International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 01 , January 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

RECENT ADVANCES IN HIGH PERFORMANCE COMPUTING

Amrita Sekhon

Assistant Professor, Department of Computer Science, Bibi Sharan Kaur Khalsa College, Sri Chamkaur Sahib (Ropar) Punjab (India)

ABSTRACT

This paper traces out the history of supercomputers, where it started from and where it has reached now, what all technological changes it underwent along the way. It compiles the various benchmarks available in the arena of high performance computing to judge the power to these giant colossal devices. It also gives the details for the supercomputers that are currently ruling the globe and indicates their future trends.

Keywords: Cluster, graph 500, green 500, grid computing, high performance computing, linpack benchmark, massive parallel processing, teraflop, top 500, vector processing

I. INTRODUCTION

High Performance Computing uses parallel processing to perform advanced application programs in a fast, efficient, reliable manner. This term applies especially to those systems that perform above teraflop that is 10^{12} floating-point operations per second and are used for solving large problems in engineering, science and business. High Performance Computing machines are called supercomputers. These machines are designed to attain the highest possible performance and their computing ability is measured in terms of number of FLOPS that is 64-bit floating-point operations per second. FLOPS are calculated using the following equation:

$$FLOPS = sockets * \left(\frac{cores}{socket}\right) * clock * \left(\frac{FLOPS}{cycle}\right)$$

Going by the famous saying – 'The only thing that is constant is Change', the arena of High Performance Computing has evolved tremendously with steady and continuous developments in the last few decades. According to Moore's Law, the circuit density and hence the processor performance doubles in every 18 months [1]. If the peak performances attained by the fastest computers of their times of the last six decades is plotted, the justification of the accuracy of Moore's Law in the entire lifespan of modern computing can be obtained. An increase in performance of the order of two is visualized on an average with every decade.

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 01 , January 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

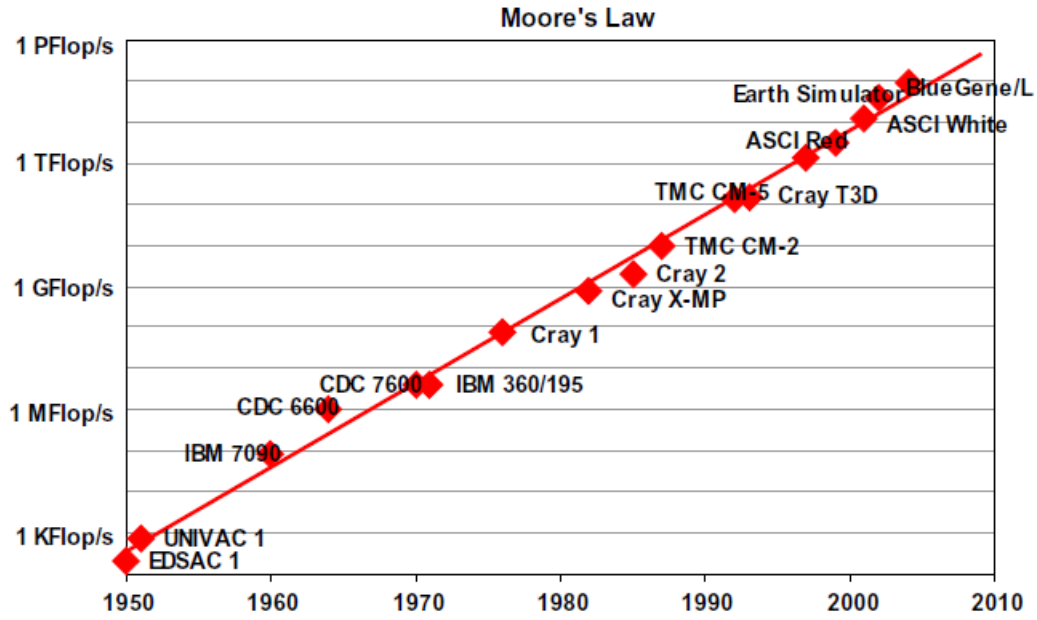

Figure 1: Performance of fast computing systems in the previous six decades compared with Moore's Law [2]

II. BRIEF HISTORICAL JOURNEY OF SUPERCOMPUTERS

The term 'supercomputer' shot to limelight with the release of CDC 6600 (Control Data Corporation) by Seymour Cray of the Colombia University in 1964 as the first supercomputer. This computer used parallelism and innovative design to achieve computational speed of 500 kiloFLOPS for standard mathematical operations. It used FORTRAN compiler.

Supercomputers market is characterized by fast changing vendors, usage of different architectures and technologies [3]. In the late seventies vector computers driven by raw performance marked the beginning of a new era of computing which is the base of modern supercomputers. This vector processing idea was inspired by APL programming language. STAR-100 developed by CDC in 1974 was one of the first computers to use vector processing to achieve 100 megaflops speed [4]. In early eighties, vector computers and conventional computing got integrated. Performance was further increased by improvements in chip technologies and production of shared memory multiprocessor systems. By the end of the eighties massive parallel computing (MPP) with scalable systems and distributed memory were common. They used thousands of processors together. With the RISC revolution, reduced cost and increased performance microprocessors provided better price/performance ratio [5]. In 1980s there was an increase in parallel computing and the integration of supercomputers with conventional computing environment gained importance [2]. The CM-2 supercomputers having 64000 one-bit processors of the Thinking Machines were one of the first main Massive Parallel Processor (MPP) systems. Compared to Symmetric Multiprocessor Systems (SMPs), in the MPPs each processor had its own memory. Then came the transition of CMOS (complementary metal-oxide semiconductor) chip technology from ECL (emitter-coupled logic). In the 1990s, the supercomputer market shifted from vector market to parallel computer market. The usage of RISC design in the microprocessors made the MPPs cheaper and better performer than the vector computers. Computer scientists use the term

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

'Embarrassingly parallel computations' (EPC) to complex problems which can be split easily into smaller parts and solved separately to combine the results later. Amdahl's law (a law of computing) explains that the serial part of the problem determines the maximum speed improvement obtained using a parallel system. In the nineties came the shared memory concept in microprocessors (symmetric multiprocessors SMP) for high end systems, which formed the base for cluster computing in the early 2000s. Cluster is a group of off-the-shelf computers placed locally and connected with high speed LAN. For the first half of 2000s, clusters of workstations and PCs formed the architecture of High Performance Computing, while the later half witnessed new hardware architecture and programming paradigms for HPC to reach where it is today [6]. Grid computing involves several separate computers placed at different places and connected using the Internet. It uses distributed computing and opportunistic supercomputing approach ie. all the computers are not actively working in the grid always but only those that are available are used.

In the year 1996 Intel of United States announced its ASCI Red (Accelerated Strategic Computing Initiative) as the world's first supercomputer to achieve 1 TeraFLOPs speed (later it reached 2 teraflops) and it provided best reliability, price and performance ever [7]. It was a MIMD massively parallel computing system having 9000 nodes, 12-terabyte hard disk and it used the Pentium Pro processors commonly found in personal computers. In 2004, the JAMSTEC's (Japan Agency for Marine Earth Science and Technology), NEC released the Earth Simulator which achieved 35.9 teraflops speed using 640 nodes each having 8 proprietary vector processors. In 2008, supercomputing reached its milestone when the American supercomputer Roadrunner, a product of IBM achieved 1 PetaFLOPs processing speed by processing over 1.026 quadrillion calculations per second.

Thereafter so many supercomputers came on stage from China, Japan, US, etc. China has made rapid progress in the field of supercomputers. In June 2003, it held 51st position in Top500 list, 10th in 2004, 5th in 2005 and the top position in 2010 [8]. In 2016, China's TaihuLight has earned the position of the world's fastest supercomputer with 93 PetaFLOPS out of 125 peak PetaFLOPS on the Linpack benchmark.

The FLOP journey can be summarized started from a few hundred FLOPS in 1940s (eg. Eniac), KiloFLOPS in 1950s (eg. IBM 704), MegaFLOPS in 1960s (eg. CDC 6600), GigaFLOPS in 1980s (eg. Cray-2), TeraFLOPS in 1990s (eg. ASCI Red) to PetaFLOPS in 2010s (eg. Jaguar).

The current architectural trends of the technologies used in the world's top ranked machines of today can be obtained from the study of a few organizations like the TOP 500, GREEN 500 and GRAPH 500.

III. TOP 500

The TOP-500 uses the Linpack benchmark system to solve a certain dense system of linear equations of Mathematics using the LU decomposition method. Since it is a widely used application available for many relevant systems, it works well and gives a good measure of the system's peak performance. The Top-500 list uses the number of floating point operations per second (flops) as a metric to rank the systems, which is calculated by the implementation of the Linpack benchmark [9]. The TOP-500 uses that version of the benchmark which allows users to scale the problem's size and optimize the software for achieving the best performance for any given machine. This performance number gives a reasonable correction of peak

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

performance measure to reflect the overall performance of any given system. The Top-500 list has been giving a list of the world's fastest supercomputers twice a year since 1993.

The latest edition of TOP-500 list of the world's top supercomputers was announced on June 20, 2016 at the 2016 International Supercomputer Conference in Frankfurt, Germany [10]. This is the 47th edition of the eagerly anticipated, closely watched and richly debated list produced twice a year since1993. The TOP-500 list is compiled by the Lawrence Berkeley National laboratory, California, the University of Tennessee, Knoxville and the ISC Group, Germany.

The world's fastest supercomputer position of the second half of the year 2016 was bagged by **Sunway TaihuLight** (Sunway MPP, Sunway SW26010 260C 1.45 GHz) of China which performs 93 Petaflops per second (quadrillions of calculations per second) on the Linpack benchmark. It has been developed by the NRCPC (National Research Center for Parallel Computer Engineering and Technology) and is installed at the NSC (National Supercomputing Center) in Wuxi, China. It has 10,649,600 computing chores and 40,960 nodes. Its peak power consumption under load is at 6GFLOPS/Watt or 15.37 MW.

The Sunway TaihuLight is three times as efficient and twice as fast as the **Tianhe-2**, which holds the second position in the TOP-500 list. The Tianhe-2 (Milky Way-2: TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT) of China performs at 33.86 petaflops per second with 3,120,000 computing chores and peak power consumption under load 17.81 MW.

The third position in the TOP-500 list is bagged by **Titan** which is a Cray XK7 system of the United States installed at the DOE (Department of Energy's) Oak Ridge National Laboratory. It achieves a speed of 17.59 petaflops per second. It has 560,640 computing chores and its peak power consumption under load is 8.21 MW.

Sequoia (BlueGene / Q, Power BQC 16C 1.60 GHz, Custom IBM) of US installed at Department of Energy's National Nuclear Security Administration -Lawrence Livermore National Laboratory is at fourth position which performs at 17.17 petaflops per second with 1,572,864 computing chores and peak power consumption under load 7.89 MW.

K Computer (SPARC 64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu) of Japan installed at RIKEN AICS (Advanced Institute for Computational Science), Japan is at fifth position which performs at 10.51 petaflops per second with 705,024 computing chores and peak power consumption under load 12.66 MW.

Mira (BlueGene / Q, Power BQC 16C 1.60GHz, Custom IBM) of US installed at Department of Energy's Office of Science –Argonne National Laboratory is at sixth position which performs at 8.58 petaflops per second with 786,432 computing chores and peak power consumption under load 3.945 MW.

Trinity (Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc.) of US installed at Department of Energy's National Nuclear Security Administration-Los Alamos National Laboratory – Sandia National Labs is at seventh position which performs at 8.10 petaflops per second with 301,056 computing chores.

Piz Daint |(Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect , NVIDIA K20x Cray Inc.) of the Swiss National Supercomputing Centre (CSCS), Switzerland is at eighth position which performs at 6.27 petaflops per second with 115,984 computing chores and peak power consumption under load 2.325 MW.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

Hazel Hen (Cray XC40, Xeon E5-2680v3 12C 2.5GHz, Aries interconnect Cray Inc.) of HLRS - Höchstleistungsrechenzentrum Stuttgart, Germany is at ninth position which performs at 5.64 petaflops per second with 185,088 computing chores.

Shaheen II (Cray XC40, Xeon E5-2698v3 16C 2.3GHz, Aries interconnect Cray Inc.) installed at King Abdullah University of Science and Technology, Saudi Arabia is at tenth position which performs at 5.537 petaflops per second with 196,608 computing chores and peak power consumption under load 2.83 MW.

These were the top 10 high performance computing devices as per the TOP-500 list chart of the second half of the year 2016.

IV. GREEN 500

The GREEN 500 list is also evaluated using the Linpack benchmark system which is a compute-intensive application. It ranks HPCs according to the energy efficiency. The first Green 500 list was announced on 15 November, 2007. To measure the efficiency of systems used for GREEN 500 list, the power consumption factor is included along with the performance factor. It considers the total power consumption used for executing the program in terms of Flops per Watt.

According to the June 2016 report of the GREEN 500 list [11], **Shobu** of Japan, which is a PEZY Computing /Exascaler ZettaScaler 1.6 System that achieves a speed of 6.67 GFlops/watt has secured the first position. It is installed at the Advanced Center for Computing and Communication at RIKEN in Japan. The second position has been awarded to **Satsuki** of Computational Astrophysics Laboratory, RIKEN, Japan. The third position has been awarded to **Sunway TaihuLight** of National Supercomputing Center in Wuxi, China.

V. GRAPH 500

The GRAPH 500 list is evaluated based on a graph based application and not the Linpack system. The graph based application system is on the contrary a data intensive application. This system consists of two kernels viz. one to build a graph from raw data and the other to operate on the graph thus built. The first kernel builds a weighted undirected graph structure whereas the second kernel works using the BFS or the Breadth First Search algorithm on the structure. The number of edges traversed per second maps closely to application domain and this is used as a metric to rank the systems in the GRAPH 500 list.

The **K Computer** (SPARC 64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu) of Japan installed at RIKEN AICS (Advanced Institute for Computational Science), Japan with 82944 nodes and 663552 cores is at numero uno rank of the GRAPH 500 list of June 2016 [12]. The second position is held by **Sunway TaihuLight** of National Supercomputing Center in Wuxi with 40768 nodes and 10599680 cores. The third position is held by DOE/NNSA/LLNL **Sequoia** of Lawrence Livermore National Lab with 98304 nodes and 1572864 cores.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

VI. FUTURE OF HIGH PERFORMANCE COMPUTING

Supercomputers have shown great relevance in history as they have played a key role in addressing problems of social importance, in scientific discoveries and in making several important advancements in critical aspects of the defense sector. Currently they are being used for climate prediction, transportation, defense intelligence, manufacturing industry, earthquake modeling, handling challenging problems in the stock-market, social health and safety and every other area requiring basic science understanding along with mainstream computing itself. Thus supercomputers are having great influence in the future progress of mankind.

Keenly observing the current trends of growth and development of High Performance Computing, they can undoubtedly be expected to reach 1 exaFLOPS (EFLOPS) level by the year 2018. In December 2009, Cray Inc. had announced about their plans to manufacture supercomputers with 1 EFLOPS before 2020. Erik P. DeBenedictis of Sandia National Laboratories had theorized the requirement of a zettaFLOPS (ZFLOPS) supercomputer for a complete weather forecasting model of a time span of two weeks [13]. Keeping in view the current high speed growth of high performance computing, we can very well expect the trotting of the globe with euphoric performance systems of unbelievably great capabilities by around 2030.

As quoted by James Bailey (New Era in Computation Ed. Metropolis & Rota) *"We are all still trained to believe that orderly, sequential processes are more likely to be true. As we were reshaping our computers, so simultaneously were they reshaping us May be when things happen in this world, they actually happen in parallel."*

The renewal of these three lists twice a year showing the rapid change in the arena of HPC yet again emphasis on the famous saying –'The only thing that is constant is change'.

REFERENCES

[1] G.E. Moore, Cramming more components onto integrated circuits, Electronics 38 (8), 1965.

[2] E. Strohmaier, J. J. Dongarra, H. W. Meuer, H. D. Simon, *Recent Trends in the marketplace of High Performance Computing*, *Parallel Computing 31*, 2005, 261-273.

[3] E. Strohmaier, J.J. Dongarra, H.W. Meuer, H.D. Simon, *The marketplace of high-performance computing*, *Parallel Computing 25*, 1999, 1517–1544.

[4] L. D. Fosdick, *An Introduction to high-performance scientific computing* (US: MIT Press, 1996), 418.

[5] G.V. Wilson, *Chronology of major developments in parallel computing and supercomputing*, www.unipaderborn.de/fachbereich/AG/agmadh/WWW/GI/History/history.txt.gz.

[6] P.R. Woodward, *Perspectives on supercomputing, Computer (10)*, 1996, 99–111.

[7] M. T. Heath,A. Ranade, R. S. Schreiber, *Algorithms for parallel processing* (Volume 105, Springer, 1998), 323.

[8] Graham, Susan L.; Snir, Marc; Patterson, A. Cynthia, *Getting up to speed: the future of supercomputing* (National Research Council, Division on Engineering and Physical Sciences, Computer Science and Telecommunications Board, Committee on the Future of Supercomputing, 2005), 188.

[9] http://www.netlib.org/benchmark/hpl/

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

[10] https://www.top500.org/lists/2016/06/

[11] http://www.green500.org/?q=lists/green201606

[12] http://www.graph500.org/results_jun_2016

[13] DeBenedictis, P. Erik, *Reversible logic for supercomputing*, Proceedings of the 2nd conference on Computing frontiers. New York (NY: ACM Press, 2005), 391–402.