From Server Access Communications to End-to-end Session Communications

Shinya Tachimoto[†] *and Naotaka Morita*

Abstract

Services using Internet protocol (IP) are shifting from server-dominated communications such as e-mail and web browsing to direct communication between users as in IP telephony. This article explores new issues created by this shift in communication style and discusses the need for and role of new network functions for dealing with these issues.

1. Shift in IP-based communication style

Using the Internet to gather information and communicate is fast becoming indispensable to life and business. Up to now, services on the Internet have been centered on the server-access type of communications such as e-mail, web browsing, and content delivery, in which the server plays the leading role. In contrast, the recent spread of file-exchange, IP-telephony (also known as voice over Internet protocol (VoIP)), and video-chat services signals the birth of a new form of interactive services based on direct communication between end users (end-to-end session communications) [1] (**Fig. 1**). It has become clear, however, that full-scale expansion of end-to-end session communications will require the deployment of new network and terminal functions and solutions to new problems that did not exist with server-access communications.

2. Server-access communications

Server-access communications has been supported by high-performance, high-reliability servers provided by Internet service providers (ISPs) and various kinds of service enterprises. This communication style has evolved dramatically along with the Internet. The services it provides are based on one-way access from the end user to a server, and many of the security, quality-control, and fee-charging problems faced by these services have been solved by equipping servers with a full range of functions. This scheme has therefore minimized the functions needed by user terminals, and this is thought to be a major reason for the dramatic expansion of these services.

3. Birth of end-to-end session communications

From the very beginning, the IP network was assumed to be a means of direct communication between terminals. In reality, however, this required certain terminal functions, network capabilities, and user skills. As a result, direct communication never filtered down to general users to any great degree. In recent years, however, advances in network technology have made it possible to provide high-speed and large-capacity access lines. This, combined with advances in terminals and devices, is making it possible to achieve high-capacity file exchange and bidirectional, realtime communication using high-quality video and audio at a relatively low cost. Services using end-to-end session communications such as IP telephony, remote control, file exchange, and multipoint video chatting are becoming a reality. At the same time, it must be kept in mind that in end-to-end session communications, the end points are terminals of ordinary users, so there are many features that differ from those in server-access communications (**Fig. 2**). These differences lead to new requirements for end-to-end session communications. The requirements need to be understood, and technology for meeting them must be investigated.

† NTT Network Service Systems Laboratories
Musashino-shi, 180-8585 Japan
E-mail: tachimoto.shinya@lab.ntt.co.jp

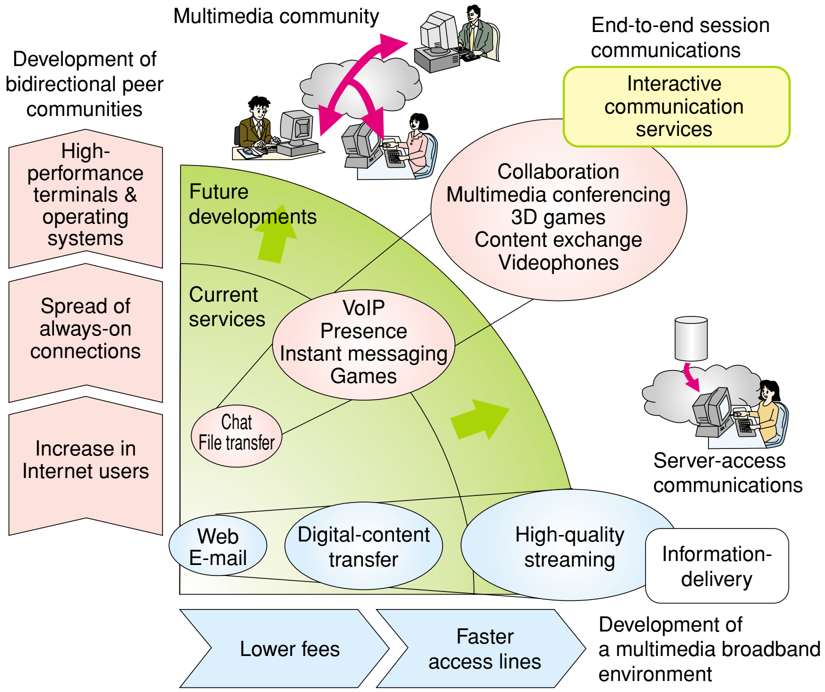
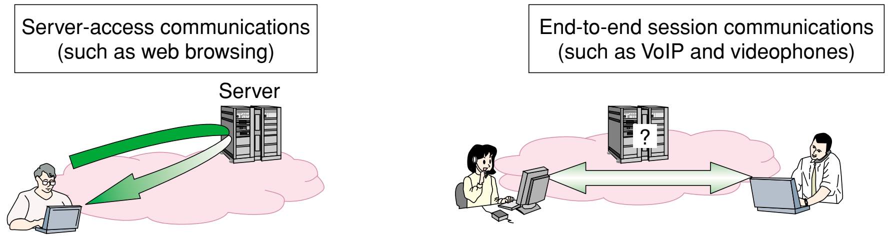

Multimedia community
End-to-end session communications
Interactive communication services
Development of bidirectional peer communities
High-performance terminals & operating systems
Spread of always-on connections
Increase in Internet users
Future developments
Current services
Collaboration
Multimedia conferencing
3D games
Content exchange
Videophones
VoIP
Presence
Instant messaging
Games
Chat
File transfer
Server-access communications
Web
E-mail
Digital-content transfer
High-quality streaming
Information-delivery
Lower fees
Faster access lines
Development of a multimedia broadband environment

Fig. 1. Shift in communication style.


Server-access communications (such as web browsing)
Server
End-to-end session communications (such as VoIP and videophones)
?

Server-access communications	Communication style	End-to-end session communications
Machine (server)	Connected party	Indeterminate number of people
Fixed, public	(IP) address of other party	Indefinite, limited public knowledge
Server presents and user selects	Communication means	Requires negotiation (differs according to users)
User always initiates communication.	Communication scheme	User may initiate communications and may also be called.
Server is presumed to be accessible at any time.	Communication availability	Uncertain (e.g., the terminal's power might be off or the user might be absent or busy)

Fig. 2. Comparison of communication styles.

4. Issues in end-to-end session communications

Table 1 lists the main communication functions and associated issues in end-to-end session communications. In server-access communications, most of these functions are supported on the server side, but in end-to-end session communications, new technologies will be needed to provide them.

Server authentication has traditionally been performed using server certificates while user authentication has been carried out using IDs and passwords managed by servers. But in end-to-end session communications, the number of parties that one can communicate with during one session becomes indeterminate as the system scale increases. This makes it difficult to perform mutual authentication based on

Table 1. Issues in end-to-end session communications.

Phase		Required functions	Current situation in server-access communications	Issues in end-to-end session communications	Solution
Authentication	Search for other party	(1) Provide authentication	Identification is provided by server certificate or user-ID/password.	Communicating with an arbitrary number of parties would require a huge number certificates.	Use network as an authentication intermediary
		(2) Specify address of other party	Server address is fixed and public knowledge.	Fixed and public user addresses could allow illegal access and invasion of privacy.	Use network-based address resolution
Status check		(3) Ensure anonymity (privacy protection)	Operation using anonymous user IDs is relatively simple.	Anonymity might be difficult because IP addresses and certificates would have to be exchanged with other parties.	Use network-based ID management (conversion)
		(4) Check other party's status	Server is assumed to be accessible at any time.	Other party may be unavailable.	Use network-based status management
	Establish communications	(5) Perform congestion control	Access control and load-distribution control are performed on the server side in accordance with current circumstances.	Installing various means of congestion control on an ordinary terminal would be costly.	Use network-based connection control
		(6) Provide authorization	Decision to enable/disable service is made on the server side based on user management information.	Communicating with an arbitrary number of parties would require a huge number of approval conditions to be managed.	Have the network manage approval
		(7) Negotiate communication means	Communication means is specified unilaterally on the server side	Communication means must be negotiated based on the terminals/applications of parties involved.	Provide protocol for negotiation
		(8) Prevent leaking/ alteration of signaling information	Encryption methods are provided on the server side.	Encryption would require procedures for exchanging and checking certificates and for exchanging keys.	Use network-based connection control
Communication	Manage communications	(9) Prevent illegal access	User side can defend against illegal access by using simple policy settings.	Communicating with an arbitrary number of parties would require complex policy settings.	Perform access control linked with connection control
		(10) Achieve quality control	Server side controls amount of transmitted information (content delivery, etc.).	Advanced terminals equipped with flow-control functions and monitoring functions would be required.	Perform quality control linked with connection control
		(11) Prevent leaking/ alteration of transmitted content	Encryption methods are provided on the server side.	Encryption would require procedures for exchanging and checking certificates and for exchanging keys.	Exchange encrypted information linked with connection control

the sharing of certificates and passwords, for example. One style that could provide more scalability than such an inter-user mesh-type authentication style is star-type authentication, in which the network acts as an authentication intermediary (function (1) in Table 1).

The address of a server is usually fixed and public knowledge. For a client, though, a fixed public address could allow illegal access as well as invasions of privacy through name identification. This could be prevented by using two kinds of addresses—public user addresses and addresses to be used only during actual communication—and having the network manage their correspondence (functions (2) and (3)).

Servers are generally assumed to be available at any time, but in end-to-end session communications, it cannot be assumed that a particular person or his/her terminal is always in a contactable state. However, the network could manage user status, notify a user of another user's current state, determine whether a particular user is accepting or rejecting incoming calls, etc., and thereby provide support for congestion control, connection approval, and the like (functions (4), (5), and (6)).

In server-access communications, the applications and protocols available for use are usually specified on the server side. In end-to-end session communications, on the other hand, a set of clients will have to agree upon what application or protocol to use through negotiation, and a protocol for conducting the negotiation will have to be provided (function (7)).

On an IP network, protecting information to be transferred is the responsibility of the parties involved in the transfer. For this purpose, many servers provide various means of encrypting transmissions such as transport layer security. To provide encrypted communication between users in end-to-end session communications, users would have to exchange certificates or common keys beforehand, which, as in function (1), would present a problem in terms of scalability. This problem can be solved by having the network act as a connection-control intermediary between users to provide control signals and encryption functions (functions (8) and (11)).

Using firewalls is an effective means of preventing illegal access. In server-access communications, one-way access from the client side is normal, and this makes it relatively simple to create policy settings. In addition, using commercially available broadband routers without changing the default settings usually provides sufficient security. In end-to-end session communications, though, the user side must process an access attempt from another party, so performing dynamic firewall control to ascertain the identity of other parties would require advanced functions and techniques (function (9)).

Technology for quality control has not yet reached a mature level in server-access communications. In end-to-end session communications too, it will be very difficult because communication paths are set in a full mesh among end users. Proposals for technology that can simplify quality control by linking with user connection control are therefore expected (function (10)).

5. Platform for end-to-end session communications

From Table 1, we can see that existing server-access communications operate on the basis of server-side support for many functions. In end-to-end session communications, deploying these functions in standard user terminals would hardly be realistic from both the cost and operation perspectives. However, using the network as a communications platform to provide these functions might make it possible to promote the development of new services.

Considering the needs for user freedom in performing communications and flexibility in extending services, it is not desirable to have the network participate in inter-user communications at the service level. Ideally, the network, as a communications platform, should provide virtual pipes (sessions) to support inter-user communications in a versatile manner independent of services. Accordingly, providing the solutions shown in Table 1 as a central aspect of such sessions can be said to be one requirement of a network in the era of end-to-end session communications.

6. Network architecture for end-to-end session communications

Session initiation protocol (SIP), as proposed by the Internet Engineering Task Force (IETF), is rapidly becoming popular for controlling sessions [2]. Here, as an architecture for supporting end-to-end session communications, we introduce a model [3] consisting of three planes centered on SIP-based session control (**Fig. 3**).

The lowest plane is the transport plane correspond-

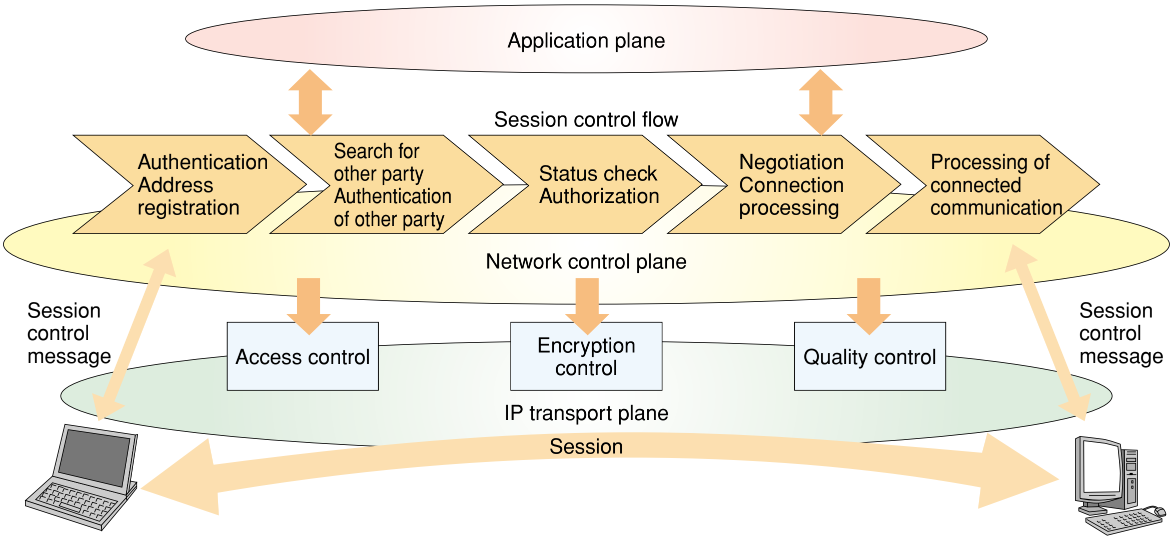


Fig. 3. Network functions supporting end-to-end session communications.

ing to an IP transport network. This plane provides firewall, encryption, and QoS (quality of service) functions with an IP routing function, which has traditionally been provided as a basic service by ISPs.

The top plane is the application plane where diversified communication services are provided interlinked with the sessions established between users. On this plane, we can envision application services that automatically control sessions in accordance with user and service states.

The network control plane in the middle accepts SIP signals from end users and sets up sessions between them. This plane ensures session security, convenience, and quality by concentrating on access control, encryption control, QoS control, and other forms of control associated with session management. A session is set up as a virtual pipe directly connecting end users on the transport plane. Since network equipment is not concerned with the content of any communication here, what information is to be exchanged within a session can be determined by the end users as they see fit.

Based on the above model, the other two articles in this special feature introduce security and quality-control functions that are provided in conjunction with session control and present technologies for achieving them.

References

[1] T. Murakami, "SIP Technology for Driving End-to-end Communication Business and Its Outlook for the Future," 6987th Special Research Forum of the Japan Planning Institute (JPI), Dec. 2002 (in Japanese).

[2] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler, "SIP Session Initiation Protocol," IETF RFC3261, June 2002.

[3] S. Tachimoto, A. Hiramatsu, and T. Murakami, "Network Architecture for Support of End-to-end Communication Services," Technical Report of IEICE, NS2002-259, Mar. 2003 (in Japanese).

Shinya Tachimoto

Senior Research Engineer, Supervisor, Network Software Service Project, NTT Network Service Systems Laboratories.

He received the B.E. and M.E. degrees in mechanical engineering from Tokyo Institute of Technology, Tokyo in 1988 and 1990, respectively. He joined NTT in 1990. He is currently researching session management. His other research interests include the next-generation network architecture, secure end-to-end communications, and high-availability middleware for reliable node systems. He is a member of IEEE.

Naotaka Morita

Senior Research Engineer, Supervisor, Network Software Service Project, NTT Network Service Systems Laboratories.

He received the B.E. and M.E. degrees from Nagoya University, Nagoya, Aichi, in 1985 and 1987, respectively. Since joining NTT Laboratories in 1987, he has been engaged in work on communication protocols and traffic management for ATM and B-ISDN. After a two-year assignment in the strategic network planning department at NTT Headquarters, he started researching IP telephony and interactive multimedia service provisioning over carrier-grade IP networks. He has been an active participant in ITU-T since the early 1990s. Since 1997, he has been a rapporteur of SG 13. He is a member of the Institute of Electronics, Information and Communication Engineers of Japan.