**Lecture Notes for AWS Solution Architect**

**MODULE-V**

**AWS Auto Scaling**

AWS Auto Scaling monitors your applications and automatically adjusts capacity to maintain steady, predictable performance at the lowest possible cost. Using AWS Auto Scaling, it’s easy to setup application scaling for multiple resources across multiple services in minutes. The service provides a simple, powerful user interface that lets you build scaling plans for resources including Amazon EC2 instances and Spot Fleets, Amazon ECS tasks, Amazon DynamoDB tables and indexes, and Amazon Aurora Replicas. AWS Auto Scaling makes scaling simple with recommendations that allow you to optimize performance, costs, or balance between them. If you’re already using Amazon EC2 Auto Scaling to dynamically scale your Amazon EC2 instances, you can now combine it with AWS Auto Scaling to scale additional resources for other AWS services. With AWS Auto Scaling, your applications always have the right resources at the right time.

**Benefits**

**SETUP SCALING QUICKLY**

AWS Auto Scaling lets you set target utilization levels for multiple resources in a single, intuitive interface. You can quickly see the average utilization of all of your scalable resources without having to navigate to other consoles. For example, if your application uses Amazon EC2 and Amazon DynamoDB, you can use AWS Auto Scaling to manage resource provisioning for all of the EC2 Auto Scaling groups and database tables in your application.

**MAKE SMART SCALING DECISIONS**

AWS Auto Scaling lets you build scaling plans that automate how groups of different resources respond to changes in demand. You can optimize availability, costs, or a balance of both. AWS Auto Scaling automatically creates all of the scaling policies and sets targets for you based on your preference. AWS Auto Scaling monitors your application and automatically adds or removes capacity from your resource groups in real-time as demands change.

**AUTOMATICALLY MAINTAIN PERFORMANCE**

Using AWS Auto Scaling, you maintain optimal application performance and availability, even when workloads are periodic, unpredictable, or continuously changing. AWS Auto Scaling continually monitors your applications to make sure that they are operating at your desired performance levels. When demand spikes, AWS Auto Scaling automatically increases the capacity of constrained resources so you maintain a high quality of service.

**PAY ONLY FOR WHAT YOU NEED**

AWS Auto Scaling can help you optimize your utilization and cost efficiencies when consuming AWS services so you only pay for the resources you actually need. When demand drops, AWS Auto Scaling will automatically remove any excess resource capacity so you avoid overspending. AWS Auto Scaling is free to use, and allows you to optimize the costs of your AWS environment.

**How it works**



**Amazon Simple Notification Service (Amazon SNS)**

Amazon Simple Notification Service (Amazon SNS) is a fully managed messaging service for both application-to-application (A2A) and application-to-person (A2P) communication.

The A2A pub/sub functionality provides topics for high-throughput, push-based, many-to-many messaging between distributed systems, microservices, and event-driven serverless applications. Using Amazon SNS topics, your publisher systems can fanout messages to a large number of subscriber systems including Amazon SQS queues, AWS Lambda functions and HTTPS endpoints, for parallel processing, and Amazon Kinesis Data Firehose. The A2P functionality enables you to send messages to users at scale via SMS, mobile push, and email.

**Benefits**

**Modernize and decouple your applications**

Amazon SNS enables you to modernize your applications and decouple them into smaller, independent components that are easier to develop, deploy and maintain. Leveraging a pub/sub event-driven architecture for your application improves performance, reliability, and allows each component to scale independently.

**Send messages directly to millions of users**

Amazon SNS enables you to send messages or notifications directly to users with SMS text messages to over 200 countries, mobile push on Apple, Android, and other platforms or email (SMTP). Amazon SNS provides redundancy across multiple SMS providers and allows you to push mobile notifications using a single endpoint for all platforms.

**Reliably deliver messages**

Amazon SNS uses a number of strategies that work together to provide message durability. To start, published messages are stored across multiple, geographically-separated servers and data centers. If a subscribed endpoint isn't available, Amazon SNS executes a message delivery retry policy. To preserve any messages that aren't delivered before the delivery retry policy ends, you can create a dead-letter queue. You can also subscribe Amazon Kinesis Data Firehose delivery streams to SNS topics, which allows messages to be sent to durable endpoints such as Amazon S3 buckets or Amazon Redshift tables.

**Automatically scale your workload**

Amazon SNS leverages the proven AWS cloud to dynamically scale with your application. Amazon SNS is a fully managed service, taking care of the heavy lifting related to capacity planning, provisioning, monitoring, and patching. The service is designed to handle high-throughput, bursty traffic patterns and enables you to send millions of messages per second.

**Ensure accuracy with message ordering and deduplication**

Amazon SNS FIFO topics work with Amazon SQS FIFO queues to ensure messages are delivered in a strictly ordered manner and are only processed once (deduplicated). This enables you to maintain consistency when processing transactions across a single or multiple independent services where it’s critical that messages are in the correct order. It also allows you to offload the effort of writing custom code for ordering and message deduplication.

**Simplify your architecture with Message Filtering**

Amazon SNS helps you simplify your pub/sub messaging architecture by offloading the message filtering logic from your subscriber systems, and message routing logic from your publisher systems. With Amazon SNS message filtering, subscribing endpoints receive only the messages of interest, instead of all messages published to the topic. Amazon CloudWatch gives visibility into your filtering activity, and AWS CloudFormation allows you to deploy subscription filter policies in an automated and secure manner.

**How it works**

**PUB/SUB**



**SMS**



**MOBILE PUSH**

