SONY

VISUAL COMMUNICATION SYSTEM-TECHNICAL DOCUMENTATION

Streaming and Recording Capabilities

PCS-XC1	All
PCS-XG77/XG77S	All
PCS-XG100/XG100S	All

IPELA

Introduction

The Sony HD Visual Communication System (hereafter referred to as PCS *1) comes standard with the ability to multicast streaming video and audio over an IP network, and to record such data into a locally mounted USB flash memory. The streaming capability allows conferences to be viewed in real time from rooms that lack a PCS system, since all that is required is a PC and the ability to receive multicast packets.
The recording capability allows the conference to be recorded so that it can be reviewed later – analogous to the keeping of conference minutes. The ability to record minutes directly into a USB flash memory is an original Sony feature.

This document explains some of the technical details underlying these streaming and recording capabilities.

*1: The supported models are referred to on the front cover.

Streaming Capability

What Is IP Streaming?

IP streaming is the transmission of audio and video data over an IP network for purpose of immediate replay, where the sending side issues IP packets containing the data, and the recipients replay the data as it comes in. Streaming can be implemented either in real time, where audio and video data are sent out "live," or from storage, where the server holds the data as content that is available for transmission on demand. PCS systems implement the first type: "live" real-time streaming.

IP streaming can be carried out in either of two ways: by multicasting, or by unicasting. The main features of each are as follows.

[Multicasting]

In multicasting, the sending side issues a stream of identically-addressed IP packets, where the address identifies a multicast group participated by multiple end users. Routers on the network replicate the stream as necessary and direct it to the group's end users. Since the sending side must send only a single stream of packets, the processing load on the server is low. And since the routers send the stream only to clients that want it, overall network efficiency is maintained. But multicasting is possible only if it is supported by all of the routers between the sender and the recipients. Accordingly, multicasting requires the appropriate network environment in order to succeed.

[Unicasting]

With unicasting, the server issues a separate stream to each recipient, addressing each IP packet with the address of its final destination. Since the server must send separate packets for each final destination, the load on it increases in proportion with the number of destinations. Greater bandwidth is also required to handle the increased number of packets.

On the plus side, unicasting does not require multicast-capable routers and is generally easier to implement.

Multicast Settings

PCS systems utilize multicast rather than unicast streaming, as the reduced processing load helps to maintain audio and video quality from ongoing conferences. To enable successful multicasting, it may be necessary to set relevant parameters to appropriate values ahead of time. These various parameters are explained below.

[Multicast Address]

A predetermined range of IP addresses has been set aside for multicasting. Specifically, IP address space has been divided into 5 classes, A to E (where E is reserved for testing), with Class D assigned to multicasting. Class D space runs from 224.0.0.0 to 239.255.255.255. Addresses from 224.0.0.0 to 224.0.0.255 are reserved, however, which means that the space actually available for multicasting runs from 224.0.1. 0 to 239.255.255.255.
An IP address in this range identifies not a single user, but rather a multicast group. The stream server sends a data stream to this IP address, while interested end users recognize the stream as directed to them and receive it. The scheme allows end users to selectively receive various IP packets transmitted over the same IP network.

[Port Number]

Audio and video data is transmitted in UDP (User Datagram Protocol) packets. By setting the appropriate port number in the packet header, it is possible to transmit the packet through packet-filtering firewalls and routers. If a different port number is specified, it is also possible to send a different stream to the same multicast address.
Separate port numbers can be set for audio data and video data.

[Hop Count]

Each leg traveled by an IP packet on its way from source to destination is referred to as a hop. For example: If a message from a server must travel through a single router before reaching its final destination, the hop count is two-the first leg from server to router, and the second from router to destination.
The server side sets the hop count into the TTL (Time to Live) field of the IP packet header. Each router through which the packet passes then decrements the count value by 1. If the count value hits zero while the packet is at a router, the router discards the packet. This mechanism is an effective way to prevent stray IP packets from looping indefinitely through the network.

Data Compression

The PCS system generates both audio data and video data. Both types of data are compressed before sending, as follows.

[Audio]

Encoding:	ISO/IEC 14496-3 (MPEG-4 Audio) AAC LC
Bit rate:	64 kbps
Channels:	1 (monaural)
Sampling frequency:	32 kHz

[Video]

Encoding:	H.264 High Profile
Bit rate:	512 or 1024 kbps
Resolution:	720P (1280 x 720)
Frame rate:	Max. 15 fps

Communication Protocols

In order for a player to correctly replay a PCS stream (audio and video data), it must first know the encoding method, transfer protocol, and other such details. Accordingly, the PCS system uses SDP (Session Description Protocol) to notify recipients of the encoding method and other relevant settings. SDP is defined by RFC2327, a specification maintained by the IETF (Internet Engineering Task Force) as part of its effort to develop and promote Internet standards.

After using SDP to notify recipients of the relevant parameter values, the PCS begins sending the audio and video data streams. Since data must be sent in real time, it is sent using RTP (Real-time Transport Protocol). RTP is defined by RFC3550.

The actual data carried by the RTP packets is referred to as the payload. The payload format varies according to the type of encoding that is used. The RTP payload format for the MPEG-4 encoding used by the PCS system is defined by RFC3016 and RFC3640, and that for the H.264 encoding is defined by RFC3984.

To sum up, the RFC documents that define the communication protocols used by the PCS system are as follows.

- RFC2327: "SDP: Session Description Protocol"
- RFC3550: "RTP: A Transport Protocol for Real-Time Applications"
- RFC3016: "RTP Payload Format for MPEG-4 Audio/Visual Streams"
- RFC3640: "RTP Payload Format for Transport of MPEG-4 Elementary Streams"
- RFC 3984, RTP Payload Format for H.264 Video

Recording Capability

Recording to a USB flash memory

The PCS system can record audio and video data from a conference into a USB flash memory. This data is identical to PCS streaming data. If streaming and recording functions are both enabled, the system will record the audio and video data at the same time as it streams it.
Audio and video data are stored on the USB flash memory in standard MPEG-4 file format (in "MP4" files). Accordingly, the data can be played on any player that is capable of parsing MP4 files and decoding the audio and video compression used by the PCS system. As of this writing, correct operation has been verified on QuickTime Player version 7.

Data Compression

The compression schemes used for recorded audio and video data are the same as those used for streamed data, as enumerated earlier in this document.

Intraframe Recording

It is possible that a player may wish to begin video replay at any arbitrary point within the recorded data. Just as with streaming data, therefore, the PCS records intraframe data about once every two seconds. This allows the player to display correct video data relatively quickly, regardless of where it starts playback.

File Formats

Audio and video data are recorded into the USB flash memory using the MPEG-4 standard formats indicated below.

- ISO/IEC 14496-10: MPEG-4 Part 10 Advanced Video Coding

- This document contains registered trademarks and trademarks that are owned by their respective companies.

SONY

IPCT001V1