Opportunities for Use and Development of Collaborative Tools in ATLAS

Steven Goldfarb
Dept. of Physics, University of Michigan, Ann Arbor MI, USA

Jeremy Herr
Dept. of Physics, University of Michigan, Ann Arbor MI, USA

Shawn P. McKee
Dept. of Physics, University of Michigan, Ann Arbor MI, USA

Homer A. Neal
Dept. of Physics, University of Michigan, Ann Arbor MI, USA

Thomas A. Finholt
School of Information, University of Michigan, USA

Gary M. Olson
School of Information, University of Michigan, USA

Jeremy P. Birnholtz
School of Information, University of Michigan, USA

Erik Hofer
School of Information, University of Michigan, USA

Mick Storr
Technical Training, CERN, Geneva, Switzerland

Giosue Vitaglione
Telecom Italia Laboratory, Naples, Italy

Joseph B. Hardin
Media Union, University of Michigan, Ann Arbor MI, USA

Charles Severance
Media Union, University of Michigan, Ann Arbor MI, USA

Abstract

This document presents an assessment of the current and expected needs of the ATLAS Collaboration in the development, deployment, usage, and maintenance of collaborative tools to facilitate its internal and external communications, member training, education, and public outreach. It is prepared in response to a request by the ATLAS management to investigate these needs, to survey the current status, and to propose solutions where needed. We conclude the document with a set of recommendations designed to address selected immediate needs and to position the Collaboration for the anticipated growing demands for collaborative tools in a Grid-enabled analysis environment.

TABLE OF CONTENTS

EXECUTIVE SUMMARY:

This document is written in response to a request by the ATLAS Deputy Spokesman to provide an assessment of the state of collaborative tools relevant to the experiment, an assessment of the views of the Collaboration about current support for facilitating the interactions between its members, and recommendations as to what steps should be taken prepare the experiment for anticipated future needs in this area.

Information was collected through informal interactions with members of the Collaboration, a formal survey issued to members of the Collaboration identified in consultation with the Chair of the ATLAS Computing Board, published literature, and through internal discussions among the paper's authors, many of whom have had experience in studying and facilitating major active international scientific collaborations and in advising national agencies on future directions in collaborative tool research.

Our principal findings are:

- high energy physics has made tremendous contributions to the infrastructure for collaborative tools (e.g., the world wide web and early videoconferencing systems)
- current large high energy physics collaborations present perhaps the greatest demands of any field on the breadth of human interactions required to achieve articulated scientific goals
- the deployment of current generation collaborative tools and technologies in support of the large high energy physics collider experiments lags that of several other disciplines with similar needs
- the principal tools now supporting our daily collaboration with colleagues spread around the world are e-mail, telephones, basic file transfers, and videoconferencing hardware and software, much of which is outdated and inadequate
- many members of the Collaboration have expressed significant frustration with the tools currently available
- the emergence of the Grid as the principal analysis vehicle for LHC experiments will raise a new set of demands for collaborative tools, whose nature and scope are presently unknown
- there are academic and private sector centers around the world who have the capability and track record for assisting entities like ATLAS in designing structures to improve the efficiency of internal communication and collaborative infrastructure and their expertise has not been appropriately tapped by ATLAS to date

We believe that there are strong reasons for ATLAS to carefully consider at this stage what steps should be taken to guarantee it is optimally prepared to fully involve its impressive human resource base in the commissioning and data analysis stages of the experiment. There is time available to achieve this goal and there is evidence that several institutions in the experiment understand the value of efforts in this area and are prepared to contribute to the realization of this goal.

Based on studies and analyses of the situation in ATLAS, we present below a set of recommendations for action.

- an entity should be created in ATLAS to support, improve and extend current collaborative services, and to serve as a facilitating center for institute collaborative R/D activities that will need some level of CERN linkage. This entity should immediately take steps to properly instrument key rooms in Building 40 for basic videoconferencing and web lecture archiving
- a professional assessment should be made of the detailed needs of ATLAS for collaborative tools
- a small set of demonstration projects should be launched to address specific needs, to begin to develop the network of individuals to implement future projects, and to demonstrate that modest investments in collaborative tools can advance the effectiveness of the Collaboration
- an appropriate working group should take on the task of specifying what collaborative tools are needed to make the Grid physics analysis environment truly useful to members of the Collaboration, regardless of their physical location
- a study should be implemented to examine the desirability of developing a single framework for ATLAS collaborative tools, similar to that being developed by the Earthquake Engineering community
- institutes and national groups of institutes should be encouraged by ATLAS Management to seek funding for collaborative tools through their local agencies, with a degree of coordination being provided by the ATLAS Computing Board or some other appropriate entity

Background information supporting this proposed set of actions is provided in the materials which follow. Section 1 provides an expanded overview of the problem we seek to address. Section 2 presents a vision as to how large scientific collaborations should work. It points out prior studies in this area, notes the evolution of the new cyberinfrastructure and its capabilities, and summarizes what has been learned over the past decade about collaboratories. Section 3 discusses some specific needs in ATLAS as identified by this preliminary study, reviews the perceived need for a common collaborative framework and separately considers the need for synchronous and asynchronous tools. It also reviews the importance of creating an informal collaborative environment where participants feel as if they are co-located.

Section 4 examines our current usage of collaborative tools, and discusses the results of a survey administered to the Collaboration. Section 5 reviews several current collaborative technologies, providing both an assessment of the state of the art and anticipated future developments.

Section 6 addresses the issue of the special collaborative tool requirements associated with the Grid environment. Section 7 references the outreach and education obligations of the experiment and the contributions that can be made to these efforts by advances in collaborative tools.

Section 8 then provides more details supporting the recommendations specified above.

This document is not meant to be prescriptive regarding funding decisions that ATLAS, the LHC or CERN should take. Its intent is to make sure that members of the ATLAS Collaboration have a single document that a) stresses the importance of collaborative tools for the success of the experiment, b) provides a glimpse of the progress made in this area by other scientific communities, and c) suggests a set of steps we believe should be taken soon to insure we are positioned to take advantage of one of our greatest strengths when the LHC turns on -- that of the combined intellectual power of hundreds of talented physicists in the Collaboration when they are provided the means to seamlessly interact.

This set of recommendations is made by a group of individuals, not all of whom are members of the ATLAS Collaboration, who are deeply interested in the power of a well structured collaborative environment to advance the effectiveness of scientific collaborations. We are fully aware that the precise collaborative plan developed by a scientific community must take into account the culture, practice, and unique features of that community. For these reasons we urge the ATLAS management, as one of its next steps, to ask a group of involved ATLAS scientists and CERN staff to review the contents of this whitepaper from the viewpoint of determining what steps are appropriate and best for ATLAS.

1 OVERVIEW

Ensuring efficient communication between its members is one of the most critical organizational challenges faced by large-scale experiments such as ATLAS. With over 1800 members in approximately 150 institutes spread throughout the world, the very success of the experiment will depend on having an appropriate spectrum of high quality communication and analysis tools available to individuals and working groups in the Collaboration.

It is clearly not feasible to conduct the bulk of planning, construction, computing and analysis for the experiment solely in Geneva. Indeed, this point was long ago recognized by CERN and the LHC collaborations when it was mandated that two-thirds of the computing be performed at institutes outside of CERN. Massive exchanges of technical data as well as near continuous interpersonal communication will be required for the globally dispersed members of the Collaboration not only to solve the technical problems associated with designing and building the experiment, but ultimately to analyze the data and to publish the physics results. In addition to the need for a high level of internal communication, the Collaboration also has responsibilities for outreach and education, especially for those nations that contribute directly to the support the LHC.

The World-Wide Web was invented at CERN in order to facilitate the collaboration of scientists pursuing frontier research. It has certainly proven its usefulness in addressing many of such issues faced by high energy physics. Moreover, as a result of its tremendous growth, usage, and development, new web-based tools offer a unique and growing source of solutions to HEP collaborative needs. In addition, the emergence of computational and data grids represent new potential benefits -- but also new challenges for collaboration. It is natural for us to look to these tools, coupled with technical developments in the fields of telecommunications and networking, and the advanced concepts of distance learning, as we seek to address the growing challenges of conducting experiments at the LHC.

This paper originated as a request of the ATLAS Deputy Spokesman to H. Neal for advice on what steps the Collaboration should take to prepare for its future collaborative tool needs. Neal and his colleagues in the Web Lecture Archive Project (WLAP), who had successfully recorded the GEANT4 training tutorials for ATLAS, as well as several years of the CERN Summer Student Lectures, began the process of collecting data that might inform a set of recommendations to ATLAS Management. This included information acquired from a survey of several members of the ATLAS Collaboration itself. Preliminary analysis of these results demonstrated that a deeper examination of the subject was called for. This led to the decision to request various members of the University of Michigan (UM) School of Information and the University of Michigan Media Union to join in the study and to share their perspectives on what options were available to large experiments like ATLAS.

These individuals are not new to the subject of large collaboratories. Gary Olson, Professor in the UM School of Information, is the Project Director of the

only ITR grant by the U.S. National Science Foundation devoted to a study of collaboratories. Olson is involved with the Global Accelerator Network (GAN) project, one of the ICFA initiatives, and was one of the co-organizers of a successful recent GAN workshop which had participants from most of the accelerator laboratories around the world, including the LHC. Tom Finholt is a UM Research Scientist, heads the Collaboratory for Research on Electronic Work (CREW), and leads the UM component of the George E. Brown, Jr. Network for Earthquake Engineering Simulation Grid (NEESGrid) Project, one that is deeply integrating collaborative tools into the work of scientists around the world studying earthquake phenomena. Olson and Finholt are co-PIs in the upper atmospheric research collaboratory (SPARC), which was one of the first efforts to provide advanced computer based collaborative tools to an entire community. Erik Hofer is a Collaborative Systems Specialist in the CREW Center, and Jeremy Birnholtz is a doctoral student in that Center.

Charles Severance is the author of Synco-mat, the heart of the WLAP Project, and has been a regular visitor to CERN as ATLAS has captured many of its software training tutorials. He is presently working on the development of the CHEF collaborative tools portal at the University of Michigan (UM) Media Union that will be described later in the text. Giosue Vitaglione was a CERN Technical Student from the University of Naples, and a UM staff member, who worked on the Web Lecture Archive Project along with Mick Storr, who is head of CERN Technical Training and also an active participant in the WLAP Project. Giosue, though now a staff member at Telecom Italia, is still very active in collaborative tool matters. Joseph Hardin is Clinical Assistant Professor in the UM School of Information, and Director of the CHEF Project at the UM Media Union. He is probably best known to the readers as the person who led the project at the National Center for Supercomputing Applications (NCSA) at the University of Illinois that resulted in the early browser Mosaic (licensed by Microsoft for Internet Explorer, and the prototype for Netscape's navigator). Approximately four years ago, Joseph Hardin and Tom Finholt were part of a University of Michigan team that spent a week at CERN, as the guest of Robert Cailliau, to interview leaders from the LHC and LEP experiments to learn more about the perceived collaborative needs at CERN.

This fairly extensive set of introductions for the non-ATLAS members on the author list is felt to be important so that the ATLAS readers will understand the context in which these individuals have contributed to the study reported herein.

We must also note that, in spite of the preponderance of members on the author list from the University of Michigan, there are several other ATLAS universities and laboratories with strong expertise in collaborative tools that should be called upon to participate in any final designs and implementation actions ultimately taken by ATLAS. Examples of such institutes include Lawrence Berkeley Laboratory and Cambridge University, just to mention two.

The presentation that follows is intended to give the reader some insights as to what might be possible with today's technologies, a sense as to how other globally dispersed scientific communities are dealing with problems similar to those in high energy physics, and a sense of the importance of having a strong collab-

orative framework in place as ATLAS moves into the commissioning, data taking and analysis phases. Also, in keeping with the original request of the Deputy Spokesman, we recommend a few basic initiatives that we feel should be implemented soon to insure that near-term opportunities are not missed.

In the past the high energy physics community has been able to proceed with its work, even with only modest attention being paid to collaborative tools and technology. One of the questions the modern LHC experiments pose for the community is whether this approach is viable in the context of thousands of active scientists spread around the globe utilizing a new Grid physics analysis environment.

The ATLAS Management will be faced with several major questions as it attempts to decide what attention it should give to the need for addressing the collaborative tool issues in an environment when budget demands make it difficulty to even complete key hardware components needs for the experiment. They will want to know why collaborative tools, as a separate entity, now wish to lay claim to importance rivaling that of detector construction. They will want to know why the experiment can not just get by as all prior experiments have done, why they can not just count on the inventiveness of physicists to solve such mundane problems as they always have done.

The reader will note that we are claiming that a certain threshold has been crossed, in terms of the scope of the scientific endeavor being pursued, the decentralization of the analysis efforts, and the global distribution of the participating scientists. We are claiming that it makes no sense to invest huge sums in creating the physical framework for an experiment in the current era and to tolerate a situation where those investments are not capitalized on by facilitating the interactions amongst those scientists who will ultimately be able to exploit them.

We note that the discussion of collaborative tools is at the heart of several national discussions underway about how individual groups will be able to most efficiently participate in LHC commissioning and analysis activities. We also note, with considerable enthusiasm, that many institutions are beginning to seek national funding to address these issues, with the recent USCMS/USATLAS request to the US National Science Foundation being just one example.

2 BACKGROUND AND VISION

Today is a time of great opportunity as research in high-energy physics continues to harness the advances in information technology, the grid, and emerging cyberinfrastructure. These developments can enable not only more powerful data analysis strategies but also broader and more intense collaborations within large-scale and data-rich experiments, such as ATLAS. However, improved collaboration is not an automatic consequence of adopting emerging information technologies. Specifically, without careful attention to the fit between tools and how physicists perform their work there is a high probability that tool introductions will fail. Given the importance of collaboration to ATLAS, decisions about collaborative tools should assume a high priority for the ATLAS leadership. In particular, we believe consideration of collaborative tools should focus on the evolution of a comprehensive ATLAS collaborative environment. Specifically, the text that follows asks the reader to envision the work of hundreds of scientists around the world organized and coordinated through a common, Web-based interface. This interface, accessed through a single sign-on, would provide support for human-to-human interaction, search and retrieval of data, and operation of analysis and visualization software. Successful realization of this integrated collaborative environment, as this paper suggests, involves both description of collaborative requirements and the use of these requirements to develop and select collaborative tools.

A key starting place when thinking about the form of an ATLAS collaborative environment could be the work in the U.S. and in Europe to develop network-based virtual laboratories, such as those described in the National Research Council (NRC) report *National collaboratories: Applying information technology for scientific research* and in the European Technology Assessment Network (ETAN) report *Transforming European Science through Information and Communication Technologies: Challenges and Opportunities of the Digital Age*. As described in these reports, collaboratories are expected to improve the speed and output of scientific research through Internet access to tools, data, and colleagues -- independent of time and place. In practice, most early collaboratory projects focused on remote control of distant equipment with less emphasis on data and colleagues (Finholt, 2002). Changes since 1993 in both the character of research, particularly in the high-energy physics community, and in underlying computer and network technologies suggest an opportunity to update and elaborate the original collaboratory concept. Increased attention must be directed to the end-to-end flow of data from acquisition to deposition in a data repository, and from data reuse to integration across various repositories and databases. Specifically, the time is right for a new round of collaborative investment in high-energy physics that will: deal with the unprecedented volume of data generated by high-energy experiments; broaden, via the grid, access anywhere to the growing availability and power of online research and training tools; and increase the throughput of useful data product generation and derivation.

The importance of a functional collaborative environment for ATLAS is implicit in decisions that have already been made about the structure of the experiment.

For example, as stated above, it is clearly not feasible to conduct the bulk of planning, construction, computing and analysis for ATLAS solely in Geneva. In the past, members of the physics community rose to similar challenges, and the invention of the World-Wide Web is one of the most significant examples. It is unlikely, though, that solutions for an ATLAS comprehensive collaborative environment can be produced solely by physicists. Fortunately, there is a growing body of work within other disciplines, within the open source movement, and in the private sector that provides a strong foundation for moving forward. In particular, new possibilities for integration of data, instruments, scientists, networks, software, and high-performance computing -- collectively termed "cyberinfrastructure" -- provide a context where the high-energy physics community can borrow from and contribute to larger development efforts. The importance of cyberinfrastructure for researchers is indicated both by the NSF in its recent blue ribbon commission on cyberinfrastructure (Atkins et al. 2003) and by the proliferation of projects in the US, Europe, and Asia that are attempting to build tools to produce cyberinfrastructure (e.g., the Globus Project, the National Middleware Initiative, Access Grid, and the semantic web).

2.1 THE COLLABORATIVE FUNCTION OF CYBERINFRASTRUCTURE

It is crucial to recognize that a main function of cyberinfrastructure will be to knit together disparate communities of expertise and practice required to generate breakthroughs in high-energy physics. That is, if investigations are increasingly data intensive, and expertise to produce and interpret these data is geographically distributed, and data storage itself is dispersed then short of massive co-location -- scientists will demand effective means for conducting research at a distance. This means researchers will need network-based tools for quickly and transparently locating and accessing data as well as applications for manipulating data, such as visualization and simulation codes. Similarly, researchers will need support for communicating about data and results-- not just via refereed publications and conferences-- but also through informal interactions that do not depend on physical proximity (e.g., the virtual equivalent of unplanned hallway encounters).

An important outcome of the development and use of cyberinfrastructure is the elimination of arbitrary limits on the scope and scale of research activity. Prior to the Internet, projects faced hard constraints on size, typically determined by the availability of physical space to house collaborators and equipment. Modern communication technologies, such as the phone and electronic mail, and modern transportation -- particularly jet air travel -- relax some space constraints. Yet frequent travel can produce dislocation, such as when a U.S.-based physicist is forced to leave her students and colleagues behind while conducting experiments at CERN. Similarly, media like electronic mail may lack sufficient richness to build and sustain levels of interpersonal trust required for successful collaboration. Collaboratories, by augmenting and virtually bridging shared physical space -- while preserving the functionality of co-location (e.g., presence awareness) --open the possibility for larger collaborations.

An equally critical outcome of the development and use of cyberinfrastructure is the increased potential for participation beyond researchers at mainstream institutions. For example, a number of collaboratory projects have demonstrated the value of collaboratories for outreach to K-12 audiences, such as the Bugscope system at the University of Illinois, where classrooms can reserve time on a scanning electron microscope to view insect specimens over the Web. At the undergraduate level, collaboratories can be used to introduce students to authentic research experiences at an earlier stage -- and in particular, provide exposure to data and instruments that may be unavailable on a typical undergraduate campus. Collaboratories may have special appeal to faculty and students at colleges and universities that have scarce research resources, both in terms of equipment and activity. Finally, for graduate students, collaboratories may provide a flexible means to obtain needed data and instrument time, as well as a mechanism for diversifying collegial networks. For instance, in space physics, collaboratory use has allowed first year students to engage immediately with experienced scientists in collecting data, compared with the more traditional model where students collected data on their own and later in their graduate careers.

2.2 LESSONS FROM A DECADE OF COLLABORATORY DEVELOPMENT

The Science of Collaboratories project at the University of Michigan, funded under the NSF ITR program, has identified over 70 collaboratory projects launched since 1993 (see www.scienceofcollaboratories.org). These projects fall roughly into five categories. First, the greatest proportion of projects focused on instrument sharing. For example, the NSF-funded Upper Atmospheric Research Collaboratory linked an international community of several hundred space physicists to instruments at an observatory in Greenland. Similarly, the DOE-funded EMSL collaboratory at the Pacific Northwest National Lab allows remote use of NMR machines by distant researchers. Second, a significant number of projects used collaboratory technology to create virtual research centers, such as the NIH-funded Great Lakes Regional Center for AIDS Research, which joins clinicians and scientists at several universities -- Northwestern, Minnesota, Wisconsin, and Michigan. Third, a number of the most visible projects are community data systems, such as the Protein Data Bank (PDB) and the Biomedical Informatics Research Network (BIRN). Fourth, many collaboratories have been used as virtual learning facilities, such as the CoVis project at Northwestern. Finally, a number of collaboratory efforts can be described as oriented to product development, such as the use of collaboratories for software engineering within Bell Labs.

A critical challenge for early collaboratory developers was developing tools for teleoperation of remote instruments. As early as 1992, the Collaboratory for Microscopic Digital Anatomy (a joint NSF and NIH-funded lab at UCSD) demonstrated successful operation of an intermediate voltage transmission electron microscope in San Diego by operators in Chicago. In 1993, the Upper Atmospheric Research Collaboratory (UARC, an NSF-funded project at the Univer-

sity of Michigan) demonstrated the first non-biomedical instance of teleoperation to control all-sky imagers and interferometers in Greenland from various locations in North America and Europe. Since these early demonstrations, teleoperation has become a routine element of instrument operation for many labs. For example, while early testbeds used custom-developed software, many instances of teleoperation can now be supported with commercial applications, such as Microsoft NetMeeting and Polycom videoconferencing units, or through public domain applications like VNC (Virtual Network Computing) and the Access Grid.

While not yet universal, most technical barriers to teleoperation are understood and have solutions -- assuming the cooperation of instrument manufacturers. Today, with modest effort, most labs that want to can support remote or automatic control of instruments. That is, following a philosophy at the Stanford Synchrotron Lab: If a device is computer controlled within an integrated systems architecture, then this creates the possibility for a completely new mode of data-centric operation where instrument instructions are executed by "smart" systems (i.e., in response to data being collected).

2.3   THE BROADER COLLABORATORY VISION: KNOWLEDGE ENVIRONMENTS FOR SCIENCE

The frontier for collaboratory development thus has moved to address the broader aspects of the original collaboratory vision, such as data integration, data access, and tools to support collaboration with data. A recent NSF workshop characterized this broader vision in terms of "knowledge environments for science" (Atkins, Olson, & Killeen, 2002). For example, the UARC system described above has evolved from a primary focus on real-time data gathering using remote facilities -- to a collaborative environment used largely to build views from multiple data sources over interesting data intervals.

Implementing the capabilities to realize knowledge environments for science will mean attention to an additional set of requirements. First, because data are generated from procedures or sources that may be unknown or invisible (i.e., not monitored by a person -- such as automatic data collection), there is a new need for assurance about data integrity and validity. Second, the increased emphasis on automated data and secondary use of data leads to a much greater effort to define and implement meta-data formats (i.e., the data about data, such as who collected the data, calibration values and so forth). Third, because by-products of data analysis can have significant economic and intellectual value, much greater attention must be placed on implementing security, such as access control and verification, as well as on authentication (i.e., assuring that a given data file does contain what it claims to contain, and these contents have not been altered). Finally, because data analysis and interpretation demands a much higher level of interaction among collaborators, tools to support collaboration at a distance need to meet expectations for scientific dis-

course, which are often formed from experience in co-located settings. For example, there is a higher demand in virtual environments to convey information that helps two distant collaborators successfully orient their attention to the same observation or region of data, since cues such as pointing and direction of gaze may be absent. Specifically, the most elegant remote control and data transfer capabilities will be wasted if collaboratories and related technologies don't allow convenient and normal human interaction. The importance of supporting normal interaction at a distance, such as for management and coordination of geographically distributed projects, is only likely to increase with the increased emphasis on projects that span institutions (e.g., the U.S. National Institutes of Health recent announcement of "regional centers of excellence").

2.4 ITERATIVE DESIGN AND EVALUATION TO GUARANTEE SUCCESS

Use of a user-centered, iterative design strategy is the best way to ensure the successful design and deployment of a comprehensive collaborative environment for entities like ATLAS. This strategy consists of the following steps.

First, at the outset of a project the current needs and requirements of users are assessed via interviews and observations of collaborative work practices. After establishing the needs and requirements of the users, a suite of technologies is recommended. Upon agreement of the appropriateness of these technologies they are deployed, users are trained, and support personnel are trained on how to maintain the tools.

Second, the uses of the technologies are then monitored on a continuous basis. Where appropriate, logs of usage within the technologies themselves can be gathered, in order to understand what functions were used and what troubles, if any, were encountered. This is standard usability engineering practice (Olson & Olson, 1991; Nielsen, 1993; Mayhew, 1999). Of course it is important to be sensitive to the privacy concerns of the users. Such logs of usage should be augmented by direct observation, surveys, and interviews. On the basis of this ongoing evaluation of usage the technologies are revised as needed to ensure that appropriate tools are available for the kind of work being done.

3 SPECIAL NEEDS FOR PROJECTS LIKE ATLAS

A successful knowledge environment for ATLAS must meet needs in four critical areas:

a) Data sharing -- which includes the creation and curation of data repositories, security and authentication, and tools for collaborative visualization and analysis of data;

b) Communications and control -- which includes video, audio, and data conferencing, and also remote control and observation of instruments (e.g., teleoperation and teleobservation);

c) Coordination -- which includes scheduling experiments and computer runs; and

d) Technology development -- continued elaboration of hardware and software, particularly to incorporate collaboration, such as ensuring that new collaborative technologies are compatible with emerging middleware standards (e.g., the Globus toolkits).

3.1 A COMMON COLLABORATIVE FRAMEWORK

While the needs described above could be met through separate but related stand-alone applications, a less risky approach focuses on core functionality provided through an integrated environment. Specifically, there are great advantages to working within a system of common and standardized protocols. For example, single sign-on and shared look and feel across applications improves usability. Also, the creation of documented and consistent application programming interfaces increases the likelihood of adoption by providing a common framework for integrating disparate user applications. For instance, the NEESgrid project, which is the system integration element of the George E. Brown, Jr. Network for Earthquake Engineering Simulation (NEES, an NSF Major Research Equipment project in structural engineering), uses the CompreHensive CollaborativE Framework (CHEF) developed at the University of Michigan as a uniform user environment for a grid-enabled system to acquire and store physical and numerical simulation data within the earthquake engineering community. Through CHEF, NEESgrid obtains a broad range of collaborative tools (e.g., presence awareness, scheduling, threaded discussion, document version control, shared data viewers) that are accessed through a common user interface and are controlled using a single sign-on that uses grid security protocols. The main benefit of CHEF for NEESgrid users is the availability of a Web-based environment that provides customizable configuration of collaborative spaces for teams of collaborators. The interoperability with software produced by the various middleware initiatives means that NEESgrid provides secure group workspaces that include the data and resource discovery services obtained through computational grids.

One vision for the ATLAS collaborative environment, then, would be to follow the NEESgrid example and use CHEF to coordinate user access to collabora-

tive tools across the breadth of ATLAS activities. These activities might include the following representative uses:

- coordination of formal meetings and production time lines through a group scheduling application;
- concurrent videoconferencing and data conferencing so that participants in a virtual meeting can see and hear each other, and also view common documents, presentations, or diagrams;
- searching an email archive to locate a URL recently shared within a collaborating group;
- using a shared and secure group file space to access common files -- thus eliminating dependence on email attachments for file exchange;
- notification services that alert members of a collaboration about new content on a threaded discussion list or about additions to the group file space;
- presence awareness so that collaborators can meet spontaneously in a virtual workspace and then converse via chat, or start an audio, video, or data conference;
- mechanisms for distributed presentations -- both live broadcast (e.g., with slides and voice over IP) and captured for later replay.

Given these potential uses for collaborative technology within the ATLAS experiment, three kinds of functionality emerge as particularly important: a) tools for synchronous collaboration; b) tools for awareness; and c) tools for asynchronous collaboration.

3.1.1   Synchronous collaborative tools

Videoconferencing tools. The physics researchers are already regular users of the Virtual Room Videoconferencing System (VRVS, www.vrvs.org/Doc/faq.html). This is a system put into production within the physics community in 1997 as an alternative to high-cost H.320 (ISDN) systems. VRVS provides "a low cost, bandwidth-efficient, extensible means of videoconferencing and remote collaboration over networks within the High Energy and Nuclear Physics communities." For example, equipment requirements are minimal (e.g., a workstation with microphone, speakers, camera, and video card) and data flows over the network -- as opposed to dedicated phone lines in the H.320 case. The VRVS software is a component-based package, integrating its own reservation system with public-domain audio, video and application-sharing packages.

Survey responses from ATLAS institutions, discussed in detail below, indicate support for and participation in distributed meetings that use VRVS. However, responses also indicate low tolerance for poor quality audio and video and frustration with the lack of established social protocols (e.g., turn taking, conventions for distributing advance meeting materials, and norms for camera placement and operation). A comprehensive collaborative environment for ATLAS should ensure that standardized hardware, software, and room environments are deployed for video conferencing at each site -- and that these configurations are optimized. Standardization is the first step toward isolating problems (i.e., determining what is due to local network congestion and what is due to idiosyncrasies of video and audio codecs) and toward development of

shared knowledge and experience (i.e., for training and maintenance). In terms of optimization, there are technological advances in networking that can significantly improve the quality and robustness of videoconferencing, such as the authenticated Quality of Service (QoS) efforts described below.

Given the existing use of VRVS, it is likely one would want to provide ways of accessing VRVS in the collaborative environment rather than replicating its functionality. However, we recommend the appointment of a formal conferencing facilitator at each site. Such a facilitator would have responsibility for setting up equipment, supervising operations, and participating in testing and development across the ATLAS Collaboration. For example, the Commons Project operated by Internet2 is one model for a facilitator program. Under the Commons guidelines, institutions that wish to use shared resources (such as Internet2-supplied multipoint connection units for H.323 conferencing) must designate an institutional representative who will attend mandatory training, join in large-scale tests of Commons systems (e.g., the annual Mega-Conferences), and act as a liaison and support person for local video conferencing users.

Application sharing and data conferencing. Experience from other collaboratories suggests the importance of tools to share material in real time. The materials to be shared could include documents, images, drawings, spreadsheets, data bases, web pages, source code, in short, any digital work object. What is key is the ability to share the same view among all participants, for all participants to be able to point to things in the materials, and under appropriate controls, for all participants to have the possibility of editing. There are a variety of commercial and public domain products available in this area. Specifically, there is significant evidence that the data conferencing industry is poised for dramatic growth. Indeed, the market research company Frost and Sullivan (Edwards, 200) predict that data conferencing services in the US will jump from $62 million in 2000 to $238 million in 2003. For example, within Honeywell, a Fortune 500 avionics and control systems manufacturer based in the U.S., data conferencing applications are used to conduct Web-based seminars and project meetings within geographically dispersed engineering teams (Edwards, 2001). In one group, data conferencing is used to conduct up to 20 meetings per month -- with average savings of US$1,200 per participant.

Whiteboard. A special kind of collaborative application is a shared whiteboard. This is a space in which the team members can sketch and write informally, much as on a real whiteboard.

Instant messaging. Instant messaging (IM), or chat, is one of the fastest growing collaborative tools. It offers informal, text-based conversations that are more interactive than e-mail but less demanding than voice. There has been considerable publicity about the rise of instant messaging for social chat, but now in a number of companies it is being used to help coordinate work (Handel and Herbsleb, 2002; Isaacs, et al., 2002). Messaging is also possible from PDAs and even cell phones, making it a simple, portable coordination tool. A variety of supplementary tools to help manage multi-party messaging are appearing in the research literature (e.g. Ribak, et al. 2002). Finally, the "buddy lists" associated with instant messaging have turned out to be a very powerful awareness

tool when used by key members of a team or organization. A major problem in the widespread adoption of instant messaging tools is the difficulty in matching handles to particular users. Since many of these services are very heavily used, it is often impossible for a user to obtain the same handle he/she uses for email on a public IM system like AOL. The rise of Grid computing provides a solution to this problem as all users will be unique within a Grid context allowing tools to be built that map unique Grid identities to individuals.

3.1.2   Awareness information and services.

"Awareness information," in a collaborative sense, consists of the cues used to determine whether someone is available to talk or meet. For example, in a col-located work setting, a simple walk down an office corridor provides numerous indications about the availability of co-workers, such as whether doors are open or closed. These cues are important because they form the basis for unplanned encounters. Research suggests that unplanned encounters, like a hallway conversation, are critical as occasions for collaborators to update each other about progress on a joint task -- or as opportunities to explore new directions for future collaboration. Obtaining the cues that lead to encounters in a virtual setting is difficult. As a result, a number of efforts have attempted to develop technology to provide awareness information among remote collaborators. These applications are collectively termed "awareness services."

One class of awareness service allows users to visually roam the hallways at remote locations using remote video cameras to make brief glances into each office (Bellotti & Dourish, 1997; Fish, Kraut, Root & Rice, 1993). Another similar service, called Portholes, provides periodic snapshots instead of full-motion video (Dourish & Bly, 1992). Because of privacy implications, these systems have had mixed success. The places in which this succeeds are those in which the individuals seem to have a reciprocal need to be aware of each other's presence, and a sense of cooperation and coordination. A contrasting case is instant messaging (IM) where users have control as to what state they wish to advertise to their co-workers about their availability. The video systems are much more light-weight to the user but more intrusive; IM gives the user more control but require intention in action.

As mentioned earlier, instant messaging systems provide an awareness capability. Most systems display a list of "buddies" and whether they are currently on line or not. Nardi and her colleagues (2000) found that people liked this aspect of IM. And, since wireless has allowed constant connectivity of mobile devices like PDAs, this use of tracking others is likely to grow. But again, there are issues of monitoring for useful or insidious purposes, and the issues of trust and privacy loom large (see Godefroid, Herbsleb, Jagadeesan & Li, 2000).

3.1.3   Asynchronous collaborative tools

Group calendaring. A number of organizations have now adopted on-line calendars, mainly in order to view people's schedules to arrange meetings. The calendars also allow a form of awareness, allowing people to see if a person who is not present is expected back soon. Individuals benefit only insofar as they offload scheduling meetings to others, as to an administrative assistant,

who can write as well as read the calendar. And, in some systems the individual can schedule private time, blocking the time but not revealing to others their whereabouts. By this description, on-line calendaring is a classic case of what Grudin (1989) warns against, a misalignment of costs and benefits: the individual puts in the effort to record his/her appointments so that another, in this case a manager or coworker, can benefit from ease of scheduling. However, since the early introduction of electronic calendaring systems, many organizations have found successful adoption (Mosier & Tammaro, 1997; Grudin & Palen, 1995; Palen & Grudin, in press). Apparently such success requires a culture of sharing and accessibility, something that exists in some organizations and not in others (Lange, 1992; Ehrlich, 1987). CHEF provides tools and services for group calendaring.

E-mail coordination. This community is already very familiar with email, but we anticipate that many groups will choose to deploy other tools to help manage group communication over email. Group email lists that generate a web-accessible archive of both the message and any multimedia attachments will likely be a major benefit. The CHEF environment currently supports e-mail list management and archiving.

Code and document repositories. As with most software development communities of practice, high energy physicists engaged in developing software are fluent in configuration management systems, specifically CVS. CVS automatically generates logs of activity -- both who acts on code and the nature of the changes that have been made -- and these logs are an information source that can be tapped to make the versioning system into a true collaborative tool. Outside of development work, document repositories are another important coordination tool. These repositories are often web-based with technologies like WebDAV providing versioning and usability enhancements, differentiating them from simple shared file spaces.

Lecture capture and distributed training. Web based archiving has been shown to be a very useful approach for capturing content-rich information within large enterprises like ATLAS for training and outreach purposes (Schocken, 2001; Emery, n.d.). Complicated applications such as GEANT4 and ATHENA may be well understood initially by only a few experts within the experiment, who inherit the dual burden of continuing to develop the packages while simultaneously trying to educate hundreds of colleagues spread around the globe on the intricacies of using the packages. In such cases a direct investment in carefully preparing the needed web based training materials can yield large returns. Examples of the role that web based tools can play in this area are described below.

4   CURRENT COLLABORATIVE TOOL USE WITHIN ATLAS

The current base of the ATLAS Collaboration's routine internal communications consists of e-mail, telephone calls, publications, and meetings. In addition, some subset of the ATLAS Collaboration uses more exotic technologies, such as VRVS (described earlier). Each of these technologies is useful, essential and, in certain specific environments, sufficient. But there are needs not being met by the existing technologies, either because of limitations of the technologies themselves or the lack of their careful integration. In this section we describe some current needs for ATLAS, first by describing a set of representative use cases and then through a description of results from a collaborative technology survey completed by ATLAS institutions.

4.1 TYPICAL USE-CASE SCENARIOS FOR ATLAS

We identify the following as representative contemporary use-cases for collaborative tools by the ATLAS Collaboration.

4.1.1   One-to-one synchronous conferencing

This typically involves real-time phone or video communication between two individuals or two very small groups of individuals. It might also include the use of application sharing tools or white boards to enhance communication. Possible scenarios include the debugging of a detector hardware problem by a remote expert or the remote instruction of a graduate student by her/his advisor.

4.1.2   Remote synchronous conferencing

This includes the remote participation of individuals in a regularly scheduled group meeting. This might include active or passive participation of the remote participants, with the minimal requirement of the broadcasting of a high-quality audio stream from the meeting. It often includes the need to view slide presentations made in the conference room and might require the integration of phone and videoconferences.

4.1.3   Multipoint synchronous conferencing

This is used for a regularly scheduled group meeting with no central meeting location. This is a relatively new, but common scenario for the LHC collaborations. It requires high quality audio, preferably with duplex microphone capabilities, and would often benefit from the ability for remote sites to be able to share slide presentations or other video signals with the other participants.

4.1.4 One-to-many synchronous conferencing

This might include a passive remote audience, attending a seminar, but it might also require active remote participation, including remote presentation of slides. This scenario differs from the conference room scenario in that the emphasis is

typically on the broadcast and/or archiving of a formal presentation, such as a seminar or series of presentations, as in an ATLAS plenary session.

4.1.5 Individual viewing of archived meetings, lectures, or training tutorials

In this scenario, there is a single participant viewing a lecture or taking a tutorial at her/his leisure. Caching of the audio and/or video signal is possible in order to obtain a high quality presentation. The ability to search the archives is desirable and the computing platform, operating system, viewing applications, and network capabilities of each user often varies.

4.2 SURVEY OF ATLAS COLLABORATIVE TOOL USE

As the use cases above suggest, there are a number of collaborative tools currently in use within ATLAS. To acquire as much information as possible about the current practices and future aspirations of the members of the ATLAS Collaboration, we circulated a survey among the participating institutes to determine the nature and adequacy of existing collaborative tools. The survey questions and the results from the survey are presented in Appendix I. In this paper, we base our statements primarily on the survey responses, as well as on our own personal observations of those of us who are members of the Collaboration.

4.2.1   Method

In May 2002 a representative from each of the 151 participating institutions was invited to complete a 24 item Web-based questionnaire. Institutions received no incentive for completing the questionnaire, but the invitation stressed the importance of responding in order to improve the performance of collaborative tools in use within ATLAS. In June 2002 a follow-up message was sent to a more selected sample of participating institutions who are involved with the ATLAS Computing Board in an effort to increase the level of response. In this follow-up, institutions were given the option of completing the original Web-based questionnaire, or of filling out an email or hardcopy version. After the follow-up the usable response rate was 16% (24 out of 151 institutions completed a response). Of the responding institutions, 67% were European, 25% were North American, and 7% were elsewhere -- compared to the total sample of individuals asked to complete the survey, where 58% of the institutions are European, 27% are North American, and 15% are elsewhere.

4.2.2 Results

4.2.2.1   Geography and ATLAS participation

As the first line of Table 1 shows, European institutes reported more Ph.D.-level personnel affiliated with ATLAS ($M_{European}$ = 10.8) than did non-European institutes ($M_{Non-European}$ = 6.4). The next three lines in the table indicate the reported number of trips to CERN each year, per ATLAS scientist at each

responding institute, classified in terms of: one day meetings; visits shorter than a week; and visits longer than a week. On each measure, European institutes reported more frequent travel to CERN than non-European institutes, with greater disparity in terms of day trips ($M_{European}$ = 1.2 vs. $M_{Non-European}$ = 0.2) and visits shorter than one week ($M_{European}$ = 3.0 vs. $M_{Non-European}$ = 1.5), and less disparity in terms of visits longer than one week ($M_{European}$ = 1.5 vs. $M_{Non-European}$ = 1.1). The fifth line shows the number of resident staff at CERN, where European institutes reported more staff ($M_{European}$ = 3.0) than non-European institutes ($M_{Non-European}$ = 0.8). Finally, the sixth line shows the number of remote meetings attended via VRVS (the video conferencing technology described previously). Paradoxically, European institutes reported a greater frequency of VRVS meetings per scientist ($M_{European}$ = 1.7) than non-European institutes ($M_{Non-European}$ = 0.6).

The disparities in the number of ATLAS-affiliated scientists and amount of travel to CERN, as summarized in Table 1, suggest that participation by non-European institutions in ATLAS is qualitatively different from participation by European institutions. In order to more fully realize the benefits brought by non-European institutes to the Collaboration, their ability to participate should be improved -- such as through the deployment and use of a comprehensive collaborative environment (i.e., that might reduce differences due to geographic distance). There is some evidence that non-European institutions recognize the potential helpfulness of collaborative tools in that 86% of non-European institutes reported a willingness to pay a small annual fee to support collaborative capabilities within ATLAS. Though only 50% of European institutes reported a willingness to pay a similar fee, this may be due to the large amount of CERN support already being provided by the home country governments of these institutes. On the other hand, it may also be the case that these institutes are content with their proximity to CERN and do not feel that an investment in collaborative tools would yield substantial benefits.

Table 1

Comparison of participation between non-European and European Institutes (n=24)

	Non-European			European			Total		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
1. Number of scientists at your institute affiliated with ATLAS (Ph.D. level)	6.3	2.4	8	10.8	6.9	16	9.3	6.2	24
2. Number of trips to CERN each year to attend a 1 day meeting (per scientist)	.2	.4	8	1.2	1.6	16	.9	1.4	24
3. Number of trips to CERN each year of less than one week duration (per scientist)	1.5	1.3	8	3.0	2.9	16	3.0	2.6	24
4. Number of trips to CERN each year of greater than one week duration (per scientist)	1.1	1.2	8	1.5	1.3	16	1.4	1.3	24
5. Number of staff from your institute at CERN at any given time	.8	.9	8	3.0	3.0	16	2.3	2.7	24
6. Number of remote meetings attended via VRVS (per scientist)	.6	.3	5	1.7	2.4	16	1.5	2.1	21

4.2.3   Current use of videoconferencing

Turning to the current use of videoconferencing, there are three important dimensions to consider: 1) the overall usage of videoconferencing by ATLAS institutes, 2) the variation in videoconferencing perceptions as a function of the

level of involvement with ATLAS; and 3) the effect of having local VRVS experts on perceptions and use of videoconferencing.

4.2.3.1 Overall usage

In terms of overall frequency of videoconferencing, institutes reported using VRVS to participate in ATLAS-related meetings more than once a month ($M_{VRVS\ per\ year}$ = 14.0). In terms of attitudes toward use of VRVS, a majority of institutes (61%) reported that they found VRVS helpful. However, institutes reported a low percentage of satisfactory VRVS sessions ($M_{\%\ sessions\ judged\ satisfactory}$ = 33%). Specifically, institutes reported a number of technical and social problems. On the technical side, 56% of the institutes reported dissatisfaction with audio and video quality during videoconferences. In addition, significant numbers of the responding institutes reported problems with network reliability (33%), equipment availability (33%), and usability (25%). Comments from institutes reflected these difficulties:

"The technology is poor. It is hard to take part in a videoconference as often the background noise is higher than the voices (this is especially true when an individual is connected with a group meeting at CERN)."

"Contributions from the audience are often inaudible."

"Resolution is a problem, it is not easy to see any details in figures."

"Some question if reliability is good enough when many sites are involved."

On the social side, 46% of the institutes reported dissatisfaction with the organization and conduct of videoconferences, such as awkward turn taking, incomplete distribution of materials before meetings, and poor camera control (e.g., not pointing at speakers). Again, comments from the institutes reflected these problems:

"The ethos has to change to make ALL slides available electronically before ALL meetings."

"We need simple education of videoconference participants in videoconference use (the voice-activated priority system sometimes leads to shouting matches -- we need protocols, just like a formal meeting)."

"The cameras in meeting rooms need to be operated professionally."

Despite a number of reported deficiencies and overall poor experience using videoconferencing, the institutes were unanimous in their willingness to use videoconferencing -- if technical and social problems are fixed.

4.2.3.2 Variation in videoconferencing perceptions with number of ATLAS-affiliated scientists

Table 2 compares perceptions of videoconferencing by number of ATLAS-affiliated scientists. Institutes classified as "largest number" were in the top third of

responding institutes, in terms of affiliated scientists, while institutes classified as "smallest number" were in the bottom third. The first line of Table 2 shows that a higher percentage of the "largest number" institutes (67%) reported having someone familiar with VRVS compared to the "smallest number" institutes (33%). The second line shows that a lower percentage of the "large number" institutes (14%) reported that Building 40 at CERN is adequately equipped for videoconferencing compared to the "smallest number" institutes (75%). Finally, the third line shows that a higher percentage of the "largest number" institutes (71%) reported problems with social aspects of videoconferences compared to

Table 2

Comparison of videoconferencing perceptions by number of scientists affiliated with ATLAS

	Smallest Number		Largest Number		Total	
	Percent of institutes	N	Percent of institutes	N	Percent of institutes	N
1. A member of your group has good familiarity with VRVS	33%	6	67%	6	50%	12
2. Building 40 at CERN is adequately equipped with video-conferencing facilities	75%	4	14%	7	36%	11
3. Social protocol issues were reported as a videoconferencing problem	29%	7	71%	7	50%	14

the "smallest number" institutes (29%). This result is particularly interesting because the "largest number" institutes reported better access to a local video-conferencing expert -- suggesting that dissatisfaction with social aspects of videoconferencing may be a second order effect that becomes salient only after an institute reaches a certain level of technical videoconferencing capability. Similarly, if social problems are apparent only above a given frequency of use, the survey may underestimate social difficulties with videoconferencing due to under-sampling of high use institutes.

Table 2

Comparison of videoconferencing perceptions by number of scientists affiliated with ATLAS

	Smallest Number Percent of institutes	N	Largest Number Percent of institutes	N	Total Percent of institutes	N
1. A member of your group has good familiarity with VRVS	33%	6	67%	6	50%	12
2. Building 40 at CERN is adequately equipped with video-conferencing facili-ties	75%	4	14%	7	36%	11
3. Social protocol issues were reported as a videoconferenc-ing problem	29%	7	71%	7	50%	14

4.2.3.3   The effects of local VRVS expertise

As indicated above, institutes reported differences in terms of the availability of local VRVS experts. Therefore, this section examines the effect of having a local VRVS expert on both use of videoconferencing and on travel. The first line of Table 3 shows more frequent use of VRVS to participate in ATLAS meetings, per scientist, at institutes with a VRVS expert ($M_{VRVS\ expert}$ = 2.2) compared to institutes without an expert ($M_{No\ VRVS\ expert}$ = 0.3). The second and third lines show less travel to CERN for short visits, per scientist, at institutes with a VRVS expert ($M_{VRVS\ expert,\ 1\ day}$ = 0.7; $M_{VRVS\ expert,\ <\ 1\ week}$ = 2.3) compared to insti-tutes without an expert ($M_{No\ VRVS\ expert,\ 1\ day}$ = 1.3; $M_{No\ VRVS\ expert,\ <\ 1\ week}$ = 3.7). Finally, the last line shows roughly comparable levels of travel to CERN for long visits, per scientist, at institutes with and without a VRVS expert. These results suggest that the presence of a VRVS expert can be beneficial in encour-aging use of the technology, and that the presence of such an individual can fur-ther lead to reduced amounts of travel, at least for short duration visits to CERN.

Table 3

Comparison of videoconferencing and meeting behavior between institutes with and without a group member familiar with VRVS

	Nobody familiar with VRVS			At least 1 person familiar with VRVS			Total		
	Mean	SD	N	Mean	SD	N	Mean	SD	N
1. Approximate number of meetings attended via VRVS (in person meetings, per scientist)	.3	.6	7	2.2	2.4	13	1.5	2.1	20
2. Approximate number of person-trips to CERN per year, per scientist primarily to attend a 1 day meeting	1.3	1.6	8	.74	1.3	14	.9	1.4	22
3. Approximate number of person-trips to CERN per year, per scientist that are less than 1 week in duration	3.7	3.5	8	2.3	1.3	14	2.8	1.4	22
4. Approximate number of person-trips to CERN per year, per scientist that are greater than 1 week in duration	1.3	1.1	8	1.2	1.2	14	1.2	1.1	22

5 THE CURRENT STATE OF THE ART

5.1 INTRODUCTION

This section describes three technology efforts within the current ATLAS Collaboration that are likely to shape the direction and functionality of a comprehensive collaborative environment for ATLAS. These efforts are in the areas of videoconferencing, authenticated quality of service (QoS), and web capture of presentations. While these technology developments are crucial they do not constitute a complete approach to collaboration within the ATLAS experiment. The roadmap for that development was described earlier in Section 3. Rather, the point here is to explore in detail a set of tools that have been deployed in the ATLAS context with the intent of showing both the sophistication of these technologies and the effort required to produce reliable and usable systems.

Videoconferencing, authenticated QoS, and web capture are under intense development within the commercial sector and some of the needs of ATLAS will likely be covered by off-the-shelf software in the years ahead. However, it is likely that the special requirements of the ATLAS Collaboration will require initiatives -- such as those described below -- that are undertaken within the high-energy physics community and in collaboration with other scientific communities. Special requirements include a high dependence on networking across institutions with connections of varying quality and reliability, and a need to support applications and protocols across a diverse mix of computing platforms and operating systems. In addition, resources for deploying and upgrading information technology are limited, such that the equipment replacement cycle is much longer than in the private sector. Therefore, within the high-energy physics community there is an emphasis on extracting maximal utility out of systems that are inexpensive, while at the same time demanding a high degree of stability and reliability in a heterogeneous environment. Finally, given the long time periods that are spanned by experiments like ATLAS, there is concern about the potential playback obsolescence of recorded media.

5.2 VIDEOCONFERENCING

5.2.1 *Introduction*

Within the global research community several driving and enabling factors are pushing videoconferencing into the mainstream of scientific practice. That is, while travel remains an unavoidable reality within globally distributed collaborations, the events of September 11th have fueled a renewed interest in adequate substitutes for travel. Indeed, a September 19, 2001 study by the National Business Travel Association revealed that 88% of companies surveyed planned to increase their utilization of videoconferencing (NBTA Press Release, 2001).

The movement from meetings to virtual meetings is not without hurdles, however. Despite the emergence of numerous high-speed regional research networks and connections between them, high quality network-based videoconferencing remains elusive. In many cases, the problems are not in moving data across a wide-area network, but rather in local institutional networks. In addition, negative perceptions of network-based videoconferencing often reflect insufficient lighting, inadequate microphone coverage and balance, poor echo cancellation, and haphazard camera positioning (e.g., see the earlier quotes in Section 4 pulled from the survey data). Finally, building systems that support a variety of videoconferencing protocols - as is done currently using VRVS within ATLAS - makes it difficult to support all of the capabilities of any individual conferencing system (e.g. proprietary algorithms for high quality video and audio, unique data collaborative capabilities, etc.). Therefore, improving the experience of network-based videoconferencing involves a coordinated approach across a number of dimensions, including network performance, the design of videoconferencing rooms, and the selection of videoconferencing systems that are standards compliant and do not rely on proprietary extensions for added functionality or quality. The following section describes efforts to date in the context of VRVS.

5.2.2　State of the Art in ATLAS: VRVS

The strength of VRVS is that it has managed to leverage existing effort in both industry and in the research community to provide high quality IP-based video conferencing. VRVS has focused on integrating conferencing tools and standards to allow users to conference using whatever technology is optimal at a local site rather than generating new tools from scratch. These tools are tied together using a "room" metaphor so that an end user is able to conference from a web-based portal without being encumbered by the underlying technology. VRVS currently supports three different, widely adopted video conferencing tools. All are well suited to different environments and have different strengths and weaknesses. This section reviews each of the technologies and highlights the best applications for each.

5.2.2.1 Mbone tools: vic and rat

The best-known method for connecting to VRVS is through the public domain vic (video conference) and rat (robust audio tool) tools, that run on a number of platforms (i.e., Windows, Unix and Mac OS X). These tools were originally developed to support multicast-based MBone conferencing, but do not require multicast to connect to VRVS. When used with an inexpensive web-camera and a headset from a desktop workstation, vic and rat work well at low cost. However, vic and rat are not the most usable tools in the desktop videoconferencing category. Further, to equip even a small conference room requires high quality and expensive cameras, displays and sound equipment (e.g., echo canceling devices).

5.2.2.2 H.323 -- "Codec" video

The first widespread adoption of video conferencing equipment involved television set-top terminals that connected to each other over ISDN networks and communicated using protocols specified in the ITU H.320 umbrella standard. With the advent of reliable high speed packet-based networks, the ITU defined a successor standard for video conferencing known as H.323. The vast majority of commercially produced video conferencing equipment is now H.323 compatible. Conferencing endpoints based on this technology are prevalent across academia and industry and consist of both set-top stations and workstation-based equipment. Largely due to the widespread commercialization of this technology, products are available from vendors such as Polycom that are extremely easy to use and offer a rich set of conferencing features such as built in echo cancellation, support for multiple omni-directional microphones for quickly equipping a small conference room and built in, high quality cameras with Pan-Tilt-Zoom capability. H.323 is best suited to point-to-point video, but can be multipoint enabled through the use of bridges -- a feature fully supported by VRVS. These bridged conferences, however, usually only allow an H.323 site to view the video stream of the active speaker, though some bridges can show up to 9 sites at a time. These products are still rather expensive for equipping individual offices, but can be easily deployed in small conference rooms where their audio and video capture and display capabilities are best suited. Separate workstations for data collaboration and VRVS initiation are needed to complete these conferencing environments.

5.2.2.3 Access Grid

While not simply a video conferencing environment, the Access Grid is heavily used for this purpose. The Access Grid is an immersive environment to support group-to-group collaboration by providing tools for multipoint video conferencing, a suite of presentation and interaction applications, interfaces to Grid middleware and interfaces to visualization environments coupled with customized systems for capturing and displaying audio and video. The video conferencing tools incorporated in the access grid are based on *vic* and *rat*. Unlike the workstation-based VRVS deployment, Access Grid installations typically have more significant hardware resources dedicated to encoding and decoding media streams making it more feasible to support multiple video streams and more advanced collaborative tools (e.g., shared slide presentations). Aside from being rich in computational resources, Access Grid nodes also feature custom audio subsystems to capture, balance and echo cancel both speaker and ambient conversation using a number of microphones. High in cost in terms of installation and the personnel needed to run an installation, the Access Grid can be of tremendous value in locations where a sizeable local group needs to collaborate with a number of distant groups or individuals. CERN, for instance, could dramatically improve the end-user video conferencing experience by equipping heavily used conference rooms and auditoriums as Access Grid nodes. Deploying the Access Grid in these locations would make it easier to simultaneously view and interact with multiple distant sites and provide much cleaner audio and video to those distant participants.

5.3 AUTHENTICATED QOS

5.3.1 *Introduction*

Many research tasks of current and future interest require the transfer of large amounts of data across computer networks in a timely fashion. Today's Internet provides only a "best effort" data delivery service where all network traffic receives the same priority, whether it contains time-sensitive research data or is a MP3 file. When network traffic exceeds the capacity of some network segments to carry all of the traffic, the network responds to the congestion by dropping some packets, which must then be re-sent with corresponding delay. Well-behaved network applications are expected to slow their rate of data transmission when they re-send packets, and thus adapt to the network resources that are available. This allows large numbers of users and applications to share the network, but the speed of delivery and the amount of network bandwidth available to a specific user and application varies, which can seriously compromise some applications -- such as network-based interactive video or tightly coordinated distributed computations. The variability in network performance is typically beyond the control of any single user or application.

5.3.2 Current Situation

Two approaches are being pursued to overcome the limitations of "best effort" delivery. The first and currently most common approach, termed "over provisioning," is to build specialized portions of the network that have controlled access and are capable of very high performance. In this approach the expectation is that over-provisioned portions of the network will result in little or no network congestion, and hence no need to re-send packets at slower rates. Another approach, termed "Quality of Service (QoS)," is to implement different levels of network service that may be requested by an application, so that important or time-sensitive traffic is given preference over other traffic, much like first-class mail is given preference over bulk mail. In this approach, packets are marked to indicate the service level they require and capable networks give the packets different priorities. There have been very successful tests illustrating the power of QoS and automated bandwidth brokering in improving videoconferencing quality between the IT Division at CERN and Ann Arbor, Michigan in Spring, 2002.

5.3.3 Needed research

These two approaches, over provisioning and QoS, are not mutually exclusive. In fact it seems likely that significant progress will require that both approaches be pursued. A particular example of the need for alternatives to "best effort" delivery is in ATLAS. For example, ATLAS participants need guaranteed high-bandwidth and low latency for interactive collaboration (e.g., high quality video-conferencing) and for testing and implementing distributed grid computing. As the ATLAS detector becomes operational it will generate, on average, several terabytes of data per day (on the order of petabytes per year), and distributed grid computing is seen as the best way that meaningful physics can be

extracted by dispersed researchers from such a large volume of data. ATLAS researchers are beginning initial work to implement both QoS and distributed grid computing. Although the idea behind QoS is simple, implementing it in practice is a focus for further research. The research challenges include: a) creation of an authenticated mechanism to create and release bandwidth allocations in a dynamic environment; b) gaining control over intermediate network segments, such as commercial commodity networks, which may not support QoS; and c) incomplete understanding of the effect of adding QoS flows to existing network traffic.

5.4 LECTURE CAPTURE

5.4.1 *Introduction*

"Lecture capture" is the concept of recording a lecture or presentation for later viewing on the web. This is accomplished by blending relatively low-bandwidth video (i.e., of the speaker or instructor) with high quality still images (i.e., of the slides), which can be viewed successfully even over a dial-up connection. This is of interest to ATLAS because of the need to archive and convey detailed information within and among groups in the Collaboration on an asynchronous basis. Using this vehicle, even complex issues initially only understood by "experts" can be "taught" to colleagues who are either newcomers to the Collaboration or simply not yet experts.

The field of lecture capture has matured from the mid-1990s, where it was an activity conducted by professionals using various multimedia applications, to the present, where popular commercial applications (e.g., Microsoft PowerPoint) allow end users to create narrated Web presentations. Despite the rapid evolution over the past decade, development of lecture capture capabilities has ignored two critical areas. First, applications remain difficult in terms of usability. Second, it is important to allow captured materials to be reused and repurposed in different formats over time. Work relevant to these two areas has been done within a number of projects, notably the Web Lecture Archive Project (WLAP) (Bousdira, et al., n.d.). For example, WLAP uses a Windows application called Sync-O-Matic, developed at the University of Michigan, to collect together PowerPoint slides, video and audio streams, and timing information for slide transitions to produce a synchronized presentation that can be played back from any web browser with the RealPlayer plug-in. The WLAP archive now consists of over 300 on-line lectures, including CERN summer student lectures, ATLAS plenary meetings, and ATLAS and LHC software training seminars

5.4.2 Current Situation

In terms of the WLAP project, a significant challenge is to find ways to reduce the effort required to process presentations and to manage the archive of stored presentations. For example, the Wilma database tool allows hierarchical arrangement and display of lectures and provides a search mechanism for lec-

ture content. A significant use of WLAP has been in support of ATLAS software training activities, such as for GEANT4 -- where presentations by GEANT4 developers before a live audience were recorded, including questions from the audience -- and similar presentations for CMT (the configuration management system adopted within ATLAS), and Athena (the framework for ATLAS analysis software).

In terms of commercial development, the release of PowerPoint XP and Microsoft Media version 9 join separate development trajectories. First, PowerPoint XP offers significant improvements, in terms of usability, for capturing presentation content and replaying this content. For example, using PowerPoint XP, users can edit presentations and add new material, and then republish the modified presentation. Second, Microsoft Media 9 provides VHS/DVD playback quality while using a highly compressed transmission/storage format. This eliminates the previous tradeoff where formats suitable for limited bandwidth (e.g., streaming formats) were low quality -- while higher quality formats (e.g., AVI or QuickTime) required more bandwidth.

5.4.3 Needed research

As commercial offerings for presentation capture, such as those from Microsoft, approach (and in some areas eclipse) the capabilities of non-commercial research software, there remain a number of areas for additional research: a) creation of mechanisms to publish a captured presentation in a variety of formats suitable for a wide range of connection speeds, viewing applications, and devices; b) creation of standardized archival formats that would allow exchange of content across different systems (Vitaglione, et al., 2000); c) addition of new channels of information to captured presentations, such as voice or graphical annotations; and d) exploration of methods for capturing content from other presentation formats (e.g., PS or PDF) and from other presentation systems (e.g., PolyCom H.323 videoconferencing stations).

6 THE GRID AND COLLABORATIVE TOOLS

The challenge to physicists presented by the ATLAS experiment are enormous. With a huge number of physicists globally distributed and with data sets starting at the Petabyte range and growing to Exabyte scales by ~2012, we are forced to look for new ways to accomplish the articulated physics goals. The problem is compounded by the complexity of the data and the analysis software which must be shared by multiple, dynamic working groups within the Collaboration.

Fortunately, technology has also made significant strides in network bandwidth and ubiquity, computer processing power and storage density that have given rise to the grid computing model. The central concept of grid computing, that of harnessing a collaboration's globally distributed resources, wherever they may exist, has great appeal and promises to efficiently and effectively utilize all of a collaboration's computing assets. We are in the midst of a revolution in computing with middleware and standards being developed and deployed to enable higher level applications to deliver a grid computing environment for physics.

Unfortunately, while this developing grid computing environment will deliver unprecedented access to computational cycles and massive datasets it will create a host of new issues which must be addressed to actually deliver efficient, effective use of a collaboration's resources. Assuming such a physics grid will deliver transparent access to the collaboration's resources, one must immediately confront the problems the success has wrought: which working groups or individuals are allowed to use what fraction of the Collaboration's resources? How will the Collaboration decide such issues? What about the myriad of topical working groups within the Collaboration; how will they coordinate efforts to efficiently undertake their physics analyses? How will the Collaboration verify the accuracy and reliability of potential discoveries made in smaller working groups? A truly effective grid environment requires much more than transparent access to resources; it requires transparent access to collaborators.

Partly because the Grid will make possible the rapid processing of computational tasks, the ability of key experts on a given topic to quickly consult will become an increasingly important limitation on the production of scientific results. In a Grid-enable physics environment collaborators, though isolated from each other, will want to be able to view just-generated tabular and graphical results while talking. They will want to share files and work space like never before. They will need ready access to calibration and versioning information at a level that exceeds what has been anticipated. With the broad dispersion of scientists, many of whom may be pursuing similar research paths, the results of prior work should be readily available to avoid needless duplication of effort. There should be a robust scheme for reporting application bugs. A single collaborator actively involved in several related analyses will have difficulty keeping up with the work of the respective working groups unless there is a reliable and robust web capture archiving system in place. Chat rooms will be vital in debating anomalous results -- especially during periods where new effects may be seen at a very high rate, given the large number of collaborators and the exciting new physics discovery opportunities.

The challenge for ATLAS, and any other large distributed collaboration needing grids, is to insure that collaborative capabilities are a fundamental part of all grid components, from their inception.

7 OUTREACH AND EDUCATION

There is worldwide an intense interest in seeing that frontier research facilities like the LHC contribute to enhancing the educational experiences of children and improving the general public's understanding of science. Indeed, many funding agencies insist that every major research undertaking include components of education and outreach.

High energy physics has always held a special role in this area. The public is very interested in what we do and in the scientific puzzles we try to solve. The enormous number of visitors at CERN in the course of a year is a manifestation of this interest, with some guests having to wait for weeks to get to go on a tour of the Laboratory.

Technology advances, of the very same nature that facilitates the work of our LHC scientists, can have a major impact on our ability to share with people all over the world the excitement and wonder of doing science.

Students in a third-world country having only access to an old computer with an old modem and a free browser and free RealPlayer can listen to lectures on ATLAS given by the ATLAS Physics Coordinator to students at CERN. If we choose, an adult with an abiding interest in science can watch events displayed in our control room, again with no incremental investment in computer equipment.

For selected audiences, we could provide the true experience of what it is like to be in the control room, through total immersion technology.

It is not the purpose of this paper to delineate all of the ways that collaborative tools can impact outreach and education. We note that each of the LHC experiments has an active Education and Outreach Committee, and that they are quite aware of the importance of such activities to CERN, to our experiments, and to the many nations that support the operations at the Laboratory. We do, however, urge these committees to keep in close touch with advances being made in collaborative tool deployment in ATLAS for new idea about how to regularly update their education and outreach functions.

8 RECOMMENDATIONS

8.1 INTRODUCTION

The discussion above has described the evolution of collaborative tools, and pointed to some of the frontiers that may be of relevance to ATLAS in the future. A key question for ATLAS is what, if any, actions should be taken at this time.

Relevant to this question, there is a new era emerging where the importance of updating HEP collaborative tools is being recognized in several communities. Projects with collaborative components in which CERN is involved have been funded through the EU. Work is underway in the U.S. directed toward the submission of a NSF Information Technology Research (ITR) proposal to support U.S. CMS and U.S. ATLAS collaborative environment needs. This is precisely the kind of activity that some of the authors of this note have been calling for over the past year and we are very pleased to note the current steps. But we also recognize the need for such efforts to be well coordinated with CERN.

With the material in the above sections as background, and in the context of the various new initiatives that are underway, we respond to the request of the ATLAS Management with the following recommendations:

8.2 CREATE A FOCAL POINT FOR COLLABORATIVE ACTIVITIES

We recommend that an entity be created in ATLAS to support, improve and extend current collaborative services, and to serve as a facilitating center for remote institute activities oriented toward support of the ATLAS experiment.

It will not be possible for ATLAS to take advantage of, or to leverage the contributions of, outside groups wishing to help the experiment expand its collaborative tools if there is no one to contact at CERN, even for the most rudimentary purposes. Moreover, a small amount of support at CERN could make significant differences in the quality of various collaborative services (e.g., better videoconferencing). Also, we note that there is no one in ATLAS with responsibility for training.

We recognize that budget considerations would necessarily limit how much could be done in this area. But we also recognize that the current total lack of such support will be an enormous obstacle to progress. We urge the examination of possible funding opportunities through national or super-national sources.

We have provided in Appendix II a more detailed analysis of a proposed ATLAS Communication and Training Program and a Collaborative Tool Service.

Embedded in this analysis is also a suggestion that several of our meeting rooms in Building 40 be equipped at least to minimal standards. We offer this recommendation as an example of what could be done immediately. As noted at the outset, such a recommendation should be vetted through an appointed ATLAS group reporting to ATLAS Management.

8.3 IDENTIFY AND REPAIR HEAVILY USED AND UNDEREQUIPPED COLLABORATIVE FACILITIES AT KEY SITES

Looking at day to day work at CERN and other key sites there will be facilities that are heavily used, but improperly equipped. An example of this are large conference rooms at CERN that are used for video conferencing, but do not have adequate microphone, speaker or camera systems installed. Dramatic gains can be made in the quality of conferencing as perceived by remote users by upgrading the audio (microphones, mixers, echo cancellation devices, etc.) and camera systems.

8.4 CONDUCT A PROFESSIONAL ASSESSMENT OF ATLAS COLLABORATIVE NEEDS

We recommend that a professional assessment should be made of the detailed needs for collaborative tools in ATLAS. The scope of LHC experiments is such that the analysis of how the work of scientists might best be facilitated is worthy of an examination by experts.

This would require one person approximately 6 months to do the actual data collection, including interviews and sitting in on meetings at CERN and abroad. An additional 6 months would be required to develop a suggested action plan for presentation to the ATLAS management.

8.5 INITIATE SELECTED DEMONSTRATION PROJECTS IMMEDIATELY

We urge the initiation of a carefully selected set of demonstration projects immediately. These would be designed to address particular needs, while also beginning to build the network of individuals across the Collaboration (and outside) who may ultimately be required to implement larger projects. An example might be calling upon entities like NCSA or Internet2 to help install AccessGrid facilities for key sites of one or two of the ATLAS working groups.

8.6 ASSESS THE IMPLICATIONS OF THE GRID FOR COLLABORATIVE TOOL NEEDS

We recommend that an appropriate working group be assigned the task of specifying what collaborative tools are needed to make the Grid physics analysis environment truly useful to members of the Collaboration, regardless of their physical location.

This question will undoubtedly be addressed at regional or national levels by one or more of the current initiatives underway. We stress, however, that for progress in this area to be of true benefit to ATLAS, a degree of coordination will be needed at the ATLAS level.

Furthermore, we note that the overall topic of how physics analysis is to be conducted in ATLAS is of paramount importance as one evaluates what tools are required to support the modalities adopted. It is time for such topics to be engaged by the Collaboration.

8.7 ASSESS DESIRABILITY OF A SINGLE PORTAL FOR ATLAS COLLABORATIVE TOOLS

We recommend that ATLAS commission a small study on the desirability of the development of a single framework and single portal for all ATLAS collaborative tools.

There are true advantages of being able to have e-mail, calendars, analysis groups, data files, Grid jobs, videoconferences, etc. share a common framework. But there are dangers of disturbing longstanding work habits. Some scientific collaborative communities are adopting the single portal approach, with some evident success. The recommendation above only suggests that this is a question that should be examined by ATLAS as to whether, for it, the benefits outweigh the risks.

8.8 ENCOURAGE QUEST OF COLLABORATIVE TOOLS FUNDING SUPPORT THROUGH INSTITUTES AND NATIONAL GROUPS

We recommend that the ATLAS Management take overt steps to encourage institutes and national groups of institutes to seek funding for collaborative tools through their local agencies, with coordination being provided by the ATLAS Computing Board or some other appropriate entity.

Underlying this recommendation is the fact that while the importance of funding support for specialized instruments and computing is internationally accepted, the need for support to make the collaboration among scientists more efficient is not. The leadership of ATLAS management could be important in making the point that the success of the experiment also depends critically on the ability of the Collaboration members to easily interact.

Many more recommendations could be derived from the material presented above. But we believe that a detailed action plan should be developed by an appropriate ATLAS Working Group.

9 SUMMARY

We hope that this report will assist members of the Collaboration as they reflect on how they wish to organize themselves for collaboration as they approach the commissioning and datataking and analysis phases of the experiment, and how to decide on what frameworks, tools and services will be needed to facilitate the vision they embrace.

We believe the magnitude of the scientific challenge facing the LHC experiments, the number of the actively participating scientists, and the global distribution of these scientists, calls for a renewed attention on just how to most effectively facilitate the human-to-human and human-to-data interactions required to optimally exploit the enormous investments made in these experiments. Moreover, the overlay of the new computational schema of the Grid, will present new challenges and opportunities for enhancing the richness of collaboration and for accelerating the pace of discovery.

We suggest that this is an appropriate item to be considered by ATLAS Management for the health of the Collaboration and that now is an appropriate time to initiate action.

APPENDIX I - SURVEY RESULTS

The following people participated in the ATLAS Collaborative Tools Survey

- Leonardo Rossi, INFN - Genova
- G. Mikenberg, Weizmann Institute
- Kenneth W. McFarlane, Hampton University
- Bob Orr, University of Toronto
- Maris Abolins, Michigan State University
- John Renner Hansen, Niels Bohr Institute
- J. Ernwein, Dapnia-Saclay
- Dave Charlton, The University of Birmingham
- Kaushik De, Univ. of Texas at Arlington
- Schwemling, LPNHE-Paris
- Claus Goessling, Universitaet Dortmund
- Siggi Bethke MPI of Physics, Munich
- Tony Doyle, University of Glasgow
- Bengt Lund-Jensen, KTH (Royal In stitute of Technology) Stockholm
- Andreu Pacheco, IFAE Barcelona
- Martin Sevior, University of Melbourne
- Roger Jones, Lancaster
- Zhongliang Ren, Inst. of Physics, Academia Sinica, Taiwan
- John Huth, Harvard
- T.Davidek and M.Lokajicek, Inst.of Particle and Nuclear Physics, Charles Univ. + Inst.of Physics, AS CR
- Andrey Minaenko, IHEP (Protvino) Institute for High Energy Physics
- Piotr Malecki, Institute of Nuclear Physics, Krakow, Poland
- Lamberto Luminari, INFN - Roma
- Farid Ould-Saada Department of Physics (particle physics), University of Oslo

4. Number of Ph.D. scientists in group affiliated with ATLAS: (24 of 24 answered) average: 9.29 min: 3, max: 24 [12, 7, 3, 8, 4, 7, 17, 14, 9, 4, 3, 19, 8, 3, 6, 4, 8, 9, 6, 13, 23, 8, 24, 4]

5. Total number of graduate students in group affiliated with ATLAS: (24 of 24 answered) average: 2.79 min: 0, max: 9 [3, 4, 1, 1, 0, 2, 0, 3, 6, 2, 5, 1, 1, 1, 5, 2, 1, 0, 2, 6, 6, 9, 4, 2]

6. What videoconferencing equipment is available at your institute? (23 of 24 answered)

- Pleiadi Twin (Aetra), based on ISDN and Codec
- ISDN U.S. DOE DCS-ESNET compatible -- Picture Live200.
- ISDN video40

- Polycom System
  Picturetel System
- VCON - 2 x 64 ISDN
- CODEC H320 and H323
  Multi-connexion 1 + 3
- HEP group: VRVS on one PC; H323 on another PC; H320/H323 system being installed soon.
  University facility: H320 system
- ISDN based DCS room, few VRVS capable workstations
- Viewstation 128 Polyspan
- videoconference rooms in a central university-area (ISDN)
  some point-to point videoconferencing from individual PCs
- Codec
  VR
- VRVS, H323, MCU
- VRVS (with very limited use of camera)
- Picturetel 200
- ISDN - based video conference.
  gnome-meeting and webcams.
- Vidacom box, H323 connection. Some H323 on desktops
- both web VRVS and ISDN
- Polycomm, VRVS
- Nearly all personal computers (Linux and Windows) have speakers,many have microphone, some have web camera, digital Sony videocamera, 3 wireless microphones with mixer, camera for documents(visualiser), 2 PC projectors
- Individual PCs with speakers and microphones at the moment
- About 15 PC computers equipped with digital cameras and audio tools.
- ISDN
  VRVS nothing (we use equipment at IT nearby)

7. What is the approximate number of person-trips to CERN per year from your institute primarily to attend a 1 day meeting? (24 of 24 answered)

average: 6.45   min: 0, max: 30

[20, 0, 0, 5, 0, 12, 5, 3, 0, 20, 3, 16, 5, 1, 30, 0, 2, 8, 2, 4, 0, 3, 10, 6]

8. What is the approximate number of person-trips to CERN each year from your institute of duration less than 1 week? (24 of 24 answered) average: 27.77 min: 3, max: 175 [40, 30, 5, 5, 10, 25, 105, 10, 3, 25, 30, 175, 20, 10, 25, 4, 20, 8, 6, 13, 3.5, 20, 70, 4]

9. What is the approximate number of person-trips to CERN each year from your institute of duration longer than 1 week? (24 of 24 answered)

average: 15.62 min: 0, max: 79

[20, 5, 1, 30, 2, 10, 15, 5, 3, 4, 5, 79, 20, 2, 6, 6, 3, 3, 6, 60, 35, 5, 50, 0]

10. What is the typical number of staff from your institute at CERN at any given time? (24 of 24 answered)

average: 2.25 min: 0, max: 12

[2, 2, 0, 0, 1, 1, 12, 1.5, 0, 0.5, 0.5, 5, 2, 0, 4, 1, 2, 2, 0.2, 3, 7.5, 3, 3, 1]

11. Which ATLAS Working Groups is your institute affiliated with? (24 of 24 answered)

- IDWG, EB, IDSG, Pixel working meetings, TMB, Trigger working groups, Software working groups.
- MUON, LV1, Higgs, SUSY
- TRT
  Grid
- Liquid Argon
  Software
  top physics
- Trigger/DAQ
- TRT, TDAQ and GRID
- Liquid argon calorimeter
  Muon spectrometer
  Software
- Trigger/DAQ
  SCT
  Software
  PESA
- TileCal, Computing
- Liquid Argon, Exotica, QCD, Electron-gamma
- ATLAS-pixel
- SCT, MDT, HEC
- Inner Detector,
  Higgs,
  Software
- LARG, SUSY physics, NCB
- Physics and Selection Software Architecture
- SCT, software
- B physics
  Offline, Grid, Data Challenges
- optical links for several sub-detectors
  Offline Data Challenges
  Atlas software and infrastructure tasks

- Muon, Construction, Electronics, Computing
- TILECAL, ID (Strips and Pixels) , Outreach, Physics groups, DC
- HEC, TILE calorimeter, Muon system, Inner detector
- SCT, TRT and T/DAQ
- - TDAQ (LVL1, HLT, DAQ, DIG, Testbeam, ...)
  - Muon Spectrometer (MDT, Simul. & Reco, Prod. DB,Testbeam,...)
  - Computing (Data Challenges, Muon softw., NCB, ...)
- SCT (hardware), software data challenges, Exotic group

12. What is a rough estimate of the financial investment your group is making in collaborative tools each year (in CHF)? (22 of 24 answered)

average: 3586.36 min: 0, max: 30000

[30000, 1000, 0, 1000, 3000, 0, 1000, 5000, 0, 0, 10000, 2000, 500, 1000, 1000, 2000, 5000, 5000, 0, 400, 1000, 10000]

13. How many FTE's has your group devoted to support collaborative technology? (23 of 24 answered)

average: 0.15 min: 0, max: 1

[0.5, 0, 0, 0, 0, 0, 0.05, 0, 0, 0, 0.5, 0.2, 0.005, 0.5, 0.01, 0.2, 1, 0.1, 0.1, 0, 0.3, 0.2, 0]

14. How many ATLAS Working Group Meetings do one or more group members participate in by video conferencing each year? (24 of 24 answered)

average: 19.79 min: 0, max: 70

[12, 5, 70, 4, 20, 5, 30, 20, 2, 0, 1, 40, 50, 6, 10, 0, 10, 52, 12, 30, 35, 10, 50, 1]

15. Please estimate the reduction in travel expenses for trips to CERN (in CHF) that might be possible for your group if high quality, reliable, on-demand videoconferencing capabilities existed. (19 of 24 answered)

average: 21000 min: 0, max: 100000

[20000, 30000, 6000, 10000, 15000, 0, 0, 20000, 50000, 2000, 100000, 5000, 50000, 0, 0, 26000, 5000, 10000, 50000]

16. Please estimate how many additional members of your group might participate in key meetings by videoconference each year if high quality, reliable, on-demand videoconferencing capabilities existed. (Please express answer in terms of person-meetings per year). (23 of 24 answered)

average: 41.26 min: 0, max: 400

[30, 10, 1, 20, 1, 0, 15, 10, 10, 5, 50, 50, 10, 10, 10, 40, 3, 50, 400, 100, 20, 100, 4]

17a. Please estimate how much your group uses VRVS to participate in any of the ATLAS meetings (in person-meetings per year): (21 of 24 answered) average: 14 min: 0, max: 50 [10, 5, 1, 1, 0, 0, 20, 10, 0, 0, 10, 50, 6, 50, 12, 4, 40, 35, 10, 30, 0]

17b. Is it helpful? (18 of 24 answered)

yes: 61 % (11 / 18) no: 38 % (7 / 18)

18a. Roughly what percent of the time do you have a satisfactory video session with VRVS? (17 of 24 answered) average: 32.94 min: 0, max: 85

[10, 60, 0, 0.0, 65, 30, 50, 50, 50, 30, 10, 10, 30, 85, 50, 10, 20]

18b. What improvements do you wish to see made? (14 of 24 answered)

- drop it or make it work.
- More bandwidth; dedicated bandwidth.
- Reliability (particularly sound problems some months ago - have not used it recently because migrated to H320 or phone, some others report it has got better)
- Audio reliability, routing through firewall
- better band width, improved video and audio quality, professional management of camera at CERN,
- Reliability
  Web pages prepared in advance
  Notes on what is expected for participants
- Sound quality (often impossible to hear what is said during the meeting)
  Switch camera to people during discussions when transparacies are not really used.
- Offline access to meeting recordings
- A properly equipped v/c suite with microphones that capture the ambient comments.
  Use of standard protocols for meetings
  (H323). The ability to make standard non-H323 meetings from CERN V/C rooms.
  More use of exclusively telephone meetings (*ALL* participants) Effective unification of the VRVS sound with a telephone dial-in
- ease of software download,
  ease of use (Plug and Play),
  quality of service
- improve quality of sound, somtimes the equipment is good but is not correctly used, stability
- General improvement of audio quality.
  Quality of VRVS service is very poor when number of active users is larger than 5
- Better average quality of audio and video, greater homogeneity of connections (quality of images, noise and volume vary too much), longer lifetime of the sessions.

- higher quality,
  better availability (more facilities in the building),
  don't want to have to request a time slot

18c. Is there a member of your group who has a good familiarity with this application? (22 of 24 answered)

yes: 63 % (14 / 22) no: 36 % (8 / 22)

19a. Will members of your institute require any type of GEANT4 training within the next two years? (23 of 24 answered)

yes: 78 % (18 / 23) no: 21 % (5 / 23)

19b. Has the existing training web archive been used for this purpose? (20 of 24 answered)

yes: 20 % (4 / 20) no: 80 % (16 / 20)

19c. Was it helpful? (5 of 24 answered)

yes: 80 % (4 / 5) no: 20 % (1 / 5)

20a. Do you believe building 40 at CERN is adequately equipped with video conferencing facilities? (19 of 24 answered)

yes: 31 % (6 / 19) no: 68 % (13 / 19)

20b. If not, do you have any suggestions? (14 of 24 answered)

- equip more rooms.
  have a support person available 'on call' at all times.
- Although I say 'yes' above, this is for conventional VC rooms. IF there were a way of easily doing desktop VC with most ATLAS members, that would be helpful (as easy as a phone call or email).
- audio quality sent from CERN needs to be improved;
- Equip all basement rooms and some fraction of other rooms. Wireless LAN and Polyspan VIA-Video = Universal
- VRVS must also be possible for small meetings.
- Better standard phone conferencing integration support with the videoconference.
- Sorry I don't know what video conferencing facilities building 40 has. We've used video conferecing at Belle very,successfuly for 4 years. It has helped a lot for that experiment.
- Wiring 40RD10 properly for sound. Replace the poorly-working microphones in the 3rd floor conference room

- clarification: the existing rooms are adequately equipped, but the demand for meetings exceeds the number of rooms.
- It is sufficient for receiving videoconference, broadcasting is probably sufficiently equipped but people do not know how to use it well or correct usage is too demanding
- increase the number
- First, number of rooms properly equipped for VRVS should increase.
- More small equipped rooms are needed
- it often takes 10-20 minutes to get a session started,
  it's hard to hear everyone

21. Would your institute be willing to pay a small annual fee of the order of 1000 CHF to support a web archiving/ videoconferencing effort within ATLAS if the reliability and quality factors could be significantly improved? (21 of 24 answered)

yes: 57 % (12 / 21) no: 42 % (9 / 21)

22. At present, what are the factors you believe limit the effectiveness of videoconferencing in support of your ATLAS activities? (22 of 24 answered)

- technology is poor. It is hard to take part in a videoconf.
  as often the background noise is higher than the voices (this
  is expecially true when an individual is connected with a group meting
  at Cern and also depends on people 'education' to videoconf.)
- Resolution -- not easy to see any details in figures.

  No universal means of putting figures on screen.
  In our system, we use an auxiliary camera, or a laptop with video output. A VC system running on a computer that could take screen graphics and send them out would be useful.

  Need simple education of VC participants in VC use (the voice-activated priority system sometimes leads to shouting matches -- need protocols, just like a formal meeting.)

  Inability to do a VC without pre-arrangement of rooms and equipment.
- People have not developed the necessary 'skills' for video conferencing
- Contributions from the audience are often inaudible. Distributed microphones would help.
  Resolution limits usefulness of overheads. A factor of 4 improvement (e.g. 4x0 kbs) would be a significant improvement.
- 1) Poor technical quality (cf; recent meetings of the National Contact Physicists which I attended)
  2) Most ATLAS coordinators (in the systems we are involved) call meetings to take place at CERN

- Reliability is inadequate. Quality can be quite good when it works. Ethos has to change to make ALL slides available electronically before ALL meetings. Some question if reliability good enough when many sites involved.
- quality
- The availibility of the local equipment is a problem, due to intensive use by USA-based experiments (D0/Babar)
- lack of usage
- limited video resolution and quality of audio;
  missing professionality of operating cameras in meeting room;
  lack of distributing copies of documents diplayed at the meeting beforehand...
- Reliability - Repeated breaks in continuity.
  Difficult to break in from a remote site.
- Availability of VRVS in small meeting rooms at CERN and
  the poor sound quality (microphones?)
- The quality of the network
- Lack of knowlege on our part.
- Very poor sound, restriction to VRVS; the rest of the world (including most HEP) use standard H323 via gatekeepers and most conference suites do not have (and are unwilling to install) VRVS
- limited bandwidths, low video quality, hard to see the projected contents of transparencies shown.
- availability and ease of use, not enough virtual rooms available
- low quality, low stability, high demand for complicated equipment
- network bandwidth
- 1. quality of VRVS
  2. number of rooms with VRVS available during meetings
  3. human reasons.
- The inadequacy of the present tools (except ISDN)
- quality/availability (see question 18b)

23. Would you use more desktop-desktop videoconferencing if the implementation of the technology were improved? (23 of 24 answered)

yes: 100 % (23 / 23) no: 0 % (0 / 23)

24. Please note any other issues involving collaboratory tools you would wish to bring to our attention. (14 of 24 answered)

- Does 'collaboratory tools' include web-enabled system like LabView? Database? PAW? If not, why not? Seems to me we often discuss performance of apparatus and data as part of our collaboration.
  Voice with high-resolution data would be useful, and perhaps more so than
- Sound quality needs improvement.
- It is not possible to make guesses about how 'high quality video conferencing' would affect the effectiveness of our work.
  CODEC conferences have proven useful when they were well focused and prepared

in advance with a limited agenda and written material available (over the net e.g.). This has happened in the trigger/DAQ system, in particular with US labs..

On questions 12 & 13: our CODEC equipment at Saclay is being installed and maintained by our local technical software support. The ATLAS group is not being charged for this service.
This should be true also at CERN who should act as a host lab in this instance.

On question 14: 30 video conferences in 1999 for our trigger/DAQ activities. None since.

On question 15: This cannot be guessed. It depends on the CONTENT of the meetings. One must try with good technical installations and a meeting organization which takes into account its nature (video conference).

- General feeling here is that a good, reliable, voice connection, and rigorous circulation of slides in advance, are the keys. When working, VC is as good (or better) for this as phone conferencing, which is still generally more reliable. For the most part being able to see people is rather unimportant.

- ESNET used at Fermilab is of far higher quality to Glasgow. Use of VIA-Video to be encouraged (cost approx. 1000 CHF).

Question 15 comment:
Savings already made, improvements in quality real (VRVS -> H323)

Question 19c. comment:
Lecture access.

- By now, the preferred method by physicists to attend remote meetings is the standard phone conference. The reason is the quality of the sound in VRVS is not good enough.

- Please don't put unreasonable demands on the facilities required for desktop video conferencing. We run Linux workstations almost exclusively here. gnome-meeting works very well to Belle (after Belle gate-keeper issues were fixed.)

- The telephone is underused. Newsgroups are not used enough in ATLAS (although the technology was set-up on request) and mailing lists are overused, with the result the signal is often lost in noise.

- Question 15: Hard to estimate, it might be offset by the higher telephone bills if ISDN (which uses reserved bankwidth and has a better quality) video conference has to be used

Question 17a: many

- Tools should be easy to use and stable.
  It would be nice to use collaborative tools on my laptop so that information is at the fingertips.
  Most conferences are teleconferences.
  One should be realistic: VRVS can help with a certain class of meetings, but not with others (protracted personal interaction, hallway conversations).

- telephone conferences (even over ip - VOIP) have much better quality and stability and are used often, have simple equipment with simple usage

- Question 20a: some meetings don't get translated because of the lack of equipped rooms
- Phone-conferences show very useful (the audio technology seems adequate and presentations published on the web can be used if needed). Equipments should be available on demand.
- Our group does not have adequate facilities for videoconferencing, but we are interested in changing this.
- Question 15: it is difficult to give a monetary value, but we estimate a reduction in travel expenses by a factor of 2.

APPENDIX II - DETAILS OF THE RECOMMENDED ATLAS COMMUNICATION AND TRAINING PROGRAM AND THE ATLAS COLLABORATIVE TOOL SERVICE

II.1 INTRODUCTION

The results of the ATLAS survey on collaborative tools, summarized in Section 4, have clearly identified an increasing need within the Collaboration for improvements to the availability and efficiency of its communication infrastructure, between and within working groups, member institutes, and CERN. There is a growing desire within the Collaboration for the effective training of its current and future members in the complexities of the experiment, and an increasing urgency of this issue as the start-up date approaches. In Section 3, we have presented a summary of the growing abundance of the rapidly improving (albeit ever-changing) tools available to address these needs. Yet, no entity currently exists within ATLAS or CERN to oversee the development, deployment, and integration of these technologies to the specific needs of the Collaboration.

To address the major organizational issues, we recommend the creation of an ATLAS Communication and Training Program, and we provide, as an exemplar, the skeleton of such a program. Its purpose would be to propose, develop, and coordinate communication and training projects within the Collaboration. In support of this program and other programs, working groups, and projects dependent on the communication infrastructure, we recommend the creation of an ATLAS Collaborative Tool Service. Such a service would be responsible for developing, deploying, and maintaining tools to be used by the Collaboration, with a focus primarily on identifying technical issues and finding solutions.

We believe that, ultimately, much more will need to be done to effectively transform ATLAS's capabilities in the area of collaborative tools. We are encouraged by national efforts (e.g., in the EU and US). But we also note that a common pillar in all of the bridges tying institutes to ATLAS is CERN, and that CERN will need to undertake certain steps, examples of which are outlined here, to insure the success of these other leveraging opportunities.

II.2 THE ATLAS COMMUNICATION AND TRAINING PROGRAM

We recommend that a Communication and Training Director be selected to oversee the planning and organization of the program. The goal of the program would be to facilitate communication within the experiment, through improved coordination of the existing infrastructure, and the planning, development, and maintenance of new projects and facilities. The director would be responsible for an annual budget covering the expenses of the program. We will not attempt

to specify the exact nature and source of the funding for the program here, but will provide an outline of estimated costs for potential projects.

The communication aspects of the program would include the development and maintenance of facilities at CERN, and the provision of guidelines and limited technical support for facilities at the home institutes. Following an assessment of the existing facilities, the director would provide the short-term (approximately one-year) and long-term (approximately three-year) goals of the program and request the funding and other resources necessary to achieve the program. Specific projects would probably include, but not be limited to, the equipping or upgrading and maintenance of video and phone conference rooms, desktop conference systems, lecture recording and archiving facilities, and web-based tools, such as white boards and application-sharing programs. The exact nature of the program will depend on the needs and priorities of the Collaboration, the technology and expertise available, and funding.

The training aspects of the program would focus on the organization of a comprehensive and coordinated training program for ATLAS. The program would complement the existing CERN technical and academic training programs by emphasizing topics specific to ATLAS, such as tutorials on the technical aspects of hardware construction or software development in the ATLAS framework. In addition, the program would not limit itself to organizing courses at CERN, but would also bring tutorials and lectures to the outside institutes and laboratories, either by conventional means or through the employment of collaborative tools.

It is not strictly necessary to combine the responsibilities of a communications officer and a training officer. However, we feel there is a significant overlap of scope of the two projects. In particular, the need for close coordination between activities at the home institutes and laboratories with those at CERN, and between the various ATLAS projects warrants the coupling of the projects, if only for the sake of efficiency. In addition, we expect both entities to benefit more and more from the deployment and usage of collaborative tools within ATLAS. The development of these tools would require feedback from communication and training, in order to best determine the overall needs and priorities of the Collaboration.

It would be natural for there to be a significant overlap of this program with that of outreach, especially concerning usage of communication technology. However, the goals and special requirements of outreach differ significantly from the intra-collaborative nature of the communication and training program described above, and we feel that close coordination between the projects, rather than a merging of the projects, is the appropriate course of action. In addition, we feel the breadth of scope of the projects would be too large for a single director to perform effectively, and would thus risk compromising their success.

II.3 THE ATLAS COLLABORATIVE TOOL SERVICE

The Collaborative Tool Service would provide technical support for the development, installation, and maintenance of collaborative tools in ATLAS. While the clientele of the service would include all working groups -- indeed all members -- of the Collaboration, its primary direction ought to come from the Communication and Training Program. We therefore recommend that the service be run by a Collaborative Tool Coordinator who is selected (or at least approved) by and who answers directly to the Communication and Training Officer.

The main focus of the Collaborative Tool Service would be on finding or developing solutions to technical problems related to communication within the Collaboration. This could include, for example, the equipping of videoconferencing facilities in the ATLAS meeting rooms and auditoria in a manner optimized for the needs of the users. Typical technical projects could include:

- digital transmission of projected transparencies;

- coupling of phone conference systems to videoconference audio systems;

- general improvements of sound quality;

- effective recording and archival of meeting presentations;

- increased automation; etc.

Anyone who has already attended a typical videoconference in one of the CERN meeting rooms would probably be happy to continue such a list. Indeed, one of the primary duties of the service would be to periodically survey the users of the systems to locate problems and, whenever possible, proposed solutions.

The Collaborative Tool Service would also be responsible for providing technical assistance and/or advice to local and remote users within the Collaboration. This could include the testing and documentation of standard hardware and software configurations for personal desktop systems, or the provision of expertise and assistance in the equipping of facilities at external institutes or laboratories. The responsibilities of the service would be limited to advice and a reasonable amount of hands-on service; it would not be responsible for the purchasing of equipment unless authorized by the Communication and Training Director.

The recording and archival of lectures presented at CERN and at the member institutes and laboratories will be an important responsibility of the Collaborative Tool Service. Under the guidance of the Communication and Training Director, the service will construct a recording facility, complete with equipment for recording lectures given publicly in meeting rooms, tutorial labs and auditoria, and privately either in a user's office or, ideally, in a modest recording studio at CERN. A small recording facility would aid in the development of training tutorials, and would allow for the recording and transmission of student-teacher mentoring sessions during the course of the scholastic year. It is expected that these

services would be available on demand, through a booking service such as CRBS (Conference Room Booking System) [ref to CRBS]. Special services might require modest fees, in order to reimburse the costs of manpower and equipment. This ought to be taken into consideration when determining the cost of the overall proposal.

Finally, it is important that the Collaborative Tool Service keep in touch with the state-of-the-art in collaborative tool technology, not only for the effective development of tools for the Collaboration, but also for the maintenance of the existing facilities. As the field is in a state of rapid development, there are frequent changes to the requirements of the tools, including their adaptation to new standards. In addition, certain tools have been developed for environments which differ than those of large collaborations, involving global participation. For example, the usage of proprietary software and hardware solutions may adapt well to the economic climate of a single business, but is often unacceptable for academic institutes, with limited funding for maintenance and licensing. As such, it is expected that the service spend a significant amount of time performing the research and development necessary to implement and maintain solutions optimized for ATLAS.

II.4 COORDINATION WITH THE LHC COLLABORATIONS AND CERN

Many of the issues raised above are common to the other collaborations of the LHC and to CERN, in general. While it is important for the ATLAS program to focus on the immediate concerns of the Collaboration, common technical solutions often help to reduce overall costs and benefit from the combined expertise of the groups involved. In addition, the installation of permanent video and audio infrastructure at CERN will require the laboratory's permission and technical coordination. It is therefore expected that the Communication and Training Director coordinate the ATLAS program closely with similar programs of the other LHC experiments and of CERN, and that common solutions be sought whenever overall benefits have been clearly defined.

II.5 RECOMMENDED PROJECTS

The following is a list of projects we feel merit immediate attention as potential projects to be coordinated by the Communication and Training Program:

1. evaluation and reparation of problems in existing ATLAS CERN phone and video-conferencing facilities;

2. equipping of the two ATLAS Auditoria in CERN Building 40 (Salle Curie and Salle Dirac) for phone and videoconferencing, and lecture recording;

3. equipping of room 40-R-C10 for phone and videoconferencing, and lecture recording;

4. evaluation of existing network status, investigation and implementation of improvements, as needed;

5. preparation of a Lecture Archival Facility to record ATLAS meetings, plenary sessions, tutorials, and teacher-student communications;

6. development of an ATLAS database of recorded lectures, with a web interface, including the ability to search lectures and the potential for mirroring at external institutes.

There are a significant number of other projects that ATLAS could benefit from. The ones listed above are considered a top priority. Table II.1 presents a breakdown of the estimated costs of these projects.

Table II.1 Breakdown of estimated costs for proposed communication and training projects

Project Equipment	Cost Per Unit	Number Units	Total Cost
Reparation of Existing Building 40 Video and Phone Conferencing Facilities	**10 kCHF**	**2**	**20 kCHF**
Equipping of Building 40 Auditoria for Video and Phone Conferencing	**40 kCHF**	**2**	**80 kCHF**
Equipping of 40-R-C10 Conference Room for Video and Phone Conferencing	**20 kCHF**	**1**	**20 kCHF**
Equipping of Videoconferencing Rooms for Web Lecture Recording	**10 kCF**	**5**	**50 kCHF**
Preparation of Lecture Archival Facility for Recording and Processing Lectures	**10 kCHF**	**1**	**10 kCHF**
Computing Facilities for Lecture Archival, Web Servers, Collaborative Tool and Network R&D	**20 kCHF**	**1**	**20 kCHF**
Total			**200 kCHF**

Note that manpower is not included in these estimates, only the cost of equipment. There is a certain flexibility in the estimates, based on the desired functionality and quality of the facilities. In our judgment, the costs are small when compared to the amount of savings to be gained due to decreased travel and improved communication, and quality should not be sacrificed at the risk of returning to today's unacceptable situation.

Manpower includes the Communication and Training Officer and the Collaborative Tool Coordinator, described above, as well as their staffs. The two coordination positions must each be at a minimum half-time. Most likely, they will need to be full-time at the beginning, but might be eventually scaled down to half-time,

depending on the success of the projects, the abilities of the coordinators and their availability. We expect the staffs to these positions be comprised of one or two full-time positions at the equivalent of a CERN technical student, as well as one or two temporary students (e.g. CERN Summer Student), as available and as needed.

APPENDIX III - EXCERPTS OF THE ITR PROPOSAL FROM US ATLAS AND US CMS FOR A PROGRAM OF R/D ON COLLABORATIVE TOOLS IN A GRID-ENVIRONMENT (FEB. 12, 2003)

Project Summary

ITR: A Global Grid-Enabled Collaboratory for Scientific Research

We propose to develop, prototype and deploy the *first Grid-Enabled Collaboratory for Scientific Research (GECSR)* on a global scale. A distinguishing feature of this proposal is the tight integration between the science of collaboratories, a globally scalable working environment built on the foundation of a powerful fully functional set of working collaborative tools, and an agent-based monitoring and decision-support system that will allow collaborating scientists to perform data intensive analysis tasks efficiently. Assessment of the methodology of scientific collaborations and the iterative evaluation of the tools by a team independent of the developers will be a critical element ensuring the success of the proposed work. The assessment will focus on the ability of the target communities (HENP, other sciences and eventually other fields of research) to collaborate one-on-one and in peer groups on a variety of scales across long distance networks; it will abstract requirements and barriers to effective communication and shared work; and it will aim to develop standardized, broadly applicable guidelines and tools for effective collaboration in a variety of working contexts. The major collaborations of experimental High Energy and Nuclear Physics (HENP), who face unprecedented challenges in accessing, processing, and sharing Petabyte-scale data, and who have successfully developed some of the most ambitious Grid systems in cooperation with leading computer scientists in the US as well as Europe, as well as developing scalable video-conferencing tools and the early components of a Grid-enabled Analysis Environment (GAE), will be the initial targeted early-adopter community. The recent report of the NSF Blue-Ribbon Advisor Panel on Cyberinfrastructure identifies five key service categories that will provide a foundation for the comprehensive knowledge environments that will enable individuals, teams and organizations to revolutionize scientific practice. The HENP community, working with computer science and communities in astrophysics and other disciplines, has addressed four of these services, based largely on common underlying middleware, in the widespread deployment of grid-enabled high perfor-

mance computing resources, through several data grid projects that facilitate management of data, information and knowledge, through instruments that are monitored remotely through eLogs and other interfaces, and through the developing GAE. The proposed Grid-Enabled Collaboratory for Scientific Research will provide the collaboration services element (the missing 5th element) of the cyberinfrastructure that is required for the HENP community to fully realize a functionally complete environment, with the potential to transform the conduct of research.

Broader Impact This proposal is distinguished by a tight integration between 1) the science of collaboratories, 2) a globally scalable working environment built on the foundation of a powerful fully functional set of working collaborative tools, 3) an agent-based monitoring and decision-support system that will allow collaborating scientists to perform data intensive analysis tasks efficiently, and 4) an education and outreach agenda that is inclusive of minorities, and reaches high-school physics programs through vehicles such as QuarkNet and success models such as ThinkQuest. Assessment of the methodology of scientific collaborations and the iterative evaluation of the tools by a team independent of the developers will be a critical element ensuring the success of the proposed work and insuring its value beyond HENP. The broader implications of this proposal, and its mission to grant full partnership to groups in all world regions are driven by (1) The right of groups and governments that have contributed to the global Collaborations to share and collaborate in the data analysis, and thus in the process of search and discovery that is at the root of all basic research. (2) The need to justify global projects, and global Collaborations as a valid and vital means of conducting future leading-edge scientific research and other endeavors, without exploitation of the poorer nations and world regions and (3 The need to develop effective means collaboration and cooperation among sub-groups with vastly different academic and cultural backgrounds, and especially to allow effective collaboration by university-based students in the native cultural milieu of their home countries. Seven percent of this medium ITR budget is allocated to Education and Outreach that specifically targets minority populations: this should produce a sea change in how high energy physics experimentation, and the necessary global collaboration, impacts the quality of science education and reaches to under-served populations across the US, and around the world.

BIBLIOGRAPHY

Atkins, D. E., G. M. Olson, et al. (2002). Symposium on Knowledge Environments for Science and Engineering: Past Present and Future website. Retrieved January 10, 2003 from http://www.scienceofcollaboratories.org/Workshops/WokshopNov252002/index.php.

Atkins, D. E., K. K. Droegemeier, et al. (2003). "Revolutionizing science and engineering through cyberinfrastructure: Report of the National Science Foundation blue-ribbon advisory panel on cyberinfrastructure" Retrieved January 20, 2003 from http://www.communitytechnology.org/nsf_ci_report/

Bellotti, V. & Dourish, P. (1997) Rant and RAVE: Experimental and Experiential Accounts of a Media Space. In Finn, Sellen & Wilbur (Eds.), Video-Mediated Communication (pp. 245-272). New Jersey: LEA.

Bousdira N., S. Goldfarb, E. Myers, H.A. Neal, C. Severance, M.Storr, & G.Vitaglione (n.d.) "WLAP: The Web Lecture Archive Project - The Development of a Web-Based Archive of Lectures, Tutorials, Meetings and Events at CERN and at the University of Michigan," retrieved from http://weblib.cern.ch/abstract?CERN-OPEN-2001-066

Dourish, P. & Bly, S. (1992) Portholes: Supporting Awareness in a Distributed Work Group. Paper presented at the ACM Conference on Human Factors in Computing Systems, Monterey, CA.

Edwards, J. (2001) "Don't Hang Up," Emerging Technology, CIO Magazine, 1 October, 2001 http://www.cio.com/archive/100101/et_article.html

Ehrlich, S. F. (1987). "Strategies for encouraging successful adoption of office communication systems." ACM Transactions on Office Information Systems **5**: 340-357.

Emery, G.R. (n.d.) "E-learning lessons learned, Standards reduce compatibility problems." Government Computer News 21(34), retrieved from http://www.gcn.com/21_34/tech-report/20627-1.html

ETAN (1999). Transforming European science through information and communication technologies: Challenges and opportunities of the digital age. Retrieved from ftp://ftp.cordis.lu/pub/etan/docs/ict-report.pdf.

Finholt, T. (2002). Collaboratories. Annual Review of Information Science and Technology. B. Cronin. Medford, NJ, Information Today. **36**.

Fish, R., Kraut, R. Root, R. & Rice, R. (1993) Video as a technology for informal communication. Communications of the ACM, 36 (1), 48-61.

Godefroid, P., J. Herbsleb, et al. (2000). <u>Ensuring privacy in presence awareness systems: An automated verification approach</u>. ACM Conference on Computer Supported Cooperative Work, Philadephia, PA.

Grasso, A. and J. Meunier (2002). <u>All ways aware: Who can claim abstinence from peeking at print jobs?</u> Paper presented at the ACM Conference on Computer Supported Cooperative Work, New Orleans, LA.

Grudin, J. (1989). "Why groupware applications fail: Problems in design and evaluation." <u>Office: Technology & People</u> **4**(3): 245-264.

Grudin, J. and L. Palen (1995). <u>Why groupware succeeds: Discretion or mandate</u>. European conference on Computer Supported Cooperative Work, Dordrecht, The Netherlands, Kluwer.

Handel, M. and J. Herbsleb (2002). <u>What is chat doing in the workplace?</u> Proceedings of ACM Conference on Computer Supported Cooperative Work, New Orleans, LA, ACM Press.

Herbsleb, J., A. Mockus, et al. (2000). <u>Distance, dependencies and delay in a global collaboration</u>. CSCW 2000, Philadelphia, PA.

Isaacs, E., A. Walendowski, et al. (2002). <u>Hubbub: A Sound-enhanced Mobile Instant Messenger that Supports Awareness and Opportunistic Interactions.</u> ACM Conference on Human Factors in Computing Systems, Minneapolis, MN.

Lange, B. M. (1992). Electronic group calendaring: Experiences and expectations. In D. Coleman, (Ed.), <u>Groupware</u>. San Mateo, CA, Morgan Kaufmann.

Mayhew, D. J. (1999). <u>The usability engineering lifecycle.</u> San Francisco, Morgan Kaufmann.

Mockus, A. & Herbsleb, J. (2002). <u>Expertise browser: a quantitative approach to identifying expertise.</u> Paper presented at the ACM International Conference on Software Engineering, Orlando, FL.

Mosier, J. and S. Tammaro (1997). "When are group scheduling tools useful?" <u>Computer Supported Cooperative Work</u> **6**(1): 53-70.

Nardi, B., S. Whittaker, et al. (2000). <u>Interaction and outeraction: Instant Messaging in action</u>. ACM Conference on Computer Supported Cooperative Work, Philadelphia, PA.

National Business Travel Association Press Release, 19 September 2001, http://www.nbta.org/info/pressreleases_survey_future.htm

Nielsen, J. (1993). <u>Usability engineering.</u> New York, Academic Press.

NRC (1993). National Collaboratories: Applying information technology to scientific research. Washington DC, National Research Council.

Olson, G. M. and J. S. Olson (1991). "User-centered design of collaboration technology." Journal of Organizational Computing **1**: 61-83.

Palen, L. and J. Grudin (2002). Discretionary adoption of group support software: Lessons from calendar applications. In B.E. Munkvold (Ed.) Organizational implementation of collaboration technology.

Ribak, A., Jacovi, M., Soroka, V. (2002). Ask before you search: peer support and community building with reachout. Proceedings of ACM Conference on Computer Supported Cooperative Work, New Orleans, LA, ACM Press.

Schocken, S. (2001). "Standardized Frameworks for Distributed Learning", Journal of Asynchronous Learning Networks 5 (2).

Vitaglione, G., M.K.Storr, H.A. Neal, C.Severance, & S.Goldfarb (2000) "Lecture Object: an architecture for archiving lectures on the web." 2nd Topical Seminar on Global and Local Network for Research and Education, Pontignano (Siena, Italy) Nov.2000