A Characterization of Multi-Party Interactive Multimedia Applications

Clemens Szyperski and Giorgio Ventre

The Tenet Group
International Computer Science Institute
and
Computer Science Division
Department of EECS
University of California, Berkeley

e-mail: {szyperski,ventre}@icsi.berkeley.edu

High Performance Network Research Report
February 1993

Abstract

This document tries to define and characterize a class of applications called Multi-Party Interactive Multimedia (MIM), for which many examples are given. This class includes applications such as CSCW, teleconferencing, and remote education; its consideration in this report is based on the observation that MIM applications are both important and representative for the area of high-performance real-time communication. Purely functional criteria are used to capture the MIM class, i.e. ones that are not related to any particular way of implementation. Thus, future directions are sketched that give some indications on what a network architecture will need to provide, in order to effectively support such applications.

1 Introduction

Among real-time multimedia applications, those requiring (and providing) interaction among a number of parties via a combination of communication media are of great interest to the network research community, because of their particular network support requirements. These applications, which may be called multi-party interactive multimedia (MIM) [8], ask for the availability of a communication system able to efficiently support the amount of data transfer they need, and do so with a client-specified quality of service (QoS). A last but not less important requirement is to provide the clients of such applications with a clear and simple interface to access the services offered by the network, to hide the internal architecture and mechanisms of the communication subsystem. Such services should be specifiable in purely functional terms, i.e. without reference to any particular implementation.

Other proposals have been presented in the literature for the classification of MIM applications from the viewpoint of the support they require from the communication subsystem [2, 7, 9, 10, 11, 13, 14]. In this report we propose a different approach to such classification based on the identification of a set of basic functional characteristics that this class of multimedia applications exhibits. We start from examples in the area of interpersonal communication, to extract a set of attributes we believe to be the most significant for the support of these examples by means of computer-based tools. We focussed our attention on those attributes related to the logical aspects of how the information exchange happens, and were not associated with any assumption on possible implementations over a real-time communication network. However, throughout the report we keep the eventual support by a reasonable network architecture in mind. Starting with a classification based on these attributes, we analyze the scale issue and then present the services that designers and users of MIM will require from a real-time communication system. Finally, we give a brief description of the mechanisms that real-time network protocols should provide to offer such services.

2 Definition of Terms

Due to their origin from interpersonal communication models, multi-party interactive applications have been, in general, rather vaguely defined. To avoid confusion, we start by defining some fundamental terms that will be extensively used in the rest of the paper. The proposed terminology has been derived from the one usually adopted for the definition and characterization of typical patterns in the area of interpersonal communication.

- *Conference*: Even though the term *conference* is generally associated to a specific type of multiparty interaction, we will interchangeably use both terms *MIM application* and *conference* to identify applications involving a number of participants exchanging multimedia data under real-time constraints.

- *Participant*: The parties involved in a conference are called *participants*. Each participant is an addressable network entity related to the conference in at least one point in time. A participant typically is a human being; however, we believe that our classification might be extended to consider also non-human participants, as is the case of some multimedia applications such as teleprogramming [5] and multimedia database access [12]. A basic criterion could be that the participants involved are assumed to be subject to real-time constraints, and are, in some sense, "interactive". Involvement in a conference is understood here in a rather general sense; e.g., it includes related activities taking place either before the conference's starting time or after its conclusion.

- *Role*: The participants relate to the conference by taking certain roles. Each role represents a contribution to the conference itself, its organization, or its management. In our opinion, four basic roles can be distinguished:

  Organizer (or Initiator): A participant who initiates/establishes the conference, decides on invitations, reserves in advance the resources needed to run the conference, specifies security and accessibility requirements, specifies the quality of service requirements (harmonizing them to the characteristics of the expected participants), and is responsible for paying the bill.

  Manager (or Chair): A participant who *runs* the conference, decides admission of late comers, moderates the discussion, makes decisions about the possible restructuring of the conference after admissions and departures, and closes the conference.

  Receiver: A participant who can access the multimedia data produced within the conference.

  Sender: A participant who can produce part of the multimedia data that may be consumed by the receiving participants of the conference.

  Note that a role can be shared by multiple participants. For example, in some case all participants (or a subset) can take the manager's role, effectively making the conference management fully (or partially) distributed.

- *Role Association*: A participant may qualify for multiple roles. In principle, participants may request or give up a role at any point in time. A participant is defined to have a certain role, even if it is not actually playing that particular role all the time. The classification we propose is related to whether a participant is to be allowed to assume a role during its lifetime. If it is, proper communication support must be provided by the network. Of course, certain restrictions might apply, depending on the application's characteristics. For example, the number of participants assuming a certain role and the dynamics of role association may be limited by rules or agreements, explicitly or implicitly accepted by the participants themselves.

- *Media*: Conferences are based on participants communicating through the use of a number of media. To clarify the scope of this document, the following are the media being considered here:

  Motion Video

  Audio

  Still Images

  Image Streams

  Traditional Data, possibly to be distinguished further into:

    Low-Bandwidth, Interactive

    High-Bandwidth, Bulk-Transfer

3 Towards a Functional Characterization of MIM

Multimedia applications, and particularly MIM, show an inherent complexity, derived from the different media they include and the large number of participants they might involve. To define a functional characterization of MIM, we first need a better understanding of the models of interpersonal communication these applications should be able to support.

Below, a classification is proposed for some typical examples of social and professional communication and interaction, based on their attributes that are primarily related to the functional aspects of the information exchange:

- *Model of Interaction*: This issue relates to how interaction may occur among participants. Depending on the ability of a participant to produce and transmit multimedia data, we can identify three different kinds of interactions:

  Dynamic: All the participants are allowed to freely send data without any form of restriction, even though some sort of control is usually implicit for this model of interaction. For example, social (e.g. parliamentary) rules or simply good manners apply to such applications like meetings. However, in this model we assume that no control is enforced by the application. A meeting of executives, a project design session, or an informal gathering of people are typical cases that may be included in this category.

  Static: This is the case when the participants that are allowed to send data are a statically determined subset that will not change through the application's lifetime. As an example, we can consider a public hearing or a musical performance, where only a small number of the participants (i.e. the official speakers and the music players respectively) produce information, while the majority of the attendees can only receive it.

Controlled: The subset of participants allowed to send data can change during the evolution of the application, but such a change is conditioned to the occurrence of conditions or on rules depending on the application itself. In this category fall examples like a course lecture, a seminar, or a conference, where the role of speaker can be temporarily assumed, in principle, by all the participants, as is the case of a course lecture, where the instructor will occasionally allow the students to ask questions.

- *Data Flow*: This characteristic is related to how the information flows among the participants. In some cases, a one-to-many data-flow can be observed, while others are characterized by a more symmetric, all-to-all type of flow. An example of the first type is a seminar, where both the speaker and the attendees are expected to intervene only one at a time, while an example of the second case can be found in meetings where multiple communications can freely and informally coexist. In other models of interpersonal communication, like panels and debates, a $M \rightarrow N$ data flow can be observed, where usually $M \ll N$.

- *Accessibility*: In some cases, the access to the information flow is restricted, and some sort of control is enforced for billing or security reasons. This implies that the participants must undergo an admission or monitoring process.

  Controlled: A course lecture is a typical example of a communication model where access is *controlled*. In fact, to be allowed to attend a lecture, a participant should be enrolled in that class. In other cases, such as business meetings, access is restricted for obvious security reasons.

  Uncontrolled: For some models, no kind of control must be enforced over the participants. In a public hearing, as is the case of a City Council audition, everyone is usually permitted (and encouraged) to attend. There might be examples, as in a seminar-like communication, where access rules are imposed on a case by case basis.

  Sometimes, access control is enforced by the existence of participant lists. For example, this might be the case when the number of persons allowed to participate in a conference is restricted. Sometimes a participant list is prepared before the communication occurs. In other cases, this list is dynamic, and may be modified during the event. However, it should be noted that for a large class of applications a hybrid situation is expected, where a certain subset of the participants is known when the communication is planned (*bounded subset*), while it is anticipated that the list of participants may be expanded during the communication lifetime to allow people originally not included in the list to join it. An example is a conference where a number of attendees announced their participation in advance, while the others have joined it only at the last moment. The two groups differ in the way they are treated. Participants who are known in advance are guaranteed to be accepted, while the others have to rely on the residual availability of resources (i.e. seats).

- *Event Scheduling*: There are a number of interpersonal communication examples where time limits are essential parts of the model. For example, conferences and collaborative work sessions are to be scheduled in advance, to allow participants to avoid conflicts with other commitments and to reserve the facilities needed. Therefore, we believe that the following two categories can be distinguished:

  Planned: This is the case when the communication is regularly scheduled, or more generally, just scheduled in advance. In addition to allow the participants to schedule in advance to attend the event, the availability of information like the starting time and the expected duration of the communication may be used for advance reservation of the facilities needed (e.g., a room or a particular resource).

  Unplanned: In some cases, nothing is known about when the communication will occur. A typical example is a casual encounter of coworkers that develops into a discussion over work-related matters, or any meeting that is organized impromptu. Of course, the possibility of successfully having a communication of this kind is subject to the current availability of needed resources.

Table 1 summarizes what we have presented in this section. It shows how a number of typical models of interpersonal communication fit into the classification proposed above. It should be noted that the table entries are meant to describe the dominant cases, and should not be considered as discriminative in the mathematical sense.

Even though other important examples have been probably left out, we believe that the ones included in the table are general enough to represent the characteristics of most of the cases. In the next section we will show what kind of network support we need so that such models can be effectively supported by computer-based multimedia tools.

4 Model of Interaction

We will now switch our attention from the physical world of interpersonal communication, to the virtual world of electronically supported, multimedia interaction. A number of MIM applications have been proposed to implement and support the communication models listed in Table 1. In addition to those, other applications have been presented, that do not have an exact correspondent in the traditional examples of social and professional interaction.

As we have seen in the previous section, interpersonal communication models show differences in the way data is exchanged among the participant involved. In order to offer computer-based multimedia tools implementing these and other models, we need to analyze how such differences influence the kind of communication service these tools require from the

Type	Description	Model of Interaction	Data Flow	Accessibility	Event Scheduling
Course Lecture	Long term setup, lecture w/ managed discussion	Controlled	1 -> N	Controlled	Planned
Seminar	Special event version of Course Lecture	Controlled	1 -> N	Controlled Uncontrolled	Planned
Casual Meeting	All participants actively send/receive data	Dynamic	N -> N	Uncontrolled	Unplanned
Business Meeting	All participants actively send/receive data	Dynamic	N -> N	Controlled	Planned
Public Hearing	Small subset of senders many receivers	Static	M -> N (M << N)	Uncontrolled	Planned
Conference Panel/Debate	Moderated Q&A from a subset of participants	Controlled	M -> N (M << N)	Controlled	Planned

Tab.1 Functional Characteristics of Interpersonal Communication Models

network. This can be done by considering, as the prevailing aspect, the *model of interaction* criterion, as it has been previously defined.

We will present our conjectures by considering a computer supported multimedia environment where participants to a multi-party communication are geographically distributed entities (*nodes*), interacting through communication abstractions (*channels*) over a real-time internetwork.

A *dynamic* interaction is required for MIM applications in which the amount of information that can be exchanged among all the participants is almost equivalent and the characteristics of the applications do not allow the determination of a particular subset of participants that will be the main source of the data flow. Since the coexistence of multiple flows of information should be, in these cases, permitted, support for simultaneous communications among the nodes should be provided by the network, in spite of any arbitration or control mechanisms that might be imposed by the users.

A *static* interaction is required in all the MIM applications where the information flows from a statically determined subset of nodes, or where the amount of data flowing from the other nodes does not justify the creation of a dedicated, real-time communication channel. Typical applications requesting such kind of interaction are those generally referred to as *telepresence* [6]. An example of this kind of application is *teleprogramming* [5], where a user receives a number of different media from remote devices, sensors or computers, typically to control or manipulate a remote apparatus, according to a many-to-one information flow.

A *controlled* interaction is required when the application specifically permits the identification among all the participants of a dynamically variable subset that can contribute to the data flow. The subset may depend on the specific attributes of the participants. For example, in a conference, there exists a small subset including the conference speakers and another subset of the conference managers. Both sets are expected to contribute to the data flow more than the attendees. The identification of the subset of senders can also depend on application-specific rules and agreements. For example, in a class lecture students are expected to ask for permission to interrupt the professor. In such cases, those rules and agreements dynamically modify the subset of participants authorized to transmit data. In the lesson example, the student who is authorized to intervene will become the speaker and the multimedia data produced by him will be transmitted to the other participants in addition to or in alternation with the data flowing from the professor.

An important example characterized by this model of interaction is the one generally classified as CSCW, or *computer supported collaborative work*. The goal of this service is to support the complex interactions required to allow effective meetings among geographically distributed professionals. In addition to video and audio data, participants in this model of interaction share applications such as editors, spreadsheets and drawing spaces. Information flows according to a highly dynamic, all-to-all pattern. Since the target environment is often that of technical and commercial enterprises, access control has usually to be enforced.

Another example is the so-called Virtual Café. The idea is to allow users at different sites to communicate informally through an electronic common space [1]. However, due to the informal nature of this model, access control is generally not of concern and scheduling is usually limited to the specification of some "opening hours".

Table 2 lists these new applications, and how they fit into the taxonomy we proposed in section 3. As with Table 1, the entries are meant to indicate the dominant cases, but should not be taken as discriminative in the mathematical sense. Two more examples have been added to this table: *broadcasting services* and *distributed computing*. Even though they do not directly fall into the definition of MIM applications, we included them for their relevance in terms of network support requirements.

5 Aspects of Size

So far we considered purely functional aspects of MIM applications. In the environment considered so far, no limits of any sort were assumed for the capability of a real-time communication system to provide the services required by these applications.

We believe, however, that the important issue of size of MIM applications should be taken into account [10]. By *size* we refer to the expected bound on the number of participants. In particular, we are interested in how support from a real-time communication subsystem has to cope with this aspect.

Type	Description	Model of Interaction	Data Flow	Accessibility	Event Scheduling
CSCW	All participants actively send/receive data	Dynamic	N -> N	Controlled	Planned
Telepresence	Many senders, one receiver	Static	N -> 1	Controlled	Unplanned
Virtual Cafe	All participants actively send/receive data	Dynamic	N -> N	Uncontrolled	Unplanned
Broadcasting Services	One sender many receivers	Static	1 -> N	Controlled	Planned
Distributed Computing	All participants actively send/receive data	Dynamic	N -> N	Controlled	Unplanned

Tab.2 Functional Characteristics of Selected MIM Applications

Indeed, the maximum number of participants that a particular application may support depends on when it becomes infeasible either for the human user or for the service provider to add another participant. For example, from a logical viewpoint, the dynamic model of interaction of some MIM applications might require a fully interconnected communication structure. This requirement seems to dictate a small bound on the maximum number of participants, since the connection complexity grows quadratically, and the information to be consumed by each participant grows linearly with the number of participants. Therefore, the existence of such a bound is due not only to implementation reasons, as those related to the complexity of the required communication structure, but also to the limits inherent in the human capability of effectively receive and elaborate several simultaneous information flows.

We studied how size relates to the various applications introduced in the previous two sections (Tables 1 and 2). The result is presented in Figure 1, where the number of participants and the model of interaction have been used as criteria.

From Figure 1 we observe that the upper bound on the number of participants is somehow related to the specific application. It should be noted that this figure is intended to be only indicative. Still, we can distinguish among small, large, and very large scale applications. As a first guess, we expect the bounds for each of the three classes to be around 20, around 100, and practically unbounded, respectively. Of course, more precise numbers require an analysis of individual applications.

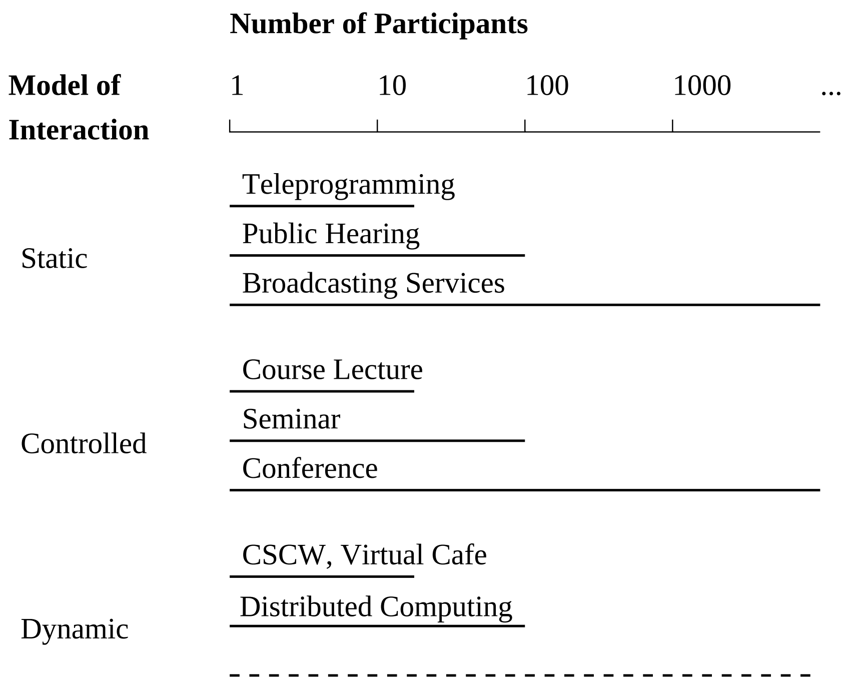

Fig.1 Size Ranges of MIM Applications

6 Requirements of MIM Application Clients

We can now try to list the MIM clients' wish-list, i.e. the list of services that a client would like to obtain from the network service provider, in order to access and manage MIM applications:

- *Directory of MIM Services*: A client should be able to obtain informations about the services being offered over the internetwork at a particular time. Such information should be of two classes: general and technical. The former class should be related to data regarding the kind of application, its subject, the scheduled date, time, and duration, according to parameters listed in Table 1. The latter should be related to the technical aspects of the applications, such as the one that are listed in this section: application management and control, quality of service, security and accessibility constraints, billing.

- *Joining and Leaving Applications*: The information provided by the Directory Service presented above will be used by a client to dynamically join or leave MIM applications. Therefore, mechanisms should be provided by the network to allow such dynamic behavior of clients, for example to contact the application organizer beforehand to get an invitation if possible.

- *Conference Organization*: This service should permit a real-time network client to set up, schedule, reserve in advance, and advertise its own MIM applications.

- *Conference Management*: After a client organizes a MIM application, the network must provide it with management mechanisms, including access control, authentication, invitation, initiation.

- *QoS Mechanisms*: MIM applications involve the coordination of a number of different media, each one of them characterized by its own QoS specifications. In some cases, a separate set of specifications might be allowed to each participant. The network should provide mechanisms for: *traffic requirement specification* (e.g. minimum and average throughput, burstiness), *performance requirement specification* (e.g. delay and jitter bounds), *synchronization control* (e.g. among destinations and among different media), *reliability specification* (e.g. loss probabilities and fault-tolerance requirements). We believe that the approach proposed in [4] to providing guaranteed quality of service for unicast real-time channels can be successfully extended to multi-party, multimedia communications.

- *Security Mechanisms*: Mechanisms should be provided to help network clients manage security and accessibility of MIM applications.

- *Scalability and Optimization*: MIM applications involve a number of different media, connecting a potentially large number of participants. Due to the large scale of some of the applications, some optimization in the allocation and management of resources is required. This functionality also requires the availability of load-balancing capabilities.

All the above mentioned functionalities are being considered as the basis for the design of a client-network interface, whose complexity increases dramatically when the network is able to offer quality of service in real-time communication [3]; some of these functionalities, however, are intended to be accessible only to a subset of the application participants or to the network administrators.

7 Directions for Future Research

In the previous sections we have listed the functionality that should be provided by a communication network to effectively support MIM applications. In this section we will briefly describe the mechanisms that should be implemented by the network to offer the desired service. Some of these mechanisms are already provided (albeit sometimes only partially) by some of the real-time protocols proposed in the literature, while others are yet to be incorporated.

First of all, the network should provide mechanisms for *resource control* and *reservation*, in order to offer quality-of-service protection and management to its clients. Since future multimedia applications will be very dynamic, management should be intended as a dynamic procedure, allowing the modification of the QoS parameters of MIM applications during their entire lifetime. For example, the dynamic inclusion into, and departure of participants from a conference might require a dynamic adjustment of the real-time channels to compensate their effect on the quality of service provided. Synchronization is probably the most important of the new services. Another important issue concerns the management of a directory, listing the currently available network services and applications.

A second issue relates to the *scalability* and *optimization* of multimedia communication over internetworks. Since MIM applications might involve different media and a large numbers of participants, possibly spread over a large internetwork, mechanisms should be provided to allow for optimization in the allocation and management of resources. Such mechanisms should also include load-balancing capabilities of the network. Clients should be able to accurately characterize their requirements to the service provider in order to implement such mechanisms in the network.

Another important capability is to handle heterogeneity among both machines and networks. *Interoperability* should be transparently provided to the clients. However, it is our opinion that in some cases heterogeneity should be explicitly addressed, to be potentially exploited by the network. For example, if some machines do not have the necessary facilities to handle video, they should not have video data delivered. Likewise, support for *internetworking* has to be provided in such a way that networks offering different approaches to QoS might be accessed through a common client-network interface. The definition of such an interface determines how a client can negotiate with the network the quality of service of a requested communication service.

In addition, some other properties should be available to real-time network clients, including *fault tolerance* and *security control.* Fault tolerance, in particular, is required to allow a reliable use of the network by critical multimedia applications such as teleprogramming. Networks should offer mechanisms to allow applications to cope with events such as malfunctioning nodes and links, and also failures concerning the protocol-level entities. Last but not least, support is required for *accounting*, i.e. the ability to determine (in advance) the costs of using network resources. This includes the possibility for an application to negotiate quality of service parameters within a given budget.

Acknowledgments

This report is based on a series of discussions held within the Tenet Group and, in particular, during the meetings of Tenet's Multicast SIG. The authors are especially grateful to Domenico Ferrari, Amit Gupta, Bruce Mah, Mark Moran, Colin Parris, Makiko Yoshida, and Hui Zhang, for their ideas, comments, and criticisms. Wendy Heffner and Ginger Ogle are acknowledged for their contribution to an earlier version of the MIM applications taxonomy.

This research is supported by the National Science Foundation and the Defense Advanced Research Projects Agency (DARPA) under Cooperative Agreement NCR-8919038 with the Corporation for National Research Initiatives, by AT&T Bell Laboratories, Digital Equipment Corporation, Hitachi, Ltd., Hitachi America, Ltd., Pacific Bell, the University of California under a MICRO grant, and the International Computer Science Institute.

References

[1] S.A. Bly, S.R. Harrison, and S. Irwin. Media Spaces: Bringing People Together in a Video, Audio and Computing Environment. *Communications of ACM*, 36(1), January 1993.

[2] Yee-Hsiang Chang. Remote Conferencing: Statement of Desired Functions. *Internet Engineering Task Force, Remote Conferencing BOF Draft*, October 1992.

[3] D. Ferrari, J. Ramaekers, and G. Ventre. Client-Network Interactions in Real-Time Communication Environments. In *Proceedings of HPN'92, International Conference on High Performance Networking*, Liege, December 1992.

[4] D. Ferrari and D. Verma. A Scheme for Real-Time Channel Establishment in Wide-Area Networks. *IEEE Journal on Selected Areas in Communications*, 8(3), April 1990.

[5] J. Funda. Teleprogramming: Towards Delay-Invariant Remote Manipulation. Technical Report MS-CIS-91-40, University of Pennsylvania, Dept. of Computer and Information Science, May 1991.

[6] D.R. Hofstadter and D.C. Dennett. *The Mind's I: Fantasies and Reflections on Self and Soul.* Basic Books, New York, 1981.

[7] T. D. C. Little and A. Ghafoor. Network Considerations for Distributed Multimedia Object Composition and Communication. *IEEE Network Magazine*, November 1990.

[8] M. Moran and R. Gusella. System Support for Efficient Dynamically-Configurable Multi-Party Interactive Multimedia Applications. In *Proceedings of the Third International Workshop on Network and Operating System Support for Digital Audio and Video*, San Diego, November 1992.

[9] J.C. Pasquale et al. The Multimedia Multicast Channel. In *Proceedings of the Third International Workshop on Network and Operating System Support for Digital Audio and Video*, San Diego, November 1992.

[10] E.M. Schooler. The Impact of Scaling on a Multimedia Communication Architecture. In *Proceedings of the Third International Workshop on Network and Operating System Support for Digital Audio and Video*, San Diego, November 1992.

[11] H. Schulzrinne. A Transport Protocol for Audio and Video Conferences and other Multiparticipant Real-Time Applications. *Internet Engineering Task Force, Internet-Draft*, October 1992.

[12] M. Stonebraker and J. Dozier. Large Capacity Object Servers to Support Global Change Research. Technical Report 91/1, Sequoia 2000, Berkeley, 1991.

[13] C. Topolcic, editor. *Experimental Internet Stream Protocol, Version 2 (ST-II).* Request for Comments. Networking Working Group, N. 1190, October 1990.

[14] R. Yavatkar. Issues of Coordination and Temporal Synchronization in Multimedia Communication. In *Proceedings of IEEE Multimedia '92*, Monterey, 1992.