ORACLE

ZFS STORAGE APPLIANCE

Oracleテクニカル・ホワイト・ペーパー
2013年10月

Oracle® ZFS Storage ApplianceとOracle Fabric Interconnectを使用したVMware vSphere 5.1の構成

Oracle® ZFS Storage ApplianceとOracle Fabric Interconnectを使用したVMware vSphere 5.1のIP over InfiniBand構成概要

ORACLE

はじめに 2
システム・コンポーネントの構成例の概要 3
Oracle ZFS Storage Applianceの設定 4
コントローラ、ソフトウェア・リリース、ディスク・プール 4
Oracle ZFS Storage Appliance向けのInfiniBandネットワーク設定 6
VMware用のOracle Virtual Networkingホスト・ドライバのインストール 10
IP over InfiniBand（IPoIB）インタフェースの有効化 11
VMware ESXi 5.1のiSCSI用マルチパス設定に対する推奨事項 16
結論 18
付録A：Mellanox HCAファームウェアの更新 19
付録B：参考資料 23
Oracle ZFS Storage Applianceドキュメント 23

はじめに

本書では、IP over InfiniBandプロトコル（IPoIB）を使用して、Oracle ZFS Storage ApplianceとOracle Fabric Interconnectに対してVMware ESXi 5.1ハイパーバイザを構成する方法を説明します。また、IPoIB環境でのiSCSIおよびファイバ・チャネルLUNの構成向けマルチパスのベスト・プラクティスと推奨事項を説明します。

本書のおもな内容は次のとおりです。

- Oracle ZFS Storage ApplianceとOracle Fabric Interconnectを使用したVMware ESXi 5.1環境の概要
- VMware ESXi 5.1用のOracle Virtual Networkingホスト・ドライバのダウンロード、インストール、有効化の方法
- IP over InfiniBandのVMwareアップリンク・ネットワーク・インタフェースの有効化とVMware仮想スイッチの構成方法
- IPoIB向けのiSCSIおよびファイバ・チャネル（FC）マルチパスのベスト・プラクティスと推奨事項
- VMware ESXi 5.1環境でのMellanox InfiniBandホスト・チャネル・アダプタ（HCA）カードのファームウェア更新方法

注：Sun ZFS Storage Appliance、Sun ZFS Storage 7000、ZFS Storage Applianceという用語はすべて同じOracle ZFS Storage Applianceファミリを指します。

システム・コンポーネントの構成例の概要

次の各表に、本書の構成例で使用したハードウェア構成、オペレーティング・システム、ソフトウェア・リリースを示します。

表1に使用したハードウェアを示します。

表1：リファレンス・アーキテクチャで使用したハードウェア

機器	数量	構成
ストレージ	1クラスタ （2コントローラ）	Oracle ZFS Storage ZS3-2 コントローラあたり256GBのDRAM コントローラあたり1つのデュアルQuad Data Rate（QDR）InfiniBand HCA M2カード 4xDE2-24Cディスク・エンクロージャ（3TB 7200RPM） 2xDE2-24Pディスク・エンクロージャ（900GB 10k RPM） 1xデュアル・ポート10GbEのネットワーク・インタフェース・カード（NIC） 1xデュアル・ポート8GbpsのFCホスト・バス・アダプタ（HBA） 4x73GBのログ・デバイス
ネットワーク	1	Oracle Fabric InterconnectモデルVP780-CH-QDR
サーバー	2	オラクルのSun Blade X6275 M2 10GbEサーバー・モジュール* 98GB DRAM 2x内蔵HDD 1xNEMスイッチ（1Gbps） 1xデュアルHCA Mellanox Technologies MT26428（ConnectX VPI - 10GigE/IB QDR、PCIe 2.0 5GT/秒） – Firmware 2.11.2010

*このサーバー・モジュールはオラクルのSun Blade X3-2Bサーバー・モジュール（旧称Sun Blade X6270 M3）に更新されていますが、新しいサーバー・モジュールは上記リファレンス・アーキテクチャと完全に互換性があります。

表2に使用した仮想マシン・コンポーネントを示します。

表2：リファレンス・アーキテクチャで使用した仮想マシン・コンポーネント

オペレーティング・システム	数量	構成
Oracle Linux 6.3	100	標準の仮想マシン・ゲスト

表3に使用したソフトウェアを示します。

表3：リファレンス・アーキテクチャで使用したソフトウェア

ソフトウェア	バージョン
Oracle® ZFS Storageソフトウェア	OS8
Oracle Virtual Networking Fabric Switch O.S	3.9.0
Vdbench	5.0.3
SWAT	305 Beta 21
VMware vCenter Server	5.1.0（Build 880146）
VMware ESXiハイパーバイザ・ソフトウェア	5.1.0（Build 799733）

Oracle ZFS Storage Applianceの設定

VMware vSphere5.xおよびOracle Fabric Interconnect IPoIB環境では、次のOracle ZFS Storage Appliance構成が使用されました。

コントローラ、ソフトウェア・リリース、ディスク・プール

- Oracle ZFS Storage Appliance ZS3-2クラスタ、アクティブ/アクティブ・モード
- Oracle ZFS Storage OS8ソフトウェア
- Oracle ZFS Storage Applianceコントローラあたり2つの8コア2.10GHzインテル®□Xeon® □CPU E5-2658 0（2.10GHz）
- Oracle ZFS Storage Applianceコントローラあたり256GBのDRAM（ARC）L1キャッシュ
- 1xデュアルQDR InfiniBand HCA M2カード（Oracle Fabric Interconnectに対する4つの物理スタンドアロンInfiniBand接続）
- 84個のデータとパリティ用ディスク、4つのスペア・ディスク、2つのログ・デバイス用SSDで構成されたミラー化ディスク・プール（容量用）。合計111TBの使用可能ディスク領域。容量用ディスク・プールのディスク・レイアウトは図2を参照。
- 38個のデータとパリティ用ディスク、2つのスペア・ディスク、2つのログ・デバイス用SSDで構成されたミラー化ディスク・プール（パフォーマンス用）。合計15TBの使用可能ディスク領域。パフォーマンス用ディスク・プールのディスク・レイアウトは図3を参照。

図1 Oracle ZFS Storage Appliance BUIに表示された容量用ディスク・プールのレイアウト

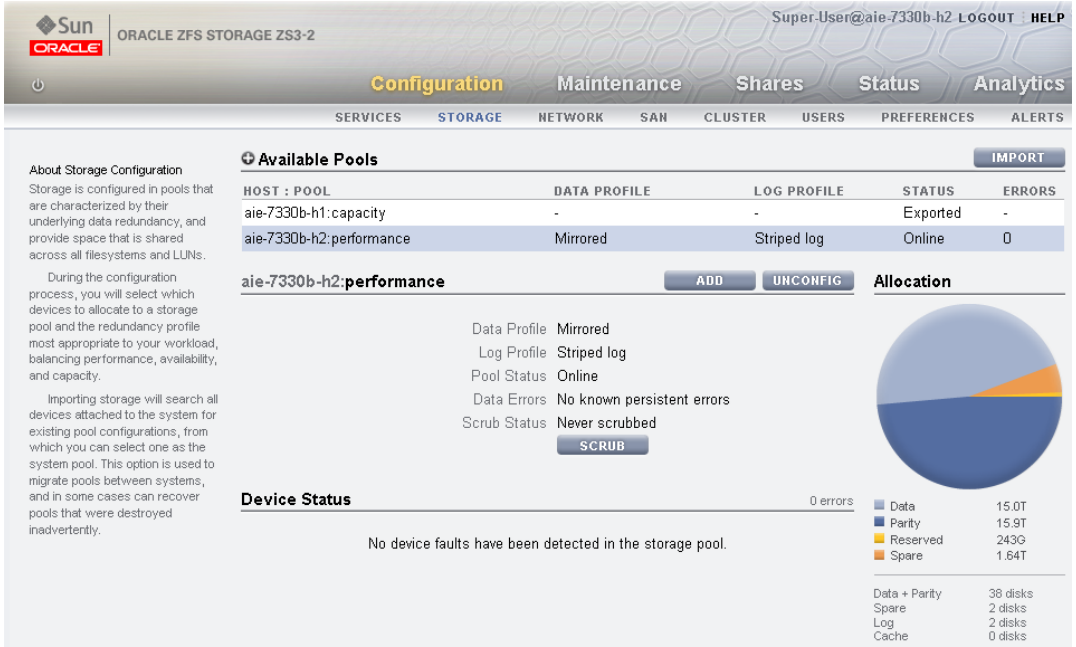


図2 Oracle ZFS Storage Appliance BUIに表示されたパフォーマンス用ディスク・プールのレイアウト

注：Oracle ZFS Storage Applianceのクラスタ構成について、詳しくは『*Sun ZFS Storage 7000システム管理ガイド*』（http://docs.oracle.com/cd/E26765_01/html/E26397/index.html）を参照してください。

Oracle ZFS Storage Appliance向けのInfiniBandネットワーク設定

本書の例では、Oracle ZFS Storage ApplianceとVMware ESXi 5.1ホストの間にIP over InfiniBand（IPoIB）とスタンドアロンのInfiniBandインタフェースを使用しています。この例で使用したOracle ZFS Storage Applianceには4つの物理InfiniBandインタフェースが搭載されており、InfiniBand接続モード・リンク・タイプ（別名CMリンク・モード）を使用してOracle Fabric Interconnectに接続されています。パフォーマンスとスループットを向上するため、それぞれのInfiniBandインタフェースには40Gbpsと4種類のIPアドレスが構成されています。これらはOracle ZFS Storage ApplianceのiSCSIターゲットとしてVMware ESXi 5.1ホストに認識されます。また、Oracle Fabric Interconnectには、VMware ESXi 5.1ホストへの4つの物理InfiniBand接続が構成されており、これらのホストにはMellanox ConnectX-2 HCA InfiniBandカードが搭載されています。

この環境でサポートされるのはスタンドアロンのInfiniBand構成のみであり、Oracle ZFS Storage Appliance向けのIPMP構成はサポートされていません。

Oracle Fabric InterconnectはVMware ESXi 5.1向けのホスト・ドライバを提供しており、このドライバがVMwareカーネル内のInfiniBandモジュールとIPoIBモジュールの有効化を行います。これにより、Oracle ZFS Storage Applianceと連携するための40Gbpsアップリンク・インタフェースを構成できます。VMware ESXi 5.1向けのリリース5.3.1ホスト・ドライバはOracle ZFS Storage Applianceに対して認定されており、FC、iSCSI、NFSプロトコルで使用できます。

注：Oracle ZFS Storage Appliance向けに提供されているOracle Virtual Networkingホスト・ドライバはありません。この環境では、Oracle ZFS Storage Appliance向けのホスト・ドライバは不要です。このシナリオ例では、Oracle Fabric Interconnect、VMware ESXi 5.1ホスト、アプライアンス間の接続は標準のInfiniBandネットワーク・トポロジと同じになります。

Oracle Fabric Interconnectは、仮想ネットワーク・インタフェース・カード（vNIC）と仮想ホスト・バス・アダプタ（vHBA）向けの構成オプションを提供します。これらのvNICとvHBAはその他のNICやHBAとまったく同様に動作し、Oracle Fabric InterconnectのCLI環境に対して、またはOracle Fabric Managerを使用して動的にプロビジョニングが可能です。

Oracle Fabric Interconnectにファームウェアの一部として組み込まれたサブネット・マネージャは、InfiniBandファブリックに接続された各ポートにローカル識別子（LID）を割り当て、割り当てられたLIDに基づいてルーティング表を構成するため、InfiniBandネットワークを管理するために別のサブネット・マネージャやスタンドアロンのサブネット・マネージャをインストールする必要はありません。

次の図に、VMware ESXi 5.1とともにOracle Fabric InterconnectとOracle ZFS Storage Applianceを使用したIP over InfiniBand環境の高水準のアーキテクチャを示します。このアーキテクチャでは、VMware ESXi 5.1ホストとOracle ZFS Storage Appliance間でエンド・ツー・エンドのIPoIBソリューションを使用する方法が複数あります。

注：仮想ネットワーク・インタフェース・カード（vNIC）と仮想ホスト・バス・アダプタ（vHBA）の構成は、本書の対象範囲には含まれません。これらの構成要素について、詳しくはOracle Fabric Interconnectのドキュメント（http://docs.oracle.com/cd/E38500_01/）を参照してください。

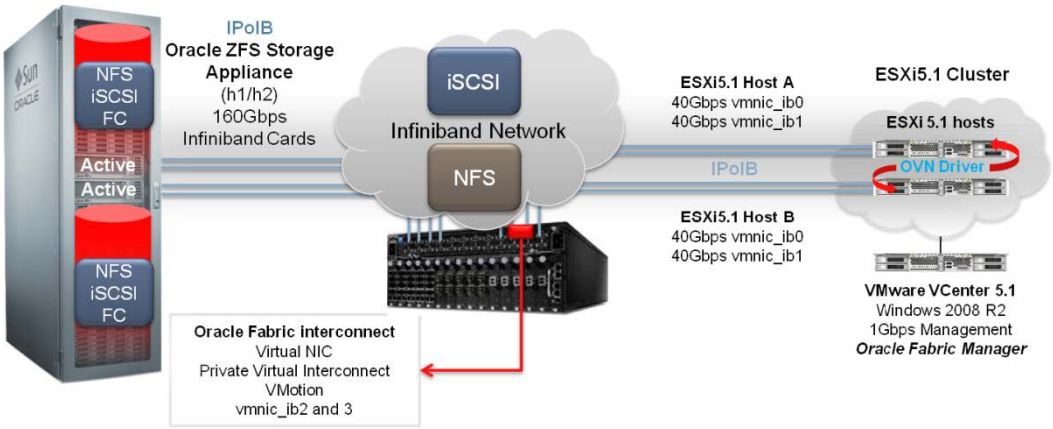


図3 Oracle Fabric InterconnectとOracle ZFS Storage Applianceを使用したVMware ESXi 5.1のIPoIB環境の高水準のアーキテクチャ

Oracle ZFS Storage Applianceには、コントローラごとに1つのデュアルInfiniBandカードが搭載されており、図4、5、6、7に示すとおりに構成されています。

図4 Oracle ZFS Storage Appliance BUIに表示されたInfiniBandインタフェースib0のネットワーク構成

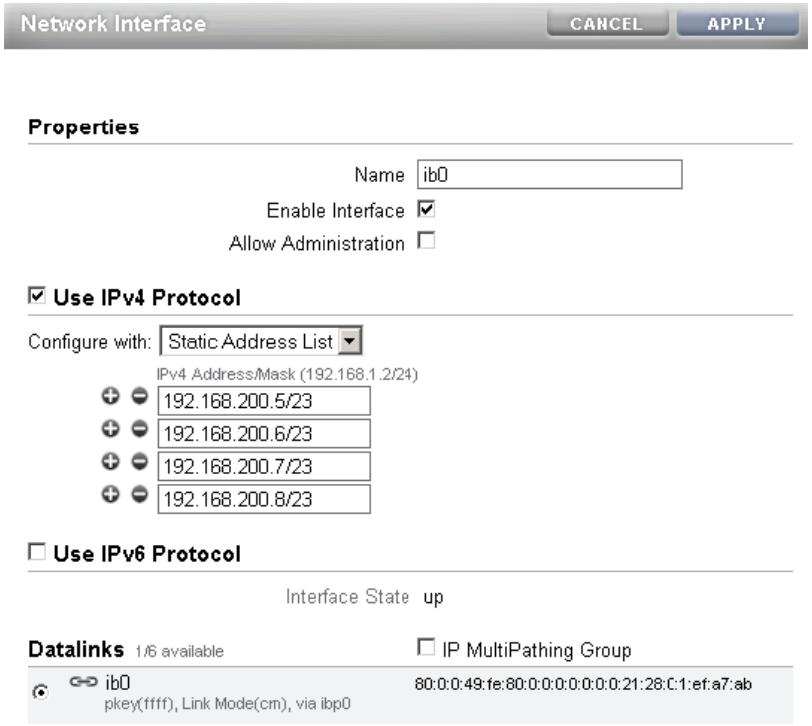


図5 Oracle ZFS Storage Appliance BUIに表示されたInfiniBandインタフェースib0のIPアドレス構成

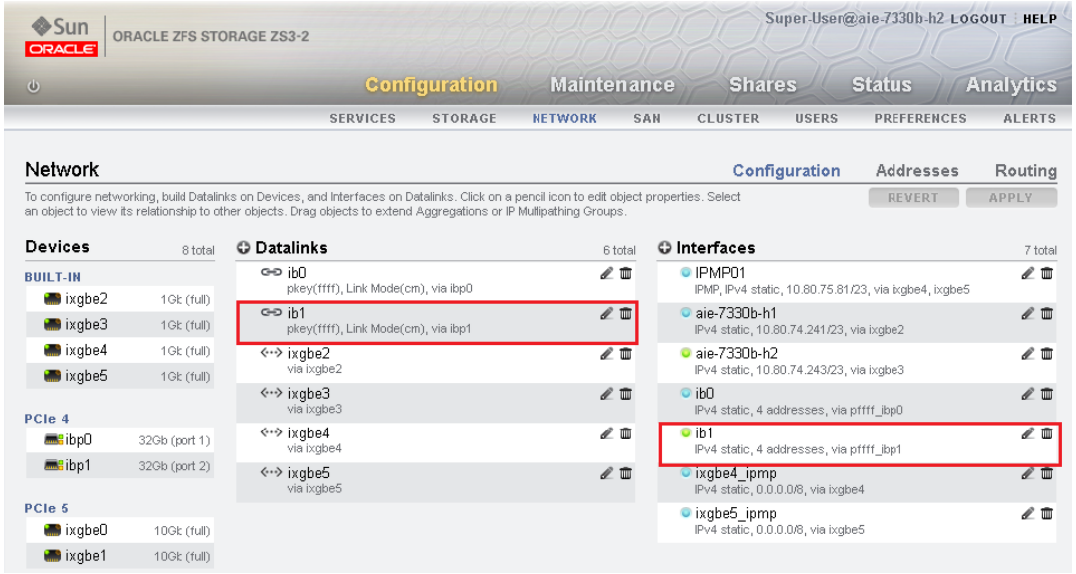


図6 Oracle ZFS Storage Appliance BUIに表示されたInfiniBandインタフェースib1のネットワーク構成

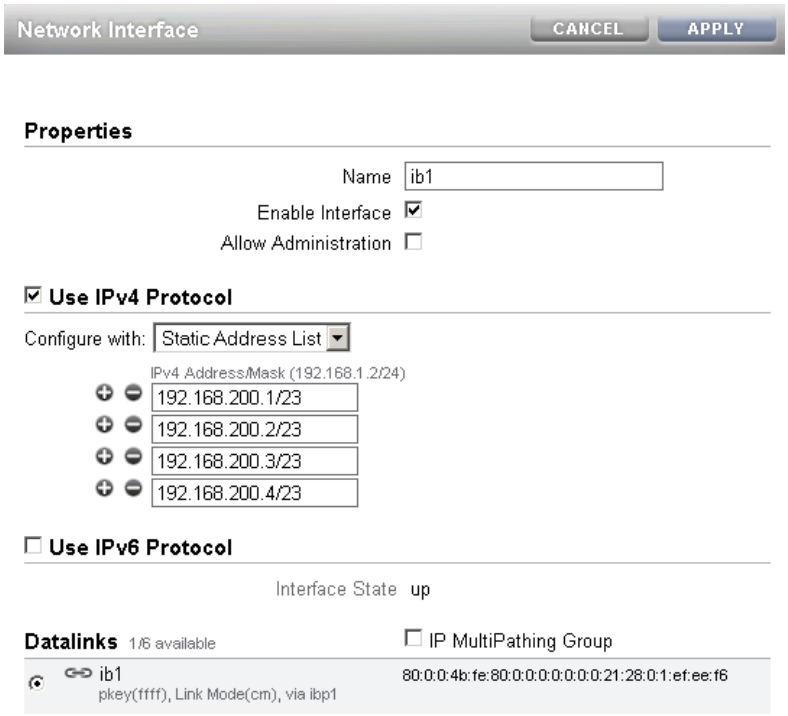


図7 Oracle ZFS Storage Appliance BUIに表示されたInfiniBandインタフェースib1のIPアドレス構成

ソリューションの一部として、Oracle Fabric InterconnectはVMware ESXi 5.1ホスト向けのホスト・ドライバを提供しており、このドライバが、VMwareカーネル内のInfiniBandアップリンク物理インタフェースと、Oracle Fabric Interconnectに対するIPoIB通信を有効化します。

次に、Oracle Fabric Interconnectで`show physical-server`コマンドラインを実行した出力を示します。この例では、ホストaie-6275m2a0およびホストaie-6275m2a1が、ESXi-5.3.1ホスト・ドライバとHCA Firmware 2.11.2010を使用しています。

注：VMware環境でのHCAファームウェアのアップグレード手順とOracle Virtual Networkingドライバのインストール手順は、次の項で説明します。

```
admin@qdr1[xsigo] show physical-server aie-6275m2a0
name            guid                descr       port                os                                  version             server-profile
-----------------------------------------------------------------------------------------------------------------------------------------
aie-6275m2a0    21280001efe456                  qdr1:ServerPort12   VMware/ESXi-5.3.1.ESX.1/x86_64      2.11.2010/3.0.0
aie-6275m2a0    21280001efe455                  qdr1:ServerPort13   VMware/ESXi-5.3.1.ESX.1/x86_64      2.11.2010/3.0.0
2 records displayed
admin@qdr1[xsigo]
```

図8 Oracle Fabric Interconnect CLIに表示された`show physical-server`コマンドラインの出力

Oracle Fabric InterconnectのESXホスト・ドライバについて、詳しくはOracle Fabric Interconnectのドキュメント一覧を参照してください。

http://docs.oracle.com/cd/E38500_01/

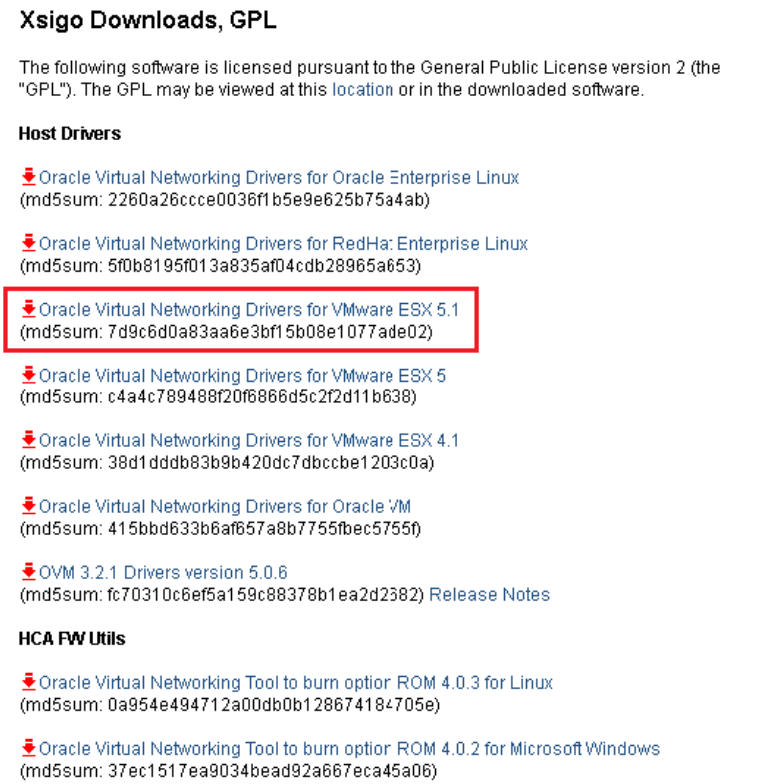
重要：Oracle ZFS Storage ApplianceのIPMP（IPマルチパス・グループ）構成は、この環境ではサポートされません。代わりに、Oracle ZFS Storage ApplianceとOracle Fabric Interconnect間でInfiniBandのスタンドアロン物理ネットワーク構成を使用します。

VMware用のOracle Virtual Networkingホスト・ドライバのインストール

VMware ESXi 5.1ホストにOracle Virtual Networkingホスト・ドライバをインストールするには、次の手順を実行します。

1. 次のリンクからVMware ESXi 5.1用のOracle Virtual Networkingホスト・ドライバをダウンロードします。

   http://www.oracle.com/technetwork/server-storage/xsigo-1870185.html



Xsigo Downloads, GPL

The following software is licensed pursuant to the General Public License version 2 (the "GPL"). The GPL may be viewed at this location or in the downloaded software.

Host Drivers

Oracle Virtual Networking Drivers for Oracle Enterprise Linux
(md5sum: 2260a26ccce0036f1b5e9e625b75a4ab)

Oracle Virtual Networking Drivers for RedHat Enterprise Linux
(md5sum: 5f0b8195f013a835af04cdb28965a653)

Oracle Virtual Networking Drivers for VMware ESX 5.1
(md5sum: 7d9c6d0a83aa6e3bf15b08e1077ade02)

Oracle Virtual Networking Drivers for VMware ESX 5
(md5sum: c4a4c789488f20f6866d5c2f2d11b638)

Oracle Virtual Networking Drivers for VMware ESX 4.1
(md5sum: 38d1dddb83b9b420dc7dbccbe1203c0a)

Oracle Virtual Networking Drivers for Oracle VM
(md5sum: 415bbd633b6af657a8b7755fbec5755f)

OVM 3.2.1 Drivers version 5.0.6
(md5sum: fc70310c6ef5a159c88378b1ea2d2382) Release Notes

HCA FW Utils

Oracle Virtual Networking Tool to burn option ROM 4.0.3 for Linux
(md5sum: 0a954e494712a00db0b128674184705e)

Oracle Virtual Networking Tool to burn option ROM 4.0.2 for Microsoft Windows
(md5sum: 37ec1517ea9034bead92a667eca45a06)

図9 VMware ESXi 5.1用のOracle Virtual Networkingホスト・ドライバのダウンロード・リンク

2. ESXi 5.1ホストにzipファイルをコピーしてインストールします。

```
/tmp # esxcli software vib install -d /tmp/xsigo_5.3.1.ESX.1-1vmw.510.0.0.613838.zip
```

注：必ずzipファイルのフル・パスを指定します。この例では/tmpフォルダを使用しています。

画面出力にはインストールの正常終了を示すメッセージが表示されますが、変更を有効にするにはシステムを再起動する必要があります。

```
   Reboot Required: true
   VIBs Installed: Mellanox_bootbank_net-ib-core_5.3.1.ESX.1-
1OEM.510.0.0.613838, Mellanox_bootbank_net-ib-ipoib_5.3.1.ESX.1-
1OEM.510.0.0.613838, Mellanox_bootbank_net-ib-mad_5.3.1.ESX.1-
1OEM.510.0.0.613838, Mellanox_bootbank_net-ib-sa_5.3.1.ESX.1-
1OEM.510.0.0.613838, Mellanox_bootbank_net-mlx4-core_5.3.1.ESX.1-
1OEM.510.0.0.613838, Mellanox_bootbank_net-mlx4-ib_5.3.1.ESX.1-
1OEM.510.0.0.613838, VMware_bootbank_net-xscore_5.3.1.ESX.1-
1vmw.510.0.0.613838, VMware_bootbank_net-xsvnic_5.3.1.ESX.1-
1vmw.510.0.0.613838, VMware_bootbank_net-xve_5.3.1.ESX.1-1vmw.510.0.0.613838,
VMware_bootbank_scsi-xsvhba_5.3.1.ESX.1-1vmw.510.0.0.613838
   VIBs Removed:
   VIBs Skipped:
```

3. ESXi 5.1ホストを再起動したら、次のコマンドラインを実行してVMware用のOracle Virtual Networkingホスト・ドライバが正しくインストールされていることを確認します。

```
/tmp # esxcli software vib list | grep –i 5.3.1.ESX.1
```

```
/tmp # esxcli software vib list | egrep -i "PartnerSupported|Mellanox"
net-ib-core                    5.3.1.ESX.1-1OEM.510.0.0.613838     Mellanox  PartnerSupported  2013-07-02
net-ib-ipoib                   5.3.1.ESX.1-1OEM.510.0.0.613838     Mellanox  PartnerSupported  2013-07-02
net-ib-mad                     5.3.1.ESX.1-1OEM.510.0.0.613838     Mellanox  PartnerSupported  2013-07-02
net-ib-sa                      5.3.1.ESX.1-1OEM.510.0.0.613838     Mellanox  PartnerSupported  2013-07-02
net-mlx4-core                  5.3.1.ESX.1-1OEM.510.0.0.613838     Mellanox  PartnerSupported  2013-07-02
net-mlx4-ib                    5.3.1.ESX.1-1OEM.510.0.0.613838     Mellanox  PartnerSupported  2013-07-02
net-xscore                     5.3.1.ESX.1-1vmw.510.0.0.613838     VMware    PartnerSupported  2013-07-02
net-xsvnic                     5.3.1.ESX.1-1vmw.510.0.0.613838     VMware    PartnerSupported  2013-07-02
net-xve                        5.3.1.ESX.1-1vmw.510.0.0.613838     VMware    PartnerSupported  2013-07-02
scsi-xsvhba                    5.3.1.ESX.1-1vmw.510.0.0.613838     VMware    PartnerSupported  2013-07-02
/tmp #
```

図10 VMware CLI環境に表示されたesxcli software vib listの出力

IP over InfiniBand（IPoIB）インタフェースの有効化

ESXi 5.1ホスト用のVMwareホスト・ドライバをインストールしたら、総合的なインストール・プロセスの一部として、VMware ESXi 5.1ホストでInfiniBandのIPoIBアップリンク・インタフェースを有効化する必要があります。追加の設定を行う前に、次の簡単な手順を完了しておく必要があります。

1. `root`ユーザーとしてESXi 5.xホストにログインします。
2. `/sbin/xg_ctrl_ipoib enable`コマンドを実行します。
3. ESXi 5.xホストを再起動します。

注：この設定を完了するにはESXi 5.xホストを再起動する必要があります。

IPoIBドライバを無効化するには、disableコマンドを使用し、その後でESXi 5.xホストを再起動します。

```
/sbin/xg_ctrl_ipoib disable
reboot
```

ESXi 5.xホストのCLIで`esxcfg-nics –l`コマンドを実行し、新しいInfiniBandのvmnicが正しく作成されていることを確認します。次の図には、vmnic_ib0とvminc_ib1という2つの新しいvmnicが表示されています。

```
~ # esxcfg-nics -l
Name      PCI           Driver      Link Speed     Duplex MAC Address       MTU    Description
vmnic0    0000:07:00.00 e1000e      Down 0Mbps     Half   00:1e:68:2f:24:9a 1500   Intel Corporation 80003ES2LAN Gigabit Ethernet Controller
vmnic1    0000:07:00.01 e1000e      Up   1000Mbps  Full   00:1e:68:2f:24:9b 1500   Intel Corporation 80003ES2LAN Gigabit Ethernet Controller
vmnic2    0000:0d:00.00 e1000e      Up   1000Mbps  Full   00:c0:dd:0a:b0:2e 1500   Intel Corporation 82571EB Gigabit Ethernet Controller
vmnic3    0000:0d:00.01 e1000e      Up   1000Mbps  Full   00:c0:dd:0a:b0:2f 1500   Intel Corporation 82571EB Gigabit Ethernet Controller
vmnic_ibC 0000:03:00.00 ib_ipoib    Up   40000Mbps Full   00:21:28:a1:47:2d 1500   Mellanox Technologies MT26428 [ConnectX VPI - 10GigE / IB QDR, PCIe 2.0 5G
T/s]
vmnic_ib1 0000:03:00.00 ib_ipoib    Up   40000Mbps Full   00:21:28:a1:47:2e 1500   Mellanox Technologies MT26428 [ConnectX VPI - 10GigE / IB QDR, PCIe 2.0 5G
T/s]
```

図11 VMware CLIでのVMwareのesxcfg-nics –lコマンドの出力

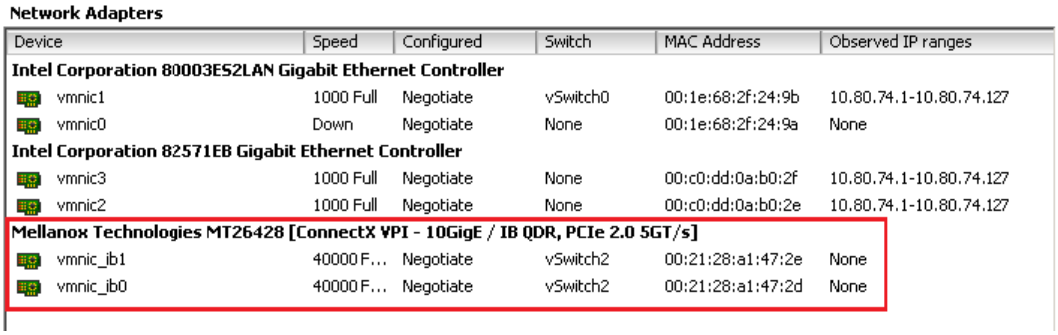


図12 VMware vSphereクライアントのネットワーク・アダプタ・ページに表示されたInfiniBandアップリンク・インタフェース

この時点で、IP over InfiniBandアップリンク・インタフェースのvmnic_ib0とvmnic_ib1に加えて、既存のOracle ZFS Storage ApplianceのiSCSI LUNとVMware ESXi 5.1ホストのNFS共有に対して、新しいVMware仮想スイッチ（vSwitch）を作成できます。次の例では、IPoIBと呼ばれるVMkernelポートを使用してvSwitch01を作成します。vSwitchを構成するには次の手順とスクリーンショット例に従います。

1. VMware vSphereクライアントで、新しい仮想スイッチの構成先としてESXi 5.1ホストを選択します。次の図に示すとおり、「Configuration」タブを選択して、「Networking」と「Add Networking」を選択します。

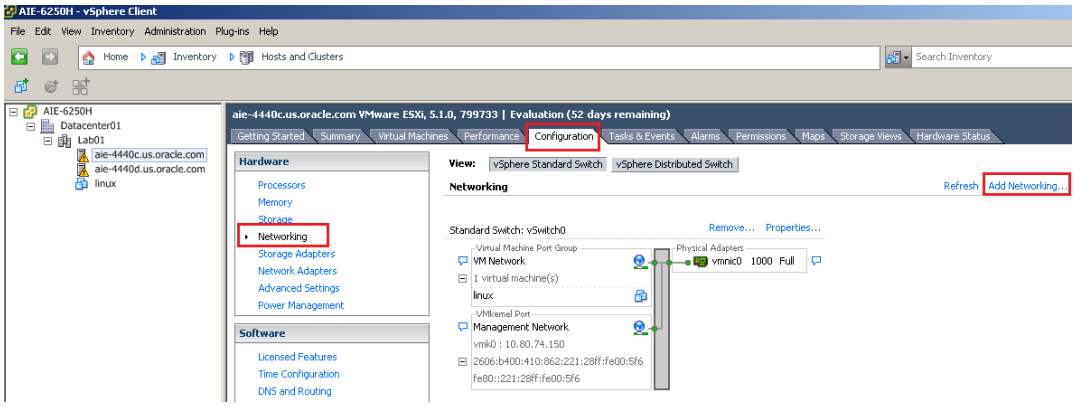


図13 VMware vSphereクライアントでのVMware仮想スイッチの構成

残りの構成手順で使用するウィンドウはすべて、次の図に表示します。

2. 次の画面のConnection Typesボックスで「VMkernel」を選択し、「Next」をクリックします。
3. VMkernel Network Access画面で、新しい仮想スイッチの一部になるネットワーク・アダプタを選択します。この例では、vmnic_ib0とvmnic_ib1という2つのInfiniBandネットワーク・アダプタを選択しています。「Next」をクリックします。
4. Port Group Propertiesセクションでネットワーク・ラベルを入力し、使用する仮想化環境にもっとも適したVLAN IDを選択します。「Use the port group for vMotion」、「Use the port group for fault tolerant logging」、「Use the port group for management traffic」のオプションのチェックは外したままにします。「Next」をクリックします。
5. 次のウィンドウで、新しいネットワーク接続のIPアドレスとサブネット・マスクを入力し、「Next」をクリックします。
6. ネットワーク構成を確認したら（このウィンドウは次の図に含まれません）、「Finish」をクリックして新しいVMware仮想スイッチを作成します。

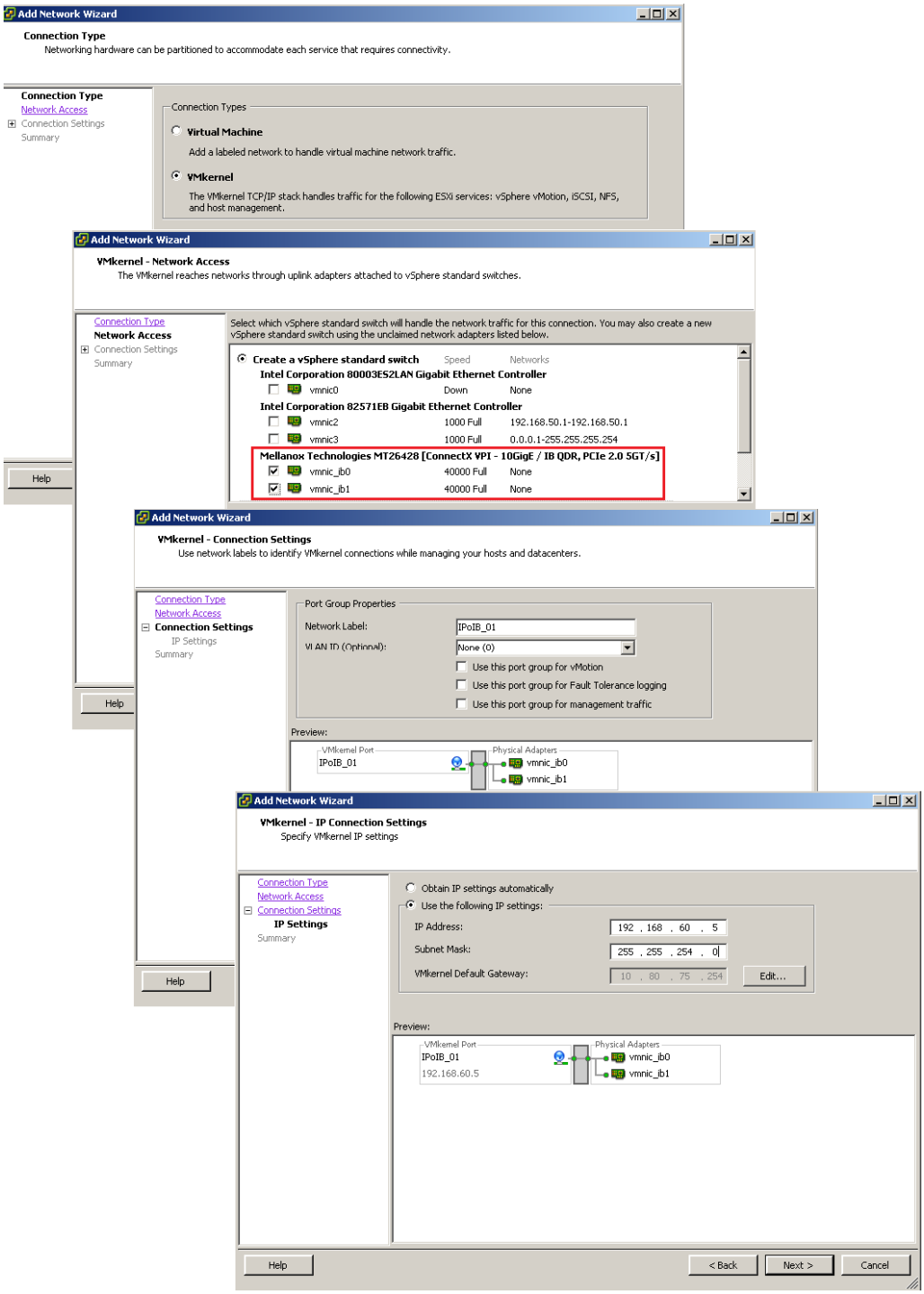


図14 VMware vSphereクライアントでのVMware仮想スイッチの構成手順

仮想スイッチの構成が問題なく完了したら、次の図と手順に示すとおりに、ネットワーク・チーミング、例外ポリシー、ネットワーク・カードのフェイルオーバー順序を設定します。

1. VMware vSphereクライアントで、新しい仮想スイッチの「Properties」をクリックし、Portsタブの構成から「IPoIB」を選択して「Edit」をクリックします。
2. IPoIB Properties画面が表示されます。「NIC Teaming」タブを選択します。Policy Exceptionsセクションで次の項目にチェックを付けます。
   - Load Balancing：Route based on the originating virtual port ID
   - Network Failover Detection：Link status only
   - Notify Switches：Yes
   - Failback：Yes
   - Failover Order：「Override switch failover order」を選択。両方のInfiniBand over IPアップリンク・アダプタをアクティブにします。

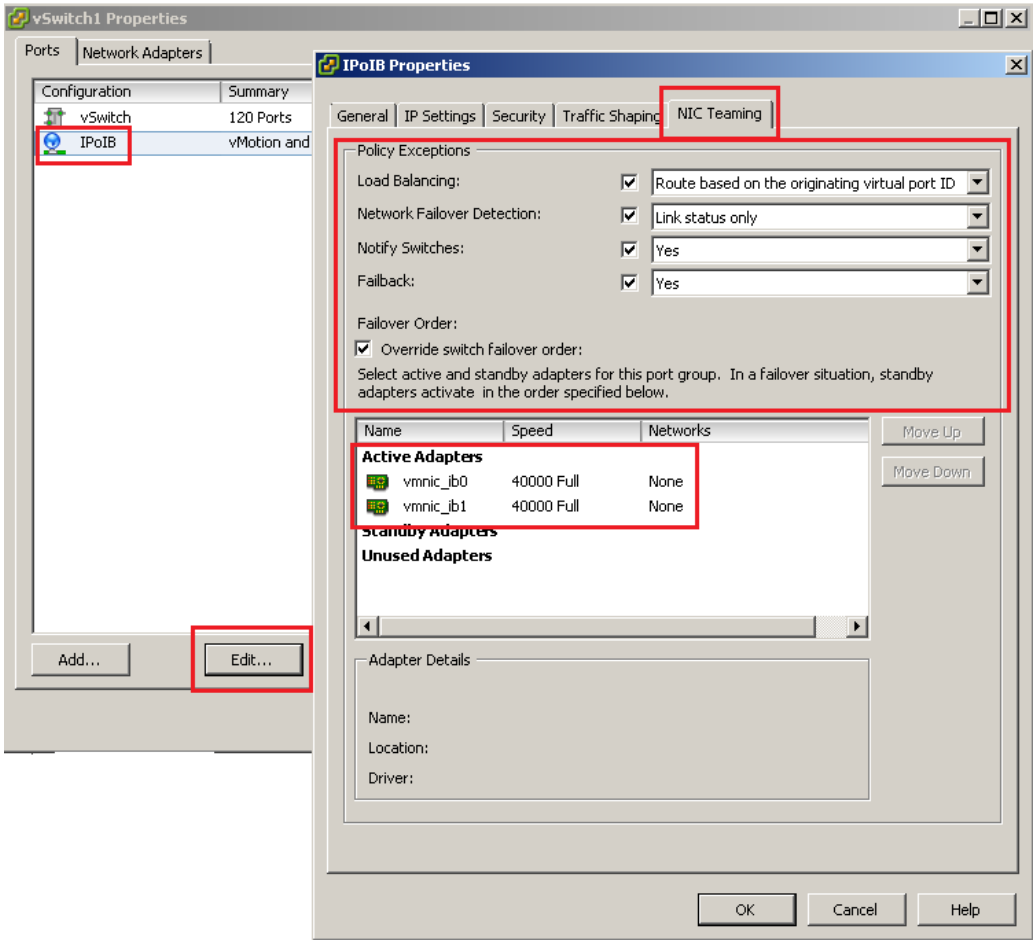


図15 VMware仮想スイッチのプロパティの構成

VMware ESXi 5.1のiSCSI用マルチパス設定に対する推奨事項

Oracle ZFS Storage ApplianceのiSCSI LUNに対してこの環境でサポートされるVMware Storage Array Type Plug-in（SATP）はVMW_SATP_DEFAULT_AAであり、サポートされるVMware Path Selection Policy（PSP）はラウンドロビンのVMW_PSP_RRです。新しいOracle ZFS Storage Appliance LUNを認識またはスキャンする前にマルチパス・ポリシーを追加することが重要です。

次のVMwareマルチパス・ポリシーをVMware ESXi 5.1ホストに追加する必要があります。具体的には、ESXi 5.1ホストに対するSSHセッションを開いて次の手順を実行します。

1. 次のコマンドを使用して、ESXi 5.1ホストにOracle ZFS Storage Appliance向けのポリシーが設定されているかどうかを確認します。コマンドから値が返されない場合、Oracle ZFS Storage ApplianceにはSATPストレージ・ポリシーが設定されておらず、新しいルールを追加できます。

   ```
   esxcli storage nmp satp rule list | grep -i ZFS
   ```

2. Oracle ZFS LUNに対してVMwareマルチパス・ルールを追加します。

   ```
   esxcli storage nmp satp rule add --vendor="SUN" --model="^ZFS*" --
   satp=VMW_SATP_DEFAULT_AA --psp=VMW_PSP_RR
   ```

3. 新しいルールが正常に追加されたことを確認したら、Oracle ZFS Storage ApplianceのiSCSI LUNを表示します。

   ```
   esxcli storage nmp satp rule list | grep -i ZFS
   ```

   ```
   VMW_SATP_DEFAULT_AA                SUN    ^ZFS*  user
   VMW_PSP_RR
   ```

iSCSI LUNを表示した後で、新しいSATPルールによって、すべての新規iSCSI LUNが正しいSATPポリシーとPSPマルチパス・ポリシーの組合せを使用するように指定されていることを確認します。

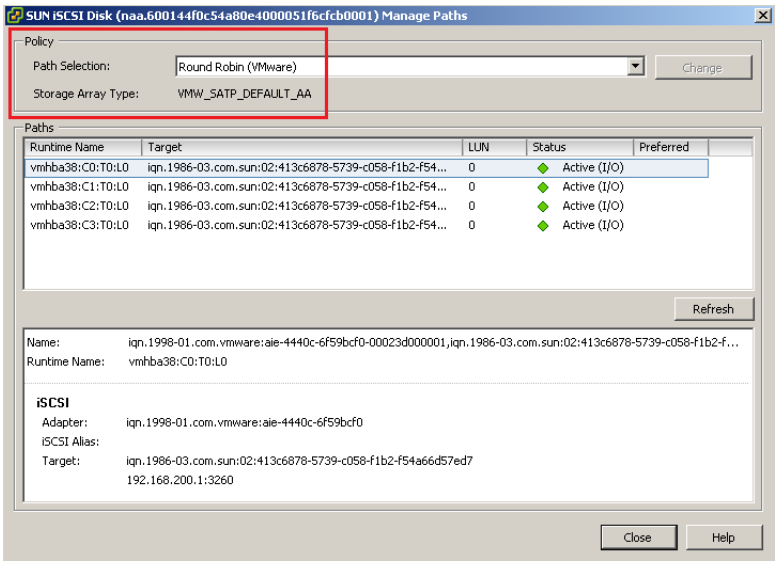


図16 VMware vSphereクライアントに表示されたSCSI LUN向けのVMware PSPおよびSATPマルチパス・ポリシー

重要：前の図に示した構成例はOracle ZFS Storage ApplianceのiSCSI LUNに対してのみ推奨されており、ファイバ・チャネルLUNでは正しく動作しません。上述のルールを追加すると、すべてのOracle ZFS Storage Appliance LUNがVMW_SATP_DEFAULT_AAポリシーとVMW_PSP_RR（ラウンドロビン）マルチパス・ポリシーの組合せを使用するように設定されます。ただし、Oracle Fabric InterconnectはVMware環境での仮想ホスト・バス・アダプタ（vHBA）のプロビジョニング・オプションを複数提供しているため、場合によっては、同じ環境で同時に両方のプロトコル（FCとiSCSI）を使用している可能性があります。これを修正するには、Oracle ZFS Storage Applianceファイバ・チャネルLUN向けのSATPポリシーを変更し、VMW_SATP_ALUAポリシーを適用する必要があります。

この環境では、NFSプロトコルに対する追加構成はありません。同じ40GbpsのIPoIBを使用して、NFSに基づくVMwareデータストアを設定できます。ただし、『How to Get Best Performance From the Oracle ZFS Storage Appliance: Best Practices for Oracle ZFS Storage Appliance and VMware vSphere5.x』で説明されているNFSのベスト・プラクティスや推奨事項、チューニング・オプションを適用していることを確認してください。

http://www.oracle.com/technetwork/articles/servers-storage-admin/sto-recommended-zfs-settings-1951715.html

すべてのiSCSI LUNとファイバ・チャネルLUNに対して、IOPのチューニング・オプションを適用します。これらのオプションは、ダウンロード可能なホワイト・ペーパー版の『Best Practices for Oracle ZFS Storage Appliance and VMware vSphere5.x』にも記載されています。

http://www.oracle.com/technetwork/server-storage/sun-unified-storage/documentation/bestprac-zfssa-vsphere5-1940129.pdf

ファイバ・チャネルLUNのみに対してSATPマルチパス・ポリシーを変更するには、次の手順を実行します。

1. ESXi 5.1ホストに接続されたすべてのファイバ・チャネルLUNを特定し、これらのLUNに対してSATPマルチパス・ポリシーを変更します。次の例では、naa.600144f0d230d6c5000051ffde250001というファイバ・チャネルLUNが特定されています。

   ```
   esxcli storage nmp device list | egrep -i "SUN Fibre Channel Disk" |
   awk '{ print $8 }' | cut -c 2-37
   ```

   ```
   naa.600144f0d230d6c5000051ffde250001
   ```

2. このLUNに対するSATPマルチパス・ポリシーを変更し、ポリシーを有効化します。

   ```
   esxcli storage nmp satp rule add --device=
   naa.600144f0d230d6c5000051ffde250001--satp=VMW_SATP_ALUA --
   psp=VMW_PSP_RR
   ```

3. 新しい構成を再適用します。

```
esxcli storage core claiming reclaim –
device=naa.600144f0d230d6c5000051ffde250001
```

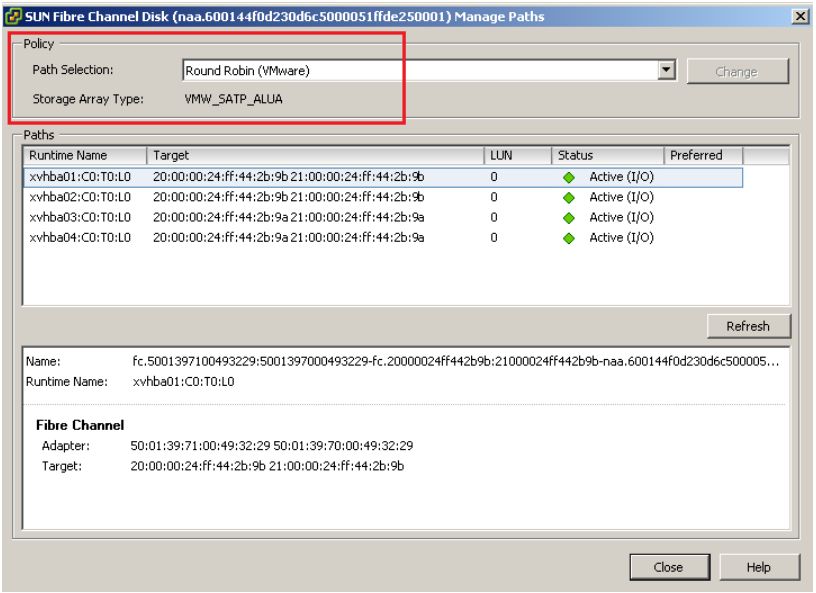


図17 VMware vSphereクライアントに表示されたファイバ・チャネルLUN向けのVMware PSPおよびSATPマルチパス・ポリシー

結論

Oracle Fabric Interconnectのアーキテクチャとネットワーク仮想化向け機能は高速なスループット、低レイテンシ、インテリジェントな管理を提供する目的で開発されており、仮想化環境に最高のパフォーマンスをもたらします。Oracle ZFS Storage Applianceの提供するインテリジェント・キャッシング・テクノロジーは、仮想化環境で数千IOPSを実現し、仮想化されたアプリケーションとデータベースに最高のスループットとレスポンスを実現するように設計されています。Oracle ZFS Storage ApplianceとOracle Fabric InterconnectにVMwareを組み合わせることで、高パフォーマンス、低レイテンシの仮想化環境が実現し、非常に重要な仮想化アプリケーションの要求を満たすことができます。

付録A：Mellanox HCAファームウェアの更新

ここでは、VMware環境でのMellanox ConnectX/ConnectX-2 HCAカードの更新方法について説明します。

重要：InfiniBand HCAのファームウェアを更新する前に、使用するHCAモデルに合った正しいファームウェアであることを確認してください。誤ったファームウェアをインストールすると、HCAが損傷し、使用できなくなる場合があります。カードに合った正しいファームウェアを入手するには、HCAベンダーに問い合わせてください。Oracle InfiniBand HCAを使用している場合、最新のHCAファームウェア・リリースについてはsupport.oracle.comを参照してください。

使用するMellanox ConnectX®またはConnectX®-2のファームウェア・バージョンが2.9.1000以上であり、ConnectX®-3のバージョンが2.10.0000以上であることを確認します。

上記のファームウェア・リリースは、多機能モード（10GbEおよびIB）でConnectXカードを使用するための条件です。詳しくは、MellanoxのWebサイトを参照してください。

また、使用するハードウェア（サーバーまたはホスト）でHCAカードがサポートされていることを確認します。詳しくは、oracle.support.comまでお問い合わせください。

InfiniBand HCAのファームウェアを更新する方法は、"プッシュ"オプションとLinuxオープンソースLive CDの2種類があります。プッシュ・オプションを使用すると、Oracle Fabric Interconnectディレクタを使用してESXiホストにHCAファームウェアをプッシュし、HCAファームウェアを更新することができます。Live CDオプションを使用すると、オープンソースのLinux Live CDディストリビューションからESXiホストを起動し、Oracle Virtual Networkingユーティリティをインストールして HCAファームウェアを更新できます。

"プッシュ"更新を実行するには、次の手順を実行します。

1. ESXi 5.1ホストにOracle Virtual Networkingドライバがインストールされていることを確認します。
2. 現在使用しているHCAカードのファームウェア・リリースを特定します。Oracle Fabric Interconnect CLIで次のコマンドを実行します。

   show physical-server *your_server_name* **hca**

出力結果は次のようになります。HCAカードが現在使用しているファームウェア・リリースはversion列に表示されます。

```
admin@qdr1[xsigo] show physical-server aie-6275m2a0
name          guid              descr     port                os                          version          server-profile
-------------------------------------------------------------------------------------------------------------------------
aie-6275m2a0  21280001efe456                qdr1:ServerPort12   VMware/ESXi-5.3.1.ESX.1/x86_64   2.11.2010/3.0.0
aie-6275m2a0  21280001efe455                qdr1:ServerPort13   VMware/ESXi-5.3.1.ESX.1/x86_64   2.11.2010/3.0.0
2 records displayed
admin@qdr1[xsigo]
```

図18.HCAカードのファームウェア・リリースの特定

3. SSHクライアントを使用して、HCAカードに合った正しいファームウェアをOracle Fabric Interconnectスイッチの/tmpにコピーします。
4. Oracle Fabric Interconnect CLIで次のコマンドラインを実行し、ESXi 5.1ホストに新しいファームウェアをプッシュします。

set physical-server *your_server_name upgrade-hca your_new_firmware_name*

'ライブCD'更新を実行するには、次の手順を実行します。

1. Linuxオープンソース Live CDを使用して、ESXi 5.1ホストを起動します。
2. 新しいLinuxライブ環境のネットワークを設定します。
3. Oracle Virtual Networking Tool to burn option ROM 4.0.3 for LinuxとHCAカードに適したファームウェアをダウンロードします。SSHクライアントを使用して、これらをLinuxライブ環境にコピーします。
4. Oracle Virtual Networking Tool to burn option ROMの.RPMを解凍し、インストールします。CLIセッションは次のようになります。

```
[root@localhost ~]# cd /tmp/
[root@localhost tmp]# ls -ltr
total 21636
-rw-r--r--. 1 root root 21488017 Jul  9 00:24 xsigo-hca-firmware-
4.0.3.r177.XGBOOT-1.i386.zip
-rw-r--r--. 1 root root   661228 Jul  9 12:24 SUN0170000009_fw.img
-rw-------. 1 root root        0 Jul 25 06:27 yum.log

[root@localhost tmp]# unzip xsigo-hca-firmware-4.0.3.r177.XGBOOT-1.i386.zip
Archive:  xsigo-hca-firmware-4.0.3.r177.XGBOOT-1.i386.zip
  inflating: xsigo-hca-firmware-4.0.3.r177.XGBOOT-1.i386.rpm

[root@localhost tmp]# rpm -ivh xsigo-hca-firmware-4.0.3.r177.XGBOOT-
1.i386.rpm
Preparing...                ##########################################
[100%]
   1:xsigo-hca-firmware     ##########################################
[100%]

 [root@localhost tmp]# /opt/xsigo/bin/xg_config

##############################################################################
###
# (C) 2007,2008 XSIGO SYSTEMS Inc. All rights reserved. This material may not
be
# reproduced, displayed, modified or distributed without the express prior
# written permission of the copyright holder.
##############################################################################
###


##############################################################################
###
# Main menu
##############################################################################
###

Selected card:
   Node GUID          : '0021280001efe454'
   Board ID           : 'SUN0170000009'
```

```
   CA type            : 'InfiniBand: Mellanox Technologies MT26428 [ConnectX
VPI PCIe 2.0 5GT/s - IB QDR / 10GigE] (rev b0)'
   Firmware version   : '2.7.8130'
   Hardware version   : 'b0'
   Option ROM version : 'unknown'

1) Flash HCA Firmware
2) Burn Option ROM
3) Erase Option ROM
4) Change selected card
0) Quit
Select option> 1

##############################################################################
###
# Flash HCA Firmware Menu
##############################################################################
###

Selected card:
   Node GUID          : '0021280001efe454'
   Board ID           : 'SUN0170000009'
   CA type            : 'InfiniBand: Mellanox Technologies MT26428 [ConnectX
VPI PCIe 2.0 5GT/s - IB QDR / 10GigE] (rev b0)'
   Firmware version   : '2.7.8130'
   Hardware version   : 'b0'
   Option ROM version : 'unknown'

Press any key to continue
```

注："Found no matching firmware files to use"というエラー・メッセージが表示される場合は、次の手順に従い、mstflintユーティリティを使用して手動でHCAファームウェアを更新します。

1. 同じLinuxライブ環境で、次のコマンドを使用してMellanoxカードを特定します（PCIバスIDの確認）。

   lspci | grep -i "mellanox"

   ```
   00:03:00.0 Serial bus controller:Mellanox Technologies MT26428
   [ConnectX VPI PCIe 2.0 5GT/s - IB QDR / 10GigE]
   ```

2. SSHクライアントを使用してLinuxライブ環境に適切なHCAファームウェアをコピーし、`unzip` ***firmware file name.zip***を実行してファームウェア・イメージを解凍します。mstflintユーティリティを使用してHCAカードのファームウェアを更新します。

3. ファームウェア・ファイルと同じディレクトリで次のコマンドを実行します。

   /opt/xsigo/bin/mstflint –d <*PCI bus ID03:00.0*> -i <*.bin file*> b”.

   [root@localhost tmp]# **/opt/xsigo/bin/mstflint -d 03:00.0 -i fw-ConnectX2-rel-2_11_2010-be-375-3696-01.bin b**

   Current FW version on flash: 2.7.8130

```
   ew FW version:                    2.11.2010
Burning FW image without signatures - OK
Restoring signature                 - OK
```

詳しくは、Mellanoxファームウェア更新ドキュメントとMy Oracle Support（support.oracle.com）を参照してください。

http://www.mellanox.com/page/firmware_HCA_FW_update

mstflintユーティリティの使用法について、詳しくはmstflint FW Burning Tool READMEを参照してください。

4. **ibstat**を実行して新しいファームウェア・バージョンを確認します。
5. Linux Live CDを取り外してシステムを再起動します。ESXi 5.1ホストが再度実行中になったら、Oracle Fabric Interconnect CLIでコマンドライン**show physical-server *your_server_name hca***を使用して、新しいファームウェア・バージョンを確認します。

付録B：参考資料

Oracle ZFS Storage Applianceドキュメント

参考資料	パス
『*Sun ZFS Storage 7000システム管理ガイド*』 『*Sun ZFS Storage 7000 Analyticsガイド*』 『*Sun ZFS Storage 7x20 Appliance設置マニュアル*』 『*Sun ZFS Storage 7x20 Applianceカスタマ・サービス・マニュアル*』	Oracle ZFS Storage Applianceドキュメント・ライブラリ（http://www.oracle.com/technetwork/documentation/oracle-unified-ss-193371.html）
Vmware	http://www.vmware.com
Oracle Fabric Interconnectドキュメント	http://docs.oracle.com/cd/E38500_01/
『Configuring Xsigo Systems I/O Directors on ESX 4.x, ESXi 4.x and ESXi 5.x (Partner Verified and Support) (2005239)』	http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2005239
VMware vSphere 5.1ドキュメント	http://www.vmware.com/support/pubs/vsphere-esxi-vcenter-server-pubs.html
Mellanox	http://www.mellanox.com

Oracle® ZFS Storage Applianceと
Oracle Virtual Networkingを使用した
VMware vSphere 5.xの構成
2013年10月、バージョン1.0
著者：Anderson Souza

Oracle Corporation
World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065
U.S.A.

海外からのお問い合わせ窓口：
電話：+1.650.506.7000
ファクシミリ：+1.650.506.7200

oracle.com

Oracle is committed to developing practices and products that help protect the environment

Copyright © 2013, Oracle and/or its affiliates. All rights reserved.本文書は情報提供のみを目的として提供されており、ここに記載される内容は予告なく変更されることがあります。本文書は一切間違いがないことを保証するものではなく、さらに、口述による明示または法律による黙示を問わず、特定の目的に対する商品性もしくは適合性についての黙示的な保証を含み、いかなる他の保証や条件も提供するものではありません。オラクル社は本文書に関するいかなる法的責任も明確に否認し、本文書によって直接的または間接的に確立される契約義務はないものとします。本文書はオラクル社の書面による許可を前もって得ることなく、いかなる目的のためにも、電子または印刷を含むいかなる形式や手段によっても再作成または送信することはできません。

OracleおよびJavaはOracleおよびその子会社、関連会社の登録商標です。その他の名称はそれぞれの会社の商標です。

IntelおよびIntel Xeonは Intel Corporationの商標または登録商標です。すべてのSPARC商標はライセンスに基づいて使用されるSPARC International, Inc.の商標または登録商標です。AMD、Opteron、AMDロゴおよびAMD Opteronロゴは、Advanced Micro Devicesの商標または登録商標です。UNIXはX/Open Company, Ltd.によってライセンス提供された登録商標です。0611

Hardware and Software, Engineered to Work Together