

Parallel Computing: From Inexpensive Servers to Supercomputers

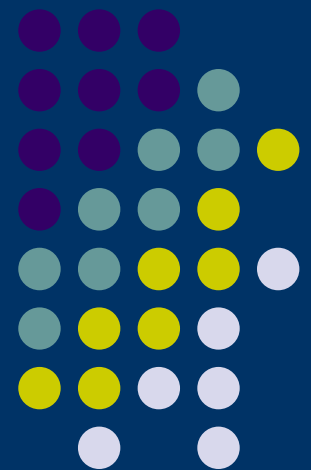
Lyle N. Long

The Pennsylvania State University &
The California Institute of Technology

Seminar to the Koch Lab

<http://www.personal.psu.edu/lnl>

February 1, 2008



Outline



- Overview of parallel computer hardware and software
- Discussion of some existing parallel computers
- New inexpensive, yet powerful, desktop computers
- Some performance results

Warning!



- I will present a lot of “Peak” or “Linpack” numbers for computer performance.
- These are nothing more than a performance level that you will never reach !
- You might get 10 – 20% of peak speed
- Many years ago, I achieved 50% of the peak speed using 4096 processors (CM-5) and won a Gordon Bell prize



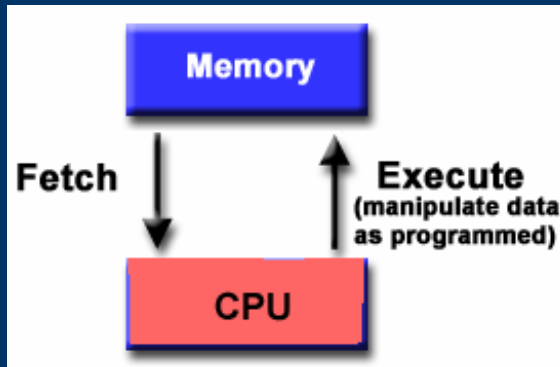
Introduction

- Traditional computers have one processor connected to the main memory (von Neumann)
- Symmetric Multi-Processor (SMP) machines have typically <64 processors in one cabinet all connected to the same memory (with high speed, expensive interconnect, e.g. cross-bar switch)
- Massively parallel (MP) computers (and PC clusters) use network connections (even up to 200,000 processors)
- Chips now have more than one processor on them: multi-core or “SMP on a chip” (MP machines can be built using them too)
- Also, 64-bit operating systems, now allows large amounts of memory (128 GB) on your desktop (or at least next to it!)

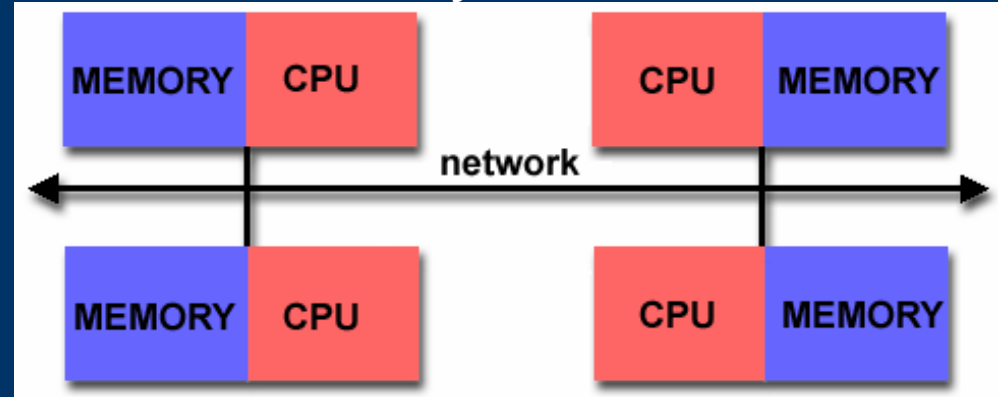


Parallel Computer Architectures

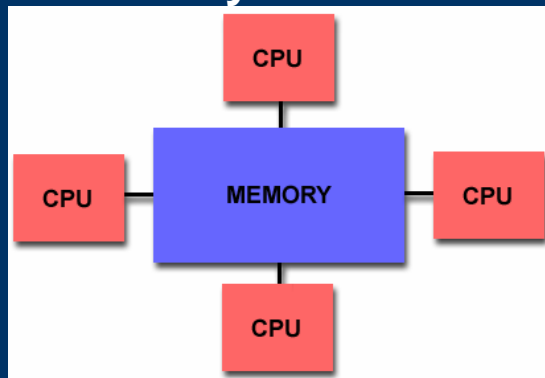
Traditional (von Neumann)



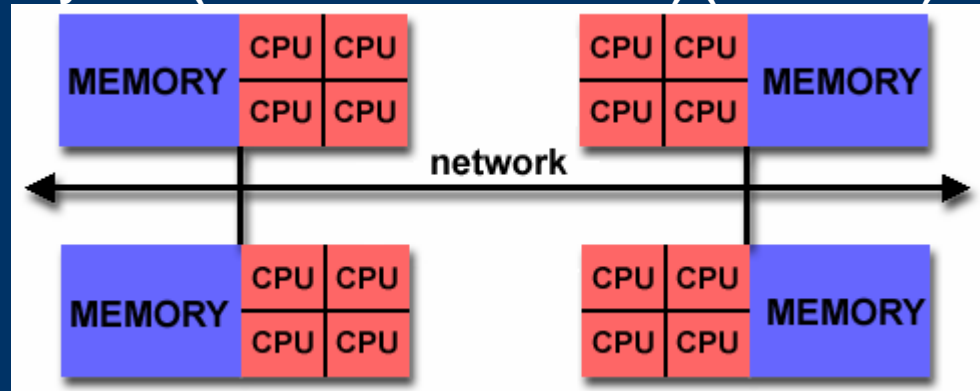
Distributed Memory



Shared Memory



Hybrid (shared & distributed) (the trend)



Easy to use, but not scalable

Difficult to use, but scalable

Parallel Computing Software Approaches

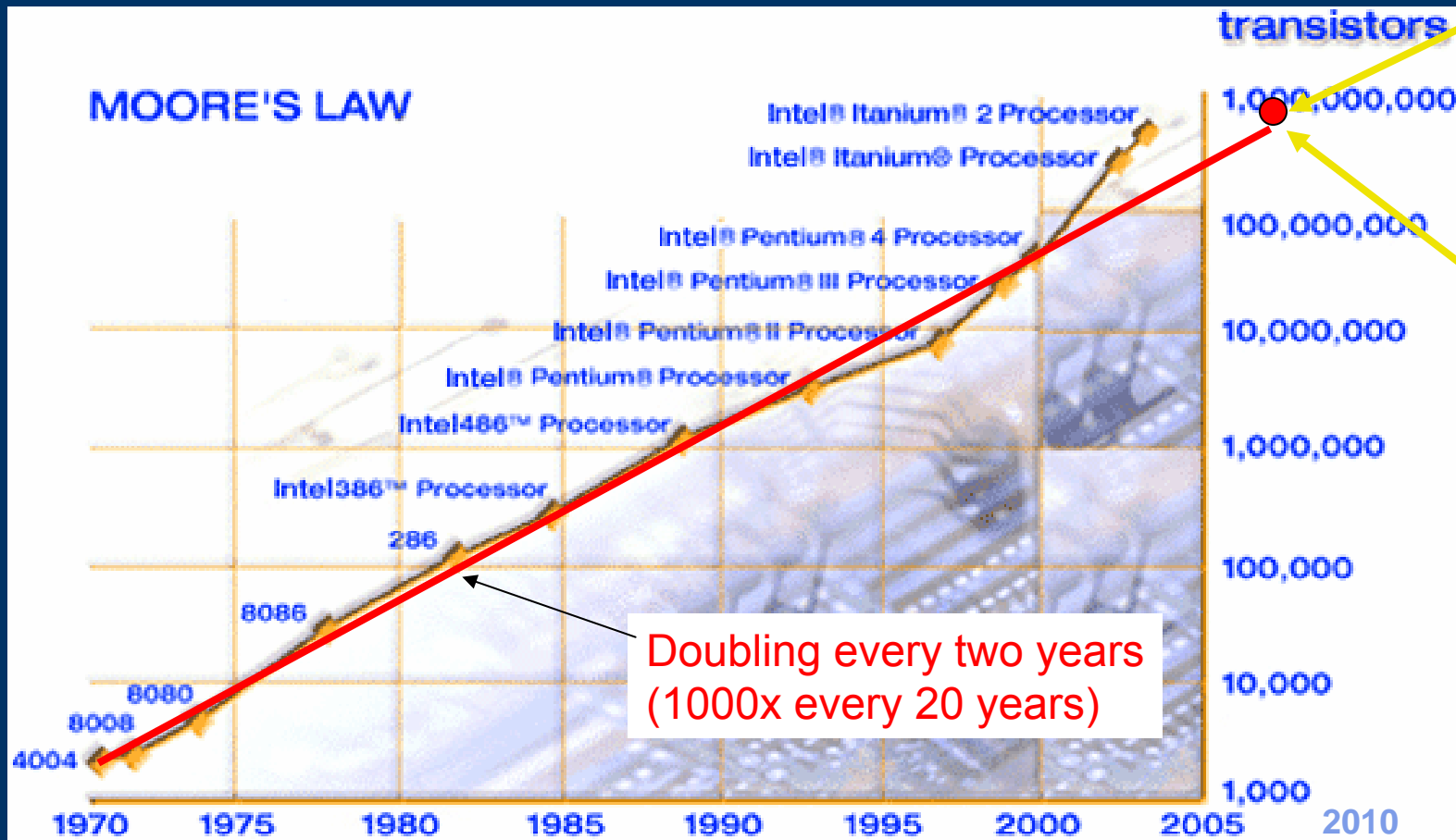


- Message passing (MPI)
 - Dominant approach
 - Unfortunately, very difficult for many problems
 - Must hand-code all inter-processor communications
- OpenMP
 - Very easy software development
 - Not available on MP
- Threads
 - Fairly easy
 - Java has threads built in
 - C/C++ with Posix threads
- Data Parallel
 - Used on old Connection Machines (~4096 processors)
 - Unfortunately, out of favor
- Hybrid
- Others ...

The market for supercomputers is so small, that there is little incentive for industry to develop good compilers for Massively Parallel computers.

Moore's Law

“no. of transistors/chip doubles every year”, 1965, “every two years”, 1975)
 (Co-Founder Intel, Ph.D., Chemistry, Caltech, 1954)



- IBM Power6
- 790 million transistors
- 2007
- 65 nm

- Intel Xeon 5400
- 820 million transistors
- 2007
- 45 nm

This is about 400 molecules wide !!

2 K transistors

2 M transistors

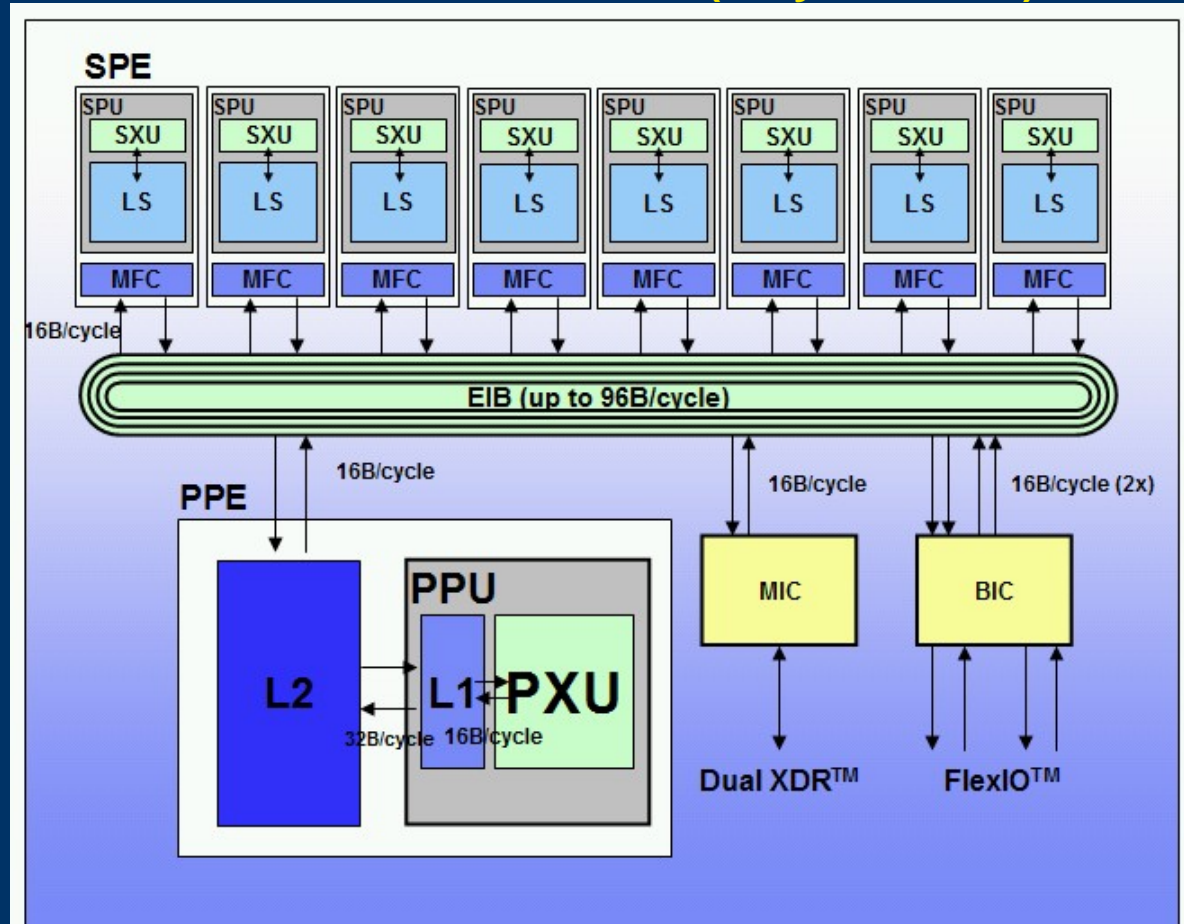
2 B transistors



Multi-Core Chips

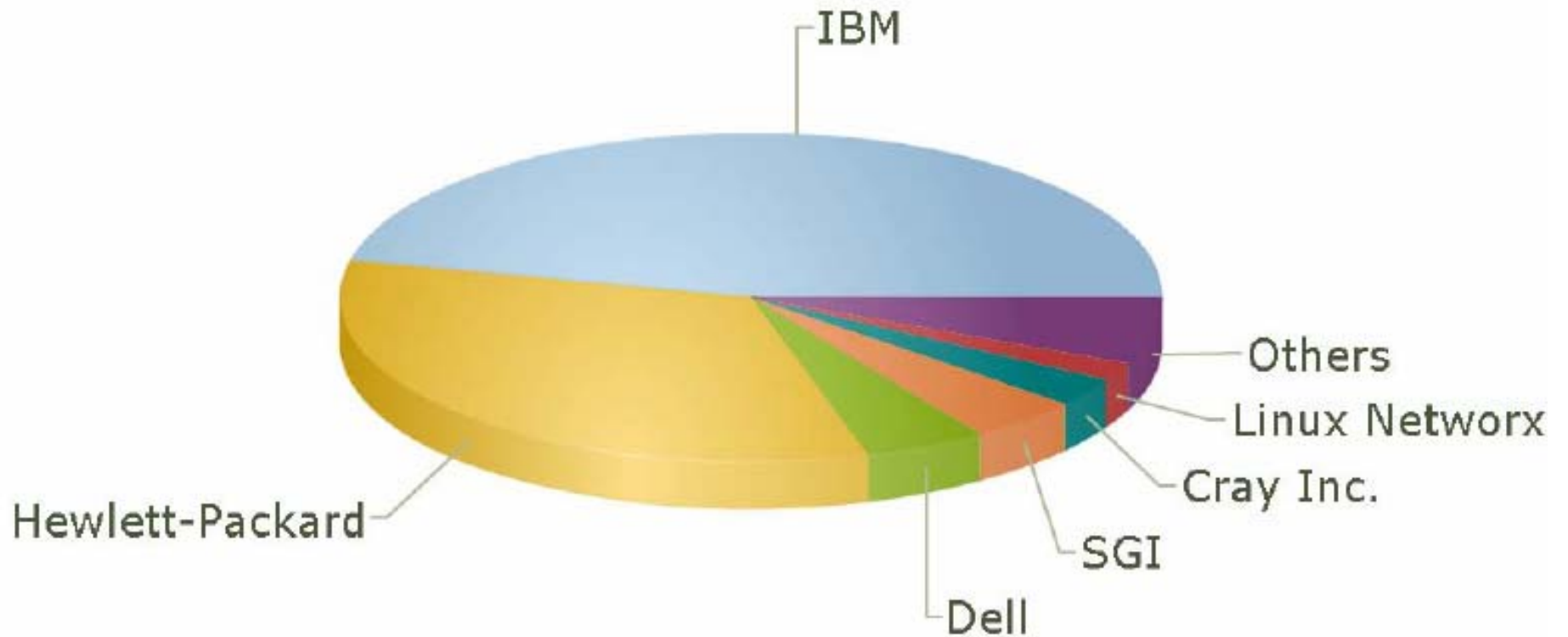
- Intel
 - Xeon Quad-Core
- AMD
 - Phenom Quad-Core
- Sun
 - T2 8 core
- IBM
 - Cell (8 spe + 1 cpu)
 - with Sony

IBM Cell Processor (PlayStation 3)



Source: M. Gschwind et al., Hot Chips-17, August 2005

Top 500 Largest Supercomputers



www.top500.org

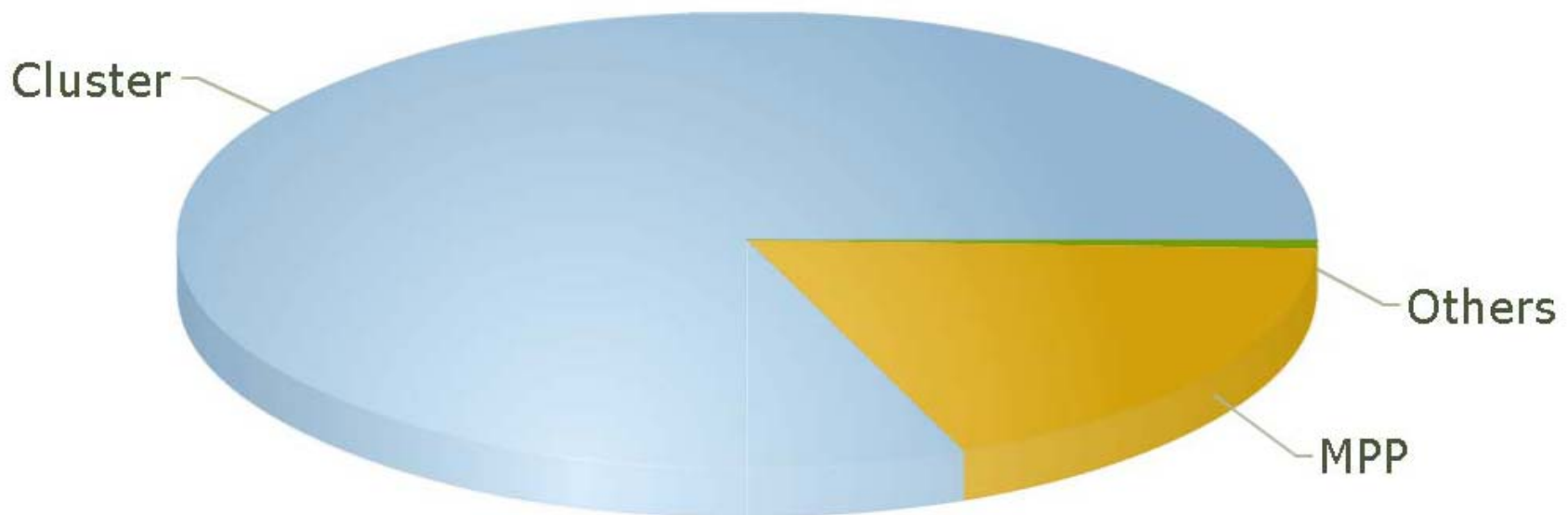
Nov., 2007

Top 500 Largest Supercomputers

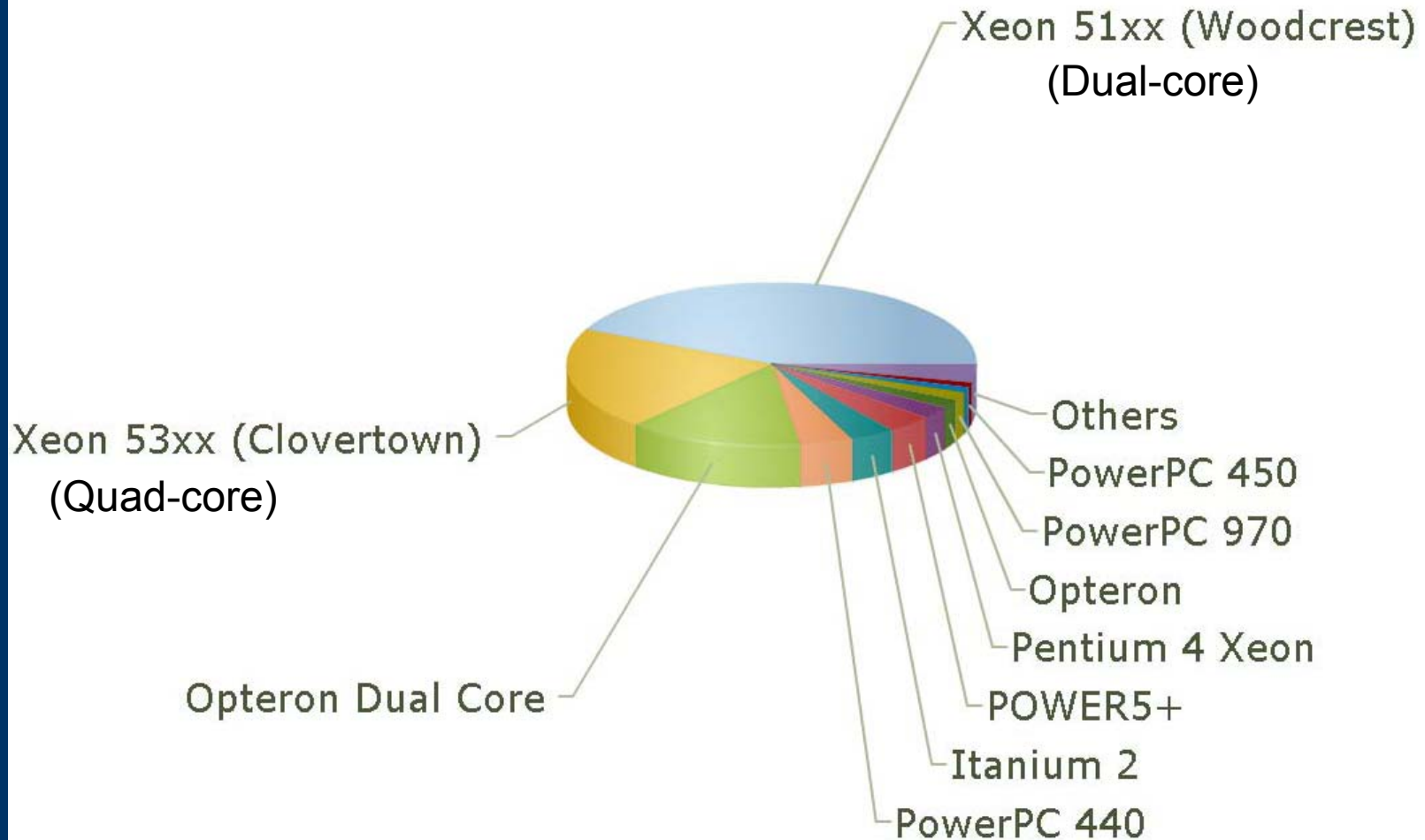


Power and A/C are huge concerns these days.

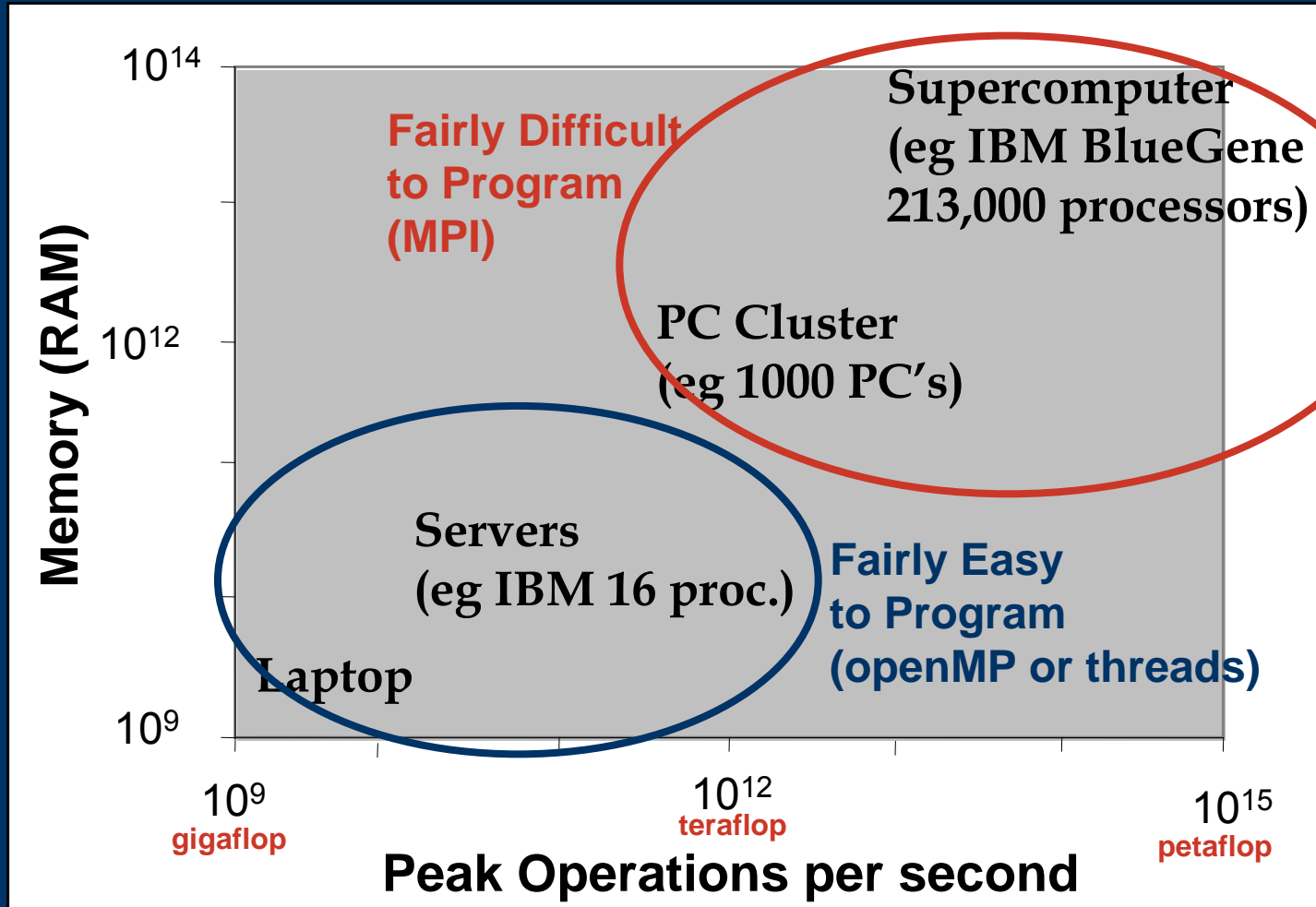
A 131,000 processor BlueGene/L requires 1.5 megawatts (\$ ~1M/year) and 300 tons (4 M BTU / hour) of cooling.



Processors used in Top 500 Largest Supercomputers



Range of Computer Systems



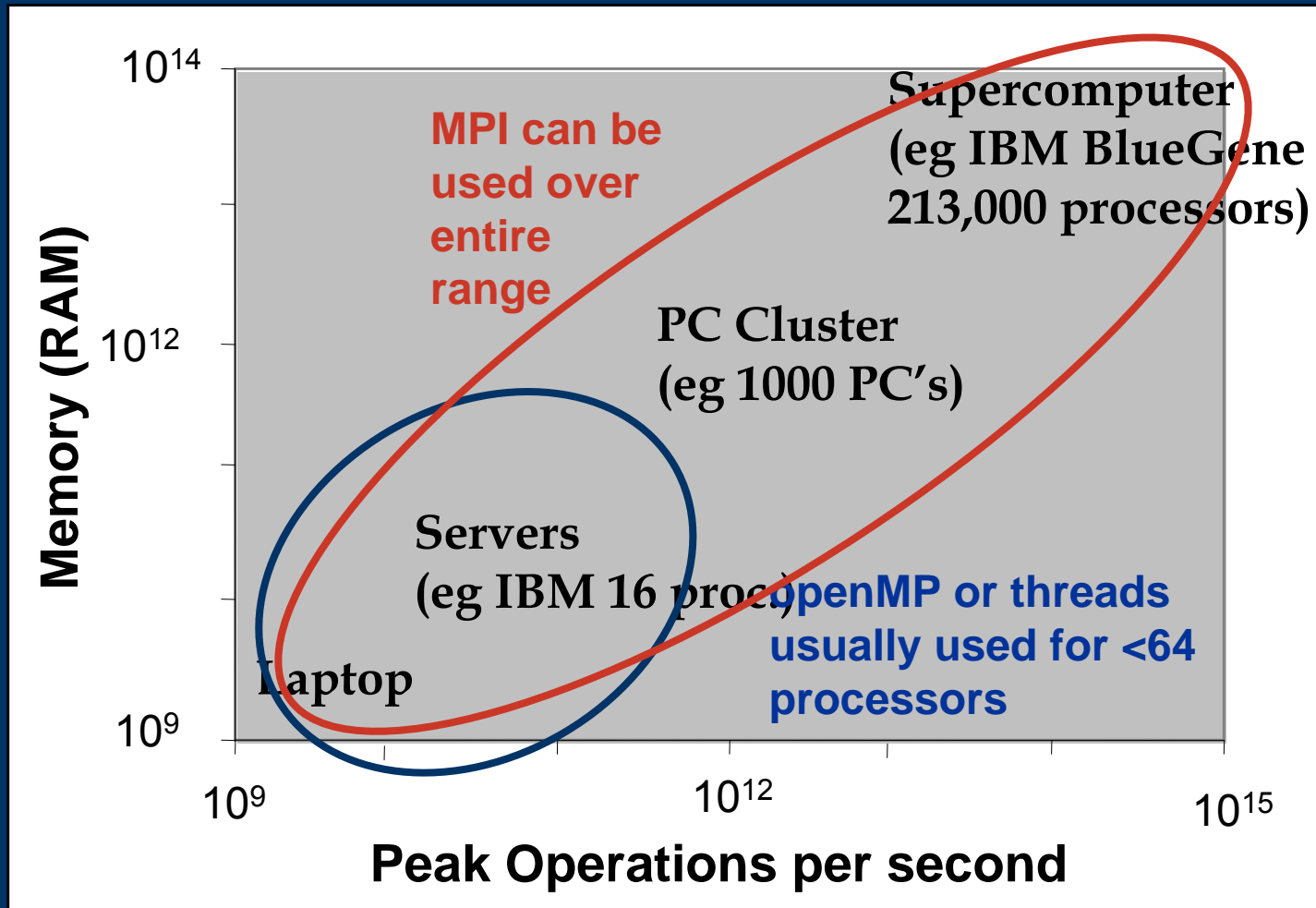
\$ 200 M ?

\$ 10 M ?

\$ 1 M ?

\$ 2 K

Range of Computer Systems



As will become clear later:

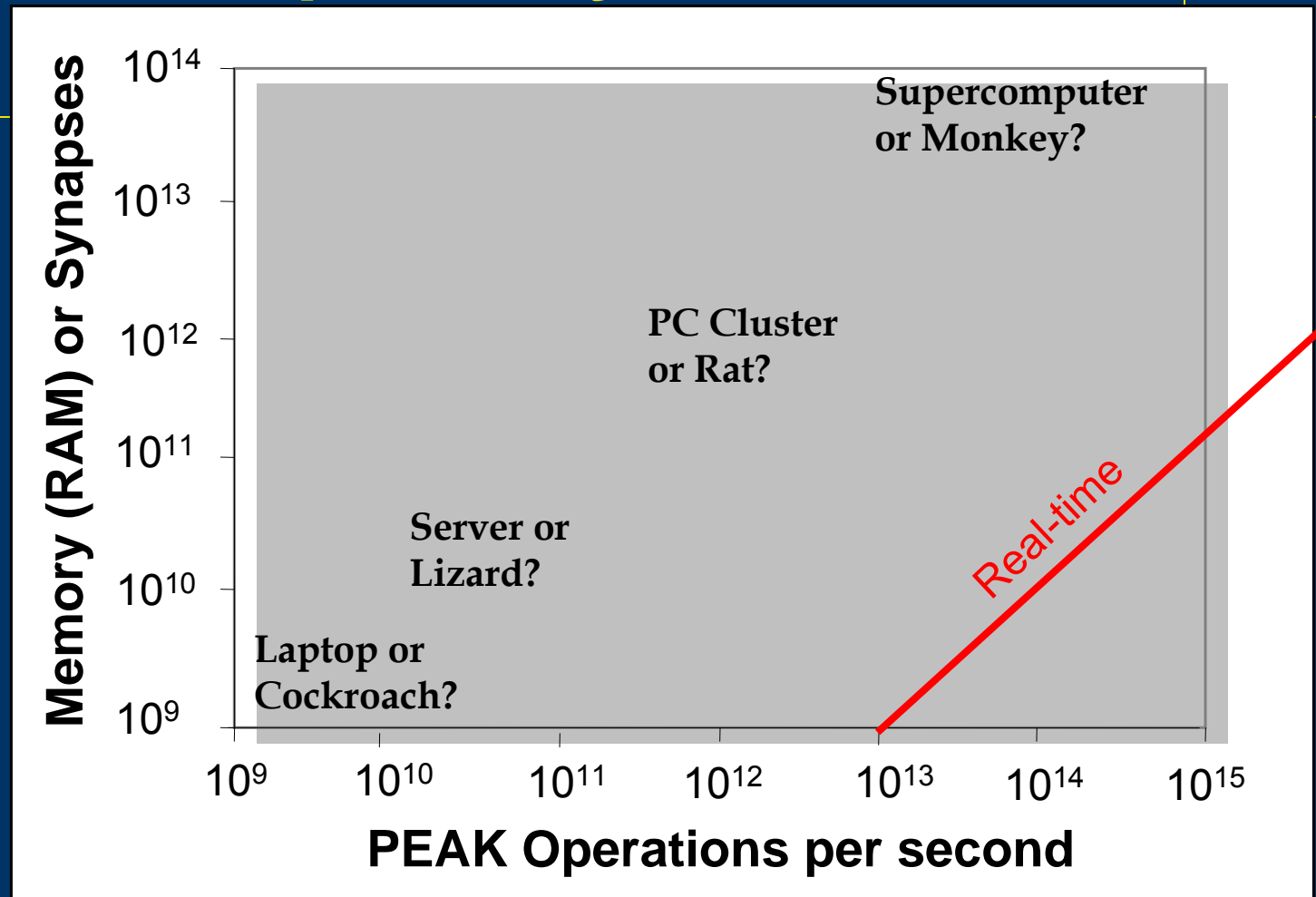
If you need to use more than ~8 processors or more than ~128 GB RAM, then you probably need to use MPI.

But if you have LOTS of money (\$4M), you could go to 64 processors and 2 TB RAM without using MPI.

Range of Computer Systems

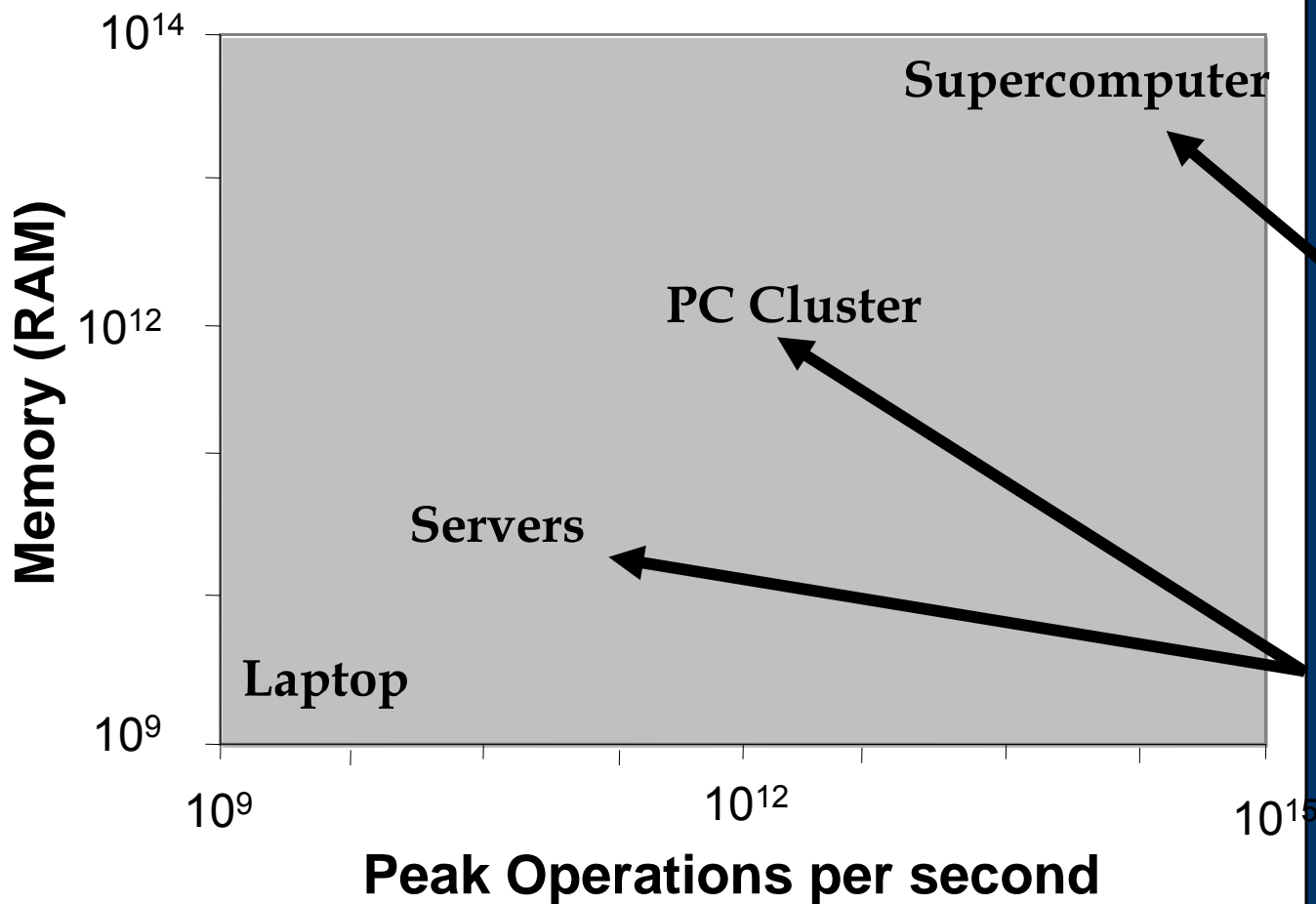


If you have NN software that requires ~1 byte per synapse, then this axis can represent the max number of synapses that you can fit in memory



If you have NN software that requires ~1 operation per synapse/timestep, then this axis represents the max number of timesteps / second

Range of Computer Systems



- Often U.S. citizen only
- Security Checks
- SecurId cards
- Complex login
- Batch Processing
- Queuing system
- Graphics difficult
- Can't install software or compilers
- Remote access
- Often Limited to small no. of nodes
- Very difficult for code development

- Useful for MPI code development

Supercomputer Centers in U.S.



- DOD: <http://www.hpcmo.hpc.mil/> :
 - Maryland: <http://www.arl.hpc.mil/>
 - Mississippi: <http://www.ercd.hpc.mil/>
 - Mississippi: <http://www.navo.hpc.mil/>
 - Ohio: <http://www.asc.hpc.mil/>
- NSF:
 - San Diego: <http://www.sdsc.edu/>
 - Illinois: <http://www.ncsa.uiuc.edu/>
 - Pittsburgh: <http://www.psc.edu/>
- DOE:
 - Argonne: <http://www.alcf.anl.gov/>
 - LLNL: https://asc.llnl.gov/computing_resources/
 - LANL: <http://www.lanl.gov/orgs/hpc/index.shtml>
- Caltech:
 - <http://citerra.gps.caltech.edu/> (512 nodes: each node is Xeon dual quad-core)
 - <http://www.cacr.caltech.edu/>
- Other: NSA, CIA, ORNL, Sandia, NERSC, MHPCC, LBNL, NASA Ames, NRO, ...

If you have DOD grants or contracts you can use these.

You can write proposals to get access to these.

More difficult to access these



Inexpensive 8-Processor Server

- Systemax at www.tigerdirect.com
 - Dual quad-core Intel Xeon processors
 - 8 cores (or processors)
 - 1.6 GHz (but can get 3.2 GHz)
 - 4 GB RAM, but can go to 16 GB
 - Supermicro X7DVL-E motherboard
 - **(the X7DWN+ motherboard supports 128 GB RAM)**
 - Dual gigabit ethernet
 - 600W and can have 6 fans
 - Software:
 - 64-bit Suse Linux OS
 - Java, C++, Matlab
 - MPI
 - \$ 2000
- } Free!
- (to get 16 GB RAM and 3.2 GHz processors would cost \$3000)



Screen Shot from Dual Quad-Core



The screenshot shows a Linux desktop with a MATLAB 7.5.0 (R2007b) window and a System Monitor window. The MATLAB Command Window contains the following code and output:

```
error =  
-1.5509e-13  
>> clear  
>> matrix1  
nx =  
5000  
Elapsed time is 16.559016 seconds.  
time =  
16.5592  
ops =  
8.3333e+10  
megaflops =  
5.0324e+03  
gigabytes =  
1.6000  
error =  
-2.7309e-13  
>> clear  
>>
```

The System Monitor window shows the following data:

CPU History:

CPU1: 0.0%	CPU2: 1.0%	CPU3: 18.2%	CPU4: 0.0%
CPU5: 0.0%	CPU6: 95.0%	CPU7: 0.0%	CPU8: 8.9%

Memory and Swap History:

User memory: 1.9 GiB of 3.9 GiB 48.4%
Used swap: 0 bytes of 2.0 GiB 0.0%

Network History:

Received: 0 bytes/s Total: 226.8 KiB
Sent: 0 bytes/s Total: 59.9 KiB

Red arrows point from the MATLAB Command Window to the System Monitor graphs, indicating the start and end of the MATLAB code execution.

Java-based NN code started here

Matlab code started here

matlab code ends

matlab clear



Apple Mac Pro



- Dual quad-core Intel Xeon 5400's
 - 8 processors
 - 2.8 – 3.2 GHz
 - 64-bit
- Up to 32 GB RAM
- \$ 12,000 with 32 GB



For Comparison: Dell & IBM Servers



Dell PowerEdge 6800

- Quad dual-core Xeon processors
 - 8 cores (processors)
 - 3.2 GHz
- 64 GB RAM
- Software:
 - 64-bit Suse Linux OS
 - Java, C++, Matlab
- \$ 27,000

} Free!

IBM P-595

- In 2006 Penn State got:
 - IBM P-570
 - 12 Power5 Proc.
 - 100 GB RAM
 - \$ 500,000 in 2006
- Could buy today:
 - IBM P 595
 - 64 Power5+ proc.
 - 2000 GB RAM (2 TB RAM !)
 - \$ 4,000,000

(These are really amazing machines, and should not really be compared to PC's. These are incredibly reliable and could support thousands of users.)



For Comparison: PC Cluster

- You could also build your own cluster
- For example:
 - 48 dual quad-processor PC's (384 processors)
 - Peak speed of ~300 gigaflops
 - 800 GB RAM
 - Simple gigabit ethernet network switch (\$3K)
 - \$ 150,000 ?
 - Linux, MPI, c/c++, ...
- Would need a server front-end for user disk storage and login
- Someone would need to run / manage it (not trivial)

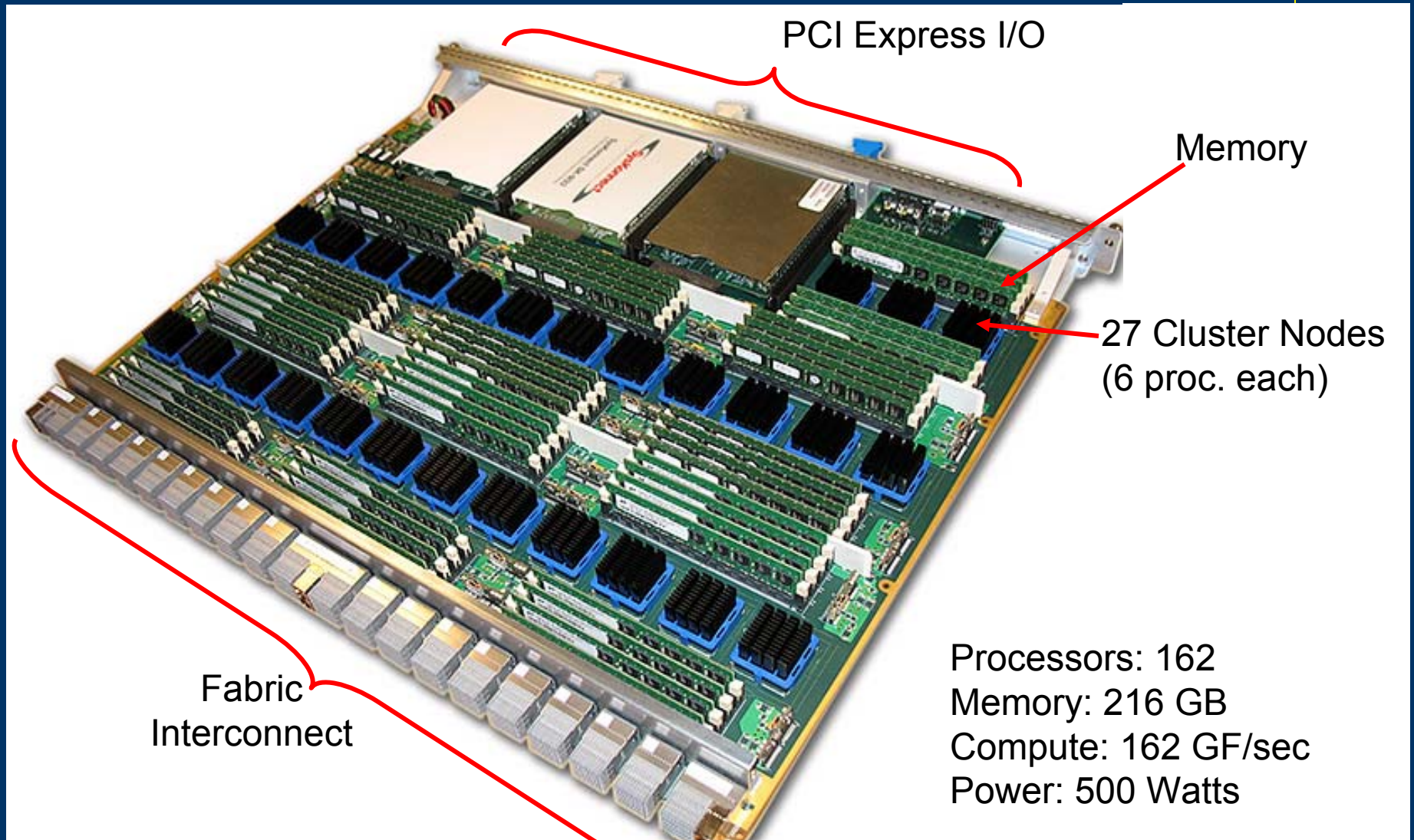
New HPC Company (www.SiCortex.com)



- Formed by Thinking Machines, DEC, and Cray people
- Linux, MPI, C/C++,...
- Lower Watts/Gigaflop (3) compared to PC Clusters (10)
- SC-648 model:
 - 648 500-MHz processors
 - 648 gigaflops (peak) in one rack
 - 900 gigabytes memory
 - \$ 180 K
- SC-72 model:
 - 72 500-MHz processors
 - 72 gigaflops (peak)
 - 48 gigabytes memory
 - \$ 15 K
- They've offered to present a seminar



SiCortex CPU Module



Summary of Some Computers



Name (# proc.)	Machine Type	Memory (GB)	Peak Speed (Gflops)	Price (\$)
Dual Quad Server (8)	SMP	16	50	\$ 3 K
Dell Server (8)	SMP	64	50	\$ 27 K
PC Cluster (96)	Distributed Memory	800	300	\$ 150 K
SciCortex (648)	Distributed memory	900	648	\$ 180 K (72 proc. for \$15K)
IBM Server (64)	SMP	2,000	300	\$ 4,000 K ?
IBM BlueGene (200,000)	Distributed memory	74,000	600,000	\$ 200,000 K ?



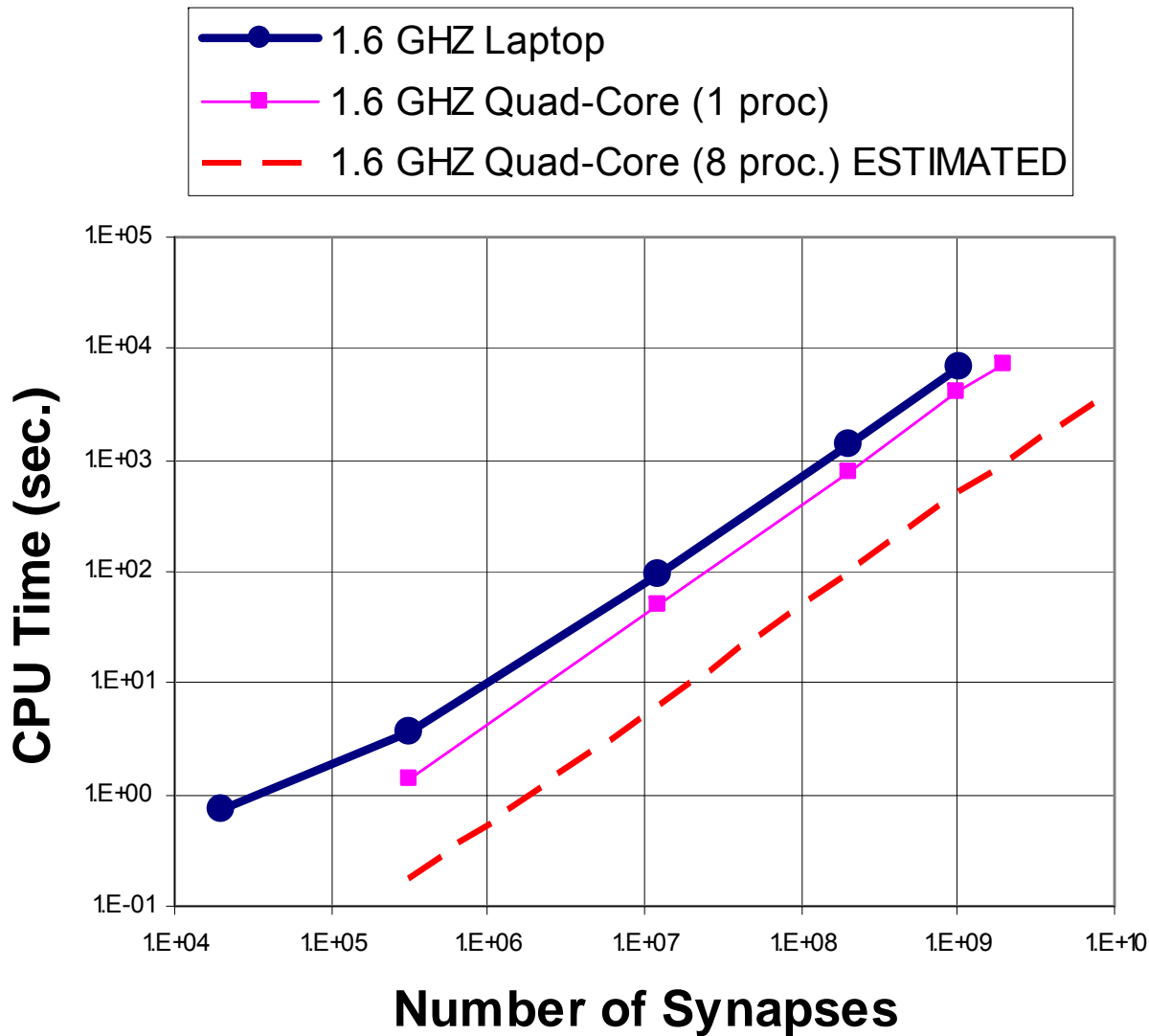
Some Results



New LIF NN Code

- I'm developing this code now, will just show performance results here
- Java based
- Object oriented: Neuron, Layer, and Network objects
- Feed-forward layered network (but could do recurrent)
- Arbitrary neuron connections between layers (all-to-all, stencil, ...)
- Network input coupled to webcam
- Hebbian learning
- Hoping to use this for object recognition
- This will also be developed in C++/MPI for massively parallel computers
- Recent conference paper discussing initial software development:
 - <http://www.personal.psu.edu/lnl/papers/aiaa20080885.pdf>
- Paper on massively parallel rate-based neural networks:
 - Long & Gupta, www.aiaa.org/jacic, Vol. 5, Jan., 2008

Neural Network Code Performance on One Processor (500 time steps or 0.1 sec.)



3 layers of 2-D
arrays of neurons

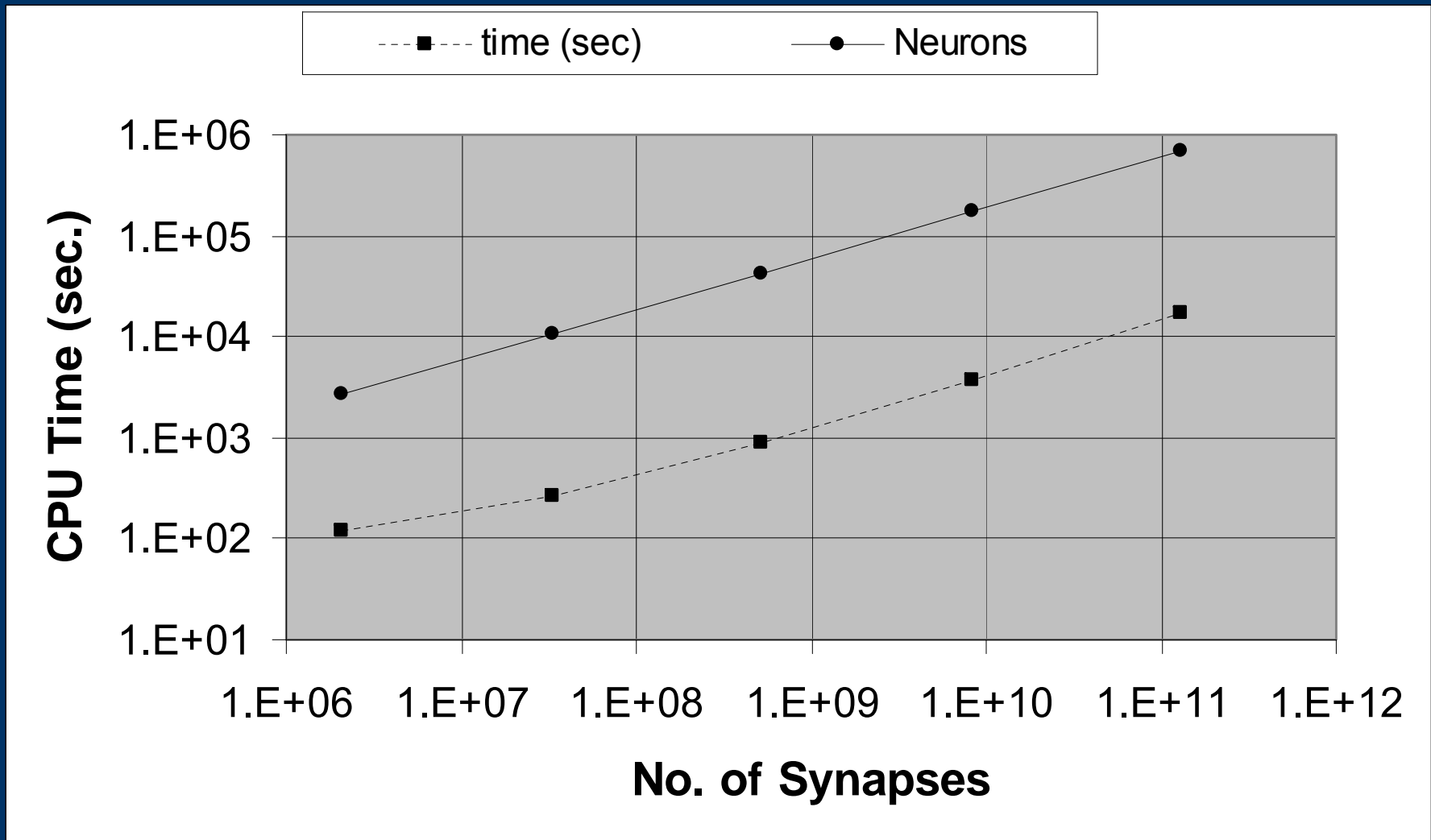
$N * N$ neurons per
layer

Synapses $\approx N^4$

(300 to 94,000
total neurons)

← “Real-time”

Benchmarking of Ali Soltani's Code (LIF using FFT's & Matlab)

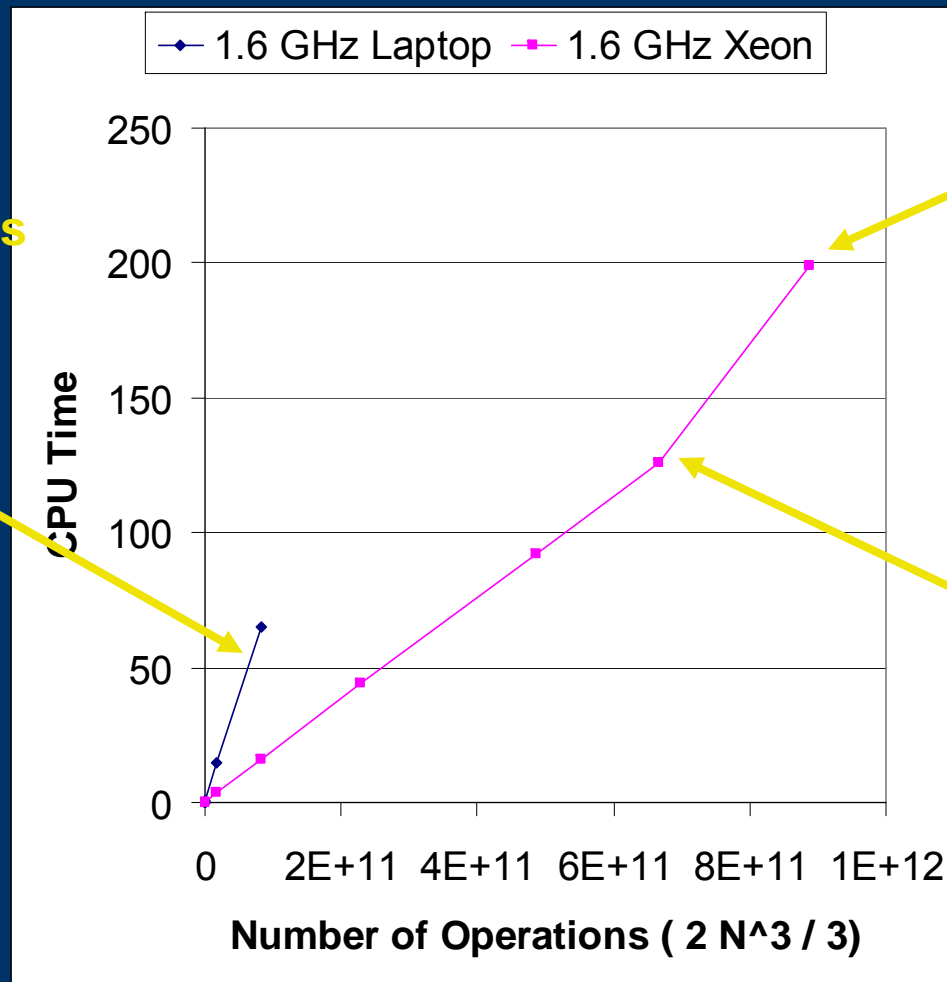


Gaussian Elimination using Matlab and One Xeon Processor



1.6 GHz LAPTOP:
5,000 x 5,000 matrix
0.2 GB for matrix
0.1 trillion operations
65 CPU seconds
1300 megaflops

Shows more diff. between laptop and Xeon since it this problem more effectively uses processors



Started using virtual memory, so performance was reduced

1.6 GHz Xeon:
10,000 x 10,000 matrix
~1 GB for matrix
~1 trillion ops
126 CPU seconds
5300 megaflops

No. of Eqtns: 100 7000 9000 11,000



Conclusions

- 64-bit operating systems finally allow us to have more than 4 GB RAM in desktops and laptops
- Multi-core chips will require new approaches to software development
- Its easy to build small PC clusters
- Very large SMP machines are very expensive
- One new exciting massively parallel computer (SiCortex)
- If you want to try this 8-proc. machine, just let me know (LNL@caltech.edu)



References

- computing.llnl.gov/tutorials/parallel_comp/
- www.top500.org
- www.sicortex.com
- www.beowulf.org
- www.personal.psu.edu/lnl/
- www.csci.psu.edu (grad minor in Computational Sci.)
- Books:
 - “Parallel Computing in C++ and MPI,” Karniadakis & Kirby
 - “Parallel Programming with MPI,” Pacheco
 - “Java for Engineers and Scientists,” Chapman
 - “C++ for Scientists and Engineers,” Yang