INTERNET DRAFT
draft-ohta-multi6-8plus8-00.txt

M. Ohta
Tokyo Institute of Technology

January 2004

# 8+8 Addressing for IPv6 End to End Multihoming

## Status of this Memo



## Abstract

This memo describes 8+8 address format, which is an IPv6 address format with locator/ID separation useful for end to end multihoming. A 16 byte address of an end is separated into two 8 byte fields: locator, which is used to route packets to a link to which the end is attached, and ID, which is used to globally identify the end.

Locators are assigned from (top level) ISPs to sites (and lower level ISPs) in hierarchical and aggregatable manner that a multihomed site (and ISPs) receive multiple locators from upstream ISPs.

A usual host in a multihomed site (or a singly homed site under a multihomed ISP) is expected to have an ID and multiple locators and transport layer protocols are expected to handle multiple locators of the host and its peer.

## 1. Introduction & Terminologies

This memo describes 8+8 address format, which is an IPv6 address format with locator/ID separation useful for end to end multihoming [ARCH]. A 16 byte address of an end is separated into two 8 byte fields: locator, which is used to route packets to a link to which the end is attached, and ID, which is used to globally identify the end.

Locators are assigned from (top level) ISPs to sites (and lower level ISPs) in hierarchical and aggregatable manner that a multihomed site (and ISPs) receive multiple locators from upstream ISPs.





A usual host in a multihomed site (or a singly homed site under a multihomed ISP) is expected to have an ID and multiple locators and transport layer protocols are expected to handle multiple locators of the host and its peer.

Multicast is not considered in this memo.

Anycast is treated identically to unicast.

The followings are terminologies used in this memo:

8+8 Address
: An address composed of an 8 byte locator and an 8 byte ID.

Address
: 16 byte information to identify and locate an end. An end may have multiple addresses.

Destination ID
: ID of a destination address

Destination Locator
: Locator of a destination address

End
: The primary unit, of the end to end principle [ARCHINTERNET], also called "end point" in [SALTZER].

ID
: 8 byte information to identify an end. An end may have multiple IDs.

Locator
: 8 byte information to locate an end. An end may have multiple locators.

Source ID

ID of a source address

Source Locator
Locator of a source address

## 2. Address Format

An 8+8 address format is identical to that of the existing IPv6 unicast address [ADDRARCH, UNIADDR] derived from EUI-64, except that it uses unused part of the IPv6 unicast address space.

### 2.1 Locator Format

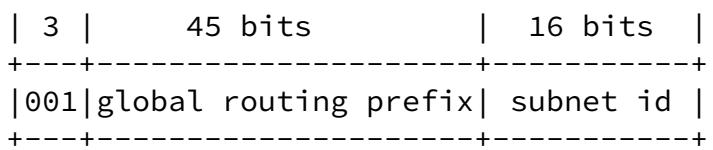
Locator format is identical to the upper 8 bytes of existing IPv6 unicast address [ADDRARCH, UNIADDR].

That is, as defined in [UNIADDR], a locator of an end have the





following format:

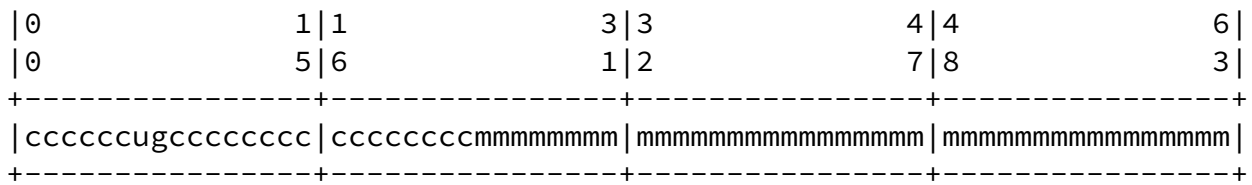



### 2.2 ID Format

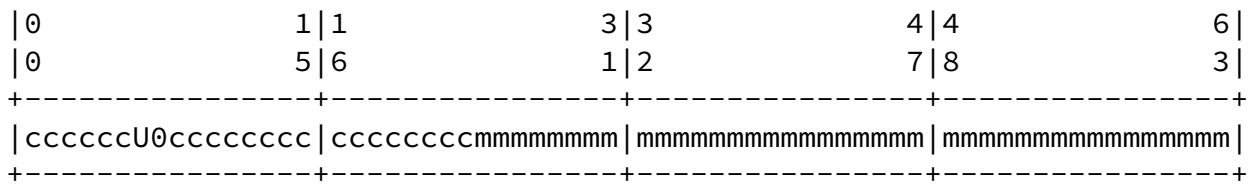
ID format is identical to the lower 8 bytes of existing EUI-64 based IPv6 unicast address [ADDRARCH, UNIADDR].

However, when an 8+8 address is viewed as an EUI-64 based address, individual/group bit of EUI-64 is turned on, which is not a case with a real EUI-64 based address.

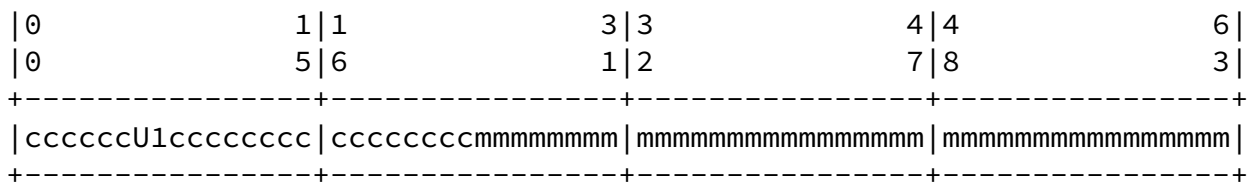
As a globally unique IEEE EUI-64 identifier has the following form:



and existing IPv6 interface identifier has the following form ("U" bit is a inversion of "u" bit):




an ID of an 8+8 address has the following form:




"U" bit is reserved and should be, for the time being, filled with 0 though either 0 or 1 should be accepted. See "3.1 Mobility Consideration" for possible use of the bit.

## 2.3 ID scope and semantics

An ID is expected to be globally unique.

An ID is to identify an end.

As such, all the interfaces of an end share an ID(s) of the end, as if an EUI-64 address corresponding to the ID is shared by all the interfaces, which is consistent with [ADDRARCH, UNIADDR].





## 2.4 ID Space Structure

An ID of an end is structured to have hierarchical mapping by DNS from ID to DNS name of the end, just as IPv4 addresses under in-addr.arpa. If such mapping does not configured or wrongly configured, DNS-based confirmation of association between addresses and hostnames will not be available, which is the case of IPv4 addresses.

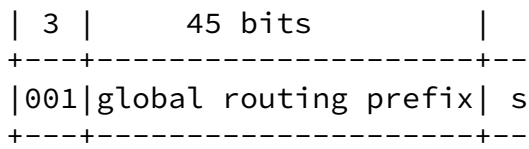
### 2.4.1 Locator Derived ID

To ease initial assignment of IDs to ends, IDs may be derived from

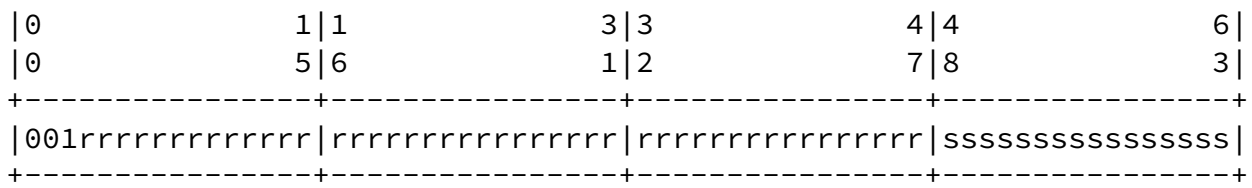
locators allocated to sites.

An ID is a locator derived ID, if its first bit is 1.

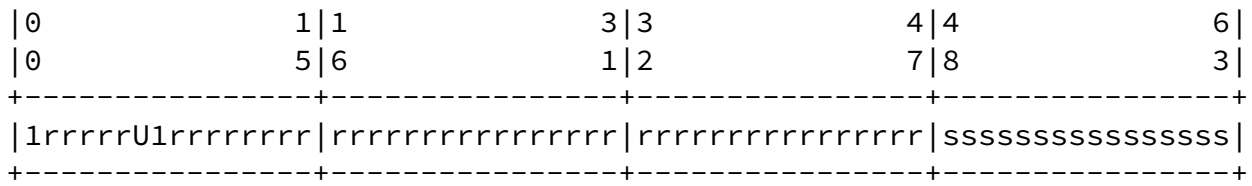
With a locator with the following format:




or




a locator derived ID has the following format:




Reverse mapping of a locator derived ID of an end is performed under ipv6.arpa. domain, using locator format corresponding to the ID [IPV6DNS], though only 64 bits of the locator format are used to reach an PTR record corresponding to the locator format.

That is, a site is assured to have 65,536 IDs assigned, though the locator nature of the ID makes IDs not stable against site renumbering.

Note that ID assignment within a site can be arbitrary and will not be consistent of intra site link structure. That is, an end with a locator derived ID including a certain subnet id may be located anywhere in the site, not necessarily in the subnet corresponding to the subnet id.

If one want to hide its privacy in ID, one should use locator derived

ID for one's ends.

### 2.4.2 Location Independent ID

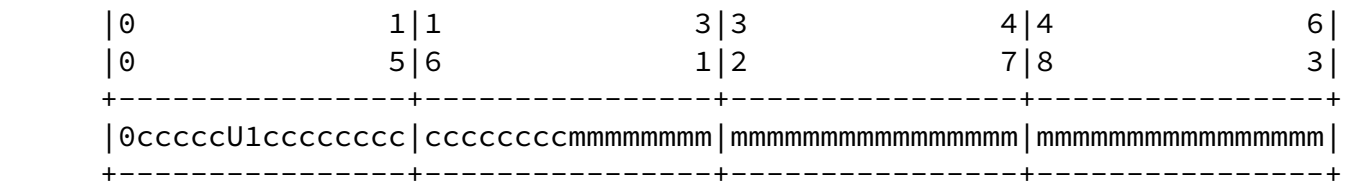
Location independent ID is introduced to overcome the difficulties of locator derived IDs, that

- locator derived IDs are location dependent
- a site can not be assigned IDs a lot more than 65,536.

Location independent ID is, just like DNS names, assigned with logical structure independent of network topology, including site structure. That is, location independent IDs are assigned to organizations and individuals.

As location independent ID space does not incur inefficiency due to route aggregation, entities equivalent to site is expected to be able to receive a lot more than 65536 IDs.

An ID is a location independent ID, if its first bit is 0 and has the following structure.




Further details of operations and management of location independent IDs are TBD, though they should be almost identical to those for domain names, except that there is no trademark issues expected for 16 hexadecimal digits, which is not expected to be visible with most user interface.

## 3. Internetworking Layer Protocol

At the Internetworking layer, an 8+8 address has no special meaning and packets containing 8+8 addresses in their IP headers is treated as a normal IPv6 packets for all the processing, including, but not limited to, forwarding by routers and ICMP [ICMPv6].

As is discussed in [ARCH], the Internetworking layer, having no notion of connection nor time out, has little to do with multihoming, except that IP addresses are part of transport layer identifiers.

However, following subsections give some considerations for performance enhancements.

## 3.1 Mobility Considerations

While [MIPV6] should work with 8+8 addressing as is, 8+8 addressing offers chance to have better mobility protocol. One is that DAD, which is the primary cause of delay of [MIPV6] is not seriously





unnecessary for globally unique ID. The other is that tunneling and associate MTU change of [MIPV6] can be eliminated if locators can be rewritten.

[LIN6MOBI] is expected to exploit such possibilities.

To mark a mobile end as mobile without first communicating with the end, "U" bit may be used.

Further discussion on mobility is out of scope of site multihoming.

## 3.2 Destination Locator Selection

A problem with end to end multihoming is that an end have multiple locators and proper locator must be chosen to reach the end.

While [ADDRSELECT] tries to define some guideline for destination address (locator) selection, it is not very useful to select one from multiple global unicast addresses.

If an end has no idea on what is the best destination address, selection should be at random.

However, if an end have routing information, it can use it to determine which locator is unreachable and which locator have least metric (such as type 2 metric of OSPF) [ARCH].

To obtain metric information of external routes of IGP, BGP AS path length may be used as the metric. Or BGP administrators may adjust IGP metric more finely to control load of outgoing traffic.

As end to end multihoming is expected to remove the only reason to bloat global routing table size, save laziness of assignment authorities, the global routing table of IPv6 should be kept small and all hosts should have default free full routing table for efficient selection of destination addresses. Note that an end should

receive, but may not necessarily generate, routing information.

While the Internetworking layer gives information on preference of locators, the Internetworking layer does not perform retransmission. Thus, if some locator fails, it is detected by a transport or an application layer and the layer takes care of retransmission with next best destination locator candidates. Implementations are encouraged to let upper layers look at the routing table for efficiency, which is not a layering violation.

## 3.3 Source Locator Selection

While [ADDRSELECT] tries to define some guideline for source address (locator) selection, it is not very useful to select one from multiple global unicast addresses.

Given highly asymmetric nature of Internet routing, a host basically has no knowledge on what is the best destination locator used by its





peer for reply. In this sense, source locator selection is not necessary and source locator can be chosen randomly.

However, source locator selection becomes important for ingress filtering on source addresses, in which case, proper source locator corresponding to a destination locator must be chosen. Otherwise, most packets will be filtered and a lot of time is wasted for a transport or an application protocol retransmit and find the proper source locator.

As is discussed in [ARCH], such corresponding can be best obtained from proper routing protocol. For example, with OSPFv6, locator part of forwarding address field of AS-external-LSA can designate the locator for the LSA.

When routing protocols supply source locator, source locator selection is performed purely by the Internetworking layer without involving inefficient retransmission by a transport or an application layer.

# 4. Transport and Application Layer Protocols

For backward compatibility, if either a source or a destination

address is not an 8+8 address, transport and application layer protocols behave as if it is a legacy end.

Otherwise, that is, if both source and destination address are 8+8 addresses, transport layer protocols ignore source and destination locators, except for security and performance enhancement. That is, transport layer protocols does not use the locators for checksum and identify their connections using IDs only.

As is discussed in the previous section, a transport or an application protocol is responsible for selection of destination locator and associated retransmission.

However, as is discussed in [ARCH], such retransmission dependent on the nature of applications that no generic mechanism can be discussed in this memo. Each application has different notion of connection or loss of it that it is not meaningful to give a generic time out value.

Nevertheless, as is discussed in [ARCH], TCP has defact default timeout values. So, it is possible to have generic TCP with default behavior for locator selection and retransmission.

It should be noted that most applications over the Internet works over TCP and that such applications can run with end to end multihoming without modifying application programs.

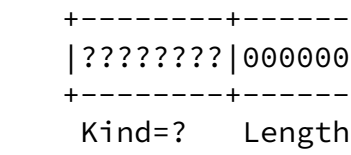
In addition, as DNS is a basic infrastructure and has its own timeout values, it is necessary to investigate possible modifications to DNS with end to end multihoming.





## 4.1 Modification to TCP

A new TCP option, Multi Address option, is introduced.




Kind is to be assigned by IANA.

The option has one argument, the number of locators, value of which must be between 1 and 9. If the option appears in a TCP header, data field just after the TCP header contains, with network byte order, locators of the source host, the number of which is specified by the argument.

It is expected that 9 locators are enough for most ends, as a site of the end can be multihomed to three lower level ISPs each multihomed to 3 top level ISPs. However, if an end has more than 9 locators, which is a case with routers with more than 9 interfaces, TCP or upper layer modules should be responsible to select 9 or less locators to be used for the TCP connection.

All TCP modules of ends supporting 8+8 addressing must recognize the Multi Address option.

TCP modules memorize current most source locators of its peer and reject TCP packets with unknown source locators.

TCP modules should have interface to application modules to let the application modules check whether the set of locators supplied by the Multi Address option is valid or not.

Multi Address options must be used by packets for the initial three way handshaking and may appear in any other TCP packet.

Multi address option is also useful for performance reasons.

Note that, as route of the Internet is highly asymmetric, a source locator of a packet, which is chosen for ingress filtering, may not work for reply that it is essential to provide all the candidate locators for the reply.

In addition, when SYN times out, TCP should retry with new destination locators. When SYN ACK times out, TCP should retry with new destination locators contained in a set of locators provided by Multi Address option of SYN packet. Thus, with N locators, it is expected that O(N), not O(N*N), attempt is enough to find a working pair of source and destination locators. If TCP modules detect SYN flood attack, they do not have to allocate state for SYN packets to memorize the set of locators in the Multi Address option of the packets. Instead, it should randomly choose one from the Multi Address option of the SYN packet.

### 4.2 Modification to DNS

As is discussed in [ARCH], DNS tries all the addresses of name servers that it is already an application with end to end multihoming.

A problem is that unlike TCP, DNS servers do not expect acknowledgment and do not retransmit. So, if a client can not get a response, it should retry with alternative destination locators of a server with corresponding source locators. But, even if the server somehow knows all the locators of the client, the server send just one reply and does not try all of them. Thus, with N locators and in the worst case, O(N*N) attempt is necessary to find a working pair of locators.

But, the problem is not serious, because usual clients of DNS today, gives up with small number of attempts and because there are multiple servers are provided for each zone.

### 4.3 8+8 Adaptation Layers for Applications over TCP and DNS

The 8+8 adaptation layers make end to end multihoming invisible to applications over TCP and DNS, that is, applications using TCP transport only without hard coded IP addresses.

Some multihoming proposals try to introduce an adaptation layer in the Internetworking layer to hide locator changes. However, this attempt does not make sense. As demonstrated by NAT, rewriting of addresses is, in general, not transparent to transport and application protocols. For example, transport layer checksum computation involves IP addresses and is different for each transport protocol that address rewriting can not be confined in the Internetworking layer. Insertion or deletion of IP headers affects MTU, which is also visible to the transport layer. Applications like FTP transmit raw addresses in application data streams.

However, it is possible to confine modifications in 4.1 in TCP and let applications get a fixed locator (LIN6 locator) [LIN6ARCH]. In addition, it is possible to modify DNS library to let such applications get addresses with the LIN6 locator.

Then, it is possible to make end to end multihoming invisible from applications over TCP and DNS, including applications transmit raw addresses in application data streams.

## 5. Assessment Against RFC3582

Redundancy
   Fully supported. That is, ends should be able to keep communicating against all failure modes of locators as long as a pair of working source and destination locators exists, though details are transport and application layer dependent.

Load Sharing, Performance





   Fully supported. That is, if site administrators elaborate on BGP configuration, they have as much control as possible with BGP. Site administrators, instead, can just rely on ASpathlen, in which case, there will be no configuration effort.

Policy
   Fully supported. If a site is multihomed to ISP A and B and an end has locators of ISP A only, all traffic to the end will be through ISP A. Traffic from the end can be controlled by policy of accepting route.

Simplicity
   As for deployment, the proposal fully interoperate with legacy systems. Applications over TCP and DNS does not need any coding changes. As for operation, as long as IDs may change with rehoming, it is as simple as the current IPv4 multihoming practices.

Transport-Layer Survivability
   Fully supported, though details are transport layer specific, of course.

Impact on DNS
   The proposal is fully compatible with the observed dynamics of the current DNS.

Packet Filtering
   The proposal is compatible with packet filtering on source addresses.

Scalability
   Fully scales. That is, the number of multihomed site does not affect the number of routing table entries.

Impact on Routers
: Nothing, except that routers may have 8+8 addresses and behave accordingly.

Impact on Hosts
: Communications with legacy hosts is kept unchanged, though most, if not all, the benefit of multihoming will be lost.

  API of applications using TCP and DNS remain unchanged and the applications can still enjoy full multihoming support.

  Applications over UDP, where all the packets can and must be controlled by "user", need changes, if it want transport-layer survivability (even the meaning of "transport-layer survivability" is defined by the "user"). Otherwise, the current transport-layer protocol may be used as is.

Interaction with Hosts and Routing System.
: Ends expect to passively receive routing information from the routing system, which is simple, scalable and securable.





Operations and Management
: Site's multihoming system, with the proposal, is site's routing system that it is possible for staff responsible for the operation of a site to monitor and configure it.

Cooperation between Transit Providers
: No cooperations are required.

Multiple Solutions
: Proposal contains a single solution only.

Security Considerations
: Multihomed sites and ends are not more valunerable to traditional IPv4-multihomed sites and ends.

  Routing practice change to carry source locator does not affect security of non multihomed site.

## 6. References

[ADDRARCH] R. Hinden, S. Deering, "Internet Protocol Version 6 (IPv6)

Addressing Architecture", RFC3513, April 2003.

[ADDRSELECT] R. Draves, "Default Address Selection for Internet Protocol version 6 (IPv6)", RFC3484, February 2003.

[ARCH] M. Ohta, "The Architecture of End to End Multihoming", Work in Progress as <draft-ohta-e2e-multihoming-05.txt>, June 2003.

[ARCHINTERNET] B. Carpenter, Ed., "Architectural Principles of the Internet", RFC1958, June 1996.

[ICMPv6] A. Conta, S. Deering, "ICMP for the Internet Protocol Version 6 (IPv6)", RFC 2463, December 1998.

[IPV6] S. Deering, R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.

[IPV6DNS] S. Thomson, C. Huitema, V. Ksinant, M. Souissi, "DNS Extensions to Support IP Version 6", RFC3596, October 2003.

[LIN6ARCH]

[LIN6MOBI]

[UNIADDR] R. Hinden, S. Deering, E. Nordmark, "IPv6 Global Unicast Address Format", RFC3587, August 2003.

## 7. Security Considerations

The Internetworking is identical to the legacy IPv6 that there is no new security concern.

The transport layer is modified that there is possibility of wrongly





recognized source locator and transmit a lot of packets to wrong places.

While details of source locator authentication and packet retransmissions are transport and application dependent, there can be some guideline to prevent the problem.

When a packet arrives with a source locator, the validity of the

locator can be confirmed with reasonable security with three way handshaking or cookies.

When a packet arrives with multiple locators, the validity of one of a locator can still be confirmed with reasonable security with three way handshaking or cookies. As long as all the locators are contained in a single packet, it is reasonable to treat the set of the locators have reasonable security. As the destination locator of reply packet is chosen by the replying host and retransmission is expected to be infrequent, DoS attack using the multiple locators for reply is only as efficient as DoS attack with single locators.

TCP modification in 4.3 is expected to work this way.

However, to avoid interference between connections, each connection module should maintain the set of locators separately even if several connections exists to a single ID.

## 8. Author's Address

Masataka Ohta
Graduate School of Information Science and Engineering
Tokyo Institute of Technology
2-12-1, O-okayama, Meguro-ku, Tokyo 152-8552, JAPAN

Phone: +81-3-5734-3299
Fax: +81-3-5734-3299
EMail: mohta@necom830.hpcl.titech.ac.jp