Network Working Group
Request for Comments: 4966
Obsoletes: 2766
Category: Informational

C. Aoun
Energize Urnet
E. Davies
Folly Consulting

July 2007

Reasons to Move the Network Address Translator - Protocol Translator (NAT-PT) to Historic Status

Status of This Memo

Copyright Notice

Abstract

This document discusses issues with the specific form of IPv6-IPv4 protocol translation mechanism implemented by the Network Address Translator - Protocol Translator (NAT-PT) defined in RFC 2766. These issues are sufficiently serious that recommending RFC 2766 as a general purpose transition mechanism is no longer desirable, and this document recommends that the IETF should reclassify RFC 2766 from Proposed Standard to Historic status.

Table of Contents

1. Introduction

The Network Address Translator - Protocol Translator (NAT-PT) document [RFC2766] defines a set of network-layer translation mechanisms designed to allow nodes that only support IPv4 to communicate with nodes that only support IPv6, during the transition to the use of IPv6 in the Internet.

[RFC2766] specifies the basic NAT-PT, in which only addresses are translated, and the Network Address Port Translator - Protocol Translator (NAPT-PT), which also translates transport identifiers, allowing for greater economy of scarce IPv4 addresses. Protocol translation is performed using the Stateless IP/ICMP Translation Algorithm (SIIT) defined in [RFC2765]. In the following discussion, where the term "NAT-PT" is used unqualified, the discussion applies to both basic NAT-PT and NAPT-PT. "Basic NAT-PT" will be used if points apply to the basic address-only translator.

A number of previous documents have raised issues with NAT-PT. This document will summarize these issues, note several other issues carried over from traditional IPv4 NATs, and identify some additional issues that have not been discussed elsewhere. Proposed solutions to the issues are mentioned and any resulting need for changes to the specification is identified.

Whereas NAT is seen as an ongoing capability that is needed to work around the limited availability of globally unique IPv4 addresses, NAT-PT has a different status as a transition mechanism for IPv6. As such, NAT-PT should not be allowed to constrain the development of IPv6 applications or impose limitations on future developments of IPv6.

This document draws the conclusion that the technical and operational difficulties resulting from these issues, especially the possible future constraints on the development of IPv6 networks (see Section 5), make it undesirable to recommend NAT-PT as described in [RFC2766] as a general purpose transition mechanism for intercommunication between IPv6 networks and IPv4 networks.

Although the [RFC2766] form of packet translation is not generally applicable, it is likely that in some circumstances a node that can only support IPv4 will need to communicate with a node that can only support IPv6; this needs a translation mechanism of some kind. Although this may be better carried out by an application-level proxy or transport-layer translator, there may still be scenarios in which a revised, possibly restricted version of NAT-PT can be a suitable solution; accordingly, this document recommends that the IETF should reclassify RFC 2766 from Proposed Standard to Historic status to

avoid it from being used in inappropriate scenarios while any replacement is developed.

The following documents relating directly to NAT-PT have been reviewed while drafting this document:

- Network Address Translation - Protocol Translation (NAT-PT) [RFC2766]
- Stateless IP/ICMP Translation Algorithm (SIIT) [RFC2765]
- NAT-PT Applicability Statement [NATP-APP]
- Issues with NAT-PT DNS ALG (Application Layer Gateway) in RFC 2766 [DNS-ALG-ISSUES]
- NAT-PT DNS ALG Solutions [DNS-ALG-SOL]
- NAT-PT Security Considerations [NATPT-SEC]
- Issues when Translating between IPv4 and IPv6 [TRANS-ISSUES]
- IPv6-IPv4 Translation Mechanism for SIP-Based Services in Third Generation Partnership Project (3GPP) Networks [3GPP-TRANS]
- Analysis on IPv6 Transition in 3GPP Networks [RFC4215]
- Considerations for Mobile IP Support in NAT-PT [NATPT-MOB]
- An IPv6-IPv4 Multicast Translator based on Internet Group Management Protocol / Multicast Listener Discovery (IGMP/MLD) Proxying (mtp) [MTP]
- An IPv4-IPv6 Multicast Gateway [MCASTGW]
- Scalable mNAT-PT Solution [MUL-NATPT]

Because the majority of the documents containing discussions of the issues are documents that are unlikely to become RFCs, the issues are summarized here to avoid the need for normative references.

Some additional issues can be inferred from corresponding issues known to exist in 'traditional' IPv4 NATs. The following documents are relevant:

- Protocol Complications with the IP Network Address Translator [RFC3027]

- IP Network Address Translator (NAT) Terminology and Considerations [RFC2663]

There is some ambiguity in [RFC2766] about whether the Application Layer Gateway (ALG) for DNS (referred to as DNS-ALG in this document) is an integral and mandatory part of the specification. The ambiguity arises mainly from the first section of the applicability section (Section 8), which appears to imply that 'simple' use of NAT-PT could avoid the use of the DNS-ALG.

This is important because a number of the major issues arise from the interactions between DNS and NAT-PT. However, detailed inspection of [RFC2766] shows that the 'simple' case has not been worked out and it is unclear how information about the address translation could be passed to the hosts in the absence of the DNS-ALG. Therefore, this document assumes that the DNS-ALG is an integral part of NAT-PT; accordingly, issues with the DNS-ALG must be considered as issues for the whole specification.

Note that issues not specifically related to the use of the DNS-ALG will apply to any network-layer translation scheme, including any based on the SIIT algorithm [RFC2765]. In the event that new forms of a translator are developed as alternatives to NAT-PT, the generic issues relevant to all IPv6-IPv4 translators should be borne in mind.

Issues raised with IPv6-IPv4 translators in general and NAT-PT in particular can be categorized as follows:

- Issues that are independent of the use of a DNS-ALG and are, therefore, applicable to any form of an IPv6-IPv4 translator:

  * Disruption of all protocols that embed IP addresses (and/or ports) in packet payloads or apply integrity mechanisms using IP addresses (and ports).

  * Inability to redirect traffic for protocols that lack demultiplexing capabilities or are not built on top of specific transport-layer protocols in situations where one NAPT-PT is translating for multiple IPv6 hosts.

  * Requirement for applications to use keepalive mechanisms to workaround connectivity issues caused by premature NAT-PT state timeout.

  * Loss of information due to incompatible semantics between IPv4 and IPv6 versions of headers and protocols.

  * Need for additional state and/or packet reconstruction in NAPT-PT translators dealing with packet fragmentation.

  * Interaction with SCTP and multihoming.

  * Need for NAT-PT to act as proxy for correspondent node when IPv6 node is mobile, with consequent restrictions on mobility.

  * NAT-PT not being able to handle multicast traffic.

- Issues that are exacerbated by the use of a DNS-ALG and are, therefore, also applicable to any form of an IPv6-IPv4 translator:

  * Constraints on network topology.

  * Scalability concerns together with introduction of a single point of failure and a security attack nexus.

  * 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.

  * Creation of a DoS (Denial of Service) threat relating to exhaustion of memory and address/port pool resources on the translator.

- Issues that result from the use of a DNS-ALG and are, therefore, specific to NAT-PT as defined in [RFC2766]:

  * Address selection issues when either the internal or external hosts implement both IPv4 and IPv6.

  * Restricted validity of translated DNS records: a translated record may be forwarded to an application that cannot use it.

  * Inappropriate translation of responses to A queries from IPv6 nodes.

* Address selection issues and resource consumption in a DNS-ALG with multi-addressed nodes.

* Limitations on DNS security capabilities when using a DNS-ALG.

Section 2, Section 3 and Section 4 discuss these groups of issues. Section 5 examines the consequences of deploying NAT-PT for application developers and the long term effects of NAT-PT (or any form of generally deployed IPv6-IPv4 translator) on the further development of IPv6.

The terminology used in this document is defined in [RFC2663], [RFC2766], and [RFC3314].

2. Issues Unrelated to an DNS-ALG

2.1. Issues with Protocols Embedding IP Addresses

It is well known from work on IPv4 NATs (see Section 8 of [RFC2663] and [RFC3027]) that the large class of protocols that embed numeric IP addresses in their payloads either cannot work through NATs or require specific ALGs as helpers to translate the payloads in line with the address and port translations. The same set of protocols cannot pass through NAT-PT. The problem is exacerbated because the IPv6 and IPv4 addresses are of different lengths, so that packet lengths as well as packet contents are altered. [RFC2766] describes the consequences as part of the description of the FTP ALG. Similar workarounds are needed for all protocols with embedded IP addresses that run over TCP transports.

The issues raised in Sections 2 and 3 of [RFC2663], relating to the authentication and encryption with NAT, are also applicable to NAT-PT.

Implementing a suite of ALGs requires that NAT-PT equipment includes the logic for each of the relevant protocols. Most of these protocols are continuously evolving, requiring continual and coordinated updates of the ALGs to keep them in step.

Assuming that the NAT-PT contains a colocated ALG for one of the relevant protocols, the ALG could replace the embedded IP addresses and ports. However, this replacement can only happen if no cryptographic integrity mechanism is used and the protocol messages are sent in the clear (i.e., not encrypted).

A possible workaround relies on the NAT-PT being party to the security association used to provide authentication and/or encryption. NAT-PT would then be aware of the cryptographic

algorithms and keys used to secure the traffic. It could then modify and re-secure the packets; this would certainly complicate network operations and provide additional points of security vulnerability.

Unless UDP encapsulation is used for IPsec [RFC3498], traffic using IPsec AH (Authentication Header), in transport and tunnel mode, and IPsec ESP (Encapsulating Security Payload), in transport mode, is unable to be carried through NAT-PT without terminating the security associations on the NAT-PT, due to their usage of cryptographic integrity protection.

A related issue with DNS security is discussed in Section 4.5.

2.2. NAPT-PT Redirection Issues

Section 4.2 of [RFC3027] discusses problems specific to RSVP and NATs, one of which is actually a more generic problem for all port translators. When several end-hosts are using a single NAPT-PT box, protocols that do not have a demultiplexing capability similar to transport-layer port numbers may be unable to work through NAPT-PT (and any other port translator) because there is nothing for NAPT-PT to use to identify the correct binding.

This type of issue affects IPsec encrypted packets where the transport port is not visible (although it might be possible to use the Security Parameter Index (SPI) as an alternative demultiplexer), and protocols, such as RSVP, which are carried directly in IP datagrams rather than using a standard transport-layer protocol such as TCP or UDP. In the case of RSVP, packets going from the IPv4 domain to the IPv6 domain do not necessarily carry a suitable demultiplexing field, because the port fields in the flow identifier and traffic specifications are optional.

Several ad hoc workarounds could be used to solve the demultiplexing issues, however in most cases these solutions are not documented anywhere, which could lead to non-deterministic and undesirable behavior (for example, such workarounds often assume particular network topologies, etc., in order to function correctly; if the assumptions are not met in a deployment, the workaround may not work as expected).

This issue is closely related to the fragmentation issue described in Section 2.5.

2.3. NAT-PT Binding State Decay

NAT-PT will generally use dynamically created bindings to reduce the need for IPv4 addresses both for basic NAT-PT and NAPT-PT. Both

basic NAT-PT and NAPT-PT use soft state mechanisms to manage the address and, in the case of NAPT-PT, port pools are used for dynamically created address bindings. This allows all types of NAT-PT boxes to operate autonomously without requiring clients to signal, either implicitly or explicitly, that a binding is no longer required. In any case, without soft state timeouts, network and application unreliability would inevitably lead to leaks, eventually causing address or port pool exhaustion.

For a dynamic binding to persist for longer than the soft state timeout, packets must be sent periodically from one side of the NAT-PT to the other (the direction is not specified by the NAT-PT specification). If no packets are sent in the proper direction, the NAT-PT binding will not be refreshed and the application connection will be broken. Hence, all applications need to maintain their NAT-PT bindings during long idle periods by incorporating a keepalive mechanism, which may not be possible for legacy systems.

Also, [RFC2766] does not specify how to choose timeouts for bindings. As discussed in [RFC2663] for traditional NATs, selecting suitable values is a matter of heuristics, and coordinating with application expectations may be impossible.

2.4. Loss of Information through Incompatible Semantics

NAT-PT reuses the SIIT header and protocol translations defined in [RFC2765]. Mismatches in semantics between IPv4 and IPv6 versions can lead to loss of information when packets are translated. Three issues arising from this are:

- There is no equivalent in IPv4 for the flow label field of the IPv6 header. Hence, any special treatment of packets based on flow label patterns cannot be propagated into the IPv4 domain.

- IPv6 extension headers provide flexibility for future improvements in the IP protocol suite and new headers that do not have equivalents in IPv4 may be defined. In practice, some existing extensions such as routing headers and mobility extensions are not translatable.

- As described in Section 2.2 of [NATP-APP], there are no equivalents in IPv6 for some ICMP(v4) messages, while for others (notably the 'Parameter Problem' messages) the semantics are not equivalent. Translation of such messages may lead to the loss of information. However, this issue may not be very severe because the error messages relate to packets that have been translated by NAT-PT rather than by arbitrary packets. If the NAT-PT is

functioning correctly, there is, for example, no reason why IPv6 packets with unusual extension headers or options should be generated.

Loss of information in any of these cases could be a constraint to certain applications.

A related matter concerns the propagation of the Differentiated Services Code Point (DSCP). NAT-PT and SIIT simply copy the DSCP field when translating packets. Accordingly, the IPv4 and IPv6 domains must have equivalent Per-Hop Behaviors for the same code point, or alternative means must be in place to translate the DSCP between domains.

2.5. NAT-PT and Fragmentation

As mentioned in [RFC3027], simple port translators are unable to translate packet fragments, other than the first, from a fragmented packet, because subsequent fragments do not contain the port number information.

This means that, in general, fragmentation cannot be allowed for any traffic that traverses a NAPT-PT. One attempted workaround requires the NAPT-PT to maintain state information derived from the first fragment until all fragments of the packet have transited the NAPT-PT. This is not a complete solution because fragment misordering could lead to the first fragment appearing at the NAPT-PT after later fragments. Consequently, the NAPT-PT would not have the information needed to translate the fragments received before the first.

Although it would not be expected in normal operation, NAPT-PT needs to be proofed against receiving short first fragments that don’t contain the transport port numbers. Note that such packets are a problem for many forms of stateful packet inspection applied to IPv6 packets. The current specifications of IPv6 do not mandate (1) any minimum packet size beyond the need to carry the unfragmentable part (which doesn’t include the transport port numbers) or (2) reassembly rules to minimize the effects of overlapping fragments. Thus, IPv6 is open to the sort of attacks described in [RFC1858] and [RFC3128].

An additional concern arises when a fragmented IPv4 UDP packet, which does not have a transport-layer checksum, traverses any type of NAT-PT box. As described in [RFC2766], the NAT-PT has to reconstruct the whole packet so that it can calculate the checksum needed for the translated IPv6 packet. This can result in a significant delay to the packet, especially if it has to be re-fragmented before transmission on the IPv6 side.

If NAT-PT boxes reassembled all incoming fragmented packets (both from the IPv4 and IPv6 directions) in the same way they have to for unchecksummed IPv4 UDP packets, this would be a solution to the first problem. The resource cost would be considerable apart from the potential delay problem if the outgoing packet has to be re-fragmented. In any case, fragmentation would mean that the NAT-PT would consume extra memory and CPU resources, making the NAT-PT even less scalable (see Section 3.2).

Packet reassembly in a NAT-PT box also opens up the possibility of various fragment-related security attacks. Some of these are analogous to attacks identified for IPv4. Of particular concern is a DoS attack based on sending large numbers of small fragments without a terminating last fragment, which would potentially overload the reconstruction buffers and consume large amounts of CPU resources.

2.6. NAT-PT Interaction with SCTP and Multihoming

The Stream Control Transmission Protocol (SCTP) [RFC2960] is a transport protocol, which has been standardized since SIIT was specified. SIIT does not explicitly cover the translation of SCTP, but SCTP uses transport port numbers in the same way that UDP and TCP do, so similar techniques can be used when translating SCTP packets.

However, SCTP also supports multihoming. During connection setup, SCTP control packets carry embedded addresses that would have to be translated. This would also require that the types of the options fields in the SCTP control packets be changed with consequent changes to packet length; the transport checksum would also have to be recalculated. The ramifications of multihoming as it might interact with NAT-PT have not been fully explored. Because of the 'chunked' nature of data transfer, it does not appear that that state would have to be maintained to relate packets transmitted using the different IP addresses associated with the connection.

Even if these technical issues can be overcome, using SCTP in a NAT-PT environment may effectively nullify the multihoming advantages of SCTP if all the connections run through the same NAT-PT. The consequences of running a multihomed network with separate NAT-PT boxes associated with each of the 'homes' have not been fully explored, but one issue that will arise is described in Section 4.4. SCTP will need an associated "ALG" -- actually a Transport Layer Gateway -- to handle the packet payload modifications. If it turns out that that state is required, the state would have to be distributed and synchronized across several NAT-PT boxes in a multihomed environment.

SCTP running through NAT-PT in a multihomed environment is also incompatible with IPsec as described in Section 2.1.

2.7. NAT-PT as a Proxy Correspondent Node for MIPv6

As discussed in [NATPT-MOB], it is not possible to propagate Mobile IPv6 (MIPv6) control messages into the IPv4 domain. According to the IPv6 Node Requirements [RFC4294], IPv6 nodes should normally be prepared to support the route optimization mechanisms needed in a correspondent node. If communications from an IPv6 mobile node are traversing a NAT-PT, the destination IPv4 node will certainly not be able to support the correspondent node features needed for route optimization.

This can be resolved in two ways:

- The NAT-PT can discard messages and headers relating to changes of care-of addresses, including reverse routing checks. Communications with the mobile node will continue through the home agent without route optimization. This is clearly sub-optimal, but communication should remain possible.

- Additional functionality could be implemented in the NAT-PT to allow it to function as a proxy correspondent node for all IPv4 nodes for which it has bindings. This scheme adds considerably to the complexity of NAT-PT. Depending on the routability of the IPv6 PREFIX used for translated IPv4 addresses, it may also limit the extent of mobility of the mobile node: all communications to the IPv4 destination have to go through the same NAT-PT, even if the mobile node moves to a network that does not have direct IPv6 connectivity with the NAT-PT.

In both cases, the existing NAT-PT specification would need to be extended to deal with IPv6 mobile nodes, and neither is a fully satisfactory solution.

2.8. NAT-PT and Multicast

SIIT [RFC2765] cannot handle the translation of multicast packets and NAT-PT does not discuss a way to map multicast addresses between IPv4 and IPv6. Some separate work has been done to provide an alternative mechanism to handle multicast. This work uses a separate gateway that understands some or all of the relevant multicast control and routing protocols in each domain. It has not yet been carried through into standards.

A basic mechanism, which involves only IGMP on the IPv4 side and MLD on the IPv6 side, is described in 'An IPv6-IPv4 Multicast Translator based on IGMP/MLD Proxying (mtp)' [MTP]. A more comprehensive approach, which includes proxying of the multicast routing protocols, is described in 'An IPv4 - IPv6 multicast gateway' [MCASTGW]. Both approaches have several of the issues described in this section, notably issues with embedded addresses.

[NATPT-SEC] identifies the possibility of a multiplicative reflection attack if the NAT-PT can be spoofed into creating a binding for a multicast address. This attack would be very hard to mount because routers should not forward packets with multicast addresses in the source address field. However, it highlights the possibility that a naively implemented DNS-ALG could create such bindings from spoofed DNS responses since [RFC2766] does not mention the need for checks on the types of addresses in these responses.

The issues for NAT-PT and multicast reflect the fact that NAT-PT is at best a partial solution. Completing the translation solution to cater for multicast traffic is likely to carry a similar set of issues to the current unicast NAT-PT and may open up significant additional security risks.

3. Issues Exacerbated by the Use of DNS-ALG

3.1. Network Topology Constraints Implied by NAT-PT

Traffic flow initiators in a NAT-PT environment are dependent on the DNS-ALG in the NAT-PT to provide the mapped address needed to communicate with the flow destination on the other side of the NAT-PT. Whether used for flows initiated in the IPv4 domain or the IPv6 domain, the NAT-PT has to be on the path taken by the DNS query sent by the flow initiator to the relevant DNS server; otherwise, the DNS query will not be modified and the response type will not be appropriate.

The implication is that the NAT-PT box also has to be the default IPv6 router for the site so that the DNS-ALG is able to examine all DNS requests made over IPv6. On sites with both IPv6 and dual-stack nodes, this will result in all traffic flowing through the NAT-PT with consequent scalability concerns.

These constraints are described in more detail in [DNS-ALG-ISSUES].

[DNS-ALG-SOL] proposes a solution for flows initiated from the IPv6 domain, but it appears that this solution still has issues.

For IPv6-only clients, the solution requires the use of a DNS server in the IPv4 domain, accessed via an IPv6 address which uses the NAT-PT PREFIX (see [RFC2766]). Queries to this server would necessarily pass through the NAT-PT. Dual-stack hosts would use a separate DNS server accessed through a normal IPv6 address. This removes the need for the NAT-PT box to be the default IPv6 gateway for the domain.

The primary proposal suggests that the IPv6-only clients should use this DNS server for all queries. This is expensive on NAT-PT resources because requests relating to hosts with native IPv6 addresses would also use the NAT-PT DNS-ALG.

The alternate suggestion to reduce this burden appears to be flawed: if IPv6-only clients are provided with a list of DNS servers including both the server accessed via NAT-PT and server(s) accessed natively via IPv6, the proposal suggests that the client could avoid using NAT-PT for hosts that have native IPv6 addresses.

Unfortunately, for the alternate suggestion, there is no a priori way in which the initiator can decide which DNS server to use for a particular query. In the event that the initiator makes the wrong choice, the DNS query will return an empty list rather than failing to respond. With standard DNS logic, the initiator will not try alternative DNS servers because it has received a response. This means that the solution would consist of always using DNS servers having the NAT-PT PREFIX. This imposes the burden of always requiring the DNS RR (Resource Record) [RFC1035] translation.

For flows initiated from the IPv4 network, the proposal recommends that the advertised DNS servers for the IPv6 network would have the IPv4 address of the NAT-PT. Again there is no deterministic way to choose the correct DNS server for each query resulting in the same issues as were raised for flows initiated from the IPv6 domain.

Although the engineering workaround, just described, provides a partial solution to the topology constraints issue, it mandates that DNS queries and responses should still go through a NAT-PT even if there would normally be no reason to do so. This mandatory passage through the NAT-PT for all DNS requests will exacerbate the other DNS-related issues discussed in Section 3.4 and Section 4.1.

3.2. Scalability and Single Point of Failure Concerns

As with traditional NAT, NAT-PT is a bottleneck in the network with significant scalability concerns. Furthermore, the anchoring of flows to a particular NAT-PT makes the NAT-PT a potential single

point of failure in the network. The addition of the DNS-ALG in NAT-PT further increases the scalability concerns.

Solutions to both problems have been envisaged using collections of cooperating NAT-PT boxes, but such solutions require coordination and state synchronization, which has not yet been standardized and again adds to the functional and operational complexity of NAT-PT. One such solution is described in [MUL-NATPT].

As with traditional NAT, the concentration of flows through NAT-PT and the legitimate modification of packets in the NAT-PT make NAT-PTs enticing targets for security attacks.

3.3. Issues with Lack of Address Persistence

Using the DNS-ALG to create address bindings requires that the translated address returned by the DNS query is used for communications before the NAT-PT binding state is timed out (see Section 2.3). Applications will not normally be aware of this constraint, which may be different from the existing lifetime of DNS query responses. This could lead to "difficult to diagnose" problems with applications.

Additionally, the DNS-ALG needs to determine the initial lifetime of bindings that it creates. As noted in Section 2.3, this may need to be determined heuristically. The DNS-ALG does not know which protocol the mapping is to be used for, and so needs another way to determine the initial lifetime. This could be tied to the DNS response lifetime, but that might open up additional DoS attack possibilities if very long binding lifetimes are allowed. Also, the lifetime should be adjusted once the NAT-PT determines which protocol is being used with the binding.

As with traditional NATs (see Section 2.5 of [RFC3027]), NAT-PT will most likely break applications that require address mapping to be retained across contiguous sessions. These applications require the IPv4 to IPv6 address mapping to be retained between sessions so the same mapped address may be reused for subsequent session interactions. NAT-PT cannot know this requirement and may reassign the previously used mapped address to different hosts between sessions.

Trying to keep NAT-PT from discarding an address mapping would require either a NAT-PT extension protocol that would allow the application to request the NAT-PT device to retain the mappings, or an extended ALG (which has all the issues discussed in Section 2.1) that can interact with NAT-PT to keep the address mapping from being discarded after a session.

3.4. DoS Attacks on Memory and Address/Port Pools

As discussed in Section 2.3, a NAT-PT may create dynamic NAT bindings, each of which consumes memory resources as well as an address (or port if NAPT-PT is used) from an address (or port) pool. A number of documents, including [RFC2766] and [NATPT-SEC] discuss the possible denial of service (DoS) attacks on basic NAT-PT and NAPT-PT that would result in a resource depletion associated with address and port pools. NAT-PT does not specify any authentication mechanisms; thus, an attacker may be able to create spurious bindings by spoofing addresses in packets sent through NAT-PT. The attack is more damaging if the attacker is able to spoof protocols with long binding timeouts (typically used for TCP).

The use of the DNS-ALG in NAT-PT introduces another vulnerability that can result in resource depletion. The attack identified in [DNS-ALG-ISSUES] exploits the use of DNS queries traversing NAT-PT to create dynamic bindings. Every time a DNS query is sent through the NAT-PT, the NAT-PT may create a new basic NAT-PT or NAPT-PT binding without any end-host authentication or authorization mechanisms. This behavior could lead to a serious DoS attack on both memory and address or port pools. Address spoofing is not required for this attack to be successful.

[DNS-ALG-SOL] proposes to mitigate the DoS attack by using Access Control Lists (ACLs) and static binds, which increases the operational cost and may not always be practical.

The ideal mitigation solution would be to disallow dynamically created binds until authentication and authorization of the end-host needing the protocol translation has been carried out. This would require that the proper security infrastructure be in place to support the authentication and authorization, which increases the network operational complexity.

4. Issues Directly Related to Use of DNS-ALG

4.1. Address Selection Issues when Communicating with Dual-Stack End-Hosts

[DNS-ALG-ISSUES] discusses NAT-PT DNS-ALG issues with regard to address selection. As specified in [RFC2766], the DNS-ALG returns AAAA Resource Records (RRs) from two possible sources, to the IPv6 host that has made an AAAA DNS query.

If the query relates to a dual-stack host, the query will return both the native IPv6 address(es) and the translated IPv4 address(es) in AAAA RRs. Without additional information, the IPv6 host address

selection may pick a translated IPv4 address instead of selecting the more appropriate native IPv6 address. Under some circumstances, the address selection algorithms [RFC3484] will always prefer the translated address over the native IPv6 address; this is obviously undesirable.

[DNS-ALG-SOL] proposes a solution that involves modification to the NAT-PT specification intended to return only the most appropriate address(es) to an IPv6 capable host as described below:

- When a DNS AAAA query traverses the NAT-PT DNS-ALG, the NAT-PT will forward the query to the DNS server in the IPv4 domain unchanged, but using IPv4 transport. The following two results can occur:
  - If the authoritative DNS server has one or more AAAA records, it returns them. The DNS-ALG then forwards this response to the IPv6 host and does not send an A query as the standard NAT-PT would do.
  - Otherwise, if the DNS server does not understand the AAAA query or has no AAAA entry for the host, it will return an error. The NAT-PT DNS-ALG will intercept the error or empty return and send an A query for the same host. If this query returns an IPv4 address, the ALG creates a binding and synthesizes a corresponding AAAA record, which it sends back to the IPv6 host.
- The NAT-PT thus forwards the result of the first successful DNS response back to the end-host or an error if neither succeeds. Consequently, only AAAA RRs from one source will be provided instead of two as specified in [RFC2766], and it will contain the most appropriate address for a dual-stack or IPv6-only querier.

There is, however, still an issue with the proposed solution:

- The DNS client may timeout the query if it doesn’t receive a response in time. This is more likely than for queries not passing through a DNS ALG because the NAT-PT may have to make two separate, sequential queries of which the client is not aware. It may be possible to reduce the response time by sending the two queries in parallel and ignoring the result of the A query if the AAAA returns one or more addresses. However, it is still necessary to delay after receiving the first response to determine if a second is coming, which may still trigger the DNS client timeout.

Unfortunately, the two queries cannot be combined in a single DNS request (all known DNS servers only process a single DNS query per request message because of difficulties expressing authoritativeness for arbitrary combinations of requests).

An alternative solution would be to allow the IPv6 host to use the NAT-PT PREFIX [RFC2766] within its address selection policies and to assign it a low selection priority. This solution requires an automatic configuration of the NAT-PT PREFIX as well as its integration within the address selection policies. The simplest way to integrate this automatic configuration would be through a configuration file download (in case the host or Dynamic Host Configuration Protocol for IPv6 (DHCPv6) server did not support vendor options and to avoid a standardization effort on the NAT-PT PREFIX option). This solution does not require any modification to the NAT-PT specification.

Neither of these solutions resolves a second issue related to address selection that is identified in [DNS-ALG-ISSUES]. Applications have no way of knowing that the IPv6 address returned from the DNS-ALG is not a 'real' IPv6 address, but a translated IPv4 address. The application may therefore, be led to believe that it has end-to-end IPv6 connectivity with the destination. As a result, the application may use IPv6-specific options that are not supported by NAT-PT. This issue is closely related to the issue described in Section 4.2 and the discussion in Section 5.

4.2. Non-Global Validity of Translated RR Records

Some applications propagate information records retrieved from DNS to other applications. The published semantics of DNS imply that the results will be consistent to any user for the duration of the attached lifetime. RR records translated by NAT-PT violate these semantics because the retrieved addresses are only usable for communications through the translating NAT-PT.

Applications that pass on retrieved DNS records to other applications will generally assume that they can rely on the passed on addresses to be usable by the receiving application. This may not be the case if the receiving application is on another node, especially if it is not in the domain served by the NAT-PT that generated the translation.

4.3. Inappropriate Translation of Responses to A Queries

Some applications running on dual-stack nodes may wish to query the IPv4 address of a destination. If the resulting A query passes through the NAT-PT DNS-ALG, the DNS-ALG will translate the response inappropriately into a AAAA record using a translated address. This happens because the DNS-ALG specified in [RFC2766] operates statelessly and hence has no memory of the IPv6 query that induced the A request on the IPv4 side. The default action is to translate the response.

The specification of NAT-PT could be modified to maintain a minimal state about queries passed through the DNS-ALG, and hence to respond correctly to A queries as well as AAAA queries.

4.4. DNS-ALG and Multi-Addressed Nodes

Many IPv6 nodes, especially in multihomed situations but also in single homed deployments, can expect to have multiple global addresses. The same may be true for multihomed IPv4 nodes. Responses to DNS queries for these nodes will normally contain all these addresses. Since the DNS-ALG in the NAT-PT has no knowledge which of the addresses can or will be used by the application issuing the query, it is obliged to translate all of them.

This could be a significant drain on resources in both basic NAT-PT and NAPT-PT, as bindings will have to be created for each address.

When using SCTP in a multihomed network, the problem is exacerbated if multiple NAT-PTs translate multiple addresses. Also, it is not clear that SCTP will actually look up all the destination IP addresses via DNS, so that bindings may not be in place when packets arrive.

4.5. Limitations on Deployment of DNS Security Capabilities

Secure DNS (DNSSEC) [RFC4033] uses public key cryptographic signing to authenticate DNS responses. The DNS-ALG modifies DNS query responses traversing the NAT-PT in both directions, which would invalidate the signatures as (partially) described in Section 7.5 of [RFC2766].

Workarounds have been proposed, such as making the DNS-ALG behave like a secure DNS server. This would need to be done separately for both the IPv6 and IPv4 domains. This is operationally very complex and there is a risk that the server could be mistaken for a conventional DNS server. The NAT-PT specification would have to be altered to implement any such workaround.

Hence, DNSSEC is not deployable in domains that use NAT-PT as currently specified. Widespread deployment of NAT-PT would become a serious obstacle to the large scale deployment of DNSSEC.

5. Impact on IPv6 Application Development

One of the major design goals for IPv6 is to restore the end-to-end transparency of the Internet. Therefore, because IPv6 may be expected to remove the need for NATs and similar impediments to transparency, developers creating applications to work with IPv6 may be tempted to assume that the complex expedients that might have been needed to make the application work in a 'NATted' IPv4 environment are not required.

Consequently, some classes of applications (e.g., peer-to-peer) that would need special measures to manage NAT traversal, including special encapsulations, attention to binding lifetime, and provision of keepalives, may build in assumptions on whether IPv6 is being used or not. Developers would also like to exploit additional capabilities of IPv6 not available in IPv4.

NAT-PT as specified in [RFC2766] is intended to work autonomously and be transparent to applications. Therefore, there is no way for application developers to discover that a path contains a NAT-PT.

If NAT-PT is deployed, applications that have assumed a NAT-free IPv6 environment may break when the traffic passes through a NAT-PT. This is bad enough, but requiring developers to include special capabilities to work around what is supposed to be a temporary transition 'aid' is even worse. Finally, deployment of NAT-PT is likely to inhibit the development and use of additional IPv6 capabilities enabled by the flexible extension header system in IPv6 packets.

Some of these deleterious effects could possibly be alleviated if applications could discover the presence of NAT-PT boxes on paths in use, allowing the applications to take steps to workaround the problems. However, requiring applications to incorporate extra code to workaround problems with a transition aid still seems to be a very bad idea: the behavior of the application in native IPv6 and NAT-PT environments would be likely to be significantly different.

6. Security Considerations

This document summarizes security issues related to the NAT-PT [RFC2766] specification. Security issues are discussed in various sections:

- Section 2.1 discusses how IPsec AH (transport and tunnel mode) and IPsec ESP transport mode are broken by NAT-PT (when IPsec UDP encapsulation is not used [RFC3498]) and authentication and encryption are generally incompatible with NAT-PT.

- Section 2.5 discusses possible fragmentation related security attacks on NAT-PT.

- Section 2.8 discusses security issues related to multicast addresses and NAT-PT.

- Section 3.3 highlights that NAT-PT is an enticing nexus for security attacks.

- Section 3.4 discusses possible NAT-PT DoS attacks on both memory and address/port pools.

- Section 4.5 discusses why NAT-PT is incompatible with DNSSEC [RFC4033] and how deployment of NAT-PT may inhibit deployment of DNSSEC.

7. Conclusion

This document has discussed a number of significant issues with NAT-PT as defined in [RFC2766]. From a deployment perspective, 3GPP networks are currently the only 'standardised' scenario where NAT-PT is envisaged as a potential mechanism to allow communication between an IPv6-only host and an IPv4-only host as discussed in the 3GPP IPv6 transition analysis [RFC4215]. But RFC 4215 recommends that the generic form of NAT-PT should not be used and that modified forms should only be used under strict conditions. Moreover, it documents a number of caveats and security issues specific to 3GPP. In addition, NAT-PT has seen some limited usage for other purposes.

Although some of the issues identified with NAT-PT appear to have solutions, many of the solutions proposed require significant alterations to the existing specification and would likely increase operational complexity. Even if these solutions were applied, we have shown that NAT-PT still has significant, irresolvable issues and appears to have limited applicability. The potential constraints on the development of IPv6 applications described in Section 5 are particularly undesirable. It appears that alternatives to NAT-PT exist to cover the circumstances where NAT-PT has been suggested as a solution, such as the use of application proxies in 3GPP scenarios [RFC4215]

However, it is clear that in some circumstances an IPv6-IPv4 protocol translation solution may be a useful transitional solution,

particularly in more constrained situations where the translator is not required to deal with traffic for a wide variety of protocols that are not determined in advance. Therefore, it is possible that a more limited form of NAT-PT could be defined for use in specific situations.

Accordingly, we recommend that:

- the IETF no longer suggest its usage as a general IPv4-IPv6 transition mechanism in the Internet, and
- RFC 2766 is moved to Historic status to limit the possibility of it being deployed inappropriately.

8. Acknowledgments

This work builds on a large body of existing work examining the issues and applicability of NAT-PT: the work of the authors of the documents referred to in Section 1 has been extremely useful in creating this document. Particular thanks are due to Pekka Savola for rapid and thorough review of the document.

9. References

9.1. Normative References

[RFC2765] Nordmark, E., "Stateless IP/ICMP Translation Algorithm (SIIT)", RFC 2765, February 2000.

[RFC2766] Tsirtsis, G. and P. Srisuresh, "Network Address Translation - Protocol Translation (NAT-PT)", RFC 2766, February 2000.

[RFC2663] Srisuresh, P. and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", RFC 2663, August 1999.

[RFC3027] Holdrege, M. and P. Srisuresh, "Protocol Complications with the IP Network Address Translator", RFC 3027, January 2001.

[RFC3314] Wasserman, M., "Recommendations for IPv6 in Third Generation Partnership Project (3GPP) Standards", RFC 3314, September 2002.

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

[RFC1035] Mockapetris, P., "Domain names - implementation and specification", STD 13, RFC 1035, November 1987.

[RFC4294] Loughney, J., "IPv6 Node Requirements", RFC 4294, April 2006.

[RFC4215] Wiljakka, J., "Analysis on IPv6 Transition in Third Generation Partnership Project (3GPP) Networks", RFC 4215, October 2005.

[RFC4033] Arends, R., Austein, R., Larson, M., Massey, D., and S. Rose, "DNS Security Introduction and Requirements", RFC 4033, March 2005.

9.2. Informative References

[RFC1858] Ziemba, G., Reed, D., and P. Traina, "Security Considerations for IP Fragment Filtering", RFC 1858, October 1995.

[RFC3128] Miller, I., "Protection Against a Variant of the Tiny Fragment Attack (RFC 1858)", RFC 3128, June 2001.

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

[RFC3498] Kuhfeld, J., Johnson, J., and M. Thatcher, "Definitions of Managed Objects for Synchronous Optical Network (SONET) Linear Automatic Protection Switching (APS) Architectures", RFC 3498, March 2003.

[NATP-APP] Satapati, S., "NAT-PT Applicability", Work in Progress, October 2003.

[DNS-ALG-ISSUES] Durand, A., "Issues with NAT-PT DNS ALG in RFC2766", Work in Progress, February 2002.

[DNS-ALG-SOL] Hallingby, P. and S. Satapati, "NAT-PT DNS ALG solutions", Work in Progress, July 2002.

[NATPT-MOB] Shin, M. and J. Lee, "Considerations for Mobility Support in NAT-PT", Work in Progress, July 2005.

[NATPT-SEC] Okazaki, S. and A. Desai, "NAT-PT Security Considerations", Work in Progress, June 2003.

[TRANS-ISSUES] Pol, R., Satapati, S., and S. Sivakumar, "Issues when translating between IPv4 and IPv6", Work in Progress, January 2003.

[3GPP-TRANS] Malki, K., "IPv6-IPv4 Translation mechanism for SIP-based services in Third Generation Partnership Project (3GPP) Networks", Work in Progress, December 2003.

[MTP] Tsuchiya, K., Higuchi, H., Sawada, S., and S. Nozaki, "An IPv6/IPv4 Multicast Translator based on IGMP/MLD Proxying (mtp)", Work in Progress, February 2003.

[MCASTGW] Venaas, S., "An IPv4 - IPv6 multicast gateway", Work in Progress, February 2003.

[MUL-NATPT] Park, S., "Scalable mNAT-PT Solution", Work in Progress, May 2003.

Authors' Addresses

Cedric Aoun
Energize Urnet
Paris
France

EMail: ietf@energizeurnet.com

Elwyn B. Davies
Folly Consulting
Soham, Cambs
UK

Phone: +44 7889 488 335
EMail: elwynd@dial.pipex.com

Full Copyright Statement

Intellectual Property

Acknowledgement

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