Internet Engineering Task Force M. Xu
Internet-Draft Z. Ming
Intended status: Experimental Tsinghua University
Expires: March 21, 2011 J. Ubillos
Swedish Institute of Computer Science
C. Vogt
Ericsson
September 17, 2010

Name Based Sockets - Shim6
draft-xu-name-shim6-00

Abstract

This document describes and defines shim6 as a mobility solution for name-based sockets. Using names rather than pseudo IP addresses, shim6 can handle a more diverse set of mobility scenarios. These changes allow a shim6 session to persist even through cases where one node has no working locators to its correspondent node. If the name is also a resolvable fully qualified domain name, the connection can be kept alive even if neither node have a working locator to the corresponding node. As can be the case if both nodes are mobile simultaneously.

Table of Contents

1. Conventions 3
2. Terminology 3
3. Overview 3
3.1. Introduction 3
4. Mobility support 4
4.1. Shim6 4
4.1.1. Brief overview of changes 4
4.1.2. Identity change 5
4.1.3. The hand-shake with name exchange 5
4.1.4. Triggers of shim6 6
4.1.5. Establishing Shim6 context 6
4.1.6. Problems for Shim6 to support mobility 6
4.1.7. Changes to REAP 8
4.1.8. Changes to STATE Machine 8
5. Security Considerations 9
6. IANA Considerations 9
7. Contributors 9
8. References 9
8.1. Normative References 9
8.2. Informative References 9
8.3. URL References 9

1. Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Terminology

Locator - An IP address (v4 or v6) on which a host can be reached.

Multi-home - A host which is reachable through multiple locators (on one interface or more)

Name - A character string (max 255 chars long) on which an endpoint can be identified. A name maps to zero or more locators.

3. Overview

3.1. Introduction

Mobility can be treated as a special case of multihoming. Shim6 provides a promising way to implement multihoming for upper layer protocols, thus it is reasonable to use Shim6 to provide NBS with mobility functionality.

The traditional Shim6 defined in RFC5533 [RFC5533] does not aim to solve mobility problem, so changes need to be made to the existing Shim6 protocol. One of the reasons for not supporting mobility is that Shim6 uses a specific IP address as the identifier of the upper lay protocol. To avoid confusion, communication must be stopped when this IP address becomes unavailable. With the presence of name based socket, Shim6 can use the name as the upper layer identifier rather than IP. In such a way, Shim6 will not encounter the problem that connection must be terminated when the locator used as ULID becomes invalid.

To allow connections to survive during movement, Shim6 context should be preserved for a certain period of time when there are no available locator to use. When both nodes move simultaneously, Shim6 path exploration may fail due to lost update messages. To keep the connection from being terminated in this case, DNS is involved to provide address information.

Technical details will be covered in this document.

4. Mobility support

4.1. Shim6

"... The Shim6 protocol, a layer 3 shim for providing locator agility below the transport protocols, so that multihoming can be provided for IPv6 with failover and load-sharing properties, without assuming that a multihomed site will have a provider-independent IPv6 address prefix announced in the global IPv6 routing table. The hosts in a site that has multiple provider- allocated IPv6 address prefixes will use the Shim6 protocol specified in this document to set up state with peer hosts so that the state can later be used to failover to a different locator pair, should the original one stop working. " RFC5533 [RFC5533]

4.1.1. Brief overview of changes

To the upper layers, shim6 provides a stable IP address-like identifier (ULID) to identify the remote host and make the IP addresses (locators) transparent to the application. This way of providing a pseudo-address (ULID) does however invite confusion. The ULID selected by Shim6 is actually the IP address which is available for the application when the connection is being established. This address (ULID) may become invalid during the connection (RFC5533 Section 1.5 [RFC5533]). ULID invalidation is beyond the control of the individual hosts, it is controlled by the network. This might cause confusion if the applications continues to use the ULIDs which are no longer valid. Shim6s solution to this problem to terminate the communication immediately when ever any ULID becomes invalid. This is definitely inappropriate in a mobile scenario as connections are expected to be preserved during the mobile period. Moving between two distinct networks, changing your complete locator set is the common scenario (e.g. entirely switching from one WiFi-provider to another.)

Name Based Sockets suggest using the name of a host as the identifier. This solves the above problems, as a name is valid for as long as a host wishes it to be. Also, as Name Based Sockets provide a new explicit interface (names rather than 'pseudo IP addresses'), applications that use it will be aware of the available features, and may make correct assessments of the underlying IP stack and its enhancements.

This document describes a set of changes and improvements to shim6 that are to be incorporated with the Name Based Sockets.

Briefly, the changes are:

- o Name is used as ULID rather than an IP (Section 4.1.2).

- o Node inter-reachability resilience, for when both nodes are simultaneously mobile using DNS (Section 4.1.6.1).

4.1.2. Identity change

Shim6 selects a locator (IP) in the initial contact with the remote peer and uses this locator as an upper-layer identifier (ULID). To support NBS, we use name (or some structure related to name) as ULID instead of IP addresses.

Because the end point identifier is no longer a locator but rather a name, the initial name exchange is performed by NBS and Shim6 uses this name to construct the ULID.

Shim6 requires that any communication using a ULID MUST be terminated when the ULID becomes invalid. Using names as ULIDs instead of IP address is more in line with the transport semantic. Having names as ULIDs means that the session may still exist even if both communicating hosts locator lists are empty at a given point of time. This is particularly important when one or both peer(s) are moving.

Note that replacing a ULID with a name does not necessarily mean representing the ULID as a string or a string-like structure internally. In order to lower the complexity, one should make the least possible modification to both the Shim6 protocol (where ULID is a 128-bit IPv6 address) and its definition implementation, we propose to internally represent the name as a 128-bit MD5-hash and use this MD5-hash as the corresponding ULID.

4.1.3. The hand-shake with name exchange

As is described in RFC5533 [RFC5533], Shim6 does not need to react immediately when connections start up. The initial name exchange is performed by NBS and it requires no help from Shim6. The name exchanged by NBS will be further used as ULID by Shim6. At some time during the communication, some heuristic may determine that it is appropriate to use shim6 to support mobility/multi-homing, so the communicating hosts initiate a 4-way, context-establishment exchange. As a result, both hosts get a locator list of each other.

As an extension to Shim6, we do not change the operation sequence of the 4-way exchange, namely the order of I1, R1, I2, R2 will not be changed. What is changed is that the IP-based ULID is replaced by a name-based ULID and the hand shake no longer requires ULID negotiation because it has already been done by NBS.

4.1.4. Triggers of shim6

It is not necessarily worth paying the overhead of setting up a shim context when e.g. only a small number of packets are exchanged between two hosts. As a result, Shim6 functionality will not be started immediately as a new communication is initiated.

NBS uses some heuristic for determining when to perform a deferred context establishment. This heuristic might be that more than 50 packets have been sent or received, or that a timer expires while active packet exchange was in place RFC 5533 [RFC5533]. Exactly how the heuristic is designed is beyond the scope of this document.

4.1.5. Establishing Shim6 context

At a certain time during the connection, some heuristic on host A or B (or both) determine that it is appropriate to pay the Shim6 overhead to improve host-to-host communication. This makes the Shim6 initiate the 4-way, context-establishment exchange (defined in RFC 5533).

As a result, both A and B get a list of each others locators. In name-based Shim6, the ULID is represented as a MD5-hash of name rather than IP.

4.1.6. Problems for Shim6 to support mobility

When only one host moves to a new network, a REAP Update is triggered to prevent connection from being terminated. Under normal circumstances, connection will be smoothly preserved during the REAP Update process.

However, REAP itself is not sufficient to support full mobility functionality, as when both hosts move simultaneously, neither of them will receive the update message, which will lead to a connection loss. To deal with this problem, DNS should be involved to provide address information.

4.1.6.1. DNS querying

An effective solution for the mobility problem is to have a "stationary infrastructure" to provide address information for all mobile devices. We propose to use DNS as the stationary infrastructure as it associates addresses with names and has enough capability. How DNS incorporates with name-based Shim6 is described in the following part.

4.1.6.2. One peer moves

In the case that only one host moves, the moving host starts a REAP Update process to re-establish Shim6 context with the correspondent host. At the same time, DNS should be updated by the moving host. This procedure is the normal REAP [RFC5534] procedure with the added update to DNS.

The following sequence illustrates the details:

1. Two hosts, A and B are communicating using NBS and Shim6.

2. At certain moment, A moves to a new network and changes its IP address (locator).

3. A updates the authoritative DNS with its new IP address. In parallel, A starts the REAP Update process by sending B an Update Request and the REAP Update process is invoked.

4. New operational locator pair is found by REAP Update process.

5. Handover process is completed and connection is preserved during the process.

4.1.6.3. Both peers move

When both hosts moves simultaneously, neither host will receive the REAP Update Request, thus REAP will fail in finding the new operational locator pair. Under such circumstances, both hosts need to query DNS for the correspondent hosts addresses. When new address is retrieved, both hosts initiate REAP Update process as specified in RFC5534 [RFC5534].

The following sequence illustrates the details:

1. Two hosts, A and B are communicating using NBS and Shim6.

2. At certain moment, both A and B move simultaneously and both hosts change their respective IP addresses.

3. Both hosts update DNS with their new addresses and send REAP Update Request to their correspondent peer.

4. Due to concurrent move, Update Requests are lost for both directions.

5. Both hosts experience an Update Timeout and query DNS for correspondent hosts' locators using their names.

6. New addresses are returned by the respective DNS queries and REAP Update is now able to operate. A and B re-invoke a REAP Update process using the new addresses.

7. New operational locator pair is found by REAP Update process.

8. Handover process is completed and connection is preserved during the handover process.

4.1.7. Changes to REAP

We extend REAP by adding DNS Querying into its Path Exploration. For the sake of backwards compatibility, DNS can be implemented as a separate module and has no impact on the other part of REAP which is specified in RFC 5534. The DNS functionality can be turned off for stationary hosts and be turned on for mobile devices.

In the mobile scenario, DNS Query and REAP Path Exploration may work together to provide stronger reliability. DNS query might be lost due to link failure or timeout due to high network delay. Under such circumstances, REAP Path Exploration will be triggered because of a SEND TIMEOUT and tries to find an available path. This is meaningful when a host has multiple available interfaces (for instance Wi-Fi and 3G) and the address change for one interface does not lead to the change for others.

4.1.8. Changes to STATE Machine

In order to implement mobility, we propose to add a new state ?C NO_LOC into the shim6 state machine. With this state, Shim6 does not directly transmit to a dead state when there is no available address, but transmits to NO_LOC state to wait for a new locator. There is a minor difference between the Exploring state and the InBoundOK state. For simplicity, in this draft we use the Exploring state to represent both Exploring state and InBoundOK state.

Upon realizing that there is no available address to use, Shim6 sets the state to NO_LOC, starts a time and wait for an available address. If the host successfully attaches to a new network, Shim6 performs the following operations:

1. Update the context.

2. Send Update Request to the remote peer.

3. Transitions into the Exploring state.

4. Stops the timer.

If the timer expires, Shim6 removes the context.

Upon entering the Exploring state, Shim6 waits for the Update ACK from the peer. If the update ACK is received, Shim6 can perform the normal operation as defined by RFC 5533 and RFC 5534 to continue the connection. Otherwise, Shim6 infers that the peer may also move and asks the DNS for the peers locator, insert the locator into the locator list of the peer (if not already present in the list) and starts the Reap Path Exploration process. As the DNS stores the latest address of the peer, we believe this locator is most likely to be available, thus in the Reap Path Exploration process, we first try the locator returned by DNS and then other locators (if any).

5. Security Considerations

6. IANA Considerations

7. Contributors

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

[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.

8.3. URL References

Authors' Addresses

Mingwei Xu
Tsinghua University
FIT Building 4-104, Tsinghua University
Beijing
China

EMail: xmw@cernet.edu.cn

Zhongxing Ming
Tsinghua University
FIT Building 4-104, Tsinghua University
Beijing
China

EMail: mingzx@126.com

Javier Ubillos
Swedish Institute of Computer Science
Kistagangen 16
Kista
Sweden

Phone: +46767647588
EMail: jav@sics.se

Christian Vogt
Ericsson
200 Holger Way
San Jose, CA 95134-1300
USA

EMail: christian.vogt@ericsson.com