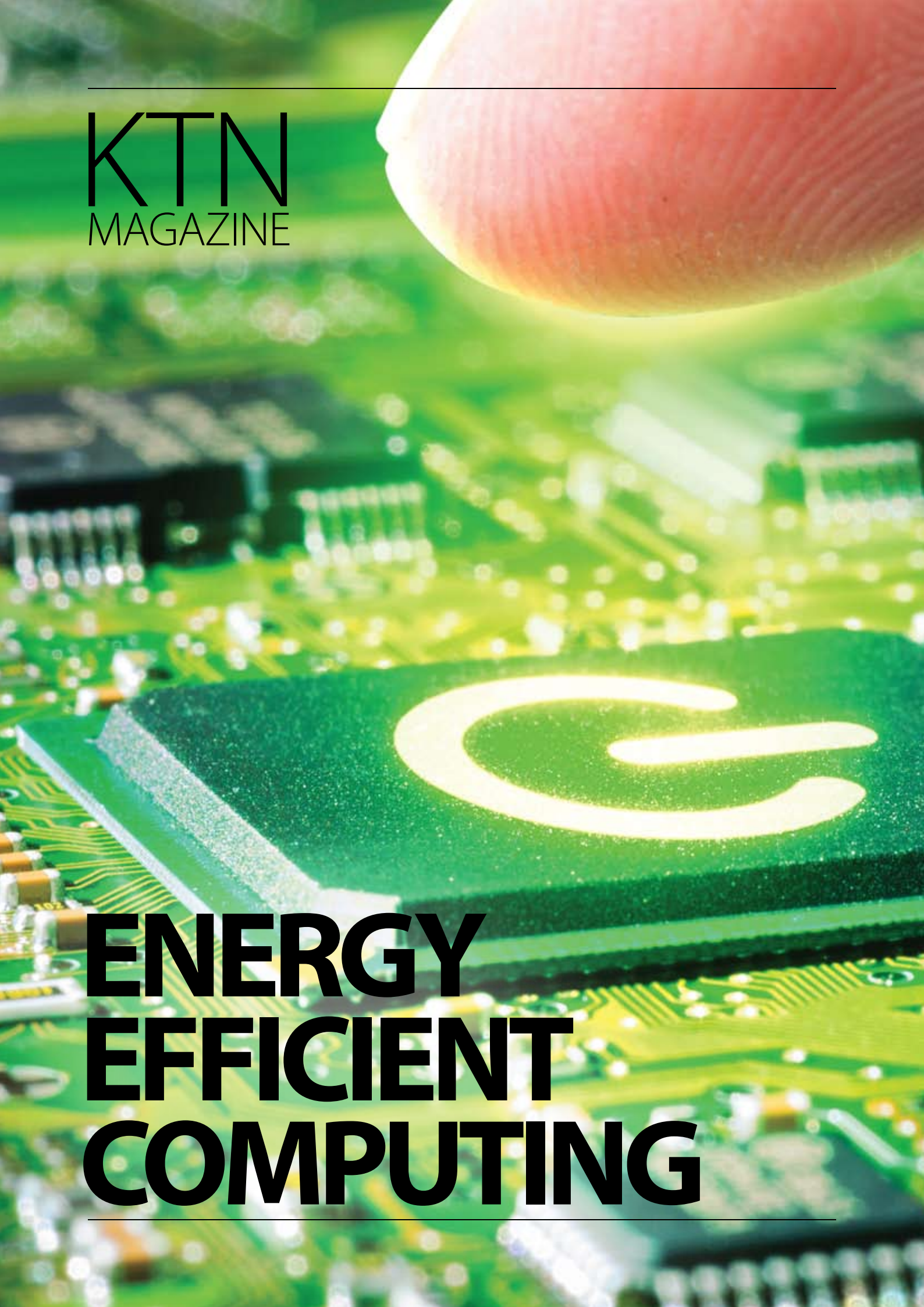
KTN
MAGAZINE



# ENERGY EFFICIENT COMPUTING

# Contents


| | |
|---|---|
| **Foreword** | **3** |
| **Introduction** | **4** |
| **Mobile Devices** | **6** |
| **Embedded systems** | **15** |
| **Multicore Computing** | **16** |
| **Data Centres** | **17** |
| **HPC** | **21** |
| **Measuring** | **25** |
| **New Energy Efficient Materials** | **30** |
| **Conclusion** | **32** |


## Contributors


**MAGAZINE EDITOR AND LEAD AUTHOR:** Adriana Hamacher

**WRITING AND RESEARCH:** Ben Lewis

**DESIGN:** Helen Boosey


## Contact


**EMAIL:** enquiries@ktn-uk.org

**PHONE:** 01403 251354

**ADDRESS:** Bailey House,
4-10 Barttelot Road, Horsham,
West Sussex, RH12 1DQ, UK

# Foreword

**When we think about electricity we think of the cost of domestic fuel bills - but energy is not just about the cost of a fuel bill - it is about keeping the world going.**

Mobile phones mean that now we are never out of contact. Texts, emails and even phone calls keep us in touch with the world. Modern phones also contain GPS and sensors that tell us about our environment; satellite and 3D images of cities help us to know where we are and where we are going.

The cost of these increasingly essential tools and information is battery life! Remember when your phone lasted over a week on a single charge? Manufacturers and designers now struggle to keep a phone running for just a working day. Without invisible advances in energy efficiency our phones wouldn't even last an hour.

Today much is said about the ability to connect everything to the internet (the Internet of Things or IoT). This vision promises to supply us with data on the world about us - enabling increased efficiency, better use of resources and allowing us to lead easier lives. All these smart devices need energy to capture information and transmit it to the internet - but who is going to change the batteries?

Many sensor nodes - the eyes, ears and noses of the IoT - will have to run for 5-10 years without a battery change (the cost of replacing a battery can exceed the original cost of installation). Energy Harvesting technologies are working to generate power from the natural environment (wind, light, heat, vibration etc.) to recharge batteries, but the amount of energy you can generate is still very small so reducing energy consumption is the critical factor.

At the other end of the scale, the supercomputers used to process data and run our national services, billing systems and logistics are significant energy users. Again, it's not just the cost of the electricity, it's about being able to get sufficient power into a building, and removing sufficient heat, in order to run the computers. The development of more and more powerful computers is demonstrated by the fact that the first teraflop computer was achieved in 1997, the first petaflop computer in 2008 and it's projected that an exaflop computer could be possible by 2018. BUT a teraflop computer built on today's design rules would require 200 megawatts (MW) of power - the belief is that this needs to be reduced to 20MW for a practical computing platform. Hence energy efficiency is critical to achieve the desired performance.

Today's most exciting developments are at either end of the performance range - IoT and wearable computers, such as fitbit, at the low power end, and exaflop computers (to process the data) at the top end.

In both cases power reduction, in the form of **ENERGY EFFICIENT COMPUTING**, is vital to achieve the required performance.

The rest of this guide explores many examples of **ENERGY EFFICIENT COMPUTING** in more detail and we hope that it provides useful, and interesting, information in this critical sector of technical developments.


**Nigel Rix**

**EEC SIG (bit.ly/EECportal )**

**THE KNOWLEDGE TRANSFER NETWORK (ktn-uk.org)**

**Surging demand from consumers and businesses, soaring energy costs, advancing supercomputer power and nanotechnology have opened up a new frontline in the global market in computing technology.**

Energy efficient computing, or EEC for short, may until recently have played only a bit part in the industry's history but the issues it deals with are becoming too important to ignore any longer.

It's as if, for the computing industry, the oil shock of the seventies, a population explosion and the age of environmentalism have all arrived at once.

The statistics for the increase in use of energy for computing are formidable. According to one study (US analysts the Digital Power Group), the global digital economy consumes as much electricity today as it took to illuminate the entire planet in 1985 and uses 10% of the world's electricity.

"If you look at how computing is evolving," says David May, founder of chip designer XMOS, and a computer scientist at Bristol University, "at one end everything is moving into mobile and wireless and the IoT {Internet of Things) and increasingly it will be wearable. And at the other end it's all heavy-duty storage and computing moving into the cloud. Sales for desktops and laptops are declining, in the future they will just be for professional use. Batteries don't last long enough, and the Cloud's data centres cost a fortune to run. So for both the cloud and mobile devices, energy efficiency is the big challenge."

The race is now on to bring the digital world's staggering energy consumption figures down. Many different factors motivate this push. It's a question of cost but also of basic technological progress. Energy inefficiency is now the single biggest obstacle to the advance of computing. It's not just that the electricity bill is the biggest expense in running a data centre, but also that the next generation of supercomputers will need to be much more energy efficient, or they will overheat. Meanwhile new markets in wearables and the IoT mingle with consumer frustration at the shortness of battery life on their mobile devices.

Government intervention is becoming a factor too. The European Union has made EEC a priority for grants and interdisciplinary project support, with supercomputing highlighted in its Ecodesign Directive, because of the "significant" energy savings potential. On the other side of the pond, the US Department of Energy wants to establish minimum energy efficiency standards for all computers and servers sold there, just as they have already done for air conditioners, TV sets and dishwashers.

**Energy inefficiency is now the single biggest obstacle to the advance of computing.**

In Britain we have a reputation for producing low energy computer products. ARM Holdings - a world leader in microprocessor Intellectual Property - is headquartered in Cambridge; much world class research is taking place, from operating system code-writers inventing new algorithms with low energy sleep states to apps developers who reduce the energy consumption for GPS and engineers who are designing new cooling systems for data centres. The British government has also recognised the importance of the issue and pledged £189m to Big Data and energy efficient computing projects. With government support and sector shaping businesses, the UK is well placed to become a market-leader in this field.

MOBILE DEVICES
PORTABILITY COMES AT A COST

It's a phrase that's regularly bandied around these days and, for sure, the market in mobile devices has exploded in the past couple of years. But there's not just one market. There are, in fact, three: an established market, an emerging one and another at embryonic stage. This activity has placed energy efficiency at the forefront of technological developments today.

The established market is for mobile devices like phones and tablets and is still growing exponentially. The emerging market is for the IoT, all those machines and gadgets, from cars to toasters to ingestible health sensors, that access the internet, feeding data back to their manufacturers and obtaining new firmware and software updates. The embryonic market is in wearables, of which the advance column is Google Glass.

Each of these markets presents its own challenge. In the arena of mobiles, tablets and laptops, customers have high demands that make battery life one of the key selling points. Consumers are using their devices for longer and longer each day; they like applications that consume larger and larger amounts of power, like games, videos, Instant Messaging (IM) apps and mobile location-sensing apps (MLSAs). Meanwhile the devices themselves are increasingly fitted with energy-consuming sensors. Software designers want devices to do more and more by themselves without prompting, which means they need to be responsive all the time. And, of course, we all want our devices to work faster, and be ever smaller and thinner. One survey found that only 8% of mobile users were satisfied with the battery performance of their smartphones.

The IoT presents a fresh range of problems. Many sensors that supply the information that makes up the Big Data revolution require batteries that can last years before they have to be changed or charged, not just a few hours. Examples include sensors such as those that monitor traffic and are embedded in tarmac, or the swallowable ones appearing in diagnostic healthcare. Other IoT devices - our set-top boxes, DVD players and (coming soon) smart thermostats - rely on electricity. This suggests that if the IoT continues to grow at its current speed, without massive advances in energy efficiency, then the world's power stations simply won't be able to supply enough fuel to cope.

And then there are the wearables - the next 'revolution' in computer devices. Here we will need power sources that are not only long lasting but also ultra-thin and often bendable. Early reviews of Google Glass are already highlighting power issues, with people finding they can only use the device for five hours between charges.

But these issues aren't insurmountable and developers, innovators and entrepreneurs from different disciplines are working on them from a variety of perspectives.

## WHEN THE CHIPS ARE DOWN...

The British company ARM has long been a leader in the market for low energy processors and chip architecture and has been designing ever more efficient versions since the mid-nineties, when it began supplying them to Apple and Nokia. Now it's pursuing a strategy of designing different processors for different tasks, each one using the lowest possible energy for its purpose. "It's not just a question of low energy, it's low energy for what?" says John Goodacre, ARM's head of technology. ARM's IoT arsenal includes processors that are de-signed for roadside sensors. "Our lowest energy processor, the Cortex M0, uses microwatts (uW) and can be operated through power-scavenging," says Goodacre. "These could be powered by the vibrations from passing traffic. They use such lower power that a key cell battery could keep one going for 15 years."

ARM's designs are not just about extending battery life, but also reducing electricity bills on plug-in IoT devices. ARM's processors are targeted at specific power usages: rechargeable mobile devices, which require three watts (W) of power; low-power tethered consumer devices (such as a HDTV or set top box), which are capped at 10-15W, and purely tethered devices, which are limited by cooling to around 100W. Goodacre predicts that soon there will be further pressure to reduce energy, as the EU and the US continue to impose regulations on electronic devices.

Most existing mobile device chips maintain a high standard output for receiving and transmitting data,

yet the strength of signal required for those processes varies considerably depending on what one is doing with the device. This is the focus of studies by US-based ETA Devices and Qualcomm. The research is aiming to raise or lower the radio signal according to how much power is required, something known as 'envelope tracking'. Qualcomm has already installed such chips in Nexus and Samsung phones. ETA Devices' Mattias Astrom says, "Envelope tracking is basically a continuous variable transmission, comparable to a manual gearbox." he says. "Fuel consumption is always better when you have a manual gearbox."

Screens are another big energy drain and can consume up to half the battery power of a mobile device, so they have become another area of interest for developers. Companies like Israel's Genoa Colour Technologies (GCT) and Unpixel, from the US, have made considerable advances in the last five years, from the US. GCT previously developed technology to reduce power consumption in TVs and have now adapted this technology for mobile devices to reduce screen power consumption by 40%.

But it's a British company, Plastic Logic, based in Cambridge, which is at the forefront of current advances here - they have designed a 'bistable' display which retains its image without using power.

The company specialises in robust, flexible displays, for everything from smartphone accessories to large-area digital signage. Plastic Logic's secondary display for smartphones functions as a protective cover for the LCD or OLED display on the phone. The idea is that people continue to use their primary, colour, mobile display to watch videos and for the internet, but can easily access a secondary display for saving information, such as plane tickets or to check messages. Because it doesn't require battery power, it can always be on and can be integrated into the plastic casing on the phone.

## TARGETED REDUCTION

Each and every power-consuming aspect of a mobile device is becoming the target of research in energy reduction.

**SLEEP STATES** Developers at Intel Labs are working on something called 'Energy Efficient Sleep State Selection' which takes advantage of lulls in activity to create power-naps for your mobile device, while exploiting long periods of downtime to enter ultra-low power consumption deep sleeps. Their experiments suggest they can reduce power consumption in active states by 50%, without slowing down the performance of the Operating System (OS).

**WI-FI** Scanning for Wi-Fi requires oodles of power. New algorithms, such as those in a system called WIFIsense, developed by researchers from Deutsche Telkom R&D, use lower-powered Wi-Fi scans, which can reduce the energy used in scanning by almost 80%.

**LOCATION SERVICES** One of the highest energy consuming features on a mobile device because it requires constant reference to cell phone towers. The less the device communicates with a locating tower the less power is used, but also the less accurate the positioning. Developers at Deutsche Telekom are trying to find a balance between energy efficiency and accuracy, by programming the phone to look up its nearest tower less often, and fill in the gaps with trajectory mapping and user histories. They call it CAPS - Cell-ID aided positioning system - and claim that it can save on 90% of the energy and provide at least 80% of the accuracy of conventional GPS.

**SENSORS** A key enabling technology, sensors are likely to be increasingly used to record our vital signs, monitor our exercise and report that information in real time to health professionals. But much data, which is streamed continuously, is not very useful. Acquisition Cost-Aware Query Adaptation (ACQUA) is a new approach, developed by researchers in Singapore, which seeks to remedy this. Instead of sending all data from the sensor to the smart phone, the framework cherry-picks only valuable material, making choices or judgements with algorithms that assess the usefulness of the data against the cost of sending it. The developers say their tests have shown that these algorithms can reduce the use of energy in continuous query processing by 70%.

**APPLE** For optimum energy efficiency, hardware and software need to work together. Here Apple has long been considered the market leader - an edge it owes, analysts agree, to its control over hardware and software design.

Apple was the first major company to take an interest in ARM's processors (for the Newton personal assistant device) precisely because of the British company's low power architecture. Among the Apple innovations that any user will be familiar with are the ambient light sensor in their devices that automatically adjusts the brightness of the screen to its surroundings, and the restrained use of power-hungry graphics processors (they only become active when you are using graphics programmes). Its OS even tries to save on energy in the moments between keystrokes.

Apple continues to advance the cause with its latest M7 processor - the first of a new generation of 'ambient chips', which process sensor data away from the main CPU processor, thus saving energy. It is purposed to collect data from sensors and uses less battery power to track a user's movements, even when the phone is in sleep mode. Overall, Apple says it has reduced the average power consumed by its products by 40% since 2008. Nevertheless, even for Apple there is a trade-off between user experience and energy saving: the touchscreen is a relatively power-hungry innovation.

## BATTERY POWER

In America, various start-ups are working on making batteries last longer. At the forefront of these is Amprius, which is already selling super-powerful lithium ion batteries, delivering up to 700W per litre, almost double that of other leading batteries.

Imprint Energy, another stateside firm, is developing screen-printed ultra-thin batteries based on zinc instead of lithium. They can be created in any shape and are perfect for the new market in wearables.

There are also a number of long-term projects that remain far from production stage but look promising.

- Scientists at Northwestern University and the University of Illinois are collaborating on stretchable and even foldable batteries for devices you might wear on a wrist or headband.
- Researchers at Harvard are working on a 3D printed battery that is smaller than a grain of rice yet can power the new technology of ingestible activity trackers.
- IBM is working on a technology for a lightweight, long-lasting, faster-charging battery which they call lithium air.
- Vorbeck Materials in America is researching the potential of combinations of lithium-seawater and lithium-graphene.

Other developers are looking at an entirely different solution. If you can't extend the battery life, why not make charging easier to do? The Wireless Power Consortium is developing a wireless charging system called Qi, based on universal protocols, so anyone can continuously charge up their mobile device in a public place without wiring. At the start of 2014, 334 products were carrying the Qi logo. Some of the newer Toyotas already have Qi capability and Mercedes Benz are putting Qi chargers in their 2014 models.

## NETWORKS

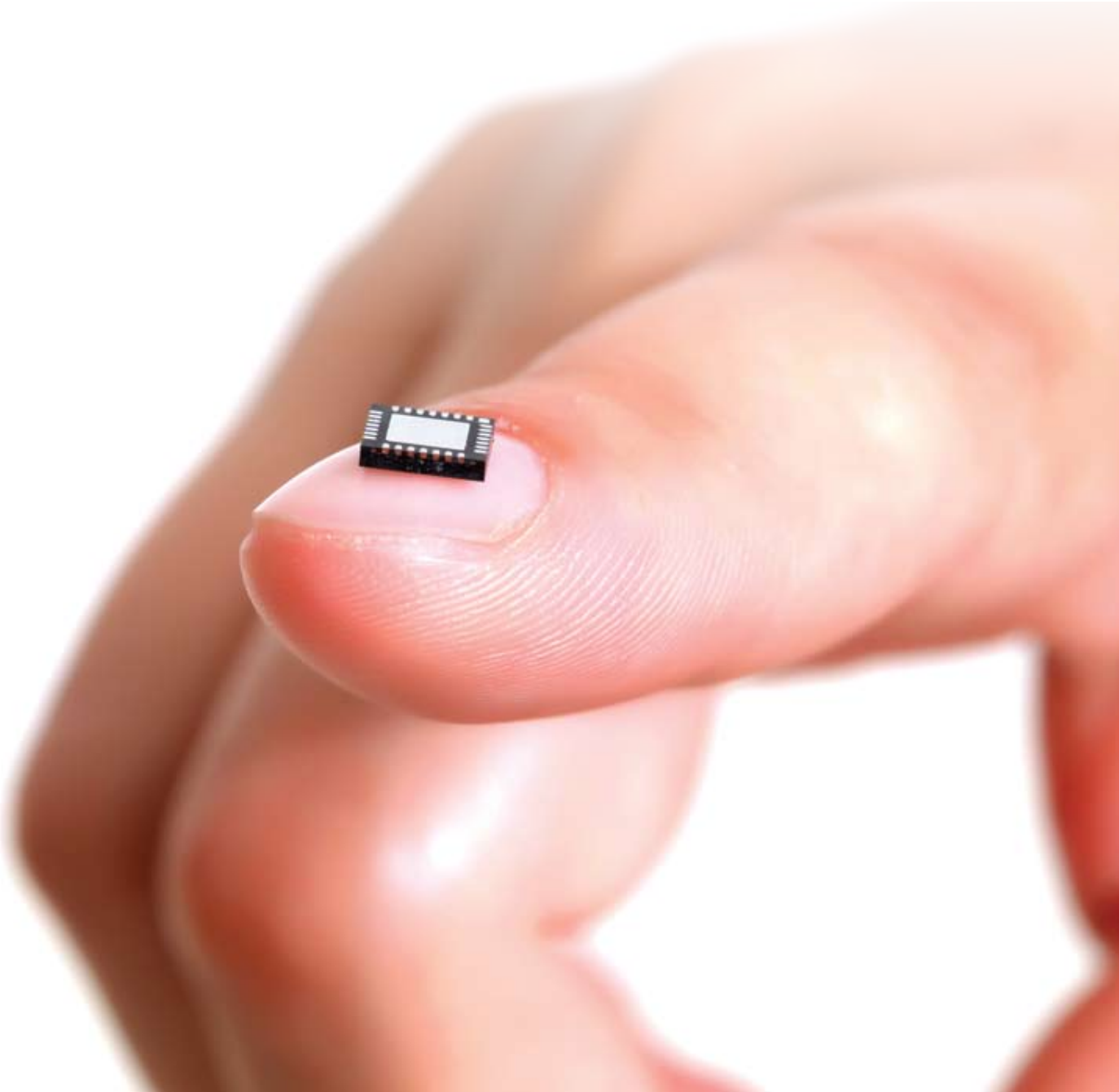
Finally, reducing the power usage of mobile network cell towers and base stations is also a pressing issue. In Africa, limited electricity supplies are a huge barrier to the expansion of mobile phone use. A company from Massachusetts, VANU, has developed the lowest-power cellular base station in history - weighing just five kilograms and consuming only 50W. The gadget provides connectivity for over 1,000 people and is being deployed in Zambia.

Closer to home, British network providers are building frameworks for pooling resources with new small-cell deployments in towns. Sharing infrastructure between different operators and developing a more responsive and flexible allocation of network capacity would further reduce energy use.

The mobile phone business is very optimistic about the progress that can be made. A 2013 study by the GreenTouch consortium (which consists of over 50 telecoms companies and research bodies around the world, including Bell Labs, China Mobile, France Telecom Orange, Fujitsu, Swisscom and Vodafone) asserted that they could cut their own energy consumption by 90% using new technology, and deliver a 1000 times improvement in their energy efficiency, despite the projected increase in phone users and phone usage.

# ARM

## BRITISH LOW-POWER GIANT



With their chips powering over ten billion new devices in just the last year, ARM Holdings, pioneer of low energy processor design, is arguably Britain's most influential technology company.

Games consoles, DVD players, set-top boxes, mobile phones and the gamut of the Internet of Things - the uses for ARM's low power chip designs are seemingly limitless, a factor no doubt reflected in the company's soaring share price. However, walking around the ARM Campus, on the edge of a pretty village a few miles outside Cambridge, you're struck by the fact that the maths decorating the HQ of this multinational giant doesn't quite add up. There's ARM Building 1 and ARM Building 2 and ARM Building 3 - and then suddenly… ARM Building 6. So what happened to 4 and 5?

"The reason behind that," says John Goodacre, the company's jovial, red-haired head of technology, with disarming candour, "is the numbering on our processors. We had a processor design one and then a processor design two and three, but when we finished designing our fourth, everyone else was releasing their 'Version 6', so we thought we'd better call ours '6' too. We decided we'd do the same with the numbering on our buildings."

Nowadays ARM doesn't have to toe the line so much to attract attention to its products. The company has enjoyed a meteoric rise since 2004, built on its know-how in energy efficient chip design, which it licenses to a surprisingly small number of firms around the world. "We have fewer than 100 customers and around the same number of people in marketing," says Goodacre, "this is a personal business."

Being a world leader also counts for something. Goodacre claims ARM's smallest processor is four times smaller than its nearest rival's. The company is working on transistors for their chips that are just 28 nanometers, and they expect them to be even smaller in the near future.

The origins of the company lie in Acorn, the eighties British home computer company which built the legendary BBC Micro. In 1990 Acorn spun off its chip design wing as a separate company that came to be known as ARM. The tale behind the infamous low-power chip's design is equally fabulous: "The first processor we designed was simple, not low energy," says Goodacre. "We only had twelve engineers working on processor design, so it had to be simple. The low power element almost happened by accident. One of the power connectors got loose from the chip but the chip still worked so we built that into the design. It was only later that we realised we had a competitive edge in low power, so we ended up branding the company along those lines."

**"We have fewer than 100 customers and around the same number of people in marketing."**

**JOHN GOODACRE, ARM**

But because most computer companies in the early nineties designed and built their own chips, it took a while for ARM's products to take off. The breakthrough came when two high-profile companies became customers: Texas Instruments and Nokia. Goodacre recounts how ARM's chips in Nokia's new GSM phones came to attract attention: "Everyone asked why is this phone so small? Then they realised it must be the small processor."

# ALSO HELPING THE UK EXCEL IN ENERGY EFFICIENCY

## IMAGINATION

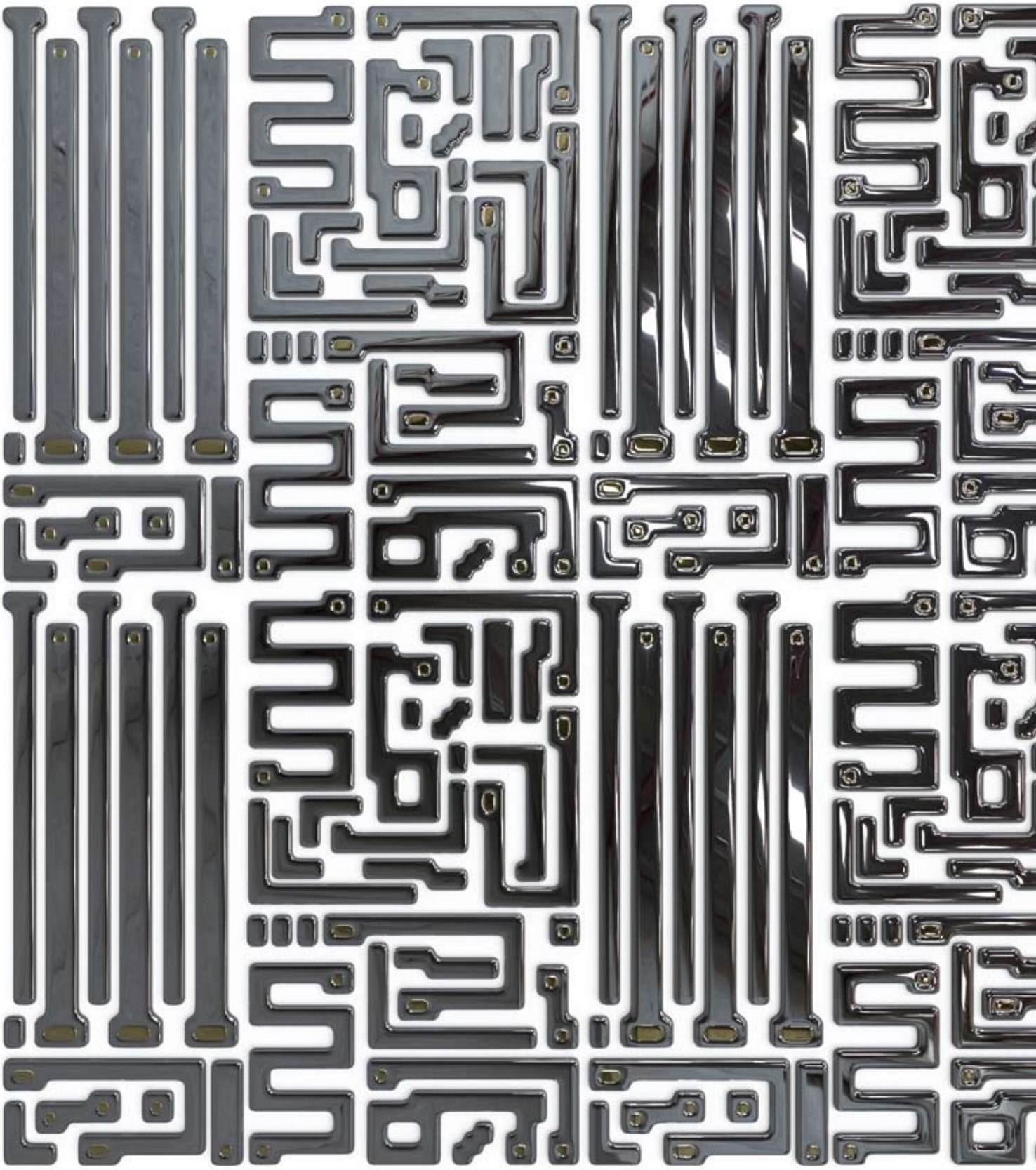
Imagination Technologies is a multimedia and communications IP company with headquarters in the UK and a listing in the FTSE top 250. Among their shareholders are Apple and Intel and their CEO Hossein Yassaie was knighted in the 2012 New Year Honours list. The company's technology is in a remarkable number of products associated with the IoT, including Google glass, smart watches, wireless audio, as well as cars and UltraHD 4K TV.

## XMOS

XMOS is a fabless (i.e. without a fabrication plant) chip designer whose blueprints for semi-conductor chips excel in demanding embedded applications that are time-critical and require flexibility. It was founded in 2005 by Hitesh Mehta (of Acacia Capital Partners) and the renowned David May, famous for designing the transputer, the first processor built for parallel computing.

The company's CTO, Simon Knowles says: "Embedded systems are becoming much more sophisticated and will require much higher levels of compute. Applications such as robotics, machine vision and natural language interfaces will drive innovation in processor architecture and energy efficiency for decades. XMOS is perfectly poised to build a leadership position in these new growth markets."

# EMBEDDED SYSTEMS>



# >SEEKING SOLUTIONS

Energy efficiency is increasingly being recognised as a key performance attribute of many systems, but for embedded designs, where energy consumption can be imperative, it's particularly significant.

When we talk about embedded systems, people often think of the IoT where they are present in products ranging from smart meters to wireless sensor nodes and mobile health monitors. Some need to operate reliably over a period of years or even decades while using limited energy resources (some of these are explored in the preceding section). But what many don't realise is that systems are also embedded in their home PCs and mobile phones - ticking away in the background, unnoticed - monitoring, backing up material and using up energy. We take a look at efforts to measure these in a subsequent section.

The key characteristic is that an embedded system is dedicated to handling a particular task. This means that design engineers can optimise it to reduce the size and cost of the product and increase the reliability and performance.

### Power consumption and sleep modes

At the heart of efforts to achieve energy efficiency in these devices are a new generation of microcontrollers (MCUs). And there are two important features to consider in their design.

Firstly, lower active power consumption doesn't necessarily equate to efficiency. In the past there's been an assumption that, because of their lower active power consumption, 8 or 16-bit solutions must be the best choice for low-energy applications. However, a modern 32-bit CPU core like the ARM Cortex-M3, can perform a task in around one quarter of the time required by older 16-bit CPU cores. In short, 'low power' does not necessarily mean 'low energy'.

Secondly, sleep modes are one of the most important features in an MCU. It must deliver real-time (or close to real-time) performance throughout a service life that is spent predominantly in a low-power sleep mode that reduces its speed of response to external events. A device such as a smart meter will often be difficult to access and is likely to spend in excess of 99% of its life in some kind of sleep state.

Most modern MCUs have more than one low-power mode (these will typically be something like: "Sleep/Standby"; "Deep Sleep"; and "Off"). Although it may appear that it would make sense to use the deepest possible sleep mode in every case, waking too frequently from deep sleep in order to perform a small amount of processing could conceivably burn more power than staying in a "higher" activity state, with less time wasted on the wake-up process.

### When everything is interactive

The Internet of Things has, naturally, awakened a great deal of interest in embedded systems and the ultra-low power integrated circuits, energy efficient communications and networking protocols need to make them work.

Adding energy harvesting techniques to the mix, academics at Columbia University in the US have developed a nanoscale chip that requires so little energy to transmit wireless signals that the batteries may never need to be replaced.

The project is part of a larger, award-winning research programme called EnHANTs and the goal is to make thin, flexible, energy efficient tags that can be attached to common objects and used for applications such as the IoT, logistics, tracking and search or disaster recovery.

### Embedded multicore systems

The UK currently leads the world in the embedded multi-core systems field, thanks to the achievements of ARM and Imagination, whose products power the on-going mobile revolution.

The two companies are also collaborating with Altera, Microsoft Research, Freescale and a group of UK universities to develop more efficient computer architectures for dependable embedded systems with many-core processors.

The work is funded by a five-year £5.6m Engineering and Physical Sciences Research Council (EPSRC) grant and has been named PRiME - Power-efficient Reliable Many-core Embedded systems.

"Embedded systems, many of which will be low-power mobile devices, will be one of the most powerful tools in tackling global societal challenges," says Professor Bashir Al-Hashimi from the University of Southampton, director of PRiME. "However, while many-core processing is viewed as a way to improve the performance of computing systems, the energy consumption and reliability of these systems with hundreds or thousands of cores has yet to be fully understood.

# MULTICORE COMPUTING

## A PARALLEL UNIVERSE

Computer scientists and developers are experimenting with myriad combinations of chips to perform tasks both more effectively and more efficiently. Parallel computing - where calculations are carried out simultaneously - has become the dominant paradigm in computer architecture, mainly in the form of multicore processors, each performing different tasks. According to ARM's John Goodacre: "Mobile phones went dual-core a few years ago. Now they are quad-core, and we're now seeing eight-core, but there is also a reverse process going on, in that developers are constantly trying to make these smaller processors more powerful."

David May has plenty of credentials in this area and believes that it's an area with scope for much more exploration. "Until recently the trend in computing has been to get big complex processors to run programmes faster and faster," he says. "But trying to get processors to run faster is not very energy efficient. Computing doesn't have to be tackled this way. Do you have a few big processors to run many different tasks, or do you have a lot of smaller simpler processors, each with a different task? That is the big question in computing today."

Improvements in the architecture that links the various chips, so that they work more efficiently together, is key. For instance, up to now the central processor - the CPU - and graphics processor - the GPU - have not worked very efficiently together which leads to significant dispatch latency, with overheads that makes the process worthwhile only when the application requires a very large amount of parallel computation.

In the UK, ARM is aiming to lead the fray in solving some of the issues, with a power optimisation system named big.LITTLE, already installed in the Samsung Galaxy S4. The technology works by combining a high performance processor with a super efficient one. When the device needs a lot of muscle, tasks are automatically given to the more powerful processor but when the compute load is lighter (for instance when monitoring signals) the powerful CPU is switched off and the low power one used. This can save up to 75% of energy and increase performance by 40%.

While HSA - Heterogeneous System Architecture - brings several different processors together, placing specific tasks on a particular type, in big.LITTLE, they work together as a common team. "This is the ARM project that will have the greatest impact for general purpose computing," says Goodacre. ARM and Bristol-based Imagination are, along with Texas Instruments, AMD and Qualcomm, founders of an international HSA foundation, a non-profit association of tech companies dedicated to creating an HSA ecosystem with universal open source industry standards.

The reasons behind the popularity of this trend are obvious. "Many modern apps are very easily parallelised (i.e. run with parallel computing)," says David May, and he believes it may well be the shape of things to come: "In robotics, for example, the next big tech breakthrough will be naturally parallelised."

It also helps that many of the functions needed by the programmes run by the technology giants lend themselves to the process.

"Google and Facebook spend most of their time mining very large graphs, keeping relationships and looking for connections between them" says May. "This is also easily parallelised. But the answer to the question "lots of little processors or one big one?" is not obvious. We need to do a lot more research in this area."

# DATA CENTRES

## COOLING THE SERVER FARMS

# DATA CENTRES

are among the major consumers of electricity and are thought to use between 1-2% of the world's total. The big Internet businesses - including Apple, Amazon, Google and Facebook - have recently become deeply engaged in the issue of energy efficiency in their data centres. The challenges here are to find new, cheap and green ways to keep the thousands of servers in their data farms running without over-heating.

"How Clean is your Cloud?" - a Greenpeace report published in 2012 - was highly influential in persuading the technology giants to take energy efficiency seriously. It praised Yahoo, Facebook and Google for their commitment to using renewable energy in their server farms, and castigated Amazon, Apple and Microsoft. Since then, Apple has announced that all its data centres are 100% powered by renewable energy. That has meant building some big solar farms, including, most recently, one on a 137 acre site in Nevada. Meanwhile, Microsoft has signed a 20-year deal on a 110 megawatts (MW) wind farm in Texas. Amazon has apparently taken little action. Its Web Services (AWS) division includes Netflix, Pinterest, Spotify, and SmugMug and is estimated to be accessed by a third of all internet users every day, so its impact is substantial.

With imminent government regulation, environmental pressure to economise on energy bills is likely to be further amplified and the ICT industry is putting effort into keeping sustainable energy to the fore of policy. In 2008, Britain's technology trade association, Intellect (now known as techUK) released its "High-Tech: Low-Carbon" report, which articulates an action plan on how the UK technology sector is going to reduce its emissions.

Further, the trade organisation

Digital Europe (formerly EICTA) has committed to reduce Europe's ICT-related carbon emissions by 20% leading up to 2020. Many think that target is achievable by 2015.

The UK, in particular the BCS Data Centre Specialist Group, has taken a global lead in advancing the field of energy efficiency within data centres, and was instrumental in developing the European Union's Code of Conduct, which stipulates a range of best practices for every layer of the IT service delivery stack (from mechanical and electrical to software selection). One of the Code's authors, Kate Craig-Wood, is also Managing Director of Memset, a UK based data centre.

The market for more efficient data centres will rise from $17.1 bn in 2012 to $45.4bn by 2016.

Organisations are looking at other measures too, with the EU funding its own low energy data centre projects. Under the Ecodesign Directive, the European Commission is expanding its minimum energy efficiency regulations to servers and data storage equipment. Another project, Eurocloud, has developed a chip that uses just 10% of the energy of those usually found in data centres, capitalising on ARM's low energy processors. "Today's power-hungry cloud data centres are not sustainable in the long run," said European Commission Vice-President Neelie Kroes. "The Eurocloud chip addresses the core of this energy consumption problem. I hope further development of the Eurocloud chip will boost the position of European businesses in a sector currently dominated by non-Europeans."

## Keeping cool

The biggest slice of a data centre's energy bill is not running the servers, but powering the cooling system for them. An increasingly popular solution has been to situate new European centres inside the Arctic Circle. In 2013, after four years of construction, Google switched on its data centre in Hamina on the Finnish coast, where freezing seawater is piped in to chill the server racks. In March 2014, Google also announced plans to buy the entire electricity output of a Swedish wind farm to power the centre. Microsoft, Eriksson, Yandex, Digiplex and other players have also been investing heavily in data centres in Sweden and Finland.

In June 2013, Facebook turned on its first European server farm in Lulea, Northern Sweden. It inducts icy Arctic air to cool the servers for eight months of the year. For the rest of the year, instead of the usual air-conditioning approach, a cooling system based on evaporating water, powered by energy from a nearby hydroelectric dam, is employed. The centre boasts a complementary system that removes the heat produced by all the racks to warm nearby offices. Facebook's Lulea farm consumes just 120MW of energy per year, which the company says is less than half the energy needed to run an average data centre.

Iceland is the latest Northern European country trying to lure in the data centre business with the offer of a cold climate and a nation that uses geothermal energy to provide much of its power needs.

But according to the annual Data Centre Risk Index 2013 (published by consulting firms Cushman & Wakefield, Hurleypalmerflatt and Source), Britain and the US are still the safest places to build data centres, although the Nordics are beginning to offer fierce competition and British data centres have some catching up to do on the energy efficiency front. According to Pike Research, the market for more efficient data centres and those using energy from renewable sources will rise from $17.1bn in 2012 to $45.4bn by 2016.

## Air versus water

Servers in the majority of data centres are cooled by air, but research indicates that liquid is much more effective in transferring heat and uses significantly less energy to do so. Liquid cooling is now becoming an established trend, although the technology has been around for some time - the first water-cooled data centre went live in 2005. In Cambridge, ARM has built one that uses rainwater.

As a measure of the success of its patented approach, in January 2014, the Yorkshire-based liquid cooling start-up Iceotope announced an investment of £6m from Aster Capital, who are sponsored by Alstom, Solvay, Schneider Electric and the European Investment Fund. Iceotope's cooling system submerges computer racks in a special liquid, M™ Novec™ Engineered Fluid, which removes the heat either for disposal or recycling. The company says that its technology lowers cooling costs by 97%, ICT power load by 20% and overall ICT infrastructure costs by 50%. The deal should give Iceotope access to the large server client base that Schneider already supplies.

Britain and the US are still the safest places to build data centres.

Proponents of liquid cooling say that it allows racks to be built at high density. Since air doesn't need to circulate, the problem of too many cables blocking the supply is eradicated, and the racks are also much quieter without the air-con approach. According to eBay's data centre expert Dean Nelson: "Liquid cooling will remove the locks on chips, and take away constraints."

## Data centre design

In 2011 Facebook launched its 'Open Compute' project, an effort to get tech businesses to share their data-centre design know-how, with a view to pooling knowledge to increase energy efficiency. It's a bid to move data centre operators away from buying general purpose servers. Today, Open Compute has 1,500 members including Rackspace, Intel, AMD, EMC, VMware, HP, Dell, ARM and Goldman Sachs - but not the big rivals Google, Amazon, Apple or Microsoft. Using Open Compute, Facebook designs its own efficient servers and other data centre equipment, the blueprints for which it then shares as open source.

Apple has also implemented an impressive range of energy efficient design strategies at its flagship data centre in Maiden, North Carolina:

- A chilled water storage system that maximises the use of off-peak electricity.
- Importing cool night air on cold weather days, which allows the cooling system to be switched off for 75% of the time.
- 'Cold air containment pods' with fans whose speed can be set according to needs.
- High-voltage power, which has less wastage than low-voltage.
- A roof that maximises solar reflectivity.
- LED lights.
- Real-time energy consumption analysis.
- The use of 15% recycled materials in construction.

## Virtualisation

Despite all the evidence pointing to cooling as a vital element in data centre efficiency, there are major companies - like Google - who don't think the cooling system is the key to energy reduction. They think that a better route lies through the redesign of server architecture and software.

Here, similar energy-saving innovations to those used in HPC are applicable: 3D architecture for chips and low energy solid state drives (since they have no moving parts). But by far the most important innovation is the virtualisation of servers - also known as the software-defined data centre - a major new trend enjoying widespread implementation.

Ordinarily, servers run one operating system or application at a time, which compels data centres to operate scores of servers, each working at 5-15% of their capacity. Virtualisation creates a data centre within the server, allowing several different OS's and apps to run on one host. Each 'virtual server' is separate from others and draws only as much computing power as it needs.

The energy savings of virtualisation are staggering. In September 2013, the Energy Efficient Computing Special Interest Group (EEC SIG) published its "Towards the estimation of the energy costs of Internet mediated transactions" report, which calculated that virtualisation is allowing the consolidation of physical machines by a factor of more than ten to one.

There are major companies – like Google – who don't think the cooling system is the key to energy reduction.

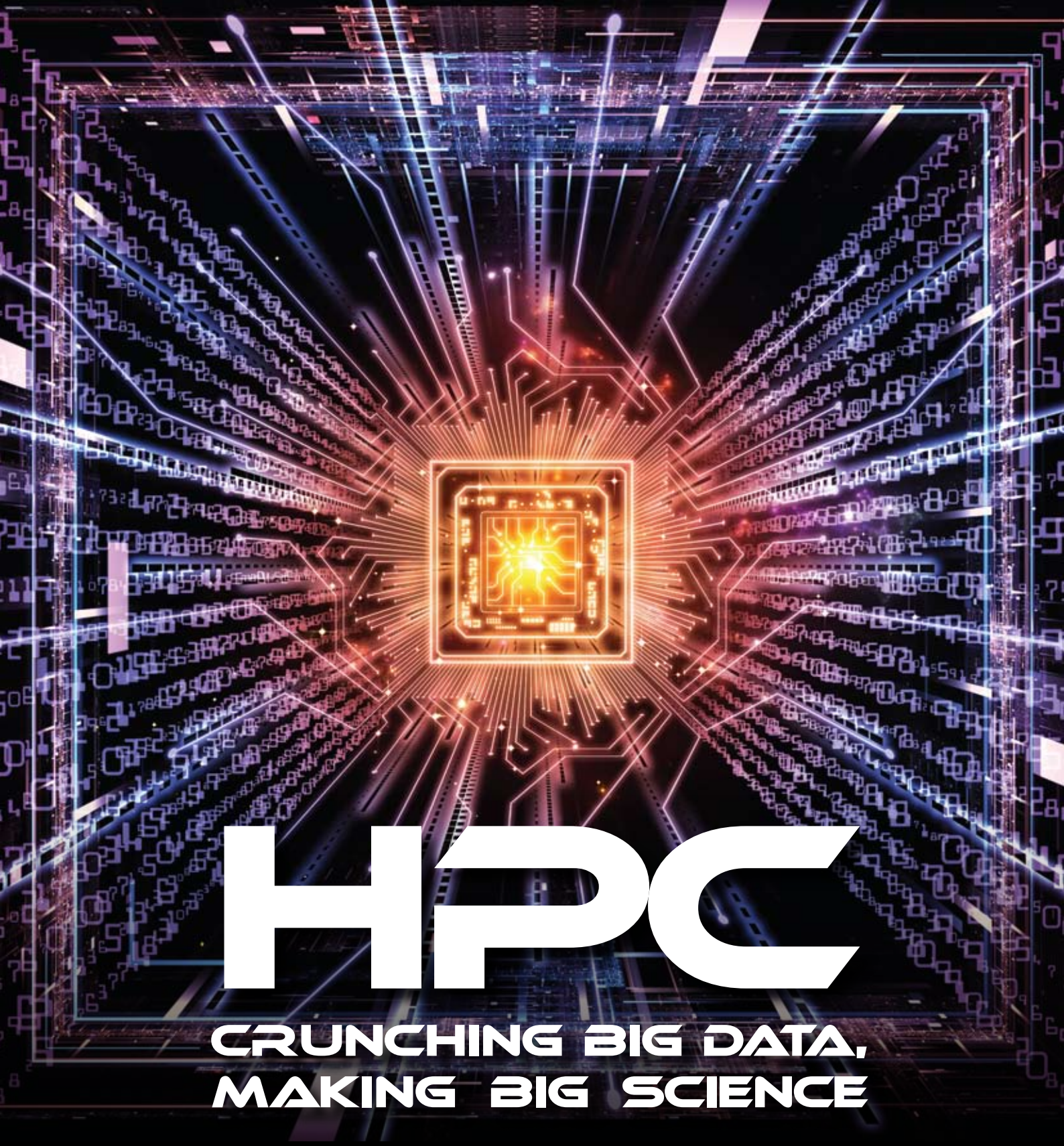
The British server and hosting data centre company Memset uses Xen-based Miniserver VMs (Virtual Machines or virtual servers). In 2009, the company deployed 1,000 virtual servers, each consuming 5-10W, but doing the same work as a normal server using 90-100W.

With this in mind, a broad 100,000W saving can be calculated, so, according to the report "one small-scale data centre in just one small location in the world has helped avoid over 30 tonnes per month, or 360 tonnes per year, of carbon dioxide emissions. To put that in context, each British citizen is responsible for about 9 tonnes of CO2 emissions per year."

## Beginning to pay off?

There are signs that all these energy saving efforts are beginning to pay off. The number of data centres is expanding rapidly, however their energy consumption grew far more slowly in the last five years than in the early part of the twenty first century.

Stanford University professor Jonathan Koomey conducted a study on growth in data centre electricity use from 2005 to 2010. He found that the rapid rates of growth that had prevailed from 2000 to 2005 slowed significantly from 2005 to 2010. This seems to indicate that the industry is becoming more efficient. Or does it? Some would say that these lower levels of use are due to a combination of factors - possibly the recent recession as well as advances in energy efficiency.



# HPC

## CRUNCHING BIG DATA, MAKING BIG SCIENCE

To thrive and excel in the emerging age of Big Data, governments, scientists, businesses and services - large or small - will have to analyse unprecedentedly vast amounts of information about their citizens, customers, clients and the cosmos.

According to a Computer Weekly study, Route to Real Time, "75% of UK organisations are currently investing in big data analytics, and 80% of these are deploying solutions." Information, as they say, is the new oil. Yet the British government believe that businesses, particularly SMEs, are under using the opportunities presented by High Performance Computing (HPC) - the vehicle necessary for the number crunching which super accurate analysis often requires.

The main obstacles identified by experts - including the International Exascale Software Project (IESP) and the European Exascale Software Initiative - are availability and the economic and ecological energy cost needed to power and cool the new, high performing supercomputers. This has also been identified as one of the major challenges in designing and operating future multi-petascale and exascale HPC systems.

## GOVERNMENT STRATEGY

The British government's strategy to make Britain a centre for energy efficient HPC is to invest millions in Big Science projects which will have direct and indirect commercial applications.

The largest of these is the SKA project, the Square Kilometre Array, which is the largest radiotelescopy enterprise in history and has attracted funding from many nations. It involves placing thousands of dishes and three million antennae in remote parts of South Africa and the Australian desert.

The SKA telescope is 100 times more sensitive and one million times faster at mapping the sky than anything that has gone before. When the first phase of the project is completed in 2020, the entire network will be handling around 20,000 petabytes of data a day - a number that will rise by a factor of 10 once the full SKA is completed in 2028. By comparison, the internet currently transfers about 300 petabytes a day.

This volume of data requires a huge increase in data and processing. "We've wanted to do this for a long time," says Professor Paul Alexander, who is leading the British contribution to SKA. "But we can only do it now because the tech will be here tomorrow."

For the uninitiated this could be construed as wishful thinking, but Alexander is certain that technology will keep apace of the needs of the project. "This is a massive computing problem and massive EEC problem," he says. "At the moment we simply cannot afford to process this amount of information. HPC needs to become five times, but ideally ten times, more energy efficient by 2020. But I think we will probably get an increase of twenty times."

**"75% of UK organisations are currently investing in big data analytics, and 80% of these are deploying solutions."**

In Cambridge, Alexander and his colleagues are working on the low energy architecture of the SKA supercomputers, some of which will have to be built in the desert. As part of the new "SKA Open Architecture Lab," Wilkes, the fastest GPU supercomputer in the UK and the world's most energy efficient air cooled HPC system, has been developed in partnership with Dell, nVidia and Mellanox. It comes an impressive second in the worldwide Green500 ranking. The team are planning to use fluid to cool the next one they build. Cambridge University is also, separately, building a cutting edge data centre, to be opened later this year. This will also offer state of the art ways to test cooling systems.

The British government sees long term direct and indirect commercial benefits to be drawn from making a major investment in the SKA project.

Firstly, it places Britain at the forefront of the growth market in EEC.

Secondly, it draws in major commercial tech entities; in the UK, SKA is working with the biggest industry players including Intel, nVidia, IBM, Dell, and HP, who are all designing different parts of the system.

**"HPC needs to become five times, but ideally ten times, more energy efficient by 2020."**

Thirdly, the technology used to process Big Data from the cosmos will be useful close to home. "We are trying to solve Big Data problems," says Alexander. "Streaming data off the telescope is very similar to the problems of Big Data here on earth - dealing with sensor networks, for example, that analyse the traffic flows in the city."

Finally, the technology that comes from Big Science projects has a history of creating unforeseen, indirect economic stimulus. This policy is central to Alexander's beliefs and something he's emphasised to the UK's Chancellor George Osborne. "Often ICT growth doesn't come from direct sales alone," he explains. "For example, wi-fi was developed from a radio astronomy project in Australia. CSIRO [the Commonwealth Scientific and Industrial Research Organisation] own the patent and get a tiny percentage of sales, but the value for the economy is not in sales, but in all the things wi-fi is used for."

## SOFTWARE LEADS THE WAY

The UK's supercomputing hub, at the Hartree Centre in Daresbury, Cheshire, opened in February 2013 with a £37m government investment. It's involved in the SKA project but also caters to commercial clients. Two supercomputers - Blue Joule and Blue Wonder - churn through mountains of data, conducting virtual experiments on shampoo to produce a structure that gives you 'that silky feel' for Unilever; aiding the design process for Bentley car interiors and working out how to make vehicles lighter for Jaguar Land Rover. Hartree also works with a number of SMEs, including ACAL Energy who make platinum free fuel cells.

Energy efficiency has become a priority here. "Everything in the Hartree Centre had been focussed on speeding time-to-solution", says Michael Gleaves, the Centre's Head of Business. "Now we have a separate category - the cost of computing. If you're going to move to exascale, you need a 50-fold improvement in your systems to make it affordable for science and business. We're starting to think about the algorithms and types of architecture that will allow this to happen."

"The real technological and commercial opportunity here lies not in the hardware but in developing fast but energy-efficient algorithms," he continues. "Most of the codes written nowadays are made as scalable as possible and as quick as possible. That may be in an attempt to be more energy efficient, but it's not necessarily the case."

Hartree is pursuing a software-led approach. HP claim to have built the first operational software-defined server with their Moonshot system, but Hartree are following in their wake, running tests with the latest chips which allow you to change the speed of the processor to match its task. "The low-energy chip might need only a third of a gigahertz (GHz)," says Gleaves, "and take only 1-2% longer to perform the task."

Hartree has recently spent £19m on new equipment, with much anticipated state-of-the-art ARM processors due to arrive in the Spring (2014). "That will be very interesting to scientists because the next ARM servers are 64-bit," says Gleaves. It's the most viable option for efficiently handling vast amounts of data.

**"The real technological and commercial opportunity here lies not in the hardware but in developing fast but energy-efficient algorithms."**

Hartree has also acquired a FPGA supercomputer from Maxeler, an HPC solutions specialist, based in London's Hammersmith. FPGA stands for Field-programmable gate array and consists of an integrated circuit designed to be configured for ultimate efficiency by a customer or a designer after manufacturing.

Meanwhile, the recycling of energy produced by HPC is also taken into account, and Hartree have acquired a liquid cooled system, "to evaluate how much of the heat we could extract from HPC and what we could do with it. To heat an office you need high-grade heat of 40 degrees, but we only have 16-18 degrees of heat coming off our processors at the moment. So we need to improve that."

### NEURAL NETWORKS AND 3D PACKAGING

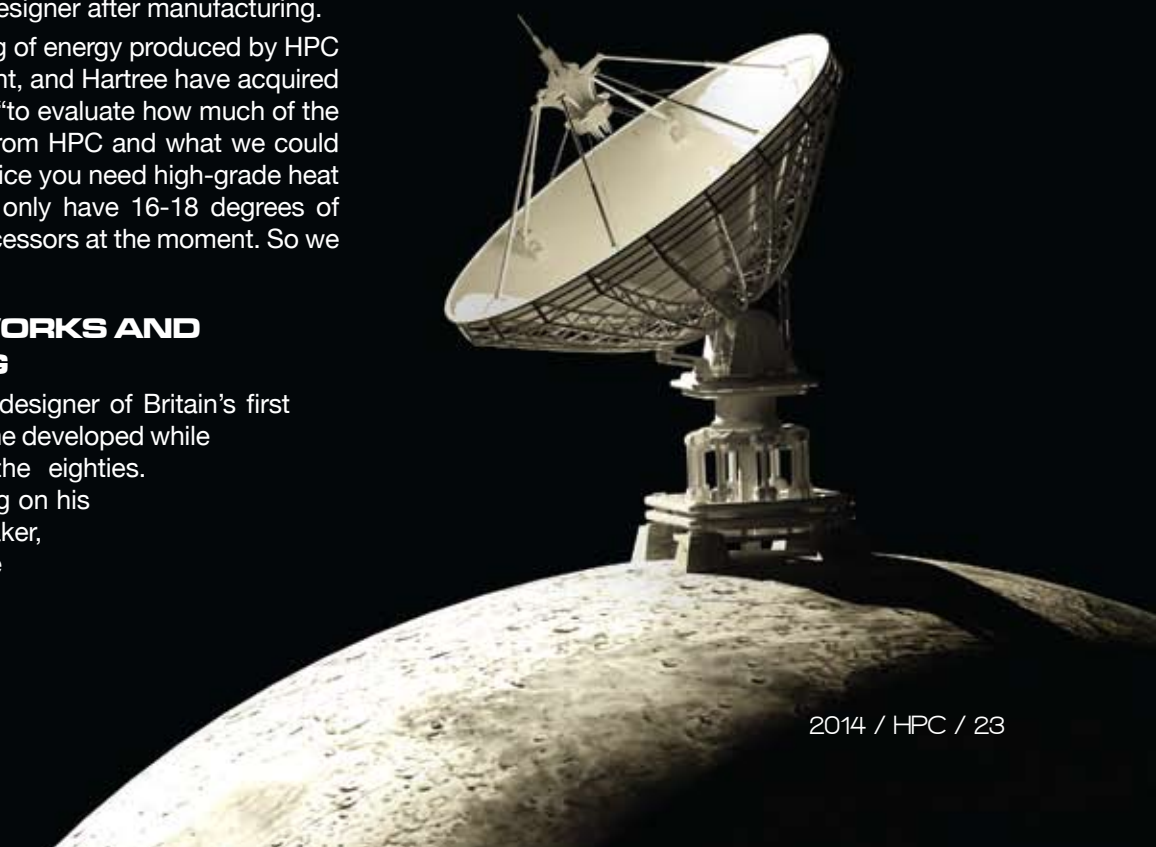
Steve Furber is the co-designer of Britain's first low energy chip, which he developed while working for ARM in the eighties. These days he's working on his supercomputer, Spinaker, which reproduces the neural networks of the mind. Spinaker uses one million low power ARM processors, which consume 100 kilowatts at full blast. "We've traded off performance for energy efficiency in a number of ways to keep the electricity bill under control," he says. "Part of energy efficiency is making it proportional to the computer and appropriate for the task."

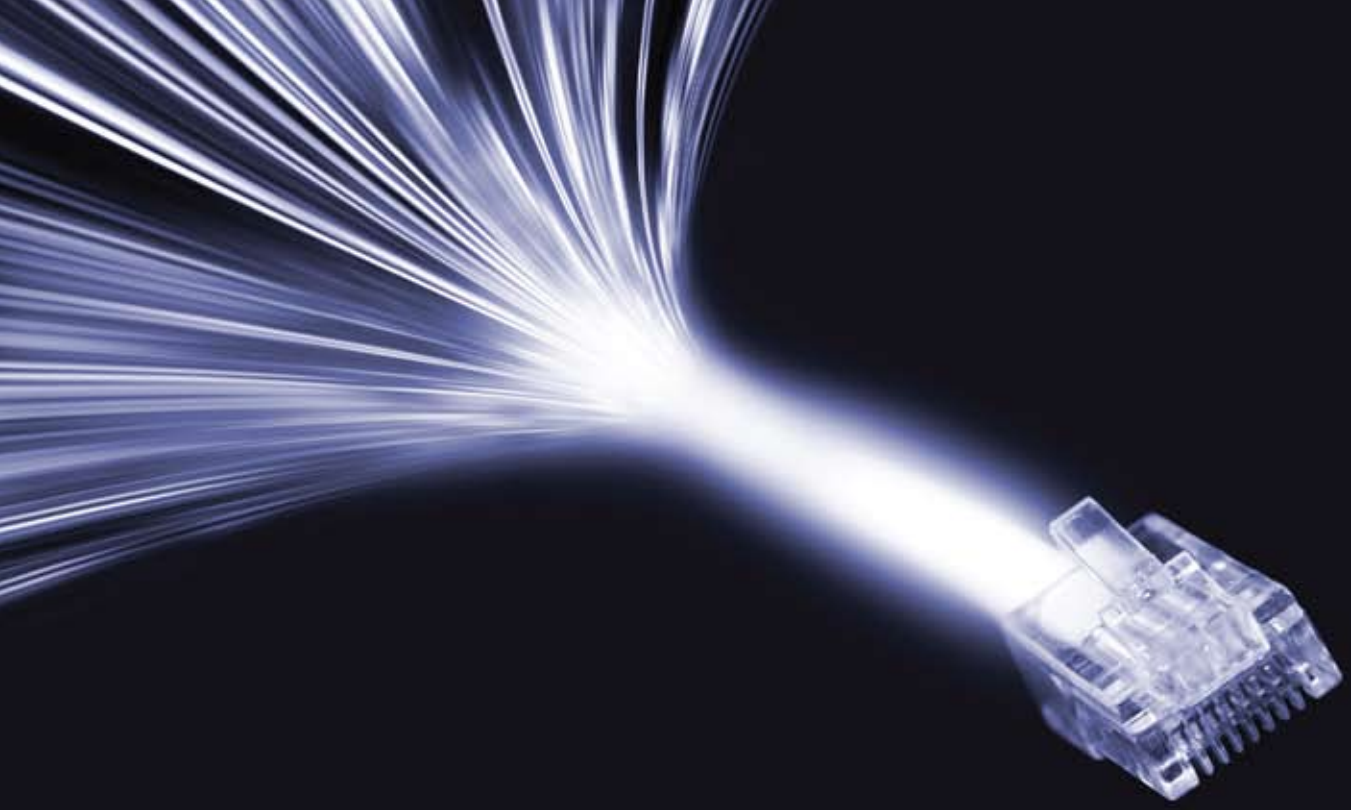
"Throughout most of the history of computing," says Furber, "energy efficiency has been delivered through transistor shrinkage, but now that has reached its limits. There are several interesting technologies using new materials for chips, like graphene or light, but they are a long way from being able to market a product. Water-cooling is effective but we should work harder to stop generating the heat."

Furber believes the key to energy efficiency lies not just in the kind of chips one uses, but also in their arrangement. The process of 3D packaging stacks chips one on top of another, instead of in rows next to each other. This allows them to be placed much closer together.

"In engineering terms," says Furber, "the key to energy efficient computing is to keep data as close to the processor, where it is used, as possible. So you have to move it as little as possible. Movement requires energy. So the technology of 3D chip stacking, which is already on SD [secure digital] cards, is a big step forward."

Furber's belief is that the UK should capitalise on its strength in energy efficient computing and a focus on 3D packaging may be a good way to proceed: "The UK has a strong niche in silicon chip design, thanks to ARM," he says, "and now David May's company XMOS in Bristol - but we don't do anything above that. There is a lot of development in this space, with a lot of high value added. We could really benefit from a company that does packaging."





# LOW POWER HPC RESEARCH

Having estimated that a combined energy saving approach in HPC would amount to a saving of 1157 terawatt-hours per year by 2030 - twice the total annual energy consumption in Sweden - the European Commission has made energy advances in HPC a priority in the Ecodesign Directive.

Taking a software-defined approach, ADEPT - ADdressing Energy in Parallel Technologies - is a three-year EU funded programme run by the Edinburgh Parallel Computing Centre in association with Swedish and Belgian teams. Its aim is to develop a tool that will allow "users to model and predict the power consumption and performance of their code" through the efficiencies of parallel computing.

Developers in America are pursuing similar R&D lines to the UK. Scientists at Case Western Reserve University announced in December 2013 that they had built nanoscale switches and logic gates whose energy efficiency dwarves those currently installed in computers, tablets and smart phones.

# MEASURING ENERGY USE

## Getting hardware and software to work together

Accurate measurement of energy use is a prerequisite to its reduction - in computing as in all things.

Scientists and business communities want to know how much cheaper it is to use the cloud than their office servers. They also want to know how much energy is saved by using the Internet instead of buying things AFK - away from the keyboard. Does it ultimately save resources to order online, instead of driving to the supermarket or download MP3s instead of buying a CD?

How do we calculate all this? How do we find out exactly how much energy mobiles, home PCs, data centres and HPC are using? The problem here is developing adequate measuring equipment and industry standards for the task.

Many different models for EEC measurements are competing with each other and are often controversial in their methods and conclusions. Now efforts are underway to harmonise measuring systems and create universal standards.

### DATA CENTRES AND HPC: IMPROVING ON PUE

Until recently, the standard industry tool for measuring the efficiency of data centres has been Power Usage Effectiveness (PUE), which contrasts server power consumption against that of the whole data centre. It includes lighting, chillers and other cooling equipment needed to keep the computers performing as designed. The perfect PUE rating would be 1. Today the average PUE of a data centre is 2.5, but now many data centres, especially the ones run by the largest image-conscious internet brands, achieve PUE's of between 1.1 and 1.6. ARM's data centre operates at even less than that.

**Many different models for EEC measurements are competing with each other and are often controversial in their methods and conclusions.**

Recently, PUE has come in for some criticism. Firstly, it's argued that it reveals nothing in terms of the actual energy consumption of a data centre and, secondly, it's all too often used by data centre owners as part of their 'green credentials' campaigns. In certain configurations, it actually penalises better performance. For

example, if a cloud manager identified servers in their data centre that were not being used, and elected to shut them off and create virtual servers, this could result in a decrease in the power consumption rate (good) but an increase in the facility's PUE (bad). The same amount of power is being used by the facility to cool and light the building even though there's less power being used by IT equipment. This illustrates a fundamental point: it's vital to address how efficient hardware and software are themselves as they relate to performing actual compute tasks.

The computer giants are striving to find their own measures of efficiency. Google has set itself up as something of a pioneer in this area, becoming, in 2011, the first internet corporation to publish details of its energy use and alerting the rest of the industry to this need.

**It's vital to address how efficient hardware and software are themselves as they relate to performing actual compute tasks**

Google funded a further report in June 2013, which factored in the energy use and carbon footprint of not only the data centre, but also of transportation, manufacturing, waste management, network equipment, appliances and gadgets used to access the data. Using the Cloud Energy and Emissions Research Model (CLEER), the study concluded that if American businesses moved their software, customer databases and email to the cloud, there could be savings of almost 90% in the energy used, enough to generate the electricity used by Los Angeles each year.

Meanwhile, eBay has its own energy-measuring model, Digital Service Efficiency (DSE) metrics, unveiled in March 2013, which it is trying to push forward as an industry standard. It essentially measures work divided by power.

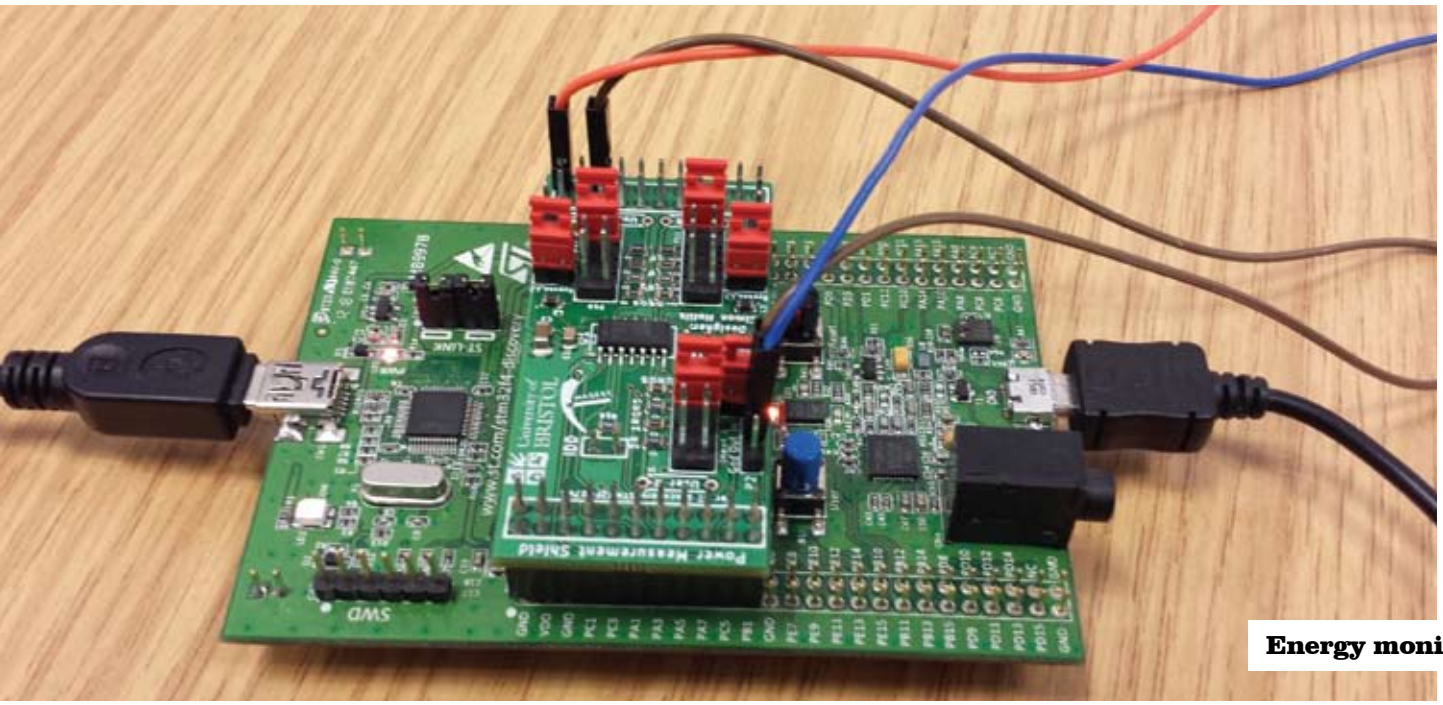
The company boasts that the metric has already allowed it to remove 400 servers from a pool (out of a total of over 50,000) by slightly decreasing the memory allocated for one application. It also eliminated 1MW of power and saved itself $2m (calculated as the costs of replacing the servers.) Another way the company measures efficiency is against profits, however the energy required to ship its user products does not appear to be something it takes into account.

A more 'holistic' energy-measuring model is being developed at the Leibniz Supercomputing Centre of the Bavarian Academy of Sciences (BAdWLRZ). Their Power Data Aggregation Monitor (PowerDAM) combines measuring HPC systems with data from cooling and building infrastructure.

## STANDARD SOLUTIONS

With such differing standards, it's very hard to effectively judge efficiency, but the Green500, Top500, Green Grid and Energy Efficient HPC Working Group are working on a solution. They're collaborating to create standards for measuring energy produced by HPC and data-centres with a focus on "improving the methodology for measuring energy in order to get an apples to apples comparison between system architectures."

And there's also an aid for unravelling the HPC needs of companies engaged in engineering complex products. They have a bewildering array of options: software, hardware and algorithms and the choice of using in-house or cloud infrastructure. In order to help them make sense of these, a Technology Strat-




**Energy monitor**

egy Board (TSB) supported project, GREENS (Global Register for Energy Efficient Numerical Simulation), has been set up - it's basically an EEC version of a price comparison website. GREENS is also now developing a standard measure of energy efficiency, so that devices, systems and software energy use can all be measured and characterised.

In the UK, the EEC SIG has produced its own exhaustive report, "Towards the estimation of the energy cost of internet mediated transactions," attempting to measure conclusively and comprehensively how much energy the internet is using. The report points out that any assessment of the energy efficiency of ICT should take into account all the journeys, by car or plane, saved by people who perform a task online. Their modelling suggests that the Internet is a considerably more efficient place to store and access data than previous studies claimed.

Everything points to the fact that the cloud is vastly more efficient than home computers and office networks but the report also warns that efficiency gains might be offset by increased resource consumption.

## TIME TO TACKLE THE SOFTWARE

Up to now, hardware has been the focus of most scrutiny but many, including researchers Kaushik Roy and Mark C. Johnson, have argued that this emphasis has been misplaced:

"It is tempting to suppose that only hardware dissipates power, not software," they wrote in 1997. "However, that would be analogous to postulating that only automobiles burn gasoline, not people. In microprocessor, micro-controller, and digital signal processor based systems, it is software that directs much of the activity of the hardware. Consequently, the software can have a substantial impact on the power dissipation of a system. Without a power estimate there is no way to reliably optimise software to minimise power".

The development of energy efficient software requires metrics, which measure energy consumption, as well as models to monitor and minimise it.

With the aim of giving coders the tools to measure the energy their programmes use, two projects - MAGEEC (Machine Guided Energy Efficient Compilation) and ENTRA (whole systems energy transparency) - are being developed.

Supported by the TSB Energy Efficient Computing Initiative, the MAGEEC project is led by Simon Hollis from Bristol University and, on the industrial side, by Jeremy Bennett at Embecosm, a company that specialises in compiling software - the code that translates programmes into the binary ones and zeroes that are put onto chips.

**"Historically, the people who programmed a computer never thought about energy"**

"Historically, the people who programmed a computer never thought about energy," says Bennett. No one ever looked at a £500 computer and thought that it could burn through £2000 of electricity in its life, because that was just the household bill. But there is a second issue too: until now, on the whole, software engineers also haven't had the tools to see how much energy is being used. The measuring kits available cost thousands and thousands of pounds."

This, according to Bennett, has led to an imbalance. "The hardware guys are good at working out where energy is wasted, but the software guys are not. A software designer might think it's a great idea to have a clock ticking away in the corner of the screen when the programme is running, telling you how long you've been using it for, but this uses up a lot of energy. It's

**The ENTRA team, led by Kerstin Eder**

not something that software designers have been thinking about until now, when the energy efficiency of battery life has become the key differentiator of products on the market."

MAGEEC is making a cheap (£30) open source board to allow software engineers to measure energy use. The consortium has been buying chips from all the major manufacturers to make sure their board can work across the industry. In addition, Bennett and his colleagues are developing software that will translate the code into binaries more efficiently, i.e. with fewer of those ones and zeroes. "I reckoned over the long term, the value of the energy saving potential from MAGEEC will be equivalent to closing three power stations," says Bennett. "Though I admit that doesn't factor in the increase in use that usually comes with energy efficiency."

Meanwhile, ENTRA, led by Kerstin Eder, Reader in Design Automation and Verification at Bristol University, is about aligning software design decisions with energy efficient goals.

"Lack of energy transparency in today's software development tools means that much of the potential energy saving available from power efficient hardware is not utilised," she explains.

The three-year EU funded project aims to push energy efficiency to the fore as a first class design goal and the plan is to use the project's results within a mainstream compiler and create demand on the end user side.

## MAJOR STEPS

So it's looking as if things are heading in the right direction. The rise in energy use by data centres has slowed significantly in recent years and standards and software measuring systems are slowly beginning to appear. But are we doing enough?

"The simple answer is that software is where biggest potential gains are," says XMOS' David May. "The efficiency of well-implemented software can be vastly greater than slapping stuff together. In simple areas, like the way you sort data, a well-chosen algorithm can be orders of magnitude more efficient than a bad one. Tweaking the hardware gets you improvements of a factor of two, but the software efficiencies possible are of a magnitude of orders of up to 100."

But, to gain maximum efficiency, software and hardware need to work together and, according to TSB Lead Technologist Jonathan Mitchener, one famous entity, in particular, has proved the efficacy of this approach: "Occasionally, there is a company, like Apple, where the hardware and software departments do talk to each other, and that's why it works," he says. "But typically engineers and computer scientists don't work together. Technology is nowhere near as integrated as manufacturing and this is a result of education, because people do either electronic engineering or they do computer science. They don't really cross over that much. But if you have separate hardware and software all the time - this can cause problems, especially for energy efficiency."

**"Lack of energy transparency in today's software development tools means that much of the potential energy saving available from power efficient hardware is not utilised"**

# ENERGY SHOCK

**Google used 2.26 million MW hours of electricity in 2010. This, according to the New York Times, is about a quarter of the output of a nuclear power plant.**

**The American-based Digital Power Group claims that an iPhone uses more power than a fridge once you add in the cost of the data centres, wireless connections and networks it needs to function.**

**Jon Summers, from the University of Leeds, has attempted to measure the energy expended in users downloading the Youtube video hit "Gangnam Style," one of the most popular videos of all time. He calculated that the energy consumed by 1.7 billion people viewing it over a period of 12 months was more than the power used by the 9.8 million population of Burundi in the same period of time.**

**Cost effective investments in energy efficiency could save the UK 196TWh in 2020, equivalent to the output from 22 power stations.**

# NEW ENERGY EFFICIENT MATERIALS

One way to save energy is to change the materials used to carry it. In computing, up to now, silicon has been used almost exclusively, but recently new materials are being touted as the future of EEC.

## Electronic Blood

IBM has built a new computer powered by what it calls "electronic blood" - a liquid, containing the element Vanadium, which both fuels and cools the computer. The electrically charged fluid is pumped through the computer. It carries power and it removes heat. Essentially, it's technology mimicking nature, where blood supplies energy and minerals to the body, while taking away excess heat.

The technology has been developed at the company's Zurich lab by Dr Patrick Ruch and Dr Bruno Michel.

IBM was inspired by the super efficient human brain, which uses approximately the power of a 20-watt light bulb yet is around 10, 000 times more efficient than any computer today.

In a BBC interview, Michel explained, "Ninety-nine per cent of a computer's volume is devoted to cooling and powering. Only 1% is used to process information. And we think we've built a good computer? The brain uses 40% of its volume for functional performance - and only 10% for energy and cooling."

Computers today are effectively air-cooled by the air that circulates between them, but as more and more transistors get placed closer and closer together this modus operandi becomes increasingly difficult. It's now limiting the development of 3D computing, which requires a great deal of processor power.

IBM says its new technology is a solution to this problem - it interlays the chips with minuscule water pipes. The company claims that the new technology could make HPC machines 10,000 times more efficient and one day allow a one petaflop computer - currently the size of half a football field - to fit onto a desktop

## Light

Since 2010, Intel has been working on new processors that use light beams instead of electric signals to move data. They call this process “silicon photonics.” Now the corporation says that it can send data at a speed of up to 100 gigabits per second along a cable with a diameter of five millimetres - compare that to the latest USB cables which do five gigabits per second, or network cables which can do 40.

Silicon photonics has the potential to liberate designers. Today’s servers are self-contained computers with processors, RAM and hard drives, arranged on racks in layers one on top of the other. But the speed and compactness of silicon photonics makes it possible to reorganise all this so you get a layer with processors, another with memory, and a third with storage.

There are also huge potential cost-savings for data centres. Just one of Intel’s new optical cables can replace over ten of the copper cables that connect servers stacked on the same rack. Since those cables reduce the flow of cooling air, getting rid of them will also help reduce energy consumption.

Amazon, Facebook and the US National Security Agency (NSA) are already queuing for the Intel technology. IBM, HP and some smaller companies are working on silicon photonics too.

## Quantum Computing

Taking advantage of quantum laws also holds great promise for faster, more efficient computing because atoms and photons can have more than two quantum states as opposed to the two states used in conventional computers. This means that calculations can be carried out faster. But quantum technology necessitates dealing with single photons that are well isolated from the environment - a major challenge. Up to now, although researchers have developed parts that produce or detect single photons, they were always located on separate chips. In order to make a practical quantum computer it is important that all the components be packed onto one chip. Now, Bristol University researchers have achieved this goal and, in January 2014, published research detailing their development of the most complex quantum computing system of its kind to date, and it can be made on mainstream electronics manufacturing equipment.

## Graphene

Graphene is a contender for ultra-efficient conductor of the future. The new material, made from a one-atom thick layer of graphite, was developed by Nobel Prize winners Andre Geim and Konstantin Novoselov from the University of Manchester.

It’s been suggested as a candidate for use in the new light-based processors, which Intel is currently developing. Among its attractions are the possibilities of faster data processing, the ability to absorb a broader range of wavelengths, cheaper development and less energy needs.

But there’s a drawback - graphene is not very good at absorbing light so a number of re-search teams are trying to crack this problem.

## Carbon Nanotubes

At the University of Southern California (USC), researchers are trying their luck with carbon nanotubes, tubular arrangements of carbon atoms akin to unimaginably thin wires. Thanks to their nanosize, computer operations that use less energy than silicon semi-conductors will be possible.

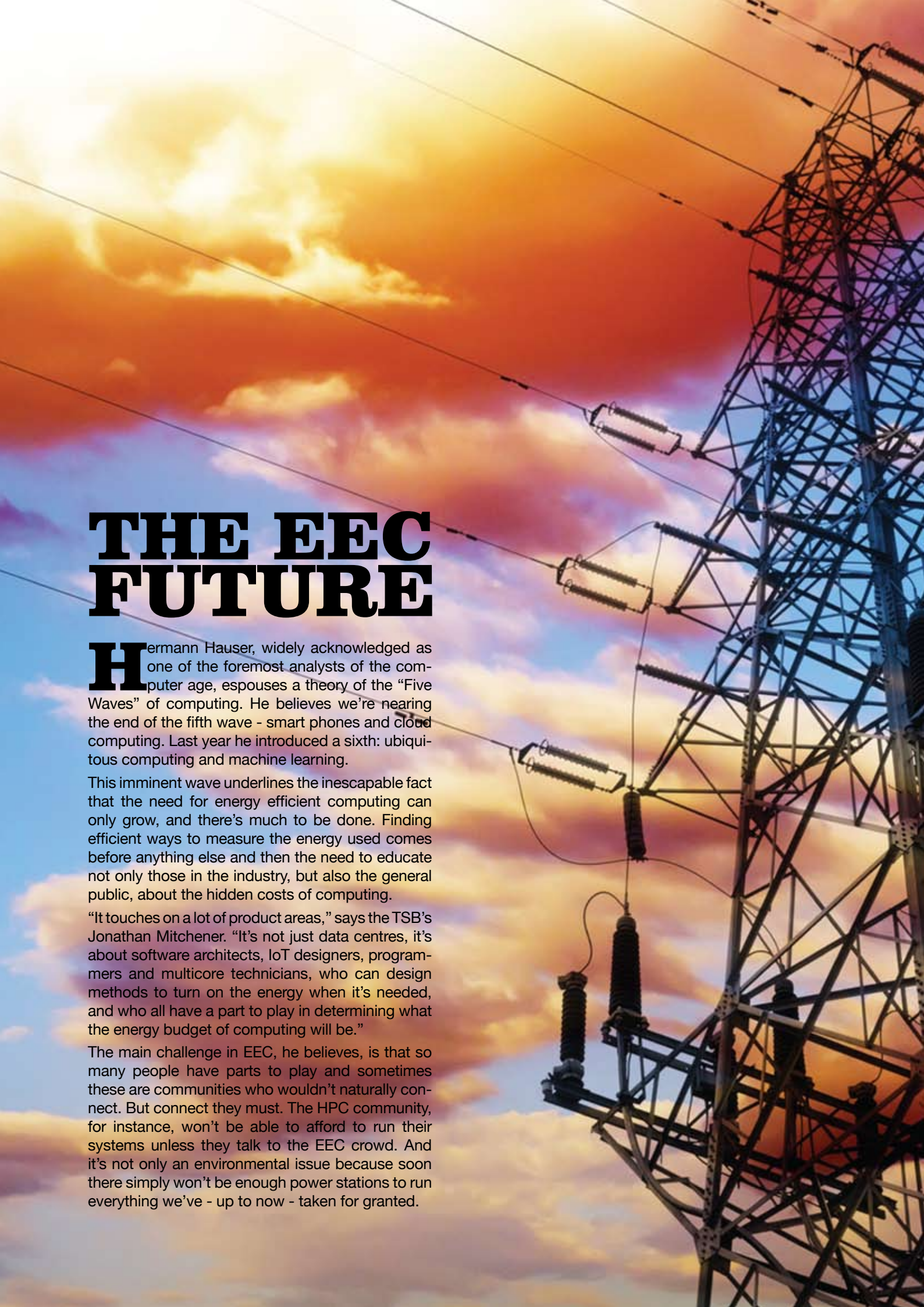
USC professor Chongwu Zhou says his team have almost realised the long cherished dream of nanotube developers - controlling atomic structure - and can now mass-produce the technology.

## Multiferroics

An emerging class of magnetic materials called multiferroics could make future devices far more energy-efficient by diminishing the wasted energy emitted as heat. The material has the effect of reducing the power consumed by logic devices - effectively a circuit on a computer chip dedicated to performing functions such as calculations.

UCLA researchers have demonstrated that using multiferroic materials to generate spin waves - of electrons - could reduce wasted heat and therefore increase power efficiency for processing by up to 1,000 times.

It’s helpful to imagine this as similar to an ocean wave, which keeps water molecules in the same place while the energy is carried through the water. Crucially, a multiferroic can be switched on or off by applying alternating voltage.



# THE EEC FUTURE

Hermann Hauser, widely acknowledged as one of the foremost analysts of the computer age, espouses a theory of the "Five Waves" of computing. He believes we're nearing the end of the fifth wave - smart phones and cloud computing. Last year he introduced a sixth: ubiquitous computing and machine learning.

This imminent wave underlines the inescapable fact that the need for energy efficient computing can only grow, and there's much to be done. Finding efficient ways to measure the energy used comes before anything else and then the need to educate not only those in the industry, but also the general public, about the hidden costs of computing.

"It touches on a lot of product areas," says the TSB's Jonathan Mitchener. "It's not just data centres, it's about software architects, IoT designers, programmers and multicore technicians, who can design methods to turn on the energy when it's needed, and who all have a part to play in determining what the energy budget of computing will be."

The main challenge in EEC, he believes, is that so many people have parts to play and sometimes these are communities who wouldn't naturally connect. But connect they must. The HPC community, for instance, won't be able to afford to run their systems unless they talk to the EEC crowd. And it's not only an environmental issue because soon there simply won't be enough power stations to run everything we've - up to now - taken for granted.