

# 排除HX节点PCI地址重新排序故障

## 目录

[简介](#)

[背景信息](#)

[VMWare OS系统影响](#)

[分辨率](#)

[解决方法](#)

[缺陷](#)

## 简介

本文档介绍各种第3代C系列虚拟接口卡(VIC)适配器上的网络接口卡(NIC)重新排序问题。通过参考虚拟网络接口卡(vNIC)/虚拟主机总线适配器(vHBA)的放置、设备顺序、外围组件互联(PCI)枚举，注意到行为变化，并且从各种操作系统(OS)中看到PCI重新排序问题。

## 背景信息

此问题的原因是：

- 添加/删除vNIC
- 重新启动服务器
- 服务器升级

此问题是由NIC枚举引起的，该枚举在您应用BIOS、固件、补丁升级或为添加新NIC而发生。根据虚拟机网络接口控制器(vMNIC)名称配置向外部世界提供网络通信的虚拟交换机(vSwitch)。如果所有NIC都重命名，vSwitch会将数据包路由到不再存在的接口。

统一计算系统管理器(UCSM)在处理VIC13XX主机端口时遵循标准负载均衡算法。

请注意：

1. vNIC的总数按其相对容量比例分配。
2. 系统尝试在具有更高vNIC容量的适配器卡上放置更多vNIC，在具有较低容量的适配器卡上放置更少vNIC。因此，系统将实际放置适当地分配给vNIC，vNIC又被映射到预期的外围组件互联枚举(PCIe)适配器卡。此位置按其实际顺序进行。
3. 在混合夹层设置中，只要所有适配器卡都支持所需的功能，负载分配算法就可确保每个适配器卡上至少放置一个vNIC和一个vHBA。

## VMWare OS系统影响

PCI ID到vMNIC编号关系在启动时确定，并自动输入到位于/etc/vmware/的esx.conf文件中，以保持持久性。ESX/ESXi主机首先扫描seg编号，然后扫描总线编号、插槽编号，最后扫描功能编号。此顺序可确保同一多端口网卡上的端口按顺序编号。

安装ESXi/ESX时，VMNIC顺序是连续的。随着NIC的删除和其他NIC的添加，这种情况会随时间而改变。这会导致VMNIC顺序不理想，并且与其他ESX/ESXi主机的命名约定不同步。

ESXi枚举VMNIC和vHBA设备时，并不总是遵循PCI地址顺序。在初始安装时，会遵循PCI地址。但是，随着设备的删除和添加，有一种特定的算法来处理顺序。这可能导致VNIC/vHBA订单不正确或服务配置文件与操作系统之间不同步。

请参阅VMware KB:<http://kb.vmware.com/kb/2019871>

ESXi 5.5和6.0已更改了其顺序算法，即使这些症状持续存在。请参阅：<http://kb.vmware.com/kb/2091560>

## 分辨率

在初始安装时，新名称方案会尝试按物理位置以可预测的顺序分配名称。之后，即使在系统中添加或移除其他设备时，它也能保持设备名称稳定。在安装时，通常集成在系统主板上的设备会按其输出端口在系统机箱上编号的顺序接收编号最低的名称，而可插拔插槽中的设备会按插槽在机箱上编号的顺序接收编号最高的名称。读者应注意，此过程会忽略找不到驱动程序的设备。

## 解决方法

此问题可通过以下解决方案解决：

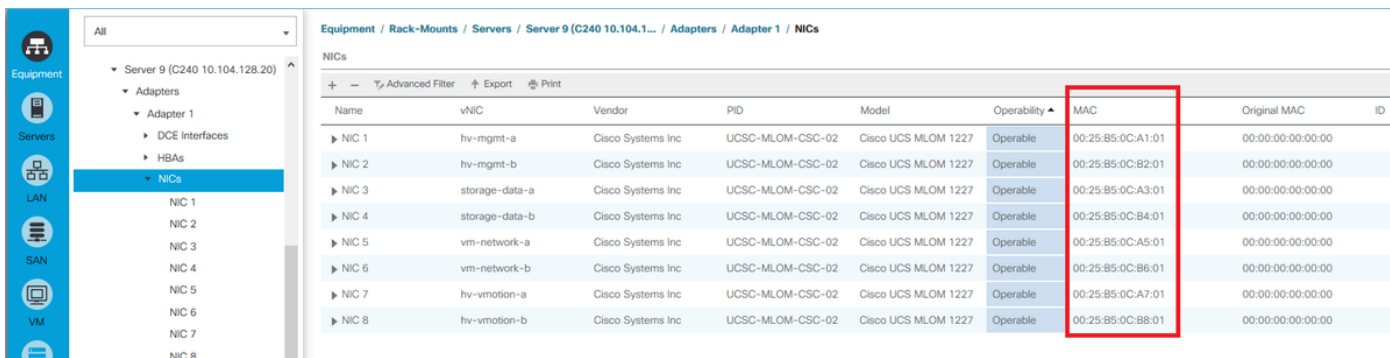
步骤1.重新安装ESXi。

步骤2.手动更正ESXi安全外壳(SSH)的重新排序。

步骤3.在UCSM中创建vNIC/vHBA放置策略。

步骤2.手动更正ESXi SSH的重新排序：

1.从UCSM GUI获取服务器的MAC地址映射，如图所示：



Name	vNIC	Vendor	PID	Model	Operability	MAC	Original MAC	ID
NIC 1	hv-mgmt-a	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:85:0C:A1:01	00:00:00:00:00:00	
NIC 2	hv-mgmt-b	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:85:0C:B2:01	00:00:00:00:00:00	
NIC 3	storage-data-a	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:85:0C:A3:01	00:00:00:00:00:00	
NIC 4	storage-data-b	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:85:0C:B4:01	00:00:00:00:00:00	
NIC 5	vm-network-a	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:85:0C:A5:01	00:00:00:00:00:00	
NIC 6	vm-network-b	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:85:0C:B6:01	00:00:00:00:00:00	
NIC 7	hv-vmotion-a	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:85:0C:A7:01	00:00:00:00:00:00	
NIC 8	hv-vmotion-b	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:85:0C:B8:01	00:00:00:00:00:00	

MAC
00:25:B5:0C:A1:01
00:25:B5:0C:B2:01
00:25:B5:0C:A3:01
00:25:B5:0C:B4:01
00:25:B5:0C:A5:01
00:25:B5:0C:B6:01
00:25:B5:0C:A7:01
00:25:B5:0C:B8:01

2.从ESXi CLI获取MAC地址映射。

```
[root@ucs-116:~] esxcli network nic list
[root@ucs-116:~] esxcli network nic list
Name      PCI Device  Driver  Admin Status  Link Status  Speed  Duplex  MAC Address  MTU  Description
-----
vmnic0    0000:05:00.0  enic   Up           Up           10000  Full    00:25:b5:0c:a1:01  1500 Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic1    0000:06:00.0  enic   Up           Up           10000  Full    00:25:b5:0c:b2:01  1500 Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic2    0000:09:00.0  enic   Up           Up           10000  Full    00:25:b5:0c:a5:01  9000 Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic3    0000:08:00.0  enic   Up           Up           10000  Full    00:25:b5:0c:b4:01  9000 Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic4    0000:07:00.0  enic   Up           Up           10000  Full    00:25:b5:0c:a3:01  1500 Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic5    0000:0a:00.0  enic   Up           Up           10000  Full    00:25:b5:0c:b6:01  1500 Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic6    0000:0b:00.0  enic   Up           Up           10000  Full    00:25:b5:0c:a7:01  9000 Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic7    0000:0c:00.0  enic   Up           Up           10000  Full    00:25:b5:0c:b8:01  9000 Cisco Systems Inc Cisco VIC Ethernet NIC
[root@ucs-116:~]
```

```
MAC Address
-----
00:25:b5:0c:a1:01
00:25:b5:0c:b2:01
00:25:b5:0c:a5:01
00:25:b5:0c:b4:01
00:25:b5:0c:a3:01
00:25:b5:0c:b6:01
00:25:b5:0c:a7:01
00:25:b5:0c:b8:01
```

3.验证UCSM的MAC地址是否与esxcli网络nic列表输出匹配。

在此输出中，MAC地址不匹配，vmnic2和vmnic4被交换。这是由于PCI总线重新订购问题。

4.检查esxcli上的PCI总线排序。

```
localcli --plugin-dir /usr/lib/vmware/esxcli/int/ deviceInternal alias list | grep -i vmnic | sort -k3
[root@ucs-116:~] localcli --plugin-dir /usr/lib/vmware/esxcli/int/ deviceInternal alias list | grep -i vmnic | sort -k3
pci      s00000000:03.00      vmnic0
pci      s00000000:04.00      vmnic1
pci      s00000000:07.00      vmnic2
pci      s00000000:06.00      vmnic3
pci      s00000000:05.00      vmnic4
pci      s00000000:08.00      vmnic5
pci      s00000000:09.00      vmnic6
pci      s00000000:0a.00      vmnic7
[root@ucs-116:~] █
```

vmnic2和vmnic4之间的PCI总线已交换。

5.记下PCI总线编号，并使用以下命令来修复重新排序：

```
localcli --plugin-dir /usr/lib/vmware/esxcli/int/ deviceInternal alias store --alias vmnic2 --bus-address s00000000:05.00 --bus-type pci
localcli --plugin-dir /usr/lib/vmware/esxcli/int/ deviceInternal alias store --alias vmnic4 --bus-address s00000000:07.00 --bus-type pci
[root@ucs-116:~] localcli --plugin-dir /usr/lib/vmware/esxcli/int/ deviceInternal alias store --alias vmnic2 --bus-address s00000000:05.00 --bus-type pci
[root@ucs-116:~] localcli --plugin-dir /usr/lib/vmware/esxcli/int/ deviceInternal alias store --alias vmnic4 --bus-address s00000000:07.00 --bus-type pci
[root@ucs-116:~] █
```

6.重新启动，主机启动后，验证PCI订购和网卡列表：

```
[root@ucs-116:~]
[root@ucs-116:~] reboot
[root@ucs-116:~] █
```

```
[root@ucs-116:~] localcli --plugin-dir /usr/lib/vmware/esxcli/int/ deviceInternal alias list | grep -i vmnic | sort -k3
pci      s00000000:03.00      vmnic0
pci      s00000000:04.00      vmnic1
pci      s00000000:05.00      vmnic2
pci      s00000000:06.00      vmnic3
pci      s00000000:07.00      vmnic4
pci      s00000000:08.00      vmnic5
pci      s00000000:09.00      vmnic6
pci      s00000000:0a.00      vmnic7
[root@ucs-116:~]
[root@ucs-116:~]
[root@ucs-116:~] esxcli network nic list
Name      PCI Device      Driver      Admin Status      Link Status      Speed      Duplex      MAC Address      MTU      Description
-----
vmnic0    0000:05:00.0    enic       Up                Up                10000     Full       00:25:b5:0c:a1:01 1500     Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic1    0000:06:00.0    enic       Up                Up                10000     Full       00:25:b5:0c:b2:01 1500     Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic2    0000:07:00.0    enic       Up                Up                10000     Full       00:25:b5:0c:a3:01 9000     Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic3    0000:08:00.0    enic       Up                Up                10000     Full       00:25:b5:0c:b4:01 9000     Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic4    0000:09:00.0    enic       Up                Up                10000     Full       00:25:b5:0c:a5:01 1500     Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic5    0000:0a:00.0    enic       Up                Up                10000     Full       00:25:b5:0c:b6:01 1500     Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic6    0000:0b:00.0    enic       Up                Up                10000     Full       00:25:b5:0c:a7:01 9000     Cisco Systems Inc Cisco VIC Ethernet NIC
vmnic7    0000:0c:00.0    enic       Up                Up                10000     Full       00:25:b5:0c:b8:01 9000     Cisco Systems Inc Cisco VIC Ethernet NIC
[root@ucs-116:~] █
```

您已成功更正PCI重新排序。

### 步骤3.在UCSM中创建vNIC/vHBA放置策略

在UCSM中，导航至服务配置文件>网络>修改vNIC/vHBA放置。

1.初始安装期间：导航至修改vNIC/vHBA放置策略，如图所示。

## Modify vNIC/vHBA Placement

Specify how vNICs and vHBAs are placed on physical network adapters

vNIC/vHBA Placement specifies how vNICs and vHBAs are placed on physical network adapters (mezzanine) in a server hardware configuration independent way.

Select Placement:

Virtual Network Interface connection provides a mechanism of placing vNICs and vHBAs on physical network adapters. vNICs and vHBAs are assigned to one of Virtual Network Interface connection specified below. This assignment can be performed explicitly by selecting which Virtual Network Interface connection is used by vNIC or vHBA or it can be done automatically by selecting "any". vNIC/vHBA placement on physical network interface is controlled by placement preferences.

Please select one Virtual Network Interface and one or more vNICs or vHBAs

Name	Order	Admin Host Port	Selection Preference
vCon 1			All
vNIC vnic0	1	ANY	
vNIC vnic1	2	ANY	
vCon 2			All
vCon 3			All
vCon 4			All

2.现在，您将管理主机端口视为**ANY**。这将是轮询分配，将vnic0 置于主机端口1，将vnic1 置于主机端口2。如果要向服务配置文件添加另外两个vNIC，请首先匹配已存在的vNIC的已分配主机端口：

Name	Order	Admin Host Port	Selection Preference
vCon 1			All
vNIC vnic0	1	1	
vNIC vnic1	2	2	

3.然后，只将任何新vNIC添加到主机端口2。这将保留现有vNIC的PCI地址，不会导致任何重新分配：

Name	Order	Admin Host Port	Selection Preference
vCon 1			All
vNIC vnic0	1	1	
vNIC vnic1	2	2	
vNIC vnic2	3	2	
vNIC vnic3	4	2	

**注意：**现在，主机端口1上有1个vNIC，主机端口2上有3个vNIC。负载均衡的点是，负载和带宽均匀分配。如果一个主机端口负载较高，则可能会影响性能/带宽。如果必须创建2个以上的

附加vNIC，建议您计划重新安装操作系统，以使主机端口均匀分布。

## 缺陷

- [CSCut78943](#) - VIC1340/1380，有错误的vNIC和vHBA放置顺序。
- [CSCuv19605](#) — 使用连接策略时无法为vNIC设置主机端口。

已提出增强缺陷。HX安装程序应创建vNIC/vHBA放置策略以避免PCI NIC重新排序。

**注意：**对vCON放置策略的修改（vCON或管理主机端口）可能导致不同的PCIe枚举更改。在vCON2 vNIC之前枚举vCON1 vNIC。