

Trabalho 1: Ordenação Externa

Emanoel de Moura

edmoura@inf.ufsm.br

1. Introdução

O trabalho proposto foi implementar um algoritmo de ordenação externa baseado em Merge Sort, assumindo que o arquivo de entrada não cabe em memória. A entrada deve ser um arquivo texto. O algoritmo foi desenvolvido na linguagem C++.

2. Algoritmo

O Merge Sort, é um exemplo de algoritmo de ordenação por comparação do tipo dividir para conquistar. Sua ideia básica consiste em Dividir (o problema em vários sub-problemas e resolver esses sub-problemas através da recursividade) e Conquistar (após todos os sub-problemas terem sido resolvidos ocorre a conquista que é a união das resoluções dos sub-problemas).

Funcionamento do algoritmo:

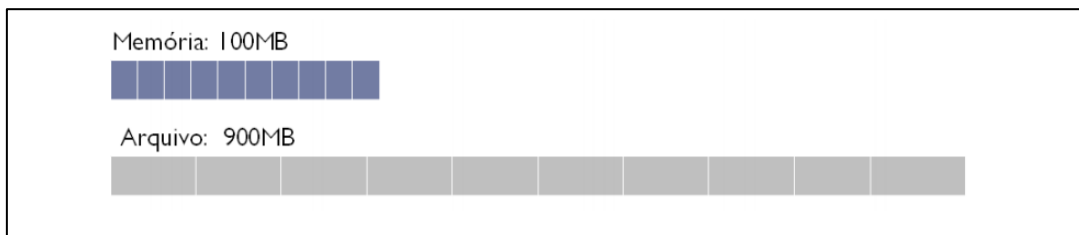
- Memória RAM comporta “N” registros de dados;
- Só podemos carregar parte do arquivo na RAM;
- Ordenar os dados na RAM com o algoritmo de Merge Sort;
- Salvar os dados ordenados em um arquivo separado;
- Repetir os anteriores até terminar o arquivo original;
- Ao final temos K arquivos ordenados;
- Criamos K+1 buffers de tamanho $N/(K+1)$;
- Carregar parte dos arquivos ordenados no buffer de entrada, intercalar no buffer de saída;
- Se buffer de entrada for vazio, carregamos mais dados;
- Se buffer de saída estiver cheio, salvamos os dados.

3. Aplicações

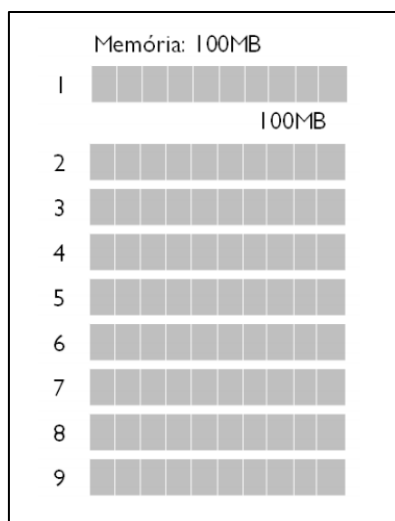
Um exemplo de aplicações é na busca online, onde os dados chegam de blocos em blocos, são ordenados por qualquer método e depois fundidos pela mesma. No entanto, esta abordagem pode demandar muito tempo e espaço de armazenamento se os blocos recebidos forem pequenos em comparação com os dados ordenados. Outro quando o número de registros a ser ordenado é maior do que o computador pode suportar em sua memória principal.

Supondo um arquivo a ordenar de 900 megabytes de dados e uma máquina com 100 megabytes de memória RAM.

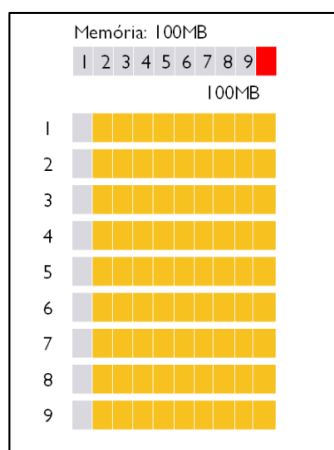
- Memória e o arquivo a ordenar



- Memória e o arquivo a ordenar particionado.



- Memória contendo a primeira parte de cada arquivo ordenado



- Para finalizar gera um novo arquivo temporário resultante da ordenação e termina quando existir somente um arquivo temporário que contém a relação inteira ordenada.

Referências

https://pt.wikipedia.org/wiki/Merge_sort

<https://www.youtube.com/watch?v=RZbg5oT5Fgw>

<https://www.youtube.com/watch?v=sVGbj1zgvWQ>

<http://recantodocodigo.blogspot.com/2016/02/algoritmo-de-ordenacao-mergesort-c.html>

http://www.uvv.br/edital_doc/ORDENA%C3%87%C3%83O%20DE%20DADOS%20EM%20MEM%C3%93RIA%20EXTERNA%20UTILIZANDO%20PROGRAMA%C3%87%C3%83O%20PARALELA_6d8f385b-d05c-494a-b805-009b57fcb015.pdf

<https://www.vivaolinux.com.br/artigo/Manipulacao-de-arquivos-em-C++>