e science TALK

e-ScienceBriefings

Talking about e-science

Transferring Technology & Knowledge

In 1991, Tim Berners-Lee, a computer scientist working at CERN, gave us the World-Wide Web. This tool, designed to make navigating information easy, is perhaps the most famous example of technology transfer to come out of e-science. It has revolutionised how we communicate; the global economy; even how we live our lives. Its pervasiveness in society was even celebrated in the London 2012 Olympic Games opening ceremony.

Though predating the foundation of an official e-science funding programme in the UK by a number of years, the tremendous computational requirements of experiments at CERN tie the foundation of the Web to what has since become known as e-science. This somewhat well-worn example does not stand alone, however. From the developments in supercomputing in the 1980s, which gave us the multiple processor cores found in the advanced mobile electronics of today; to improved information storage for huge experiments, which have helped create 'the cloud', the backbone of the Web 2.0 economy – e-science impacts many areas of industry and commerce. Bioinformatics, a field largely made possible by technological advances in e-science, has revolutionised R&D in the pharmaceutical industry. Still other advances impact on earthquake prediction and oil exploration. But it's a two-way street for both ideas and people: computer games, now multibillion-dollar business, make use of the technologies and talent of the e-science world, just as games technologies feed back into e-science..

Supercomputer in the palm of your hand

Seymour Cray had a virtual monopoly on supercomputers for two decades, with his eponymous series of machines famed for their number-crunching prowess. His mantra, 'anyone can build a fast processor, the trick is to build a fast system', coined at a time when processing power was expensive, had held true for two decades. Eventually, his designs were superseded by a paradigm shift in computer architecture: multi-processor machines, which started being built in the 1980s. At the time, whether these clusters of processors counted as single supercomputers was debated, partly because the mechanics of getting the processors to work together had not been fully worked out.

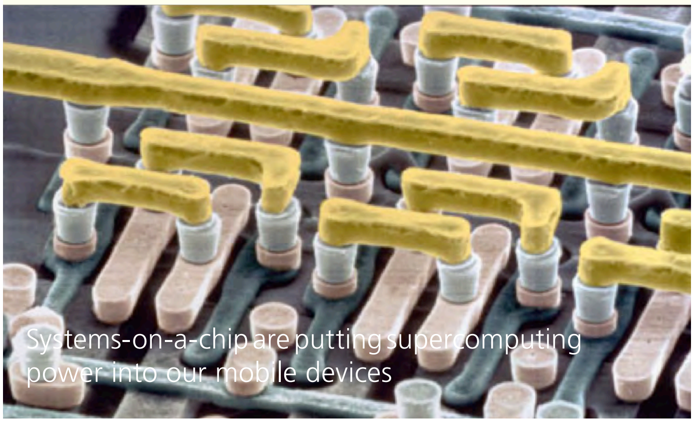
The message passing interface (MPI) standard, introduced in the early 1990s, allows processors to effectively 'talk to each other' and distribute computational load. The first implementation was worked on, among others, by Tony Hey, later head of the UK e-science programme. It has allowed software engineers to develop the concept by introducing bridging tools to make use of multiple processors, including commercial solutions such as Microsoft Compute cluster and e-infrastructure middleware such as gLite.

Chip manufacturers such as Intel, IBM and ARM have been switching to 'multicore' for several years, as heat restrictions have meant that cramming more transistors onto a die, or increasing the number of cycles per second, simply wouldn't work. More recently, system-on-chip (SoC) has meant that mobile technologies such as smart phones and tablets are virtual 'supercomputers in the hand', allowing such devices to run high-powered

The Web was celebrated in the 2012 Olympic opening ceremony

1/ Today, there would be no question – not only supercomputers, but even home computers, games consoles and mobile devices can contain more than one CPU core.

software. VisIVO, a project featured in the e-science Briefing on visualisation, has ported its astrophysics simulation system to Apple's iOS, for instance, and tablets are widely used in the medical sphere.


Systems-on-a-chip are putting supercomputing power into our mobile devices


Tony King-Smith, Imagination Technologies – *“Advancements in multicore and parallel processing technologies will continue to be the key driver in the mobile space as heterogeneous computing keeps blurring the lines between the PC and embedded market in terms of overall system performance. The graphics processor - the quintessence of what parallel computing is all about - has now become the most important part of a standard System-on-Chip. This has enabled smartphones and tablets to display resolutions higher than your regular TV or bring the quality of console gaming into the handheld realm while allowing consumer electronic products to get smaller yet faster with each generation. With new APIs, like OpenCL, that parallel processing capability will also enable more image recognition, photo enhancement and augmented reality applications to become mobile”*

Peopleware: Transferring Knowledge

Silvia Olabarriaga worked in industry before moving into academia, but is insistent that her peers, including those like her who work in e-science, have also learned to think like entrepreneurs: “To survive in academia, just doing good science isn't enough. You have to be able to lead people, to write grant proposals, and to manage the overall operations of a research group. All of those skills are useful to industry.”

Olabarriaga leads the e-Bioscience group at the Academic Medical Center in Amsterdam, having come from the software and computer graphics industries. She holds a doctorate in medical imaging. She is also passionate about peopleware – her idea that people are absolutely critical to the functioning of e-science; as important as hardware, middleware and software. “Communication is extremely important,” she says, “some people are more adept at solving complex issues, but you also need good communicators, just like in any organisation.”

‘Capacity-building’ is an often bandied-about term, but the skills gained by working in large, multidisciplinary and geographically dispersed teams are manifold, she says: “the business of the future will also be very distributed. It's going to be about teams of people with complementary skills and ideas working together. That's where the nice ‘apps’ are coming from.”

Olabarriaga thinks cloud computing is being adopted much more quickly by industry than grid (a model successful in academia where research institutes share computing power over a network) because it equates to a service model that business leaders understand. “In commercial settings you need a very clear business model. With cloud, the business model is clear.”


Silvia D Olabarriaga, Academic Medical Center, Amsterdam, The Netherlands – *“People who are exposed to big e-science projects undergo a change in the way they communicate. When recruiting for my research group, I look for people who have worked in these projects – they give you the mindset to see that, ‘OK – it's not not just my part of the programme, it's bigger than that…I have to understand the whole, and know how to communicate my part of the problem with others.’ To be able to go beyond the desktop is important because this is also what the industries of the future will need.”*

Cloud: Product to Service

Box, Dropbox, Google Docs and Apple iCloud: consumer platforms for cloud services. The key word here is ‘service’. In many cases, that service is storage, whether that's from online backup utilities such as Crashplan, or simply document sharing. But it doesn't stop at storage or even document and calendar synchronisation, although these are the services that, as consumers, many of us use day-to-day and see as being synonymous with cloud. Indeed, the Centre for Economics and Business Research in the UK estimated in 2011 that cloud computing could add €763bn to the European economy before 2015.

Simon Wardley, geneticist, former CEO for a division of Canon, and now researcher at the Leading Edge Forum, sees cloud as simply representing “the evolution of a bunch of activities from across the computing stack from products to services”[2]. While businesses might be told by that replacing their IT departments with services cloud might save them money, Wardley warns against this. Referencing William Stanley Jevons' consideration of more efficient steam engines in 1865, Wardley explains that businesses adapt to changing economics of operation. Cloud is more efficient, because it allows businesses to focus on their defining activity – “[but] does this reduce IT spend? No, because consumption of services goes up to compensate.” What cloud does offer, he suggests, is “greater agility. It's faster…and you can do more stuff”.

Yelp, Foursquare, and Hipstamatic all make use of Amazon's elastic cloud, otherwise known as EC2. Compared to the cost of setting up their own data centres, this has undoubtedly reduced the time needed to bring their services to market. There are also long-term benefits to using standard systems in a business model, akin to the adoption of standard screw fittings developed in the industrial revolution. If

---

[2] *OSCON 2010: Simon Wardley, «Situation Normal, Everything Must Change»*

Cloud fits business well because of its service model

competitors are using standard services, "ubiquity means that such activities will have diminishing strategic value, and become just a cost of doing business," explains Wardley. The competitive gap that opens up by not using standard services means that you lose focus on your defining activity.

e-science is feeding directly into this cloud-based computing ecosystem. HPC-in-the-cloud allows scientists to access virtual supercomputers running on Amazon's EC2. Just as for businesses, using standard services gives researchers a competitive advantage This is something that has particular resonance in the world of corporate R&D, but whether privately or publicly-funded, scientists and those working in science policy need to understand the entrepreneurial requirement for competing on a world stage. At a recent conference on e-infrastructure standards in Denmark, EGI Director Steven Newhouse framed the issue succinctly: "How many people here are flying home after this conference… [OK] – and how many people are building their own airport and aeroplanes to do so?" Concentrating on core activities doesn't only give private and public-sector researchers an economic competitive advantage, it also gives them more opportunities to innovate: "The main benefit of cloud is accelerated innovation," explains Wardley.

Tony Hey, **Corporate Vice President for Technical Computing, Microsoft –** *"Just as the scientific community sees an emerging data deluge that will fundamentally alter disciplines in areas throughout academic research, a wide range of industrial and commercial businesses are now beginning to recognize that data-driven business decision-making and analysis will be necessary to stay competitive. The magnitude of the data and the complexity of the analytical computations will drive both scientists and business analysts to explore the use of cloud computing, high performance computing systems, multicore processors, and parallel programming to manage the computing workload for these efforts."*

Computer Graphics: Innovation driving business

Computer graphics were once seen as a frippary by many computer scientists, but they have become increasingly important tools to communicate ideas and findings across disciplines (for example, see Briefing on Visualisations). e-science makes use of developments in computer graphics chips originally designed to be better for more immersive games. It's an example of commodity hardware, with prices driven down by economies of scale, feeding back into the e-Science ecosystem.

GPUs really began to be adopted by e-science with the advent of separate graphics 'cards', which were designed to meet demand for ever more realistic 3D worlds. The graphics co-processors in these cards were optimised to draw lots of polygons needed for 3D games very fast, making them highly suited to the parallel processing of mathematical algorithms. This makes them very useful for performing simulations of natural processes for e-science applications, such as molecular simulations for drug discovery in medicine and astrophysics simulations ever more realistic 3D worlds for computer games. This is called GPGPU (General Purpose computing on Graphics Processing Units).

Many of the techniques used for 3D visualisations in fields such as medicine were first developed for the entertainments industry. Realistic shading and depth of field can be crucial for clinicians deciding on a course of action when examining 3D recreations of internal organs from MRI scans, but the tools used to do this were perfected bringing characters like Woody from Disney's Toy Story to life.

Renderfarm.fi takes some of the ideas of desktop grids and volunteer computing and moves them into an entertainment sphere. The philosophy is that people want to contribute and like to help, whether that's towards solving scientific problems in citizen cyberscience projects such as SETI@Home or volunteering free computing cycles to creative works for entertainment purposes .

Renderfarm offers a way for small computer animation companies to get off the ground by providing computing cycles donated by volunteers at home.

GPU History

The first wave was discrete graphics processing units (GPUs) in some 8-bit and 16-bit 'home computers' in the early 1980s were produced to offload some of the burden of drawing objects on the screen from the CPU to a separate custom graphics chip (including the Atari 400/800 and Amiga[3]). These chips were very good for moving around 2D objects on-screen at a time when processor power was expensive, and 16-bit iterations were capable of broadcast video production and photorealistic static graphics. The second wave came with IBM PCs running DOS and Windows. Economies of scale associated with ubiquity meant their CPUs became much faster and cheaper, negating the advantage given by including graphics co-processors. Games such as Wolfenstein 3D and Doom, which ran on 386 and 486 processors,

[3] *Maher, J. 2012, The Future Was Here*

opened the floodgates in terms of what people expected from games, eventually leading towards high-powered dedicated graphics hardware, including graphics cards by ATI and NVIDIA. Today these are being used to build GPU supercomputers, such as EMERALD at the Rutherford Appleton Laboratory in the UK, which is used to model everything from pandemics to galactic simulations.

Ian Osborne, Knowledge Transfer Network – *"There are clearly big opportunities for those working in e-science to teach us in industry how to cope with large data – which has suddenly become a hot topic and which will only become more important with the 'internet of things'. There is suddenly a shared working space that industry and the e-science sphere can work together effectively in. Small companies can also use technologies developed in academia – properly licensed – to help fulfil their product pipeline without so much investment in R&D "*

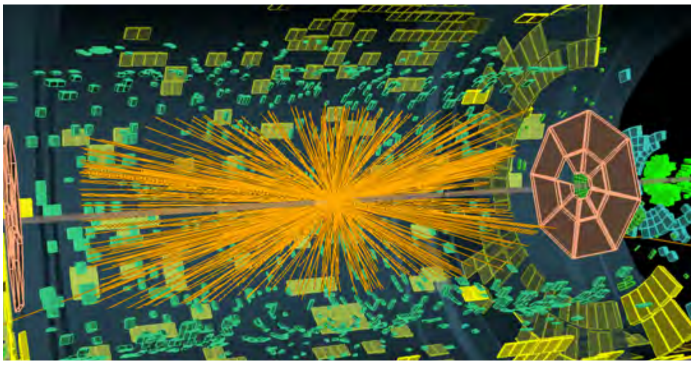
Intellectual Property in Tech Transfer

Intellectual property (IP) can be an issue for technologies coming out of e-science because it is largely embedded in academia. It is not always obvious who owns the intellectual property developed by an academic working in a university, although some institutions are now addressing this with agreements written into job contracts. Exactly who should own IP is still being debated, although a growing number of those working in science policy believe that it should be owned by the university rather than an individual professor. As engines of knowledge generation, its is argued, universities are well placed to properly licence technologies to industry and use the profits for future societal benefit. Indeed, one of the outcomes of individuals owning IP is that academic publishing could capitalise on it by asking them to sign away copyright on published research, which has hindered open access publishing. (This couldn't happen if the institute owned the IP).

The European Association of National Research Facilities (ERF) is a collection of European research facilities that have adopted an ERA open access policy, publishing their research for free for the benefit of those working outside of their walls. It includes members such as CERN, DESY and research centres of the UK Science and Technology Facilities Council

ATLAS: Not just measuring the massive

The technologies being developed for big experiments at CERN are, of course, widely disseminated thanks to an innovation in computer technology that went unpatented: the Web. But just as many of the advances for medicine came from the physical sciences (x-ray crystallography elucidating protein structures, radioisotopes from cyclotrons), experiments like ATLAS at CERN, which explores superheavy particles heavier than the Higgs boson and top quark, are themselves contributing to wider areas of human endeavour.



- Superconducting coils and cables are being explored in higher-density magnetic storage media, and could lead to even better power lines.
- Diamond-based sensors used to detect the fragments of particles yielded by high energy collisions could be used to monitor doses in hadron therapy, used to treat childhood cancers.
- 3D silicon detectors might be used for advanced medical imaging
- Pattern recognition technologies could be used for augmented reality, to allow automation of complex procedures by machines.
- Drift tubes can detect scattering of cosmic rays by dangerous materials, including those that could be used to make a bomb, hidden in containers.
- Medipix, a project based on ATLAS and other experiments at CERN, has been operating since the 1990s. The ability of detectors at CERN to count single photons with no noise is highly applicable to fields such as medical imaging, hence medipix.

Summary

Transfer or people, of ideas and of technologies continues to feed into and out of the e-science ecosystem. There are sometimes challenges in commercialising ideas coming out of academia, but scientists are becoming more adept at doing so as larger cultural changes take hold. Commercial models like cloud equally finding a place in public research settings. There are certainly exciting times ahead as academia and industry learn to work more closely to achieve their goals.

For more information:

- *ktn.innovateuk.org*
- *www.astp.net*
- *research.microsoft.com*
- *setiathome.berkeley.edu*
- *www.renderfarm.fi*
- *www.atlas.ch*
- *www.cloudbroker.com/*
- *www.stfc.ac.uk/About+STFC/51.aspx*

***EGI :** www.egi.eu*
***Real Time Monitor:** rtm.hep.ph.ic.ac.uk*
***iSGTW:** www.isgtw.org*
***e-ScienceTalk:** www.e-sciencetalk.org*
***email:** info@e-sciencetalk.org*

Scan this QR code into your smart phone for more on this e-ScienceBriefing

e-ScienceTalk is co-funded by the EC under FP7
INFSO-RI-260733