INFOSOFT IT SOLUTIONS

Training | Projects | Placements

Revathi Apartments, Ameerpet, 1st Floor, Opposite Annapurna Block, Infosoft It solutions, Software Training & Development Institute, +91-9059683947|91-9182540872

CUDA Programming

Introduction to GPU Computing and CUDA

- Overview of GPU architecture and CUDA programming model
- Evolution and benefits of GPU computing
- CUDA programming paradigm and basic concepts

CUDA Programming Basics

- Setting up CUDA development environment (CUDA Toolkit, IDE)
- Writing and compiling CUDA programs
- Understanding CUDA threads, blocks, and grids

Memory Hierarchy in CUDA

- Overview of CUDA memory model (global, shared, constant, and local memory)
- Memory allocation and management in CUDA
- Optimization techniques for memory access patterns

CUDA Thread Coordination

- Synchronization and communication between CUDA threads
- Thread divergence and warp execution model
- Utilizing thread synchronization primitives (e.g., barriers, locks)

CUDA Kernel Optimization

- Techniques for optimizing CUDA kernels (memory coalescing, loop unrolling)
- Performance considerations and profiling tools (nvprof)
- Hands-on exercises in optimizing CUDA code

Advanced CUDA Memory Management

- Unified Memory and managed memory in CUDA
- Asynchronous memory operations and data transfers
- Best practices for memory usage in CUDA applications

CUDA Libraries and Utilities

- Overview of CUDA-accelerated libraries (cuBLAS, cuFFT, cuDNN)
- Integrating CUDA libraries into applications
- Using CUDA Thrust for high-level GPU programming

Multi-GPU Programming with CUDA

- Scalable parallelism with multiple GPUs

- CUDA Multi-GPU programming techniques (MPI, CUDA-aware MPI)

CUDA Applications and Case Studies

- Real-world applications of CUDA in various domains (e.g., scientific computing, deep learning)
- Case studies of CUDA-accelerated projects and success stories

CUDA and Deep Learning

- Overview of CUDA support in deep learning frameworks (e.g., TensorFlow, PyTorch)
- Accelerating neural network training and inference with CUDA
- Implementing custom layers and optimizations in CUDA

CUDA and Image Processing

- GPU-accelerated image processing techniques with CUDA
- Implementing filters, transformations, and feature extraction using CUDA
- Case studies in CUDA-powered image processing applications

CUDA and Parallel Algorithms

- Parallel algorithm design and implementation in CUDA
- Implementing parallel sorting, reduction, and other algorithms
- Analyzing performance and scalability of parallel algorithms

CUDA and Real-Time Systems

- CUDA applications in real-time and embedded systems
- Challenges and considerations for real-time CUDA programming
- Case studies in real-time CUDA applications

CUDA and High Performance Computing (HPC)

- CUDA in high-performance computing (HPC) clusters
- Optimizing CUDA applications for large-scale distributed computing
- Managing data locality and communication overhead in CUDA HPC

Future Trends in CUDA and GPU Computing

- Emerging technologies and advancements in CUDA
- GPU architectures and trends in parallel computing
- Exploring CUDA for AI, IoT, and other emerging fields

Capstone Project (if applicable)

- Design and implementation of a CUDA-accelerated application
- Project-based learning with mentorship and feedback