

**AN OVERVIEW TO MULTI-CORE PROCESSORS****Jasveen Kaur* and Amandeep Kaur**

Assistant Professor, Guru Nanak Dev University, Amritsar, India.

Article Received on 24/04/2016

Article Revised on 14/05/2016

Article Accepted on 03/06/2016

Corresponding Author*Jasveen Kaur**Assistant Professor, Guru
Nanak Dev University,
Amritsar, India.**ABSTRACT**

Multi-core processors represent a major evolution in computing technology. Microprocessors have revolutionized the world we live in and continuous efforts are being made to manufacture not only faster chips but also smarter ones. A number of techniques such as data level

parallelism, instruction level parallelism and hyper threading (Intel's HT) already exists which have dramatically improved the performance of microprocessor cores.^[1, 2] This paper briefs on evolution of multi-core processors followed by introducing the technology and its advantages in today's world. The paper concludes by detailing on the challenges currently faced by multi-core processors and how the industry is trying to address these issues.

KEYWORDS: Microprocessor, Multi core, multi-threading technology, software challenges and High performance computing (HPC).

1. INTRODUCTION

The microprocessor industry continues to have great importance in the course of technological advancements ever since their coming to existence in 1970s.^[3] The growing market and the demand for faster performance drove the industry to manufacture faster and smarter chips. One of the most classic and proven techniques to improve performance is to clock the chip at higher frequency which enables the processor to execute the programs in a much quicker time^[4, 5] and the industry has been following this trend from 1983 – 2002.^[6] Additional techniques have also been devised to improve performance including parallel processing, data level parallelism and instruction level parallelism which have all proven to be very effective.^[1] One such technique which improves significant performance boost is

multi-core processors. Multi-core processors have been in existence since the past decade, but however have gained more importance off late due to technology limitations single-core processors are facing today^[7] such as high throughput and long lasting battery life with high energy efficiency.^[6] Once upon a time, when human realized that in a normal atmosphere, it is not possible to increase the CPU's clock speed after above standard level (generally at a frequency about 4 GHz the CPU would get too hot and take a lot of electricity). Because of too much heat generated by CPU both Intel & AMD faced problem. High clock speed means high heat & resulted boom the CPU. Faster processing is the basic reason, so that's why technology shifted towards the Multicore processor rather than increase the clock frequency. It was decided to put more than one processor on a single die. That is what we call two core or dual core.

2. History of Multi-Core Processor

The first commercial dual core processor was IBM's Power 4 processor for its RISC servers in 2001. The first dual core processor for home use was Intel's Pentium Processor Extreme Edition 840, released in early April 2005. Less than two weeks later AMD brought their Opteron 800 Series and Athlon 64 X2 processors to market.

Driven by a performance hungry market, microprocessors have always been designed keeping performance and cost in mind. Gordon Moore, founder of Intel Corporation predicted that the number of transistors on a chip will double once in every 18 months to meet this ever growing demand which is popularly known as Moore's Law in the semiconductor industry.^[5, 8, and 9] Advanced chip fabrication technology alongside with integrated circuit processing technology offers increasing integration density which has made it possible to integrate one billion transistors on a chip to improve performance.^[7] However, the performance increase by micro-architecture governed by Pollack's rule is roughly proportional to square root of increase in complexity.^[1] This would mean that doubling the logic on a processor core would only improve the performance by 40%. With advanced chip fabrication techniques comes along another major bottleneck, power dissipation issue. Studies have shown that transistor leakage current increases as the chip size shrinks further and further which increases static power dissipation to large values as shown in Figure 1.^[1] One alternate means of improving performance is to increase the frequency of operation which enables faster execution of programs.^[4, 5] However the frequency is again limited to 4GHz currently as any increase beyond this frequency increases power dissipation again.^[5]

“Battery life and system cost constraints drive the design team to consider power over performance in such a scenario”.^[1] Power consumption has increased to such high levels that traditional air-cooled microprocessor server boxes may require budgets for liquid-cooling or refrigeration hardware.^[3] Designers eventually hit what is referred to as the power wall, the limit on the amount of power a microprocessor could dissipate.^[4]

Semiconductor industry once driven by performance being the major design objective, is today being driven by other important considerations such chip fabrication costs, fault tolerance, power efficiency and heat dissipation.^[9] This led to the development of multi-core processors which have been effective in addressing these challenges.

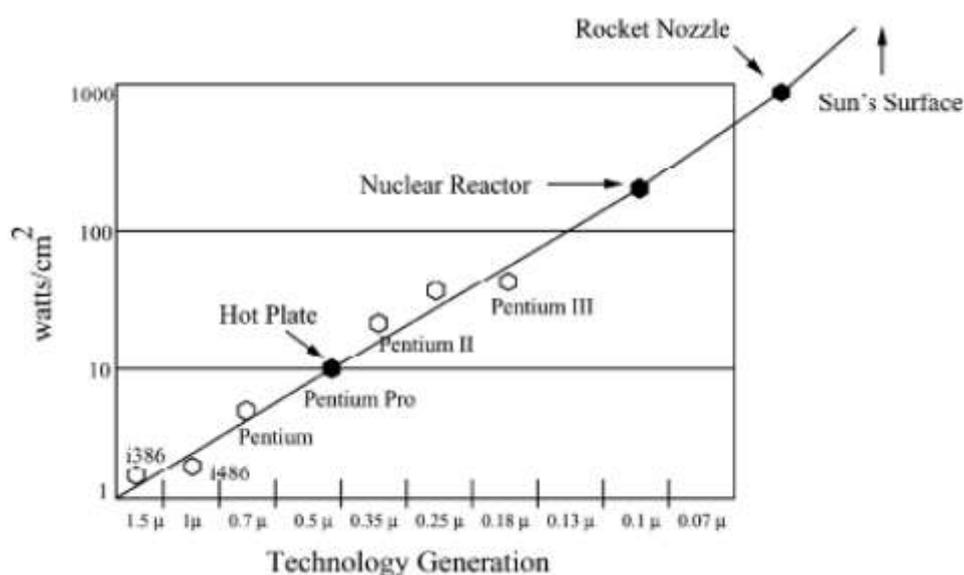


Figure 1: Power density rising^[5]

3. Multi-core processors

A multi-core processor is a processing system composed of two or more independent cores (or CPUs). The cores are typically integrated onto a single integrated circuit die (known as a chip multiprocessor or CMP), or they may be integrated onto multiple dies in a single chip package. A many-core processor is one in which the number of cores is large enough that traditional multi-processor techniques are no longer efficient — this threshold is somewhere in the range of several tens of cores — and likely requires a network on chip.

“A Multi-core processor is typically a single processor which contains several cores on a chip”.^[7] The cores are functional units made up of computation units and caches.^[7] These multiple cores on a single chip combine to replicate the performance of a single faster processor. The individual cores on a multi-core processor don't necessarily run as fast as the

highest performing single-core processors, but they improve overall performance by handling more tasks in parallel.^[6] The performance boost can be seen by understanding the manner in which single core and multi-core processors execute programs. Single core processors running multiple programs would assign time slice to work on one program and then assign different time slices for the remaining programs. If one of the processes is taking longer time to complete then all the rest of the processes start lagging behind.^[6] However, In the case of multi-core processors if you have multiple tasks that can be run in parallel at the same time, each of them will be executed by a separate core in parallel thus boosting the performance as shown in figure 2.^[6]

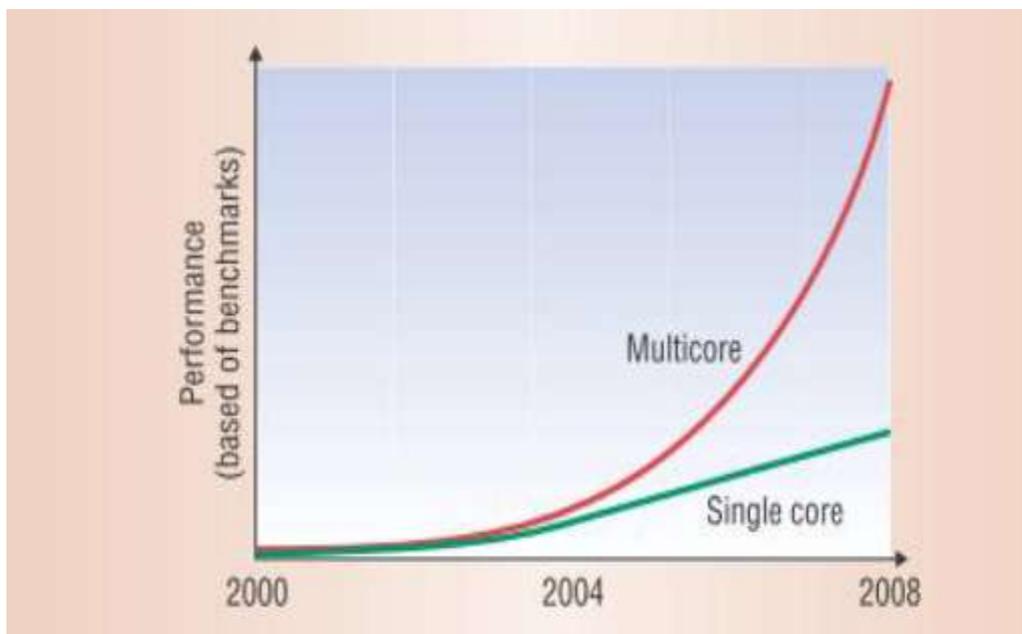


Figure 2: Multicore chips perform better – based on Intel tests using the SPECint2000 and SPECfp2000 benchmarks – than single-core processors.^[6]

The multiple cores inside the chip are not clocked at a higher frequency, but instead their capability to execute programs in parallel is what ultimately contributes to the overall performance making them more energy efficient and low power cores as shown in the figure below.^[6] Multi-core processors are generally designed partitioned so that the unused cores can be powered down or powered up as and when needed by the application contributing to overall power dissipation savings.

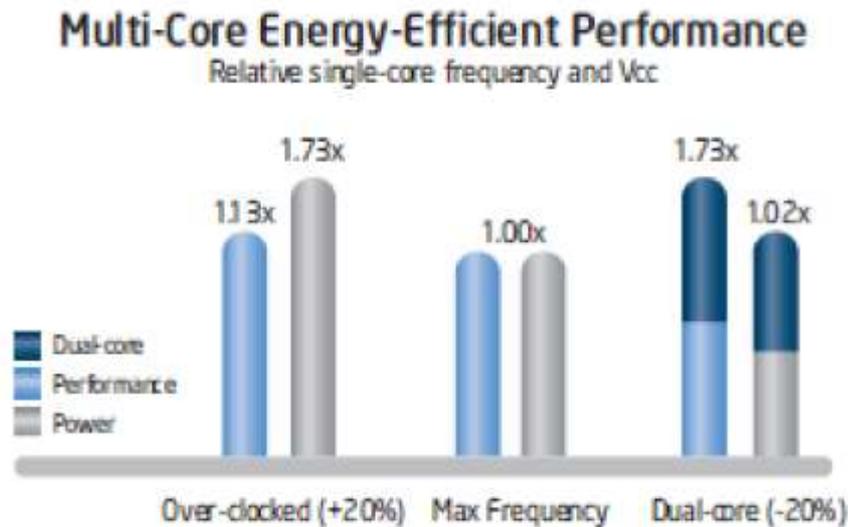


Figure 3: Dual core processor at 20% reduced clock frequency effectively delivers 73% more performance while approximately using the same power as a single-core processor at maximum frequency.^[6]

Multi-core processors could be implemented in many ways based on the application requirement. It could be implemented either as a group of heterogeneous cores or as a group of homogenous cores or a combination of both. In homogeneous core architecture, all the cores in the CPU are identical^[7] and they apply divide and conquer approach to improve the overall processor performance by breaking up a high computationally intensive application into less computationally intensive applications and execute them in parallel.^[5] Other major benefits of using a homogenous multi-core processor are reduced design complexity, reusability, reduced verification effort and hence easier to meet time to market criterion.^[8] On the other hand heterogeneous cores consist of dedicated application specific processor cores that would target the issue of running variety of applications to be executed on a computer.^[5] An example could be a DSP core addressing multimedia applications that require heavy mathematical calculations, a complex core addressing computationally intensive application and a remedial core which addresses less computationally intensive applications.^[5]

Multi-core processors could also be implemented as a combination of both homogeneous and heterogeneous cores to improve performance taking advantages of both implementations. CELL multi-core processor from IBM follows this approach and contains a single general purpose microprocessor and eight similar area and power efficient accelerators targeting for specific applications has proven to be performance efficient.^[5]

“Another major multi-core benefit comes from individual applications optimized for multi-core processors. These applications when properly programmed, can split a task into multiple smaller tasks and run them in parallel.^[6]” Due to the multiple advantages that multi-core processors come along with, most of the processor manufactures started developing them. Intel announced that all its future processors will be multi-core when they realized that this technology can get past the power wall to improve performance.^[9] Other popular processor manufactures namely AMD, IBM and TENSILICA all have started developing multi-core processors. The number of cores in a processor is expected to increase and some even predict it to follow Moore’s law.^[6] TENSILICA’s CEO has expressed his view on this technology stating that “System on a Chip will become a sea of processors. You will have ten to maybe a thousand processors on a chip”.^[2]

However it was observed that there is no throughput improvement while executing sequential programs/single threaded applications on multi-core chips due to underutilization of cores. Compilers are being developed which automatically parallelize applications so that multiple independent tasks can run simultaneously on different cores.^[1]

4. Multicore Architecture

Multicore refers to an architecture in which a single physical processor incorporates the core logic of more than one processor. A single integrated circuit is used to package or hold these processors. These single integrated circuits are known as a die. Multicore architecture places multiple processor cores and bundles them as a single physical processor. The objective is to create a system that can complete more tasks at the same time, thereby gaining better overall system performance. This technology is most commonly used in multicore processors, where two or more processor chips or cores run concurrently as a single system. Multicore-based processors are used in mobile devices, desktops, workstations and servers. The architecture of a multicore processor enables communication between all available cores to ensure that the processing tasks are divided and assigned accurately. At the time of task completion, the processed data from each core is delivered back to the motherboard by means of a single shared gateway. This technique significantly enhances performance compared to a single-core processor of similar speed. When a processor has more than one core to execute all the necessary functions of a computer, it’s processor is known to be a multi core architecture.

A dual-core processor contains two cores, and a quad-core processor contains four cores. A multi-core processor implements multiprocessing in a single physical package. Cores in a

multi-core device may be coupled together tightly or loosely. For example, cores may or may not share caches, and they may implement message passing or shared memory inter-core communication methods. Common network topologies to interconnect cores include: bus, ring, 2-dimensional mesh, and crossbar. All cores are identical in homogeneous multi-core systems and they are not identical in heterogeneous multi-core systems. Just as with single-processor systems, cores in multi-core systems may implement architectures such as superscalar, VLIW, vector processing, SIMD, or multithreading.

A dual-core device contains only two independent microprocessors.

Ideally, a dual core processor is nearly twice as powerful as a single core processor. In practice, performance gains are said to be about fifty percent: a dual core processor is likely to be about one-and-a-half times as powerful as a single core processor.

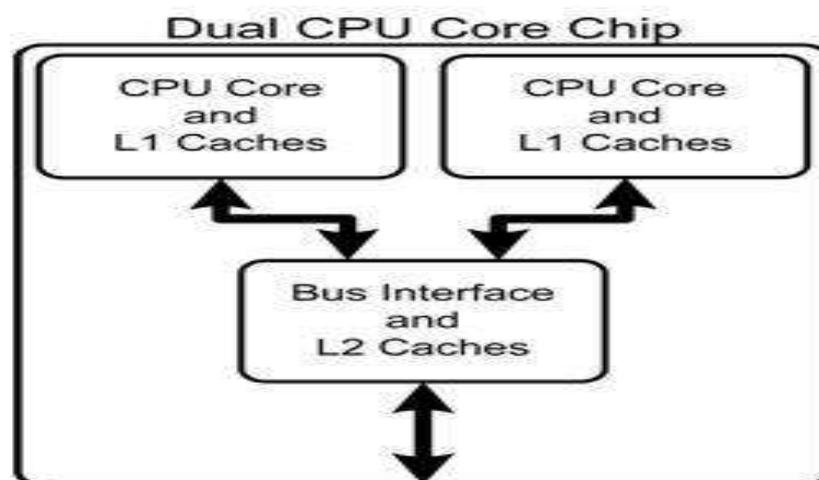


Fig: Dual core

5. Major challenges faced by multi-core processors.

In spite of the many advantages that multi-core processors come with, there are a few major challenges the technology is facing. One main issue seen is with regard to software programs which run slower on multi-core processors when compared to single core processors. It has been correctly pointed out that “Applications on multi-core systems don’t get faster automatically as cores are increased”.^[4] Programmers must write applications that exploit the increasing number of processors in a multi-core environment without stretching the time needed to develop software.^[1] Majority of applications used today were written to run on only a single processor, failing to use the capability of multi-core processors.^[2] Although software firms can develop software programs capable of utilizing the multi-core processor to the fullest, the grave challenge the industry faces is how to port legacy software programs developed years ago to multi-core aware software programs.^[3] Redesigning programs

although sounds possible, it's really not a technological decision in today's environment. It's more of a business decision where in companies have to decide whether to go ahead redesigning software programs keeping in mind key parameters such as time to market, customer satisfaction and cost reduction.^[2]

The industry is addressing this problem by designing compilers which can port legacy single core software programs to 'multi-core aware' programs which will be capable of utilizing the power of multi-core processors. The compilers could perform "code reordering", where in compilers will generate code, reordering instructions such that instructions that can be executed in parallel are close to each other.^[1] This would enable us to execute instructions in parallel improving performance. Also compilers are being developed to generate parallel threads or processes automatically for a given application so that these processes can be executed in parallel.^[10] Intel released major updates for C++ and FORTRAN tools which aimed at programmers exploiting parallelism in multi-core processors. Also alongside Open MP (Open Multiprocessing), an application programming interface which supports multiprocessing programming in C, C++ and FORTRAN provides directives for efficient multithreaded codes.^[3] 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".^[6]

Secondly, on-chip interconnects are becoming a critical bottle-neck in meeting performance of multi-core chips.^[4] With increasing number of cores comes along the huge interconnect delays (wire delays) when data has to be moved across the multi-core chip from memories in particular.^[5] 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.^[6] Buffering and smarter integration of memory and processors are a few classic techniques which have attempted to address this issue. Network on a chip (NoCs) are IPs (Intellectual Property) being developed and researched upon which are capable of routing data on a SoC in a much more efficient manner ensuring less interconnect delay.^[7]

Increased design complexity due to possible race conditions as the number of cores increase in a multi-core environment. "Multiple threads accessing shared data simultaneously may lead to a timing dependent error known as data race condition".^[2] In a multi-core environment data structure is open to access to all other cores when one core is updating it. In the event of a secondary core accessing data even before the first core finishes updating the

memory, the secondary core faults in some manner. Race conditions are especially difficult to debug and cannot be detected by inspecting the code, because they occur randomly. Special hardware requirement implementing mutually exclusion techniques have to be implemented for avoiding race conditions.^[9]

Another important feature which impacts multi-core performance is the interaction between on chip components viz. cores, memory controllers and shared components viz. cache and memories^[9] where bus contention and latency are the key areas of concern. Special crossbars or mesh techniques have been implemented on hardware to address this issue.^[9]

6. Advantages

The proximity of multiple CPU cores on the same die allows the cache coherency circuitry to operate at a much higher clock rate than is possible if the signals have to travel off-chip. Combining equivalent CPUs on a single die significantly improves the performance of cache snoop (alternative: Bus snooping) operations. Put simply, this means that signals between different CPUs travel shorter distances, and therefore those signals degrade less. These higher quality signals allow more data to be sent in a given time period since individual signals can be shorter and do not need to be repeated as often. The largest boost in performance will likely be noticed in improved response time while running CPU-intensive processes, like antivirus scans, ripping/burning media (requiring file conversion), or searching for folders. For example, if the automatic virus scan initiates while a movie is being watched, the application running the movie is far less likely to be starved of processor power, as the antivirus program will be assigned to a different processor core than the one running the movie playback.

Assuming that the die can fit into the package, physically, the multi-core CPU designs require much less Printed Circuit Board (PCB) space than multi-chip SMP designs. Also, a dual-core processor uses slightly less power than two coupled single-core processors, principally because of the decreased power required to drive signals external to the chip. Furthermore, the cores share some circuitry, like the L2 cache and the interface to the front side bus (FSB). In terms of competing technologies for the available silicon die area, multi-core design can make use of proven CPU core library designs and produce a product with lower risk of design error than devising a new wider core design. Also, adding more cache suffers from diminishing returns.

7. Disadvantages

In addition to operating system (OS) support, adjustments to existing software are required to maximize utilization of the computing resources provided by multi-core processors. Also, the ability of multi-core processors to increase application performance depends on the use of multiple threads within applications. The situation is improving: for example the Valve Corporation's Source engine, offers multi-core support, and Crytek has developed similar technologies for Cry Engine 2, which powers their game, Crysis. Emergent Game Technologies' Gamebryo engine includes their Floodgate technology^[3] which simplifies multicore development across game platforms. See Dynamic Acceleration Technology for the Santa Rosa See platform for an example of a technique to improve single-thread performance on dual-core processors.

Integration of a multi-core chip drives production yields down and they are more difficult to manage thermally than lower-density single-chip designs. Intel has partially countered this first problem by creating its quad-core designs by combining two dual-core on a single die with a unified cache, hence any two working dual-core dies can be used, as opposed to producing four cores on a single die and requiring all four to work to produce a quad-core. From an architectural point of view, ultimately, single CPU designs may make better use of the silicon surface area than multiprocessing cores, so a development commitment to this architecture may carry the risk of obsolescence. Finally, raw processing power is not the only constraint on system performance. Two processing cores sharing the same system bus and memory bandwidth limits the real-world performance advantage. If a single core is close to being memory bandwidth limited, going to dual-core might only give 30% to 70% improvement. If memory bandwidth is not a problem, a 90% improvement can be expected. It would be possible for an application that used two CPUs to end up running faster on one dual-core if communication between the CPUs was the limiting factor, which would count as more than 100% improvement.

8. Applications

Multi-core processors are widely used across many application domains including: general-purpose, embedded, network, digital signal processing, and graphics. Video editing, Encoding, 3D gaming, Powerful graphics solution, Optimized productivity titles, such as AVG's virus-scanning software .The full effect and the advantage of having a multi-core processor, when it is used together with a multithreading operating.

CONCLUSION

Power and frequency limitations observed on single core implementations have paved the gateway for multi-core technology and will be the trend in the industry moving forward. However the complete performance throughput can be realized only when the challenges multi-core processors facing today are fully addressed. A lot of technological breakthroughs are expected in this area of technology including a new multi-core programming language, software to port legacy software to “multi-core aware” software programs. Although it has been one of the most challenging technologies to adopt to, there is considerable amount of research going on in the field to utilize multi-core processors more efficiently.

REFERENCES

1. Goodacre, J. and Sloss, A. N., "Parallelism and the ARM instruction set architecture", *Computer*, 2005; 38(7): 42-50.
2. D. T. Marr, F. Binns, D. L. Hill, G. Hinton, D. A. Koufaty, J. A. Miller and M. Upton, "Hyper-Threading Technology Architecture and Microarchitecture," *Intel Technology Journal*, 2002; 6(1): 02/14.
3. Aspray, W. (1997), "The Intel 4004 microprocessor: what constituted invention?", *Annals of the History of Computing*, IEEE, 1997; 19(3): 4-15.
4. Cass, S. Multicore Processors Create Software Headaches. *Technology Review*, 2010; 113(3): 74-75.
5. Parkhurst, J., Darringer, J. & Grundmann, B., "From Single Core to Multi-Core: Preparing for a new exponential", *Computer-Aided Design*, 2006. ICCAD '06. IEEE/ACM International Conference on, 2006; 67.
6. R. Ramanathan., “Intel multi-core processors: Making the move to quad-core and beyond. Technology”@Intel Magazine., 2006.
7. Lizhe Wang, Jie Tao, Gregor von Laszewski, Holger Marten., “Multicores in Cloud Computing: Research Challenges for Applications”, *Journal of Computers*, 2010; 5(6).
8. Schaller, R. R., "Moore's law: past, present and future", *Spectrum*, IEEE, 1997; 34(6): 52-59.
9. Sodan, A., Machina, J., Deshmeh, A., Macnaughton, K. & Esbaugh, B., "Parallelism via Multithreaded and Multicore CPUs", *Computer*, 2010; 99: 1-1.