Computer Challenges to emerge from e-Science

Malcolm Atkinson (NeSC), Jon Crowcroft (Cambridge), Carole Goble (Manchester), John Gurd (Manchester), Tom Rodden (Nottingham),Nigel Shadbolt (Southampton), Morris Sloman (Imperial College), Ian Sommerville (Lancaster), Tony Storey (IBM)

Executive Summary

The UK e-Science programme has initiated significant developments that allow networked grid technology to be used to form virtual colaboratories. The e-Science vision of a globally connected community has broader application than science with the same fundamental technologies being used to support eCommerce and e-Government.
The broadest vision of e-Science outlines a challenging research agenda for the computing community. New theories and models will be needed to provide a sound foundation for the tools used to specify, design, analyse and prove the properties of future grid technologies and applications. Fundamental research is needed in order to build a future e-Science infrastructure and to understand how to exploit the infrastructure to best effect.

A future infrastructure needs to be dynamic, universally available and promote trust. Realising this infrastructure will need new theories, methods and techniques to be developed and deployed. Although often not directly visible these fundamental infrastructure advances will provide the foundation for future scientific advancement, wealth generation and governance.

• We need to move from the current data focus to a *semantic grid* with facilities for the generation, support and traceability of knowledge.
• We need to make the infrastructure more available and more trusted by developing *trusted ubiquitous systems*.
• We need to reduce the cost of development by enabling the *rapid customised assembly of services.*
• We need to reduce the cost and complexity of managing the infrastructure by realising *autonomic computing systems*.

Exploiting a future infrastructure to best effect will require us to engage with users, understand and support their needs and the wider socio-economic issues. Essentially, the development of the infrastructure and its application needs to be informed by a human-centred approach to computing.
• We need to understand how we might best *support new forms of community* and how the infrastructure can be made available to users.
• We need to understand the *socio economic impact* of the infrastructure its implications for future public policy.

It is essential that the computing community tackle this research agenda collectively and in partnership with users. This will require a constructive approach in which the new foundational theories needed to underpin this work are developed in tandem with trials and experiments. The community needs to manage the process of engaging with users exploiting existing services while also realising future generations of these services. This will require:

- *An e-Science experimental platform*, isolated from the current service facilities, to support research on fundamental services and computing infrastructure.
- *An archival repository* that will allow experimental results to be shared across the computing community and made more generally available.

Research results to emerge will improve the exploitation of knowledge, increase the availability and reliability of on-line facilities, reduce the cost of service development and management and allow new arrangements of work across distributed organisations.

Introduction

The UK e-Science programme aims to develop support for large-scale science through distributed global collaborations. A significant focus of the e-Science programme is the development of a communication and computational infrastructure to underpin the work of scientists termed the Grid. The Grid promotes the rapid formation of virtual colaboratories allowing scientists to work together and share resources irrespective of the location of the scientists or the resources they are using.

The broadest vision of e-science is inclusive in nature and seeks to allow those in sophisticated research labs to work together with scientists in the field. For example, an environmental field officer monitoring plant growth in a remote jungle can be connected through wireless devices to scientist in the lab with access to sophisticated climate modelling software. Developing the technology to realise this vision represents an adventurous research agenda for IT. It also has implications beyond the support of the scientist. We can envisage maintenance engineers being seamlessly connected to those with access to large-scale simulations, or the salesperson having access to the rest of a virtual organisation.

The computer science research underpinning the e-science programme will also be relevant to e-Business, e-Commerce, e-Medicine and e-Government as computers increasingly become part of our everyday life. The Computing Community needs to respond to this trend and undertake the fundamental research needed to underpin the development and deployment of ubiquitous systems that people can trust and rely upon. To take just one example, such a programme of research could provide the infrastructure support for future large-scale continuous monitoring of medical data via implanted or on-body sensors and wireless transmission to monitoring bureaus. This would provide medical researchers with the data required to implement automated early detection and notification of adverse reactions to drugs, the detection of a whole range of problems such as adverse cardiac events, blood chemistry imbalance and so on. Another potential ubiquitous application is the mobile monitoring of traffic conditions via sensors and wireless transmitters on all new cars.

As e-Science, e-Business and e-Commerce mature we will see the emergence of a digital infrastructure that is pervasive in nature and accessible through a diverse collection of devices. A future grid will provide a digital fabric to support a broad range of activities and will be open, flexible and heterogeneous in nature. Essentially IT will disappear into the background and those involved in these activities will focus on the work at hand rather than the technology they use.

Realising this future vision questions some of the fundamental assumptions about the nature of computing and requires us to reassess many of the existing theories, concepts and practices at the core of computing. To deliver the e-science vision requires us to answer a number of important questions.

- How can we develop an infrastructure that promotes and supports the management and traceability of knowledge?
- How can we develop an infrastructure that lets people trust ubiquitous computing systems?
- How can we develop techniques to rapidly design, configure and evolve these systems?
- How can we deal with the inherent complexity of these systems and develop infrastructures that help us manage them?
- How can we make these systems usable and what new forms of organisation can be supported?
- How can we understand the broader social and economic impact of these technologies?

Answering these questions outlines a far-reaching research agenda that will require significant progress across many aspects of computing. Issues of heterogeneity, scale, mobility, dependability and privacy pervade all aspects of the research agenda outlined in this document.

New theories and models will be needed to provide a sound foundation for the tools needed to specify, design and analyse these dynamic, self-organising ubiquitous systems. This research will have *significant benefits* including:

- Improving the ability for organisations to manage knowledge including both scientific and business knowledge.
- Allowing critical applications to confidently exploit the pervasive availability of a trusted infrastructure.
- Reduce the development cost of applications by providing rapid assembly techniques from well-known services.
- Reduce the cost of managing both applications and infrastructure by developing dynamic self-managing techniques.
- Improve cooperation and communication across a distributed organisation and allow new arrangements of work.
- Inform UK IT policy and strategy by considering the socio-economic impact of a pervasive e-Science infrastructure.

The means of tackling this research is as important as the topics themselves. It is essential that the computing community address these problems collectively and in partnership with eventual users. This will require a constructive approach in which new theories and concepts are being developed in tandem with trials and experiments. The computing community also needs to manage the process of engaging with users exploiting existing services while also realising future generations of these services. This will require:

- **An e-Science experimental platform**, isolated from the current service facilities, is needed to support research on network and computing infrastructure.
- **An archival repository** is necessary so that experimental results can be shared across the computing community and made more generally available.

Fundamental research will need to be undertaken into developing the core science needed to **build a future infrastructure** that meets the demands of e-science. This will need to be complemented by research that focuses on how best to **use the infrastructure** in practice. In the rest of this document we outline the research issues that need to be tackled to realise our ambitious and broad vision of e-science.

Building a Future Infrastructure

e-Science, e-Commerce and e-Government are all predicated on the existence of a digital infrastructure that is universally available and promotes trust. The presumption is that the complexity of this infrastructure will essentially be invisible at the point of use. Making complexity disappear requires considerable work by the computing community and will need new theories, methods and techniques to be developed and deployed. Although often not directly visible these infrastructure advances will provide the foundation for future scientific advancement, wealth generation and governance.

Developing a Semantic Grid

Lying at the core of scientific development is the discovery of new knowledge. Indeed the generation, support and maintenance of knowledge provides the foundation of most scientific endeavour. The rapid increase in the volume and variety of data inherent within e-Science and mirrored by e-Commerce and e-Government means that any supporting infrastructure must provide a set of semantic services. These core services must be able to equip data with meaning and generate a surrounding semantic context in which data can be meaningfully interpreted. Fundamental research on knowledge systems and services is needed to allow us to move from the current data centric view supporting the grid to a semantic grid with a powerful set of knowledge services. Research that needs to take place includes work on:

- Techniques to manage the traceability and integrity of information and trace provenance all the way from initial data through information to knowledge structures.
- New theories and techniques to allow tolerant, safe and scalable reasoning over uncertain and incomplete knowledge where it embraces data, metadata and knowledge activities.
- Tools, methods and techniques to support the design, development and deployment of large-scale ontologies.

- Support for collaboration and sharing across different knowledge repositories at varying scales including working across personalised knowledge structures and larger organisational and disciplinary structures.
- Support for semantic directed knowledge discovery to complement data mining methods
- The development of lightweight and incidental knowledge capture techniques.
- Development of network based reasoning and decision support services that can be tailored to meet the demands of different domains and users.

Trusted Ubiquitous Systems

In the future a pervasive digital infrastructure will allow computing facilities to be always available via a heterogeneous range of devices. The infrastructure will seamlessly combine reliable high performance computing and communication networks and variable low performance embedded or portable devices with integrated wireless facilities. This will connect scientists in resource rich labs to field scientists with limited resources or to remote automated experiments to form a distributed ubiquitous system. The supporting infrastructure will need to be open to all legitimate users, promote heterogeneity and be extremely flexible. Resources will vary in their availability, their certification of quality and their reliability. However, services need to be trusted at the point of use. Fundamental computing research is needed to enable the realisation of trusted ubiquitous systems formed from the coalition of these potentially uncertain components. Research that needs to take place includes work on:
- New theories and techniques to model, specify and analyse trust in distributed ubiquitous systems where partial and uncertain knowledge pervades all aspects of the system.
- New quality of service and service-based models for ubiquitous systems that embrace issues of dependability, security, privacy and risk, and to support the formation of service-level agreements between systems across ubiquitous environments.
- New design guidelines and practices to enable the development of reusable trusted components that support independent testing and capability accreditation.
- New understanding of the practical engineering trade-offs required to realise trusted ubiquitous systems.
- Theories and techniques to address the combined requirements of open and dynamic environments, scale and mobility.

Rapid Customised Assembly of Services

Matching the facilities provided by a flexible infrastructure to the demands of a particular situation lies at the core of successful computational support. It is essential that users have facilities that enable the rapid assembly of a number of services to allow them to undertake their work. Currently, the cost of assembly is prohibitively high requiring specialised computing expertise and often taking considerable time. In the case of e-Science, where many experiments will be entirely digital, the role of assembly is crucial. The cost of assembly must be sufficiently low that the future user will focus on the work at hand rather than the services and infrastructure supporting it. The goal is to develop easy to use techniques for rapidly assembling distributed components to form a trusted service that can still be analysed to determine that it meets its requirements. This achievement would be of widespread value in e-industry and e-government. Research that needs to take place includes work on:

- New theories and techniques to describe and reason about the semantics and behaviour of services and the compositional effects of putting services together.
- Techniques for representing the capability of services and the quality of service they provide.
- New tools to support the discovery, composition and use of services based on high-level descriptions of requirements.
- Orchestration systems that manage the coordination, scheduling and recovery strategies for composed systems.
- Agent and service representations that promote adaptability and the emergent, opportunistic and implicit arrangement of services.
- Techniques to supported directed automatic composition, decomposition and recomposition of services.
- Knowledge structures that provide representations of cooperative activities, workflow and resources thus informing service assembly.
- Techniques for the management and version control of services version and lifetime management.
- New models of both the markets and charging structures that are appropriate to the ownership of services.

Autonomic Computing

Complex assemblies of open systems lie at the heart of future support for e-Science, e-Commerce and e-Government. As more of our activities depend on the use of digital services the very complexity of these systems threatens progress. We must develop a supporting open digital infrastructure that is able to handle rapid and potentially radical changes with minimum systems administration. Our current approach to system building and configuration is overly dependant on human intervention and simply does not scale. We must shift to self-configuring systems that are able to act autonomously and adapt to changes in application or user needs. This requires a shift toward autonomic computing[1] where computing systems are self-managed with a minimum of human interference. Key challenges that need to be addressed include:

- New theories and techniques to analyse, describe and reason about adaptive systems that are self-organising, self managing and self decommissioning.
- New models and metaphors to allow semi-autonomous systems to be managed through a combination of specified policies, negotiated agreements between users, services, software agents and regulatory structures.
- Techniques to allow interoperability across and between different autonomous domains and to reason about the combination of different domains.

---

[1] The term derives from the body's autonomic nervous system, which controls key functions without conscious awareness or involvement

- Techniques to model and measure performance and ensure quality of services when they depend on autonomic computational structures.
- Techniques to capture and represent context relating to location, device capabilities, history of the computation, user activity, ambient environment and to build applications which can automatically adapt to their context.

Putting the Infrastructure to work

A future ubiquitous computing infrastructure will be part of our everyday life and interactions and needs to be informed by a human-centred approach to computing. We need to understand the impact this will have at work, in the home and how it will effect public interaction with services such as government, education and healthcare. Current approaches to commerce, scientific research and entertainment will have to adapt to support mobile users interacting with a ubiquitous infrastructure. We also need to understand the economic model for funding and supporting this infrastructure. This raises a set of challenges for the computing community to address in partnership with other disciplines.

Support for New Forms of Community

Scientific community lies at the core of scientific investigation. As new challenges emerge science has relied on the formation of new communities to address them. The facilities inherent within a digital infrastructure allow communities to form more rapidly and offer the possibility of new forms of community. Facilities are needed to easily form dynamic and transient collaborative groups for applications such as scientific cooperation, virtual businesses or patient health-care. Key challenges that need to be addressed include:
- New theories to represent and understand transient organisations and dynamic, potentially mobile work arrangements including virtual teams, collaboratories and virtual organisations
- New techniques and metaphors to understand and support communities that combine both human and computational agents including new models of role and new forms of augmented cognition.
- Techniques to manage the work processes associated with a wide range of dynamic collaborative communities.
- The development of human factors techniques for the design, development and assessment of systems to support distributed, transient and mobile communities of users.
- New forms of support for community memories that ensure the consolidation and preservation of the content within these community memories along with its originating provenance.
- New forms of naturalistic and multimodal interface to support cooperative interaction whether it takes place across a distance supported by a variety of different interaction devices or in a co-located manner.
- Personalisation techniques that allow users to tailor the services and infrastructure to meet their individual needs and current activity while still maintaining an active role in the community they belong to.

- Techniques to support and promote interaction across and between communities including the discovery of new communities.
- Techniques to support introduction and induction into different communities ranging from ad-hoc communities to highly regulated and managed communities.

Socio Economic Impact

The deployment and use of a dynamic and open infrastructure that pervades all aspects of work will have significant social and economic impact. Understanding this impact and responding to the challenges and opportunities it presents will be essential for the future social and economic health of the UK. Understanding these potential impacts will involve a broad range of disciplines (e.g. economists, social scientists and philosophers) and will ask questions of future regulatory bodies and public policy. Key challenges that need to be addressed include:

- The economic and legal issues to emerge from combining open shared systems and closed owned systems to form different partnerships arrangements including public private partnerships.
- New models for the valuation of computational knowledge assets and content.
- New forms of business model and charging structures suitable for ubiquitous, multi-service environments.
- Techniques and policies to encourage and foster new forms of entrepreneurship made possible by the emergence of ubiquitous systems and services.
- Assessing the impact of ubiquitous computing and communication technology on the work practice, the home environment and entertainment.
- The issues of ethics, privacy, liability, risk and responsibility and the impact on future public policies, when our everyday activities are based on the use of an assembly of services supported by an ubiquitous infrastructure able to monitor movement and activity.
- Exploiting a distributed pervasive environment to support and promote public awareness of science and interactive government.
- Approaches and organisational structures that allow multinational collaboration to be simultaneously encouraged and regulated.
- Cost benefit analysis techniques for the use and deployment of ubiquitous technology to support different communities.