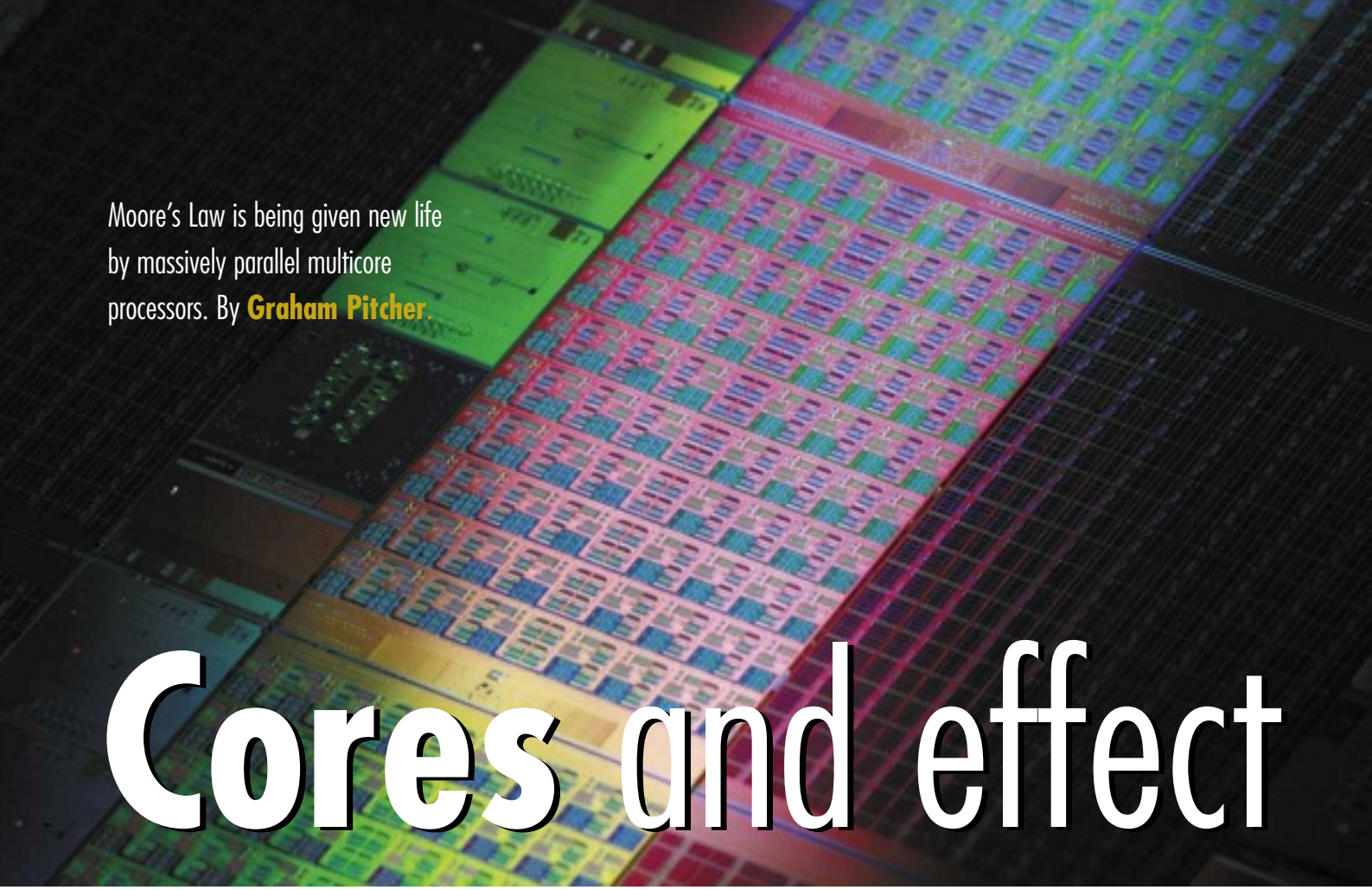


Moore's Law is being given new life by massively parallel multicore processors. By **Graham Pitcher**.

Cores and effect

The future for microprocessors is, without doubt, multicore. The fundamental physical challenges which chip designers face – with power dissipation close to the top of the list – mandate that the only way to get more computing power from a piece of silicon is to increase the number of processor cores.

Recently, Intel has taken a lead in this area, with the launch of devices featuring four processor cores – its so called Quad-Core family. But a project underway in its research operation is taking that a bit further; it has demonstrated what is essentially a supercomputer on a chip. The device integrates 80 processor cores and is said to bring teraflop performance from a small piece of silicon and to consumer just 62W in the process. In fact, the company claims the device just breaks the teraflop barrier at 1.01Tflops when running at 3.16GHz.

"Our researchers have achieved a wonderful and key milestone in terms of being able to drive multicore and parallel computing performance forward," said Justin Rattner, Intel's senior fellow and chief technology officer. "It points the way to the near future when Teraflops capable designs will be commonplace and reshape what we can all expect from our computers and the internet at home and in the office."

Each of the simple cores in the research chip contain two programmable floating point engines and Intel says this is the largest number to have be integrated on one chip. There are two reasons for choosing floating point engines for the demonstration. Firstly, the engines are used for accurate calculations, including graphics and scientific modelling. But there is also a design challenge: these cores are more complex than integer engines.

In 1996, ASCI Red was the first computer to benchmark at a teraflops. The system used around 10,000 Pentium Pro processors running at 200MHz and consumed 500kW of power plus an additional 500kW just to cool the room that housed it. Now, Intel has integrated similar computing power on one piece of silicon drawing just 62W.

The chip consists of 80 tiles laid out in an 8 x 10 block array. Each tile includes a small core, or compute element, with a simple instruction set for processing floating point data. However, the device is not compatible with the Intel Architecture as yet. The tile also includes a router connecting the core to an on chip network that links all the cores to each other and gives them access to memory.

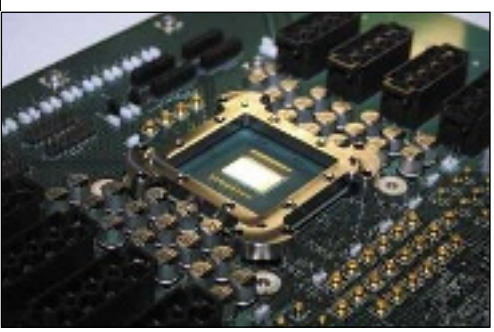
Having said that, the 80 core device is very much experimental; Intel itself has no plans to release a commercial implementation of this chip for some years – or won't admit to them if it has.

Rattner observed: "These experimental chips address the three major requirements for terascale computing – teraops

of performance, terabytes per second of memory bandwidth and terabits per second of I/O capacity," said Rattner. "Whilst any commercial application of these technologies is years away, it is an exciting first step in bringing terascale performance to pcs and servers."

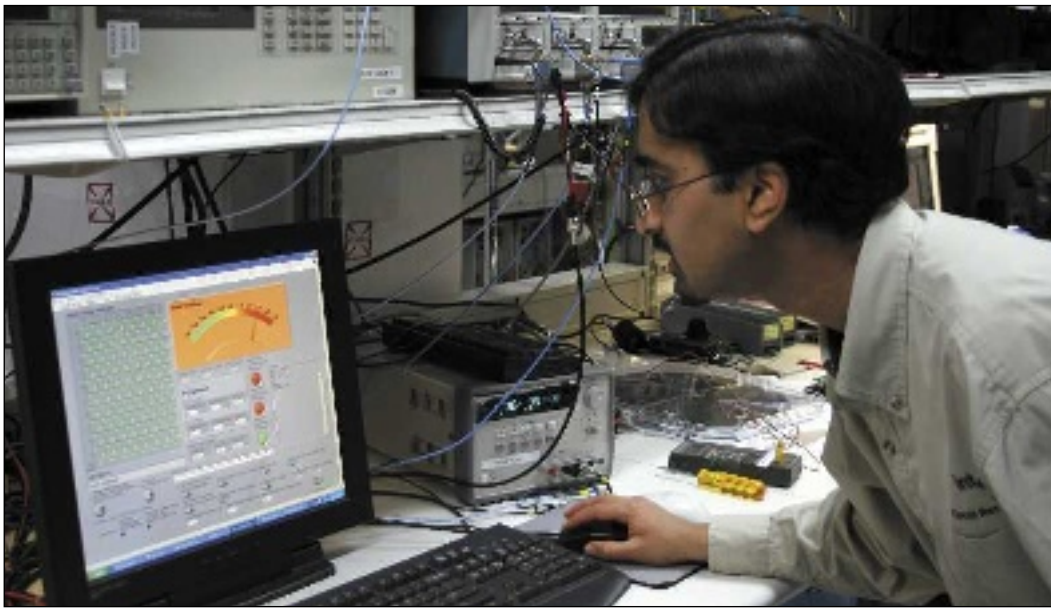
Multicore processing brings a number of challenges; not only to chip designers, but also to those who need to program them. Part of Intel's multicore work is aimed at trying to solve these problems. In all, Intel's Terascale Computing Research Program has more than 100 projects underway, exploring a range of architectural, software and system design challenges.

One of the attractions of a multicore approach is a reduction in power consumption. According to Intel, a larger number of simple cores can be built within the same area as a smaller number of large complex cores. In addition, power consumption can be optimised by using multiple types of cores tuned to match the needs of different usage models. Power consumption is further improved by powering down or turning off those cores that are not busy.



By varying the mix of functional elements, the company believes, the architecture can allow designers to create multiple implementations that match specific market needs. If the processor uses regular tiles that consist of one or more cores, a cache, a router and other supporting hardware, different aspects of a layout could be reused from processor to processor.

Alongside the computing elements, each processor core features a five port message passing router and these elements can be connected in a 2d mesh. Intended to provide more scalability than the interconnects found in current multicore devices, the mesh will allow for speedier communications between processor cores and a boost in overall performance.



But multicore devices – even the quad core parts which Intel has released of late – are of no use unless they can be programmed to operate in the most efficient manner.

Intel admits that, without optimised software, terascale platforms will not work as well as they could. Platforms, it notes, must address the needs of new and existing programming models.

Here, Intel is looking at ways in which legacy software can be accommodated. A substantial amount of existing software has not only been written for single core processors, it has also been written for single threaded execution. Intel says its researchers are now exploring ways to incorporate heterogeneous general purpose cores within its terascale architectures.

In a terascale white paper, Intel notes that one of the features of terascale architectures will be dedicated partitions that appear as devices to regular software. These partitions will provide functions such as system management and I/O acceleration (such as network protocol processing). The architectures may also include hardware support for lightweight message passing for a distributed computation model. This type of streaming, Intel continues, has proven to be an effective programming model for workloads such as graphics and media processing.

Software is a major part of the multicore program, but you still have to design the chip in the first place. One of the approaches used by Intel is tiling. According to the company, this design philosophy allows smaller cores to be created, which can then be repeated across the chip.

The 80 core demonstration chip comprises around 100million transistors. Intel says that designing a similar sized single core processor would take twice as long and require twice as many people.

For the future, Intel will be looking at such areas as adding 3d stacked memory to the chip and an indication of the research's importance is given by the fact that the Terascale Computing Research Program now has more than 100 projects underway.

As chip manufacturing processes continue to shrink, the Laws of Physics are making it harder to deliver the progress called for by Moore's Law. Engineers being the creative people they are have continued to find solutions – Intel's recent announcement of new materials for its 45nm process is an example.

These developments, alongside massively parallel multicore processing, are said by the company to lay out a way forward to make chips with billions of transistors in the future. Moore's Law appears to have a few years left in it.