# 排除ACI中的虚拟端口通道(vPC)故障

## 目录

简介 背景信息 缩写 vPC端口通道故障排除的前提条件 vPC验证 1. vPC状态: show vpc 2. vPC角色、vPC系统mac和LAG ID:show vpc role 3. Port-channel Status : show port-channel extended 4. TEP详细信息和逻辑对等链路状态: show system internal epm vpc 5. ZMQ连接详细信息: show system internal vpcm zmq statistics 排除VPC端口通道故障 1.物理端口关闭 2.由LACP暂停 3.由vPC暂停 <u>4. LACP暂停个人</u> 其他错误 1. mcp-loop-err-disable 2. bpdu-guard-error-disable

## 简介

本文档介绍如何识别和解决ACI中的vPC可能出现的问题。

## 背景信息

虚拟端口通道(vPC)允许物理上连接到两个不同ACI枝叶节点的链路显示为连接到第三台设备(即网络交换机、服务器、支持链路聚合技术的任何其他网络设备)的单个端口通道。vPC包括两个指定为vPC对等交换机的ACI枝叶交换机。在vPC对等设备中,一个为主设备,另一个为辅助设备。交换机构成的系统称为vPC域

vPC对等设备之间没有专用对等链路;交换矩阵本身充当MCT。

·对等可达性协议 — 使用ZMQ代替CFS。

·ZMQ是一个开源的高性能消息库,使用TCP作为传输。

·此库在交换机上打包为libzmq,并链接到需要与vPC对等设备通信的每个应用程序。

对等连通性不是通过物理对等链路来处理;而是使用路由触发器来检测对等连通性。

·vPC Manager向URIB注册对等路由通知。

·当ISIS发现到对等体的路由时,URIB会通知vPC管理器,然后尝试打开与对等体的ZMQ套接字

0

·当对等路由被ISIS撤销时,vPC管理器会再次被URIB通知,从而使MCT链路断开。

作为升级最佳实践的一部分,建议升级每个Pod中至少两个独立组的交换机,以使每个Pod中的一半 枝叶和主干节点在任何给定时间都处于工作状态。例如,一个组具有偶数编号的枝叶和主干节点 ,另一个组在每个pod中具有奇数编号的枝叶和主干。通过配置vPC的设备,我们可以将至少一台设 备置于不同的组中,以确保在升级过程中该设备处于工作状态。这可以防止升级期间出现任何中断 ,因为至少有一个设备在另一个设备升级期间保持运行状态。

### 缩写

ACI:以应用为中心的基础设施

- vPC:虚拟端口通道
- MCT:多机箱EtherChannel中继
- CFS:思科交换矩阵服务
- ZMQ:零消息队列
- LACP:链路汇聚控制协议
- PDU:协议数据单元
- LAG:链路聚合

### vPC端口通道故障排除的前提条件

有关vPC配置,请参阅

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验证

#### 1. vPC状态: show vpc

输出显示,使用vPC域ID 101形成对等邻接关系,注意vPC保持活动状态在ACI中禁用,因为不需要 专用链路。具有活动vlan 86的vPC中的Po3处于UP状态。请注意,vPC对交换机上的端口通道号可 能不同。

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

#### 由于终端设备连接到两台不同的vPC交换机,因此它们必须有一种机制将vPC对等设备标识为一个 逻辑设备。这可以通过在对等体之间共享的LAG ID中使用vPC系统MAC来实现。这使终端设备将 vPC对等设备视为一个逻辑单元。

N3K# show lacp interface ethernet 1/24 Interface Ethernet1/24 is up Channel group is 1 port channel is Po1 PDUs sent: 31726 PDUs rcvd: 31634 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, 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 Po1 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)=

输出显示LAG ID(7f9b, 0-23-4-ee-be-65, 82ae, 8000, 4121),它是作为系统ID的优先级(32667 in Hex)、vPC系统mac(00:23:04:ee:be:65)、操作密钥(33454 in Hex)和端口标识符的组合。

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

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

Show port-channel extended显示有关属于端口通道捆绑的物理链路状态的更多信息。

#### 4. TEP详细信息和逻辑对等链路状态: 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连接详细信息: show system internal vpcm zmq statistics

FAB3-L1# show system internal vpcm zmq statistics ------ 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统计信息显示ZMQ会话的状态、连接和断开的次数以及发生的任何错误。

### 排除VPC端口通道故障

#### 1.物理端口关闭

#### 输出鞋Po3下降。

FAB3-L1# show port-channel summary Flags: D - Down P - Up in port-channel (members) I -
Individual H - Hot-standby (LACP only) s - Suspended r - Module-removed S - Switched R - Routed
II - IIn (port-channel) M - Not in use Min-links not mot E - Configuration failed
0 - op (port-channer) M - Not in use. Min-rinks not met F - Configuration failed
Group Port- Type Protocol
Member Ports Channel
3 Po3(SD) Eth LACP Eth1/33(D)

#### 我们进一步查看属于port-channel一部分的接口状态。此时,Eth1/33处于Down(关闭)状态。 LACP配置为捆绑协议。

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输出提供了有关接口e1/33的详细信息。我们可以看到E1/33处于notconnect状态。

#### 建议操作:

确保端口连接正确并且配置正确。

#### 2.由LACP暂停

-- 3 Po3(SD) 101-102 LACP Eth1/33(s)

#### 输出显示Eth1/33处于挂起状态。接下来,我们查看show interface Eth1/33以了解更多详细信息。

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建议端口暂停,因为没有LACP PDU。我们可以进一步查看LACP计数器并确定是否 正在发送和接收LACP PDU。

输出显示,计数器只递增友达的LACPDU,而接收计数器保持个变。这表明我们没有收到米目远程 终端的LACP PDU。

我们还可以查看LACP协商参数、计数器等。对于特定接口,请使用"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,还可以在枝叶上为LACP数据包捕获数据包,您可以使用特定过滤器过滤出有问题的接口。

tcpdump -vvvi kpm\_inb ether proto 0x8809

建议操作:

确保在远程端正确配置了LACP,并且设备在正确的接口上发送了LACP PDU。

#### 3.由vPC暂停

#### 此输出显示vPC端口通道因vPC配置错误而关闭。让我们进一步了解端口通道状态。

---- 3 Po3(SD) Eth LACP Eth1/33(D)

#### 此处Eth1/33处于Down状态,我们进一步查看"show interface e1/33"以了解更多详细信息。

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确定vPC对等体是否连接到同一主机,如果LAG ID不匹配,则接口由vPC挂起。 "Show vpc brief"显示vPC对等体上的端口通道中的物理链路未连接到同一远程设备。

#### 可以使用"show vpc consistency-parameters interface port-channel 3"检查LAG ID比较。

#### 建议操作:

确保port-channel中的物理链路连接到同一远程设备。

#### 4. LACP暂停个人

LACP将端口设置为挂起状态(如果它未收到来自对等设备的LACP PDU)。这会导致某些服务器 无法启动,因为它们需要LACP以逻辑方式启动端口。您可以通过禁用LACP挂起个体将行为调整为 单独使用。为此,请在vPC策略组中创建一个端口通道策略,在将LACP活动模式设置为后,删除 Suspend Individual Port。现在,vPC中的端口保持活动状态,并继续发送LACP数据包。

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)

输出显示,即使删除LACP Suspend-Individual标志后,我们未在Eth1/33上收到LACP PDU,但端 口作为单个端口处于启用状态。请注意,我们仍然使用此配置从ACI枝叶发送LACP PDU,一旦收 到LACP PDU,端口将回到捆绑模式。

## 其他错误

更多并非特定于vPC但仍适用于vPC接口的接口错误。有关详细信息,请参阅链接。

#### 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-error-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

#### 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)。