Inside Parallel Computing: present and future.

Roberto Hoyos

Emerging Information Technologies march, 2006

Abstract:

In this paper we will discuss how the new trends in parallel computing are a changing scenario where funding will be more limited and is putting pressure to obtain the most from a scarce-resources schema. The most promising scientific research activities depend heavily on computers, but is in itself, in computer science, where a major task has been uncovered: to improve in the quality of the parallel computing, rather than the quality. There are limits imposed to the evolution of hardware, so the next choice has to come from the software and architecture. To this end, there has been a revival of old methods, and the need for better algorithms to suit the growing needs for limited computer resources.

1. Introduction

To make a difference between Supercomputing and Parallel Computing is extremely difficult. In fact, modern supercomputers work in a parallel fashion, making the concepts super and parallel almost undistinguishable. Because of that, we have split this research into two main areas: this will be the technical approach to supercomputing, where we will present in section 2 general information of parallel computing, and in section 3 the architectures and specific problematic involved, both hardware and software, in a more down-to-earth way; whereas we will outline the social and political implications of supercomputing on a further work. As a matter of fact, supercomputing and parallel computing are so intricately entwined that we may only mention some important aspects of their history (like the role of the Cold War in their development), section 5, in the understanding that we will go deeper in further research [that is the second presentation: Supercomputing].

Because of time limitations we will not present detailed mathematical analysis or algorithms –in fact we won't present any– but we will mention why these are important and why there has been a disparity in their development. We will see how it has evolved and explain why it has reached a level where it can no longer depend on hardware to continue.

2. A brief Introduction to Parallel Computing

According to Ian Foster [1] Parallel Computing is the "use of a *set of processors* that are able to work *cooperatively* to solve a computational problem.

The cutting-edge of scientific research requires heavy computational power. The most promising endeavors for the next century rely, more and more, on computers. These are some of the disciplines that use parallel computing[2].

- Astrophysics: *Modeling the Big Bang*
- Mathematics: *Fractals*
- Medicine: *Cerebral Function Mapping, Molecular Dynamics of Cancer, Protein Mapping*
- Physics: *Superconductors*
- Meteorology: *Modeling the Weather and Natural Phenomena*

2.1 Parallel Computing Classification

We can divide Parallel Computing in three main areas:

1.- *Cluster Computers*
2.- *Parallel Vector Processors*
3.- *Network Computing*

Cluster Computers

Rely on large groups of processors, called *clusters*, which are positioned in groups, close to each other, and can be to the size of several hundred. Because of the increase in their performances, they were the leading trend trough the nineties [3].

An example of the implementation of clusters is the European Centre for Medium Range Weather Forecasts [4], buying two clusters of the IBM 1600 series and is presented in the next page.

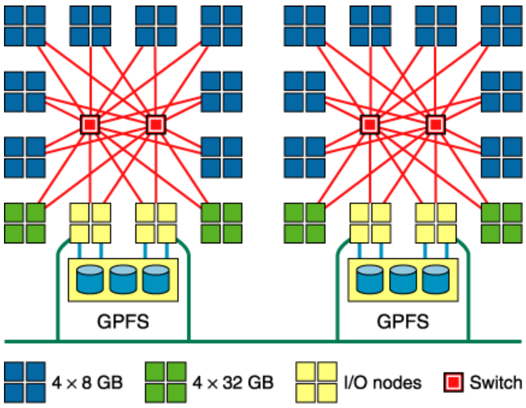

Fig 1. Two-Cluster Array. IBM series 1600. ECMWF.

As shown in the picture, each cluster has 40 interconnected processors, and eight nodes.

Two nodes is the minimum computer hardware required to build a cluster [5].

Parallel Vector Processors

Process the data in chunks of information, of *fixed* size, instead of bit by bit. It is especially handy for large numeric calculations, and its architecture can perform more than one operation in just one instruction [6], but are more expensive than the cluster computers [7]. It was a good idea in the sixties, because it was centered on the data rather than the instructions. It can prove to be good today, especially in multimedia.

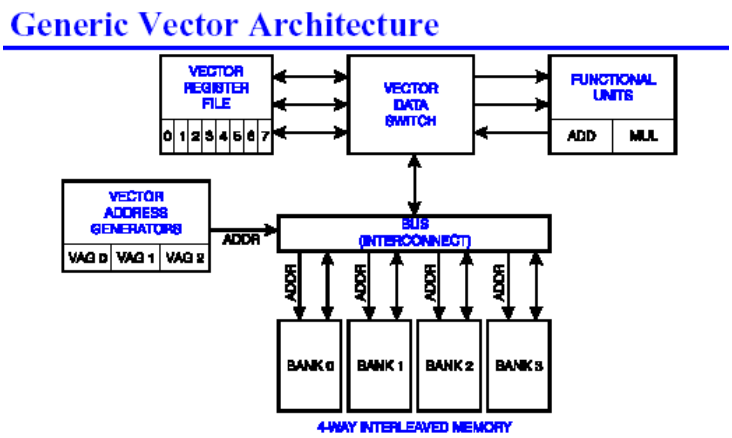

Fig 2. Generic Vector Architecture. Carnegie Mellon [8].

Vector architecture is expensive when compared to the Von Neumman model, because it has very specialized components. But it requires much less processors.

Network Computing

Refers to the use of the Internet to join computers from different geographical regions to perform a common task, where each computer does the processing of a fraction of the problem.

Google is involved in a project to use the idle time of personal computers to compute pieces of scientific problems [9], in cooperation with Stanford University, to map the protein structure [10]. There were other projects like SETI (Search for Extra-Terrestrial Intelligence), from Berkeley, to find patterns in cosmic radio signals [11].

3. Problematic: Hardware Approach

Computer Speed.- The first problem we encounter is that there is a growing gap between clock cycle times and memory access in a computer. We are reaching speeds beyond 1 gigahertz, which has been growing exponentially the last ten years while memory access is being growing only linearly [12].

1.13 GHz meaning clock cycles of around 0.88 nanoseconds.

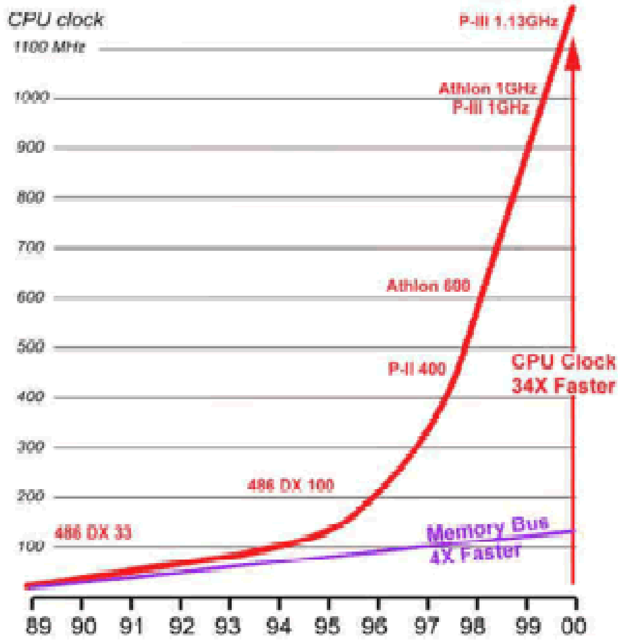

Fig. 3 Gap between CPU speed and Memory Access' Speed. University y of Texas.

There has to be a way of improving the fetch cycle so that operations cannot be wasted because that disparity. Several attempts to overcome this difficulty are sophisticated techniques of *caching* and pipelining (create subsets of instructions that can be done in a parallel way).

Massive Scalar Architecture.- The fastest computer in the world, Blue Gene, has 32,768 processors and can perform more than 70,000 billion of floating point operations per second, while the second fastest, and the rest of the 499 fastest computers in the world, use below 10,000 processors and still compete around with 50, 000 Gigaflops [13]. But a huge number of processors is not a solution while there is plenty of opportunity in algorithm design.

4. Software Approach

The Von Neumman paradigm (the general purpose machine of a single processor) has given the illusion of creating algorithms without regard of the machine used. There is plenty of books who treat parallel computing, but most of them depend exclusively on a single design, for example, José Torres Jiménez has wrote a book based only in the IBM-SP2 [14], and has shown some examples of algorithmic techniques to get from a sequential set of instructions to an algorithm with subsets that can be done at the same time.

Just as efficient sequential algorithms have proven to upgrade the computing power of a single processors in remarkable times (consider sorting, or binary search), well-applied parallel algorithms could improve dramatically the efficiency of a parallel architecture. Numbers cannot be given, because much of the work done is merely theoretical.

5.0 General Considerations for the new trends in Parallel Computing.

According to our research the era of supercomputers starts with World War II and gained a considerable importance during the Cold War [15]. Because the nature of the projects was mainly of military concern and funding, there were no restraints for the investments and maintaining of the equipment. Nowadays, there is still a government concern, but as supercomputers had reached the market, economy is a new issue for which we have to be acquainted of.

Japanese Companies, like *Fujitsu*, *Hitachi*, and *NEC* are becoming the leading selling companies for parallel technology, heading against *Intel*, *Cray* and *IBM* in the market [16].

5.1 Who is behind?

Companies

- Fujistu
- Hitachi
- NEC
- Pacific Northwest Laboratory.
- Intel's Supercomputing Systems Division

Universities

- Caltech
- University of Chicago
- In Mexico: UNAM and ITESM

Government

- Argonne National Laboratory
- Defense Advance Research Projects Agency – DARPA
- Department of Energy – DOE
- NASA
- Center for Research on Parallel Computation

4.2 What is being done?

The Department of Advance Research Projects Agency, the one who developed the Internet, along with the Department of Energy and companies such as SUN Microsystems are working in a project to revive Vector Computers [17]. The novelty about this news is that the this project is expected to reach one *petaflops*, and enormous amount of floating point operations per second.

6.0 Conclusions:

The new trend is to get the more computational power with the less resources. Since WWII until the mid 90's major parallel computing project funding has come from government agencies and universities, making any economic concerns not a factor in the search for speed and computing power. Nowadays the trend is privately funded, and its goal is to obtain the greatest computer power with the less resources.

Improve the Fetch Cycle. CPU's are becoming faster and faster, but memory access is still limited. If the gap is not resolved systems will become unstable, losing lots of clock cycles just to synchronize the process.

To go to a practical Parallel Architecture. There are several theoretical models of how a parallel environment should work; given their complexity it is better to this day to use 'clusters of clusters', but relying completely on a large number of processors will not be an option in the next 10 years, although it has been cheaper that way.

Improve Parallel Algorithms. The sequential, Von Neumann paradigm has given the 'illusion' of algorithms practically independent of the machine used. Intensive research is been done to create algorithms that can be used in a parallel fashion.

The return of vector computing. Efficiency of the vector computers has proven to be up to the challenge to overthrow the massive cluster array of processors, especially operations regarding the manipulation of numbers.

7.0 References:

[1] Ian Foster, Argonne National Laboratory: Mathematics and Computer Science Division. [online] University of Chicago. http://www-unix.mcs.anl.gov/dbpp/text/node1.html [2005, Jan, 26]

[2] Electronic Visualization Laboratory [online] http://www.evl.uic.edu/EVL/SHOWCASE/cave/vrpaper/applications.html [2005, Jan 31]

[3] *Vector Processing*, Wikipedia.org [online] Wikimedia Org http://en.wikipedia.org/wiki/Vector_processing [2005, Jan 31]

[4] European Centre for Medium-Range Weather Forecasts (Ecmwf) [online] http://www.ecmwf.int/services/computing/overview/ibm_cluster.html [2005, Feb 1]

[5] *Two-Node Cluster*, Wikipedia.org, idem.

[6] John Morris, Centre for Intelligent Information Processing Systems [online] The University of Western Australia http://ciips.ee.uwa.edu.au/~morris/CA406/vector.html [2005, Jan, 26]

[7] Steven J. Vaughan-Nichols, Linux Network Inc [online] www.linuxnetworx.com/news/pdf/IEEEComputer.pdf [2005, Jan, 31]

[8] Philip Koopman. Vector Architecture. [online] Carnegie Mellon University. http://www.ece.cmu.edu/~ece548/handouts/16v_arch.pdf [2005, Feb, 2]

[9] Google Labs. [online] Google Corporation. http://toolbar.google.com/dc/offerdc.html [2005, Jan 28]

[10] Folding@home [online] Stanford University http://folding.stanford.edu/ [2005, Jan 30]

[11] Seti@home [online] Berkeley University. http://setiathome.ssl.berkeley.edu/ [2005, Jan 30]

[12] Armando Rodriguez and John Salinas [online] Department of Computer Science, University of Texas, http://www.cs.utsa.edu/~jsalinas/cs5513/selectionCache.pdf [2005, Feb 02]

[13] Top500. The list of the fastest supercomputers in the world. [online] Top500.org http://www.top500.org/lists/plists.php?Y=2004&M=11 [2005, Feb, 2]

[14] José Torres Jiménez and Eduardo Arturo Rodríguez Tello, *Conceptos de Cómputo Paralelo*. Trillas, México, 2000.

[15] Time Table of Supercomputers. http://en.wikipedia.org/wiki/Supercomputer Idem

[16] Christian Bischof and Horst D. Simon [online] National Energy Research Scientific Computing Center, Department of Energy http://www.nersc.gov/~simon/Papers/NASA/implementation.pdf [2005, Jan 29]

[17] Linuxnetworx [online] www.linuxnetworx.com/news/pdf/IEEEComputer.pdf [2005, Jan, 31]

Appendix:

A1 to A2.- *Images of the Two-Cluster IBM 1600 series. ECMWF.*

A3.- *Performance of Supercomputers, where: ‘o’ is uniprocessor; ‘+’ is vector computers; and ‘x’ massive-scalar computers. Argonne National Laboratory.*