

7주차 2강

# 중간고사 대체 과제 데모 및 발표

(카카오엔터프라이즈 CDC Tutorial 및 기타)

오영석

2023년 1학기



## Demo 기능 리스트 (1/2)

- 1. (필수) 카카오엔터프라이즈 CDC Tutorial 단계별 데모 및 발표
  - MySQL to MySQL CDC 파일럿
  - Straw Format은 카카오엔터프라이즈 자체 데이터 Format으로 필수 진행사항은 아니고, 추가적으로 자신이 Straw JSON 파일과 유사한 개발 언어를 사용하여 **파일이나, DB에 해당 내용을 저장하는 경우에는 추가 Advantage를 부여한다.**
  - **기본 : MySQL(source) → MySQL(Target)**
  - **(선택사항) : Source DB : 1 개 → Target DB : 1개 이상**
- 2. (필수) 자신이 만든 custom table을 Source에 생성하고, target table에 데이터 변경사항을 Capture하는 부분을 데모하고 발표
  - CDC의 전체적인 흐름에 대해서 개념을 이해하고, 변화되는 부분에 대해서 데모를 하고, 발표한다.
  - CDC의 흐름별 데이터의 변화 상황을 설명 → 변화를 보여주는 데모 또는 해당 부분을 프로그램을 통해서 구현(프로그램 구현시 Advantage를 부여)
  - 파일럿의 프로그램을 수행하시면 해당 부분을 확인할 수 있습니다.

## Demo 기능 리스트 (2/2)

- 3. (필수) 카카오엔터프라이즈 인프라 사용을 local PC나 다른 인프라로 변경하고, CDC 관련 프로그램은 동작하도록 수정한다.
  - 기본 : 인프라(local or cloud or 학교 플랫폼) + 카카오 CDC 프로그램 수행
  - (선택사항) : 기본적인 파일럿 CDC flow가 대응이 안되는 경우, 수업시간에 진행했던 기본 아키텍처로 구성하여 파일럿과 유사하게 구성한다.
    - Source DB + Debezium connector + Kafka + sink connector + Target DB
    - DBMS는 Any 가능합니다.
    - Kafka는 분산 모드로 수행할 것
- 4. (Advantage) 필수사항외의 사항에 대한 추가 점수 부여 항목
  - Target DB가 1개 이상인 경우
  - 파일럿 프로그램 외에 Kafka Connect API를 통한 CDC 기능 구현
  - 프로그램 언어 : Any 가능

## 과제 발표

- 과제 발표일 : 4월 24일(월) or 4월 26일(수)
- 1회차에 전체 인원 발표 진행
- 발표는 3분이내로
- 발표 이후 질문 또는 피드백
- 개인별로 발표
  - 기본 : 노션 기반 발표
  - 발표 사항 : 소스 사항의 데모를 진행
- 발표 순서 : 4월 20일 공지사항 참조

## 추가사항

- 1. (기말 발표) Kafka streams의 실시간 streaming data에 대한 데이터 처리 및 저장에 대한 데모 및 발표 (9주차)
  - 필수 요건 : Kafka의 REST API를 활용
  - 개발 언어는 Any
  - CLI 단계에서의 데모는 2번과 유사하지만, 점수를 부여함. Try에 대한 점수
- 2.(팀프로젝트) Redpanda를 통한 구현
  - Kafka의 CDC 구현을 Redpanda로 구현하는 것은 Advantage를 부여함.
  - 팀 프로젝트는 팀이 같이 진행하지만, 각 팀원별로 특정 부분을 SQL 부분을 개별적으로 구성할것