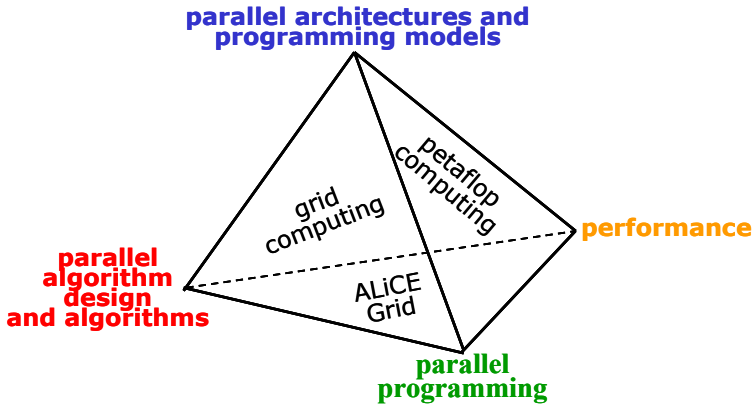
SINGAPORE-MIT ALLIANCE (COMPUTER SCIENCE)

SMA5505 / 6.338J/ 18.337J – Applied Parallel Computing

Department of Computer Science, National University of Singapore (NUS)
Spring 2004 (updated: 20 April 2004)



Lecture Slides

L0 – Course Admin

L1 – Introduction
- What, Why, How
- Classical (Modern) Science
- Definitions
- Evolution of Supercomputing
- Limits of Serial Computing
- Programming Parallel Computers
- Computing in the Internet Age
- Summary

L2 – Parallel Architectures and Programming Models
- Architectures – von Neumann, Flynn's taxonomy, memory model, interconnect
- Parallel Programming Models – Shared, Distributed and Hybrid
- Designing Parallel Programs
- Overheads of Parallelism
- Summary

L3 – Message Passing Computing
- Overview
- What is MPI?
  - Six Basic functions
  - Basic Program Structure
  - Basic Send and Receive
  - Types of Program
- Summary

L4 – Shared- Memory Programming

- OpenMP
- Data Parallelism
  - Shared memory Model
  - Parallel for Loops
  - Declaring Private Variables
  - Critical Sections
  - Reductions
  - Performance Improvements
  - More General Data Parallelism
- Functional Parallelism
- Summary

L5 – Parallel Algorithm Design

- Motivation
- Task/Channel Model
- Algorithm Design Methodology
  - Partitioning
  - Communication
  - Agglomeration
  - Mapping
- Examples
  - Finding the Maximum
  - N-Body Problem
- Summary

Lx – Parallel Algorithms (see Edelman's slides)

L6 – Principles of Scalable Performance

- Arguments against the Merit of Parallelism
- Performance Metrics
  - Average Program Parallelism
  - Harmonic Mean Performance
  - Efficiency, Utilization and Quality of Computation
- Applications / Algorithms
  - Application Models
  - Scalability of Parallel Algorithms
- Speedup Performance Laws
  - Fixed Workload - Amdahl's Law (1967)
  - Scaled Problems - Gustafson's Law (1987)
  - Memory-bounded Speedup Model – Sun and Ni (1993)
- Scalability Analysis and Approaches
- Summary

L7 – Grid Computing

- Internet – 3 Generations
- The Grid Problem
- Why Grid and Why Now?
- What is Grid Computing?

- Types of Grid and Grid Computing Models
- Main Grid Computing Problems
  - Systems Problem
  - Programming Problem
- Where are we today?
  - Global Grid Forum
  - OGSA
  - OGSI
  - Globus Project, GT2, GT3
- GT3 Implementation & Terminology
- Summary

L8 – Case Study of ALiCE Grid
- Cost of Idle Computing Cycles
- What and Why ALiCE?
- ALiCE
  - Design
  - Implementation
  - Grid Programming
  - Some Applications
- What's next?
  - Coping with System Complexities and Challenges
  - Desirable Properties of Large Distributed Systems
  - Ongoing Works
- Grid Computing Activities – Worldwide and in Singapore
- References

L9 – Conclusions
- Supercomputing today
- Cost Comparison - Supercomputer vs a Physical Cluster and a Virtual Grid of 100,000 PCs
- Distributed Computing Economics
- Petaflop Computing
  - Parallel Architectures
  - Application Requirements
- Predictions
- What have we covered?
- References

Reference Texts

1. Fundamentals of Parallel Processing, Harry Jordan, Gita Alaghband, Prentice-Hall, 2003.
2. Parallel Programming in C with MPI and OpenMP, Michael J. Quinn, Mc-Graw Hill, 2003.
3. Parallel Scientific Computing, Alan Edelman, Spring 2002 (draft).
4. Edelman's Notes