第19部

JBプロジェクト

第 1 章 JB Project

As the Internet moves forward into the next century, the Internet is going to serve as an information infrastructure for everyone from the information infrastructure only for scientists or professionals. The Project JB is a joint high-performance nation-wide network research project in Japan. This project consists of a couple of Network Computing Research Project, such as WIDE Project, ITRC, Cyber Kansai Project and so on. They collaborate and co-operate each other. Furthermore, it has collaborated with international networking research groups, such as Internet2 Project, NGI Project and APAN project. The initial technological targets are IP version 6, quality of service control (e.g., diffserv), realtime high speed stream multicast, reliable multicast and new applications (e.g., School on the Internet). The test-bed is going to integrate all of these functions and protocol stacks into an unified platform based on the PC-UNIX based advanced network stack provided by KAME project.

1. IP version 6
   KAME project and TAHI project have developed high quality reference codes (KAME code) for various UNIX system and system evaluation tools. All the technologies described below have been unified into the KAME code and operates on the JB project network.
2. Quality of Service Control (QoS)
   We have integrates the Diff-Serve (Differentiated Service) functions into the ALTQ packet scheduling module. Also, in order to manage the QoS policies across the Internet, we develop the Bandwidth Broker (B.B.) using the COPS.
3. High Speed Stream Multicast
   We have developed the PIM-SM multicast routing protocol for KAME code, to deploy Digital Video (DV) class high speed stream multicast over JB project network.
4. Reliable Multicast
   The integration of FEC (Forward Error Correction) and packet retransmission has been researched and developed for a large scale reliable multicast.
5. New Applications
   School on the Internet is a applications using the advanced technologies listed above. JB project has developed and operated the realtime and interactive international remote lecture system among University of Wisconsin and JB project network across the pacific ocean.

Practical network operation with the integration of the above functions and protocol stacks will disclose some new issues and requirements for the next generation Internet infrastructure. The deployed test-bed internetworks with international research networks, such 6REN and Q-Bone.

第 2 章 Introduction

Internet technology provides global, ubiquitous and universal connectivity for all computers and everyone connecting through various datalink platforms. Since connectivity is the Internet's own reward, the Internet has been growing numerically and geographically at a more than exponential rate. The core technology for the Internet is IP (Internet Protocol) and TCP (Transmission Control Protocol). A network in the Internet contains computers that are interconnected using IP. As well as interconnecting computers within a network using IP, IP further interconnects these networks. This is reason why it is said that the Internet is a network of networks. Reliable data transmission is achieved by the end-to-end TCP control mechanism between source and destination hosts. Therefore, it is said that Internet with TCP/IP is inherently a distributed system. As the Internet moves forward into the next century, the Internet transits to an information infrastructure for everyone from the information infrastructure only for scientists or professionals. This means that the next generation Internet has to achieve the following features.

- Internet for everyone
- Internet for everything
- Internet everywhere
- Internet at anytime
- Internet any way

Now, what are the technological requirements and challenges for the next generation Internet. It is "scalability". Scalability has various aspects from various quantitative and qualitative view points. The initial technological target of the "JB" project is IP version 6 (IPv6), quality of service control, multicast and new applications.

- IPv6

  IPv6 provides sufficiently larger address space (128 bits) for the next generation Internet. IPv6 also provides various new functions compared with the legacy of IPv4 (IP version 4). For example, IPv6 enables autoconfigurations for end hosts to accommodate everyone into the Internet.

- Quality of service control

  Each person and each application may require different communication quality levels. Differentiated Service (Diff-Serv), Integrated Service (Int-Serv) and RSVP are the key architecture models and protocol stacks which provide various communication qualities for each particular packet flow over the Internet. The quality of service control must be scalable, regarding the heterogeneity of the required service class and of the condition of the network.

- Multicast

  The unicast based multicast emulation, that is the general solution in the existing Internet environment, may not be numerically scalable to need to multicast capability. Multicast has two different technological challenges. One is for realtime high speed stream data transmission, e.g., digital video (DV) data multicasting. The other is for error free data delivery for large number of receivers, i.e., so-called as reliable multicast.

- New applications

  Using the new technologies described above, new applications should be researched and developed. For example, in the "JB" project, the School on the Internet (SOI) applications using the above technologies will be deployed. Also, we have demonstrated the global and realtime collaborative music concert (i.e., LIFE).

We, the "JB" project, will integrate the above mentioned functions and protocol stacks into a unified platform based on the KAME stack Also, we operates these above mentioned functions over a newly deployed research test-bed, called the "JB" project test-bed. The "JB" project is a joint research project to explore next generation Internet technologies. The project is organized by several network computing projects, such as the WIDE project, ITRC project and Cyber Kansai Project. Therefore, we build a new nation-wide high performance research test-bed, that is jointly operated by the participating research projects and organizations. The deployed test-bed also internetworks with international research networks, such as 6REN (IPv6 Research and Educational Network) and Q-Bone (end-to-end Quality of service backBone) Network operation with the integration of these functions and protocol stacks described above will disclose some of the new issues and of requirements for the next generation Internet infrastructure. Section "Project overview" gives brief description for project overview, test-bed topology and its architecture. Section "IPv6", "Multicast", "QoS" and "New application" give an background, roadmap and work items of each sub-project. Section "Conclusion" gives brief conclusion.

第 3 章 Project overview: design and architecture

The goals of the "JB" project are development new technologies to construct a nation-wide next generation Internet infrastructure and to point out the problems that will be disclosed through its operation. Furthermore, it has collaborated with international networking research groups, such as Internet2 Project, NGI Project and APAN project in IPv6, Reliable Multicasting, QoS area and new applications. Initially, the "JB" project focuses on the following three themes, IPv6, Multicasting and Quality of Service, as three independent sub-projects that collaborate with each other. As an early phase objective of the "JB" project, those sub-projects aim to establish a test-bed network that has integrated these three themes. Then, research and development of new applications will begin after the test-bed is available. In order to build a high-speed and high-performance test-bed, "JB" has been designed adopting high-speed lines such as ATM and SDH leased line as a datalink. The rest of this section, will describe the following:

1. Formation of this project
2. Network topology and technologies

3.1 Formation of JB project

The "JB" project has several sub-projects, such as IPv6, QoS, Multicast and new application, and these are able to increase or decrease depending on the situation. In other words, these sub-projects are able to integrate and divide. As this behavior can be likened to piling up various layers, we designated that each sub-project be called "plane" in the "JB" project. The figure fig 3.1 ("multiple-planes unite into one plane") shows that these "planes" will in the end be unified into one single "plane". This is exactly what the next generation Internet infrastructure is all about.

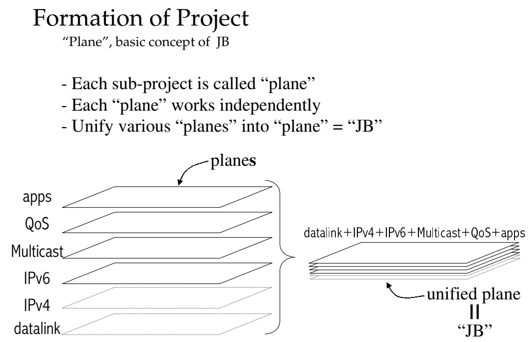

図 3.1 multiple-planes unite into one plane

In a technological viewpoint, it shows a typical model of the process that the Internet grows to the Infrastructure for "Universal Service". Especially, things that related with all planes are important, such as a fusion between old technologies and new technologies, a network management that covers each planes.

3.2 Network Topology

There are at least 10 NOCs (Network Operation Center), such as KDD Otemachi, University of Tokyo, Keio SFC(Shonan Fujisawa Campus), Osaka University, Kyoto University, NAIST(Nara Advanced Institute of Science and Technology), JAIST(Japan Advanced Institute of Science and Technology), CRL(Communication Research Lab., MPT), Kurashiki and Kyushu University, and at least 20 leaf-sites have connected to the test-bed network. The figure fig 3.2 ("Japan Backbone: nation-wide backbone") shows the initial topology of the "JB" test-bed network. Most of connectivity has been provided

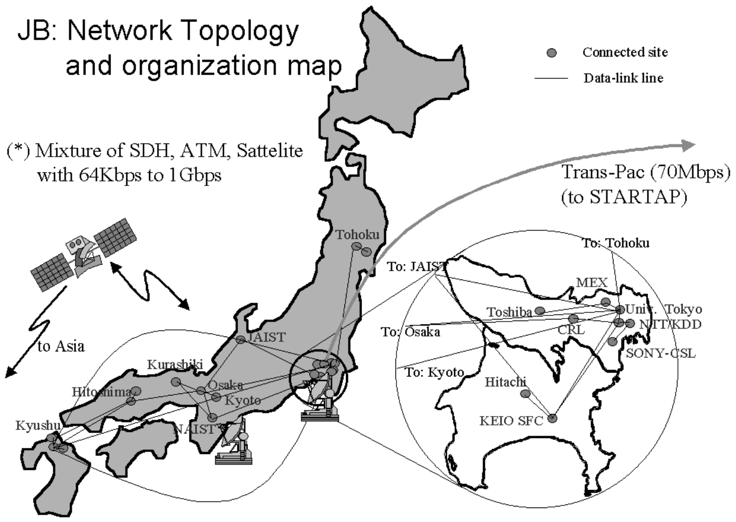

図 3.2　“Japan Backbone”: nation-wide backbone

by ATM links(OC-3 to OC-48), though SDH links have also been used in parts. The KAME stack based PC-UNIX boxes have been installed as both hosts and routers to all ”JB” sites.

第 4 章　JB/6: the “IP version 6” plane

The goal of ”IPv6” plane is building a IPv6 network environment for next generation Internet infrastructure.

4.1 Background

The IPv6 (IP version 6) is a core protocol of the next generation Internet. It has a 128bit address space that is enough to cover all worldwide networks. IPv6 core stacks have already been in development for four years. During these years, a lot of network equipment vendors and research and development organizations have developed IPv6 stacks on various platforms. As a result, most of those stacks are available as products or public domain. Likewise, three years have passed since the significant test bed for IPv6 stack evolution, development and deployment, known as the 6bone, has started operation. As a result of the 6bone, we have developed important technologies such as dynamic routing, source address selection and multihoming as well as having acquired operational tips and experience. However, if those technologies are not deployed generally throughout the world, IPv6 can not be called ”Internet Protocol for Next Generation”in its true sense. Therefore, ”IPv6 plane” is going to focus on a establishment of a network environment using these technologies as commodity operation.

4.2 Roadmap

Through the constructing and operating of a test-bed network for IPv6, that is the next generation core protocol for the Internet, on ”JB”, the ”IPv6 plane” is going to tackle the following technical issues:

- Establishment of transition technologies from IPv4 to IPv6
- Verification of interconnectivity among various IPv6 products
- Verification of IPv6 unicast and multicast

routing

- Building up applications for IPv6

Furthermore, these issues are also applicable to the deployment of IPv6 in the worldwide Internet. Especially with the IPv4-IPv6 transition issues, through actively providing an IPv4 and IPv6 coexisting environment for all "JB" sites, we will positively find out and solve new problems, and provide feed back to the IETF ngtrans working group. Finally, we believe that the current 6bone-jp should be replaced with "JB" as its successor, in order to change it from a simple test-bed to a IPv6 nation-wide, general and commodity infrastructure. This is because for the most part, it consists of complex and ad hoc IPv6 over IPv4 tunnels. At present, we are going to install KAME stacks, as both of router and end-host, "Toshiba CSR", that has been improved based on KAME stack, and "Hitachi GR" as routers to all "JB" sites.

4.3 Work items

Starting from May 1999, we have focused on whether these three individual research planes, IPv6, Multicasting and Quality of Service were able to work with each other on the same network. In other words, IPv6 sub-project had to provide IPv6 connectivity in order that the other planes can work on IPv6. All "JB" sites have already had IPv6 capablities each other by July 1999. Currently, we can develop a new application on IPv6 platform at all of the "JB" site. For example, DVTS is one of the most interest application which has developed on IPv6. It is able to provide a function to transfer "Digtal Video stream" over IPv4/IPv6 with/without IPsec/Mutilcast. We often use it as Demonstration of "JB" project or remote lecture between United States and Japan. We have cooperated with operation experiments for commercial IPv6 products by allowing them deployment as soon as they are ready for use. Our initial experiences for experiment were for the "Toshiba CSR" and "Hitachi GR" which are able to operate with high-speed lines such as OC-3 and OC-12. Several OSPF version 3 implementations are in progress. They have been tested starting from 4Q of 1999. At present, there are two candidates for developing platform, one is "gated", the other is "zebra". As both of them are available in the public domain, we will also make available results of development, such as codes and documents. As the Internet will have to go through the same transition in near future, the transition of the main IPv6 test-bed of Japan from 6bone-jp to "JB" is a most interesting event, not only for us, but also for IPv6 researchers all over the world. Actually, this transition is going on with at present, and we get several interesting experiences. Therefore, we have to feed back these experiences to the IETF community.

4.4 Performance Evaluation of Data Transmission using IPSec

This section evaluates the performance of data transmission using the ordinary PC both for large sized data transmission and for the actual application. As an actual application, this paper picks up the DV (Digital Video) transmission. This is because the Digital Video will be a common and widely used application media in the emerging internet. Also, the video contents requires the secured and reliable data transmission. For large sized data, when we apply the authentication (AH) and encryption (ESP), the throughput degrades to 1/9 comparing with the throughput without AH nor ESP. With AH and ESP, we obtains about 10 Mbps for UDP data transmission and about 6 Mbps for a simple TCP transmission. As for the DV transmission, the end-to-end throughput was again about 10Mbps. With 10 Mbps end-to-end throughput, the 1/10 of video information can be successfully transferred from the source node to the destination node to obtain a sufficient quality of DV transmission.

4.4.1 Research Backgroud

The next generation Internet has to achieve the scalable and reliable data transmission. The IPv6 (IP version6) and IPSec (IP Security) is a core protocol suite for it. IPv6 has a 128bit address space that is enough to cover all worldwide networks and equipment, and IPSec technology provides a essential functions for reliable and secured data exchange over the Internet.

The purpose of this section is a performance evaluation of data transmission with the IPSec over IPv6 networks using an ordinary PC platform. People wants to perform a high speed multimedia communications, such as high quality video communications, with a low cost PC platform. Also, when we go to the production level of multimedia services over the Internet, we have to apply the authentication and encryption to protect themselves and the information exchanged among them.

4.4.2 IPSec (IP Security)

RFC 2401 describes the architecture framework of IPSec (IP Security). IPSec protocol suite provides the functional suite for secured and reliable data exchange over the Internet. IPSec has the following two functions, i.e., Authentication and Encryption.

- Authentication

  The authentication is a user validation of the communication peer. In order to validate (authenticate) the user, IPSec defines the AH (Authentication Header) in RFC2402. AH field contains the digital signature calculated by the sender node. The receiver node validates the digital signature in the received AH field. When the AH contains the valid digital signature, the received packet is correctly delivered to the corresponding application module. On the other hand, when the AH does not contains the valid digital signature, the received packet is discarded. With the AH, the receiver can receive the packets only from the authenticated node/user.

- Encryption

  The encryption is to allow to be able to read a data by only person(s) who has (have) the correct encryption key. In order to encrypt the user data in the IP packet, IPSec defined the ESP (Encapsulating Security Payload) in RFC2403. ESP field contains the encryption parameters to identify the encryption scheme between the communicating peer nodes. The payload of IP packet contains an ESP header, that has the encryption parameter, and the encrypted user data. In other word, the encrypted user data is encapsulated into the container, whose header is an IP header and an ESP header.

Both for IPv4 and for IPv6, the IPSec is independent from type of data transmission medium. Also, the application does not care whether the IPSec is applied to or not. For IPv6, IPSec is defined as a mandatory option, i.e., every node has to have the IPSec function.

We have a concerning with regard to the performance of IPSec. As well known, the required processing for security functions are not light, rather would be large. When the execution of security function (i.e., IPSec) requires very large processing power, we could not obtain an enough throughput for many applications. Or, we have to implement the special hardware to handle those security functions. When the ordinary PC platform can provide enough processing power to handle the IPSec for major applications, we can deploy the secured and reliable information infrastructure, cost effectively. Now, therefore, the purpose of this paper is performance evaluation of IPSec with ordinary PC platform.

4.4.3 Performance Evaluation of Bulk Data Transmission

In this section, we evaluate the performance of bulk data transmission. The performance is evaluated with the STREAM data transmission and with the REQUEST/RESONSE data transmission. Regarding the transport protocol, both TCP and UDP are applied to.

Evaluation System

The end-to-end throughput was evaluated using the netperf with the KAME IPv6 protocol stack. The patch for to use netperf 2.1pl3 with KAME IPv6 stack is available through the following ftp directory.

```
ftp://ftp/kame.net/pub/kame/misc/netperf-21pl3-19990824.diff.gz.
```

The end host are connected through the two routers. All nodes have the fast Ethernet interfaces. The followings are the specification of hosts and routers.

- Host
  - CPU : Intel Pentium II 450 MHz
  - Memory : 128MB
  - NIC : Intel EtherExpress Pro 100
  - OS : FreeBSD 2.2.8 with KAME 19990908-stable
- Router
  - CPU : Intel Pentium III 500MHz
  - Memory : 256MB
  - NIC : Intel EtherExpress Pro 100, DEC 21040 PCI Ethernet
  - OS : FreeBSD 2.2.8 with KAME 19990908-stable

The end-to-end throughput is evaluated in the following cases.

- without IPSec
- Only with AH (Authentication Header, HMAC-SHA1 160 bits)
- Only with ESP (Encapsulated Security Payload, 3DES-CBC 192bits)

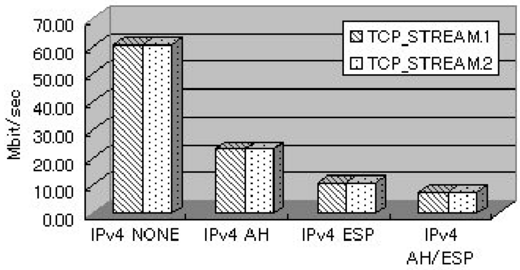

図 4.1 IPv4 TCP STREAM

- Both with AH and ESP

For all cases, the performance is evaluated, in the cases of using IPv6 and using IPv4. Also, the performance is evaluated with TCP and UDP transmission with two modes. One is STREAM data transmission, and the other is REQUEST/RESPONSE data transmission. Sender host executes netserver, and the receiver host executes netperf. The data is transmitted for 60 minutes.

Evaluation Results

[1] TCP STREAM

Figure 4.1 shows the end-to-end throughput using IPv4, and figure 4.2 shows that using IPv6. Here, the MTU size is 4,096 Bytes, the socket size is 57,344 Bytes and is 32,768 Bytes.

- TCP-STREAM.1 : socket buffer size = 57,344 Bytes
- TCP-STREAM.2 : socket buffer size = 32,769 Bytes

The end-to-end throughput is degraded by the processing of IPSec. With the AH, the end-to-end throughput degrades to about 1/2. With the ESP, the end-to-end throughput degrades to about 1/4. With both the AH and the ESP, the end-to-end throughput is slightly less than with only ESP.

Regarding the IP version, the end-to-end throughput with IPv6 is almost the same as that with IPv4.

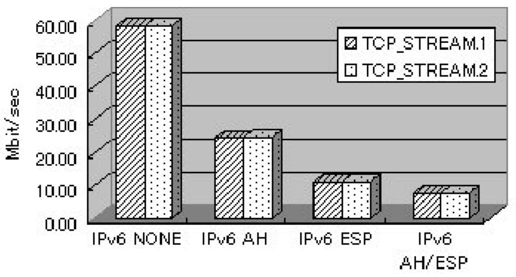

図 4.2　IPv6 TCP STREAM

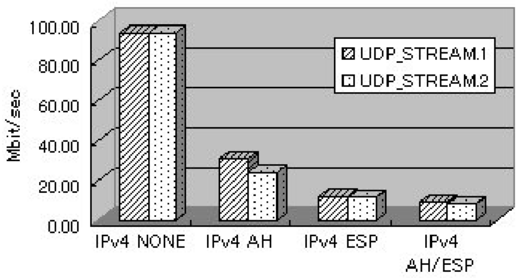

図 4.3　IPv4 UDP STREAM

[2] UDP STREAM

Figure 4.3 shows the end-to-end throughput using IPv4, and figure 4.4 shows that using IPv6. Here, the MTU size is 4,096 Bytes and is 1,024 bytes, and the socket size is 32,768 Bytes.

- UDP-STREAM.1 : MTU size = 4,096 Bytes
- UDP-STREAM.2 : MTU size = 1,024 Bytes

The end-to-end throughput is degraded by the processing of IPSec. With the AH, the end-to-end throughput degrades to about 1/3. With the ESP, the end-to-end throughput degrades to about 1/9. With both the AH and the ESP, the end-to-end throughput is slightly less than with only ESP. Also, when the MTU size is larger, the end-to-end throughput is slightly improved. Again, regarding the IP version, the end-to-end throughput with IPv6 is almost the same as that with IPv4.

[3] REQUEST/RESPONSE

Figure 4.5 shows the end-to-end throughput using IPv4, and figure 4.6 shows that using IPv6.

- TCP-RR : Request message = 1 Byte, Response message = 1 Byte

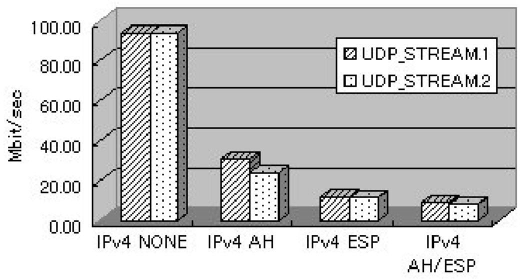

図 4.4　IPv6 UDP STREAM

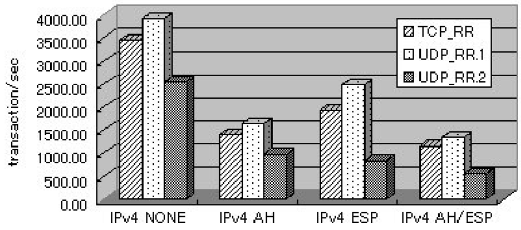

図 4.5　IPv4 Request/Response

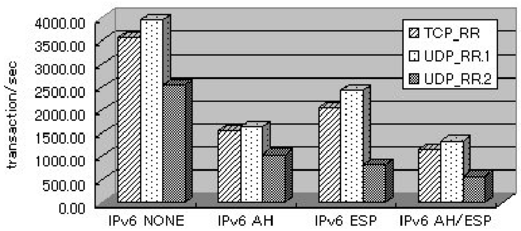

図 4.6　IPv6 Request/Response

- UDP-RR.1 : Request message = 1 Byte, Response message = 1 Byte
- UDP-RR.2 : Request message = 516 Byte, Response message = 4 Byte

With TCP-RR and UDP-RR.1, the throughput with ESP is larger than the throughput with AH. With UDP-RR.2 (i.e., large message size), the throughput with ESP is smaller than the throughput with AH. And, again, regarding the IP version, the end-to-end throughput with IPv6 is almost the same as that with IPv4.

Discussion

- STREAM Data Transmission

  By the AH and ESP processing at the end hosts, the end-to-end throughput degrades, when we use the IPSec. AH performs the hash function, and ESP performs encryption.

Without applying the IPSec, the end-to-end throughput over TCP is less than the end-to-end throughput over UDP. This is because the TCP requires larger processing than the UDP does. However, when we use the IPSec, the end-to-end throughput over TCP and over UDP is almost the same. This is the simple proof, that the processing for IPSec is far larger then that for TCP and UDP.

When we compare the end-to-end throughput with AH and with ESP, the throughput with AH is about twice larger than that with ESP. This is because, with STREAM data transmission, the packet size is large enough than the header length (basic IP header field and AH field), and the required processing for ESP is far larger than that for AH. As shown in figures, when the MTU size becomes larger, the end-to-end throughput degrades. The degradation with ESP is smaller than with AH, since the AH uses IP packet header field and the ESP uses a whole of payload in IP packet. Also, since the processing for ESP is large enough than that for AH, the performance degradation from only ESP system to ESP/AH system is not large.

In summary, for large sized data (i.e., STREAM data transmission), when we apply the authentication (AH) and encryption (ESP), the throughput degrades to 1/9 comparing with the throughput without AH nor ESP. With AH and ESP, we obtains about 10 Mbps for UDP data transmission and about 6 Mbps for TCP transmission. Also, the throughout was compared with the data transmission with IPv4. The degradation of throughput at the end system, due to the use of IPv6 instead of IPv4, was significantly small.

- REQUEST/RESPONSE Data Transmission

  With the REQUEST/RESPONSE data transmission, the end-to-end throughput degradation by applying the IPSec is less than that with the STREAM data transmission. This is because the processing overhead for request messages is not significantly small, compared to the processing overhead for IPSec, and because the packet size is not large.

4.4.4 Performance Evaluation of DV Data Transmission

In this section, we evaluate the end-to-end DV data transmission over IPv6 network. The reason why we pick up the DV transmission is that the DV will be a common and widely used application media in the emerging internet. Also, the video contents requires the secured and reliable data transmission.

Evaluation System

The end host are connected through the three routers and wide area high speed ATM links. Nodes have the fast Ethernet interfaces and the ATM interfaces. The followings are the specification of hosts and routers. In order to send and receive the DV data, we use the DVTS developed by Keio University.

- Sender Host
  - CPU : Intel Pentium II 400 MHz
  - Memory : 64MB
  - NIC : Intel EtherExpress Pro 100
  - OS : FreeBSD 3.2 with KAME 19990810-stable
  - Sender software : DVTS-0.0.9

- Receiver Host
  - CPU : Intel Pentium II 450 MHz
  - Memory : 128MB
  - NIC : Intel EtherExpress Pro 100
  - OS : FreeBSD 3.2 with KAME 19990810-stable
  - Sender software : DVTS-0.0.9
- Router
  - CPU : Intel Pentium III 500MHz
  - Memory : 256MB
  - NIC : Intel EtherExpress Pro 100, DEC 21040 PCI Ethernet
  - OS : FreeBSD 2.2.8 with KAME 19990908-stable
- ATM Link
  - Bandwidth : 155 Mbps

The end-to-end throughput is evaluated in the following cases.

- NONE : without IPSec
- AH1 : only with AH (Authentication Header) using HMAC-SHA1 160bits
- AH2 : only with AH using KEYED-SHA1 160bits
- ESP1 : only with ESP (Encapsulated Security Payload) using 3DES-CBC 192bits
- ESP2 : only with ESP using BLOWFISH-CBC 192bits
- AH/ESP 1 : both with AH1 and ESP1
- AH/ESP 2 : both with AH2 and ESP2

Evaluation Results

Figures 4.7 and 4.8 shows the end-to-end throughput.

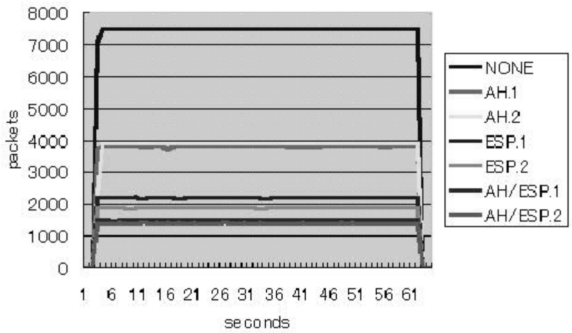

図 **4.7** Performance of DV data transmission using IPSec (packets)

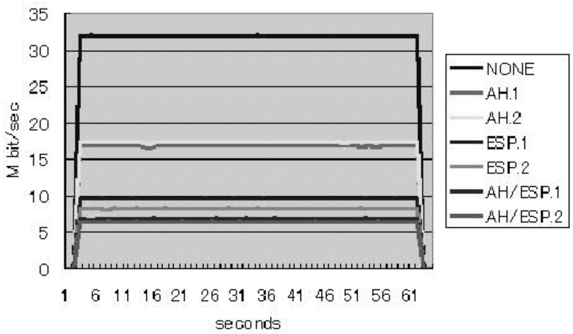

図 **4.8** Performance of DV data transmission using IPSec (MBytes)

As for the DV transmission, the end-to-end throughput was about 7 Mbps to 17 Mbps with IPSec. With 10 Mbps, we can not transmit the full rate quality of DV data. However, we can transfer the DV data with reducing the sending frame rate. The DVTS, that is DV transmission and receiving software module, can control the frame rate sent out from the sender node. With the evaluation result, the 1/10 of video frames are transferred from the source node to the destination node, so as to the required bandwidth is around 10Mbps. Even with 1/10 of video frame rate, we can not obtain a fine quality of video transmission. However, we can obtain a sufficient video quality for many applications, such as video conferencing, with this reduced frame rate.

The result above shows that the existing ordinary PC platform could handle the high quality video transmission using the DV technology, without any special hardware assistance. This is even when we apply the IPsec technology to provide the secured and reliable multimedia communication over the Internet. The current ordinary PC platform can not handle the plain DV data with the full frame rate while applying the IPSec functions. However, due to the fast technological improvement for the PC components (e.g., CPU), it would be expected the ordinary PC platform can handle the full frame rate DV data without any special hardware.

第 5 章　JB/M: the “Multicast” plane

The goal of ”Multicast” plane is establishment of reliable technologies at a worldwide scale. For example, to provide multicast communication capabilities for high-quality multimedia streaming such as video (cf. ”Digital Video”) and audio(cf. ”Internet Telephony”), and for guarantee of reliability.

5.1 Background

The most suitable situation to make use of the striking features of multicast such as ”reduction of redundant traffic” and ”similarity to existing broadcasting media” is high band width communication within a widely distributed area. Therefore, the development of IP multicast technology on super high-speed networks is necessary for the deployment of multicast communication in general. The MBone, which is a worldwide test-bed for IP multicasting, has long supported the development of a lot of IP multicasting technologies such as ”DVMRP” which is a multicast routing protocol based on ”distance vector” algorithms, ”RTP” which is a multimedia transport protocol for realtime communication and some reliable multicast protocols for reliable communication. In this way, the MBone has grown into worldwide test-bed and has been connected with hosts of various performance and nature. However, there is an important problem. The MBone consists of many tunnels laid out on an Internet based on ”best effort”. Because of the overhead needed to maintain these tunnels the Mbone is not suited for functions as a high band width multimedia communication test-bed. Therefore, the ”JB” has built a live test-bed for multicasting to make use of ”JB” test-bed, and experiment on multicast communication with the next generation Internet environment in mind. Furthermore, another important study concerning IP multicasting technology is the development and deployment of a scalable multicast routing protocol such as ”PIM”(Protocol Independent Multicasting) and to find out and solve the problems discovered during this process. In other words, it is for these reasons that IP multicasting has not seen worldwide deployment despite its superior technological value.

5.2 Roadmap

On the ”JB”, IPv6 is ready to operate from the beginning as stated at section ”IPv6”. We assumes that it is the core protocol for development and operation. When considering scalability, we hypothesize several different size scales for the Internet and develop and deploy several appropriate multicast routing protocols for each accordingly. Practical reliable multicast protocols are also going to be designed and implemented by making use of QoS guarantee technology. And as application of these, technology for multi-point delivering will be developed. Finally, the ”Multicast” is going to focus on the establishment of high speed datalinks, IP technology, Multicast routing and applications taken as a whole on the next generation Internet Infrastructure. In other words, the technologies that have been developed by ”JB” will be adoptable on the next generation Internet immediately.

5.3 Evaluation of high speed multicast data transmission

This section evaluates the robustness of multicast routing and the performance evaluation of multicast service using the experimental network. For the performance evaluation of multucast service, we use the the DV (Digital Video) transmission service, as the practical application for the next generation Internet. The evaluatoin system uses the PIM (Protocol Independent Multicast), that is an RPF (Reverse Path Forwarding) multicast routing protocol. The packet loss observed at the end host is evaluated, compared with the case where the system provide the multicast service using N of the multiple unicast packet transmission. We can make sure the multicast system has better system scalability, regarding the number of receivers, than the unicast based multicast system has. PIM operates correctly and stably, even when the network has a routing loop.

5.3.1 Backgroud of this research

The most of the data communications over the currently operating Internet is using the unicast data transmission with TCP or UDP. And, many exisiting applications do transfer exactly the same data to large number of receivers. Also, the upcomming Internet applications will includes a realtime multicast services, such as DV (Digital Video) program multicasting service or interactive multi-party multimedia conferences or games. We have to establish a stable and cost effective multicast data transmission infrastructure for the next generation Internet applications.

With a multicast data transmission, when a sender transfers a single packet toward the particular multicast group, the network copies the packet to transfer them only to the particularreceivers belonging to the corresponding multicast group. Since the sender does not need to send the same data to every receivers, the network can provide an efficient multicast data transmission services. With the multicast data transmission, the required processing power at the sender and the required bandwidth for the sender can be smaller than those with the unicast data transmission. Since the total number of packtes transferred by the network with multicast data transmission is smaller than that with unicast data transmission, the possibility of inappropriate routing protocol operation due to the large amount of user packet transmission with multicast data transmission would be less than that with unicast data transmission. Also, the diversity of the data reception time by the receivers must be smaller that with the unicast data transmission, and the data reception delay by the reviewers must be smaller than that with the unicast transmission.

On the other hand, the multicast service may have the following technical concernings.

- the performance of packet transmission at the router may degrade, due to the copying of multicast packets. This is because the copying the received packet at the router is a new function for the router.

- the network congestion due to the routing loop with multicast packet will be worse than that with unicast packet. This is because the number of links and nodes affected by the routing loop with multicast packet transmission should be larger than that with unicast packet transmission.

The evaluated system in this paper has the following features;

- Digital Video (DV) multicasting

  DV is one of the popular applications with high speed data transmission used in the next generation Internet. Also, the DV is an application, that is sensitive to the packet loss and the delay jitter.

- PC based router and host

  For flexible and cost effective network operation, the experimental system uses the conventional PCs for routers and hosts.

- PIM-SM multicast routing protocol

  For effective and scalable multicast service, we used the PIM-SM (Protocol Independent Multicast - Sparse Mode) routing protocol.

- Native IPv6 network

  Since the IPv6 technology is the fundamental core protocol for the upcomming Internet, the experimental system operated over the IPv6 network.

We implemented the PC router, that can operate the PIM-SM with IPv6 and the PC host, that can send and receive the DV data using the PIM-SM with IPv6. Using these PC based routers and hosts, we developed both the local area and the wide area experimantal networks. The wide area exeprimental network accommodate the live traffic with the conventional applications, as well as the experimental DV multicast application. Also, the experimental network was not a loop free topology. Therefore, the experimental network would create a transitional routing loop with some network status.

Using the experimental network, we evaluate the performence of packet transmission, compared with the performance where the multicast service is provided by the multiple unicast packet transmissions. Also, we evaluate the stability of routing protocol operation with high speed DV data transmission. It was shown that the experimental system works well with the high speed DV data multicast application, even when the system has a routing loop. As a further study item, it is recognised that we should evaluate the effectiveness of the packet scheduling mechanism to provide different quality of service among system control information (e.g., routing protocol information) and user information.

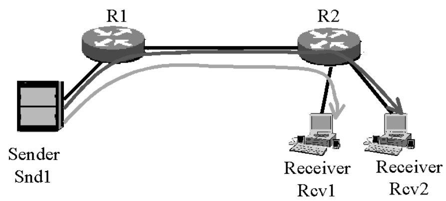

図 5.1 Experimantal Configuration for Multicast by Multiple Unicasts

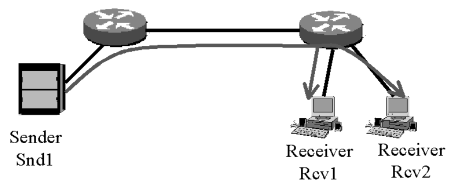

図 5.2 Experimantal Configuration for Multicast by Multicast Session

5.3.2 Performance Evaluation of Packet Transmission in Multicast Network

Evaluation System

Figures 5.1 and 5.2 show the evaluation system. Sender (Snd1) sends the packets to two receiver hosts (Rcv1 and Rcv2) through the routers (R1 and R2). Rcv1 and Rcv2 receive the same data from Snd1, i.e., multicating from Snd1 to Rcv1 and Rcv2. R1 and R2 are connected through the OC-3 ATM link, and the hosts (Snd1, Rcv1 and Rcv2) are connected to the routers through the 100MBase-T links.

In the configuration of figure 5.1, Snd1 estblish two (unicast) packet flows to each receiver host. This is the case where the multicast service is provided by the multiple unicast connections. Routers (R1and R2) do not need to copy the received data, and the sender host (Snd1) copies the sending data. In the configuration of figure 5.2, Snd1 establish one (multicast) packet flow to the receiver hosts. This is the case where the multicast service is provided by a single multicast connection. The router R2 copies the received data to deliver the data to receivers (Rcv1 and Rcv2). Here, with the configuration of figure 5.2, PIM-

SM is applied to as the multicast routing. Sender host (Snd1) can control the packet transmission rate to the network, with the rate control (i.e., shaping).

All nodes, that are routers and hosts, are the ordinary IBM compatible PC with the following specification.

- Host
  - CPU : Intel Pentium II 450 MHz
  - Memory : 128MB
  - NIC : Intel EtherExpress Pro 100
  - OS : FreeBSD 3.4 with KAME 19990908-stable
- Router
  - CPU : Intel Pentium III 450MHz
  - Memory : 128MB
  - NIC : Intel EtherExpress Pro 100, ENI OC-3 256MB ATM
  - OS : FreeBSD 3.4 with KAME 19990908-stable

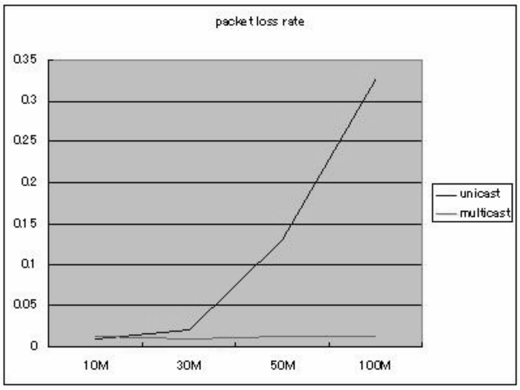

図 **5.3**　Packet Loss Rate at the Router

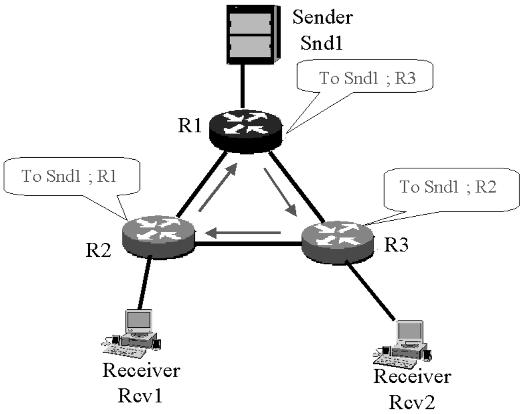

図 **5.4**　Experimantal Configuration for Evaluation of Multicast Routing Against Routing Loop

Evaluation Results

Figure 5.3 show the packets loss ratio at the router (R1), parameterizing the packet transmission rate from the sender host (Snd1). As shown, the router can not relay the unicast packets, in accordance with the increase of packet transmission rate from the sender with unicast based multicast. However, with the multiast service, the router can relay the packets and the packet dropped rate does not increase even when the packet transmission-rate from the sender host (Snd1) increases.

5.3.3 Evaluation of System Robustness of Multicast Routing Protocol

Overview of RPF Multicast Routing

The RPF represents Reverse Path Forwarding. RPF is commonly used in the major multicast routing protocols, such as PIM that is used in the evaluation system discussed in this paper. In RPF system, the multiast packet transmission is executed using the unicast routing information. The multicast session is recognized and managed by the pair of sourece node IP address (”S”) and IP multicast address assigned for the multucast group (”G”). When the router receives the multicast packet, the router checks the source IP address in it.

```
if (Recv_I/F = Nxt-Hop_R_I/F)
  {forward the received packet to the rest of interfaces}
 else
  {silently discard the received packet}
endif

%  Here,
%   (1) the Recv_I/F : the interface (including the virtual
%       interface such as ATM or Frame Relay),
```

```
%       where receives the multicast.
%   (2) Nxt-Hop_R_I/F : the interface, where the packet to
%       the source IP address shown in the received multicast
%       packet should relay based on the unicast routing
%       information
```

Therefore, the RPF multicast routing does not need any additional routing mechanism for a multicast service, i.e., it only needs the unicast routing information. Also, when the IP multicast packet is received from the wrong interface compared to the unicast routing information, the received multicast packets are automatically and silently discarded.

This mechanism avoids the packet forwarding to the routing loop. When the unicast routing system has the routing loop, the packet transferred into the loop shall discarded due to the mismatch of unicast routing information to the source IP address and the interface receiving the IP multicast packet. This function of RPF multicast routing has a significant benefit, compared with the unicast based multicast service. With the unicast based multicast service, the packet has to be looped in the network until the TTL expires. However, with the multicast service using the RPF, it is expected that the packets are discared at the entry router to the routing loop.

Evaluation System

Figure 5.4 shows the evaluation system. Sender (Snd1) sends the multicast packets toward two receiver hosts (Rcv1 and Rcv2) through the routers (R1, R2 and R3). All routers run the PIM-SM multicast routing protocol. All the data links in the system are 100MBase-T.

In the evaluation system, the R1 has the inappropriate unicast routing information. For R1, the next hop router to transfer the packet toward Snd1is the R3. Therefore, the routing toward the Snd1 has the routing loop among R1, R2 and R3.

Evaluation Results

The multicast packets are transferred to the first hop router R1. R1 has to discard the all the received multicast IP packet from the Snd1 node, since the interface receiving the IP multicast packet is different from the interface RPF expected. We manually modify the routing information at R1, so that the evaluation system has a routing loop among R1, R2, and R3. By the applicatin of the RPF mechanism, it is expected that the R1 silently drop the multicast packets.

The experimental system actually discard the IP multicast packets at the R1, when the R1 has the wrong unicast routing information to form a routing loop. When we use the multiple unicast connection to provide the multicast service from Snd1 to Recv1 and Rcv2, the unicast packets do loop among the R1, R2 and R3. Moreover, the packet to be looped shall be two packet flows (both toward Recv1 and toward Rcv2). We did make sure the routing loop has generate the congestion with unicast based multicast service. And, also, we did make sure the avoid of packet transmission at the entry router of the (unicast) routing loop, to avoid the network congestion due to the unicast routing loop.

5.4 DV Multicast using PIM-SM

At the beginning, we are going to focus on establishing a multicast capable test-bed of a scale of about 20 sites on "JB", using PIM/SM as the multicast routing protocol. Also, we are going to focus on a establishment of technologies for high band width multimedia transport such as Digital Video(DV) streaming. Starting from April 1999, we had focused on a simple implementation of PIM/DM multicast routing protocol daemon. It had been available by June 1999, and then we deployed it to all "JB" sites. From July 1999, we has began to develop PIM/SM daemons and continue to experiment through the test-bed. Currently, we use PIM/SM as primary multicast routing protocol. November 1999, we had succeeded "Digital Video" multicast streaming over IPv6 from Kurashiki to several JB NOC/leaf site for WIDE Project meeting. It provided reasonable quality to discuss and did not occur any fatal problem at

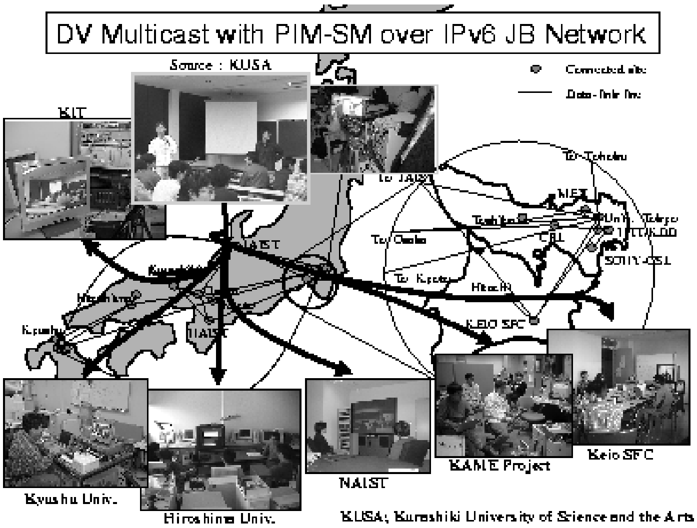

図 5.5　Overview of DV Multicasting over the JB Project Network

all.

Figure 5.5 shows the overview of DV multicast service over the JB project IPv6 network.

第 6 章　JB/Q: the Quality of Service plane

The goal of ”QoS” plane is building a quality control capable network for next generation Internetinfrastructure.

6.1 Background of QoS Technology

There is a growing demand to multiplex voice, video and data into a single infrastructure, without sacrificing scalability of the Internet. The IETF’s differentiated services (diffserv) working group has been trying to solve the problem and standardize protocols for the differentiated services mechanism. In the diffserv framework, interpretation of the IP Type-of-Service field has been redefined so that various queueing mechanisms and packet-drop mechanisms can be deployed at routers. However, the standardization of an IP-based QoS framework does not tell us how best to classify packets and prioritize them within the given traffic mix of voice, video and data. While a lot of work has been done in the area of QoS, very little work focuses on flow aggregation and their probabilistic QoS-guarantees. Little is known about the probabilistic QoS-guarantee on aggregation of flows; there is a large possibility that network engineers must fill the gaps between application requirements and the underlying mechanisms, such as traffic class, drop preference, queueing mechanism and so on. The opaque nature of the diffserv framework leaves a good number of design alternatives for network engineers. While the standard can be directly adaptable to enterprising networks demanding assured forwarding of voice traffic or expedited forwarding of business transactions, the scale and multiplexing nature of the Internet makes it difficult to apply the standard to the backbones of the Internet. For example, an ISP must be able to handle such questions as if we prioritize a specific customer, what kind of quality will the other customers receive? If diffserv is going to be widely available across the Internet, we must have at least one working service model for the differentiated services.

6.2 Roadmap

"JB" is going to bridge the gaps between application practitioners and protocol designers by providing a live QoS infrastructure. Our focus is on both software implementation and network deployment, since it enables fast feedback to software/protocol development process. Our goal is as follows:

1. provide open-source implementations for a QoS infrastructure
2. accumulate working knowledge of QoS-enabled network design
3. develop algorithms as well as implementations for intra-domain QoS-routing, admission control, bandwidth allocation across multiple administrative domains, and flow aggregation at domain boundary
4. provide live QoS-enabled infrastructures to researchers in various science fields

6.3 Work items

Starting from April 1999, we has focused on intra-domain QoS frameworks. Our work spans across routing, forwarding, QoS signalling, and admission control. The "Bandwidth Broker" is name of one of the framework for QoS has developed by a couple of organization, such as "Internet2". After a one-year focus on intra-domain QoS frameworks, we will be able to move on to implementation and deployment of inter-domain QoS frameworks, starting from around June 2000. We are going to develop a service model incrementally through actual deployment. Our initial applications are Internet telephony and Digital Video transmission, as well as live IPv6 data traffic. Several people are studying the effects of flow aggregation at DS boundary routers. The primineary COPS (Common Open Policy Service) implementations has been completed. They has been tested at the WIDE camp experimantal network in March 2000. Interactions between different types of clients and servers, as well as service location mechanism, will be tested here. DSCP marking at DS boundary routers is another issue. While it has been easy to identify packets based on TCP ports, it becomes impossible with IPsec encapsulated packets. Interaction between COPS, IPv6 flow-label and DSCP will be studied here.

第７章 New application, "Internet Application for Next Generation plane"

The goal of "New application" plane is implementation and experiment of application system/software for Next Generation Internet. There are two themes are in progress. One is "DVTS" which consists of two programs, "Digital Video Sender" and "Digital Video Receiver". The other is School of Internet, which is learning and education program based on realtime and/or archived remote lecture system.

7.1 DVTS

The "DVTS" is system which provide Digital Video transmission ability using normal PC and DV VTR/Camera. DV streaming needs 25Mbps bandwidth to transmit it as full rate. It is reasonable application to evaluate high-speed/high-performance network environment for next generation infrastructure. Since DVTS is well-design and often used, it can work on both IPv4 and IPv6, and also can work with both IPSec and Multicast. So, we are able to apply it various situations.

7.2 SOI

The SOI (School on the Internet) project, Working Group in WIDE Project was started in September 1997, considering the change of role in the Internet. Our goal is "to provide higher education and opportunity for all the people in the world who have the will to study using Internet based technologies, eliminating traditional limitations." In October of the same year, we have started WIDE University "School of Internet" as a testbed of our new concept. (http://www.sfc.wide.ad.jp/soi) Since then, we have been doing researches such as education based on the Internet, development of the university environment, and research about the education system of the new age. "School of Internet" is the studying environment to learn about the Internet on the Internet. It is difficult for just one educational organization to gather enough teachers that can teach about this whole new subject and also provide sufficient educational environment for people who want to learn about the Internet systematically. The establishment of "School of Internet" will be an important guideline to set up this new educational field by coordination of different universities. The activities of "School of Internet" are all on the Internet. Lectures and speeches provided on the Internet by demand are all about Internet and computers, which are given by the professors of the WIDE Project. 1,415 people, including actual university students and adults, entered this university. The average number of access to the lecture page per month goes up to 200,000.

第 8 章　Conclusion

The project "JB" is an advanced network research and educational project in Japan. It has focused on "high speed", "high bandwidth", "well-designed" and "reliable" technologies for the next generation Internet infrastructure. It especially focuses on "IPv6", "Multicast", "QoS" and "New application" as a "plane" for this project. It is therefore called "Japan Backbone" which is one of the things which "JB" stands for. Although, we described the background, roadmap and work items at each section of "plane", the "JB" project's work is still in progress. And, it will provide robust, high-quality reliable and high-speed service for everyone in near future.