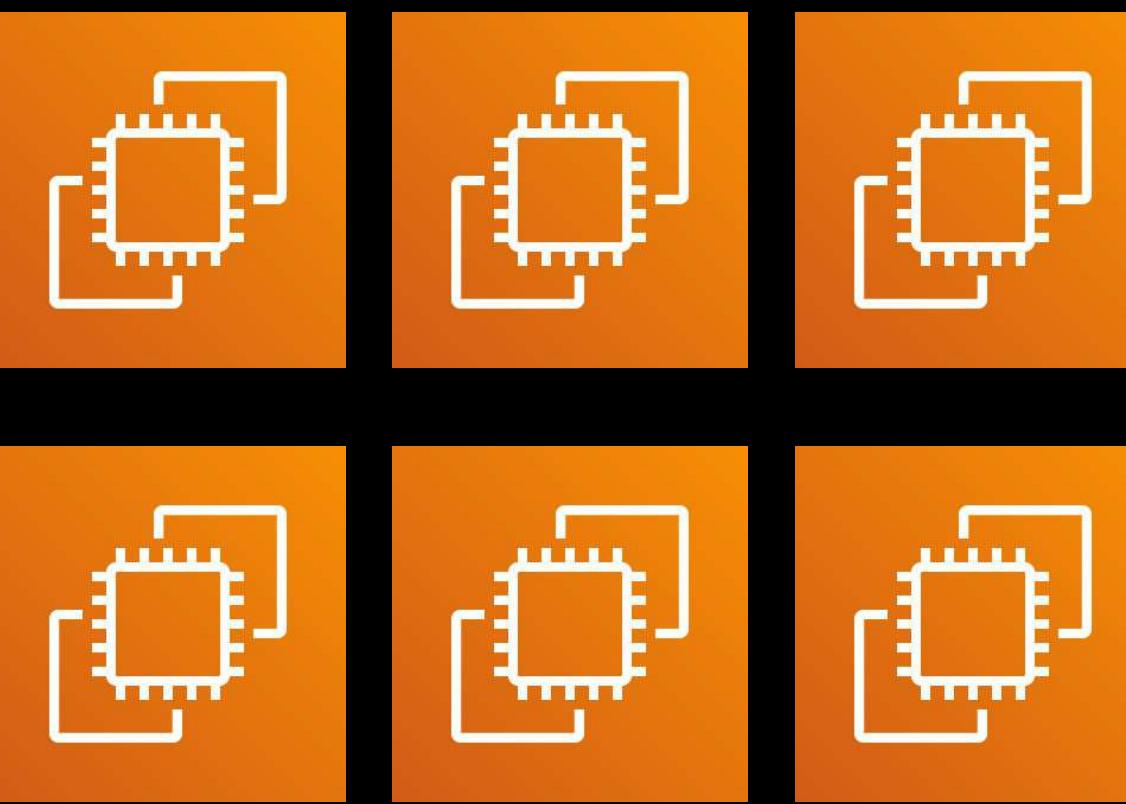
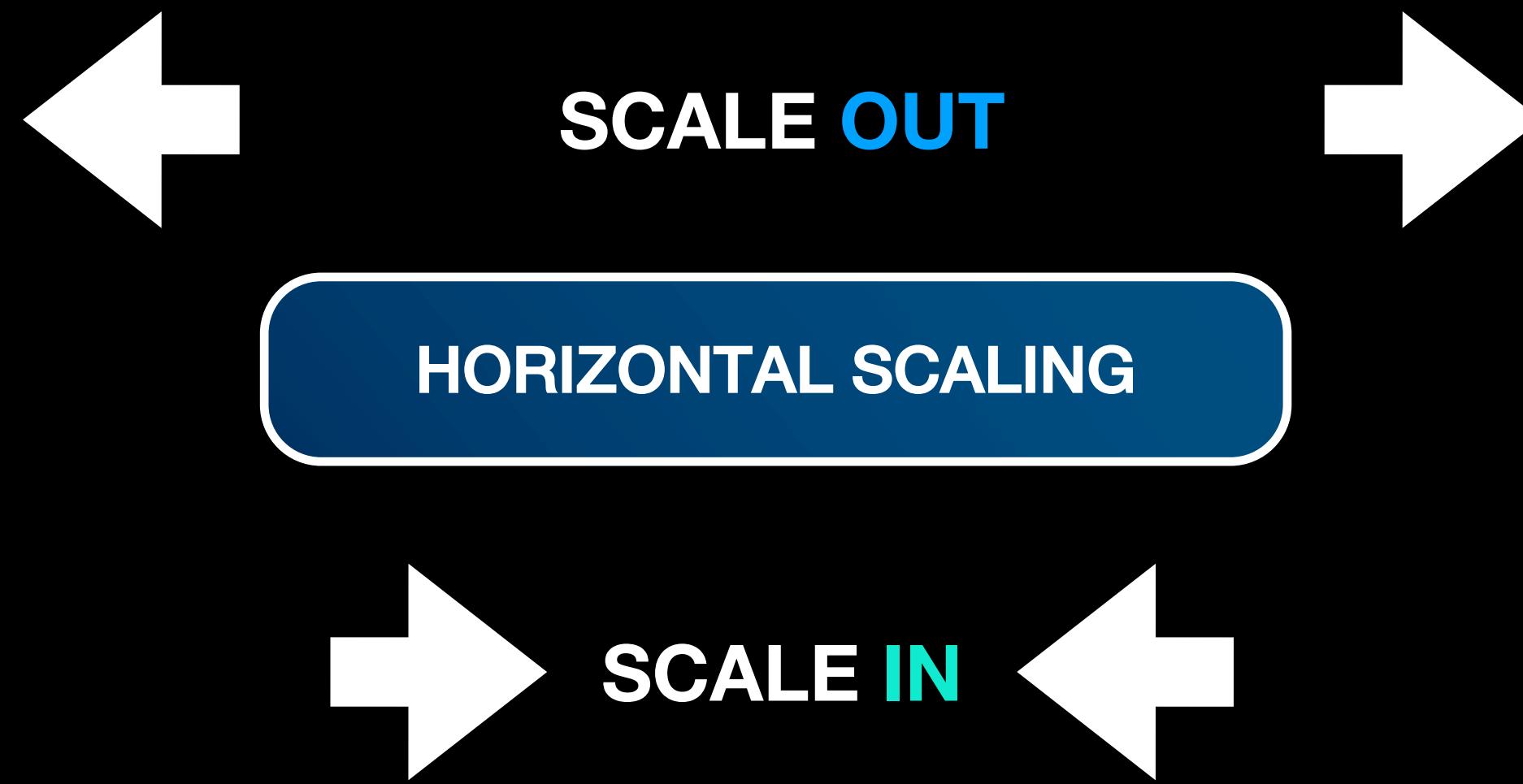
**100 GB**

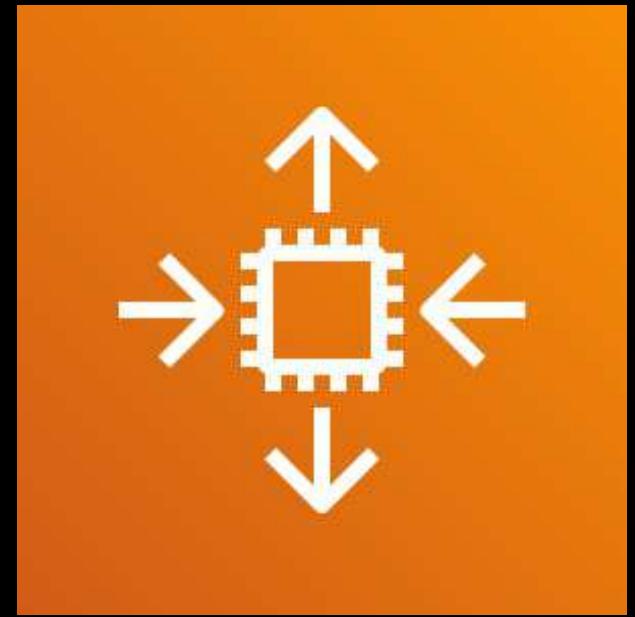




Amazon Machine  
Image (AMI)

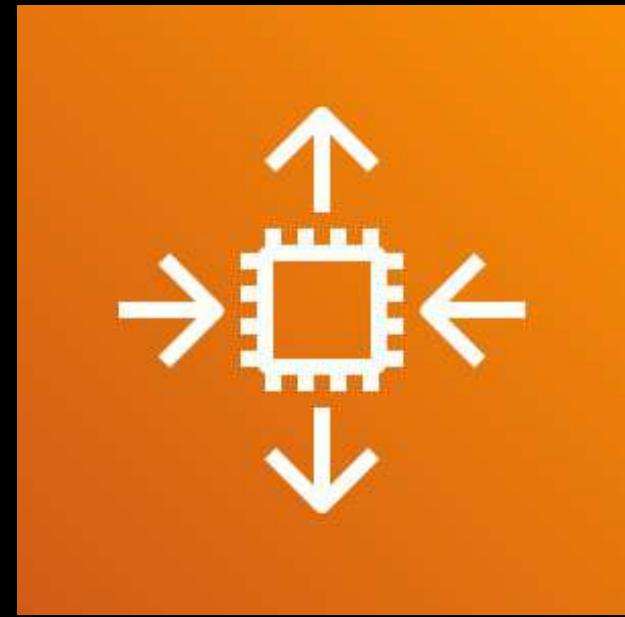






## Amazon EC2 Auto Scaling

HORIZONTAL SCALING



## Amazon EC2 Auto Scaling

AUTO SCALING GROUP

CONFIGURATION TEMPLATE

SCALING OPTION

## AUTO SCALING GROUP

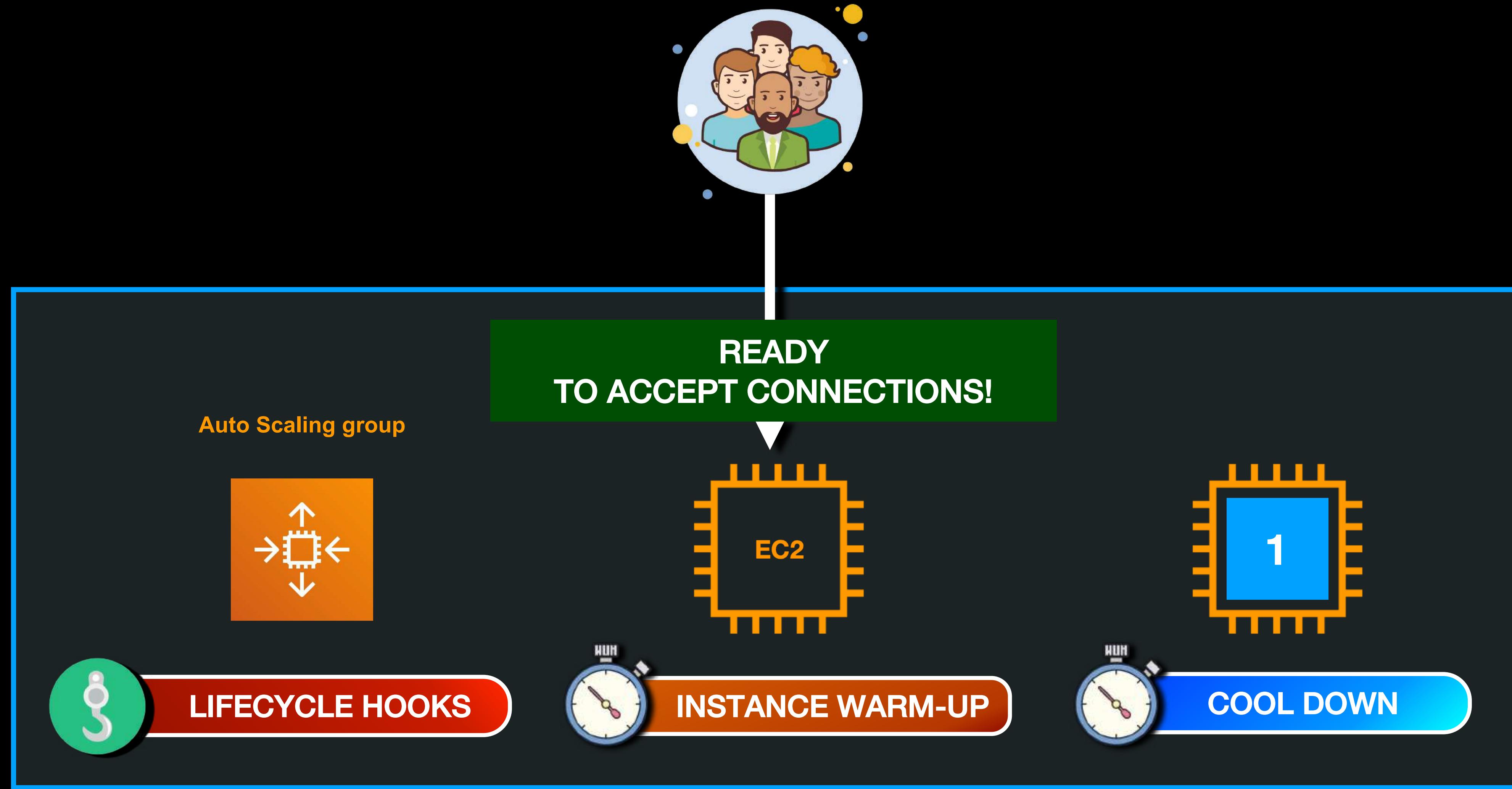
- **Organizes your Amazon EC2 instances into groups**
- **A logical unit for scaling and management**
- **Must have a setting for the minimum, maximum, and desired number of Amazon EC2 instances**

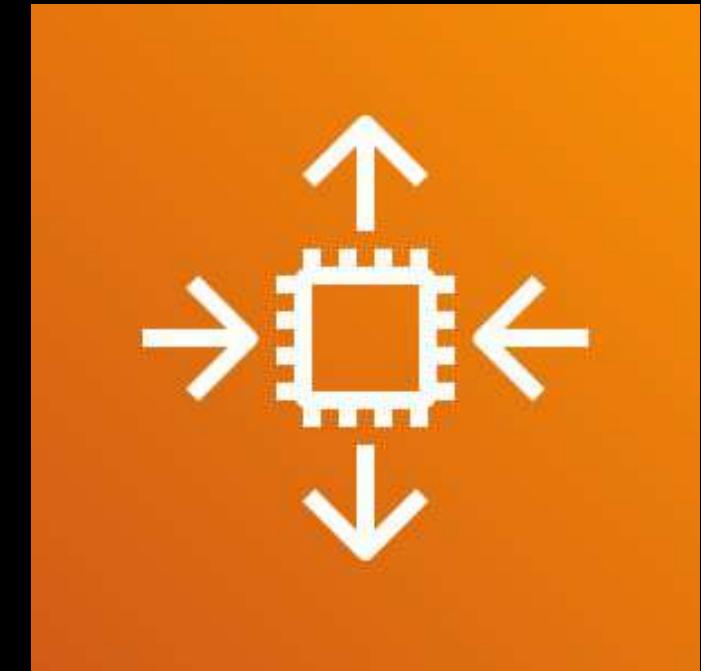
## CONFIGURATION TEMPLATE

- **Types:**
  - **Launch Template**
  - **Launch Configuration**
- **Acts as a template for your Auto Scaling Group, containing the AMI ID, the instance type, the key pair, the security groups, block device mapping and others**
- **It is recommended to use a Launch Template, rather than a Launch Configuration, as the latter only offers limited features**

## SCALING OPTION

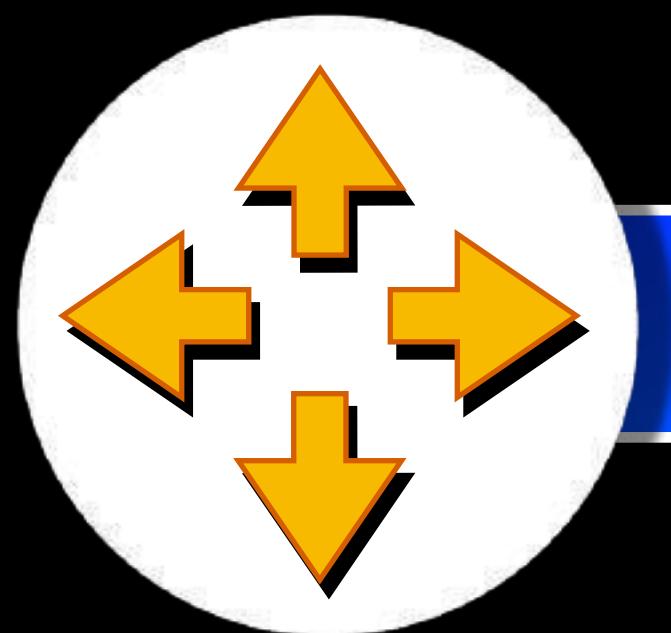
- Allows you to choose the suitable **scaling behavior** of your Auto Scaling Group.
- Types:
  - **Dynamic**
  - **Predictive**
  - **Scheduled**



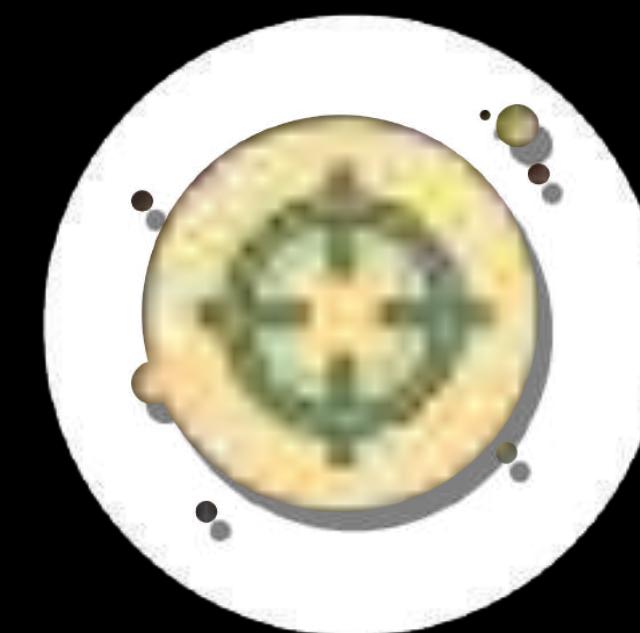


# Amazon EC2 Auto Scaling Types

# AMAZON EC2 AUTO SCALING TYPES



SIMPLE SCALING



TARGET TRACKING

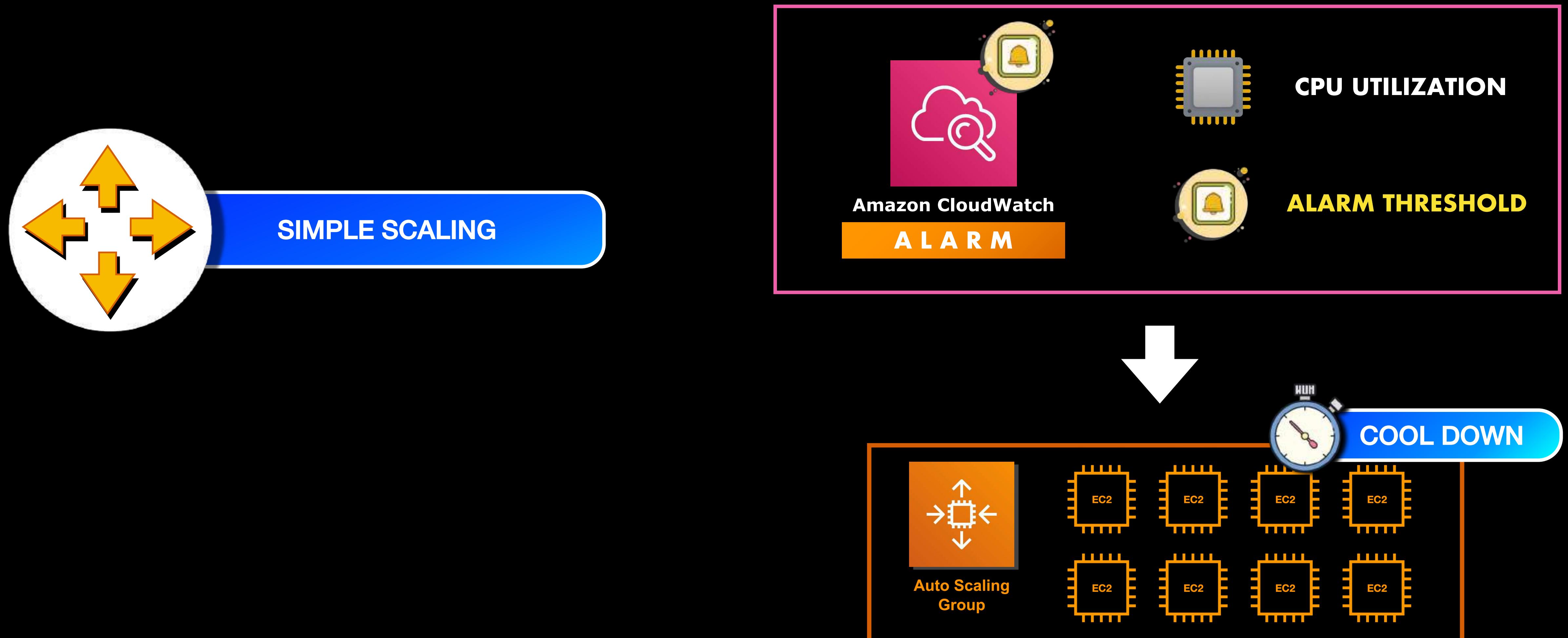


STEP SCALING

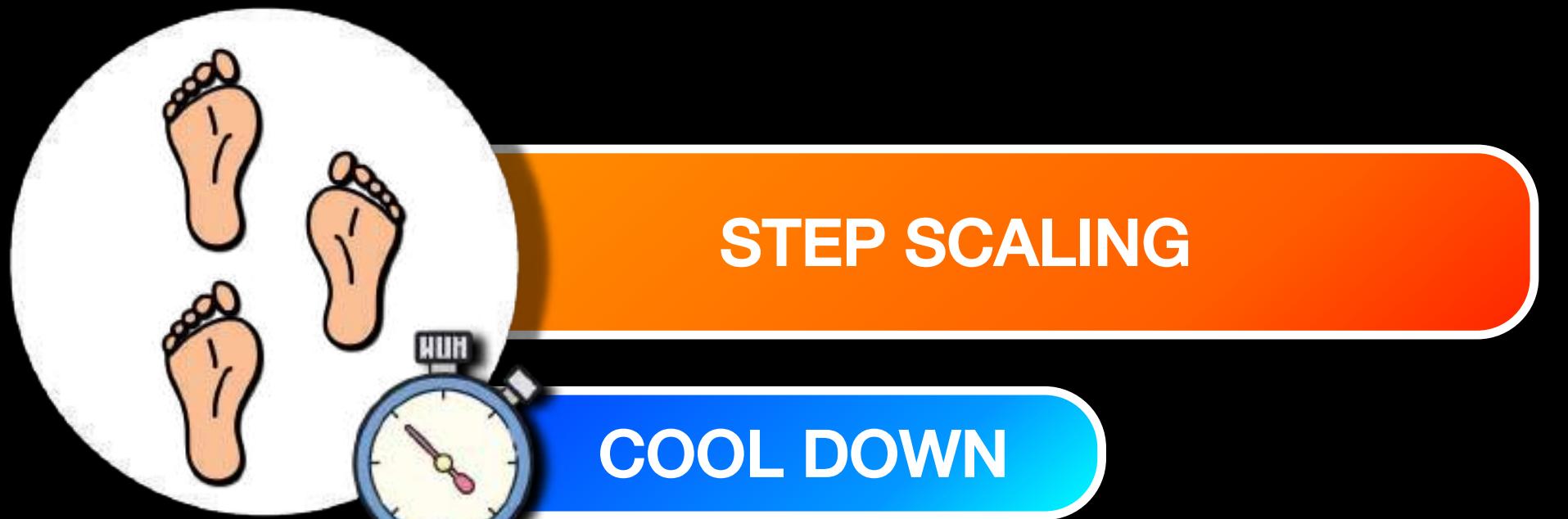


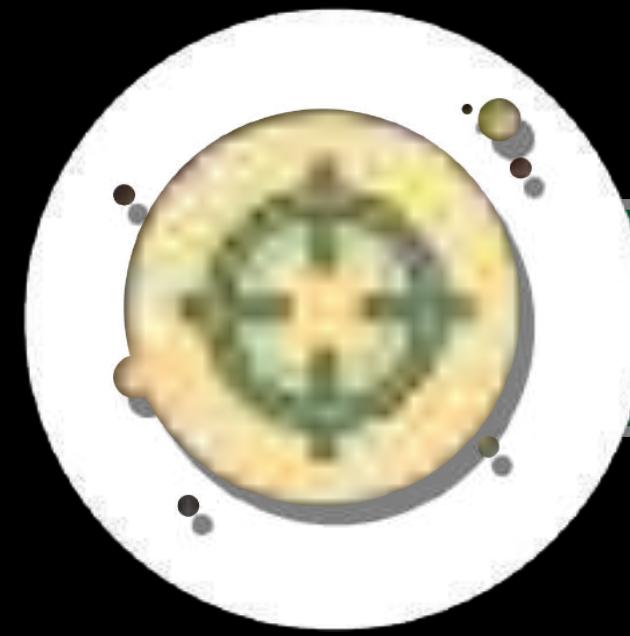
SCHEDULED SCALING

- Automatically increases or decreases the current capacity of your Auto Scaling Group based on a **single scaling adjustment**



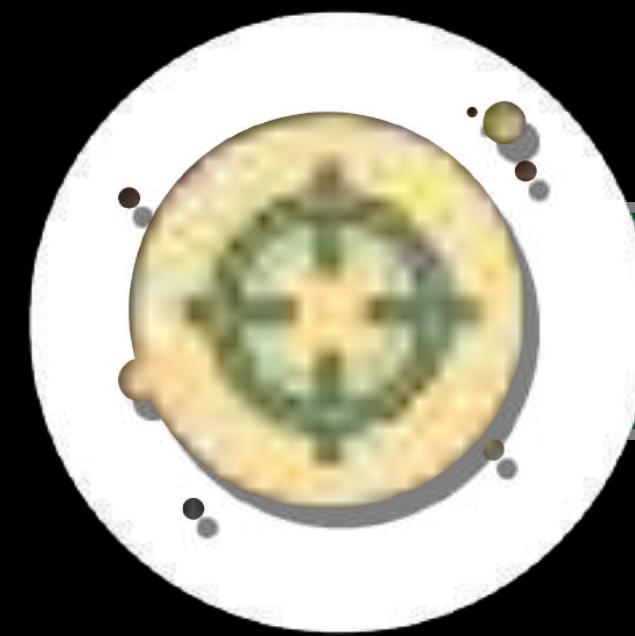
- Automatically increases or decreases the current capacity of your Amazon EC2 Auto Scaling group **based on a set of scaling adjustments**, also known as **step adjustments**
- Also requires the use of CloudWatch alarms with specified high and low thresholds as well as a defined action that either adds or removes instances
- Also supports setting the Auto Scaling group to an **exact size or a fixed capacity unit** in the event that your CloudWatch alarm threshold was breached
- Unlike Simple Scaling policy, it **can continue to respond to additional CloudWatch alarms**, even if the current scaling activity or health check replacement is already in progress



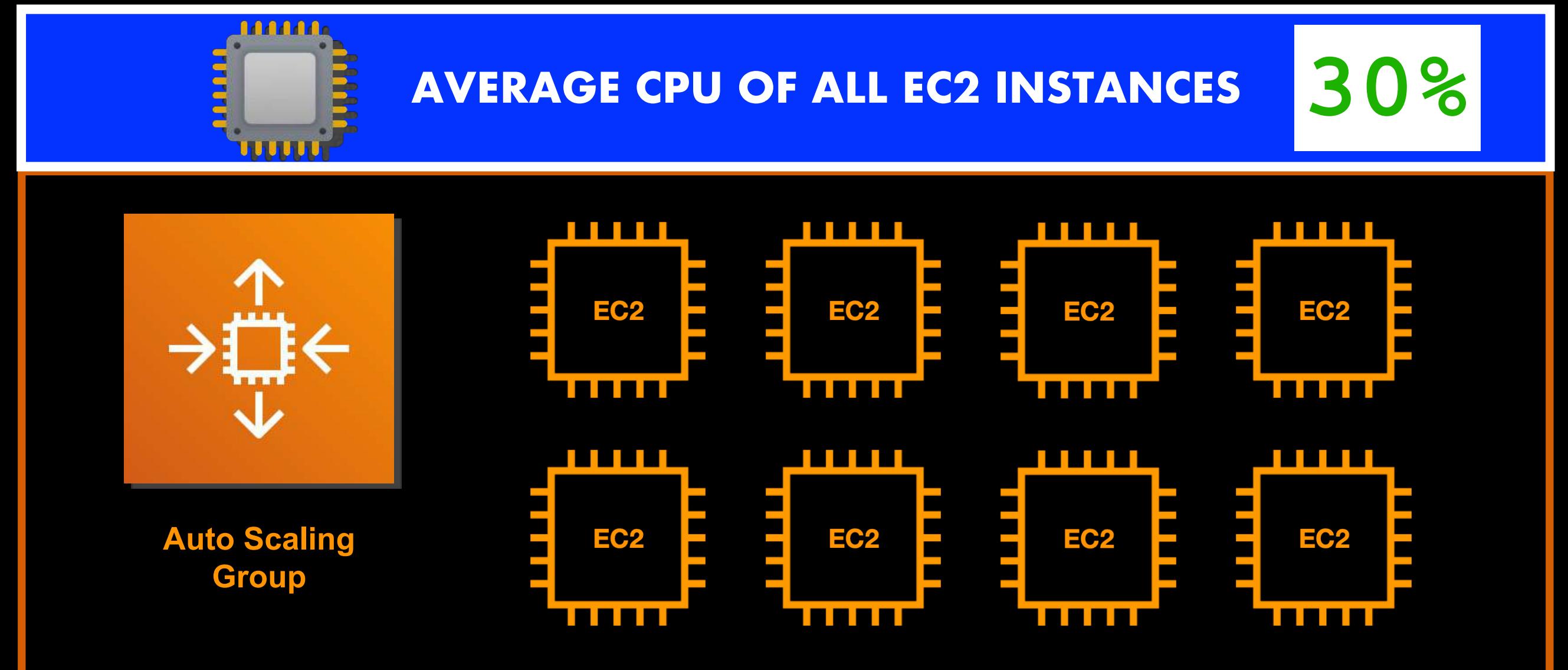
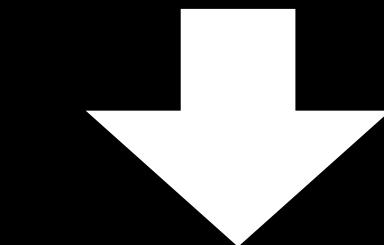
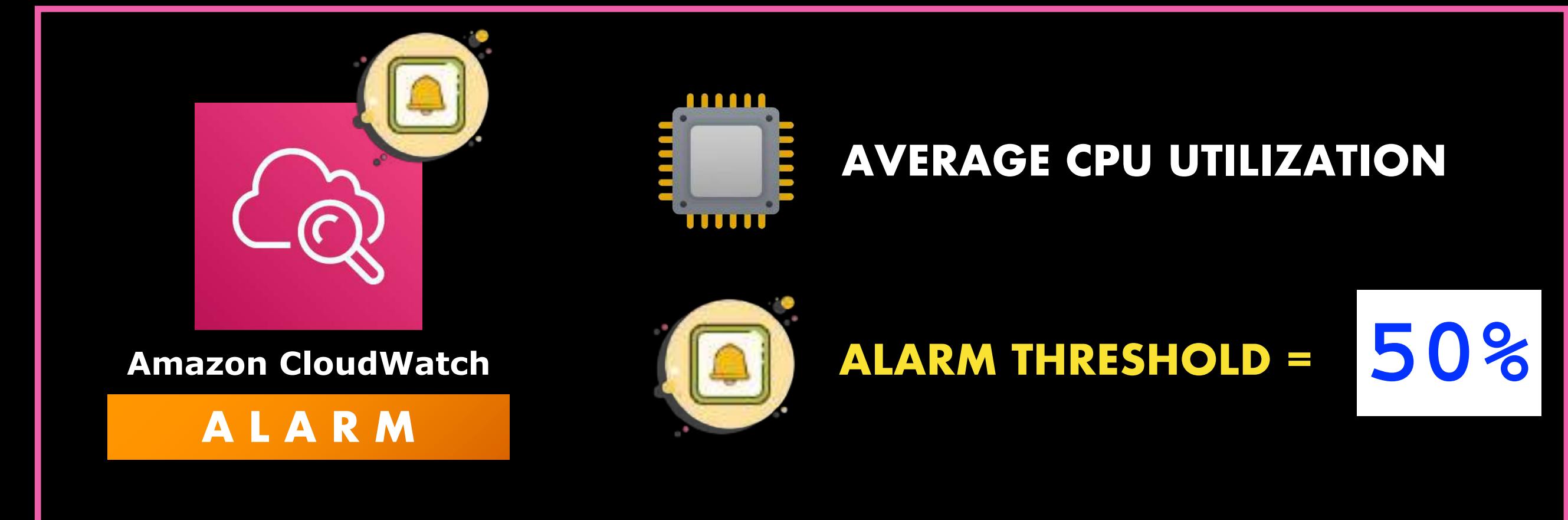


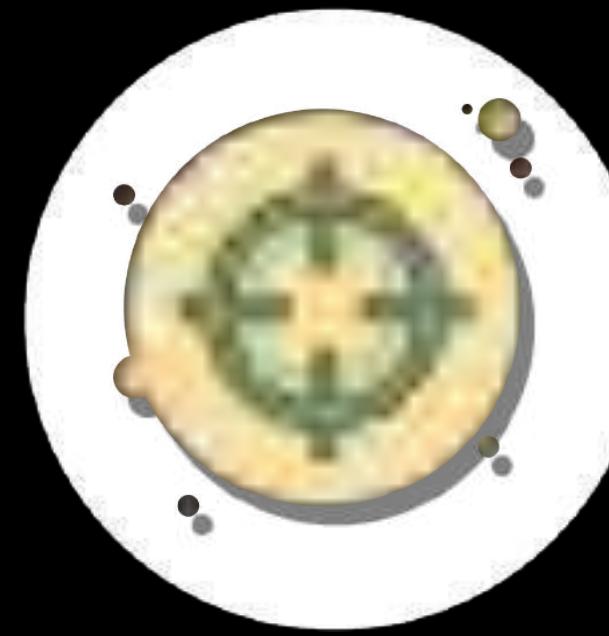
## TARGET TRACKING

- Automatically increases or decreases the current capacity of your Auto Scaling group based on a target value for a specific metric
- Maintains and adjusts the number of EC2 instances in your Auto Scaling group based on the target that you specify



TARGET TRACKING





TARGET TRACKING

WORKS LIKE A **THERMOSTAT!**



## TARGET TRACKING USE CASES

- If you've determined the **optimal performance** of your web application and you want to **maintain its desired performance across** all EC2 instances of your Auto Scaling group
- If your application works best when the combined CPU utilization of your Amazon EC2 instances is at or near a certain percentage (e.g. 40% ). You can set up a target tracking policy with a metric type of "**Average CPU utilization**" and a 40% target value



## TARGET TRACKING

## USE CASES

- Tracking of a certain metric that is produced by your application. You can track the **average network in or network out of all your instances**
- You can use the **request count per target** (`ALBRequestCountPerTarget`) metric of your Application Load Balancer as the metric type for your Target Tracking policy



## SCHEDULED SCALING

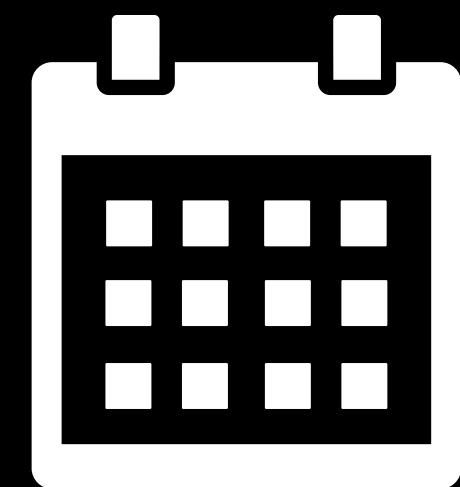
- Automatically increases or decreases the current capacity of your Auto Scaling group based on a set schedule that you define
- Allows you to set up your own scheduled scaling based on the predictable load changes of your application.

## Month-end Batch Processing Scenario



SCHEDULED SCALING

USE CASES



- **Performs significantly slower** when the month-end financial calculation batch executes
- Causes the CPU utilization of your Amazon EC2 instances to immediately peak to 100% on that period
- **Always happens on the first day of every month** at the stroke of midnight.

- Set a **scheduled scaling policy with a monthly schedule**
- **Scale out before the clock hits 12 midnight** on the first day of the month so there would be **more EC2 instances deployed to handle the peak load**

## Holidays and Public Announcements



### SCHEDULED SCALING

### USE CASES

- Provides a **consistent user experience** by scaling your Auto Scaling group a few hours before your event or specific holidays
- Scaling out your compute capacity takes time due to the **cooldown period**. It may take an hour or more to fully scale your compute capacity to match the current load. This is the reason why **you have to scale-out early!**
- Setting up a scheduled scaling activity beforehand can reduce the performance issues of your application

**Slow site every morning when work day begins...**



## SCHEDULED SCALING

## USE CASES



- **Sluggish application performance right when the workday begins (e.g. 8 AM) but usually runs well by mid-morning (e.g. 10 AM) or at lunchtime**
- **There is a delay in launching new instances as opposed to the number of incoming requests**
- **For example, your Auto Scaling group scales up to 20 or 25 instances during work hours, but scales down to just 2 instances overnight**
- **In the morning, it takes a few hours for the scaling process to complete – extending to mid-morning or till lunchtime, since there are only 2 instances at the start of the day**



# Amazon EC2 Lifecycle Hooks

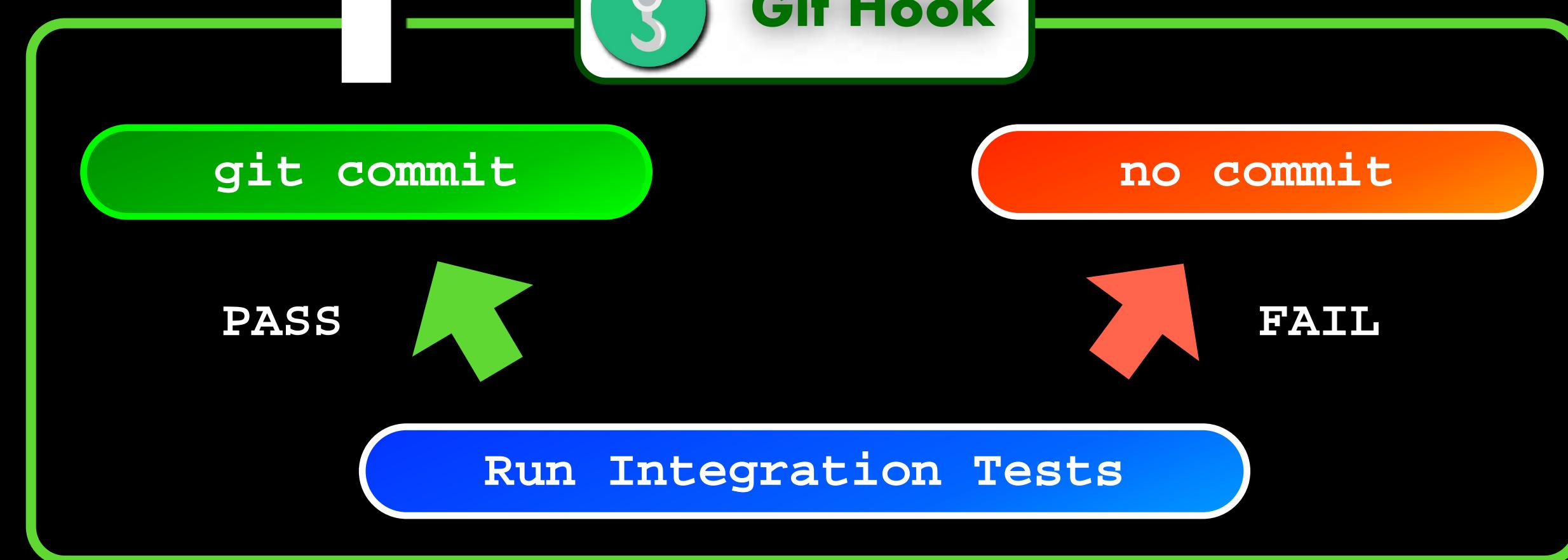
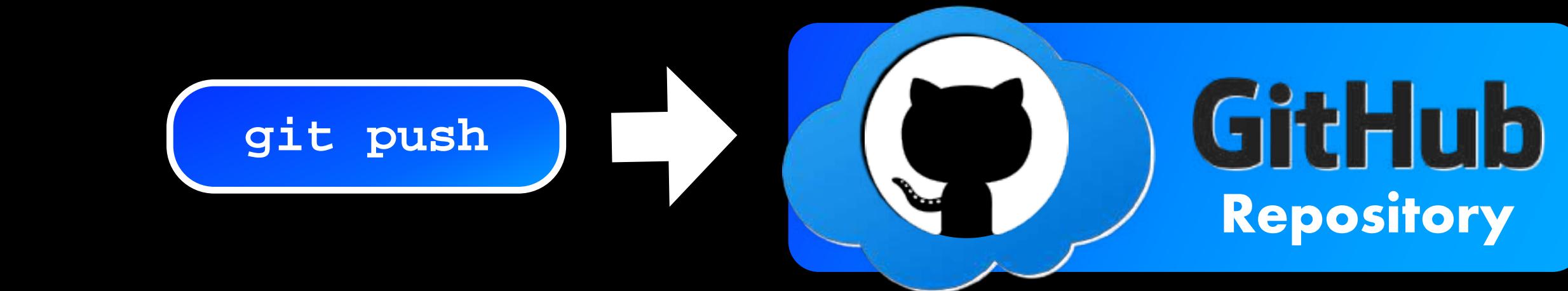


# Hooks

- A **function** that gets executed automatically on a certain event
- Provides the ability to **influence the outcome of your workflow** based on the criteria that you define
- Can stop, skip, or replace the other **function** that is supposed to run on a particular lifecycle
- Also used in some programming languages, version control, and other programs



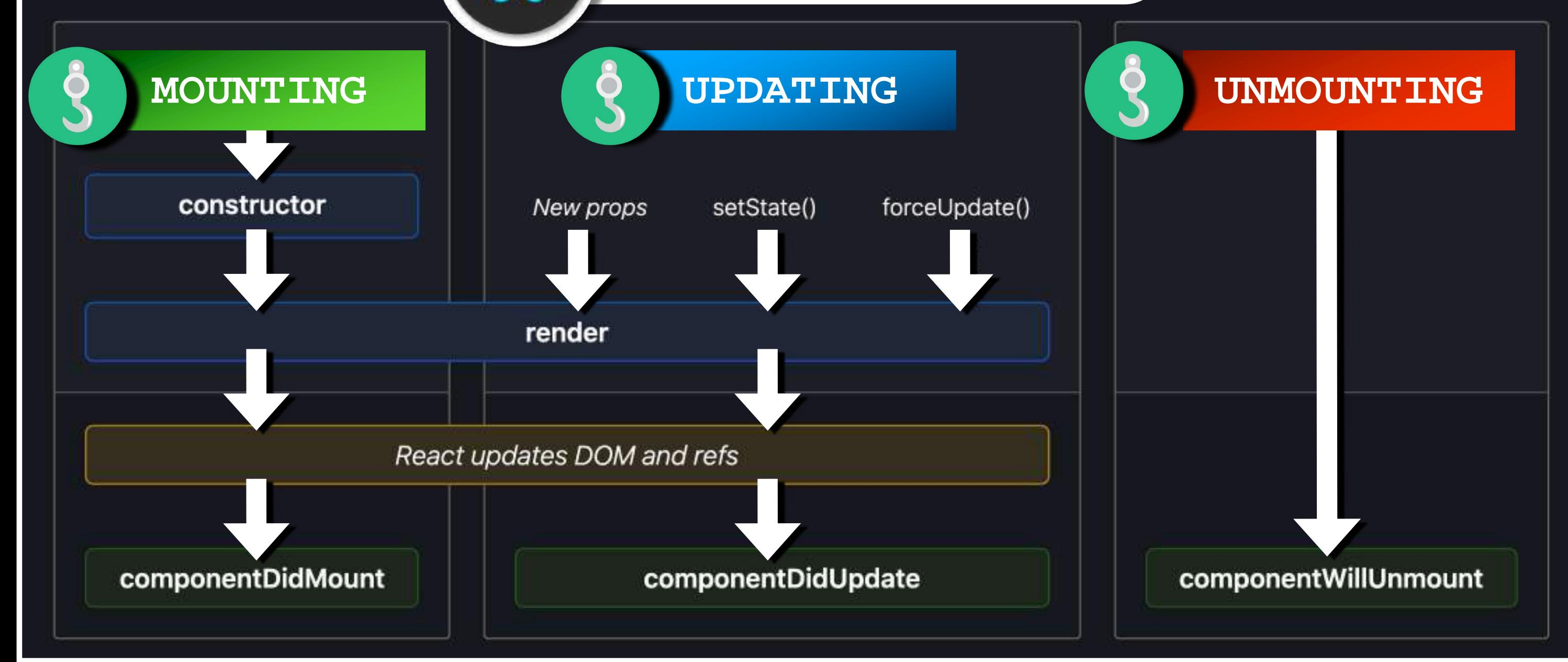
# Hooks



```
git commit -m "TD-1898: Fix javascript race condition in the Tutorials Dojo Portal"  
# `npm test` will run each and everytime time you commit
```



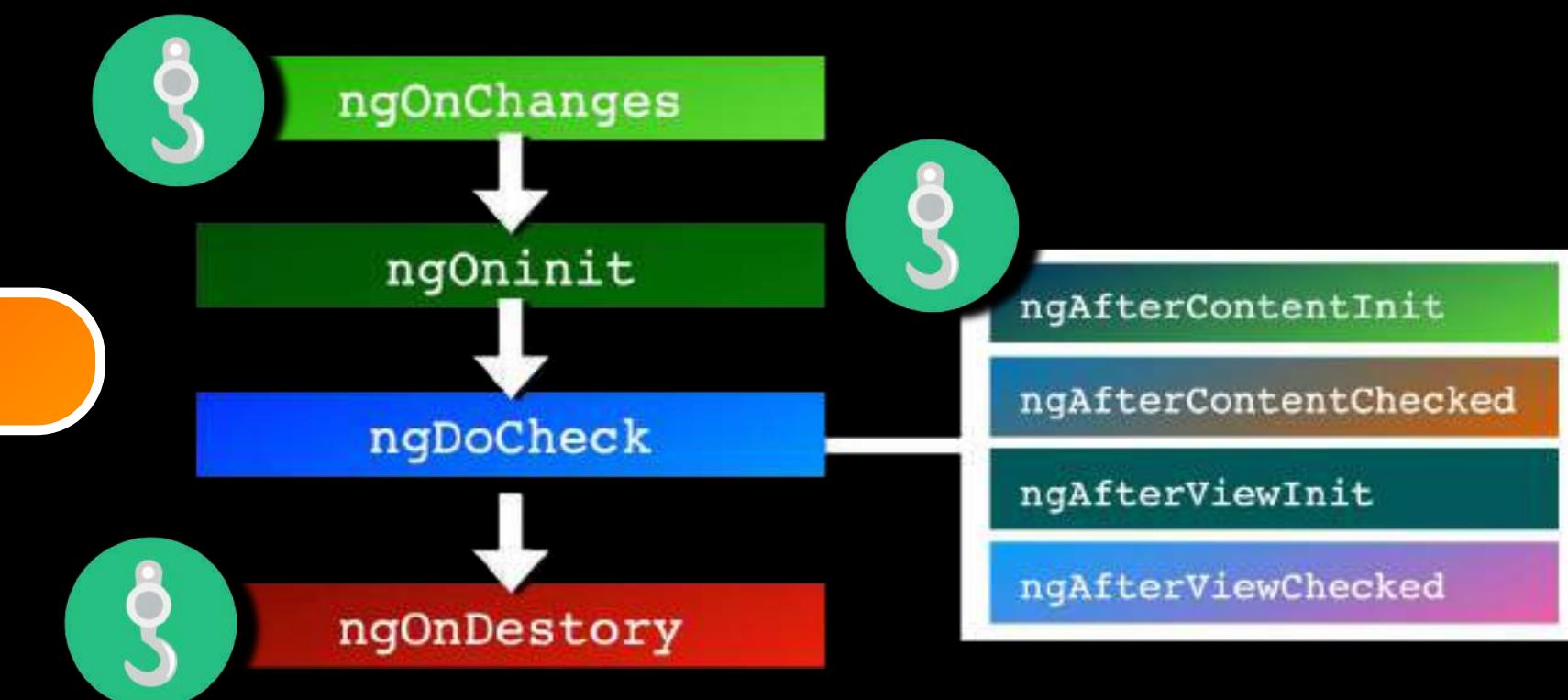
### REACT COMPONENT LIFECYCLE



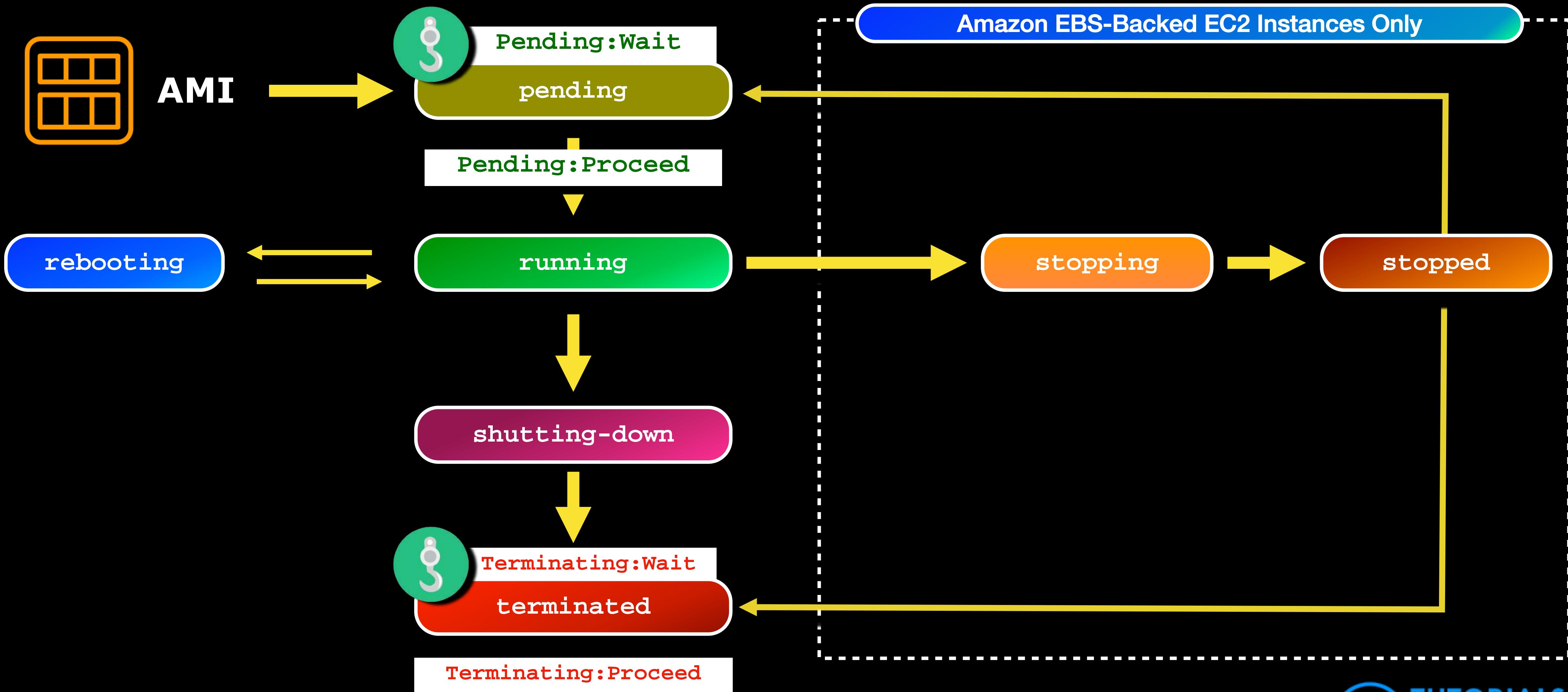
# Hooks



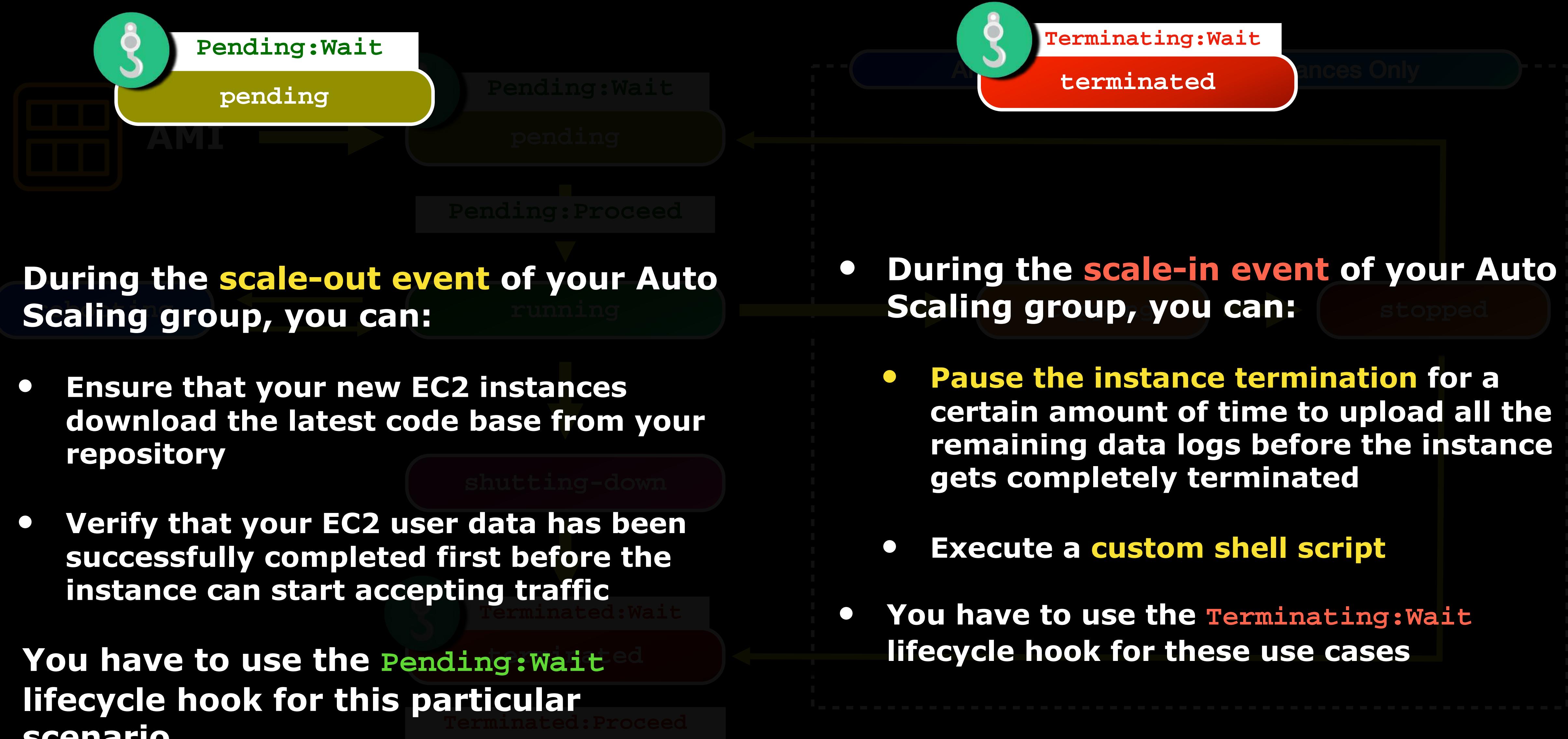
### ANGULAR COMPONENT LIFECYCLE

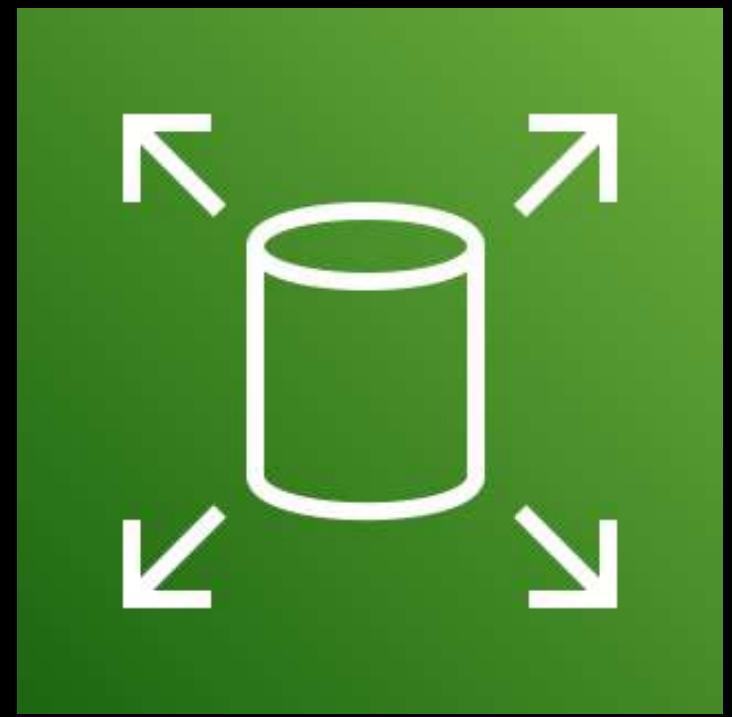


# Amazon EC2 Instance Lifecycle



## Amazon EC2 Instance Lifecycle

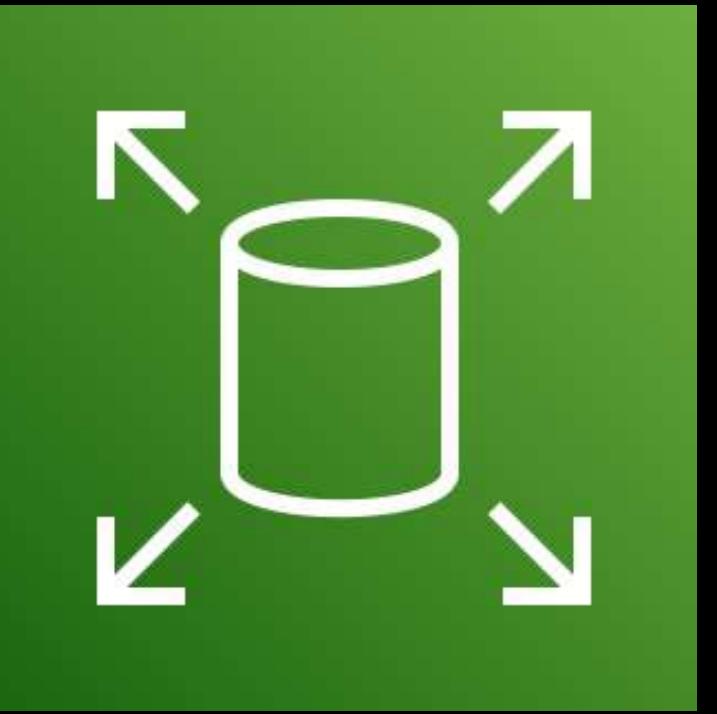




# Amazon EBS Overview

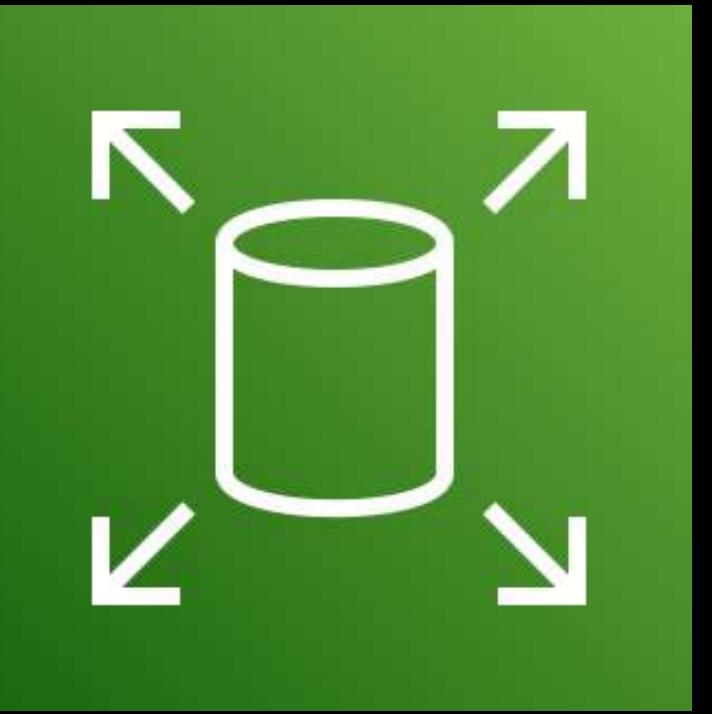
---

- EBS stands for **Elastic Block Store**
- A type of a **block storage** like the Amazon EC2 Instance Store
- Its data is more persistent and **will not get lost even if the EC2 instance was stopped, restarted, or terminated**
- Zonal in scope, which means it **only exists in a single Availability Zone**
- Can be attached to any EC2 instances in the same Availability Zone only
- Can be encrypted at rest using **AWS KMS**
- You can attach one or more Amazon EBS volumes in a single EC2 instance



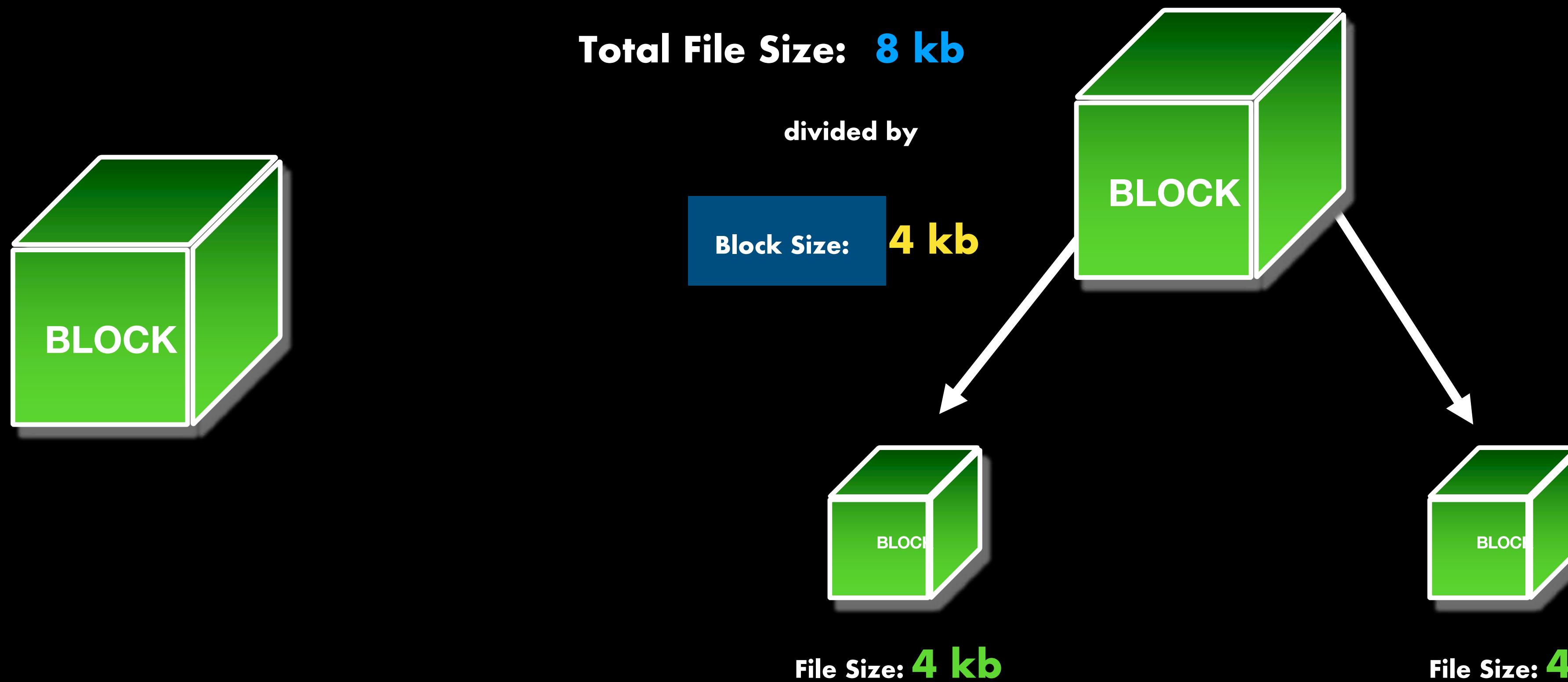
## Amazon EBS

- Suitable for a **variety of workloads** such as databases, enterprise applications, big data analytics engines, file systems, media workflows, and others
- Allows you to store and retrieve your data with **high throughput and low latency**



## Amazon EBS

- Since the underlying physical resources that power your Amazon EC2 instance and EBS volumes are located within the same city or geographic area, Amazon EBS is **capable of providing low latency read or write access to your data**
- Mainly operates on the **hardware level**



Administrator: Windows PowerShell

Windows PowerShell  
Copyright (c) Microsoft corporation. All rights reserved.

Try the new cross-platform PowerShell <https://aka.ms/pscore6>

```
PS C:\WINDOWS\system32> Get-ciminstance -Class Win32_Volume | Select-Object Label, BlockSize | Format-Table -AutoSize
```

Label	BlockSize
DATA	4096
OS	4096
RECOVERY	4096
SYSTEM	4096

**Block Size: 4096 bytes**

```
PS C:\WINDOWS\system32>
```

A screenshot of a macOS Terminal window titled "PILIPINAS — Terminal — -zsh — 75x5". The window shows the command `jonbonso@tutorialsdojo > diskutil info / | grep "Device Block Size"`. The output displays the line `Device Block Size: 4096 Bytes`, which is highlighted with a green rectangular border. To the right of the terminal window, there are two colored boxes: a blue one containing the text "Block Size:" and an orange one containing the text "4096 bytes" in yellow.

```
[jonbonso@tutorialsdojo > diskutil info / | grep "Device Block Size"]
Device Block Size: 4096 Bytes
jonbonso@tutorialsdojo >
```

**Block Size:** **4096 bytes**

# RAID

**Redundant Array of Independent Disks**

## RAID 0

- **Stripes multiple volumes together**
- **Provides greater I/O performance**
- **Divides a body of data into blocks and then spreads the data blocks across multiple storage devices**
- **Suitable if I/O performance is your priority**

## RAID 1

- **Mirrors two or more volumes together**
- **Provides on-instance redundancy**
- **Duplicates data to provide more durability and availability**
- **Suitable if data redundancy is your focus**

**Select RAID type**

Select the type of RAID you would like to create using Disk Utility.  
Different RAID types can provide different levels of data protection or performance.

**• Striped (RAID 0)**

Splits data evenly across two or more disks, without parity of information, with speed as the intended goal.

**Mirrored (RAID 1)**

An exact copy of a set of data on two or more disks. This type is useful when read performance or reliability is more important.

**Concatenated (JBOD)**

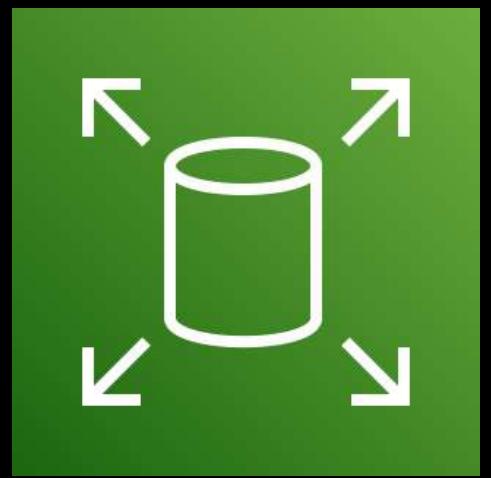
Concatenated disks is not a RAID, it is a group of disks connected together for the purpose of creating a larger disk.

?

Cancel

Previous

Next



**Amazon Elastic Block Store  
(Amazon EBS)**

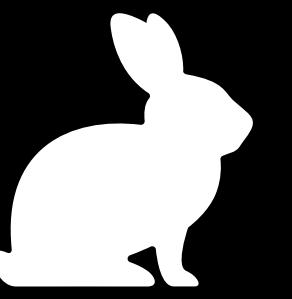


**Solid State Drive  
(SSD)**

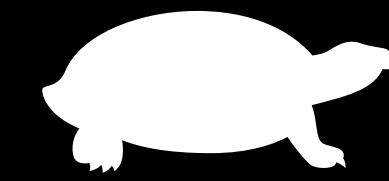


**Hard Disk Drive  
(HDD)**

### Read & Write Speeds



**Fast !**



**Slow...**

### Use Case

For workloads with  
frequent read/write operations

For **data archiving, backups**  
or throughput-oriented storage

### Dominant Performance Attribute

#### IOPS

Input/Out operations Per Second

#### Throughput

Megabit per second (Mbps)

Can be used as  
Boot Volume for



**Yes**

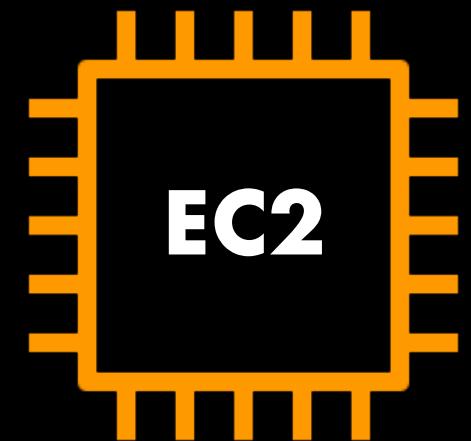
**No**

- An incremental backup that internally uses Amazon S3 to persist your data
- It only saves the data blocks that have changed after your most recent snapshot
- Allows you to restore the state of your EBS volume in the event of data loss
- Enables you to copy your EBS volume to another AWS Region for your data migration, disaster recovery activities
- Can be used to encrypt an unencrypted Amazon EBS volume.
- Automate the creation, retention, and deletion of your EBS snapshots and EBS-backed AMIs using the Amazon Data Lifecycle Manager (Amazon DLM) service



## Amazon EBS Snapshots

### ENCRYPTION IN TRANSIT

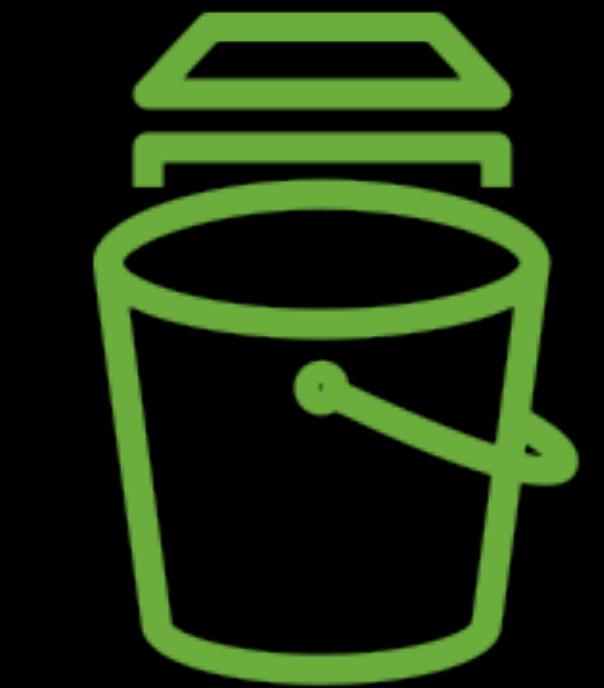
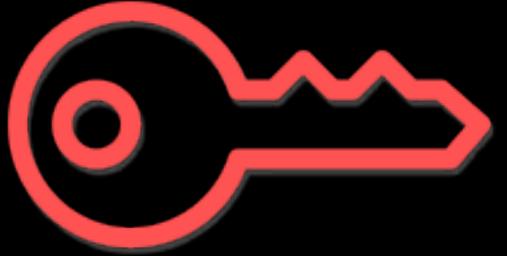


**AMAZON EBS  
VOLUME**

### ENCRYPTION AT REST



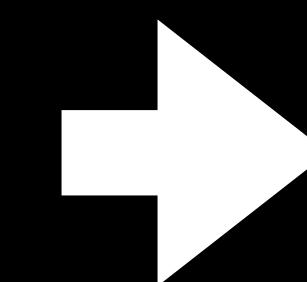
AWS KMS Keys



**AMAZON EBS  
SNAPSHOT**

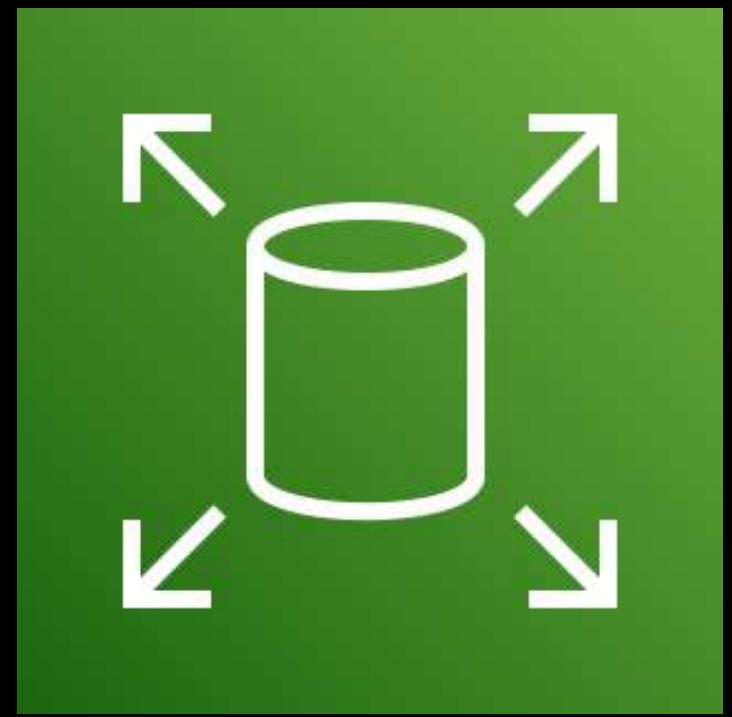
**Amazon EBS Encryption  
by Default**

Must be manually enabled per AWS Region



**INTERNAL AMAZON  
S3 BUCKET**

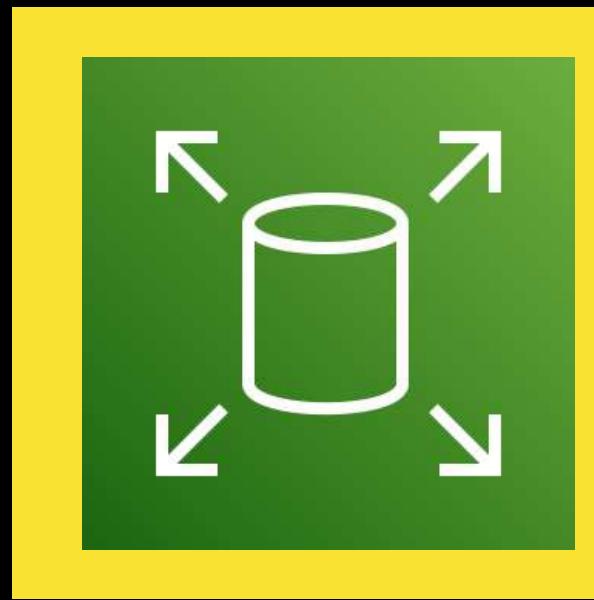
Exclusively managed by AWS



# Amazon EBS Types

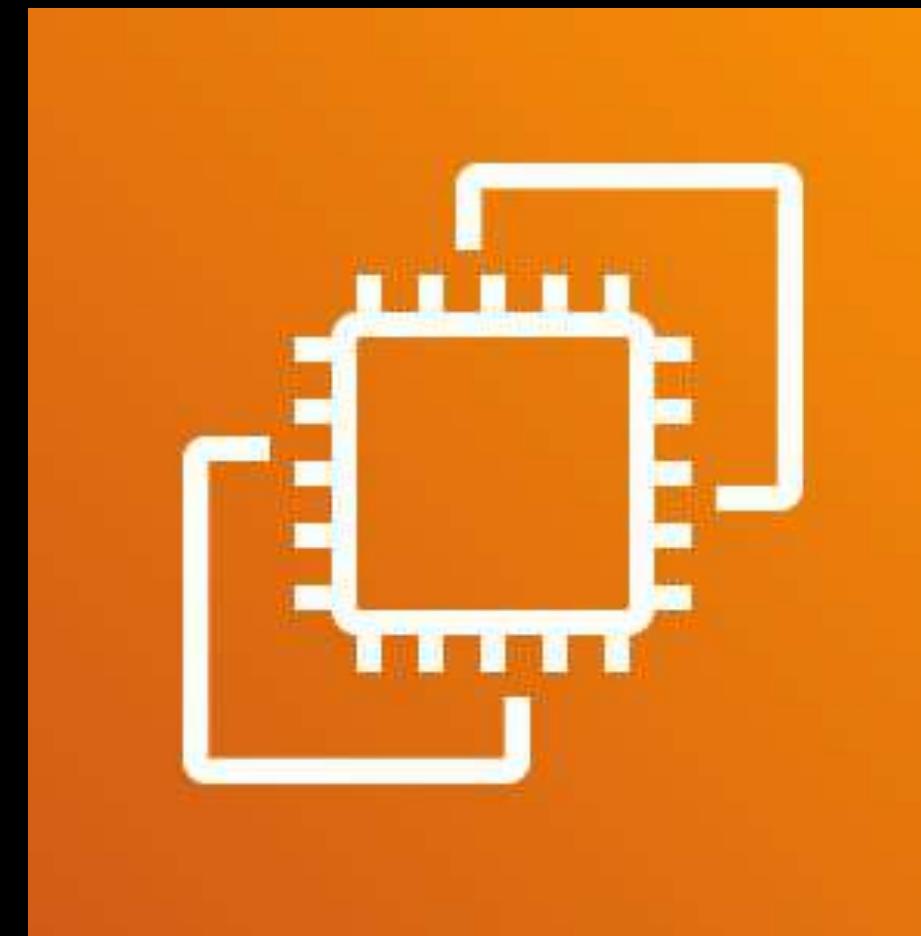
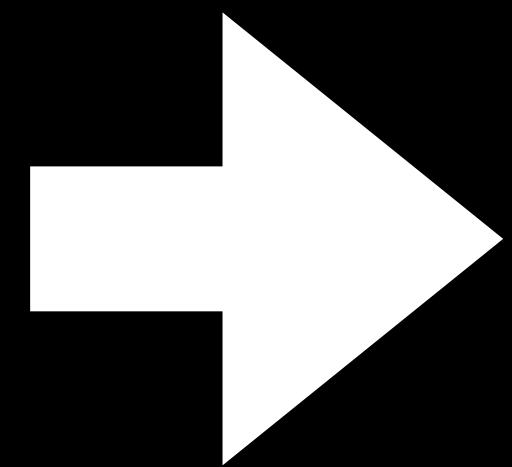
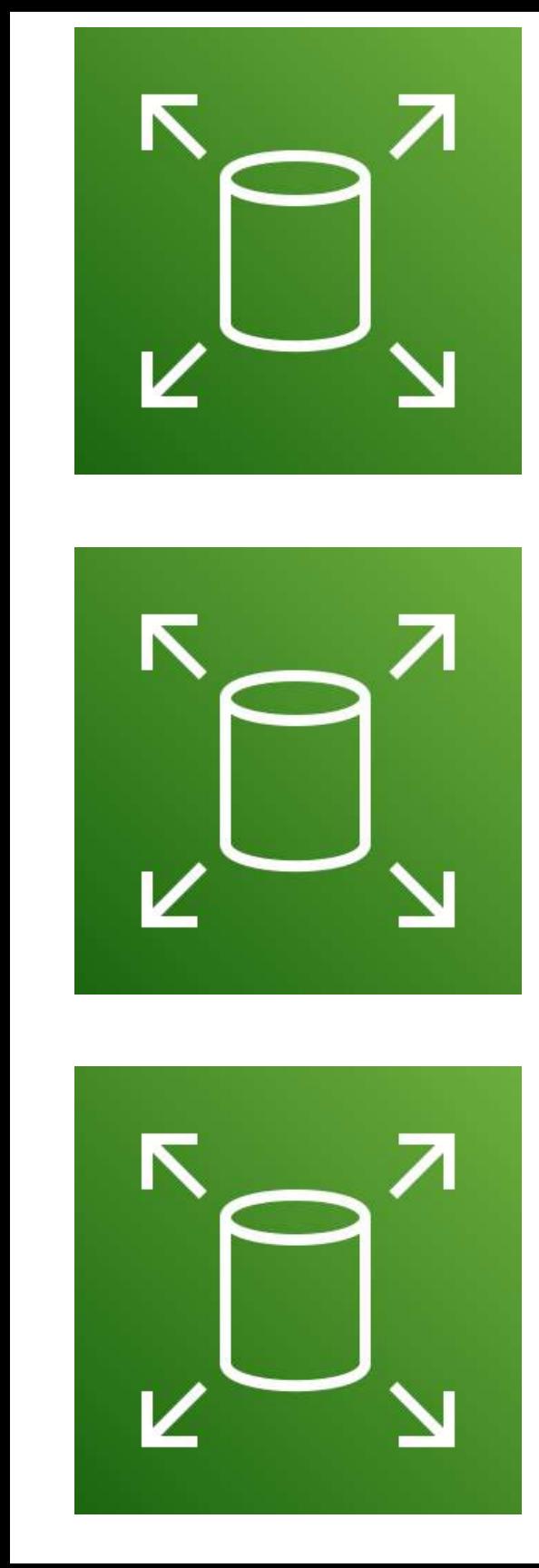
---

## ROOT EBS VOLUME



\* contains the system image for  
booting the EC2 instance

## OTHER DATA VOLUMES



Amazon EC2 Instance



**Solid State Drive  
(SSD)**



**Hard Disk Drive  
(HDD)**



## Solid State Drive (SSD)

- Suitable for **transactional workloads**
- For various types of applications and systems with **frequent read/write operations with small I/O sizes**
- Performance Attribute: **IOPS**



**Solid State Drive  
(SSD)**



**General Purpose SSD**



**Provisioned IOPS SSD**



## Solid State Drive (SSD)



### General Purpose SSD



### Provisioned IOPS SSD

- Provides a **balance of price and performance for your workloads**
- Recommended for **most workloads**
- Also suitable for apps with **unpredictable or unknown access patterns**
- **Provides a configurable and consistent IOPS to allow you to accommodate the changes in your data storage requirements**



## Solid State Drive (SSD)

- Suitable for **low-latency interactive apps** in production as well as your development and test environments



## General Purpose SSD

- For your **infrequently accessed applications or systems that:**

- Only peaks during certain times of the day
  - Has a varying Disk I/O operations



## Provisioned IOPS SSD

- Provides ample IOPS for your applications but **not on par** with what a Provisioned IOPS type can give
- The **most cost-effective storage option that does NOT sacrifice performance**



## Solid State Drive (SSD)



### General Purpose SSD

- Primarily used for **mission-critical, low-latency, or high-throughput workloads**

- Provides **sub-millisecond latency and consistent IOPS performance**

- Allows you to set the amount of available IOPS of your EBS volume



### Provisioned IOPS SSD

- For hosting data to your application that makes small reads and writes to a small file system

- For applications that require high read and write IOPS



## Solid State Drive (SSD)

- For hosting data to your **applications** that makes **small reads and writes to a small file system**



## General Purpose SSD

- For applications that require a number of **high read and write IOPS performance**



## Provisioned IOPS SSD

- For fixing **latency issues**
- For scenarios where your **database storage performance is the bottleneck**
- For storage systems that require a **configurable and consistent IOPS**
- . . . and many more!

## Amazon EBS Multi-Attach



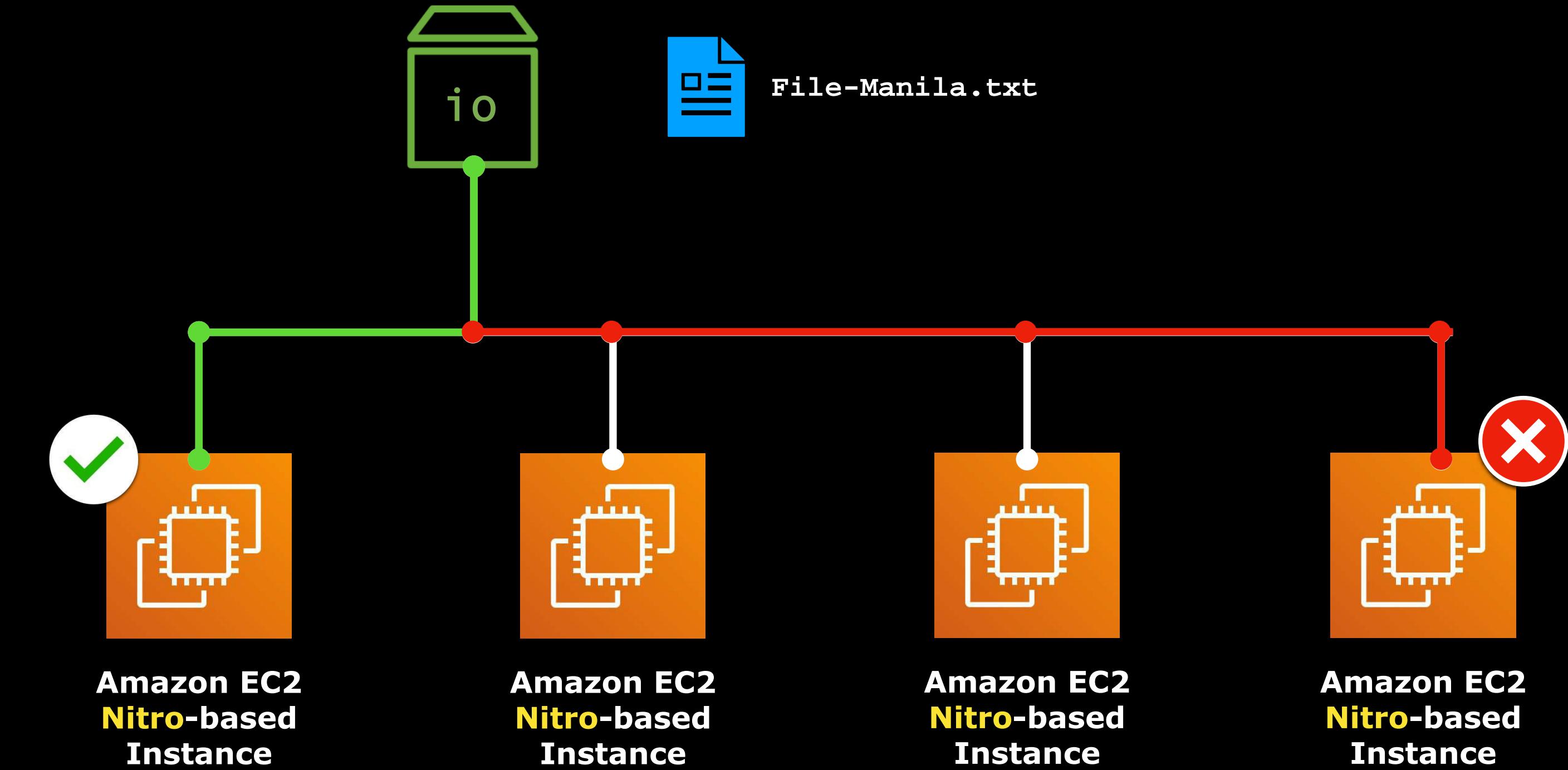
**Solid State Drive  
(SSD)**



**General Purpose SSD**



**Provisioned IOPS SSD**



**No concurrent file modification**



## Hard Disk Drive (HDD)

- Optimized for **large streaming workloads**
- For various types of applications and systems with **large, sequential I/O operations**
- Performance Attribute: **Throughput (MB/s)**



## Hard Disk Drive (HDD)



Throughput Optimized HDD



Cold HDD



## Hard Disk Drive (HDD)



### Throughput Optimized HDD



### Cold HDD

- A low-cost HDD designed for frequently accessed, throughput-intensive workloads
- Can be used for your Big data applications, Data Warehouses, and Log Processing
- Cannot be used as your boot (root device) volume



## Hard Disk Drive (HDD)



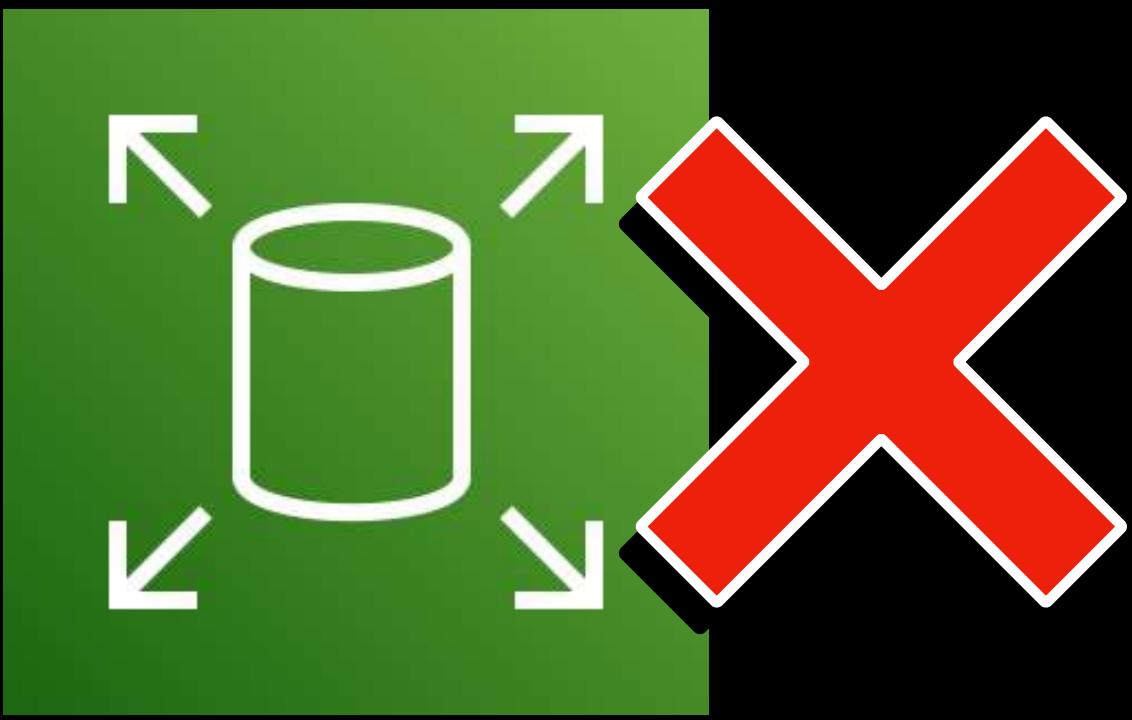
### Throughput Optimized HDD



### Cold HDD

- **Lowest-cost HDD storage type**
- **Meant for storing less frequently accessed workloads**
- **The most cost-effective storage EBS type** option for data archiving only since its throughput performance is substantially low
- **Suitable for throughput-oriented storage for data that is infrequently accessed**
- **Perfect for scenarios where the lowest storage cost is of the utmost importance**

- If you just need a **temporary storage** for your data, use EC2 Instance Store instead
- If you have to store your application or system data in a **POSIX-compliant hierarchical directory structure** (use Amazon EFS instead)
- If you have multiple applications that are **concurrently accessing the same files at the same time**, it is better to use the Amazon EFS or Amazon FSx service instead
- If you need to **store your static data in the most cost-effective way**, it's more appropriate and cheaper to store them in Amazon S3



## ANTI-PATTERNS

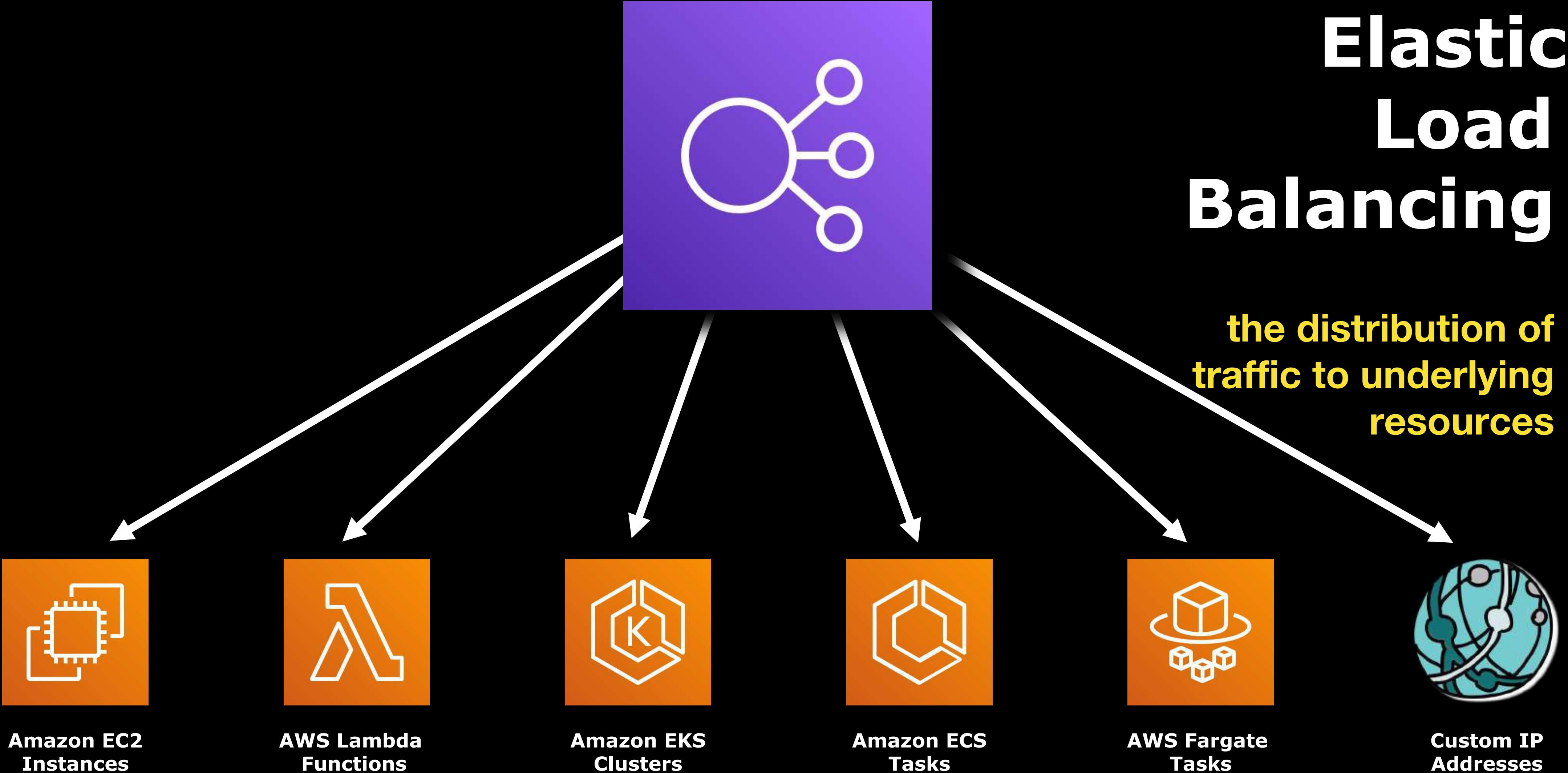


# Amazon Elastic Load Balancing Overview

---

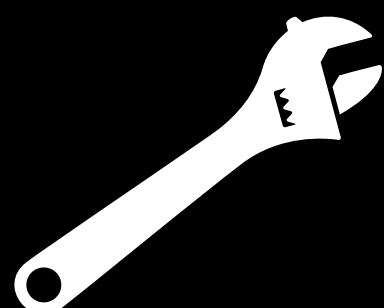
# Elastic Load Balancing

the distribution of traffic to underlying resources





52.44.107.223

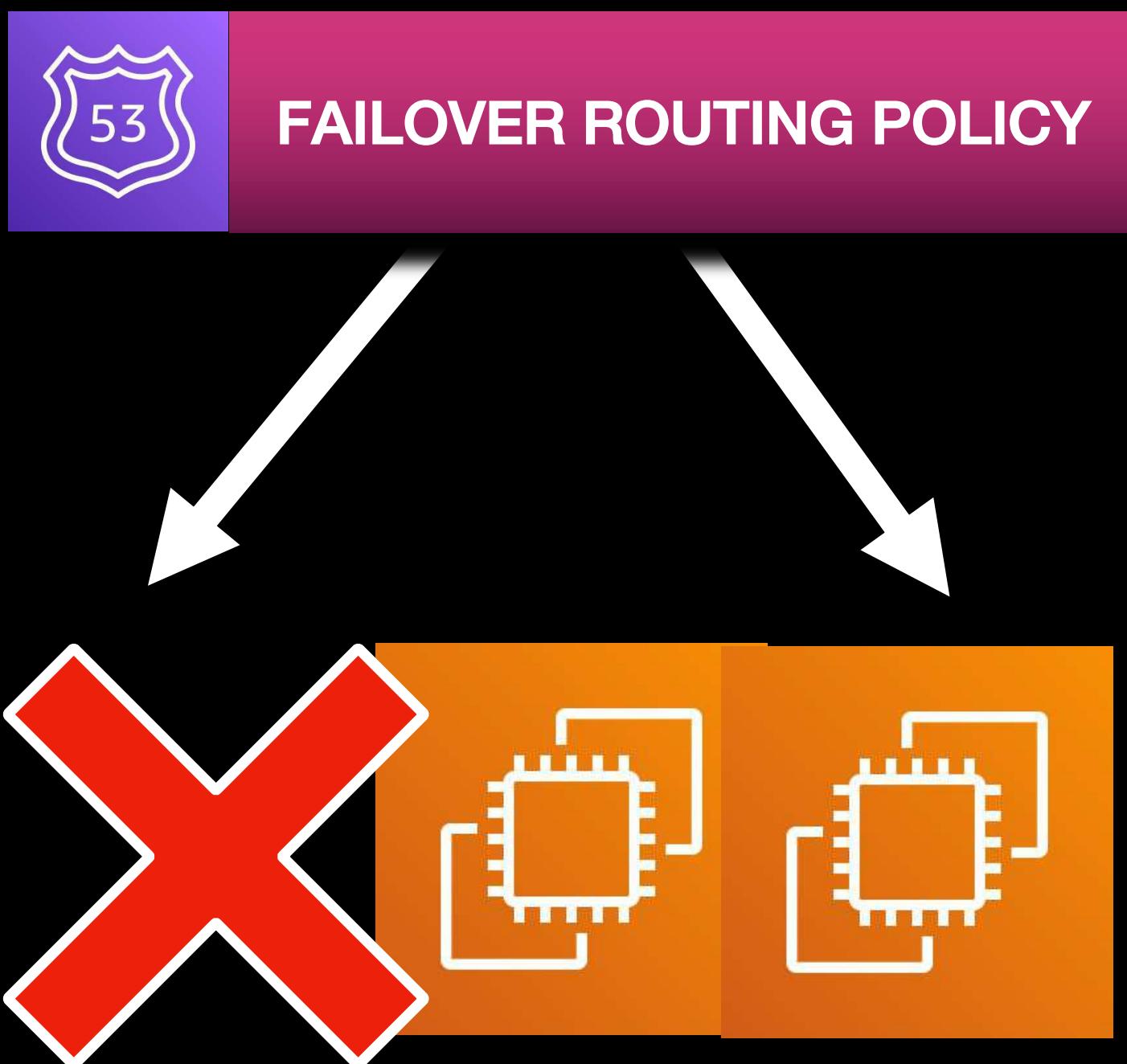
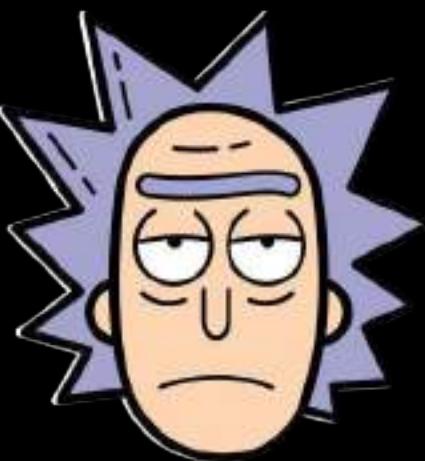


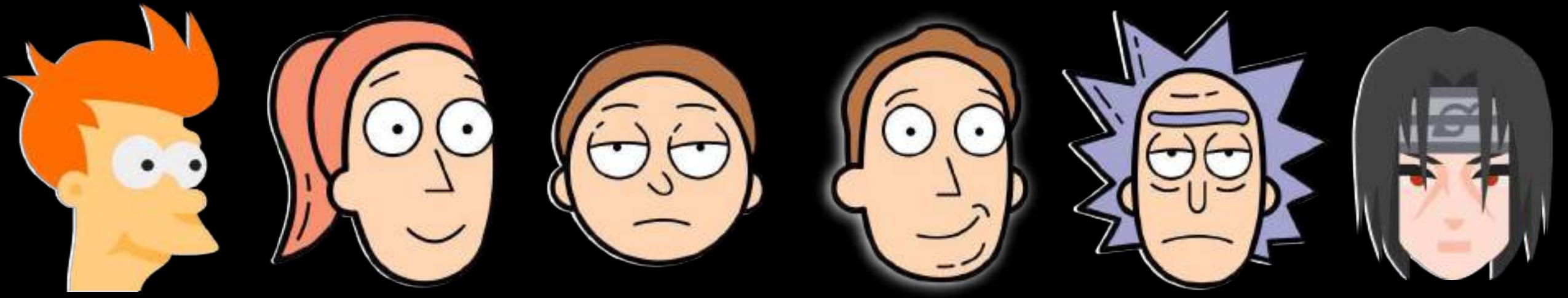
OS Patching or  
System Maintenance



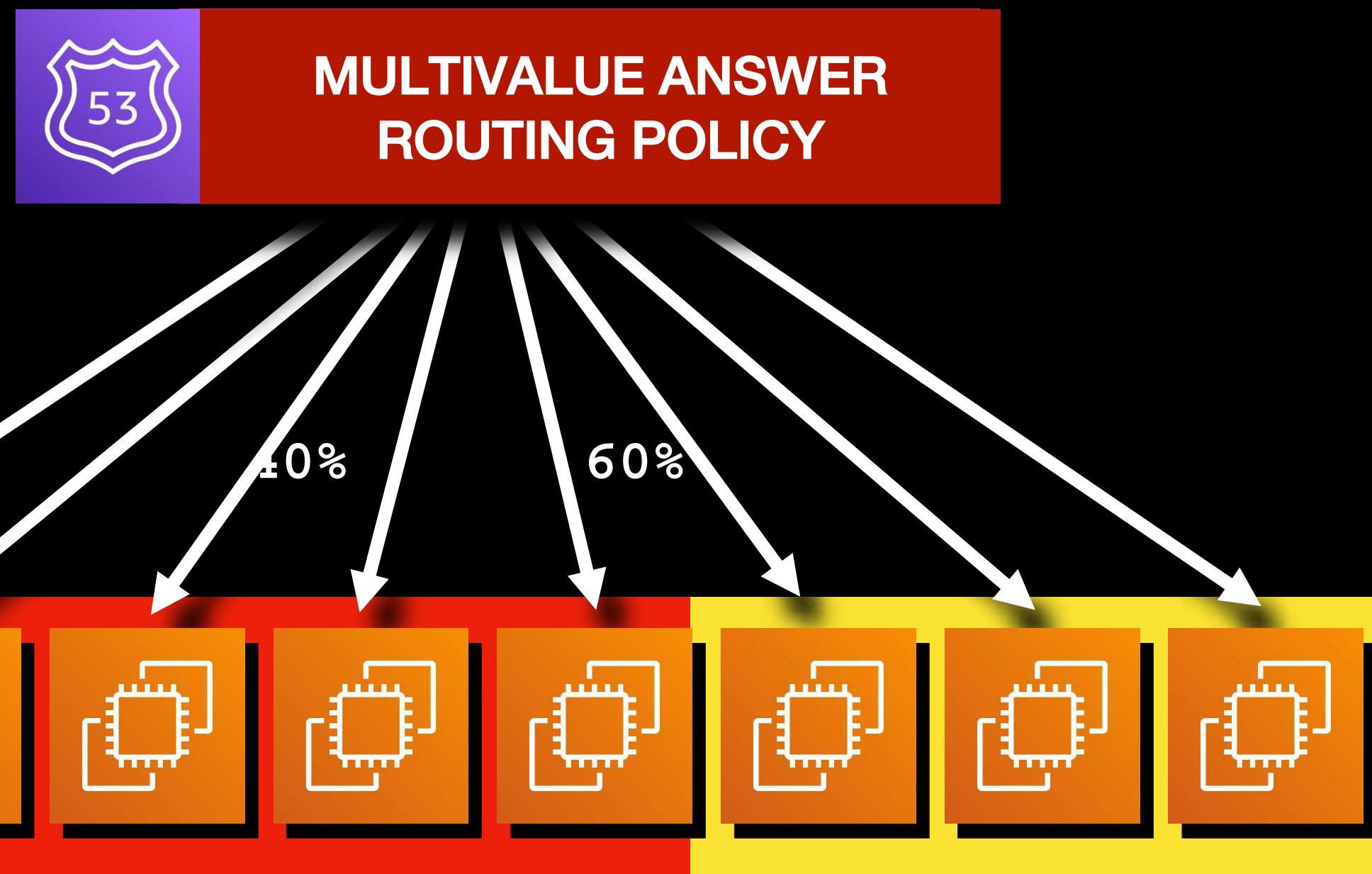
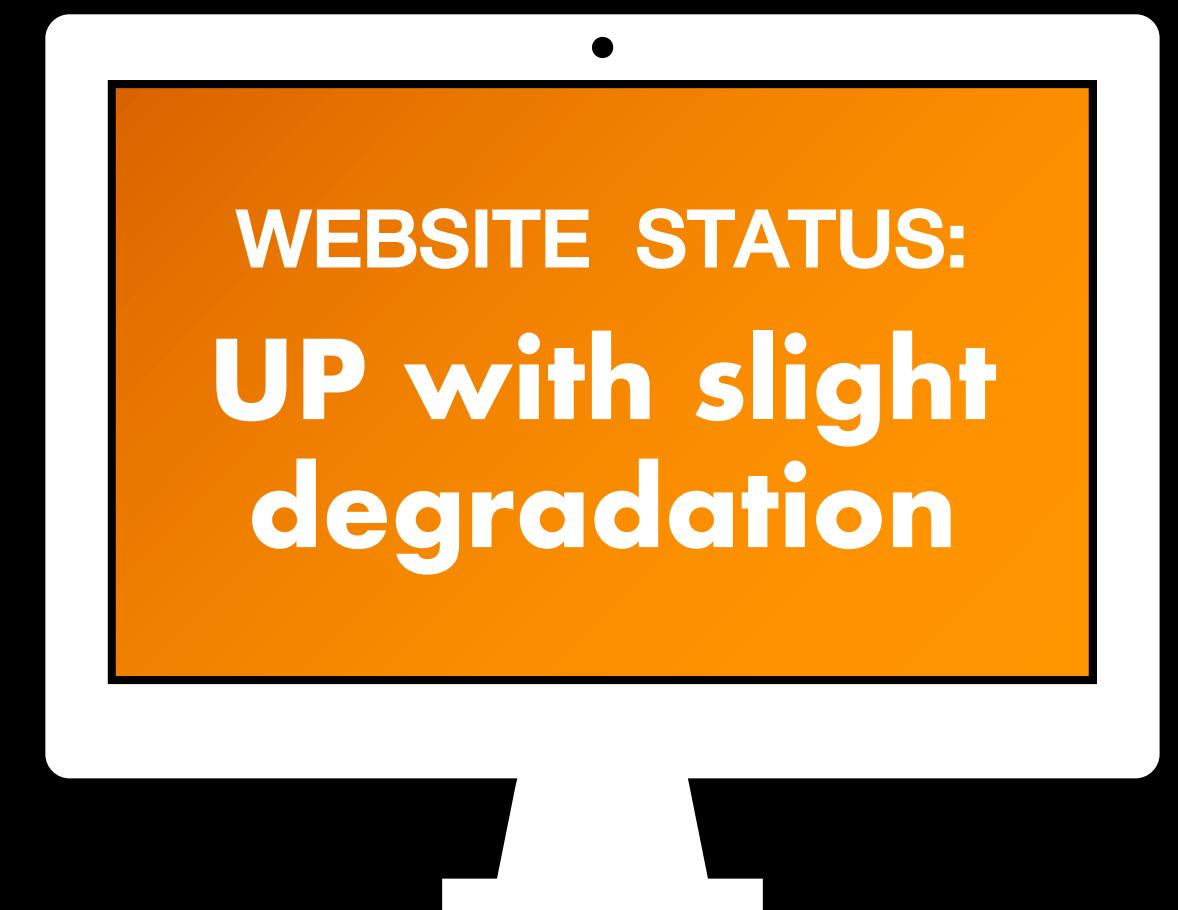
Critical Application  
or System Errors







Incoming Load of Traffic



- The distribution of the **incoming load traffic is not balanced** across the underlying servers
- The traffic is distributed **randomly**
- **Unbalanced** - Some servers are **overutilized** while others are **underutilized**
- **No routing algorithm**
- **Lacks security features**



CPU Utilization: **Over 100%**



CLOUD

REGION



AZ 1



AZ 2

**Load Balancer**

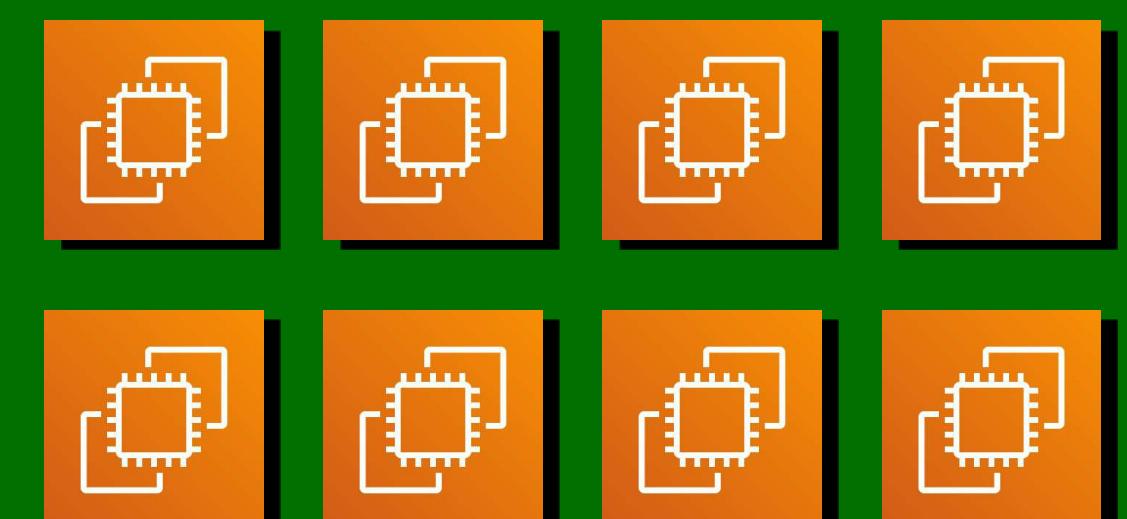
Public subnet A

10.0.1.0/24

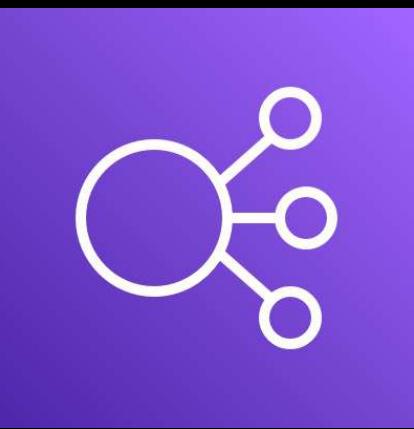


Public subnet B

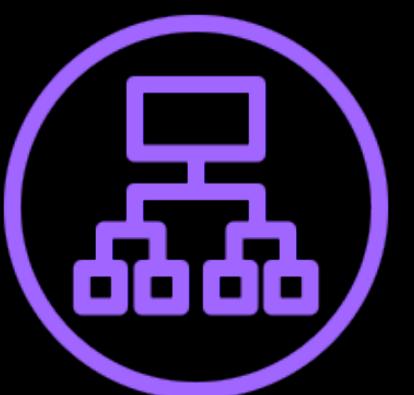
10.0.1.0/24



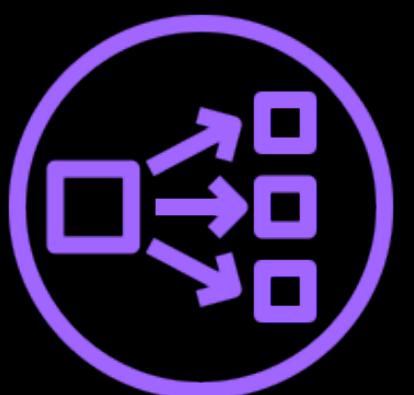
**Balanced distribution  
of incoming traffic  
through the use of  
routing algorithm**



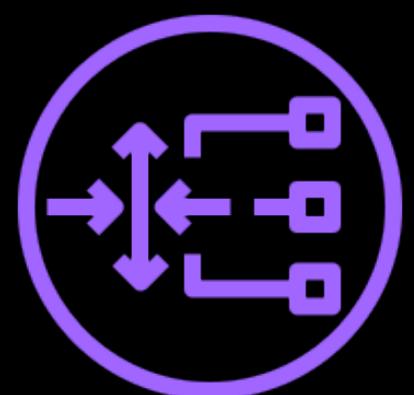
## Elastic Load Balancing TYPES



### Application Load Balancer ( ALB )



### Network Load Balancer ( NLB )



### Gateway Load Balancer ( GWLB )



### Classic Load Balancer ( CLB )

#### ROUTING ALGORITHM

Round Robin  
Least Outstanding  
Requests (LOR)

Flow Hash

IP Listener Routing that  
leverages on GENEVE  
protocol

Round Robin  
Least Outstanding  
Requests (LOR)

#### PROTOCOL LISTENERS

HTTP / HTTPS  
gRPC

TCP / UDP  
TLS

IP

HTTP / HTTPS  
TCP  
SSL/TLS

#### USE CASES

For web apps,  
microservices  
& containers

Handling  
millions of requests  
per second while  
maintaining  
ultra-low latencies

Running third-party  
virtual appliances  
in AWS

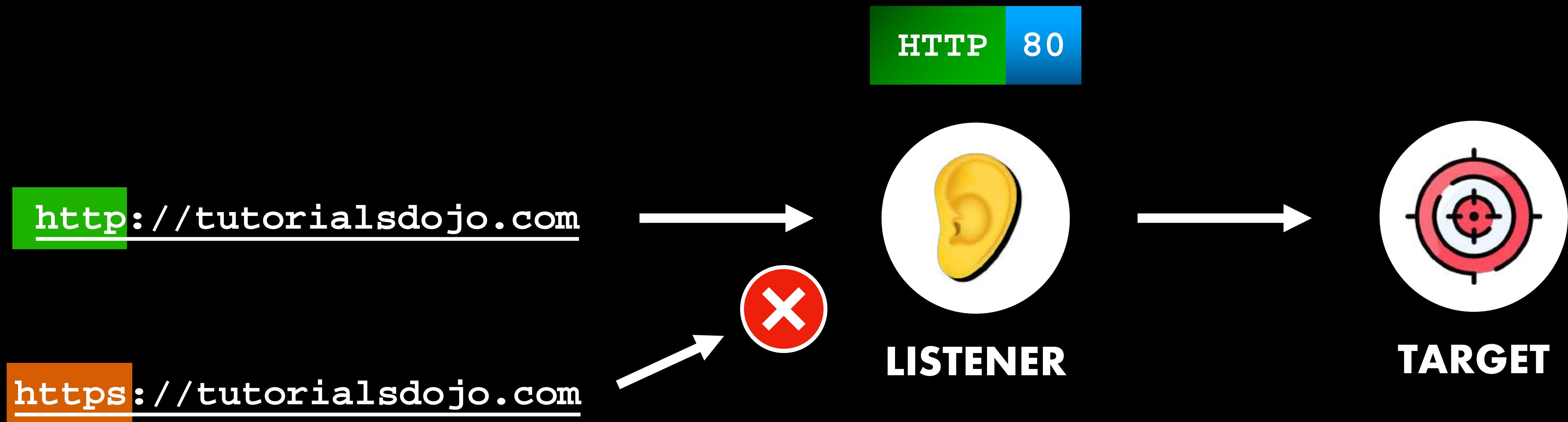
For legacy applications  
in AWS  
  
For implementing  
Custom Security Policies  
and  
TCP passthrough  
configuration

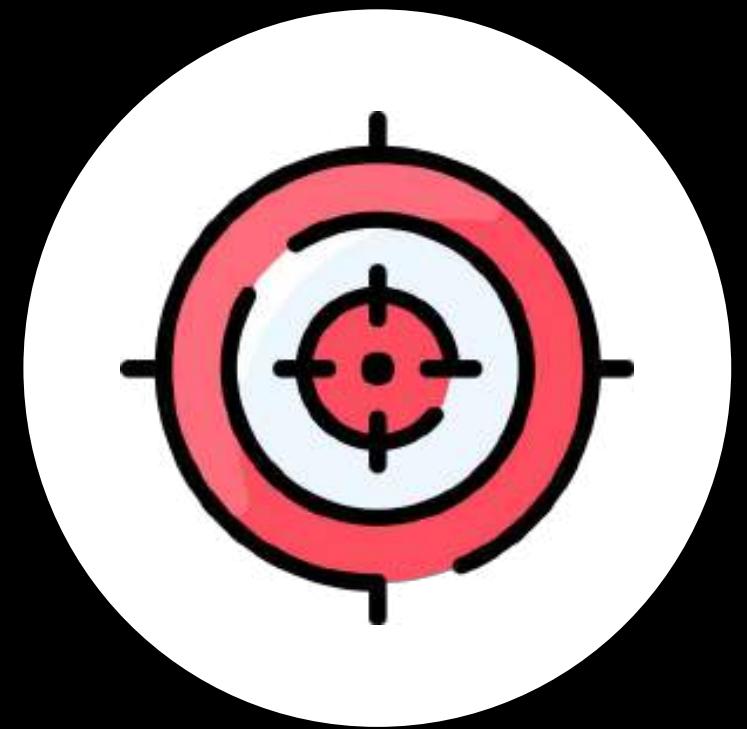


**LISTENER**

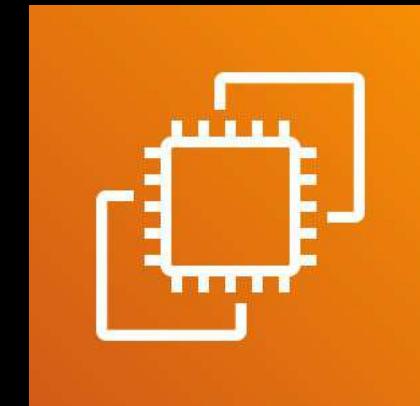


**TARGET**





**TARGET**



**Amazon EC2 Instances**



**AWS Lambda Functions**



**Amazon EKS Clusters**



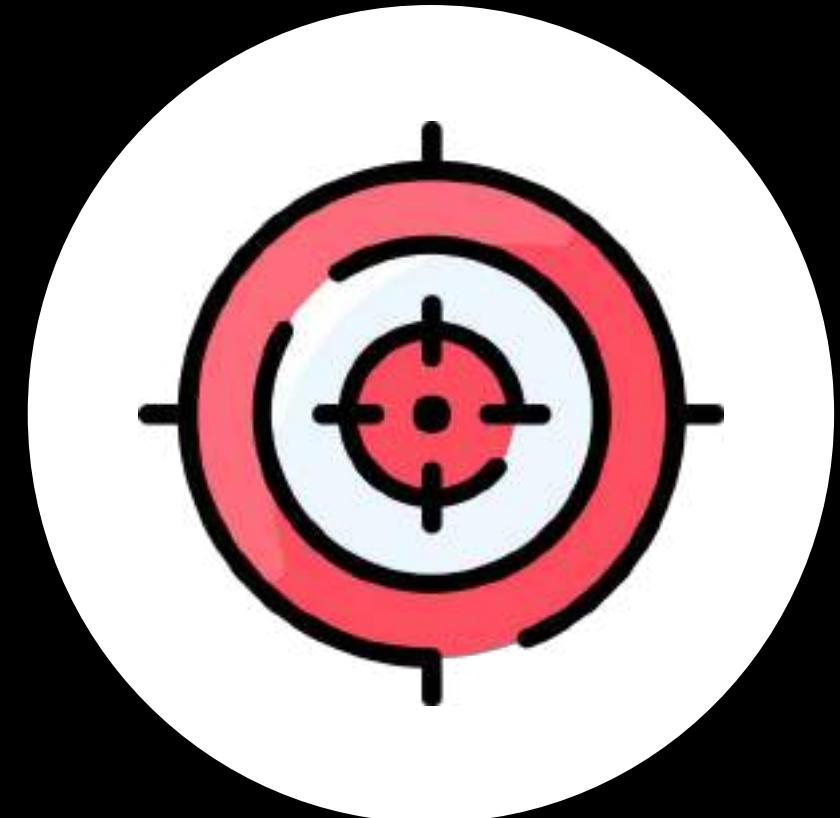
**Amazon ECS Tasks**



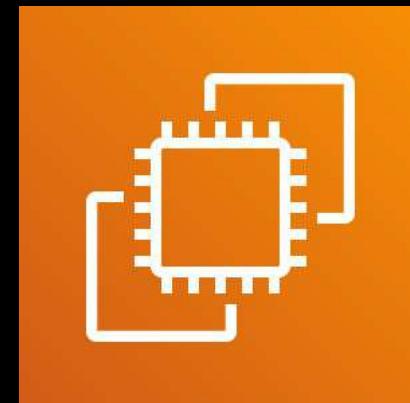
**AWS Fargate Tasks**



**Custom IP Addresses**



## TARGET GROUP



Amazon EC2  
Instances



AWS Lambda  
Functions



Amazon EKS  
Clusters



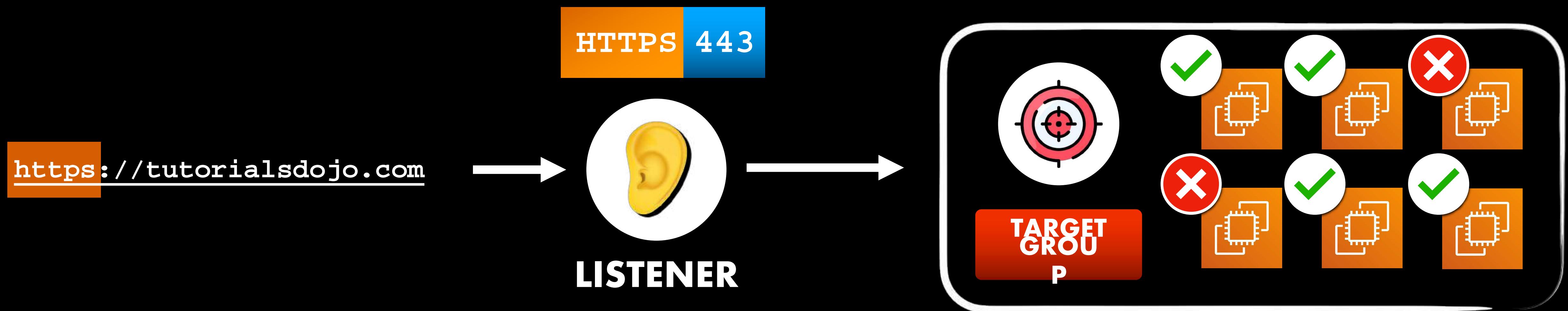
Amazon ECS  
Tasks



AWS Fargate  
Tasks



Custom IP  
Addresses





CLOUD

## US-EAST-1 REGION



ELB

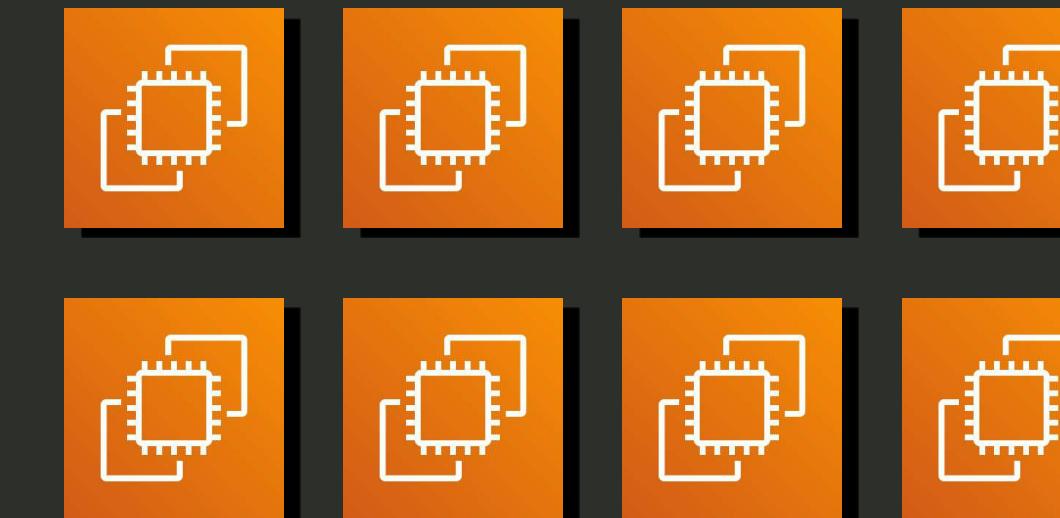
Public subnet A 10.0.1.0/24

TARGET GROUP

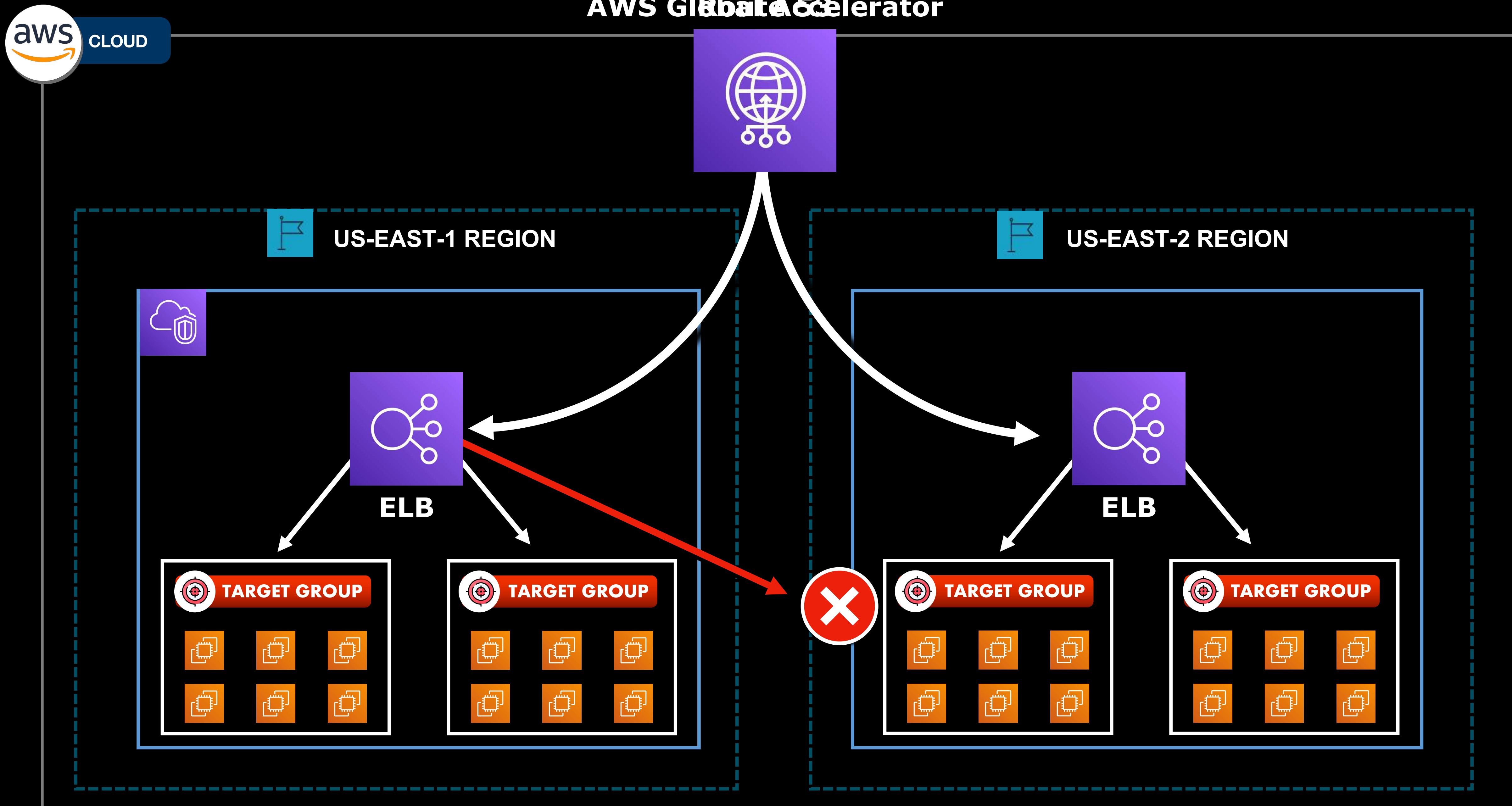


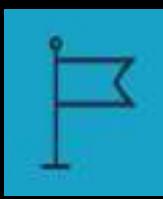
Public subnet B 10.0.1.0/24

TARGET GROUP



# AWS Global Accelerator

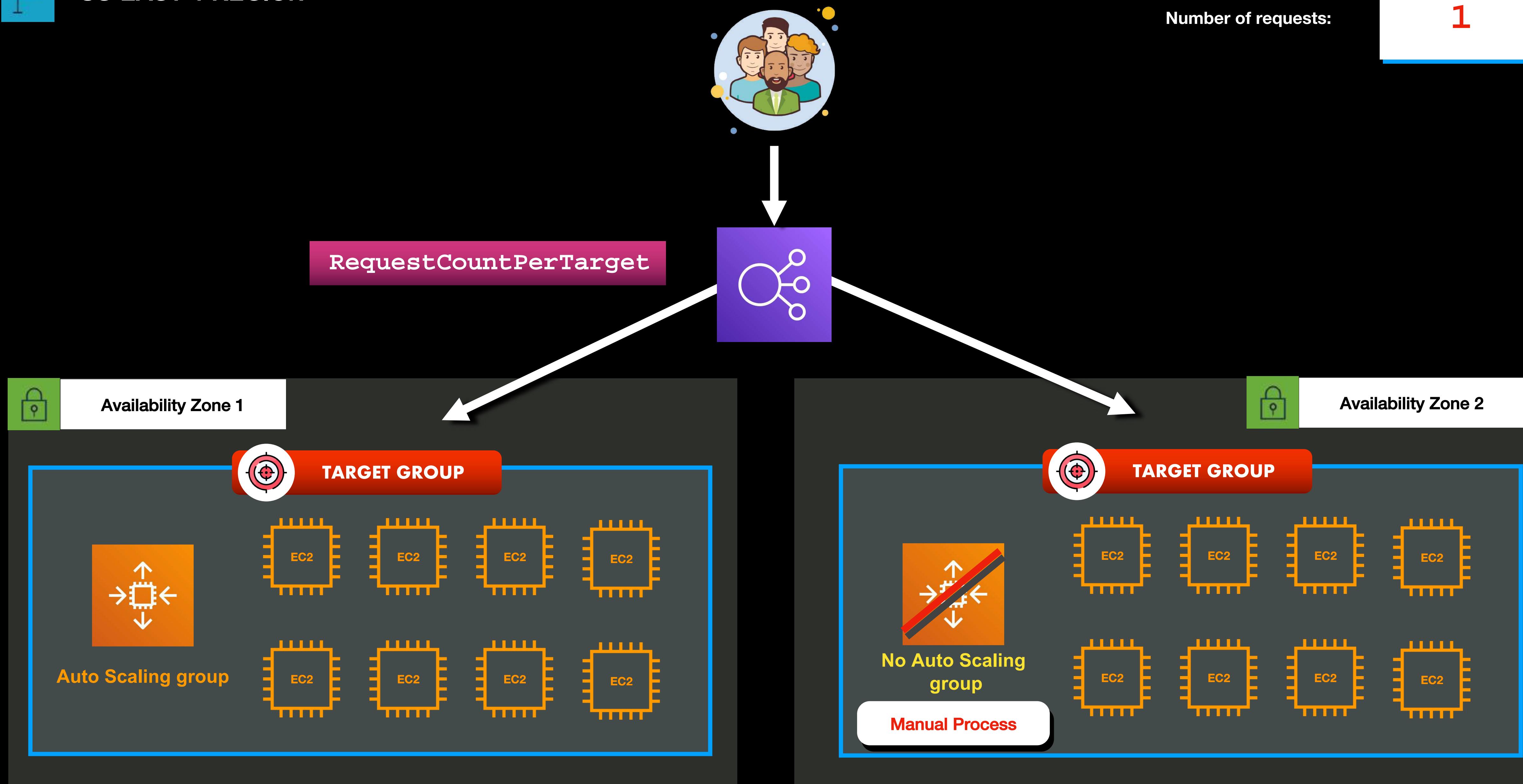




US-EAST-1 REGION

Number of requests:

1

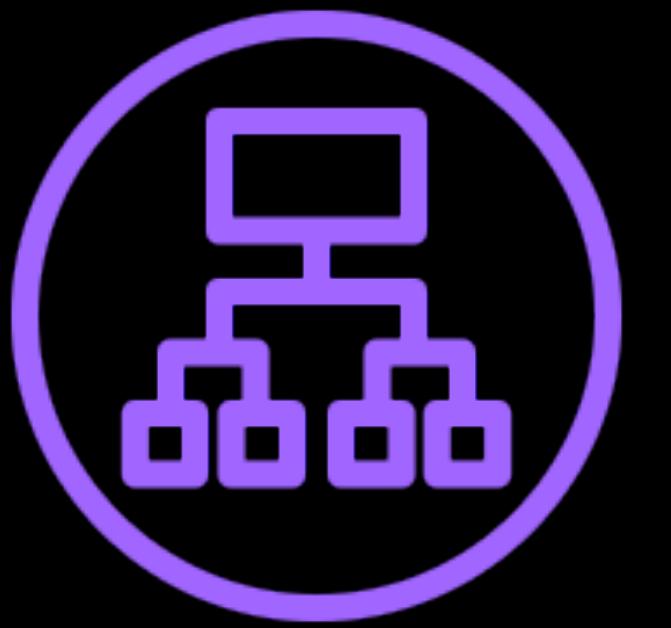




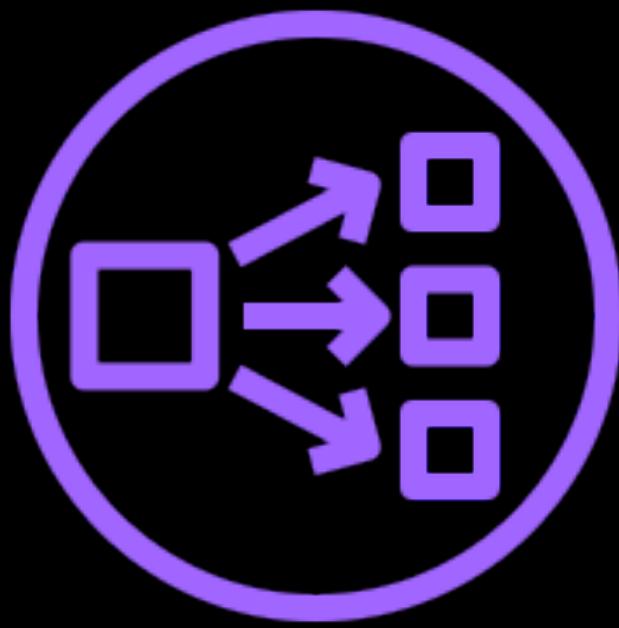
# Amazon Elastic Load Balancing

## TYPES

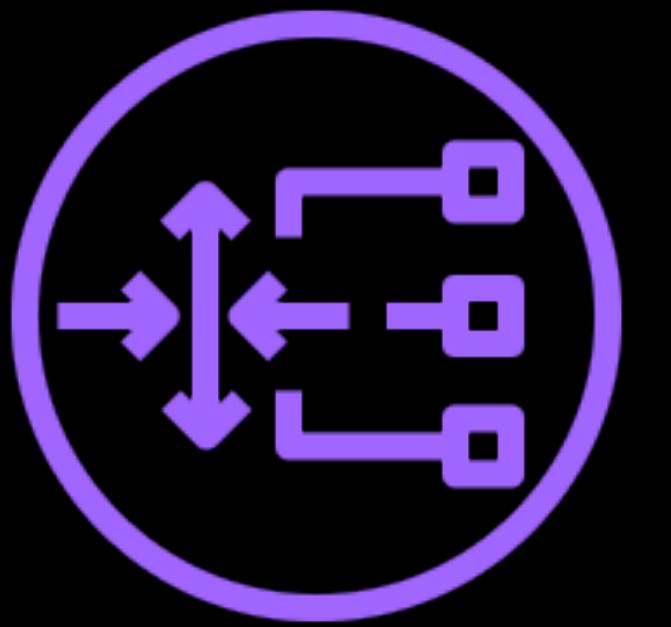
---



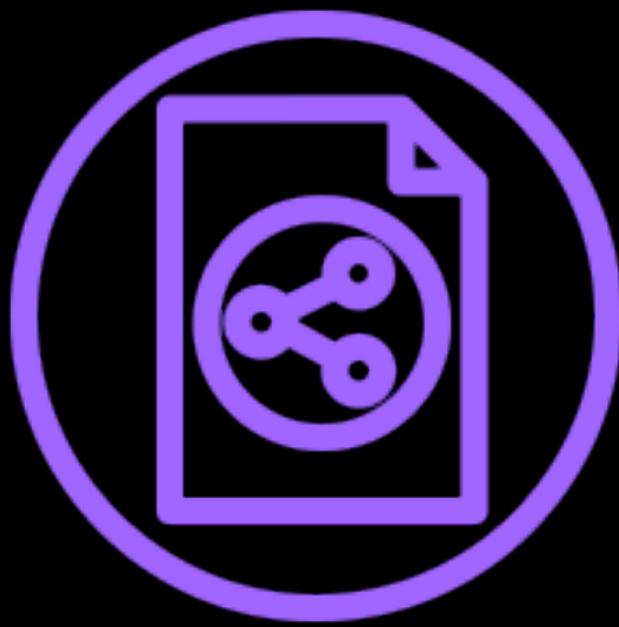
**Application Load Balancer**



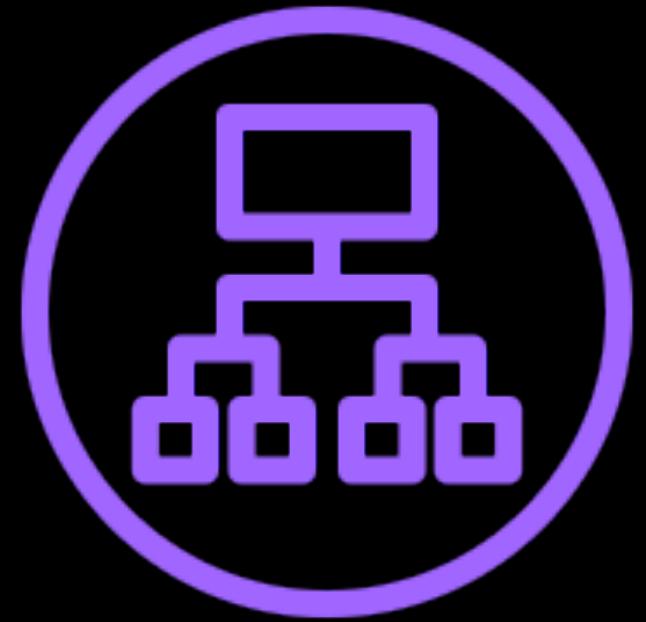
**Network Load Balancer**



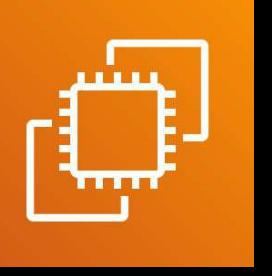
**Gateway Load Balancer**

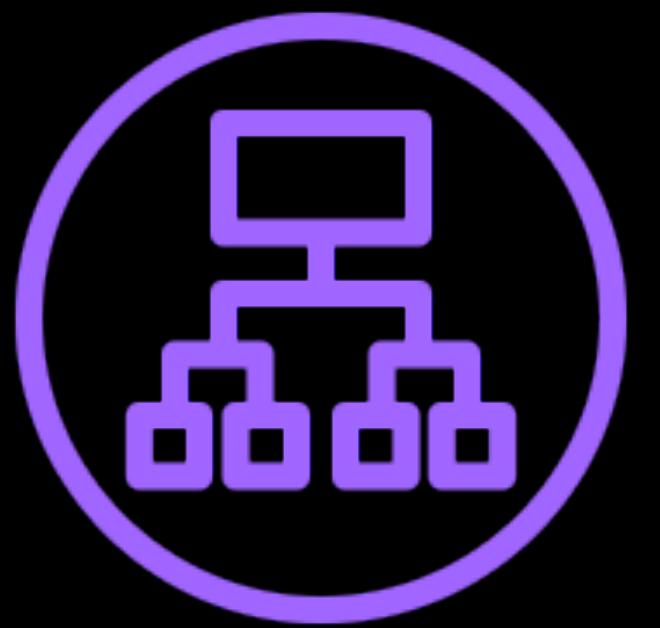


**Classic Load Balancer**



## Application Load Balancer

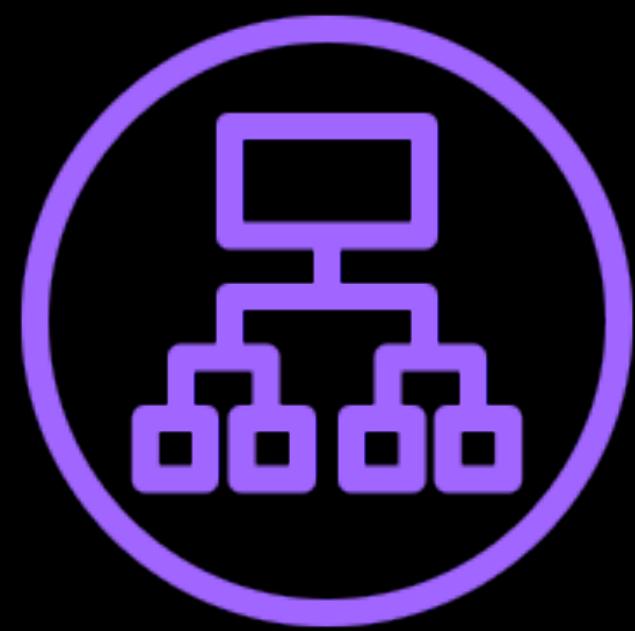
- Primarily used for load balancing **HTTP and HTTPS traffic**
- Suitable for web applications
- Works on the **Layer 7 (Application Layer) of the OSI Model**
- Supports Round Robin (default) and Least Outstanding Requests (LOR) routing algorithms
- Target types:
  -  Amazon EC2 Instance
  -  AWS Lambda Function
  -  IP Address
- Supported Protocol listeners: **HTTP, HTTPS, and gRPC**
- Also supports **WebSockets and HTTP2**
- Can be integrated with **AWS Global Accelerator, AWS Config, AWS WAF and other features**



## Application Load Balancer

- **Notable features:**
  - **Advanced routing via listener rule condition types**
  - **Connection Draining**
  - **Idle connection timeout**
  - **Cross-zone Load Balancing**
  - **Preserving Source IP address**
  - **Slow Start**
- **Has different security features such as:**
  - **SSL Offloading**
  - **Server Name Indication (SNI)**
  - **Back-end Server Encryption**
  - **User Authentication**
  - **Application-Layer Protocol Negotiation (ALPN)**
  - **Integration with Security Group and AWS WAF**

## LISTENER RULE CONDITION TYPES



**Application Load Balancer**

- **Host condition**

tutorialsdojo.com  
portal.tutorialsdojo.com  
app.tutorialsdojo.com  
.tutorialsdojo.com

- **HTTP Header**

User-Agent  
Content-Type

- **HTTP Request Method**

GET, POST, PUT, DELETE

- **Path**

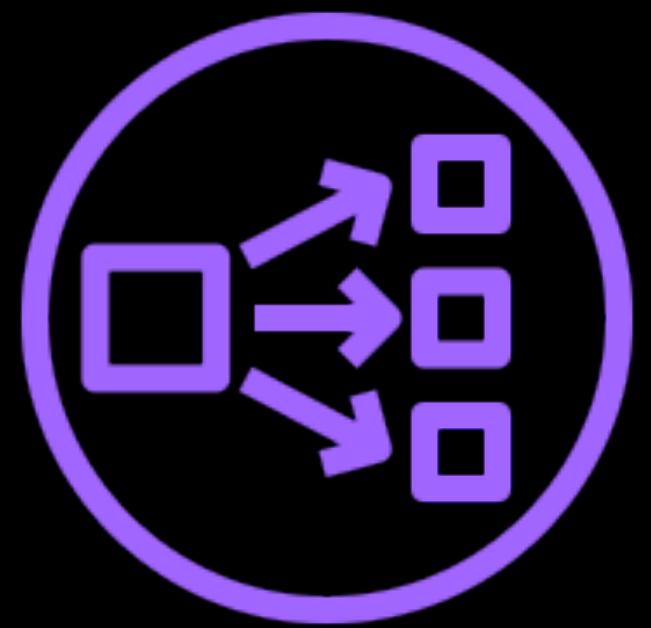
/img/  
/doc/cebu  
/pdf/\*/\*report

- **Query String**

/info?version=1  
/health?status=manila  
/account?id=123&alias=pogi

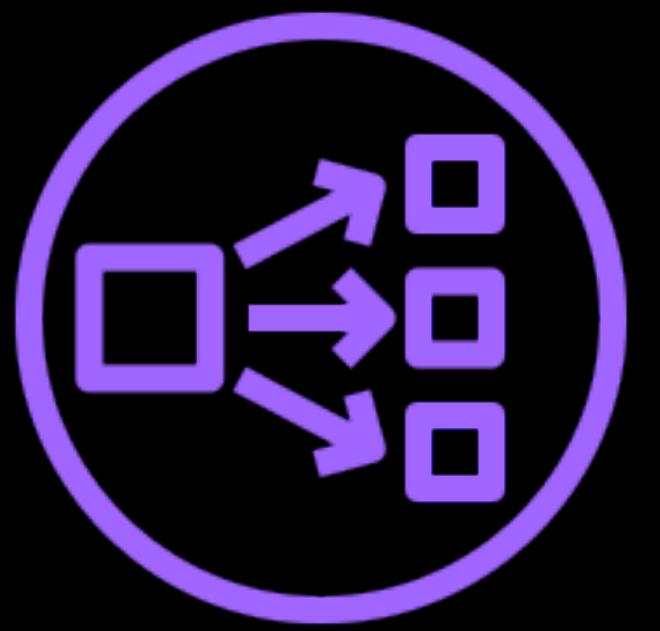
- **Source IP**

192.0.2.0, 198.51.100.10



## Network Load Balancer

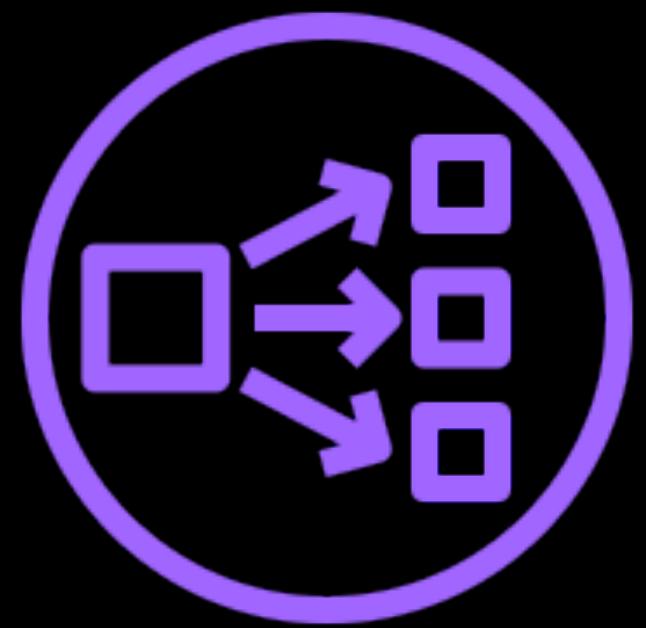
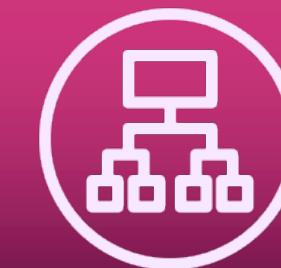
- For load balancing TCP, UDP, and TLS traffic
- Can handle millions of requests per second
- Routes the traffic while maintaining ultra-low latencies
- Works on the Layer 4 (Transport Layer) of the OSI Model
- Uses the flow hash routing algorithm
- Can be directly associated with an Elastic IP address
- Supports direct integration with: AWS Global Accelerator, AWS Config, VPC Endpoint Services and Traffic Mirroring



## Network Load Balancer

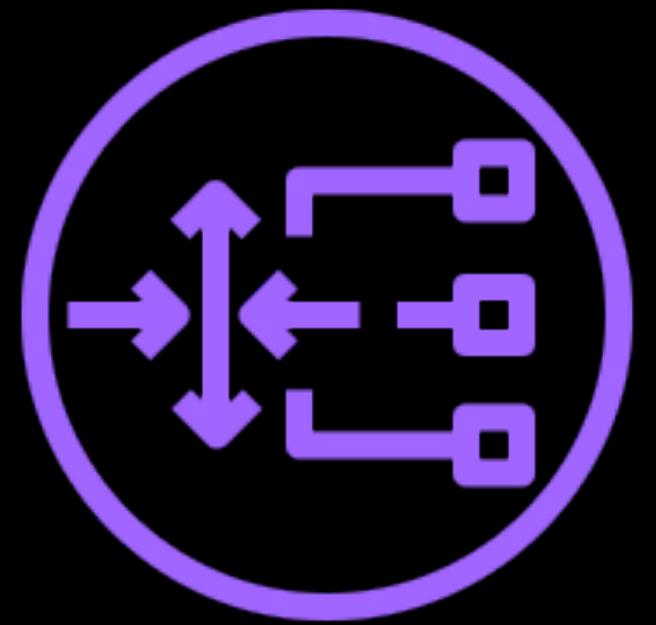
- **Notable features:**
  - **Connection Draining**
  - **Cross-zone Load Balancing**
  - **Preserving Source IP address**
  - **WebSockets support**
  - **Long-lived TCP connection**
- **Has different security features such as:**
  - **SSL Offloading**
  - **Server Name Indication (SNI)**
  - **Back-end Server Encryption**
  - **Application-Layer Protocol Negotiation (ALPN)**
  - **Integration with AWS Global Accelerator**

## Notable differences between ALB and NLB



### Network Load Balancer

- Does not have a selection of rule condition types unlike ALB
- Uses the TCP and UDP transport protocols not HTTP and HTTPS
- Suitable for various networking use cases, or for real-time multiplayer games that uses UDP
- Can support millions of requests per second while maintaining ultra-low latencies unlike ALB
- Can be directly integrated with an Elastic IP address, unlike ALB



## Gateway Load Balancer

- Primarily used for running **third-party virtual appliances**
- Suitable for custom firewalls, deep packet inspection systems, intrusion detection & prevention systems and many other virtual appliances
- Uses the Internet Protocol (IP) to pass the OSI Layer 3 traffic to its registered targets
- Works on both Layer 3 (Network Layer) and Layer 4 (Transport Layer) of the OSI Model
- Uses the **Generic Network Virtualization Encapsulation (GENEVE)** protocol to exchange application traffic
- You can use **GWLB endpoints** to exchange traffic across different VPC boundaries
- The access is **configured using the route tables of your VPC**, instead of virtual IP addresses



## Classic Load Balancer

- Intended for legacy applications that are still using the EC2-Classic network
- Not recommended for modern applications
- Supports both the transport layer protocols (TCP, SSL) as well as the application layer protocols (HTTP, HTTPS)
- Works on both Layer 4 (Transport Layer) and Layer 7 (Application Layer) of the OSI Model
- For applications with custom security policies and TCP passthrough configuration
- Can provide end-to-end security for your data-in-transit



# Amazon S3 Overview



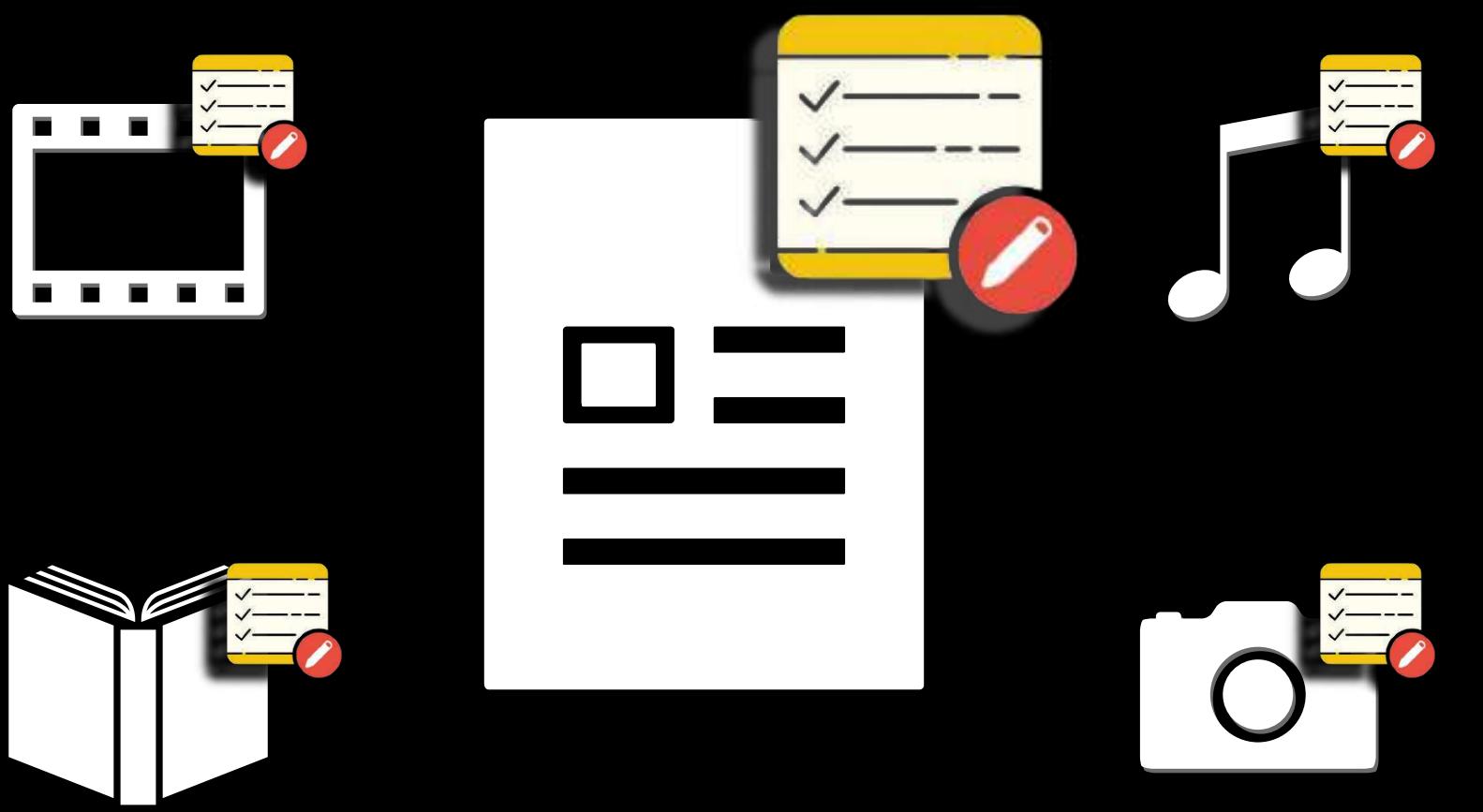
## Amazon S3

- An **object storage service**
- S3 stands for “**Simple Storage Service**”
- **Highly durable, available & scalable storage service**
- Primarily **used to store static data that does not change frequently**
- Allows your files to be publicly available via the Internet



a set of name-value pairs

Highly scalable and allows you to store  
virtually unlimited amounts of files





## BUCKET NAMING GUIDELINES

- The S3 bucket name is **globally unique**
- The namespace is **shared by all AWS accounts around the world**
- Example:
  - If you created an S3 bucket named “tutorialsdojo”, then no other AWS user can **create a bucket with that same name**
  - If someone tries to create a new bucket called “tutorialsdojo”, then that request will fail



## Amazon S3 Folders and Prefixes

- Helps you organize or group your objects
- S3 has a flat structure
- The concept of a “folder” is not hierarchical unlike Amazon EFS
- Example:

Object key name  
**tutorialsdojo/aws.jpeg**

Prefix

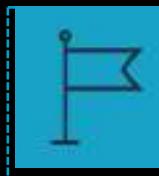
Filename



- Amazon S3 does **NOT** support POSIX, including:
  - Concurrent file modification
  - File system access semantics
  - File locking

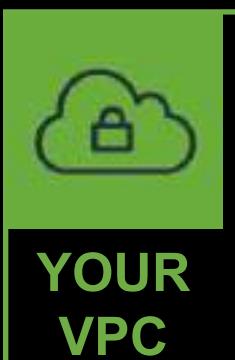


## AWS Cloud



N. Virginia Region

Automatically replicates your objects to all Availability Zones of the AWS region by default



YOUR  
VPC

Availability Zone (AZ) 2



Availability Zone (AZ) 3



AVAILABILITY

99.99%

DURABILITY

99.99999999%

- The probability that an object remains intact and accessible after a period of one year



DURABILITY

99.99999999%

100%

Absolutely no data loss per year

99%

1% chance of data loss per year

99.99%

0.01% chance of data loss per year

0.00000001% chance  
of data loss per year or one lost data  
every 10 million years



## Amazon S3 Storage Classes



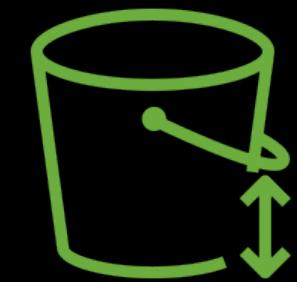
For **frequently accessed data**

**S3 Standard**



For **changing or unknown access patterns**

**S3 Intelligent-Tiering**



**S3 Standard-IA**  
(Infrequent Access)



**S3 One Zone-IA**  
(Infrequent Access)

For **storing long-lived, yet less frequently accessed data**



**S3 Glacier**

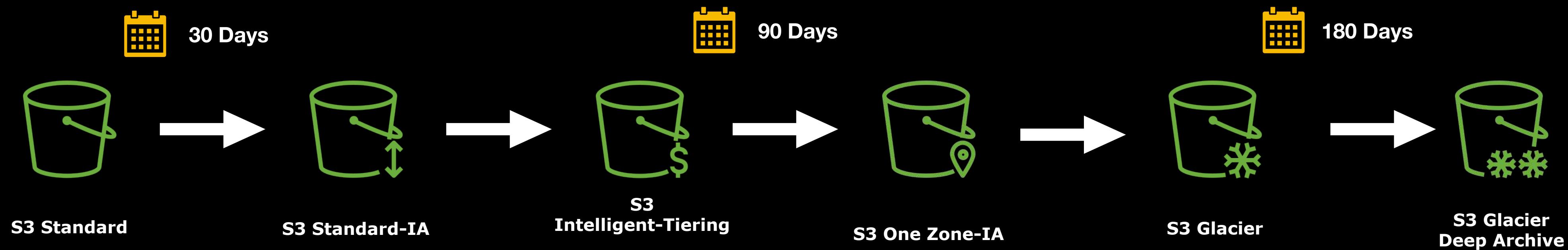


**S3 Glacier Deep Archive**

For **low-cost long-term storage and data archiving**



## Lifecycle Policy



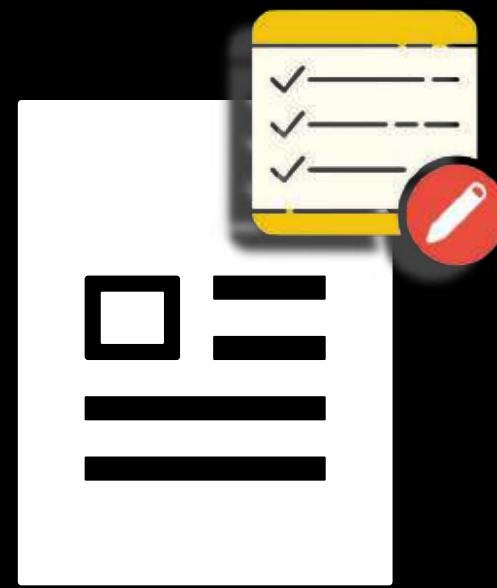


## Static Website Hosting

- Launch a **static website** with **HTML pages, downloadable packages, images, media files, or other client-side scripts**
- **Cost-effective solution** for hosting your **static websites with no server management required (serverless)**
- **Cannot be used for running server-side scripts** such as **PHP, JSP, ASP.NET etc...**



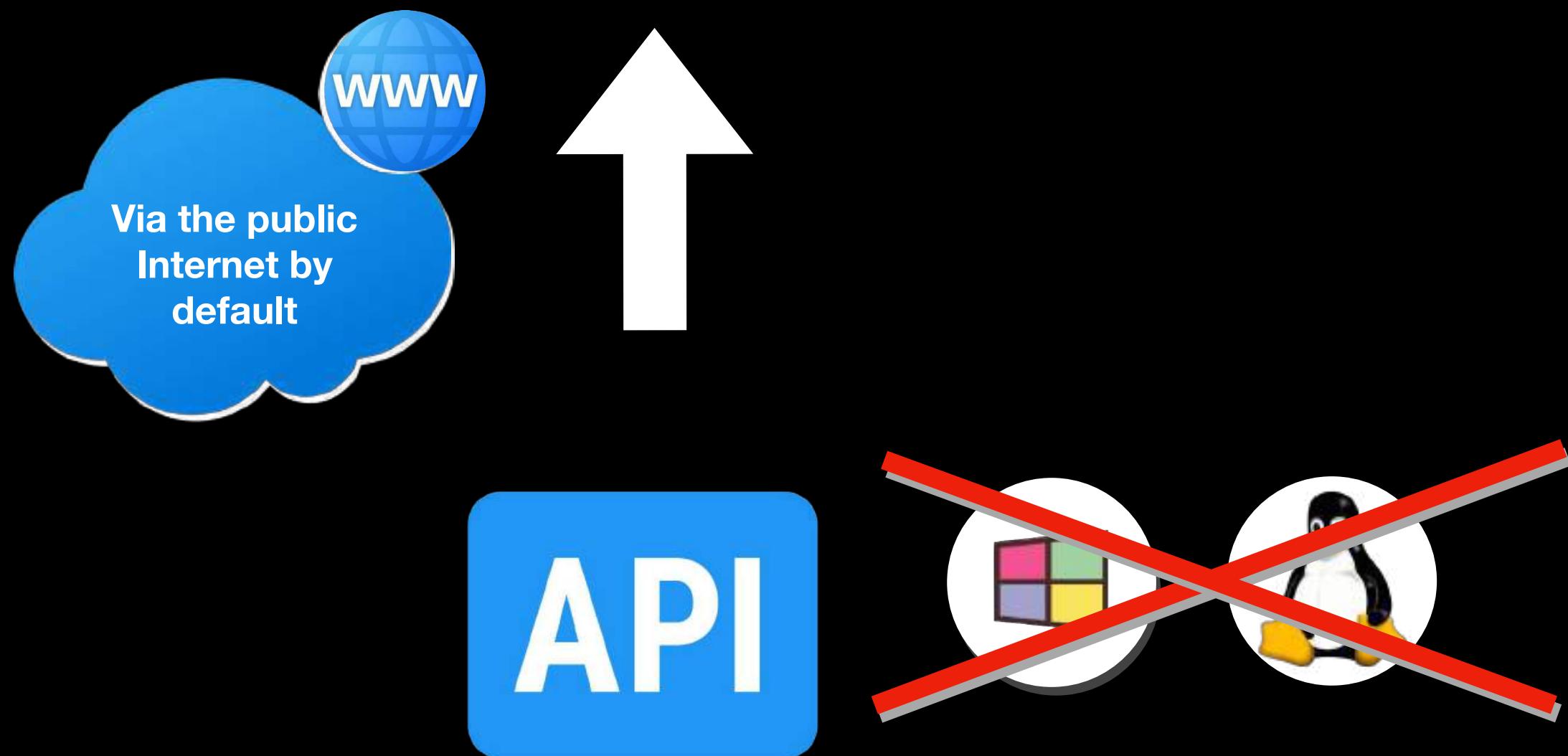
Amazon S3



Amazon EBS



Amazon EFS



- Invoked via a **REST API** request call

- **Attached/Mounted** to the Amazon EC2 instance



S3 Versioning



Multi-Factor Authentication  
(MFA)

- Prevent accidental data deletion in Amazon S3



Access Control List  
(ACL)

- Secure access to your S3 buckets and objects



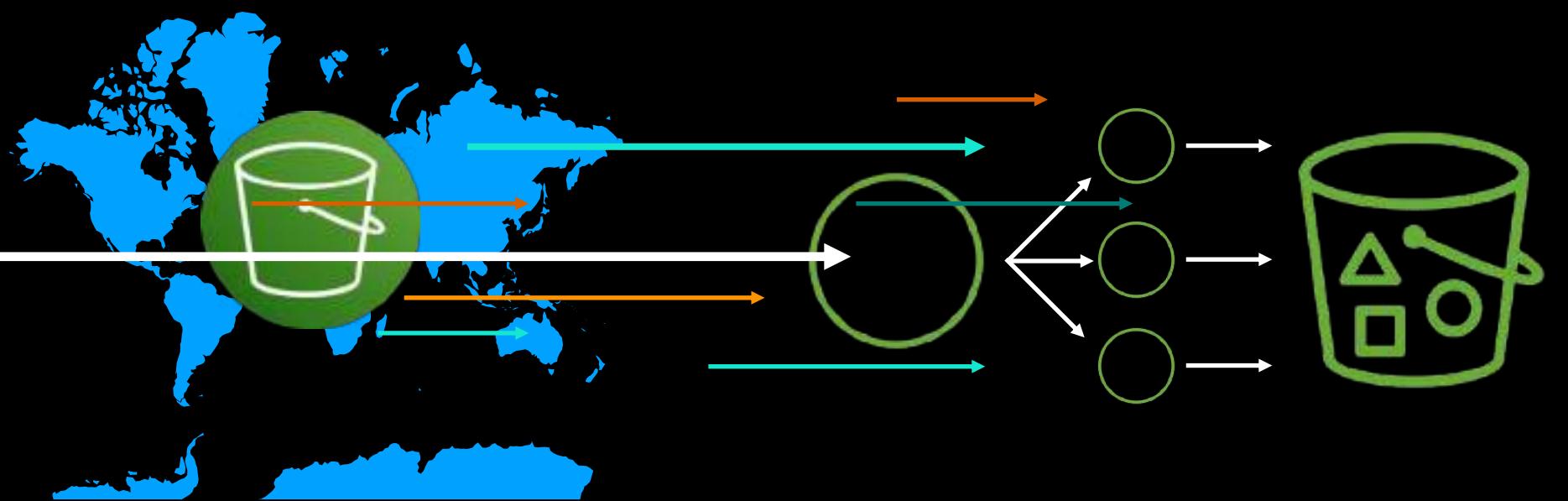
Bucket Policy

- Control external access to your Amazon S3 bucket



Cross Region Replication (CRR)

- Automatically replicate objects to a different AWS Region for backup purposes



Transfer Acceleration

Multipart Upload

- Accelerate or expedite the data transfer (upload/download) of S3 objects

...and many other S3 features!



# Amazon S3 Storage Classes



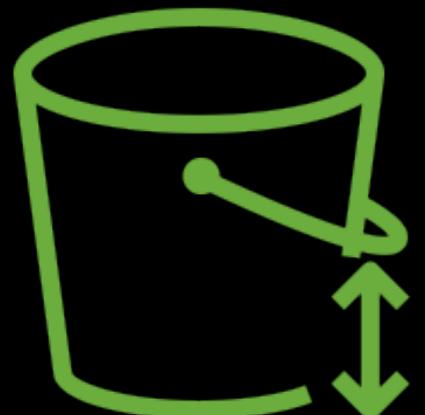
## Amazon S3 Storage Classes



**S3 Standard**



**S3 Intelligent-Tiering**



**S3 Standard-Infrequent Access** (Standard-IA)



**S3 One Zone-Infrequent Access** (One Zone-IA)



**S3 Glacier**



**S3 Glacier Deep Archive**



## S3 Standard

- Primarily used for storing your data that are frequently accessed
- Highly durable, highly available, and high performance object storage
- Replicates your data to 3 or more Availability Zones
- 99.99% Availability
- No minimum storage duration charge
- No data retrieval fee



## S3 Standard

USE CASES

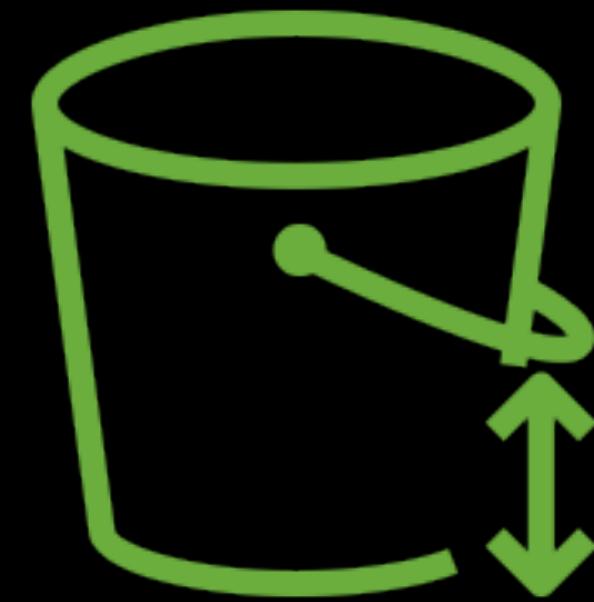
- For setting up a highly available and durable **static web hosting**
- As a **temporary storage service** for storing the nightly log processing of your application, where the logs are meant to be stored for 1 day (24 hours) only. It is a cost-effective option for this case since it has **no minimum storage duration charge**



## S3 Standard

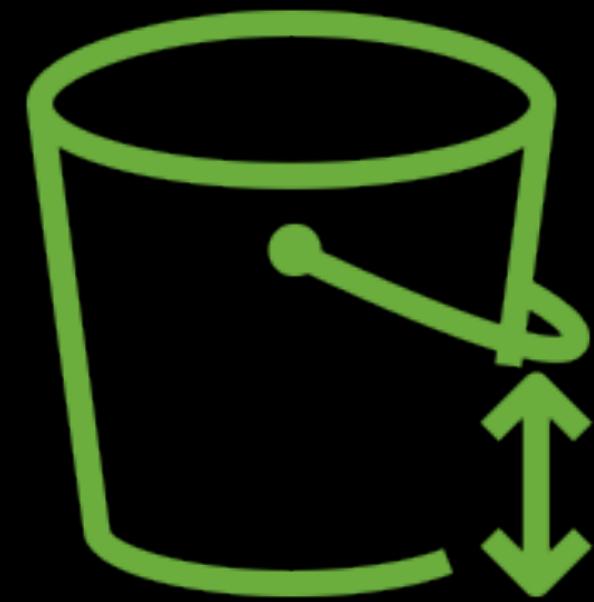
LIMITATIONS

- Not cost-effective as this storage class is the most expensive among all other classes
- Not recommended for data archiving, for infrequently access files or for any workloads that require a cost-effective storage



## S3 Standard-IA

- **Primarily used for storing infrequently accessed data but provides a way to rapidly retrieve the stored files**
- **Replicates your data to 3 or more Availability Zones**
- **99.99% Availability**
- **30-day minimum storage duration charge**
- **Has a data retrieval fee that is measured per gigabyte (GB)**



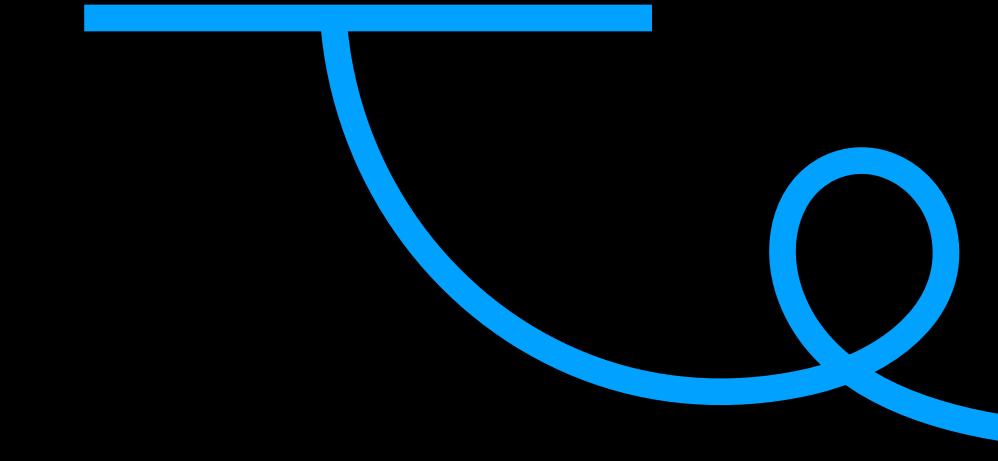
## S3 Standard-IA

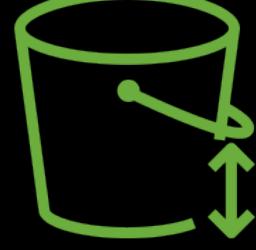
USE CASES

- As a long-term storage for long-lived, but infrequently accessed data
- For data backups
- As a data store for your Disaster Recovery (DR) files
- For storing the primary backup copies of your on-premises dataset



## S3 One Zone-IA



- For storing **less frequently accessed and easily reproducible data** that requires **immediate retrieval** when needed
- **30-day minimum storage duration charge**
- **Cheaper than:**  S3 Standard-IA
- **Only uses 1 Availability Zone**
- **99.95% Availability (the *lowest* among all other Amazon S3 storage classes)**



## S3 One Zone-IA

USE CASES

- If you require a **cost-effective option to store infrequently accessed data**
- For **workloads that do not require the availability and resilience of the Amazon S3 Standard or S3 Infrequent Access class**
- For **storing secondary backup copies of rarely-accessed on-premises dataset**
- For **storing easily recreatable data**



## S3 One Zone-IA

LIMITATIONS

- The data is replicated in a **single AZ** only
- Not recommended for **storing your company's primary backup copies or any critical business data that is difficult to reproduce**



## S3 Intelligent-Tiering

- Delivers **automatic cost savings**
- **Automatically moves your objects between different access tiers whenever your access pattern changes**
- **30-day minimum storage duration charge**
- **No data retrieval fee**
- Moves data to the most cost-effective access tier **without any operational overhead**
- Stores the objects in four access tiers:
  - 2 **low-latency** access tiers
  - 2 optional **archive** access tiers



## S3 Intelligent-Tiering

USE CASES

- Suitable if your data has an **unpredictable access pattern**
- For buckets with a mix of frequent and infrequent accessed data
- If the access patterns to **your data vary all the time**
- If some of your files are accessed frequently while the **others are rarely accessed (move to Glacier)**
- If some of your data are accessed less frequently than others (**move to IA tier**)
- If you are unsure of how frequently your data will be accessed



## S3 Intelligent-Tiering

USE CASES

- If you want to **keep costs low by automatically moving your data to the appropriate S3 storage class**
- If your data will be accessed by **users over variable periods of time**
- If you need storage with **no management overhead**
- If you want to **avoid lifecycle policies that are not consistently implemented or are partially implemented**



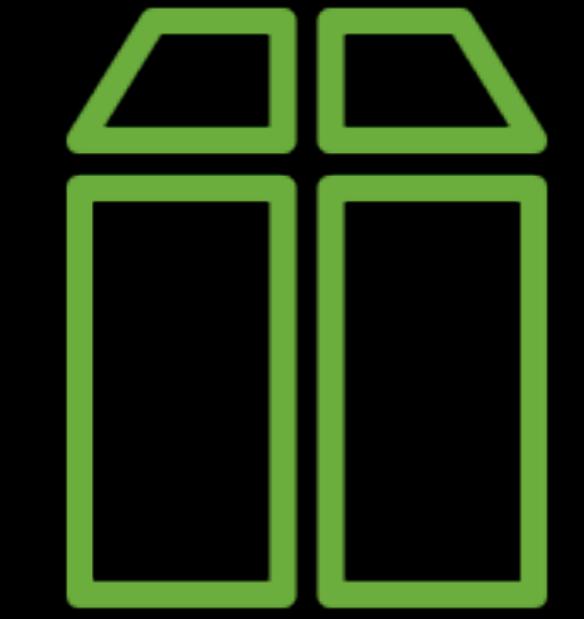
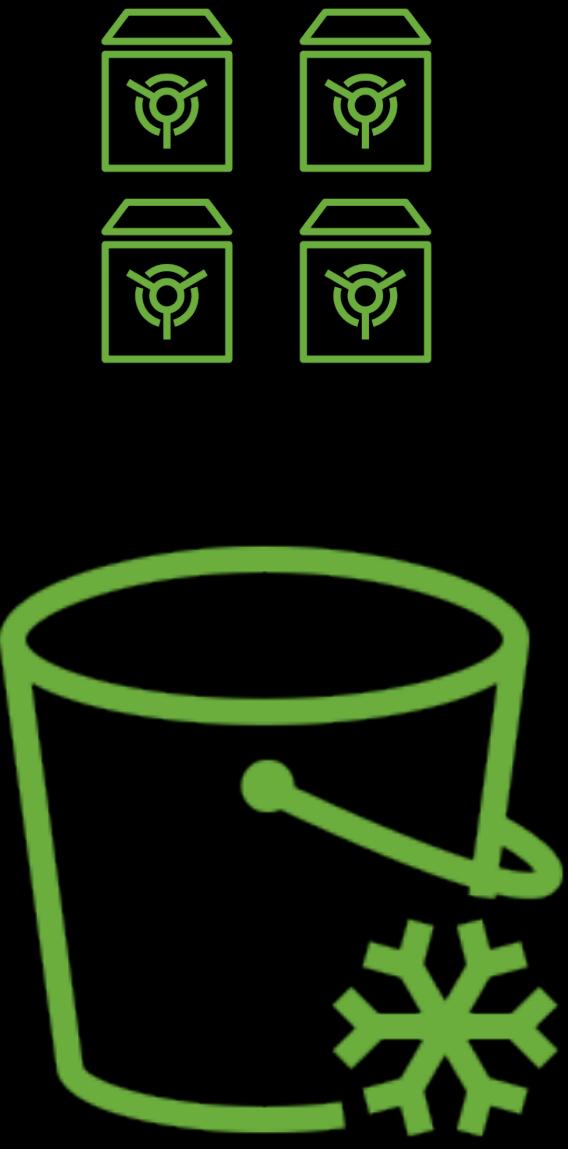
## S3 Glacier

- A secure, durable, and low-cost storage
- Suitable for **data archiving**
- A cost-effective storage solution for rarely accessed data and does not require a fast retrieval time
- Replicates your data to 3 or more Availability Zones
- 99.99% Availability
- 90 day-minimum storage duration charge
- High data retrieval fee (expensive)



## S3 Glacier

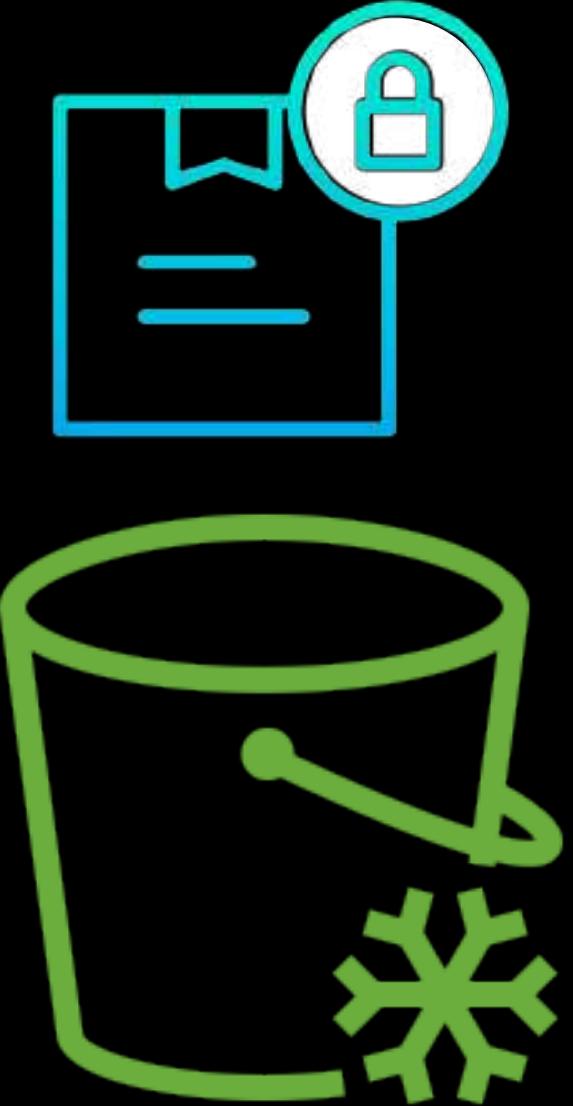
- **Has its own management console apart from the regular Amazon S3 console**
- **2 Ways to store your data:**
  - **Using the Amazon S3 console**
  - **Using the Amazon Glacier console**
- **Automatically move your data from S3 Standard or S3 Standard-IA to Amazon S3 Glacier by using a lifecycle policy**



S3 Glacier  
**Vault**

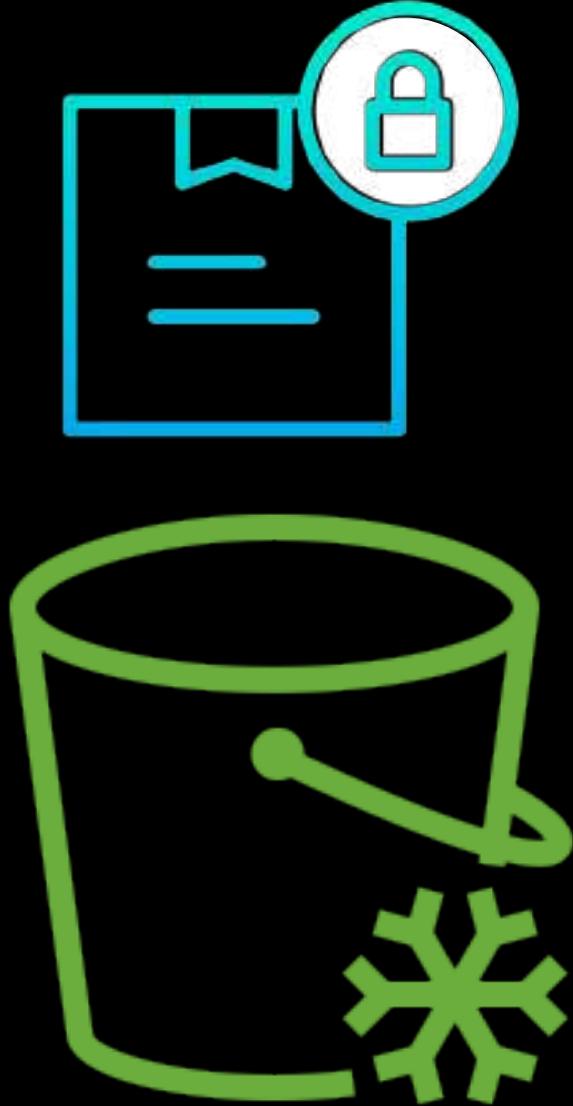
A dark grey rectangular button with the word "Vault" in white, surrounded by small diamond shapes.

- Has a resource called: **Vault**
- A vault is a **container for storing your data archives**
- Base unit of storage in S3 Glacier, containing a **unique ID and an optional description**
- Can only be created in the **Amazon S3 Glacier console**
- You must provide the **vault name and its corresponding AWS Region**

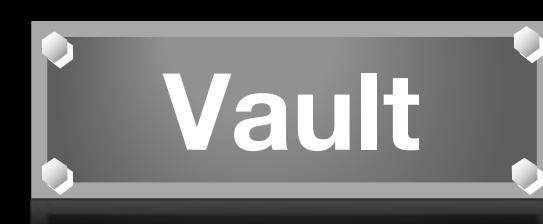


S3 Glacier  
**Vault**

- Use a **Vault Lock** to ensure data integrity and access control to your Amazon S3 Glacier Vaults
- A **Vault Lock** is an access policy that helps you **enforce regulatory and compliance requirements**
- You can specify a “**Write Once Read Many**” (**WORM**) control to lock your Glacier vault policy from future edits
- A **Glacier vault access policy** can no longer be changed when the vault lock process has been completed after 24 hours



## S3 Glacier



USE CASES

- Applicable if your company wants to retain its archives for a specific number of years before the files can be deleted
- If you want to deny users from modifying or deleting an archive until after 1 year, 3 years, 7 years et cetera



## S3 Glacier Archival Retrieval Options

### EXPEDITED

- Quickly access a subset of your data archives
- Allows you to access your archived data within 1 - 5 minutes ( file size should NOT exceed 250 MB )
- Ensure sufficient retrieval capacity for your *Expedited* retrieval operations by purchasing provisioned capacity

### STANDARD

- Default option for retrieval requests
- Allows you to access any of your glacier archives within 3 – 5 hours

### BULK

- Lowest-cost retrieval option
- Retrieves large amounts of data archive in less than half a day
- Typically completes the process within 5 – 12 hours



## S3 Glacier Deep Archive

- The **lowest-cost storage class** in Amazon S3.
- Supports **long-term retention and digital preservation** for your data
- Primarily used to retain your data sets for **7 to 10 years or longer** to meet **regulatory compliance requirements**
- Replicates your data to **3 or more Availability Zones**
- **99.99% Availability**



## S3 Glacier Deep Archive

- **180-day minimum storage duration charge ( *roughly 6 months* )**
- **Should be used for data archiving only**
- **The data stored here should be rarely accessed with no strict retrieval time**



## S3 Glacier Deep Archive - Retrieval Options

STANDARD

- Default option for retrieval requests
- Data will be restored within 12 hours

BULK

- Costs lower than the Standard retrieval option
- Data will be restored within 48 hours



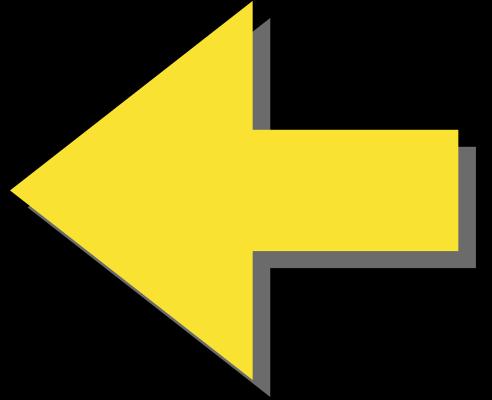
# Amazon S3 **Minimum Storage Duration**

---

Which is more **cost-effective?**



**S3 Glacier**



**S3 Standard**



## Minimum Storage Duration

- The specific amount of time that your objects must be stored in a particular storage class
- Deleting your objects won't affect the minimum storage duration. You will still have to pay the remaining days of the mandatory minimum period
- A minimum storage duration of 30 days means that you will be charged for the entire 30 days even if you deleted or changed the storage class of your objects before that period



## Minimum Storage Duration



- An object was uploaded in an Amazon S3 Standard Infrequent Access (S3 Standard-IA) storage class
- You deleted the object after 10 days
- You're still billed for the entire 30 days
- Also applicable if you changed the storage class to another class



	S3 Standard	S3 Intelligent-Tiering	S3 Standard-IA	S3 One Zone-IA	S3 Glacier	S3 Glacier Deep Archive
Designed for durability	99.999999999% (11 9's)					
Designed for availability	99.99%	99.9%	99.9%	99.5%	99.99%	99.99%
Availability SLA	99.9%	99%	99%	99%	99.9%	99.9%
Availability Zones	≥3	≥3	≥3	1	≥3	≥3
Minimum capacity charge per object	N/A	N/A	128KB	128KB	40KB	40KB
Minimum storage duration charge	N/A	30 days	30 days	30 days	90 days	180 days
Retrieval fee	N/A	N/A	per GB retrieved	per GB retrieved	per GB retrieved	per GB retrieved
First byte latency	milliseconds	milliseconds	milliseconds	milliseconds	select minutes or hours	select hours
Storage type	Object	Object	Object	Object	Object	Object
Lifecycle transitions	Yes	Yes	Yes	Yes	Yes	Yes

# Which is more cost-effective?

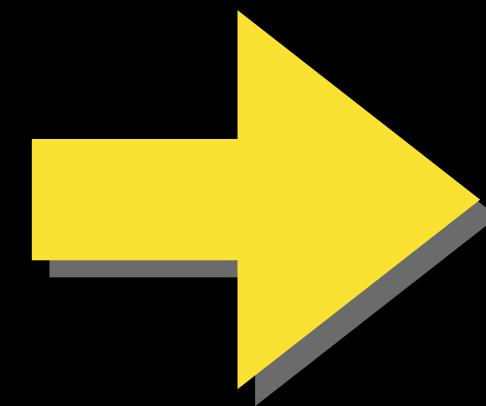
Non-reproducible and frequently-accessed data that needs to be **temporarily** stored for



hours only



S3 Glacier



S3 Standard



**90-Day Minimum Storage Duration**



**NO Minimum Storage Duration**

$$180 - 10 = 170 \text{ Days!}$$



S3 Glacier

VS



S3 Glacier  
Deep Archive

LOW    \$ \$



90 Days

You will be billed for the entire 90 Days

Normal storage usage charge

Normal storage usage charge

COST

MINIMUM STORAGE DURATION

DATA DELETED AFTER  
10 DAYS

DATA DELETED AFTER  
90 DAYS

DATA DELETED AFTER  
180 DAYS

LOWEST    \$



180 days

You will be billed for the entire 180 Days

You will be billed for the entire 180 Days

Normal storage usage charge



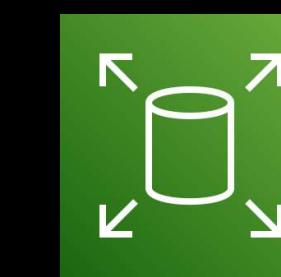
# Amazon S3 Event Notification



AWS CloudTrail  
Logs



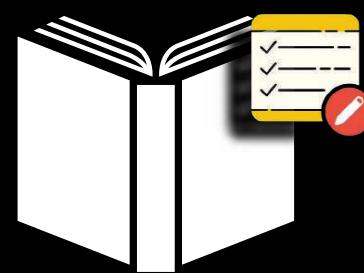
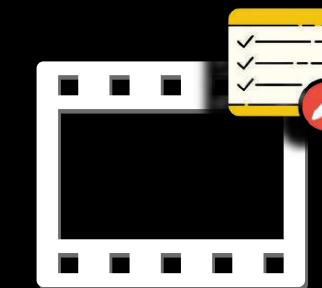
AWS CloudFormation  
Templates



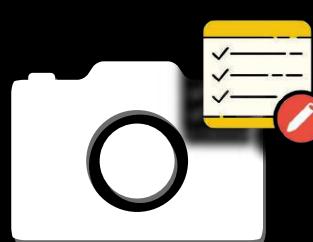
Amazon EBS  
Snapshots



ELB Access  
Logs



DATA LAKE



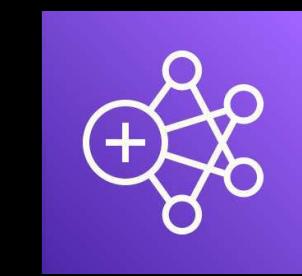
Amazon Redshift  
Spectrum



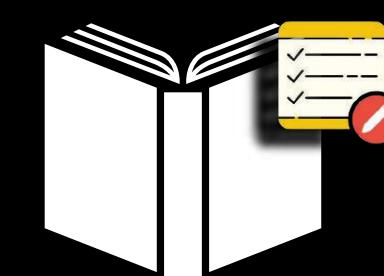
AWS Glue



Amazon Athena



Amazon EMR





**S3 Event Notifications**



## S3 Event Notifications

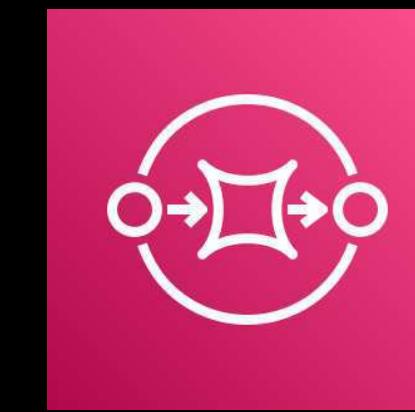
- **New Object Creation**
- **Object Deletion**
- **Object Restoration from the Amazon S3 Glacier storage class**
- **Reduced Redundancy Storage (RRS) object lost events**
- **Replication events**



## S3 Event Notifications



Amazon SNS



Amazon SQS



AWS Lambda

- Transmitted **within seconds**
- Delivered **at least once**
- Enable **object versioning** to ensure that an event notification is always sent whenever you upload an object



# Amazon RDS Overview

- A **relational database service**
- **Managed by both you (limited access) and AWS**
- **Allows you to run various database engines:**



## Amazon RDS



ORACLE®



Amazon  
Aurora

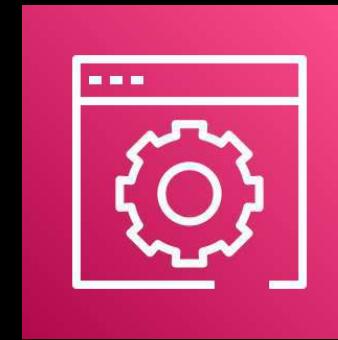
# Amazon RDS



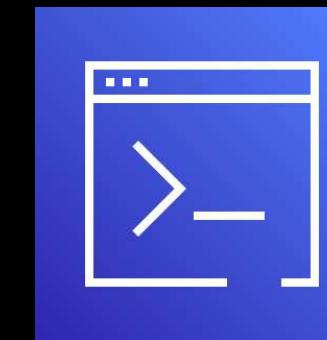
- Can be deployed using:



AWS  
CloudFormation



AWS Management  
Console



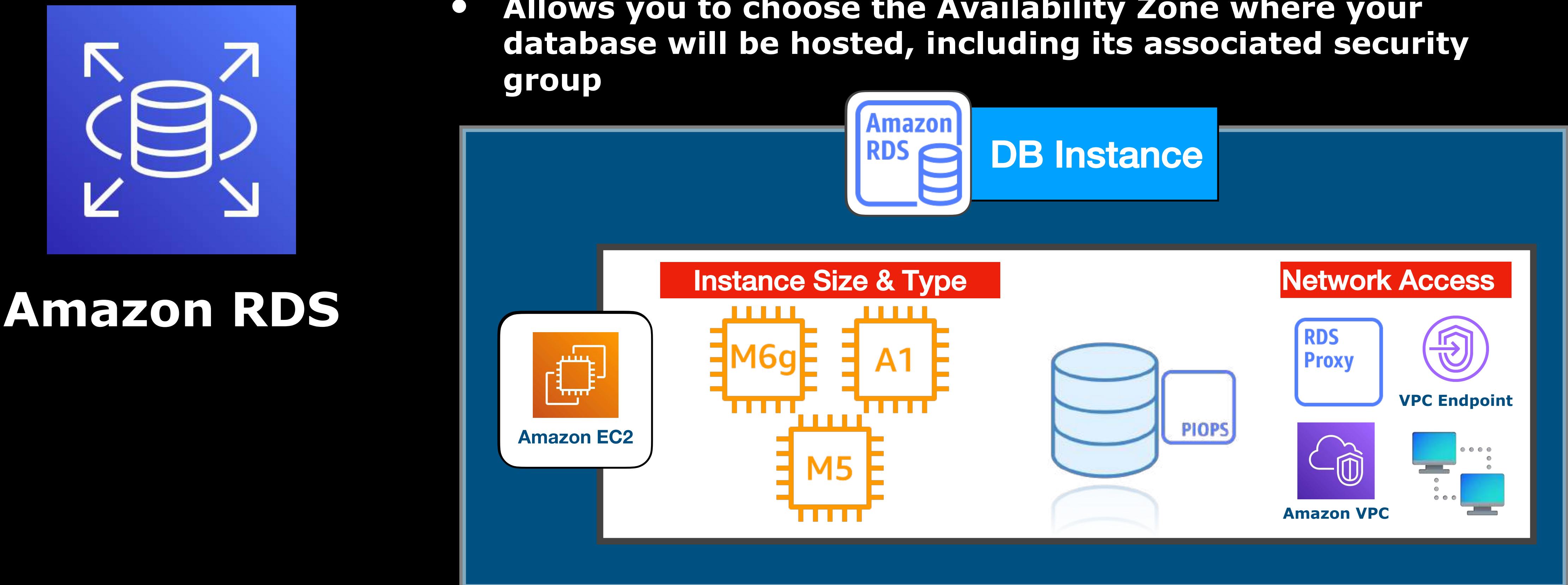
AWS CLI

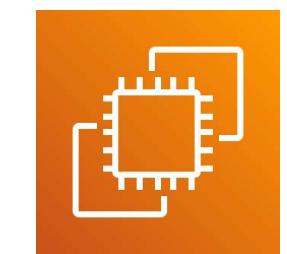


Amazon RDS  
API

- **Eliminates the time-consuming tasks of hardware provisioning, patching, backups, and maintenance for your database**

- You can **configure the underlying EC2 instance** used by your Amazon RDS database such as its size, instance type & storage
- Purchase a **Reserved DB instance** to lower down your RDS costs
- Allows you to choose the Availability Zone where your database will be hosted, including its associated security group





Amazon EC2

**Self-Hosted Database**



**Amazon RDS Database**



MANAGED BY



Amazon EC2

Self-Hosted Database



Amazon RDS Database

YOU  
*(AWS Customer)*

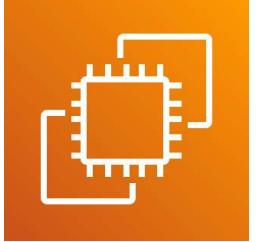


- Patching
- Scaling
- Taking database backups
- Ensuring high availability
- Replication
- Monitoring

- Minimal maintenance work

- Physical Infrastructure
- Virtualization layer
- Host OS of the EC2 instance

- Patching
- Scaling
- Taking database backups
- Ensuring high availability
- Replication
- Monitoring



Amazon EC2

## Self-Hosted Database

- Can be directly accessed via SSH, RDP or other connections
- Allows direct access and modification of your database configuration files such as:

/etc/mysql/my.cnf

- ConfigurationFile.ini
- INIT.ORA, TNSNAMES.O



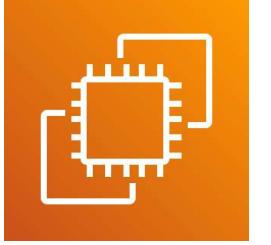
## Amazon RDS Database

- The underlying EC2 instance CANNOT be directly accessed via SSH or RDP

```
# Default MySQL Server config for Tutorials Dojo Tarlac
[mysqld]
# Only allow connections from localhost
bind-address = 127.0.0.1
mysqlx-bind-address = 127.0.0.1

read_only = 1
```

Read-Only setting



Amazon EC2

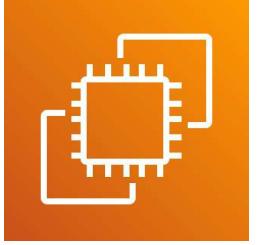
## Self-Hosted Database

- You have **full access** to the virtual machine and the underlying database
- You are responsible for making your database highly available, fault-tolerant and secure
- You have to apply the OS patches as well as the Database Engine patches regularly
- You will handle all of the database administrative tasks



## Amazon RDS Database

- Modify the database configuration via:
  - Parameter Group
  - Options Group
- You can choose the actual time when Amazon RDS will apply the DB patches in its **maintenance window**
- Database maintenance tasks are handled automatically



Amazon EC2

**Self-Hosted Database**



**Amazon RDS Database**



Microsoft  
**SQL Server**

```
# Default MySQL Server config for Tutorials Dojo Tarlac  
[mysqld]  
  
# Only allow connections from localhost  
bind-address = 127.0.0.1  
mysqlx-bind-address = 127.0.0.1  
  
read_only = 1
```

read\_only

aws Services ▾ Q #KayangKayaNgPinoy X 🔍 Tutorials Dojo N. Virginia Support ▾

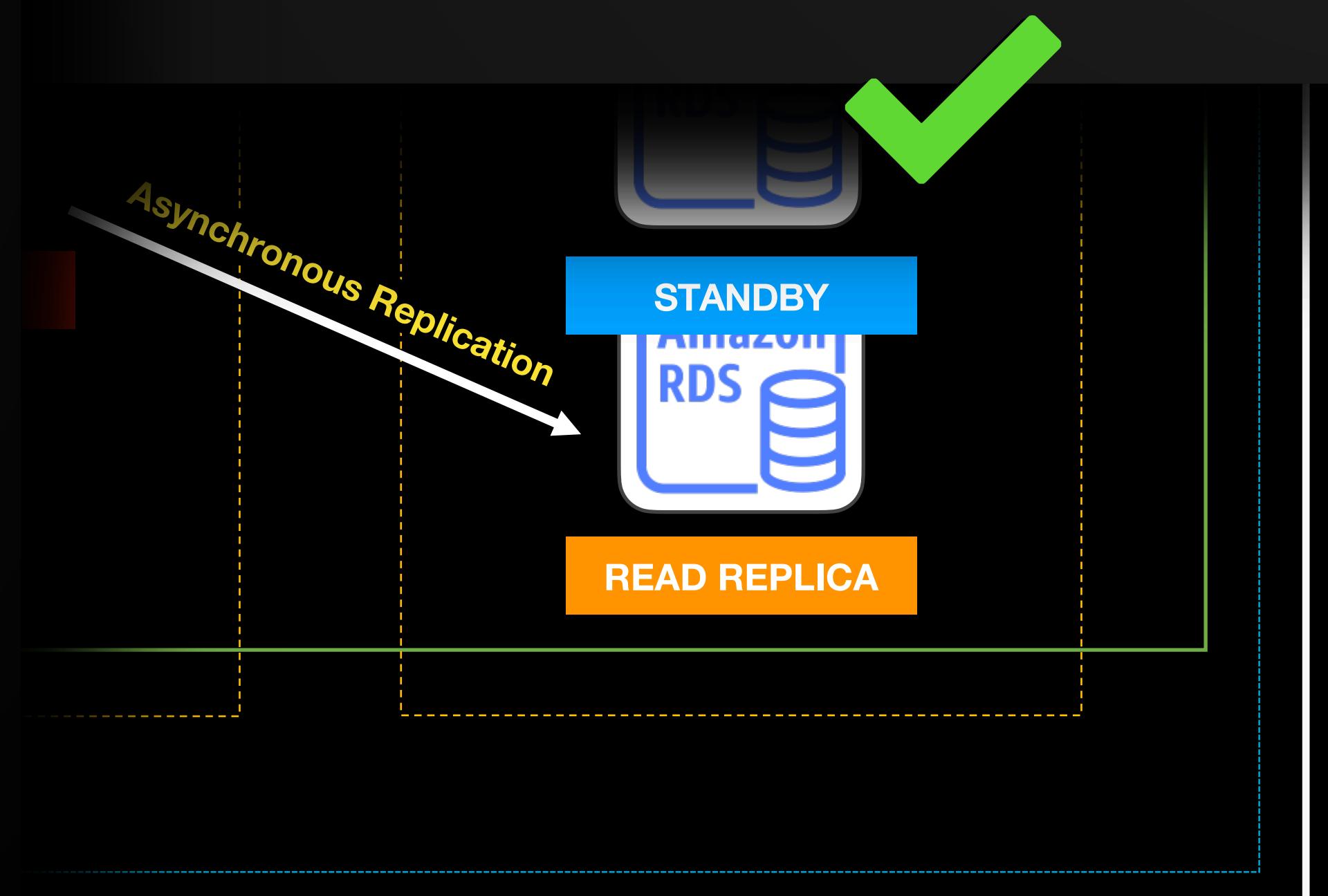
RDS > Parameter groups > Tutorials-Dojo-Manila

## Tutorials-Dojo-Manila

Parameters

Name	Values	Allowed values	Description
innodb_read_only		0, 1	Starts the server in read-only mode.
read_only	1	0, 1, {TrueIfReplica}	When it is enabled, the server permits no updates except from updates performed by slave threads.
super_read_only		0, 1	Whether client connections to the server are required to use some form of secure transport.

Cancel editing Preview changes Reset Save changes

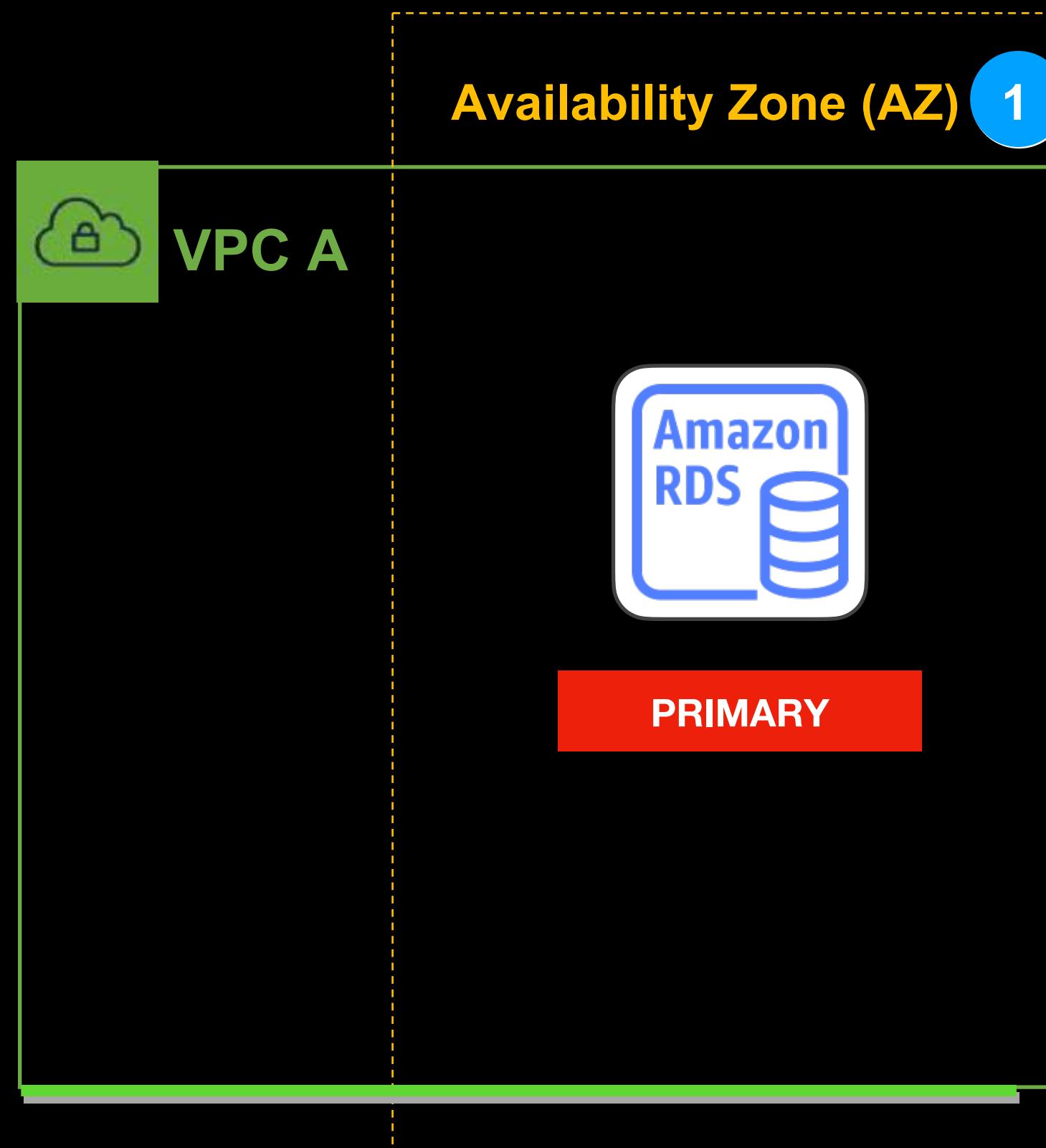




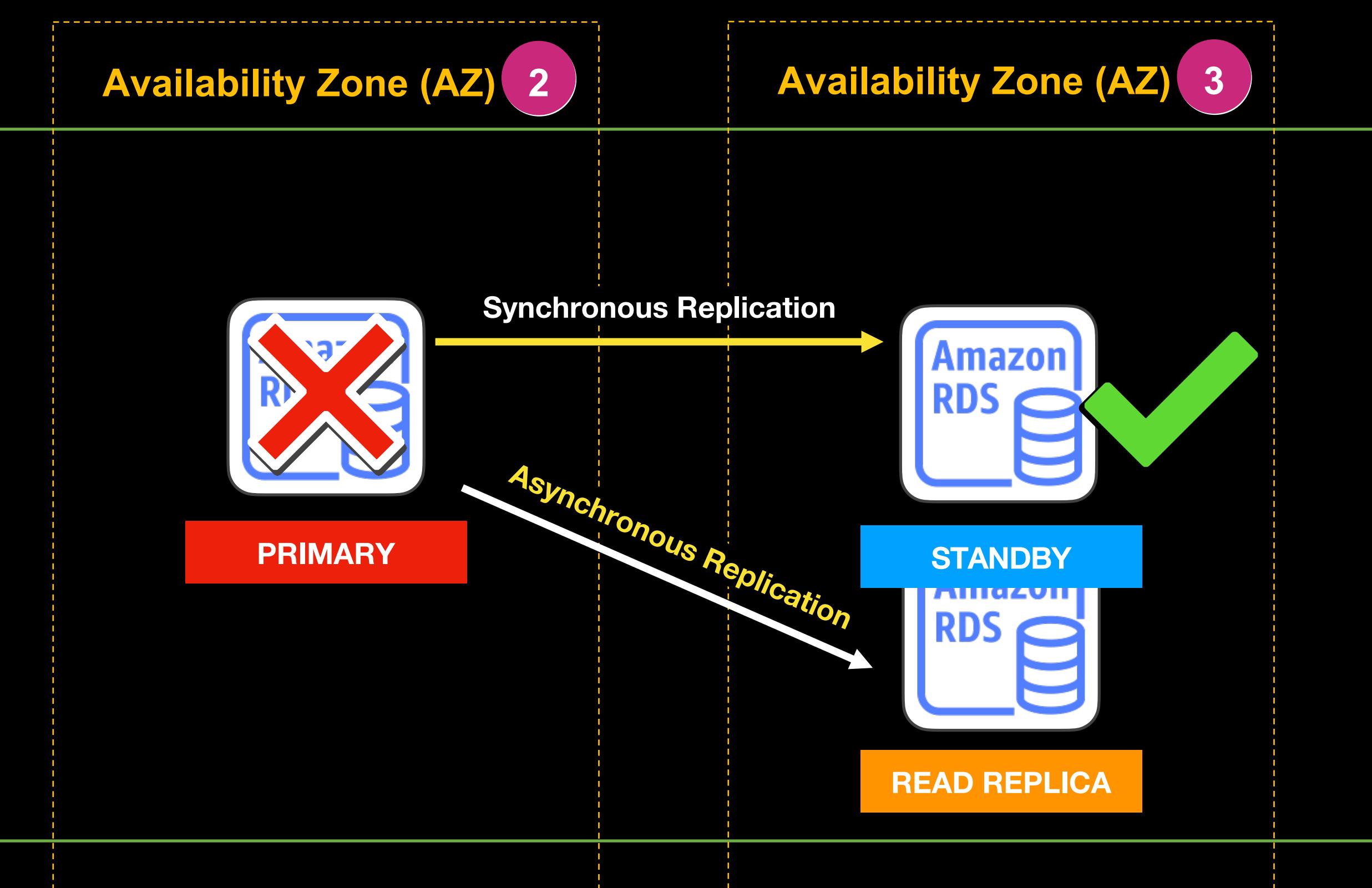
AWS  
Cloud

N. Virginia Region

## Single AZ

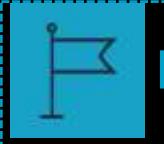


## Multi-AZ





## AWS Cloud



N. Virginia Region

### Single AZ

Availability Zone (AZ) 1



VPC A



PRIMARY

Availability Zone (AZ) 2



PRIMARY

### Multi-AZ

Availability Zone (AZ) 3



STANDBY

Asynchronous Replication



READ REPLICA



VPC B



READ REPLICA



Ohio Region



## Amazon RDS



## OLTP Applications

- Suitable for applications that read or write **constantly changing data**, such as **Online Transaction Processing** OLTP applications



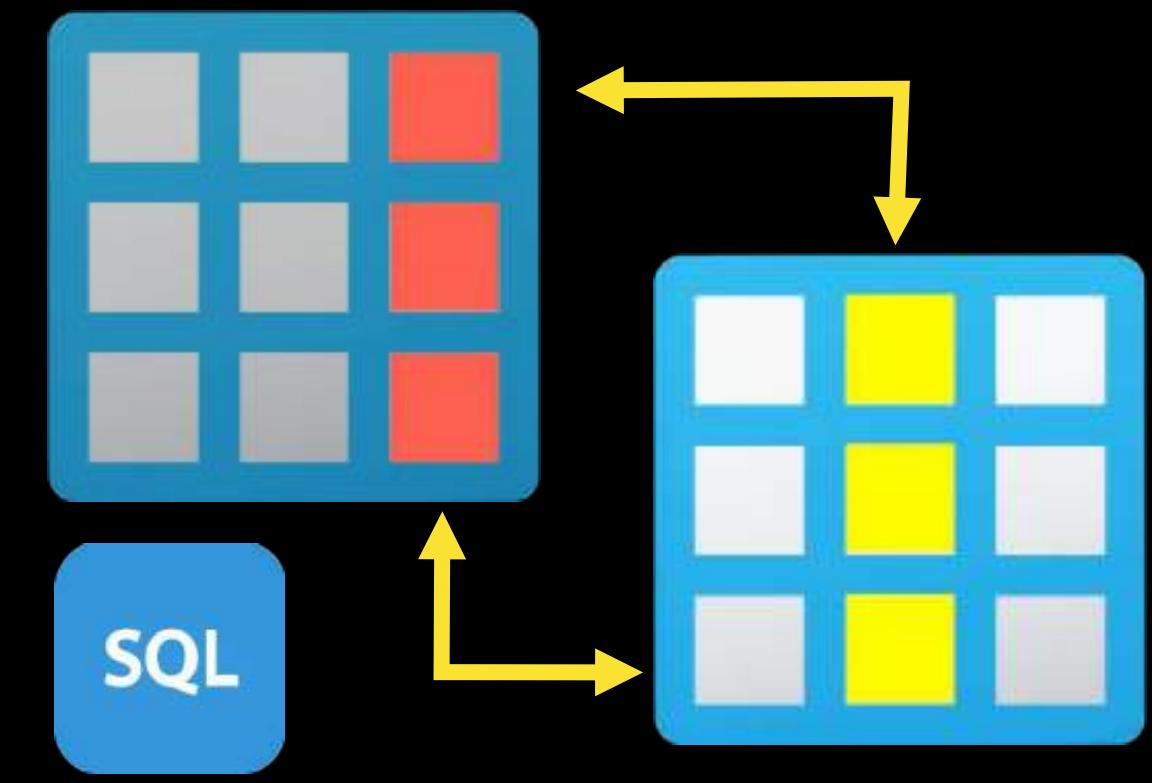
Amazon RDS

A TOMIC

C ONSISTENT

I SOLATED

D URABLE





- A fully managed, highly available database proxy
- Automatically connects your application to a new DB instance while preserving its application connections
- Minimizes downtime by instantly routing the incoming requests directly to the new database instance

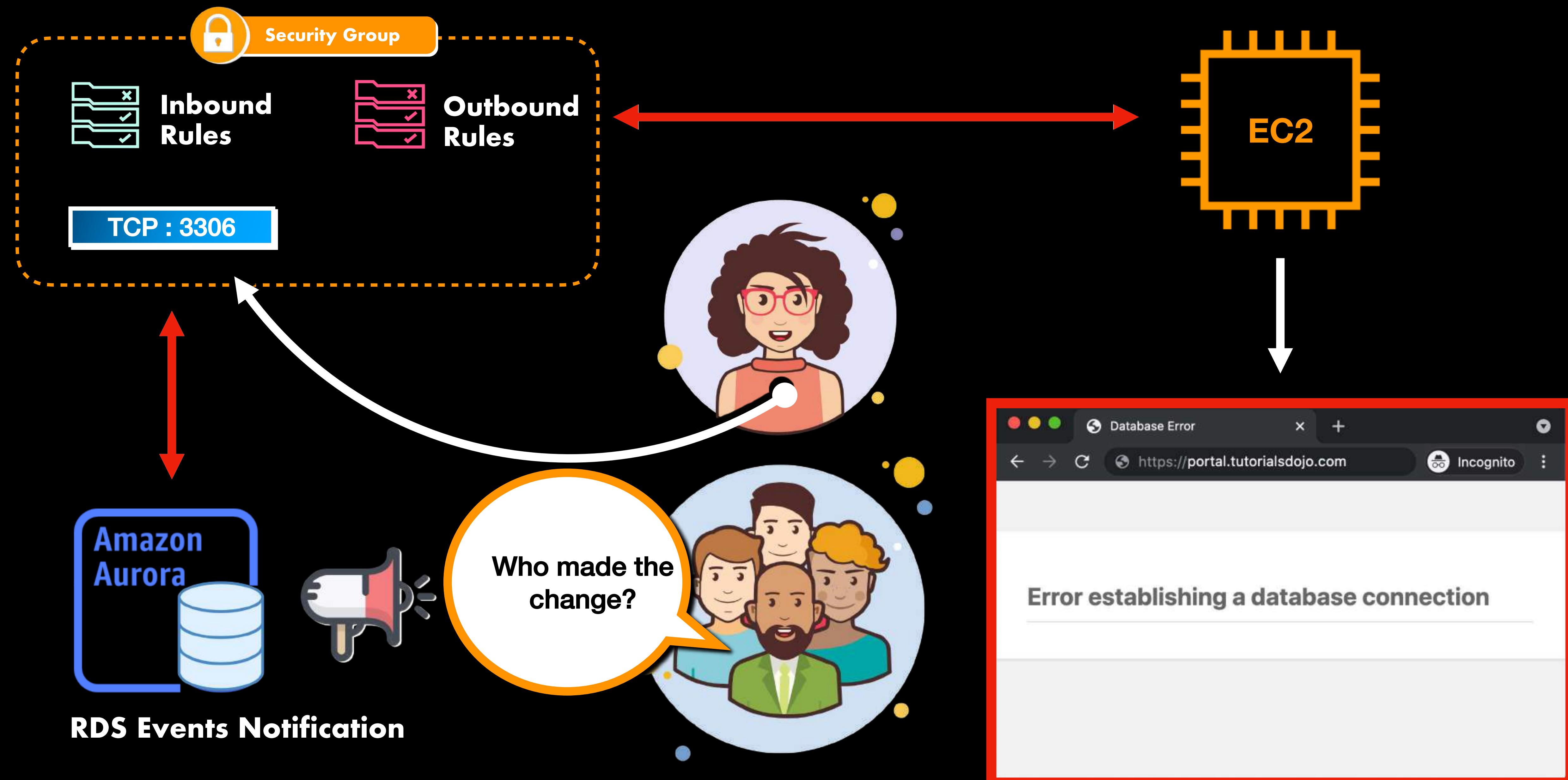


# Amazon RDS **Event Notification**

---

## Your Database







**RDS Events Notification**



Services ▾

Q Maghanap ng services, features, marketplace products, at mga docs [Option+S]



Tutorials Dojo ▾

N. Virginia ▾

Support ▾

## Amazon RDS



RDS &gt; Events

### Events (17)

Filter events

&lt; 1 &gt;



- Dashboard
- Databases
- Query Editor
- Performance Insights
- Snapshots
- Automated backups
- Reserved instances
- Proxies
- Subnet groups
- Parameter groups
- Option groups
- Custom Availability Zones
- Events**
- Event subscriptions
- Recommendations 6
- Certificate update

Source	Type	Time	Message
tutorialsdojo-manila	Instances	July 18, 2023, 3:37:56 AM UTC	DB instance stopped
tutorialsdojo-bengaluru	Instances	July 18, 2023, 3:30:25 AM UTC	DB instance restarted
tutorialsdojo-new-jersey	Instances	July 18, 2023, 3:29:56 AM UTC	Disabled automated backups
tutorialsdojo-new-delhi	Instances	July 18, 2022, 3:02:53 AM UTC	DB instance shutdown
tutorialsdojo-tacloban	Instances	July 18, 2023, 1:12:12 AM UTC	Finished DB Instance backup
tutorialsdojo-divisoria	Snapshots	July 18, 2023, 1:12:11 AM UTC	Automated snapshot created
tutorialsdojo-new-clark-city	Snapshots	July 18, 2023, 1:10:48 AM UTC	Creating automated snapshot
tutorialsdojo-leeds	Instances	July 18, 2023, 1:10:48 AM UTC	Backing up DB instance
tutorialsdojo-berlin	Instances	July 18, 2023, 1:10:22 AM UTC	Finished DB Instance backup
tutorialsdojo-france	Snapshots	July 18, 2023, 1:10:21 AM UTC	Automated snapshot created

## SOURCE TYPE

- Instances
- Security Groups
- Parameter Groups
- Snapshots
- Clusters
- Cluster Snapshots

**Source**

**Source Type**  
Source type of resource this subscription will consume events from

Parameter groups

**Parameter groups to include**  
Parameter groups that this subscription will consume events from

All parameter groups  
 Select specific parameter groups

**Specific parameter groups**

Select parameter groups

tutorialsdojo-divisoria X

**Source**

**Source Type**  
Source type of resource this subscription will consume events from

Instances

**Instances to include**  
Instances that this subscription will consume events from

All instances  
 Select specific instances

**Specific instances**

Select instances

tutorialsdojo-pasig X

## SOURCE TYPE

## EVENT CATEGORIES

### Source

#### Source Type

Source type of resource this subscription will consume events from

Instances

#### Instances to include

Instances that this subscription will consume events from

- All instances
- Select specific instances

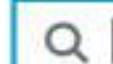
#### Event categories to include

Event categories that this subscription will consume events from

- All event categories
- Select specific event categories

#### Specific event categories

Select event categories



availability

backtrack

backup

configuration change

creation

deletion

failover

failure

low storage

maintenance

notification

read replica

recovery

restoration

SOURCE TYPE

EVENT CATEGORIES

**Source**

**Source Type**  
Source type of resource this subscription will consume events from

Clusters

**Clusters to include**  
Clusters that this subscription will consume events from

All clusters  
 Select specific clusters

**Event categories to include**  
Event categories that this subscription will consume events from

All event categories  
 Select specific event categories

**Specific event categories**

Select event categories

Q |

- configuration change
- creation
- deletion
- failover
- failure
- global-failover
- maintenance
- notification

## TARGET TYPE



**Amazon SNS**

### Target

Send notifications to

- ARN
- New email topic

Topic name

Name of the topic

With these recipients

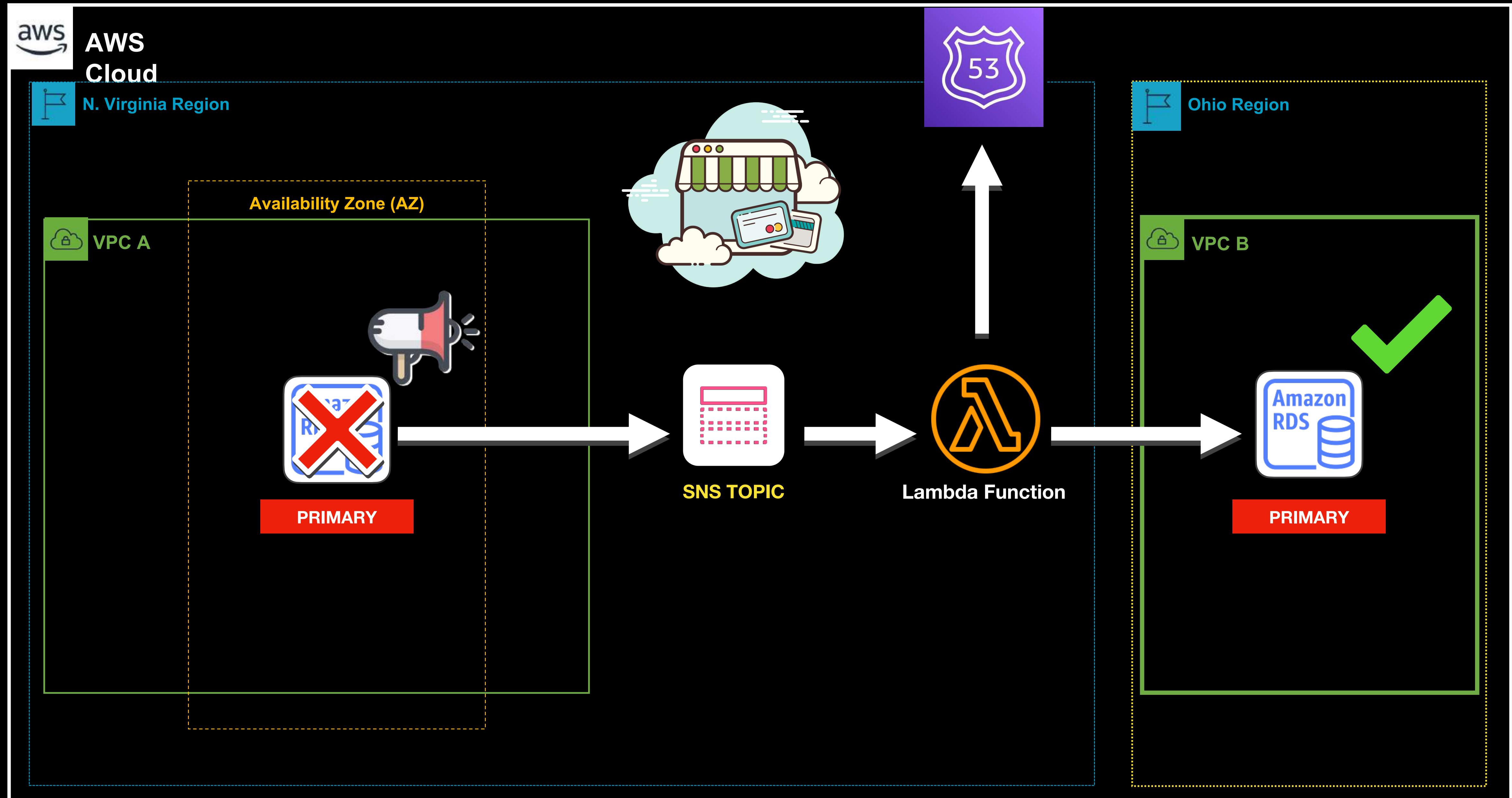
Email addresses to send the notifications to

e.g. user@domain.com



**FANOUT EVENT NOTIFICATIONS**



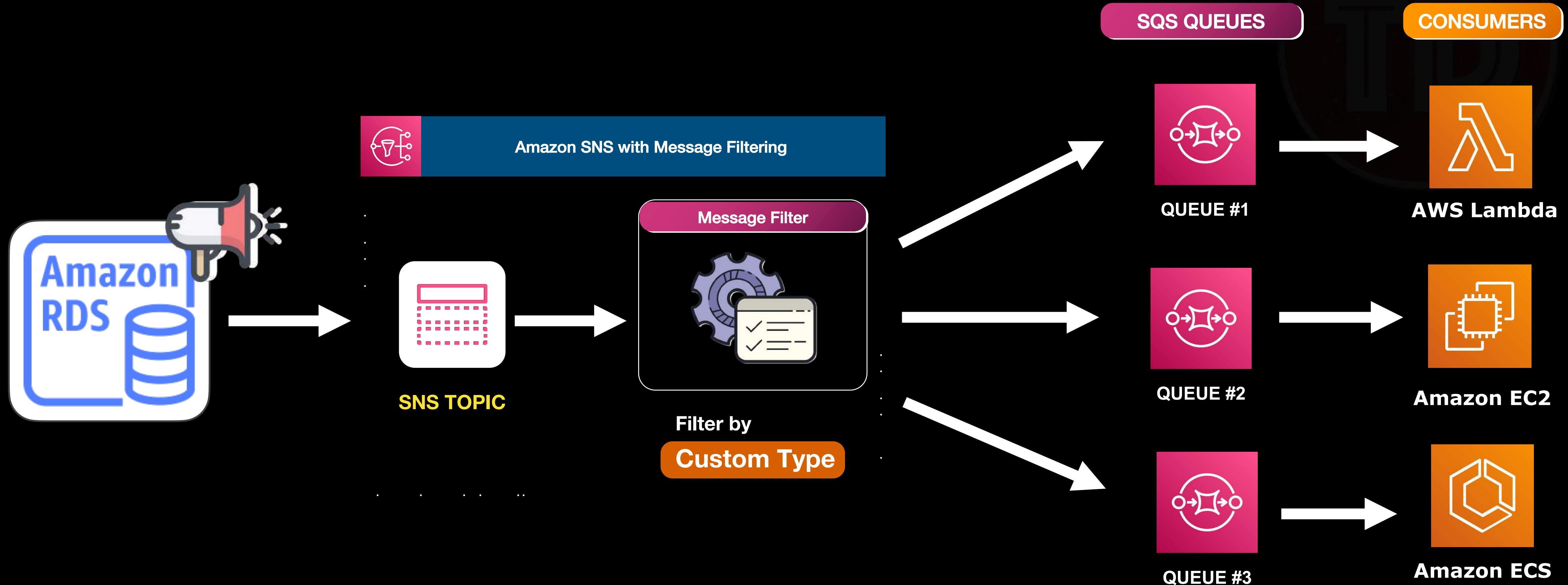




FANOUT EVENT NOTIFICATIONS



## FANOUT EVENT NOTIFICATIONS





# Amazon RDS **Multi-AZ Deployments**

---

## SYNCHRONOUS REPLICATION



STANDBY REPLICA

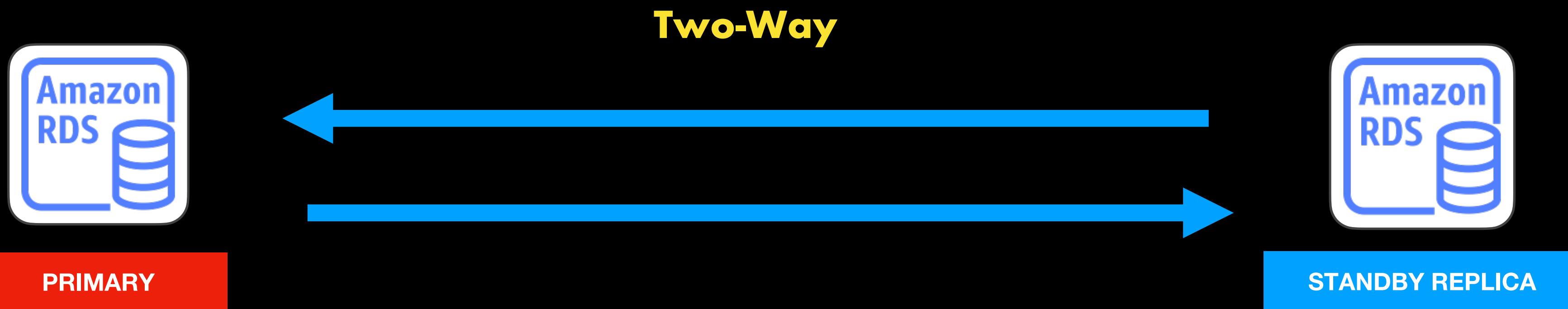
## ASYNCHRONOUS REPLICATION



READ REPLICA

**REPLICA**  
a copy of your primary database

# SYNCHRONOUS REPLICATION

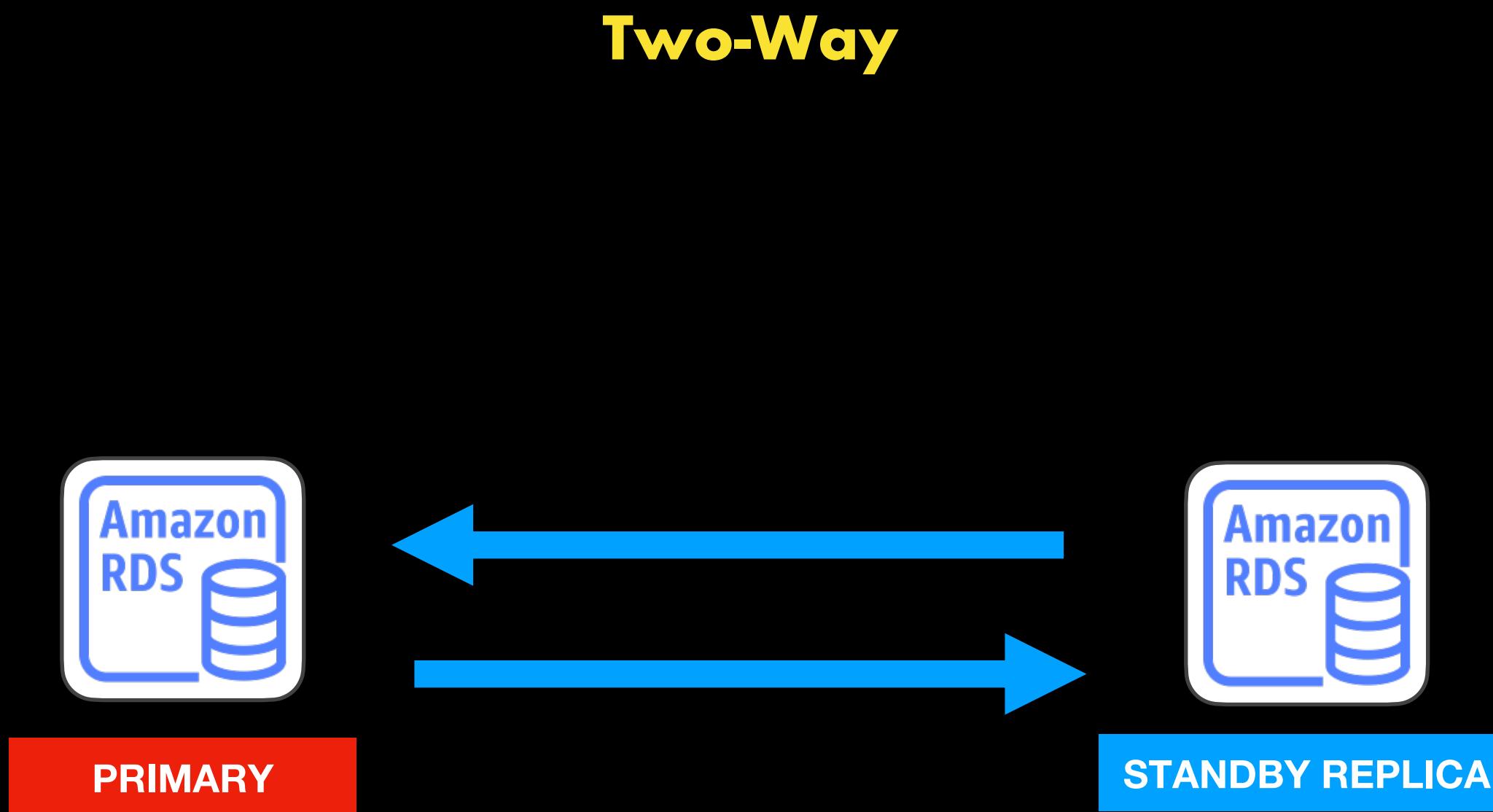


```
INSERT INTO CITIES (Name, Country)  
VALUES ('Manila', 'Philippines');
```

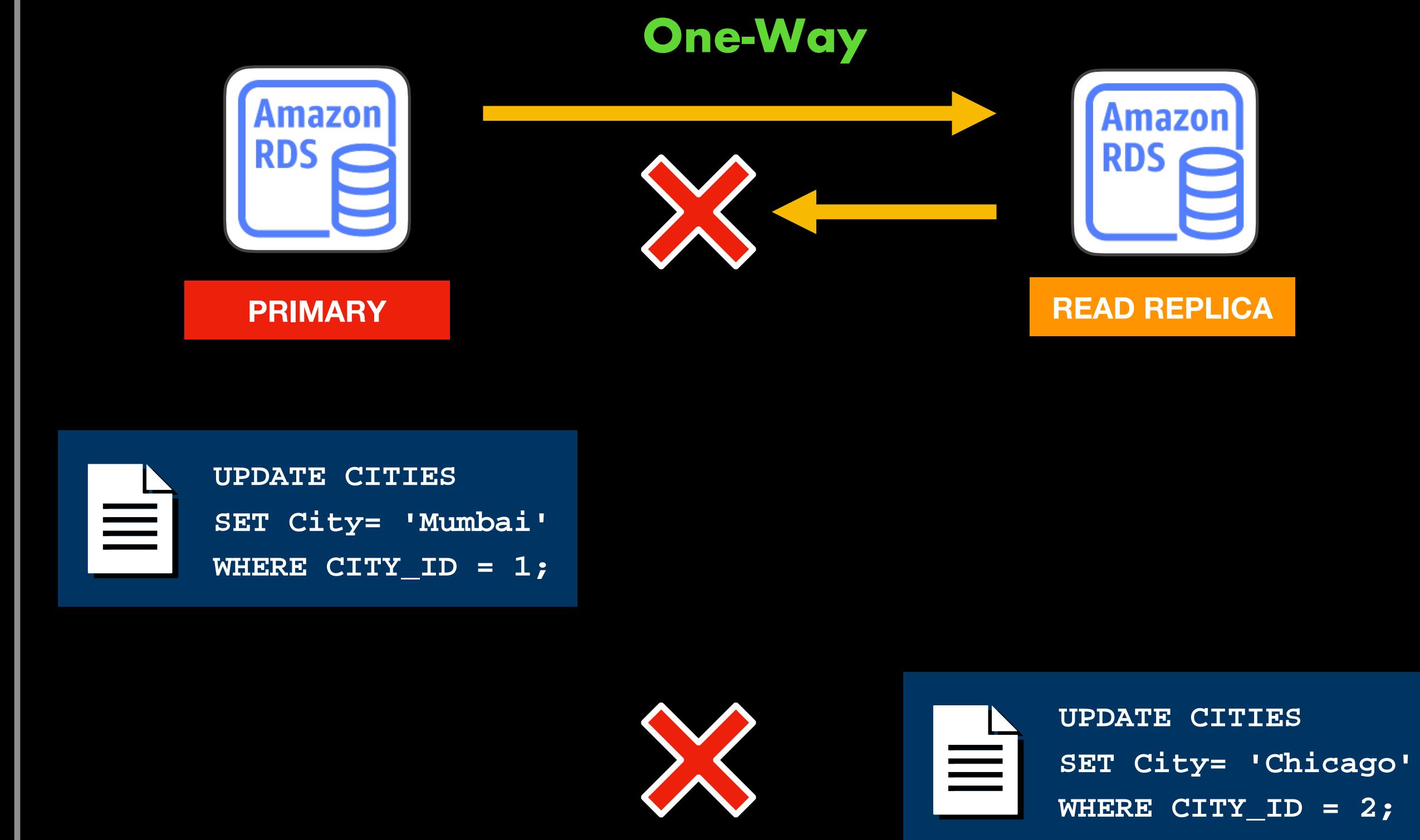


```
INSERT INTO CITIES (Name, Country)  
VALUES ('Toronto', 'Canada');
```

# SYNCHRONOUS REPLICATION



# ASYNCHRONOUS REPLICATION



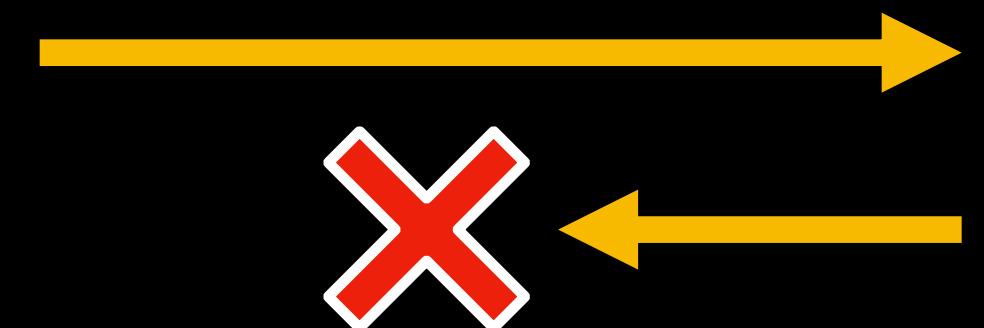
**STANDALONE**

**MASTER-SLAVE  
CONFIGURATION**

**Single DB Instance (Single AZ)**



**Read Replica**

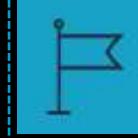


**Multi-AZ Deployments**

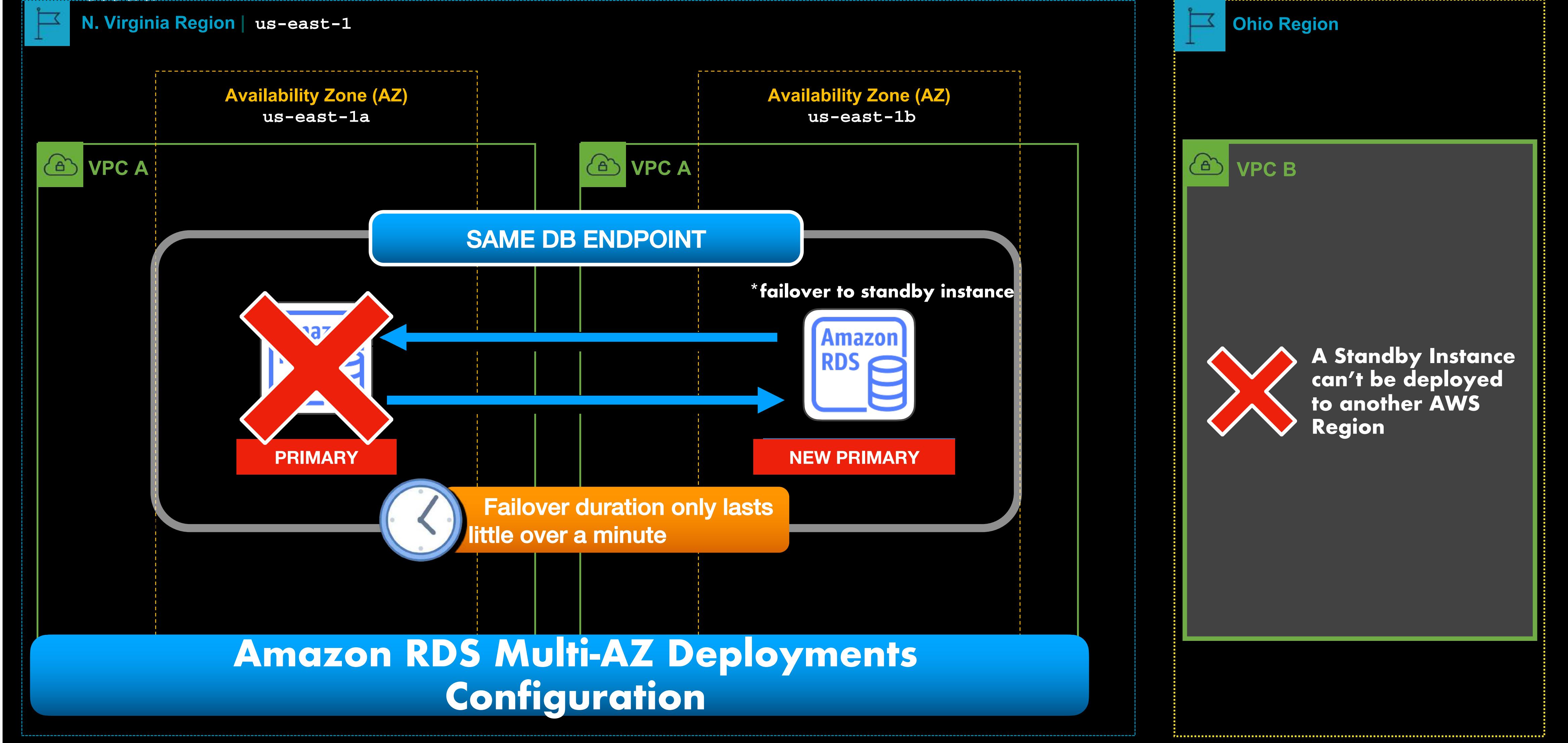


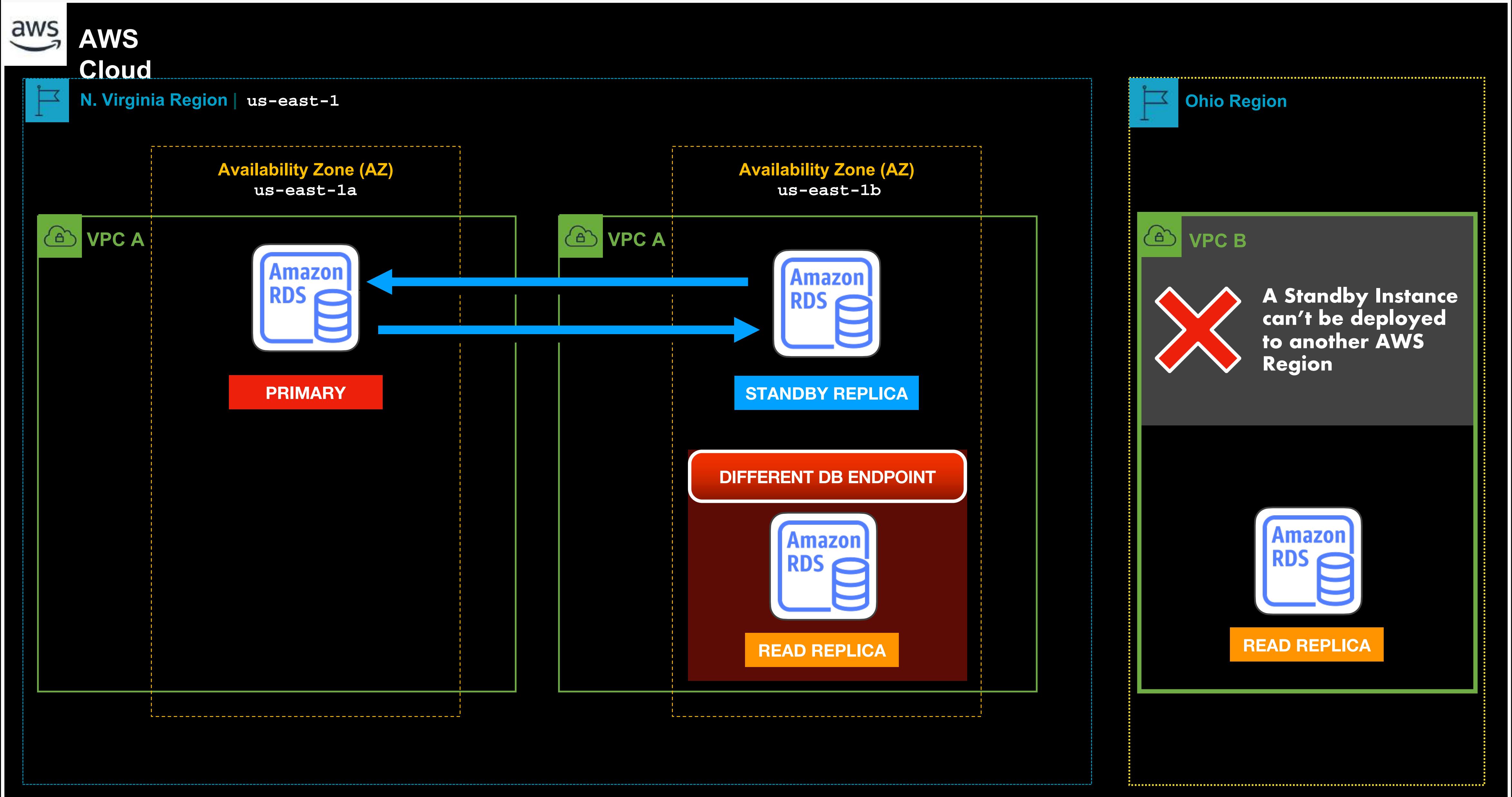


AWS  
Cloud



N. Virginia Region | us-east-1

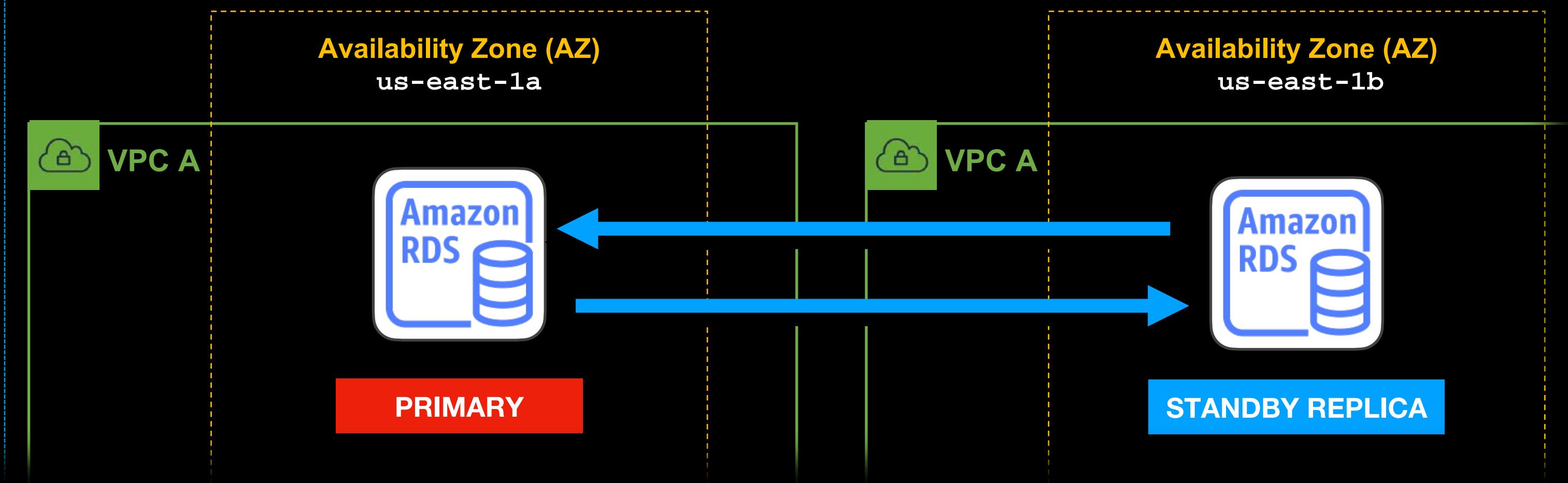




Cloud



N. Virginia Region | us-east-1



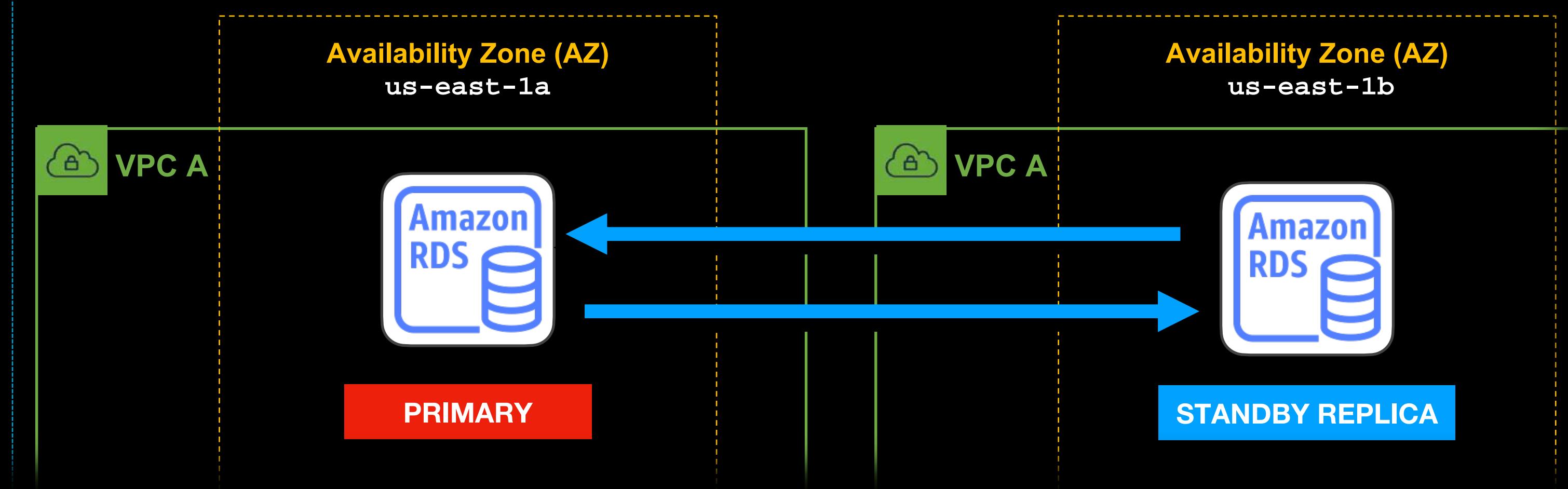
## Amazon RDS Multi-AZ Deployments Configuration

- Provides High Availability
- Improves Data Redundancy
- Minimizes latency spikes during system backups
- Keeps your database available on your planned system maintenance or DB Engine upgrade
- Protects your database against DB instance failure and disruptions when an Availability Zone outage occurs

Cloud



N. Virginia Region | us-east-1

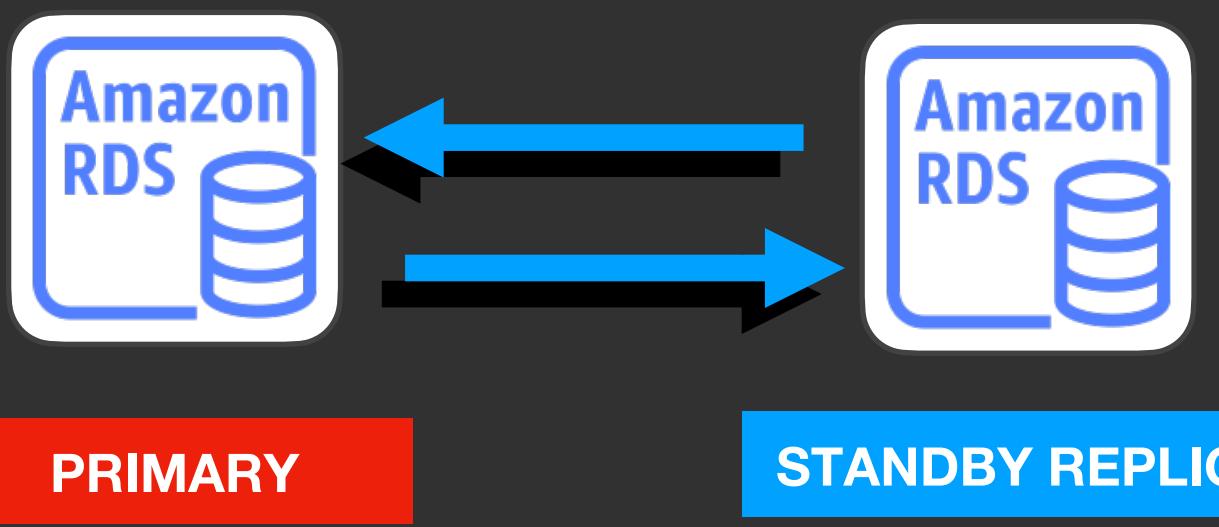


## Amazon RDS Multi-AZ Deployments Configuration

### Multi-AZ Deployments Configuration – Internal Steps

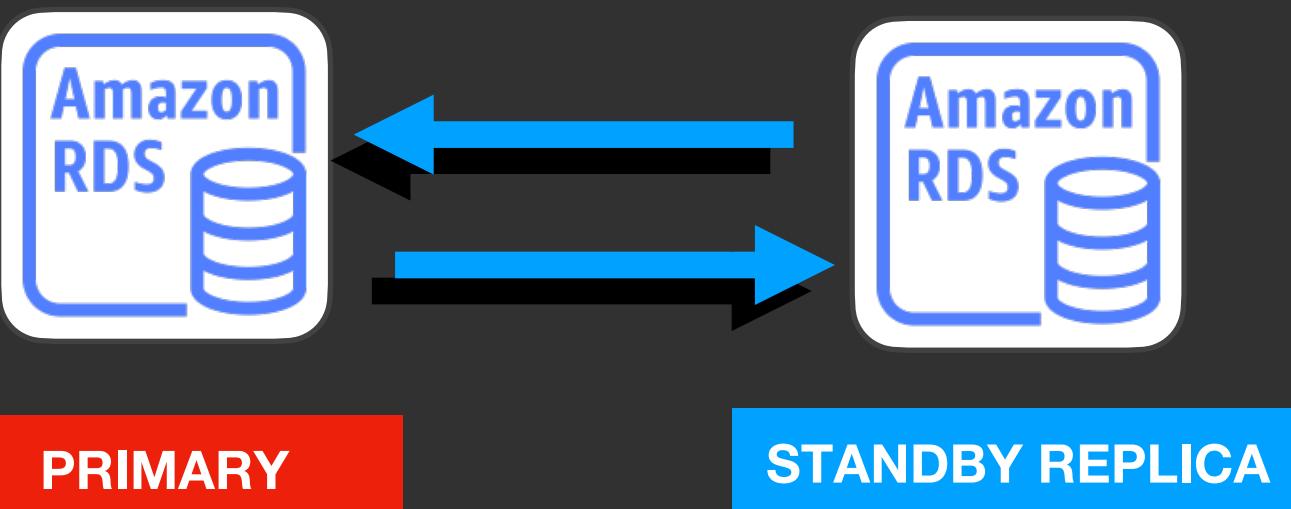
1. Takes a snapshot of your primary DB instance
2. Launch a new Standby Instance in a different Availability Zone
3. Automatically configure synchronous replication between the primary and standby instances

## Multi-AZ Deployments



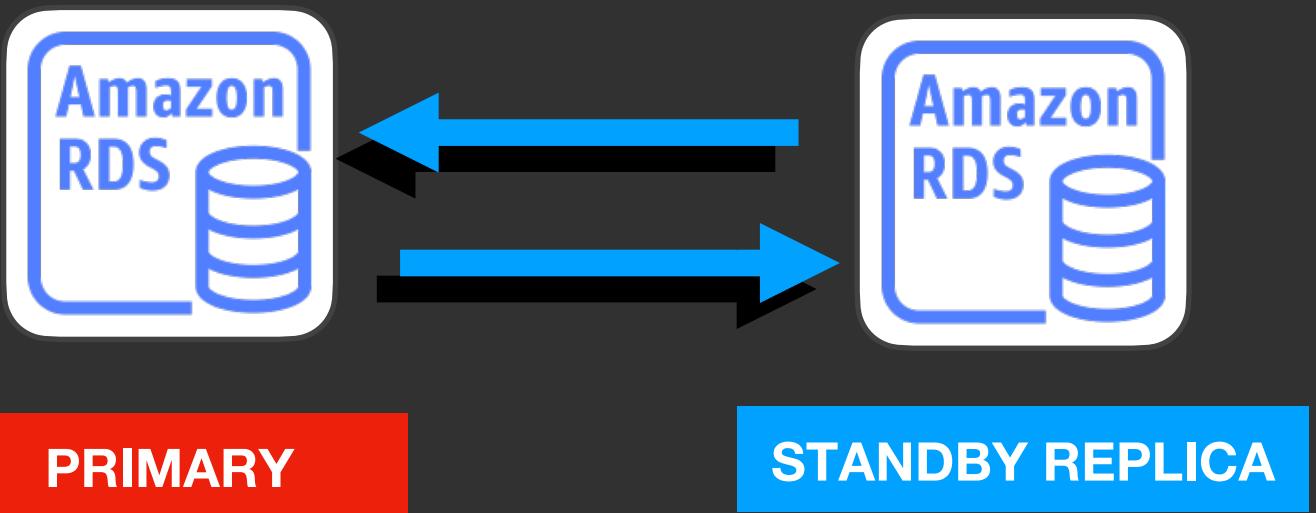
- Amazon RDS uses an internal Amazon EC2 instance that has its own operating system and attributes
- Maintains database performance while the regular process of patching the database engine is on-going
- Ensures the availability of your database when the OS and its underlying hardware go through its scheduled maintenance activities

## Multi-AZ Deployments



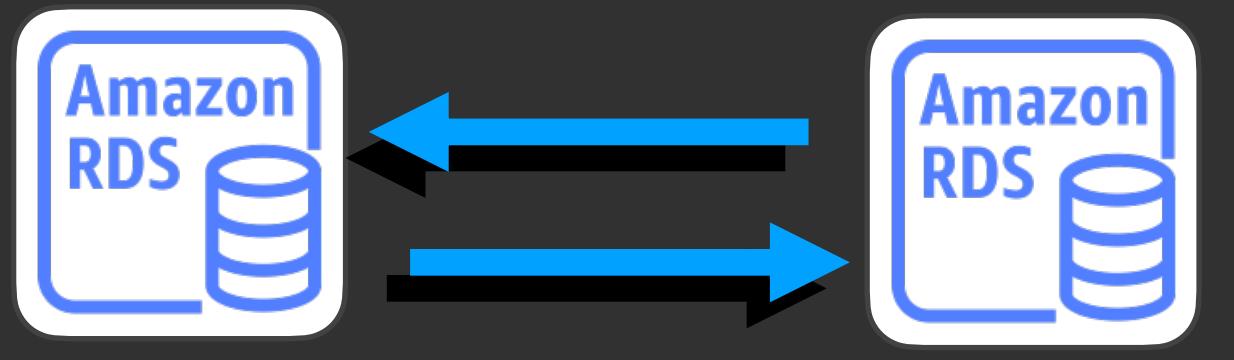
- During an AWS-initiated hardware maintenance, a Multi-AZ database will only have a **minimal disruption** unlike a Single-AZ database
- Your database will only be unavailable during the primary DB instance **failover to the Standby Replica**
- The duration of the failover process to the Standby Replica is only about 1 minute or so

## Multi-AZ Deployments



- When the automatic failover in Amazon RDS occurs, the **Canonical Name record (CNAME)** of your DB instance is automatically altered to point to the newly promoted Standby Instance
- If AWS conducts a hardware maintenance on the Availability Zone where your Standby Replica is hosted, your Multi-AZ RDS database will not experience any failover or downtime
- The **Operating System (OS) patch will be applied to the Standby Replica first before it is installed to the primary instance**
- The only downtime would be the failover process

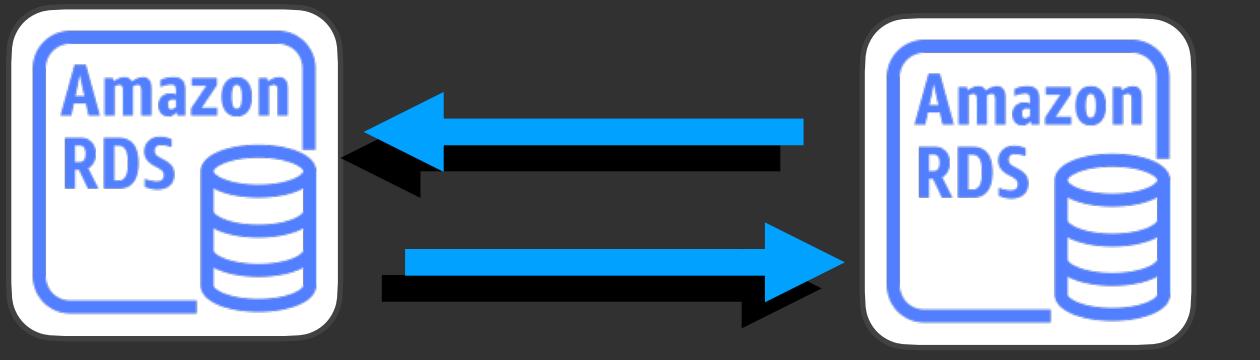
## Multi-AZ Deployments



### U S E   C A S E S

- Suitable for **mission-critical applications where you need the highest availability while minimizing your operational and management overhead.**
- Applicable if you have an application running in your production environment that uses a single-instance RDS database
- If you want to migrate your existing database running on your on-premises network, that is running on a single database configuration
- If you are required to **eliminate single points of failure in your architecture**

## Multi-AZ Deployments



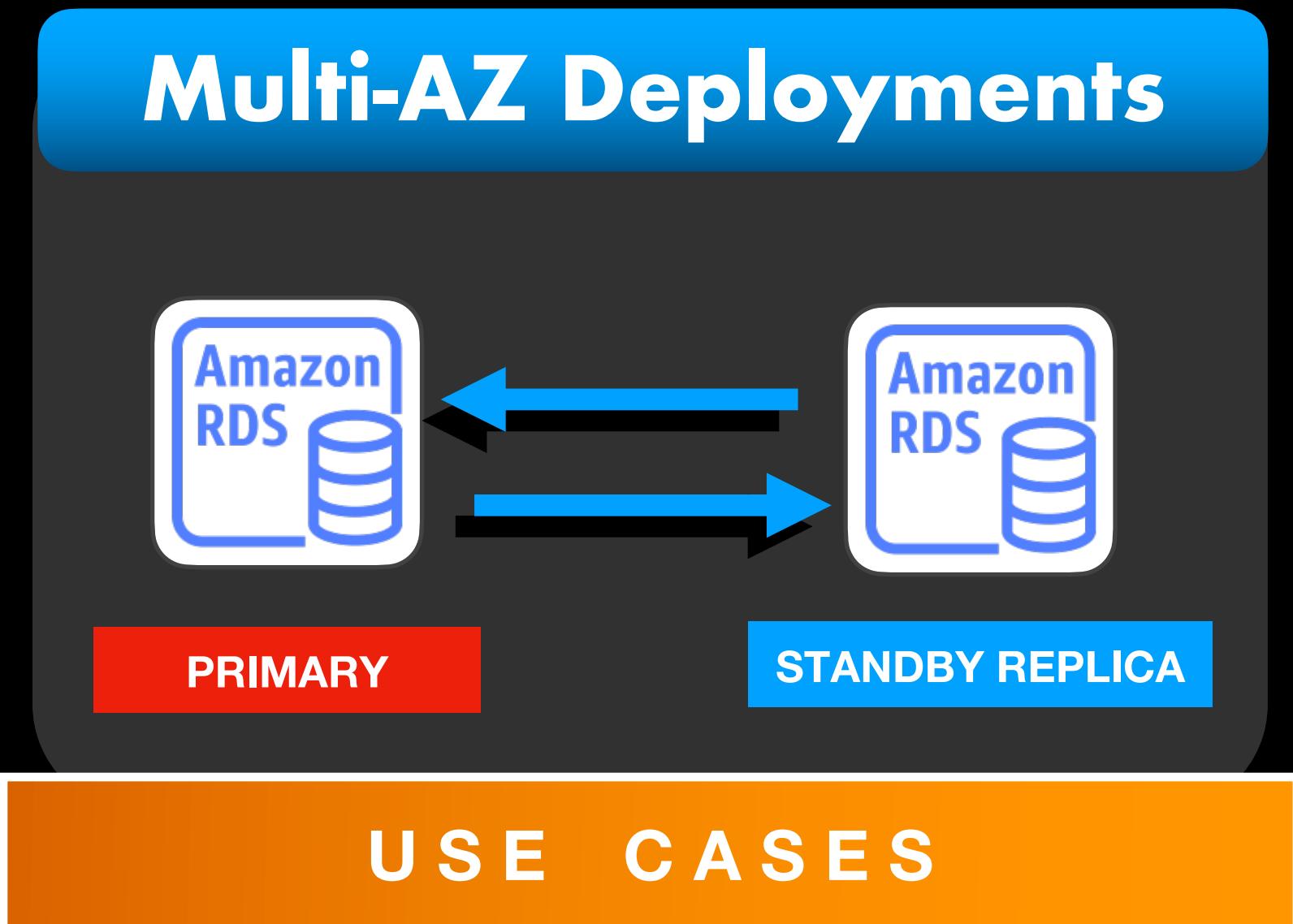
PRIMARY

STANDBY REPLICA

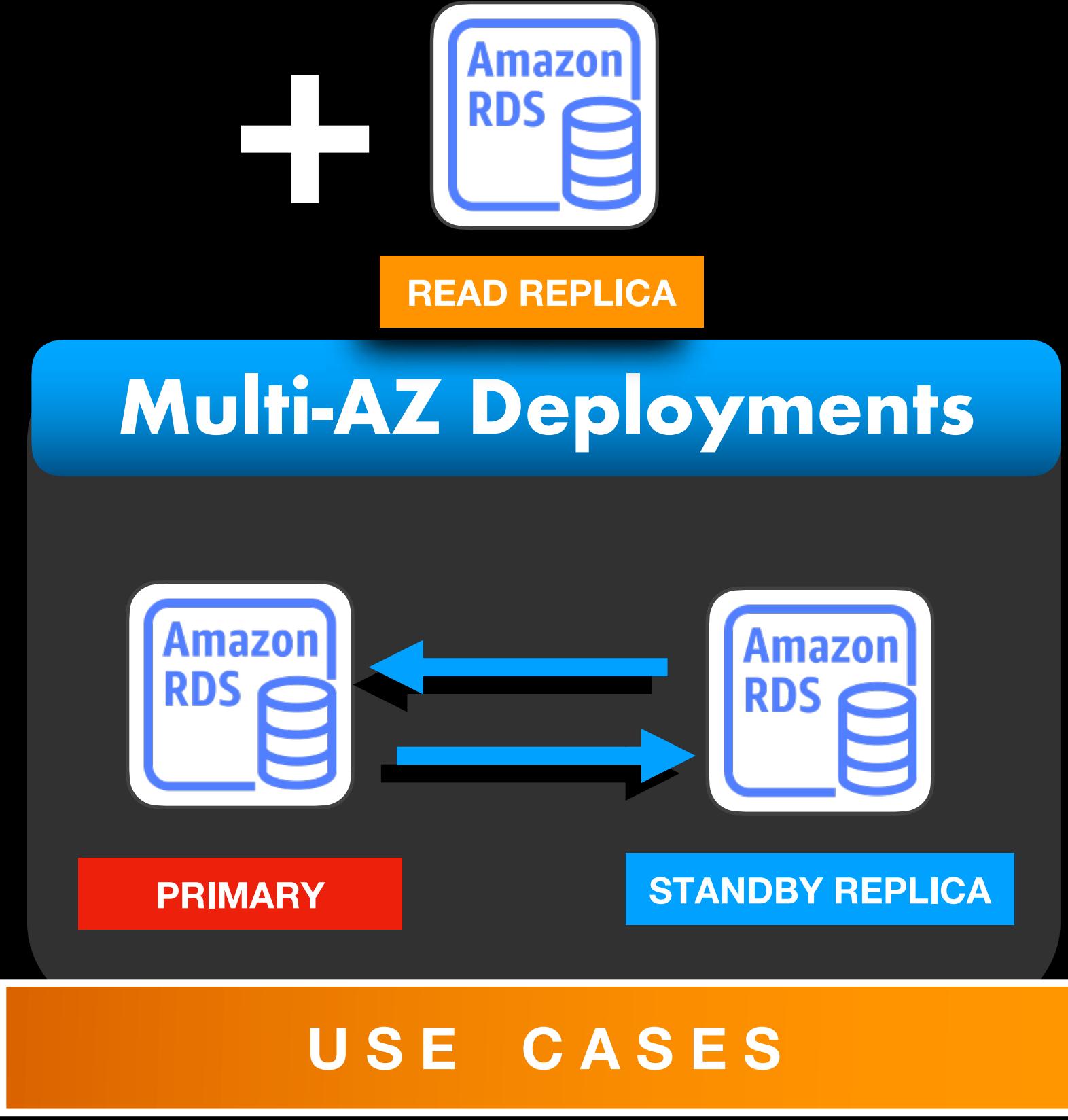
U S E   C A S E S

- For minimizing database downtime **without requiring any changes to your application code**
- For enterprise systems that need to be **highly available with low operational complexity**
- For any scenario where **the availability of your database is the highest priority/most important requirement and not its scalability**

- For poorly-designed architectures that needs to be re-designed/refactored, such as:
  - A three-tier application architecture runs in public and private subnets
  - The application is running on a **single Amazon EC2 instance** that is hosted in the public subnet
  - A **single Amazon RDS database** running on the private subnet

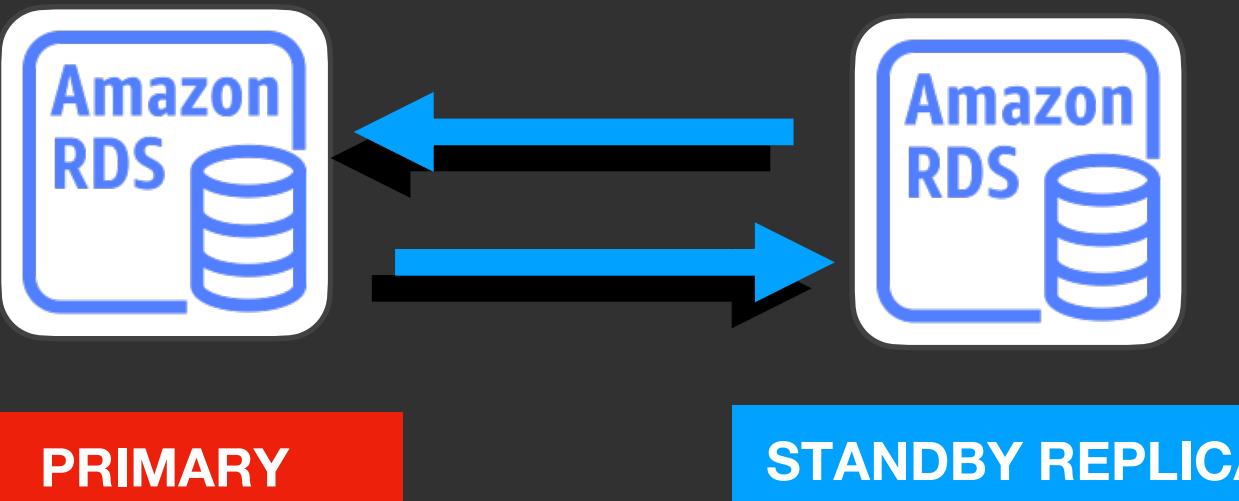


- Improved architecture:
  - Launch an Auto Scaling group of EC2 instances behind an Application Load Balancer that spans multiple AZs
  - Enable the Multi-AZ Deployments configuration in Amazon RDS to **make the database tier highly available**



- You can **combine Multi-AZ Deployments configuration with Read Replicas**
- A Read Replica can provide **cross-region database replication for multi-Region disaster recovery**, which a Multi-AZ Deployment configuration can't provide
- Having both Standby and Read Replica ensures both **high availability and scalability** of your database tier

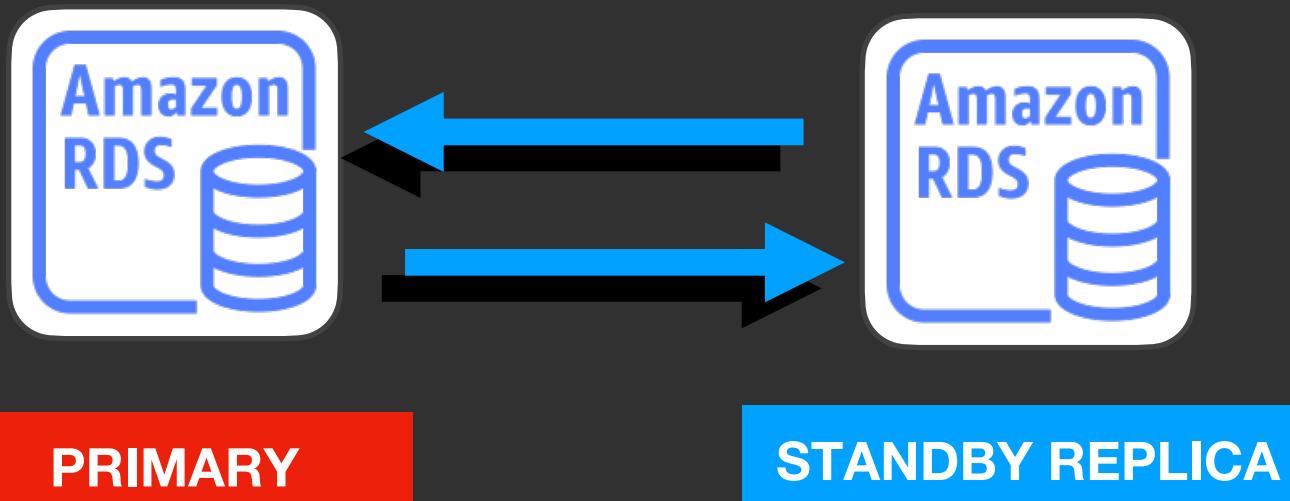
## Multi-AZ Deployments



## LIMITATIONS

- A Multi-AZ database can provide high availability in a **single AWS Region only**
- You cannot deploy a Standby Replica to another AWS Region
- Does not provide multi-region disaster recovery
- The Standby Replica cannot be used to read or write your application data, or accept live traffic
- Cannot be used this to scale your application in terms or read performance or handle the increased number of queries to your database

## Multi-AZ Deployments



### LIMITATIONS

- Not suitable if the required Recovery Point Objective (RPO) and Recovery Time Objective (RTO) are quite short
- It cannot provide an RPO of 1 second and an RTO of 1 minute
- If you have this requirement, you have to use:



Amazon Aurora  
Global Databases

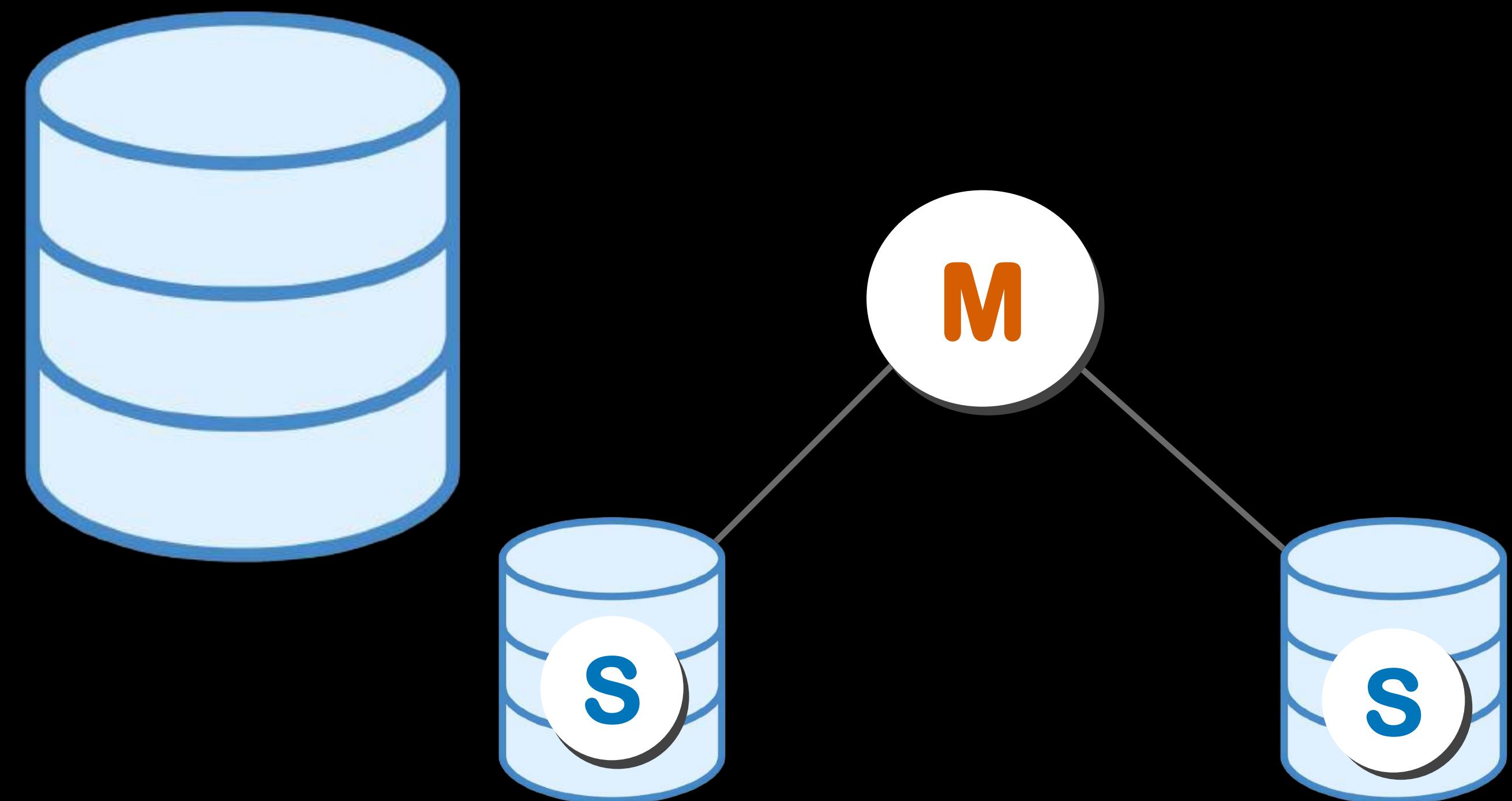


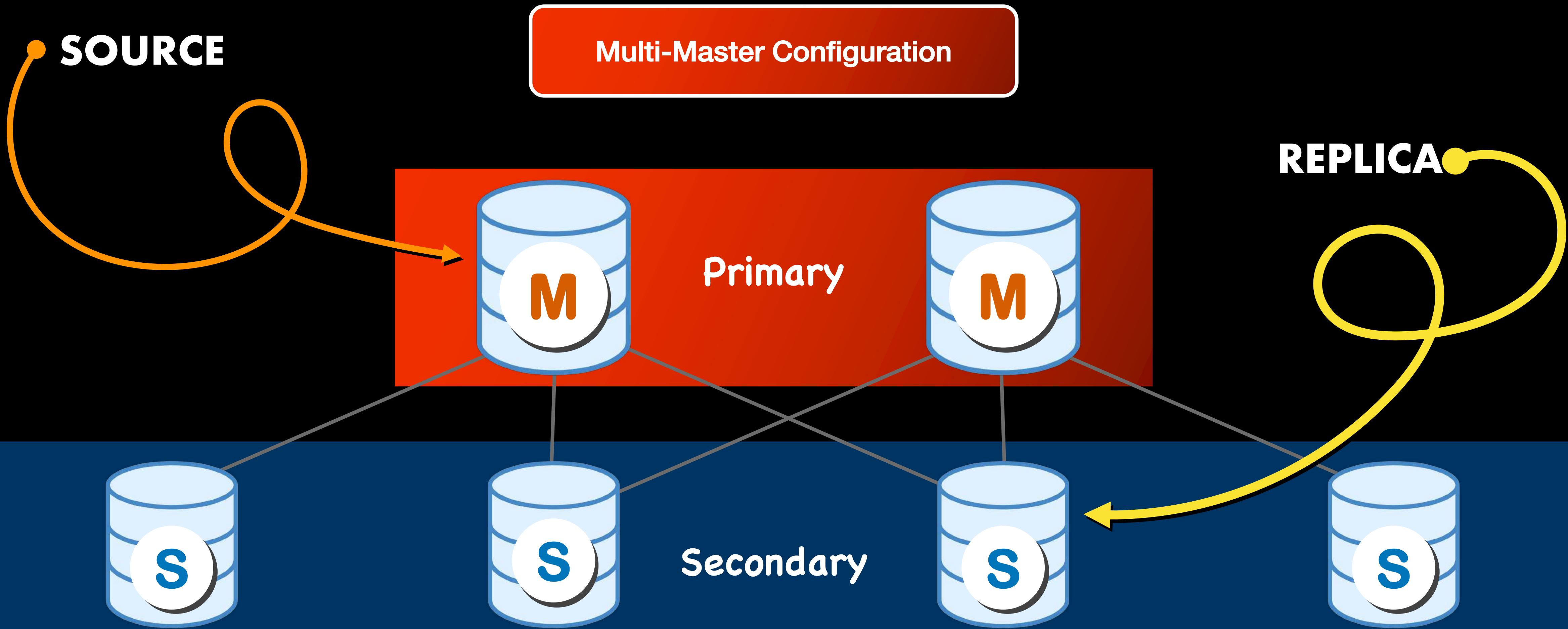
# Amazon RDS Read Replica

---

**STANDALONE**

**MASTER-SLAVE  
CONFIGURATION**





# **READ REPLICA**



STANDBY REPLICA

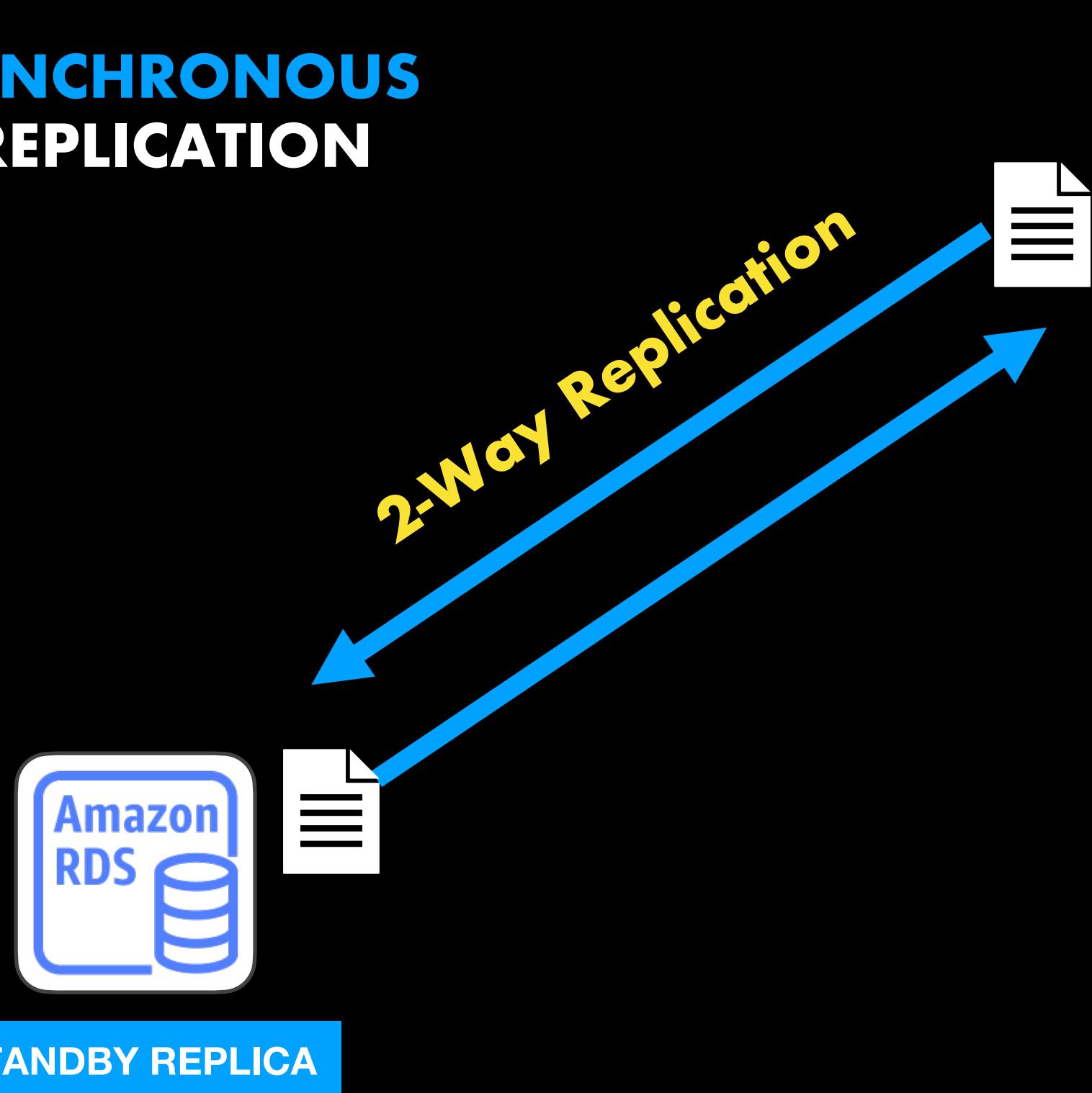
# REPLICA

a **copy of something**

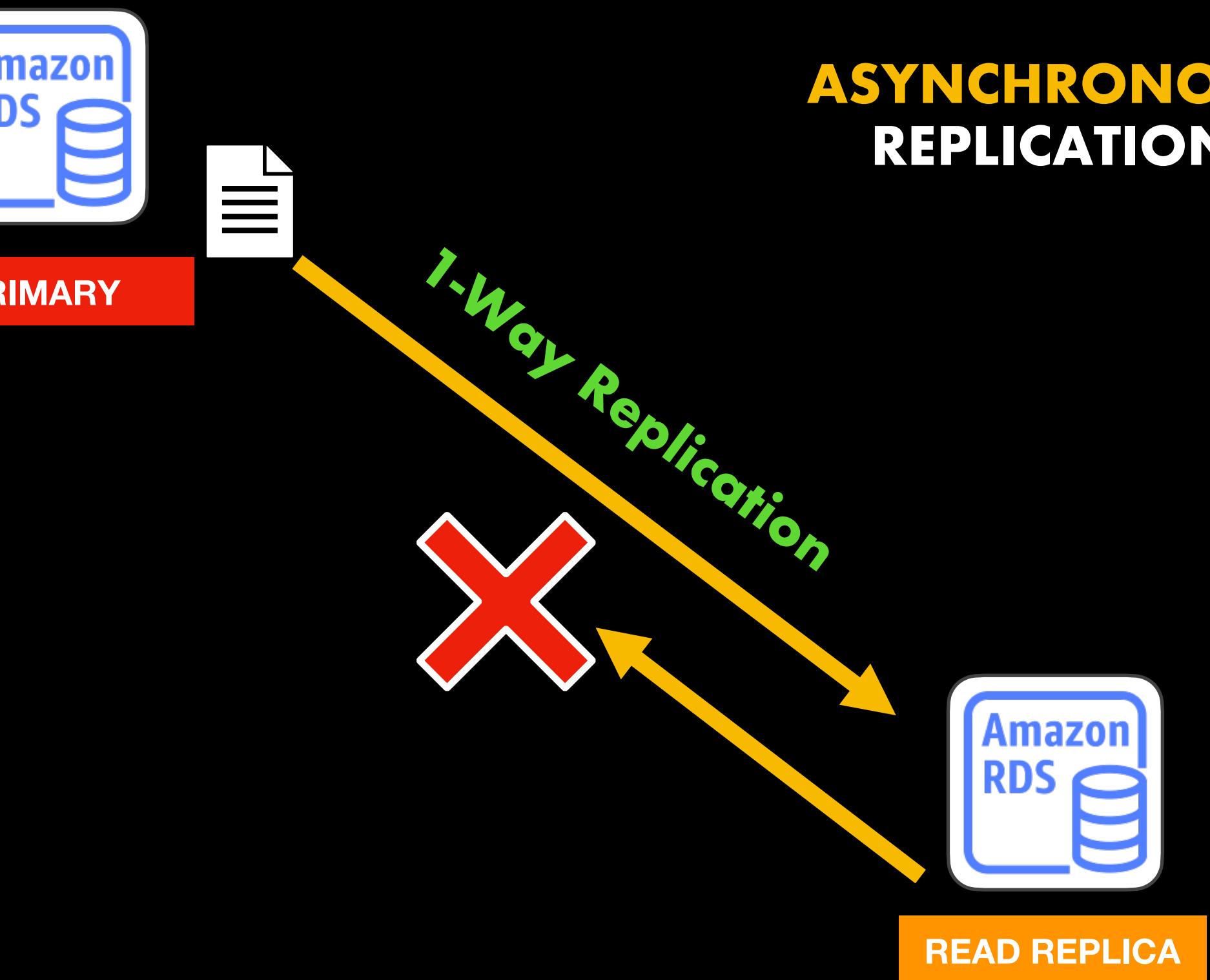


READ REPLICA

## SYNCHRONOUS REPLICATION



## ASYNCHRONOUS REPLICATION



- Does not accept live traffic without failover
- Cannot be seen in the Amazon RDS Console as a separate DB instance
- The DB Endpoint is the same as the primary DB instance

- Can accept live traffic
- Can be seen in the Amazon RDS Console as a separate DB instance
- The DB Endpoint is different from the primary DB instance



- Based on the **built-in replication functionality** of:



- Just a regular database with a **read-only configuration**
- Under the hood, Amazon RDS creates this by cloning your source database, setting up the replication parameters, and disabling any write operations



READ REPLICA

my.cnf

```
# Default MySQL Server config for Tutorials Dojo Divisoria
[mysqld]

# Only allow connections from localhost
bind-address = 127.0.0.1
mysqlx-bind-address = 127.0.0.1

# Set the database to read-only
read_only = 1

# Block any write operations from DB Administrators
super_read_only = 1
```

**X**

CREATE  
INSERT  
UPDATE  
DELETE



READ REPLICA

## Other required parameters for **binary logging** to be set:

- **log\_bin**
- **binlog-format**
- **sync\_binlog**
- ...and many more!

```
# Default MySQL Server config for Tutorials Dojo Divisoria
[mysqld]

# Only allow connections from localhost
bind-address = 127.0.0.1
mysqlx-bind-address = 127.0.0.1

# Set the database to read-only
read_only = 1

# Block any write operations from DB Administrators
super_read_only = 1
```

- **A binary log**

- **Also known as '*binlog*'**
- **A set of log files that contain information about the recent SQL modifications**
- **Contains all of the CREATE, INSERT, UPDATE, DELETE, ALTER, and other SQL statements that were made in your primary database**
- **The actual data that is being transferred from the source database to the database replica**



READ REPLICA



READ REPLICA

- Can be launched two ways:
  - On the same AWS Region of your primary DB
  - On a different AWS region
- Does NOT provide the capability of directly accessing the actual configuration files – `my.cnf` (MySQL), `ConfigurationFile.ini` (MS SQL) and others in Amazon RDS
- View and modify the DB configuration of the replica using a **parameter group**



READ REPLICA



**READ REPLICA**

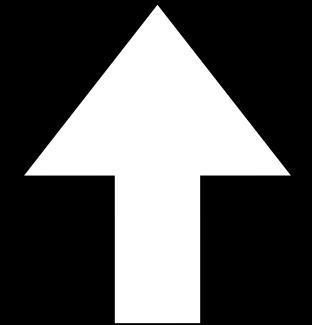
aws Services ▾ #KayangKayaNgPinoy X Tutorials Dojo ▾ N. Virginia ▾ Support ▾

RDS > Parameter groups > Tutorials-Dojo-Manila

## Tutorials-Dojo-Manila

Parameters				
		Cancel editing Preview changes Reset Save changes		
<input type="text"/> read_only X < 1 >				
Name	Values	Allowed values	Description	
innodb_read_only		0, 1	Starts the server in read-only mode.	
read_only	1	0, 1, {TrueIfReplica}	When it is enabled, the server permits no updates except from updates performed by slave threads.	
super_read_only		0, 1	Whether client connections to the server are required to use some form of secure transport.	

**PRIMARY**



**READ REPLICA**

- Can be promoted to be a standalone DB instance
- Useful for:
  - Database sharding
  - Implementing failure recovery
  - Performing Data Definition Language (DDL) operations
- Lessens the impact to the primary DB instance brought by rebuilding indexes, scheduled jobs, and other processing
- Helpful if your primary AWS Region experiences an outage
- Can be deployed to a different AWS Region and be promoted as the primary DB instance in the event that the AWS Region of your source/primary database experiences a downtime



- **Cannot directly create an encrypted Read Replica from an unencrypted database instance**
- **Can be created from your encrypted database instances but not from the unencrypted ones**
- **An encrypted cross-region read replica can be launched as long as the target region and an encryption key in AWS KMS for that particular region are supplied**
- **Allows the use of a custom encryption key or the default encryption key for Amazon RDS that is created by AWS KMS in each region**



READ REPLICA

USE CASES

- Suitable if your company has a web application with a **built-in reporting module**
- If your department or application runs large SQL queries every month that impact your database's performance due to high usage
- If you need to minimize the impact that the reporting activity has on your application by offloading the read requests



## READ REPLICA

USE CASES

- If you need to separate the read requests from the write requests of your application
- If you have an application wherein the read operations are causing high I/O usage to your primary RDS database instance which then results in high latency to the write requests in your production environment
- If you have application modules or reporting tools that only send SELECT queries. You can configure the reporting module to use the Read Replica endpoint and direct the transactional operations to the primary database instance



## READ REPLICA

USE CASES

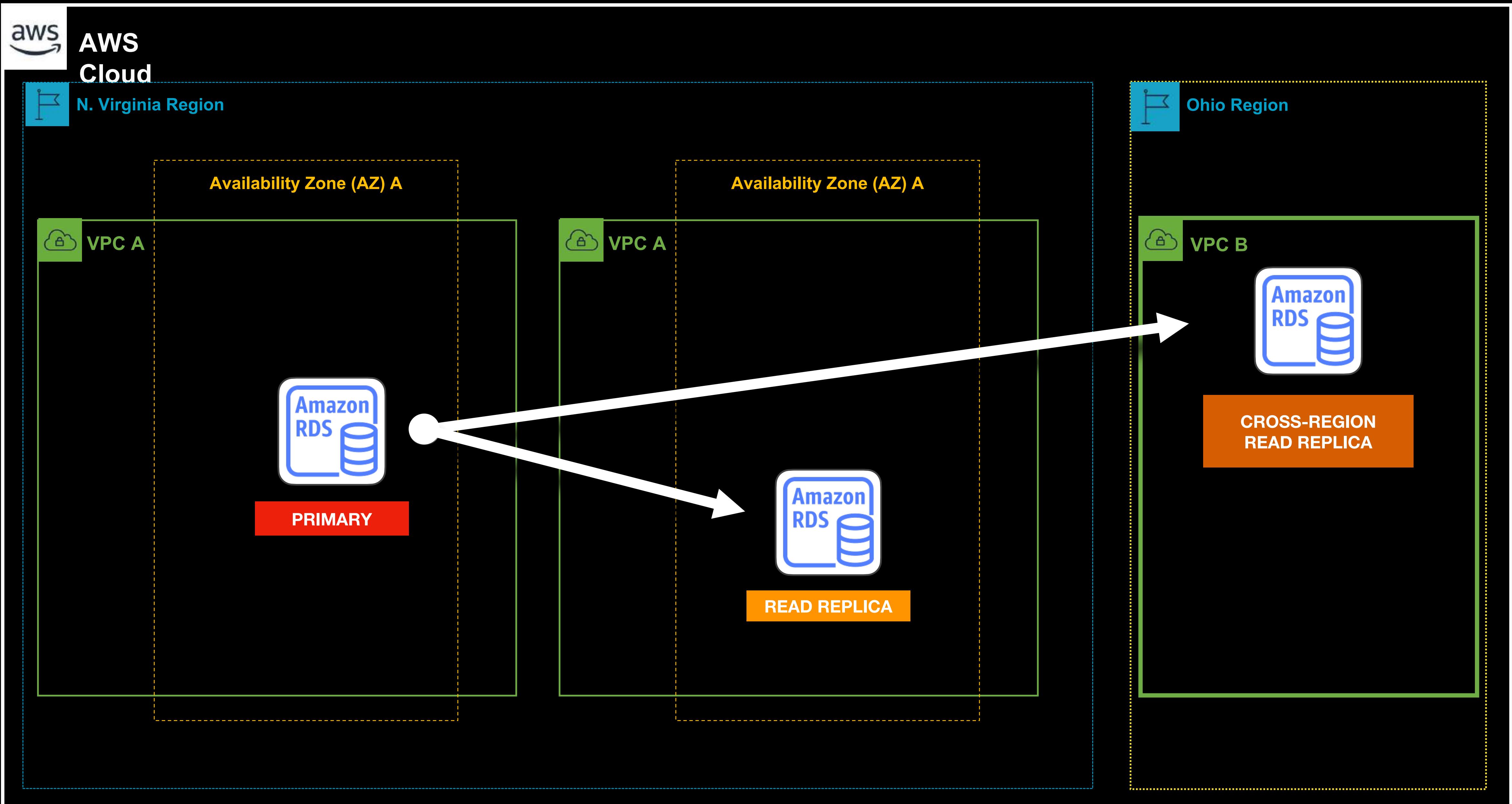
- If you have **3rd-party applications** or other internal systems that query your database instance heavily
- If you have an internal batch processing job that fetches reporting data from your RDS DB instance.
- If your entire database slows down significantly whenever your batch runs which impacts the overall read and write performance of your application
- If you need to configure your internal systems to fetch data from the replica instead of the primary instance



## READ REPLICA

ANTI-PATTERNS

- A Read Replica is primarily used to **improve the scalability of your application** in terms of read operations and not for improving the availability of your database
- Cannot be used for ensuring that the database will be highly available in the event of an outage. You have to use the Multi-AZ Deployments configuration instead
- Unlike Multi-AZ RDS, a Read Replica doesn't have an automatic failover. If the primary DB instance experienced an outage, the incoming requests are not automatically routed to the Read Replica by default





# Amazon Aurora Overview



Amazon RDS



Amazon Aurora

- A **fully managed database service** and also a type of **database engine** within Amazon RDS
- Scales automatically, performs faster, and costs lower
- A relational database that is compatible with:





## Amazon Aurora

- Can automatically grow or scale its storage
- Usually deployed as a **database cluster**
- A cluster consists:

ONE PRIMARY



MULTIPLE REPLICAS



## CLUSTER TYPES



**Amazon Aurora**

Single-master

Multi-master

## STANDALONE TYPE

Single primary DB instance  
with no replica



## Amazon Aurora

- **Performs faster than other databases**
- **Can scale the computing components and storage automatically without any manual intervention**
- **The database cluster typically lags behind the primary instance by a few milliseconds only**
- **Provides less than 1 second of read replication latency for Aurora Replicas in the same or different AWS Region**

- Group the individual DB instances and associate them with a particular endpoint



## Amazon Aurora ENDPOINTS

Cluster endpoint

Reader endpoint

Custom endpoint

Instance endpoint



## Amazon Aurora Serverless

- Recommended for **sporadic usage workloads or with unpredictable usage**
- Pay your database usage **on a per-second basis**
- Provides a **more cost-effective option** than the regular Amazon RDS or Amazon Aurora databases

- For **migrating legacy applications** hosted on-premises that needs to be re-architected and reduce operating costs
- If it is required to re-architect your application by using technologies that **do not require any IT administration team to regularly manage your servers or clusters**
- If you **need to turn your monolithic application into microservices architecture with serverless resources**
- **Can be used for serverless stack** with the application containers running on AWS Fargate and your database on Aurora Serverless



## Amazon Aurora Serverless



## Amazon Aurora Serverless

- For **sporadic usage patterns**
- If your application has:
  - **Extremely high usage at the beginning of each month**
  - **An unpredictable usage at the start of each week**
  - **A moderate usage over the weekend**
- For situations where it is **difficult to predict the application demand or to choose the most suitable instance size of your database due to the constantly changing usage**
- If a **cost-effective database platform is required which does not require any database modifications**
- If you **need to automatically scale the capacity up or down based on your application's needs**



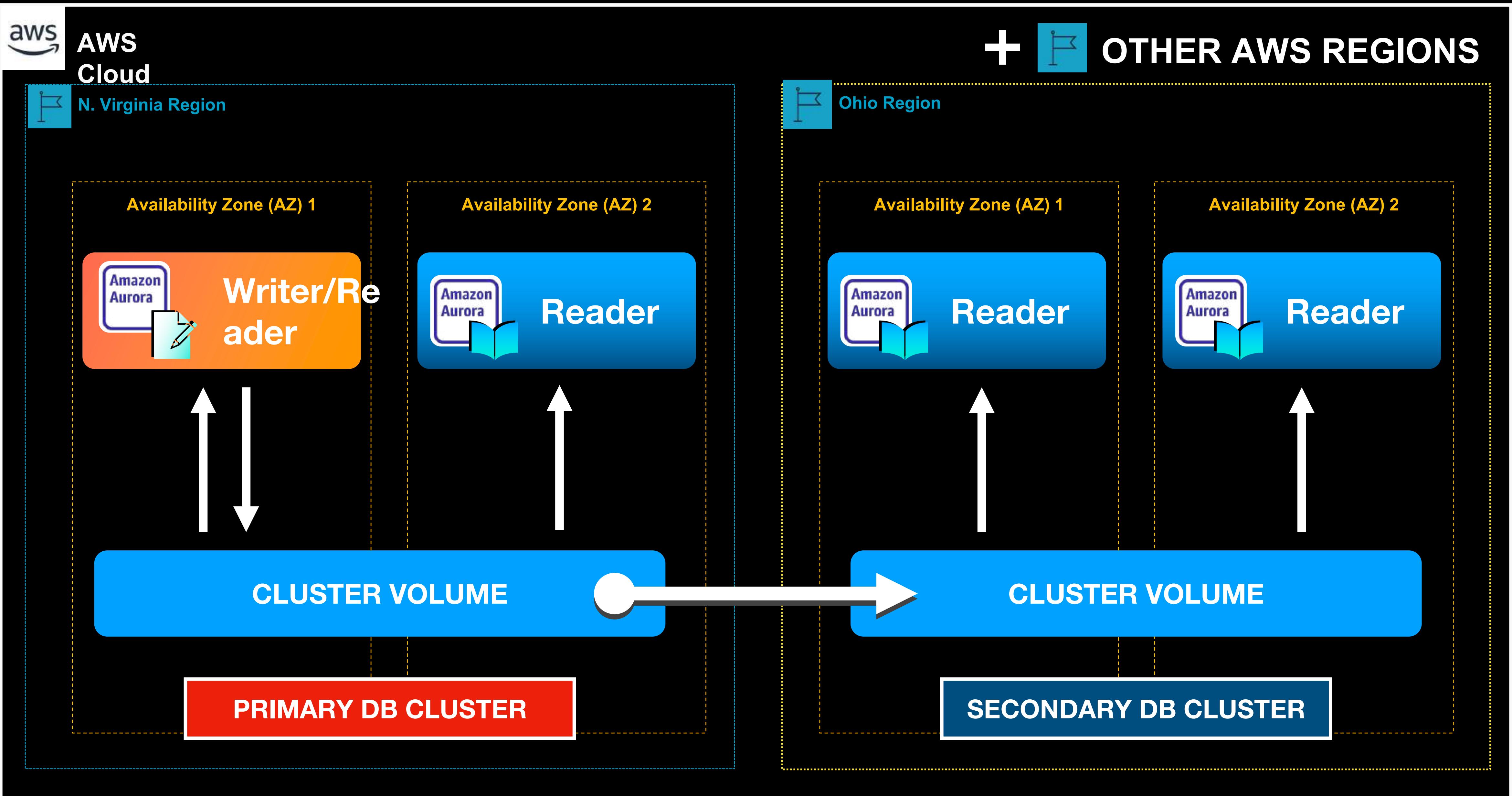
## Amazon Aurora **Serverless**

- For applications with **infrequent access patterns**
- **Automatically scales down your database capacity if there's less incoming traffic coming in, without any manual intervention**
- For migrating your on-premises database to AWS Cloud **without having to worry about its particular database instance type**
- If you need to **eliminate the need to manually modify your database instance type in anticipation of the changes in the number of your users or workloads**



## Amazon Aurora Global Databases

- **Designed for globally distributed applications**
- **Allows a single Amazon Aurora database to span multiple AWS Regions**
- **Offers faster physical replication between Aurora clusters**
- **Eliminates the need to manually create cross-region Aurora Replicas yourself**





**Amazon Aurora  
Global Databases**



**Recovery Point Objective**

**RPO**

= **1** second



**Recovery Time Objective**

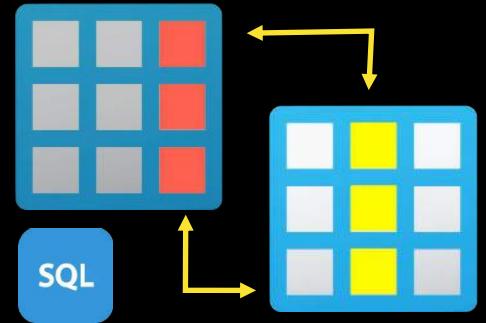
**RTO**

= **1** minute



# Amazon DynamoDB

## Overview



## Relational Database

- For applications with well-defined schema that does NOT change too often
- Has hundreds or thousands of tables
- Multiple table joins
- Tables having foreign keys
- Support complex SQL queries
- Tables having a relationship with other tables
- Has ACID properties
- Perfect for transactional workloads



A<sub>tomicity</sub>  
C<sub>onsistency</sub>  
I<sub>solation</sub>  
D<sub>urability</sub>



## NoSQL Database

- For applications that require a flexible schema that changes too often
- Does not have any related tables or table joins
- Usually has one table only
- Provides high throughput and performance for your global applications
- Can scale better than relational databases
- Can be used if you are unsure of the database schema that you will implement
- Suitable if you expect to make a lot of database changes as your website or application grows
- Does not have ACID properties by default



## Amazon DynamoDB

- A **fully managed NoSQL database**
- **Highly scalable storage and read/write capacity**
- Provides **single-digit millisecond performance**
- **Serverless**
- **Highly durable database**
- **Has built-in security, backup features as well as in-memory caching**

- Has the **least amount of operational overhead** than other types of databases



## Amazon **DynamoDB**

- Eliminates the **manual database management tasks, provisioning and scaling activities**
- Capable of automatically scaling its **read and write capacity** without the need for advanced capacity planning
- Can be queried using **simple key-value requests** via its APIs
- Can handle **millions of requests per second**

# Dynamo

## Dynamo: Amazon's Highly Available Key-value Store

Giuseppe DeCandia, Deniz Hastorun, Madan Jampani, Gunavardhan Kakulapati,  
Avinash Lakshman, Alex Pilchin, Swaminathan Sivasubramanian, Peter Vosshall  
and Werner Vogels

Amazon.com

### ABSTRACT

Reliability at massive scale is one of the biggest challenges we face at Amazon.com, one of the largest e-commerce operations in the world; even the slightest outage has significant financial consequences and impacts customer trust. The Amazon.com platform, which provides services for many web sites worldwide, is implemented on top of an infrastructure of tens of thousands of servers and network components located in many datacenters around the world. At this scale, small and large components fail continuously and the way persistent state is managed in the face of these failures drives the reliability and scalability of the software systems.

This paper presents the design and implementation of Dynamo, a highly available key-value storage system that some of Amazon's core services use to provide an "always-on" experience. To achieve this level of availability, Dynamo sacrifices consistency under certain failure scenarios. It makes extensive use of object versioning and application-assisted conflict resolution in a manner that provides a novel interface for developers to use.

### Categories and Subject Descriptors

D.4.2 [Operating Systems]: Storage Management; D.4.5 [Operating Systems]: Reliability; D.4.2 [Operating Systems]: Performance;

### General Terms

Algorithms, Management, Measurement, Performance, Design, Reliability.

### 1. INTRODUCTION

Amazon runs a world-wide e-commerce platform that serves tens of millions customers at peak times using tens of thousands of servers located in many data centers around the world. There are strict operational requirements on Amazon's platform in terms of performance, reliability and efficiency, and to support continuous growth the platform needs to be highly scalable. Reliability is one of the most important requirements because even the slightest outage has significant financial consequences and impacts customer trust. In addition, to support continuous growth, the platform needs to be highly scalable.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SOSP '07, October 14–17, 2007, Stevenson, Washington, USA.

Copyright 2007 ACM 978-1-59593-591-5/07/0010...\$5.00.

One of the lessons our organization has learned from operating Amazon's platform is that the reliability and scalability of a system is dependent on how its application state is managed. Amazon uses a highly decentralized, loosely coupled, service oriented architecture consisting of hundreds of services. In this environment there is a particular need for storage technologies that are always available. For example, customers should be able to view and add items to their shopping cart even if disks are failing, network routes are flapping, or data centers are being destroyed by tornados. Therefore, the service responsible for managing shopping carts requires that it can always write to and read from its data store, and that its data needs to be available across multiple data centers.

Dealing with failures in an infrastructure comprised of millions of components is our standard mode of operation; there are always a small but significant number of server and network components that are failing at any given time. As such Amazon's software systems need to be constructed in a manner that treats failure handling as the normal case without impacting availability or performance.

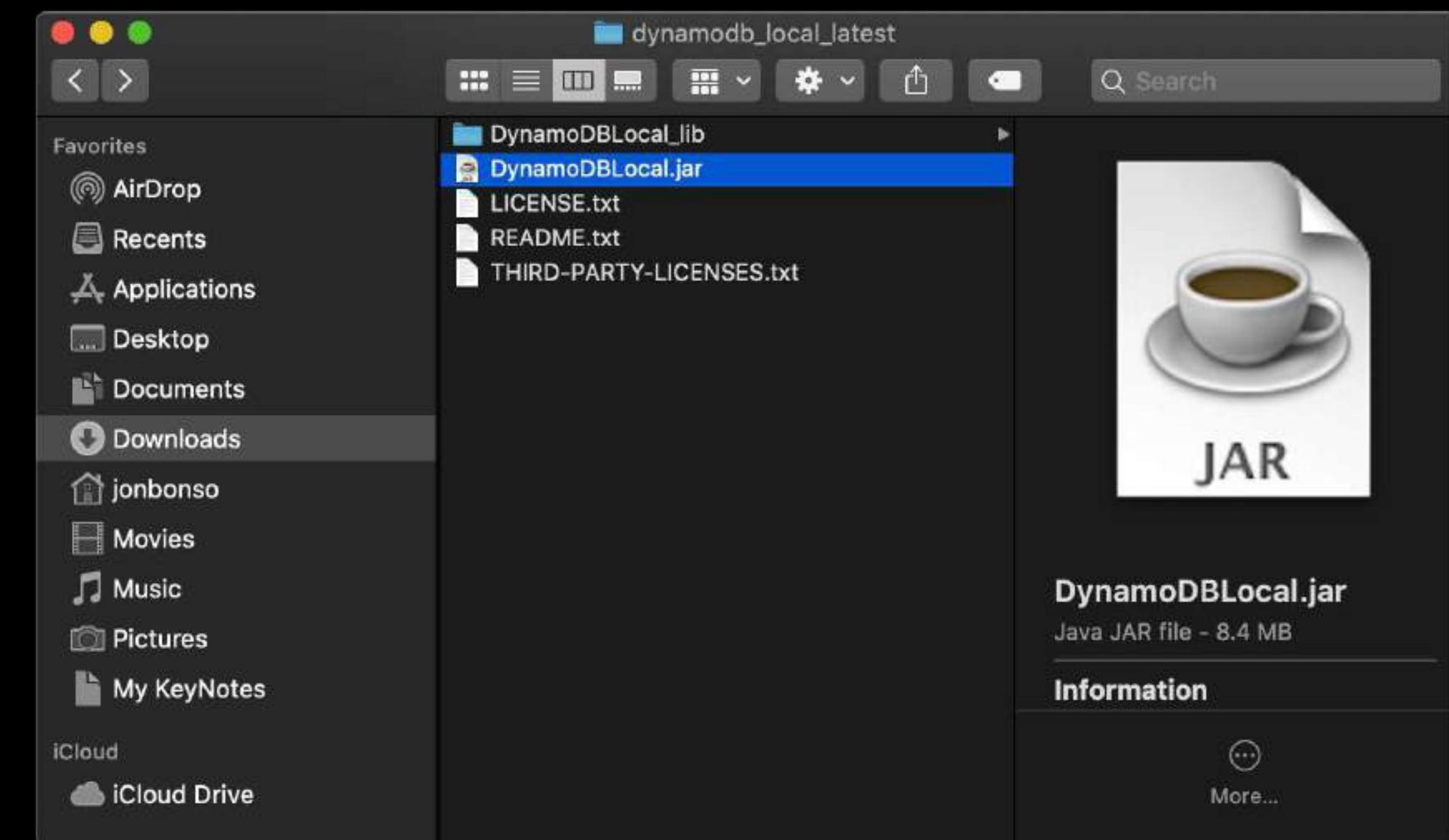
To meet the reliability and scaling needs, Amazon has developed a number of storage technologies, of which the Amazon Simple Storage Service (also available outside of Amazon and known as Amazon S3), is probably the best known. This paper presents the design and implementation of Dynamo, another highly available and scalable distributed data store built for Amazon's platform. Dynamo is used to manage the state of services that have very high reliability requirements and need tight control over the tradeoffs between availability, consistency, cost-effectiveness and performance. Amazon's platform has a very diverse set of applications with different storage requirements. A select set of applications require a storage technology that is flexible enough to let application designers configure their data store appropriately based on these tradeoffs to achieve high availability and guaranteed performance in the most cost effective manner.

There are many services on Amazon's platform that only need primary-key access to a data store. For many services, such as those that provide best seller lists, shopping carts, customer preferences, session management, sales rank, and product catalog, the common pattern of using a relational database would lead to inefficiencies and limit scale and availability. Dynamo provides a simple primary-key only interface to meet the requirements of these applications.

Dynamo uses a synthesis of well known techniques to achieve scalability and availability: Data is partitioned and replicated using consistent hashing [10], and consistency is facilitated by object versioning [12]. The consistency among replicas during updates is maintained by a quorum-like technique and a decentralized replica synchronization protocol. Dynamo employs



# Amazon **DynamoDB**



# Amazon **DynamoDB**



HIGHLY SCALABLE

ULTRA-FAST PERFORMANCE

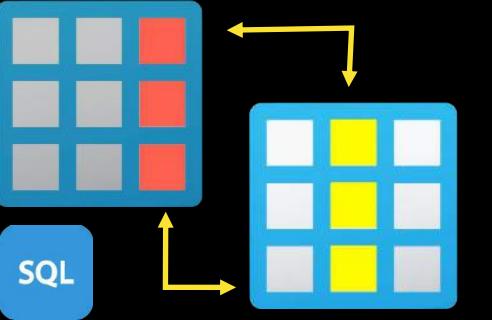


Response times in a matter of **milliseconds** or even in **microseconds!**

## DynamoDB Table



- All data is stored in a **single table only**
- Capable of accepting **millions of requests per second globally**
- Faster and more scalable than traditional relational databases
- Does not have a relationship with other DynamoDB tables



## Relational Database

**TABLE**

**ROW**

**COLUMN**

**PRIMARY KEY**

**INDEX**

**VIEW**

**NESTED TABLE/OBJECT**

**ARRAY**



## Amazon DynamoDB

**TABLE**

**ITEM**

**ATTRIBUTE**

**PRIMARY KEY / PARTITION KEY**

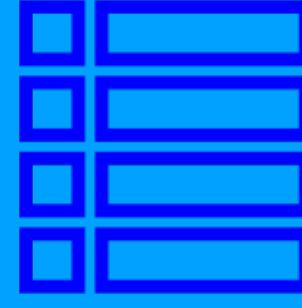
**SECONDARY INDEX**

**GLOBAL SECONDARY INDEX**

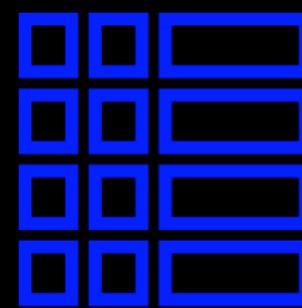
**MAP**

**LIST**

MAKES YOUR QUERIES  
RUN **FASTER!**



## LOCAL SECONDARY INDEX



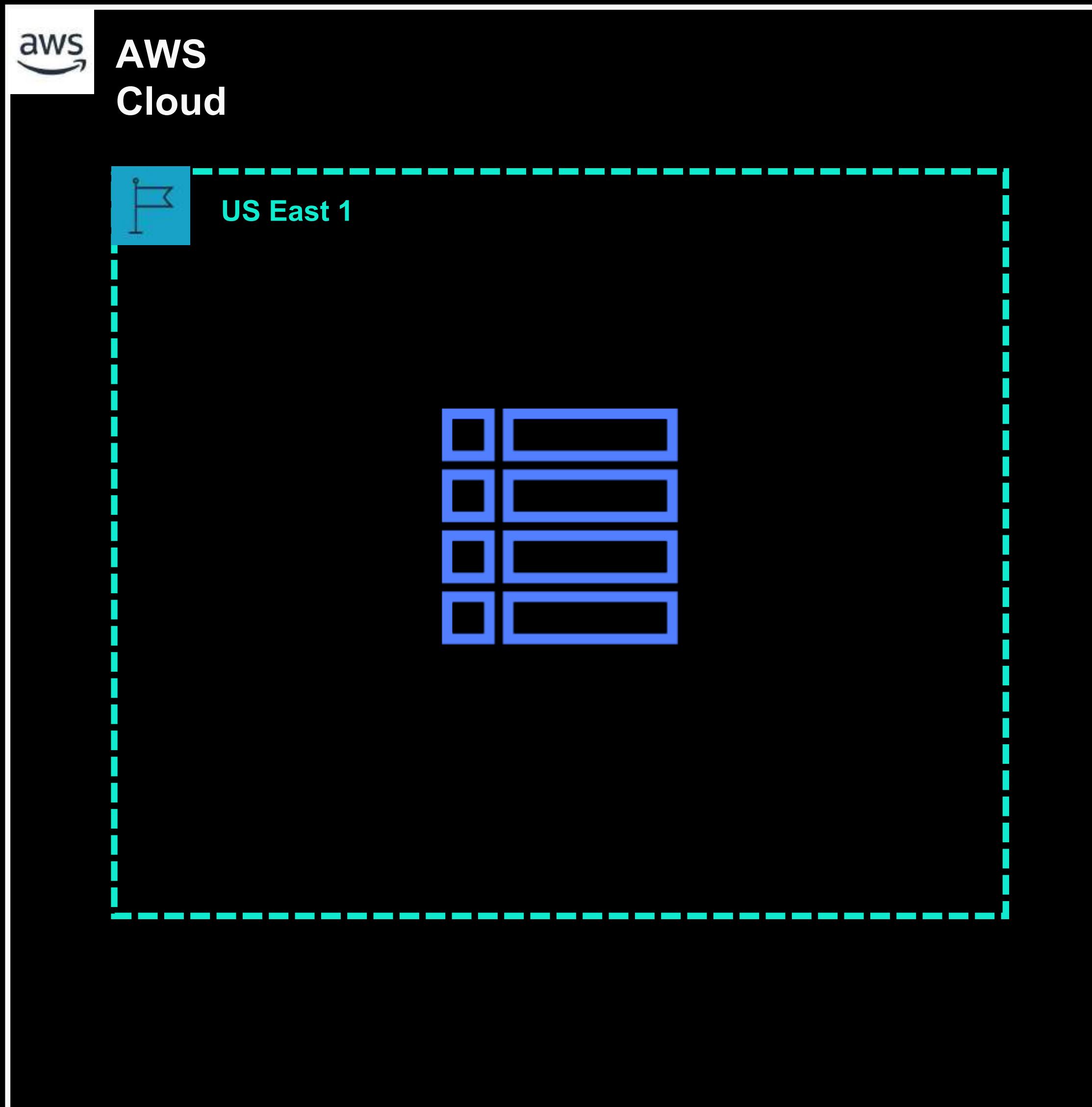
## GLOBAL SECONDARY INDEX

- Queries data over a **single partition only (localized)**
- Supports both **eventual consistency or strong consistency**
- Can only be added at the same time that you create the base table
- Queries data across **all partitions of the entire table**
- Only supports **eventual consistency**
- Can be added or deleted at any time

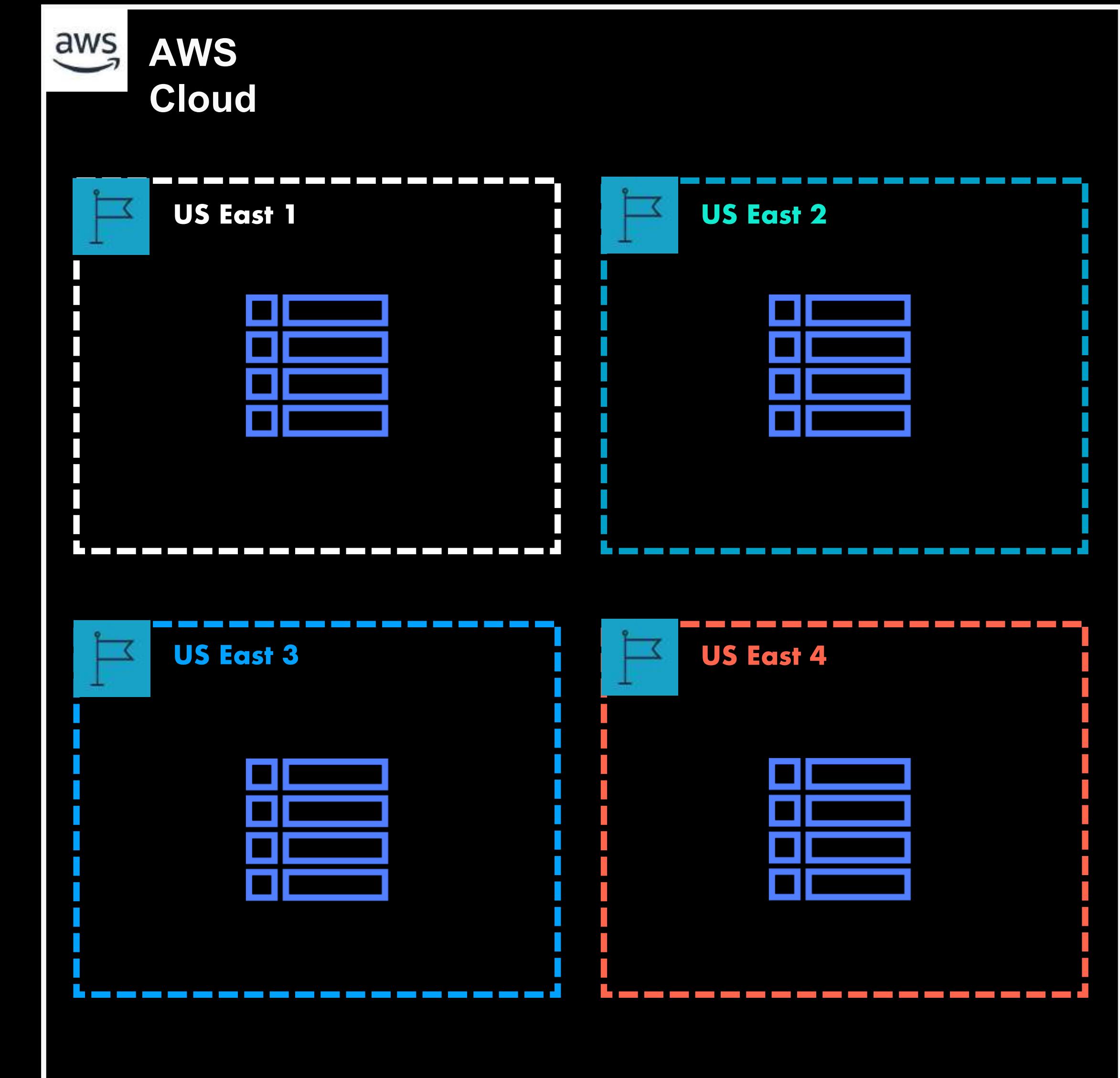


## Amazon DynamoDB Features

# Single DynamoDB Table



# DynamoDB Global Tables



## **Amazon DynamoDB Streams**

- A data stream that **captures each and every data change made to the items**
- If an item was added, modified, or deleted, then that item will be included in the **DynamoDB stream**
- Can be associated with **AWS Lambda**. The function can poll the stream and execute a set of actions whenever it detects new stream records
- Can also be integrated with **Kinesis Data Streams**
- Important component that needs to be enabled when using **Amazon DynamoDB Global Tables**

## Amazon DynamoDB TTL

- Automatically expire the items based on their timestamp and the TTL value that you specify
- TTL stands for Time to Live
- Allows you to define a timestamp per item
- Deletes the item from your table after the date and time of the specified timestamp
- Reduces the number of obsolete data in your table which can also lower down your costs

## Amazon DynamoDB Transactions

- Provides **ACID properties** to your DynamoDB table for your **transactional workloads**
- Provides an **all-or-nothing change** to multiple items both **within and across DynamoDB tables**
- Consists of **DynamoDB transactional read and write APIs**
  - `TransactWriteItems`
  - `TransactGetItems`
- Empowers you to **manage complex business workflows** that require adding, updating, or deleting multiple items as an **atomic operation**



## Amazon DynamoDB Accelerator (DAX)

- An **in-memory cache** for Amazon DynamoDB that is fully managed and highly available
- Launches a **DAX cluster** that can be run in your default or custom **Amazon VPC**
- Provides response time in **microseconds** and not just in milliseconds
- Delivers fast response times for accessing eventually consistent data
- Significantly reduces the response times of your DynamoDB database

# Amazon DynamoDB Scaling

- Measured in terms of:
  - Read Capacity Unit or RCU
  - Write Capacity Unit or WCU

## Provisioned Capacity Mode

- Suitable if your application has **predictable traffic** that doesn't vary over time
- Allows you to **manually set** or provision the RCU and WCU of your DynamoDB table
- Has an **Auto Scaling feature** that you can configure
- Can set the target utilization, minimum provisioned capacity, and maximum provisioned capacity values in the Auto Scaling settings
- At risk of over-provisioning and having unnecessary costs when the incoming traffic is way lower than expected

## On-Demand Capacity Mode

- For applications with **inconsistent traffic** or has varying access patterns
- Suitable if you expect that there'll be more **traffic with sharp spikes in the future**
- **No manual Auto Scaling setting** that you can configure. The RCU & WCU are automatically scaled without any intervention
- Can be used if your application has a combination of predictable and **variable traffic**
- Suitable if you have clearly defined access patterns throughout the year but with variable amounts of traffic on certain days only

# Amazon DynamoDB Scaling

# Amazon DynamoDB Security

- Protects your data both in transit and at rest
- All data stored in Amazon DynamoDB is fully encrypted at rest by default
- The API calls from your private Amazon EC2 instances that go to DynamoDB can be configured to not traverse the public Internet by creating a VPC Gateway Endpoint and adding a new route table entry

# Amazon DynamoDB Identity & Access Management

```
{  
    "Id": "TutorialsDojoPhilippineBooksPolicy1",  
    "Version": "2012-10-17",  
  
    "Statement": [  
        {  
            "Sid": "AllowAccessToBooksTable",  
            "Effect": "Allow",  
  
            "Action": [  
                "dynamodb:Get*",  
                "dynamodb:Query",  
                "dynamodb:Scan",  
                "dynamodb:BatchWrite*",  
                "dynamodb CreateTable",  
                "dynamodb>Delete*",  
                "dynamodb:Update*",  
                "dynamodb:PutItem" ],  
  
            "Resource": "arn:aws:dynamodb:us-west-2:12345:table/Books"  
        }  
    ]  
}
```



# Amazon DynamoDB Backups

Point-in-time Recovery  
(PITR)

- Automated backup process
- Enables continuous backups to your table
- Allows you to restore your table at a point in time that you specify
- Entails additional costs

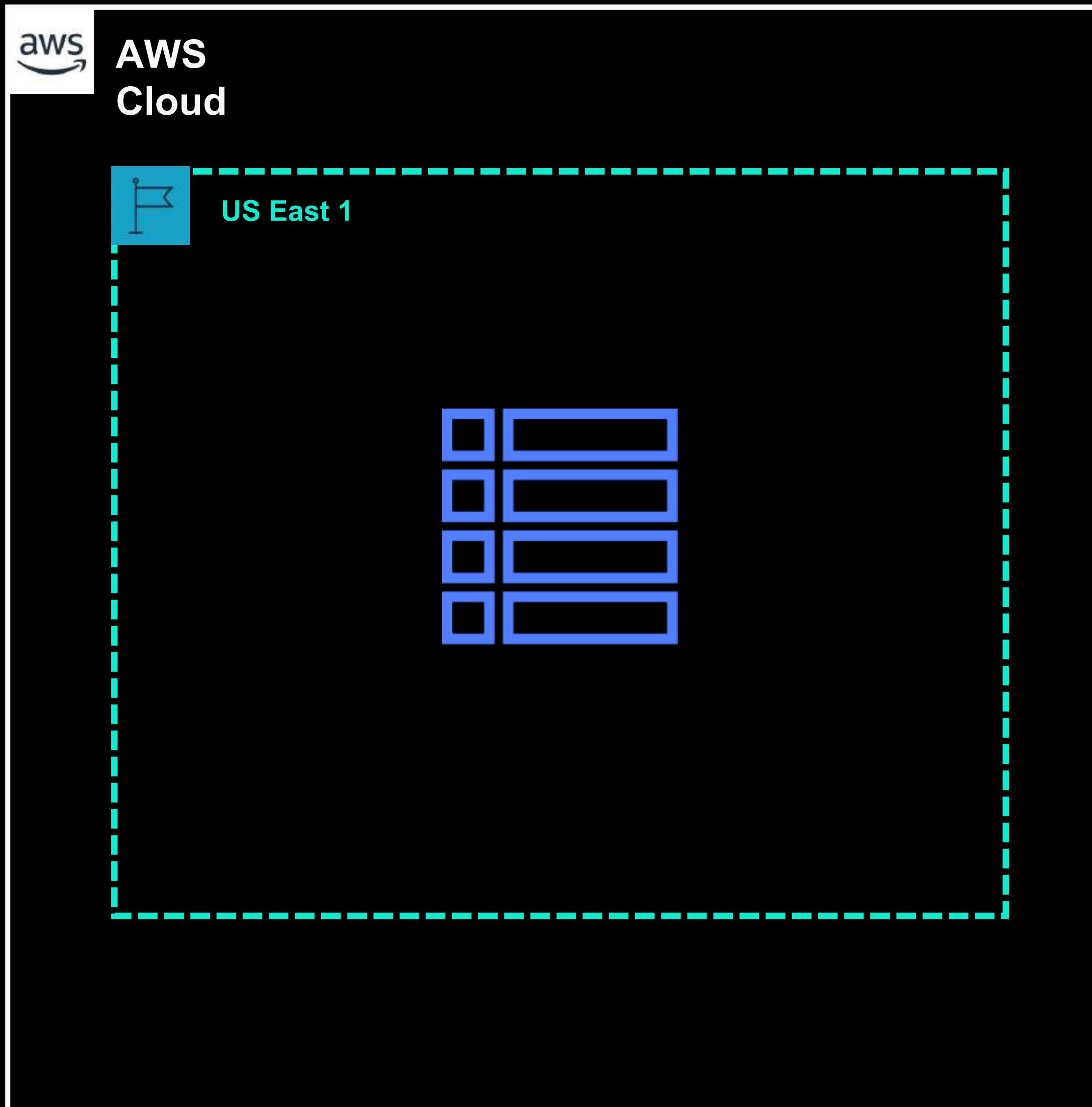
On-Demand Backup  
and Restore

- Manual backup process
- No continuous backups
- Can only restore to a particular backup that you've taken
- A cost-effective yet limited backup option feature for your data

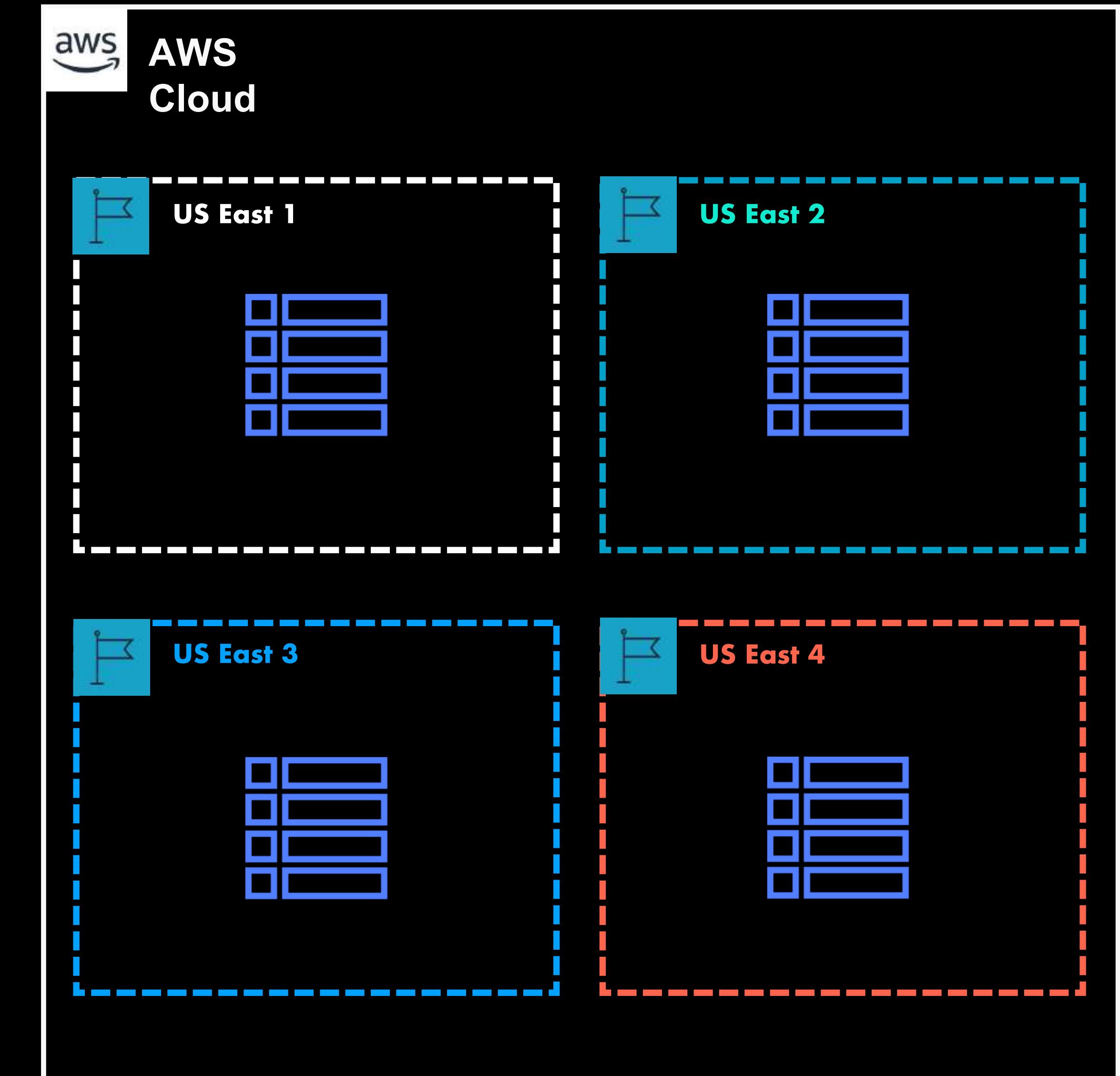


# Amazon DynamoDB Core Components

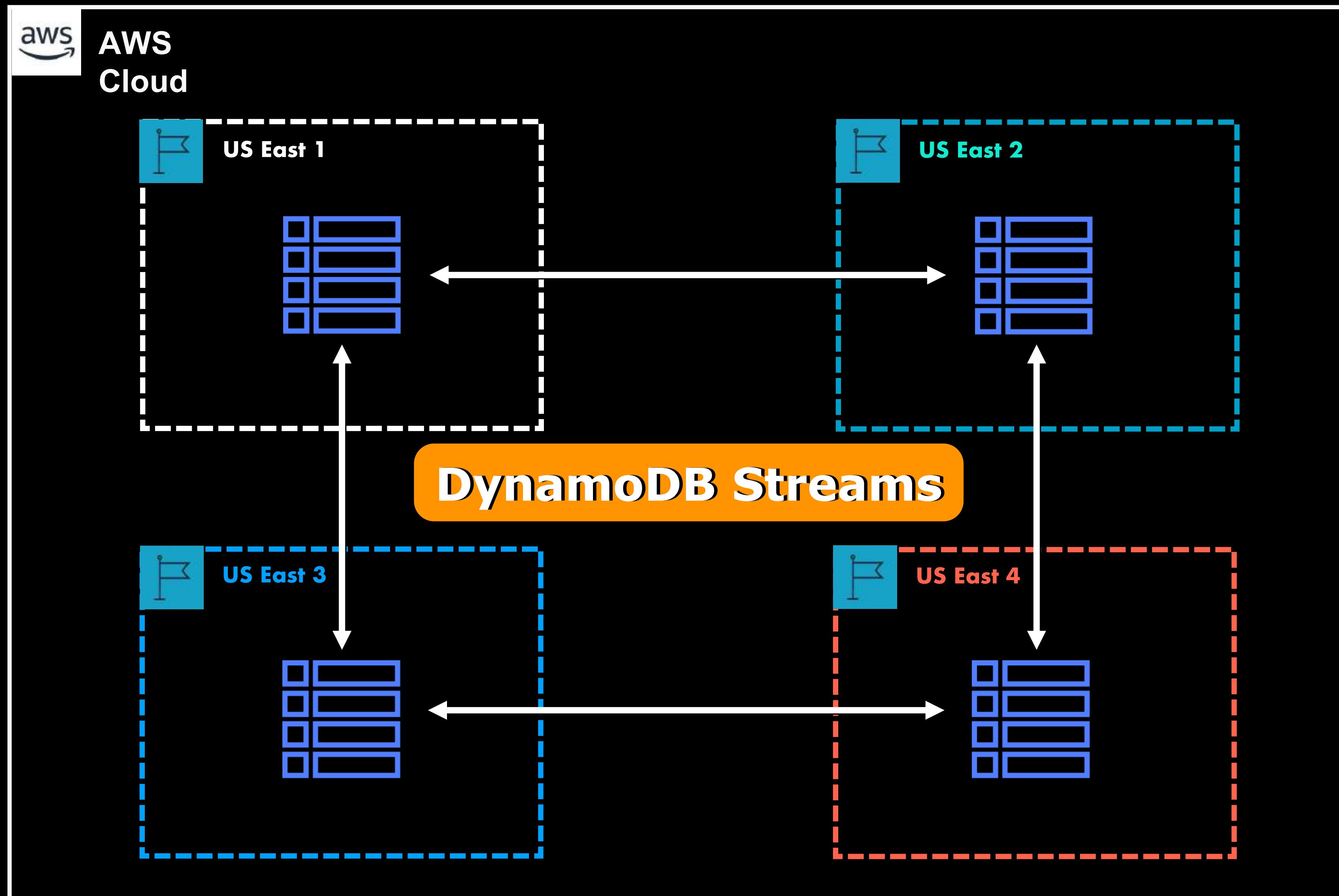
# Single DynamoDB Table

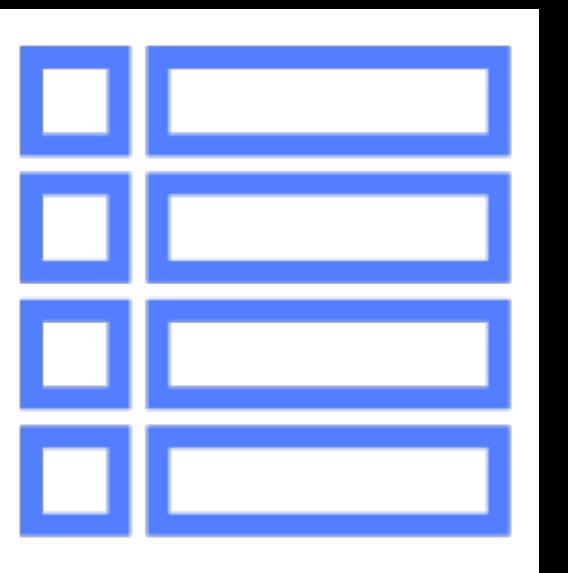


# DynamoDB Global Tables

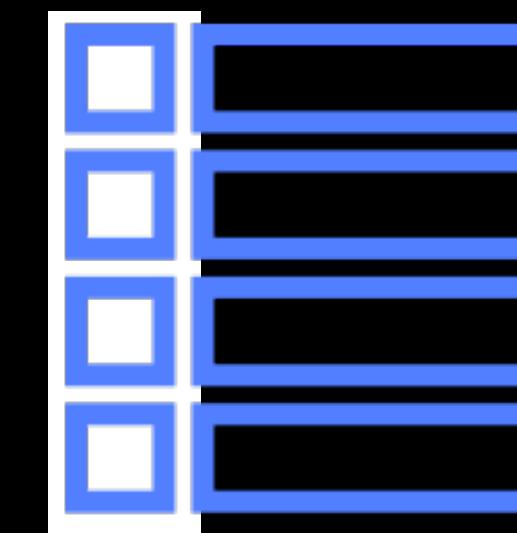


# DynamoDB Global Tables

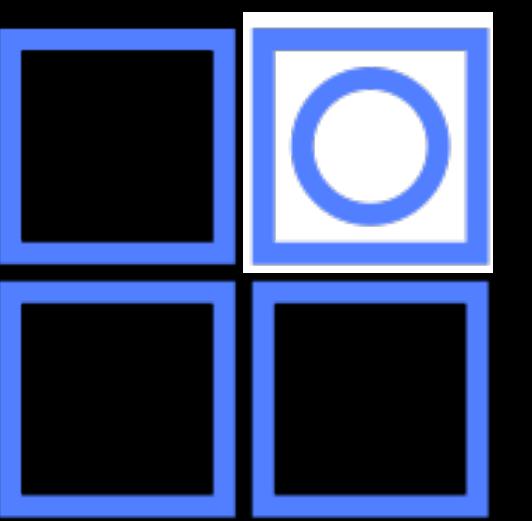




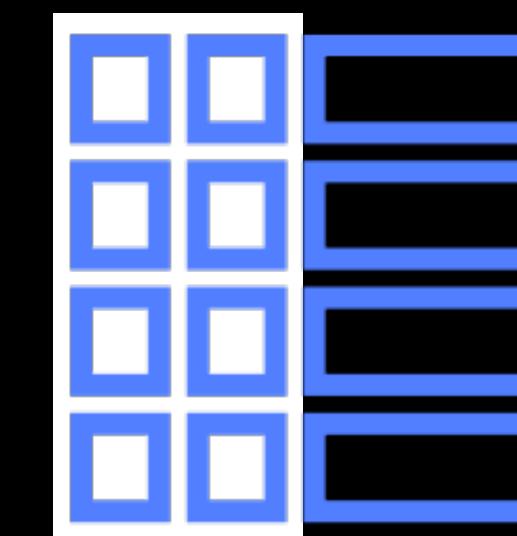
**TABLE**



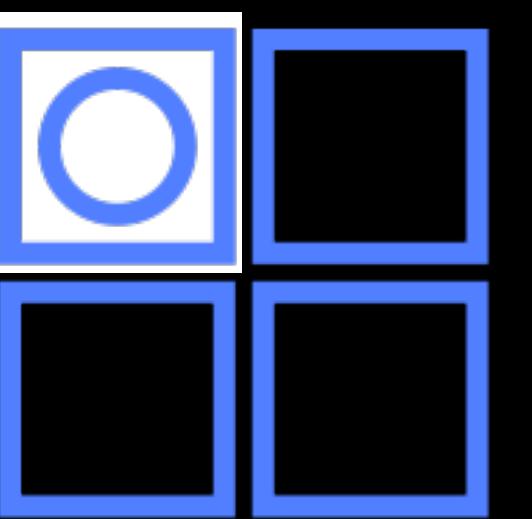
**PRIMARY KEY**



**ITEM**



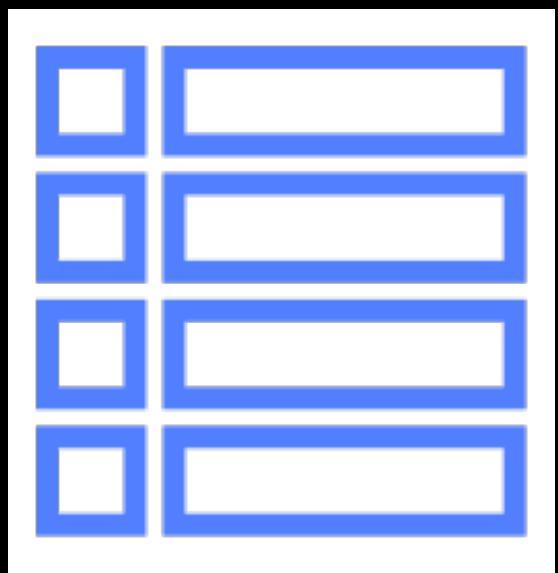
**SECONDARY INDEX**



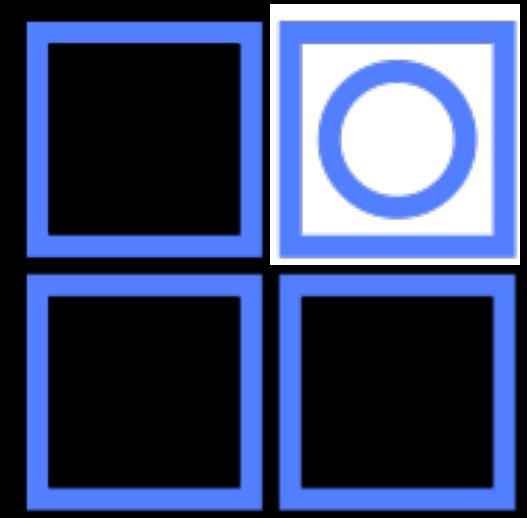
**ATTRIBUTE**

**AND OTHER COMPONENTS...**

# TABLE

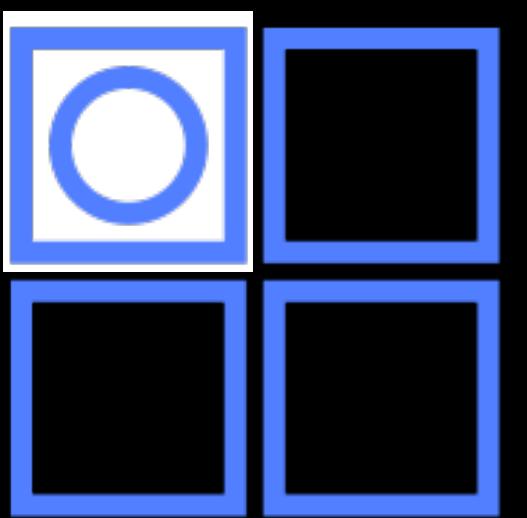


- Similar to the table of other database systems
- A **collection of related data** that can represent an object, an idea, a role, or an abstract concept
- In DynamoDB, the entire NoSQL database is within a **single DynamoDB table** only



## ITEM

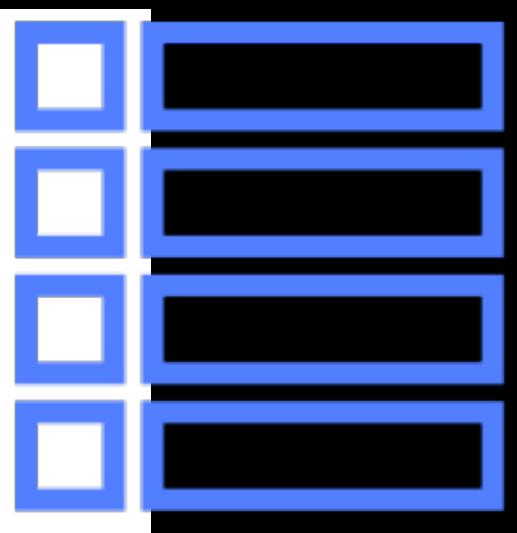
- Each table contains zero or more items
- Similar to the rows, records, or tuples in other database systems
- The “**Row**” of the DynamoDB Table
- Can have a nested attribute, which contains another item or another nested attribute
- Can be automatically expired based on its timestamp using TTL, or Time to Live



## ATTRIBUTE

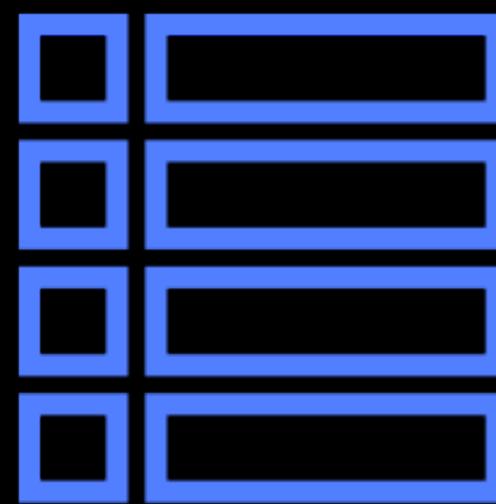
- Each item contains zero or more attributes
- Similar to the fields or columns in other data stores
- The “**Column**” of the DynamoDB Table

## PRIMARY KEY



- Also known as the **partition key**
- Acts as the primary index that **uniquely identifies each item in your DynamoDB table**
- Provides the ability to search for a particular item in your table
- Used as an input to the **internal hash function** in DynamoDB. The output from that function determines the physical internal storage in which the item will be stored
- The primary key attribute must be a **scalar**

## **PRIMARY KEY**



**Simple**

**PARTITION KEY**

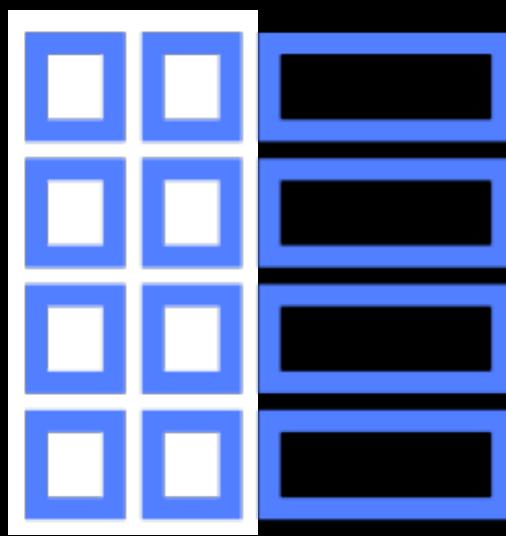
**Composite**

**PARTITION KEY**

**+**

**SORT KEY**

## SECONDARY INDEX



- Makes your queries run **faster!**
- Provides more flexibility and performance improvement to your queries
- Supports your advanced queries to access your stored data faster
- Allows you to query the data in the table using an **alternate key** other than the primary key

## MUSIC TABLE

PARTITION KEY: SongId

```
{  
    "SongId": 1,  
    "Artist" : "Jon Bonso",  
    "SongTitle" : "Brand New Memories",  
    "Genre": "Rock",  
    "Year" : 2009  
}
```

```
{  
    "SongId": 2,  
    "Artist" : "Ariel Rivera",  
    "SongTitle" : "Sana Kahit Minsan",  
    "Genre": "R&B",  
    "Year" : 1991  
}
```

```
{  
    "SongId": 3,  
    "Artist" : "Rey Valera",  
    "SongTitle" : "Kung Kailangan Mo Ako",  
    "Genre": "Jazz",  
    "Year" : 1980  
}
```

```
{  
    "SongId": 4,  
    "Artist" : "Gino Padilla",  
    "SongTitle" : "Closer You and I",  
    "Genre": "R&B",  
    "Year" : 2000  
}
```

## SECONDARY INDEX LOGICAL TABLE

•  
•  
•  
•  
{  
 "SongId": 1,  
 "Artist" : "Jon Bonso",  
 "SongTitle" : "Brand New Memories",  
 "Genre": "Rock"  
}

{  
 "SongId": 2,  
 "Artist" : "Ariel Rivera",  
 "SongTitle" : "Sana Kahit Minsan",  
 "Genre": "R&B"  
}

{  
 "SongId": 4,  
 "Artist" : "Gino Padilla",  
 "SongTitle" : "Closer You and I",  
 "Genre": "R&B"  
}

•

PARTITION KEY: Artist  
SORT KEY: Genre

## MUSIC TABLE

PARTITION KEY: SongId

SORT KEY: Artist

```
{  
    "SongId": 1,  
    "Artist" : "Jon Bonso",  
    "SongTitle" : "Brand New Memories",  
    "Genre": "Rock",  
    "Year" : 2009  
}
```

```
{  
    "SongId": 2,  
    "Artist" : "Ariel Rivera",  
    "SongTitle" : "Sana Kahit Minsan",  
    "Genre": "R&B",  
    "Year" : 1991  
}
```

```
{  
    "SongId": 3,  
    "Artist" : "Rey Valera",  
    "SongTitle" : "Kung Kailangan Mo Ako",  
    "Genre": "Jazz",  
    "Year" : 1980  
}
```

```
{  
    "SongId": 4,  
    "Artist" : "Gino Padilla",  
    "SongTitle" : "Closer You and I",  
    "Genre": "R&B",  
    "Year" : 2000  
}
```

## GLOBAL SECONDARY INDEX

PARTITION KEY: Artist

SORT KEY: Genre

```
{  
    "SongId": 1,  
    "Artist" : "Jon Bonso",  
    "SongTitle" : "Brand New Memories",  
    "Genre": "Rock"  
}
```

```
{  
    "SongId": 2,  
    "Artist" : "Ariel Rivera",  
    "SongTitle" : "Sana Kahit Minsan",  
    "Genre": "R&B"  
}
```

## LOCAL SECONDARY INDEX

PARTITION KEY: SongId

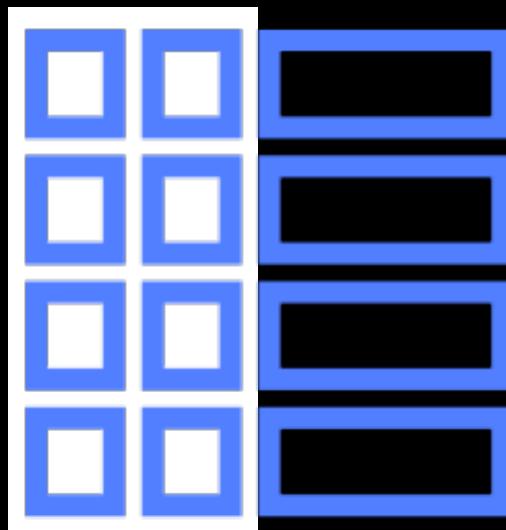
SORT KEY: Genre

```
{  
    "SongId": 3,  
    "Artist" : "Rey Valera",  
    "SongTitle" : "Kung Kailangan Mo Ako",  
    "Genre": "Jazz"  
}
```

```
{  
    "SongId": 4,  
    "Artist" : "Gino Padilla",  
    "SongTitle" : "Closer You and I",  
    "Genre": "R&B"  
}
```

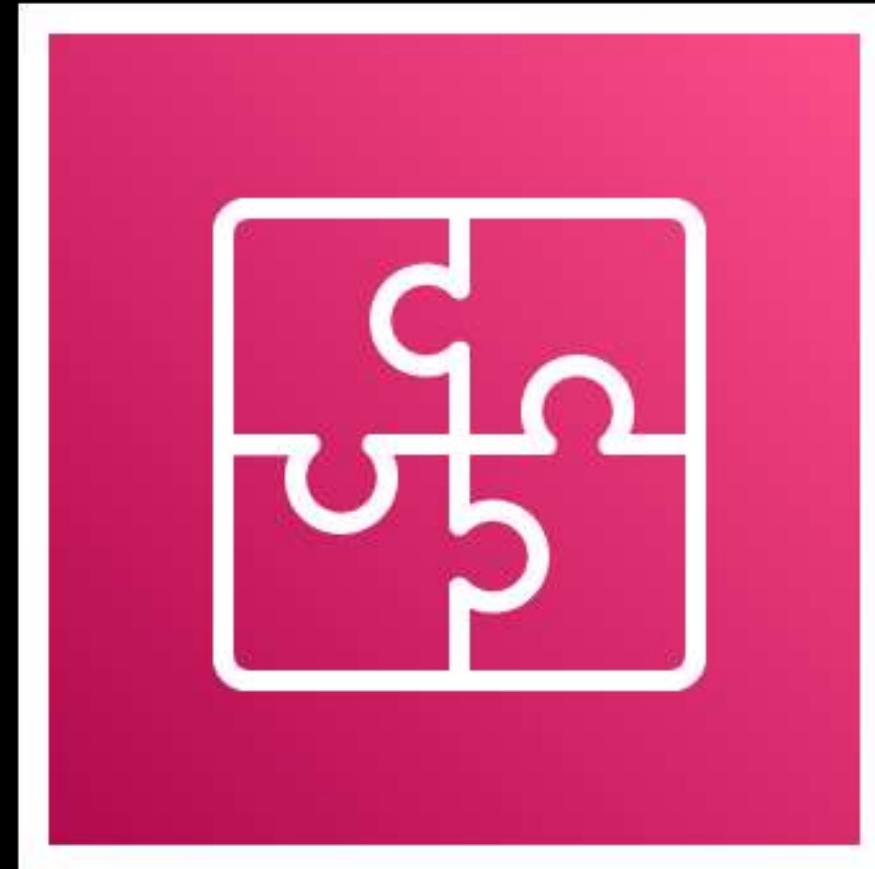


## SECONDARY INDEX



- Similar to the **INDEX** of MySQL, Oracle, SQL Server, and other relational databases
- Primarily used to make your queries **FASTER!**

```
1
2
3 * CREATE INDEX TUTORIALDOJO_INDEX
4 ON COURSES (ID, COURSE_NAME);
5
6
```



# **Application Integration**

## Overview

---

# Application Integration



# **Application Integration**

*Empowers the migration from*

**Monolithic Architecture**



**Distributed Architecture**

**Monolithic Architecture**

**Distributed Architecture**

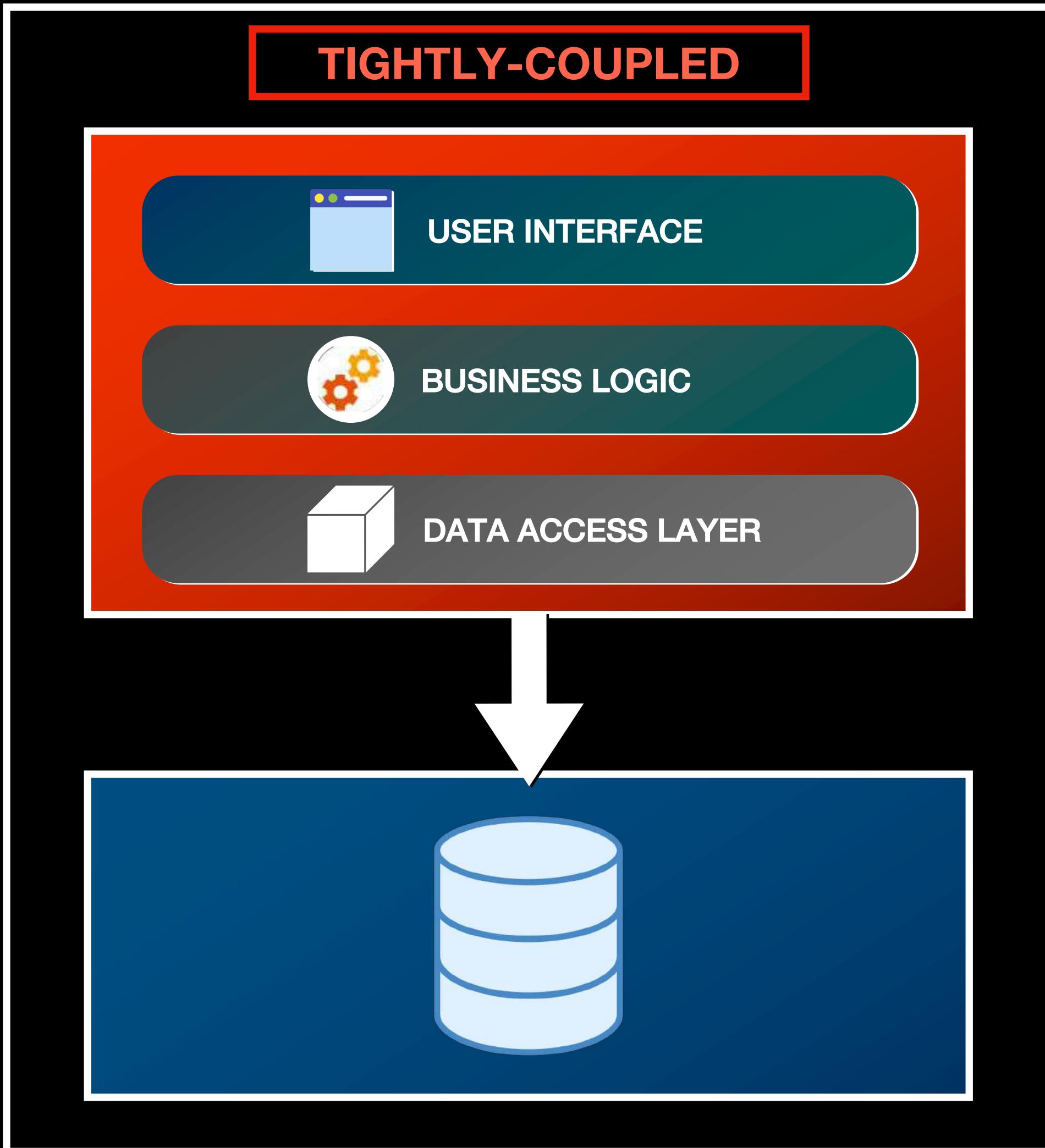
**1**

**MONO**

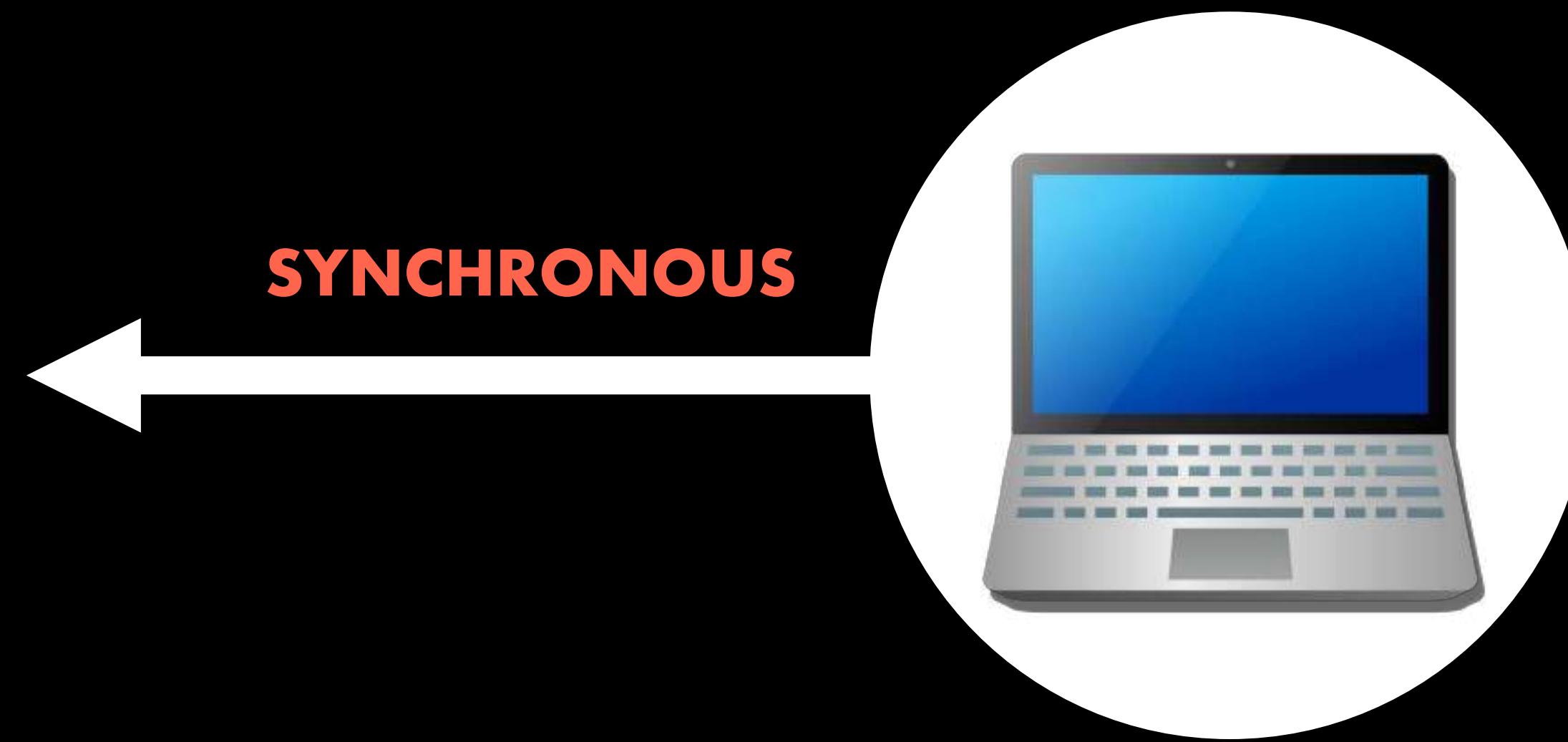


**LITH**

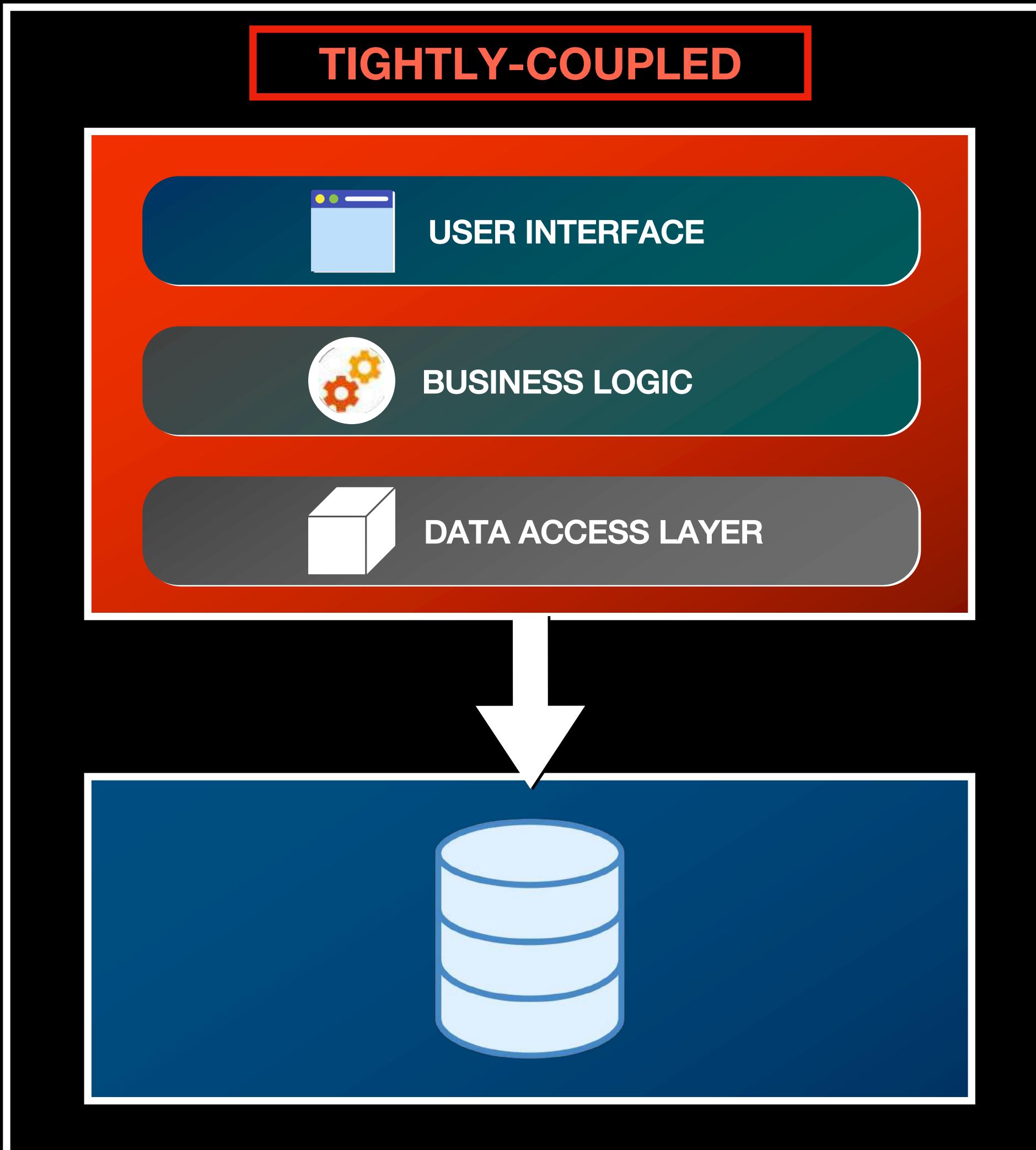
## Monolithic Architecture



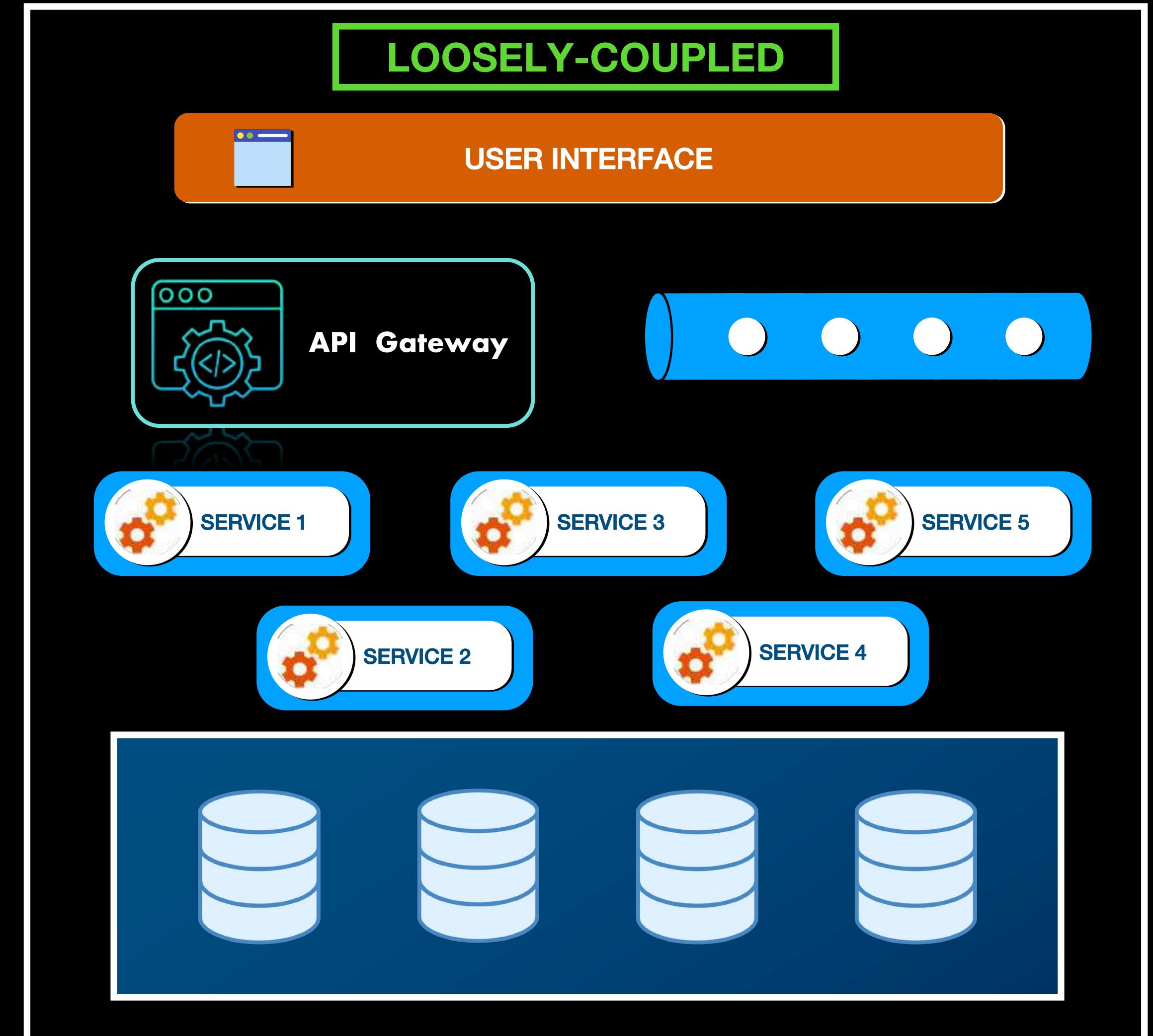
## Distributed Architecture



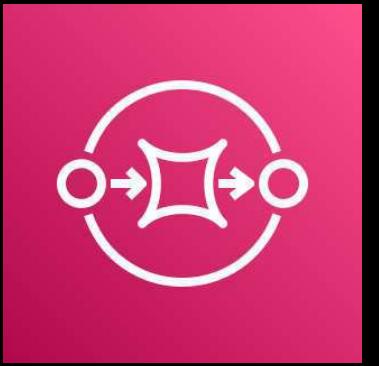
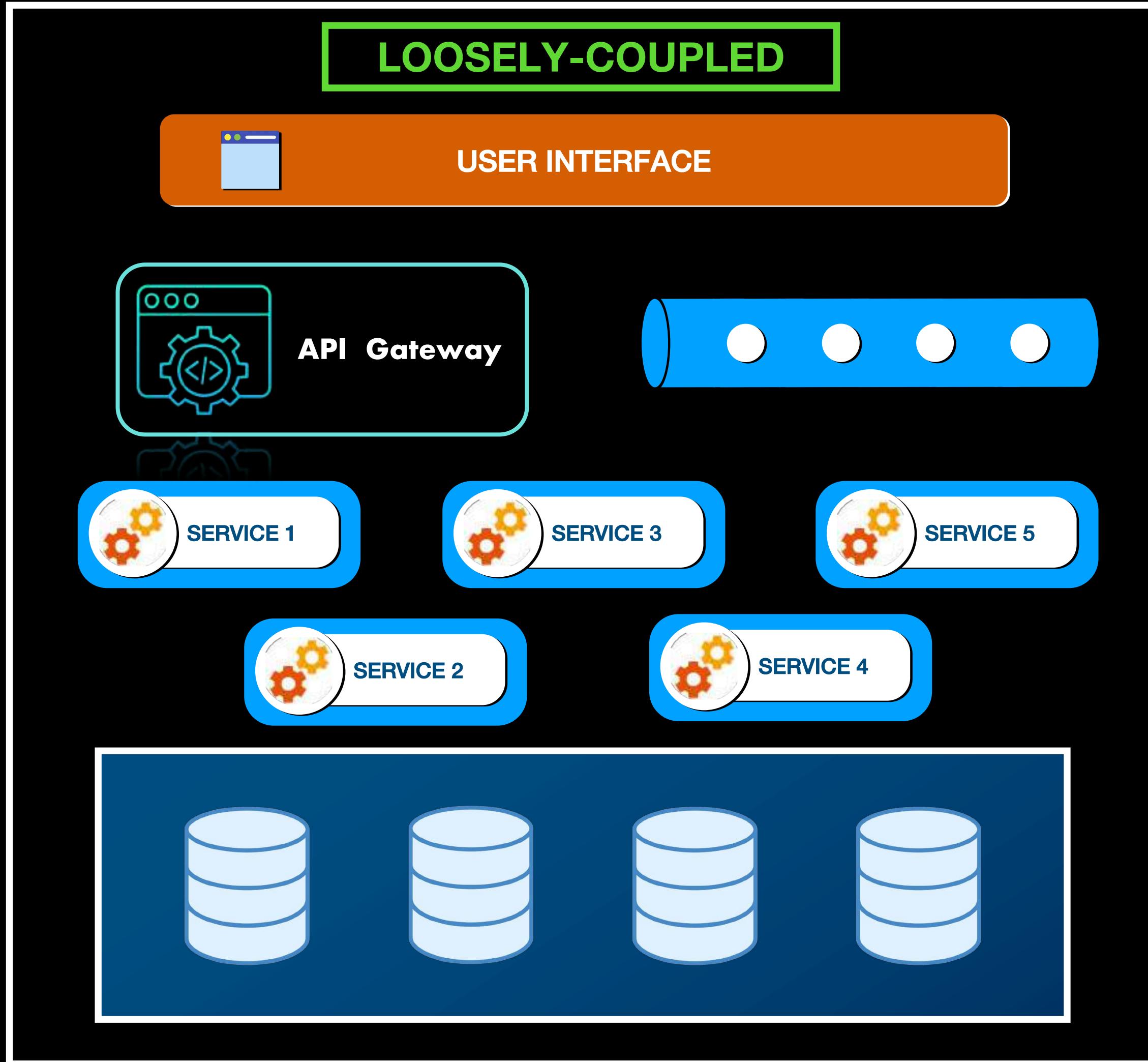
## Monolithic Architecture



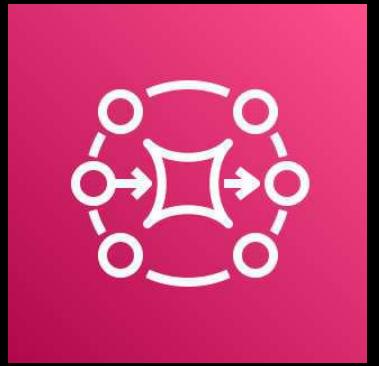
## Distributed Architecture



# Distributed Architecture



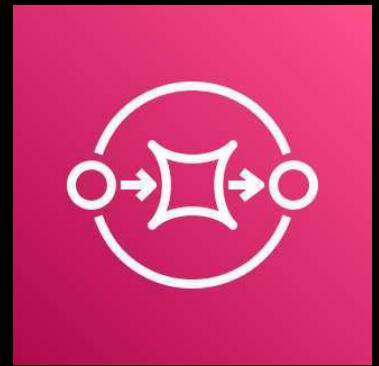
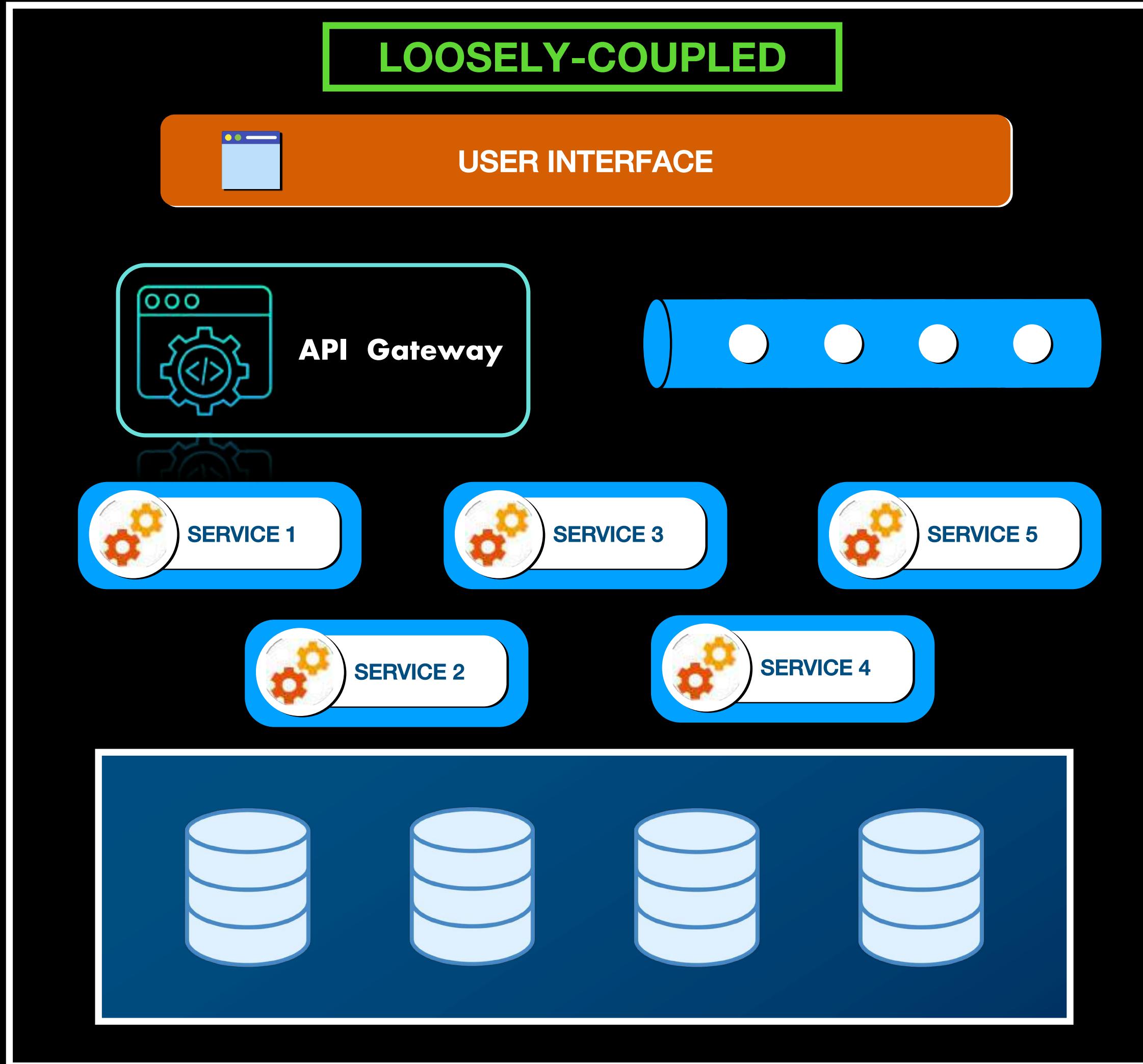
Amazon SQS



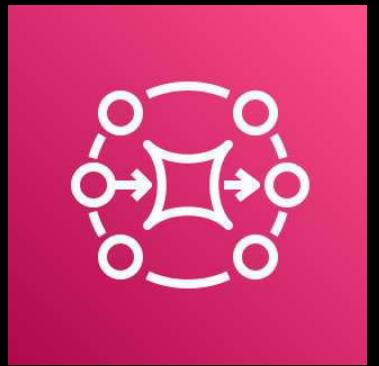
Amazon MQ



# Distributed Architecture



**Amazon SQS**



**Amazon MQ**



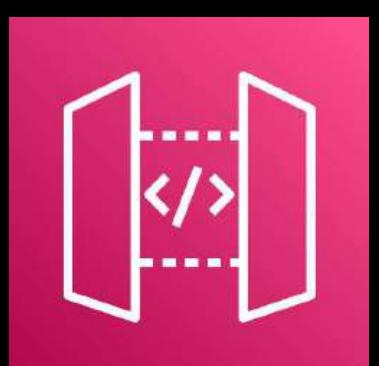
**Amazon SNS**



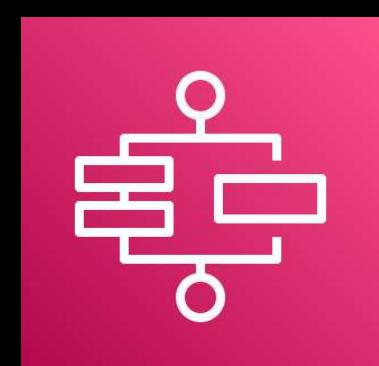
**Amazon EventBridge**



**AWS AppSync**

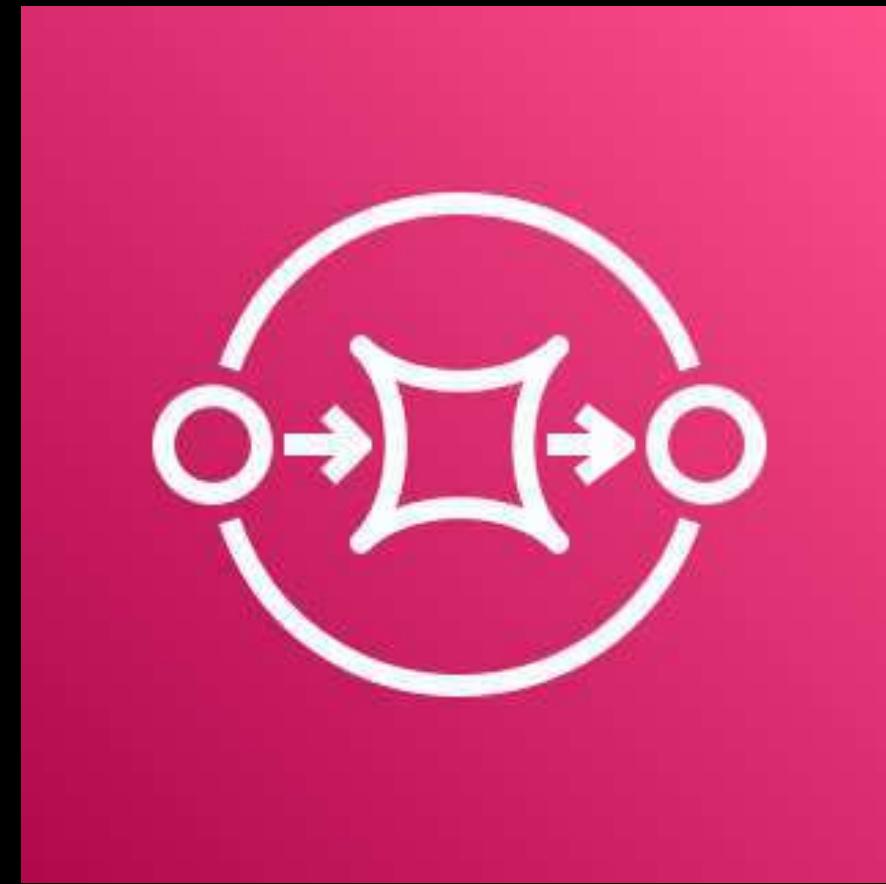


**Amazon API Gateway**



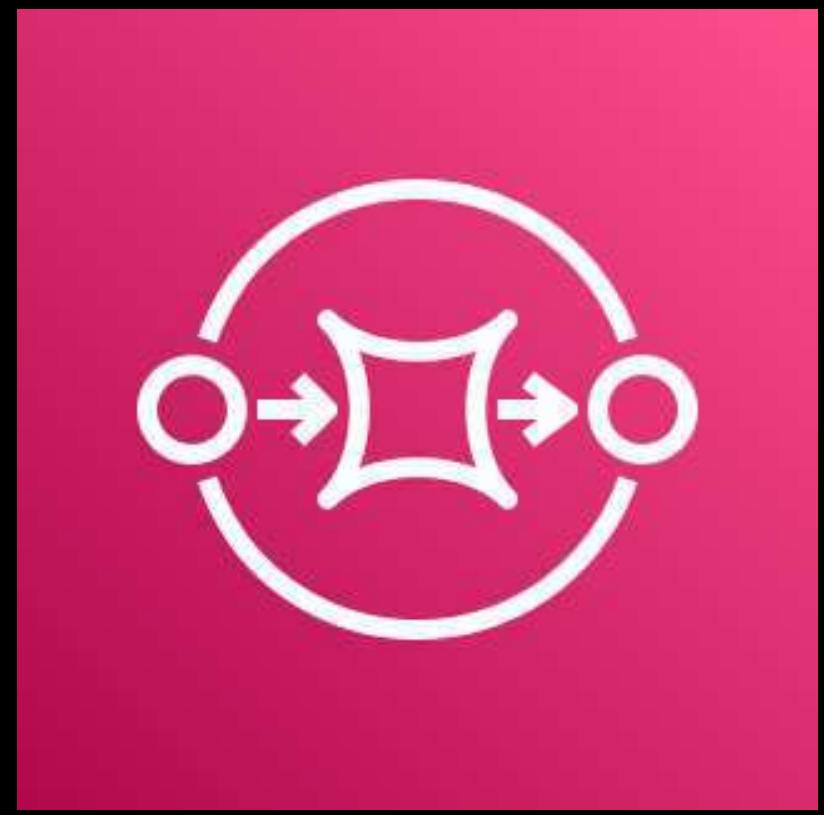
**AWS Step Functions**





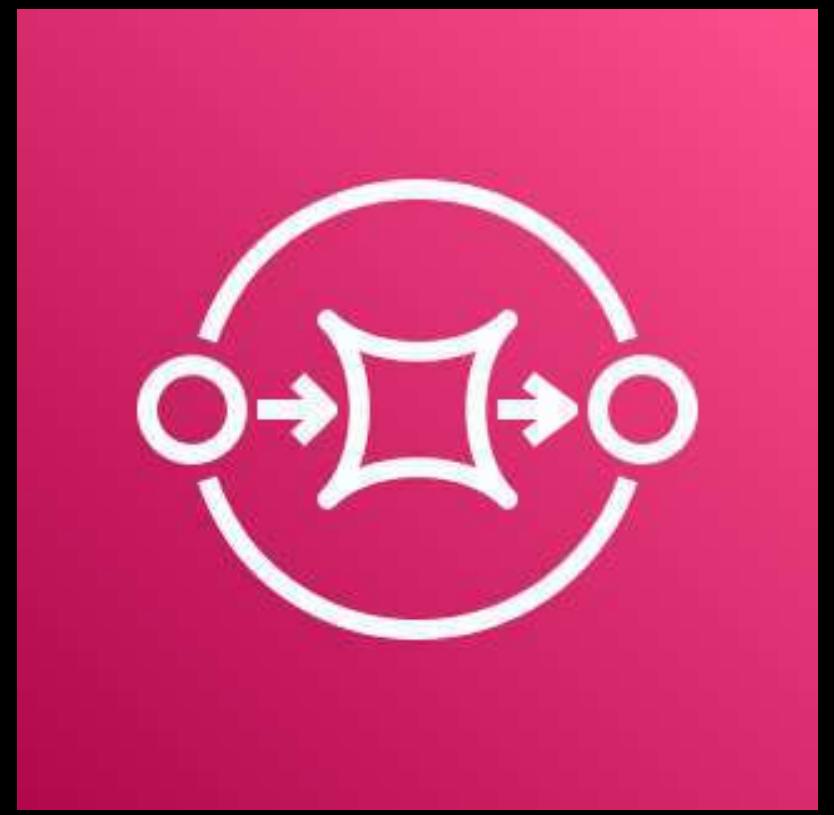
# Amazon SQS Overview

---



## Amazon SQS

- Decouple tightly-coupled architecture
- Process workloads asynchronously



=

**MESSAGE  
QUEUE**

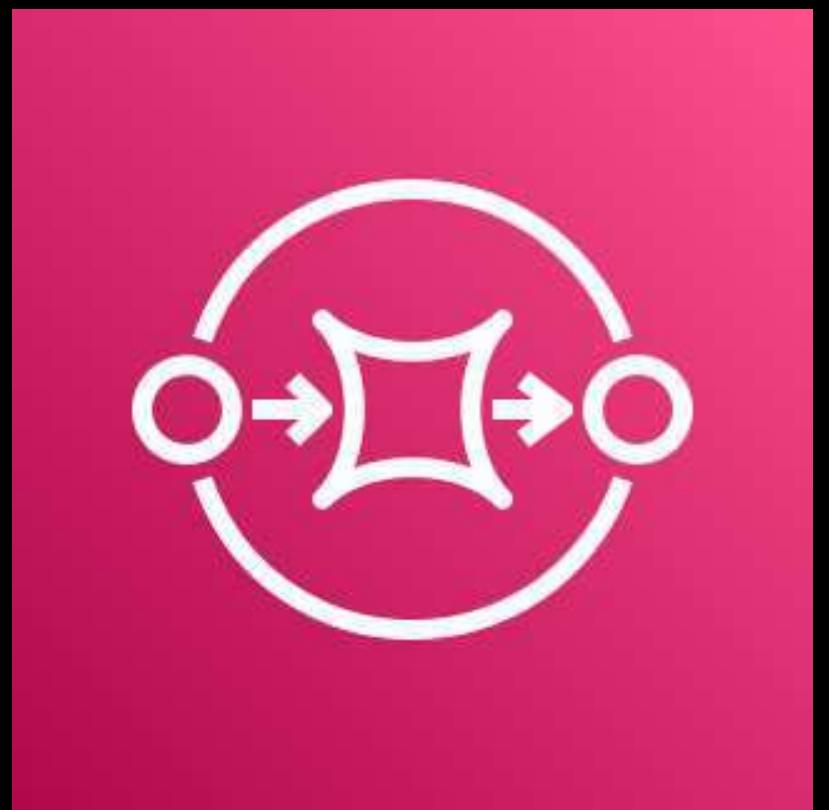
**Amazon SQS**

# QUEUE

- The order of processing is **First-In, First-Out (FIFO)**
- Items are stored **sequentially**
- The processing is done by a **Consumer**

# MESSAGE QUEUE

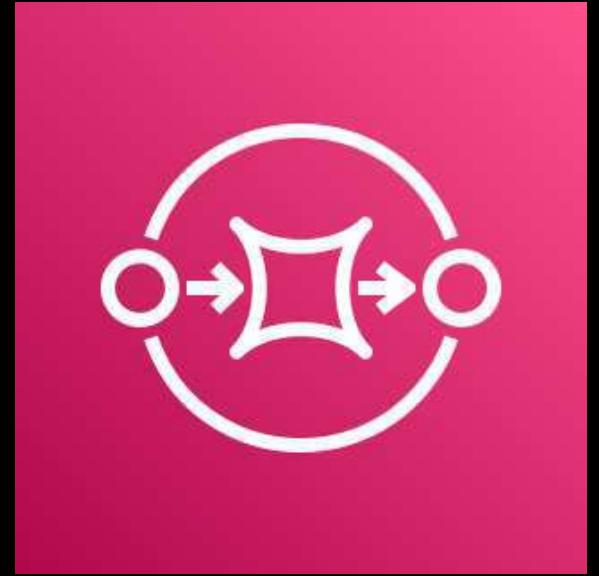
- Handles the incoming messages of your application
- Sends the items to the consumers for processing
- Asynchronous service-to-service communication
- Messages can be HTTP or an API request
- For workloads that take several minutes to complete
- Fetching messages for processing is called Polling



## Amazon SQS

- Fully-managed message queue
- For workloads with long-running requests
- Assists in scaling your compute resources
- Can be integrated with other AWS services

## Amazon SQS TYPES



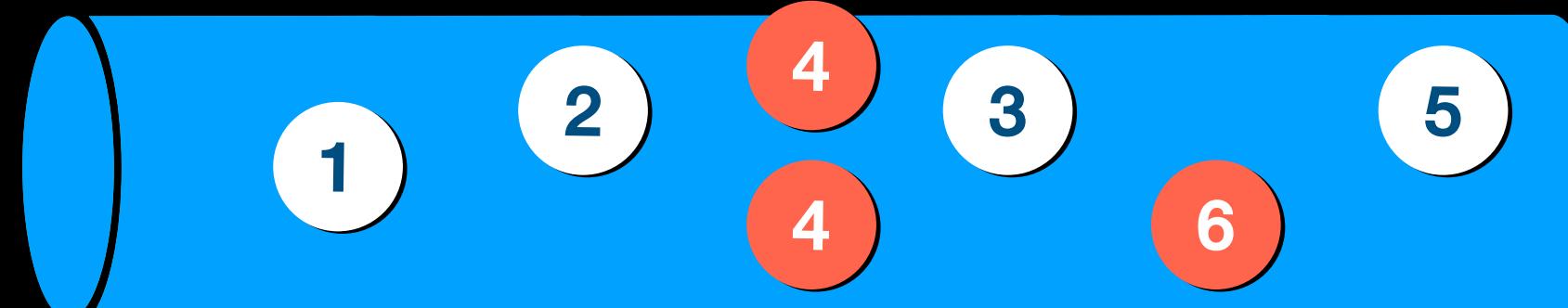
DELIVERY

ORDERING

THROUGHPUT

## STANDARD

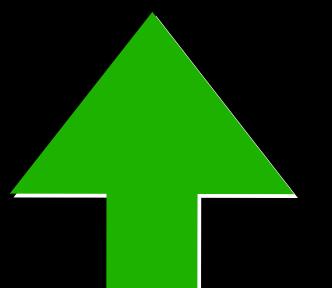
ChangeMessageVisibility API



At Least Once

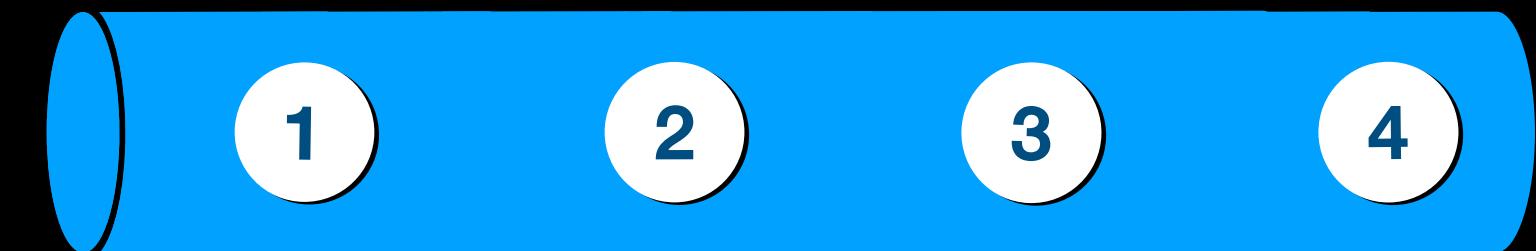
Best Effort  
*Messages might be delivered in a different order*

HIGH



## FIFO

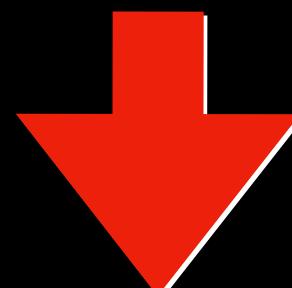
Deduplication

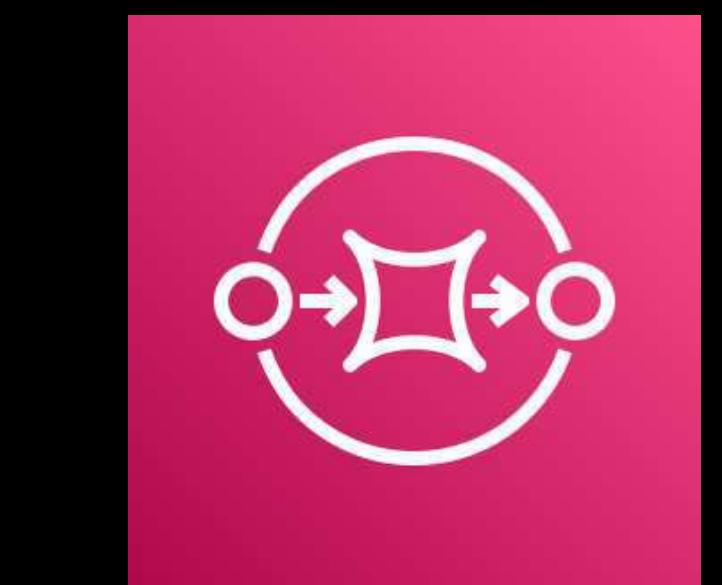


Exactly Once

Preserves the exact order  
in which the messages are received

LIMITED





## Amazon SQS SETTINGS

VISIBILITY TIMEOUT

MESSAGE RETENTION PERIOD

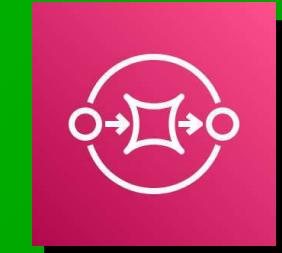
DELIVERY DELAY

*RECEIVEMESSAGE* WAIT TIME

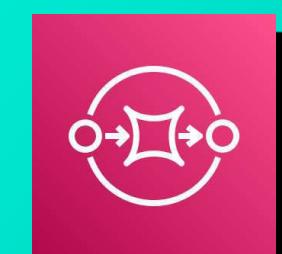
ACCESS POLICY



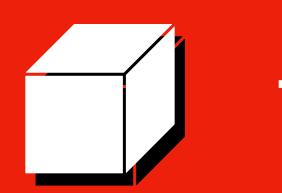
DEAD-LETTER  
QUEUE



DELAY  
QUEUE



TEMPORARY  
QUEUE

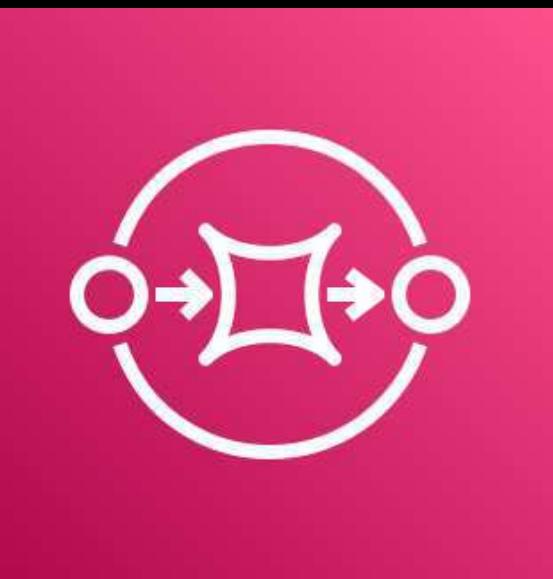


TEMPORARY QUEUE CLIENT

# ENCRYPTION

DATA IN-TRANSIT

DATA AT-REST



## Amazon SQS SECURITY

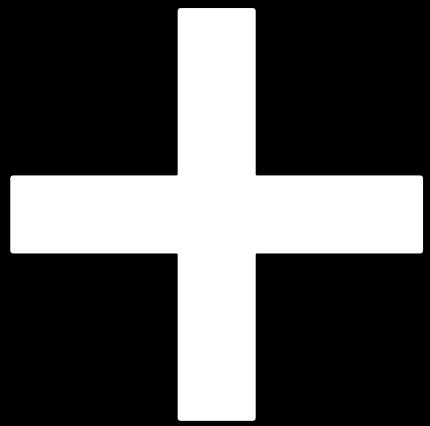
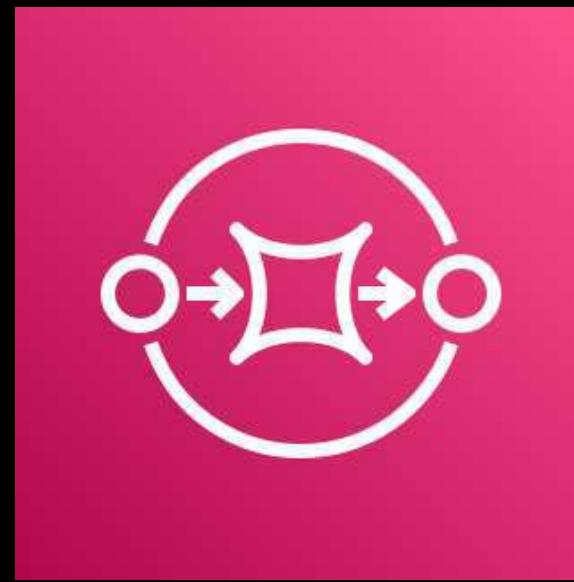
ACCESS POLICY

```
{  
    "Version": "2012-10-17",  
    "Id": "Banana_Queue1_Policy_UUID",  
    "Statement": [{  
        "Sid": "JonBonsoQueue1_SendMessage",  
        "Effect": "Allow",  
        "Principal": {  
            "AWS": [  
                "111122223333"  
            ]  
        },  
        "Action": "sns:Publish",  
        "Resource": "arn:aws:sns:us-east-2:1234:bananaqueue"  
    }]  
}
```



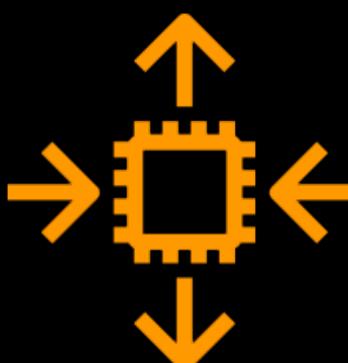
AWS Lambda

LAMBDA TRIGGER



Amazon SNS

FAN-OUT EVENT NOTIFICATION



Amazon EC2 Auto Scaling

AGE OF OLDEST MESSAGE



Amazon S3

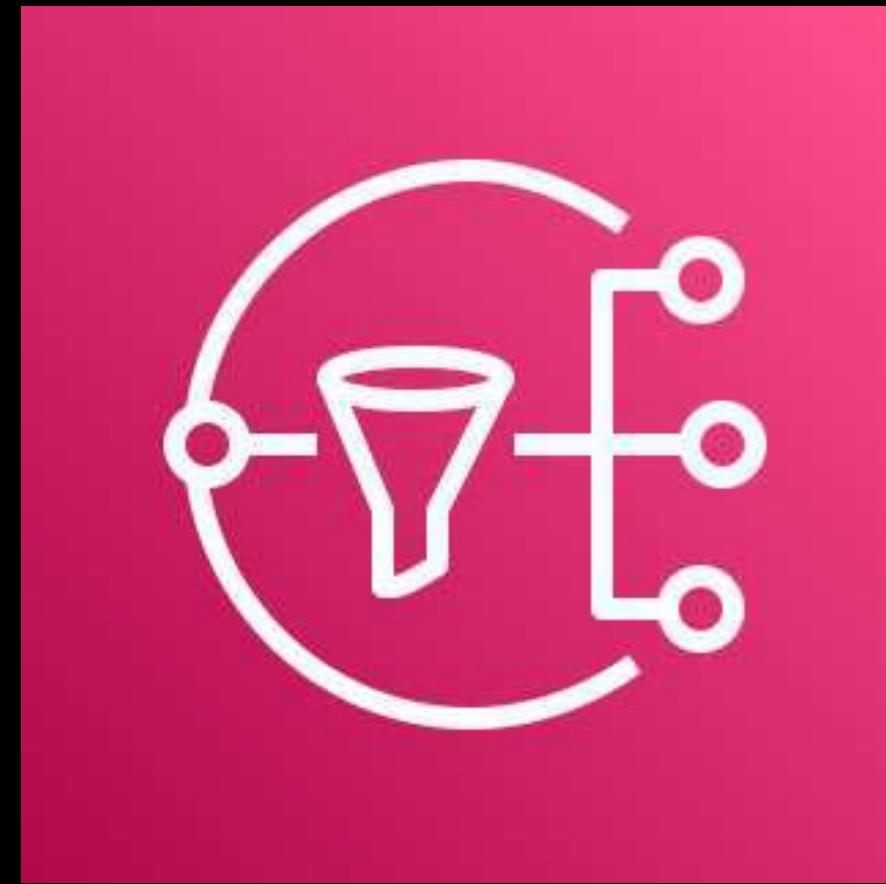
NUMBER OF SQS MESSAGES



Amazon ECS & EKS

S3 EVENT NOTIFICATION

INTER-CONTAINER  
COMMUNICATION



# Amazon **SNS** Overview

---



# NOTIFICATION



# NOTIFICATION

?!?





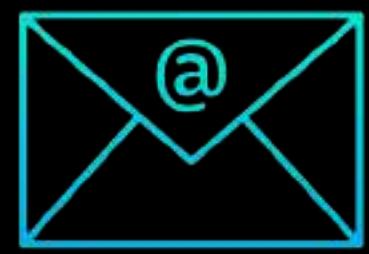
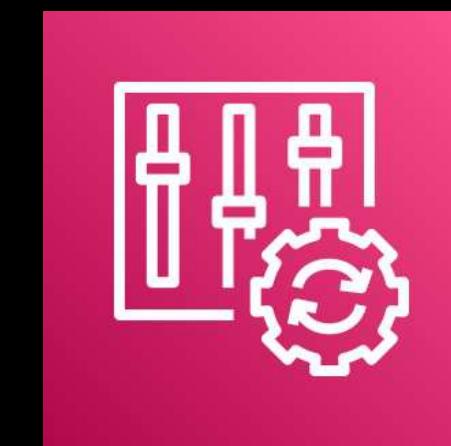
## NOTIFICATION

# Amazon SNS

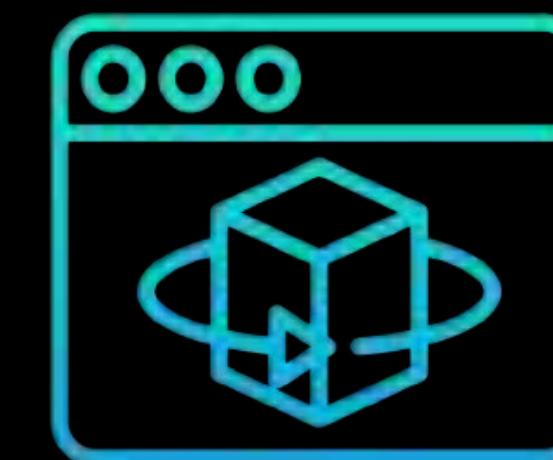
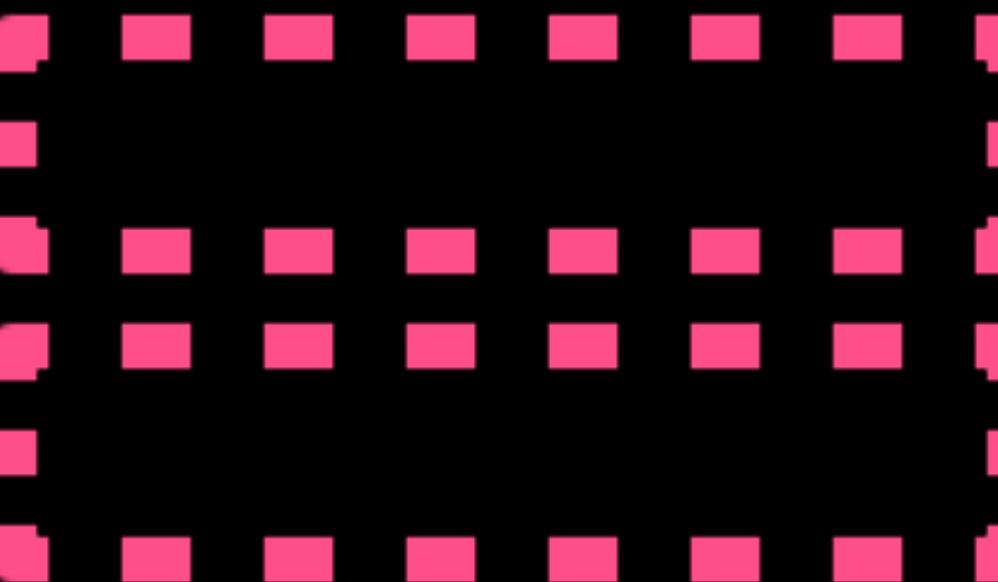
FULLY-MANAGED MESSAGING & **NOTIFICATION SERVICE**

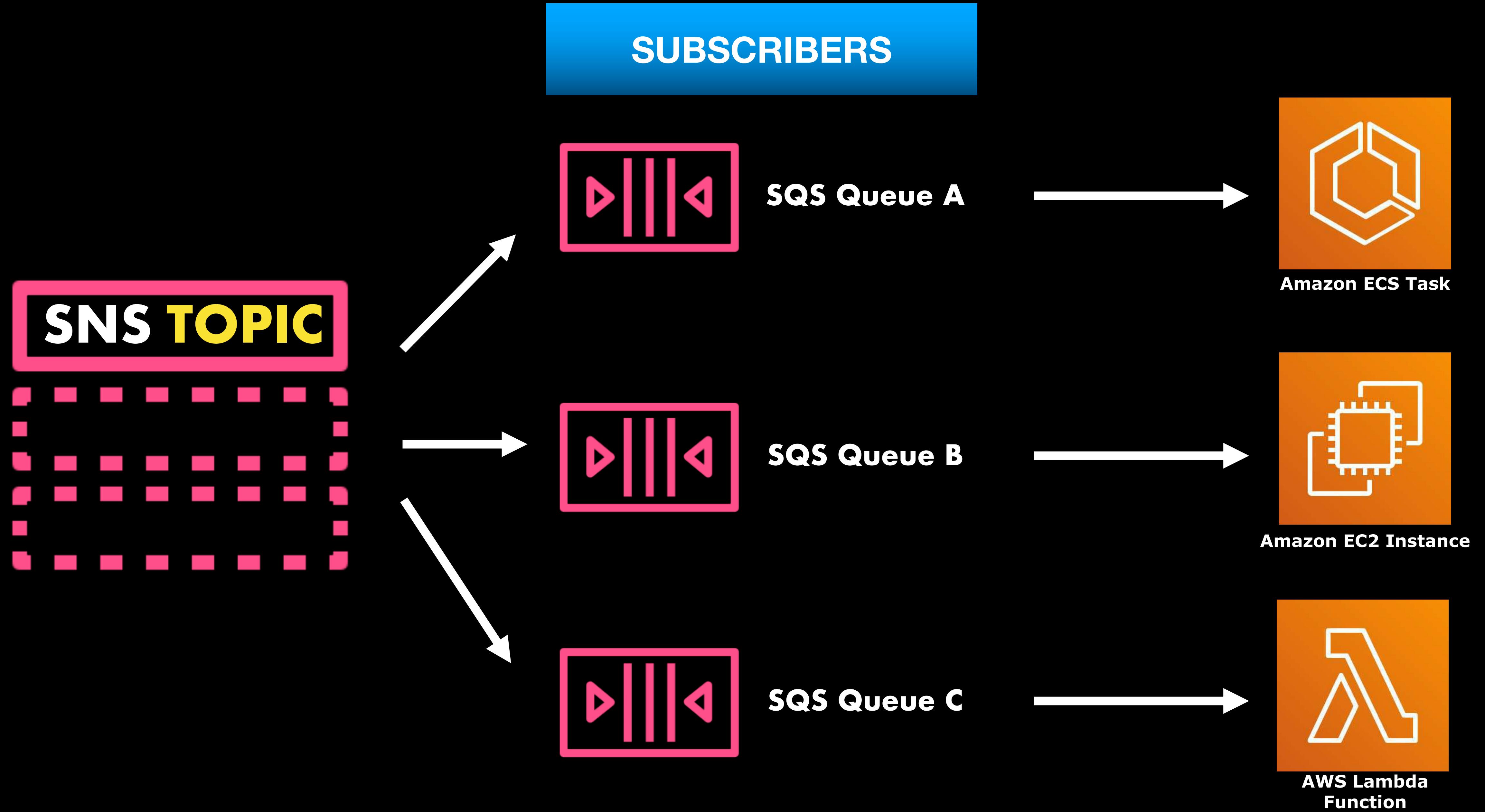
## PUBLISHERS

## SUBSCRIBERS

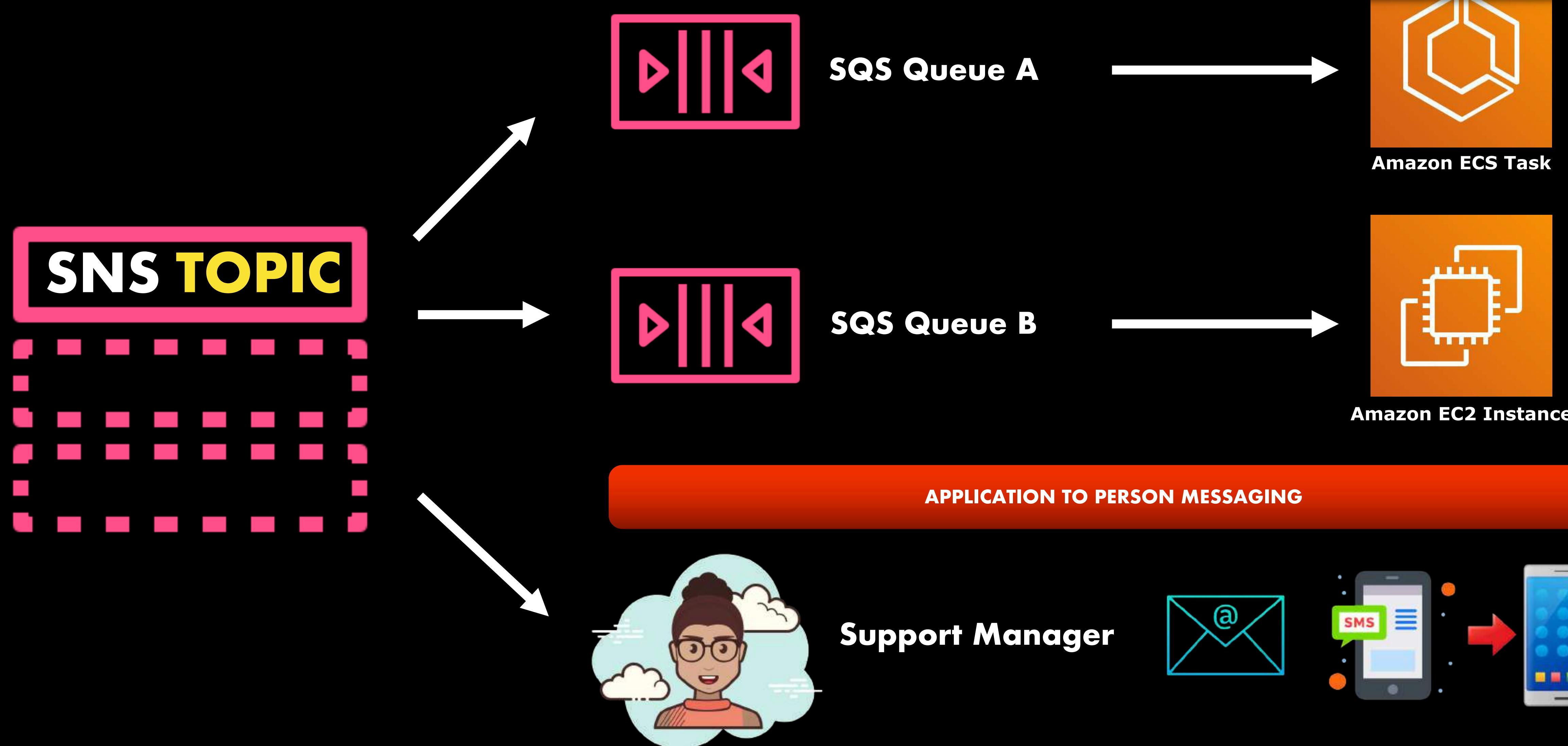


**SNS TOPIC**





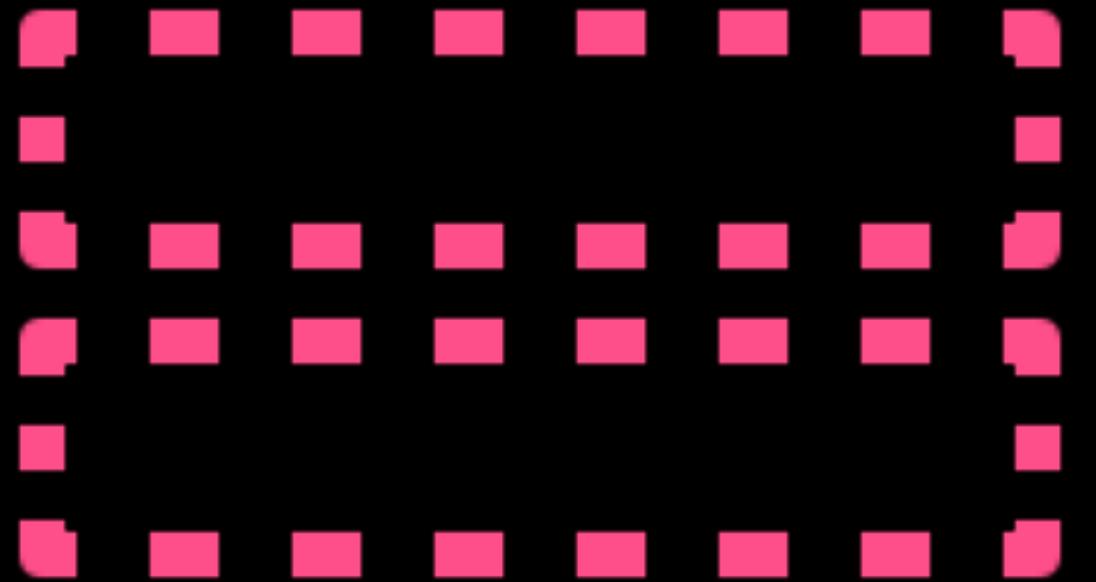
## APPLICATION TO APPLICATION MESSAGING



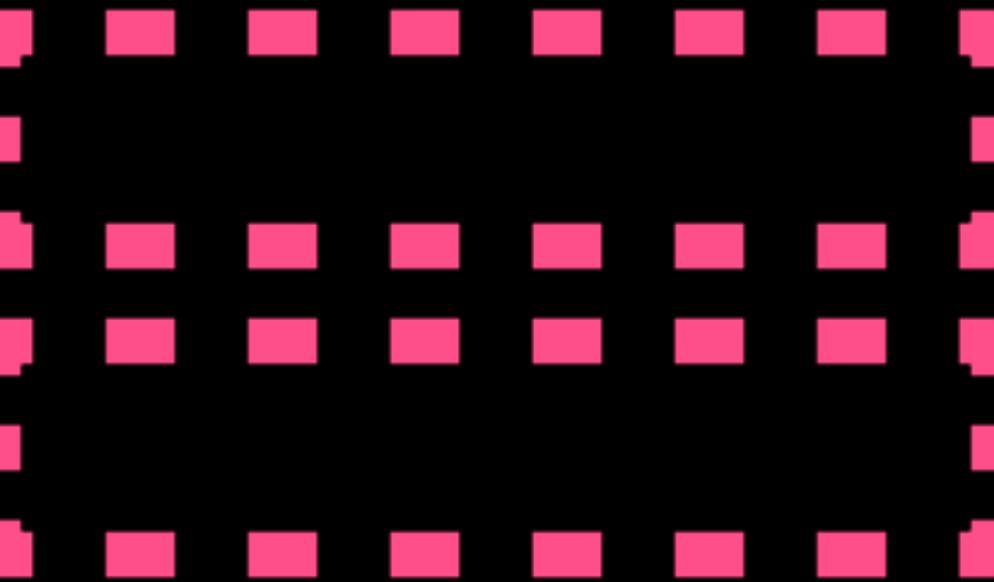


# Amazon SNS Types

**Standard**



**F I F O**





# Amazon SNS Encryption

## ENCRYPTION

DATA IN-TRANSIT

DATA AT-REST

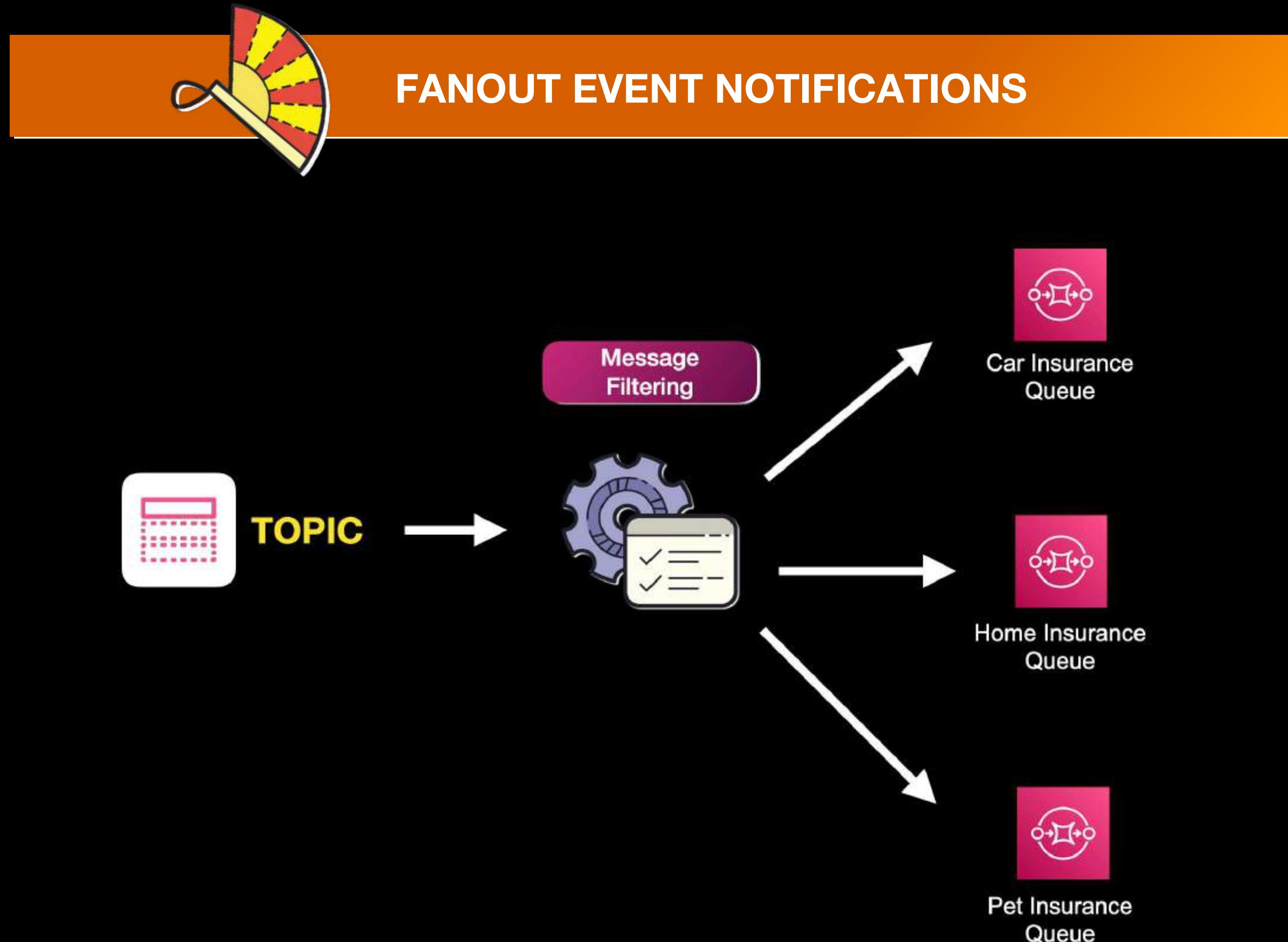
ACCESS POLICY

```
{  
  "Statement": [{}  
    "Sid": "TutorialsDojo-Allow-SNS-SendMessage",  
    "Effect": "Allow",  
    "Principal": {  
      "Service": "sns.amazonaws.com"  
    },  
    "Action": ["sns:SendMessage"],  
    "Resource": "arn:aws:sns:us-east-2:444455556666:TutorialsDojoTopic",  
    "Condition": {  
      "ArnEquals": {  
        "aws:SourceArn": "arn:aws:sqs:us-east-2:444455556666:BananaQueue"  
      }  
    }]  
}
```



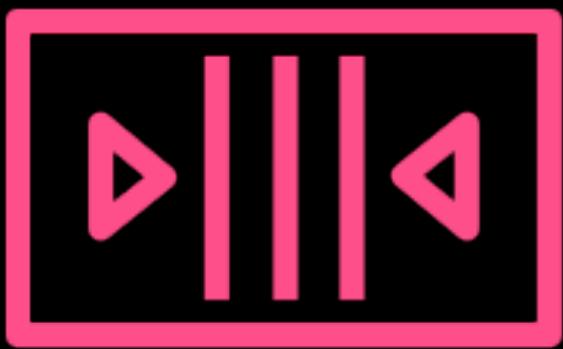
# Amazon SNS Features

- MESSAGE FILTERING
- MESSAGE FANOUT
- MESSAGE DURABILITY
- MESSAGE ENCRYPTION
- MESSAGE ARCHIVING





# Amazon SNS Features



## Dead-Letter Queue (DLQ) for Amazon SNS



Redrive Policy

### ▼ Delivery status logging - optional

These settings configure the logging of message delivery status to CloudWatch Logs. [Info](#)

Log delivery status for these protocols

- AWS Lambda
- Amazon SQS
- HTTP/S
- Platform application endpoint
- Amazon Kinesis Data Firehose

Success sample rate

The percentage of successful message deliveries to log.

0

%



# AWS Amplify Overview

---



## AWS Amplify

- One of the development services in AWS
- Allows you to build extensible, full-stack web and mobile apps faster
- Automates the deployment, scaling and management of your applications and underlying resources
- Provides Machine Learning integration to your apps



## AWS Amplify MODULES

### AWS Amplify Studio

### AWS Amplify Libraries

### AWS Amplify CLI

### AWS Amplify Hosting

Amplify Studio

tutorials-dojo-manila > staging

Welcome back to **Tutorials Dojo Manila App's – staging** environment

Things to do next

Data model Iterate on your app's data model. Create relationships between models and set up authorization rules. <a href="#">Create data model</a>	View and edit app content Use our content editor to create new records or manage existing ones. <a href="#">Manage app content</a>	Configure log in and sign up Configure password-protected login for your app or leverage 3rd party authentication providers. <a href="#">Enable authentication</a>	Accelerate UI development Export UI designs made in Figma to clean React code. Bind UI to backend data. <a href="#">Build UI</a>
--	--	--	--

Deployment activity

Category	Timestamp	Status	Reason
amplify-amplifycb86d1f8efb34-staging-55618	1/8/2024, 5:56:48 AM	<span>✓ CREATE_COMPLETE</span>	-
amplify-amplifycb86d1f8efb34-staging-55618	1/8/2024, 5:56:19 AM	<span>CREATE_IN_PROGRESS</span>	User Initiated
Create backend	1/8/2024, 5:56:15 AM	<span>✓ COMPLETED</span>	-

AWS Amplify Studio is supported by Amazon Web Services © 2022, Amazon Web Services, Inc. and its affiliates. All rights reserved. View the site terms and privacy policy.



# AWS Amplify M O D U L E S

# AWS Amplify Studio

# AWS Amplify Libraries

# AWS Amplify CLI

# AWS Amplify Hosting

Amplify Dev Center Docs Learn UI Library Pricing About AWS Amplify

Amplify UI

Angular Flutter React Vue

Getting started Components

Overview

BASE

Divider Heading Icon Image ScrollView Text View

FEEDBACK

Alert Loader Placeholder

NAVIGATION

Link Menu Tabs

INPUTS

Button

# Button

Click me!

Variation

Default

Size

Default

isFullWidth

isDisabled

isLoading

loadingText

<Button loadingText="" onClick={() => alert('hello')} ariaLabel="">  
Click me!  
</Button>

Copy

## Usage

Import the Button primitive and styles.

Hello world

```
import { Button } from '@aws-amplify/ui-react';
import '@aws-amplify/ui-react/styles.css';

<Button>Hello world</Button>;
```

Copy



# AWS Amplify MODULES

## AWS Amplify Studio

## AWS Amplify Libraries

## AWS Amplify CLI

## AWS Amplify Hosting

Data > Local setup instructions

Local setup instructions for

React

Skip and setup later

### 1 Install Amplify CLI to pull the data model

Open your existing React project. If you do not have one, create a new React project:

```
npx create-react-app@latest myapp  
cd myapp
```



Install the Amplify CLI. The Amplify CLI is a command line toolchain that runs locally in order to communicate with your app backend.

```
curl -sL https://aws-amplify.github.io/amplify-cli/install | bash && $SHELL
```



Run the following command from your project's root folder (myapp):

```
amplify pull --sandboxId 0e111a3e-a370-45f5-95df-9e572b815a70
```



Back

Next

### 2 Install Amplify library and initialize Amplify

### 3 Test CRUD APIs locally with Amplify DataStore



## AWS Amplify MODULES

### AWS Amplify Studio

### AWS Amplify Libraries

### AWS Amplify CLI

### AWS Amplify Hosting

The screenshot shows the AWS Amplify Hosting setup interface. At the top, there's a navigation bar with the AWS logo, a search bar, and options for Services, Help, and Account (N. Virginia). Below the navigation is a sidebar titled "AWS Amplify" containing a list of existing apps: Manila-App, Mandaluyong-App, Caloocan-App, and Makati-App. Under the sidebar, there are links for Documentation and Support.

The main content area is titled "Get started with Amplify Hosting" and describes Amplify Hosting as a fully managed hosting service for web apps. It provides options to connect source code from a Git repository or upload files. A section titled "From your existing code" lists providers: GitHub (selected), Bitbucket, GitLab, AWS CodeCommit, and Deploy without Git provider. Each provider has a corresponding icon. At the bottom of this section, it notes that "Amplify Hosting requires read-only access to your repository." A "Continue" button is located at the bottom right of the main content area.



# **Serverless Computing**

## Overview

---

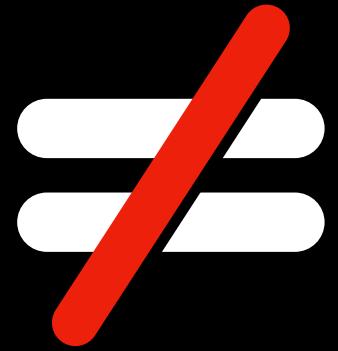
What is  
**Serverless Computing?**



## **Serverless Computing**

- **On-Demand Service**
- **Less Server Management**

**Serverless**



**No Server ?**

**Serverless**



**Less Server Management**

# Serverless



# FaaS



AWS Lambda



AWS Fargate



Amazon Aurora  
Serverless



Amazon DynamoDB



Amazon S3

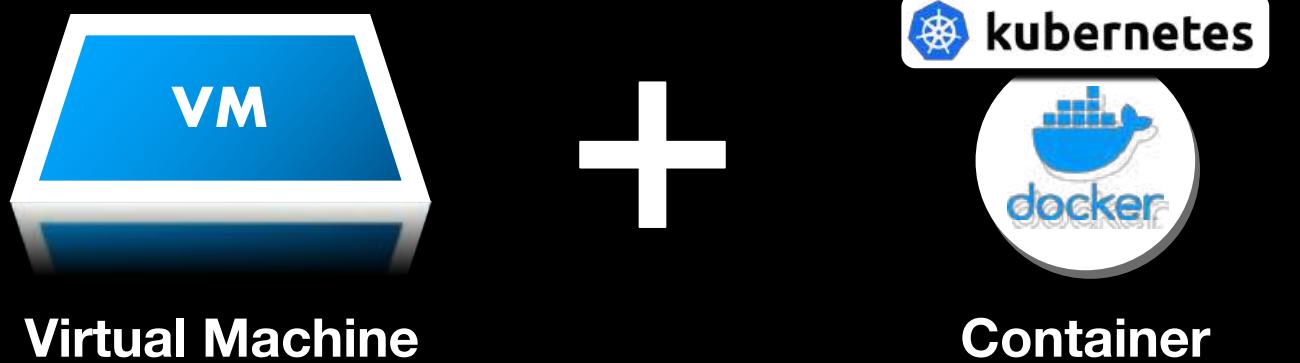
# Serverless

powered by



# microVMs

# microVMs



# Serverless



```
function fetchDojoDataAndStoreToS3(){

  /*
   * Fetch data from Tutorials Dojo API
   */
  let UserData = await fetch(
    `https://data.tutorialsdojo.com/getUsers/manila`
    ).response.json();

  /*
   * Store user data to my S3 bucket
   */
  let upload = new AWS.S3.ManagedUpload({
    params: {
      Bucket: albumBucketName,
      Key: dataKey.generate(),
      Body: UserData
    }
  });
}
```

**Serverless**



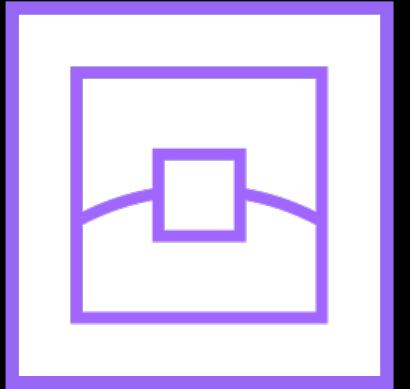
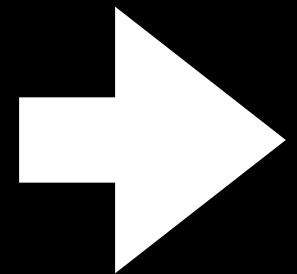
**Edge Computing**

# Edge Computing

Serverless



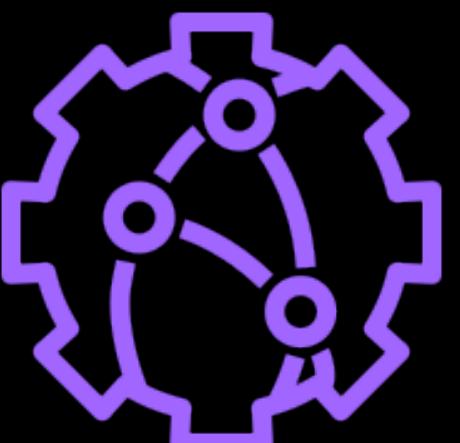
```
function fetchDojoDataAndStoreToS3(){  
    /*  
     * Fetch data from Tutorials Dojo API  
     */  
    let UserData = await fetch(`https://data.tutorialsdojo.com/getUsers/manila`)  
        .response.json();  
  
    /*  
     * Store user data to my S3 bucket  
     */  
    let upload = new AWS.S3.ManagedUpload({  
        params: {  
            Bucket: albumBucketName,  
            Key: dataKey.generate(),  
            Body: UserData  
        }  
    });  
}
```



Edge Location

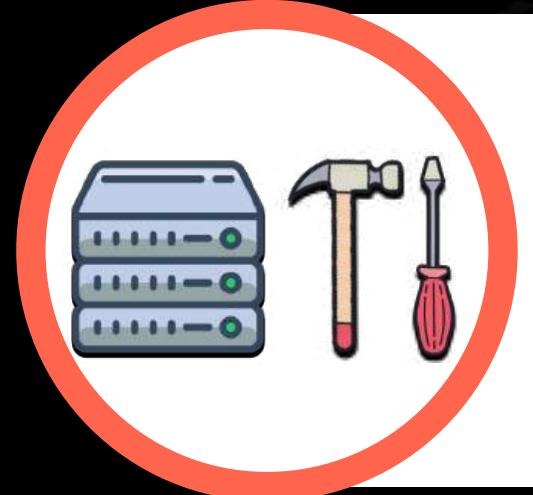


Lambda@Edge



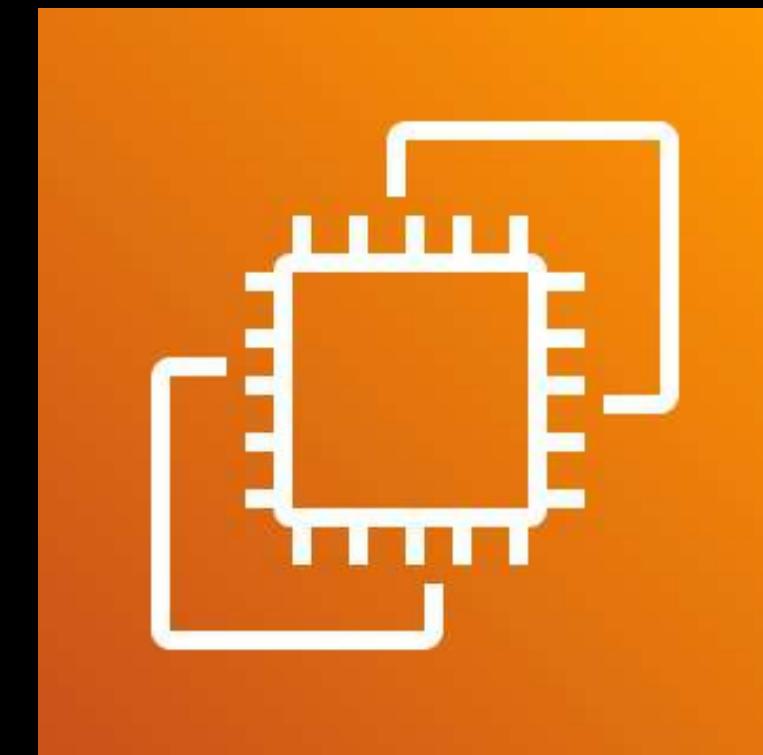
CloudFront Function

## Traditional



- - Virtual Server Deployment
- - OS Patching
- - Storage Management

## Infrastructure-as-a-Service (IaaS)



- - Virtual Server Management
- - Virtual Server Maintenance
- - Scaling

## Function-as-a-Service (FaaS)



## Serverless



**Serverless**

- Does **NOT** run all the time unlike a traditional virtual machine
- Will only run **once you invoked it**
- Start up time ranges from several milliseconds to less than a second
- Can only run your function continuously for **15 minutes**



**Amazon DynamoDB**



**Amazon Aurora  
Serverless**



# **Serverless Computing**

## Architectures

---

**Serverless**

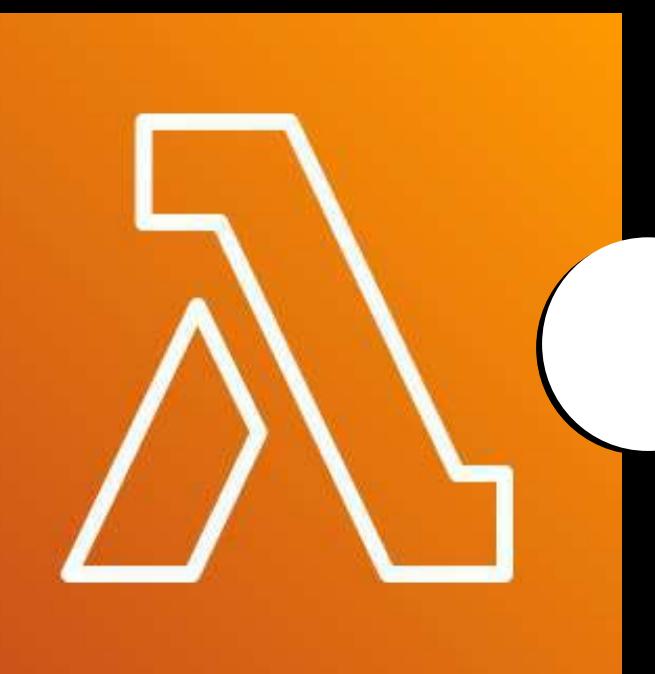
=

**Less Server Management**

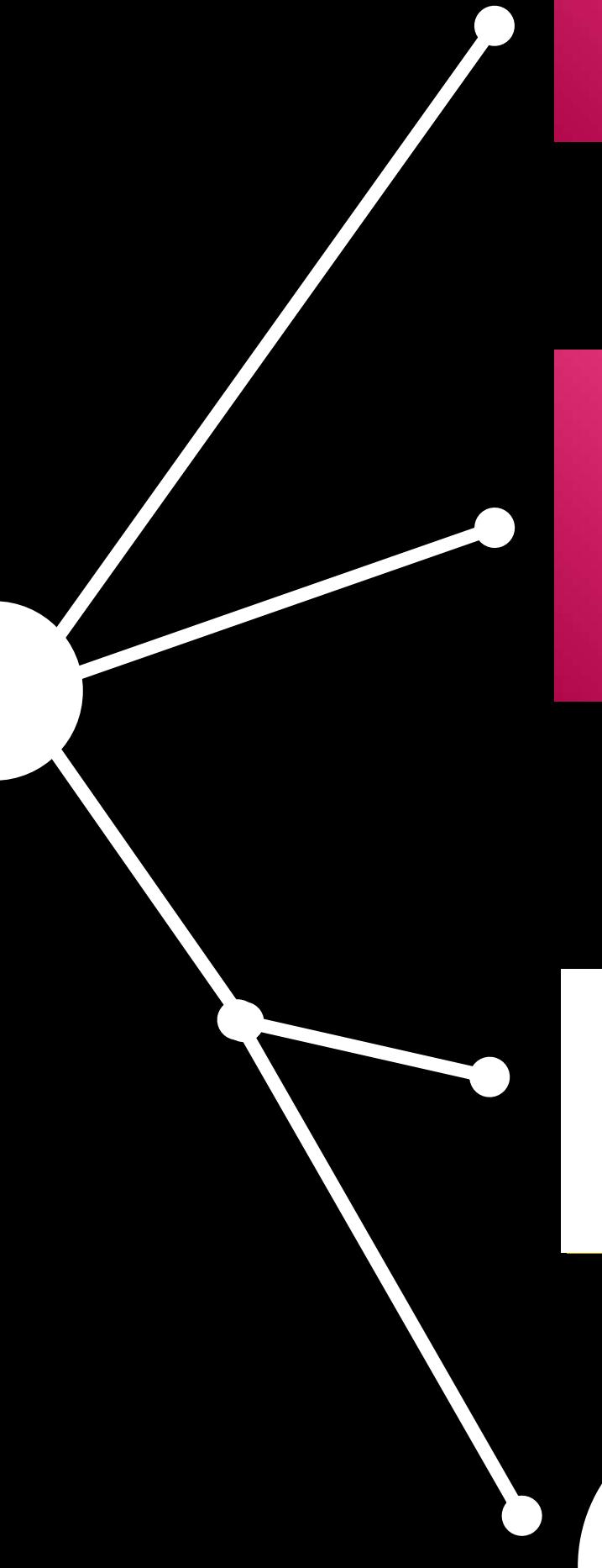


# Function as a Service (FaaS)

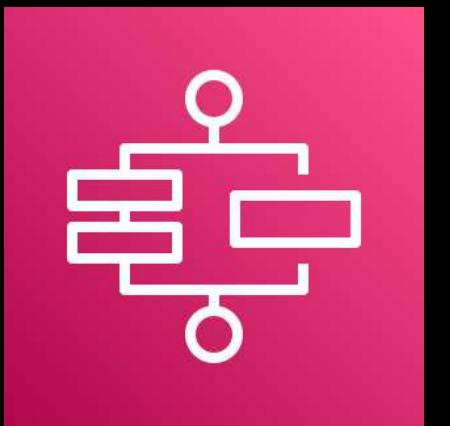
```
function fetchDojoDataAndStoreToS3(){  
    /*  
     * Fetch data from Tutorials Dojo API  
     */  
    let UserData = await fetch(  
        `https://data.tutorialsdojo.com/getUsers/manila`  
    ).response.json();  
  
    /*  
     * Store user data to my S3 bucket  
     */  
    let upload = new AWS.S3.ManagedUpload({  
        params: {  
            Bucket: albumBucketName,  
            Key: dataKey.generate(),  
            Body: UserData  
        }  
    });  
}
```



AWS Lambda



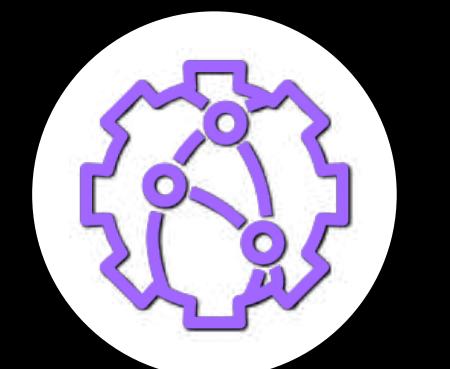
Amazon EventBridge



AWS Step Functions



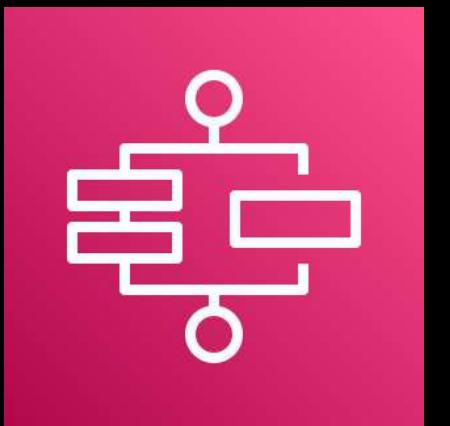
AWS Lambda@Edge



CloudFront Functions



Scheduled Actions



Orchestration



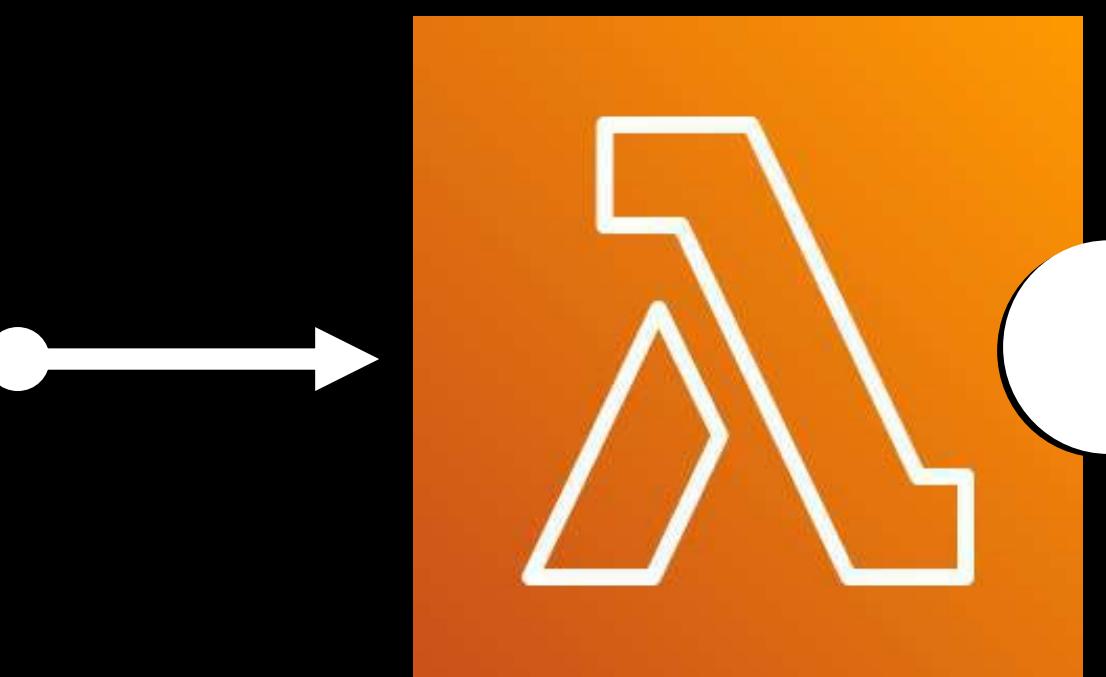
Edge Computing at  
Regional Edge Locations



Edge Computing at  
Edge Locations

# Function as a Service (FaaS)

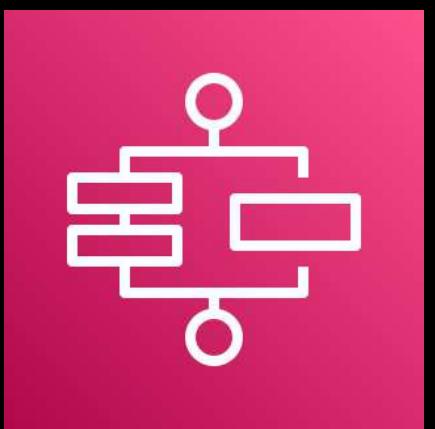
```
function fetchDojoDataAndStoreToS3(){  
  
    /*  
     * Fetch data from Tutorials Dojo API  
     */  
    let UserData = await fetch(  
        `https://data.tutorialsdojo.com/getUsers/manila`  
    ).response.json();  
  
    /*  
     * Store user data to my S3 bucket  
     */  
    let upload = new AWS.S3.ManagedUpload({  
        params: {  
            Bucket: albumBucketName,  
            Key: dataKey.generate(),  
            Body: UserData  
        }  
    });  
}
```



AWS Lambda



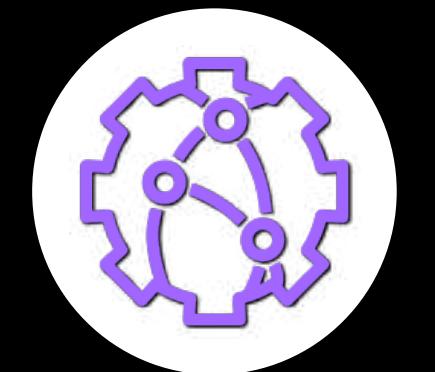
Amazon EventBridge



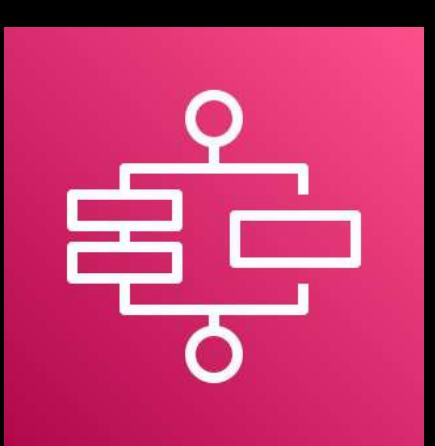
AWS Step Functions



AWS Lambda@Edge



CloudFront Functions

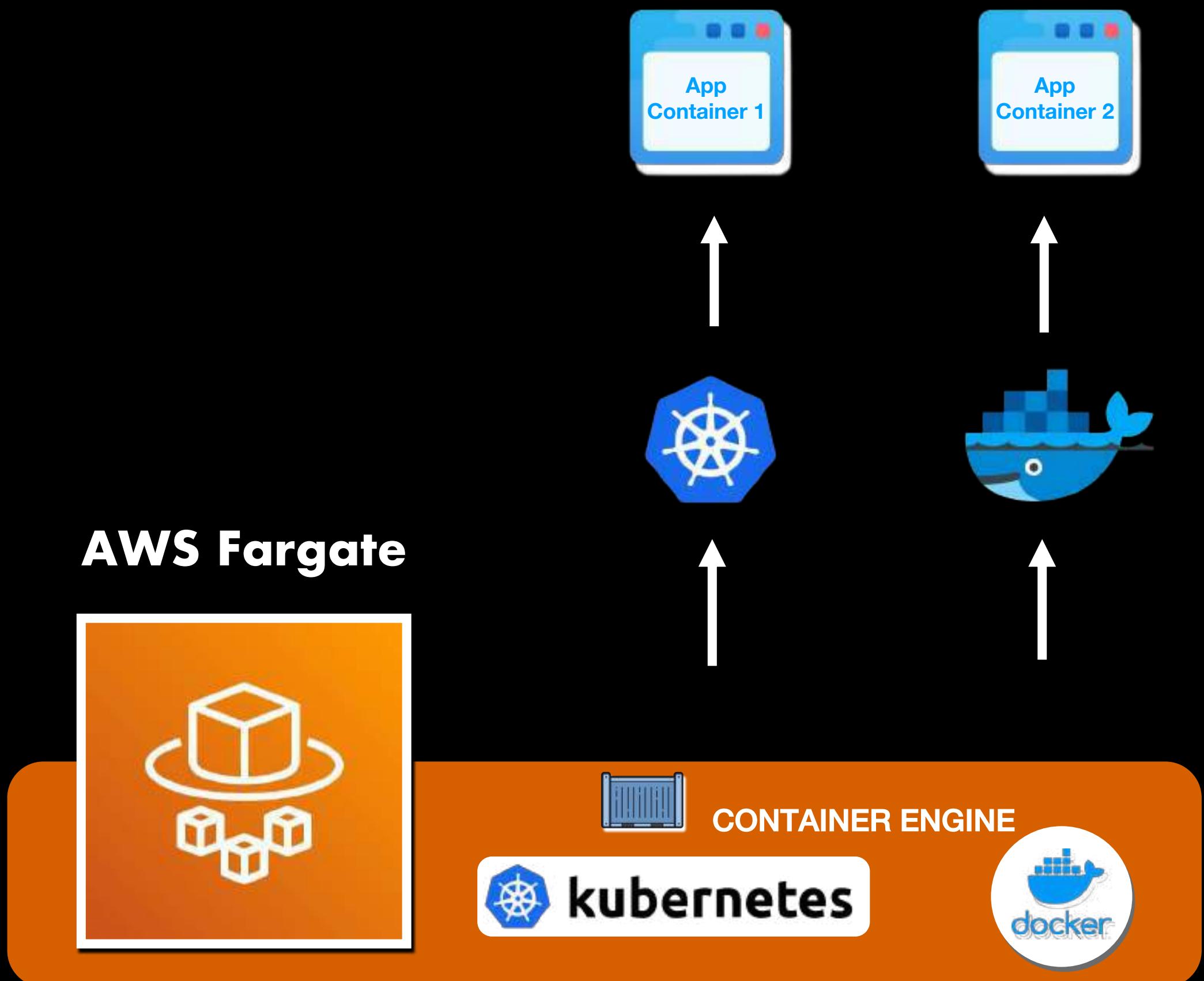


Orchestration

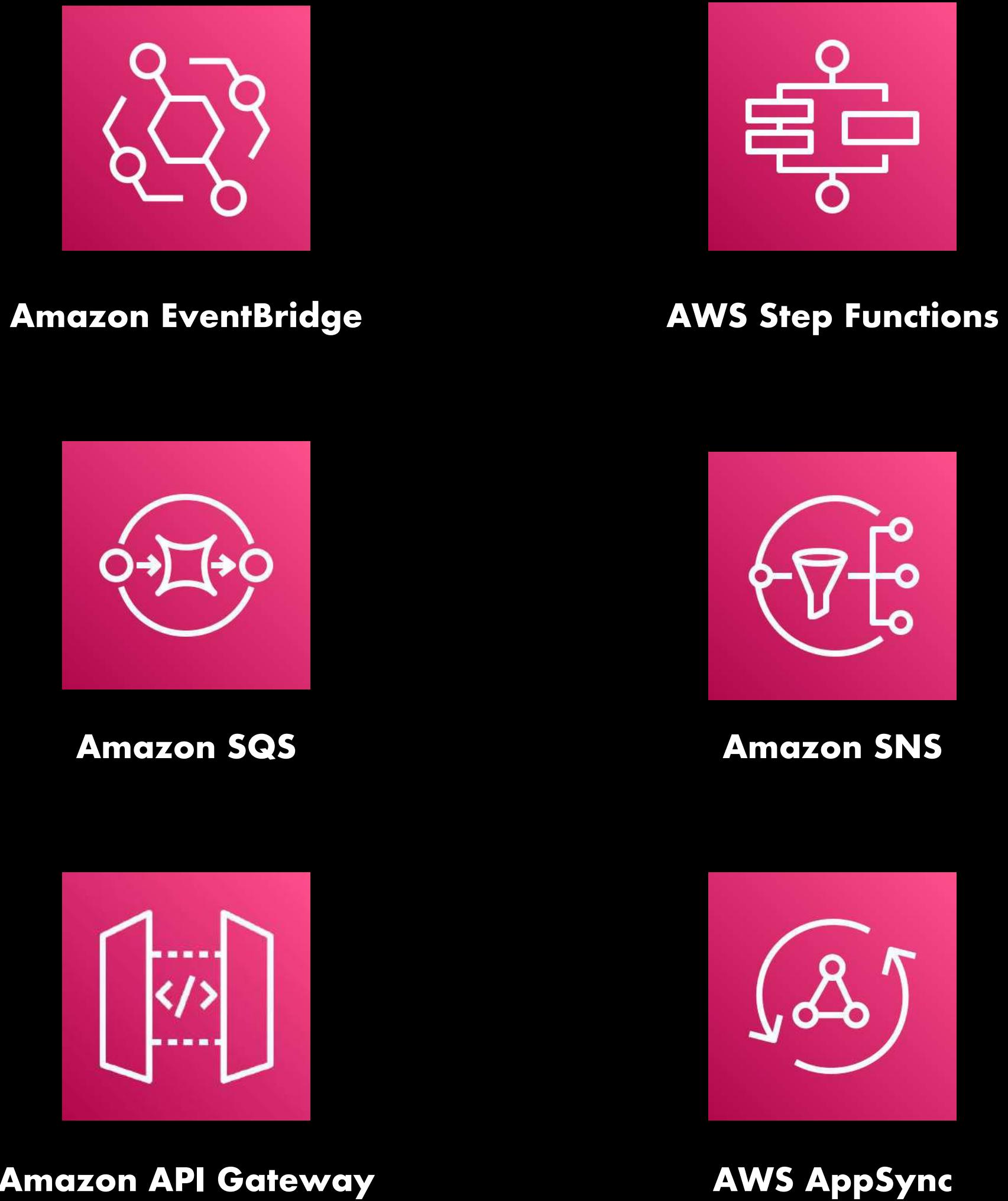


Scheduled Actions

## SERVERLESS CONTAINERS



## SERVERLESS APPLICATION INTEGRATION



## SERVERLESS DATA STORES

STATIC DATA



Amazon S3

DYNAMIC DATA



Amazon  
DynamoDB



Amazon Aurora  
Serverless

DATA  
WAREHOUSE



Amazon Redshift Spectrum

## SERVERLESS ETL & ANALYTICS

Extract, Transform &  
Load (ETL)



AWS Glue

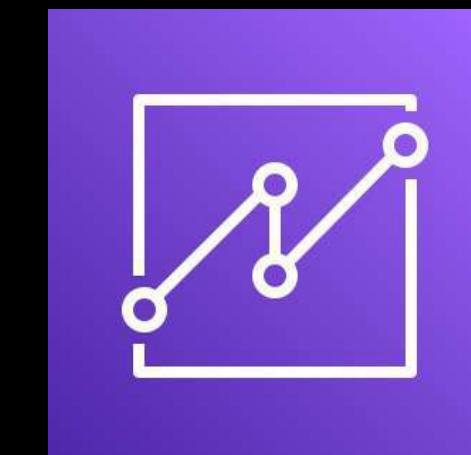
Analytics Services



Amazon Athena



Amazon Kinesis  
Data Analytics



Amazon QuickSight



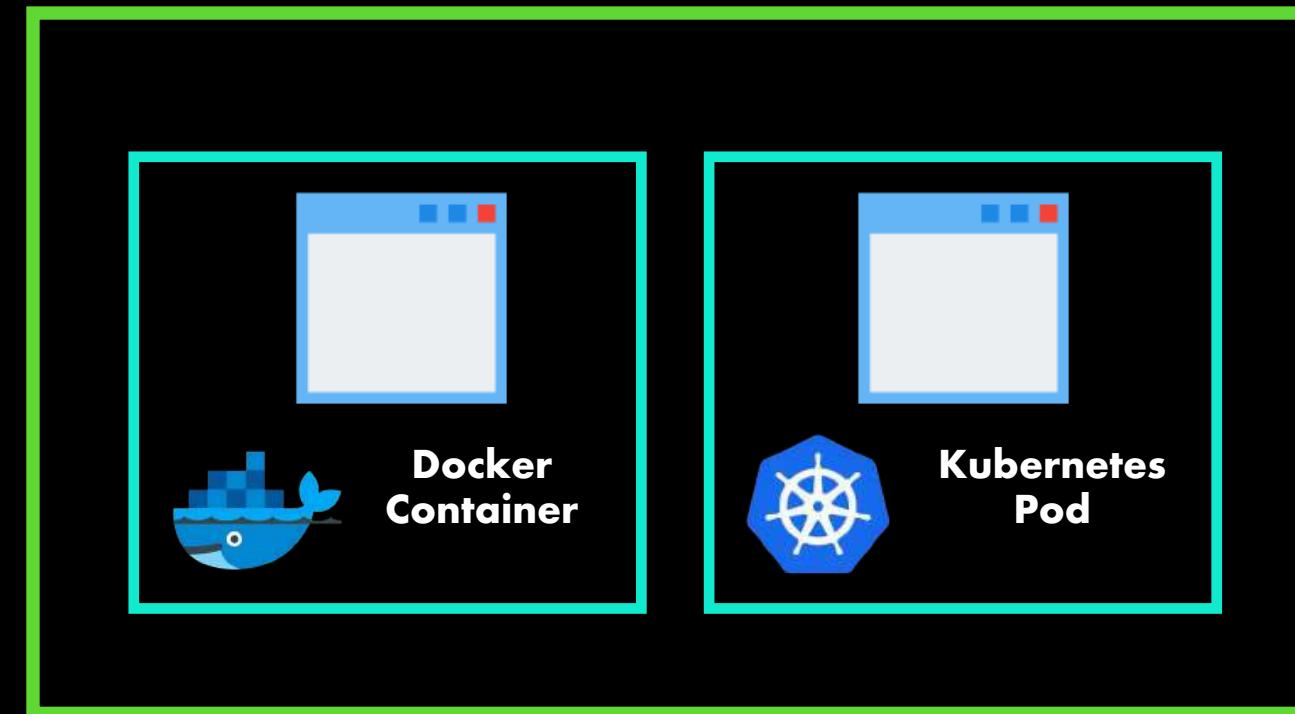
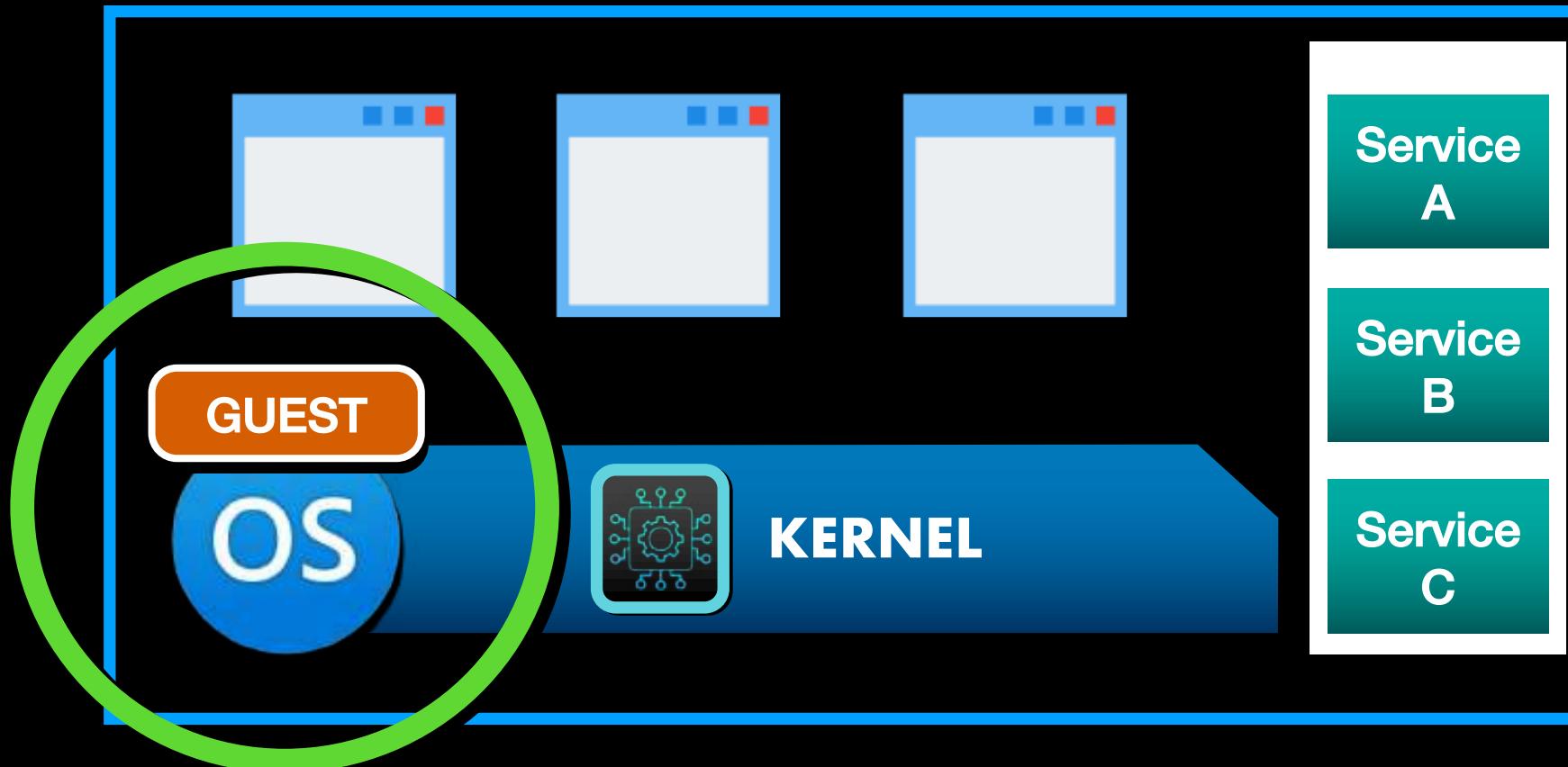
## Virtual Machine



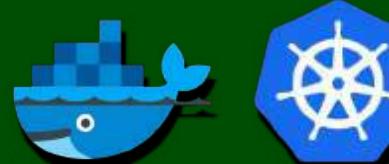
## Container



## MicroVM



**AWS NITRO HYPERVISOR /  
VIRTUAL MACHINE MONITOR (VMM)**



**CONTAINER ENGINE**

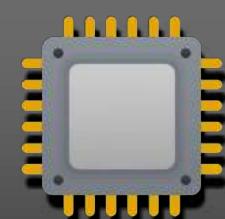


**Firecracker Virtualization /  
VIRTUAL MACHINE MONITOR (VMM)**

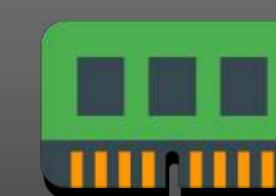
**HOST**



**HARDWARE /  
BARE-METAL SERVER**



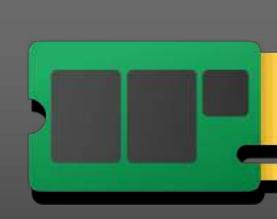
**CPU**



**MEMORY  
(RAM)**



**NETWORK**



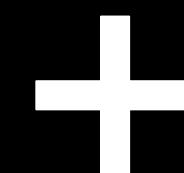
**SSD/HDD STORAGE**



**KERNEL**

# Serverless Architecture Types

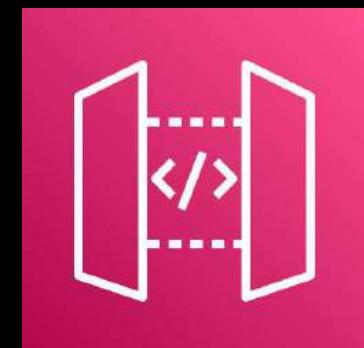
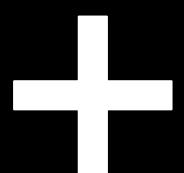
## Static Single Page Application



Amazon S3

Amazon CloudFront

## Service-Oriented Architecture



AWS Lambda

API Gateway

## Containerized Application

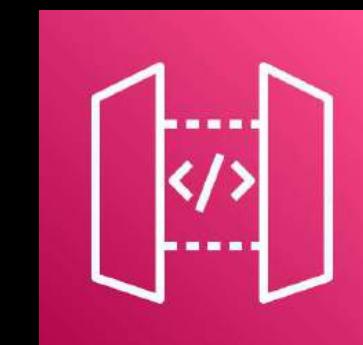


AWS Fargate

## Serverless Architecture



AWS Lambda



API Gateway



AWS Fargate



Amazon  
DynamoDB



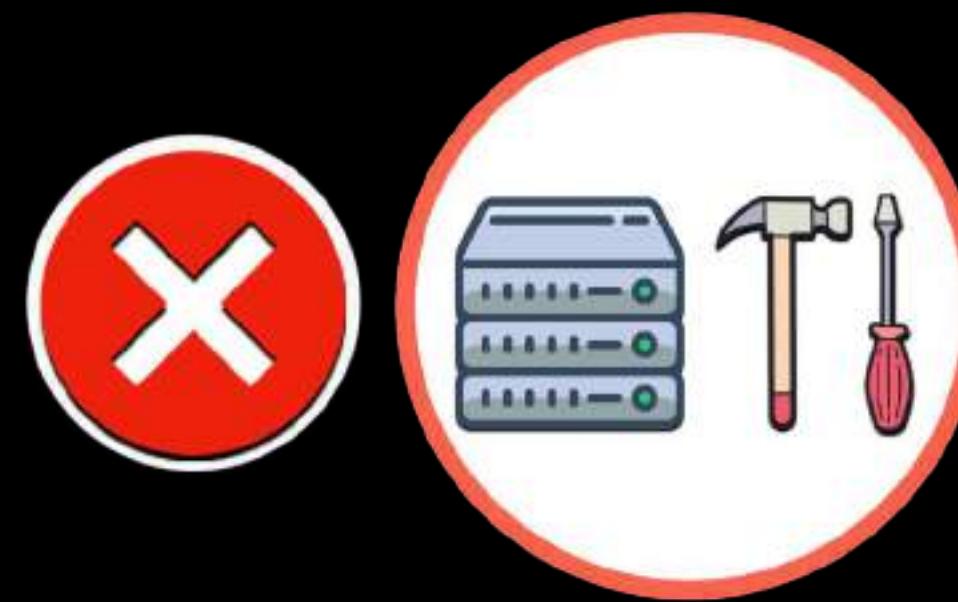
Amazon  
DynamoDB



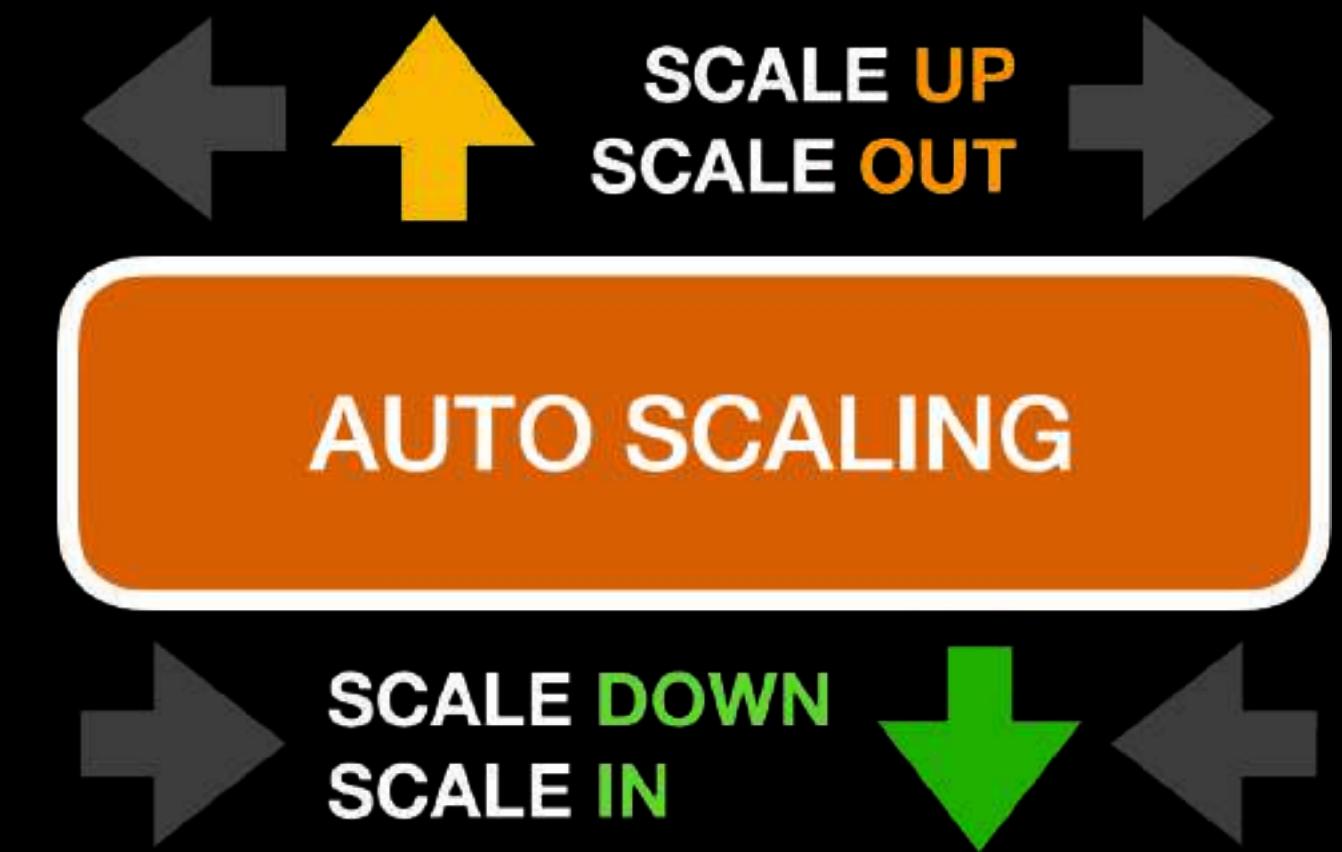
Amazon Aurora  
Serverless

- For applications that have **sporadic or infrequent database usage patterns**
- **No need to choose a particular DB instance type or do any advanced capacity planning**
- **Automatically increases and decreases the compute and storage capacity of your database**
- **Unlike RDS, there's no need to downgrade your database instance if your demand decreases**
- **Costs way less than a regular server-based database**

# Serverless



**Less Server Management**

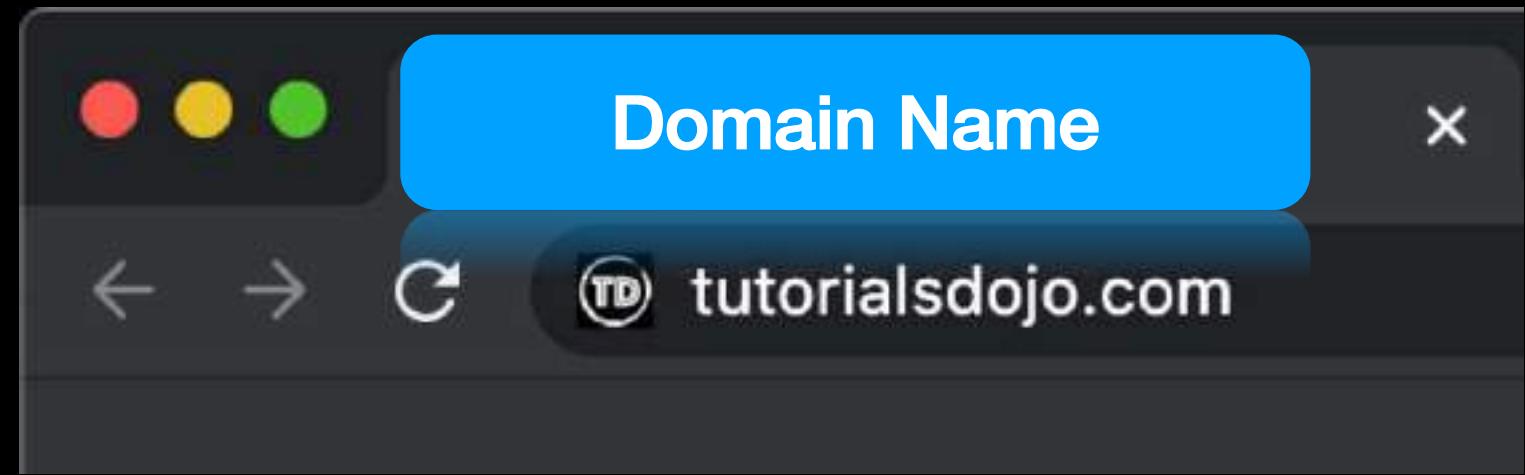




# Amazon Route 53 Overview

- A global **Domain Name System (DNS)** service
- Provides different **Routing Policies**
- Allows you to **register your own domain name**
- Transfer a domain from another domain registrar
- Create **health checks**
- Route **traffic flows**
- Configure **DNS resolvers**
- . . . and many more!





## Domain Name System (DNS)



Elastic IP  
address



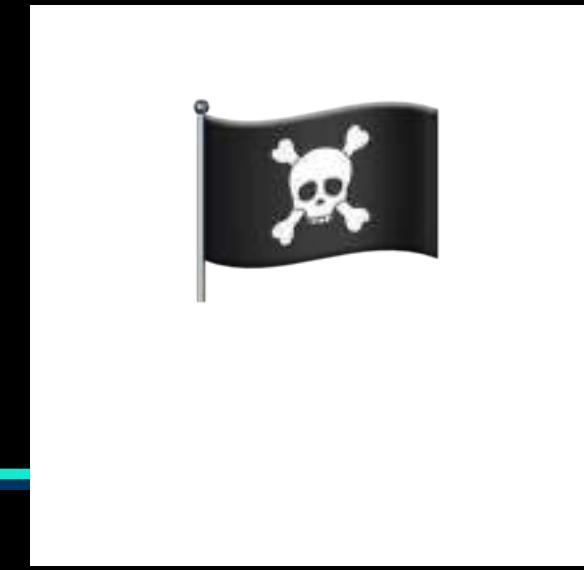
Amazon S3  
Static Website



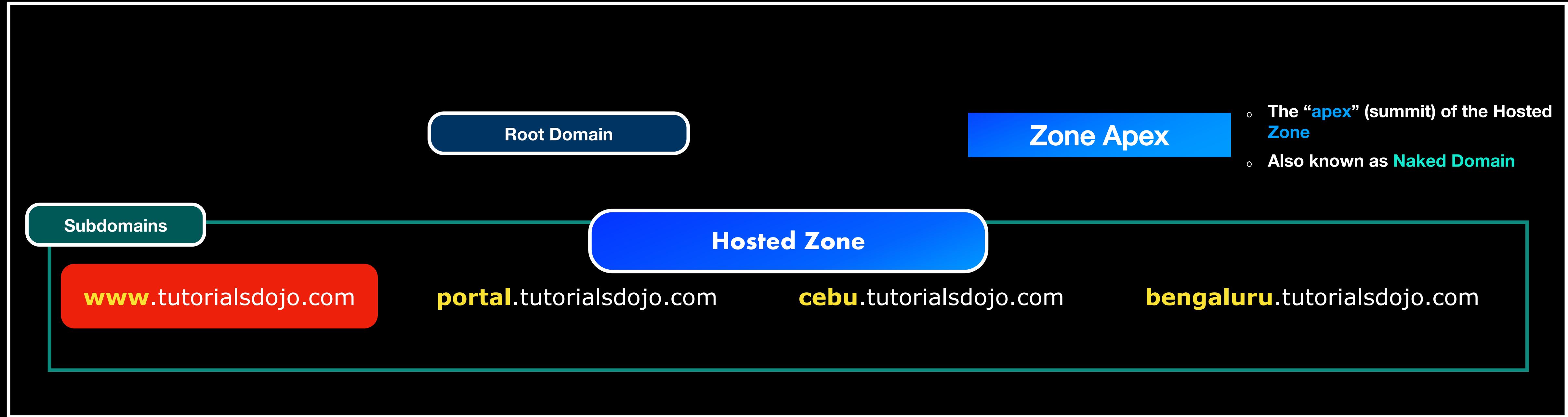
Elastic Load  
Balancers

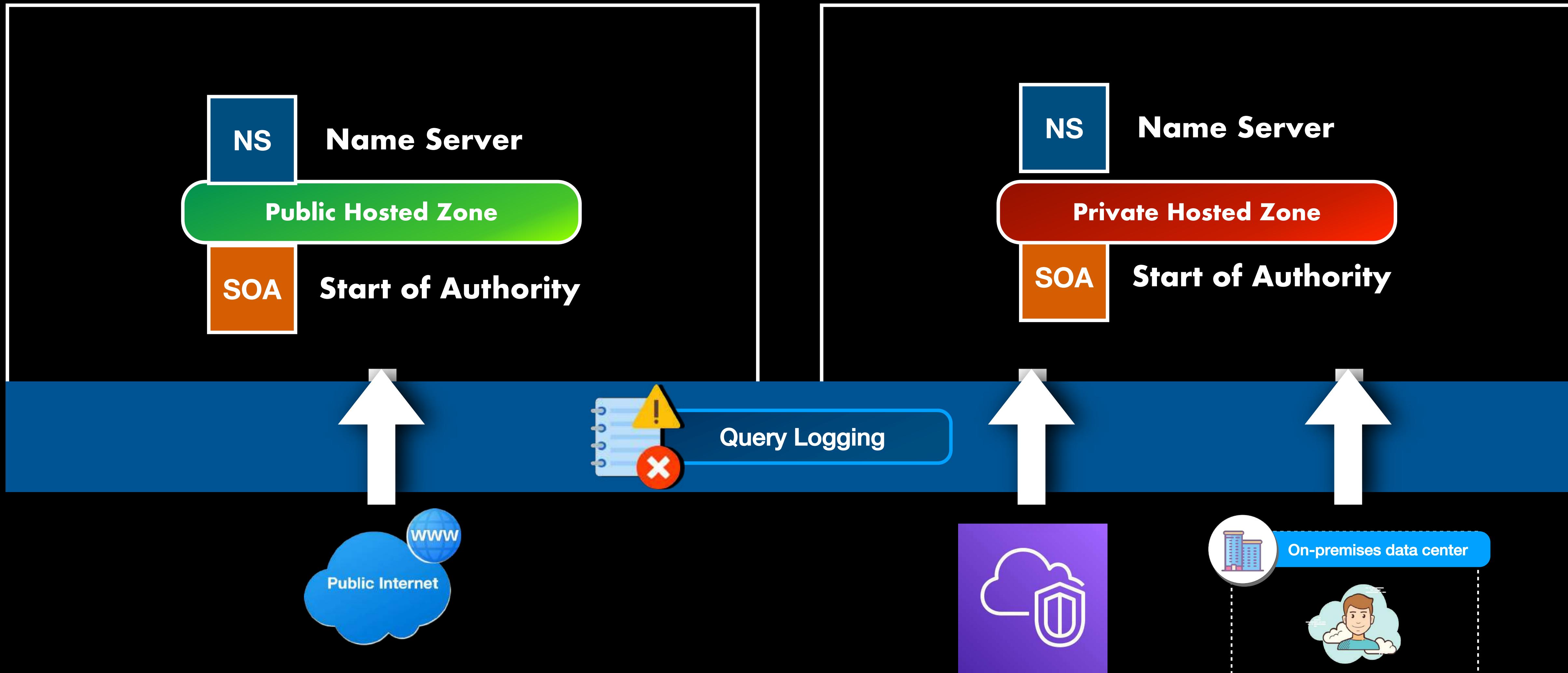


Amazon CloudFront  
Web Distributions



Man-In-The-Middle Attacks





## ALIAS RECORD



## NON-ALIAS RECORD

49.143.173.201

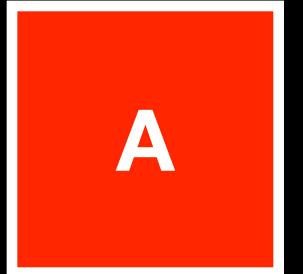
- Route traffic to selected **AWS resources**
- Works like a **CNAME (Canonical Name) Record**
- Not visible to DNS resolvers
- Points to a **specific AWS resource**

- Allows you to specify the **IP addresses or the custom domain names of your servers or resources**
- Visible to DNS resolvers
- Points to a **particular IP address**



## DNS RECORD T Y P E S

ALIAS



**IPv4 Host Address**

ALIAS



**IPv6 Host Address**



**Pointer**



**Service Locator**



**Name Server**



**Canonical Name**



**Sender Policy Framework**



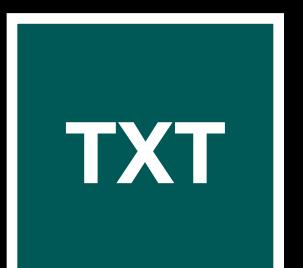
**Start of Authority**



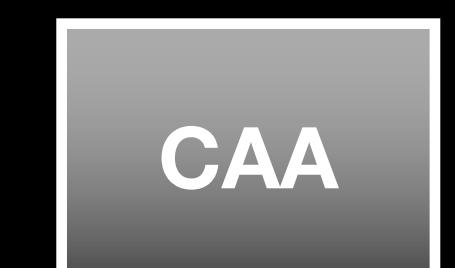
**Mail Exchange**



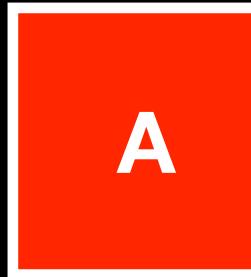
**Naming Authority Pointer**



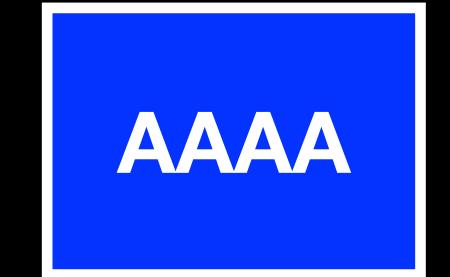
**Text**



**Certification Authority  
Authorization**



**IPv4 Host Address**



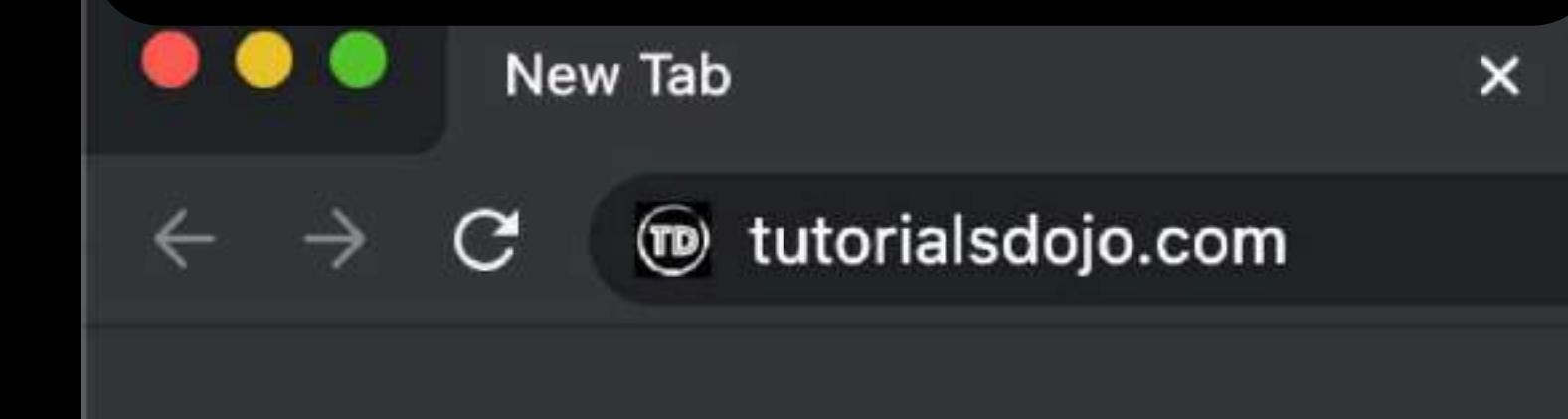
**IPv6 Host Address**



**Canonical Name**



**Root Domain / Zone Apex**



- An open-source program that you can use as a fully customizable domain name server
- Usually launched by companies as their internal DNS service
- Stands for Berkeley Internet Name Domain server
- Has a BIND DNS forwarder that allows you to resolve the domain names in the private hosted zones in AWS from your on-premises network
- Can be migrated to Amazon Route 53 by importing the BIND zone file



**ACTIVE**

Live Traffic

**PASSIVE**

FailOver

ACTIVE

ACTIVE

ACTIVE

PASSIVE

- Improves fault tolerance and performance of your applications
- Entails additional cost
- Has several active environments that accepts live production traffic
- Ensures the high availability and resiliency of your global applications
- Can be implemented by using a single policy, or a combination of routing policies such as:

Geolocation

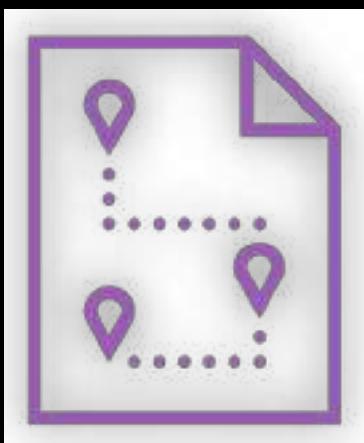
Geoproximity

Latency

Multivalue Answer

Weighted

...other routing types!

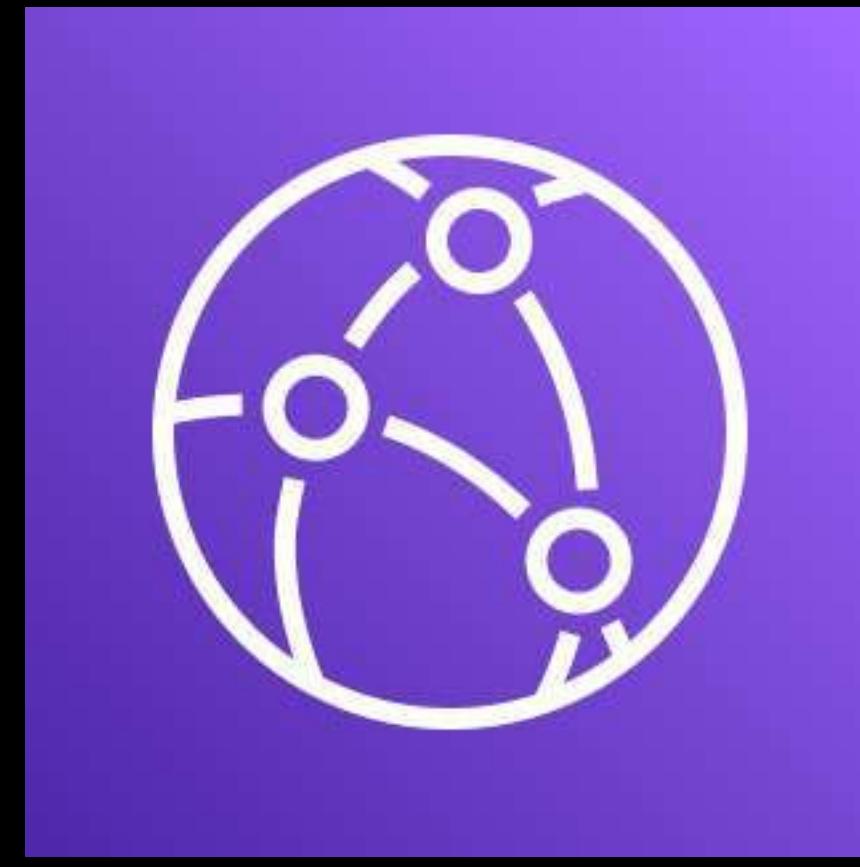


Failover Policy

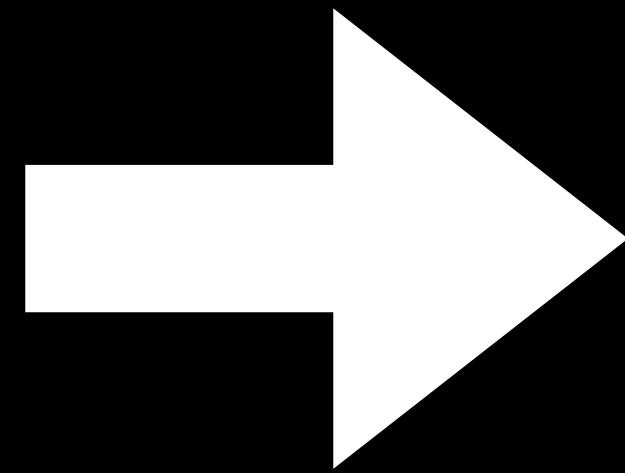


# Amazon CloudFront Overview

---

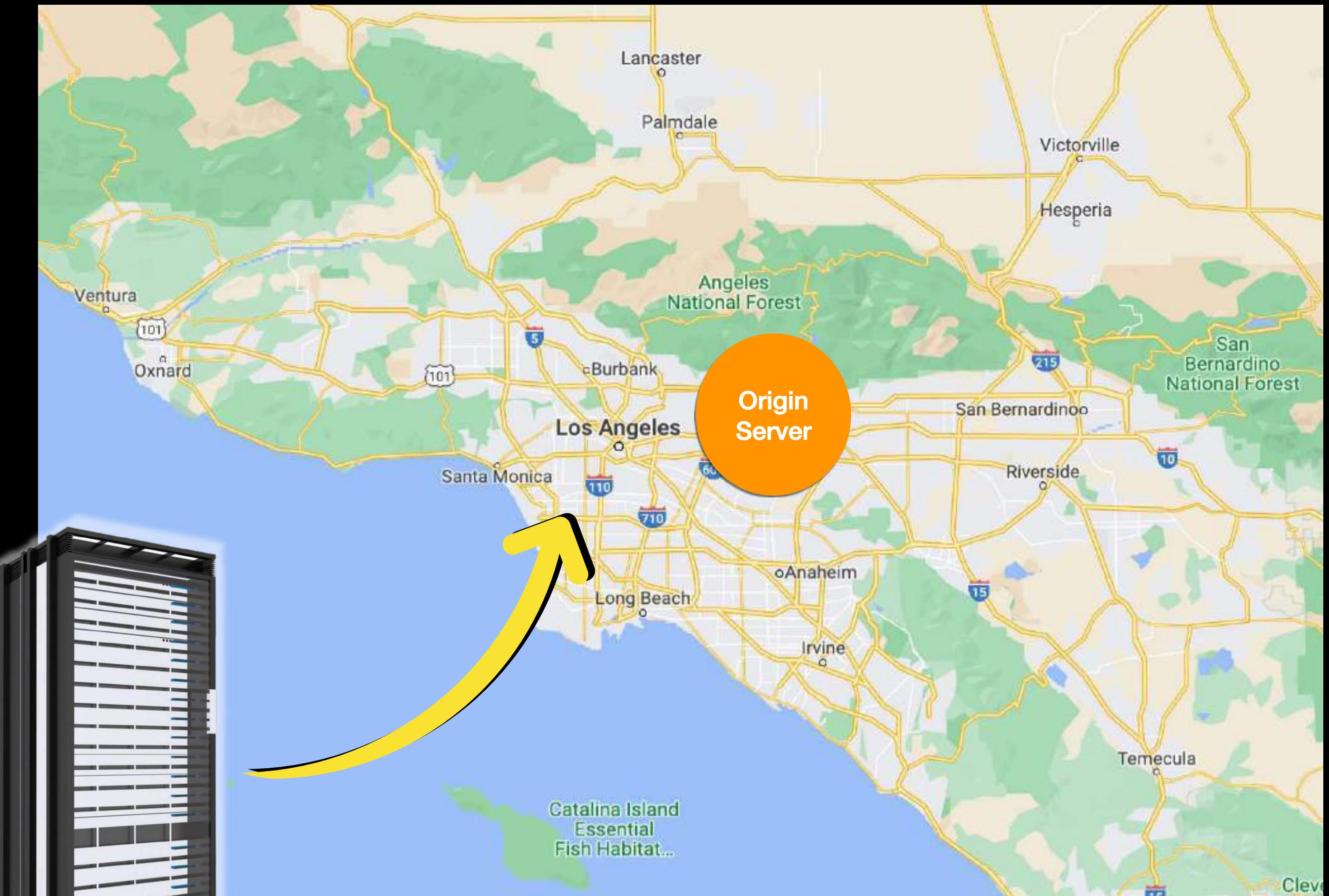


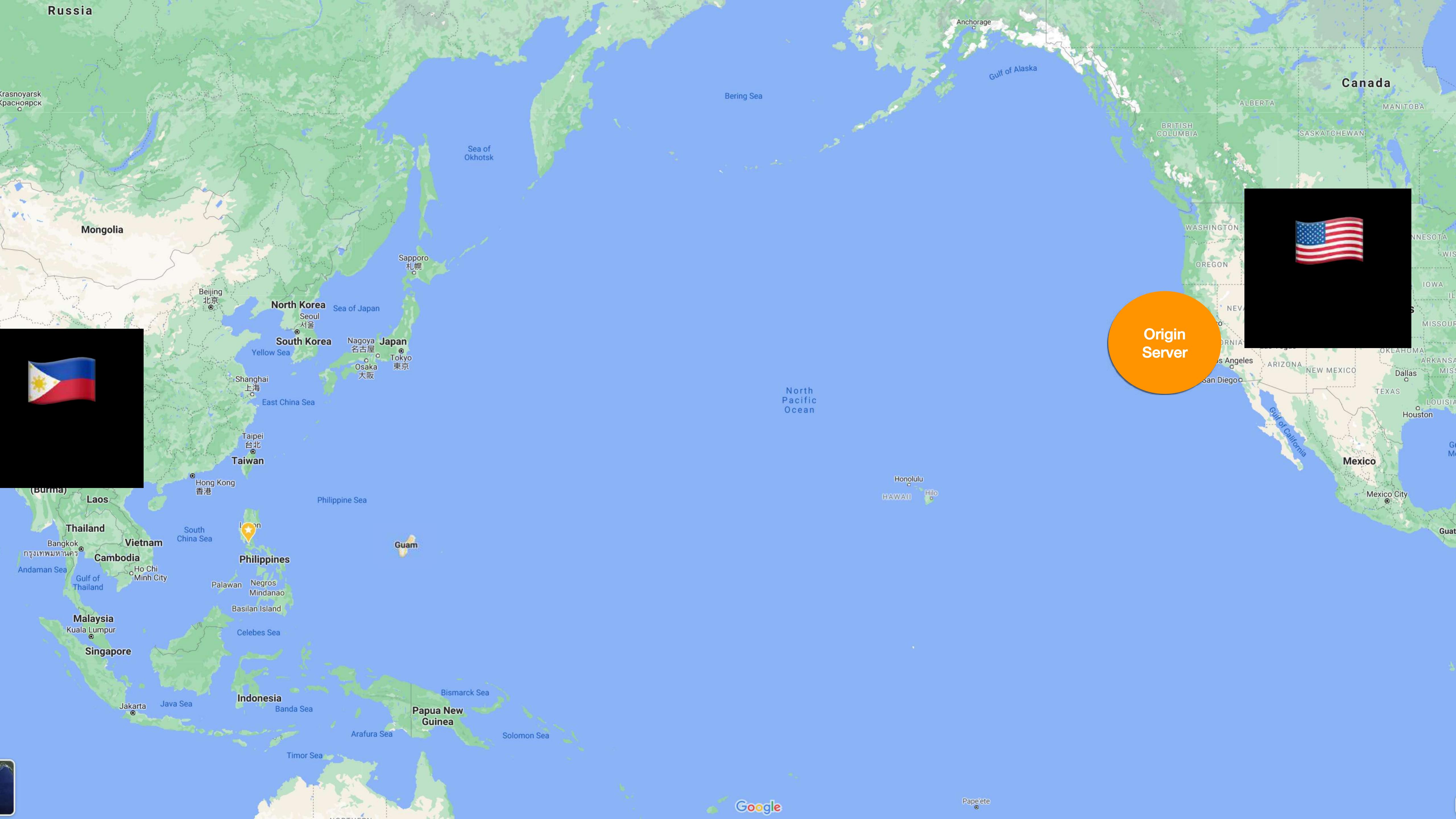
**CloudFront**



**Content  
Delivery  
Network**

# Content Delivery Network

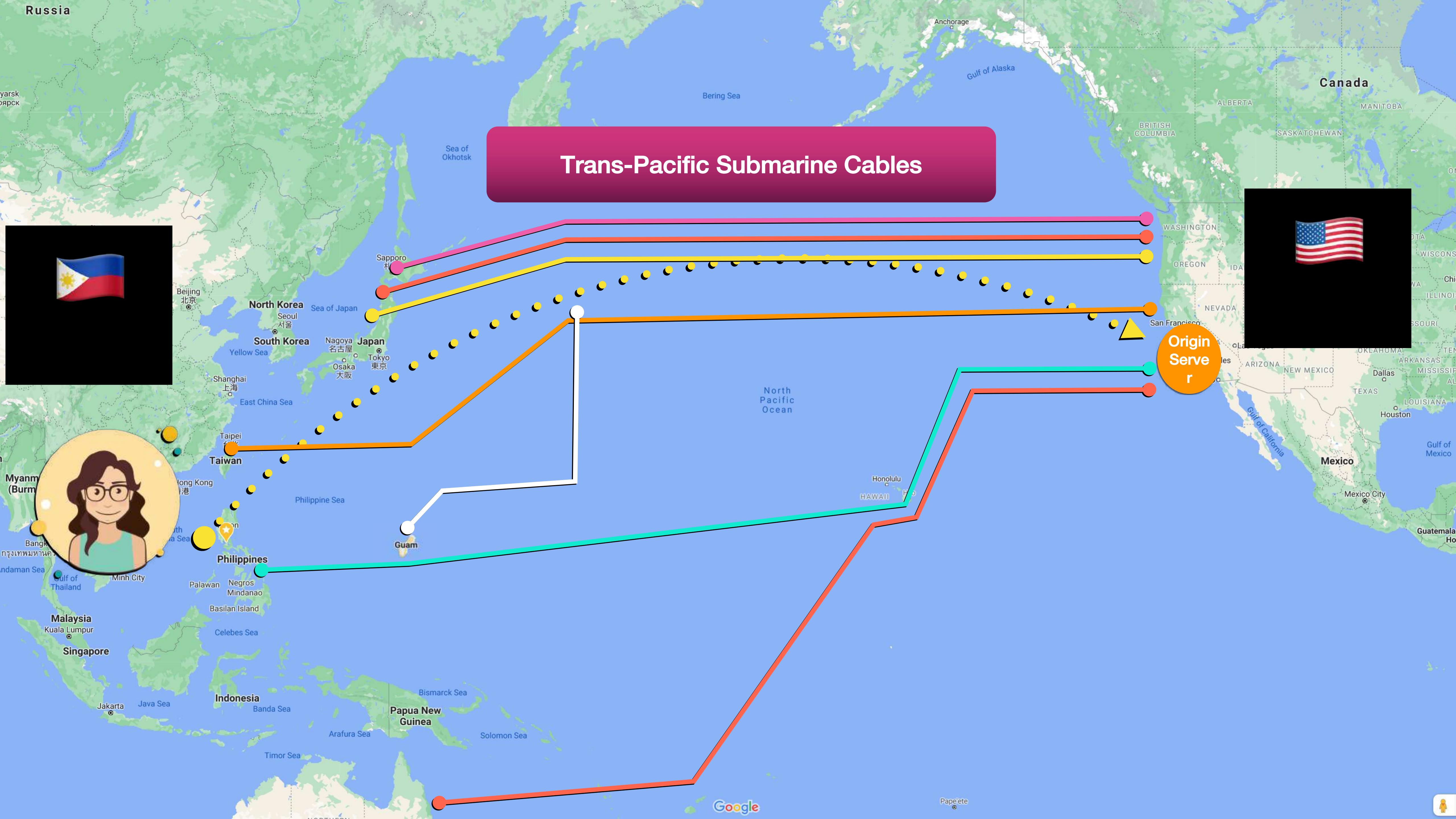


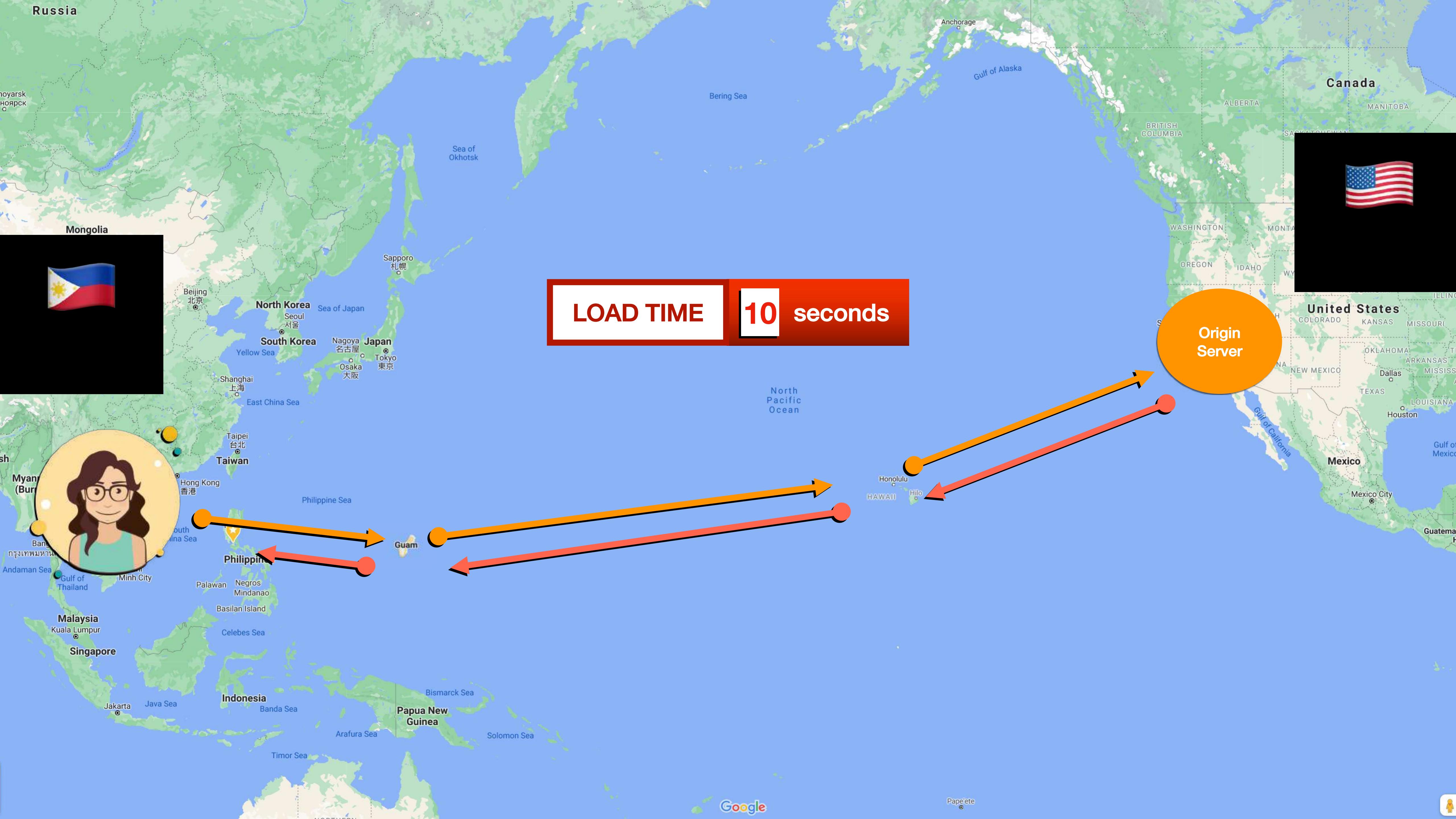


Origin  
Server



## Trans-Pacific Submarine Cables



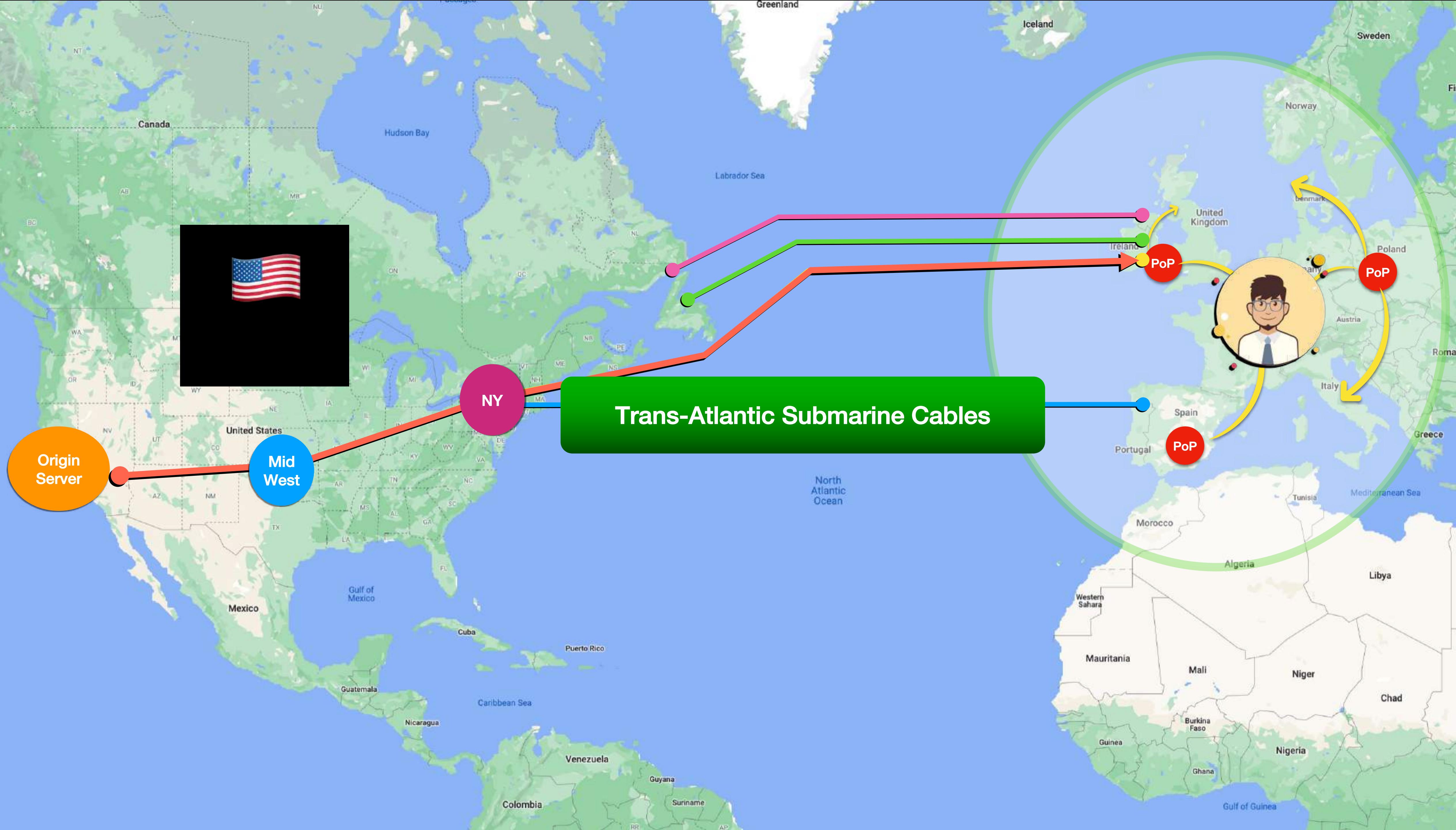


LOAD TIME

10 seconds

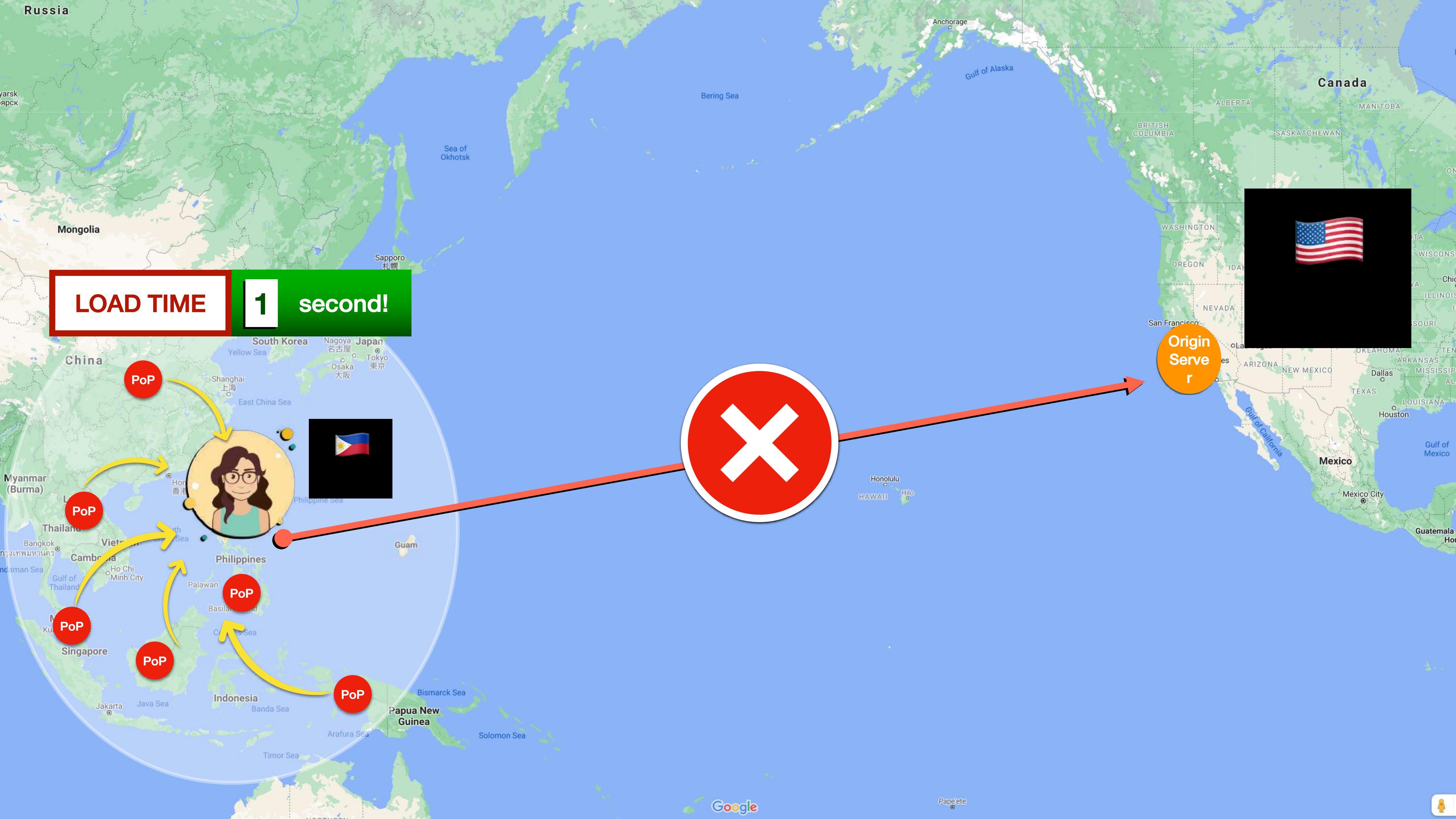
Origin  
Server

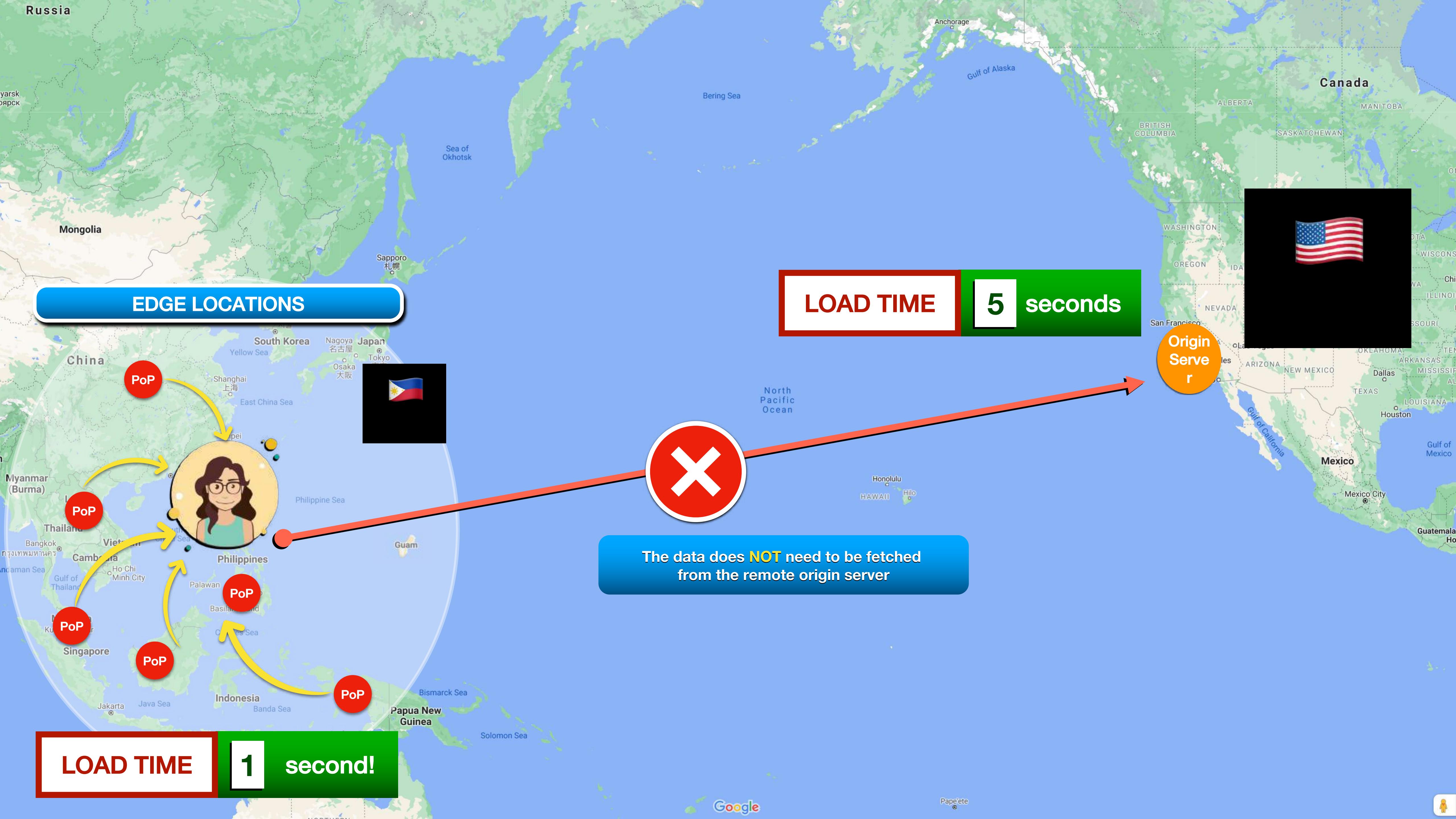


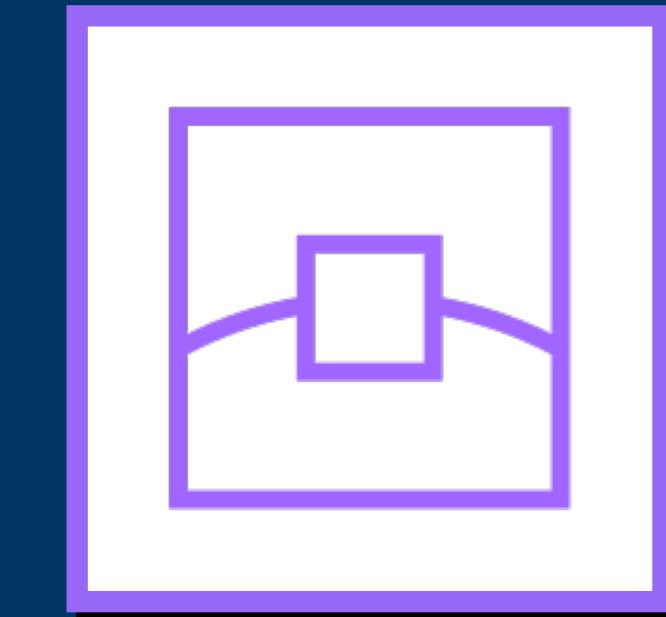
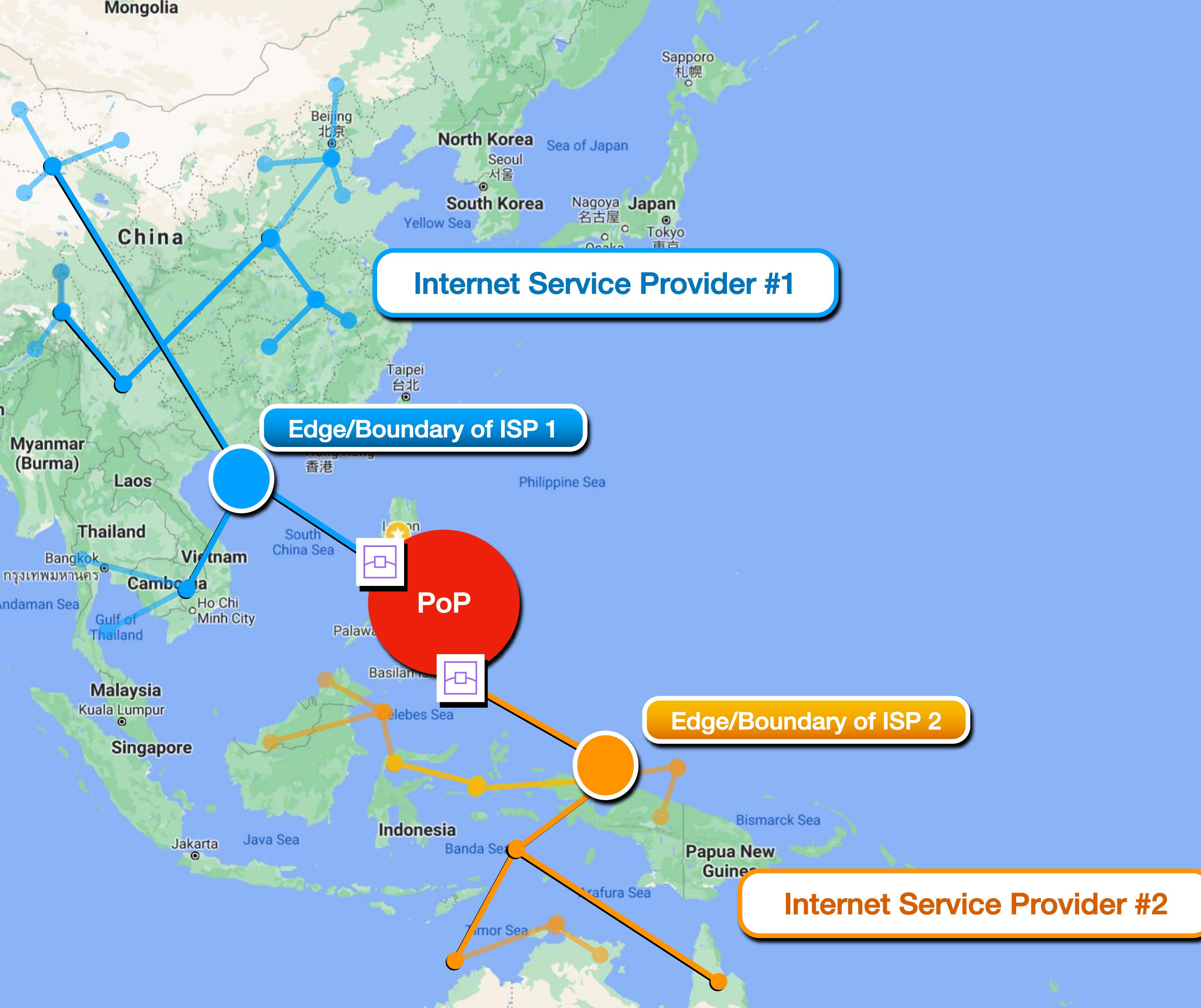


# LOAD TIME

# 1 second

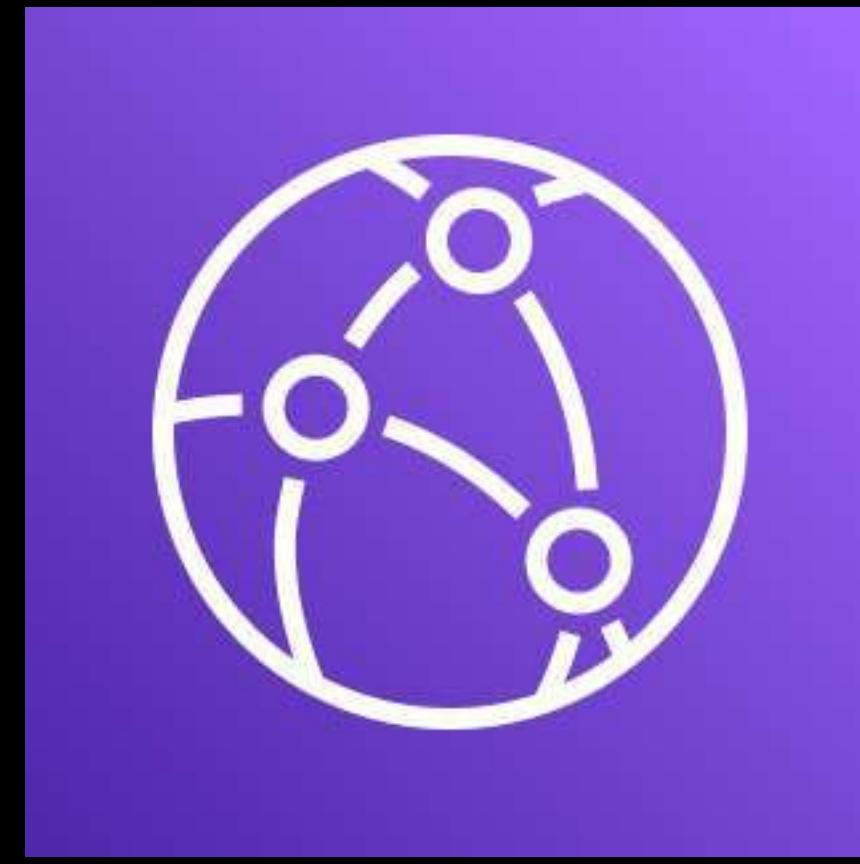




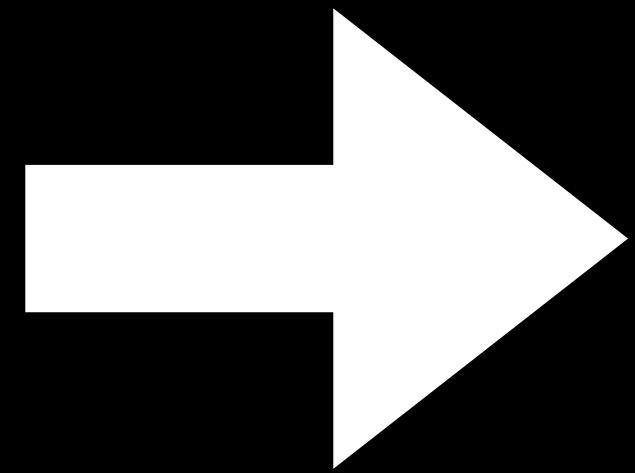


## Edge Location

- Refers to the 'edge' or the boundary of the network
- Connects the different networks of various Internet Service Providers (ISPs) or Telecommunications companies



**CloudFront**



**Content  
Delivery  
Network**



# CloudFront

**ORIGIN**

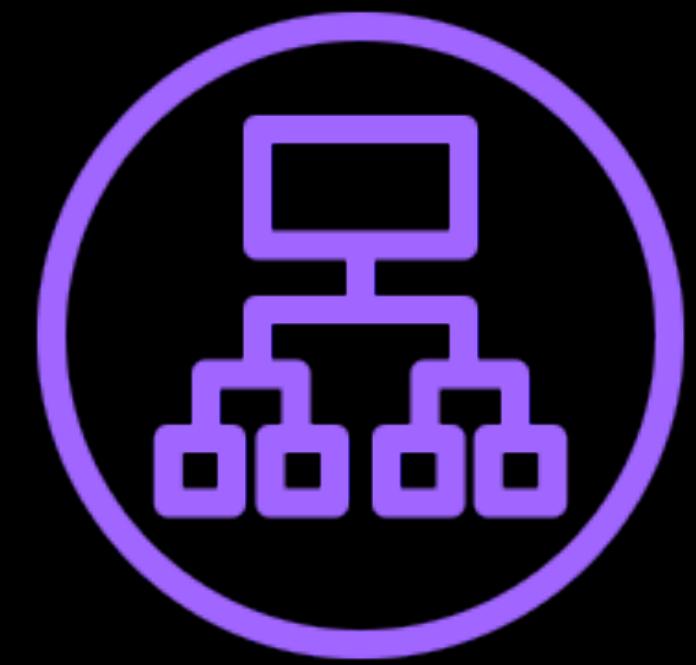
**DISTRIBUTION**

**VIEWER**

## ORIGIN



**Amazon S3 Bucket**



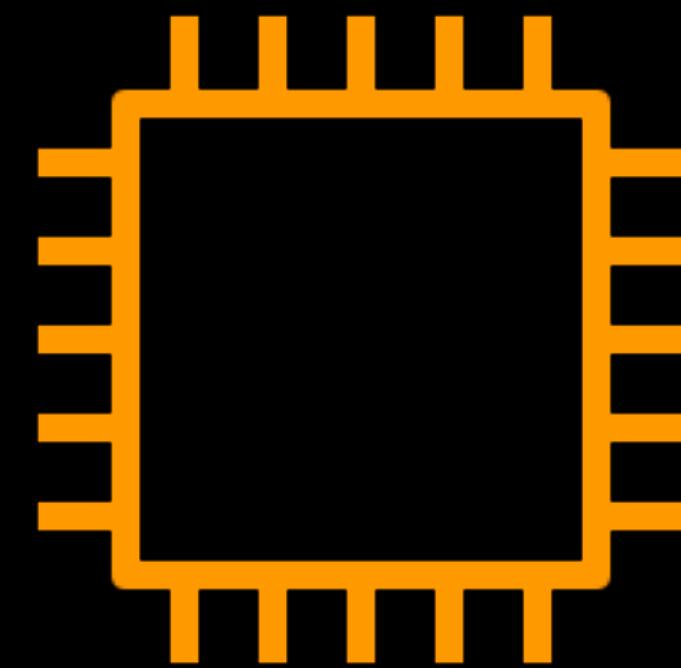
**Elastic Load Balancer**



**AWS Elemental  
MediaPackage Endpoint**



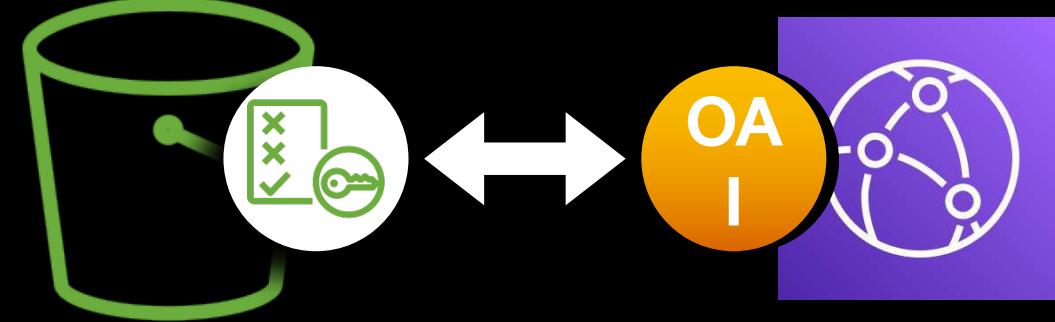
**AWS Elemental  
MediaStore Container**



**Amazon EC2 Instance or  
Your On-Premises Server**



# Amazon CloudFront Features



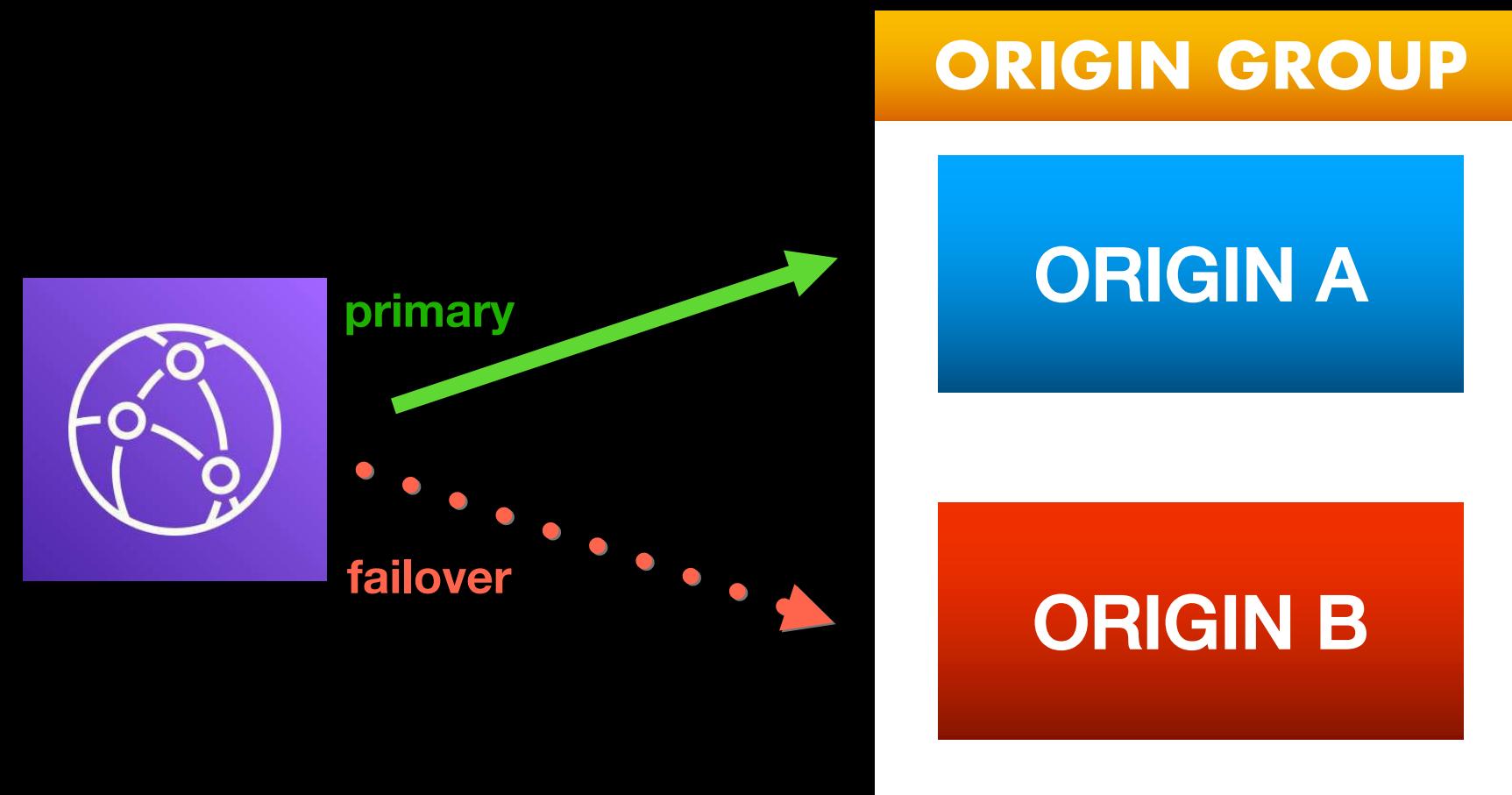
**ORIGIN ACCESS IDENTITY  
(OAI)**



**GEO-RESTRICTION**



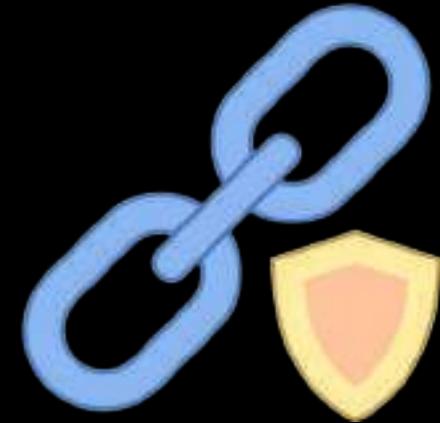
**Lambda@Edge  
and  
CloudFront Functions**



**ORIGIN GROUP and ORIGIN FAILOVER**



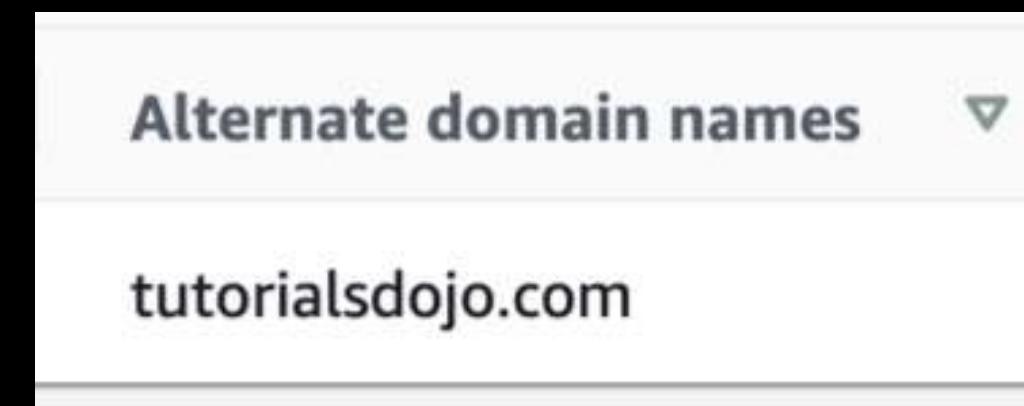
# Amazon CloudFront Features



**Signed URLs**



**Signed Cookies**



**Custom Domain Name and Custom SSL  
(SNI / Dedicated IP)**



## AWS WAF - CloudFront Integration

AWS WAF web ACL - *optional*

Choose the web ACL in AWS WAF to associate with this distribution.

*Choose web ACL*



# Amazon CloudFront Security Features

---

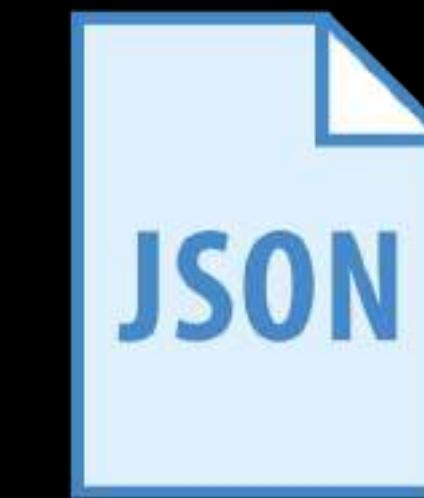
## Content

## Delivery

## Network

STATIC

DYNAMIC

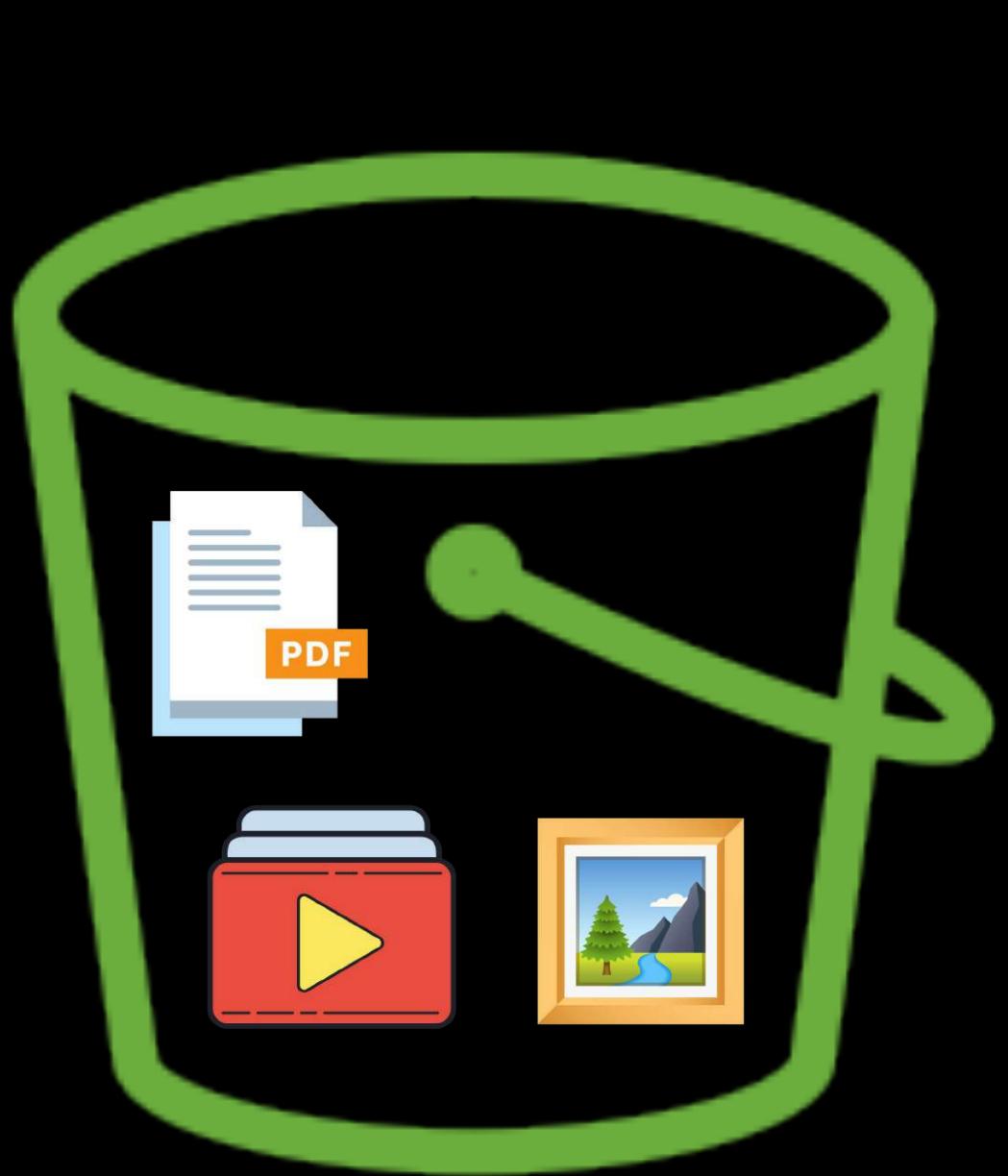


## Content

## Delivery

## Network

### AWS ORIGINS



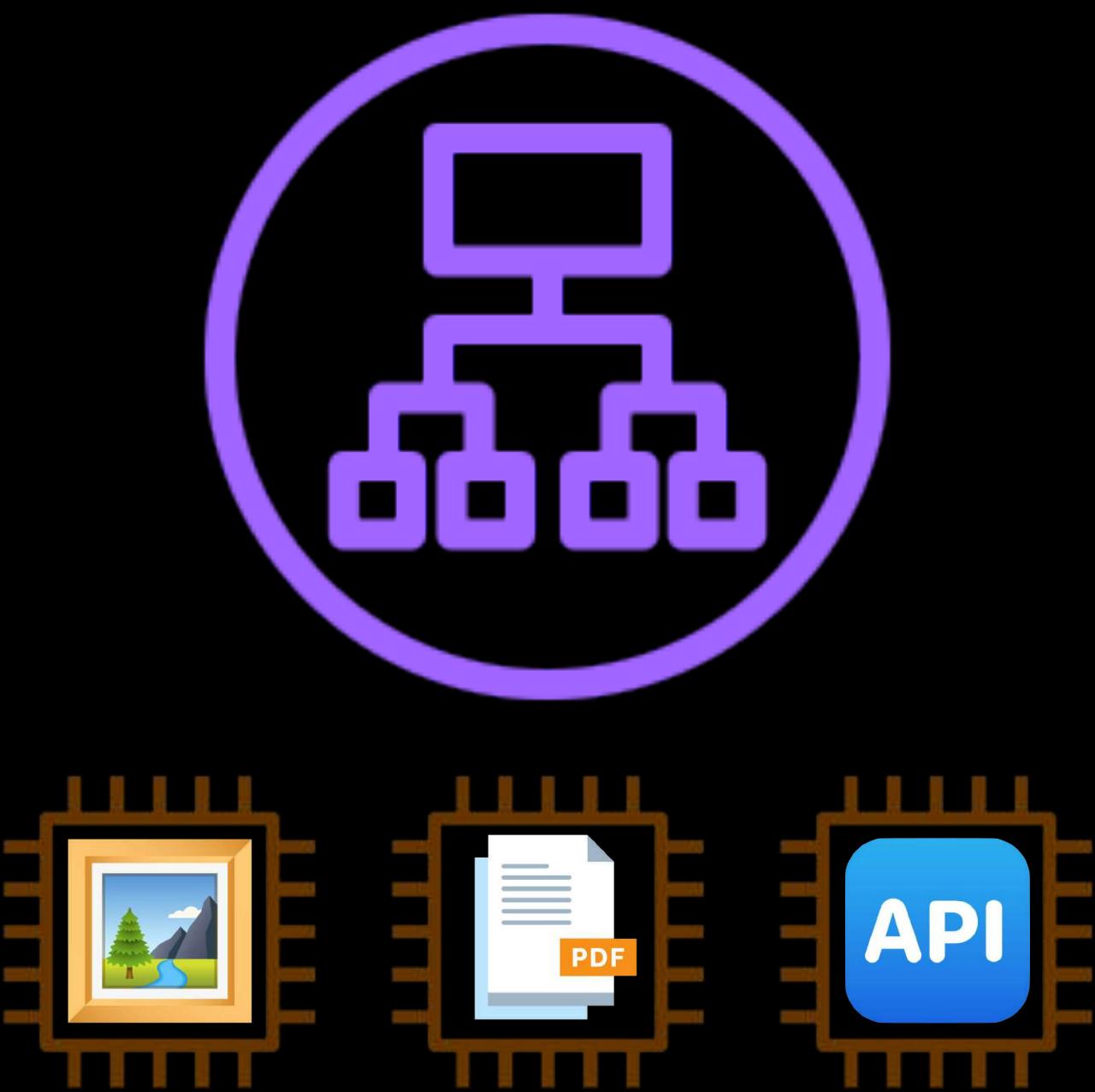
Amazon S3 Bucket



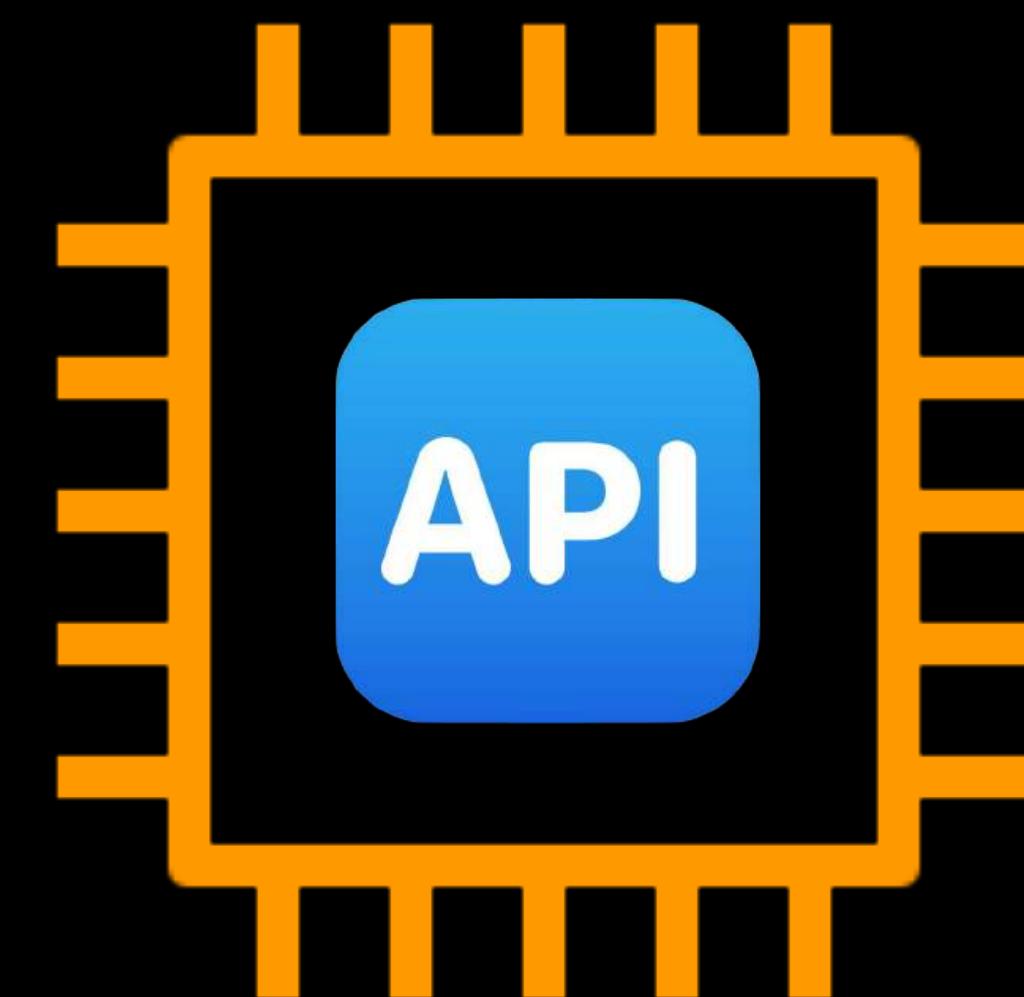
AWS Elemental  
MediaPackage



AWS Elemental  
MediaStore



Elastic Load Balancer



Amazon EC2 Instance or  
Your On-Premises Server

# Content

# Delivery

# Network

## Amazon S3 Origin

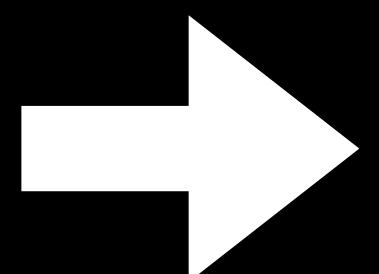


Origin Protocol Policy



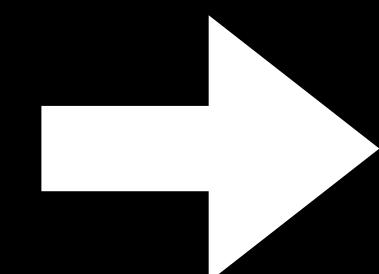
HTTP

HTTPS



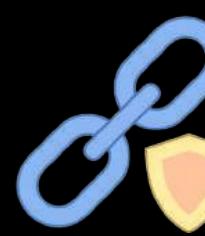
HTTP

HTTPS



Viewers

Viewer Protocol Policy



Signed URL



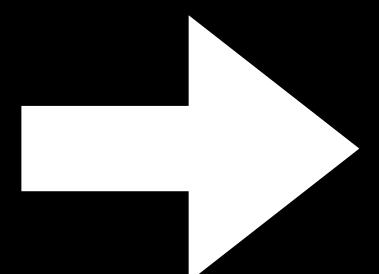
Signed Cookies

# Content Delivery Network

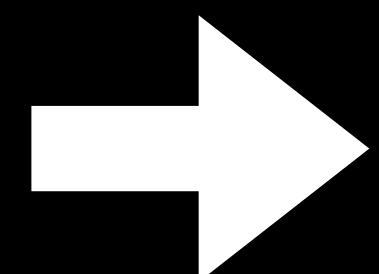
Amazon S3 Origin  
Origin Protocol Policy



HTTP  
HTTPS



HTTP  
HTTPS



Viewers  
Viewer Protocol Policy



Signed URL



Signed Cookies

**Content**

**Delivery**

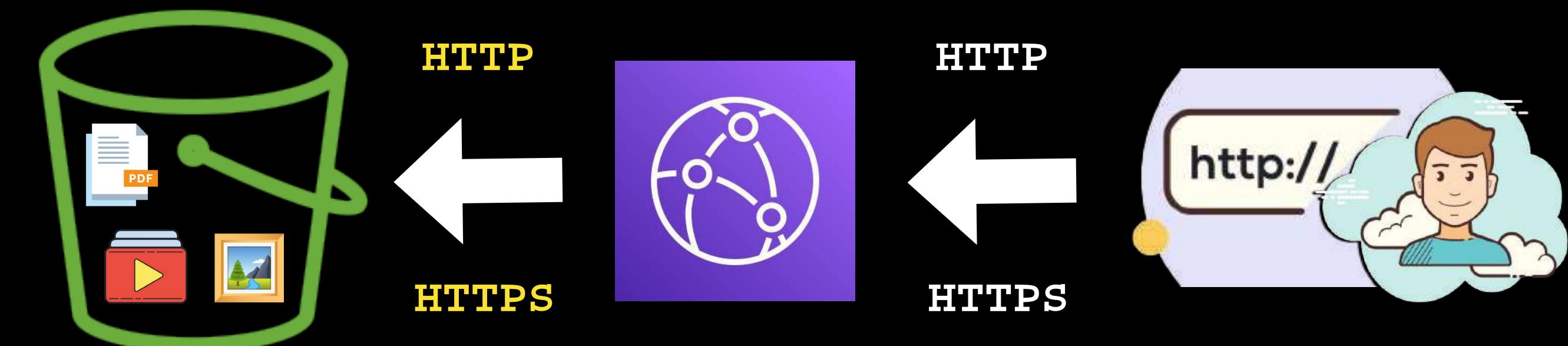
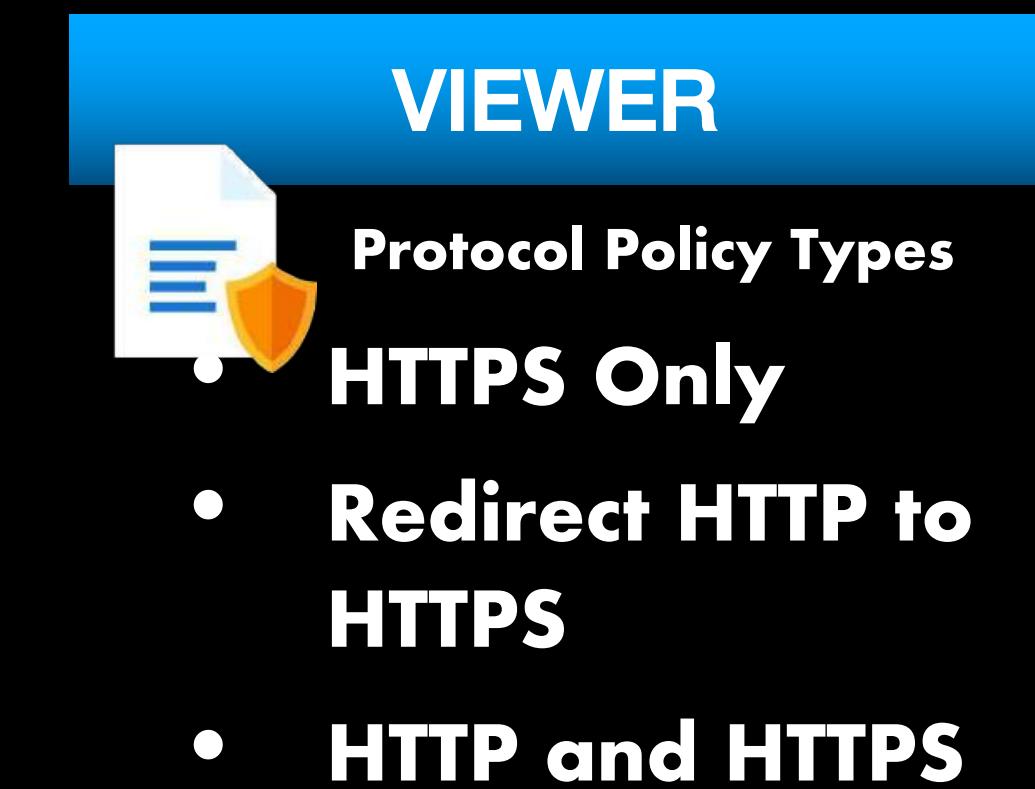
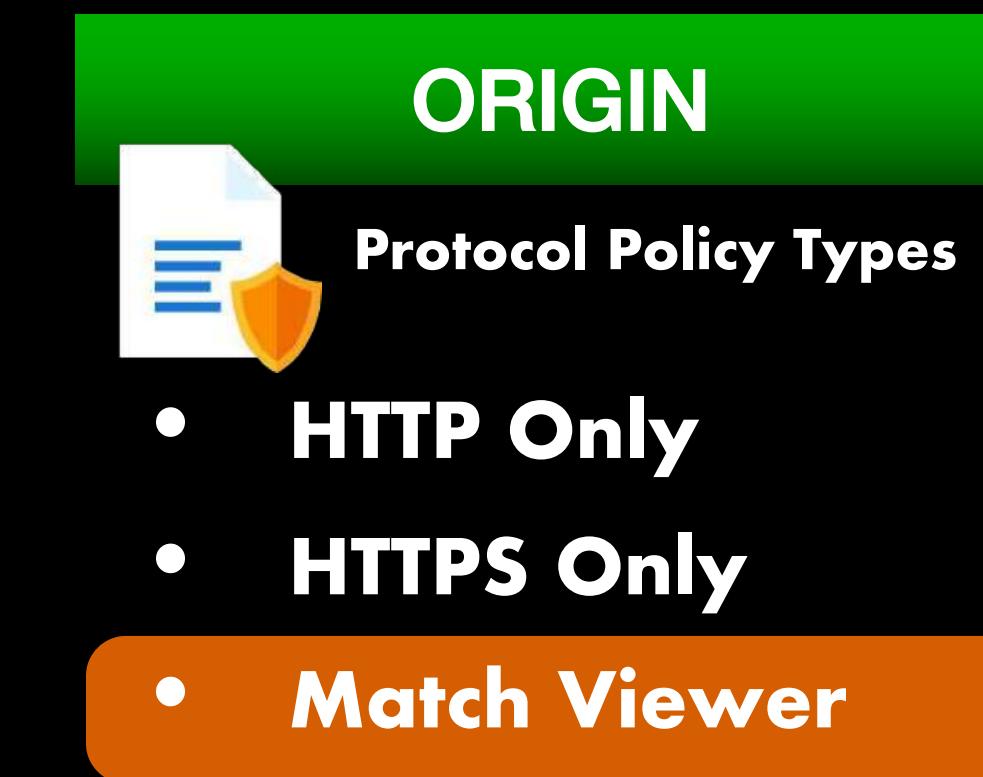
**Network**

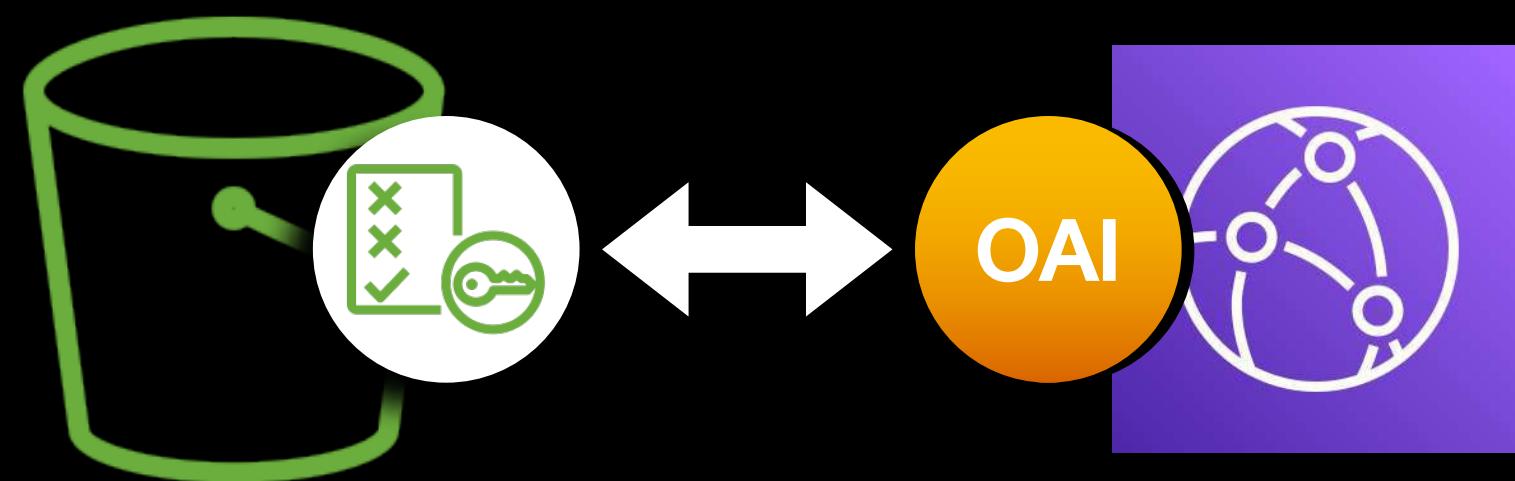
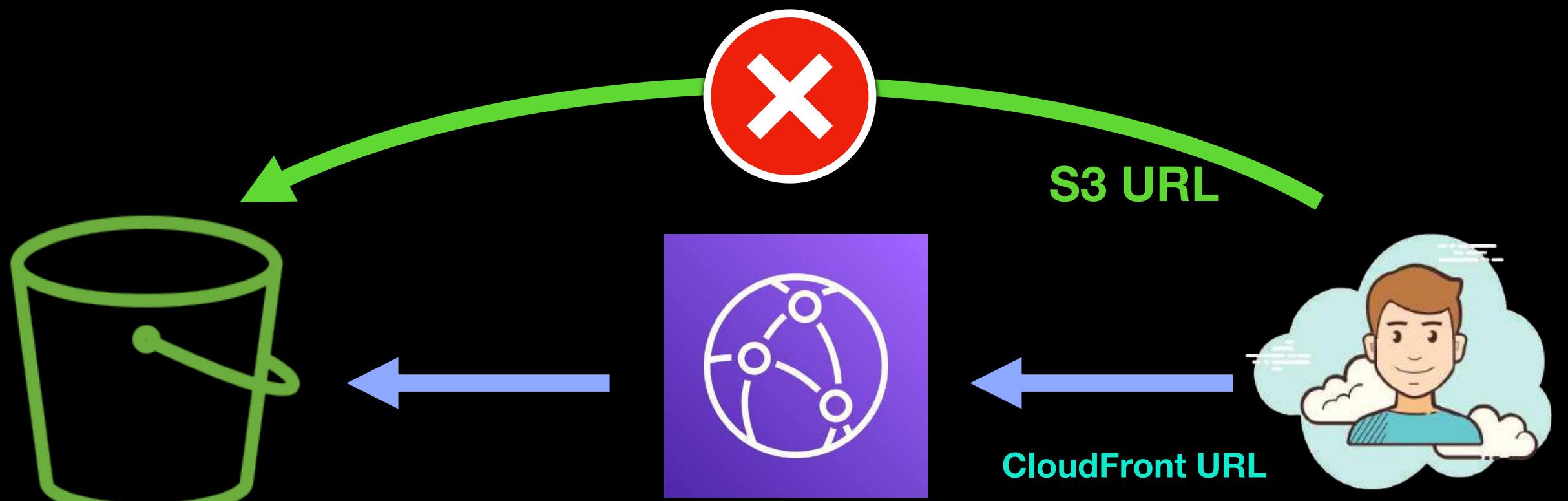


- Specifies the allowed **protocols** for the Origin and the Viewer (end users)
- Configures the CloudFront distribution to use **HTTP, HTTPS or both**



## PROTOCOL POLICY





## ORIGIN ACCESS IDENTITY (OAI)

- **Primarily used for CloudFront distributions with an Amazon S3 bucket as the origin**
- **Restricts access to the content that you serve from your S3 bucket**
- **Works like an IAM User which you can associate to the Origin or Origin Group of your CloudFront distribution**
- **After OAI has been created, the Amazon S3 bucket policy must be configured too**

## FIELD-LEVEL ENCRYPTION

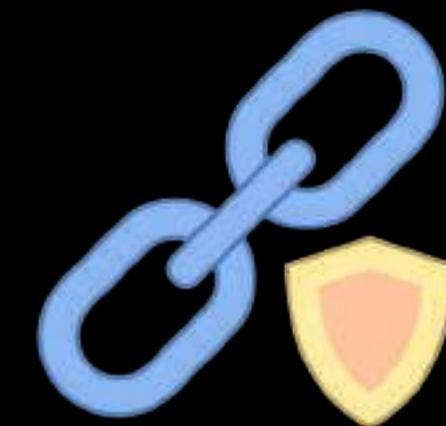


- Allows you to **encrypt the specific data fields**
- Protects sensitive information in your origin and the data being sent by your customers
- Suitable for securing Credit Card numbers, Personal Health Information (PHI) and Personally Identifiable Information (PII)
- Encrypts the sensitive fields using a **public key**
- Provides you with a **private key** that can be used to decrypt the protected fields

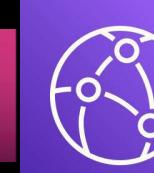


## SIGNED URLs & SIGNED COOKIES

- Primarily used for **distributing private content over the Internet**
- **Restrict access to your confidential or private data to authorized users only**



## SIGNED URLs



CloudFront Distribution with Custom Domain Name

<https://tutorialsdojo.com/report.pdf>

?Expires=13570344005

&Signature=nitfHRCrtziwO2HwPfWw~yYDhUF5EwRunQA...

&Key-Pair-Id=K2JCJMDEHXQW5F

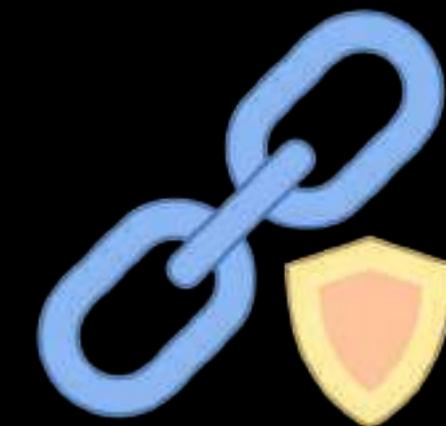


## SIGNED COOKIES

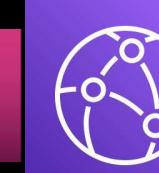
The screenshot shows the Chrome DevTools Application tab. A green box highlights the 'Set-Cookie' button in the main panel, which is labeled 'HEADER'. Another green box highlights the 'Cookies' section in the Storage sidebar. The Cookies list shows a single entry for the domain <https://portal.tutorialsdojo.com>. To the right, a list of cookie names is shown, also highlighted with a green box: \_ga, signed\_user\_cookie\_1, and user\_id. Below this list are buttons for 'Cookie Value' and 'Show URL decoded', and the text 'Tutorials Dojo Private Content'.

Name
_ga
signed_user_cookie_1
user_id

Cookie Value  Show URL decoded  
Tutorials Dojo Private Content



## SIGNED URLs



CloudFront Distribution with Custom Domain Name

<https://tutorialsdojo.com/report.pdf>

?Expires=13570344005

&Signature=nitfHRCrtziwO2HwPfWw~yYDhUF5EwRunQA...

&Key-Pair-Id=K2JCJMDEHXQW5F



## SIGNED COOKIES

The screenshot shows the Chrome DevTools Application tab. A green box highlights the 'Set-Cookie' button in the main panel, which is labeled 'HEADER'. Another green box highlights the 'Cookies' section in the Storage sidebar. The Cookies list shows a single entry for the domain <https://portal.tutorialsdojo.com>. To the right, a list of cookie names is shown, also highlighted with a green box: \_ga, signed\_user\_cookie\_1, and user\_id. Below this list are buttons for 'Cookie Value' and 'Show URL decoded', and the text 'Tutorials Dojo Private Content'.

Name
_ga
signed_user_cookie_1
user_id

Cookie Value  Show URL decoded  
Tutorials Dojo Private Content



## GEO-RESTRICTION

- **Restricts access to your content based on the specific country (geographic location) of your users**
- **Allows you to select the specific countries where you want to deliver your content and which countries to block**

Restriction type

No restrictions

Allow list

Block list

Countries

Select countries ▾

Australia X Canada X India X Philippines X

United Kingdom X United States X



## ALTERNATE DOMAIN NAME & SSL CERTIFICATE

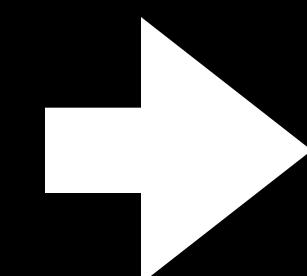
CloudFront > Distributions

Distributions (5) [Info](#)

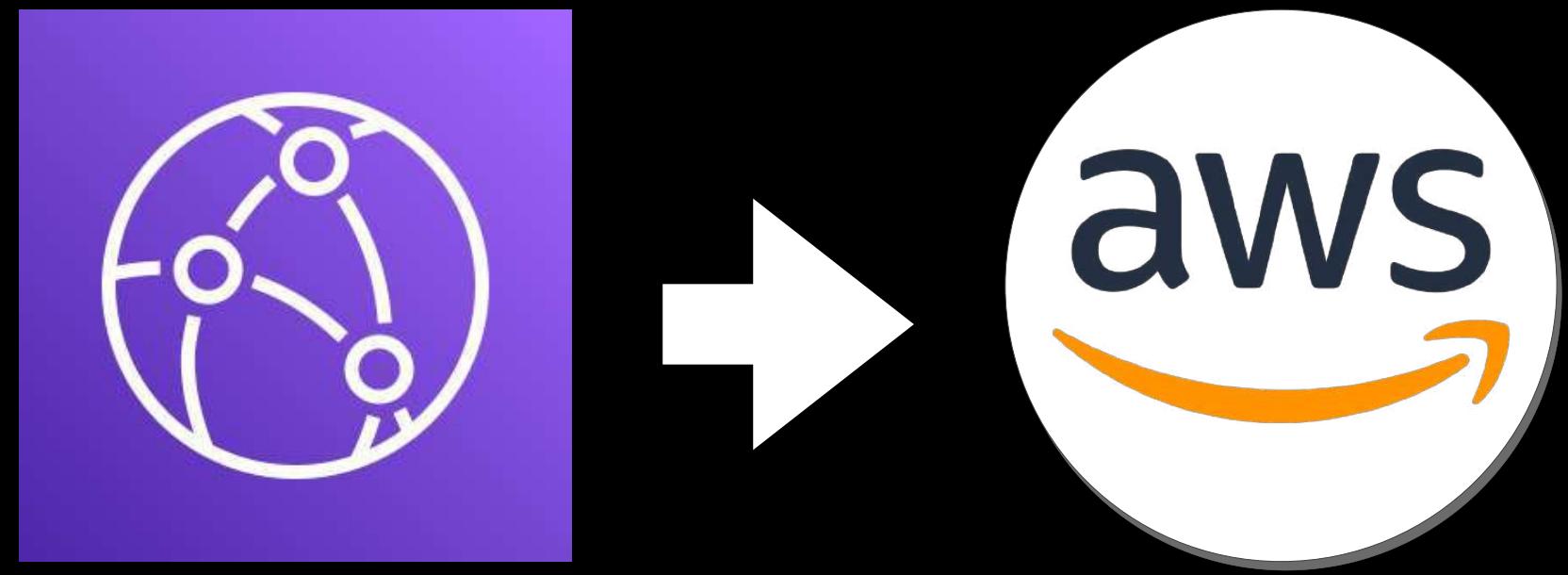
Q SQ X 1 match

ID	Domain name	Alternate domain names
ERMIT4M4N1L4SQ	d123456abcdef.cloudfront.net	tutorialsdojo.com

- **SNI (Server Name Indication)**
- **Dedicated IP address**



AWS Certificate Manager



## INTEGRATIONS TO OTHER AWS SERVICES



### AWS WAF - CloudFront Integration

AWS WAF web ACL - *optional*  
Choose the web ACL in AWS WAF to associate with this distribution.

Choose web ACL ▾



### AWS Shield



- Different from the **Origin Shield** feature

Enable Origin Shield [Info](#)  
Origin Shield is an additional caching layer that can help reduce the load on your origin and help protect its availability.

No  
 Yes

**HIGH AVAILABILITY**

**VS**

**FAULT TOLERANCE**

Are these two **exactly the same?**

**HIGH AVAILABILITY**

**FAULT TOLERANCE**

# HIGH AVAILABILITY



# FAULT TOLERANCE

## SAME OBJECTIVE

Both of them aims to ensure the application runs all the time without any system degradation, data loss or outage

## DESIGN

## SINGLE SERVER ARCHITECTURE

HIGH  
AVAILABILITY

FAULT  
TOLERANCE

UPTIME

LOW

99.99%

100%

REDUNDANCY

NONE

HAS AT LEAST ONE  
REDUNDANT RESOURCE  
FOR FAILOVER

HAS A LOT  
OF REDUNDANT  
RESOURCES

COST

LOW

MODERATE

HIGH

# HIGH AVAILABILITY

99.99% UPTIME

HAS AT LEAST ONE  
REDUNDANT RESOURCE  
FOR FAILOVER

MODERATE COST

# FAULT TOLERANCE

100% UPTIME

HAS A LOT  
OF REDUNDANT  
RESOURCES

MORE RESOURCES  
CAUSES



HIGH COST



# RTO

Recovery **Time** Objective

VS



# RPO

Recovery **Point** Objective

# DISASTER RECOVERY OBJECTIVES



## RTO

Recovery **Time** Objective



## RPO

Recovery **Point** Objective



9:00 AM

10:00 AM

11:00 AM

12:00 NN

1:00 PM

2:00 PM

3:00 PM

4:00 PM

5:00 PM

ALL DATA  
**BEFORE 11 AM**  
MUST BE  
RECOVERABLE

D I S A S T E R

SERVICE RESTORED



**RPO**

1 HOUR

Recovery Point Objective

11 AM - 12 NN

ACCEPTABLE  
DATA LOSS



**RTO**

3 HOURS

Recovery Time Objective



12:00 NN

1:00 PM

2:00 PM

3:00 PM

4:00 PM

5:00 PM

6:00 PM

7:00 PM

ALL DATA  
**BEFORE 2 PM**  
MUST BE  
RECOVERABLE

D I S A S T E R

SERVICE RESTORED

3:00 PM

1 HOUR

=  
02:00 PM



RPO

1 HOUR

Recovery Point Objective

ACCEPTABLE DATA LOSS  
2 PM - 3 PM



RTO

2 HOURS

Recovery Time Objective

3:00 PM

+  
2 HOURS

=  
05:00 PM



**Network Access Control List  
(Network ACL)**

**VS**



**Security Group**



## Network ACL



## Security Group

- **Created by default when you launch a new VPC and on your default VPC**
- **Acts as a virtual firewall that protects your AWS resources from unauthorized traffic**
- **Inbound & Outbound rules can be set to have one IP address or a CIDR range as a source**
- **Allows you to control the incoming and outgoing traffic to and from your network**



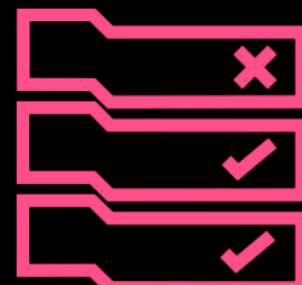
## Network ACL



## Security Group

Rule #	Type	Protocol	Port range	Source	Allow/Deny
10	HTTP	TCP	80	172.31.1.2/32	DENY
20	HTTPS	TCP	443	192.0.2.0/24	DENY
30	All IPv4 traffic	All	All	0.0.0.0/0	ALLOW
40	All IPv6 traffic	All	All	::/0	ALLOW

### Ephemeral Ports



- 1024 – 65535
- 32768 – 61000
- 49152 – 65535

### Outbound Rules

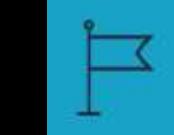
# STATE

STATEFUL

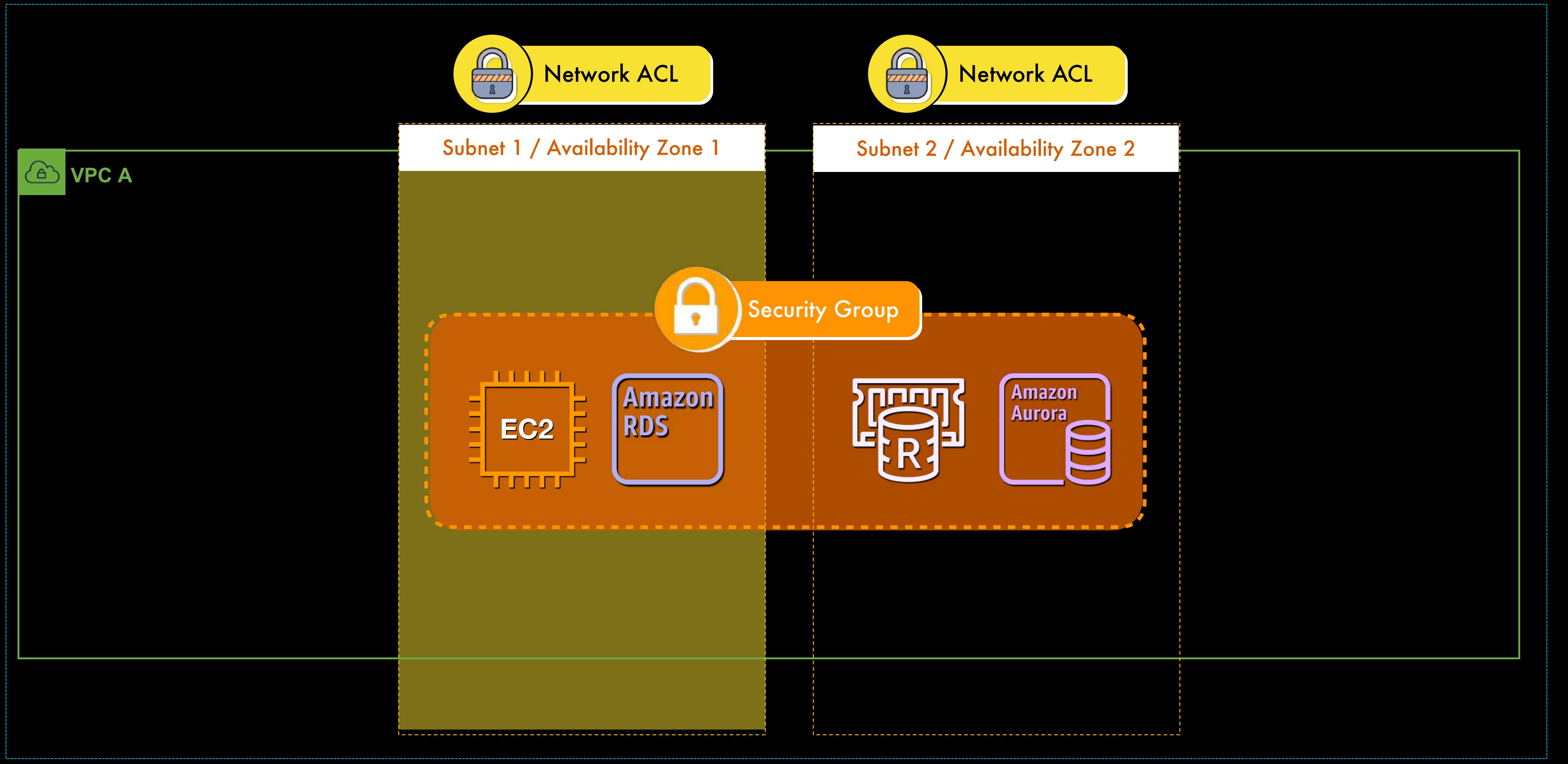
STATELESS



AWS Cloud



N. Virginia Region





## Network ACL

- Can explicitly DENY traffic

The screenshot shows the 'Edit inbound rules' page for a specific Network ACL. It displays two sets of rules:

- Inbound rule 1:** Rule number 100, Type: All traffic, Protocol: All. The 'Allow/Deny' dropdown menu is open, showing three options: Allow (selected), Allow, and Deny. This set is highlighted with a yellow box.
- Inbound rule 2:** Rule number +, Type: All traffic, Protocol: All. The 'Allow/Deny' dropdown menu is set to Deny.



## Security Group

- Cannot explicitly DENY traffic

The screenshot shows the 'Edit inbound rules' page for a specific Security Group. A blue banner at the top right states 'WHITELISTING only!'. It lists two inbound rules:

- Inbound rule 1:** Security group rule ID: sg-002dd084man1l4, Type: HTTP, Protocol: TCP. Port range: 80. Source type: Custom, Source: ::/0. This rule is highlighted with a red box.
- Inbound rule 2:** Security group rule ID: sgr-052960d0e8b1332db, Type: Custom ICMP - IPv6, Protocol: IPv6 ICMP. Port range: All. Source type: Custom, Source: ::/0.





Network ACL

STATELESS



Security Group

STATEFUL

- Does not track the status of the request
- The inbound traffic that has already been permitted before is still subject to the rules for the outbound traffic, and vice versa
- Provides a more fine-grained control to configure both the inbound and outbound rules of your Network ACL

- Tracks all the status of the incoming requests
- If a traffic is a response to a particular request, then it will be allowed automatically regardless of any rules in your Outbound Rules
- It is aware if the outgoing traffic is:
  - Initiated from the EC2 instance itself
  - A response to the request that was initiated externally
- Its Outbound Rule can filter:
  - An API call initiated by an application hosted in the EC2 instance
  - A scheduled OS Patch that is initiated by the EC2 instance which automatically fetches updates from a designated repository



## Network ACL

- Each rule has a corresponding **rule number**
- **Evaluates the rules in order**, starting with the lowest numbered rule

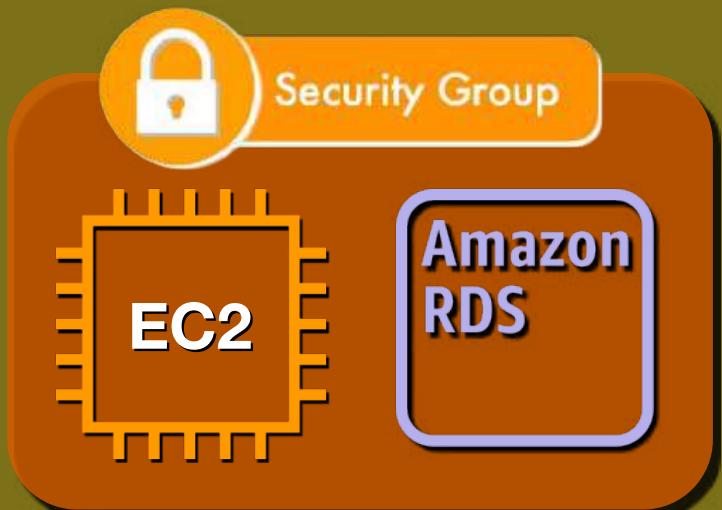


## Security Group

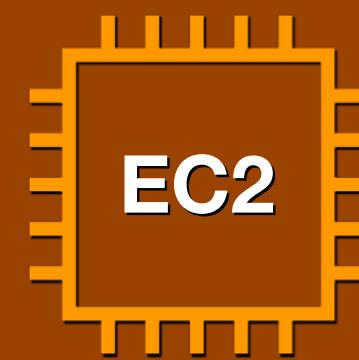
- No rule number
- Evaluates ALL of the rules **at the same time** (*no order of precedence*)



## Network ACL



## Security Group



- Applies the rules to **all EC2 instances and other AWS resources in the subnets** that it's associated with

- Applies the rules to a **single EC2 instance only or to a group of AWS resources** where it is associated with



## Network ACL



1024 – 65535



32768 – 61000



49152 – 65535

### Outbound Rules



## Security Group

- Does NOT use

### Ephemeral Ports



## Network ACL

Inbound rule

Rule number <a href="#">Info</a>	Type <a href="#">Info</a>	Protocol <a href="#">Info</a>
100	All traffic	All
Port range <a href="#">Info</a>	Source <a href="#">Info</a>	Allow/Deny <a href="#">Info</a>
All	0.0.0.0/0	Allow
<a href="#">Remove</a>		

Inbound rule

Rule number <a href="#">Info</a>	Type <a href="#">Info</a>	Protocol <a href="#">Info</a>
101	Custom TCP	TCP (6)
Port range <a href="#">Info</a>	Source <a href="#">Info</a>	Allow/Deny <a href="#">Info</a>
4000	110.238.109.37/32	Deny
<a href="#">Remove</a>		



## Security Group

Source <a href="#">Info</a>	Description - optional <a href="#">Info</a>
My IP	Jon Bonso's IP Address
6.12.18.98/32	<a href="#">Delete</a>
Custom	CIDR Block
192.0.0.0/24	<a href="#">Delete</a>
Custom	Prefix List
pl-02cd2c6b	<a href="#">Delete</a>
Custom	Another Security Group
sg-e1023de3	<a href="#">Delete</a>



A N O T H E R



Security Group



## AWS Storage Gateway Types Comparison



**File  
Gateway**

**VS**

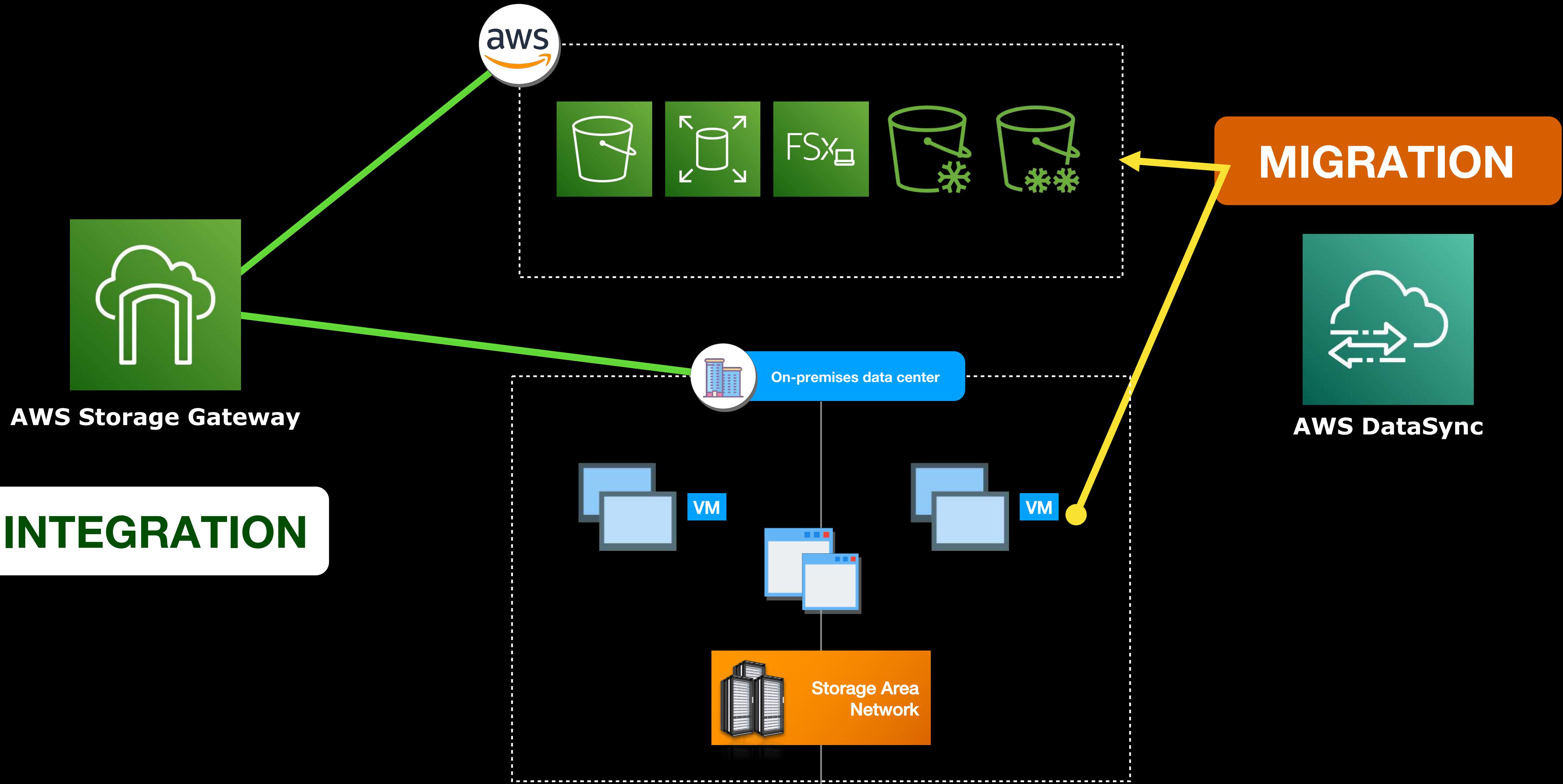


**Volume  
Gateway**

**VS**



**Tape  
Gateway**





\* **File storage**

## File Gateway

### NFS file share

- Also known as **Amazon S3 File Gateway**



- Stores data in: **Amazon S3**

- Provides a **local cache** for low-latency access to your most recently used data



\* **Block storage**

## Volume Gateway



### CACHED

- Uses **Amazon S3** as the primary storage
- Stores a **subset** of frequently accessed data locally



\* **Tape storage**

## Tape Gateway

### GLACIER POOL



### Amazon S3 Glacier



### STORED

- Retains the **entire** dataset in your on-premises data center
- Asynchronously backs up your data to **Amazon S3**



### DEEP ARCHIVE POOL



### Amazon S3 Glacier Deep Archive



**Amazon FSx for Windows File Server**

- Also known as **Amazon FSx File Gateway**

- Stores data in: **Amazon FSx for Windows File Server**

- Provides a **low-latency on-premises** access to **Windows SMB file shares** of the Amazon FSx for Windows File Server service in AWS



## File Gateway



## Volume Gateway

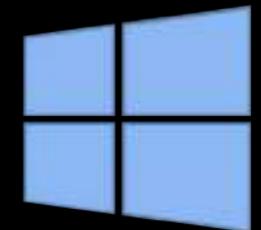


## Tape Gateway



### Amazon FSx for Windows File Server

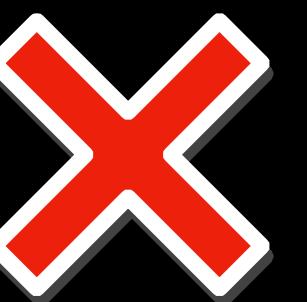
Can be integrated with:



Microsoft  
**Active Directory**



**AWS Managed Microsoft AD**



No Active Directory Support



No Active Directory Support



**File Gateway**



**Volume Gateway**



**Tape Gateway**

**NFS**

**iSCSI**

**SMB**

**VTL**



**File Gateway**



**Volume Gateway**



**Tape Gateway**



An image of an actual AWS Storage Gateway Hardware Appliance

**INTEGRATION**



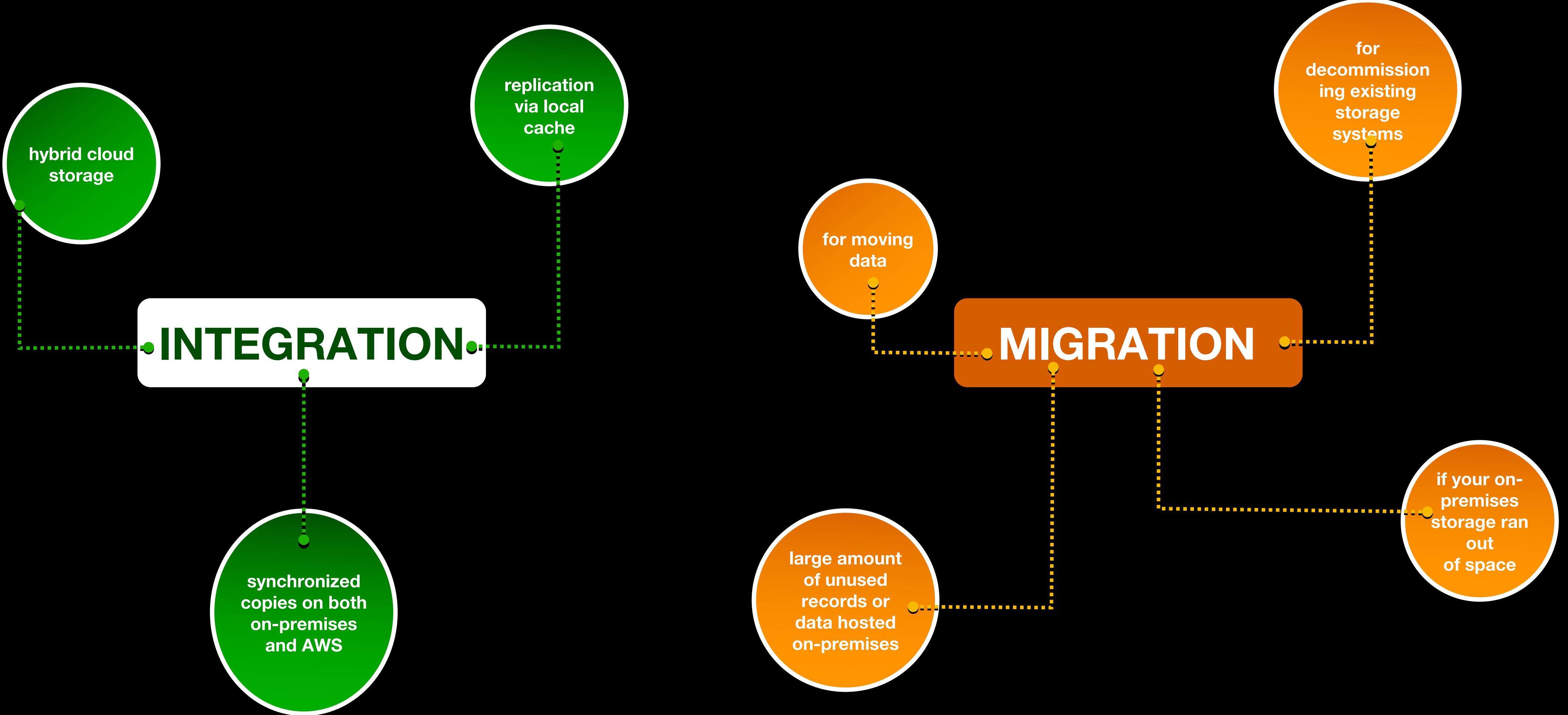
**AWS Storage Gateway**

**MIGRATION**



**AWS DataSync**

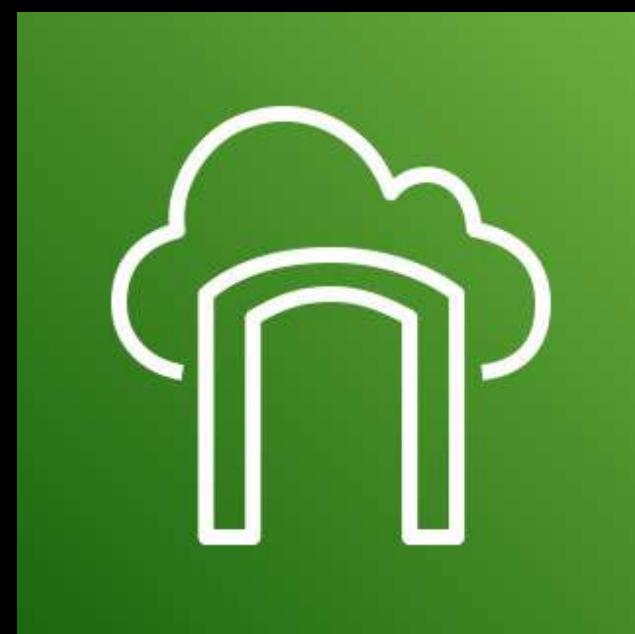
**VS**



**REPLICATE DATA**



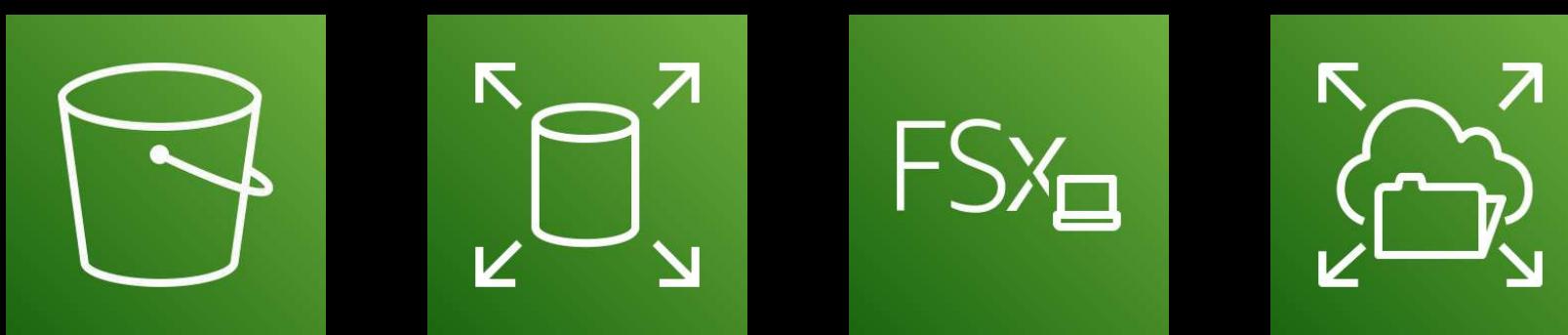
**INTEGRATION**



**AWS Storage Gateway**



*On-premises data will  
still be actively used*

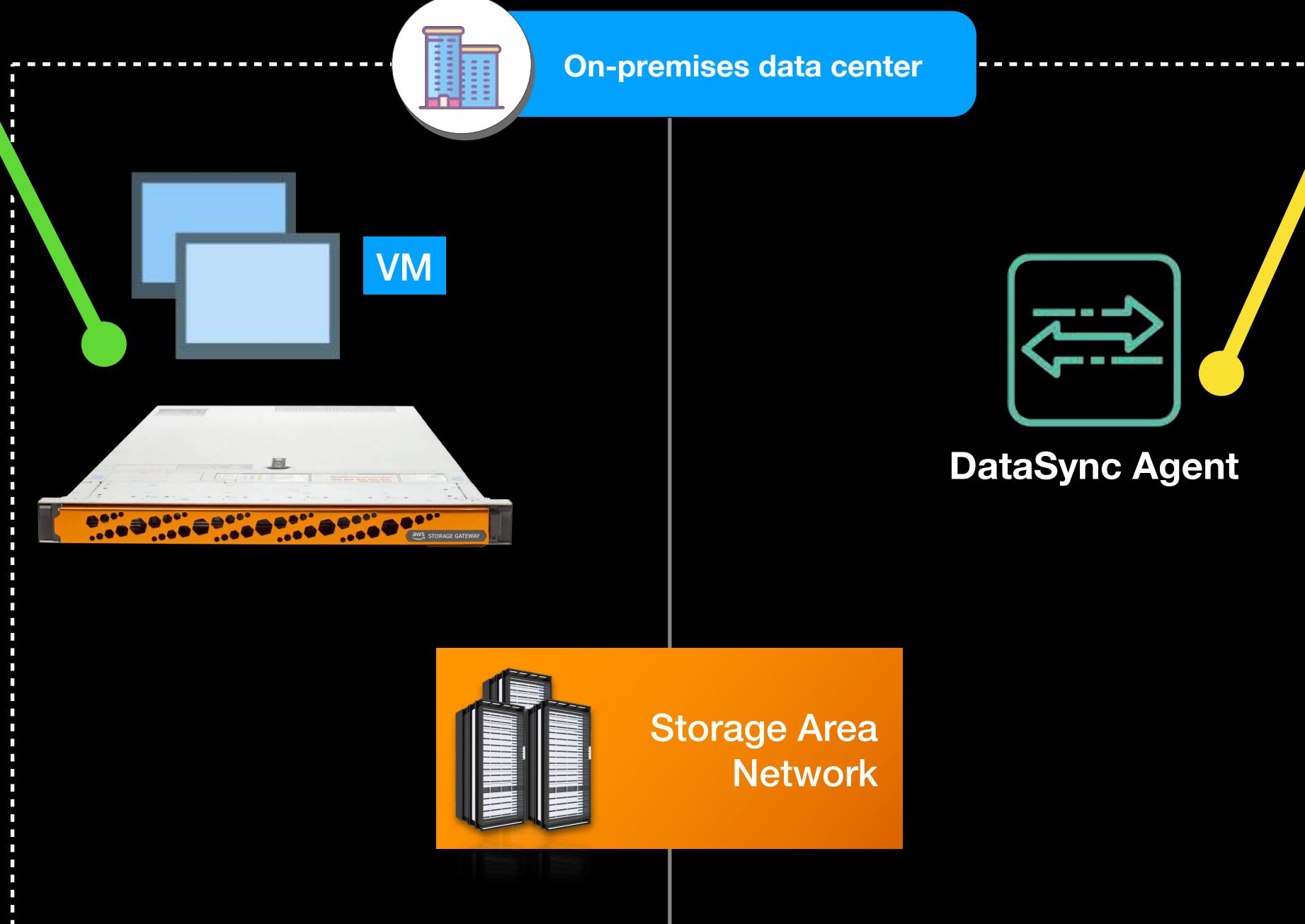


**MOVE DATA**

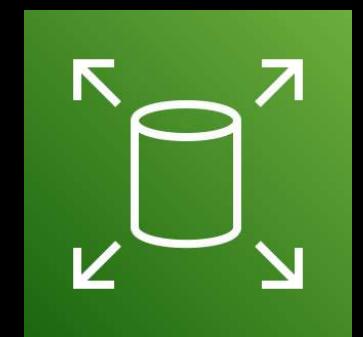
**MIGRATION**



**AWS DataSync**



*On-premises data would not  
be utilized anymore/will be  
decommissioned*



**Amazon EBS**

**VS**

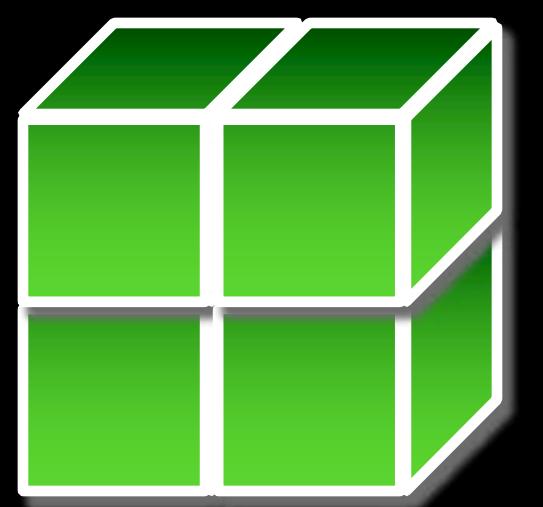


**Amazon EFS**

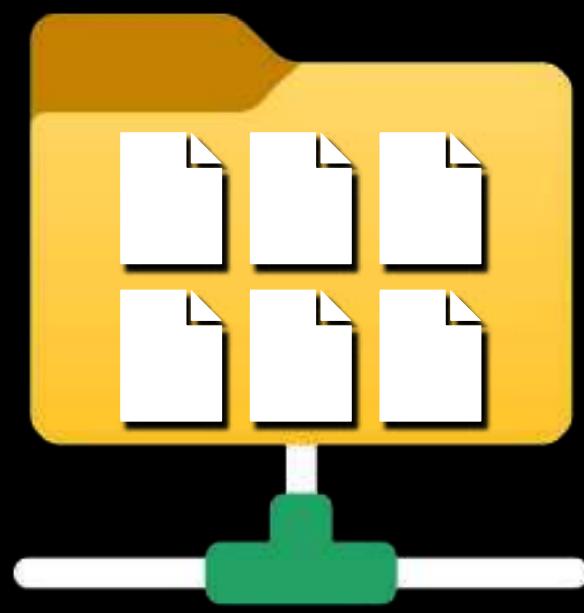
**VS**



**Amazon S3**



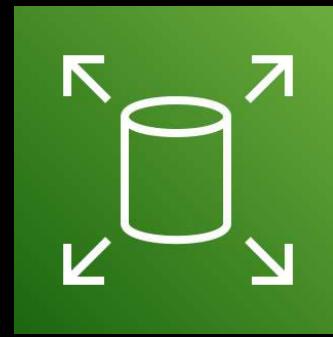
**BLOCK STORAGE**



**FILE STORAGE**



**OBJECT STORAGE**



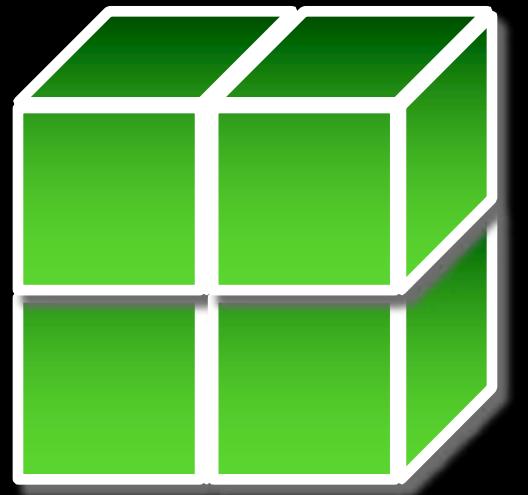
**Amazon Elastic Block  
Store**



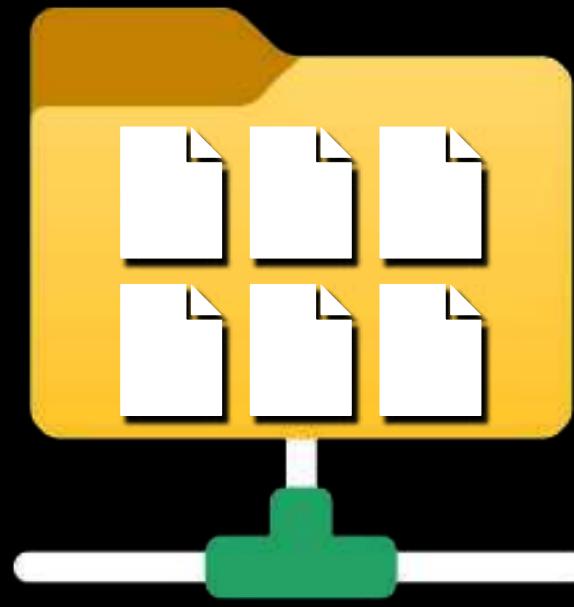
**Amazon Elastic File  
System**



**Amazon Simple Storage  
Service**



**BLOCK STORAGE**

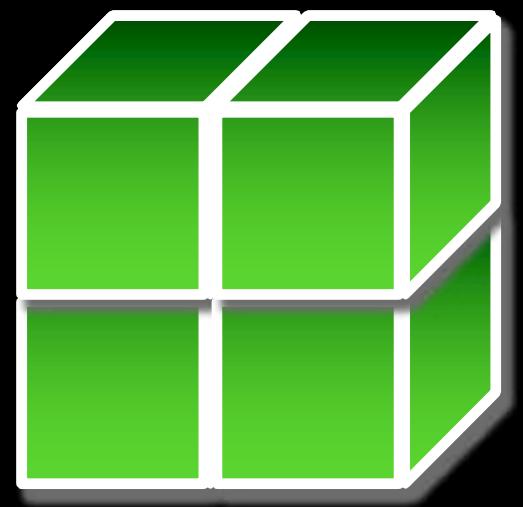


**FILE STORAGE**



**OBJECT STORAGE**

**Total File Size = 16 kb**  
**Block Size = 4 kb**



**BLOCK STORAGE**

```
dev — Terminal — -zsh — 67x10

jonbonso@tutorialsdojo > ls -lrt disk1 rdisk1 stdin dev
dr-xr-xr-x  3 root  wheel          4393 11 Aug 10:35 dev
lr-xr-xr-x  1 root  wheel          0 11 Aug 10:35 stdin -> fd/0
brw-r----- 1 root  operator      1,   3 11 Aug 10:35 disk1
crw-r----- 1 root  operator      1,   3 11 Aug 10:35 rdisk1
```

```
[jonbonso@tutorialsdojo ~] cd /dev  
[jonbonso@tutorialsdojo ~] ls -l  
brw-r----- 1 root operator 1, 0 11 Aug 10:35 disk0  
brw-r----- 1 root operator 1, 1 11 Aug 10:35 disk0s1  
brw-r----- 1 root operator 1, 2 11 Aug 10:35 disk0s2  
brw-r----- 1 root operator 1, 3 11 Aug 10:35 disk1  
brw-r----- 1 root operator 1, 4 11 Aug 10:35 disk1s1  
brw-r----- 1 root operator 1, 5 11 Aug 10:35 disk1s2  
brw-r----- 1 root operator 1, 6 11 Aug 10:35 disk1s3  
brw-r----- 1 root operator 1, 8 11 Aug 10:35 disk1s4  
crw-r----- 1 root operator 1, 0 11 Aug 10:35 rdisk0  
crw-r----- 1 root operator 1, 1 11 Aug 10:35 rdisk0s1  
crw-r----- 1 root operator 1, 2 11 Aug 10:35 rdisk0s2  
crw-r----- 1 root operator 1, 3 11 Aug 10:35 rdisk1
```

```
portal — Terminal — ssh -i tutorialsdojo.pem andresbonifacio@portal.tutorialsdojo.com — 135x24

drwxr-xr-x 3 root root          140 Aug  3 05:20 input
brw-rw---- 1 root disk          7,   1 Aug  3 05:20 loop1
brw-rw---- 1 root disk          7,   5 Aug  3 05:20 loop5
brw-rw---- 1 root disk          7,   4 Aug  3 05:20 loop4
brw-rw---- 1 root disk          7,   2 Aug  3 05:20 loop2
brw-rw---- 1 root disk          7,   3 Aug  3 05:20 loop3
brw-rw---- 1 root disk        259,   1 Aug  3 05:21 nvme0n1p1 ←
crw----- 1 root root          5,   1 Aug  3 05:21 console
crw--w---- 1 root tty           4,  64 Aug  3 05:21 ttys0
crw--w---- 1 root tty           4,   1 Aug  3 05:21 ttys1
brw-rw---- 1 root disk          7,   6 Aug 11 21:26 loop6
brw-rw---- 1 root disk          7,   0 Aug 11 21:26 loop0
crw-rw-rw- 1 root tty           5,   2 Aug 12 02:16 ptmx
root@ip-172-31-7-83:/dev#
```



Amazon EBS Volume

**Lower latency than**

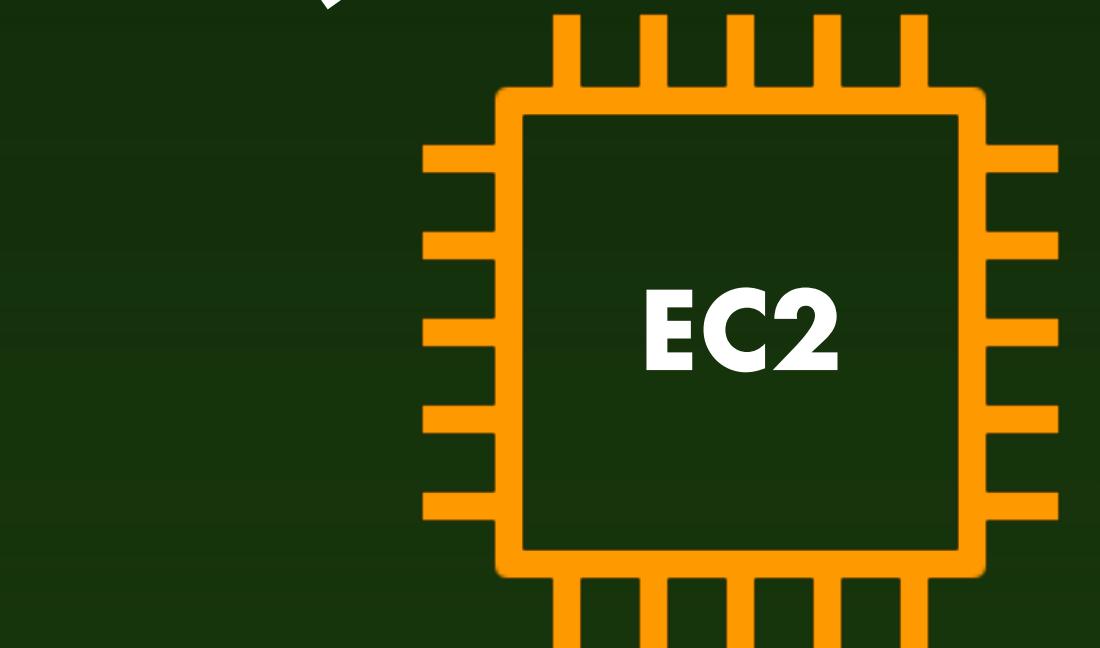


**Amazon S3**

**Amazon EBS**



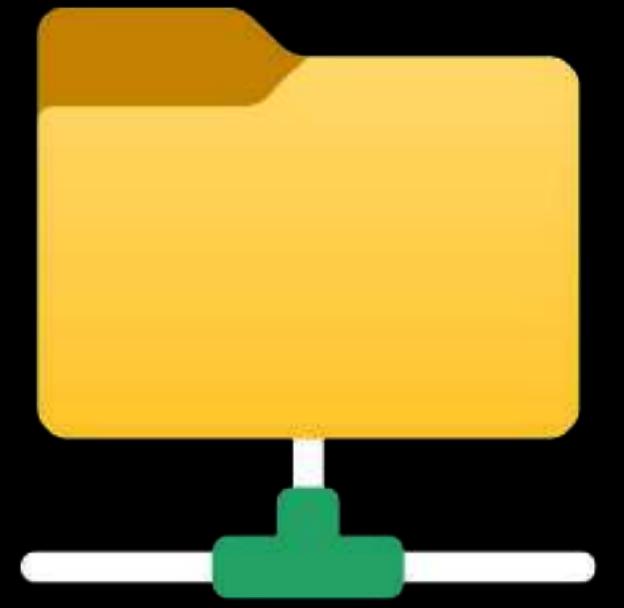
**Amazon EFS**



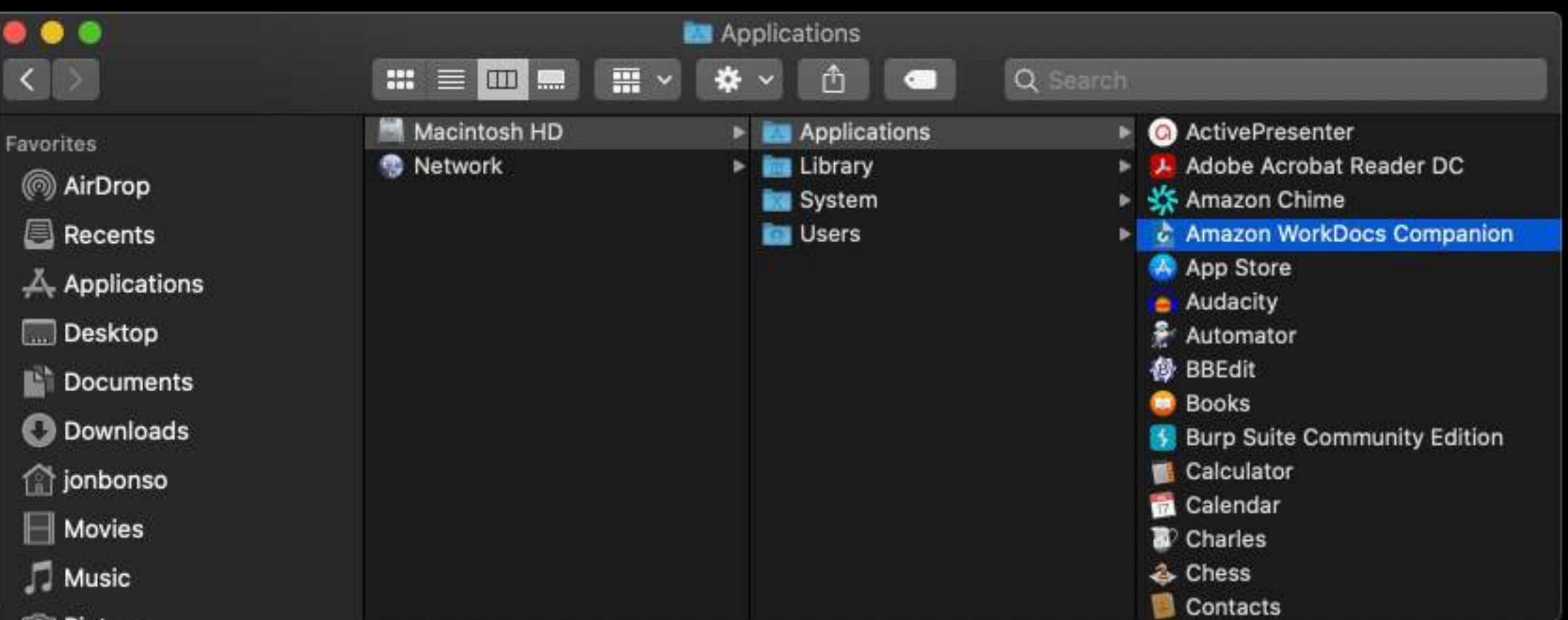
**Attached/Mounted to the  
Amazon EC2 instance**

- The **block storage or file storage** is **physically attached** to the host/server or located in **close proximity**

- The **latency** is **low** when transferring data between 2 systems



# FILE STORAGE



```
jonbonso@tutorialsdojo > cd /
jonbonso@tutorialsdojo > ls -lrt
total 9
drwxr-xr-x  2 root  wheel   64 29 Feb  2020 cores
drwxr-xr-x  2 root  wheel   64 29 Feb  2020 opt
drwxr-xr-x  5 root  admin  160  7 Apr  2020 Users
drwxr-xr-x@ 8 root  wheel  256  7 Apr  2020 System
lrwxr-xr-x@ 1 root  admin   11  7 Apr  2020 etc -> private/etc
lrwxr-xr-x@ 1 root  admin   11  7 Apr  2020 var -> private/var
lrwxr-xr-x@ 1 root  admin   11  7 Apr  2020 tmp -> private/tmp
drwxr-xr-x@ 11 root  wheel  352  7 Apr  2020 usr
drwxr-xr-x@ 38 root  wheel 1216 22 May 17:16 bin
drwxr-xr-x@ 63 root  wheel 2016 22 May 17:16 sbin
drwxr-xr-x  6 root  wheel  192 22 May 17:17 private
drwxr-xr-x  66 root  wheel 2112 22 May 17:17 Library
drwxrwxr-x+ 49 root  admin 1568 10 Aug 07:15 Applications
dr-xr-xr-x  3 root  wheel 4438 11 Aug 10:35 dev
lrwxr-xr-x  1 root  wheel   25 11 Aug 10:36 home -> /System/Volumes/Data/home
drwxr-xr-x  3 root  wheel   96 12 Aug 09:32 Volumes
jonbonso@tutorialsdojo >
```



- Commonly used by **multiple servers**
- Uses the **Portable Operating System Interface (POSIX)**



## OBJECT STORAGE

- Every object usually includes a globally unique identifier, its custom metadata and the data itself
- Doesn't depend on the operating system of the host/ EC2 instance
- Upload or fetch objects using RESTful web APIs and NOT by mounting it to the host

# DURABILITY



**Amazon EBS**



**Amazon EFS**

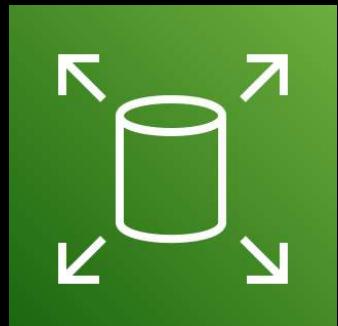


**Amazon S3**

**Data is stored  
redundantly in a single  
AZ only**

**Data is stored redundantly across multiple AZs**

# ACCESS METHOD



**Amazon EBS**

**Usually attached/mounted to a single EC2 instance**

**A single EBS volume can be attached to multiple EC2 instances by using the Multi-Attach feature**

(available on certain EBS types only)

**Two or more applications/EC2 instances can't access the exact same file concurrently**



**Amazon EFS**

**Can be mounted to thousands of EC2 instances or on-premises servers across multiple AZs**

**Allows multiple applications or servers to concurrently access the same files at the same time**

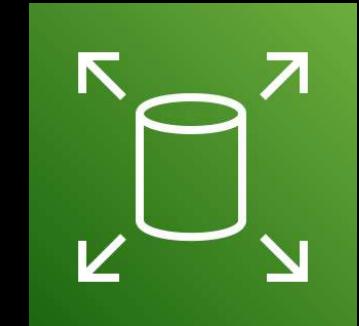


**Amazon S3**



**Invoked via a REST API request call**

# SCALABILITY



**Amazon EBS**

**Not highly scalable**

**Need to manually  
resize the EBS Volume  
to increase storage  
capacity**



**Amazon EFS**

**Both Amazon EFS and Amazon S3 are highly scalable**

**Automatically grows  
and shrinks the file  
system as you add and  
remove files**

**Can store virtually unlimited  
amounts of data**



**Amazon S3**

# L A T E N C Y



**Amazon EBS**



**Amazon EFS**



**Amazon S3**

**LOWEST**

**MODERATE**

**HIGH**

**MODERATE**

*if the request  
goes through the  
public Internet*

*if the request goes  
through the  
S3 Gateway Endpoint or  
S3 Interface Endpoint*

# BACKUPS



**Amazon EBS**



**Amazon EFS**



**Amazon S3**

**Back up data using  
Amazon EBS Snapshots  
(incremental backups)**

**Allows you to copy your  
EBS snapshot to another  
AWS Region**

**Transfer your file system to  
another EFS file system using**



**AWS DataSync**

**Perform incremental  
backups of your EFS file  
system using**



**AWS Backup**

**Cross-Region  
Replication (CRR)**

# DATA ENCRYPTION



**Amazon EBS**



**Amazon EFS**



**Amazon S3**

Encrypt your volume using  
**Amazon EBS Encryption**  
which is powered by



**AWS KMS**

**Encryption in Transit**  
Via TLS and the [EFS mount helper](#)

**Amazon EBS Encryption By Default**  
(Regional Setting)

**Encryption at Rest**

**Client-side Encryption**

**Server-side Encryption**

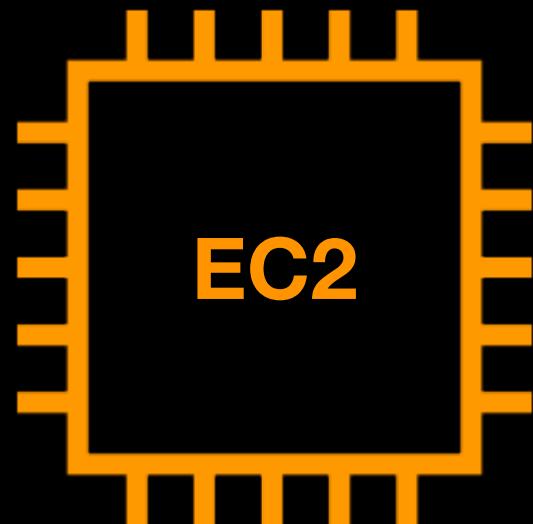
**Enforce HTTPS connection  
by setting up the Bucket  
Policy**

# ACCESS CONTROL



**Amazon EBS**

**Controlled by the associated security groups and Network ACL of the EC2 instance that the volume is mounted to**



**Amazon EFS**

**Can associate a security group to the file system mount target**



**NFSv4 endpoint**



**Amazon S3**



**Access Control List (ACL)**

**Bucket Policy**

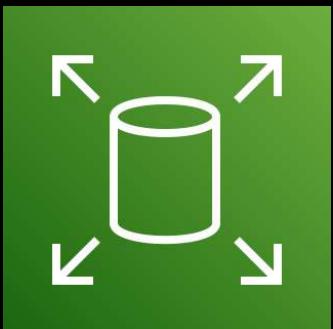


**S3 Access Points**



**S3 Object Lambda Access Points**

# NFSv4 Protocol Support



**Amazon EBS**



**Amazon EFS**



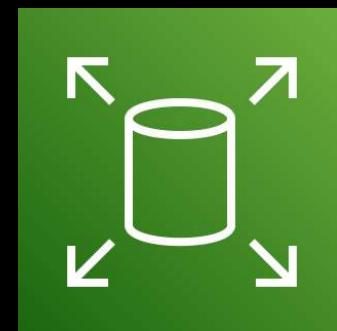
**Amazon S3**



**NFSv4 Support**  
**POSIX-compliant**



# DATA LIFECYCLE



**Amazon EBS**

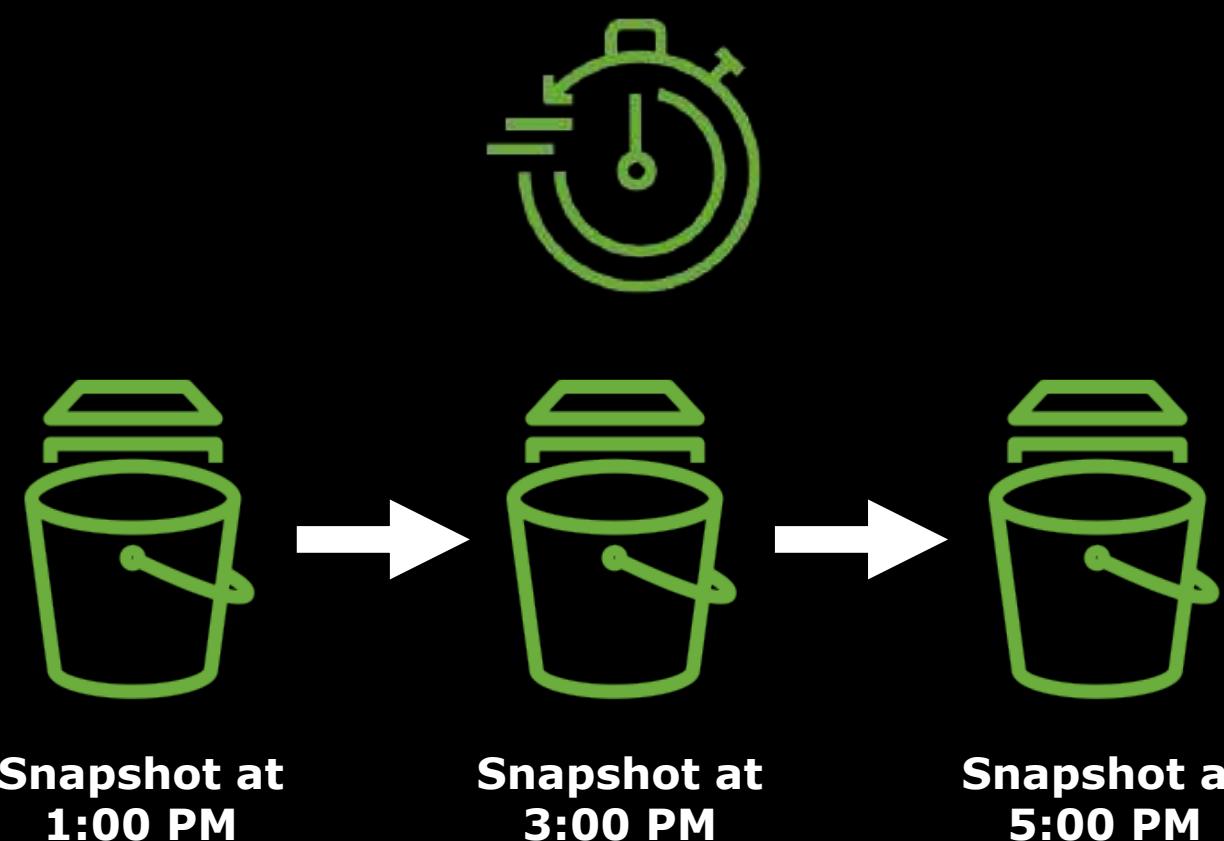


**Amazon EFS**

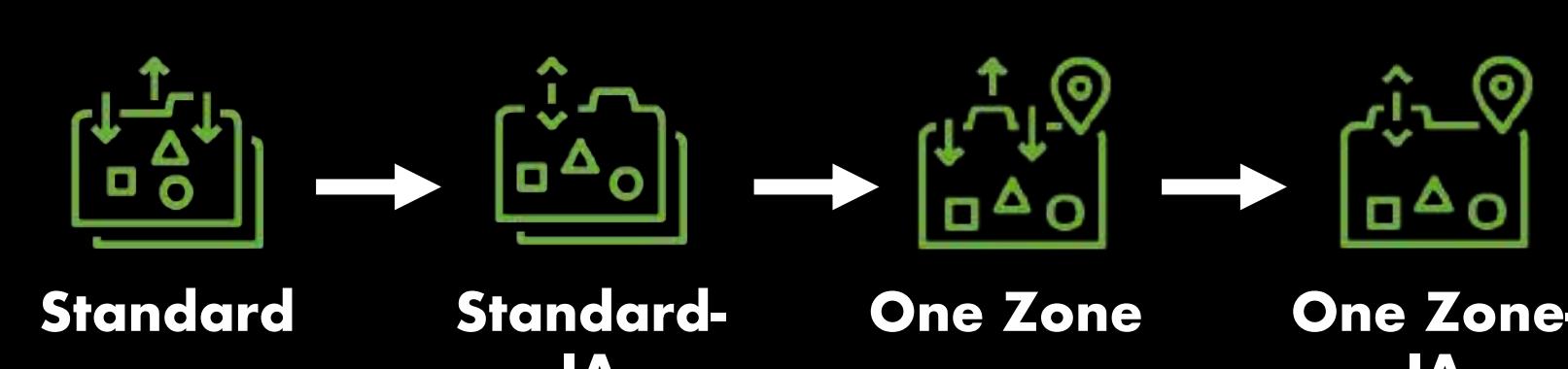


**Amazon S3**

## Amazon Data Lifecycle Manager (DLM)



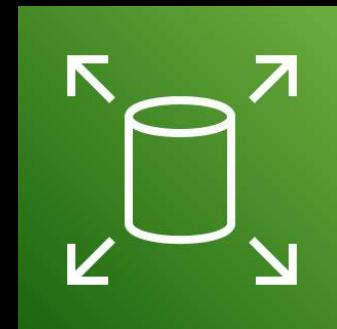
## Amazon EFS lifecycle management



## Amazon S3 Lifecycle Policy



# U S E C A S E S



**Amazon EBS**



**Amazon EFS**



**Amazon S3**

For storing **dynamic data** that are frequently accessed and updated

**LOWEST** Latency

A storage system accessed by multiple servers that need concurrent access to the same set of files at the same time

**POSIX-compliant**

For **static data** or for files that are NOT usually modified regularly

For a cost-effective & serverless static web hosting that can be integrated with:



**Amazon CloudFront**