

使用Netplan配置LACP C885A-M8网卡

目录

[简介](#)

[背景信息](#)

[先决条件](#)

[使用的组件](#)

[使用的网卡](#)

[验证NIC未处于DPU模式](#)

[步骤 1：确定PCIe地址](#)

[步骤 2：检查NIC配置](#)

[步骤 3：转换为NIC模式](#)

[步骤 4：重新启动并检验配置](#)

[配置LACP](#)

[步骤 1：创建新的Netplan配置文件](#)

[步骤 2：应用配置](#)

[验证](#)

[1.检查网络接口详细信息](#)

[2.确认绑定详细信息](#)

[故障排除](#)

[相关信息](#)

简介

本文档介绍如何在C885A-M8服务器上使用Netplan配置LACP的信息。

背景信息

在BMC中，无法在网络侧进行任何修改，所有更改都需要在OS Guest中完成。

先决条件

使用的组件

UCS C885A M8机架式服务器

PID:UCSC-885A-M8-H11

固件版本:1.0.28

Nexus 9k交换机

Ubuntu 22.04.5 LTS

使用的网卡

BlueField-3 P系列DPU 400GbE/NDR VPI双端口

BlueField-3 E系列SuperNIC 400GbE/NDR单端口

用于OCP 3.0的Intel(R)以太网网络适配器X710-T2L

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。

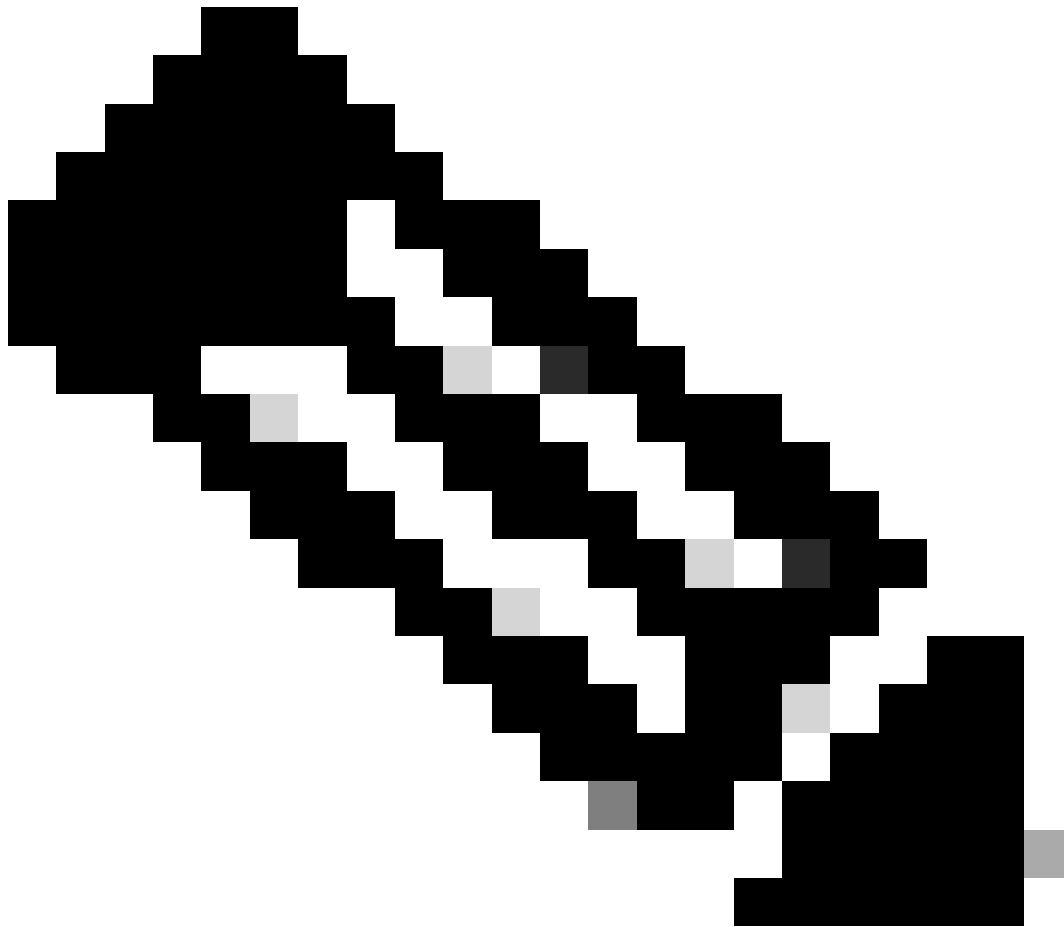
验证NIC未处于DPU模式

为确保功能正常，请确认网卡未处于DPU模式。当网卡处于DPU模式时，不支持链路汇聚控制协议(LACP)。如有必要，将其转换为网卡模式。

步骤 1：确定PCIe地址

运行命令列出所需端口的PCIe地址：

```
cisco-tac@C885A-M8:~$ sudo mst start
Starting MST (Mellanox Software Tools) driver set
Loading MST PCI module - Success
Loading MST PCI configuration module - Success
Create devices
Unloading MST PCI module (unused) - Success
```



注意：这不是必需的，但有助于在使用mlxconfig进行更改时指定正确的设备

```
cisco-tac@C885A-M8:~$ sudo mst status -v
MST modules:
-----
MST PCI module is not loaded
MST PCI configuration module loaded
PCI devices:
-----
DEVICE_TYPE MST PCI RDMA NET NUMA
BlueField3(rev:1) /dev/mst/mt41692_pciconf9 f1:00.0 mlx5_6 net-ens205f0 1
BlueField3(rev:1) /dev/mst/mt41692_pciconf8 c5:00.0 mlx5_7 net-ens206f0 1
BlueField3(rev:1) /dev/mst/mt41692_pciconf7 a4:00.0 mlx5_11 net-ens208f0 1
BlueField3(rev:1) /dev/mst/mt41692_pciconf6 97:00.0 mlx5_10 net-ens207f0 1
BlueField3(rev:1) /dev/mst/mt41692_pciconf5.1 91:00.1 mlx5_9 net-ens214f1np1 1
BlueField3(rev:1) /dev/mst/mt41692_pciconf5 91:00.0 mlx5_8 net-ens214f0np0 1
BlueField3(rev:1) /dev/mst/mt41692_pciconf4 69:00.0 mlx5_0 net-ens202f0 0
BlueField3(rev:1) /dev/mst/mt41692_pciconf3 4b:00.0 mlx5_3 net-ens201f0 0
BlueField3(rev:1) /dev/mst/mt41692_pciconf2.1 45:00.1 mlx5_2 net-ens214f1np1 0
BlueField3(rev:1) /dev/mst/mt41692_pciconf2 45:00.0 mlx5_1 net-ens211f0np0 0
```

```
BlueField3(rev:1) /dev/mst/mt41692_pciconf1 2b:00.0 mlx5_5 net-ens203f0 0
BlueField3(rev:1) /dev/mst/mt41692_pciconf0 09:00.0 mlx5_4 net-ens204f0 0
```

此命令提供详细的设备列表，允许您识别适当的MST和PCI标识符。

在此场景中，使用NET net-ens214f1np1和NET net-ens211f0np0。

步骤 2：检查NIC配置

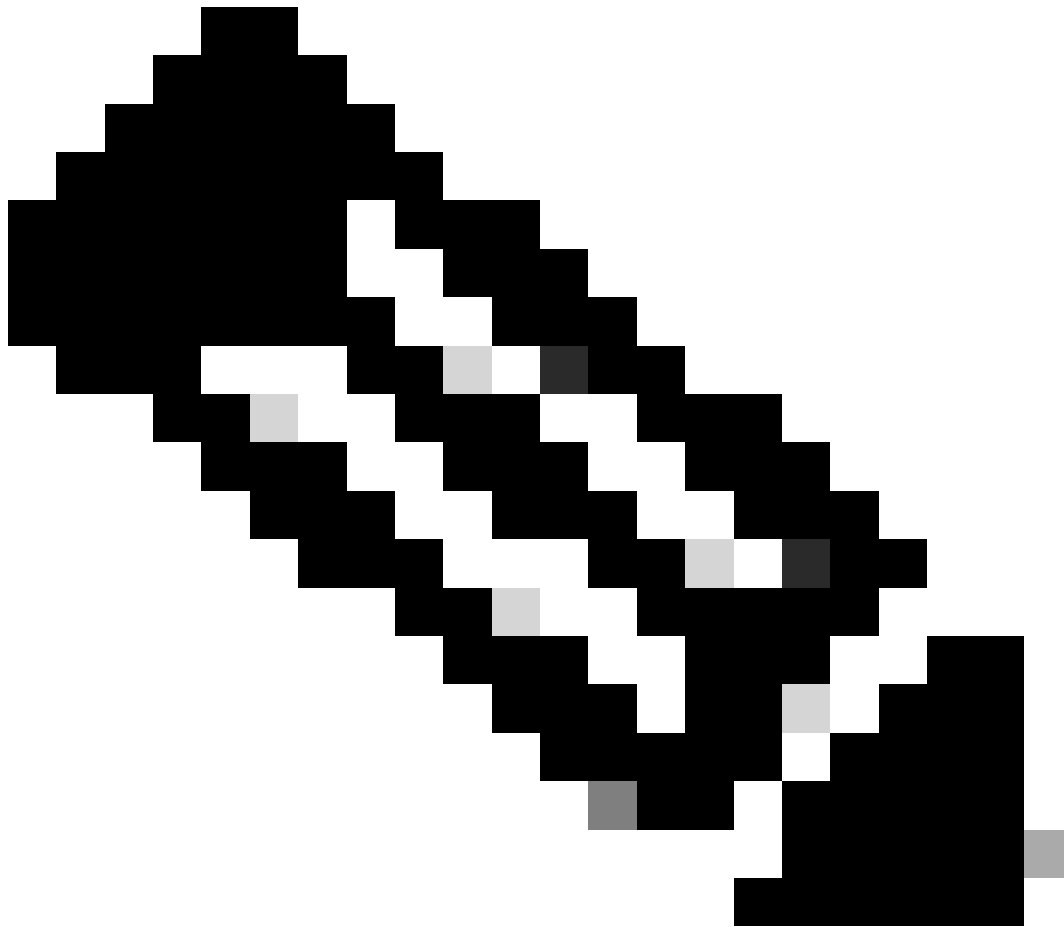
使用已识别的MST或PCI标识符，使用以下命令之一验证当前的NIC配置：

```
sudo mlxconfig -d 0000:<PCI_ID> -e q | grep 'INTERNAL_CPU_MODEL\|EXP_ROM_UEFI_ARM_ENABLE\|INTERNAL_CPU_
```

或

```
sudo mlxconfig -d /dev/<MST_DEVICE> -e q | grep 'INTERNAL_CPU_MODEL\|EXP_ROM_UEFI_ARM_ENABLE\|INTERNAL_
```

```
cisco-tac@C885A-M8:~$ sudo sudo mlxconfig -d dev/mst/mt41692_pciconf1 -e q | grep \ 'INTERNAL_CPU_MODEL
Configurations: Default Current Next Boot
INTERNAL_CPU_MODEL EMBEDDED_CPU(1) EMBEDDED_CPU(1) EMBEDDED_CPU(1)
INTERNAL_CPU_OFFLOAD_ENGINE ENABLED(0) ENABLED(0) ENABLED(0)
EXP_ROM_UEFI_ARM_ENABLE True(1) True(1) True(1)
```



注意：CPU模型和ARM UEFI使能变量向我们确认卡支持DPU模式

在上面的示例中：

INTERNAL_CPU_OFFLOAD_ENGINE设置为ENABLED\0\，适用于当前和下次引导配置。

这表示网卡当前在DPU模式下运行，需要转换为NIC模式才能继续LACP配置。

步骤 3：转换为NIC模式

要将NIC转换为NIC模式，请发出命令：

```
sudo mlxconfig -d 0000:<PCI_ID> s INTERNAL_CPU_OFFLOAD_ENGINE=0
```

```
cisco-tac@C885A-M8:~$ sudo mlxconfig -d 0000:45:00.0 s INTERNAL_CPU_OFFLOAD_ENGINE=0
```

```
Device #1:
```

```
-----
```

```
Device type: BlueField3
```

```
Name: 900-9D3B6-00SN-A_Ax
```

```
Description: NVIDIA BlueField-3 B3240 P-Series Dual-slot FHHL DPU; 400GbE / NDR IB (default mode); Dual
```

```
Device: 0000:45:00.0
```

```
Configurations: Next Boot New
```

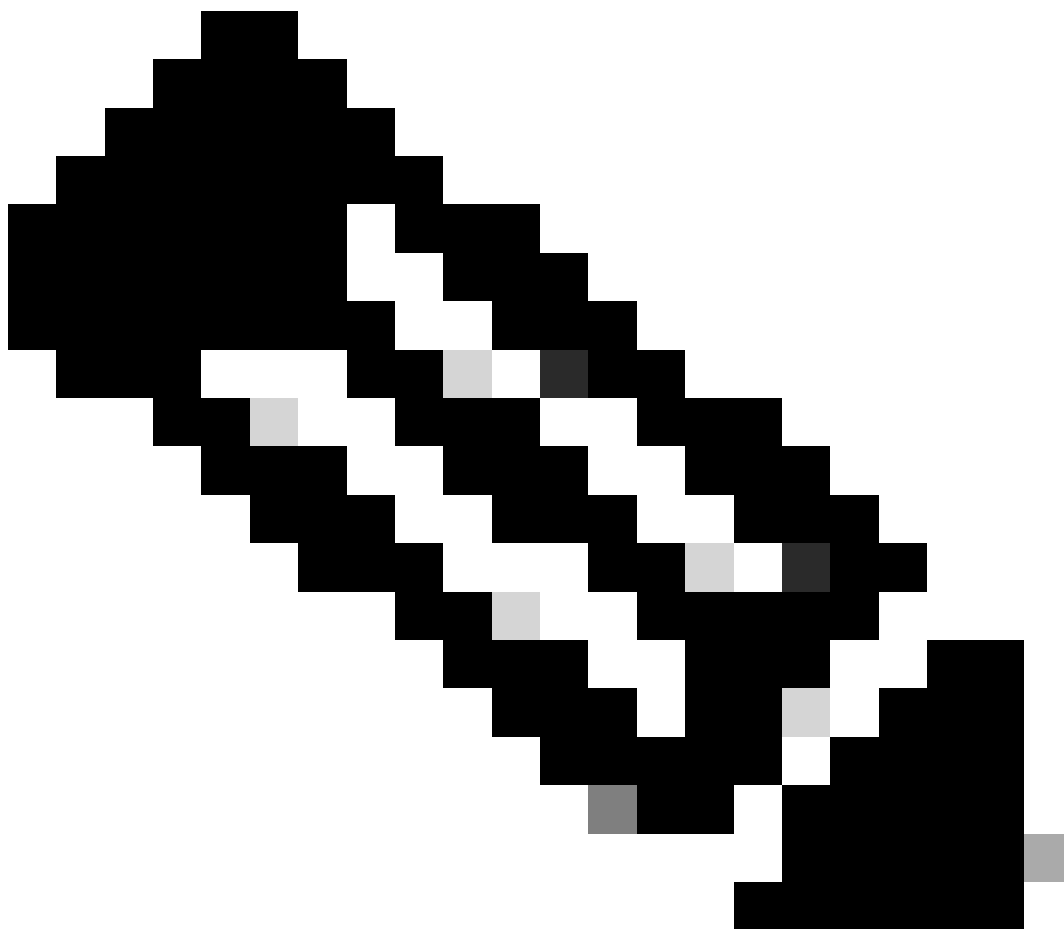
```
INTERNAL_CPU_OFFLOAD_ENGINE ENABLED(0) DISABLED(1)
```

```
Apply new Configuration? (y/n) [n] : y
```

```
Applying... Done!
```

```
-I- Please reboot machine to load new configurations.
```

```
cisco-tac@C885A-M8:~$ sudo init 0
```



注意：使用sudo init 0，关闭主机以确保正确加载配置。

步骤 4：重新启动并检验配置

执行sudo init 0命令后，主机系统将关闭。使用底板管理控制器(BMC)重新启动主机操作系统。

系统重新联机后，请运行前面所示的相同验证命令，以验证配置更改是否已成功应用：

```
sudo mlxconfig -d 0000:<PCI_ID> -e q | grep 'INTERNAL_CPU_MODEL\|EXP_ROM_UEFI_ARM_ENABLE\|INTERNAL_CPU_
```

或

```
sudo mlxconfig -d /dev/<MST_DEVICE> -e q | grep 'INTERNAL_CPU_MODEL\|EXP_ROM_UEFI_ARM_ENABLE\|INTERNAL_
```

如果配置已成功应用，您可以继续创建所需的端口通道、绑定或LAG配置。

配置LACP

步骤 1：创建新的Netplan配置文件

在Ubuntu上配置接口的当前标准是Netplan。首先，请运行以下命令检查现有的Netplan YAML配置文件：

```
ls /etc/netplan/
```

默认情况下，/etc/netplan/目录中通常只有一个YAML文件。Netplan配置文件使用命名约定##-<relevant name>.yaml，其中##是介于01和99之间的数字。请考虑在存在冲突时，编号较小的配置文件会被编号较高的文件覆盖。

为确保更改具有持久性和组织性，建议创建新的YAML文件，而不是编辑默认文件。如果修改默认文件，则存在重新引导后更改不会持续的风险。在本指南中，我们将创建一个名为

51-lACP-config.yaml。

要创建新文件，请使用命令：

```
sudo nano /etc/netplan/51-lACP-config.yaml
```

确保文件编号不会与现有文件冲突，并使用适当的层次结构来避免意外行为。

这是用于设置VLAN的LACP的Netplan YAML配置文件(51-lACP-config.yaml)的示例。

```
cisco-tac@C885A-M8:~$ cat /etc/netplan/51-lACP-config.yaml
network:
  ethernets:
    ens211f0np0:
      mtu: 9000
    ens211f1np1:
      mtu: 9000
  bonds:
    bond0:
      mtu: 9000
  dhcp4: false
  dhcp6: false
```



```
interfaces:
- ens211f0np0
- ens211f1np1
parameters:
mode: 802.3ad
mii-monitor-interval: 100
vlans:
bond0.2001:
id: 2001
link: bond0
```

价值	考虑事项	备注
MTU设置	对于ens211f0np0和ens211f1np1,MTU设置为9000。	建议将MTU设置为避免将来出现潜在MTU问题所需的最高值。每当提到接口时，始终指定MTU
LACP 模式	模式参数设置为802.3ad	这是LACP所需的模式。存在其他键合模式，但不适用于LACP
VLAN 配置	bond0.2001部分创建子接口，用于通过绑定发送多个VLAN。id字段指定VLAN ID（本例中为2001）。链路字段将子接口与绑定接口(bond0)关联	如果您在同一个YAML文件中有多个绑定，这可确保将正确的配置应用到正确的绑定
IP地址和路由	地址字段指定VLAN的IP地址和子网掩码(例如10.10.200.33/27)	Routes部分允许您定义此接口处理的所有路由。尽管示例包括10.10.200.32/27的路由，但在本实验设置中它是冗余的，但演示了如何添加多个路由
DNS Servers	nameservers部分允许您指定DNS服务器（地址）和搜索域（搜索）	需要根据工作环境提供DNS服务器详细信息。
VLAN的MTU	VLAN配置中的mtu字段设置为9000	如果需要，您可以逐个VLAN配置不同的MTU值
版本字段	文件末尾的version字段可以包含并设置为2	确保使用两个空格缩进以维护YAML语法

步骤 2：应用配置

要应用新配置，请运行：

```
sudo netplan apply
```

如果遇到权限警告，请调整文件权限：

```
sudo chmod 600 /etc/netplan/51-lACP-config.yaml
sudo netplan apply
```

```
cisco-tac@C885A-M8:~$ sudo netplan apply
** (generate:140701): WARNING **: 17:10:58.313: Permissions for /etc/netplan/51-lACP-config.yaml are too
```

```
cisco-tac@C885A-M8:~$ ls -l /etc/netplan/
total 16
-rw----- 1 root root 869 Apr 24 14:26 50-cloud-init.yaml
-rw-r--r-- 1 root root 586 Apr 25 17:10 51-lACP-config.yaml
```

```
cisco-tac@C885A-M8:~$ sudo chmod 600 /etc/netplan/51-lACP-config.yaml
cisco-tac@C885A-M8:~$ ls -l /etc/netplan/
total 16
-rw----- 1 root root 869 Apr 24 14:26 50-cloud-init.yaml
-rw----- 1 root root 586 Apr 25 17:10 51-lACP-config.yaml
```

```
cisco-tac@C885A-M8:~$ sudo netplan apply
```

验证

应用Netplan配置并重新启动系统后，使用命令确认LACP设置是否按预期运行。

1.检查网络接口详细信息

使用此命令可显示网络接口的状态：

```
cisco-tac@C885A-M8:~$ ip link show
4: ens211f0np0: <BROADCAST,MULTICAST,SLAVE,UP,LOWER_UP> mtu 9000 qdisc mq master bond0 state UP mode DE
link/ether 86:a2:e3:01:3d:f2 brd ff:ff:ff:ff:ff:ff permaddr 5c:25:73:3a:9a:fc
altname enp69s0f0np0
5: ens211f1np1: <BROADCAST,MULTICAST,SLAVE,UP,LOWER_UP> mtu 9000 qdisc mq master bond0 state UP mode DE
link/ether 86:a2:e3:01:3d:f2 brd ff:ff:ff:ff:ff:ff permaddr 5c:25:73:3a:9a:fd
altname enp69s0f1np1
```

确保接口 (ens211f0np0和ens211f1np1) 显示：

- 主设备设置为bond0，确认它们属于绑定接口。
- 状态为UP，确认接口处于活动状态。
- 确认mtu值已设置为9000（在YAML文件中指定）。
- 检验两个接口是否都具有有效的硬件地址(permaddr)。

2.确认绑定详细信息

使用此命令可显示绑定配置并确保LACP正常工作：

```
cisco-tac@C885A-M8:~$cat /proc/net/bonding/bond0
Ethernet Channel Bonding Driver: v5.15.0-141-generic
```

```
Bonding Mode: IEEE 802.3ad Dynamic link aggregation
Transmit Hash Policy: layer2 (0)
MII Status: up
MII Polling Interval (ms): 100
Up Delay (ms): 0
Down Delay (ms): 0
Peer Notification Delay (ms): 0
```

```
802.3ad info
LACP active: on
LACP rate: slow
Min links: 0
Aggregator selection policy (ad_select): stable
```

```
Slave Interface: ens211f1np1
MII Status: up
Speed: ***** Mbps
Duplex: full
Link Failure Count: 0
Permanent HW addr: 5c:25:73:3a:9a:fd
Slave queue ID: 0
Aggregator ID: 1
Actor Churn State: none
Partner Churn State: none
Actor Churned Count: 0
Partner Churned Count: 0
```

```
Slave Interface: ens211f0np0
MII Status: up
Speed: ***** Mbps
Duplex: full
Link Failure Count: 0
Permanent HW addr: 5c:25:73:3a:9a:fc
Slave queue ID: 0
Aggregator ID: 1
Actor Churn State: none
Partner Churn State: none
Actor Churned Count: 0
Partner Churned Count: 0
```

确保绑定模式设置为IEEE 802.3ad动态链路聚合：

- LACP活动：亮起：表示已启用LACP。
- LACP速率：slow反映配置的LACP速率，尽管它可能因设置而异。
- 检验从接口：ens211f0np0和ens211f1np1都可以显示为绑定下的从接口。
- MII状态：up确认链路处于活动状态。
- 永久硬件地址值必须与相应接口的硬件地址匹配。
- 确保两个从接口属于同一个聚合器ID(例如，聚合器ID:1)。

故障排除

如果验证结果不符合预期：

- 检查Netplan YAML配置文件是否存在错误或缺少参数。
- 确保上游交换机上正确LACP支持和配置。
- 确认sudo netplan apply命令已成功完成

相关信息

[TACDCN-2018](#)

[Cisco UCS C885A M8机架式服务器产品手册](#)

[UCS C885A M8机架式服务器规格表](#)

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）。