

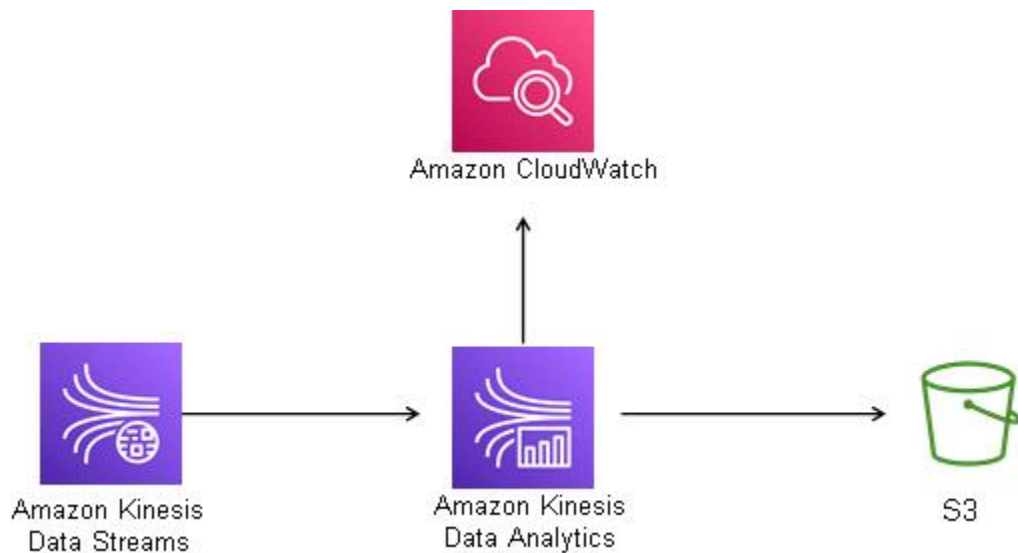
# 2022년도 지방기능경기대회 과제

직 종 명	클라우드컴퓨팅	과 제 명	Automation	과제번호	제2과제
경기시간	4시간	비 번 호		심사위원 확 인	(인)

## 1. 요구사항

클라우드 플랫폼은 구현 시 사용자 로그 파일, 전자 상거래 구매 내역, 게임 플레이 활동, 소셜 네트워크의 정보 등 수많은 스트리밍 데이터들을 처리해야 합니다. AWS 상에는 이러한 스트리밍 정보를 처리하기 위한 Amazon Kinesis Data Analytics(이하 KDA)라는 서비스가 있습니다. KDA는 Apache Flink를 통해 실시간으로 스트리밍 데이터를 변환하고 분석할 수 있는 가장 쉬운 방법입니다. Flink는 데이터 스트림 처리를 위한 오픈소스 프레임워크 및 엔진입니다. KDA와 Flink를 이용해 Kinesis로부터 데이터를 받아 처리하는 Architecture를 구현해야 합니다. 구현해야 할 Cloud Architecture는 다음과 같습니다.

### 다이어그램



### Software Stack(AWS)

- Kinesis data stream
- Kinesis data analytics (Flink)
- S3
- Cloud Watch

## 2. 선수 유의사항

- 1) 경기 시작 전 가벼운 스트레칭 등으로 긴장을 풀어주시고, 작업도구의 사용 시 안전에 주의하십시오.
- 2) 경기 중 완전하게 공개된 자료만 인터넷에서 열람하여 참고할 수 있습니다.  
(공개된 자료가 아니면, 실격으로 처리될 수 있습니다.)
- 3) 선수의 계정에는 비용제한이 존재하며, 이보다 더 높게 과금될 시 계정 사용이 불가능할 수 있습니다.
- 4) 문제에서 특별하게 지정되지 않은 옵션 또는 설정, 값은 시스템이 추천하는 기본값을 사용하며, 대소문자에 주의하여 주시기 바랍니다.
- 5) 과제에서 제공된 AWS resource 외에 다른 resource가 없어야 합니다.
- 6) Region은 버지니아 북부(us-east-1)을 사용합니다.
- 7) 과제 종료 시 진행 중인 테스트를 모두 종료하여 서버에 부하가 발생하지 않도록 합니다.
- 8) 채점을 위해, AdministratorAccess 권한을 가지면서 Console Access, Programmatic Access가 가능한 IAM User를 생성하고, 해당 IAM User의 ID, Password, Access Key, Secret Access Key를 제공해야 합니다.
- 9) 선수는 자신의 채점이 완료될 때까지 외부와 접촉할 수 없으며, 퇴실할 수 없습니다.  
대신 경기 시간 내에 과제가 종료된 선수는 채점 후 퇴실할 수 있습니다.

### 3. 기본설계

- 1) Kinesis data stream을 생성합니다.
  - 이름은 “kda\_flink\_kinesis\_stream” 으로 생성합니다.
- 2) S3를 생성합니다.
  - 이 과정에 생성된 Bucket name은 KDA 에 runtime properties에 입력해야 합니다.
- 3) 생성된 S3 bucket에 “/kda\_flink\_jar/” 위치로 첨부된 jar(kinesis-data-analytics-flink.jar)를 upload합니다.
- 4) IAM을 생성합니다.
  - IAM 이름은 “kda\_flink\_role” 로 생성합니다.
  - IAM policy는 cloudwatch, cloudwatch\_log, Kinesis data streaming, S3에 대한 접근 권한을 갖습니다. (cloud watch 관련 policy 첨부 json file을 참조합니다.)
  - IAM은 KDA에 할당할 수 있어야 합니다.
- 5) 생성된 IAM을 이용하여 Kinesis data analytics application을 생성합니다.
  - Flink 선택합니다. (version 1.13)
  - Application 이름은 “kda\_flink\_application” 로 생성합니다.
- 6) 생성된 KDA Application 구성 설정을 합니다.
  - Bucket 설정은 위에서 생성한 S3 bucket으로 설정합니다.
  - 객체 설정은 Upload 한 jar file 위치로 path를 설정합니다.
  - Runtime properties 추가하고, Group ID는 “FlinkAppProperties” 로 입력합니다.

Key	value
s3_output_path	s3a://<bucket_name>/kda_flink_output

- 7) 설정 완료 후 KDA를 실행시킵니다.
- 8) 실행된 KDA는 kinesis data stream을 통해 data를 받아 s3에 데이터를 저장합니다.
- 9) 실행된 KDA는 Cloud watch에 log를 남깁니다.
- 10) 생성한 S3 bucket name을 제출합니다.

S3 Bucket name	
----------------	--

- 11) 첨부
  - Kinesis data analytics Application : kinesis-data-analytics-flink.jar
  - Cloud Watch IAM policy : flink\_iam\_policy\_cloudwatch.json