BEHAVE
Internet-Draft
Expires: January 9, 2005

F. AUDET
Nortel Networks
C. Jennings
Cisco Systems
July 11, 2004

NAT/Firewall Behavioral Requirements
draft-audet-nat-behave-00

Abstract

This document defines basic terminology for describing different types of behavior for NATs and firewalls. It also defines a set of requirements for NATs and firewalls that would allow many applications, such as multimedia communications or online gaming, to work consistently. Developing NATs and firewalls that meet this set of requirements will greatly increase the likelihood that applications will function properly.

Table of Contents

1. Introduction

Network Address Translators (NAT) and firewalls are well known to cause very significant problems with applications that carry IP addresses in the payload [5]. Applications that suffer from this problem include Voice Over IP and Multimedia Over IP (e.g., SIP [6] and H.323 [11]), as well as online gaming.

Many techniques are used to attempt to make realtime multimedia applications, online games, and other applications work across NATs and firewalls. Application Level Gateways [3] are one such mechanism. STUN [7] describes a UNilateral Self-Address Translation (UNSAF) mechanism[2]. Media Relays have also been used to enable applications across NATs and firewalls, but these are generally seen as a solution of last resort. ICE [9] describes a methodology for using many of these techniques and avoiding a Media Relay unless the type of NAT/firewall is such that it forces the use of such a Media Relay.

This specification defines requirements for NATs and firewalls aimed at ensuring that a NAT or firewall that satisfies these requirements will avoid forcing the use of a Media Relay for supporting applications. "Peer-to-Peer (P2P) communication across middle boxes" [10] made several recommendations regarding NATs and firewalls for Peer-to-Peer media; this specification derives a lot of its requirements from that draft.

As pointed out in UNSAF [2], "From observations of deployed networks, it is clear that different NAT boxes implementation vary widely in terms of how they handle different traffic and addressing cases." This wide degree of variability is one part of what contributes to the overall brittleness introduced by NATs and makes it extremely difficult to predict how any given protocol will behave on a network traversing NATs. Discussions with many of the major NAT vendors have made it clear that they would prefer to deploy NATs that were deterministic and caused the least harm to applications while still meeting the requirements that caused their customers to deploy NATs in the first place. The problem the NAT vendors face is they are not sure how best to do that or how to document how their NATs behave. The situation is not as problematic for firewalls but still exists: there is no good common terminology even to describe the behavior of firewalls.

The goals of this document are to define a set of common terminology for describing the behavior of NATs and firewalls and to produce a set of requirements on a specific set of behaviors for NATs and firewalls. The requirements represent what many vendors are already doing, and it is not expected that it should be any more difficult to build a NAT that meets these requirements or that these requirements should affect performance.

The authors strongly believe that if there were a common set of requirements that were simple and useful for voice, video, and games, the bulk of the NAT vendors would choose to meet those requirements. This document will simplify the analysis of protocols for deciding whether or not they work in this environment and will allow providers of services that have NAT traversal issues to make statements about where their applications will work and where they will not, as well as to specify requirements for NATs.

This specification only covers Traditional NATs [4]. Bi-directional, Twice NAT, and Multihomed NAT [3] are outside the scope of this document. Approaches using directly signaled control off the middle boxes such as midcom, UPNP or in-path signaling are also out of scope. Media Relays are out of the scope of this document as well.

This document only covers the UDP aspects of NAT/firewall traversal and does not cover TCP, ICMP, IPSEC, or other protocols.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [1].

It is assumed that the reader is familiar with the terminology described in RFC 2663 [3] and RFC 3022 [4]. This specification attempts to preserve the terminology used in those RFCs.

This document uses the term "session" as defined in RFC 2663 [3]: "TCP/UDP sessions are uniquely identified by the tuple of (source IP address, source TCP/UDP ports, target IP address, target TCP/UDP Port)."

This document uses the term "address binding" as defined in RFC 2663 [3] and RFC 3022 [4]: "Address binding is the phase in which a local IP address is associated with an external address, or vice versa, for purpose of translation."

The term NAT is used to refer to both traditional address translation and address port translation. The authors understand that there was a time when these were considered different, but terminology has changed over time, and the term NAT has subsumed port translation as part of it.

RFC 3489 [7] defines a terminology for different NAT variations. In particular, it uses the terms "Full Cone", "Restricted Cone", "Port Restricted Cone" and "Symmetric" to refer to different variations of NATs/firewalls. Unfortunately, this terminology has been the source of much confusion. This terminology does not distinguish between the NAT and the firewalling behavior of NAT/firewall devices. It was found that many devices' behaviors do not exactly fit into the described variations. For example, a device could be symmetric from a firewall point of view and Cone from a NAT point of view. Other aspects of NAT/firewall are not covered by this terminology: for example, many NATs will switch over from basic NAT (preserving ports) to NAPT (mapping ports) in order to preserve ports when possible.

This specification will therefore not use the Cone/Symmetric terminology. Furthermore, many other important behaviors are not fully described by the Cone/Symmetric terminology. This specification refers to specific individual NAT/Firewall behaviors instead of using the Cone/Symmetric terminology.

Note: RFC 3489 [7] defines a "Symmetric NAT" in effectively two parts:

1. All requests from the same internal IP address and port to a specific destination IP address and port are mapped to the same external IP address and port. If a host sends a packet with the same source address and port to different destination addresses or ports, a different mapping is used for each.
2. Furthermore, only the external host that receives a packet can send a UDP packet back to the internal host.

Condition 1 is the NAT behavior and condition 2 is the firewall behavior. However, they are not necessarily dependent: we have observed NATs that will conform to condition (1) but not to (2). Using RFC 3489, this type of NAT would be detected as a "Cone NAT" since it uses condition (2). Using a different algorithm such as the one described in NATCECK [12] which uses condition (1), it would be detected as a "Symmetric NAT". If the endpoint receiving the media has a permissive policy on accepting media, condition (2) is more appropriate, but if it has a restrictive policy, condition (1) is more appropriate.

3. UDP NAT Behavior

This section describes the various NAT behaviors applicable to dynamic NAT; static NAT is outside the scope of this document.

3.1 Address and port binding

When an internal endpoint opens an outgoing UDP session through a

NAT, the NAT assigns the session an external IP address and port number so that subsequent response packets from the external endpoint can be received by the NAT, translated and forwarded to the internal endpoint. This is a binding between an internal IP address and port (IP:port) and external IP:port tuple. It establishes the translation that will be performed by the NAT for the duration of the session. For many applications, it is important to distinguish the behavior of the NAT when there are multiple simultaneous sessions established to different external endpoints.

The key behavior to describe is the criteria for re-use of a binding for new sessions to external endpoints, after establishing a first binding between an internal X:x address and port and an external Y1:y1 tuple. Let's assume that internal IP address and port X:x is mapped to X1':x1' for this first session. The endpoint then sends from X:x to an external address Y2:y2 and gets a mapping X2':x2' on the NAT. The relationship between X1':x1' to X2':x2' for various combinations of the relationship between Y1:y1 to Y2:y2 is critical for describing the NAT behavior. This arrangement is illustrated in the following diagram:

```
                                          E
       +------+                 +------+  x
       |  Y1  |                 |  Y2  |  t
       +--+---+                 +---+--+  e
          | Y1:y1           Y2:y2  |     r
          +----------+   +----------+     n
                     |   |                a
             X1':x1' |   | X2':x2'        l
                  +--+---+-+
       ...........|   NAT  |...............
                  +--+---+-+              I
                     |   |                n
                 X:x |   | X:x            t
                    ++---++               e
                    |  X  |               r
                    +-----+               n
                                          a
                                          l
```

The following address and port binding behavior are defined:

External NAT binding is endpoint independent: The NAT reuses the port binding for subsequent sessions initiated from the same internal IP address and port (X:x) to any external IP address and port. Specifically, X1':x1' equals X2':x2' for all values of Y2:y2.

(From a RFC 3489 NAT perspective, this is a "Cone NAT" where the sub-type is really based on the firewall behavior.)

External NAT binding is endpoint address dependent: The NAT reuses the port binding for subsequent sessions initiated from the same internal IP address and port (X:x) only for sessions to the same external IP address, regardless of the external port. Specifically, X1':x1' equals X2':x2' if, and only if, Y2 equals Y1. (From an RFC 3489 NAT perspective, but not necessarily a firewall perspective, this is a "Symmetric NAT".)

External NAT binding is endpoint address and port dependent: The NAT reuses the port binding for subsequent sessions initiated from the same internal IP address and port (X:x) only for sessions to the same external and port. Specifically, X1':x1' equals X2':x2' if, and only if, Y2:y2 equals Y1:y1. (From an RFC 3489 NAT perspective, but not necessarily a firewall perspective, this is a "Symmetric NAT".)

The three possibilities are abbreviated as NB=I, NB=AD, and NB=APD, respectively. NB stands for Nat Binding, I for independent, AD for Address Dependent, and APD for Address Port Dependent.

It is important to note that these three possible choices make no difference to the security properties of the NAT. The security properties are fully determined by which packets the NAT allows in and which it does not. This is determined by the firewall behavior in the firewall portions of the NAT/FW.

3.2 Port assignment

Some NATs attempt to preserve the port number used internally when assigning a binding to an external IP address and port (e.g., X:x to X':x). A basic NAT, for example, will preserve the same port and will assign a different IP address from a pool of external IP addresses in case of port collision (e.g. X1:x to X1':x and X2:x to X2':x). This is only possible as long as the NAT has enough external IP addresses. If the port x is already in use on all available external IP addresses, then the NAT needs to switch from Basic NAT to a Network Address and Port Translator (NAPT) mode (i.e., X1:x to X':x and X2:x to X':x'). This is referred-to as "port preservation". It does not guarantee that the external port x' will always be the same as the internal port x but only that the NAT will preserve the port if possible.

A NAT that does not attempt to make the external port numbers match the internal port numbers in any case (i.e., X1:x to X':x1', X2:x to X':x2') is referred to as "no port preservation".

Tools such as network sniffers identify traffic based on the destination port, not the source port, so port preservation does not help these tools.

Some particularly nasty NATs use Port overloading, i.e. they always use port preservation even in the case of collision (i.e., X1:x to X':x, and X2:x to X':x). These NATs rely on the source of the response from the external endpoint (Y:y, Z:z) to forward a packet to the proper internal endpoint (X1 or X2). Port overloading fails if the two internal endpoints are establishing sessions to the same external destination. This is referred to as "Port overloaded".

Most applications fail in some cases with "Port overloaded". It is clear that "Port overloaded" behavior will result in many problems.

3.3 Bind Refresh Direction

NAT UDP binding timeouts implementations vary but include the timer's value and the way the binding timer is refreshed to keep the binding alive.

The binding timer is defined as the time a binding will stay active without packets traversing the NAT. There is great variation in the values used by different NATs.

Some NATs keep the binding active (i.e., refresh the timer value) when a packet goes from the internal side of the NAT to the external side of the NAT, but do not take into account packets from the external side of the NAT to the internal side of the NAT. This is referred to as having a NAT refresh direction behavior of "Outbound".

Other NATs keep the binding active when packets go in any direction. This is referred to as "Bidirectional" NAT refresh direction behavior.

Yet other NATs keep the binding active when a packet goes from the external side of the NAT to the internal side of the NAT but do not take into account packets from the internal side of the NAT to the external side of the NAT. This is referred to as having a NAT refresh direction behavior of "Inbound".

3.4 Bind Refresh Scope

If the binding is refreshed for all sessions on that bind by any outbound traffic, the NAT is said to have a NAT refresh method behavior of "Per binding". If the binding is refreshed only on a specific session on that particular bind by any outbound traffic, the NAT is said to have a "Per session" NAT refresh method behavior.

4. UDP Firewall Behavior (Filtering)

This section describes various firewall behaviors.

4.1 Filtering of unsolicited packets

When an internal endpoint opens an outgoing UDP session through a firewall, the firewall assigns a filtering rule for the binding between an internal IP:port (X:x) and external IP:port (Y:y) tuple.

The key behavior to describe is what criteria are used by the firewall to filter packets originating from specific external endpoints.

External filtering is open: The firewall does not filter any packets.

External filtering is endpoint independent: The firewall filters out only packets not destined to the internal address and port X:x, regardless of the external IP address and port source (Z:z). The firewall forwards any packets destined to X:x. In other words, sending packets from the internal side of the firewall to any external IP address is sufficient to allow any packets back to the internal endpoint. (From an RFC 3489 Firewall perspective, this is a "Full Cone Firewall".)

External filtering is endpoint address dependent: The firewall filters out packets not destined to the internal address X:x. Additionally, the firewall will filter out packets from Y:y destined for the internal endpoint X:x if X:x has not sent packets to Y previously (independently of the port used by Y). In other words, for receiving packets from a specific external endpoint, it is necessary for the internal endpoint to send packets first to that specific external endpoint's IP address. (From an RFC 3489 Firewall perspective, this is a "Restricted Cone Firewall".)

External filtering is endpoint address and port dependent: This is similar to the previous behavior, except that the external port is also relevant. The firewall filters out packets not destined for the internal address X:x. Additionally, the firewall will filter out packets from Y:y destined for the internal endpoint X:x if X:x has not sent packets to Y:y previously. In other words, for receiving packets from a specific external endpoint, it is necessary for the internal endpoint to send packets first to that external endpoint's IP address and port. (From an RFC 3489 Firewall perspective, this is both a "Port Restricted Cone Firewall" and a "Symmetric Firewall" as they have the same firewall behavior.)

These are abbreviated at EF=O, EF=I, EF=AD, EF=APD respectively. In the case of a NAT, "open" cannot forward a packet unless there is a NAT binding, so for all practical purposes, a NAT will never be "open" but will be one of the others.

4.2 Firewall Filter Refresh

The time for which a firewall filter is valid can be refreshed based on packets that are inbound, outbound, or going either direction. In the case of a EF=AD or EF=APD firewall, the scope of the refresh could include the filters for just the particular port and destination or for all the ports and destinations sharing the same external address and port on the firewall.

5. Hairpinning Behavior

If two hosts (called X1 and X2) are behind the same NAT and exchanging traffic, the NAT may allocate an address on the outside of the NAT for X2, called X2':x2'. If X1 sends traffic to X2':x2', it goes to the NAT, which must relay the traffic from X1 to X2. This is referred to as hairpinning and is illustrated below.

```
                                   NAT
  +----+ from X1:x1 to X2':x2'   +-----+ X1':x1'
  | X1 |>>>>>>>>>>>>>>>>>>>>>>>>>>>>--+---
  +----+                         |  v  |
                                 |  v  |
                                 |  v  |
                                 |  v  |
  +----+ from X2':x2' to X1:x1   |  v  | X2':x2'
  | X2 |<<<<<<<<<<<<<<<<<<<<<<<<<<<<--+---
  +----+                         +-----+
```

Hairpinning allows two endpoints on the internal side of the NAT to communicate even if they only use each other's external IP addresses and ports.

More formally, a NAT/firewall that supports hairpinning forwards packets originating from an internal address, X1:x1, destined for an external address X2':x2' that has an active binding to an internal address X2:x2, back to that internal address X2:x2. (Note that typically, X1'=X2'.)

Furthermore, the NAT may present the hairpinned packet with either an internal or an external source IP address and port. The hairpinning

NAT behavior can therefore be either "External source IP address and port" or "Internal source IP address and port". "Internal source IP address and port" may cause problems by confusing an implementation that is expecting an external IP address and port.

6. Deterministic Properties

The diagnosis is further complicated by the fact that under some conditions the same NAT will exhibit different behaviors. This has been seen on NATs that preserve ports or have specific algorithms for selecting a port other than a free one. If the external port that the NAT wishes to use is already in use by another session, the NAT must select a different port. This results in different code paths for this conflict case, which results in different behavior.

For example, if three hosts X1, X2, and X3 all send from the same port x, through a port preserving NAT with only one external IP address, called X1', the first one to send (i.e., X1) will get an external port of x but the next two will get x2' and x3' (where these are not equal to x). There are NATs where the External NAT Binding characteristics and the External Filter characteristics change between the X1:x and the X2:x binding. To make matters worse, there are NATs where the behavior may be the same on the X1:x and X2:x binds but different on the third X3:x binding.

Some NATs that try to reuse external ports flow from two internal IP addresses to two different external IP addresses. For example, X1:x is going to Y1:y1 and X2:x is going to Y2:y2, where Y1:y1 does not equal Y2:y2. Some NATs will map X1:x to X1':x and will also map X2:x to X1':x. This works in the case where the NAT Binding is address port dependent. However some NATs change their behavior when this type of port reuse is happening. The NAT may look like it has NAT Bindings that are independent when this type of reuse is not happening but may change to Address Port Dependent when this reuse happens.

Any NAT that changes the NAT Binding or the External Filtering at any point in time or under any particular conditions is referred to as a "non-deterministic" NAT. NATs that don't are called "deterministic".

Non-deterministic NATs generally change behavior when a conflict of some sort happens, i.e. when the port that would normally be used is already in use by another bind. The NAT binding and External Filtering in the absence of conflict is referred to as the Primary behavior. The behavior after the first conflict is referred to as Secondary and after the second conflict is referred to as Tertiary. No NATs have been observed that change on further conflicts but additional testing may be required.

7. ICMP Behavior

There are cases in which a host inside the NAT sends a packet to the NAT that gets relayed towards a host on the external side of the NAT that results in an ICMP Destination Unreachable message being returned to the NAT. Most NATs and firewalls will send an appropriate ICMP Destination Unreachable message to the internal host that sent the original packet. NATs and firewalls that do not filter out this ICMP Destination Unreachable message when it is in reply to a IP packet sent are referred to as "Support Destination Unreachable" (abbreviated SU).

Incoming Destination Unreachable messages can be ignored after some period of time after the packet which elicited the Destination Unreachable message. This IMCP timeout needs to be greater than the RTT for any destination the NAT may attempt to send IP packets to. Keep in mind satellite links when setting this timeout.

Applications use the destination unreachable message to decide that they can stop trying to retransmit to a particular IP address and can fail over to a secondary address. If a destination unreachable message is not received, the fail over will take too long for many applications. Another key use of this message is for MTU discovery (described in RFC 1191 [14]). MTU discovery is important for allowing applications to avoid the fragmentation problems discussed in the next section.

There is no significant security advantage to blocking these ICMP Destination Unreachable packets.

8. Fragmentation Behavior

When a fragmented packet is received on the external side, some NATs forward the packet to the same location as a recent initial fragment packet with the same identifier in the IP header. Other NATs reassemble fragmented packets and forward them after reassembly. NATs that do either of these are referred to as "Support Fragmentation" (abbreviated SF).

When a fragmented packet is received from the external side and the packets are out of order so that the initial fragment does not arrive first, many systems simply discard the out of order packets. Moreover, since some networks deliver small packets ahead of large ones, there can be many out of order fragments. NATs that are capable of delivering these out of order packets are possible but they need to store the out of order fragments, which opens up a DOS opportunity.

Fragmentation has been a tool used in many attacks, some involving passing fragmented packets through firewalls and others involving DOS attacks based on the state needed to reassemble the fragments. Firewall implementers should be aware of RFC 3128 [17] and RFC 1858 [16].

NATs that do not remap the identification field in the IP header run the risk that two hosts behind the NAT will choose the same value, and a host receiving packets from them will not be able to correctly reassemble the packets. It seems unlikely that this will happen often in practice.

9. TCP Behavior

TCP connections are often long lived with long periods of no traffic. The timeouts for the NAT bindings and firewall filters need to be set appropriately. In the initial stage where a SYN has been sent but there is no ACK, the timeouts can be fairly short. Typically they are set at around one minute. After an ACK is received the session is connected and needs to have a very long timer, typically hours. This time is called the TCP timeout. After a RST or FIN packet is seen, the timeout can be reduced to a short time such as one minute.

10. Multicast and IGMP Behavior

This section is weak and requires more discussion, thought, and experimentation with existing systems. Take it with a grain of salt and expect it to change significantly as this document matures.

Some NATs support multicast while others block it. In general to support multicast, the NAT needs to process the source address as it would processes other UDP packets but not modify the destination address. It also needs to process IGMP packets as a normal router would. Multicast is used by various applications including some that deliver video to residences.

The simplest implementation would forward packets that are addressed to a multicast destination and would proxy IGMP messages in the same way that a NAT can proxy ICMP messages. A more complex implementation would fully process the IGMPv3 RFC 3376 [15] messages and only forward multicast packets based on the information IGMP has provided.

If a device inside the NAT can receive multicast traffic from a sender outside the NAT after the device inside sends an appropriate IGMP message, the NAT is said to "Support Multicast" (abbreviated SM).

11. Requirements

The requirements in this section are aimed at minimizing the damage caused by NATs and firewalls to applications such as realtime communications and online gaming.

It should be understood, however, that applications normally do not know in advance if the NAT or firewall conforms to the recommendations defined in this section. Peer-to-peer media applications still need to use normal procedures such as ICE [9].

REQ-1: A NAT MUST have an "External NAT Binding is endpoint independent" behavior (NB=I).

REQ-2: It is RECOMMENDED that a NAT have a "No port preservation" behavior.

REQ-2a: A NAT MAY use a "Port preservation" behavior.
REQ-2b: A NAT MUST NOT have a "Port overloaded" behavior.

REQ-3: A dynamic NAT UDP binding timer MUST NOT expire in less than 2 minutes.

REQ-3a: The value of the NAT UDP binding timer MAY be configurable.
REQ-3b: A default value of 5 minutes for the NAT UDP binding timer of 5 minutes is RECOMMENDED.

REQ-4: The NAT UDP timeout binding MUST have a NAT refresh direction behavior of "Outbound" (i.e. based on outbound traffic only).

REQ-4a: The NAT UDP timeout binding MUST have a NAT refresh method behavior of "Per binding" (i.e. refresh all sessions active on a particular bind).

REQ-5: It is RECOMMENDED that a firewall have an "External filtering is endpoint address dependent" behavior. (EF=AD)

REQ-5a: A firewall MAY have an "External filtering is endpoint independent" behavior. (EF=I)
REQ-5b: A firewall MAY have an "External filtering is endpoint address and port dependent" behavior. (EF=APD)

REQ-6: The firewall UDP filter timeout behavior MUST be the same as the NAT UDP binding timeout.

REQ-7: A NAT/FW MUST support "Hairpinning" behavior.

REQ-7a: A NAT/FW Hairpinning NAT behavior MUST be "External source IP address and port".

REQ-8: A NAT MUST have the capability to turn off individually all ALGs it supports, except for DNS and IPsec.

REQ-8a: Any NAT ALG for SIP MUST be turned off by default.
REQ-9: A NAT/firewall MUST have deterministic behavior.
REQ-10: The TCP binding timeout for NATs and the filter rule timeout for firewalls MUST be greater than 7800 seconds.
REQ-11: A NAT/firewall SHOULD support forwarding fragmented packets (SF).
REQ-12: A NAT/FW MUST support ICMP Destination Unreachable (SU).

REQ-12a: The ICMP timeout SHOULD be greater than 2 seconds.
REQ-13: A NAT/FW SHOULD support forwarding multicast packets (SM).

11.1 Requirement Discussion

This section describes why each of these requirements was chosen and the consequences of violating any of them:

REQ-1: In order for UNSAF methods to work, REQ-1 needs to be met. Failure to meet REQ-1 will force the use of a Media Relay which is very often impractical.

REQ-2: NATs that implement port preservation have to deal with conflicts on ports, and the multiple code paths this introduces often result in nondeterministic behavior.

REQ-2a: Port preservation can work, but the NAT implementors need to be very careful that it does not become a nondeterministic NAT.

REQ-2b: REQ-2b must be met in order to enable two applications on the internal side of the NAT both to use the same port to try to communicate with the same destination.

REQ-3: This requirement is to ensure that the timeout is long enough to avoid too frequent timer refresh packets.

REQ-3a: Configuration is desirable for adapting to specific networks and troubleshooting.

REQ-3b: This default is to avoid too frequent timer refresh packets.

REQ-4: This requirement is a security concern: it is not secure to let inbound traffic refresh the timer, as an outside party could use it to keep a port open on the NAT/firewall.

REQ-4a: Using the refresh on a per binding basis avoids the need for separate keep-alives for all the available sessions.

REQ-5: Filtering based on the IP address is felt to have the maximum balance between security and usefulness. See below.

REQ-5a: Filtering independently of the external IP address and port is not as secure: an unauthorized packet could get at a specific port while the port was kept open if it was lucky enough to find the port open.

REQ-5b: In theory, filtering based on both IP address and port is more secure than filtering based only on the IP address (because the external endpoint could in reality be two endpoints behind another NAT, where one of the two endpoints is an attacker). However, such a restrictive policy could interfere with certain applications that use more than one port.

REQ-6: This is to avoid overly complex applications.

REQ-7: This requirement is to allow communications between two endpoints behind the same NAT/firewall when they are trying each other's external IP addresses.

REQ-7a: Using the external IP address is necessary for applications with a restrictive policy of not accepting packets from IP addresses that differ from what is expected.

REQ-8: NAT ALGs may interfere with UNSAF methods.

REQ-8a: A SIP NAT ALG will interfere with UNSAF methods.

REQ-9: Non-deterministic NATs are very difficult to troubleshoot because they require more intensive testing. This non-deterministic behavior is the root cause of much of the uncertainty that NATs introduce about whether or not applications will work.

REQ-10: Most operating systems have a default TCP keep alive time of 2 hours, plus it can take 10 minutes for the keep alive to happen or fail with all the default timeouts. The sum of these leads to the recommendation of 7800 seconds.

Req-11: Fragmented packets become more common with large video packets and should continue to work. Applications can use MTU discovery to work around this.
Req-12: This is easy to do, is used for many things including MTU discovery and rapid detection of error conditions, and has no negative consequences.
Req-13: Minimal support of multicast for NATs is simple and allows interesting applications.

12. Security Considerations

Firewalls and NATs are often deployed to achieve security goals. Most of the recommendations and requirements in this document do not affect the security properties of these devices, but a few of them do have security implications and are discussed in this section.

This work recommends that the timers for binding be refreshed only on outgoing packets and that inbound packets should not update the timers. If inbound packets update the timers, an external attacker can keep the binding alive forever and attack future devices that may end up with the same internal address. Some devices today do update the timers on inbound packets.

This work recommends that the firewall filters be specific to the external IP only and not the external IP and port. It can be argued that this is less secure than using the IP and port. Devices that wish to filter on IP and port do still comply with these requirements.

Non-deterministic NATs and firewalls are risky from a security point of view. They are very difficult to test because they are, well, non-deterministic. Testing by a person configuring one may result in the person thinking it is behaving as desired, yet under different conditions, which an attacker can create, it may behave differently. These requirements recommend that devices be deterministic.

The work requires that NATs have an "external NAT binding is endpoint independent" behavior. This does not reduce the security of devices. Which packets are allowed to flow across the device is determined by the external filtering behavior, which is independent of the binding behavior.

13. IANA Considerations

There are no IANA considerations.

14. IAB Considerations

The IAB has studied the problem of "Unilateral Self Address Fixing", which is the general process by which a client attempts to determine its address in another realm on the other side of a NAT through a collaborative protocol reflection mechanism [2].

This specification does not constitute in itself an UNSAF application. It consist of a series of requirements for NATs and firewalls aimed at minimizing the negative impact that those devices have on peer-to-peer media applications, especially when those applications are using UNSAF methods.

Section 3 of UNSAF lists several practical issues with solutions to NAT problems. This document makes recommendations to reduce the uncertainty and problems introduced by these practical issues with NATs. In addition, UNSAF [2] lists five architectural considerations. Though this is not an UNSAF proposal, it is interesting to consider the impact of this work on these architectural considerations.

Arch-1: The scope of this is limited to UDP packets in NATs like the ones widely deployed today. The "fix" helps constrain the variability of NATs for true UNSAF solutions such as STUN.

Arch-2: This will exit at same rate that NATs exit. It does not imply any protocol machinery that would continue to live after NATs were gone or make it more difficult to remove them.

Arch-3: This does not reduce the overall brittleness of NATs but will hopefully reduce some of the more outrageous NAT behaviors and make it easer to discuss and predict NAT behavior in given situations.

Arch-4: This work combined with the test results [13] of various NATs represent the most comprehensive work at IETF on what the real issues are with NATs for applications like VoIP. This work and STUN have pointed out more than anything else the brittleness NATs introduce and the difficulty of solving these issues.

Arch-5: This work and the test results [13] provide a reference model for what any UNSAF proposal might encounter in deployed NATs.

15. Acknowledgments

The editor would like to acknowledge Bryan Ford, Pyda Srisuresh and Dan Kegel for the NATP2P [10] draft, from which a lot of the material in this specification is derived. Thanks to Rohan Mahy for many discussions on this and much helpful text. Jonathan Rosenberg provided key suggestions and corrections, and Mary Barnes provided very helpful review.

16. References

16.1 Normative References

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

[2] Daigle, L. and IAB, "IAB Considerations for UNilateral Self-Address Fixing (UNSAF) Across Network Address Translation", RFC 3424, November 2002.

16.2 Informational References

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

[4] Srisuresh, P. and K. Egevang, "Traditional IP Network Address Translator (Traditional NAT)", RFC 3022, January 2001.

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

[6] 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.

[7] Rosenberg, J., Weinberger, J., Huitema, C. and R. Mahy, "STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)", RFC 3489, March 2003.

[8] Schulzrinne, H., Casner, S., Frederick, R. and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", RFC 3550, July 2003.

[9] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Methodology for Network Address Translator (NAT) Traversal for the Session Initiation Protocol (SIP)", draft-ietf-mmusic-ice-00 (work in progress), February 2004.

[10] Ford, B., "Network Address Translation and Peer-to-Peer Applications (NATP2P)", draft-ford-natp2p-00 (work in progress), April 2003.

[11] "Packet-based Multimedia Communications Systems (includes Annex C - H.323 on ATM)", ITU-T Recommendation H.323v3, September 1999.

[12] Ford, B. and D. Andersen, "Nat Check Web Site: http://midcom-p2p.sourceforge.net", June 2004.

[13] Jennings, C., "NAT Classification Results using STUN", draft-jennings-midcom-stun-results-00 (work in progress), February 2004.

[14] Mogul, J. and S. Deering, "Path MTU discovery", RFC 1191, November 1990.

[15] Cain, B., Deering, S., Kouvelas, I., Fenner, B. and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, October 2002.

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

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

Authors' Addresses

Francois AUDET
Nortel Networks
4655 Great America Parkway
Santa Clara, CA 95054
USA

Phone: +1 408 495 3756
EMail: audets@nortelnetworks.com

Cullen Jennings
Cisco Systems
170 West Tasman Drive
MS: SJC-21/2
San Jose, CA 95134
USA

Phone: +1 408 902-3341
EMail: fluffy@cisco.com

Intellectual Property Statement

Disclaimer of Validity

Copyright Statement

Acknowledgment

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