

Query Processing

Monday, 22 September 2025 22:44

Diketahui relasi dalam sistem berdata relasional sebuah tabel penyeleksi buku sebagai berikut (yang digaris bawahi adalah primary key):
 Pelanggan (idpl, nama, alamat)
 Buku = (bkd, judul, penulis)
 Peminjaman = (bkd, bkd, tanggal)
 FK: Peminjaman (pid) → Pelanggan (pid);
 Peminjaman (bkd) → Buku (bkd)

Assumsi:

- Relasi Pelanggan terdiri atas 10000 tuple yang tersimpan dalam 1000 blok.
- Relasi Buku terdiri atas 5000 tuple yang tersimpan dalam 500 blok.
- Relasi Peminjaman terdiri atas 30000 tuple yang tersimpan dalam 15000 blok.
- Ada 500 orang penulis buku berbeda
- Usia pelanggan bervariasi antara 7 s.d. 74 tahun.
- Distribusi usia dan jumlah pelanggan perusia dan perusia.

Diketahui ukuran tuple hasil join antara relasi Buku dan Peminjaman adalah 2 kali ukuran tuple relasi Peminjaman.

$\Pi_{\text{Nama}} ((\sigma_{\text{Umur} \geq 21}(\text{Pelanggan})) \bowtie ((\sigma_{\text{Penulis} = \text{'Andrea Hirata'}}(\text{Buku})) \bowtie \text{Peminjaman}))$

Tujuan sebuah evaluasi plan query tersebut dan estimasi cost untuk memperbaiki plan tersebut. Estimasi pada jumlah tuple yang akan dihasilkan.

Format penilaian jumlah tuple tabel std:

Operasi	Cost	Banyaknya tuple
...
Total (banyaknya tuple final)

Assume kiri:

• Mekanisme evaluasi ekspresi adalah pipeline (secara distribusional memanfaatkan).

• Assumsi worst case scenario untuk operasi join jika dilakukan, yaitu hanya terdapat memori cukup menampung 1 blok dari setiap relasi.

① seleksi (σ)

a) $\sigma_{\text{Umur} \geq 21}(\text{pelanggan})$

$$\text{fraksi} = \frac{\text{jumlah umur} \geq 21}{\text{jumlah umur}} = \frac{(24-21+1)}{(24-7+1)} = \frac{1}{18} = 0,2222$$

estimasi hasil tuple = fraksi \times tuple pelanggan = $0,2222 \times 10.000 = 2.222$

estimasi jumlah blok = fraksi \times blok = $0,2222 \times 1000 = 222$

query melakukakan linear search sehingga cost = 1000 block transfers + 1 seek

b) $\sigma_{\text{Penulis} = \text{'Andrea Hirata'}}(\text{Buku})$

$$\text{fraksi} = \frac{\text{penulis}}{\text{tersedia}} = \frac{1}{500} = 0,002$$

estimasi hasil tuple = $0,002 \times 50.000 : 100$

estimasi jumlah blok = $0,002 \times 5.000 : 10$

query melakukakan linear search sehingga cost = 5000 block transfers + 1 seek

② Join buku dan peminjaman ($\sigma_{\text{Penulis} = \text{'Andrea Hirata'}}(\text{Buku}) \bowtie \text{Peminjaman}$)

karena buku Andrea Hirata ada di 100 tuple dalam 10 blok dan peminjaman \rightarrow setiap join antara Buku & Peminjaman menghasilkan tuple Peminjaman, maka ukuran tuple baru adalah 2 kali ukuran tuple Peminjaman

blok = 10 blok

bs = 15.000 blok

cost = br \times bs + br block transfer + 2 \times br seeks

$$= 10 \times 15.000 + 10 + 2 \times 10$$

= 150.010 block transfers + 20 seeks

③ Join dengan pelanggan ($(\sigma_{\text{Umur} \geq 21}(\text{Pelanggan})) \bowtie ((\sigma_{\text{Penulis} = \text{'Andrea Hirata'}}(\text{Buku})) \bowtie \text{Peminjaman})$)

Karena tuple peminjaman = 60.000 & blok peminjaman = 15.000

tuple baru = 2 kali nya, sehingga :

terdapat 600.000 tuple & 30.000 block

br = 222 block ($\text{pelanggan} \geq 21$)

bs = 300.000 block

cost = br \times bs + br block transfers + 2 \times br seeks

$$= 222 \times 300.000 + 222 + 2 \times 222$$

= 6.660.222 block transfers + 444 seeks

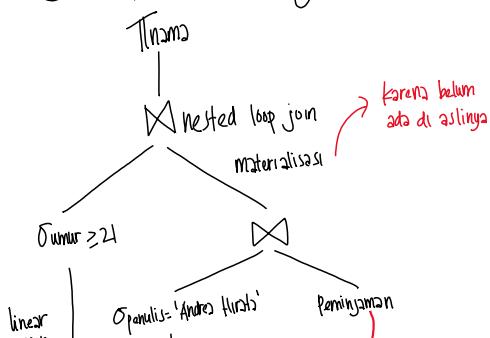
TABEL

operasi	
$\sigma_{\text{Umur} \geq 21}(\text{pelanggan})$	
$\sigma_{\text{Penulis} = \text{'Andrea Hirata'}}(\text{Buku})$	
$\sigma_{\text{Penulis} = \text{'Andrea Hirata'}}(\text{Buku}) \bowtie \text{Peminjaman}$	
$\sigma_{\text{Umur} \geq 21}(\text{pelanggan}) \bowtie (\sigma_{\text{Penulis} = \text{'Andrea Hirata'}}(\text{Buku}) \bowtie \text{Peminjaman})$	

cost	tuple
1000 block transfers + 1 seek	2.222
5000 block transfers + 1 seek	100
150.010 block transfers + 20 seeks	600.000
6.660.222 block transfers + 444 seeks	6.660.222
6.816.232 block transfers + 462 seeks	600.000

(1) Step = bikin query tree

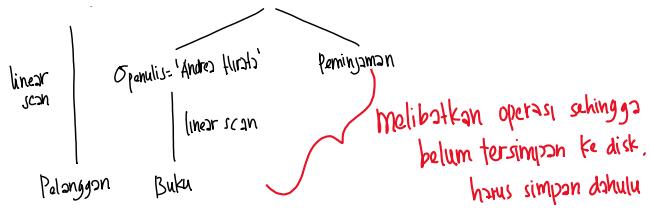
TOTAL



banyak tuple
ukuran tuple \times bs

$$N_B < 360.000 \\ 300.000 - 600$$

$(br) + nr \times bs$
↓
Karena ukuran tuple
di memory
nilanya 0
pipeline nilai juga 0



Latihan Alternatif Solusi (lanjutan)		
Operasi	Cost (#block transfer)	Banyaknya tuple
$\sigma_{\text{umur} \geq 21}(\text{Pelanggan})$	$D_{\text{Pelanggan}} = 1000$	$n_p = 4/18 * 10000$ $= 2222$ karena hanya melibatkan data Pelanggan umur ≥ 21
$C \bowtie B$ Asumsi: nested loop join	$n_C * b_B = 2222 * 60$ $= 133320$	$n_B < 600 = 4/18 * 600 = 133$ pid adalah key pada Pelanggan di C. 4/18 karena hanya melibatkan data Pelanggan umur ≥ 21
Π_{Nama} (D)	0	$n_D = n_B = 133$
Total	$= 5000 + 1500060 + 1000 + 133320$ $= 1639380$	

Operasi	Cost (#block transfer)	Banyaknya tuple
$\sigma_{\text{penulis} = \text{'Andrea Hirata'}}(\text{Buku})$	$D_{\text{Buku}} = 5000$	$n_B = 50000/500 = 100$
Linear scan Hasil operasi = A		
$A \bowtie D$ Asumsi: nested loop join	$n_A * b_D + b_B$ $= 100 * 15000 + 60$ $= 150060$	$n_Y < 300000$ $= 300000 / 500$ $= 600$
Hasil perlu dimaterialisasi karena akan digunakan sbg inner relation ultk operasi berikutnya: b_2 adalah # block yang ditulis ke disk.	$t_B = 10$ $t_D = 15$ $t_B * t_D = 150$	jadi # tuple hasil operasi <# tuple Peminjaman: 300000/500 karena A hanya melibatkan penulis = 'Andrea Hirata' dan ada 500 penulis berbeda
Hasil operasi = B	$b_B = n_A * t_B = 1500 * 10 = 15000$	

