Indian Institute of Information Technology Sri City, Chittoor
(An Institute of National Importance under an Act of Parliament)

High Performance Computing Course Plan

(CSE Program Elective)

Spring – 2021 **L-T-P-C: 3 - 1 - 0 – 4**

Outline

The Objective of the course is to understand, analyze, and implement parallel programming paradigms such as openMP, MPI and CUDA.

Course Objectives

1. Provide a basic foundation on memory hierarchy design and performance tradeoffs.
2. Provide systematic and comprehensive understanding of the hardware and the software high performance techniques in present day.
3. Introduce the fundamentals of high performance computing with the graphics processing units (GPU) and graphics processing programming paradigm.
4. Introduce concepts to design high performance versions of standard single threaded algorithms.

Prerequisites

Basic knowledge in computer programming, Basic Computer Organization.

Who Can Take The Course

Ph. D., M.Tech and Advanced UG students.

Course Outcomes

1. The learner will be able to design, formulate, solve and implement high performance versions of standard single threaded algorithms
2. The learner will be able to design programs to extract maximum performance in a multicore, shared memory execution environment processor.
3. The learner will be able to design and deploy large scale parallel programs on tightly coupled parallel systems using the message passing paradigm and advanced parallel algorithms.
4. The learner will know and will be able to demonstrate the architectural features of the GPU.

Syllabus

Module 1: Introduction: Introduction to HPC, The Memory System: Memory Hierarchy, Cache tradeoffs, Technology Trends: Moore's Law. Delay, Power, Energy. Dependability. Performance Quantification. Performance Measures of Parallel Algorithms.

2- Weeks

Module 2: Parallel Computer Memory Architectures: Parallel Platforms Models (SIMD, MIMD, SPMD), Communication (Shared Address Space vs. Message Passing) PRAM.

2- Weeks

Module 3: Shared Memory Parallel Programming: Symmetric and Distributed architectures. OpenMP Introduction. Thread creation, Parallel regions. Work-sharing, Synchronization.

3- Weeks

Module 4: Distributed Memory Parallel Programming – MPI, Collective communication. Data grouping for communication.

3- Weeks

Module 5 : Introduction to Heterogeneous Parallel Computing and Interconnection Networks - Case studies.

2- Weeks

Text Books

1. John Hennessy and David Patterson, Computer Architecture - A Quantitative Approach. 5ed. Morgan Kaufmann.
2. Ananth Grama, Vipin Kumar, Anshul Gupta, George Karypis, Introduction to Parallel Computing, Addison-Wesley, 2003
3. Wen-Mei W Hwu, David B Kirk, Programming Massively Parallel Processors A Hands-on Approach, Morgann Kaufmann, 3e.
4. William Dally and Brian Towles. Principles and Practices of Interconnection Networks, MK, 2004

Reference Books

1. Michael J Quinn, Parallel Programming in C with MPI and OpenMP, TMH, 2003
2. Barbara Chapman, Gabriele Jost, Ruud van der Pas, Using OpenMP, MIT Press, 2008.
3. Peter S. Pacheco, An Introduction to Parallel Programming, Morgan Kaufmann, 2011.
4. Gropp, Lusk, Skjellum, Using MPI, Using MPI, 2014

Web References

1. https://nptel.ac.in/courses/106/102/106102114/
2. https://nptel.ac.in/courses/106/102/106102163/#
3. https://docs.nvidia.com/cuda/cuda-c-programming-guide/index.html
4. GPU university courses list: https://developer.nvidia.com/educators/existing-courses
5. Tim Mattson. Introduction to OpenMP. SC11. (Available on Youtube)
6. MPI Video Tutorials by Open-MPI. https://www.open-mpi.org/video/

Course Evaluation

Course grades will be based on the following weightage pattern.

Sl. No.	Mode of Assessment	Marks
1	Mid Exam	20%
2	End Semester Exam	30%
3	Assignments*	30%
4	Class Participation (Surprise Quiz)	10%
5	Scheduled Quiz	10%

*Programming assignments will include serial performance, MPI, OpenMP, GPU programming.