

NoOffset Optimization

목적: 이 강의에서는 MySQL 쿼리에서 **LIMIT** 절에 **Offset** 을 사용하지 않고 성능을 최적화하는 방법을 배웁니다.

포인트: **Offset** 사용의 문제점과 이를 회피하여 데이터를 효율적으로 조회하는 방법을 이해합니다.

▼ Problems with Using Offset

성능 저하 원인: **Offset** 을 사용하면 지정된 수만큼의 데이터를 읽고 버려야 하므로 성능이 저하됨.

```
SELECT * FROM blog_posts ORDER BY post_id LIMIT 100000, 10
```

▼ NoOffset Optimization Strategies

Offset 없이 조회하는 법: 마지막으로 처리된 레코드의 키를 기준으로 다음 데이터를 조회

배치 시스템을 예시: 현재 Chunk 를 처리하고 다음 번 Chunk 를 읽어오는 상황

- LIMIT 절에 Offset 을 사용하면 간단히 이를 구현할 수 있음.
- 그러나 Offset 을 사용하지 않고도 가능하다: 마지막으로 처리된 레코드의 키를 사용해서 다음 데이터를 읽으면 되니까.

```
# 첫 번째 Chunk 읽기 작업: 마지막 Id 값은 15800 이라고 가정
SELECT * data FROM my_table
WHERE status = 'pending'
ORDER BY id
LIMIT 10000
```

```
# 다음 번 Chunk 읽기 작업
SELECT * data FROM my_table
WHERE id > 15800
ORDER BY id
LIMIT 10000;
```

▼ Praticce

Task 1: Reading Data Using 'NoOffset'

목적: Offset 방식과 NoOffset 방식을 이용한 읽기 성능 비교

시나리오: 사용자 활동 로그 정보를 읽어와서 처리하는 배치 시스템에서 데이터를 읽어 오는 부분

테이블: `user_activity_logs`

- `id` : 기본 키
- `user_id` : 사용자 식별키
- `activity_type` : 활동 로그 타입
- `activity_timestamp` : 로그 타임스탬프
- `additional_info` : 추가정보

1. 사용자 활동 로그 테이블 설계

```
create table user_activity_logs
(
    id                int auto_increment
        primary key,
    user_id           int          null,
    activity_type      varchar(50) null,
    activity_timestamp datetime    null,
    additional_info    text        null
);
```

2. 테스트 데이터 삽입

3. Offset 방식으로 읽어보기:

```
SELECT *  
FROM user_activity_logs  
WHERE activity_type = 'click'  
ORDER BY id  
LIMIT 500;
```

```
SELECT *  
FROM user_activity_logs  
WHERE activity_type = 'click'  
ORDER BY id  
LIMIT 500, 500;
```

4. Offset 방식의 쿼리 소요 시간 확인해보기:

```
EXPLAIN ANALYZE SELECT *  
FROM user_activity_logs  
WHERE activity_type = 'click'  
ORDER BY id  
LIMIT 500, 500;
```

5. NoOffset 방식으로 읽어보기:

```
SELECT *  
FROM user_activity_logs  
WHERE id > 0 AND activity_type = 'click'  
ORDER BY id  
LIMIT 500;
```

```
SELECT *  
FROM user_activity_logs  
WHERE id > 2029 AND activity_type = 'click'  
ORDER BY id  
LIMIT 500;
```

6. NoOffset 방식의 쿼리 소요 시간 확인해보기:

```
EXPLAIN ANALYZE SELECT *  
FROM user_activity_logs  
WHERE id > 2029 AND activity_type = 'click'  
ORDER BY id  
LIMIT 500;
```