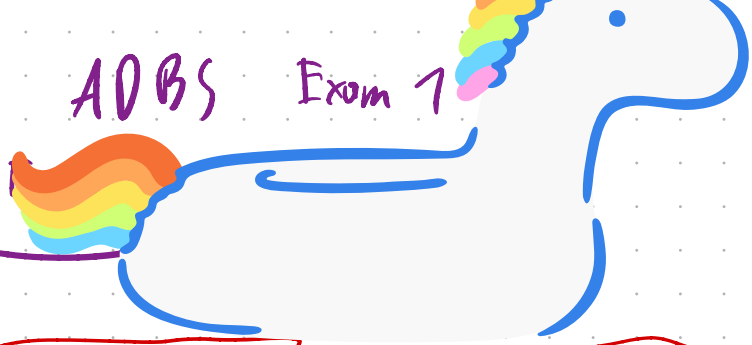
Cheat-Sheet ADBS Exam 1



I/O-Times for HDD

t_s ... seek time

t_r ... rotation delay

↳ $0.5 \cdot \frac{1\,min}{rpm}$ or $0.5 \cdot \frac{60\,s}{rpm}$

ie $\frac{30}{7500} = 2.00\,ms$

t_{tr} ... transfer (rate) time

↳ $\frac{1}{tr} * \text{block size}$ ie $\frac{8\,kB}{163000\,kB/s} \approx 0.05\,ms$

random access time

↳ $t_s + t_r + t_{tr} = t_a$

$t = t_s + t_r + nr.blocks \cdot t_{tr} + (nr.blocks - 1) \cdot t_{st}$

for sequential

$t = nr.blocks \cdot t_a$

for random access

→ p% by $\frac{t_{seq}}{t_{random}}$ when it's better to seq read whole file

I/O-Times for SSD

t_a = access time for one page $= \frac{\text{block size}}{tr}$ ie $\frac{128\,kB}{170000\,kB/s} = 0.73\,ms$

$t = t_a \cdot nr.of.blocks$ for random access

$t = \frac{\lceil nr.of.blocks \cdot page\text{-}size \rceil}{block\ size} \cdot t_a$ ie $\frac{\lceil 1000 \cdot 8\,kB \rceil}{128\,kB} \cdot 0.73\,ms$

for seq read

Blocking Factor

$bfr = \left\lfloor \frac{B}{R} \right\rfloor$ B → Block size, R → Record size

Sort cost

b = # pages

$2 \cdot b \cdot (1 + \lceil \log_2(b) \rceil)$ for 3 buffers

$2 \cdot b \cdot (1 + \lceil \log_{B-1}(b/B) \rceil)$ for multi-way with B buffers

Cost fn variables

r_X or $|X|$ ... # tuples in rel. X

b_X ... # blocks occupied by X

bfr_X ... blocking factor (# records/block)

sel_p ... 0..1 - selectivity of predicate p

S_p ... cardinality of p ($sel_p * r_X$)

Search cost:

b ... for non unique

$b/2$ ... for unique

$\log_2 b + \left(\lceil \frac{sel \cdot r}{bfr} \rceil - 1\right)$ ... for bin search of indexed file

x or $x+1$ ... for B+ search where x is depth

1 ... for hash search

Join cost:

$b_R + (b_R \cdot b_S)$ ... for nested loops

$b_R \cdot \lceil \frac{b_R}{B-2} \rceil * b_S$ ... for block nloops

idly nested loops

$b_R + (|R| \cdot x)$ for dense on impl. of index B+ idxnl

$b_R + b_S$ for sortmerge join if one relation is unique otherwise might be quadratic

$2(b_R + b_S) + (b_R + b_S)$ for hash join

$(3 - 2 \cdot \frac{k}{B-1})(b_R + b_S)$ ... for hybrid hash

k: r/v

$b_R + b_S$ for nhj if R fits into buffer entirely

Selectivity

$sel_P = \frac{|\sigma_P(R)|}{|R|}$ ... for predicate P

$sel_{R \bowtie S} = \frac{|R \bowtie S|}{|R \times S|} = \frac{|R \bowtie S|}{|R| \cdot |S|}$ for join

$sel_{R.A=c} = \frac{1}{|R|}$ for unique attribute c

$sel_{\bowtie R.A=S.B} = \frac{1}{|S|}$ ... for eq-join via foreign key and R.A not null

$sel_{R.A=c} = \frac{1}{NDV(A,R)}$ for non unique attribute A

NDV(A,R) ... # diff values of A in R

$sel_{R.A>c} = \frac{High(A,R) - c}{High(A,R) - Low(A,R)}$ for $c \le High(A,R)$ where High(A,R) is highest value A and Low(A,R) is lowest

$sel_{\bowtie R.A=S.B} = \frac{1}{\max(NDV(A,R), NDV(B,S))}$ for all equi-joins

for $A \le c$: $\frac{n-1}{|buckets|} + \frac{1}{|buckets|} \cdot r$

$r = 1 - \frac{max - c}{max - min}$

in Bucket if max = 500, min = 400

for $A > c$: $\frac{|buckets| - n}{|buckets|} + \frac{1}{|buckets|} \cdot r$

$r = \frac{max - c}{max - min}$

Rel Algebra

De'morgan: $R \cap S = R - (R - S)$

$((\pi_B(R) \cap \pi_B(S)) \bowtie T = (\pi_B(R) \bowtie T) \bowtie (\pi_B(S) \bowtie T)$

for $R(AB), S(AB), T(BC)$

$(R \bowtie S) \bowtie T = (R \bowtie S) \cap (R \bowtie T)$

for $R(AB), S(AC), T(BC)$

⊗ $\pi_c(U - V) = \pi_c(U) - \pi_c(V)$

no: $U(CD), V(CD)$

$\pi_{AB}(R \times S) = \pi_A(R) \times \pi_B(S)$

for $R(A \ldots), S(B \ldots)$

⊗ $\pi_A(R \cap S) = \pi_A(R) \cap \pi_A(S)$

for $R(AB), S(AB)$