

Software User Manual 57712-k



# NIC Partitioning (NPAR) Setup Guide

### **Revision History**

Revision	Date	Change Description
2CS57712-SWUM102-R	12/08/11	Updated:
		<ul> <li>T6.4 software release by adding FCoE, DCB, VMWare<sup>®</sup> ESX/ ESXi 4.1, and BACS4 information.</li> </ul>
2CS57712-SWUM101-R	3/31/11	Updated:
		<ul> <li>"NPAR" in title and footer</li> </ul>
		<ul> <li>"NIC Partition", "NIC Partitioned", and "NIC Partitioning" to "NPAR" throughout the document</li> </ul>
		<ul> <li>"Single Function" to "SF" throughout the document</li> </ul>
2CS57712-SWUM100-R	03/15/11	Initial release

Broadcom Corporation 5300 California Avenue Irvine, CA 92617

© 2011 by Broadcom Corporation All rights reserved Printed in the U.S.A.

Broadcom<sup>®</sup>, the pulse logo, Connecting everything<sup>®</sup>, and the Connecting everything logo are among the trademarks of Broadcom Corporation and/or its affiliates in the United States, certain other countries and/ or the EU. Any other trademarks or trade names mentioned are the property of their respective owners.

# **Table of Contents**

About This Document
Purpose
Audience
Acronyms and Abbreviations
Technical Support
Configuring NPAR
Using the Unified Server Configurator
Supported Operating Systems1
Viewing and Configuring the Partitions10
Windows Server 2008 R210
Installing the Latest Dell Drivers16
Viewing the Enabled Devices in Device Manager19
Broadcom Advanced Control Suite 4 (BACS4)23
Microsoft Windows Network Connections
Device PCIe Bus Location
Red Hat Enterprise Linux
VMWare ESX/ESXi 4.142
Setting MTU Sizes40
Setting MTU Sizes in Windows40
Setting MTU Sizes in Linux49
Setting MTU Sizes in VMWare ESX/ESXi 4.153
Examples
Equal Oversubscription Example5
Partitioned Oversubscription Example64
Weighted Oversubscription Example6
Oversubscription With One High Priority Partition Example69
Default Fixed Subscription Example7
Mixed Fixed Subscription and Oversubscription Example73
Mixed Weights and Subscriptions Example7

# **List of Tables**

Table 1:	Protocols Available in Operation Systems Versus SF and NPAR Mode	.15
Table 2:	Port, Function, MAC Address Example	35
Table 3:	Non-DCB Equal Oversubscription	57
Table 4:	DCB Equal Oversubscription	60
Table 5:	DCB Equal Oversubscription with one Lossless FCoE Offload	62
Table 6:	Non-DCB Partitioned Oversubscription	64
Table 7:	Non-DCB Weighted Oversubscription	67
Table 8:	Non-DCB Oversubscription With One High Priority Partition	69
Table 9:	Non-DCB Default Fixed Subscription	71
Table 10:	Non-DCB Mixed Fixed Subscription and Oversubscription	73
Table 11:	DCB Mixed Fixed Subscription and Oversubscription with Lossless FCoE Offload	75
Table 12:	Non-DCB Mixed Fixed Subscription and Oversubscription	.77

# About This Document

## Purpose

This document provides instructions on how to enable NIC Partitioning (NPAR) on a Dell<sup>®</sup> PowerEdge<sup>®</sup> M710HD and M915 Blade Servers installed with the Broadcom<sup>®</sup> 57712-k Converged Network Daughter Card (NDC) Dual Port 10 GbE A Fabric option (see Figure 1).



Figure 1: Broadcom 57712-k Converged NDC

## Audience

This document is written for the network administrator who wishes to partition the Broadcom network controller on a Microsoft<sup>®</sup> Windows Server 2008 R2, VMWare<sup>®</sup> ESX/ESXi 4.1, Oracle<sup>®</sup> Solaris, SUSE Linux Enterprise Server (SLES), and Red Hat Enterprise Linux<sup>®</sup> (RHEL) system with:

- up to eight functions (four per port) Ethernet enabled in addition to:
  - up to four functions (two per port) iSCSI HBA enabled (in operating systems where the specific HBA can be enabled).

or

• up to two functions (one per port) FCoE HBA enabled plus up to two functions (one per port) iSCSI HBA enabled (in operating systems where the specific HBA can be enabled).

## **Acronyms and Abbreviations**

In most cases, acronyms and abbreviations are defined on first use.

For a comprehensive list of acronyms and other terms used in Broadcom documents, go to: http://www.broadcom.com/press/glossary.php.

## **Technical Support**

Broadcom provides customer access to a wide range of information, including technical documentation, schematic diagrams, product bill of materials, PCB layout information, and software updates through its customer support portal (<u>https://support.broadcom.com</u>). For a CSP account, contact your Sales or Engineering support representative.

In addition, Broadcom provides other product support through its Downloads & Support site (<u>http://www.broadcom.com/support/</u>).

# **Configuring NPAR**

## Using the Unified Server Configurator

Use Dell's Unified Server Configurator (USC) to configure Broadcom's 57712-k NPAR parameters.

#### To configure NPAR with the USC

Enter USC during system boot up by selecting the UEFI boot option. See the Dell website (<u>http://www.dell.com/content/topics/global.aspx/power/en/simplify\_management?c=us&l=en&cs=555</u>) for more information on USC.



2. From the USC, select Hardware Configuration and the HII Advanced Configuration option.

DOLL UNIFIED SERVER CONFIGURATOR   LIFECYCLE CONTROLLER ENABLED							
	Hardware Configuration						
Home Lifecycle Log Platform Update	Configuration Wizards guide system device set up (Ex: RAID, iDRAC, Encryption). HII Advanced Configuration configures Human Interface Infrastructure(HII)– enabled devices (Ex: BIOS, NICs). Hardware Inventory is used to view or export the server's current hardware inventory or the factory shipped hardware inventory. Delete Configuration and						
Hardware Configuration OS Deployment Platform Restore	Reset Defaults deletes the Lifecycle Controller configuration and restores factory defaults.						
Hardware Diagnostics USC Settings About	Configuration Wizards HII Advanced Configuration Hardware Inventory						

**3.** All of the Broadcom Ethernet Controller devices should be displayed on this page. Select the desired 57712-k device port from the displayed list.

ardware Co	onfiguratio	on					
dvanced Co	onfiguratio	on					
10 13 1000	100 No. 100 No. 10						
System BI	IOS Setting	gs					
<u>System BI</u> Broadcom	IOS Setting NetXtreme	gs II 10	Gigabit	Ethernet	- 00:10:1	18:6F:D2:A4	T

**4.** From the **Broadcom Main Configuration Page**, select **Device Configuration Menu** to turn on or off the NPAR mode of operation for the selected devices.

-	UNIFIED SERVER CONFIGURATOR   LIFECYCLE CONTROLLER ENABI
I	Broadcom NetXtreme II 10 Gigabit Ethernet – 00:10:18:6F:D2:A4
1	Broadcom Main Configuration Page
	Firmware Image Menu
	Device Configuration Menu
	MBA Configuration Menu
	iSCSI Boot Configuration Menu
	NIC Partitioning Configuration Menu
1	Chip Type BCM57712E A1
1	PCI Device ID 1663
l	Bus:Dev:Func 06:00:00
ļ	Link Status UP
l	Permanent MAC Address
	00:10:18:6F:D2:A4

5. From the **Device Configuration** window, select either **Enabled** (NPAR mode, where each port has four functions) or **Disabled** (Single Function (SF) mode, where each port has one function) for the device's NPAR mode.

Broadcom	NetXtreme	II 10	Gigabit	Ethernet	- 00	:10:18:6	F:D2:A4	
Device Co	nfiguratio	n						
Device MA	C Addresse	s:						
Broadcom	NetXtreme	II 10	Gigabit	Ethernet	- 00	:10:18:6	F:D2:A4	
Broadcom	NetXtreme	II 10	Gigabit	Ethernet	- 00	:10:18:6	F:D2:A6	
NIC Parti	tion							
Enable	d							
Disab1	ed							
Enable	d							

6. Return to the Broadcom Main Configuration Page to edit the four partitions' attributes by selecting the NIC Partitioning Configuration Menu.



7. This window gives access to the Global Bandwidth Allocation Menu, the Flow Control settings, and each of the four partition's protocol settings. First select the Global Bandwidth Allocation Menu option.



8. The Global Bandwidth Allocation Menu window controls the Relative Bandwidth Weight and Maximum Bandwidth parameters for all four partitions. See "Broadcom Advanced Control Suite 4 (BACS4)" on page 23 for more information on how BACS4 can also be used to control these settings.

Broadcom I	NetXtreme II 10 Gigabit Ethernet – 00:10:18:6F:D2:A4
Global Bar	ndwidth Allocation Menu
Partition	1 Relative Bandwidth Weight
0	
Partition	1 Maximum Bandwidth
10	
Partition	2 Relative Bandwidth Weight
0	
Partition	2 Maximum Bandwidth
10	
Partition	3 Relative Bandwidth Weight
0	
Partition	3 Maximum Bandwidth
80	
Partition	4 Relative Bandwidth Weight
0	
Partition	4 Maximum Bandwidth
80	

The **Relative Bandwidth Weight** is the value the port gives to that single partition's send or outgoing traffic with respect to any other actively sending partitions on that port when there is more send traffic pending on the four partitions than send bandwidth available on that port. It is more than just a minimum bandwidth setting. This setting follows these rules:

- The individual configurable value range is 0 to 100.
- The **SUM** of a single port's four partitions values **MUST** be either exactly **100** or exactly **0** (which means all four of the partitions are set to 0).
- If one or more of a partition's weight is set to **0**, but the sum is **100** (i.e. not all of the partitions are set to zero) then that partition's relative bandwidth weight value is effectively **1** with respect to allocation calculations.
- Setting all four partition's values to **0** will give every traffic flow on every partition equal access to the ports available bandwidth without regard to which partition they are on unless restricted by the partition's Maximum Bandwidth settings.
- If the sum of the relative bandwidth weights is **100** and there is more than one type of traffic flow on a specific partition (i.e. iSCSI and L2 Ethernet or FCoE and L2 Ethernet) then the traffic on that specific partition will share the bandwidth being allocated as if there was only one traffic flow on that partition.
- The weight applies to all enabled protocols on that partition.
- The Relative Bandwidth Weight is not applicable when in Data Center Bridging (DCB) mode. In DCB mode, all traffic flows act as if their Relative Bandwidth Weight is set to all 0s.
- The NPAR transmit direction traffic flow rates are affected by the three main modes in the following ways:
  - In non-DCB mode where the sum of the partition's Relative Bandwidth Weights equal 100, each Partition's combined traffic flow is equally scheduled to transmit within the limitations of the partition's Relative Bandwidth Weight and Maximum Bandwidth settings and the overall connection's link speed. This means a specific partition's Relative Bandwidth Weight value will restrict the traffic flows sharing that partition's bandwidth allocation, as if one combined traffic flow with respect to the other actively sending partitions. The partition's send flow rate is based on the ratio of that partition's individual weight verses the aggregated weights of all the other actively sending partitions. Furthermore, each partition's combined traffic flow will be capped by that partition's Maximum Weight setting. See the User Guide's examples for more details. The actual inter-partition ratio of the two sharing traffic flows is controlled by the host OS. Think of the dynamic weight ratio as a variable sized funnel that could be further restricted by the Maximum Bandwidth fixed sized funnel with the OS determining how the sharing traffic types are pouring into the combined funnels.
  - In non-DCB mode where the sum of the partition's Relative Bandwidth Weights equals zeros (i.e., each partition's Relative Bandwidth Weight is set to zero), each individual traffic flow (i.e. Ethernet or iSCSI Offload or FCoE Offload) is equally scheduled to transmit within the limitations of the partition's Maximum Bandwidth and the overall connection's link speed. This means if the Maximum Bandwidth of a specific partition is set to less than 100%, then the traffic flows sharing that partition will be further restricted to where their combined traffic flow bandwidth will be capped by that per partition setting. If all four partition's individual Maximum Bandwidths are set to 100% (i.e. they are unrestricted), then each actively sending traffic flow (without regard to which partition they are on) will equally share the transmit directions total bandwidth (i.e. TX link speed). The actual inter-partition ratio of the two sharing traffic flows is controlled by the host OS. Think of the Maximum Bandwidth as a fixed sized funnel with the OS determining how the two sharing traffic types are pouring into that funnel.
  - In DCB mode, all of the Partition's Relative Bandwidth Weights are disregarded and the individual traffic flows are scheduled to transmit within the limitations of the Priority Group's ETS value

(determined by it's Traffic Type) and each partition's Maximum Bandwidth setting and the overall connections link speed. For example, the FCoE traffic type could be assigned to Priority Group 1 (PG1) and all of the other traffic types (iSCSI and Ethernet) could be assigned to another Priority Group (such as PG0). Each Priority Group has it's own ETS value (which works similarly to a minimum bandwidth setting). DCB Lossless iSCSI (iSCSI-TLV) could be used in place of FCoE for a similar effect where the Lossless iSCSI Offloaded traffic would go through it's assigned Priority Group while the Lossy Ethernet traffic would go through another. Similarly to the other two rate controlling modes, the host OS determines the actual inter-partition traffic ratio for the cases where two traffic types share the same partition.



**Note:** A traffic type's send flow rate will be approximately the ratio of its individual partition's relative bandwidth weight setting divided by the sum of the relative bandwidth weights of all the partitions currently actively sending on that port or that partition's maximum bandwidth setting, whichever is lower. In the case where the Relative Bandwidth Weights are all zeros OR in DCB mode, each traffic type will have an equal "weight" with respect to one another (see "Examples" on page 56).



**Note:** DCB mode is supported in Windows and some Linux (RHEL v6.x and SLES11 SP1) OS's on the 57712-k. VMWare ESX/ESXi 4.1 does not support DCB (which includes both FCoE and DCB Lossless iSCSI) on the 57712-k.

Each partition's **Maximum Bandwidth** settings can be changed in the same way and has a range of 1 to 100% in increments of 1% of the port's current Link Speed (at 10 Gbps this would be in ~100 Mbps increments and at 1 Gbps this would be in ~10 Mbps increments). This setting limits the most send bandwidth this partition will use and will appear as its approximate link speed in various places in the respective operating system even though the four partition's are sharing the same overall connection - i.e. the four partitions may advertise in the OS that their link speed is 10Gbps each, but they all share the same single 10Gbps connection. Displayed values may be rounded off by various applications. The Maximum Bandwidth value is applicable to both DCB and non-DCB modes of operation. The Maximum Bandwidth value is applicable to the send (TX) direction only.



**Note:** A partition's send Maximum Bandwidth setting does not affect a partition's receive direction traffic bandwidth, so the link speed displayed for the partition is for the send/transmit/outgoing direction only. All partitions receive direction maximum bandwidth is always the ports current Link Speed and is regulated by the attached switch port just as it is in SF mode when multiple (L2 Ethernet and iSCSI Hardware Offload and FCoE Hardware Offload) traffic protocol types are enabled.

The Maximum Bandwidth settings can be used to "oversubscribe" a port. This is done by setting the four partitions of that single port to having a total Maximum Bandwidth setting **SUM** of more than 100% (i.e., 10000 Mbps or 10 Gbps). That just means the various partitions will attempt to take as much bandwidth as allowed (by their specific setting maximum limits and weights) as their individual traffic flow needs change. In an oversubscription situation, the 57712-k will ration out free bandwidth based on the weights (sum is 0 verses sum is 100) and maximum settings and the mode (DCB verses non-DCB) it is in. The above example shows the first port's four partitions being set to 10+10+80+80 = 180, which means the port is 180% subscribed (18 Gbps) or 80% oversubscribed (i.e., 18 Gbps subscribed – 10 Gbps line rate available = 8 Gbps oversubscribed). The Maximum Bandwidth setting applies to all protocols enabled on that partition.



**Note:** When NPAR mode is first enabled or after a reset, the default values for all four partitions is **Relative Bandwidth Weight** = 0 and **Maximum Bandwidth = 25**.

See "Examples" on page 56 for more details on both the **Relative Bandwidth Weight** and **Maximum Bandwidth** settings affect traffic flow in DCB and non-DCB modes of operation.

1. Return to the previous **NIC Partition Configuration** window to change any of the four partitions protocol settings by selecting the specific partition here.



 In the specific Partition window, select which protocols it will support and also view the partitions assigned Networking MAC address, the Windows used iSCSI MAC address, the FCoE FIP MAC address, the FCoE Node WWN, and FCoE Port WWN values. BACS4 can also be used to control these settings.

Protocol selection follows these rules:

- A maximum of two iSCSI or one FCoE and one iSCSI Offload Protocols (HBA) can be enabled over any two of the four available partitions of a single port.
- The FCoE Offload Protocol is only available if DCB is also enabled and active on that port (i.e., the 57712-k port is connected to a DCB compliant and enabled link partner).
- The iSCSI Offload Protocol can function without DCB but if DCB Lossless iSCSI (iSCSI-TLV) is required, then DCB must be enabled and active on that port (i.e., the 57712-k port is connected to a DCB compliant and enabled link partner).
- Only one Offload Protocol (either iSCSI or FCoE) can be enabled per single partition in NPAR mode.
- For simplicity, using the first partition of a port for FCoE offload protocol is recommended since the FCoE port WWN will be the same for both SF and NPAR mode on the same port. This will make your Fiber Channel Forwarder (FCF) switch configuration much simpler.
- For Windows operating systems, you can have the Ethernet Protocol enabled on all, some, or none of the four partitions on an individual port simultaneously with any enabled offload protocols.
- For Linux OSs, the Ethernet protocol will always be enabled (even if disabled in USC).
- For simplicity, we recommend always using the first two partitions of a port for any iSCSI offload protocols.
- For Windows OSs, the Ethernet protocol does not have to be enabled for the iSCSI or FCoE offload protocol to be enabled and used on a specific partition.
- For VMWare ESX/ESXi 4.1, in NPAR mode, the host and hosted Virtual Machines (VMs) should only

connect to enabled Ethernet protocol adapters.

Par	tition 1
Eth	ernet Protocol
1	Enabled
isc	SI Offload Protocol
	Disabled
FCc	E Offload Protocol
1	Enabled
Net	work MAC Address
1	00:10:18:6F:D2:A4
Vir	tual Network MAC Address
1	00:10:18:6F:D2:A4
150	SI MAC Address
[	00:10:18:6F:D2:A5
Vir	tual iSCSI MAC Address
1	00:10:18:6F:D2:A5
FIP	MAC Address
1	00:10:18:6F:D2:A5
Vir	tual FIP MAC Address
1	00:10:18:6F:D2:A5

### **Supported Operating Systems**

The 57712-k SF and NPAR mode supported operating systems are shown in Table 1.



**Note:** The drivers may not be in the box.

		-	-				
		SF Mode		NPAR Mode			
Operating System	Ethernet	iSCSI Offload	FCoE Offload	Ethernet	iSCSI Offload	FCoE Offload	
Windows 2008 <sup>a</sup>	Yes	Yes	Yes	Yes	Yes	Yes	
Windows 2008 R2 <sup>a</sup>	Yes	Yes	Yes	Yes	Yes	Yes	
Windows 2008 R2 Hyper-V <sup>a</sup>	Yes	Yes	Yes	Yes	Yes	Yes	
Oracle Solaris 10u9	Yes	No	No	Yes	No	No	
Linux <sup>b</sup>	Yes	Yes	Yes <sup>c</sup>	Yes	Yes	Yes <sup>c</sup>	
VMWare ESX/ESXi 4.0	Yes	No	No	Yes	No	No	
VMWare ESX/ESXi 4.1	Yes	Yes	No	Yes	Not certified.	No	

Table 1: Protocols Available in Operation Systems Versus SF and NPAR Mode

a. DCB (DCBX/PFC/ETS) supported.

b. DCB (DCBX/PFC/ETS) supported in RHEL v6.x and SLES11 SP1 only.

c. FCOE offload supported in RHEL v6.x and SLES11 ASP1 only.

## Viewing and Configuring the Partitions

- Windows Server 2008 R2
- Red Hat Enterprise Linux
- VMWare ESX/ESXi 4.1

### Windows Server 2008 R2

#### **Installing the Latest Dell Drivers**

When the 57712-k cNDC is first installed, the iSCSI and FCoE devices may not appear. If the latest Dell driver is already present on the system, the 57712-k will be identified and the NDIS personality/protocol will be installed. If the latest NetXtreme II drivers are not present on the system, go to the Dell driver download web site (<u>http://support.dell.com/support/downloads/</u> under the specific Dell blade server platform) and install the latest NetXtreme II network drivers for your specific installation system.



Figure 2: Windows Device Manager

To enable the devices to be detected by the operating system, start BACS4 and while in 57712-k SF mode, select the port **System Device>Configuration>Resource Reservations**, check the boxes to enable the applicable iSCSI and FCoE Offload Engines protocols and then click **Apply**. Click **Yes** when the temporary network connection interruption warning displays and wait for the discovered devices to be installed by Windows - no reboot is necessary while in SF mode.

File View Action Filter Context Tools Teams iSCSI Help		
Filter: ALL VIEW 🔽 Configurations 🔽 🔽 Resource Reservations 🔽 Licenses 🔽 Ac	dvanced	
Explorer View 🗗	Information Configurations Statis	tics
Hosts     Hosts     MUN-RA3QJRK0731     Adapter1 (BCM5709 C0)     Adapter2 (BCM5709 C0)     Adapter3 (BCM5771 A0)     Adapter4 (57712 A1)     Port1     Double	Property     Advanced       • Advanced       • Resource Reservations       □• Pre-Allocated Resources       □• NDIS       □• TCP Offload Engine (TOE)       □• FCOE       • Maximum ICP Offload Engine       □• FCOE       • Maximum ICS Offload Engine       • Maximum ICS Offload Engine       • FCOE       • Memory consumption(%)       • TOE RSS       • FCOE Pending Tasks       • FCOE Pending Tasks Sessions       • Licenses	Value         Ivalue         Ivalue
		Reset

Figure 3: Broadcom Advanced Control Suite 4

If you are in NPAR mode, go to the **57712-k Adapter>Configuration>Property** window and click the **+** next to the NIC Partition to expand the menu. In the expanded menu, if the NIC Partition setting is unchecked (**Disabled**), change it to checked (**Enabled**) and then set the desired partition's Ethernet/NDIS, iSCSI and FCOE protocols. Also set the Relative Bandwidth Weights, Maximum Bandwidth settings and then click **Apply**. You must reboot the system for Windows to discover and install the device drivers.

Broadcom Advanced Control Suite 4							
File View Action Filter Context Tools Teams iSCSI Help							
Filter: ALL VIEW	ation 🔽 🔽 NIC Partition						
Explorer View 🗗	Information Configuration						
⊡. 000 Hosts	Property	Value					
🗄 📳 WIN-RA3QJRK0731	- NIC Partition						
🕀 🔝 Adapter 1 (BCM5709 C0)	Reset Configuration to Default						
🕀 😅 Adapter 2 (BCM5709 C0)	NIC Partition	Enabled					
🗄 😅 Adapter 3 (BCM 57711 A0)	- Port 0						
- Adapter 4 (57712 A1)	Flow Control	Auto					
- Port1	E Function 0						
	Etheret/Ndis	Enable					
	iSCSI	☐ Disable					
[0012] Broadc	FCoE	Finable					
🧿 [0027] Broadc	Relative Bandwidth Wei	5					
[0051] Broadcom I	Maximum Bandwidth (%)	10					
🖨 📅 [0053] Broadcom I	Function 2						
5 [0030] Broadc	Etheret/Ndis	☐ Disable					
E [0054] Broadcom L	··· iSCSI	☐ Disable					
	FCoE	☐ Disable					
	···· Relative Bandwidth Wei	5					
····· 💽 [0031] Broadd	Maximum Bandwidth (%)	10					
. E. → Port2	Evention 4						

Figure 4: Broadcom Advanced Control Suite Adapter Settings

After the devices are installed, the enabled devices (L2 Ethernet NDIS, FCoE and iSCSI) will be visible in the Windows Device Manager and BACS4. The following is the 57712-k's Device Manager's display in SF mode (see Figure 5).



Figure 5: Windows Device Manager

### Viewing the Enabled Devices in Device Manager

Windows shows all of the enabled devices in Device Manager with the respective USC-enabled NPAR protocols. The following example shows:

- Eight-enabled Ethernet protocol partitions (four possible per port) as the Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #50 through #57 in the Network Adapters section.
- Two enabled iSCSI protocol partitions (up to two are possible per port if no FCoE is enabled) as the Broadcom BCM57712 NetXtreme II 10 GigE Multifunction iSCSI Adapters #51 and #55 AND two enabled FCoE protocol partitions (one possible per port) as the Broadcom BCM57712 NetXtreme II 10 GigE Multifunction FCoE Adapters #52 and #56 in the Storage Controllers section
- Eight Broadcom BCM57712 NetXtreme II 10 GigE Multifunction virtual bus devices #50 through #57 in the System Devices section. These eight virtual bus system devices are always present and are not controlled by what protocol is enabled in USC.

Server Manager	
File Action View Help	
Server Manager (BROADCOM-ISCS	Device Manager
<ul> <li>Peatures</li> <li>Diagnostics</li> <li>Event Viewer</li> <li>Performance</li> <li>Device Manager</li> <li>Configuration</li> <li>Storage</li> </ul>	<ul> <li>Broadcom BCM5709C NetXtreme II GigE (NDIS VBD Client) #46</li> <li>Broadcom BCM5709C NetXtreme II GigE (NDIS VBD Client) #47</li> <li>Broadcom BCM5709C NetXtreme II GigE (NDIS VBD Client) #48</li> <li>Broadcom BCM5709C NetXtreme II GigE (NDIS VBD Client) #49</li> <li>Broadcom BCM5712 NetXtreme II 10 GigE (NDIS VBD Client) #50</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #51</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #52</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #52</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #53</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #54</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #57</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #57</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #57</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #57</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #57</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function iSCSI Adapter #51</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function FCOE Adapter #52</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function iSCSI Adapter #52</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function iSCSI Adapter #52</li> </ul>
	<ul> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function FCoE Adapter #56</li> <li>System devices</li> <li>Broadcom BCM5709C NetXtreme II GigE #46</li> <li>Broadcom BCM5709C NetXtreme II GigE #47</li> <li>Broadcom BCM5709C NetXtreme II GigE #48</li> <li>Broadcom BCM5709C NetXtreme II GigE #49</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #50</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #51</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #51</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #52</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #53</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #54</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #55</li> <li>Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #55</li> </ul>

Right click the specific device and select its **Properties** to access some of the advanced features of the device.

evice Manager	
	Update Driver Software Disable Uninstall Scan for hardware changes Properties

This brings up that device's property window. The following is the BCM57712 iSCSI device showing the HBA and connected target session information, send/receive statistics and ping test results.

roadcom BCM57712 NetXtreme II 10 GigE	Multi Function is	5C51 Ad 🗙
General Advanced Driver Details		
iSCSI MAC: 00:10:18:6F:D3:BF		
HBA Portals:		
Address	Mask/Scope	DHCP
1.1.1.7 FE80:0000:0000:0000:0210:18FF:FE6F:D3BF	255.255.0.0 2	No No
Sessions:	Address	Port
ign 2010-10 com broadcom tot 15a ram 15a	11115	3260
ign.2010-10.com.broadcom.tgt15b.ram15b	1.1.2.15	3260
iqn.2011-05.com.broadcom.tgt14a.ram14a	1.1.1.14	3260
iqn.2011-05.com.broadcom.tgt14b.ram14b	1.1.2.14	3260
Remote IP Address:		
1.1.1.12		Ping
Success!!		
Total Connections: 4		
Bytes Sent: 20692885 MB		
Bytes Received: 872225 MB		

The following shows the properties window for the FCoE device showing World Wide IDs, connected Fabric and Target information.

	LM57712 NetXt	reme II	no engle n	LUC AUG	JUEF # 79	Prop P
General A	dvanced Driver	Details				
Po	rt WWN: 20:00:0	0:10:18:8	8:E8:C5			
Nod	e WWN: 10:00:0	0:10:18:8	8:E8:C5			
FC	oE MAC: 00:10:1	8:88:E8:0	25			
abrics:						
Fabric Nar	me	NPIV	FCID	VLAN	Port WW	N
27:D9:00:0	D:EC:CA:04:81	No	62005B	2009	20:00:00:	10:18:88
-						
I argets:						Þ
I argets:	Port WWN		Node	wwn		Disc
I argets: FCID C60500	Port WWN 50:06:01:60:44:	:60:0C:CE	Node	WWN 01:60:C4:	60:0C:CE	Disc 20:00
I argets:           FCID           C60500           C61100           C62500	Port WWN 50:06:01:60:44: 50:06:01:69:44: 20:09:00:05:11	:60:0C:CE :60:0C:CE :12:70:42	Node 50:06: 50:06: 20:07	WWN 01:60:C4: 01:60:C4:	60:0C:CE 60:0C:CE 12:70:42	Disc 20:00 20:00
I argets:           FCID           C60500           C61100           C62500	Port WWN 50:06:01:60:44: 50:06:01:69:44: 20:08:00:0E:11:	:60:0C:CE :60:0C:CE :12:70:A2	Node 50:06: 50:06: 20:07:	WWN 01:60:C4: 01:60:C4: 00:0E:11:	60:0C:CE 60:0C:CE 12:70:A2	Disc 20:00 20:00 20:00
I argets:           FCID           C60500           C61100           C62500           ↓	Port WWN 50:06:01:60:44: 50:06:01:69:44: 20:08:00:0E:11	:60:0C:CE :60:0C:CE :12:70:A2	Node 50:06: 50:06: 2 20:07:	WWN 01:60:C4: 01:60:C4: 00:0E:11:	60:0C:CE 60:0C:CE 12:70:A2	Disc 20:00 20:00 20:00

The following shows the property window for the NDIS (Ethernet) device:

Broadcom BCM57712 NetXtreme	II 10 GigE (NDIS VBD Client) #50 🗙
General Advanced Driver Deta	als Power Management
BROADCOM. BROADCOM.	operties are available for this network e property you want to change on the lect its value on the right.
Property:	<u>V</u> alue:
IPv6 Large Send Offload Jumbo Packet Locally Administered Address Number of RSS Queues Pause On Exhausted Host Ring Priority & VLAN Receive Buffers (0=Auto) Receive Side Scaling TCP Connection Offload (IPv4) TCP Connection Offload (IPv6) Transmit Buffers (0=Auto) Wrtual Machine Queues VLAN ID VMQ Lookahead Split	Enabled
	OK Cancel

**Note:** In NPAR mode, MS Window's TCP Chimney Offload or TOE functionality can be enabled or disabled on a per partition granularity in this Advanced Properties control window and in BACS4's NDIS Advanced Properties control window.

The number of currently active Windows TOE connections can be viewed by using the netstat -not" command in a DOS window.

C:\Users\Administrator>netsta	at -not		
Active Connections			
Proto Local Address Offload State	Foreign Address	State	PID
TCP 1.1.1.24:49301	1.1.1.25:5010	ESTABLISHED	3352
TCP 1.1.1.24:49302	1.1.1.25:5012	ESTABLISHED	3352
TCP 1.1.1.24:49303	1.1.1.25:5001	ESTABLISHED	3352
TCP 1.1.1.24:49304	1.1.1.25:5013	ESTABLISHED	3352
TCP 1.1.1.24:49305	1.1.1.25:5008	ESTABLISHED	3352
TCP 1.1.1.24:49306	1.1.1.25:5003	ESTABLISHED	3352
TCP 1.1.1.24:49307	1.1.1.25:5014	ESTABLISHED	3352
TCP 1.1.1.24:49308	1.1.1.25:5000	ESTABLISHED	3352
TCP 1.1.1.24:49309	1.1.1.25:5004	ESTABLISHED	3352
0ffloaded TCP 1.1.1.24:49310	1.1.1.25:5006	ESTABLISHED	3352
Offloaded TCP 1.1.1.24:49311	1.1.1.25:5009	ESTABLISHED	3352
Offloaded TCP 1.1.1.24:49312	1.1.1.25:5015	ESTABLISHED	3352
Offloaded TCP 1.1.1.24:49313	1.1.1.25:5002	ESTABLISHED	3352
Offloaded TCP 1.1.1.24:49314	1.1.1.25:5011	ESTABLISHED	3352
Offloaded TCP 1.1.1.24:49315	1.1.1.25:5005	ESTABLISHED	3352
Offloaded TCP 1.1.1.24:49316	1.1.1.25:5007	ESTABLISHED	3352
Offloaded TCP 3.3.3.24:49289	3.3.3.24:49293	ESTABLISHED	3716
InHost TCP 3.3.3.24:49291	3.3.3.24:49294	ESTABLISHED	3364
InHost TCP 3.3.3.24:49293	3.3.3.24:49289	ESTABLISHED	2120
InHost TCP 3.3.3.24:49294 InHost	3.3.3.24:49291	ESTABLISHED	2120

### **Broadcom Advanced Control Suite 4 (BACS4)**

The BACS4 utility provides useful information about each network adapter that is installed in your system, including partitioned adapters. BACS4 enables you to perform detailed tests, diagnostics, and analyses, as well as allows you to view and modify various property values and view traffic statistics for each adapter, including other vendor devices.

BACS4 allows the enabling and configuring of both ports NPAR flow control/protocols/Relative Bandwidth Weights/Maximum Bandwidth settings.

The following figure shows the per partition NPAR settings (see Figure 6). This is where BACS4 can enable or disable NPAR mode. This is also where BACS4 controls the NPAR per port IEEE 802.3x Link-Level Flow Control settings (used when DCB's PFC is disabled), enabled protocols (Ethernet or iSCSI or FCoE), the **Relative Bandwidth Weight** values, and the **Maximum Bandwidth** values per partition.

Broadcom Advanced Control Suite		
le View Action Filter Context Too	Is Teams iSCSI Help	
Filter: ALL VIEW Config	guration 🔽 🔽 NIC Partition	
plorer View	Information Configuration	
- 000 Hosts	Property	Value
🗄 📳 WIN-RA3QJRK0731	- NIC Partition	
🗄 😅 Adapter 1 (BCM5709 C0)	Reset Configuration to Default	
	NIC Partition	Enabled
T: Adapter 3 (BCM57711 AD	Port 0	
Adapter 4 (57712 A 1)	Flow Control	Auto
	E. Function 0	
	Etheret/Ndis	Enable
E [0050] Broadcon	- iSCSI	☐ Disable
😥 [0012] Broa	dc FCoE	Enable
🚺 [0027] Broa	dc Relative Bandwidth Wei	5
[0051] Broadcon	Maximum Bandwidth (%)	10
🖨 📅 [0053] Broadcon	Function 2	
🛐 [0030] Broa	dc Etheret/Ndis	Disable
E [0054] Broadcon	iSCSI	Disable
	FCoE	T Disable
	Relative Bandwidth Wei	5
····· 💽 [0031] Broa	Maximum Bandwidth (%)	10
t± → Port2	E Function 4	
	Etheret/Ndis	Enable
	iSCSI	Finable
	FCoE	Disable
	Relative Bandwidth Wei	45
	Maximum Bandwidth (%)	80
	E. Function 6	
	Etheret/Ndis	Enable
	iSCSI	Disable
	FCoE	Disable
	··· Relative Bandwidth Wei	45
	Maximum Bandwidth (%)	80

Figure 6: BACS4 NPAR Settings

BACS4 displays the per partition Virtual Bus Device (VBD) information.

Filter: ALL VIEW 💽 Information 🔽 🔽 Vital Signs 🔽 Driver Information 🖾 NIC Partition	n		
plorer View 8	Information	Configurations Stati	stics
E- 800 Hosts	Property		Value
🖻 🗻 WIN-RA3QJRK0731	Driver Infor	mation	
🕀 🔝 Adapter 1 (BCM5709 C0)	- Driver	lersion	6.4.39.0
🕀 🔝 Adapter 2 (BCM5709 C0)	- Driver I	Date	4/21/2011
🟵 😅 Adapter3 (BCM57711 A0)	- Driver 1	Name	evbda.sys
🖻 🌌 Adapter 4 (57712 A1)	Driver S	Status	Loaded
B Port1	Vital Signs		
E III [0050] Broadcom BCM57712 NetXtreme II 10 GirE Multi Function #50	- Ndis M/	AC Address	0010186FD3C4
- The foot of Broadcom BCM57712 NetVitreme II 10 GinE Multi Exection ISCSI Adapter	- ISCSI M	AC Address	0010186FD3C5
[002] Brandram BCME7713 NetWeater II 10 Gire AIDIS (Boot) #50	NIC Partito	n	
Cost and a contract of the cost of the cos	- Physica	Network MAC Address	0010186FD3C4
[UUS1] Broadcom BCMS//12 NetAtreme II 10 Gige Multi Function #51	Physica	I ISCSI MAC Address	0010186FD3C5
E [0053] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #53	Relative	e Bandwidth Weight (%)	45
[0030] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #53	Maximu	m Bandwidth (%)	80
E [0054] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #54			
[0016] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function FCoE Adapter			

The following shows the per partition FCoE device information.

File         View         Action         Filter         Context         Tools         Teams         ISCSI         Help           I         Filter:         ALL VIEW         Imformation         Imformation         Vital Signs         Driver Information		
Explorer View 8	Information Configuration St	atistics
E- 000 Hosts	Property	Value
🖻 📗 WIN-RA3QJRK0731	Vital Signs	
🗄 😅 Adapter 1 (BCM5709 C0)	- MAC Address	0010186FD38D
E- Mapter 2 (BCM5709 C0)	MTU	2500
Adapter 3 (BCM57711 A0)	- FCF Selection	First Received
- Adapter 4 (57712 A1)	WWNN	10:00:00:10:18:6f:d3:bd
E Port1	WWPN	20:00:00:10:18:6f:d3:bd
Control Providence DOMETTED NotWithout IT 10 Clar Multi E-motion (ED)	Driver Information	
[iii] [doso] prodocom bcHs7712 Neokarene 11 to Gige Hola Function #50	- Driver Name	bxfcoe.sys
[0051] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #51	- Driver Status	Loaded
E- [] [0053] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #53	- Driver Version	6.4.18.0
[0030] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #53	Driver Date	4/20/2011
[D] [0054] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #54		
😑 🎫 [0016] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function FCoE Adapter		

The following shows the configuration of FcoE device settings.

File View Action Filter Context Tools Teams SCSI Help		_8
Explorer View Ø	Information Configuration Statistic	cs
Image: State Stat	Advanced     Reset to Factory Defaults     Fabric disconnect timeout     Target removal timeout     Symbolic Node Name     VLAN ID for VLAN discovery     VLAN ID	Click here to set to default 30 30 Broadcom FCoE 6.4.18.0 6.4.39.0 16 0

The following shows the per partition FCoE device statistics.

Broadcom Advanced Control Suite 4 File View Action Filter Context Tools Teams ISCSI Help		
Filter: ALL VIEW   Statistics  General  General		
Explorer View Ø	Information Configuration	Statistics
E- 000 Hosts	Property	Value
E WIN-RA3QJRK0731	General	
E Adapter 1 (BCM5709 C0)	- Input requests	1713102088
E- Adapter2 (BCM5209 C0)	Output requests	12332
Advanter 2 (BCM57711 A0)	- Control requests	438
Adapted (CTT12 A1)	- Input megabytes	843970
Adapter4 (S//12 AL)	- Output megabytes	3083
e Porti	-FCoE transmitted frame	5 1714693711
[0050] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #50	- FCoE transmitted bytes	174651619204
<ul> <li>[0051] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #51</li> </ul>	- Transmitted FCP frames	3430043518
E [0053] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #53	- FCoE received frames	3430043895
[0030] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #53	- FCoE received bytes	1158447010868
E- 10054] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #54	- Received FCP frames	3430043518
😑 🚍 [0016] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function FCoE Adapter	- FIP VLAN negotiations p	performed 1
□ 0 1 Port 20:00:00:10:18:6F:D3:8D	<ul> <li>FIP fabric discoveries per</li> </ul>	erformed 1
E- C Taroet0	- FLOGIs performed	1
- O LUNO	<ul> <li>FDISCs performed</li> </ul>	0
- O LUN1	-Packets received with w	rong FC 0
- O LUN2	- Wrong FCoE version co	unt 0
- O LUN3	- Wrong delimiter count	0
- O LUN4	<ul> <li>Missing frame count</li> </ul>	0
- Q LUNS	-Receive sequence times	out count 0
- Q LUN6	- ULP_TOV expiration cou	int 0
LUN7	-REC expiration count	0
[0031] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #54	- ABTS count	20
D Port2	SRR count	0
	- Reset LUN count	0
	Reset target count	0
	<ul> <li>Session recovery count</li> </ul>	0
	FCoE dropped frames	0
	- Dropped sequences	0

The following shows the per partition FCoE device connection information.

Stroadcom Advanced Control Suite 4		The second s
File View Action Filter Context Tools Teams ISCSI Help		
Filter: ALL VIEW		
Explorer View 8	Information	
B- 000 Hosts	Property	Value
E WIN-RA30/RK0731	- Fabric Information	TOUL .
E- Adapter 1 (BCM5709 C0)	- Port Type	NPort
Advance 2 (PCM5200 C0)	- WW Port Name	20:00:00:10:18:65:d3:bd
(a) (a) Adapter 2 (BCM57711 A0)	- WW Node Name	10:00:00:10:18:6f:d3:bd
	- Port FCID	0x5f006e
Er Cataloger 4 (5//12 A1)	- Fabric Priority	0
E- Port1	LinkKeepAliveInterval	8000
E [0050] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #50	FPMA MacAddress	0EFC00SF006E
<ul> <li>[0051] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #51</li> </ul>	- Fabric Name	20:01:00:0d:ec:e3:ae:01
E [0053] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #53	- Fabric VLAN	1001
<ul> <li>[0030] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #53</li> </ul>	-FC-Map	0e:fc:00
E- 10 [0054] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #54	- FCF MacAddress	000DECE3AE00
😑 🚾 [0016] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function FCoE Adapter	- FCF WW Node Name	20:01:00:0d:ec:e3:ae:01
Port 20:00:00:10:18:6F:D3:8D	FCF WW Port Name	20:0c:00:0d:ec:e3:ae:3f

The following shows the per partition per FCoE target information.

Broadcom Advanced Control Suite 4		
Filter: ALL VIEW		
Explorer View 8	Information	
Ė- 000 Hosts	Property	Value
E WIN-RA3QJRK0731	Target Information	
Adapter 1 (BCM5709 C0)	FCID	0x5f0007
E Adapter2 (BCM5709 C0)		10:00:00:00:c9:92:c6:56
Adapter 3 (BCM57711 A0)	- WW Node Name	20:00:00:00:c9:92:c6:56
Adapter 4 (57712 A1)	SCSI Bus Number	0
	- SCSI Target Number	0

The following shows the per port DCB protocol information

Broadcom Advanced Control Suite 4		8	1		
File New Action Filter Context Tools Tools SCSL Filep	PV Istanced				
	bit Advanced	Lee - L		(	
oper vev e	phormation	Configurations	Statetics	Diagnostics	
8 000 Hosts	Property			Value	<u>ف</u>
E- 👔 100+65495(316+P 8- 😅 Adapter 1 (ICM5709 C0)					
				Enable	
8 Adapter 2 (8CH5709 C0)	8 Priorit	y Tagging		Operational	
B: Adapter3 (BCM57712 A1)	- Ne	etworking PRI		0	
Adapter 4 (BCM57712 A1)	-PC	Coll PR1		3	
😑 🤟 Porta	- 6	CSEPRE		4	
B- 10051] Broadcom BCM57712 NetXtreme II 10 GigE #51	8 Priorit	y Flow Control (PFC)		Operational	
R (2005) Broadcom BCM57712 Netthreme II 10 Goll (SCSI Adapter #51	- 14	C Enabled on Priorite	8	3	
E- 10000 Broadrom 8/167212 Net/Drame II 10 Golf 5/ of Advator #51	O. Paker	C Disabled on Phone	S description	0124567	
50178 Broadrow B/14577113 artitizena II 10 Golf ADIS UBD / Last) #51	0.N	ced transmission see	Needland (E15)	Operational	
Image (and a second action action action and a second action		Principles in Principle	inacceng	0124567	
C The state of the	Priorities in Priority Group U		011-507		
E. [8] [0025] prospcou prv/2/17 verviewe II to ride +25		Priorities in Priority (	kan 2	3	
[0007] Broadcon BCMS77127NetXtheme II 10 GigE PCoE Adapter #52	D-Pr	ionity Group Bandwidth	1		
[0008] Broadcom BCH57712 NetXtreme II 10 Ggt SCSI Adapter #52		Priority Group		0.1.2.3.4.5.6.7	
[0028] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #52		Bandwidth %		50. 0. 50. 0. 0. 0. 0. 0	
8 Adapter 5 (IICM57712 A1)	DOBX Adva	anced			A REAL PROPERTY AND A REAL PROPERTY.
B: Adapter6 (BCH57712 A1)	8-Local M	-03			
	- ET	5		Enable	10 million (10 mil
	- Pf	c		Enable	
	- Configuration mis-match		No		
	- Ne	etworking PRI		0	
	- FC	CoE PRI		3	
	-64	CSE PR1		4	
	- PF	C Enabled on Prioritie	5	3	
	- Pf	C Disabled on Prioritie	s	0124567	
	- Ne	etworking PGID		0	
	-PC	Coll PGID		2	
	- 64	CSI PGID		0	
	- PG	3D(0) BW(%)		50	-
	-				
Α					
BROADCOM.					ENT STA

BACS4 allows per port enabling and configuring of the DCB protocol default and initially advertised settings and port "willingness" to change to the received DCBx values as shown below.

Se Broadcom Advanced Control Suite 4				_IO ×
File View Action Filter Context Tools Teams ISCSI Help				
Filter: ALL VIEW  Configurations				
Explorer Vew Ø	Information Configurations State	stics   De	agrestics	
B- 000 Hosts	Property	Value		19 A 19 A 19
8- 1 WINFSFERTER	- DOW			
8: 😅 Adapter 1 (8CM5709 C0)	- 008	F Enabl	ie .	
E Adapter 2 (BCM5709 C0)	- Overwrite default DCBX settings	Yes		
B S Adapter 3 (ICM57712 A1)	-Reset Configuration to Default			Set to default value
E Adapter 4 (BCM57712 A1)	ETS	Enable		
B -+ Port0	- PFC	Enable		
(a) 100513 Broadcom BCM57712 NetXtreme II 10 Gigl: #51	Local machine willing	Yes		
R: UR 100051 Broadcom BCM57712 Net/Streme II 10 Got ISCS1 Adapter #51	PRI/PFC/PGID settings		📚 PR1/PTC/PGID settings 💦 🔀	Edit Settings
R- 100067 Resadown RCM57712 NetWhene II 10 GinF PCoF Adapter #51	- Networking PRI	0	Name follow the state balance is configured	
[0027] Broadcon BCM57712 Net/Strene II 10 Golf (NDIS VBD Clent) #51	- PCoE PR1	3	Heade rokow the steps below to compare.	
R-+ byt	- OCSLPRI	O	Step 1: Assign Priority (PRI)	
E 100501 Resadows RCM57712 NetWrene II 10 GoF #52	- Reverses	Enable	·	
- INDER Broadrow BONS7712 National To 10 Golf BColf Adapter #52	- ACSU PEC	Disable	Networking PR2: 0	
- Control Restations RCMC7713 Matthema T 10 Coll SCVI Advance #53	Networking PGID	1	-	
10038] Broadcon BCH57712 Natiltrane III 10 GoE (NDIS VED Clant) #52	- FCoE PGID	2	PCOE PRU: 3	
E Martine C (ICME2712 & 1)	- 6CSI PGID	1	entration	
E Manager (Constrained)	- PGID(1) 8W(%)	50	bosinai iv I	
the state of the s	- PGID(2) 8W(%)	50	And a second second second	-
	PGID(3) BW(%)	0	Step2: Assign Priority Group (PGID)	
			Step3: Set Friority Flow Control (PFC)	
			Step-I: Assign PGID Bandwidth	
			OK Cancel	
	1			
		Apply	Reset	
			FIE	174
~~~	and the second	_	-10 ac	

The following shows the per port DCB protocol statistics.

File View Action Filter Context Tools Teams iSCSI Help				
Filter: ALL VIEW 💽 Statistics 🔽 🔽 DCBX Statistics 🔽 Custom	and the second			
Explorer View 8	Information Configurations Statis	tics Diagnostics		
E- 000 Hosts	Property	Value		
E WIN-RA3QJRK0731	- DCBX Statistics			
Adapter 1 (BCM5709 C0)	- DCBX Frames Sent	15398		
E Adapter 2 (BCM5709 C0)	- DCBX Frames Received	15381		
Adapter3 (BCM57711 A0)	- PFC Frames Sent	44441544		
- Adapter 4 (57712 A1)	PFC Frames Received	0		
R - Dort1	- Custom			
Control Decoders DOMETTICA Northerne TI 10 Cont Multi Eventice and	- Frames size less than 64 bytes	0		
[0050] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #50	- MAC Rx. Xoff	0		
<ul> <li>[0012] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function ISCSI Adapter</li> </ul>	- MAC Rx. Xon	0		
<ul> <li>[0027] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #50</li> </ul>	MAC Rx. w/ no Pause Command	0		
<ul> <li>[0051] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #51</li> </ul>	- MAC Sent Xon	0		
[] [0053] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #53	- MAC Sent Xoff	0		
[0030] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #53	<ul> <li>Pause frames received</li> </ul>	0		
[0054] Broadcom BCM57712 NetXtreme II 10 GidE Multi Function #54	Pause frames sent	0		
[0016] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function FCoE Adapter [0031] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #54				

The following shows the per partition iSCSI device information.



The per partition iSCSI Offload device's VLAN, MTU and IP address (IPv4 and IPv6) settings are configured in BACS4 as shown below.

Broadcom Advanced Control Suite 4		6 ×
File View Action Filter Context Tools Teams ISCSI Help		
Explorer View 8	Information Configurations Diagnostics Statistics	
E 000 Hosts	Property Value	
E WIN-RA3QJRK0731	ISCSI Management	
Adapter 1 (BCM5709 C0)	- VLAN ID	
Adapter 2 (BCM5709 C0)	MTU 1500	
Adapter 3 (BCM57711 A0)	E- IPv4 Configuration	Edit
- Adapter 4 (57712 A1)	- IPv4DHCP Disable	
E - Port1	- IP Address 1.1.1.4	
C. T [0050] Brandow BOME7712 NetWorks II 10 Cold Mult Exection #50	Subnet Mask 255.255.0.0	
C [] [UUSU] broadcom bCMS7712 NetAtleme II 10 Gige Multi Punction #30	- Default Gateway	
[0012] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function ISCSI Adap	E-IPv6 Configuration	Edit
- eca Portal 1.1.1.4	- IPv6 DHCP Disable	
B Qign. 2010-10.com.broadcom:tgt15a.ram15a	Process Router Advertiseme Disable	

The following shows the per partition iSCSI device traffic statistics.



The following shows the per partition iSCSI device per Portal traffic statistics.

	Information Statistics	
t00 Hosts	Property	Value
Hosts	Property	Value 7372080416256 16911 56244633 112489200 0 0 0 7372215682560 16911 56245655 112491264 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0

The following shows the per partition iSCSI device per Portal per target information.



The following shows the per device/partition's current offload licenses.

Broadcom Advanced Control Suite 4	1 2			
File View Action Filter Context Tools Teams iSCSI Help				
Filter: ALL VIEW  Configurations  Resource Reservations  Licenses  A	dvanced			
Explorer View 8	Informatio	on Configurations	Statist	ics
⊡- 000 Hosts	Property			Value
🖻 📳 WIN-RA3QJRK0731	Advan	ced		
🕀 🔝 Adapter 1 (BCM5709 C0)	<ul> <li>Resource</li> </ul>	rce Reservations		
E- Adapter2 (BCM5709 C0)	= License	es		
Adapter3 (BCM57711 A0)	- TC	CP Offload Engine (TO	E)	Maximum Connections
- Adapter 4 (57712 A1)	-is(	CSI Offload Engine		Maximum Connections
B-→- Port1	-FC	CoE Offload Engine		Maximum Connections
Financial (10045) Broadcom BCMS7712 NetVitreme II 10 GinE #45				
FOR [0007] Broadcon BCM57712 NetWrame II 10 GirE (DDIS VRD Clast) #45				
[0007] brodocom bcH37712 Neokalement to dige (ND13 YOD Client) #45				
er Portz				
E [0046] Broadcom BCM57712 NetXtreme II 10 GigE #46				
[0012] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #46				

An iSCSI Discovery and Management Wizard is also included with BACS4. This can be used to connect and manage the iSCSI hardware offload enabled initiator devices to various iSCSI targets as shown below.

Broadcom Advanced Control Suite 4	
File View Action Filter Context Tools Teams	iSCSI Help
Filter: ALL VIEW  Information  Explorer View	Discovery Wizard Manage Targets Wizard
E 000 Hosts	Manage iSNS Servers
🗄 🗍 WIN-RA3QJRK0731	Manage Discovery Portals

#### **Microsoft Windows Network Connections**

These devices can be used by any application, as if they were a separate adapter port. They appear as separate Ethernet devices in **Network Connections** from the Windows Control Panel. The following shows eight USC-enabled Ethernet Protocol partitions as eight separate Ethernet network connections and these are arranged in port order (0 and 1) and partition order (1 through 4).



Each of these network connection devices can be accessed, individually, as if they where separate adapters. The connection status shows the USC **Maximum Bandwidth** setting as the **Speed** of the connection.

1-57712-NP0-Port 0 Partition 1-1.1.1.20 Status	🖞 🖞 2-57712-NP2-Port 0 Partition 2-3.3.3.20 Status 🔀
General	General
Connection	Connection
IPv4 Connectivity: No Internet access	IPv4 Connectivity: No Internet access
IPv6 Connectivity: No network access	IPv6 Connectivity: No network access
Media State: Enabled	Media State: Enabled
Duration: 17:02:12	Duration: 17:02:10
Speed: 400.0 Mbps	Speed: 1.7 Gbps
Details	Details
Activity	Activity
Sent — Received	Sent — Received
Bytes: 91,756,308,698 7,455,871,680,584	Bytes: 55,622,032,520   7,456,531,486,823
Properties Diagnose Diagnose	Properties Stable Diagnose
3-57712-NP4-Port 0 Partition 3-5.5.5.20 Status	🖞 4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status
3-57712-NP4-Port 0 Partition 3-5.5.5.20 Status     General	4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status
3-57712-NP4-Port 0 Partition 3-5.5.5.20 Status      General      Connection	4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status      General      Connection
General Connection IPv4 Connectivity: No Internet access	General Connection IPv4 Connectivity: No Internet access
General Connection IPv4 Connectivity: No Internet access IPv6 Connectivity: No network access	Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Provide the second system       Image: Provide the second system         Image: Provide the second system       Image: Prov
General Connection IPv4 Connectivity: No Internet access IPv6 Connectivity: No network access Media State: Enabled	Image: Provide the second state       Image: Provide the second state         Image: Provide the second state       Image: Provide the second state         Image: Provide the second state       Image: Provide the second state         Image: Provide the second state       Image: Provide the second state         Image: Provide the second state       Image: Provide the second state         Image: Provide the second state       Image: Provide the second state
General Connection IPv4 Connectivity: No Internet access IPv6 Connectivity: No network access Media State: Enabled Duration: 17:02:07	Image: Provide the second state in
General Connection IPv4 Connectivity: No Internet access IPv6 Connectivity: No network access Media State: Enabled Duration: 17:02:07 Speed: 3.3 Gbps	4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status       ×         General          Connection          IPv4 Connectivity:       No Internet access         IPv6 Connectivity:       No network access         Media State:       Enabled         Duration:       17:02:08         Speed:       4.6 Gbps
General Connection IPv4 Connectivity: No Internet access IPv6 Connectivity: No network access Media State: Enabled Duration: 17:02:07 Speed: Details	4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status       ×         General
General Connection IPv4 Connectivity: No Internet access IPv6 Connectivity: No network access Media State: Enabled Duration: 17:02:07 Speed: Details Activity	4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status       ×         General          Connection          IPv4 Connectivity:       No Internet access         IPv6 Connectivity:       No network access         Media State:       Enabled         Duration:       17:02:08         Speed:       4.6 Gbps         Details
General Connection IPv4 Connectivity: No Internet access IPv6 Connectivity: No network access Media State: Enabled Duration: 17:02:07 Speed: Details Activity Sent — Received	4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status       ×         General
General General Connection IPv4 Connectivity: No Internet access IPv6 Connectivity: No network access Media State: Enabled Duration: 17:02:07 Speed: Details Activity Sent — Received Bytes: 132,542,805,914 3,920,501,721,494	4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status       ×         General       Connection         IPv4 Connectivity:       No Internet access         IPv6 Connectivity:       No network access         Media State:       Enabled         Duration:       17:02:08         Speed:       4.6 Gbps         Details       Sent — Received         Bytes:       135,167,179,352       3,868,136,383,260

The previous Link Speeds are the result of the following USC **Maximum Bandwidth** settings, and show its 100 Mbps (1%) configurable granularity.

ØLL	UNIFIED SERVER CONFIGURATOR	LIFECYCLE CONTROLLER ENA
Broadcom Ne	etXtreme II 10 Gigabit Ethernet	- 00:10:18:6F:D2:A4
Global Band	dwidth Allocation Menu	
Partition	1 Relative Bandwidth Weight	
0		
Partition :	1 Maximum Bandwidth	
4		
Partition :	2 Relative Bandwidth Weight	
0		
Partition :	2 Maximum Bandwidth	
17		
Partition 3	3 Relative Bandwidth Weight	
0		
Partition 3	3 Maximum Bandwidth	
33		
Partition 4	4 Relative Bandwidth Weight	
0		
Partition 4	4 Maximum Bandwidth	
46		

#### **Device PCIe Bus Location**

The PCIe interface Location, Bus, and Device position numbers are the same for both ports and all eight of the partitions on those ports. The only PCIe interface location values that are different are the **Function** numbers. In non-partitioned Single Function (SF) mode, you would only have functions 0 and 1. In partitioned (NPAR) mode, you have functions 0 through 7, with functions 0-2-4-6 existing on the first port and functions 1-3-5-7 existing on the second port. The actual numbering position an adapter is assigned by Windows is not entirely related to the PCIe interface numbering and is more related to what open location position numbers are available in the registry when the adapters get enumerated. Therefore, port 0 partition 1 may not always occupy the first position in the Windows Device Manager's Network Adapters or Storage Controllers or System Devices sections.

Dell Unified Server Configurator	PCIe Location / Bus / Device / Function The first three are the same for both ports	Ethernet MAC Address	Example IP Address
Port 0 Partition 1	x/y/z/ <b>0</b>	00:10:18:88:E7:A8	1.1.1.1
Port 0 Partition 2	x/y/z/ <b>2</b>	00:10:18:88:E7:AC	2.2.2.1
Port 0 Partition 3	x/y/z/ <b>4</b>	00:10:18:88:E7:B0	3.3.3.1
Port 0 Partition 4	x/y/z/6	00:10:18:88:E7:B4	4.4.4.1
Port 1 Partition 1	x/y/z/ <b>1</b>	00:10:18:88:E7:AA	5.5.5.1
Port 1 Partition 2	x/y/z/ <b>3</b>	00:10:18:88:E7:AE	6.6.6.1
Port 1 Partition 3	x/y/z/ <b>5</b>	00:10:18:88:E7:B2	7.7.7.1
Port 1 Partition 4	x/y/z/ <b>7</b>	00:10:18:88:E7:B6	8.8.8.1

The partition's MAC addresses interleave the two ports, as do the function numbers (see Table 2).

Port 0, Partition 1	Function 0 = MAC address:A8
Port 0, Partition 2	Function 2 = MAC address:AC
Port 0, Partition 3	Function 4 = MAC address:B0
Port 0, Partition 4	Function 6 = MAC address:B4
Port 1, Partition 1	Function 1 = MAC address:AA
Port 1, Partition 2	Function 3 = MAC address:AE
Port 1, Partition 3	Function 5 = MAC address:B2
Port 1, Partition 4	Function 7 = MAC address:B6

#### Table 2: Port, Function, MAC Address Example

One way to locate PCIe information is to open the individual Network Connection's Properties.



In the Properties window, select the device **Configure** button.

5771	12-NP-Port 0 Partition 1-1.1.1.10 Properties	>
Netwo	orking Sharing	
Conn	nect using:	
9	Broadcom BCM57712-NP NetXtreme II 10 GigE (NDIS VB	
	Configure.	7
	Contigues.	

In the NDIS client device properties, you will find that connection's PCIe bus, device, function location information. For partitioned adapters, locate the function number that provides the partition that this connection is connected. The same can be done with Device Manager, especially for iSCSI Storage devices. All of the enabled devices on the same partition have identical PCIe interface location information, with only the function number varying. The following shows the eight Ethernet-partitioned adapter's PCIe device location information.

Notice the Windows enumerated device numbering (#) values do not follow the PCIe bus function numbering, nor the port to partition numbering.
	2-57712-NP2-Port 0 Partition 2-3.3.3.20 Status
Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #76	Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #78
General Advanced Driver Details Power Management	General Advanced Driver Details Power Management
Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #76	Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #78
Device type: Network adapters	Device type: Network adapters
Manufacturer: Broadcom Corporation	Manufacturer: Broadcom Corporation
Location: Location 4 (PCI bus 6, device 0, function 0)	Location: Location 4 (PCI bus 6, device 0, function 2)
Device status	Device status
This device is working property.	This device is working property.
3-57712-NP4-Port 0 Partition 3-5.5.5.20 Status	🛊 4-57712-NP6-Port 0 Partition 4-7.7.7.20 Status 🔀
Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #80	Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #79
General Advanced Driver Details Power Management	General Advanced Driver Details Power Management
Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #80	Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #79
Device type: Network adapters	Device type: Network adapters
Manufacturer: Broadcom Corporation	Manufacturer: Broadcom Corporation
Location: Location 4 (PCI bus 6, device 0, function 4)	Location: Location 4 (PCI bus 6, device 0, function 6)
Device status	Device status
This device is working property.	This device is working properly.
5-57712-NP1-Port 1 Partition 1-2.2.2.20 Status	6-57712-NP3-Port 1 Partition 2-4.4.4.20 Status
Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75	Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74 🗙
General Advanced Driver Details Power Management	General Advanced Driver Details Power Management
General Advanced Driver Details Power Management Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75	General Advanced Driver Details Power Management Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74
General Advanced Driver Details Power Management Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75 Device type: Network adapters	General Advanced Driver Details Power Management Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74 Device type: Network adapters
General       Advanced       Driver       Details       Power Management         Image: Stress of the str	General Advanced Driver Details Power Management Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74 Device type: Network adapters Manufacturer: Broadcom Corporation
General       Advanced       Driver       Details       Power Management         Image: Strategy of the strategy	General       Advanced       Driver       Details       Power Management         Image: Strategy of the strategy
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)	General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       This device is working property.	General       Advanced       Driver       Details       Power Management         Image: Strategy St
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       Image: Tris device is working property.         Image: T-S7712-NPS-Port 1 Partition 3-6.6.6.20 Status       X	General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74       Device type:       NetWork adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status       This device is working property.         Image: Sector 1 Port 1 Partition 4-8.8.8.20 Status       X
General       Advanced       Driver       Details       Power Management         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS         VBD Client) #75         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       This device is working property.         This device is working property.	General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74       Device type:       NetWork adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status       This device is working property.         Image: R-57712-NP7-Port 1 Partition 4-8.8.8.20 Status       X         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #73       X
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       This device is working property.         Image: T-S7712-NPS-Port 1 Partition 3-6.6.6.20 Status       Image: Status         Broadcom BCH57712 NetXtreme II 10 GigE (NDIS VBD Client) #77         General       Advanced       Driver         Details       Power Management       Advanced	General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status       This device is working property.         Image: B-57712-NP7-Port 1 Partition 4-8.8.8.20 Status       X         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #73 X         General       Advanced       Driver
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       Image: Constraint of the status         This device is working property.       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constraint of the status         Image: Constraint of the status       Image: Constatus         Image: Constraint of the	General       Advanced       Driver       Details       Power Management         Image: Strate Control       Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status       This device is working property.         Image: Status       Image: Status
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       Image: This device is working property.         Image: T-57712-NP5-Port 1 Partition 3-6.6.6.20 Status       Image: This device 1 Partition 3-6.6.6.20 Status         Broadcom BCH57712 NetXtreme II 10 GigE (NDIS VBD Client) #77       General         Advanced       Driver       Details         Power Management       Image: This device type:       NetWork adapters	General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status       This device is working property.         Image: B-57712-NP7-Port 1 Partition 4-8.8.8.20 Status       X         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #73 X         General       Advanced         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #73         Device type:       Network adapters
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       Image: Advanced       Image: Advanced         This device is working property.       Image: Advanced       Image: Advanced         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #77       General         General       Advanced       Driver       Details         Power Management       Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #77         Device type:       Network adapters         Manufacturer:       Broadcom Corporation	General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status       This device is working property.         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #73 ×         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #73 ×         General       Advanced         Driver       Details         Power Management       Manufacturer:         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #73 ×         Device type:       Network adapters         Manufacturer:       Broadcom Corporation
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       Image: Status         This device is working properly.       Image: Status         Image: Status       Image: Status         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #77         General       Advanced         Driver       Details         Power Management       Image: Status         Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #77         General       Advanced         Driver       Details         Power Management       Image: Status         Image: Status       Image: Status	General       Advanced       Driver       Details       Power Management         Image: Status       Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status       This device is working property.         Image: Status       Image: Status         Image: Status       Image: Status         Image: Status       Status         Image: Status       Image: Status         Image
General       Advanced       Driver       Details       Power Management         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #75       Device type:       Network adapters         Device type:       Network adapters       Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 1)         Device status       Image: Constraint of the status         This device is working property.       Image: Constraint of the status         Prosadcom BCH57712 NetXtreme II 10 GigE (NDIS VBD Client) #77         General       Advanced         Driver       Details         Power Management       Image: Constraint of the status         Image: Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #77         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 5)	General       Advanced       Driver       Details       Power Management         Image: Strate Strate       Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status
General       Advanced       Driver       Details       Power Management         Image: Strategy St	General       Advanced       Driver       Details       Power Management         Image: Status       Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #74         Device type:       Network adapters         Manufacturer:       Broadcom Corporation         Location:       Location 4 (PCI bus 6, device 0, function 3)         Device status       Image: Status         This device is working property.       ▲         Image: Status       Image: Status         Second BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #73 X         General       Advanced         Driver       Details         Power Management         Image: Status       Image: Status         Image: Status       Image: Sta

## **Red Hat Enterprise Linux**

Linux shows the respective device protocols that were enabled in USC if the NetXtreme II device drivers are installed. Go to the Dell Driver Download web site for the latest Linux device drivers and insure they are installed on your system. In Linux, the Ethernet Protocol is always enabled on all eight partitions, and iSCSI Offload (HBA) Protocol was enabled on the first two partitions of each port.

The following shows the RHEL Network Configuration page with the eight enabled Ethernet protocol partitions (always four per port) as the Broadcom 57712-k NetXtreme II 10 GigE Ethernet devices Eth2 through Eth9 (in this example). Linux enumerates the partitions in order of the PCI function numbers, which is slightly different from Windows where port 0 has functions/partitions 0/2/4/6 (which are eth 2/4/6/8) and port 1 has functions/ partitions 1/3/5/7 (which are eth 3/5/7/9).

ma	Ð			Ne	etwork Configuration
yste	<u>F</u> ile <u>P</u> rofi	ile <u>H</u> elp			and the state of the second
laces S	New	Edit Co	ppy Dele	te Activat	e Deactivate
ч.	Dev <u>i</u> ces	Hard <u>w</u> are	IPsec DM	S Hosts	
pplication		You may co physical ha be associat	onfigure ne rdware he ed with a	etwork devices ere. Multiple log single piece of	associated with gical devices can hardware.
A	Profile S	tatus	Device	Nickname	Туре
	V \$	🗴 Active	🗃 eth1	ethl	Ethernet
0		X Active	eth0	eth0	Ethernet
1000		Active	eth2	1_57712_NP0	Ethernet
		Active	eth3	5_57712_NP1	Ethernet
	. ⊻ \$	Active	eth4	2_57712_NP2	Ethernet
	<b>☑</b> 🐒	Active 🗸	📑 eth5	6_57712_NP3	Ethernet
	. ☑ 🐒	Active	📑 eth6	3_57712_NP4	Ethernet
	☑ 🐒	Active	🗃 eth7	7_57712_NP5	Ethernet
		Active	🗃 eth8	4_57712_NP6	Ethernet
	☑ ૹ	Active	📑 eth9	8_57712_NP7	Ethernet

Linux's ifconfig command shows the partition's eight Ethernet protocol devices and various statistics.

eth2	Link encap:Ethernet HWaddr 00:10:18:88:E7:A8 inet addr:10.1.1.200 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7a8/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:9600 Metric:1 RX packets:27937 errors:0 dropped:0 overruns:0 frame:0 TX packets:40 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:5030387 (4.7 MiB) TX bytes:9923 (9.6 KiB) Interrupt:169 Memory:d1800000-d1fffff
eth3	Link encap:Ethernet HWaddr 00:10:18:88:E7:AA inet addr:10.1.1.201 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7aa/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:9600 Metric:1 RX packets:27500 errors:0 dropped:0 overruns:0 frame:0 TX packets:36 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:4946656 (4.7 MiB) TX bytes:9456 (9.2 KiB) Interrupt:225 Memory:d2800000-d2ffffff
eth4	Link encap:Ethernet HWaddr 00:10:18:88:E7:AC inet addr:10.1.1.202 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7ac/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:9600 Metric:1 RX packets:27034 errors:0 dropped:0 overruns:0 frame:0 TX packets:42 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:4860331 (4.6 MiB) TX bytes:10027 (9.7 KiB) Interrupt:225 Memory:d3800000-d3ffffff
th5	Link encap:Ethernet HWaddr 00:10:18:88:E7:AE inet addr:10.1.1.203 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7ae/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:9600 Metric:1 RX packets:26782 errors:0 dropped:0 overruns:0 frame:0 TX packets:34 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:4807217 (4.5 MiB) TX bytes:8866 (8.6 KiB) Interrupt:204 Memory:d4800000-d4ffffff
th6	Link encap:Ethernet HWaddr 00:10:18:88:E7:B0 inet addr:10.2.2.200 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7b0/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:9600 Metric:1 RX packets:2682 errors:0 dropped:0 overruns:0 frame:0 TX packets:34 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:472737 (461.6 KiB) TX bytes:9139 (8.9 KiB) Interrupt:204 Memory:d5800000-d5ffffff
th7	Link encap:Ethernet HWaddr 00:10:18:88:E7:B2 inet addr:10.2.2.201 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7b2/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:9600 Metric:1 RX packets:2233 errors:0 dropped:0 overruns:0 frame:0 TX packets:32 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:393371 (384.1 KiB) TX bytes:8381 (8.1 KiB) Interrupt:181 Memory:d6800000-d6ffffff

eth8 Link encap:Ethernet HWaddr 00:10:18:88:E7:B4 inet addr:10.2.2.202 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7b4/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:9600 Metric:1 RX packets:1701 errors:0 dropped:0 overruns:0 frame:0 TX packets:31 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:290494 (283.6 KiB) TX bytes:8110 (7.9 KiB) Interrupt:181 Memory:d7800000-d7ffffff eth9 Link encap:Ethernet HWaddr 00:10:18:88:E7:B6 inet addr:10.2.2.203 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7b6/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:9600 Metric:1 RX packets:1730 errors:0 dropped:0 overruns:0 frame:0 TX packets:33 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:299886 (292.8 KiB) TX bytes:8450 (8.2 KiB) Interrupt:169 Memory:d8800000-d8ffffff

Check the Linux system message logs to see if the iSCSI HBAs on the first two partitions of each port (Port 0, Partition 1 = eth2; Port 0, Partition 2 = eth4; Port 1, Partition 1 = eth3; and Port 1 Partition 2 = eth5) are available.



For iSCSI, also check the iscsi\_host folder to see if your iSCSI devices are present.

[root@localhost]# cd /sys/class/iscsi\_host/

[root@localhost iscsi\_host]# pwd

/sys/class/iscsi\_host

[root@localhost iscsi\_host]# 11

total 0

lrwxrwxrwx 1 root root 0 Jun 29 11:15 host3 -> ../../devices/pci0000:00/0000:00:09.0/0000:07:00.1/ host3/iscsi\_host/host3

lrwxrwxrwx 1 root root 0 Jun 29 11:15 host4 -> ../../devices/pci0000:00/0000:00:09.0/0000:07:00.0/ host4/iscsi\_host/host4

Each installed iSCSI device will appear here. The iSCSI devices uses the **bnx2i** driver which can be checked for with the lsmod | grep bnx2 command.

Additionally, check the **fc\_host** folder to see if your FCoE devices are present.

[root@localhost]# cd /sys/class/fc\_host/

[root@localhost fc\_host]# pwd

/sys/class/fc\_host

[root@localhost fc\_host]# 11

total 0

lrwxrwxrwx 1 root root 0 Jun 29 11:11 host8 -> ../../devices/pci0000:00/0000:00:07.0/0000:05:00.0/ host5/fc\_host/host5

lrwxrwxrwx 1 root root 0 Jun 29 11:11 host8 -> ../../devices/pci0000:00/0000:00:07.0/0000:06:00.0/ host6/fc\_host/host6

Each installed FCoE device will appear here. FCoE uses the **bnx2fc** driver which can be checked for with the lsmod | grep bnx2 command.

Another useful command is sg\_map -i -x which will show all SCSI LUN devices visible to the host.

For Fiber Channel, another useful application is FCInfo that is part of the Broadcom Linux driver release utilities and displays the FCoE HBA port information.

## VMWare ESX/ESXi 4.1

VMWare ESX/ESXi 4.1 shows the respective device protocols that were enabled in USC if the NetXtreme II device drivers are installed. Go to the Dell Driver Download web site (<u>http://support/del.com</u>) for the latest device Firmware - if not already installed. Go to the VMWare web site (<u>http://downloads.vmware.com</u>) for the latest device drivers and insure they are installed on your system. In VMWare, the Ethernet Protocol is always enabled on all eight partitions. VMWare ESX/ESXi 4.1 does not support the iSCSI Offload Protocol in NPAR mode. VMWare ESX/ESXi 4.1 does not support the FCoE Offload Protocol in SF or NPAR modes.



**Note:** VMWare ESX/ESXi 4.1 only supports four 10 GbE ports. Using NPAR mode allows you to expand the number of ports usable from 4 physical ports to 16 virtual ports. This allows better port flexibility, traffic isolation, service quality, and bandwidth tuning for your management/backup/migration/ production networks.

The following shows the vSphere Network Adapters Configuration page with the eight enabled Ethernet protocol partitions (always four per port) as the Broadcom Corporation NetXtreme II BCM57712 10 Gigabit Ethernet MultiFunction devices vmnic6 through vmnic13 (in this example). VMWare enumerates the partitions in the order of the PCI function numbers where port 0 has functions 0/2/4/6 (which are vmnics 6/8/10/12) and port 1 has functions 1/3/5/7 (which are vmnics 7/9/11/13).

Hardware	Netw	ork Adap	ters			
Upplik Chatra	Devi	ce	MAC Address	Speed	Configured	
Processors	Broadcom Corporation Broadcom 577125					
Memory		vmnic9	00:10:18:6f:d3:be	2000 Full	Negotiate	
Storage		vmnic8	00:10:18:6f:d3:bc	10000 Full	10000 Full	
Networking		vmnic7	00:10:18:6f:d3:ba	1000 Full	1000 Full	
Networking		vmnic6	00:10:18:6f:d3:b8	10000 Full	10000 Full	
Storage Adapters	<b>E9</b>	vmnic13	00:10:18:6f:d3:c6	4000 Full	Negotiate	
<ul> <li>Network Adapters</li> </ul>		vmnic12	00:10:18:6f:d3:c4	10000 Full	10000 Full	
Advanced Settings		vmnic11	00:10:18:6f:d3:c2	3000 Full	Negotiate	
Power Management		vmnic10	00:10:18:6f:d3:c0	10000 Full	10000 Full	

These Ethernet Protocol adapters in each partition are configurable like a normal port adapter in vSphere's Networking Configuration pages as shown below.

ting Started Summary Virtual Mac	hines Resource Allocation Performance	Configuration Local Users & Grou
rdware	View: Virtual Switch	
Health Status	Networking	
Processors		
Memory	Michael Cuitada - Cuitada	Remove Properties
Storage	Virtual Switch: vSwitch0	Renoven Propercesti
Networking	Virtual Machine Port Group	Physical Adapters
Storage Adapters	Villeral Dec	
Network Adapters	Management Network	
Advanced Settings	vmk0:1.10.41.2	-11
Power Management		
-ower Management		
ware	Virtual Switch: vSwitch1	Remove Properties
Linear d Frank and	Virtual Machine Port Group	Physical Adapters
Licensed Features	VMK_NIC_1	💁 🔶 🕳 📰 vmnic6 10000 Full 🖓
Time Configuration	1 virtual machine(s)	
DNS and Routing	VM41_P1_1	₿+
Authentication Services	-VMkernel Port	
/irtual Machine Startup/Shutdown	VMK_I_1	2+-
Virtual Machine Swapfile Location	vmk1:1.1.41.1	
Security Profile		
System Resource Allocation	Virtual Switch: vSwitch2	Remove Properties
Advanced Settings	- Virtual Machine Port Group	Physical Adapters
	VMK_NIC_2	• • • • • • • • • • • • • • • • • • •
	1 virtual machine(s)	
	VM41_P2_1	B+
	- VMkernel Port	
	VMK_I_2	2.4
	vmk2:2.2.41.1	-
		Remove Properties
	Virtual Switch: vSwitch3	Renove Properdes
	- Virtual Machine Port Group	Physical Adapters
	VMK_NIC_3	
	I virtual machine(s)	
	10144 01 0	

VMWare's vSphere 4.1 (or vCenter) can be used to view a selected VM's Networking performance in the specific Host's Performance sub-tab selecting. In the first example, the first port's four partitions (on VMNIC6/ 8/10/12) are set to 0% Relative Bandwidth Weight and 100% Maximum Bandwidth each and the second port's partitions (on VMNIC7/9/11/13) are similarly set to 0% Relative Bandwidth Weight but the Maximum Bandwidth values are set to 10%/20%/30%/40% respectively which results in VMNIC7's link speed (for the transmit direction only) indicating 1000 Mbps, VMNIC9's link speed indicating 2000 Mbps, VMNIC11's link speed indicating 3000 Mbps and finally VMNIC13's link speed indicating 4000 Mbps. This is indicated in the vSphere Host's Configuration - Network Adapter page.

The first port's network performance indicated in the vSphere Host's Performance page shows each individual VM's send traffic rate, when none to all four are sending, sharing the available bandwidth between each other.



The second port's network performance indicates each VM is limited to it's specific top end setting (Maximum Bandwidth) and does not expand into the unused area.



## **Setting MTU Sizes**

- Setting MTU Sizes in Windows
- Setting MTU Sizes in Linux
- Setting MTU Sizes in VMWare ESX/ESXi 4.1



**Note:** In all cases, the connecting switch port that a 57712-k NPAR port is connected to must have the switch's MTU size set to the largest MTU size of those four partitions of the port if the user wants to support all four partitions MTU size settings. Additionally, the remaining network that the traffic flows though must also support the desired MTU sizes for that sized frames to be used without being dropped/truncated/fragmented by the network.

## **Setting MTU Sizes in Windows**

The MTU size for each individual Ethernet protocol-enabled partition can be independently set from Normal (1500 bytes) up to Jumbo (9600 bytes) in several places in Windows.

One place to set the Ethernet protocol-enabled partition's adapter MTU size is in the Window's Networking Adapter - Advanced Properties - Jumbo Packet properties.



Another place to set the MTU size is in the BACS4 NDIS device Configurations page.

Broadcom Advanced Control Suite 4 Ele View Action Filter Context Tools Teams SCSI Help			_0,
Filter: ALL VIEW  Configuration  Advanced			
Explorer View Ø	Information	Configuration Diagr	ostics Statistics
⊡- 000 Hosts	Property		Value
B- 🚺 WIN-RA3QJRK0731	- Advanced		
Adapter 1 (BCM5709 C0)	- Set pro	perties to default Values	Set to default value
Adapter2 (BCM5709 C0)	Intern	pt Moderation	Enabled
Adapter 3 (BCM57711 A0)	- IPv4C	hecksum Officad	Tx/Rx enabled
E- C Adapter 4 (57712 A1)	-IPv4Li	arge Send Offioad	Enable
B Derti	- IPv6 C	hecksum Offload	Tx/Rx enabled
COST Providence POWE7712 NetWorks TI 10 Cost Multi Evention #50	-IPv6L	arge Send Offioad	Enable
Er [[ [0050] Broadcom BCM57712 Neckberne II 10 Gige Mola Function #50	Jumbo	Packet	1500 ÷
E- 9 [0012] Broadcom BCM57712 NetXtreme II 10 GgE Multi Function ISCSI Ar	-Locally	Administered Address	Not present
- 608 Portal 1.1.1.4	- Numbe	r of RSS Queues	Auto
<ul> <li>— BO3 Portal fe80::210:18ff:fe6f:d3c5</li> </ul>	- Pause	On Exhausted Host Ring	Disable
[0027] Broadcom BCM57712 NetXtreme II 10 GigE (NDIS VBD Client) #50	- Priority	& VLAN	Priority & VLAN enabled
[0051] Broadcom BCM57712 NetXtreme II 10 GigE Multi Function #51	Receiv	e Buffers (0=Auto)	0
E [0053] Broadcom BCM57712 NetXtreme II 10 GoE Multi Function #53	Receiv	e Side Scaling	Enable
[0030] Broadcom BCM57712 NetVitreme II 10 CivE AUTIC USD Classi) #53	-TCP Co	onnection Offload (IPv4)	Enable
	- TCP Co	nnection Offload (IPv6)	Enable
E E E E E E E E E E E E E E E E E	Transm	it Buffers (0=Auto)	0

The MTU size for each individual iSCSI Offload (HBA) protocol-enabled partition can be independently set from Normal (1500 bytes) up to Jumbo (9600 bytes) in the BACS4 iSCSI Management Configuration page.



The FCoE device MTU frame size is fixed at 2500 bytes and is not adjustable but is viewable in the FCoE device Information page.



In Windows, each individual partition's Ethernet and iSCSI protocol-enabled adapter MTU size setting can be different. For example:

- Port 0, partition 1 Ethernet can be set to 9600 bytes.
- Port 0, partition 1 iSCSI Offload HBA can be set to 2500 bytes.
- Port 0, partition 2 **Ethernet** can be set to 1500 bytes.
- Port 0, partition 2 iSCSI Offload HBA can be set to 9600 bytes.
- Port 0, partition 3 Ethernet can be set to 3500 bytes.
- Port 0, partition 4 Ethernet can be set to 5500 bytes.

In Windows, use the ping command with the "-f" option to set the Don't Fragment (DF) flag AND the "-l size" option (small I) to verify that Jumbo Frame support is configured throughout the desired network path - i.e. "ping -f -l 8972 A.B.C.D". The unfragmentable ping packet size is the desired MTU size to be checked (9000 bytes) minus the automatically added overhead (28 bytes) or 8972 bytes.

C:\> ping -f -1 8972 192.168.20.10
Pinging 192.168.20.10 from 192.168.20.50 with 8972 bytes of data:
Reply from 192.168.20.10: bytes=8972 time<1ms TTL=64
Reply from 192.168.20.10: bytes=8972 time<1ms TTL=64
Reply from 192.168.20.10: bytes=8972 time<1ms TTL=64
Ping statistics for 192.168.20.10:
Packets: Sent = 4, Received = 4, Lost = 0 (0% loss),
Approximate round trip times in milli-seconds:
Minimum = 0ms, Maximum = 0ms, Average = 0ms</pre>

If it does not work, you might see the following reply (if there is connectivity – try 1472 byte standard frames to see if the non-jumbo frame size is passing through):

C:\> ping -f -l 8972 192.168.20.10

Pinging 192.168.20.10 from 192.168.20.50 with 8972 bytes of data:

Packet needs to be fragmented but DF set.

Ping statistics for 192.168.20.10:

Packets: Sent = 4, Received = 0, Lost = 4 (100% loss)

## **Setting MTU Sizes in Linux**

In Linux, the MTU size for each individual Ethernet protocol-enabled partition can be independently set from Normal (1500 bytes) up to Jumbo (9600 bytes).

Both the Ethernet protocol and iSCSI Offload (HBA) enabled partition's adapter MTU size is adjusted at the same time using the ifconfig command.

ifconfig eth3 mtu NNNN up

From above, eth3 is the port identification of the specific 57712-k partition to adjust the MTU size. The NNNN is the new size of the MTU for that partition, and can be set from 1500 to 9600 bytes.

The following shows all eight partitions being set to different MTU values.

[root@localhost	~]#					
[root@localhost	~]#	ifconfig	eth2	mtu	2200	up
[root@localhost	~]#	ifconfig	eth3	mtu	3300	up
[root@localhost	~]#	ifconfig	eth4	mtu	4400	up
[root@localhost	~]#	ifconfig	eth5	mtu	5500	up
[root@localhost	~]#	ifconfig	eth6	mtu	6600	up
[root@localhost	~]#	ifconfig	eth7	mtu	7700	up
[root@localhost	~]#	ifconfig	eth8	mtu	8800	up
[root@localhost	~]#	ifconfig	eth9	mtu	9600	up
[root@localhost	~]#					

eth2	Link encap:Ethernet HWaddr 00:10:18:88:E7:A8
	inet addr:10.1.1.200 Bcast:10.255.255.255 Mask:255.0.0.0
	inet6 addr: fe80::210:18ff:fe88:e7a8/64_Scope:Link
	UP BROADCAST RUNNING MULTICAST MTU:2200 Metric:1
	RX packets:251 errors:0 dropped:0 overruns:0 frame:0
	TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
	collisions:0 txqueuelen:1000
	RX bytes:47008 (45.9 KiB) TX bytes:0 (0.0 b)
	Interrupt:169 Memory:d1800000-d1ffffff
eth3	Link encap:Ethernet HWaddr 00:10:18:88:E7:AA
	inet addr:10.1.1.201 Bcast:10.255.255.255 Mask:255.0.0.0
	inet6 addr: fe80::210:18ff:fe88:e7aa/64 Scope:Link
	UP BROADCAST RUNNING MULTICAST MTU:3300 Metric:1
	RX packets:418 errors:0 dropped:0 overruns:0 frame:0
	TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
	collisions:0 txqueuelen:1000
	RX bytes:80164 (78.2 KiB) TX bytes:0 (0.0 b)
	Interrupt:225 Memory:d2800000-d2ffffff

eth4 Link encap:Ethernet HWaddr 00:10:18:88:E7:AC inet addr:10.1.1.202 Bcast:10.255.255.255 Mask:255.0.0.0 inet6 addr: fe80::210:18ff:fe88:e7ac/64 Scope:Link UP BROADCAST RUNNING MULTICAST MTU:4400 Metric:1 RX packets:376 errors:0 dropped:0 overruns:0 frame:0 TX packets:0 errors:0 dropped:0 overruns:0 carrier:0 collisions:0 txqueuelen:1000 RX bytes:73184 (71.4 KiB) TX bytes:0 (0.0 b) Interrupt:225 Memory:d3800000-d3ffffff

eth5	Link encap:Ethernet HWaddr 00:10:18:88:E7:AE
	inet6 addr: fe88::218:18ff:fe88:e7ae/64 Scope:Link
	UP BROADCAST RUNNING MULTICAST MTU:5500 Metric:1
	RX packets:373 errors:0 dropped:0 overruns:0 frame:0
	TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
	collisions:0 txqueuelen:1000
	RX bytes:72266 (70.5 KiB) TX bytes:0 (0.0 b)
	Interrupt:204 Memory:d4800000-d4ffffff
eth6	Link encap:Ethernet HWaddr 00:10:18:88:E7:B0
	inet addr:10.2.2.200 Bcast:10.255.255.255 Mask:255.0.0.0
	inet6 addr: fe80::210:18ff:fe88:e7b0/64 Scope:Link
	UP BROADCAST RUNNING MULTICAST MTU:6600 Metric:1
	RX packets:371 errors:0 dropped:0 overruns:0 frame:0
	TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
	collisions:0 txqueuelen:1000
	RX bytes:72138 (70.4 KiB) TX bytes:0 (0.0 b)
	Interrupt:204 Memory:d5800000-d5ffffff
eth7	Link encap:Ethernet HWaddr 00:10:18:88:E7:B2
	inet addr:10.2.2.201 Bcast:10.255.255.255 Mask:255.0.0.0
	inet6 addr: fe80::210:18ff:fe88:e7b2/64 Scope:Link
	UP BROADCAST RUNNING MULTICAST MTU:7700 Metric:1
	RX packets:336 errors:0 dropped:0 overruns:0 frame:0
	TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
	collisions:0 txqueuelen:1000
	RX bytes:63944 (62.4 KiB) TX bytes:0 (0.0 b)
	Interrupt:181 Memory:d6800000-d6ffffff
eth8	Link encap:Ethernet HWaddr 00:10:18:88:E7:B4
	inet addr:10.2.2.202 Bcast:10.255.255.255 Mask:255.0.0.0
	inet6 addr: fe80::210:18ff:fe88 <u>:e7b4/64 S</u> cope:Link
	UP BROADCAST RUNNING MULTICAST MTU:8800 Metric:1
	RX packets:316 errors:0 dropped:0 overruns:0 frame:0
	TX packets:0 errors:0 dropped:0 overruns:0 carrier:0
	collisions:0 txqueuelen:1000
	RX bytes:60760 (59.3 KiB) TX bytes:0 (0.0 b)
	Interrupt:181 Memory:d7800000-d7ffffff
othQ	link encan Ethernet Hwaddr 88,18,18,88,57,86
C (115	inet addr. 10 2 2 203 Reast. 10 255 255 Mack. 255 0 0 0
	inet6 addr: fe80: 210:18ff:fe80:e7h6/64 Scope:Link
	UP BROADCAST RUNNING MULTICAST MILLOGAR Metricel
	RY packate: 181 errors: 8 dropped: 8 overrups: 8 frame: 8
	TV packate: 0 arrors: 0 dropped: 0 overruns: 0 frame: 0
	collicions:0 trauquelon:1000
	DV butcc:22702 (22 0 KiP) TV butcc:0 (0 0 b)
	KA bytes:55/02 (52.9 KIB) IX bytes:0 (0.0 b)
	TUTELLAD WENOLÀ:08800000-08111111

Additionally, the Ethernet and iSCSI protocol-enabled adapter MTU sizes can be simultaneously adjusted in the Ethernet Devices window from the Network Configuration GUI (if available in the desired version of Linux).

	Ethernet Device X
General Boute Hardware De	evice
Nickname: 1_57712_NP0	
Activate device when con	nputer starts
Allow all users to enable a	and disable the device
Enable IPv <u>6</u> configuration	for this interface
O Automatically obtain IP ad	dress settings with: dhcp \$
DHCP Settings	
Hostname (optional):	
Automatically obtain D	NS information from provider
<ul> <li>Statically set IP addresses</li> </ul>	2
Manual IP Address Settings	
Address:	10.1.1.200
Subnet mask:	255.0.0.0
Default gateway address:	
🗹 Set MTU to: 9600 🌲	
Set MRU to: 0	



**Note:** In Linux, both the Ethernet and iSCSI Offload HBA MTU sizes are changed simultaneously and will have the same value. In other words, setting eth2 to MTU size of 9600 bytes using the ifconfig eth2 mtu 9600 up sets this example's Port 0, Partition 1 Ethernet adapter to MTU = 9600 bytes and the Port 0, Partition 1 iSCSI Offload HBA adapter to MTU = 9600 bytes.



**Note:** Each partition's MTU size setting can be different. Using the above protocol-enabled partition example:

- Port 0, Partition 1 Ethernet and iSCSI Offload HBA can be set to 9600 bytes.
- Port 0, Partition 2 Ethernet and iSCSI Offload HBA can be set to 5500 bytes.
- Port 0, Partition 3 Ethernet can be set to 1500 bytes.
- Port 0, Partition 4 Ethernet can be set to 9000 bytes.

In Linux, use the ping command with the "-s size" option to verify that Jumbo Frame support is configured throughout the desired network path - i.e. "ping -s 8972 A.B.C.D". The Don't Fragment (a.k.a. DF) flag is AUTOMATICALLY set in Linux. The unfragmentable ping packet size is the desired MTU size to be checked (9000 bytes) minus the automatically added overhead (28 bytes) or 8972 bytes. Note that the ping will return the send size plus the 8 bytes of the ICMP header. The "-c" switch sets the number of times this ping command is sent. You can also use the optional source interface (Capital I) command "-I a.b.c.d" or "-I devName" if you have multiple adapters that are connected to the same target IP address and you want to check a specific interface.

[root@server]# ping -c 3 -s 8972 192.168.20.10

PING 192.168.20.10 (192.168.20.10) from 192.168.20.200:8972(9000) bytes of data. 8980 bytes from 192.168.20.10 (192.168.20.10): icmp\_seq=0 ttl=255 time=0.185 ms 8980 bytes from 192.168.20.10 (192.168.20.10): icmp\_seq=1 ttl=255 time=0.177 ms 8980 bytes from 192.168.20.10 (192.168.20.10): icmp\_seq=2 ttl=255 time=0.180 ms
--- 192.168.20.10 ping statistics --3 packets transmitted, 3 packets received, 0% packet lost, time 3043ms
rtt min/ave/max/mdev = 0.177/0.181/0.185/0.005 ms
If it does not work, you will see:
[root@server]# ping -c 3 -s 8972 192.168.20.10
PING 192.168.20.10 (192.168.20.10) from 192.168.20.200:8972(9000) bytes of data.
--- 192.168.20.10 ping statistics --3 packets transmitted, 0 packets received, 100% packet lost, time 3010ms

### Setting MTU Sizes in VMWare ESX/ESXi 4.1

In VMWare ESX/ESXi 4.1, the MTU size for each individual Ethernet protocol-enabled partition can be independently set from Normal (1500 bytes) up to Jumbo (9600 bytes). This MTU size change will affect both regular L2 Ethernet and iSCSI software non-offload pathway generated traffic. Unlike other Linux OS's, you can not directly adjust the MTU size using the "ifconfig" command. Instead you must use various ESX 4.1 "esxcfg" commands.

The Ethernet protocol enabled partition's adapter MTU size is adjusted using SSH command line, for example:

List all current vSwitch's information using the esxcfg-vswitch -1 command.

Modify a specific vSwitch's MTU size with the following command - this modifies vSwitch1 (which is in port group VMK\_I\_1) to 9000 bytes:

esxcfg-vswitch -m 9000 vSwitch1 esxcfg-vswitch -1 witch Name Num Ports Used Ports Configured Ports MTUUplinks Switch0 128 128 1500 vmnic0 3 PortGroup Name VLAN ID Used Ports Uplinks VM Network vmnic0 Management Network vmnic0 witch Name Num Ports Used Ports Configured Ports Uplinks MTU Switch1 9000 128 128 vmnic6 PortGroup Name VLAN ID Used Ports Uplinks VMK NIC 1 vmnic6 VMK I 1 vmnic6

esxcfg-vswitch -m 9000 vSwitch1

The following command will list all of the current VMKernel NIC settings:

esxcfg-vmknic -1

Modify the VMKernel NIC's MTU size using the following command. This modifies **vmnic1** (which is in port group VMK\_I\_1) to 9000 bytes:

esxcfg-vmknic -m 9000 VMK\_I\_1

· # esxofg-	-vmknic -m 2000 VMK 1	1 1							
(2011-09-22	18:35:45 'NotifyDC	I' warning	] Notifying the DCUI of configuration c	hange					
<ul> <li># esxolg-</li> </ul>	-vmknic -1								
Interface	Port Group/DVPort	IP Family	IP Address	Netmask	Broadcast	MAC Address		TSO MSS	Enabled T
/pe									
rmikt)	Management Network	IPv4		255.255.0.0	1.10.255.255	78:2b:cb:27:b2:18	1500	65535	true S
TATIC									
rmk2	VMK_I_2	IPv4		255.255.0.0	2.2.255.255	00:50:56:76:08:£4	1500	65535	true S
CATIC									
rmk3	V20C_3	IPv4				00:50:56:76:c0:a9		65535	true S
DATIC									
rm/c4	VMK_4	IPv4		255.255.0.0	4.4.255.255	0015015617b10e100			true S
TATIC									
rmk5	VMK_5	IPv4	5.5.41.1	255.255.0.0	5.5.255.255	00:50:56:70:3c:dc		65535	true S
TATIC									
rmk6	V205_6	IPv4			6.6.255.255	00:50:56:71:90:a1			true S
TATIC									
rmie7	VMK_I_7	IPv4				00:50:56:77:5d:71	1500		true S
TATIC									
rm.lett	VMK_I_8	IPv4	8.8.41.1			001501561731a01d7			true S
DITATIC									
mk1	VMK_I_1	IPv4				00:50:56:73:2a:1d	9000		true S
TATIC									



Note: VMWare ESX uses the term vmnic# instead of the typical Linux term eth# in the command line.

Finally, don't forget to set the associated VM's virtual adapter setting as desired. In a Windows Server 2008 R2 VM, the **Network Connection Advanced Properties Jumbo Packet** setting would be either 9014 Bytes or DISABLED for standard size 1500 Bytes if the Virtual Adapter is the default E1000.

The following properties are available for the property you want to change on the on the right.	r this network adapter. Click left, and then select its value
Property:	<u>V</u> alue:
Adaptive Inter-Frame Spacing Row Control Interrupt Moderation Interrupt Moderation Rate IPv4 Checksum Offload Jumbo Packet Large Send Offload (IPv4) Link Speed & Duplex Locally Administered Address Number of Coalesce Buffers Priority & VLAN Receive Buffers TCP Checksum Offload (IPv4) Transmit Buffers	Disabled The second sec

If you are using the VMXNET3 Virtual Adapter, then change it's **Advanced Properties Jumbo Packets** setting to **Jumbo 9000**.

vmxnet3 Ethernet Adapter Properties	×
General Advanced Driver Details Pow The following properties are available for this the property you want to change on the left, on the right.	er Management
Property: Interrupt Moderation IPv4 Checksum Offload IPv4 Giart TSO Offload IPv4 TSO Offload IPv6 TCP Segmentation Offload Umbo Packet Large Rx Buffers MAC Address Max Rx Queues Max Tx Queues Max Tx Queues Offload IP Options Offload TCP Options Priority / VLAN tag	Value: Jumbo 9000
	OK Cancel

In a RHEL VM, the virtual adapter MTU sizes can be adjusted in the Ethernet Devices window from the **Network Configuration** GUI (or command line using the ifconfig ethX mtu NNNN up command).

				Et	2	n	110	1	111	et	2 D	IGA1	lce							0
eneral	Route	Hardware D	evice																	
licknam	e: 573	712_eth2		_					_	_										٦
Activ	ate dev	vice when co	mpute	er	s	st	ta	art	ts											_
Allow	all <u>u</u> se	ers to enable	and di	isa	al	b	ole	e t	th	ie (	de	vic	e							
Enab	le IPv <u>6</u>	configuration	for th	his	s	ir	ini	te	erfa	ac	ce									
Autor	matical	ly obtain IP a	ddress	s s	se	e	ett	tin	ngs	s v	with	th:								
-OHCP	natical	ly obtain IP a	ddress	s s	se	e	ett	tin	ngs	s v	with	h:								
-OHCP Hosto	natical	ly obtain IP a	ddress	s s	se	e	ett	tin	ngs	is v	with	h:								
-CHCP Hostn	natical	ly obtain IP a	ddress	5 5	se	e	ett	tin	ngs	s v	with	h:								
	natical Setting ame (o	ly obtain JP a ptionale	ddress	5 5	se	e	ett	tin	ngs	s v	with	h: (								
Autor     Autor     Autor     Autor     Autor     Autor     Autor     Static	natical Setting ame to tomate ally se	ly obtain JP a	ddress (HS Inf	S S	se	e	ett	tin	ngs	is v	fre	h: (								
Autor     A	anatical Setting annation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation connation	ly obtain IP a ptionale cally obtain I t IP addresse dress Setting	ddress	5 5	se	e	ett	tin	ngs	is v	with	h:	d	hicp With	fel	•	]			
Static Addre	natical Setting ame (o comati cally se i IP Ado ss:	ly obtain JP a ptionate cally obtain g t IP addresse dress Setting	ddress (45 m 5 1.1.1	1.	1	12	2	tin	ngs tio	is v	with	th: (	d	hop Wio	ie j	•	]			
<ul> <li>Autor</li> <li>OHCP</li> <li>Hosts</li> <li>Manua</li> <li>Addre</li> <li>Subne</li> </ul>	natical Setting ame (o comati cally se il IP Ado ss: et mask	ly obtain JP a ptionale cally obtain g t IP addresse dress Setting	ddress (15 m 5: 5 1.1.1 255.	1.2	1	12	2	.0	ng:		with	h: (	a pro	100	ie)	•	)			

In Linux, use the following ping command to verify that Jumbo Frame support is configured throughout the desired network path:

**Note:** The odd numbering for the pink packet size (-s NNNN), which is the desired MTU size of 9000 bytes minus the 28 bytes of automatically added OVERHEAD, but the outputted displayed size is plus 8 bytes which are the included ICMP header data bytes. The Don't Fragment (a.k.a. DF) flag (-d) must also be set, else the packet could be fragmented somewhere along the way and you would not know it. The optional count (-c) is the number of times you will send this ping. You don't have to use the optional source interface command "-I a.b.c.d" or "-I devName" unless you have multiple adapters that lead to the same target IP address and you want to check a specific one. You can also use the ESX "vmkping" command.

```
[root@sieoral network-scripts]# ping -d -c 3 -I eth1 -s 8972 192.168.20.10
PING 192.168.20.10 (192.168.20.10) from 192.168.20.101 eth1:8972(9000) bytes of data.
8980 bytes from 192.168.20.10: icmp_seq=1 ttl=255 time=0.185 ms
8980 bytes from 192.168.20.10: icmp_seq=2 ttl=255 time=0.177 ms
8980 bytes from 192.168.20.10: icmp_seq=2 ttl=255 time=0.180 ms
---- 192.168.20.10 ping statistics ----
3 packets transmitted, 3 packets received, 0% packet lost
round-trip min/ave/max = 0.177/0.181/0.185ms
If it does not work you will see a failure message:
[root@esx41]# ping -d -c 3 -I eth1 -s 8972 192.168.20.10
PING 192.168.20.10 (192.168.20.10) from 192.168.20.101 eth1:8972(9000) bytes of data.
sendto () failed (Message too long)
sendto () failed (Message too long)
sendto () failed (Message too long)
--- 192.168.20.10 ping statistics ---
```

3 packets transmitted, 0 packets received, 100% packet lost

You should also check if the desired VMs can send Jumbo Frames all the way to their various end points using the applicable OS ping commands described earlier in this document.

# **Examples**

- Equal Oversubscription Example
- Partitioned Oversubscription Example
- Weighted Oversubscription Example
- Oversubscription With One High Priority Partition Example
- Default Fixed Subscription Example
- Mixed Fixed Subscription and Oversubscription Example
- Mixed Weights and Subscriptions Example

# **Note:** All bandwidths given in these examples are approximations. Protocol overhead, application send-rate variances, and other system limitations may give different bandwidth values, but the ratio relationship between the send bandwidths of the four partitions on the same port should be similar to the ones given.

Depending upon OS requirements, the traffic types for each partition could be L2 Ethernet or iSCSI Offload or FCoE Offload where L2 Ethernet can be on any of the four partitions, AND up to two iSCSI Offload can be on any two of the four partitions OR one FCoE Offload can be on any one of the four partitions plus one iSCSI Offload can be on any one of the remaining three partitions. The partitions data flows on Windows can be a combination of L2 Ethernet traffic (with or without TOE) and/or HBA traffic (iSCSI OR FCoE). For Linux RHEL v5.x, the L2 Ethernet protocol is always enabled and up to two iSCSI Offloads can be enabled (but no FCoE Offload). For Linux RHEL v6.x and SLES 11 SP1 the L2 Ethernet protocol is always enabled and one iSCSI Offload can be enabled. For Solaris 10u9 and VMWare ESX/ESXi 4.1 only the L2 Ethernet protocol is available for the partitions.

## **Equal Oversubscription Example**

The following is an example of oversubscribed bandwidth sharing in non-DCB mode. All traffic types over the four partitions of the port have an equal weight (i.e., they are all set to 0%) and can individually use the maximum bandwidth of the connection (i.e., 10 Gbps, in this case). In addition to the Ethernet Protocol's being enabled on all four partitions, the iSCSI Offload protocol is enabled on partition 1 and 4. The iSCSI Offload Protocol can be enabled in any two partitions. When all of the partitions have zero relative bandwidth weights, each traffic flow will act as if in it's own separate partition, each taking an equal share of the available bandwidth up to that partitions maximum bandwidth (which is 100% in this example so does not further limit any of the traffic flows).

FCoE Offload is not available in non-DCB mode so two iSCSI Offload protocols are used here.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1i)	0	100	iSCSI Offload	Green
Port 0, Partition 1 (P1e)	0	100	Ethernet	Brown
Port 0, Partition 2 (P2e)	0	100	Ethernet with TOE	Purple
Port 0, Partition 3 (P3e)	0	100	Ethernet	Yellow
Port 0, Partition 3 (P4e)	0	100	Ethernet with TOE	Blue
Port 0, Partition 4 (P4i)	0	100	iSCSI Offload	Red

#### Table 3: Non-DCB Equal Oversubscription

The following plot shows how all of the partitions would share a ports available send bandwidth. Each traffic type flow (such as P1i and P1e) is expanded in to its own bandwidth trace for ease of understanding. The send traffic flows are independent in each partition and the individual traffic type flow rate is balanced with each of the other traffic type flow rates when traffic demands exceed the available bandwidth

• Starting at t0, the first partition's iSCSI Offload traffic flow (P1i) initially takes ~100% of the available port's

TX send bandwidth when an iSCSI test application is flooding that port by itself.

- When P1's L2 Ethernet (P1e) starts to send at **t1**, both stabilize to half of the bandwidth or ~5 Gbps each even thought they are in the same partition, they share the total available bandwidth.
- When P2e starts to send at **t2**, all three traffic flows (P1i, P1e and P2e) will stabilize to 1/3rd of the bandwidth or ~3.3 Gbps each (they all equally share the available bandwidth).
- When P3e starts to send at **t3**, all four traffic flows (P1i, P1e, P2e and P3e) will stabilize to 1/4th of the bandwidth or ~2.5 Gbps each (effectively sharing the available bandwidth).
- When P4e starts to send at **t4**, all five traffic flows (P1i, P1e, P2e, P3e and P4e) will stabilize to 1/5th of the bandwidth or ~2 Gbps each (again equally sharing the available bandwidth).
- When P4i starts to send at **t5**, all six traffic flows (P1i, P1e, P2e, P3e, P4e and P4i) will stabilize to 1/6th of the bandwidth or ~1.65 Gbps each (all sharing the available bandwidth).
- When P1i stops sending at **t6**, the five currently active traffic flows (P1e, P2e, P3e, P4e and P4i) will readjust to ~2 Gbps each (equally absorbing the freed up bandwidth).
- As a previously sending traffic flow stops sending (**t7**, **t8**, **t9** and **t10**) the remaining active flows will readjust to equally fill any available bandwidth.
- Notice the symmetry of the BW allocation. No matter which traffic type is currently running, each will get an equal share with respect to the other currently transmitting traffic type flows. This assumes the application creating the transmitted traffic type flow can fill the allocated amount of BW it is given if not, the other traffic flows will equally absorb the unused BW.



The following two examples are of oversubscribed bandwidth sharing with DCB enabled. The first example is similar to the above non-DCB example with four Ethernet and two iSCSI Offload protocols enabled with the same two traffic types in six distinct "flows" with the partitions similarly configured. The main differences here is the iSCSI traffic type is assigned to DCB Priority Group 2 and is Lossless (i.e. iSCSI-TLV) with an ETS setting of 50%; while the L2 Ethernet traffic type is still assigned to Priority Group 0 and is Lossy with an ETS setting of 50%. If the iSCSI Offload protocol traffic flows had been assigned to the same PG as the Ethernet protocol traffic flows, then the traffic BW would have looked very similar to the previous Non-DCB example since ETS would never be activated for traffic flows belonging to the same PG.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1i)	N/A	100	iSCSI Offload	Green
Port 0, Partition 1 (P1e)	N/A	100	Ethernet	Brown
Port 0, Partition 2 (P2e)	N/A	100	Ethernet with TOE	Purple
Port 0, Partition 3 (P3e)	N/A	100	Ethernet	Yellow
Port 0, Partition 4 (P4e)	N/A	100	Ethernet	Blue
Port 0, Partition 4 (P4i)	N/A	100	iSCSI Offload	Blue

#### Table 4: DCB Equal Oversubscription

The following plot shows how the two iSCSI traffic streams in PG2 act verses the L2 Ethernet traffic streams in PG0. The traffic in the two PGs will almost act independently of each other when their aggregated traffic bandwidth demands exceed the available bandwidth - each taking it's half of the ETS managed pie.

- Starting at **t0**, only P1i (iSCSI Offload) is sending, so it takes ~100% or all of the 10 Gbps bandwidth.
- When P1e (L2 Ethernet) starts to send at **t1**, both flows stabilize to ~5 Gbps each (P1i in PG2 takes it's allocated 50% bandwidth and P1e in PG0 takes it's allocated bandwidth of 50%).
- When P2e starts to send at **t2**, the traffic in P1i is not affected it remains at ~5 Gbps due to it being in a different Priority Group. Both P1e and P2e will stabilize to ~2.5 Gbps each (P1e and P2e equally share PG0's allocated portion of the bandwidth).
- When P3e starts to send at t3, P1i is still unaffected and remains at ~5 Gbps. The three L2 Ethernet traffic types will split their 50% of PGO's share between themselves, which is ~1.65 Gbps (each takes 1/3rd of ~5 Gbps).
- When P4e starts to send at **t4**, the four Ethernet traffic flows take 1/4th of PGO's bandwidth (~1.25 Gbps each) while P1i is still unaffected and remains at ~5 Gbps.
- When P4i starts to send at **t5**, the four Ethernet traffic flows remain the same but the two iSCSI traffic flows split PG2's allocated bandwidth (~2.5 Gbps each).
- Then when P1i stops sending at **t6**, the traffic flows in PGO are unaffected while P4i's share increases to all of PG2's allocated bandwidth of ~5 Gbps.
- As each of the traffic flows stops sending in PG0, the traffic flows of the remaining member's of PG0 equally increase their respective shares to automatically occupy all of the available bandwidth remain in PG0 until all of PG0's Ethernet flows stop. At **t7**, there are three active PG0 flows (P2e, P3e and P4e) so each gets 1/3rd of PG0's 5 Gbps or ~1.65 Gbps. At **t8**, there are two active PG0 flows (P3e and P4e) so each gets ½ of PG0's 5 Gbps or ~2.5 Gbps. At **t9**, there is only one active PG0 flow (P4e) so it gets all of PG0's bandwidth or ~5 Gbps. Through all of this, the Lossless iSCSI flow of P4i remains at 5 Gbps since it takes all

of PG2's portion of the overall bandwidth (ETS of 50%). Finally, at **t10**, there is only one active flow after P4e stops sending so at this point P4i gets 100% of all the bandwidth or ~10 Gbps.



Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1f)	N/A	100	FCoE Offload	Green
Port 0, Partition 1 (P1e)	N/A	100	Ethernet with TOE	Brown
Port 0, Partition 2 (P2e)	N/A	100	Ethernet	Purple
Port 0, Partition 3 (P3e)	N/A	100	Ethernet	Yellow
Port 0, Partition 4 (P4e)	N/A	100	Ethernet with TOE	Blue
Port 0, Partition 4 (P4i)	N/A	100	iSCSI Offload	Red

Table 5: DCB Equal Oversubscription with one Lossless FCoE Offload

The following plot shows how the first partition's FCoE traffic PG1 acts verses the other traffic types in PG0. Just like the previous example, the traffic in the two PGs will almost act independently of each other when their aggregated traffic bandwidth demands exceed the available bandwidth.

- Starting at **t0**, only P1f (FCoE Offload) is sending, so it takes ~100% or all of the 10 Gbps bandwidth.
- When P1e (L2 Ethernet) starts to send at **t1**, both flows stabilize to ~5 Gbps each (P1f in PG1 takes 50% and P1e in PG0 takes the other 50%).
- When P2e starts to send at t2, the traffic in P1f is not affected it remains at ~5 Gbps due to it being in a different PG. Both P1e and P2e will stabilize to ~2.5 Gbps each (P1e and P2e equally share PG0's portion of the bandwidth so they each get ~2.5 Gbps (ETS of 50% \* total 10G BW \* 1 / 2).
- When P3e starts to send at **t3**, P1f is still unaffected and remains at ~5 Gbps. The three L2 Ethernet traffic types will split their 50% of PGO's share between themselves which is ~1.65 Gbps (each takes 1/3rd of 5G).
- When P4e starts to send at **t4**, the four Ethernet traffic flows take 1/4th of PGO's bandwidth (~1.25 Gbps each) while P1f is still unaffected and remains at 5 Gbps.
- When P4i starts to send at **t5**, the four Ethernet traffic flows plus the new iSCSI traffic flow take 1/5th of PG0's bandwidth (~1 Gbps each) while P1f is still unaffected and remains at 5 Gbps.
- Then when P1f stops sending at **t6**, the five traffic flows in PGO now take all of the ports bandwidth, so now their 1/5th of PGO's bandwidth doubles to ~2 Gbps each the available bandwidth went from 5 Gbps to 10 Gbps.
- As each traffic flow stops sending in PGO, the remaining member traffic flows equally increase their respective shares to automatically occupy all of the available bandwidth. At **t7**, there are four active PGO flows so each gets 1/4th or ~2.5 Gbps. At **t8**, there are three active PGO flows so each gets 1/3rd or ~3.3 Gbps. At **t9**, there are two active PGO flows so each gets half or ~5 Gbps. Finally, at **t10**, there is only one active PGO flow (P4i) so it gets 100% or ~10 Gbps.
- Any of the traffic flows will take 100% of the available bandwidth if it is the only sending traffic flow.



# **Partitioned Oversubscription Example**

The following is an example of oversubscribed bandwidth sharing in non-DCB mode where all four partitions of the port have their weight set to 25% and can individually use the maximum bandwidth of the connection (i.e., 10 Gbps, in this case). In addition to the Ethernet Protocol's being enabled on all four partitions, the iSCSI Offload protocol is enabled on Partition 1 and 4. By setting the partition's relative bandwidth weights to 25%, each partition's traffic flows (i.e. P1's iSCSI (P1i) + L2 Ethernet (P1e) and P4's iSCSI (P4i) + L2 Ethernet (P4e)) will be contained in their respective partition while each partition over all takes an equal share of the available bandwidth. The traffic flows within that partition can only expand into that partition's allocated by weight portion.

There would be no difference between the previous examples and this example with DCB enabled since it essentially sets the Relative Bandwidth Weight values to all ZEROs and the PG's ETS would come into play.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1i)	25	100	iSCSI Offload	Green
Port 0, Partition 1 (P1e)	25	100	Ethernet	Brown
Port 0, Partition 2 (P2e)	25	100	Ethernet with TOE	Purple
Port 0, Partition 3 (P3e)	25	100	Ethernet	Yellow
Port 0, Partition 3 (P4e)	25	100	Ethernet with TOE	Blue
Port 0, Partition 4 (P4i)	25	100	iSCSI Offload	Red

#### Table 6: Non-DCB Partitioned Oversubscription

The following plot shows how each traffic type flow must remain within a partition's share of a ports available send bandwidth - i.e. if there are two different traffic type flows (such as P1i and P1e) in a single partition, they are combined, as if one flow, for determining the amount of bandwidth allocated to them.

- Starting at **t0**, the first partition's iSCSI Offload traffic flow (P1i) initially takes ~100% of the available port's TX send bandwidth when an iSCSI test application is flooding that port by itself.
- When P1's L2 Ethernet (P1e) starts to send at **t1**, both stabilize to ~5 Gbps each even thought they are in the same partition, they share the total available bandwidth. This is because no other partition's traffic flow is sending.
- When P2e starts to send at **t2**, the traffic flows in P1 (P1i and P1e) will reduce to ~2.5 Gbps each while the P2 traffic flow (P2e) will take ~5 Gbps. This is because the bandwidth is initially split by partition and then traffic flows within each individual partition.
- When P3e starts to send at **t3**, the two traffic flows in P1 (P1i and P1e) are further reduced to ~1.65 Gbps (half of the partition's 1/3rd allocation going to the three active partitions) while P2e and P3e each stabilize at ~3.3 Gbps.
- When P4e starts to send at **t4**, the three single partition traffic flows (P2e, P3e and P4e) will stabilize to 1/ 4th of 10 Gbps or ~2.5 Gbps each while the P1 partition shares it allocated bandwidth between it's two users (P1i and P1e) so each gets half of the allocated 1/4th of 10 Gbps or 1/8th which is ~1.25 Gbps.
- When P4i starts to send at **t5**, the two single traffic flows (P2e and P3e) will remain at ~2.5 Gbps each, as well as the P1 partition's traffic flows (P1i and P1e) each still getting ~1.25 Gbps while P4's allocated

bandwidth will now be split into two traffic flows (P4e and P4i) which means each get ~1.25 Gbps.

- When P1i stops sending at **t6**, the only partition P1 traffic flow (P1e) will readjust to ~2.5 Gbps and all of the others will remain the same.
- When P1e stops sending at **t7**, the other traffic flows will readjust to ~3.3 Gbps (P2e and P3e) and ~1.65 Gbps each for partition P4's shared P4e and P4i traffic flows.
- When P2e stops sending at **t8**, partition P3's single traffic flow will readjust to ~5 Gbps (P3e) and partition P4's shared P4e and P4i traffic flows will increase to half that or ~2.5 Gbps.
- When P3e stops sending at **t9**, the remaining partition (P4) will now receive all of the available bandwidth so it's two traffic flows (P4e and P4i) will equally share it for ~5 Gbps each.
- Finally, when P4e stops sending at t10, the remaining traffic flow (P4i) will now receive all of the available bandwidth or ~10 Gbps.
- If there is only one flow in a partition, it would take all of the bandwidth allocated for that partition.
- The main difference between setting all four partition's relative bandwidth weight to 0% and setting them to all 25% is that 0%'s causes the send bandwidth to be shared between all active traffic flows while 25%'s cause the send bandwidth to be shared between the active sending partitions first and then the active sending traffic type flows in a two step manner. Setting them to all 0%'s causes the logic to work similarly to the way it does in DCB mode when all traffic types are in the same PG.
- If there was only one traffic flow in each partition, then the results would be similar to setting each partition's relative bandwidth weight to 0%, since the single traffic flow would not be sharing a partition's bandwidth with another traffic type.



# Weighted Oversubscription Example

The following is an example of weighted oversubscribed bandwidth sharing with different weights assigned to each partition in non-DCB mode. This example has each partition taking the maximum bandwidth when no other partition is active, plus as each partition starts and stops sending, the amount of bandwidth is shared as an approximate ratio of the currently sending partitions weight values.

There would be no difference between the previous examples and this example with DCB enabled since the weights are ignored and if all traffic types are in the same single PG.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1i)	10	100	iSCSI Offload	Green
Port 0, Partition 1 (P1e)	10	100	Ethernet with TOE	Orange
Port 0, Partition 2 (P2e)	20	100	Ethernet with TOE	Purple
Port 0, Partition 3 (P3e)	30	100	Ethernet with TOE	Yellow
Port 0, Partition 4 (P4e)	40	100	Ethernet with TOE	Blue
Port 0, Partition 4 (P4i)	40	100	iSCSI Offload	Red

#### Table 7: Non-DCB Weighted Oversubscription

The following plot shows:

- The first partition's traffic flow (P1i) initially takes ~100% of the available bandwidth at t0 when an iSCSI test application is sending traffic out that port by itself.
- When P1e starts to send Ethernet traffic at **t1**, the two active traffic flows have equal weights with respect to each other so they are allocate half of the total bandwidth available (~10 Gbps) to partition P1 which equates to ~5 Gbps each for P1i and P1e.
- When P2e starts sending at t2, the partition's relative bandwidth weights come into effect. Partition P1 has a weight of 10% while P2 has twice as much at 20%, so P1's two sending traffic flows are reduced to half of the partition's assigned 1/3rd (derived from P1's weight of 10% / (P1's weight of 10% + P2's weight of 20%)) or ~1.65 Gbps each for P1i and P1e. P2e starts at ~6.7 Gbps (it's relative weight is 20% / 30% total active weights) it is not halved since it is the only traffic-flow on partition P2.
- When partition P3e starts sending Ethernet traffic at t3 with a relative weight of 30%, it takes ~5 Gbps (30/ 60 of 10 Gbps), P2e drops to ~3.3 Gbps (20/60) and partition P1's total drops to ~1.65 Gbps (10/60) so that means P1i and P1e each get half of that or ~0.825 Gbps each.
- When P4e starts (40% relative weight) at **t4**, it takes ~4 Gbps (40/100) and the three other partition's send traffic drop; partition P1 is reduced to ~1 Gbps (10/100) so that means P1i and P1e split that for ~0.5 Gbps; partition P2 drops to ~2 Gbps (20/100) and since there is only one send traffic flow (P2e) it takes all of that assigned bandwidth; and finally partition P3 (with it's single traffic flow P3e) drops to ~3 Gbps (30/100).
- When the second traffic flow on partition P4 (P4i) starts at **t5**, the two flows (P4e and P4i) on the same partition (P4) split the partition's assigned bandwidth of ~4 Gbps, so each gets ~2 Gbps. The other send traffic on the other three partitions remains the same.

- If P1i stops at **t6**, the remaining traffic flow on partition P1 (P1e) absorbs that partition's share of the send bandwidth to ~1 Gbps. The remaining traffic flows on the other three partitions are unaffected.
- If P1e stops at t7, the others adjust slightly upwards to fill that newly available bandwidth; partition P2 (and it's traffic flow P2e) increases to ~2.2 Gbps (20/90); partition P3 (and it's traffic flow P3e) raises to ~3.3 Gbps (30/90); and partition P4 raises to ~4.5 Gbps (40/90) which means P4e and P4i split that for ~2.25 Gbps each.
- If P2e stops at **t8**, the others again adjust upwards to fill that newly available bandwidth; partition P3 (and it's traffic flow P3e) raises to ~4.3 Gbps (30/70), and partition P4 raises to ~5.7 Gbps (40/70) which means P4e and P4i split that for ~2.85 Gbps each.
- If P3e stops at **t9**, partition P4's share raises to ~100% of the bandwidth (40/40) so it's two traffic flows (P4e and P4i) split this for ~5 Gbps each.
- 1 1

t6

t7

t8

• When P4e stops at **t10**, the only remaining traffic flow P4i takes all of the bandwidth at ~10 Gbps.

ť2

t3

t4

t5

t0

t1

t10

t9

## **Oversubscription With One High Priority Partition Example**

Next is an example of a single high priority partition with all of the relative bandwidth weight, but all four of the partitions are still oversubscribing the available bandwidth. The first three partitions of the port have 0% weight and the last partition (P4) has all of the weight (100%). All four partitions are set to use the maximum amount of the connection's bandwidth (i.e., 100%, which is 10 Gbps, in this case).

Again, since the maximum bandwidths are set to 100% for all four partitions, there is no difference between the earlier DCB mode example and this one, in DCB mode.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1i)	0 (effectively 1)	100	iSCSI Offload	Green
Port 0, Partition 1 (P1e)	0 (effectively 1)	100	Ethernet	Brown
Port 0, Partition 2 (P2e)	0 (effectively 1)	100	Ethernet	Purple
Port 0, Partition 3 (P3e)	0 (effectively 1)	100	Ethernet with TOE	Yellow
Port 0, Partition 4 (P4e)	100	100	Ethernet	Blue
Port 0, Partition 4 (P4i)	100	100	iSCSI Offload	Red

#### Table 8: Non-DCB Oversubscription With One High Priority Partition

The following plot shows a similar effect to some of the previous examples, except that the fourth partition takes as much of the bandwidth as it needs (up to ~100%) when it starts to transmit. In this example the three 0% Relative Bandwidth Weight partitions have an effective Relative Bandwidth Weight of 1% instead of 0%.

- The first partition's traffic flow (P1i) initially takes ~100% of the available bandwidth when the test application starts to transmit traffic on that port by itself at **t0**.
- When P1e starts to send Ethernet traffic at t1, both will stabilize to ~5 Gbps each.
- When partition P2 (traffic flow P2e) starts to send traffic at **t2**, since only partition's P1 and P2 are now sending each partition gets half of the send bandwidth, so partition P2 (P2e) gets all of the allocated ~5 Gbps and partition P1's two traffic flows (P1i and P1e) will share it's allocated ~5 Gbps or ~2.5 Gbps each.
- When partition P3 (P3e) starts to send at **t3**, all three partitions will be allocated 1/3rd of the available bandwidth P3e will received ~3.3 Gbps, P2e will receive the same allocation of ~3.3 Gbps and P1i an P1e will approximately split it's partition's bandwidth for ~1.65 Gbps each.
- But when P4e starts to send Ethernet traffic at **t4**, it will take almost all of the ~10 Gbps bandwidth, regardless of the bandwidth needs of the other three partitions four traffic flows. P1i and P1e will each get approximately half of 1/103 of the available bandwidth or ~0.05 Gbps while P2e and P3e will receive ~0.1 Gbps and P4e will take ~9.7 Gbps.
- When P4i starts to send Ethernet traffic at **t5**, it will take half of the allocated bandwidth for partition P4. Therefore P4e will drop to ~4.75 Gbps and P4i will start at ~4.75 Gbps. The other three partitions four traffic flows will be unaffected. P1i and P1e will each get approximately half of 1/103 of the available bandwidth or ~0.05 Gbps while P2e and P3e will receive ~0.1 Gbps and P4e will take ~9.7 Gbps
- When P1i stops sending traffic at t6, it's freed up 0.05 Gbps bandwidth will be reallocated to the other traffic flows according to their relative bandwidth weight settings. The same is true for when P1e (at t7), P2e (at t8) and P3e (at t9) stop sending traffic.

- Finally when P4e stops sending traffic at **t10**, P4i will take all of the available bandwidth for ~10 Gbps.
- Whenever the fourth partition's bandwidth needs drop off, the other actively sending partitions will equally increase their respective shares to automatically occupy all of the available bandwidth.



# **Default Fixed Subscription Example**

This is an example of the default partition settings that has all of the relative bandwidth weights set to 0% and the maximum bandwidths set to 2.5 Gbps. Since the total of the maximum bandwidth values are set to exactly 100% (i.e. never can reach an oversubscription situation), the traffic flows in each partition will share that partition's bandwidth allocation with respect to it's overall maximum bandwidth ceiling and the relative bandwidth weights are never used.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1i)	0	25	iSCSI Offload	Green
Port 0, Partition 1 (P1e)	0	25	Ethernet	Brown
Port 0, Partition 2 (P2e)	0	25	Ethernet	Purple
Port 0, Partition 3 (P3e)	0	25	Ethernet	Yellow
Port 0, Partition 4 (P4e)	0	25	Ethernet	Blue
Port 0, Partition 4 (P4i)	0	25	iSCSI Offload	Red

 Table 9: Non-DCB Default Fixed Subscription

This following plot shows how the four partition's send traffic is independent of each other. Unlike the previous examples, none of the partitions (and their associated traffic flows) take more than their designated bandwidth portion; the total bandwidth of all four partitions of the port is equal to or less than the total available bandwidth of the port. In this example, each partition takes only ~25% or 2.5 Gbps of the total available bandwidth when their test application starts to transmit traffic. Furthermore, if a partition has more than one active traffic flow, these flows will share that partition's allowed bandwidth. Unused port bandwidth is not reallocated to any partition above it's own maximum bandwidth setting.

- When P1i starts to send traffic at **t0**, it only takes the subscribed 25% of the 10 Gbps bandwidth available which is ~2.5 Gbps.
- when P1e starts to send at **t1**, it will share partition P1's 25% with P1i. Each is allocated ~1.25 Gbps and neither expands into the unused ~7.5 Gbps remaining.
- When P2e starts to send at **t2**, it only takes it's partitions subscribed ~2.5 Gbps and does not affect either of partition P1's sending traffic flows.
- When P3e starts to send at **t3**, it again only takes it's partitions subscribed ~2.5 Gbps and does not affect P2e or either of partition P1's sending traffic flows.
- When P4e starts to send at **t4**, it also only takes it's partitions subscribed ~2.5 Gbps and does not affect any of the other partition's sending traffic flows.
- When P4i starts to send at **t5**, it will share partition P4's 25% with P4e. Each is allocated ~1.25 Gbps and the other partitions are unaffected.
- When P1i stops sending at **t6**, it will release it's 12.5% share of the bandwidth and the other remaining partition P1 traffic flow (P1e) will increase to 2.5 Gbps while the other traffic flows are unaffected.
- When P1e stops sending at **t7**, there will only be three partitions, but each is still assigned only 25% of the overall bandwidth. The other traffic flows (P2e, P3e, P4e and P4i) will not change.
- When P2e stops sending at t8, again there will be no change to the other traffic flows.
- When P3e stops sending at **t9**, there will still be no change to the other traffic flows.

- When P4e stops sending at **t10**, the remaining traffic flow on P4 (P4i) will absorb the freed 12.5% of the partition P4's allocated bandwidth and will increase to 2.5 Gbps.
- Each partition's flows on the same port are logically isolated from the others as if they were on separate ports and a partition's send flows stopping or restarting will not affect its fellow partition's send traffic flows except where the flows are on the same partition and then they will only take the freed bandwidth for their respective partition.


## Mixed Fixed Subscription and Oversubscription Example

This example shows partitions with all of the relative bandwidth weights set the same, but with the partitions partially oversubscribing the available bandwidth, unequally. Two of the partitions are set to use 10% or 1 Gbps each of bandwidth and the other two (the ones with the hardware offload protocol's enabled) are set to use 80% or 8 Gbps of the connection's bandwidth, thus oversubscribing the connection by 80%.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1i)	0	80	iSCSI Offload	Green
Port 0, Partition 1 (P1e)	0	80	Ethernet	Orange
Port 0, Partition 2 (P2e)	0	10	Ethernet	Purple
Port 0, Partition 3 (P3e)	0	10	Ethernet	Yellow
Port 0, Partition 4 (P4e)	0	80	Ethernet	Blue
Port 0, Partition 4 (P4i)	0	80	iSCSI Offload	Red

Table 10: Non-DCB Mixed Fixed Subscription and Oversubscription

This is a combination example of a fixed subscription (three of the partitions sum to 100%), but all four sum to 180%. When all four, or at least the two larger partitions, are running traffic, they share the space with each other, up to their partition maximum bandwidth values; otherwise, they act as if they are independent connections.

- The first partition's traffic flow (P1i) initially takes its designated ~8 Gbps when the test application starts to transmit traffic at **t0** to that port by itself, not expanding into the remaining unused ~2 Gbps bandwidth.
- When the second traffic flow on the first partition (P1e) starts to send at **t1**, the two active traffic flows on the same partition share its ~8 Gbps bandwidth for ~4 Gbps each.
- When the third traffic flow (P2e) starts sending at **t2**, it only takes its partitions maximum bandwidth allowed ~1 Gbps. Partition P1's two traffic flows are unaffected.
- When the fourth traffic flow (P3e) starts sending at **t3**, it again only takes its partitions maximum bandwidth allowed ~1 Gbps. Partition P1's two traffic flows and the traffic flow on partition P2 (P2e) are unaffected.
- But when P4e starts to send traffic at t4, the condition is now oversubscribed. Since P2e and P3e uses only ~1 Gbps of their allocated 2 Gbps (10 Gbps / 5 equally weighted traffic flows) that leaves ~8 Gbps free for the other three traffic flows. The remaining traffic flows (P1i, P1e and P4e) would then be allocated ~2.6 Gbps each (8 Gbps / 3 equally weighted traffic flows).
- But when P4i starts to send traffic at **t5**, it shares the available bandwidth within it's maximums with the other traffic flows. P2e and P3e are still using only ~1 Gbps of their allocated 1.6 Gbps (10 Gbps / 6 equally weighted traffic flows) which again leaves ~8 Gbps free for the other four traffic flows. Therefore these four (P1i, P1e, P4e and P4i) are allocated ~2 Gbps each (8 Gbps / 4 equally weighted traffic flows).
- When P1i stops at **t6**, it releases it's bandwidth to the available pool and since P2e and P3e are capped by their maximum bandwidth value to 1 Gbps, the three other traffic flows (P1e, P4e and P4i) automatically take equal shares and increase their bandwidth used to ~2.6 Gbps each.
- When P1e subsequently stops sending at **t7**, P4e and P4i grab up the extra available bandwidth and go to ~4 Gbps each. Both P2e and P3e are unaffected and continue sending at ~1 Gbps each.
- When P2e stops sending at t8, P4e and P4i are not able to make use of the freed up bandwidth since they

are both in partition P4 which as a maximum bandwidth ceiling of 8 Gbps. Therefore none of the traffic flows increase their sending rates and this unused bandwidth is ignored.

- When P3e stops sending at **t9**, the same condition is still in effect. Therefore none of the remaining active traffic flows increase their sending rates to use this extra bandwidth.
- Finally, P4e stops at **t10** and this allows it's companion traffic flow (P4i) to increase to ~8 Gbps which is partition P4's maximum top end. The remaining ~2 Gbps is unassigned.



The following example is the same as the previous example, but with FCoE in the first partition. Additionally, the FCoE traffic flow is Lossless and in DCB Priority Group 1 with an ETS = 50% and the other traffic flows are Lossy and in Priority Group 0 with an ETS = 50%.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1f)	N/A	80	FCoE Offload	Green
Port 0, Partition 1 (P1e)	N/A	80	Ethernet	Orange
Port 0, Partition 2 (P2e)	N/A	10	Ethernet	Purple
Port 0, Partition 3 (P3e)	N/A	10	Ethernet	Yellow
Port 0, Partition 4 (P4e)	N/A	80	Ethernet	Blue
Port 0, Partition 4 (P4i)	N/A	80	iSCSI Offload	Red

Table 11: DCB Mixed Fixed Subscription of	and Oversubscription with	Lossless FCoE Offload
-------------------------------------------	---------------------------	-----------------------

This is a similar combination example of a fixed subscription (three of the partitions sum to 100%), but all four sum to 180%. When all four, or at least the last two partitions, are running traffic, they share the space with each other, up to their partition maximum bandwidth values and their PG's ETS settings; otherwise, they act as if they are independent connections.

- The first partition's traffic flow (P1f) initially takes its maximum bandwidth designated ~8 Gbps when the test application starts to transmit traffic at t0 to that port by itself, not expanding into the remaining unused ~2 Gbps bandwidth.
- When the second traffic flow on the first partition (P1e) starts to send at **t1**, the two active traffic flows on the same partition share its ~8 Gbps bandwidth for ~4 Gbps each. ETS does not take effect since the traffic in PGO and PG1 are still less than the amount prescribed by their respective ETS values.
- When the third traffic flow (P2e) starts sending at **t2**, it only takes its partitions maximum bandwidth allowed which is ~1 Gbps. Partition P1's two traffic flows are unaffected and the unassigned 1 Gbps bandwidth remains free.
- When the fourth traffic flow (P3e) starts sending at t3, it only takes its partitions maximum bandwidth allowed which is ~1 Gbps. Now the first partition's two traffic flows readjust so that PG 0 does not get more than 50% of the overall bandwidth i.e. PG0's P1e+P2e+P3e = 40%+10%+10% which is greater than 50%. The P1e traffic flow is reduced to 30% or ~3 Gbps and the P1f traffic flow (in PG1) is adjusted upwards to 50% or ~5 Gbps.
- When P4e starts to send traffic at t4, it equally shares PGO's ETS assigned bandwidth with P1e, P2e and P3e but since P2e and P3e use only ~1 Gbps of their allocated 1.25 Gbps (5 Gbps / 4 equally weighted traffic flows) this leaves ~3 Gbps free (5 Gbps available 2 GBps assigned to P2e and P3e) for the other two traffic flows (P1e and P4e) and they are both allocated ~1.5 Gbps each (3 Gbps / 2 equally weighted traffic flows). P1f is in PG1 so it is unaffected and keeps sending at ~5 Gbps.
- When P4i starts to send traffic at **t5**, it also equally shares PGO's bandwidth (5 Gbps / 5 equally weighted traffic flows) which means P1e, P2e, P3e, P4e and P4i all send at ~1 Gbps. P1f in PG1 is still unaffected and keeps sending at ~5 Gbps.
- When P1f stops at **t6**, it releases all of PG1's bandwidth to the available pool and since P2e and P3e are capped by their maximum bandwidth value to 1 Gbps, the three other traffic flows (P1e, P4e and P4i) automatically take equal shares of 8 Gbps and bump up their bandwidth used to ~2.6 Gbps each.
- When P1e subsequently stops sending at **t7**, P4e and P4i grab up the extra available bandwidth and go to ~4 Gbps each. Both P2e and P3e are unaffected and continue sending at ~1 Gbps each.
- When P2e stops sending at **t8**, P4e and P4i are not able to make use of the freed up bandwidth since they are both in partition P4 which has a maximum bandwidth ceiling of 8 Gbps. Therefore none of the traffic

flows increase their sending rates and the unused bandwidth is ignored.

- When P3e stops sending at **t9**, the same condition is still in effect. Therefore none of the remaining active traffic flows increase their sending rates to use this extra bandwidth.
- Finally, P4e stops at **t10** and this allows it's companion traffic flow (P4i) to increase to ~8 Gbps which is partition P4's maximum top end. The remaining ~2 Gbps is unassigned.



## **Mixed Weights and Subscriptions Example**

This example shows partitions with different relative bandwidth weights and maximum bandwidths, but with the same partitions partially oversubscribing of the available bandwidth as the previous example. The first pair of partitions are set to use 10% or 1 Gbps each of bandwidth and both of their weights are set to 5% while the second pair of partitions are set to use 80% or 8 Gbps of the connection's bandwidth each with both of their relative bandwidth weights set to 45%. The total is still oversubscribing the connection by 80%.

In DCB mode, there would be no difference between the previous DCB mode example and this one, since the Relative Bandwidth Weights are not applicable (in DCB mode) and also if all of the traffic types are in the same Priority Group; the results would be similar.

Port, Partition	Relative Bandwidth Weight (%)	Maximum Bandwidth	Protocol	Plot Color
Port 0, Partition 1 (P1i)	45	80	iSCSI Offload	Green
Port 0, Partition 1 (P1e)	45	80	Ethernet	Orange
Port 0, Partition 2 (P2e)	5	10	Ethernet	Purple
Port 0, Partition 3 (P3e)	5	10	Ethernet	Yellow
Port 0, Partition 4 (P4e)	45	80	Ethernet	Blue
Port 0, Partition 4 (P4i)	45	80	iSCSI Offload	Red

## Table 12: Non-DCB Mixed Fixed Subscription and Oversubscription

This is a combination example of a fixed subscription (three of the partitions sum to 100%), and oversubscription (all four sum to 180%) with different weights and maximum bandwidths. When all four, or at least the two larger partitions, are running traffic, they share the space with each other with respect to their partition's weight and maximum bandwidth values; otherwise, the partition's continue to act as if they are independent connections.

- The first partition's traffic flow (P1i) initially takes its designated ~8 Gbps when the test application starts to transmit traffic at **t0** to that port by itself, not expanding into the remaining unused ~2 Gbps bandwidth.
- When the second traffic flow on the first partition (P1e) starts to send at **t1**, the two active traffic flows on the same partition share its ~8 Gbps bandwidth for ~4 Gbps each.
- When the third traffic flow (P2e) starts sending at **t2**, it only takes its partitions maximum bandwidth allowed ~1 Gbps. Partition P1's two traffic flows are unaffected.
- When the fourth traffic flow (P3e) starts sending at t3, it again only takes its partitions maximum bandwidth allowed ~1 Gbps. Partition P1's two traffic flows (P1i and P1e) and the traffic flow on partition P2 (P2e) are unaffected.
- But when P4e starts to send traffic at **t4**, the traffic needs are oversubscribed so the available bandwidth is redistributed based on each partition's individual weights and maximums settings. P2e and P3e use 5% each (5/100) so their traffic flows are reduced to ~0.5 Gbps which leaves ~9 Gbps free for the other three traffic flows. The two other partition's traffic flows are allocated ~3 Gbps each (9 Gbps / 3 equally weighted traffic flows) the total bandwidth for P1i and P1e is 6 Gbps which is less than partition P1's maximum of 80% of 10 Gbps.
- When P4i starts to send traffic at **t5**, the bandwidth is again redistributed. P2e and P3e are still using only ~0.5 Gbps (5/100). This again leaves ~9 Gbps free for the remaining four equally weighted traffic flows, therefore these four (P1i, P1e, P4e and P4i) all are allocated ~2.25 Gbps each (9 Gbps / 4 flows) where P1i

plus P1e and P4e plus P4i totals are 4.5 Gbps each which is less than their respective partition's maximum bandwidth settings.

- When P1i stops at t6, it releases it's bandwidth to the available pool and since P2e and P3e are capped by their relative bandwidth weight values to 0.5 Gbps, the three other traffic flows (P1e, P4e and P4i) automatically take equal shares of the remaining bandwidth and bump up their portion to ~3 Gbps each where P4e plus P4i total is 6 Gbps which is still less than their respective partition's maximum bandwidth value.
- When P1e subsequently stops sending at **t7**, P4e and P4i grab up some of the extra available bandwidth and go to ~4 Gbps each, where they reach their partition's maximum bandwidth value of 80% or 8 Gbps. The remaining bandwidth is shared equally by P2e and P3e at ~1 Gbps each.
- When P2e stops sending at **t8**, P4e and P4i are not able to make use of the freed up bandwidth since they are both in partition P4 which as a maximum bandwidth ceiling of 8 Gbps. The same is true for P3e which is also at it's bandwidth maximum. Therefore none of the remaining traffic flows increase their sending rates and this unused bandwidth is ignored.
- When P3e stops sending at **t9**, the same maximum ceiling condition is still in effect. Therefore neither P4e or P4i increase their sending rates to use this extra bandwidth.
- Finally, P4e stops at **t10** and this allows it's companion traffic flow (P4i) to increase to ~8 Gbps which is partition P4's maximum top end. The remaining ~2 Gbps is unassigned.

\_

Dell has tested and certified the Broadcom 57712-k Dual-port 10 GbE Converged Network Daughter Card with NIC Partitioning, TOE, iSCSI, and FCoE ready technology. Dell specifically disclaims knowledge of the accuracy, completeness, or substantiation for all statements and claims made in this document regarding the properties, speeds, or qualifications of the adapter.

Broadcom<sup>®</sup> Corporation reserves the right to make changes without further notice to any products or data herein to improve reliability, function, or design.

Information furnished by Broadcom Corporation is believed to be accurate and reliable. However, Broadcom Corporation does not assume any liability arising out of the application or use of this information, nor the application or use of any product or circuit described herein, neither does it convey any license under its patent rights nor the rights of others.

Connecting everything<sup>®</sup>

**BROADCOM CORPORATION** 

5300 California Avenue Irvine, CA 92617 © 2011 by BROADCOM CORPORATION. All rights reserved.

2CS57712-SWUM102-R December 8, 2011



Phone: 949-926-5000 Fax: 949-926-5203 E-mail: info@broadcom.com Web: www.broadcom.com