瞭解Nexus 9300上的NAT

目錄

簡介

本文檔介紹在配備運行NX-OS軟體的Cisco Cloud-Scale ASIC的Nexus 9000交換機上的NAT功能。

必要條件

需求

思科建議您先熟悉Cisco Nexus作業系統(NX-OS)和基本Nexus架構，然後再繼續使用本文檔中介紹的資訊。

採用元件

本文中的資訊係根據以下軟體和硬體版本：

- N9K-C93180YC-FX3

- nxos64-cs.10.4.3.F

本文中的資訊是根據特定實驗室環境內的裝置所建立。文中使用到的所有裝置皆從已清除（預設）的組態來啟動。如果您的網路運作中，請確保您瞭解任何指令可能造成的影響。

在N9K上引入NAT支援

技術

- NAT - NAT是一種用於聯網的技術，用於修改IP資料包的源或目標IP地址。
- PAT -埠地址轉換，也稱為「NAT過載」，多個內部IP地址共用一個外部IP地址，以唯一埠號區分。
- TCP感知NAT - TCP感知NAT支援使NAT流條目與TCP會話的狀態匹配，並相應地建立和刪除。

NAT TCAM資源

預設情況下，不會為Nexus 9000上的NAT功能分配任何TCAM條目。您必須透過減小其他功能的TCAM大小來分配NAT功能的TCAM大小。

NAT操作涉及三種型別的TCAM：

- NAT區域

NAT使用TCAM NAT區域進行基於IP地址或埠的資料包匹配。

內部或外部源地址的每個NAT/PAT條目都需要兩個NAT TCAM條目。

預設情況下，啟用ACL原子更新模式，支援60%的非原子級數。

- TCP感知區域

對於具有「x」 ace的每個NAT內部策略，需要「x」個條目。

每個配置的NAT池都需要一個條目。

啟用原子更新模式時，TCP-NAT TCAM大小必須加倍。

- NAT重寫表

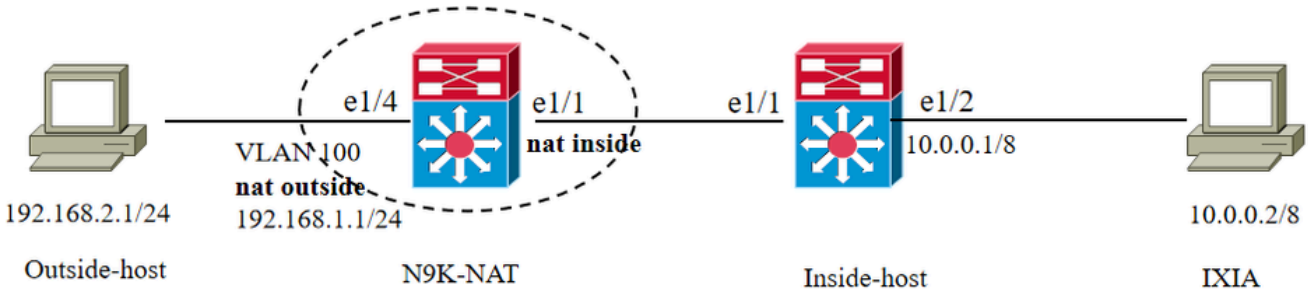
NAT 重寫 和 翻譯 是 已儲存 在 其 "NAT 重寫 表格、" 哪個 存在 outside 的 其 NAT TCAM 地區。其 'NAT 重寫 表格' 有 答 已修正 大小 的 2048 專案 用於 思科 9300-EX/FX/FX2/9300C 和 4096 專案 用於 思科 9300-FX3/GX/GX2A/GX2B/H2R/H1。 此 表格 是 獨佔 已使用 用於 NAT 翻譯。

內部或外部源地址的每個靜態NAT/PAT條目都需要一個「NAT重寫表」條目。

有關Nexus 9000上的TCAM的詳細資訊，您可以參考 《使用Cisco CloudScale ASIC對Nexus 9000系列交換機進行分類TCAM白皮書》。

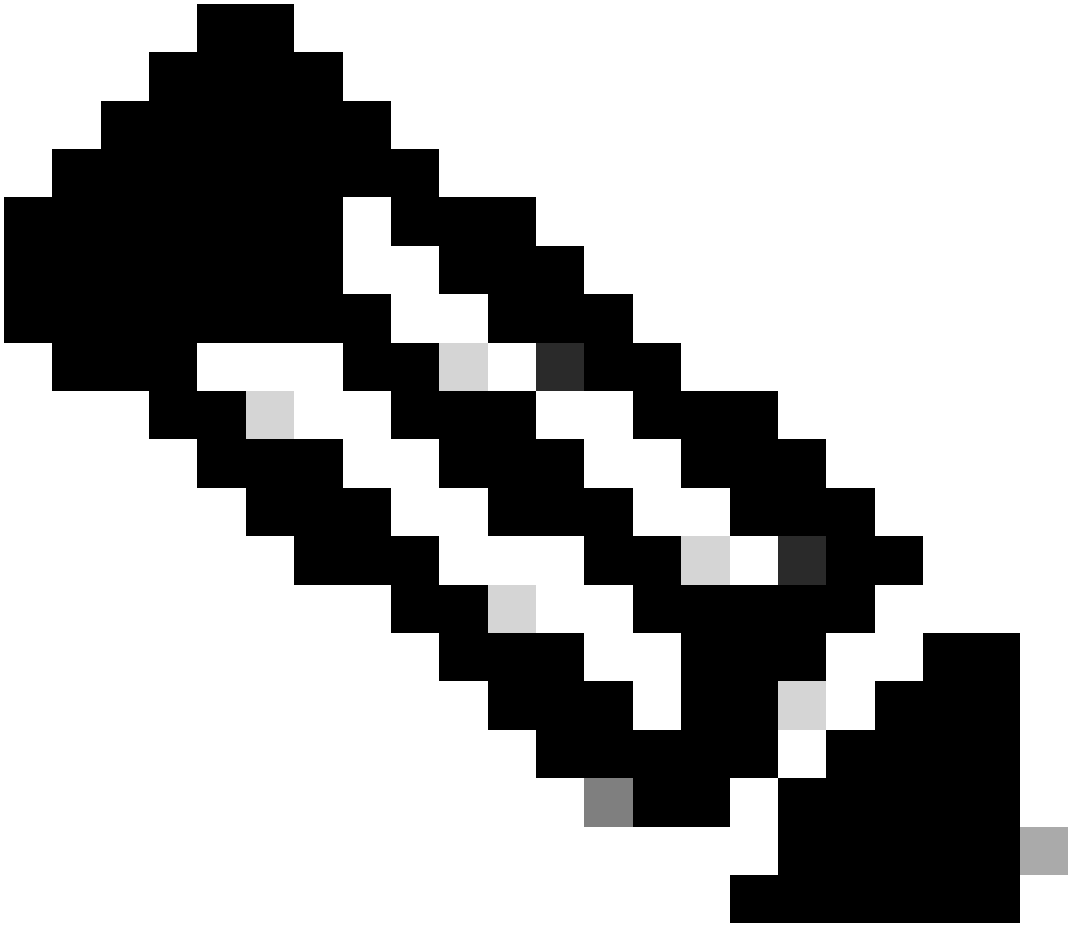
組態和驗證

拓撲



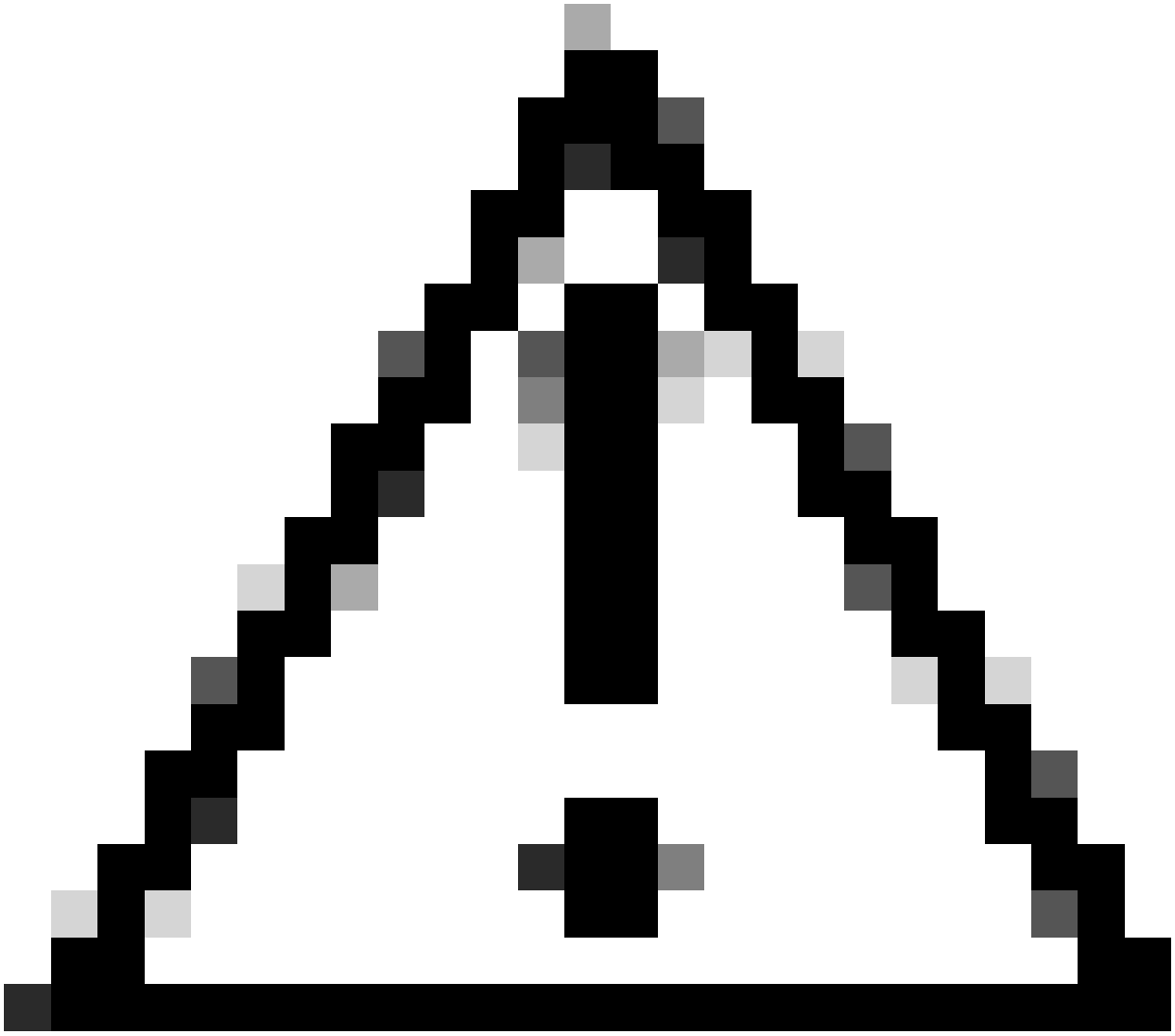
N9K-NAT配置

hardware access-list tcam region nat 1024 hardware access-list tcam region tcp-nat 100 ip nat translation max-entries 80



注意：預設情況下，動態nat轉換max-entries為80。

ip access-list TEST-NAT 10 permit ip 10.0.0.1/8 192.168.2.1/24 ip nat pool TEST 192.168.1.10 192.168.1.10 netmask 255.255.255.0 ip nat inside source list TEST-NAT pool TEST overload



注意：Cisco Nexus 9200、9300-EX、9300-FX 9300-FX2、9300-FX3、9300-FXP和9300-GX平台交換機上不支援用於內部和外部策略的介面過載選項選項

interface Vlan100 no shutdown ip address 192.168.1.1/24 ip nat outside

interface Vlan100 no shutdown ip address 192.168.1.1/24 ip nat outside

驗證

主機內部Ping

資料包的源IP：10.0.0.1轉換為IP：192.168.1.10

目的IP：192.168.2.1

Inside-host# ping 192.168.2.1 source 10.0.0.1 PING 192.168.2.1 (192.168.2.1): 56 data bytes 64 bytes from 192.168.2.1: icmp_seq=0 ttl=63 time=0.784 ms 64 bytes from 192.168.2.1: icmp_seq=1 ttl=63 time=0.595 m

NAT轉換表檢查

N9K-NAT# show ip nat translations icmp 192.168.1.10:60538 10.0.0.1:48940 192.168.2.1:0 192.168.2.1:0 icmp 192.168.1.10:60539 10.0.0.1:0 192.168.2.1:0 192.168.2.1:0

NAT統計資訊

N9K-NAT# show ip nat statistics IP NAT Statistics == Stats Collected since: Tue Sep 3 14:33:01 2024 ---------- Total active translations: 82 / Number of translations active in the system. This number is incremented each time a translation is created and is decremented each time a translation is cleared or times out. No.Static: 0 / Total number of static translations present in the system. No.Dyn: 82 / Total number of dynamic translations present in the system. No.Dyn-ICMP: 2 ---------- Total expired Translations: 2 SYN timer expired: 0 FIN-RST timer expired: 0 Inactive timer expired: 2 ---------- Total Hits: 10475 / Total number of times the software does a translations table lookup and finds an entry. Total Misses: 184884 / Total number of packet the software dropped Packet. In-Out Hits: 10474 In-Out Misses: 184884 Out-In Hits: 1 Out-In Misses: 0 ---------- ---------- Total SW Translated Packets: 10559 / Total number of packets software does the translation. In-Out SW Translated: 10558 Out-In SW Translated: 1 ---------- Total SW Dropped Packets: 184800 / Total number of packet the software dropped Packet. In-Out SW Dropped: 184800 Out-In SW Dropped: 0 Address alloc. failure drop: 0 Port alloc. failure drop: 0 Dyn. Translation max limit drop: 184800 / Total number of packets dropped due to configured maximum number of dynamic translation entry limit reached. (ip nat translation max-entries <1-1023>) ICMP max limit drop: 0 Allhost max limit drop: 0 -------- ---------- Total TCP session established: 0 Total TCP session closed: 0 ---------- NAT Inside Interfaces: 1 Ethernet1/1 NAT Outside Interfaces: 1 Vlan100 ---------- Inside source list: ++++++++++++++++++++ Access list: TEST-NAT RefCount: 82 / Number of current references to this access list. Pool: TEST Overload Total addresses: 1 / Number of addresses in the pool available for translation. Allocated: 1 percentage: 100% Missed: 0

常見問題

當NAT TCAM用盡時，會發生什麼情況？

如果TCAM資源用盡，則會報告錯誤日誌。

2024 Aug 28 13:26:56 N9K-NAT %ACLQOS-SLOT1-2-ACLQOS_OOTR: Tcam resource exhausted: Feature NAT outside [nat-outside] 2024 Aug 28 13:26:56 N9K-NAT %NAT-2-HW_PROG_FAILED: Hardware programming for NAT failed:Sufficient free entries are not available in TCAM bank(3)

達到Max-entries時會發生什麼？

預設情況下，NAT轉換max-entries為80。 一旦動態NAT轉換條目超過最大限制，流量將被傳送到CPU，從而導致錯誤日誌和丟棄。

Ping test failure: Inside-host# ping 192.168.2.1 source 10.0.0.1 count unlimited interval 1 PING 192.168.2.1 (192.168.2.1): 56 data bytes Request 0 timed out N9K-NAT Error log: 2024 Sep 5 15:31:33 N9K-NAT %NETSTACK-2-NAT_MAX_LIMIT: netstack [15386] NAT: Can't create dynamic translations, max limit reached - src:10.0.0.1 dst:192.168.2.1 sport:110 dport:110 Capture file from CPU: N9K-NAT# ethanalyzer local interface inband limit-captured-frames 0 Capturing on 'ps-inb' 15 2024-09-05 15:32:44.899885527 10.0.0.1 → 192.168.2.1 UDP 60 110 → 110 Len=18

為什麼某些NAT資料包被傳送到CPU？

通常，流量路由到CPU有兩種情況。

第一種情況發生在尚未將NAT條目程式設計到硬體時，此時流量需要由CPU處理。

頻繁的硬體程式設計會給CPU帶來壓力。為了減少硬體中NAT條目的程式設計頻率，NAT將轉換程式設計為一秒鐘。commandip nat translation creation-delay會延遲會話的建立。

第二個場景涉及在建立TCP會話的初始階段和終止會話互動期間傳送到CPU處理的資料包。

為什麼NAT在Nexus 9000上不使用代理ARP工作？

從9.2.X版中增加了一個稱為nat-alias的功能。此功能預設啟用，可解決NAT ARP問題。除非手動停用，否則不需要啟用ip proxy-arp或ip local-proxy-arp。

NAT裝置擁有內部全局(IG)和外部本地(OL)地址，並負責響應定向到這些地址的任何ARP請求。當IG/OL地址子網與本地介面子網匹配時，NAT會安裝IP別名和ARP條目。在這種情況下，裝置使用local-proxy-arp響應ARP請求。

如果地址範圍與外部介面位於同一子網中，則無別名功能將響應給定NAT池地址範圍內所有轉換後的IP的ARP請求。

add-route引數在N9K上的工作原理以及為什麼它是必需的？

在Cisco Nexus 9200和9300-EX、-FX、-FX2、-FX3、-FXP、-GX平台交換機上，由於ASIC硬體限制，內部和外部策略均需要add-route選項。使用此引數，N9K將增加一個主機路由。從外部到內部的TCP NAT流量被傳送到CPU，並且可以在沒有此引數的情況下丟棄。

之前：

192.168.1.0/24, ubest/mbest: 1/0, attached *via 192.168.1.1, Vlan100, [0/0], 10:23:08, direct 192.168.1.0/32, ubest/mbest: 1/0, attached *via 192.168.1.0, Null0, [0/0], 10:23:08, broadcast 192.168.1.1/32, ubest/mbest: 1/0, attached *via 192.168.1.1, Vlan100, [0/0],10:23:08, local

之後：

192.168.1.2/32, ubest/mbest: 1/0 *via 10.0.0.2, [1/0], 00:02:48, nat >>route created by NAT feature 10.0.0.2/32, ubest/mbest: 1/0 *via 192.168.100.2, [200/0], 06:06:58, bgp-64700, internal, tag 64710 192.168.1.0/24, ubest/mbest: 1/0, attached *via 192.168.1.1, Vlan100, [0/0], 20:43:08, direct

為什麼NAT最多支援100個ICMP條目

通常，ICMP NAT會在配置的sampling-timeout和translation-timeout到期後超時。但是，當交換機中存在的ICMP NAT流變為空閒時，在配置的sampling-timeout到期後立即超時。

從Cisco NX-OS版本7.0(3)I5(2)開始，在Cisco Nexus 9300平台交換機上引入了ICMP硬體程式設計。因此，ICMP條目會消耗硬體中的TCAM資源。由於ICMP在硬體中，Cisco Nexus平台系列交換機

中NAT轉換的最大限制更改為1024。最多允許100個ICMP條目以最佳方式使用資源。此命令是固定的，並且沒有選項可用於調整最大ICMP條目。

相關資訊

Cisco Nexus 9000 系列 NX-OS 介面組態指南（ 10.4(x) 版 ）

用於Nexus 9000系列交換機的Cisco CloudScale ASIC的分類TCAM白皮書

Cisco Nexus 9000系列NX-OS驗證可擴充性指南

關於此翻譯

思科已使用電腦和人工技術翻譯本文件，讓全世界的使用者能夠以自己的語言理解支援內容。請注意，即使是最佳機器翻譯，也不如專業譯者翻譯的內容準確。Cisco Systems, Inc. 對這些翻譯的準確度概不負責，並建議一律查看原始英文文件（提供連結）。