Outline

- Course Overview
  - What is CIS 631
  - What is expected of you?
  - What will you learn in CIS 631?

- 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 12:00 – 13:20
  - Final: Thursday 8:00, June 14, project presentations

- Undergraduate course prerequisite
  - CIS 431/531 – Introduction to Parallel Computing

- Webpage:
  [http://www.cs.uoregon.edu/Classes/18S/cis631](http://www.cs.uoregon.edu/Classes/18S/cis631)

- Do Assignment 0
  - Survey of skills and projects
  - See schedule
Heritage of CIS 631

- CIS 631 has been graduate course since I began at UO
- Parallel computing in undergraduate curriculum
  - CIS 607
    - Winter 2014 seminar to plan undergraduate course
    - Develop 410/510 materials, exercises, labs, …
    - Intel Parallel Computing Center (IPCC)
      [http://ipcc.cs.uoregon.edu](http://ipcc.cs.uoregon.edu)
  - CIS 410/510
    - Spring 2014 and Winter 2015 “experimental” course
  - CIS 431/531 – Introduction to Parallel Computing
    - Proposed in AY 2015/16 as a formal course
    - Spring 2017 (Norris) and Winter 2018 (Malony)
- Ok, so what about CIS 631?
CIS 631 – Advanced Parallel Computing

- CIS 631 can now be more “advanced”
  - Assume basic parallel computing knowledge
    - CIS 431/531 or equivalent is a prerequisite
  - Assume some experience with shared and/or distributed memory parallel programming
    - multi-threading and message passing
  - Assume experience with Unix/Linux program development environments

- Focus on more state-of-the-art methodologies, systems, environments, tools, ….
- Take on more sophisticated projects
- Consider current research and development directions
- Would like your help to develop CIS 631
People in Parallel Computing at CIS/UO

- Allen D. Malony
  - Scalable parallel computing
  - Parallel performance analysis
  - Taught CIS 631 for many years

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

- Hank Childs
  - Large-scale, parallel scientific visualization
  - Visualization of large data sets
Course Plan

- Cover main areas of parallel computing
  - Architecture (1 week)
  - Programming models (paradigms) (1 week)
  - High-level programming frameworks (1 week)
  - Runtime systems (1 week)
  - Performance tools (1 week)
  - Libraries, algorithms and applications (1 week)
  - Heterogeneous computing (1 week)
  - Parallel systems environments (1 week)

- I will be away on the week of May 14
- I might miss another Thursday for eye surgery
Course Assignments

- Parallel programming training
  - Gain proficiency in parallel programming
- Individual project
- Team term project
  - Proposal, presentation, report
  - Teams TBD
  - Presentation on Thursday, 8:00, June 14 (finals period)
  - Report due on Thursday, 8:00, June 14 (finals period)
- Research summary paper
  - Document due Friday, June 8 at 18:00
Parallel Programming Term Project

- Major programming project for the course
  - Non-trivial parallel application
  - Apply parallel programming and include performance analysis

- Project teams
  - 2-4 person teams

- Project dates
  - Proposal due on ???
  - Presentation on Thursday, 8:00, June 14 (finals period)
  - Report due on Thursday, 8:00, June 14 (finals period)

- Need to get system accounts!!
  - NIC: https://systems.nic.uoregon.edu/account
    - indicate that this is for CIS 631 in Spring 2018
  - Talapas: should have been contacted by system administration
Term Paper

- 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 topic due Friday, April 27, 18:00
  - Abstract and 8 research references
  - Final term paper due Friday, June 8 at 18:00
<table>
<thead>
<tr>
<th>Percentage</th>
<th>Component</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>parallel programming training</td>
</tr>
<tr>
<td>25%</td>
<td>individual programming project</td>
</tr>
<tr>
<td>25%</td>
<td>team term project</td>
</tr>
<tr>
<td>25%</td>
<td>research paper</td>
</tr>
<tr>
<td>5%</td>
<td>CIS 631 course help</td>
</tr>
</tbody>
</table>
Parallel Programming Technologies

- Strong focus on learning certain parallel programming methodologies and technologies
  - Objective is to increase your skill to a high level

- Shared memory multi-threading
  - OpenMP ([http://openmp.org/](http://openmp.org/))

- Multicore (accelerator, coprocessor) programming
  - OpenACC ([http://www.openacc.org](http://www.openacc.org))

- Distributed memory message passing

- Do this is a structured training environment
OpenMP Training

- Several excellent tutorials
- A lot of resources online
- Need to identify what specific training materials we want to use

http://www.openmp.org/
NVIDIA Educator Resources

- Part of NVIDIA Developer
- NVIDIA teaching kits
  - Accelerated computing training
    - lecture slides, lecture videos, e-books
    - hands-on labs and coding projects
    - source code solutions
  - Access to online labs using cloud-based GPUs
- NVIDIA Deep Learning Institute (DLI)
  - DLI self-paced labs and workshops
- CUDA and OpenACC
- Need help figuring this out
PGI Community Edition

- PGI products deliver world-class multicore CPU performance, an easy on-ramp to GPU computing with OpenACC directives, and performance portability across all major HPC platforms
  - Version 17.10 is available now
  - Free PGI Community Edition updated to version 17.10

- Features include:
  - OpenMP 4.5 for multicore CPUs
  - OpenACC for multicore CPUs and manycore GPUs
    - CUDA Unified Memory, Tesla V100 GPU support
    - C++ performance optimizations and C++14 lambdas
    - OpenACC PGI Unified Binary for Multicore/Tesla
MPI Training

- Several excellent tutorials
- A lot of resources online
- Need to identify what specific training materials we want to use

http://mpi-forum.org

https://www.open-mpi.org
http://parallelbook.com/

- 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
- Used in CIS 431/531
Reference Algorithms Books

- Leverage parallelism on multicore processors and manycore processors
- Examples of successful programming efforts from industries and domains such as chemistry, engineering, environmental science
- Provides detailed explanations of the programming techniques used, with source code available
Recent OpenACC Books

- Both edited books with chapters by leaders in the parallel computing field.
WOPR Cluster (OACISS)

- WOPR (What Operational Parallel Resource)
- Built as a Next Unit of Computing (NUC) cluster with Intel funds
  - 16x Intel NUC
    - Haswell i5 CPU (2 cores, hyperthreading)
    - Intel HD 4000 GPU (OpenCL programmable)
    - 1 GigE, 16 GB memory, 240 GB mSATA
    - Logitech keyboard and mouse
  - Head node and GigE switch
- Cerberus head node
  - Dell 2x4 core E5-2603v2 CPU, 1.8 GHz, 32GB
  - Linux environment
  - Compilers: GCC, Intel 17, PGI
- WOPR accessed through Cerberus using SLURM
- ISO image with HPC Linux environment
  - Available for booting or running in VirtualBox
NVIDIA Jetson TX1 Cluster (OACISS)

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

[Link to NVIDIA Jetson Developer Site](https://developer.nvidia.com/embedded-computing)

[Link to Jetson TX1 Development Kit](http://www.nvidia.com/object/jetson-tx1-dev-kit.html)
Accelerator Resources (OACISS)

- Variety of NVIDIA GPUs
  - GTX 980
    - Maxwell architecture, 2048 cores
  - K80
    - Kepler architecture, 4992 cores (dual-GPU)
  - Tesla P100
    - Pascal architecture, 3584 cores
  - Quadro GV100
    - Volta architecture, 5120 cores

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

- New HPC cluster at UO
  - 250 Tflops and 1.5 Petabytes
- Maintained by RACS
  - Research Advanced Computing Services (RACS)
    - [https://hpcf.uoregon.edu](https://hpcf.uoregon.edu)
- Talapas specifications

<table>
<thead>
<tr>
<th>Qty</th>
<th>Node Type</th>
<th>Processors (total cores)</th>
<th>Memory</th>
<th>Local Storage</th>
<th>Networking</th>
<th>Accelerator</th>
</tr>
</thead>
<tbody>
<tr>
<td>96</td>
<td>Standard Nodes</td>
<td>dual E5-2690v4 (28 cores)</td>
<td>128GB</td>
<td>200GB SSD</td>
<td>Single Port EDR InfiniBand</td>
<td>N/A</td>
</tr>
<tr>
<td>24</td>
<td>GPU Nodes</td>
<td>dual E5-2690v4 (28 cores)</td>
<td>256GB</td>
<td>200GB SSD</td>
<td>Single Port EDR InfiniBand</td>
<td>Dual NVIDIA Tesla K80</td>
</tr>
<tr>
<td>8</td>
<td>Large Memory Nodes</td>
<td>quad E7-4830v4 (56 cores)</td>
<td>1TB, 2TB, or 4TB</td>
<td>dual 480GB SSD</td>
<td>Single Port EDR InfiniBand</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Repurposed ACISS (to be renamed) Cluster

- **Applied Computational Instrument for Scientific Synthesis (ACISS)**
  - 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

- We are hoping ACISS (to be renamed) will be repurposed soon
What will you get out of CIS 631?

- In-depth understanding of parallel computer design
- Knowledge about how to program parallel computer systems
- Understanding of parallel programming paradigms
- 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 computer science
  - It was there from the very beginning
  - Involves all aspects: architecture, HW/SW systems, languages, programming paradigms, algorithms, theoretical models, ...
  - Parallelism is everywhere

- 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 to 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 default case)
  - Computing problems differ in level / type of parallelism

- Parallelism is all about performance! Really?
Concurrency (a more fundamental concept)

- 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 (constraints on concurrency)
  - 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 (Why?)
  - Must find concurrent execution opportunities
  - Must develop application to execute in parallel
  - Then run application on parallel hardware and hope
- 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, …
  - What software elements 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)
  - Increasing throughput is a related reason (more secondary)

- 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 (big time!)
- 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?
Leveraging Moore’s Law (http://en.wikipedia.org/wiki/Moore's_law)

- 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?
    - try Googling “end of Moore’s Law” and see what you get!

- More transistors = more parallelism opportunities

- Microprocessors
  - Implicit parallelism
    - Pipelining, multiple functional units, superscalar
  - Explicit parallelism
    - SIMD instructions, long instruction works
Gordon Moore predicted that transistor counts would double every two years.

Increased transistor counts translate to increases in processing power.

- What happens to power with more transistors?
- Dennard scaling
  - Power requirements are proportional to area
  - As transistors get smaller, power density stays constant (as a result of the IC materials physics)
  - Effectively smaller transistors consume less power
  - Performance per watt grows with transistor density
  - This is a good thing, until the materials hit a limit
- Unfortunately, Dennard scaling has broken down
  - At very small scales, the same properties do not hold
Power Density Growth

- Power density has significantly increased recently
- Breakdown of Dennard scaling is part of the story

Figure courtesy of Pat Gelsinger, Intel Developer Forum, Spring 2004
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 super hot!
  - Greater clock frequency $\Rightarrow$ greater electrical power
  - Combine this with 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
Memory Wall

Processor-DRAM Memory Gap (latency)

“Moore’s Law”

Processor-Memory Performance Gap:
(grows 50% / year)

μProc
60%/yr.
(2X/1.5yr)

von Neumann bottleneck!!
(memory wall)

CIS 631: Advanced Parallel Computing, University of Oregon
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 to see performance gains
- Need to greater memory bandwidth to keep up
What’s driving parallel computer architecture?

- Single core performance from frequency increase
- Hardware problem becomes a software problem

Credit: National Research Council, Report: The Future of Computing Performance: Game Over or Next Level?
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 (type of SIMD)
  - 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, right?

- 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 supercomputers
- Yardstick for high-performance computing (HPC)
- Benchmark problem (Linpack benchmark)
  - Solve a dense linear system of equations (Ax = b)
  - Scale problem size to achieve maximum
- Report
  - $R_{max}$: maximal performance
  - $R_{peak}$: theoretical peak performance
  - $N_{max}$: problem size needed to achieve $R_{max}$
  - $N_{1/2}$: problem size needed to achieve 1/2 of $R_{max}$
- Updated twice a year at SC and ISC conferences

https://www.top500.org
Supercomputer Performance (1937-2016)

Credit: Wikipedia, Topic: World's most powerful computer?
# Top 10 Supercomputers (November 2017)

<table>
<thead>
<tr>
<th>#</th>
<th>Site</th>
<th>Manufacturer</th>
<th>Computer</th>
<th>Country</th>
<th>Cores</th>
<th>Rmax [Pflops]</th>
<th>Power [MW]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>National Supercomputing Center in Wuxi</td>
<td>NRCPC</td>
<td>Sunway TaihuLight</td>
<td>China</td>
<td>10,649,600</td>
<td>93.0</td>
<td>15.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NRPC Sunway SW26010, 260C 1.45GHz</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>National University of Defense Technology</td>
<td>NUDT</td>
<td>Tianhe-2</td>
<td>China</td>
<td>3,120,000</td>
<td>33.9</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>NUDT TH-IVB-FEP, Xeon 12C 2.2GHz, IntelXeon Phi</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Swiss National Supercomputing Centre (CSCS)</td>
<td>Cray</td>
<td>Piz Daint</td>
<td>Switzerland</td>
<td>361,760</td>
<td>19.6</td>
<td>2.27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cray XC50, Xeon E5 12C 2.6GHz, Aries, NVIDIA Tesla P100</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Japan Agency for Marine-Earth Science and Technology</td>
<td>ExaScaler</td>
<td>Gyoukou</td>
<td>Japan</td>
<td>19,860,000</td>
<td>19.1</td>
<td>1.35</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>ZettaScaler-2.2 HPC System, Xeon 16C 1.3GHz, IB-EDR, PEZY-SC2 700Mhz</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Oak Ridge National Laboratory</td>
<td>Cray</td>
<td>Titan</td>
<td>USA</td>
<td>560,640</td>
<td>17.6</td>
<td>8.21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cray XK7, Opteron 16C 2.2GHz, Gemini, NVIDIA K20x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Lawrence Livermore National Laboratory</td>
<td>IBM</td>
<td>Sequoia</td>
<td>USA</td>
<td>1,572,864</td>
<td>17.2</td>
<td>7.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BlueGene/Q, Power BQC 16C 1.6GHz, Custom</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Los Alamos NL / Sandia NL</td>
<td>Cray</td>
<td>Trinity</td>
<td>USA</td>
<td>979,968</td>
<td>14.1</td>
<td>3.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cray XC40, Intel Xeon Phi 7250 68C 1.4GHz, Aries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Lawrence Berkeley National Laboratory</td>
<td>Cray</td>
<td>Cori</td>
<td>USA</td>
<td>622,336</td>
<td>14.0</td>
<td>3.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cray XC40, Intel Xeons Phi 7250 68C 1.4 GHz, Aries</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>JCAHPC Joint Center for Advanced HPC</td>
<td>Fujitsu</td>
<td>Oakforest-PACS</td>
<td>Japan</td>
<td>556,104</td>
<td>13.6</td>
<td>2.72</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>PRIMERGY CX1640 M1, Intel Xeons Phi 7250 68C 1.4 GHz, OmniPath</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>RIKEN Advanced Institute for Computational Science</td>
<td>Fujitsu</td>
<td>K Computer</td>
<td>Japan</td>
<td>795,024</td>
<td>10.5</td>
<td>12.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>SPARC64 Vllf x 2.0GHz, Tofu Interconnect</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## 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 2.20GHz, 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>3</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>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 (SCSCS), 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 E5-2698v3 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 2.600GHz, 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>
## 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><strong>Tianhe-2 (MilkyWay-2)</strong> - 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><strong>Titan</strong> - 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><strong>Sequoia</strong> - 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><strong>Mira</strong> - 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><strong>Piz Daint</strong> - Cray XC30, Xeon E5-2670 8C 2.600GHz, 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><strong>Stampede</strong> - 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>
<tr>
<td>8</td>
<td>Forschungszentrum Juelich (FZJ) Germany</td>
<td><strong>JUQUEEN</strong> - 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><strong>Vulcan</strong> - 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>
## Top 10 (November 2013)

<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>1,572,884</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 - iDataPlex 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>

Same as 2014!
Top 500 Performance History
Country Distribution History

COUNTRIES

- China
- Korea, South
- Italy
- Canada
- France
- United Kingdom
- Germany
- Japan
- United States
#1: NRCPC: Sunway TaihuLight

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

- **Compute Nodes** have 3.432 Tflop/s per node
  - 16,000 nodes
  - 32,000 Intel Xeon CPU
  - 48,000 Intel Xeon Phi

- **Operations Nodes**
  - 4,096 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
#4: ZettaScaler 2.2

- 19,840,000 cores
  - 10,000 PEZY-SC2 processors
- PEZY-SC2 processor
  - 1,984 cores, 8-way SMT
  - 15,872 threads
- Liquid immersion cooling
#5: ORNL Titan 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
#6: 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**
#7: LANL Trinity (Cray XC40)

- Cray XC40
- Intel Xeon Phi 7250 68C (Knights Landing)
  - 1.4GHz 979,968 cores
- Aries interconnect
  - 0.25 us to 3.7 us MPI latency
  - ~8 GB/s bandwidth
- 14.137 PF
#8: NERSC Cori (Cray XC40)

- **Theoretical peak performance:**
  - Phase I Haswell: 1.92 PFlops
  - Phase II KNL: 29.1 PFlops

- **Sustained performance on NERSC codes:**
  - Phase I Haswell: 83 TFlops
  - Phase II KNL: 562.31 TFlops

- **Total compute nodes:**
  - Phase I Haswell: 2,004 compute nodes, 64,128 cores
  - Phase II KNL: 9,688 compute nodes, 658,784 cores

- **Cray Aries high-speed interconnect with Dragonfly topology**

- **Aggregate memory:**
  - Phase I Haswell: 203 TB
  - Phase II KNL: 1 PB
Intel Xeon Phi (Knights Corner)

- Many Integrated Cores (MIC) architecture

![Diagram of Intel Xeon Phi (Knights Corner) architecture]

**CIS 631: Advanced Parallel Computing, University of Oregon**
Intel Xeon Phi (Knights Landing)

- Changed interconnection architecture
- Beefed up the core

Note: Knights Hill processor was killed and Xeon Phi future is uncertain.
#10: Japanese K Computer – Interconnect

- 80,000 CPUs (SPARC64 VIIIfx), 640,000 cores
- 800 racks
- 8.6 Petaflops (Linpack)

Tofu: Fujitsu’s original 6D mesh/torus interconnect
- High communication performance
- High system scalability
- High fault-tolerance
# Top 500 Linpack Benchmark List (June 2002)

<table>
<thead>
<tr>
<th>Computer (Full Precision)</th>
<th>Number of Processors</th>
<th>$R_{max}$ Gflop/s</th>
<th>$N_{max}$ order</th>
<th>$N_{1/2}$ order</th>
<th>$P_{peak}$ Gflop/s</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Earth Simulator, NEC processors</strong>**</td>
<td>esc</td>
<td>5104</td>
<td>35610</td>
<td>1041216</td>
<td>265408</td>
</tr>
<tr>
<td>ASCI White-Pacific, IBM SP Power 3 (375 MHz)</td>
<td>ilnl</td>
<td>8000</td>
<td>7226</td>
<td>518096</td>
<td>179000</td>
</tr>
<tr>
<td><strong>Compaq AlphaServer SC ES45/EV68 1GHz</strong></td>
<td>psc</td>
<td>3016</td>
<td>4463</td>
<td>280000</td>
<td>85000</td>
</tr>
<tr>
<td>Compaq AlphaServer SC ES45/EV68 1GHz</td>
<td>psc</td>
<td>3024</td>
<td>4059</td>
<td>525000</td>
<td>105000</td>
</tr>
<tr>
<td><strong>Compaq AlphaServer SC ES45/EV68 1GHz</strong></td>
<td>cea</td>
<td>2560</td>
<td>3980</td>
<td>360000</td>
<td>85000</td>
</tr>
<tr>
<td>IBM SP Power3 208 nodes 375 MHz</td>
<td>lbnl</td>
<td>3328</td>
<td>3052.2</td>
<td>371712</td>
<td>4992</td>
</tr>
<tr>
<td><strong>Compaq Alphaserver SC ES45/EV68 1GHz</strong></td>
<td>lanl</td>
<td>2048</td>
<td>2916</td>
<td>272000</td>
<td>4096</td>
</tr>
<tr>
<td><strong>IBM SP Power3 158 nodes 375 MHz</strong></td>
<td>lbnl</td>
<td>2528</td>
<td>2526.9</td>
<td>102400</td>
<td>3792</td>
</tr>
<tr>
<td>ASCI Red Intel Pentium II Xeon core 333MHz</td>
<td>snl</td>
<td>9632</td>
<td>2379.6</td>
<td>362880</td>
<td>75400</td>
</tr>
<tr>
<td>ASCI Blue-Pacific SST, IBM SP 604E (332 MHz)</td>
<td>ilnl</td>
<td>5808</td>
<td>2144.4</td>
<td>431344</td>
<td>432344</td>
</tr>
<tr>
<td>ASCI Red Intel Pentium II Xeon core 333MHz</td>
<td>snl</td>
<td>9472</td>
<td>2121.3</td>
<td>251904</td>
<td>66000</td>
</tr>
<tr>
<td>Compaq Alphaserver SC ES45/EV68 1GHz</td>
<td>lanl</td>
<td>1520</td>
<td>2096</td>
<td>390000</td>
<td>71000</td>
</tr>
<tr>
<td><strong>IBM SP 112 nodes (375 MHz POWER3 High)</strong></td>
<td>ibm</td>
<td>1792</td>
<td>1791</td>
<td>275000</td>
<td>275000</td>
</tr>
<tr>
<td>HITACHI SR8000/MPP/1152 (450 MHz)</td>
<td>u tokyo</td>
<td>1152</td>
<td>1709.1</td>
<td>141000</td>
<td>16000</td>
</tr>
<tr>
<td><strong>HITACHI SR8000-F1/168 (375 MHz)</strong></td>
<td>leibniz</td>
<td>168</td>
<td>1653.0</td>
<td>160000</td>
<td>19560</td>
</tr>
<tr>
<td>ASCI Red Intel Pentium II Xeon core 333MHz</td>
<td>snl</td>
<td>6720</td>
<td>1633.3</td>
<td>306720</td>
<td>52500</td>
</tr>
<tr>
<td>SGI ASCI Blue Mountain</td>
<td>lanl</td>
<td>5040</td>
<td>1608.8</td>
<td>374400</td>
<td>138000</td>
</tr>
<tr>
<td>IBM SP 328 nodes (375 MHz POWER3 Thin)</td>
<td>noo</td>
<td>1312</td>
<td>1417.1</td>
<td>374000</td>
<td>374000</td>
</tr>
<tr>
<td>Intel ASCI Option Red (200 MHz Pentium Pro)</td>
<td>snl</td>
<td>9152</td>
<td>1338.8</td>
<td>235000</td>
<td>63000</td>
</tr>
<tr>
<td>NEC SX-5/128M8 (3.2ns)</td>
<td>osaka</td>
<td>128</td>
<td>1192.0</td>
<td>129536</td>
<td>10240</td>
</tr>
<tr>
<td>CRAY T3E-1200 (600 MHz)</td>
<td>us government</td>
<td>1488</td>
<td>1127.0</td>
<td>148800</td>
<td>28272</td>
</tr>
<tr>
<td>HITACHI SR8000-F1/112 (375 MHz)</td>
<td>leibniz</td>
<td>112</td>
<td>1035.0</td>
<td>120000</td>
<td>15160</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!!!
### Top 500 Top 10 (2006)

<table>
<thead>
<tr>
<th>Manufacturer</th>
<th>Computer</th>
<th>R\text{max} [TF/s]</th>
<th>Installation Site</th>
<th>Country</th>
<th>Year</th>
<th>#Proc</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 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>2 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>3 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>4 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>5 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>6 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>7 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>8 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>9 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>10 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 10 (Top500 List, June 2011)

<table>
<thead>
<tr>
<th>Rank</th>
<th>Site</th>
<th>Computer</th>
<th>Country</th>
<th>Cores</th>
<th>Rmax [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>
<td>93</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>
<td>55</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>
<td>75</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>
<td>43</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>
<td>52</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>
<td>81</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>
<td>83</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>
<td>82</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>
<td>84</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>
<td>76</td>
</tr>
</tbody>
</table>

Figure credit: [http://www.netlib.org/utk/people/JackDongarra/SLIDES/korea-2011.pdf](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 VIIIfx)</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</td>
<td></td>
</tr>
<tr>
<td></td>
<td>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>
<tr>
<td>Node</td>
<td>Configuration</td>
<td>1 CPU / Node</td>
</tr>
<tr>
<td></td>
<td>Memory capacity</td>
<td>16GB (2GB/core)</td>
</tr>
<tr>
<td>System board(SB)</td>
<td>No. of nodes</td>
<td>4 nodes /SB</td>
</tr>
<tr>
<td>Rack</td>
<td>No. of SB</td>
<td>24 SBs/rack</td>
</tr>
<tr>
<td>System</td>
<td>Nodes/system</td>
<td>&gt; 80,000</td>
</tr>
</tbody>
</table>

Interconnect

<table>
<thead>
<tr>
<th>Topology</th>
<th>6D Mesh/Torus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance</td>
<td>5GB/s. for each link</td>
</tr>
<tr>
<td>No. of link</td>
<td>10 links/ node</td>
</tr>
<tr>
<td>Additional feature</td>
<td>H/W barrier, reduction</td>
</tr>
<tr>
<td>Architecture</td>
<td>Routing chip structure (no outside switch box)</td>
</tr>
<tr>
<td>Cooling</td>
<td>CPU, ICC*</td>
</tr>
<tr>
<td>Other parts</td>
<td>Direct water cooling</td>
</tr>
<tr>
<td></td>
<td>Air cooling</td>
</tr>
</tbody>
</table>

System

LINPACK 10 PFlops over 1PB mem.
800 racks
80,000 CPUs
640,000 cores

New Linpack run with 705,024 cores at 10.51 Pflop/s (88,128 CPUs)
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 programming lab this afternoon
- Parallel computer architectures (Thursday)
- Parallel performance models (Tuesday next week)