Name: CS443-683
Bilkent ID: Midterm

Exam Rules

- In person exam held on Saturday, April 8, 2023 15:30 - 17:30.
- The exam rooms are booked between 15:30 - 18:30.
- The exam is not an open book exam. Cheat sheets are not accepted.
- Write your answers in the spaces provided below each problem. If you make a mess, clearly indicate where and which one is your final answer. If that is not clear, you will receive your points based on the first answer visible under the question.
- If you have to make an assumption, write it down clearly.
- The exam has a maximum score of 100 points.
- No electronic devices are allowed (this means no laptops/computers/smart-phones)
- Manage your time wisely

The proctor will notify the students verbally when they have 30, 10, and 5 minutes left before the exam is over.

Q1: Enumerate 3 cloud service models and explain them briefly and give one example scenario or example service for each. **(6 Points)**

IaaS: Customers have more control over the underlying infrastructure, including the operating system, storage, and networking components. Users can rent virtualized hardware resources on a pay-as-you-go basis.

PaaS: PaaS offers a platform allowing customers to develop, run, and manage applications without worrying about the underlying infrastructure. PaaS providers manage the underlying infrastructure, including servers, storage, and networking, allowing developers to focus solely on building and deploying applications.

SaaS: SaaS delivers software applications over the internet on a subscription basis.
Users access the software through a web browser or API without needing to install, configure, or maintain any software locally. SaaS providers handle all aspects of software maintenance, including updates, patches, security, and infrastructure management.

Q2: How can IT professionals benefit from the cloud? Write 5 main advantages of using cloud and explain them with 1 sentence. **(5 Points)**

Agility: can experiment and innovate quickly and frequently **(1 Point)**

Cost Saving: Only pay for what you use, leveraging economies of scale **(1 Point)**

Elasticity: Easily scale up or down with the needs of the business **(1 Point)**

Innovation: Focus on business differentiators, not infrastructure **(1 Point)**

Global Footprint: Extensive, reliable, and secure global cloud infrastructure **(1 Point)**

Q3: What is **shared responsibility** model for AWS? According to shared responsibility model who is responsible for **updating operating system** that is running **in** EC2 of the customer? **(5 Points)**

AWS is responsible for the security **of** the cloud, including the infrastructure, hardware, software, and networking that runs AWS services. **(2 Points)**
The customer is responsible for security **in** the cloud, including data, identity and access management, configurations, and applications. . **(2 Points)**
Customer will be responsible for operating system updates. **(1 Point)**

Q4: What are these terms: Explain them shorty according to AWS; **(4 Points)**

Region: geographical area where AWS has multiple data centers **(1 Point)**

Availability Zone: An Availability Zone (AZ) is a distinct data center within an AWS region, **(1 Point)**

Direct Connect: AWS Direct Connect is a service that provides a dedicated network connection between a customer's data center or network and AWS, bypassing the public internet. **(1 Point)**

Points of Presence (PoP): physical locations where AWS extends its network infrastructure to improve performance and reduce latency for end users. **(1 Point)**

Q5: What is the difference between using containers and virtual machines on cloud? And what benefits can be achieved by using containers? **(9 Points)**

- VMs have higher resource overhead compared to containers because they require a separate OS for each VM. Containers virtualize the operating system, enabling multiple isolated user-space instances (containers) to run on a single host OS kernel.
- Containers package applications and their dependencies into a single unit, making them lightweight and portable. Containers share the host OS kernel, leading to faster startup times and lower resource consumption compared to VMs.
- They offer consistent environments across development, testing, and production, reducing the "it works on my machine" problem.
- Containers are highly scalable and can be easily orchestrated using tools like Kubernetes, enabling efficient resource utilization and auto-scaling.
- Portability, Efficiency, Scalability, Consistency, Isolation, Microservices Architecture

Q6: Match the customer scenarios with the best AWS Compute Services **(4 Points) (1 point each)**

1. I want to configure servers, storage, networking, and my OS
2. I want to run servers, configure applications, and control scaling
3. I want to run my containers
4. Run my code when it's needed

AWS Compute Service Name	**Customer Need**
Amazon ECS	2
AWS Lambda	4
Amazon EC2	1
AWS Fargate	3

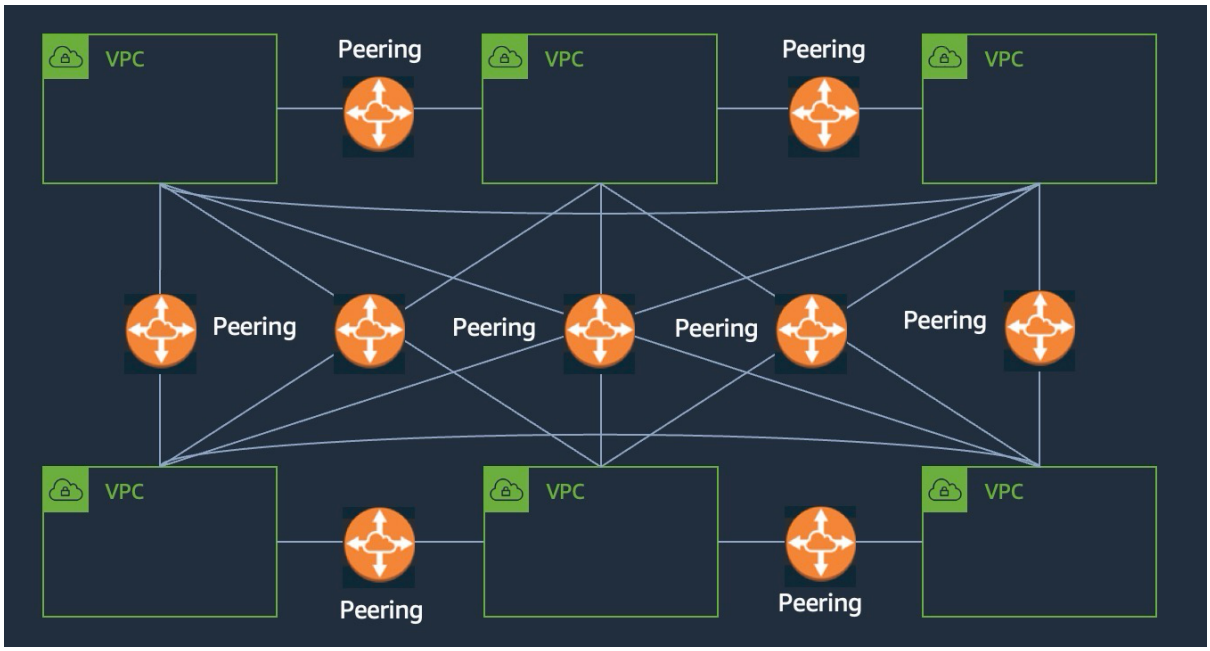
Q7: For the applications/microservices below, write the best compute service name for each according to given information. (EC2, ECS, Lambda) **(3 Points) (1 point each)**

Monolithic application that runs everything in one place with multiple functions included:	EC2
Microservices that runs when they called but can run maximum 15 minutes:	Lambda
Microservices that runs when they called and can run long time:	ECS

Q8: Which processor model cannot be used in AWS EC2? **(4 Points)**

A. ARM Based processors (Graviton)
B. x86 Based Intel Processors
C. **Zhaoxin**
D. AMD Processors

Q9: You are given network diagram below; with which AWS Service you could simplify this architecture? **(5 Points)**



A. S3 B. Fargate C. Virtual Machine **D. AWS Transit Gateway** E. AWS VPN

Q10: Which operating systems can be used in AWS EC2? (Multiple selection) **(5 Points)**
7 true: 5 points
6 true: 4 points
5 true: 3 points
4 true: 2 points
3 true: 1 point

5 false: -5 points
4 false: -4 points
3 false: -3 points
2 false: -2 point

Windows 7 and later	
Windows XP	
Amazon Linux	X
Debian	X
SUSE	X
CentOS	X
iOS	
Red Hat Enterprise Linux	X
Ubuntu	X
MSDoS	
MacOS	X
Android OS	

Q11: You are working as a cloud solutions architect in a tech company that specializes in developing messaging applications. The company is planning to launch a new messaging platform, and you've been tasked with designing the core messaging functionality. You chose to use AWS services for this project. **(50 Points)**

a. As part of your design process, you're considering whether to implement **synchronous** or **asynchronous** messaging for handling user interactions and data processing.
Explain the differences between synchronous and asynchronous applications in the context of a messaging platform. Provide examples of scenarios where each approach would be beneficial.
Finally, propose which approach you would recommend for the new messaging platform and justify your choice. **(10 Points)**

synchronous or **asynchronous can be choosen, justification is important for grading.**
But Asynchronous is better way to do here.

b. As part of your design process in the next steps; you are considering to use **containers** for your microservices. Which service/services do you need to use for running your containers in your architecture and how would you manage your containers? What are the 3 ways for autoscaling of your containers in AWS Cloud for your containers with the container management service tha you choose? **(10 Points)**
ECS or Fargate, **(3 points)**
Kubernetes if ECS, Fargate if Fargate for managing. **(3 Points)**
Horizontal Pod Scaling, Vertical Pod Scaling, Cluster Autoscaling for Kubernetes. **OR**
Automatically Horizontal and Vertical for Fargate. **(4 Points)**

c. You are asked to limit your application for only specific IP addresses from public. Your application should be accessed from only addresses starting with 139.179.X.X and port 80 and. 443. How should your Security Group Inbound Rule look like? **(6 Points) (1 points each blank)**

Source	Type	Protocol	Port Range	Rule
139.179.0.0/16	HTTP	TCP	80	Allow
139.179.0.0/16	HTTPS	TCP	443	Allow

d. After some considerations you came up with that you will need 98 Internal IP addresses for your architecture. How should you specify your Subnet IPv4 CIDR addresses when you are creating this private network on VPC? **(4 Points)**
10.0.0.0/25 or 10.0.0.128/25 **(4 points)**

e. Compare 3 storage types you can use in AWS, Amazon EBS (Elastic Block Store), Amazon EFS (Elastic File Store) and Amazon S3, in terms of their use cases.
In your messaging app you need to keep some programming data **shared** between your containers and **some image** files sent in your messaging app. Which 2 AWS Storage services above you would use in your architecture? (**10 Points)**
Amazon EBS is ideal for block-level storage, persistent storage, and I/O-intensive workloads.
Amazon EFS is suitable for shared file storage, elasticity, scalability, and simplified management.
Amazon S3 is best suited for object storage, static website hosting, data lakes, analytics, and content distribution. **(4 points)**
Choose:
EFS for Shared Data **(3 points)**
S3 for Image files **(3 points)**

f. What is the difference with Hot Data, Warm Data and Cold Data? And what can you do manually or automatically to minimize cost for your storage according to your data? **(5 Points)**
Hot data refers to frequently accessed or actively used data. This type of data typically requires low-latency access and high-performance storage. **(1 point)**
Warm data refers to moderately accessed or less frequently used data that may still be needed for operational or analytical purposes. **(1 point)**

Cold data refers to rarely accessed or inactive data that is retained for compliance, regulatory, or archival purposes. (**1 point)**

Any of two of these below (2 points)

- Implement a data lifecycle management strategy to automatically move data across storage tiers based on its access frequency and importance.
- Design your storage infrastructure with tiered storage architectures that leverage different storage classes or tiers based on data access patterns and requirements.
- Optimize object storage solutions like Amazon S3 by utilizing storage classes designed for infrequently accessed data, such as S3 Standard-IA (Infrequent Access) or S3 One Zone-IA.
- Use data compression and deduplication techniques to reduce storage space and optimize storage costs, especially for warm and cold data that may not require immediate access.
- Regularly review and archive inactive or obsolete data to free up storage space and reduce costs.

g. Calculate the approximate Availability percentage we should provide for our systems that can handle at most 32 seconds interruption for our application in a year. **(5 Points)**

%99,9999 (5 points)

=((((365*24)+6)*60*60)-30)/ ((((365*24)+6)*60*60) = 99,9999