CHIP MAKERS HIT HEAT BARRIER

Power consumption is the hot issue in this winter's conferences. Luke Collins puts some possible solutions under the microscope

RISING POWER consumption in semiconductor devices is threatening the improvement in microelectronics that has underpinned growth in the IT sector for the past forty years.

That's the message to emerge from a pair of technical conferences being held this winter in San Francisco, California. If the issues cannot be resolved quickly then the accepted rule that electronics will get faster and cheaper year after year, and hence more pervasive in our daily lives, will have to be rewritten.

Andrew Grove, chairman of Intel, the world's largest semiconductor company, said at the International Electron Devices Meeting in December that power consumption is "becoming a limiter of integration," in other words, that "faster, cheaper, smaller" was under threat.

THE PROBLEM NOW

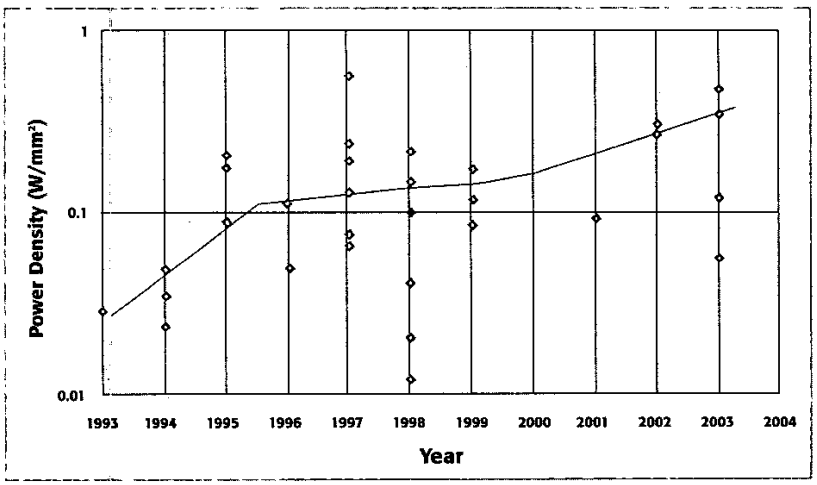
Power consumption will also be the focus at February's International Solid State Circuits Conference, whose organisers have chosen "Power aware systems" as the theme for the five-day conference. Clearly, something is afoot.

But does it matter? There are indications that chip power consumption does matter in the real world, right now. The third-generation mobile phone network operators will only be able to pay off the massive debts they incurred for licences if they can deliver compelling new services; no one needs a 3G phone to make a phone call. But those new services, such as video clips of your favourite team's Saturday afternoon goals, will only be possible if highly integrated handsets with long battery lives are available. First-generation 3G handsets have fallen well short of this goal.

Chip power consumption is already an issue in today's technology. Laptop computers are notorious battery hogs. But a recent letter to medical journal The Lancet revealed that a patient had been badly burned by working with his computer in his lap for an hour. Power consumption is already close to limiting innovation in areas such as the 3D graphics chips beloved of computer gamers. The latest offering from nVidia has 125 million transistors, three times as many as the Pentium 4. But cooling the chip demands a ducted fan, copper heatspreaders on top and bottom and heatpipes connected to a radiator. While one set of engineers struggle with sub-micron process issues another is working on plumbing.

Graphics chips are fun but not vital. But processors for the web servers, databases engines and general computing functions that now rule our lives are important. Intel's third-generation server processor, the Itanium, will pack 410 million transistors onto one chip. It will dissipate 130W, or more energy than a household light-bulb, from an area of 310mm2 – about the size of a small postage stamp. Patrick Gelsinger, now chief technology officer of Intel, even used a keynote at 2001's ISSCC to compare the energy density of chips over the next ten years to hot plates, rocket nozzles, nuclear reactors and finally the surface of the sun.

So can the power consumption problem be solved? The issue is one of increasing transistor density and decreasing transistor efficiency. As the minimum dimensions of transistors have fallen below 1μm it has become more difficult to turn them completely off as the insulating layer between the gate which controls current flow and the channel through which it passes has thinned. At minimum dimensions



of 100μm or less, the Off state leakage current can be 10% of the On state current flow. When your chip has hundreds of millions of transistors, that can be significant.

WAYS AND MEANS

One way to control current leakage is to improve the insulating properties of the dielectric layers that separate the gate from the channel. To date these have been made out of polysilicon. But in devices when the insulating layers are just atoms thick, silicon's insulating properties fall short. Alternatives using oxides of hafnium and lanthanium will be presented at IEDM, although the industry is bound to be cautious in moving away from polysilicon, a central pillar of its manufacturing strategy.

Another way to reduce leakage is to move the transistor up out of the plane of the chip surface into a vertical arrangement. These FinFETs (field effect transistors) build the conducting channel of the transistor in a fin of material above the chip's surface. Using gate electrodes on either side of the channel can help turn the transistor off more effectively, reducing leakage current. Although this is a radical change in chip design, research into FinFETs has moved from university to commercial research rapidly in the past five years.

These approaches to reducing leakage current may solve the problem, but still need to prove they are compatible with other parts of the process, manufacturable and reliable. Improving other chip characteristics, such as speed, may also have a detrimental effect on leakage. At IEDM, for example, Intel will describe a process built on 'strained silicon', that is silicon alloyed with germanium to create a strained crystal lattice through which electrons can travel more quickly than in pure silicon. Although strained silicon offers a performance improvement over ordinary silicon, defects in the lattice can promote leakage. So Intel has had to take care to ensure there are no defects in the transistor's conducting channel.

Chip manufacturing processes demand trade-offs all the time. But if power consumption cannot be tamed at the process level, engineers will have to make trade-offs at other levels, making better use of the transistors that a limited power budget allows. Engineers will have to think about architectural issues, such as whether they run all of a chip all of the time, and about algorithmic issues such as better ways of doing certain tasks. Finally engineers will have to investigate systems issues, for example splitting work between battery-powered clients and mains-powered servers.

THE SEMICONDUCTOR INDUSTRY'S SUCCESS HAS SOWN THE SEEDS OF ITS DOWNFALL

HIGH EXPECTATIONS

The semiconductor industry's success has sown the seeds of its downfall. The bi-annual doubling of chip density has created the expectation that electronics will go on getting faster, cheaper and smaller forever. When that growth slows or comes to an end, the knock-on effect on other industries could be immense.

But there is hope. IBM Fellow Russell Lange pointed out a number of years ago that the chip industry is essentially driven by exponentials – zdensity, price/performance and so on. When a major barrier to the continued exponential growth of the industry emerges, he argued, an exponentially increasing amount of engineering effort tends to be applied to solve it. ■

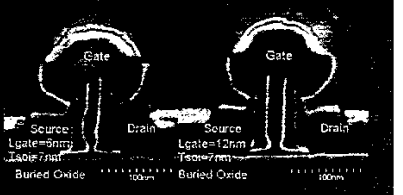
MOLECULAR'S BURNING ISSUE

IBM has developed a working transistor whose critical dimension is fifteen times smaller than the leading-edge commercial processes available today.

The transistor has a gate length of 6nm, whereas today's leading processes have 90nm gate lengths. Gate length is the critical dimension of transistors since it defines the distance over which electrons have to travel and hence the maximum switching speed.

Devices with these dimensions need to be available in 2016 in order to keep semiconductor technology on the development path defined by the International Technology roadmap created by a consortium of chip companies in 2001.

Dr Randy Isaac, vice president of science and technology, IBM Research, said: "This achievement underscores the fundamental challenges of scaling, namely power density, that must be addressed as silicon is pushed to molecular dimensions."



IBM has been able to address the problem by using silicon layers just 4 to 8nm thick, on top of a buried oxide layer, to form the device's channel. Both the drain and source were formed above the substrate and the gate's length was controlled by making a very narrow aperture through a dielectric layer. But even with a gate this short the researchers were able to turn the device off properly, a critical factor in controlling power consumption in future chip processes.