2.1 merge Algorithm

(a).

$N = 24.000,000$

$B = 60$

Pass 0: $\frac{N}{B} = 400,000$

Merge Passes.

$B-1=59$

$\log_{59} 400,000 = 3.16$

$\Rightarrow$ we need 4.

$4+1=5 \Rightarrow$ 5 passes needed.

(2). $N=24,000,000$

$B=60$

$k=B-1=59$

Pass: 0

$\frac{N}{B} = 400,000$

Pass 1:

$\lceil \frac{400,000}{59} \rceil \approx 6780$

Pass 2:

$\lceil \frac{6780}{59} \rceil \approx 115$

Pass 3

$\lceil \frac{115}{59} \rceil = 2$

Pass 4

$\lceil \frac{2}{59} \rceil = 1$

$\Rightarrow \Sigma = 400,000 + 6780 + 115 + 2 = 406897$

(3).

$B=60$

$k=B-1=59$

for Pass 2:

$60 \times 59 \times 59 = 208860$

(4) $N=24,000,000$

$B=120$

Pass 0: $\frac{N}{B} = 200,000$

$k=B-1=119$

$\therefore \log_{119}(200,000) = 2.55$

$\therefore$ need 3

$3+1=4$

$\Rightarrow$ 4 Passes

$\therefore$ I/O cost $= 4 \times 2 \times 24,000,000 = 192,000,000$

(5). ∵ only 3 passes

$\therefore \log_{B-1}(\frac{N}{B}) \leq 2$

$\therefore \log_{B-1}(\frac{24,000,000}{B}) \leq 2$

$\Rightarrow \frac{24,000,000}{B} \leq (B-1)^2$

$\therefore B \cdot (B-1)^2 \geq 24,000,000$

$\therefore B \geq 290$

$\therefore$ smallest B is 290.

(6). $\log_{24}(\frac{N}{25}) \leq 3$

$\therefore N \leq 24^3 \cdot 25$

$\therefore N \leq 345600$

$\Rightarrow$ 345600 pages.

22. join algorithm:

(a) $n = 4000$, 50 tuple per page

$N = 800$

$T_A = 4000 \times 50 = 200,000$ tuples.

Cost = $M + (T_A \times N) = 160,024,000$ I/Os

(b). $F = 500$

$O = 3000$

$N = 800$

Block-size = $F - 2 = 498$

blocks = $\lceil \frac{O}{\text{Block-size}} \rceil = \lceil \frac{3000}{498} \rceil = 7$

Cost = $O + (\lceil \frac{O}{F-2} \rceil \times N) = 8600$ I/Os.

(c). $N = 800$

$F = 500$

$\lceil \frac{N}{F} \rceil = 2 \Rightarrow$ IO: $2 \times N = 1600$

$k = F - 1 = 499$

$2 < 499$

$\therefore I/O = 2 \times N = 1600$

$\therefore \Sigma = 1600 + 1600 = 3200$

for cost of merge phase in the worst case senario

Cost = $N + O = 3800$

(d). $P = 800$

$O = 3000$

no duplicates in the join attribute

Cost = $P + O = 3800$

(e).

$N = 800$

$O = 3000$

$F = 500$

for Partition Phase.

Cost = $2 \times (N+O) = 7800$

for probe phase

Cost = $N+O = 3800$