

module 2 test

Module 2 Intro

- ▼ What are the three advantages of EC2?
 1. highly flexible
 2. cost effective
 3. quick
- ▼ What does AWS alr take care of so you can use EC2s?
 1. datacenters
 2. servers
 3. alr online to use
 4. massive amt of compute capacity
- ▼ What is a hypervisor?

It's running on the host machine and is responsible for sharing instances (or VMs) on the machine.
- ▼ What is multitenancy?

Sharing underlying hardware between VMs. This is done so that EC2 instances are not aware of one another.
- ▼ What are the three steps to EC2?
 1. launch
 2. connect
 3. use
- ▼ What basic configs options are included in the launch step?
 1. Operating System
 2. Application Servers
 3. Applications

4. Instance Types

5. Specific Hardware Configs of your Instances

Module 2 Part 2

▼ What are the different families that instance types can fall under?

1. general purpose
2. compute optimized
3. memory optimized
4. accelerated computing
5. storage optimized

▼ Describe the general purpose family:

Balanced resources; diverse workloads;

▼ What can you use general purpose family instances for?

Web servers or code repos

▼ Describe the compute optimized family:

Provides more intensive instances

▼ What can you use compute optimized family instances for?

Gaming servers; high performance computing; scientific modeling

▼ Describe the memory optimized family:

Faster performance for workloads

▼ What can you use memory optimized family instances for?

Used to process large datasets in memory

▼ Describe the accelerated computing family:

Better for calculations, processing, data pattern matching, and accelerators

▼ What can you use storage optimized family instances for?

Locally stored data that has high reads and writes (OLTP systems)

▼ What does IOPS stand for?

Input output per second

▼ What are storage optimized family instances designed for?

Designed for delivering tens of thousands of low latency random IOPS to applications

Which Amazon EC2 instance type is suitable for data warehousing applications?

☐ Memory optimized

☒ Storage optimized

☐ General purpose

☐ Compute optimized

Which Amazon EC2 instance type balances compute, memory, and networking resources?

☐ Memory optimized

☐ Storage optimized

☒ General purpose

☐ Compute optimized

Which Amazon EC2 instance type is ideal for high-performance databases?

☒ Memory optimized

☐ Storage optimized

☐ General purpose

☐ Compute optimized

Which Amazon EC2 instance type offers high-performance processors?

☐ Memory optimized

☐ Storage optimized

☐ General purpose

☒ Compute optimized

Module 2 Part 3

▼ Explain on demand pricing:

Pay for duration and this is what you use when you just wanna get started and spin up servers to test out workloads or play around

▼ Give an example of when you would use on demand pricing?

When you are developing and testing applications that have unpredictable usage patterns - not good for years or longer.

▼ Explain savings plan pricing:

It can help you save up to 72 percent. This is good if you don't have commitment issues. It's hourly and gives flexibility in your amazon EC2 usage over the duration of the commitment term.

▼ Explain reserved instances pricing:

It can help you save up to 75 percent. And you can pay whenever you want to. There's two types of reserved instances - standard convertible.

▼ Explain standard reserved instances pricing:

This is a good fit if you know the specs for ur EC2 instance type and which region you want to specify

▼ Explain convertible reserved instances pricing:

This is used when you want to run your EC2 instances in different availability zones or use different instance types

▼ What's the difference between standard reserved instances and saving plans instances?

In saving plans instances you don't need the specs.

▼ After a term ends for reserved instances, what can you do?

Either terminate the instance or purchase a new reserved instance

▼ Explain spot instances pricing:

Where you can request spare EC2 computing capacity for 90 percent of the demand pricing. Allows for flexible start and end times. But if there's no capacity they'll interrupt ur instance.

▼ Explain dedicated hosts instances pricing:

A physical server fully dedicated to your use

Which Amazon EC2 pricing option provides a discount when you make an hourly spend commitment to an instance family and Region for a 1-year or 3-year term?

☐ On-demand

☒ EC2 Instance Savings Plans

☐ Spot Instances

☐ Reserved Instances

Which Amazon EC2 pricing option provides a discount when you specify a number of EC2 instances to run a specific OS, instance family and size, and tenancy in one Region?

☐ Convertible Reserved Instances

☐ EC2 Instance Savings Plans

☐ Spot Instances

☒ Standard Reserved Instances

Module 2 Part 4

▼ How do you make systems highly available?

By making instances redundant

▼ What is scalability?

Involves beginning only w resources u need and then changing based on Demand

▼ What service does AWS provide for this?

Amazon EC2 auto scaling

▼ What does EC2 Auto Scaling do?

It enables you to automatically add or remove EC2 instances in response to changing application demand.

▼ What are the two types of scaling that are available?

Dynamic and predictive

▼ What is dynamic scaling?

Responds to changing demand

▼ What is predictive scaling?

Automatically schedules the right number of amazon EC2 instances based on predicted demand

▼ Explain the difference between scaling up and scaling out?

Scaling up adds more power to the machines that are running and scaling out adds more instances

▼ If you don't specify the desired number of EC2 instances in an auto scaling group, what does the desired capacity default to?

Your minimum capacity

Module 2 Part 5

▼ When you have multiple EC2 instances all running the same program, to serve the same purpose, and a request shows up, how do you know which EC2 instance to send that request?

You can use ELB - elastic load balancing - this runs regionally and is highly available to you. It takes in requests and routes them to instances to be

processed.

▼ Is there any change to the hourly cost for ELB?

No change.

▼ How does the auto scaling service relate to the ELB service?

When the EC2 fleet auto scales out, the auto scaling service just lets the ELB service know it's ready to handle the traffic.

▼ What happens once the fleet scales in?

ELB first stops all new traffic and wait for the existing requests to complete, to drain out. Afterwards, the auto scaling engine can terminate the instances without disruption to existing customers.

▼ Why do we need to consider communication between instances?

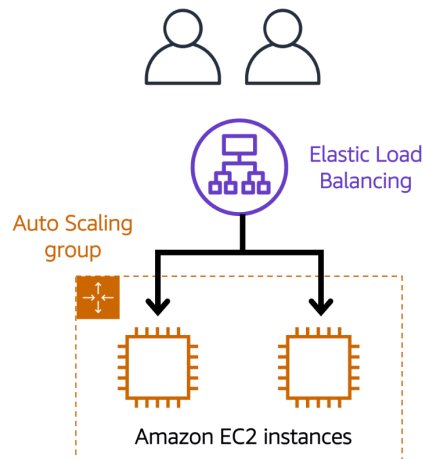
Because in current architecture, we would have to mention this to every frontend instance everytime we make a backend instance. And this is complicated.

▼ Why should we decouple this architecture?

This way the front end doesn't need to know how many back end instances are running, that is complicated to do.

▼ What is a load balancer?

It's a single point of contact for all incoming web traffic to your auto scaling group. You can add or remove EC2 instances in response to the amount of incoming traffic, these requests route to the load balancer first.



Module 2 Part 6

▼ What is the advantage of keeping things loosely coupled?

If one comp fails, it is isolated and therefore won't cause cascading failures → Amazon simple queue service or SQS and SNS

▼ What is SQS?

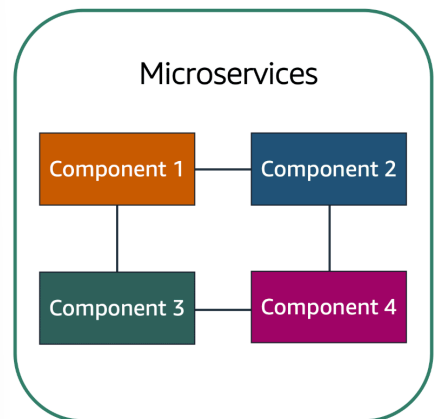
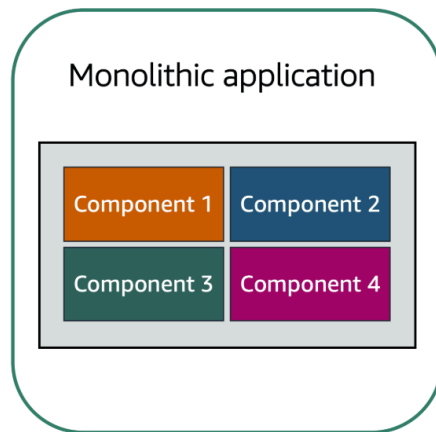
a way to send, store, and receive messages between software components at any volume, without losing messages or requiring other services to be available.

▼ What is a payload?

It's data that is protected under delivery.

▼ What is SNS?

It's a way to send out messages to services but it can also send out notifications to end users.



Which AWS service is the best choice for publishing messages to subscribers?

☐ Amazon Simple Queue Service (Amazon SQS)

☐ Amazon EC2 Auto Scaling

☒ Amazon Simple Notification Service (Amazon SNS)

☐ Elastic Load Balancing

Module 2 Part 7

▼ What are serverless compute options?

A way to go through development without seeing infra or instances that are hosting ur application.

▼ What is AWS Lambda?

A serverless compute option

▼ Explain how AWS Lambda works?

You upload your code into a lambda function, configure a trigger, then the service waits for that trigger.

▼ What if you have 1000 incoming triggers?

AWS Lambda will scale your function to meet demands.

▼ Is Lambda for long running processes?

No

▼ What are the two services that handle containers?

ECS - Elastic Container Service. EKS - Elastic Kubernetes Service

▼ Where do containers run?

On top of EC2 instances and run in isolation from each other similar to how VMs work.

▼ So, where do these services run?

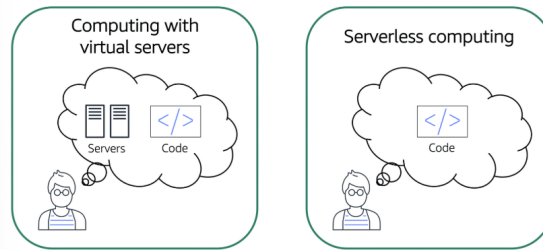
On top of EC2

▼ What if you don't want to think about the underlying OS?

Use Fargate - a serverless compute platform for ECS and EKS.

▼ If you're looking to host short running functions, service-oriented or event driven apps, which service should you use?

AWS Lambda



Comparison between computing with virtual servers (thinking about servers and code) and serverless computing (thinking only about code).