了解Nexus 9300上的NAT

目录

简介

本文档介绍配备运行NX-OS软件的思科云规模ASIC的Nexus 9000交换机上的NAT功能。

先决条件

要求

思科建议您先熟悉Cisco Nexus操作系统(NX-OS)和基本Nexus架构，然后再继续阅读本文档中介绍的信息。

使用的组件

本文档中的信息基于以下软件和硬件版本：

- N9K-C93180YC-FX3

- nxos64-cs.10.4.3.F

本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

介绍N9K上的NAT支持

术语

- NAT - NAT是网络中用于修改IP数据包的源或目标IP地址的一种技术。
- PAT -端口地址转换，也称为“NAT过载”，多个内部IP地址共享一个外部IP地址，用唯一的端口号加以区分。
- TCP感知NAT - TCP感知NAT支持使NAT流条目能够匹配TCP会话的状态，并相应地创建和删除。

NAT TCAM资源

默认情况下，不会为Nexus 9000上的NAT功能分配TCAM条目。您必须通过减小其他功能的TCAM大小来为NAT功能分配TCAM大小。

NAT操作涉及三种类型的TCAM：

- NAT区域

NAT利用TCAM NAT区域进行基于IP地址或端口的数据包匹配。

内部或外部源地址的每个NAT/PAT条目需要两个NAT TCAM条目。

默认情况下，启用ACL原子更新模式，支持60%的非原子级位数。

- TCP感知区域

对于具有“x”个ace的每个NAT内部策略，需要“x”个条目。

对于每个配置的NAT池，需要一个条目。

启用原子更新模式时，TCP-NAT TCAM大小必须加倍。

- NAT重写表

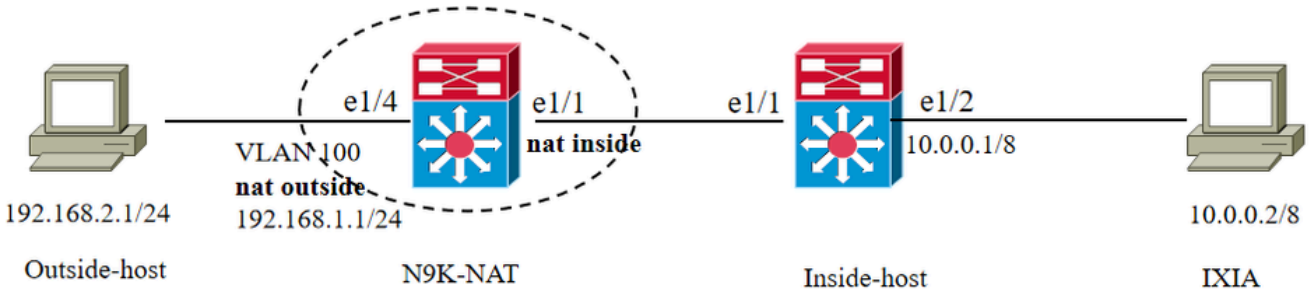
NAT 重写 和 转换 是 已存储 在 此 "NAT 重写 表、" 哪个 存在 外部 / 此 NAT TCAM 区域。 此 'NAT 重写 表' 有 a fixed（已修复） 大小 / 2048 条目 对于 Nexus 9300-EX/FX/FX2/9300C 和 4096 条目 对于 Nexus 9300-FX3/GX/GX2A/GX2B/H2R/H1。 此 表 is 独占 已使用 对于 NAT 翻译。

内部或外部源地址的每个静态NAT/PAT条目都需要一个“NAT重写表”条目。

有关Nexus 9000上的TCAM的更多详情，请参阅 《Classification TCAM with Cisco CloudScale ASICs for Nexus 9000 Series Switches》白皮书。

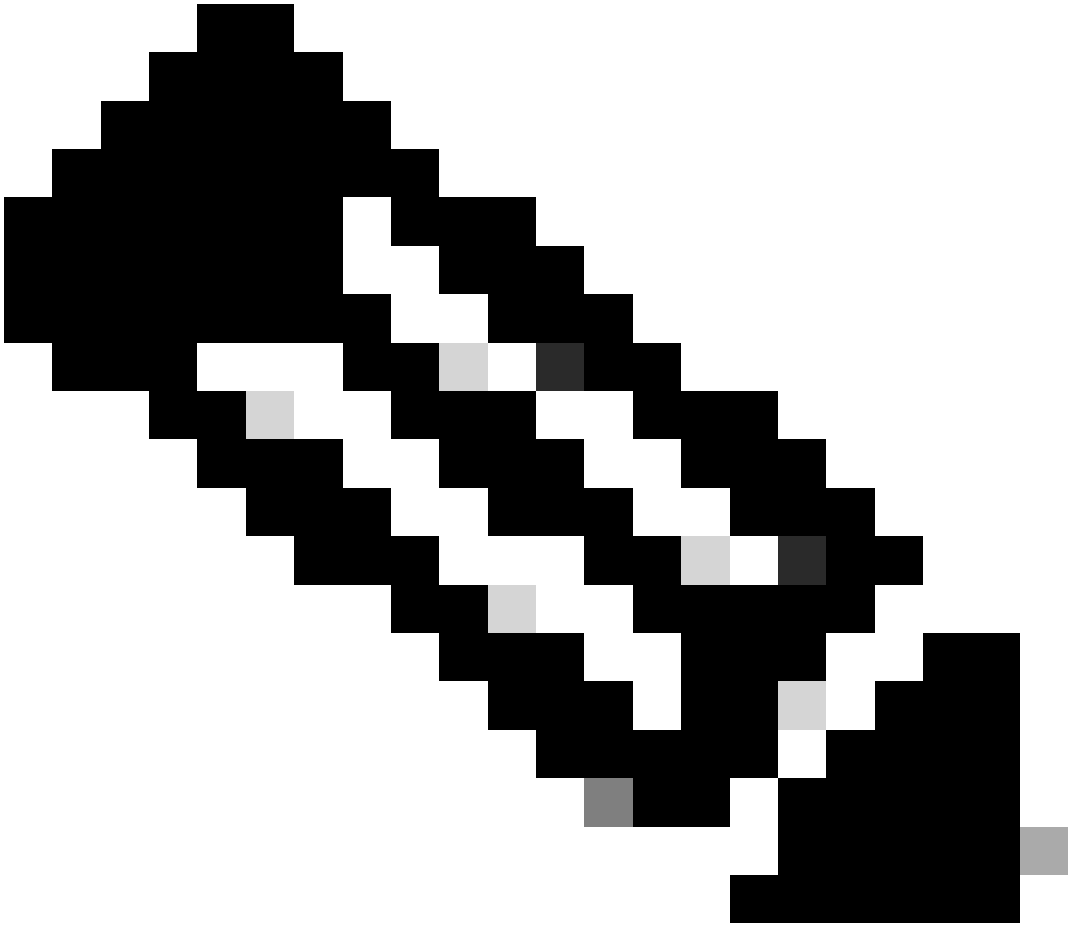
配置和验证

拓扑



N9K-NAT配置

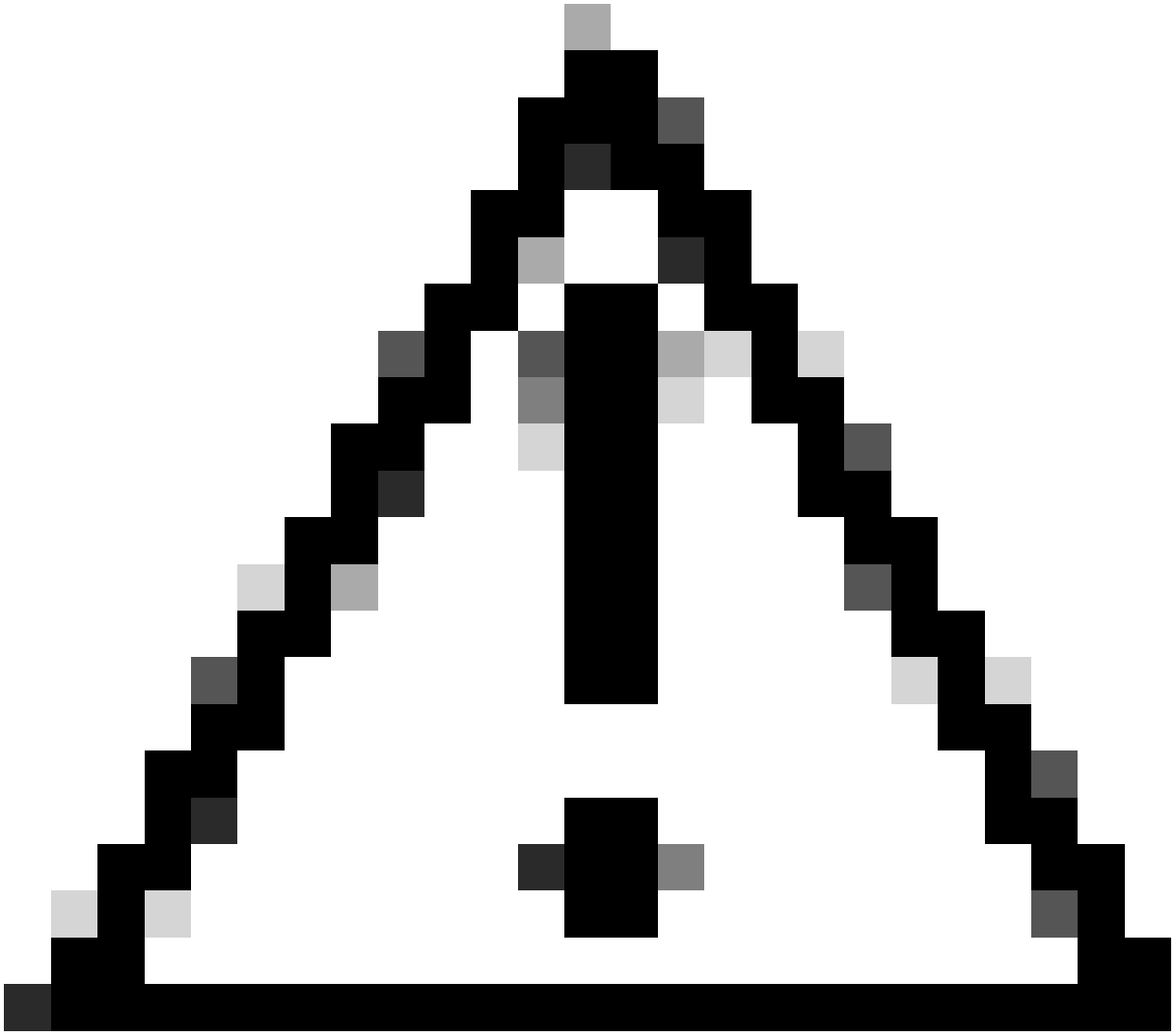
hardware access-list tcam region nat 1024 hardware access-list tcam region tcp-nat 100 ip nat translation max-entries 80



注意：默认情况下，动态nat转换max-entries为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平台交换机上不支持用于内部和外部策略的接口过载选项

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

确认

内部主机Ping

数据包的源IP：10.0.0.1转换为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转换表检查

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 统计信息

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

常见问题解答

NAT TCAM耗尽时会发生什么情况？

如果TCAM资源耗尽，则会报告错误日志。

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)

达到Max-entries时会发生什么？

默认情况下，NAT转换max-entries为80。 一旦动态NAT转换条目超过最大限制，流量将被传送到CPU，从而导致错误记录和丢弃。

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

为什么有些NAT数据包被传送到CPU？

通常，有两种情况会将流量路由到CPU。

当NAT条目尚未编程到硬件时，会首先发生这种情况，此时需要由CPU处理流量。

频繁的硬件编程会给CPU带来压力。为了减少硬件中NAT条目的编程频率，NAT将转换编程为一秒批。命令dip nat translation creation-delay延迟会话建立。

第二个场景涉及在建立TCP会话的初始阶段以及在该会话的终止交互期间发送到CPU进行处理的数据包。

为什么NAT在Nexus 9000上不使用proxy-arp也能运行？

从版本9.2.X中添加了称为nat-alias的功能。此功能默认启用并解决NAT ARP问题。除非手动禁用，否则您不需要启用ip proxy-arp或ip local-proxy-arp。

NAT设备拥有内部全局(IG)和外部本地(OL)地址，并负责响应定向到这些地址的任何ARP请求。当IG/OL地址子网与本地接口子网匹配时，NAT会安装一个IP别名和一个ARP条目。在这种情况下，设备使用local-proxy-arp响应ARP请求。

如果地址范围与外部接口位于同一子网中，则无别名功能会响应来自给定NAT池地址范围的所有转换IP的ARP请求。

add-route参数如何在N9K上运行，为什么它是必需的？

在Cisco Nexus 9200和9300-EX、-FX、-FX2、-FX3、-FXP、-GX平台交换机上，由于ASIC硬件限制，内部和外部策略均需要add-route选项。使用此参数，N9K将添加一个主机路由。从外部到内部的TCP NAT流量被传送到CPU，并且可以在没有此参数的情况下丢弃。

攻击前：

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

在:

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

为什么NAT最多支持100个ICMP条目

通常，ICMP NAT在配置的sampling-timeout和translation-timeout过期后流超时。但是，当交换机中存在的ICMP NAT流变为空闲时，它们会在配置的采样超时到期后立即超时。

从Cisco NX-OS版本7.0(3)I5(2)开始，Cisco Nexus 9300平台交换机上引入了ICMP硬件编程。因此

，ICMP条目会占用硬件中的TCAM资源。由于ICMP在硬件中，因此Cisco Nexus平台系列交换机中NAT转换的最大限制更改为1024。最多允许100个ICMP条目充分利用资源。它是固定的，并且没有调整最大ICMP条目的选项。

相关信息

Cisco Nexus 9000 系列 NX-OS 接口配置指南 10.4(x) 版

《Nexus 9000系列交换机使用Cisco CloudScale ASIC的分类TCAM白皮书》

Cisco Nexus 9000系列NX-OS验证可扩展性指南

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。