

High-Quality Video Services for the Access Grid

Sangwoo Han and JongWon Kim

Networked Media Lab., Dept. of Information and Communications
Gwangju Institute of Science & Technology (GIST), Gwangju, 500-712 Korea
{*swhan, jongwon*}@gist.ac.kr

Abstract

The Access GRID (AG), commonly categorized as one of three major GRID areas, refers to an emerging advanced collaboration environment, with which multi-party group of people from remote can collaborate over high-performance networks by using conferencing as well as semi-immersive visualization capabilities. In this paper, we introduce an effort to improve the AG by adding support for high-quality media such as digital video (DV) and high definition video (HDV). An improved media service named as 'Extended Video Service' for existing AGTk (AG toolkit) is designed and implemented to support multi-party delivery of high-quality video over multicast-enabled IP networks. The enhanced user experience is verified by demonstrating the proposed service over international R&D networks in recent AG Retreat 2004.

Keywords: Access GRID, next-generation video conferencing, high-quality video over IP multicast, and advanced collaboration environment.

1. Introduction

The Access GRID (AG), targeting advanced collaboration environment over high-performance networks, is commonly categorized as one of three major GRID areas (i.e., computing, data, and access GRID) [1]. As shown in Fig. 1, the AG provides group-to-group interactions with an ensemble of resources including multimedia large-format displays, interactive conferencing tools, and interfaces to Grid middleware [2]. It also assists group collaboration by building persistent virtual venues [3]. The AG project aims to keep up with the latest trends in remote collaboration thanks to recent international joint efforts for the AG. Thus, at present, thousands of certificates have been issued for the AG nodes and more than 150 institutes worldwide now use the AG for their remote meetings, distributed education, and other e-science applications.

With the advent of Gbps-range R&D (research and development) networks, it is now possible to accommodate the increasing demand toward high-quality immersive media services that requires tens of Mbps bandwidth for a single media flow. However, with current AGTk (AG toolkit) 2.3 [4], the video services for the AG still relies on outdated VIC (Video Conferencing

Tool) [5], which originated from MBONE (multicast backbone) tools. The VIC adopts a modified H.261 codec to support CIF (common intermediate format) sized video around 300Kbps rate. The video quality of VIC falls short of user expectation and it cannot cover properly large-scale view of involved people as well as detailed close-up of meeting documents.



Fig. 1. The AG remote collaboration environment [4].

To solve this weakness, we propose an 'Extended Video Service' for the AGTk to enable high-quality video such as DV¹ and HDV² by extending existing 'Video {Producer, Consumer} Service' tied heavily with outdated VIC. The proposed 'Extended Video Service' takes charge of video session management based on SAP (session announcement protocol) defined by RFC 2974 [6] and SDP (session description protocol) defined by RFC 2327 [7]. In addition, the proposed 'Extended Video Service' can employ versatile media tools for various video codecs available.

In the followings, we will discuss about the proposed 'Extended Video Service' for the AG. First, Section 2 presents the architecture of the AGTk and the applicable high-quality video codecs. Then, we point out technical issues to be considered to enable a high-quality video service for the AG. Next, in Section 3, we detail the proposed framework and implementation of 'Extended Video Service'. System setup and evaluation result will follow in Section 4. Finally we conclude this paper in Section 5.

¹ DV (digital video): around 30Mbps 720x480 DV encoding.

² HDV (high-definition digital video): around 20Mbps MPEG-2-encoded 1280x720 30progressive.

2. Access GRID and High-Quality Video

2.1 The Architecture of the AGTk

As shown in Fig. 2, current architecture of the AGTk 2.3 consists of the Venue Server and the Venue Client. The Venue Server provides a central infrastructure for doing advanced collaboration: persistent data repository, access control based on authentication and authorization, synchronized venue state for coherent distributed experience, and capability brokering service³. The Venue Server can contain number of virtual venues, which provide virtual spaces among collaborators. By logging into one of the venues and sharing services, users can communicate with others. The Venue Management is an administrative tool used to create and maintain venues located on the Venue Server. It includes information about present venues on the server, server administrators, and the encryption type used for media communication. The Network Services manipulate streams, possibly converting a stream to a different format, or saving a stream for later playback.

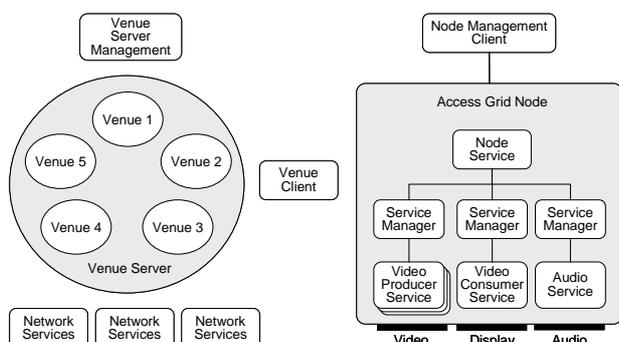


Fig. 2. The architecture of the AGTk 2.3.

The Venue Client provides end-user an interface to the virtual venues. With the Venue Client interface, end-users share data, services, and applications. They also exchange text-based messages with other participants. To enable collaboration, the Venue Client software manages the resources of the AG node such as microphones, cameras, and so forth. It also gives users an ability to travel from one venue to another by using a navigation interface for the virtual venue space.

An AG node consists of a node service, one or more service managers, and one or more services. A node configuration, pictured in Fig. 2, uses three machines; one for video creation, one for video display, and one responsible for audio. The services, in this case, are used to produce and receive audio and video. Each machine runs service managers communicating with services on that specific machine. Note that the service managers are controlled by the node service, which can run on any machine. Default services used by the Venue

³ The venue server provides a capability brokering service because all users do not have the same capabilities (meaning network, audio, video, application, and so forth).

Client are based on VIC for video and RAT for audio, respectively.

In Fig. 3, the default configuration of a typical AG node is illustrated. A display computer, connected to three projectors, provides large-scale display capability. In this machine, a 'Video Consumer Service' processes the reception of multiple remote-site video streams and the display of them. Each 'Video Producer Service' in a video capture computer captures a video source and places the packetized video on the multicast-enabled network. Similarly an audio capture computer performs bi-directional audio streaming with the RAT (Robust Audio Tool) and is coupled with an echo canceller manipulated by a control computer.

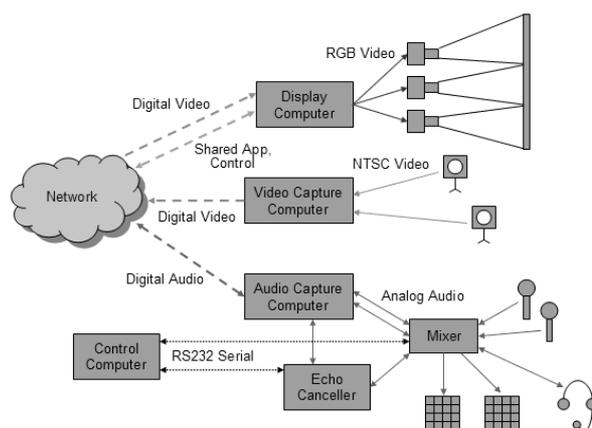


Fig. 3. Typical configuration for a room-based AG node.

2.2 High-Quality Video Codecs and Tools

Digital Video (DV): DV Format is specially optimized to support the recording of digital A/V data using helical-scan magnetic tapes [8]. The DV format adopts the intra-frame only DCT encoding compression for video. The DV encoding generates a structured stream in three-level hierarchy. A picture frame is divided into several DCT super blocks, shaped rectangle or clipped-rectangle. Each super block contains 27 macro blocks, which are compounds of 4 or 6 DCT block units. A video frame in the DV-format stream is divided into several DIF (Digital Interface Format) sequences. A DIF sequence contains 150 chunks of 80-byte DIF blocks. Each DIF block, the primitive unit for the DV stream, contains a 3-byte ID header that specifies the block type and its position in the DIF sequence. Five types of DIF blocks are defined: DIF sequence header, subcode, video auxiliary information (VAUX), audio, and video.

Digital Video Transport System (DVTS): The DVTS software supports real-time streaming of DV stream over unicast/multicast IP networks [9]. The DVTS receives DV frames from a consumer DV camcorder via IEEE1394 interface and encapsulates them to IP/UDP/RTP packets. By default, the DVTS consumes around 30 Mbps bandwidth. However, the DVTS can reduce frame rates by dropping frames and thus it can be

utilized to match the available bandwidth of underlying networks⁴.

High Definition Video (HDV): HDV stands for high-definition upgrade of DV format to cover the HD recording of MPEG2-encoded (in MP@HL) video in mini DV cassette tapes [10]. It provides extremely high-quality video with better sharpness, detail, and clarity by supporting a range of HD picture formats distinguished by frame size and rate, aspect ratio, and scanning. Typically it can support both 1280x720 progressive and 1920x1080 interlaced HD videos. The HDV format supports the storage of MPEG2 TS (transport stream) data that consumes around 20Mbps bandwidth. The MPEG2-TS is composed of 188-byte fixed-size segments. Each MPEG2-TS segment has a header with PID (Packet ID) that identifies the type of component stream (e.g., audio, video, and others).

VideoLAN: The VideoLAN stands for an open-source software project to provide multimedia streaming of MPEG-1, MPEG-2, MPEG-4, DivX files, DVDs, digital satellite channels, digital terrestrial television channels, and live videos over high-bandwidth IPv4/IPv6 networks [11]. The VideoLAN software package is capable of supporting both unicasting and multicasting under most operating systems [11]. It consists of VLS (VideoLAN Server) and VLC (VideoLAN Client) softwares. Note however that the most popular VLC, originally developed for client, is actually supporting both the sender and the receiver roles. The used version of VLC, VLC 0.7.2, supports the announcement of multimedia session based on the SAP and SDP standards. The mini-SAP-server component of VLC sends announces about the multicast session over the network, and the VLC receives the announcement and automatically adds the session information to its playlist. The MPEG2-TS streaming capability of VLC allows us to interface the live HD video from HDV camcorder (via IEEE 1394). In this case, the VLC can directly transport 19.34Mbps MPEG2-TS. Also, with the transcoding capability of VLC, it is possible to deliver the transcoded video at less than 3Mbps.



Fig. 4. The images of HDV, DV and H.261.

⁴ In this case, the DVTS sender should be modified to take out the audio data from the discarded frames and sends them to the DVTS receiver(s).

In Fig. 4, we compare the size and quality of videos such as HDV, DV, and H.261 (i.e., VIC). From the right to the left, 1280x720 HDV video, 720x480 DV video, and 325x288 H.261 video are shown on the projected screen. We can feel the merit of using high-quality video with the AG.

2.3 Challenges in High-Quality Video Services for the AG

The first technical issue to enable high-quality video services for the AG is the method to allocate the multicast address. As mentioned earlier, the VIC shares a multicast address to deliver one or more video streams simultaneously. Each video is identified by the SSRC (synchronizing source identifier) of RTP packet header. But this approach cannot isolate video streams from selected senders and the amount of occupied bandwidth just depends on the number of video streams. That is, regardless of how many video senders are transmitting the video over the commonly shared multicast address, the video tool on the receiver side should have to receive all streams⁵. It may be acceptable when each sender sends 300Kbps traffic. However, when the bandwidth-demanding HD video traffic (consumes 20~30Mbps per stream) is used, total required bandwidth can easily go beyond the limit of typical receivers (e.g., 100Mbps). To solve the weakness, the video tool needs to receive video streams selectively. Furthermore, the video quality controls such as rate adaptation and transcoding are needed to reduce the required network bandwidth.

The second issue is to generalize the video services for the AG so that they can easily accommodate selected variety of open-source video applications (i.e., codecs). In order to support video applications with different codec capabilities and proprietary interfaces, the AGTk should be able to match the diversity of them. To handle this, the video interface for the AG needs to be generalized while keeping the simplicity of current interface as much as possible.

3. The Proposed Framework for the AG Video Enhancement

The proposed 'Extended Video Service' is designed to address the two challenges: resolve multicast address allocation and support versatile media tools. Fig. 5 shows the modified architecture of the AG node, where 'Extended Video {Producer, Consumer} Service' are added by enhancing the existing 'Video {Producer, Consumer} Service'.

To enable selective reception of multicast video streams, the ExtendedVideoProducerService (EVPS) utilizes two multicast addresses. A shared multicast address, obtained from the Venue Server, is used to announce the

⁵ This is true for currently deployed IP multicast based on any source multicast (ASM). When the recent SSM (source specific multicast) becomes available, the situation will change.

session information. Also, another unique multicast address is created by the EVPS itself through a multicast address allocation call to the Venue Server. This unique multicast address is used to transport a video stream for the EVPS. With these two kinds of multicast addresses ready, the EVPS operates its SAP server to announce its session information. The announcement over the shared announcement-only multicast address includes session title, unique video multicast address, video codec, name of video application, and running arguments for video application (including required bandwidth). Then, the EVPS starts to stream the video destined to the unique multicast address by executing the video application configured in advance.

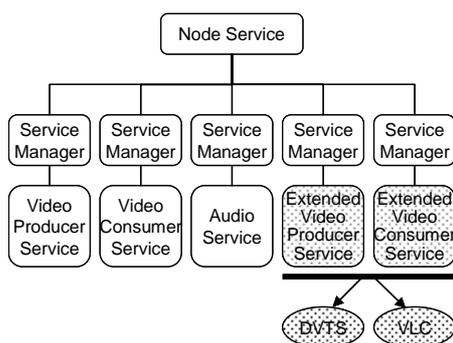


Fig. 5. The AG node with the proposed modification.

The ExtendedVideoConsumerService (EVCS) takes the announcement-only multicast address from the Venue Server. This service receives and parses all SAP packets, broadcasted by the EVPS, via the announcement-only multicast address. The obtained session information is displayed using the graphical user interface (GUI). Then a user selectively decides to join video streams considering their bandwidth and computational budget.

To employ versatile media tools, the video interface recognizes the properties of them. The DVTS and the VLC both require different option commands. To handle this, the video interface enables the EVPS/EVCS to understand the proprietary option commands. For instance, the VLC needs to get the friendly name of video devices to acquire video source and whether transcoding mode is required or not. Additionally, the DVTS demands the level of frame discarding. In the property dialog box of EVPS, users can configure the media properties manually.

EVPS Implementation: This service is built by extending the existing ‘Video Producer Service’ and produces high-quality video service. It takes care of following jobs: building the session information, announcing it, and executing the media tool (application). The EVPS is composed of two classes: EVPS class and SAPPacket class. The EVPS class interacts with the node service by handling events such as service start, service stop, and service configuration via SOAP (simple object access protocol). The EVPS

class receives the event and performs the corresponding handler. In addition, the EVPS class defines the properties of video tool such as video codec, frame rates, transcoding mode, time-to-live (TTL), and so forth.

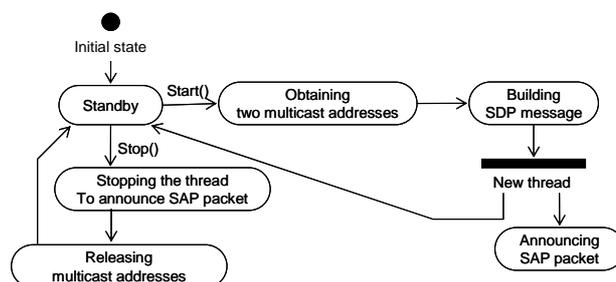


Fig. 6. State chart diagram of the EVPS.

From Fig. 6, we can consider two scenarios: service start and service stop. When the node service lets the EVPS class start, the Start() of the EVPS class is invoked and obtains two multicast addresses as explained. Next, the SAP message is built with SDP syntax to announce the session information. Here, the EVPS class uses session name, owner name, video codec, video tools, video bandwidth, and camera information to describe the AG-specific media information. The example SDP message is shown in Table 1. It means that a user at ‘GIST’ with IP address 203.239.52.123 intends to do media streaming with multicast address (224.1.3.121 with 127 TTL). For this, the DVTS software with DV codec is sending RTP-compatible video to port 50624 based on the RTP A/V profile (with static payload number 111). Note that the required network bandwidth is set to 30Mbps. It also informs that a camera interface named Microsoft DV/VCR Camcorder is used to get the video.

Table 1. An example of SDP message.

```

v=0
o='GIST' 31203218 31203218 203.239.52.123
s=DVTS
i=Microsoft DV/VCR Camcorder
e=ag-gist@gist.ac.kr
c=IN IP4 224.1.3.121/127
b=AS:30000
m=video 50624 RTP/AVP 111
  
```

This SDP message is packetized to a SAP packet. Defined by RFC2974, the SAP standard provides the features of session management: adding session, deleting session, and modifying session. Also, the SAP standard defines a way to authenticate the multimedia session, which is not utilized in current implementation yet. We implement the SAPPacket class to announce the SAP packets. By invoking a new thread, the EVPS periodically announces the SAP packets via announcement-only multicast address. When the node service runs the Stop() of the EVPS, the service sends the stop message, and the field of message type is set 1

to stop the session. Then it releases the allocated video multicast address.

EVCS Implementation: The EVCS is the counterpart of the EVPS, extended by the existing ‘Video Consumer Service’. As depicted in Fig. 7, there are three main jobs of the EVCS: handling the received SAP packets, checking sessions, and interacting with the node service. When the Venue Client is connected to the virtual venue, the node service invokes Start() of the EVPS to open a SAP announcement channel. After completing the initial job, the EVPS creates two threads to receive SAP packets and to check sessions.

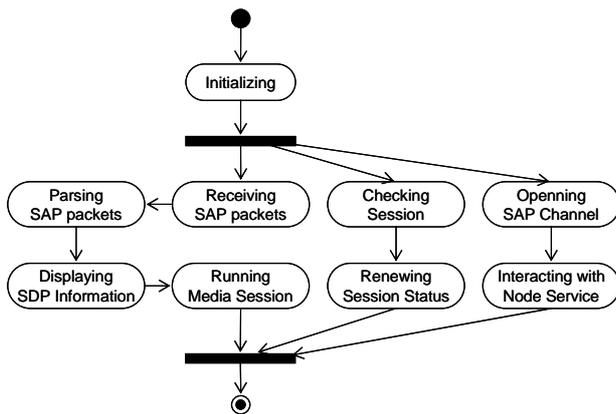


Fig. 7. Activity diagram of the EVCS.

The first thread of EVCS receives SAP packets periodically via the announcement-only multicast address. When the SAP announcement packet (the message type of SAP header is 0) is received, it is parsed. The obtained session information is added to the session list and displayed on the GUI. When a user selects the reception of a certain session, the corresponding media tool is launched, and its process identifier is added to the process list. On the other hand, when the SAP stop packet (the message type of SAP header is 0) is received, the corresponding session is terminated. Every session is associated with timeout mechanism to detect implicit session termination. In our implementation, the EVPS announces the SAP packet every second and the timeout is set to 60 seconds.

The second thread monitors the status of SAP sessions and executes the video applications. First, this thread monitors actually takes care of SAP timeout. That is, when the timeout happens, it removes the SAP session from session list and terminates the video application by force. This thread also monitors whether the processes of video tools are executing or not. If a user terminates a certain process of video tool, the process monitor changes the process status of GUI about the process.

Invoked by the node service, Stop() ends both threads and closes the SAP channel. By this call, all SAP sessions and the media applications are terminated.

4. System Setup and Evaluation

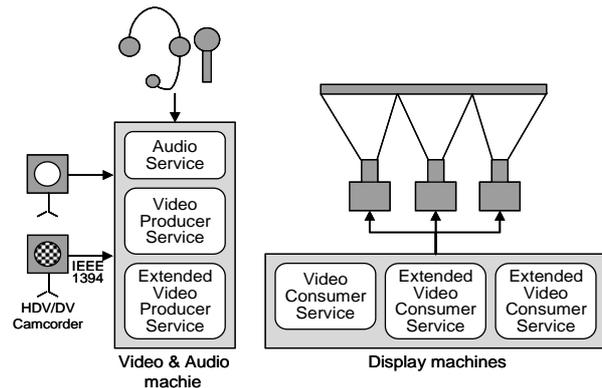


Fig. 8. A possible configuration of the AG for the high-quality video service.

In Fig. 8, one possible configuration method is illustrated to convert a room-based AG node for high-quality video service. A video & audio capture machine has a role of both a video capture computer and an audio capture computer. The video & audio capture machine receives video originated from a HDV/DV camcorder via the IEEE1394 interface; instead of using the audio from the HDV/DV camcorder, separate microphones managed by the audio service cover the audio part in order to make echo cancellation easy. The EVPS consumes lower computing power in comparison with the EVCS. Therefore a machine can perform the EVPS along with the VideoProducerService and the AudioService. Resource-demanded instances of EVPS are run with several display machines.

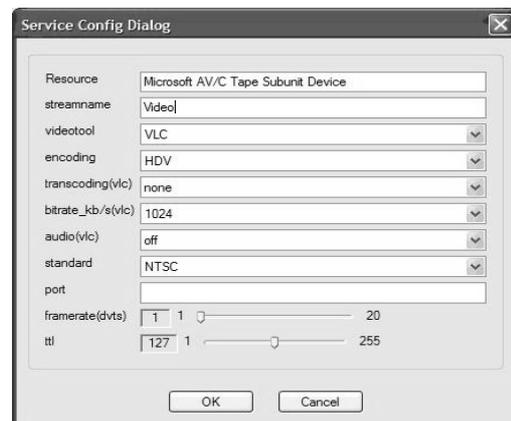


Fig. 9. The EVPS configuration.

Before the Venue Client connects the Venue Server, the service configuration should be done in advance. In this step, the applicable video codec (and video tool) is selected, and the related options are configured such as transcoding, frame rates and so forth. In Fig. 9, an example shows the properties set for the EVPS. The selected video tool is the VLC and the applicable video codec is HDV. The audio originated from the VLC is turned off. In Fig. 10, the snapshot of the proposed service is depicted. The GUI of the EVCS is placed on

the right-lower side. An operator selects the desired session and executes the media tool. The executed VLC media application is placed on the right-upper side and display larger video in comparison with the VIC placed on the left-upper side.



Fig. 10. The snapshot of the Extended Video Service.

To evaluate the proposed service, we have demonstrated the AG with high-quality video support between Ryerson University in Canada and GIST in Korea at AG Retreat 2004. The demonstration situation is depicted in Fig. 11. For the AG node hardware, we use two Dell GX270s equipped with 2.8 GHz CPU plus 512 MB main memory and a JVC GR-HD1 camcorder to enable DV and HDV-encoding in real-time. Participants in each node can communicate with each other using high-quality video with the AG, as shown in Fig. 12.

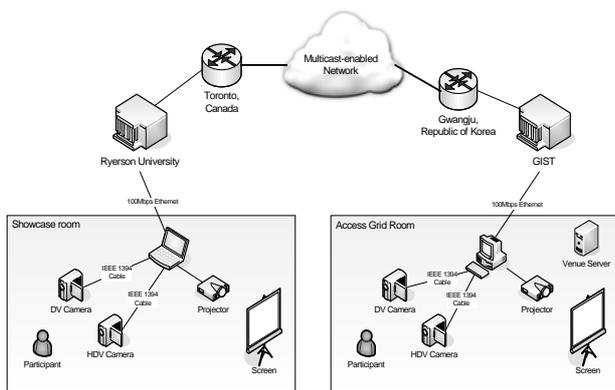


Fig. 11. The demonstration setup of AG Retreat 2004.



Fig. 12. The demonstration in the AG Retreat 2004.

5. Conclusions and Future Works

In this paper, we have proposed the high-quality video service for the AG to enable DV and HDV delivery. The proposed system provides the improved video service with the enhanced method of multicast address allocation and the video interface to employ versatile video codecs. The improved AG by adding support for high-quality media aids an importance class of collaboration services that properly cover large-scale view of collaboration space as well as detailed close-up of meeting.

For a successful continuous media delivery, we should consider how to provide high-quality media service for the AG nodes only with limited bandwidth and computing power. To address this problem, we are currently investigating a network-adaptive version of media transport so that it can be tailored to the given network/system constraints.

Acknowledgements

This research was supported in part by the KISTI (Korea Institute of Science and Technology Information) and in part by GIST (Gwangju Institute of Science and Technology). We also would like to thank the AG development team at ANL Futures Lab. for helpful comments.

References

1. R. Stevens et. al., "Advanced collaboration environments and scientific workplaces of the future," in *Proc. Workshop on Advanced Collaborative Environments (WACE 2004)*, 2004.
2. R. Stevens, M. E. Papka, and T. Disz, "Prototyping the workspaces of the future," *IEEE Internet Computing*, vol. 7, pp. 51-58, July/Aug. 2003.
3. L. Childers, T. Disz, R. Olson, M. E. Papka, R. Stevens, and T. Udeshi, "Access Grid: Immersive group-to-group collaborative visualization," in *Proc. Immersive Projection Technology Workshop*, 2000.
4. ANL Futures Laboratory, *Access Grid Toolkit v2.3*, <http://www.accessgrid.org/>.
5. UCB/LBNL, *Video Conferencing Tool (VIC)*, <http://www-nrg.ee.lbl.gov/vic/>.
6. M. Handley, C. Perkins, and E. Whelan, "Session announcement protocol," IETF RFC 2974, Oct. 2000.
7. M. Handley and V. Jacobson, "SDP: Session description protocol," IETF RFC 2327, July 2003.
8. K. Kobayashi, A. Ogawa, S. L. Casner, and C. Bormann, "RTP payload format for DV (IEC 61834) video," IETF RFC 3189, Jan. 2002.
9. WIDE Project, *Digital Video Transport System (DVTS)*, <http://www.sfc.wide.ad.jp/DVTS/>.
10. HDV Format, <http://www.hdv-info.org/>.
11. VideoLAN, <http://www.videolan.org/>.