

6.

Conclusão, Contribuição e Trabalhos Futuros

O presente trabalho, além de mostrar alguns dos fundamentos relacionados a *cluster*, *grids* (Anexo 2) e às ferramentas BLAST (Capítulo 3) voltadas ao ambiente distribuído, apresentou também alguns dos fatores e critérios a considerar quando efetuando a avaliação e comparação entre essas ferramentas (Capítulo 2 e Anexo 3).

Foram efetuadas duas avaliações. A primeira teve por finalidade certificar se o balaBLAST realmente efetuava o balanceamento de carga durante a sua execução, enquanto a segunda teve como propósito avaliar e comparar três ferramentas BLAST em paralelo, incluindo o próprio balaBLAST, o mpiBLAST e o gradeBLAST, esta última desenvolvida durante a elaboração do trabalho.

O conteúdo teórico, que compõe o Capítulo 1, assim como as avaliações efetuadas, constantes dos Capítulos 4 e 5, e aquelas discutidas no Capítulo 2, permitiram traçar um conjunto de ações, incluídas no Anexo 3, destinado a auxiliar na avaliação de desempenho de ferramentas BLAST em paralelo, assim como efetuar a comparação do desempenho dessas ferramentas, e que foi utilizado durante a avaliação e comparação das ferramentas BLAST acima citadas.

6.1

Contribuição.

Este trabalho procurou mostrar que é possível efetuar a avaliação e a comparação entre ferramentas BLAST voltadas ao ambiente distribuído, não somente observando os tempos de execução alcançados pelas ferramentas, mas também demonstrando à luz de outros fatores, tais como a quantidade de leituras e gravações em disco efetuadas, tempos de latência, etc., o porquê de uma

determinada ferramenta ter se comportado melhor do que alguma outra, durante a realização dos testes em ambientes distribuídos.

Adicionalmente, foram mostrados também alguns dos fatores a considerar quando se efetuando uma análise de balanceamento de carga para uma ferramenta BLAST em paralelo. O fato de se obter, ou não, um melhor resultado, a partir de um balanceamento, no presente trabalho, não pareceu tão importante, pois o objetivo era tão somente verificar se tal balanceamento era efetuado pela ferramenta balaBLAST, o que de fato foi constatado.

6.2

Trabalhos Futuros.

Uma sugestão sobre possíveis trabalhos futuros vem do fato de ser possível, dentro de um contexto de gerência de memória, otimizar ainda mais a execução do balaBLAST em um ambiente distribuído. A idéia, além de efetuar o balanceamento de carga, é otimizar a execução do BLAST em cada um dos equipamentos de um *cluster*, ou de um *grid*, e dessa maneira angariar um ganho significativo no conjunto das execuções.

Noronha, em sua dissertação de mestrado [32], apresentou uma estratégia de gerência de memória em anel com o objetivo minimizar o número de leituras em disco para um conjunto de execuções do BLAST, obtendo com isso um auspicioso ganho em termos de tempo de execução da ferramenta em uma única máquina. Por outro lado, o balaBLAST utiliza uma estratégia de alocação de fragmentos em memória, para a execução do programa “blastall” [34], que evita leituras desnecessárias em disco, acelerando com isso a execução do BLAST. No entanto, o procedimento adotado no balaBLAST, apesar de funcionar perfeitamente, pode ser otimizado quando utilizando um pequeno número de máquinas e grandes bases de dados, quando o tamanho de cada fragmento pode superar a memória RAM disponível, através da implementação, em conjunto com a estratégia atual, daquela estratégia em anel utilizada por Noronha em seu trabalho

Por outro lado, a estratégia Sob Demanda utilizada no balaBLAST, pode evoluir para um conjunto de procedimentos que utilize dados estatísticos de execução, para incrementar o desempenho da ferramenta. O problema a ser estudado, e que pode incluir algumas das estratégias utilizadas, por exemplo, no package BLAST, terá de considerar, em adição, a natureza intrínseca dos dados envolvidos, a qual possui um peso bastante considerável na determinação do tempo de execução da ferramenta.

De qualquer maneira, os estudos envolvidos nas alternativas apresentadas, deverão considerar a sua viabilidade técnica e econômica, além dos ganhos de desempenho esperados.