

UNIDAD 5: ALGORITMOS DE MEMORIA EXTERNA

COTAS DE OPERACIONES FUNDAMENTALES DE ENTRADA Y SALIDA

Gibran Fuentes Pineda

Mayo 2020

- Complejidad E/S de muchos algoritmos depende de las siguiente operaciones
 - *Escaneo*: lectura secuencial de N elementos
 - *Ordenamiento*: acomodar N elementos en un orden dado
 - *Búsqueda*: encontrar elementos específicos en N elementos ordenados
 - *Salida*: regresar los Z elementos de una consulta
- En el modelo de discos paralelos, un algoritmo de memoria interna normalmente realiza $\Omega(N)$ operaciones E/S

Problema	Memoria interna	Memoria externa
Escaneo	$O(N)$	$O\left(\frac{N}{B}\right)$
Ordenamiento	$O(N \log N)$	$O\left(\frac{N}{B} \log_{M/B} \frac{N}{B}\right)$
Búsqueda	$O(\log_2 N)$	$O(\log_B N)$

Operación	$D = 1$	$D \geq 1$
Escaneo	$\Theta(\frac{N}{B}) = \Theta(n)$	$\Theta(\frac{N}{DB}) = \Theta(\frac{n}{D})$
Ordenamiento	$\Theta(\frac{N}{B} \log_{\frac{M}{B}} \frac{N}{B})$ $= \Theta(n \log_m n)$	$\Theta(\frac{N}{DB} \log_{\frac{M}{B}} \frac{N}{B})$ $= \Theta(\frac{n}{D} \log_m n)$
Búsqueda	$\Theta(\log_B N)$	$\Theta(\log_{DB} N)$
Salida (Z)	$\Theta(\max\{1, \frac{Z}{B}\})$ $= \Theta(\max\{1, Z\})$	$\Theta(\max\{1, \frac{Z}{B}\})$ $= \Theta(\max\{1, \frac{Z}{D}\})$

DIFERENCIAS ENTRE N Y N/B

- Para escanear un arreglo de N elementos en disco, tendríamos una complejidad E/S:
 - $O(N)$, si los elementos contiguos se encuentran en bloques distintos
 - $O(N/B)$, si los elementos contiguos se encuentran en un mismo bloque
- Considera un arreglo de 600,000,000 elementos (N) almacenados en bloques de 6,000 (B) y un tiempo de acceso a disco de 1 milisegundo
 - N operaciones E/S tomarían 600,000 segundos (~166 horas)
 - N/B operaciones E/S tomarían 100 segundos