CIS 607 – Winter 2010
General Purpose
GPU (GPGPU) Computing

Lecture 1
Introduction

University of Oregon
Department of Computer and Information Science
Course Information

- Welcome!
- CIS 607 seminar
- Instructor
  - Prof. Allen D. Malony (malony@cs.uoregon.edu)
- Course webpage
  - http://www.cs.uoregon.edu/classes/11W/cis607gpu
- Class:
  - Wednesday, 14:00-15:20 pm, 200 Deschutes
Course Text and Reference

- Programming Massively Parallel Processors: A Hands-on Approach
  - David B. Kirk, NVIDIA, and Wen-mei Hwu, University of Illinois, Urbana-Champaign
  - Morgan Kaufmann, 2010

- GPU Computing Gems
  - Wen-mei Hwu (Editor)
  - Morgan Kaufmann, 2010
Courses Resources

- Text webpage
  - http://www.elsevierdirect.com/morgan_kaufmann/kirk/
- There are a lot of other materials and resources
  - References, specifications, courses, tutorials, …
  - Check course webpage
- Slides form book prepared by authors
Seminar Plan

- Goals
  - Understand multicore/manycore architecture
  - Focus on CUDA and OpenCL programming
- Utilize existing course material and tutorials
- Use seminar to jointly organize training materials
  - Install tools for CUDA and OpenCL programming
  - Create webpage/repository of GPU resources
- Follow “hands-on” approach
  - Get experience with NVIDIA GPGPUs
- Help prepare for the ACISS system
Assignments and Grades

- Grades are P/NP
- Everyone give an presentation
- Everyone will work on a GPU programming project
- Everyone will help in collection materials and putting together an online resource
CUDA

- Computer Unified Device Architecture (CUDA)
- CUDA is the computing architecture of NVIDIA’s GPU
  - Supported across all GPU line (only NVIDIA!)
- CUDA is also a programming environment
  - C extensions
  - nvcc compiler
    - based on Open64 compiler
  - Driver and runtime library
  - Several language bindings
    - Fortran, Python, …
  - Several libraries/packages
    - CUBLAS, CUFFT, …
OpenCL

- Open Computing Language (OpenCL)
- Framework for programming heterogeneous platforms
  - Task-based and data-based parallelism
  - C-based language for writing kernels
  - Data structures, APIs, and libraries
- OpenCL is cross-platform
  - NVIDIA and AMD GPUs
  - Also can run on multi-core processors
GPU Resources

- **Mist cluster**
  - 2x S1070 GPU computer servers
  - Each server contains 4x Tesla GPUs
  - Each Tesla GPU contains:
    - 240 streaming processor cores
    - 4 GB memory

- **GPU workstations with Fermi architecture**
  - Frankenstein
    - 4x GTX 480
  - Vampir and Werewolf
    - each 2x C2050