

Relatório Programação Paralela

Bruno Vilela - 2016.1.08.042

Abstract—Este relatório ilustra a execução de um algoritmo cujo objetivo é identificar e contar estrelas em uma imagem, em paralelo.

I. INTRODUÇÃO

O objetivo deste projeto é produzir um relatório que demonstre o comportamento de um sistema executado em aglomerados de computadores em termos de tempo de execução, tamanho do bloco de dados e quantidade de processos.

II. DESCRIÇÃO DO PROBLEMA

A tarefa a ser realizada é contar a quantidade de estrelas numa imagem astronômica (dividida em 20 arquivos PNG), que será dividida em imagens menores pelo computador master e por ele distribuída para os demais computadores slaves.

III. SOLUÇÃO

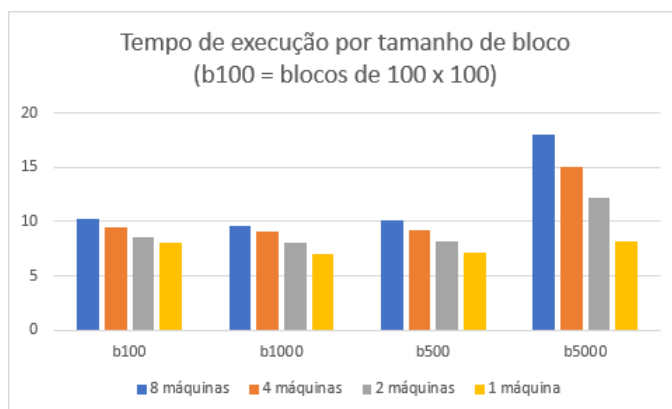
Cada computador slave deve processar a parte da imagem que receber e devolver ao master a quantidade de estrelas daquela parte da imagem. O computador master deve resumir a quantidade de estrelas de todas as partes.

Para contar a quantidade de estrelas, cada computador slave deve binarizar a parte da imagem que receber e aplicar elementos estruturantes de várias dimensões (da maior para menor), contando as estrelas. O problema de estrelas nas bordas devem ser desprezado.

IV. RESULTADOS

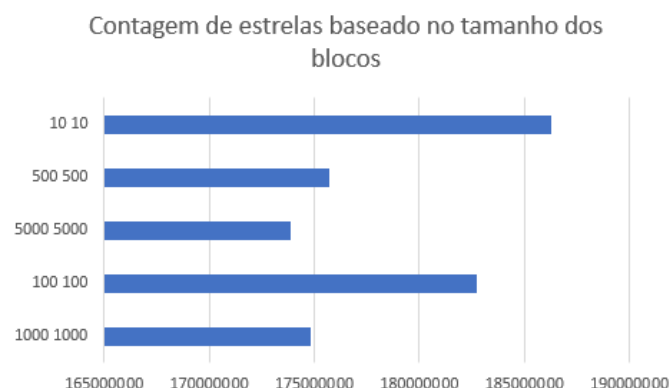
O algoritmo apresentou diferentes comportamentos com a variação dos principais parâmetros. Vamos a alguns gráficos

Execução em blocos de tamanho 100, 1000, 500, 5000, utilizando 8 CPU's de cada máquina:

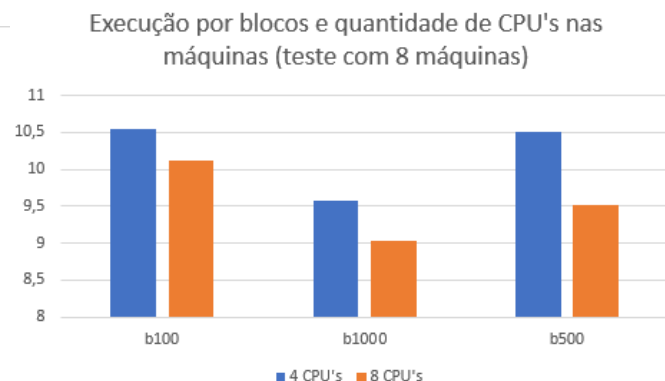


Como podemos ver, com menos máquinas conseguimos mais desempenho. Em relação ao tamanho dos blocos, conseguimos notar que um aumento muito grande, ocasiona em um gargalo no algoritmo. Também foi realizado testes para com blocos de tamanho 10x10, porém o tempo de execução a partir de 2 máquinas ultrapassou os 40 minutos, assim sendo inviável concluir os testes. Porém, com apenas uma máquina e utilizando blocos 10x10 foi obtido a maior contagem de estrelas (186318588), e um tempo de execução aceitável, 10 minutos e 12 segundos.

Em relação a quantidade blocos, o quão mais reduzimos o tamanho dos mesmos, mais aumentamos a precisão. Abaixo, um gráfico onde conseguimos visualizar esse comportamento:



Diminuindo a quantidade de CPU em cada máquina temos uma significativa piora em desempenho como podemos ver abaixo:



V. CONCLUSÕES

A contagem de estrelas em uma imagem astronômica apresentou um comportamento peculiar: Conforme adicionamos máquinas para ajudar na resolução do problema (contagem das estrelas), o tempo de execução aumenta. Existem duas causas prováveis para esse evento: ou uma máquina, para esse problema, é a quantidade ideal, ou o algoritmo atual necessita de uma melhora na implementação.

Olhando do ponto de vista do tamanho dos blocos que são separados, conseguimos abstrair que conforme diminuimos o tamanho dos blocos, melhoramos a precisão da contagem das imagens, porém pioramos o tempo de execução. Esse aumento no tempo de execução não é tão significativo, portanto é válido deixar o tamanho do bloco bem pequeno para uma contagem mais específica. Por outro lado, conforme aumentamos o tamanho dos blocos, perdemos desempenho, precisão e aumentamos o tempo. Existe um meio tempo caso o que se mais necessite seja otimizar o tempo de execução.

Do ponto de vista da quantidade de processadores, melhoramos gradativamente a performance, e consequentemente o tempo de execução, conforme aumentamos a quantidade de processadores. Ou seja provavelmente o gargalo desse algoritmo está na utilização da rede para o paralelismo, visto que quando aumentamos a quantidade de CPU's paralelas possíveis localmente, temos um ganho significativo.