# SORTING IN 2 PASSES

ASSUME ① FILE F HAS N PAGES

② # BUFFERS $\geq \sqrt{N} + 1$

PHASE #1    READ F, $\sqrt{N}$ PAGES AT A TIME,

QUICKSORT, WRITE SORTED

RUN OF LENGTH $\sqrt{N}$ PAGES

# I/Os

= 2×N

# SORTED RUNS

= $\sqrt{N}$

REPEAT FOR ALL PAGES IN F

# REPEATS = $\sqrt{N}$

PHASE #2    DO $\sqrt{N}$ WAY MERGE



1   $R_1$

2   $R_2$

$\sqrt{N}$   $R_{\sqrt{N}}$

MERGR →

SORTED F
OF LENGTH
$\sqrt{N} \cdot \sqrt{N} = N$

EACH RUN
$\sqrt{N}$ PAGE

---
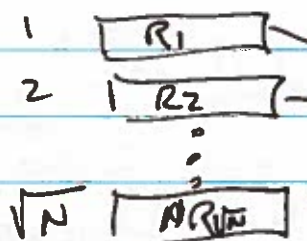
EXAMPLE #1

ASSUME $N = 10^6$ pages    $\dfrac{10^6 \times 10^4 \text{ Bytes}}{\text{pg}} = 10^{10}$ Bytes $= 10\,GB$

$\sqrt{N} = 10^3$ pages

$\dfrac{10^3 \times 10^4 \text{ Bytes}}{\text{pg}} = 10\,MB$

---

EX#2    $\sqrt{N} = 10^5$ pages    $\dfrac{10^5 \times 10^4}{\text{Bytes/pg}} = 10^9 = 1\,GB$

$N = 10^{10}$ pages

$\times 10^4 = 10^{14} = 100 \times 10^{12} = 100\,TB$

Sort merge
JOIN

R     a                    S.b

$\Rightarrow$

| a |
|---|
| 1 |
| 2 |
| 3 |
| 3 |
| 4 |
| 5 |
| 5 |
| 5 |
| 6 |

$\Rightarrow$

| S.b |
|---|
| 0 |
| 2 |
| 3 |
| 3 |
| 3 |
| 4 |
| 5 |
| 5 |
| 7 |
| 7 |

DO {

  DO {

    ADVANCE R SCAN
      UNTIL $R.a \geqslant S.b$

    ADVANCE S SCAN
      UNTIL $S.b \geqslant R.a$

  } UNTIL $R.a = S.b$

  JOIN $\{R_i\}$ WITH $\{S_i\}$ WITH CURRENT $R.a, S.b$ values

} UNTILL END OF R OR END OF S