AWS Architect Associate Course Content

- **Introduction to Cloud Computing**
  - Introduction to Cloud Computing
  - Service Models in Cloud Computing
  - Deployment Model in Cloud Computing
  - Introduction to AWS
  - AWS Account Creation & Free Tier Limitations Overview
- **Compute**
  - EC2 Instance Launch Wizard
  - EC2 Instance Type
    - Windows EC2
    - Linux EC2
  - Security Group
  - Webservers
    - Linux Webserver
  - Bootstrap Script
  - Elastic load Balancer
    - Classic Load Balancer
  - Auto Scaling
  - Status Checks
    - Instance Status Check
    - System Status Check
  - Protect the Instance From Termination
  - Scale Up and Down
  - Scale Out and In
  - Snapshots
    - Moving Snapshot to Another Region
  - Creating Customized Amazon Machine Images (AMI)
  - Create EC2 Instance Alarm
  - Metadata and Tags
  - AWS Command Line Interface (CLI)
  - Elastic Beanstalk

AWS Architect Associate Course Content

- **Database**
    - Introduction of Database
    - Types of AWS Databases
    - Launching a RDS Instances (MySQL)
    - Multi Availability Zone For RDS Instances
    - Read replicas For RDS Instances
- **Security Options**
    - Cloud Trails
    - Cloud Front (S3 Transfer Acceleration)
- **Application Service**
    - Simple Notification Service
    - Simple Queue Service
    - AWS Calculator
- **Monitoring Tools**
    - Cloud Watch
        - Create Billing Alarm
- **Migration and Transfer**
    - AWS Snow (Theory)

AWS Architect Associate Course Content

- **Storage**
    - Elastic Block Storage (Volume)
        - Attach & Detach Volume
    - Introduction of Simple Storage Service (S3)
    - Versioning in S3
    - Static Webhosting with S3 Bucket
    - Amazon S3 Storage Classes
- **Identity & Access Management**
    - Root Account VS IAM User
    - Create IAM User & Attach Policy
    - Create User Groups & Attach Policy
    - Creating Custom IAM Policies
    - IAM Password Policies
- **Amazon Virtual Private Cloud**
    - Introduction about VPC & Create VPC
    - Introduction about Subnet & Create Subnet
        - Public Subnet
        - Private Subnet
    - Internet Gateways & Route Tables
    - Bastion Hosts
    - NAT Gateways
    - Network Access Control List (NACL)
    - Ephemeral Ports
- **Route 53**
    - DNS Record Overview
    - Routing Policies
        - Simple Routing Policy
        - Latency Routing Policy
        - Failover Routing Policy
        - Weighted Routing Policy
        - Geolocation Routing Policy
    - Hosting Sample Website and Configuring Policies

Auto Scaling - Lab

- Create Load Balancer (Steps already given in previous class)

- Go to Launch Templates - Create launch template
- Launch template name - MyLTMP(Can give any name)
- Template version description - 1
- Select "Provide guidance to help me set up a template that I can use with EC2 Auto Scaling"
- Amazon machine image (AMI) - Amazon Linux 2
- Instance type - T2-Micro
- Key pair - Select existing Key Pair
- Security groups - Select existing Security group (SSH & HTTP must be opened)
- Storage (volumes) - Make it as 9 GB
- Resource tags (Key - Name & Value - MyLTMP)
- Advanced details - User data

```
#!/bin/bash
sudo su
yum update -y
yum install httpd -y
cd /var/www/html
echo "MyGoogle" > index.html
service httpd start
chkconfig httpd on
```

- Create launch template - View launch templates
- Can see launch template has been created successfully

- Go to Auto Scaling Groups - Create an Auto Scaling group
- Auto Scaling group name - MyASG
- Launch template - Select MyLTMP - MyLTMP
- Subnets - Select all 3 subnets - Next
- Select Enable load balancing
- Classic Load Balancer - Select MyLB
- Select ELB
- Health check grace period - 150 - Next
- Desired capacity - 3
- Minimum capacity - 3
- Maximum capacity - 10
- Target tracking scaling policy
- Target value - 90
- Instances need - 300 - Next - Next
- Tags - Add Tag (Key - Name & Value - Web Server) - Next
- Create Auto Scaling group
- Can see Auto Scaling has been created successfully
- Can see all 3 instances running successfully.

You can verify by terminating some instances to check whether Auto Scaling is working fine or not.

Terminate all after finishing lab

- Delete Auto Scaling
- Delete Launch Template
- Delete Load Balancer

Cloud Computing

What is Cloud Computing?

- Cloud computing is the on-demand delivery of IT resources (compute, storage, applications....) through a cloud services platform (AWS) via the internet with pay-as-you-go pricing.
- Accessing IT resources provided by cloud provider (AWS) through web.
- Three key words
  1. On – Demand (Whenever/whatever we need, we get immediately)
  2. Scalable (Increase and Decrease the configuration as per requirement)
  3. Pay only whatever you use

EC2

- EC2 is one of the famous Amazon web services by using which we can launch any number of Instances (Servers) as per our required configuration with in fraction of minutes.
- After launching Instances, we can increase and decrease the configuration as per our requirement without stopping the Instance.
- This service (EC2) we use more in companies.
- To launch an Instance, we need to select below

AMI (Amazon Machine Image):-

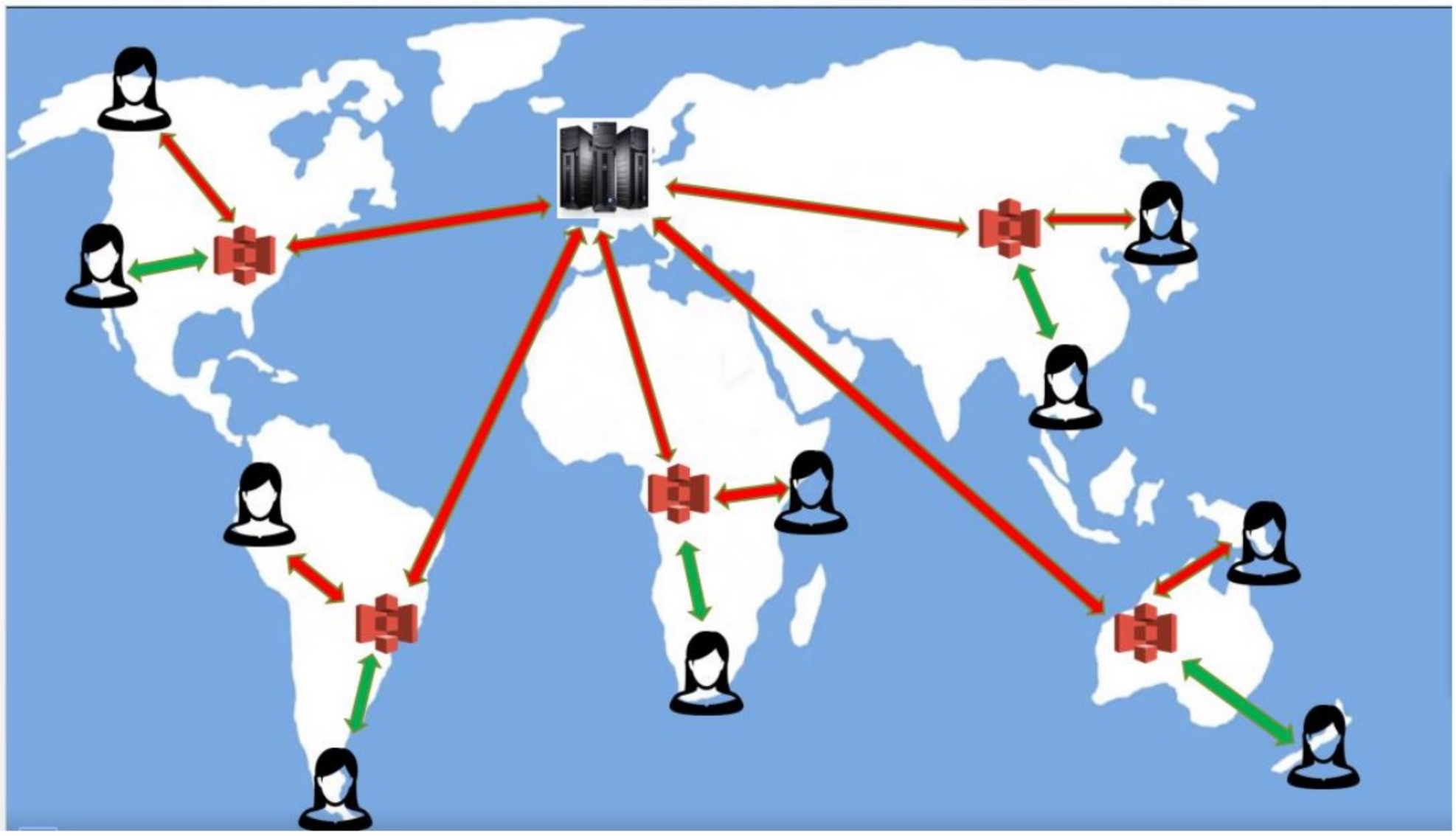
. AMI is simply an Operating system in AWS. AWS has provided some pre-defined Operating Systems of both windows and Linux. We have to choose OS from AWS provided list only. We can't bring any OS form outside into AWS.

CloudFront

- A Content delivery network (CDN) is a system of distributed servers (network) that delivers webpages and other web content to a user based on the different geographic locations.

CloudFront - Important Terminology

- **Edge Location** - This is the location where content will be cached. This is separate to an AWS Region/AZ
- **Origin** - This is the origin of all files that the CDN will distribute. This can be either an S3 Bucket, an EC2 Instance, an Elastic Load Balancer or Route53
- **Distribution** - This is the name given to CDN which consists of a collection of Edge Locations.



CloudFront - Important Points

- Edge locations are not just read only, you can write to them too.
- Objects are cached for the life of the TTL (Time To Live).

CloudFormation

- CloudFormation allows you to convert code into infrastructure.
- CloudFormation gives developers and system administrators an easy way to create and manage a collection of related AWS resources, provisioning and updating them in an orderly and predictable fashion.

- You don't need to figure out the order for provisioning AWS services. CloudFormation takes care of this for you.
- After the AWS resources are deployed, you can modify and update them in a controlled and predictable way, in effect applying version control to your AWS infrastructure the same way you do with your code.

- A CloudFormation Template is essentially an architectural diagram and a CloudFormation Stack is the end result of that diagram
- You create, update, and delete a collection of resources by creating, updating, and deleting stacks using CloudFormation templates.
- CloudFormation templates are in the JSON format or YAML.

A Sample Template

JSON

```
{
   “Resources” : {
      “HelloBucket” : {
         “Type” : “AWS::S3::Bucket”
      }
   }
}
```

YAML

```
Resources:
  HelloBucket:
    Type: AWS::S3::Bucket
```

Important points

- By default, the "automatic rollback on error" feature is enabled
- You are charged for rolled back infrastructure.
- CloudFormation is free

Elastic Beanstalk

- With Elastic Beanstalk, you can deploy, monitor, and scale an application quickly
- It provides developers or end users with the ability to provision application infrastructure is an almost transparent way
- It has a highly abstract focus towards infrastructure, focusing performance - not configuration and specifications
- It attempts to remove, or significantly simplify infrastructure management, allowing applications to deploy into infrastructure environments easily
- Either your entire application is one EB application or each logical component of your application, can be a EB application or a EB environment within an application

Important Points

- You can have multiple versions of your applications
- You can update your application
- You pay for the resources that you use, but Elastic Beanstalk is free
- If elastic beanstalk creates your RDS database then it will delete it when you delete your application. If not then the RDS instance stays

Supported Languages

1. Java
2. PHP
3. Python
4. Node.js
5. .NET
6. Go
7. Ruby

CADEMY

Instance type:-

Here we are going to choose CPU Cores and RAM. AWS is giving them as pairs. AWS paired best possible combinations form which we get maximum performance. These pairs we call as instance types.

EBS (Elastic Block Storage):-

EBS is simply a Hard disk that we attach to Instances. We can choose any amount of hard disk. We can have any no of drives as well. In this hard disk, we can keep both Operating System as well as Objects (MP3, MP4, Pictures, Documents.....)

- When we combine above all, we will get our required configured Server. That server in AWS we call as Instance.

- Apart from above, we select many other things as well like

Tag:-

Tag is just a meaning full name that we give to EC2 Instance for identification purpose.

Security Groups:-

1. Security Groups deals with Ports.
2. Port is like a door to your Instance.
3. We have total 0 - 65535 number of ports are there. Each port will have both incoming and outgoing options.
4. All these ports are dedicated for some special purpose
5. Most Important ports are RDP (3389), SSH (22), HTTP (80) and HTTPS (443).
6. RDP port is a dedicated port for windows. If you want to access any windows server, you need to open RDP port of that server. Through that port only we can access windows server. If it is Linux, then we have to open SSH (22) Port.
7. Remaining ports will be discussed in next classes.

Key:-

Key pair is just like a password. But here it is a file. By default we get .pem (Privacy-Enhanced Mail) key. To access windows instance, we need to convert that .pem into Password. Because, windows instance supports password only.

- Best practice is after finishing work, either terminate or stop your Instance. If we stop your instance, we can start at any time. But if we terminate instance, we can't start. Termination means losing instance forever.

Elasticache

- Elasticache is a web service that makes it easy to deploy, operate, and scale an in-memory cache in the cloud.
- The service improves the performance of web applications by allowing you to retrieve information from fast, managed, in-memory caches, instead of relying entirely on slower disk-based databases.
- It can be used to significantly reduce latency and throughput for many read-heavy application workloads (such as social networking, gaming, media sharing and Q&A portals) or compute-intensive workloads.
- Caching improves application performance by storing critical pieces of data in memory for low-latency access.

Types of Elasticache

- Memcached
- Redis (Open source)

EC2

- Amazon Elastic Compute Cloud (EC2) is a web service that provides resizable compute capacity in the cloud.
- EC2 reduces the time to launch new servers to minutes, allowing you to quickly scale capacity, both up and down.

UNIX flavors

- Linux
- Mac OS
- AIX
- Solaris
- HP-UX

Linux Flavors

- RHEL (Red Hat Enterprise Linux)
- Cent OS
- Ubuntu
- Amazon Linux
- Fedora
- Linux Mint
- Opens SUSE

CADEMY

EC2

EC2 (Linux Instance) - Lab

- EC2-Launch Instance
- Amazon Linux-Select
- T2-Micro-Next
- Number of instances-1-Next
- Next(storage section)
- click to add a Name tag: Linux - Next
- Security group name: LinuxSG
- Description: LinuxSG
- Review and Launch-Launch
- Create a new keypair-demo-Download key pair-Launch Instances-Click on Instance ID
- Click on Instances(Can see Linux machine running)

- Download Putty & PuttyGen tools
- Open PuttyGen tool-Load-demo.pem-ok-Save private key-yes-Desktop-demo-save
- Select Linux machine-connect-copy(user name@dns name)
- Open Putty tool-paste(user name@dns name)-ssh-Auth-Browse-demo.ppk-Open

EC2 Instance Types

	Family	Type	vCPUs	Memory (GiB)
☐	General purpose	t2.nano	1	0.5
☑	General purpose	t2.micro Free tier eligible	1	1
☐	General purpose	t2.small	1	2
☐	General purpose	t2.medium	2	4
☐	General purpose	t2.large	2	8
☐	General purpose	t2.xlarge	4	16
☐	General purpose	t2.2xlarge	8	32
☐	General purpose	t3.nano	2	0.5
☐	General purpose	t3.micro	2	1

EBS (Elastic Block Store)

- EBS allows you to create storage (Volumes)

  (Object storage: MP3, MP4, pictures, documents…..)

  (Block storage: OS, Data Bases)

EC2

EC2 (Windows) - Lab

- EC2-Launch Instance
- Windows Server 2019-Select
- T2-Micro-Next
- Number of instances-1-Next
- Next(storage section)
- click to add a Name tag: Windows - Next
- Security groups - Next
- Review and Launch-Launch
- Create a new key pair-demo-Download key pair-Launch Instances-Click on Instance ID
- Click on Instance ID (Can see Windows machine running)

- Actions-Connect-RDP client-Get Password-Choose File-demo.pem (upload)-Decrypt password
- Download Remote Desktop Connection application and run it.
- Enter Password-connect-yes
- Can see windows server 2019 screen in your laptop

Real Time Use Cases

(Where we use Windows Instance)

- To Install and work on DevOps tools
- To Test applications in different configuration servers
- Used by Big E-Commerce companies & Marts to do larger calculations
- To develop & run Games which need High configuration servers
- To work on High definition Graphics and Visual effects while making movies
- And many more….

CAIDEMY

EC2 Pricing Options

- **On Demand –** Whenever you want, launch server. Once you are done with your work, stop/terminate it.

  Eg:- uber car

- **Reserved** – Reserve servers for some duration.

  Eg:- train berth reservation

- **Spot** – Based upon market price.

  Eg:- share market price

- **Dedicated Hosts** – Where physical machine will be dedicated to you.

  Eg:- govt project

EBS Volume Types

1. SSD, General Purpose - GP2/GP3 - up to 10,000 IOPS
2. SSD, Provisioned IOPS – IO1/IO2 - >10,000 IOPS
3. HDD, Throughput Optimised - ST1 - Frequently accessed data
4. HDD, Cold - SC1 - Less frequently accessed data
5. HDD, Magnetic - Standard - Cheapest, Infrequently accessed storage

Elastic Transcoder

- Media Transcoder in the cloud
- Convert media files from their original source format in to different formats that will play on smartphones, tablets, PC's etc.
- Provides transcoding presets for popular output formats, which means that you don't need to guess about which settings work best on particular devices.
- Pay based on the minutes that you transcode and the resolution at which you transcode.

How To Create an Amazon AWS Free Tier Account

1. Go to https://aws.amazon.com

2. Click on "Create an AWS Account"

3. Enter Email address, Password, Confirm password & Account name. Click Continue when done.

4. Select Account type (Click on Personal)

5. Fill in the listed fields & click on acknowledge, then click Create Account and Continue.

6. Enter your credit/debit card details

7. Amazon will confirm your identity by sending you a text message or giving you a voice call. Fill in the details and click on Send SMS

8. You will receive a verification code either via text message or voice call, depending on what you selected in the previous step. Enter the code and select Verify Code.

9. Once the code has been entered in, you will be shown 'Your identity has been verified successfully'.

10. There are 3 Support Plans you can choose. Click on Free.

11. That's it! Your Amazon AWS Free Tier account has now been created.

12. Wait for one hour.

13. You are now ready to login to the AWS console. You can click on 'Sign in to the Console'.

14. Enter the email address that you used for your Free Tier account. Click next and then enter in your password.

15. You have now successfully signed in to the AWS Account.

HTTP & HTTPS

- **HTTP**:- Customers web requests will go in plain text. HTTP port no is 80.

- **HTTPS**:- Customers web requests will go in an encrypted manner. HTTPS port no is 443. To have this port opened, we need to buy SSL (Secure Socket Layer) certificate form registered organizations which comes with encryption and decryption keys.

Status Checks

There are two status checks

1. System status check
2. Instance status check

- **System status check**: - System status check means underlying physical machine check. Its status will be displayed as "0/2 checks passed" if this check fails. To rectify this issue, we may need to "stop & start" instance so that, At AWS Availability Zone, entire EC2 instance will be migrated to some other physical system which is running fine.

- **Instance status check**: - Instance status check means EC2 Instance check. Its status will be displayed as "1/2 checks passed" if this check fails. To rectify this issue, we may need to "reboot" instance so that, At AWS Availability Zone, entire EC2 instance OS will be reloaded and then it will work fine.

- **Note****: - If still unable to troubleshoot any of above issues, we need to contact AWS support.

Scaling

- We can scale up and down Instance type (CPU Cores & RAM). For this we have to stop instance. We can't scale Instance type while Instance in running mode.
- We can scale up EBS (Hard Disk). For this we need not to stop instance. We can scale up while Instance is in running mode. We can't scale down EBS as AWS is not allowing us to do so due to data security reasons.

Roles

- Without using credentials, we can manage aws services through aws cli by using roles
- Roles give secure way to access all aws resources
- Can access one aws service with another aws service without credentials.

Important Points

- IAM is universal. It does not apply to regions at this time.
- The "root account" is simply the account created when first setup your AWS account. It has complete Admin access.
- New Users have NO permissions when first created.
- New Users are assigned Access Key ID & Secret Access Keys when first created.
- These are not the same as a password, and you cannot use the Access key ID & Secret Access Key to login in to the console. You can use this to access Command line.
- You can create and customise your own password rotation policies.

IAM
(Identity and Access Management)

- IAM allows you to manage Users, Groups and their level of access to the AWS Services.

What does IAM give you?

- Centralised control of your AWS account
- Shared access to your AWS account
- Granular Permissions
- Multifactor Authentication
- Allows you to set up your own password rotation policy

Important Terms

- **Users** - End Users (People)
- **Groups** - A Collection of users under one set of permissions
- **Policies(Permissions)** - A document that defines one (or more) permissions
- **Roles** - You create roles and can then assign them to AWS resources

---

Load Balancer

There are two advantages of Load Balancer

1. It will distribute the load among all web servers
2. It will check the health of web servers. If at all health check fails of any web server, then load balancer will stop sending traffic to that particular web server until that server becomes healthy.

Load Balancer - Lab

- Go to Load Balancers-Create Load Balancer-Classic Load Balancer-Create
- Load Balancer name: MyLB-Next
- Select an existing security group: WebSG-Next-Next
- Response Timeout:2
- Interval:5
- Unhealthy threshold:2
- Healthy threshold:2 - Next
- Select EC2(both)-Next-Review and Create-Create-Close
- Take DNS name of LB and paste in browser(Can see content in web page)

Commands

- #!/bin/bash
- sudo su
- yum update -y
- yum install httpd -y
- cd /var/www/html
- echo "MyGoogle" > index.html
- ls
- service httpd start
- chkconfig httpd on

IP Addresses Classes

- **IP Address Classes**

The IP addresses are further broken down into classes. These classes are A, B, C, D, E and their possible ranges can be seen in Figure below.

Class	Start	End	Default subnet mask	CIDR
Class A	0.0.0.0	127.255.255.255	255.0.0.0	/8
Class B	128.0.0.0	191.255.255.255	255.255.0.0	/16
Class C	192.0.0.0	223.255.255.255	255.255.255.0	/24
Class D (multicast)	224.0.0.0	239.255.255.255		
Class E (reserved)	240.0.0.0	255.255.255.255		

*CIDR - Classless Inter-Domain Routing

* 127.0.0.0 to 127.255.255.255 is reserved for loopback address

CAIDE

Network

Networking: It's a connection between two or more machines to communicate with each other.

Network components:

- NIC (Network Interface Card)
- Media
- Topology
- Protocol
- IP Addresses

Basic requirements for Networking

NIC (Network Interface Card):

- It is a computer hardware component that connects a computer to a computer network.
- Each NIC will have a unique MAC address to avoid conflicts b/w same NIC adapters.
- We represent these by the word "eth" or "ens".

Media: It is a medium via which two different computer's NIC card will be connected.

E.g: RJ 45

Topology: Design in which the computers in the network will be connected to each other.

Eg : Bus, Ring, Star, Mesh and Tree

Protocol: Defines rules and conventions for communication b/w network devices.

IP Address: An Internet Protocol **address** is a numerical label assigned to each device connected to a computer network for communication.

Eg : Like Phone number

Protocol

TCP/IP:

- Transmission Control Protocol
- It is connection oriented
- TCP acknowledgement will be sent/received
- Slow Communication
    - Eg: HTTP, HTTPS

UDP:

- User Datagram Protocol
- Connectionless
- No Acknowledgement for UDP
- Faster Communication
    - Eg: DNS, DHCP

RDS (Relational Database Service)

- Relational databases are what most of us are all used to.
- A database structured to recognize relations between stored items of information.
  eg: Excel sheet

Relational Database Types

- SQL Server
- Oracle
- MySQL Server
- PostgreSQL
- Aurora
- MariaDB

RDS Back-ups

Two types of Backups

- Automated Backups
- DB Snapshots

Automated Backups:

- Automated backups allow you to recover your database to any point in time within a "retention period". The retention period can be between one and 35 days.

- Automated Backups will take a full daily snapshot and will also store transaction logs throughout the day. When you do a recovery, AWS will first choose the most recent daily back up, and then apply transaction logs relevant to that day.
- Automated Backups are enabled by default.

Snapshots:

- DB Snapshots are done manually (ie they are user initiated). They are stored even after you delete the original RDS instance, unlike automated backups.

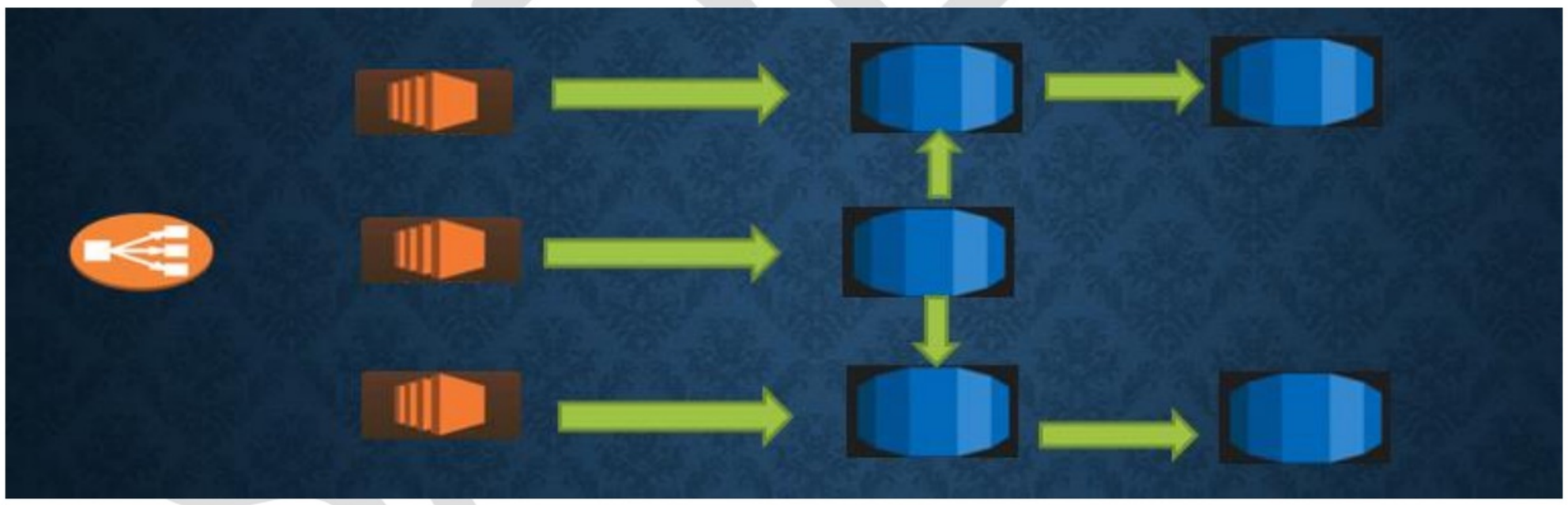
Restoring Backups

- Whenever you restore either an Automatic Backup or a manual Snapshot, the restored version of the database will be a new RDS instance with a new DNS endpoint.

Multi-AZ

- Multi-AZ allows you to have an exact copy of your production database in another AZ. AWS handles the replication for you, so when your production database is written to, this write will automatically be synchronized to the stand by database.
- In the event of planned database maintenance, DB Instance failure, or an AZ failure, Amazon RDS will automatically failover to the standby so that database operations can resume quickly without administrative intervention.
- Both DB servers have same DNS endpoints.

Read Replica



- Read replicas allow you to have a read-only copy of your production database. This is achieved by using Asynchronous replication from the primary RDS instance to the read replica.
- You use read replicas primarily for very read-heavy database workloads

- Use for scaling. Not for DR.
- You can have up to 5 RR copies of any database
- You can have read replicas of read replicas (But latency will be there)
- Each RR will have its own DNS end point

CADEMY

Redshift

Data Warehousing

- Used to pull in very large and complex data sets. Usually used by management to do queries on data.
    - eg: Cognos, Jaspersoft, SQL Server Reporting Services....

OLTP vs OLAP

OLTP(Online Transaction Processing) vs OLAP (Online Analytics Processing)

- OLTP: eg- I need particular value in so and so row and column
- OLAP: eg- I need sum of all values in table

Redshift

Amazon Redshift is a fast and powerful, fully managed, petabyte-scale data warehouse service in the cloud.

Redshift Configuration

- Single Node (160Gb)
- Multi-Node
    - Leader Node (manages client connections and receives queries)
    - Compute Node (store data and perform queries and computations). Up to 128 Compute Nodes

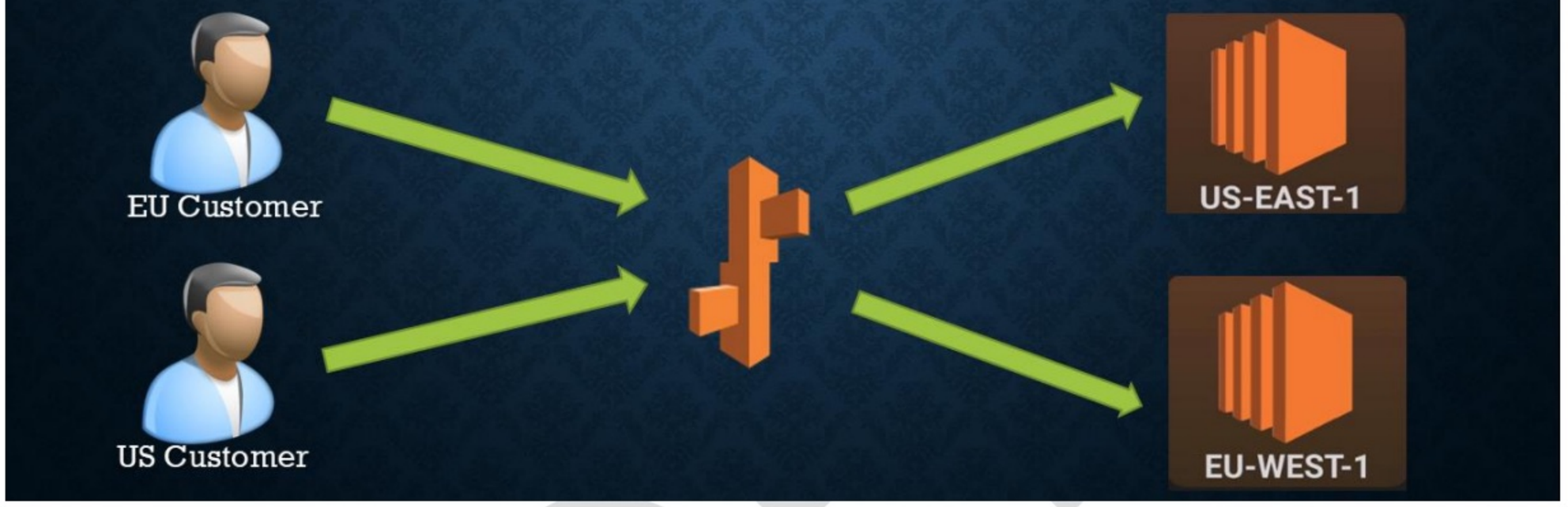
Massive Parallel Processing (MPP)

- Amazon Redshift automatically distributes data and query load across all nodes.
- Redshift makes it easy to add nodes to your data warehouse and enables you to maintain fast query performance as your data warehouse grows.

CADEMY

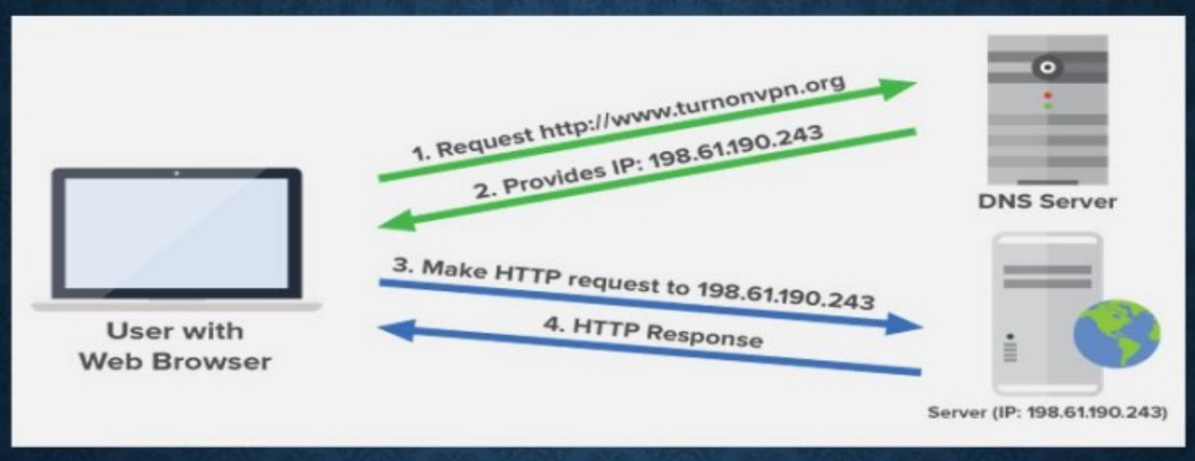
Geolocation Routing Policy

- Geolocation routing lets you choose where your traffic will be sent based on the geographic location of your users (ie the location from where DNS queries originate)
- eg: You might want all queries from Europe to be routed to a fleet of EC2 instances that are specially configured for your European customers. These servers may have the local language of your European customers and all prices are displayed in Euros.



Route 53

What is DNS?



- DNS is used to convert human friendly domain names into an Internet Protocol (IP) address and vice versa.
- IP addresses are used by computers to identify each other on the network.
- Two types of IPs
  - IPv4
  - IPv6

Domain Names

- **Top Level Domain**: The last word in a domain name represents the "Top level domain"

  eg: .com

- **Second Level Domain**: The second last word in a domain name represents the "Second level domain name"

  eg: .gov.in

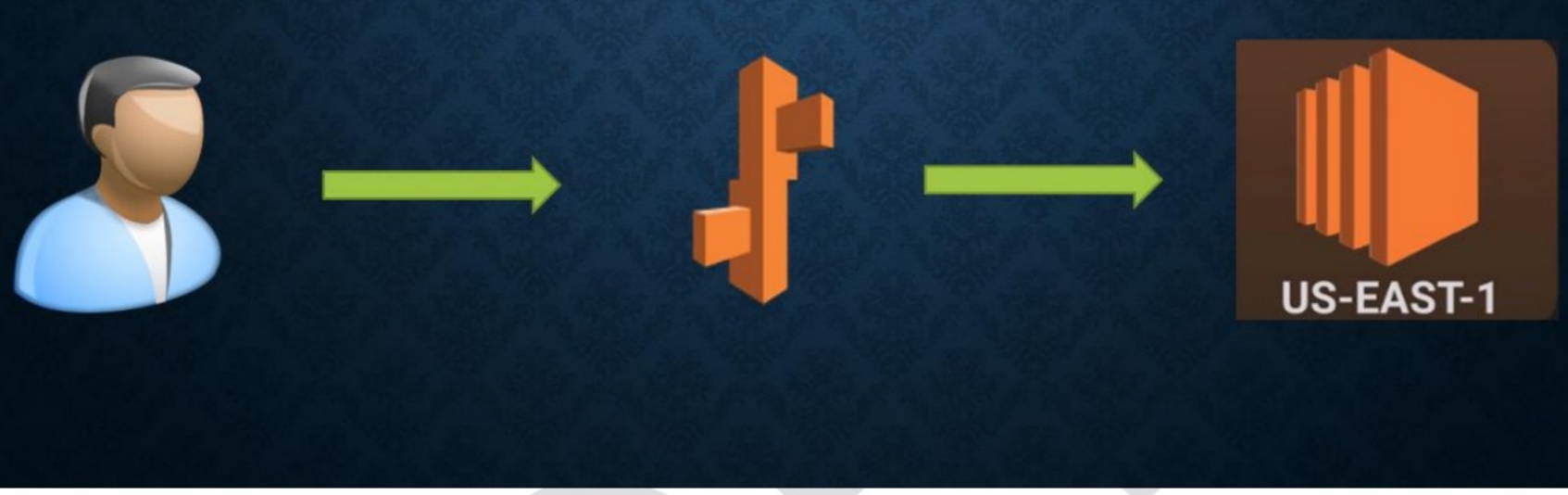
  .edu.in

Route53 Routing Policies

- Simple
- Weighted
- Latency
- Failover
- Geolocation

Simple Routing Policy

- This is the default routing policy.
- This is most commonly used when you have a single region that performs a given function for your domain

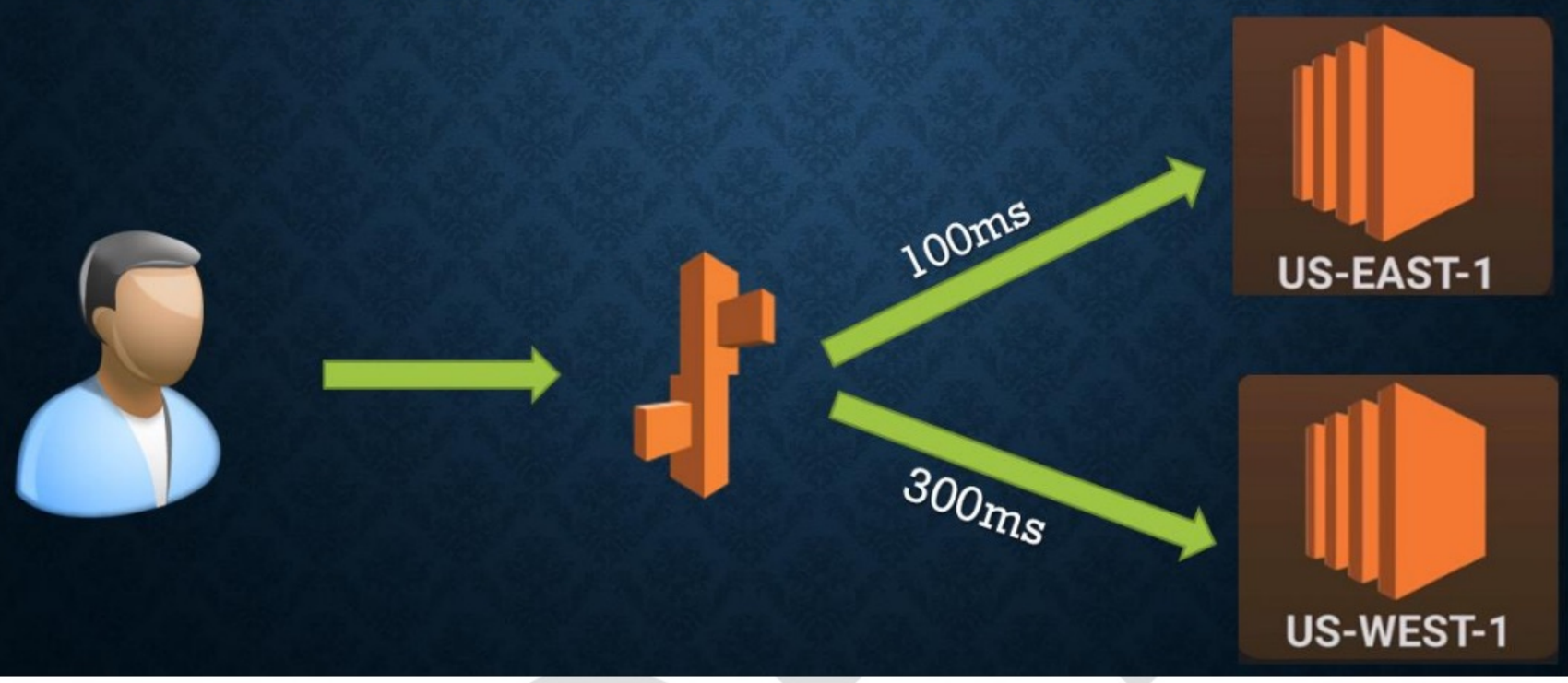


Weighted Routing Policy

- Weighted Routing Policies let you split your traffic based on different weights assigned.
- eg: 20% of your traffic to go to US-EAST-1
- 80% of your traffic to go to EU-WEST-1

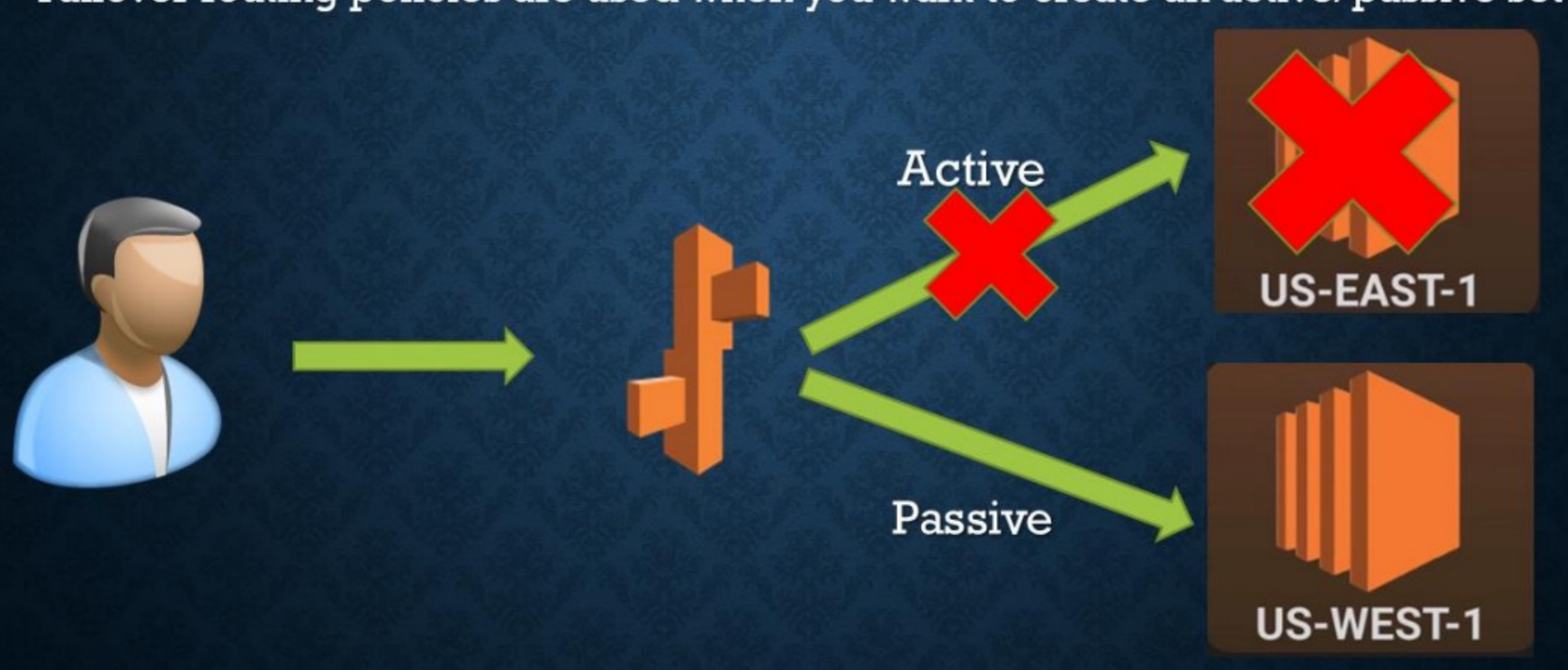
Latency Routing Policy

- Latency based routing allows you to route your traffic based on the lowest network latency for your end user (ie which region will give them the fastest response time)



Failover Routing Policy

- Failover routing policies are used when you want to create an active/passive set up.



- eg: you may want your primary site to be in US-EAST-1 and your secondary DR Site in US-WEST-1
- Route53 will monitor the health of your primary site using a health check

S3 Storage Classes/Tiers

- S3 Standard
- S3 Intelligent Tiering
- S3 Standard IA
- S3 One Zone-IA (Infrequently Access)
- S3 Glacier (To get data, need to wait for 2-5 hours)
- S3 Glacier Deep Archive (To get data, need to wait for 12 hours)

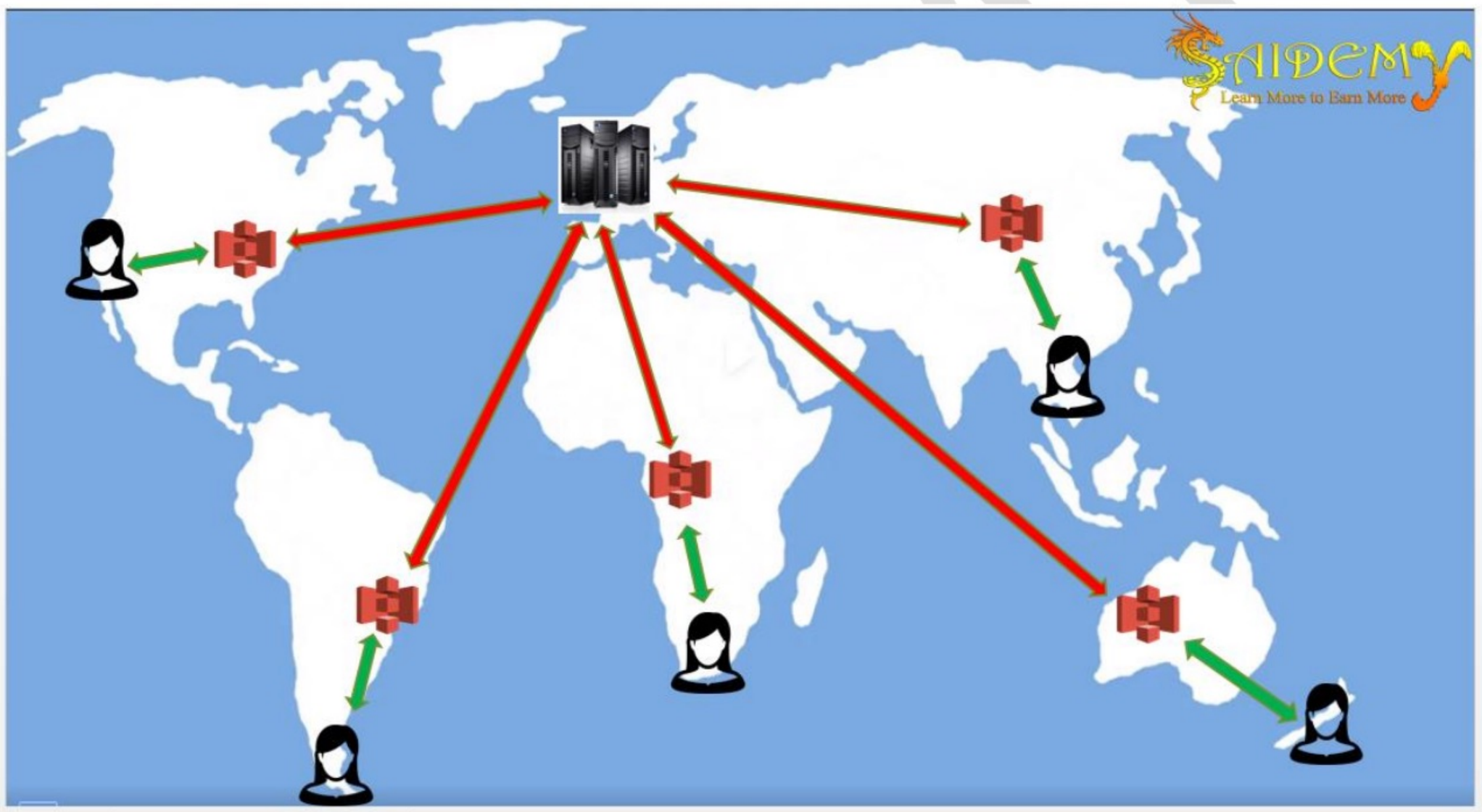
S3 - Charges

- Storage
- Requests
- Storage Management Pricing (Tiers)
- Data Transfer Pricing (Manual)
- Transfer Acceleration
- Cross Region Replication (Automatic)

	Storage pricing
S3 Standard - General purpose storage for any type of data, typically used for frequently accessed data	
First 50 TB / Month	$0.023 per GB
Next 450 TB / Month	$0.022 per GB
Over 500 TB / Month	$0.021 per GB
S3 Intelligent - Tiering * - Automatic cost savings for data with unknown or changing access patterns	
Frequent Access Tier, First 50 TB / Month	$0.023 per GB
Frequent Access Tier, Next 450 TB / Month	$0.022 per GB
Frequent Access Tier, Over 500 TB / Month	$0.021 per GB
Infrequent Access Tier, All Storage / Month	$0.0125 per GB
Monitoring and Automation, All Storage / Month	$0.0025 per 1,000 objects
S3 Standard - Infrequent Access * - For long lived but infrequently accessed data that needs millisecond access	
All Storage / Month	$0.0125 per GB
S3 One Zone - Infrequent Access * - For re-createable infrequently accessed data that needs millisecond access	
All Storage / Month	$0.01 per GB
S3 Glacier ** - For long-term backups and archives with retrieval option from 1 minute to 12 hours	
All Storage / Month	$0.004 per GB
S3 Glacier Deep Archive ** - For long-term data archiving that is accessed once or twice in a year and can be restored within 12 hours	
All Storage / Month	$0.00099 per GB

Transfer Acceleration

- S3 Transfer Acceleration enables fast, easy and secure transfers of files over long distances between your end users and as S3 bucket.
- Transfer Acceleration takes advantage of Amazon CloudFront's globally distributed edge locations.
- As the data arrives at an edge location, data is routed to Amazon S3 over an optimized network path.



S3 (Simple Storage Service)

- S3 is a secure, durable and highly-scalable object storage. S3 is easy to use, with a simple web service interface to store and retrieve any amount of data from anywhere on the web

Important Points

- S3 is a safe place to store your files.
- It is Object based storage.
- Files can be from 0 Bytes to 5 TB (Graphical/Console)
- Files can be from 0 Bytes to 5 GB (CLI)
- There is unlimited storage
- Objects are stored in buckets.
- S3 bucket names must be unique globally

Other Advantages of S3

- Built for 99.99% availability for the S3 platform
- Amazon Guarantee 99.999999999% durability
- Tiered Storage Available
- Lifecycle Management
- Versioning
- Encryption
- Secure your data using Access control lists & Bucket policy

SES (Simple Email Service)

- SES (Simple Email Service) is a email service designed to send and receive marketing, notification and transactional emails to their customers using a pay as you go model.

- Automated emails.
- Purchase confirmations, shipping notifications and order status updates.
    - e.g. when we do online purchase, we get mails regarding confirmation, status, expected delivery .....
- Marketing communications, advertisements, newsletters, special offers.

SES

- Email messaging service
- Can be used for both incoming and outgoing email
- An email address is all that is required to send emails.

SNS

- Messaging service formats include SMS, HTTP, HTTPS, Phone messages, email and email JSON.
- Can fan out messages to large no of recipients

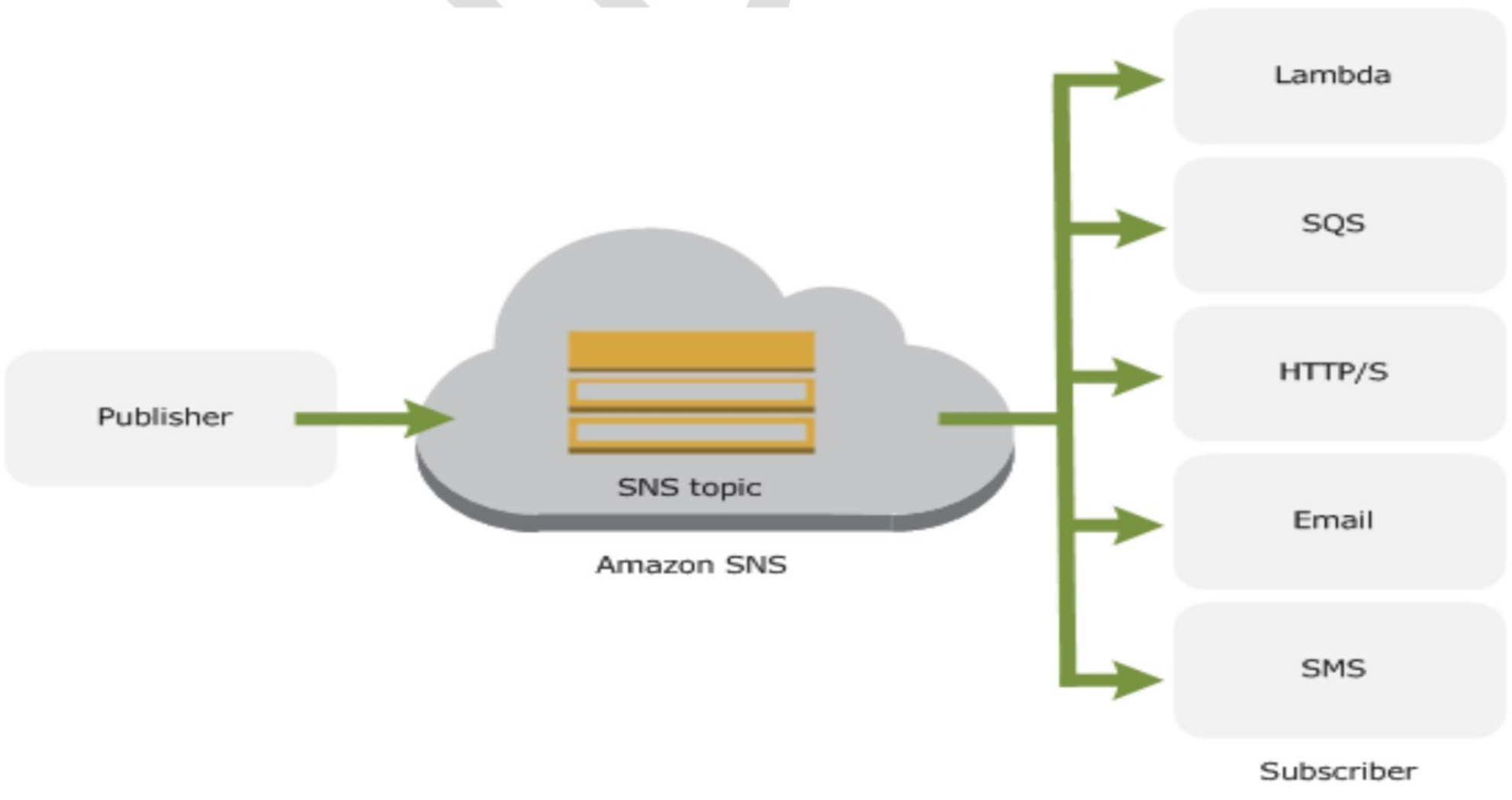
- Employees must subscribe to a topic to receive the notifications.

Important Points to Note

- Remember that SES is for email only
- SES can be used for incoming and outgoing mail
- SES is not subscription based, you only need to know the email address.
- SNS supports multiple formats (SMS, HTTP, email)
- Push notification only (SNS & SES)
- Consumers must subscribe to a topic (SNS)
- You can fan-out messages to large number of recipients (SNS)

SNS (Simple Notification Service)

- Amazon SNS (Simple Notification Service) is a web service that makes it easy to send notifications from cloud.
- It provides users with a highly scalable, flexible, and cost-effective capability to publish messages and immediately deliver them to subscribers.
- SNS allows you to group multiple recipients using topics. A topic is an "access point" for allowing recipients to dynamically subscribe for identical copies of the same notification.
- Besides pushing cloud notifications directly to mobile devices, SNS can also deliver notifications by SMS text message or email or any end point.
- To prevent messages from being lost, all messages published to SNS are stored redundantly across multiple availability zones.



SNS Benefits

- Instantaneous, push-based delivery (No polling)
- Flexible message delivery over multiple transport protocols
- Inexpensive, pay-as-you-go model with no up-front costs
- Web-based AWS Management Console offers the simplicity of a point-and-click interface.

SQS vs SNS

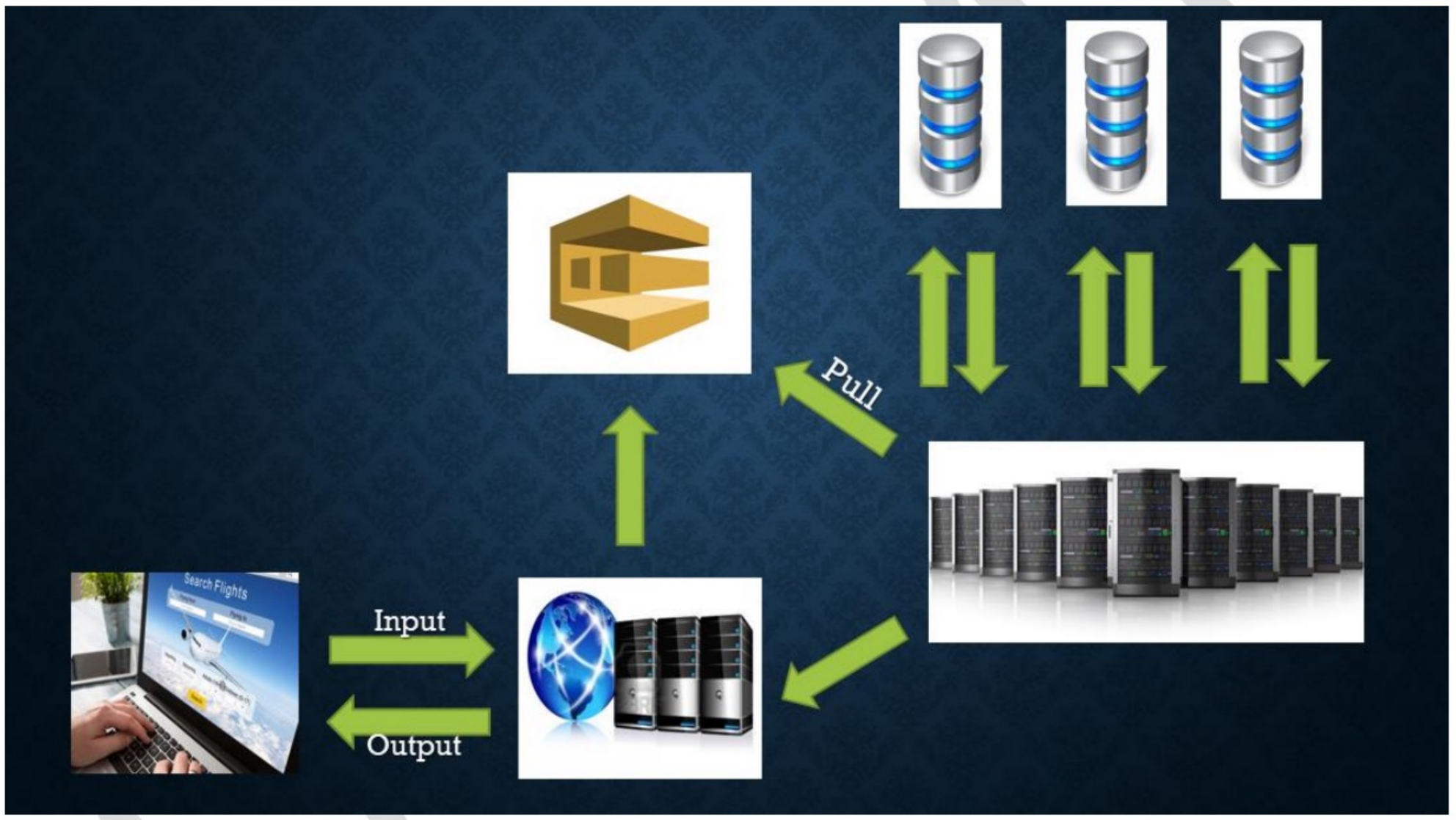
- Both Messaging Services in AWS
- SNS - Push
- SQS - Polls (Pulls)

CAIDEMY

SQS

(SIMPLE QUEUE SERVICE)

- SQS is a message queue used to store messages while waiting for a computer to process them
- Queue is a temporary repository for messages that are awaiting process



Queue Types

Standard Queue

- Default queue
- Messages can go out of order
- Unlimited no of transactions per second

FIFO (First-In-First-Out)

- Not Default queue
- Messages will go in exact order
- 300 Transactions per second

Important points to note

- SQS is pull based. not push based.
- Decoupling mechanism
- Messages can contain up to 256KB of text in any format
- Messages can be kept in the queue for 14 days max
- Default retention period is 4 days
- SQS guarantee that your messages will be processed at least once

Visibility Time Out

- The Visibility Timeout is the amount of time that the message is invisible in the SQS queue after a reader picks up that message. Provided the job is processed before the visibility time out expires, the message will then be deleted from the queue. If the job is not processed within that time, the message will become visible again and another reader will process it.
- Default Visibility Timeout is 30 Seconds

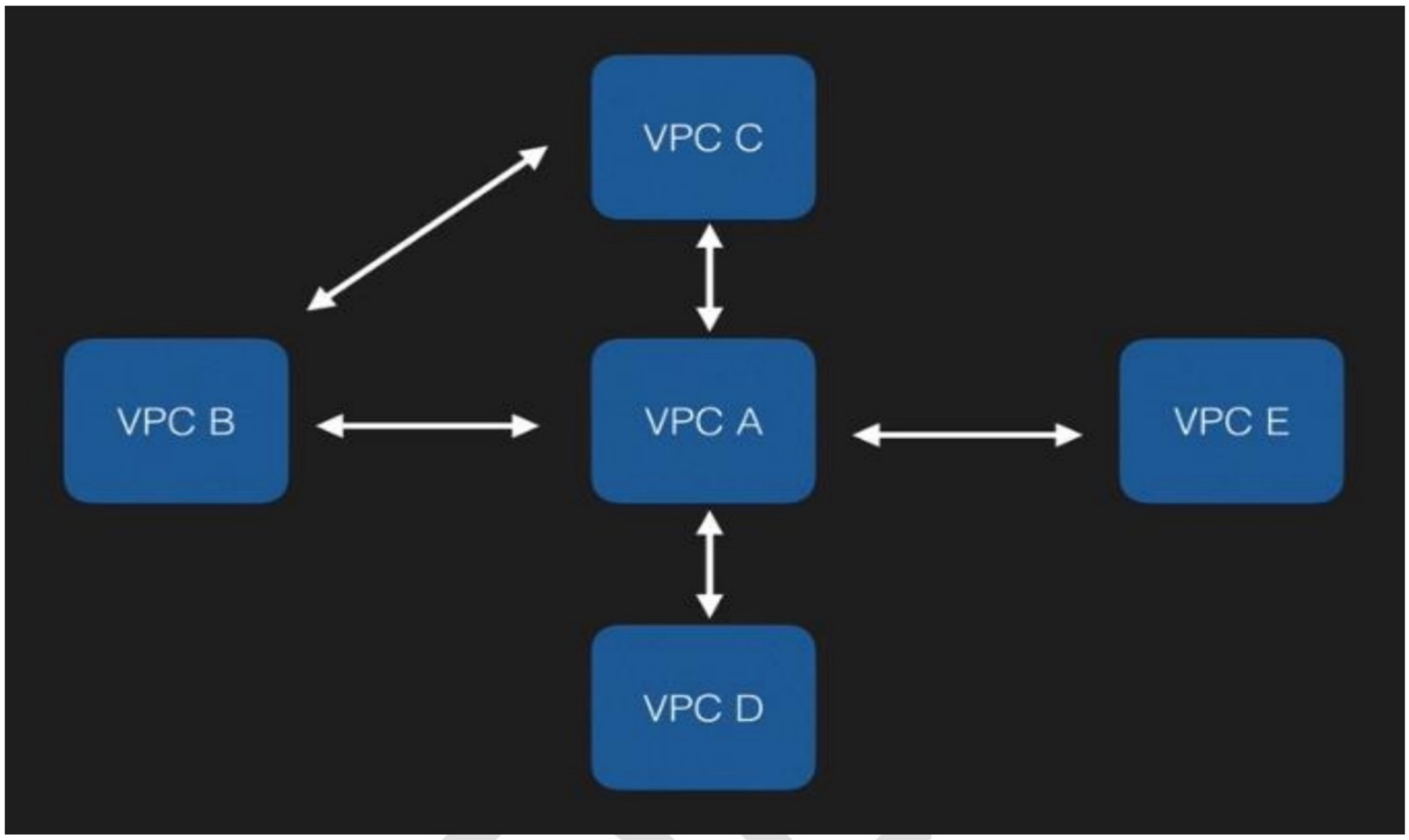
- Increase it if your task takes >30 seconds
- Maximum is 12 hours

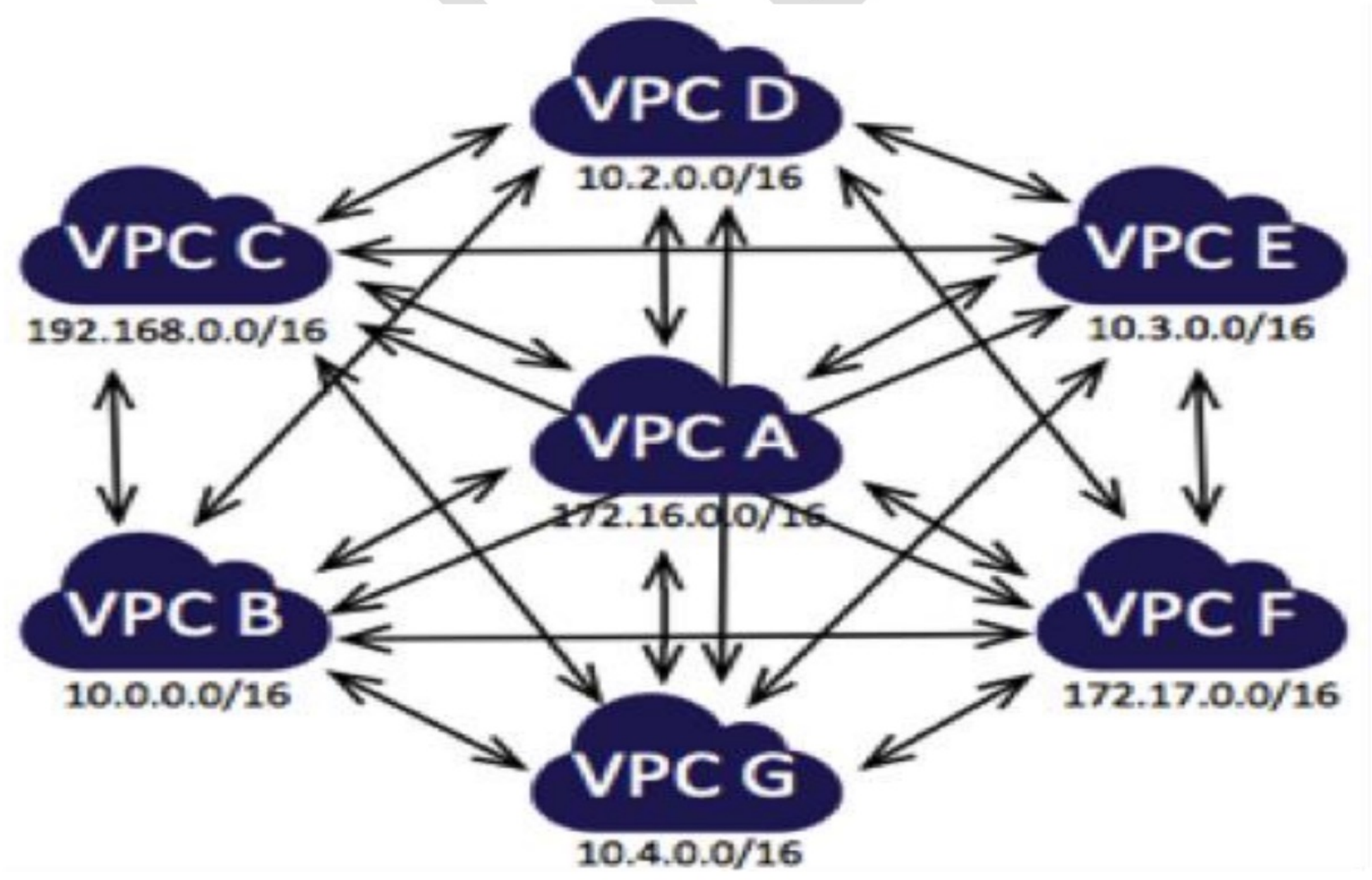
SQS Long Polling

- It is a way to retrieve messages from your SQS Queue
- Polls the queue periodically as per the time interval we set.
- Long polling can save your money.

CAIDEMY

VPC Peering

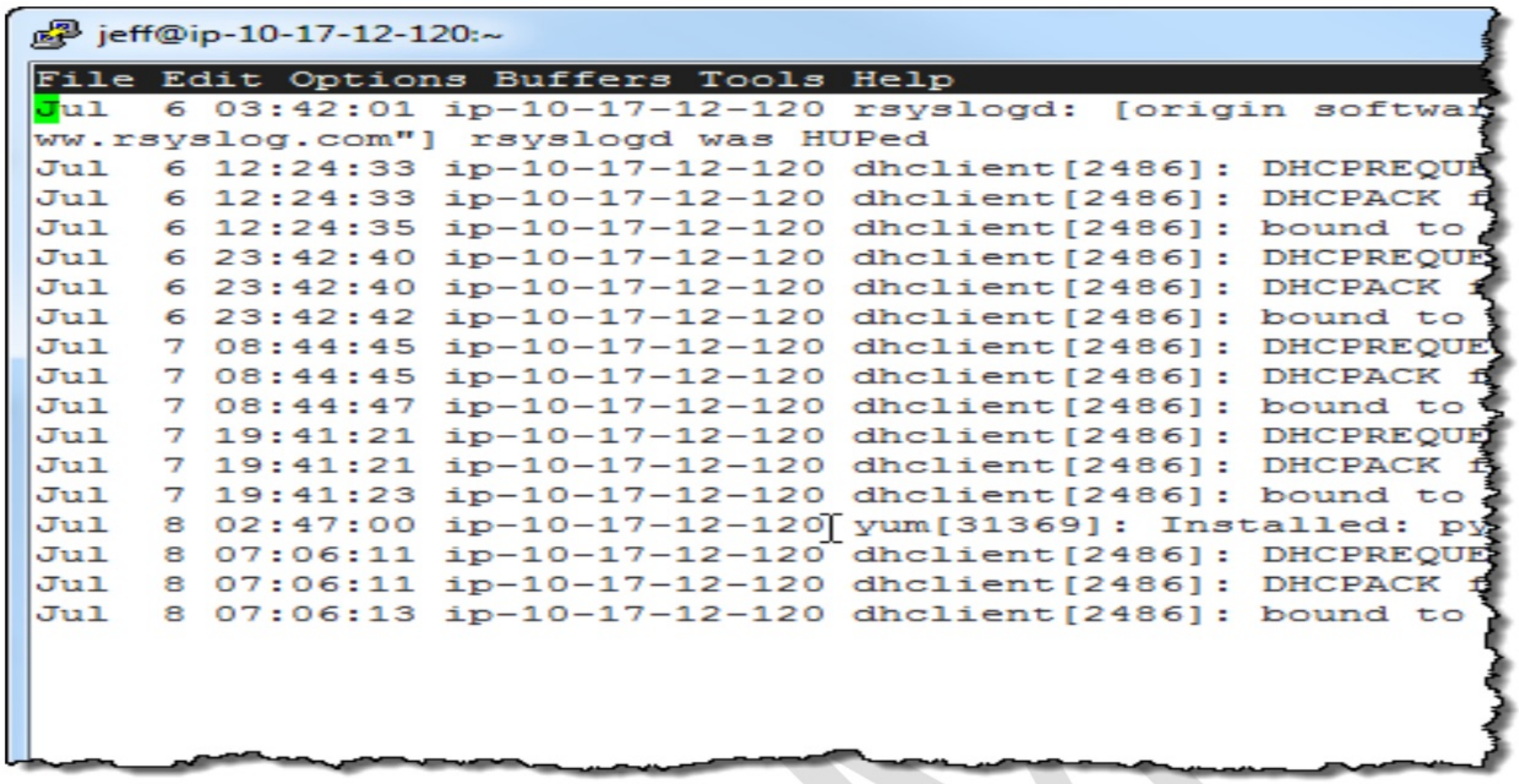

VPC C
VPC B
VPC A
VPC E
VPC D


VPC D
10.2.0.0/16
VPC C
192.168.0.0/16
VPC E
10.3.0.0/16
VPC A
172.16.0.0/16
VPC B
10.0.0.0/16
VPC F
172.17.0.0/16
VPC G
10.4.0.0/16

- Allows you to connect one VPC with another via a direct network route using private IP addresses.
- Instances behave as if they are on the same private network
- You can peer VPC's with other AWS accounts as well as with other VPCs in the same account
- NO TRANSITIVE PEERING!!

VPC Flow Logs

- VPC Flow logs is a feature that enables you to capture information about the IP traffic going to and from network interfaces in your VPC.
- It's a way of tracking all network traffic to and from within your VPC.
- Flow logs can be created at 2 levels;
  - VPC
  - Subnet



VPC Flow Logs limitations

- You cannot enable flow logs for VPCs that are peered with your VPC unless the peer VPC is in your account
- You cannot tag a flow log.

Not all IP Traffic is monitored

- Traffic generated by instances when they contact the Amazon DNS server.
- Traffic to and from 169.254.169.254 for instance metadata.
- DHCP traffic
- Traffic to the reserved IP address

VPC End Points

- It is AWS private network used to connect EC2 instance which is there in private SN of VPC with S3 bucket to transfer the data.

SRI CADEMY

VPC (Virtual Private Cloud)

- VPC is a virtual data centre in the cloud
- VPC lets you provision a logically isolated section of the Amazon Web Services cloud where you can launch AWS resources in a virtual network that you define.

Important Points

- VPC consists of IGWs, Route Tables, NACL, Subnets and Security Groups
- 1 Subnet = 1 AZ
- Security Groups are Stateful
- NACLs are Stateless
- Your VPC automatically comes with a default network ACL and by default it allows all outbound and inbound traffic
- You can create a custom network ACL. By default, each custom network ACL denies all inbound and outbound traffic until you add rules.
- Each subnet in your VPC must be associated with a network ACL. If you don't explicitly associated a subnet with a network ACL, the subnet is automatically associated with the default network ACL

- You can associate a network ACL with multiple subnets; however, a subnet can be associated with only one network ACL at a time. When you associate a network ACL with a subnet, the previous association is removed
- A network ACL contains a numbered list of rules that is evaluated in order, starting with the lowest numbers rule
- A network ACL has separate inbound and outbound rules, and each rule can be either allow or deny traffic.

Security Group vs Network ACL

Security Group	Network ACL
Operates at the instance level	Operates at the subnet level
Supports allow rules only	Supports allow rules and deny rules
Is stateful: Return traffic is automatically allowed, regardless of any rules	Is stateless: Return traffic must be explicitly allowed by rules
We evaluate all rules before deciding whether to allow traffic	We process rules in number order when deciding whether to allow traffic
Applies to an instance only if someone specifies the security group when launching the instance, or associates the security group with the instance later on	Automatically applies to all instances in the subnets it's associated with (therefore, you don't have to rely on users to specify the security group)

WebServer

1. It is a normal server having web package installed in it.
2. It's OS could be either Windows or Linux.
3. To make any server as web server, three things we need to do
    - Install web package (HTTPD)
    - Create a file inside html directory called "index.html" and put some content whatever you want to expose to your customers.
    - Start web service (HTTPD)
4. Make sure that you have opened http (80) port. Otherwise you can't able to access website.

WebServer Commands

- sudo su -
- yum update -y
- yum install httpd -y
- cd /var/www/html
- echo "MyGoogle" > index.html
- ls
- service httpd start
- chkconfig httpd on

Whitepapers

Whitepapers & Well Architected Framework

1. Security
2. Reliability
3. Performance Efficiency
4. Cost Optimization
5. Operational Excellence

Security

- 24/7 Electronic Surveillance & Multi-Factor Access control system
- 24/7 Security Guards
- Access is authorized on “Least Privilege Basis”

Compliance

- PCI DSS (For billing)
- ISO 27001
- ISO 9001

- VPN

Amazon Corporate Segregation

Network Monitoring & Protection

- Port Scanning

AWS Credentials

- Passwords
- MFA
- Access & Secret Keys
- Key Pair

AWS Trusted Advisor

Security Considerations

- OS (You are Admin. Not AWS)
- Firewall
- ELB
- Direct Connect

Strategic Business Plan

- Bi-Annually

Design for Failure (Be a Pessimistic)

Decouple Your Components

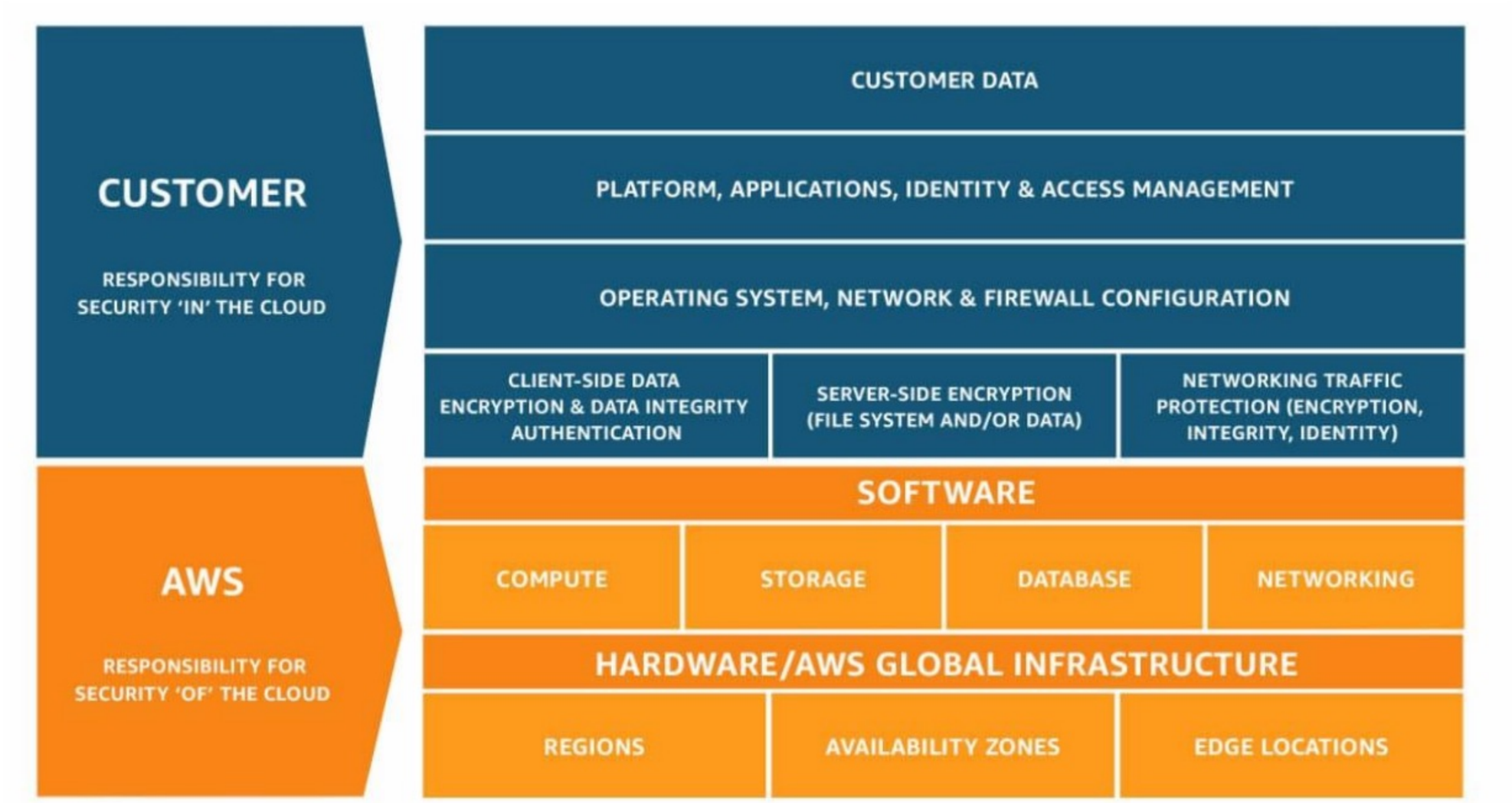
Implement Elasticity

- Proactive Cyclic Scaling
- Proactive Event-based Scaling
- Auto-Scaling based on demand

Scale Up & Scale Out

CADEMY

Shared Responsibility Model



Customer Security Responsibilities

- IAAS
- PAAS
- SAAS

Storage De-commissioning

- **Deleting (Disk Zeroing)**
- **Destroying**

Network Security

- HTTPS
- VPC

AWS Certification Questions

You have an application running in us-west-2 requiring 6 EC2 Instances running at all times. With 3 Availability Zones in the region viz. us-west-2a, us-west-2b, and us-west-2c, which of the following deployments provides fault tolerance if an Availability Zone in us-west-2 becomes unavailable? Choose 2 answers from the options given below.

Please select :

- A. 2 EC2 Instances in us-west-2a, 2 EC2 Instances in us-west-2b, and 2 EC2 Instances in us-west-2c
- B. 3 EC2 Instances in us-west-2a, 3 EC2 Instances in us-west-2b, and no EC2 Instances in us-west-2c
- C. 4 EC2 Instances in us-west-2a, 2 EC2 Instances in us-west-2b, and 2 EC2 Instances in us-west-2c
- D. 6 EC2 Instances in us-west-2a, 6 EC2 Instances in us-west-2b, and no EC2 Instances in us-west-2c
- E. 3 EC2 Instances in us-west-2a, 3 EC2 Instances in us-west-2b, and 3 EC2 Instances in us-west-2c

---

A Solutions Architect is developing a document sharing application and needs a storage layer. The storage should provide automatic support for versioning so that users can easily roll back to a previous version or recover a deleted account.

Which AWS service will meet the above requirements?

Please select :

- A. Amazon S3
- B. Amazon EBS
- C. Amazon EFS
- D. Amazon Storage Gateway VTL

You have an application running in us-west-2 requiring 6 EC2 Instances running at all times. With 3 Availability Zones in the region viz. us-west-2a, us-west-2b, and us-west-2c, which of the following deployments provides fault tolerance if an Availability Zone in us-west-2 becomes unavailable?
Choose 2 answers from the options given below.

Please select :

- A. 2 EC2 Instances in us-west-2a, 2 EC2 Instances in us-west-2b, and 2 EC2 Instances in us-west-2c
- B. 3 EC2 Instances in us-west-2a, 3 EC2 Instances in us-west-2b, and no EC2 Instances in us-west-2c
- C. 4 EC2 Instances in us-west-2a, 2 EC2 Instances in us-west-2b, and 2 EC2 Instances in us-west-2c
- D. 6 EC2 Instances in us-west-2a, 6 EC2 Instances in us-west-2b, and no EC2 Instances in us-west-2c
- E. 3 EC2 Instances in us-west-2a, 3 EC2 Instances in us-west-2b, and 3 EC2 Instances in us-west-2c

An administrator runs a highly available application in AWS. A file storage layer is needed that can share between instances and scale the platform more easily.
Which AWS service can perform this action?

Please select :

- A. Amazon EBS
- B. Amazon EFS
- C. Amazon S3
- D. Amazon EC2 Instance store

A Solutions Architect is designing a highly scalable system to track records. These records must remain available for immediate download for up to three months and then must be deleted.
What is the most appropriate decision for this use case?

Please select :

- A. Store the files in Amazon EBS and create a Lifecycle Policy to remove files after 3 months.
- B. Store the files in Amazon S3 and create a Lifecycle Policy to remove files after 3 months.
- C. Store the files in Amazon Glacier and create a Lifecycle Policy to remove files after 3 months.
- D. Store the files in Amazon EFS and create a Lifecycle Policy to remove files after 3 months.

An application currently stores all its data on Amazon EBS Volumes. All EBS volumes must be backed up durably across multiple Availability Zones.
What is the MOST resilient way to backup the volumes?

Please select :

- A. Take regular EBS snapshots.
- B. Enable EBS volume encryption.
- C. Create a script to copy data to an EC2 Instance store.
- D. Mirror data across 2 EBS volumes.

There is a requirement to host a database on an EC2 Instance. It is also required that the EBS volume should support 12,000 IOPS.

Which Amazon EBS volume type meets the performance requirements of this database?

Please select :

- A. EBS Provisioned IOPS SSD
- B. EBS Throughput Optimized HDD
- C. EBS General Purpose SSD
- D. EBS Cold HDD

Development teams in your organization use S3 buckets to store log files for various applications hosted in AWS development environments. The developers intend to keep the logs for a month for troubleshooting purposes, and subsequently purge the logs.

What feature will enable this requirement?

Please select :

- A. Adding a bucket policy on the S3 bucket.
- B. Configuring lifecycle configuration rules on the S3 bucket.
- C. Creating an IAM policy for the S3 bucket.
- D. Enabling CORS on the S3 bucket.

What options can be used to host an application that uses NGINX and is scalable at any point in time?

Choose 2 correct answers.

Please select :

- A. AWS EC2
- B. AWS Elastic Beanstalk
- C. AWS SQS
- D. AWS ELB

There is a requirement for Block-level storage to store 500GB of data. Data Encryption is also required.

Which of the following can be used in such a case?

Please select :

- A. AWS EBS Volumes
- B. AWS S3
- C. AWS Glacier
- D. AWS EFS

An application needs to access data in another AWS account in the same region. Which of the following can be used to ensure that the data can be accessed as required?

Please select :

- A. Establish a NAT instance between both accounts.
- B. Use a VPN between both accounts.
- C. Use a NAT Gateway between both accounts.
- D. Use VPC Peering between both accounts.

There is a requirement for EC2 Instances in a private subnet to access an S3 bucket. It is required that the traffic does not traverse to the Internet. Which of the following can be used to fulfill this requirement?

Please select :

- A. VPC Endpoint
- B. NAT Instance
- C. NAT Gateway
- D. Internet Gateway

A database is being hosted using the AWS RDS service. This database is to be made into a production database and is required to have high availability. Which of the following can be used to achieve this requirement?

Please select :

- A. Use Multi-AZ for the RDS instance to ensure that a secondary database is created in another region.
- B. Use the Read Replica feature to create another instance of the DB in another region.
- C. Use Multi-AZ for the RDS instance to ensure that a secondary database is created in another Availability Zone.
- D. Use the Read Replica feature to create another instance of the DB in another Availability Zone.

A company wants to host a web application and a database layer in AWS. This will be done with the use of subnets in a VPC.
Which of the following is a proper architectural design for supporting the required tiers of the application?

Please select :

- A. Use a public subnet for the web tier and a public subnet for the database layer.
- B. Use a public subnet for the web tier and a private subnet for the database layer.
- C. Use a private subnet for the web tier and a private subnet for the database layer.
- D. Use a private subnet for the web tier and a public subnet for the database layer.

A company has a requirement for archival of 6TB of data. There is an agreement with the stakeholders for an 8-hour agreed retrieval time. Which of the following can be used as the MOST cost-effective storage option?

Please select :

- A. AWS S3 Standard
- B. AWS S3 Infrequent Access
- C. AWS Glacier
- D. AWS EBS Volumes

Your company has a requirement to host a static web site in AWS. Which of the following steps would help implement a quick and cost-effective solution for this requirement? Choose 2 answers from the options given below. Each answer forms a part of the solution.

Please select :

- A. Upload the static content to an S3 bucket.
- B. Create an EC2 Instance and install a web server.
- C. Enable web site hosting for the S3 bucket.
- D. Upload the code to the web server on the EC2 Instance.

A company currently storing a set of documents in the AWS Simple Storage Service, is worried about the potential loss if these documents are ever deleted. Which of the following can be used to ensure protection from loss of the underlying documents in S3?

Please select :

- A. Enable Versioning for the underlying S3 bucket.
- B. Copy the bucket data to an EBS Volume as a backup.
- C. Create a Snapshot of the S3 bucket.
- D. Enable an IAM Policy which does not allow deletion of any document from the S3 bucket.

A company has a set of EC2 Linux based instances hosted in AWS. There is a need to have a standard file interface for files to be used across all Linux based instances. Which of the following can be used for this purpose?

Please select :

- A. Consider using the Simple Storage Service.
- B. Consider using Amazon Glacier.
- C. Consider using AWS RDS.
- D. Consider using AWS EFS.

Your company is planning on using Route 53 as the DNS provider. There is a need to ensure that the company's domain name points to an existing CloudFront distribution. How can this be achieved?

Please select :

- A. Create an Alias record which points to the CloudFront distribution.
- B. Create a host record which points to the CloudFront distribution.
- C. Create a CNAME record which points to the CloudFront distribution.
- D. Create a Non-Alias Record which points to the CloudFront distribution.

Your current setup in AWS consists of the following architecture: 2 public subnets, one subnet which has web servers accessed by users across the Internet and another subnet for the database server. Which of the following changes to the architecture adds a better security boundary to the resources hosted in this setup?

Please select :

- A. Consider moving the web server to a private subnet.
- B. Consider moving the database server to a private subnet.
- C. Consider moving both the web and database servers to a private subnet.
- D. Consider creating a private subnet and adding a NAT Instance to that subnet.

Instances in your private subnet hosted in AWS, need access to important documents in S3. Due to the confidential nature of these documents, you have to ensure that this traffic does not traverse through the internet. As an architect, how would you you implement this solution?

Please select :

- A. Consider using a VPC Endpoint.
- B. Consider using an EC2 Endpoint.
- C. Move the instances to a public subnet.
- D. Create a VPN connection and access the S3 resources from the EC2 Instance.

Instances in your private subnet hosted in AWS, need access to important documents in S3. Due to the confidential nature of these documents, you have to ensure that this traffic does not traverse through the internet. As an architect, how would you you implement this solution?

Please select :

- A. Consider using a VPC Endpoint.
- B. Consider using an EC2 Endpoint.
- C. Move the instances to a public subnet.
- D. Create a VPN connection and access the S3 resources from the EC2 Instance.

You have a set of EC2 Instances that support an application. They are currently hosted in the US Region. In the event of a disaster, you need a way to ensure that you can quickly provision the resources in another region. How could this be accomplished? Choose 2 answers from the options given below.

Please select :

- [ ] A. Copy the underlying EBS Volumes to the destination region.
- [ ] B. Create EBS Snapshots and then copy them to the destination region.
- [ ] C. Create AMIs for the underlying instances.
- [ ] D. Copy the metadata for the EC2 Instances to S3.

A company wants to have a NoSQL database hosted on the AWS Cloud, but do not have the necessary staff to manage the underlying infrastructure. Which of the following choices would be ideal for this requirement?

Please select :

- ( ) A. AWS Aurora
- ( ) B. AWS RDS
- ( ) C. AWS DynamoDB
- ( ) D. AWS Redshift

You plan on creating a VPC from scratch and launching EC2 Instances in the subnet. What should be done to ensure that the EC2 Instances are accessible from the Internet?

Please select :

- A. Attach an Internet Gateway to the VPC and add a route for 0.0.0.0/0 to the Route table.
- B. Attach an NAT Gateway to the VPC and add a route for 0.0.0.0/0 to the Route table.
- C. Attach an NAT Gateway to the VPC and add a route for 0.0.0.0/32 to the Route table.
- D. Attach an Internet Gateway to the VPC and add a route for 0.0.0.0/32 to the Route table.

Your company currently has an entire data warehouse of assets that needs to be migrated to the AWS Cloud. Which of the following services should this be migrated to?

Please select :

- A. AWS DynamoDB
- B. AWS S3
- C. AWS RDS
- D. AWS Redshift

Your company has confidential documents stored in the Simple Storage Service. Due to compliance requirements, there is a need for the data in the S3 bucket to be available in a different geographical location. As an architect, what change would you make to comply with this requirement?

Please select :

- A. Apply Multi-AZ for the underlying S3 bucket.
- B. Copy the data to an EBS Volume in another region.
- C. Create a snapshot of the S3 bucket and copy it to another region.
- D. Enable Cross-Region Replication for the S3 bucket.