Behavior Engineering for Hindrance Avoidance

Internet-Draft

Intended status: Informational Expires: September 3, 2012

R. Penno Juniper Networks T. Saxena Cisco Systems M. Boucadair France Telecom S. Sivakumar Cisco Systems March 2, 2012

Analysis of Stateful 64 Translation draft-ietf-behave-64-analysis-07

Abstract

Due to specific problems, NAT-PT was deprecated by the IETF as a mechanism to perform IPv6-IPv4 translation. Since then, new efforts have been undertaken within IETF to standardize alternative mechanisms to perform IPv6-IPv4 translation. This document analyzes to what extent the new stateful translation mechanisms avoid the problems that caused the IETF to deprecate NAT-PT.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of \underline{BCP} 78 and \underline{BCP} 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 3, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to $\underline{\mathsf{BCP}}$ 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

| ${	extstyle 1}$. Introduction | . <u>3</u> |
|---|------------|
| <u>1.1</u> . Definition | <u>3</u> |
| <u>1.2</u> . Context | <u>3</u> |
| <u>1.3</u> . Scope | <u>4</u> |
| 2. Analysis of 64 Translation Against Concerns of RFC4966 | . <u>4</u> |
| 2.1. Problems Impossible to Solve | <u>5</u> |
| 2.2. Problems Which Can be Solved | . <u>5</u> |
| 2.3. Problems Solved | . <u>7</u> |
| $\underline{3}$. Conclusions | <u> 10</u> |
| 4. IANA Considerations | <u>11</u> |
| 5. Security Considerations | . 12 |
| 6. Acknowledgements | |
| 7. References | . 12 |
| <u>7.1</u> . Normative References | . 12 |
| 7.2. Informative References | <u>13</u> |
| Authors' Addresses | . 15 |

1. Introduction

1.1. Definition

This document uses stateful 64 (or 64 for short) to refer to the mechanisms defined in the following documents:

- o IP/ICMP Translation Algorithm [RFC6145]
- o Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers [RFC6146]
- o DNS64: DNS extensions for Network Address Translation from IPv6 Clients to IPv4 Servers [RFC6147]
- o IPv6 Addressing of IPv4/IPv6 Translators [RFC6052]
- o Framework for IPv4/IPv6 Translation [RFC6144]

1.2. Context

Stateful 64 is widely seen as a major interconnection technique designed to enable communications between IPv6-only and IPv4-only networks. One of the building blocks of the stateful 64 is decoupling the DNS functionality from the protocol translation itself.

This approach is pragmatic in the sense that there is no dependency on DNS implementation for the successful NAT handling. As long as there is a function (e.g., DNS64 [RFC6147] or other means) that can construct an IPv6-embedded IPv4 address with a pre-configured IPv6 prefix, an IPv4 address and a suffix (refer to [RFC6052]), NAT64 will work just fine.

The focus of the stateful 64 is on the deployment and not the implementation details. As long as a NAT64 implementation conforms to the expected behavior, as desired in the deployment scenario, the details are not very important as mentioned in this excerpt from [RFC6146]:

"A NAT64 MAY perform the steps in a different order, or MAY perform different steps, but the externally visible outcome MUST be the same as the one described in this document."

1.3. Scope

This document provides an analysis of how the proposed set of documents that specify stateful IPv6-only to IPv4-only translation and replace NAT-PT [RFC2766] address the issues raised in [RFC4966].

As a reminder, it is worth mentioning the analysis is limited in the sense that hosts from IPv6 networks can initiate a communication to IPv4 network/Internet, but not vice-versa. This corresponds to Scenario 1 and Scenario 5 described in [RFC6144]. Hence, the scenario of servers moving to IPv6 while clients remaining IPv4 remains unaddressed. Of course, IPv6 to IPv4 communications can also be supported if static or explicit bindings (e.g., [I-D.ietf-pcp-base]) are configured on the stateful NAT64.

Stateful 64, just like any other technique under development, has some positives and some drawbacks. The ups and downs of the proposal must be clearly understood while going forward with its future development.

The scope of this document does not include stateless translation.

2. Analysis of 64 Translation Against Concerns of RFC4966

Of the set of problems pointed out in [RFC4966], the stateful 64 addresses some of them, whereas leaves others unaddressed.

Some issues mentioned in [RFC4966] were solved by [RFC4787], [RFC5382] and [RFC5508]. At the time when NAT-PT was published these recommendations were not in place but they are orthogonal to the translation algorithm per se, therefore they could be implemented with NAT-PT. On the other hand, NAT64 [RFC6146] explicitly mentions that these recommendations need to be followed and thus should be seen as a complete specification.

It is also worth pointing out that the scope of the stateful 64 is reduced when compared to NAT-PT. Following is a point by point analysis of the problems. This document classifies the issues listed in [RFC4966] into three categories:

- 1. Problems impossible to solve.
- 2. Problems which can be solved.
- 3. Problems solved.

Penno, et al. Expires September 3, 2012 [Page 4]

2.1. Problems Impossible to Solve

Problems discussed in [RFC4966], which are impossible to solve:

 Inability to redirect traffic for protocols that lack demultiplexing capabilities or are not built on top of specific transport-layer protocols for transport address translations (Section 2.2 of [RFC4966]).

Analysis: This issue is not specific to 64 but to all NAT-based solutions.

 Loss of information due to incompatible semantics between IPv4 and IPv6 versions of headers and protocols (<u>Section 2.4 of</u> [RFC4966]).

Analysis: This issue is not specific to 64 but due to the design of IPv4 and IPv6.

3. Need for the NAT64-capable device to act as proxy for correspondent node when IPv6 node is mobile, with consequent restrictions on mobility (Section 2.7 of [RFC4966]).

Analysis: This is not specific to NAT64 but to all NAT flavors. Refer to [I-D.haddad-mext-nat64-mobility-harmful] for an early analysis on mobility complications encountered when NAT64 is involved.

2.2. Problems Which Can be Solved

Problems discussed in [RFC4966], which can be solved:

 Disruption of all protocols that embed IP addresses (and/or ports) in packet payloads or apply integrity mechanisms using IP addresses (and ports) (Section 2.1 of [RFC4966]).

Analysis: In the case of FTP [RFC0959] this problem can be mitigated in several ways (e.g., use a FTP64 ALG [RFC6384] or in the FTP client (e.g., [I-D.ietf-ftpext2-ftp64])).

In the case of SIP [RFC3261], no specific issue is induced by 64; the same techniques for NAT traversal can be used when a NAT64 is involved in the path (e.g., ICE [RFC5245], maintain SIP-related NAT bindings as per Section 3.4 of [RFC5853], media latching [I-D.ietf-mmusic-media-path-middleboxes], embedded SIP ALGs, etc.). [RFC6157] provides more discussion on how to establish SIP sessions between IPv4 and IPv6 SIP user agents.

The functioning of other protocols is left for future study. Note that the traversal of NAT64 by application embedding IP address literal is not specific to NAT64 but generic to all NAT-based solutions.

2. Interaction with SCTP [RFC4960] and multihoming (Section 2.6 of [RFC4966]).

Analysis: Only TCP and UDP transport protocols are within the scope of NAT64 [RFC6146]. SCTP is out of scope of this document.

3. Inability to handle multicast traffic (Section 2.8 of [RFC4966]).

Analysis: This problem is not addressed by the current 64 specifications.

 Scalability concerns together with introduction of a single point of failure and a security attack nexus (<u>Section 3.2 of</u> [RFC4966]).

Analysis: This is not specific to NAT64 but to all stateful NAT flavors. The presence of single point of failures is deployment-specific; some service providers may deploy state synchronization means while others may only rely on a distributed NAT64 model.

5. Restricted validity of translated DNS records: a translated record may be forwarded to an application that cannot use it (Section 4.2 of [RFC4966]).

Analysis: If a node on the IPv4 side forwards the address of the other endpoint to a node which cannot reach the NAT box or is not covered under the endpoint-independent constraint of NAT, then the new node will not be able to initiate a successful session.

Actually, this is not a limitation of 64 (or even NAT-PT) but a deployment context where IPv4 addresses managed by the NAT64 are not globally reachable. The same limitation can be encountered when referrals (even without any NAT in the path) include reachability information with limited reachability scope (See [I-D.carpenter-behave-referral-object] for more discussion about issues related to reachability scope).

6. IPsec traffic using AH (Authentication Header) [RFC4302] in both transport and tunnel modes cannot be carried through NAT-PT without terminating the security associations on the NAT-PT, due

to the inclusion of IP header fields in the scope of AH's cryptographic integrity protection [RFC3715] (Section 2.1 of [RFC4966]). In addition, IPsec traffic using ESP (Encapsulating Security Payload) [RFC4303] in transport mode generally uses UDP encapsulation [RFC3948] for NAT traversal (including NAT-PT traversal) in order to avoid the problems described in [RFC3715] (Section 2.1 of [RFC4966]).

Analysis: This is not specific to NAT64 but to all NAT flavors.

7. Address selection issues when either the internal or external hosts implement both IPv4 and IPv6 (Section 4.1 of [RFC4966]).

Analysis: This is out of scope of 64 since Scenarios 1 and 5 of $[\mbox{RFC6144}]$ assume IPv6-only hosts.

Therefore this issue is not resolved and mitigation techniques outside the 64 need to be used (e.g., [I-D.ietf-6man-addr-select-opt]). These techniques may allow to offload NAT64 resources and prefer native communications which do not involve address family translation. Avoiding NAT devices in the path is encouraged for mobile nodes in order to save power consumption due to keepalive messages which are required to maintain NAT states ("always-on" services). An in-depth discussion can be found in [I-D.wing-behave-dns64-config].

2.3. Problems Solved

Problems, identified in [RFC4966], which are solved:

 Constraints on network topology (as it relates to DNS-ALG; see <u>Section 3.1 of [RFC4966]</u>).

Analysis: The severity of this issue has been mitigated by the separation of the DNS from the NAT functionality.

Nevertheless, a minimal coordination may be required to ensure that the NAT64 to be crossed (the one to which the IPv4-Converted IPv6 address returned to a requesting host) must be in the path and has also sufficient resources to handle received traffic.

 Need for additional state and/or packet reconstruction in dealing with packet fragmentation. Otherwise, implement no support for fragments. (Section 2.5 of [RFC4966])

Analysis: This issue is not specific to 64 but to all NAT-based solutions. [RFC6146] specifies how to handle fragmentation; appropriate recommendations to avoid fragmentation-related DoS attacks are proposed (e.g., limit resources to be dedicated to out of order fragments).

3. Inappropriate translation of responses to A queries from IPv6 nodes (Section 4.3 of [RFC4966]).

Analysis: DNS64 [RFC6147] does not alter A queries.

4. Address selection issues and resource consumption in a DNS-ALG with multi-addressed nodes (Section 4.4 of [RFC4966]).

Analysis: Since no DNS-ALG is required to be co-located with NAT64, there is no need to maintain temporary states in anticipation of connections. Note that explicit bindings (See Section 3 of [I-D.ietf-pcp-base]) are required to allow for communications initiated from an IPv4-only client to an IPv6-only server.

5. Limitations on DNS security capabilities when using a DNS-ALG (Section 2.5 of [RFC4966]).

Analysis: A DNSSEC validating stub resolver behind a DNS64 in server mode is not supported. Therefore if a host wants to do its own DNSSEC validation, and it wants to use a NAT64, the host has to also perform its own DNS64 synthesis. Refer to Section 3 of[RFC6147] for more details.

6. Creation of a Denial of Service (DoS) threat relating to exhaustion of memory and address/port pool resources on the translator (Section 3.4 of [RFC4966]).

Analysis: This specific DoS concern on Page 6 of [RFC4966] is under a DNS-ALG heading in that document, and refers to NAT-PT's creation of NAT mapping state when a DNS query occurred. With the new IPv6-IPv4 translation mechanisms, DNS queries do not create any mapping state in the NAT64.

To mitigate the exhaustion of port pool issue (<u>Section 3.4 of [RFC4966]</u>), 64 must enforce a port limit similar to the one defined in [<u>I-D.ietf-behave-lsn-requirements</u>].

Thus, this concern can be fully eliminated in 64.

7. Requirement for applications to use keepalive mechanisms to workaround connectivity issues caused by premature timeout for session table and BIB entries (Section 2.3 of [RFC4966]).

Analysis: Since NAT64 follows some of the [RFC4787], [RFC5382] and [RFC5508] requirements, there is a high lower bound for the lifetime of sessions. In NAT-PT this was unknown and applications needed to assume the worst case. For instance, in NAT64, the lifetime for a TCP session is approximately 2 hours, so not much keep-alive signaling overhead is needed.

Application clients (e.g., VPN clients) are not aware of the timer configured in the NAT device. For unmanaged services, a conservative approach would be adopted by applications which issue frequent keepalive messages to be sure that an active mapping is still be maintained by any involved NAT64 device even if the NAT64 complies with [RFC4787][RFC5382][RFC5508].

Note that keepalive messages may be issued by applications to ensure that an active entry is maintained by a firewall, with or without a NAT in the path, which is located in the boundaries of a local domain.

8. Lack of address mapping persistence: Some applications require address retention between sessions. The user traffic will be disrupted if a different mapping is used. The use of the DNS-ALG to create address mappings with limited lifetimes means that applications must start using the address shortly after the mapping is created, as well as keep it alive once they start using it. (Section 3.3 of [RFC4966])

Analysis: In the following, address persistence is used to refer to the support of "IP address pooling" behavior of "Paired" [RFC4787].

In the context of 64, the external IPv4 address (representing the IPv6 host in the IPv4 network) is assigned by the NAT64 machinery and not the DNS64 function. Address persistence can be therefore easily ensured by the NAT64 function (which complies with NAT recommendations [RFC4787][RFC5382]). Address persistence should be guaranteed for both dynamic and static bindings.

In the IPv6 side of the NAT64, the same IPv6 address is used to represent an IPv4 host; no issue about address persistence is raised in IPv6 network.

3. Conclusions

The above analysis of the solutions provided by the stateful 64 shows that the majority of the problems that are not directly related to the decoupling of NAT and DNS remain unaddressed. Some of these problems are not specific to 64 but are generic to all NAT-based solutions.

This points to several shortcomings of stateful 64 which must be addressed if the future network deployments have to move reliably towards 64 as a solution to IPv6-IPv4 interconnection.

Some of the issues, as pointed out in [RFC4966], have possible solutions. However these solutions will require significant updates to the stateful 64, increasing its complexity.

The following table summarizes the conclusions based on the analysis of stateful 64.

| + | + | | | | ++ |
|---|--------------------------|-----------------------|------------------|----------------|-----------------------------|
| Issue | NAT-PT Specific | Exists in NAT64 | DNS ALG | Generic NAT | Can be solved? |
| Protocols embedding addresses | No | Yes | No | Yes | Yes |
| Protocols without demux capability | No | Yes | No | Yes | No |
| Binding state decay | No | Yes | No | Yes | Yes |
| Loss of information | No | Yes | No | No | No |
| Fragmentation | No | No No | No | Yes | Yes |
| SCTP and Multihoming interaction | No | Yes | No | Yes | Yes |
| Proxy correspondent node for MIPv6 | No | Yes | No | No | No |

Penno, et al. Expires September 3, 2012 [Page 10]

| Multicast | No | Yes | No | Yes | Yes |
|---|----------------------|-----|---------------------|----------|-------------------------|
| Topology constraints with DNS-ALG | Yes | No | Yes | No | Yes Yes |
| Scale and Single point of failure | No | Yes | No | Yes | Yes Yes |
| Lack of address persistence | No | Yes | No | Yes | Yes |
| DoS attacks | No | Yes | No | Yes | Yes |
| Address selection issues with Dual stack hosts | Yes | No | Yes | No | Yes |
| Non-global validity of Translated RR records | Yes | No | Yes | No | Yes |
| Incorrect translation of A responses | Yes Yes | No | Yes | No | Yes Yes |
| DNS-ALG and Multi- addressed nodes | No | Yes | No | Yes | Yes |
| DNSSEC limitations | No | Yes | No | Yes | Yes Yes |

Table 1: Summary of NAT64 analysis

4. IANA Considerations

This document makes no request of IANA.

Penno, et al. Expires September 3, 2012 [Page 11]

Note to RFC Editor: this section may be removed on publication as an $\ensuremath{\mathsf{RFC}}$.

5. Security Considerations

This document does not specify any new protocol or architecture. It only analyzes how BEHAVE WG 64 documents mitigate concerns raised in [RFC4966] and which ones are still unaddressed.

6. Acknowledgements

Many thanks to M. Bagnulo, D. Wing, X. Li, D. Anipko and S. Moonesamy for their review and comments.

D. Black provided the IPsec text.

7. References

7.1. Normative References

- [RFC0959] Postel, J. and J. Reynolds, "File Transfer Protocol", STD 9, RFC 959, October 1985.
- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston,
 A., Peterson, J., Sparks, R., Handley, M., and E.
 Schooler, "SIP: Session Initiation Protocol", RFC 3261,
 June 2002.
- [RFC3948] Huttunen, A., Swander, B., Volpe, V., DiBurro, L., and M. Stenberg, "UDP Encapsulation of IPsec ESP Packets", RFC 3948, January 2005.
- [RFC4302] Kent, S., "IP Authentication Header", <u>RFC 4302</u>, December 2005.
- [RFC4303] Kent, S., "IP Encapsulating Security Payload (ESP)", RFC 4303, December 2005.
- [RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", <u>BCP 127</u>, <u>RFC 4787</u>, January 2007.
- [RFC4960] Stewart, R., "Stream Control Transmission Protocol", RFC 4960, September 2007.

- [RFC4966] Aoun, C. and E. Davies, "Reasons to Move the Network Address Translator Protocol Translator (NAT-PT) to Historic Status", RFC 4966, July 2007.
- [RFC5382] Guha, S., Biswas, K., Ford, B., Sivakumar, S., and P. Srisuresh, "NAT Behavioral Requirements for TCP", <u>BCP 142</u>, <u>RFC 5382</u>, October 2008.
- [RFC5508] Srisuresh, P., Ford, B., Sivakumar, S., and S. Guha, "NAT Behavioral Requirements for ICMP", <u>BCP 148</u>, <u>RFC 5508</u>, April 2009.
- [RFC6052] Bao, C., Huitema, C., Bagnulo, M., Boucadair, M., and X. Li, "IPv6 Addressing of IPv4/IPv6 Translators", RFC 6052, October 2010.
- [RFC6144] Baker, F., Li, X., Bao, C., and K. Yin, "Framework for IPv4/IPv6 Translation", RFC 6144, April 2011.
- [RFC6145] Li, X., Bao, C., and F. Baker, "IP/ICMP Translation Algorithm", RFC 6145, April 2011.
- [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.
- [RFC6147] Bagnulo, M., Sullivan, A., Matthews, P., and I. van Beijnum, "DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers", RFC 6147, April 2011.

7.2. Informative References

- [I-D.carpenter-behave-referral-object]
 - Carpenter, B., Boucadair, M., Halpern, J., Jiang, S., and K. Moore, "A Generic Referral Object for Internet Entities", draft-carpenter-behave-referral-object-01 (work in progress), October 2009.

draft-ietf-6man-addr-select-opt-03 (work in progress),
February 2012.

[I-D.ietf-behave-lsn-requirements]

Perreault, S., Yamagata, I., Miyakawa, S., Nakagawa, A., and H. Ashida, "Common requirements for Carrier Grade NATs (CGNs)", draft-ietf-behave-lsn-requirements-05 (work in progress), November 2011.

[I-D.ietf-ftpext2-ftp64]

Liu, D., Beijnum, I., and Z. Cao, "FTP consideration for IPv4/IPv6 transition", draft-ietf-ftpext2-ftp64-02 (work in progress), January 2012.

[I-D.ietf-mmusic-media-path-middleboxes]

Stucker, B., Tschofenig, H., and G. Salgueiro, "Analysis of Middlebox Interactions for Signaling Protocol Communication along the Media Path", draft-ietf-mmusic-media-path-middleboxes-04 (work in progress), January 2012.

[I-D.ietf-pcp-base]

Wing, D., Cheshire, S., Boucadair, M., Penno, R., and P. Selkirk, "Port Control Protocol (PCP)", draft-ietf-pcp-base-23 (work in progress), February 2012.

[I-D.wing-behave-dns64-config]

Wing, D., "IPv6-only and Dual Stack Hosts on the Same Network with DNS64", <u>draft-wing-behave-dns64-config-03</u> (work in progress), February 2011.

- [RFC2766] Tsirtsis, G. and P. Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT)", <u>RFC 2766</u>, February 2000.
- [RFC3715] Aboba, B. and W. Dixon, "IPsec-Network Address Translation (NAT) Compatibility Requirements", <u>RFC 3715</u>, March 2004.
- [RFC5853] Hautakorpi, J., Camarillo, G., Penfield, R., Hawrylyshen,
 A., and M. Bhatia, "Requirements from Session Initiation
 Protocol (SIP) Session Border Control (SBC) Deployments",
 RFC 5853, April 2010.

[RFC6384] van Beijnum, I., "An FTP Application Layer Gateway (ALG) for IPv6-to-IPv4 Translation", RFC 6384, October 2011.

Authors' Addresses

Reinaldo Penno Juniper Networks 1194 N Mathilda Avenue Sunnyvale, California 94089 USA

Email: rpenno@juniper.net

Tarun Saxena Cisco Systems

Email: tasaxena@cisco.com

Mohamed Boucadair France Telecom Rennes 35000 France

Email: mohamed.boucadair@orange.com

Senthil Sivakumar Cisco Systems 7100-8 Kit Creek Road Research Triangle Park, North Carolina 27709 USA

Email: ssenthil@cisco.com