IPv6 and Neighbor Discovery over ATM
<draft-ietf-ion-ipv6nd-00.txt>

Status of this Memo

This document was submitted to the IETF Internetworking over NBMA (ION) WG. Publication of this document does not imply acceptance by the ION WG of any ideas expressed within. Comments should be submitted to the ion@nexen.com mailing list.

Distribution of this memo is unlimited.

This memo is an internet draft. Internet Drafts are working documents of the Internet Engineering Task Force (IETF), its Areas, and its Working Groups. Note that other groups may also distribute working documents as Internet Drafts.

Internet Drafts are draft documents valid for a maximum of six months. Internet Drafts may be updated, replaced, or obsoleted by other documents at any time. It is not appropriate to use Internet Drafts as reference material or to cite them other than as a "working draft" or "work in progress".

To learn the current status of any Internet-Draft, please check the "lid-abstracts.txt" listing contained in the Internet-Drafts shadow directories on ds.internic.net (US East Coast), nic.nordu.net (Europe), ftp.isi.edu (US West Coast), or munnari.oz.au (Pacific Rim).

Abstract

This document attempts to describe and summarize some current proposals for running IPv6 over ATM, and identifies open issues that require resolution by one or more IETF working groups. The frame formats for unicast and multicast transmission of IPv6 packets in a UNI3.1 based ATM environment are specified. Some issues regarding the construction of IPv6 link-local addresses are identified, and a proposal made. A format for source and target link-layer address options in Neighbor Discovery messages is suggested, and the interactions between IPv6 Neighbor Discovery and existing IP over ATM models are outlined.
This revision looks at the three models that were presented at the Los Angeles IETF in March 1996, in a joint session of the IPATM, IPNG, and ROLC working groups. No firm decisions were made at this joint meeting. Readers are encouraged to locate and review the Internet Drafts describing each model in greater detail.

Further discussion of the issues raised in this document is requested, as not all questions are currently answered satisfactorily.

Revision History

[This part to be removed when I-D is completed.]

June 1996, ION version 00.
Re-release of draft-ietf-ipatm-ipv6nd-02.txt as draft-ietf-ion-ipv6nd-00.txt, after the merger of IPATM and ROLC into ION. Retains a certain ATM-centricity which will be refined in later releases.

April 1996, IPATM version 02.
Added further description of the orthogonal issues of interface ID selection and the specification and identification of one’s Neighbor in an IPv6 sense. Pointers to all three models presented at the March 1996 IETF. No firm consensus from this meeting - most open issues are still open.

February 1996, IPATM version 01.
Re-write to deprecate some contentious suggestions issues, and provides pointer to work being presented at the March 1996 IETF meeting.

August 1995, IPATM version 00.
Poses the original question of how the IPng assumption of ‘cheap’ link level multicasting makes the IPng Neighbor Discovery protocol hard to support when the underlying technology is an ATM network. Suggests a straw-man model based on MARS to identify its limitations. Solicits ideas from the community.

1. Introduction.

This document deals with a number of issues associated with running IPv6 [1] over UNI3.1 [2] based ATM services. These may be characterised as:
- Packet framing and multicasting.
- Link Local address specification.
- Neighbor Discovery source/target link-layer address option.
- Interactions between ND and underlying IP/ATM architectures.

Packet framing is dealt with in section 2, applying the newly assigned IPv6 Ethertype [3] to the encapsulation models developed for IPv4 unicast [4] and multicast [5]. Section 2 will also note the specific behaviour required of an IPv6-ATM interface when using the MARS protocol defined in [5] to support IPv6 multicast over UNI3.1 ATM networks.

Section 3 outlines the requirements for the structure of IPv6 Link Local addresses [6], and provides pointers to some current ideas on the creation of link-local tokens.

The format of the source and target link-layer address options in Neighbor Discovery [7] messages is described in section 4.

Section 5 summarizes the current discussion of how the IPv6 Neighbor Discovery service and/or protocol may be applied to ATM environments. Primarily it points to three models that were presented to the March 1996 IETF [10], [11], and [13].

It is expected that the models in this document may be applied to a wider community of NBMA networks, with suitable refinement of the text.

[Editors note: Further discussion of the issues raised in this document is requested on the ip-atm@nexen.com mailing list.]

2. Multicast support, and packet framing.

2.1 Using MARS for multicast support.

Multicasting is an integral part of IPv6. However, most NBMA networks (and UNI3.0/3.1 based ATM networks in particular) do not provide sufficient native multicast support to allow a trivial mapping. The IP over ATM working group is nearing completion of a convergence function, known as the ‘MARS model’ [5], which builds the required multicast support using a point to multipoint VC service. A MARS based IP/ATM device driver emulates link level multicast support for the IP layer.

IPv4 is used as the main example in [5]. What follows are the main changes required to use [5] for IPv6.
The encapsulation of MARS control messages (between MARS and MARS Clients) remains the same:

```
[0xAA-AA-03][0x00-00-00-00][0x08-06][MARS control message]
(LLC)       (OUI)     (PID)
```

The mar$afn field in MARS messages remains 0x0F.

The mar$pro field in MARS messages SHALL be 0x86DD.

The mar$spln and mar$tpln fields (where relevant) are either 0 (for null or non-existent information) or 16 (for the full IPv6 protocol address).

When a host starts up it SHALL issue a single group MARS_JOIN for the following groups:

- Its derived Solicited-node address(es) with link-local scope.
- The All-nodes address with link-local scope.
- Other multicast groups with at least link-local scope.

For example the IPv6 node with address 4037::01:800:200E:8C6C would issue the following MARS_JOIN to register as a member of its Solicited-Node multicast group:

```
MARS_JOIN(FF02::1:200E:8C6C, FF02::1:200E:8C6C)
```

Joining or leaving a multicast group with node-local scope (scope 1) MUST NOT cause MARS_JOIN or MARS_LEAVE messages to be transmitted. (The smallest scope managed by a MARS is scope 2 (link-local), and so this is the smallest scope that MARS message are issued for.)

IPv6 mrouters may be considered to be built of two parts - a forwarding engine, and an endpoint. The forwarding engine needs to be listening promiscuously across all multicast groups that need forwarding outside the link scope. The endpoint within a router needs to listen only on specific groups that have scope of link-local or larger.

To support the forwarding engine:

- IPv6 mrouters SHALL perform a block MARS_JOIN for the range(s) of IPv6 multicast addresses they require each ATM interface to listen on (described in section 9 of [5] for IPv4).
- They MUST NOT issue a block join for multicast addresses with scope of 1 (node-local) or 2 (link-local).

To support any internal endpoints, IPv6 mrouters SHALL perform a
single group MARS_JOIN for the following groups:

- Their derived Solicited-node address(es).
- The All-nodes address with link-local scope.
- The All-routers address with link-local scope.
- Other multicast groups to which endpoints within the mrouter belong with at least link-local scope.

It should be noted that the use of MARS for supporting the general case of IPv6 multicasting is independent of how Neighbor Discovery is implemented. This will be discussed further in section 5.

2.2 Unicast packet encapsulation.

The Ethertype assigned to IPv6 is 0x86DD [3]. Following the convention of RFC1483 for IPv4 unicast transmissions, the default encapsulation for a unicast IPv6 packet SHALL be:

```
[0xAA-AA-03][0x00-00-00][0x86-DD][IPv6 packet]
(LLC)       (OUI)     (PID)
```

Local administrators MAY choose to discard the LLC/SNAP encapsulation and use ‘VC multiplexing’. In this case an IPv6 packet is placed directly into an AAL5 AAL_SDU.

An IP/ATM interface SHALL accept IPv6 packets whose IP destination address is a multicast address, even if encapsulated as shown above. It SHALL only transmit packets using the above encapsulation if the IP destination is a unicast or anycast address.

2.3 Multicast packet encapsulation.

The encapsulation used for multicast IPv6 packets by MARS based IP/ATM interfaces SHALL be:

```
[0xAA-AA-03][0x00-00-5E][0x00-01][CMI][0x86-DD][IPv6 packet]
(LLC)       (OUI)     (PID)
```

The 2 octet Cluster Member ID (CMI) field is defined in [5].

Local administrators MAY choose to discard the LLC/SNAP encapsulation and use ‘VC multiplexing’. In this case the [CMI][0x86-DD][IPv6 packet] is placed directly into an AAL5 AAL_SDU.

An IP/ATM interface SHALL accept IPv6 packets whose IP destination address is not a multicast address, even if encapsulated as shown above. It SHALL only transmit packets using the above encapsulation if the IP destination is a multicast address.
3. IPv6 Link-Local address.

IPv6 nodes are required to generate a unique Link-Local IPv6 address for every link layer interface they have [6, 9]. Constructing these addresses requires an interface ID (or link-local token) that is at least unique amongst all the interfaces attached to the same link. Routers are not allowed to forward packets with Link-Local destinations, so it is not necessary for the interface ID to be unique across multiple independent links.

3.1 A Logical Link Group.

The ATM environment complicates the sense of the word ‘link’ in much the same way as it complicated the sense of ‘subnet’ in the IPv4 case. For IPv4 this required the definition of the Logical IP Subnet (LIS) – an administratively constructed set of hosts that would share the same routing prefixes (network and subnetwork masks).

For want of a better term this document considers the IPv6 analog to be a Logical Link Group – LLG.

An LLG consists of nodes administratively configured to be ‘on link’ with respect to each other. (This is described further in section 5.1)

Sets of hosts that are members of the same MARS Cluster [5] SHALL be taken from the membership of an LLG. (This is the analog of the current restriction that an IPv4 MARS Cluster is constructed from the multicast capable members of an LIS.)

It should be noted that whilst members of an LLG are IPv6 Neighbors, it is possible for Neighbors to exist that are not, administratively, members of the same LLG. This is discussed later in this document.

3.2 Choice of Interface ID.

The choice of interface ID is a compromise. You need to uniquely identify IPv6 interfaces that share the same Link, and a number space large enough to keep down the probability of different IPv6 interfaces generating identical Link-Local addresses. On the other hand, you want to keep the width (in bits) of the interface ID down because it impinges on the number of bits remaining to use as routing prefixes. It is preferable to choose the smallest unique identifier possible to maximise our ability to build hierarchy into the routing prefixes.

Using the model in section 3.1, the scope of uniqueness for a Link-
Local address is the LLG.

The IPv6 over Ethernet world suggests that a 48 bit interface ID is large enough for uniqueness and small enough to leave a useful routing prefix width. This 48 bit value is taken directly from the 48 bit MAC address associated with a node’s Ethernet interface - each Ethernet interface supporting a single IPv6 interface.

However, in the ATM environment we can find logical IP interfaces layered over logical ATM interfaces, themselves layered over a single physical ATM interface. If these logical IP interfaces are members of the same IPv6 Link (e.g. virtual hosts on a single physical machine) then each one needs a different interface ID in order to generate a different Link-Local address.

This issue currently lacks a consensus solution. The previous version of this ID proposed an oversized interface ID of 8 octets to cover all the possible ATM based virtual interfaces. This has been deprecated, but is described in Appendix A for reference.

Of the three Internet Drafts proposing solutions at the March 1996 IETF meeting, only Section 2 of [10] contained a concrete proposal for generating interface IDs. The author’s goal was to constrain the interface ID to be 6 octets wide, equivalent to the width of interface IDs on media such as Ethernet. A mechanism for generating 6 octet interface IDs is provided for the following cases:

- When a single IP interface is layered over a single ATM interface, and an IEEE MAC address (or a unique ESI field from an ATM Forum NSAP Address) uniquely identifies the ATM interface.
- When a single IP interface is layered over a single ATM interface, and an E.164 number uniquely identifies the ATM interface.

The suggested mechanisms for Duplicate Address Detection, and handling multiple logical IP interfaces per physical ATM interface, are closely coupled to the authors’ mechanism for Neighbor Discovery. Readers should consider section 2 of [10] while bearing in mind that at this stage there is no consensus on which ND approach is appropriate, and that there was no discussion either way on mechanisms for interface ID selection at the March 1996 meeting. This is a good start, but still an open issue.
4. ND link-layer address options.

Neighbor Discovery defines two options for carrying link-layer specific source and target addresses. In this case these options must carry full ATM addresses.

The source and target link-layer address options must carry any one of the three possibilities, and indicate which one it is.

The format for these two options when in an ATM environment is adapted from the MARS [] and NHRP [] specs, and SHALL be:

\[
\begin{array}{c}
\text{[Type]}
\end{array}
\begin{array}{c}
\text{[Length]}
\end{array}
\begin{array}{c}
\text{[NTL]}
\end{array}
\begin{array}{c}
\text{[STL]}
\end{array}
\begin{array}{c}
\text{[..ATM Number..]}
\end{array}
\begin{array}{c}
\text{[..ATM Subaddress..]}
\end{array}
\]

\[ | \hspace{1em} \text{Fixed} \hspace{1em} | \hspace{1em} \text{Link layer address} \hspace{1em} |
\]

[Type] is a one octet field.

1 for Source link-layer address. 2 for Target link-layer address.

[Length] is a one octet field.

The total length of the option in multiples of 8 octets. Zeroed bytes are added to the end of the option to ensure its length is a multiple of 8 octets. (For example, a single ATM address in NSAPA format would result in 24 bytes of real data, require no padding, and result in [Length] being set to 3.)

[NTL] is a one octet ‘Number Type & Length’ field.

Defines the type and length of the ATM number immediately following the [STL] field. The format is as follows:

\[
\begin{array}{c}
7 \ 6 \ 5 \ 4 \ 3 \ 2 \ 1 \ 0
\end{array}
\begin{array}{c}
|0| \hspace{1em} x \hspace{1em} | \hspace{1em} \text{length} \hspace{1em} |
\end{array}
\begin{array}{c}
|\text{0}+|\text{0}+|\text{0}+|\text{0}+|\text{0}+|\text{0}+|\text{0}+|\text{0}+
\end{array}
\]

The most significant bit is reserved and MUST be set to zero. The second most significant bit (x) is a flag indicating whether the ATM number is in:

- ATM Forum NSAPA format (x = 0).
- Native E.164 format (x = 1).

The bottom 6 bits is an unsigned integer value indicating the length of the associated ATM address in octets.
The [STL] is a one octet ‘Subaddress Type & Length’ field.

Format is the same as the [NTL] field. Defines the length of the subaddress field, if it exists. If it does not exist this entire octet field MUST be zero. If the subaddress exists it will be in NSAPA format, so flag x SHALL be zero.

[ATM Number] is a variable length field. It is always present.

[ATM Subaddress] is a variable length field. It may or may not be present. When it is not, the option ends after the [ATM Number] (or any additional padding for 8 byte alignment).

5. The wider implications of Neighbor Discovery.

The Neighbor Discovery protocol makes some assumptions about the underlying link layer service that are not immediately applicable in the ATM environment. ND assumes that multicast support is trivially available from the IP/link-layer interface. It also makes no clear statements about how ‘cut through’ unicast connections might be achieved - a concept that has acquired some prominence in the IP over ATM area through the development of NHRP [8] for IPv4 deployment.

As noted in section 2, multicast support needs to be emulated in UNI 3.0/3.1 environments.

The ‘on-link/off-link’ distinction for Neighbors might seem to lend itself to a Classical IP model of IPv6 over ATM (where IPv6 interfaces would only use direct ATM connections between members of their Logical Link Groups). However, as for IPv4, such administrative boundaries need to be ‘cut through’ to provide maximal use of the underlying ATM service.

A simplistic approach to ND would be to treat one’s MARS based IP/ATM interface as a black box that magically supports IP multicasting. The ND protocol and service will then appear to ‘work’ as designed. A key limitation is that there is no obvious way to achieve ‘cut through’ connections.

5.1 Neighbor Discovery and ‘cut through’ routing.

IPv6 contains a concept of on-link and off-link. Neighbors are those nodes that are considered on-link and whose link-layer addresses may therefore be located using Neighbor Discovery. Borrowing from the terminology definitions in the ND text:

on-link   - an address that is assigned to a neighbor’s interface on
a shared link. A host considers an address to be on-link if:
- it is covered by one of the link’s prefixes, or
- a neighboring router specifies the address as the target of a Redirect message, or
- a Neighbor Advertisement message is received for the target address, or
- a Router Advertisement message is received from the address.

off-link — the opposite of "on-link"; an address that is not assigned to any interfaces attached to a shared link.

Off-link nodes are considered to only be accessible through one of the routers directly attached to the link.

The preceding descriptions may need refinement in the context of Logical Link Groups (or equivalent concept). The LLG is the same set of hosts that make up a given MARS Cluster — an administratively defined group. These are an IPv6 interface’s initial set of neighbors, and each interface’s Link-Local address only needs to be unique amongst this set.

Events such as the receipt of ND advertisement messages, or the operation of some alternative discovery protocol, may result in the expansion of an IPv6 interface’s set of Neighbors. However, this should not be considered to have changed the set of interfaces that make up its LLG. This approach leads to three possible relationships between any two IPv6 interfaces:

- On LLG, Neighbor.
- Off LLG, Neighbor.
- Off LLG, not Neighbor.

Off LLG Neighbors are the ‘cut through’ connections, where some dynamic protocol activity has ascertained that although a target IPv6 interface is not a member of the source’s LLG, it is possible to achieve link level connectivity.

Whatever protocol we choose to locate IPv6 Neighbors should address the following issues:

- How do you perform Duplicate Address Detection?
- How do you decide who is on or off link (or LLG)?
- ND allows the targetted Neighbor to return different link layer addresses to every ND query. How do you retain this capability?
5.2 Solutions as of the March 1996 IETF.

There was no clear consensus on any one of the three ideas documented in [10], [11], and [13]. What follows is a brief summary of each proposal’s salient points. Readers are encouraged to locate the original (or subsequent) versions of these documents for more specific details. (Subsequent versions are identified by a numerical suffix higher than the ones listed here.)

5.2.1 draft-schulter-ipv6atm-framework-01.txt

The author builds upon the premise that the IPv6 stack’s interaction with an IP/ATM interface should be no different than its interaction with something like an IP/Ethernet interface. This means that Neighbor Discovery, as both a service and a protocol, should be run unchanged. The underlying IP/ATM service itself is required to perform certain special processing of ND messages to emulate the required functionality.

The author uses Logical Link (LL) as an analog of the LLG. To solve the need for multicasting ND messages around the LL, the author introduces Neighbor Discovery Servers (NDSs – essentially a multicast server for ND messages). A hierarchy of NDSs is then constructed to allow discovery messages to propagate outside an LL when necessary. This provides the ability to establish ‘cut through’ connections by discovering Off-LL neighbors.

Other IPv6 protocols, such as Router Discover, Duplicate Address Detection, and autoconfiguration are also supported transparently by the author’s hierarchy of NDSs.

The document currently makes no proposal for a mechanism to generate unique Link Local addresses.

5.2.2 draft-ahl-ipv6-nbma-00.txt
(or draft-ietf-ion-ipv6-nbma-00.txt)

The authors propose solutions to both the discovery of Off LLG Neighbors, and the generation of Link Local addresses.

A distinction is made between the ND service, and the ND protocol defined in [7]. The authors bypass the ICMPv6 based ND protocol itself, and provide a number of functional equivalents using extensions to NHRP. As distinct from [11], this model implies that IPv6 will perform ND according to [7] for some link technologies, and delegate a number of ND services to the link layer interface for other technologies such as NBMA networks.
The point is made that resolving next hops, and discovering neighbors, amounts to the same thing. General NBMA environments do not lend themselves to multicast based discovery mechanisms, so a logical alternative is the client-server based NHRP. Some extensions to NHRP are suggested in order for the NHRP client registration process to provide duplicate address detection.

While NHRP is demonstrated to replace the use of ICMPv6 messages for a number of ND services, the host protocol stack continues to process and act on incoming ICMPv6 based ND messages (e.g. for Neighbor Unreachability Detection, Redirects, etc).

5.2.3 draft-armitage-ipatm-tn-00.txt
(or draft-armitage-ion-tn-00.txt)

This document refines a proposal that the author presented in half-baked form during the IETF itself. It attempts to synthesize a solution for ND that utilizes parts of the NHRP model for discovering neighbors outside one’s LLG, and uses MARS emulation of multicast to allow the ND protocol described in [7] to run without change within an ATM based LLG.

The author postulates that egress routers from an LLG (which hosts use in the absence of more direct information) are in a position to detect the existence of IP traffic flows. Such flows are presumably amenable to ‘cut through’. The router generates a NHRP query in an effort to establish a ‘better’ link level point to cut through to. Once the query is resolved, the router multicasts an ND Redirect message (containing the discovered ATM address) to the LLG from which the traffic is originating. The source(s) of the traffic then have the option of cutting over to the ATM destination supplied in the Redirect message. These are considered to be Transient Neighbors.

Intra-LLG IPv6 Discovery services operate as defined in [7], and IPv6 hosts do not run NHRP to achieve cut-through. Identification of suitable flows is considered an open issue. No proposal is made for generating unique Link Local addresses.

Security Consideration

Security considerations are not addressed in this memo.

Acknowledgments

Sue Thomson (Bellcore) patiently answered my more inane questions during the initial stages of version 00. Peter Schulter, Ran Atkins, and others have ensured that the issues are being studied carefully.
within the wider IETF environment. Unless noted otherwise, errors are my own.

Author’s address

Grenville Armitage
Internetworking Research Group,
Bellcore.
445 South Street
Morristown, NJ, 07960-6438
USA

Email: gja@bellcore.com

References.


Appendix A: A deprecated form of Interface ID generation.

[This text originally proposed an 8octet field, but has been modified slightly to demonstrate the basic idea with a 7.5 octet field.]

An interface to an LLG is itself logical, and is supported by a logical ATM interface. The ATM Forum allows three possible variations of ATM addresses. These are:

- Native E.164 number.
- 20 byte ATM Forum NSAP format number.
- Native E.164 number with NSAP format subaddress.

When NSAP format addresses are in use, logical ATM interfaces are constructed over physical ATM interfaces by using different SEL within the context of a given ESI, or even having multiple ESIs route to the same physical interface. When native E.164 addresses are in use, each logical ATM interface requires its own E.164 number.

Therefore, the unique interface ID for the construction of Link-Local IPv6 addresses could be 7.5 octets wide and be constructed in one of two ways:

If the link interface has an NSAPA assigned, the 7 byte ESI+SEL value of the logical ATM interface being used by the IPv6 node is extracted and placed into the rightmost octets of the 7.5 octet interface ID. The leftmost semi-octet is reserved and MUST be set to zero.

If the link interface has only a native E.164 number assigned to it then a 7.5 octet BCD encoded version of the E.164 number is used to fill the field. The semi-octet value 1111 is used to pad out the field in cases where the E.164 number was less than the maximum 15 digits.

The Link-Local IPv6 address thus appears as:

```
| 10 bits | 58 bits | 60 bits |
+---------+---------+---------+
|1111111010| 0       | interface ID |
+---------+---------+----------+
```

58 zero bits pad out the IPv6 address between the interface ID and the 10 bit Link-Local prefix.