

Outline

- Course Administration 18.337/6.338/SMA5505
- Parallel Machines in 2003
 - Overview
 - Details
- Applications
- Special Approaches
- Our Class Computer

Parallel Computer Architectures

- MPP – Massively Parallel Processors

- Top of the top500 list consists of mostly mpps but clusters are “rising”

- Clusters are there!

- Earth Simulator: Old-old style making news again

- ASCI Machines: Big companies, special purpose

- Beowulf Clusters: Popping up everywhere

- Software

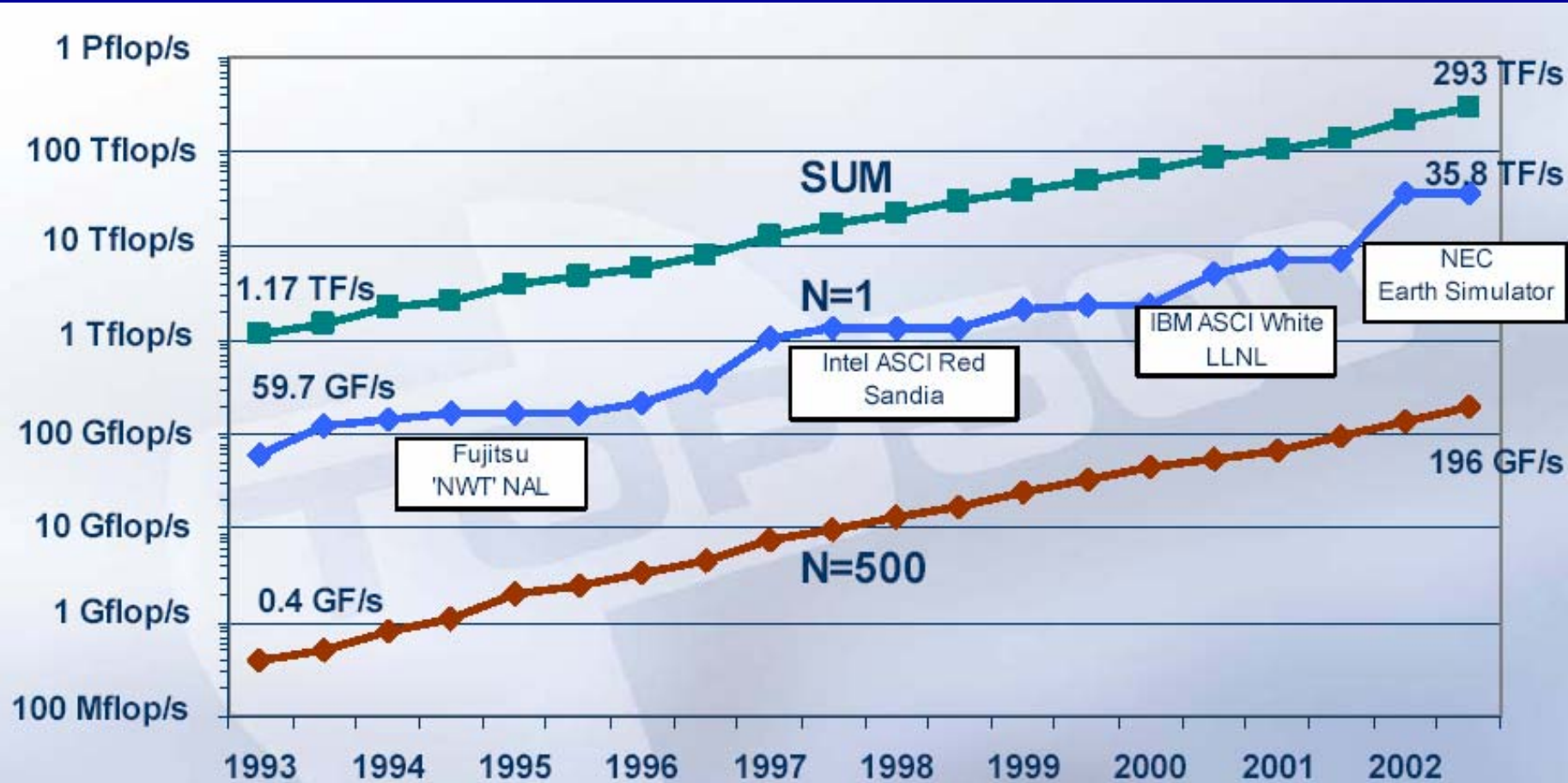
- Embarrassingly parallel or sacrifice a grad student

- MATLAB*p (our little homegrown project)

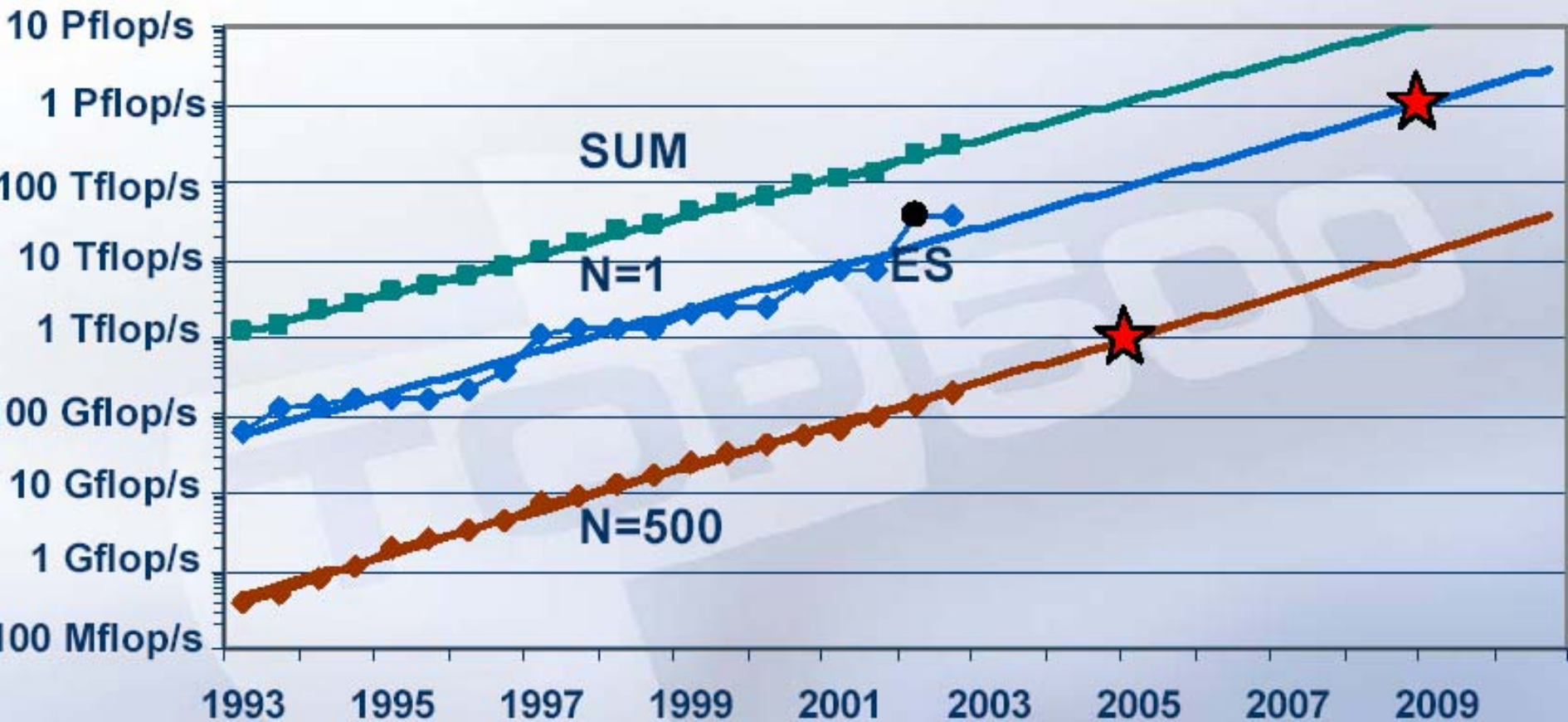
Always interesting to check out <http://www.top500.org/dlist/2002/11/>

Rank	Manufacturer	Computer	R_{\max} [TF/s]	Installation Site	Country	Year	Area of Installation	# Proc
1	NEC	Earth-Simulator	35.86	Earth Simulator Center	Japan	2002	Research	5120
2	HP	ASCI Q, AlphaServer SC	7.73	Los Alamos National Laboratory	USA	2002	Research	4096
3	HP	ASCI Q, AlphaServer SC	7.73	Los Alamos National Laboratory	USA	2002	Research	4096
4	IBM	ASCI White SP Power3	7.23	Lawrence Livermore National Laboratory	USA	2000	Research	8192
5	Linux NetworX	MCR Cluster – Quadrics	5.69	Lawrence Livermore National Laboratory	USA	2002	Research	8192
6	HP	AlphaServer SC ES45 1 GHz	4.46	Pittsburgh Supercomputing Center	USA	2001	Academic	3016
7	HP	AlphaServer SC ES45 1 GHz	3.98	Commissariat a l'Energie Atomique (CEA)	France	2001	Research	2560
8	HPTi	Xeon Cluster - Myrinet2000	3.34	Forecast Systems Laboratory - NOAA	USA	2002	Research	1536
9	IBM	pSeries 690 Turbo	3.16	HPCx	UK	2002	Academic	1280
10	IBM	pSeries 690 Turbo	3.16	NCAR (National Center for Atmospheric Research)	USA	2002	Research	1216

Performance Trends



Extrapolations





Beowulf Clusters



Current Beowulfs (2)



Parallel Computer Architectures

- MPP – Massively Parallel Processors
 - Top of the top500 list consists of mostly mpps but clusters are “rising”
- Clusters
 - “simple” cluster (1 processor in each node)
 - Cluster of small smp’s (small # processors / node)
 - Constellations (large # processors / node)
- Older Architectures
 - SIMD – Single Instruction Multiple Data (CM2)
 - Vector Processors (Old Cray machines)

Architecture Details

1. **MPPs** are built with specialized networks by vendors with the intent of being used as a parallel computer. **Clusters** are built from independent computers integrated through an aftermarket network.
 - Buzzwords: “COTS” Commodity off the shelf components rather than custom architectures.
 - Clusters are a market reactions to MPPS with the thought of being cheaper
 - Originally considered to have slower communications but are catching up.

More details

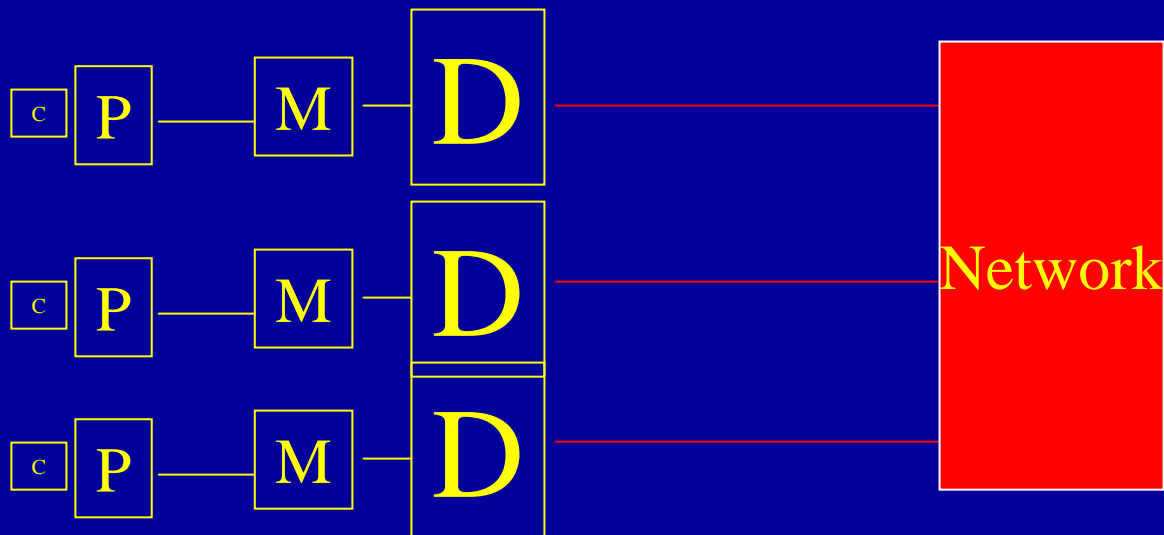
- 2. “NOW”-Networks of Workstations
- Beowulf (Goddard in Greenbelt MD) –
- Clusters of a small number of pc’s
- (pre 10th century poem in Old English about a Scandinavian warrior from the 6th century)

The First Beowulf (2)

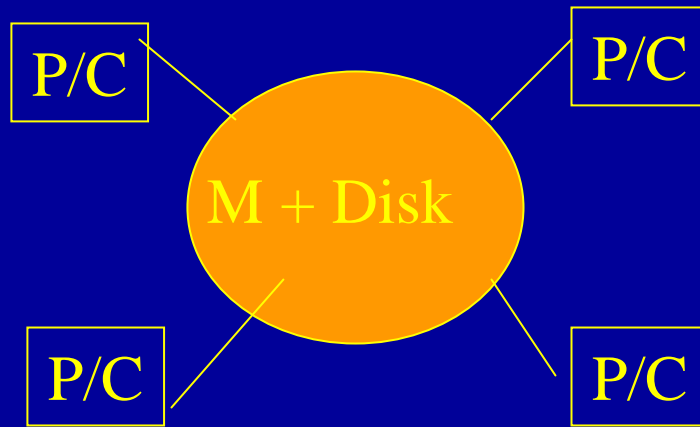


Architecture Details

- Computers → MPPs



Architecture Details SMP (Symmetric Multiprocessor)



NUMA – Non uniform memory access

4. Constellation: Every node is a large smp

More details

- 5) SIMD – SIngle instruction multiple data
 - data parallel vs mimd
- 6: Speeds:
 - Megaflops 10^6 flops
 - Gigaflops 10^9 flops workstations
 - Teraflops 10^{12} top 17 supercomputers
by 2005 every supercomputer in the top 500
 - Petaflops 10^{15} 2010?

Trends

- Moore's Law: Number of Transistors per square inch in an integrated circuit doubles every 18 months
- Every decade – computer performance increases 2 order of magnitude

Applications of Parallel Computers

- Traditionally: government labs, numerically intensive applications
- Research Institutions
- Recent Growth in Industrial Applications
 - 236 of the top 500
 - Financial analysis, drug design and analysis, oil exploration, aerospace and automotive

Goal of Parallel Computing

- Solve bigger problems faster
- Often bigger is more important than faster
- P-fold speedups not as important!

Challenge of Parallel Computing

Coordinate, control, and monitor the computation

Easiest Applications

- Embarrassingly Parallel – Lots of work that can be dived out with little coordination or communication
- Example: integration, Monte Carlo methods, Adding numbers

Special Approaches

- Distributed Computing on the internet
 - signal processing 15 Teraflops
 - Distributed.net factor product of 2 large primes
 - Parabon – biomedical, protein folding, gene expression
- Akamai Network – Tom Leighton, the late Danny Lewin
- Thousands of servers spread globally that caches web pages and routes traffic away from congested areas
- Embedded Computing : Mercury (inverse to the worldwide distribution)

The Computational Grid

- Computers are everywhere. Can we plug into the power? Challenges: managing machines that are distributed, might fail, be turned off, or there might be malicious intent.

MATLAB*P demo