

# Internet Metronome: An Experimental Remote Jazz Jam Session with Uncompressed HDTV Transmission over Lightpaths

Osamu NAKAMURA<sup>†a)</sup>, Kazunori SUGIURA<sup>††</sup>, Seiichi YAMAMOTO<sup>†††</sup>,  
Noriyuki SHIGECHIKA<sup>†</sup>, *Nonmembers*, Akira KATO<sup>††††</sup>, *Member*,  
Katsuyuki HASEBE<sup>†††††</sup>, *Nonmember*, and Jun MURAI<sup>†</sup>, *Member*

**SUMMARY** An experimental remote jazz jam session with uncompressed HDTV over the Internet was conducted on September 21st as a Grand Final event of the Aichi Exposition 2005. Professional jazz musicians located at the venue of Aichi Exposition and at SARA in Amsterdam have made the jazz jam session with new mechanisms called as “Internet Metronome” and “delay-control unit” using an international “lightpath.” This was the first music collaboration using a new methodology and, one of the challenging demonstrations to transport the uncompressed HDTV streams with timing control under the current software and hardware architectures. “Internet Metronome” and “delay-control unit” enabled to make a tempo using and controlling delay, and “lightpath” minimized the network jitter. Using these new mechanisms and technology, the musicians could play with new music collaboration environment over the Internet with long communication delay, and enjoyed remote jazz jam session at both ends.

**key words:** *Internet Metronome, delay control, lightpath, network collaboration, HDTV over Internet*

## 1. Introduction

There are several researches on the music collaboration over the Internet. Large delay, jitter, other communication parameters, and audio video quality are main focuses on those researches. For example, SoundWIRE Project [1]–[3] has been discussed about the method of audio synchronization and reliable professional quality audio streaming. This experimental jazz jam session on the Aichi Exposition 2005 was the first music collaboration with long-haul Internet with lightpath technology for minimized the jitter. While the trials of previous music collaborations like SoundWIRE Project tried to identify the allowance delay over the Internet, they didn’t accept the large delay in a long distance communication. Our strategy was different from those previous music collaborations. We explicitly consider the delay over the Internet, and did not ignore the delay. “Internet Metronome” with “delay-control unit” provided a new music collaboration environment over the Internet with long

delay. Internet Metronome made a tempo using the delay. It was demonstrated that the musicians were able to play with this new music collaboration environment. In this paper, the “Internet Metronome” and “delay-control unit” for “lightpath” technology are described in the following sections. The experimental remote jazz jam session at the Aichi Exposition 2005 between the venue and Netherlands with result of the international lightpath configuration is also described. The “lightpath” technology provides the provisioning services for the Internet applications. We used i-Visto [7] as uncompressed HDTV streaming for the experimentation. The i-Visto has buffer of only two frames. This means that the i-Visto can not work in a large jitter environment. In the normal situation, there is large jitter on the Internet between Japan and Netherlands. The success on the experimental jazz jam session realized and proofed the availability of the “lightpath” technology.

## 2. Internet Metronome

The beat is very important element of music. In remote music collaboration over the Internet, it is very hard to collaborate with beat, because, the communication delay can not be ignored in many cases. In 1999, Ryuichi Sakamoto gave a performance called as “Internet Opera, LIFE” and we supported its communication path over the Internet. While we had been discussed about remote collaboration deeply before the performance, we failed to find a good solution at that time. In the “Internet Opera” performance, we made collaboration between the music that played by Ryuichi Sakamoto at Tokyo and the dancer who are dancing with the music in New York and Frankfurt. The music was very slow without beat and the choreography of the dance was also flow without beat. Another remote music collaboration was that Seiji Ozawa gave a performance at opening ceremony of the Nagano Winter Olympic Games 1998 with Beethoven, symphony No. 9. “Chorus.” In this performance, the music of the orchestra conducted by Seiji Ozawa was transmitted to five cities in the world, and the chorus groups sang with this music and returned to the Nagano Olympic venue. Then mixing operation was performed with adjusting the timing of that music that played at each city. This performance was not collaboration with the orchestra and the chorus group. Just sang with the orchestra like a music minus one CD. To realize the music collaboration with remote musician, “Internet Metronome” was provided.

Manuscript received October 17, 2005.

Manuscript revised November 24, 2005.

<sup>†</sup>The authors are with the Faculty of Environmental Information, Keio University, Fujisawa-shi, 252-8520 Japan.

<sup>††</sup>The author is with Graduate School of Media and Governance, Keio University, Fujisawa-shi, 252-8520 Japan.

<sup>†††</sup>The author is with School of Information Science and Technology, The University of Tokyo, Tokyo, 113-8666 Japan.

<sup>††††</sup>The author is with Information Technology Center, The University of Tokyo, Tokyo, 113-8658 Japan.

<sup>†††††</sup>The author is with Corporate Planning Department, NTT Communications Corporation, Yokosuka-shi, 239-0847 Japan.

a) E-mail: osamu@wide.ad.jp

DOI: 10.1093/ietcom/e89-b.4.1052

## 2.1 Related Works

The latency problem inherent in long-haul network paths is well known but is less well understood in terms of its effect on real-time musical collaboration. There are several researches on the music collaboration over the Internet. Large delay, jitter, other communication parameters, and audio video quality are main focuses on those researches. A method on audio synchronization and reliable professional quality audio streaming has been discussed on SoundWIRE Project. SoundWIRE Project is fundamentally concerned with two areas of research: streaming professional-quality audio to remote destinations using the Internet, and developing practical initiative methods for assessment of network reliability. The namesake for SoundWIRE (Sound Waves on the Internet from Real-time Echoes) is a utility for measuring the reliability of bidirectional network connections. This technique involves the “ping” between a pair of host [4], [5]. It is predicted that the tempo of a performer will be the most affected by delay. Tempo is commonly referred to as the speed of a composition. Composers could define a more specific tempo which can be generated artificially by combining additional delay to the network “ping.” The direction of the tempo is a very useful indicator for composer to synchronize with other end. Our application requires a composer noticeable delay “ping” also including encoding and decoding delay in codecs. Our goal is to hide the level of delay at which effective real-time musical collaboration shifts from possible to impossible. As RTT varies, we want to control the acoustic effect of composer by adjusting a delay period to perform a good tempo.

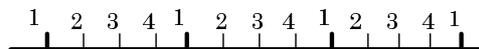
There are artistic experiments, such as “Ping” installed in the outdoor atrium of SFMOMA [6]. This practical installation is using the RTT as a tempo.

## 2.2 Concept of the “Internet Metronome”

The basic concept of the “Internet Metronome” is that making the beat based on the delay.

Figure 1 shows the basic beat with original metronome. Selecting 120 for the beat, the metronome ticks each 500 ms interval. Setting the original metronome at one side of distance locations, ticks of the beat are transmitted as shown Fig. 2. If there is no transmission delay, ticks make a beat at both sides. But with some transmission delay, ticks cannot make a beat at both sides. Suppose that a musician listens only the other side of music and play own music. (Like a normal remote conference setting just considers minus one.) A musician starts the play at side A. Then another musician plays at side B with listen the music transmitted from side A. The musician at side B can collaborate with the music played by side A. But the musician at side A cannot collaborate with the music played by side B. Because the musician at side A listen the music with gap of beat.

The “Internet Metronome” makes a beat based on transmission delay such as Fig. 3. Using the “Internet



1/120: 1 beat is 500ms interval

Fig. 1 Beat with original metronome.

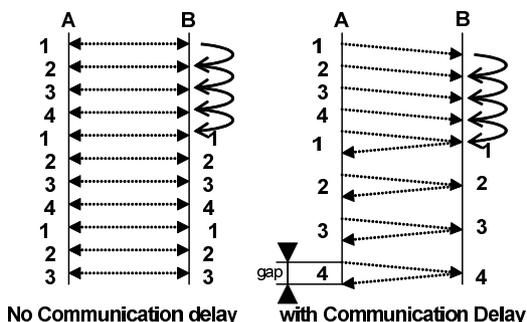


Fig. 2 Gap on the communication delay.

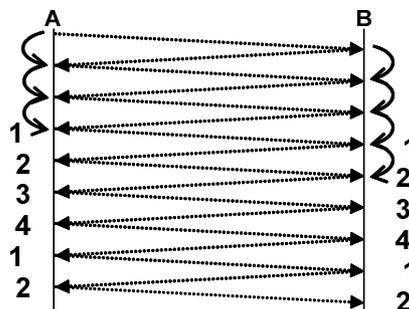


Fig. 3 Beat on the Internet Metronome.

Metronome,” the players can share the beat both side.

## 2.3 Implementation of the Internet Metronome

Figure 4 shows the overview of Internet Metronome. If a player in Amsterdam started performance at the same time of the players in the Aichi venue, the stream arrives at the venue with some delay.

A simple ICMP-like packet is used to measure delay time between Amsterdam and Japan. There are 3 steps of measuring the delay.

1. Send a packet from Amsterdam to Aichi and ring bell.
2. Upon reception of the packet in Aichi, send a response packet back to Amsterdam.
3. Upon reception of the packet in Amsterdam, calculate time since the last packet sent. Calculated time is delay time between Amsterdam and Japan.

Consideration of packet-loss and jitter caused by a packet-loss is required for Internet Metronome. To solve this problem, we implemented a structure called “average counter.” Figure 5 shows the average counter. Previous counter is based on a packet retrieval base system, not effective when packet loss occurs. To solve packet loss problem,

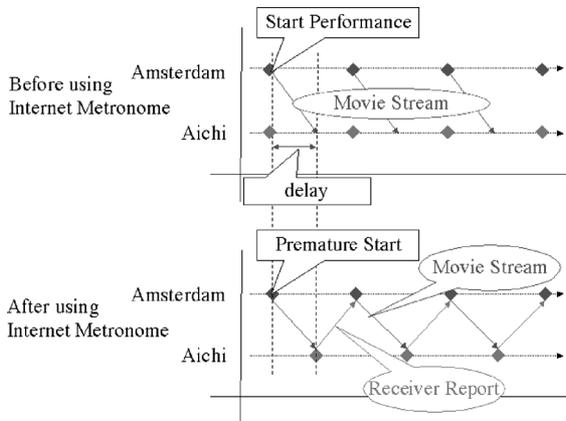


Fig. 4 System overview.

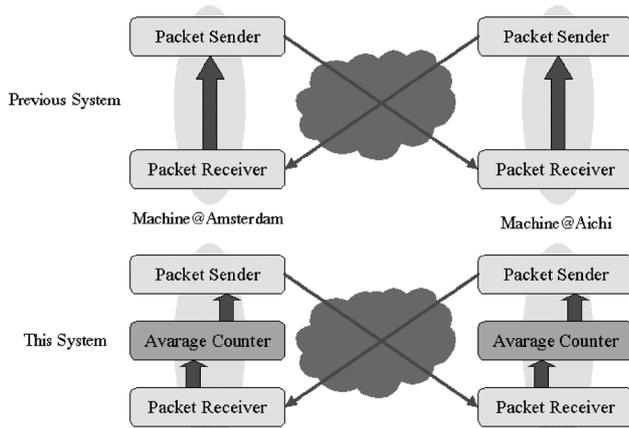


Fig. 5 Average counter.

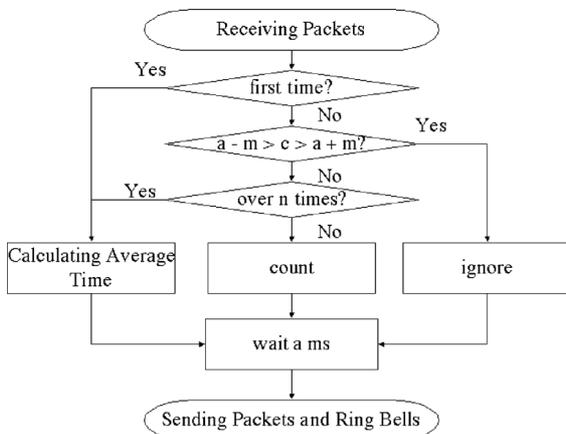


Fig. 6 Calculation method of average delay time.

we implemented first version of “average counter.” This version calculates average time from previous sending time stamp and receiving time stamp. The calculating formula for average  $a$  (ms) is shown below. To calculate average time, current delay time  $c$  and previous average  $p$  are only used.

$$a \text{ (ms)} = \frac{c + p}{2}$$

We focused that network delay jitter can be divided into two types, accidental delay and consecutive delay. Figure 6 shows the calculation method of average time.

“Internet Metronome” is implemented on Windows XP Platform. In the Actual event, Fujitsu LOOX notebook PC was used for the metronome.

### 3. Delay-Control

To evaluate the effectiveness of “Internet Metronome,” 2 types of delay should be considered: 1) Network delay, 2) Equipment codec latency. 1) as described in former section, a delay constructed by the continental location differences, from Netherlands to Japan. 2) is a delay caused by a codec’s of HD digital video streams. For applied evaluation, we constructed a delay-control unit within the network to artificially increase the delay of the network and prove the effectiveness of “Internet Metronome.” Network bandwidth consumed by i-Visto [7] is more than 1.5 Gbps per a HD stream. In this remote jam session event, a bidirectional streams is required for collaboration, consuming total of more than 3 Gbps bandwidth.

There are 3 ways to generate packet delays.

1. Using a bulk optical fiber to establish a physical delay.
2. Using a network memory delay box.
3. Using a network packet forwarding control application.

There are several fixed delay box with fiber connectors capable of generating precise delay using optical bulk fibers. Somehow, it is hard to configure variable set of delays using the combination of boxes, together with the attenuation of laser signals caused by the distance of fiber itself. Solution using network memory boxes will be applicable to generate variety of delay timings, but assuming the bandwidth consumed by HD, network memory box will require a big memory buffer to allocate delay packets. In our events, we constructed a PC based delay-control unit using network packet forwarding control application. We enhanced DUMMYNET [8] which can be implemented in most of UNIX based operating systems. Using DUMMYNET will effectively generate a delay within the packet streams.

DUMMYNET has been used in many applications, and researches for evaluation of bandwidth control. Somehow, most of the applications are using DUMMYNET as a packet flow control and bandwidth limitation. Generating a delay for a Gbps (Gigabit per seconds) class network bandwidth is not considered in current DUMMYNET design.

To establish a delay for a Gbps class traffic using DUMMYNET, optimization of hardware coupled with operating system is required. The operating system should collaborate with hardware for optimal specifications. The main goals for the design are: efficient code / minimum overhead.

We designed the delay-control unit with two types of network interfaces: 10 Gigabit Ethernet and Gigabit Ethernet. 2 delay-control units are required for Gigabit Ethernet to forward an HD stream.

In the traditional DUMMYNET implementation, due

to network buffer availability as well as kernel memory map space limitation, construction of delay is limited. Our initial implementation with DUMMYNET results not more than 300 Mbps of bandwidth capability for 500 ms.

### 3.1 FreeBSD Kernel Optimization

Using DUMMYNET as a delay-control unit, kernel optimization is required. One of the most effective solutions to enhance the bandwidth and delay capability of DUMMYNET is to allocate as many memories as possible for delay buffers. Some of the kernel state can be configured through sysctl interface. The sysctl utility retrieves kernel variables and allows a process with appropriate privilege to set them.

*kern.ipc.nmbclusters* has been increased to enlarge the number of network mbufs in the system allocations. Each network cluster represents approximately 2 kbyte of memory.

We increased *nmbclusters* to 819200, allocating 1.6 Gbyte of memory space to network memory buffers. Configuration of *vm.kmem\_size* is also required for kernel memory space allocation.

Using 10 Gbit Ethernet effectively, modification of device driver was required. Intel PRO/10 GbE LR Server Adapter has configurable parameters for optimal network throughput under conditions. Parameter *RxIntDelay* generates receives interrupts in units of 0.8192 microseconds. Receive interrupt reduction improves processor efficiency under burst UDP traffic streams. Since FreeBSD has been defined HZ as 1000. The default configuration of Hz in FreeBSD is 100, which is larger than single *RxIntdelay*. We tuned this parameter to 123 generating receive interrupt approximately each 100 ms.

We also considered device polling configuration for kernel interface. In the normal, interrupt-based mode, network interface generates an interrupt whenever packet is received. Under heavy traffic load of an HD stream, the system might persistently process the interrupts without being able to completing packet that needs to be sent. When using DUMMYNET in HD delay-control, configuring device based interrupt with interrupt delay was most effective.

Optimization of PCI card timing specification can results in expansion of bandwidth limitations. Configuration of MMRBC (Maximum memory Read Byte Count) field from 512 byte to 4 kbyte in PCI-X configuration space increases transmit burst data length on the bus.

Since we used DUMMYNET in bridged mode, we had a difficulty in configuring DUMMYNET through tagged VLAN based network infrastructure. Whenever the delay configuration change occurs, tagged VLAN packet leak occurs within the bridge causing router to temporary shutting down the network interface. To solve the situation, in the actual event, we configured two delay-control unit for each HD network constructed by Gigabit Ethernet.

**Table 1** Packet throughput of the delay-control unit.

| Interface       | Traffic (MB/s) | Peak (MB/s) |
|-----------------|----------------|-------------|
| Interface 1 in  | 191.290        | 191.611     |
| Interface 1 out | 191.260        | 191.931     |
| Interface 2 in  | 191.123        | 191.886     |
| Interface 2 out | 191.203        | 191.668     |

### 3.2 Implementation

2 types of delay-control box were implemented.

#### 1. 10 Gigabit Ethernet Delay-control Box

**Processor:** AMD Opteron Model 270 (2 GHz) x2

**Motherboard:** TYAN Thunder K8 W

**Memory:** DDR400 4 GB

**Chipset:** AMD 8131 Chipset

**Network Interface:** Intel Pro/10 GbE LR Server Adapter x 2, Broadcom Gigabit Ethernet

#### 2. Gigabit Ethernet Delay-control Box

**Processor:** Intel Xeon 2.0 GHz L2 2 MB 800 MHz FSB

**Motherboard:** Dell Power Edge 1850

**Memory:** DDR400 3 GB

**Chipset:** Intel 7520 Chipset

**Network Interface:** Intel PRO/1000 Network Connection SC x 2, Intel 82541 Dual port Gigabit Ethernet

### 3.3 Packet Throughput

Table 1 show the status of packet throughput in delay-control box sending bi-directional HD stream. Delay-control box shows no packet losses during the event, with constant processor load average of under 0.3. In this status, we probed the bandwidth using sysstat. The period of probing of status was taken for 600 seconds. The size of input and output differ because of the packet dump occurred during the phase of changing the delay period.

## 4. Lightpath Configuration

Recently the Internet research community is starting research on a next generation Internet architecture called as lambda networking. In the lambda networking, one or more lightpaths are provisioned between two locations, ideally the locations where the end systems reside. A lightpath provides a guaranteed bandwidth with a fixed delay and very low packet loss rate. This property can hardly be implemented in a QoS enabled Internet. A number of national and international research networks including CANARIE, IEEAF, Abilene, SURFnet, and WIDE formed a group called GLIF (Global Lambda Integrated Facility) [9] in order to provide lightpaths which may spawn over multiple network infrastructures. Each lightpath may have large bandwidth such as GigabitEthernet (STS-24 equivalent), 2.4 Gbps SONET, or

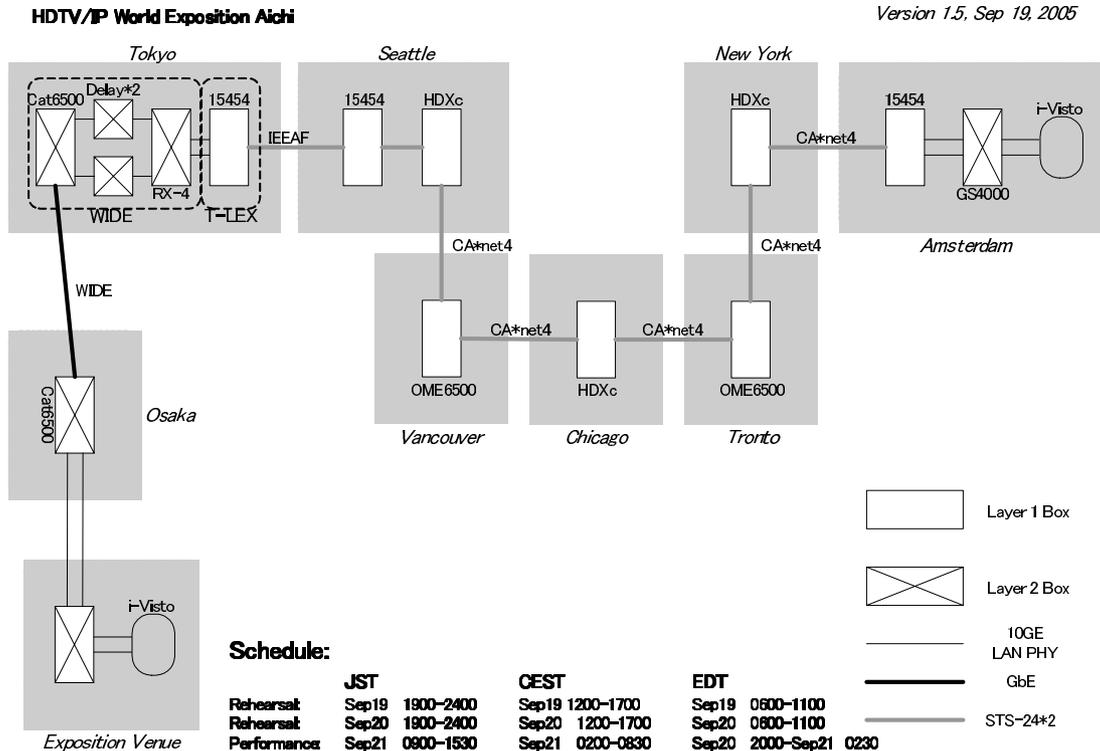


Fig. 7 World Exposition Aichi network diagram.

even 9.6 Gbps SONET or 10 Gigabit Ethernet WANPHY. A couple of STS-24 lightpaths were provided for the Internet Metronome demonstration.

4.1 Configured Lightpaths

A stream of HD-SDI consumes 1.485 Gbps. In order to transfer a bi-directional uncompressed HD streams between the venue and Amsterdam, a couple of Gigabit Ethernet links were provided, because the system we used was NTT's i-Visto which used two Gigabit Ethernet ports for an uncompressed HD transfer. There were several fiber links to make the lightpaths between Japan and Netherlands. We prepared two gigabit ethernets over a STM-16 circuit between the venue and Osaka. Then packets were carried by the WIDE's 10 Gbps backbone primary and JGN2 10 Gbps backbone for backup to Tokyo via two VLANs. In Tokyo, the VLANs on 10GE were mapped into two STS-24 lightpaths. The IEEAF Transpacific circuit provided the lightpaths to Seattle, WA and CA\*net4 provided them their extensions to Amsterdam via Chicago, IL and New York, NY. All of the international links prepared through the international collaboration between among academic and research networks as a part of the GLIF activity.

Figure 7 illustrates the network diagram used for the demonstration.

5. Music Collaboration

We selected the jazz jam session for proving the Internet

metronome from much music collaboration. Jazz jam session has beat and chord. But it is not so tight rather than classic music. As the timing of the translating the chord around the players is not so tight in jazz jam session, is suitable for the remote music collaboration with the Internet metronome.

In this experimental jazz jam session, four jazz musicians were staffed such as Piano/Vocal, Tenor Sax/clarinet, Violin and Drums/Percussion. The violin player played in Amsterdam and other musicians played in the Aichi Exposition venue. This staffing and locations of the players were defined through the previous excises. We had excises at local studio that has two soundproof rooms connected by DVTS/i-Visto streaming technologies twice. We have been any combination of the staffing of the players. The lead piano and rhythm section were located at same place and made main music stream. Then violin over came onto that music. This configuration was easy way to collaboration.

We demonstrated three type of the collaboration by controlling the speed of the Internet metronome on the stage as following.

1. Classic ensemble with one beat gap
2. Kanon with one subsection gap
3. Professional Jazz Collaboration (free jam session)

Classic ensemble with one beat gap was very hard to collaborate by remote players, because remote player (as violin player in this experimental session) should play the melody with one beat before.

Kanon is music that has originally one subsection gap.

This matched the environment well. We controlled the Internet metronome as one subsection. Remote player just played with other side player, caused to kanon to the audience.

And finally professional jazz players played free jam session. They had several experiences to play with a delay environment and the Internet metronome. They were able to play and make good collaborated jazz jam session with one beat gap.

## 6. Conclusion

This experimental jazz jam session on the Aichi Exposition 2005 was the first music collaboration with long haul Internet with lightpath technology. Internet Metronome with delay-control units could provide the new music collaboration environment over the Internet with long delay. Previous music collaboration's trials tried to identify the allowance delay over the Internet. Our strategy was different from those previous music collaborations. We explicitly consider the delay over the Internet, and did not ignore the delay. Internet Metronome could make a tempo using the delay. We could proof that the musicians were able to play with this new music collaboration environment. It was not normal music collaboration environment, but we believe that musician at different locations will be enjoy the collaboration over the Internet with Internet Metronome technology.

In this experimental, we deal with two new high-end technologies over the Internet. One was an uncompressed HD transmission system and the other one was lightpath technology. Using those technologies, we could proof that the availability and operation-ability of the lightpath technology, especially to provide provisioning services such as constant delay and guaranteed bandwidth for high data volume streaming services. Uncompressed HD streaming used 1.5 Gbps bandwidth. And i-Visto was very sensitive the jitter. They did not enough buffers for absorbing the jitter. We could success this experimental by using lightpath technology.

## Acknowledgements

We sincerely thank the people who have greatly contributed to this particular project from IEEEAF, Pacific Northwest Gigapop, CANARIE, SURFnet and SARA, especially Dr. Egon Verharen who coordinated Amsterdam side for this project even David Verharen (his son)'s delivery expected at the same time, then he was born just before the performance. Also, we thank to the Japan Association for the 2005 World Exposition and National Institute of Information and Communications Technology who gave us the opportunity to have this demonstration as grand final event.

## References

- [1] SoundWIRE Project. [http://ccrma.stanford.edu/groups/soundwire/delay\\_p.html](http://ccrma.stanford.edu/groups/soundwire/delay_p.html), As of Oct. 2005.
- [2] C. Chafe, S. Wilson, R. Leistikov, D. Chisholm, and G. Scavone, "A simplified approach to high quality music and sound over IP," Proc. Workshop on Digital Audio Effects (DAFx-00), pp.159–163, Verona, Italy, 2000.
- [3] N. Schuett, The Effects of Latency on Ensemble Performance, CCRMA Department of Music, Stanford Univ., May 2002.
- [4] R.A. Rasch, "Synchronization in performed ensemble music," *Acustica*, vol.43, pp.121–131, 1979.
- [5] T. Kuhn, "Discrimination of modulated beat tempo by professional musicians," *J. Research in Music Education*, vol.22, no.4, pp.270–277, 1974.
- [6] C. Chafe and G. Niemeyer, "Ping, 2001," Installation at SFMOMA, 2001.
- [7] T. Mochida, K. Harada, and M. Maruyama, "The i-Visto XG: Uncompressed HDTV multiple transmission system," IEICE Technical Report, NS2005-44, June 2005.
- [8] L. Rizzo, "Dumynet: A simple approach to the evaluation of network protocols," *ACM SIGCOMM Computer Communication Review*, vol.27, no.1, pp.31–41, Jan. 1997.
- [9] K. Neggers, "Global lambda integrated facility," Spring 2004 Internet2 Meeting, April 2004.



**Osamu Nakamura** graduated Keio University in 1983, Department of Mathematics, Faculty of Science and Technology, MS for Computer Science from Keio University in 1985, received Ph.D in Computer Science, Keio University, 1993. He is currently associate professor at Faculty of Environmental Information, Keio University.



**Kazunori Sugiura** received the B.E. degree in Faculty of Environmental Information, M.E. degree in Media and Governance, and Ph.D. in Media and Governance from The Keio University, Japan, 1994, 1996, and 2003 respectively. He is currently assistant professor at Keio University, Graduate School of Media and Governance. His research interests are network streaming, multicast network, network engineering and operating system architecture



**Seiichi Yamamoto** received the B.E. degree in Mathematical Science and Information Physics and the M.E. degree in Information and Communication Engineering from The University of Tokyo, Tokyo, Japan, in 2001 and 2003, respectively. He is currently a Ph.D. student of Information Science and Technology at the University of Tokyo. His research interests are network traffic measurement, multicast network and practical network operation. He is working on network operation at WIDE Project.



**Noriyuki Shigechika** graduated Keio University in 1994, Faculty of Policy Management, received master's degree for Media and Governance in 1996, and Ph.D. in Media and Governance from Keio University in 2003. He is currently assistant professor at Faculty of Environmental Information, Keio University. His research interests are network operation and internet communication.



**Akira Kato** received his master degree from Graduate School of Science and Engineering, Tokyo Institute of Technology in 1986, and left the school after earning credits in 1990. He received Ph.D. of Media and Governance from Graduate School of Keio University in 2004. He worked as a research associate in Faculty of Environmental Information, Keio University since 1990, a research associate at Computer Centre of the University of Tokyo since 1994, and an associate professor at Information Technology

Center of the University of Tokyo since 2002. He has been working on network research and development of JUNET since 1985 and later in WIDE Project since 1988, when the project bootstrapped.



**Katsuyuki Hasebe** graduated at Keio University and joined NTT in 1993 and is currently Technical Manager at NTT Communications Corporation. He also belongs to National Institute of Information and Communications Technology, Japan. Until 1999, he had researched distributed information network at NTT Labs in Tokyo, Japan and Palo Alto, California, U.S.A. Since 1999, he has researched distributed IX environment and international research network environments. His research in-

terests are international collaboration on the Internet and to make people happy on the Internet.



**Jun Murai** Vice-President, Keio University, Professor, Faculty of Environmental Information, Keio University. Graduated Keio University in 1979, Department of Mathematics, Faculty of Science and Technology, MS for Computer Science from Keio University in 1981, received Ph.D in Computer Science, Keio University, 1987. Executive Director, Keio Research Institute at SFC, 1999-2005. He is appointed as one of the advisory Member of IT Strategy Headquarters established within the Cabinet

since August 2000.