High Performance Computing

Lecture 4: Introduction

Nayda G. Santiago

August 15, 2008

Administrative Details

- Make up class for Monday Aug 11, 2008 class
  - Reserved S 203 for Thursday, Aug 21, 2008
  - 5:30pm to 6:30pm
- Additional details
  - Traveling from Sept 30 to Oct 4, Colorado – 2 lectures
  - Traveling from Oct 9 to Oct 12, Houston – 1 lecture
  - Traveling from Oct 20 to Oct 22, Ann Arbor – 2 lectures
  - Traveling from Oct 27 to Oct 29, Boston – 2 lectures
- We need to schedule additional make up classes

Basic Terminology and Concepts

- The definitions are fuzzy, many terms are not standardized, definitions often change over time.
- Many algorithms, software, and hardware systems do not match the categories, often blending approaches.
- No attempt to cover all models and aspects of parallel computing. For example, quantum computing not included.

Parallel Computing Thesaurus

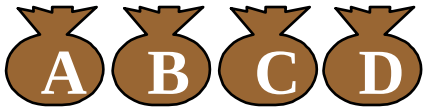
- **Embarrassingly Parallel**
  - Solving many similar, but independent, tasks. E.g., parameter sweeps. Also called *farming*.
- **Parallel Computing**
  - Solving a task by simultaneous use of multiple processors, all components of a unified architecture.
- **Symmetric Multiprocessing (SMP)**
  - Multiple processors sharing a single address space and access to all resources.
- **Multi-core Processors**
  - An SMP with multiple processors (cores) in a single chip. Also known as *many-core*. Heterogenous multi-core chips are being developed.
- **Cluster Computing**
  - Hierarchical combination of commodity units (processors or SMPs) to build parallel system.

Parallel Computing Thesaurus

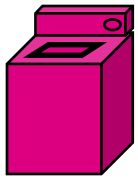
- **Constellation**
  - A combination of clusters.
- **Supercomputing**
  - Use of the fastest, biggest machines to solve large problems. Historically vector computers, but now are parallel or parallel/vector.
- **High Performance Computing**
  - Solving problems via supercomputers + fast networks + visualization.
- **Pipelining**
  - Breaking a task into steps performed by different units, with inputs streaming through, much like an assembly line.
- **Vector Computing**
  - Use of vector processors, where operation such as multiply broken into several steps and applied to a stream of operands (“vectors”).

Pipelining is Natural!

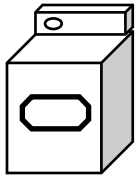
- **Laundry Example**
- **Ann, Brian, Cathy, Dave each have one load of clothes to wash, dry, fold, and put away**



- **Washer takes 30 minutes**



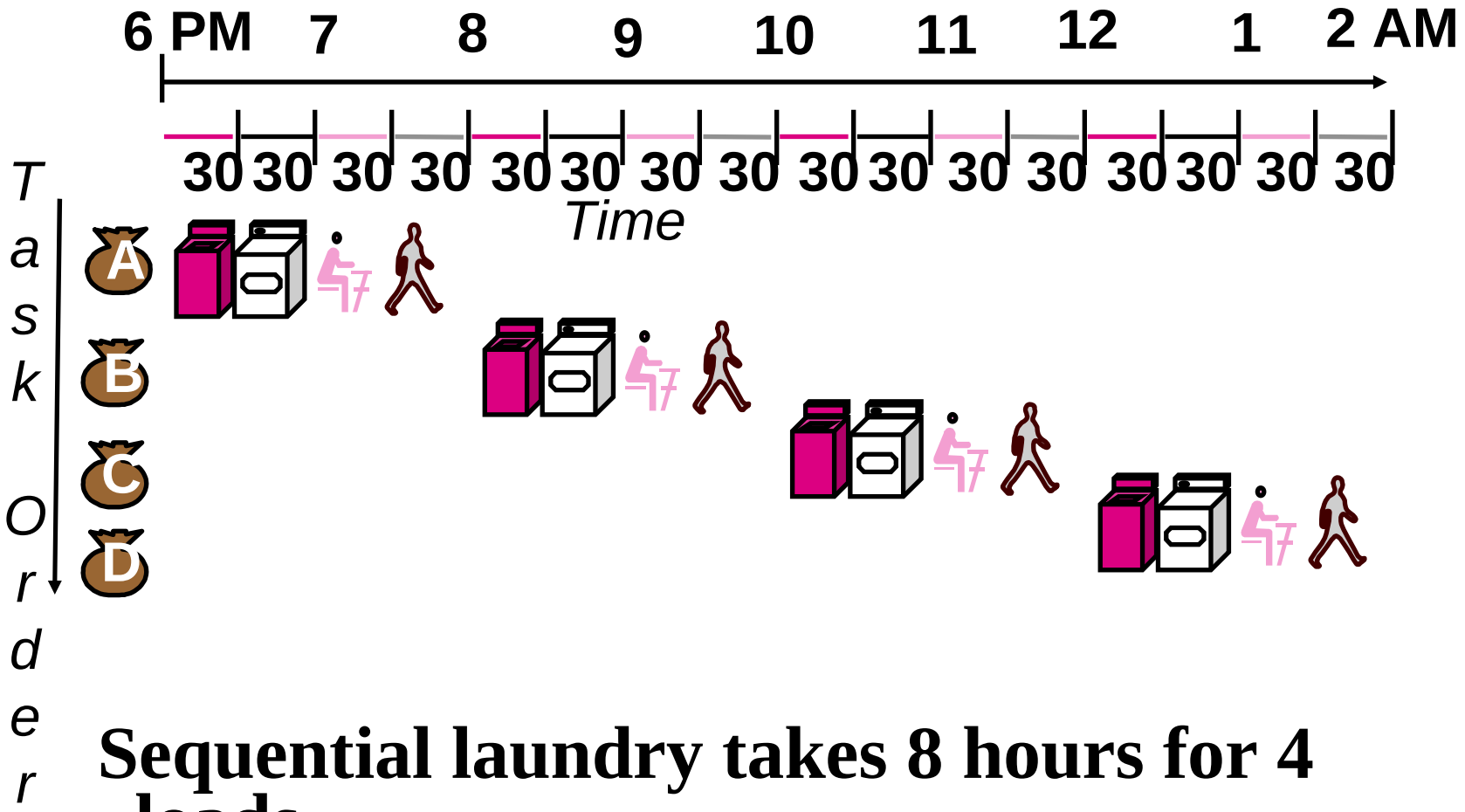
- **Dryer takes 30 minutes**



- **“Folder” takes 30 minutes**

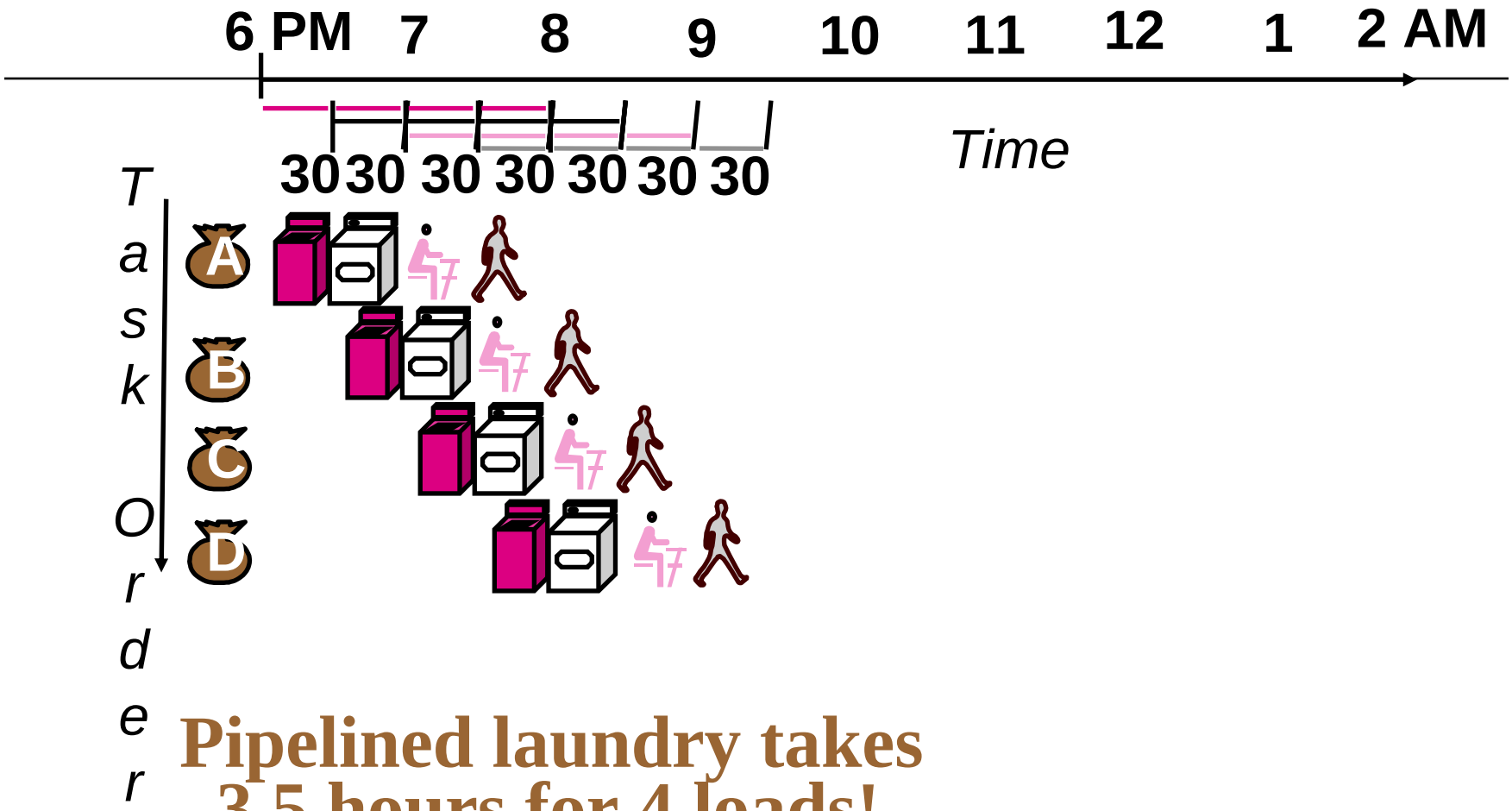
- **“Stasher” takes 30 minutes to put clothes into drawers**

Sequential Laundry



Sequential laundry takes 8 hours for 4 loads (Example taken from Patterson & Hennessey book)

Pipelined Laundry: Start work ASAP



Pipelined laundry takes 3.5 hours for 4 loads!

Who uses supercomputers?

- Historically, the military (nuclear simulations, cryptography).
- Weather forecasting was the civilian application.
- These continue to be major users but now many more civilian users.

Top 500

- A list of the 500 most powerful computer systems.
  - Started 1993
  - Compiled twice a year

http://www.top500.org/

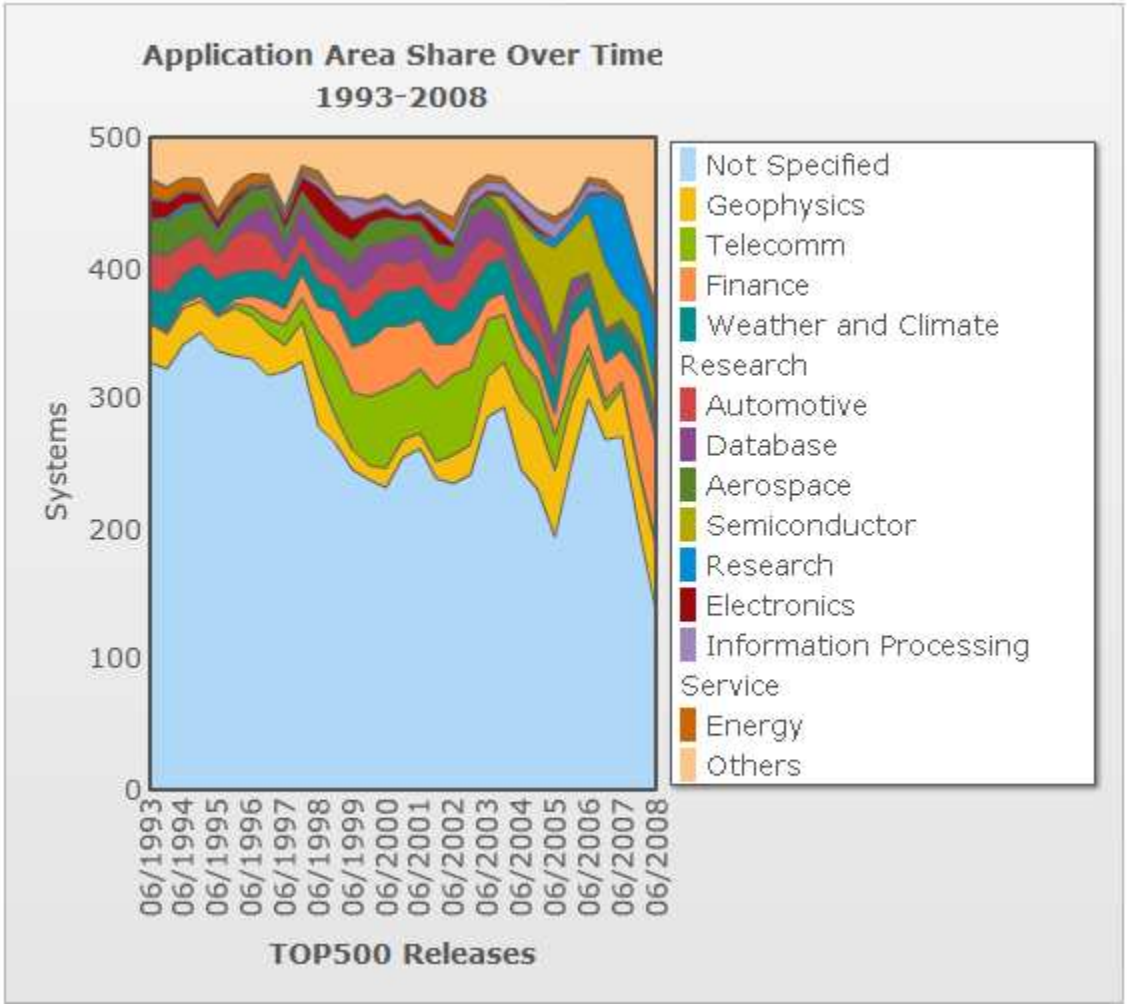
- List computers ranked by their performance on the LINPACK Benchmark.
  - Solve a dense system of linear equations

The following charts are from the Top 500 list, showing the status as of June 08

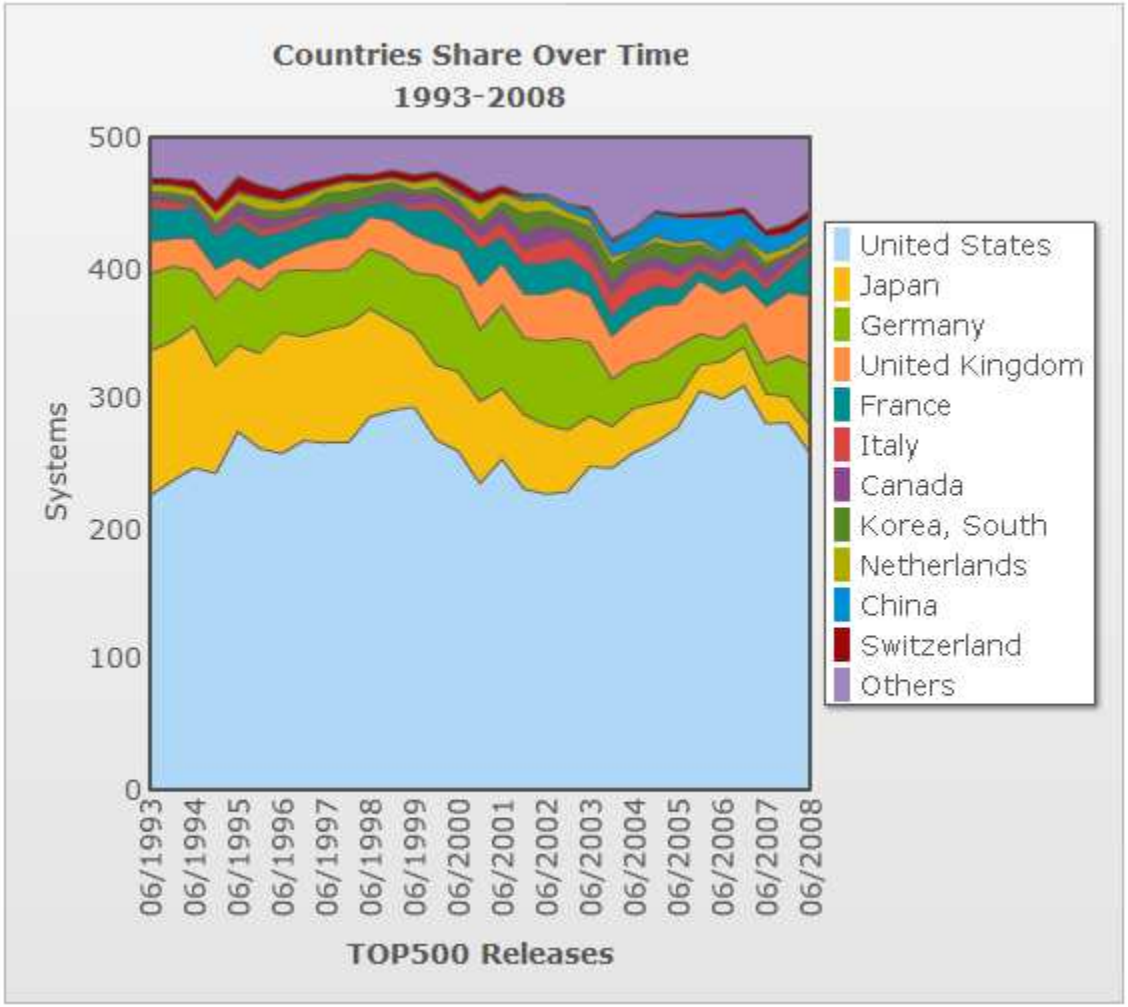
Top 500 performance

Top 500 Projected Performance

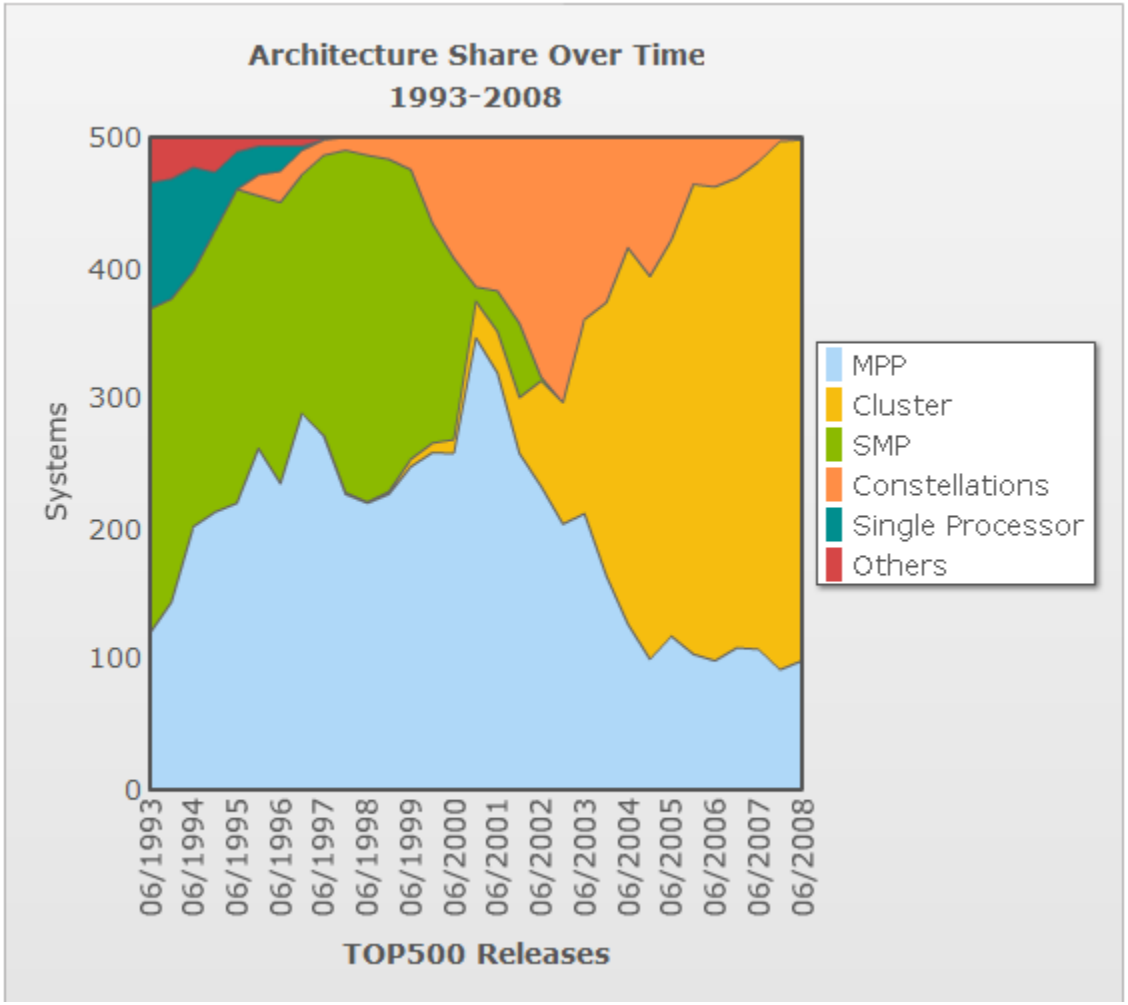
Top 500: Application Systems



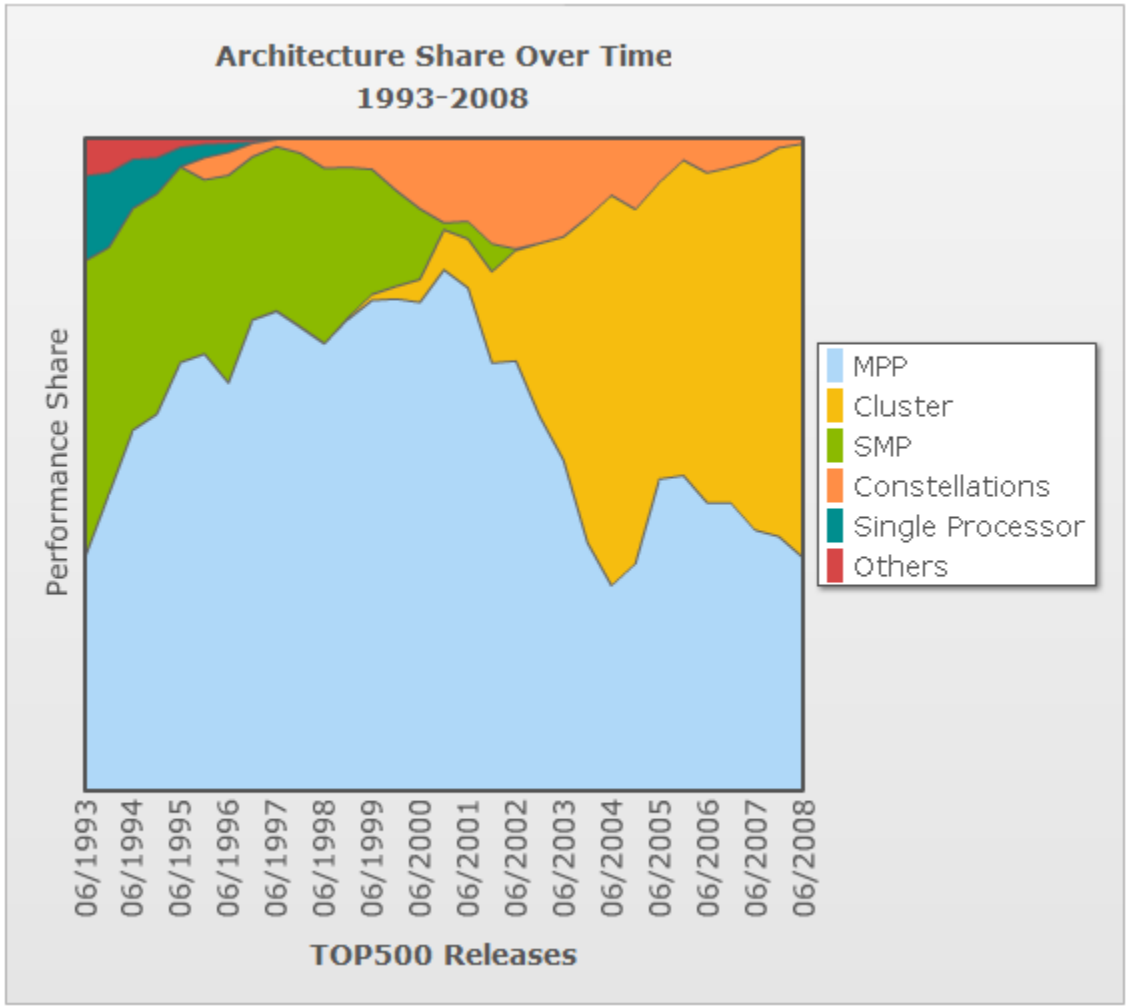
Top 500: Countries



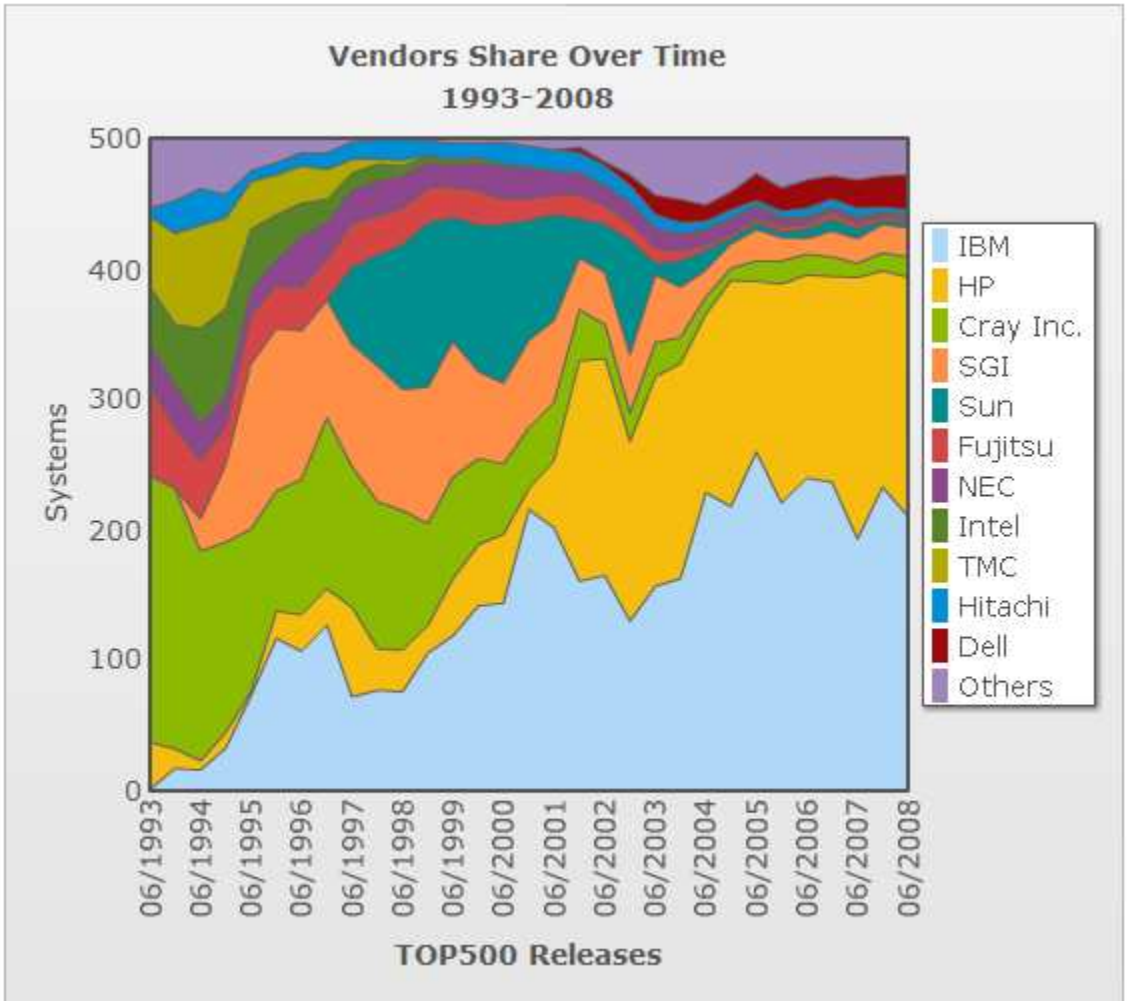
Top 500: Architecture



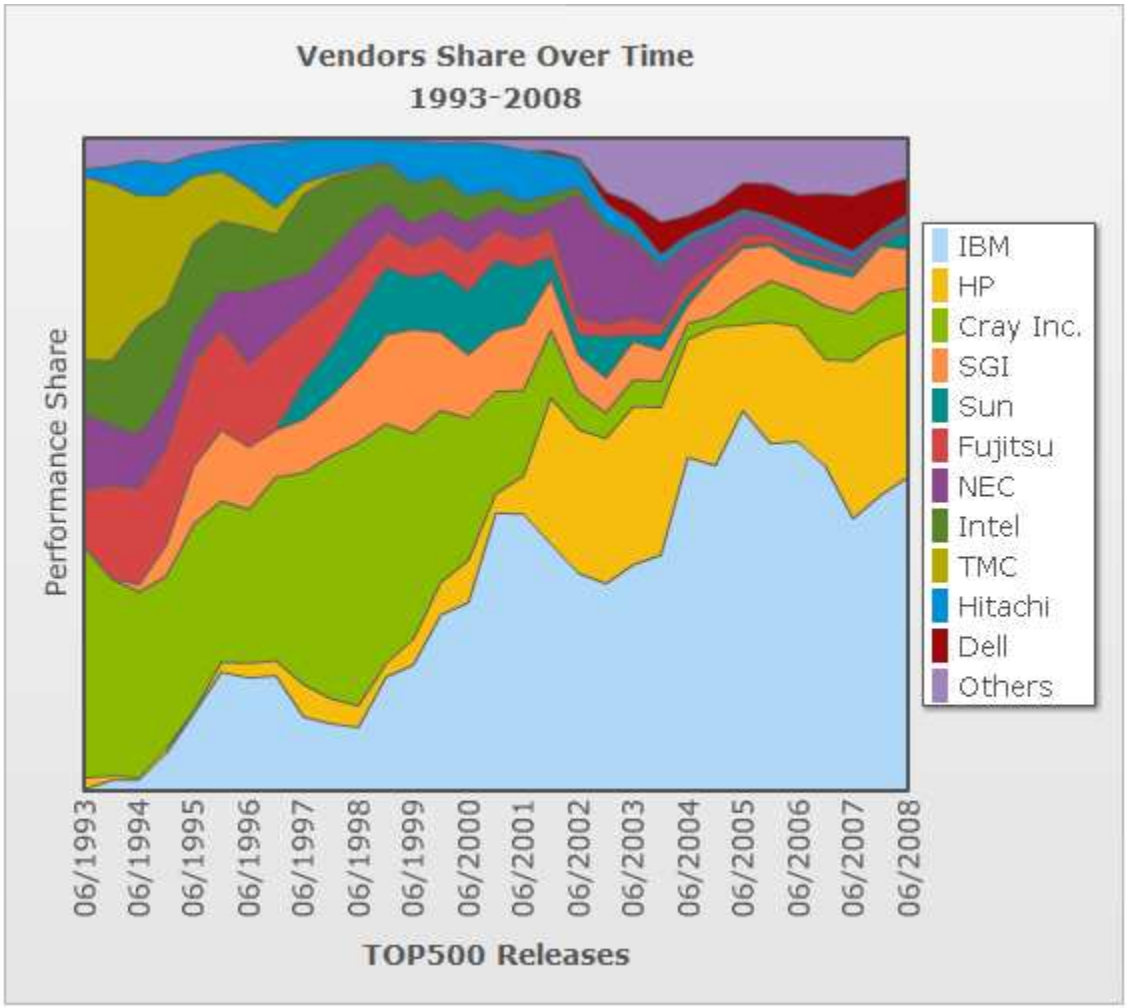
Top 500: Architecture Performance



Top 500: Vendors



Top 500: Vendors performance



Measuring

- A “flops” is an acronym meaning floating-point operations per second.
- One petaflop/s is 1,000 trillion operations per second.

Top System as of June 08

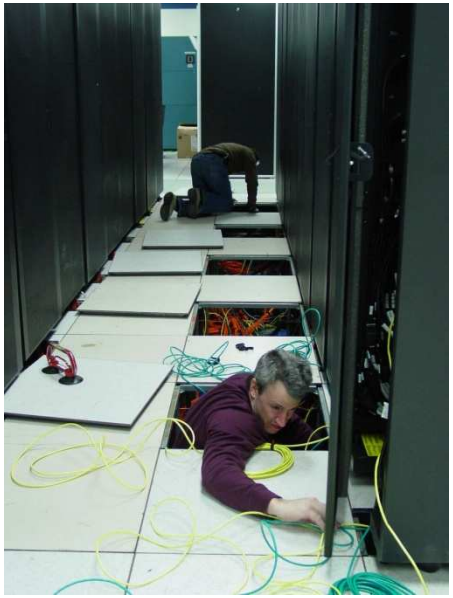
- Built by IBM for the U.S. Department of Energy's Los Alamos National Laboratory and called "Roadrunner,"
  - Achieved performance of 1.026 petaflop/s
    - To put this into perspective, if each of the 6 billion people on earth had a hand calculator and worked together on a calculation 24 hours per day, 365 days a year, it would take 46 years to do what Roadrunner would do in one day.
    - The first supercomputer ever to reach this milestone.
    - Roadrunner is also one of the most energy efficient systems on the TOP500.

Roadrunner

- System Name
  - **Roadrunner**
- Site
  - Department of Energy's National Nuclear Security Administration, Los Alamos National Laboratory
  - IBM
- Cluster System Model
  - BladeCenter QS22 Cluster
- Computer
  - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband
- Vendor
  - IBM
- Application area
  - Not Specified
- Installation Year
  - 2008
- Operating System
  - Linux
- Interconnect
  - Infiniband
- Processor
  - PowerXCell 8i 3200 MHz (12.8 GFlops)

Roadrunner


Roadrunner first to run at petaflop/s speed!




Roadrunner 10^{15}
First to sustained Petaflop/sec
IBM

Top 10 sites: Top 500

TOP 10 Sites for June 2008

For more information about the sites and systems in the list, click on the links or view the complete list.

Rank	Site	Computer
1	DOE/NNSA/LANL United States	Roadrunner - BladeCenter QS22/LS21 Cluster, PowerXCell 8i 3.2 Ghz / Opteron DC 1.8 GHz , Voltaire Infiniband IBM
2	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution IBM
3	Argonne National Laboratory United States	Blue Gene/P Solution IBM
4	Texas Advanced Computing Center/Univ. of Texas United States	Ranger - SunBlade x6420, Opteron Quad 2Ghz, Infiniband Sun Microsystems
5	DOE/Oak Ridge National Laboratory United States	Jaguar - Cray XT4 QuadCore 2.1 GHz Cray Inc.
6	Forschungszentrum Juelich (FZJ) Germany	JUGENE - Blue Gene/P Solution IBM
7	New Mexico Computing Applications Center (NMCAC) United States	Encanto - SGI Altix ICE 8200, Xeon quad core 3.0 GHz SGI
8	Computational Research Laboratories, TATA SONS India	EKA - Cluster Platform 3000 BL460c, Xeon 53xx 3GHz, Infiniband Hewlett-Packard
9	IDRIS France	Blue Gene/P Solution IBM
10	Total Exploration Production France	SGI Altix ICE 8200EX, Xeon quad core 3.0 GHz SGI

References

- Most of the material is taken from “Parallel Computing 101, by Quentin F. Stout and Christiane Jablonowski, University of Michigan, Supercomputing 2007 tutorial on Sunday, Nov 11, 2007, 8:30am to 5:00pm, Reno, Nevada”
- www.top500.org