BAEgrid: From e-Science to e-Engineering
A progress report on Grids in the Aerospace Industry.

Alan Gould, Sean Barker, Elizabeth Carver, David Golby, Michael Turner

BAE SYSTEMS Advanced Technology Centre
Sowerby Building,
FPC267, PO Box 5, Filton,
BRISTOL BS34 7QW. UK.
Email: Alan.Gould@baesystems.com

The BAE SYSTEMS Advanced Technology Centre is currently constructing a prototype Grid infrastructure - “BAEgrid”. BAEgrid is designed as an in-house “laboratory” in which we can exercise grid concepts, capabilities and tools to develop and evaluate a range of Grid-enabled business scenarios. This is facilitated by our close relationship with the world-leading UK e-Science programme, and is being implemented in collaboration with several major hardware and software vendors. Grid technology in its current form is ready to be deployed in areas of compute-intensive activity and can demonstrably improve efficiency of deployed IT assets within a single organisational domain. However, the big pay-off from Grid is in the collaborative “Virtual Organisation” area. Here we are engaging with several e-Science projects to test existing capabilities, and gain insight into the requirements for future work. This paper gives a brief overview of the current status of the BAEgrid, and some of the collaborations now running. Our experience demonstrates the importance of the involvement of the end-user community in the development of Grid. We identify three topics as crucial to the success of Grid in a business environment: *security* (the fundamental enabler), *semantics* (the key to interoperability at the applications level) and *human factors* (central to effective collaboration).

Introduction

Once computers reached a certain cost and power in the early 1980’s engineers switched en-masse to a new way of working, and computing in engineering became routine within a few years. More recently, the commercial benefits of the Internet were viewed with healthy skepticism. Once protocols were agreed and adopted certain applications (particularly email and the web) lead to a huge uptake. The impact of those technologies on BAE SYSTEMS has been obvious, and it would be inconceivable to do business today without Intranet or Internet capabilities.

It is now accepted that Grid protocols and applications are becoming sufficiently capable and heterogeneous (in the academic domain, at least) that their persistence is assured[1]. Of course, like computers and the Internet before, it is difficult to predict the timing and eventual impact of Grid. There is a growing expectation that the impact will be large and soon!

Having established that the Grid concept is taking shape, and has a good chance of success, we need to examine what the role of an engineering “end-user” should be. Even if processors, networks and operating systems are delivered “off the shelf” they do not constitute a capability that can immediately help our business. We still need to consider the applications that we wish to run on the systems, the policies under which they are operated and the degree to which we wish to rely on others (e.g. suppliers) to meet our requirements. The interactions of our applications, processes and workflows with the Grid standards and the emerging Grid-services architecture are not necessarily straightforward to implement.

It is therefore essential that we develop an understanding of Grid – both from the technical and socio-political viewpoints. From this understanding we can make informed choices as to how we adopt Grid concepts, which software and hardware solutions best fit our needs, and which areas of the business will benefit most, and when. Based on past experience it is important that we are able to define our own requirements and then deploy a combination of internal effort, contractors and out-sourcing to meet those requirements.

The BAEgrid

Having determined that the growth of "Grid" on a global scale merits investigation (and spurred on by the launch of the UK e-Science programme), the Company's Advanced Technology Centre has been tasked with establishing a prototype Grid-enabled infrastructure. This has the aim of establishing the business case for Grid deployment across a wide range of corporate activity. The infrastructure on which our collaborative research programme is built is known as "BAEgrid". The work on the BAEgrid is being performed in each of the four layers of our simplified Grid architecture (figure 1).

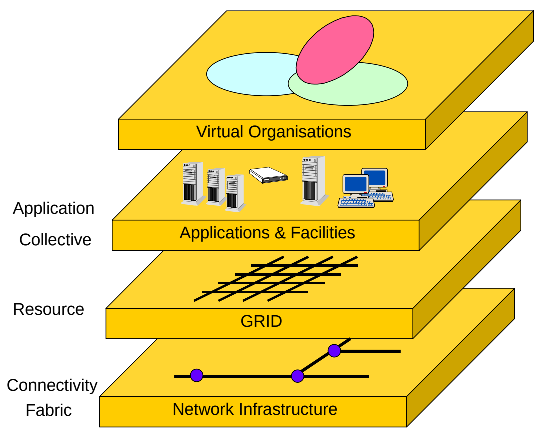

Figure 1 A simplified layer representation. The corresponding Grid architecture elements[11] are indicated on the left.

Network layer:
The network layer consists entirely of existing computing and connectivity resources. No specialist hardware has been purchased for the BAEgrid initiative - an important point to stress: since "Grid" is built on top of existing protocols it requires no additional or specialist equipment. In fact, the computers within the company are largely operating with Windows NT (or successor O/S types) and we currently maintain only a small number of Unix/Linux machines as compute facilities and to enable Globus deployment.

Grid layer:
The BAEgrid is constructed using the Globus Toolkit[2] (version 2.4). It uses components of the toolkit such as GSI, GRAM, GridFtp and MDS as well as other Globus based tools like SimpleCA and MyProxy[3]. Globus Toolkit version 3.0 is currently being installed and some applications are being ported to the gridService architecture. In addition to this we have made use of the freeware AccessGrid[4] system (version 2.0), and used the Condor[5] software to set up the first version of our desktop aggregation pool. We are keen to utilise commercial solutions where possible, and have recently signed partnering agreements with Entropia[6] and Platform[7]. Some additional tools have been written in-house to support deployment of a number of applications. These include portals that provide access to grid resources using tools such as the Globus Java CoG, in-house Globus interfaces and wrappers and toolkits such as the MatLab interface developed in the GEODISE[8] project.

Applications layer:
A schematic representation of the BAEgrid is shown in Figure 2. This is intended to illustrate the organisations both within and outside the company that are participating in one or more trial "virtual organisations". (See below). Of significant importance is the separation of the "Intra-grid" deployed on the corporate Intranet and the "Extra-grid" that utilises a separate link to the Internet. This is discussed further in the Security section (below).

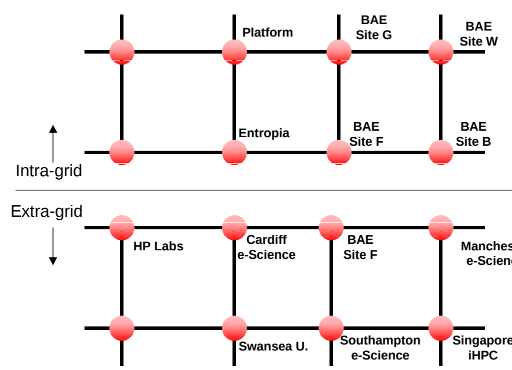

Figure 2 A topological view of the BAEgrid showing some of the members of our Intranet and Internet based Virtual Organisations.

The technology deployed at the "nodes" of the BAEgrid includes the following:

- Processing power in the form of multi-processor clusters.

- A Condor pool made up of 50+ Desktop processors, stand alone Unix machines and a MPI enabled alpha cluster. A condor view server is used to monitor the pool usage.
- A “Datagrid" space of around 100GB configured, built using the Globus toolkit and interfaced with analysis codes.
- AccessGrid facilities (in the process of being deployed) including small meeting room and PIG facilities.
- Software applications in the form of a growing list of engineering applications and computationally expensive analysis codes.
- Applets and portal software to permit access from lightweight clients through browsers.
- Experimental measurement equipment. BAE SYSTEMS’ proprietary pressure sensitive paint system[9] is used for data gathering in wind-tunnel experiments, and is being configured to operate remotely across the Grid.

Virtual Organisation layer:

Our research is centred of the deployment of several trial Virtual Organisations (VO’s). Foster et al[10] define Virtual Organisations as a set of individuals and/or institutions whose relationship is defined by sharing rules. In practice we have found it necessary to be more specific in defining the boundaries of a VO. Our VO’s are constituted around a specific technical goal. This goal is owned by one or more of the VO participants, and only resources, organisations and people that are engaged towards meeting that goal are included. This also helps to define and control the duration of the VO, and makes deployment and management of shared resources easier. These organisations are known as Virtual Organisations formed by Goal-Oriented Networks (VOGONs).

The Research Programme

A traditional view of the growth of grids is an assumption that there will be a progression from “cluster grids”, through “enterprise grids” to “global grids”. This probably makes sense if the goal is to aggregate computing power, but as we point out here, it is the collaborative capabilities of Grid that can pay the greatest dividends. Our involvement in external collaborative grids is at least equal to (if not ahead of) our internal cluster grid capabilities. In fact it is the joining of these two components of the BAEgrid that possibly presents the greatest challenge.

The current internal research programme is split into several distinct areas. “*Infrastructure*” and “*Distributed Problem Solving*” deal with the deployment of Grid and its immediate application, “*Security*” and “*Semantics and Standards*” cover areas we see as among the most critical for further development of Grid, and the “*Socio-technical*” – up to now a severely under-resourced yet crucial component.

Infrastructure

Our work in this area has been restricted to the collection and deployment of existing tools and processes – such as Globus, Condor and AccessGrid. We are also interested in testing and evaluating the state-of-the-art in commercially produced software. The main goals here are to develop familiarity with Grid capabilities, provide in-house best-practice, assist with policy development and provide expertise and connectivity to support a number of external collaborations (including those within the e-Science programme).

Distributed problem solving

Because of its similarities with “big science” we see the area of multi-disciplinary problem solving as one of the first areas of Grid deployment within the Company. We are therefore looking to configure engineering workflows that involve compute-intensive applications together with the associated processing power such that they can be operated more efficiently. In the current economic climate there is ever-increasing pressure to reduce fixed costs associated with IT. Grid-like access to distributed resources, together with capability-as-service architectures, offer new ways of delivering these savings. In the longer term it is clear that a Grid-services approach to multi-disciplinary analysis will lead to new ways of combining currently disparate capabilities. The management of, and access to, distributed data, information and knowledge sources could have a huge impact on the decision-making processes deployed during a product development lifecycle. The challenges are therefore not just in the areas of connectivity and analysis delivery, but also in the development of new tools and processes to fully exploit the new distributed systems.

Security

Security covers a huge range of issues that go right to the heart of a Grid system. One of the reasons why Grid has been so successful is the awareness of the need for a security infrastructure in the early standards and implementations. However, it has become clear that these measures are inadequate to support the deployment of even a moderately complex academic VO. The UK e-Science security task force[11] has identified a dozen different aspects of security (from authentication and authorisation, through to assurance and usability), and only some of these are handled adequately by the existing middleware.

The BAEgrid deployment uses the full range of existing corporate and national requirements and procedures. Using Grid within this environment is clearly possible but results in a loss of "performance" that all but the largest VO's cannot tolerate. Resources from multiple sites required by a rapidly initiated (or dynamically changing VO) cannot be configured or protected quickly enough using current procedures since these procedures are largely based on specifically negotiated, written agreements. The current "split" that is shown in the BAEgrid topology (figure 2) is a good example of this issue. Only resources (hardware, software etc) that are specifically licensed and authorised can be accessed from outside the organisation – and those resources cannot subsequently be re-deployed to other networks easily.

Some pertinent examples of the problems encountered when deploying Grid middleware in conjunction with firewalls are given in Surridge's Rough Guide to Grid Security[12].

Grid cannot work effectively under existing commercial-strength security systems and procedures. It is essential that additional tools and methods are developed very quickly to remove a major barrier to the future deployment of Grid-enabled VO's.

Standards and Semantics

Aircraft and their support environment increasingly rely on embedded computer systems, not only for flying the aircraft, but also for support systems such as structural health monitoring, engine usage recording etc. At the same time the design of the platforms and systems themselves is a highly collaborative endeavour requiring significant levels of integration. It is impossible to connect different systems together unless a data interface for the embedded software has been agreed in advance. This defines the format of the information exchanged, but assumes that the meaning of each of the data elements is hard coded into the applications. Defining and implementing the interface generally takes months or even years. The development of semantic standards such as STEP has greatly improved interoperability – in the design office at least. STEP is used to define data exchanges, with the meaning of the information again hard-coded in the interface. It does not cover all areas of interest, nor – being an international standard – can it keep up with rapid technical developments. We are working to develop systems that will automatically map parts of existing data models to each other, and interpret them consistently. This becomes truly "Grid-enabled" when the data model mapping techniques are sufficiently robust that automatic negotiation can take place and systems can rapidly establish grounds for communication of understanding with a high degree of automation.

Socio-technical

We will not succeed as a global company unless we create effective virtual organisations who are able to share knowledge and information and to work together collaboratively in the context of risk sharing partnerships. The business environment means that we are exposed to situations where we are both partnering and competing with the same companies at the same time. It takes a more sophisticated shared working environment to enable supply chain partners to work collaboratively with us in a secure way. A high degree of partnership maturity is needed to enable these relationships to be successful.

As a company we have invested in connectivity throughout the company and within projects so that shared working environments can be enabled. However *connectivity* does not equal *collaboration* and a concentration on the technological aspects alone will not achieve success. It is very important to adopt an integrated approach that considers the people issues, process requirements and technological aspects together in order to deliver effective knowledge sharing and knowledge management.

A recent industry survey[13] supports the view that some of the major barriers to the deployment of Grid concepts and solutions are organisational and political, not technological.

Grid architectures promise many benefits but it is important that we are ready as a company to reap those benefits as and when the technology is mature. Looking at the human factors of collaboration is crucial in assuring the success of Grid. Therefore the human factors component of our Grid research portfolio aims to address issues of collaborative technology[14] as well as communication and knowledge-sharing issues within multidisciplinary teams.

Current Application Examples

Three examples from our portfolio of operational research partnerships are now given to illustrate current capabilities and issues.

VOGON – Analysis service

An engineering function within one of the company's business units has a requirement to use full-physics fluid dynamics analysis during a particular part of a product development lifecycle. This capability, together with a substantial body of knowledge and experience in the use of such tools, exists within the wider company, but is not easily accessible. A licence for the relevant software, and an appropriately sized computing resource must be purchased, installed and maintained locally – even if the capability is not required continually.

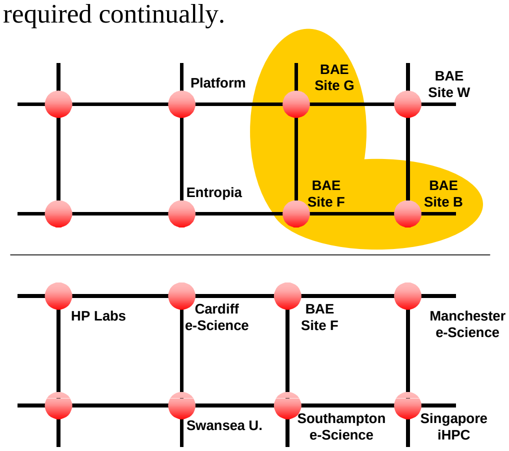

Figure 3 The Analysis service VOGON

Using existing Intranet connectivity we have established a Globus-enabled process whereby engineers can access a fluid dynamics analysis code, and the associated computing power. The problem to be solved is set up within an existing problem-solving environment, and the input files are transferred to the remote "service" using GridFTP authenticated by the GSI. The job(s) are then scheduled using Condor and results returned via the same route. One of the most important aspects of this system is the lightweight client. The client interface is based on the Globus Java CoG and can be accessed through any Java-enabled browser on any O/S – providing valid credentials can be presented. Another aspect concerns the "service" provision. It is now much easier to access support and information on the use of the code since access to the problem in hand can be organised much more easily by a previously remote site. Future plans include the use of "knowledge-enabled" capabilities such as those being scoped by the GEODISE project.

VOGON – GECEM

GECEM[15] is an e-Science project that is concerned with Computational Electromagnetic analysis for product performance prediction – where the component technologies needed to solve very large problems are owned by different organisations. The partners are BAE SYSTEMS, Cardiff and Swansea universities, HPLabs and the iHPC in Singapore. This provides a trial system whereby different capabilities (mesh generators, solvers, compute resources, visualisation and VR facilities) can be supplied and co-ordinated by a "customer". Important aspects of the proposed research include the deployment and use of Grid-based mesh generation and manipulation services, the secure remote execution of computational electromagnetic simulations on the various hardware resources available at the Grid nodes (including that of the Singapore Institute of High Performance Computing), and collaborative visualisation and analysis of both meshes and the simulation results by geographically dispersed participants. A critical issue in the proposed work, and in the use of the Grid in general, is the ability to make effective use of the distributed resources in the face of dynamically changing network bandwidth constraints.

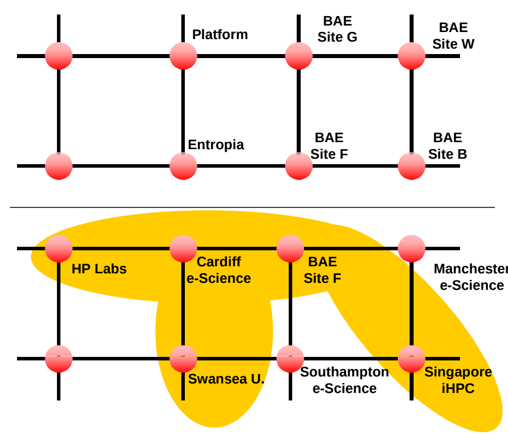

Figure 4 The GECEM VOGON

VOGON – GEWiTTS

GEWiTTS (Grid Enabled Wind Tunnel Test Services) is another e-Science project and is a collaboration between BAE SYSTEMS, Manchester University and the Aircraft Research Association. As before we are interested in scoping the delivery of capability as a service, but in this case we are attempting to “Grid-enable” a physical test facility (in the form of an instrumented wind-tunnel). The orchestration of the necessary resources takes on a slightly different flavour in this instance. We are now dealing with new issues such as safety during remote access to steerable experiments, and the definition of the appropriate architecture for the apparatus (how much processing is to be done at the server or client locations, for example).

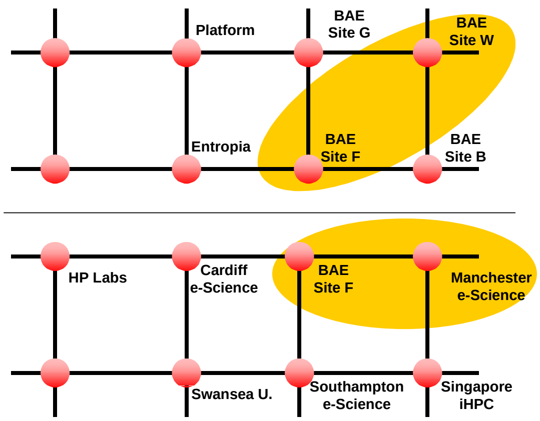

Figure 5 The GEWiTTS VOGON

Forward Look

It is clear from our early exposure to Grid technology that there is a lot of useful capability offered by existing open-source and proprietary solutions. This capability is mainly centred on the aggregation and distribution of IT resources. Where our requirements match those of “big science” – such as computationally intensive processes – Grid technology can be (and is) declared “operational”[16].

In a very simple interpretation of the "Grid" paradigm, there are two distinct issues:

- The use of Grid concepts, architectures, standards and technologies to "Grid-enable" resources (hardware, software, facilities, people etc).
- The operation of Virtual Organisations.

The current generation of Grid technology is almost entirely related to the issue of enabling existing resources – from the networking initiatives such as GTRN, SuperJanet4, GEANT, through the middleware development in Globus, Unicore, Condor to applications in areas of "big science": CERN LHC, NASA IPG, Bio-informatics etc.

Grid for "big science" permits only "symmetrical" VO's. These VO's are constituted as coalitions of equals where each participant has a similar degree of ownership over the objectives of the VO. Resources are contributed and utilised by the VO members in a way that is commensurate with what they expect to get out. Negotiation and verification of assets and their utilisation is largely based on established relationships. Where services are provided by one member to another without direct reciprocation it is usually part of a wider context where the resource provider is motivated by central funding. They may even part of the same establishment and are simply using Grid to provide a resource that would normally be provided anyway by other means.

Current-generation Grid is not able to cope with a more pragmatic and realistic "asymmetric" VO model. This refers to a coalition of partners whose motivation for participation varies - from those who have ownership of the problem in hand, to those who are interested in contributing capability towards that goal in return for payment or some other redress.

A lifecycle model for a generic Virtual Organisation can be summarised as four phases: *Identification*, *Formation*, *Operation* and *Dissolution*. Each of these phases requires concepts, architectures, standards and technologies, and almost none of this capability is available right now. For example: we are currently attempting to operate and/or participate in several VOs. Each has a different purpose but may share the same physical resources and staff. Formal policies for the operation of the VOs that can address quality-of-service issues, trustworthy sharing and access policies and

This suggests that there is a further layer to be added to the Grid model – called "VO Operation" – and this sits between the Grid-enabled resources layer and the VO itself. This layer is the location for many services and capabilities identified in various discussion papers on the future of Grid[17,18]. It also helps explain why many of the shortfalls identified by users of current-generation Grid are not easy to address at the middleware level. Trust, negotiation, contracts, financial processing, auditing, business system integration are a handful of the issues we need to build into this layer.

In some cases there is an interesting inter-relationship between the current middleware and the VO Operation layer. For example, we need to consider interoperability within the proprietary applications that we wish to bundle up and offer to a VO. This requires many aspects of GGF standards and involves us in issues such as semantics and ontologies, workflows, directory services, security etc. This scales up when we join a VO where all these issues are needed this time at the partner/partner level. However, there are also factors that are unique to this new layer. In particular the financial, contracting, negotiating and auditing tools are unlikely to be needed within a single organisational entity, but are key enablers for an asymmetric VO.

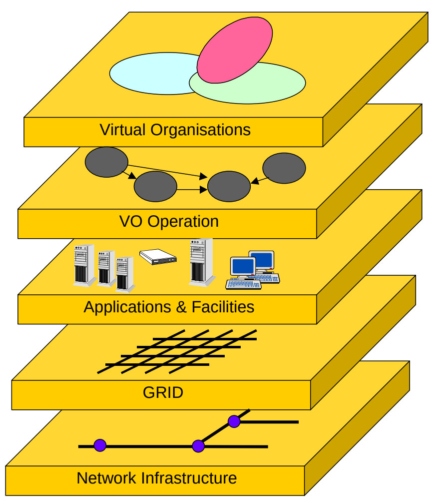

Figure 6 A modified layer architecture with the addition of a "VO Operation" category.

Conclusions

Grid, and its related concepts, are just the latest steps in a series of innovations in information technology over the past few decades. From an "end-user" perspective, most of these innovations have led to changes and (usually) improvements in the way business is conducted. It may therefore seem sensible to "sit back" and let the field mature before investing time and effort. However, Grid goes deeper and encompasses a much broader range of business activity than previous IT developments since what is at first glance a system for aggregating supercomputers turns out to be a whole new way of collaborating and interacting.

Right now there is, of course, a large gap between the vision and the reality, and the extent of the current deployment of Grid techniques is limited to those areas of business that have similar demands to large-scale science. The compute-intensive engineering activities (such as automotive crash simulations, aerodynamic and structural design of aircraft) are already starting to benefit. An additional motivation is the current economic climate that is driving demands to share resources and distribute capability with the aim of reducing fixed costs and improving efficiency.

There are two general areas of Grid benefits:

1. The delivery of capability over the Internet in the form of services. This offers a new way to access software, hardware, data, and information.
2. The ability to form and manage dynamic, scalable virtual organisations. This offers a new way to organise, manage, control and participate.

The first is largely achievable through extensions to existing capabilities, and careful and well-validated deployment of new systems and processes. Assuming, of course, that we can also solve the security problems! The second is much more difficult, but could have an order of magnitude more impact. Key enablers for the "next generation" of Grid technology are security (the fundamental enabler), semantics (the key to interoperability at the applications level) and human factors (central to effective collaboration).

References

[1] Ian Foster. *The Grid: A New Infrastructure for the 21st Centrury*. Physics Today, 2002.

[2] Globus Tookit:
www.globus.org

[3] MyProxy is a credential repository for the Grid. See http://grid.ncsa.uiuc.edu/myproxy

[4] The AccessGrid collaboration system:
www.accessgrid.org

[5] Condor High Throughput Computing capability, www.cs.wisc.edu/condor

[6] Entropia's DCGrid product: www.entropia.com

[7] Platform Computing's grid technology: www.platform.com/grid

[8] The GEODISE e-Science project: www.geodise.org

[9] AG Davies, D Bedwell, M Dunleavy, N Brownjohn *Pressure Sensitive Paint Limitations and Solutions* ICIASF Record 1997.

[10] Ian Foster, Carl Kesselman, Steve Tueke. *The Anatomy of the Grid*. Intl. Journal of HPC Applications, 2001.

[11] Anne Trefethen et al. Presentation to the e-Science Town Meeting, May 2003.

[12] M Surridge *A Rough Guide to Grid Security*. IT Innovation Centre, 2002.

[13] Platform Computing Inc. *The Non-Technical Barriers to Implementing Shared Computing in a Commercial Environment*, Market Study Report, March 2003. www.platform.com/barriers

[14] Multi-Site Videoconferencing for the UK e-Science Programme: *A Roadmap for the Future of Videoconferencing within e-Science*. e-Science technical report, 2002.

[15] J Giddy, A Gould, J Jones, D Rowse, M Turner, D Walker, N Weatherill. *Grid-Enabled Computational Electromagnetics*. Proceedings of the e-Science All-Hands meeting, Sept 2003.

[16] A summary of some operational Grids deployed by IBM can be found at: www.ibm.com/grid

[17] Expert Group Report. *Next Generation Grid(s) – European Grid Research 2005-2010.* European Commission, June 2003.

[18] D De Roure, M Baker, N Jennings, N Shadbolt *The Evolution of the Grid.* In Grid Computing – Making the Global Infrastructure a Reality. Berman, Hey, Fox eds. 2003.