ASR1K NAT Intermittently Fails to Translate Some Packets

Contents

Introduction

This document describes a situation where packets that should be translated by Network Address Translation (NAT) on a Cisco 1000 Series Aggregation Services Router (ASR1K) are not translated (NAT is bypassed). This could result in traffic failure as the next hop is likely not configured to allow the untranslated packets to be processed.

Background Information

In Software Version 12.2(33)XND a feature called NAT Gatekeeper was introduced and enabled by default. NAT Gatekeeper was designed to prevent non-NAT-ed flows from using excessive CPU in an effort to create a NAT translation. In order to achieve this, two small caches (one for in2out direction and one for out2in direction) are created based on the source address. Each cache entry consists of a source address, a virtual routing and forwarding (VRF) ID, a timer value (used to invalidate the entry after 10 seconds), and a frame counter. There are 256 entries in the table that make up the cache. If there are multiple traffic flows from the same source address where some packets require NAT and some do not, it could result in packets not being NAT-ed and sent through the router untranslated. Cisco recommends that customers should avoid having NAT-ed and non-NAT-ed flows on the same interface wherever possible.

> **Note**: This has nothing to do with H.323.

Demonstration of NAT Being Bypassed

This section describes how NAT can be bypassed due to the NAT gatekeeper feature. Review the

diagram in detail. You can see there is a source router, an Adaptive Security Appliance (ASA) firewall, the ASR1K, and the destination router.

Traffic Flows to Non-NAT-ed Destination

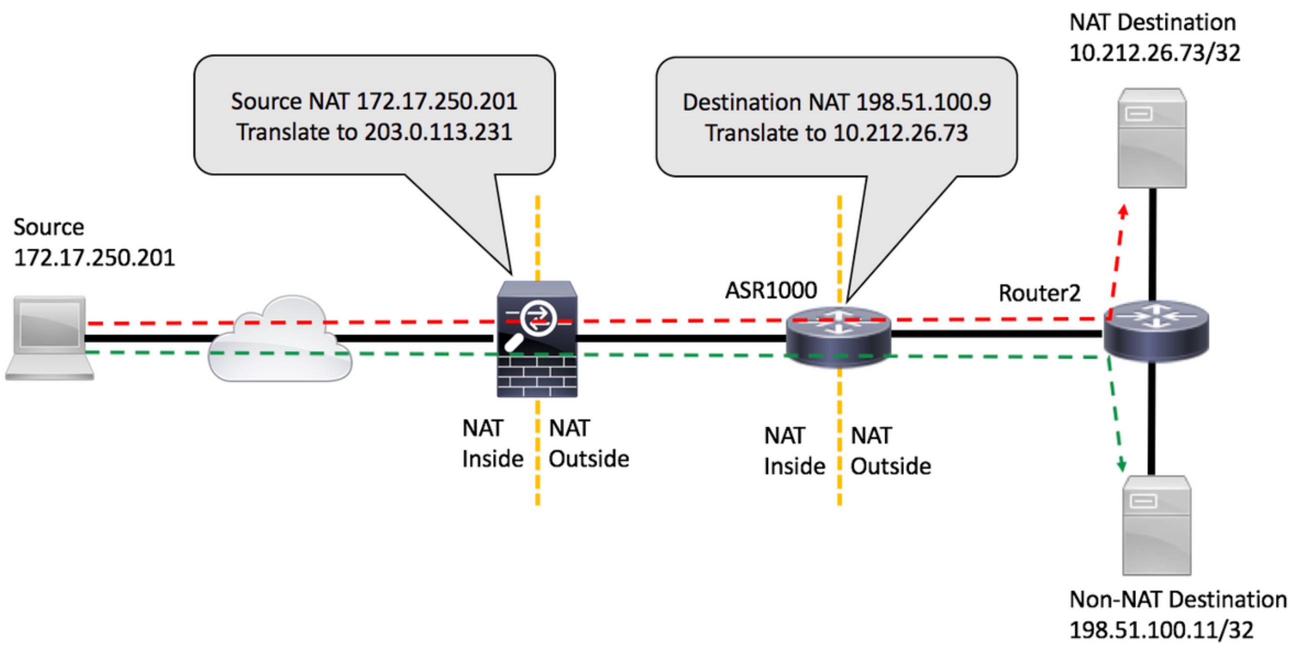
1. Ping is initiated from the source: Source: 172.17.250.201 Destination: 198.51.100.11.
2. The packet arrives on the inside interface of the ASA which performs source address translation. The packet will now have Source: 203.0.113.231 Destination: 198.51.100.11.
3. The packet arrives at the ASR1K on the NAT outside to inside interface. NAT translation finds no translation for the destination address and so the gatekeeper "out" cache is populated with the source address 203.0.113.231.
4. The packet arrives at the destination. The destination accepts the Internet Control Message Protocol (ICMP) packet and returns an ICMP ECHO Reply which results in ping success.

Traffic From Same Source Attempts to Send NAT-ed Destination

1. .Ping is initiated from the source: Source: 172.17.250.201 Destination: 198.51.100.9.
2. The packet arrives on the inside interface of the ASA which performs source address translation. The packet will now have Source: 203.0.113.231 Destination: 198.51.100.9.
3. The packet arrives at the ASR1K on the NAT outside to inside interface. NAT first looks for a translation for the source and destination. As it does not find one, it checks the gatekeeper "out" cache and finds the source address 203.0.113.231. It (erroneously) assumes that the packet does not need translation and either forwards the packet if a route exists for the destination or drops the packet. Either way, the packet will not reach the intended destination.

Restoration of NAT-ed Traffic

1. After 10 seconds, the entry for source address 203.0.113.231 times out in the gatekeeper out cache. **Note**: The entry still physically exists in the cache, but because it has expired it is not used.
2. Now if the same source 172.17.250.201 sends to NAT-ed destination 198.51.100.9.When the packet arrives at the out2in interface on the ASR1K, no translation will be found. When you check the gatekeeper out cache, you will not find an active entry and so you will create the translation for the destination and packets willl flow as expected.
3. Traffic in this flow will continue as long as translations are not timed out due to inactivity. If, in the meantime, the source again sends traffic to a non-NAT-ed destination, which causes another entry to be populated in the gatekeeper out cache, it will not affect established sessions but there will be a 10 second period in which new sessions from that same source to NAT-ed destinations will fail.



Example of the Issue

1. Ping is initiated from the source router : Source: 172.17.250.201 Destination: 198.51.100.9. The ping is issued with repeat count of two, over and over [FLOW1].
2. Then ping a different destination which is not being NAT-ed by the ASR1K: Source: 172.17.250.201 Destination:198.51.100.11 [FLOW2].
3. Then send more packets to 198.51.100.9 [FLOW1]. The first few packets of this flow will bypass NAT as seen by the access-list matching on the destination router.

```
source#ping 198.51.100.9 source lo1 rep 2

Type escape sequence to abort.
Sending 2, 100-byte ICMP Echos to 198.51.100.9, timeout is 2 seconds:
Packet sent with a source address of 172.17.250.201
!!
Success rate is 100 percent (2/2), round-trip min/avg/max = 1/1/1 ms
source#ping 198.51.100.9 source lo1 rep 2

Type escape sequence to abort.
Sending 2, 100-byte ICMP Echos to 198.51.100.9, timeout is 2 seconds:
Packet sent with a source address of 172.17.250.201
!!
Success rate is 100 percent (2/2), round-trip min/avg/max = 1/1/1 ms
source#ping 198.51.100.11 source lo1 rep 200000

Type escape sequence to abort.
Sending 200000, 100-byte ICMP Echos to 198.51.100.11, timeout is 2 seconds:
Packet sent with a source address of 172.17.250.201
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!.
Success rate is 99 percent (3007/3008), round-trip min/avg/max = 1/1/16 ms
```

```
source#ping 198.51.100.9 source lo1 rep 10

Type escape sequence to abort.
Sending 10, 100-byte ICMP Echos to 198.51.100.9, timeout is 2 seconds:
Packet sent with a source address of 172.17.250.201
...!!!!!!!
Success rate is 70 percent (7/10), round-trip min/avg/max = 1/1/1 ms
source#
```

The ACL match on the destination router shows the three packets that failed were not translated:

```
Router2#show access-list 199
Extended IP access list 199
   10 permit udp host 172.17.250.201 host 198.51.100.9
   20 permit udp host 172.17.250.201 host 10.212.26.73
   30 permit udp host 203.0.113.231 host 198.51.100.9
   40 permit udp host 203.0.113.231 host 10.212.26.73 (4 matches)
   50 permit icmp host 172.17.250.201 host 198.51.100.9
   60 permit icmp host 172.17.250.201 host 10.212.26.73
    70 permit icmp host 203.0.113.231 host 198.51.100.9 (3 matches) <<<<<<<
   80 permit icmp host 203.0.113.231 host 10.212.26.73 (42 matches)
   90 permit udp any any log (2 matches)
   100 permit icmp any any log (4193 matches)
   110 permit ip any any (5 matches)
Router2#
```

On the ASR1K you can check the gatekeeper cache entries:

```
PRIMARY#show platform hardware qfp active feature nat datapath gatein
Gatekeeper on
sip 203.0.113.231 vrf 0 cnt 1 ts 0x17ba3f idx 74
sip 10.203.249.226 vrf 0 cnt 0 ts 0x36bab6 idx 218
sip 10.203.249.221 vrf 0 cnt 1 ts 0x367ab4 idx 229

PRIMARY#show platform hardware qfp active feature nat datapath gateout
Gatekeeper on
sip 198.51.100.11 vrf 0 cnt 1 ts 0x36db07 idx 60
sip 10.203.249.225 vrf 0 cnt 0 ts 0x36bb7a idx 217
sip 10.203.249.222 vrf 0 cnt 1 ts 0x367b7c idx 230
```

Workaround/Fix

In most environments the NAT gatekeeper functionality works fine and does not cause issues. However, if you do run into this problem there are a few ways to resolve it.

Solution 1

The preferred option would be to upgrade Cisco IOS® XE to a version that includes the gatekeeper enhancement:

Cisco bug ID CSCun06260 XE3.13 Gatekeeper Hardening

This enhancement allows for NAT gatekeeper to cache the source **and** the destination addresses, as well as makes the cache size configurable. In order to turn on the extended mode, you need to increase the cache size with these commands. You can also monitor the cache to see if you need to increase the size.

```
PRIMARY(config)#ip nat settings gatekeeper-size 1024
PRIMARY(config)#end
```

Extended mode can be verified by checking these commands:

```
PRIMARY#show platform hardware qfp active feature nat datapath gatein
Gatekeeper on
sip 10.203.249.221 dip 10.203.249.222 vrf 0 ts 0x5c437 idx 631

PRIMARY#show platform hardware qfp active feature nat datapath gateout
Gatekeeper on
sip 10.203.249.225 dip 10.203.249.226 vrf 0 ts 0x5eddf idx 631

PRIMARY#show platform hardware qfp active feature nat datapath gatein active
Gatekeeper on
ext mode Size 1024, Hits 2, Miss 4, Aged 0 Added 4 Active 1

PRIMARY#show platform hardware qfp active feature nat datapath gateout active
Gatekeeper on
ext mode Size 1024, Hits 0, Miss 1, Aged 1 Added 2 Active 0
```

Solution 2

For releases that do not have the fix for Cisco bug ID CSCun06260, the only option is to turn off the gatekeeper feature. The only negative impact will be slightly reduced performance for non-NAT-ed traffic as well as higher CPU utilization on the Quantum Flow Processor (QFP).

```
PRIMARY(config)#no ip nat service gatekeeper
PRIMARY(config)#end
PRIMARY#PRIMARY#Sh platform hardware qfp active feature nat datapath gatein
Gatekeeper off

PRIMARY#
```

QFP utilization can be monitored with these commands:

```
show platform hardware qfp active data utilization summary
show platform hardware qfp active data utilization qfp 0
```

Solution 3

Separate traffic flows so that NAT and non-NAT packets do not arrive on the same interface.

Summary

The NAT Gatekeeper command was introduced in order to enhance performance of the router for non-NAT-ed flows. Under some conditions the feature might cause problems when a mix of NAT and non-NAT packets arrive from the same source. The solution is to use the enhanced gatekeeper functionality, or if that is not possible, disable the gatekeeper feature.

References

Software changes which allowed gatekeeper to be turned off:
Cisco bug ID CSCty67184 ASR1k NAT CLI - Gatekeeper On/Off
Cisco bug ID CSCth23984 Add cli capability to turn on/off nat gatekeeper functionality

NAT Gatekeeper enhancement
Cisco bug ID CSCun06260 XE3.13 Gatekeeper Hardening