



Nombre Completo del Programa de Posgrado		Maestría en Ciencias en Ingeniería Eléctrica		
Nombre Completo del Curso		Computación III: Cómputo Paralelo		
Tipo de Curso		Electivo	Créditos	8
Número de horas	Teóricas:	60	Prácticas:	0
		Presenciales		No presenciales
Profesores que impartirán el curso				
Arturo Díaz Pérez				
Objetivos del curso:	General	El propósito de este curso es discutir técnicas y aplicaciones de la computación paralela y de las formas en que se construyen aplicaciones para este paradigma.		
	Específicos	De manera particular, se proporcionarán los conceptos y las herramientas necesarias para explotar el enfoque de cómputo paralelo multi/many-core, guiando a los estudiantes en la preparación incremental de implementaciones para memoria compartida, uso de GPUs que aceleren núcleos de procesamiento y para intercambio de mensajes.		
Contenidos temáticos				
1. Computación Paralela				
1.1. ¿Por qué computación paralela?				
1.2. Tipos de computadoras paralelas				
1.3. Fundamentos del diseño de computadoras paralelas				
1.4. Redes de Interconexión				
2. Programación con memoria compartida				
2.1. Procesos				
2.2. Sincronización y coordinación de procesos				
2.3. Programación Multithreading con pthreads				
2.4. Programación Multithreading con OpenMP				
2.5. Caso de estudio: Algoritmo K-means.				
2.6. El proceso de paralelización				
2.7. Caso de estudio: Solución de sistemas lineales				
3. Programación para GPUs				
3.1. Arquitectura de los GPUs				
3.2. Introducción a Cuda				
3.3. Patrones de acceso a memoria en CUDA				
3.3.1. Transpuesta				
3.3.2. Reducción				
3.3.3. Histograma				
3.3.4. Operación Scan				
3.3.5. Ordenamiento				

3.4. Algoritmos avanzados para GPU
3.4.1. El problema de los n-cuerpos
3.4.2. Matrices dispersas
3.4.3. Breadth first search
3.4.4. Last node, list ranking
3.5. Programación avanzada en GPUs
3.5.1. Uso de tensor cores
4. Programación paralela mediante intercambio de mensajes
4.1. Principios básicos
4.2. Message Passing Interface
4.3. Algoritmo de Cannon para multiplicación de matrices
5. Programación paralela para mejorar el rendimiento
5.1. Descomposición
5.2. Mapeo
5.3. Calendarización
5.4. Evaluación del rendimiento
5.5. Medidas de eficiencia

Bibliografía

1. Culler, David, Singh, J. P. And Gupta, A. Parallel Computer Architecture: A Hardware/Software Approach. Morgan Kauffman Publishers, 1998. ISBN 1-55860-343-3.
2. Czech, Z. Introduction to Parallel Computing. Cambridge University Press. Cambridge, UK. 2016. ISBN-13: 978-1107174399.
3. Ananth Grama, George Karypis, Vipin Kumar, Anshul Gupta. An Introduction to Parallel Computing: Design and Analysis of Algorithms (2nd Ed.). Addison-Wesley 2003.
4. Pacheco, Peter. An Introduction to Parallel Programming. Morgan Kaufmann. 2011.
5. Quinn, Michael J. Parallel Programming in C with MPI and OpenMP. McGraw-Hill. 2003. ISBN 0072822562.
6. Wilkinson, Barry and Allen, Michael. Parallel Programming: Techniques and Applications Using Networked Workstations and Parallel Computers (2nd Ed.). Prentice-Hall Inc. 2004, ISBN 0131405632.
7. Wilt, Nicholas. The CUDA Handbook: A Comprehensive Guide to GPU Programming. Addison-Wesley Professional, 2013.
8. Herlihy, Maurice. The Art of Multiprocessor Programming. Morgan Kaufman, 2012.
9. Chapman, Barbara. Using OpenMP: Portable Shared Memory Parallel Programming. The MIT Press, 2007.
10. Balaji, P. Programming models for parallel computing. The MIT Press. Cambridge, Mass. 2015.
11. Toumanen, B. Hands-On GPU Programming with Python and CUDA: Explore high-performance parallel computing with CUDA. Packt Publishing Co. Birmingham, UK. 2018.
12. Ben-Nun, T., Levy, E., Barak, A., & Rubin, E. (2015, November). Memory access patterns: The missing piece of the multi-gpu puzzle. In Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (p. 19). ACM.
13. Bridges, R. A., Imam, N., & Mintz, T. M. (2016). Understanding GPU Power: A Survey of Profiling, Modeling, and Simulation Methods. ACM Computing Surveys (CSUR), 49(3), 41.
14. Che, S., Basu, A., & Gallmeier, J. (2016, October). Challenges of Programming a System with Heterogeneous Memories and Heterogeneous Processors: A Programmer's View. In Proceedings of the Second International Symposium on Memory Systems (pp. 99-103). ACM.
15. Diaz, J., Munoz-Caro, C., & Nino, A. (2012). A survey of parallel programming models and tools in the multi and many-core era. IEEE Transactions on parallel and distributed systems, 23(8), 1369-1386.

16. Lopez-Novoa, U., Mendiburu, A., & Miguel-Alonso, J. (2015). A survey of performance modeling and simulation techniques for accelerator-based computing. *IEEE Transactions on Parallel and Distributed Systems*, 26(1), 272-281.
17. Mittal, S., & Vetter, J. S. (2015). A survey of CPU-GPU heterogeneous computing techniques. *ACM Computing Surveys (CSUR)*, 47(4), 69.
18. Xu, Q., Jeon, H., & Annavaram, M. (2014, October). Graph processing on GPUs: Where are the bottlenecks? In *Workload Characterization (IISWC), 2014 IEEE International Symposium on* (pp. 140-149). IEEE.

Criterios de evaluación

Examen 1	0%
Examen 2	0%
Examen 3	0%
Proyecto 1	0%
Proyecto 2	0%
Apreciación (puntualidad, participación, calidad en la entrega de trabajos, etc.)	0%
Total	100%

Contribución del curso al perfil de egreso del programa

Conocimientos:

Habilidades:

Actitudes y valores: