Network Working Group
Internet-Draft
Expires: December 7, 2006

J. Abley
Afilias Canada
M. Bagnulo
UC3M
June 5, 2006

Applicability Statement for the Level 3 Multihoming Shim Protocol (Shim6)
draft-ietf-shim6-applicability-01

Status of this Memo

This Internet-Draft will expire on December 7, 2006.

Copyright Notice

Abstract

This document discusses the applicability of the Shim6 IPv6 protocol element and associated support protocols to provide site multihoming capabilities in IPv6.

Table of Contents

1. Introduction

Site multi-homing is an arrangement by which a site may use multiple paths to the rest of the Internet, to provide better reliability for traffic passing in and out of the site than would be possible with a single path. Some of the motivations for operators to multi-home their network are described in [RFC3582].

In IPv4, site multi-homing is achieved by introducing the additional state required to allow session resilience over re-homing events to the global Internet routing system (sometimes referred to as the Default-Free Zone, or DFZ) [RFC4116]. There is concern that this approach will not scale [RFC3221].

In IPv6, site multi-homing in the style of IPv4 is not generally available to end sites due to a strict policy of route aggregation in the DFZ. Site multi-homing for sites without PI addresses is achieved by assigning multiple addresses to each host, one or more from each provider. This multi-homing approach provides no transport-layer stability across re-homing events.

Shim6 introduces transport-layer mobility across re-homing events using a layer-3 shim approach. State information relating to the multi-homing of two endpoints exchanging unicast traffic is retained on the endpoints themselves, rather than in the network. Communications between Shim6-capable hosts and Shim6-incapable hosts proceed as normal, but without the benefit of transport-layer stability. The Shim6 approach is thought to have better scaling properties with respect to the state held in the DFZ than the IPv4 approach.

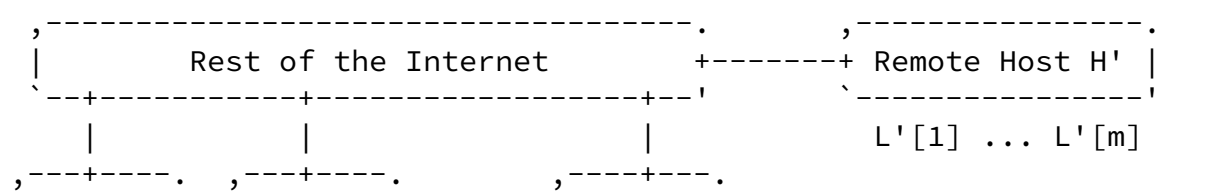
This note describes the applicability of the Level 3 multihoming (hereafter Shim6) protocol defined in [I-D.ietf-shim6-proto] and the failure detection mechanisms defined in [I-D.ietf-shim6-failure-detection].

2. Application scenarios

The goal of the Shim6 protocol is to support locator agility in established communications: different layer-3 endpoint addresses may be used to exchange packets as part of the same transport-layer session, all the time presenting a consistent identifier pair to upper-layer protocols.

In order to be useful, the Shim6 protocol requires that at least one of the peers has more than one address (locator). In the event of communications failure between an active pair of addresses, the Shim6 protocol will attempt to reestablish communication by trying different combinations of locators.

While the Shim6 protocol does not impose any requirements on the disposition of the locators involved in this process, the scenario in which the Shim6 protocol is expected to operate is that of a multi-homed site which is connected to multiple transit providers, and which receives an IPv6 prefix from each of them. This configuration is intended to provide protection for the end-site in the event of a failure in some subset of the available transit providers without requiring the end-site to acquire provider-independent (PI) address space.



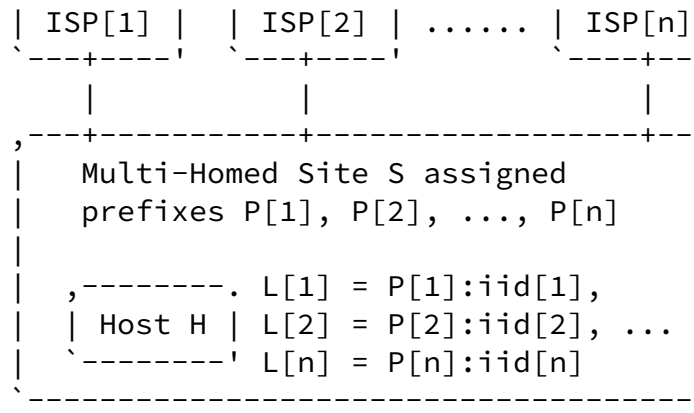


Figure 1

In the scenario illustrated in Figure 1 host H communicates with some remote host H'. Each of the addresses L[i] configured on host H in the multi-homed site S can be reached through provider ISP[i] only, since ISP[i] is solely responsible for originating a covering prefix for P[i] to the rest of the Internet.

The use of locator L[i] on H hence causes inbound traffic towards H to be routed through ISP[i]. Changing the locator used from L[i] to L[j] will have the effect of re-routing inbound traffic to H from ISP[i] to ISP[j]. This is the central mechanism by which the Shim6 protocol aims to provide multi-homing functionality: by changing locators, the H can change the upstream ISP used to route inbound packets towards itself. Corresponding control of the outbound path for packets from H towards H' is shared between the locator L'[j] chosen by H', and the administrative exit selection policy of site S.

The Shim6 protocol has other potential applications beyond site multi-homing. For example, since Shim6 is a host-based protocol, it can also be used to support session mobility between interfaces on a multi-homed host. A failure in communication between a multi-homed host and some other, remote host might be repaired by selection of a locator associated with a different interface.

3. Address Configuration

3.1. Protocol Version (IPv4 vs. IPv6)

The Shim6 protocol is defined only for IPv6. However, there is no fundamental reason why a Shim6-like approach could not support IPv4 addresses as locators, either to provide multi-homing support to IPv4-numbered sites, or as part of an IPv4/IPv6 transition strategy. Some extensions to the shim6 protocol for supporting IPv4 locators have been proposed in [I-D.nordmark-shim6-esd].

The Shim6 protocol, as specified for IPv6, incorporates cryptographic elements in the construction of locators (see [RFC3972], [I-D.ietf-shim6-hba]). Since IPv4 addresses are insufficiently large to contain addresses constructed in this fashion, direct implementation

of Shim6 as specified for IPv6 for use with IPv4 addresses might require protocol modifications.

3.2. Prefix Lengths

The Shim6 protocol does not assume that all the addresses assigned to the multihomed site have the same prefix length.

The use of CGA [RFC3972] and HBA [I-D.ietf-shim6-hba] involve encoding information in the lower 64 bits of locators. This imposes the requirement on address assignment to Shim6-capable hosts that all interface addresses should be able to accommodate 64-bit interface identifiers. This requirement is also imposed by CGA [RFC3972].

3.3. Address Generation

The security of the Shim6 protocol is based on the use of CGA and HBA addresses.

CGA and HBA can be generated through the stateless auto-configuration mechanism defined in [RFC2462] with the additional considerations presented in [RFC3972] and [I-D.ietf-shim6-hba].

Stateful address auto-configuration using DHCP [RFC3315] is not currently supported, because there is no defined mechanism to convey the CGA Parameter Data Structure and other relevant information from the DHCP server to the host. The definition of such mechanisms seems to be quite straightforward in the case of the HBA, since only the CGA Parameter Data Structure needs to be delivered from the DHCP server to the Shim6 host, and that data structure does not contain any secret information. In the case of CGAs, however, private key information must be exchanged as well as the CGA Parameter Data Structure.

3.4. Use of CGA vs. HBA

The choice between CGA and HBA is a trade-off between flexibility and performance.

The use of HBA is more efficient in the sense that addresses require less computation than CBA, involving only hash operations for both the generation and the verification of locator sets. However, with

HBA the locator set is determined during the generation process, and cannot be subsequently changed; addition of new locators to that initial set is not supported, except by re-generation of the entire set which will cause all addresses to change.

Use of CGA is more computationally expensive, involving public key cryptography in the verification of locator sets. However, CGAs are more flexible in the sense that they support the dynamic modification of locator sets.

CGAs are well suited to support dynamic environments such as mobile hosts, where the locator set must be changed frequently. HBAs are better suited for static sites where the prefix set remains relatively stable.

It should be noted that, since HBAs are defined as a CGA extension, it is possible to generate hybrid HBA/CGA structures that incorporate the strengths of both: i.e. that a single address can be used as an HBA, enabling computationally-cheap validation amongst a fixed set of addresses, and also as a CGA, enabling dynamic manipulation of the locator set. For additional details, see [I-D.ietf-shim6-hba].

4. Shim6 Capabilities

4.1. Fault Tolerance

4.1.1. Establishing Communications After an Outage

If a host within a multihomed site attempts to establish communication with a remote host outside the site while one of the site's transit paths has failed, and selects an local locator from which to source packets which corresponds to the failed transit path, bidirectional communication between the two hosts will not succeed. The failure of the transit path will not, in general, be known in advance to the host.

In order to establish communication, the initiating host must try different combinations of (source, destination) locator until it finds a pair that works. The mechanism for this default address selection is described in [RFC3484]; commentary on this mechanism in the context of multi-homed environments can be found in [I-D.bagnulo-ipv6-rfc3484-update].

Since Shim6 context is normally only established between two hosts after initial communication has been established, there is no opportunity for shim6 to participate in the discovery of a suitable, initial (source, destination) locator pair.

4.1.2. Short-Lived Communications

The Shim6 context establishment operation requires a 4-way packet exchange, and involves some overhead on the participating hosts in memory and CPU.

For short-lived exchanges between two hosts, the benefit of establishing a Shim6 context might not exceed the cost, perhaps because the protocols concerned are tolerant of failure and can arrange their own recovery (e.g. DNS) or because the frequency of re-homing events is sufficiently low that the probability of such a failure occuring during a short-lived exchange is not considered significant.

It is anticipated that the exchange of Shim6 context will provide most benefit for exchanges between hosts which are long-lived. For this reason the default behaviour of Shim6-capable hosts is expected to employ deferred context setup. This default behaviour will be able to be overridden by applications which prefer immediate context establishment regardless of transaction longeivity.

It must be noted that all the above considerations refer to lifetime

of the contact between the peers and not about the lifetime of the particular connection (e.g. TCP connection). In other words, the shim6 context is established between ULID pairs and it affects all the communication between these ULIDs. So, two nodes that perform multiple short lived communications with the same ULID pair would benefit as much from the shim features as two nodes having a single long-lived communication.

4.1.3. Long-Lived Communications

As discussed in Section 4.1.2, hosts engaged in long-lived communications will suffer lower proportional overhead, and greater probability of benefit than those performing brief transactions.

Deferred context setup ensures that session establishment time will not be increased by the use of Shim6.

4.2. Load Balancing

The Shim6 protocol does not support load balancing within a single context: all packets associated with a particular context are exchanged using a single locator pair per direction, with the exception of forked contexts which involve the upper-layer protocol.

It may be possible to extend the shim6 protocol to use multiple locator pairs in a single context, but the impact of such an extension on upper-layer protocols (e.g. on TCP congestion control) should be considered carefully.

When many contexts are considered together in aggregate, e.g. on a single host which participates in many simultaneous contexts or in a site full of hosts, some degree of load sharing should occur naturally due to the selection of different locator pairs in each context. There is no mechanism defined to ensure that this natral load sharing is arranged to provide a statistical balance between transit providers, however.

4.3. Traffic Engineering

The Shim6 protocol provides some lightweight traffic engineering capabilities in the form of the Locator Preferences option, which allows a host to inform a remote host of local preferences for locator selection.

This mechanism is only available after a Shim6 context has been established, and is a host-based capability rather than a site-based

capability. There is no defined mechanism which would allow use of the Locator Preferences option amongst a site full of hosts to be

managed centrally.

5. Interaction with Other Protocols

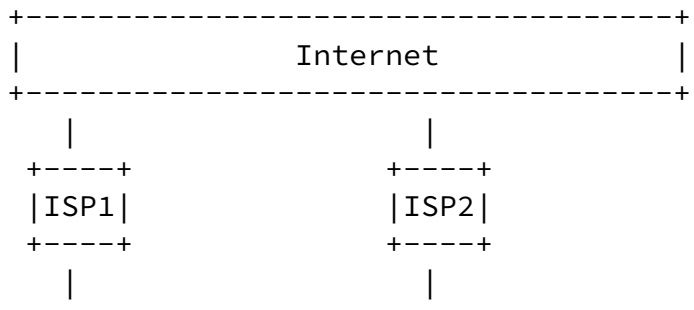
5.1. Shim6 and Mobile IPv6

Three scenarios where the Shim6 protocol and the MIPv6 protocol MIPv6 protocol [RFC3775] might be used simultaneously have been considered.

5.1.1. Multi-homed Home Network

In this case, the Home Network of the Mobile Node (MN) is multi-homed. This implies the availability of multiple Home Network prefixes, resulting on multiple HoAs for each MN. Since the MN is a node within a multihomed site, it seems reasonable to expect that the MN should able to benefit from the multihoming capabilities provided by the Shim6 protocol. Moreover, the MN needs to be able to obtain the multihoming benefits even when it is roaming away from the Home Network: if the MN is away from the Home Network while the Home Network suffers a failure in a transit path, the MN should be able to continue communicating using alternate paths to reach the Home Network.

The resulting scenario is the following:



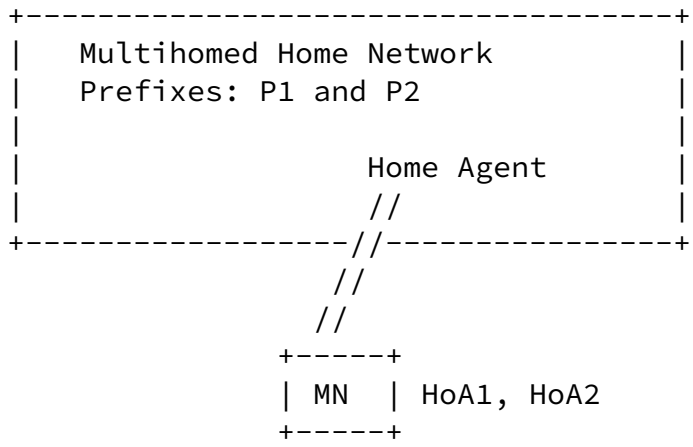


Figure 2

So, in this configuration, the shim6 protocol is used to provide

multihoming supports to all the nodes within the multihomed sites (including the mobile nodes) and the MIPv6 protocol is used to support mobility of the mobile nodes of the multihomed site.

The proposed protocol architecture would be the following:

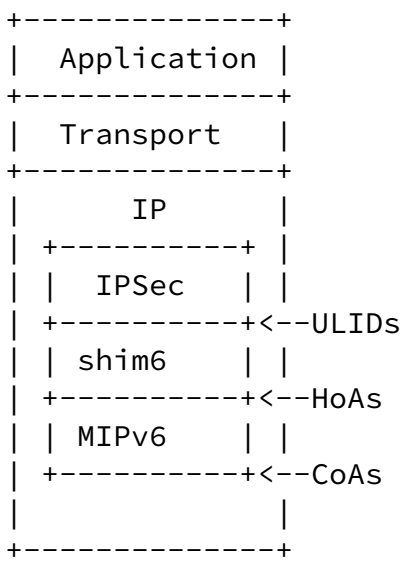


Figure 3

In this architecture, the upper layer protocols and IPSec would use ULIDs of the shim6 protocol. Only the HoAs will be presented to the shim6 layer as potential ULIDs. The shim6 protocol will then be used to provide failover between different HoAs. This is useful to preserve established communications when an outage affects the path through the ISP that has delegated the HoA used for initiating the communication (similarly to the case of a host within a multihomed site). The CoAs are not presented to the shim6 layer and are not included in the local locator set in this case. The CoAs are managed by the MIPv6 layer, that binds each HoA to a CoA.

So, in this case, the ULP select a ULID pair for the communication. The shim6 protocol translates the ULID pair to an alternative locator is case that is needed. Both the ULIDs and the alternative locators are HoAs. Next, the MIPv6 layer maps the selected HoA to the corresponding CoA, and this is the actual address included in the wire.

The shim6 context is established between the MN and the CN, and it would allow the communication to use all the available HoAs to provide fault tolerance. The MIPv6 protocol is used between the MN and the HA in the case of the bidirectional tunnel mode and between

the MN and the CN in case of the RO mode.

5.1.2. Shim6 between the HA and the MN

Another scenario where a shim6-MIPv6 interaction may be useful is the case where a shim6 context is established between the MN and the Home Agent (HA) in order to provide fault tolerance capabilities to the bidirectional tunnel between them.

Consider the case where the HA has multiple addresses (whether because the Home Network is multihomed or because the HA has multiple interfaces) and/or the MN has multiple addresses (whether because the visited network is multihomed or because the MN has multiple interfaces). In this case, if a failure affects the address pair that is being used to run the tunnel between the MN and HA, additional mechanisms need to be used to preserve the communication.

One possibility would be to use MIPv6 capabilities, by simply

changing the CoA used as the tunnel endpoint. However, MIPv6 lacks of failure detection mechanisms that would allow the MN and/or the HA to detect the failure and trigger the usage of an alternative address. Shim6 provides such failure detection protocol, so one possibility would be re-use the failure detection function from the shim6 failure detection protocol in MIPv6. In this case, the shim6 protocol wouldn&#65533;t be used to create shim6 context and provide fault tolerance, but just the failure detection functionality would be re-used.

The other possibility would be to use the shim6 protocol to create a shim6 context between the HA and the MN so that the shim6 detects any failure and re-homes the communication in a transparent fashion to MIPv6. In this case, the shim6 protocol would be associated to the tunnel interface

5.1.3. Shim6-based Route Optimization

Another scenario where it may be reasonable to support the simultaneous operation of MIPv6 and the shim6 protocol is to achieve some form of shim6-based Route Optimization. This case is similar to the one described in the multihomed home network section, the difference being that both the HoAs and the CoAs available in the MN are presented to the shim6 layer. The result is that the shim6 layer can select a CoA as an alternative address for an ongoing communication resulting in a route optimization mechanism. In this case, the shim6 protocol is used to provide both fault tolerance and route optimization, while the MIPv6 protocol is used for initial contact and non-optimized communications.

5.2. Shim6 and SeND

Secure Neighbour Discovery (SeND) [RFC3971] uses CGAs to prove address ownership for Neighbour Discovery [RFC2461]. The Shim6 protocol can use either CGAs or HBAs to protect locator sets included in Shim6 contexts. It is expected that some hosts will need to participate in both SeND and Shim6 simultaneously.

In the case that both the SeND and Shim6 protocols are using the CGA technique to generate addresses, then there is no conflict: the host will generate addresses for both purposes as CGAs, and since it will

be in control of the associated private key, the same CGA can be used for the different protocols.

In the case that a Shim6-capable host is using HBAs to protect its locator sets, the host will need to generate hybrid HBA/CGA addresses as defined in [I-D.ietf-shim6-hba] and discussed briefly in Section 3.4. In this case, the CGA Parameter Data Structure containing a valid public key and the Multi-Prefix extension is included as inputs to the hash function.

5.3. Shim6 and SCTP

The SCTP [RFC2960] protocol provides a reliable, stream-based communications channel between two hosts which provides a superset of the capabilities of TCP. One of the notable features of SCTP is that it allows the exchange of endpoint addresses between hosts, and is able to recover from the failure of a particular endpoint pair in a manner which is conceptually similar to locator selection in Shim6.

SCTP is a transport-layer protocol, higher in the protocol stack than Shim6, and hence there is no fundamental incompatibility which would prevent a Shim6-capable host from communicating using SCTP.

However, since SCTP and Shim6 both aim to exchange addressing information between hosts in order to meet the same general goal, it is possible that their simultaneous use might result in unexpected behaviour, e.g. due to race conditions.

The capabilities of SCTP with respect to path maintenance of a reliable, connection-oriented stream protocol are more extensive than the more general layer-3 locator agility provided by Shim6. It is recommended that Shim6 is not used for SCTP sessions, and that path maintenance is provided solely by SCTP.

5.4. Shim6 and NEMO

The NEMO [RFC3963] protocol extensions to MIPv6 allow a Mobile

Network to communicate through a bidirectional tunnel via a Mobile Router (MR) to a NEMO-compliant Home Agent (HA) located in a Home Network.

If either or both of the MR or HA are multi-homed, then a Shim6 context established between them preserves the integrity of the bidirectional tunnel between them in the event that a transit failure occurs between them. The MR in this case can be considered to be immobile either side of the failure event, and the Shim6 protocol provides a stable pair of ULIDs for the tunnel endopints.

Once the tunnel between MR and HA is established, hosts within the Mobile Network which are Shim6-capable can establish contexts with remote hosts in order to receive the same multi-homing benefits as any host located within the Home Network.

5.5. Shim6 and HIP

Placeholder.

6. Security considerations

This section will be completed before publication is requested.

7. Change History

This section should be removed prior to publication.

The list of Normative References to this document includes internet drafts; publication of those documents on the standards track is a prerequisite for the publication of this document, as-is.

draft-ietf-shim6-applicability-01: Added text for section 2 (Application scenarios), section 3 (About Address Configuration), section 4 (Resulting shim6 capabilities) and section 5 (Interactions with other protocols).

draft-ietf-shim6-applicability-00: First draft, largely incomplete, submitted to facilitate comments on general structure and approach.

8. Contributors

The anlysis on the interaction between the Shim6 protocol and the other protocols presented in this note benefited from the advice of various people including Erik Nordmark, Hesham Soliman, Vijay Devarpalli, John Loughney and Dave Thaler.

9. Acknowledgements

Joe Abley's work was supported in part by the US National Science Foundation (research grant SCI-0427144) and DNS-OARC.

Marcelo Bagnulo worked on this document while visiting Ericsson Research Laboratory Nomadiclab.

10. References

10.1. Normative References

[I-D.ietf-shim6-failure-detection]
Arkko, J. and I. Beijnum, "Failure Detection and Locator Pair Exploration Protocol for IPv6 Multihoming", draft-ietf-shim6-failure-detection-03 (work in progress), December 2005.

[I-D.ietf-shim6-hba]
Bagnulo, M., "Hash Based Addresses (HBA)", draft-ietf-shim6-hba-01 (work in progress), October 2005.

[I-D.ietf-shim6-proto]
Bagnulo, M. and E. Nordmark, "Level 3 multihoming shim protocol", draft-ietf-shim6-proto-05 (work in progress), May 2006.

[I-D.nordmark-shim6-esd]
Nordmark, E., "Extended Shim6 Design for ID/loc split and Traffic Engineering", draft-nordmark-shim6-esd-00 (work in progress), February 2006.

[RFC2461] Narten, T., Nordmark, E., and W. Simpson, "Neighbor Discovery for IP Version 6 (IPv6)", RFC 2461, December 1998.

[RFC2462] Thomson, S. and T. Narten, "IPv6 Stateless Address Autoconfiguration", RFC 2462, December 1998.

[RFC2960] Stewart, R., Xie, Q., Morneault, K., Sharp, C., Schwarzbauer, H., Taylor, T., Rytina, I., Kalla, M., Zhang, L., and V. Paxson, "Stream Control Transmission Protocol", RFC 2960, October 2000.

[RFC3315] Droms, R., Bound, J., Volz, B., Lemon, T., Perkins, C., and M. Carney, "Dynamic Host Configuration Protocol for IPv6 (DHCPv6)", RFC 3315, July 2003.

[RFC3484] Draves, R., "Default Address Selection for Internet Protocol version 6 (IPv6)", RFC 3484, February 2003.

[RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, June 2004.

[RFC3963] Devarapalli, V., Wakikawa, R., Petrescu, A., and P. Thubert, "Network Mobility (NEMO) Basic Support Protocol",

RFC 3963, January 2005.

[RFC3971] Arkko, J., Kempf, J., Zill, B., and P. Nikander, "SEcure Neighbor Discovery (SEND)", RFC 3971, March 2005.

[RFC3972] Aura, T., "Cryptographically Generated Addresses (CGA)", RFC 3972, March 2005.

10.2. Informative References

[I-D.bagnulo-ipv6-rfc3484-update]
Bagnulo, M., "Updating RFC 3484 for multihoming support",

draft-bagnulo-ipv6-rfc3484-update-00 (work in progress), December 2005.

[RFC3221] Huston, G., "Commentary on Inter-Domain Routing in the Internet", RFC 3221, December 2001.

[RFC3582] Abley, J., Black, B., and V. Gill, "Goals for IPv6 Site-Multihoming Architectures", RFC 3582, August 2003.

[RFC4116] Abley, J., Lindqvist, K., Davies, E., Black, B., and V. Gill, "IPv4 Multihoming Practices and Limitations", RFC 4116, July 2005.

Authors' Addresses

Joe Abley
Afilias Canada, Inc.
Suite 204

4141 Yonge Street
Toronto, Ontario M2P 2A8
Canada

Phone: +1 416 673 4176
Email: jabley@ca.afilias.info
URI: http://afilias.info/

Marcelo Bagnulo
Universidad Carlos III de Madrid
Av. Universidad 30
Leganes, Madrid 28911
Spain

Phone: +34 91 6248814
Email: marcelo@it.uc3m.es
URI: http://www.it.uc3m.es/

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

Funding for the RFC Editor function is currently provided by the Internet Society.