Fred L. Templin
SRI International
17 May 2001

Intra-Site Automatic Tunnel Addressing Protocol (ISATAP)

draft-ietf-ngtrans-isatap-01.txt

Abstract

This document specifies an intra-site automatic tunneling protocol (ISATAP) for connecting IPv6 hosts and routers (nodes) within predominantly IPv4-based networks. This method is based on an IPv6 aggregatable global unicast address format (described herein) that embeds the IPv4 address of a node within the EUI-64 format interface identifier. This document assumes that, during the IPv4 to IPv6 co-existence and transition phase, many sites will deploy IPv6 incrementally within their IPv4 interior routing domains; especially those sites which have large and complex pre-existing IPv4 infrastructures. Within such sites, the address format and methods described in this document will enable IPv6 deployment for nodes that do not share a common data link with an IPv6 gateway for their site.

While other works in progress in the NGTRANS working group propose mechanisms for assigning globally-unique IPv6 address prefixes to sites and methods for inter-domain routing between such sites, the approach outlined in this memo enables large-scale incremental deployment of IPv6 for nodes within a site's pre-existing IPv4 infrastructure without incurring aggregation scaling issues at the border gateways nor requiring site-wide deployment of special IPv4 services such as multicast. The approach proposed by this document supports IPv6 routing within both the site-local and global IPv6 routing domains as well as automatic IPv6 in IPv4 tunneling across portions of a site's IPv4 infrastructure which have no native IPv6 support. Additionally, this approach supports automatic tunneling within sites which use non globally-unique IPv4 address assignments, such as when Network Address Translation [NAT] is used.

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

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 and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet- Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

1. Introduction

The IETF NGTRANS working group anticipates an heterogeneous IPv4/IPv6 infrastructure in the near future and thus is chartered to develop mechanisms to support IPv4/IPv6 coexistence and transition toward global IPv6 deployment. For the most part, existing NGTRANS approaches focus on inter-domain routing between IPv6 islands using the existing global IPv4 backbone as transit. But, these islands may themselves comprise complex heterogeneous IPv4/IPv6 networks (e.g. large academic or commercial campus intranets) that require intra-domain IPv4 to IPv6 transition mechanisms and strategies as well. In order to address this requirement, this document presents a simple and scalable approach that enables incremental deployment of IPv6 nodes within predominantly IPv4-based intranets. We refer to this approach as the Intra-Site Automatic Tunnel Addressing Protocol, or ISATAP (pronounced: "ice-a-tap").

The ISATAP approach is based on an aggregatable global unicast address format that carries a standard 64-bit IPv6 address prefix [ADDR][AGGR] with a specially-constructed 64-bit EUI-64 Interface Identifier [EUI64]. This address format is fully compatible with both native IPv6 and NGTRANS routing practices (e.g. [6to4],[6BONE]). But, the interface identifier in an ISATAP address employs a special construction (using the IEEE Organizationally Unique Identifier (OUI) reserved by the Internet Assigned Numbers Authority [IANA]) that encapsulates an IPv4 address suitable for automatic IPv6-in-IPv4 tunneling. Since tunneling occurs only within the site-level prefix of the ISATAP address, the embedded IPv4 address NEED NOT be globally unique; rather, it need only be topologically correct for (and unique within) the context of the site.

This approach allows dual-stack nodes that do not share a common

datalink with an IPv6 gateway to join the global IPv6 network by automatically tunneling IPv6 messages through the IPv4 routing infrastructure within their site. Two methods for automatic discovery of an off-link IPv6 gateway for ISATAP address autoconfiguration are provided. This approach allows large-scale intra-site deployment without incurring aggregation scaling issues at the border gateways, since only a single IPv6 address prefix is used for the entire site. Finally, this approach supports intranets which use non-globally unique IPv4 addresses, such as when private address allocations [PRIVATE] and/or Network Address Translation [NAT] are used.

2. Changes

Major changes from version -00 to version -01:

- Revised draft to require *different* /64 prefixs for ISATAP addresses and native IPv6 addresses. Thus, a node's ISATAP interface is assigned a /64 prefix that is distinct from the prefixes assigned to any other interfaces attached to the node - be they physical or logical interfaces. This approach eliminates ISATAP-specific sending rules presented in earlier draft versions.

- Changed sense of 'u/l' bit in the ISATAP address interface identifier to indicate "local scope", since ISATAP interface identifiers are unique only within the scope of the ISATAP prefix. (See section 4.)

Major changes from version personal draft to NGTRANS WG version -00:

- Title change to provide higher-level description of field of use addressed by this draft. Removed other extraneous text.

- Major new section on automatic discovery of off-link IPv6 routers when IPv6-IPv4 compatibility addresses are used.

3. Terminology

The terminology of [IPv6] applies to this document. Additionally, the following terms are used extensively throughout this document:

ISATAP prefix:
  Any globally aggregatable 64-bit IPv6 routing prefix (whether from a native IPv6 assigned numbers authority or from a special-purpose numbering scheme such as [6BONE][6TO4]) reserved by a local network administrator

specifically for ISATAP purposes. ISATAP prefixes are used to configure ISATAP addresses ONLY; native IPv6 addresses SHOULD NOT be configured using an ISATAP prefix.

ISATAP address:
An IPv6 address with an ISATAP prefix and having an IPv4 address embedded in the interface identifier in the manner described in section 4 below.

ISATAP pseudo-interface:
ISATAP encapsulation of IPv6 packets inside IPv4 packets occurs at a point that is logically equivalent to an IPv6 interface, with the link layer being the IPv4 unicast network. This point is referred to as a pseudo-interface. An ISATAP pseudo-interface is assigned an ISATAP address through address autoconfiguration.

ISATAP router:
An IPv6 router supporting an ISATAP pseudo-interface. It is normally an interior router within an heterogeneous IPv6/IPv4 network.

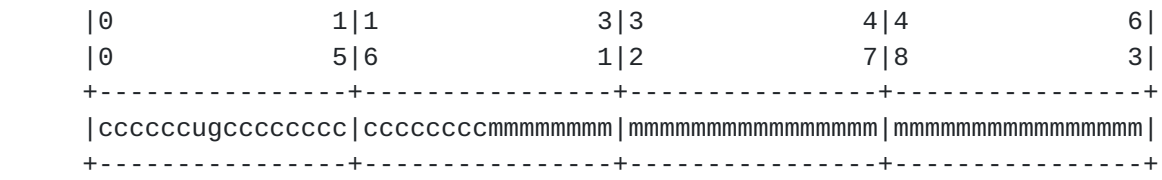
ISATAP host:
An IPv6 host which has an ISATAP pseudo-interface.

4. ISATAP Address Format

In sections 4.1 and 4.2, we will motivate our proposed extensions of the existing IEEE OUI reserved by IANA to support IEEE EUI-64 format addresses. While these proposed extensions are intended support the ISATAP address format, they also provide a flexible framework for future IANA use. Therefore, the extensions proposed in sections 4.1 and 4.2 may provide beneficial future use to IANA beyond the scope of ISATAP addresses. We present the ISATAP address format itself in sections 4.3 and 4.4.

4.1. IEEE EUI-64 Interface Identifiers in IPv6 Addresses

IPv6 aggregatable global and local-use unicast addresses [ADDR] include a 64-bit interface identifier in IEEE EUI-64 format [EUI64], which is specified as the concatenation of a 24-bit company_id value (also known as the OUI) assigned by the IEEE Registration Authority (IEEE/RAC) and a 40-bit extension identifier assigned by the addressing authority for that OUI. (Normally, the addressing authority is the organization to which the IEEE has allocated the OUI). IEEE EUI-64 interface identifiers are formatted as follows:

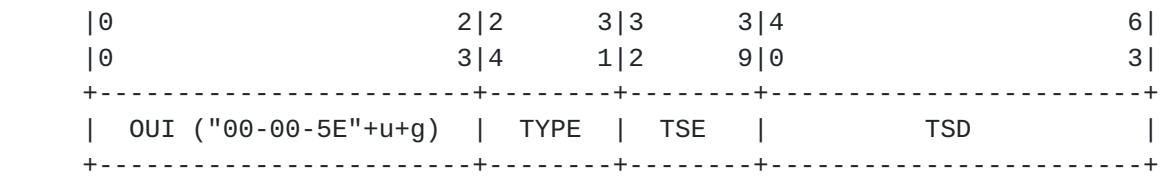


Where 'c' are the company-specific bits of the OUI, 'u' is the universal/local bit, 'g' is the individual/group bit and 'm' are the extension identifier bits. (NOTE: [ADDR] specifies that the 'u' bit is inverted from its normal sense in the IEEE context; therefore u=1 indicates global scope and u=0 indicates local scope).

In order to support encapsulation of legacy IEEE EUI-48 (24-bit) extension identifier values, [EUI64] specifies that the first two octets of the EUI-64 40-bit extension identifier (bits 24 through 39 of the EUI-64 address itself) SHALL BE 0xFFFE if the extension identifier encapsulates an EUI-48 value. [EUI64] further specifies that the first two octets of the extension identifier SHALL NOT be 0xFFFF, since this value is reserved by the IEEE/RAC. However, all other 40-bit extension identifier values are available for assignment by the OUI addressing authority.

4.2. An EUI-64 Interface Identifier Format for IANA

The IANA owns IEEE OUI: 00-00-5E, and [IANA] specifies EUI-48 format (24-bit) interface identifier assignments within that OUI. But, [IANA] does not specify how these legacy EUI-48 assignments will be written in EUI-64 format, nor does it specify a format for future 40-bit extension identifier assignments. We propose the following format for EUI-64 addresses within IANA's OUI reservation:



Where the fields are:

OUI IANA's OUI: 00-00-5E with 'u' and 'g' bits (3 octets)

TYPE Type field; indicates how (TSE, TSD) are interpreted (1 octet)

TSE Type-Specific Extension (1 octet)

TSD Type-Specific Data (3 octets)

And the following interpretations are defined based on TYPE:

TYPE	(TSE, TSD) Interpretation
0x00-0xFD	RESERVED for future IANA use
0xFE	(TSE, TSD) together contain an embedded IPv4 address
0xFF	TSD is interpreted based on TSE as follows:

TSE	TSD Interpretation
0x00-0xFD	RESERVED for future IANA use
0xFE	TSD contains 24-bit EUI-48 intf id
0xFF	RESERVED by IEEE/RAC

Essentially, if TYPE=0xFE, TSE is treated as an extension of TSD. If TYPE=0xFF, TSE is treated as an extension of TYPE. Other values for TYPE (and hence, other interpretations of TSE, TSD) are reserved for future IANA use. This format conforms to all requirements specified in [EUI64] and supports encapsulation of EUI-48 interface identifiers in the manner described by that document. For example, an existing IANA EUI-48 format multicast address such as:

    01-00-5E-01-02-03

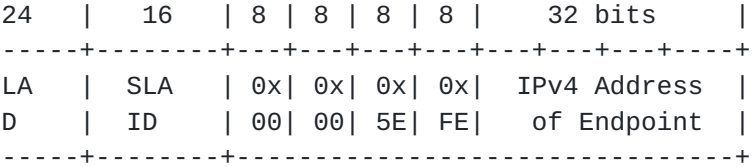
would be written in the IANA EUI-64 format as:

    01-00-5E-FF-FE-01-02-03

But, this proposed format also provides a special TYPE (0xFE) for embedding IPv4 addresses within the IANA 40-bit extension identifier. This special TYPE forms the basis for the ISATAP address format as described in the following sections.

4.3. ISATAP Address Construction

Using the proposed IANA-specific method for interface identifier construction discussed in sections 4.1 and 4.2 (with TYPE=0xFE), and with reference to [ADDR], we can construct an ISATAP address as follows:



(NOTE: since ISATAP address interface identifiers are interpreted

only within the local scope of the /64 ISATAP prefix, we set the u/l bit in the least significant octet of the OUI to '0' to indicate local scope.)

By way of example, an existing node with IPv4 address 140.173.129.8 might be assigned an IPv6 64-bit prefix of 3FFE:1a05:510:200::/64. We can then construct an ISATAP address for this node as:

3FFE:1a05:510:200:0:5EFE:8CAD:8108

or (perhaps more appropriately) written as the alternative form for an IPv6 address with embedded IPv4 address found in [ADDR]:

3FFE:1a05:510:200:0:5EFE:140.173.129.8

Similarly, we can construct the link-local and site-local variants (respectively) of the ISATAP address as:

FE80::0:5EFE:140.173.129.8
FEC0::200:0:5EFE:140.173.129.8

4.4. Advantages

By embedding an IPv4 address in the interface identifier portion of an IPv6 address as described in section 4.3, we can construct aggregatable global unicast IPv6 addresses that can either be routed globally via the IPv6 infrastructure or automatically tunneled locally across portions of a site's IPv4 infrastructure which have no native IPv6 support. Additionally, a node with an ISATAP address could act as a gateway for nodes with native IPv6 addresses with which it shares a common physical link, since the ISATAP node could automatically tunnel messages across a site's IPv4 domain on behalf of the native IPv6 nodes. An example would be deployment of IPv6 on some subset of the hosts attached to a workgroup's LAN. In this case, one host could configure an ISATAP address and act as a gateway for other hosts on the LAN which use native IPv6 addresses.

An additional advantage for our proposed method of embedding an IPv4 address in the interface identifier portion of an IPv6 address not found in other approaches such as [6TO4] is that large numbers of ISATAP addresses could be assigned within a common IPv6 routing prefix, thus providing maximal aggregation at the border gateways. For example, the single 64-bit IPv6 prefix:

3FFE:1a05:510:2412::/64

could include literally millions of nodes with ISATAP addresses.

This feature would allow a "sparse mode" IPv6 deployment such as the deployment of sparse populations of IPv6 hosts on large numbers of independent links throughout a large corporate Intranet.

A final important advantage is that this method supports both sites that use globally unique IPv4 address assignments and those that use non-globally unique IPv4 addresses, such as when private address assignments and/or Network Address Translation are used. By way of analogy to the US Postal system, inter-domain transition approaches such as [6TO4] provide means for routing messages "cross-country" to the "street address" of a distant site while the approach outlined in this document provides localized routing information to reach a specific (mailstop, apartment number, post office box, etc) WITHIN that site. Thus, the site-level routing information need not have relevance outside the scope of that site.

5. ISATAP Deployment Considerations

ISATAP addresses should only be used by nodes which do not share a common datalink with a native IPv6 router. At least one ISATAP router must be configured within the site which advertises an administratively- assigned ISATAP prefix in response to an Rtsol message from an off-link host. Such off-link hosts will configure an ISATAP pseudo-interface and assign it an address using the ISATAP prefix it receives in an Rtadv message solicited from an ISATAP router.

Following ISATAP address configuration, ISATAP hosts automatically and transparently communicate the IPv4 address of their *own* end of the ISATAP tunnel to any ISATAP host or router which uses the same ISATAP prefix. While nodes may optionally use stateful configuration to set an ISATAP prefix and a "default" route that points to an ISATAP router, a greatly preferred alternative is to provide for automatic intra-site IPv6 router discovery and stateless address autoconfiguration [DISCUSS]. The following section presents a means for the automatic discovery of ISATAP routers.

5.1. Automatic Discovery of ISATAP Routers

As described in [AUTO], a node that does not share a common multiple access datalink with an IPv6 router will NOT receive unsolicited Router Advertisements (Rtadv's), nor will Router Solicitations (Rtsol's) from that node reach an IPv6 router on the local link. But, the node may still be able to connect to the global IPv6 Internet if an ISATAP router for the site exists. Hence, a means for ISATAP router discovery is required. We present the following procedure for

a node to initiate ISATAP router discovery (and for an ISATAP router to respond) when an on-link IPv6 router is not available:

- The node constructs an ISATAP link local address for itself (as described in section 4.) as:

  FE80::0:5EFE:V4ADDR_NODE

- The node discovers the IPv4 address for an ISATAP router as: V4ADDR_RTR (**)

- The node sends an Rtsol to the IPv6 "all-routers-multicast" address tunneled through the IPv4 infrastructure to the ISATAP router's IPv4 address. The addresses used in the IPv6 and IPv4 headers are:

  ipv6_src: FE80::0:5EFE:V4ADDR_NODE
  ipv6_dst: FF02::2
  ipv4_src: V4ADDR_NODE
  ipv4_dst: V4ADDR_RTR

- Upon receiving the tunneled Rtsol, the ISATAP router sends a unicast Rtadv to the unicast address of the node which sent the Rtsol; again, by tunneling the Rtadv through IPv4. The addresses used in the IPv6 and IPv4 headers are:

  ipv6_src: FE80::0:5EFE:V4ADDR_RTR
  ipv6_dst: FE80::0:5EFE:V4ADDR_NODE
  ipv4_src: V4ADDR_RTR
  ipv4_dst: V4ADDR_NODE

- Upon receiving the Rtsol, the originating node performs address autoconfiguration as described in [AUTO] and constructs:

  - a fully-qualified ISATAP address for use as the source address for an ISATAP pseudo-interface

  - a default route that points to the ISATAP router

Note (**) that the above procedure assumes a means for discovering V4ADDR_RTR. We present two alternative methods for the automatic discovery of V4ADDR_RTR:

5.2. DNS Well-Known Service Name

The first method for discovering V4ADDR_RTR employs a new DNS Well-Known Service (WKS) name [DNS1,DNS2]. With the establishment of a new

well-known service name (e.g. "ISATAPGW"), administrators could publish the IPv4 address of a gateway which implementations could use to discover V4ADDR_RTR. This method has the advantage that it can be deployed immediately using existing mechanisms. However, it requires name service lookups and may not always provide the optimum V4ADDR_RTR resolution for isolated hosts if multiple ISATAP routers are available.

5.3. IPv4 Anycast for ISATAP routers

[6TO4ANY] proposes an IPv4 anycast prefix for 6to4 relay routers. The proposal suggests an IPv4 prefix assignment 'x.x.x.0/nn' ('nn' is currently proposed as 16) where the single address 'x.x.x.1' is assigned as the "6to4 IPv6 relay anycast address". We propose analogous assignments for the purpose of an "ISATAP router anycast address". (Whether the reservation of a second /32 assignment from the 6to4 IPv4 anycast prefix proposed in [6TO4ANY] would be possible, or a separate prefix assignment would be required is a matter of debate and TBD.)

ISATAP routers would advertise the ISATAP router anycast prefix via the intra-domain IPv4 routing infrastructure. Isolated IPv6 nodes would then use the ISATAP router anycast address as the V4ADDR_RTR IPv4 destination for off-link Rtsol's. This approach has the significant advantages that:

- implementations could hard-code the well-known ISATAP anycast address, thus avoiding service discovery via DNS

- an optimum path to an ISATAP router would be ensured by intra-domain IPv4 routing

As described above, the IPv4 anycast method for locating ISATAP routers provides significant functional advantages over the DNS approach, while the DNS approach can be implemented immediately pending the registration of a WKS name with IANA. While either method will work, the decision of which to push for standardization is TBD pending discussion at upcoming NGTRANS WG meetings.

6. Sending Rules and Routing Considerations

Since each node will be assigned an ISATAP prefix which is administratively reserved for use ONLY by ISATAP nodes, no special sending rules are needed. In particular, correspondent nodes that share a common ISATAP prefix will always exchange messages using their ISATAP pseudo-interfaces, whereas nodes that do not share a common ISATAP

prefix will always exchange messages via standard IPv6 routing. When sending a message on an ISATAP pseudo-interface, an implementation SHOULD verify that the IPv6 destination address employs the ISATAP address construction rules described in section 4 in order to detect mis-configured addresses. No other sending rules are necessary.

7. Address Selection

No special address selection rules are necessary.

8. Automatic Deprecation

ISATAP addresses are intended for use only by nodes which do not receive native IPv6 Rtadv's due to not sharing a common datalink with an IPv6 router. When native IPv6 Rtadv's become available (such as when an IPv6 router is deployed on a node's datalink), the node should construct a non-ISATAP aggregatable global IPv6 unicast address using address auto-configuration [AUTO] for a non-ISATAP IPv6 prefix discovered through normal means [DISC]. After the node's native IPv6 address is populated in the DNS, the node should eventually cease sending Rtsol's to the ISATAP router and discontinue use of its ISATAP pseudo-interface. In this way, ISATAP addresses will gradually (and automatically) disappear as IPv6 routers are widely deployed within sites.

9. Multicast Considerations

Other works in progress [6TO4MULTI] are currently investigating multicast addressing issues for [6TO4]. The address format discussed in this document is expected to be compatible with those emerging approaches.

10. IANA considerations

In order to support the EUI-64 address form described in this document, we propose that IANA adopt the EUI-64 Interface Identifier format specified in section 4.2 for the existing 00-00-5E OUI owned by IANA. No other actions are required by the IANA.

11. Security considerations

The ISATAP address format does not support privacy extensions for stateless address autoconfiguration [PRIVACY]. However, such privacy

extensions are intended primarily to avoid revealing one's MAC address, and the ISATAP address format described in this document accomplishes this same goal.

Additional security issues are called out in [6TO4] and probably apply here as well.

12. Implementation status

The author has implemented the mechanisms described in this draft through modifications to the FreeBSD 3.2-RELEASE [FBSD] operating system with the INRIA [INRIA] IPv6 distribution. A Linux implementation is planned for the June, 2001 timeframe.

Additionally, Windows XP RC1 will implement elements of the mechanism proposed in this paper.

Acknowledgements

The original ideas presented in this draft were derived from SRI contractual work. The author recognizes that ideas similar to those in this document may have already been presented by others and wishes to acknowledge any other such authors. The author also wishes to acknowledge the government contract administrators who sponsored the projects from which these works derived as well as his SRI colleagues with whom he has discussed and reviewed this work, including Monica Farah-Stapleton, Dr. Mike Frankel, J. Peter Marcotullio, Lou Rodriguez, and Dr. Ambatipudi Sastry.

The author acknowledges valuable input from numerous members of the NGTRANS community which has helped guide the direction of the draft. The list of contributors is too long to enumerate, but the input from the community has been vital to the draft's evolution. Alain Durand deserves special mention for contributing the title of this draft and the ISATAP acronym.

The author finally wishes to provide special acknowledgement to Dave Thaler, Art Shelest, Richard Draves, and others at Microsoft Research for their ideas on automatic discovery of off-link IPv6 routers. Much of the text in section on deployment considerations derives directly from discussions with Dave, Art, Rich and others.

References

[AGGR] Hinden., R, O'Dell, M., and Deering, S., "An IPv6 Aggregatable Global Unicast Address Format", RFC 2374, July 1998.

[ADDR] Hinden, R., and S. Deering, "IP Version 6 Addressing Architecture", RFC 2373, July 1998.

[AUTO] Thomson, S., and T. Narten, "IPv6 Stateless Address Autoconfiguration", RFC 2462, December 1998.

[DISC] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[DNS1] Mockapetris, P. "Domain names - concepts and facilities", STD 13, RFC 1034, November 1987.

[DNS2] Mockapetris, P. "Domain names - Implementation and Specification", STD 13, RFC 1035, November 1987.

[DNSSRV] Gulbrandsen, A., Vixie, P., and L. Esibov, "A DNS RR for specifying the location of services (DNS SRV)", RFC 2782, February 2000.

[EUI64] IEEE, "Guidelines for 64-bit Global Identifier (EUI-64) Registration Authority", http://standards.ieee.org/regauth/oui/tutorials/EUI64.html, March 1997

[IANA] Reynolds, J., and J. Postel, "Assigned Numbers", STD 2, USC/Information Sciences Institute, October 1994.

[IPV4] Postel, J., "Internet Protocol", RFC 791

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

[6TO4] Carpenter, B., and K. Moore, "Connection of IPv6 Domains via IPv4 Clouds", RFC 3056, February 2001.

[6TO4ANY] Huitema, C., "An anycast prefix for 6to4 relay routers", draft-ietf-ngtrans-6to4anycsat-02.txt (work in progress)

[6TO4MULTI] Thaler, D., "Support for Multicsat over 6to4 Networks", draft-ietf-ngtrans-6to4-multicast-00.txt (work in progress)

[MECH] Gilligan, R., and E. Nordmark, "Transition Mechanisms for IPv6 Hosts and Routers", RFC 2893, August 2000.

[SELECT] Draves, R., Default Address Selection for IPv6, draft-

ietf-
ipngwg-default-addr-select-00.txt (work in progress)

[FBSD] http://www.freebsd.org

[INRIA] ftp://ftp.inria.fr/network/ipv6/

[6BONE] Rockell, R., and R. Fink, RFC 2772, February 2000.

[PRIVATE] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G. J.,
and E. Lear, "Address Allocation for Private Internets", RFC 1918, February 1996.

[PRIVACY] Narten, T., R. Draves, "Privacy Extensions for Stateless Address
Autoconfiguration in IPv6", RFC 3041, January 2001.

[NAT] Egevang, K., and P. Francis, "The IP Network Address Translator (NAT)", RFC 1631, May 1994.

[DISCUSS] private discussions with Dave Thaler, Art Shelest, et al.

Authors Addresses

Fred L. Templin
SRI International
333 Ravenswood Ave.
Menlo Park, CA 94025, USA

Email: templin@erg.sri.com

Intellectual Property

PLACEHOLDER for full IETF IPR Statement if needed.

Full Copyright Statement