

Understand the difference between a region, an Availability Zone (AZ) and an Edge Location.

- A Region is a physical location in the world which consists of two or more Availability Zones (AZ's).
- An AZ is one or more discrete data centers, each with redundant power, networking and connectivity, housed in separate facilities.
- Edge Locations are endpoints for AWS which are used for caching content. Typically this consists of CloudFront, Amazon's Content Delivery Network (CDN)

What Have We Learnt So Far?

- **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.** You cannot use the Access key ID & Secret Access Key to Login in to the console. You can use this to access AWS via the APIs and Command Line, however.
- **You only get to view these once.** If you lose them, you have to regenerate them. So, save them in a secure location.



What Have We Learnt So Far?

- Always setup Multifactor Authentication on your root account.
- You can create and customise your own password rotation policies.



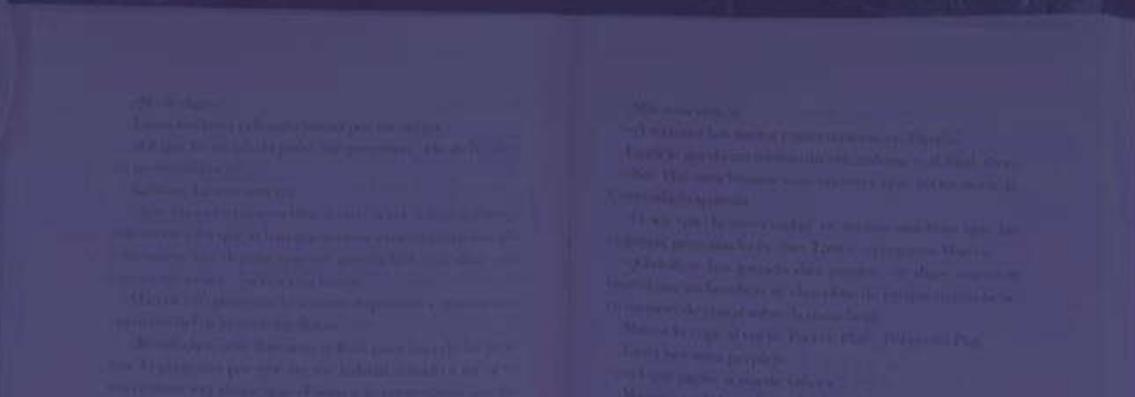
The Key Fundamentals of S3 Are;

- Key (This is simply the name of the object)
- Value (This is simply the data and is made up of a sequence of bytes).
- Version ID (Important for versioning)
- Metadata (Data about data you are storing)
- Subresources;

Access Control Lists

Torrent

- Read after Write consistency for PUTS of new Objects
- Eventual Consistency for overwrite PUTS and DELETES (can take some time to propagate)



Exam Tips

1

S3 Standard

99.99% availability
99.99999999% durability,
stored redundantly across
multiple devices in multiple
facilities, and is designed to
sustain the loss of 2 facilities
concurrently.

4

S3 - Intelligent Tiering

Designed to optimize costs
by automatically moving
data to the most cost-
effective access tier, without
performance impact or
operational overhead.

2

S3 - IA

(Infrequently Accessed):
For data that is accessed
less frequently, but requires
rapid access when needed.
Lower fee than S3, but you
are charged a retrieval fee.

3

S3 One Zone - IA

For where you want a
lower-cost option for
infrequently accessed data,
but do not require the
multiple Availability Zone
data resilience.

5

S3 Glacier

S3 Glacier is a secure, durable,
and low-cost storage class for
data archiving. Retrieval times
configurable from minutes to
hours.

6

S3 Glacier Deep Archive

S3 Glacier Deep Archive is
Amazon S3's lowest-cost
storage class where a
retrieval time of 12 hours is
acceptable.

- Read the S3 FAQs before taking the exam. It comes up A LOT!





- Buckets are a universal name space
- Upload an object to S3 receive a **HTTP 200 Code**
- S3, S3 - IA, S3 - IA (One Zone), Glacier

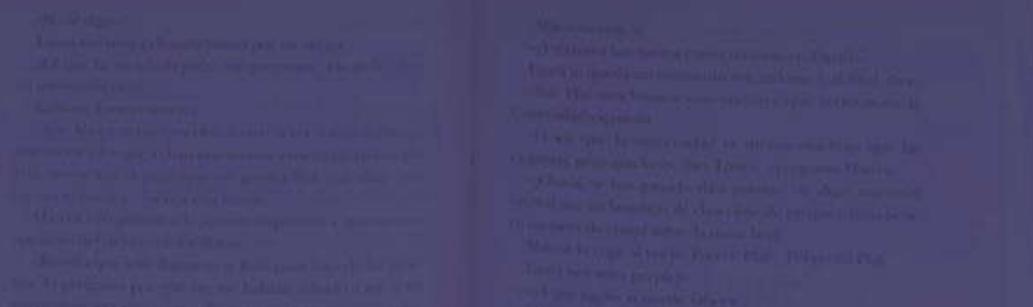
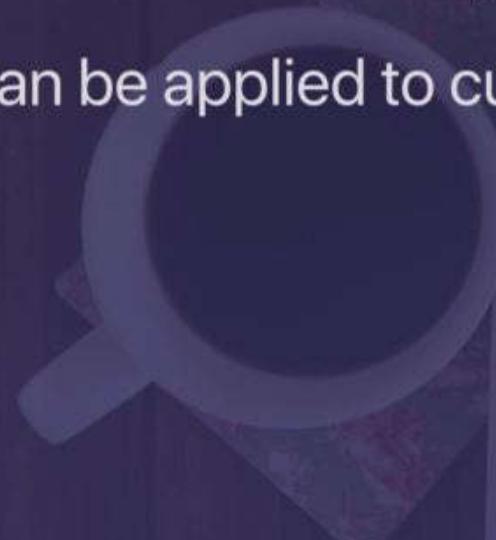


- Control access to buckets using either a **bucket ACL** or using **Bucket Policies**



- Stores all versions of an object (including all writes and even if you delete an object)
- Great backup tool.
- Once enabled, Versioning cannot be disabled, only suspended.
- Integrates with Lifecycle rules
- Versioning's MFA Delete capability, which uses multi-factor authentication, can be used to provide an additional layer of security.

- Automates moving your objects between the different storage tiers.
- Can be used in conjunction with versioning.
- Can be applied to current versions and previous versions.



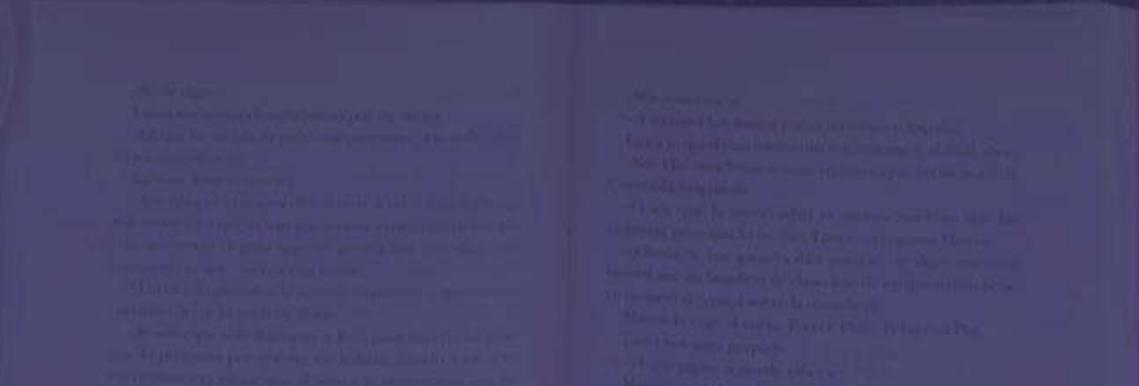
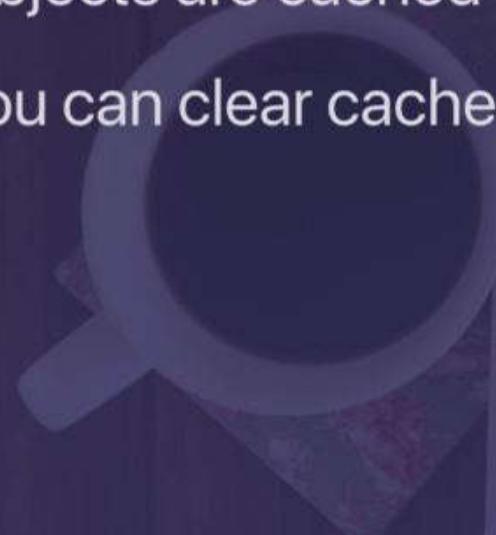
- Versioning must be enabled on both the source and destination buckets.
- Regions must be unique.
- Files in an existing bucket are not replicated automatically.
- All subsequent updated files will be replicated automatically.
- Delete markers are not replicated.
- Deleting individual versions or delete markers will not be replicated.



- **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 the 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 the CDN which consists of a collection of Edge Locations.
- **Web Distribution** - Typically used for Websites.
- **RTMP** - Used for Media Streaming.



- Edge locations are not just READ only — you can write to them too. (ie put an object on to them.)
- Objects are cached for the life of the **TTL (Time To Live.)**
- You can clear cached objects, but you will be charged.





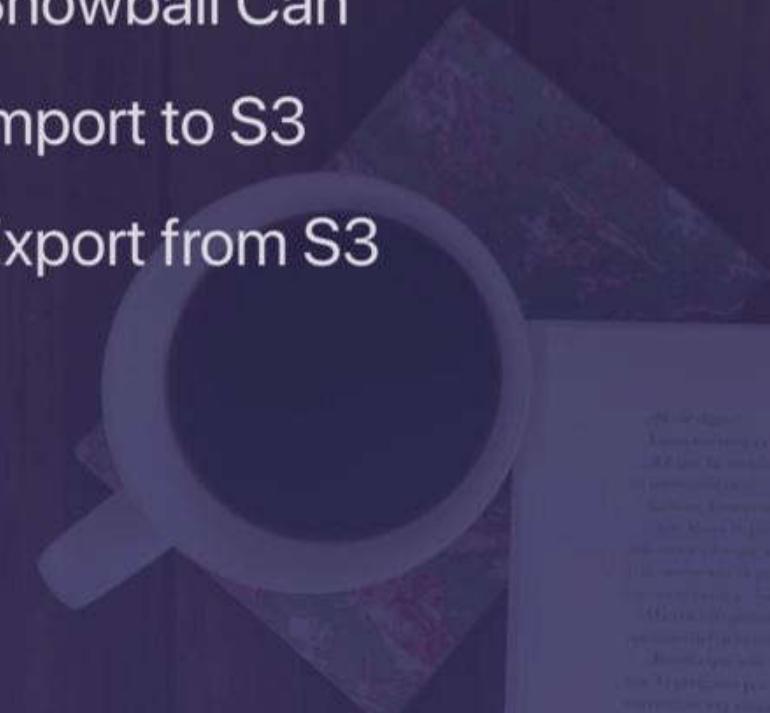
- **Edge Location** - This is the location where content will be cached. This is separate to an AWS Region/AZ.
- **Origin** - This is the origin of all the files that the CDN will distribute. This can be either an S3 Bucket, an EC2 Instance, an Elastic Load Balancer, or Route53.
- **Distribution** -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



- Edge locations are not just READ only — you can write to them too. (ie put an object on to them.)
- Objects are cached for the life of the **TTL (Time To Live.)**
- You can invalidate cached objects, but you will be charged.



- Understand what Snowball is
- Snowball Can
- Import to S3
- Export from S3



QUESTION

Amazon Snowball is a durable device used for moving large amounts of data between your on-premises environment and Amazon S3. It can import data from S3 and export data to S3.

QUESTION

Amazon Snowball is a durable device used for moving large amounts of data between your on-premises environment and Amazon S3. It can import data from S3 and export data to S3.

ANSWER

Amazon Snowball is a durable device used for moving large amounts of data between your on-premises environment and Amazon S3. It can import data from S3 and export data to S3.

ANSWER

Amazon Snowball is a durable device used for moving large amounts of data between your on-premises environment and Amazon S3. It can import data from S3 and export data to S3.



File Gateway

- File Gateway - For flat files, stored directly on S3.

Volume Gateway

- **Stored Volumes** - Entire Dataset is stored on site and is asynchronously backed up to S3.
- **Cached Volumes** - Entire Dataset is stored on S3 and the most frequently accessed data is cached on site.

Gateway Virtual Tape Library

Identity Access Management Consists Of The Following;

- Users
- Groups
- Roles
- Policies

```
{  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Effect": "Allow",  
      "Action": "*",  
      "Resource": "*"  
    }  
  ]  
}
```

Exam Tips

What Have We Learnt So Far?

- **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.** You cannot use the Access key ID & Secret Access Key to Login in to the console. You can use this to access AWS via the APIs and Command Line, however.
- **You only get to view these once.** If you lose them, you have to regenerate them. So, save them in a secure location.



What Have We Learnt So Far?

- Always setup Multifactor Authentication on your root account.
- You can create and customise your own password rotation policies.



Amazon S3 is a highly reliable, low-cost, and flexible cloud storage service. It provides secure, durable, and fast access to your data. Amazon S3 offers a range of features, including multi-region support, server-side encryption, and automatic data replication across multiple regions. It also supports various file formats and provides a simple API for interacting with your data. Amazon S3 is used for storing and retrieving any amount of data at any time, from anywhere, securely and efficiently.



Amazon IAM is a highly secure, flexible, and cost-effective way to manage access to AWS services and resources. It provides a central place to define who has permission to do what, and where. IAM uses a fine-grained access control model, allowing you to specify exactly what actions a user or group can perform on specific resources. It also supports multi-factor authentication, which adds an extra layer of security to your accounts. IAM is used for managing access to AWS services, such as S3, Lambda, and CloudWatch, and for integrating with other AWS services, such as AWS Lambda and AWS Lambda@Edge.



- Remember that S3 is **Object-based**: i.e. allows you to upload files.
- Files can be from 0 Bytes to 5 TB.
- There is unlimited storage.
- Files are stored in Buckets.
- **S3 is a universal namespace.** That is, names must be unique globally.
- <https://s3-eu-west-1.amazonaws.com/acloudguru>

- Not suitable to install an operating system on.
- Successful uploads will generate a **HTTP 200** status code.



Amazon S3 is not suitable to install an operating system on. It is a storage service designed for storing and retrieving data at any time. It is not intended to be used as a primary storage solution for your application. Instead, it is designed to be used as a secondary storage solution for backup, archiving, and other non-primary purposes. Amazon S3 is highly reliable and can handle large amounts of data. It is also very cost-effective, making it a popular choice for many organizations. However, it is not suitable for use as a primary storage solution for your application. Instead, it is designed to be used as a secondary storage solution for backup, archiving, and other non-primary purposes. Amazon S3 is highly reliable and can handle large amounts of data. It is also very cost-effective, making it a popular choice for many organizations.

By default, all newly created buckets are **PRIVATE**. You can setup access control to your buckets using;

- **Bucket Policies**
- **Access Control Lists**

S3 buckets can be configured to create access logs which log all requests made to the S3 bucket. This can be sent to another bucket and even another bucket in another account.

The Key Fundamentals of S3 Are;

- Key (This is simply the name of the object)
- Value (This is simply the data and is made up of a sequence of bytes).
- Version ID (Important for versioning)
- Metadata (Data about data you are storing)
- Subresources;

Access Control Lists

Torrent



QUESTION 1: Which AWS service allows you to store data in an object-based format?

AWS S3 is an object storage service that allows you to store data in an object-based format. It provides a highly durable and reliable way to store and retrieve any amount of data, from a few bytes to several terabytes.

QUESTION 2: What is the primary purpose of AWS Lambda?

AWS Lambda is a serverless compute service that allows you to run code without provisioning or managing servers. It triggers your code in response to events, such as new data being added to an Amazon S3 bucket or a user logging in to an application.

QUESTION 3: How does AWS IAM manage access to AWS services?

AWS IAM (Identity and Access Management) is a service that provides secure, fine-grained access control for AWS services and other partner services. It uses a central set of policies and users to define who can do what in your AWS account. You can use IAM to grant specific permissions to individual users or groups, and then assign those groups to roles or attach them directly to users.

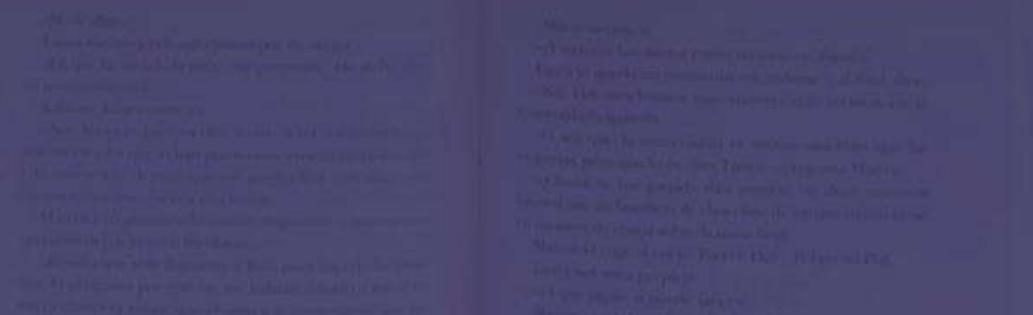
QUESTION 4: What is the difference between AWS S3 and AWS Glue?

AWS S3 is an object storage service that stores data in an object-based format. It is designed for storing large amounts of unstructured data, such as images, videos, and documents. AWS Glue, on the other hand, is a serverless ETL (Extract, Transform, Load) service that makes it easy to crawl, transform, and load data from various sources, such as databases, data lakes, and data warehouses, into AWS Data Lakes or data processing pipelines.

QUESTION 5: What is the difference between AWS Lambda and AWS Step Functions?

AWS Lambda is a serverless compute service that runs your code in response to events. It is designed for small, discrete tasks that don't require a full server. AWS Step Functions, on the other hand, is a serverless workflow service that allows you to build complex workflows by combining multiple Lambda functions and other AWS services. It provides a visual interface for defining the flow of data and control between different steps in a workflow.

- Read after Write consistency for PUTS of new Objects
- Eventual Consistency for overwrite PUTS and DELETES (can take some time to propagate)



Exam Tips

1

S3 Standard

99.99% availability
99.99999999% durability,
stored redundantly across
multiple devices in multiple
facilities, and is designed to
sustain the loss of 2 facilities
concurrently.

4

S3 - Intelligent Tiering

Designed to optimize costs
by automatically moving
data to the most cost-
effective access tier, without
performance impact or
operational overhead.

2

S3 - IA

(Infrequently Accessed):
For data that is accessed
less frequently, but requires
rapid access when needed.
Lower fee than S3, but you
are charged a retrieval fee.

3

S3 One Zone - IA

For where you want a
lower-cost option for
infrequently accessed data,
but do not require the
multiple Availability Zone
data resilience.

5

S3 Glacier

S3 Glacier is a secure, durable,
and low-cost storage class for
data archiving. Retrieval times
configurable from minutes to
hours.

6

S3 Glacier Deep Archive

S3 Glacier Deep Archive is
Amazon S3's lowest-cost
storage class where a
retrieval time of 12 hours is
acceptable.

Encryption In Transit is achieved by

- **SSL/TLS**

Encryption At Rest (Server Side) is achieved by

- **S3 Managed Keys - SSE-S3**
- **AWS Key Management Service, Managed Keys - SSE-KMS**
- **Server Side Encryption With Customer Provided Keys - SSE-C**

Client Side Encryption

Cross Region Replication

- Versioning must be enabled on both the source and destination buckets.
- Regions must be unique.
- Files in an existing bucket are not replicated automatically.
- All subsequent updated files will be replicated automatically.
- Delete markers are not replicated.
- Deleting individual versions or delete markers will not be replicated.
- Understand what Cross Region Replication is at a high level.



Lifecycle Policies

- Automates moving your objects between the different storage tiers.
- Can be used in conjunction with versioning.
- Can be applied to current versions and previous versions.

CloudFront

- **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 the 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 the CDN which consists of a collection of Edge Locations.
- **Web Distribution** - Typically used for Websites.

CloudFront

- **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 the 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 the CDN which consists of a collection of Edge Locations.
- **Web Distribution** - Typically used for Websites.
- **RTMP** - Used for Media Streaming.

CloudFront

- Edge locations are not just READ only — you can write to them too. (ie put an object on to them.)
- Objects are cached for the life of the **TTL (Time To Live.)**
- You can clear cached objects, but you will be charged.

Snowball

- Understand what Snowball is
- Snowball Can
- Import to S3
- Export from S3





File Gateway

- File Gateway - For flat files, stored directly on S3.

Volume Gateway

- **Stored Volumes** - Entire Dataset is stored on site and is asynchronously backed up to S3.
- **Cached Volumes** - Entire Dataset is stored on S3 and the most frequently accessed data is cached on site.

Gateway Virtual Tape Library

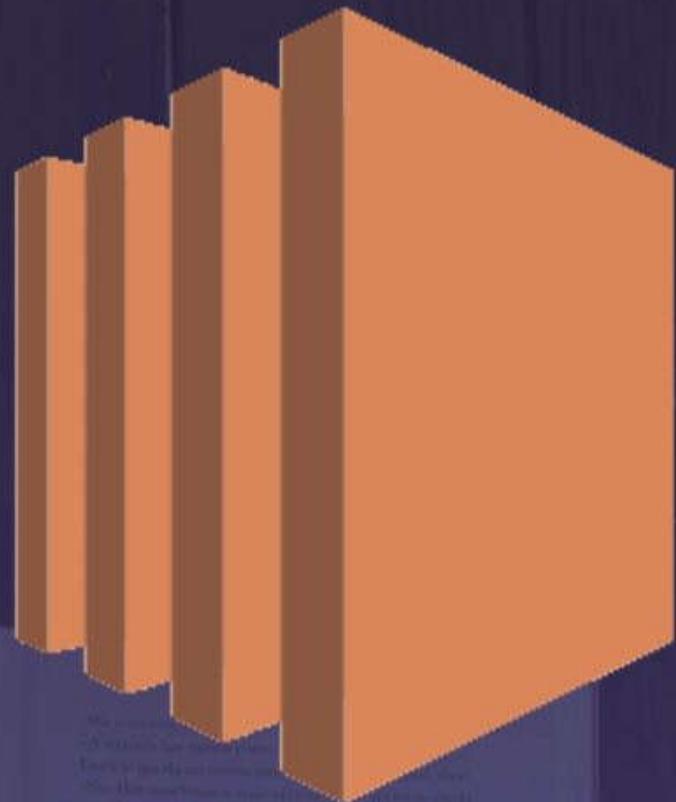
- Used for backup and uses popular backup applications like NetBackup, Backup Exec, Veeam etc.



- Read the S3 FAQs before taking the exam. It comes up A LOT!



Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. Amazon EC2 reduces the time required to obtain and boot new server instances to minutes, allowing you to quickly scale capacity, both up and down, as your computing requirements change.





1

On Demand

Allows you to pay a fixed rate by the hour (or by the second) with no commitment.

2

Reserved

Provides you with a capacity reservation, and offer a significant discount on the hourly charge for an instance. Contract Terms are 1 Year or 3 Year Terms.

3

Spot

Enables you to bid whatever price you want for instance capacity, providing for even greater savings if your applications have flexible start and end times.

4

Dedicated Hosts

Physical EC2 server dedicated for your use. Dedicated Hosts can help you reduce costs by allowing you to use your existing server-bound software licenses.



If the Spot instance is terminated by Amazon EC2, you will not be charged for a partial hour of usage. However, if you terminate the instance yourself, you will be charged for any hour in which the instance ran.



EC2 Instance Types - Mnemonic

- **F** - For FPGA
- **I** - For IOPS
- **G** - Graphics
- **H** - High Disk Throughput
- **T** - Cheap general purpose (think T2 Micro)
- **D** - For Density
- **R** - For RAM
- **M** - Main choice for general purpose apps
- **C** - For Compute
- **P** - Graphics (think Pics)
- **X** - Extreme Memory
- **Z** - Extreme Memory AND CPU
- **A** - Arm-based workloads
- **U** - Bare Metal



- Termination Protection is **turned off** by default, you must turn it on.
- On an EBS-backed instance, the **default action is for the root EBS volume to be deleted** when the instance is terminated.
- EBS Root Volumes of your DEFAULT AMI's cannot be encrypted. You can also use a third party tool (such as bit locker etc) to encrypt the root volume, or this can be done when creating AMI's (lab to follow) in the AWS console or using the API.
- Additional volumes can be encrypted.



- All Inbound traffic is blocked by default.
- All Outbound traffic is allowed.
- Changes to Security Groups take effect immediately.
- You can have any number of EC2 instances within a security group.
- You can have multiple security groups attached to EC2 Instances.

- Security Groups are **STATEFUL**.
- If you create an inbound rule allowing traffic in, that traffic is automatically allowed back out again.
- You cannot block specific IP addresses using Security Groups, instead use Network Access Control Lists.
- You can specify allow rules, but not deny rules.



- Volumes exist on EBS. Think of EBS as a virtual hard disk
- Snapshots exist on S3. Think of snapshots as a photograph of the disk.
- Snapshots are point in time copies of Volumes.
- Snapshots are incremental — this means that only the blocks that have changed since your last snapshot are moved to S3.



- To create a snapshot for Amazon EBS volumes that serve as root devices, you should stop the instance before taking the snapshot.
- However you can take a snap while the instance is running.
- You can create AMI's from both Volumes and Snapshots
- You can change EBS volume sizes on the fly, including changing the size and storage type.
- Volumes will ALWAYS be in the same availability zone as the EC2 instance.



- To move an EC2 volume from one AZ to another, take a snapshot of it, create an AMI from the snapshot and then use the AMI to launch the EC2 instance in a new AZ.
- To move an EC2 volume from one region to another, take a snapshot of it, create an AMI from the snapshot and then copy the AMI from one region to the other. Then use the copied AMI to launch the new EC2 instance in the new region.

- Instance Store Volumes are sometimes called Ephemeral Storage.
- Instance store volumes cannot be stopped. If the underlying host fails, you will lose your data.
- EBS backed instances can be stopped. You will not lose the data on this instance if it is stopped.
- You can reboot both, you will not lose your data.
- By default, both ROOT volumes will be deleted on termination. However, with EBS volumes, you can tell AWS to keep the root device volume.



- Snapshots of encrypted volumes are encrypted automatically.
- Volumes restored from encrypted snapshots are encrypted automatically.
- You can share snapshots, but only if they are unencrypted.
- These snapshots can be shared with other AWS accounts or made public.



- Create a Snapshot of the unencrypted root device volume
- Create a copy of the Snapshot and select the encrypt option
- Create an AMI from the encrypted Snapshot
- Use that AMI to launch new encrypted instances

Remember;

- CloudWatch is used for monitoring performance.
- CloudWatch can monitor most of AWS as well as your applications that run on AWS.
- CloudWatch with EC2 will monitor events every 5 minutes by default.
- You can have 1 minute intervals by turning on detailed monitoring.
- You can create CloudWatch alarms which trigger notifications.
- CloudWatch is all about performance. CloudTrail is all about auditing.



What Can I do With CloudWatch?

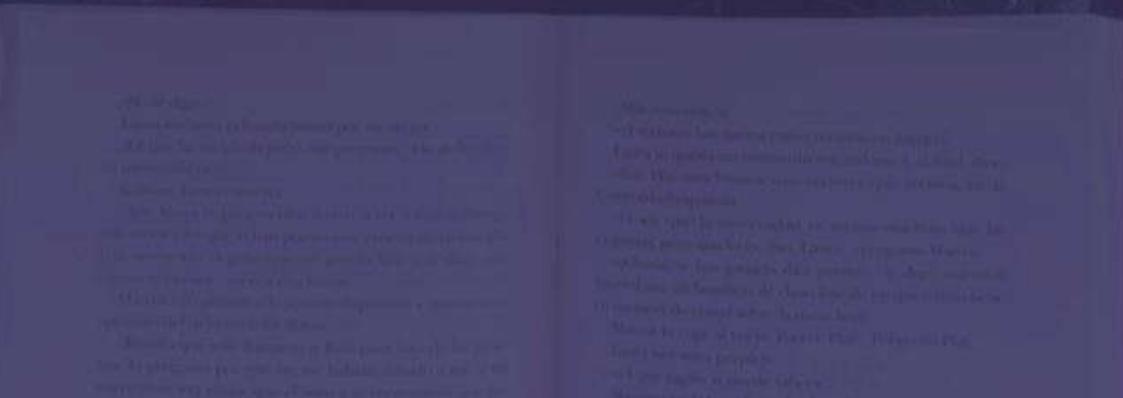
- Dashboards - Creates awesome dashboards to see what is happening with your AWS environment.
- Alarms - Allows you to set Alarms that notify you when particular thresholds are hit.
- Events - CloudWatch Events helps you to respond to state changes in your AWS resources.
- Logs - CloudWatch Logs helps you to aggregate, monitor, and store logs.

CloudTrail vs CloudWatch

- CloudWatch monitors performance.
- CloudTrail monitors API calls in the AWS platform.



- You can interact with AWS from anywhere in the world just by using the command line (CLI).
- You will need to set up access in IAM
- Commands themselves are not in the exam, but some basic commands will be useful to know for real life.



- Roles are more secure than storing your access key and secret access key on individual EC2 instances.
- Roles are easier to manage.
- Roles can be assigned to an EC2 instance after it is created using both the console & command line.
- Roles are universal — you can use them in any region.



- Used to get information about an instance (such as public ip)
- curl <http://169.254.169.254/latest/meta-data/>
- curl <http://169.254.169.254/latest/user-data/>

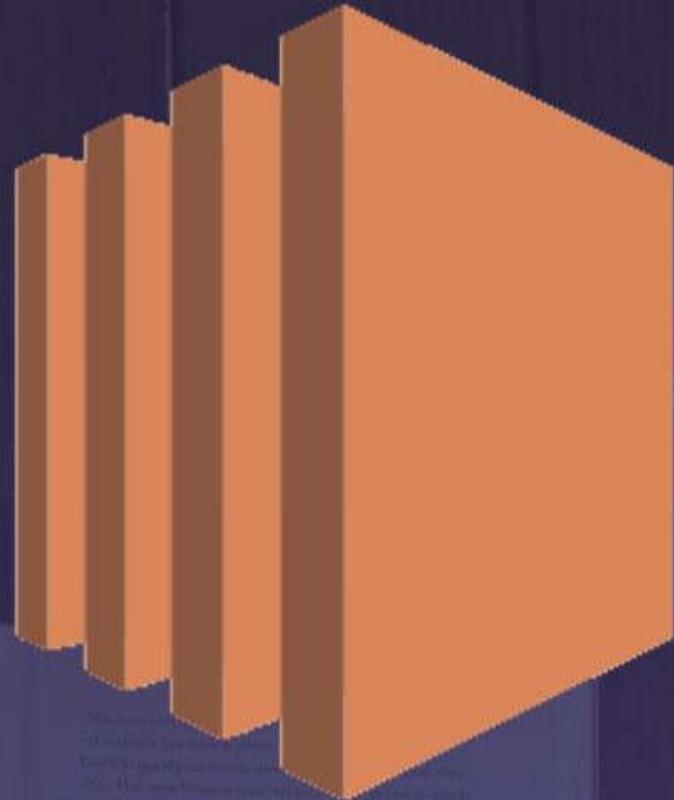


- Supports the Network File System version 4 (NFSv4) protocol
- You only pay for the storage you use (no pre-provisioning required.)
- Can scale up to the petabytes
- Can support thousands of concurrent NFS connections
- Data is stored across multiple AZ's within a region
- Read After Write Consistency

Placement Groups Exam Tips

- A clustered placement group can't span multiple Availability Zones.
- A spread placement group can.
- The name you specify for a placement group must be unique within your AWS account.
- Only certain types of instances can be launched in a placement group (Compute Optimized, GPU, Memory Optimized, Storage Optimized)
- AWS recommend homogenous instances within placement groups.
- You can't merge placement groups.
- You can't move an existing instance into a placement group. You can create an AMI from your existing instance, and then launch a new instance from the AMI into a placement group.

Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. Amazon EC2 reduces the time required to obtain and boot new server instances to minutes, allowing you to quickly scale capacity, both up and down, as your computing requirements change.



EC2 Exam Tips



1

On Demand

Allows you to pay a fixed rate by the hour (or by the second) with no commitment.

2

Reserved

Provides you with a capacity reservation, and offer a significant discount on the hourly charge for an instance. Contract Terms are 1 Year or 3 Year Terms.

3

Spot

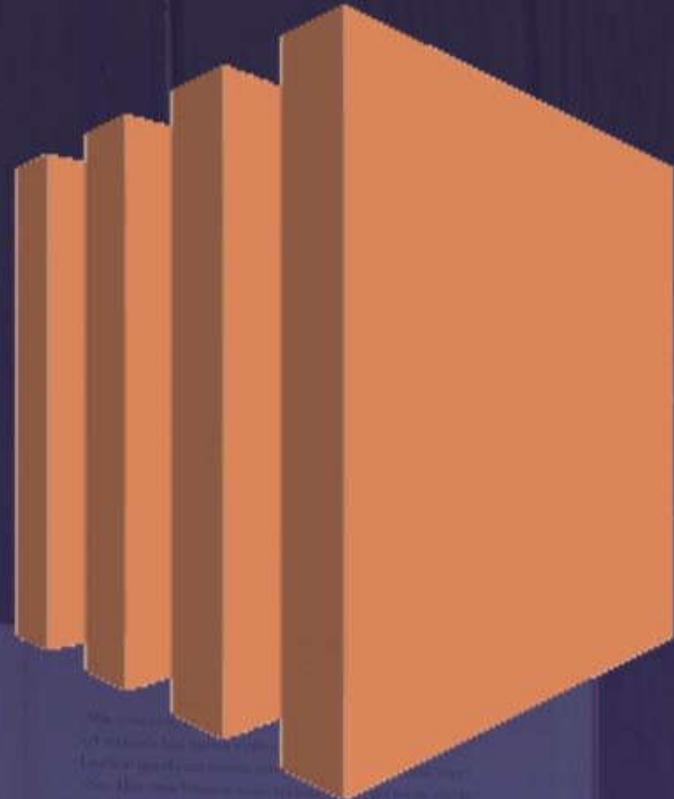
Enables you to bid whatever price you want for instance capacity, providing for even greater savings if your applications have flexible start and end times.

4

Dedicated Hosts

Physical EC2 server dedicated for your use. Dedicated Hosts can help you reduce costs by allowing you to use your existing server-bound software licenses.

If the Spot instance is terminated by Amazon EC2, you will not be charged for a partial hour of usage. However, if you terminate the instance yourself, you will be charged for any hour in which the instance ran.



EC2 Instance Types

- **F** - For FPGA
- **I** - For IOPS
- **G** - Graphics
- **H** - High Disk Throughput
- **T** - Cheap general purpose (think T2 Micro)
- **D** - For Density
- **R** - For RAM
- **M** - Main choice for general purpose apps
- **C** - For Compute
- **P** - Graphics (think Pics)
- **X** - Extreme Memory
- **Z** - Extreme Memory AND CPU
- **A** - Arm-based workloads
- **U** - Bare Metal



- Termination Protection is **turned off** by default, you must turn it on.
- On an EBS-backed instance, the **default action is for the root EBS volume to be deleted** when the instance is terminated.
- EBS Root Volumes of your DEFAULT AMI's cannot be encrypted. You can also use a third party tool (such as bit locker etc) to encrypt the root volume, or this can be done when creating AMI's (lab to follow) in the AWS console or using the API.
- Additional volumes can be encrypted.



- All Inbound traffic is blocked by default.
- All Outbound traffic is allowed.
- Changes to Security Groups take effect immediately.
- You can have any number of EC2 instances within a security group.
- You can have multiple security groups attached to EC2 Instances.

Compare EBS Types

Solid-State Drives (SSD)			Hard disk Drives (HDD)		
Volume Type	General Purpose SSD	Provisioned IOPS SSD	Throughput Optimized HDD	Cold HDD	EBS Magnetic
Description	General purpose SSD volume that balances price and performance for a wide variety of transactional workloads	Highest-performance SSD volume designed for mission-critical applications	Low cost HDD volume designed for frequently accessed, throughput-intensive workloads	Lowest cost HDD volume designed for less frequently accessed workloads	Previous generation HDD
Use Cases	Most Work Loads	Databases	Big Data & Data Warehouses	File Servers	Workloads where data is infrequently accessed
API Name	gp2	io1	st1	sc1	Standard
Volume Size	1 GiB - 16 TiB	4 GiB - 16 TiB	500 GiB - 16 TiB	500 GiB - 16 TiB	1 GiB-1 TiB
Max. IOPS**/ Volume	16,000	64,000	500	250	40-200

- Volumes exist on EBS. Think of EBS as a virtual hard disk
- Snapshots exist on S3. Think of snapshots as a photograph of the disk.
- Snapshots are point in time copies of Volumes.
- Snapshots are incremental — this means that only the blocks that have changed since your last snapshot are moved to S3.
- If this is your first snapshot, it may take some time to create.

- To create a snapshot for Amazon EBS volumes that serve as root devices, you should stop the instance before taking the snapshot.
- However you can take a snap while the instance is running.
- You can create AMI's from both Volumes and Snapshots
- You can change EBS volume sizes on the fly, including changing the size and storage type.
- Volumes will **ALWAYS** be in the same availability zone as the EC2 instance.

- To move an EC2 volume from one AZ to another, take a snapshot of it, create an AMI from the snapshot and then use the AMI to launch the EC2 instance in a new AZ.
- To move an EC2 volume from one region to another, take a snapshot of it, create an AMI from the snapshot and then copy the AMI from one region to the other. Then use the copied AMI to launch the new EC2 instance in the new region.

- Snapshots of encrypted volumes are encrypted automatically.
- Volumes restored from encrypted snapshots are encrypted automatically.



- Instance Store Volumes are sometimes called Ephemeral Storage.
- Instance store volumes cannot be stopped. If the underlying host fails, you will lose your data.
- EBS backed instances can be stopped. You will not lose the data on this instance if it is stopped.
- You can reboot both, you will not lose your data.
- By default, both ROOT volumes will be deleted on termination. However, with EBS volumes, you can tell AWS to keep the root device volume.

Encrypting Root Device Volumes

- Create a Snapshot of the unencrypted root device volume
- Create a copy of the Snapshot and select the encrypt option
- Create an AMI from the encrypted Snapshot

Remember;

- CloudWatch is used for monitoring performance.
- CloudWatch can monitor most of AWS as well as your applications that run on AWS.
- CloudWatch with EC2 will monitor events every 5 minutes by default.
- You can have 1 minute intervals by turning on detailed monitoring.
- You can create CloudWatch alarms which trigger notifications.
- CloudWatch is all about performance. CloudTrail is all about auditing.

What Can I do With CloudWatch?

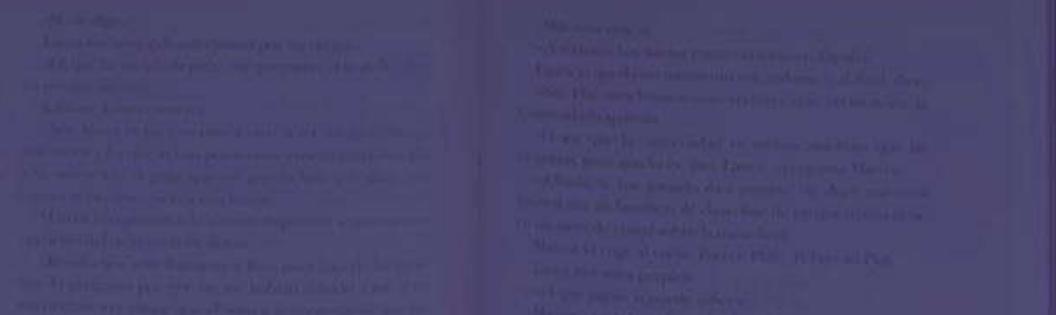
- Dashboards - Creates awesome dashboards to see what is happening with your AWS environment.
- Alarms - Allows you to set Alarms that notify you when particular thresholds are hit.
- Events - CloudWatch Events helps you to respond to state changes in your AWS resources.
- Logs - CloudWatch Logs helps you to aggregate, monitor, and store logs.

CloudTrail vs CloudWatch

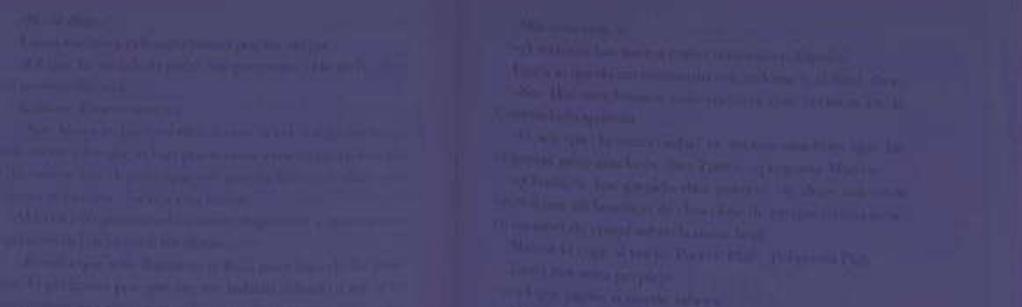
- CloudWatch monitors performance.
- CloudTrail monitors API calls in the AWS platform.



- You can interact with AWS from anywhere in the world just by using the command line (CLI).
- You will need to set up access in IAM
- Commands themselves are not in the exam, but some basic commands will be useful to know for real life.



- Roles are more secure than storing your access key and secret access key on individual EC2 instances.
- Roles are easier to manage.
- Roles can be assigned to an EC2 instance after it is created using both the console & command line.
- Roles are universal — you can use them in any region.



- Bootstrap scripts run when an EC2 instance first boots.
- Can be a powerful way of automating software installs and updates.



Amazon Elastic Compute Cloud (Amazon EC2) is a web service that provides resizable compute capacity in the cloud. It is designed to make it easier for you to build, run, and manage web-scale applications without thinking about the underlying infrastructure. You can quickly boot up new instances of your application and scale them up or down based on demand. You can also choose from a variety of instance types to fit your specific needs. Amazon EC2 makes it easy to add more resources to your application as it grows, so you can focus on what's important: building your application. With Amazon EC2, you can quickly spin up new instances of your application and scale them up or down based on demand. You can also choose from a variety of instance types to fit your specific needs. Amazon EC2 makes it easy to add more resources to your application as it grows, so you can focus on what's important: building your application.



- Used to get information about an instance (such as public ip)
- curl <http://169.254.169.254/latest/meta-data/>
- curl <http://169.254.169.254/latest/user-data/>

- Supports the Network File System version 4 (NFSv4) protocol
- You only pay for the storage you use (no pre-provisioning required.)
- Can scale up to the petabytes
- Can support thousands of concurrent NFS connections
- Data is stored across multiple AZ's within a region
- Read After Write Consistency



Two Types of Placement Groups;

- Clustered Placement Group
- Spread Placement Group



Placement Groups Exam Tips

- A clustered placement group can't span multiple Availability Zones.
- A spread placement group can.
- The name you specify for a placement group must be unique within your AWS account.
- Only certain types of instances can be launched in a placement group (Compute Optimized, GPU, Memory Optimized, Storage Optimized)
- AWS recommend homogenous instances within placement groups.
- You can't merge placement groups.
- You can't move an existing instance into a placement group. You can create an AMI from your existing instance, and then launch a new instance from the AMI into a placement group.



RDS (OLTP)

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

DynamoDB (No SQL)

Red Shift OLAP



RDS (OLTP)

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

DynamoDB (No SQL)

Red Shift OLAP



Redshift for Business Intelligence or Data Warehousing

Redshift is a highly scalable, petabyte-scale data warehouse service that makes it easy to analyze large amounts of data. It provides fast query performance, automatic scaling, and simple administration. Redshift is designed for business intelligence (BI) and data warehousing workloads, allowing users to store and analyze large volumes of data in a cost-effective way.

Redshift's architecture is based on a distributed system that uses a master-worker model. The master node manages the cluster and performs administrative tasks, while the worker nodes handle data storage and computation. This allows Redshift to handle large amounts of data and perform complex queries quickly.

One of the key features of Redshift is its ability to automatically manage data distribution across multiple nodes. This means that data is automatically partitioned and replicated across the cluster, making it easy to scale up or down as needed. Redshift also supports a variety of data formats, including CSV, JSON, and Parquet, making it easy to load data from a variety of sources.

Redshift is designed to be used with a variety of tools and languages, including Python, Java, and SQL. It also includes a range of built-in functions and operators that make it easy to perform complex data analysis. Redshift is a powerful tool for businesses that need to analyze large amounts of data quickly and efficiently.



Elasticache to speed up performance of existing databases (frequent identical queries).



Remember the following points;

- RDS runs on virtual machines
- You cannot log in to these operating systems however.
- Patching of the RDS Operating System and DB is Amazon's responsibility
- RDS is NOT Serverless
- Aurora Serverless IS Serverless

There are two different types of Backups for RDS:

- Automated Backups
- Database Snapshots



Read Replicas

- Can be Multi-AZ.
- Used to increase performance.
- Must have backups turned on.
- Can be in different regions.
- Can be Aurora or MySQL.
- Can be promoted to master, this will break the Read Replica

MultiaZ

- Used For DR.
- You can force a failover from one AZ to another by rebooting the RDS instance.



Multi-AZ
Amazon RDS Multi-AZ is a feature that provides automatic failover protection for your database instances. It uses two separate Amazon RDS instances located in different Availability Zones (AZs) within the same AWS Region. This ensures that if one AZ experiences a failure, the database can automatically switch to the other AZ without downtime.

Multi-AZ
Multi-AZ failover is a feature that allows you to manually or automatically switch between the two Multi-AZ instances. This is useful for disaster recovery scenarios where you want to ensure that your database remains available even if one AZ fails. You can force a failover from one AZ to another by rebooting the RDS instance.

Encryption at rest is supported for MySQL, Oracle, SQL Server, PostgreSQL, MariaDB & Aurora. Encryption is done using the AWS Key Management Service (KMS) service. Once your RDS instance is encrypted, the data stored at rest in the underlying storage is encrypted, as are its automated backups, read replicas, and snapshots.



The basics of DynamoDB are as follows;

- Stored on SSD storage
- Spread across 3 geographically distinct data centres
- Eventual Consistent Reads (Default)
- Strongly Consistent Reads



- Redshift is used for business intelligence.
- Available in only 1 AZ



Redshift is used for business intelligence. It is available in only one AZ. It is a columnar storage system. It is highly parallelized. It is a distributed system. It is a managed service. It is a serverless service. It is a pay-as-you-go service. It is a fully managed service. It is a highly reliable service. It is a highly available service. It is a highly performant service. It is a highly scalable service. It is a highly cost-effective service. It is a highly secure service. It is a highly compliant service. It is a highly audited service. It is a highly monitored service. It is a highly optimized service. It is a highly automated service. It is a highly integrated service. It is a highly flexible service. It is a highly customizable service. It is a highly extensible service. It is a highly reliable service. It is a highly available service. It is a highly performant service. It is a highly scalable service. It is a highly cost-effective service. It is a highly secure service. It is a highly compliant service. It is a highly audited service. It is a highly monitored service. It is a highly optimized service. It is a highly automated service. It is a highly integrated service. It is a highly flexible service. It is a highly customizable service. It is a highly extensible service.

Redshift is used for business intelligence. It is available in only one AZ. It is a columnar storage system. It is highly parallelized. It is a distributed system. It is a managed service. It is a serverless service. It is a pay-as-you-go service. It is a fully managed service. It is a highly reliable service. It is a highly available service. It is a highly performant service. It is a highly scalable service. It is a highly cost-effective service. It is a highly secure service. It is a highly compliant service. It is a highly audited service. It is a highly monitored service. It is a highly optimized service. It is a highly automated service. It is a highly integrated service. It is a highly flexible service. It is a highly customizable service. It is a highly extensible service. It is a highly reliable service. It is a highly available service. It is a highly performant service. It is a highly scalable service. It is a highly cost-effective service. It is a highly secure service. It is a highly compliant service. It is a highly audited service. It is a highly monitored service. It is a highly optimized service. It is a highly automated service. It is a highly integrated service. It is a highly flexible service. It is a highly customizable service. It is a highly extensible service.



Backups

- Enabled by default with a 1 day retention period.
- Maximum retention period is 35 days.
- Redshift always attempts to maintain at least three copies of your data (the original and replica on the compute nodes and a backup in Amazon S3).
- Redshift can also asynchronously replicate your snapshots to S3 in another region for disaster recovery.



- 2 copies of your data is contained in each availability zone, with minimum of 3 availability zones. 6 copies of your data.
- You can share Aurora Snapshots with other AWS accounts.
- 2 types of replicas available. Aurora Replicas and MySQL replicas. Automated failover is only available with Aurora Replicas.
- Aurora has automated backups turned on by default. You can also take Snapshots with Aurora. You can share these snapshots with other AWS accounts.



- Use ElastiCache to increase database and web application performance.
- Redis is Multi-AZ
- You can do back ups and restores of Redis
- If you need to scale horizontally, use Memcached



RDS (OLTP)

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

DynamoDB (No SQL)

Red Shift OLAP



Elasticache

- Memcached
- Redis



Remember the following points;

- RDS runs on virtual machines
- You cannot log in to these operating systems however.
- Patching of the RDS Operating System and DB is Amazon's responsibility
- RDS is NOT Serverless
- Aurora Serverless IS Serverless

There are two different types of Backups for RDS:

- Automated Backups
- Database Snapshots



Read Replicas

- Can be Multi-AZ.
- Used to increase performance.
- Must have backups turned on.
- Can be in different regions.
- Can be Aurora or MySQL.
- Can be promoted to master, this will break the replication with the Read Replica

MultiAZ

- Used For DR.
- You can force a failover from one AZ to another by rebooting the RDS instance.

Encryption at rest is supported for MySQL, Oracle, SQL Server, PostgreSQL, MariaDB & Aurora. Encryption is done using the AWS Key Management Service (KMS) service. Once your RDS instance is encrypted, the data stored at rest in the underlying storage is encrypted, as are its automated backups, read replicas, and snapshots.

DynamoDB

- Stored on SSD storage
- Spread Across 3 geographically distinct data centres
- Eventual Consistent Reads (Default)
- Strongly Consistent Reads



- Redshift is used for business intelligence.
- Available in only 1 AZ

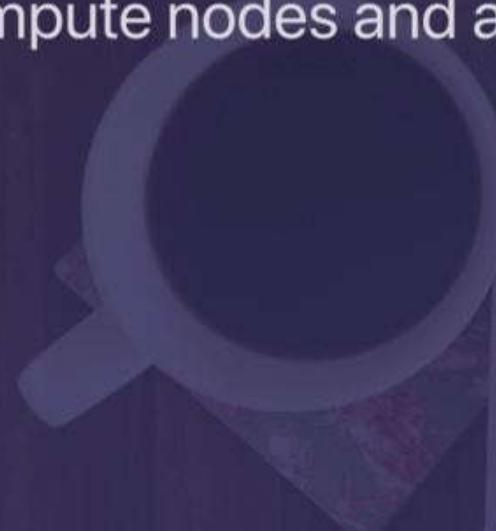


Redshift is used for business intelligence. It is available in only 1 AZ.

Redshift is used for business intelligence. It is available in only 1 AZ.

Redshift Backups

- Enabled by default with a 1 day retention period.
- Maximum retention period is 35 days.
- Redshift always attempts to maintain at least three copies of your data (the original and replica on the compute nodes and a backup in Amazon S3).



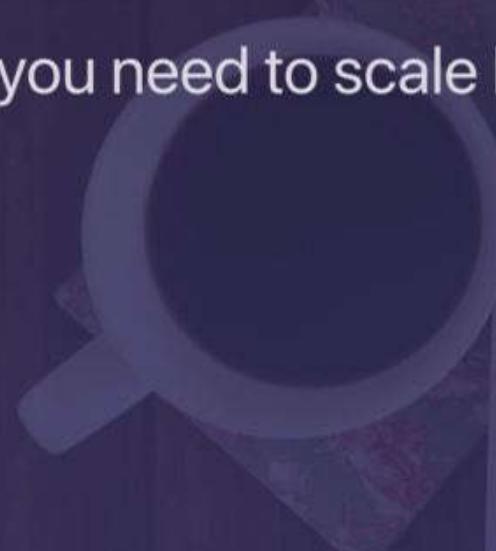
Redshift automatically creates and maintains multiple copies of your data. This ensures that your data is always available and can be restored quickly if needed. Redshift uses a combination of local disk storage on the compute nodes and Amazon S3 for backups. By default, Redshift retains one day of backups, but you can increase this to up to 35 days. Redshift also attempts to maintain at least three copies of your data: one copy on the compute nodes, one replica on the compute nodes, and one backup in Amazon S3. This provides redundancy and ensures that your data is always available even if one of the copies is lost or corrupted.

Aurora

- 2 copies of your data is contained in each availability zone, with minimum of 3 availability zones. 6 copies of your data.
- You can share Aurora Snapshots with other AWS accounts.
- 2 types of replicas available. Aurora Replicas and MySQL replicas. Automated failover is only available with Aurora Replicas.
- Aurora has automated backups turned on by default. You can also take Snapshots with Aurora. You can share these snapshots with other AWS accounts.

Elasticache

- Use ElastiCache to increase database and web application performance.
- Redis is Multi-AZ
- You can do back ups and restores of Redis
- If you need to scale horizontally, use Memcached



ElastiCache
Amazon ElastiCache is a managed service for Redis and MySQL that makes it easy to set up, operate, and scale your own database instances. It provides automatic failover, monitoring, and backups. It also integrates with other AWS services like Lambda and Step Functions.

Redis
Redis is a fast, open-source, in-memory data store that can be used as a database, cache, or message broker. It supports various data structures like strings, hashes, lists, sets, and sorted sets. Redis is highly available and can be used in a multi-AZ configuration.

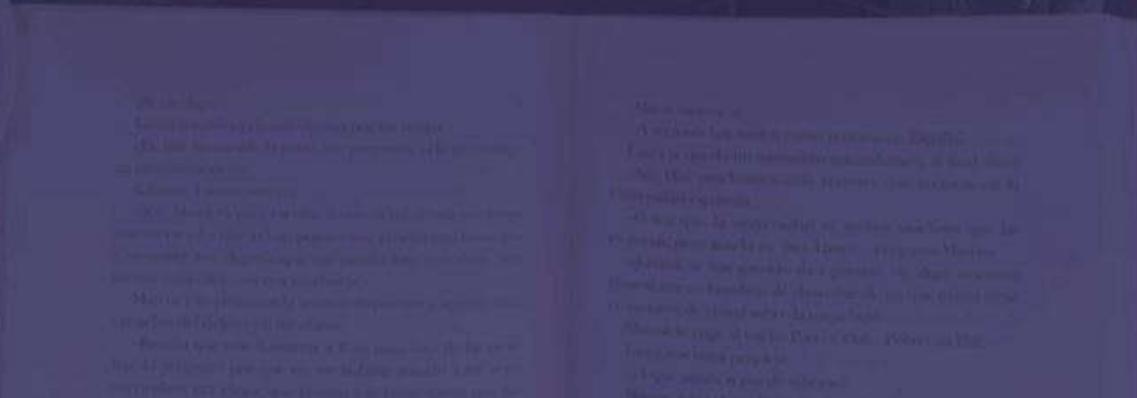
Route53 Exam Tips

- ELBs do not have pre-defined IPv4 addresses; you resolve to them using a DNS name.
- Understand the difference between an Alias Record and a CNAME.
- Given the choice, always choose an Alias Record over a CNAME.

Common DNS Types

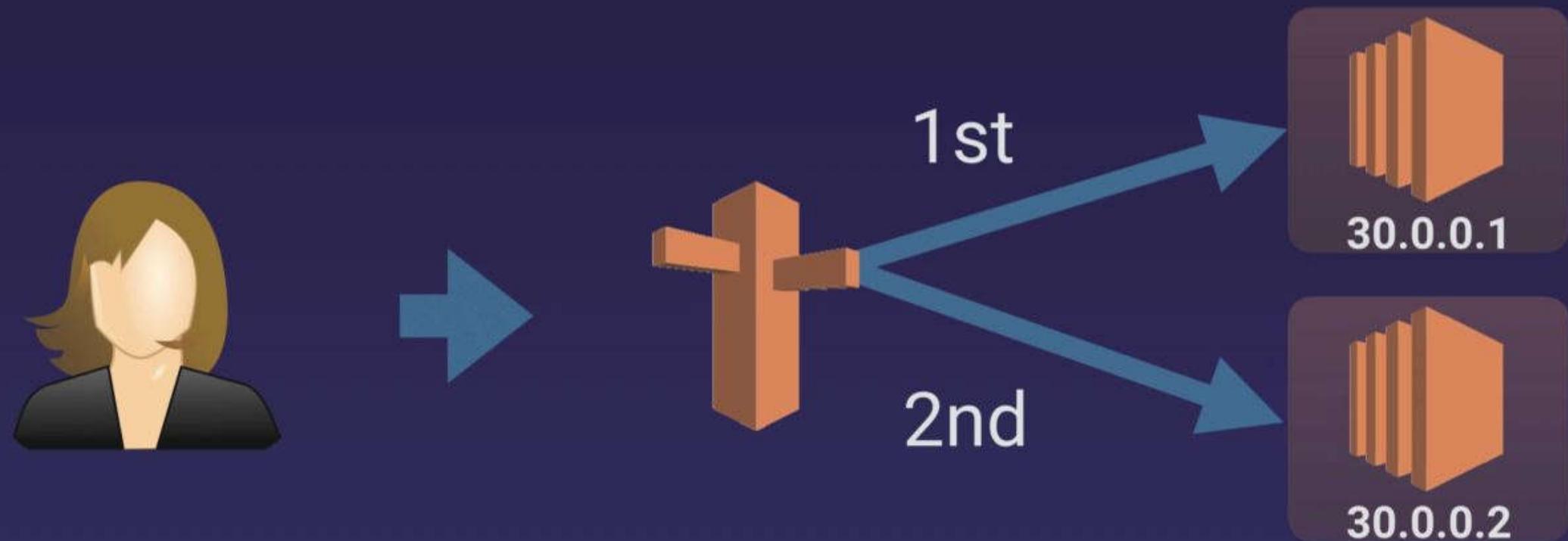
- SOA Records
- NS Records
- A Records
- CNAMEs
- MX Records
- PTR Records

- You can buy domain names directly with AWS.
- It can take up to 3 days to register depending on the circumstances.



Simple Routing Policy

If you choose the simple routing policy you can only have one record with multiple IP addresses. If you specify multiple values in a record, Route 53 returns all values to the user in a random order.

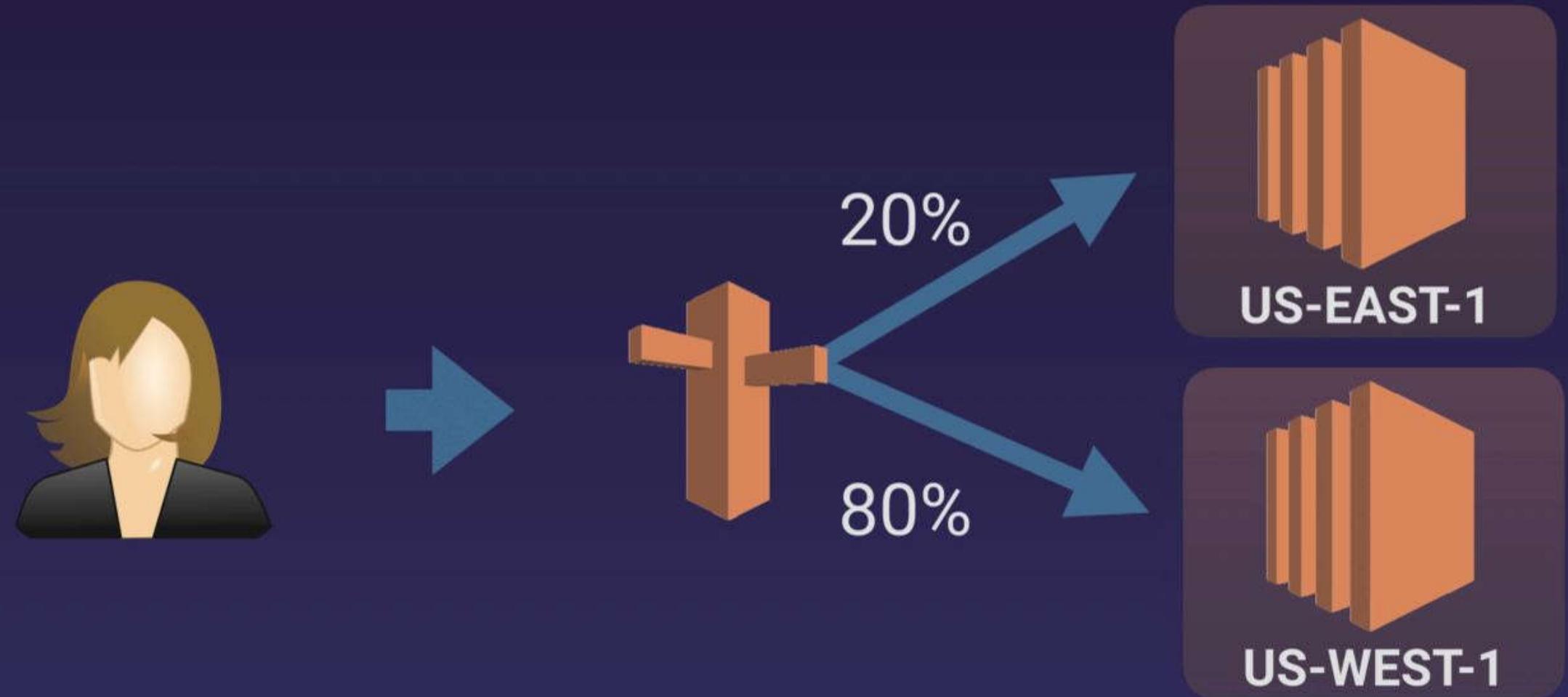


Simple Routing Policy

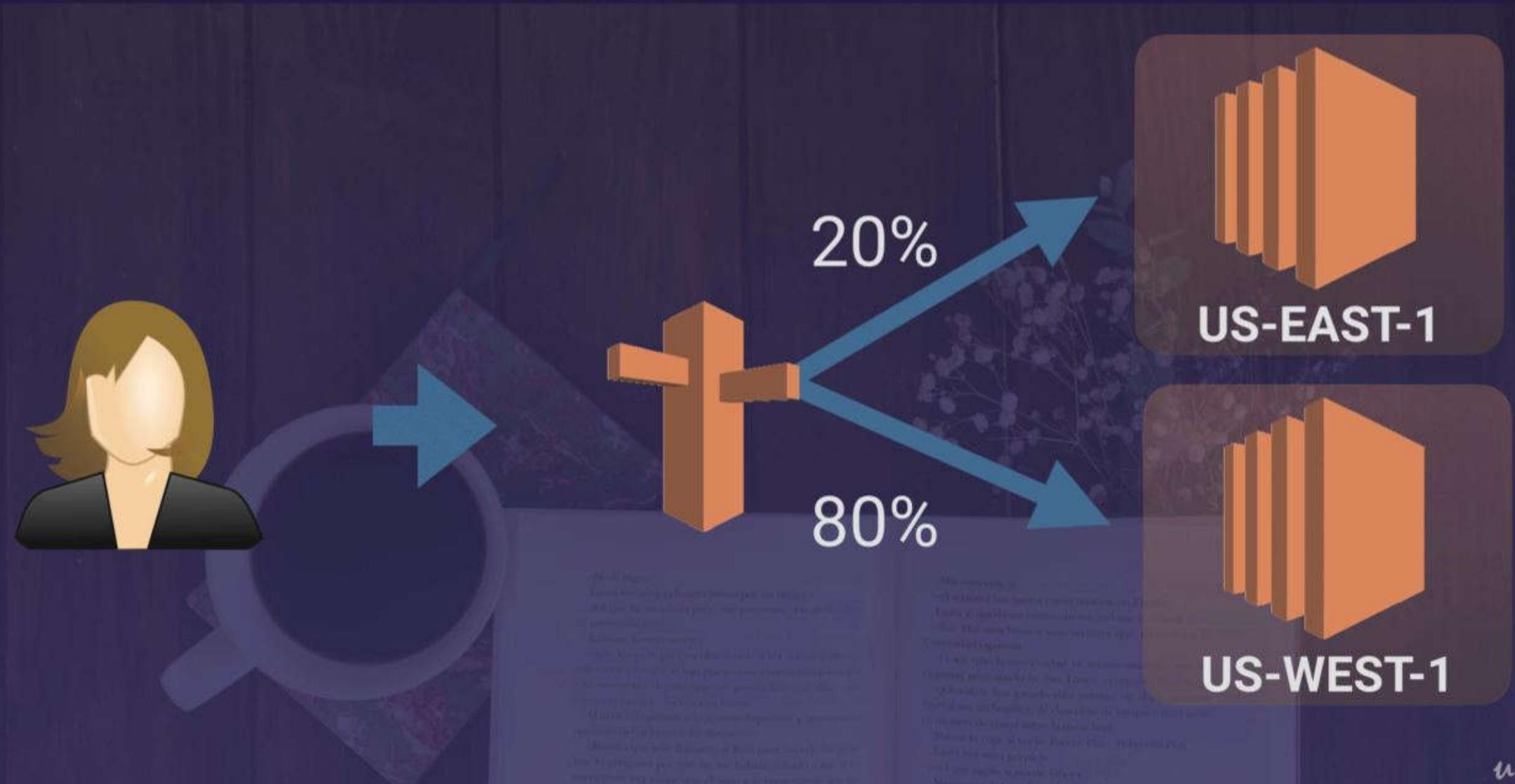
If you choose the simple routing policy you can only have one record with multiple IP addresses. If you specify multiple values in a record, Route 53 returns all values to the user in a random order.



Exam Tips

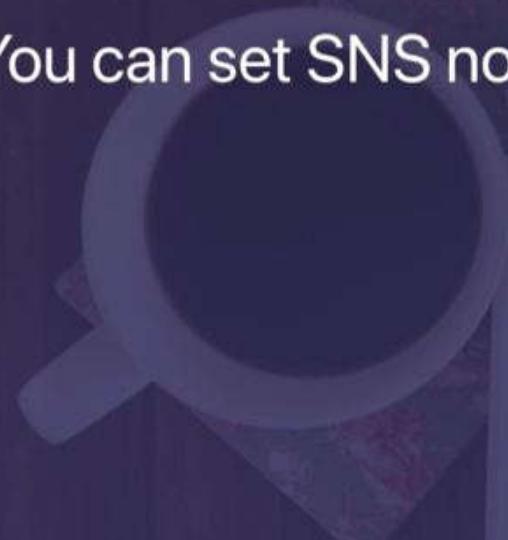


Exam Tips

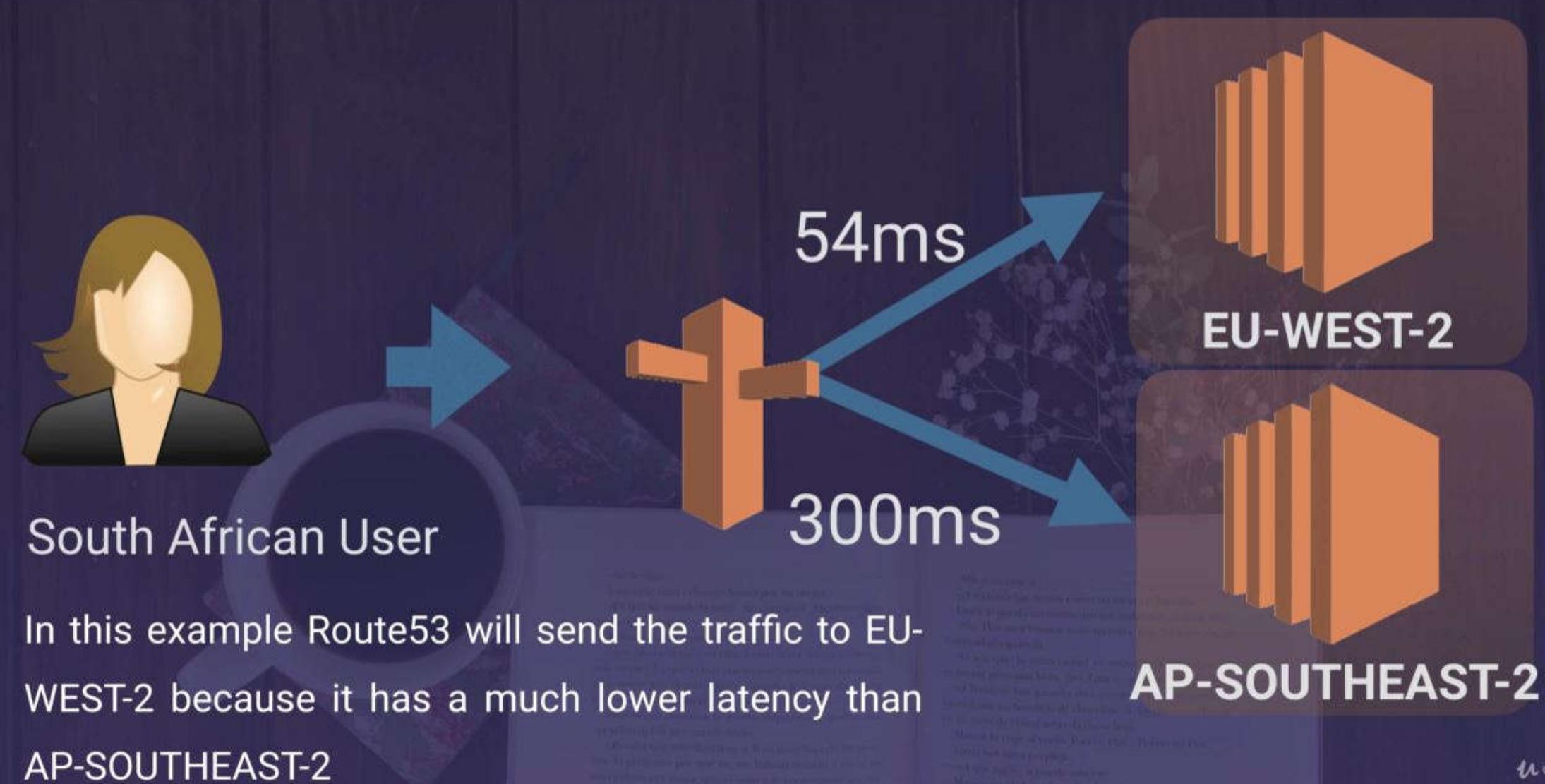


Health Checks

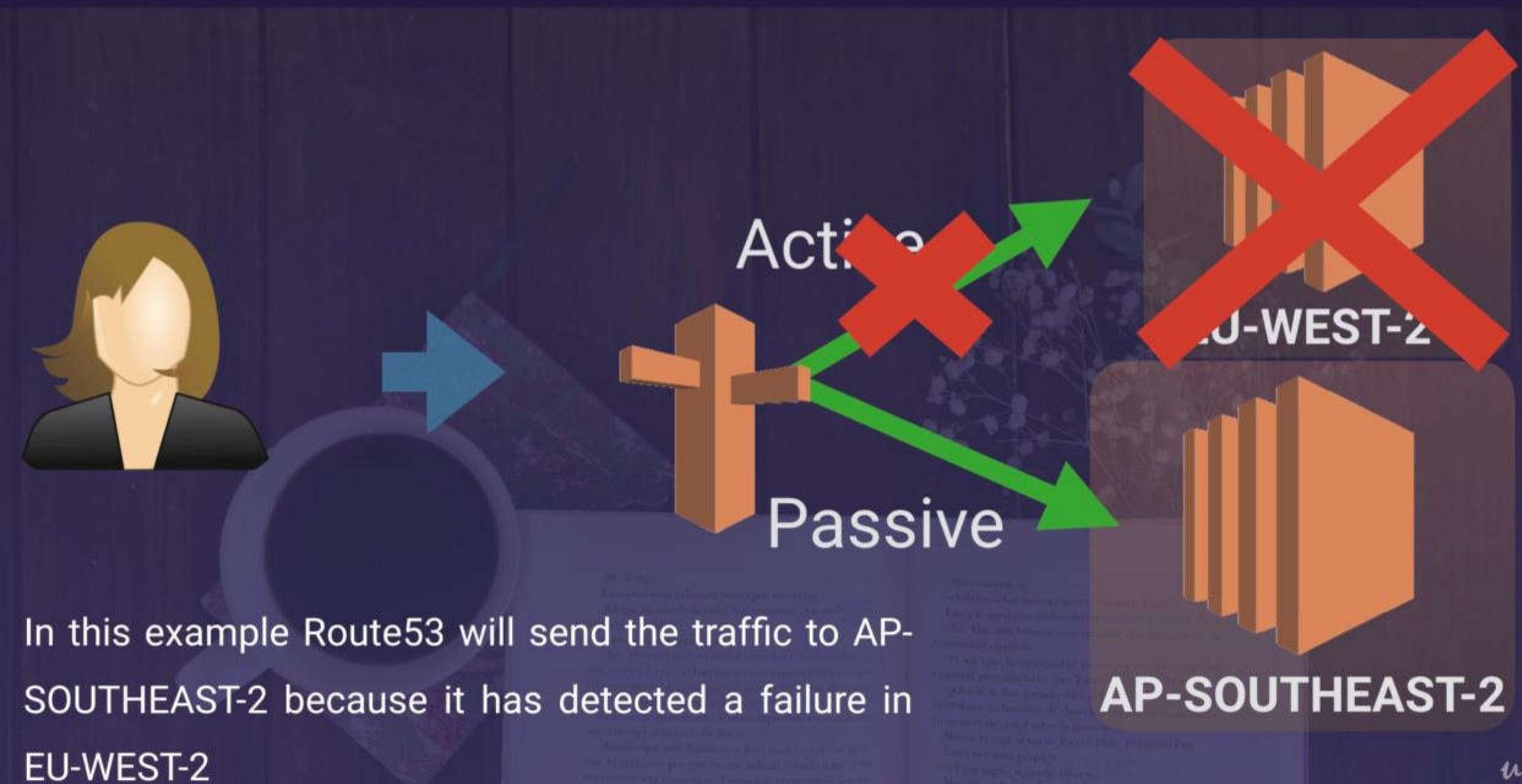
- You can set health checks on individual record sets.
- If a record set fails a health check it will be removed from Route53 until it passes the health check.
- You can set SNS notifications to alert you if a health check is failed.



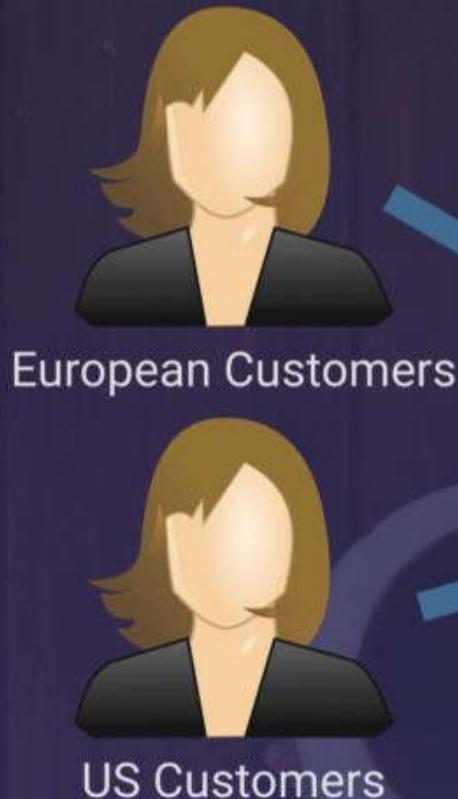
Health checks are used to determine the status of individual resources in a Route53 health check configuration. You can set up health checks for individual record sets in a routing policy. If a record set fails a health check, it will be removed from Route53 until it passes the health check. You can also set up SNS notifications to alert you if a health check fails.



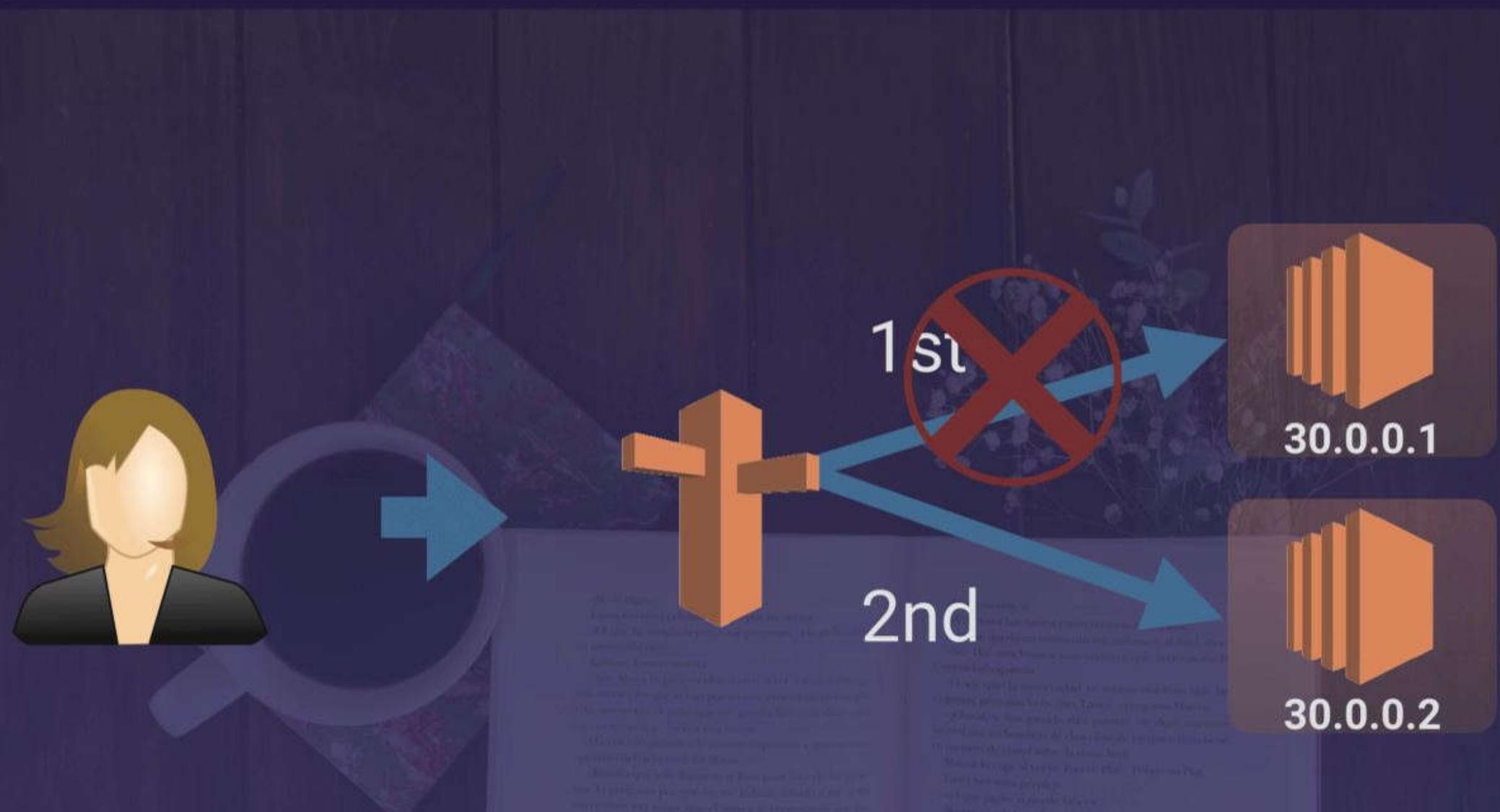
In this example Route53 will send the traffic to EU-WEST-2 because it has a much lower latency than AP-SOUTHEAST-2



Exam Tips



In this example, Route53 will send the European customers to EU-WEST-1 and the US customers to US-EAST-1



Route53 Exam Tips

- ELBs do not have pre-defined IPv4 addresses; you resolve to them using a DNS name.
- Understand the difference between an Alias Record and a CNAME.
- Given the choice, always choose an Alias Record over a CNAME.

Common DNS Types

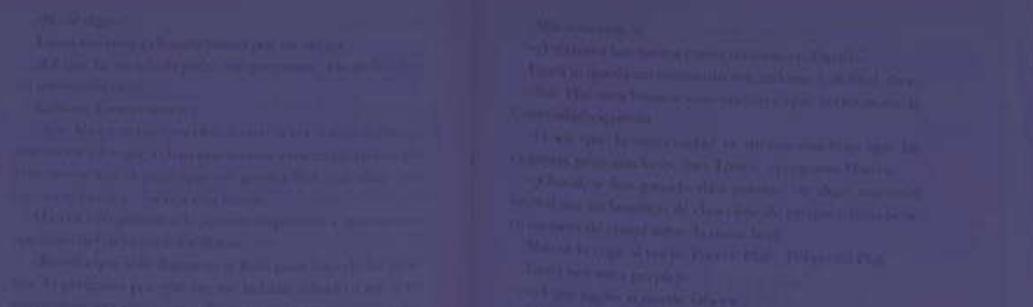
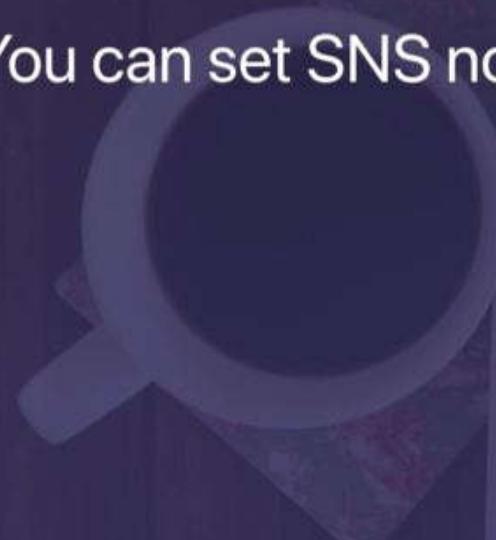
- SOA Records
- NS Records
- A Records
- CNAMEs
- MX Records
- PTR Records

The Following Routing Policies Are Available With Route53:

- Simple Routing
- Weighted Routing
- Latency-based Routing
- Failover Routing
- Geolocation Routing
- Geoproximity Routing (Traffic Flow Only)
- Multivalue Answer Routing

Health Checks

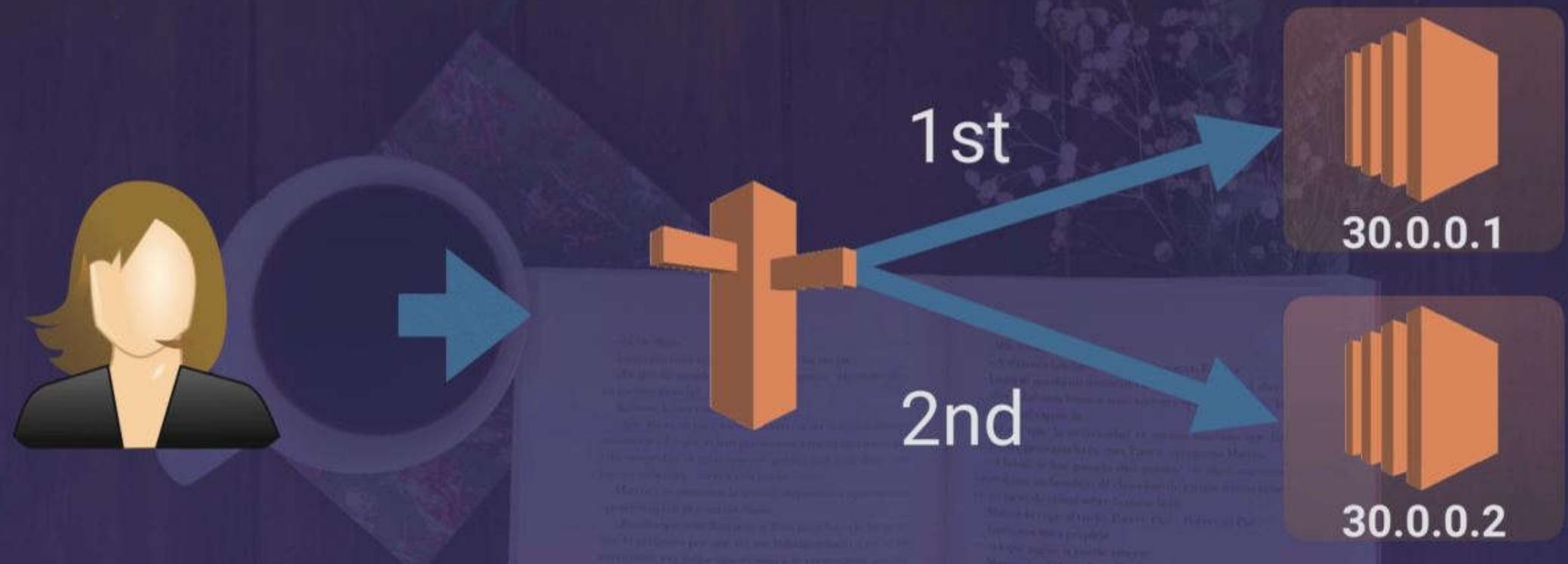
- You can set health checks on individual record sets.
- If a record set fails a health check it will be removed from Route53 until it passes the health check.
- You can set SNS notifications to alert you if a health check is failed.



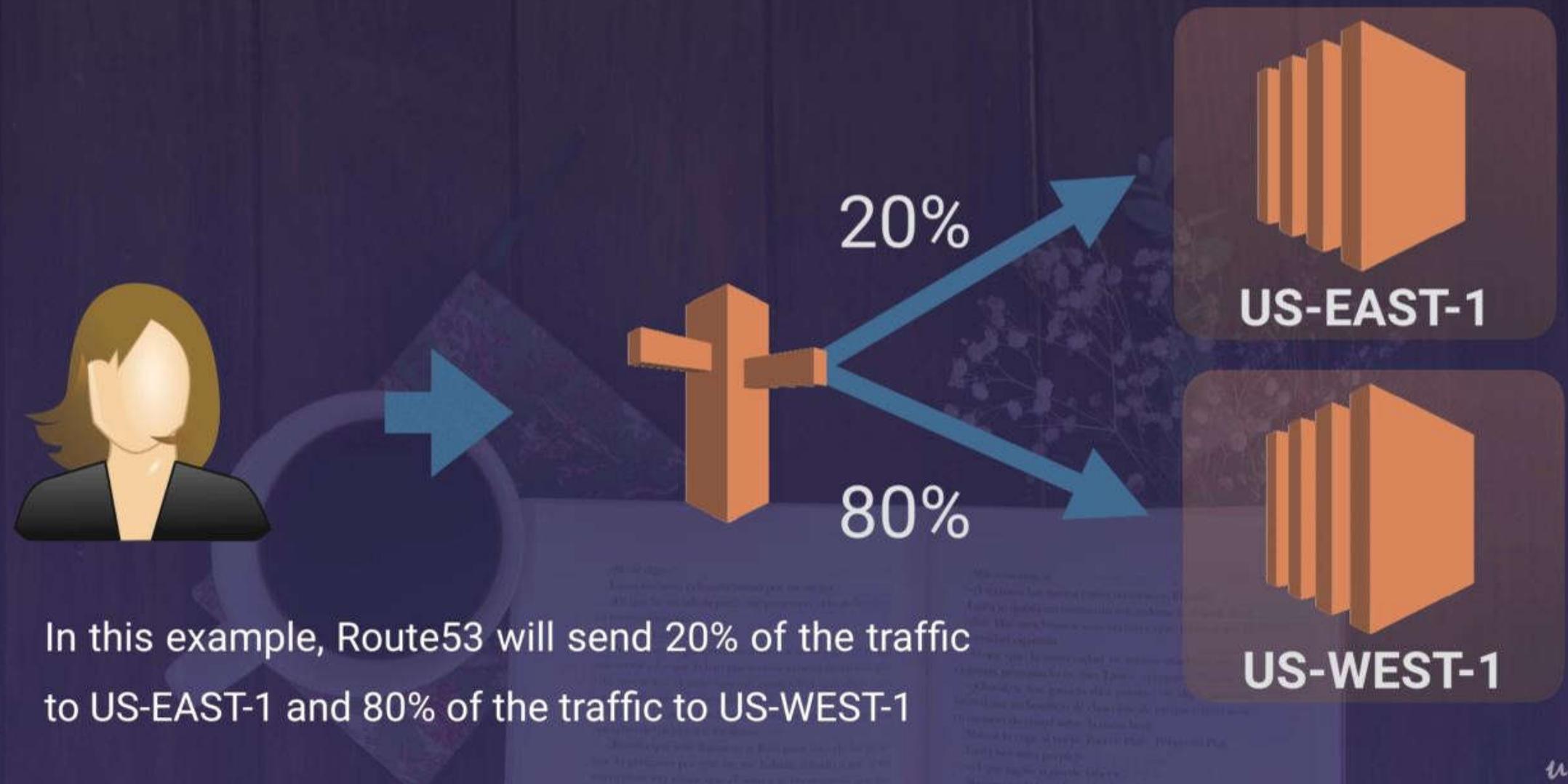
Simple Routing Policy

Simple Routing Policy

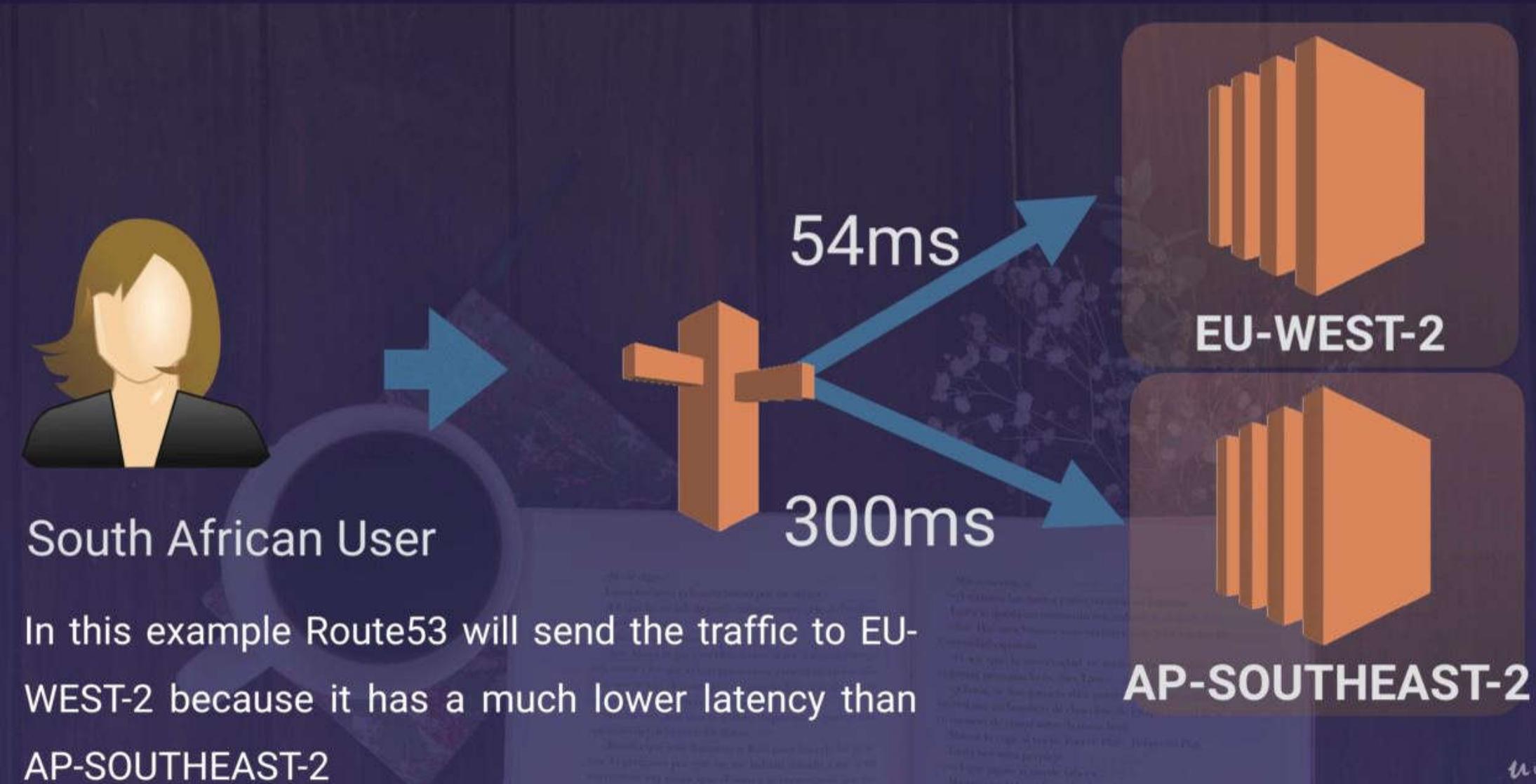
If you choose the simple routing policy you can only have one record with multiple IP addresses. If you specify multiple values in a record, Route 53 returns all values to the user in a random order.



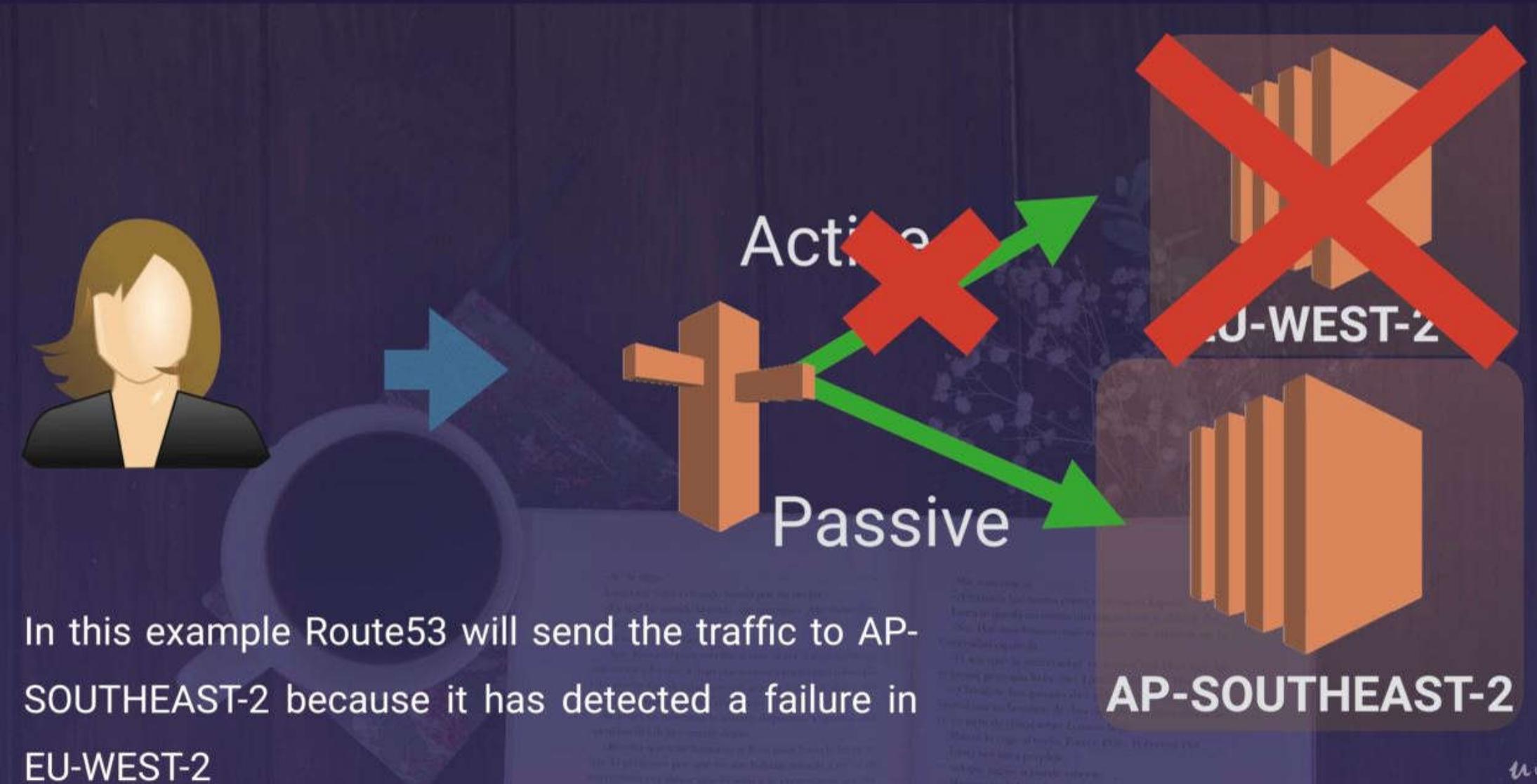
Weighted Routing Policy



Latency Routing Policy



Failover Routing Policy



Geolocation Routing Policy





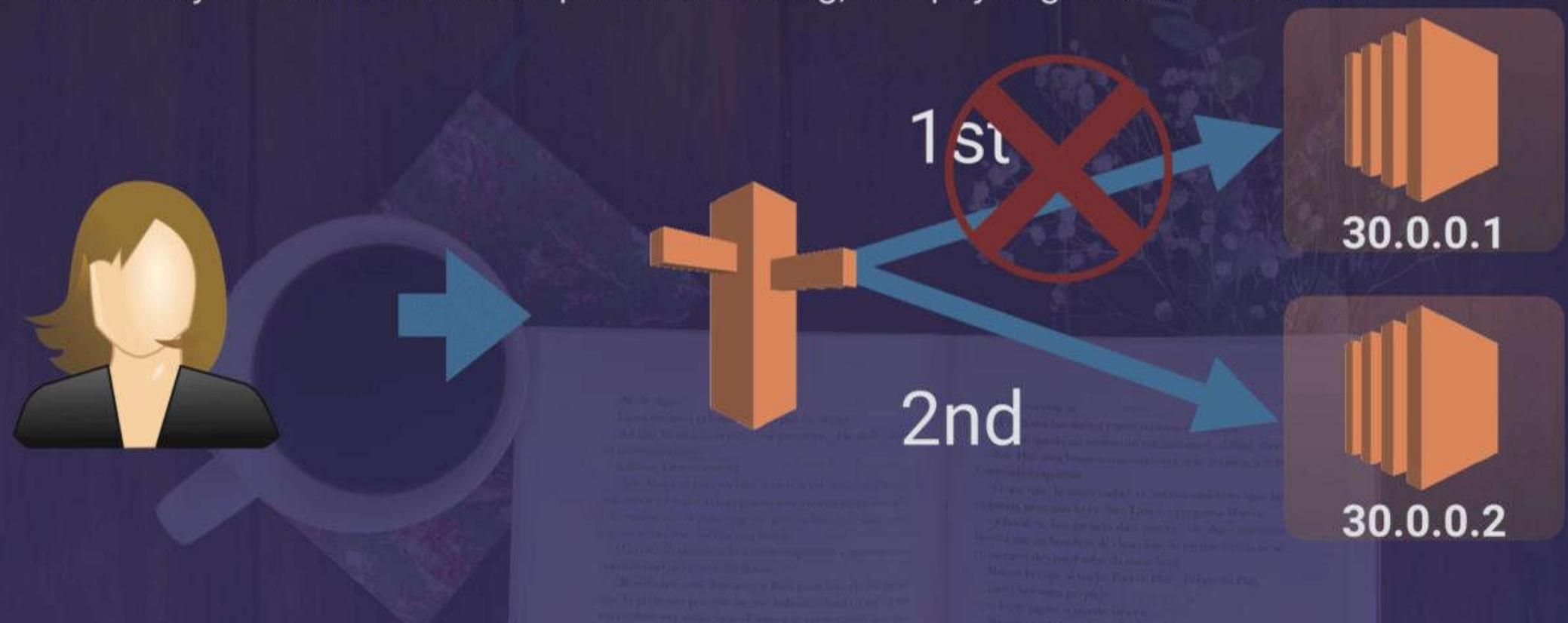
Geoproximity Routing (Traffic Flow Only)

Geoproximity routing lets Amazon Route 53 route traffic to your resources based on the geographic location of your users and your resources. You can also optionally choose to route more traffic or less to a given resource by specifying a value, known as a bias. A bias expands or shrinks the size of the geographic region from which traffic is routed to a resource.

To use geoproximity routing, you must use Route 53 traffic flow.

Multivalue Answer Policy

Essentially the same as with Simple based routing, except you get **Health Checks**.





Remember the following;

- Think of a VPC as a logical datacenter in AWS.
- Consists of IGWs (Or Virtual Private Gateways), Route Tables, Network Access Control Lists, Subnets, and Security Groups
- 1 Subnet = 1 Availability Zone
- Security Groups are Stateful; Network Access Control Lists are Stateless
- NO TRANSITIVE PEERING



Remember the following;

- When you create a VPC a default Route Table, Network Access Control List (NACL) and a default Security Group.
- It won't create any subnets, nor will it create a default internet gateway.
- US-East-1A in your AWS account can be a completely different availability zone to US-East-1A in another AWS account. The AZ's are randomized.
- Amazon always reserve 5 IP addresses within your subnets.
- You can only have 1 Internet Gateway per VPC.
- Security Groups can't span VPCs.



Nat Instances Exam Tips

- When creating a NAT instance, Disable Source/Destination Check on the Instance.
- NAT instances must be in a public subnet.
- There must be a route out of the private subnet to the NAT instance, in order for this to work.
- The amount of traffic that NAT instances can support depends on the instance size. If you are bottlenecking, increase the instance size.
- You can create high availability using Autoscaling Groups, multiple subnets in different AZs, and a script to automate failover.
- Behind a Security Group.



Nat Gateways

- Redundant inside the Availability Zone
- Preferred by the enterprise
- Starts at 5Gbps and scales currently to 45Gbps
- No need to patch
- Not associated with security groups
- Automatically assigned a public ip address
- Remember to update your route tables.
- No need to disable Source/Destination Checks



Nat Gateways

- If you have resources in multiple Availability Zones and they share one NAT gateway, in the event that the NAT gateway's Availability Zone is down, resources in the other Availability Zones lose internet access. To create an Availability Zone-independent architecture, create a NAT gateway in each Availability Zone and configure your routing to ensure that resources use the NAT gateway in the same Availability Zone.



Exam Tips

Remember the following for your exam;

- Your VPC automatically comes with a default network ACL, and by default it allows all outbound and inbound traffic.
- You can create custom network ACLs. 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 associate a subnet with a network ACL, the subnet is automatically associated with the default network ACL.
- Block IP Addresses using network ACLs not Security Groups

Exam Tips



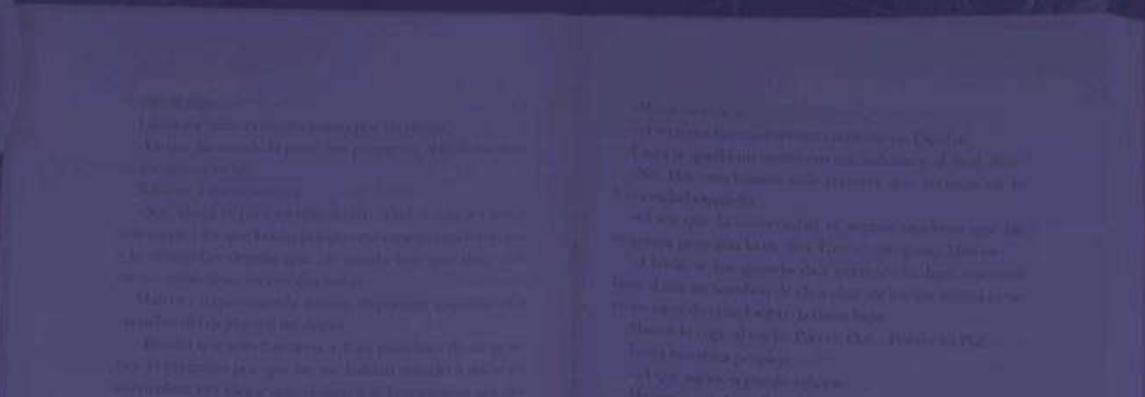
Remember the following for your exam;

- 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.
- Network ACLs contain a numbered list of rules that is evaluated in order, starting with the lowest numbered rule.
- Network ACLs have separate inbound and outbound rules, and each rule can either allow or deny traffic.
- Network ACLs are stateless; responses to allowed inbound traffic are subject to the rules for outbound traffic (and vice versa.)



Remember the following;

- 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.
- After you've created a flow log, you cannot change its configuration; for example, you can't associate a different IAM role with the flow log.





Not all IP Traffic is monitored;

- Traffic generated by instances when they contact the Amazon DNS server. If you use your own DNS server, then all traffic to that DNS server is logged.
- Traffic generated by a Windows instance for Amazon Windows license activation.
- Traffic to and from 169.254.169.254 for instance metadata.
- DHCP traffic.
- Traffic to the reserved IP address for the default VPC router.



Remember the following;

- A NAT Gateway or NAT Instance is used to provide internet traffic to EC2 instances in a private subnets.
- A Bastion is used to securely administer EC2 instances (Using SSH or RDP). Bastions are called Jump Boxes in Australia.
- You cannot use a NAT Gateway as a Bastion host.

**Remember the following;**

- Direct Connect directly connects your data center to AWS
- Useful for high throughput workloads (ie lots of network traffic)
- Or if you need a stable and reliable secure connection.



An interface endpoint is an elastic network interface with a private IP address that serves as an entry point for traffic destined to a supported service. The following services are supported:

- Amazon API Gateway
- AWS CloudFormation
- Amazon CloudWatch
- Amazon CloudWatch Events
- Amazon CloudWatch Logs
- AWS CodeBuild
- AWS Config
- Amazon EC2 API
- Elastic Load Balancing API
- AWS Key Management Service
- Amazon Kinesis Data Streams
- Amazon SageMaker and Amazon SageMaker Runtime
- Amazon SageMaker Notebook Instance
- AWS Secrets Manager
- AWS Security Token Service
- AWS Service Catalog
- Amazon SNS
- Amazon SQS
- AWS Systems Manager
- Endpoint services hosted by other AWS accounts
- Supported AWS Marketplace partner services

Exam Tips



A VPC Endpoint:

A VPC endpoint enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by PrivateLink without requiring an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Instances in your VPC do not require public IP addresses to communicate with resources in the service. Traffic between your VPC and the other service does not leave the Amazon network.

Endpoints are virtual devices. They are horizontally scaled, redundant, and highly available VPC components that allow communication between instances in your VPC and services without imposing availability risks or bandwidth constraints on your network traffic.



There are two types of VPC endpoints:

- Interface Endpoints
- Gateway Endpoints

Currently Gateway Endpoints Support:

- Amazon S3
- DynamoDB

Remember the following;

- Think of a VPC as a logical datacenter in AWS.
- Consists of IGWs (Or Virtual Private Gateways), Route Tables, Network Access Control Lists, Subnets, and Security Groups
- 1 Subnet = 1 Availability Zone
- Security Groups are Stateful; Network Access Control Lists are Stateless
- NO TRANSITIVE PEERING



Remember the following;

- When you create a VPC a default Route Table, Network Access Control List (NACL) and a default Security Group.
- It won't create any subnets, nor will it create a default internet gateway.
- US-East-1A in your AWS account can be a completely different availability zone to US-East-1A in another AWS account. The AZ's are randomized.
- Amazon always reserve 5 IP addresses within your subnets.
- You can only have 1 Internet Gateway per VPC.
- Security Groups can't span VPCs.

NAT Instances vs NAT Gateways

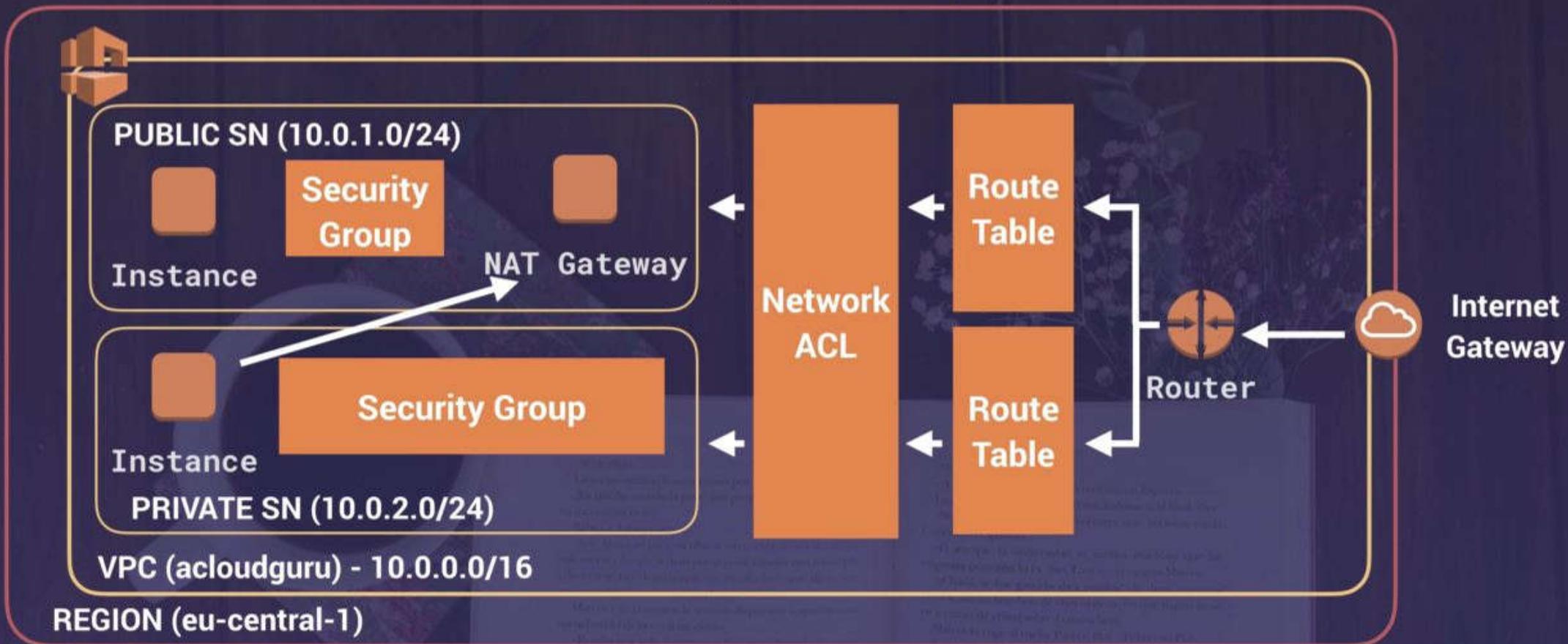


Nat Instances Exam Tips

- When creating a NAT instance, Disable Source/Destination Check on the Instance.
- NAT instances must be in a public subnet.
- There must be a route out of the private subnet to the NAT instance, in order for this to work.
- The amount of traffic that NAT instances can support depends on the instance size. If you are bottlenecking, increase the instance size.
- You can create high availability using Autoscaling Groups, multiple subnets in different AZs, and a script to automate failover.
- Behind a Security Group.

NAT Gateways Explained

VPC with Public & Private Subnet(s)



NAT Instances vs NAT Gateways

Nat Gateways

- Redundant inside the Availability Zone
- Preferred by the enterprise
- Starts at 5Gbps and scales currently to 45Gbps
- No need to patch
- Not associated with security groups
- Automatically assigned a public ip address
- Remember to update your route tables.
- No need to disable Source/Destination Checks

NAT Instances vs NAT Gateways



Nat Gateways

- If you have resources in multiple Availability Zones and they share one NAT gateway, in the event that the NAT gateway's Availability Zone is down, resources in the other Availability Zones lose internet access. To create an Availability Zone-independent architecture, create a NAT gateway in each Availability Zone and configure your routing to ensure that resources use the NAT gateway in the same Availability Zone.

Remember the following for your exam;

- Your VPC automatically comes a default network ACL, and by default it allows all outbound and inbound traffic.
- You can create custom network ACLs. 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 associate a subnet with a network ACL, the subnet is automatically associated with the default network ACL.
- Block IP Addresses using network ACLs not Security Groups



Remember the following for your exam;

- 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.
- Network ACLs contain a numbered list of rules that is evaluated in order, starting with the lowest numbered rule.
- Network ACLs have separate inbound and outbound rules, and each rule can either allow or deny traffic.
- Network ACLs are stateless; responses to allowed inbound traffic are subject to the rules for outbound traffic (and vice versa.)



Remember the following for your exam;

- You need a minimum of two public subnets to deploy an internet facing loadbalancer.

Remember the following;

- 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.
- After you've created a flow log, you cannot change its configuration; for example, you can't associate a different IAM role with the flow log.

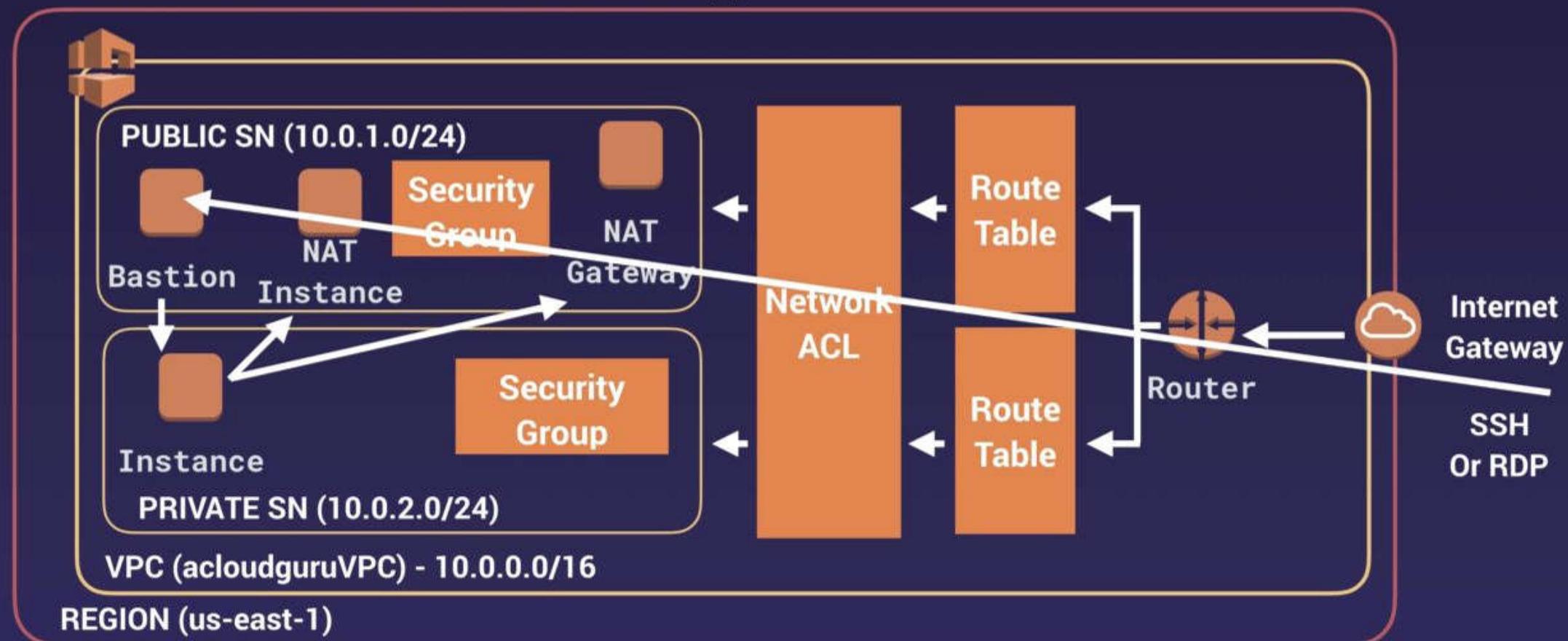


Not all IP Traffic is monitored;

- Traffic generated by instances when they contact the Amazon DNS server. If you use your own DNS server, then all traffic to that DNS server is logged.
- Traffic generated by a Windows instance for Amazon Windows license activation.
- Traffic to and from 169.254.169.254 for instance metadata.
- DHCP traffic.
- Traffic to the reserved IP address for the default VPC router.

Bastions In Action

VPC with Public & Private Subnet(s)



Bastions vs NAT Gateways/Instances



Remember the following;

- A NAT Gateway or NAT Instance is used to provide internet traffic to EC2 instances in a private subnets.
- A Bastion is used to securely administer EC2 instances (Using SSH or RDP). Bastions are called Jump Boxes in Australia.
- You cannot use a NAT Gateway as a Bastion host.

Remember the following;

- Direct Connect directly connects your data center to AWS
- Useful for high throughput workloads (ie lots of network traffic)
- Or if you need a stable and reliable secure connection.

VPC Endpoints

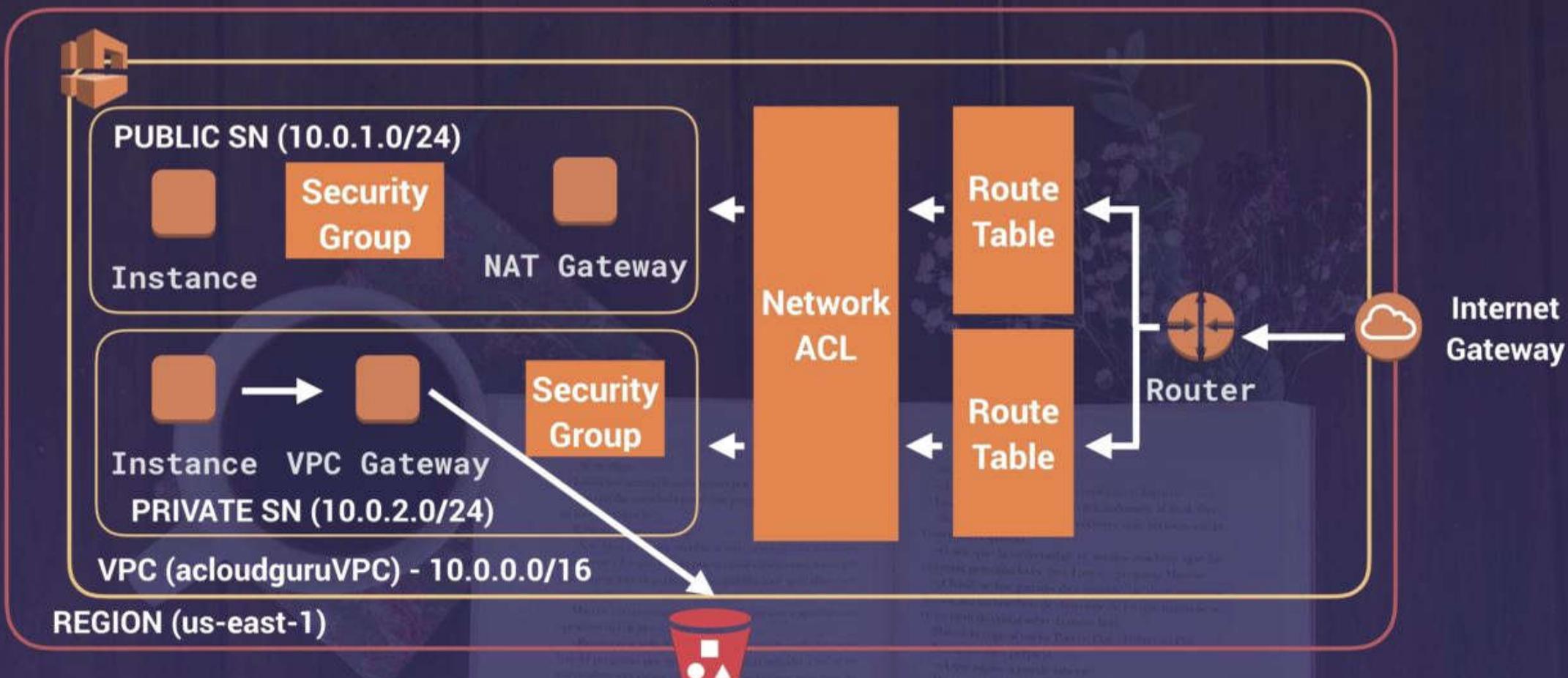


A VPC Endpoint:

A VPC endpoint enables you to privately connect your VPC to supported AWS services and VPC endpoint services powered by PrivateLink without requiring an internet gateway, NAT device, VPN connection, or AWS Direct Connect connection. Instances in your VPC do not require public IP addresses to communicate with resources in the service. Traffic between your VPC and the other service does not leave the Amazon network.

Endpoints are virtual devices. They are horizontally scaled, redundant, and highly available VPC components that allow communication between instances in your VPC and services without imposing availability risks or bandwidth constraints on your network traffic.

VPC with Public & Private Subnet(s)





There are two types of VPC endpoints:

- Interface Endpoints
- Gateway Endpoints

Currently Gateway Endpoints Support:

- Amazon S3
- DynamoDB



3 Different Types Of Load Balancers;

- Application Load Balancers
- Network Load Balancers
- Classic Load Balancers

- 504 Error means the gateway has timed out. This means that the application not responding within the idle timeout period.
- Trouble shoot the application. Is it the Web Server or Database Server?





- If you need the IPv4 address of your end user, look for the **X-Forwarded-For** header.



- 504 Error means the gateway has timed out. This means that the application not responding within the idle timeout period.
- Trouble shoot the application. Is it the Web Server or Database Server?





- If you need the IPv4 address of your end user, look for the **X-Forwarded-For** header.





- Instances monitored by ELB are reported as ; InService , or OutofService
- Health Checks check the instance health by talking to it.
- Load Balances have their own DNS name. You are never given an IP address.
- Read the ELB FAQ for Classic Load Balancers.
- Want to deep dive on application load balancers? Check out our deep dive course!

Advanced Load Balancer Theory

- Sticky Sessions enable your users to stick to the same EC2 instance.
Can be useful if you are storing information locally to that instance.
- Cross Zone Load Balancing enables you to load balance across multiple availability zones.
- Path patters allow you to direct traffic to different EC2 instances based on the URL contained in the request.

HA Sample Exam Question

Scenario: You have a website that requires a minimum of 6 instances and it must be highly available. You must also be able to tolerate the failure of 1 Availability Zone. What is the ideal architecture for this environment while also being the most cost effective?

- 2 Availability Zones with 2 instances in each AZ.
- 3 Availability Zones with 3 instances in each AZ.
- 1 Availability Zone with 6 instances in each AZ.
- 3 Availability Zones with 2 instances in each AZ.

Remember the following

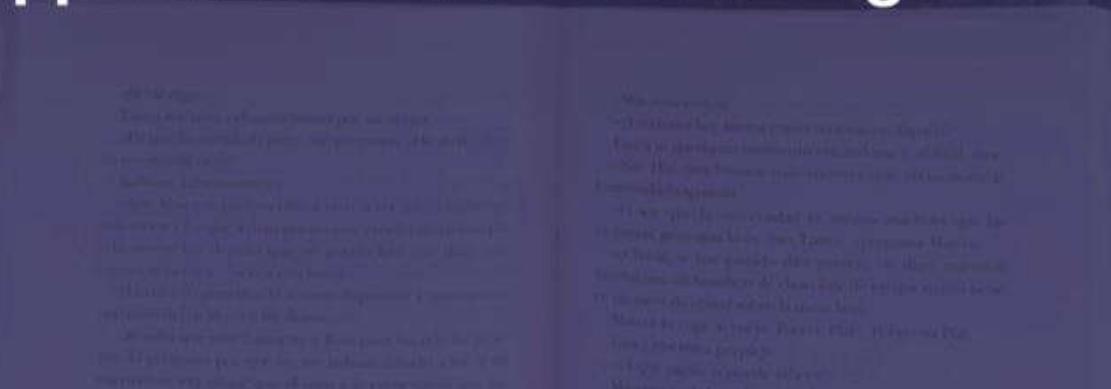
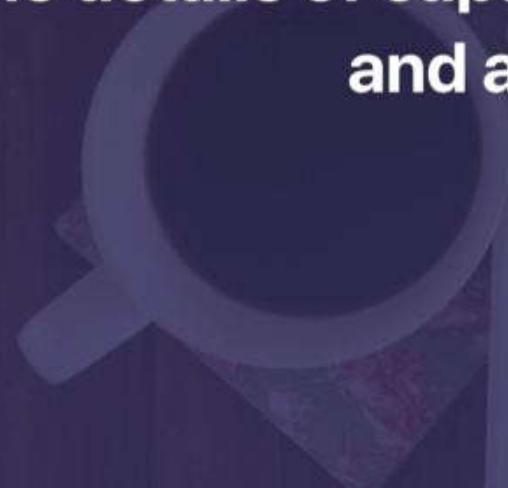
- Always Design for failure.
- Use Multiple AZ's and Multiple Regions where ever you can.
- Know the difference between Multi-AZ and Read Replicas for RDS.
- Know the difference between scaling out and scaling up.
- Read the question carefully and always consider the cost element.
- Know the different S3 storage classes.



CloudFormation

- Is a way of completely scripting your cloud environment
- Quick Start is a bunch of CloudFormation templates already built by AWS Solutions Architects allowing you to create complex environments very quickly.

With Elastic Beanstalk, you can quickly deploy and manage applications in the AWS Cloud without worrying about the infrastructure that runs those applications. You simply upload your application, and Elastic Beanstalk automatically handles the details of capacity provisioning, load balancing, scaling, and application health monitoring.





3 Different Types Of Load Balancers;

- Application Load Balancers
- Network Load Balancers
- Classic Load Balancers



Load balancers are used to distribute traffic across multiple servers. They can be used to improve performance, availability, and scalability. There are three main types of load balancers:

- Application Load Balancers: These are typically implemented at the application layer (Layer 7) and can handle both HTTP and HTTPS traffic. They can route requests based on various criteria such as IP address, port, or URL path.
- Network Load Balancers: These are typically implemented at the network layer (Layer 4) and can handle both TCP and UDP traffic. They can route requests based on source and destination IP addresses and ports.
- Classic Load Balancers: These are typically implemented at the transport layer (Layer 4) and can handle both TCP and UDP traffic. They can route requests based on source and destination IP addresses and ports.

- 504 Error means the gateway has timed out. This means that the application not responding within the idle timeout period.
- Troubleshoot the application. Is it the Web Server or Database Server?





- Instances monitored by ELB are reported as ;
InService , or OutofService
- Health Checks check the instance health by talking to it.
- Load Balances have their own DNS name. You are never given an IP address.
- Read the ELB FAQ for Classic Load Balancers.
- Want to deep dive on application load balancers? Check out our deep dive course!

Advanced Load Balancer Theory

- Sticky Sessions enable your users to stick to the same EC2 instance.
Can be useful if you are storing information locally to that instance.
- Cross Zone Load Balancing enables you to load balance across multiple availability zones.
- Path patterns allow you to direct traffic to different EC2 instances based on the URL contained in the request.



CloudFormation

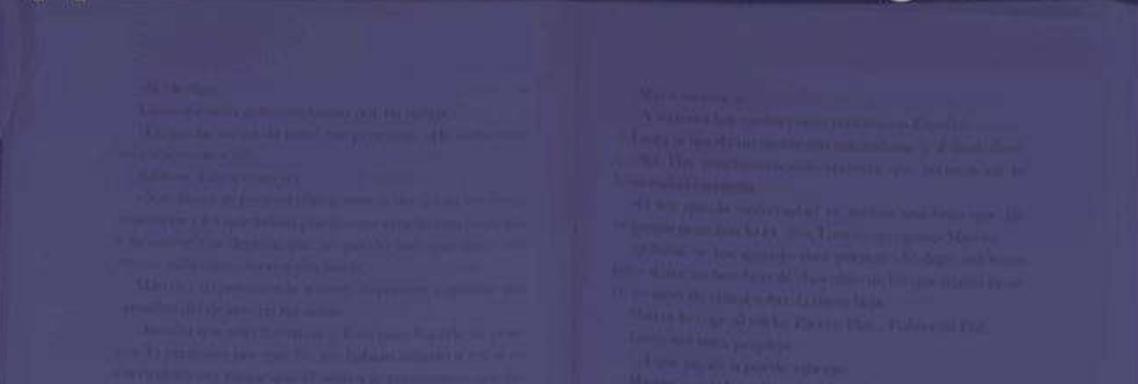
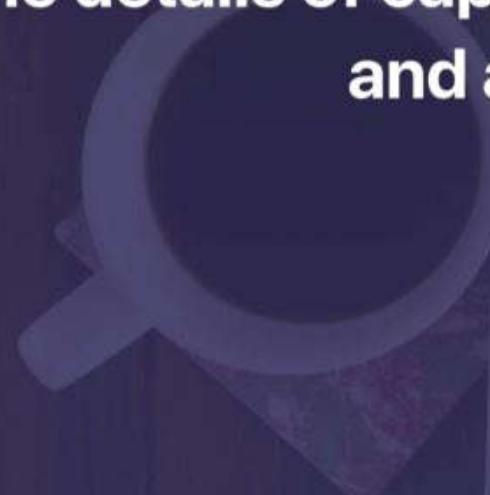
- Is a way of completely scripting your cloud environment
- Quick Start is a bunch of CloudFormation templates already built by AWS Solutions Architects allowing you to create complex environments very quickly.



CloudFormation is a service that allows you to define your AWS resources using a JSON or YAML-based template. This template describes the resources you want to create, their properties, and how they interact with each other. Once you've created a CloudFormation template, you can use it to provision your resources in a repeatable and predictable way. You can also use CloudFormation to update existing resources or delete them entirely. CloudFormation is particularly useful for creating complex environments, such as databases, web servers, and storage systems, in a consistent and reliable manner. It's also great for automating the deployment of infrastructure, so you can focus on building your application instead of managing the underlying infrastructure.



With Elastic Beanstalk, you can quickly deploy and manage applications in the AWS Cloud without worrying about the infrastructure that runs those applications. You simply upload your application, and Elastic Beanstalk automatically handles the details of capacity provisioning, load balancing, scaling, and application health monitoring.





SQS Exam Tips

- SQS is pull based, not pushed based.
- Messages are 256 KB in size.
- Messages can be kept in the queue from 1 minute to 14 days; the default retention period is 4 days.



SQS Exam Tips

- Visibility Time Out 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. This could result in the same message being delivered twice.
- Visibility timeout maximum is 12 hours.

SQS Exam Tips

- SQS guarantees that your messages will be processed at least once.
- Amazon SQS long polling is a way to retrieve messages from your Amazon SQS queues. While the regular short polling returns immediately (even if the message queue being polled is empty), long polling doesn't return a response until a message arrives in the message queue, or the long poll times out.

SWF vs SQS

- SQS has a retention period of up to 14 days; with SWF, workflow executions can last up to 1 year.
- Amazon SWF presents a task-oriented API, whereas Amazon SQS offers a message-oriented API.
- Amazon SWF ensures that a task is assigned only once and is never duplicated. With Amazon SQS, you need to handle duplicated messages and may also need to ensure that a message is processed only once.
- Amazon SWF keeps track of all the tasks and events in an application. With Amazon SQS, you need to implement your own application-level tracking, especially if your application uses multiple queues.



SWF Actors

- Workflow Starters — An application that can initiate (start) a workflow. Could be your e-commerce website following the placement of an order, or a mobile app searching for bus times.
- Deciders — Control the flow of activity tasks in a workflow execution. If something has finished (or failed) in a workflow, a Decider decides what to do next.
- Activity Workers — Carry out the activity tasks.

SNS Benefits

- Instantaneous, push-based delivery (no polling)
- Simple APIs and easy integration with applications
- 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

SNS vs SQS?

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



Just remember that Elastic Transcoder is a media transcoder in the cloud. It converts media files from their original source format in to different formats that will play on smartphones, tablets, PCs, etc.



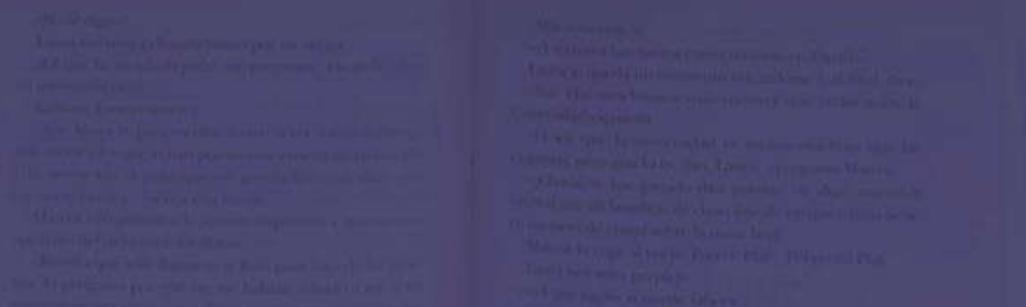


API Gateway Exam Tips

- Remember what API Gateway is at a high level
- API Gateway has caching capabilities to increase performance
- API Gateway is low cost and scales automatically
- You can throttle API Gateway to prevent attacks
- You can log results to CloudWatch
- If you are using Javascript/AJAX that uses multiple domains with API Gateway, ensure that you have enabled CORS on API Gateway
- CORS is enforced by the client

Kinesis Exam Tips

- Know the difference between Kinesis Streams and Kinesis Firehose. You will be given scenario questions and you must choose the most relevant service.
- Understand what Kinesis Analytics is.



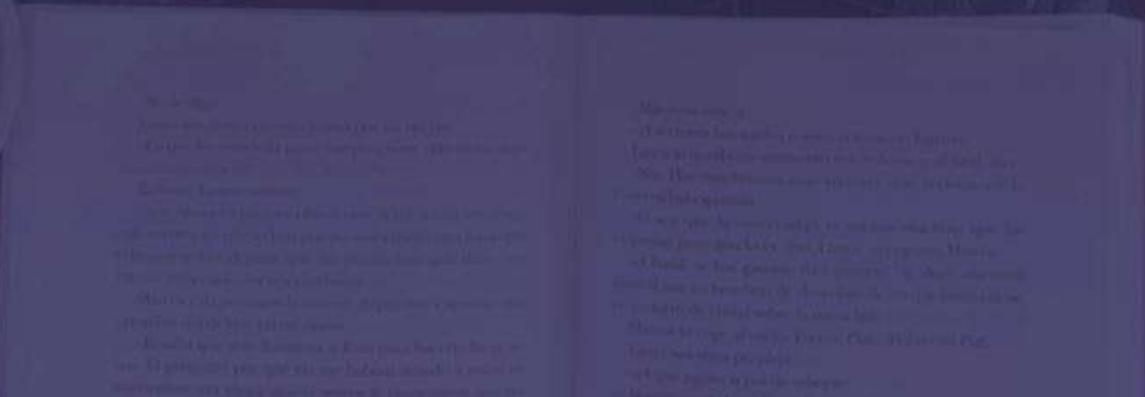
Cognito Exam Tips

- Federation allows users to authenticate with a Web Identity Provider (Google, Facebook, Amazon)
- The user authenticates first with the Web ID Provider and receives an authentication token, which is exchanged for temporary AWS credentials allowing them to assume an IAM role.
- Cognito is an Identity Broker which handles interaction between your applications and the Web ID provider (You don't need to write your own code to do this.)



Cognito Exam Tips

- User pool is user based. It handles things like user registration, authentication, and account recovery.
- Identity pools authorise access to your AWS resources.



SQS Exam Tips

- SQS is a way to de-couple your infrastructure
- SQS is pull based, not pushed based.
- Messages are 256 KB in size.
- Messages can be kept in the queue from 1 minute to 14 days; the default retention period is 4 days.
- Standard SQS and FIFO SQS
- Standard order is not guaranteed and messages can be delivered more than once.
- FIFO order is strictly maintained and messages are delivered only once.

SQS Exam Tips

- Visibility Time Out 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. This could result in the same message being delivered twice.



Visibility Time Out
When a message is received from an SQS queue, it becomes invisible for the duration of the visibility timeout. If the message is not processed by the end of the visibility timeout, it becomes visible again and can be received by another consumer. This allows for failover and ensures that messages are not lost if a consumer fails or takes longer than expected to process them.

SQS Exam Tips

- SQS guarantees that your messages will be processed at least once.
- Amazon SQS long polling is a way to retrieve messages from your Amazon SQS queues. While the regular short polling returns immediately (even if the message queue being polled is empty), long polling doesn't return a response until a message arrives in the message queue, or the long poll times out.

SWF vs SQS

- SQS has a retention period of up to 14 days; with SWF, workflow executions can last up to 1 year.
- Amazon SWF presents a task-oriented API, whereas Amazon SQS offers a message-oriented API.
- Amazon SWF ensures that a task is assigned only once and is never duplicated. With Amazon SQS, you need to handle duplicated messages and may also need to ensure that a message is processed only once.
- Amazon SWF keeps track of all the tasks and events in an application. With Amazon SQS, you need to implement your own application-level tracking, especially if your application uses multiple queues.

SWF Actors - Exam Tips

- Workflow Starters — An application that can initiate (start) a workflow. Could be your e-commerce website following the placement of an order, or a mobile app searching for bus times.
- Deciders — Control the flow of activity tasks in a workflow execution. If something has finished (or failed) in a workflow, a Decider decides what to do next.
- Activity Workers — Carry out the activity tasks.

SNS Benefits

- Instantaneous, push-based delivery (no polling)
- Simple APIs and easy integration with applications
- 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

SNS vs SQS?

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

Just remember that Elastic Transcoder is a media transcoder in the cloud. It converts media files from their original source format in to different formats that will play on smartphones, tablets, PCs, etc.



API Gateway Exam Tips

- Remember what API Gateway is at a high level
- API Gateway has caching capabilities to increase performance
- API Gateway is low cost and scales automatically
- You can throttle API Gateway to prevent attacks
- You can log results to CloudWatch
- If you are using Javascript/AJAX that uses multiple domains with API Gateway, ensure that you have enabled CORS on API Gateway



Kinesis Exam Tips

- Know the difference between Kinesis Streams and Kinesis Firehose. You will be given scenario questions and you must choose the most relevant service.
- Understand what Kinesis Analytics is.



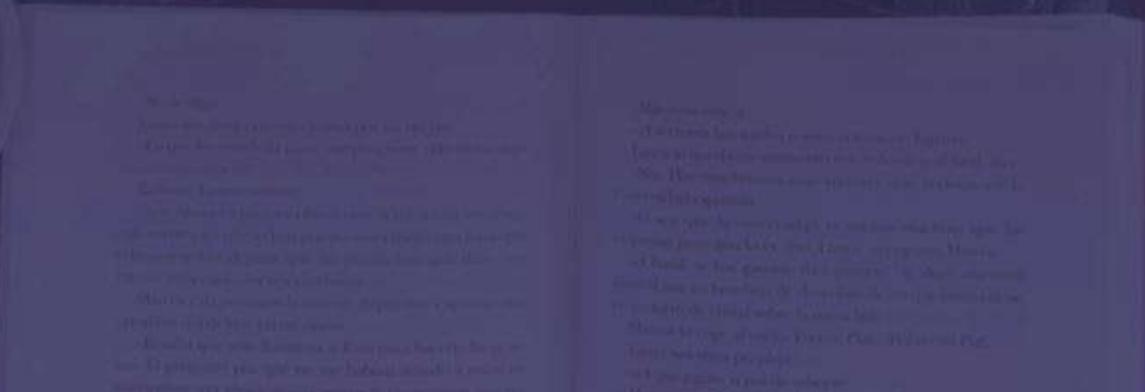
Cognito Exam Tips

- Federation allows users to authenticate with a Web Identity Provider (Google, Facebook, Amazon)
- The user authenticates first with the Web ID Provider and receives an authentication token, which is exchanged for temporary AWS credentials allowing them to assume an IAM role.
- Cognito is an Identity Broker which handles interaction between your applications and the Web ID provider (You don't need to write your own code to do this.)



Cognito Exam Tips

- User pool is user based. It handles things like user registration, authentication, and account recovery.
- Identity pools authorise access to your AWS resources.





Lambda Exam Tips

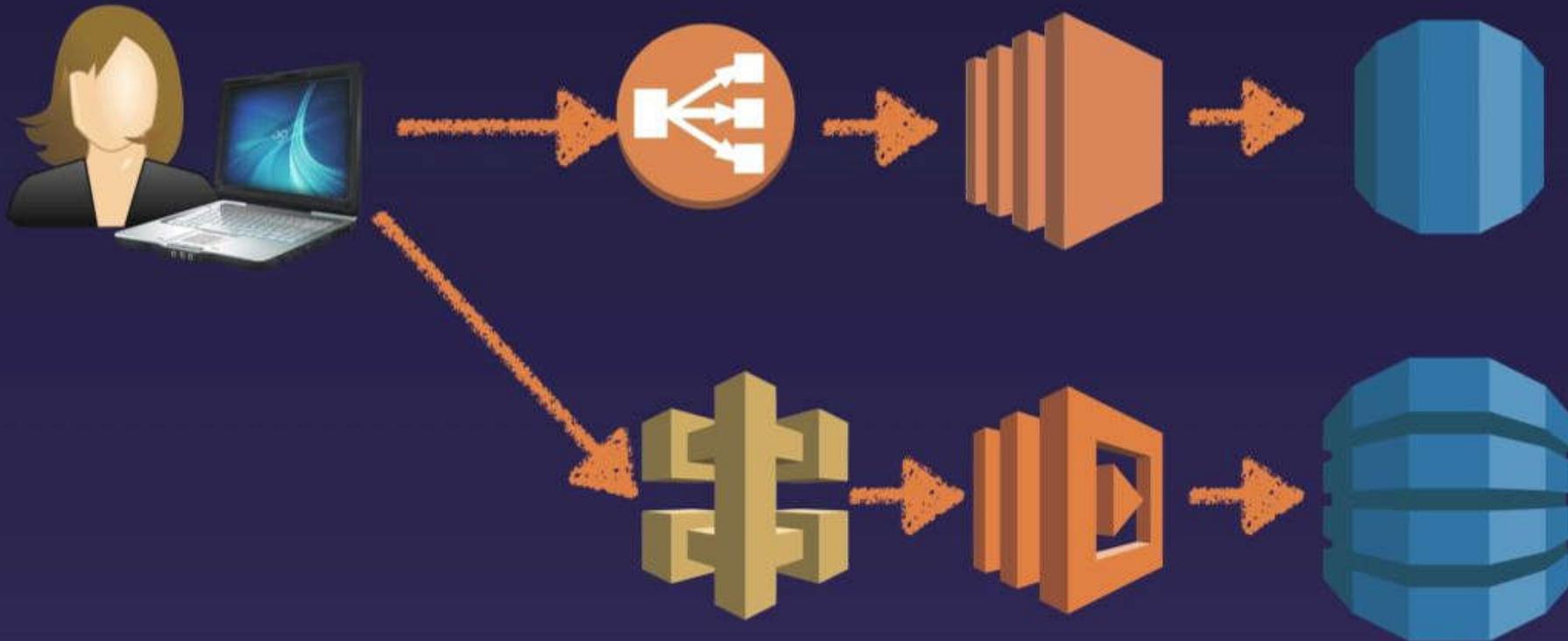
- Lambda scales out (not up) automatically
- Lambda functions are independent, 1 event = 1 function
- Lambda is serverless
- Know what services are serverless!
- Lambda functions can trigger other lambda functions, 1 event can = x functions if functions trigger other functions



Lambda Exam Tips

- Architectures can get extremely complicated, AWS X-ray allows you to debug what is happening
- Lambda can do things globally, you can use it to back up S3 buckets to other S3 buckets etc
- Know your triggers

Traditional Architecture



Serverless Architecture



Lambda Exam Tips

- Lambda scales out (not up) automatically
- Lambda functions are independent, 1 event = 1 function
- Lambda is serverless
- Know what services are serverless!
- Lambda functions can trigger other lambda functions, 1 event can = x functions if functions trigger other functions

Lambda Exam Tips

- Architectures can get extremely complicated, AWS X-ray allows you to debug what is happening
- Lambda can do things globally, you can use it to back up S3 buckets to other S3 buckets etc
- Know your triggers

