Functions of PC Communicator: PC-to-FOMA IP Videophone Technology

Hideaki Takeda†, Dai Ando, Sumitaka Sakauchi, Shigehiko Onishi, and Hirohisa Jozawa

Abstract

This article describes the point-to-point communication functions of the "PC Communicator" Internet protocol (IP) videophone/videoconferencing system developed by NTT Cyber Solutions Laboratories. The PC Communicator provides voice-over-IP telephony between a personal computer (PC) and various types of terminals such as another PC, PSTN (public switched telephone network) phone, cellular phone, or PHS (personal handy-phone system) phone. It also provides IP videophone connections between a PC and a FOMA or ".Phone Personal V" phone from NTT DoCoMo and NTT Communications, respectively.

1. PC communicator

NTT Laboratories and NTT Resonant Inc. have developed the PC Communicator (PCC), an Internet videophone/videoconferencing system that conforms to SIP*1 (session initiation protocol). PCC is the NTT Group's first wholesale type service*2. It is being provided to NTT Communications Corp. by NTT Resonant Inc. as ".Phone Business V" [1]. Commercial services began in March of this year and were upgraded at the end of May.

PCC handles both point-to-point and multipoint calls among personal computers (PCs) and phone terminals, both fixed-line and mobile ones. A detailed explanation of multipoint calls is given in the third article in this Special Feature "Videoconferencing Technology for High-quality PC Multipoint Connections" [2]. Here, we explain the point-to-point communication functions of the PCC.

2. PCC system configuration

PCC is integrated with a video communication service to form the PCC system. This comprises several servers, a FOMA network gateway (FOMA-GW), a public switched telephone network gateway (PSTN-GW), and client software that runs on a Windows personal computer (**Fig. 1**). The servers include a SIP server that administers SIP call control, a presence server that provides a presence service that indicates whether or not a user is online and other such user information, an instant messaging service, and a multipoint videoconferencing server that provides a videoconferencing service. The SIP server (called Type1-CA(SS)), presence server (called Type1-CA(PS)), and PSTN-GW were developed by NTT Network Service Systems Laboratories. The network technology that supports the PCC is explained in detail in the fourth article in this Special Feature "Network Technology to Support PC Communication Services" [3].

3. PCC functions

The functions of PCC are listed in **Table 1**. This

† NTT Cyber Solutions Laboratories
Yokosuka-shi, 239-0847, Japan
E-mail: takeda.hideaki@lab.ntt.co.jp

*1 SIP: An IETF (internet engineering task force) standard protocol for establishing sessions for telephony, chatting, etc. between terminals on an IP network.

*2 Wholesale service: A form of service in which a company uses the services and facilities of a service wholesaler to provide services to others.

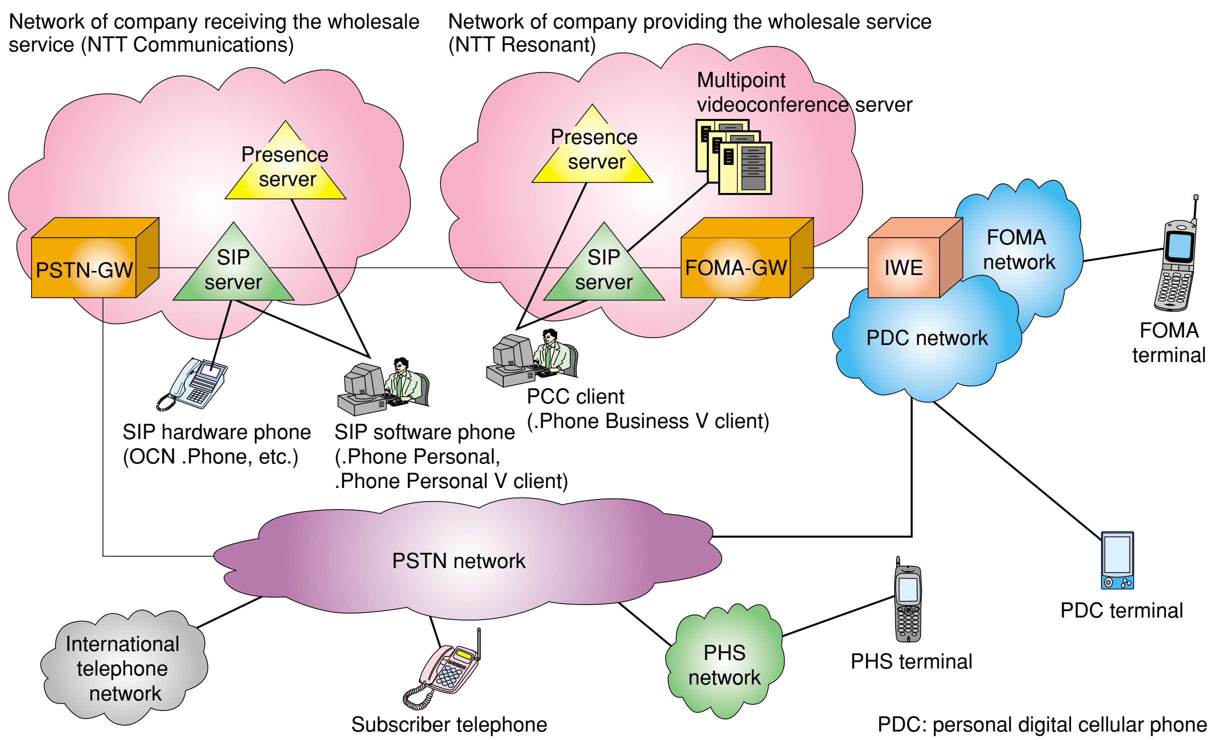


Fig. 1. PCC system and network configuration.

Table 1. Functions of PCC.

Point-to point communication	Voice phone		Provides voice calls to analog phones, IP phones, cellular phones, etc.
	Videophone		Provides video and voice communication between PCC terminals, between PCC and .Phone Personal (V) terminals, and between PCC and FOMA terminals
	Instant messaging (IM)		Provides point-to-point text-based conversation
	Active member list (presence)		Displays log-in status (presence) and history
	Address book		Stores addresses and other user information (calls can also be made from the address book)
	Call history		Saves the history of incoming and outgoing telephone, videophone, IM, and videoconference calls (calls can also be initiated from the history)
	Access control list		Assigns control of telephone and videophone calls and incoming instant messages and presence queries
	Whiteboard		Provides a point-to-point text and diagram sharing function
	File transfer		Provides point-to-point file transfer
Multipoint communication	Videoconferencing		Provides multipoint videoconference
		Videoconference chat	Enables text messaging among participants
		Videoconference materials sharing	Shares materials, whiteboards, and applications among videoconference participants

section describes the main functions used in the point-to-point communication.

(1) Voice phone function

The PCC client conforms to SIP and contains a SIP stack*3 developed by NTT Network Service Systems Laboratories. In addition, an audio echo and noise canceller developed by NTT Cyber Space Laboratories reduces echoes and noise to ensure that the quality of voice calls is high.

*3 SIP stack: A software library that implements the SIP communication protocol.

Table 2. Services compatible with PCC (.Phone Business V).

Service type	Connection type	Destination
Voice phone	IP-IP	OCN .Phone
		OCN .Phone Office
		.Phone IP Centrex
		OCN provider IP phone (050 number) (connection is free of charge)
		.Phone Personal
		.Phone Personal V
	Via PSTN-GW	Other provider's IP phone (050 number) (connection is not free of charge)
		Ordinary phone
		Cellular phone (PDC, FOMA, and voice calls)
		PHS
		Ordinary telephone outside Japan (international call)
Videophone	IP-IP	.Phone Personal
		.Phone Personal V
	Via FOMA-GW	FOMA

The SIP standard enables PCC to provide IP (Internet protocol) telephony between PCC clients and SIP hardware/software phones connected to the SIP servers of the company providing customers with the wholesale service (e.g., OCN .Phone or .Phone Personal) as well as between PCC clients (Fig. 1). The PCC supports connections to both SIP hardware phones (OCN .Phone, OCN .Phone Office, the IP phones of OCN providers, etc.) and SIP software phones (.Phone Personal and .Phone Personal V) registered in the SIP servers of NTT Communications.

Voice communication with analog phones, cellular phones (PDC (personal digital cellular) or FOMA) and PHS terminals is also possible by connecting to the PSTN via the PSTN-GW. The services that are compatible with PCC are listed in **Table 2**.

(2) Videophone function

This function provides IP videophone communication between two PCC clients and between a PCC client and a FOMA terminal. Video data is compressed and decompressed by an MPEG-4 codec developed by NTT Cyber Space Laboratories to reduce the processing load while maintaining high video quality. Between PCC clients, communication with screen sizes of CIF (352 pixels × 288 lines) or QCIF (176 pixels × 144 lines) is possible. In addition, a function for VGA-size (640 pixels × 480 lines) communication has also been implemented (but is not provided for commercial use). During communication, the image size can be changed and communication can also be switched over to voice-only mode. The videophone windows of the PCC are shown in **Fig. 2**.

Video communication between PCC and FOMA terminals passes through the FOMA-GW. Small-scale FOMA-GWs for enterprises have been on the market for some time, but large-scale systems for communication carriers have recently been introduced to produce the world's first carrier service for videophone connections between PC and cellular phone terminals. The function for videophone connections with FOMA terminals, which is another major feature of the PCC, is described in detail in the next section.

The PCC client can make videophone connections to both the .Phone Personal and .Phone Personal V clients of NTT Communications. Because MPEG-4 encoding parameters are exchanged between the sending and receiving terminals in addition to the SIP specifications, mutual MPEG-4 video connections between different types of IP videophone terminals are implemented through combination with the SDP[*4] (session description protocol) description specifications and interpretation method.

To exchange video codec information, SDP only specifies the method for declaring the capabilities; the operations at the receiving terminal side are not specified at all. NTT Resonant Inc. has created common specifications for exchanging capability information for MPEG-4 connections between all the SIP

*4 SDP: An IETF standard protocol for the exchange of session content information. It describes the types of video and voice codecs and their parameters and is used for connection between terminals.

Fig. 2. Example of videophone screen.

terminals offered by companies in the NTT Group rather than just between the PCC client and .Phone Personal V. We intend to work towards greater use of the common specifications among all of the companies in the NTT Group.

(3) Active member list function

This function indicates the ‘presence’ of a person registered in the active member list. That is, it indicates whether the person is online or offline, engaged in a call, etc. (**Fig. 3**).

A member can use the access control list function to specify whether or not his/her presence can be accessed by others. If queries are blocked, no presence information is displayed. In addition to icons that indicate presence status, this function allows text messages to be entered in the presence field.

(4) Whiteboard function

This function enables text and diagrams to be shared with a specified party. It can display materials prepared in advance and lets users write or draw on top of them (**Fig. 4**). The user who initially activated the whiteboard has control of the functions for creating, adding, or deleting shared screens (pages) and incorporating materials, but he/she can choose to pass control over to the other user. The shared screen can also be saved as a bit map or in some other format.

The PCC whiteboard can also be used independently when communication with another terminal is not

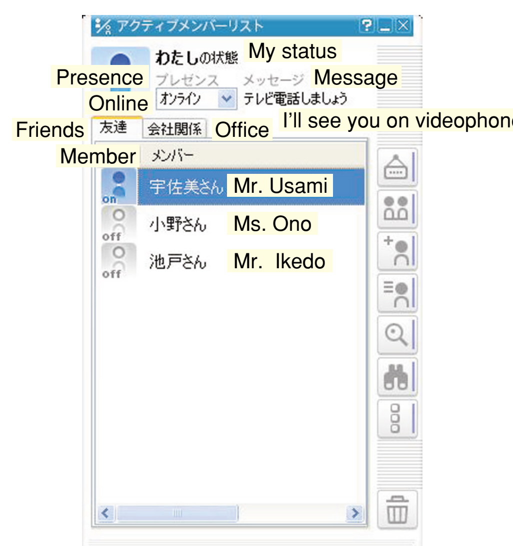


Fig. 3. Screen showing the active member list.

taking place. The call connection is made using SIP/SIMPLE*5 via the presence server.

(5) File transfer function

This function transfers files to a specified other party. In the same way as for the whiteboard function,

*5 SIMPLE: An IETF standard protocol for the exchange of presence and message information.

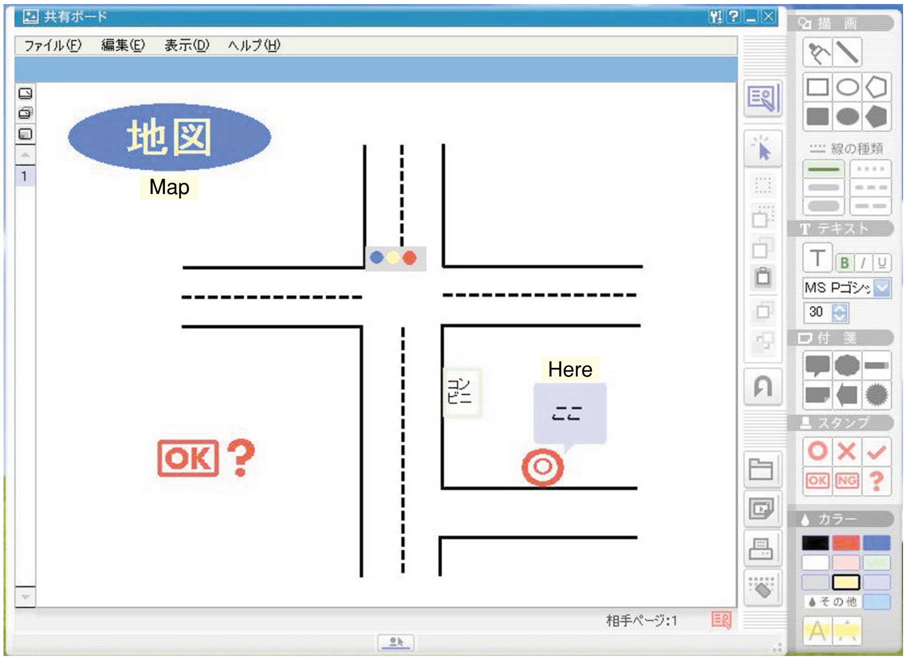
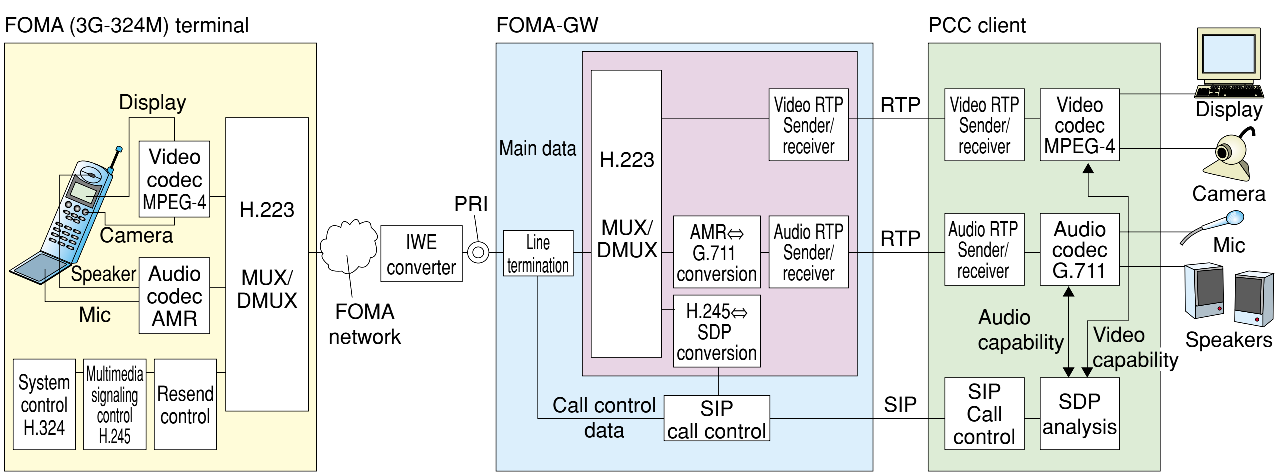


Fig. 4. Screen showing the whiteboard.



H.324: ITU-T Recommendation. A multimedia communication system that uses the analog telephone network. 3G-324M is a specification that is based on the H.324 mobile network version H.324/M.
H.223: ITU-T Recommendation. A low bit rate multimedia multiplexing and demultiplexing protocol.
RTP: realtime transport protocol. An IETF standard protocol for realtime transfer of voice and video over a network.
H.245: ITU-T Recommendation. A procedure for performing format and protocol negotiation for calls between terminals.
AMR: advanced multirate codec. A voice compression and encoding method for flexible variation of transfer rate according to line type and conditions. It is used for third-generation cellular phones based on W-CDMA (wireless code division multiple access).
G.711: ITU-T Recommendation. The PCM (pulse code modulation) voice coding specification that is most widely used for IP phones.

Fig. 5. FOMA connection system configuration.

Table 3. Functions of FOMA-GW.

Call processing conversion	H. 324 (Q. 931) ↔ SIP
Transport conversion	H. 223 ↔ RTP
Media description conversion	H. 245 ↔SDP
Voice transcoding	AMR ↔ G. 711
Video transcoding	None (MPEG-4 through)

Q.931: ITU-T Recommendation. A protocol for establishing connections between terminals.

it can also be used independently, whether or not a call is in progress. The call connection is also made via the presence server. It is possible to transfer files directly (up to 100 Mbytes at a time) without an intermediate file server.

4. FOMA connection service

(1) Connection method

The PCC uses Internet Protocol based on packet switching which differs fundamentally from FOMA's 64-kbit/s circuit switching, so a connection to either of them must go via a FOMA-GW. The system configuration for a FOMA connection is illustrated in **Fig. 5**. The FOMA-GW is connected by a leased line primary rate interface (PRI) to an NTT DoCoMo IWE (interworking equipment). The FOMA-GW converts between the specifications of the 3G-324M and SIP IP phone standards. As shown in **Table 3**, media description conversion for call processing conversion, transport conversion, and codec information conversion between terminals and voice transcoding are performed. The MPEG-4 codec of the PCC also conforms to the FOMA specifications, so transcoding at the FOMA-GW is not necessary. Thus, the processing load on the FOMA-GW is reduced and the channel cost can also be reduced.

The screen size for a FOMA connection is QCIF, which is the upper limit for the 3G-324M specifications. The specifications of the MPEG-4 encoder of FOMA terminals vary from manufacturer to manufacturer, but the PCC has demonstrated connectivity with all devices of the FOMA 2102V series and later (900i, 901i, and 700i).

(2) Concept of use

Videophone connections between PC and FOMA terminals make possible ubiquitous video communication that is unrestricted in time or space. A few examples of this are given below.

- Remote support and instruction from the office for company employees who are out of the office.
- Remote consultation between customers and a call center (e.g., explanation of how to use home appliances and solve problems)
- Remote monitoring to understand the situation at another place of business from the office or to check on conditions at home while away on business
- Broadcasting to the home or office the conditions at an exhibition or during a trip
- Distance learning while at home or traveling

The fusion of the Internet and cellular phones is expected to increase the opportunities for video communication.

5. Future development

We will continue to improve the functions of PCC and add new functions in response to customer needs and utilize NTT research results as well as technology on the market to increase the added value.

References

[1] http://www.dpbv-service.com/ (in Japanese).

[2] H. Takeda, M. Moriuchi, T. Imaeda, S. Honda, and T. Yasuno, "Videoconferencing Technology for High-quality PC Multipoint Connections," NTT Technical Review, Vol. 3, No. 10, pp. 25-30, 2005.

[3] Y. Yamanaka, S. Esaki, A. Shibata, S. Takada, and S. Seto, "Network Technology to Support PC Communication Services," NTT Technical Review, Vol. 3, No. 10, pp. 31-35, 2005.

Hideaki Takeda
Senior Research Engineer, Supervisor, Promotion Project 1, NTT Cyber Solutions Laboratories.
He received the B.E. and M.E. degrees in electrical engineering from Hokkaido University, Sapporo, Hokkaido in 1979 and 1981, respectively. In 1983, he joined the Electrical Communication Laboratories of Nippon Telegraph and Telephone Public Corporation (now NTT), Kanagawa. He has been engaged in R&D of information systems. He is a member of the Information Processing Society of Japan.

Dai Ando
Research Engineer, Promotion Project 1, NTT Cyber Solutions Laboratories.
He received the B.E. degree in information technology from Tohoku University, Sendai, Miyagi in 1989. In 1989, he joined NTT Human Interface Laboratories and has mainly been engaged in developing video communication systems. He is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan.

Sumitaka Sakauchi
Associate Manager, NTT Advanced Technology Corporation.
He received the B.S. degree in physics from Yamagata University, Yamagata in 1993, the M.S. degree in physics from Tohoku University, Sendai, Miyagi in 1995, and the Ph.D. degree in engineering from Tsukuba University, Tsukuba, Ibaraki in 2005. He joined NTT in 1995. Until 2005, he was engaged in R&D of acoustic echo cancellation and noise reduction as a Research Engineer, Media Processing Project, NTT Cyber Space Laboratories. He moved to the present position in 2005. He received the Best Paper Award from IEICE in 2001. He is a member of the Acoustical Society of Japan (ASJ) and IEICE.

Shigehiko Onishi
Manager, Visual Communications Division, NTT IT Corp.
He received the B.E. and M.E. degrees in electrical & computer engineering from Yokohama National University, Yokohama, Kanagawa, in 1989 and 1991, respectively. He joined NTT LSI Laboratories in 1991. From 2000 to 2003, he was with ATR Spoken Language Translation Research Laboratories, Kyoto, Japan. From 2003 to 2005, he was engaged in research on video coding systems at NTT Cyber Space Laboratories. His research interests include audio/speech signal processing and image processing. He is a member of IEICE and ASJ.

Hirohisa Jozawa
Senior Manager, NTT Resonant Inc.
He received the B.E. and M.E. degrees in electrical engineering from Waseda University, Tokyo, in 1987 and 1989, respectively. He joined NTT Human Interface Laboratories in 1989 and was engaged in R&D of video coding systems. He was a Senior Research Engineer, Supervisor of NTT Cyber Space Laboratories, and moved to NTT Resonant Inc. in 2004. He is a member of IEICE and the Institute of Image Information and Television Engineers of Japan.