Network Working Group
Internet-Draft
Expires: April 16, 2004

S. Satapati
S. Sivakumar
Cisco Systems, Inc.
P. Barany
No Affiliation
S. Okazaki
NTT Multimedia Communications Labs
H. Wang
Nokia
October 17, 2003

NAT-PT Applicability
draft-satapati-v6ops-natpt-applicability-00

Status of this Memo

Copyright Notice

Abstract

This document discusses the applicability RFC 2766, Network Address Translation - Protocol Translation (NAT-PT) employing the Stateless IP/ICMP Translation (SIIT) algorithm, as an IPv4-to-IPv6 transition and co-existence mechanism. It highlights the NAT-PT/SIIT operational principles and the network context for which NAT-PT was designed. Known limitations of NAT-PT/SIIT are presented. Applicable scenarios

along with guidelines for deployment are being offered.

Table of Contents

1. Introduction

NAT-PT [1], or Network Address translation Protocol translation, is the standard mechanism that allows communication between IPv6-only and IPv4-only nodes. RFC 2766 also defines NAPT-PT, a variant of NAT-PT that translates transport identifiers (TCP/UDP port numbers and ICMP identifiers) in addition to IP header, allows multiple V6 nodes to communicate with V4 nodes using a single public IPv4 address.

SIIT [2] specifies an algorithm for translation between IPv4 and IPv6 packet headers (including ICMP headers) in separate translator boxes in the network without requiring any per-connection state in those boxes. An IPv6-only node acquires a temporary IPv4 address (known as an IPv4-translated address) when communicating with an IPv4-only node. IPv6 nodes, that use SIIT, would need additional logic in the network layer so that an application inserts IPv4 type literals in the payload if the destination is an IPv4-mapped IPv6 address. SIIT does not specify a mechanism for the IPv6-only node to acquire a temporary IPv4 address, nor does it specify a mechanism for providing routing (perhaps using tunneling) to and from the temporary IPv4 address assigned to the IPv6-only node.

NAT-PT uses SIIT algorithm for protocol independent translation of IPv4 and IPv6 datagrams. NAT-PT describes the dynamic address assignment of IPv4 address to IPv6 nodes and vice-versa, as sessions are initiated across IPv4/IPv6 boundaries. The IPv4 addresses are assigned from an IPv4 address pool and are used to transparently replace the original IPv6 addresses used by the IPv6-only nodes. Consequently, the IPv6-only nodes need not be changed in any way. However, NAT-PT must track sessions (i.e., state must be maintained) and the inbound and outband packets pertaining to a session must traverse the same NAT-PT device.

RFC 2766 specifies DNS-ALG for address discovery to support bidirectional session initiation. RFC 2766 also specifies FTP-ALG to provide application level transparency between IPv4 and IPv6. For applications (like FTP) that carry IP addresses and/or transport layer port numbers in their application-level data, additional Application Layer Gateways (ALGs) would need to be deployed along with NAT-PT.

This document discusses the applicability of Network Address Translation - Protocol Translation as an IPv4-to-IPv6 transition and co-existence mechanism. It highlights the NAT-PT/SIIT operational principles and the network context for which NAT-PT was designed. Known limitations of NAT-PT/SIIT are presented. Applicable scenarios along with guidelines for deployment are being offered.

Sections 2 and 3 talk about limitations of SIIT and NAT-PT (including DNS-ALG) in its current form. Section 4 presents the applicable scenarios. Section 5 summarizes the applicability and proposes some recommendations when deploying NAT-PT.

2. SIIT Limitations

In this section, we present the limitatons that are specific to SIIT. It is important to analyze SIIT because it is an integral part of NAT-PT.

2.1 Overloading the Use of IPv4-mapped addresses

IPv4-mapped IPv6 addresses are of the form ::ffff:a.b.c.d, and are used to refer to IPv4-only nodes. An IPv6 packet, going through SIIT from an IPv6-only node, will have IPv4-mapped address as destination in the IPv6 header. Problems associated with usage of IPv4-mapped address have been discussed in an expired draft [3].

Most notably, as described in RFC 3493 [8], the IPv4-mapped address representation is used in the basic API to denote IPv4 addresses using the AF_INET6 socket. However, at the same time, the IPv4-mapped address is also used by SIIT to denote IPv6 communication using AF_INET6 sockets. Consequently, a security threat exists due to the ambiguous dual use of IPv4-mapped addresses.

2.2 Translation Semantics Difficult to Use

Some of the translation semantics defined by SIIT are difficult to implement/interpret. The ICMPv4 "Parameter Problem" (type 12 code 0) and the ICMPv6 "Parameter Problem" (Type 4 Code 0, 1, 2) error messages are classic examples. Some ICMP error messages do not have IPv6 counterpart and hence there were no semantics defined. A particular problem being reported in the IPv4-side may go unrecognized from the IPv6 perspective.

In order to translate from ICMPv6 to ICMPv4, if the ICMPv6 code field is set to 1 (unrecognizable Next Header Type encountered), then the ICMPv6 message is translated into an ICMPv4 Destination Unreachable Message (Type 3 code 2) Error Message. There is no loss of information in this case. However, if the ICMPv6 code field is set to 0 (erronenous header field encountered) or 2 (unrecognized IPv6 option encountered), then the ICMPv6 message is translated into an ICMPv4 Parameter Problem (Type 12 Code 0) Error Message and the pointer needs to be updated to point to the corresponding field in the translated and included IP header. There is a loss of information in this case.

2.3 Multicast Mapping Impossible

IPv4 multicast addresses cannot be mapped to IPv6 multicast addresses (e.g., ::ffff:224.1.2.3 is an IPv4-mapped address with a Class D IPv4 address, but it is not an IPv6 multicast address). This limitation makes it impossible for SIIT to support multicast between IPv6 and IPv4 networks.

2.4 SCTP and Multihoming

SCTP [9] supports multihoming. If an IPv6-only node sets up SCTP connections to an IPv4-only node with one or more SIITs between them, there could be a problem. If more than one IPv4-translated address is used, these addresses are transported to the IPv4-only node during association setup in the payloads of the INIT and INIT-ACK chunks. SIIT, as specified in [1], cannot translate IP addresses included in INIT and INIT-ACK chunks. Even if it could, there would be a problem with SCTP/IP packets being carried through the translator using ESP in Transport-mode.

Also, there is ongoing work which allows the modification of the IP address once an SCTP association has been established (for non-multihoming purposes). The modification messages have IP addresses in the SCTP packet [14].

2.5 IP Security (IPsec)

2.5.1 IPsec Encapsulating Security Protocol (ESP)

ESP provides both confidentiality and integrity for the packet that it is protecting [10]. ESP in Transport-mode encrypts the transport layer header, the data, and the ESP trailer (optional Padding, Pad Length, and Next Header). ESP in Transport-mode authenticates the ESP header (SPI, Sequence Number, and IV) in addition to the transport layer header, the data, and the ESP trailer.

SIIT does not modify any headers above the logical IP layer (IP headers, IPv6 fragment headers). Therefore, TCP/UDP packets encrypted using ESP in Transport-mode can pass through, assuming that key management (manual or IKE) can operate somehow between an IPv6-only node and IPv4-only node. The reason being the notation of IPv4-translatable addresses, which is ::ffff:0:a.b.c.d, was chosen to avoid any changes to the transport layer protocol's pseudo-header checksum. Also, SCTP calculates a CRC-32c checksum only on the SCTP packet (SCTP common header and one or more control or DATA chunks). Therefore, as long as SCTP control chunks like INIT and INIT-ACK do not contain IP addresses, SCTP packets encrypted using ESP in Transport-mode can pass through SIIT.

ICMP messages, on the other hand, cannot be carried through SIIT for a number of reasons:

- the ICMP checksum field must be updated as part of the translation since ICMPv6, unlike ICMPv4, has a pseudo-header checksum like TCP or UDP
- ICMP packets need to have the Type value translated
- for ICMP error messages, the included IP header also needs translation

ESP in Tunnel-mode encrypts the IP header of the encapsulated IP packet in addition to the transport layer header, the data, and the ESP trailer of the encapsulated IP packet. ESP in Tunnel-mode authenticates the ESP header in addition to the encapsulated IP packet and the transport layer header, the data, and the ESP trailer of the encapsulated IP packet. Consequently, TCP/IP packets, UDP/IP packets, SCTP/IP packets, and ICMP packets cannot be carried through the SIIT translator using ESP in Tunnel-mode.

2.5.2 IPsec Authentication Header (AH)

AH provides integrity for the packet that it is protecting [11]. AH is explicitly designed to detect alterations to IP packet headers. AH in Transport-mode authenticates the immutable fields of the IP header (setting the mutable fields to zero prior to calculating the Integrity Check Value (ICV)), the AH header (Next Header, Payload Length, Reserved, SPI, and Sequence Number), and the data. Consequently, TCP/IP packets, UDP/IP packets, SCTP/IP packets, and ICMP packets, or any such packets cannot be carried through the SIIT translator using AH in Transport-mode.

AH in Tunnel-mode authenticates the immutable fields of the outer IP header (setting the mutable fields to zero prior to calculating the ICV), the AH header, the encapsulated IP header, the transport layer header, and the data. Consequently, TCP/IP packets, UDP/IP packets, SCTP/IP packets, and ICMP packets, or any such packets cannot be carried through the SIIT translator using AH in Tunnel-mode.

2.5.3 Key Management

Note that this entire discussion begs the question that key management can operate between the IPv6-only node and the IPv4-only node. While this may not be an issue for IPsec implementations that are integrated with the host's Operating System (OS), there could be substantial challenges that need to be overcome with either a Bump-In-the-Stack (BIS) or Bump-In-The-Wire (BITW) IPsec

implementation [12].

For example, the ISAKMP Identification payload [13] that is used in IKE Phase 1 Main Mode/Aggressive Mode (pre-shared secret, digital signatures with certificates, public key encryption, and revised public key encryption) may contain the IP address of the host as an identifier (e.g., ID_IPV6_ADDR, ID_IPV4_ADDR). The same holds true for IKE Phase 2 Quick Mode and the optional usage of the Client Identification payload. While the use of other identifiers such as ID_FQDN and ID_USER_FQDN are possible, there are usage scenarios that cannot be accommodated in this manner (e.g., SPD entries describing subnets).

3. NAT-PT Limitations

The adverse effects of NAT [4] have been studied to a great extent in the past [7]. While NAT-PT has exactly the same implications as NAT, the DNS-ALG as specified in RFC 2766 introduces additional failure modes. Most applications may fail for exactly the same reasons as those cited for IPv4 NAT. All SIIT limitations mentioned above hold true for NAT-PT, except those due to usage of IPv4-mapped IPv6 addresses. NAT-PT proposes to use a /96 prefix, which need not be a IPv4-mapped IPv6 address type. The limitations of DNS-ALG, as an inherent address discovery mechanism, are also discussed in this section.

3.1 Application deployment

Applications that embed literals, such as IP addresses and/or TCP/UDP port numbers, will break across NAT-PT in the absence of a supporting ALG. ALGs are required to setup bindings during signalling, which would subsequently be used for data (or media) traffic. Each application requires its own specific ALG to be deployed that should work in tandem with header-translation functionality. It is hard to cope up with this situation, as new applications become popular or existing applications define new extensions or message types.

3.2 Scalability/Single-point-of-failure

A single device implementing NAT-PT (and ALGs) can easily become a bottleneck, when traffic from a large IPv6 network is forced to go through a single device. The device builds and maintains binding state as traffic flows traverse. Therefore all subsequent traffic of the flow must pass through it, turning the device into a single point of failure.

3.3 IP Security

3.3.1 IPsec ESP/AH

All of the IPsec ESP and AH limitations that apply to SIIT also apply to NAT-PT because NAT-PT(and NAPT-PT) make use of SIIT translation algorithm to translate IP headers and transport identifiers. However, unlike SIIT, NAT-PT need not use IPv4-translated addresses and IPv4-mapped addresses. Therefore, TCP/IP, UDP/IP, and ICMP packets cannot be carried through NAT-PT in ESP Transport-mode because TCP, UDP, and ICMPv6 checksums have a dependency on the IP source and destination addresses via the pseudo-header. However, SCTP calculates a CRC-32c checksum only on the SCTP packet (SCTP common header and one or more control or DATA chunks). Therefore, as long as SCTP control chunks like INIT and INIT-ACK do not contain IP addresses, SCTP packets encrypted using ESP in Transport-mode can pass through NAT-PT.

3.3.2 Key Management

All of the key management limitations that apply to SIIT also apply to NAT-PT because NAT-PT (and NAPT-PT) make use of SIIT translation algorithm to translate IP headers and transport identifiers. Also, since NAT-PT maintains state (unlike SIIT), even if IP addresses are not used in the ISAKMP Identification payload during IKE Phase 1 Main Mode/Aggressive Mode and IKE Phase 2 Quick Mode, there is still difficulty with retaining the IPv4-to-IPv6 address mapping on NAT-PT from the time that IKE completes negotiation to the time that IPsec uses the key on an application.

3.4 DNS-ALG

Address discovery is the mechanism by which an IPv4-only or an IPv6-only node discovers the "translated address" to which the packets should be sent. The NAT-PT specification proposes to solve address discovery, for applications that use DNS, by using a DNS-ALG, as specified in section 4 of RFC-2766.

Address discovery through DNS-ALG is broken when dual-stack nodes attempt to talk to either IPv6-only or IPv4-only nodes. The result here is the possibility of traffic flow taking a translated path as opposed to native. Additionally an IPv4 node that sends an A query, through DNS-ALG, may receive a AAAA reply.

Since DNS-ALG modifies queries (from A to AAAA) and replies (IPv6 address to IPv4 address literals), DNSSEC cannot be deployed in the presence of DNS-ALG.

3.5 Denial of Service

The DoS attacks being discussed here are mostly due to address

spoofing.

3.5.1 Address Pool Depletion Attacks

Suppose that the attacker resides in the same IPv6 stub domain as NAT-PT, and is sending packets to an IPv4-only destination. If the IPv6 attacker sends multiple packets with different source address (that are topologically inside the stub domain), then the pool of IPv4 addresses managed by NAT-PT will quickly exhaust resulting in a Denial of Service to legitimate IPv6-only nodes.

3.5.2 Reflection Attacks

There is another address spoofing attack, wherein the attacker forges the IPv6 source address to be a broadcast or multicast or the source address of an existing node. The reply IPv4 packets that get translated by NAT-PT will result in an attack.

4. Applicability

4.1 SIIT

SIIT assumes v4 address assignment to IPv6 nodes, and routing to that assigned v4 address. These aspects are actually requirements for SIIT deployment. Additionally, IPv6 nodes need some modification as mentioned in Section 1. Any host modifications to support a certain transition mechanism are strongly discouraged. Applications cannot interoperate between pure IPv6-only and IPv4-only nodes using SIIT, unless there exists an ALG accompanying SIIT.

If SIIT were to de deployed for IPv6-only nodes (i.e. without the modifications proposed by SIIT), it is mandatory to have supporting ALGs for applications that need interoperability. The device implementing SIIT must maintain binding state somehow. This combination is not very different from deployment model of NAT-PT. Hence applicable scenarios would be the same as the ones explained below.

The SIIT algorithm may be useful in header translation if some specific-purpose translators are specified which can live with the shortcomings of SIIT.

4.2 NAT-PT

It is theoretically possible to deploy NAT-PT at the border of an IPv6-only stub network, either translating specific IPv6-only services to IPv4 nodes or by translating IPv4-only services to IPv6 nodes. The latter comes in two flavors: translation by the party

providing such an IPv4-only service, or someone servicing them, or by the party which is planning to deploy IPv6-only infrastructure. There are some very specific cases, where one may be able to consider its use.

4.2.1 Legacy IPv4-only nodes in an IPv6 network

This is the scenario where the entire network infrastructure is IPv6 and the majority of nodes attached to the network are IPv6-only. But there are some existing (legacy) IPv4-only hosts that cannot be upgraded (or abandoned) and that need connectivity to the neighboring IPv6 nodes.

It is unrealistic to continue deploying/supporting dual-stack nodes network-wide to support a small set of legacy IPv4-only hosts. Ideally, legacy nodes must be retired as quickly as possible. Proxy/relay mechanism, such as TRT[5], could be used to enable basic applications that use TCP/UDP. Since host upgrade is ruled out, NAT-PT can be used when traffic to/from these legacy hosts is minimal. NAT-PT must be considered only when proxies cannot be deployed or deemed unfit for the specific application being enabled.

While this scenario may be acceptable within an organization/domain, the same may not be true between independent domains. The consequences of NAT-PT, as discussed above, must be kept in mind during deployment.

4.2.2 Special Nodes/Networks

This scenario applies to situations where constraints arise that are either imposed by the network or by the end nodes. These constraints do not apply to a general purpose IP network or generic TCP/IP hosts connected to a general purpose IP network. Few special cases are listed below:

- devices are IPv6-only due to memory and code size constraints
- devices are built to run a specific well-defined set of applications
- devices could be dual-stack; but resource consumption in the network should be kept at a minimum

3GPP networks is an example of the latter, where IMS is IPv6-only by design and PDP context is supposedly a scarce resource.

4.2.2.1 Restricted use in 3GPP Networks

An important requirement for 3GPP networks is that 3GPP hosts, running SIP-based IMS applications over IPv6, must be able to communicate with IPv4 SIP hosts on the Internet [6]. To minimize the number of active PDP Contexts in the mobile network, it should be possible for a 3GPP host to access both IMS applications and other non-IMS applications (non-SIP-based) over a single IPv6 connection (IPv6 PDP Context in 3GPP).

NA(P)T-PT may be used for translating traffic pertaining to specific (non-IMS) applications, for which dual-stack application proxies are not suitable. NA(P)T-PT may be used for header translation of IMS media traffic, provided the binding is acquired through different means. In other words, the device implementing header translation must not implement ALGs. The mechanism of acquiring the binding from an external entity is beyond the scope of this document.

IMS media translation may prove to be a burden and inefficient if a single device is implementing header translation. Distributed approaches/techniques may be considered to offload the translation responsibility to multiple devices, which requires exchange of binding state between devices. The exact operation of such an approach is beyond the scope of this document.

4.2.2.2 3GPP2 Networks

For further study.

4.2.3 IPv6-only Networks

The deployment of IPv6-only networks in general is not recommended, and therefore this scenario is only tentatively described in Appendix A.

5. Summary

This document discourages NAT-PT (RFC 2766) as a general purpose transition mechanism, except for the scenarios mentioned above. The limitations cited in this document must be observed during deployment. Only if the shortcomings of NAT-PT are acceptable in a particular deployment, may the use of NAT-PT be considered as a short-term solution. In no case should NAT-PT be viewed as a generic solution for IPv6/IPv4 transition in an IPv6-only network.

It is to be noted that DNS-ALG has severe failure modes when processing AAAA queries. However, DNS-ALG could still be used for IPv4-initiated connections.

For special networks, targeted translation using NAT-PT is acceptable

for short term. Migrating entirely to IPv6 is considered beneficial in the long run.

RFC 2766 specifies NAT-PT using the SIIT algorithm for header conversion, DNS-ALG for address discovery, and FTP-ALG as one application level gateway mechanism. The terminology in RFC 2766 can easily be misinterpreted, in that NAT-PT mandates ALGs. RFC 2766 does not mandate DNS-ALG to be implemented along with header translation mechanism.

6. Security Considerations

This memo discusses the limitations and the applicability of NAT-PT. One important consideration in this analysis is the security related to packet translation. As shown, IPsec AH and ESP generally do not work through the NAT-PT translators. Similarly, the NAT-PT DNS-ALG and similar algorithms cause bottlenecks and a concern for availability if deployed. The generic translation mechanisms seem to be causing a number of problems. Therefore, mass deployment of translators is not recommended; some use may be possible in very specific cases identified in this memo.

7. Acknowledgements

The authors gratefully acknowledge the contributions of Pekka Savola and Rob Austein.

Normative References

[1] Tsirtsis and Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT)", RFC 2766, February 2000.

[2] Nordmark, "Stateless IP/ICMP Translator (SIIT)", RFC 2765, February 2000.

[3] Metz and Hagino, "IPv4-Mapped Addresses on the Wire Considered Harmful", draft-itojun-v6ops-v4mapped-harmful-01 Expired, October 2002.

[4] Srisuresh and Egevang, "The IP Network Address Translator", RFC 3022, January 2001.

[5] Hagino and Yamamoto, "An IPv6-to-IPv4 Transport Relay Translator", RFC 3142, June 2001.

[6] Soininen et al., "Transition Scenarios for 3GPP Networks", RFC 3574, August 2003.

Informative References

[7] Hain, "Architectural Implications of NAT", RFC 2993, November 2000.

[8] Gilligan, R. et al., "Basic Socket Interface Extensions for IPv6", RFC 3493, February 2003.

[9] Stewart, R. et al., "Stream Control Transmission Protocol", RFC 2960, October 2000.

[10] Kent and Atkinson, "IP Encapsulating Security Payload", RFC 2406, November 1998.

[11] Kent and Atkinson, "IP Authentication Header", RFC 2402, November 1998.

[12] Kent and Atkinson, "Security Architecture for the Internet Protocol", RFC 2401, November 1998.

[13] Maughan, D. et al., "Internet Security Association and Key Management Protocol", RFC 2408, November 1998.

[14] Stewart, R. et al., "Stream Control Transmission Protocol (SCTP) Dynamic Address Reconfiguration", draft-ietf-tsvwg-addip-sctp-08 Work in progress, September 2003.

Authors' Addresses

Suresh Satapati
Cisco Systems, Inc.
170 West Tasman Dr.
San Jose, CA 95134
USA

EMail: satapati@cisco.com

Senthil Sivakumar
Cisco Systems, Inc.
170 West Tasman Dr.
San Jose, CA 95134
USA

EMail: ssenthil@cisco.com

Peter Barany
No Affiliation

EMail: Baranyp9@aol.com

Satomi Okazaki
NTT Multimedia Communications Labs
250 Cambridge Avenue
Palo Alto, CA 94306
USA

EMail: satomi@nttmcl.com

Hao H. Wang
Nokia
5th Floor, House 1, No.11, Hepingli Dongjie
Beijing
China

EMail: hao.h.wang@nokia.com

Appendix A. IPv6-only Networks

To roll out a new network, the options are IP4-only, IPv6-only or dual-stack. The most important requirement would be connectivity to surrounding internal networks or an external network such as the Internet. This connectivity has to be IPv4, due to existing deployed base. Therefore the options are:

- If the new network is deployed as IPv4-only, there needs to be a NAT at the edge of the network due to private addressing.
- If the network is dual-stack, there is still a need for NAT again due to private addressing.
- If the network is deployed as IPv6-only, then there is a need for mechanism such as NAT-PT.

Some are of the opinion that dual-stack network is most preferred manner of introducing IPv6, until such time when there are no IPv4-only nodes. Several questions arise like: what are the costs involved in running a dual-stack network; will IPv4-only implementations cease to appear in the marketplace; will the true potential of IPv6 be realized on a dual-stack network; how do protocols interact in a dual-stack network; are there known problems/ limitations/failure modes; do we have enough experience running a

dual-stack network etc.

Intellectual Property Statement

Full Copyright Statement

Acknowledgment

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