Cisco VM-FEX Basic Troubleshooting Guide in VMware ESX Environment

July 2012

Introduction

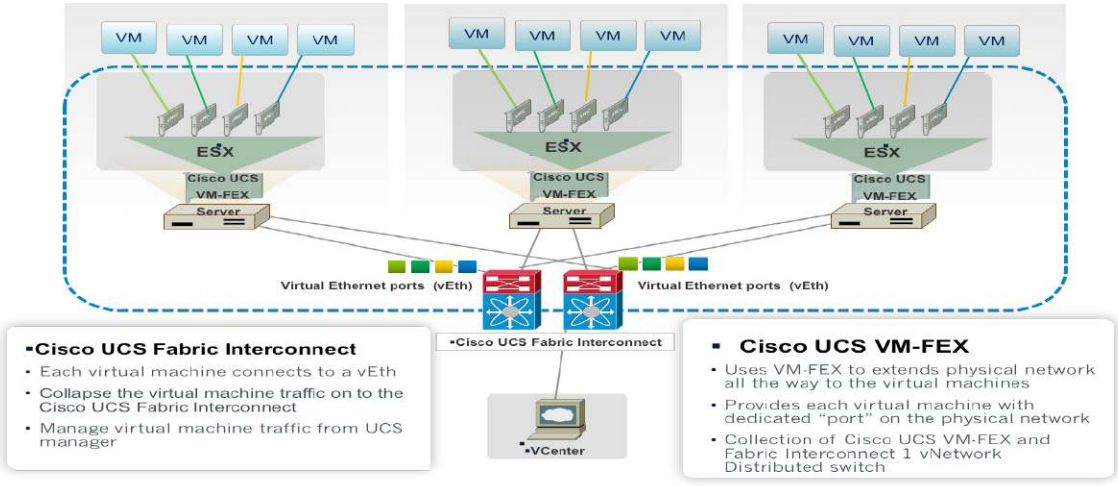
Cisco Virtual Machine Fabric Extender (VM-FEX) is a Cisco technology that addresses management and performance concerns in a data center by unifying physical and virtual switch management. The Cisco VM-FEX collapses virtual and physical networking into a single infrastructure. This unified infrastructure enables data center administrators to provision, configure, manage, monitor, and diagnose virtual machine network traffic and bare metal network traffic.

The Cisco VM-FEX significantly reduces the number of network management points, enabling both physical and virtual network traffic to be treated in a consistent policy driven manner.

The VM-FEX software extends Cisco fabric extender technology to the virtual machine with the following capabilities:

- Each virtual machine includes a dedicated interface on the parent switch
- All virtual machine traffic is sent directly to the dedicated interface on the switch
- The software-based switch in the hypervisor is eliminated

Figure 1. Extension of Fabric Extender technology with Fabric interconnects using VM-FEX



The Figure 1 shows the extension of Fabric Extender technology with Fabric interconnects using VM-FEX. The Cisco Virtual Machine Fabric Extender (VM-FEX) technology extends Cisco Fabric Extender technology all the way to the virtual machine. Each virtual machine gets a dedicated interface on the parent switch (virtual Ethernet port). All virtual machine traffic is sent directly to the dedicated interface on the switch. VM-FEX eliminates the software based switch within the hypervisor by providing individual virtual machine virtual ports on the physical network switch. Virtual machine traffic is sent directly to the upstream physical network switch, which takes full responsibility for virtual machine switching and policy enforcement. This approach leads to consistent treatment for all network traffic, virtual or physical. VM-FEX consolidates virtual and physical switching layers into a single layer and reduces the number of network management points by an order of magnitude. The following are the benefits of VM-FEX:

- Simplicity

- One infrastructure for virtual and physical resource provisioning, management, monitoring and troubleshooting
- Consistent features, performance and management for virtual and physical infrastructure

- Robustness
  - Programmability, ability to re-number VLANs without disruptive changes
  - Trouble shooting & Traffic engineering VM traffic from the physical network
- Performance
  - VMDirectPath with vMotion provides near bare metal I/O performance
  - Line rate traffic to the virtual machine

UCS VM-FEX Basic Troubleshooting

The document covers the basic troubleshooting methods to be followed in the VM-FEX DVS environment.

Troubleshooting

ESX Host Side

This section explains the method to check the VM-FEX switch status on ESX Host:

Step 1. Check the VEM switch status on the ESX host using the following command.

```
~ # vem status -v
Package vssnet-esxmn-ga-release
Version 4.2.1.1.4.1.0-3.0.4
Build 4
Date Thu Aug 25 10:47:10 PDT 2011
Number of PassThru NICs are 15
VEM modules are loaded
DVS Name        Num Ports   Used Ports  Configured Ports  MTU   Uplinks
DVS-PTS-VNFEX   256         17          256               1500  vmnic0,vmnic1
Number of PassThru NICs are 15
```

Step 2. Check if the VEM Module is loaded by the ESX Kernel using the following command.

```
# vmkload_mod -l | grep vem
vem-v132-svs-mux   12   32
vem-v132-pts       1    144
```

Step 3. You can view information on VM-FEX maximum number of ports, connectivity status, used ports IDs which are mapped toVirtual Machines Dynamic vNIC ,vmkernel etc., using the following command.

```
esxcfg-vswitch –l
Switch Name       Num Ports   Used Ports  Configured Ports  MTU     Uplinks
vSwitch0           128         2           128               1500
PortGroup Name    VLAN ID     Used Ports  Uplinks
Service Console    0           0
VM Network         0           0           DVS Name          Num Ports   Used
Ports  Configured Ports  MTU     Uplinks
UPT_DVS            256         5           256        1500    vmnic0,vmnic1
DVDV Port ID 1500         In Use      Client
1516               0
```

```
1523                1           vmnic1
1539                1           vmnic0
1                   1           vmk0
1500                1           Windows2K8R2 ethernet0
```

Step 4. Check the Virtual Machine's Dynamic vNIC network connectivity issues from the ESX host side using the following command. The commands and figures help you to narrow down the network connectivity issues.

The following command allows you to get the login details from the vmkernel log file. Figure 2 shows the UCS Graphic User Interface (GUI) and Figure 3 shows the vCenter GUI. These figures show the DV Port ID 1500 to which the Virtual Machine is connected.

```
~ # cat /var/log/vmkernel.log  | grep 1500 | grep hardPts
2011-10-03T09:57:03.314Z cpu12:3875)hardPts: hard_niv_dev_authenticate:527: Auth
STARTED on devName vf_vmnic0 08:00.2 MAC: 00:0c:29:b7:8c:95 port_profile:
dvportgroup-44 name: Windows2K8R2 ethernet0 portID: 0x2000011 dvportname: 1500
dvsname: 59 63 21 5$
2011-10-03T09:57:03.314Z cpu12:3875)hardPts: hard_niv_dump_provinfo:377: String:
1500
```

Figure 2. UCS Manager showing the Virtual Machine and the DV Port ID 1500 connectivity

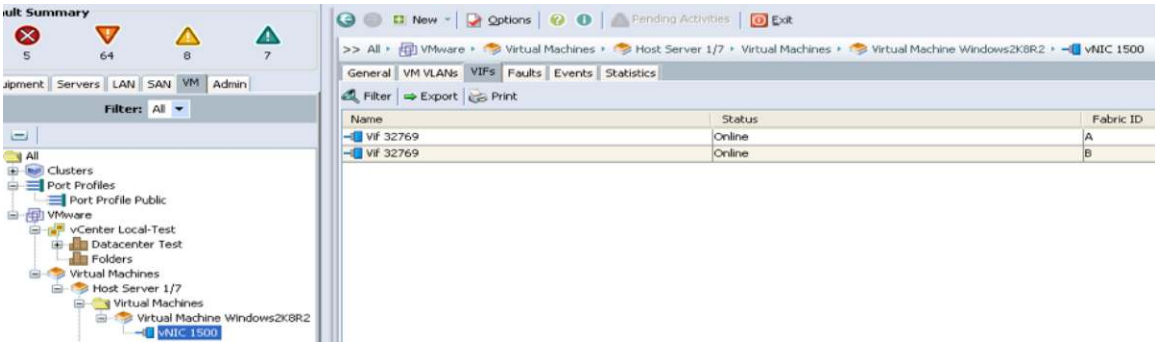
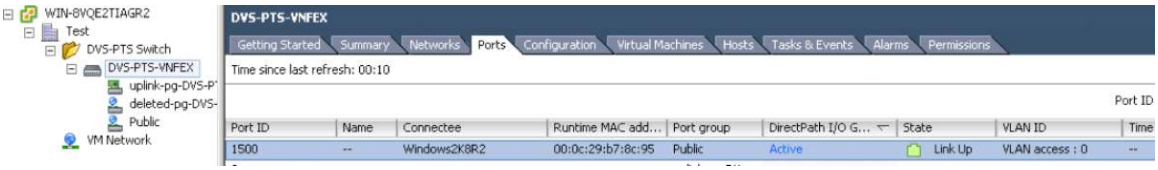


Figure 3. vCenter showing the DV Port ID 1500 connectivity



The following command allows you to get the login details from the vmkernel log file. Figure 4 shows the UCS Graphic User Interface (GUI) and Figure 5 shows the vCenter GUI. These figures show the DV Port ID 1500 to which the Virtual Machine is connected.

```
~ # cat /var/log/vmkernel.log  | grep 1500
2011-10-03T10:04:01.239Z cpu6:2093)Net: 2191: dissociate dvPort 1500 from port
0x2000011
```

Figure 4. The MAC address status after dissociation with the VM-FEX DV Port ID 1500

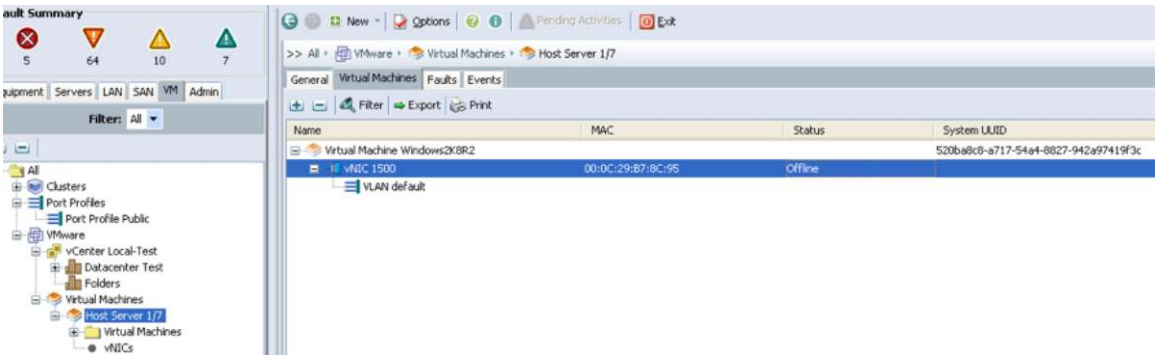
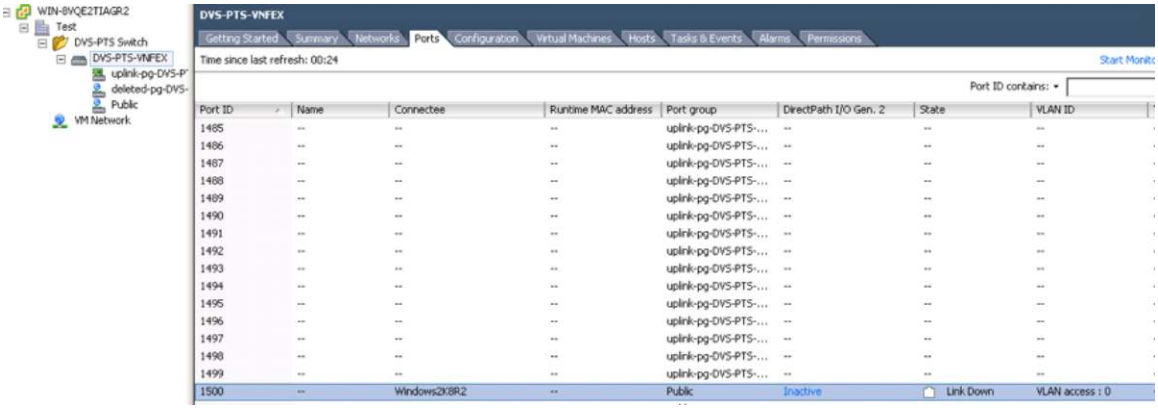


Figure 5. DirectPath I/O status after the dissociation of MAC address with the VM-FEX DV Port ID 1500



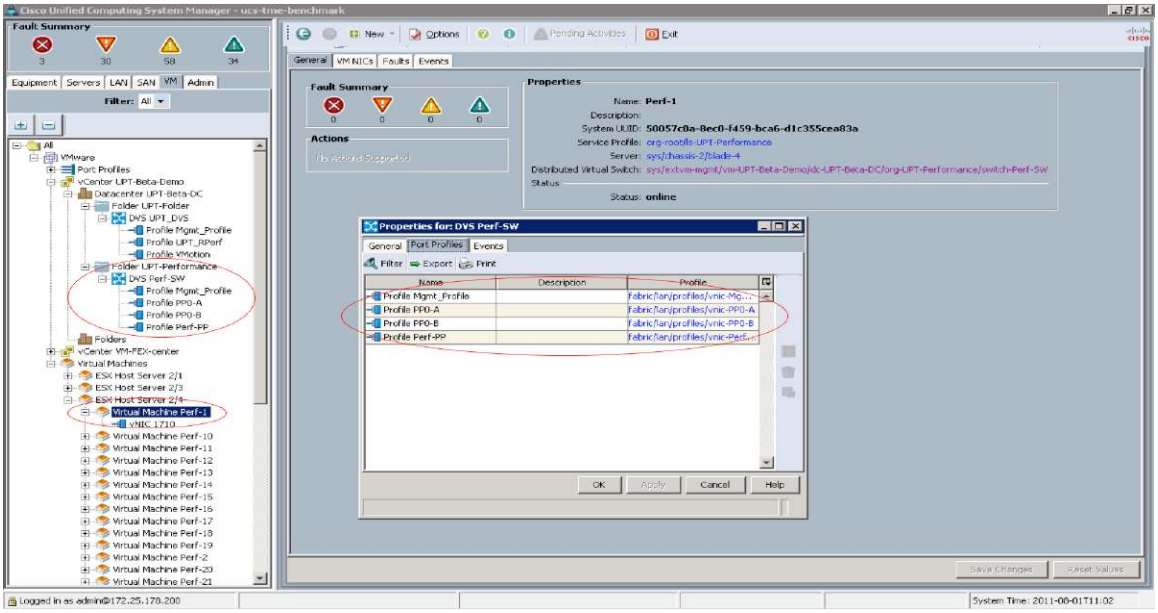
The following command allows you to get the login details from the vmkernel log file.

```
~ # cat /var/log/vmkernel.log  | grep 1500
2011-10-03T10:12:52.931Z cpu4:4687)hardPts: hard_niv_dev_authenticate:527: Auth
STARTED on devName vf_vmnic0 08:00.2 MAC: 00:0c:29:b7:8c:95 port_profile:
dvportgroup-44 name: Windows2K8R2 ethernet0 portID: 0x2000012 dvportname: 1500
dvsname: 59 63 21 50$
2011-10-03T10:12:52.931Z cpu4:4687)hardPts: hard_niv_dump_provinfo:377: String:
1500                                                              2011-10-
03T10:12:52.939Z cpu4:4687)NetPort: 2600: resuming traffic on DV port 1500
```

UCS VM-FEX / vCenter View

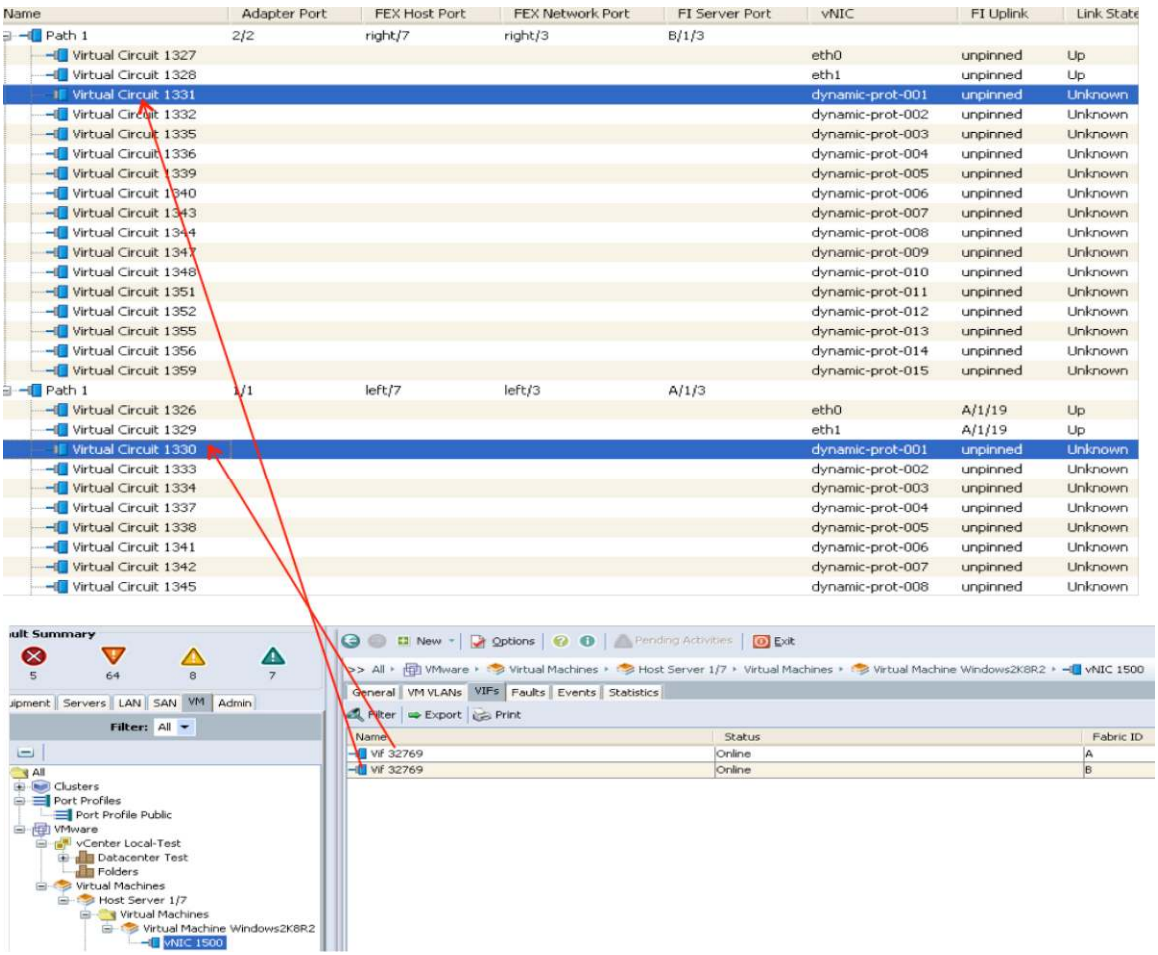
The UCS VM tab allows you to view all available VM-FEX switches defined in the Fabric Interconnect and also allows you to define port profiles with network settings. You can apply the port profiles to multiple VM-FEX switches running on the Fabric Interconnect. The UCS VM tab provides a logical view of all virtual machines' dynamic vNICs with the corresponding VM-FEX DV Port ID 1500 connectivity information.

Figure 6. DVS perf-SW window



To support automatic Fabric based failover on Virtual Machine Dynamic vNICs, the VM-FEX creates active and standby Virtual Interfaces (VIF). These interfaces are placed on Fabric Interconnect A and B. The corresponding Dynamic port names of the virtual interfaces are displayed on the UCS service profile. Figure 7 shows the mapping information.

Figure 7. Mapping information of VIF with Dynamic vNICs



The UCS CLI view provides information on Virtual Interface of Virtual Machine Dynamic vNIC mapping to the UCS vEthernet Interface. You can verify the mapping information using the following commands.

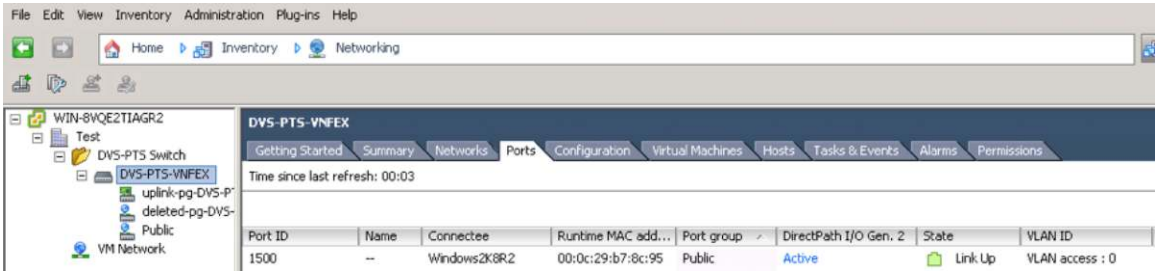
```
Fabric-Cus1-A /system/vm-mgmt/vmware/virtual-machine # show vnic
vNIC:
    Name        Status  MAC Address
    ----------- ------- -----------
    1501        Online  00:50:56:A1:6A:56

Fabric-Cus1-A /system/vm-mgmt # show virtual-machine Virtual Machine:
 UUID: 520ba8c8-a717-54a4-8827-942a97419f3c
 Name: Windows2K8R2
Service Profile: org-root/ls-ESX-Blade7-ESX
       Server: sys/chassis-1/blade-7
Status: Online
Fabric-Cus1-A /system/vm-mgmt # scope virtual-machine 520ba8c8-a717-54a4-8827-
942a97419f3c
```

```
Fabric-Cus1-A /system/vm-mgmt/vmware/virtual-machine # show detail expand
Virtual Machine:
    UUID: 520ba8c8-a717-54a4-8827-942a97419f3c
    Name: Windows2K8R2
    Service Profile: org-root/ls-ESX-Blade7-ESX
    Server: sys/chassis-1/blade-7
    hv-type: Esx
    Status: Online
    Description:
 vNIC:
    Name: 1501
    Status: Online
    MAC Address: 00:50:56:A1:6A:56
    Profile Name: Public
    Virtual Adapter:
    Vnic Dn:
    Current Task:
 VIF:
    Vif Id: 32771
    Adapter Virtual If ID: 1330
    Status: Online
    Phys Fabric ID: A
    linkState: Up
    operState: Active
    Phys Border Card ID: 1
    Phys Border DV Port ID 1500: 19
    Vif Id: 32771
    Adapter Virtual If ID: 1331
    Status: Online
    Phys Fabric ID: B
    linkState: Up
    operState: Active
    Phys Border Card ID: 1
    Phys Border DV Port ID 1500: 19
Fabric-Cus1-A(nxos)# sh mac address-table vlan 1 | grep 0050.56a1.6a56
* 1        0050.56a1.6a56    static    0          F    F  Veth32771
```

The VMWare vCenter Networking tab displays all the available VM-FEX DVS switches from different Data Centers. The tab also provides the corresponding Virtual Machine's Dynamic vNIC DV Port ID 1500 which is a part of VM-FEX DVS switches across Data Center.

Figure 8. Availability of VM-FEX DVS switches



ESX Host / Cisco VIC / UCS VM-FEX Mapping

This section describes the method to check the mapping from ESX host to Cisco VIC and displays the entire mapping of UCS VM-FEX:

ESX Host Vmkernel log

The Vmkernel in ESX Host logs the control work flow of the Virtual Machine Dynamic vNIC's association and disassociation during boot and shutdown of the Virtual Machine. ESX Host is a part of the VM-FEX DVS switch, running within the Fabric Interconnect.

The following example shows the method to check the DV Port ID 1500 connected to the Dynamic vNIC, the VIF IDs with their paths and the DV Port ID 1500s on both the Fabric Interconnects A and B associated with the paths.

~ # cat /var/log/vmkernel.log | grep 1500 | grep hardPts

2011-10-03T09:57:03.314Z cpu12:3875)hardPts: hard_niv_dev_authenticate:527: Auth STARTED on **devName vf_vmnic0 08:00.2 MAC: 00:0c:29:b7:8c:95** port_profile: dvportgroup-44 name: Windows2K8R2 ethernet0 portID:0x2000011 dvportname: 1500 dvsname:59 63 21 5$

2011-10-03T09:57:03.314Z cpu12:3875)hardPts: hard_niv_dump_provinfo:377: String:1500

In the above command, hard_niv_dev_authernticate is the control signal which does the authentication process on Dynamic vNIC vf_vmnic0 MAC Address **(08:00.2)**. This address is assigned to the Virtual Machine MAC address **(00:0c:29:b7:8c:95)**.

Cisco VIC Adapter

To see the corresponding Dynamic vNIC vf_vmnic0 MAC Address **08:00:2** associated with the vNIC on the adapter, you need to connect to the corresponding adapter by using the following command. In this example blade 1 is connected on chassis 2 which has the Cisco VIC adapter 1.

UCS-A# connect adapter 2/1/1

adapter 2/1/1 # connect

adapter 2/1/1 (top):1# attach-mcp

adapter 2/1/1 (mcp):1# vnic
vif uif : bound uplink 0 or 1, =:primary, -:secondary, >:current

---

vnic lif vifid name type bb:dd.fstate lifstate uif ucsm idx vlanstate

```
-------------------------
 5vnic_1 enet   08:00.0 UP 2 UP  =>0  1326   187    1 UP
                           - 1  1327   181    1 UP
 6vnic_2 enet   08:00.1 UP 3 UP  - 0  1329   186    1 UP
                           =>1  1328   180    1 UP
 7vnic_3 enet_pt08:00.2 UP 4 UP  =>0  1330   189    1 UP
                           - 1  1331   183    1 UP
adapter 1/7/1 (mcp):14# vnic 7
vnicid : 7
name : vnic7
type : enet_pt
state : UP
adminst : UP
slot : 0
bdf : 03:00.2
mac : 00:00:00:00:00:00
vifid : 1330
vifcookie : 1297761
uif : 0
stby_vifid : 1331
profile :
stdby_profile :
```

In the above command output, devName vf_vmnic 0 08:00.2 MAC: 00:0c:29:b7:8c:95 has two VIF DV Port ID 1500s, 1330 and 1331 created on both the Fabric Interconnects. 1330 VIF ID is active and created on Fabric Interconnect A (UP =>0 1330). 1331 VIF ID is standby (stby_vifid :1331) and created on secondary Fabric Interconnect B (-1 1331).

In the output shown below, Guest VM Dynamic vNIC MAC Address 00:0c:29:b7:8c:95 (VIF ID 1330) is registered on the active Fabric Interconnect A and VIF ID 1331 is in standby and becomes active when there is a fabric failover event.

```
Fabric-Cus1-A(nxos)# sh mac address-table address   000c.29b7.8c95
  VLAN     MAC Address      Type      age     Secure NTFY    Ports
---------+-----------------+--------+---------+------+----+-----------------
* 1       000c.29b7.8c95    static    0         F    F  Veth32769
```

Error Messages

This section provides error messages that are seen when enabling VMDirectPathmode on the Dynamic vNICs. The Table 1 gives the section wise information on these errors.

Table 1. Error Messages

Error Messages	Sections
“The NIC does not support DirectPath I/O Gen. 2.”	Refer to section Guest Virtual Machine Network Adapter Type on page 13
“The hardware of the virtual machine’s host does not support DirectPath I/O Gen. 2.” “The NIC is disabled or disconnected which prevents it participating in DirectPath I/O Gen. 2.”	Refer to section ESX Host BIOS Setting for enabling VMDirectPath on page 14
“To activate DirectPath I/O Gen. 2, go to the Resources tab and select Memory Settings to reserve all guest memory”	Refer to section ESX Guest VM Memory Reservation Setting on page 16
“The vSphere distributed switch to which the NIC’s port belongs does not support DirectPath I/O Gen. 2.” “DirectPath I/O is unavailable due to insufficient host interrupt vectors.”	Refer to section Insufficient Resources in Dynamic Adapter Configuration on page 17
“The NIC has features enabled (For example INT-x or PXE boot) that prevent it from partcipating in DirectPath IO “	Refer to section ESX Guest VM PXE Operation on page 19

VM-FEX with VMDirectPath Basic Prerequisites

You need to follow the checklist before enabling High Performance on the Port profile. Enabling High Performance turns on VMDirectPath (UPT) mode on the Dynamic vNICs which are part of Virtual Machines.

If you cannot set the VMDirectPath to active state and get error as explicitly disabled NIC’s ports as shown in Figure 9, then there can be several causes as described below:

Figure 9. Properties windows to check the DirectPath I/O status

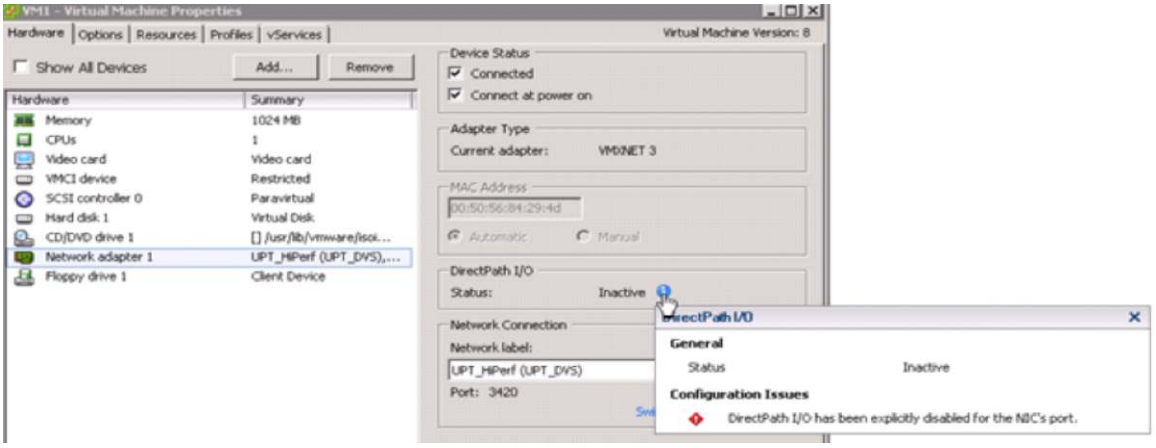
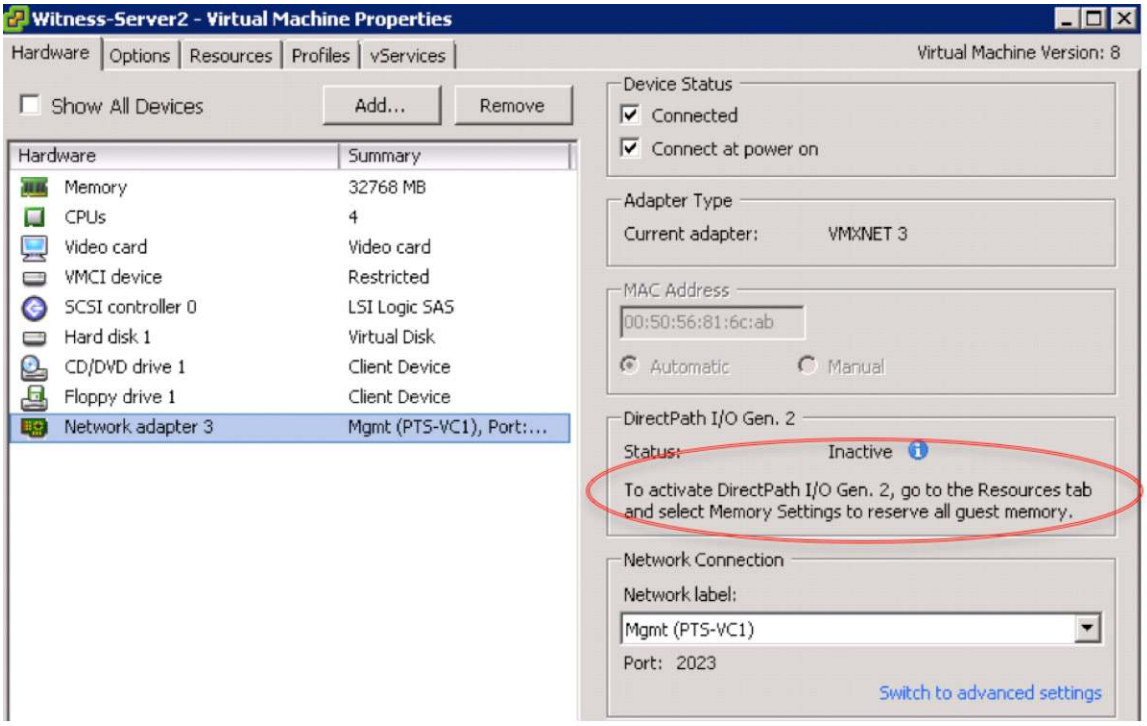


Figure 10. Virtual Machine properties window to activate DirectPath I/O



Guest Virtual Machine OS and VMXNET3 supported Versions

The following table shows the VMXNET3 driver versions to support VMDirectPath mode for different Guest Virtual Machine Operating Systems.

Table 2. VMXNET3 driver versions to support VMDirectPath mode

SL No	Guest OS	VMXNET3 Driver version
1	Windows 2008 SP2	1.2.22.0
2	Windows 2008 R2	1.2.22.0
3	RHEL 6.0	1.0.14.0-k
4	SLES11 SP1	1.0.14.0-k
5	SLES11	1.0.36.0

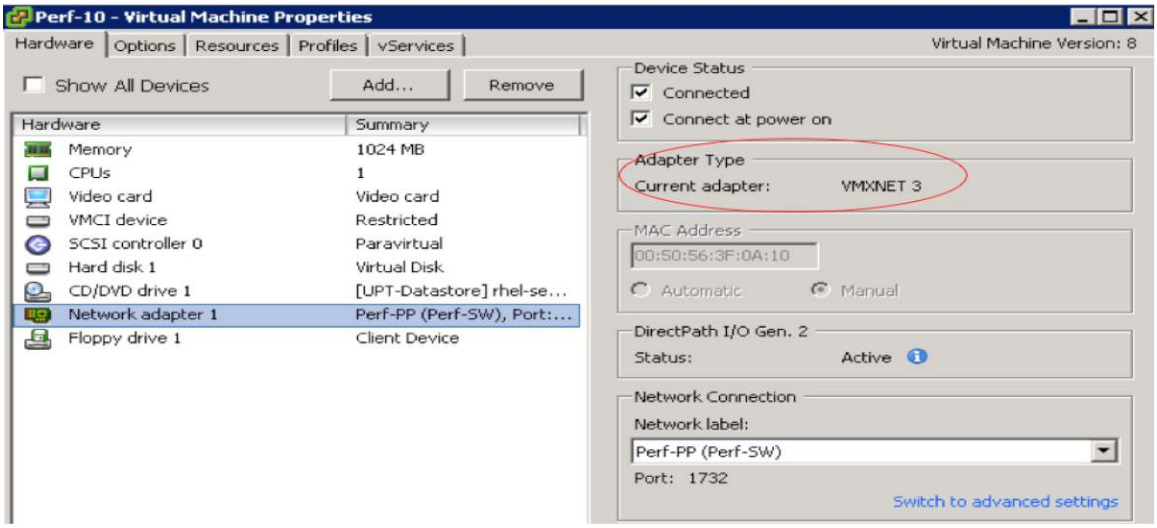
Guest Virtual Machine Network Adapter Type

VM-FEX VMDirectPath mode is enabled and operational only if the Guest Virtual Machine is loaded with VMXNET 3 emulated Network driver. By default the UPT mode will be turned off.

Ensure that when you are adding a Network Adapter to a Virtual Machine in vCenter, the adapter type of the selected adapter is VMXNET 3. Also, you need to ensure that the VMXNET 3 driver is loaded on the Guest Virtual machine by using the following command:

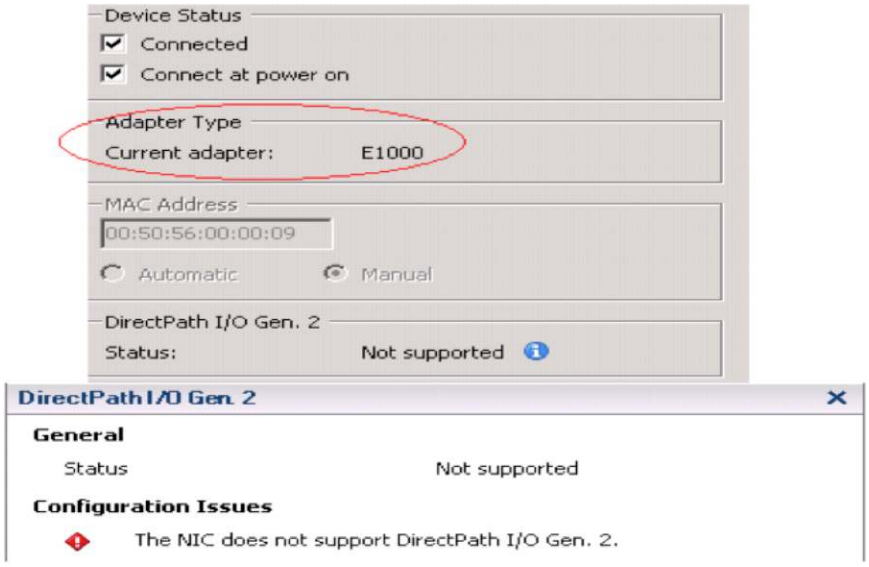
```
lsmod | grep vmxnet3
```

Figure 11. Virtual Machine properties window showing the current adapter in use



When the adapter attached to the Virtual Machine is not VMXNET3, Dynamic vNIC with VMDirectPathMode cannot be enabled. Click the icon to view the error message as shown in the Figure 12.

Figure 12. VM properties window showing the error message

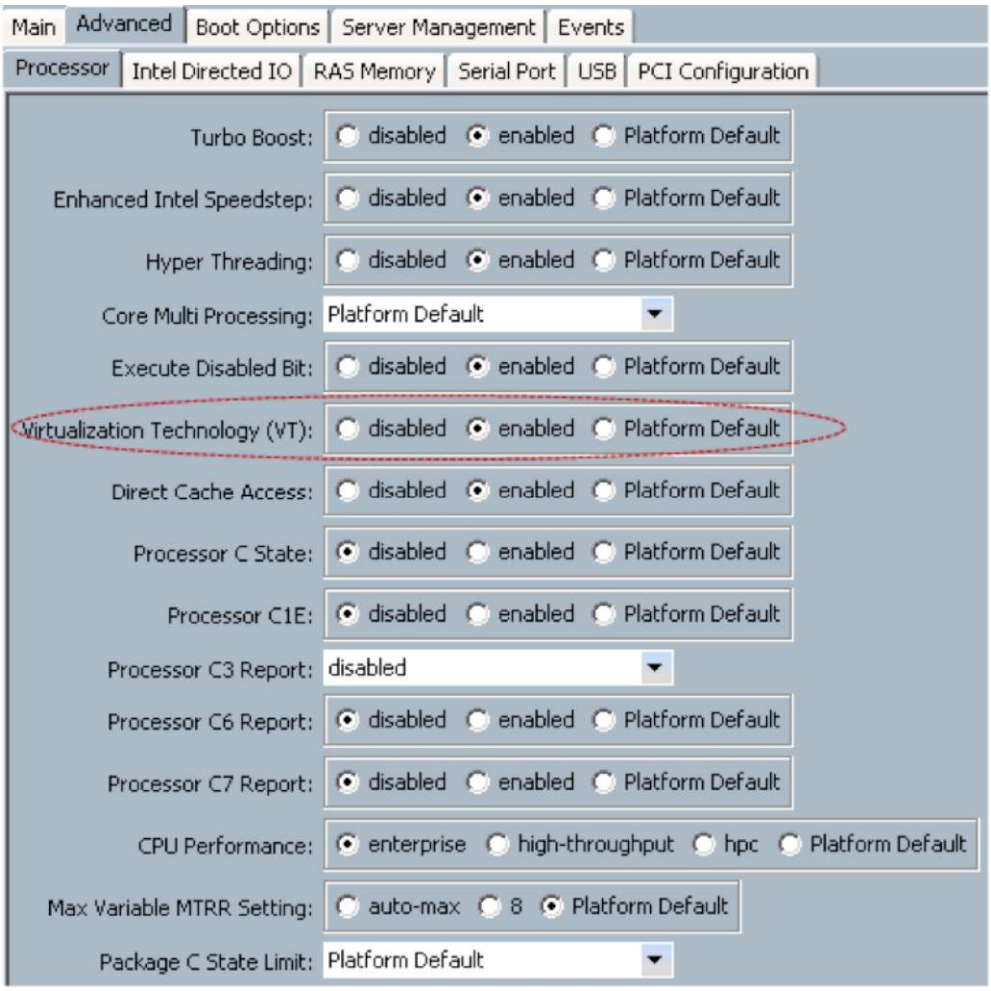


ESX Host BIOS Setting for enabling VMDirectPath

To set VMDirectPath mode in VM-FEX, you need to enable Intel VT-D in the BIOS. When Intel VT-D is enabled, the status of Direct IO on Virtual machine properties page changes to active.

You can enable Virtualization Technology (VT) in the BIOS policy as shown in Figure 13.

Figure 13. Virtualization Technology settings in BIOS Policy



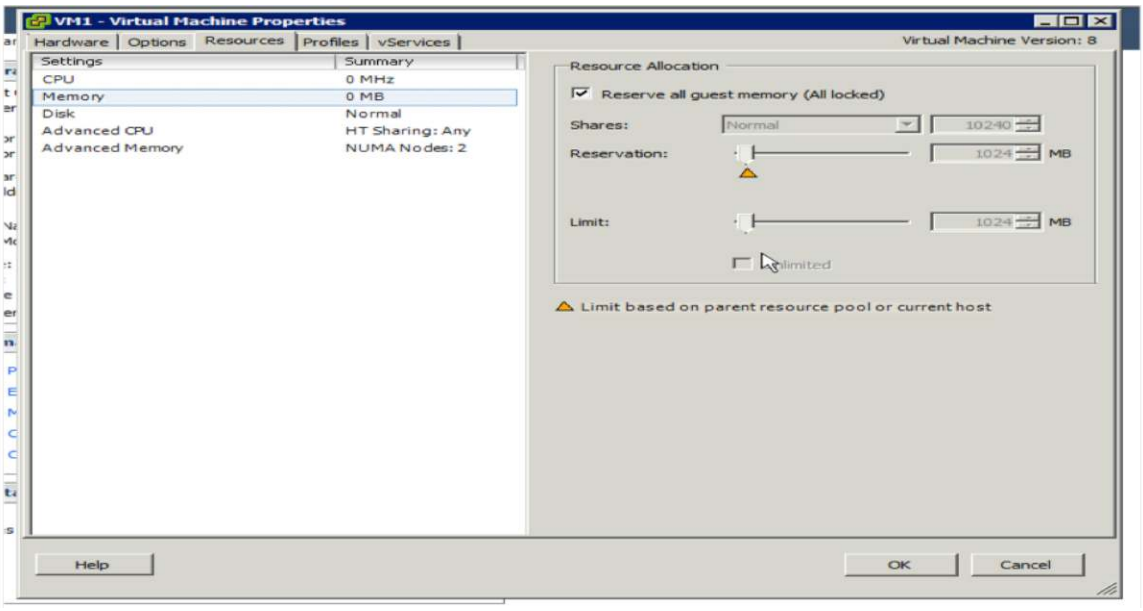
When the Intel VT-D feature is disabled in the ESX host Server, the VMDirectPathMode cannot be enabled. Click the icon to view the error message as shown in the Figure 14.

Figure 14. VM properties window showing the error message

ESX Guest VM Memory Reservation Setting

To enable VM-FEX VMDirectPath mode on Guest VMs, you need to reserve memory in advance in the VM configuration setting. The Figure 15 shows the guest memory reservation details.

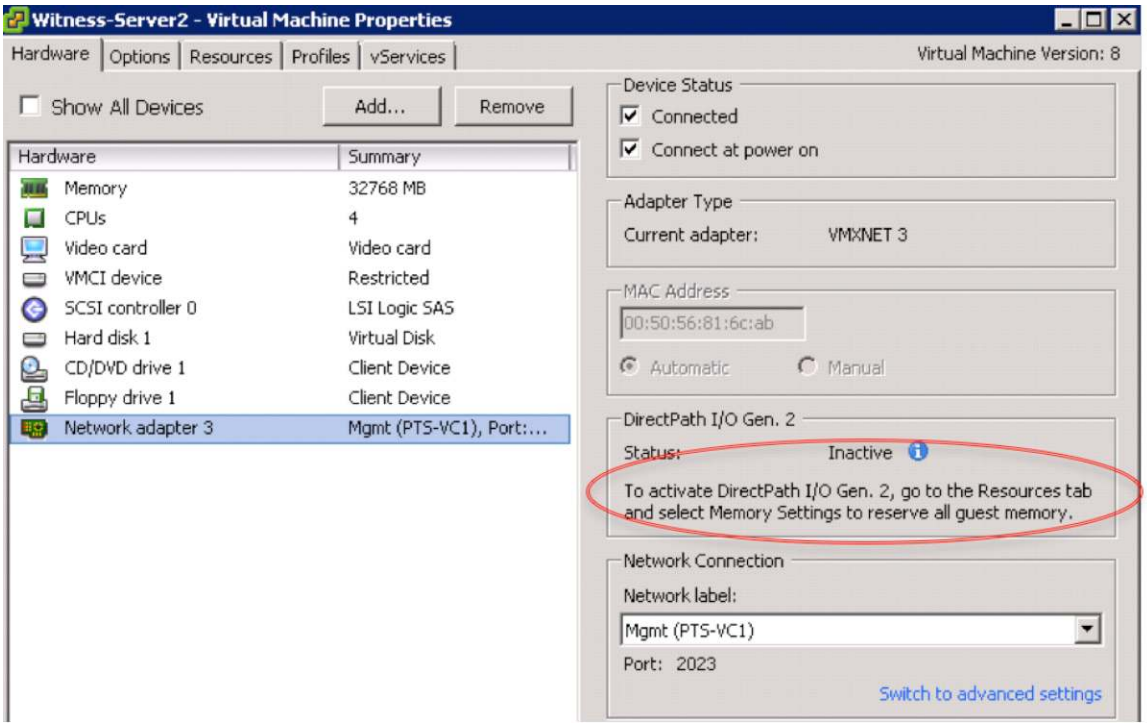
Figure 15. Virtual Machine properties window showing resource allocation



When there is no memory reserved for the virtual machine, the VMDirectPathMode cannot be enabled.

Click the icon to view the error message as shown in the Figure 16.

Figure 16. VM properties window showing the error message



Insufficient Resources in Dynamic Adapter Configuration

- Interrupts

  VMDirectPath mode by default Interrupt mode is set to MSI-X which consumes 4 Interrupts vectors from ESX Host Interrupt pool of 128 maximum limit, if there is a scenario where ESX host reports Interrupt pool is full and cannot allocate any interrupt vectors to any Virtual Machines' to change to VMDirectPath mode from emulated mode then will fail and reports in active status in vCenter on Virtual Machine network properties pane. The Figure 17 shows the default Interrupt values.

- Transmit/Receive/Completion Queues

  To enable VMDirectPath mode with RSS/multi-queue, you need to ensure that the adapter policy is configured accordingly. The values can be set in Eth Adapter Policy VMWare PassThru and in VMXNET3 driver. The values for these queues must meet the following criteria:

  - The values set for Transmit Queues (TQs) in VMXNET3 should be greater than the values set for TQs in Eth Adapter Policy VMWare PassThru (largest VMXNET3.NumTQs on host <= dynamicPolicy.NumTQs).
  - The values set for Recieve Queues (RQs) in VMXNET3 should be greater than the values set for RQs in Eth Adapter Policy VMWare PassThru (largest VMXNET3.NumRQs on host <= dynamicPolicy.NumRQs).
  - The values set for Transmit Queues (CQs) in VMXNET3 should be greater than the values set for CQs in Eth Adapter Policy VMWare PassThru (largest VMXNET3.NumCQs on host <= dynamicPolicy.NumCQs).

- The values set for Interrupts (INTRs) in VMXNET3 should be greater than the values set for INTRs in Eth Adapter Policy VMWare PassThru (largest VMXNET3.NumINTRs on host <= dynamicPolicy.NumINTRs)

Note: If any one of these criterias are not met, then the VMDirectPath will remain in inactive mode.

Figure 17. Ring Size Values for Transmit and Receive Queues

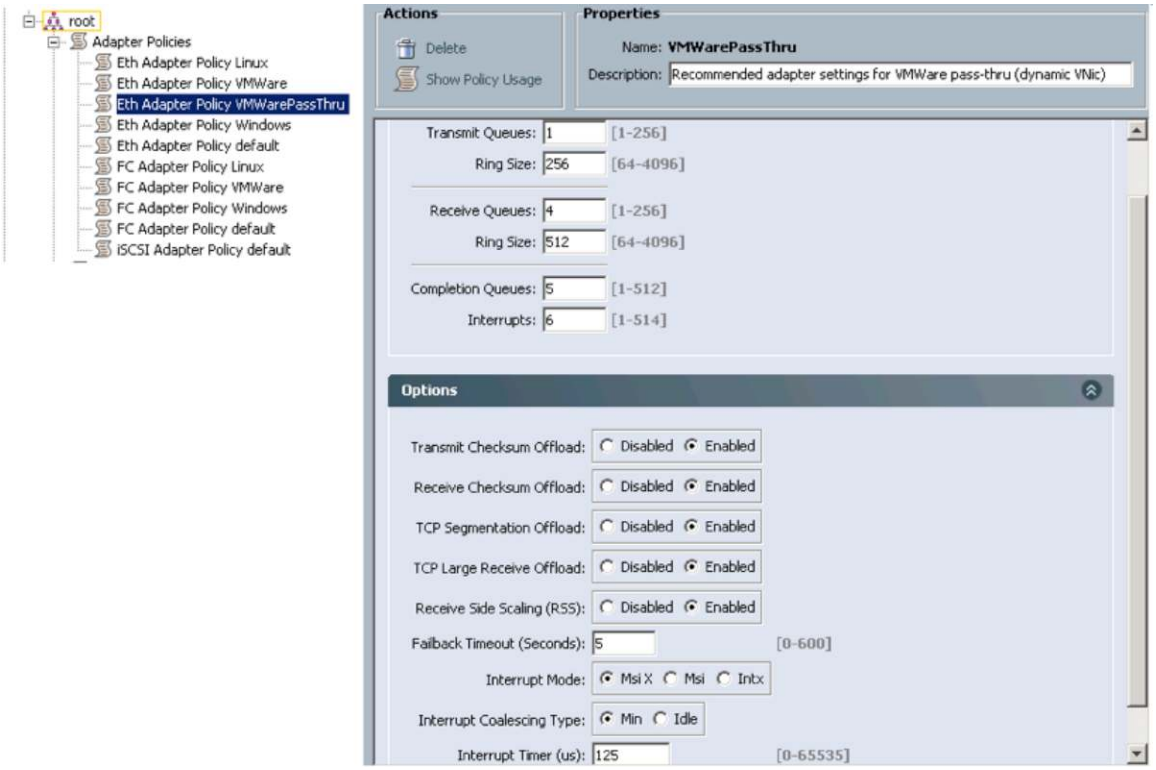
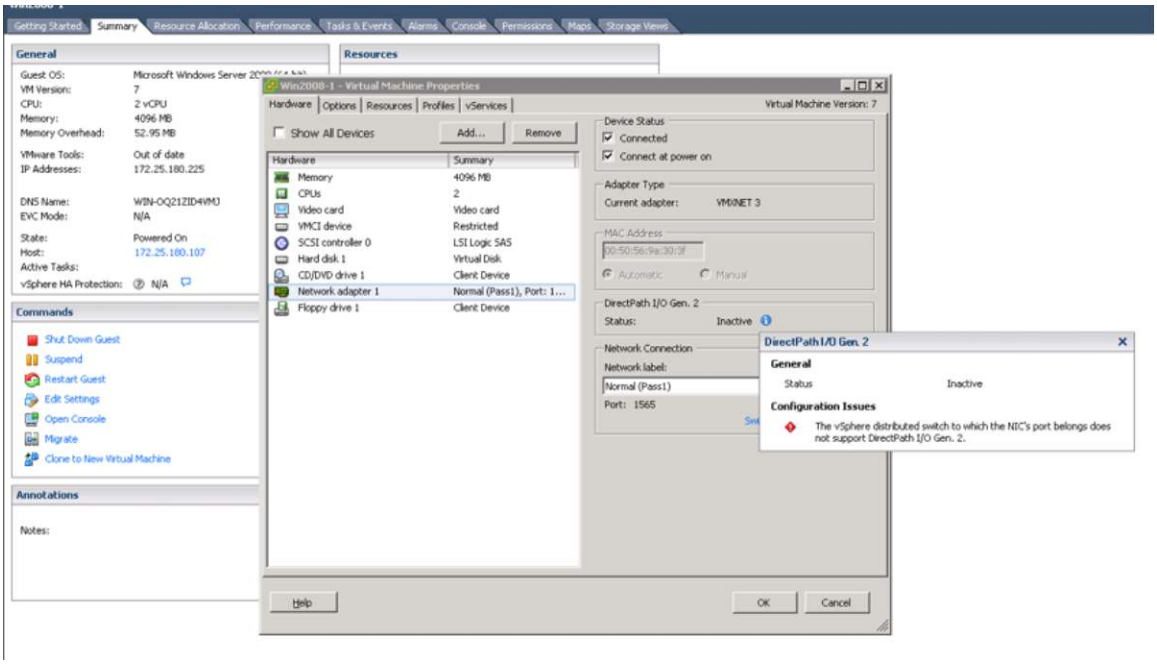


Figure 18. VM properties window showing the error message



When the Dynamic vNIC does not have sufficient resources like Write Queues (WQs), Receive Queues (RQs), Completion Queues (CQs), and intrrupts (intrs) to support VMDirectPathMode, you will get an error message as shown in the Figure 18.

Click the icon to view the error message as shown in the Figure 18.

ESX Guest VM PXE Operation

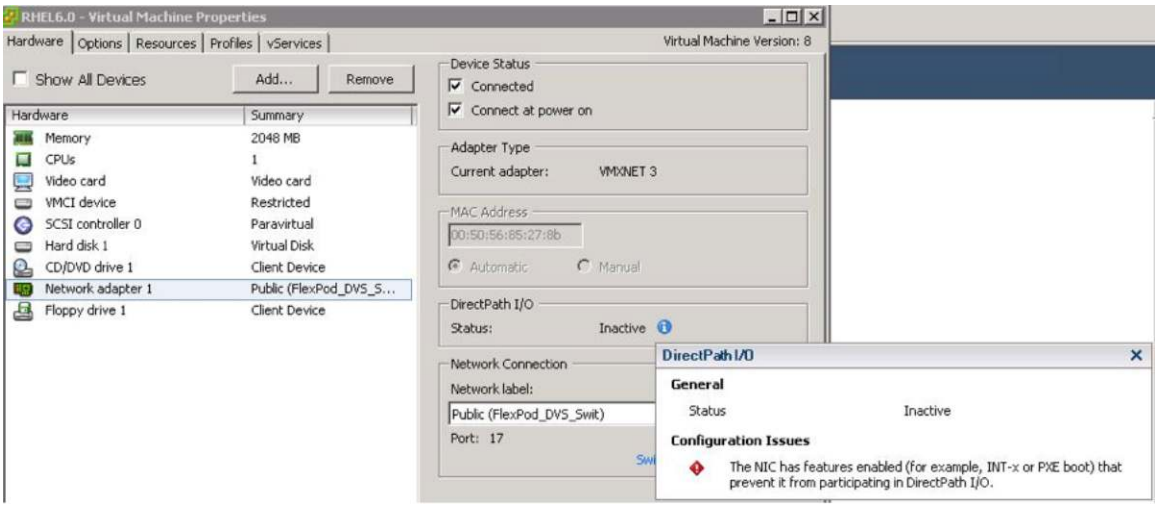
VMDirectPath IO will be in “Active” on Guest VMs only after Virtual Machince PXE install and boot operatrions are completed with correct VMXNET3 Driver is loaded.

For Example: RHEL 6.0 with VMXNET3 driver version 1.0.14.0-k, default Virtual Machince setting on vSphere 5 may encounter this behavior during PXE installtion or boot operation.

For more detail refer Guest Operating System Installtion Guide http://partnerweb.vmware.com/GOSIG/home.html

The Figure 19 shows VMDirectPath IO configuration error message during the Guest VM PXE installation or the boot phase. Click the icon to view the error message as shown in the Figure 19.

Figure 19. VM properties window showing the error message



References

Hardware and Software Interoperability Matrix
http://www.cisco.com/en/US/docs/unified_computing/ucs/interoperability/matrix/r_hcl_B_rel2_0.pdf

Cisco UCS Manager VM-FEX for VMware GUI Configuration
http://www.cisco.com/en/US/docs/unified_computing/ucs/sw/vm_fex/vmware/gui/config_guide/b_GUI_VMware_VM-FEX_UCSM_Configuration_Guide.pdf

vSphere Networking ESXi 5.0
http://pubs.vmware.com/vsphere-50/topic/com.vmware.ICbase/PDF/vsphere-esxi-vcenter-server-50-networking-guide.pdf

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C07-712224-00 07/12