

Intel® Ethernet and Configuring Single Root I/O Virtualization (SR-IOV) on Microsoft* Windows* Server 2012 Hyper-V

Technical Brief v1.0

September 2012

Legal Statements

INFORMATION IN THIS DOCUMENT IS PROVIDED IN CONNECTION WITH INTEL PRODUCTS. NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. EXCEPT AS PROVIDED IN INTEL'S TERMS AND CONDITIONS OF SALE FOR SUCH PRODUCTS, INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO SALE AND/OR USE OF INTEL PRODUCTS INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

A "Mission Critical Application" is any application in which failure of the Intel Product could result, directly or indirectly, in personal injury or death. SHOULD YOU PURCHASE OR USE INTEL'S PRODUCTS FOR ANY SUCH MISSION CRITICAL APPLICATION, YOU SHALL INDEMNIFY AND HOLD INTEL AND ITS SUBSIDIARIES, SUBCONTRACTORS AND AFFILIATES, AND THE DIRECTORS, OFFICERS, AND EMPLOYEES OF EACH, HARMLESS AGAINST ALL CLAIMS COSTS, DAMAGES, AND EXPENSES AND REASONABLE ATTORNEYS' FEES ARISING OUT OF, DIRECTLY OR INDIRECTLY, ANY CLAIM OF PRODUCT LIABILITY, PERSONAL INJURY, OR DEATH ARISING IN ANY WAY OUT OF SUCH MISSION CRITICAL APPLICATION, WHETHER OR NOT INTEL OR ITS SUBCONTRACTOR WAS NEGLIGENT IN THE DESIGN, MANUFACTURE, OR WARNING OF THE INTEL PRODUCT OR ANY OF ITS PARTS.

Intel may make changes to specifications and product descriptions at any time, without notice. Designers must not rely on the absence or characteristics of any features or instructions marked "reserved" or "undefined". Intel reserves these for future definition and shall have no responsibility whatsoever for conflicts or incompatibilities arising from future changes to them. The information here is subject to change without notice. Do not finalize a design with this information.

The products described in this document may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Contact your local Intel sales office or your distributor to obtain the latest specifications and before placing your product order.

Intel and Intel logo are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2012, Intel Corporation

Table of Contents

| | |
|---|-----------|
| Legal Statements | 3 |
| Table of Contents | 4 |
| Revision History | 4 |
| 1 Introduction | 5 |
| 1.1 Deploying IO Virtualization in Microsoft Windows Server 2012 Hyper-V..... | 5 |
| 1.1.1 SR-IOV Overview in Microsoft Windows Server 2012 Hyper-V | 5 |
| 1.1.2 Virtual Function Overview | 6 |
| 1.1.3 Live Migration & SR-IOV | 6 |
| 1.2 Live Migration & SR-IOV | 7 |
| 1.3 Hardware Requirements | 7 |
| 1.4 Software Requirements..... | 7 |
| 2 Installation and Configuration | 7 |
| 3 Summary | 14 |
| 4 Customer Support | 14 |
| 5 Product Information | 14 |

Revision History

| Date | Revision | Reference # | Description |
|----------------|----------|-------------|-----------------|
| September 2012 | 1.0 | | Initial Release |
| | | | |

1 Introduction

The Intel® Ethernet Converged Network Adapter and Intel® Ethernet Server Adapter family of adapters introduced numerous industry-leading features that are helping data center administrators implement innovative solutions for difficult and challenging connectivity problems. I/O Virtualization is one of the fastest growing usage models within the data center.

1.1 Deploying IO Virtualization in Microsoft Windows Server 2012 Hyper-V

Many of the 10Gb Converged Network Adapters and 1Gb Server Adapters that are based on Intel® Ethernet Controllers include Intel® Virtualization Technology for Connectivity (Intel® VT-c) which is a collection of I/O virtualization technologies that enables lower CPU utilization, reduced system latency, and improved networking throughput. Intel® VT-c is one of the key component technologies of Intel® Virtualization Technology that have been included as part of the Microsoft Windows Server 2012 Hyper-V.

Virtual Machine Device Queues (VMDq) provides hardware assists to the hypervisor that improves traffic management within the server by offloading traffic sorting and routing from the hypervisor's virtual switch to the Intel® Ethernet Controller. Working in conjunction with Microsoft Virtual Machine Queues*, VMDq enables traffic steering and balanced bandwidth allocation across the Intel Ethernet Controller's multiple hardware queues.

PCI-SIG Single Root I/O Virtualization and Sharing (SR-IOV) support allows multiple virtual machines to bypass the hypervisor networking stack and talk directly to unique resources on Intel® Ethernet Converged Network Adapters. The PCI-SIG SR-IOV standard offers an advanced, nonproprietary approach to providing I/O resources directly to virtual machines that helps significantly reduce latency, reduce CPU utilization and increase throughput.

1.1.1 SR-IOV Overview in Microsoft Windows Server 2012 Hyper-V

The single root I/O virtualization (SR-IOV) interface is an extension to the PCI Express (PCIe) specification. SR-IOV allows a network adapter to separate access to its resources among various PCIe hardware functions. These functions consist of the following types:

- A PCIe Physical Function (PF). This function is the primary function of the device and advertises the device's SR-IOV capabilities. The PF is associated with the Hyper-V parent partition in a virtualized environment.
- One or more PCIe Virtual Functions (VFs). A VF is a lightweight PCIe function on a network adapter that supports the SR-IOV interface. The VF is associated with the PCIe Physical Function (PF) on the network adapter, and represents a virtualized instance of the network adapter. Each VF has its own PCI Configuration space. Each VF also shares one or more physical resources on the network adapter, such as an external network port, with the PF and other VFs. Each VF is associated with a Hyper-V child partition in a virtualized environment.

1.1.2 Virtual Function Overview

A VF is not a full-fledged PCIe device. However, it provides a basic mechanism for directly transferring data between a Hyper-V child partition and the underlying SR-IOV network adapter. Software resources associated for data transfer are directly available to the VF and are isolated from use by the other VFs or the PF. However, the configuration of most of these resources is performed by the PF miniport driver that runs in the management operating system of the Hyper-V parent partition.

A VF is exposed as a virtual network adapter (VF network adapter) in the guest operating system that runs in a Hyper-V child partition. After the VF is associated with a virtual port (VPort) on the NIC switch of the SR-IOV network adapter, the virtual PCI (VPCI) driver that runs in the VM exposes the VF network adapter. Once exposed, the PnP manager in the guest operating system loads the VF miniport driver.

Each PF and VF is assigned a unique PCI Express Requester ID (RID) that allows an I/O memory management unit (IOMMU) to differentiate between different traffic streams and apply memory and interrupt translations between the PF and VFs. This allows traffic streams to be delivered directly to the appropriate Hyper-V parent or child partition. As a result, non-privileged data traffic flows from the PF to VF without affecting other VFs.

SR-IOV enables network traffic to bypass the software switch layer of the Hyper-V virtualization stack. Because the VF is assigned to a child partition, the network traffic flows directly between the VF and child partition. As a result, the I/O overhead in the software emulation layer is diminished and achieves network performance that is nearly the same performance as in non-virtualized environments.

1.1.3 Live Migration & SR-IOV

In Windows Server 2012, Live Migration can be performed with SR-IOV being used by a VM. If the source and target systems support SR-IOV and the target has an available VF, the VM will use the virtual function. If not, the VM will revert to the traditional path (VM-Bus). Each SR-IOV capable network adapter exposes a fixed number of Virtual Functions, which can be obtained by running the PowerShell command "Get-NetAdapterSriov".

Each PF and VF is assigned a unique PCI Express Requester ID (RID) that allows an I/O memory management unit (IOMMU) to differentiate between different traffic streams and apply memory and interrupt translations between the PF and VFs. This allows traffic streams to be delivered directly to the appropriate Hyper-V parent or child partition. As a result, non-privileged data traffic flows from the PF to VF without affecting other VFs.

SR-IOV enables network traffic to bypass the software switch layer of the Hyper-V virtualization stack. Because the VF is assigned to a child partition, the network traffic flows directly between the VF and child partition. As a result, the I/O overhead in the software emulation layer is diminished and achieves network performance that is nearly the same performance as in non-virtualized environments.

1.2 Live Migration & SR-IOV

In Windows Server 2012, Live Migration can be performed with SR-IOV being used by a VM. If the source and target systems support SR-IOV and the target has an available VF, the VM will use the virtual function. If not, the VM will revert to the traditional path (VM-Bus).

Each SR-IOV capable network adapter exposes a fixed number of Virtual Functions, which can be obtained by running the PowerShell command "Get-NetAdapterSriov".

The document shows how to make use of VFs using Windows Server 2012, which is included in the Windows Server 2012 distribution.

1.3 Hardware Requirements

- The following list shows appropriate components
 - Intel® Ethernet Controller X540 Family
 - Intel® 82599 10 Gigabit Ethernet Controller Family
 - Intel® Ethernet Controller I350 Family
 - Intel® 82576 Gigabit Ethernet Controller

 - Intel® Ethernet Converged Network Adapter X540 Family
 - Intel® Ethernet Converged Network Adapter X520 Family
 - Intel® Ethernet Server Adapter I350 Family
 - Intel® Gigabit ET/EF Server Adapter Family
- A server platform that supports Intel® Virtualization Technology for Directed I/O (Intel® VT-d) and the PCI-SIG* Single Root I/O Virtualization (SR-IOV)
- A server platform with an available PCI Express* v2.0 eight-lane, 5.0 GT/s (Gen2) PCI slot (PCI Express v2.0 and v1.0 lane requirements for 2 port 10Gb and 4 port 1Gb)

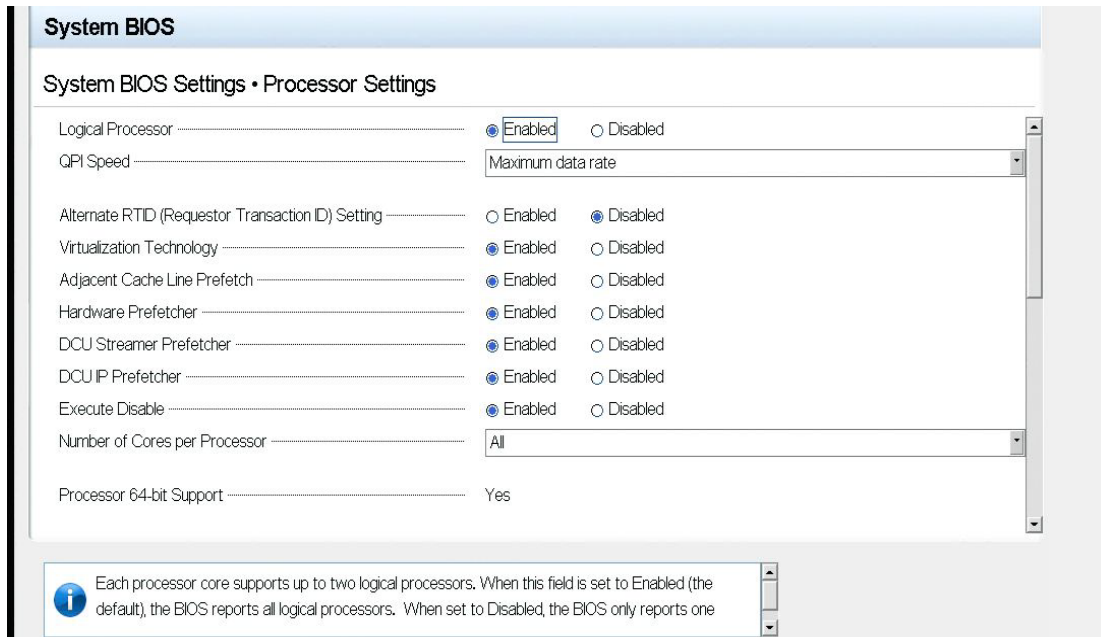
1.4 Software Requirements

- Windows Server 2012* Build 8370

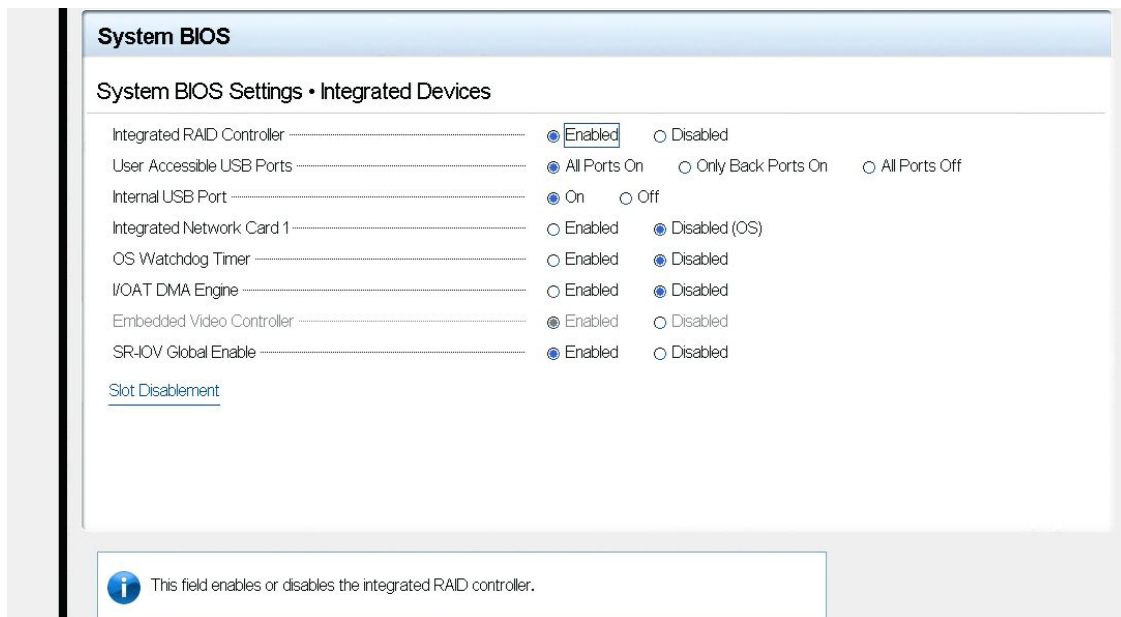
2 Installation and Configuration

- Install the Intel® Converged Network Server Adapter X520 in an available PCI-Express v2.0 x8 slot. (Ensure that the x8 slot is electrically connected as a x8, some slots are physically x8 but electrically support only x4. Verify this with your server manufacturer or system documentation.)
- Power up the server and enter the server's BIOS setup making sure the virtualization technology and Intel® VT-d features are enabled. The screen shots below are examples of the options to change.

Note: The location is BIOS-dependent.



Make sure Virtualization Technology is enabled.

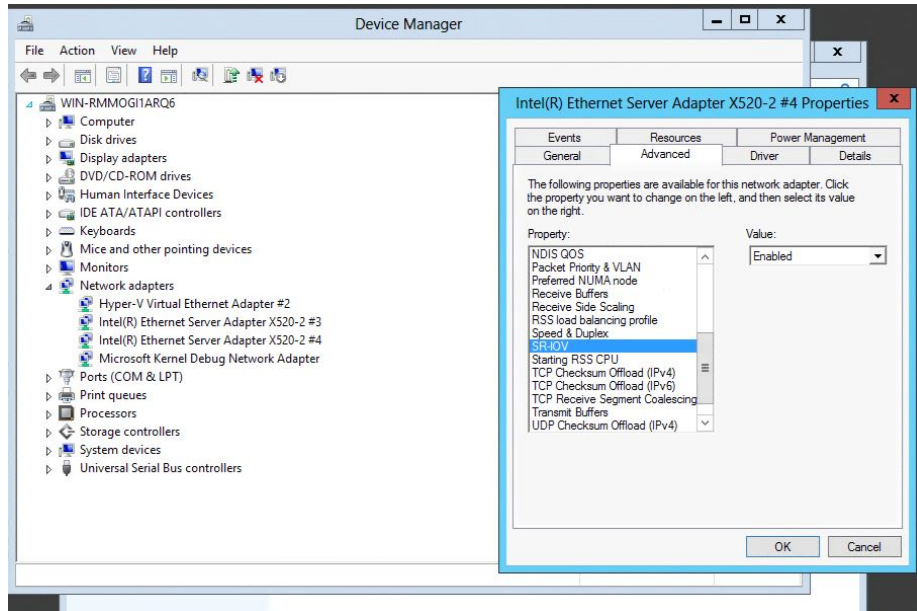


Make sure SR-IOV is enabled

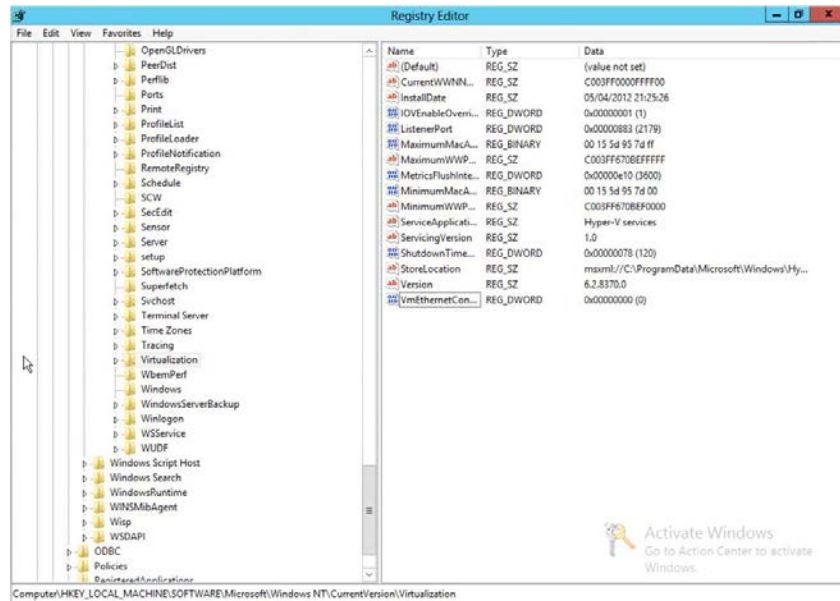
- Install Windows Server 2012 on the server platform.
- Make sure all the drivers for the adapter have been installed during the

Operation System installation.

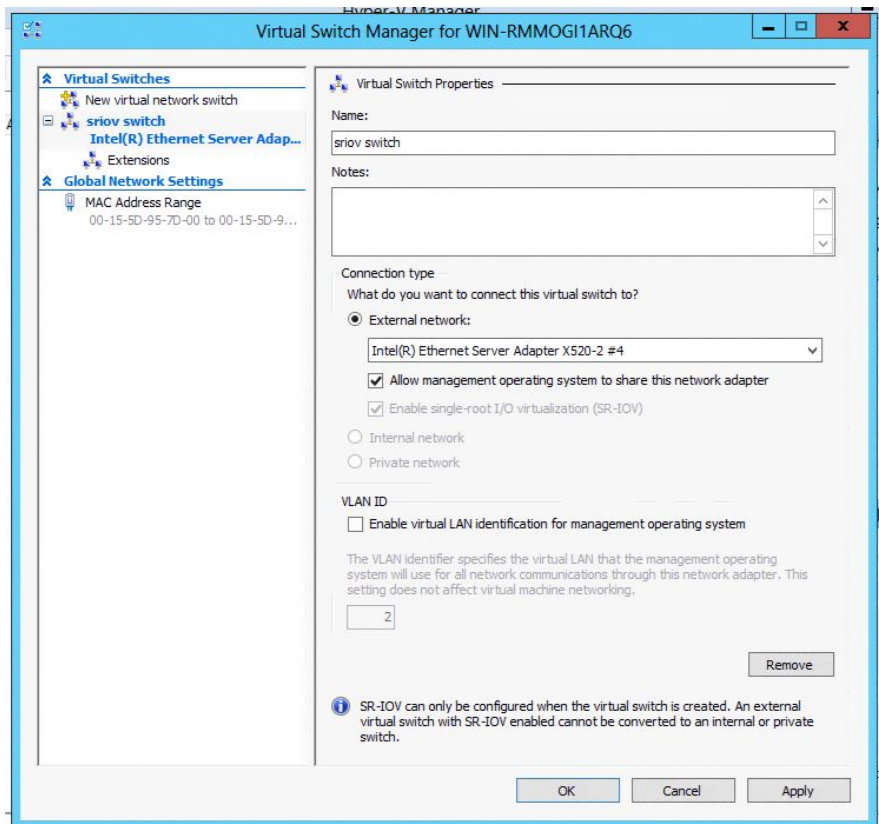
- The Windows Server 2012 installation will require updates to be installed. It will also require rebooting the system.
- Log in to the Server with Administrator privileges.
- Open the Device Manager and select the adapter interface you are using.
- Make sure SR-IOV is enabled.



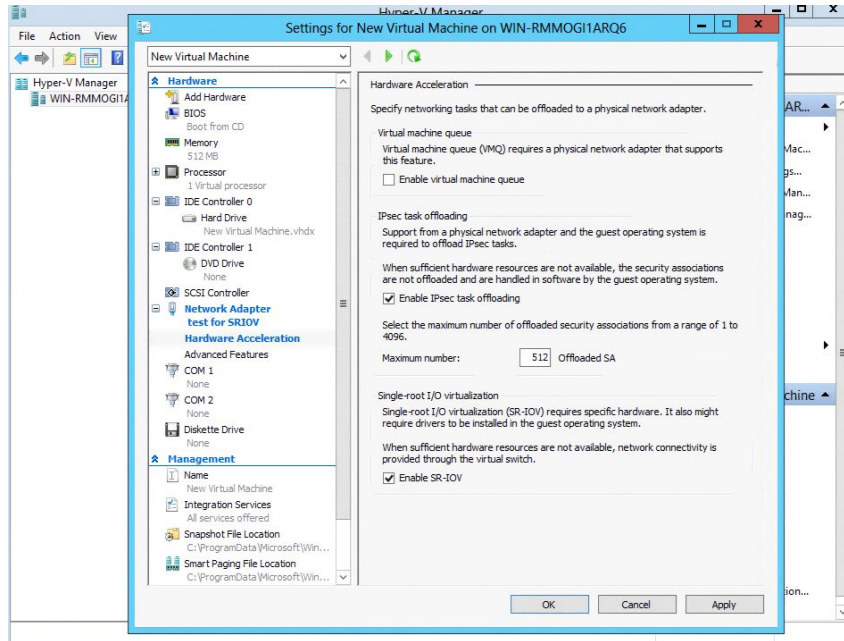
- Enable the `IovEnableOverride` field in the Registry with a Key value of 1 by performing the following:
 1. Open `regedit.exe` and browse to the following location:
`HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion\Virtualization`
 2. Insert a registry key of type DWORD (32-bit):
 3. `IovEnableOverride`
 4. Set the value of that key to 1.



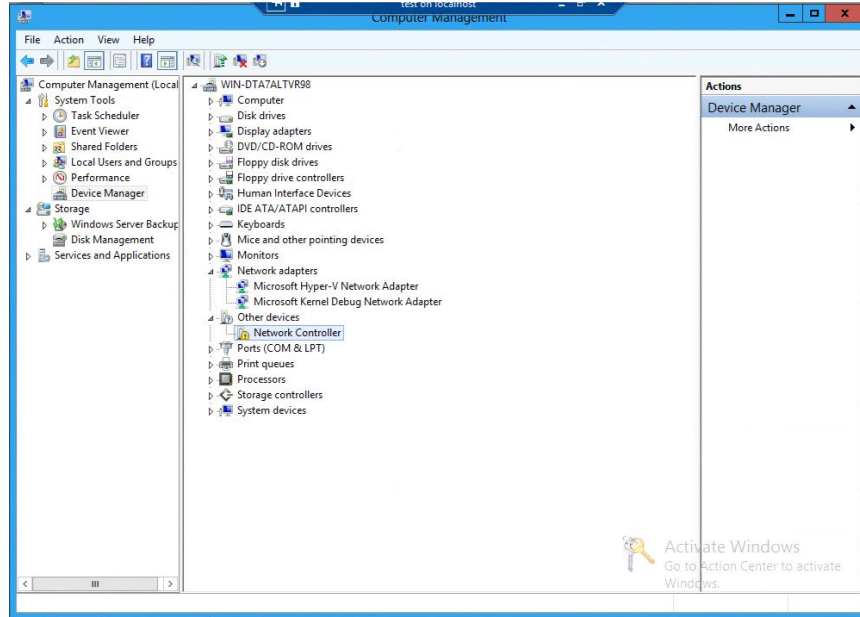
- The vmms (Virtual Machine Management Service) is then restarted . Restart the system or by using the net stop/start command from Powershell:
 - *net stop vmms*
 - *net start vmms*
- Open the Hyper-V Virtual Switch Manager. Using the GUI, create a new virtual switch from a supported Intel® Ethernet Converged Network Adapter X520, checking the "Enable Single Root I/O Virtualization (SR-IOV)" option.



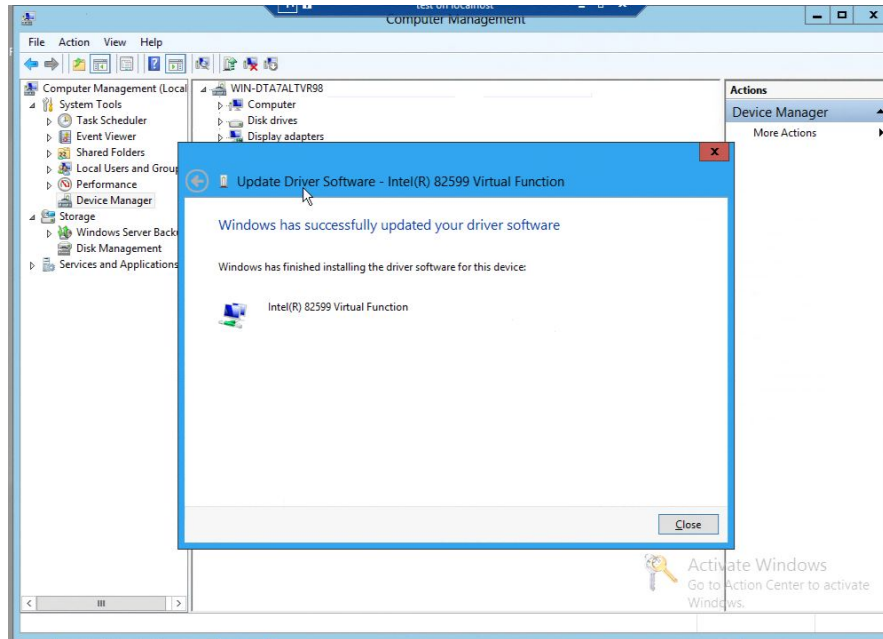
- Enable SR-IOV in the Virtual Machine settings. If there are insufficient hardware resources (due to non-SR-IOV-capable hardware or lack of resources) or, if 'Enable SR-IOV' is not checked, the Virtual Machine will default to a synthetic network adapter.
- VMQ and SR-IOV capabilities are mutually exclusive on the same instance of a Network Adapter. However, it is possible to have two instances of the same physical network adapter in the same VM—one that is VMQ-enabled and one that is SR-IOV-enabled.



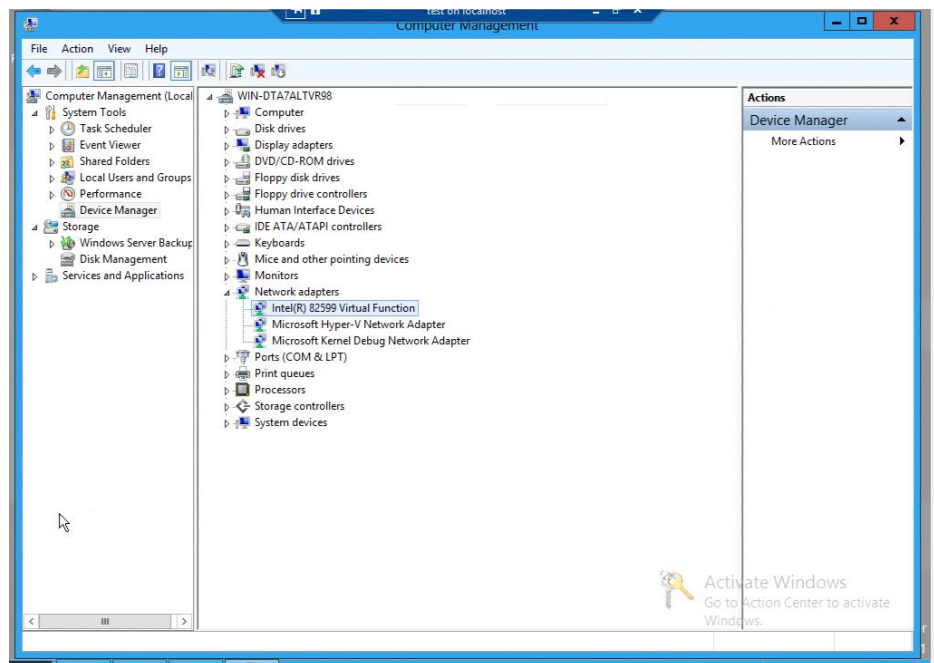
- Start or create a virtual machine. If SR-IOV has been correctly setup and supported on your system, an Unknown device will appear in the Device Manager. This is most likely the Virtual Function for your Intel® Ethernet Server Adapter.



- Obtain the drivers from the Windows Update site. Otherwise, download the drivers from a trusted source and install them via Plug and Play.



- This is the Virtual Function driver that provides the enablement of SR-IOV. This provides the virtual machines with direct access to dedicated resources in the Ethernet controller.



- The Virtual Function is configured for DHCP but you can assign a static IP address if needed. The VF is then ready to communicate

3 Summary

With Intel Ethernet Converged Network Adapters and the Microsoft Windows Server 2012 with Hyper-V applications running in Virtual Machines, the system can benefit from the low latency provided by SR-IOV. In addition, virtualized network appliances used for security, load balancing, and other functions are prime candidates for the SR-IOV approach.

4 Customer Support

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

<http://www.intel.com/support/go/network/adapter/home.htm>

(Service and availability may vary by country.)

5 Product Information

To speak to a customer service representative regarding Intel products, please call 1-800-538-3373 (U.S. and Canada) or visit

<http://www.support.intel.com/support/go/network/contact.htm>

for the telephone number in your area.