Probleemoplossing voor multisite VXLAN met CloudSec in vierkante topologie

Inhoud

Inleiding

Dit document beschrijft de configuratie en probleemoplossing van VXLAN op meerdere locaties met CloudSec tussen grensgateways die in een vierkante topologie zijn verbonden.

Voorwaarden

Vereisten

Cisco raadt u aan bekend te zijn met deze onderwerpen:

- Nexus NXOS © software.
- VXLAN EVPN-technologie.
- BGP- en OSPF-routeringsprotocollen.

Gebruikte componenten

De informatie in dit document is gebaseerd op de volgende software- en hardwareversies:

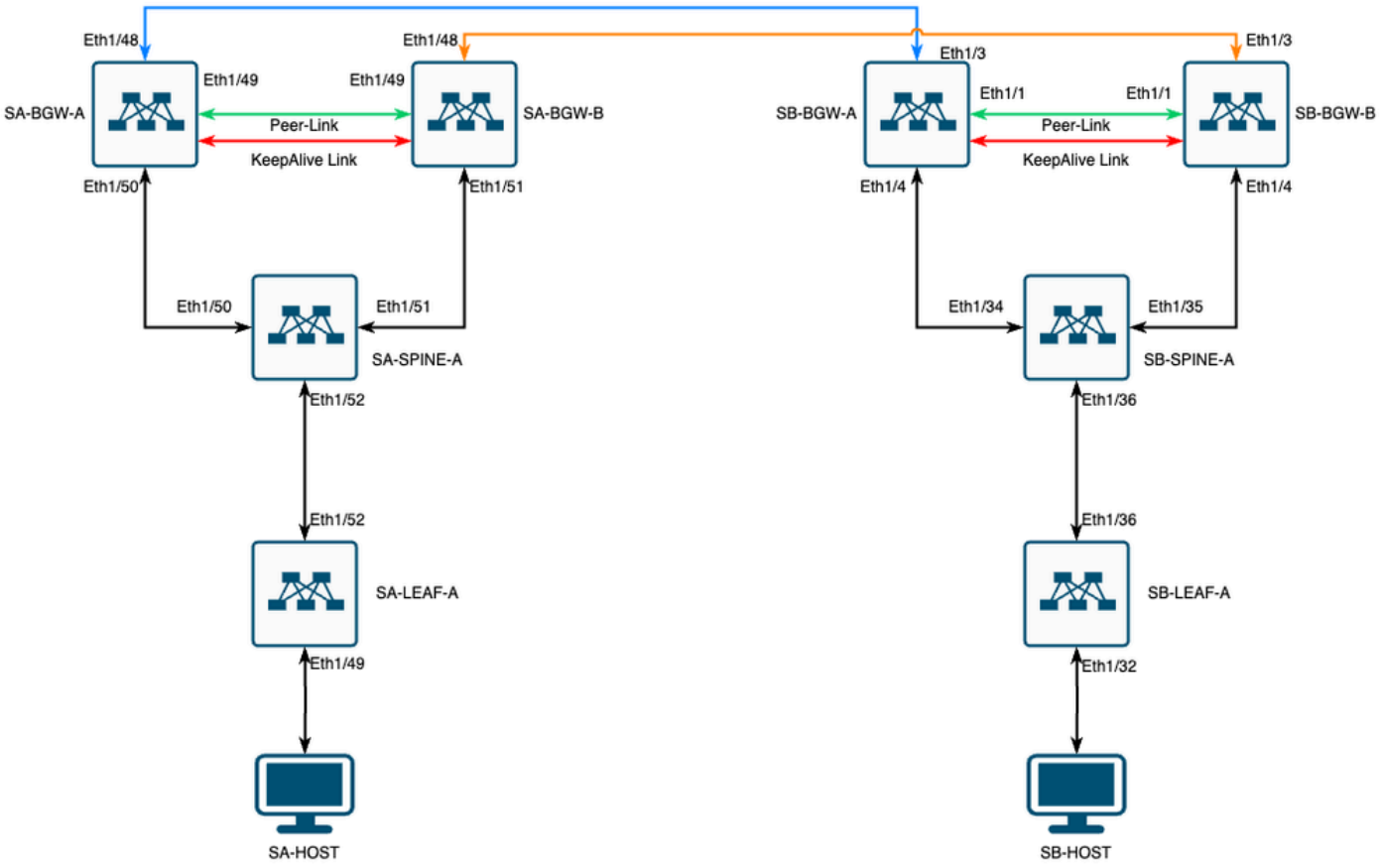
- Cisco Nexus 9000 switch

- NXOS versie 10.3(4a).

De informatie in dit document is gebaseerd op de apparaten in een specifieke laboratoriumomgeving. Alle apparaten die in dit document worden beschreven, hadden een opgeschoonde (standaard)configuratie. Als uw netwerk live is, moet u zorgen dat u de potentiële impact van elke opdracht begrijpt.

Configureren

Netwerkdiagram



VXLAN MultiSite met CloudSec in vierkante topologie

Details van de topologie

- Twee-site multisite VXLAN VPN-fabric.
- Beide locaties zijn geconfigureerd met vPC Border Gateways.
- De endpoints worden gehost in VLAN 1100.
- Grensgateways op elke site hebben IPv4 iBGP-buurten tussen elkaar via de SVI-interface Vlan3600.
- Grensgateways op de ene site hebben alleen een eBGP IPv4-buren met direct verbonden border-gateway op de andere site.
- Grensgateways op site A hebben een eBGP L2VPN EVPN-buurt met grensgateways op site B.

Geadresseerdplan

De IP-adressen in de tabel worden tijdens de configuratie gebruikt:

	SITE A	SITE B				
Apparaatrol	Interface-ID	Fysieke Intel IP	RID-loop IP	NVE-lus-IP	MSITE-VIP	Ba
BLAD	Eth1/52	192.168.1.1/30	192.168.2.1/32	192.168.3.1/32	N.v.t.	
RUGGENGRAAT	Eth1/52	192.168.1.2/30			N.v.t.	
Eth1/50	192.168.1.5/30	192.168.2.2/32	N.v.t.	N.v.t.	N.v.t.	
Eth1/51	192.168.1.9/30			N.v.t.		
BGW-A	Eth1/51	192.168.1.6/30	192.168.2.3/32	192.168.3.2/32	192.168.100.1/32	192
Eth1/48	10.12.10.1/30		192.168.3.254/32			
BGW-B	Eth1/51	192.168.1.10/30	192.168.2.4/32	192.168.3.3/32	192.168.100.1/32	192
Eth1/48	10.12.10.5/30		192.168.3.254/32			

Configuraties

- Merk op dat in deze handleiding alleen multisite-gerelateerde configuratie wordt getoond. Voor de volledige configuratie kunt u de officiële documentatiegids van Cisco voor VXLAN Cisco Nexus 9000 Series NX-OS VXLAN-configuratiehandleiding, release 10.3(x) gebruiken

Om CloudSec mogelijk te maken moet de dci-advertise-pip opdracht worden geconfigureerd onder de evpn multisite border-gateway:

SA-BGW-A en SA-BGW-B	SB-BGW-A en SB-BGW-B
evpn multisite border-gateway 65001 dci-advertise-pip	evpn multisite border-gateway 65002 dci-advertise-pip

BGP-configuratie

Deze configuratie is sitespecifiek.

SA-BGW-A en SA-BGW-B	SB-BGW-A en SB-BGW-B
router bgp 65001 address-family ipv4 unicast maximum-paths 64 address-family l2vpn evpn maximum-paths 64 additional-paths send additional-paths receive	router bgp 65002 address-family ipv4 unicast maximum-paths 64 address-family l2vpn evpn maximum-paths 64 additional-paths send additional-paths receive

- Met de opdracht **maximum-path** kunnen meerdere eBGP L2VPN EVPN-paden van de buur worden ontvangen.

- De opdracht **Extra-Path** draagt het BGP-proces op om te adverteren dat het apparaat in staat is om extra paden te verzenden/ontvangen

Voor alle L3VNI VRF's op grensgateways moet multipath ook worden geconfigureerd:

SA-BGW-A en SA-BGW-B	SB-BGW-A en SB-BGW-B
router bgp 65001 vrf tenant-1 address-family ipv4 unicast maximum-paths 64 address-family ipv6 unicast maximum-paths 64	router bgp 65002 vrf tenant-1 address-family ipv4 unicast maximum-paths 64 address-family ipv6 unicast maximum-paths 64

Tunnel-encryptie configuratie

Deze configuratie moet hetzelfde zijn op alle grensgateways:

key chain CloudSec_Key_Chain1 tunnel-encryption key 1000 key-octet-string Cl0udSec! cryptographic-algorithm AES_128_CMAC feature tunnel-encryp

Deze configuratie is sitespecifiek. tunnel-encryption De opdracht moet alleen worden toegepast op de interface die de evpn multisite dci-trackingopdracht heeft.

SA-BGW-A en SA-BGW-B	SB-BGW-A en SB-BGW-B
tunnel-encryption peer-ip 192.168.13.2 keychain CloudSec_Key_Chain1 policy CloudSec_Policy1 tunnel-encryption peer-ip 192.168.13.3 keychain CloudSec_Key_Chain1 policy CloudSec_Policy1 interface Ethernet1/48 tunnel-encryption	tunnel-encryption peer-ip 192.168.3.2 keychain CloudSec_Key_Chain1 policy CloudSec_Policy1 tunnel-encryption peer-ip 192.168.3.3 keychain CloudSec_Key_Chain1 policy CloudSec_Policy1 interface Ethernet1/3 tunnel-encryption

Na het inschakelen van de tunnelencryptie worden extra attributen toegevoegd aan de lokale loopback terwijl advertentieroutes naar de buurman en alle eBGP IPv4 unicast-buren deze eigenschap moeten zien:

<#root>

SA-BGW-A# show ip bgp 192.168.2.3 BGP routing table information for VRF default, address family IPv4 Unicast BGP routing table entry for 192.168.2

!---

This is a new attribute

```
 Path type: redist, path is valid, not best reason: Locally originated, no labeled nexthop AS-Path: NON
```

Voor Route Type-2 is er ook een nieuw kenmerk:

<#root>

SA-BGW-A# show bgp l2vpn evpn 00ea.bd27.86ef BGP routing table information for VRF default, address family L2VPN EVPN Route Distinguisher: 65

!---

Ethernet Segment Identifier (ESI) is also new attribute

```
 Path-id 1 (dual) advertised to peers: 192.168.2.2 SA-BGW-A#
```

Verifiëren

Alvorens cloudsec in te schakelen, is het goed om te controleren of de installatie werkt prima zonder:

SA-BGW-A(config)# show clock Warning: No NTP peer/server configured. Time may be out of sync. 10:02:01.016 UTC Fri Jul 19 2024 Time source is N

Na de configuratie van de cloudsec, moet endpoint op SA met succes pingelen het endpoint op site B. Maar in sommige gevallen kan ping niet succesvol zijn. Het is afhankelijk van welke cloudsec peer door het lokale apparaat wordt geselecteerd om cloudsec versleuteld verkeer te verzenden.

SA-HOST-A# ping 10.100.20.10 PING 10.100.20.10 (10.100.20.10): 56 data bytes Request 0 timed out Request 1 timed out Request 2 timed out Request 3

Problemen oplossen

Controleer de lokale ARP-tabel op het bron-endpoint:

SA-HOST-A# ping 10.100.20.10 count unlimited interval 1 Request 352 timed out Request 353 timed out Request 354 timed out 356 packets transmitted, 0

Deze output bewijst dat, het verkeer van BUM overgaat en controle-Vlak werkt. De volgende stap is het controleren van de tunnelcoderingsstatus:

SA-BGW-A# show tunnel-encryption session Tunnel-Encryption Peer Policy Keychain RxStatus TxStatus ---------- ----------

Deze output toont aan dat de CloudSec-sessie is ingesteld. Als volgende stap kunt u onbeperkt ping op SA-HOST-A uitvoeren:

SA-HOST-A# ping 10.100.20.10 count unlimited interval 1

Vanaf dit punt moet u de apparaten op site A controleren en zien of het verkeer deze apparaten bereikt. U kunt deze taak met ELAM op alle apparaten langs het pad op site A. Veranderend in-select van standaardwaarde van 6 tot 9 staat toe om aan te passen op basis van binnenkopballen. Lees meer over ELAM op deze link: Nexus 9000 Cloud Scale ASIC (Tahoe) NX-OS ELAM.

ELAM op SA-LEAF-A

In een productienetwerk zijn meer dan één spinapparaat aanwezig. Om te begrijpen naar welke ruggengraat het verkeer werd verzonden, moet je eerst een ELAM op LEAF nemen. Ondanks dat dat in-select 9 gebruikt, bij het LEAF verbonden met de bron, moet de router ipv4 header worden gebruikt, omdat het verkeer bereikt dit LEAF niet VXLAN versleuteld. In echt netwerk, kan het moeilijk zijn om het nauwkeurige pakket te vangen u produceerde. In dergelijke gevallen kunt u ping met specifieke lengte uitvoeren en de pakketdiagrammen gebruiken om uw pakket te identificeren. Standaard is het ICMP-pakket 64 bytes lang. Plus 20 bytes van IP-header, die in samenvatting gaf u 84 bytes PKT Len:

<#root>

SA-LEAF-A# debug platform internal tah elam SA-LEAF-A(TAH-elam)# trigger init in-select 9 Slot 1: param values: start asic 0, start slice 0, lu-a2d 1, in

!---Note dpid value

Dst Idx : 0xcd, Dst BD : 1100 Packet Type: IPv4 Outer Dst IPv4 address: 10.100.20.10 Outer Src IPv4 ad

Pkt len = 84

, Checksum = 0xb4ae

!---64 byte + 20 byte IP header Pkt len = 84

Inner Payload Type: CE L4 Protocol : 1 L4 info not available Drop Info: ---------- LUA: LUB: LUC: LUD:

!---

Put dpid value here

IF_STATIC_INFO: port_name=Ethernet1/52,if_index:0x1a006600,ltl=5940,slot=0, nxos_port=204,dmod=1,dpid=

Van deze output kunt u zien dat het verkeer SA-LEAF-A wordt bereikt en uit de interface Ethernet1/52 door:sturen, die met SA-SPINE-A van de topologie wordt verbonden.

ELAM op SA-SPINE-A

Op SPINE zal de waarde van de Pkt Len meer zijn, aangezien de 50 bytes VXLAN header ook toegevoegd. Standaard kan SPINE niet overeenkomen op interne kopregels zonder vxlan-parse of feature nv overlay . Dus moet u vxlan-parse enable opdracht op SPINE gebruiken:

<#root>

SA-SPINE-A(config-if)# debug platform internal tah elam SA-SPINE-A(TAH-elam)# trigger init in-select 9 Slot 1: param values: start asic 0, start slice 0,

!---

84 bytes + 50 bytes VXLAN header Pkt len = 134

Inner Payload Type: IPv4 Inner Dst IPv4 address: 10.100.20.10 Inner Src IPv4 address: 10.100.10.10 L4

SA-SPINE-A stuurt verkeer naar de SA-BGW-A volgens de output.

ELAM op SA-BGW-A

SA-BGW-A(TAH-elam-insel9)# set inner ipv4 src_ip 10.100.10.10 dst_ip 10.100.20.10 SA-BGW-A(TAH-elam-insel9)# start SA-BGW-A(TAH-elam-ins

Volgens de output van SA-BGW-A, ging het verkeer uit Ethernet1/48 naar SB-BGW-A. De volgende stap is het controleren van SB-BGW-A:

<#root>

SB-BGW-A# debug platform internal tah elam SB-BGW-A(TAH-elam)# trigger init in-select 9 Slot 1: param values: start asic 0, start slice 0, lu-a2d 1, in-

!---Reset the previous filter and start again just in case if packet was not captured.

SB-BGW-A(TAH-elam-insel9)# reset SB-BGW-A(TAH-elam-insel9)# set inner ipv4 src_ip 10.100.10.10 dst_ip

Volgens de gegevens van SB-BGW-A is ELAM niet eens geactiveerd. Dit betekent dat ofwel SB-BGW-B de pakketten ontvangt en ze niet correct kan ontcijferen en ontleden, of ze helemaal niet ontvangt. Om te begrijpen wat er met het cloudsec-verkeer is gebeurd, kunt u een ELAM op SB-BGW-A opnieuw uitvoeren, maar het trigger-filter moet worden ingesteld op het externe IP-adres dat voor cloudsec wordt gebruikt, omdat er geen manier is om de interne header van cloudsec versleuteld transit-pakket te zien. Van de vorige output weet je dat de SA-BGW-A het verkeer verwerkt, wat betekent dat SA-BGW-A verkeer versleutelt met cloudsec. Zo, kunt u NVE IP van SA-BGW-A als trekkerfilter voor ELAM gebruiken. Van de vorige uitgangen is de VXLAN versleutelde ICMP-pakketlengte 134 bytes. Plus 32 byte cloudsec header in samenvatting geeft u 166 bytes:

<#root>

SB-BGW-A(TAH-elam-insel9)# reset SB-BGW-A(TAH-elam-insel9)# set outer ipv4 src_ip 192.168.3.2 SB-BGW-A(TAH-elam-insel9)# start SB-BGW-

192.168.13.3 ***!---NVE IP address of SB-BGW-B***

Outer Src IPv4 address: 192.168.3.2 Ver = 4, DSCP = 0, Don't Fragment = 0 Proto = 17, TTL = 254, More

!---134 byte VXLAN packet + 32 byte cloudsec header Pkt len = 166

Inner Payload Type: CE L4 Protocol : 17 L4 info not available Drop Info: ---------- LUA: LUB: LUC: LUD

!---To reach SB-BGW-B NVE IP traffic was sent out of Ethernet1/4 which is connected to SB-SPINE-A

SB-BGW-A(TAH-elam-insel9)# show system internal ethpm info all | i i "dpid=130" IF_STATIC_INFO: port_n

SB-BGW-A(TAH-elam-insel9)# show cdp neighbors interface ethernet 1/4 Capability Codes: R - Router, T - Trans-Bridge, B - Source-Route-Bridge S - S

192.168.13.3/32

, ubest/mbest: 1/0 *via 192.168.11.5,

Eth1/4

, [110/6], 00:56:13, ospf-UNDERLAY, intra via

192.168.14.2

, [200/0], 01:13:46, bgp-65002, internal, tag 65002

!---The device still have a route for SB-BGW-B NVE IP via SVI

SB-BGW-A(TAH-elam-insel9)# show ip route 192.168.14.2 IP Route Table for VRF "default" '*' denotes best

***via 192.168.14.2, Vlan3600**

, [250/0], 01:15:05, am SB-BGW-A(TAH-elam-insel9)# show ip arp 192.168.14.2 Flags: * - Adjacencies lear

ecce.1324.c803

Vlan3600

SB-BGW-A(TAH-elam-insel9)# show mac address-table address ecce.1324.c803 Legend: * - primary entry, G

3600

ecce.1324.c803

static - F F

vPC Peer-Link(R)

SB-BGW-A(TAH-elam-insel9)#

Van deze output, kunt u zien, dat het cloudseconverkeer naar SB-BGW-B via de interface Ethernet1/4 door:sturen, die op de routeringstabel wordt gebaseerd. Volgens de configuratiehandleiding voor Cisco Nexus 9000 Series NX-OS VXLAN, release 10.3(x) gidsen en beperkingen:

- CloudSec-verkeer dat voor de switch is bestemd, moet de switch via de DCI-uplinks invoeren.

Volgens de vPC Border Gateway Support for Cloudsec sectie van dezelfde handleiding zullen de BGP-padkenmerken van zowel vPC BGW als vPC BGW aan DCI-kant hetzelfde zijn als wanneer vPC BGW het IP-adres van peer-vPC BGW leert en adverteert. Daarom kunnen de DCI tussenknooppunten uiteindelijk de weg kiezen van vPC BGW die niet het PIP-adres bezit. In dit scenario wordt de MCT-link gebruikt voor versleuteld verkeer vanaf de externe site. Maar in dit geval wordt de interface naar de RUGGENGRAAT gebruikt, ondanks dat, hebben BGWs ook een OSPF nabijheid via BackUp SVI.

SB-BGW-A(TAH-elam-insel9)# show ip ospf neighbors OSPF Process ID UNDERLAY VRF default Total number of neighbors: 2 Neighbor ID Pri State

Reden van het probleem en oplossing

De reden is de OSPF-kosten van de SVI-interface. Standaard is op NXOS de bandbreedte van de automatische kostenreferentie 40G. SVI-interfaces hebben bandbreedte van 1 Gbps, terwijl de fysieke interface een bandbreedte van 10 Gbps heeft:

<#root>

SB-BGW-A(TAH-elam-insel9)# show ip ospf interface brief OSPF Process ID UNDERLAY VRF default Total number of interface: 5 Interface ID Area C

<Output omitted>

Eth1/4 5 0.0.0.0 1 P2P 1 up

In een dergelijk geval kan de administratieve wijziging van de kosten voor SVI het probleem oplossen. De afstemming moet op alle grensgateways gebeuren.

<#root>

SB-BGW-A(config)# int vlan 3600 SB-BGW-A(config-if)# ip ospf cost 1 SB-BGW-A(config-if)# sh ip route 192.168.13.3 IP Route Table for VRF "defau

via 192.168.14.2

, Vlan3600, [110/2], 00:00:08, ospf-UNDERLAY, intra via 192.168.14.2, [200/0], 01:34:07, bgp-65002, int

!---The ping is started to work immediately

Request 1204 timed out Request 1205 timed out Request 1206 timed out 64 bytes from 10.100.20.10: icmp_seq=1207 ttl=254 time=1.476 ms 64 bytes from

Over deze vertaling

Cisco heeft dit document vertaald via een combinatie van machine- en menselijke technologie om onze gebruikers wereldwijd ondersteuningscontent te bieden in hun eigen taal. Houd er rekening mee dat zelfs de beste machinevertaling niet net zo nauwkeurig is als die van een professionele vertaler. Cisco Systems, Inc. is niet aansprakelijk voor de nauwkeurigheid van deze vertalingen en raadt aan altijd het oorspronkelijke Engelstalige document (link) te raadplegen.