How to Configure Intel® Ethernet Converged Network Adapter-Enabled Virtual Functions on VMware® ESXi® 5.1

Technical Brief v1.0

February 2013
## Revision History

<table>
<thead>
<tr>
<th>Date</th>
<th>Revision</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>February 2013</td>
<td>1.0</td>
<td>Initial release</td>
</tr>
</tbody>
</table>
1.0 Introduction

As data centers deploy more virtualized servers to consolidate infrastructure and reduce TCO, providing scalable networking performance within a virtualized environment becomes a significant challenge for IT administrators. The Intel® Ethernet Converged Network Adapter X520 and Intel® Ethernet Converged Network Adapter X540 families support industry-leading features that are accelerating the performance and implementation of 10 Gigabit Ethernet in a virtualized environment. As part of Intel® Virtualization Technology for Connectivity (Intel® VT-c), Intel® Ethernet Converged Network and Server Adapters support PCI-SIG® Single Root I/O Virtualization and Sharing (SR-IOV), which offers the ability to create multiple lightweight Virtual Functions (VFs) per physical adapter port. These VFs are lightweight PCIe functions that contain all the resources necessary for data movement but have a carefully minimized set of configuration resources available to host operating systems. They are capable of providing 10 Gb Ethernet connectivity to Virtual Machines (VMs) within a virtualized operating system framework, such as provided by Microsoft® Windows® Server 2012 with Hyper-V®, Linux® KVM and XEN®, and VMware® ESXi® 5.1.

In VMware® ESXi 5.1, SR-IOV enables administrators to scale the lower latency and higher throughput performance benefits of VMware® DirectPath I/O* pass-through technology to multiple guest VMs without increasing physical port count. To learn more about VMware's implementation of SR-IOV, including restrictions on Guest OS and feature support, please consult VMware's documentation at the following address:


This document explains configuration and use SR-IOV VFs on Intel® Ethernet CNA X540 and X520 in a VMware ESXi® 5.1 environment.

1.1 Hardware Requirements

A server platform that meets the following requirements:

- Compatibility with the ESXi release
- An Intel processor (with no AMD processors present)
- Support for an input/output memory management unit (IOMMU) with IOMMU enabled in the BIOS.
- Support for Intel® Virtualization Technology for Directed I/O (VT-d) and the PCI-SIG® Single Root I/O Virtualization and Sharing (SR-IOV) specification
- Support for SR-IOV enabled in the BIOS. Contact the server vendor to determine whether the host supports SR-IOV.
- An available PCI Express* v2.0 (5.0GT/s) x8 Lanes with ARI and ACS support slot.
- An Intel® Ethernet Converged Network Adapter X520/X540

Note: To verify compatibility of physical hosts and NICs with ESXi releases, see the VMware Compatibility Guide.

1.2 Software Requirements

- VMware ESXi 5.1 and VMware vSphere* 5.1 Client and/or VMware vCenter 5.1 Server
1.2.1 Upgrading from earlier versions of vSphere

- If upgrading from vSphere 5.0 or earlier to vSphere 5.1 or later, SR-IOV support is not available unless NIC drivers have been updated for the vSphere 5.1 release. NICs must enable firmware and drivers supporting SR-IOV for SR-IOV functionality to operate properly.

1.2.2 Supported Guest configurations (from VMware documentation)

1.2.2.1 Guest OS

- Red Hat Enterprise Linux 6.x
- Windows Server 2008 R2 with SP2

1.2.2.2 VF driver in the guest OS

- Must be compatible with the NIC
- Must be supported on the guest OS release according to the technical documentation from the NIC vendor
- Must be Microsoft WLK or WHCK certified for Windows virtual machines.
- Must be installed on the OS

The OS release contains a general, default driver for NICs. For specific NICs, download and install the appropriate driver provided by the NIC’s manufacturer.

Note: When upgrading from vSphere 5.0 or earlier to vSphere 5.1 or later, SR-IOV support is not available until NIC drivers for the vSphere release are updated. NICs must have firmware and drivers supporting SR-IOV enabled to get SR-IOV functionality.

2.0 Installation and Configuration

2.1 Host Server Enablement and OS Installation

1. Install the Intel Ethernet CNA in an available PCI-Express x8 slot. (Ensure that the x8 slot is electrically connected as an x8. Some slots are physically x8 but electrically support x4 only. Verify this, and that the slot connects to a PCI Root Complex supporting ACS and ARI, with your server manufacturer or system documentation.)

2. Power up the server.

3. Enter the server’s BIOS setup and make sure the virtualization technology, Intel® VT-d, and SR-IOV features are enabled on the server.

4. Install VMware ESXi 5.1 on the server.

ESXi 5.1 host inbox drivers support SR-IOV, and are the recommended drivers for SR-IOV.

5. Configure vCenter server and/or vSphere client to connect to ESXi host for GUI management.

Note: Attempts to enable or configure unsupported features with SR-IOV in the vSphere Web Client will result in unexpected behavior in your environment.
2.2 VF Creation

There are two ways to create VFs on Intel Ethernet adapters in VMware ESXi 5.1:

- via command line console
- via host profile

Both of these methods are described below.

2.2.1 VF Creation via CLI Console Configuration

1. Enable ESXi Shell, either remote or direct console, for CLI configuration. Consult VMware documentation for more details on shell enablement.
2. Log into ESXi Shell.
3. Before you begin, make a note of the server adapter port numbering on PCI bus.

   ~ # lspci | grep -i intel | grep -i 'ethernet\|network'

4. Set ixgbe max_vfs module parameter to create VFs:

   ~ # esxcfg-module ixgbe -s max_vfs=W,X,Y,Z

   max_vfs is a comma-separated list, where each number in the list corresponds to the number of virtual functions enabled on a physical port. The numbers are ordered according to the order of the NIC ports on the PCI bus.

   Note: Use `lspci | grep -i intel | grep -i 'ethernet\|network'` to see port ordering on the bus. Example: max_vfs=0,10,0,10 corresponds to 10 VFs on adapter 1, port 2, and 10 VFs on adapter 2, port 2.

   The maximum number of VFs supported per ESXi host is a function of the maximum number of MSI-X vectors available for VF use. Intel Ethernet CNAs require three MSI-X vectors per VF. VMware ESXi 5.1 hosts have at most 128 MSI-X vectors available for VF use, so the theoretical maximum is 42 VFs per host when using Intel Ethernet CNAs. The actual limit depends on your server configuration.

5. Check ixgbe module parameters to verify setting:

   ~ # esxcfg-module -g ixgbe

   Example output:

   `ixgbe enabled = 1 options = 'max_vfs=0,10,0,10'`

6. Reboot the ESXi host for VF settings to take effect.

7. Check that VFs are created on the PCI bus through the shell commands and in the vSphere GUI under Host>Configuration>Advanced settings
Example:

```
~ # lspci | grep -i intel | grep -i 'ethernet\|network'
```

2.2.2 VF Creation via Host Profile Configuration

For completeness, instructions for configuring SR-IOV via Host Profiles in ESXi are included in this section. Most of this information is also included in the VMware vSphere 5.1 documentation. For more information about host profiles, see the vSphere Host Profiles documentation.
1. From vSphere Client, click View>Management>Host Profiles

![Screenshot of vSphere Client interface showing Host Profiles]

2. Create a profile or select one that has already been created and select Edit Host Profile.

![Screenshot of vSphere Client interface showing Edit Host Profile]

3. In the host profile edit screen, scroll down to **Kernel Module Configuration > Kernel Module > ixgbe**
5. Set `max_vfs` module parameter to create VFs. The value of `max_vfs` is a comma separated list, where each number in the list corresponds to the number of virtual functions enabled on a physical port. The numbers are ordered according to order of the NIC ports on the PCI bus.

From the console command line, use `lspci | grep -i intel | grep -i 'ethernet|network'` to see port ordering on the bus. Example: `max_vfs=0,10,0,10` corresponds to 10 VFs on adapter 1, port 2, and 10 VFs on adapter 2, port 2.

The maximum number of VFs supported per ESXi host is a function of the maximum number of MSI-X vectors available for VF use. The Intel® Ethernet Converged Network Adapter X520 and Intel® Ethernet Converged Network Adapter X540 families require 3 MSI-X vectors per VF. VMware ESXi 5.1 hosts have, at most, 128 MSI-X vectors available for VF use, so the theoretical maximum is 42 VFs per host when using the above adapters. The actual limit depends on your server configuration.

Set the host in Maintenance Mode before applying the Profile.
6. Attach the profile to the host by returning to the Inventory view, right-clicking on the host, and selecting **Host Profile > Manage Profile**. Select the SR-IOV enabled profile from the list and apply.
7. Set the host in Maintenance Mode before applying the Profile.
8. Apply the Profile by right-clicking on the host, and selecting **Host Profile > Apply Profile**.

9. You must reboot the ESXi host for the changes to take effect.

10. Check VFs are created in the vSphere GUI under **Host > Configuration > Advanced**. The numbers are ordered according to order of the NIC ports on the PCI bus.

### 2.3 VF Assignment to Guest VMs

#### 1. VF Naming

VFs are shown in the vSphere GUI under **Host > Configuration > Advanced Settings**

*Note:* After the virtual functions become enabled on the host, the physical NIC no longer shows up as a host network adapter in the Physical Adapters list within the Networking or Storage tab for the host. It appears in the PCI Devices list in the Settings tab for the host.
Example:

VF naming conventions are based on the Alternative Routing ID Interpretation (ARI) section of the PCIe Specifications. The VF name is dependent on the PCI slot number and port number of the physical adapter it is attached to. The rule of thumb is as follows:

VF# = {first number of PCI slot#}.10.{2*VF#+port#}

Example 1:
PF: PCI slot 03, Port 1
00:03:00.1 Network controller: Intel Corporation 82599EB 10-Gigabit SFI/SFP+ Network Connection [vmnic1]
VFss 1-3:
00:03:10.1 Network controller: Intel Corporation 82599 Ethernet Controller Virtual Function [PF_0.3.1_VF_0]
00:03:10.3 Network controller: Intel Corporation 82599 Ethernet Controller Virtual Function [PF_0.3.1_VF_1]
00:03:10.5 Network controller: Intel Corporation 82599 Ethernet Controller Virtual Function [PF_0.3.1_VF_2]

Example 2:
PF: PCI slot 04, Port 0
00:04:00.0 Network controller: Intel Corporation 82599EB 10-Gigabit SFI/SFP+ Network Connection [vmnic6]
VFss 1-3:
00:04:10.0 Network controller: Intel Corporation 82599 Ethernet Controller Virtual Function [PF_0.4.0_VF_0]
Configuring Intel® Ethernet Converged Network Adapter-Enabled Virtual Functions on VMware® ESXi® 5.1

00:04:10.2 Network controller: Intel Corporation 82599 Ethernet Controller Virtual Function [PF_0.4.0_VF_1]

00:04:10.4 Network controller: Intel Corporation 82599 Ethernet Controller Virtual Function [PF_0.4.0_VF_2]

2. Assignment to VMs
   1. Power off the virtual machine.
   2. Click the Manage tab of the virtual machine, and select Settings > VM Hardware.
   3. Click Edit.
   4. Expand the Memory section, and set the Limit to Unlimited.
   5. In vSphere or vCenter Server 5.1, in the Inventory > Hosts and Clusters view, right-click on the desired VM to assign a VF and select Edit Settings.
7. For Device Type, Select PCI Device.
8. Specify the VF to add to the VM from the drop-down box and click Next

Adding a virtual function as a PCI device to a virtual machine sets memory reservation to the memory size of the virtual machine.

**Note:** Once a VF has been assigned to a virtual machine, it does not depopulate from the list of available pass-through devices. However, it should not be assigned again. Also, keep in mind the VF naming convention explained earlier in this section.

9. Complete the configuration change and start the VM to begin using the VF in the guest OS.
3. **Warnings on VF assignment**

- Assigning one VF to multiple VMs is *not* supported. Note that this might inadvertently happen if a VM already having a VF assigned to it is cloned or if a VF previously assigned during configuration of the VM for VF connectivity is selected.

- If a network adapter has been configured for SR-IOV and has created VFs, its physical function should NOT be configured for pass-through mode. In ESXi 5.1, although the physical functions configured with VFs are not greyed out from pass-through mode selection under **Advanced Settings > edit**, they should *not* be selected for passthrough mode.

If the PF is used, all VFs move to the next PF, change PCIe IDs and impact all VMs already assigned to VFs. Also, if a new PCIe adapter is added to the system, it can also adversely affect DirectPath assignments if the new adapter forces the PCIe ID to change on the SR-IOV enabled port.

- VMware limits support for SR-IOV to specified Guest operating systems only. As an extension of DirectPath I/O, all limitations and rules that apply to VMware DirectPath I/O also apply to SR-IOV in ESXi 5.1. Please consult the VMware documentation for OS support and supported VF functionality.
2.4 VF Use in Guest Operating Systems

2.4.1 Driver Support

- For Microsoft Windows Server 2008 R2 guests, please download and install the latest network adapter drivers from Intel.com for VF support.
  
  [Link to drivers]

- For RHEL6.x guests, VF support is included in the Intel inbox drivers.

2.4.2 Verification

Once VFs have been assigned to VMs and the VMs are loaded with a VMware SR-IOV supported guest OS, confirm the presence of the VF in the Guest OS.
In Windows Server 2008 R2 guest console open Device Manager or Network Connections to view Virtual Function assigned.

And/or Network Connections:
• In RHEL 6.x guest console:

Use `lspci | grep -i intel | grep -i 'ethernet\|network'` to show Virtual Functions

```
[root@rhel61vm1 ~]# lspci | grep -i intel | grep -i 'ethernet\|network'
13:00.0 Ethernet controller: Intel Corporation 82599 Ethernet Controller (rev 01)
1b:00.0 Ethernet controller: Intel Corporation 82599 Ethernet Controller (rev 01)
```

Or `ethtool -i ethX` to show Virtual Function driver

```
[root@rhel61vm1 ~]# ethtool -i eth1
driver: ixgbevf
version: 1.0.19-k8
firmware-version: N/A
bus-info: 0000:13:00.0
```

**2.4.3 Traffic monitoring of NICs in pass-through mode:**

When a network adapter is in pass-through mode, its physical device is no longer visible to ESXi host monitoring tools. Network traffic monitoring in VMware vSphere 5.1 Client or VMware vCenter 5.1 Server is, therefore, not possible for adapters in pass-through mode. Use of ethtool on the ESXi CLI or tools provided in the guest OS are recommended for monitoring network traffic for VFs.
2.5 Advanced Topics

- In Red Hat Enterprise Linux 6.x, the Linux channel bonding driver is agnostic to the difference between a network interface on a PF and one on a VF. This enables bonding multiple VFs from multiple unique ports to the same VM. However, since teaming assigns a single MAC address to multiple ports, MAC address anti-spoofing security blocks traffic when bonded ports are configured for load balancing or failover.

- SR-IOV and Storage Over Ethernet: While FCoE and iSCSI are not supported over VFs in ESXi 5.1, an interesting use case for Intel's dual port 10Gb CNAs is to use one port of the CNA to provide FCoE or iSCSI storage to the ESXi host, while the other port is divided into VFs using SR-IOV and provides low latency/high throughput traffic to guest VMs.

- If using systems with the Intel Xeon processor E5-2600, the network adapter to be configured for SR-IOV must be in a PCIe slot provided by the Intel Integrated I/O PCI root complex directly off the processor.
3.0 Summary

Intel's “best-of-breed” 10 GbE solutions are now available with I/O Virtualization capabilities. SR-IOV features supported by Intel Ethernet CNAs offer the increased performance, lower latency and decreased CPU utilization of a direct I/O device while still allowing one physical port to provide connectivity to multiple VMs. SR-IOV on Intel Ethernet CNAs is now supported by VMware ESXi 5.1, Windows Server 2012 with Hyper-V, and mainstream Linux distributions.

Other Intel documentation of interest:

Which Intel® Ethernet Adapters and controllers support SR-IOV?
FAQs: Using SR-IOV with Intel® Ethernet Converged Network Adapters

4.0 Customer Support

Intel Customer Support offers a broad selection of programs including phone support and warranty service. For more information, contact us at:

support.intel.com/support/go/network/adapter/home.htm

Note: Service and availability may vary by country.

5.0 For Product Information

To speak to a customer service representative regarding Intel products, please call:
1-800-538-3373 (U.S. and Canada) or visit
support.intel.com/support/go/network/contact.htm
for the telephone number in your area.