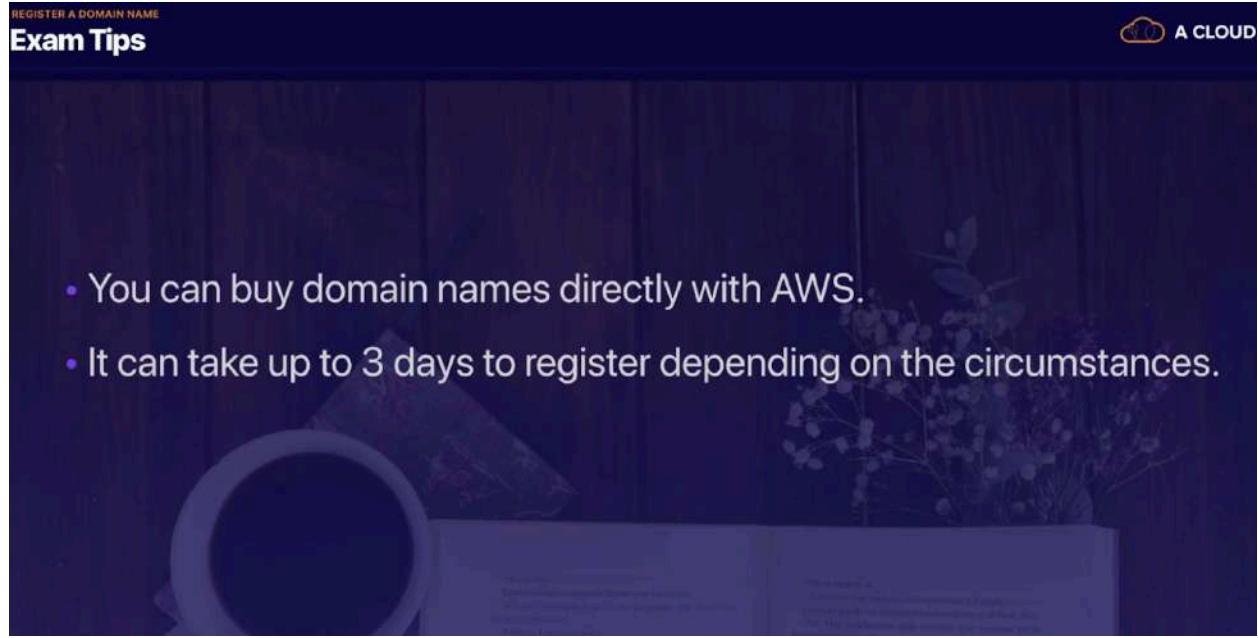


## Route53 Exam Tips

- 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



## Route53 Routing Policies Available On AWS

**The Following Routing Policies Are Available With Route53:**

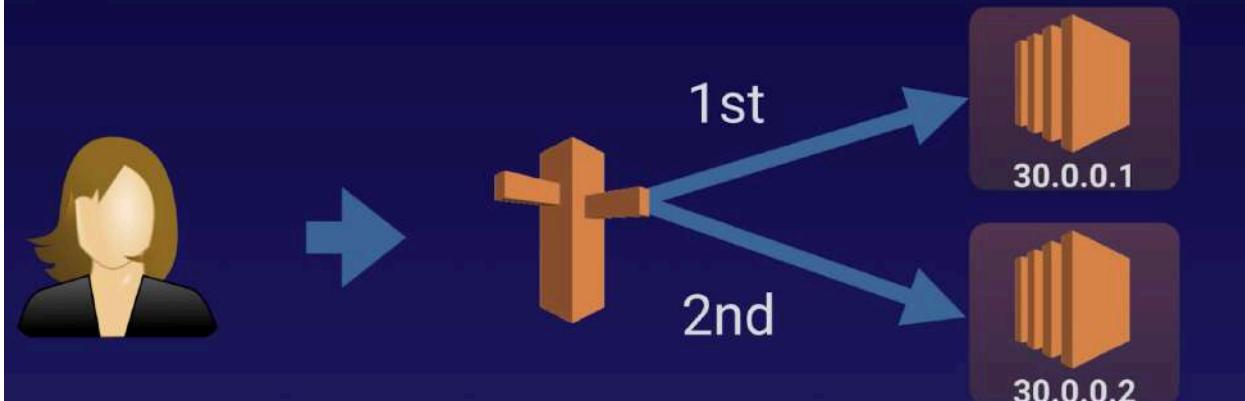
- Simple Routing
- Weighted Routing
- Latency-based Routing
- Failover Routing
- Geolocation Routing
- Geoproximity Routing (Traffic Flow Only)
- Multivalue Answer Routing

## Simple Routing Policy Lab

### 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.

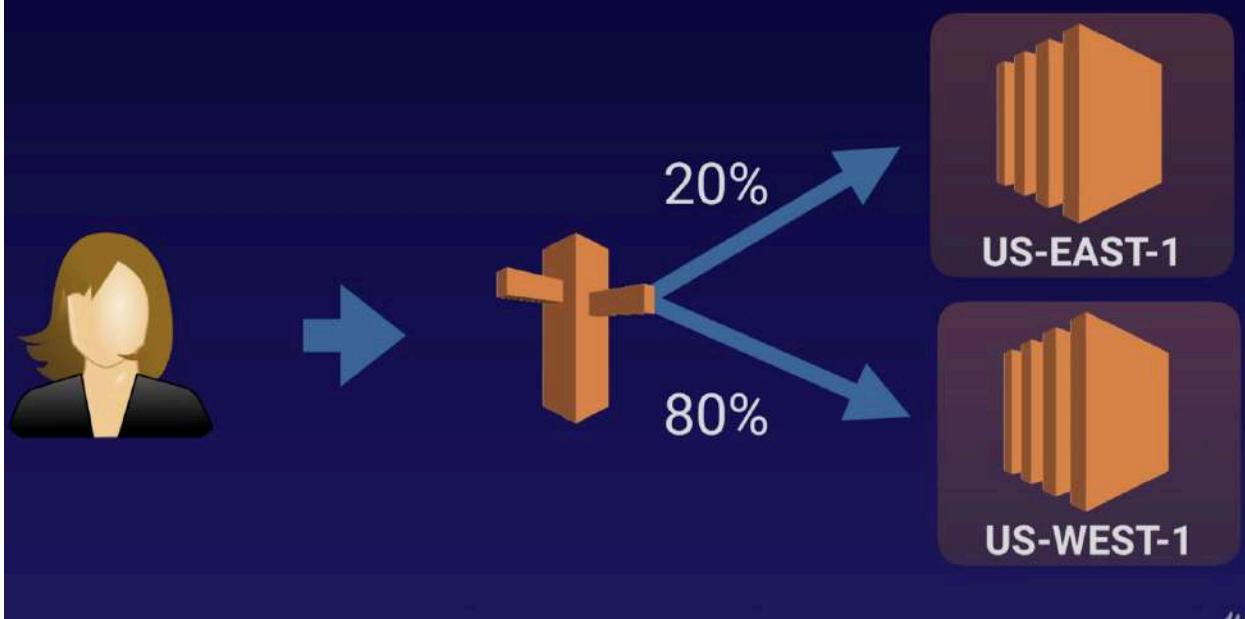
If you specify multiple values in a record, Route 53 returns all values to the user in a random order.

### Weighted Routing Policy Lab

## Weighted Routing Policy

Allows you split your traffic based on different weights assigned.

For example, you can set 10% of your traffic to go to US-EAST-1 and 90% to go to EU-WEST-1.



## 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.

## Latency Routing Policy

ROUTE53 LATENCY ROUTING POLICY

## Latency-Based Routing

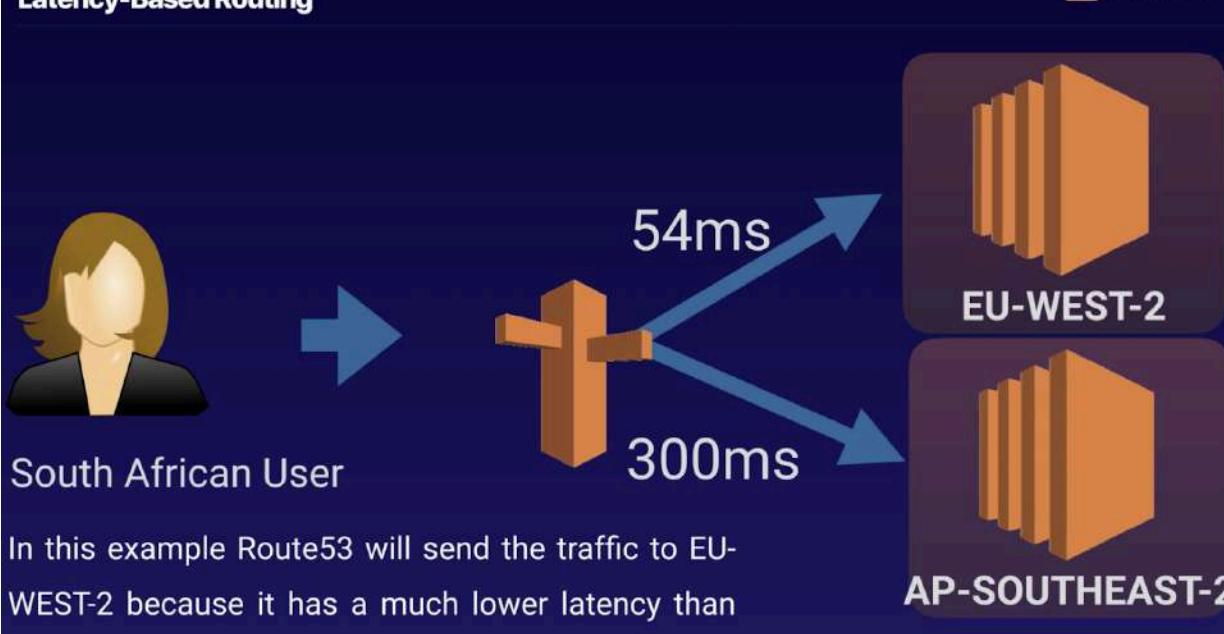
Allows you to route your traffic based on the lowest network latency for your end user (ie which region will give them the fastest response time).

To use latency-based routing, you create a latency resource record set for the Amazon EC2 (or ELB) resource in each region that hosts your website. When Amazon Route 53 receives a query for your site, it selects the latency resource record set for the region that gives the user the lowest latency. Route 53 then responds with the value associated with that resource record set.



ROUTE53 LATENCY ROUTING POLICY

## Latency-Based Routing



South African User

54ms

300ms

EU-WEST-2

AP-SOUTHEAST-2

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

## Failover Routing Policy

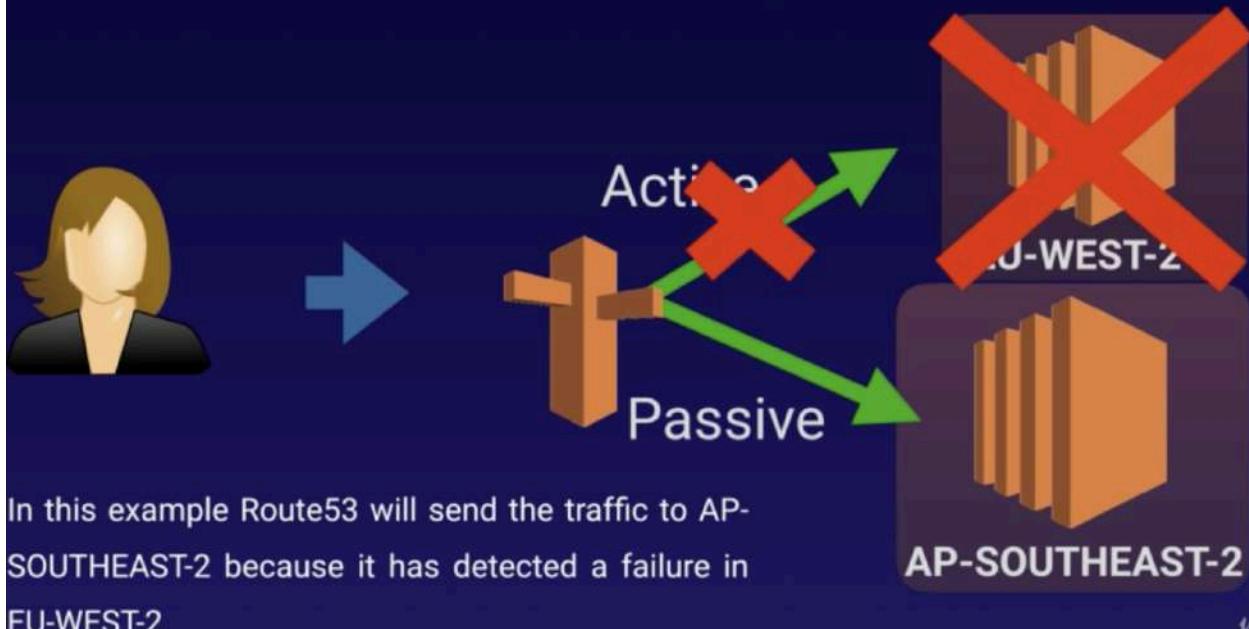
## Failover Routing Policy

Failover routing policies are used when you want to create an active/passive set up. For example, you may want your primary site to be in EU-WEST-2 and your secondary DR Site in AP-SOUTHEAST-2.

Route53 will monitor the health of your primary site using a health check.



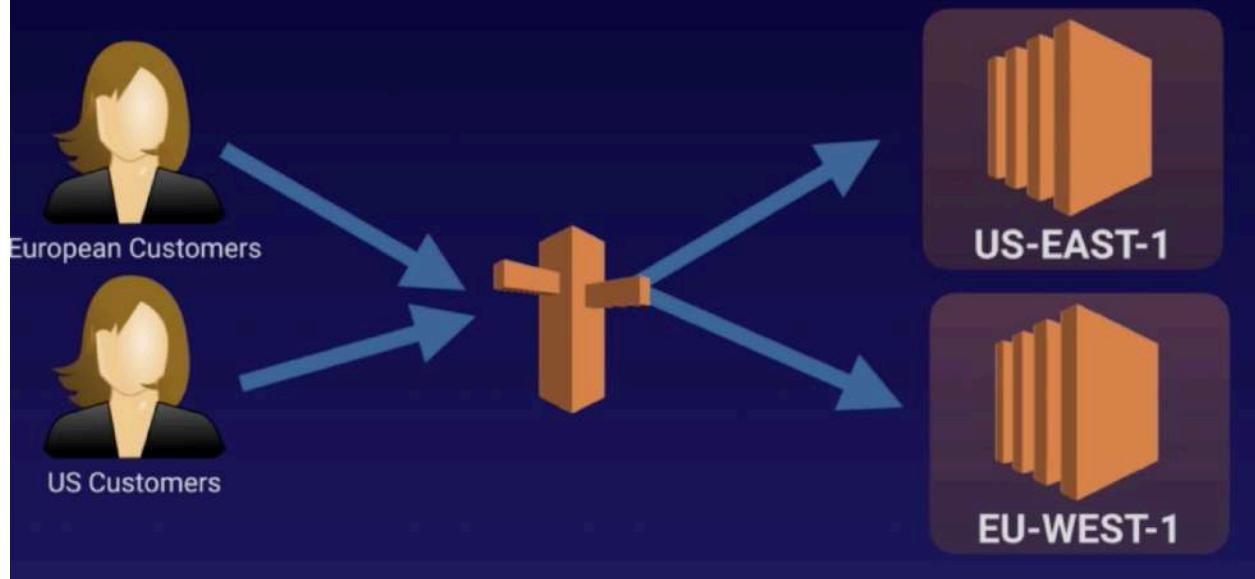
A health check monitors the health of your end points.



## Geolocation Routing Policy

## 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 which DNS queries originate). For example, you might want all queries from Europe to be routed to a fleet of EC2 instances that are specifically configured for your European customers. These servers may have the local language of your European customers and all prices are displayed in Euros.



## Geoproximity Routing Policy (Traffic Flow Only)

## 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

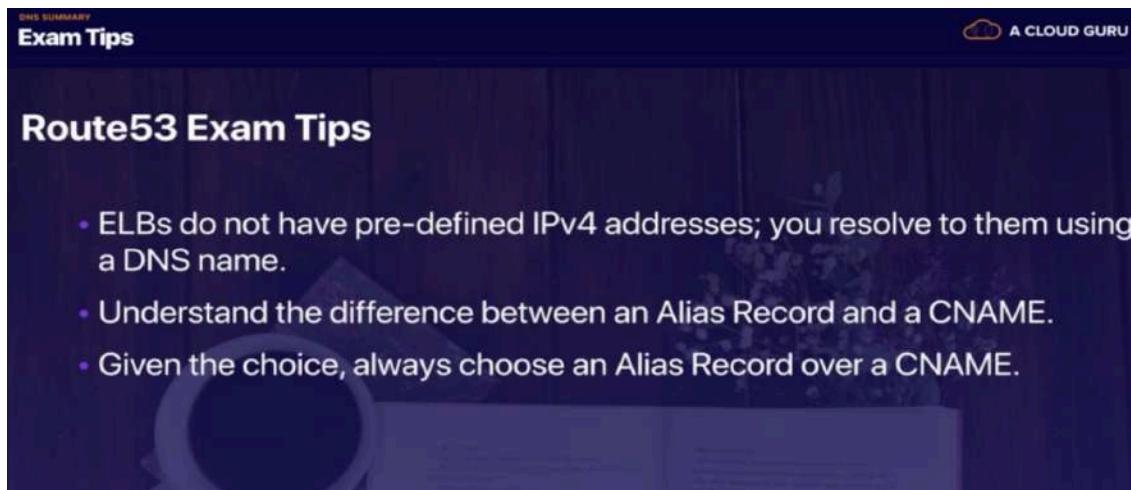
## Multivalue Answer Policy

Multivalue answer routing lets you configure Amazon Route 53 to return multiple values, such as IP addresses for your web servers, in response to DNS queries. You can specify multiple values for almost any record, but multivalue answer routing also lets you check the health of each resource, so Route 53 returns only values for healthy resources.

**This is similar to simple routing however it allows you to put health checks on each record set.**



## Route53 Summary



## 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

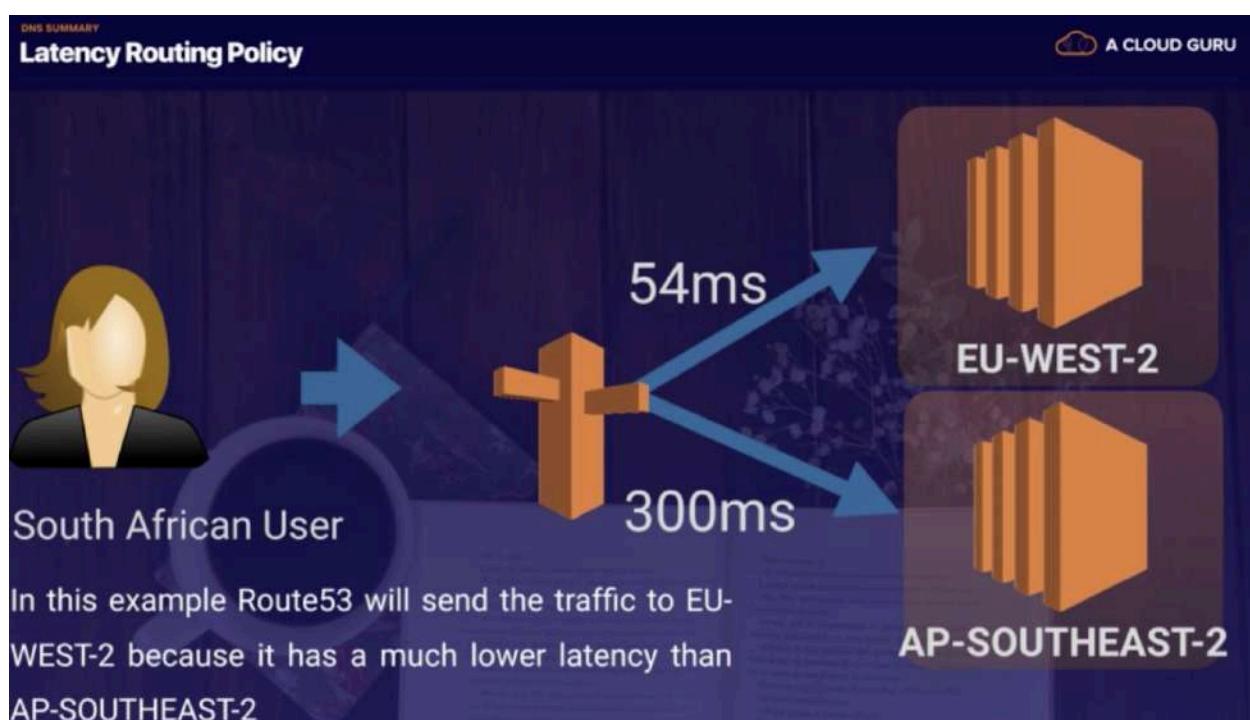
## Health Checks

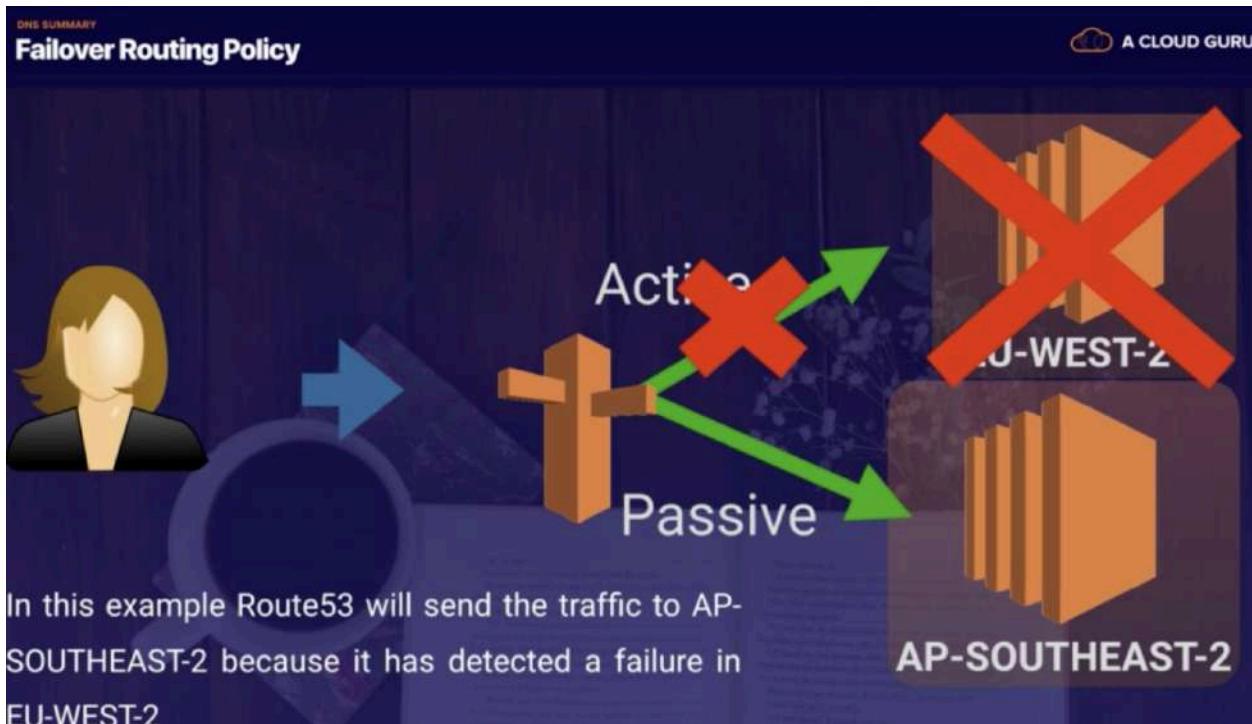
- You can set health checks on individual record sets.
- If a record set fails a health check it will be removed from Route53 until it passes the health check.
- You can set SNS notifications to alert you if a health check is failed.

### Simple Routing Policy

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.







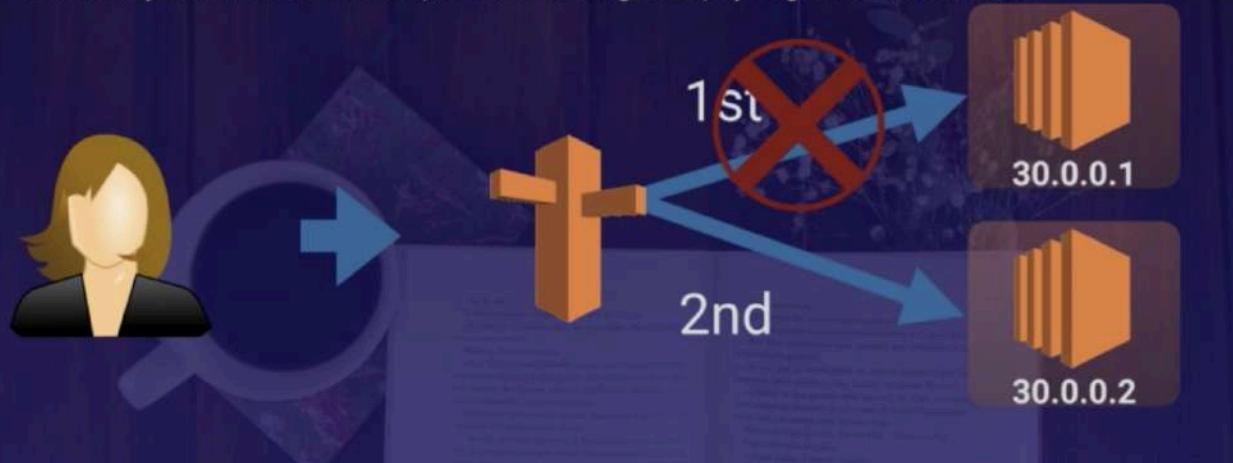
## Geoproximity Routing (Traffic Flow Only)

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**.



Quiz on Route53

✓ Good job!

Failover Routing and Latency-based Routing are the only two correct options, as they consider routing data based on whether the resource is healthy or whether one set of resources is more performant than another. Any answer containing location based routing (Geoproximity and Geolocation) cannot be correct in this case, as these types only consider where the client or resources are located before routing the data. They do not take into account whether a resource is online or slow. Simple Routing can also be discounted as it does not take into account the state of the resources.

Question 1:

Which of the following Route 53 policies allow you to a) route data to a second resource if the first is unhealthy, and b) route data to resources that have better performance?

Failover Routing and Simple Routing

Geoproximity Routing and Geolocation Routing

Geolocation Routing and Latency-based Routing

Failover Routing and Latency-based Routing

 Good job!

Question 2:

**Route 53 is Amazon's DNS Service.**

TRUE

FALSE

 Good job!

Question 3:

**Route 53 is named so because \_\_\_\_\_.**

It was invented in 1953.

Route 66 was already registered with Microsoft.

The DNS Port is on Port 53 and Route 53 is a DNS Service.

Beats me – only people in marketing can tell you the reason behind its name.

✓ Good job!

Question 4:

You have created a new subdomain for your popular website, and you need this subdomain to point to an Elastic Load Balancer using Route53. Which DNS record set should you create?

A

AAAA

MX

CNAME

✓ Good job!

Question 5:

True or False: There is a limit to the number of domain names that you can manage using Route 53.

True. There is a hard limit of 10 domain names. You cannot go above this number.

True and False. With Route 53, there is a default limit of 50 domain names. However, this limit can be increased by contacting AWS support.

False. By default, you can support as many domain names on Route 53 as you want.

✓ Good job!

Multivalue answer routing lets you configure Amazon Route 53 to return multiple values, such as IP addresses for your web servers, in response to DNS queries. Route 53 responds to DNS queries with up to eight healthy records and gives different answers to different DNS resolvers. The choice of which to use is left to the requesting service effectively creating a form or randomization.

Question 7:

Your company hosts 10 web servers all serving the same web content in AWS. They want Route 53 to serve traffic to random web servers. Which routing policy will meet this requirement, and provide the best resiliency?

Simple Routing

Weighted Routing

Multivalue Routing

Latency Routing

**VPC**

**Think of a VPC as a virtual data centre in the cloud.**



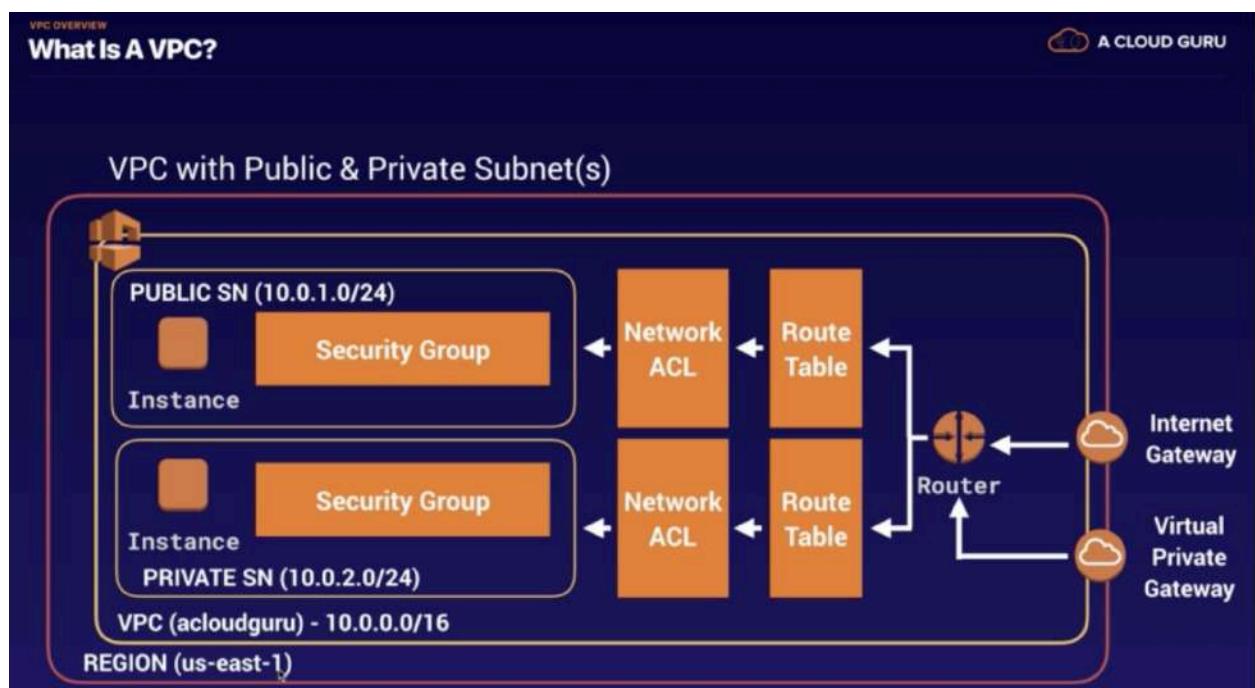
Amazon Virtual Private Cloud (Amazon VPC) lets you provision a logically isolated section of the Amazon Web Services (AWS) Cloud where you can launch AWS resources in a virtual network that you define. You have complete control over your virtual networking environment, including selection of your own IP address range, creation of subnets, and configuration of route tables and network gateways.

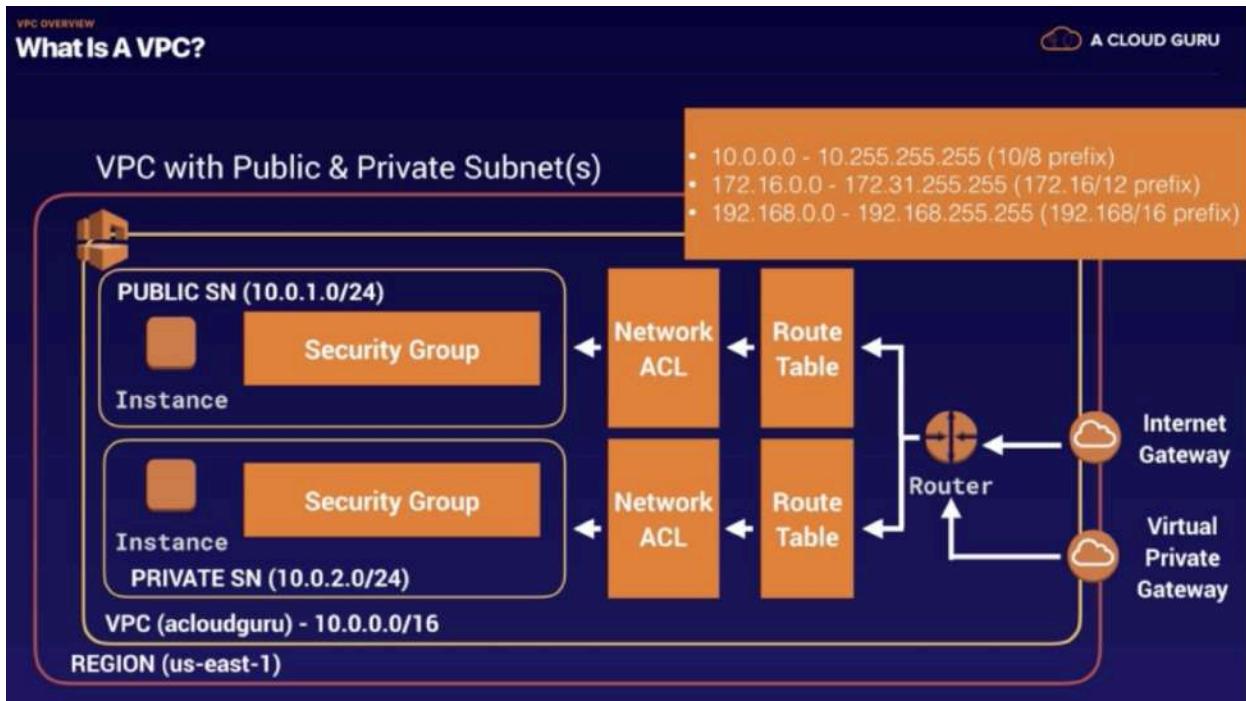


## What Is A VPC?

You can easily customize the network configuration for your Amazon Virtual Private Cloud. For example, you can create a public-facing subnet for your webservers that has access to the Internet, and place your backend systems such as databases or application servers in a private-facing subnet with no Internet access. You can leverage multiple layers of security, including security groups and network access control lists, to help control access to Amazon EC2 instances in each subnet.







VPC OVERVIEW

## VPC Features

A CLOUD GURU

### Default VPC vs Custom VPC

- Default VPC is user friendly, allowing you to immediately deploy instances.
- All Subnets in default VPC have a route out to the internet.
- Each EC2 instance has both a public and private IP address.



VPC OVERVIEW

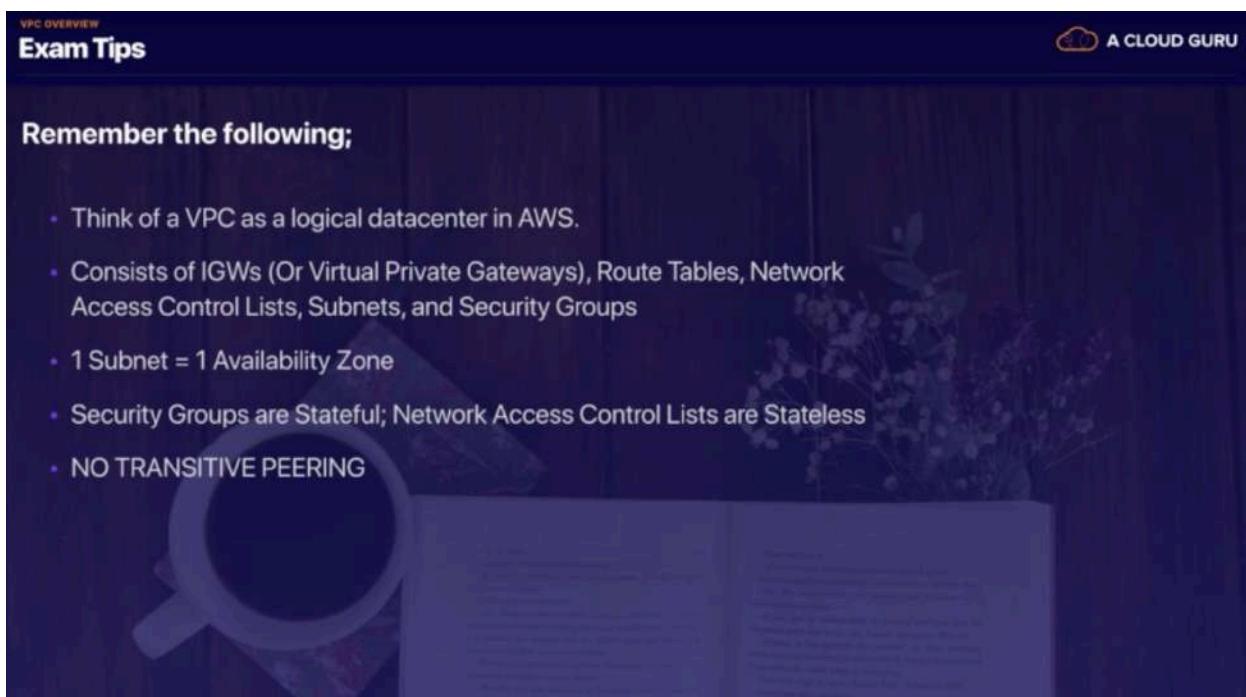
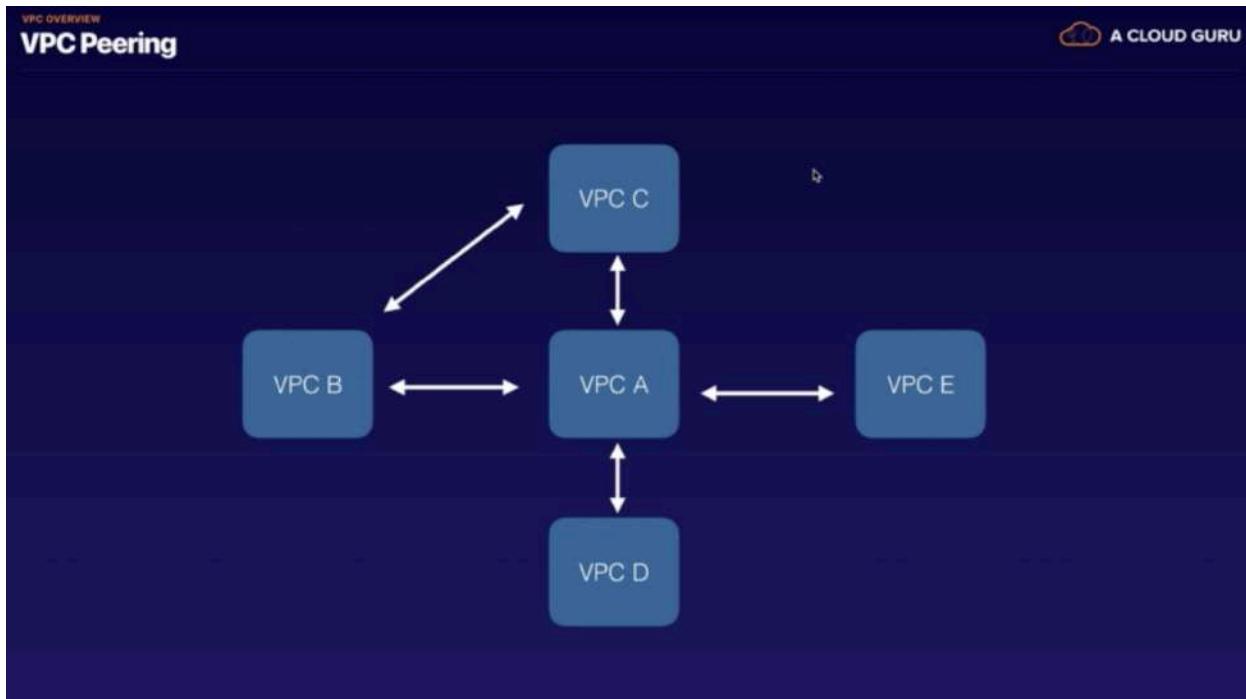
## VPC Features

A CLOUD GURU

### VPC Peering

- Allows you to connect one VPC with another via a direct network route using private IP addresses.
- Instances behave as if they were on the same private network
- You can peer VPC's with other AWS accounts as well as with other VPCs in the same account.
- Peering is in a star configuration: ie 1 central VPC peers with 4 others.  
NO TRANSITIVE PEERING!!!





## VPC-Lab

## What Is A VPC?

### VPC with Public & Private Subnet(s)

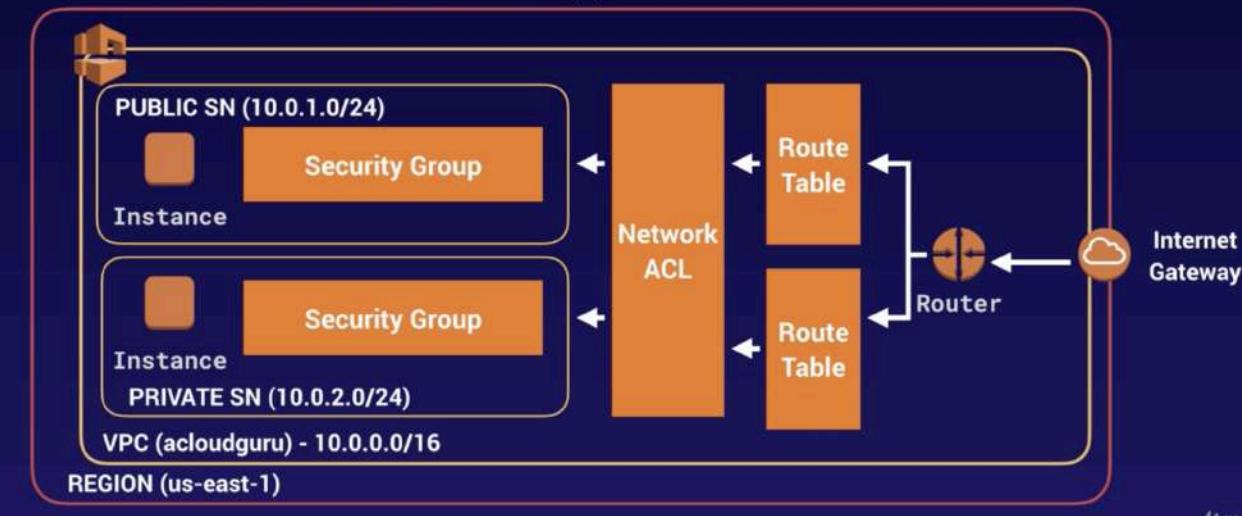


## What Is A VPC?

### VPC with Public & Private Subnet(s)



## VPC with Public &amp; Private Subnet(s)

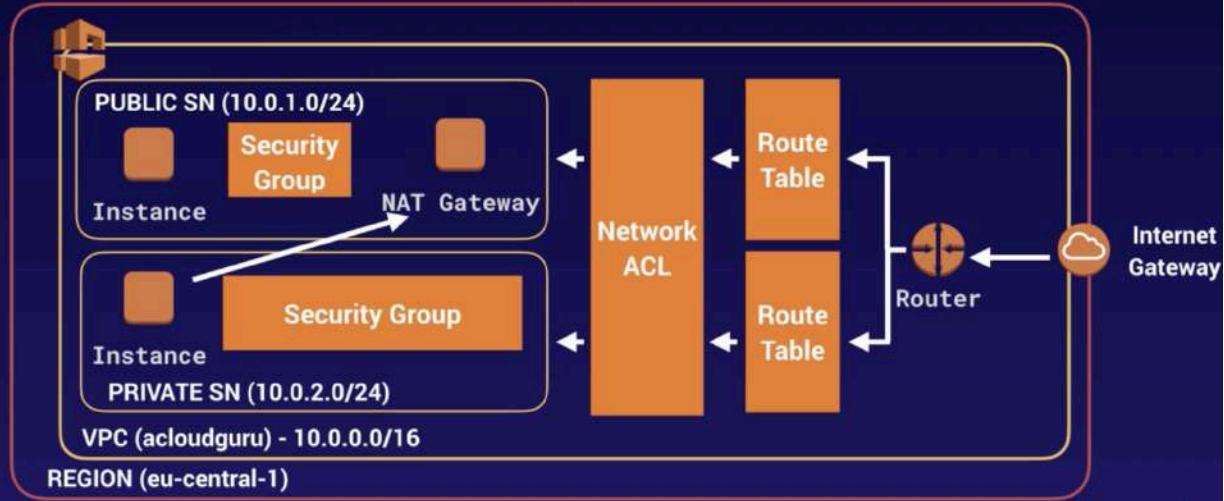


## Remember the following:

- 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 Gateway

## VPC with Public &amp; Private Subnet(s)



## 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

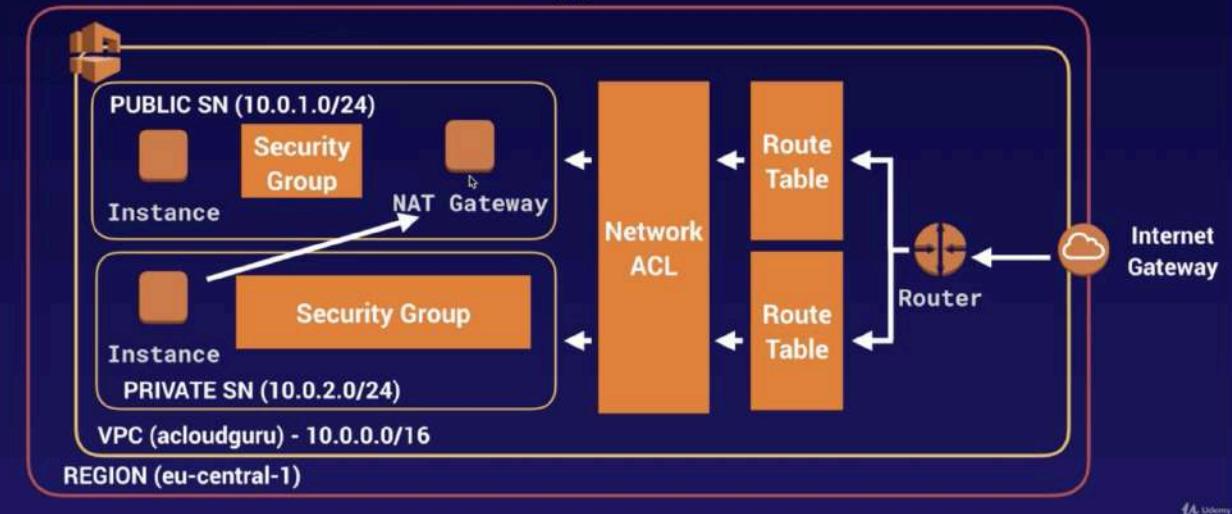
- Redundant inside the Availability Zone
- Preferred by the enterprise
- Starts at 5Gbps and scales currently to 45Gbps
- No need to patch
- Not associated with security groups
- Automatically assigned a public ip address
- Remember to update your route tables.
- No need to disable Source/Destination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

## Access Control Lists (ACL)

## VPC with Public &amp; Private Subnet(s)



## Exam Tips

## Remember the following for your exam:

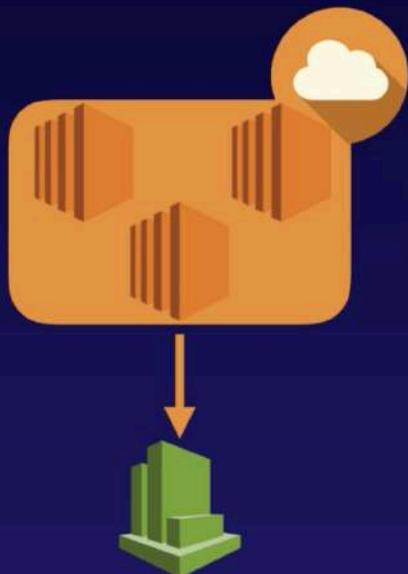
- Your VPC automatically comes with a default network ACL, and by default it allows all outbound and inbound traffic.
- You can create custom network ACLs. By default, each custom network ACL denies all inbound and outbound traffic until you add rules.
- Each subnet in your VPC must be associated with 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

**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.)

**VPC Flow Logs****What Are VPC Flow Logs?**

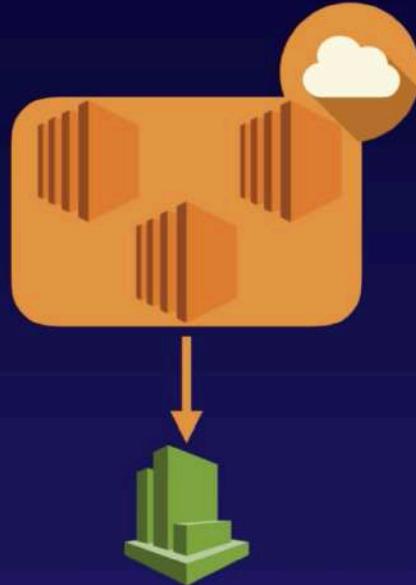
VPC Flow Logs is a feature that enables you to capture information about the IP traffic going to and from network interfaces in your VPC. Flow log data is stored using Amazon CloudWatch Logs. After you've created a flow log, you can view and retrieve its data in Amazon CloudWatch Logs.



## VPC Flow Logs Levels

Flow logs can be created at 3 levels;

- VPC
- Subnet
- Network Interface Level



## Exam Tips

**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 can tag flow logs.
- 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;**

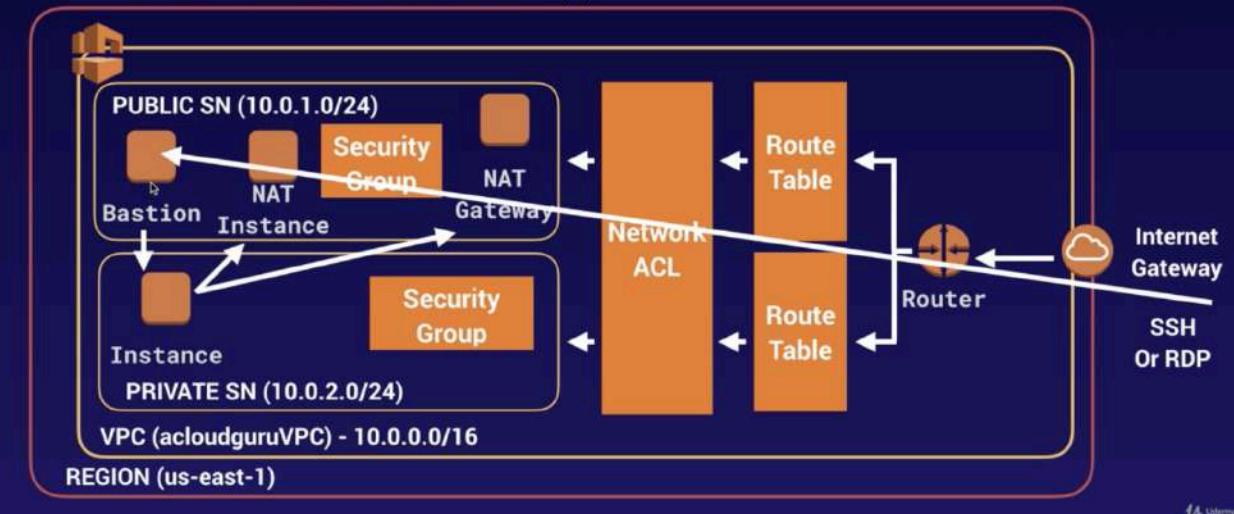
- Traffic generated by instances when they contact the Amazon DNS server. If you use your own DNS server, then all traffic to that DNS server is logged.
- Traffic generated by a Windows instance for Amazon Windows license activation.
- Traffic to and from 169.254.169.254 for instance metadata.
- DHCP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

**Bastion****A Bastion Host:**

A bastion host is a special purpose computer on a network specifically designed and configured to withstand attacks. The computer generally hosts a single application, for example a proxy server, and all other services are removed or limited to reduce the threat to the computer. It is hardened in this manner primarily due to its location and purpose, which is either on the outside of a firewall or in a demilitarized zone (DMZ) and usually involves access from untrusted networks or computers.



## VPC with Public &amp; Private Subnet(s)

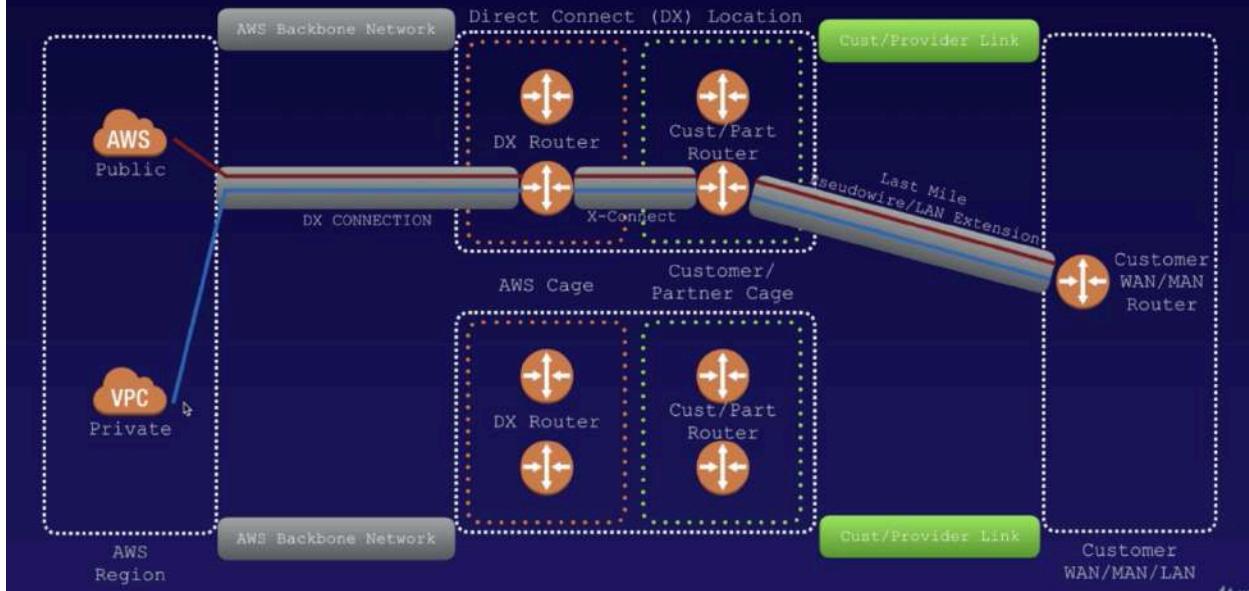
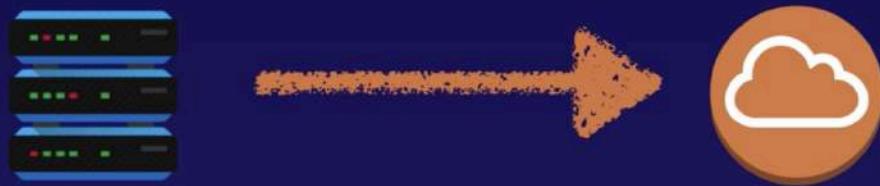
**Remember the following:**

- 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.

**Direct Connect**

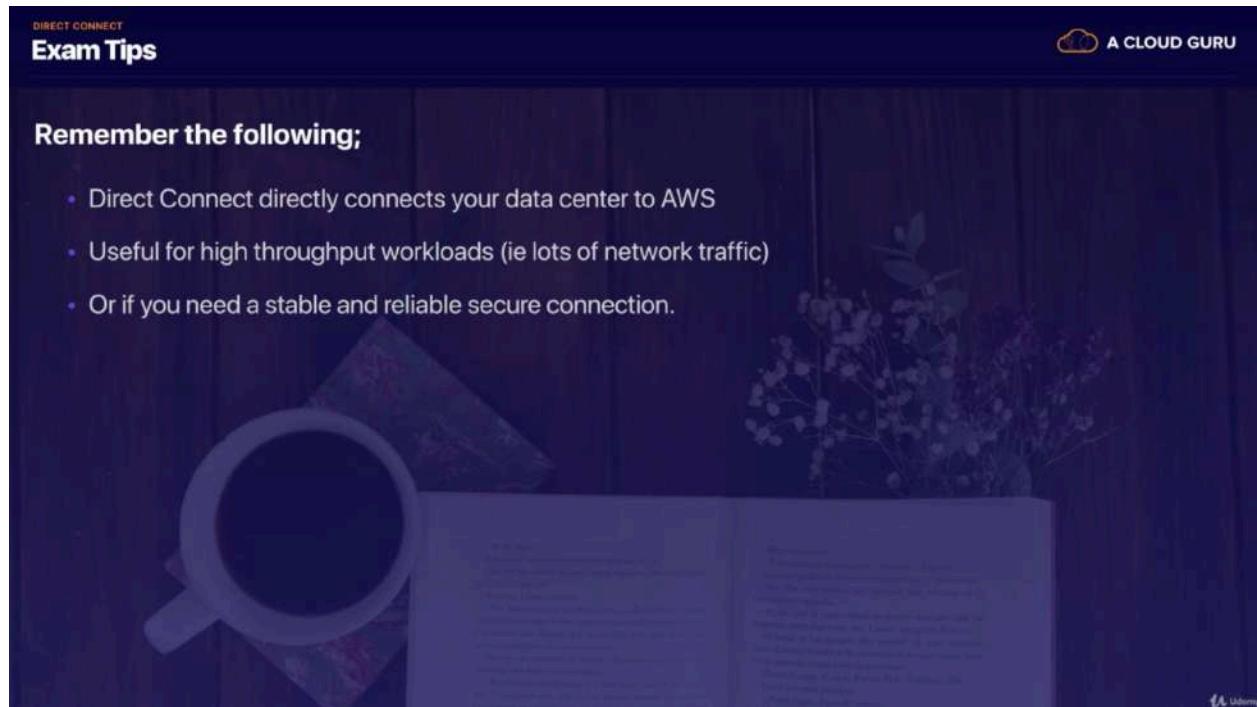
## Direct Connect

AWS Direct Connect is a cloud service solution that makes it easy to establish a dedicated network connection from your premises to AWS. Using AWS Direct Connect, you can establish private connectivity between AWS and your datacenter, office, or colocation environment, which in many cases can reduce your network costs, increase bandwidth throughput, and provide a more consistent network experience than Internet-based connections.



**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.

**Setting up Direct Connection**

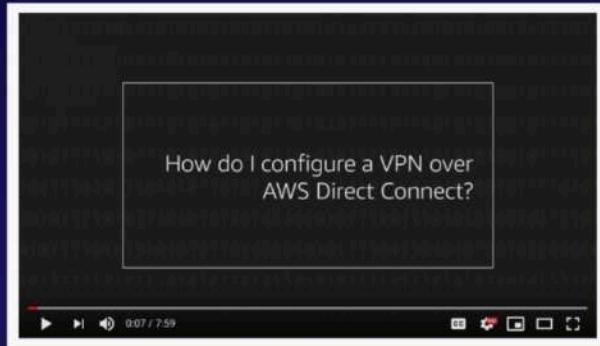
## Steps to setting up Direct Connect

- Create a virtual interface in the Direct Connect console. This is a **PUBLIC Virtual Interface**.
- Go to the VPC console and then to VPN connections. Create a Customer Gateway.
- Create a Virtual Private Gateway
- Attach the Virtual Private Gateway to the desired VPC.
- Select VPN Connections and create new VPN Connection.
- Select the Virtual Private Gateway and the Customer Gateway
- Once the VPN is available, set up the VPN on the customer gateway or firewall.



## Watch this AWS Youtube video

<https://www.youtube.com/watch?v=dhpTTT6V1So>



<https://www.youtube.com/watch?v=dhpTTT6V1So&feature=youtu.be>

## Remember the Steps to Creating a Direct Connect Connection.

- Create a virtual interface in the Direct Connect console. This is a PUBLIC Virtual Interface.
- Go to the VPC console and then to VPN connections. Create a Customer Gateway.
- Create a Virtual Private Gateway
- Attach the Virtual Private Gateway to the desired VPC.
- Select VPN Connections and create new VPN Connection.
- Select the Virtual Private Gateway and the Customer Gateway
- Once the VPN is available, setup the VPN on the customer gateway or firewall.

## Global Accelerator



AWS Global Accelerator is a service in which you create accelerators to improve availability and performance of your applications for local and global users.

Global Accelerator directs traffic to optimal endpoints over the AWS global network. This improves the availability and performance of your internet applications that are used by a global audience.



By default, Global Accelerator provides you with two static IP addresses that you associate with your accelerator.

Alternatively, you can bring your own.

## AWS Global Accelerator includes the following components

- Static IP addresses
- Accelerator
- DNS Name
- Network Zone
- Listener
- Endpoint Group
- Endpoint



### Static IP addresses

By default, Global Accelerator provides you with two static IP addresses that you associate with your accelerator.

Or, you can bring your own.

1 1.2.3.4

2 5.6.7.8

## Accelerator

An accelerator directs traffic to optimal endpoints over the AWS global network to improve the availability and performance of your internet applications.

Each accelerator includes one or more listeners.



## DNS Name

Global Accelerator assigns each accelerator a default Domain Name System (DNS) name - similar to **a1234567890abcdef.awsglobalaccelerator.com** - that points to the static IP addresses that Global Accelerator assigns to you.

Depending on the use case, you can use your accelerator's static IP addresses or DNS name to route traffic to your accelerator, or set up DNS records to route traffic using your own custom domain name.



## Network zone

A network zone services the static IP addresses for your accelerator from a unique IP subnet. Similar to an AWS Availability Zone, a network zone is an isolated unit with its own set of physical infrastructure.

When you configure an accelerator, by default, Global Accelerator allocates two IPv4 addresses for it. If one IP address from a network zone becomes unavailable due to IP address blocking by certain client networks, or network disruptions, client applications can retry on the healthy static IP address from the other isolated network zone.



## Network zone

A network zone services the static IP addresses for your accelerator from a unique IP subnet. Similar to an AWS Availability Zone, a network zone is an isolated unit with its own set of physical infrastructure.

When you configure an accelerator, by default, Global Accelerator allocates two IPv4 addresses for it. If one IP address from a network zone becomes unavailable due to IP address blocking by certain client networks, or network disruptions, client applications can retry on the healthy static IP address from the other isolated network zone.



## Listener

A listener processes inbound connections from clients to Global Accelerator, based on the port (or port range) and protocol that you configure. Global Accelerator supports both TCP and UDP protocols.

Each listener has one or more endpoint groups associated with it, and traffic is forwarded to endpoints in one of the groups.

You associate endpoint groups with listeners by specifying the Regions that you want to distribute traffic to. Traffic is distributed to optimal endpoints within the endpoint groups associated with a listener.



## Endpoint group

Each endpoint group is associated with a specific AWS Region.

Endpoint groups include one or more endpoints in the Region.

You can increase or reduce the percentage of traffic that would be otherwise directed to an endpoint group by adjusting a setting called a traffic dial.

The traffic dial lets you easily do performance testing or blue/green deployment testing for new releases across different AWS Regions, for example.



## Endpoint

Endpoints can be Network Load Balancers, Application Load Balancers, EC2 instances, or Elastic IP addresses.

An Application Load Balancer endpoint can be an internet-facing or internal. Traffic is routed to endpoints based on configuration options that you choose, such as endpoint weights.

For each endpoint, you can configure weights, which are numbers that you can use to specify the proportion of traffic to route to each one. This can be useful, for example, to do performance testing within a Region.



Web1

Web2

Web3

## Know what a Global Accelerator is and where you would use it.

- AWS Global Accelerator is a service in which you create accelerators to improve availability and performance of your applications for local and global users.
- You are assigned two static IP addresses (or alternatively you can bring your own).
- You can control traffic using traffic dials. This is done within the endpoint group.

## VPC Endpoint

## What Is A VPC Endpoint?



### 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.



## Exam Tips



**There are two types of VPC endpoints:**

- Interface Endpoints
- Gateway Endpoints

## Exam Tips

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

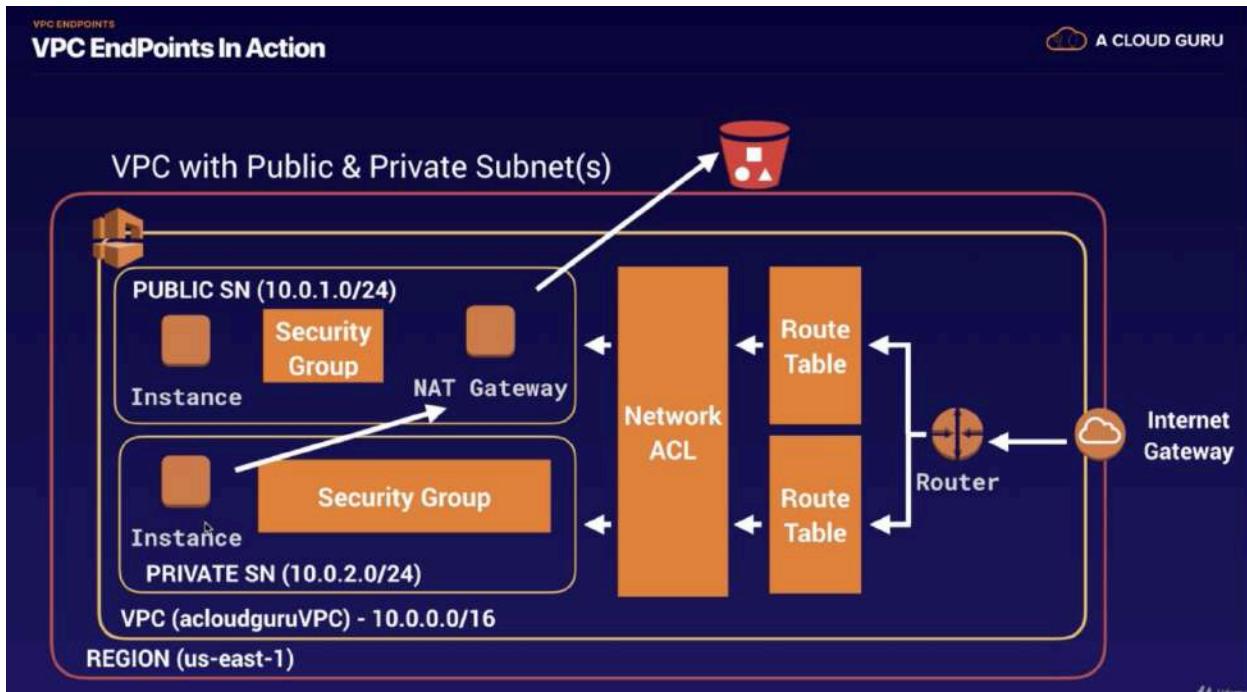
A. Udemy

## Gateway Endpoints

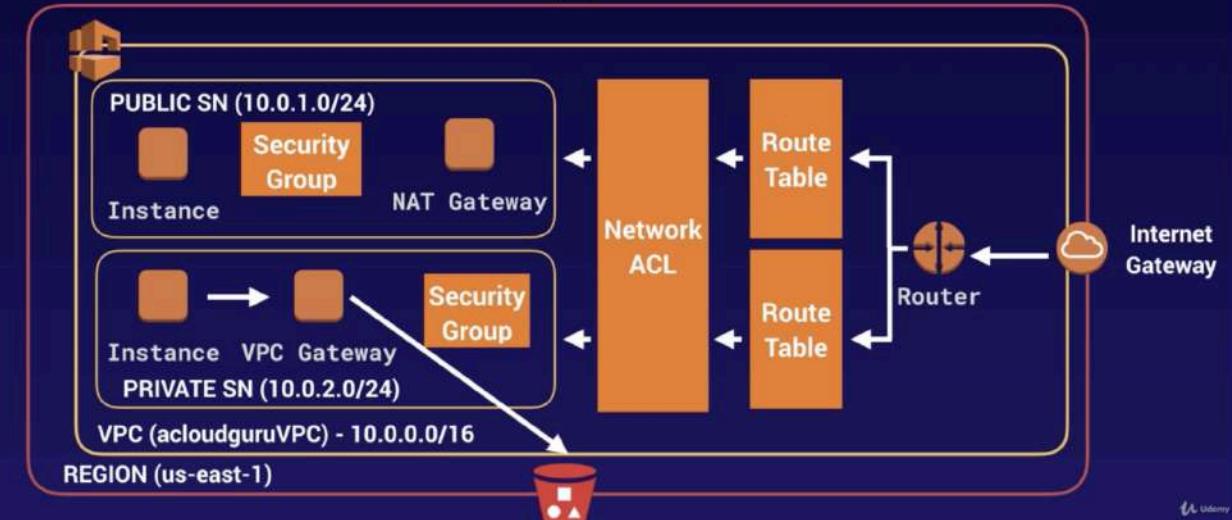
### Currently Gateway Endpoints Support

- Amazon S3
- DynamoDB

A. Udemy



## VPC with Public &amp; Private Subnet(s)

**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:**

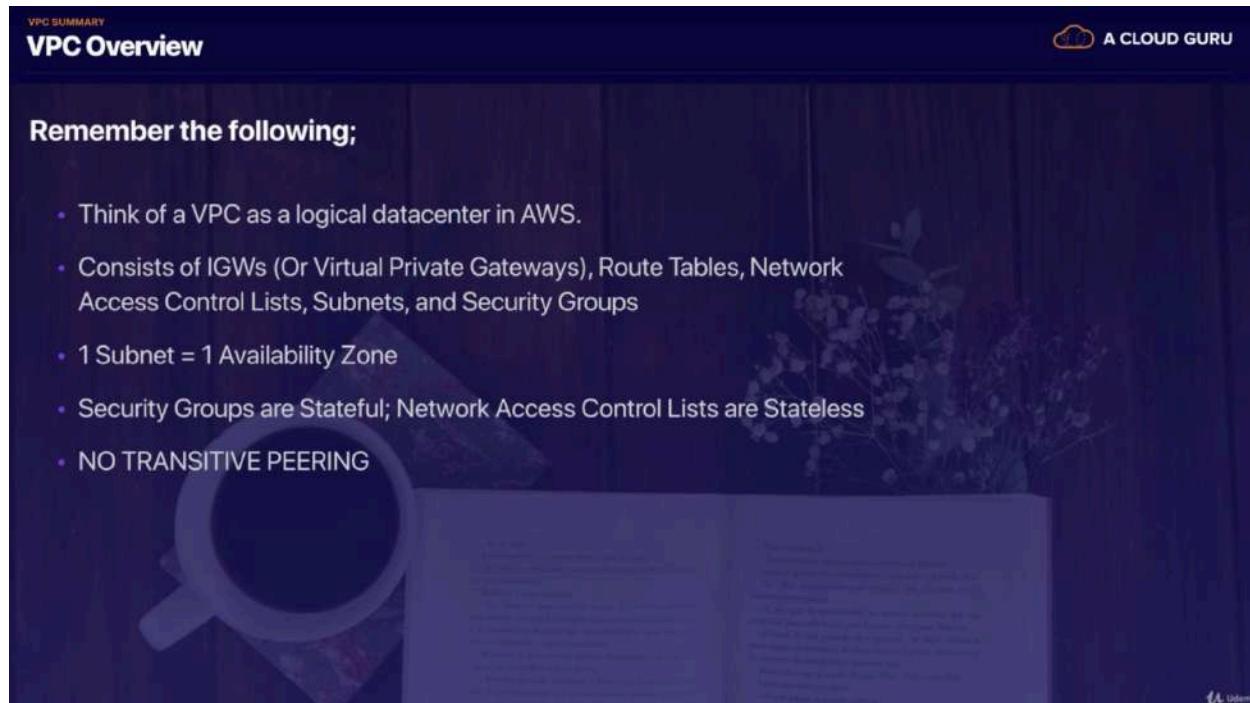
- Amazon S3
- DynamoDB

**Summary of VPC**

VPC SUMMARY

## VPC Overview

A CLOUD GURU



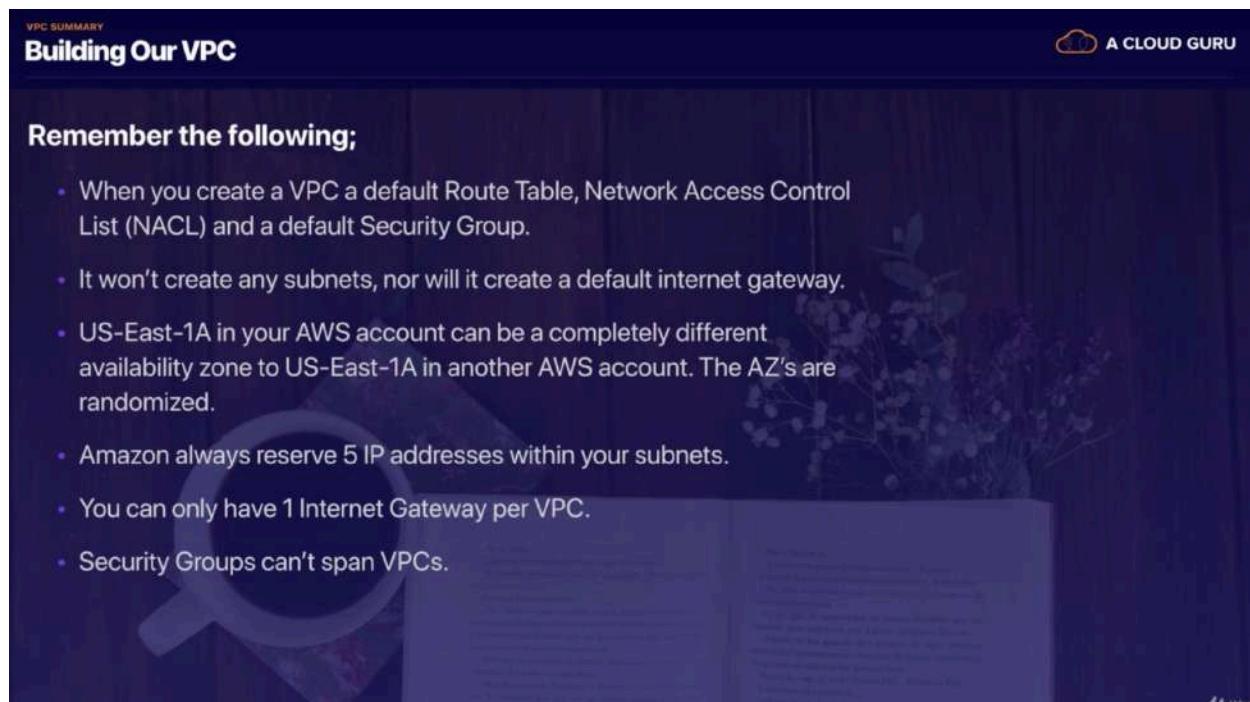
**Remember the following;**

- 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

VPC SUMMARY

## Building Our VPC

A CLOUD GURU



**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.

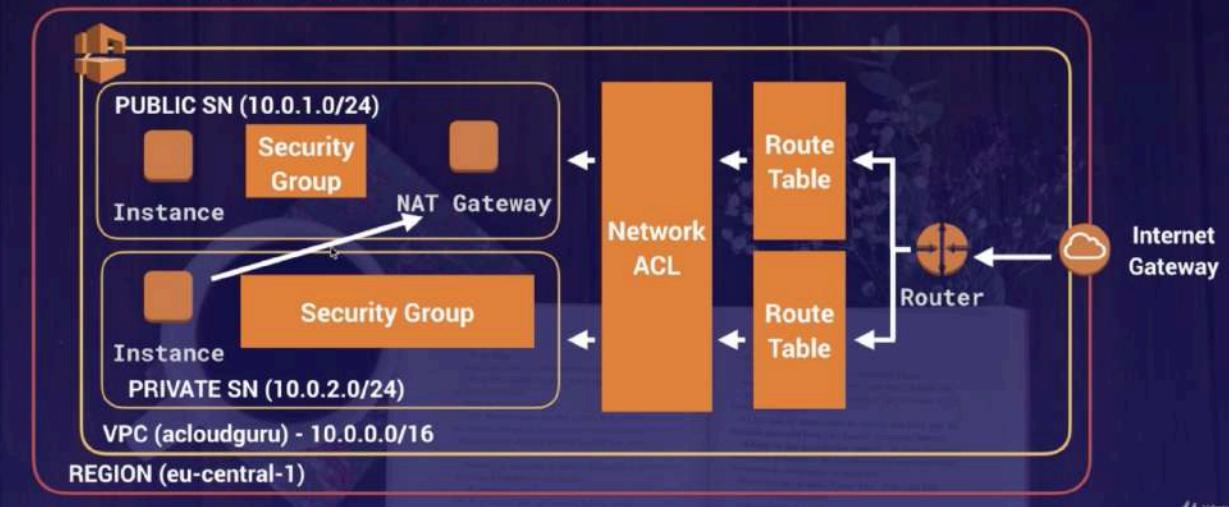
## NAT Instances vs NAT Gateways

### 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 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.

**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;**

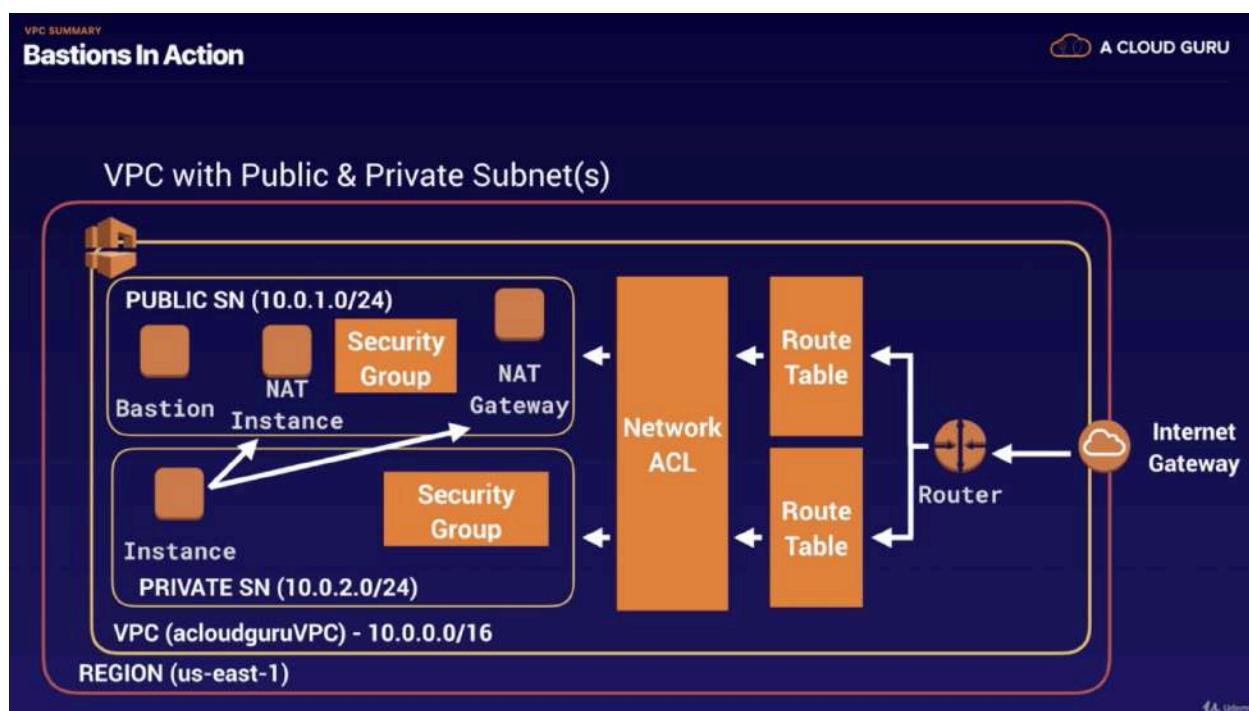
- You cannot enable flow logs for VPCs that are peered with your VPC unless the peer VPC is in your account.
- You can tag flow logs.
- After you've created a flow log, you cannot change its configuration; for example, you can't associate a different IAM role with the flow log.

VPC SUMMARY

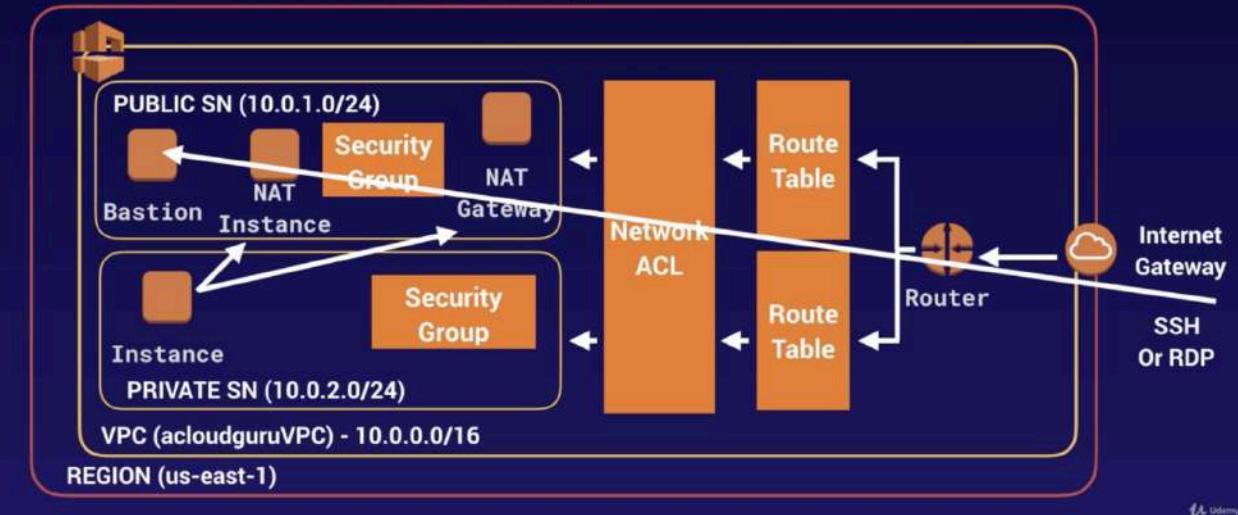
## VPC Flow Logs

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.



## VPC with Public &amp; Private Subnet(s)



## 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.

VPC SUMMARY

## Direct Connect

A CLOUD GURU

**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.



SETTING UP DIRECT CONNECT

## Direct Connect - Exam Tips

A CLOUD GURU

**Remember the Steps to Creating a Direct Connect Connection.**

- Create a virtual interface in the Direct Connect console. This is a PUBLIC Virtual Interface.
- Go to the VPC console and then to VPN connections. Create a Customer Gateway.
- Create a Virtual Private Gateway
- Attach the Virtual Private Gateway to the desired VPC.
- Select VPN Connections and create new VPN Connection.
- Select the Virtual Private Gateway and the Customer Gateway
- Once the VPN is available, setup the VPN on the customer gateway or firewall.

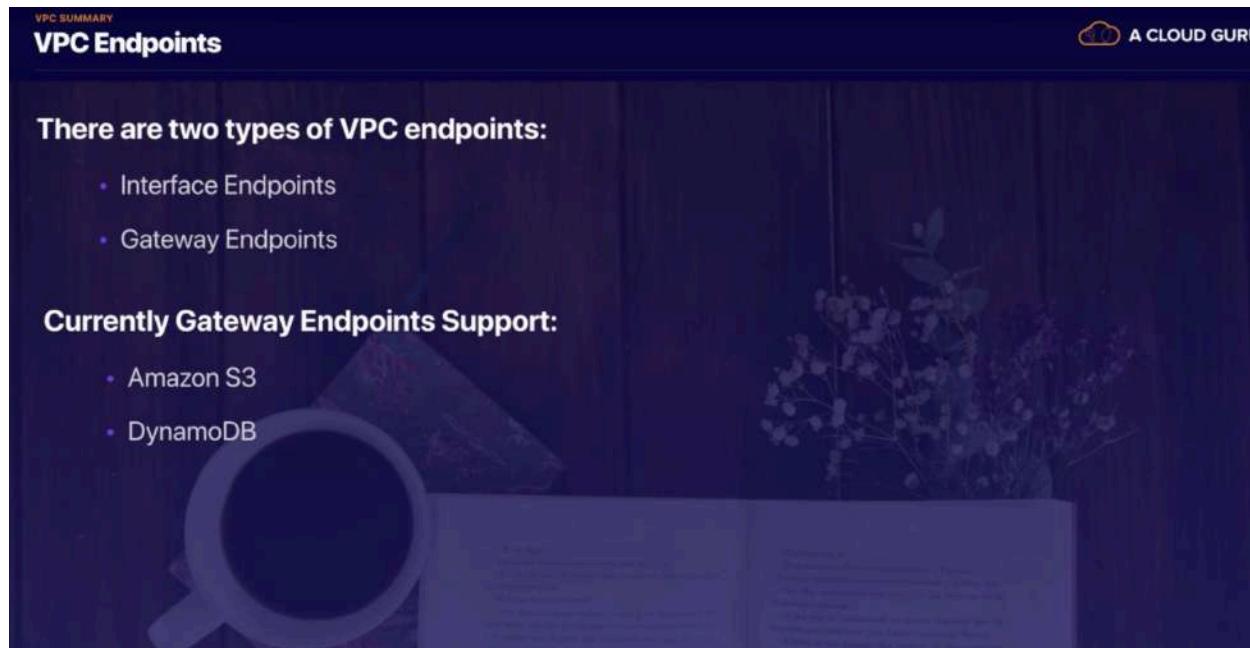
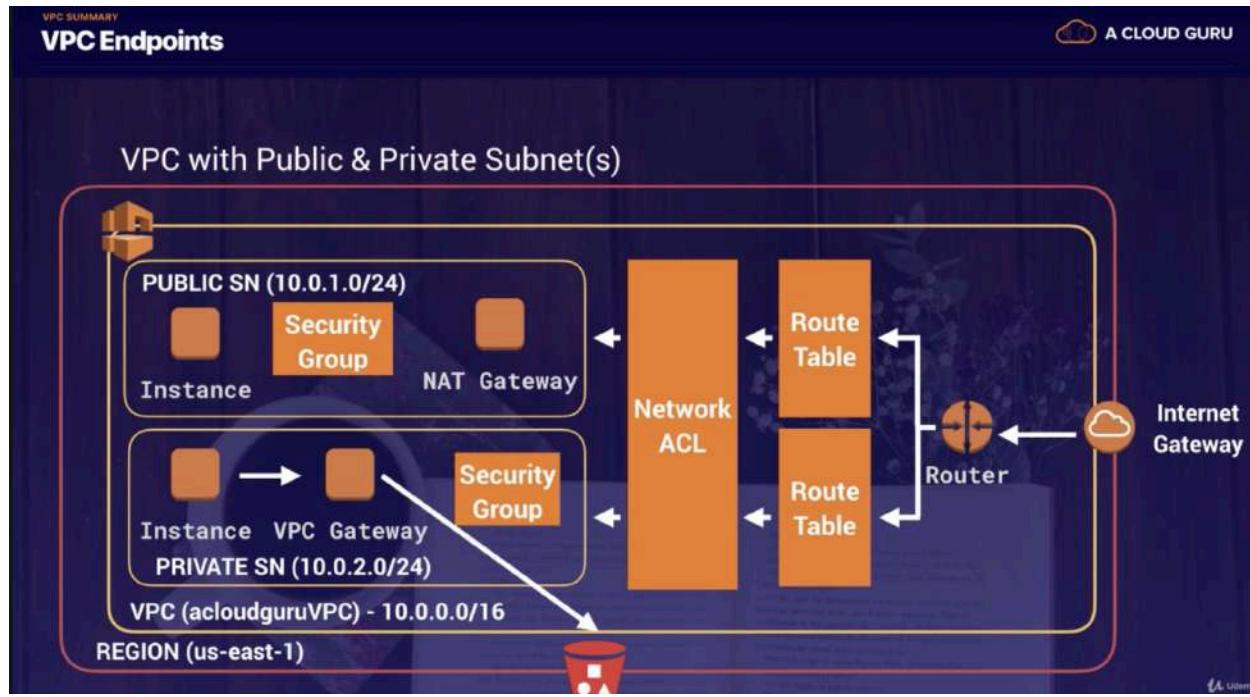
## Know what a Global Accelerator is and where you would use it.

- AWS Global Accelerator is a service in which you create accelerators to improve availability and performance of your applications for local and global users.
- You are assigned two static IP addresses (or alternatively you can bring your own).
- You can control traffic using traffic dials. This is done within the endpoint group.
- You can control weighting to individual end points using weights.

### 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.



## Databases on AWS

**Relational databases are what most of us are all used to. They have been around since the 70's.**  
**Think of a traditional spreadsheet:**

- Database
- Tables
- Row
- Fields (Columns)



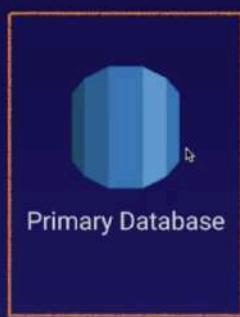
## Relational databases on AWS;

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

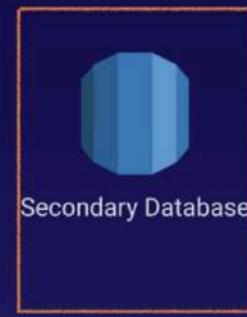


## RDS has two key features;

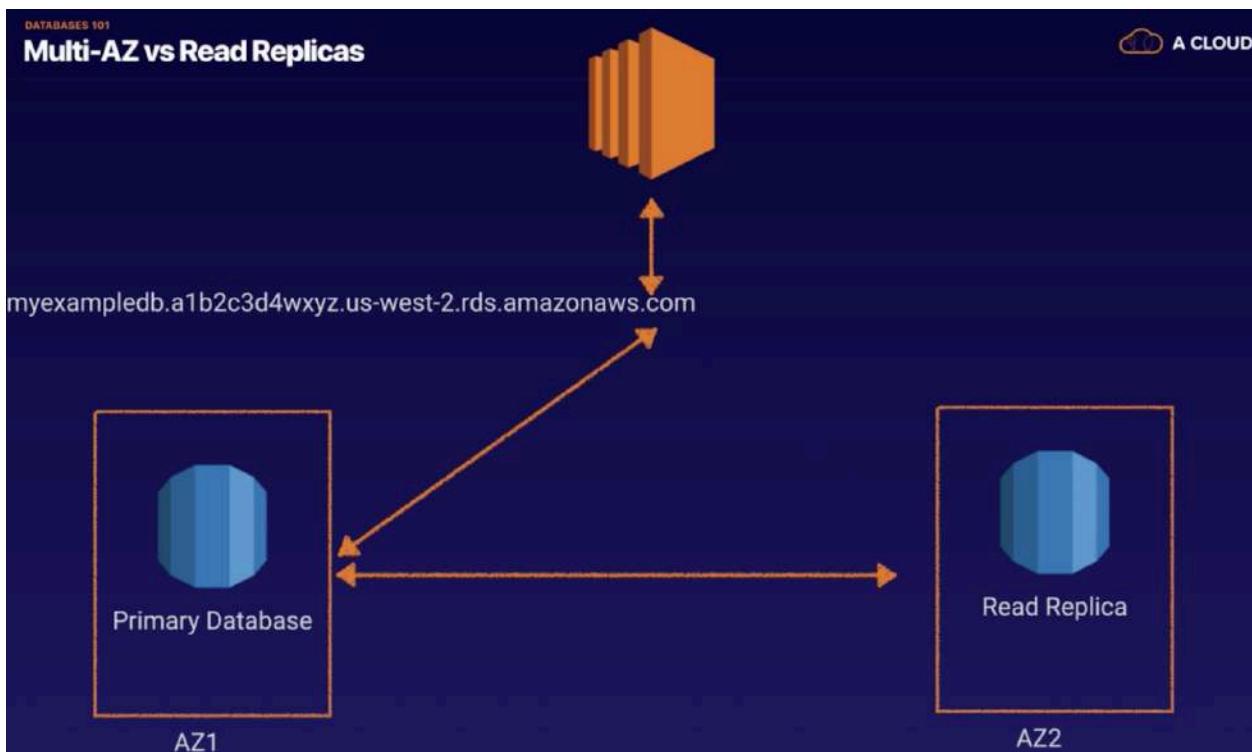
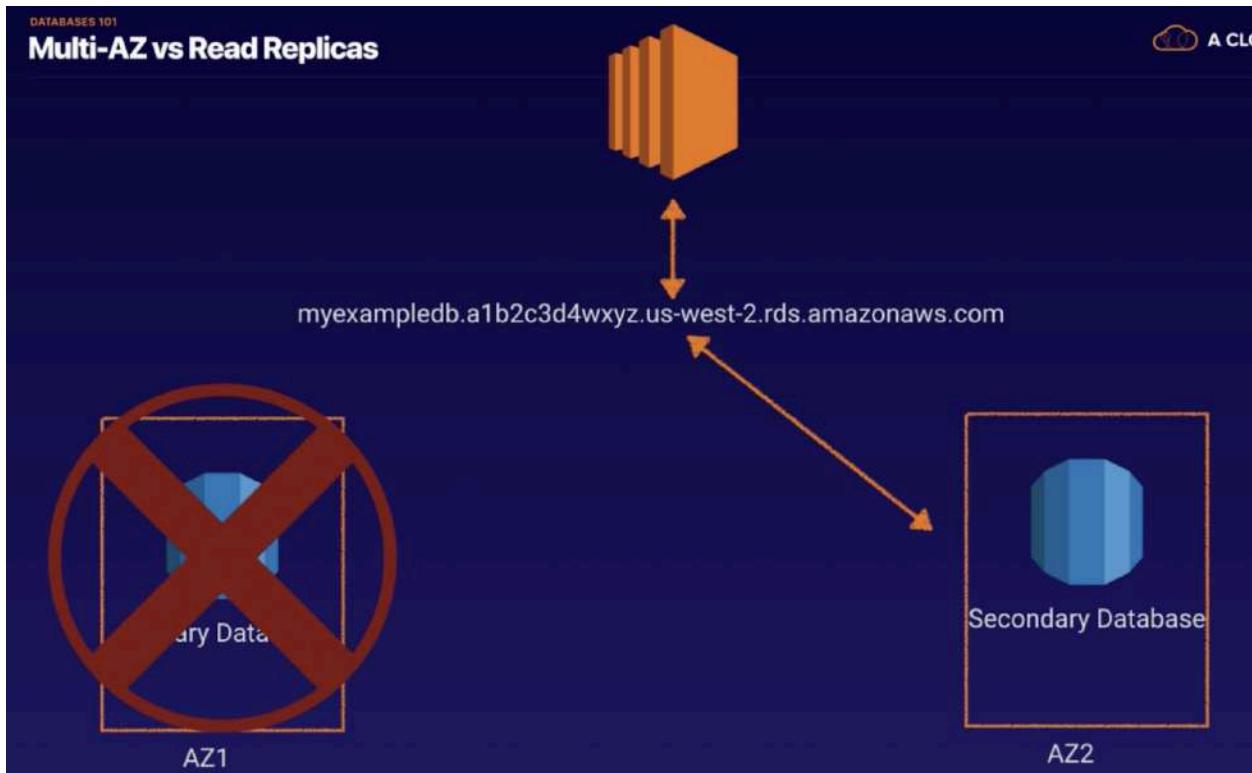
- Multi-AZ - For Disaster Recovery
- Read Replicas - For Performance

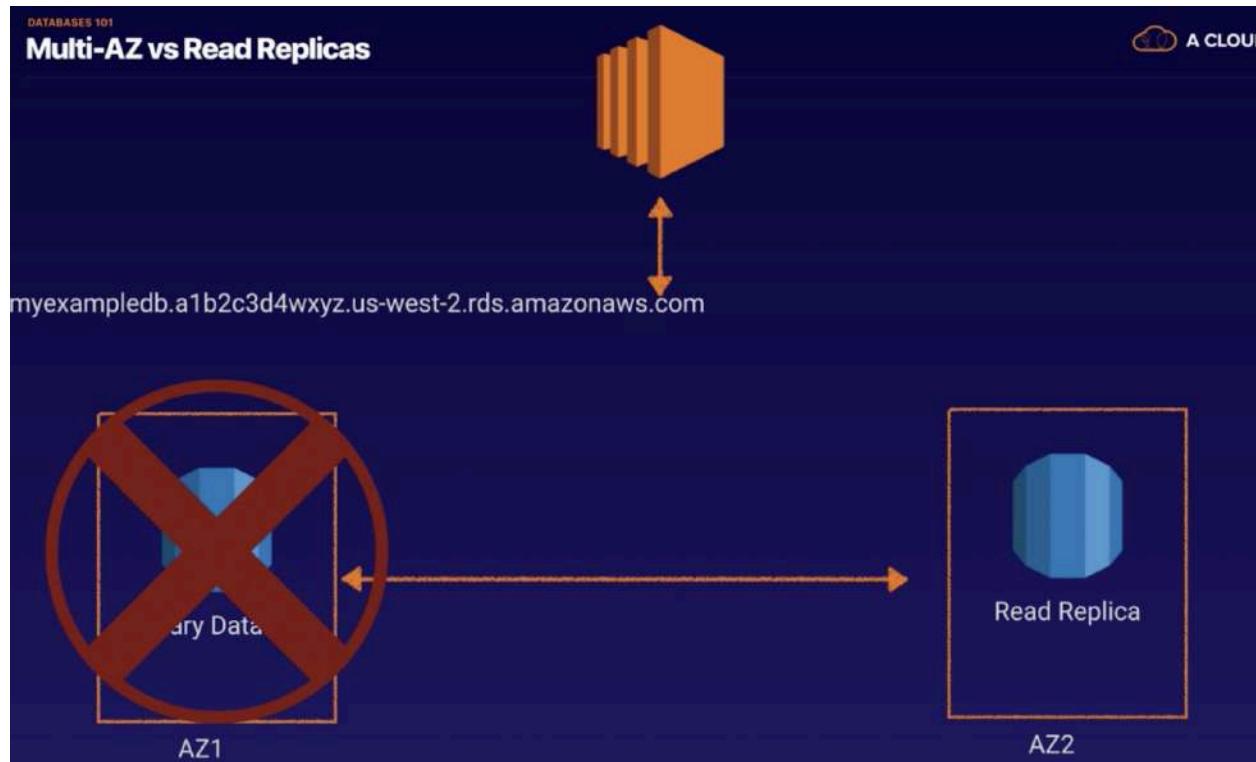


AZ1

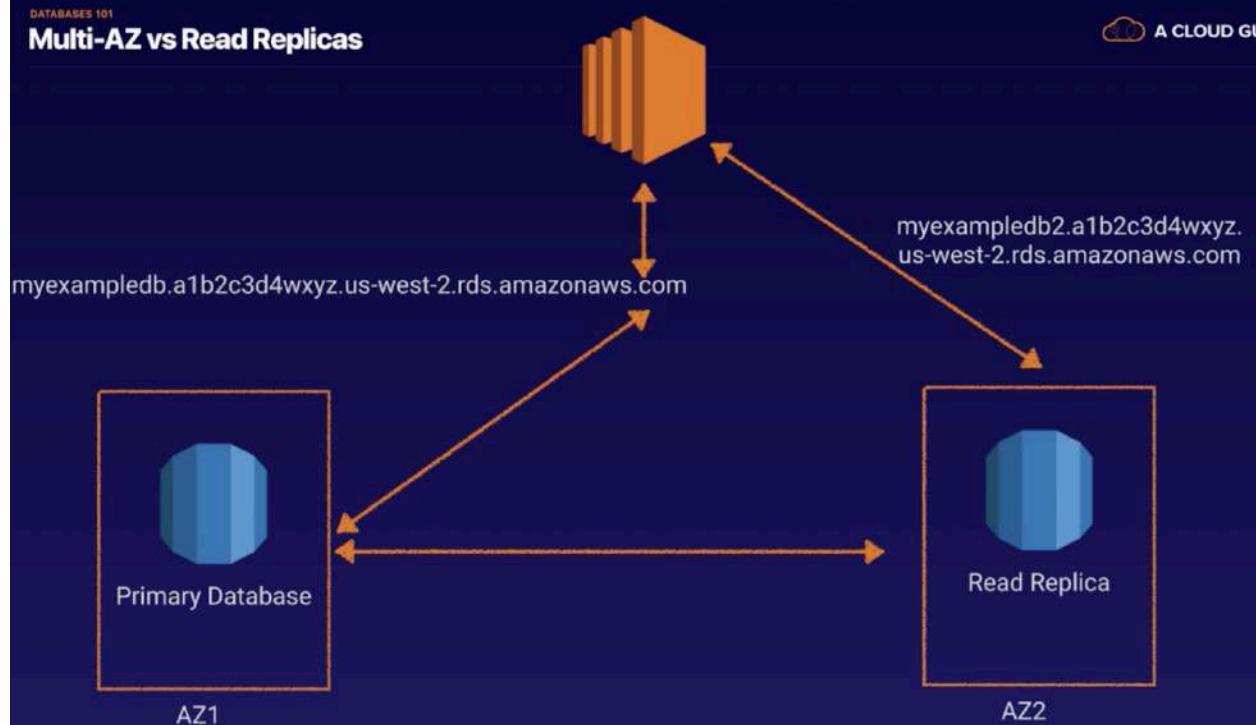


AZ2



**Multi-AZ vs Read Replicas**

## Multi-AZ vs Read Replicas



## Non Relational Databases are as follows;

- Collection = Table
- Document = Row
- Key Value Pairs = Fields



```
{  
  "_id" : "51262c865ca358946be09d77",  
  "firstname" : "John",  
  "surname" : "Smith",  
  "Age" : "23",  
  "address" : [  
    {"street" : "21 Jump Street",  
     "suburb" : "Richmond"}  
  ]  
}
```

**Used for business intelligence. Tools like Cognos, Jaspersoft, SQL Server Reporting Services, Oracle Hyperion, SAP NetWeaver.**

**Used to pull in very large and complex data sets. Usually used by management to do queries on data (such as current performance vs targets etc)**

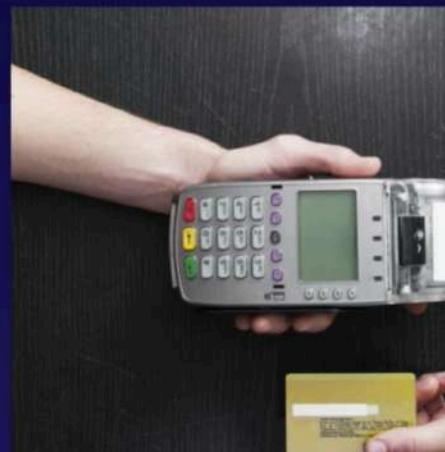


**Online Transaction Processing (OLTP) differs from OLAP Online Analytics Processing (OLAP) in terms of the types of queries you will run.**

**OLTP Example:**

**Order number 2120121**

**Pulls up a row of data such as Name, Date, Address to Deliver to, Delivery Status etc.**



**OLAP transaction Example:**  
**Net Profit for EMEA and Pacific for the Digital Radio Product.**  
**Pulls in large numbers of records**

**Sum of Radios Sold in EMEA**

**Sum of Radios Sold in Pacific**

**Unit Cost of Radio in each region**

**Sales price of each radio**

**Sales price - unit cost.**



**Data Warehousing databases use different type of architecture both from a database perspective and infrastructure layer.**



**Data Warehousing databases use different type of architecture both from a database perspective and infrastructure layer.**

**Amazon's Data  
Warehouse  
Solution Is Called  
Redshift**



**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.**

**ElastiCache supports two open-source in-memory caching engines:**

**ElastiCache supports two open-source in-memory caching engines:**

- Memcached
- Redis



## AWS Database Types - Exam Tips

### RDS (OLTP)

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

### DynamoDB (No SQL)

### Red Shift OLAP

## AWS Database Types - Exam Tips

**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



## 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. This allows you to do a point in time recovery down to a second, within the retention period.



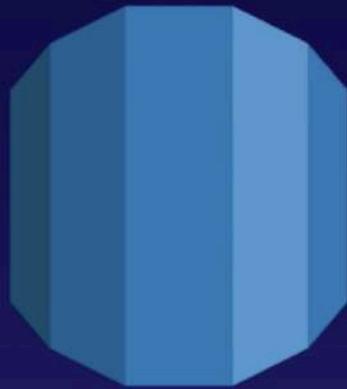
## Automated Backups

**Automated Backups** are enabled by default. The backup data is stored in S3 and you get free storage space equal to the size of your database. So if you have an RDS instance of 10Gb, you will get 10Gb worth of storage.

Backups are taken within a defined window. During the backup window, storage I/O may be suspended while your data is being backed up and you may experience elevated latency.



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



original.eu-west-1.rds.amazonaws.com



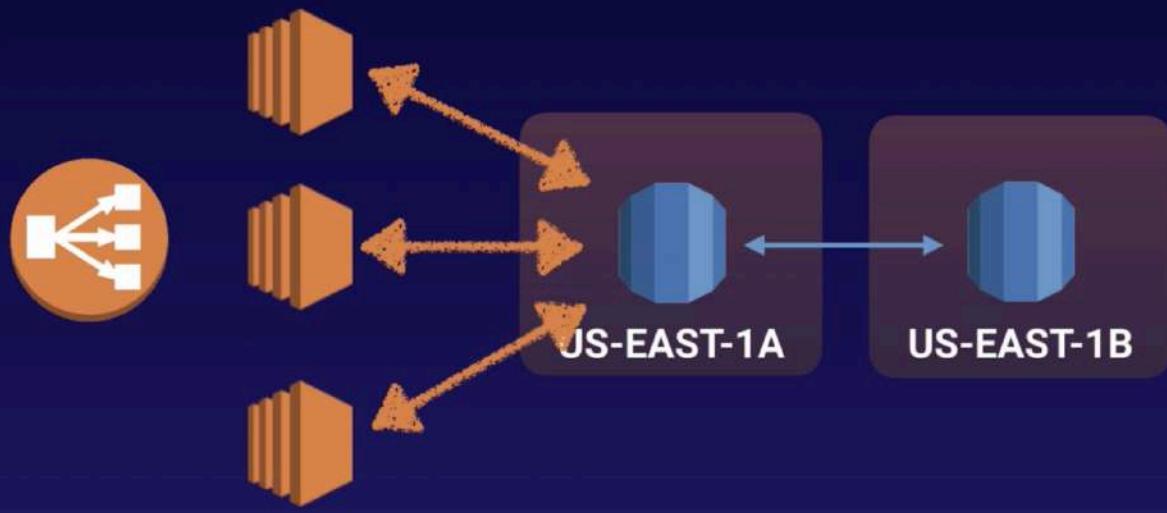
restored.eu-west-1.rds.amazonaws.com

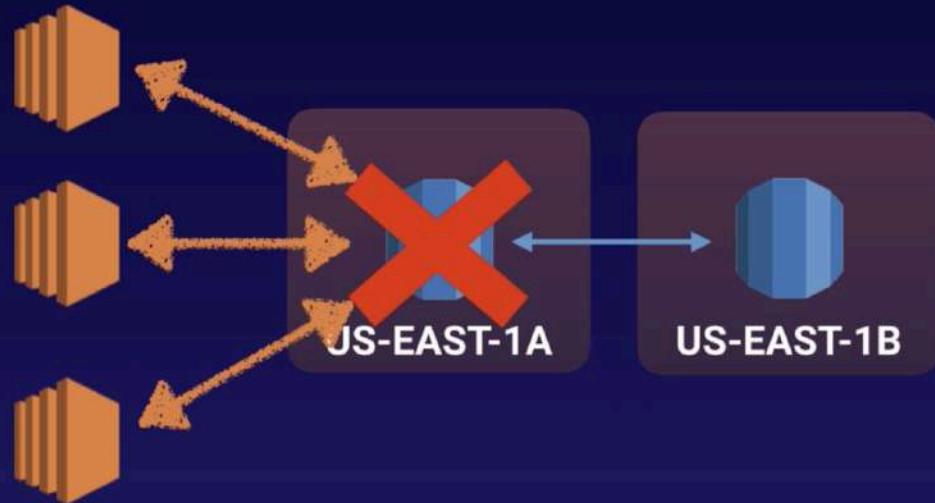
## Encryption At Rest

**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.**



## What Is Multi-AZ





**Multi-AZ allows you to have an exact copy of your production database in another Availability Zone. 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 Availability Zone failure, Amazon RDS will automatically failover to the standby so that database operations can resume quickly without administrative intervention.**

## Multi-AZ is for Disaster Recovery only



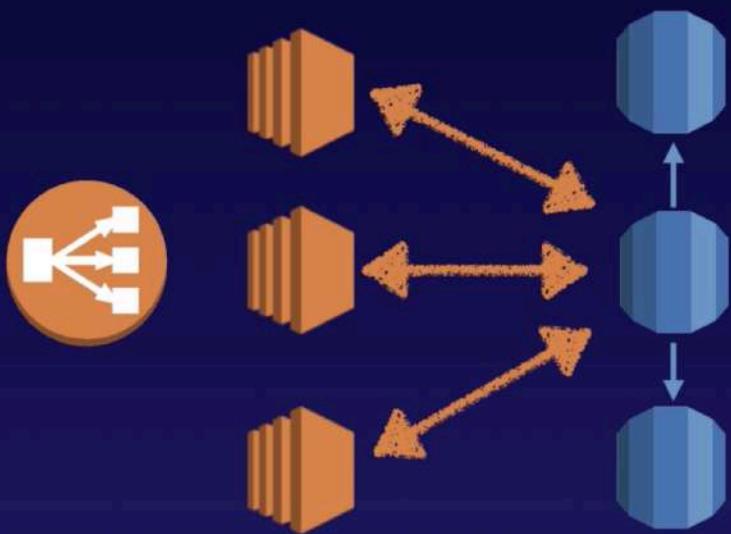
**It is not primarily used for improving performance. For performance improvement, you need Read Replicas.**

## Multi-AZ is available for the following databases

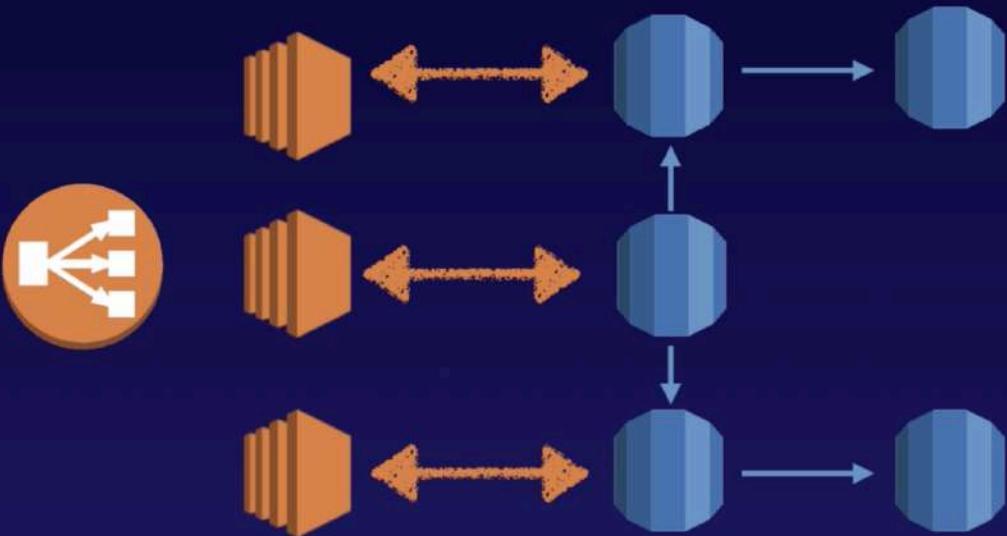
- SQL Server
- Oracle
- MySQL Server
- PostgreSQL
- MariaDB



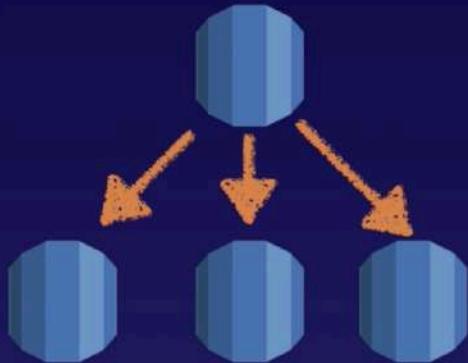
## What Is A Read Replica



## What Is A 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.**



**Read Replicas are available for the following databases**

- MySQL Server
- PostgreSQL
- MariaDB
- Oracle
- Aurora



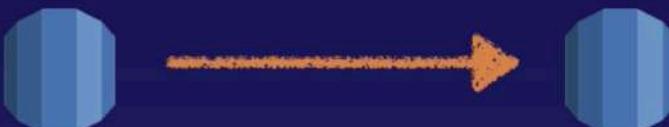
RDS - BACK UPS, MULTI-AZ & READ REPLICAS

Back Ups With RDS

A CLOUD GURU

## Things to know about Read Replicas;

- Used for scaling, not for DR!
- Must have automatic backups turned on in order to deploy a read replica.
- You can have up to 5 read replica copies of any database.
- You can have read replicas of read replicas (but watch out for latency.)



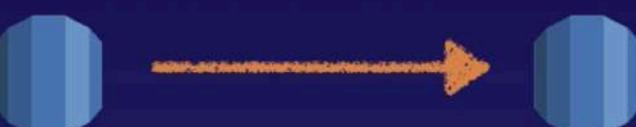
RDS - BACK UPS, MULTI-AZ & READ REPLICAS

Back Ups With RDS

A CLOUD GURU

## Things to know about Read Replicas;

- Each read replica will have its own DNS end point.
- You can have read replicas that have Multi-AZ.
- You can create read replicas of Multi-AZ source databases.
- Read replicas can be promoted to be their own databases. This breaks the replication.
- You can have a read replica in a second region.



## Exam Tips

**There are two different types of Backups for RDS:**

- **Automated Backups**
- **Database Snapshots**

## Exam Tips



A CL

## Read Replicas

- **Can be Multi-AZ.**
- **Used to increase performance.**
- **Must have backups turned on.**
- **Can be in different regions.**
- **Can be MySQL, PostgreSQL, MariaDB, Oracle, Aurora.**
- **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.

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

## What Is DynamoDB?

Amazon DynamoDB is a fast and flexible NoSQL database service for all applications that need consistent, single-digit millisecond latency at any scale. It is a fully managed database and supports both document and key-value data models. Its flexible data model and reliable performance make it a great fit for mobile, web, gaming, ad-tech, IoT, and many other applications.



## DynamoDB Reads

### Eventual Consistent Reads

- Consistency across all copies of data is usually reached within a second. Repeating a read after a short time should return the updated data. (Best Read Performance)



DYNAMODB  
**DynamoDB Reads**

A CLOUD

## Strongly Consistent Reads

- A strongly consistent read returns a result that reflects all writes that received a successful response prior to the read.

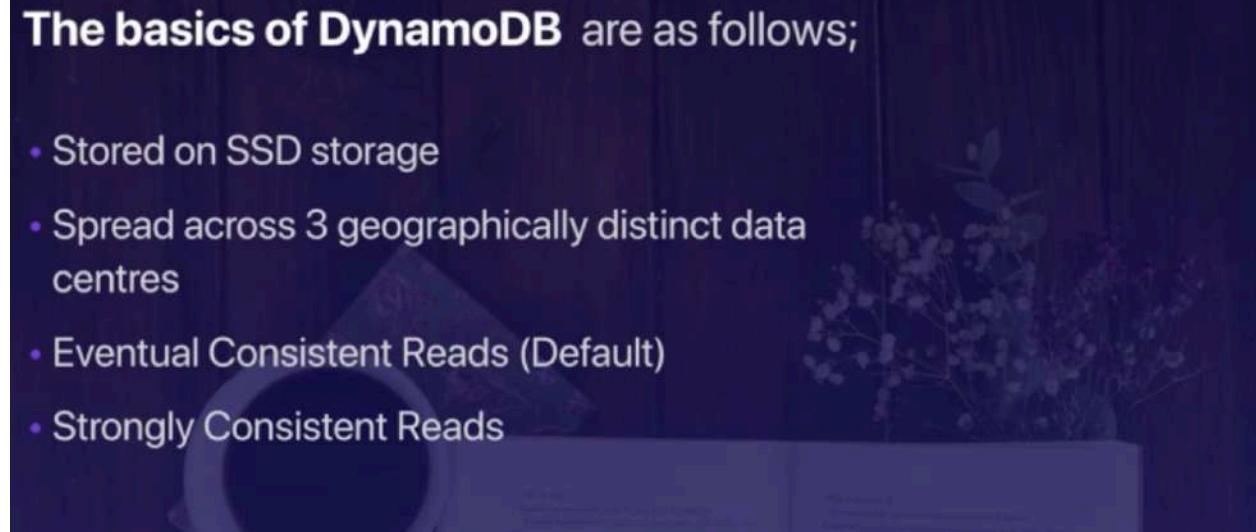


DYNAMODB  
**DynamoDB Exam Tips**

A CLOUD

### 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**

**Amazon Redshift is a fast and powerful, fully managed, petabyte-scale data warehouse service in the cloud. Customers can start small for just \$0.25 per hour with no commitments or upfront costs and scale to a petabyte or more for \$1,000 per terabyte per year, less than a tenth of most other data warehousing solutions.**



**OLAP transaction Example:  
Net Profit for EMEA and Pacific for the Digital Radio Product.  
Pulls in large numbers of records**

**Sum of Radios Sold in EMEA  
Sum of Radios Sold in Pacific  
Unit Cost of Radio in each region  
Sales price of each radio  
Sales price - unit cost.**



**Data Warehousing databases use different type of architecture both from a database perspective and infrastructure layer.**

**Amazon's Data  
Warehouse  
Solution Is Called  
Redshift**



## Redshift can be configured as follows

- 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.

### Advanced Compression:

Columnar data stores can be compressed much more than row-based data stores because similar data is stored sequentially on disk. Amazon Redshift employs multiple compression techniques and can often achieve significant compression relative to traditional relational data stores. In addition, Amazon Redshift doesn't require indexes or materialized views, and so uses less space than traditional relational database systems. When loading data into an empty table, Amazon Redshift automatically samples your data and selects the most appropriate compression scheme.



## **Massively Parallel Processing (MPP):**

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



## 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.



## Redshift is priced as follows;

- Compute Node Hours (total number of hours you run across all your compute nodes for the billing period. You are billed for 1 unit per node per hour, so a 3-node data warehouse cluster running persistently for an entire month would incur 2,160 instance hours. You will not be charged for leader node hours; only compute nodes will incur charges.)
- Backup
- Data transfer (only within a VPC, not outside it)

## Security Considerations:

- Encrypted in transit using SSL
- Encrypted at rest using AES-256 encryption
- By default RedShift takes care of key management.
  - Manage your own keys through HSM
  - AWS Key Management Service



## Redshift Availability

### Redshift Availability:

- Currently only available in 1 AZ
- Can restore snapshots to new AZs in the event of an outage.



## Exam Tips

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

## Exam Tips

### 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.

### Aurora

## What is Aurora?

**Amazon Aurora is a MySQL and PostgreSQL-compatible relational database engine that combines the speed and availability of high-end commercial databases with the simplicity and cost-effectiveness of open source databases.**

**Amazon Aurora provides up to five times better performance than MySQL and three times better than PostgreSQL databases at a much lower price point, whilst delivering similar performance and availability.**





## Things to know about Aurora

- 1 Start with 10GB, Scales in 10GB increments to 64TB (Storage Autoscaling)
- 2 Compute resources can scale up to 32vCPUs and 244GB of Memory.
- 3 2 copies of your data is contained in each availability zone, with minimum of 3 availability zones. 6 copies of your data.



## The Basics of Aurora

### Scaling Aurora

- Aurora is designed to transparently handle the loss of up to two copies of data without affecting database write availability and up to three copies without affecting read availability.
- Aurora storage is also self-healing. Data blocks and disks are continuously scanned for errors and repaired automatically.



## The Basics of Aurora

### Three Types of Aurora Replicas are available:

- Aurora Replicas (currently 15)
- MySQL Read Replicas (currently 5)
- PostgreSQL (currently 1)



## The Basics of Aurora



Feature	Amazon Aurora Replicas	MySQL Replicas
Number of replicas	Up to 15	Up to 5
Replication type	Asynchronous (milliseconds)	Asynchronous (seconds)
Performance impact on primary	Low	High
Replica location	In-region	Cross-region
Act as failover target	Yes (no data loss)	Yes (potentially minutes of data loss)
Automated failover	Yes	No
Support for user-defined replication delay	No	Yes
Support for different data or schema vs. primary	No	Yes

## BACKUPS WITH AURORA

- Automated backups are always enabled on Amazon Aurora DB Instances. Backups do not impact database performance.
- You can also take snapshots with Aurora. This also does not impact on performance.
- You can share Aurora Snapshots with other AWS accounts.



**Amazon Aurora Serverless is an on-demand, autoscaling configuration for the MySQL-compatible and PostgreSQL-compatible editions of Amazon Aurora. An Aurora Serverless DB cluster automatically starts up, shuts down, and scales capacity up or down based on your application's needs.**

**Aurora Serverless provides a relatively simple, cost-effective option for infrequent, intermittent, or unpredictable workloads.**



AURORA

## Aurora Exam Tips



- 2 copies of your data are 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.
- 3 types of replicas available. Aurora Replicas, MySQL replicas & PostgreSQL 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 Aurora Serverless if you want a simple, cost-effective option for infrequent, intermittent, or unpredictable workloads.

## Elasticache

ELASTICACHE  
What Is 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.**

ELASTICACHE  
What Is ElastiCache?



**ElastiCache supports two open-source in-memory caching engines:**

- Memcached
- Redis



**redis**

Requirement	Memcached	Redis
Simple Cache to offload DB	Yes	Yes
Ability to scale horizontally	Yes	Yes
Multi-threaded performance	Yes	No
Advanced data types	No	Yes
Ranking/Sorting data sets	No	Yes
Pub/Sub capabilities	No	Yes
Persistence	No	Yes
Multi-AZ	No	Yes
Backup & Restore Capabilities	No	Yes

- Use Elasticache to increase database and web application performance.
- Redis is Multi-AZ
- You can do back ups and restores of Redis

## Database Summary

## AWS Database Types - Exam Tips

### RDS (OLTP)

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

### DynamoDB (No SQL)

### Red Shift OLAP

## AWS Database Types - Exam Tips

### 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