



[Snowflake] 4-4. Continuous Data Processing



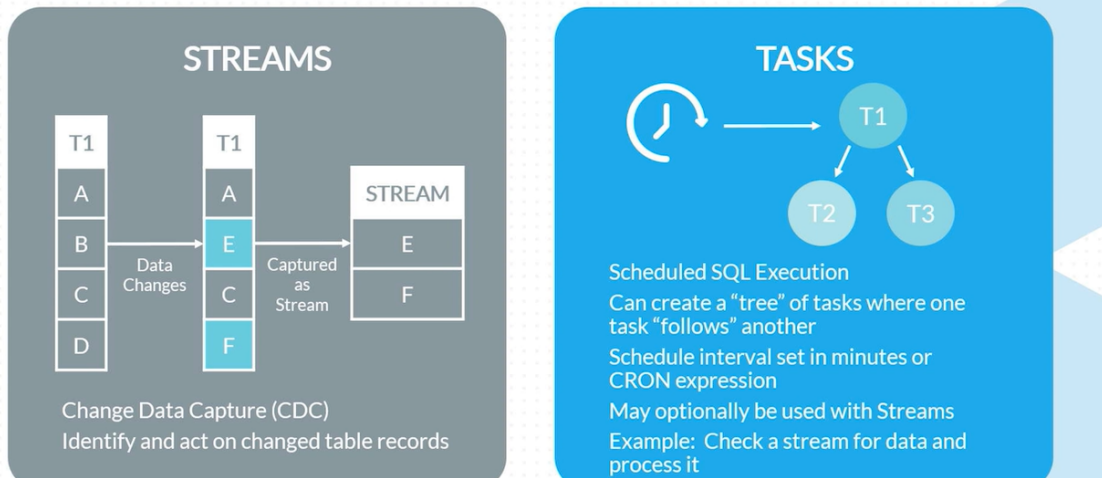
[노션 웹 공유 링크 \(댓글 & 상세설명 참고\)](#)

References

- [Snowflake Learn \(SnowPro PREP-CORE Course\)_4장 5강](#)
- [Snowflake 설명서 \(스트림\)](#)
- [Snowflake 설명서 \(작업\)](#)

-
- 스트림 (Stream) 과 작업 (Task) 개요

STREAMS AND TASKS



- 스트림 (STREAMS) 은 Snowflake의 변경 데이터 캡처 기능
 - 테이블의 변경된 레코드를 식별하고 처리
 - 작업 (TASKS) 는 스케줄 된 SQL 실행
 - 스케줄로 단일 SQL문 또는 프로시저를 실행할 수 있고, 외부 CRON을 사용할 필요 없음
 - 두 기능은 독립적인 Snowflake 오브젝트지만, 같이 사용하면 좋음
-
- 스트림 (Stream)
 - 테이블의 변경 데이터 캡처기능으로 데이터 변경 시 (DML) 그 시점의 오프셋을 저장, 일종의 데이터 북마크 기능이자 오프셋 저장소 역할
 - 오프셋 (Offset) 이란 스트림이 생성 될 때 시점을 의미
 - 스트림은 원본 오브젝트에 대한 오프셋만 저장하며 실제 데이터는 저장하지 않음
 - 스트림 생성 시 (CREATE STREAM) 원본 테이블에 한쌍의 숨겨진 열이 추가되고 변경내용 추적 메타데이터가 저장되기 시작
 - 원본 오브젝트에 대한 버전 관리 내역을 활용하여 CDC 레코드를 반환
 - 스트림 사용과 DDL

WORKING WITH STREAMS AND STREAMS DDL

WORKING WITH STREAMS

Streams hold no data
Think of it like a bookmark
Standard and Append-only
Stream Columns:

METADATA\$ACTION

METADATA\$ISUPDATE

METADATA\$ROW_ID

STREAMS DDL

CREATE

ALTER

DROP

DESCRIBE

SHOW

GRANT...TO ROLE

REVOKE...FROM ROLE

SHOW GRANTS

- 스트림의 유형은 표준(Standard), Append-Only, Insert-Only 가 있음.
 - Standard : 테이블, 디렉터리 테이블 또는 뷰의 스트림에 지원, 원본 오브젝트에 대한 모든 DML 변경 사항 추적 (삽입, 삭제, 수정)
 - Append-Only : 표준 테이블, 디렉터리 테이블 또는 뷰 스트림에 지원, 행 삽입만 추적
 - Insert-Only : 외부 테이블의 스트림에만 지원되며, 행의 삽입만 추적
- **작업 (TASK)**
 - 단일 SQL 문, 저장 프로시저, Snowflake Scripting 개발자 가이드를 사용한 절차 논리 중 하나를 실행
 - 작업을 스트림과 결합하여 최근에 변경된 테이블의 행을 처리하는 지속적인 ELT 워크 플로를 생성 가능
 - 작업을 독립적으로 사용하여 테이블에 행을 삽입하거나, 병합하여 정기적 보고서 생성 및 정기적 작업 가능, (일종의 Cron, 스케줄링)