CIS 431/531 Parallel Computing Introduction

Prof. Boyana Norris
Department of Computer and Information Science
Spring 2017
Outline

- Course Overview
  - What is CIS 431/531
  - What is expected of you?
  - What will you learn in CIS 431/531?

- Parallel Computing
  - What is it?
  - What motivates it?
  - Trends that shape the field
  - Large-scale problems and high-performance
  - Parallel architecture types
  - Scalable parallel computing and performance
Course Logistics

- Lecture and final times
  - Lecture: Tuesday, Thursday: 8:30 – 9:50
  - No Final, the only exam is on Tuesday of Week 10

- Undergraduate course prerequisite
  - CIS 330 (C/C++ and Unix)

- Webpage:
  - [http://www.cs.uoregon.edu/Classes/17S/cis431](http://www.cs.uoregon.edu/Classes/17S/cis431)

- Canvas
  - Do Assignment 0 (no points, but get it done ASAP)
How did the idea for CIS 431/531 originate?

- Prior to 2 years ago, there has never been an undergraduate course in parallel computing at UO
  - Only 1 course taught at the graduate level (CIS 631)
- Goal was to bring parallel computing education to CIS undergraduate curriculum, starting at senior level
  - CIS 410/510 (Spring 2014/15/16, “experimental” course)
  - CIS 431/531 (AY 2016/17, “new” course, real course #)
- CIS 607 – Parallel Computing Course Development
  - Winter 2014 seminar to plan undergraduate course
  - Develop 410/510 materials, exercises, labs, …
- Intel gave a generous donation ($100K) to the effort!!!
- Created Intel Parallel Computing Center (IPCC)
  
  http://ipcc.cs.uoregon.edu
Who’s involved?

- Instructor
  - Boyana Norris
    - High-performance computing
    - Automated software analysis and transformation
    - Performance analysis and optimization

- GTF
  - Daniel Ellsworth

- Faculty colleagues and course co-designers
  - Allen D. Malony
    - scalable parallel computing
    - parallel performance analysis
    - taught CIS 631 for many years
  - Hank Childs
    - Large-scale, parallel scientific visualization
    - Visualization of large data sets

- Intel scientists (next)
Intel Partners

- James Reinders
  - Director, Software Products
  - Multi-core Evangelist

- Michael McCool
  - Software architect
  - Former Chief scientist, RapidMind
  - Adjunct Assoc. Professor, University of Waterloo

- Arch Robison
  - Architect of Threading Building Blocks
  - Former lead developers of KAI C++

- David MacKay
  - Manager of software product consulting team

Presents parallel programming from a point of view of patterns relevant to parallel computation

Focuses on the use of shared memory parallel programming languages and environments
  - Intel Thread Building Blocks (TBB), Intel Cilk Plus, OpenMP
Other Reference Textbooks

- *Introduction to Parallel Computing*, A. Grama, A. Gupta, G. Karypis, V. Kumar, Addison Wesley, 2nd Ed., 2003
  - Lecture slides from authors online
  - Excellent reference list at end
  - Used for CIS 631 before
  - Getting old for latest hardware

- *Designing and Building Parallel Programs*, Ian Foster, Addison Wesley, 1995.
  - Entire book is online!!!
  - Historical book, but very informative

- *Patterns for Parallel Programming*, T. Mattson, B. Sanders, B. Massingill, Addison Wesley, 2005.
  - Targets parallel programming
  - Pattern language approach to parallel program design and development
  - Excellent references
Course Plan

- Cover main areas of parallel computing in the lectures
  - Architecture (1 week)
  - Performance models and analysis (2 weeks)
  - Programming patterns (paradigms) (5 weeks)
  - Algorithms (1 week)
  - Exam, project demos (1 week)

- Augment lecture with a programming lab
  - Students will take the lab with the course
    - graded assignments and term project will be posted
  - Lab assignments targeted to learning parallel programming methods and technologies
  - A term programming project will be required
Lectures

- Book and online materials are you main sources for broader and deeper background in parallel computing
- Lectures should be more interactive
  - Supplement other sources of information
  - Covers topics of more priority
  - Intended to give you some of my perspective
  - Will provide online access to lecture slides
- Lectures will complement programming component, but intended to cover other parallel computing aspects
- Try to arrange a guest lecture or 2 during quarter
Resource: ACISS Cluster

- **Applied Computational Instrument for Scientific Synthesis**
  - NSF MRI R² award (2010)
- **Basic nodes** (1,536 total cores)
  - 128 ProLiant SL390 G7
  - Two Intel X5650 2.66 GHz 6-core CPUs per node
  - 72GB DDR3 RAM per basic node
- **Fat nodes** (512 total cores)
  - 16 ProLiant DL 580 G7
  - Four Intel X7560 2.266 GHz 8-core CPUs per node
  - 384GB DDR3 per fat node
- **GPU nodes** (624 total cores, 156 GPUs) (16 nodes reserved for course!)
  - 52 ProLiant SL390 G7 nodes, 3 NVidia M2070 GPUs (156 total GPUs)
  - Two Intel X5650 2.66 GHz 6-core CPUs per node (624 total cores)
  - 72GB DDR3 per GPU node
- ACISS has 2672 total cores
- ACISS is located in the UO Computing Center
NVIDIA Jetson TX1 Cluster

- Jetson Embedded Platform
  - NVIDIA Tegra SoC
  - Maxwell architecture GPU
- 16x TX1 development kits
- Can use the Jetson cluster for parallel programming projects


Variety of NVIDIA GPUs
- GTX 980
  - Maxwell architecture
  - 2048 cores
- K80
  - dual-GPU
  - 4992 cores

Several Intel Xeon MIC
- Many Integrated Cores (MIC)
- Knights Landing architecture
- 72 cores each with
  - 2 vector processing units (AVX512)
Course Assignments

- **Q&A assignments**
  - Exercises/questions primarily to prepare exam
- **Individual programming assignments**
  - 4-5 programming assignments
- **No midterm**
- **Team term project**
  - Programming, presentation, report
  - Teams TBD
- **Term exam in beginning of Week 10**
- **Research summary paper (only graduate students)**
- **No final exam during finals week**
  - Team project presentations during final period
Parallel Programming Term Project

- Major programming project for the course
  - Non-trivial parallel application
  - Include performance analysis
  - Use whatever parallel machine resources available
  - Project ideas will be provided in 4th week

- Project teams
  - 2-4 person teams
  - Maybe a whole class project

- Project dates
  - Proposal: due end of 5th week
  - Project talks during last lecture of week 10
  - Project report due Friday (11:59pm) of week 10

- Need to get system accounts!!!
Term Paper (for graduate students)

- Investigate parallel computing topic of interest
  - More in depth review
  - Individual choice
  - Summary of major points

- Requires minimum of ten references
  - Book and other references has a large bibliography
  - Google Scholar, Keywords: parallel computing
  - NEC CiteSeer Scientific Literature Digital Library

- Paper abstract and references due Friday of week 8
- Final term paper due Friday of week 10
- Individual work
What will you get out of CIS 431/531?

- In-depth understanding of parallel computer design
- Knowledge about how to program parallel computer systems
- Understanding of pattern-based parallel programming
- Exposure to different forms of parallel algorithms
- Practical experience using parallel technology
- Background on parallel performance modeling
- Techniques for empirical performance analysis
- Fun and new friends
Overview

- Parallel computing is both an old and broad field of computer science
  - It was there from the very beginning
  - Involves all aspects: architecture, HW/SW systems, languages, programming paradigms, algorithms, and theoretical models
  - Computing in parallel

- Performance is the *raison d’être* for parallelism
  - High-performance computing
  - Drives computational science revolution

- Topics of study
  - Parallel architectures
  - Parallel programming
  - Parallel algorithms
  - Parallel performance models and tools
  - Parallel applications
Parallel Processing – What is it?

- A *parallel computer* is a computer system that uses multiple processing elements simultaneously in a cooperative manner to solve a computational problem.

- *Parallel processing* includes techniques and technologies that make it possible to compute in parallel:
  - Hardware, networks, operating systems, parallel libraries, languages, compilers, algorithms, tools, …

- Parallel computing versus sequential computing:
  - Parallelism is natural (sequential computing is base case)
  - Computing problems differ in level / type of parallelism

- Parallelism is all about performance! Really?
Concurrency

- Consider multiple tasks to be executed in a computer
- Tasks are concurrent with respect to each if
  - They *can* execute at the same time (*concurrent execution*)
  - Implies that there are no dependencies between the tasks
- Dependencies
  - If a task requires results produced by other tasks in order to execute correctly, the task’s execution is *dependent*
  - If two tasks are dependent, they are not concurrent
  - Some form of synchronization must be used to enforce (satisfy) dependencies
- Concurrency is fundamental to computer science
  - Operating systems, databases, networking, …
Concurrency and Parallelism

- Concurrent is not the same as parallel! Why?
- Parallel execution
  - Concurrent tasks *actually* execute at the same time
  - Multiple (processing) resources have to be available
- Parallelism = concurrency + “parallel” hardware
  - Both are required
  - Find concurrent execution opportunities
  - Develop application to execute in parallel
  - Run application on parallel hardware
- Is a parallel application a concurrent application?
- Is a parallel application run with one processor parallel? Why or why not?
Parallelism

- There are granularities of parallelism (parallel execution) in programs
  - Processes, threads, routines, statements, instructions, …
  - Think about what are the software elements that execute concurrently
- These must be supported by hardware resources
  - Processors, cores, … (execution of instructions)
  - Memory, DMA, networks, … (other associated operations)
  - All aspects of computer architecture offer opportunities for parallel hardware execution
- Concurrency is a necessary condition for parallelism
  - Where can you find concurrency?
  - How is concurrency expressed to exploit parallel systems?
Why use parallel processing?

- Two primary reasons (both performance related)
  - Faster time to solution (response time)
  - Solve bigger computing problems (in same time)
- Other factors motivate parallel processing
  - Effective use of machine resources
  - Cost efficiencies
  - Overcoming memory constraints
- Serial machines have inherent limitations
  - Processor speed, memory bottlenecks, …
- Parallelism has become the future of computing
- Performance is still the driving concern
- Parallelism = concurrency + parallel HW + performance
Course Themes

- Parallel programming
  - Parallel thinking and algorithms
  - Parallel programming paradigms and abstractions
  - Parallel programming environments

- Parallel architectures and machines
  - Parallel programming models for parallel architectures
  - Architecture determines performance concerns
  - Characteristics of parallel machine matter

- Performance
  - Faster does not necessarily mean more efficient
  - Tradeoffs between programmability and performance
Perspectives on Parallel Processing

- Parallel computer architecture
  - Hardware needed for parallel execution?
  - Computer system design
- (Parallel) Operating system
  - How to manage systems aspects in a parallel computer
- Parallel programming
  - Libraries (low-level, high-level)
  - Languages
  - Software development environments
- Parallel algorithms
- Parallel performance evaluation
- Parallel tools
  - Performance, analytics, visualization, …
Why study parallel computing today?

- Computing architecture
  - Innovations often drive to novel programming models
- Technological convergence
  - The “killer micro” is ubiquitous
  - Laptops and supercomputers are fundamentally similar!
  - Trends cause diverse approaches to converge
- Technological trends make parallel computing inevitable
  - Multi-core processors are here to stay!
  - Practically every computing system is operating in parallel
- Understand fundamental principles and design tradeoffs
  - Programming, systems support, communication, memory, …
  - Performance
- Parallelism is the future of computing
Inevitability of Parallel Computing

- Application demands
  - Insatiable need for computing cycles
- Technology trends
  - Processor and memory
- Architecture trends
- Economics
- Current trends:
  - Today’s microprocessors have multiprocessor support
  - Servers and workstations available as multiprocessors
  - Tomorrow’s microprocessors are multiprocessors
  - Multi-core is here to stay and #cores/processor is growing
  - Accelerators (GPUs, gaming systems)
Application Characteristics

- Application performance demands hardware advances
- Hardware advances generate new applications
- New applications have greater performance demands
  - Exponential increase in microprocessor performance
  - Innovations in parallel architecture and integration

- Range of performance requirements
  - System performance must also improve as a whole
  - Performance requirements require computer engineering
  - Costs addressed through technology advancements
Broad Parallel Architecture Issues

- Resource allocation
  - How many processing elements?
  - How powerful are the elements?
  - How much memory?

- Data access, communication, and synchronization
  - How do the elements cooperate and communicate?
  - How are data transmitted between processors?
  - What are the abstractions and primitives for cooperation?

- Performance and scalability
  - How does it all translate into performance?
  - How does it scale?

- Moore’s Law (Gordon E. Moore, Intel co-founder)
  - # transistors in an integrated circuit doubles every 2 years
  - Observation or conjecture, not a physical or natural law
  - End of Moore’s Law?
    ◦ expected to continue to ~2020

- More transistors = more parallelism opportunities

- Microprocessors
  - Implicit parallelism
    ◦ pipelining
    ◦ multiple functional units
    ◦ superscalar
  - Explicit parallelism
    ◦ SIMD instructions
    ◦ long instruction works

- Dennard scaling
  - Power requirements are proportional to area
  - Performance per watt grows with transistor density
  - Dennard scaling has broken down since 2007!
    ◦ *Dennard scaling states that as transistors get smaller, their power density stays constant, or effectively smaller transistors consume less power.*
Microprocessor Transistor Counts (1971-2011)

Data from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanović
Slide from Kathy Yelick
What’s Driving Parallel Computing Architecture?

von Neumann bottleneck!!
(memory wall)
What has happened in the last 15 years?

- Processing chip manufacturers increased processor performance by increasing CPU clock frequency
  - Riding Moore’s law
- Until the chips got too hot!
  - Greater clock frequency $\Rightarrow$ greater electrical power
  - Breakdown of Dennard scaling
  - Pentium 4 heat sink  ☀ Frying an egg on a Pentium 4

- Add multiple cores to add performance
  - Keep clock frequency same or reduced
  - Keep lid on power requirements
Power Density Growth

Figure courtesy of Pat Gelsinger, Intel Developer Forum, Spring 2004
What’s Driving Parallel Computing Architecture?

15 Years of exponential growth ~2x year has ended

- Transistors (in Thousands)
- Frequency (MHz)
- Cores

Data from Kunle Olukotun, Lance Hammond, Herb Sutter, Burton Smith, Chris Batten, and Krste Asanović

Slide from Kathy Yelick
What’s Driving Parallel Computing Architecture?

Power is the root cause of all this

A hardware issue just became a software problem

power wall
Single-core Performance Scaling

- Rate of single-instruction stream performance scaling has decreased to almost 0
- Frequency scaling is limited by power
- ILP scaling is tapped out
- CPU architects are now building faster processors by adding more execution units to run in parallel
- Parallel software must be written to see performance gains
Classifying Parallel Systems – Flynn’s Taxonomy

- Distinguishes multi-processor computer architectures along the two independent dimensions
  - Instruction and Data
  - Each dimension can have one state: Single or Multiple

- SISD: Single Instruction, Single Data
  - Serial (non-parallel) machine

- SIMD: Single Instruction, Multiple Data
  - Processor arrays and vector machines

- MISD: Multiple Instruction, Single Data (weird)

- MIMD: Multiple Instruction, Multiple Data
  - Most common parallel computer systems
Parallel Architecture Types

- Instruction-Level Parallelism
  - Parallelism captured in instruction processing
- Vector processors
  - Operations on multiple data stored in vector registers
- Shared-memory Multiprocessor (SMP)
  - Multiple processors sharing memory
  - Symmetric Multiprocessor (SMP)
- Multicomputer
  - Multiple computer connect via network
  - Distributed-memory cluster
- Massively Parallel Processor (MPP)
Phases of Supercomputing (Parallel) Architecture

- Phase 1 (1950s): sequential instruction execution
- Phase 2 (1960s): sequential instruction issue
  - Pipeline execution, reservations stations
  - Instruction Level Parallelism (ILP)
- Phase 3 (1970s): vector processors
  - Pipelined arithmetic units
  - Registers, multi-bank (parallel) memory systems
- Phase 4 (1980s): SIMD and SMPs
- Phase 5 (1990s): MPPs and clusters
  - Communicating sequential processors
- Phase 6 (2000s): many cores, accelerators, scale, …
- Phase 7 (2010s): many more cores, heterogeneity, …
Performance Expectations

- If each processor is rated at k MFLOPS and there are p processors, we should expect to see k*p MFLOPS performance? Correct?
- If it takes 100 seconds on 1 processor, it should take 10 seconds on 10 processors? Correct?
- Several causes affect performance
  - Each must be understood separately
  - But they interact with each other in complex ways
    - solution to one problem may create another
    - one problem may mask another
- Scaling (system, problem size) can change conditions
- Need to understand performance space
Scalability

- A program can scale up to use many processors
  - What does that mean?
- How do you evaluate scalability?
- How do you evaluate scalability goodness?
- Comparative evaluation
  - If double the number of processors, what to expect?
  - Is scalability linear?
- Use parallel efficiency measure
  - Is efficiency retained as problem size increases?
- Apply performance metrics
Top 500 Benchmarking Methodology

- Listing of the world’s 500 most powerful computers
- Yardstick for high-performance computing (HPC)
  - $R_{\text{max}}$: maximal performance Linpack benchmark
    - dense linear system of equations ($Ax = b$)
  - $R_{\text{peak}}$: theoretical peak performance
- Data listed
  - $N_{\text{max}}$: problem size needed to achieve $R_{\text{max}}$
  - $N_{1/2}$: problem size needed to achieve $1/2$ of $R_{\text{max}}$
  - Manufacturer and computer type
  - Installation site, location, and year
- Updated twice a year at SC and ISC conferences
### Top 10 (November 2016)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>System</th>
<th>Cores</th>
<th>Rmax (TFlop/s)</th>
<th>Rpeak (TFlop/s)</th>
<th>Power (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>National Supercomputing Center in Wuxi, China</td>
<td>Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRPC</td>
<td>10,649,600</td>
<td>93,014.6</td>
<td>125,435.9</td>
<td>15,371</td>
</tr>
<tr>
<td>2</td>
<td>National Super Computer Center in Guangzhou, China</td>
<td>Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT</td>
<td>3,120,000</td>
<td>33,862.7</td>
<td>54,902.4</td>
<td>17,808</td>
</tr>
<tr>
<td>3</td>
<td>DOE/SC/Oak Ridge National Laboratory, United States</td>
<td>Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.</td>
<td>560,640</td>
<td>17,590.0</td>
<td>27,112.5</td>
<td>8,209</td>
</tr>
<tr>
<td>4</td>
<td>DOE/NNSA/LLNL, United States</td>
<td>Sequoia - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM</td>
<td>1,572,864</td>
<td>17,173.2</td>
<td>20,132.7</td>
<td>7,890</td>
</tr>
<tr>
<td>5</td>
<td>DOE/SC/LBNL/NERSC, United States</td>
<td>Cori - Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries interconnect Cray Inc.</td>
<td>622,336</td>
<td>14,014.7</td>
<td>27,880.7</td>
<td>3,939</td>
</tr>
<tr>
<td>6</td>
<td>Joint Center for Advanced High Performance Computing, Japan</td>
<td>Oakforest-PACS - PRIMERGY CX1640 M1, Intel Xeon Phi 7250 68C 1.4GHz, Intel Omni-Path Fujitsu</td>
<td>556,104</td>
<td>13,554.6</td>
<td>24,913.5</td>
<td>2,718.7</td>
</tr>
<tr>
<td>7</td>
<td>RIKEN Advanced Institute for Computational Science (AICS), Japan</td>
<td>K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu</td>
<td>705,024</td>
<td>10,510.0</td>
<td>11,280.4</td>
<td>12,659.9</td>
</tr>
<tr>
<td>8</td>
<td>Swiss National Supercomputing Centre [CSCS], Switzerland</td>
<td>Piz Daint - Cray XC50, Xeon E5-2690v3 12C 2.6GHz, Aries interconnect, NVIDIA Tesla P100 Cray Inc.</td>
<td>206,720</td>
<td>9,779.0</td>
<td>15,988.0</td>
<td>1,312</td>
</tr>
<tr>
<td>9</td>
<td>DOE/SC/Argonne National Laboratory, United States</td>
<td>Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM</td>
<td>786,432</td>
<td>8,586.6</td>
<td>10,066.3</td>
<td>3,945</td>
</tr>
<tr>
<td>10</td>
<td>DOE/NNSA/LANL/SNL, United States</td>
<td>Trinity - Cray XC40, Xeon E5-2690v3 16C 2.3GHz, Aries interconnect Cray Inc.</td>
<td>301,056</td>
<td>8,100.9</td>
<td>11,078.9</td>
<td>4,232.6</td>
</tr>
</tbody>
</table>
## Top 10 (November 2015)

<table>
<thead>
<tr>
<th>RANK</th>
<th>SITE</th>
<th>SYSTEM</th>
<th>CORES</th>
<th>RMAX (TFLOP/S)</th>
<th>RPEAK (TFLOP/S)</th>
<th>POWER (KW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>National Super Computer Center in Guangzhou China</td>
<td>Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT</td>
<td>3,120,000</td>
<td>33,862.7</td>
<td>54,902.4</td>
<td>17,808</td>
</tr>
<tr>
<td>2</td>
<td>DOE/SC/Oak Ridge National Laboratory United States</td>
<td>Titan - Cray XK7, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.</td>
<td>560,640</td>
<td>17,590.0</td>
<td>27,112.5</td>
<td>8,209</td>
</tr>
<tr>
<td>3</td>
<td>DOE/NNSA/LLNL United States</td>
<td>Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM</td>
<td>1,572,864</td>
<td>17,173.2</td>
<td>20,132.7</td>
<td>7,890</td>
</tr>
<tr>
<td>4</td>
<td>RIKEN Advanced Institute for Computational Science (AICS) Japan</td>
<td>K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu</td>
<td>705,024</td>
<td>10,510.0</td>
<td>11,280.4</td>
<td>12,660</td>
</tr>
<tr>
<td>5</td>
<td>DOE/SC/Argonne National Laboratory United States</td>
<td>Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM</td>
<td>786,432</td>
<td>8,586.6</td>
<td>10,066.3</td>
<td>3,945</td>
</tr>
<tr>
<td>6</td>
<td>DOE/NNSA/LANL/SNL United States</td>
<td>Trinity - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc.</td>
<td>301,056</td>
<td>8,100.9</td>
<td>11,078.9</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Swiss National Supercomputing Centre (CSCS) Switzerland</td>
<td>Piz Daint - Cray XC30, Xeon E5-2670 8C 2.60GHz, Aries interconnect, NVIDIA K20x Cray Inc.</td>
<td>115,984</td>
<td>6,271.0</td>
<td>7,788.9</td>
<td>2,325</td>
</tr>
<tr>
<td>8</td>
<td>HLRS - Höchstleistungsrechenzentrum Stuttgart Germany</td>
<td>Hazel Hen - Cray XC40, Xeon E5-2680v3 12C 2.5GHz, Aries interconnect Cray Inc.</td>
<td>185,088</td>
<td>5,640.2</td>
<td>7,403.5</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>King Abdullah University of Science and Technology Saudi Arabia</td>
<td>Shaheen II - Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc.</td>
<td>196,608</td>
<td>5,537.0</td>
<td>7,235.2</td>
<td>2,834</td>
</tr>
<tr>
<td>10</td>
<td>Texas Advanced Computing Center/Univ. of Texas United States</td>
<td>Stampede - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell</td>
<td>462,462</td>
<td>5,168.1</td>
<td>8,520.1</td>
<td>4,510</td>
</tr>
</tbody>
</table>

Different architectures
# Top 10 (November 2014)

<table>
<thead>
<tr>
<th>RANK</th>
<th>SITE</th>
<th>SYSTEM</th>
<th>CORES</th>
<th>RMAX (TFLOP/S)</th>
<th>RPEAK (TFLOP/S)</th>
<th>POWER (KW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>National Super Computer Center in Guangzhou, China</td>
<td>Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.20GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT</td>
<td>3,120,000</td>
<td>33,862.7</td>
<td>54,902.4</td>
<td>17,808</td>
</tr>
<tr>
<td>2</td>
<td>DOE/SC/Oak Ridge National Laboratory, United States</td>
<td>Titan - Cray XK7, Opteron 6274 16C 2.20GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.</td>
<td>560,640</td>
<td>17,590.0</td>
<td>27,112.5</td>
<td>8,209</td>
</tr>
<tr>
<td>3</td>
<td>DOE/NNSA/LLNL United States</td>
<td>Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM</td>
<td>1,572,864</td>
<td>17,173.2</td>
<td>20,132.7</td>
<td>7,890</td>
</tr>
<tr>
<td>4</td>
<td>RIKEN Advanced Institute for Computational Science (AICS), Japan</td>
<td>K computer, SPARC64 VIIIx 2.0GHz, Tofu interconnect, Fujitsu</td>
<td>705,024</td>
<td>10,510.0</td>
<td>11,280.4</td>
<td>12,660</td>
</tr>
<tr>
<td>5</td>
<td>DOE/SC/Argonne National Laboratory, United States</td>
<td>Mira - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM</td>
<td>786,432</td>
<td>8,586.6</td>
<td>10,066.3</td>
<td>3,945</td>
</tr>
<tr>
<td>6</td>
<td>Swiss National Supercomputing Centre (CSCS), Switzerland</td>
<td>Piz Daint - Cray XC30, Xeon E5-2670 8C 2.60GHz, Aries interconnect, NVIDIA K20x Cray Inc.</td>
<td>115,984</td>
<td>6,271.0</td>
<td>7,788.9</td>
<td>2,325</td>
</tr>
<tr>
<td>7</td>
<td>Texas Advanced Computing Center/Univ. of Texas, United States</td>
<td>Stampede - PowerEdge C8220, Xeon E5-2680 8C 2.70GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell</td>
<td>462,462</td>
<td>5,168.1</td>
<td>8,502.1</td>
<td>4,510</td>
</tr>
<tr>
<td>8</td>
<td>Forschungszentrum Juelich (FZJ), Germany</td>
<td>JUQUEEN - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM</td>
<td>458,752</td>
<td>5,008.9</td>
<td>5,872.0</td>
<td>2,301</td>
</tr>
<tr>
<td>9</td>
<td>DOE/NNSA/LLNL United States</td>
<td>Vulcan - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM</td>
<td>393,216</td>
<td>4,293.3</td>
<td>5,033.2</td>
<td>1,972</td>
</tr>
<tr>
<td>10</td>
<td>Government, United States</td>
<td>Cray CS-Storm, Intel Xeon E5-2660v2 10C 2.2GHz, Infiniband FDR, Nvidia K40 Cray Inc.</td>
<td>72,800</td>
<td>3,577.0</td>
<td>6,131.8</td>
<td>1,499</td>
</tr>
</tbody>
</table>

Different architectures

Same systems!!!
## Top 10 (November 2013)

Same as 2014!

<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>System</th>
<th>Cores</th>
<th>Rmax (TFlop/s)</th>
<th>Rpeak (TFlop/s)</th>
<th>Power (kW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>National Super Computer Center in Guangzhou, China</td>
<td>Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2682 12C 2.20GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT</td>
<td>3,120,000</td>
<td>33,862.7</td>
<td>54,902.4</td>
<td>17,808</td>
</tr>
<tr>
<td>2</td>
<td>DOE/SC/Oak Ridge National Laboratory, United States</td>
<td>Titan - Cray XK7, Opteron 6274 16C 2.20GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.</td>
<td>560,640</td>
<td>17,590.0</td>
<td>27,112.5</td>
<td>8,209</td>
</tr>
<tr>
<td>3</td>
<td>DOE/NNSA/LLNL, United States</td>
<td>Sequoia - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM</td>
<td>1,572,864</td>
<td>17,173.2</td>
<td>20,132.7</td>
<td>7,890</td>
</tr>
<tr>
<td>4</td>
<td>RIKEN Advanced Institute for Computational Science (AICS), Japan</td>
<td>K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu</td>
<td>705,024</td>
<td>10,510.0</td>
<td>11,280.4</td>
<td>12,660</td>
</tr>
<tr>
<td>5</td>
<td>DOE/SC/Argonne National Laboratory, United States</td>
<td>Mira - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM</td>
<td>786,432</td>
<td>8,586.6</td>
<td>10,066.3</td>
<td>3,945</td>
</tr>
<tr>
<td>6</td>
<td>Swiss National Supercomputing Centre (CSCS), Switzerland</td>
<td>Piz Daint - Cray XC30, Xeon E5-2670 8C 2.60GHz, Aries interconnect, NVIDIA K20x Cray Inc.</td>
<td>115,984</td>
<td>6,271.0</td>
<td>7,788.9</td>
<td>2,325</td>
</tr>
<tr>
<td>7</td>
<td>Texas Advanced Computing Center/Univ. of Texas, United States</td>
<td>Stampede - PowerEdge C8220, Xeon E5-2680 8C 2.70GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell</td>
<td>462,462</td>
<td>5,168.1</td>
<td>8,520.1</td>
<td>4,510</td>
</tr>
<tr>
<td>8</td>
<td>Forschungszentrum Juelich (FZJ), Germany</td>
<td>JUQUEEN - BlueGene/Q, Power BQC 16C 1.60GHz, Custom Interconnect IBM</td>
<td>458,752</td>
<td>5,008.9</td>
<td>5,872.0</td>
<td>2,301</td>
</tr>
<tr>
<td>9</td>
<td>DOE/NNSA/LLNL, United States</td>
<td>Vulcan - BlueGene/Q, Power BQC 16C 1.60GHz, Custom Interconnect IBM</td>
<td>393,216</td>
<td>4,293.3</td>
<td>5,033.2</td>
<td>1,972</td>
</tr>
<tr>
<td>10</td>
<td>Leibniz Rechenzentrum, Germany</td>
<td>SuperMUC - IDDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR IBM</td>
<td>147,456</td>
<td>2,897.0</td>
<td>3,185.1</td>
<td>3,423</td>
</tr>
</tbody>
</table>
Top 500 – Performance (up to November 2016)
**#1: NRCPC: Sunway TaihuLight**

- 93 petaflop/s (Pflop/s) on the Linpack benchmark. Power: 15 MW
- Nodes: 40,960 Chinese-designed SW26010 manycore 64-bit RISC processors based on the Sunway architecture; each processor contains:
  - 256 processing cores
  - an additional 4 auxiliary cores for system management
- Memory: 1.31
- Storage: 20 PB
#2: NUDT Tiahne-2 (Milkyway-2)

- **Compute Nodes** have 3.432 Tflop/s per node
  - 16,000 nodes
  - 32000 Intel Xeon CPU
  - 48000 Intel Xeon Phi

- **Operations Nodes**
  - 4096 FT CPUs

- **Proprietary interconnect**
  - TH2 express

- **1PB memory**
  - Host memory only

- **Global shared parallel storage** is 12.4 PB

- **Cabinets**: 125+13+24 = 162
  - Compute, communication, storage
  - ~750 m²
#3: ORNL Titan Hybrid System (Cray XK7)

- Peak performance of 27.1 PF
  - 24.5 GPU + 2.6 CPU
- 18,688 Compute Nodes each with:
  - 16-Core AMD Opteron CPU
  - NVIDIA Tesla “K20x” GPU
  - 32 + 6 GB memory
- 512 Service and I/O nodes
- 200 Cabinets
- 710 TB total system memory
- Cray Gemini 3D Torus Interconnect
- 8.9 MW peak power

4,352 ft²
#4: LLNL Sequoia (IBM BG/Q)

- **Compute card**
  - 16-core PowerPC A2 processor
  - 16 GB DDR3
- **Compute node has 98,304 cards**
- **Total system size:**
  - 1,572,864 processing cores
  - 1.5 PB memory
- **5-dimensional torus interconnection network**
- **Area of 3,000 ft²**
#5: NERSC Cori (Cray XC40 supercomputer)

- **Theoretical peak performance:**
  - Phase I Haswell: 1.92 PFlops/sec; Phase II KNL: 29.1 PFlops/sec.

- **Sustained application performance on NERSC SSP codes:**
  - Phase I Haswell: 83 TFlop/s (vs. 129 TFlop/s for Edison and 52.1 TFlop/s for Hopper); Phase II KNL: 562.31 TFlop/s.

- **Total compute nodes:**
  - Phase I Haswell: 2,004 computes nodes, 64,128 cores in total (32 cores per node);
  - Phase II KNL: 9,688 compute nodes, 658,784 cores in total (68 cores per node).

- **Cray Aries high-speed interconnect with Dragonfly topology as on Edison**
  - (0.25 µs to 3.7 µs MPI latency, ~8GB/sec MPI bandwidth)

- **Aggregate memory:**
  - Phase I Haswell: 203 TB; Phase II KNL: 1 PB.

- **Scratch storage capacity:** 30 PB
## Contemporary HPC Architectures

<table>
<thead>
<tr>
<th>Date</th>
<th>System</th>
<th>Location</th>
<th>Comp</th>
<th>Comm</th>
<th>Peak (PF)</th>
<th>Power (MW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Jaguar; Cray XT5</td>
<td>ORNL</td>
<td>AMD 6c</td>
<td>Seastar2</td>
<td>2.3</td>
<td>7.0</td>
</tr>
<tr>
<td>2010</td>
<td>Tianhe-1A</td>
<td>NSC Tianjin</td>
<td>Intel + NVIDIA</td>
<td>Proprietary</td>
<td>4.7</td>
<td>4.0</td>
</tr>
<tr>
<td>2010</td>
<td>Nebulae</td>
<td>NSCS Shenzhen</td>
<td>Intel + NVIDIA</td>
<td>IB</td>
<td>2.9</td>
<td>2.6</td>
</tr>
<tr>
<td>2010</td>
<td>Tsubame 2</td>
<td>TiTech</td>
<td>Intel + NVIDIA</td>
<td>IB</td>
<td>2.4</td>
<td>1.4</td>
</tr>
<tr>
<td>2011</td>
<td>K Computer</td>
<td>RIKEN/Kobe</td>
<td>SPARC64 VIIIfx</td>
<td>Tofu</td>
<td>10.5</td>
<td>12.7</td>
</tr>
<tr>
<td>2012</td>
<td>Titan; Cray XK6</td>
<td>ORNL</td>
<td>AMD + NVIDIA</td>
<td>Gemini</td>
<td>27</td>
<td>9</td>
</tr>
<tr>
<td>2012</td>
<td>Mira; BlueGeneQ</td>
<td>ANL</td>
<td>SoC</td>
<td>Proprietary</td>
<td>10</td>
<td>3.9</td>
</tr>
<tr>
<td>2012</td>
<td>Sequoia; BlueGeneQ</td>
<td>LLNL</td>
<td>SoC</td>
<td>Proprietary</td>
<td>20</td>
<td>7.9</td>
</tr>
<tr>
<td>2012</td>
<td>Blue Waters; Cray</td>
<td>NCSA/UIUC</td>
<td>AMD + (partial) NVIDIA</td>
<td>Gemini</td>
<td>11.6</td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>Stampede</td>
<td>TACC</td>
<td>Intel + MIC</td>
<td>IB</td>
<td>9.5</td>
<td>5</td>
</tr>
<tr>
<td>2013</td>
<td>Tianhe-2</td>
<td>NSCC-GZ (Guangzhou)</td>
<td>Intel + MIC</td>
<td>Proprietary</td>
<td>54</td>
<td>~20</td>
</tr>
</tbody>
</table>
## Top 10 (Top500 List, June 2011)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>Computer</th>
<th>Country</th>
<th>Cores [Pflops]</th>
<th>% of Peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RIKEN Advanced Inst for Comp Sci</td>
<td>K Computer Fujitsu SPARC64 VIIIfx + custom</td>
<td>Japan</td>
<td>548,352</td>
<td>8.16</td>
</tr>
<tr>
<td>2</td>
<td>Nat. SuperComputer Center in Tianjin</td>
<td>Tianhe-1A, NUDT Intel + Nvidia GPU + custom</td>
<td>China</td>
<td>186,368</td>
<td>2.57</td>
</tr>
<tr>
<td>3</td>
<td>DOE / OS Oak Ridge Nat Lab</td>
<td>Jaguar, Cray AMD + custom</td>
<td>USA</td>
<td>224,162</td>
<td>1.76</td>
</tr>
<tr>
<td>4</td>
<td>Nat. Supercomputer Center in Shenzhen</td>
<td>Nebulea, Dawning Intel + Nvidia GPU + IB</td>
<td>China</td>
<td>120,640</td>
<td>1.27</td>
</tr>
<tr>
<td>5</td>
<td>GSIC Center, Tokyo Institute of Technology</td>
<td>Tusbame 2.0, HP Intel + Nvidia GPU + IB</td>
<td>Japan</td>
<td>73,278</td>
<td>1.19</td>
</tr>
<tr>
<td>6</td>
<td>DOE / NNSA LANL &amp; SNL</td>
<td>Cielo, Cray AMD + custom</td>
<td>USA</td>
<td>142,272</td>
<td>1.11</td>
</tr>
<tr>
<td>7</td>
<td>NASA Ames Research Center/NAS</td>
<td>Plelades SGI Altix ICE 8200EX/8400EX + IB</td>
<td>USA</td>
<td>111,104</td>
<td>1.09</td>
</tr>
<tr>
<td>8</td>
<td>DOE / OS Lawrence Berkeley Nat Lab</td>
<td>Hopper, Cray AMD + custom</td>
<td>USA</td>
<td>153,408</td>
<td>1.054</td>
</tr>
<tr>
<td>9</td>
<td>Commissariat a l'Energie Atomique (CEA)</td>
<td>Tera-10, Bull Intel + IB</td>
<td>France</td>
<td>138,368</td>
<td>1.050</td>
</tr>
<tr>
<td>10</td>
<td>DOE / NNSA Los Alamos Nat Lab</td>
<td>Roadrunner, IBM AMD + Cell GPU + IB</td>
<td>USA</td>
<td>122,400</td>
<td>1.04</td>
</tr>
</tbody>
</table>

Figure credit: http://www.netlib.org/utk/people/JackDongarra/SLIDES/korea-2011.pdf
Japanese K Computer (#1 in June 2011)

### K computer Specifications

<table>
<thead>
<tr>
<th>CPU (SPARC64 VIII/ffx)</th>
<th>Cores/Node</th>
<th>8 cores (@2GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>128GFlops</td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td>SPARC V9 + HPC extension</td>
<td></td>
</tr>
<tr>
<td>Cache</td>
<td>L1(I/D) Cache : 32KB/32KB L2 Cache : 6MB</td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td>58W (typ. 30 C)</td>
<td></td>
</tr>
<tr>
<td>Mem. bandwidth</td>
<td>64GB/s.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Node</th>
<th>Configuration</th>
<th>1 CPU / Node</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Memory capacity</td>
<td>16GB (2GB/core)</td>
</tr>
</tbody>
</table>

| System board(SB) | No. of nodes | 4 nodes /SB |
| System           | No. of SB    | 24 SBs/rack |

| System           | Nodes/system  | > 80,000     |

<table>
<thead>
<tr>
<th>Inter-connect</th>
<th>Topology</th>
<th>6D Mesh/Torus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>5GB/s. for each link</td>
<td></td>
</tr>
<tr>
<td>No. of link</td>
<td>10 links/node</td>
<td></td>
</tr>
<tr>
<td>Additional feature</td>
<td>H/W barrier, reduction</td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td>Routing chip structure (no outside switch box)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cooling</th>
<th>CPU, ICC*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Direct water cooling</td>
</tr>
<tr>
<td></td>
<td>Air cooling</td>
</tr>
</tbody>
</table>

**CPU**

128GFlops

SPARC64™ VIII/ffx

8 Cores@2.0GHz

**System**

LINPACK 10 PFlops

over 1PB mem.

800 racks

80,000 CPUs

640,000 cores

**Node**

128 GFlops

16GB Memory

64GB/s Memory bandwidth

**System Board**

512 GFlops

64 GB memory

**Rack**

12.3 TFlops

15TB memory

* ICC : Interconnect Chip

New Linpack run with 705,024 cores at 10.51 Pflop/s (88,128 CPUs)
# Top 500 Top 10 (2006)

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Computer</th>
<th>Rmax [TF/s]</th>
<th>Installation Site</th>
<th>Country</th>
<th>Year</th>
<th>#Proc</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBM</td>
<td>BlueGene/L eServer Blue Gene</td>
<td>280.6</td>
<td>DOE/NNSA/LLNL</td>
<td>USA</td>
<td>2005</td>
<td>131,072</td>
</tr>
<tr>
<td>Sandia/Cray</td>
<td>Red Storm Cray XT3</td>
<td>101.4</td>
<td>NNSA/Sandia</td>
<td>USA</td>
<td>2006</td>
<td>26,544</td>
</tr>
<tr>
<td>IBM</td>
<td>BGW eServer Blue Gene</td>
<td>91.29</td>
<td>IBM Thomas Watson</td>
<td>USA</td>
<td>2005</td>
<td>40,960</td>
</tr>
<tr>
<td>IBM</td>
<td>ASC Purple eServer pSeries p575</td>
<td>75.76</td>
<td>DOE/NNSA/LLNL</td>
<td>USA</td>
<td>2005</td>
<td>12,208</td>
</tr>
<tr>
<td>IBM</td>
<td>MareNostrum JS21 Cluster, Myrinet</td>
<td>62.63</td>
<td>Barcelona Supercomputing Center</td>
<td>Spain</td>
<td>2006</td>
<td>12,240</td>
</tr>
<tr>
<td>Dell</td>
<td>Thunderbird PowerEdge 1850, IB</td>
<td>53.00</td>
<td>NNSA/Sandia</td>
<td>USA</td>
<td>2005</td>
<td>9,024</td>
</tr>
<tr>
<td>Bull</td>
<td>Tera-10 NovaScale 5160, Quadrics</td>
<td>52.84</td>
<td>CEA</td>
<td>France</td>
<td>2006</td>
<td>9,968</td>
</tr>
<tr>
<td>SGI</td>
<td>Columbia Altix, Infiniband</td>
<td>51.87</td>
<td>NASA Ames</td>
<td>USA</td>
<td>2004</td>
<td>10,160</td>
</tr>
<tr>
<td>NEC/Sun</td>
<td>Tsubame Fire x4600, ClearSpeed, IB</td>
<td>47.38</td>
<td>GSIC / Tokyo Institute of Technology</td>
<td>Japan</td>
<td>2006</td>
<td>11,088</td>
</tr>
<tr>
<td>Cray</td>
<td>Jaguar Cray XT3</td>
<td>43.48</td>
<td>ORNL</td>
<td>USA</td>
<td>2006</td>
<td>10,424</td>
</tr>
</tbody>
</table>
## Top 500 Linpack Benchmark List (June 2002)

<table>
<thead>
<tr>
<th>Computer (Full Precision)</th>
<th>Number of Processors</th>
<th>$R_{\text{max}}$ Gflop/s</th>
<th>$N_{\text{max}}$ order</th>
<th>$N_{1/2}$ order</th>
<th>$R_{\text{peak}}$ Gflop/s</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earth Simulator, NEC processors**** esc</td>
<td>5104</td>
<td>35610</td>
<td>1041216</td>
<td>265408</td>
<td>40832</td>
</tr>
<tr>
<td>ASCI White-Pacific, IBM SP Power 3 (375 MHz) llnl</td>
<td>8000</td>
<td>7226</td>
<td>518096</td>
<td>179000</td>
<td>12000</td>
</tr>
<tr>
<td>Compaq AlphaServer SC ES45/EV68 1GHz psc</td>
<td>3016</td>
<td>4463</td>
<td>280000</td>
<td>85000</td>
<td>6032</td>
</tr>
<tr>
<td>Compaq AlphaServer SC ES45/EV68 1GHz psc</td>
<td>3024</td>
<td>4059</td>
<td>525000</td>
<td>105000</td>
<td>6048</td>
</tr>
<tr>
<td>Compaq AlphaServer SC ES45/EV68 1GHz psc</td>
<td>2560</td>
<td>3980</td>
<td>360000</td>
<td>85000</td>
<td>5120</td>
</tr>
<tr>
<td>IBM SP Power3 208 nodes 375 MHz llnl</td>
<td>3328</td>
<td>3052</td>
<td>371712</td>
<td>4992</td>
<td></td>
</tr>
<tr>
<td>Compaq Alphaserver SC ES45/EV68 1GHz lanl</td>
<td>2048</td>
<td>2916</td>
<td>272000</td>
<td>4096</td>
<td></td>
</tr>
<tr>
<td>IBM SP Power3 158 nodes 375 MHz llnl</td>
<td>2528</td>
<td>2526</td>
<td>371712</td>
<td>102400</td>
<td>3792</td>
</tr>
<tr>
<td>ASCI Red Intel Pentium II Xeon core 333MHz snl</td>
<td>9632</td>
<td>2379.6</td>
<td>362880</td>
<td>75400</td>
<td>3207</td>
</tr>
<tr>
<td>ASCI Blue-Pacific SST, IBM SP 604E(332 MHz) llnl</td>
<td>5808</td>
<td>2144.</td>
<td>431344</td>
<td>432344</td>
<td>3868</td>
</tr>
<tr>
<td>ASCI Red Intel Pentium II Xeon core 333MHz snl</td>
<td>9472</td>
<td>2121.3</td>
<td>251904</td>
<td>66000</td>
<td>3154</td>
</tr>
<tr>
<td>Compaq Alphaserver SC ES45/EV68 1GHz lanl</td>
<td>1520</td>
<td>2096</td>
<td>390000</td>
<td>71000</td>
<td>3040</td>
</tr>
<tr>
<td>IBM SP 112 nodes (375 MHz POWER3 High) ibm</td>
<td>1792</td>
<td>1791</td>
<td>275000</td>
<td>275000</td>
<td>2688</td>
</tr>
<tr>
<td>HITACHI SR8000/MPP/1152(450MHz) u tokyo</td>
<td>1152</td>
<td>1709.1</td>
<td>141000</td>
<td>16000</td>
<td>2074</td>
</tr>
<tr>
<td>HITACHI SR8000-F1/168(375MHz) leibniz</td>
<td>168</td>
<td>1653</td>
<td>160000</td>
<td>19560</td>
<td>2016</td>
</tr>
<tr>
<td>ASCI Red Intel Pentium II Xeon core 333MHz snl</td>
<td>6720</td>
<td>1633.3</td>
<td>306720</td>
<td>52500</td>
<td>2238</td>
</tr>
<tr>
<td>SCI ASCI Blue Mountain lanl</td>
<td>3040</td>
<td>1608.</td>
<td>374400</td>
<td>138000</td>
<td>2520</td>
</tr>
<tr>
<td>IBM SP 328 nodes (375 MHz POWER3 Thin) noo</td>
<td>1312</td>
<td>1417.</td>
<td>374000</td>
<td>374000</td>
<td>1968</td>
</tr>
<tr>
<td>Intel ASCI Option Red (200 MHz Pentium Pro) snl</td>
<td>9152</td>
<td>1338</td>
<td>235000</td>
<td>63000</td>
<td>1830</td>
</tr>
<tr>
<td>NEC SX-5/128M8(3.2ns) osaka</td>
<td>128</td>
<td>1192.0</td>
<td>129536</td>
<td>10240</td>
<td>1280</td>
</tr>
<tr>
<td>CRAY T3E-1200 (600 MHz) us government</td>
<td>1488</td>
<td>1127.</td>
<td>148800</td>
<td>28272</td>
<td>1786</td>
</tr>
<tr>
<td>HITACHI SR8000-F1/112(375MHz) leibniz</td>
<td>112</td>
<td>1035.0</td>
<td>120000</td>
<td>15160</td>
<td>1344</td>
</tr>
</tbody>
</table>
Japanese Earth Simulator

- World’s fastest supercomputer!!! (2002)
  - 640 NEC SX-6 nodes
    - 8 vector processors
  - 5104 total processors
  - Single stage crossbar
    - ~2900 meters of cables
  - 10 TB memory
  - 700 TB disk space
  - 1.6 PB mass storage
  - 40 Tflops peak performance
  - 35.6 Tflops Linpack performance
Computer Scientists at the Japanese ES!!!
Performance Development in Top 500

Figure credit: http://www.netlib.org/utk/people/JackDongarra/SLIDES/korea-2011.pdf
Exascale Initiative

- Exascale machines targeted for 2019 (now 2022)
- What are the potential differences and problems?

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>System peak</td>
<td>8.7 Pflop/s</td>
<td>1 Eflop/s</td>
<td>O(100)</td>
</tr>
<tr>
<td>Power</td>
<td>10 MW</td>
<td>~20 MW</td>
<td></td>
</tr>
<tr>
<td>System memory</td>
<td>1.6 PB</td>
<td>32 - 64 PB</td>
<td>O(10)</td>
</tr>
<tr>
<td>Node performance</td>
<td>128 GF</td>
<td>1,2 or 15TF</td>
<td>O(10) - O(100)</td>
</tr>
<tr>
<td>Node memory BW</td>
<td>64 GB/s</td>
<td>2 - 4TB/s</td>
<td>O(100)</td>
</tr>
<tr>
<td>Node concurrency</td>
<td>8</td>
<td>O(1k) or 10k</td>
<td>O(100) - O(1000)</td>
</tr>
<tr>
<td>Total Node Interconnect BW</td>
<td>20 GB/s</td>
<td>200-400GB/s</td>
<td>O(10)</td>
</tr>
<tr>
<td>System size (nodes)</td>
<td>68,544</td>
<td>O(100,000) or O(1M)</td>
<td>O(10) - O(100)</td>
</tr>
<tr>
<td>Total concurrency</td>
<td>548,352</td>
<td>O(billion)</td>
<td>O(1,000)</td>
</tr>
<tr>
<td>MTTI</td>
<td>days</td>
<td>O(1 day)</td>
<td>- O(10)</td>
</tr>
</tbody>
</table>
Major Changes to Software and Algorithms

- What were we concerned about before and now?
- Must rethink the design for exascale
  - Data movement is expensive (Why?)
  - Flops per second are cheap (Why?)
- Need to reduce communication and synchronization
- Need to develop fault-resilient algorithms
- How do we deal with massive parallelism?
- Software must adapt to the hardware (autotuning)
Supercomputing and Computational Science

- By definition, a supercomputer is of a class of computer systems that are the most powerful computing platforms at that time.
- Computational science has always lived at the leading (and bleeding) edge of supercomputing technology.
- “Most powerful” depends on performance criteria:
  - Performance metrics related to computational algorithms
  - Benchmark “real” application codes
- Where does the performance come from?
  - More powerful processors
  - More processors (cores)
  - Better algorithms
Computational Science

- Traditional scientific methodology
  - Theoretical science
    - Formal systems and theoretical models
    - Insight through abstraction, reasoning through proofs
  - Experimental science
    - Real system and empirical models
    - Insight from observation, reasoning from experiment design

- Computational science
  - Emerging as a principal means of scientific research
  - Use of computational methods to model scientific problems
    - Numerical analysis plus simulation methods
    - Computer science tools
  - Study and application of these solution techniques
Computational Challenges

- Computational science thrives on computer power
  - Faster solutions
  - Finer resolution
  - Bigger problems
  - Improved interaction
  - BETTER SCIENCE!!

- How to get more computer power?
  - Scalable parallel computing

- Computational science also thrives better integration
  - Couple computational resources
  - Grid computing
Scalable Parallel Computing

- Scalability in parallel architecture
  - Processor numbers
  - Memory architecture
  - Interconnection network
  - Avoid critical architecture bottlenecks

- Scalability in computational problem
  - Problem size
  - Computational algorithms
    - Computation to memory access ratio
    - Computation to communication ration

- Parallel programming models and tools

- Performance scalability
Next Lectures

- Parallel computer architectures (Thursday)
- Parallel performance models (Tuesday next week)