Individual Submission
Internet-Draft
Intended status: Informational
Expires: August 13, 2012

J. Abley
ICANN
M. Bagnulo
A. Garcia-Martinez
UC3M

February 10, 2012

Applicability Statement for the Level 3 Multihoming Shim Protocol (Shim6)

draft-garcia-shim6-applicability-03

Abstract

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

Status of this Memo

Copyright Notice

Table of Contents

1. Introduction

Site multihoming 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 multihoming is achieved by injecting into the global Internet routing system (sometimes referred to as the Default-Free Zone, or DFZ) the additional state required to allow session resilience over re-homing events [RFC4116]. There is concern that this approach will not scale [RFC3221], [RFC4984].

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

Shim6 provides layer-3 support for making re-homing events transparent to the transport layer by means of a shim approach. Once a Shim6 session has been established, the failure detection mechanism defined for Shim6 allows finding new valid locator combinations in case of failure, and using these locators to continue the communication. However, Shim6 does not provide failure protection to the communication establishment, so if a host within a multihomed site attempts to establish a communication with a remote host and selects an address which corresponds to a failed transit path, the communication will fail. State information relating to the multihoming 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. In order to successfully deploy Shim6 in a multihomed site, additional mechanisms may be required to solve issues such as selecting the source address appropriate to the destination and to the outgoing provider, or to allow the network manager to perform traffic engineering. Such problems are not specific of Shim6, but are suffered by the hosts of any site which is connected to multiple transit providers, and which receives an IPv6 prefix from each of the providers. Some of these mechanisms are not defined today. However, note that once a Shim6 session has been established, Shim6 reduces the impact of these problems, because if a working path exists, Shim6 will find it.

This note describes the applicability of the Level 3 multihoming (hereafter Shim6) protocol defined in [RFC5533] and the failure detection mechanisms defined in [RFC5534].

The terminology used in this document, including terms like locator, and ULID, is defined in [RFC5533].

2. Deployment 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 belonging to 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 which could be used on the wire (as locators). In the event of communications failure between an active pair of addresses, the Shim6 protocol attempts to reestablish communication by trying different combinations of locators.

While other multi-addressing scenarios are not precluded, the scenario in which the Shim6 protocol is expected to operate is that of a multihomed 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 PI address space or requiring any particular cooperation between the transit providers.

```
      ,-----------------------------------.       ,-----------------.
      |        Rest of the Internet       +-------+ Remote Host R  |
      `--+-----------+------------------+--'       `-----------------'
         |           |                  |           LR[1] ... LR[m]
     ,---+----.  ,---+----.        ,----+---.
     | ISP[1] |  | ISP[2] | ...... | ISP[n] |
     `---+----'  `---+----'        `----+---'
         |           |                  |
     ,---+-----------+------------------+---.
     |   Multi-Homed Site S assigned        |
     |   prefixes P[1], P[2], ..., P[n]     |
     |                                      |
     |  ,--------. L[1] = P[1]:iid[1],      |
     |  | Host H | L[2] = P[2]:iid[2], ...  |
     |  `--------' L[n] = P[n]:iid[n]       |
     `--------------------------------------'
```

Figure 1

In the scenario illustrated in Figure 1 host H communicates with some remote host R. Each of the addresses L[i] configured on host H in the multihomed site S can be reached through provider ISP[i] only, since ISP[i] is solely responsible for advertising 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 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 multihoming functionality: by changing locators, host H can change the upstream ISP used to route inbound packets towards itself. Regarding the outbound traffic to H, the path taken in this case depends on both the actual locator LR[j] used for R, and the administrative exit selection policy of site S. As discussed in Section 4, the site should deliver outgoing packets having a source address derived from the prefix of ISP[i] to that particular provider, in order to prevent those packets to be filtered due to Ingress Filtering [RFC2827] being applied by the providers. It is worth noting that in an scenario such as the one depicted in Figure 1, the paths followed by inbound and outbound traffic are determined to a large extent by the locators in use for the communication. This is not a particular issue of Shim6, but it is common to any deployment in which hosts are configured with addresses received from different providers. Traffic Engineering in such sites will likely involve proper configuration of address selection policies in the hosts, by means of mechanisms such as the ones discussed in Section 4.

The Shim6 protocol has other potential applications beyond site multihoming. For example, since Shim6 is a host-based protocol, it can also be used to support host multihoming. In this case, a failure in communication between a multihomed host and some other remote host might be repaired by selecting a locator associated with a different interface.

3. Addresses and Shim6

3.1. Protocol Version (IPv4 vs. IPv6)

The Shim6 protocol is defined only for IPv6. While some Shim6-like approaches have been suggested to support IPv4 addresses as locator [I-D.nordmark-shim6-esd], at this time it is not clear if such extensions are feasible.

The Shim6 protocol, as specified for IPv6, incorporates cryptographic elements in the construction of locators (see [RFC3972], [RFC5535]). 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 is not possible.

In addition, there are other factors to take into account when considering the support of IPv4 addresses, in particular IPv4 locators. Using multiple IPv4 addresses in a single host in order to support Shim6 style of multihoming would result in an increased IPv4 address consumption, which would be problematic considering the current situation in which IPv4 address space has been exhausted. Besides, Shim6 may suffer additional problems if locators become translated on the wire. Address translation is more likely to involve IPv4 addresses. IPv4 addresses can be translated to other IPv4 addresses (for example, private IPv4 address into public IPv4 address and vice versa) or to/from IPv6 addresses (for example, as defined by NAT64 [RFC6146]). When address translation occurs, a locator exchanged by Shim6 could be different to the address needed to reach the corresponding host, either because the translated version of the locator exchanged by Shim6 is not known or because the translation state does not exist any more in the translator device. Besides, the translated locators will not be verifiable with the current CGA and HBA verification mechanisms, which protect the locators as seen by the node for which they are configured.

3.2. Prefix Lengths

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

However, the use of CGA [RFC3972] and HBA [RFC5535] involve encoding information in the lower 64 bits of the 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. It should be noted that this is imposed by RFC4291 [RFC4291].

3.3. Address Generation and Configuration

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

CGA and HBA generation process can use the information provided by the stateless auto-configuration mechanism defined in [RFC4862] with the additional considerations presented in [RFC3972] and [RFC5535].

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 mechanism 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 this data structure does not contain any secret information. In the case of CGAs, the difficulty is increased, since private key information should be exchanged as well as the CGA Parameter Data Structure. However, with appropriate extensions a DHCP server could inform to a host about the SEC value to use when generating an address, or DHCP could even be used by the host to delegate to the server the CPU-intensive task of computing a Modifier for a given <prefix, public key, SEC> combination [I-D.ietf-csi-dhcpv6-cga-ps].

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 CGA, involving only hash operations for both the generation and the verification of locator sets. However, the locators of an HBA set are determined during the generation process, and cannot be subsequently changed; the addition of new locators to that initial set is not supported, except by re-generation of the entire set which will in turn cause all addresses to change.

The 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.

Therefore, 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 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 [RFC5535].

There is no current mechanism designed to allow an administrator to enforce the configuration of a CGA or an HBA in a host.

4. Shim6 in Multihomed Nodes

Shim6 multihomed nodes are likely to suffer problems related to the attachment to different provision domains. Note that these problems are not specific of Shim6. [I-D.ietf-mif-problem-statement] discusses the problems associated to nodes with multiple interfaces, which may involve difficulties in

- o managing the configuration associated to different providers
- o finding the appropriate DNS server to resolve a query and to match DNS answers to providers
- o routing the packets to the right provider
- o selecting the source address appropriate to the destination and to the outgoing provider
- o performing session management appropriately

Some of these problems may also arise in single-interface hosts connected to multiple networks, for example in configurations in which a customer network receives multiple Provider Aggregatable prefixes. These problems are suffered by other solutions supporting multihoming such as SCTP [RFC4960] or HIP [RFC4423]. Note also that single-homed nodes implementing Shim6 to improve communications with other nodes having multiple addresses will not suffer these problems.

The compatibility of Shim6 with configurations or mechanisms developed to solve any multihoming problem has to be carefully considered in a case-by-case basis. However, the interaction of Shim6 with some of the solutions discussed in [I-D.ietf-v6ops-ipv6-multihoming-without-ipv6nat] is commented in the next paragraphs.

In order to configure source and destination address selection, tools such as DHCPv6 can be used to disseminate a [RFC3484] policy table to a host. The impact to Shim6 of a solution which disseminates the policy table to the hosts is the following: Shim6 selects the ULID pair to use in a communication according to the mechanism described in [RFC3484]. In case different locator pairs need to be explored, nodes also use the rules defined by [RFC3484] to identify valid pairs, and to establish an order among them, as described in [RFC5534].

Shim6 has no means to enforce neither host nor network forwarding for a given locator to be used as source address. For IPv6 nodes, the next hop router to use for a given set of destinations can be configured through Extensions to Router Advertisements through Default Router Preference and More-Specific Routes [RFC4191], the use of a DHCPv6 option, or the use of a routing protocol. It is also possible to rely on routers considering source addresses in their forwarding decisions in addition to the usual destination-based forwarding. All these solutions are compatible with Shim6 operation. Note that an improper matching of source address and egress provider may result in packets being dropped if the provider performs Ingress Filtering [RFC2827], i.e. dropping packets which come from customer networks with source addresses not belonging to the prefix assigned to them, to prevent address spoofing.

For some particular configurations, i.e. for a walled-garden or closed service, the node may need to identify the most appropriate DNS server to resolve a particular query. For an analysis of this problem, the reader is referred to [I-D.ietf-v6ops-ipv6-multihoming-without-ipv6nat].

Finally, it is worth to note that Shim6 is built to handle communication problems, so it may recover from the misconfiguration (or lack) of some of the mechanisms used to handle the aforementioned problems. For example, if any notification is received from the router dropping the packets with legitimate source addresses as a result of ingress filtering, the affected locator could be associated to a low preference (or not being used at all). But even if such notification is not received, or not processed by the Shim6 layer, defective source address or next-hop selection will be treated as a communication failure, and Shim6 re-homing could finally select a working path in which packets are not filtered, if this path exists. This behavior results from the powerful end-to-end resilience properties exhibited by REAP.

5. Shim6 Capabilities

5.1. Fault Tolerance

5.1.1. Establishing Communications After an Outage

If a host within a multihomed site attempts to establish a communication with a remote host and selects a locator which corresponds to a failed transit path, bidirectional communication between the two hosts will not succeed. In order to establish a new communication, the initiating host must try different combinations of (source, destination) locator pairs until it finds a pair that works. The mechanism for this default address selection is described in [RFC3484]. As a result of the use of this mechanism, some failures may not be recovered even if a valid alternative path exists between two communicating hosts. For example, assuming a failure in ISP[1] (see Figure 1), and host H initiating a communication with host R, the source address selection algorithm described in [RFC3484] may result in the selection of the source address corresponding to ISP[1] for every destination address being tried by the application. However, note that if R is the node initiating the communication, it will find a valid path provided that the application at R tries every available address for H.

Since a Shim6 context is normally established between two hosts only after initial communication has been set up, there is no opportunity for Shim6 to participate in the discovery of a suitable, initial (source, destination) locator pair. The same consideration holds for referrals, as it is described in Section 6.

5.1.2. 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 communications between two hosts, the benefit of establishing a Shim6 context might not exceed the cost, perhaps because the protocols concerned are fault tolerant 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 occurring 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-establishment. This default behaviour

will be able to be overridden by applications which prefer immediate context establishment regardless of transaction longevity.

It must be noted that all the above considerations refer to the lifetime of the interaction between the peers and not about the lifetime of a 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 with multiple short-lived communications using the same ULID pair would benefit as much from the Shim6 features as two nodes having a single long-lived communication. One example of such scenario would be a web client software downloading web contents from a server over multiple TCP connections. Each TCP connection is short-lived, but the communication/contact between the two ULID could be long-lived.

5.1.3. Long-Lived Communications

As discussed in Section 5.1.2, 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.

5.2. 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 are created upon explicit requests from 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 aggregation, 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. However, there is no mechanism defined to ensure that this natural load sharing is arranged to provide a statistical balance between transit providers.

It is worth to note that the use of transport-layer solutions enhanced with mechanisms to allow the use of multiple paths for a transport session are more amenable for achieving load-balancing.

One such solution is being developed at the MP-TCP Working Group.

5.3. Traffic Engineering

For sites with prefixes obtained from different providers, the paths followed by inbound and outbound traffic are determined to a large extent by the locators selected for each communication. This is not a particular issue of Shim6, but it is common to any deployment in which hosts are configured with addresses received from different providers. Traffic Engineering in such sites will likely involve proper configuration of the address selection policies defined by [RFC3484].

Besides, 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. In this way, the host can influence in the incoming path for the communication. This mechanism is only available after a Shim6 context has been established, and it 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 by the administrator of the site.

6. Application Considerations

Shim6 provides multihoming support without forcing changes in the applications running on the host. The fact that an address has been generated according to the CGA or HBA specification does not require any specific action from the application, e.g. it can obtain remote CGA or HBA addresses as a result of a getaddrinfo() call to trigger a DNS Request. The storage of CGA or HBA addresses in DNS does not require also any modification of this protocol, since they are recorded using AAAA records. Moreover, neither the ULID/locator management [RFC5533] nor the failure detection and recovery [RFC5534] functions require application awareness.

However, a specific API [RFC6316] is developed for those applications which might require additional capabilities in ULID/locator management, such as the locator pair in use for a given context, or the set of local or remote locators available for it. This API can also be used to disable Shim6 operation when required.

It is worth noting that callbacks can benefit naturally from Shim6 support. In a callback, an application in B retrieves IP_A, the IP address of a peer A, and B uses IP_A to establish a new communication with A. As long as the address exchanged, IP_A is the ULID for the

initial communication between A and B, and B uses the same address as in the initial communication, and this initial communication is alive (or the context has not been deleted), the new communication could use the locators exchanged by Shim6 for the first communication. In this case, communication could proceed even if the ULID of A is not reachable.

However, Shim6 does not provide specific protection to current applications when they use referrals. A referral is the exchange of the IP address IP_A of a party A by party B to party C, so that party C could use IP_A to communicate with party A. In a normal case, the ULID IP_A would be the only information sent by B to C as referral. But if IP_A is no longer valid as locator in A, C could have trouble in establishing a communication with A. Increased failure protection for referrals could be obtained if B exchanged the whole list of alternative locators of A, although in this case the application protocol should be modified. Note that B could send to C the current locator of A, instead of the ULID of A, as a way of using the most recent reachability information about A. While in this case no modification of the application protocol is required, some concerns arise: host A may not accept one of its locator as ULID for initiating a communication, and if CGA are used, the locator may not be a CGA so a Shim6 context among A and C could not be created.

7. Interaction with Other Protocols and Mechanisms

In this section we discuss the interaction between Shim6 and other protocols and mechanisms. Before starting the discussion, it is worth noting that at the time of this writing there is a lack of experience with the combination of Shim6 and these protocols and mechanisms. Therefore, the conclusions stated should be review as real experience is gained in the use of Shim6.

7.1. Shim6 and Mobile IPv6

We next consider some scenarios in which the Shim6 protocol and the MIPv6 protocol [RFC6275] might be used simultaneously.

7.1.1. Multihomed Home Network

In this case, the Home Network of the Mobile Node (MN) is multihomed. 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 be 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:

```
       +-------------------------------------+
       |               Internet              |
       +-------------------------------------+
          |                   |
        +----+             +----+
        |ISP1|             |ISP2|
        +----+             +----+
          |                   |
       +-------------------------------------+
       |   Multihomed Home Network           |
       |   Prefixes: P1 and P2               |
       |                                     |
       |                   Home Agent        |
       |                   //                |
       +------------------//-----------------+
                         //
                        //
                      +-----+
                      | MN  | HoA1, HoA2
                      +-----+
```

Figure 2

So, in this configuration, the Shim6 protocol is used to provide multiple communication paths 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:

```
                        +--------------+
                        |  Application |
                        +--------------+
                        |  Transport   |
                        +--------------+
                        |      IP      |
                        | +----------+ |
                        | |  IPSec   | |
                        | +----------+<--ULIDs
                        | | Shim6    | |
                        | +----------+<--HoAs
                        | | MIPv6    | |
                        | +----------+<--CoAs
                        |              |
                        +--------------+
```

Figure 3

In this architecture, the upper layer protocols and IPSec would use ULIDs of the Shim6 protocol. Only the HoAs will be presented by the upper layers to the Shim6 layer as potential ULIDs. Two Shim6 entities will exchange their own available HoAs as locators. Therefore, Shim6 provides failover between different HoAs and allows preserving 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, which binds each HoA to a CoA.

So, in this case, the upper layer protocols select a ULID pair for the communication. The Shim6 protocol translates the ULID pair to an alternative locator in 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, which 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 (Route Optimization) mode.

7.1.2. 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 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-using the failure detection function from the Shim6 failure detection protocol in MIPv6. In this case, the Shim6 protocol wouldn't be used to create Shim6 context and provide fault tolerance, but just its 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.

7.2. Shim6 and SEND

Secure Neighbor Discovery (SEND) [RFC3971] uses CGAs to prove address ownership for Neighbor Discovery [RFC4861]. 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 [RFC5535] 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 are included as inputs to the hash function.

7.3. Shim6 and SCTP

The SCTP [RFC4960] 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 generic goal, it is possible that their simultaneous use might result in unexpected behaviour, e.g. lead 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. Therefore, it is recommended that Shim6 is not used for SCTP sessions, and that path maintenance is provided solely by SCTP. There are at least two ways to enforce this behaviour. One option would be to make the stack, and in particular the Shim6 sublayer, aware of SCTP sockets and in this case refrain from creating a Shim6 context. The other option is that the upper layer, SCTP in this case, informs using a Shim6-capable API like the one proposed in [RFC6316] that no Shim6 context must be created for this particular communication.

Note that the issues described here for SCTP may also arise for a multipath TCP solution.

7.4. 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 multihomed, then a Shim6 context established preserves the integrity of the bidirectional

tunnel between them in the event that a transit failure occurs in the connecting path.

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 multihoming benefits as any host located within the Home Network.

7.5. Shim6 and HIP

Shim6 and the Host Identity Protocol (HIP [RFC4423]) are architecturally similar in the sense that both solutions allow two hosts to use different locators to support communications between stable ULIDs. The signaling exchange to establish the demultiplexing context on the hosts is very similar for both protocols. However, there are a few key differences. First, Shim6 avoids defining a new namespace for ULIDs, preferring instead to use a routable locator as a ULID, while HIP uses public keys and hashes thereof as ULIDs. The use of a routable locator as ULID better supports deferred context establishment, application callbacks, and application referrals, and avoids management and resolution costs of a new namespace, but requires additional security mechanisms to securely bind the ULID with the locators. Second, Shim6 uses an explicit context header on data packets for which the ULIDs differ from the locators in use (this header is only needed after a failure/rehoming event occurs), while HIP may compress this context-tag function into the ESP SPI field [RFC5201]. Third, HIP as presently defined requires the use of public-key operations in its signaling exchange and ESP encryption in the data plane, while the use of Shim6 requires neither (if only HBA addresses are used). HIP by default provides data protection, while this is a non-goal for Shim6.

The Shim6 working group was chartered to provide a solution to a specific problem, multihoming, which minimizes deployment disruption, while HIP is considered more of an experimental approach intended to solve several more general problems (mobility, multihoming and loss of end-to-end addressing transparency) through an explicit identifier/locator split. Communicating hosts that are willing and interested to run HIP (perhaps extended with Shim6's failure detection protocol) likely have no reason to also run Shim6. In this sense, HIP may be viewed as a possible long-term evolution or extension of the Shim6 architecture, or one possible implementation of the extended Shim6 design ESD [I-D.nordmark-shim6-esd].

7.6. Shim6 and Firewalls

The ability of Shim6 to divert the communication to different paths may be affected by certain firewall configurations. For example,

consider a deployment in which one of the peers of a Shim6 session is protected by a firewall (i.e. all the paths to the locators of that peer traverse the firewall). The firewall implements the Simple Security model [RFC4864], in which incoming packets are checked against a state resulting from outgoing traffic, either associated to the locator of the internal node ('endpoint independent filtering') or to both the locators of the internal and external nodes ('address dependent filtering' or 'address and port dependent filtering'). If the external node changes the locator associated to the internal node, the packet will be discarded by the firewall. In addition, if the firewall implements 'address dependent filtering' or 'address and port dependent filtering', any change by the external node in the locator used to identify itself will also result in the packet being discarded by the firewall.

This issue could be mitigated by making the firewalls aware of the different locators which could be associated to a given communication. If the firewall is implemented in the communication node itself, the firewall could inspect the Shim6 control packet exchange to obtain this information, or the Shim6 software module could explicitly inform the firewall software module. For firewalls located outside the node, the Shim6 control packet exchange can be used to associate the alternate locators to the communication state, although it may not work for topologies in which both directions for the communication do not traverse the firewall, or in which the firewall is not traversed after a locator change. The detail of any of such mechanisms is out of the scope of this document.

7.7. Shim6 and IPv6 NAT

Address translation techniques such as Network Prefix Translation [RFC6296] may be used until workable solutions to avoid renumbering or facilitate multihoming are developed [RFC5902]. We now consider the impact of IPv6 NATs in Shim6 operation.

The security provided by Shim6 against redirection attacks is built upon the cryptographic properties of CGA and HBA addresses. When a CGA address of a node is used as the local ULID, the locators configured in the node can be signed with the private key associated to the CGA. When a HBA address of a node is used as the local ULID, the HBA address securely chains the ULID and other locators of the node by means of a hash. Shim6 allows a node to securely convey the alternative locators to a peer node to which it wants to communicate to, or to which it is already communicating with.

To illustrate the problems introduced by the use of NAT, we will consider two nodes communicating, with one node, the internal node, located behind an IPv6 NAT.

Regarding to the start of the communication, we can distinguish two cases:

- o If the four-way handshake which establishes a Shim6 session traverses an IPv6 NAT, and no ULID-pair option is exchanged, the Shim6 context is associated to the source and destination addresses contained in the packets exchanged. Since no cryptographic check is performed on these addresses, the Shim6 context and the communication will be established. As long as these addresses are kept, communication can proceed. Peers can exchange locators, which can be successfully validated and stored in the locator list for future use, although these locators may not be usable, as we will discuss later.
- o If the four-way handshake which establishes a Shim6 session traverses an IPv6 NAT and an ULID-pair option is exchanged, the locators received would not correspond to the ULID-pair option exchanged the session establishment should fail.

Consider now that the communication has been successfully established, after an exchange in which no ULID-pair option was exchanged. There are two cases in which the IPv6 NAT would prevent the communication:

- o The internal node changes the locator used as source address for a packet, and the packet traverses an IPv6 NAT. In this case the NAT will translate the locator into an address belonging to its address pool. Suppose that this address is different to the address used to establish the communication. This address will be different to any locator sent by the internal node to the peer using Shim6. Therefore, the peer node will not be able to identify this packet as belonging to the communication.
- o The external node changes the locator used as destination address to send packets to the internal node. This locator is one of the locators sent by the internal node to the peer, so this locator may not be meaningful out of the domain being served by the NAT. Then this packet may not be routable, and would be discarded.
- o The external node changes the locator used as source address to send packets to the internal node. On one hand, if the mapping or filtering states of the NAT depends on both the locators of the internal and external nodes, the NAT would drop the packet. On the other hand, if the mapping and filtering states of the NAT depend only on the locator of the internal node, the packet would be delivered to the internal node.

Since Shim6 provide some means to recover from the selection of unusable locator pairs, it may recover from the situations in which locators are changed after a successfull Shim6 session initiation. However, the use of NATs reduces the ability of Shim6 of using all available paths.

Note that an Application Level Gateway designed to modify the Shim6 control packets would not be able to generate a valid signature, in case a CGA is being used, or a Parameter Data Structure binding the translated locator to the other locators of a node, in case a HBA is being used. Therefore, the same failure cases described before would remain.

8. Security Considerations

This section considers the applicability of the Shim6 protocol from a security perspective, i.e. which security features can expect applications and users of the Shim6 protocol.

First of all, it should be noted that the Shim6 protocol is not a security protocol, like for instance HIP. This means that as opposed to HIP, it is an explicit non-goal of the Shim6 protocol to provide enhanced security for the communications that use the Shim6 protocol. The goal of the Shim6 protocol design in terms of security is not to introduce new vulnerabilities that were not present in the current non-Shim6 enabled communications. In particular, it is an explicit non-goal of the Shim6 protocol security to provide protection from on-path attackers. On-path attackers are able to sniff and spoof packets in the current Internet, and they are able to do the same in Shim6 communications (as long as the communication flows through the path they are located on). So, summarizing, the Shim6 protocol does not provide data packet protection from on-path attackers.

However, the Shim6 protocol does use several security techniques. The goal of these security measures is to protect the Shim6 signaling protocol from new attacks resulting from the adoption of the Shim6 protocol. In particular, the use of HBA/CGA prevents on-path and off-path attackers injecting new locators into the locator set of a Shim6 context, thus preventing redirection attacks [RFC4218]. Moreover, the usage of probes before re-homing to a different locator as a destination address prevents flooding attacks from off-path attackers. Note that for nodes using CGA addresses, security depends on the secure handling of the private key associated to the signature and validation of locators. In particular, any address configuration method must assure that the private key remains secret, as discussed in Section 3.3.

In addition, the usage of a 4-way handshake for establishing the Shim6 context protects against DoS attacks, so hosts implementing the Shim6 protocol should not be more vulnerable to DoS attacks than regular IPv6 hosts.

Finally, many Shim6 signaling messages contain a Context Tag, meaning

that only attackers that know the Context Tag can forge them. As a consequence, only on-path attackers can generate false Shim6 signaling packets for an established context. The impact of these attacks would be limited since they would not be able to add additional locators to the locator set (because of the HBA/CGA protection). In general the possible attacks have similar effects to the ones that an on-path attacker can launch on any regular IPv6 communication. The residual threats are described in the Security Considerations of the Shim6 protocol specification [RFC5533].

8.1. Privacy Considerations

The Shim6 protocol is designed to provide some basic privacy features. In particular, HBAs are generated in such a way, that the different addresses assigned to a host cannot be trivially linked together as belonging to the same host, since there is nothing in common in the addresses themselves. Similar features are provided when the CGA protection is used. This means that it is not trivial to determine that a set of addresses is assigned to a single Shim6 host.

However, the Shim6 protocol does exchange the locator set in clear text and it also uses a fixed Context Tag when using different locators in a given context. This implies that an attacker observing the Shim6 context establishment exchange or seeing different payload packets exchanged through different locators, but with the same Context Tag, can determine the set of addresses assigned to a host. However, this requires that the attacker is located along the path and that it can capture the Shim6 signaling packets.

9. IANA Considerations

This document has no actions for IANA.

10. Contributors

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

11. 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.

Shinta Sugimoto reviewed this document and provided comments and text.

Iljitsch van Beijnum, Brian Carpenter, Sam Xia and Jari Arkko reviewed this document and provided comments.

12. References

12.1. Normative References

[RFC2827] Ferguson, P. and D. Senie, "Network Ingress Filtering: Defeating Denial of Service Attacks which employ IP Source Address Spoofing", BCP 38, RFC 2827, May 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.

[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.

[RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

[RFC4423] Moskowitz, R. and P. Nikander, "Host Identity Protocol (HIP) Architecture", RFC 4423, May 2006.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, September 2007.

[RFC4862] Thomson, S., Narten, T., and T. Jinmei, "IPv6 Stateless Address Autoconfiguration", RFC 4862, September 2007.

[RFC4960] Stewart, R., "Stream Control Transmission Protocol", RFC 4960, September 2007.

[RFC5201] Moskowitz, R., Nikander, P., Jokela, P., and T. Henderson, "Host Identity Protocol", RFC 5201, April 2008.

[RFC5533] Nordmark, E. and M. Bagnulo, "Shim6: Level 3 Multihoming Shim Protocol for IPv6", RFC 5533, June 2009.

[RFC5534] Arkko, J. and I. van Beijnum, "Failure Detection and Locator Pair Exploration Protocol for IPv6 Multihoming", RFC 5534, June 2009.

[RFC5535] Bagnulo, M., "Hash-Based Addresses (HBA)", RFC 5535, June 2009.

[RFC6146] Bagnulo, M., Matthews, P., and I. van Beijnum, "Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers", RFC 6146, April 2011.

[RFC6275] Perkins, C., Johnson, D., and J. Arkko, "Mobility Support in IPv6", RFC 6275, July 2011.

[RFC6316] Komu, M., Bagnulo, M., Slavov, K., and S. Sugimoto, "Sockets Application Program Interface (API) for Multihoming Shim", RFC 6316, July 2011.

12.2. Informative References

[I-D.ietf-csi-dhcpv6-cga-ps]
Jiang, S., Shen, S., and T. Chown, "DHCPv6 and CGA Interaction: Problem Statement", draft-ietf-csi-dhcpv6-cga-ps-07 (work in progress), May 2011.

[I-D.ietf-mif-problem-statement]
Blanchet, M. and P. Seite, "Multiple Interfaces and Provisioning Domains Problem Statement", draft-ietf-mif-problem-statement-15 (work in progress), May 2011.

[I-D.ietf-v6ops-ipv6-multihoming-without-ipv6nat]
Troan, O., Miles, D., Matsushima, S., Okimoto, T., and D. Wing, "IPv6 Multihoming without Network Address Translation", draft-ietf-v6ops-ipv6-multihoming-without-ipv6nat-03 (work in progress), November 2011.

[I-D.nordmark-shim6-esd]
Nordmark, E., "Extended Shim6 Design for ID/loc split and Traffic Engineering", draft-nordmark-shim6-esd-01 (work in progress), February 2008.

[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.

[RFC4191] Draves, R. and D. Thaler, "Default Router Preferences and More-Specific Routes", RFC 4191, November 2005.

[RFC4218] Nordmark, E. and T. Li, "Threats Relating to IPv6 Multihoming Solutions", RFC 4218, October 2005.

[RFC4864] Van de Velde, G., Hain, T., Droms, R., Carpenter, B., and E. Klein, "Local Network Protection for IPv6", RFC 4864, May 2007.

[RFC4984] Meyer, D., Zhang, L., and K. Fall, "Report from the IAB Workshop on Routing and Addressing", RFC 4984, September 2007.

[RFC5902] Thaler, D., Zhang, L., and G. Lebovitz, "IAB Thoughts on IPv6 Network Address Translation", RFC 5902, July 2010.

[RFC6296] Wasserman, M. and F. Baker, "IPv6-to-IPv6 Network Prefix Translation", RFC 6296, June 2011.

Authors' Addresses

Joe Abley
ICANN
4676 Admiralty Way, Suite 330
Marina del Rey, CA 90292
USA

Phone: +1 519 670 9327
Email: joe.abley@icann.org

Marcelo Bagnulo
U. 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/

Alberto Garcia-Martinez
U. Carlos III de Madrid
Av. Universidad 30
Leganes, Madrid 28911
Spain

Phone: +34 91 6248782
Email: alberto@it.uc3m.es
URI: http://www.it.uc3m.es/