R&D Activities and Testbed Operation in WIDE Project

Hiroshi Esaki
The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku,
Tokyo, 113-8656, Japan
hiroshi@wide.ad.jp,

Akira Kato
The University of Tokyo,
2-11-16 Yayoi, Bunkyo-ku,
Tokyo, 113-8658, Japan
kato@wide.ad.jp,

Jun Murai
Keio University
5322 Endo, Fujisawa-shi
Kanagawa, 252-8520
junsec@wide.ad.jp

Abstract

The WIDE project, www.wide.ad.jp, is a research consortium among industry and academia. WIDE project is consist of more than four hundred active researchers from more than hundred private companies and more than forty universities. The WIDE project operates the nation-wide R&D testbed, which is connected with many other R&D testbeds, and the NSPIXPs, that are the largest IXes in Japan. The main focus of WIDE project has been IPv6 technology and has established IPv6 testbed since 1998. Basically, all of R&D activities in WIDE project are based on the IPv6. This paper describes the overview of WIDE project R&D testbed and the R&D activities, such as KAME/USAGI/TAHI projects.

1. Introduction

The Internet technology provides the global and ubiquitous digital communication platform using the various types of datalinks. At the first stage, the Internet had been established and operated for the researchers and scientists using the computer technology. Through the deployment of Internet technology to the industrial activities and people's daily life, the Internet technology has mutated to a common and essential infrastructure and platform for all of people, including peoples who would not be familiar with the computer technology. The digital infrastructure assumes the IP technology as the core and common platform, that flexibly adopts various needs required by the various digital media and applications. To achieve goal, the Internet must continue the mutation and revolution through the invention and the introduction of new technologies. In order to develop these new technologies, we need a practical live testbed, which accommodates the traffic generated by users' (researchers, engineers and ordinary users) daily activities. On this testbed, the researchers and engineers can evaluate and validate the new technologies invented and developed by themselves. Through this practical operation, the technical issues will be newly realized and recognized, so as to work on the new research and development items. This is a positive spiral among the testbed and R&D. The emerging Internet shall have the following features :

- Internet for everyone
- Internet for everything
- Internet everywhere
- Internet at any time
- Internet anyway

Since 1998, the WIDE project has established the full scale nation-wide IPv6 testbed, in order to accelerate and integrate the R&D activity on the IPv6 technology. On the WIDE IPv6 network, the operating system, middleware and applications developed by the researchers in the WIE project has been evaluated and has been widely used by various research projects and industry in the global internet community. This paper describes the overview of the WIDE IPv6 project and the testbed

2. WIDE IPv6 Project

2.1 Structure of WIDE IPv6 tested

The WIDE IPv6 testbed must be a part of global IPv6 network. The testbed has been interconnected with various other R&D network, e.g., APAN, APII, AI3, Abilene. The WIDE IPv6 testbed as been tightly collaborate with the ITRC (Internet Technology Research Consortium) and the CKP (Cyber Kansai Project). This collaboration, including the collaborating testbed operation, has been called as JB project. The WIDE IPv6 testbed and the JB project testbed has been overlaid on the following platforms.

- JGN(Japan Gigabit Network, www.shiba.tao.go.jp)
- APII testbed (www.apii.net/)
- APAN/TransPac (www.apan.net/)
- AI3 testbed (www.ai3.net/)
- BBCC (www.bbcc.or.jp/)

2.2 R&D activities on the Platform Technology

2.2.1 IPv6 Core Technology

(1) IPv6 Protocol Stack

WIDE project has been progressed IPv6 R&D activities since middle of 1990's. Especially, the following three core project has been established.

- KAME project (www.kame.net) ;
  IPv6 protocol stack for BSD Unix
- USAGI project (www.linux-ipv6.org/);
  IPv6 protocol stack for Linux
- TAHI project (www.tahi.org/);
  IPv6 evaluation and test suit

KAME project has merged and integrated the IPv6 protocol stack development activities among WIDE, INRI and NRL, and has been adopted by almost all of BSD UNIX distribution, e.g., BSDI, FreeBSD. The protocol stack developed by the USAGI project has started to be merged into the Linux main software tree since April 2002. TAHI project has hosted interoperability test events, three time, and has been cooperated with PLUG group of ETSI and Connectathon (connectathon.org)..

(2) Routing Protocol Suit

KAME project works on routing protocol for IPv6, as well as core IPv6 protocol stack. We have developed our own routing protocol suite (i.e., RIPng, OSPFv3 and BGP4+), collaborating with IP Infusion Inc. (www.ipinfusion.com/), and we have daily operated on the (large scale) testbed.

(3) DNS System

DNS system is a global scale distributed directory service between IP address and FQDN (Fully Qualified Domain Name). As for DNS system, we have following two projects. One is the design and operation of thirteen route DNS servers, collaborating with USC-ISI and ICANN. Here, WIDE project has the responsibility on the operation of M route server. The other project is "bind" for IPv6. "bind" is a free of source program running in on DNS server. We have collaborated with ISC ISC(Internet Software Consortium, www.isc.org/), which distribute the "bind".

Fig.1 Logo; WIDE/KAME/TAHI/USAGI

2.2.2 Network Management

We have worked on the network management system based on SNMP, so as to collect the various management information effectively and promptly in the large scale network with large number of nodes. It is a challenge for this system that the synchronization of time, the synchronization of poling and the reduction of traffic volume generated through the collection of management information. Our network management system has been operated since December of 1999. This system can collect the traffic information on both IPv4 and IPv6 (Fig.2). For the dual stack operation, we have worked on the enhancement of SNMP. One is the SNMP transport on dual stack environment, and the other is an SNMP agent that can transform among multi-protocol environment. These results have been proposed to the IETF.

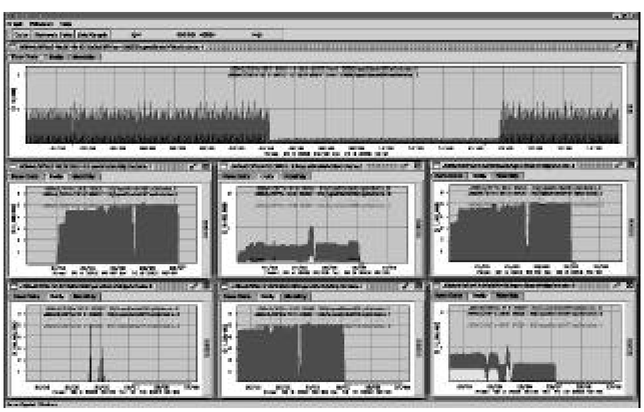

Fig.2 Example of Management View

2.2.3 Multicast Technology

We have two types of testbed operations on multicast technologies; one is for large scale multicast using PIM (Protocol Independent Multicast) and the other is for small group multicast using XCAST. Also, we have worked on the research on a reliable multicast using the FEC (Forward Error Correction) technology.

(1) Large scale multicast using PIM

The first operation on WIDE project testbed was on November 27 in 1999, as shown in Fig.3. The technical conference held at Kurashiki University of Science and Arts had struggled to be multicasted toward more than ten sites in nation-wide Japan, using the DV (Digital Video) stream. We have realized a lot of technical issues, e.g., memory run out at the PC routers or multicast packet flooding at the Ethernet switches. The DV stream multicasting was an unidirectional. In March of 2002, the interactive multicast remote conference among USA and Japan, collaborating by Internet2/Abeline, FLA (Fujitsu Lab. America) in Maryland, NTT Communication MCL in Palo Alto and WIDE project. Now, we are considering how and when we introduce the SSM (Source Specific Multicast).

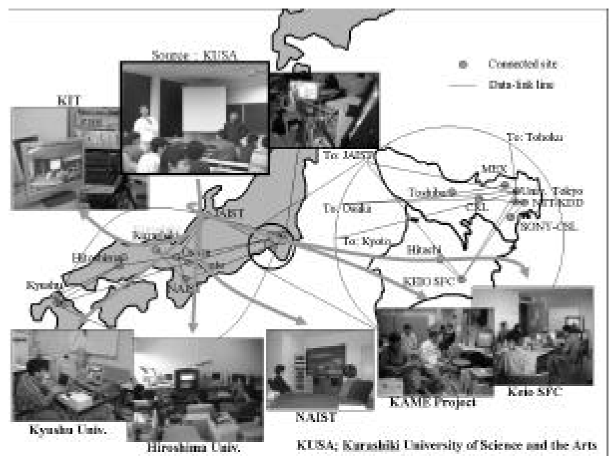

Fig. 3. DV Multicast with PIM-SM

(2) XCAST (Explicit Multicast)

XCAST (www.alcatel.com/xcast/) explicitly describes the receivers' IP addresses in every IP packet header, rather than using ISM (Internet Standard Multicast) address. We have proposed XCAST as MDO6 (Multicast Destination Option on IPv6), that uses IPv6 destination option. XCAST has a scalability associated with the number of multicast groups, which does not includes large number of receivers. XCAST has operated collaborating with research organizations in various countries (e.g., Korea).

2.2.4 MPLS for Distributed IX (http://www.distix.net)

An IX (Internet eXchange) is a system to interconnect many AS(Autonomous System) networks to each other efficiently. Each ISP installs their BGP boarder router at each IX point so as to establishes large number of peerings with other ISPs at each IX point. The issues of existing IX are (1) all boarder routers connected with the IX point have to use the same link technology (e.g., Ethernet or ATM), (2) each ISP has to prepare a high speed link to connect with the IX. When the IX points were not widely distributed, the link cost for remote ISP from IX point would be expensive. This is typical issue in Japan, since the IX is centralized in Metropolitan Tokyo (and Osaka). In order to solve these issues, we have proposed and operated the distributed IX system using the MPLS technology, called as MPLS-IX. We have operated the nation-wide testbed, that is over-layed over the JGN (Japan Gigabit Network) and is interconnected with commercial and regional ISPes. MPLS -IX provides a peer-to-peer LSP link using any datalinks among ISP's boarder routers. The MPLS -IX is going to be extended to support IPv6 capability. IPv6 networks interconnect each other over the IX network (Fig.4). After Edge LSRs establish LSPs between each other, those Edge LSRs assign a LSP to some IPv6 destination. In this figure, LSR-1 assigns a LSP for the destination of network B. When LSR-1 transmit data packets, all packets travel through the LSP. That is, Core LSRs have no need to understand IPv6 protocols nor IPv6 routing information, in this model.

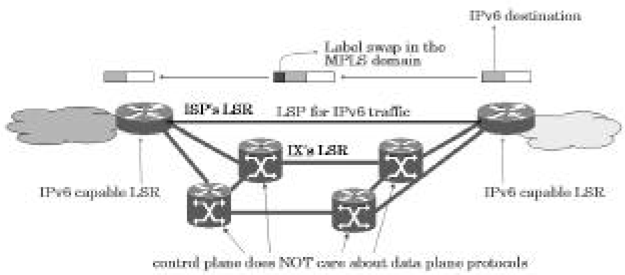

Fig.4 MPLS-IX operation for IPv6

2.2.5 DVTS (http://www.sfc.wide.ad.jp/dvts/)

Digital Video (DV) is a high quality video image compression format for TV quality. DV uses an intra-frame compression for video signal, and uses an 12bit or 16 bits PCM for sound signal. From this point of view, the DV signal would be tolerant for packet losses regarding the video image transmission, comparing with the video transmission using the other compression mechanism (e.g., MPEG2). We have developed the DVTS (Digital Video Transmission System), that encapsulates the DV signal via IEEE1394 interface into the IP packets to transmit over the Internet. DVTS runs with RTP (Real-time Transmission Protocol) both over IPv4 and IPv6 protocol stack. DVTS consumes about 35Mbps bandwidth. However, when the available bandwidth between source and destination nodes are not enough large, the video frame can be periodically discarded at the sender node to reduce the required bandwidth for DV transmission. DVTS runs over various platforms, e.g., BSD-UNIX and Linux with X-window system, Windows with direct-X and Macintosh.

2.2.4 Traffic Measurement

MAWI (Measurement and Analysis on the WIDE Internet) working group has monitored live traffic in WIDE project and developed AGURI (Aggregation-based Traffic Profiler) [6], that effectively analyzes the long period traffic data. AGURI can aggregate and summarize the traffic data, while remaining the particular feature of traffic, for arbitrary period of monitored traffic data. AGURI uses Patricia -Tree algorithm and LRU (Least Recent Used) methodology, in order to keep the traffic data constant volume. Figure 5 shown an example of ARURI's output. AGURI could be used for DoS detection or traffic engineering.

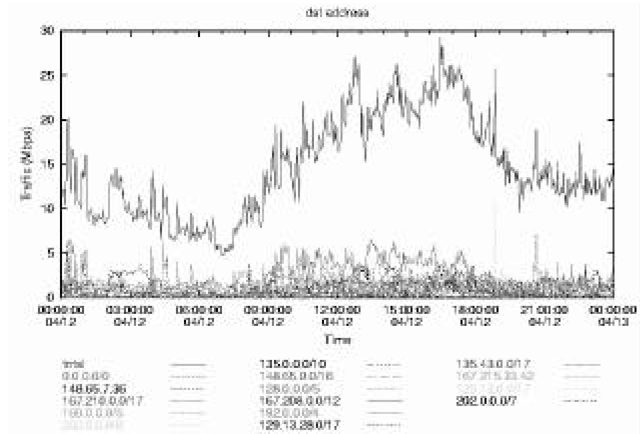

Fig.5 Example of AGURI Output

2.2.5 Packet Scheduling

QoS (Quality of Service) and CoS (Class of Service) is one of policy instances for the policy-aware network. In the policy-aware network, there are various network policy including the QoS or CoS. We have researched on the policy networking and has developed the intra-domain bandwidth broker system, that can provide QoS or CoS for each application flow. The developed system uses an COPS (Common Open Policy Service) server to manage the policy and control information. For a packet scheduling at every node, we use the ALTQ system [6], that has been developed by WIDE project and has been widely used in various projects, such as Q-Bone. ALTQ is a platform that can easily integrate various packet scheduling mechanism, such as CBQ, WFQ or RED. ALTQ has been integrated into KAME implementation, to be available Diff-Serv functionality in IPv6 environment.

2.2.6 Mobile Technology

(1) LIN6(Location Independent Networking for IPv6
LIN6(www.lin6.org/) applies the LINA (Location Independent Network Architecture), that has the node identifier field and the node locator field. Since the TCP/IP session can be managed only by the node identifier field information, while not including the locator field information, the LIN6 can work in a mobile environment and in a multi-home environment. This is because the locator identifier is only for the routing of IP packet and can be changed during a given session. LIN6 has several advantage, compared to the Mobile IPv6 discus at IETF.

(2) GLI (Geographical Location Information)
GLI is a middleware system to manage the location information of each mobile node. Mobile node registers it's location information to the GLI server. The node accessing to the mobile node sends the query messages to the GLI server. There are two types of query. One is asking the location by the node identifier. The other is asking the node identifier and it's location by the geographical area. Global GLI system can operate with

hierarchical configuration and have access control policy for privacy consideration. In order to obtain the geographical information, we have worked on a pseudo-light system with differentiated GPS mechanism. In a pseudo light system, we installed pseudo satellite station on the earth (e.g., at the buildings) so that mobile gear can catch multiple GPS (satellite) signals.

(3) Network Mobility

We have initiated the network mobility work at the IETF, collaborating with INRIA. The architecture would be an extension and modification of mobile IP, that can provide a node mobility. All of these research activities would be integrated into the InternetCAR project, described below.

2.2.7 Authentication and Certification

Technology moCA working group, that is a working group working on CA (Certificate Authority), has operated the WIDE ROOT CA since September of 1998. This CA is a parent CA for several CA system, e.g., moCA for WIDE project members and SOI CA for students for the School of Internet project, running around WIDE project. The WIDE ROOT CA has participated in the "Challenge PKI", that is interoperability test among multiple CAs, in order to accelerate the researches on bridging and cross certific ation.

2.2.8 UDLR (Uni-Directional Link Routing)

The routing protocol running in the Internet assumes that the link is bi-directional and symmetric. However, when we look at satellite link, it is uni-directional. We have discussed the UDLR technology at the IETF collaborating with INRIA, since 1997. Simultaneously, we have established the satellite internet testbed using the UDLR technology. The testbed is called as AI3(www.ai3.net/), that is collaborating with CRL(www.crl.go.jp) and JSAT (www.iijnet.or.jp/JSAT/), and is covering the Asian countries as shown in Fig.6. The new research item on UDLR system is IPv6 and multicasting for production quality operation.

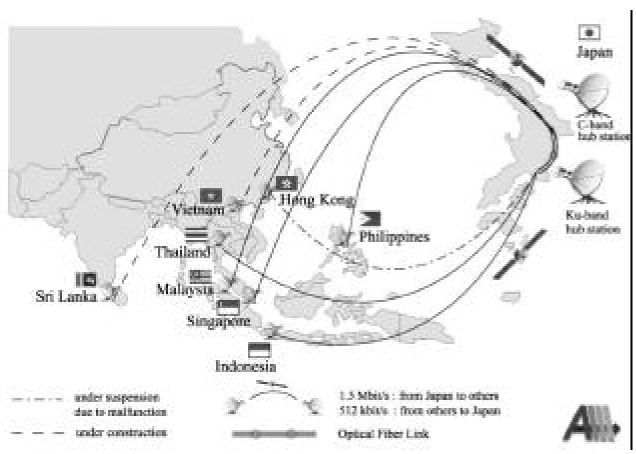

Fig.6 AI3 Service Footprint

2.3 Application

2.3.1 Remote Collaboration on Microscopy [7]

The researchers at NCMIR(School of Medicine's National Center for Microscopy and Imaging Research) of USC San Diego and the researchers at the Research Center for Ultra -High Voltage Electron Microscopy (UHVEM) of Osaka University are remotely collaborating in real-time using the high resolution video images. In order to provide the infrastructure for this collaboration, ITRC, WIDE project, Abilene, vBNS and SDSC(San Diego Supercomputer Center) has worked together. This project is one of iGrid project.

2.3.2 SOI (School of Internet) Project

The virtual university, SOI, has been established on the Internet since 1997. The students of SOI has participated in through various network platform. As shown in Fig.7, the SOI testbed has been across the Asian Pacific countries. SOI operated two studios in USA, i.e., FLA (Fujitsu Lab. America) in Maryland and NTT Communications MCL in Palo Alto, and is covering the Asian Pacific countries via satellite link collaborating with AI3 project. This SOI project collaborating Asian-Pacific area is called as SOI-AP (SOI - Asian Pacific) Project.

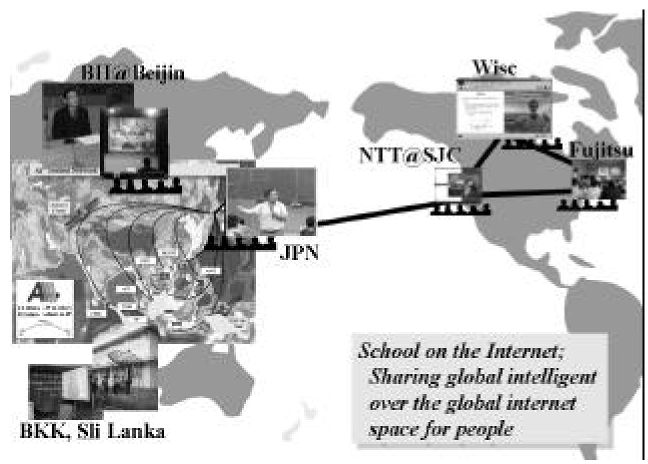

Fig.7 Global SoI System

2.3.3 IAA (I Am Alive) Project

Internet system should be able to provide a lot of contributions and has to securely operate, when a disaster occurs. IAA system is a distributed database system, that collects and provides the information of survivors. IAA system must be dependable platform, as well as privacy-aware system. IAA project has a large scale drill once in every year.

2.3.3 InternetCAR

InternetCAR Project is connecting the automobile, which has autonomous internet node function, to the internet. Nodes in the automobile can be accessed from the Internet and can access the Internet. We have operated several testbed using real automobiles (e.g., taxi or bus), in large metropolitans, such as Yokohama or Nagoya. Some applications, such as a remote sensing using the speed mater and wiper's position, has been operated and evaluated. The testbed system adopts the mobile technologies, such as Mobile IPv6, NEMO (Network in Motion), dynamic tunneling or vertical hand-over. In InternetCAR project, we have installed the differential GPS system with IP to obtain the order of few cm location information. How works the differentiated GPS system is shown in Fig.8.

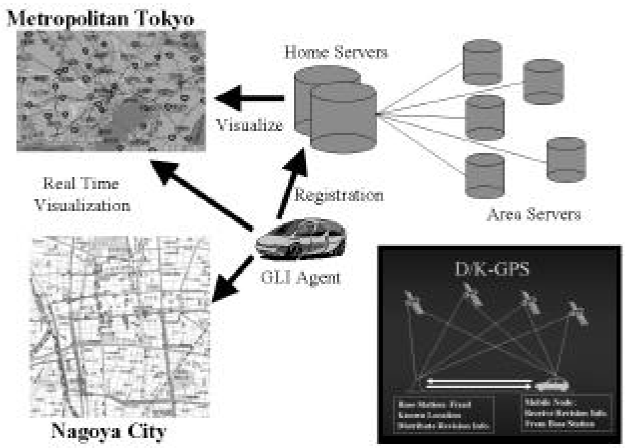

Fig.8 Differentiated GPS with IP

2.3.4 IETF2002 Yokohama

WIDE Project has hosted the IETF2002 held in Yokohama from July 14 to 19. At the meeting, we have installed the full dual stack (IPv6 and IPv4) wireless IP network all the way from the Narita international airport to the conference hotel in Yokohama. IPv6 system with IEEE802.11b has been available, when a participant arrived at the airport lobby, when a participant is on super express train from the airport to conference location, when a participant is at the conference rooms. The express train, called Narita Express (NEX), runs more than 100 km/h, and obtains the IP connectivity using the 3G cellular gear, called FOMA by NTT DoCoMo. With this system, almost a seamless wireless LAN access with IPv6 has been provided to every participant.

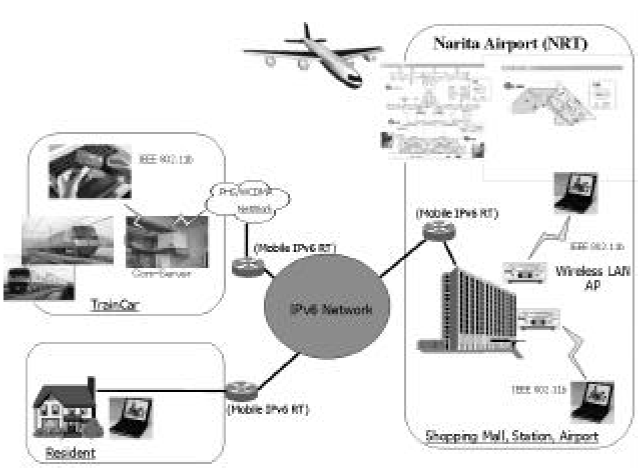

Fig.9 Wireless IPv6 access from NRT to Yokohama

3. WIDE Project Network Configuration

3.1 WIDE Internet Backbone

Figure 10 shows the WIDE Internet backbone topology. Major portion of datalinks are ATM links provided by JGN (Japan Gigabit Network, http://www.shiba.tao.go.jp/), however, we use various types of links, such as WDM, SDH or Satellite link. WIDE internet is internetworking with other networks, mainly at NSPIXP2, 3 and 6. WIDE project operates these three IXes, and the NSPIXP6 is a distinguished IX only for IPv6. At these IXes, we have internetworked with various IPv6 testbeds, such as IPv6 promotion council (www.v6pc.jp) IPv6 testbed. Also, WIDE Internet internetworks with APAN Tokyo-XP, where a lot of R&D network in Asia internetwork. With regards to Abilene of Internet2, we interconnects at Abilene Sunnyvale POP. Figure 11 show the core backbone configuration of IPv6 WIDE Project testbed.

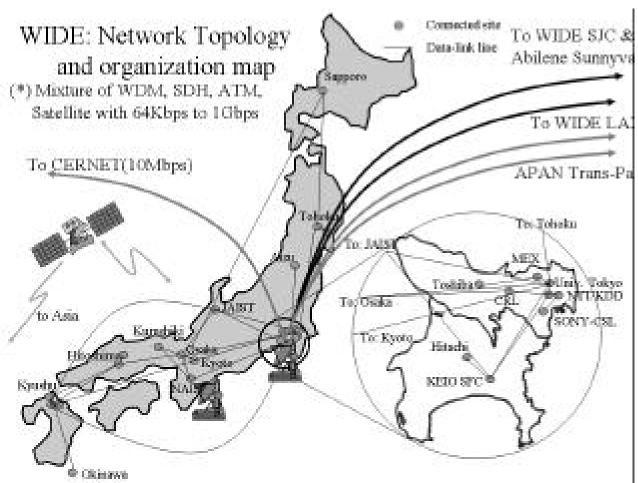

Fig. 10 WIDE Project Internet Backbone

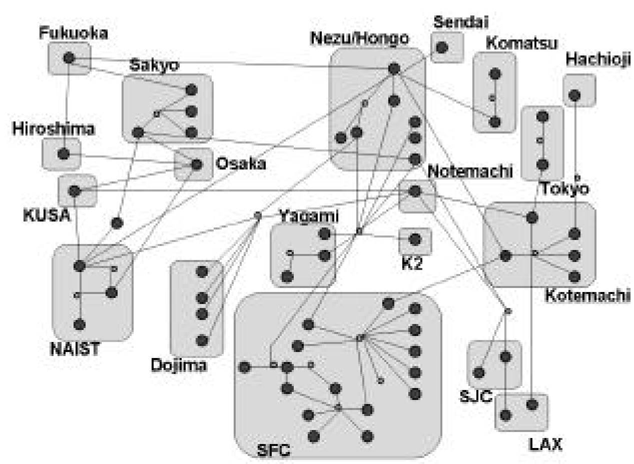

Fig.11 WIDE Project IPv6 Core Configuration

The followings are lost of 22 WIDE-NOC (Network Operation Center);

NTT-Otemachi NOC, KDDI-Otemachi NOC,
Osaka NOC, Kyoto NOC, Kurashiki NOC,
Komatsu NOC, Sakyo NOC, Sendai NOC,
Tokyo NOC, Dojima NOC, Nara NOC,
Hachioji NOC, Horoshima NOC, Fukuoka NOC,
Fujisawa NOC, Yagami NOC, Nezu NOC,
Los Angels NOC, San Jose NOC, Maryland NOC

Figure 12 shows the logical configuration of NSPIXP6 and Fig.13 shows the picture of NSPIXP6. Also, Fig 14 shows the recent traffic volume switched at NSPIXP6.

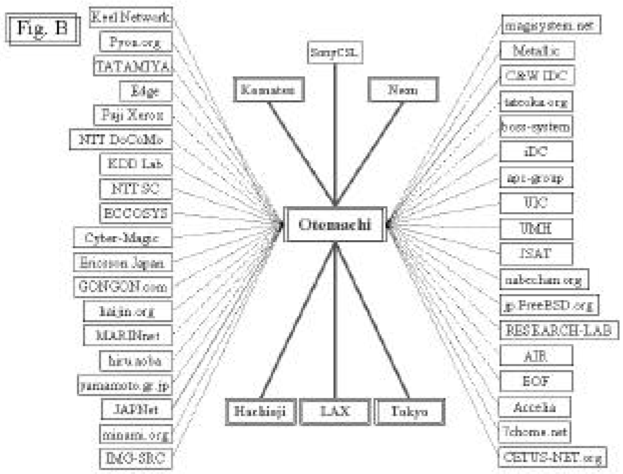

Fig.12 Logical Configuration of NSPIXP6

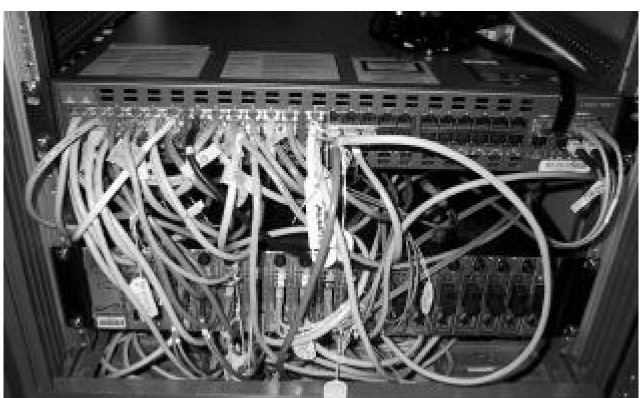

Fig.13 Picture of NSPIXP6 Switch

It has been sometime more then 100Mbps. This means that we have already live and daily IPv6 traffic. Also, we have operated NSPIXP2 and NSPIXP3, which carry large volume of commercial traffic. Since some commercial ISP have already IPv6 commercial operation, NSPIXP2 has already accommodated IPv 6 commercial traffic.

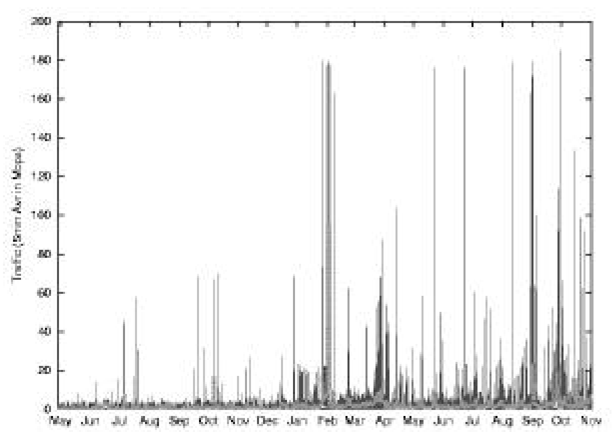

Fig.14 Traffic Volume at NSPIXP6 the NSPIXP6.

4. Conclusion

WIDE project is a research forum among industry and academia regarding the internet technology, covering wide variety of research area. Almost all of R&D activities assume the IPv6 technology as a common platform as layer 3 protocol. The R&D are on platform technologies, middleware technologies, applications and operational technologies. WIDE project operates the R&D testbed, which uses state-of-art technologies, in order to evaluate the new technologies and to realize the further R&D items through the practical operation.

Acknowledgement

WIDE internet operates collaboration with and contribution from various organizations, such as TAO JGN, CRL APII testbed, AI3, PoweredCom/TTnet, NTT Group.

References

[1] http://www.internet2.edu/
[2] http://www.ucaid.edu/abilene/
[3] http://www.6ren.net/
[4] http://www.net-snmp.org/
[5] K.McCloghrie, D.Farinacci, D.Thaler, B. Fenner; "Protocol Independent Multicast MIB for IPv4", IETF RFC2934, October 2000
[6] http://www.csl.sony.co.jp/person/kjc/papers/
[7] Kiyokazu Yoshida, Hirotaro Mori, Shinji Shimojo, Youki Kadobayashi, Toyokazu Akiyama and Mark H. Ellisman, "Design of a remote operation system for trans-Pacific microscopy via international advanced networks", Journal of Electron Microscopy 51(Supplement): S253-S257(2002)