



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PARANÁ
SETOR DE CIÊNCIAS DE EXATAS
Departamento de Informática

Ficha 2 (variável)

Disciplina: Computação Paralela com GPUs						Código:CI1009						
Natureza:			(<input checked="" type="checkbox"/>) Semestral					(<input type="checkbox"/>) Anual		(<input type="checkbox"/>) Modular		
(<input type="checkbox"/>) Obrigatória			(<input checked="" type="checkbox"/>) Optativa									
Pré-requisito: CI1055 / CI1068 / CI1003 / CMA111 / CM304 / CI1056 / CI1210 / CI1001 / CMA211 / CM303 / CI1057 / CI1212 / CI1002 / CI1237 / CE009 / CI1215			Co-requisito:			Modalidade: (<input checked="" type="checkbox"/>) Presencial (<input type="checkbox"/>) Totalmente EAD (<input type="checkbox"/>) CH em EAD: _____						
CH Total: 60	Padrão (PD): 30	Laboratório (LB): 30	Campo (CP): 0	Estágio (ES): 0	Orientada (OR): 0	Prática Específica (PE):	Estágio de Formação Pedagógica (EFP):					
CH Semanal: 4												
EMENTA												
GPU:processadores gráficos para computação de propósito geral. Arquitetura de GPUs. Linguagens, bibliotecas e técnicas para programação paralela em GPUs. Modelos de arquitetura e execução. Desempenho e Escalabilidade. Sincronização. Primitivas para paralelismo. Algoritmos paralelos em GPUs. Estudo de casos de algoritmos e/ou aplicações paralelas em GPUs. Processamento heterogêneo CPU-GPU												
PROGRAMA												
PROGRAMA: Computação Paralela com GPUs:												
1. Introdução												
<ul style="list-style-type: none"> • Arquitetura de GPUs, Processadores orientados à vazão. Modelos de arquitetura e execução. Aceleração e escalabilidade. Desafios. 												
2. Computação com paralelismo de dados (Data Parallel)												
<ul style="list-style-type: none"> • Paralelismo de dados, Estrutura de programas em linguagem C/C++/CUDA. Tipos e declarações de memória em GPUs. Transferências de dados. Kernels e modelos de threads. 												

3. Escalabilidade transparente em CUDA

- Organização e mapeamento de blocos de threads. Mapeamento de threads em estruturas de dados multidimensionais. Sincronização entre threads e operações atômicas. Escalonamento e tolerância a latências.

4. Localidade de memórias e dados.

- Eficiência de acesso a memória. Divisão em quadros ou regiões (Ladrilhamento/tiling), verificações de índices de contorno. Exemplo em multiplicação de matrizes. Acessos a Memória como limitante de paralelismo (gargalos).

5. Considerações de desempenho

- Vazão de memórias. Warps e hardware SIMD. Divergência de threads. Granularidade de threads. Análise de Eficiência e Desempenho

6. Uso de biblioteca paralela Thrust

7. Padrões de paralelismo: Introdução a computações com estêncil

- Exemplos com algoritmos simples de filtros em imagens.

8. Padrões de paralelismo: Algoritmo reduce (redução)

9. Padrões de paralelismo: Soma-de-prefixos (Prefix-Sum)

- Algoritmo Prefix-Sum (Parallel Scan) em GPU.
- Introdução a eficiência de trabalho (work efficiency) em algoritmos paralelos.

10. Padrões de paralelismo: Computação paralela de histogramas

- Algoritmo Computação paralela de histogramas em GPU. Operações atômicas. Particionamento.

Latência de vazão de operações atômicas. Privatização de computações. Agregação.

11. Paralelismo irregular

- Estudo de caso em algoritmo Barnes-Hut. Layout de dados para aceleração. Caminhamento eficiente em árvore octree.

12. Estudo de Caso:

- Algoritmo massivamente paralelo para geração de árvores octree.

13. Estudo de Caso: Criptografia em GPU.

14. Outros estudos de casos: apresentações de artigos em algoritmos e aplicações paralelas em GPU

15. Trabalho Final de Implementação

OBJETIVO GERAL

A disciplina visa estudar arquiteturas, linguagens, técnicas ou algoritmos para computação, processamento e programação paralela em GPUs (ou processamento heterogêneo CPU-GPU)

OBJETIVO ESPECÍFICO

O aluno deverá ser capaz implementar algoritmos e programas em linguagem C/C++/CUDA bem como usar biblioteca de primitivas paralelas Thrust. Analisar e medir desempenho desses programas.

PROCEDIMENTOS DIDÁTICOS

A disciplina será desenvolvida mediante aulas expositivo-dialogadas, quando serão apresentados os conteúdos curriculares teóricos e através de atividades de laboratório. Serão utilizados os seguintes recursos: quadro de giz, notebook e projetor multimídia, insumos de laboratório e softwares livres específicos em Linux.

FORMAS DE AVALIAÇÃO

Uma prova teórica (P), pequenos trabalhos práticos de implementação e usos de bibliotecas de programação em Linux (tp1 até tp5, tipicamente), apresentação (A) de artigos de interesse e um trabalho final (TF) com relatório de apresentação do mesmo.

$$\text{Média} = 0.1 * P + 0.4 * (t1 + t2 + \dots + t5) + 0.1 * A + 0.4 * TF$$

BIBLIOGRAFIA BÁSICA (mínimo 03 títulos)

- [1] David B. Kirk e Wen-mei W. Hwu. Programming Massively Parallel Processors: A Hands-on Approach. 3a ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2017. ISBN : 978-0-12-811986-0.
- [2] Raphael Couturier. Designing Scientific Applications on GPUs. Chapman & Hall/CRC, 2013. ISBN: 1466571624, 9781466571624.
- [3] Wen-mei W. Hwu. GPU Computing Gems Emerald Edition. 1st. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011. ISBN : 0123849888, 9780123849885.

BIBLIOGRAFIA COMPLEMENTAR (mínimo 05 títulos)

- [4] Jakub Kurzak, David A. Bader e Jack Dongarra. Scientific Computing with Multicore and Accelerators. Boca Raton, FL, USA: CRC Press, Inc., 2010. ISBN : 143982536X, 9781439825365.
- [5] Ananth Grama et al. Introduction to Parallel Computing (2nd Edition). 2 a ed. Addison Wesley, 2003.
- [6] Hager G. e Wellein G. Introduction to High Performance Computing for Scientists and Engineers. Boca Raton, FL, USA: CRC Press, Inc., 2010. ISBN : 9781439811931.
- [7] John L. Hennessy e David A. Patterson. Computer Architecture, Fifth Edition: A Quantitative Approach. 5th. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011. ISBN : 012383872X, 9780123838728.
- [8] John D. Owens et al. «GPU Computing». Em: Proceedings of the IEEE 96.5 (mai. de 2008), pp. 879–899. DOI : 10.1109/JPROC.2008.917757 .

**OBS: ao assinalar a opção CH em EAD, indicar a carga horária que será à distância.*



Documento assinado eletronicamente por **WAGNER MACHADO NUNAN ZOLA, PROFESSOR DO MAGISTERIO SUPERIOR**, em 13/11/2018, às 03:26, conforme art. 1º, III, "b", da Lei 11.419/2006.



A autenticidade do documento pode ser conferida [aqui](#) informando o código verificador **1304438** e o código CRC **D0DC2ED8**.