Rich Media *from* the Masses

Sujata Banerjee, Jack Brassil, Amy C. Dalal, Sung-Ju Lee,
Ed Perry, Puneet Sharma, Andrew Thomas
Internet Systems and Storage Laboratory
HP Laboratories Palo Alto
HPL-2002-63 (R.1)
May 20th , 2002*

E-mail: team -media@groups.hp.com

rich media, streaming media, media distribution networks, Internet

By putting the tools for media creation in the hands of the consumer, there will be exponential growth in the distribution and hosting of media on the Internet. We expect this growth to parallel that of HTTP traffic. Our position is that we need to confront the issues related to the potential exponential growth in the number of media sources and services. Further, we believe that significantly higher quality streaming and context-sensitive personalization of multimedia services are key enablers of this whole technology area. Our vision differs from the delivery of video-on-demand services because we believe that there will be a vast number of sources of copyleft content. In this paper we focus on four key areas: the distributed discovery of content; the concurrent manipulation of multiple media streams; media sensitive mechanisms for providing end-to-end services over the Internet; and the subjective assessment of delivered media quality.

* Internal Accession Date Only Approved for External Publication

Rich Media *from* the Masses

Sujata Banerjee, Jack Brassil, Amy C. Dalal, Sung-Ju Lee, Ed Perry, Puneet Sharma, and Andrew Thomas
Internet Systems & Storage Lab
Hewlett-Packard Laboratories
Palo Alto, CA 94304
team-media@groups.hp.com

Abstract

By putting the tools for media creation in the hands of the consumer, there will be exponential growth in the distribution and hosting of media on the Internet. We expect this growth to parallel that of HTTP traffic. Our position is that we need to confront the issues related to the potential exponential growth in the number of media sources and services. Further, we believe that significantly higher quality streaming and context-sensitive personalization of multimedia services are key enablers of this whole technology area. Our vision differs from the delivery of video-on-demand services because we believe that there will be a vast number of sources of copyleft content. In this paper we focus on four key areas: the distributed discovery of content; the concurrent manipulation of multiple media streams; media sensitive mechanisms for providing end-to-end services over the Internet; and the subjective assessment of delivered media quality.

1 The Vision

We envision a future where Internet users will interact with applications, other humans and perhaps even objects in the environment, primarily via multimedia objects and streams. We further envision a world where every user of a service can receive live, stored or interactive media content that is personalized, custom-made, context- and location-aware for that specific user or interest group. In this environment, service providers will be able to compete for customers on primarily their value-added services offering. Examples of value-added services include advertisement insertion and commentary overlay in live or stored media streams. These services will require the manipulation of multiple streams sourced from different origins.

Furthermore, with emerging multimedia authoring tools, there is the potential for every Internet user to be a media content author and source, much like the prior WWW trend. Many future services will require high bandwidth, high quality, large format media streams. Today's postage-stamp format video will not suffice. The final component of the vision that presents the most significant technical challenge is that the above be achieved over a cost-effective shared infrastructure and be available to the public at large over both wired and wireless access links. A natural progression to the above vision is the wide-spread availability of highly interactive *tele-experience* services.

This vision is in direct contrast to the conventional notion of mass dissemination of entertainment-quality video-on-demand, the goal of which is to deliver highly polished, copyrighted, rights- managed media at a price. We believe that there are some critical differences in the infrastructure required to support our vision (see Section 2).

First we highlight the technology drivers that we think will make our vision possible. We then postulate some scenarios which differ from today's content authoring and delivery. We discuss the challenges that our vision presents us with in Section 2. Section 3 lists some novel research problems associated with this vision, followed by concluding remarks in Section 4.

1.1 Technology Drivers

Technology advances are paving the way for rapid growth in the distribution of real-time, multimedia streams. Consider the following:

Backbone Capacity Growth Advances in optical communications technology including dense wavelength division multiplexing and emerging metropolitan area network systems continue to increase Internet backbone and access transmission rates.

Broadband Access Links The deployment of cable modem and DSL service has greatly increased the access bandwidth available to the home, with some five million homes wired in the U.S. [12].

Improved Peering for QoS Backbone providers now offer end-to-end Service Level Agreements on which network QoS can be built.

Video and Audio Coding The desire to stream and download audio and video over the Internet has fueled interest in low bit rate video and audio encoders. The initial results are somewhat crude, but gradual improvements in codecs along with increasing bandwidth will lead to a richer media experience, with audio and video quality becoming similar to broadcast television.

Handheld Digital Capture Devices The consumer electronics industry is rapidly producing digital appliances to supplant the "old" analog equivalents. MP3 players are replacing portable cassette tape players, digital cameras are replacing film cameras, and digital video is challenging analog tape formats in the camcorder space. These digital devices are readily connected to the PC which acts as a bridge between the various devices and the Internet.

Home Video Production Tools Improvements in PC technology, particularly processor speed and disk space, have made it possible to store and edit digital video at a very attractive price point. The capabilities are already built into many high end personal computers and will soon be available on all new PCs. The net result being that anyone with a PC and a digital camera will have the technology to be a cinematographer, editor and director of his or her own movie.

Wireless Systems Improvements in wireless technologies (both LAN and mobile telephony), video coding, and miniaturization of audio and video capture technologies make possible a world where media can be streamed from anywhere into the Internet. This should open up a new world for the spontaneous creation and dissemination of media.

1.2 Scenarios

Media-centric services will dramatically and fundamentally change our future way of life. Below we take four seemingly familiar examples and expand them to demonstrate how we think future media services will be constructed.

Home Director Increases in available network bandwidth mean that future commercial Internet television stations will not face the channel (bandwidth) limitations faced by conventional broadcasters. For live broadcasts (e.g., news, sports), conventional broadcasters have historically created and transmitted a single "moderated," or produced, program. Regardless of the number of cameras covering an event, at each instant only one is selected for broadcast, with all other camera feeds terminating in the local production room. With Internet backbone bandwidth plentiful, these alternate points-of-view could be transmitted in parallel with the primary, moderated stream, and each of these streams might find its own audience. Further, interaction between the remote audience and the event site can direct camera activity. Alternatively, different points of view might be of interest to certain audiences (e.g., attention on specific players or positions for training purposes).

Sportscast Commentary American football fans in the U.S. have grown accustomed to seeing the electronic yellow "first down" line superimposed on the field in the television broadcast of games, or the commentators drawing "play analysis" that is superimposed on the image of the field. The speed-skating lanes in the Winter Olympics are superimposed with the flag of the country that the athlete represents. The broadcast of NASCAR races have virtual flags attached to each of the cars for easy identification by the viewers. These are examples of value-added features to the original content that have tremendous support from the audience. However, production of these seemingly simple features is a big task. For example, a crew of four people (including one on the field) and five computers is dedicated to the "first down" line feature [46]. In the future, perhaps two well-renowned commentators will be brought in on the fly to discuss the game as it is happening and "draw" play analysis on the screen, when neither is actually at the game or the production site.

Media Web Television advertisements already carry URLs in the picture, but this limits the seamless transition from the world of television to the world of the Web. The Advanced Television Enhancement Forum has defined a way to carry URLs in the NTSC video picture [15]. Widespread adoption of a mechanism like this would greatly ease the integration of television and the Internet. For instance, during the broadcast of a program there may be a piece of background music playing. The program metadata could contain a URL about that music. Metadata might be rendered upon viewer request, like subtitles.Noticing an associated URL, the viewer could then seek out product information and perhaps make an online purchase of that music. After the browsing session is complete, the user returns to the streamed media to resume viewing.

Pledge Week Imagine that it is pledge week on the Internet. One hundred live performers along with some pre-recorded content are simultaneously accessible through 1000 different charitable organizations, each of which gives a user access to a different production of the performances. Charitable donors select and tune in to the performances through a media finder. Unlike today, after the user has pledged funds to a charity, he or she does not have to view any of the pledge breaks.

There are several more challenging scenarios as well. What if you could receive time-shifted, personalized "radio" in your car, home, office, or hotel room? What if remote collaboration was really simple to use and readily available to businesses and consumers? What if you could cheer your child's soccer team from your office or on the road stuck in traffic?

Today, media services such as broadcast quality television (RTPtv [14]) are available over Internet-2 [22]. The real challenge is to make them available to the masses over resource-constrained content networks at an affordable cost to both the service provider and the consumer. In the next section, we provide some background on the research challenges alluded to in this section.

2 Elements of the Vision

Our position is that we need to confront the issues related to the potential exponential growth in the number of media sources and services, so we are not taken by surprise again, like the growth of the Web in the 90s. Further, we believe that significantly higher quality streaming and context-sensitive personalization of multimedia services are key enablers of this whole technology area. To summarize, our vision rests on the following four pillars of thought:

- Huge number of rich media sources
- Requirement for high quality media
- Personalized custom-made and context-aware media streams for individuals or interest groups
- Cost-effective evolution of the Internet infrastructure

2.1 Huge Number of Rich Media Sources

We have wondered whether the world of streaming media will come about. Where will this content come from? Who will author it? These are exactly the same questions we asked ourselves before the World Wide Web surpassed everyone's expectations on how it would be used.

It seems highly likely that there will be two camps: the entrenched copyright owners (such as the large studios) who will provide copyrighted movies and audio to large numbers of consumers; and the home users who will take videos of their family and vacations and the like, probably not for financial gain, and without copyright concerns, who will produce a substantial amount of content for consumption by a different demographic than the copyright owners. One family in a hundred producing a one hour home video per year (one million hours for the US alone) will dwarf the output of the major studios. Finding and cataloging this material and making it available to its audience is a significant challenge.

The growth of multimedia on the Internet in some ways parallels the rise of the World Wide Web in the 1990s. Just as the Web increased the number of content "publishers," as shown by the growth in the number of Web sites in Figure 1, we expect the number of media "producers" to rise as well.

While numbers for the popularity and growth of all forms of multimedia on the Internet are not readily available, there is some data on the rise in popularity of one form of streaming media: Internet radio. Figure 2 shows how the total time spent listening to Internet radio stations has risen by a factor of five in the past year alone.

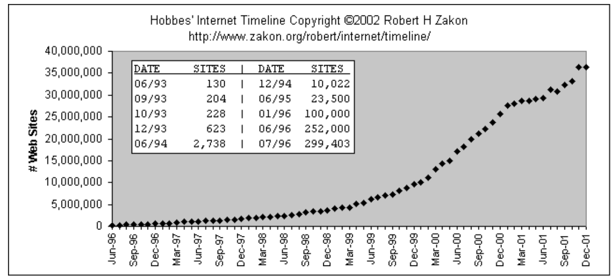

Figure 1: The growth in the number of World Wide Web sites, 1993-2001 [58].

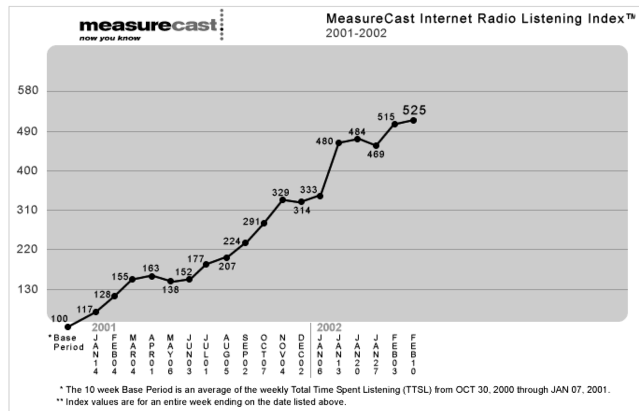

Figure 2: Weekly Internet radio listening index, February 10, 2002 [31].

2.2 High Quality Media

Quality is a highly subjective issue, and discussion of what might be an acceptable level of quality is beyond the scope of our work. However, we are motivated to provide an experience similar to what people are most familiar with: television. We can think of this as the source quality for the media, which will be relayed over the Internet, perhaps suffering delay and loss along the way. Below, we analyze the stress that high-quality media sources will place on the network.

High quality audio means CD-quality stereo sound. A rate of 128 kb/s (MP3) is considered "near CD quality." Multi-channel (5.1) systems run at higher bitrates, with Dolby Digital using 448 kb/s.

Video can consume substantial bandwidth as shown in Table 1. There has been substantial interest in compressing video, using DivX and MPEG-4 to fit 90 minutes of video and audio on readily available CDR material. It is likely that streaming at around 1 Mb/s will be very attractive because of the amount of material being encoded in this way.

Table 1: Sample high quality video encoding rates.

DV (Digital Camera)	35 Mb/s
Hi Definition (broadcast) TV	20 Mb/s
DVD	10 Mb/s
Standard Definition TV	5 Mb/s
Video CD (VCD)	1.15 Mb/s
DivX (90 minutes/700MB)	1.03Mb/s

For the near term it is likely that standard definition television is as high as we are likely to set the bar. The consumer DV video format is widely available, but it will almost certainly have to be transcoded for streaming. The demand for HiDefinition is very limited because of the (prohibitively) high cost of consumer access to video cameras and editing equipment.

Encoding media at higher quality levels is meaningless if the media cannot be delivered at the same quality levels. When transmitting high quality media over the Internet or shared infrastructure, we risk losing the gains in quality from these encodings at the source. Ensuring proper delivery requires correlating objective metrics with subjective, end-user perceptual quality analysis, which we discuss in Section 3.4.

2.3 Personalized, Context-Aware Media Streams

Media today flows directly from the content creator to the audience without any intermediate manipulation. In the future we see this situation changing by the addition of value added content and services. These additions may be performed by a third party, and we are interested in supporting this model.

We can see examples of repurposed content today.An interesting example is “pop-up video” [47] which is shown in the US on the VH1 television channel. The original content is a pop music video which provides the backdrop for speech bubbles which pop-up frequently during the video. The material in each bubble is some anecdote about the video, the artist, the location, the clothes and so on. In addition to modifying the video, the sound is augmented with a bubble popping noise or other appropriate jingle.

Mystery Science Theater [44] is another example of repurposed content, where movies are broadcast along with an additional overlay of cartoon characters in a cinema. The added value here is the lampooning of the feature by the cartoon characters who contribute a witty dialog for the duration of an otherwise dull movie.

Taking these two examples, we can postulate a similar phenomenon for media on the Internet. Moreover, the Internet is an open delivery mechanism where content can, at least in theory, be modified and republished by anyone. Today many sites exist which are devoted to televisions shows, offering plot summaries and character features. In the future these sites could repurpose content with fans of show adding their own audio or video commentary, or perhaps producing their own “directors’ cuts.”

Streams could also be customized based on personal profiles and preferences. Examples include insertion of location aware advertisements or specific camera feeds as described in the home director scenario in Section 1.2.

2.4 Cost-Effective Infrastructure

End-to-end bandwidth is still not plentiful or cheap in the Internet as well as in many corporate intranets today. At the very least, it is not cheap enough to afford massive overprovisioning. For the near- to medium-term future, a shared infrastructure will be the only cost-effective way to achieve some of the objectives of our vision.

A rough model of the Internet is that a consumer connects to an ISP over an ADSL, cable modem, or dial-up connection. The ISP buys transit from a backbone provider. The backbone is attached to another ISP who hosts the target service (and servers) that the consumer is trying to reach. This model has many problems and is not authoritative in any way, but is useful for the following illustration.

A rough model of the Internet is that a consumer connects to an ISP (on ADSL, cable or dial-up). The ISP buys transit from a backbone provider. The backbone is attached to another ISP who hosts the target service (and servers) that the consumer is trying to reach. This model has many problems and is not authoritative in any way, but is useful for the following illustration.

A simplistic back of the envelope calculation provides insight into the magnitude of the costs involved. Consider a 90-minute, 10 Mb/s DVD movie download (not streaming) service over the Internet. Table 2 compares the cost to download a single DVD movie over various communication link types. The monthly cost of the link is normalized with respect to bandwidth in “Megabit month,” which is the cost to send 1 Mb of traffic continually for a month on the link. The Megabit/hour costs are simply the megabit month costs divided by (30 days $\times$ 24 hours).

Using the rough model of the Internet, downloaded media traffic must cross two highly expensive transit links (target ISP to backbone, backbone to consumer ISP) and this has to be paid for by the fee the consumer pays for access. Table 2 demonstrates that this model is not economically feasible. This argues for, at least in the near term, *content delivery networks* and caches where content is transferred into a consumer’s ISP once and consumed many times in order to mitigate the transit costs. Sharing of the infrastructure by multiple users is essential to achieve the required cost benefit.

Rather than wait for the whole Internet to evolve, and in reaction to transit costs, large ISPs who have a large customer base may find the “walled garden” business model successful, in which media can only be streamed between

Table 2: Typical costs to download a DVD movie over various link types.

Link	Monthly cost	Mb/month	Mb/hour	DVD movie
AT&T 2.4 Gb/s OC-48	$840,000	$350	$0.48	$7.2
AT&T 155 Mb/s OC3	$72,200	$465	$0.64	$9.6
AT&T 1.544 Mb/s T1	$1,970	$1313	$1.82	$27.3
Typical 56 kb/s Dial	$20	$357	$0.49	$7.35
Typical 40 Mb/s Cable	$40	$1	$0.001	$0.015
Pac Bell 1.5 Mb/s ADSL	$50	$33	$0.04	$0.60
Pac Bell 6.0 Mb/s ADSL	$179	$30	$0.04	$0.60

customers of the same ISP. That there is not ubiquitous access to service (a cornerstone of the Internet) may be a little disturbing, but nevertheless this may be a profitable way to make progress.

Some ISPs may face considerable infrastructure upgrades should high-quality streaming media take off. To illustrate the magnitude of the issue, in the 650 area code Boardwatch [8] lists 109 ISPs, only ten of which have a peering bandwidth of more than 100 Mb/s.

We have mentioned that Web caching and content delivery networks can help reduce the cost of access to streaming media. In addition to Web hosting we believe that hosting streaming media content will be an important function of ISPs. This is driven by an observation about access technology: most links (ADSL, cable, 56k dial-up) have asymmetric bandwidth, with a lower speed uplink from the home. Consequently, it is better to upload content into the network and stream it from a staging post rather than streaming directly from the home.

3 Novel Research Problems

We next identify four important research areas list the novel problems that need to be addressed for each. The four areas are (1) announcement and discovery of *interesting* media content, (2) program manipulation, (3) end-to-end performance issues and (4) media quality assessment. These research areas cover the problem space alluded to in the previous section.

To understand why these areas are so important to us, it is necessary to understand what we have chosen *not* to do. We are not building a video-on-demand system for movie and television content. A commercial system like this would most likely be *over-provisioned* with *dedicated resources* to ensure a good customer experience at possibly a high cost to the user. Further, in such a system, content manipulation would be forbidden, while announcement and discovery would be reduced to an electronic TV Guide. By starting with a large number of sources we have a very different problem. The sources will necessarily be widely distributed, and hence require a decentralized location service. Access to the media will be over a network with varying degrees of service requiring media assessment, adaptive protocols and media manipulation.

We acknowledge that there are several other important research areas that will have an impact on the viability of the services we envision. These include media encoding, privacy, security, digital rights management, and pricing mechanisms, and are out of scope of the paper. It is our belief that the kinds of applications we envision will first emerge in corporate intranets or "walled gardens" where these issues are less problematic.

3.1 Announcement and Discovery

In each of the scenarios described in Section 1.2 a *directory system* is the presumed basis for both announcing and discovering multimedia content. The challenges associated with this system are formidable. The system must scale to enable a vast audience to identify content of interest. Unlike traditional information discovery systems (e.g., Web search engines), the system must be capable of handling short duration, real-time communications. These sessions can be pre-scheduled, or occur spontaneously and be announced at the session start. The directory system should also be lightweight, consuming only small amounts of computing and bandwidth resources. Directory service users should have the benefit of powerful searching tools, yet search times must be kept low. Announcements of events should be sufficiently expressive to facilitate discovery, without putting undue burden on the announcer, who may in fact be little more than a camera connected to a communications link.

Further, the directory system might take on additional features when used in an enterprise setting. Here the directory could benefit from integration with other meeting directory services, such as that for Microsoft's NetMeeting. This integration could further extend to announcements for both Internet and conventional telephone system conferences. For example, an Internet voice conference call initiated by the Session Initiation Protocol (SIP) [21] might conceivably issue an INVITE message to our directory system for the purposes of a public announcement. Similarly, an announcement of a conventional telephone conference call, now often distributed semi-privately by mechanisms such as electronic mail, might also be listed. Indeed, if connected to a telephone switching system, the directory would in principle be able to announce any active connection in progress.

3.1.1 Related Work

Though research in conventional information discovery systems has been extensive, relatively little attention has been paid to systems capable of discovering live multimedia sessions. As an example of a conventional system, the Domain Name System (DNS) is a robust, heavily-used decentralized database primarily for name/address translation. DNS has a variety of desirable properties that we seek, including robustness, global scale, and caching. DNS queries are quite limited in nature, however, as is the data stored in the system (e.g., CNAMES, PTR). Moreover, even with advances such as dynamic DNS, management of system data is restricted, and propagating new records throughout the system can take hours, far too slow for our application.

The announcement of live multicast sessions on the MBone is achieved through use of the Session Announcement Protocol (SAP) [20], and discovery is realized by client-based software 'session directory' tools such as *sd*, *sdr*, and *rsdr*. SAP relies on the presence of IP multicast, and essentially implements a soft-state protocol; clients monitor an announcement channel, and perform all discovery operations (e.g., searching) locally. Announced session are described by the somewhat unexpressive Session Description Protocol (SDP) [19]. SAP/SDP was later extended to support a limited session directory hierarchy.

Though shown to be useful as a discovery system for small-scale, prescheduled MBone sessions, a SAP based system suffers certain limitations. Efforts to conserve bandwidth by making infrequent announcements caused relatively long startup delays for users joining the multicast group at program start time. The system was not intended for use by unscheduled sources, nor would it be effective for very short-lived sessions. SDP is also quite limited as a descriptive language for programming metadata.

A collection of research attempts have addressed these limitations. SDPng [27], a successor to SDP, is an XML-based session description language under consideration within the IETF. Also, tools to permit Web browsers to interact with directory information have tended to displace dedicated client-side session directory tools.

One proposal has specifically addressed a number of the scaling limitations within the SAP-based discovery system design. The Information Discovery Graph (IDG) [50] seeks to impose a hierarchical network of directory managers—each responsible for a semantically distinct session topic—to balance system load, facilitate searches, and reduce system overhead.

3.1.2 Research Challenges

No existing directory service appears capable of coping with the scale and the dynamics of the media distribution system we are now considering. Among the difficult questions we seek to answer are:

Announcement Creation Each announcement must effectively describe its associated multimedia session with sufficient expressiveness such that a simple search by a potential recipient will discover the session. Yet suppose that there are two closely located sources, each unaware of the other, transmitting live video from the same event. How can we ensure that their independent announcements facilitate discovery? The challenge becomes greater if we consider sources with extremely limited data input capabilities (e.g., digital cameras). Can metadata not requiring human input, such as unique device identifiers or GPS coordinates, be used to facilitate content searches?

Access Control and Privacy The directory system we envision will be most powerful if it supports public, private and semi-private announcements. How can we ensure that the only parties to receive an announcement are those we desire? In addition, a receiver might desire personalized directory presentation according to pre-defined preferences. How will filtering of listings take place, and where should this filtering occur?

Directory System Architecture The architecture of the directory system must be scalable yet still enable rapid directory dissemination. Should such a system combine both *push* and *pull* distribution mechanisms? If so, how do we prevent a program directory listing from being stale? Will multicasting be part of the solution?

3.2 Program Manipulation

Future media streams will have audio, video, textual and markup components, and interesting multimedia programs will require several semantically related streams to be composed and customized. We think of *program manipulation* as the control, modification and coordination of these streams.

3.2.1 Related Work

Defining the layout, interaction and temporal and semantic relationships among various media streams is very important for composable and customizable multimedia presentations. SMIL [3] (Synchronized Multimedia Integration Language) is a language developed by the SYMM (Synchronized Multimedia) Working Group of W3C to describe time-based structure for multimedia *documents*. CMIF (CWI Media Interchange Format) [13] also allows composition of multiple video and audio streams.

Various inter- and intra-stream synchronization algorithms have been proposed to smooth out the network delay jitter. A comparison of several such techniques has been presented in [24]. Several video compression standards such as

MPEG-4 and MPEG-7 also allow for synchronization markers.

3.2.2 Research Challenges

We discuss key research problems in program manipulation.

Stream Manipulation Along the path between the original media and the final audience, the stream may pass through several processing steps to add value. These processing steps occur on *mediators* which perform the following functions:

Switching This component would take several incoming streams and generate a single outgoing stream. An example is switching between various camera feeds in the home director scenario described in Section 1.2.

Semantic Manipulation A simple example of semantic manipulation is inserting a logo into an existing stream. More complex examples include editing "pop-up" text or selecting an audio stream appropriate for a user's preferences.

Transcoding Transcoding would transform material from one format to another (e.g., MPEG-2 to MPEG-1), increase compression to allow transmission over a lower bandwidth link, or add appropriate link level forward error recovery.

Synchronization We might wish to coordinate the arrival of multiple streams at clients to reduce the client buffering load (especially for mobile devices).

We believe that content distribution networks (such as Speedera and Akamai) are the pre-cursors to mediators. We argue that CDNs work for on-demand video, but that much of the research on optimal placement of mediators will be revisited because of the advanced processing and requirement for "live" low latency video.

Copyright also has an interesting architectural impact on how and where streams may be manipulated. In repurposing content a third party is producing a derived work, which may be a breach of copyright (Section 106(2) of the US Copyright Act does not allow unauthorized derived works). For example, a third party may take a video and create a commentary but be prohibited from mixing the two and then republishing it. On the other hand, it may be perfectly reasonable for the third party to pass a reference to the original work and for the commentary to the audience to be rendered in an appropriate browser.

Stream Markup Program manipulation requires enhancing the streams with metadata not only about their identities but also about structure and semantics. Cueing protocols [9] provide one mechanism for transporting metadata in-band with the stream.

Repackaging audio and video streams with additional markup allows us to personalize content. The markup allows the expression of the type of content which could be matched against the user's preferences to create a personalized experience of the media. We see this today in parental controls, but in the future we imagine a much wider expression of what is in the streams.

Stream markup will also aid synchronization among multiple streams with different origins. Though some video compression standards have proposed synchronization markers, we require a format-independent mechanism for stream synchronization.

Multi-Stream Program Composition The W3C standard SMIL [3] can be used for defining the layout and temporal structure of multimedia documents. Several SMIL-based presentation authoring tools are now available. Most of these presentations are limited to static (predefined) streams. For example, it would not be possible to author a sportscast using SMIL where a viewer is interested in watching only the camera feed that has a particular player. Novel mechanisms are required to be able to dynamically compose media presentations based on stream metadata and user preferences.

Navigation The ability to effectively navigate streams within multi-stream programs is key to good viewer experience. The single stream control protocols such as RTSP [43] are restricted to simple commands like *start*, *fast-forward*, *rewind*, *pause* etc. Such primitives are not adequate for proposed program manipulation. Navigation among the different components of a multi-stream program will require new control primitives and protocols for exchanging these commands. For instance, new primitives are needed to express a viewer's intention to switch between TV and the Web in the media Web scenario described in Section 1.2.

Personalization Users will need some way of expressing their preferences so that streams can be personalized for them. Stream metadata/cues will provide the mechanism for personalization, but how users will specify their preferences is an open problem. We note that many useful services could be constructed by making the profile available publicly, and many more services would be enabled by including location information as part of a user's profile. However, history has shown us that there are privacy and security issues related to dissemination of user information.

3.3 End-to-End Performance Issues

The problem of network performance for media is a well-researched topic and addresses the problem of how media services can be delivered with acceptable quality over resource-constrained content networks. Quality of Service (QoS) requirements of media streams are very stringent in terms of the acceptable delay jitter and packet loss. Typically, existing Internet media players attempt to hide the delay jitter in large playback buffers (often 30 seconds or larger) and scale down (the already low) transmission rates to battle losses. However, if the goal of deploying very high bandwidth media streams that may be transmitted live or as part of interactive applications, requiring very small or negligible playback buffers, is to be realized, significant strides in QoS deployment will need to be made. All of the scenarios described earlier require multiple semantically related, high quality media streams to be delivered over a cost-effective shared infrastructure. The area of coordinating the performance of multiple correlated streams is a new area of research that has not received any attention. Since end-to-end bandwidth is likely to be a scarce and highly variable commodity, at least in the near- to medium-term, we argue that innovative and non-conventional methods are required to continually search out the necessary resources that can provide high media quality.

3.3.1 Related Work

Predictable end-to-end performance is still an unsolved problem because the Internet is a heterogeneous, federated environment connected together at private and public peering points. Some parts of the Internet have very good QoS and traffic management mechanisms implemented, while other parts may have no such mechanisms. In such an environment, it is difficult to obtain any end-to-end performance guarantees. The offered load on the Internet exhibits unexpected long-term as well as short-term variations, which has the impact of making the overall performance quite unpredictable and unsuitable for media applications. We briefly describe the research done to mitigate performance problems.

Over-provisioning Stemming from the observation that QoS is achieved automatically in an under-utilized network, a primary approach to achieving good network performance is by gross over-provisioning. However, massive over-provisioning of network resources is very expensive, particularly at the levels required to accommodate global high quality media streams. Due to the purported high costs of lighting dark fiber, less than 5% of fiber in the ground has been lit [41], and in the current market conditions, making more bandwidth available is not perceived as being urgent. Further, in spite of good growth levels, broadband home access penetration is still fairly low. That said, Internet backbone bandwidth has been increasing dramatically over the last decade.

While Internet performance has indeed improved from the early 90's, the emergence of the Web and increased Internet usage has caused a proportionate increase in network traffic flow. The net result is that serious performance issues still exist, and this trend is likely to continue for some time to come, particularly as media applications grow in popularity [30]. Also, depending on the traffic mix within the network and the often imperfect nature of peering arrangements, congestion points can develop dynamically and unpredictably, causing performance impacts, particularly on media streams that suffer greatly from short-term fluctuations in quality of service.

Resource Reservation The most common technical approach to providing network QoS guarantees to applications has been based on a resource reservation approach. However, the success of this approach is heavily dependent on the capability of applications and services to accurately predict their demands on short as well as long time scales. This is inherently a difficult problem for most applications/services and has either of two effects: the resource reservation is inadequate and fails to meet the desired application QoS level, or there is resource overbooking leading to large costs and resource under-utilization. Further, statistical bandwidth allocation formulas such as equivalent bandwidth computations (e.g., [18]) have been found to be overly pessimistic. For media applications, this is made even harder because the resource requirements are content-dependent and compressed media is highly bursty. Technical challenges aside, as alluded to in Table 2, the economic cost of resource reservation at this point in time is just too high for the end user. Thus it is likely that users will continue to select the cheaper best effort service that promotes higher levels of resource sharing.

Scalability An important issue with QoS provisioning is scalability concerns. Providing QoS on a per-application flow basis requires the maintenance of per-flow state information. In the core of the network, scaling to thousands of flows passing through a single switch/router is a technical challenge. A newer framework for QoS in IP networks is the Differentiated Services (DiffServ) architecture [7] which ameliorates the scaling problem by providing QoS to flow aggregates only in the network core. However, preliminary research on the mapping from aggregate QoS to individual per-flow QoS indicates that under some circumstances, there might be a significant performance difference between the individual QoS and the aggregate QoS [57]. Further, flows in the same traffic aggregate do not have performance isolation, which is likely to impact the QoS of a stream, particularly on short time scales.

Media Transport Without Reservations The downside of using non-dedicated resources is that congestion points can develop often and suddenly. For most Internet applications, congestion control is primarily exerted in the form of TCP congestion control, where one of the techniques is Additive Increase Multiplicative Decrease (AIMD), which halves the current TCP window in response to a lost packet [1]. For media applications, the primary form of congestion control has been rate control applied at the end system [56], sometimes using AIMD-like rate control [39]. The main assumption is that the applications under consideration are able to scale back their packet transmissions without loss of functionality. While this is certainly true for applications such as ftp, emerging applications such as streaming audio and video cannot comply with such harsh control without losing usability. However, if no control is exerted on such applications, then responsive TCP applications can suffer unfairly. There has been a host of new congestion control algorithms proposed recently [6, 54] that are less responsive than TCP but react in a manner that is TCP-friendly or TCP-compatible. TCP-friendly schemes such TCP-Friendly Rate Control (TFRC) [16] and Binomial Congestion Control [5] provide for a more smooth transmission rate as compared to TCP, but do not take into account the inherent dynamic source rate that the application may have. Providing a constant rate to an application is not adequate if the application requires a dynamic rate. End system rate control alone cannot manage congestion adequately, and Active Queue Management (AQM) techniques such as Random Early Detection (RED) [17] and its variations have been proposed that attempt to control queue lengths inside the network. In the above mentioned DiffServ environment, media flows can use an enhanced service class (e.g., expedited forwarding as in [2]) but the mapping between the user level quality and network parameters such as the token rate need to be studied further for different classes.

Content Distribution Networks A powerful approach to providing good application QoS is via content distribution networks or other content delivery mechanisms, including replication and caching at the network edges [40, 45, 52]. These approaches are successful to some extent in bypassing the network by priming caches close to the receiver ahead of the time it may be required. However, these techniques work best for on-demand usage and have yet to prove their merits for live or interactive multimedia applications.

3.3.2 Research Challenges

In this section, we discuss four open issues whose resolution is imperative to the success of the future vision.

Semantically Correlated Streams Many of the scenarios described earlier involve multiple semantically related media streams. Traditional QoS research has focused on single streams, connections, or sessions. The research space is significantly altered with multiple diverse but semantically related streams. The individual streams may originate from different points in the network and may traverse paths with very different QoS characteristics to the merge/manipulation points. There are several research problems involving stream coordination that emerge in this environment. How would one define QoS metrics and targets for multiple semantically related streams? Should the routing, scheduling and congestion control mechanisms be coordinated across the correlated streams? For instance, if the primary stream suffers a loss burst, it is not useful to transmit the temporally correlated packets from the value-added stream as well. Similarly, it makes sense to pick routing paths for the streams that have approximately the same path characteristics to avoid unnecessary buffering in some cases. Due to congestion, if the primary stream is rate controlled, it may not be useful to transmit the other stream at a higher rate. Can this coordination lead to more efficient bandwidth allocation and better dimensioning of playback buffers?

Short-term QoS Traditional QoS mechanisms have focused on long-term QoS metrics.Despite this, there has been relatively little work (e.g., [34]) on solving this important problem and offers a significant research space. To solve this problem, it is likely that innovative video coding techniques would need to be coupled with network mechanisms such as selective packet dropping and bandwidth renegotiation. Recently, codes resilient to bandwidth and delay variations have been developed [11], which need to be coupled to appropriate mechanisms inside the network to achieve acceptable short-term QoS levels.

Media-Aware Network Control As we push towards higher bandwidth and higher quality media streams, it becomes imperative to design media-aware network mechanisms that are also friendly to other traffic types. Traditional research pushes towards one goal or the other, and there needs to be a good balance between the two. Two big differences between media streams and other traffic types are that (1) media flows are of longer duration than many of the other dominant Internet traffic types (e.g., WWW traffic) and (2) media flows are adversely affected by sudden and large changes to their natural bandwidth demands. Current rate control mechanisms for media advocate TCP friendliness and either subject media traffic to harsh rate decreases or provide an “almost constant rate” or “slowly varying rate” that is fair to other TCP traffic [6, 16]. Although the second proposal is better than the first, it dictates a rate based on other shorter-lived traffic flows and does not work for compressed variable rate encoding. Towards this end, new measures of fairness are required, taking into account long-lived as well as short-lived flows. Perhaps media streams can be subjected to smaller degrees of rate control on a continuous basis (as a function of the queue length at the bottleneck node using early congestion notifications (ECN) [38] from routers) rather than only when packet losses are detected. A related idea is to link priority of network packets based on the duration of the flow, not just the type of application. Particularly in best-effort networks, users have no expectation of QoS levels when they initiate a session, and this fact could be exploited to ensure that existing long duration flows continue to receive the performance level they entered the network with, at the expense of new sessions. Current media players decrease the source rate in steps when congestion is detected, but have limited methods built in to increase their rates when either the congestion period has ended or alternate higher bandwidth paths are available. Probing for higher bandwidth paths and/or lightly loaded paths will be important to support high quality media flows. Further, methods to dynamically and collaboratively aggregate resources on the network to support high bandwidth flows is an open research problem.

Latency Masking An important hurdle in deploying global tele-presence and tele-experience applications is the large end-to-end latencies. However, interactive users can typically deal with longer latencies if the variability of the latency is relatively low. The design of low variance, almost constant delay pipes over shared infrastructure could perhaps be one step towards deploying such highly interactive applications.

3.4 Quality Assessment

Much research effort over the past several years has addressed the general problem of constructing scalable and relevant measurement infrastructures, network fault diagnosis methods, and fault prediction methods, particularly in the context of the Internet. However, conducting quality assessment for streaming media services, particularly from the end user perspective, has not been widely addressed by the network research community and remains a hard problem.

Several factors make measuring media more difficult than measuring, for example, Web transactions or file transfers. For one, media sessions tend to be of a longer duration than file transfer sessions or Web sessions (see [32], for example, for RealAudio traffic characteristics). Media files are larger than the typical data file on the Web. Most significantly though, media metrics are much more context-specific and temporal. For example, which packet is lost or where in the stream bandwidth throttling occurs is significant; 5% packet loss may be detrimental for a movie trailer clip, but may not produce any significant degradation in a newscast clip. Traditional metrics of network quality, such as average received bandwidth and average packet loss rate, may not be adequate for assessing media quality.

As an example, consider Figure 3, which illustrates the received bandwidth as measured by a commercially available media player application for an eleven-minute video clip streamed over a lossy connection. The received bandwidth is sampled at ten-second intervals during the stream; the dashed line shows the average bandwidth as measured over the entire stream, a common network quality metric. It is difficult to determine from this plot the end user's viewing experience: excellent, passable, unacceptable, etc. Also, the average received bandwidth tells little about the received

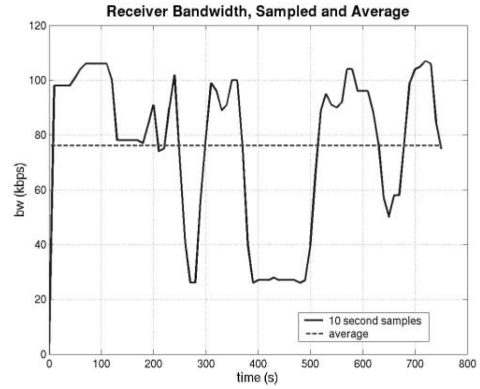

Figure 3: Received bandwidth vs. time, sampled at ten second intervals and averaged over the entire stream.

bandwidth at various points during the stream, failing to show the many peaks and valleys over the duration of the stream. In addition, it is not clear from the graph if the variation in sampled bandwidth is due to network loss or if the bandwidth variation is normal for this particular encoding of the stream. This example illustrates the need for observing both short-term and long-term metrics as well as the need to define appropriate metrics for the media services environment.

In this section, we explore the measurement and analysis spaces and discuss the repercussions for high-quality, multi-stream media delivery over both closed and open infrastructures.

3.4.1 Related Work

Several software solutions exist to measure and assess the quality of streaming media. For example, NetIQ Corporation's Chariot [35] uses specialized endpoints that send and receive synthetic traffic to emulate, in packet size and rate, real media servers and receivers. The tool cannot assess characteristics such as stream start-up delay or player stall that are of relevance to actual media clients, and has scalability issues. Broadstream's service [10] obtains assessment of actual end-user activities via a Java applet and provides aggregated usage and quality data to content-producers. It requires clients to communicate to a single central service location on the Internet. Streamcheck [49] operates in a similar manner, using synthetic clients built on top of existing players in a test-mode basis. A "subjective" measurement in the form of a letter grade is generated from a measure of initial connection and total stream buffering time as compared to total playback time. The test clients reside at "desktop" level; i.e., at the end of cable modem and DSL connections. Keynote [26] measures streaming media quality at locations along the Internet backbone (usually at peering points) using actual client applications, but it is limited to "test-mode" functionality only and does not measure "desktop" performance as seen by the end-user.

A variety of proposed techniques and existing products, such as Lariat [28], obtain information from the log files of media servers in an attempt to provide an assessment of stream quality. These approaches are proprietary, not open to general integration into a complete assessment service, do not provide for testing and fault isolation independent of end-user cooperation, and cannot accurately assess quality as seen at the client.

Real-time Transport Protocol (RTP) [42] for media streaming has a related control protocol, RTCP, for communication from the client about the quality of received streams. This protocol suffers from severe scalability issues, and is generally disabled in real production deployments of the media servers and players that utilize standard RTP. The commercially popular servers and players do not fully support standard RTP.

Some proprietary media players (such as RealPlayer and Windows Media Player) support agents that provide quality feedback in an attempt to adjust stream server characteristics. Principally these agents work in conjunction with the server to enable transmission and reception of a lower bit-rate encoding of the media stream. To provide log file-based reports of quality assessment, they utilize TCP-based streams with the attendant scalability problems.

Some work has focused on instrumenting client-side media players in order to assess the quality of received video streams. For example, [53] studies user-perceived video quality using an instrumented version of RealPlayer. The application is completely user-controlled, and requires the user to manually rate his or her perception of video quality on a numeric scale once a clip has played. Similarly, [29] describes a measurement study of streaming video to dial-up modem clients that utilizes instrumented MPEG-4 receivers.

The mapping between objective measurements and subjective video quality is attempted in [2] for MPEG-1 and Windows Media clips over a DiffServ network. Their assessment system is based on earlier work described in [55]. The mapping works by measuring both the sent and received video streams to extract key parameters; these parameters are then compared to compile a subjective quality score.

The general problem of developing large-scale network assessment and measurement architectures have been widely studied. A survey of existing efforts and a description of the problem space can be found in [33]. Examples of large-scale network measurement architectures include [25], [37], [36], and [48]. Inter-domain assessment and troubleshooting have also been addressed previously. Thaler and Ravishankar [51] describe a troubleshooting methodology for coordinating problem diagnosis across multiple management domains. Baek *et al.* [4] describe a multi-domain measurement system appropriate for SLA use.

3.4.2 Research Challenges

Several areas within the space of end-to-end streaming media quality assessment remain open problems. We discuss them here.

Using Objective Metrics to Quantify Subjective Expectations of Quality End users must have the ability to interact with the network, both automatically and manually, to monitor their level of received service quality. To do so requires mapping subjective ideas of media quality, such as picture and sound clarity, to objective metrics such as frame rate, packet loss, and jitter. One such solution is given in [55] and [2], but it requires correlating measurements on both the sender and receiver sides. A more useful solution would assess received quality by taking a select set of measurements from both the receiver and the network and using these measurements to deduce the user's received quality.

A significant challenge is to derive assessments from these collected metrics with sufficient information to permit either the adaptation of quality of service parameters or the correct diagnostic action. For example, if an end user's quality suffers because of inadequate bandwidth for a particular media stream, how can the network and/or media source address the problem in real time? Can the network find an alternate, less-congested path with more available bandwidth? Or can the source reduce the rate at which it streams, by sending fewer encoded layers? Achieving this goal requires developing new test tools that can interact with client-side players and existing network measurement tools.

We claim that the recipient of a media stream is obviously the best authority to assess its quality. Humans, however, cannot be relied upon to provide consistent, timely information. Thus it becomes necessary to make assessments of streaming media quality within the receiving client. For such a system to work necessitates the cooperation of end users as well as service providers. We must ensure that the system is secure, in that it protects both data integrity and the privacy of the end-users and network providers.

Multi-Domain Operation An end-to-end assessment system must operate among domains rather than within a single domain. This necessitates inter-domain cooperation among service providers. Inter-domain cooperation is a difficult problem because historically service providers have been reluctant to share measurement data among themselves. Also, the correlation of measurements/assessment across domains is a non-trivial problem. Even if providers were open with their data, it is hard to construct an end-to-end picture of network performance. There is an effort [23] to standardize the collection and sharing of data among network service providers, but the problem remains.

The problem is made more urgent when applied to a media services environment. Because of the longer duration of media sessions as compared to much of the traffic traversing the Internet today, and the amount of data transferred per session, it is important that a clear end-to-end picture emerge for proper fault diagnosis and troubleshooting. In addition, cooperating parties still need to determine an adequate set of metrics with which to evaluate the quality of media sessions, for the reasons described in Section 3.4.

A wide range of parties value the assessment of streaming media quality for diverse reasons, as indicated in Table 3. When a number of entities have a need for common information, an open standard interface to such information is ideal. End users and providers both benefit, as both sets of consumers are able to get a clearer picture of network performance and where problems lay.

System Scalability The end-to-end quality assessment system we envision will have difficult scalability issues. The system will have a very large number of measurement points, as each end user is a potential measurement point. Because of the complexities involved in evaluating streaming media quality, as discussed in Section 3.4, each measurement point is capable of generating a significant amount of measurement data. Additionally, we envision a huge number of media sources which will lead to a general increase in the amount of media traffic to monitor.

An important challenge is to determine how to scale this measurement data. The assessment system needs to decide from which monitoring points to collect data, on what timescales, and which metrics to measure and return to assessment entities. Measurements from different end-user monitoring points must be correlated in order to produce a relevant picture of streaming media quality. Data sampling and data mining are possible solutions here. The assessment system will also require some level of "intelligence," or automation, that allows it to switch between passive measurements for normal system monitoring and active measurements for troubleshooting, prediction, and proactive monitoring without human intervention.

4 Conclusions

We have taken the position that there will be a vast number of content creators scattered across the Internet creating copyleft media. That the content is widely distributed and of temporal value has forced us to look at the discovery problem. That the content is copyleft allows us to be creative in the ways in which content is manipulated to create new works, perhaps by merging and mixing multiple streams. In a context where multiple streams need to be coordinated we have considered multi-stream quality of service. We have also reviewed subjective quality assessment which will play a part in the provisioning of quality of service.

References

[1] M. Allman, V. Paxson, and W. Stevens, "TCP Congestion Control," *IETF RFC 2581*, April 1999.

[2] W. Ashmawi, R. Guerin, S. Wolf, and M. H. Pinson, "On the Impact of Policing and Rate Guarantees in Diff-Serv Networks: A Video Streaming Application Perspective," *Proceedings of the ACM SIGCOMM'01*, San Diego, CA, August 2001, pp. 83-95.

[3] J. Ayars, *et al.* (editors), "Synchronized Multimedia Integration Language (SMIL 2.0)," *W3 Recommendation*, April 2001, `http://www.w3.org/TR/smil20`.

[4] J. W. Baek, J. T. Park, and D. Seo, "End-to-End Internet/Intranet Service Management in Multi-Domain Environment Using SLA Concept," *Proceedings of the IEEE/IFIP NOMS'00*, Honolulu, HI, April 2000, pp. 989-990.

[5] D. Bansal and H. Balakrishnan, "Binomial Congestion Control Algorithms," *Proceedings of the IEEE INFOCOM'01*, Anchorage, AK, April 2001, pp. 631-640.

[6] D. Bansal, H. Balakrishnan, S. Floyd, and S. Shenker, "Dynamic Behavior of Slowly-Responsive Congestion Control Algorithms," *Proceedings of the ACM SIGCOMM'01*, San Diego, CA, August 2001, pp. 263-274.

[7] S. Blake, D. Black, M. Carlson, E. Dazies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services," *IETF RFC 2475*, December 1998.

[8] Boardwatch Magazine, "Directory of Internet Service Providers," Spring 2001, Penton.

[9] J. Brassil and H. Schulzrinne, "Enhancing Internet Streaming Media with Cueing Protocols," *Proceedings of the IEEE INFOCOM'01*, Anchorage, AK, April 2001, pp. 95-103.

[10] Broadstream, `http://www.broadstream.com`.

[11] Y.-C. Chang and D. G. Messerschmitt, "Improving Network Video Quality with Delay Cognizant Video Coding," *Proceedings of the IEEE IPCIP'98*, Chicago, IL, October 1998, pp. 27-31.

[12] W. S. Ciciora, "Cable Modem Traffic Jam," *IEEE Spectrum*, June 2001 pp. 48-53.

[13] CWI Media Interchange Format, `http://www.cwi.nl`.

[14] M. R. Delco, "Production Quality Internet Television," *Berkeley Multimedia Research Center Technical Report*, TR 161, August 2001.

Table 3: Consumer-Metrics matrix.

Consumers of Quality Assessment	Uses
Content Owners	Did the viewing stop because content was uninteresting or because it was unwatchable (poor received quality)? What content is watched, and for how long? What content has been unwatched for some time? What can be said about viewer behavior (pause, skip, rewind) for the content?
End Users	To whom do I complain when my reception is unwatchable or inaudible? Am I getting the streaming media quality I paid for? What information do I need when I call customer support?
Network Operators, Media Service Providers, Media Distribution Networks	Is my operation supporting service level agreements for streaming media service? What causes poor quality for "this user's" media stream? Is the trend of my operation's behavior headed towards unsatisfactory streaming media service? How do I determine my operational parameters that will improve streaming media service? How do I test my operation in advance of need, especially in advance of a large scale streaming media event?
Product Developers	Will this improvement in my product (e.g. encoder, transport mechanism, decoder, etc.) result in improved media quality as I expect?
Management System Providers	What are the key alert events for streaming media quality? What methods can be used to "drill down" and correlate or diagnose streaming media faults, all the way to the receiver when possible?

[15] Enhanced Content Specification, "Advanced Television Enhancement Forum," `http://www.atvef.com/library/ATVEF%20spec1_1a_Final.pdf`.

[16] S. Floyd, M. Handley, J. Padhye, and J. Widmer, "Equation-Based Congestion Control for Unicast Applications," *Proceedings of the ACM SIGCOMM'00*, Stockholm, Sweden, August 2000, pp. 43-56.

[17] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, August 1993, pp. 397-413.

[18] R. Guerin, H. Ahmadi, and M. Naghshineh, "Equivalent Bandwidth and Its Application to Bandwidth Allocation in High-Speed Networks," *IEEE Journal on Selected Areas in Communications*, vol. 9, no. 7, September 1991, pp. 968-981.

[19] M. Handley and V. Jacobson, "SDP: Session Description Protocol," *IETF RFC 2327*, April 1998.

[20] M. Handley, C. Perkins, and E. Whelan, "Session Announcement Protocol," *IETF RFC 2974*, October 2000.

[21] M. Handley, H. Schulzrinne, E. Schooler, and J. Rosenberg, "SIP: Session Initiation Protocol," *IETF RFC 2543*, March 1999.

[22] Internet2, `http://www.internet2.org`.

[23] IP Performance Metrics Working Group, IETF, `http://www.ietf.org/html.charters/ippm-charter.html`.

[24] Y. Ishibashi and S. Tasaka, "A Comparative Survey of Synchronization Algorithms for Continuous Media in Network Environments," *Proceedings of the IEEE LCN'00*, Tampa, FL, November 2000, pp. 337-348.

[25] S. Kalidindi and M. Zekauskas, "Surveyor: An Infrastructure for Internet Performance Measurements," *Proceedings of the ISOC INET'99*, San Jose, CA, June 1999.

[26] Keynote Streaming Perspective, `http://www.keynote.com/solutions/html/streaming_perspective1.html`.

[27] D. Kutscher, J. Ott, and C. Bormann, "Session Description and Capability Negotiation," *IETF Internet Draft*, `draft-ietf-mmusic-sdpng-03.txt`, November 2001, work in progress.

[28] Lariat Media Reports, `http://www.lariat.com`.

[29] D. Loguinov and H. Radha, "Measurement Study of Low-Bitrate Internet Video Streaming," *Proceedings of the ACM SIGCOMM Internet Measurement Workshop 2001*, San Francisco, CA, November 2001.

[30] S. McCreary and K. Claffy, "Trends in Wide Area IP Traffic Patterns: A View from Ames Internet," *Proceedings of the 13th ITC Specialist Seminar on Internet Traffic Measurement and Modelling*, Monterey, CA, September 2000.

[31] MeasureCast Internet Radio Listening Index, MeasureCast, Inc., February 10, 2002, `http://www.measurecast.com/news/weeklyindex.html`.

[32] A. Mena and J. Heidemann, "An Empirical Study of Internet Audio Traffic," *Proceedings of the IEEE INFOCOM'00*, Tel Aviv, Israel, March 2000, pp. 101-110.

[33] M. Murray and K. Claffy, "Measuring the Immeasurable: Global Internet Measurement Infrastructure", *Proceedings of PAM'01*, Amsterdam, Netherlands, April 2001.

[34] R. Nagarajan, J. Kurose, D. Towsley, "Finite-Horizon Statistical Quality-of-Service Measures for High-Speed Networks," *Journal of High Speed Networks*, vol. 3, no. 4, December 1994, pp. 351-374.

[35] NetIQ's Chariot Software, `http://www.netiq.com`.

[36] NLANR, Network Analysis Infrastructure, `http://moat.nlanr.net/nai/`.

[37] V. Paxson, J. Mahdavi, A. Adams, and M. Mathis, "An Architecture for Large-Scale Internet Measurement," *IEEE Communications Magazine*, vol. 36, no. 8, August 1998, pp. 48-54.

[38] K. Ramakrishnan and S. Floyd, "A Proposal to Add Explicit Congestion Notification (ECN) to IP," *IETF RFC 2481*, January 1999.

[39] R. Rejaie, M. Handley, and D. Estrin, "RAP: An End-to-End Rate-Based Congestion Control Mechanism for Realtime Streams in the Internet," *Proceedings of the IEEE INFOCOM'99*, New York, NY, March 1999, pp. 1137-1145.

[40] R. Rejaie, H. Yu, M. Handley, and D. Estrin, "Multimedia Proxy Caching Mechanism for Quality Adaptive Streaming Applications in the Internet," *Proceedings of the IEEE INFOCOM'00*, Tel Aviv, Israel, March 2000, pp. 980-989.

[41] S. Romero, "Once Bright Future of Optical Fiber Dims," *NY Times Article*, June 18, 2001.

[42] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," *IETF RFC 1889*, January 1996.

[43] H. Schulzrinne, A. Rao, and R. Lanphier, "Real Time Streaming Protocol (RTSP)," *IETF RFC 2326*, April 1998.

[44] SCIFI.com, Mystery Science Theater 3000, `http://www.scifi.com/mst3000`.

[45] S. Sen, J. Rexford, and D. Towsley, "Proxy Prefix Caching for Multimedia Streams," *Proceedings of the IEEE INFOCOM'99*, New York, NY, March 1999, pp. 1310-1319.

[46] C. Shih, "Drawing the line," *Mountain View Voice Article*, `http://www.mv-voice.com/morgue/2002/2002_02_08.sportvision.html`, February 2002.

[47] Spin The Bottle Dot Com, Pop Up Video, `http://www.popupvideo.com`.

[48] Stanford Linear Accelerator Center's IEPM Project, `http://www-iepm.slac.stanford.edu`.

[49] Streamcheck, `http://www.streamcheck.com`.

[50] N. R. Sturtevant, N. Tang, and L. Zhang, "The Information Discovery Graph: Towards a Scalable Multimedia Resource Directory," *Proceedings of the IEEE WIAPP'99*, San Jose, CA, July 1999, pp. 72-79.

[51] D. Thaler and C. V. Ravishankar, "An Architecture for Inter-Domain Troubleshooting," *Proceedings of the IEEE ICCCN'97*, Las Vegas, NV, September 1997, pp. 516-523.

[52] J. Wang, "A Survey of Web Caching Schemes for the Internet," *ACM SIGCOMM Computer Communications Review*, vol. 29, no. 5, October 1999, pp. 36-46.

[53] Y. Wang, M. Claypool, and Z. Zuo, "An Empirical Study of RealVideo Performance Across the Internet," *Proceedings of the ACM SIGCOMM Internet Measurement Workshop 2001*, San Francisco, CA, November 2001.

[54] J. Widmer, R. Denda, and M. Mauve, "A Survey on TCP-Friendly Congestion Control," *IEEE Network*, vol. 15, no. 3, May/June 2001, pp. 28-37.

[55] S. Wolf and M. H. Pinson, "Spatial-Temporal Distortion Metrics for In-Service Quality Monitoring of Any Digital Video System," *Proceedings of the SPIE VVDC'99*, Boston, MA, September 1999.

[56] D. Wu, Y. T. Hou, W. Zhu, Y.-Q. Zhang, and J. M. Peha, "Streaming Video over the Internet: Approaches and Directions," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, March 2001, pp. 282-300.

[57] Y. Xu and R. Guerin, "Individual QoS Versus Aggregate QoS: A Loss Performance Study," to appear in *Proceedings of the IEEE INFOCOM'02*, New York, NY, June 2002.

[58] R. Zakon, "Hobbes' Internet Timeline," `http://www.zakon.org/robert/internet/timeline/`.