Design of the VoD System for High-Quality Video and Audio with D1 Over IP

Michiaki KATSUMOTO

Our research group have researched and technological development concerning integrated into communications and broadcasting, partially of the next generation Internet technology development which used high speed networks. The technology which transmitted the digital video and audio of non-compression D1 format which was the business quality as a high-quality, digital video transmission technology on IP was developed as one project in that. In this paper, we report on the design of the Video-on-Demand system which can be used on Internet which uses high speed networks by using this technology.

Keywords
Next generation Internet, D1 over IP, DV over IP, Broadband networks

1 Introduction

Internet broadcasting and streaming contents have recently been attracting a great deal of attention, despite their inadequate content quality. The demand for such services is projected to continue to increase in the near future, and streaming contents are expected to play a major role among applications for the next-generation Internet [1] [2]. With the further development of DVTS, it will certainly be possible to provide high-quality content through the Internet [3] [4] [5]. On the other hand, digital broadcasting and compressed audio/video such as DVD (MPEG2) and DV are considered high-quality services and have become increasingly popular at home. However, they are not high in quality compared with the non-compressed audio/video used in professional applications such as commercial broadcasting and movies. As it becomes increasingly difficult to draw a distinction between communication and broadcasting, and the high-speed Internet gains in popularity among both content providers and home users under the circumstances described above, the need for higher-quality contents will grow and technologies for providing such high-quality contents will become essential.

We have therefore developed a transmission technology for super-high-quality digital video and multi-channel audio that can transmit data on a real-time basis. It employs the IP technology that is a core technology in the Internet world. This paper describes the technological aspects of and future prospects for this super-high-quality media transmission.

2 Outline of System Design

In order to establish super-high-quality video transmission and archive systems, we have investigated the following: (1) Video of professional quality (D1) and audio synchronization, (2) video of home quality (DV) and audio synchronization, (3) flexibility in the handling of a variety of video and audio formats for synchronization of the distribution of super-high-quality digital data plus real-time transmission technology using the TCP/IP protocol to establish IP communication. In addition, we investigated (1) management of the bandwidth of large-capacity archive disks, (2) management of the network bandwidth, (3) management of the bandwidth within the system, (4) management of the video interface bandwidth, as functions necessary to ensure stable data transmission across entire systems.

3 Implementation of D1 Over IP

The technology required for the implementation of the above functions, which we have decided to refer to as "D1 over IP," transmits non-compressed D1 video data loss-free by TCP/IP as super-high-quality media. The non-compressed D1 video consists of 60 fields of pictures per second, based on the interlace mode. In general, its video resolution has two modes: 8-bit and 10-bit. For the real-time display of non-compressed D1 video, a substantial transmission bandwidth of at least 270 Mbps (at a resolution of 10 bits; the bandwidth must be at least 217 Mbps at a resolution of 8 bits) is required over the entire bandwidth range of disk transmission, network transmission, and in-system transmission in typical cases. We have succeeded in displaying non-compressed D1 video sent from one server to two clients through the use of the TCP/IP protocol, which is employed on the Internet, in order to maintain synchronization between the non-compressed D1 video and the production quality AES/EBU audio. In this paper, we describe its implementation technologies.

3.1 Multi-AV format

The following problems are involved in the transmission of both high-quality digital video data and multi-channel audio data by the TCP/IP protocol from high-speed disks through high-speed networks.

(1) If audio and video data are sent separately by IP transmission, the synchronization between audio and video fails when delay processing occurs.

(2) Stable transmission fails due to collisions and shortages in bandwidth when audio and video data are sent separately by IP transmission.

The above problems make data transmission unstable over entire systems, including disks and networks.

In addition, the following problems also arise in the transmission of high-quality digital transmission data of various video formats.

(1) We must design a transmission block capacity suited for the capacity of the video field each time data of various video formats is transmitted. Therefore, the optimal transmission block capacity must be set at each disk transmission, network transmission, and in-system transmission site.

(2) If the data transmission becomes unstable at the site of disk transmission, network transmission, or in-system transmission, the entire bandwidth becomes unstable in terms of data transmission.

In order to solve these problems and realize stable operation of the systems, we have developed a multi-AV format as our original video data format, as shown in Fig. 1.

■ Multi-AV format:

The multi-AV format consists of an AV header and multi-AV data, as shown in Fig. 1.

■ AV header:

This header consists of the following information:

- AV data identifier
- Initial settings for the video interface
- Initial settings for the audio interface
- Determined values of the multi-AV data format

■ Multi-AV data:

This data consists of the following information sets:

- Field number
- Video field data
- Audio-frame number
- Audio-frame data
- Padding data

The format of the multi-AV data is determined based on the information included in the AV header. The format of multi-AV data of the structure shown in Fig. 1 is that for non-

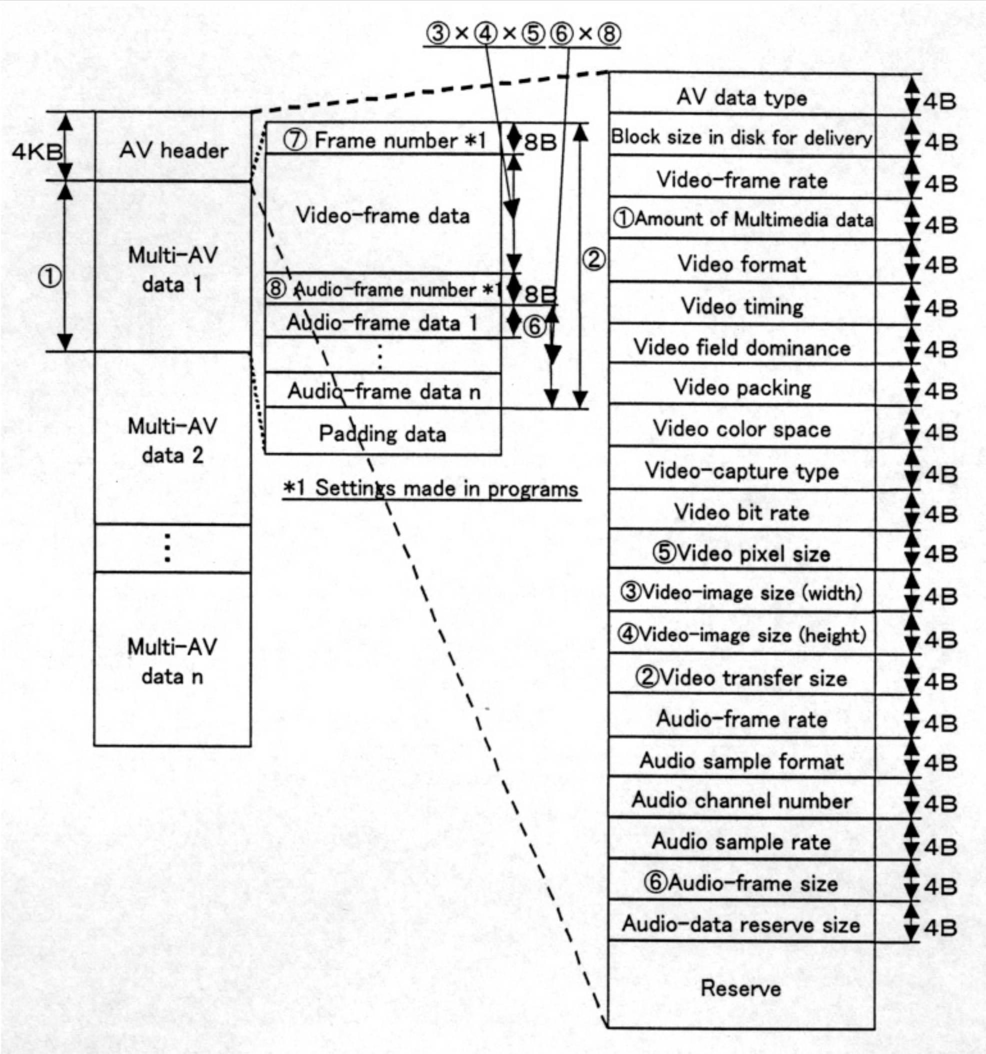

Fig. 1 *Multi-AV format*

compressed D1 data. In the case of DV and other data of different video formats, other parameter values are set in the AV header to suit the various video formats. We have thus adopted a format structure for multi-AV data that can be changed according to each transmission-data type. Therefore, it can correspond to a wider array of video formats.

3.2 Securing bandwidth and the multi-AV format

We have set the following functions using the management daemon at each site at which bandwidth must be secured by the system (Fig. 2).

■ Disk bandwidth management function: Reservation of the bandwidth for data I/O transmission from the disks in the VOD server to memory, and management of the transmission capacity

■ Network bandwidth management function: Reservation of the bandwidth for data I/O transmission from the VOD server and client memory to networks, and management of the transmission capacity

■ Video bandwidth management function: Reservation of the bandwidth for data I/O transmission from the ring-buffer memory of the VOD clients to the video interface, and management of the transmission capacity

In our system, each AV header of the multi-AV-format data stores a number of transmission blocks that provide data transmission optimized through bandwidth man-

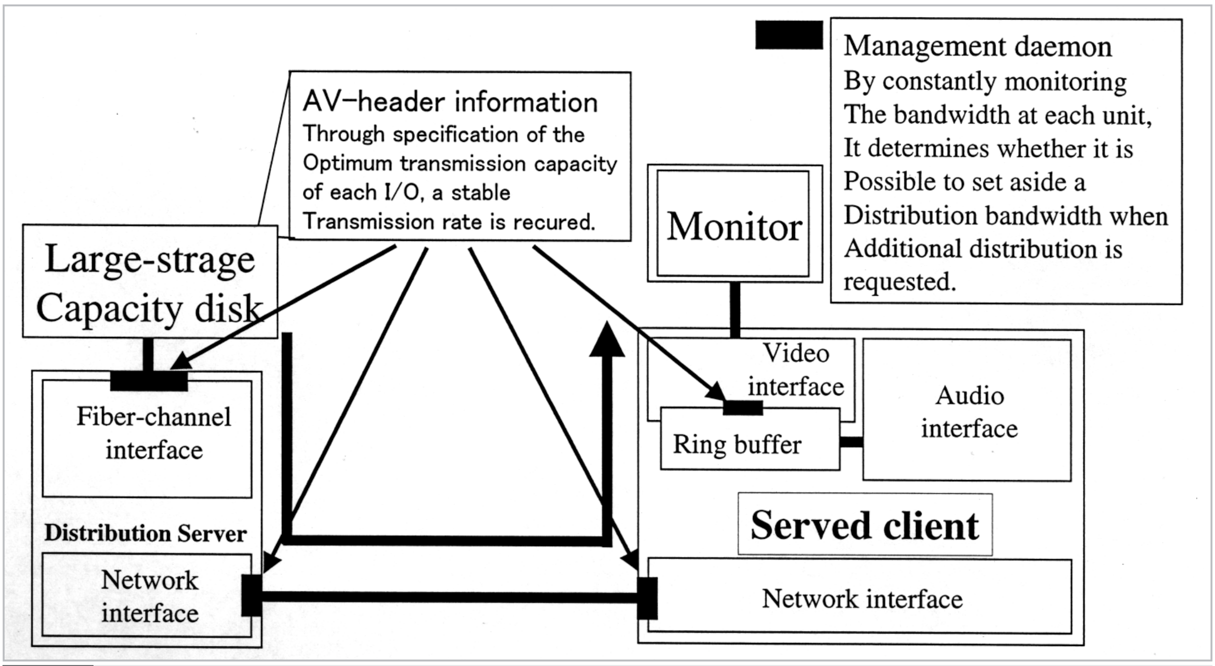

Fig.2 System Overview

agement. This transmission-block number is set for each bandwidth management function of the VOD server and client systems, and makes data transmission stable over the entire system. This implies that data transmission can be stabilized over the entire system by controlling the settings of the AV header, even in the transmission of high-quality digital data of a variety of multi-AV formats. The management daemon controls the transmission speed for each bandwidth management function and, if it is determined that it is impossible to secure the necessary transmission bandwidth in the case of a transmission request from a new client, notify the client of the refusal of transmission and continue to secure the optimal bandwidth.

3.3 Non-compressed D1 active-area transmission

Even if a wide-band network is available, less transmission data provides lighter loads on the network and higher stability, in some cases making it possible to transmit a larger amount of stream data. In our prototype system, when non-compressed D1 video data is transmitted, only the D1 active-area data shown in Fig. 3 is stored in the video-frame data area shown in Fig. 1, and the D1 control-area data is not transmitted. The non-compressed D1 video data consists of control signals such as output timing signals for the output of video data based on the D1 control-area data and the video data itself in the D1 active area. The control-area data is the data necessary for the video output of non-compressed D1 video. The D1 control-area data of a field is basically the same as that of another field. In the case of network transmission, video output is allowed only when the D1 active-area data has been sent out and the client system incorporates the D1 control-area data after receiving the D1 active-area data from networks. Through the use of this method, the amount of data transmission can be reduced by approximately 23% per stream compared with the method of transmitting all D1 video data simultaneously.

3.4 Boundary settings for stable disk transmission

The efficiency of disk access is increased when disk access is made on a block-by-block basis. However, data is not always read out in exact disk-block units. When padding data suited for the disk-blocking size is added to

AV data for faster access to the disk, the data I/O becomes faster and the transmission speed is stabilized, although the amount of data increases slightly. We have adopted a mechanism in the server that removes the padding data from the AV data saved in memory, and does not send padding data to networks. In this way, high-speed, stable disk transmission and a reduction in the load placed on the network have been attained.

3.5 Synchronization between audio and video

The AV data area stores non-compressed D1 video data and audio data on a field-by-field basis. The system uses the inner real-time clock employed by IRIX, the SGI operation system, and synchronizes the audio and video data on a field-by-field basis through multi-processing. As the audio and video data are stored on a field-by-field basis, data can be successfully synchronized and transmitted even in the event of data loss due to a delay during data transmission, which may be caused by an excessive load on back-bone networks.

3.6 Jitter control by the ring buffer

Despite the above measures for ensuring the disk bandwidth and network bandwidth, slight delays in transmission still occur due to fragmentations in disks and re-transmission processing on networks. The transmission of non-compressed D1 video data could lead to a field loss unless specified field data is sent in intervals of exactly 1/60 s. The client system includes a ring buffer to deal with such possible jitter problems in disks and on networks. The ring buffer is a mechanism for reading the data delivered from networks in advance, storing it in the ring buffer, and then transferring it to the video option by DMA, thereby ensuring the delivery of data. Even if a delay is caused by disks and networks, field loss does not occur during video output, as the data that has already been stored in the ring buffer is sent to the video option. The current system includes a buffer of an amount equivalent to one second of data (approximately 22 MB, or approximately 4 MB of DV data) for the transmission of non-compressed D1 data.

4 Overview of the actual system

Fig.3 illustrates the high-quality video-on-demand system we created. The primary system was an SGI workstation; the VOD server was equipped with six R10000 250-MHz CPUs, 3 GB of memory, and high-speed fiber-channel disks; and the VOD client was equipped with two R10000 225-MHz CPUs, 1 GB of memory, a D1 serial interface, and an IEEE1394 interface. Device drivers, daemons, and application software were created using an SGI compiler.

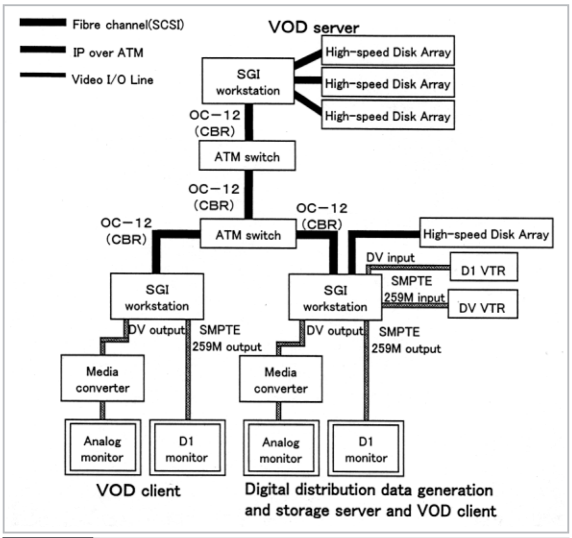

Fig.3 System configuration

4.1 Experimental results

We delivered two hours of high-quality digital data of non-compressed D1 video (8-bit) and audio data of 2-channel AES/EBU. The LAN used as the network in this experiment featured a configuration of only three PCs of one VOD server and two client PCs, as shown in Fig. 3. It was found through the measurement of data-transmission speeds at the VOD server and client PCs that transmission with no frame loss can be attained.

4.2 Field tests for verification

In addition to the successful experiment

using a LAN, field tests using a WAN were conducted for verification as described below.
(1) Research presentation at CRL (September 1999): Three buildings a short distance apart were connected by ATM OC-12 (partly using optical non-line transmission), and one stream each of D1 and DV were distributed. No frame loss occurred.
(2) N+I INTERROP (June 2000): Real-time transmission from a digital BetaCam camera was conducted, while an ATM LAN was employed at the site. One stream of D1 and two streams of DV were transmitted successfully with no frame loss.
(3) The Exhibition of Dream Technologies for the 21st Century [Yume Tech] (July-August 2000): Otemachi and Odaiba were connected by ATM OC-12, and one stream of D1 and six streams of DV were delivered over a period of approximately two weeks. While D1 video data was successfully transmitted with no frame loss, a failure occurred in the synchronization by IEEE1394 during DV transmission.
(4) INET2000(July 2000): Yokohama and Osaka (approximately 500 km apart) were connected by a gigabit Ethernet, and one stream of D1 was transmitted. Although time was required to achieve a stable transmission state due to TCP/IP slow starts resulting from improper settings, the transmission itself was made successfully with no frame loss. During slow starts, there was no real-time playback.
(5) In-house demonstration: Koganei and Otemachi (approximately 30 km apart, 60 km roundtrip) were connected using a gigabit Ethernet (Fig. 4), and two streams of D1 and one stream of DV were transmitted. All streams were successfully transmitted with no frame loss.

Thereby, D1 over IP has proven itself to be a practical system.

5 Discussion and future challenges

The present field tests have proven that the developed transmission technology for high-

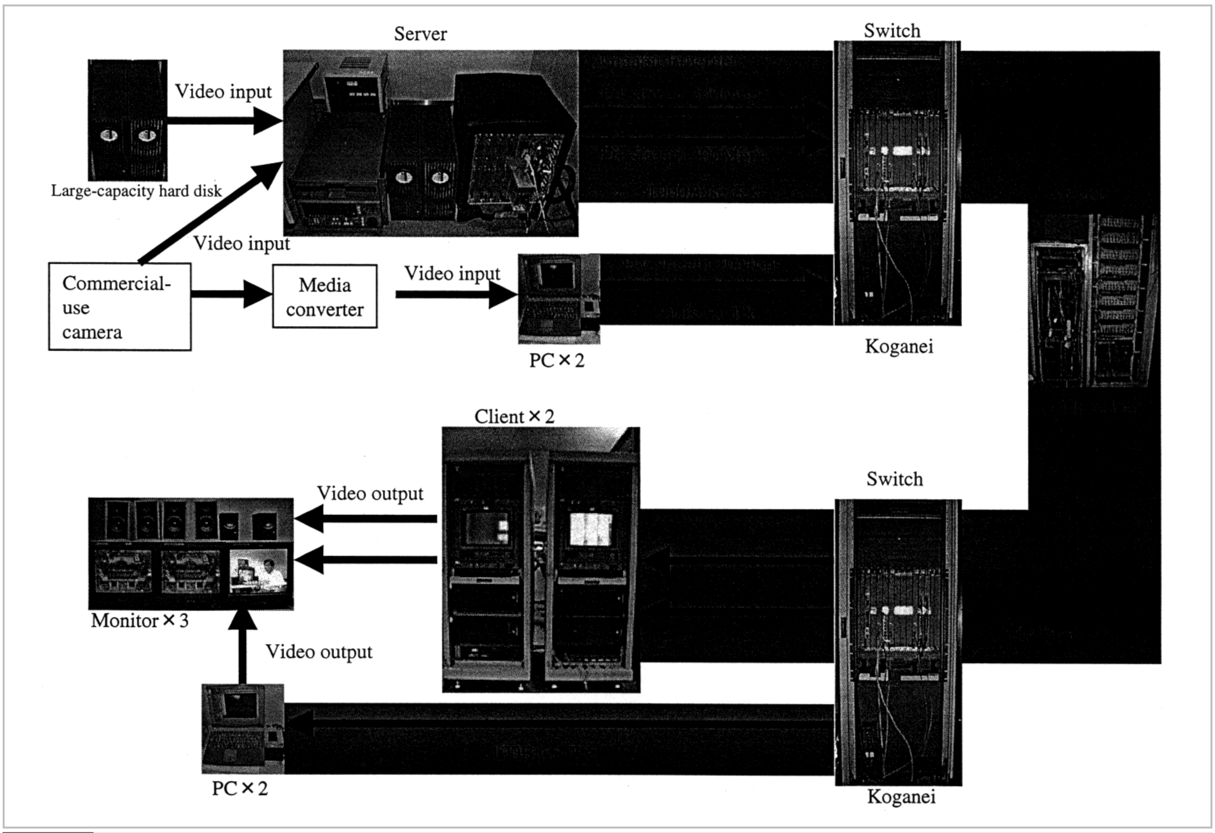

Fig.4 *Outline of the demonstration*

quality digital video and multi-channel audio data can be used in IP-based high-speed networks with professional levels of service. The operation of the present system was also checked by a three-point connection test using ATM switchers through the JGN (Japan Gigabit Network). The test results have shown that the developed technology can also be used in high-speed Internet setups. As the need for a high-speed Internet using a gigabit Ethernet and routers will grow in the future, we feel that further field tests are necessary. In addition, we feel that specifications corresponding to multi-cast transmission that takes network bandwidth and professional agent service into account should be investigated.

6 Conclusions

Communication and broadcasting are expected to be fused in the near future, requiring large amount of content to be prepared in order to satisfy demand. Existing data-handling technologies relying on tape media and video-dedicated lines may not provide the capabilities necessary to meet that demand. Our developed technology will enable high-quality media to be edited and distributed on a real-time basis, greatly contributing to the era of multiple channels and multimedia. Based on the successful results achieved by HDTV over IP, despite the fact that they were achieved within a LAN in a laboratory, we feel confident that the next-generation Internet will soon be a reality in areas in which super-high-quality media are distributed through an established super-high-speed-network infrastructure.

References

1 Yuki Kadobayashi, *"The ITRC working group report: Science and Next Generation Internet"*, SIG HQI IPSJ, Oct., 1999.

2 Kazuyosi Kobayashi, *"Construction of OKAYAMA Information High Way"*, SIG HQI IPS, Oct., 1999.

3 Kazunori Sugiura, Akimichi Ogawa, Osamu Nakamura, and Jun Murai, *"Internet Video Conference System based on Digital Video"*, Journal of IPSJ, Vol.40, No.7, 1999.

4 Kazunori Sugiura, Akimichi Ogawa, Osamu Nakamura, and Jun Murai, *"Interoperability of Home Networks and IEEE1394"*, Journal of Measurement and Control, Vol.39, No.8, pp.522-527, Aug., 2000.

5 A. Ogawa, K. Sugiura, K. Kobayashi, O. Nakamura, J. Murai, *"Design and Implementation of DV Stream Over Internet"*, IWS Internet Workshop, Feb., 1999.

***Michiaki KATSUMOTO**, Dr.Eng.*

Leader, High-Speed Network Group, Information and Network Systems Division

High-Quality Contents Transmission Technology, IP