

Técnicas de otimização em Aceleradores Vetoriais NEC SX-Aurora*

Félix D. P. Michels¹, Matheus S. Serpa¹, Danilo Carastan-Santos¹
Lucas M. Schnorr¹, Philippe O. A. Navaux¹

¹ Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970, Porto Alegre – RS – Brasil
Programa de Pós-Graduação em Computação

{felix.junior, msserpa, danilo.csantos, schnorr, navaux}@inf.ufrgs.br

Resumo. *Avalia-se nesse trabalho a utilização de técnicas de otimização clássicas na nova arquitetura NEC SX-Aurora. Utilizou-se como estudo de caso o benchmark NAS e uma aplicação real de migração sísmica, utilizada pela indústria de petróleo e gás. Os resultados experimentais finais mostram a melhora no desempenho, em FLOPS, utilizando as técnicas de otimização loop unrolling e inlining, no benchmark NAS em até 7,8× e na aplicação real de migração sísmica em até 1,9×, em comparação com o desempenho das versões originais.*

1. Introdução

A computação de alto desempenho provê diversos benefícios para as mais variadas áreas do conhecimento, como por exemplo, a análise de dados criando novas possibilidades com novas modelagens e simulações [Ezell and Atkinson 2016]. Diferentes tipos de arquiteturas são utilizados atualmente, tais como aceleradores vetoriais. Para aprimorar o desempenho, técnicas de otimização são empregadas. Técnicas para melhorar a vetorização e aumentar a taxa de acerto da memória *cache* são utilizadas regularmente.

Semelhante a este trabalho a publicação de Komatsu et al. mostra o potencial da arquitetura SX-Aurora TSUBASA. Uma comparação com outras arquiteturas, NVIDIA Tesla V100 e SX-Ace, entre outras, mostra resultados em que a SX-Aurora tem capacidade de executar eficientemente, com até 3,5× de desempenho, além de obter um speedup maior de até 2,8× [Komatsu et al. 2018].

Neste trabalho, otimiza-se o desempenho de uma aplicação real de *Reverse Time Migration* (RTM), utilizada na indústria de petróleo e acústica, e o conjunto de *benchmarks* NAS utilizando técnicas de *loop unrolling* e *inlining* na nova arquitetura da NEC, a SX-Aurora TSUBASA, procurando melhorar a vetorização automática desempenhada pelo compilador.

2. Metodologia e Experimentos

Para a avaliação da SX-Aurora, utilizou-se o *benchmark* NAS 3.4 e uma aplicação real utilizada pela indústria de petróleo para migração sísmica, denominada RTM. Utilizando

*Este trabalho foi parcialmente financiado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, pelo projeto Petrobras (2016/00133-9, 2018/00263-5) e pelo projeto “GREEN-CLOUD: Computação em Cloud com Computação Sustentável” (#16/2551-0000 488-9), da FAPERGS e do CNPq, programa PRONEX 12/2014.

os contadores de hardware da máquina vetorial SX-Aurora, pode-se constatar uma baixa taxa de vetorização e uma alta taxa de erro da *cache*. Portanto, aplicou-se duas técnicas de otimização: *inlining* que substitui uma chamada de função pela própria função; *loop unrolling* que consiste em desenrolar laços. Utilizou-se o sistema operacional CentOS 7.9 e os compiladores exclusivos da NEC.

A execução dos experimentos utilizou os recursos da infraestrutura PCAD, <http://gppd-hpc.inf.ufrgs.br>, no INF/UFRGS, principalmente o ambiente SX-Aurora TSUBASA. O acelerador vetorial possui um processador de 8 *cores* a 1408 MHz, com memória global de 48 GB a 900 MHz, três níveis de cache, sendo a cache L1 e L2 para cada *core* e cache L3, uma unidade de processamento vetorial e uma unidade de processamento escalar [Komatsu et al. 2018]. Todos os experimentos foram executados 10 vezes, demonstrando os valores médios dessas execuções, utilizando uma distribuição de erro padrão.

3. Resultados das otimizações e Conclusão

A seguir são apresentados os resultados dos experimentos. Inicialmente, são expostos os resultados dos experimentos sendo a principal métrica avaliada a de operações de ponto flutuante por segundo (FLOPS). Obtiveram-se os dados através de contadores de hardware do acelerador SX-Aurora, e para visualização dos dados uma rotina em R foi desenvolvida. Finalmente, os resultados obtidos são expostos, comparando o otimizado e original.

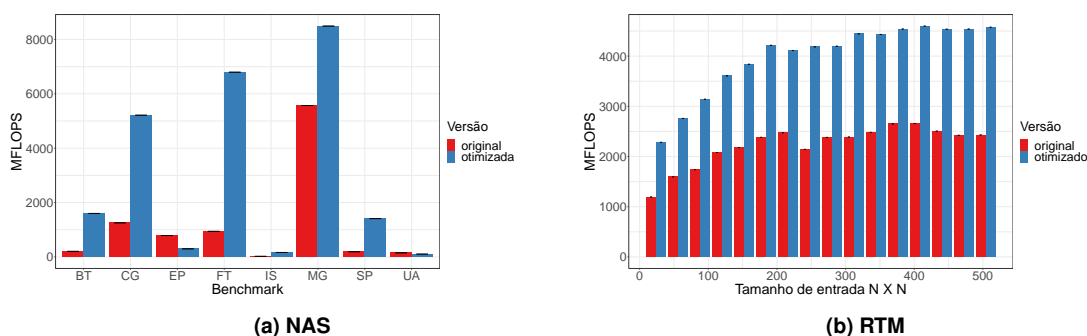


Figura 1. MFLOPS X Tamanho da entrada

Percebe-se na Figura 1 que as técnicas influenciaram em um aumento em FLOPS considerável, chegando até $7,8\times$ maior para a aplicação otimizada do *benchmark* NAS. Comparando o aumento entre a versão original e a otimizada, temos uma melhora máxima de, respectivamente, 204,42 MFLOPS e 1599,18 MFLOPS em referência ao *benchmark* BT. Para o RTM o aumento máximo foi de aproximadamente $1,9\times$, de 2429,83 MFLOPS para 4574,75 MFLOPS referindo-se ao tamanho de entrada de 504×504 .

Referências

- Ezell, S. J. and Atkinson, R. D. (2016). The vital importance of high-performance computing to us competitiveness. *Information Technology and Innovation Foundation, April*.
- Komatsu, K., Momose, S., Isobe, Y., Watanabe, O., Musa, A., Yokokawa, M., Aoyama, T., Sato, M., and Kobayashi, H. (2018). Performance evaluation of a vector super-computer sx-aurora tsubasa. In *SC18: International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 685–696. IEEE.