

Universidad de Carabobo

Facultad Experimental de Ciencias y Tecnologías

José J. Díaz y María Moreno

Arquitectura del Computador

22 de Mayo de 2023



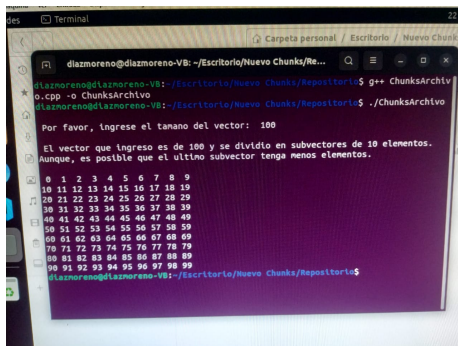
¿Qué es Chunks en Programación?

En programación, un chunk (trozo) se refiere a una porción de datos o un bloque de información relacionada. El concepto de chunking, es importante en programación, ya que, permite dividir grandes cantidades de datos en fragmentos más pequeños y manejables. Esto facilita el trabajo con conjuntos más grandes de información sin abrumarse por el volumen y permite una recuperación y análisis más rápidos cuando sea necesario, al tiempo que reduce la cantidad de memoria necesaria para almacenarlo todo de una vez.

Librerías empleadas:

- **fstream (file stream)**, lee o escribe del mismo fichero en el mismo programa.
- **string**, utiliza cadenas: copiar, cambiar caracteres, comparar cadenas, etc.
- **random**, genera números aleatorio.
- **fcntl**, trabaja con el descriptor de ficheros.
- **unistd**, define constantes, tipos simbólicos/misceláneos y declara funciones misceláneas.
- **sys/stat**, llamada al sistema fstat.
- **iostream**, lectura y escritura en los flujos estándar (cin, cout, cerr y clog, etc).
- **vector**, plantilla de clase para contenedores de secuencia.
- **limits**, esta librería contiene parámetros de entorno y rangos para tipos enteros.
- **sys/mman**, permite declaraciones de gestión de memoria.
- **Ctime**, controla el manejo del tiempo.

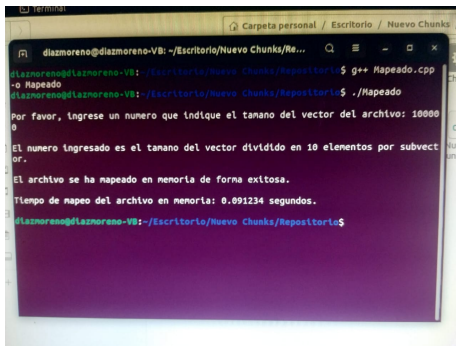
Vector dividido en Chunks



```
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunks/Re...  
diazmoreno@diazmoreno-VB:~/Escritorio/Nuevo Chunks/Repositorio$ g++ ChunksArchiv  
o.cpp -o ChunksArchivo  
diazmoreno@diazmoreno-VB:~/Escritorio/Nuevo Chunks/Repositorio$ ./ChunksArchivo  
  
Por favor, ingrese el tamaño del vector: 100  
  
El vector que ingreso es de 100 y se dividió en subvectores de 10 elementos.  
Aunque, es posible que el último subvector tenga menos elementos.  
  
0 1 2 3 4 5 6 7 8 9  
10 11 12 13 14 15 16 17 18 19  
20 21 22 23 24 25 26 27 28 29  
30 31 32 33 34 35 36 37 38 39  
40 41 42 43 44 45 46 47 48 49  
50 51 52 53 54 55 56 57 58 59  
60 61 62 63 64 65 66 67 68 69  
70 71 72 73 74 75 76 77 78 79  
80 81 82 83 84 85 86 87 88 89  
90 91 92 93 94 95 96 97 98 99  
diazmoreno@diazmoreno-VB:~/Escritorio/Nuevo Chunks/Repositorio$
```

Figure: 1.- Imagen por terminal: división de vector con Chunks de $n=100$

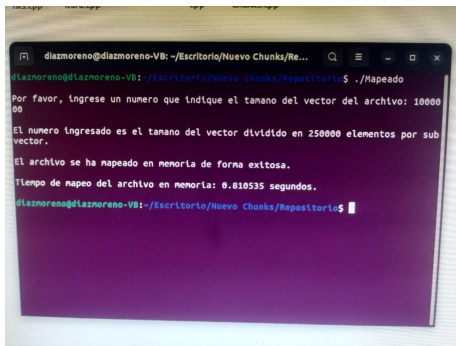
Mapecto de un archivo en memoria utilizando mmap



```
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunks/Re...  
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunks/Repositorio$ g++ Mapeado.cpp  
-o Mapeado  
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunks/Repositorio$ ./Mapeado  
Por favor, ingrese un numero que indique el tamaño del vector del archivo: 10000  
0  
El número ingresado es el tamaño del vector dividido en 10 elementos por subvector.  
El archivo se ha mapeado en memoria de forma exitosa.  
Tiempo de mapeo del archivo en memoria: 0.091234 segundos.  
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunks/Repositorio$
```

Figure: 2.- Imagen por terminal del mapeo de un archivo con Chunks

CASOS PEQUEÑOS: Tamaño del Vector menor a 500.000



```
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunks/Re...
diazmoreno@diazmoreno-VB:~/Escritorio/Nuevo Chunks/Repositorio$ ./Mapeado
Por favor, ingrese un numero que indique el tamaño del vector del archivo: 100000
00

El numero ingresado es el tamaño del vector dividido en 250000 elementos por sub
vector.

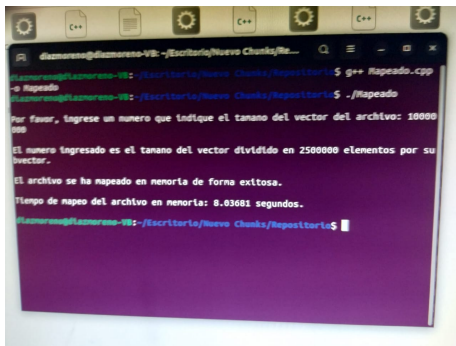
El archivo se ha mapeado en memoria de forma exitosa.

Tiempo de mapeo del archivo en memoria: 0.810535 segundos.

diazmoreno@diazmoreno-VB:~/Escritorio/Nuevo Chunks/Repositorio$
```

Figure: 3.- Imagen por terminal de Casos Pequeños. Tamaño del Vector $n = 100.000$

CASOS MEDIANOS: Tamaño del Vector mayor a 500.000



```
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunk/Repositorio$ g++ Rapeado.cpp
-o Rapeado
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunk/Repositorio$ ./Rapeado

Por favor, ingrese un numero que indique el tamaño del vector del archivo: 1000000

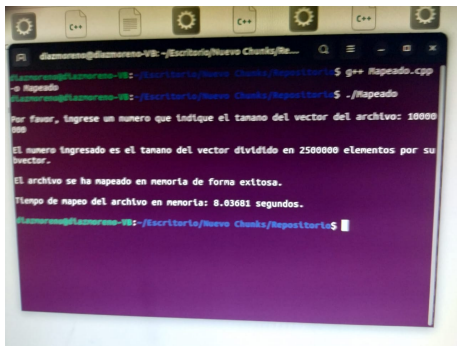
El número ingresado es el tamaño del vector dividido en 2500000 elementos por su
divisor.

El archivo se ha rapeado en memoria de forma exitosa.

Tiempo de rapeo del archivo en memoria: 8.03681 segundos.
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunk/Repositorio$
```

Figure: 4.- Imagen por terminal de Casos Medianos. Tamaño del Vector $n = 1.000.000$

CASOS GRANDES: Tamaño del Vector mayor a 500.000



```
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunk/Repositorio$ g++ Rapeado.cpp
-o Rapeado
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunk/Repositorio$ ./Rapeado

Por favor, ingrese un numero que indique el tamaño del vector del archivo: 10000000

El número ingresado es el tamaño del vector dividido en 2500000 elementos por su
divisor.

El archivo se ha rapeado en memoria de forma exitosa.

Tiempo de rapeo del archivo en memoria: 8.03681 segundos.
diazmoreno@diazmoreno-VB: ~/Escritorio/Nuevo Chunk/Repositorio$
```

Figure: 5.- Imagen por terminal de Casos Grandes. Tamaño del Vector $n = 10.000.000$

Muchas Gracias!