Module 3

Global Infrastructure and Reliability in AWS

AWS provides a global infrastructure that ensures reliability and high availability for your applications. With AWS, you can choose from multiple regions and availability zones, allowing you to place resources closer to your customers for faster access and better performance. Additionally, AWS offers a range of services to help you maintain the reliability of your applications, such as Auto Scaling, Elastic Load Balancing, and AWS CloudFormation. These tools allow you to automate the scaling and management of your resources, ensuring that your applications can handle changes in traffic and demand while maintaining high availability and reliability.

Edge locations

An edge location is a site that amazon CloudFront uses to store cached copies of your content closer to your customers for faster delivery.

AWS services

1. AWS management console:

The AWS management console is a web based interface for accessing and managing AWS services. You can quickly access recently used services and search for other services by name, keyboard, or acronym.

2. AWS command line interface:

To save time when making API requests, you can use the AWS command line interface. It enables you to control multiple AWS services directly from the command line withing one tool.

3. Software development kits:

Another option for accessing and managing AWS services is the software development kits. It make it easier for you to use AWS services through an API designed for your programming language or platform. SDKs enables you to use AWS services with your existing applications or create entirely new application that will run on AWS.

AWS Elastic Beanstalk

With AWS elastic beanstalk, it provide code and configuration settings and elastic and it deploys the resources necessary yo perform the following tasks:

1. adjust capacity
2. load balancing
3. automatic scaling
4. application health monitoring

AWS CloudFormation:

With AWS CloudFormation, you can treat your infrastructure as code. You can build an environment by writing lines of code instead of using the AWS management console to individually provision resources.

AWS Auto Scaling:

AWS Auto Scaling enables you to automatically adjust the capacity of your resources to maintain steady, predictable performance at the lowest possible cost. This service automatically scales your resources up or down based on demand, ensuring that you always have the resources you need to handle your application's traffic. Additionally, you can use AWS Auto Scaling with other AWS services, such as Amazon EC2 and Amazon ECS, to ensure that your resources are always available and performing optimally.

AWS CloudFront:

AWS CloudFront is a content delivery network that securely delivers data, videos, applications, and APIs to customers globally with low latency, high transfer speeds, all within a developer-friendly environment. You can use AWS CloudFront to cache and distribute your content from edge locations, improving the performance of your applications and reducing the load on your origin servers. Additionally, AWS CloudFront integrates with other AWS services, such as Amazon S3 and Elastic Load Balancing, to provide a comprehensive solution for delivering your content to customers.

AWS Direct Connect:

AWS Direct Connect is a dedicated network connection between your on-premises infrastructure and AWS. This service bypasses the public internet and provides a more reliable and secure connection to your AWS resources. With AWS Direct Connect, you can establish a private, high-speed connection to AWS, making it easier to transfer large amounts of data and run latency-sensitive applications. Additionally, AWS Direct Connect integrates with other AWS services, such as Amazon VPC and AWS Transit Gateway, to provide a seamless and scalable solution for connecting your on-premises infrastructure to the cloud.