



## Ficha 2 (variável) / Plano

Documento elaborado com base nas Res. 22/2021 34/2021 e 52/2021 – CEPE, considerando o contexto das medidas de enfrentamento da pandemia de Covid-19 no País, que permitem, com aprovação dos colegiados do curso, modalidade 100% remota de ensino.

Disciplinas: Computação Paralela com GPUs Tópicos em Computação Paralela						Códigos: CI1009 (BCC) CI318 (IBM)		
Natureza: ( ) Obrigatória ( x ) Optativa	( x ) Semestral ( ) Anual ( ) Modular							
Pré-requisito:	Co-requisito:	Modalidade: ( ) Presencial ( x ) Totalmente EaD ( ) *c.H.EaD						
CH Total: 60 CH semanal: 4	Padrão (PD): 60	Laboratório (LB): 30	Campo (CP): 00	Estágio (ES): 00	Orientada (OR): 00	Prática Específica (PE): 00		
Estágio de Formação Pedagógica (EFP):	Extensão (EXT): 00	Prática como Componente Curricular (PCC): 00						
<p><b>Indicar a carga horária semestral (em PD-LB-CP-ES-OR-PE-EFP-EXT-PCC)</b>  <b>*Indicar a carga horária que será à distância.</b></p>								
<h3>EMENTA (Unidade Didática)</h3> <p>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</p>								
<h3>PROGRAMA (itens de cada unidade didática)</h3> <p>PROGRAMA: Computação Paralela com GPUs:</p> <ol style="list-style-type: none"> <li>Introdução       <ul style="list-style-type: none"> <li>Arquitetura de GPUs, Processadores orientados à vazão. Modelos de arquitetura e execução. Aceleração e escalabilidade. Desafios.</li> </ul> </li> <li>Computação com paralelismo de dados (Data Parallel)       <ul style="list-style-type: none"> <li>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.</li> </ul> </li> <li>Escalabilidade transparente em CUDA       <ul style="list-style-type: none"> <li>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.</li> </ul> </li> <li>Localidade de memórias e dados.       <ul style="list-style-type: none"> <li>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).</li> </ul> </li> <li>Considerações de desempenho       <ul style="list-style-type: none"> <li>Vazão de memórias. Warps e hardware SIMD. Divergência de threads. Granularidade de threads. Análise de Eficiência e Desempenho</li> </ul> </li> <li>Uso de biblioteca paralela Thrust</li> <li>Padrões de paralelismo: Introdução a computações com estêncil       <ul style="list-style-type: none"> <li>Exemplos com algoritmos simples de filtros em imagens.</li> </ul> </li> <li>Padrões de paralelismo: Algoritmo reduce (redução)</li> <li>Padrões de paralelismo: Soma-de-prefixos (Prefix-Sum)       <ul style="list-style-type: none"> <li>Algoritmo Prefix-Sum (Parallel Scan) em GPU.</li> <li>Introdução a eficiência de trabalho (work efficiency) em algoritmos paralelos.</li> </ul> </li> <li>Padrões de paralelismo: Computação paralela de histogramas       <ul style="list-style-type: none"> <li>Algoritmo Computação paralela de histogramas em GPU. Operações atômicas. Particionamento. Latência de vazão de operações atômicas. Privatização de computações. Agregação.</li> </ul> </li> </ol>								



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 a serem desenvolvidas via acesso remoto ao sistema computacional Linux do DINF/UFPR ou em computadores pessoais. Serão utilizados os seguintes recursos: slides com áudio ou aulas pré-gravadas, apresentação de artigos para estudos de casos quando feitas por alunos deverão ser pré-gravadas para possibilitar acesso assíncrono, insumos de laboratório e softwares (gratuitos/livres) específicos em Linux. A disciplina é compatilhada com alunos de Pós-Graduação do PPGINF.

**Datas:** Início em 22/09/2021 término em 17/12/2021 conforme cronograma abaixo.

**Aulas Síncronas semanais:** quartas e sexta 16:00h, duração máxima de 1h:10min cada

**Atividades Assíncronas semanais:** 2h:50min

**Total de horas em 12 semanas: 60 h** (atividades síncronas+assíncronas)

#### Cronograma

Setembro/2021	
Datas	Atividades
22,24, 29	Síncronas

Outubro/2021	
Datas	Atividades
1, 6,8, 13,15, 20,22, 27,29	Síncronas

Novembro/2021	
Datas	Atividades
3,5, 10,12, 17,19	Síncronas
22 a 26	semana SIEPE sem atividades

Dezembro/2021	
Datas	Atividades
1,3, 8,10, 15,17	Síncronas
10	Entrega (assíncrona) do trabalho Final
17	Exame Final

**1.Sistema de comunicação:** Também nesta fase serão realizados levantamentos junto aos alunos quanto ao melhor ambiente de videoconferência para atividades síncronas (BigBlueButton, Discord, Teams/UFPR ou Google Meet). O material do curso será disponibilizado em WWW ou na UFPR virtual.

**2. Atividades síncronas:** As atividades síncronas consistirão de aulas ao vivo online até duas vezes por semana, nos horários da disciplina, com duração máxima de 01:10h cada. O objetivo principal será apresentar resumo do assunto já abordado assíncronamente e esclarecimento de dúvidas com a execução de exercícios.

As atividades síncronas poderão ser gravadas e ficarão disponíveis em local restrito ao alunado e tem como objetivo principal permitir que este tenha acesso posterior à atividade, caso por algum motivo não lhe tenha sido possível a participação no dia e horário programados.

**3. Material didático específico:** Serão utilizados slides digitalizados como material de referência básico. O professor também poderá produzir vídeos próprios onde serão esclarecidos aspectos específicos e avançados que possam surgir no decorrer da disciplina.

**4. Infraestrutura de suporte tecnológico, científico e instrumental à disciplina:** A UFPR lançou um plano para disponibilizar computadores e acesso a Internet a alunos com problemas de acessibilidade digital.



**5. Identificação do controle de frequência das atividades:** O controle de frequência será feita com base na participação do aluno nas atividades síncronas e exercícios e trabalhos entregues. Em eventuais casos de impossibilidade de participação síncrona o aluno poderá responder a um questionário online ou fazer um resumos do material gravado sobre o assunto.

#### 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.

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

Exame Final: Data 17/12/2021

O exame final consiste na re-entrega, para re-avaliação, de exercícios e trabalhos que tiveram nota insatisfatória durante o curso. Essa re-entrega pode ser feita na data final acima, para reconsideração.

#### BIBLIOGRAFIA BÁSICA (mínimo 03 títulos)

- [1] Wen-mei W. Hwu. "GPU Computing Gems" Emerald Edition. 1st. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011. ISBN : 0123849888, 9780123849885. **Disponível de forma gratuita Online.**
- [2] Vandeir Eduardo, "Arquitetura de GPUs NVIDIA e a plataforma CUDA", Capítulo 4, em "Sistema de Arquivos Criptográfico com Aceleração Especulativa em GPU", Dissertação de Mestrado, PPGINF/UFPR, Curitiba, Setembro de 2018, **Disponível de forma gratuita Online.**
- [3] 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. **acesso gratuito online para a UFPR disponibilizado via portal CAPES periódicos.**
- [4] Artigos Científicos recentes diversos, para estudo de casos com técnicas interessantes no contexto da disciplina, selecionados pelo professor e alunos, conforme interesses da turma, **acesso online gratuita online para a UFPR disponibilizado via portal CAPES periódicos.**
- [5] David B. Kirk e Wen-mei W. Hwu. Programming Massively Parallel Processors: A Hands-on Approach. 3 a ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2017. ISBN : 978-0-12-811986-0.

#### BIBLIOGRAFIA COMPLEMENTAR (mínimo 05 títulos)

- [6] Raphael Couturier. Designing Scientific Applications on GPUs. Chapman & Hall/CRC, 2013. ISBN : 1466571624, 9781466571624.
- [7] 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.
- [8] Ananth Grama et al. Introduction to Parallel Computing (2nd Edition). 2 a ed. Addison Wesley, 2003.
- [9] Hager G. e Wellein G. Introduction to High Performance Computing for Scientists and Engineers. Boca Raton, FL, USA: CRC Press, Inc., 2010. ISBN : 9781439811931.
- [10] 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.

**Professor da Disciplina:** \_\_\_\_\_Wagner Machado Nunan Zola\_\_\_\_\_

**Assinatura:** \_\_\_\_\_

**Chefe de Departamento ou Unidade equivalente:** \_\_\_\_\_Fabiano Silva\_\_\_\_\_

**Assinatura:** \_\_\_\_\_