Technology White Paper

In this section, we have included a white paper that was written to describe the PCS-1 videoconferencing system technology in greater detail.

Because the PCS-11 system is based on the same technology, much of this paper is also applicable to the PCS-11. In fact, the PCS-11 uses exactly the same top-quality Sony camera and microphone as the PCS-1. The differences that allow the entry-level pricing for the PCS-11 are all in the CODEC, which was not built to host multi-point conferences.

So, please use this paper as a general technical reference for the PCS-1 and the PCS-11 systems, and just be aware that not all the functionality of the PCS-1 is included in the PCS-11. At the highest level, only the PCS-1:

— can support hosting of multi-point calls (Internal MCU Function)

— supports concurrent dual stream including live video and live data

— has the dedicated whiteboard input jack

Otherwise, the functions of the two systems are much the same, including:

— excellent quality audio and video

— easy to set up and use

— high level security functions

— sharing complex collaboration

— IP or ISDN connectivity

Technology White Paper

CORE TECHNOLOGIES OF SONY'S PCS-1 VIDEOCONFERENCING SYSTEM

Table of Contents

Introduction

Videoconferencing has a surprisingly long history. Experimental analogue-video-based systems were developed decades ago. The first commercial system based on digital technology, using digital- circuit-switched networks, was put on the market in the early 1980s. However the poor quality of video and audio and the exorbitant price of the equipment hindered its wide acceptance. The high tariff on circuit-switched lines created another bottleneck to the acceptance process. Videoconferencing systems were used only in wealthy organizations and in schools and until the end of the 90s they were not considered to be popular or user-friendly.

The recent rapid progress in IC technology has changed this situation. Most significantly, it has enabled dramatic cost reductions in videoconferencing equipment. Moreover, the penetration of broadband IP networks has enabled a cost-effective IP-based videoconferencing infrastructure.

Sony's new IP-based videoconferencing system PCS-1 encompasses the emerging IP-based broadband network infrastructure and provides a highly reliable, cost-effective videoconferencing environment for corporate users, educators, teleworkers, and others. The system provides not only a conventional audio and videoconferencing environment, but also a collaboration environment that enables data sharing of RGB images output from a PC as well as white board images.

This paper first illustrates the features of the new PCS-1 Videoconferencing System. Some of the highlights of the PCS-1's attributes are introduced in Chapter 2. The key features of the system – QoS, MCU function, advanced video and audio coding technology, whiteboard function, security, PPPoE, and dual video streaming – are described in detail in Chapter 3.
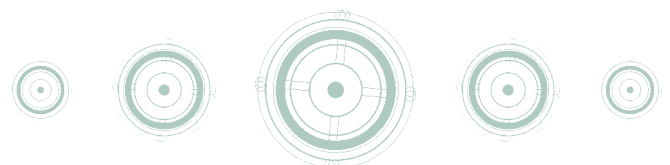
Product features

The PCS-1 is a feature-rich, high-performance videoconferencing system available at a very affordable price, developed to address a variety of demands in the conferencing market. Having incorporated advanced AV and IT technologies into its compact B5-footprint body, the PCS-1 can address not only small-to-large-group conferencing applications but also distance education, telemedicine applications, and much more. The PCS-1 will bring a new level of participant interaction, creativity, and productivity to videoconferencing communication. Listed below are just some of the highlights of the PCS-1 features.

Free-Layout Design

One of the innovations that distinguish the PCS-1 from other systems is its compact two-piece design. The unit consists of a codec and a separate camera. This "free layout" enables the PCS-1 not only to sit on top of a conventional TV monitor but also to function alongside a flat panel display, such as a plasma or projector screen. Installed on a camera stand (PCS-STP1 and PCS-STG1), the PCS-1 can be located just beside the flat panel display or projector screen easily and without restrictions. With the PCS-1's flexibility of installation, videoconferencing is no longer limited to a dedicated room with a TV monitor, but instead can be extended to executive offices with plasma displays or any meeting space or desk with a projector or LCD monitor.

Data Sharing

PC Desktop Sharing

The ability to share presentation data from applications like Microsoft® Power Point®, Excel, and Word has became a very important element of videoconferencing. In response to this demand, the PCS-DSB1 (Data Solution Box), a separate optional device, was developed. In conjunction with PCS-DSB1, the PCS-1 is able to transmit a PC desktop image in XGA resolution to PCS-1s at the far ends so that remote participants can view the presentation material in real time with the original resolution.

Digital Whiteboard I/F

Other types of data to be shared in videoconferencing are notes and drawings on the whiteboard. To meet this demand, the PCS-1 is equipped with direct I/F for the Mimio Xi™ digital whiteboard recorder from Virtual Ink Corporation. Using Mimio Xi, notes and drawings on a conventional whiteboard can be electronically converted and then transferred in real time to remote sites for onscreen display. This is a unique feature of the PCS-1 and is described in detail in Chapter 3.

Superb Video Quality

High-Speed Connection

Responding to the proliferation of broadband IP networks, the PCS-1 can achieve up to 2Mbps at a 30 fps IP connection, bringing superb video quality to videoconferencing. Using the PCS-B768, an optional ISDN unit, the PCS-1 can maintain videoconferencing at up to 768Kbps via an ISDN line.

Advanced Video Coding

MPEG-4, H.264, and interlaced SIF are supported

The PCS-1 is the first in the industry to support MPEG-4 Video as standard. MPEG-4 capability provides the potential for connection with a variety of consumer video communication devices in the future.

H.264, the newly ratified standard, is capable of providing the same picture quality on half the bandwidth. For instance, a video conference call placed using a bit rate of 384Kbps can now deliver the same picture quality of a 768Kbps bit rate call. Network congestion problems and bandwidth limitations can also be overcome due to the lower bandwidth demand of H.264 coding.

Interlaced SIF video coding operates at 60 fields per second rather than coding and decoding 30 frames of video per second, resulting in smoother, more life-like images. Interlaced video coding is enabled at bit rates higher than 512Kbps on H.323 calls in "motion priority" mode.

Superb Audio Quality

The PCS-1 supports MPEG-4 AAC. With this audio codec, the PCS-1 is able to transfer 14 kHz wideband audio, which is double the bandwidth of conventional videoconferencing audio. This wideband audio raises clarity of sound to a new level and greatly facilitates recognition of the speaker's voice. The details are described in Chapter 3.

Enhanced Internal MCU

The PCS-1 supports high-performance internal MCU as an option. Videoconferences among up to 6 sites can be held using IP or ISDN connections with PCS-323M1 and H.323 MCU software respectively. In this internal MCU operation, a 4-way or 6-way split-display mode is available depending upon the number of sites. In addition both H.320 (ISDN) and H.323 (IP) calls can be connected on a single multi-point conference call if both PCS-323M1 and PCS-320M1 are installed. In most cases, this increased flexibility eliminates the need for an outside bridging company or additional expensive network hardware when deploying mixed-network multi-point videoconferencing. By cascading two systems with H.323 MCU software, up to 10 sites can be connected on a multi-point call. The details are described in Chapter 3.

Enhanced QoS

The QoS, Quality of Service, is one of the key issues in IP-based videoconferencing. In order to address the needs for QoS enhancement, the PCS-1 adopts not only Network-level QoS, IP Precedence, Type of service, and DiffServ, which are commonly used to deliver high-quality video and audio services over IP networks, but also two advanced technologies, ARC (Adaptive Rate Control) and Real-time ARQ (Auto Repeat reQuest). ARC is used to adjust the data rate in response to network conditions. ARQ is used for packet loss recovery. By employing the optimal algorithm that provides the best combination of the two technologies, the PCS-1 is able to maintain the picture and sound quality in varying network conditions. The details are described in Chapter 3.

Private Phonebook

The phonebook stored in the PCS-1 is a common one for all users of the PCS-1. For convenience and improved functionality, private phonebook listings can be created and maintained on a Memory Stick™. When the Memory Stick is inserted, the PCS-1 automatically recognizes the private phone book as the primary one, thereby eliminating the need for menu adjustments. As an added convenience, one of the video phone numbers can be registered for auto-transmission so that when the Memory Stick is inserted into the PCS-1 unit, the call is executed instantaneously.

Memory Stick Support

The PCS-1 is equipped with an M/S (Memory Stick) slot on the codec unit. Notes and drawings or any image displayed on the screen – either near- or far-side – can be captured in the "SnapShot" mode, saved in a JPEG file, and recorded to Memory Stick® media inserted in the PCS-1 codec unit. Additionally, pre-recorded images saved on Memory Stick media can be displayed in 4CIF format and transferred during the call.

Core Technologies

3.1. QoS

On the current Internet, an estimated 5 % of packets are lost. Packet loss causes video frame freeze, error propagation, and audio glitches in videoconferencing sessions. The QoS (Quality of Service) and packet loss compensation features are very important when videoconferencing in less than optimal "best-effort" IP network environments.

The main objective of QoS features used in conventional videoconferencing endpoints is to "conceal" the error. For example, certain technology is used to shuffle image data for transmission so that packet loss does not destroy wide areas of the video image. Other technology is used to increase "intra" macroblocks requiring no previous frame information, so that the packet loss error does not propagate over a long period, but the increased amount of intra macroblocks requires more bits to carry video frames, resulting in a lower frame rate.

Sony has developed new QoS features using a different approach. Sony's approach is to "recover" the error – not to "conceal" it – thereby maintaining the quality of real-time communication. With Sony PCS-1 QoS features, users need not be concerned about bandwidth and network quality; the PCS-1 automatically adjusts its bandwidth, buffering size, and algorithm to maintain high-quality conferencing.

The following two major functions are implemented on the PCS-1 to provide the above-mentioned features:

— Real-time ARQ (Automatic Repeat reQuest)

— ARC (Adaptive Rate Control)

First Network level QoS, IP Precedence, Type of Service, and Differentiated Services are briefly introduced; these are commonly used to deliver high-quality video and audio service over IP networks and are adopted in PCS-1. The details of the newly developed functions are covered in the following sections.

3.1.1. Network level QoS

The PCS-1 can enter the values of IP Precedence, Type of Service, and Differentiated Services. The TOS (Type of Service) field in IP header is used for either defining IP Precedence and Type of Service, or DSCP (Differentiated Services Code Point) bits of Differentiated Services. The usage of the field for either service is up to the service administrator of the network.

• IP Precedence and Type of Service

0	1	2	3	4	5	6	7
Precedence			Delay	Through-put	Reliability	Minimum Cost	CU

CU: Currently Unused

The value of IP Precedence and Type of Service can be defined in Setup menu.

IP Precedence bits The IP precedence bits indicates the priority with which a packet is handled as follows:

1 1 1 : Network control

1 1 0 : Internetwork control

1 0 1 : Critical

1 0 0 : Flash Override

0 1 1 : Immediate

0 1 0 : Priority

0 0 1 : Routine

Delay bit Used when the time it takes for a datagram to travel from the source PCS-1 to the destination PCS-1 or host (latency) is most important. A network provider is requested to select a route with the minimum delay when this bit is set to on.

Throughput bit Used when the volume of data transmitted in any period of time is important. A network provider is requested to select a route producing the maximum throughput when this bit is set to on.

Reliability bit Used when it is important to be certain that the data will arrive at the destination without requiring retransmission. A network provider is requested to select a route with maximum reliability when this bit is set to on.

Minimum cost bit Used when minimizing the cost of data transmission is important. A network provider is requested to have datagrams routed via the lower-cost route when this bit is set to on.

Unless the value of this field is set by the administrator, the default values of this field are all 0.

• Differentiated Services

The Differentiated Services (Diffserv) architecture is based on a network model implemented over a complete Autonomous System (AS) or domain. By putting this domain under administrative control, it is possible to establish clear and consistent rules to manage traffic entering and flowing through the networks that conform the domain. Diffserve defines a field in the IP header called the Differentiated Services Code Point (DSCP), which is a six-bit field as shown below.

0	1	2	3	4	5	6	7
DS5	DS4	DS3	DS2	DS1	DS0	CU	CU

CU: Currently Unused

Type of service field of the IP header is used to define DSCP. Hosts in a network that supports DiffServe make each packet with a DSCP value. Routers within the DiffServe network use these values to classify the traffic into distinct service classes according to DSCP value. Thus, Routers control packets not on a flow-by-flow basis but by traffic classes based on DSCP marking. Because the routers are not required to maintain any elaborate state information to identify the flows, the routers can handle a large number of flows. PHB (Per Hop Behavior) is defined according to the traffic classes based on DSCP marking. For example, if the routers receive packets with DSCP = 101110 that means expedited forwarding (EF), the routers are requested to forward the packet for low latency and low-loss service. Assured Forwarding (AF) that defines three levels of drop precedence is also defined and is used to guarantee the minimum bandwidth.

When this field is used as Differentiated Services, the default value is 000000, meaning simply that the best possible service is being provided.

3.1.2. Real-time ARQ (Automatic Repeat request)

The Real-time ARQ is a mechanism that automatically resends lost packets and reorders received packets. Thanks to Real-time ARQ, the PCS-1 can recover almost completely from a packet loss rate of 10%, automatically maintaining the minimum latency according to the network environment.

Figure 3.1 describes how Real-time ARQ works. The packet loss of an RTP packet of video/audio bit stream is detected on the receiver side. A "resend request" is then sent to the transmission side using an RTCP (Real-time Transport Control Protocol) packet. On the transmission side, the transmitted packet is held, preparing for resending according to the resend request.

By using an RTCP packet, the Round-Trip Time (RTT) – or network latency – can be measured. If the RTT is large, it would be a waste of network traffic to send the "resend request" knowing that the resent packet would not arrive in time for decoding. Sony's Real-time ARQ is capable of determining whether or not the resending packet would arrive in time, and adaptively selects the optimum algorithm according to the RTT and packet loss rate.

Figure 3.1 Real-time ARQ resending diagram

Tpi: Packet Interval
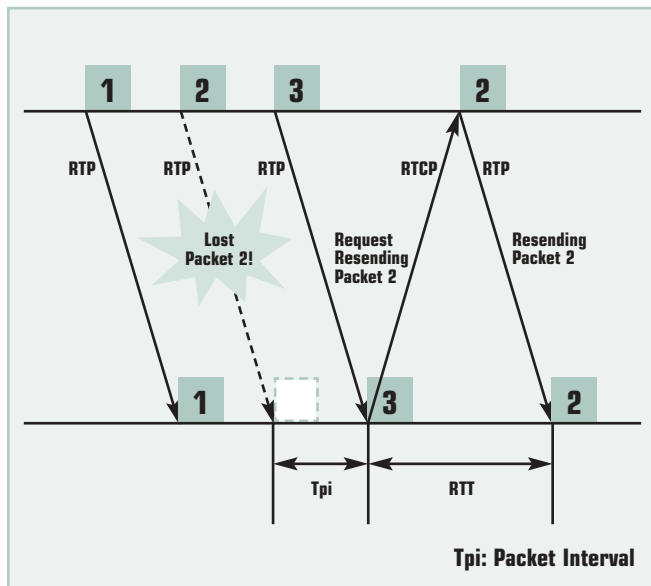
A. Minimum latency with optimum packets buffering
In general, resending packets requires a receiver buffer for reordering of packets, and hence increases the system latency. Sony's PCS-1 has a variable-length reordering buffer for rearranging a packet in order of the sequence number of RTP headers. The size of the reordering buffer is optimized according to the measured network RTT and PLR (Packet Loss Rate). In an environment without packet loss, the reordering buffer is minimal and hence the communication latency is not affected.

B. Adaptive algorithm switching
In addition to the buffer size, the resending algorithm itself is adaptively switched according to the RTT and PLR. For example, if the network latency (RTT) is very large, redundant audio transmission is used instead of packet resending to maintain conversation quality.

C. MCU and presentation data
Real-time ARQ works in combination with the internal MCU function. When the MCU function is engaged, each link between the terminals and the PCS-1 with MCU function uses ARQ to recover from packet losses. Additionally, when the PCS-DSB1 (Data Solution Box) is used, presentation data can also be recovered from packet losses.

3.1.3. Adaptive Rate Control
Adaptive Rate Control is a mechanism used to slow down the video bit rate if network congestion occurs. Unlike similar conventional methods, the PCS-1 employs a TCP-friendly rate control which adjusts to other TCP traffic (e.g. FTP) having its own rate control functions. TCP-Friendly Rate Control (TFRC) is designed for unicast flows operating in an Internet environment and competing with TCP traffic. The TCP throughput equation in IETF RFC3448 is a function of packet loss rate, and RTT should be suitable for use in TFRC. Figure 3.2 is an example of the calculated target video bit rate.

Figure 3.2 Adaptive Rate Control (for example, max_video_rate is 960Kbps)
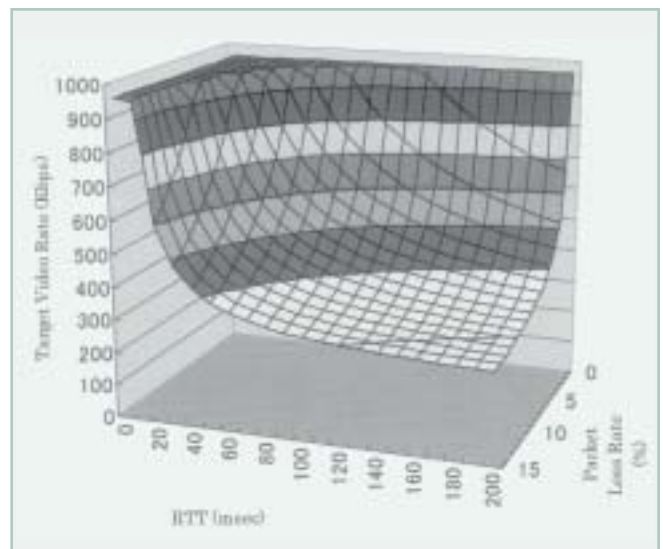
In the worst case, Adaptive Rate Control decreases the video bit rate to 1/4 of the initial bit rate, or the minimum video rate of 64Kbps.

3.2. Internal MCU Function
Outline

By using the internal MCU function, up to five videoconferencing terminals can be connected to the PCS-1 in both H.320 and H.323 modes. Small-scale multi-point conferences are supported without external MCU equipment. In addition, both H.320 (ISDN) and H.323 (IP) calls can be connected on a single multi-point conference call using the H.320/323 bridging function.

The MCU function of the PCS-1 offers all basic functions as MCU, such as FECC (Far End Camera Control), broadcast mode, and continuous presence/voice activation function. What's more, if only PCS-1s are connected, it is possible to support the digital whiteboard function and PC data-sharing function in a multi-point conference environment.

The maximum 10-site meeting is attained through the cascade connection of two PCS-1s with H.323 MCU function installed.

Main features

Supported protocols	H.323(LAN), H.320(ISDN), H.323/H.320 Mixed
Number of connected terminals	Up to 6 terminals in H.320, H.323, H.323/H.320 Mixed mode Up to 10 terminals by cascading two PCS-1s with MCU function. The connection between two PCS-1s with MCU function must be based on LAN (H.323).
Video Coding	H.263, H.261 CIF/QCIF
Audio Coding	G.711, G.722, G.728

Maximum bit rate and frame rate	**H.323**		
	Number of Sites	**Max Bandwidth (Kbps)/each**	**Max Frame Rate**
	1	1920	15
	2	960	15
	3	640	15
	4	480	15
	5	384	15

	H.320		
	Number of Sites	**Max Bandwidth (Kbps)/each**	**Max Frame Rate**
	1	384	15
	2	384	15
	3	256	15
	4	128	10
	5	128	7.5

	H.323/H.320 Mixed In mixed mode, the video rates (not total rate) are matched between the ISDN and LAN interface. For example, if a PCS-1 set at G.722 audio is connected via two B channels of an ISDN line to a PCS-1 set at 1024Kbps LAN bandwidth and G.728 audio, the video rate becomes approximately 96Kbps, which is the video rate of the ISDN connection. The total bandwidth for PCS-1s that are connected via LAN becomes approximately 160Kbps (video: 96Kbps, audio: 64Kbps).

Display Mode Continuous presence (Split), voice activation, and broadcast mode	Continuous Presence Layout Four-way and six-way split-display modes are supported depending upon the number of sites connected. **4Site Up to 3 terminals:** 	FAR1	FAR2
Near	FAR3	 **6 Site 4 to 5 terminals.** There are two modes, voice-activated and fixed-mode, for selecting the site in the large window. **1 voice-activated mode** When FAR5 is talking: When FAR3 is talking: FAR1 FAR2 FAR3 / FAR4 / Near / FAR5 FAR1 Near FAR2 / FAR4 / FAR5 / FAR3 **2 fixed mode** The MCU terminal selects the site in the large window. When FAR1 is selected: When FAR3 is selected: Near FAR2 FAR3 / FAR4 / FAR1 FAR1 Near FAR2 / FAR4 / FAR5 / FAR3
FECC	H.281 FECC (Far End Camera Control) is available in H.320, H.323, and H.323/320 mixed modes. In H.320 mode, H.224-LSD 6.4Kbps is used. In H.323 mode, H.224 logical channel is used. The camera at the site currently displayed in full-screen mode can be controlled at the terminal with MCU or at other terminals. In the continuous-presence mode, FECC is not supported.
Whiteboard	Whiteboard data is transmitted using H.224-LSD in H.320 mode or using H.224 logical channel in H.323 mode. Sony's original data format and protocol are used. Moreover, the MCU terminal controls other terminals using original token management protocol in order to select one terminal for transmitting whiteboard data in conferencing.

| Data sharing function | The PCS-1 MCU enables data sharing of XGA data. XGA data from one site, which can be the terminal with MCU function, is automatically distributed to all the other terminals. The PCS-1 MCU function controls other PCS-1s so that XGA data from two or more terminals is not transmitted simultaneously and transmission of XGA data is prevented while the PCS-1 is receiving XGA data. Moreover, when XGA data is transmitted, the video signal is controlled so that the video of the terminal to which XGA data is sent, is selected and transmitted to all other terminals in H.323 mode. On the other hand, the split mode can be maintained even when XGA data is transmitted in H.320 mode.

Availability of a dual stream of video and XGA data depends on the modes.

In the case of H.323 mode, dual streaming of video and XGA data is possible if all the terminals are PCS-1s with Version 2.0 firmware installed or if the terminals are a combination of PCS-1s with Version 2.0 firmware or with Version 1.0/1.1 firmware and RGB output ON. Otherwise continuous transmission of 4CIF still picture is used.

In H.320 mode, dual streaming is possible only when all the terminals are PCS-1 with Version 2.0 firmware.

In H.323/H320 mixed MCU mode, only the continuous transmission of 4CIF still picture is supported. XGA data transmission will become available when the transmission format of dual streaming fully supports the H.239 ITU-T standard in H.323 mode. |
|---|---|
| Cascading | Connection of PCS-1s with MCU function is supported in H.323 mode; up to 10 terminals can be connected.
— The functions of FECC/Whiteboard/RGB Data sharing are supported.
— When two PCS-1s with MCU function have a cascade connection and another PCS-1 with MCU function attempts to connect, it is disconnected automatically.
— Continuous-presence mode is not supported. When a cascading connection is established, the split-screen mode automatically changes to full-screen mode. |

3.3. Advanced Video and Audio Coding Technology
3.3.1 H.264 (MPEG-4 AVC)

Also known as MPEG-4 Advanced Video Coding (AVC), H.264 is an emerging video compression standard. This is the product of the Joint Video Team of ITU-T and MPEG, officially ratified as H.264 on May 2003. Although the H.264 algorithm requires huge computation power, approximately 40-50% lower bit rate is needed to achieve the same picture quality as any of the previous encoding standards, such as MPEG-4 and H.263. H.264 produces significantly better picture quality at a low bit rate.

A very powerful digital signal processor (TI C64x) used in the Sony PCS-1 provides 30 fps H.264 encoder and decoder capability. Software Version 2.0 or later supports H.264 capability, which is used as a default codec except in the "Motion Priority" mode, or when the other side does not support H.264.

Table: History of Video Coding Standards

	1990	1992	1994	1995	1998	1999	2003
ITU-T	H.261		H.262	H.263	H.263(v2)		H.264
ISO/IEC (MPEG)		MPEG-1	MPEG-2			MPEG-4	MPEG-4 (AVC)

3.3.2. MPEG-4 AAC Wideband Audio

In order for all of the participants in a video conference to feel as if they are in the same room, high-quality audio is critical. Conventional videoconferencing equipment covers only the "speech bandwidth". Most of the voice CODEC algorithms support 3.4 kHz telephone-line-quality bandwidth (FS = 8 kHz), and even using "wideband" G.722, the bandwidth only extends to 7 kHz (FS = 16 kHz).

The Sony PCS-1 offers MPEG-4 Advanced Audio Coding (AAC) among its various audio CODEC choices. The AAC is designed to compress high-quality audio, such as music, with extremely high coding efficiency. Sony's implementation in the PCS-1 uses the bandwidth of 14 kHz (FS = 32 kHz), enabling the conference to share the atmosphere of the remote site. The PCS-1 turns on the MPEG-4 AAC audio when using an IP-based connection.

Sony's videoconferencing endpoints use well-proven high-quality acoustic echo and noise cancellers, which quickly adapt to the environment of each conference room. For the PCS-1, Sony developed a wideband echo canceller algorithm to accommodate 14 kHz audio bandwidth.

Table: Audio CODEC Algorithms on PCS-1

Audio Bandwidth	Sampling Frequency (FS)	Algorithm	Data Rate
3.4kHz	8kHz	G.711	56Kbps, 64Kbps
		G.723.1	5.3Kbps, 6.3Kbps
		G.728	16Kbps
		G.729	8Kbps
7kHz	16kHz	G.722	48, 56, 64Kbps
		G.722.1	24, 32Kbps
14kHz	32kHz	MPEG-4 AAC	64, 96Kbps

3.3.3. Interlace Video

Software Version 2.0 or later supports Interlace SIF video with 60 fields/sec in NTSC and 50 fields/sec in PAL. The SIF (Source Input Format) is 352 x 240 (NTSC) and 352 x 288 (PAL), and interlaced SIF results in twice as much vertical resolution with 352 x 480 (NTSC) and 352 x 577 (PAL). As with the 60 fields/sec frame rate, the motion is much smoother than in the 30 fps conventional CIF.

This mode uses ITU-T recommendation H.263 Annex W and is provided only in point-to-point connections. Please note that in this mode, the amount of data transmitted is twice as much as in the 30 fps CIF, and hence a higher bit rate is required to provide sufficient picture quality.

3.3.4 MPEG-4 Video

MPEG-4 (ISO14496) is an ISO/IEC standard developed by MPEG (Moving Picture Experts Group), and it became an international standard at the beginning of 1999. MPEG-4 addresses the need for distributing rich interactive media over a variety of networks. MPEG-4 is widely used for high-quality video over IP networks, including streaming media over the Internet using PC, PDA, and 3G mobile phones.

While MPEG-4 has many profiles and levels to accommodate various applications, the PCS-1 has adopted the most widely used profile, Visual Simple Profile. The MPEG-4 Visual Simple Profile (SP) is based on the H.263 compression algorithm. It also shares many of the H.263+ tools defined in the Annexes of H.263, such as Deblocking Filter, Advanced Intra Prediction, and Unrestricted Motion Vectors. With some additional tools, MPEG-4 SP gives slightly better picture quality than H.263+.

3.4. Whiteboard Function
Outline
The PCS-1 is equipped with direct I/F for the Mimio Xi™ digital whiteboard recorder from Virtual Ink Corporation. Using Mimio Xi, notes and drawings on a conventional whiteboard can be electronically converted and then transferred in real time to remote sites for onscreen display. This function is supported in various modes, point-to-point and multi-point configuration, including cascading connection in H.323 and H.320 modes. It is possible to share notes and drawings on a whiteboard on the screens of all the connected terminals as well as to save the data of the drawings to a Memory Stick.

Functions
A. Real-time drawing
You can draw a picture using a 4-color pen, large eraser, and small eraser. A picture of the drawing is displayed on all the terminals in the conference in real time.

B. Transmitting data
All drawing data is transmitted using H.224-LSD in H.320 mode or H.224 logical channel in H.323 mode. The maximum data transmission rate is 6.4Kbps in H.320 mode and 64Kbps in H.323 mode. Data sharing of the drawings through a network with limited capacity is made possible by processing and compressing the data of the drawing.

C. Display on RGB monitor or TV in dual-screen mode
One monitor is used for pictures of the conference and the other monitor, an RGB monitor or TV, is used for whiteboard.

D. Saving images to Memory Stick as JPEG files
You can save a picture of a drawing to Memory Stick as a JPEG format file at all terminals in the conference. The size of a JPEG format file is 1024 x 768 pixels.

E. Whiteboard
The only whiteboard the PCS-1 supports is the Mimio Xi from Virtual Ink Corporation.

3.5. Security
Needless to say, it is of the utmost importance that corporate users have a conference system secure enough to hold confidential meetings and exchange confidential data. Because videoconferencing now uses both an Intranet, which is installed only inside companies and organizations, as well as the Internet, which malicious hackers can easily attack, a function to encrypt the data transferred through the IP network is essential. It is also important to prevent unauthorized users from logging in to the PCS-1; otherwise malicious users could eavesdrop on a conference as a web client, or secretly change the setting of the PCS-1 by logging in using Telnet or as a web client. A password protection mechanism is therefore provided to prevent access by unauthorized users.

Encrypting data
It is possible to encrypt data only with a LAN connection, and in Point-to-Point, MCU, and cascade mode, only if the connection is H.323-based and all the end points connected are PCS-1s. In H.320 and mixed-MCU mode, the encryption function is not provided because ISDN connection is a closed connection and the security level is much higher than with an IP connection.

Video, audio, and DSB data are encrypted. Data for far-end camera control and whiteboard images are not encrypted.

There are two factors that must be considered when implementing the encryption function: the encryption algorithm itself and how the encryption key is delivered.

In regard to the encryption algorithm, we adopt the AES (Rijndael) algorithm, which is the abbreviation for Advanced Encryption Standard, selected by the U.S. National Security Agency as its next-generation algorithm. The encrypted part of the IP diagram is limited to the payload part of RTP (real-time protocol) packets based on UDP.

AES uses a 128-bit symmetric key, and the sender of a message and the receiver must hold the same encryption key. Known by both the sender and receiver, the password is used to generate the common encryption key. Since a proprietary Sony algorithm is used to generate the encryption key from the password, it is not possible to communicate with videoconferencing terminals from other vendors, like Tandberg, even if the encryption mode is set to ON.

To initiate encrypted videoconferencing, the password with 13 to 20 alphabetical characters, numerals, or symbols must be registered at both ends. Encrypted videoconferencing can be established only if the passwords at both ends match.

Password protection
The PCS-1 supports three levels of the password protection system. It is assumed that the system administrator sets up the passwords on the setup menu if necessary.

— **Administrator Password**
    For cases in which the administrator password is set by the administrator, users must enter the password in order to change any settings in the setup menu or phonebook. Users must also enter a user ID and password to log in by Telnet.

— **Super User Password**
    For cases in which the Super User Password is set by the administrator, users must enter the password in order to make any changes to the phonebook.

— **Remote Access Password**
    For cases in which the Remote Access Password is set by the administrator, users must enter the user ID and the password to access the PCS-1 via a web browser in order to control the PCS-1. It is also possible to access the PCS-1 via a web browser as administrator or super user.

3.6. PPPoE
H.323 protocol requires that the global IP address be embedded in the data area of the IP packet. Therefore, when an NAT router is used to connect the PCS-1 to the Internet, users must enter the global IP address assigned to the NAT router on the PCS-1 setup menu in order to have it in the data area. However, it is not easy for non-technical end users to understand the meaning of entering a global IP address. PPPoE eliminates this troublesome procedure and enables the PCS-1 to be connected directly to the Internet.

Standardized at IETF and defined in RFC1661, PPP is the most proven architecture, having worked well for dial-up for more than a decade. By requiring a password/ID handshaking procedure before network configuration is granted, this approach inherently supports the authentication process required to track usage and to bill for services accordingly. The line can be ADSL, FTTH, or CATV, if the modem interface is Ethernet.

The procedure for obtaining the IP address by PPPoE consists of two steps as shown in the chart below.

1. User Authentication
2. Global IP Address Allocation

Figure 3.3 The procedure for obtaining the IP address by PPPoE
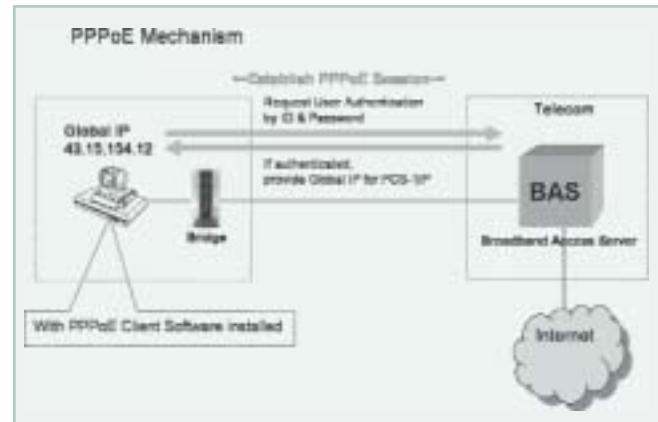
The PCS-1 is connected to the Bridge via Ethernet. The broadband access server of the ISP resides at the facility of the telecommunications company. First the PCS-1 will request user authentication from the BAS with the User ID and Password. If authenticated, the BAS will provide the Global IP Address to the PCS-1. Then the PCS-1 can communicate with remote terminals through the broadband network.

3.7. Dual Video Stream
In both H.323 and H.320 mode, RGB data is transmitted using additional video channels. In the case of H.320 mode, the protocol is fully compatible with the H.239 ITU-T (Role Management and Additional Media Channels for H.300-series Terminals) standard, and terminals from other vendors should be able to connect to the PCS-1 if those terminals support the H.239 standard. In the case of H.323 mode, the dual-stream protocol is Sony's original format and the PCS-1 cannot be connected to videoconferencing terminals with dual-streaming function from other vendors. Standardization activities to enable interoperability among videoconferencing terminals from various vendors have started in Study Group 16 of the ITU-T. The first baseline document on the standard was jointly proposed by Polycom and Tandberg at the Rapporteur's Meeting in February, 2003. At the ITU-T meeting held in May, 2003, H.239 has been approved by the Alternate Approval Process (AAP).