排除Nexus上的链路汇聚控制协议(LACP)故障

目录

简介

本文档介绍如何对Nexus 9000 cloudscale系列上的链路聚合控制协议(LACP)进行故障排除。

先决条件

要求

思科建议您了解以下主题：

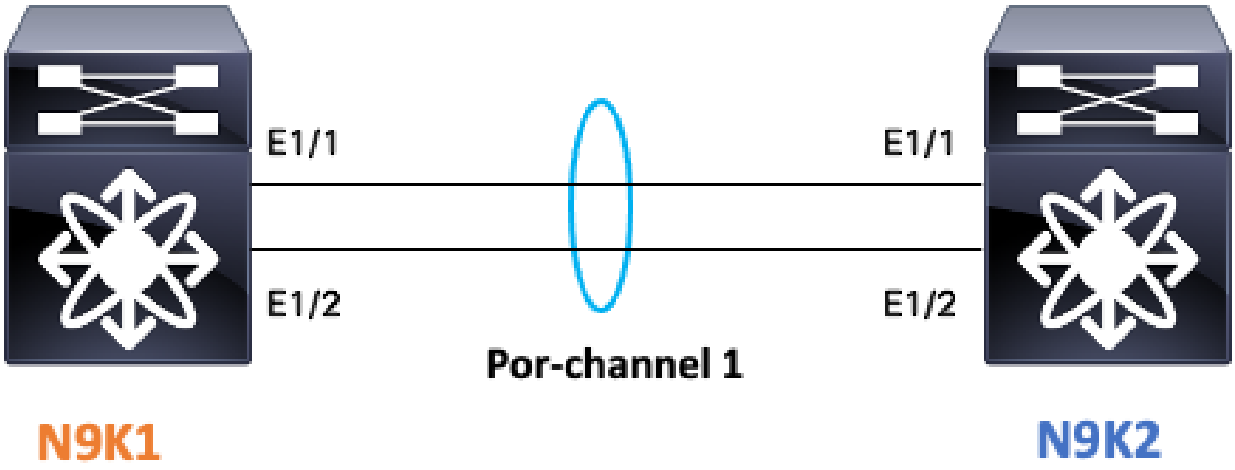
- LACP协议
- NXOS平台
- ELAM了解
- Ethanalyzer了解

使用的组件

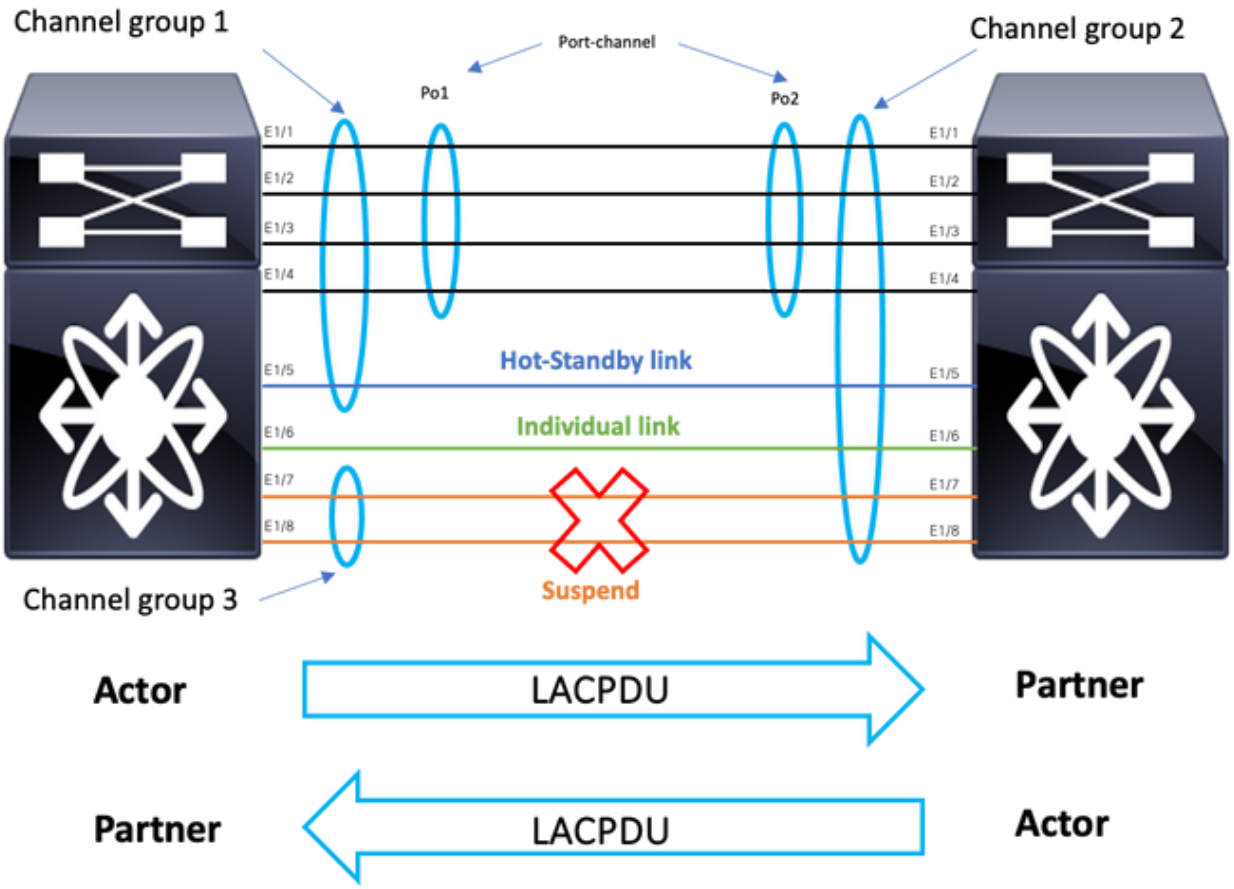
名称	平台	version
N9K1	N9K-C93108TC-EX	9.3(10)
N9K2	N9K-C93108TC-EX	9.3(10)

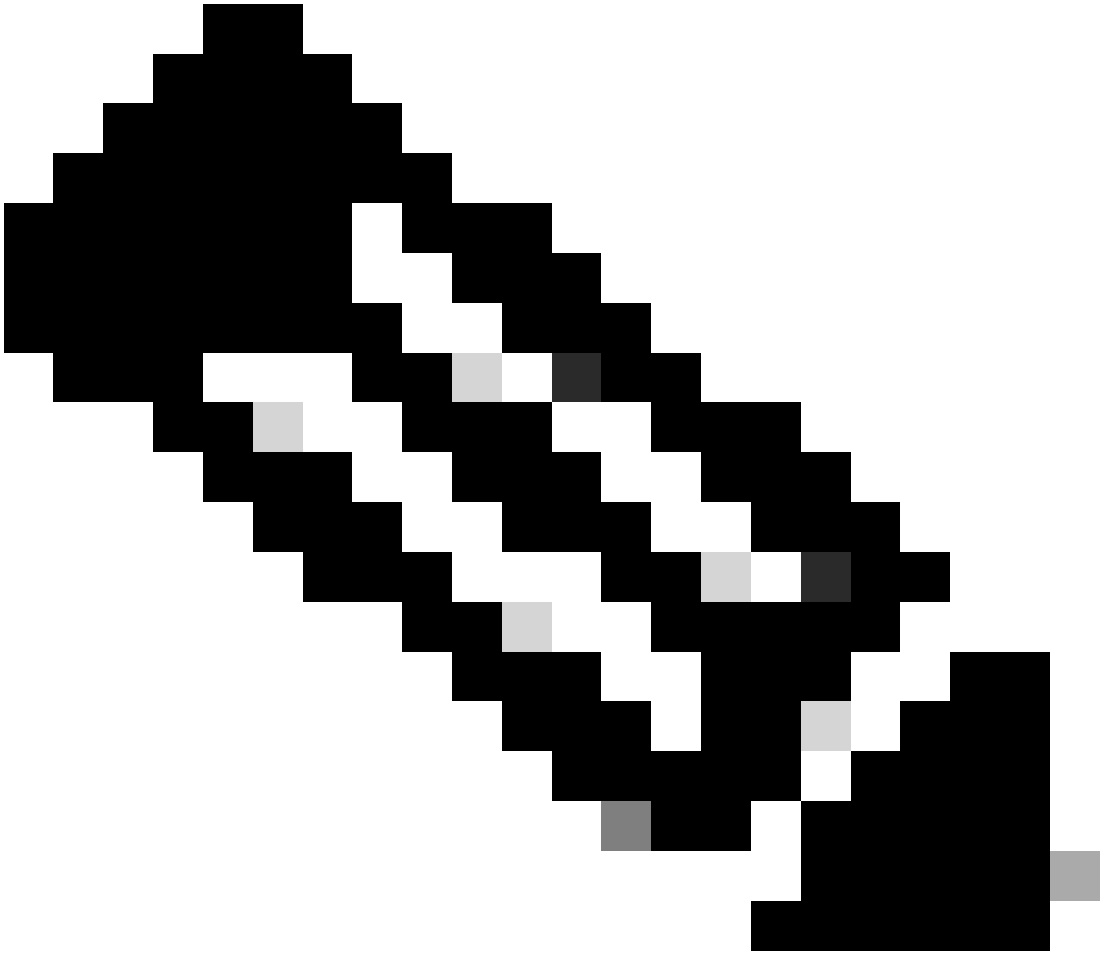
本文档中的信息都是基于特定实验室环境中的设备编写的。本文档中使用的所有设备最初均采用原始（默认）配置。如果您的网络处于活动状态，请确保您了解所有命令的潜在影响。

拓扑


E1/1
E1/2
E1/1
E1/2
Por-channel 1
N9K1
N9K2

检验LACP链路状态


Channel group 1
Port-channel
Channel group 2
Po1
Po2
E1/1
E1/2
E1/3
E1/4
E1/5
E1/6
E1/7
E1/8
Hot-Standby link
Individual link
Suspend
Channel group 3
Actor
LACPDU
Partner
Partner
LACPDU
Actor



注意：映像1.1 LACP链路状态。

配置LACP：

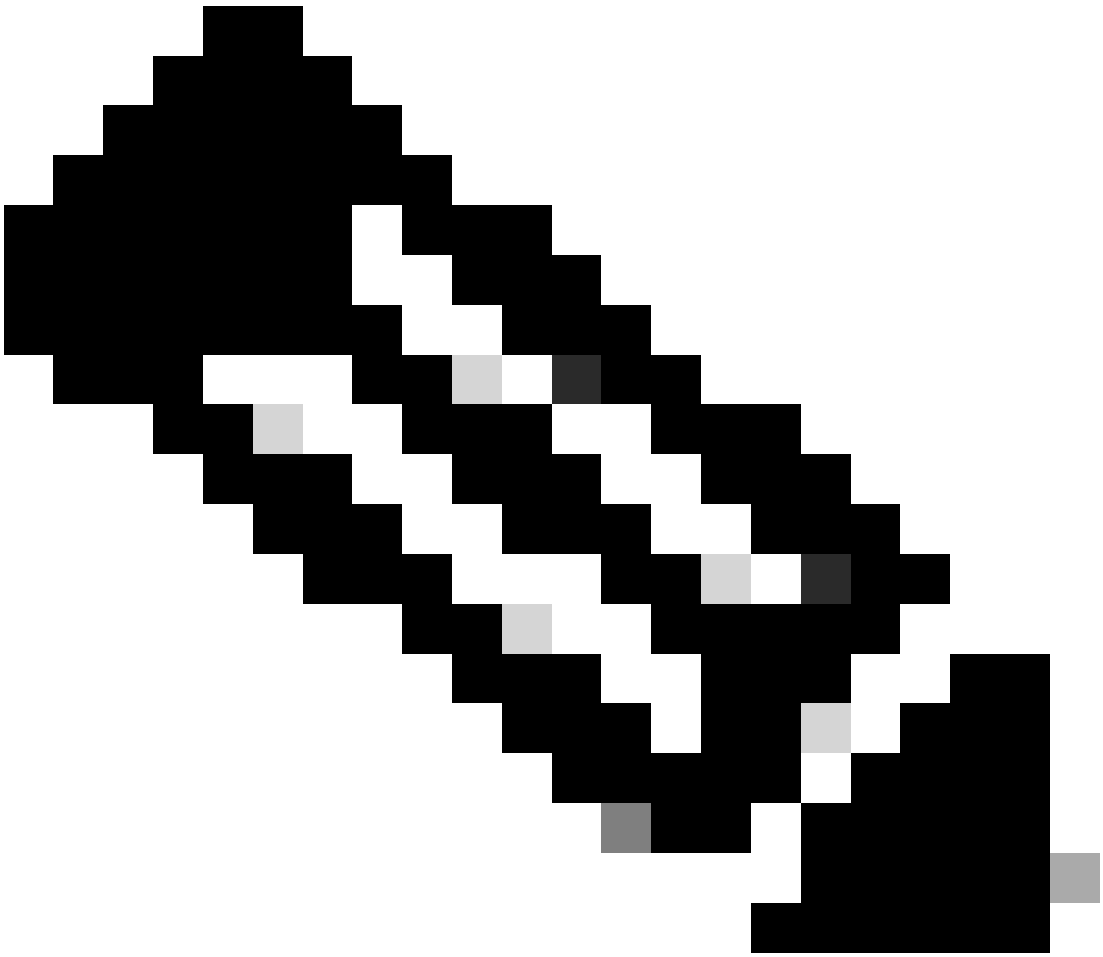
N9K1	N9K2

show run interface port-channel 1 membership
 	
interface port-channel1
  switchport
  switchport mode trunk
 	
interface Ethernet1/1
  switchport
  switchport mode trunk
  channel-group 1 mode active
  no shutdown
 	
interface Ethernet1/2
  switchport
  switchport mode trunk
	show run interface port-channel 1 membership
	
	interface port-channel1
	  switchport
	  switchport mode trunk
	
	interface Ethernet1/1
	  switchport
	  switchport mode trunk
	  channel-group 1 mode active
	  no shutdown
	
	interface Ethernet1/2
	  switchport
	  switchport mode trunk

channel-group 1 mode active no shutdown	channel-group 1 mode active no shutdown

验证端口通道状态

N9K1	
sh port-channel summary interface port-channel 1 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) p - Up in delay-lacp mode (member) M - Not in use. Min-links not met --- Group Port- Type Protocol Member Ports Channel --- 1 Po1(SU) Eth LACP Eth1/1(P)	sh port-channel summary Flags: D - Down I - Individual s - Suspended b - BFD Session S - Switched U - Up (port-ch p - Up in delay M - Not in use. --- Group Port- Type Channel --- 1 Po1(SU) Eth

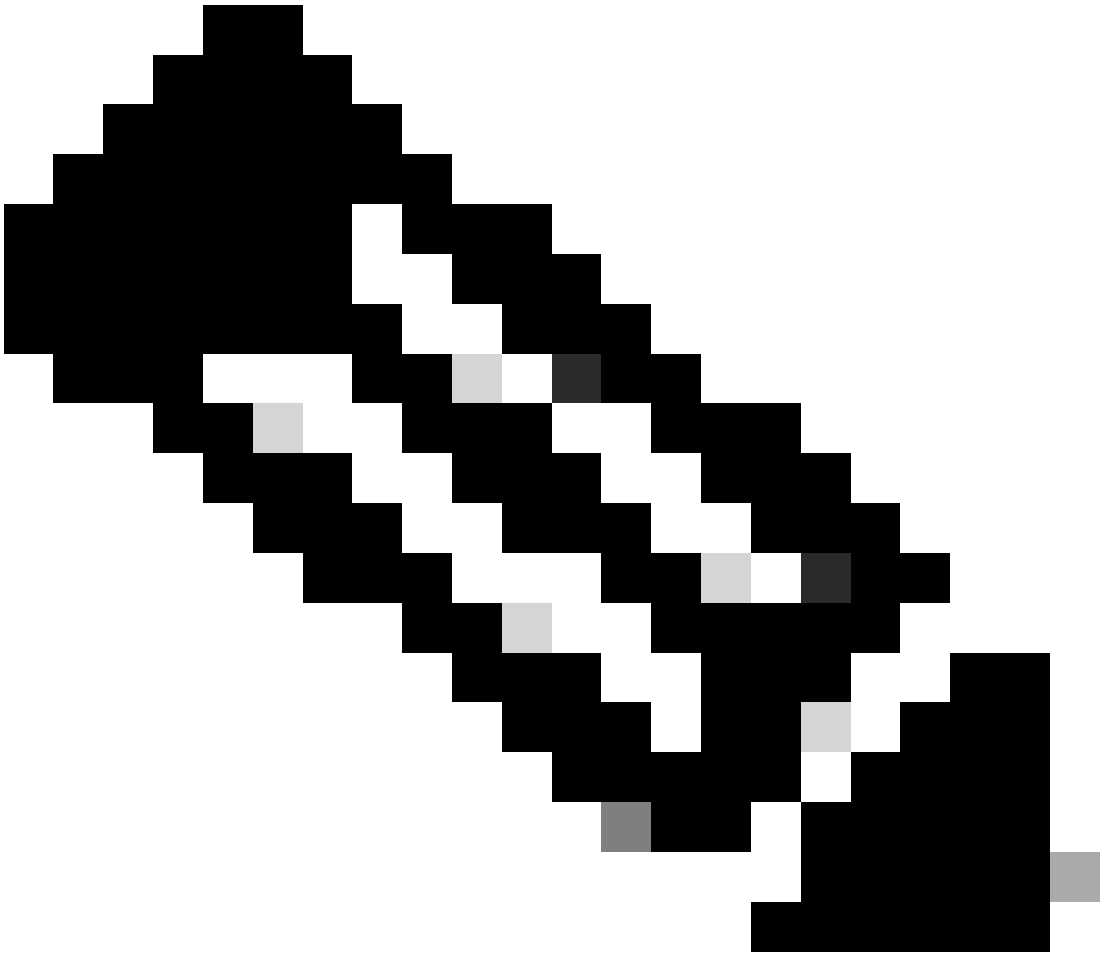


注意：最常见的故障场景是Nexus暂停接口，这在LACP暂停接口部分中介绍。

检验LACP挂起接口

```
sh port-channel summary interface port-channel 1

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)
        p - Up in delay-lacp mode (member)
        M - Not in use. Min-links not met
--------------------------------------------------------------------------------
Group Port-       Type     Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
1     Po1(SD)     Eth      LACP      Eth1/1(s)
```

```
sh int e1/1 status
--------------------------------------------------------------------------------
Port          Name           Status     Vlan    Duplex   Speed     Type
--------------------------------------------------------------------------------
Eth1/1        --             suspended  trunk    auto     auto   10Gbase-SR


sh int e1/1
Ethernet1/1 is down (suspended(no LACP PDUs))
admin state is up, Dedicated Interface
  Belongs to Po1
  Hardware: 100/1000/10000/25000 Ethernet, address: 003a.9c08.68ab (bia 003a.9c08.68ab)
  MTU 9216 bytes, BW 10000000 Kbit , DLY 10 usec
  reliability 255/255, txload 1/255, rxload 1/255

  <Snipped>
```



注意：当面临此情况时，Nexus没有从合作伙伴接收LACP PDU，可以验证LACP接口计数器，如所写部分：可以执行验证LACP接口计数器或数据包捕获，例如SPAN或ELAM(在LACP ELAM部分中说明)。

检验LACP接口计数器

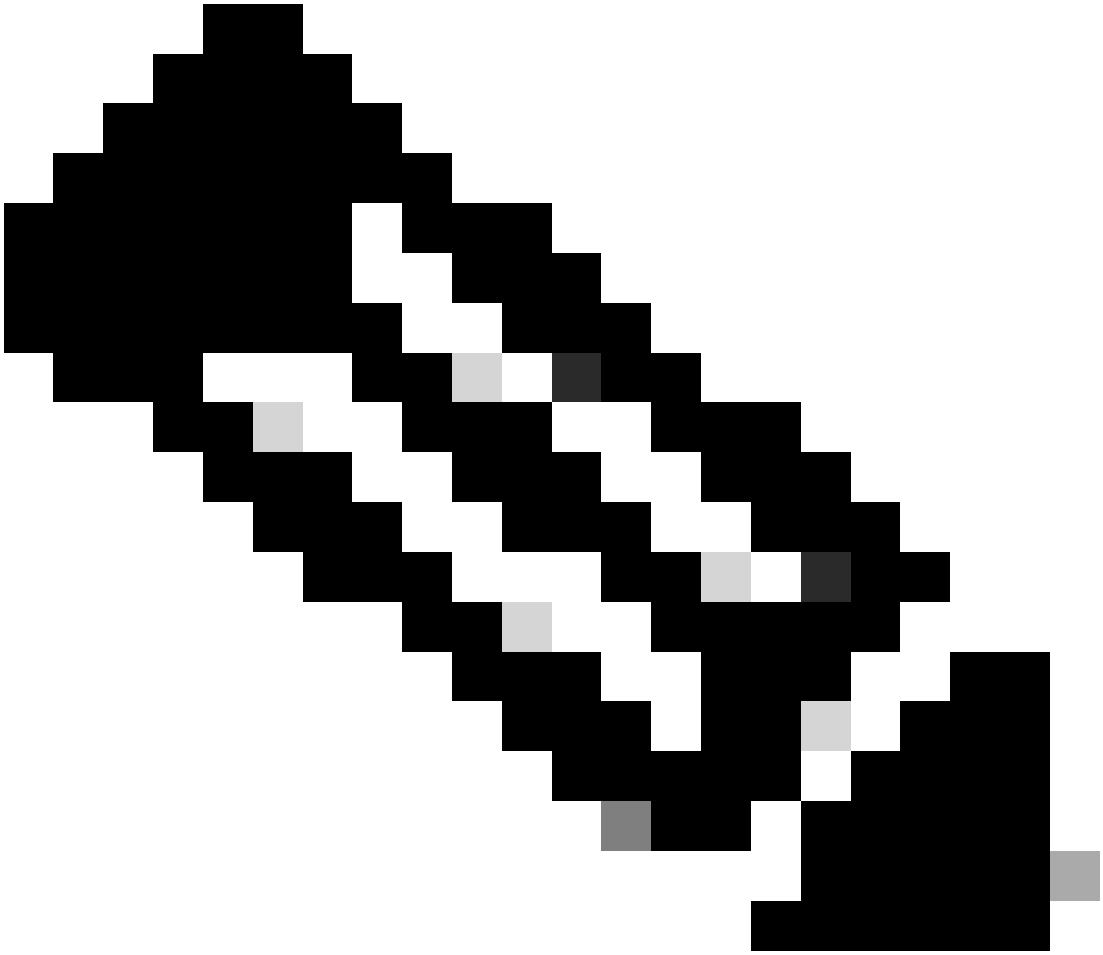
每台设备必须以相同的速率发送和接收LACPDU，端口通道才能打开。

N9K1	
sh lacp counters interface port-channel 1 NOTE: Clear lacp counters to get accurate statistics --- LACPDUs Markers/Resp LACPDUs Port Sent Recv Recv Sent Pkts Err --- port-channel1 Ethernet1/1 445 445 0 0 0 Ethernet1/2 445 445 0 0 0	sh lacp counters interfa NOTE: Clear lacp counter --- Port Sent --- port-channel1 Ethernet1/1 445 Ethernet1/2 445

检验LACP参与者状态位

在每个LACP PDU参与者中，合作伙伴与参与者之间交换状态信息。

活动	1：主动模式	0：被动模式
超时	1：短超时	0：长超时
汇聚	1：可聚合	0：个人
同步	1：同步	0：不同步
收集	1：已启用收集	0：已禁用收集
分发	1：已启用分发	0：已禁用分发
已违约	1：对合作伙伴使用默认值	0：为合作伙伴使用rx LACPDU
已到期	1：合作伙伴PDU已过期	0：未过期

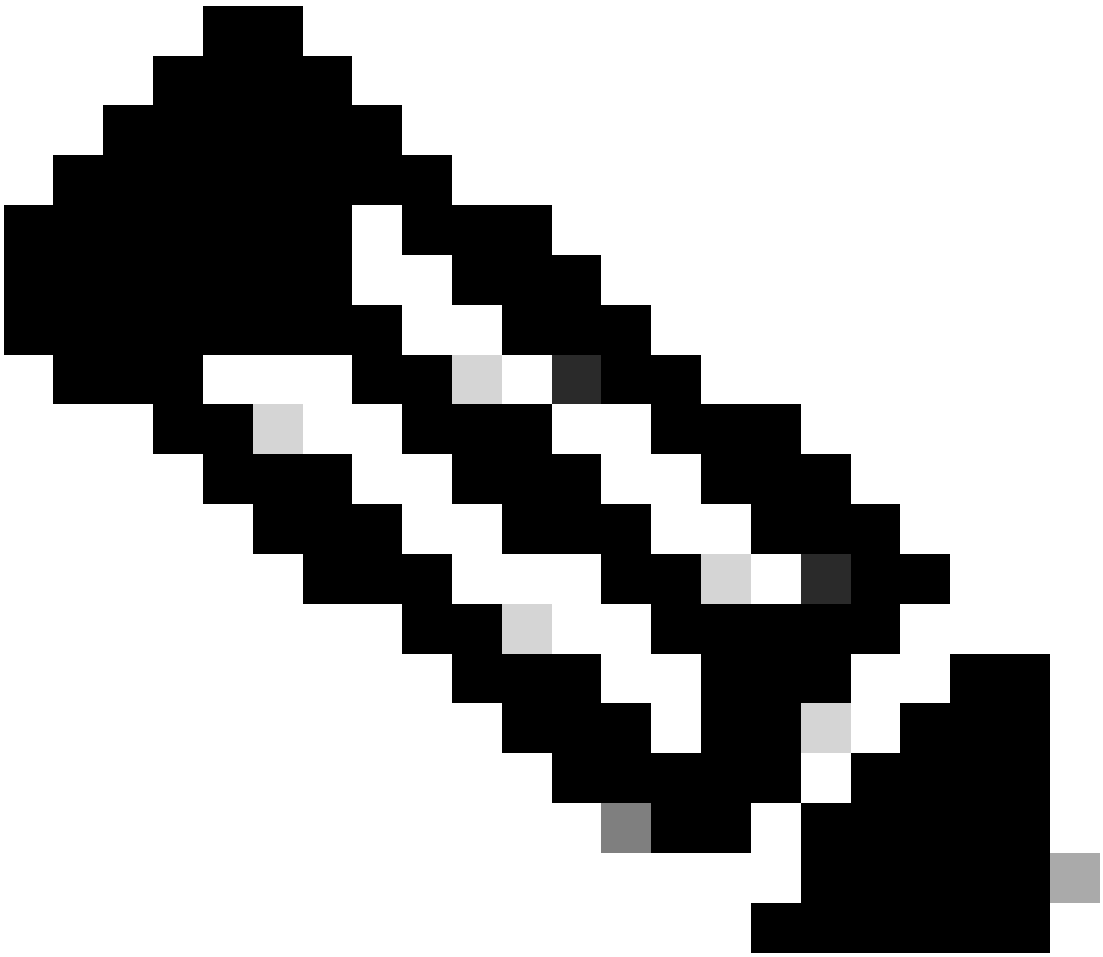


注意：表2.0参与者状态位的含义

检验LACP状态十六进制值：

state:0x3d (Ac-1:To-0:Ag-1:Sy-1:Co-1:Di-1:De-0:Ex-0)

	State							
	Ex	De	Di	Co	Sy	Ag	To	Ac
0x3d=	0	0	1	1	1	1	0	1



注意：映像3.0 LACP状态从二进制转换为十六进制

验证LAG ID

链路汇聚标识符是同一个端口通道的每个物理接口成员共享的信息，显示为单个“虚拟接口”。可以使用命令进行验证。

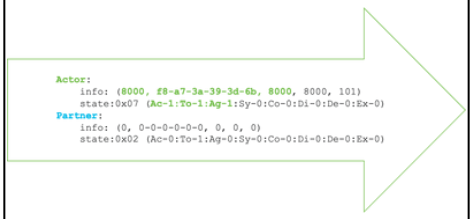
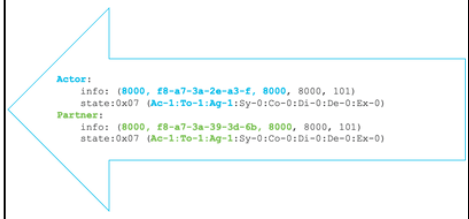
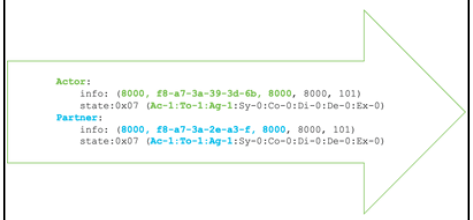
N9K1 LAG ID	N9K2 LAG ID
sh lacp interface e1/1 | include iignore local|lag Lag Id: [[(1770, a8-c-d-96-43-7f, 5, 8000, 1c9), (1770, a8-c-d-96-c9-bf, 5, 8000, 1c9)]] Local Port: Eth1/1 MAC Address=a8-c-d-96-c9-bf	sh lacp inter Lag Id: [[(1 Local Port: E
sh lacp interface e1/2 | include iignore local|lag Lag Id: [[(1770, a8-c-d-96-43-7f, 5, 8000, 1ca), (1770, a8-c-d-96-c9-bf, 5, 8000, 1ca)]] Local Port: Eth1/2 MAC Address=a8-c-d-96-c9-bf	sh lacp inter Lag Id: [[(1 Local Port: E

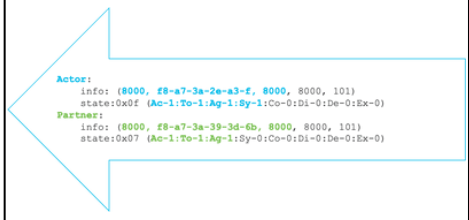
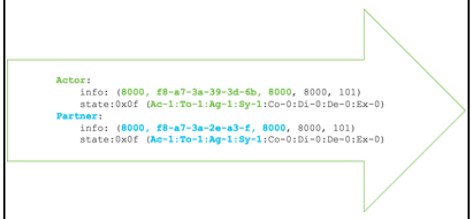
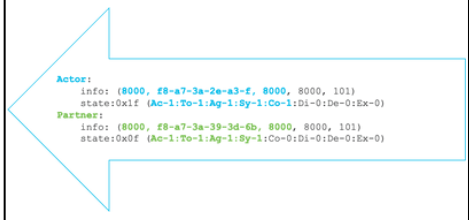
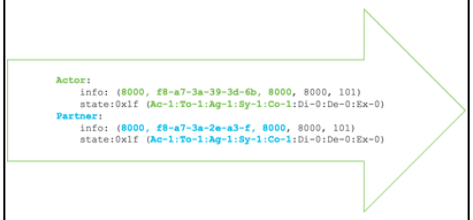
检验LACP PDU交换

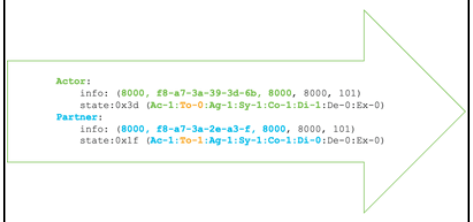
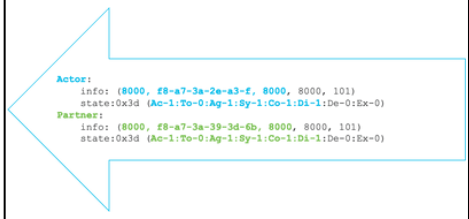
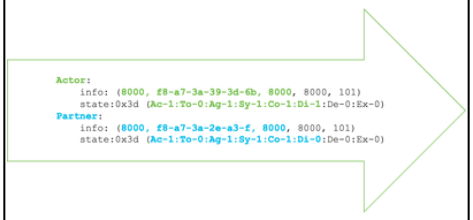
在某些场景中，即使Nexus以正确的速率交换LACP PDU，端口信道也不会接通。

这可能是由于LACP协商失败。

下表显示了一个端口通道正常运作的LACP事务示例。

1	N9K1参与者	N9K2合作伙伴
• N9K1发送包含参与者信息和状态位的LACP。 • 由于N9K1没有收到来自伙伴的任何LACP PDU，因此伙伴信息为0。		
2	N9K1合作伙伴	N9K2参与者
• N9K2接收来自N9K1的信息。 • N9K2发送LACP PDU及其信息，并确认N9K1信息（LACP PDU中包含N9K1信息）。		
3	N9K1参与者	N9K2合作伙伴
• N9K1确认N9K2信息。		
4	N9K1合作伙伴	N9K2参与者

• N9K2收到来自N9K1的确认。 • N9K2发送LACP PDU添加同步位，位置为1。		
5	N9K1参与者	N9K2合作伙伴
• N9K1确认来自N9K2的SYNC位。 • N9K1将第1处的SYNC位添加到LACP PDU。		
6	N9K1合作伙伴	N9K2参与者
• 来自N9K1的N9K2确认同步位。 • N9K2将1处的收集位添加到LACP PDU。		
7	N9K1参与者	N9K2合作伙伴
• N9K1从N9K2确认收集位。 • N9K1将收集位添加到其LACP PDU。		
8	N9K1参与者	N9K2合作伙伴

• N9K1确定已准备好转换到分发状态，因此它现在将超时位从1 (fast)更改为0 (Slow)，并将分发位设置为1。		
9	N9K1合作伙伴	N9K2参与者
• N9k2确认N9k1 PDU并将其“超时”位从1更改为0，并将distribute bite设置为1。 • 此时，两台nexus都已准备好在port-channel上发送数据。		
10	N9K1参与者	N9K2合作伙伴
• N9K1确认来自N9K2的LACP PDU。 • 此时，端口通道将全部转换到运行状态。		

检验LACP FSM日志

LACP有限状态机有一个专用日志，其中存储了接口LACP状态的所有事件，LACP PDU可在此日志中找到：

sh lacp internal info interface e1/1 detail fsmlog

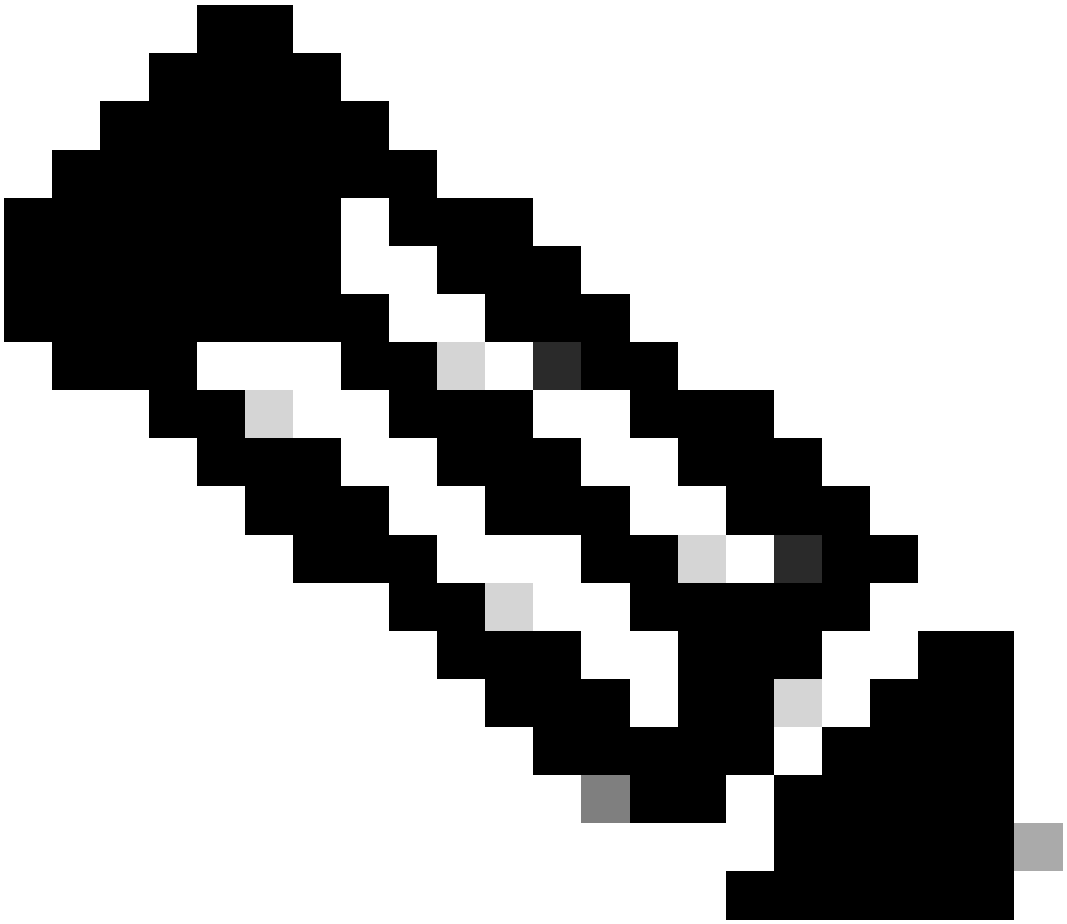
在新版本中，您还可以使用：

sh lacp internal event-history interface e1/1
第一部分介绍接口LACP转换

```
>>>>FSM: <Ethernet1/1> has 61 logged transitions<<<<<
<Snipped>

58) FSM:<Ethernet1/1> Transition at 127198 usecs after Mon Aug 14 22:34:42 2023
    Previous state: [LACP_ST_WAIT_FOR_HW_TO_PROGRAM_RECEIVE_PATH]
    Triggered event: [LACP_EV_PORT_RECEIVE_PATH_ENABLED_AS_CHANNEL_MEMBER_MESSAGE]
    Next state: [LACP_ST_PORT_MEMBER_RECEIVE_ENABLED]

59) FSM:<Ethernet1/1> Transition at 127227 usecs after Mon Aug 14 22:34:42 2023
    Previous state: [LACP_ST_PORT_MEMBER_RECEIVE_ENABLED]
    Triggered event: [LACP_EV_PARTNER_PDU_IN_SYNC_COLLECT_ENABLED_DISTRIBUTING_DISABLED]
    Next state: [LACP_ST_WAIT_FOR_HW_TO_PROGRAM_TRANSMIT_PATH]

60) FSM:<Ethernet1/1> Transition at 128265 usecs after Mon Aug 14 22:34:42 2023
    Previous state: [LACP_ST_WAIT_FOR_HW_TO_PROGRAM_TRANSMIT_PATH]
    Triggered event: [LACP_EV_PERIODIC_TRANSMIT_TIMER_EXPIRED]
    Next state: [FSM_ST_NO_CHANGE]

61) FSM:<Ethernet1/1> Transition at 134352 usecs after Mon Aug 14 22:34:42 2023
    Previous state: [LACP_ST_WAIT_FOR_HW_TO_PROGRAM_TRANSMIT_PATH]
    Triggered event: [LACP_EV_PORT_HW_PATH_ENABLED]
    Next state: [LACP_ST_PORT_MEMBER_COLLECTING_AND_DISTRIBUTING_ENABLED]
    Curr state: [LACP_ST_PORT_MEMBER_COLLECTING_AND_DISTRIBUTING_ENABLED]
```

第二部分显示已发送或接收的所有LACP PDU信息。

```
<Snipped>
(1) Send LACP PDU: len:110 at 492243 usecs after Tue Aug 15 00:02:13 2023
01010114 8000f8a7 3a393d6b 80008000 01013d00 00000214 8000f8a7 3a2ea30f
80008000 01013d00 00000310 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 0000
Actor:
    info: (8000, f8-a7-3a-39-3d-6b, 8000, 8000, 101)
    state:0x0f (Ac-1:To-1:Ag-1:Sy-1:Co-0:Di-0:De-0:Ex-0)
Partner:
    info: (8000, f8-a7-3a-2e-a3-f, 8000, 8000, 101)
    state:0x0f (Ac-1:To-1:Ag-1:Sy-1:Co-0:Di-0:De-0:Ex-0

(2) Recv LACP PDU: len:124 at 708749 usecs after Tue Aug 15 00:02:12 2023
0180c200 0002f8a7 3a2ea310 88090101 01148000 f8a73a2e a30f8000 80000101
3d000000 02148000 f8a73a39 3d6b8000 80000101 3d000000 03100000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000 00000000
00000000 00000000 00000000 00000000 00000000 00000000 00000000
Actor:
    info: (8000, f8-a7-3a-2e-a3-f, 8000, 8000, 101)
    state:0x0f (Ac-1:To-1:Ag-1:Sy-1:Co-0:Di-0:De-0:Ex-0)
Partner:
    info: (8000, f8-a7-3a-39-3d-6b, 8000, 8000, 101)
    state:0x07 (Ac-1:To-1:Ag-1:Sy-0:Co-0:Di-0:De-0:Ex-0)
```



注意：请考虑LACP PDU事件历史记录时间线是从较新的日志转到较旧的日志，但是LACP会说明事件历史记录时间线是从较旧的日志转到较新的日志。

配置和验证LACP Ethanalyzer

LACP PDU需要由CPU处理（Cisco IOS软件），nexus已安装在硬件内部访问列表中，以将LACP数据包重定向到CPU，所有LACP PDU都可通过Ethanalyzer进行观察，以便使用wireshark过滤器“slow”对其进行过滤。

```
ethanalyzer local interface inband display-filter "slow and eth.addr==04:76:b0:b2:00:20 and eth.addr==0
Capturing on inband
2023-07-03 23:37:14.420839 04:76:b0:b2:00:20 -> 01:80:c2:00:00:02 LACP Link Aggregation Control Protoco

Detailed:


Frame 19 (124 bytes on wire, 124 bytes captured)
```

```
    Arrival Time: Jul 3, 2023 23:38:14.425502000
    [Time delta from previous captured frame: 0.836575000 seconds]
    [Time delta from previous displayed frame: 11.246799000 seconds]
    [Time since reference or first frame: 11.246799000 seconds]
    Frame Number: 19
    Frame Length: 124 bytes
    Capture Length: 124 bytes
    [Frame is marked: False]
    [Protocols in frame: eth:slow]
Ethernet II, Src: 04:76:b0:b2:00:20 (04:76:b0:b2:00:20), Dst: 01:80:c2:00:00:02 (01:80:c2:00:00:02)
    Destination: 01:80:c2:00:00:02 (01:80:c2:00:00:02)
        Address: 01:80:c2:00:00:02 (01:80:c2:00:00:02)
        .... ...1 .... .... .... .... = IG bit: Group address (multicast/broadcast)
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)
    Source: 04:76:b0:b2:00:20 (04:76:b0:b2:00:20)
        Address: 04:76:b0:b2:00:20 (04:76:b0:b2:00:20)
        .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory default)

    Type: Slow Protocols (0x8809)
Link Aggregation Control Protocol
    Slow Protocols subtype: LACP (0x01)
    LACP Version Number: 0x01
    Actor Information: 0x01
    Actor Information Length: 0x14
    Actor System Priority: 32768
    Actor System: 04:76:b0:b2:00:1f (04:76:b0:b2:00:1f)
    Actor Key: 32768
    Actor Port Priority: 32768
    Actor Port: 257
    Actor State: 0x3d (Activity, Aggregation, Synchronization, Collecting, Distributing)
        .... ...1 = LACP Activity: Yes
        .... ..0. = LACP Timeout: No
        .... .1.. = Aggregation: Yes
        .... 1... = Synchronization: Yes
        ...1 .... = Collecting: Yes
        ..1. .... = Distributing: Yes
        .0.. .... = Defaulted: No
        0... .... = Expired: No
    Reserved: 000000
    Partner Information: 0x02
    Partner Information Length: 0x14
    Partner System Priority: 32768
    Partner System: 70:0f:6a:d7:d0:fb (70:0f:6a:d7:d0:fb)
    Partner Key: 32768
    Partner Port Priority: 32768
    Partner Port: 449
    Partner State: 0x3d (Activity, Aggregation, Synchronization, Collecting, Distributing)
        .... ...1 = LACP Activity: Yes
        .... ..0. = LACP Timeout: No
        .... .1.. = Aggregation: Yes
        .... 1... = Synchronization: Yes
        ...1 .... = Collecting: Yes
        ..1. .... = Distributing: Yes
        .0.. .... = Defaulted: No
        0... .... = Expired: No
    Reserved: 000000
    Collector Information: 0x03
    Collector Information Length: 0x10
    Collector Max Delay: 0
    Reserved: 000000000000000000000000
    Terminator Information: 0x00
```

```
    Terminator Length: 0x00
    Reserved: 00000000000000000000000000000000000000000000000000...
```

配置并检验LACP管理器

需要识别远程接口中的第一个mac地址：

```
N9K2#sh int e1/1 | i i addr
Hardware: 1000/10000/25000/40000/50000/100000 Ethernet, address: 0476.b0b2.0020 (bia 0476.b0b2.0020)
```

现在，在nexus N9K1上配置了elam。

```
N9K1#debug platform internal tah elam
N9K1(TAH-elam)#trigger init
N9K1(TAH-elam-insel6)#set outer l2 src_mac 04:76:b0:b2:00:20 dst_mac 01:80:c2:00:00:02 <<<<<Dest mac is
N9K1(TAH-elam-insel6)#tart
N9K1(TAH-elam-insel6)#report
SUGARBOWL ELAM REPORT SUMMARY
slot - 1, asic - 0, slice - 1
============================Incoming Interface: Eth1/49
Src Idx : 0x601, Src BD : 1
Outgoing Interface Info: met_ptr 0Packet Type: CE

Dst MAC address: 01:80:C2:00:00:02
Src MAC address: 04:76:B0:B2:00:20
Sup hit: 1, Sup Idx: 2627. <<<<<Traffic needs to be punted to the CPU.
Drop Info:
----------LUA:
LUB:
LUC:
LUD:
Final Drops:vntag:
vntag_valid : 0
vntag_vir : 0
vntag_svif : 0
```

要对sup重定向索引进行解码，可以执行sh system internal access-list sup-redirect-stats：

```
sh system internal access-list sup-redirect-stats | i i 2627
2627 LACP 0
2627 LACP 103
```

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）。