 ISSN 1751-5971

Multi-Site Videoconferencing for the UK e-Science Programme

A Roadmap for the Future of Videoconferencing within e-Science

Submission Date: 23rd October 2002

Contributors from:

Argonne National Laboratory, BAE Systems, California Institute of Technology (Caltech), inSORS Communications Inc, Rutherford Appleton Laboratory, UKERNA and the Universities of Cambridge, Edinburgh, Glasgow, Manchester and Southampton

Important Note

A previous version of this report was submitted on 27th August 2002. This version is now superseded.

Contributors

The following people have contributed to this report. Roles and chapter headings are in parentheses and indicate the main areas of contribution:

Stephen Booth – *University of Edinburgh* (Recommendations, reviews)

John Brooke – *University of Manchester* (Executive Summary, Recommendations)

Kate Caldwell – *University of Cambridge* (Recommendations, User Requirements and Human Factors, H.323 / H.320 Studio-based Videoconferencing, reviews)

Liz Carver – *BAE Systems* (User Requirements and Human Factors)

Michael Daw – *University of Manchester* (Recommendations, Access Grid Studio-based Videoconferencing, Non-studio-based Videoconferencing, Interoperability – Technical Issues & User Experience, Summary, report management & editing, reviews)

David De Roure – *University of Southampton* (Recommendations, reviews)

Alan Flavell – *University of Glasgow* (Recommendations, VRVS Studio-based Videoconferencing, Non-studio-based Videoconferencing, reviews)

Philippe Galvez – *California Institute of Technology (Caltech)* (Recommendations, reviews)

Brian Gilmore – *University of Edinburgh* (Recommendations, Executive Summary, Interoperability – Scheduling Systems)

Henry Hughes – *UKERNA* (Recommendations, H.323 / H.320 Studio-based Videoconferencing, reviews)

Ben Juby – *University of Southampton* (reviews)

Ivan Judson – *Argonne National Laboratory* (Recommendations, Access Grid Studio-based Videoconferencing)

Jim Miller – *inSORS Integrated Communications, Inc.* (Access Grid Studio-based Videoconferencing)

Harvey Newman – *California Institute of Technology (Caltech)* (Recommendations, reviews)

Chris Osland – *Rutherford Appleton Laboratory* (Recommendations, Data Application and Sharing)

Sue Rogers – *University of Cambridge* (Recommendations, User Requirements and Human Factors, H.323 / H.320 Studio-based Videoconferencing, reviews)

Contents

1 Executive Summary

The purpose of this report is to investigate the needs and requirements of advanced videoconferencing and distributed collaborative working in the UK e-Science programme. These can be regarded separately, in that it is possible to have videoconferencing without collaborative working and vice versa. However the nature of the advanced scientific distributed projects at the heart of the UK e-Science programme requires solutions that allow maximum access to both aspects of distributed meetings.

It is important to recognise that collaborations may be of many different sizes – sometimes small ad hoc meetings of a few people at a handful of sites; at other times meetings across many sites in different locations and timezones with prepared presentations and shared visualizations of large and complex simulations or data-mining searches. In between these extremes there will be a rich array of working environments and meetings, some of which are only just beginning to be explored.

Our purpose in this report was firstly to gather the experience of those working in this field in the UK and to allow this experience to be described as objectively as possible with the strengths and weaknesses of each technology and community approach being presented by those who have a depth of experience with the chosen solution. This is important because of the prevalence of anecdotal evidence in this field. A videoconferencing session that suffers failure either through weaknesses in technology, networking or preparedness of the participants is a very distressing experience. Since users of a particular system become familiar with it and incorporate it intuitively into their conduct of the session, they can often be very disoriented when placed in another session with different technology and accepted norms of conduct. They may conclude that the unfamiliar session indicates that the underlying technology is flawed, whereas if it is used correctly it can produce highly productive meetings. For this reason there are no clear conclusions to be drawn about which system is "better". Therefore, we focus on the different solutions in the context in which they were intended to operate. Nonetheless our survey has allowed us to produce some guidelines and a proposed roadmap for further development that we present later in this document.

We examine three studio-based solutions – Access Grid, H.323/H.320 with commercially supplied codecs, and VRVS. We also look at non-studio based videoconferencing and how this might interface with studio-based sessions. All these solutions have enthusiastic and growing user communities. Indeed, the one indisputable message from this survey is that the need for advanced videoconferencing and collaborative working is growing to the point where it is becoming an essential component of many e-Science projects. We find the increasing involvement of UK researchers in large international working sessions very encouraging, indicating that the UK is well-placed to make an active contribution to scientific collaborations that are truly global in scope, covering areas such as climate research, high-energy physics, astrophysics and space physics, chemistry and materials science research, bioinformatics and computational molecular biology. We consider that the UK is also well placed to make important contributions to the interoperability of videoconferencing technologies. This would be of great benefit to

UK projects and would also enhance the UK's reputation for innovation and leadership in e-Science.

Access Grid is the solution most directly targeted to large, complex and persistent collaborations. It is also intending to utilise the developing Grid standards as currently implemented in the Globus toolkit v2.0 for security, data transfer and integration of scientific applications. VRVS shares with Access Grid these large-scale ambitions and is also targeted to participants connecting from a range of hardware, from the studio to the laptop. In the relevant sections we discuss the reasons for these choices and the potential benefits and problems of both. A major difference of emphasis between them is that Access Grid has chosen multicast networking as its solution for the problem of scaling to many sites without producing unrealistic demands for wide-area bandwidth while the current deployment of VRVS uses unicast with a system of reflectors for the same purpose (but can support multicast when and if appropriate). It is already possible to integrate Access Grid and VRVS in the same virtual meeting; however, interoperability between them could be considerably improved. Both use the metaphor of Virtual Rooms for organising the connectivity of meetings.

An alternative to these software-based systems is the use of commercially supplied codecs using H.323/H.320 standards, together with a service for establishing the MCU connections between sites. We describe how this is organised over JANET in the UK and describe flourishing UK communities who are using this on a regular basis. VRVS can operate with the H.323 protocols and thus provides a bridging point for H.323 users into VRVS or Access Grid sessions. Again the interoperability is feasible but could be considerably improved. H.320 offers security and guaranteed bandwidth but at the cost of leasing the ISDN lines. As bandwidth on IP networks continues to increase, use of H.323 will continue to grow. This solution is the simplest to use and requires least maintenance but the hardware codecs restrict the number of video feeds that can be deployed and this means that all participants in large meetings cannot be seen simultaneously. Solutions such as switching according to audio strength are utilised (loudest talker wins!). Clearly this imposes a different set of constraints on the conduct of a meeting as compared to Access Grid or certain VRVS meetings.

We have included an important section on Human Factors in videoconferencing, including input from the Human Resources unit of BAE Systems. We consider this industrial perspective to be of great importance since the UK e-Science projects typically have several industrial partners and a key aim of the programme is to promote the uptake of e-Science in industry. We note that one of the EPSRC pilot projects, Reality Grid, has specifically referred to the use of Access Grid as part of its whole methodology and a key industrial partner, Schlumberger, is incorporating virtual collaborations as a key feature of its business methodology. We note that VRVS is used in the working methodology of the High Energy Physics community – one of the largest distributed collaborations in the world – and is also an integral part of COAKTinG, an e-Science project. We also note the great importance of H.323 videoconferencing in UK educational collaboration. This shows that the lessons drawn in the different sections of our report are based on demanding production quality environments where failures of the videoconferencing system would be very serious in terms of the project mission.

These considerations are discussed fully in the rest of this report. The Technical Summary chapter provides a more in-depth overview and the details are in the various sections. Here we now describe briefly the main conclusions and propose a roadmap for the e-Science programme. The justification for these statements is in the detail of the report given the guidelines we have described above.

1. No single solution can be imposed across the whole UK e-Science programme. We show that interoperability of all solutions described is technically possible and propose that the UK take a world-leading role in providing solutions and support.

2. A major cost of videoconferencing is the projectors, cameras, room alterations etc. These are common costs across all systems described and the interoperability provided in (1) will allow maximum advantage to be taken of these.

3. There is a need to integrate scientific applications with videoconferencing. Visualization of results from e-Science simulations will be crucial to this. This means that large amounts of data will flow across the core JANET backbone. Multicasting and VRVS reflectors provide the means for managing the scalability issues. Multicast-to-Unicast bridges are needed for sites that cannot support multicast. Multicast is currently available on the core Janet backbone. There are still issues of multicast connectivity between SuperJanet4 and the regional MANs and between MANs and site-specific LANs that can prove problematic. More investigation of these issues is required.

4. There is a need to coordinate UK efforts in this field and to provide input to the Advanced Collaborative Environments research group at the Global Grid Forum. There needs to be a forum for dialogue between those working on these issues in the Grid Community and those working in the Internet Engineering Task Force and in networking standards efforts.

5. We predict that the collaborative working aspects of e-Science and the Grid will become increasingly important and will be of equal importance to the work on computational and data Grids in the very near future. The needs of collaborative working will have a major driving role in the development of Grid services. We base this conclusion on extrapolation of current practice and also on discussions with large globally distributed companies, e.g. BAE Systems, Schlumberger, Boeing, Johnson & Johnson.

6. Our collaborations will increasingly involve participants for whom English is not their native language. Experience shows that audio quality is the single most critical factor in the success of such events but also that more visual context clues need to be provided (e.g. highlighting the window of the current speaker). Such meetings need carefully considered chairing and planning.

7. New sites/centres joining the UK e-Science programme should have access to funds for advanced videoconferencing. We suggest possible amounts in the Technical Summary and suggest providing choice as to how they deploy their resources. However they need to ensure that they will provide environments

and resources that meet minimum standards for the community they wish to interact with, and this should be a condition of their funding.

8. The UK should have a clearly identified point of access from which it can collaborate with efforts such as Access Grid and VRVS and participate in standards and protocol definitions for advanced collaborative environments. This could also provide a centre for queries and advice for UK collaborations and projects. By providing reference implementations and coordinating with commercial providers this could reduce the amount of local support required and avoid duplication of resources across many institutions. The Grid Support Centre currently performs such a role for Grid Middleware deployment for computational and data Grids.

9. It is recognised that running an Access Grid centre requires ongoing resources. Some UK centres and projects will be willing to regard their facilities as part of a research programme and will provide this support from their own resources. For others, advanced videoconferencing will be seen as a tool that they wish to be as low maintenance as possible. Commercial providers, either of codecs or else of customised software solutions (e.g. *inSORS*) can provide an attractive and reasonably priced option for such sites. Some centres and projects are already collaborating and contributing to projects such as Access Grid and VRVS. They should utilise the point of access described in (8) to coordinate such effort on behalf of the UK research community.

10. Although e-Science is currently defined as the projects and centres involved in or funded by the UK e-Science programme, the need for the collaboration tools we describe has a much wider base in the UK research community. Access to advanced videoconferencing will be of key importance in maintaining the ability to be at the cutting edge of scientific collaborations on a global scale. These considerations will also apply in the commercial sector.

2 Recommendations and Roadmap

- **Create an e-Science Advanced Collaborative Environments Research and Development Effort**
- **Formalise Access Grid Support**
- **Extend VTAS Specifically to Include Deployment Advice on Access Grid and VRVS in Addition to H.323/H.320**
- **Enable Full Interoperability between Access Grid & VRVS**
- **Enable Maximum Interoperability between Access Grid & H.323/H.320**
- **Deploy and Support Multicast Bridge(s)**
- **Reduce Access Grid Resource Implications by working closely with Commercial Vendor(s)**
- **Improve Local Networking in Support of IP-based Videoconferencing**
- **Investigate Improvements for Multi-Site Booking Systems**

These are recommendations related to multi-site videoconferencing for the e-Science programme. We have tried to make our proposals consistent with on-going organisations and solutions so as not to 'reinvent the wheel' and to make the most productive use of existing resources.

We note that the H.323/H.320 solution is currently well supported by JANET. The (mainly commercial) solutions for H.323/H.320 are, for the most part, working well and have a satisfied user base. Although VRVS has a support team situated at Caltech (California Institute of Technology), because there are 600 registered VRVS nodes in the UK and global usage of the system likely to continue to grow, there is a need for support to become localised for it to remain effective.

Access Grid has no formal support structure. We believe that the experience of Access Grid to date of the e-Science programme has been a good one and that the unique contribution that this technology can make to aspects of the programme means that this experience can only be enhanced by our recommendations for the establishment for it of a more formal support structure. Because of the current predominance of Access Grid throughout the programme, many of our other recommendations focus on enabling better interoperability between it and other solutions considered by this report.

However, our recommendations do not impose a specific solution on any future e-Science project or centre. We believe that each of the solutions will play a role. Different factors need to be considered when assessing the appropriateness of the solutions to requirements. A guide of some sample scenarios where different facilities may match different sets of requirements is given in *3.1 Matching Facilities to Requirements*.

2.1 Create an e-Science Advanced Collaborative Environments Research and Development Effort

Remote collaborations are becoming an increasingly important part of working life, enabling productive work to continue between individuals who may be widely geographically dispersed. The e-Science programme has benefited greatly from research efforts made by the Access Grid Project, led by Argonne National Laboratory, with contributors from around the world.

As the country with the second highest number of Access Grid nodes, an established international reputation and influential contacts, we believe that the UK is uniquely placed to make a valued contribution to research and developments in the field of advanced collaborative environments. This will not only be of visible benefit to the e-Science programme, but also to other UK and global projects. The effort should examine ways to improve the productivity of remote collaborations through an increased sense of presence. This might involve developments in the areas of video (towards standards such as HDTV), audio (involving multi-channel and virtualised audio), interoperability, and the integration of collaborative applications, among other possibilities. The aim is not to concentrate solely on a single technology, but to look at developing each of the technologies considered by this report.

We recommend that this effort be co-ordinated via a UK-based Advanced Collaborative Environments Research Centre (ACERC). This would have the dual role of liasing with groups such as the Futures Laboratory at ANL, the VRVS team, and the Global Grid Forum to ensure UK development is fully integrated with the global effort and also to co-ordinate development in the field within the UK.

We envisage that the ACERC would co-ordinate efforts such as extending interoperability between H.323 and Access Grid (recommendation 2.5), the integration of scientific applications into collaborative environments, and many other efforts in this area that are on-going, planned and potential. Part of this body's remit would include tracking new tools and standards, such as SIP, with regard to their possible application to advanced collaborative environments. Concrete developments from ACERC would feed directly into the UK deployment of remote collaboration facilities through bodies such as the Access Grid Support Centre.

2.2 Formalise Access Grid Support

We recommend the establishment of a formal UK Access Grid Support Centre (AGSC). We recommend that the JVCS extend its QA services to UK Access Grid nodes (in close liaison with the AGSC). The role of the AGSC will be to move the Access Grid towards a full service footing.

Whilst the initial rollout programme of 12 Access Grid nodes was accomplished with no formal support structure, the facilities are currently used in a manner that is mission-critical. Without a formal support mechanism, quality cannot be assured across all sites, especially when the number of Access Grid sites in the UK is likely to increase over time. We note that there have been certain high-profile events that may have benefited from the existence of formal support services.

As UK usage of VRVS grows, a more localised support structure is likely to become necessary. The AGSC should look to be able to take on aspects of VRVS support for UK users as an offshoot of the central VRVS support team.

The AGSC would take on the following roles:

- Establish important links between existing support structures (i.e. JANET Video Technical Advisory Service [VTAS] and operators of other videoconferencing services like the JANET Videoconferencing Service [JVCS] and VRVS support team)
- Operate the Multicast Bridges (which provide conversion between unicast and multicast environments) (recommendation 2.6)
- Provide technical help and advice with scheduling and running Access Grid conferences
- Provide help and support in scheduling and interoperation between videoconferencing technologies
- Provide advice on recommended hardware, software and configuration
- Establish and promote good practice guidelines (both in terms of the technical operation and in terms of usage of the facility [see *3.2 Human Factors*])
- Run a series of workshops and one-day events to help promote best practice in the use of Access Grid technology
- Support for rollout of technology upgrades and improvements (e.g. support of better quality video and audio; integration of Access Grid with other Grid technologies – this would require close liaison with the Grid Support Centre)
- Co-ordination with commercial providers (e.g. *inSORS*) (recommendation 2.7)
- Investigate form and provision of VRVS support for UK in consultation with VRVS support team at Caltech

The AGSC would use the JANET Videoconferencing Service (JVCS) as a 'best practice' model and closely collaborate to exchange ideas for improving the level of support across both services.

We believe that Access Grid sites would greatly benefit from Quality Assurance (QA) testing. The recommendation that the JVCS extend its QA services to Access Grid nodes in close liaison with the AGSC represents a number of advantages over the AGSC taking on this role. Firstly, there is no need for duplication of a parallel infrastructure – in both equipment and staff – that would entail a considerable cost implication. Secondly, there is a vast amount of skills and experience in this area that exists at the JVCS. To leverage this for Access Grid QA testing would greatly benefit Access Grid users and would help to inform Access Grid technology development.

2.3 Extend VTAS Specifically to Include Deployment Advice on Access Grid and VRVS in Addition to H.323/H.320

Projects, teams and centres have different requirements of facilities to enable remote collaboration. It is not appropriate for us to recommend one type of facility over another for a whole programme when the requirements for disparate parts of that programme are not currently known and may well differ. However, there does need to be a well-defined process of determining the appropriateness of technologies and equipment for a given set of requirements so that this can translate directly into deployment decisions.

There already exists a centralised service that provides this type of advice – UKERNA's Video Technology Advisory Service (VTAS, http://www.video.ja.net). However, this service mostly limits itself to H.323/H.320 systems. We recommend that the scope of this service be extended to include other categories of videoconferencing covered by this report, namely Access Grid and VRVS. This would enable VTAS better to match clients' requirements to possible deployment options, especially by e-Science projects/centres.

We envisage that pre-purchase advice would continue to be impartial and would include product tests, notes on studio set up and technical assistance. Also, importantly, VTAS would provide a central point from which site visits of Access Grid nodes, H.323/H.320 studios and VRVS facilities, etc., can be arranged. We believe that this type of service is vital to anyone who must make informed deployment decisions for videoconferencing solutions.

For example, an e-Science project may need a facility that enables them to perform large multi-site international collaboration on Grid-type activities. This project may only have limited funds available for a collaborative environment. VTAS would be able to help them make an initial assessment of which type of facility may be most appropriate (Access Grid, H.323, H.320, or VRVS) and direct them to further advice (e.g. whether and which peripheral devices are required, what level of hardware is necessary) and arrange site visits of the different types of facility.

We would envisage the Access Grid Support Centre (recommendation 2.2) to be contracted by VTAS to cover advice for Access Grid. (This is similar to the distributed manner in which the service currently offers expert advice.) Methods for incorporating expert advice for VRVS require further investigation, but will probably include contracting existing VRVS support mechanisms or the AGSC (recommendation 2.2).

We also recommend that publicity for this service be improved to ensure that its usage is widespread and productive.

2.4 Enable Full Interoperability between Access Grid & VRVS

We recommend extending interoperability between Access Grid and VRVS: by extending the provision of gateways, and by enabling Access Grid nodes to participate natively in VRVS sessions.

Improve VRVS to Access Grid Interoperability

VRVS users may currently join Access Grid sessions. However, they are restricted by the fact that there is only one server supporting this connectivity, that this is sited in the US, and that it only supports a subset of Access Grid Virtual Venues.

We recommend that an investigation be performed with the aim of deploying an additional VRVS-Access Grid server, to be sited in the UK. This would require liaison with the VRVS team to determine deployment issues (e.g. who takes responsibility for its support and any technical issues) and any further development that may be required to interact with the existing server and to support additional Virtual Venues.

Enable Access Grid to VRVS Interoperability

It is not currently practicable for an Access Grid node to participate in a VRVS session, because the VRVS software installation assumes that the audio component resides on the same physical server as the display component. In a standard Access Grid configuration, the audio component resides on a separate physical server from the display component.

There may be a number of possible solutions that would result in Access Grid nodes having a dual role as VRVS studios. We recommend that these be investigated so that future investment in Access Grid results in the added value of nodes being able to participate fully in what are predominately VRVS sessions.

Establish Common Code Base for Access Grid and VRVS

Access Grid and VRVS both base their implementations on the Mbone software tools *vic* and *rat*. Both communities have made developments to these applications and now operate with separate implementations. We recommend that progress towards a common code base for the Mbone tools be established, so that developments for one set of users can benefit all. Progress towards this unified effort should be made through an appropriate mechanism, e.g. the Global Grid Forum and/or ACERC (see recommendation 2.1).

2.5 Enable Maximum Interoperability between Access Grid & H.323/H.320

Interoperation between H.323 and Access Grid is currently possible via separate gateways supported by VRVS and by JANET. However, interoperation between the two technologies is challenging because of differences between the ways Access Grid (Mbone tools) and H.323 are typically used (e.g. for many H.323 studios, continuous video presence may result in images that are too small at the H.323 end if monitors are used; the alternative – voice-switched mode – presents difficulties for interoperation because Access Grid does not currently match outgoing audio and video channels from each site).

We recommend the creation of an integrated production quality collaboration environment to enable the maximum possible interoperation between H.323 and Access Grid and overcome the problems with the existing gateways. A full investigation into technical solutions and their impact on user experience is required.

Some possible solutions are:

- Enhance and deploy widely VRVS/AG Gateways (called Virtual Access Grid [VAG]) which support interoperability between the Mbone clients (used by Access Grid) and H.323

- Fund a UK contribution to the OpenH323 project (http://www.openh323.org), which is committed to the collaborative development of an Open Source H.323 protocol stack. This project includes OpenMCU, which is an open source implementation of an H.323 conference server. Our contribution might involve development on OpenMCU in order to have a direct integration into the Access Grid Venues Server

It must be stressed that these possible solutions are used only to illustrate the kind of effort required. The actual solution will depend upon a full investigation prior to any development being carried out.

In the mean time, the current H.323/Mbone gateway offered by JANET could be a useful short-term solution and could also provide a testbed for development of H.323/Access Grid interoperation. However, the current gateway requires some further development work to enable it to be useful for actual H.323/Mbone meetings. Issues arising from this development work – both in terms of technical issues and in terms of users' experiences – would feed directly into the development of the integrated production quality H.323/Access Grid collaboration environment. We therefore recommend that this gateway be developed, deployed, tested and that user surveys are conducted as a first step towards the longer-term production solution.

2.6 Deploy and Support Multicast Bridge(s)

There are difficulties in enabling multicast to all parts of the UK academic network. These are unlikely to be fully overcome in the near future. Deployment of Access Grid is set to become more widespread. We would not wish there to be a situation where certain projects or sites may determine that Access Grid is the most appropriate technology for them to deploy (perhaps via VTAS, see recommendation 2.3) but that there are difficulties with implementing multicast on their institution's Local Area Network, and therefore connectivity is problematic.

Currently, ad hoc solutions are applied, for example by using a bridge supported by Argonne National Laboratory or requesting a bridge to be run by a nearby multicast-enabled site. However, neither of these solutions is scalable nor is appropriate with offering support of Access Grid on a service footing.

We recommend that Multicast Bridges be deployed on the JANET core network and be supported by the Access Grid Support Centre (recommendation 2.2). The exact implementation of these bridges requires further investigation to determine the basis for a solution. For example, likely candidates include the multicast bridge software currently in use by the Access Grid community (*QuickBridge*) and VRVS multicast bridging software. It is important that any solution employed be compatible with AG2.0. This investigation and subsequent deployment should proceed quickly, as a production solution is required as soon as possible.

2.7 Reduce Access Grid Resource Implications by Working Closely with Commercial Vendor(s)

As a service, Access Grid has been criticised for being resource-intensive, in terms of the initial procurement and installation effort, in the need for an operator to be present during sessions and in the level of on-site support that is necessary for the facility.

To go a long way to remedying this situation, we recommend a closer relationship with appropriate commercial vendor(s) (e.g. *inSORS*). *InSORS* supply a range of 'off-the-shelf' Access Grid solutions at around the same price it would take to self-build. This would obviate the need for institutions to undertake a large procurement and installation effort. *InSORS* provide a hard- and software support service that is included in the price. *InSORS* have also re-written the Access Grid user front-end to make it far more usable so that the facility may be operated by non-expert end-users for most sessions. In addition to this, the company is fully integrated within the Access Grid community both in terms of contributions they make and also in terms of their plans for integration of future developments of Access Grid technology.

We recognise that this solution will not be suitable for those sites that wish to develop advanced collaborative environments based on the Access Grid model; nevertheless, it goes some way to answering those who criticise the Access Grid on the grounds stated above.

Existing Access Grid nodes can be 'retro-fitted' to use the *inSORS* front-end with some minor and inexpensive modifications. (These modifications have the incidental beneficial side-effect of enabling Access Grid nodes to become VRVS studios in addition to their primary role. [See recommendation 2.4.])

2.8 Improve Local Networking in Support of IP-based Videoconferencing

IP-based videoconferencing is especially vulnerable to variations in available bandwidth that lead to a percentage of packet loss. Networking for IP-based videoconferencing often fails within end institutions where bandwidth may not be sufficient to support videoconferences simultaneously with other network traffic (see *5.7 Networking Issues*).

It is necessary to achieve adequate performance for IP videoconferencing on the local network, in terms of network hardware (to provide ample bandwidth capacity) and also in the use of appropriate network engineering.

We recommend an investigation into focused local network improvements, appropriate QoS solutions, and other possible solutions aimed at improving the quality of IP based videoconferencing sessions, to lead to the subsequent implementation of these solutions.

2.9 Investigate Improvements for Multi-Site Booking Systems

We recognise that there is a problem in the co-ordination of booking systems between and within videoconferencing communities. None of the booking systems used by the videoconferencing technologies considered by this report is perfect and interoperation between them is likely to be very difficult. There is also a problem of the interaction between any booking system that exists to enable multi-site booking and the physical room booking system that exists within an institution.

We believe that there is little to be gained from trying to invent a new booking system for the UK communities. In fact, whilst this may solve some problems in the short-term, it is likely to exacerbate the situation in the long-term. However, there may be something to be gained by investigating the possibility of standards-based systems. The authors of this report are not aware of any standards that exist to facilitate the design of multi-site booking systems. This may be an area in which the UK could take a lead. We recommend that investigation be initiated on improving systems for booking multi-site conferences, possibly including the formation of a booking system standard within the context of videoconferencing standards.

3 Technical Summary

This section compares and contrasts the findings from the five core chapters on videoconferencing systems and interoperability between them.

3.1 Matching Facilities to Requirements

Our recommendations do not impose a specific solution on any future e-Science project or centre. We believe that each of the solutions considered by this report will play a role. Despite the obvious theoretical attraction of trying to select a single solution, the practical reality makes it inevitable that each of the videoconferencing solutions considered in the report will be present on occasion, and this needs to be taken into account both in deciding how to equip conferencing facilities, and in deciding from case to case how to host a particular meeting. When setting up a meeting, it is clearly desirable to host it by a single technology throughout, if feasible; but on the other hand, some would-be participants may be unreasonably disadvantaged by an inflexible approach to mixed meeting arrangements. The following scenarios explore some likely meeting profiles and indicate feasible (but non-exclusive) solutions.

Scenario A

Requirements: Large, multi-site group-to-group collaborations, where persistence of collaboration and a sense of shared presence is important.

Suggested solution: Access Grid; Studio-based VRVS

Scenario B

Requirements: Meetings that involve a high degree of integration with specialist collaborative tools, such as visualisation applications.

Suggested solution: Access Grid; Studio-based VRVS

Scenario C

Requirements: Collaboration between groups for whom integration with other Grid technologies (e.g. using Globus certificates) is important.

Suggested solution: Access Grid; VRVS

Scenario D

Requirements: Plug and play solutions for small and large meetings with minimal technical knowledge required and where quality of video image and stability are critical.

Suggested solution: H.323/H.320

Scenario E

Requirements: Communication with other H.323-enabled academic institutions and/or H.320-enabled business partners.

Suggested solution: H.323/H.320

Scenario F

Requirements: High degree of privacy (e.g. for patent discussions, medical records, beta developers).

Suggested solution: presently, H.320: in future, secure/VPN

Scenario G

Requirements: Small one-to-one meetings from laptop or office with no need for peripherals.

Suggested solution: Non-studio-based videoconferencing: desktop H.323 clients may use point-to-point; vic/rat clients, or mixed, may use VRVS

Scenario H

Requirements: To be audience at home or in office environment (especially at unsocial hours) for distributed seminar broadcast from an Access Grid/VRVS equipped seminar/lecture theatre.

Suggested solution: VRVS

Scenario I

Requirements: Informal working meetings between developers working from home or own desks with some data sharing.

Suggested solution: VRVS

3.2 Costs

Comparing costs between the various studio solutions is not straightforward. Issues of initial outlay on hardware, on-going support, resourcing, call charges and centralised costs all play a part. It is assumed that all the various studio-based solutions in this report have a similar resource implication and therefore only the comparative initial outlay costs of hardware are considered here.

Large portions of each of the different studio-based videoconferencing solutions considered by this report may be built using similar hardware. For example, the same models of projectors, cameras, microphones and many peripheral devices can be used whether the studio is to be predominately Access Grid, H.323/H.320 or VRVS.

In order to compare like with like, it is best to consider the core, non-transferable parts of the hardware (the PC servers and audio hardware in the cases of Access Grid and VRVS, and hardware codecs in the case of H.323/H.320).

A typical H.323/H.320 codec with built-in level balancer and echo canceller costs around £10,000. Lower speed H.323/H.320 codecs with less peripheral integration possibilities can cost £3000, whilst top-end H.323/H.320 codecs, offering multipoint bridge facilities across IP and ISDN simultaneously, cost up to £20,000. All of these codecs include one microphone and one camera (which are integral in the case of the low-end codecs).

The PC servers, level balancer and echo canceller for an Access Grid node cost around £7000. (It should be noted, though, that the minimum specification for a full

Access Grid node specifies a higher number of cameras, microphones, projectors, etc. than is usual for an H.323 facility. This may offset the savings made over H.323 for the core hardware.)

The core hardware for a VRVS studio solution is probably around the same cost as for an Access Grid node for a facility that offers similar quality results.

The Access Grid Personal Interface to the Grid costs £9K (NB as well as core hardware, this cost includes all necessary equipment, e.g. cameras, microphones, etc.).

The cheapest videoconferencing solution examined by this report is clearly the non-studio-based option. Connectivity can be achieved from around £50 (as long as existing hardware, like a desktop PC, is not included in the cost). However, greater quality can be achieved by adding relatively inexpensive echo cancellation that can cost around £100, hardware codecs (which include echo cancellation) that cost from around £400 and improved audio hardware. Of course, this removes some (but not all) of the cost differential advantage that this solution has over studio-based solutions.

3.3 Ease of Use

The H.323/H.320 solution is known as the easiest to use solution amongst those considered by this report, as it is usually operated from a single remote control at each site. At the other extreme, Access Grid has a reputation for requiring the permanent presence of a trained dedicated operator.

However, the real situation is not quite as simple as this. Typically, H.323/H.320 solutions are used with the participation of fewer sites and less complex scenarios than are attempted with Access Grid. Also, most Access Grid sites in the UK use software that comes from a research and development project rather than software that is commercially available. The *inSORS* software is more usable in terms of installation, configuration and in-conference use than the software package administered by Argonne National Labs. It has a usability that is on a par with an H.323/H.320 remote control. (This has been achieved by the integration of audio and display controls from the same screen and by providing the facility easily to control the volume of individual remote sites.)

If complex meetings are attempted (complex in terms of number of sites, peripheral devices, collaborative software, etc.), then the presence of an operator may be desirable with any of the technologies considered. If simple meetings are attempted (few sites, non-high-profile, simple or no data sharing), then operators are not necessary with any of the considered solutions. This is particularly the case with experienced users. Users who are not used to using videoconferencing of any kind will always require the presence of some kind of support, at least at first.

Technical back-up support should always be available for each of the studio-based solutions, in order to remedy unexpected problems to do with the audio-visual hardware or network issues.

When VRVS is used with the Mbone tools, comments about ease of use of Access Grid and VRVS usually apply equally to both (except where the commercial version of Access Grid supplied by *inSORS* is considered). VRVS can also be used with H.323 software; this has a similar level of usability to the Mbone tools (*vic* and *rat*).

The operation of non-studio-based solutions is usually within reach of anyone who is computer-literate. This is due to the focus of design of the associated software to home and business users and its utilisation of readily available, non-specialist hardware. However, with the addition of devices to improve quality, operation becomes less easy, though still within the reach of non-expert users.

3.4 Display Quality

Access Grid

Regardless of the number of participating sites, Access Grid can display all sites at the same time (as long as there is sufficient display space). Video feeds are displayed in windows that can be shown in a number of different sizes for different purposes (e.g. life- or larger than life-size for presenters, medium for remote audiences, small for local camera views).

H.323/H.320

Proprietary H.323/H.320 systems utilise video feeds that can be near to broadcast quality, depending on the hardware used to capture and display these feeds. If data projectors are used, then the effect can be very impressive. However, only one or two video streams are typically displayed. Therefore, if multiple cameras or sites are used, then each feed must be displayed within the same video stream. For example, if there are only one or two remote sites, they can be shown full-size. To show 4 sites simultaneously, each site is displayed at one quarter of the full-size single display. This is known as *continuous presence mode*. Otherwise, the video is voice or user selected and feeds can only be shown singly. This is known as *voice activated mode*. Therefore, display space can be a constraining factor, especially when data sharing applications or devices must be displayed in addition to video feeds.

VRVS

VRVS can either be used in H.323 or Mbone modes. When used in H.323 mode, then the display is similar to that described for H.323/H.320. When used in Mbone mode, the display is similar to that described for Access Grid.

Non-studio-based Solutions

Non-studio-based solutions have a display quality dependent upon the type of monitor in use. The typical use of a single monitor as a display device severely limits the quantity of video feeds that may be displayed simultaneously, especially when there is also some form of data sharing in the collaboration that must also be shown. The use of a small display device also limits the size of each video feed. If too many feeds are shown simultaneously, they will be too small to be easily viewable.

3.5 Visual Quality

Access Grid nodes send four video feeds from each site, although only two or three usually have useful content. H.323/H.320 nodes usually send only one (sometimes two) video feeds from each site. VRVS studios can send any number of video feeds, but typically send only one. Non-studio-based solutions rarely, if ever, send more than one video feed per site.

Access Grid / VRVS

Access Grid and VRVS (when using *vic*) have a video quality that is subjectively good when displayed in medium or small windows but can look pixellated when displayed in a large window. Access Grid and VRVS video feeds are typically broadcast at 25 frames per second. The compression algorithms used in sending video data can result in a reduction of video quality. This manifests itself in degradation of the image when displaying fast movement and also in blocks of old data showing on the periphery of video feeds that can lead to strange temporary effects, such as disembodied hands.

H.323 / H.320

The visual quality for proprietary H.323/H.320 solutions can be near to broadcast quality if sufficient bandwidth is utilised.

Non-studio-based Solutions

The visual quality for non-studio-based solutions is usually subjectively judged to be poor relative to studio-based solutions. Partly, this is because relatively inexpensive cameras are used, but also it is because of the fact that the outgoing video of users is typically that of a person staring at a computer monitor, which does not lend itself to a realistic 'sense of presence'.

3.6 Audio Quality

There is not much to choose between the high-end H.323/H.320, high-end VRVS and Access Grid solutions when considering audio. Each of these utilises echo cancellation and results in a full-duplex audio quality that is close to that experienced in real life.

Despite using H.323 networking, non-studio-based solutions tend to have an audio quality that is worse than the audio quality of studio-based H.323. This is because of a combination of having low specification computers, low-quality software codecs, and using poor quality speakers and microphones. Addressing these factors can greatly improve the audio quality of non-studio-based solutions.

The effect on a meeting that involves just one site with sub-standard audio can pervade throughout all sites and detract from the productiveness of the collaboration as a whole.

3.7 Networking Issues

Access Grid

The Access Grid in its pure form utilises IP multicast. Multicasting provides a scalable solution for large-scale collaborations over IP. However, the deployment of multicast can be problematic both on a local scale between institutions and also globally between continents. Whilst it may be straightforward were the IP network to consist of homogenous types of routers and if qualified network support was ubiquitous, this is not currently the case, especially at smaller institutions. Therefore the full deployment of multicast across all UK academic LANs may be some years off. The workaround solution for sites that are not multicast-enabled is the use of multicast bridges (this is already widely-used). Currently, these bridges are deployed on an ad hoc and voluntary basis at multicast-enabled sites. If the e-Science programme were to deploy Access Grid facilities more widely, there would be a need for a more formerly supported solution.

H.323 (including Non-studio-based Solutions Using H.323)

The H.323 solution uses the standard unicast IP network. A common perception among JANET users is that H.323 is vulnerable to packet losses and mis-ordering (more so than vic/rat as used by Access Grid); if available bandwidth suddenly falls during an H.323 session, then calls can drop out. It should be stressed that this is a feature of implementations of the H.323 standard rather than the standard itself. This may be ameliorated in due course by deployment of network QoS features, or H.323 implementations designed to be more resilient.

H.320

The H.320 solution relies upon the installation of dedicated ISDN lines at institutions, usually six channels to achieve the required 384kbps call quality.

VRVS

VRVS relies on sufficient numbers of reflectors being deployed throughout the network to support the number of facilities using that part of the network. What determines a 'sufficient' number of reflectors depends upon the aggregate bandwidth of all clients connected to any particular reflector at any one time. Were the e-Science programme to make a greatly increased use of VRVS, it would be necessary to supply additional reflectors. VRVS reflector supports unicast/multicast between reflectors, and unicast/multicast between reflector and client application. Therefore, VRVS could be used to create unicast tunnels where needed to interconnect multicast domains. This has been demonstrated to be highly scalable.

3.8 Multi-Site Issues

Access Grid

Access Grid is particularly well suited to multi-site conferences because of its utilisation of multicast, the use of *vic* and *rat* software and the large display wall. The UK Engineering Task Force regularly has successful meetings over the Access Grid involving 12 sites simultaneously (about 2-4 people at each site and 2 or 3 video feeds from each site) that are almost as productive as if they were conducted face-to-face.

H.323 / H.320

Higher-end H.323/H.320 codecs incorporate MCUs to accommodate multi-site conferences up to a maximum of four or five sites. For larger multipoint conferences, an MCU bridge service is employed. JANET have run an H.320 bridge for many years and are currently running a pilot H.323 bridge. Both of these bridges incur no costs at point of use by academic institutions. There are also several commercial organisations offering H.323/H.320 bridging, but these are not free to use.

If large numbers of sites participate in an H.323/H.320 conference, then there are issues to do with the display size of each video feed, or whether more than one feed may be displayed simultaneously at all (most H.323/H.320 multipoint calls use voice-switched video because of the limits set by the display size of each video feed). For more discussion about this issue, see *3.4 Display Quality*.

VRVS

Like Access Grid, VRVS is also often used for large multi-site conferences and also performs well, though individual user's subjective experiences of such conferences are dependent upon the size of their display hardware and the method by which they connect. Connection in voice-switched or user selection mode gives a severely limited view of participants; connection in multi-video mode gives users a similar experience to users of the Access Grid. VRVS runs on average around 15 worldwide multi-site international meetings per day.

Non-studio-based Solutions

Non-studio-based solutions may not be well suited to multi-site conferences, owing to the lack of likely display space and the possibility of having audio that is not echo cancelled and may not have full duplex. However, there are many examples of successful multi-site conferences held with most users participating from laptops or office-based desktops. To reiterate what was said in the core chapter on this subject, some users prefer the informality of participating from office or home and timezone differences may make the use of a studio impossible. However, multi-site conferences involving groups or that need intensive use of specialised collaborative software or peripheral devices are still best performed from studios.

3.9 Collaborative Tools

Peripheral Devices

Peripheral devices are pieces of hardware that aid various aspects of remote collaborations, such as Document Viewers, VCRs, proprietary electronic whiteboards and laptop computers.

Most peripheral devices can be fitted seamlessly to Access Grid, H.323/H.320 and VRVS facilities. The peripheral devices considered by this section are those that primarily output video.

In Access Grid facilities, peripheral devices usually take the place of one of the video feeds (any audio component to the device is routed through the Gentner echo canceller). The method of switching between camera and device is dependent upon how the facility is configured, but is best catered for by the use of an external video switch box.

In H.323/H.320 facilities, the source is switched using the remote control to display output from these devices. The higher-specification codecs have several video feeds and all except the integral codecs can be supplemented by adding an external video switch box.

VRVS facilities can manage the integration of peripheral devices in a similar manner to that of Access Grid facilities.

Non-studio-based solutions may or may not be able to integrate peripheral devices. Any attempt to do this must be done on an ad hoc basis. However, support for peripheral devices is unlikely to be great for this type of solution, compared to studio-based solutions.

Collaborative Software

The Access Grid has at its heart the integration of collaborative software. The virtual venues model upon which the system is based is partly a concept of persistent objects (e.g. data repositories and visualisation tools) being present in virtual spaces. This element of Access Grid will come much more to the fore in future releases of the technology. Software that is currently widely used for collaborations within Access Grid includes the MUD (a text-based chat tool for side conversations), Distributed PowerPoint (for sharing presentation slides), VNC (for sharing computer desktops) and *vtk* (Visualisation Toolkit). However, many groups are developing other instances of collaborative software for integration within the Access Grid.

H.323/H.320 solutions do not support a wide range of collaborative software. A common approach is to integrate a computer as a peripheral device on which is run VNC to share the whole desktop or MS NetMeeting to share specific applications.

VRVS commonly use a Java-based text chat tool for side conversations and VNC for sharing computer desktops.

Non-studio-based solutions may use application-sharing tools that are not specific to any videoconferencing solution, such as MS NetMeeting and VNC.

3.10 Security

Few methods of communication are completely secure. It is also often the case that more security is gained at the expense of greater inconvenience. Therefore, meetings that are held over insecure media (and most media are insecure, to differing degrees) should be assessed for their possible interest to potential snoopers. It may be felt that many meetings that are held within the e-Science programme will be about issues too complex or uninteresting for casual snoopers to bother eavesdropping. Therefore, there is little point in going to great lengths to ensure these meetings are secure. However, some collaboration within the programme involves issues of Intellectual Property Rights (IPR) that have a necessity of privacy. There are still decisions to be made about the degree of security that is required for such meetings (because of the increased level of inconvenience that it can entail).

The subject of security and the various sub-issues that are subsumed within it (integrity, authentication, confidentiality, non-repudiation, authorisation, etc.) is too involved for the issue to be treated in anything but a very shallow depth in this report. This section is merely a broad overview of the different videoconferencing solutions and their varying approaches to security.

Access Grid

The Access Grid system currently approaches the security issue by having 'secure' virtual venues. If a meeting is held within this type of venue, then video and audio data are automatically encrypted. Additionally, an Access Control List (ACL) limits access to secure virtual venues. These two elements provide a reasonably high level of security. For an even greater level of privacy, the public telephone system can be used for the audio element of the collaboration. With the future incorporation of Grid technologies into the core of the Access Grid, this technology should benefit from advancements in security expertise within the general Grid community. There is also ongoing research into secure multicast networking that will also improve the situation for private meetings held over the Access Grid.

H.323

H.323 conferences are open to potential eavesdropping, as data streams are usually not encrypted and travel over the IP network. There are various ways to minimise risk (especially so that administrative passwords are not compromised), but many users of H.323 will not use this technology if security is an important consideration.

H.320

H.320 is inherently a much more secure system because audio and video travel over dedicated ISDN lines. These are difficult to tap because the networks are private. Because the hardware codecs that support H.323 conferences usually also support H.320, it is often the case that users of H.323 will switch to H.320 if a high level of confidentiality is required, for example, when discussing patients' medical histories.

VRVS

All VRVS users should be registered to the VRVS server in order to use the services. Therefore the VRVS administrator knows in realtime who is connected where, and can take any appropriate actions (delete, add, ..) in regards of security. Admission to a VRVS session is optionally password-controlled. This will prevent casual snoopers, but anyone who can obtain access to the data packets containing video and audio data will be able to eavesdrop the meeting. The *vic* and *rat* applications have a capability of encrypting streams, so if all users are using these applications within a particular VRVS meeting, it is possible for it to be secure in this respect. However, this feature is not integrated within VRVS. If one or more participants are joining via H.323, this ad hoc method cannot be applied. The VRVS team recognise that security is an important aspect for global collaboration and are currently working on major enhancements in this respect which will include methods for traversing firewall and using VPN (Virtual Private Network) software to encrypt data.

Non-studio-based videoconferencing

Different non-studio-based solutions take different approaches to the security issue. For example, MS NetMeeting allows the option of making encrypted calls. However, such a call does not encrypt video or audio, so only the data-sharing aspects of this application are available encrypted. NetMeeting does allow for meetings to be accessible only to participants who have a pre-determined password, which will prevent casual snooping, but is not appropriate when a high level of security is required.

Also, since non-studio-based videoconferencing utilises H.323, similar issues apply.

3.11 Appropriate Usage

Access Grid

The Access Grid is most suited for group collaborations, perhaps that require the use of specialised collaborative software. For example, it is well placed to support the shared development, debugging, execution and performance monitoring of Grid-based applications. It has also proved its worth in supporting conferences between large numbers of sites. The range of different cost solutions with Access Grid makes this technology accessible to a variety of institutions.

H.323/H.320

H.323/H.320, with its current support of high quality video is most suitable for collaborations that have a large visual component and where collaboration between large numbers of sites is not so important. The University of Cambridge are currently embarked on a project that utilises a proprietary H.320 solution for doctors to consult with patients and colleagues over a wide area. In this type of situation, it is important for body language and other non-verbal cues to be interpreted correctly between people who may never have met. This makes high quality video indispensable.

VRVS

Because VRVS can have a spectrum of hardware implementations from low- to high-end, it has a corresponding range of appropriate usages. As it also supports H.323 facilities and an H.323/Mbone gateway, it has the potential for hosting conferences that require significant interoperation. As with Access Grid, it is suitable for conferences between large numbers of sites.

Non-studio-based Solutions

Non-studio-based solutions are well suited for informally arranged, one-to-one meetings where it is useful for participants to see each other or to conduct limited data sharing. They are also useful as supplementary to studio-based facilities for clients to use when studios are unavailable or inconvenient, such as when large timezone differences are involved.

Non-studio-based solutions are not appropriate when most other participants are studio-based (unless used as a fallback option when there is no other option), when a large element of the meeting involves collaborative tools, or when a high level of audio and video quality is required (for example, the conference is high-profile).

3.12 Future Potential

Access Grid

The roadmap of the future of Access Grid may be seen as four complementary strands:

- To integrate more fully with Grid computing tools for security, data management and GridFTP

- To be a fully extensible environment, so that developers have specifications and tools for building integrated services and a framework in which services can be found and utilised within the Access Grid

- To extend the audio-visual experience. Video to include support for higher resolutions, more frames per second and High-Definition TV support. Audio to include multi-channel, virtualised audio support

- Utilisation of improved multicast protocols, automated network failure recovery and streaming event notifications from network devices and services.

H.323

H.323 videoconferencing has issues that require resolution to improve the quality of user experience. A roadmap of H.323 development would include:

- Enhanced security

- Resolution of Quality of Service (QoS) issues

- A rationalisation of IP addressing and gatekeeping.

H.320

H.320 has great strengths in the areas of QoS and security. However, users pay a premium price for this with ISDN call charges and line rental. As IP bandwidth increases and as and when security issues are addressed in H.323, more H.320 users will switch to H.323. This move is already foreseen and indicated by the proprietary hardware codec manufacturers, most of whose codecs now support both H.323 and H.320 protocols.

VRVS

The VRVS team plans many developments. Imminent developments include:

- User authentication
- A massive increase in the number of virtual rooms
- Packet recovery
- Chair/centralised control of audio and video
- The selection of bandwidth ranges.

Further in the future, there will be:

- More support for additional formats such as MPEG2 and High-Definition TV
- More support for future Internet protocols such as SIP and IPv6
- Architecture enhancement to support several thousands users and/or conferences in parallel
- Automatic network failure detection and automatic rerouting to the best VRVS reflector.

Non-studio-based Solutions

A definitive roadmap on the future on non-studio-based videoconferencing solutions relies upon the direction that commercial organisations that develop such solutions are willing to take. However, the fact that these solutions are, by definition, non-studio, means that they have limited future potential for the types of intensive, collaborative, multi-site conferences for which the e-Science programme will require facilities.

3.13 Interoperability

All the technologies considered by this report play a vital role in multi-site collaborations within the e-Science programme and will continue to do so:

- The telephone enables participants who cannot easily join a videoconferencing studio and is also a backup solution if other technology fails
- Access Grid has good current and future support of collaborative tools; a potential for high level of integration with general Grid technologies; is suitable for good quality, large multi-site meetings; is set to have a wide influence on the future of advanced collaborative environments; is now an established and well-used technology within the e-Science Centres

- H.323/H.320 systems dominate in industry. This is an important point, because industrial partnerships are the norm in e-Science projects

- VRVS is widely used, especially among particular communities, for example, High-Energy Physics.

Interoperability between these systems is important so as not to exclude certain users.

Interoperability with the telephone system is readily achievable with all the systems under consideration by this report. With H.323/H.320 and Access Grid, this is provided by the standard implementations of these technologies. Many facilities using VRVS have the capability to interoperate with the telephone already and the rest could achieve it with a relatively small amount of financial outlay and effort.

Interoperation between Access Grid and H.323/H.320 is currently possible. However, it is not straightforward technically, as it relies on VRVS gateways (in addition to H.323/H.320 gateways for interoperation with ISDN). Also, for the user, it results in an experience that is not as good as can be achieved using one or other of the technologies on its own. See section 2 for our recommendations to improve this situation.

Interoperation between Access Grid and VRVS is also important so that users of each kind of facility are not excluded from conferences that are predominately based on the other technology. Users of VRVS can currently join most Access Grid conferences, as long as they are held in certain Virtual Venues. However, the VRVS bridge that supports this connectivity is not ideal – its siting at a single physical location in the US means there is a noticeable lag for users in the UK and it is unlikely to be able to cope with a large increase in load. We make a recommendation in section 2 to resolve these difficulties.

Access Grid nodes supplied by *inSORS* are configured so that they are VRVS-compatible. However, the Access Grid configuration applied at the e-Science Centres separates audio from video functions, leading to difficulty participating directly in VRVS sessions: measures for resolving this minor issue are discussed in 9.2. The use of different (usually incompatible) booking systems and possible reductions in quality for all or some participants means that interoperation between technologies is currently to be avoided in most cases. However, interoperation remains an important consideration for certain scenarios and we anticipate demand for this to increase with time. Because of this, some of the recommendations arising from this report are aimed at improving the user experience and technical feasibility of interoperability between technologies.

4 User Requirements and Human Factors

Face-to-face meetings offer rich verbal and non-verbal communication. A potential pitfall of videoconferencing is that this richness is lost. Therefore, it is imperative that videoconferencing should maximise the interactivity between remote sites. Successful videoconferences are those where the technology does not interfere with the normal progression of the meeting. To be widely adopted, videoconferencing must be a seamless technology to its users. For this to happen the User Requirements must be met and the Human Factors must be addressed.

Videoconferencing can provide a valuable supplement to face-to-face meetings, if it is managed correctly, and participants acknowledge the additional effort required in participating in a distributed meeting.

User Requirements are the technicalities that result in videoconferencing being a seamless, enjoyable experience. Attaining these requirements enables customers to conduct their meeting without hindrance. Therefore the technology must appear transparent to the users.

Human Factors are recognition of the ways in which humans react to the technology presented to them, adapting their manner and behaviour to accommodate distributed participants into a virtual environment. Good distributed meetings can be achieved by helping humans to get the best from the equipment and technology through training, preparation and some forethought to the wider issues such as culture, language and the meeting’s purpose and aims beforehand.

Videoconferencing should not replace the face-to-face meeting. In the best cases it should enhance and build upon relationships already established by allowing people to meet at relatively short notice where travel is prohibitive or time is critical or restricted.

4.1 User Requirements

Preparation time before start

10-15 minutes preparation time is vital to settle users who are unfamiliar with videoconferencing. This is to familiarise them with any peripheral devices they may wish to use and to make them aware of the special characteristics of videoconferencing, e.g. being aware of what they look like to remote participants and to ensure they are near a microphone. Emergency evacuation procedures should also be covered, as participants are often unchaperoned during the conference.

Room layout and design

An acoustically insulated, uncluttered, spacious, well-lit dedicated room is the ideal, with a modular table system and stacking chairs to allow choice in furniture layout for different conference types. A light slate-blue colour for walls and curtains is best. Clocks to show local and remote sites' times are useful. Videoconferencing equipment often includes the facility to view outgoing video feeds, but users prefer to check their appearance before going on camera, so a mirror is a good idea.

Users' clothing

Pastel and plain clothes work best. Users should be advised of this beforehand and asked to avoid white or red shirts, loud stripes and patterns, which causes bleed and additional camera problems.

Tips for users

A website of tips for users, including instructions on peripherals, is a boon for putting users at ease and helping them to prepare before arrival.

Users control pads/remote controls

These must be clearly labelled, simple to understand, with an emphasis on ease of operation.

Technical support

Technical support must be readily available throughout the conference, and able to provide immediate action to rectify any problems that may occur.

Set-up issues and prior testing

Unless there are frequent meetings between sites, it is imperative to test the connection to remote sites before the meeting to check compatibility, audio and video quality, echo control, to decide connection speed and to verify conference and contact details. The test need not involve users, though they should receive notification of the results of the tests.

Audio

Audio quality makes or breaks a videoconference (whereas poor video, to some extent, can be accommodated). Audio standards must be compatible across all sites. Microphones and speakers should be checked to eliminate echo.

Videoconferencing equipment should include echo cancellation. Audio must have clarity and not break-up in transmission. Delay is inevitable with longer distances and/or slower transmission speeds, especially over IP, but it should be recognised that large audio delays detract significantly from ease of use so should be minimised as far as possible.

Full duplex, as offered by the better proprietary hardware codecs and Access Grid, should be seen as a must, not a luxury. Without full duplex, too much attention is paid by users to the technology, rather than to the substance, of the meeting.

Video

Video quality in videoconferencing is not yet television-quality. Some delay is currently inevitable. Using higher bandwidths for IP and more lines for ISDN alleviates some delay. Ensuring all sites have compatible equipment and audio-visual standards minimises delay.

The conference timing should be planned to coincide with quieter times on the chosen IP or ISDN technology. This alleviates loss of quality due to packet loss that can occur on heavily trafficked IP, and loss of quality due to high levels of international traffic through switches on ISDN. Video delay manifests itself in terms of picture pixellation, lag in screen update, temporarily frozen-video and lack of lip sync. H.320 Quality of Service standards exist. Unfortunately the same is not true of H.323 videoconferencing and further research is necessary in this area.

Cameras

Video quality is also determined by camera quality. Small desktop cameras do not return the same quality of picture as dedicated videoconferencing cameras. The more advanced cameras offer white and light balance for better picture quality. Having multiple cameras is useful for providing several close-up views of various members of a large group. Multiple cameras are essential in distance learning to enable speaker or audience to be relayed to remote sites.

Eye contact is important. To do this, the cameras should ideally be positioned on top of the incoming display and if not, as close to that display as possible.

Monitors/screens/projectors

The distance between the display and the local audience should govern the size of participants on the display. The shorter the distance the smaller the size, with life-size generally being the optimum, and larger than life-size for longer distances.

Using a data projector makes maximising the screen size easier. Monitors need an anti-glare coating to cut out reflections from room lights or windows. If using a projector, a projector screen with reflective coating will give brighter, sharper image quality than a white wall. Projectors should be sited so that users do not cause shadows on the image.

Ceiling mounting is probably best and it should be set at the manufacturer's recommended distance from the projector screen for best image quality.

Multiple screens

Using multiple screens offers continuous views of the remote site(s) plus views of outgoing video from cameras and peripherals. If using multiple screens, it is important for viewers to be able to associate correctly each audio source with the correct video. Labelling locations helps viewers identify them. DuoVideo, as offered by Tandberg, enables two video sources to be sent simultaneously in a single call, provided that all partaking sites use Tandberg equipment and have two video outputs (monitors, projectors, etc).

Where Duo Video is not possible, technicians may be required to cope with the implications of multiple incoming/outgoing video feeds and other shared data, for example, to switch between feeds and/or to utilise display space.

Multiple screens and cameras may be deemed necessary for many types of session, e.g. distance learning, to enable all participants to maintain good eye contact between each other.

Multipoint conferences

If voice-controlled, the video of whichever site is speaking is transmitted to other sites as a full-screen image. If there is continuous presence, all sites are constantly visible and transmitted, as with Access Grid. If using continuous presence with small image size, location labelling is advisable to help identify speakers.

Lighting

Videoconferencing is often held in offices where the lighting (often pointing downwards and harsh to light table-top documents) shows participants in unflattering format (e.g. bright head tops, racoon shadows under eyes and noses, dark faces). Ideal videoconference lighting should include soft front lights to illuminate the speaker's face. Lighting should be dimmable to enable the speaker to be highlighted and to facilitate use of a data or overhead projector. The lighting reflectors should not reflect light either into the camera lens(es) or onto the surfaces of the display device(s).

Blinds or lightproof curtains should be fitted to windows to prevent light pollution of images to camera lenses, screens and projectors.

Calls dropping off

This fault is extremely off-putting for users. Its causes are much the same as for video delay problems: heavy traffic, network problems and incompatible standards between equipment. It can be minimised by performing tests beforehand to maximise equipment set-up and conference timing for reliability.

Computers/laptops

Users of laptops or computers need to find videoconferencing a "plug-and-go" experience. Prior testing is advisable, as is having an experienced technician on hand with a range of cables and knowledge of the various PC and Mac operating systems, video mirroring and dual video output set-up. Laptops can be connected directly into videoconferencing systems and the outgoing video source can then be switched between camera and computer, or displayed simultaneously, as in the cases of DuoVideo or Access Grid.

Alternatively, information from laptops can be transmitted via a networked computer using NetMeeting, VNC or similar applications. If it is a large meeting, the computers should be connected to data projectors and screens for the audience.

Document cameras

It is important to have easy operation to switch between outgoing video sources (if there are not enough to make switching unnecessary), including the document camera. Document camera should have simple-to-operate zooming and focussing.

Video recording and playing

Videocassette recorders (VCRs) at each site are useful for playback and transmission of video via the codec. They are invaluable for recording conferences (one video source plus all audio is usual) but permission from all sites to record must be obtained first.

If the conference is via Access Grid, other methods than VCRs are more appropriate for recording sessions. (For example, Voyager, which is a media server that enables scalable, multi-stream, multimedia record and playback facilities.)

Presentation material

Presentation material can cause serious problems when speakers are unaware of the limitations of the technology. This applies to presentations in electronic format (e.g. using a laptop), but even more so when presentations are on paper or foils (e.g. using a visualiser). Firm guidance to speakers is often necessary to militate against some ingrained habits. A checklist can be useful for this.

Presentations should be landscape format to match the typical display surface; fonts should be large and graphics bold. Pages should not be cluttered with detail. A recommended colour scheme is dark blue text on a light cream (rather than white) background. Red should be avoided as it bleeds. Speakers should not shuffle pages around on the visualiser whilst speaking.

In the case of Access Grid, presentations are often distributed first, so the above problems are avoided. However, there may be other limitations for methods employed by this system (for example, Distributed PowerPoint only allows slide transitions if triggered by a 'Page Down' event).

Presentations should always be emailed/ftp’d or faxed ahead of the conference so that each site can produce paper copies for users to refer to and make notes on during the conference. This is also a failsafe for the speaker should the technology break.

Interactive whiteboard

NetMeeting offers a PC-based solution, Mimio and others a proprietary product. These are useful for minuting if the culture of the meeting allows on-line minuting. Whiteboards are also useful for pooling ideas.

4.2 Human Factors

There are essentially two sides to every distributed meeting, the first is the technology (providing it works, it can be transparent to the meeting), the second is the users, their behaviour and integration into a distributed meeting environment. There are several key areas that need to be addressed with human factors. They are grouped under the following headings:

- Chair Skills
- Meeting Administration

- Physical Behaviour
- Environment and Culture

These areas are not specific to videoconferencing - most are just as relevant to face-to-face meetings - but due to time limitations videoconferencing exacerbates these issues by highlighting poorly planned and delivered meetings. In addition, these key areas are not mutually exclusive, they are interdependent and when combined they give a solid foundation for a successful distributed meeting.

Chair Skills

A strong, considerate chair is an essential part to any meeting, but with videoconferencing the meeting is more of a production. This requires planning and direction by the chair to allow all remote and local participants to have equal consideration.

Hierarchies may already be present and it will usually be the chair's responsibility to ensure that a rich network of communication develops through the whole group rather than being channelled through a minority of senior individuals or dominated by those who are prominent on camera (funnelling).

- The Welcome – sites joining the conference should be greeted in a simple, but effective way by acknowledging their presence, e.g. “Hello Edinburgh”. Similarly, if sites leave part way through the meeting, or even at the end, finish with “Goodbye Edinburgh”.

- Introductions – once everybody is settled in the meeting, the Chair should introduce everybody at his/her local site and then participants at the remote sites. Alternatively, this may be done by somebody local to that site who knows these participants.

- New/late arrivals – should be introduced by the local site as they arrive. This alerts people to who is present, even if they are off camera (although the situation where participants are off-camera should be avoided, if at all possible).

- Agenda and timescales – if possible, an agenda should be drawn up for every meeting and circulated in advance. This is not always possible for informal discussions but videoconferencing is usually time critical and overruns are not always possible. A time limit could be allocated to each agenda item. This gives participants an indication of the importance of the item and can focus attention on the decisions that need to be made before moving on to the next item. This should not be used as a tool to stifle debate or hurry through unpopular decisions, so the chair needs to be aware of the audience and the nature of the meeting in deciding if this is appropriate.

- Off-line discussions – discussions not related to the meeting or agenda should be directed to be carried out “off-line” by the chair.

- Decision making – measuring the value of any meeting is difficult. With cost savings being the main driver for distributed meetings it can sometimes be difficult to quantify hidden costs such as delays in the decision making process due to technical failure or key personnel being unavailable at a particular venue. In addition, the chair should clarify decisions that are being made, to avoid false assumptions by participants who may not know each other very well or have never worked together before. Participants should be encouraged to seek clarification of issues or decisions that are not clear.

Meeting Administration

After successfully booking a venue and agreeing participants, there are some additional administrative tasks that can be undertaken to help a distributed meeting flow more smoothly. The most essential part of the meeting is the initial set-up time beforehand. Sound checks, video check, peripheral devices and seating should all be ready to go, participants should arrive 10-15 minutes prior to the start time to get comfortable, in camera shot and be prepared (as discussed in User Requirements).

- Attendance list/contact details – names and contact details of participants should be circulated in advance. This stops people interrupting the flow of the meeting to ask for these details. It can also help to gauge the audience and importance of the meeting.

- Agenda/presentation – agreed agendas and non-sensitive presentations should also be circulated in advance. This helps people prepare more effectively and provides a backup if PC's or laptops have technical problems.

- On-line minuting versus secretary – on most occasions it is better to have a designated person or secretary taking minutes "off-line". On-line minuting can quickly turn into on-line editing and proof reading.

- Timing – this is applicable to all types of meeting. 9am on a Monday morning or late Friday afternoon should be avoided. Organisers also need to be aware of timezones. It should be accepted that if meetings are during the lunch hour, participants will bring food or expect it to be provided. This can cause problems with the microphone, so the mute button should be clearly identified.

- Comfort Breaks – when meetings are scheduled to take over an hour, breaks should be built in to this time, breaking for 5-10 minutes will help to re-focus the meeting and refresh the participants.

- Microphone off – whatever type of distributed meeting is being held, it is essential that participants understand that the microphones are sensitive, not just to discrete sideline conversations, but also to finger tapping, paper shuffling and eating lunch. Remote sites should be encouraged to switch off their microphone if they want to chat or eat lunch, but still listen to the proceedings.

Physical Behaviour

People can behave very differently in a distributed meeting to one that is held face-to-face; especially if they don't know everybody or have never used the technology before.

- Breathing/speaking – participants should be encouraged to speak normally, but to allow for very slight delays in the audio. By letting other participants finish this can stop the cycle of voices interrupting and backing off. The chair should intervene if comments turn into speeches.

- Shouting/mumbling/eye contact - participants looking down at their notes, shuffling papers or not making eye contact with the camera can be frustrating. Directing participants to speak normally, but clearly, at the outset and to sit "newsreader" style to the camera should avoid these problems.

- Physical space – the videoconference room should be set-up to facilitate the most effective distributed meetings. This means putting a conference table at the front of the room around which the main participants can sit. Theatre style seating could be placed behind this for observers. The camera should have a wide-angle lens to take in as much of the audience as possible, although speakers can be zoomed into camera.

- Non-native speakers – for the UK, if there are participants whose first language is not English, the chair should be made aware of this so that decisions can be clarified if necessary. Similarly, if there are participants with very heavy accents the chair may need to repeat things, in a manner that will not cause offence to the speaker.

Organisational Culture and Dispersed Teams

Depending upon the environment in which videoconferencing will be deployed, e.g. academic, clinical or commercial, requirements will vary considerably according to timescales, financial commitment and the culture of the organisation. For example, clinical meetings may be concerned with simply facilitating clinical discussions with remote colleagues or students, rather than fostering a collaborative virtual environment.

A commercial organisation may employ a large distributed team of people who need to teamwork over large distances; they may regard the effort of building the distributed team essential to the organisation's success. The e-Science programme may be considered to have similar requirements. Organisational culture is not something that is simply highlighted by videoconferencing; the list below shows some important considerations when trying to build a distributed team. However, these may also apply to most localised teams.

- Ensure there is one clearly defined leader for the whole team that takes a proactive approach within the team

- Establish a clear sense of vision for the team and communicate this to them

- Balance distributed meetings with face-to-face meetings, making an effort to visit remote sites periodically

- Encourage open communication of problems and clarification of decisions

- Work on team integration, de-emphasising cultural differences, personalisation of meetings, not showing partiality to any one site

- Develop a strategy for dealing with time zone differences and work at standardising procedures

5 Access Grid Studio-based Videoconferencing

5.1 About Access Grid

The Access Grid was first conceived and initially developed by the Futures Laboratory at Argonne National Laboratory. The aim of Access Grid is to provide an effective environment for remote group-to-group collaboration. Whilst this includes various audio and video conferencing components, they are used within a peer-to-peer model that is coordinated via a separate peer-to-peer services layer presented as virtual spaces.

Access Grid is a research project that attempts to provide a sense of presence that approaches that experienced in face-to-face meetings. The project is largely developed and supported by a worldwide research community that consists primarily of academic institutions. However, there are also major corporations that have committed R&D funds to the project and some that offer Access Grid hardware and services as commercial products (most notably *inSORS*).

In order to build a wide scale testbed for building collaboration tools, the Access Grid project specifically defines the minimum set of requirements necessary to be considered an Access Grid Node. These requirements are specified in a manner that avoids specific hardware and software implementations, but rather encourages the use of standards based digital media tools. This provides a level playing field for research into alternative node devices and the integration of other hardware and software components.

There are 12 Access Grid nodes deployed at the UK e-Science Centres. The nodes experience a typical usage of 4 or 5 meetings a week at each site. Demand on current nodes is increasing. Examples of use are: Engineering Task Force management meetings, e-Science project meetings (*Geodise*, *myGrid*, among many others) and other ad hoc meetings between sites.

Personal Interface to the Access Grid

In order to enable scientists that are geographically isolated, the Futures Laboratory has released a version of the Access Grid Toolkit that can provide a subset of the minimum node requirements from a single workstation. This release is known as the 'Personal Interface to the Access Grid' (affectionately known as the 'PIG' to the Access Grid community). This toolkit release has a very specific hardware platform that it can utilise to provide two video streams, the single audio stream, and a desktop scale high-resolution display. The details can be found via the Access Grid website.

This chapter of the report is assuming the deployment of Access Grid Nodes, rather than Personal Interfaces to the Access Grid (PIG). Consequently, deploying Personal Interfaces to the Access Grid would significantly change the conclusions of this chapter. However, this scaled-down toolkit may be thought of as a fully integrated solution that has a greatly reduced cost.

5.2 Costs

It is possible to buy an Access Grid node piecemeal and install software, etc. 'in-house'. This may be appropriate if Access Grid development is to take place. This is the route taken by the e-Science centres that administer the current UK sites. However, following this route may take a large amount of time and money in procurement (sourcing the elements of a node) and installation (integrating the hardware, installing and configuring software).

Another route is to use an established commercial organisation that is fully involved in and accepted by the Access Grid community. The product supplied is fully integrated with the Access Grid system and the virtual spaces model and is based upon software in use by the community. The costs below are for purchasing the system as a 'product' from this organisation (*inSORS*) and include a one-year's support contract as well as software designed to be highly usable that obviates the need for an operator. As such, it is a system that is comparable with the systems examined in other areas of this report. (The costs for the equipment for a 'home-made' node are similar, but will not include support or the resources required for procuring and installing such a system, which may increase actual costs.)

There are no running costs, beyond the cost of maintaining the existing network infrastructure.

Hardware Costs

InSORS supply a range of Access Grid solutions – 'Platinum', 'Gold', 'Silver' and 'Grid Station'. The Platinum solution goes way beyond the minimum specification required to use the term 'Access Grid' (for example, it includes wireless and stalk microphones in addition to the standard table-top microphones, flat panel monitors, etc.). The Gold solution is only slightly above the minimum and the Silver solution could not term itself a full Access Grid node (although users would still be able to participate in Access Grid meetings). The Grid Station solution is designed to meet the specification for the 'Personal Interface to the Grid' (PIG).

These are the prices quoted for the various solutions in August 2002:

- Platinum £30,000
- Gold £28,000
- Silver £22,000
- Grid Station (PIG) £9,000

5.3 Ease of Use

Most Access Grid sessions do not require the services of a dedicated node operator. Users of the system can perform most operations themselves with very little training. (It should be noted that these statements refer to the commercial software interface developed by *inSORS*, which is not yet deployed throughout the UK Access Grid community.)

Like any complex technology, it is advisable to have someone 'on-call' during meetings to assist with technical problems. Being available via a text messaging system (e.g. the MUD, in common Access Grid usage), is probably sufficient for most common situations; however there are times when production quality solutions are required. *InSORS* provides support for these production quality meetings providing 24 hour, 7 day a week support in seven languages. These support services can be negotiated on an event or time (yearly or monthly) basis.

5.4 Display Quality

An Access Grid node provides a large-scale high-resolution display either by projecting onto a wall or using a rear projection system. The minimum specifications require 3072x768 pixels, at a distance between 2 and 8 times the height of the projected image. This allows 18 QCIF and 6 CIF video streams to be displayed; additionally an entire XGA screen is available for collaborative applications to be used simultaneously.

The windows that render video streams on the display can be three sizes. The largest size window is used for 'presenter'-type shots and allows facial expressions and gestures to be seen easily. The middle size windows are used when there are many video streams that need to be displayed. Remote participants viewed in middle size windows can still be seen fairly clearly. The small size is typically used so that local participants can view outgoing video. This size is too small to be useful for viewing remote participants, except when it is necessary to display a very large number of participating sites.

5.5 Visual Quality

Each Access Grid node transmits four video streams. These can be used for a variety of purposes, including multiple camera angles of few participants, whole room shots, close-up views or to transmit video for other media such as VCRs or document viewers. One frequent usage is to use one outgoing feed for a presenter, two for the local audience and one feed to show remote participants the local display.

The minimum requirements document specifies these streams are transmitted 25 frames per second encoded using H.261 via RTP at CIF resolution.

Video is encoded for efficient data transfer, which means that fast movement can result in degradation and parts of the video feed not being updated for several seconds, depending upon network performance.

5.6 Audio Quality

The audio quality in an Access Grid node is of a very high quality. The audio stream is sent uncompressed and sampled at 16 bits at 16 KHz. A single audio stream provides mono audio, but the system is capable of sending multiple audio streams if necessary.

The audio component utilises a high-end echo canceller (Gentner XAP400 or XAP800), level balancer and good quality microphones. Participants have hands-free, full-duplex audio (i.e. many people can speak simultaneously). The resulting sound is as good as that experienced when the participants are co-located.

5.7 Networking Issues

An Access Grid node requires connectivity to an IP-based, multicast-enabled network. It is possible to interact with the Access Grid via a multicast-unicast bridge (or reflector, or tunnel) even where the Local Area Network is not multicast-enabled, but this is usually used as a temporary, stopgap solution. However, the quality of experience using a bridge is indistinguishable from the experience when using full multicast.

The Local Network connection requires 100Mbps connectivity. The Wide Area Network requires 10Mbps.

5.8 Multi-Site Issues

Features of Access Grid such as a large display, multicast networking and the use of a high-end echo canceller make it especially well suited to meetings involving multiple sites. The large display means that many video feeds may be shown without losing the benefits of reasonably sized windows. Multicast networking is an inherently scalable solution to the distribution of audio-visual and other data used in meetings. The high-end echo canceller enables audio quality to be maintained even though the number of audio inputs is high.

As such, the maximum number of sites that may participate in an Access Grid meeting may be limited more by logistical considerations than limitations of the technology.

The UK e-Science Engineering Task Force has held all its meetings in 2002 using the Access Grid. The most recent meeting (30^{th} July) involved 27 participants located at 12 Access Grid nodes, utilising 19 useful video feeds and was just as effective as if participants had met face-to-face.

5.9 Collaborative Tools

Most peripheral devices can be seamlessly integrated within the Access Grid system. For example, a document viewer, VCR or similar device can take the place of a camera for any of the video feeds. Switching between camera and device is a trivial operation with the addition to the standard Access Grid set up of a relatively cheap (~£100?) video switch. (Alternatively, further video capture cards may be purchased to provide extra video feeds.) Similarly, audio devices can be input to the high-end echo canceller. Because of the large number of inputs available on the standard echo canceller, there is no need for an audio switch.

Because the system is based upon PCs, there are usually ways to integrate other devices via the PC's various input ports.

The integration into the Access Grid of software tools and applications for data and information sharing is the area where this system has its greatest strengths. The virtual venues model upon which the system is based has at its heart the concept of persistent objects being present in virtual spaces. For example, a data repository and visualisation tools with which to analyse the data might reside a virtual venue. Scientists could then agree to meet at that venue to perform their collaboration or spend their days working together in a virtual laboratory provided by the venue. The aim always is to make remote meetings of equal value to those that are collocated.

Software tools that have already been integrated into Access Grid meetings are VNC (this enables sharing of PC desktops and the resulting applications), two implementations of a distributed PowerPoint tool (this enables presentations and seminar-type meetings), The Visualization Toolkit (vtk), and others.

5.10 Security

The Access Grid Toolkit v1.0 has implemented a trivial proof of concept for security where an Access Control List is used to allow access to a virtual venue. If a particular user is allowed into the venue, the venue provides keys with which to encrypt the audio and video streams shared among the participants. Another approach is to share audio using the public telephone system. However, this has its own set of security issues that are not addressed by this chapter.

Securing shared data is tackled on an ad hoc basis. Most collaborative tools have their own security features that are utilised. For example, VNC employs a password entry system. Another example is when Distributed PowerPoint is used to share a presentation. In this case, the presentation itself is separately held at each node. Only control events, such as 'Page Down', are broadcast and not the data within the presentation itself. These control events, whilst not encrypted, are of little use in themselves.

Security is a primary concern of the Access Grid project and careful design, incorporating wide participation from the community is progressing. There are many possible security threats and attacks, and the Access Grid community is relying on the Global Grid Forum to provide a vehicle (through the Advanced Collaborative Environments and Grid Computing Environments Working Groups) to standardise security interfaces, toolkits and policies.

When v2.0 of the Access Grid Toolkit is released it will leverage the security framework that is provided via the Globus Toolkit v2.0. The Globus Toolkit will provide us with an open implementation of the IETF GSI standard, as well as single sign-on access to the Access Grid services and other Grid resources.

5.11 Appropriate Usage

The Access Grid is well suited for 3-6 participants at each of 2-12 sites. It can be used in very formal meetings, web casts, classroom style interactions, or for unstructured interactions where the desire is simply to provide a continual sense of presence with remote collaborators.

The Access Grid does not provide any formal floor control mechanisms, since that would be in conflict with the premise that if participants feel more presence then standard social norms can and will govern interactions. Similarly, formal voting mechanisms are not supported.

Moving forward, the Access Grid is placed to become a ubiquitous resource for collaborators in various disciplines. It should provide the collaboration fabric that is rich enough to increase productivity and create new ways to do science.

5.12 Future Potential

The Access Grid is positioned to become a platform for exploring the future of collaborative research; as such it is uniquely capable of providing a system that can support both legacy applications and the development of next generation, grid-enabled applications. It also provides a vital end-user experience for shared development, debugging, execution and performance monitoring of grid-based applications.

During the Global Grid Forum's 5th meeting held in Edinburgh, the Access Grid Project unveiled a plan for the second version of its toolkit. This version will integrate grid computing tools provided by the Globus Toolkit v2.0, including security, data management, and GridFTP. Further, the Access Grid Team at Argonne National Laboratory will be engaging at least two applications groups to develop the appropriate Access Grid interfaces so that applications will be able seamlessly to leverage the collaboration, data management and grid computing resources being made available through grid computing achievements.

Additionally, the Access Grid 2.0 Toolkit will expose a services model that provides extensibility. It will allow developers to extend the functionality of the Access Grid by providing not only specifications and tools for building services, but also a framework within which the services can be found and utilised. The initial goal of these services is to provide data stream processing to enable even higher quality audio and video experiences that include NTSC/PAL, HDTV and higher resolution video, multi-channel, virtualised audio, streaming event notifications from devices and services on the network, and potentially automated network failure recovery mechanisms in response to real-time monitoring.

6 H.323 / H.320 Studio-based Videoconferencing

6.1 About H.323 / H.320 Studio-based Videoconferencing

H.323 (IP) and H.320 (ISDN) based systems represent the mainstream of commercial products and development in the videoconferencing arena. H.320 based systems have been in use by the education community for over ten years. H.323 and H.320 are international 'standards' for videoconferencing as defined in recommendations published by the International Telecommunications Union.

H.323 IP-based videoconferencing systems have come more to the fore in the last 2-3 years. This has been as a result of a number of factors; principal among these is cost saving. The use of IP as a network transport technology removes the recurrent call costs and line rental associated with the use of ISDN. However, H.323 systems have additional advantages over the more conventional H.320 ISDN based systems both in terms of flexibility and scalability. H.323 systems can more easily scale to higher bandwidths and resolutions than H.320 systems. It should be noted that in practice most H.320 systems have H.323 capability and vice versa. These are sometimes referred to as hybrid codecs.

Within the education sector in the UK, the majority of videoconferencing deployments have been in studio-based applications. However, a full range of commercial products is available from a number of manufacturers including Sony, Panasonic, Polycom (PictureTel), Tandberg and VTEL. These range from small portable devices that plug into a laptop, to large portable systems (that can be carried), to rollabouts (systems that are mounted on a trolley and can be easily moved from room to room) and full room/studio based systems (typically fixed and dedicated to the room in which they are installed).

Both the commercial and education sectors have invested heavily in H.323/H.320 technology and continue to do so. Within the education sector during 2000/1 eighty-hybrid H.323/H.320 studios were deployed in Wales, this project is referred to as the Welsh Video Network (WVN) [http://www.wvn.ac.uk]. The WVN has provided at least one hybrid H.323/H.320 videoconferencing studio to every HE and FE institution in Wales. The recommendations to adopt a hybrid H.323/H.320 approach to meet the requirements of videoconferencing in Wales are described in a report [http://www.wfc.ac.uk/pubs/pdfs/welshvid.pdf]. The WVN studios were designed with the help of teachers and lectures specifically to support distance education, with the same studio equipment and user interface being deployed in every studio. H.323/H.320 systems have been deployed in the Scottish Higher Education sector that has also invested in 36 Hybrid H.323/H.320 systems to update their videoconferencing studios at all HE institutions across Scotland.

A JANET H.320 ISDN based videoconferencing service has been operational for over five years now, providing multipoint capability, booking and scheduling services and Quality Assurance testing. There are currently over 250 studios within the education community in the UK that are registered for this service. A JANET H.323 multipoint videoconferencing service is being piloted with a view to a full service being provided by the end of the year. Preparation for the provision of H.323 service and recommendations can be found in the JANET H.323 Architecture Group Report [http://www.ja.net/development/video/arch/].

As part of the central services provided by the JANET videoconferencing management centre based at Edinburgh, an online videoconferencing booking service is operated. The management centre also operates regular 6 monthly Quality Assurance testing to ensure that audio and all registered users of the system maintain video quality levels. The management centre also provides help and advice with fault-finding and diagnosis. This service is provided free at the point of use to the JANET community.

The H.323 standard describes the behaviour of "terminals and other entities that provide multimedia communications over Packet Based Networks which may not provide a guaranteed Quality of Service". Support for audio is mandatory, while data and video are optional. H.323 entities may work with H.320 equipment and the standard can be regarded as parallel to the H.320 standard for packet-based environments. The two services have much in common, including the protocols for call signalling and call set-up. Both encompass the same encoding algorithms for media streams. The main difference between the protocols is at the transport layer.

H.323 videoconferencing systems are particularly attractive to academic institutions because those institutions are already connected to the Internet via the JANET academic backbone, and so can support IP-based videoconferencing using existing infrastructure.

Challenges remain for all types of IP videoconferencing in the areas of security and quality of service. The issues faced in providing security and authentication is outlined in 5.10 Security. Quality of Service is the term given to providing some measure of guarantee that information transmitted from one end of a network connection will reach the other. Different measures of guarantee can be applied, however this is very much a development area at present.

It should be noted that in looking to deploy any videoconferencing system the largest single component that contributes to the quality of system is video resolution, which is hence the largest contributor to cost. The vast majority of videoconferencing systems in use today run at CIF resolution (352x288). Therefore the largest single contributor to cost is the bandwidth that the unit operates at as this will determine the quality of the call and the hardware processing and components required to operate the call at the bandwidth specified.

The remainder of this chapter examines commercial hardware-assisted H.323/H.320 videoconferencing systems. It does not examine desktop software-based H.323 systems, which are considered as part of the non-studio-based solutions chapter.

6.2 Costs

Set Up Costs

Different videoconferencing systems on the market vary in cost and functionality and are therefore suitable for different applications. The systems range from mobile and portable devices to rollabout and studio based systems. All systems are hardware-assisted in the encoding/decoding process for audio and video. Most come complete with a microphone package, remote control and camera.

Small studio systems typically have support for H.320 ISDN-8 (512kbit/s) and H.323 (IP) up to 2Mbit/s. These units offer a range of connectivity for peripherals such as document cameras, projectors, VCRs, interactive whiteboards, etc., and typically operate at a higher bandwidth and therefore higher quality than desktop or portable systems. These systems can also have small-inbuilt MCUs. However, it should be noted that these devices are very limited in terms of call speed and therefore the quality of each site in an MCU call using these units. For example if a unit has ISDN-6 connectivity then three other units could connect at ISDN-2 (128kbit/s). This functionality can be very useful in specific applications, but is typically rarely used due to the limitations of quality and functionality of the MCU feature set. These systems typically retail for £10,000 and are suitable for small group tutorial work. The limitations of videoconferencing group size are associated with the display medium chosen, room acoustics and the bandwidth at which the conference is operated.

Systems suitable for large studio applications typically have multiple independent camera inputs with remote, pan/tilt/zoom, support for ISDN-30 (2Mbit/s) and H.323 (IP) up to 4Mbit/s. These units are typically the size of a desktop PC and offer a larger range of connectivity for peripherals and the flexibility for remote control via LCD screens that can be custom designed for ease of use in the given application. These codecs' typical costs are around £13,000 for the base system with a camera and remote control.

Typical H.320/H.323 studios have the following peripheral equipment. Two or three large 29" TV for viewing images; projectors are used in some cases but are a compromise between larger viewable image, noise generated and the lower light levels required to see projected images. (Cameras in contrast require high light levels to operate optimally.) Document cameras, interactive electronic white boards, VCRs, graphics tablets, PCs to support data and application sharing, and touch screen control pads can all be added to enhance the communications environment.

It should be noted that if any site is deploying H.323 videoconferencing facilities they should consider running a gatekeeper. These are call set up and address translation devices that provide call routing information. A central JANET gatekeeper is provided for institutions that do not have a gatekeeper. However, this limits a sites ability to secure its network and implement any controls on videoconferencing traffic to or from a site. It is recommended that all institutions running H.323 service on JANET run a gatekeeper and take note of the recommendations from the JANET H.323 Architecture Group [http://www.ja.net/development/video/arch/].

Running Costs

The cost of managing and operating equipment are hard to quantify directly but someone has to manage the videoconferencing facility.

Maintenance costs for service equipment are always present. Equipment maintenance costs are typically around 10% of purchase price per annum, discounts can be obtained by purchasing multiple years in advance. Typically, three years' maintenance paid in advance would equal the price of two years' paid for annually.

There will always be a cost associated with operating and maintaining a network infrastructure within institutions. However, this work is usually undertaken by IT departments as part of routine activity unless specific work is necessary to connect the videoconferencing studio to the network, e.g. long cable runs etc.

H.320 videoconferencing systems incur line rental and call charges for ISDN calls. H.323 videoconferencing systems do not incur such costs even on international calls. However, because many codecs incorporate ISDN and IP connectivity, and because of security issues with IP conferencing (see 5.12 Security) it is often the case that an institution may also rent ISDN line to provide their users with multiple options for conferences.

6.3 Ease of Use

Simple H.323/H.320 videoconferencing systems do not require the services of a dedicated operator. The systems come with a simple-to-use remote control or in studio applications can be fitted with touch screen LCD panels. Once users have been given a small amount of training, they should be able to instigate and manage conferences for themselves. LCD control systems can be programmed to control other peripherals including additional cameras, VCRs, whiteboards and even room lighting giving studios an integrated feel.

As with other types of videoconferencing, it is always advisable to have someone 'on-call' to assist with technical problems. A hotline telephone in a corner of the room is sufficient, provided the person 'on-call' can come to the studio immediately if summoned.

Where there is a combination of customers new to videoconferencing and/or wishing to use many peripherals and/or it is a high level conference, then a degree of technical assistance is often requested. In these circumstances it is often desirable for the technical assistance to be present for the whole conference.

6.4 Display Quality

Commercial hardware-assisted H.323 solutions can output to television monitors, plasma screens or data projectors, according to the scale of the facility and budget available. The aspect ratio is 4:3 which does not preclude the use of modern 'wide-vision' televisions. However only a 4:3 picture can be displayed.

In many ways, a data projector (with greater than 1200 ANSI Lumens using LCD or DLP technology) is the ideal display medium. The size of the display can be enlarged or reduced, according to the number of people in the room. However, data projectors work best in reduced levels of light. This conflicts directly with cameras, which require high light levels to function optimally. This problem can be reduced by using a high output data projector and using a light reflective projection screen rather than white painted wall, however care needs to be taken not to introduce extraneous noise in to the videoconferencing environment. Plasma screens would appear to be an ideal solution however practical experience would suggest that quality and reliability are not yet on a par with projectors.

6.5 Visual Quality

Video resolution is the single largest contributing factor to the cost of videoconferencing systems. Most videoconferencing systems and all H.323/H.320 systems operate at CIF resolution (352x288). This is the limiting factor in terms of video quality. The next area is frame rate. To be useful in most applications the frame rate needs to be 15fps (frames per second) or higher with 30fps or greater being optimal. In addition commercial hardware-assisted codecs today now use complex algorithms to improve the video image that is displayed. Commercial companies who produce H.323 hardware-assisted codecs are developing new systems that will operate at 4CIF (704x576). These systems will take a large step forward in terms of image quality and will require additional bandwidth to operate optimally.

In parallel with H.320 videoconferencing systems, H.323 systems supports H.261 and H.263, both ITU-T 'standards' for video encoding.

6.6 Audio Quality

There are three components to good quality audio that apply to any form of videoconferencing. These are room acoustics, echo cancellation and the audio encoding format. A problem in any one area will make the conference uncomfortable to participate in and very tiring due to the concentration involved. Significant advances have been made over the last three years in commercial hardware-assisted H.320/H.323 systems in this area. The resultant improvements have been significant and noticeable to users as conferencing is much easier and feels more natural.

Like other videoconferencing systems, echo cancellation is required for H.320 and H.323 because of the delay created in the encoding process. The image and audio data is compressed many thousands of times so that it fits into a relatively small amount of bandwidth. This causes a delay of approximately 100ms in hardware-based systems. In software-only systems, the video encodings are not undertaken completely and several shortcuts are used to reduce the encoding time to acceptable levels but this is to the detriment of image quality. However, the encoding cycle can still take several hundred milliseconds adding to the delay noticed by conference participants. The latest commercial H.323/H.320 systems are sophisticated in their treatment of echo cancellation, which had been a problem for many H.320 systems for a number of years. Commercial systems now use a combination of adaptive echo cancellation, automatic gain control and specialized microphones to all but eliminate this problem. Good quality conferences now use full duplex audio allowing participants to 'talk over' each other as in a normal meeting. This makes for a more natural meeting/teaching environment.

H.320 and H.323 videoconferencing systems typically support the three standard audio encoding formats of G.711 (3.5kHz audio bandwidth), G.728 (3.5kHz) and G.722 (7kHz) - all three ITU-T 'standards'.

6.7 Networking Issues

Commercial hardware-assisted H.323 codecs require well-engineered unicast IP networks. There is no requirement for IP multicast. H.320 codecs require the installation of ISDN connections.

It should be noted that H.323 IP based videoconferencing is the first of a series of new real-time applications that puts significant demands on underlying network infrastructure. It also has implications for how networks in general are operated and supported. If there is a network fault in the middle of a conference, taking 15 minutes to restore service could be disastrous to the success of the conference.

It is recommended that connections to any IP videoconferencing codecs be made with 100Mbit/s switched Ethernet connections. The onward connection to an institution's main access router needs to be well-engineered. The videoconferencing traffic should not pass through any heavily loaded routers, hubs or switches. Where possible, direct connection to an institution's main access router on dedicated ports can solve significant problems of sensitive videoconferencing traffic having to transit legacy network infrastructure. The bandwidth of the onward connection into JANET and its loading is also crucial to the success of IP-based videoconferencing.

There were two issues related to H.323 IP videoconferencing that needed to be resolved before any service could be launched on JANET. These were the issue of gatekeeper hierarchy and an addressing scheme that could be used globally. The JANET H.323 architecture group was formed to provide recommendations on these two and other complex areas in preparation for a JANET service. The full recommendations are available from [http://www.ja.net/development/video/arch/]. In summary a global dialling scheme was devised in collaboration with other National Research and Education Networks in Europe and the US. SurfNET, HEAnet, JANET and Internet2 have now adopted the Global Dialling Scheme. In respect of gatekeepers that are call set up, admission control and address routing devices it is recommended that each institution deploys their own device to allow control of this traffic type locally.

In order to support H.323 videoconferences over JANET from sites that have not yet deployed a gatekeeper, a central JANET gatekeeper has been provided which is available for everyone on JANET to use.

Currently within H.323 videoconferences, the audio and video quality experience varies throughout the call. This is due to the variation of network loading being experienced during the day. This issue appears to be largely associated with campus networks where alternative network engineering is possible at relatively low cost. However, in some cases the problem resides in the backbones of regional area networks. In this case a Quality of Service mechanism could be used to prioritise real-time and therefore improve the quality of IP Videoconferences. A JANET Quality of Service trial is due to start shortly.

The audio and visual quality of H.320 ISDN based videoconferences tends to be reliable and is of a fixed quality once the call is established. This is due to the use of dedicated ISDN lines that provide guaranteed bandwidth between the two systems, at a price. However, in use there can be difficulties in connecting to remote systems overseas given differences in telecommunications infrastructure globally and again traffic loading on the international telecommunication networks. Once calls are established typically few problems are encountered.

6.8 Multi-Site Issues

There is a need for a full multipoint bridging service that may be met by the codec itself for smaller multipoint conferences (i.e. involving less than 4 sites). For larger conferences, or conferences that need to operate at higher quality, a full service is available for H.320 ISDN videoconferencing on JANET and a pilot is operational to support H.323 IP videoconferencing. The H.323 IP Videoconferencing Pilot on JANET currently supports each site's connection at up to 768kbit/s but will support 2Mbit/s over the next few months. It is also able to transcode between different bit rates and protocols and combines calls over ISDN with calls over IP. There are also commercial bridge providers though these incur a cost, unlike the JANET bridge, which is free at the point of use to the JANET community.

The current pilot service can support a maximum of two 24 site H.323 IP videoconferences with all sites operating at 2Mbit/s. If sites operate at 768kbit/s then four 24 sites conferences can be supported concurrently. Typically the largest conferences run on the ISDN service feature about 17 sites connected to a single conference. However, the average size of conference is between three and four sites. By default these conferences operate in a voice-switched mode, where whoever is talking loudest, everyone sees. It is also possible to operate conferences in a so-called ‘continuous presence’ mode of operation, where the screen is split into four or nine squares and everyone in the conference can see everyone else. However, the effective size of each video feed in this mode of operation will be small if using traditional monitors rather than projectors.

6.9 Collaborative Tools

As part of the JANET Videoconferencing Service a server is provided to support data and application sharing between all the sites involved in a conference. An ILS server is provided for all registered sites to use. This allows all sites to collaborate and share PowerPoint presentations, Word documents or any other material live and interactively during a conference. All the end site needs to do is to locate a PC in the videoconferencing studio, register on the ILS server, by contacting the management centre, and share data and/ or applications with other members within the videoconference. The application used to facilitate this is Microsoft’s free NetMeeting product. This is a quasi-compliant H.323 software-only codec. However, in this application only the data and application sharing side is used.

6.10 Security

Securing IP networks and end systems from hacking attempts is a continuous battle. Here best practice and sensible security measures apply as for any other IP connected device. H.323 is an inherently difficult protocol to deal with in networking terms, as it requires the use of random high ports within the IP stack. Sensible measures can be taken in defence of attempted snooping or hacking of H.323 systems or sessions.

Perhaps the best defence is to deploy an H.323 proxy server. In this way you are only concerned with securing one system rather than all your videoconferencing equipment. Access list controls can be used effectively but a careful approach and an understanding of H.323 are required here. Another approach is to physically separate H.323 traffic onto a separate subnet from the rest of the network traffic and isolate it. In this way a dedicated router port is available and performance improvements should be visible during conferences.

Encryption technology for IP videoconferencing applications is being developed, and should be available in the next 6 to 12 months. Encryption technology is available for ISDN H.320 systems and is used where the application demand and the additional cost can be justified. It should be noted that when communicating over JANET between connected organisations who have implemented general best practice in terms of a network security policy, the likelihood of any snooping of conferences is slight.

Currently users are concerned and many will only use H.323 videoconferencing for non-confidential videoconferences, preferring H.320 for confidential matters, whether discussion of patients' medical records or patents for products. As most IP-based proprietary codecs are also ISDN-based, this does not mean purchase of a second codec. It does however mean that the user may wish to have ISDN lines installed for such occasions, with the associated cost implications.

The International Telecommunication Union publishes a standard for the security of H.323 devices titled H.235. This describes a control protocol and framework to implement encryption and authorisation between H.323 entities. It is likely that commercial products supporting the standard will be available in the next six to twelve months.

The JANET H.323 Architecture Group Report recommends that in the absence of appropriate encryption methods, the management of H.323 equipment should be performed over a switched network path to minimise the risk of passwords being compromised. The report also recommends that a H.323 proxy or an effective H.323 firewall should protect every end system using the service. To improve security and performance, institutions should also seek to deploy switched network devices for connecting their H.323 endpoint(s) to their site access router(s). The potential for undesired snoopers to silently join a conference should be minimised by having the MCU initiate calls to participants and the service restrict connections to registered IP addresses.

6.11 Appropriate Usage

The H.320/H.323 videoconferencing technology in a studio-based environment is well-suited to all types of videoconference, whether small or large. The ease with which peripheral equipment can be added makes it suitable for distance learning and other presentation-type videoconferences with multiple video sources, as well as more traditional meetings.

The equipment is commercially produced and supported to agreed international standards. This is a well-established market place with a significant number of large commercial manufacturers and many suppliers and integrators of solutions for videoconferencing studios who have many years of experience deploying and integrating equipment of this type.

H.320/H.323 hardware-assisted codecs are the mainstay of commercial videoconferencing systems worldwide and represent the vast majority of usage within the education sector. Other systems do exist and are suitable for specialist applications, however H.320/H.323 hardware-assisted codecs represent the main stay of videoconferencing within the education, research and commercial sectors in the UK and worldwide.

6.12 Future Potential

The main issue for H.320 is the ISDN cost. Therefore the main threat is H.323 given that it is free to use, given an appropriate broadband Internet connection.

The main issues with commercial hardware-assisted H.320/H.323 videoconferencing systems are consistent audio and video quality in H.323 calls and the security of H.323 conferences.

The issue of consistent call quality can be addressed by the use of Quality of Service. The JANET H.323 Architecture Group Report recommends that H.323 videoconferencing services should make use of an appropriate Quality of Service that results from the JANET initiative on QoS development. Also that an appropriate QoS monitoring facility should be developed to monitor those QoS performance metrics that are particularly relevant to an IP videoconferencing service.

In respect of network and device security there is already a significant body of good practice. In regard to encryption and authentication there are now international standards in this area, which will see additional functionality added to commercial products in the next six to twelve months.

Codec manufacturers are producing IP-based codecs and academic institutions, always the leaders in new technology, are pioneering take-up of IP-conferencing for cost-saving reasons. The Welsh Video Network (http://www.wfc.ac.uk/pubs/welshvid.pdf) describes how a network of IP and ISDN-based facilities has been established across Welsh higher education institutions for videoconferencing collaboration.

In East Anglia, Anglia Polytechnic University, with its various campuses geographically distributed, has been a long-time exponent of both ISDN and now IP-based videoconferencing. For *it* and others, videoconferencing is a need, both for teaching and management. With the advent of EASTNET they have been able to switch largely from ISDN to IP. It is a similar story in Scotland with the Scottish MANs recently upgrading 36 studios with commercial H.323 hardware-assisted codecs.

7 VRVS Studio-based Videoconferencing

7.1 About VRVS

A principal design aim of VRVS is to make multipoint videoconferencing over IP available to a wide range of users, ranging from well-appointed studios, auditoriums and seminar rooms to inexpensive desktops and laptops. In order to achieve this, it supports various protocols: currently H.323 and the so-called "Mbone" tools (*vic*, *rat*, etc.). However, it does not rely upon the availability of IP multicast but can support it when and if appropriate.

VRVS itself can be understood as two key components:

- A central web server and reservation system
- A worldwide "reflector network", which distributes the media streams and adapts them between the supported client formats.

Clients initiate a session via a web browser page, which co-ordinates their connections to an appropriate "reflector" node. The reflectors communicate with each other via IP unicast or multicast paths In effect, the reflector network acts as a sort of "distributed MCU".

The operational model is based on "virtual rooms". At any given time, each "virtual room" can support one conference session. This therefore sets the total number of sessions that the system can accommodate simultaneously. Some rooms have worldwide scope whereas others are regional (e.g. USA-only, Europe-only, etc.). There is no such thing as a specifically "VRVS studio": any studio equipped for IP-based videoconferencing, with a web browser, has the option to utilise VRVS. VRVS also offers gateways to other videoconferencing facilities, e.g. its gateway to the Access Grid.

It should be noted that VRVS has deliberately sought to make videoconferencing available also to low-end platforms. As a consequence, it has acquired an undeserved reputation for giving low-quality results. In fact, when used with good quality equipment and with adequate network paths, the results are good and compatible with the quality of the equipment used.

There are occasional technical compatibility issues, due to non-compliant vendor implementations of the standard, e.g. *Polycom ViaVideo* version 3.0 does not work with VRVS due to some points of non-compliance with the H.323 I.T.U. standard, whereas version 2.2 does work. On the other hand, some implementation incompatibilities between different H.323 products are resolved by code in VRVS, thereby actually enhancing interoperability.

7.2 Costs

Client participation in VRVS incurs no specific cost; it requires only an appropriate client platform, an Internet connection, a web browser and the appropriate audio-visual hardware. The cost, therefore, is entirely dependent on how well equipped the studio is to be. VRVS does not specify a minimum hardware requirement.

However, if the e-Science programme were to deploy VRVS nodes extensively, then it should probably supply one or more additional reflectors. The reflectors are commodity hardware (currently Linux PCs). If necessary, virtual rooms could be created whose scope was limited to a defined community.

7.3 Ease of Use

VRVS is designed to be used by non-specialist operators (frequently the users themselves). In most meetings there is no need for a technical operator to be present. The VRVS team provides free support via an email address. Newcomers and occasional users tend to experience teething problems mainly due to local system and/or networking settings and the VRVS team recognise that although the system is easy to use when all goes according to plan, some of the error reporting and diagnostic procedures are in need of enhancement to enable users to help themselves with obvious local problems and avoid unnecessary calls to the VRVS support team. Improvements are promised in the upcoming version of the web software.

Many of the anecdotal reports of unreliability of VRVS are in fact a consequence of such teething problems, of problems with users' own equipment, and/or with local network firewalls. Regular users of VRVS tell a much more positive tale of its reliability.

When the local set-up is checked successfully then the user just has to click on a web browser to schedule, join or leave a conference.

At the scheduled time, intending clients use their web browser to voluntarily connect to a session, thereby establishing a data flow to/from their appropriate reflector node.

7.4 Display Quality

This is determined by the client equipment (and, where relevant, by any networking limitations) and not by the VRVS system itself. The decision whether to use large displays, video projectors etc. is for individual client sites: there is no specifically VRVS studio configuration, however, for the e-Science programme, projectors may be the preferred solution.

Users of H.323 by default see a single voice-switched video in their H.323 video window, but may choose other options, such as client selection of incoming video or a multi-video mode display using *vic*. The ability simultaneously to display all video streams coming from several H.323 clients is unique and is provided only via the VRVS infrastructure. Users of Mbone use *vic* and will see thumbnails of all participants who are transmitting video, and may select one or more for a multi-video display.

7.5 Visual Quality

Again this is primarily determined by the client equipment and not by VRVS itself. Although VRVS sessions are started by default at a medium bandwidth, VRVS allows different protocols at different connection rates, e.g. Mbone at 100Kbit per second or 3Mbit per second, H.323 at 128Kbit per second, or at 30 frames per second and 768Kbit per second. The closest analogy to Access Grid would be operation in Mbone mode, and then, using such equipment as Access Grid uses, a very similar result could be produced with VRVS. It is relatively simple to swap in custom versions of *vic* to allow for multiple video capture sources.

7.6 Audio Quality

Audio standard used on VRVS is normally PCM u-Law 8kHz (G.711u).

The reflector network has no problem handling other codec standards, such as the Access Grid's choice of 16kHz, if the other participants agree. VRVS does not itself transcode between standards (except in the Access Grid gateway [VAG]), but transcoding software exists, and could be incorporated if there were sufficient demand. Transcoding may be required when participants are using very different types of facilities, e.g. if a user is at home on a cable modem where uplink bandwidth is limited to about 100Kbit per second, and they wish to participate in high-bandwidth sessions.

The quality of audio is determined primarily by the client hardware, and the lowest-quality participant sets the sound quality of a conference. Well-disciplined workstation users should have the courtesy to mute their microphones when not actually speaking. This is less important in the case of studios equipped with adequate echo cancelling.

VRVS has an undeservedly poor reputation based on popular use from low-end workstation clients. When used from studio-quality equipment, the results are fine. If an unfortunate situation arises where one of the sites becomes noisy and keeps intruding into the conference, each receiving site has individual options to mute them. In future, chair control features will appear, allowing sites to be forcibly muted.

7.7 Networking Issues

VRVS conferences take place using the IP network. Networking issues can be considered in two parts: the backbone between the reflector nodes, and the local links between the clients and their local reflectors.

VRVS support can reconfigure and enhance the reflector network whenever appropriate, without it having any direct impact on the clients.

The paths between the clients and their reflectors need sufficient capacity to carry the aggregate bandwidth of all the clients who will connect to that reflector at a given time. If/when this limit is approached, then the deployment of an additional reflector node is indicated.

Since clients connect to their reflector via unicast paths, it follows that if a number of participating sites in a meeting are fully active and on a single reflector node, then the bandwidth at that reflector would scale by the square of that number. Deploying extra reflector nodes in well-chosen locations can even out concentrations of bandwidth.

It is worth noting that implementations of H.323 are intolerant of packet losses on the network: the Mbone mode of operation is relatively more resilient in that regard. This is a characteristic of implementations of the protocols, and not specific to VRVS.

7.8 Multi-Site Issues

A very common model of a conference on VRVS is of a large centre operating professional equipment (studio, seminar room, auditorium) and a number of remote spectators in groups, at desktop workstations, or even at home over cable-modems. There may be opportunities for contributions from the spectators. Sometimes the events involve more than one centre - again with numbers of remote spectators or participants.

Other configurations of meeting might be all-workstation or all-studio. VRVS is compatible with them all.

User perception is that meetings go well if sites are brought-in by turns, but that it is difficult to conduct a free-flowing multi-site discussion. That effect is common to all kinds of videoconferencing, much of it being attributable to latency. With experience, users gain familiarity in coping with it, but a residual effect inevitably remains. In comparing latency behaviour of Mbone with H.323, we must distinguish between H.323 MCUs that transcode and those that do not, because transcoding introduces additional latency. Comparing Mbone with non-transcoding H.323 MCUs, user perception is that Mbone is somewhat better. There is no reason why VRVS's Mbone should behave any differently from the Access Grid in this regard.

7.9 Collaborative Tools

All VRVS sessions include a Java-based text chat window.

Further, VRVS has chosen VNC as its portable cross-platform collaborative tool, offering sharing of application window(s) or of a complete "desktop". The VNC feature is available in any reserved virtual room if the participants wish to use it: it does not require any additional reservation.

In practice, many long-standing VRVS users have the habit of sharing web URLs to be browsed by the participants, instead of using data sharing. Those who have tried the VNC facility report good results.

We expect these results to feed back to users, resulting in increased use of shared windows or desktops. It is important to note the benefits of being able to share real application windows of any kind - and even to manipulate them 'live' and allow other participants to try them out - relative to some of the more limited sharing provisions elsewhere which only share specific proprietary format(s), such as *PowerPoint*.

7.10 Security

There are 3 levels of security within VRVS:

1. User needs to be registered to the VRVS server in order to use services. All actions are logged.

2. When a user joins a conference via VRVS, the reflector adds the user's remote host on the list of destinations needing video/audio streams. So, the administrator knows in real time which hosts receive streams and is able to remove the host if not appropriate.

3. In order to enter a particular meeting, VRVS's security provision consists of an optional password for admission to a session. If this is required, it is set up during booking.

Nevertheless, the session data streams themselves are not encrypted. Anyone connected in the same network segment as the client application or reflector to overhear the unicast streams would be able to eavesdrop the sessions.

For the future, there is a project line to investigate Firewall, NAT, Authentication and Encryption issues. Investigations into VPN are also being conducted. The next release (planned in October 2002) will include authentication via username and password.

7.11 Appropriate Usage

As a consequence of its accessibility to low-end desktops and workstations, there is a widespread but misleading perception that VRVS is "only" intended for this level of hardware. Numerically it is the case that most registered clients are desktops. However, a substantial proportion of VRVS sessions contain at least one major site (studio, auditorium, seminar-room) where the key action is taking place, and where professional-standard equipment would be in use, as well as having a number of remote participants at desktops or at home.

Some VRVS sessions, on the other hand, are entirely studio-based. In brief, VRVS can and does serve a wide range of requirements, from pure studio to pure desktop sessions, and many kinds of mix between.

As just a couple of practical examples one might take:

- The Stanford Linear Accelerator Center and Fermi National Accelerator Laboratory boardrooms are equipped with high-end H.323 *Polycom Viewstations*. These are regular VRVS users.

- CERN had equipped several meeting rooms for VRVS use by the end of 2000. Guides for self-service users of these rooms are available on the web. Their main auditorium is also equipped, and major events and other meetings are transmitted via VRVS.

Participation in VRVS is in no way exclusive of participation in other videoconferencing systems appropriate to the available equipment: MBone-capable sites might also participate in the Access Grid, while H.323-capable sites may participate in the upcoming JANET and ESnet H.323 services. VRVS neither assists - nor interferes with - any scheme of scheduling physical resources at the participating sites.

Many users of the existing VRVS are self-supported, using the web documentation and making occasional email queries; but if a wide e-Science deployment were initiated, ready access to technical support, user tutorials etc. would be essential.

7.12 Future Potential

To give some idea of the present scale of VRVS, their report as of end FY2001-2 showed:

- 6,000 individuals in 60 different countries had registered a total of over 10,000 stations; a "station" being anything from a laptop to a high-end videoconferencing suite
- The reflector network consisted of 42 reflector nodes in various places around the world.
- Total booked conference hours from May 2001 to April 2002 inclusive, were 10,288.

These figures had increased significantly over the year, and continue to increase with time. Of course, some "stations" register and become regular users, whereas others register, try it once and are never seen again.

The VRVS team are keen to see VRVS develop as a versatile interworking tool, and have presented their proposal "A Next Generation Integrated Environment for Collaborative Work Across Enhanced Networks" to the US Department of Energy (DoE).

Their next release includes:

- Faster access to web pages and functions
- User authentication
- Monitoring and tracking of usage
- Multiple communities of users
- Hundreds of virtual rooms
- Packet recovery and better aggregate bandwidth control
- Remote Mute/unmute of audio, video (chair control)
- Fixes for H.323 client incompatibilities

- Redesigned user interface to accommodate multiple communities
- Mobile user profile
- Quota management
- Mailing lists linked to booking system
- Selection of bandwidth ranges

Beyond the next release, future plans include supporting additional formats (e.g. MPEG2, MPEG4, HDTV) and future Internet protocols (SIP, IPv6).

8 Non-Studio-based Videoconferencing

8.1 About Non-Studio-based Videoconferencing

Multisite videoconferencing can be implemented with a range of hardware solutions. Solutions can involve the use of a dedicated space (studio-based) or they can make use of existing spaces, for example the desktop (non-studio-based). Whilst studio-based solutions may have cameras, microphones, speakers and other devices permanently or semi-permanently mounted, non-studio based solutions are incorporated into current working spaces in much the same way as the telephone is found on most office desks.

Non-studio-based videoconferencing can be implemented with a desktop computer, video-cam and microphone headset. More intermediate solutions can also be employed that make use of inexpensive (compared to those used in studio-based solutions) hardware codecs and echo cancelling devices. Non-studio-based videoconferencing makes use of existing network infrastructure, usually utilising the H.323 videoconferencing standard or the Mbone tools software. Examples of software applications that may be used are Microsoft NetMeeting, Mbone tools, QuickTime, CUseeMe, or Microsoft Messenger.

8.2 Costs

The cost of setting up a non-studio-based videoconferencing solution can be very low. This lack of expense results from the usage of existing infrastructure, such as desktop PCs. All that a user may need to procure are a video-cam and a microphone headset. As such, the cost may be as little as £50 for a workable solution. (£30 for a video-cam, £20 for a microphone headset and a free copy of NetMeeting bundled with Microsoft operating systems.)

Large improvements in audio quality can be attained using inexpensive echo cancellation that may be purchased for £100 or less. Hardware codec-based equipment, such as *ViaVideo* (which includes echo cancellation) may be purchased for about £400.

Similarly to the Access Grid, VRVS and H.323 solutions, running costs are subsumed into the costs for maintaining the network infrastructure.

8.3 Ease of Use

Most non-studio-based videoconferencing solutions emphasise ease of use because of the likelihood that users of this solution will be relatively isolated in terms of videoconferencing support.

8.4 Display Quality

The display quality is partially dependant upon the PC monitor used. Most monitors will suffer from a lack of pixel space for the kinds of multi-site conferences under consideration by this report. Either the software cannot display multiple video feeds simultaneously or, if it can, then the size of display for each feed will be too small to be useful.

In order to use associated collaborative tools productively with the available display space, video feeds may have to be minimised, so that the session is partially conducted with audio only.

8.5 Visual Quality

Each participant will typically send only one video feed, usually a 'head-and-shoulders' viewpoint. The quality of the video feed obviously depends on the hard- and software configuration applied, but, as an example, Microsoft NetMeeting sends CIF at 15 frames per second.

8.6 Audio Quality

The audio quality for non-studio-based solutions will depend mostly upon the hardware in use. The most frequent hardware used for this solution is a microphone headset, which obviates the need for echo cancellation. However, relatively cheap devices are available to allow for desktop echo cancellation. These only really become necessary when more than one person wishes to participate at a single site (which is difficult with this type of solution – see other sections in this chapter). Microsoft Messenger, which is bundled with Windows XP, utilises software echo cancellation and early anecdotal evidence suggests a relatively high level of audio quality.

With some applications, audio is only possible with one other site. With most hardware appropriate for this solution, full duplex is not possible.

The audio quality for non-studio-based solutions can range from approximately equivalent to that of the telephone through to much better than this with the incorporation of higher-quality speakers, microphones and computer specification, though this will of course add to the cost of this solution.

8.7 Networking Issues

Networking is via the standard existing IP infrastructure. Because most non-studio-based solutions operate using H.323 protocols, the issues discussed in *5.7 Networking Issues* apply equally here. However, some new implementations of non-studio-based solutions (e.g. Microsoft Messenger) use SIP for signalling rather than H.323.

8.8 Multi-Site Issues

Multi-site conferences using this type of solution are not as good as the results achieved when using a studio-based solution, but these type of conferences have been held with good user experience. However, not all non-studio-based solutions are suitable for this type of meeting. In order to conduct a multi-site meeting, it is usually necessary for an MCU to be used (except in the case of using VRVS, in which case software MCUs are employed).

As noted in the section on display quality, the fact that this solution typically uses a PC monitor for its display device means that it is difficult to find pixel space for multiple video feeds. This is especially the case if there are collaborative tools in use simultaneously.

Some software applications (e.g. NetMeeting) only allow audio to be conducted with one other site, but the addition of inexpensive echo cancellers and/or hardware codecs can make multi-site videoconferences outside the studio a viable option.

8.9 Collaborative Tools

Most software applications supporting non-studio-based videoconferencing have integrated within them a limited range of collaborative tools. For example, NetMeeting allows remote desktop sharing and has a shared whiteboard facility. It is also possible to use tools alongside the audio-visual tools that are used by the other videoconferencing solutions under consideration by this report, such as VNC or Distributed PowerPoint.

The integration of peripheral devices into this solution is possible. However, the cost of most such devices will exceed the cost of the other hardware necessary for this solution. Their integration is also non-trivial. Given the nature of this solution, attaching peripheral devices is not appropriate because the ease of use and cost advantages of non-studio-based videoconferencing are wiped out when these are considered. It could even be argued that integrating peripheral devices makes this solution into a studio-based one.

8.10 Security

Because most non-studio-based solutions operate using H.323 protocols, the issues discussed in *5.10 Security* apply equally here.

8.11 Appropriate Usage

One good usage for this solution is to enhance point-to-point communications to make for a richer experience than that offered by the telephone. This solution also provides a useful solution for users unable to participate in a remote collaboration via a studio and who must join a meeting via their desktop. For example, for meetings involving widely separated timezones, users may appreciate being able to participate at home late at night rather than have the inconvenience, or impossibility, of having to attend via a videoconference studio. Some users also prefer the informality of meeting from their own office environment rather than having to book a room in advance.

Due to the human factors involved, participants who join a meeting using their desktop that is largely attended by participants located in a studio may not feel as fully involved.

Non-studio-based videoconferencing is not appropriate for group-to-group collaborations.

8.12 Future Potential

It is probable that non-studio-based videoconferencing will become increasingly prevalent among desktop computer users. It is also likely that soft- and hardware to support this will continue to provide better video and audio quality. Vendors aiming at this market are mainly targeting home and small business use and may therefore concentrate upon improving ease of use and reducing price.

The support for collaborative tools for non-studio-based solutions is likely to improve. For example, it is possible that applications to do with sharing video clips and still photos between participants will be available within a couple of generations of software codecs. This is being driven by the increasing emphasis on multi-media applications such as digi-cams and scanners.

Most of the research conducted on improving sense of presence in videoconferencing scenarios advocates the use of more immersive environments that are only possible in dedicated or adapted rooms. Non-studio based videoconferencing, by its nature, is unlikely to provide a realistic sense of presence in the near or middle future.

9 Interoperability Between Videoconferencing Solutions

9.1 What is Interoperability

Interoperation is the word used to describe distributed collaborations that take place between videoconferencing solutions. For example, meetings might take place with some participants in an Access Grid node and some in an H.323 facility. The need arises for interoperability because organisations may already have invested in particular types of facility and subsequently need to collaborate with people who use a different technology. There may also be occasions where it is important for certain individuals to be present at a distributed meeting, but are unable to travel to a facility of the same type as is being used for the meeting.

However, given that interoperation may be technically difficult to achieve, as well as confusing for operators and users; that it may negatively impact upon the users' experience and that it considerably exacerbates problems of scheduling between the systems, it is a situation best avoided if at all possible. But in spite of this, there is a definite user demand for participation from office desktop/laptop or even home-based situations, and especially when international collaborations involve participation at unsocial hours. These user requirements cannot be dismissed out of hand: we need to address them as best we can, while ensuring that such mixed operation does not have an intolerable impact on the studio-based users.

There are three main areas to consider about interoperability. These are:

- Technical issues. That is, whether the interoperation can take place at all and what issues are involved in making this happen
- The user experience at each end facility and by how much this differs depending on the type of facility at which the user is present
- The issues that are involved in scheduling meetings between different types of facility.

9.2 Technical Issues

This subsection examines the technical feasibility of interoperating various combinations of conferencing solutions. It includes all the videoconferencing solutions covered by this report, including H.320 systems. Also included is interoperation with the telephone network, because a telephone may be the only form of communication technology that is available for an important participant to join a meeting.

Interoperation between systems is not necessarily a trivial task and usually involves more complex procedures than setting up meetings within systems. For example, many H.323 sites are not configured to allow traffic through the organisation's firewall. Therefore, these sites require changes to network configurations to allow sessions with other facilities.

Disclaimer

It should be noted that the authors of this report have not tested some of the interoperation scenarios examined below. In most of these instances, relevant documentation has been used that describes how interoperation takes place. In other instances (namely *Telephone and VRVS*, *Access Grid or VRVS and H.320*), we have seen no first-hand reports of these solutions in use and therefore these require further examination.

Telephone and Access Grid

Interoperation between telephone and Access Grid is accomplished using the Gentner echo canceller, which is part of the standard set up of an Access Grid node. The telephone participant phones the number of the line connected to one site's Gentner and is able to participate in the meeting with full-duplex audio between all nodes and the telephone.

Telephone and H.323

Connection to an H.323 meeting is made from a telephone by dialling the session. The voice call connection is made over ISDN. Therefore, at least one site must have their equipment connected to ISDN lines. Every videoconferencing vendor that complies with the H.323 standard can support this interoperation.

Telephone and H.320 (ISDN)

Similarly to H.323, connection to an H.320 meeting is made from a telephone by dialling the session. The only proviso is that there must be a sufficient number of ISDN channels available for all participants, including those connecting by voice only (telephone).

Telephone and VRVS

The VRVS does not itself implement a telephone gateway. However, there are various ways in which a site or community might implement a telephone gateway to VRVS, using various hardware.

These include:

- H.323/ISDN or H.323/Telephone System hardware gateway
- Some high-end H.323 stations (e.g. *Polycom Viewstation*) can make a telephone call at the same time as an H.323 connection, and thus act as the bridge between VRVS and a single telephone participant or a telephone conference service
- Quicknet cards can be used to implement an inexpensive PC-based gateway to the telephone system

Access Grid and H.323

VRVS provides a gateway that can be used to bridge between Access Grid and H.323 facilities. To use this, settings for the H.323 client (hostname or alias/gatekeeper) must be entered into the VRVS web client. Having done this, video feeds can be viewed by the H.323 client by selection, in a 'round-robin' configuration, or simultaneously (by downloading Mbone software).

Another way to interoperate between these systems is to use the JANET H.323/Mbone gateway.

Access Grid or VRVS and H.320

Access Grid and VRVS can interoperate with H.320 facilities in the same manner as each other. This section describes two possible ways that this may be achieved (but see *Disclaimer* above).

Many new H.323 commercially available hardware codecs have in-built IP/H.320 bridges. One example of this is the Tandberg 6000 product. One way to interoperate between Access Grid (or VRVS) and H.320 is for Access Grid (or VRVS) to connect to H.323 (as described above) and then use a suitable hardware codec to bridge to H.320 clients. However, by no means all H.323 hardware codecs have this facility and some, whilst able to support H.323 and H.320 protocols, cannot do both simultaneously. Such codecs are unable to interoperate with Access Grid and VRVS.

Some LANs also have LAN/H.320 gateways. These may be used in a similar way to bridge between Access Grid (or VRVS) and H.320.

Access Grid and VRVS

VRVS provides a gateway that can be used to bridge between VRVS and Access Grid facilities. As long as the meeting is predominately an Access Grid meeting, it is simple for a VRVS client to join. VRVS and Access Grid use similar metaphors of Virtual Spaces, therefore a VRVS client only has to navigate to the appropriate Access Grid Virtual Venue within VRVS to join a session. Currently not all the Access Grid Virtual Venues are supported by the VRVS gateway (and new ones have to be added via an advanced VRVS administration interface by the VRVS team). There are currently 2 VRVS gateways to support this functionality – one in ANL and a second hosted by Internet2. There is potential for further VRVS/AG Gateways to be deployed in association with the VRVS team (see recommendation 2.4).

The current standard Access Grid configuration as used in UK Access Grid nodes makes it hard for them to join VRVS meetings. We recommend investigating solutions such as:

1. Access Grid nodes to be fitted ('retro-fitted' in the case of existing facilities) with the capability to transmit and receive audio from the Display Server as well as from the Audio Server. This would be a relatively inexpensive solution (less than £50 per node) in that it would entail the installation of an additional sound card into the Display Server and connection from it to the level balancer

via 'splitter' wires. (It should be noted that Access Grid nodes supplied by *inSORS* already have a similar configuration [see recommendation 2.7].)

2. There exist daemons that relay media streams from the IP address that receives VRVS streams to another IP address and back again. These daemons may be able to relay audio streams between the Access Grid node's Display Server (which would be the default server to VRVS) and the Audio Server. A small amount of software would be developed for automatic start-up.

H.323 and H.320

As stated in the section on Access Grid and H.320 interoperation, many newer H.323 hardware codecs have the facility to bridge between the two protocols. (Older hardware codecs too can often be upgraded to have this facility.) If such a codec is present, interoperation between these systems is trivial and seamless, although as IP offers no quality of service guarantees, other network traffic may cause problems at H.323 sites. If no such codec exists at any of the participating sites, interoperation is only possible by involving a suitable site to act as a bridge.

H.323 and VRVS

VRVS also supports H.323 clients. The user requests an H.323 connection by using the VRVS web page. VRVS establishes the H.323 call to the client and negotiates the required protocols with the client (CIF H.261 video and G.711 8kHz audio, and very soon will support other codecs such as H.263 and G.721). The H.323 client can view the video in a number of different modes, including voice-activated, selected-participant, browse-video or multi-video (by downloading Mbone software). These unique features are not even available commercially from H.323 manufacturers.

9.3 User Experience

Rather than make in-depth comments on all combinations of interoperation between videoconferencing solutions (like the Technical Issues subsection), this subsection makes general comments regarding the user experience of such conferences.

One general note is that interoperability *between* systems necessarily results in lower quality conferences than staying *within* one system. This is because of the conversion algorithms that must take place to connect between different technologies. However, this quantitive reduction may not be a qualitative one, i.e. it may not be noticeable to end-users (in other words, the conversion algorithms may not impact significantly upon performance).

There are several possibilities that exist for the user experience in an inter-system videoconference. The first is that all users have an exactly similar experience to that which they would experience if the conference took place between the same types of facilities. One example of this kind of interoperation might be a VRVS facility that joins an Access Grid session. Because the systems use similar technologies, it is quite possible (though not inevitable) that users will be unaware that any interoperation is taking place at all. Users at all sites will have a similar experience of the conference.

The second is when the interoperation requires that one system, or both, must reduce the standard protocol in order for the connection to take place. This may be seen as settling on the 'lowest common denominator'. Users of one of the systems, or both, may have a reduced quality of experience to that to which they are accustomed. Another cause of reduced quality of experience for users is when telephony is mixed with videoconference calls. Telephone quality audio is substantially worse than the audio quality offered by any of the videoconferencing systems so that the participant will both hear and transmit a reduced quality of audio. This is annoying not only to the telephone participant but also to the videoconferencing participants. There is also the danger that if certain participants cannot be seen, other participants will forget that they are present.

9.4 Scheduling Systems

Any system that has a number of scarce resources that can be called upon from multiple locations needs some form of booking system. This booking system can be trivial, for example a book on a secretary’s desk, or can be very complex in the form of a fully integrated and distributed system.

A videoconferencing system that is truly desktop-to-desktop does not require a booking system per se (analogous to the telephone system) but clearly benefits from some form of shared calendar system. Studio based videoconferencing systems on the other hand are prime examples that require a robust booking system. Without such a system there can be no guarantee that studios are available, making the booking of a large multi-site conference all but impossible, as both people and facilities need to be coordinated. It is also the case that the greater number of sites that may participate in collaborative sessions, the better the booking system needs to be. As a videoconferencing system gets more heavily used then these problems get more acute, though of course a booking system can only cope with objects, rooms, MCUs etc but not people. What will work with a small community (say 10 or a dozen sites), may well not scale to 50 or 100 sites.

Each of the three studio-based solutions under consideration in this report runs a booking system. One major problem in this area is the overlap between videoconferencing solutions. For example, if a room in normally used in community A, then it will probably conform to that community’s booking system. If it is used in community B as well, it also needs to integrated into community B’s booking system. The problem then arises that it is highly unlikely that the two booking systems are capable of interaction (there no standards in common use) and problems such as double booking remain. Currently, the problem is usually ‘solved’ by having the room as a ‘guest’ of the other system but this requires manual intervention. The problem is further compounded if studios are used as normal meeting rooms in addition to remote collaborative spaces.

This problem is more extreme in the e-Science community as conferences are frequently made outside of the UK Academic community.

The goal that should be aimed for has to be to try to integrate as many of the systems as possible and to concentrate on interworking where integration is not possible

JANET Videoconferencing Booking System

The JANET videoconferencing booking system manages over 250 studios across the UK. It is highly managed in the sense that a physical room's booking should always be submitted to the booking system. Hence, if it is wanted for a non-JANET videoconferencing purpose then it ought to be booked out from the system as a whole.

The overall system is operated for UKERNA by Edinburgh and has operators that set up each videoconferencing, which removes the need for any site operators.

Access Grid Booking System

The Access Grid is based upon a virtual space metaphor. There is no need to 'book' any of these spaces, since they are implemented as multicast groups and therefore plentiful. If the metaphor is being fully employed, then collaborations are centred on a shared space (i.e. shared between a group of collaborators). These spaces are 'persistent'; perhaps holding 'objects' strongly associated with that space (e.g. data files and the soft- and hardware tools with which to examine these files).

However, Access Grid nodes are currently 'scarce', that is, demand for them outstrips supply. This means that, whilst the virtual spaces are persistent, the physical spaces are not (at least, not as Access Grid nodes). To address the resulting problem of booking scarce studios across multiple sites, the National Center for Supercomputing Applications, Illinois, has developed a web-based scheduling system specifically designed for Access Grid. This matches virtual spaces to the physical resources required for a given session (studios) and also provides a server to act as a central repository for data files associated with that session, e.g. presentation slides.

Currently, most collaboration sessions between UK sites are arranged with no formal booking system. Each site is responsible for ensuring that resources at their location are available using whatever physical room booking system is in use by that institution. The virtual space to be used for the session is agreed by negotiation between the sites (although, as all sites have a virtual space that they 'own', this is a trivial problem).

VRVS Booking System

The VRVS defines a number of "virtual rooms" (which other systems might call "conferences"), but it takes no responsibility at all for booking the physical rooms or peripheral devices necessary for a specific meeting.

The person who is making the reservation makes a booking for a "virtual room" (having chosen one which is free at the time in question, and which covers the desired geographical scope, e.g. Europe-only or worldwide) and gives the session a name and brief description with optional URL.

Optionally they may supply a password so that access may be limited to authorised participants.

The physical sites that will actually participate are made aware of the reservation out-of-band (e.g. the person booking the session contacts them via email or phone, etc., or potential participants see the session advertised in the VRVS's calendar). At the time of the session, the participants voluntarily join the meeting by using the web browser interface. It is the responsibility of users at each site to ensure that physical resources (workstations, equipment, room) are available: the VRVS system is quite unaware of what these will be, or even of which VRVS reflector they will connect to, until the interested sites initiate their connections to the session.

At present the VRVS reservation system is centralised, and provides a relatively limited number of "virtual rooms" (24 in total): this sets the current limit on the total number of simultaneous sessions which can be supported, even if the reflector network were to be scaled up. The next release will provide hundreds of virtual rooms, as well as virtual rooms dedicated for a particular community or group. The only limitation will be the network where the reflector is attached.

They have plans to support a more decentralised regional booking system. However, the principle will be similar: their booking system is for the reservation of their resources in the VRVS system, it makes no attempt to serve as a reservation system for physical resources at the participating sites.

The VRVS booking system translates date/times into each participant's preferred timezone.

10 Data and Application Sharing

10.1 Introduction

While it is obvious that real-time person-to-person communication normally requires a voice channel (as in a telephone conversation), and multi-way communication is greatly improved by one or more visual channels showing the participants, e-Science frequently demands other visual channels. These may be:

- Single pictures, charts or diagrams
- Sequences of pictures, charts or diagrams
- Shared electronic whiteboards
- 3-dimensional objects
- Video sequences
- Views of active computer applications

The source of this extra visual material may be hardcopy (paper or foils), playback hardware such as video recorders or DVD players, computer files or computers.

The way in which the material is used may require that the person presenting it can point (and be seen to point) to different parts of the material. In all cases, it may be necessary for users other than the presenter to also be able to point, or even manipulate, the material.

The approaches adopted towards providing these facilities are many and various, and vary in their applicability, platform-independence and reliability. For each of the above requirements, a number of solutions that have been used in the UK are described, together with a list of advantages and disadvantages of each. The list is not exhaustive, but covers much of current practice within the UK.

10.2 Picture Quality

One aspect has repercussions on many of the following sections, and that is picture quality. There are two main aspects to this:

- The ability to resolve detail (spatial resolution)
- The ability to handle changes in time (temporal resolution).

Whether the source is a sequence of stills (e.g. in Microsoft PowerPoint), a video, an animation or an application, the way in which different solutions pass data from the master site to the other observers is crucial.

The video output of PAL cameras has a spatial resolution of about 0.4 Mpixels and a frame update rate of 25 frames per second. A typical laptop screen has a spatial resolution of about 0.8 Mpixels and an update rate of 70 frames per second. The compressed video encodings used for video conferencing often have a spatial resolution of about 0.1 Mpixels and an update rate, for slowly changing pictures only, of about 15 frames per second. For pictures where a majority of the picture is changed from one frame to the next, the update rate can drop to seconds per frame, rather than frames per second.

As a result, while it may be easy electronically to present arbitrary video signals to the codecs or video capture system used, the result as seen at observing sites may range from useful through undesirable to unwatchable.

Most of the application-sharing systems, such as VNC, default to providing a ‘pixel-perfect’ copy of the master image at each observer site (i.e. loss-less compression is used).

While many of the systems encourage or mandate particular coding schemes that limit the bandwidth used - and so the maximum pixel/frame update rate - many do support multiple encoding schemes which, at the price of increased bandwidth, allow higher quality (in spatial or temporal resolution) images to be successfully shared. Links have been demonstrated providing full quality, stereoscopic computer-generated images between Chicago and Amsterdam; the network bandwidth used was above 500 Mbits/second, for a point-to-point connection!

10.3 Single Pictures, Charts or Diagrams

[Since sequences of such items are dealt with in section 9.4, systems that are equally happy with multiple items are deferred to that section.]

E-mail

The most common method used is for all participants to be e-mailed all items that will be required in advance.

Advantages:

Almost zero technology required

Participants can annotate (for themselves) their copies

Disadvantages:

Participants may not have seen the e-mail, or printed the material

Each participant’s notes are only seen by him/her

There is no chance of the presenter pointing to any item

Each observer must have compatible software or the material may be corrupted (e.g. wrong fonts) or not displayable at all

Visualiser (Document Camera)

It is common for permanent video conferencing systems to provide a Visualiser (also called a Document Camera), which is a pedestal-mounted camera on a baseboard. The dedicated camera captures the image of whatever is placed on the baseboard, which may be a sheet of paper, a foil (most baseboards have lighting from both above and below available), or a solid object.

Most Visualisers also have provision for the selection of one or more external video sources instead of the built-in camera. When these external inputs are provided on VGA (15-pin) connectors, the Visualiser is effectively doubling as a standards converter, and can be used as described later in this chapter.

Advantages:

Easily understood and used technology

All participants see the same object at the same time

All participants see the presenter pointing/altering the image

Pan and zoom allows relevant sections to be seen in detail

Disadvantages:

Extra cost

In the context of the Access Grid, one capture stream is used up so is not available for a camera

In the context of H320/H.323, the picture of the presenter is lost while the image from the Visualiser is selected

The quality of whole page images may not be adequate, requiring pan and zoom to be used to make sections legible

10.4 Sequences of Pictures, Charts or Diagrams

Since the majority of modern presentations use some sort of computer based picture sequencer (of which Microsoft's PowerPoint is the most commonly used, followed closely by the Web), this is perhaps the most requested type of facility. As a result a number of solutions that solely address the problem of "PowerPoint for Video Conferencing" are available in the marketplace.

Unmodified PowerPoint

It is quite possible for each presenter to e-mail the PowerPoint file(s) they will be using to all other participants (or sites) and for each site to run a local PowerPoint system. As long as a camera at the presenter's site is trained on the projected image of the talk, each site can synchronize visually.

Advantages:

Almost technology-free solution

Handles all PowerPoint features (transitions etc.)

Disadvantages:

Sites can get out of step

No chance of the presenter pointing to items on remote sites

Normal 'transported PowerPoint' problems such as missing fonts

Presenter's site has to dedicate a camera to this task

Each site has to have someone making the transitions manually

NetMeeting

When all the sites involved are using compatible Microsoft Windows systems, it is possible to use Microsoft's NetMeeting to provide a number of distributed functions. Of relevance to this paper are:

- Program Sharing
- Whiteboard
- Remote users can see presenter's mouse operation

Advantages:

Easy distribution of Microsoft Applications

Low cost

Disadvantages:

Only relevant to Windows environments (although a Unix environment may be shared via an X server such as *eXceed*)

Distributed PowerPoint (DPPT)

This is a set of wrappers provided by the Access Grid system for the Microsoft PowerPoint application that allows a single server to handle a number of presentations at a number of remote sites, with every site in a particular Virtual Venue receiving synchronized frame changes as the master for each presentation (generally the speaker giving the talk) changes slides.

The PowerPoint file itself needs to be made accessible at each remote location, and this can be achieved by e-mail or file transfer (to create a local copy of the file) or by accessing it as a Web page, in which case all sites can share a single copy.

Advantages:

Slide updates are simultaneous at all sites

No operator action required at master or client sites during talk

System is fairly robust

Picture quality is that of a local PowerPoint presentation - pixel-perfect

Disadvantages:

Operator action is required before each talk

More subtle PowerPoint transitions are not handled

Audio is possibly not handled properly

Remote PowerPoint

This is a newer development of the same sort of idea as Distributed PowerPoint and is also part of the Access Grid software suite.

Advantages:

It has a functionality that matches that of non-distributed PowerPoint

Disadvantages:

It is not yet robust enough to be recommended

Local PowerPoint relayed through Standards Converter

As with any application, the monitor signal of a computer on which a PowerPoint presentation is being run can be converted to PAL via a standards converter and this signal used as an input to the video conferencing system.

Advantages:

Simple to generate and to receive

All PowerPoint transitions accepted

No operator action required after set-up

Disadvantages:

Cost of standards converter (or Visualiser)

Need to switch video inputs to codec / video capture system

Resolution of compressed PAL may make text difficult/impossible to read

10.5 Shared Whiteboards

NetMeeting

NetMeeting provides a shared whiteboard system.

Advantages:

No extra cost

Reliable

Disadvantages:

Only works on Microsoft servers and clients

MBone tools

A shared whiteboard system was developed as part of the Mbone tools (on which the Access Grid system is based) but has been plagued by legal wrangles concerning ownership. At present, so far as the authors are aware, no decently supported shared whiteboard system is either distributed with the Access Grid toolkit, or commonly used by Access Grid users.

Other Shared Whiteboard systems

Smart hardware whiteboard systems (SmartBoard, Mimeo) can offer a robust and high-quality alternative to plain whiteboards. The difficulty is that distribution of the output may be problematic. One example solution is that many of these systems output to a PC the contents of whiteboards (either specialised equipment in the case of SmartBoard, or non-specialised in the case of Mimeo). This output may then be distributed via an application such as VNC or via a standards converter.

10.6 3-Dimensional Objects

Document Camera (Visualiser)

For solid objects, the only reasonable solution is a dedicated camera. Given that remote control of zoom and focus is required, and that the object needs to be properly lit if its image is to be clear, the simplest (if not cheapest) solution is a Visualiser (document camera). The alternative would be a regular camera and ad hoc lighting.

Advantages:

Only reasonable solution

Simple

Disadvantage:

Cost of document camera

10.7 Video Sequences

Video sources these days may either come from media dedicated to video (such as video tape in a video recorder, or DVDs in an external DVD player) or from digital sources (such as hard disk, CD or DVD drives intimately attached to a computer). In the first case the signal is available in an external form - composite, S-Video or component. In the latter, the signal is preserved in its digital form and is probably visible on a computer screen, not a TV or monitor.

From analogue (Composite or S-Video) sources over network

For composite video or S-Video sources, the natural procedure is to directly connect the available signal into the codec or video capture system. The (PAL in the UK) signal is then compressed and encoded and handled just like a camera feed.

Advantages:

Easy to connect (one off operator action)

No software required

No extra hardware required

Familiar controls for the presenter to use (PAUSE, STOP etc..)

Disadvantages:

Image quality may not be adequate after compression and encoding

Video with audio poses extra requirements of set-up and control

From digital (CD, DVD) sources over network

For sources that are logically internal to a computer (CD, DVD, hard disk) the more natural method is to display the video on a computer window and then arrange for that window to be shared with the other participants. This is dealt with in section 9.8 below, as it is in effect Application Sharing.

The alternative is to use a standards converter on a VGA output from the computer showing the video sequence, and treat that as a video source, much as described in section 10.7.1 above (if the computer has composite or S-Video outlets) or section 10.8.3 (using a standards converter) if not.

Streaming Video

A natural solution to some of the problems associated with the forms of playback mentioned above would be to use streaming video technology, which has shown that in suitable circumstances, with suitable material, good results can be obtained. At present the problem is that almost no video material that would be required in an academic meeting is available in streaming form, and the technology and knowledge required to convert from other forms into streaming video is not widely distributed.

For the present this must remain as an almost unexplored route.

From either source, using local playback

This may appear a retrograde step, but since the provision of video over the video conferencing systems is a problem, one workable solution is to distribute the material (VHS tape, computer file, etc..) in advance of the meeting and, at each site, to use local playback facilities to ensure a full video quality, full audio quality viewing.

10.8 Views of Computer Applications Running

The most general form of extra imagery for videoconference applications is the sharing of computer applications, since these may encompass stills, sequences of stills, video sequences or other imagery. A number of application-sharing systems have been developed, and many more - which seek to address the perceived deficiencies in the existing systems - are in development worldwide.

The two systems in widespread use by many user communities are NetMeeting and VNC, which are discussed below. Naturally, it is always possible to take the output from a computer running an application through a standards converter and use this as a video feed - this is also discussed.

NetMeeting

This is a Microsoft product, one component of which allows the shared viewing of applications running on Microsoft Windows systems. The presenter of the material starts the application as usual. Remote clients then become authorised to view the presenter's system. Neither the application nor the application data are moved to the client machines - just screen image changes.

Advantages:

No additional cost at either server or client

All applications are treated equivalently

Various forms of security are now supported (v3)

Multiple sites can have shared control of the application

Disadvantages:

Only works with Microsoft Windows servers and clients, or through systems like eXceed (this allows X-windows sessions to be seen)

Some types of image (e.g. video) rendered unwatchable through bandwidth limitations

VNC

This is a freely available development by AT&T that allows applications on Windows and X-Windows systems to be shared with users on Windows and X-Windows platforms. Once again all that is transferred are image changes on the machine hosting the application. VNC allows clients to have shared control of (e.g. mouse interaction with) the application.

Advantages:

No cost

Both Windows and X-Windows (Unix / Linux) supported for both server and client

Multiple sites can have shared control of the application

Disadvantages:

Some types of image (e.g. video) rendered unwatchable

At present only password security is supported

As with all software solutions, someone has to install and configure the software

Use of Standards Converter

An option that requires hardware but no software is to connect the VGA output from the computer running the application to a standards converter and thereby generate a Composite or S-Video signal that can be directly used like a camera feed. Standards converters are available zero cost (when a laptop or system unit has a video output already available), but external units vary in cost between £250 and about £3000. In addition some Visualisers have VGA to video converters built-in.

Advantages:

No software installation or configuration required

No restriction as to operating system or platform

No added time delays

Disadvantages:

Hardware may need to be purchased

Extra input on codec / video capture system is required (or a switch needs to be added)

Some types of image (e.g. video) rendered unwatchable through bandwidth / compression system limitations

Other Developments

Microsoft themselves have produced a new product called Exchange 2000 Exchange Server which contains many of the functions of NetMeeting.

The Access Grid community are actively developing extra tools in this area, under the general heading of Workstation docking. These aim to go well beyond the current features and performance of systems like VNC.

In addition, manufacturers are now producing a wider range of affordable video capture cards that will allow video signals at full PAL resolution (768 x 576 at 25 frames/sec) and even VGA/XGA (1024 x 768 at 70 frames/sec) to be captured without loss. It is therefore becoming likely that, when required, less compressed video streams may become common over IP-connected systems. This will allow both higher quality camera-captured feeds (better lip sync, more people easily accommodated on one camera) and also better handling of other sources such as digital video and computer-generated imagery.