

Query Processing

Monday, 22 September 2025 22:44

Diketahui relasi dalam skema basisdata relational sebuah tempat penjemaran buku sebagai berikut (yang digaris bawah adalah primary key).

Pelanggan = (pid, nama, umur, alamat)
Buku = (bid, judul, penulis)
Peminjaman = (pid, bid, tanggal)
FK: Peminjaman (pid) → Pelanggan (pid);
Peminjaman (bid) → Buku (bid)

Asumsi lain:

- Relasi Pelanggan terdiri atas 10000 tuple yang tersimpan dalam 1000 blok.
- Relasi Buku terdiri atas 50000 tuple yang tersimpan dalam 5000 blok.
- Relasi Peminjaman terdiri atas 300000 tuple yang tersimpan dalam 15000 blok.
- Ada 500 orang pembaca buku berbeda.
- Umur pelanggan bervariasi antara 7 s.d. 74 tahun.
- Distribusi teragam (uniform) pada nilai atribut penulis dan umur.

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

Diketahui query dalam aljabar relational sebagai berikut:

II. nama (($\sigma_{\text{umur} \geq 21}(\text{Pelanggan})$)) \bowtie (($\sigma_{\text{penulis} = \text{'Andrea Hirata'}(\text{Buku})$)) \bowtie Peminjaman))

Tentukan sebuah evaluasi plan untuk query tersebut dan estimasikan cost untuk menjalankan plan tersebut. Estimasi pada jumlah tuple yang akan dihasilkan.

Format jawaban sebagai berikut tabel di bawah.

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

Asumsi lain:

- Metode evaluasi operasi adalah pipeline ber, dibutuhkan materialisasi.
- Asumsikan worst case scenario untuk operasi join jika dibutuhkan, yaitu hanya tersedia memori cukup untuk menampung 1 blok dari setiap relasi.

① seleksi (σ)

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

$$\text{fraksi} = \frac{\text{jumlah umur} \geq 21 \text{ tahun}}{\text{jumlah nilai umur}} = \frac{(29-21+1)}{(29-7+1)} = \frac{9}{18} = 0,2222$$

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

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

query melakukan linear search sehingga cost : 1000 block transfers + 1 seek

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

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

$$\text{estimasi hasil tuple} = 0,002 \times 50.000 = 100$$

$$\text{estimasi jumlah blok} = 0,002 \times 5.000 = 10$$

query melakukan linear search sehingga cost : 5000 block transfers + 1 seek

② Join buku & 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 tuple Peminjaman, maka ukuran tuple baru adalah 2 kali ukuran tuple Peminjaman

br = 10 blok

bs = 15.000 blok

$$\text{cost} = \text{br} \times \text{bs} + \text{br} \text{ block transfers} + 2 \times \text{br} \text{ seeks}$$

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

$$= 150.010 \text{ block transfers} + 20 \text{ seeks}$$

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 : 600.000 & blok peminjaman : 15.000

tuple baru : 2 kalinya, sehingga :

terdapat 600.000 tuple & 30.000 blok

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

$$\text{bs} = 300.000 \text{ blok}$$

$$\text{cost} = \text{br} \times \text{bs} + \text{br} \text{ block transfers} + 2 \times \text{br} \text{ seeks}$$

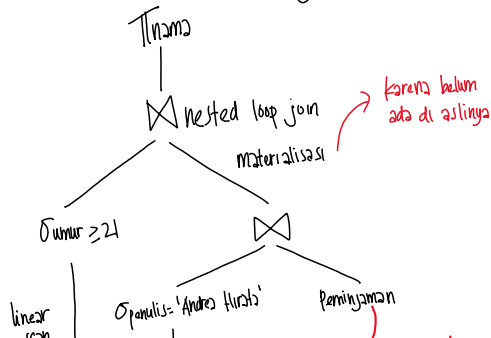
$$= 222 \times 300.000 + 222 + 2 \times 222$$

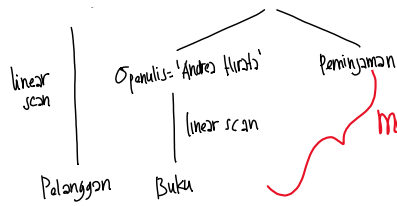
$$= 6.660.222 \text{ block transfers} + 444 \text{ seeks}$$

TABEL

operasi	Cost	tuple
$\sigma_{\text{umur} \geq 21}(\text{pelanggan})$	1000 block transfers + 1 seek	2.222
$\sigma_{\text{penulis} = \text{'Andrea Hirata'}(\text{Buku})$	5000 block transfers + 1 seek	100
$\sigma_{\text{penulis} = \text{'Andrea Hirata'}(\text{Buku}) \bowtie \text{Peminjaman}$	150.010 block transfers + 20 seeks	600.000
$\sigma_{\text{umur} \geq 21}(\text{pelanggan}) \bowtie (\sigma_{\text{penulis} = \text{'Andrea Hirata'}(\text{Buku}) \bowtie \text{Peminjaman})$	6.660.222 block transfers + 444 seeks	600.000
TOTAL	6.816.232 block transfers + 462 seeks	600.000

① step = bikin query tree





Melibatkan operasi sehingga belum tersimpan ke disk, harus simpan dahulu

Latihan Alternatif Solusi (lanjutan)

Operasi	Cost (#block transfer)	Banyaknya tuple
$\sigma_{umur \geq 21}$ (Pelanggan) Linear scan Hasil operasi = C	$b_{Pelanggan} = 1000$	$n_C = 4/18 * 10000 = 2222$ karena hanya melibatkan data Pelanggan umur ≥ 21
C \bowtie B Asumsi: nested loop join Hasil operasi = D	$n_C * b_B = 2222 * 60 = 133320$	$n_D < 600 = 4/18 * 600 = 133$ pid adalah key pada Pelanggan di C: 4/18 karena hanya melibatkan data Pelanggan umur ≥ 21 $n_{Asumsi} = n_D = 133$
Π_{nama} (D)	0 (karena pipeline)	
Total	$= 5000 + 1500060 + 1000 + 133320 = 1639380$	150

Latihan Alternatif Solusi (lanjutan)

Operasi	Cost (#block transfer)	Banyaknya tuple
$\sigma_{penulis = 'Andres Hirata'}$ (Buku) Linear scan Hasil operasi = A A \bowtie Peminjaman Asumsi: nested loop join Hasil operasi = B	$b_{Buku} = 5000$ $n_A * b_{Peminjaman} + b_B = 100 * 15000 + 60 = 1500060$ Hasil perlu dimaterialisasi karena akan digunakan sbg inner relation utk operasi berikutnya. b_B adalah # blok yang ditulis ke disk. $n_{Peminjaman} = 20$ $b_A = 1/6 n_{Peminjaman} = 10$ (karena ukuran tuple hasil join 2 x ukuran tuple peminjaman) $b_B = n_B/b_B = 600/10 = 60$	$n_A = 50000/500 = 100$ $n_B < 300000 = 300000 / 500 = 600$ bid adalah key pada Buku jadi # tuple hasil operasi $< \#$ tuple Peminjaman: 300000/500 karena A hanya melibatkan penulis = 'Andres Hirata' dan ada 500 penulis berbeda

