The *e*Minerals collaboratory: tools and experience

Martin T Dove[1,3], Mark Calleja[1], Richard Bruin[1], Jon Wakelin[1], Matthew G Tucker[1], Gareth J Lewis[2], S Mehmood Hasan[2], Vassil N Alexandrov[2],Martin Keegan[3], Stuart Ballard[3], Richard P Tyer[4], Ilian Todorov[4], Paul B Wilson[5], Maria Alfredsson[5], G David Price[5], Clovis Chapman[6], Wolfgang Emmerich[6], Stephen A Wells[7], Arnaud Marmier[8], Stephen C Parker[8], Zhimei Du[9]

1. Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ
2. Department of Computer Science, University of Reading, Whiteknights, PO Box 225, Reading RG6 6AY
3. National Institute for Environmental eScience, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0EW
4. Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD
5. Department of Earth Sciences, University College London, Gower Street, London WC1E 6BT
6. Department of Computer Science, University College London, Gower Street, London WC1E 6BT
7. Royal Institution, Albermale Street, London W1S 4BS
8. Department of Chemistry, University of Bath, Bath, BA2 7AY
9. School of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX

Abstract

Collaboratories provide an environment where researchers at distant locations work together at tackling important scientific and industrial problems. In this paper we outline the tools and principles used to form the *e*Minerals collaboratory, and discuss the experience, from within, of working towards establishing the *e*Minerals project team as a functioning virtual organisation. Much of the emphasis of this paper is on experience with the IT tools. We introduce a new application sharing tool.

Keywords

Collaboratory, Virtual Organisation, Access grid, grid technologies

1. Introduction

As discussed in the introduction to this collection of papers [1], the *e*Minerals project is a collaboration between molecular simulation scientists, molecular simulation code developers and computing specialists, with participation from a number of different departments (Earth Sciences, Computer Science and Chemistry) and UK institutes (Universities of Cambridge, Reading and Bath, University College London, Birkbeck College, the Royal Institution, and the Daresbury Laboratory). The broad aim of the project is to develop a cross-institute infrastructure to support the molecular simulation scientists using grid technologies [2]. Given that funding bodies are increasingly looking towards consortia-based funding in addition to the traditional single-investigator project model, practical research into how modern IT (information technology) tools can be harnessed to support high levels of collaboration within a consortium is now timely. In fact there is a considerable body of research into the use of IT tools to support collaborations, within a field of work known as "Computer Supported Cooperative Work" (CSCW) [3]. Work in this area began in the mid 1980's, and significant advances have been made in understanding the ways in which remote collaboration is carried out and the issues relating to dynamics of cooperative work. CSCW is technology 'independent', which means technology is not the major driving force behind the discipline. Instead, it is concerned with the way people interact and collaborate with each other, and attempts to develop guidelines for developing technology to assist in the communication process.

In this paper we report the experience of the *e*Minerals project on the use of various IT tools to support collaborations between the various members of the team. The *e*Minerals project is the first attempt to develop a large-scale collaborative infrastructure for molecular simulations based on the use of IT and grid computing tools. The core motivation for seeking ways to enhance the collaboration is to maximise the potential benefits of enabling simulation scientists and IT experts to work together. The *e*Minerals project has set up a collaborative computing and data management minigrid infrastructure to support the simulation work [2], and only through enhanced levels of collaboration can the gains of this work be fully realised. At the outset, it has been essential to support close collaboration between the scientists and the IT experts, and in the second stage the infrastructure will support scientists working together. The *e*Minerals project has used a number of collaborative tools, some very new and others more 'traditional', and here we aim primarily to report on the experiences.

The structure of the paper is as follows. In the next section we outline some of the ideas around the concepts of collaboratories and virtual organisations, and discuss how the *e*Minerals project maps onto these concepts. Then we discuss several tools we have used to support our work. Finally we look at some of the experiences.

2. Collaboratories and Virtual Organisations

2.1 Collaboratories

The concept of the Collaboratory was identified by Wulf as a "center without walls, in which the nation's researchers can perform their research without regard to geographical location – interacting with colleagues, accessing instrumentation, sharing data and computational resource, and accessing information in digital libraries" [4]. The concept is that IT will enable distributed collaborators to work together as if in the same geographical proximity (see the brief discussion on CSCW above). An early discussion (1996) on the use of collaboratories in relation to scientific research is given in reference 5. This article discusses some of the key technologies to support collaborative working (e.g. web resources, video conferencing, chat tools). However, in the intervening few years there have been rapid changes in technologies, which impact particularly on network bandwidth capabilities, which are essential for this work. Reference 5 gives a discussion on issues associated with adopting the use of the collaborative tools. At the time of writing reference 5, network speeds were not sufficient to enable videoconference tools to be satisfactory. As we discuss later in this paper, this tool now has a much better user experience, due primarily to the faster networks.

2.1 Virtual organisations

The second concept to have emerged over the past 10–15 years is that of the Virtual Organisation (VO), but many of the practices of virtual organisations can be traced back at least four decades. For example, Sor [6] has described how many of the features of virtual organisations can be discerned within the organisation of the housing construction industry in Western Australia in the early 1960's. Much has been written about the more specific concept of the 'Virtual Business', and many definitions of VO's are particularly pertinent to the industrial sector. The strong motivation for the formation of VO's in industry is the need to reduce costs – this was the driving factor that saw the drive towards collaborations in Western Australia. The core idea is that costs can be saved if partners with complementary expertise work together towards some common objective. In particular, the infrastructure costs of a group of small units are likely to be much lower in total than the corresponding cost to a large organisation.

There is no agreed definition of the term 'Virtual Organisation' in the same way that there is for a collaboratory (see quote from Wulf above), but there are a number of key characteristics that can be said to be implicit in the idea of the VO. Our working understanding of the concept of the

VO is that it is a collection of people working together within an organisational structure that is distinct from their formal allegiances (although not relevant in an academic context, some members of a VO may be freelance without any institutional allegiance). A VO will have a particular mission, and may be time limited. Its members will be geographically dispersed, and will have responsibilities on behalf of their employer institute as well as on behalf of the VO. Membership is likely to be dynamic, with members joining and leaving when their roles begin and end, rather than remaining members for the whole duration of the project around which the VO is established. Some aspects of the VO are similar to those of a more traditional working organisation, but there are other aspects, such as a flatter hierarchy and a voluntary commitment, that are more peculiar to the VO. It is being recognised that the dependence on collaborative IT tools is one of the main characteristics of a VO. However, since VO's existed in practice before the IT revolution of the past decade, the reliance on IT cannot be said to be a defining characteristic.

It is useful to identify the concepts of a VO that are most pertinent to an academic science project, noting that the project is not subject to the same constraints as would be central to the concept of a Virtual Business. In particular, there is not the equivalent of a fixed objective – research objectives have to be sufficiently flexible to be able to evolve and adapt in response to new discoveries by the project team and competing research groups, and to the common situation where proposed research meets unpredictable problems. Moreover, there is no corresponding cost-reduction motivation (apart from the advantages of sharing computing resources as outlined in reference 2), and usually it is expected that member groups will stay together for the whole lifetime of a particular project (rather than joining for short periods). The key point is the joining together of dispersed research groups to work together on a topic of common interest with a commonly agreed management structure. One might ask how this differs from a standard collaboration of distributed groups? We argue here that there are two features of a VO that are qualitatively different from features of a looser collaboration. The first is that there is a sharing of resources (as in the minigrid infrastructure established in the *e*Minerals project [2]) that is akin to the manner in which resources are shared within a formal organisation. For example, there may be semi-formal policies on access to some of the shared resources, and a commitment on the part of the donor to ensure that access is properly maintained. The second particular feature of a VO is interdependence between member groups that is built into the VO from the outset. It is possible for members of a collaboration to gain benefit from other members but to not be dependent on each other, as discussed below.

2.3 Mapping the eMinerals project onto the concept of the Virtual Organisation

As noted in the Introduction, consortium funding is increasingly becoming more common. Often such consortia are formed by groups with similar skills and related interests. In such a case, there may be no in-built interdependence on the consortia members. Thus there will be a tendency towards working within the traditional model of collaboration that is built upon regular but infrequent face-to-face meetings where progress is reviewed (a good consortium will gain a lot from these meetings), irregular but frequent email contacts where help/advice is sought (and the telephone when contact is particularly urgent), and the reading of manuscripts sent between partners. This is often as much as the partners expect out of the collaboration. On the other hand, projects such as the *e*Minerals project have the interdependence between partners built into them from the outset. Close collaboration between the science and IT team members has been essential to enable the project to achieve its goal of constructing an integrated grid structure that meets the needs of the scientists. The scientists need to inform the grid team of their requirements, and the grid team need to be helped to develop something that the scientists will find genuinely useful. The scientists also need a lot of help adapting their usual work practices to the new grid-based way of working. The simulation code developers need to select their priorities based on the requirements of the users through working closely with the scientists (as distinct from the model where groups who use a particular code formulate wish lists). The code developers can also use the grid structure to their benefit, for example for testing program builds on many platforms, and will need to interact with the grid team to ensure they get as much from the grid structure as the science team does. Thus we sought to build collaboration between all members of the *e*Minerals project team from the outset. We have had to work within the constraint that our teams are based in geographically distinct locations, and yet we want interactions to be much more frequent than would be possible if restricted to face-to-face meetings.

It should be noted that the science community from which the *e*Minerals project is drawn is not used to working within large close collaborations. It is much more characterised by individuals working with their own resources and codes. Consortia may be formed in order to gain access to high-capacity computing facilities, and partnerships may be formed between groups working on common problems. But these are a long way from the concept of the VO outlined here. Other areas of science, particularly in particle physics and astronomy, have a much stronger track record of the need to work within interdependent collaborations.

We make one final remark in this section. The success of business VO’s depends on the use of standards, so that all partners work to the same system and interoperability is built into the VO from the outset. In our case, one standard that we adopt is the use of the Chemical Markup Language (CML) to describe the simulation data. CML is an application of XML that is designed to handle the science that drives our project, and it enables data to be imported and exported into and out of many of the codes used in the eMinerals project [7–10]. Similarly, for interoperability at the grid level, we have made use of the benefits of the use of grid standards, such as the use of digital certificates to provide authentication for access to resources [2].

3. IT/Grid tools to support collaborative work

Based on the above discussion, it is obvious that the *e*Minerals project has had to pay some attention to providing the IT infrastructure in support of the operation of the VO. There are a number of tools that can be used, some of which are emerging as new grid tools within the current eScience initiative. Some available tools, such as shared diaries and project management tools, are more appropriate for short term project work with hard deadlines and inflexible deliverables, and we have not taken them on board in the eMinerals project.

3.1 Interactive collaboration tools

We now briefly review our experience of a range of tools that have the potential to enhance collaboration across the *e*Minerals VO. Each tool has its own advantages and disadvantages, which we consider under three broad categories: the 'access cost' to the communication, that is whether it is easy or difficult to initiate a new communication; the 'potential for instantaneous interaction'; and the 'setup cost', whether financial or in terms of necessary initial effort or expertise. These issues may be hard to quantify in absolute terms, but it is possible to form relative broad-brush comparisons (e.g. easy or hard). What is more difficult to quantify in any terms is the quality of the support each tool gives to the collaborative working. In this case, we give a number of remarks on some of the quality issues but we will not seek to provide a quantitative comparison between the tools on the quality factors.

The spectrum of communications tools consists of the following:

eMail: Although relatively new, email is now pervasive in the scientific community. It is cheap (zero setup cost), easy, and automatically available for everyone (low access cost). It is virtually platform independent. The difficulties we face in the use of email to support collaboration within the *e*Minerals VO are that it does not support instant communication (the speed of response will depend on people's email receipt setup, and email does not demand attention), and that with the welter of email communication (genuine and unsolicited) it is gradually losing its effectiveness as a means of communication (in general, people are no longer able to read and respond to all emails they receive, particularly emails sent to a wide circulation list). eMail is only used significantly to send announcements to the project (e.g. concerning the status of a particular resource), but it has not proven to be useful to support discussions. Quality is also affected by a number of well-known problems, such as the ease at introducing ambiguity into a message.

Instant messaging (IM): Instant messaging gives much better instantaneous chat facilities than email, and is being used within the *e*Minerals project for discussions between small groups of people. It has a very low immediate access cost and nearly zero setup cost (simple installation and initial registration), and is available for all platforms in the project. It is less useful for larger numbers (> 3) of participants for two reasons, namely the difficulty in maintaining social control (who should speak next, particularly since it takes time to type in a message), and because communication easily becomes tangled when participants follow several lines of thread in an IM discussion. IM has been particularly useful for team members developing new tools because it is easy to send quick questions and sections of code. Unfortunately some organisations formally block the use of IM tools.

Web tools: At the lowest and slowest level, information can be transmitted to members of the eMinerals project via web sites (e.g. www.eminerals.org). The value of a project website is that it is owned by the project members and can be shaped to meet the needs of the project (whether for dissemination or access to information). However, not all team members have access to the directory structure that supports the project web site, and they have no option but to rely on those who do have access to deposit information.

The development of the 'wiki' concept [11] (the name originates from the Hawaii 'wiki wiki' meaning 'superfast') removes the problem of members of the VO not having direct write access to the project website. A wiki is a web site that can be freely edited by the community (document pages altered, new documents added, links created) using a simple markup language. Changes are implemented instantaneously (without review). We have set up a private wiki site for exchange of information within the eMinerals project, to which access is limited to members of the eMinerals VO through the use of X.509 certificates. People can deposit news information, edit documents (this paper was written on the wiki, and only transformed into a document format at the final stage), deposit information (for example, in order to collate information for a project review), and post questions and answers. Unlike normal web sites, the wiki has a relatively low access cost for members of the VO to post information to. As with all web-based tools, the wiki does not support instant communication. Its main problem is that it relies upon other team members regularly looking at it, and this limits the quality of its use as a collaborative tool. In practice, we have found that the main uses of the wiki have been for project members to deposit and edit information when writing papers or grant proposals, to manage task lists, and to provide a repository for project information (particularly "frequently asked questions").

Video conferencing based on Access Grid suites: The Access Grid [12] is defined as an "Advanced Collaborative Environment". It is used mainly for high quality audio/video conferencing, allowing groups of distributed users to meet in on-line rooms known as Virtual Venues. The Access Grid is used widely in many academic institutions, and can be used for distributed lectures, meetings or conferences. Access Grid suites usually consist of a large display, with one or more projectors, several high quality cameras, conference microphones, loudspeakers, all controlled by a central console.

The Access Grid uses a network technology called multicast. This has the advantage of reduced network bandwidth requirements through sending only one copy of the signal, with that copy only being duplicated when the network routes to conference participants branch. More traditional methods would send a separate copy of the signal to each participant, which has the consequence of limiting the number of participants that can take part in a videoconference. The Access Grid can support some tens of participating sites. For example, it is being used to relay talks in conferences (both broadcasting talks to a wider audience, and allowing speakers to participate from their home institutes), and provides a much higher quality of broadcast that modern streaming video can achieve.

The Access Grid is specifically designed for group-to-group collaboration. It uses multicast to reduce the network traffic and provides a scalable group-to-group collaborative environment. Although the Access Grid and other conferencing environments cannot replace the experience of face-to-face human interaction, it does offer a viable alternative to the expense of travelling for project meetings. In fact one of the driving motivations behind the Access Grid has been to set up a structure that best reproduces the face-to-face experience of a meeting. For example, the use of more than one camera enables better use of visual clues, and the projection system to display conference partners, as opposed to a television system, gives a better visual experience. Although the initial expense for the equipment and training is high, it does allow for regular meetings of project members. The main disadvantages of using an Access Grid suite are that you may still have to travel to get to the nearest suite, arranging access to the suite, and ensuring that there is a trained operator available for the time of the meeting.

Personal Access Grid: There is a personal version of the Access Grid software, know as the Personal Interface to the Access Grid (PIG), which can run as an alternative to the use of purpose-built Access Grid suites. It can be set up on any desktop or laptop computer running Microsoft Windows or Linux, and only requires a microphone headset and a web cam. The PIG has all the functionality of the full Access Grid, and similarly relies on the use of multicast network technologies to minimise bandwidth requirements (see below for some comments on this). Although the use of a computer screen is more limited than the use of a projection wall in an Access

Grid suite, we find that it is still capable of supporting reasonably large conferences. An image from one session is shown in Figure 1. The quality of the audio and video is likely to be lower than that of an Access Grid suite, but this is mainly due to the quality of the equipment.

The PIG provides for the highest quality level of collaboration of all the tools discussed in this paper. It is easy to run, and provides collaborations with many of the advantages of face-to-face meetings (and is certainly better than telephone for long discussions). It enables instantaneous communication, and has a low access cost.

The major problem with the PIG is the high set up costs. As mentioned the Access Grid relies on Multicast capabilities, but there is a problem in that in many institutions multicast is not enabled within the Local Area Network (LAN) by default (but the capability will be in place in most cases). There are several problems associated with enabling multicast on existing networks, including setting up firewalls, routers and networks switches to handle multicast. Where multicast is not enabled within an institute, it is possible to use a unicast–multicast bridge. This is not a scalable solution, and if it is necessary for many participants to use a bridge it is necessary to have several bridges running in order to preserve quality. The *e*Minerals project has set up several such bridges. Other technical issues are discussed in some depth in reference 13.

The main alternative to the PIG are point-to-point video conferencing tools based on H.323 protocols (such as Microsoft Netmeeting and Apple's iChat). The main limitation of these tools is that they are restricted in the number of participants. However, they do have a role to play for one-to-one communications, and in some situations (e.g. slow network speeds as in asymmetric broadband from home) they work better than the PIG. For one-to-one videoconferencing the eMinerals project does make use of these alternative tools.

3.2 Other support tools

In addition to the communication tools reviewed above, we are using two other tools to support the eMinerals VO.

Tools for shared applications: In addition to holding discussions on the Access Grid, it is sometimes useful for participants in a discussion to be able to look together at representations of data, such as viewing the atomic arrangements in a molecular simulation. The VNC (Virtual Network Computing) [14] tool provides some level of support for this in that is allows users to share a single desktop and can be used to share applications. For example, a crystal structure displayed on one desktop can be seen by all participants in a discussion, and all participants are able to manipulate that structure. However, it does have two features that make it less suitable for this role. First, the user must share the entire desktop rather than a single application, which may not be desirable, and second, VNC uses unicast for data transmission rather than multicast. The main problem with using unicast in a group environment is that the server machine needs to create and send multiple copies of the same data. This increases the load on the server as well as flooding the network with duplicate packets, leading to scalability issues. To address these issues, we have developed a Multicast Application Sharing Tool (MAST). The tool has been specifically designed to enhance group-to-group collaboration within an Access Grid session. MAST shares single application windows between participants in a collaborative session. Like the Access Grid, it uses multicast to transport the application streams to the participants in the multicast group. It gains a number of speed advantages through the use of specific optimisation strategies (including the obvious one of needing to share the whole desktop). MAST supports multiple platforms. MAST is illustrated in Figures 2 and 3. MAST can be used for a range of applications, including document preparation, slideshow presentations, shared drawing, and data visualisation.

Helpdesk software: Once the *e*Minerals minigrid was operational, the science users were then helped by the grid team to port their applications across to the minigrid and start running in earnest. Recall that access to the resources of the *e*Minerals minigrid is only via Globus tools, and users needed support to switch from the more traditional login methods. At an early stage it was realised that it was essential to implement a structured support system rather than have users guess whom to ask for help. As the number and complexity of the problems encountered by the users increased, and considering the varying areas of expertise of the members, it became necessary to find means to allow groups of users to cooperate towards solving problems. It is in order to achieve this that we adopted the Open Ticket Resource System (OTRS), an open-source email support system [15]. Though email is traditionally a one-to-one form of communication, the OTRS system provides means for multiple users to manage help requests sent to a common address (i.e. helpdesk@eminerals.org). Incoming requests can be matched to the users most apt to respond, and several users can view, respond and add comments to existing help requests via an online Web interface. The OTRS is a relatively simple tool to use, that can be tailored to best match the capabilities and structure of the group. An example of this is the use of request queues and notifications. Several queues, which contain waiting help requests, have been set up to correspond to various areas of expertise (e.g. data transfer issues, grid tools, etc.). For each of these, a set of users most capable of responding was defined. Upon receipt, incoming emails are sorted into the right queues, and the corresponding helpdesk users are instantly notified of the help request by email. They can then log into the system to view, respond, comment, or move the request to a different queue.

Electronic team newsletter: Newsletters are a more traditional method to facilitate the flow of information within a collaboration, but nevertheless we have found that the use of a regular project newsletter has played a significant role in supporting the development of the eMinerals VO.

Modern desktop publishing packages make it easy to put together a high-quality document, and the web and email allow for easy dissemination. The key to the successful use of the newsletter is the editor, and his/her ability to extract contributions from the project members against a set of fixed regular deadlines. Past copies of the eMinerals newsletter are available from http://www.eminerals.org/.

4. User experiences of the personal Access Grid

The Access Grid is designed primarily to support large scale collaborative interactions, rather than the more traditional videoconferencing between two individuals. We are using the Access Grid at an intermediate level: clearly at each end there are individuals, but we are making significant use of the ability of the Access Grid to support many simultaneous users. For project meetings, we may well often have around 10 participants, and we often mix in the Daresbury Access Grid node. As noted above, we do not find that the limited screen area (compared to the large display area of an Access Grid suite) is an issue for this scale of interaction.

We use the Access Grid to support our collaborations in a number of ways:

Informal plenary project mornings: The eMinerals project team hold regular informal meetings which are used to exchange news and other information, and at which people can pose questions for discussion (Figure 1). We have specifically aimed to give the feel of an informal coffee morning were people can pop in and out. We have found that this works well, with team members engaging naturally with each other.

Team project support: Our grid team has a number of large and small projects on the go (e.g. reference 16). The team has found it to be very advantageous to be able to meet weekly to monitor and develop progress.

Support work and *ad hoc* discussions: it has been very useful to use the personal Access Grid in “on demand” mode to facilitate discussions on specific issues as they arise, such as troubleshooting other work in the project, helping people install software, planning tests of tools etc. Although this could be done with other videoconferencing tools (e.g. tools that use H.323 protocols), the personal Access Grid does just as well. The point here being that the ability to run *ad hoc* videoconferences is proving to be very useful in functioning as a VO.

Project management: When it comes to writing reports and new proposals, the Access Grid facilitates the interaction between team members that is essential. For gathering ideas, examples, summaries of work carried out, and lists of publications, the wiki tool described earlier is ideal to assist in the collation of information. However, in then working through the information, we have found that the use of the Access Grid has been invaluable.

5. Examples

Running the *e*Minerals project as a collaboratory and VO as opposed to a looser collaboration, and providing the technical support for this, has enabled the *e*Minerals project to accomplish more than would otherwise have been possible. In this section we present a few examples.

5.1 Development of the *e*Minerals compute portal

One of the ongoing tasks within the *e*Minerals project is the development of a web-based compute portal [16]. This involves developers in Daresbury, Cambridge and London. The portal team have used the PIG/Access Grid for regular meetings to review progress and plan next stages in the work. In addition, a wiki has been set up to enable information and task lists to be deposited and updated. For immediate issues, IM tools are regularly used. The experience of IM is that you know who is available at any particular instance, and it is very quick to send a message and get an immediate reply.

5.2 Support for simulation scientists

The simulation scientists in the *e*Minerals project have had to learn a number of new tools in order to use the minigrid infrastructure. A large part of the support for this work has been performed using the tools described above. The PIG/Access Grid has been particularly useful to enable group discussion sessions for training and problem solving. The helpdesk system has been made the work of the support team easier and more effective, as has the wiki for managing a rapidly developing "Frequently asked questions" section. IM has proven to be very useful for quick questions. The work of reference 10 is an example of a successful collaboration between the simulation scientists and the grid team to develop new methods of running molecular simulations. References 17–21 are particularly good examples of science that has been achieved as a result of collaborative work assisted by the tools described in this paper.

5.3 Drawing work together for publication

The tools have been particularly useful for writing collaborative documents such as publications and proposals. On several occasions team members have edited documents together, including using the PIG and a combination of IM/application sharing, while exchanging documents using the tools described below. These tools provide the immediacy that collaborative writing requires.

5.4 Use of the data grid for collaborative work

One of the components of the eMinerals minigrid is a shared distributed data grid, including the use of the Storage Resource Broker [2]. The SRB is primarily designed for managing data across the minigrid. However, it is also particularly useful to support collaborative work, as the project is able to share documents, scripts, and applications by depositing them within the distributed grid environment, rather than sending around documents by email.

6. Conclusion and summary

We have reported experience gained within the eMinerals collaboratory of using a range of IT/grid tools for the development of the infrastructure to support a virtual organisation. We have particularly emphasised the use of the personal Access Grid, providing some examples of how this has been a useful tool. Some of our points are summarised in Figure 4, which compares the various factors we have taken into account in our discussion of the different tools available to support collaboration within a VO. Clearly there is a compromise. Some tools (e.g. email and instant messaging) are very easy to set up and use, but suffer from particularly difficulties as instant communication tools. eMail as a tool is now overwhelmed by misuse and overuse, and IM supports only a small number of participants (the same criticism that can be levelled against the telephone). On the other hand, the personal version of the Access Grid has a high initial set up cost, but once setup it is easy to use and provides communication facilities that are not given by other tools. The setup costs should not be underestimated, but fortunately once it has been set up for one person at one site other users will immediately benefit.

Our experience with the personal version of the Access Grid as a collaborative tool has been very positive. We have not found that restricting the Access Grid to a computer desktop creates many disadvantages compared to the full Access Grid, and has a number of important advantages over the full Access Grid in terms of easy and instantaneous access for all members of the VO. The low financial cost also ensures that roll-out across a whole consortium is affordable.

Acknowledgements

We are grateful to the Natural Environment Research Council for financial support for the *e*Minerals project, and to the computer support personnel at our various institutes.

References

[1] Dove MT and de Leeuw N (2005) Grid computing and molecular simulations: the vision of the *e*Minerals project. Molecular Simulations **XX**, 1234–5678

[2] Dove MT, Calleja M, Bruin R, Wakelin J, Tucker MG, Lewis G, Hasan SM, Alexandrov VN, Keegan M, Ballard S, Tyer RP, Todorov I, Wilson PB, Alfredsson M, Price GD, Chapman C, Emmerich W, Wells SA, Marmier A, Parker SC, Du Z. The eMinerals collaboratory: tools and experience. *Molecular Simulations* **XX**, 1234–5678

[3] Grudin J (1994) “CSCW: History and Focus”. IEEE Computer, 27 (5), 19–27

[4] Wulf WA (1993) The collaboratory opportunity. Science, 261, 845–855.

[5] Kouzes RT, Myers JD and Wulf WA (1996). Collaboratories: Doing Science On The Internet. IEEE Computer, **29,** 40–46; Article is available from http://collaboratory.emsl.pnl.gov/presentations/papers/IEEECollaboratories.html

[6] Sor R (1999) “Virtual Organisations – Not Such a New Idea: A Case Study of the Housing Construction Industry in Western Australia”. Proc. 10th Australasian Conference on Information Systems, pp 825-836.

[7] García A, Murray-Rust P, Wakelin J. The use of XML and CML in Computational Chemistry and Physics Programs. *Proceedings of the UK e-Science All Hands Meeting 2004*, (ISBN 1-904425-21-6), pp 1111–1114, 2004

[8] Zhang Y, Murray-Rust P, Dove MT, Glen RC, Rzepa HS, Townsend JA, Tyrrell S, Wakelin J, Willighagen EG. JUMBO – An XML infrastructure for ecience. *Proceedings of the UK e-Science All Hands Meeting 2004*, (ISBN 1-904425-21-6), pp 930–934, 2004

[9] Wakelin J, Murray-Rust P, Tyrrell S, Zhang Y, Rzepa HS and García A. CML Tools and Information Flow in Atomic Scale Simulations. *Molecular Simulations* **XX**, 1234–5678

[10] Chapman C, Wakelin J, Artacho E, Dove MT, Calleja M, Bruin R and Emmerich W (2005) Workflow issues in atomistic simulations. *Molecular Simulations* **XX**, 1234–5678

[11] http://wiki.org/

[12] http://www.accessgrid.org/

[13] Dove MT, Calleja M, Bruin R, Wakelin J, Keegan M, Ballard S, Lewis G, Hasan SM, Alexandrov V, Tyer RP, Todorov I, Wilson P, Alfredsson M, Price GD, Chapman C, Emmerich W, Wells S, Marmier A, Parker S, Du Z. Collaborative tools in support of the eMinerals Virtual Organisation. *Proceedings of the UK e-Science All Hands Meeting 2004*, (ISBN 1-904425-21-6), pp 127–134 (2004)

[14] Richardson T, Stafford-Fraser Q, Wood KR and Hopper A (1998) “Virtual Network Computing”, IEEE Internet Computing **2**, 33–38

[15] Edenhofer M, Wintermeyer S, Wormser S & Kehl R (2003) “OTRS 1.2 - Manual”, http://doc.otrs.org/1.2/en/html/

[16] Tyer R, Calleja M, Bruin R, Chapman C, Dove MT, Allan RJ. Portal Framework for Computation within the eMinerals Project. *Proceedings of the UK e-Science All Hands Meeting 2004*, (ISBN 1-904425-21-6), pp 660–665 (2004)

[17] Wells S, Alfredsson M, Bowe J, Brodholt J, Bruin R, Calleja M, Catlow R, Cooke DJ, Dove MT, Du Z, Kerisit S, de Leeuw N, Marmier A, Parker S, Price GD, Smith B, Spohr H, Todorov I, Trachenko K, Wakelin J and Wright K (2004) Science outcomes from the use of

Grid tools in the *e*Minerals project. *Proceedings of the UK e-Science All Hands Meeting 2004*, (ISBN 1-904425-21-6), pp 240–247

[18] Du Z, de Leeuw NH, Grau-Crespo R, Wilson PB, Brodholt JP, Calleja M and Dove MT (2005) A computational study of the effect of Li-K solid solutions on the structures and stabilities of layered silicate materials – an application of the use of Condor pools in molecular simulation. *Molecular Simulations* **XX**, 1234–5678

[19] Marmier M, Spohr H, Cooke DJ, Kerisit S, Brodholt JP, Wilson PB and Parker SC (2005) Self Diffusion of Argon in Flexible, Single Wall, Carbon Nanotubes. *Molecular Simulations* **XX**, 1234–5678

[20] Alfredsson M, Brodholt JP, Wilson PB, Price GD, Corà F, Calleja M, Bruin R, Blanshard LJ and Tyer RP. Structural and Magnetic phase transitions in simple oxides using hybrid functionals. *Molecular Simulations* **XX**, 1234–5678

[21] Wells S, Alfe D, Blanchard LJ, Brodholt JP, Calleja M, Catlow CRA, Price GD, Tyer R and Wright K. *Ab initio* simulations of magnetic iron sulphides. *Molecular Simulations* **XX**, 1234–5678 (2005)

Figure captions

1. Snapshot of a screen during an eMinerals Access Grid team meeting. This meeting had 9 simultaneous video/audio streams. The display was on a laptop with resolution 1400 × 1050 pixel resolution.
2. Examples of the MAST application sharing tool being developed within the eMinerals project. The top images show application sharing with Microsoft Windows, and the bottom screens show application showing with Linux. In both cases, each window shows
3. Example of the MAST interface on the Linux version.
4. Rough comparison of the four main communication tools discussed in this paper. Shorter bars are better, but absolute lengths of the bars do not carry quantitative significance. Our argument is that the higher setup costs of the personal version of the Access Grid are well matched by the lower time delays for real communications. This diagram does not convey information on quality of collaboration because of the difficulty in defining even a relative scale for this.

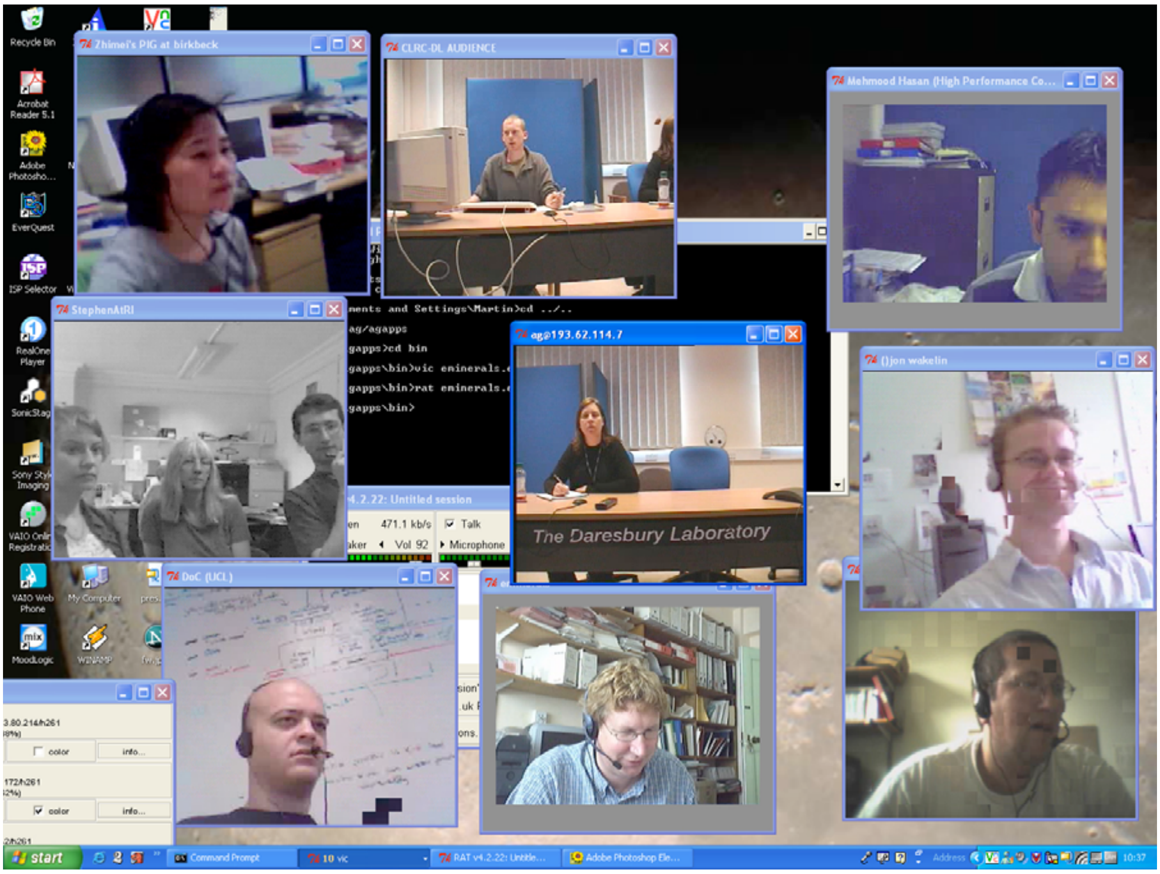
Recycle Bin
Acrobat Reader 5.1
EverQuest
ISP Selector
My Computer
WINAMP
MoodLogic
Zhimei's PIG at birkbeck
CLRC-DL AUDIENCE
Mehmood Hasan (High Performance Co...
StephenAtRI
ag@193.62.114.7
The Daresbury Laboratory
()jon wakelin
DoC (UCL)
start
Command Prompt
10 vic
RAT v4.2.22: Untitle...
Adobe Photoshop Ele...
10:37

Figure 1

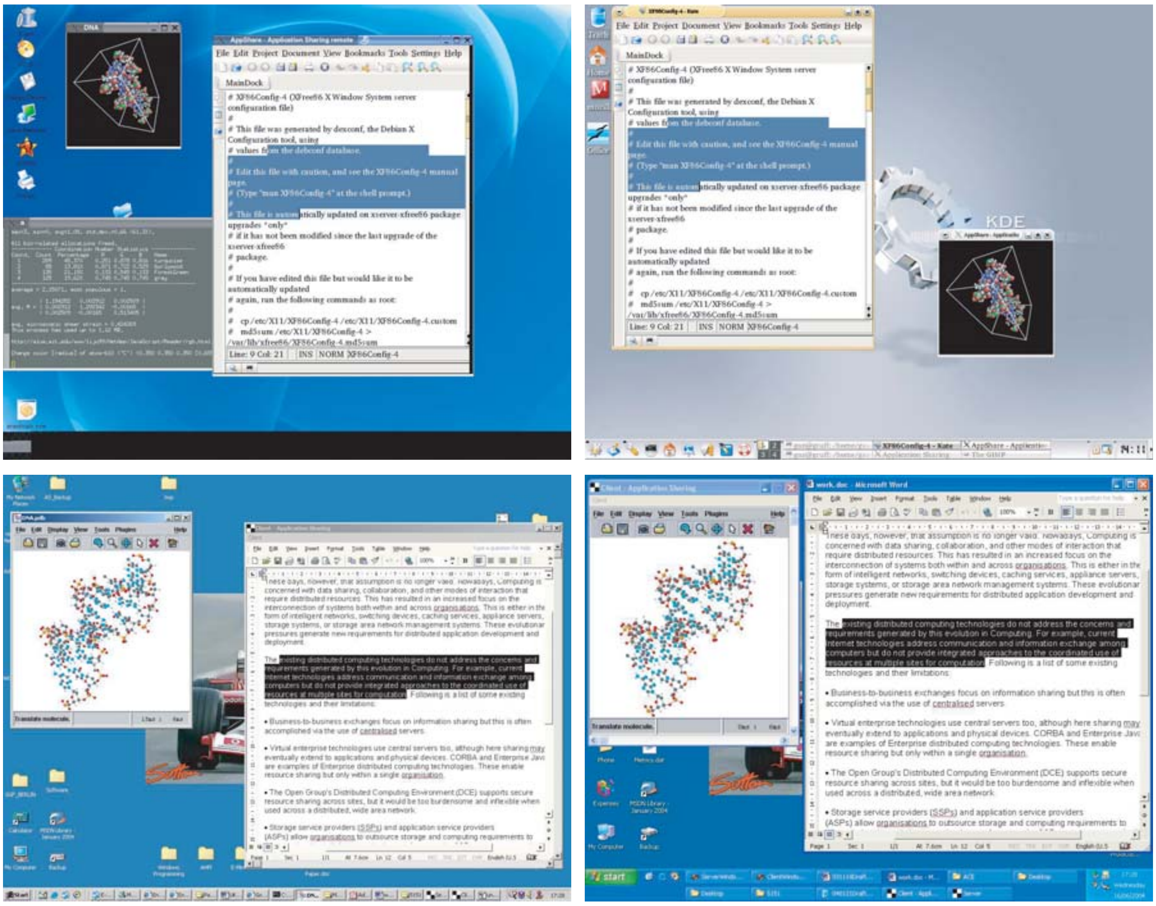

Figure 2

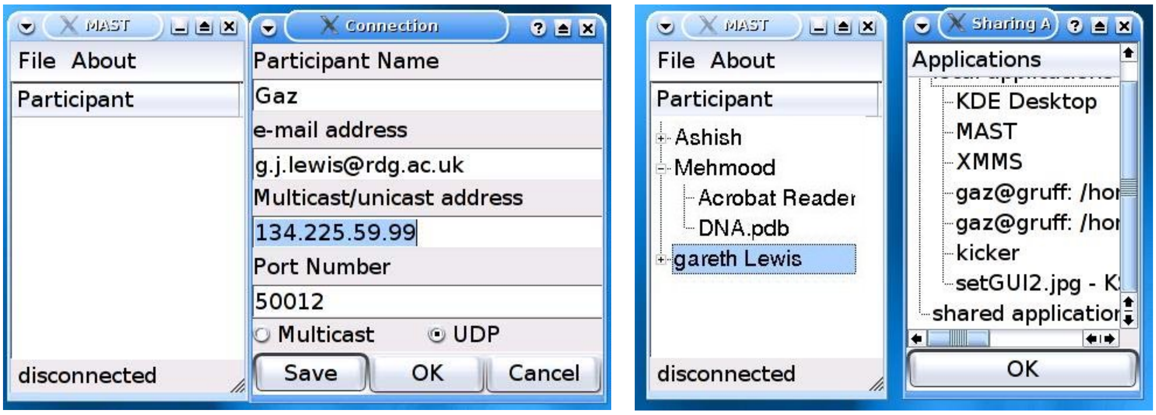
MAST
File About
Participant
disconnected
Connection
Participant Name
Gaz
e-mail address
g.j.lewis@rdg.ac.uk
Multicast/unicast address
134.225.59.99
Port Number
50012
Multicast
UDP
Save
OK
Cancel
MAST
File About
Participant
Ashish
Mehmood
Acrobat Reader
DNA.pdb
gareth Lewis
disconnected
Applications
KDE Desktop
MAST
XMMS
gaz@gruff: /hor
gaz@gruff: /hor
kicker
setGUI2.jpg - K
shared applicatior
OK

Figure 3

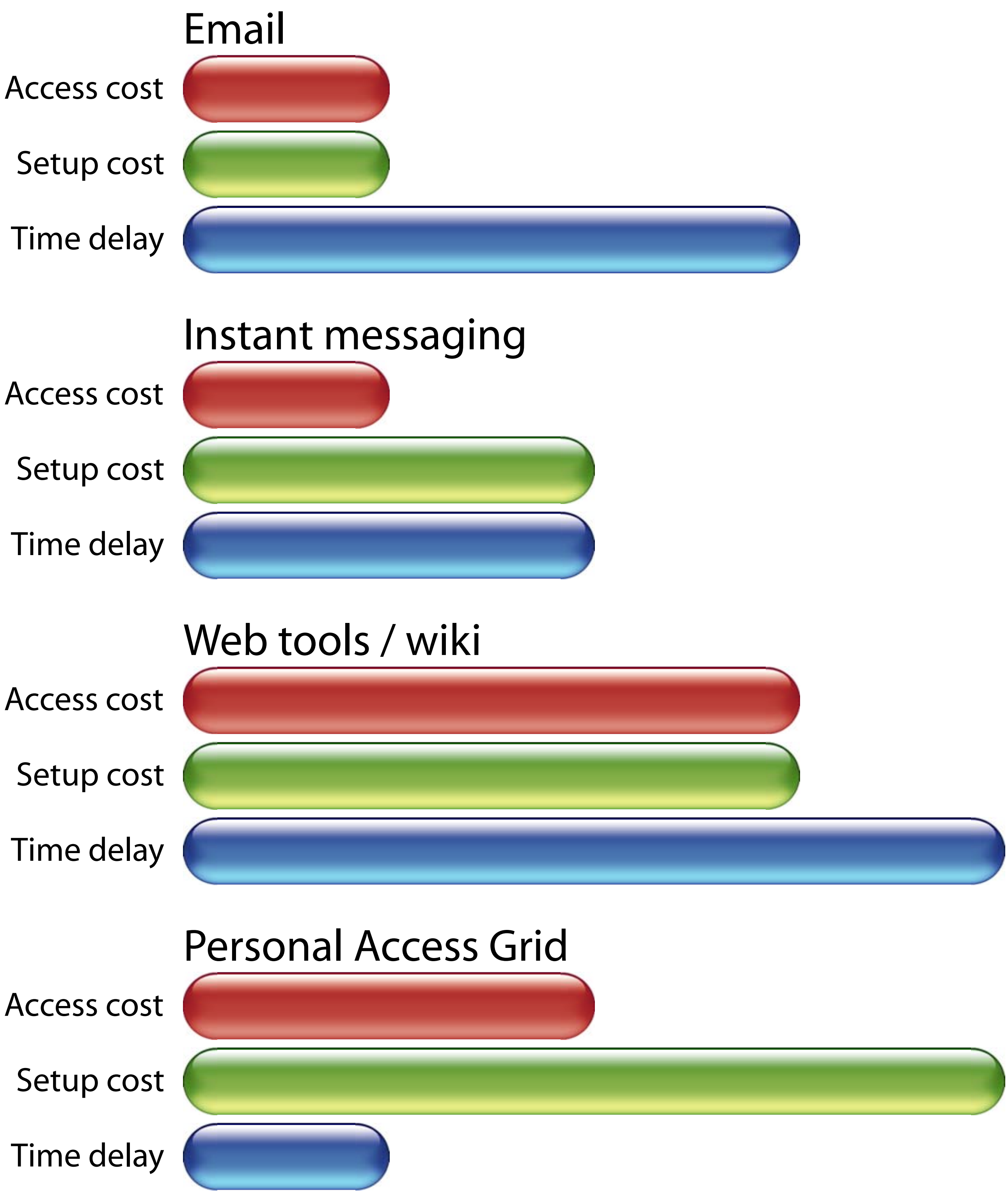
Email
Access cost
Setup cost
Time delay
Instant messaging
Access cost
Setup cost
Time delay
Web tools / wiki
Access cost
Setup cost
Time delay
Personal Access Grid
Access cost
Setup cost
Time delay

Figure 4