

Troubleshoot HX Nodes PCI Address Reorder

Contents

[Introduction](#)

[Background Information](#)

[VMWare OS System Impact](#)

[Resolution](#)

[Workaround](#)

[Defects](#)

Introduction

This document describes the Network Interface Card (NIC) reorder issue on various 3rd Gen C-series Virtual Interface Card (VIC) adapters. A behaviour change is noticed with reference to Virtual Network Interface Card (vNIC)/ Virtual Host Bus Adapter (vHBA) placement, device order, Peripheral Component Interconnect (PCI) enumeration, and contribute to PCI re-order issues are seen from various Operating Systems (OS).

Background Information

The issue is caused due to:

- Addition/Deletion of vNICs
- Reboot of the Server
- Upgrade of the Server

This issue is caused by NIC enumeration that occurs after you apply a BIOS, firmware, patch upgrade or in order to add new NIC's.

Virtual Switches (vSwitches) which provide network communication to the outside world are configured based on the Virtual Machine Network Interface Controller (VMNIC) names. If all of the NICs are renamed, the vSwitches route packets to interfaces that no longer exist.

Unified Computing System Manager (UCSM) follows the standard load balance algorithm when it deals with the VIC13XX host ports.

Note that:

1. The total number of vNICs gets distributed proportionally to their relative capacity ratio.
2. The system tries to place more vNICs on the adapter card with higher vNIC capacity and fewer vNICs on the adapter card with lower capacity. As a result, the system appropriately assigns the Actual Placement to the vNICs which in turn get mapped to the intended Peripheral Component Interconnect enumeration (PCIe) adapter card. This placement happens in order of their Actual Order.
3. In the case of Mixed-Mezz setup, so long as all the adapter cards support required capability, load distribution algorithm ensures that at least one vNIC and one vHBA would get placed on

each of the adapter cards.

VMWare OS System Impact

The PCI ID to VMNIC numbering relationship is determined at boot time and is automatically entered into the **esx.conf** file located at **/etc/vmware/** for persistence. The ESX/ESXi host first scans the seg number, then the bus number, the slot number, and finally the function number. This order ensures that ports on the same multi-port NIC are numbered sequentially.

When ESXi/ESX is installed, the VMNIC order is sequential. This changes over time as NICs are removed and other NICs are added. This results in VMNIC order that is undesirable and not in sync with the naming convention on other ESX/ESXi hosts.

ESXi does not always follow the PCI address order when it enumerates VMNIC and vHBA devices. On initial installation, PCI address is followed. However, as devices are removed and added, there is a specific algorithm which handles the order. This can cause vNIC/vHBA orders to become incorrect or out of sync between service-profile and the OS.

Refer to VMware KB: <http://kb.vmware.com/kb/2019871>

ESXi 5.5 and 6.0 has changed it's order algorithm, even though these symptoms continue. Refer to: <http://kb.vmware.com/kb/2091560>

Resolution

On initial installation, the new name scheme tries to assign names in a predictable order by a physical location. Thereafter, it keeps a device's name stable even across the addition or removal of other devices to the system. At the installation time, typically devices that are integrated on the system's motherboard receive the lowest numbered names, in the order that their output ports are numbered on the system's case, and devices that are in pluggable slots receive the next highest numbered names, in the order that the slots are numbered on the case. The reader should note that this process ignores devices for which no driver is found.

Workaround

The issue can be resolved by these solutions:

Step 1. Reinstallation of ESXi.

Step 2. Manually Correct the Re-order from the ESXi Secure Shell (SSH).

Step 3. Create a vNIC/vHBA placement policy in UCSM.

Step 2. Manually Correct the Re-order from the ESXi SSH :

1. Get the MAC address mapping from UCSM GUI for the server as shown in the image:

Equipment / Rack-Mounts / Servers / Server 9 (C240 10.104.1... / Adapters / Adapter 1 / NICs

NICs

Name	vNIC	Vendor	PID	Model	Operability	MAC	Original MAC	ID
NIC 1	hiv-mgmt-a	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:B5:0C:A1:01	00:00:00:00:00:00	
NIC 2	hiv-mgmt-b	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:B5:0C: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:B5:0C: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:B5:0C: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:B5:0C: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:B5:0C:B6:01	00:00:00:00:00:00	
NIC 7	hiv-vmotion-a	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:B5:0C:A7:01	00:00:00:00:00:00	
NIC 8	hiv-vmotion-b	Cisco Systems Inc	UCSC-MLOM-CSC-02	Cisco UCS MLOM 1227	Operable	00:25:B5:0C: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. Get the MAC address mapping from ESXi CLI.

```
[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. Verify whether the MAC addresses from UCSM matches with the **esxcli network nic** list output.

In this output, MAC addresses do not match, **vmnic2** and **vmnic4** got swapped. This is due to the PCI bus re-order issue.

4. Check the PCI bus ordering on the **esxcli**.

```
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:~]
```

PCI bus ordering between **vmnic2** and **vmnic4** got swapped.

5. Note the PCI bus number, and use these commands in order to fix the re-ordering:

```
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. Reboot and once the host comes up, verify the PCI ordering and **nic** list:

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

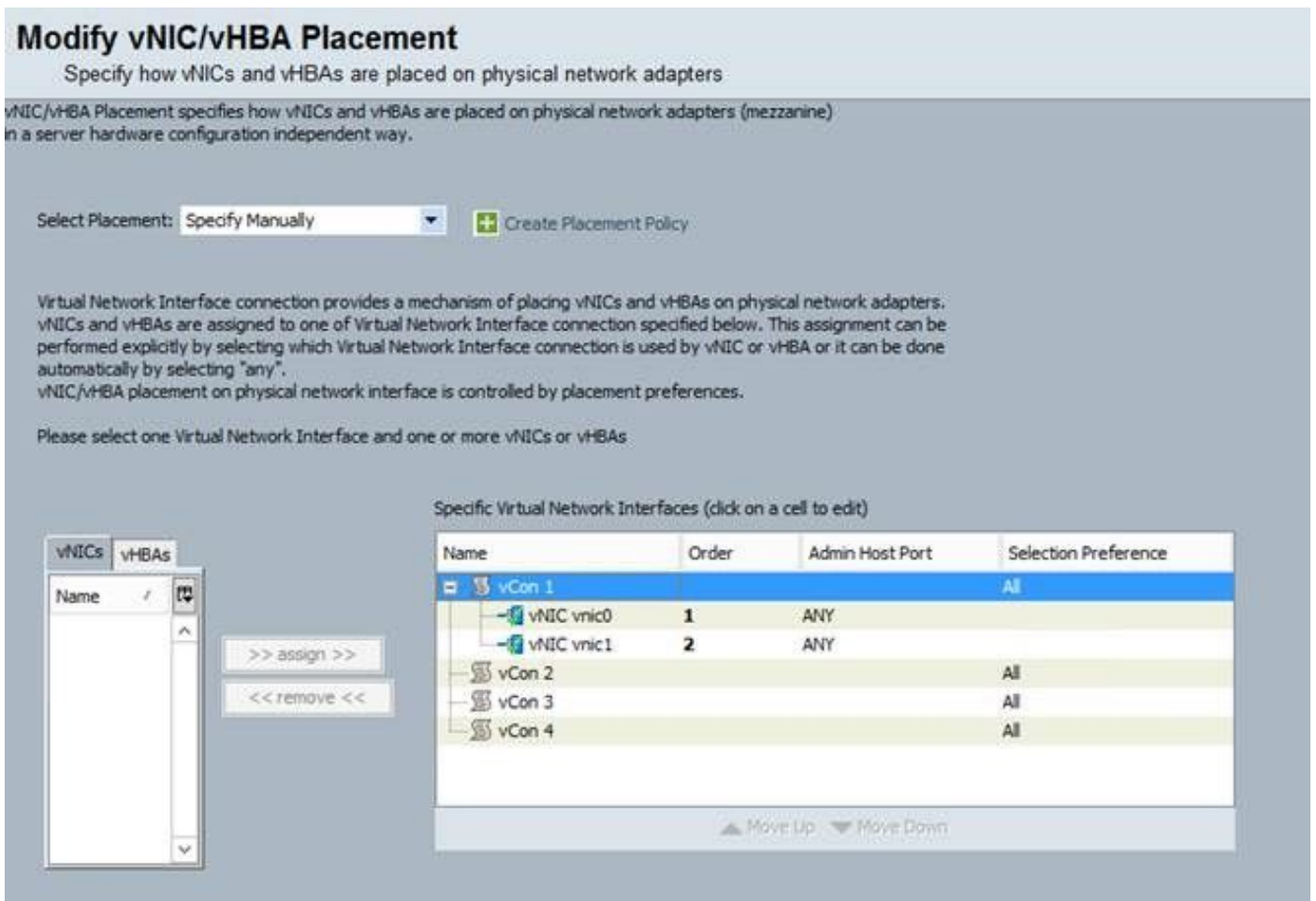
```
[root@ucs-116:~] localcli --plugin-dir /usr/lib/vmware/esxcli/int/ deviceInternal alias list | grep -i vmnic | sort -k3
pci      s00000000:03.00    vmnic0
pci      s00000000:04.00    vmnic1
pci      s00000000:05.00    vmnic2
pci      s00000000:06.00    vmnic3
pci      s00000000:07.00    vmnic4
pci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

You have successfully corrected the PCI re-ordering.

Step 3. Create a vNIC/vHBA Placement Policy in UCSM

In UCSM, navigate to **Service Profile > Network > Modify vNIC/vHBA Placement**.

1. During the initial install: navigate to **Modify vNIC/vHBA Placement** policy as shown in the image.



2. Right now you see **Admin Host Port** as **ANY**. This would be a round-robin assignment, that puts **vnic0** on Host Port 1, and **vnic1** on Host Port 2. If you want to add two additional vNICs to the service-profile, first match the assigned host-port of the vNICs that already exist:

Specific Virtual Network Interfaces (click on a cell to edit)

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

3. Then, add any new vNICs only to host-port 2. This will keep the PCI address of the pre-existing vNICs, and won't cause any re-assignment:

Specific Virtual Network Interfaces (click on a cell to edit)

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	

Note: You now have 1 vNIC on host port 1, and 3 vNICs on host port 2. The point of the load balancing is so that the load and bandwidth is evenly distributed. If you have a higher load on one host-port, it might cause a performance/bandwidth impact. If more than 2 additional vNICs must be created, it is recommended you plan a re-install of the OS in order to keep the host ports evenly distributed.

Defects

- [CSCut78943](#) - VIC1340/1380 with wrong vNIC and vHBA placement order.
- [CSCuv19605](#) - Unable to set host-port for vNICs when you use connectivity policy.

An enhancement defect has been filed. HX installer should create vNIC/vHBA placement policy in order to avoid PCI NIC re-ordering.

Note: Modifications to the **vCON** placement policy, either the vCON or the Admin Host Port might result in different PCIe enumeration changes. the **vCON1 vNICs** are enumerated before **vCON2 vNICs**.