

Metodologia PCAM - Frequência de Caracteres

Bruno Baldissera Carlotto	10724351
Bruno Gazoni	7585037
Gabriel Eluan Calado	10734453
João Villaça	10724239
Matheus Steigenberg Populim	10734710

Particionamento

O programa deve utilizar um número estático de threads. A operação potencialmente mais custosa é a transformação da string em um vetor de buckets, elementos do vetor que guardam o número de repetições de cada char de acordo com sua posição na tabela ASCII, já que é a única computação que escala conforme o tamanho da entrada. Logo, desenrolamos a execução do loop *for* que percorre a entrada em T threads, cada uma armazenando em *buckets* locais, todos de mesmo tamanho, a contagem de cada caractere dentre os 256 possíveis. O problema fica assim particionado em $O(n/T)$.

Comunicação

Na região paralela temos como elementos privados os índices do *bucket*, enquanto o *bucket* em si é compartilhado por todas as *threads*. Assim, temos *threads* que independentemente fazem a computação de sua porção do *array* de entrada, com seu próprio índice, mas que para fazerem escritas no *bucket* global entram em uma região crítica reservada apenas para a instrução de escrita, já que esta é uma região de memória compartilhada e estaria sujeita a escritas concorrentes.

Aglomeração

Após o encerramento das operações, os resultados armazenados em cada posição do vetor *bucket* por todas as threads precisam serem somados quando a execução deixa a seção paralela. Para tanto, declaramos o vetor *bucket* como uma variável *shared* na seção paralelizada, e definimos o comando de acesso e incremento a uma posição do *bucket* como uma zona crítica, com o intuito de impedir que threads diferentes tentassem acessar as mesma posição de memória ao mesmo tempo.

Mapeamento

O algoritmo está projetado para um número estático de threads considerando processamento homogêneo das EDP e algoritmo round-robin para distribuição de tarefas. Caso os nós tenham processamento heterogêneo, pode-se paralelizar a atualização do vetor de contagem de elementos de bucket também sem prejudicar a granulação desejável das EDP homogêneas responsáveis por iterar pela string. Esta alteração deve ser realizada no código, entretanto.