Discovering Multi-Core: Extending the Benefits of Moore's Law

Geoff Koch
Science Writer
Michigan State University

Table of Contents

(Click on page number to jump to sections)

Discovering Multi-Core: Extending the Benefits of Moore's Law

Geoff Koch
Science Writer
Michigan State University

Overview: Scaling Out, Not Up

In 1965, Intel co-founder Gordon Moore observed an exponential growth in the number of transistors per integrated circuit and predicted that this trend would continue—a prediction today known as Moore's Law. Through Intel's ongoing research and development efforts, the doubling of transistors every couple of years has been maintained for almost 40 years.

For most of that time, Moore's Law made it relatively easy to improve performance. Most average users simply scaled up, relying on applications that took advantage of increasing capabilities and speed of single processors. Users with the most extreme demands scaled out, turning to applications that divided processing tasks between two or more processors. However, scaling up has been harder to do because of several less-friendly laws of physics.

Working Within the Laws of Physics

First, memory speeds are not increasing as quickly as logic speeds. During the days of the i486™ CPU in the late 1980s and early 1990s, the requirements were 6 to 8 clocks per cycle to access memory. Today's Intel® Pentium® processors require 224 clocks, about a 20x increase. These wasted clock cycles can nullify the benefits of frequency increases in the processor.

Next, smaller, denser transistors on today's chips need to be threaded together with ever-increasing lengths of wire interconnects. As these interconnects stretch from hundreds to thousands of meters in length on a single processor, path delays crop up that can cancel the speed increases of the transistors.

Finally, and perhaps most important, is an unsustainable increase in power density. The number of transistors per chip has ballooned in recent years, an engineering challenge since each of these transistors is a working electrical device that consumes power and produces heat.

A 1993-era Intel Pentium processor had around 3 million transistors while today's Intel® Itanium® 2 processor has nearly 1 billion transistors. If this rate continued, Intel processors would soon be producing more heat per square centimeter than the surface of the sun—which is why the problem of heat is already setting hard limits to frequency increases.

However, immutable laws of physics don't necessarily lead to hard limits for computer users. New chip architectures built for scaling out instead of scaling up will offer enhanced performance, reduced power consumption and more efficient simultaneous processing of multiple tasks.

Intel dual/multi-core chips embrace the scale-out approach to performance. This architecture in essence reflects a "divide and conquer" strategy. By splitting the computational work performed by a single Pentium core in traditional microprocessors and among multiple Pentium execution cores, a multi-core processor can perform more work within a given clock cycle.

The result is a better overall user experience and an extension of Moore's Law well into the future.

New Computer Usage Models

In the 1970s an Intel engineer came to Intel co-founder Gordon Moore with an idea for a computer that would be used in the home. Moore asked what it would be good for and all the engineer could think of was that such a device might be used by housewives to store recipes. Moore, who couldn't imagine his wife with her recipes on a computer in the kitchen, declared the idea not to have any practical application at all.

Today, there are more than half a billion PC users worldwide doing lots more than storing recipes. Home users are relying on PCs both for delivery and creation of rich media content, placing new demands on applications for encoding and decoding multimedia files and for editing video. Business users of software for 3D modeling, scientific calculations or high-end digital content creation have their own growing list of performance-heavy requirements, as well.

Software Development Changes/Challenges

Any application that will run on a single-core Intel processor will run on an Intel dual-core processor. However, in order for an application to take advantage of the dual-core capabilities, the application should be optimized for multithreading. Microsoft Windows Movie Maker* 2.0 and Adobe Photoshop* CS are examples of widely used applications that will show significant performance improvement on multi-core systems. Users who multitask or work in environments with lots of background processing also should benefit from multi-core systems.

For both home and business users, performance means more than wringing additional benefits from a single application. Users increasingly multitask, actively toggling between two or more applications or working in an environment in which lots of background processes compete with each other and with open applications for scarce processor resources.

No one expects a leap forward in processors' core execution engines, which today are already at the edge of the manufacturing envelope. In a few years, these processor cores may have larger cache sizes and at least slightly faster clock speeds than those on today's chips. Citing physics- and engineering-related challenges, these advances will be hard won and won't keep pace with performance demands. Inevitably, meeting these demands will require a scale-out approach that divvies up computing tasks among multiple processing cores.

One way to do this divvying up is to build single computers powered by multiple processors. For years this has been the norm in high performance computing and supercomputing, industry segments in which single machines often contain hundreds or even thousands of processors.

But adding even one additional processor isn't always practical for mainstream PC users. Because of the additional socket and associated chips and circuitry, motherboards for dual-processor systems are more expensive than their single-processor counterparts. Also, since the processor is often the most expensive single component in a given computer system, trying to achieve scale-out performance by swapping a single-processor machine for a dual- or multi-processor machine can be expensive.

A New Era in Processor Architecture

Another scale-out option, and the one Intel has embraced in its move to multi-core architectures, is to add two or more "brains" to each processor. Explained most simply, multi-core processor architecture entails silicon design engineers placing two or more Pentium "execution cores," or computational engines, within a single processor. This multi-core processor plugs directly into a single processor socket, but the operating system perceives each of its execution cores as a discrete processor, with all the associated execution resources.

Multi-core chips do more work per clock cycle, and thus can be designed to operate at lower frequencies, than their single-core counterparts. Since power consumption goes up proportionally with frequency, multi-core architecture gives engineers the means to address the problem of runaway power and cooling requirements.

Applications that work on Intel single-core processors will work with Intel dual-core processors, but to make the most of a dual-core processor today, the software running on the platform must be written such that it can spread its workload across multiple execution cores. This functionality is called thread-level parallelism or "threading," and applications and operating systems (such as Microsoft Windows* XP, Windows Server and various Linux* vendor offerings) that are written to support it are referred to as "threaded" or "multithreaded."

A processor equipped with thread-level parallelism can execute completely separate threads of code. This can mean one thread running from an application and a second thread running from an operating system, or parallel threads running from within a single application. Multimedia applications are especially conducive to thread-level parallelism because many of their operations can run in parallel.

Operating System and Core Software Licensing Issues

As Intel's dual-core capability is an evolution of Intel's HT Technology, we fully expect that licensing for dual-core will work in a similar manner. So far, Microsoft, Sun, IBM, and various Linux vendors have pledged to support each dual-core chip with a single license for their respective software products. Users should check with their operating system, database or other core software vendor for additional details and confirmation.

Early in the 2000 decade, Intel introduced Hyper-Threading Technology† (HT Technology) into its Intel® NetBurst™ microarchitecture (for Intel Pentium 4 and Intel® Xeon™ processors) as an innovative means to deliver higher thread-level parallelism on volume platforms. HT Technology enables processors to execute tasks in parallel by weaving together multiple threads in a single-core processor.

As of September 2004, HT Technology has shipped on well over 50 million Intel Pentium 4 products for desktops, servers and mobile PCs, offering new incentive for software developers to design applications capable of processing information in parallel for greater efficiency. To date, more than 150 client applications have been enabled for HT Technology, on top of the many hundreds of applications that have previously been threaded. These numbers demonstrate that many developers and design tools are already in place to capitalize on Intel multi-core capability.

Thanks to $28 billion in R&D spending through the tech downturn from 2001 to 2003, Intel has the manufacturing capacity in place to quickly increase its production of multi-core chips. Intel forecasts that more than 85 percent of its server processors and more than 70 percent of its mobile and desktop Pentium family processor shipments will be dual-core-based by the end of 2006.

Further on the horizon, Intel plans to deliver additional processors with two or more cores for mobile, desktop and server platforms. At the Spring 2005 Intel Developer Forum in San Francisco, Intel senior fellow and Corporate Technology Group director Justin Rattner spoke of Intel's goals to mass-produce chips with a hundred or more processing cores. Discover how we envision Platform 2015.

Summary

Intel dual- and multi-core chips are ushering in a new era of computer architecture, embracing a scale-out approach to enhance performance and enable more-efficient simultaneous processing of multiple tasks. Multi-core chips do more work per clock cycle, and thus can be designed to operate at lower frequencies than their single-core counterparts. Since power consumption goes up proportionally with frequency, multi-core architecture gives engineers the means to address the problem of runaway power and cooling requirements. All this makes for significantly improved user experiences in both home and business environments, enabling us to extend Moore's Law well into the future.

More Info
Read the full paper on Discovering Multi-Core [PDF].

You can also learn much more at the Intel Web site:

- Intel Multi-Core Platforms
- Intel Multi-Core Processing Software Technologies
- Multi-Core Server Processors
- Multi-Core Server Resource Center
- Low Power with Multi-Core Architectures
- Moore's Law

References
"Microprocessors Circa 2000" is the October 1989 article from *IEEE Spectrum* in which Intel technologists first publicly predicted that multi-core processors would come to market soon into the next millennium: http://www.intel.com/technology/silicon/ieee/circa2000.pdf [543 KB].

"Architecting the Era of Tera" is a white paper from Intel Research and Development that defines the needs for computer architectures such as dual/multi-core that handle massive data sets: ftp://download.intel.com/technology/comms/nextnet/download/Tera_Era.pdf [298 KB].

Read about Intel's fast-track plans to deliver multi-core processors in high volume across all its platform families, including client, server and communications systems at: http://www.intel.com/business/bss/products/server/multi-core.htm.

Author Bio
Geoff Koch is a science writer for Michigan State University and a freelance science and technology reporter in Lansing, Michigan. From 1998 to 2003, Koch worked at Intel in a variety of writing, editing and communication roles, including as an editor for the forerunner of the Intel Software Network. Koch earned an M.A. in communication at Stanford in 2004. He is a contributing writer for *SD Times*. Additionally, he has written for *The Dallas Morning News*, *The Portland Business Journal*, the *St. Louis Post-Dispatch*, *The Oregonian* and *Mobilized Software*.

†Hyper-Threading Technology requires a computer system with an Intel® Pentium® 4 processor supporting Hyper-Threading Technology and an HT Technology enabled chipset, BIOS and operating system. Performance will vary depending on the specific hardware and software you use. See http://www.intel.com/info/hyperthreading/ for more information including details on which processors support HT Technology.

—End of Technology@Intel Magazine Article—