

The background of the slide features a low-angle, upward-looking shot of a modern, multi-level atrium with a glass and metal spiral staircase. The word "BROCADE" is printed in white, bold, sans-serif capital letters in the upper left corner. A large, stylized red "B" logo with a registered trademark symbol is positioned in the upper right corner.

BROCADE

ネットワーク統合ソリューション

Brocade 1000 10Gbps FCoE CNA, Brocade 8000 FCoE/CEE switch

Brocade 825/815 8Gbps FC HBA QoS 機能

評価レポート

ブロケード コミュニケーションズ システムズ(株)

2009/9/3 – 9/16

Legal Disclaimer

All or some of the products detailed in this presentation may still be under development and certain specifications, including but not limited to, release dates, prices, and product features, may change. The products may not function as intended and a production version of the products may never be released. Even if a production version is released, it may be materially different from the pre-release version discussed in this presentation.

NOTHING IN THIS PRESENTATION SHALL BE DEEMED TO CREATE A WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, STATUTORY OR OTHERWISE, INCLUDING BUT NOT LIMITED TO, ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, OR NONINFRINGEMENT OF THIRD PARTY RIGHTS WITH RESPECT TO ANY PRODUCTS AND SERVICES REFERENCED HEREIN.

Brocade, the B-wing symbol, DCX, Fabric OS, File Lifecycle Manager, MyView, and StorageX are registered trademarks, and DCFM and SAN Health are trademarks of Brocade Communications Systems, Inc., in the United States and/or in other countries. All other brands, products, or service names are or may be trademarks or service marks of, and are used to identify, products or services of their respective owners.



謝辞

Brocade 1020/1010 10Gbps CEE/FCoE CNA, Brocade 8000 CEE/FCoE switch **ならびに** 825/815 8Gbps FC HBA **の** QoS 機能の評価のために Platform Solution Center **の**設備、**ならびに**担当の皆様を適宜アサインしてご協力していただき、ありがとうございました。2週間の検証により、下記のデータを得ることができました。

VMware ESX 3.5 update4 **での** FCoE/CEE 動作

VMware ESX 3.5 update4 **での** FC QoS 動作

ご協力いただきありがとうございました。

ブロケード コミュニケーションズ システムズ (株)



本評価レポートに関する問い合わせ先

ブロード コミュニケーションズ システムズ(株)
パートナー営業本部 第一営業部 鈴木 (suzuki@brocade.com)
03-6203-9100 (代表)
03-6203-9101 (FAX)



BROCADE



Brocade 1000, 8000 評価レポートサマリ

Brocade 1000, 8000 動作検証のサマリ

- 下記の2つの環境において、Brocade 1000 10Gbps FCoE CNA, Brocade 8000 FCoE switch の動作が正常に行えることを検証しました
- サーバ、ストレージ、FCスイッチ
 - 富士通 PRIMERGY RX300-S4
 - 富士通 ETERNUS DX80 V10L11-0000
 - Brocade 300 FOS v6.2.0g (SN200 モデル610 相等)
 - Brocade 8000 FOS v6.1.2_cee1
 - Brocade 1020 driver v2.0.0.0 for VMware ESX 3.5
- OS、ミドルウェア
 - VMware ESX 3.5 update 4
 - 環境
 - Brocade Storage/Network driver for VMware v2.0.0.0
 - ゲストOS: Windows server 2003 R2
 - FCoE接続されたボリュームに対してRDMを行い、ゲストOSからアクセスさせる環境
 - IPv4 で10Gbps Ethernet 接続された環境
- 仮想マシンあたり 1Gbps 以上の IP トラフィックと 4GFC のトラフィック(ETERNUS DX80 4Gbps FC)を同時にハンドリング可能
 - 仮想マシンあたり、1Gbps IP トラフィックと 650MB/sec の FC トラフィックをハンドリング可能だが、ETERNUS DX80 が VMware との組み合わせで 8Gbps FC を未サポート

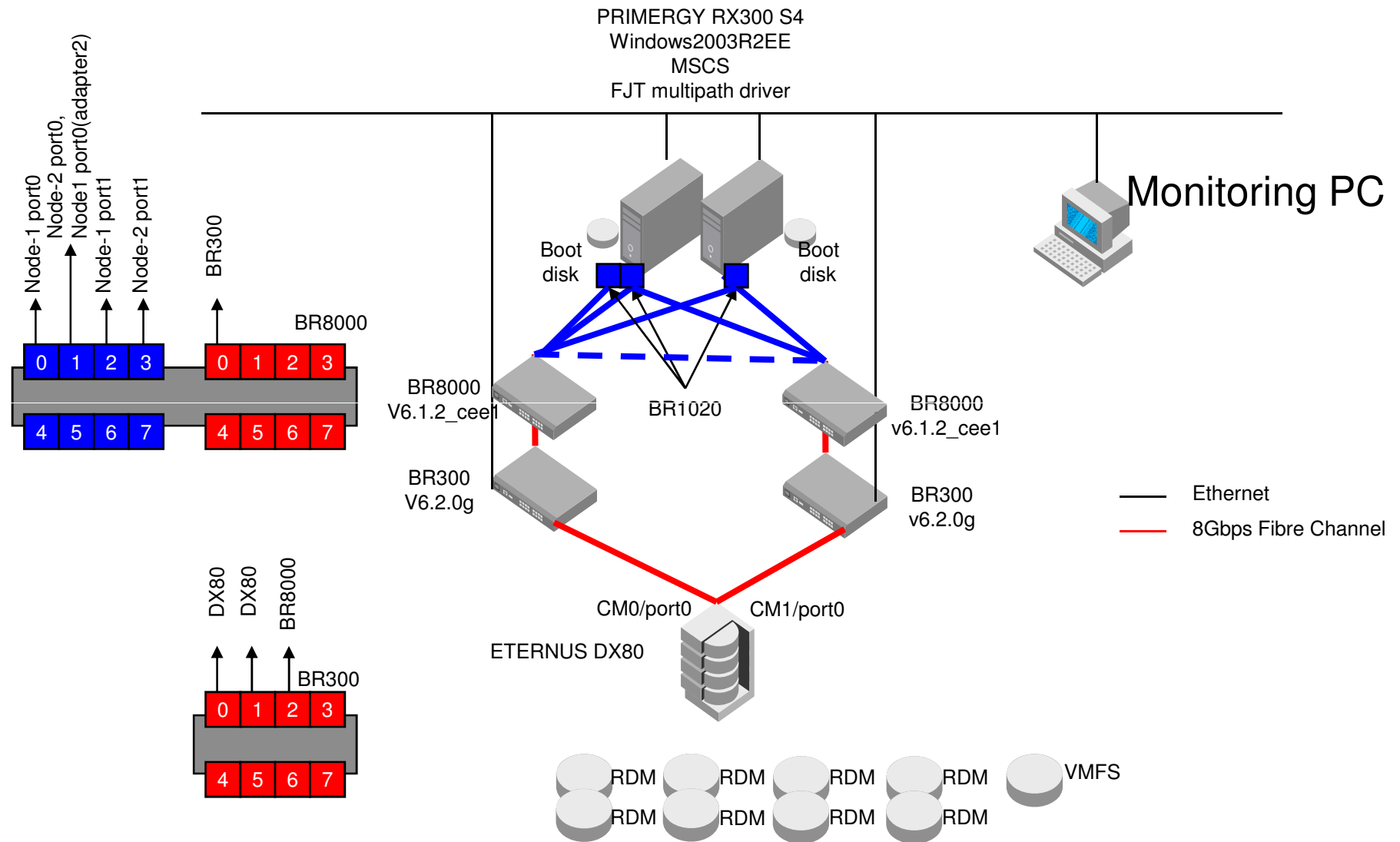


BROCADE



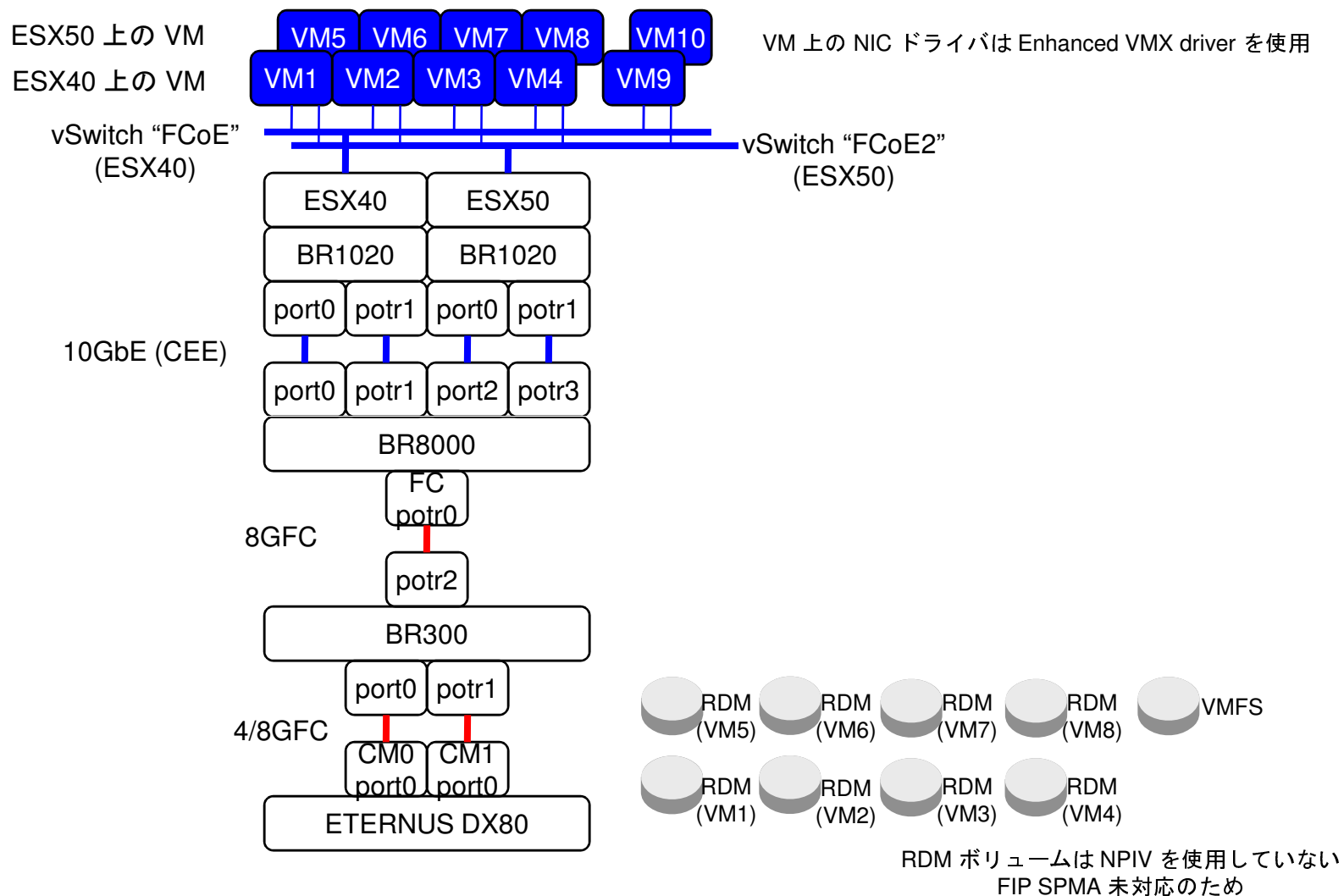
VMware 環境での FCoE 検証内容

物理構成概略



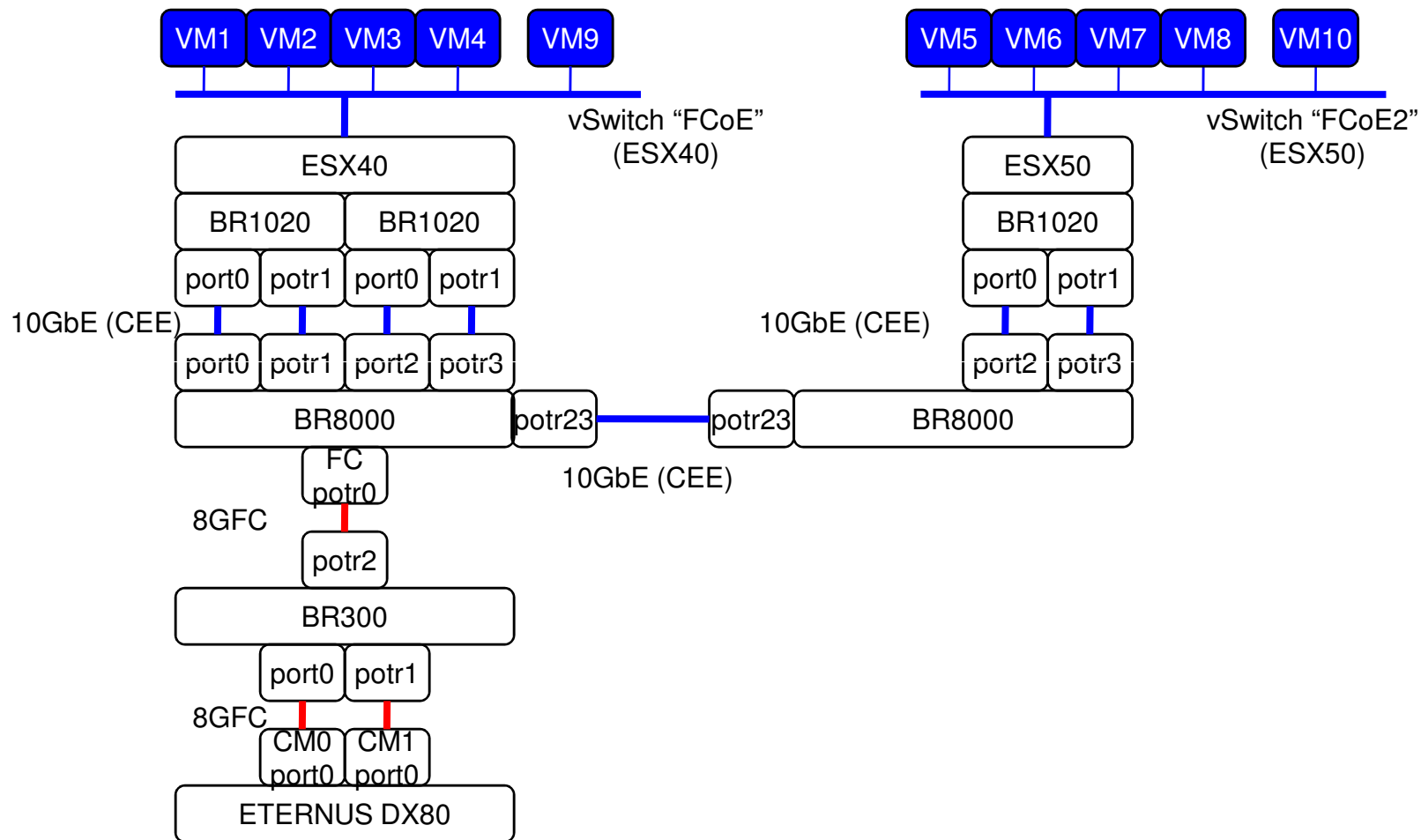
FCoE 検証構成 – 1

FCoE 通信と VM 間 IP 通信

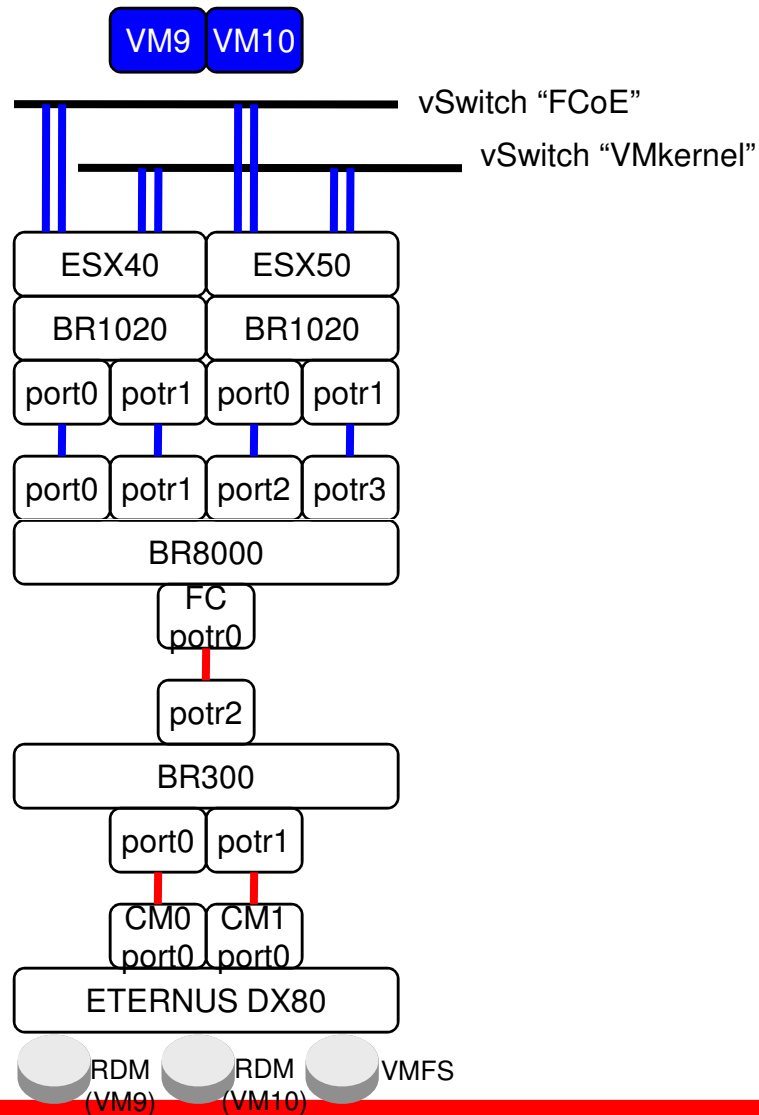


FCoE 検証構成 – 2

FCoE 通信(BR8k直接接続)と VM 間 IP 通信(スイッチカスケード)



FCoE 環境での VMotion 検証構成



RDM ボリュームは NPIV を使用していない
FIP SPMA 未対応のため

CEE環境の設定

• CEE の設定

```
!  
no protocol spanning-tree  
!  
vlan classifier rule 1 proto fcoe encap ethv2  
vlan classifier rule 2 proto fip encap ethv2  
vlan classifier group 1 add rule 1  
vlan classifier group 1 add rule 2  
!  
cee-map default  
priority-group-table 1 weight 40 pfc  
priority-group-table 2 weight 60  
priority-table 2 2 2 1 2 2 2  
!  
interface Vlan 1001  
fcof forward  
!  
protocol lldp  
advertise dcbx-fcoe-app-tlv  
advertise dcbx-fcoe-logical-link-tlv  
!
```

```
interface TenGigabitEthernet 0/0  
switchport  
switchport mode converged  
switchport converged allowed vlan all  
vlan classifier activate group 1 vlan 1001  
cee default  
qos flowcontrol pfc 3 tx on rx on  
no shutdown  
!  
interface TenGigabitEthernet 0/1  
switchport  
switchport mode converged  
switchport converged allowed vlan all  
vlan classifier activate group 1 vlan 1001  
cee default  
qos flowcontrol pfc 3 tx on rx on  
no shutdown  
!  
interface TenGigabitEthernet 0/23  
switchport  
switchport mode converged  
switchport converged allowed vlan all  
vlan classifier activate group 1 vlan 1001  
cee default  
qos flowcontrol pfc 3 tx on rx on  
no shutdown
```



検証内容

- 検証内容

- － 環境構築

- CNA の認識
 - ドライバのインストール
 - 10GbE での IP 疎通確認
 - ボリュームの認識

- － ソフトウェア機能試験

- Brocade Command Line Utility による操作
 - ハードウェア診断機能
 - サポート情報収集機能

- － リンク初期化、エラー回復

- NIC failover
 - FC path failover は HBA 基礎検証時に検証済みなので省略

- － VMotion

- CEE/FCoE 環境下での VMotion

- － 基礎性能

- NTTTCP による IP トラフィックの基礎性能測定
 - IOMeter による FCoE トラフィックの基礎性能測定

検証内容 - 1

Test number	カテゴリ	Test Item	Test 内容	クライテリア	注釈	結果	コメント
1-1	Driver software						
1-1-1		VMware へのドライバソフトウェアのインストール	VMware に Driver v2.0.0.0 をインストールする	エラーなくインストール完了する		○	
1-1-2		Adapter Discovery by VMware によるアダプタの発見	VMware から CNA を発見する(vmhba, vmnic)	CNA を発見できる	vmkload_mod -l cat /proc/vmware/version /var/log/message* , /var/log/vmkernel*, /var/log/vmkwarning*,/proc/vmware/log dmesg	○	

検証内容 - 2

Test number	カテゴリ	Test Item	Test 内容	クライテリア	注釈	結果	コメント
1-2	ネットワーク						
1-2-1		vSwitch の構成	vSwitch を構成できる	構成ができる	vSwitch1 を作成し、ポートグループ FCoE をアサイン (VM1,2,3,4を追加)	○	
1-2-2		仮想マシンに対し、IP を振り、IP 通信が可能	1台のサーバ上に2台の仮想マシンを構成し、ノード間通信をする	IP通信ができる	NTTTCIP(VM1,2間の通信)	○	
1-2-3		仮想マシンに対し、IP を振り、IP 通信が可能	2台のサーバ上に1台ずつの仮想マシンを構成し、ノード間通信をする	IP通信ができる	NTTTCIP(VM1,4間の通信)	○	

検証内容 - 3

Test number	カテゴリ	Test Item	Test 内容	クライテリア	注釈	結果	コメント
1-4	管理ソフトウェア						
1-4-1		HCM software	HCM software agent を起動する	hcmsgentservice をスタートできる		○	
1-4-2			HCM software client をインストールする	Windows 上に HCM をインストールできる		○	
1-4-3			HCM software client を起動する	Windows の HCM を起動し、パラメータを取得する		○	
1-4-4			アダプタパラメータの設定	HCM から queue depth を設定できる	40に設定	○	

検証内容 - 4

Test number	カテゴリ	Test Item	Test 内容	クライテリア	注釈	結果	コメント
1-4-5	管理ソフトウェア	CLI	各種コマンドを実行する	エラー無く実行できる	bcu adapter --list bcu adapter --disable <adapter> bcu adapter --enable <adapter> bcu port --list bcu port --disable <adapter/port> bcu port --enable <adapter/port> bcu lport --list <adapter/port> bcu diag --sfps show <adapter/port> bcu diag --temp show <adapter> bcu diag --pciloopback <adapter/port> pattern bcu diag --loopback <adapter/port> bcu diag --memtest <adapter> bcu diag beacon <adapter/port> on bcu diag beacon <adapter/port> off bcu ioc --stats <port_id> bcu ioc --query <port_id> bcu ethioc --stats <port_id> bcu ethioc --query <port_id> bcu ethport --stats <port_id> bcu cee -- query <port_id> bcu cee --stats <port_id> bcu fcoe --disable <port_id> bcu fcoe --enable <port_id> bfa_supportshow bfa_supportsave <target_directory>	○	

検証内容 - 5

Test number	カテゴリ	Test Item	Test 内容	クライテリア	注釈	結果	コメント
2-1	異常系						
2-1-1		ポートフェイルオーバー	ポート シャットダウンによるポートフェイルオーバー	10G Ethernet interface の shutdown により IP 通信が failover	ポートシャットダウンすると failover し、up すると failback する, interface tengigabitethernet shutdown, no shutdown	○	5秒程度で回復
2-1-2				10G Ethernet interface の shutdown により FC 通信が failover	ポートシャットダウンすると failover し、up すると failback する, interface tengigabitethernet shutdown, no shutdown	○	5秒程度で回復
2-1-3			10連続 port shutdown, up	正常に通信可能	ポートシャットダウンすると failover し、up すると failback する, interface tengigabitethernet shutdown, no shutdown	○	5秒程度で回復
2-1-4			ケーブルの挿抜によるポートフェイルオーバー	10G Ethernet interface の shutdown により IP 通信が failover	ケーブルを抜くと failover し、挿入すると failback する	○	5秒程度で回復
2-1-5				10G Ethernet interface の shutdown により FC 通信が failover	ケーブルを抜くと failover し、挿入すると failback する	○	5秒程度で回復
2-1-6			10連続 ケーブル抜挿	正常に通信可能	ケーブルを抜くと failover し、挿入すると failback する	○	5秒程度で回復
2-1-7			SFP+ の挿抜	10G Ethernet interface の shutdown により IP 通信が failover	SFP+ を抜くと failover し、挿入すると failback する	○	5秒程度で回復
2-1-8				10G Ethernet interface の shutdown により FC 通信が failover	SFP+ を抜くと failover し、挿入すると failback する	○	5秒程度で回復
2-1-9			10連続 SFP+ 抜挿	正常に通信可能	SFP+ を抜くと failover し、挿入すると failback する	○	5秒程度で回復



検証内容 - 6

Test number	カテゴリ	Test Item	Test 内容	クライテリア	注釈	結果	コメント
2-2-1	異常系	アダプタフェイルオーバー	ポート shutdown によるアダプタフェイルオーバー	10G Ethernet interface の shutdown により IP 通信が failover	2枚のアダプタを搭載するように構成を変更, interface tengigabitethernet 0/1 shutdown, no shutdown	○	5秒程度で回復
2-2-2				10G Ethernet interface の shutdown により FC 通信が failover	2枚のアダプタを搭載するように構成を変更, interface tengigabitethernet 0/1 shutdown, no shutdown	○	5秒程度で回復
2-2-3			10連続 ポート shutdown/up	正常に通信可能	2枚のアダプタを搭載するように構成を変更, interface tengigabitethernet 0/1 shutdown, no shutdown	○	5秒程度で回復
2-2-4		アダプタ disable によるアダプタフェイルオーバー		10G Ethernet interface の shutdown により IP 通信が failover	2枚のアダプタを搭載するように構成を変更, bcu adapter --disable 1, bcu adapter --enable 1	○	1秒程度で回復
2-2-5				10G Ethernet interface の shutdown により FC 通信が failover	2枚のアダプタを搭載するように構成を変更, bcu adapter --disable 1, bcu adapter --enable 1	○	1秒程度で回復
2-2-6			10連続 アダプタ disable/enable	正常に通信可能	2枚のアダプタを搭載するように構成を変更, bcu adapter --disable 1, bcu adapter --enable 1	○	1秒程度で回復
2-2-7		ケーブルの挿抜によるアダプタフェイルオーバー		10G Ethernet interface の shutdown により IP 通信が failover	2枚のアダプタを搭載するように構成を変更, ケーブルの挿抜	○	5秒程度で回復
2-2-8				10G Ethernet interface の shutdown により FC 通信が failover	2枚のアダプタを搭載するように構成を変更, ケーブルの挿抜	○	5秒程度で回復
2-2-9			10連続 ケーブル挿抜	正常に通信可能	2枚のアダプタを搭載するように構成を変更, ケーブルの挿抜	○	5秒程度で回復



検証内容 - 7

Test number	カテゴリ	Test Item	Test 内容	クライテリア	注釈	結果	コメント
2-3-1	異常系	FCoE デバイスロス/復旧	スイッチdisable/enable による FC デバイスの復旧	switch enable によりデバイスが復旧する	switchdisable すると IP トラフィックは failover し、switchenable すると failback する、このとき IP トラフィックは流れ続ける	○	FC トラフィックの復旧までは約10秒(ほぼ switchenableにかかる時間)
2-3-2			10連続 スイッチ disable/enable	正常に通信可能	switchdisable すると IP トラフィックは failover し、switchenable すると failback する、このとき IP トラフィックは流れ続ける	○	FC トラフィックの復旧までは約10秒(ほぼ switchenableにかかる時間)
2-3-3			ストレージポート disable/enable による FC デバイスの復旧	DX80 port failover によりデバイスが復旧する	portdisable すると FCoE トラフィックは failover し、portenable すると failback する、このとき IP トラフィックは流れ続ける	○	FC トラフィックの復旧までは1秒以内
2-3-4			10連続 スイッチ disable/enable	正常に通信可能	portdisable すると FCoE トラフィックは failover し、portenable すると failback する、このとき IP トラフィックは流れ続ける	○	FC トラフィックの復旧までは1秒以内

検証内容 - 8

<i>Test number</i>	<i>カテゴリ</i>	<i>Test Item</i>	<i>Test 内容</i>	<i>クライテリア</i>	<i>注釈</i>	<i>結果</i>	<i>コメント</i>
3-1	VMotion over BR8000	FCoE/CEE 環境での VMtion	VMotionを行う	VMkernel ネットワークを BR8000 で構成し、VMotionを行う	I/O, IP トラフィックをかけている状態で VMotion を行い、VMotion が無事完了する	○	30秒程度で完了

1020 CNA の基礎性能

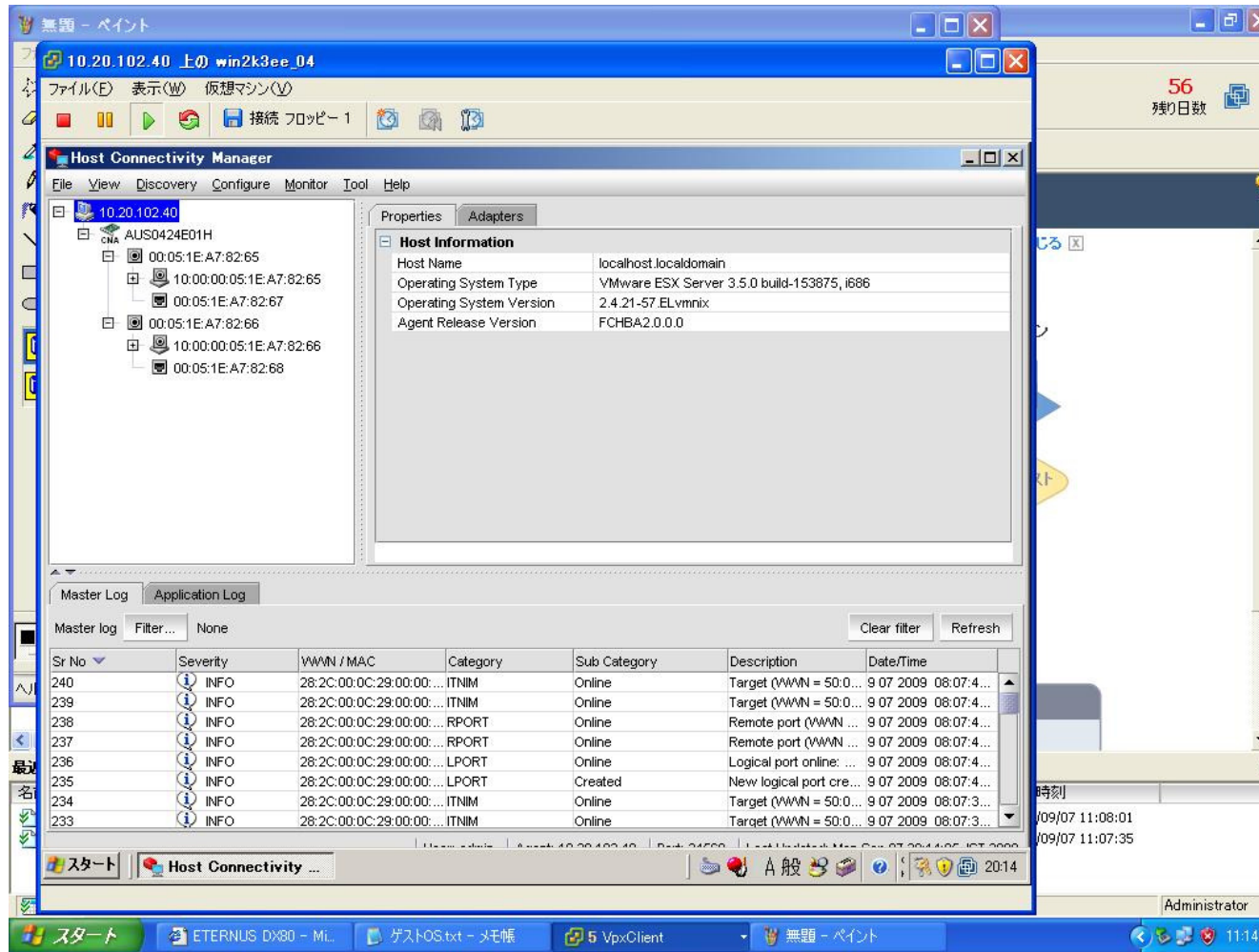
本性能値はチューニングを施していない値です

- Microsoft NTTCIPv3.0 x86 での性能
 - パラメータ
 - ntttps -m1,0,<receiver IP> -a 6
 - ntttpcpr -m1,0,<sender IP> -a 6 -fr
 - 異なる2ESXサーバ筐体間のVM間通信(10ページ参照)
 - 931.098Mbps (VM に対しては 1Gbps NIC として割り当てられる)
- I/O meter による FCoE での性能
 - パラメータ
 - 100% Read Sequential x 8 workers
 - 1VM から ETERNUS DX80 (4Gbps FC) への FCoE 通信 (9ページ参照) (*1)
 - 387.956MB/s (I/Oメータの結果ではなく、スイッチの Advanced Performance Monitoring の結果より)

*1 : 2009/9現在、VMware と ETERNUS DX80 は 8Gbps FC での接続はサポートしていませんので4GFCでの性能を測定



画面キャプチャ抜粋 - 1



Host Connectivity Manager で
CNA をモニタした様子

画面キャプチャ抜粋 - 2

VMotion の前
(vm9 が ESX40 で動作)

VMotion のための
VMkernel ポートの設定

VM 9/10で ntttp,Iometer を
動作させた状態

名前	ターゲット	ステータス	開始者	時刻	開始時刻	完了時刻
仮想マシンを移行します。	win2k3ee_09	完了	Administrator	2009/09/14 14:26:12	2009/09/14 14:26:12	2009/09/14 14:26:43
仮想マシンを移行します。	win2k3ee_09	完了	Administrator	2009/09/14 14:24:40	2009/09/14 14:24:40	2009/09/14 14:25:13
仮想マシンのパワーオン	win2k3ee_10	完了	Administrator	2009/09/14 14:21:47	2009/09/14 14:21:47	2009/09/14 14:22:01
仮想マシンのパワーオン	win2k3ee_09	完了	Administrator	2009/09/14 14:21:44	2009/09/14 14:21:44	2009/09/14 14:22:00

画面キャプチャ抜粋 - 3

The screenshot shows the VMware Infrastructure Client interface. On the left, a tree view lists virtual machines (win2k3ee_01 to win2k3ee_10) under a host named '10.20.102.40'. A red dashed box highlights the 'win2k3ee_09' VM. The main pane shows the 'win2k3ee_09' VM details, including a '仮想マシンについて' (About Virtual Machine) section and a '最近のタスク' (Recent Tasks) table.

仮想マシンについて

仮想マシンは、物理コンピュータと同様に、オペレーティングシステムおよびアプリケーションを実行するソフトウェア コンピュータです。仮想マシン上にインストールしたオペレーティングシステムのことを、ゲスト オペレーティング システム(ゲスト OS)といいます。

仮想マシンはそれぞれ隔離されたコンピュータ環境であるため、それらの仮想マシンを、デスクトップまたはワークステーション環境として、あるいはテスト環境として使用したり、サーバアプリケーションの統合に使用したりできます。

VirtualCenter Server では、仮想マシンはホスト上またはクラスター上で実行されます。同一のホストで多数の仮想マシンを実行できます。

基本タスク

- 仮想マシンのパワーオフ
- 仮想マシンのサスペンド

最近のタスク

名前	ターゲット	ステータス	開始者	時刻	開始時刻	完了時刻
仮想マシンを移行します。	win2k3ee_09	完了	Administrator	2009/09/14 14:01:57	2009/09/14 14:01:57	2009/09/14 14:03:07
仮想マシンを移行します。	win2k3ee_09	完了	Administrator	2009/09/14 13:59:01	2009/09/14 13:59:01	2009/09/14 14:00:13
仮想マシンのパワーオン	win2k3ee_09	完了	Administrator	2009/09/14 13:56:03	2009/09/14 13:56:03	2009/09/14 13:56:17
仮想マシンの再構成	win2k3ee_09	完了	Administrator	2009/09/14 13:55:45	2009/09/14 13:55:45	2009/09/14 13:55:46

On the right, a performance monitoring window shows 'IOmeter' results for 'All Managers' and 'WIN2K3EE10'. The 'Display' tab shows various metrics like Total I/Os per Second, Total MBs per Second, Average I/O Response Time (ms), Maximum I/O Response Time (ms), % CPU Utilization (total), and Total Error Count.

VMotion の後
(vm9 が ESX50 で動作)

VM 9/10で nttcp,IOMeter を
動作させた状態のまま VMotion
が完了

bcu コマンドのログ抜粋

```
root@localhost:~^G[root@localhost root]# bcu port --list
```

Port#	Type	PWWN/MAC	FC Addr/ Eth dev	Media	State	Spd
-------	------	----------	---------------------	-------	-------	-----

1/0	cee	00:05:1e:a7:82:65	--	sw	CEE Linkup	10G
	fcoe	10:00:00:05:1e:a7:82:65	030804		Linkup	
	eth	00:05:1e:a7:82:67	vmnic2		Linkup	
1/1	cee	00:05:1e:a7:82:66	--	sw	CEE Linkup	10G
	fcoe	10:00:00:05:1e:a7:82:66	030802		Linkup	
	eth	00:05:1e:a7:82:68	vmnic3		Linkup	

CEE port としてのリンクアップ

FCoE port としてのリンクアップ

Ethernet port としてのリンクアップ

```
root@localhost:~[root@localhost root]# bcu fcdiag --scsitest 1/0 50:00:00:e0:d0:43:e1:06
```

```
Sending SCSI Inquiry to lun 0...
```

```
SCSI Inquiry completed successfully
```

```
Sending SCSI Read Cap to lun 0...
```

```
SCSI Read Cap completed successfully
```

```
Sending SCSI Report Luns to lun 0...
```

```
SCSI Report Luns completed successfully
```

```
root@localhost:~[root@localhost root]#
```

ETERNUS DX80 CM0 port0 への SCSI inquiry テスト

BROCADE



Brocade 825/815 QoS 機能 評価レポートサマリ

Brocade 825/815 QoS 機能動作検証のサマリ

- 下記の2つの環境において、Brocade 825/815 8Gbps FC HBA のQoS 機能が動作が正常に行えることを検証しました
- サーバ、ストレージ、FCスイッチ
 - 富士通 PRIMERGY RX300-S4
 - 富士通 ETERNUS DX80 V10L11-0000
 - Brocade 300 FOSv6.2.0g (SN200 モデル610 相等)
 - Brocade 825 driver v1.1.0.0 for VMware ESX 3.5
- OS、ミドルウェア
 - VMware ESX 3.5 update 4
 - 環境
 - Brocade Storage driver for VMware v1.1.0.0
 - ゲストOS: Windows server 2003 R2, KB932755
 - Brocade 825 経由で FC 接続されたボリュームを RDM (NPIV 設定)でマウントし、ゲストOSから I/O を行う環境
 - 基礎的動作検証のため、シングルパス構成で行った(*1)
- VMware/ETERNUS DX80 8GFC との接続において 750GB/sec の速度を達成
 - VMware との組み合わせにおいて 8Gbps FC の接続は未サポートのため、あくまで参考値

*1: 複数パスを持つ環境の場合、パス切り替え等が発生すると VMware 側で NPIV の設定を解除してしまい、QoS ポリシの適用ができなかった(NPIV ではないが、I/O そのものは可能)

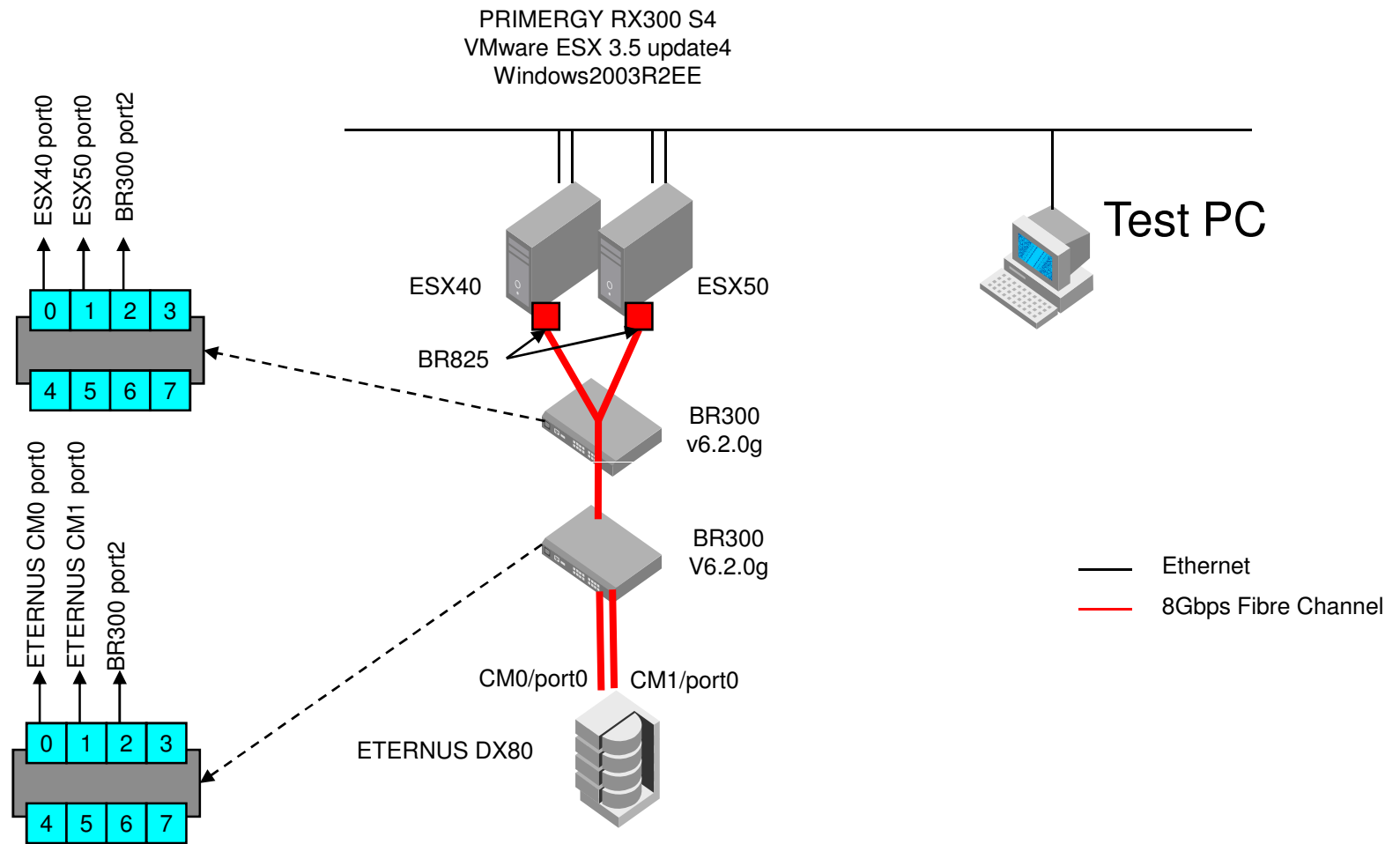


BROCADE

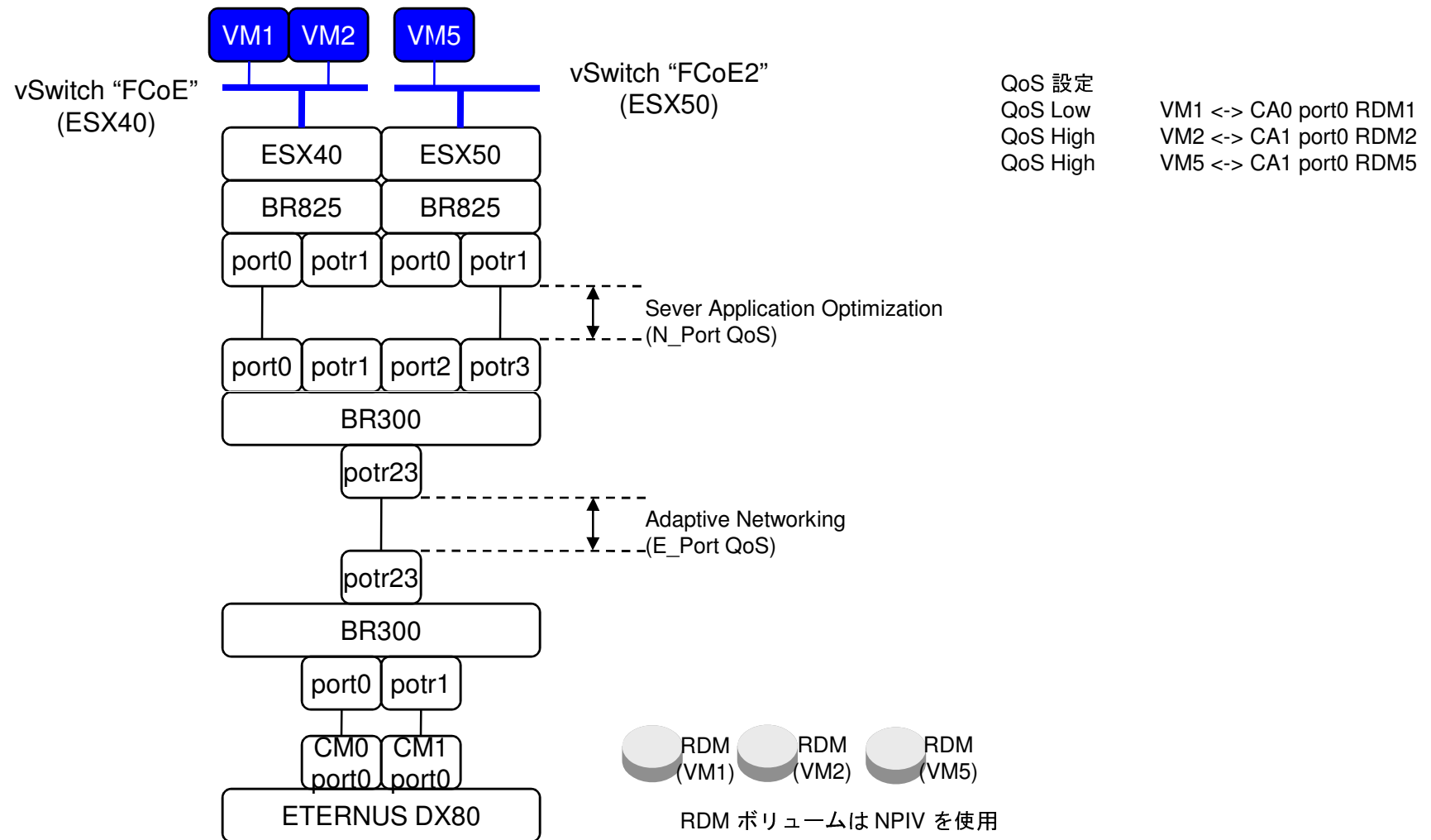


VMware 環境での FC QoS 検証内容

FC QoS 検証構成



FC QoS 検証構成



FC環境の設定

- FC スイッチの設定

- Zone Information

- alicreate "dx80_0", "50:00:00:e0:d0:43:e1:06"
 - alicreate "dx80_1", "50:00:00:e0:d0:43:e1:86"
 - alicreate "esx40_0", "10:00:00:05:1e:a7:82:69"
 - alicreate "esx50_0", "10:00:00:05:1e:a7:82:65"
 - alicreate "vm1npiv", "28:2c:00:0c:29:00:00:e8"
 - alicreate "vm2npiv", "28:2c:00:0c:29:00:00:ed"
 - alicreate "vm5npiv", "28:2c:00:0c:29:00:00:f2"
 - alicreate "vm6npiv", "28:2c:00:0c:29:00:00:f7"
 - zonecreate "e4d0", "dx80_0;esx40_0"
 - zonecreate "e4d1", "dx80_1;esx40_0"
 - zonecreate "e5d0", "dx80_0;esx50_0"
 - zonecreate "e5d1", "dx80_1;esx50_0"
 - zonecreate "QOSL_vm1", "dx80_0; vm1npiv"
 - zonecreate "QOSH_vm2", "dx80_1; vm2npiv"
 - zonecreate "QOSH_vm5", "dx80_1; vm5npiv"
 - zonecreate "QOSH_vm6", "dx80_1; vm6npiv"
 - cfgcreate "qos_1", "e4d0; e5d0; e4d1; e5d1; QOSL_vm1; QOSH_vm2; QOSH_vm5; QOSH_vm6"

- 使用したオプションライセンス

- ISL Trunking
 - Adaptive Networking
 - Server Application Optimization
 - Advanced Performance Monitoring

- HBA の設定

- Queue depth
 - 40 (富士通推奨設定)
 - Speed Negotiation
 - HBA : 8Gbps Auto Negotiation

検証内容

• 検証内容

- FC QoS ゾーンを作成し、トラフィックに対して QoS の優先度を強制できるかどうかを検証
 - Server Application Optimization 機能
 - BR8000 に一台の ESX サーバから BR1020 を1ポートで接続し、FCoE ボリュームにアクセスを行う。このときに QoS が適用されるかどうかを検証
 - Adaptive Networking 機能
 - 2台の ESX サーバから BR1020 で BR8000 に接続し、BR300 を FC でカスケードした状態で、E_Port のトラフィックに QoS が適用されるかどうかを検証

• シナリオ1

- アダプタポートを共有する VM1(QoS low), VM2(QoS high) の I/O を同時に発生させ(100% read)、QoS をかけていないときと比較する

• シナリオ2

- アダプタポートを共有する VM1(QoS low), VM2(QoS high) と、別の物理サーバの VM3(QoS high) から I/O を同時に発生させ(100% read)、QoS をかけていないときと比較する

検証結果 - 1

- シナリオ1
- FC ポートを共有する仮想マシンからI/Oを同時に発生させる
 - QoSL VM1 + QoSH VM2

```
B300-05:root> perfttmon --show 0
```

Src_WWN	Dst_WWN	MB/sec	
28:2c:00:0c:29:00:00:c5	50:00:00:e0:d0:43:e1:86	383.835	<- QoSH のトラフィック
28:2c:00:0c:29:00:00:c0	50:00:00:e0:d0:43:e1:06	285.147	<- QoSL のトラフィック

```
B300-05:root> perfttmon --show 0
```

28:2c:00:0c:29:00:00:c5	50:00:00:e0:d0:43:e1:86	383.835
28:2c:00:0c:29:00:00:c0	50:00:00:e0:d0:43:e1:06	285.147

```
B300-05:root> perfttmon --show 0
```

28:2c:00:0c:29:00:00:c5	50:00:00:e0:d0:43:e1:86	384.891
28:2c:00:0c:29:00:00:c0	50:00:00:e0:d0:43:e1:06	283.739

検証結果 - 2

- シナリオ2
- FC ポートを共有する仮想マシンからI/Oを同時に発生させ、さらに、異なる物理サーバの仮想マシンからも同種の I/O を発生させる(共有する ISL での QoS)
 - QoSL VM1 + QoSH VM2 + QoSH VM5

B300-05:root> perfittmon --show 0 <-- 同じポートを共有 (VM1, VM2)

Src_WWN	Dst_WWN	MB/sec	
28:2c:00:0c:29:00:00:c5	50:00:00:e0:d0:43:e1:86	262.451	<-- QoSH
28:2c:00:0c:29:00:00:c0	50:00:00:e0:d0:43:e1:06	248.925	<-- QoSL

B300-05:root> perfittmon --show 3 <-- 異なる物理サーバ (VM3)

Src_WWN	Dst_WWN	MB/sec	
28:2c:00:0c:29:00:00:d4	50:00:00:e0:d0:43:e1:86	271.511	<-- QoSH
28:2c:00:0c:29:00:00:d9	50:00:00:e0:d0:43:e1:86	0.000	

BROCADE



参考資料



BROCADE

VMware で FCoE 環境を設定する手順

ESX 3.5 向け CNA ドライバインストール手順

1. ESX サーバにドライバパッケージをコピー
 - ftp でも USB メモリ経由でも構わないので、/tmp にドライバパッケージ (brocade_driver_esx35_v2-0-0-0.tar.gz) をコピーし、展開する
 - tar xvf brocade_driver_esx35_v2-0-0-0.tar .gz
2. インストールシェルスクリプトを実行する
 - 展開したディレクトリ中に存在するインストールスクリプトを実行する
 - ./brocade_install.sh
3. ESX サーバをリブート
 - reboot
4. インストールの確認
 - 下記のコマンドの結果に、bfa, bna ドライバが含まれるかを確認
 - vmkload_mod -l
 - cat /proc/vmware/version
5. Host Connectivity Manager サービスの登録と起動
 - esxcfg-firewall -o 34568,tcp,in,https
 - chkconfig --add hcmagentservice
 - /usr/bin/hcmagentservice start

その他の ESX の設定

- 仮想スイッチの作成
 - 仮想マシンをアサインするための仮想スイッチをGUI にて作成
 - vSwitch1
 - Brocade 10G PCIe Ethernet Adapter ポートをアサイン
 - 仮想マシン通信用(FCoE/IP)のポートグループを構成する
 - VMkernel の通信用ポートをアサイン
 - 冗長化されていれば VMotion 時に警告がでない
 - 仮想スイッチポートの MTU サイズを変更(必要に応じて)
 - esxcli network vswitch vSwitch1 -mtu=9000

Brocade 8000 FCoE スイッチの設定 - 1

- 下記の手順は FOS の CEE Management Shell(cmsh) 上で実行するものです

- FOS 上で “cmsh” を実行することで、モードを切り替えます

- cmsh は業界で普及しているネットワーク OS と類似したオペレーションで操作可能です

- Ethernet の設定

```
vlan classifier rule 1 proto fcoe encap ethv2
vlan classifier rule 2 proto fip encap ethv2
vlan classifier group 1 add rule 1
vlan classifier group 1 add rule 2
```

- fcoe, fip プロトコルの Ether-type を group 1 に設定

- DCBX の設定

```
protocol lldp
advertise dcbx-fcoe-app-tlv
advertise dcbx-fcoe-logical-link-tlv
```

- FCoE Application TLV, FCoE Logical Link TLV を advertise するように LLDP を設定(DCBX)

- CEE map の設定

```
cee-map default
priority-group-table 1 weight 40 pfc
priority-group-table 2 weight 60
priority-table 2 2 2 1 2 2 2 2
```

- “default”という名の QoS map(CEEmap)を作成
- PGID 1 は Weight 40 で、PFC 対応の設定、PGID 2 は weight 60 で non-PFC
- Priority 3 を PGID 1 にマップし、残りは PGID 2にマップ

- FCoE Forwarder の設定

```
interface Vlan 1002
fcf forward
```

- VLAN 1002 で FCoE Forwarder を設定(FCoE VLAN)

Brocade 8000 FCoE スイッチの設定 - 2

- インタフェースの設定(使用するポート全てで個別に設定)

```
interface TenGigabitEthernet 0/0
  switchport
  switchport mode converged
  switchport converged allowed vlan all
  vlan classifier activate group 1 vlan 1002
  cee default
  no shutdown
```

- 10GbE ポート0 をスイッチポートに設定し
- FCoE トラフィックも通る設定にしている
- 全ての VLAN のトラフィックを許容
- vlan 1002 に対し group1 のルールを適用
- “default” という名前の CEE map を適用
- インタフェースを起動

- FOS 側の FCoE 設定

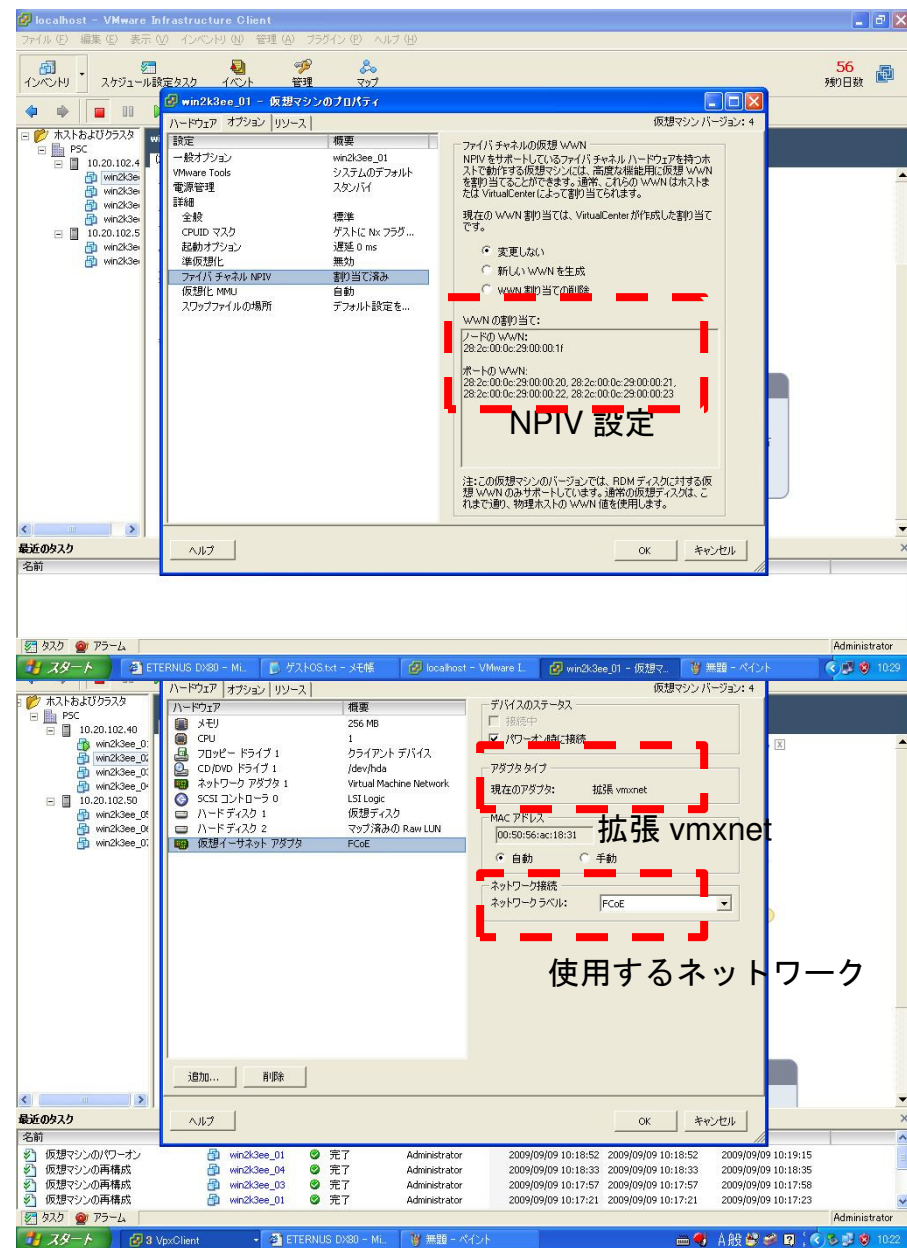
```
fcoe --fcmaset -vlan 1002 0efc00
```

- VLAN 1002 (FCoE を使用する VLAN) の FCMAP (FCoE アドレッシングで使用する値) を 0x0efc00 に設定



仮想マシンの構築

- 仮想マシンは Windows 2003 R2 EE を使用
 - 設定
 - VMotion を検証する VM については SAN 上の VMFS に構築
 - 拡張 vmxnet NIC を追加し、デフォルトの NIC は使用しないか削除
 - 今回はデフォルトの NIC は Console Network のために使用
 - MTU サイズなどは必要に応じて Guest OS 側で設定
 - FCoE 用の vSwitch のポートグループを使用するように設定
 - Raw Device Mapping (RDM) を使用し、SAN ボリュームをマッピング
 - 複数パスがあれば、VMware がパス管理を行う

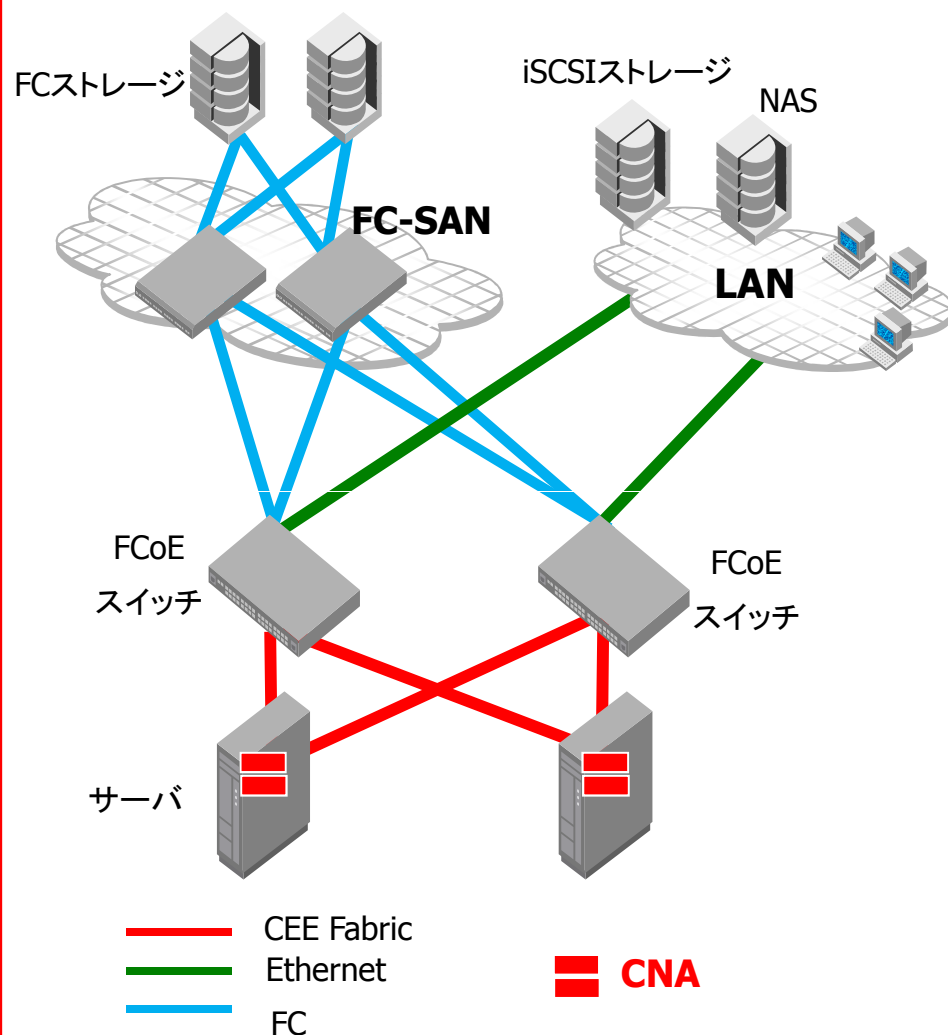


The background of the slide features a low-angle, upward-looking shot of a modern, multi-level atrium with a glass and metal spiral staircase. The word "BROCADE" is written in white, bold, sans-serif capital letters in the upper left corner. A large, stylized red "B" logo with a registered trademark symbol is positioned in the upper right corner.

BROCADE

Brocade 825/815 8Gbps FC HBA
Brocade 1020/1010 10Gbps FCoE CNA
の紹介

ネットワークI/O統合ソリューション CEE/FCoE



• CEEによるIO統合のメリット

ー 物理ネットワーク/インタフェース削減によるコスト低減

- 購入費／保守費の削減、消費電力低減、廃熱量低減、シンプルなケーブリング

ー サーバ省スペース化によるコスト低減

- 物理スペースの節約、ブレードサーバ設計への負荷軽減→価格低減

• FCoEによるメリット

- ー 高速で信頼性のあるFCPを比較的安価なイーサネット上で利用可能
- ー ストレージ管理手法に変化なし
- ー 既存資産の有効活用

サーバ仮想化を FCoE/CEE で構成するメリット

● 課題

- サーバ仮想化のネックはI/O
 - したがって高性能アダプタにより高集約可能
 - VM インスタンスごとに I/O にも QoS がかけられることが重要
- 高集約する際の問題
 - 発熱量、消費電力が低いとだめ
 - 更なる集約に Converged Network 技術が必要 (CEE/FCoE 技術)
- ストレージアダプタは特にミッションクリティカル
 - 保守のために必要な詳細なログ情報が必要
- ポート仮想化技術が重要
 - VM 間通信にはポート仮想化技術が重要

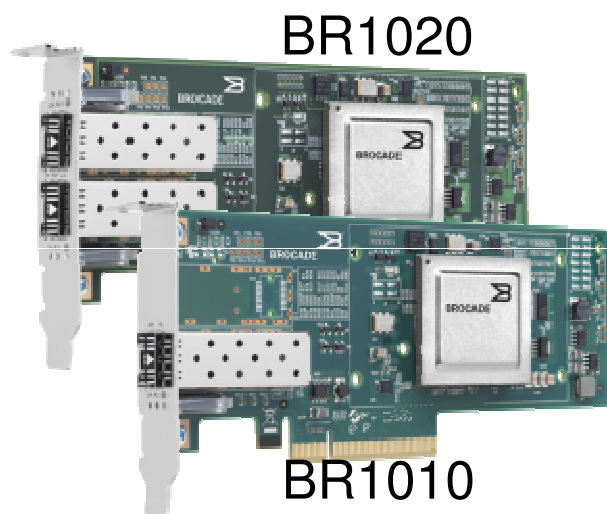
● Brocade ソリューションのメリット

- LAN/SAN スイッチを統合し、所有コスト、消費電力を半減
- Top-of-Rackスイッチからアグリゲーションまでのラインナップで大規模構成が可能
- 低消費電力(11W)かつ低発熱(45 LFM)のアダプタで小規模サーバにもフィット
- FC スイッチ同等のサポート情報取得機能により、障害時の対応がよりスムーズ
- Virtual Ethernet Bridge に対応し(予定)、VM 間通信をアダプタでスイッチすることで、低遅延かつハイパーバイザ負荷を削減

● 製品

- Brocade 8000 FCoE/CEE Switch
- Brocade FCoE10-24 FCoE/CEE blade(Brocade DCX, DCX-4S 用のポートブレード)
- Brocade 1020 10Gbps FCoE/CEE CNA

Brocade 1010/1020 10Gbps FCoE CNA



- 10Gbps Ethernet(CEE) に対応
 - CEE 機能(PFC,ETS,DCBX)に対応
 - FCoE に対応
 - IP トラフィックも FC トラフィックにも対応
- 各種 TCP/IP オフロードが可能
- シングルまたはデュアルポートのモデル
- 低消費電力アダプタ 11W
- SFP+ 光ファイバ, Twinax に対応
- 共通のFC及びFCoE ドライバ
 - FC QoS 等の FC HBA の機能も使用可能
- 統合されたSAN管理

HBAソフトウェアプラットフォーム

- **ドライバ(version2.0)**

- Windows Server 2003/2008 (x86/IA64/IE64T/AMD64) ※MPIO対応、Storportのみ
- RedHat Enterprise Linux (RHEL 4/5) (x86/IA64/IE64T/AMD64)
- SUSE (SLES 9/10) (x86/IA64/IE64T/AMD64)
- Solaris 10 (x86/IE64T/AMD64/SPARC※64ビットのみ)
 - Solaris Ready certification
- VMware ESX 3.5 U2 (x86/IE64T/AMD64)
- Citrix XenServer, OracleVM, Linux KVM

- **管理ツール**

- Brocade Host Connectivity Manager (HCM)
 - GUIベースの管理ツール
 - ローカルおよびリモートでの設定が可能
- BCU CLI

- **管理API**

- SNIA HBA API v2.0
- FDMI
- SMI-S (Version 2.0 release)

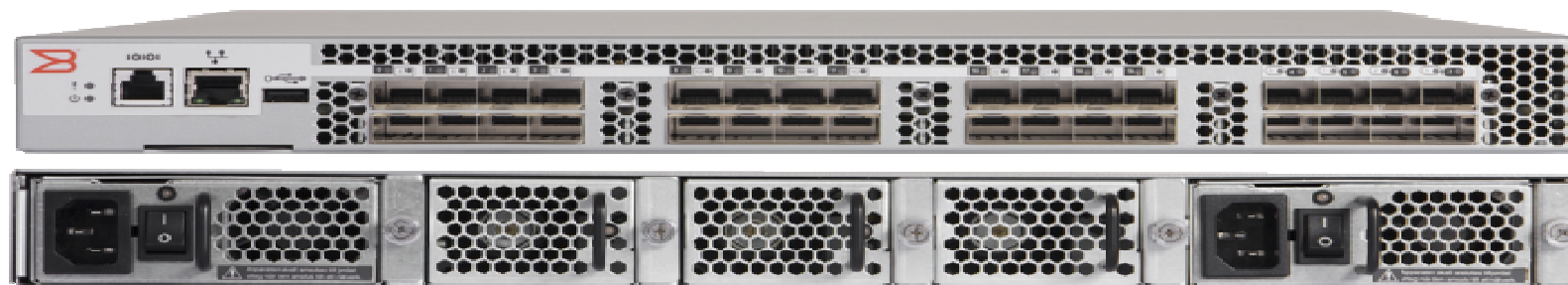


Hyper-V, XenServer 対応

BIOS, UEFI, OpenBoot 対応

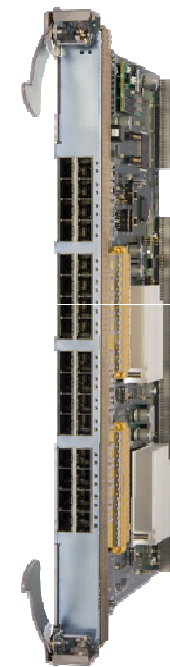
Brocade 8000 とは

- Top-of-Rack のアクセスレイヤスイッチ
- LAN トラフィック、サーバ I/O トラフィックの集約を実現
- FCoE/CEE に対応
- 24 x 10G CEE, 8 x 8Gbps FC のポートを持つ
- 1us 以下の zero over-subscription Ethernet スイッチ

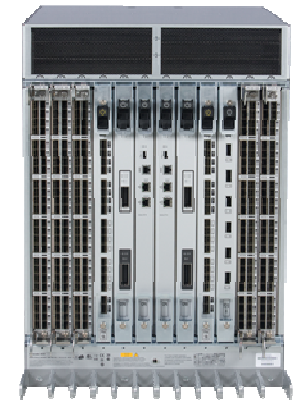


FCOE10-24: CEE/FCoE DCX ブレード

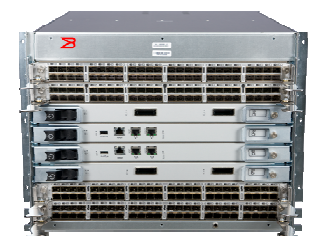
- DCX/DCX-4S 用のマルチプロトコル、マルチレイヤブレード
- End-of-Row スイッチ製品
- 機能概要:
 - 24 ポート 10 GbE CEE
 - Brocade8000 の CEE, FCoE 部分のみをブレード化



FCoE10-24



Brocade DCX
Backbone



Brocade DCX-
4S Backbone

Brocade 8000 のサービス

- Fibre Channel
 - シンプルネームサーバ (SNS)
 - Registered State Change Notification (RSCN)
 - NTP, RADIUS, LDAP, RCS (Reliable Commit Service)
 - Dynamic Path Selection (DPS)
 - Brocade Advanced Zoning
 - Web Tools™
 - オプションライセンス
 - Fabric Watch™; ISL Trunking; End-to-End Performance Monitoring (APM), Enhanced Group Management (EGM)等
- Ethernet
 - Spanning Tree Protocol (STP, MSTP, RSTP)
 - VLAN Tagging (802.1q)
 - MAC address 学習とエージング
 - FCoE Forwarder
 - Link Aggregation (802.3ad, LACP, Brocade)
 - Access Control lists
 - 8 priority levels for QoS(802.1p)
 - 4k VLAN

Brocade 8000 の CEE/FCoE サービス

- FCoE 標準
 - FC-BB-5
 - FIP v1.0.6 (2009/7現在)
 - 標準化前の FIP
- CEE 標準
 - 802.1Qbb Priority-based Flow Control
 - 802.1Qaz Enhanced Transmission Selection
 - 802.1Qaz DCBX Capability exchange protocol
 - IETF TRILL (future)
 - 802.1Qau Congestion Notification (future)



Brocade 8000 の Layer-3 サービス

- L3 IPv4/v6 ルーティング(将来リリースで対応)
 - OSPF, RIP
 - 16K exact match Route (IPv4)
 - 2K longest prefix Route (IPv4)
 - 4K exact/512 prefix (IPv6)
 - VRRP
 - BGP-4
 - IS-IS
 - L3 ACL
 - L3 multicasting



BROCADE

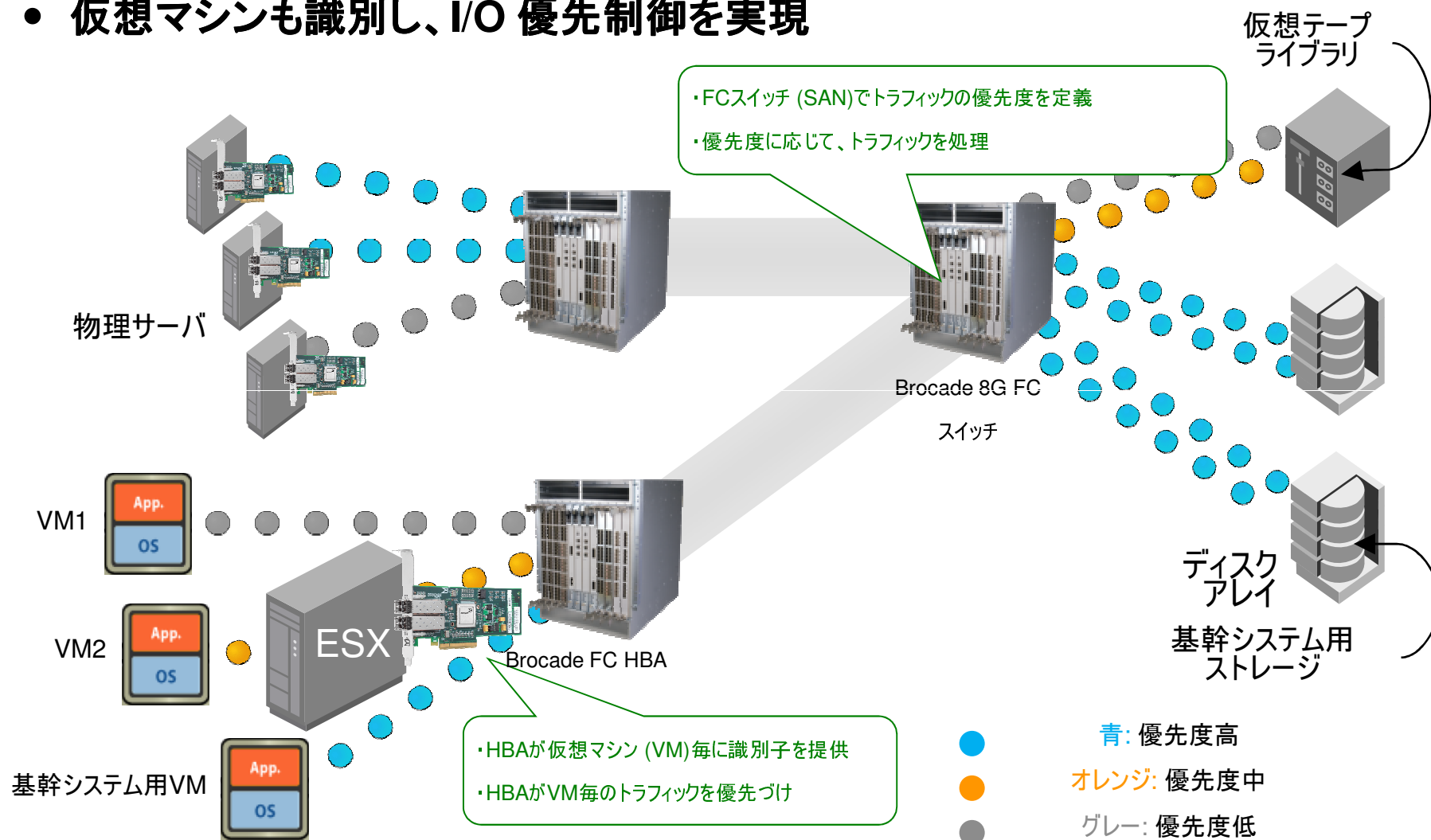
QoS 機能概要

サーバ仮想化環境のスループットを決める要因

- サーバ仮想化環境のスループットを決める要因
 - － サーバCPUコア
 - CPU コアが多ければ、システムスループットは向上する
 - － 物理サーバに搭載されたメモリ量
 - ページングが極力起きないようにメモリを搭載する(もしくは仮想マシンを割り当てる)
 - － 高速 I/O
 - 仮想サーバを複数たてる場合、IOPs 値と帯域が両方とも重要になる
- ブロケードのソリューション
 - － 8Gbps FC に対応(もしくは 10GbE FCoE に対応)し、広帯域I/Oを提供
 - － ポートあたり 50万 IOPs の I/O 性能で、複数の VM の I/O トランザクションに応答可能

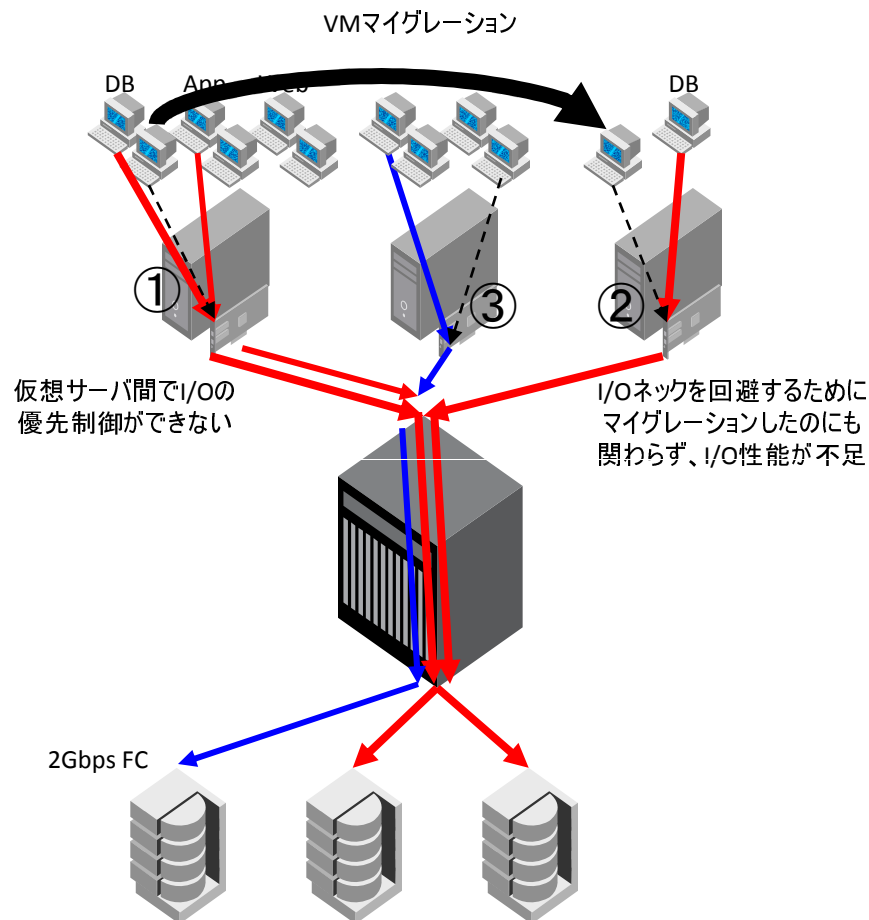
FC QoS の動き

- 仮想マシンも識別し、I/O 優先制御を実現



サーバ仮想化環境におけるボトルネックとソリューション

- サーバ仮想化では I/O 性能が重要
 - サーバ集約数を向上するには、物理メモリ量と IOPs 性能が重要
- 同じ IOPs なら VM インスタンスごとに I/O の優先度を強制できることが重要
 - ミッションクリティカルなアプリケーションとそうでないアプリケーションを同じサーバ上に統合することが困難になる
 - QoS が必要
- I/O の状態を管理ソフトから把握できることが重要
 - VM インスタンスを移動するときにマイグレーション先のサーバの I/O プロファイルを元にマイグレーションできないため、移行先でも性能の保証ができない
- Brocade 825 8Gbps FC HBA は上記をすべて実現できます



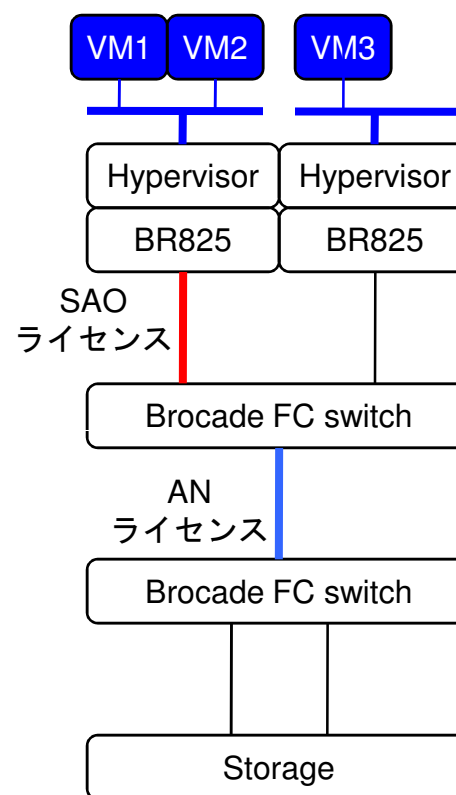
Brocade QoS 機能

1. 仮想サーバトラフィックの優先制御

- アダプタ/スイッチ間のリンクに重畳された仮想サーバ群のトラフィックを pWWN を使い優先制御する機能
- 必要なライセンス、製品
 - Server Application Optimization ライセンス
 - Brocade HBA

2. スイッチ間リンク上のトラフィックの優先制御

- スイッチ間リンクを通るトラフィックを pWWN を使い優先制御する機能
- 必要なライセンス、製品
 - Adaptive Networking ライセンス(カスケードする両側のスイッチに必要)

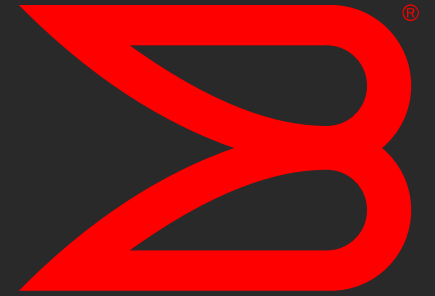


2009/9現在、FC トラフィックにのみQoSを適用できます
(FCoEの場合、ETS によりライセンスなしで実現可能)

FC QoS の原理と設定

- FC QoS は、Fibre Channel の Virtual Channel を使用し、トラフィックの優先制御を実現する技術
 - 優先度に応じた VC 数を割り当てる事で、輻輳が発生する様な環境においてトラフィックの優先制御を実現
 - Brocade FC HBA <-> Brocade FC switch 間で QoS を制御する機能(Server Application Optimization ライセンス)と Brocade FC switch <-> Brocade FC switch 間で QoS を制御する機能(Adaptive Networking ライセンス)の2種類がある
- 設定
 - SAO
 - イニシエータ/ターゲットの port WWN(NPIV の virtual port でもかまわない) を組にして zone を構成する
 - Zone の名前は “QOSx_” で始まる名前にする。x は H/M/L が入る
 - Brocade FC HBA 側で `bcu qos --enable <port_id>` を実行しておく
 - AN
 - `portcfgqos` コマンドで「トラフィックが通過する E_port」の qos を enable にする

BROCADE



THANK YOU