



[AWS] 9. 자동화 SQS, Kinesis

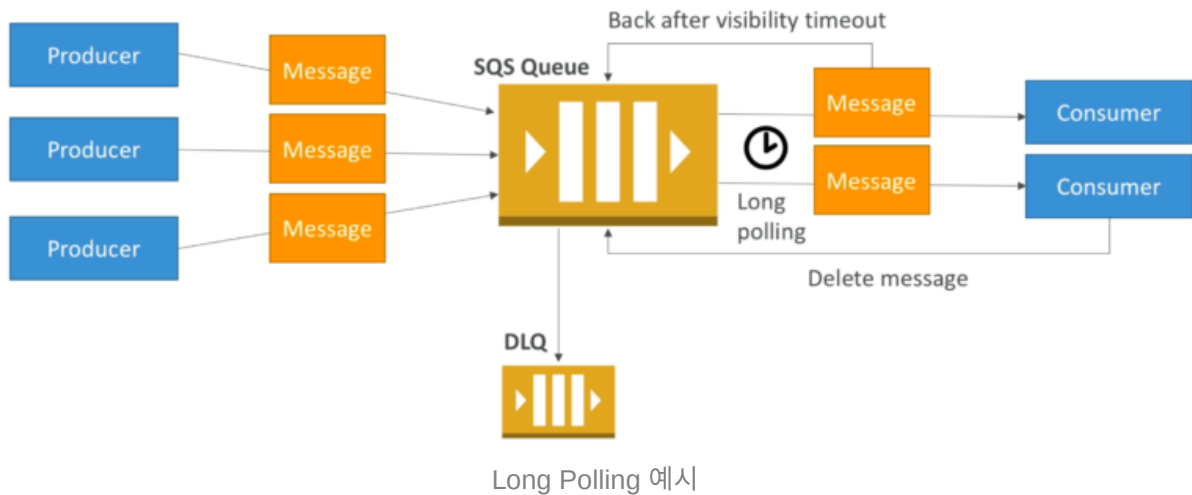
- References

- AWS 공인 솔루션스 아키텍트 스터디 가이드 - 어소시에이트 3/e - 9장
- <https://lannstark.tistory.com/88> (SQS 간단 정리)
- <https://aws.amazon.com/ko/kinesis/> (AWS Kinesis Document)

- 마이크로서비스 (Microservice) 개요

- 애플리케이션 개발자는 루즈 커플링 (Loose coupling) 이라는 설계 원칙을 이용해 애플리케이션의 확장성 및 신뢰성을 높이는데 불필요한 요소를 최소화 하여 하나의 서버에서 실행되는 단일 애플리케이션을 구현, 각 컴포넌트를 마이크로 서비스라 부르는 요소로 세분화 한다. 각 마이크로 서비스는 서로 다른 서버에서 실행되지만 **메시지**를 전송하여 서로 긴밀하게 소통한다.

※ SQS, Simple Queue Service



• SQS 란

- 애플리케이션을 구성하는 다양한 컴포넌트가 서로에게 메시지를 전송할 수 있도록 돕는 관리형 메시지 서비스
- 고가용성 및 고탄력성을 제공
- SQS로 전송되는 API 호출횟수를 줄이기 위해 최대 10개의 메시지를 묶어 일괄 처리 가능

• SQS 의 Queue와 처리 방식

- SQS 는 처리해야 할 메시지를 담은 큐를 생성하며, 큐에 메시지를 넣는 프로듀서 컴포넌트와 메시지를 읽는 컨슈머 컴포넌트로 구성
- 메시지의 최대 크기는 256KB
- 메시지 처리 방식
 - Producer 가 SendMessage 액션을 통해 큐에 하나 이상의 메시지(이동중 메시지 또는 인플라이트 메시지)를 넣음
 - Consumer가 큐에 담긴 메시지를 처리할지 확인, ReceiveMessage 액션을 통해 메시지 소비
 - Consumer는 메시지를 처리 후 DeleteMessage 액션을 통해 큐에서 삭제

- 가시성 중지기간 (Visibility Timeout)

- 컨슈머는 메시지의 삭제 여부를 결정하고, 특정 컨슈머 객체가 메시지를 확인하면 SQS는 일정시간 동안 해당 메시지를 다른 컨슈머가 확인할 수 없게 한다. 이를 가시성 중지기간 이라 함
- 기본 가시성 중지기간 30초, 최소 0초에서 12시간 까지 설정 가능

- 보유기간

- 메시지가 큐에 보유되는 기간은 기본 4일, 최소 1분에서 14일까지 설정가능

- 딜레이 큐와 메시지 타이머

- 메시지를 보낸 즉시가 아닌 일정 시간(delay) 이후에 받고 싶은 경우 사용
- 딜레이 큐 (지연 큐)
 - 큐에 메시지를 넣을 때 큐마다 지연시간 설정 가능
 - 기본 큐별 딜레이 시간은 0초, 최대 15분 까지 설정 가능
- 메시지 타이머
 - 개별 메시지는 메시지 타이머를 이용하여 지연시간 설정 가능
 - 기본 메시지 타이머는 0초 최대 15분 까지 설정 가능
- 메시지 타이머를 설정한 경우 딜레이 큐는 무시된다.

- 큐 타입

- 메시지 처리방식에 따른 큐 타입, SQS는 스탠다드 큐와 FIFO큐 두가지를 제공

스탠다드 큐	거의 무제한 처리 성능 제공, 신속하게 다수의 메시지를 처리, 메시지는 순서와 무관하게 전달되며 중복되어 전달되기도 한기에 애플리케이션은 이와같은 중복 메시지 처리 기능을 갖고 있어야 함, 기본 큐타입, 최대 12만개의 인플라이트 메시지 처리
FIFO 큐	초당 3천개의 메시지를 큐에 전달, 메시지는 도착 순서대로 큐에 전달되며 단 한번만 기록되어 중복 메시지를 피할수 있다, 약 2만개의 인플라이트 메

시지 처리 가능, 메시지 단위로 큐를 분할해 큐에 입력된 메시지의 하위 그룹을 만들 수 있도록 한다, 메시지 단위로 큐를 분할하여 입력된 메시지의 하위 그룹(메시지 그룹 ID를 부여하여 관리)을 만들수 있도록 한다,

• 폴링(Polling)

- 큐에서 메시지를 확인할 때 메시지 도착 여부 조회 옵션
- 숏폴링 또는 롱폴링으로 선택
- 숏폴링 (기본값)
 - 일부 메시지가 누락이 있더라도 즉시 메시지를 확인해야 할 때 사용
 - SQS는 대기중인 메시지 내역만 확인
 - 큐에 들어온 메시지를 즉시 확인하거나, 큐에 메시지가 없음을 즉시 확인 가능
 - 간혹 지연시간이 짧아 큐에 메시지가 있는 상황에서 없다는 응답을 받을 경우도 있어, 모든 메시지를 착오 없이 확인하려면 여러 번 조회해야 한다.
- 롱폴링
 - 지연이 있더라도 큐에 있는 모든 메시지를 정확하게 확인해야 할 때 사용
 - SQS는 큐에서 대기중인 모든 메시지를 반환
 - 모든 큐 서버를 확인하므로 응답시간이 20초가량 걸릴 수 있다.
 - 숏폴링에 비해 조회 빈도수가 작으므로 조금 더 경제적인 옵션

• 데스레터 큐

- 컨슈머가 제대로 처리하지 못한 메시지가 큐에 남을때 이 메시지를 데스레터(dead letter)라고 한다.
- 이와같은 문제를 처리하기 위해 SQS를 통해 해당 메시지를 큐에서 꺼내 데스레터 큐에 따로 보관
- 데스레터 큐를 생성하려면 기존 큐와 동일한 타입의 큐를 생성 후 maxReceiveCount 속성을 이용하여 메시지의 최대 인출 시도 횟수를 설정한다
- 데스레터 큐는 소스큐와 동일한 리전에 있어야 함

- 보유기간의 영향을 받으며, 하나의 메시지가 데스레터 큐로 이동하면 원본 생성 일자를 기준으로 삭제

※ Amazon Kinesis

• Kinesis 란

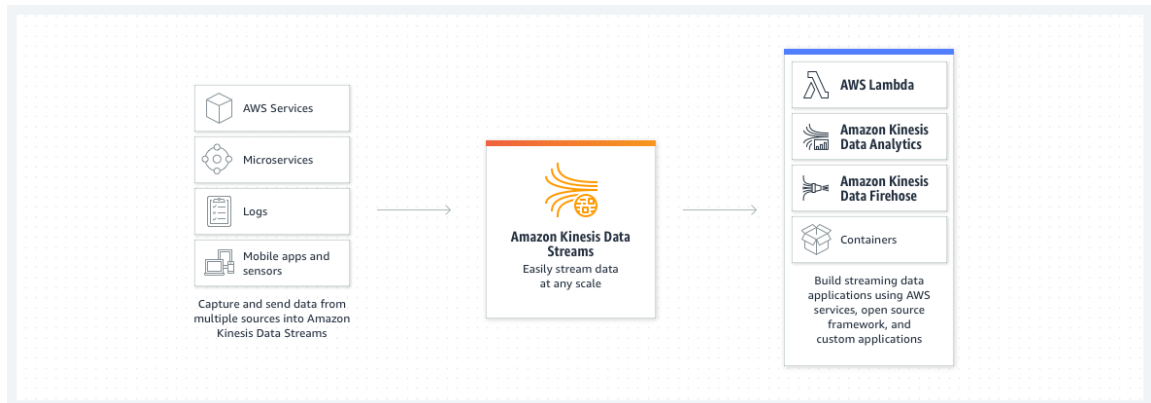
- 스트리밍 데이터의 수집, 처리, 저장, 전송을 위한 서비스 모음
- 초당 수 기가바이트 수준으로 유입되는 오디오, 비디오, 애플리케이션 로그, 원격측정 데이터 등 다양한 스트리밍 데이터를 처리 가능하다.
- 스트리밍 데이터 타입에 따라 다양한 Kinesis 서비스 제공
 - Kinesis Video Streams
 - Kinesis Data Streams
 - Kinesis Data Firehose

• Kinesis Video Streams

- 웹캠, 보안용 카메라, 스마트폰 카메라 등 다양한 스트리밍 비디오 데이터를 거의 무한대로 처리 및 분류할 수 있는 실시간 비디오 처리 서비스
- 사용 목적
 - 이미지 인식 등 컴퓨터 비전 애플리케이션
 - 스트리밍 비디오 애플리케이션
 - 양방향 비디오 회의 애플리케이션
- 프로듀서-컨슈머 모델 (Kinesis 스트림으로 데이터를 유입 시키는 데이터 소스를 프로듀서, 스트림 데이터를 읽어들이는 애플리케이션을 컨슈머)
- Kinesis 는 타임 스탬프를 이용하여 비디오 스트림을 분류(인덱싱)하고, 개별 비디오 스트림 저장시간은 기본 24시간 최대 7일 저장 가능하다

- 비디오 재생을 위한 HLS(HTTPS Live Streams), DASH(Dynamic Adaptive Streaming Over HTTPS) 규격 지원
- 피어 투 피어 비디오 회의를 위한 WebRTC(Web Real-Time Communication) 규격 지원

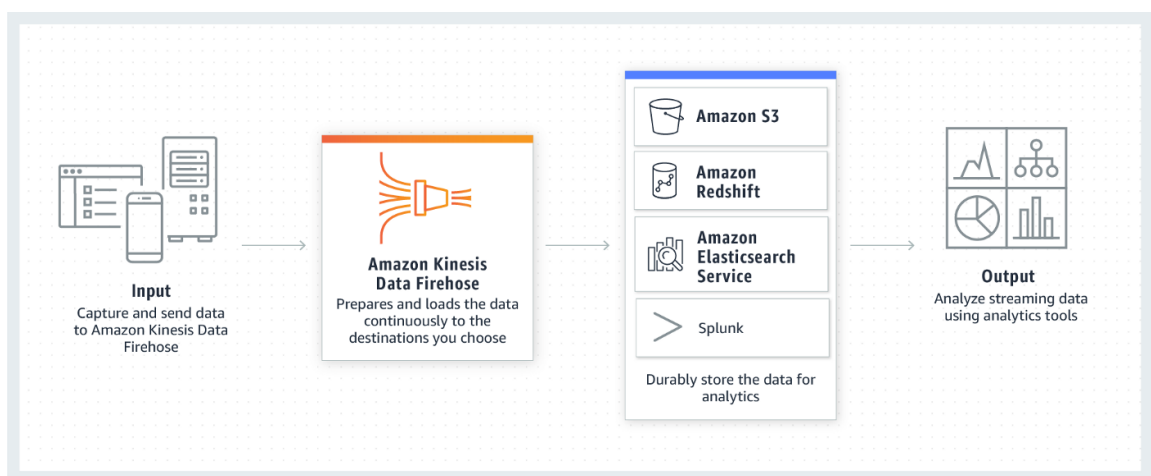
• Kenesis Data Streams



- 프로듀서가 생성한 데이터의 수집, 처리, 저장 을 위한 스트리밍 데이터 파이프라인 서비스
- 빅데이터 분석 애플리케이션 (MapReduce 와 같은) 이 컨슈머가 될 수 있다.
- 처리 및 저장할 수 있는 주요 바이너리 데이터 타입
 - 애플리케이션 로그
 - 주식 거래 데이터
 - 소셜 미디어 피드
 - 금융 거래 데이터
 - 위치 추적 데이터
- 프로듀서-컨슈머 모델
- 프로듀서가 Kenesis Data Streams 에 데이터를 전달하면, 데이터는 타입과 소스에 따라 상이한 방식으로 데이터 레코드에 기록
- KPL(Kenesis Producer Library) 를 이용하여 애플리케이션에서 생성된 데이터를 직접 Kenesis에 전송 가능
- 데이터 레코드에 파티션 키, 시퀀스 번호와 함께 저장, 타임스탬프와 같은 인덱싱 목적으로 데이터의 고유 의 순서에 따라 저장되고 처리 될 수 있다.

- 다수의 컨슈머 객체가 하나의 스트림을 동시에 읽는것 가능 (Fan-out)
- 스트림 레코드에 입력되는 시간과 컨슈머가 읽는 시간의 차이 (put-to-get-delay)는 1초 미만
- 최대 스트림 처리 용량은 샤드(Shard)값에 따라 달라짐
 - 샤드는 데이터 레코드의 시퀀스를 구분하게 해주고 나뉘고 고정된 용량을 지니며 초당 5회 읽기 작업, 초당 2MB의 데이터 처리를 제공

• Kinesis Data Firehose



- 스트리밍 데이터를 수집하고 목적지로 전송하기 전 해당 데이터를 변환하는 작업을 수행(데이터 정렬, 다른 포맷으로 변환 등)
- Lambda 함수를 이용하여 데이터 변환, 커스텀 변환 작업의 유연성 제공
- 원본 데이터 변환 실패에 대비한 S3에 전송이 가능하며 전송 데이터의 버퍼 수준도 설정 가능
- Kinesis Data Streams 의 데이터를 Kinesis Data Firehose로 전송해 최대 7일로 설정된 보유기간을 필요시 까지 늘릴 수 있으며 이를 위해 Kinesis Data Firehose 전송 스트림 설정에서 소스로 Kinesis Data Stream을, 목적지로는 S3 버킷을 설정

• Kinesis Data Streams 와 Kinesis Data Firehose의 차이

- 대량의 스트리밍 데이터를 수집한다는 점은 동일하지만 서비스 목적이 다르다

- **Kinesis Data Streams**

- 개방형 프로듀서-컨슈머 모델 사용 (어떤 애플리케이션이든 해당 스트림을 구독 가능)
- 일대다 모델 사용 (하나의 데이터 스트림을 다수의 컨슈머가 구독)
- 커스텀 애플리케이션에 스트리밍 데이터를 전송하는데 적합

- **Kinesis Data Firehose**

- 개방형 프로듀서-컨슈머 모델이 아니며, 하나 이상의 데이터 목적지를 지정해야 한다.
- 다른 AWS 서비스 및 서드파티 애플리케이션과 긴밀한 통합이 가능
- Readshift, S3, Splunk 등의 서비스에 스트리밍 데이터를 전송하는데 적합

- **SQS / Kinesis 서비스 비교**

서비스	데이터 변환	최대 보유기간	데이터 관리 모델
SQS	불가	14일	프로듀서-컨슈머
Kinesis Video Streams	불가	7일	프로듀서-컨슈머
Kinesis Data Streams	불가	7일	프로듀서-컨슈머
Kinesis Data Firehose	가능	24시간	소스-데스티네이션