Governments and other funding agencies worldwide have recently been investing in the development of research- and production-quality e-infrastructures. This investment ultimately aims not only to increase competitiveness by supporting advanced computational platforms for research but also to drive societal change—for instance, to address the “digital divide.”

This isn’t a purely technological process. It requires that the necessary infrastructure be in place and that researchers of all disciplines, and ultimately all citizens, acquire the skills required to access the available services. Investments in e-infrastructure require adequate investments in education to allow for the full development and use of these technologies.

Why e-infrastructure education?

We can identify four primary reasons why it’s vital to develop policies to support e-infrastructure education and training:

- There’s a skills and knowledge shortage in business, government, academia, and society in general.
- It would help to optimize the use of e-infrastructure, guaranteeing return on investment.
- Such education and training would advance both industry and academia.
- It promotes the transition to knowledge-based economies (fulfilling Lisbon Strategy goals in the EU context, as we’ll discuss later).

Skills and knowledge shortage

A 2007 review of IT skills and careers in the UK revealed that skills shortages and gaps still plague computer science, and this has ripple effects in other subject areas and sectors. In 2005, IT companies pointed to a European-wide “skills crisis as a shortage of computer graduates and a retiring technical workforce threaten to bite IT departments by 2006” (see www.contractoruk.com/news/002206.html). In the US, Thomas Hoffman identified similar realities at the close of 2007, which he explains in Computerworld: “there is a distinct shortage of certain IT [skills], and that shortage seems to be growing” (see www.computerworld.com/action/article.do?command=viewArticleBasic&amp;articleId=308800). Concerned computer scientists have particularly pinpointed a lack of expertise in grid computing, explaining that, “Grid may be the liberal arts of computing. It requires knowledge about many IT disciplines, a flexible management approach and acceptance of new ideas. But resumes boasting grid-specific skills and accomplishments remain rare. Grid is not widely taught, and IT workers with hands-on experience in this young field are tough to find” (see www.techworld.com/opsys/features/index.cfm?featureid=924).

Guaranteeing return on investment

The second reason to develop education and training policy relates to the first. E-infrastructure technologies such as grid computing potentially risk poor return on investment if measures to support
usage of the infrastructures aren’t put in place. A compelling example, which applies to any research infrastructure, is that it takes years of training to get the best out of facilities. E-infrastructure’s complexity, novelty, and changing nature mean that there’s a high risk of under-utilization, or non-optimal exploitation, without adequate investment in education and training. Investment in e-infrastructure to date has provided a pervasive and dependable platform on which a relatively small proportion of experts can demonstrate the high value of the research and innovation it enables. Today’s challenge is to strengthen this platform so that the realization of e-infrastructure’s benefits becomes routine—that is, any researcher in any discipline routinely uses e-infrastructure resources as fluently as an artist uses a brush or an engineer uses differential equations. This requires two concurrent and coordinated advances:

- the educational progress identified here and
- a steady improvement in the facilities, tools, and ease of use of the pervasive e-infrastructure.

At present, the second branch of this strategic requirement is limited by the lack of sufficient skills across a sufficiently broad spectrum of society and academic disciplines to deliver the advances.

**Industrial drivers**

Both industry and academia benefit from e-infrastructure or grid computing outputs and applications, another key motivation for developing policies to progress grid education. Use of e-infrastructure has already become integral to finance and online businesses, a primary reason for their economic success. In finance, grid computing can solve problems associated with large and complex computations. Data centers at online companies such as Google and Amazon use forms of grid computing to manage the vast number of searches that users request on a daily basis. Advances in scientific and other knowledge as well as new technologies have also generated vast amounts of data that require proper management.

**Transition to knowledge-based economies**

In the European context, the Lisbon Strategy is the key action plan for continued European growth and development. Its primary aim is to assist European countries in their transition to knowledge-based economies. Because the European Commission has identified information and communications technology as a vital component of any economy, this strategy has stimulated the development of policy measures involving ICT provision. Jose Manuel Durao Barrosa, president of the European Commission, stressed the importance of ICT in achieving Lisbon strategy goals during a speech in March this year: “The Information and Communication Technologies industry is a motor for our economies and for our joint prosperity … The ICT industry therefore has a major role to play in the European economy of the 21st century, crucial to the success of the EU Lisbon Strategy for growth and jobs” (see www.ict-ile.eu/news-and-events/ict-industry-has-a-major-role-to-play-in-the-european-economy-of-the-21st-century).

E-infrastructure has significant value within ICT as a means to support and promote this transition to knowledge-based economies—not merely in the European context, but worldwide. It enhances and expands capacities and capabilities, providing a platform for innovation in industry and academic research, as we mentioned earlier. The promotion of education and training policy measures ensures that a skilled workforce will be available to maintain and develop e-infrastructure for this purpose.

**Why is distributed computing education special?**

Distributed computing teaching must address issues beyond the strictly technological. Almost by definition, distributed computing has two aspects that might not be encountered in basic IT teaching but are central to distributed systems: scale and collaboration.

Although both scale and collaboration in teaching environments are well-known issues when moving from prototypes to production, some central aspects of the effects of both become readily apparent only in real-life situations. For example, in terms of scale, the implications of providing global
operational support for a 24/7 service might not be apparent at local scales. Similarly, few students can easily grasp the implications of cross-organizational collaboration without being made specifically aware of them.

These specific issues are combined with the rapidly evolving technological landscape in this field. However, many educators in computing and related fields of engineering clearly share this issue.

A further combined effect of scale and collaboration that might not be readily apparent is that few departments—or even universities—contain the complete gamut of expertise required to support the emerging production infrastructures. Additionally, many teaching environments aren’t sufficiently supportive of the inherent interdisciplinary nature of teaching e-science (the research activities across disciplines that use the underlying e-infrastructures). Therefore, the teaching environment must support not only collaborative efforts in creating and running such infrastructures but also collaborative approaches to teaching.

**Finding answers**

The question then becomes, how can we help the widest range of domains and institutions to provide education in these subjects?

This question clearly has no one answer, but a set of interlocking and mutually dependent answers. Teaching environments must provide

1. materials to let educators rapidly take up this task (such as textbooks and repositories of teaching materials),
2. education for the educators to help them bootstrap in this area,
3. seed teaching infrastructures to be used initially for teaching and to become templates for organizations in developing their own infrastructures, and
4. a supportive policy framework that encourages cooperation and sharing (for example, in managing intellectual property rights or sharing teaching infrastructures).

In successive articles, we’ll seek to address these issues individually and examine some of the existing efforts to address them. Next month, we’ll present the experiences from an attempt to address issue 1 above: a community-supported summer school that prepares individuals to become advocates and educators within their own communities.

David Fergusson is deputy director of training, outreach, and education at the National e-Science Centre, Edinburgh. Contact him at dfmac@nesc.ac.uk.

Diego Romano is a researcher in computational and computer sciences at the University of Naples Federico II. Contact him at diego.romano@dma.unina.it.

Elizabeth van der Meer is education and training policy officer at the National e-Science Centre, Edinburgh. Contact her at evmeer@nesc.ac.uk.

Malcolm Atkinson is director of the e-Science Institute and e-science envoy at the National e-Science Centre, Edinburgh. Contact him at mpa@nesc.ac.uk.

**Cite this article:**

David Fergusson, Diego Romano, Elizabeth van der Meer, and Malcolm Atkinson, "Distributed Computing Education, Part 1: A Special Case?" *IEEE Distributed Systems Online*, vol. 9, no. 6, 2008, art. no. 0806-o6002.