H1

Acronyms & Full forms

Acronym	Description
OSI Model	Open Systems Intercomm Model - a standard for network communication
CloudFront	Amazon **Cloud Front** services enable caching of content at the nearest Edge Locations for fast distribution to customers.
CIDR	Classless Inter-Domain Routing (CIDR) - Route table for eg: 10.0.0.0/16
CFN - Cloud Formation	Models & provisions a collection of related AWS resources & 3rd party resources so you can launch & configure them together as a stack. Uses Infrastructure as code. supports **JSON** or **YAML** format.
AWS	Amazon Web Service
AVM	Amazon Vending Machine - uses Service Catalogue template to setup Multi-accounts
AMI	Amazon Machine Image - Save a copy of your server
IoT	Internet of Things
RI	Reserved Instances
SSD	Solid State Device
AZ(s)	Availability Zones
SNS	Simple Notification Service
SES	Simple Email Service
SQS	Simple Queue Service
DDoS	Distributed Denial of Service - AWS managed security protection
EC2	Elastic Compute Cloud
ECS	Elastic Container Service
ES	Elasticsearch
EB	Elastic Beanstalk

EBS	Elastic Block Storage
ENI	Elastic Network Interface - Logical networking component of a VPC
EFS	Elastic File Storage
EKS	Elastic **Kubernetes** Service
MKS	Managed **Kafka** Service
KMS	Key Management Service
EMR	Elastic Map Reduce
S3	Simple Storage Service. A Key / Value based storage system built to store & retrieve huge amount of data.
SWF	Simple Workflow Service
SSM	Simple Systems Manager *(Roles assigned to developers)*
TAM	Technical Account Manager
VPC	Virtual Private Cloud
VPN	Virtual Private Network
VIF	Virtual Interfaces - Each AWS Direct connect can be configured with one or more VIFs
IAM	Identity & Access Management
IDS	Intrusion Detection System
IPS	Intrusion Protection System
MFA	Multi Factor Authentication
RDS	Relational Database Service (PostgreSQL like DB)
Amazon DynamoDB	AWS Managed NoSQL Database
Amazon DocumentDB	MongoDB compatibility AWS managed Database
Amazon Redshift	Fast, simple, cost-effective data warehousing

ASG	Auto Scaling Group - Group multiple servers
ELB	Elastic Load Balancer - puts a load balancer infront of the instances.
ALB	Application Load Balancer
NLB	Network Load Balancer
Web ACL	Web Access Control List
NACL	Network Access Control List
IGW	Internet Gateway - Enables access to the internet
AWS WAF	**Web application firewall** (**WAF**) A **web application firewall** (**WAF**) helps protect your web applications from common web exploits that could affect application availability, compromise security, or consume excessive resources.
MQ	Amazon ActiveMQ
TCO	Total Cost of Ownership

H1 AWS Global Infrastructure

H2 Features

- 69 Availability zones
- 22 Geographic regions
- More Edge locations than AZs
- Low latency
- High Availability
- **Regions** - Physical location in the world with multiple **Availability Zones (AZ)**
- **Availability Zones (AZ)** - one or more discrete data-centres
- **Edge Locations** - data-centre owned by a trusted partner of AWS

H2 Regions

❗ A **geographically distinct** location which has multiple data-centres (AZs)

- Each region has atleast **2 Availability Zones (AZs)**
- Every **region** is **physically isolated** & **independent** to every other region in terms of - location, power, water supply
- ❗ AWS largest region is **US-EAST** (ie., North Virginia)
- ❗ Services almost always available first in **US-EAST**
- Not all services are available in all regions

- ❗ **US-EAST-1** is the region where you see all your billing information
- 5 new regions - under progress
- AWS regional Endpoints
- 3D world Model - AWS Regions

❗ ❗ The following table classifies AWS services based on their scope (Global or Regional Level)

Global Scope / Service	Regional Scope / Service
AWS Route 53	AWS EC2
AWS CloudFront	AWS Lambda
AWS IAM	AWS S3
AWS STS	

> *AWS S3 has global namespace, but buckets & objects are created on regional level*
>
> *By default all AWS STS requests go to the single endpoint: https://sts.amazonaws.com*

H2

Availability Zones (AZs)

❗ An AZ is a datacenter owned & operated by AWS in which their services run

- Each region has atleast **2 Availability Zones (AZs)**
- ❗ AZs are represented by a Region Code, followed by a letter identifier

```
# a - availability zone 'a'
us-east-1a
# North Virginia has 6 Availability Zones
us-east-1a
us-east-1b
us-east-1c
us-east-1d
us-east-1e
us-east-1f
```

- **Multi-AZ:** Distributing your instances across multiple AZs allows failover configuration for handling requests when one goes down
- <10ms latency between AZz in a region
- ❗ AZs are physically separated in a typical metropolitan area and are located in a low risk flood plain
- AZs use discrete UPS & onsite backup generation facilities & are fed via different grid

from independent facilities
- AZs are all redundantly connected to mutiple tier-1 Transit Providers

H2 Edge Locations

An Edge location is a datacenter owned by a **partner trusted by AWS** which has a direct connection to the AWS network.

> *Edge locations are condent delivery network (CDN) endpoints for CloudFront*

- ❗ These locations serve requests for **CloudFront** & **Route 53** .
- Requests going to these services will be routed to the nearest edge location automatically
- ❗ **S3 Transfer Acceleration** & **API Gateway** endpoint traffic also use the AWS Edge Network
- This allows for **Low latency** no matter where the end user is geographically situated

H2 GovCloud (US)

AWS GovCloud Regions allow customers to host sensitive **Controlled Unclassified Information** & other types of regulated workloads.

> *Useful If you want to build something to a Government or Government related industry.*

- ❗ **Only** operated by employees who are U.S citizens, on U.S soil
- ❗ **Only** accessible to U.S entities & root account holders who pass a screening test
- Customers can architect secure cloud solutions which comply with
  - FedRAMP High baseline
  - DOJ's Criminal Justice Information Systems (CJIS) Security policy
  - U.S International Traffic in Arms Regulations (ITAR)
  - Export Administration Regulations (EAR)
  - Department of Defence (DoD) Cloud computing security requirements guide

H1 Getting Started

> ❗ *You cannot use AWS services without Credit Card*

H2 Billing Preferences, Budgets & Alarms

- Billing Metrics always exist in **us-east-1**
- You need to select **CloudWatch** service for billing alarms
- You get 10 free alarms as part of your basic account
- **SNS** - *Simple Notification Service* - Basic Email Alert
- Volume discounts are available, which means, The more you use, the cheaper the service gets

Free Services

Free Services include the following:

- Amazon VPC
- Elastic Beanstalk *(not the resources)*
- CloudFormation *(not the resources)*
- IAM - Identity & Access Management
- AutoScaling *(not the resources)*
- OpsWorks
- Consolidated Billing

IAM & Authentication

Identity Federation (login using Active Directory AD, Facebook) can be configured allowing secure access to AWS resources without configuring IAM accounts

- IAM supports PCI DSS compliance
- You can create only 2 Access keys to a particular user at a time
- These can be used to make programmatic calls to AWS when using the API, AWS CLI or AWS management.
- IAM users can be created to access AWS applications and these are called - "service accounts"
- You can have upto 5000 user per AWS account

IAM Roles

- IAM roles can be used for granting applications running on EC2 instances permission to AWS API requests using **instance profiles**

IAM Policies

- Policies are documents that assign permission & can be used with Users, Roles & Groups
- Policy documents are written in JSON format
- All Permissions are denied by default
- Least restrictive permissions are applied
- **IAM policy simulator tool** is to help you understand, test & validate the effects of Access Control policies.

Customize Account Alias

❗ This is an important security step when you first setup your account

> *Generally you should not always use your root account to login and setup AWS services. Also you should not share your root account credentials with developers*

- Youtube Link - Change IAM Users sign-in link
- In the Dashboard page, search for - **IAM**
- Click **customise** for - *IAM users sign-in link*
- Provide an **Account alias** - `raysiti-aws`

H3 Activate MFA on Root Account

The **Multi-Factor Authentication (MFA)** on your AWS root account adds another layer of protection to keep your account secure.

There are 3 different types of MFA device that you can assign to your root account -

- ✅ **Virtual MFA device** - An Authenticator App installed on your mobile or computer.
  - Select **Google Authenticator** - which reads QR code to authenticate
  - Scan the QR code with your mobile using the **Google Authenticator** app.
  - It provides to **MFA numbers** that you need to enter
- **U2F security key** - Yubikey
- **Other hardware MFA device** - Gemaito token

H3 Create Individual IAM user accounts

- Select the Accordian - Click on **Manage Users**
- Follow the - steps to create a new User

H3 Setup a Password Policy

- Select the Accordian - Click on Setup a Password policy
- Follow the instructions provided - in the link

H2 AWS STS

The AWS Security Token Service is to enable you to request - temporary, limited privilege for IAM users OR for the users that you authenticate (Federated access)

H3 Benefits of AWS STS

- You can provide access to your AWS resources without creating an AWS identity for them
- The temporary security credentials have limited lifetime
- After the temporary security credentials expire, they cannot be used.

The AWS STS API action returns a temporary security credential which consists of

- AWS access key ID
- Key secret
- A Session token

H2 IAM Best Practices

- Lock away the Root User access keys
- Always create individual IAM users
- Use AWS defined policies to assign permissions whenever possible
- Use Groups
- Grant least privilege
- Enable MFA for privileged users
- ❗ Dont generate access key for the root account user
- ❗ Whenever possible, use IAM roles with temporary security credentials instead of Long-term access keys
- ❗ Manage IAM user access keys properly

H1 Pricing Models

There are 4 models of payment when it comes to using AWS services

> *There are 2 ways of Paying for EC2 instances which can help you reduce spend - 1. RESERVED & 2. SPOT instances*

1. **On-Demand Instances** - Least Commitment

   When you launch an EC-2 instance, it is by default On-Demand pricing.

> *Good for users who want a low cost & flexibility of EC2 without any up-front payments or long-term commitments.*

> *Also good for short-term, spiky & unpredictable workloads that should not be interrupted*

- ❗ *no upfront payment*
- ❗ *no long-term commitment*
- *low cost & flexible*
- *only pay per hour*
- *perfect for short-term, spiky, unpredictable loads*
- ❗ *cannot be interrupted*
- *for first-time applications*

2. **Reserved Instances - RI** - upto 75% off
   1. *steady state or predictable state*
   2. *commit to EC2 over 1 or 3 year term*
   3. *can resell unused reserved instances*

> *Good for applications that have a predictable usage, that need reserved-capacity and for users who commit for a 1-3 year term.*

3. **Spot Instances** - upto 90% savings
   1. *request space computing capacity*
   2. *flexible start & end dates*
   3. *can handle interruptions (server randomly stopping & starting)*
   4. *For non critical background jobs*

4. **Dedicated Host Instances** - Most expensive
    1. *you get Dedicated servers*
    2. *can be on-demand or reserved (upto 70% off)*
    3. *when you need a guarantee of isolated hardware*
    4. *for enterprise requirements*
    5. *When you want to leverage existing out-bound server licenses such as Windows Server, Suse Linux Server...etc.,*

H2 1. On Demand Instances / Pricing

❗ Least Commitment

When you first start / create an EC2 instance, it is by default on, **On-Demand**

H2 2. Reserved Instances / Pricing - RI

❗ Best Long term value

These type of instances are designed for applications that have a steady-state, predictable usage or require "reserved capacity".

$$reduced\ \ pricing = term * class\ offering * payment\ option$$

- ***Class Offering*** - Standard, Convertible, Scheduled
- ***Term*** - You commit to a 1 year or 3 year contract. Longer the term, greater savings
- ***Payment options*** - All, Partial or No - upfront payment

1. Reserved Instances can be shared between multiple accounts within the organization
2. Unusued RI can be sold in **Reserved Instance Marketplace** .

H2 3. Spot Instances - SI

❗ upto 90% off (when it comes to savings) but can be interrupted

AWS has **unused compute capacity** that they want to utilise the max capacity of their idle servers. So they came up with **Spot Instances**

- These instance provide the most savings - upto 90% compared to **On-Demand** Pricing.
- Can be interrupted / terminated
- These instances **can be terminated** if the ***compute capacity is needed by on-demand customers***
- **AWS Batch** is an easy & convenient way to use **Spot-Pricing**
- Useful for **Load-balancing** servers, Starting / Stopping, testing, can be spun up and stopped.
- If AWS terminate your instances, you wont be charged for the partial hours that you used.
- But if You terminated these Spot-instances, you will be charged for the number of hours they were used.

H2 4. Dedicated Host Instances

❗ Most expensive pricing

- Designed to meet regulatory requirements
- When you have strict **server-bound license** that won't support multi-tenancy or cloud deployments
- Perfect for Enterprise or Large Organizations
- Enterprises / Large Organizations may have security concerns or obligations against sharing the same hardware with other AWS customers
- Comes in 2 flavours - both Reserved & On-Demand (70% off on-demand pricing)

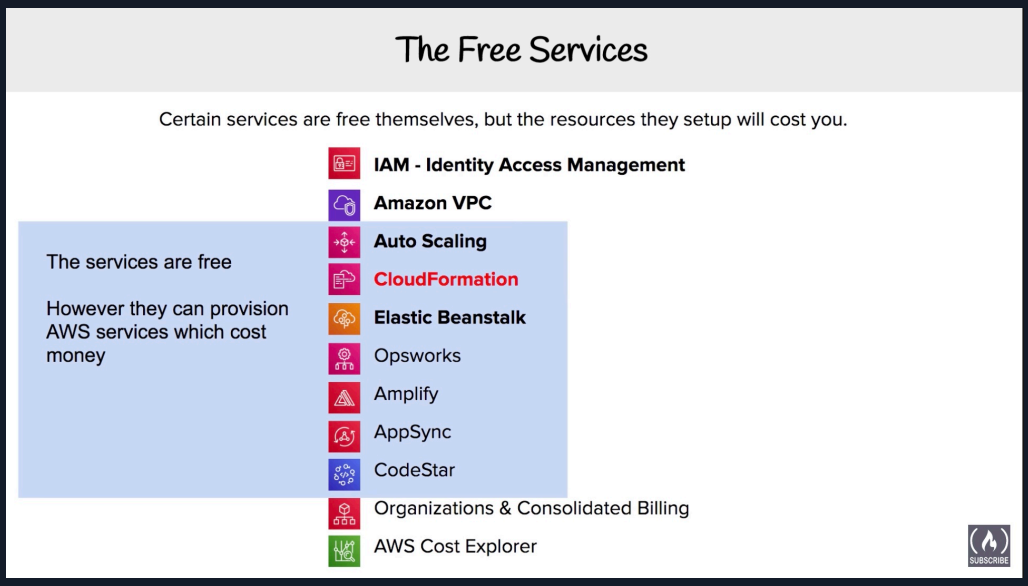
Multi tenant	Single Tenant
Multiple customers are running workloads on the same hardware	Single customer has been dedicated to the hardware
Virtual Isolation separates these customers (like a virtual apartment)	Physical isolation is what separates customers (think individual housing)
Eg: On-Demand, Reserved, Spot instances	Eg: Dedicated VPC

H1 Billing & Pricing

This module talks about billing & pricing when it comes to using AWS services

- The 3 driving forces for Billing are
  - Compute
  - Storage
  - Outbound Data Transfer
- Outbound Data is aggregated across services & then charged at Outbound data transfer rate

H2 Free Services



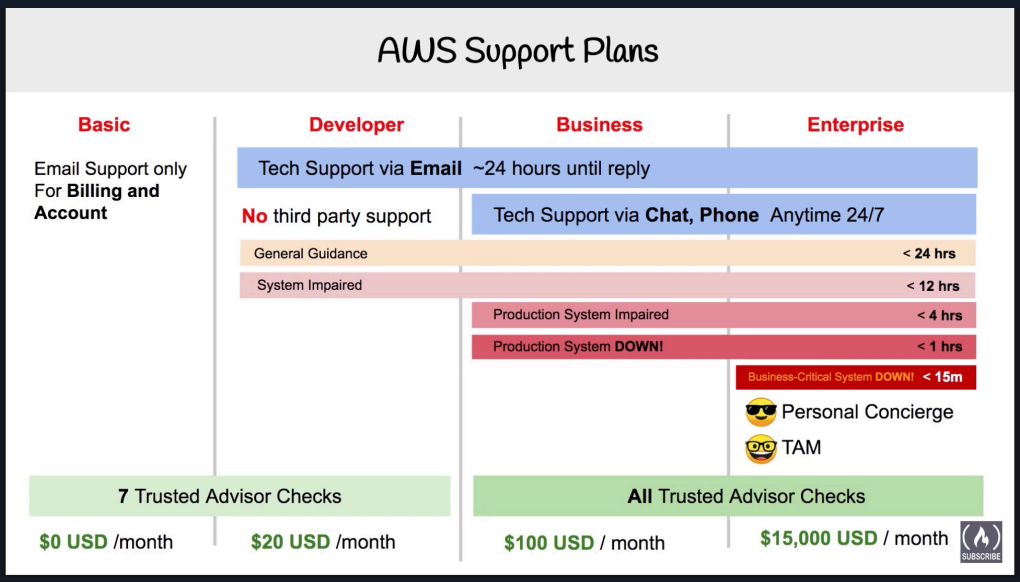
- These services are free to use
- However they can provision other AWS services which can cost money
- Pricing policies include
  - Pay as you go
  - Pay less when you reserve
  - Pay even less as your services grow
  - Pay less when AWS grows
  - Custom Pricing (for Enterprise)

H2 AWS Support Plans

AWS offers 4 different support plans

- **Basic**
  - $0 USD / month
  - Email support only
  - mainly for billing & account purposes.
  - 7 Trusted Advisor checks allowed
- **Developer**
  - $20 USD / month
  - Tech Support via Email ~24 hours only.
  - No 3rd party support (ie., no Application specific support)
  - 7 Trusted Advisor checks allowed
- **Business**
  - $100 USD / month
  - 24/7 chat & phone support.
  - Screensharing possible (includes Technical support)
  - All Trusted Advisor checks allowed
- **Enterprise**

- $15,000 USD / month
- 2 dedicated resources.
- Personal Concierge, TAM (Technical Account Manager)
- All Trusted Advisor checks allowed



H2 AWS Marketplace

AWS Marketplace is a curated digitally catalogue with 1000 of software listings from independent software vendors.

You can also search AWS marketplace under the context of a AWS service (for eg: EC2 service)

- Easily find, buy, test software that runs on AWS servers.
- Contains Products that are free to use / have an associated charge.
- The sales channel of ISV & Consulting partners allows you to sell your solutions to other customers.
- Has 1,361 (approx) vendors
- To see a list of Subscribed services ➡️ 🅰️ AWS Marketplace subscription

Products can be offered up as follows:

- Amazon Machine Images (AMIs)
- AWS cloudformation templates
- Software as a service (SAAS) offerings
- Web ACL
- AWS WAF rules

H3 What Is an Web ACL - Access Control List ?

An access control list (ACL) contains rules that grant or deny access to certain digital environments. There are two types of ACLs:

- **Filesystem ACLs:** filter access to files and/or directories. Filesystem ACLs tell operating systems which users can access the system, and what privileges the users are allowed.

- **Networking ACLs:** filter access to the network. Networking ACLs tell routers and switches which type of traffic can access the network, and which activity is allowed.

H2 AWS Trusted Advisor

This service advices you on security, saving money, performance, service limits & fault tolerance.

- Think of it as an automated checklist of AWS best practices
- On Free tiers you get 7 Trusted Advisor checks
- On Business & Enterprise tier you get all Trusted Advisor checks

This checklist can be grouped into 5 categories

- ❗ **Cost optimization**
  - Idle Load Balancers (free)
  - Unassociated Elastic IP addresses (free)
- **Performance**
  - High Utilization Amazon EC2 instance (free)
- ❗ **Security**
  - MFA on Root account (free)
  - IAM access key rotation (free)
- **Fault Tolerance**
  - Amazon RDS Backups (free)
- **Service Limits**
  - VPC (free)

H2 Consolidated Billing

This is a feature turned on by default when there is a Multiple member account.

- All billing charges are sent to the Master account
- Consolidates your billing & payment methods across multiple AWS accounts into one bill
- For Billing - AWS treats all accounts UNDER one organization as if they were one account
- This service is free of cost
- Limited to 20 linked accounts (by default)
- use **Cost explorer** to visualize your consolidated billing

H3 Volume Discounts

- AWS has volume discounts for many services
- The more you use, the more you save
- Consolidated Billing lets you take advantage of Volume discounts

Data Transfer	Price
First 10 TB	$0.17 per GB
Next 40 TB	$0.13 per GB

H2 AWS Cost Explorer

Allows you to **visualize**, **understand** & **manage** your AWS costs & usages over time.

- Default reports - Cost driver & Usage trends
- Use forcasting to get ideas on future workload & costs
- Visualize - daily or monthly level
- Can view *Past 13 months*
- Can predict *Future 3 months*

H2 AWS Cost and Usage Report

Generate a detail spreadsheet to help you analyse & understand your AWS costs

- Places reports on S3 bucket
- Uses Athena to generate the report
- Use QuickSight to generate graphs

H2 AWS Simple Monthly Calculator

Helps Customers & Prospects estimate their monthly AWS bill more efficiently

- You can add services in different regions
- Includes support for most AWS services
- Include additional costs such as - **Data ingress / egress charges, storage charges, retrieval fees**
- Support is also provided

H2 AWS Budgets

Think of AWS Budgets as - *billing alarms on steriods*!

AWS Budgets allow you to setup alerts if you are nearing / exceeding your defined budgets.

- Create Cost, Usage & Reservation budgets
- First 2 budgets are free of charge
- When you use AWS Budgets - Each budget is $0.02 per day OR ~ $0.60 USD / month
- 20,000 budget limit
- Plan your service usage, service costs & instance reservations
- Track - monthly, quarterly or yearly levels

- Get notified via - Email, Chatbot

H2 TCO Calculator

TCO Calculator stands for - Total Cost of Ownership Calculator

It estimates **"HOW MUCH YOU WOULD SAVE"** when moving from *on-premise* to *AWS cloud*.

- Generates detailed set of reports for presentation
- Built on underlying calculation which generate fair assessment of value
- Helps by reducing the need to invest on large capital expenditures
- The tool is for approximation purposes only!
- Helps reduce *"Total Cost of Ownership"* by avoiding large CAPEX on Infrstructure & hardware
- Elements can be added / removed which helps in better estimation

H2 AWS Landing Zones

Helps "Enterprise" level users quickly setup a secure, AWS multi-account architecture.

- Very expensive upfront cost
- Provides you with a baseline environment to get started with multi-account architecture
- Uses **Amazon Vending Machine (AVM)** which uses a **Service Catalogue Template** for setting multi-accounts
- Automatically services login and Single Sign On for customers
- Implement account baselines using Landing Zone configuration and pipeline

H2 AWS Quickstarts

Prebuilt Templates by AWS & AWS Partners to help you deploy your popular stacks on AWS.

- Reduce hundreds of manual procedures into few simple steps
- Composed of 3 parts
  - Reference architecture for the deployment
  - AWS Cloudformation template which automate & configure the deployment
  - Deployment guide explaining the architecture & implementation in detail

H2 Resource Groups & Tagging

H3 Tags

Words or Phrases that act as metadata for organizing your AWS resources

H3 Resource Groups

Collection of AWS resources that share one or more tags. can display based on -

- Metrics
- Alarms
- Configuration Settings

While creating a Resource Group, you can create 2 types

- Tag Based
- Cloud Formation based

H1 Question & Answers

H3 Question 01

Which of the following statements are true ? (Choose two)

- Amazon **Cloud Front** services enable caching of content at the nearest Edge Locations for fast distribution to customers ✅
- There are more Availability Zones than Edge Locations, & more Edge Locations than regions in the world ❌
- There are more Edge Locations than Availability Zones, & more Availability Zones then Regions in the world ✅
- There are more Regions than Availability Zones, & more Availability Zones than Edge Locations in the world ❌

H3 Question 02

*Which of the following is one of the design principles related to "**Reliability**" in cloud*

- Protect data on transit & on rest ❌
- Automatically recover from failure ✅
- Perform operations as code ❌
- Go global in minutes ❌

H3 Question 03

Which of the following is an - Amazon Domain Name System (DNS) web service ?

- Amazon Route 53 ✅
- Amazon Lightsail ❌
- AWS Snowball ❌
- AWS Direct connect ❌

H3 Question 04

Which of the following statements are true ? (Choose two)

- "Enterprise" AWS support plan includes Infrastructure Event Management with additional fee ❌
- "Enterprise" AWS support plan includes Infrastructure Event Management without additional fee ✅
- "Business" AWS support plan includes Infrastructure Event Management with additional fee ✅
- "Business" AWS support plan does not include Infrastructure Event Management at all. ❌

H3 Question 05

Under AWS shared responsibility model, which of the following is the responsibility of AWS

- Firewall Configuration ❌
- Physical security of Infrastructure ✅
- Patching guest operating systems ❌
- Security groups configuration ❌

H3 Question 06

Which of the following helps secure an AWS account

- Use Cloudwatch logs insight ❌
- Enable AWS Organizations ❌
- Activate Multi-Factor Authentication (MFA) ✅
- Enable AWS config ❌

H3 Question 07

When using Amazon IAM, what authentication methods are available to use (choose two) ?

- Client certificates ❌
- Access Keys ✅
- Amaon KMS ❌
- Server Certificates ✅
- AES 256 ❌

Explanation

- Supported authentication methods include
  - Console Password
  - Access Keys
  - Server Certificates (SSL / TLS)

H4

- Client certificates are not a valid authentication method
- ❗ Amazon KMS is used for managing encrypted keys and not for authentication
- AES 256, is an encryption algorithm, not an authentication method.

H3 ## Question 08

Which of the following types are valid Reserved Instances ? (Choose two)

- Convertible RI ✅
- Discounted RI ❌
- Scheduled RI ✅
- Long-term RI ❌
- Special RI ❌

Explanation

Reserved Instances are of 3 types

- Standard RI - most significant discount (upto 75% off on-demand instances)
- Convertible RI - 54% discount and the abililty to change the attributes
- Scheduled RI - Available to launch within the window timeframe you reserve

H4

H3 ## Question 09

How do you reduce latency between End-users & application / content ?

- Deploy Applications in multiple AZs ? ❌
- Deploy Applications in regions closest to end users ✅
- Use Amazon CloudFront to cache content closest to end users ✅
- Use S3 acceleration to improve application performance ❌
- User Larger EC2 instance types ❌

Explanation

- Deploying to multiple AZs may create resiliency but will not improve latency as AZs are close to each other
- S3 Transfer Acceleration is used to improve upload speeds to S3 server & does not affect application performance

H4