ADVANTAGES AND CHALLENGING ISSUES OF MICRO-MULTICORE PROCESSOR TECHNOLOGY

N. Dhasarathan Alias Raja[1], Avnish Kumar Shukla[2]
Department of Electronics and Communication Engineering,
Maharishi University of Information Technology (U.P)

ABSTRACT

A multicore processor has two or more independent computing/processing units (cores) on the same chip. Multiple cores have advantage that they run on lower frequency as compared to the single processing unit, which reduces the power dissipation or temperature. These multiple cores work together to increase the multitasking capability or performance of the system by operating on multiple instructions simultaneously in an efficient manner. This also means that with multithreaded applications, the amount of parallel computing or parallelism is increased. Multi-core processors have reformed the digital world today which is inhabited by us and still the endless efforts are being carried out to generate faster and smarter chips. Since the most significant direction to expand the performance of processor is Multi-core, the manufacturers are highly focused on this technology. Multi-core processors allow higher performance at lower energy which can boost the performance of mobile devices that operate on batteries. Also, the consorting capability of these multiple cores upturns the multitasking ability of the system. The applications or algorithms must be designed in such a way that their subroutines take full advantage of the multicore technology. Each core or computing unit has its own independent interface with the system bus.. But along with all these advantages, there are certain issues or challenges that must be addressed carefully when we add more cores. In this paper, we discuss about multicore processor technology. In addition to this, we also discuss various challenges faced such as power and temperature (thermal issue), interconnect issue etc. when more cores are added. That is the cores work together by executing multiple instructions simultaneously using parallelism in an efficient manner. In addition to this, the adjacency of multiple CPU cores on the same die grants the cache coherency to operate at a much higher speed than what would have happen if the signals had to travel off-chip. This in turn leads to the less degradation of signals since they had to travel shorter distances between different CPUs. All these factors have led to the evolution of many new multi-core processors like R65C00 and many others by Intel, AMD, etc. This paper introduces the technology of multi-core processors and its advantages in today's world. Finally, the paper concludes by detailing the challenges that are presently faced by multi-core processors.

KEYWORDS: CMP (Chip Multiprocessor), Clock, Core, ILP (Instructions Level parallelism), TLP (Thread level Parallelism), Amdahl's law, cache coherence.

1. INTRODUCTION

The multicore processor comprises of two or more cores or computational/processing units that operate in parallel to read and execute instructions. These multiple processing units or cores are fabricated on a single die. So, its also called a Chip Multiprocessor (CMP). The key factor about multicore processor is that it gives the same performance of a single faster processor at lower power dissipation and at a lower clock frequency by handling more tasks or instructions in parallel. [1] This enhances the ILP (Instructions Level Parallelism). The performance of a processor is a function of three major factors, which includes IPC (instructions per cycle), CPI (clock cycles per instruction) and clock cycle (or clock frequency). The IPC can be increased by increasing ILP and TLP (thread level parallelism). The CPI can be improved by the techniques of pipelining. But there is a limiting factor towards increasing the clock frequency. [2]

The dynamic power dissipation is given by-

$$P = QCV^2f \qquad 1.1$$

- ✓ Q- Number of Transistors
- ✓ C- Load Capacitance
- ✓ V- Supply Voltage
- ✓ f- Clock Frequency

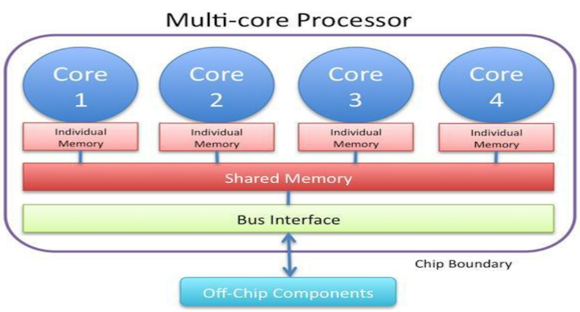

Fig.1. Quad cores sharing memory and bus interface

The flourishing market and the demand for increasing performance led the industries to manufacture quicker and smarter chips. Some of the techniques that have been advised to improve the performance include pipelining and different levels of parallel processing like data level and instruction level parallelism which has proved to be very effective. One such technique that manufacturers came up with; which improves the performance significantly is a new design of processors called as multi-core processors. This concept of multi-core processors started evolving when Gordon Moore, the co-founder of Intel and Fairchild Semiconductor, in April 1965, wrote an article for Electronics magazine titled "Cramming more components onto integrated circuits" which later came to be called as Moore's law. Obviously it should be considered as an observation or projection and not a physical or natural law. It is the consideration that the number of transistors in a dense integrated circuit doubles approximately every two years and projected this rate of growth would continue for at least another decade. In 1975, looking forward to the next decade, he revised the forecast to doubling every two years. Intel came up with its first dual core processor in 2005. A dual core processor contains two cores-(Intel Core Duo, AMD Phenom II X2), a quad core processor has four cores (Intel core i5 and i7 processors, AMD Phenom II X4), a hexa core processor has six cores(AMD Phenom II X6, Intel Core i7 Extreme Edition 980X), a Octa core processor has eight cores (AMD FX-8350, Intel Xeon E7-2820), a deca core has ten cores (Intel Xeon E7 2850) or more. There are various topologies to interconnect cores such as Ring topology, Bus topology, Two-dimensional mesh and crossbar. The choice of a particular topology is a crucial factor as it affects performance parameters like speed, latency etc. The type of topology also affects area consumed on a chip and power dissipation.

In addition to this, multicore technology uses homogeneous and heterogeneous cores. In homogeneous configuration, all the cores are identical and each core has same hardware. These cores use divide and rule approach for improving the performance by dividing a more complex application into less complex applications and execute them simultaneously. There are many other benefits of this approach such as reusability, simpler design etc.[1]

In heterogeneous cores, there are dedicated application specific cores that work on specialized applications. For example a system comprising of a DSP core that handles a multimedia application requiring intensive mathematical computations, while other cores handle some other applications simultaneously. Heterogeneous core is more complex, but has its own benefits also. Multicore Processors sometimes take advantage of both homogeneous and heterogeneous configurations to improve performance. The IBM multicore processor, CELL uses this approach. [1] Various applications that benefit from multicore technology are multimedia applications, DSP , servers, graphics, compilers etc. and also applications which exhibit thread level parallelism. The efficient use of multicore technology requires high level of parallelism.[3]

Some multicore processors allow one or more cores which are not required at times to turn off to save power.

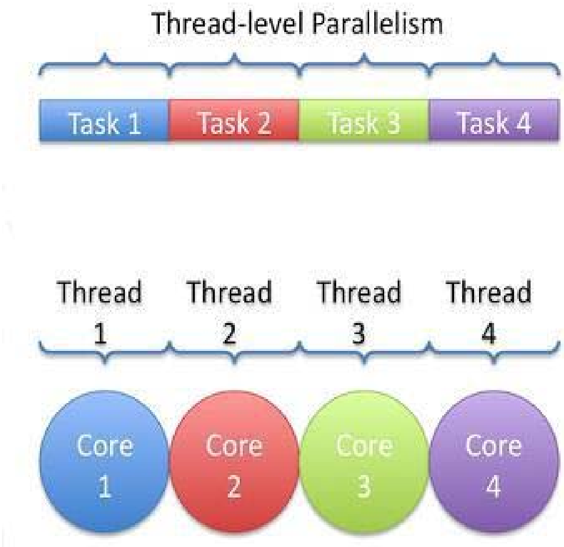

Fig.2. Multiple threads in different Cores[3]

The performance of a multicore processor strongly depends on the design of algorithm or application which is governed by Amdahl's Law. This law was given by Geve Amdahl, a computer architect. This law is used to compute the theoretical maximum speed with multiple processors in parallel computing. This law states that the speed of a process in parallel computing is limited by the time needed for the sequential part of the process. Thus A multi-core processor provides lot of advantages especially for those looking to expand their system's multitasking computing power. In this article, our main goal is to describe some of the significant challenges of multi-core processors. In addition, the paper also discusses the advantages along with its basic concept.

2. MULTICORE PROCESSORS

Multi-core processors have been in reality since the past decagon, but however it have earned more significance off late due to some of the technological limitations like large throughput and long-term battery life with tremendous energy efficiency.

A multi-core processor comprises of a single computing component with two or more independent actual computational or processing units, called as "cores", that which are the units that operate in parallel to read and execute programming instructions. These are simple CPU instructions like arithmetic operations, move data, and branch instructions, but the prime thing is that multiple cores can execute multiple instructions at the same time, thus increasing the overall speed for programs similar to parallel computing. Manufacturers typically fabricate the cores onto a single integrated circuit die or onto multiple dies in a single chip package. Hence it's also called as a Chip Multiprocessor (CMP). The prime factor about multicore processor is that the performance of these processors is not degraded as compared to a single processor even with lesser power dissipation and at a lower clock frequency. The performance of any processor mainly depends on following three major factors:

1. Instructions per cycle (IPC), which can be improved by increasing Instruction Level Parallelism and Thread Level Parallelism.

2. Clock cycles per instruction (CPI) which can be improved by the techniques of pipelining and

3. Clock frequency.

But if we increase the clock frequency beyond certain limit, the power dissipation increases which in turn causes since P is directly proportional to f. So, the idea of multicore technology is to use more than one core instead of one at a comparatively lower frequency. These cores execute multiple threads that is, multiple instructions while accessing multiple data from different parts of memory. Each core is having its own cache while the system bus and main memory is shared by all cores.

3. ADVANTAGES OF MULTICORE PROCESSORS

The Multi-core processors can deliver a very high performance benefits for multi-threaded software by adding more and more processing power with minimal latency, given the proximity of the processors.

1. The most significant benefits will be viewed in many applications such as Customer Relationship Management, larger databases, Enterprise Resource Planning, ecommerce and virtualization.

2. The more threaded applications will clearly get more benefit. In a due course of time, this trend is beginning to and will shape the future of software development towards parallel processing.

3.Meantime, we can target the quad-core servers for those types of applications which can get the maximum work out of them.

The good processing speed of the multicore processors is due to the multiple cores which operate simultaneously on instructions, at lower frequency than the single core. At the same clock frequency, the multicore processor will process more data than the single core processor. In addition to this, multicore processors deliver high performance and handle complex tasks at a comparatively lower energy or lower power as compared with a single core, which is crucial factor in appliances such as mobile phones, laptop etc. which operate on batteries. Also, since the cores are fabricated close to each other on the same chip, the signals travel shorter distance between them due to which there is less attenuation of signals. Since the signals don't attenuate much, more data is transferred in a given time and there is no need of repeating the signals.

4. CHALLENGES INVOLVED IN MULTI-CORE PROCESSORS

The two main characteristics that uniquely identify Multi-core architectures are heterogeneity and the massive parallelism in addition to any arbitrary topology for interconnection of processors or cores.

A. Memory Hierarchy

One of the most demanding challenges that are faced by multi-core processors today is its memory system. The execution of a program is often bounded by the memory bottleneck which happens not due to the non-availability of processor and its low speed but because of the fact that a heavy portion of applications always lies in main memory until it is executed by the processor. Moreover, multi-core processors can even make the situation more badly. Thus, care should be taken to conserve memory bandwidth and avoid memory contention. Also, it is significant to note that the memory hierarchies in parallel machines are more difficult than in the uni-processor's machines, especially in multi-core processors where L2 and L3 cache are shared by the multiple cores with in a chip. This leads to more complicated memory hierarchy design in CMPs.

B. Cache Levels

The current trend of multi-core processors which are general purpose use 3-level cache scheme consisting of L1, L2 and L3 cache. The present 3-level cache system is working fine with the current dual, quad and eight cores processors. In this cache model, the L2 cache may be private, shared or split. But as number of cores increases, this may cause the bottleneck and result in data traffic congestions and performance degradation.

C. Cache coherence

"Cache coherence is the discipline that ensures that changes in the values of shared operands or data are propagated throughout the system in a timely fashion" [6].

When caches of a common memory resource is maintained by the clients in a system, then complications may arise with the inconsistent data, which is especially a concern in a multi-core environment because of sharing of L1 and L2 cache. Let's take an example of a dual-core processor where each core transfers a block of data from memory into its private cache. First core writes a value to a specific memory location; when the second core tries to read that value from its cache it will not have the updated copy unless a cache miss occurs. This cache miss forces the update operation on the second cores cache entry. If this policy would not have been in place then invalid results would have been produced.

D. Developing Multicore Softwares

One of the significant issues seen with regard to software programs is that they run a bit slower on multi-core processors as compared to single core processors. It is being correctly pointed out that "Applications on multi-core systems don't get faster automatically as cores are increased" [7]. So, the solution is that the programmers must develop programs and applications that can exploit use of the increasing number of processors in a multi-core environment. Also, the time needed to write these softwares cannot be stretched. Moreover, majority of applications used today are written to run on a single core processor, which failed to use the capability of multi-core processors. Hence, the dignified challenge the industry faces is how to port these software programs which were developed years ago to multi-core adaptable software programs. Although it sounds very possible to redesign programs, it's really not a business profitable decision in today's world where in companies have to keep in mind the key parameters like time to market, customer fulfillment and cost cut. It has however been correctly pointed out that "The throughput, energy efficiency and multitasking performance of multi-core processors will all be fully realized when application code is multi-core ready" [8].

E. Level of Parallelism

The level of parallelism of the process is one of the gigantic factors that affect the performance of a multicore processor significantly. Performance will increase with the decrease of completion time of a process. Parallelism can be achieved by Instruction Level Parallelism and Thread Level Parallelism. TLP increases overall parallelism by breaking a program into many small threads and execute them concurrently. Hence to achieve a high level of parallelism, software developers must write such algorithms that can take full advantage of multicore design. Also, companies like Microsoft and Apple have designed their operating systems which can run efficiently on multicore configuration [9].

F. False-Sharing

If two different processors operate on independent data in the same memory address region that reside on the same cache line and the mechanism may force the whole line to interconnect with every data write, forcing the wastage of system bandwidth is termed as False-sharing. In multi-core processors, with the increase in number of cores, the number of private and shared caches will also increase and this might result in increase of false-sharing.

G. Power and temperature

As the number of cores placed on a single chip increases without any modification, the chip will consume more power leading to the generation of large amount of heat, which can even cause your computer to become combust; if the processor overheats. To reduce this unnecessary power consumption, the multicore designs also incorporate a power control unit that can force the unused cores to shut down that are not required at times. The core which overheats in a multicore configuration is called a hot-spot. The chip is architected such that the amount of heat generated in the chip is well distributed across the chip. The majority of the heat in the CELL processor is dissipated in the Power Processing Element and the remaining is spread out across the Synergistic Processing Elements. There are various ways to tackle the problem of power dissipation and temperature which includes thread migration, DVFS (Dynamic Voltage and Frequency Scaling) etc. In the thread migration technique, a low power consuming process or thread is moved to an overheated core. In DVFS (Dynamic Voltage and Frequency Scaling) technique, voltage and frequency of the hot core is reduced since the power dissipation or heating is a function of both voltage and frequency which slightly affects the overall performance also [9].

H. Communication Minimization or Interconnect Issues

Another important feature which impacts the program execution and its performance in multi-core processors is the interaction between on chip components viz. cores, cache, memories and if integrated—memory controllers and network controllers which are used for memory-memory and memory-processor communication; where bus contention and latency are the key areas of concern. In multi-core processors, as the number of cores and respective caches increases, the inter-memory and memory-processor communication is expected to increase exponentially. So, the memory hierarchy of multi-core processors should be designed in such a way that these communications are contained. It is rightly pointed that: "The performance of the processor truly depends on how fast a CPU can fetch data rather than how fast it can operate on it to avoid data starvation scenario" [10]. Special crossbars or mesh technologies have been enforced on hardware to solve this issue. For instance, AMD CPUs employ a crossbar, and the Tilera TILE64 implements a fast non-blocking multi-link mesh [11].

I. Level of Parallelism:

One of the biggest factors affecting the performance of a multicore processor is the level of parallelism of the process/ application.

- ✓ The lesser the time required to complete a process, better will be the performance. Performance is directly related to the amount of parallelism because more the number of processes that can be executed simultaneously more will be the parallelism. Parallelism can be considered at two levels ILP and TLP.
- ✓ TLP increases overall parallelism by breaking a program into many threads (Small Processes) and execute them simultaneously.
- ✓ To achieve a high level of parallelism and an overall high performance, software developers must write such algorithms that can take full advantage of multicore design. In other words, all the cores should be used in the most efficient manner. If the algorithms written are not compatible with the multicore design, then it may happen that one or more cores starve for data. In such a case, the process will run on one of the cores, while other cores will sit idle. So, in a nutshell, the success of multicore

technology strongly depends on the way the algorithms are written. Also, companies like Microsoft and Apple have designed their operating systems which can run efficiently on multicore configuration.

J. Interconnect Issues:

Since there are so many components on chip in a multicore processor like cores, caches, network controllers etc., the interaction between them can affect the performance if the interconnection issues are not resolved properly. In the initial processors, bus was used for communication between the components. In order to reduce the latency, crossbar and mesh topologies are used for interconnection of components. Also, as the parallelism increases at the thread level, communication also increases off-chip for memory access, I/O etc. For issues like this, packet based interconnection is actively used. This packet based interconnection has been used by Intel (Quick path interconnect) and is also being used by AMD.[1][2]

5. CONCLUSION

In this paper, we reviewed the basic concepts of the multicore technology. We studied that it is possible to increase the performance of a processor without increasing the clock frequency, by simply adding more cores. The limitation of a single core processor at a high clock frequency has opened the gates for multicore technology and has become the trend of the industries today. However, adding more cores also gives rise to certain issues or challenges that must be carefully addressed in order to get the most out of multicore technology. We also studied various challenges associated with increasing the number of cores like power and temperature (thermal issues), level of parallelism, interconnect issues etc. The limitations on power and frequency on single core systems have led the foundation for multicore processors. Also these processors differ from the Shared Memory Multi-processors (SMPs) in both hardware and software aspects and gave the solution of running chips at lower frequencies, but added interestingly some new problems. New memory hierarchy and cache designs, cores interconnection patterns and many related hardware issues including new programming models, libraries and related performance tools from software design aspect have to be worked out to use a multi-core processor at full capacity. However, in coming days, a lot of technological advancements are expected in this area of technology which may include a new multi-core programming language and there is large amount of research and development going on in this field to utilize multi-core processors more efficiently. The applications/ algorithms which run on multicore environment must be compatible with it. Research is constantly going on in the areas like developing more efficient applications/ algorithms for multicore environment and also in other areas in order to get the maximum performance throughput from multicore processors. Industries are constantly working towards achieving better and better performance from multicore processors.

REFERENCES

[1] Venu, Balaji. "Multicore processors- An overview."

[2] Jr., Ransford Hyman. "Performance issues on Multicore Processors".

[3] Barbic, Jernej. "Multicore architectures." 2007.

[4] Martin, Christian. "Multicore Processors: Challenges, Opportunities, emerging Trends." Embedded World 2014 Exhibition and Conference . 2014.

[5] http://www- 03.ibm.com /ibm/history /ibm 100/us/en/icons/power4/; access date: March 15, 2016.

[6]http://whatis.techtarget.com/definition/cache-coherence; access date: March 17,2016.

[7] Cass, S. "Multicore Processors Create Software Headaches," Technology Review, vol. 113(3), pp. 74-75, 2010.

[8] R. Ramanathan. "Intel multi-core processors: Making the move to quad-core and beyond,". Technology@Intel Magazine, Dec 2006.

[9] Anil Sethi, Himanshu Kushwah,"Multicore Processor Technology-Advantages and Challenges," International Journal of Research in Engineering and Technology, vol. 04(09), pp. 87-89, 2015.

[10] Moore, S.K., "Multicore is bad news for supercomputers", Spectrum, IEEE, vol. 45, no. 11, pp. 15-15, 2008.

[11] Yaser Ahangari Nanehkaran, Sajjad Bagheri Baba Ahmadi," The Challenges of Multi-Core Processor," International Journal of Advancements in Research & Technology, vol. 2(6), pp. 36-39, June 2013.