

* Messaging 의 정의와 필요성

- 메시징은 애플리케이션(서비스)간의 통신을 목적으로 한 매개체인 메시지를 주고 받는 것을 의미함
 - 주문을 담당하는 애플리케이션과 배송을 담당하는 애플리케이션이 각각 있다면 원활한 서비스 제공을 위해 서로 통신을 해야함
 - 주문 애플리케이션과 배송 애플리케이션이 서로 주고 받는 데이터가 바로 메시지
 - 주문 애플리케이션과 배송 애플리케이션이 직접 메시지를 주고 받는 것을 동기식 통신이라 부름
 - 주문 애플리케이션과 배송 애플리케이션 사이에 매개체(Queue)를 두고 메시지를 주고 받는 것을 비동기식 통신이라 부름
 - 주문 / 배송 애플리케이션의 데이터 처리 스펙이 다르다면 통신에 문제가 생길 수 있음, 한쪽이 메시지를 처리하는 동안 다른 한쪽은 그 메시지를 처리하지 못하고 유실할 것임
 - 각각의 애플리케이션이 직접 통신하는 것을 방지하고(Decouple), 매개체(Queue)를 통해 메시지를 주고 받는 것이 이러한 문제에 대한 대안이 될 수 있음
 - SQS, SNS, Kinesis 가 이러한 분리 매개체에 해당함
 - SQS : Queue Model
 - SNS : Public/Subscribe Model
 - Kinesis : Real-time Streaming Model

* Kinesis

- 실시간 데이터(메시지, 데이터가 빠르게 생성되는)를 수집하고 저장하여 분석하는 서비스
- kinesis 는 4 가지 서비스로 구성됨(시험 출제 항목은 "Data" 서비스가 주를 이룸)
 - Kinesis Data Streams
 - Kinesis Data Firehose
 - Kinesis Data Analytics
 - Kinesis Video Streams

* Kinesis Data Streams

- 생산자가 생성하는 대규모 실시간 데이터를 소비자로 하여금 처리하게 하기 위해 "수집 및 저장"하는 서비스
- 생산자가 데이터를 생성하여 Data Stream에 Push 하고 저장하면 소비자가 Data Stream에 적재된 데이터를 EC2에서 실행하는 애플리케이션 혹은 Data Firehose를 통해 결과를 저장할 수 있음
- 실시간으로 쌓은 데이터는 아래의 서비스에서 수집할 수 있음
 - Kinesis Data Firehose
 - Kinesis Data Analytics
 - Kinesis Client Library(KCL)
 - AWS Lambda
 - EC2
 - EMR Cluster 등
- 사용 예시
 - 다수의 웹사이트에서 생산하는 Click Stream Data를 모아 Data Stream에 저장한 후, Data Firehose를 거쳐 S3에 저장 및 분석
- Kinesis Data Streams의 주요 용어
 - Data Record(데이터 레코드) : Kinesis Data Streams에서 저장되는 데이터의 단위
 - Shard(샤드) : 일정 수 이상의 데이터 레코드가 모인 고유한 순서
 - Data Stream(데이터 스트림) : 일련의 데이터 집합, 일정 수 이상의 샤드가 모여 구성됨
 - Partition Key(파티션 키) : 샤드별로 데이터를 그룹화하는데 사용되는 키
 - 데이터 레코드가 순서를 이루어 샤드가 되고 샤드가 모여 데이터 스트림이 됨
 - 보존기간 : 데이터 레코드를 스트림에 추가한 후 데이터 레코드에 액세스할 수 있는 시간, 24 시간

* Kinesis Data Firehose

- 생산자가 생성하는 실시간 데이터를 캡처 및 변환하여 지정된 대상으로 "전송"할 수 있는 서비스
- 데이터의 변환이 가능하다는 특징을 지님
- 전송 대상에는 Amazon S3, Amazon Redshift, Amazon OpenSearch Service, Kinesis Data Analytics 뿐만 아니라 Splunk 같은 서비스 공급자가 존재함

* Kinesis Data Analytics

- 실시간 데이터를 처리하고 분석할 수 있는 서비스
- SQL, Python, Scala 등을 활용해 데이터 분석이 가능
- Data Streams나 Firehose로부터 받은 데이터를 분석한 후, 다시 동일한 소스로 보낼 수 있음
- Kinesis Data Analytics Studio Notebook을 시작하여 스트림 데이터를 대화형으로 분석할 수 있음