Harvey Mudd College

CS 156
Parallel and Real-Time Computation

Spring 2001

Prof. Bob Keller
keller@cs.hmc.edu
x 18483
Olin 1242

Outline

- Course outline, etc.
- Representative parallel computers
- Sample applications
- Architectural varieties
- Performance metrics and bounds, such as Amdahl's law

Prerequisites

- CS 110, Architecture and Operating Systems
- CS 140, Algorithms
- CS 131, Programming Languages, is helpful, but not absolutely required

Materials

- Slides will be posted to course web page:

http://www.cs.hmc.edu/courses/2001/spring/cs156

- They will be available in two forms:
  - web browser form
  - pdf, 6 slides per page

Course Emphasis Plan

- Parallel computation: about 80%
- Real-time computation: about 20%

Graded Work Plan

- Weekly assignments: 50%
- Presentation: 5%
- Longer project: 30%
- Attendance and participation: 15%

Textbook for the parallel part

Barry Wilkinson and Michael Allen,
Parallel Programming: Techniques and Applications Using Networked Workstations and Parallel Computers,
Prentice-Hall, 1999, ISBN 0-13-671710-1.

I will label this "PP".

Other sources for both parts

- Various papers, books, web pages

Outline for Parallel Part (with reading) (consider this approximate)

(1 day) Basic ideas and vocabulary of parallel computing (PP chapter 1)
(2 days) Message-passing computation, MPI (PP chapter 2)
(1 day) Embarassingly-parallel problems (PP chapter 3)
(2 days) Partitioning (PP chapter 4)
(1 day) Pipelining (PP chapter 5)
(1 day) Synchronous computations (PP chapter 6)
(1 day) Load balancing and termination detection (PP chapter 7)
(2 days) Programming with shared memory (PP chapter 8)
(2 days) Sorting algorithms (PP chapter 9)
(2 days) Numerical algorithms (PP chapter 10)
(1 day) Image processing (PP chapter 11)
(2 days) Searching and optimization, genetic algorithms (PP chapter 12)
(1 day) ZPL language, Fortran-based languages
(1 day) Functional language approach; Linda, Javaspaces
(1 day) PRAM model (PP appendix D)
(1 day) NESL model and language

Outline for Real-time Part

(1 day) Scheduling algorithms
(1 day) Rate-Monotonic scheduling
(1 day) Scheduling algorithms
(1 day) Rate-Monotonic scheduling
(1 day) Real-time languages
(1 day) Real-time operating systems
(1 day) Real-time languages
(1 day) Temporal logic and specifications

Definitions

- **Parallel computation**: Computation in which there are multiple computing activities occurring simultaneously.

- **Real-time computation**: Computation in which it is essential that *time* be taken into account in defining correctness.

Integrating parallel and real-time

- Parallel computation can be used in the context of real-time computation.

A History of Parallel Computation

- http://www.interlog.com/~gvwilson/articles/history.html (by Greg Wilson) begins at 1955

Purposes of Parallel Computation

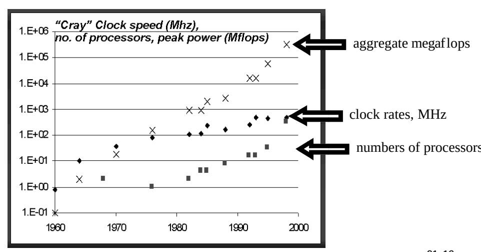
- Speedup
- Fault-tolerance
- Geographical Conformance
- Natural Structuring of a Computation

Why Parallel Computing will become *essential* in our lifetimes

- The processor performance-improvement curve of doubling every two years cannot continue forever.
- There are **physical limitations** on switching and propagation speeds.
- Saturated device capabilities leaves two alternatives:
  - Better sequential algorithms
  - Parallel algorithms

What do these curves tell us?

From Gordon Bell: http://research.microsoft.com/users/gbell/craytalk/sld013.htm



Parallel Processor: Cray T3E-1200

http://www.cray.com/products/systems/crayt3e/1200

- Distributed memory
- 32 - 2048 processing elements (PEs)
- 8 Million words (64 MB) per PE
- PE: DEC Alpha, 1200 Mflops
- 2.4 teraflops maximum
- 3-D torus interconnect
- interprocessor comm. rate: 500 MB/s
- starting at the paltry sum of $630,000

SGI Origin 3000 series

http://www.sgi.com/origin/3000/

- 16-512 processors
- up to 1 TB shared memory
- R12000 processors, 300 MHz
- NUMA architecture
- 38.4 Gflops for 64 processors
- 716 GB/s system bandwidth
- Cost starting around $300k

Parallel Processor: IBM SP3

http://www.rs6000.ibm.com/hardware/largescale/index.html

- 1-8 processors, shared memory
- up to 16 GB memory
- networkable to 512 processors
- PE: IBM RS/6000
- 2-level coherent cache
- 14.2 GB/s crosspoint switch
- Cost for 128 PE version:
  $8,614,441

01-19

Current Top 5 Sites

http://www.top500.org/list/2000/11/

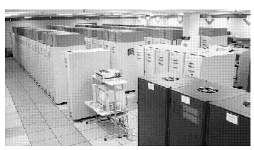
Rank	Manufacturer	Computer	Rmax (GFlops)	Installation Site	Country	Year	Area of Installation	# Proc
1	IBM	ASCI White, SP Power3 375 MHz	4938	Lawrence Livermore National Laboratory Livermore	USA	2000	Research Energy	8192
2	Intel	ASCI Red	2379	Sandia National Labs Albuquerque	USA	1999	Research	9632
3	IBM	ASCI Blue-Pacific SST, IBM SP 604e	2144	Lawrence Livermore National Laboratory Livermore	USA	1999	Research Energy	5808
4	SGI	ASCI Blue Mountain	1608	Los Alamos National Laboratory Los Alamos	USA	1998	Research	6144
5	IBM	SP Power3 375 MHz	1417	Naval Oceanographic Office (NAVOCEANO) Bay Saint Louis	USA	2000	Research Aerospace	1336

ASCI = “Accelerated Strategic Computing Initiative (DOE)

01-20

Last Year's fastest computer: ASCI Red

http://www.sandia.gov/ASCI/Red/RedFacts.htm



- Intel
- 4640 nodes
- 666 MOPs peak per node
- 2 TOPS peak Linpack performance
- 606 GB memory

01-21

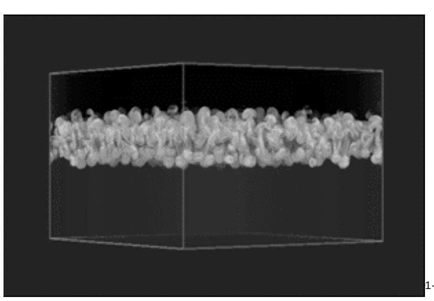
Applications Using Whopper Machines

- Hydrodynamic simulation (e.g. turbulence)
- Weather forecasting
- Modeling materials
- Pharmaceutical research
- Astrophysical simulations
- Reactor modeling
- Financial simulations
- Databases, data mining, web search

01-22

Turbulence Calculation

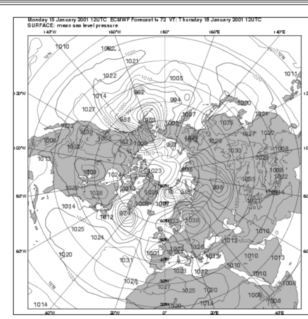
http://www.llnl.gov/casc/asciturb/simulations.shtml



01-23

3-day Forecast, Northern Hemis.

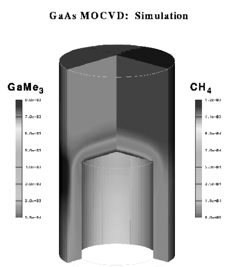
http://www.ecmwf.int



01-24

Reactor Modeling

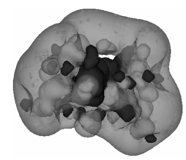
http://www.cs.sandia.gov/CRF/MPSalsa/main_pictures.htm



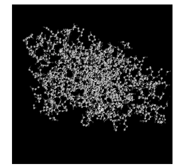
01-25

Solution of the Poission-Boltzmann equation

(from http://compbio.caltech.edu/applications/pmg/sod.html)



Isosurfaces of electrostatic potentials in the SOD (SuperOxide Dismutase) enzyme obtained by solving the nonlinear Poisson-Boltzmann equation numerically. The SOD enzyme is an antiradical or antioxident, meaning that it moves around the body binding to and then deactivating free radicals in the body, preventing them from causing cancer or other cell damage in the body.



SOD enzyme molecular structure

01-26

Parallel Web Search Engines

- inFind
- MultiCrawl

01-27

For the latest computer scores ...

TOP500 superCOMPUTER SITES

http://www.top500.org/list/2000/11/

Currently dominated by: Cray/SGI, IBM, Hitachi, Fujitsu, NEC, HP, Sun.

Beowulf clusters are no longer on the list.

01-28

For the Monetarily-Challenged

- Iditarod Multiprocessor from Terra Soft
- PE: 8-16 500 MHz G4 processors
- 256 MB per PE
- ethernet switch
- $32,000

Parallel Computing has something for everyone

- Application developers
- Language developers
- Algorithm developers
- Computer architects
- System developers
- Entrepreneurs
- Philosophers
- Engineers
- Physicists

For any topic of investigation X, one can investigate "*parallel* X".

01-30

Human-Scale Parallel Examples

- Supermarket checkout lanes
- Multi-screen movie theatres
- Multi-lane highways
- Assembly lines
- Excavations
- Orchestra

Thought Experiments

- Each person in this room is a processor.
- Decide how to make efficient use of parallelism to do the following tasks.

Thought Experiment 1

- One person is handed a white-page phone book and given a number.
- The problem is to find the individual in the book who has that number, or indicate that none exists.

Thought Experiment 2

- One person is handed a white-page phone book.
- The problem is to produce an ordering of all of the telephone numbers in increasing numeric order.

Thought Experiment 3

- Reconsider how to do the problem in Experiment 1 (find person given number),

  given that Experiment 2 has been done (ordered listing of numbers).

Thought Experiment 4

- Everyone is given a phone-book *size* list of coefficients for a set of simultaneous linear equations.
- Find a reasonably accurate solution to the equations, assuming one exists.

Conclusions from Experiments?