INTERNET-DRAFT
January 27, 2004

K. Ohira
K. Ogata
A. Matsumoto
K. Fujikawa
Y. Okabe
M. Kozuka
Kyoto University
Y. Koyama
Trans New Technology

Hierarchical IPv6 Subnet ID Autoconfiguration for Multi-Address Model Multi-Link Multihoming Site

<draft-ohira-multi6-multilink-auto-prefix-assign-00.txt>

Status of this Memo

Abstract

This document describes a method of automatic hierarchical IPv6 address assignment of a multi-link site with prefix delegation (PD) option enabled DHCPv6.

This protocol is mainly designed to provide an effective list of addresses which a host will have for multi-address model multihome solutions.

draft-ohira-multi6-multilink-auto-prefix-assign-00.txt January 27, 2004

1. Introduction

The site multihoming in IPv6 (multi6) WG explores scalable multi-homing solutions with minimum impacts on routing system and limited need for the number of prefixes advertised in the default-free zone. After long term discussions, multi-address model multihoming is watched with keen interest because it seems very scalable.

Multi-address model multihoming makes good use of the advantage of IPv6: multiple addresses can be assigned to an interface.
A host in a multihomed site has multiple addresses, one is delegated from a transit provider and another from another provider, and selects an address of the peer host and the host itself from candidate addresses. This solution gives multiple access paths without additional advertisement of prefixes in the default-free zone. As a multi-address multihoming solution, SCTP, TCP-MH (on the transport layer) and HIP, LIN6, MAST (on the shim layer which is between the network layer and the transport layer) are proposed.

[ASL] describes that source address based routing is recommended to be used in a multihomed site in order to forward an outgoing packet to the proper transit provider.

This protocol is designed to provide an effective list of addresses which a host will have for multi-address model multihome solutions.

This document describes a way of automatic hierarchical IPv6 address assignment for a multi-link site with prefix delegation (PD) option enabled DHCPv6 [DHC, PDO].

2. Terminology

This document describes how to assign address prefixes in a multi-link multihomed site with "Prefix Option" enabled DHCPv6.

Concerning about DHCPv6, this document uses the terminology defined in RFC 3315 [DHC] and RFC 3633 [PDO].

Concerning about site multihoming, this document uses the terminology defined in RFC 3582 [REQ].

3. Operation of Multihoming

3.1 Role of This Protocol

Multi address model multihoming is, roughly speaking, the model in which a host has different addresses for each path. In this model, the number of addresses which a host has may increase to infinite. Furthermore, according to the increase of the number of addresses, the cost of management of those addresses also increase. Therefore, it becomes too complicated and troublesome to manage manually.

This protocol proposes how to assign IP address prefix to each link in a site automatically.

3.2 Cooperation with Multi-Address aware Upper Layer Protocols

The main feature of multihoming, redundancy and load sharing, will be taken by some multi address aware transport layer protocols (SCTP [SCT], TCP-MH [MHO],...) or shim layer protocols (HIP [HIP], LIN6 [LIN], MAST [MAS],...).

Addresses which a node has may be renumbered. In order to be easy to rendezvous, it is recommended to use DDNS. After a connection has started, regardless of which upper layer protocol is used, the upper layer protocol should be able to handle the renumbering.

In case that an authorized DNS server of a site itself is placed inside of the site, the DNS server is recommended to be placed on the first level link of a site exit router.

This protocol does not prevent the use of MIP6 [MIP] for address portability.

4. Protocol Overview

Basically, the format of messages passed between a delegating router and a requesting router in this protocol follows the definitions in RFC 3633 [PDO] or RFC 3315 [DHC]. This protocol defines only the usage of IPv6 address. The usage of IPv6 address is described in the section 4.1.

In order addressing mechanism to take part in a routing mechanism, both delegating routers and requesting routers are assigned additional requirements besides the ones defined in RFC 3633 [PDO] or RFC 3315 [DHC]. The additional requirements are described in the

section 4.2 and 4.3.

In this section, an example network (figure 1) is considered.

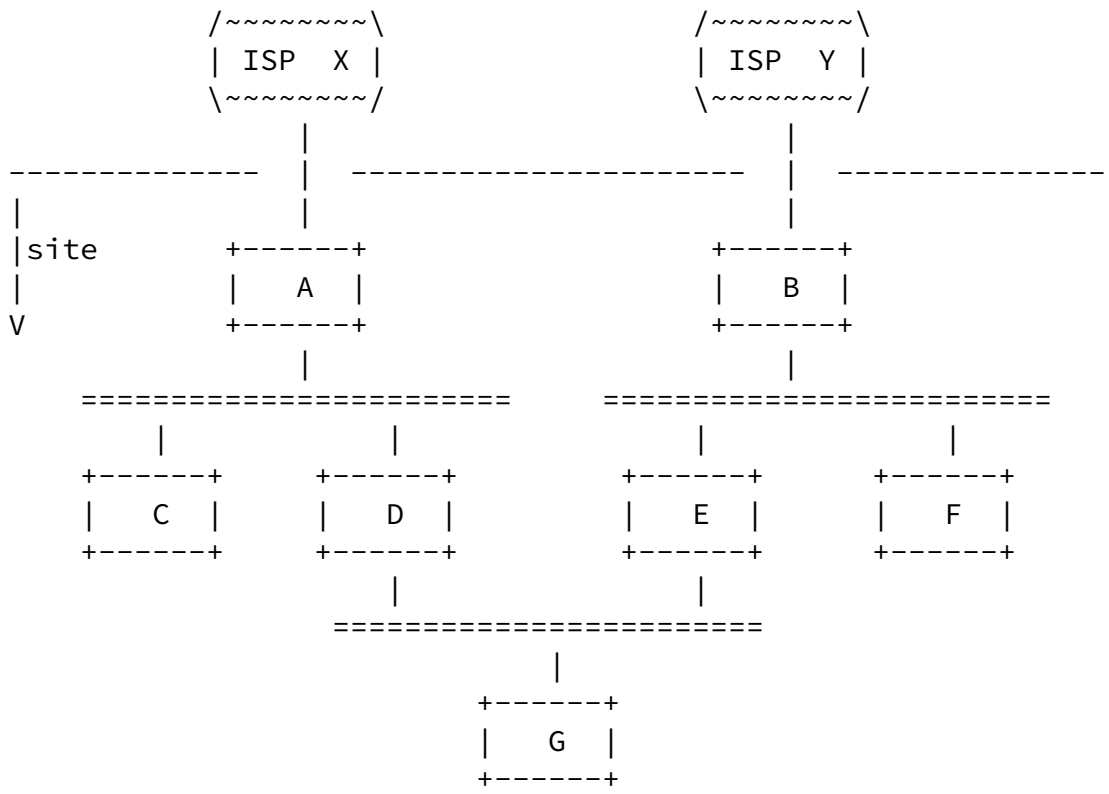


figure 1: An Example Network of a Site

4.1 Usage of IPv6 address

According to [ADD], an IPv6 address is composed as shown in figure 2. This protocol describes how to assign "Subnet Identifier" for each link in a site and how to delegate IP address block for the next hop delegating router recursively.

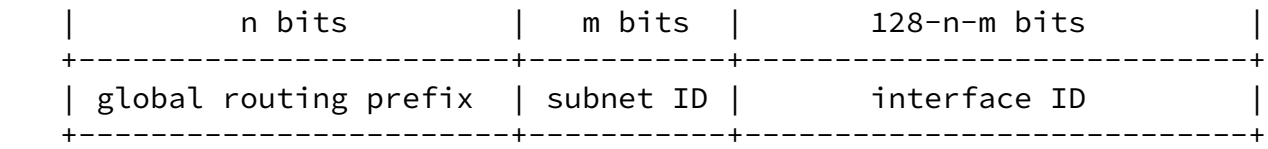


figure 2: IPv6 Global Unicast Address

In this protocol, subnet ID is splitted into four fields as shown in figure 3.

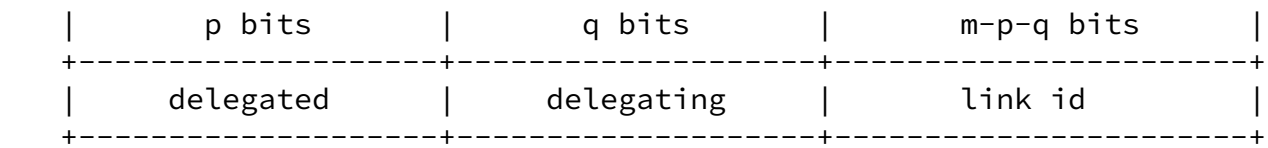


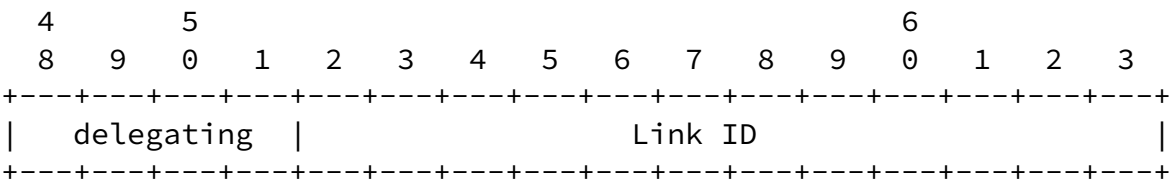
figure 3: Usage of Subnet Identifier

The meaning of each field is as below.

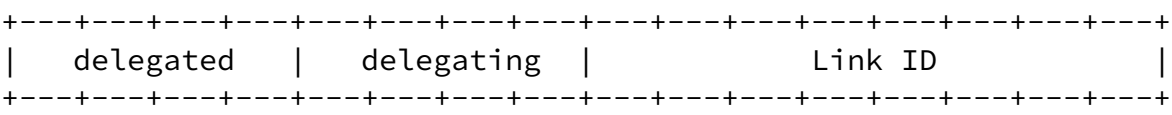
```
   delegated  : The subnet id part of prefix delegated to a
                delegating router.
   delegating :   0   - Room for subnet id of link itself.
                non-0 - A delegating router delegates to each
                        requesting router.
   link id    : Iff the delegating field is 0, this field has a
                meaning of link id.  A delegating router assigns
                an id to each link of the delegating router.
                If the delegating field is non-0, this field has
                no specific meaning.
```

For convenience of explanation, it is assumed that global routing prefix has 48 bits length and subnet id has 16 bits length (i.e.

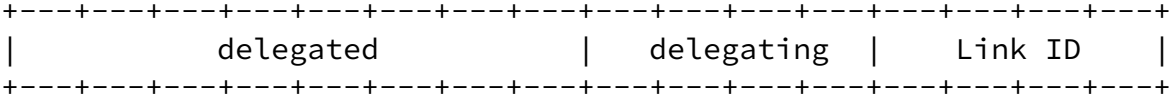
n=48, m=16) and every prefix delegating router sets the length of delegating field and link id field to 4 (i.e. q=4). As a result, subnet id field becomes as shown in figure 4.



(a) at site exit router



(b) at the 1st level delegating router

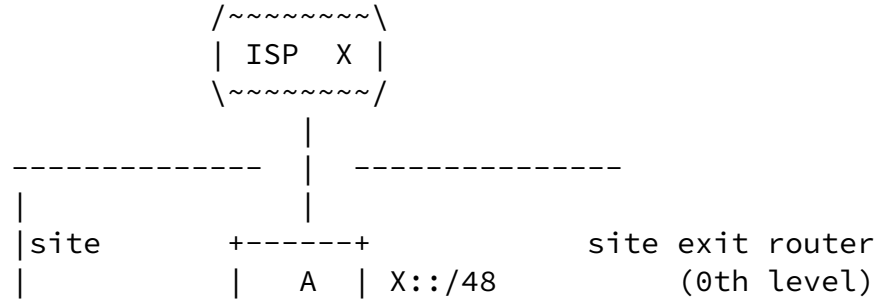


(c) at the 2nd level delegating router

figure 4: Example of Subnet Identifier

As shown in figure 4, in this example, p=0 at site exit router, p=4 at the 1st level delegating router, p=8 at the 2nd level delegating router.

As a result of prefix delegating, concerning about the prefix start with X::/48, the example network is addressed as shown in figure 5.



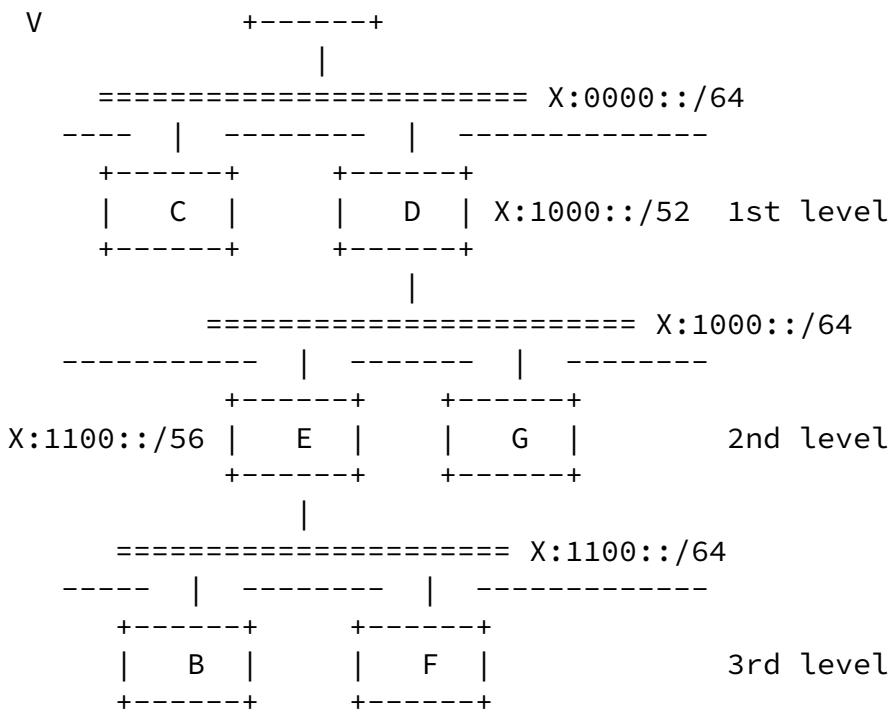
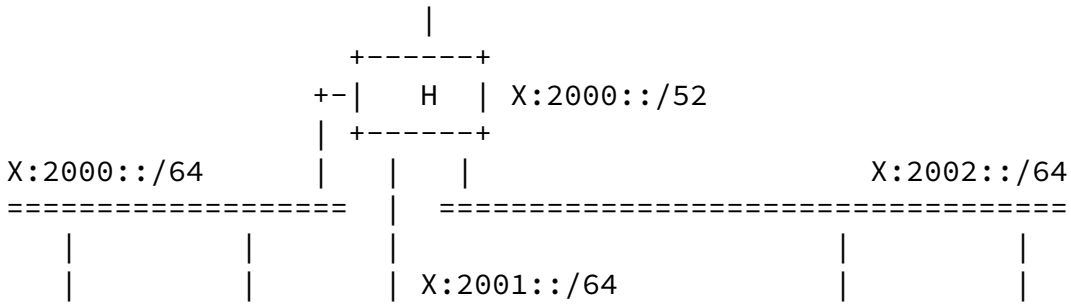


figure 5: Hierarchical expansion of the example network

The Link ID field is needed for the case as shown in figure 6. Without Link ID field, two different links have to have the same Subnet Identifier but this is not allowed.



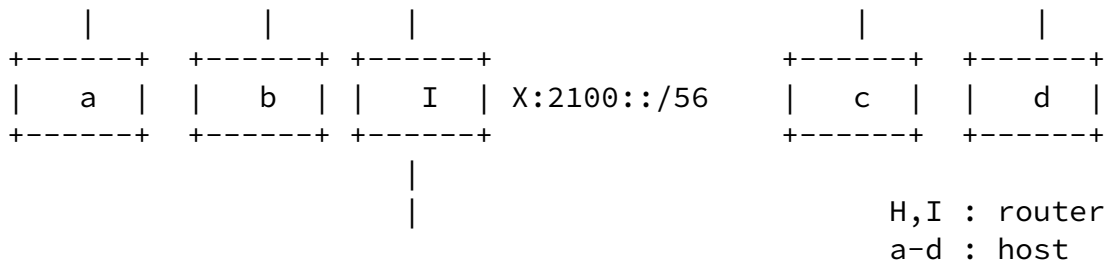


figure 6: Usage of Link Identifier

The prefix start with Y::/48 is assigned and delegated as similar as X::/48.

4.2 Delegating Router Behavior

A delegating router can delegate address blocks to (2^q - 1) requesting routers. When a request from the (2^q)th requesting router, the delegating router will reject the request.

In order to keep the management cost reasonable, it is recommended to set the number of delegated prefixes to each host to the following number:

Preferred limit : 5
Maximum limit : 7.

At the time when a delegating router has already delegate preferred limit numbers of prefix to a requesting host, if the requesting router wants to delegate another prefix which has higher priority than the prefixes which are already delegated, the delegating router must decrease the lifetime of the prefix which has the lowest priority of all delegated prefixes.

A delegating router must not delegate more than the maximum limit numbers of prefixes to a requesting router in any case.

The priority of delegated prefix is as below.

1. Shortest prefix length

2. From different delegating router (from the nearest prior

prefix)
3. Arrival order

For the multihoming feature, this priority is calculated for each global routing prefix (/n). At the time of merging, even if the priority of a prefix is low, the most prior prefix of each global routing prefix should be delegated. Note that this is the reason why p bits length delegated subnet id is separately treated from n bit length global routing prefix.

This ordering mechanism may be done with penalty based algorithm. First of all, pick up an address prefix with the shortest prefix length as the first prior prefix. Therefore, following the similarity with the first prior prefix, add penalty to the other prefixes and pick up a prefix with the least penalty as the second prior prefix. The third and less prior prefixes are picked up one after another.

In order to prevent the loop of address assignment, a delegating router must not delegate smaller address block which is fully included into an address block which the delegating router itself has delegated.

A son of site local address may be delegated and assigned. If [GLU] is used as son of site local address, a prefix and a global ID should be settled by only (candidates of) site exit routers.

A multicast address must not be delegated by this protocol.

Of course, link local address must not be delegated.

4.3 Requesting Router Behavior

For calculating the priority of prefixes, a requesting router must hold the correspondence between a delegated prefix and a delegating router.

For easy routing, it is recommended that a requesting router sets delegating router(s) as default next hop router(s). If a requesting router is delegated address prefixes from multiple delegating routers, then source address based routing should be put in operation at the requesting router.

5. Considerations on RFC 3582

In this section, assessment how much this protocol meets the requirements described in RFC 3582 [REQ].

5.1 Capabilities of IPv4 multihoming

5.1.1 Redundancy

Every external connection is treated completely separate. Therefore, our proposing method is able to continue a connection unless all external connection fail.

5.1.2 Load Sharing

A site is able to distribute both inbound and outbound traffic between multiple transit providers.

5.1.3 Performance

No information between upstream ISPs is required.

If a corresponding node can divide a stream into several destination addresses, we can accomplish to distribute inbound traffic.

5.1.4 Policy

Policies of a site should be expressed as to assign or not to assign an address to a host. If a site does not want a host to use an external connection, the site can decide not to assign an address with the prefix specific to the external connection.

If the site wants to shift traffic of a certain application to a specific ISP, the site should assign multiple addresses to a host and filter traffic whose source or destination has a specific pair of IP address and port number. Therefore the host knows the pair of IP address and port number is not usable and tries another pair. The behavior of a host should be defined in upper layer protocol, therefore it is out of scope of this proposal.

5.1.5 Simplicity

It is recommended that routers are placed as balanced tree. Once topology of routers is fixed, addresses of each routers are

automatically assigned.

5.1.6 Transport-Layer Survivability

Answer of this issue is fully depend on what upper layer protocol will be used. This is out of scope of this proposal.

5.1.7 Impact on DNS

Any change of external connections of a site cause change(s) of prefixes which the site has. Therefore, in the worst case, we may be required to change DNS information at every time.

5.1.8 Packet Filtering

Our proposing method is designed to cooperate with ingress/egress filtering. If the source address of an IP packet is valid then the packet is forwarded to the proper next hop, otherwise the packet will be discarded.

5.2 Additional requirements

5.2.1 Scalability

Only a Provider Aggregatable IP address block from upstream is required. This address is always aggregated at upstream, so even if the number of multihoming site with our proposing method increase, the number of routing information at default free zone. Still more, no AS number is required for a site to be multihomed.

In these points, our proposing method is very scalable.

5.2.2 Impact on Routers

Source address based routing is required for at least one router in a multihoming site. If there are some routers which cannot handle source address based routing, according to the position, routing loop may be occurred.

The authors think that this requirement is relatively little because source address based routing is required only for default route entry.

These modifications do not prevent normal single-homed operations. In a single-homed site, modified routers and unmodified routers can coexist.

5.2.3 Impact on Hosts

Source address based routing is required for all end hosts who want to be fully multihomed. However, a legacy (without source address based routing) host can be obtain some functions of multihome.

If you want to bind several IP addresses to a single TCP connection, TCP Extension for Multihoming may be useful.

5.2.4 Interaction between Hosts and the Routing System

Information about proper next hop for each source address prefixes is needed to be exchanged.

This interaction is quite simple and scalable.

5.2.5 Operations and Management

Administrators of a site are completely capable to monitor the state or to configure parameters of multihoming. At this time, the administrators do not have to do any cooperative work with administrators of upstream.

5.2.6 Cooperation between Transit Providers

Our proposing method does not require any cooperative work between upstream providers at all.

5.2.7 Multiple Solutions

It is recommended that this protocol is used with a multiple address aware upper layer protocol on transport layer or shim layer.

6. Security Considerations

Security considerations in DHCP are described in section 23, "Security Considerations" of RFC 3315 [DHC].

Security considerations in Prefix Options for DHCP are described in

section 15, "Security Considerations" of RFC 3633 [PDO].

The use of this protocol does not introduce any additional known security concerns.

7. IANA Considerations

There are no IANA considerations in this document.

8. References

[REQ] J. Abley, et. al., "Goals for IPv6 Site-Multihoming Architectures", RFC3582, August 2003.

[MAS] D. Crocker, "Multiple Address Service for Transport (MAST) : An Extended Proposal", Internet Draft (work in progress), draft-crocker-mast-proposal-01.txt, September 2003.

[DHC] R. Droms, Ed., et. al., "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC3315, July 2003.

[GLU] R. Hinden, et. al., "Unique Local IPv6 Unicast Address", Internet Draft (work in progress), draft-ietf-ipv6-unique-

local-addr-01.txt, September 2003.

[ADD] R. Hinden, et. al., "Internet Protocol Version 6 (IPv6) Addressing Architecture", RFC3513, April 2003.

[MIP] D. Johnson, et. al., "Mobility Support in IPv6", Internet Draft (work in progress), draft-ietf-mobileip-ipv6-24.txt, June 2003.

[MHO] A. Matsumoto, et. al., "TCP Multi-Home Options", Internet Draft (work in progress), draft-arifumi-tcp-mh-00.txt, October 2003.

[HIP] P. Nikander, et. al., "End-Host Mobility and Multi-Homing with Host Identity Protocol", Internet Draft (work in progress), draft-nikander-hip-mm-01,txt, December 2003.

[ASL] K. Ohira, et. al., "IPv6 Address Assignment and Route Selection for End-to-End Multihoming", Internet Draft (work in progress), draft-ohira-assign-select-e2e-multihome-02.txt, October 2003.

[SCT] R. Stewart, et. al., "Stream Control Transmission Protocol", RFC2960, October 2000.

[LIN] F. Teraoka, et. al., "LIN6: A Solution to Multihoming and Mobility in IPv6", Internet Draft (work in progress, expired), draft-teraoka-multi6-lin6-00.txt, December 2003.

[PDO] O. Troan, et. al., "IPv6 Prefix Options for Dynamic Host Configuration Protocol (DHCP) version 6", RFC3633, December 2003.

9. Authors' Addresses

Kenji Ohira
Graduate School of Informatics
Kyoto University
Yoshida-Hommachi, Sakyo-ku, Kyoto 606-8501 JAPAN
Tel: +81-75-753-7468
Fax: +81-75-753-7472
Email: ohira@net.ist.i.kyoto-u.ac.jp

Katsuya Ogata
Graduate School of Informatics
Kyoto University
Yoshida-Hommachi, Sakyo-ku, Kyoto 606-8501 JAPAN
Tel: +81-75-753-7468
Fax: +81-75-753-7472
Email: ogata@net.ist.i.kyoto-u.ac.jp

Arifumi Matsumoto
Graduate School of Informatics
Kyoto University
Yoshida-Hommachi, Sakyo-ku, Kyoto 606-8501 JAPAN
Tel: +81-75-753-7468
Fax: +81-75-753-7472
Email: arifumi@net.ist.i.kyoto-u.ac.jp

Kenji Fujikawa
Graduate School of Informatics
Kyoto University
Yoshida-Hommachi, Sakyo-ku, Kyoto 606-8501 JAPAN
Tel: +81-75-753-5387
Fax: +81-75-753-4961
Email: fujikawa@real-internet.org

Yasuo Okabe
Academic Center for Computing and Media Studies
Kyoto University
Yoshida-Hommachi, Sakyo-ku, Kyoto 606-8501 JAPAN
Tel: +81-75-753-7458
Fax: +81-75-751-0482
Email: okabe@i.kyoto-u.ac.jp

Masahiro Kozuka
Faculty of Law, Kyoto University
Yoshida-Honmachi, Sakyo-ku, Kyoto 606-8501 JAPAN

Tel: +81 75-753-7468
Fax: +81 75-753-7472
Email: ma-kun@kozuka.jp

Youichi Koyama
Trans New Technology, Inc.
Inohara BLDG. 2F, 72 Tanaka Monzencho, Sakyo,
Kyoto 606-8225 JAPAN
Tel: +81-75-706-9800
Fax: +81-75-706-9801
Email: koyama-y@trans-nt.com