



## SALÃO DE INICIAÇÃO CIENTÍFICA XXVIII SIC

paz no plural



<b>Evento</b>	Salão UFRGS 2016: SIC - XXVIII SALÃO DE INICIAÇÃO CIENTÍFICA DA UFRGS
<b>Ano</b>	2016
<b>Local</b>	Campus do Vale - UFRGS
<b>Título</b>	Análise da criticalidade de erros causados por radiação em dispositivos paralelos
<b>Autor</b>	VINÍCIUS FRATIN NETTO
<b>Orientador</b>	PAOLO RECH

## **Análise da criticalidade de erros causados por radiação em dispositivos paralelos**

Aluno: Vinícius Fratin Netto

Orientador: Paolo Rech

UFRGS

Durante o período da bolsa de iniciação científica, o bolsista trabalhou principalmente com as arquiteturas NVIDIA K40, NVIDIA K1 e Intel Xeon Phi. As placas K40 e Xeon Phi são utilizadas para supercomputação, enquanto a K1 é utilizada em sistemas embarcados e é especialmente interessante pela possibilidade de ser utilizada no mercado automotivo. Um dos objetivos era avaliar como estes hardwares se comportam quando são expostos à radiação de nêutrons causada pelos raios cósmicos que atingem a Terra. Os experimentos acontecem em laboratórios internacionais especializados (LANSCE, em Los Alamos, Novo México, EUA e ISIS, Oxford, Reino Unido). Outro objetivo era medir o quanto cada erro causado por radiação influencia no resultado esperado em cada um dos dispositivos, ou seja, uma medida da criticalidade de cada erro.

Nêutrons vindos do espaço podem causar picos de energia em regiões específicas do circuito que atingem. Como consequência, uma partícula destas pode provocar bit-flips (um valor 0 indo para 1 e vice-versa). Desta forma, o resultado da computação pode ser alterado sem que isso seja notado ou, em outros casos, atingir uma parte de controle do hardware e causar um *crash* (parada) do programa. A probabilidade de um erro por radiação acontecer em um único dispositivo é baixa. Porém, em um supercomputador como o Titan, por exemplo, há aproximadamente 18.000 GPUs trabalhando em conjunto, e erros deste tipo acontecem em intervalos de dezenas de horas, provando que este é um problema real que deve ser solucionado ou atenuado de alguma forma.

O bolsista avaliou a criticalidade dos erros em dispositivos paralelos, especificamente na NVIDIA K40 e Intel Xeon Phi. Foram utilizados quatro benchmarks: GEMM (multiplicação de matrizes), Hotspot (simulação de temperatura), LavaMD (simulação de partículas) e CLAMR (mecânica dos fluídos). O trabalho partiu do pressuposto de que, mesmo que ocorram erros, estes podem não modificar significativamente o resultado final do teste, dependendo da aplicação. Por exemplo, se um pixel de uma imagem relativamente grande foi muito alterado, o erro teve grande magnitude em termos deste único pixel, porém não afetou a imagem como um todo. Este tipo de situação é explorada na chamada computação imprecisa, em que se consegue maior desempenho e menor consumo de potência dando uma margem de incerteza para as operações de ponto flutuante. Utilizando métricas espaciais e que avaliam a magnitude relativa dos erros, obtivemos resultados que podem otimizar o tempo de execução de aplicações que demandam alto poder computacional e aumentar a robustez destas aplicações.

O bolsista também trabalhou com APU's (Accelerated Processing Units), dispositivos de processamento heterogêneo da empresa AMD. Nestes dispositivos, estão integradas CPU (Central Processing Unit) e GPU (Graphics Processing Unit) no mesmo chip, otimizadas para processamento serial e paralelo, respectivamente. Com as APU's, o bolsista contribuiu para testes do HSA (Heterogeneous System Architecture), arquitetura que permite o compartilhamento da memória física entre CPU e GPU, diminuindo a latência na cópia dos dados entre dispositivos e melhorando o consumo de potência. Também foi testada a tecnologia RMT (Redundant Multi-Threading), que faz uma cópia dos dados de trabalho em tempo de execução e termina com duas saídas que são comparadas entre si como redundância. Como ambas são tecnologias ainda em fase de teste, os resultados estão sendo analisados para permitir experimentos mais robustos no futuro.