AWS Solution Architect Associate

Version : C03

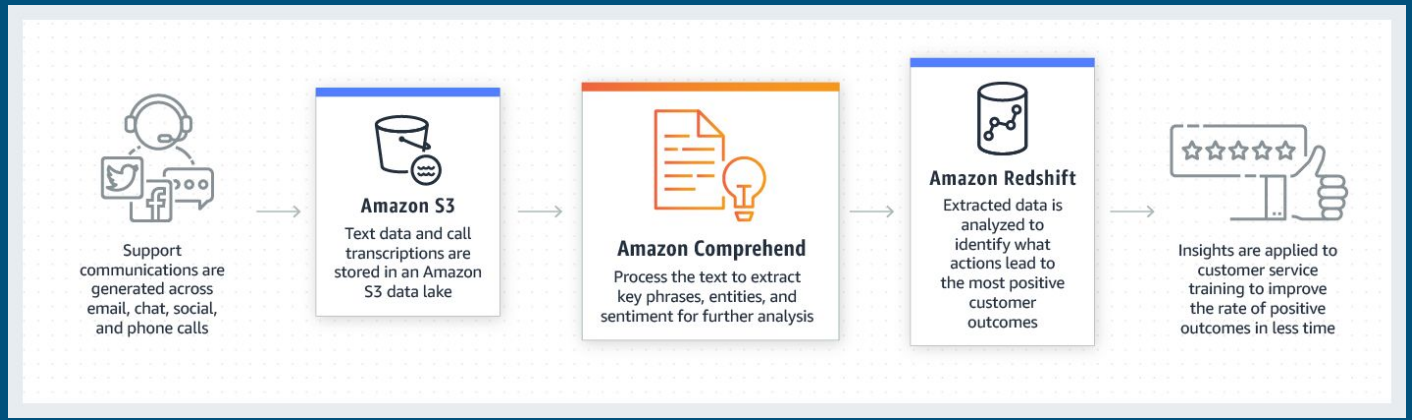
Domain 2

Task 2

Design highly available and/or fault-tolerant architectures

Amazon Comprehend

- Natural language processing (NLP) service
- Uses machine learning to find insights and relationships in text
- Identifies the language of the text , extract key phrases, places , people,brands or events
- AutoML capability can be used to build a custom set of entities or test classification model
- Perform tasks such as sentiment analysis, entity recognition,text classification and extract valuable insights from unstructured text data.



Amazon Polly

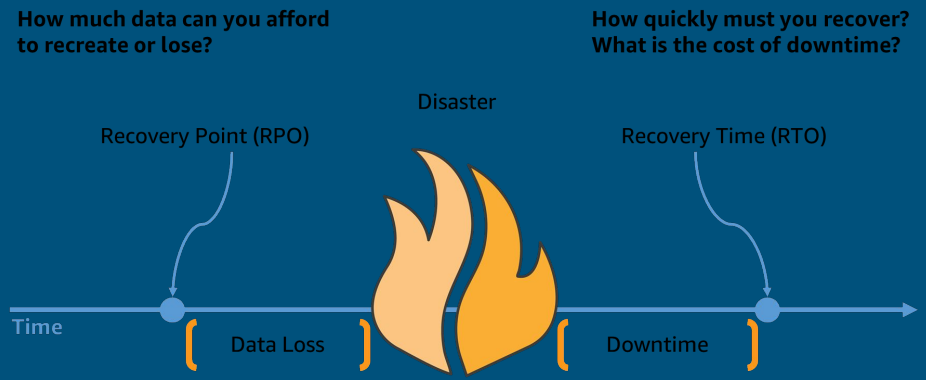
- Text-to-Speech (TTS) cloud service that converts text into lifelike speech.
- Can be used to develop applications that increase engagement and accessibility.
- Supports multiple languages and variety of lifelike voices
- Uses deep learning technologies to synthesize natural-sounding human speech
- Usecases
  - Generate speech in dozens of languages
  - Engage customers with a natural sounding voice
  - Adjust speaking style,speech rate,pitch and loudness

Route Tables

- A route table contains a set of rules called routes that determine where network traffic from your subnet or gateway is directed.
- Key concepts
  - **Main route table** : Default route table in a VPC. It controls the routing for all subnets that are not explicitly associated with any other route table
  - **Custom route table** : Explicitly created for custom VPC
  - **Destination** : Range of IP addresses where you want traffic to go (destination CIDR)
  - **Target** : The gateway, network interface, or connection through which to send the destination traffic.
  - **Route table association** : The association between a route table and a subnet, internet gateway or virtual private gateway
- **Subnet route table** : Route table associated with a subnet.
- **Local route** : Default route for communication within the VPC.
- **Propagation** : Routes are added automatically to subnet route tables for VPN connection If a virtual private gateway is attached to the VPC
- **Gateway route table** : Route table associated with an internet gateway\virtual private gateway.
- **Edge association** : Route table used to route inbound VPC traffic to an appliance.
- **Transit gateway route table** : Route table associated with a transit gateway
- **Local gateway route table** : Route table associated with an Outposts local gateway

RPO & RTO

- RPO aka Replication Point Outage
- RTO aka Recovery Time Objective
- Ensures business continuity and disaster recovery.
- RPO is the amount of data that must remain available during an outage
- RTO is the amount of time it takes for the data to be recovered after a disaster or outage
- Both RPO and RTO must be carefully considered and planned in order to ensure that data remains accessible and business operations can continue smoothly in the face of disruptions.
- RTO and RPO targets must be set on an application-by-application basis



Distributed Design Patterns

- Refers to architecture patterns that allow for efficient and scalable distributed systems
- These patterns include techniques such as microservices, serverless and data federation, among others, and address challenges such as load balancing, fault tolerance, and data management.
- Different distributed patterns
  - Microservices Architecture
  - Service Oriented Architecture
- Microservices architecture breaks down the application into small independent services that can be developed and deployed independently. AWS services - AWS Lambda, AWS API Gateway and Amazon ECS
- Service Oriented Architecture designs the system as a collection of loosely coupled services that communicate well with defined interfaces. AWS services - Amazon SQS

Immutable Infrastructure

- AWS Immutable Infrastructure refers to a type of infrastructure that uses immutability to ensure its reliability and security.
- Immutability means that the infrastructure cannot be altered once it has been created, which makes it easier to manage and more resilient to failure.
- AWS provides several services that make it easy to create and manage immutable infrastructure,
  - AWS CloudFormation
  - AWS Lambda
  - AWS WAF
- These services allow developers to define the infrastructure they need as code, which makes it easier to deploy and manage this infrastructure across multiple environments.
- Benefits
  - Enhanced security
  - Better audit and compliance
  - Increased reliability

Proxy Concepts

- An AWS proxy is a service that provides an additional layer of security and management for accessing AWS services.
- Proxies act as an intermediary between a client application and the AWS service it intends to access.
- AWS proxies route requests, sign and encrypt them and manage authentication and authorization.
- There are two main AWS proxy types
  - **AWS IAM proxy**
  - **AWS API Gateway proxy**
- IAM proxy is commonly used for web requests and is managed at the IAM service level
- API Gateway proxy is managed at the API Gateway service level and is commonly used for applications that require advanced authentication and authorization.

Amazon RDS Proxy

- Handles the network traffic between the client application and the database
- Reduces the memory and CPU overhead for connection management on the database
- The infrastructure for RDS Proxy is highly available and deployed over multiple Availability Zones (AZs)
  - Each proxy contains a target group.
  - Target group embodies the RDS DB instance that the proxy can connect to.
  - RDS DB instance associated with a proxy are called the targets of that proxy.
  - When you create a proxy through the console, RDS Proxy also creates the corresponding target group and registers the associated targets automatically.
  - An engine family is a related set of database engines that use the same DB protocol.

Amazon RDS Proxy

- **Connection Pooling**
  - Optimization that reduces the overhead associated with opening/closing connections and keeping many connections open
  - This overhead includes memory needed to handle each new connection.
  - Involves CPU overhead to close each connection and open a new one.
- **RDS Proxy Security**
  - Uses the existing RDS security mechanisms such as TLS/SSL & AWS IAM
  - Act as an additional layer of security between client applications and the underlying DB
- **Transactions**
  - All the statements within a single transaction always use the same underlying database connection.
  - The connection becomes available for use by a different session when the transaction ends.
- **Using TLS\SSL With RDS Proxy**
  - RDS proxy can be connected using TLS/SSL protocol
  - "Require Transport Layer Security" setting can be set to create and modify a proxy
  - Ensures that the session between the client and RDS proxy endpoint uses TLS/SSL
  - Supports TLS protocol version 1.0,1.1,1.2 and 1.3
- **Failover**
  - HA feature that replaces a db instance with another one when the original is unavailable
  - Applies to RDS DB instances in a Multi-AZ configuration.
  - Makes the applications more resilient to DB failovers.
  - RDS Proxy connects to the standby DB without dropping idle application connections when the original DB instance becomes unavailable
  - Helps speed up\simplify the failover process.

Service Quotas And Throttling

- AWS service quotas and throttling concepts refer to mechanisms used by AWS to manage resource usage and prevent users from exceeding specific limits.
- Quotas are fixed limits on the number of API calls or other resource usage that a user can perform within a specified time period
- Throttling is a method of adjusting the speed of API requests in a proportional manner to prevent an overuse of the system.
- It helps to ensure that users are not negatively impacting the availability of resources for others.

Storage Options/Characteristics

Storage Option	Description	Usecases	Durability	Availability	Scalability	Access Methods
Amazon S3	Object Storage	Data lakes, Backups,Static website hosting	99.999999999% (11 9s)	99.99%	Scalable to exabytes of data	Restful API,SDK CLI
Amazon EBS	Block Storage	Boot Volumes, Databases,Transactional Workloads	High Durability	99.99%	Scalable to petabytes	Attach to EC2 instance
Amazon EFS	Managed NFS file storage	Shared File Storage	High Durability	99.99%	Elastic, Scales with storage growth	NFS protocol
Amazon FSx	Fully managed file storage	Windows File Server , Lustre high performance storage	High Durability	99.99%	Elastic, Scales with storage growth	SMB/CIFS(FSx for windows). Lustre protocol
Amazon S3 Glacier	Low cost archival storage service	Long term data archiving and backup	99.999999999% (11 9s)	99.99%	Scalable to exabytes of data	Restore times from minutes to hours
Amazon Storage Gateway	Hybrid storage service connecting on-premises environments to AWS storage	Backup . Disaster recovery , File shares	High Durability	99.99%	Flexible . Integrates with AWS services	iSCSI , NFS and SMB

Workload Visibility

AWS X-Ray

- ➔ This service enables developers to analyze and debug applications and services, both on-premise and in the cloud.
- ➔ Using X-Ray,real-time performance data can be captured, health of distributed systems can be monitored and application performance issues can be troubleshooted
- ➔ Key features
  - ◆ Automated tracing
  - ◆ Distributed tracing
  - ◆ Performance cloud services
  - ◆ Supports multiple programming languages
  - ◆ Seamlessly integrate with existing applications and infrastructure

Automation strategies to ensure infrastructure integrity

Automation strategies :

- ➔ Infrastructure as Code (IaC)
- ➔ Configuration Management
- ➔ Continuous Integration/Continuous Deployment
- ➔ Automated Remediation
- ➔ Event-Driven Automation
- ➔ Patch Management
- ➔ Monitoring and Alarming Automation
- ➔ Policy Enforcement
- ➔ Backup and Disaster Recovery Automation
- ➔ Compliance as Code

Highly available and/or fault-tolerant architecture

- Key components and strategies to achieve highly available and fault-tolerant architecture in AWS
  - Multiple Availability Zones (AZs)
  - Load Balancing
  - Auto Scaling
  - Database Replication
  - Content Delivery Network (CDN)
  - Global Accelerator
  - Backup and Restore
  - Monitoring and Alarming

Designs to mitigate single points of failure

- Design considerations to mitigate Single Point Of Failure (SPOF)
  - Use Multiple Availability Zones (AZs)
  - Load Balancers
  - Multi-AZ Database Deployments
  - Auto Scaling
  - Elastic File System (EFS)
  - Cross-Region Replication
  - Multi-Region Deployments
  - Distributed DNS
  - Stateless Architecture
  - Regular Testing

AWS services for Workloads

AWS offers a comprehensive suite of services tailored to different types of workloads, enabling users to effectively deploy, manage and optimize applications and processes

- **Compute workloads**
  - AWS Lambda , AWS EC2
  - Provides solutions for scalable virtual servers
- **Big Data workloads**
  - Amazon EMR
  - For processing and analyzing large datasets
- **Application workloads**
  - AWS Elastic Beanstalk
  - Supports both Commercial off the shelf (COTS) and custom applications
- **Workload discovery**
  - Provides tools for workload discovery,aiding in understanding and optimizing workloads
- **Well-Architected workloads**
  - AWS Well-Architected Tool
  - Helps in designing reliable,secure,efficient and cost-effective workloads

AWS Solution Architect Associate

Version : C03

Domain 2

Task 2

The END