

## Exercise – Query Optimisation

A video rental store database contains the following tables:

Videos(Videoid, Title, Category, Status)

Rentals(RentalId, Videoid, MemberId, RentDate, NoNights, ReturnDate, Fine)

The table **Videos** contains 2000 records. 250 of these videos have a category 'NR' (New Release). The table Rentals contains 500 records. There are presently 200 videos out on hire, 125 of which are for new Releases, 20 of which have not yet been returned (ie. ReturnDate is NULL).

The store manager wishes to know the Videoid, Title and DateRented for all New releases presently on hire.

- (i) Give an SQL query to return the required information.
- (ii) Explain two methods (execution plans) by which this query may be executed.
- (iii) Determine the optimal method based on the above information.

## Solution

### (i) SQL Statements

```
SELECT R.VideoId, Title, RentDate
FROM Rentals R, Videos V
WHERE R.VideoId = V.VideoId AND
      Category = 'NR' AND
      ReturnDate IS NULL;
```

### (ii) Execution methods

#### Method1:

1. Extract all videos with category 'NR'
2. Join result with Rentals table
3. Extract all records with ReturnDate equal to NULL

#### Method 2:

1. Extract all rentals with ReturnDate equal to NULL
2. Join result with Videos
3. Extract all rentals with category 'NR'

### (iii) Optimisation

#### Method1:

1. Extract all videos with category 'NR'  
2000 reads (each record in the videos table)  
250 Writes (each of the 250 with category 'NR')
2. Join result with Rentals table  
 $250 \times 500 = 125,000$  reads (each of the 250 matched with each of the 500)  
125 writes (each of the 125 rental records for 'NR')
3. Extract all records with ReturnDate equal to NULL  
125 reads (all rental records for NR, even those returned)  
20 writes (20 have not yet been returned)

**TOTAL I/O's: 127,520 I/O's**

#### Method2:

1. Extract all rentals with ReturnDate equal to NULL  
500 reads (each record in the rentals table)  
200 Writes (each of the 200 out on hire)
2. Join result with Videos table  
 $200 \times 2000 = 400,000$  reads (each of the 200 matched with each of the 2000)

Query optimisation could stop at this point as method 2 has now exceeded the total number if I/O's performed by method1 → Method 1 is most optimal execution plan.

200 writes (each of the 200 rental records out on hire)

3. Extract all records with category = 'NR'
  - 200 reads (all rental records out on hire)
  - 20 writes (20 have not yet been returned)

**TOTAL I/O's: 401,120 I/O's**