排除ACI中的虛擬埠通道(vPC)故障

目錄

簡介 背景資訊 縮寫 vPC埠通道故障排除的先決條件 vPC驗證 1. vPC狀態: show vpc 2. vPC角色、vPC系統mac和LAG ID:show vpc role 3. Port-channel Status : show port-channel extended 4. TEP詳細資訊和邏輯對等鏈路狀態: show system internal epm vpc 5. ZMQ連線詳細資訊:顯示系統內部vpcm zmg統計資訊 排除VPC埠通道故障 1.物理埠關閉 <u>2.由LACP暫停</u> 3.由vPC暫停 4. LACP暫停個人 其他錯誤 1. mcp-loop-err-disable 2. bpdu-guard-err-disable

簡介

本文檔介紹如何識別和解決ACI中的vPC可能出現的問題。

背景資訊

虛擬埠通道(vPC)允許物理連線到兩個不同的ACI枝葉節點的鏈路顯示為連線到第三台裝置(即網路 交換機、伺服器和支援鏈路聚合技術的任何其他網路裝置)的單個埠通道。vPC包括兩個指定為 vPC對等交換機的ACI枝葉交換機。在vPC對等裝置中,一個為主要,另一個為輔助。由交換機構成 的系統稱為vPC域

vPC對等體之間沒有專用對等鏈路;相反,交換矩陣本身充當MCT。

·對等可達性協定 — 使用ZMQ代替CFS。

·ZMQ是一個開源的高效能消息庫,使用TCP作為傳輸。

·此庫在交換機上打包為libzmq,並連結到需要與vPC對等體通訊的每個應用程式。

對等體可達性不是通過物理對等鏈路來處理;而是使用路由觸發器來檢測對等體可達性。

·vPC Manager向URIB註冊對等路由通知。

·當ISIS發現到對等體的路由時,URIB會通知vPC管理器,然後嘗試開啟與對等體的ZMQ套接字

0

·當對等路由被ISIS撤銷時,vPC管理器將再次被URIB通知,並關閉MCT鏈路。

作為升級最佳實踐的一部分,建議至少在兩個不同的組中升級每個pod中的交換機,以便每個pod中 有一半的枝葉和主幹節點在任意給定時間啟動。例如,一個組具有偶數編號的枝葉和主幹節點,另 一個組在每個pod中具有奇數編號的枝葉和主幹。通過vPC配置裝置,我們可以將至少一個裝置置於 不同的組中,以確保在升級期間至少有一個裝置處於運行狀態。這樣可防止升級期間出現任何故障 ,因為至少有一個裝置在升級另一個裝置時保持運行。

縮寫

ACI:以應用程序為中心的基礎設施

- vPC:虚擬埠通道
- MCT:多機箱EtherChannel中繼
- CFS:思科交換矩陣服務
- ZMQ:零消息隊列
- LACP:連結彙總控制通訊協定
- PDU:協定資料單元
- LAG:鏈路聚合

vPC埠通道故障排除的先決條件

有關vPC配置,請參閱

https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/L2-configuration/Cisco-APIC-Layer2-Configuration-Guide-42x/Cisco-APIC-Layer2-Configuration-Guide-421_chapter_0111.html

https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/2x/L2 config/b Cisco APIC Layer 2 Configuration Guide/b Cisco APIC Layer 2 Configuration Guide_chapter_0100.html

vPC驗證

1. vPC狀態: show vpc

輸出顯示,使用vPC域ID 101形成對等鄰接關係,注意vPC保持活動狀態在ACI中禁用,因為不需要 專用鏈路。Po3在活動VLAN 86的vPC中處於UP狀態。請注意,vPC對交換機上的埠通道號可以不 同。

2. vPC角色、vPC系統mac和LAG ID:show vpc role

由於終端裝置連線到兩個不同的vPC交換機,因此它們必須有一種機制將vPC對等裝置標識為一個 邏輯裝置。這可以通過在對等體之間共用的LAG ID中使用vPC系統MAC來實現。這使終端裝置將 vPC對等裝置視為一個邏輯單元。

N3K# show lacp interface ethernet 1/24 Interface Ethernet1/24 is up Channel group is 1 port channel is Pol PDUs sent: 31726 PDUs rcvd: 31634 Markers sent: 0 Markers rcvd: 0 Marker response sent: O Marker response rcvd: O Unknown packets rcvd: O Illegal packets rcvd: O Lag Id: [[(7f9b, 0-23-4-ee-be-65, 82ae, 8000, 4121), (8000, 0-a6-ca-75-6f-c1, 8000, 8000, 15d)]] Operational as aggregated link since Fri Sep 2 08:05:52 2022 Local Port: Eth1/24 MAC Address= 0a6-ca-75-6f-c1 System Identifier=0x8000, Port Identifier=0x8000,0x15d Operational key=32768 LACP_Activity=active LACP_Timeout=Long Timeout (30s) Synchronization=IN_SYNC Collecting=true Distributing=true Partner information refresh timeout=Long Timeout (90s) Actor Admin State=61 Actor Oper State=61 Neighbor: 0x4121 MAC Address= 0-23-4-ee-be-65 System Identifier=0x7f9b, Port Identifier=0x8000,0x4121 Operational key=33454 LACP_Activity=active LACP_Timeout=Long Timeout (30s) Synchronization=IN_SYNC Collecting=true Distributing=true Partner Admin State=61 Partner Oper State=61 Aggregate or Individual(True=1)= 1 N3K# show lacp interface ethernet 1/25 Interface Ethernet1/25 is up Channel group is 1 port channel is Pol PDUs sent: 31666 PDUs rcvd: 31651 Markers sent: 0 Markers rcvd: 0 Marker response sent: 0 Marker response rcvd: 0 Unknown packets rcvd: 0 Illegal packets rcvd: 0 Lag Id: [[(7f9b, 0-23-4-ee-be-65, 82ae, 8000, 111), (8000, 0-a6-ca-75-6f-c1, 8000, 8000, 161)]] Operational as aggregated link since Fri Sep 2 08:00:34 2022 Local Port: Eth1/25 MAC Address= 0-a6-ca-75-6f-c1 System Identifier=0x8000, Port Identifier=0x8000,0x161 Operational key=32768 LACP_Activity=active LACP_Timeout=Long Timeout (30s) Synchronization=IN_SYNC Collecting=true Distributing=true Partner information refresh timeout=Long Timeout (90s) Actor Admin State=61 Actor Oper State=61 Neighbor: 0x111 MAC Address= 0-23-4-ee-be-65 System Identifier=0x7f9b, Port Identifier=0x8000,0x111 Operational key=33454 LACP_Activity=active LACP_Timeout=Long Timeout (30s) Synchronization=IN_SYNC Collecting=true Distributing=true Partner Admin State=61 Partner Oper State=61 Aggregate or Individual(True=1)=

輸出顯示LAG ID(7f9b, 0-23-4-ee-be-65, 82ae, 8000, 4121),它是作為系統ID的優先順序(32667 in Hex)、vPC系統mac(00:23:04:ee:be:65)、操作金鑰(33454 in Hex)和埠識別符號的組合。

3. Port-channel Status : show port-channel extended

-- 3 Po3(SU) 101-102 LACP Eth1/33(P)

Show port-channel extended顯示有關埠通道捆綁所包含的物理鏈路狀態的詳細資訊。

4. TEP詳細資訊和邏輯對等鏈路狀態: show system internal epm vpc

FAB3-L1# show system internal epm vpc Local TEP IP : 10.3.208.64 Peer TEP IP : 10.3.208.67 vPC configured : Yes vPC VIP : 10.3.16.67 MCT link status : Up Local vPC version bitmap : 0x7 Peer vPC version bitmap : 0x7 Negotiated vPC version : 3 Peer advertisement received : Yes Tunnel to vPC peer : Up vPC# 686 if : port-channel3, if index : 0x16000002 local vPC state : MCEC_STATE_UP, peer vPC state : MCEC_STATE_UP current link state : LOCAL_UP_PEER_UP vPC fast conv : Off

5. ZMQ連線詳細資訊:顯示系統內部vpcm zmq統計資訊

FAB3-L1# show system internal vpcm zmq statistics ------ ZMQ server : 1 ZmQ: Registered MCECM ZMQ counters ------ ZMQ server : 1 ZmQ: Registered ZmQ print callback ZmQ: ===== Start ZMQ statistics printing ===== ZmQ: ZMQ socket type: 5, local ID: 40d0030a ZmQ: Socket base 0x1109c3b4, #endpoints 1 ZmQ: Total 1 I/O pipes, CONNECT CNT: 0, DISCONNECT CNT: 0 ZmQ: RX CNT: 66, BYTES: 124132, ERRORS: 0 ZmQ: TX CNT: 66, BYTES: 125096, ERRORS: 0 ZmQ: Pipe tcp://10.3.208.64:5001 (ID: FD 54 flag 1 state 0): read 66 (124132 bytes) write 66 (125096 bytes) Peer I/O pipe: read 66 (125096 bytes) write 66 (124132 bytes) ZmQ: Stream engine 0xae90049c ZMQ SOCKET 0x1109c3b4 TCP FD: 54 @ 10.3.208.67:58740 ZmQ: RX CNT: 72 BYTES: 124494 ERRORS: 0 TX CNT: 73 BYTES: 125458 ERRORS: 0 ZmQ: CONNECT CNT: 0 DISCONNECT CNT: 0 ZmQ: ===== End ZMQ statistics printing =====

ZMQ統計資訊顯示ZMQ會話的狀態、連線和斷開的發生次數以及發生的任何錯誤。

排除VPC埠通道故障

1.物理埠關閉

輸出鞋Po3下降。

我們進一步瞭解作為port-channel一部分的介面狀態。這裡Eth1/33處於Down狀態。LACP被配置為 繫結協定。

FAB3-L1# show int e1/33 Ethernet1/33 is down (notconnect) admin state is up, Dedicated Interface Belongs to po3 Hardware: 100/1000/10000/auto Ethernet, address: 0081.c4b1.2521 (bia 0081.c4b1.2521) MTU 9000 bytes, BW 0 Kbit, DLY 1 usec reliability 255/255, txload 1/255, rxload 1/255 Encapsulation ARPA, medium is broadcast Port mode is trunk full-duplex, 10 Gb/s FEC (forward-error-correction) : disable-fec Beacon is turned off Auto-Negotiation is turned on Input flow-control is off, output flow-control is off Auto-mdix is turned off Switchport monitor is off EtherType is 0x8100 EEE (efficient-ethernet) : n/a Last link flapped 00:08:15 Last clearing of "show interface" counters never 9 interface resets 30 seconds input rate 0 bits/sec, 0 packets/sec 30 seconds output rate 0 bits/sec, 0 packets/sec Load-Interval #2: 5 minute (300 seconds) input rate 0 bps, 0 pps; output rate 0 bps, 0 pps Show interface output提供介面e1/33的更多詳細資訊。我們可以看到E1/33在notconnect狀態下關 閉。

建議的操作:

請確保連線埠正確連線且組態正確。

2.由LACP暫停

輸出顯示Eth1/33處於掛起狀態。接下來,我們檢視show interface Eth1/33以瞭解更多詳細資訊。

FAB3-L1# show int e1/33 Ethernet1/33 is down (suspended-due-to-no-lacp-pdus) admin state is up, Dedicated Interface Belongs to po3 Hardware: 100/1000/10000/auto Ethernet, address: 0081.c4b1.2521 (bia 0081.c4b1.2521) MTU 9000 bytes, BW 0 Kbit, DLY 1 usec reliability 255/255, txload 1/255, rxload 1/255 Encapsulation ARPA, medium is broadcast Port mode is trunk fullduplex, 10 Gb/s FEC (forward-error-correction) : disable-fec Beacon is turned off Auto-Negotiation is turned on Input flow-control is off, output flow-control is off Auto-mdix is turned off Switchport monitor is off EtherType is 0x8100 EEE (efficient-ethernet) : n/a Last link flapped 00:00:13 Last clearing of "show interface" counters never 12 interface resets 30 seconds input rate 0 bits/sec, 0 packets/sec 30 seconds output rate 1640 bits/sec, 0 packets/sec Show interface表示連線埠已暫停,因為沒有LACP PDU。我們可以進一步檢視LACP計數器,並確 定是否正在傳送和接收LACP PDU。

輸出顯示,計數器僅遞增傳送的LACPDU,Recv計數器保持不變。這表示我們沒有收到來自遠端的 LACP PDU。

我們還可以檢視LACP協商引數、計數器等。對於特定介面,請使用「show lacp interface e1/33」

FAB3-L1# show lacp interface e1/33 Interface Ethernet1/33 is suspended Channel group is 3 port channel is Po3 PDUs sent: 317 PDUs rcvd: 264 received Markers sent: 0 Markers rcvd: 0 Marker response sent: 0 Marker response rcvd: 0 Unknown packets rcvd: 0 Illegal packets rcvd: 0 Lag Id: [[(7f9b, 00-23-04-ee-be-65, 82ae, 8000, 121), (0, 0-0-0-0-0, 0, 0, 0)]] Operational as aggregated link since Mon Aug 22 09:29:53 2022 Local Port: Eth1/33 MAC Address= 00-81-c4-b1-25-4f System Identifier=0x8000,00-81-c4-b1-25-4f Port Identifier=0x8000,0x121 Operational key=33454 LACP_Activity=active LACP_Timeout=Long Timeout (30s) Synchronization=NOT_IN_SYNC Collecting=false Distributing=false Partner information refresh timeout=Long Timeout (90s) Actor Admin State=(Ac-1:To-0:Ag-1:Sy-0:Co-0:Di-0:De-1:Ex-0) Actor Oper State=Ac-1:To-0:Ag-1:Sy-0:Co-0:Di-0:De-1:Ex-0 Neighbor: 0x0 MAC Address= 0-0-0-0-0 System Identifier=0x0,0x0 Port Identifier=0x0,0x0 Operational key=0 LACP_Activity=unknown LACP_Timeout=Long Timeout (30s) Synchronization=NOT_IN_SYNC Collecting=false Distributing=false Partner Admin State=(Ac-0:To-0:Ag-0:Sy-0:Co-0:Di-0:De-0:Ex-0) Partner Oper State=(Ac-0:To-0:Ag-0:Sy-0:Co-0:Di-0:De-0:Ex-0) Aggregate or Individual(True=1)= 2

此外,還可以在枝葉上為LACP資料包捕獲資料包,您可以使用特定過濾器過濾出有問題的介面。

tcpdump -vvvi kpm_inb ether proto 0x8809 建議的操作:

確保在遠端端正確配置了LACP,並且裝置在正確的介面上傳送了LACP PDU。

3.由vPC暫停

此輸出顯示vPC埠通道因vPC配置錯誤而關閉。現在來進一步瞭解埠通道狀態。

這裡Eth1/33處於Down狀態,我們進一步檢視「show interface e1/33」以瞭解更多詳細資訊。

FAB3-L1# show int e1/33 Ethernet1/33 is down (suspend-by-vpc) admin state is up, Dedicated Interface Belongs to po3 Hardware: 100/1000/10000/auto Ethernet, address: 0081.c4b1.2521 (bia 0081.c4b1.2521) MTU 9000 bytes, BW 0 Kbit, DLY 1 usec reliability 255/255, txload 1/255, rxload 1/255 Encapsulation ARPA, medium is broadcast Port mode is trunk full-duplex, 10 Gb/s FEC (forward-error-correction) : disable-fec Beacon is turned off Auto-Negotiation is turned on Input flow-control is off, output flow-control is off Auto-mdix is turned off Switchport monitor is off EtherType is 0x8100

vPC使用LAG ID確定vPC對等體是否連線到同一主機,如果LAG ID不匹配,則介面由vPC掛起。 「Show vpc brief」顯示vPC對等體上的埠通道中的物理鏈路未連線到同一遠端裝置。

可以使用「show vpc consistency-parameters interface port-channel 3」檢查LAG ID比較。

建議的操作:

確保port-channel中的物理鏈路連線到同一遠端裝置。

4. LACP暫停個人

LACP將埠設定為掛起狀態(如果它沒有從對等方接收LACP PDU)。這可能會導致某些伺服器無 法啟動,因為它們需要LACP來邏輯啟動埠。您可以通過禁用LACP掛起個人將行為調整為單獨使用 。為此,請在vPC策略組中建立一個埠通道策略,並在將LACP活動模式設定為後,刪除Suspend Individual Port。現在,vPC中的埠保持活動狀態,並繼續傳送LACP資料包。

FAB3-L1# show port-channel extended Flags: D - Down P - Up in port-channel (members) I -Individual H - Hot-standby (LACP only) s - Suspended r - Module-removed b - BFD Session Wait S -Switched R - Routed U - Up (port-channel) M - Not in use. Min-links not met F - Configuration failed ------ Group Port-BundleGrp Protocol Member Ports Channel ------- Group Port------ 1 Pol(SD) 101-102 LACP Eth1/33(I)

輸出顯示,即使刪除LACP Suspend-Individual標誌後,我們沒有在Eth1/33上接收LACP PDU,埠 作為單個埠處於UP狀態。請注意,我們仍然通過此配置從ACI枝葉傳送LACP PDU,一旦收到 LACP PDU,埠將返回到捆綁模式。

其他錯誤

更多並非特定於vPC但仍適用於vPC介面的介面錯誤。有關詳細資訊,請參閱連結。

1. mcp-loop-err-disable

https://www.cisco.com/c/dam/en/us/solutions/collateral/data-center-virtualization/applicationcentric-infrastructure/aci-guide-using-mcp-mis-cabling-protocol.pdf

2. bpdu-guard-err-disable

https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/aci_virtual_edge/configuration/1x/b_Virtual_Edge_Config_Guide_1_2_1/b_Virtual_Edge_Config_Guide_1_2_1_chapter_0101.pdf

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)。