Master Program (Laurea Magistrale) in Computer Science and Networking

University of Pisa and Scuola Superiore Sant'Anna

Course Notes, 2012-13

High Performance Computing

Marco Vanneschi

Department of Computer Science, University of Pisa

Course objectives

This course deals with two interrelated issues in high-performance computing:

1. fundamental concepts and techniques in parallel computation structuring and design, including parallelization methodologies and paradigms, parallel programming models, their implementation, and related cost models;
2. architectures of high-performance computing systems, including shared memory multiprocessors, distributed memory multicomputers, clusters, and others.

Both issues are studied in terms of structural model, static and dynamic support to computation and programming models, performance evaluation, capability for building complex and heterogeneous applications and/or enabling platforms, also through examples of application cases. Technological features and trends are studied, in particular multi-/many-core technology and high-performance networks. An initial part is dedicated to review basic concepts and techniques in structured computer architecture, in order to render the different backgrounds of students as uniform as possible.

Course outline

Part 0: Background - Structured Computer Architecture

Structuring by levels and processing modules

The firmware level

The assembler machine and its basic interpreter

Processes and virtual memory

Memory hierarchies and cache architecture

Interprocess communication mechanisms and their run-time support

Part 1: Structuring and Design Methodology for Parallel Applications

Structured parallelism at applications and process levels

Cost models

Impact of communications

Parallel computations as queueing systems / queueing networks

Parallel paradigms: Pipeline, Data-flow, Farm, Function partitioning, Data parallel

Parallel systems at the firmware level: Instruction level parallelism: Pipeline, superscalar, multithreaded CPUs; SIMD architectures and GPUs

Part 2: Parallel Architectures

Shared memory multiprocessors: SMP and NUMA architectures

Distributed memory multicomputers: Clusters and MPP architectures

Run-time support to interprocess communication

Interconnection networks

Performance evaluation

Multicore architectures