

Express5800/100 シリーズ InfiniBand と Windows Server 2012 SMB3.0 を利用した仮想基盤

商標について

EXPRESSBUILDER、ESMPRO、EXPRESSSCOPE は日本電気株式会社の商標または登録商標です。

Microsoft、Windows、Windows Server、Hyper-V、Active Directory、MS-DOS は米国 Microsoft Corporation の米国およびその他の国における商標または登録商標です。

Intel、Xeon は、アメリカ合衆国及びその他の国における Intel Corporation、またはその子会社の商標または登録商標です。

InfiniBand は、InfiniBand Trade Association の商標またはサービスマークです。

Mellanox®とそのロゴおよび Connect-X®, Switch-X®, MLNX-OS™などの製品はメラノックステクノロジーズ社のイスラエルおよびその他の国における登録商標または商標です。

StarWind、StarWind Software、the StarWind は米国 StarWind Software の米国およびその他の国における商標または登録商標です。

その他、記載の会社名および商品名は各社の商標または登録商標です。

本書についての注意、補足

- 1 本書の内容の一部または全部を無断転載することは禁じられています。
- 2 本書の内容に関しては将来予告なしに変更することがあります。
- 3 弊社の許可なく複製・改変などを行うことはできません。
- 4 本書は内容について万全を期して作成いたしましたが、万一ご不審な点や誤り、記載もれなどお気づきのことがありましたら、お買い求めの販売店または NEC 営業にご連絡ください。
- 5 運用した結果の影響については 4 項にかかわらず責任を負いかねますのでご了承ください。

目次

商標について	2
本書についての注意、補足	2
1 概要	4
1.1 本書の目的.....	4
1.2 InfiniBand とは.....	4
1.3 略語	4
1.4 InfiniBand 環境の構築.....	4
1.5 ハードウェア接続イメージ	5
2 事前準備	6
2.1 ハードウェアの手配について	6
2.1.1 Express5800 関連	6
2.1.2 InfiniBand スイッチ関連	6
2.2 マニュアルの入手	7
2.3 ハードウェア緒元一覧	8
3 検証環境の事前設定	9
3.1 ホスト OS(Windows Server 2012)のインストール	9
3.1.1 Windows Server 2012 のインストール	9
3.2 InfiniBand の事前設定	10
3.2.1 InfiniBand ドライバのインストール.....	10
3.2.2 InfiniBand スイッチ(サブネットマネージャ)	14
3.2.2.1. 概要.....	14
3.2.2.2. InfiniBand スイッチ(サブネットマネージャ)の設定	14
3.2.3 InfiniBand 接続ボードの設定.....	15
3.3 Hyper-V のインストール	16
4 動作検証	17
4.1 スループット.....	17
4.2 SMB マルチ チャネル.....	20
4.3 透過フェールオーバー.....	21
4.4 Hyper-V over SMB3.0	35
5 注意・制限事項	37
5.1 サーバ	37
5.1.1 最新 BIOS の適用について	37
5.2 RDMA の設定について	37
5.2.1 パフォーマンスカウンター、リソースカウンターについて	37
5.2.2 設定の確認.....	38
5.2.3 電源オプション	39
6 検証環境について.....	40
6.1 ソフトウェア	40
6.2 ハードウェア.....	40

1 概要

1.1 本書の目的

本書は、Express5800/100 シリーズでサポートしている InfiniBand 接続ボードと Windows Server 2012 の SMB 3.0 機能を利用した仮想基盤環境の構築、及び性能評価結果について記したものです。

本書では、InfiniBand 接続ボードと Windows Server 2012 の SMB 3.0 機能を利用した仮想基盤環境の構築手順の明確化と、構築した仮想基盤環境での性能評価を目的としています。

なお、本書は性能/可用性を保証するものではありません。構築時には必ずシステム要件設計に基づいた、性能/可用性設計を行い、適切なシステムテストを実施するようにお願いします。

1.2 InfiniBand とは

InfiniBand とは、米 IBTA(InfiniBand Trade Association)により策定された、オープンスタンダードな I/O 規格です。サーバのクラスタリングや外部ストレージの I/O など高バンド幅、低レイテンシー、高信頼性が求められる分野においても活用されています。

FDR 4x では 56Gbps(実レートは 54Gbps)の全二重通信が可能です。また、Remote Direct Memory Access (RDMA) と呼ばれるサーバ間での CPU を介さないメモリ間のデータ転送にも対応しており、Windows Server 2012 以降の Windows OS と組み合わせることにより、そのメリットを最大限に活用することができます。

1.3 略語

本書で記載している略語は以下の通りです。

略語	正式名称
IB	InfiniBand
HCA	Host Channel Adapter
TCA	Target Channel Adapter
RDMA	Remote Direct Memory Access
FDR	Fourteen Data Rate
LID	Local Identifier
GUID	Global Unique Identifier
SMB	Server Message Block

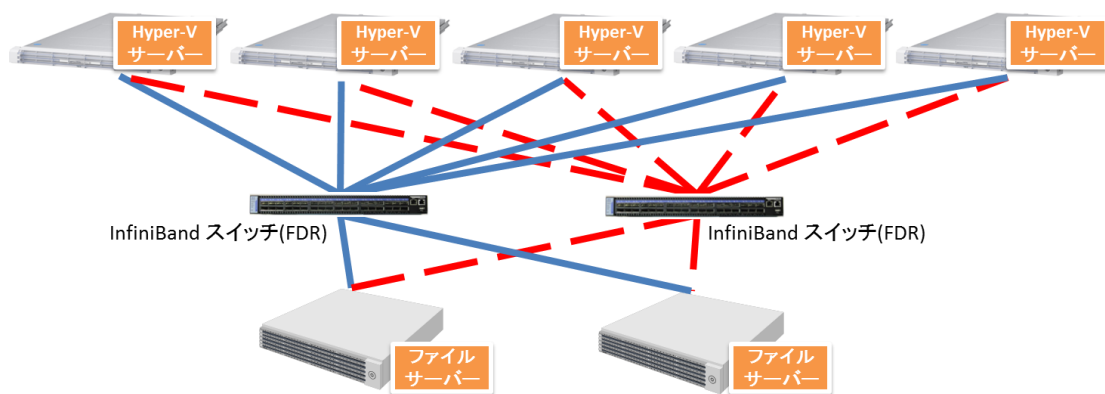
1.4 InfiniBand 環境の構築

- 各サーバ、ソフトウェアの構成ガイドまたは製品通知の動作要件につきましてもあわせて確認してください。
- InfiniBand 接続ボードと Windows Server 2012 の SMB 3.0 を利用した仮想基盤環境を構築する際には、必ず「5. 注意・制限事項」を確認したうえで構築してください。

1.5 ハードウェア接続イメージ

InfiniBand の RDMA 機能と、Windows Server 2012 の SMB 3.0 を利用した仮想基盤環境のハードウェア接続構成例を以下に示します。

InfiniBand(RDMA)と Windows Serve2012(SMB 3.0)を利用した仮想基盤環境の構成例



2 事前準備

2.1 ハードウェアの手配について

2.1.1 Express5800 関連

2.1.1.1. サーバ本体と InfiniBand 接続ボード

InfiniBand(RDMA)と Windows Server 2012 (SMB 3.0)を利用した仮想基盤を構築する際には、InfiniBand 接続ボード (N8104-146/ N8104-147), InfiniBand スイッチ(36port, FDR) (NE3707-061) を必ず手配願います。Windows Server 2012 のサポート状況、InfiniBand 接続ボードを搭載可能なサーバ、及び搭載枚数については、Express5800/100 シリーズのシステム構成ガイド、およびテクニカルガイドをご参照願います。

Express5800 シリーズ システム構成ガイド

<http://www.nec.co.jp/products/express/systemguide/index.shtml>

→ PC サーバ : Express5800 シリーズ > “タワーサーバ、ラックサーバ” あるいは”ECO CENTER”

→ 対象モデル名を選択

Express5800 シリーズ サポート情報 [PC サーバ]

<http://support.express.nec.co.jp/pcserver/>

→ PC サーバ : 情報を検索する > “型番・モデル名から探す” > “N8104-146”あるいは”N8104-147”を入力

→ テクニカルガイド “InfiniBand 接続ボード”を選択

2.1.2 InfiniBand スイッチ関連

2.1.2.1. InfiniBand スイッチと InfiniBand ケーブル

InfiniBand 接続ボードを利用の際には、必ず InfiniBand スイッチと InfiniBand ケーブルの手配してください。

2.2 マニュアルの入手

本書の中では、各製品のマニュアルの該当箇所を示しながら導入の手順を説明します。マニュアルについては各製品に添付されているほか、最新版が Web 上に公開されています。構築作業を行う際は、Web 上に掲載されている最新版を入手して作業されることをお勧めします。

Express5800 シリーズ ユーザーズガイド

<http://support.express.nec.co.jp/pcserver/>

- カテゴリから選択する > 製品マニュアル(ユーザーズガイド)
- Express100 シリーズ / ECO CENTER(i モデル)
- 対象モデル名(R120d-1M / R120d-2M / R120d-1E / R120d-2E / E120d-M / E120d-1 など)を選択
- 製品マニュアル(ユーザーズガイド)をクリックし、検索結果より対象モデルのユーザーズガイドを選択

InfiniBand 接続ボード ユーザーズガイド

<http://support.express.nec.co.jp/pcserver/>

- 型番・モデル名から探す
- 製品型番(InfiniBand 接続ボード : N8104-146[1ch] / N8104-147[2ch]) を入力し「製品型番で検索」を実施
- 検索結果より対象型番を選択
- 「すべてのカテゴリ 検索結果」内のユーザーズガイドを選択

InfiniBand スイッチ(36port, FDR) ユーザーズガイド

<http://www.nec.co.jp/products/pcserver/manycore/support/index.shtml>

- オプション製品 ユーザーズガイド
- InfiniBand スイッチ
- “NE3707-061 InfiniBand スイッチ(36port, FDR) ユーザーズガイド” を選択

Express5800 シリーズ Microsoft® Windows Server® 2012 サポート情報

<http://support.express.nec.co.jp/os/w2012/>

- ラックサーバ / ECO CENTER
- 対象モデルを選択

2.3 ハードウェア緒元一覧

各製品に関するお問い合わせは、弊社営業もしくはファーストコンタクトセンターへお問い合わせください。

Express5800 シリーズに関するお問い合わせ

『NEC ファーストコンタクトセンター』 TEL:03-3455-5800

受付時間 : 9:00～12:00/13:00～17:00 月曜日～金曜日(祝日を除く)

(電話番号をよくお確かめの上、おかけください)

『ご購入、導入についてのご相談』

http://www.nec.co.jp/products/express/question/top_sv1.shtml

→ 導入のご相談

iStorage シリーズに関するお問い合わせ

『NEC プラットフォームビジネス本部』 TEL:03-3798-9740

受付時間: 9:00～12:00/13:00～17:00 月曜日～金曜日(祝日を除く NEC 営業日)

(電話番号をよくお確かめの上、おかけください)

3 検証環境の事前設定

本章では、検証の前に設定が必要な、下記の 3 項目について説明します。

- ホスト OS(Windows Server 2012)のインストール
- InfiniBand の事前設定
- Hyper-V のインストール

3.1 ホスト OS(Windows Server 2012)のインストール

3.1.1 Windows Server 2012 のインストール

Express5800 シリーズに、Windows Server 2012 をインストールする方法について説明します。

3.1.1.1. 事前準備

➤ インストールに必要なもの

- Windows Server 2012 対応差分モジュール
R120e-1M/-2M に関しては、装置に添付されている EXPRESSBUILDER を利用した Windows Server 2012 のセットアップに対応しています。
- OS インストールメディア
✓ NEC 製 OS インストールメディア (以降、「バックアップ DVD」と呼ぶ)
- Windows Server 2012 インストール補足説明書

➤ 必要なモジュールとドキュメントの入手方法について

以下のウェブサイトが必要なモジュールとドキュメントを入手することができます。

- **Windows Server 2012 サポート情報**

<http://support.express.nec.co.jp/os/w2012/index.html>

→ 使用するモデルのページを開く

3.1.1.2. ホスト OS(Windows Server 2012)のインストール・セットアップ

「Windows Server 2012 インストール補足説明書」の手順に従って、OS のインストール・セットアップを実施してください。

3.1.1.3. 修正モジュールの適用

OS のインストール・セットアップ完了後に以下の修正モジュールを適用してください。

- **Windows Server 2012 用更新プログラム(KB2811660)**

<http://www.microsoft.com/ja-jp/download/details.aspx?id=36916>

- ◆ KB2800088
- ◆ KB2811660
- ◆ KB2812829
- ◆ KB2815769
- ◆ KB2823233

KB2823233 は OS の機能である「Media Foundation」を有効にしている時のみ適用が必要です。

「Media Foundation」機能はデフォルトでは無効になっています。

- **KB2853466**

<http://support.microsoft.com/kb/2853466/ja>

この修正モジュールは KB2811660 適用後にインストールしてください。

3.2 InfiniBand の事前設定

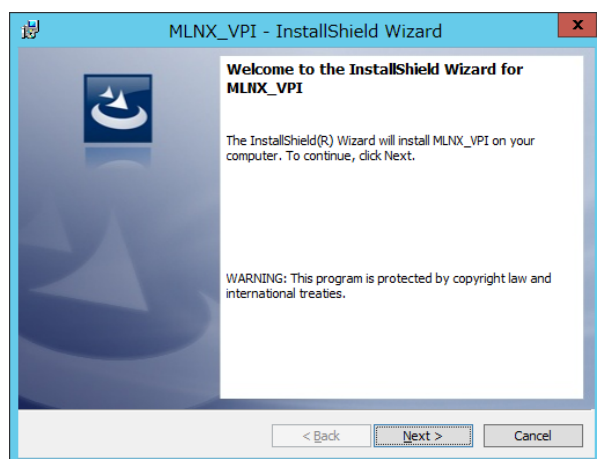
3.2.1 InfiniBand ドライバのインストール

Windows Server 2012 には、OS 標準の InfiniBand ドライバが存在し、自動的にドライバが適用されますが、この OS 標準ドライバは、弊社のサポート対象外です。

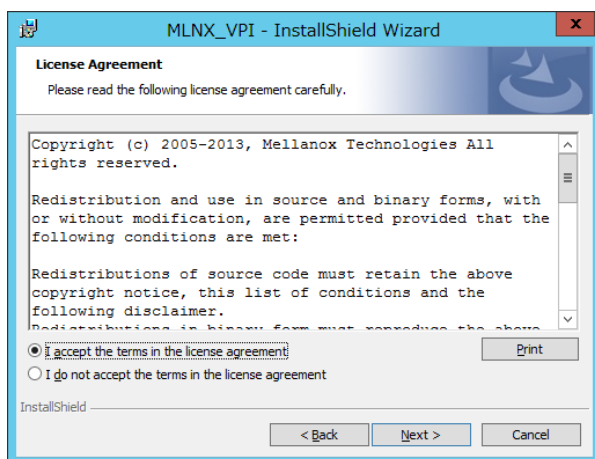
InfiniBand 接続ボード同梱の CD 媒体からドライバをインストールする必要があります。その適用手順を下記に記載します。

1. WinOF(ドライバ・管理アプリケーションパッケージ)のインストール

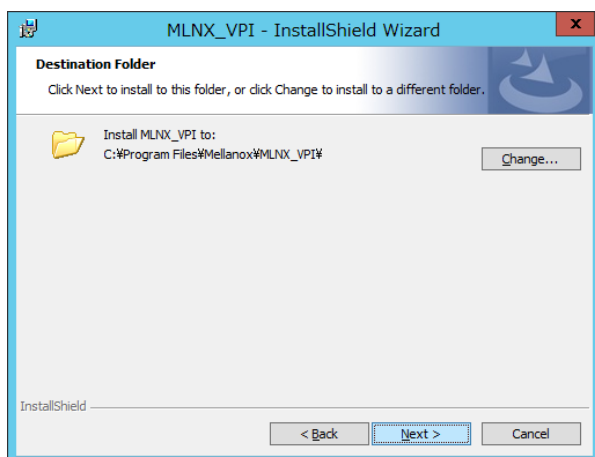
製品添付のドライバ CD 媒体から “MLNX_VPI_WinOF-4_40_0_All_win8_x64.exe” を実行し、“Next>” を左クリックします。



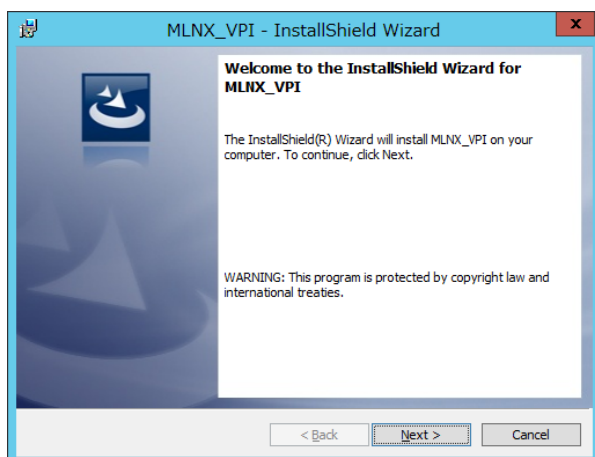
2. “I accept the terms in the license agreement” をチェックし、“Next >” を左クリックします。



3. インストールフォルダを選択し、“Next >” を左クリックします。
ここではデフォルトのファイルパス “C:\Program Files\Mellanox\MLNX_VPI\” にインストールします。

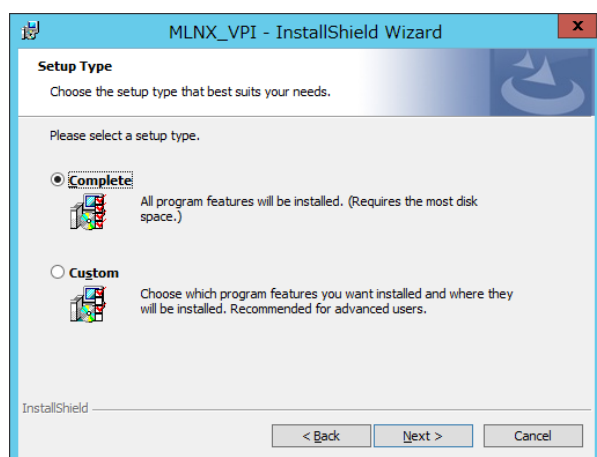


4. インストールウィザードの“Next >” ボタンを押します。

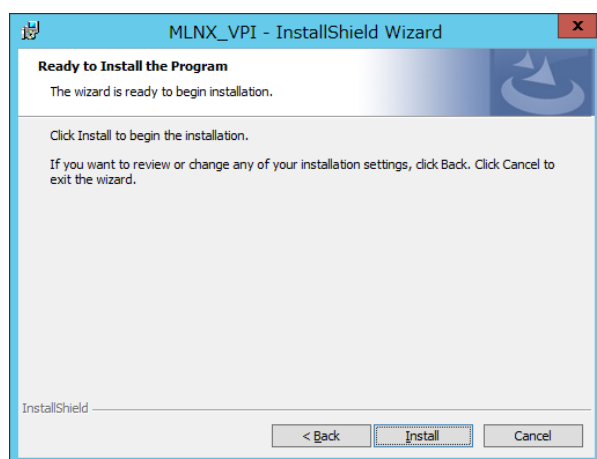


5. セットアップタイプを選択し、“Next >” を左クリックします。

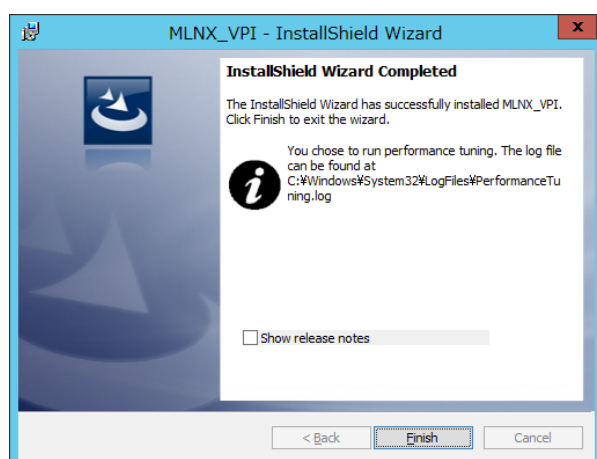
ここでは”Complete”を選択しました。



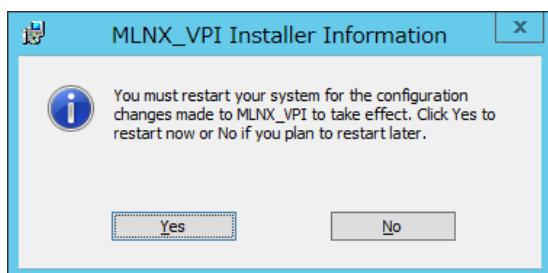
6. “Install” を左クリックし、インストールを開始します。



7. インストールを完了を待つて、“Finish” を左クリックしドライバのインストールを終了します。

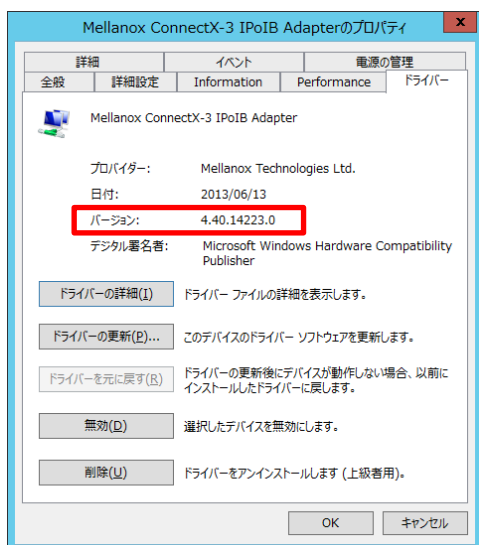


8. 設定を反映させるため、“Yes” を左クリックし、マシンを再起動します。



9. インストール後にバージョンを確認します。

“コントロール パネル¥ネットワークとインターネット¥ネットワーク接続” からインストールされた HCA のプロパティの “ドライバー” タブを見ることで、ドライバのバージョンが確認できます。



バージョンが 4.40.14223.0 以降にアップデートされていることを確認ください。

以上でドライバのインストールは終了です。

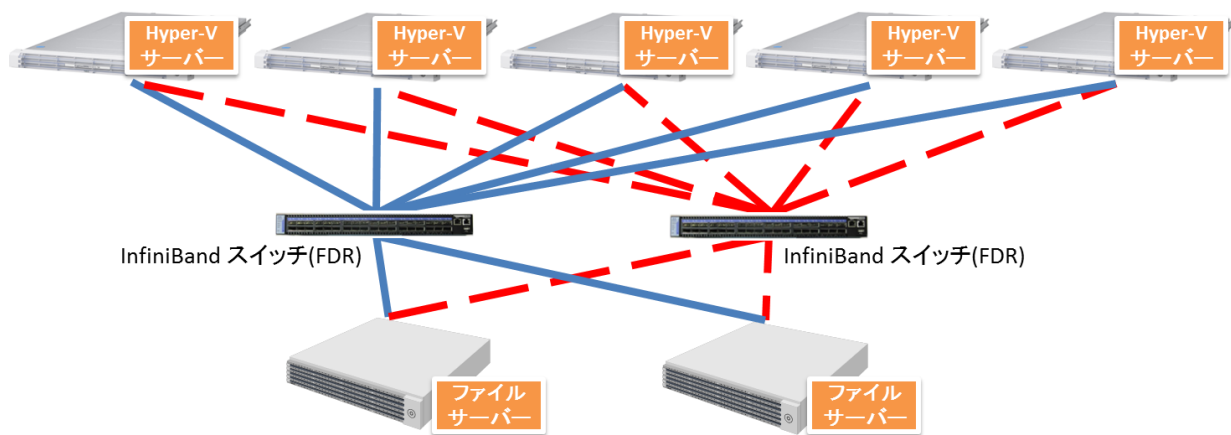
3.2.2 InfiniBand スイッチ(サブネットマネージャ)

3.2.2.1. 概要

InfiniBand の動作には InfiniBand ファブリック毎に最低1つ以上のサブネットマネージャが動作していることが必要です。サブネットマネージャは、InfiniBand ファブリック内の接続トポロジの検出や監視、及び通信の管理を実施しています。

InfiniBand スイッチ(36port, FDR) [NE3707-061]には、このサブネットマネージャの機能が実装されており、設定を enable にすることで、サブネットマネージャの機能を利用することができます。

下記のシステム構成例では、InfiniBand ファブリックが2個(実線と点線の InfiniBand ファブリック)存在することになり、それぞれのファブリックを、それぞれの InfiniBand スイッチ上のサブネットマネージャが管理しています。



システム構成例
(InfiniBand ファブリックが2つ存在))

3.2.2.2. InfiniBand スイッチ(サブネットマネージャ)の設定

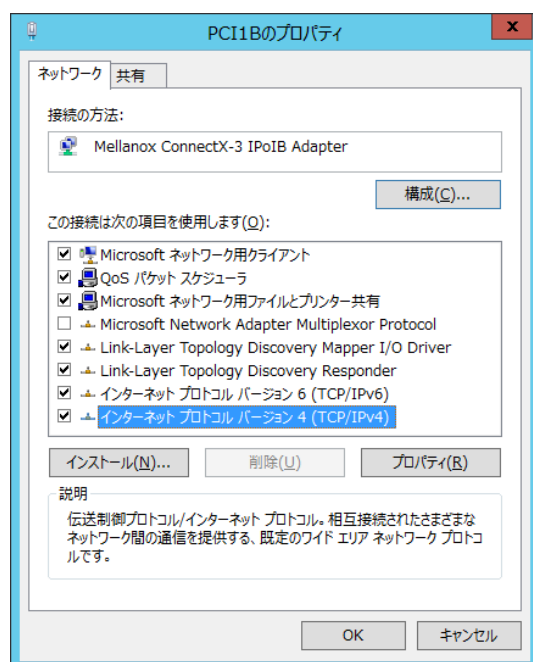
InfiniBand スイッチ(36port, FDR) [NE3707-061]の工場出荷時のサブネットマネージャ設定は Disable に設定されていますので、シリアルポート経由もしくはイーサネットポート経由でスイッチコンソール等にアクセスし、enable に設定してください。(必要に応じて、priority 等の関連設定も設定してください)。

詳細な手順は InfiniBand スイッチ(36port, FDR) ユーザーズガイド をご覧ください。

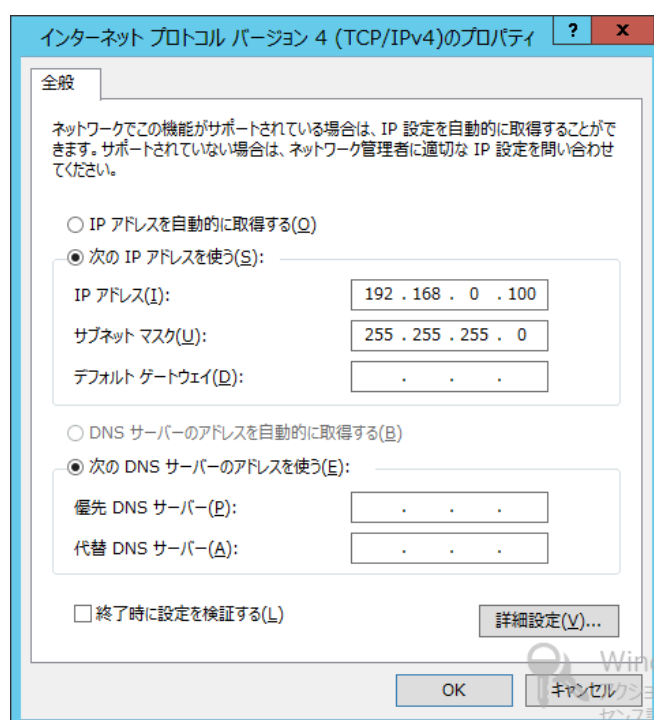
3.2.3 InfiniBand 接続ボードの設定

通常のネットワークカードと同様に設定いただけます。

1. “コントロール パネル¥ネットワークとインターネット¥ネットワーク接続” から InfiniBand アダプターを選択し、プロパティを開きます。



2. プロパティから “インターネット プロトコル バージョン 4 (TCP/IPv4)” を選択し IP アドレスとサブネットマスクを設定します。



上記は、192.168.0.100/ 255.255.255.0 に設定する場合の例

3.3 Hyper-V のインストール

Hyper-V の機能を利用する場合は、以下ウェブサイトに掲載されている「Windows Server 2012 Hyper-V インストール手順書」に従って、Hyper-V のインストールを実施してください。

- **Express5800 シリーズにおける Windows Server 2012 Hyper-V サポートについて**

<http://support.express.nec.co.jp/os/w2012/hyper-v.html>

→ インストール手順

→ Hyper-V のインストール手順

→ Windows Server 2012 Hyper-V インストール手順書 [WS2012_Hyper-V_install.pdf]

以下のウェブサイトに掲載されている注意・制限事項も確認してください。

- **Express5800 シリーズにおける Windows Server 2012 Hyper-V サポートについて**

<http://support.express.nec.co.jp/os/w2012/hyper-v.html>

→ 注意制限事項

4 動作検証

InfiniBand 動作や性能を検証します。

検証項目

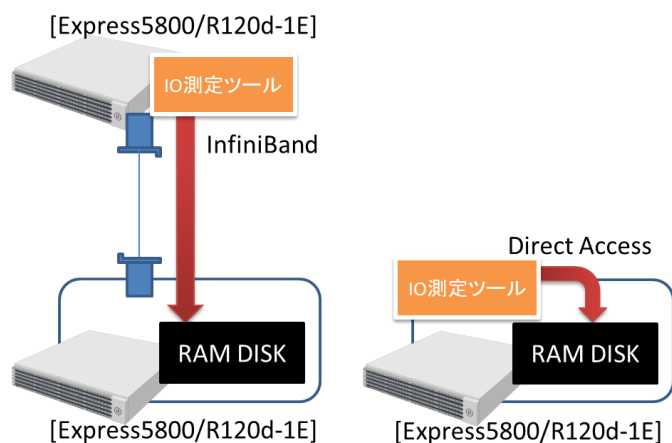
- ・ スループット
- ・ SMB 3.0 マルチチャネル
- ・ 透過フェールオーバー
- ・ Hyper-V over SMB 3.0

4.1 スループット

InfiniBand 接続ボードを用いて、SMB3.0 ファイル共有フォルダに対する Read/Write の性能を測定しました。高速な InfiniBand の性能測定では、従来、配慮不要であったストレージドライブ速度がボトルネックになり得るため、十分な Read/Write 性能を発揮可能なストレージドライブを利用する必要があります。本検証では、HDD や SSD などのディスクドライブを用いる代わりに RAM ディスク上に共有フォルダを作成することで InfiniBand の速度測定に十分な速度を確保しました。

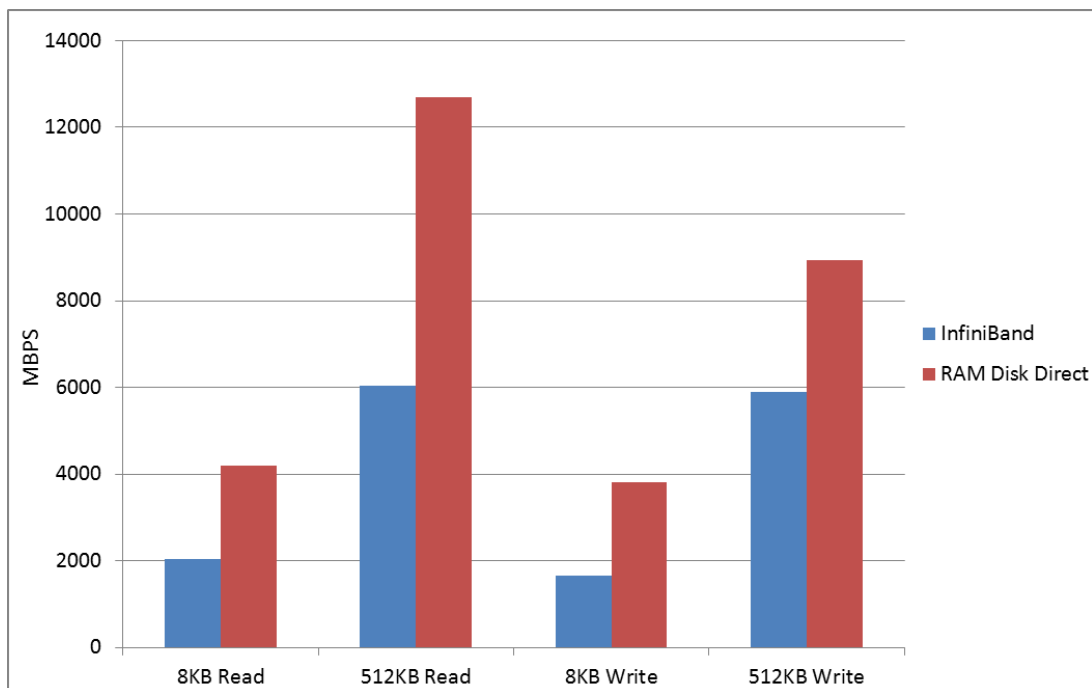
【検証内容】

ディスク IO の性能検証ツール (iometer) を用い、InfiniBand 経由でのディスク IO とローカルでのディスク IO を計測し、比較しました。



[結果]

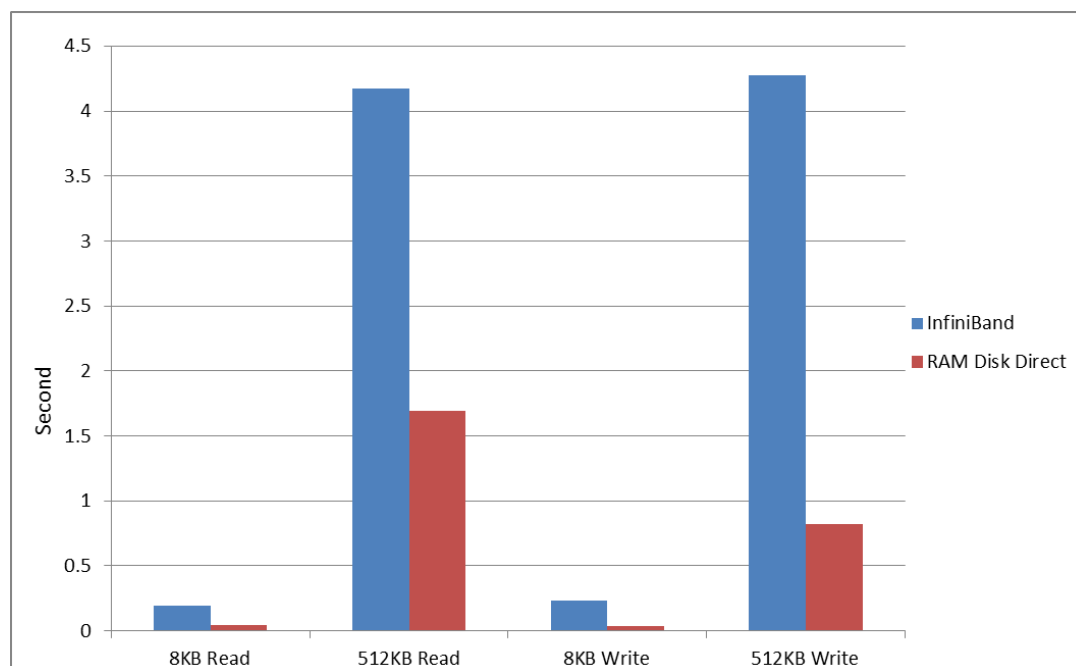
RAM ディスクへの InfiniBand 経由と直接アクセスの速度比較



RAM ディスクへの InfiniBand 経由と直接アクセスの CPU 使用率の比較



RAM ディスクへの InfiniBand 経由と直接アクセスのレスポンスタイムの比較



Iometer による計測結果。6Worker 8 Outstanding I/O 512KB Sequential Read

I/O サイズ	アクセス方式	MBPS (Decimal)	IOPS	CPU Utilization Total (%)	Average response time (Second)
8KB Read	InfiniBand(FDR)	2,047.97	249,996.01	50.63	0.1917
	ローカル (RAM ディスク直接アクセス)	4,185.27	510,896.66	47.06	0.0448
8KB Write	InfiniBand(FDR)	1,664.01	203,126.68	45.33	0.2359
	ローカル (RAM ディスク直接アクセス)	3,809.57	465,034.90	47.33	0.0385
512KB Read	InfiniBand(FDR)	6,028.26	11,497.99	9.81	4.1739
	ローカル (RAM ディスク直接アクセス)	12,691.18	24,206.50	50.34	1.6934
512KB Write	InfiniBand(FDR)	5,885.96	11,226.57	14.54	4.2750
	ローカル (RAM ディスク直接アクセス)	8,932.25	17,036.91	50.10	0.8215

InfiniBand FDR x4 の速度の理論値は 54Gbps で、Read 速度(512KB)の結果は、6028.26MBPS (47.01Gbps) で、理論値の 87%程度の転送レートが出ました。現在、販売されている SSD でも Read 速度は最高 3.0Gbps (24Gbps) 程度のもので多く、InfiniBand では単体の SSD などのディスク速度を大幅に超えた高速で通信が可能です。今回は高速な RAM ディスク上に共有ディスクを作成しましたが、実際の環境で InfiniBand の性能を活かすには、高速なストレージを複数用いて十分なディスクアクセス速度の確保が必要となります。

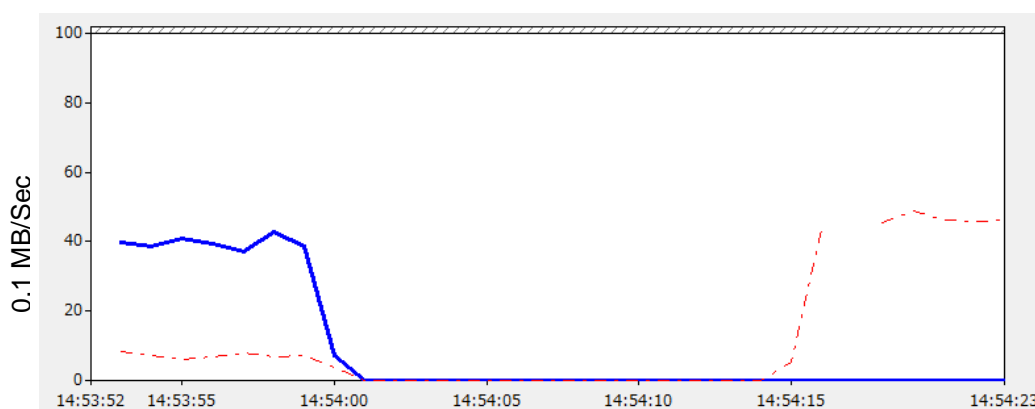
4.2 SMB マルチチャネル

SMB マルチチャネルは、Windows Server 2012 で新規にサポートされた機能です。

SMB マルチチャネル機能は、InfiniBand 接続ボードの複数ポートを利用することで、NIC チーミングのような、ネットワーク速度の高速化、物理ネットワークの切断でも動作を続ける障害時の信頼性向上などの効果があります。

[検証内容]

デュアルポートの InfiniBand RDMA FDR x4 接続での InfiniBand アダプターを経由したネットワークファイルコピー中に、ポート 1 のケーブルを抜き、ファイル転送が継続することを確認しました。



RDMA Activity¥RDMA Inbound Bytes/sec (Scale 0.0000001)

青(実践): ポート1 最大値: 429,197,481 Bytes/Sec 最小値: 0 Bytes/Sec

赤(点線): ポート2 最大値: 487,880,446 Bytes/Sec 最小値: 0 Bytes/Sec

[結果]

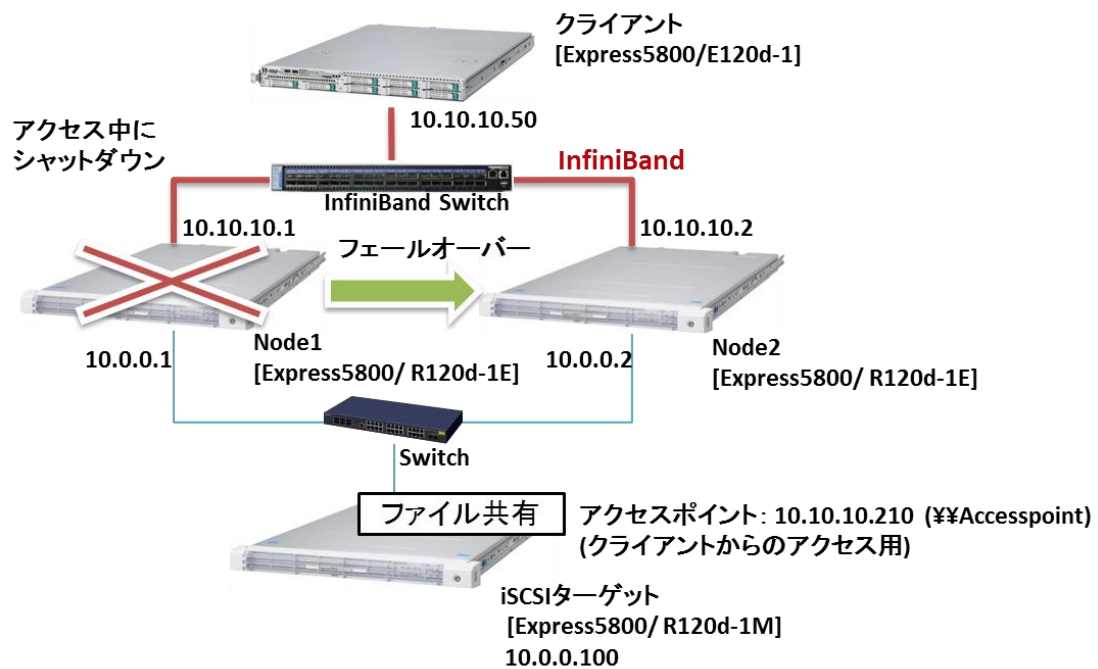
切断後 19 秒間、RDMA によるネットワークが停止し、コピーも中断しましたが、コピーの処理自体がタイムアウトになることはなく、残りのポートを通して転送が継続することを確認しました。ファイルコピー完了後のファイルについても破損がなく正常にコピーが行われました。

4.3 透過フェールオーバー

透過フェールオーバーは、フェールオーバー時にアプリケーションからの継続的なファイルアクセスを可能にする機能です。ノード障害などでフェールオーバーの発生時には、ネットワーク遅延は発生しますが、アプリケーションからのファイルアクセスは継続され、アプリケーションからは、フェールオーバーを認識することなく、サービスを継続することができます。

[検証内容]

SMB 3.0 クライアントから共有ディスクへ 10GByte ファイルをコピー中に、アクティブノードのファイルサーバをシャットダウンすることで、フェールオーバーが発生させ、フェールオーバーにかかる時間を計測しました。



[利用条件]

- ・ iSCSI などの共有ストレージがクラスターに参加する 2 台以上のノードから利用可能なこと。
- ・ クライアントと、ファイルサーバークラスターを構成するサーバがアクティブディレクトリドメインに参加していること。

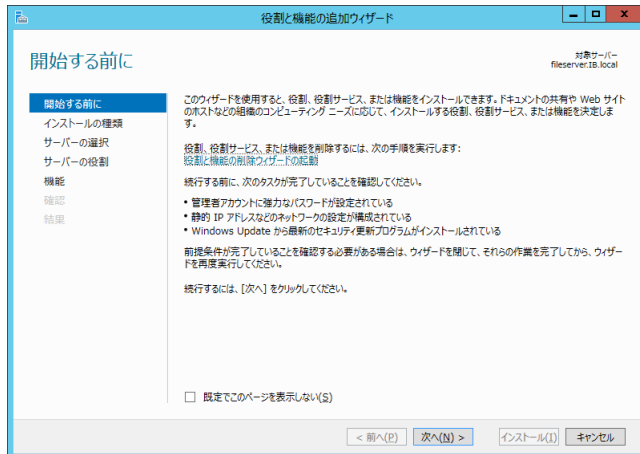
本検証では、事前に Windows Server 2012 に標準で搭載されている Microsoft iSCSI ターゲットを用いて 100GB と 1GB の共有ディスクを作成し、Node1、Node2 の 2 台のサーバーからマウントしました。

[設定]

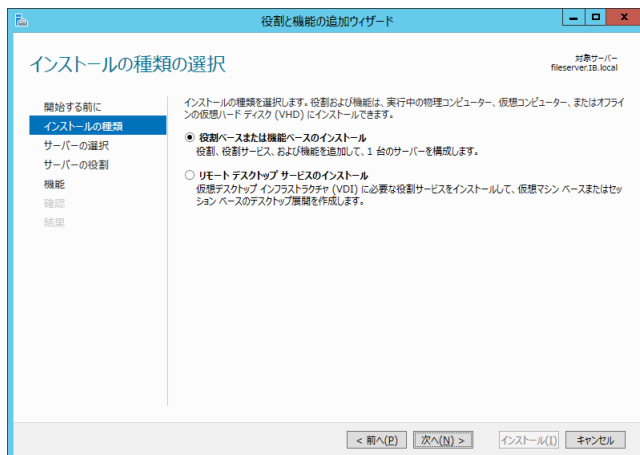
1. フェールオーバークラスターの機能のインストール

注：ファイルサーバクラスターに含まれるすべてのサーバーで実行する必要があります。本検証では、図の Node1 と Node2 の 2 台のサーバーにフェールオーバークラスターの機能をインストールしています。

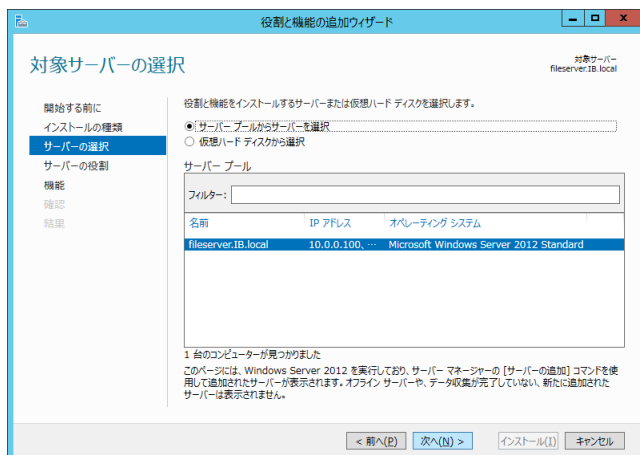
1.1. サーバーマネージャの “役割と機能の追加” を選択すると、役割と機能の追加ウィザードが実行されるので、“次へ” を左クリックします。



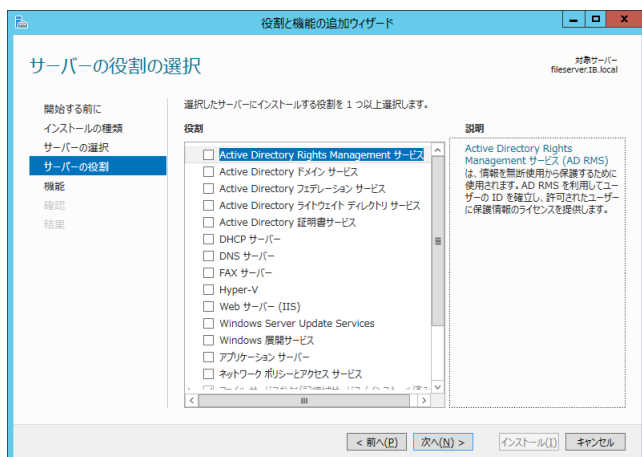
1.2. 役割ベースまたは機能ベースのインストールにチェックが入っていることを確認して “次へ” を左クリックします。



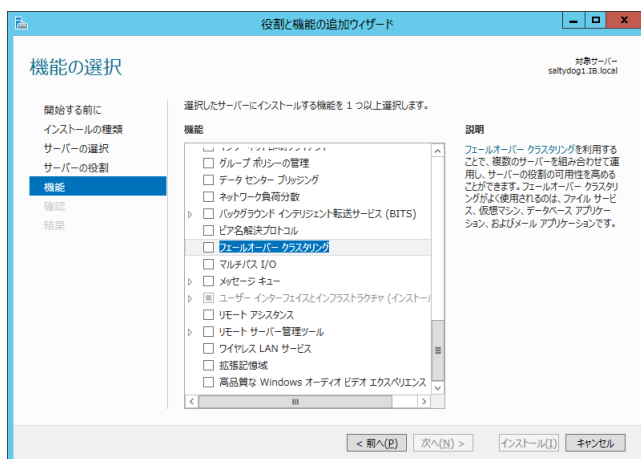
1.3. 機能をインストールするサーバーを選択し、“次へ” を左クリックします。



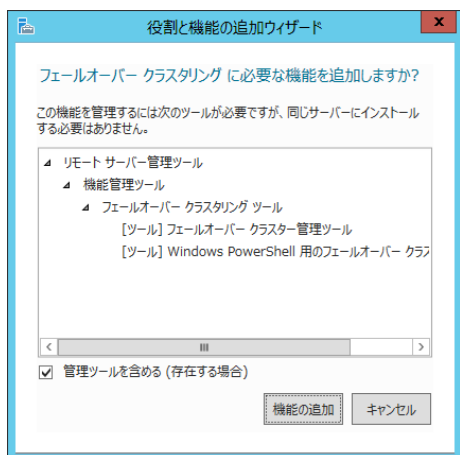
- 1.4. 役割の選択画面になりますが、なにもチェックせずにそのまま “次へ” を左クリックします。



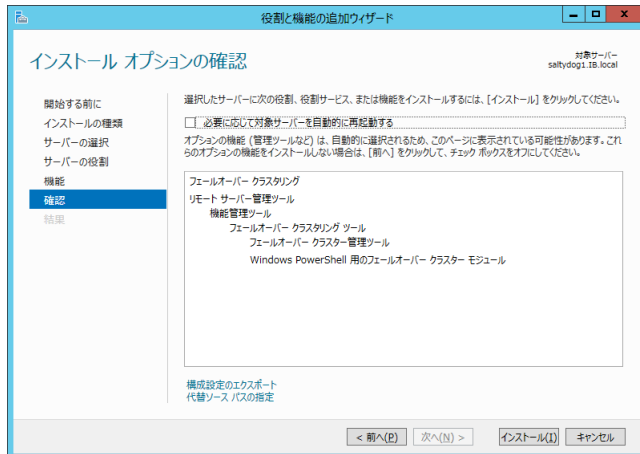
- 1.5. 機能の選択画面になるので、“フェールオーバークラスタリング” にチェックを入れます。



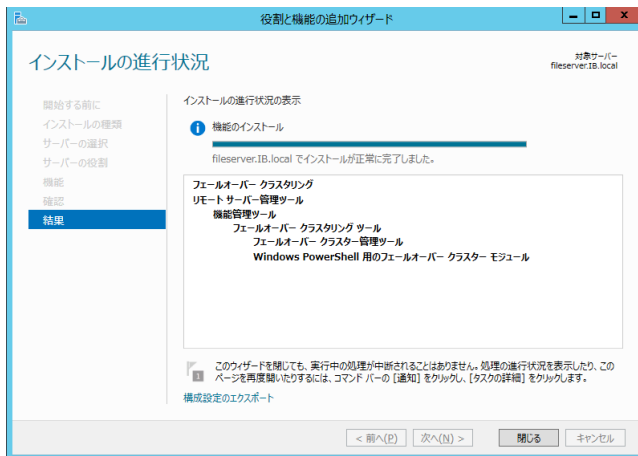
- 1.6. 「フェールオーバー クラスタリングに必要な機能を追加しますか？」と聞かれるので、“機能の追加”を左クリックします。



1.7. インストール オプションの確認を行い、“インストール” を左クリックします。



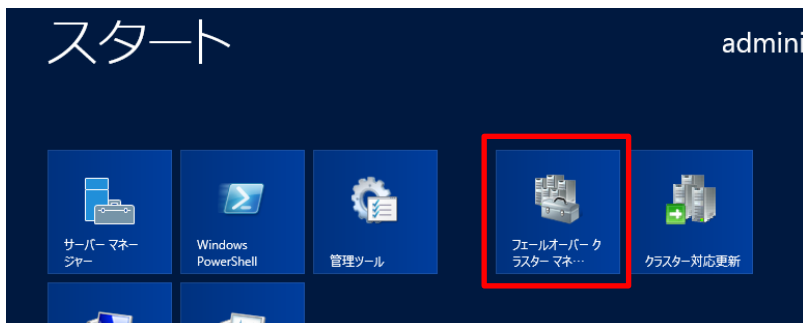
1.8. インストールの完了を待ち “閉じる” を左クリックします。



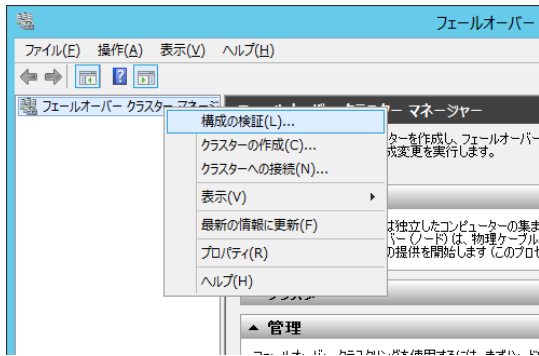
2. フェールオーバークラスターを構成します。

注意：以降の手順は、クラスターに参加する 2 台以上のサーバーで、前述のフェールオーバークラスターの機能のインストールが終了していることを前提としています。

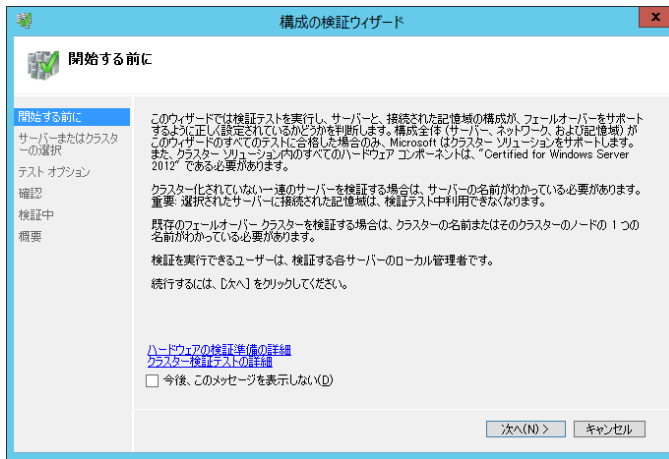
2.1. スタートスクリーンからクラスタマネージャを起動します。



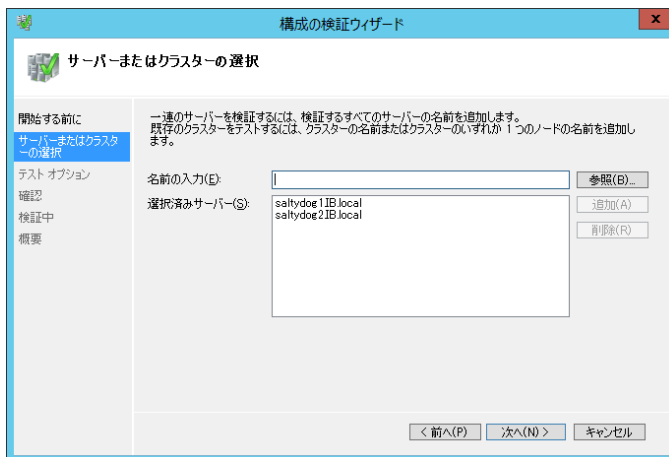
- 2.2. 左ペインの”フェールオーバー クラスター マネージャー”を右クリックし、”構成の検証”を左クリックします。



- 2.3. 構成検証ウィザードが開始するので、”次へ”をクリックします。



- 2.4. クラスターを構成するフェールオーバークラスターの機能がインストールされたサーバーの名前(最低2つ)を入力し、”次へ”をクリックします。



ここでは、クラスターを構成する 2 台のサーバーの名前をそれぞれ
Saltydog1.IB.local
Saltydog2.IB.local としています。

2.5. すべてのテストを実行するを選択し、“次へ”をクリックします。



2.6. 検証の準備が完了したら、“次へ”をクリックします。

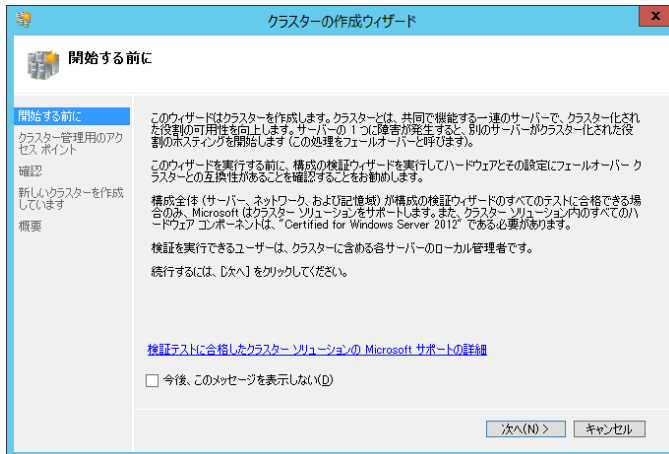


2.7. テストの結果が表示されるのを待ちます。

警告が出ることがありますが、警告が問題がないものであるか、レポートの表示ボタンを押して、内容を確認してください。確認後、検証されたノードを使用してクラスターを今すぐ作成するにチェックが入っていることを確認し、“完了ボタン”を押します。



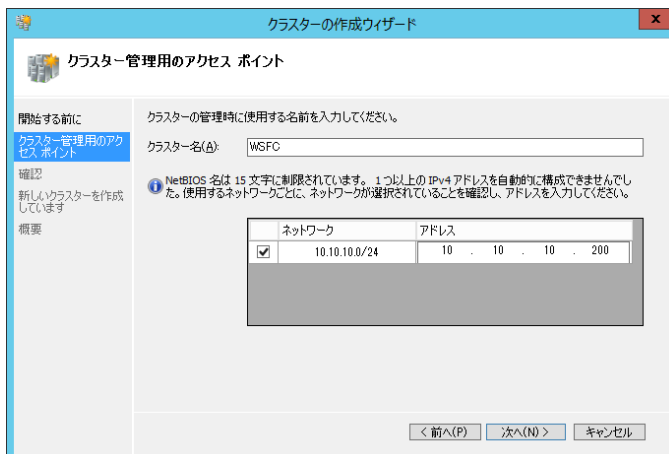
2.8. クラスターの作成ウィザードが開始するので、“次へ”をクリックします。

2.9. クラスター名と、クラスターの管理時にアクセスするための IP アドレスを入力し、“次へ”をクリックします。
ここでは、例として

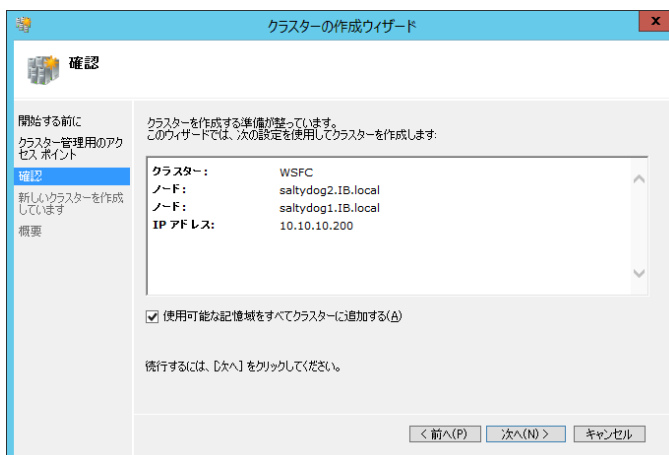
クラスター名「WSFC」

IP アドレスは “10.10.10.200”

としました。



2.10. 設定を確認し、使用可能な記憶域をすべてクラスターに追加するにチェックが入っていることを確認し、“次へ”をクリックします。

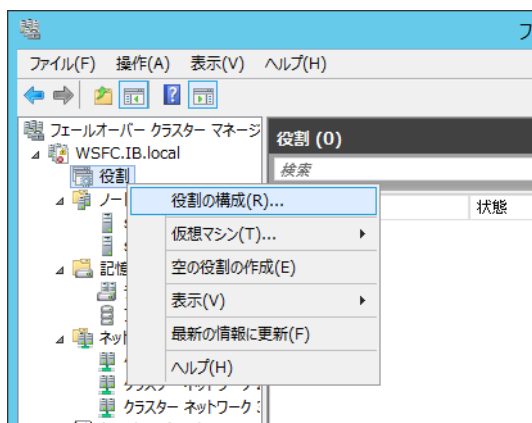


2.11. 作成が正常に完了すると以下のようなメッセージが表示されるので、“完了”を左クリックします。

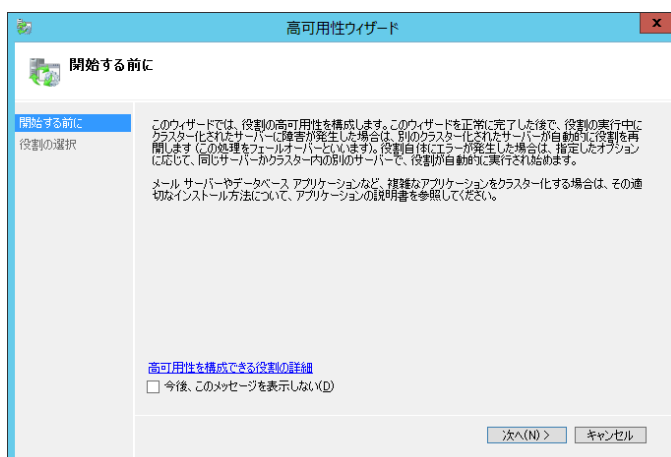


3. ファイルサーバの役割を追加します

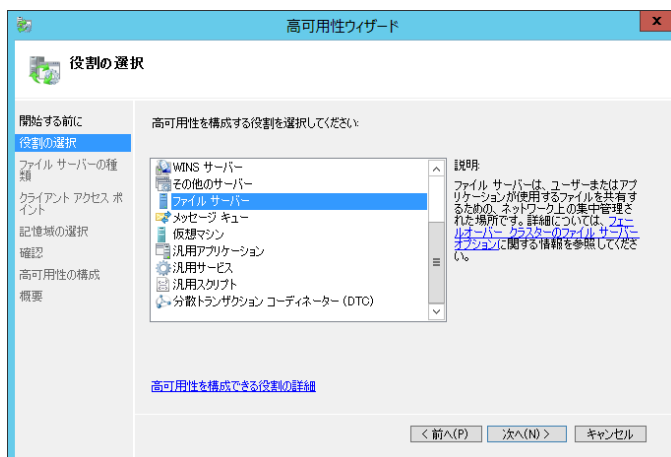
3.1. “役割”を右クリックし、“役割の構成”を左クリックします。



3.2. 高可用性ウィザードが開始するので、“次へ”を左クリックします。

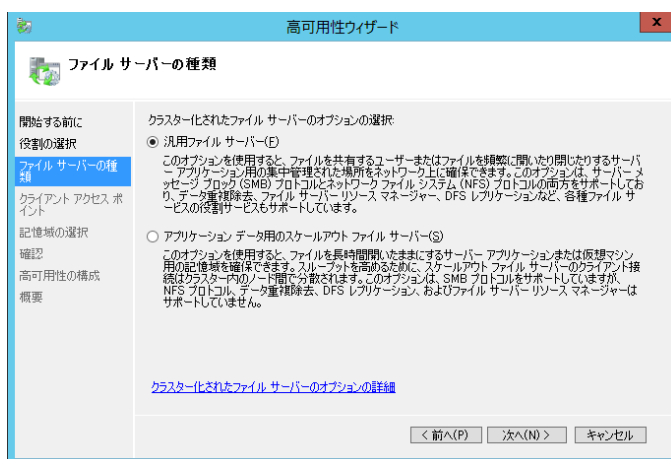


3.3. ファイルサーバを選択し、“次へ”をクリックします。



3.4. ファイルサーバのオプションを選択します。

ここでは、「汎用ファイルサーバ」を選択しました。

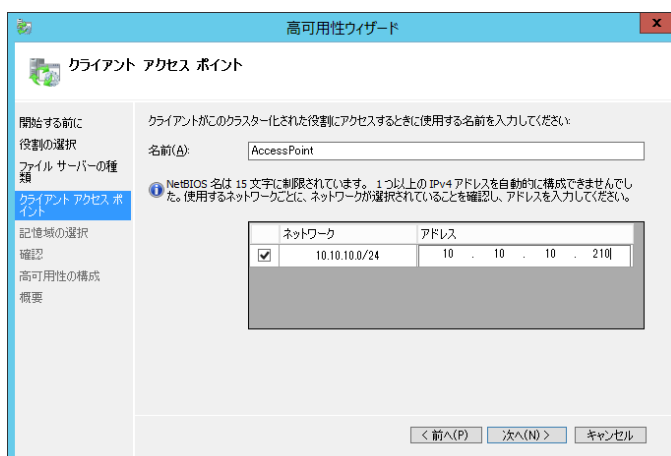


3.5. クライアントがファイルサーバにアクセスするための共有フォルダー名と IP アドレスを設定します。

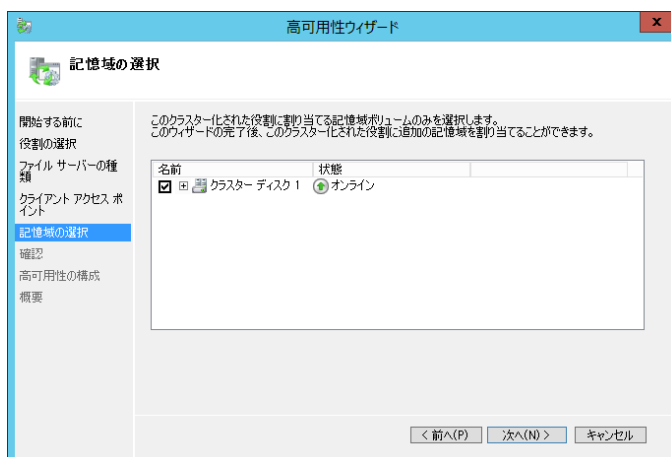
ここでは、例として

アクセスに使用する名前として「AccessPoint」

IP アドレス “10.10.10.210”を設定しました。



3.6. クラスター化する共有ディスクにチェックを入れて、“次へ”を左クリックします。



3.7. ネットワーク名と IP アドレスなどの情報が正しいかを確認し、“次へ”を左クリックします。



3.8. ファイルサーバの構成が完了するので、“完了”を左クリックし、ウィザードを終了します。

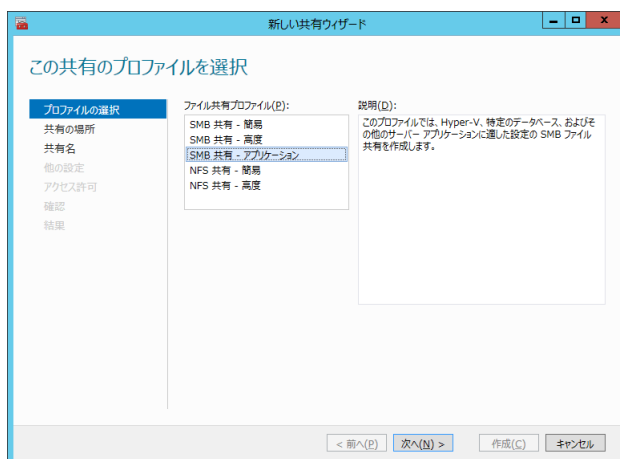


4. 透過フェールオーバー(継続的可用性)を設定します。

- 4.1. “フェールオーバークラスターマネージャー WSFC 役割 AccessPoint” を右クリックし、“ファイル共有の追加”を左クリックします。



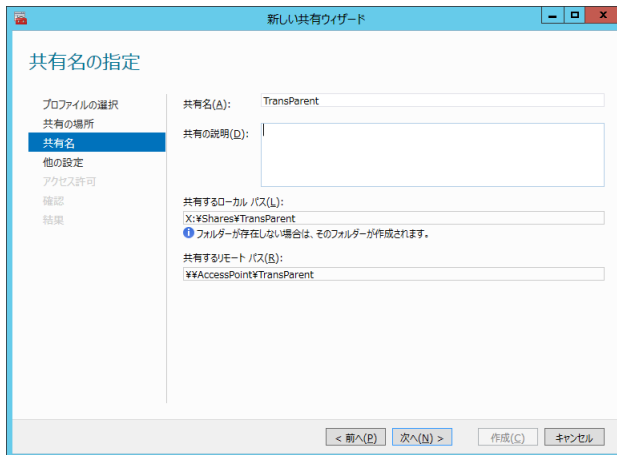
- 4.2. 新しい共有ウィザードが開始するので、“SMB 共有 – アプリケーション” を選択し、“次へ”を左クリックします。



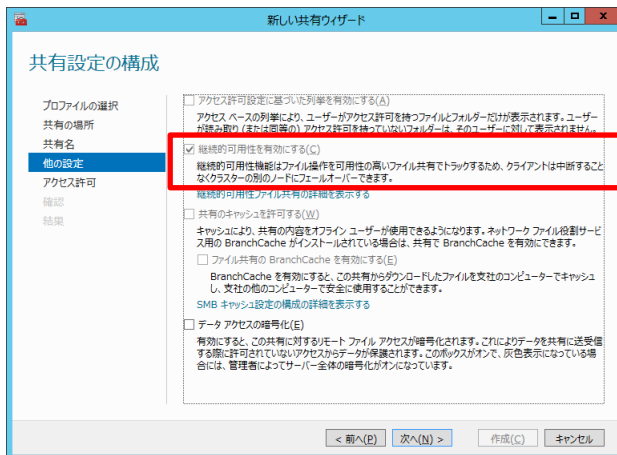
- 4.3. 透過フェールオーバーを適用する新しい共有のサーバとボリュームを選択し、“次へ”をクリックします。



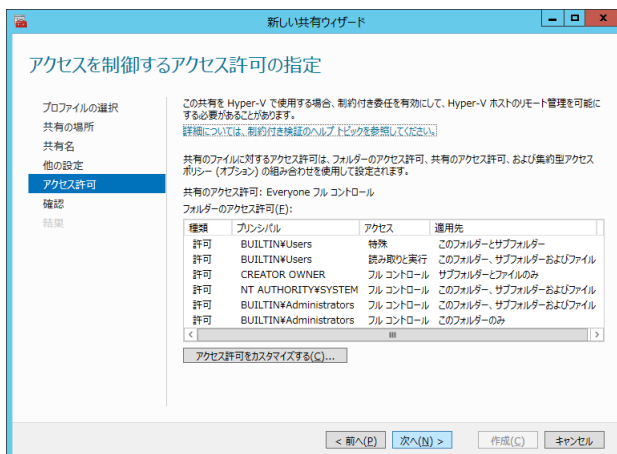
- 4.4. フォルダアクセスに使う共有名を入力し、“次へ”を左クリックします。
ここでは、共有名を「Transparent」として設定しました。



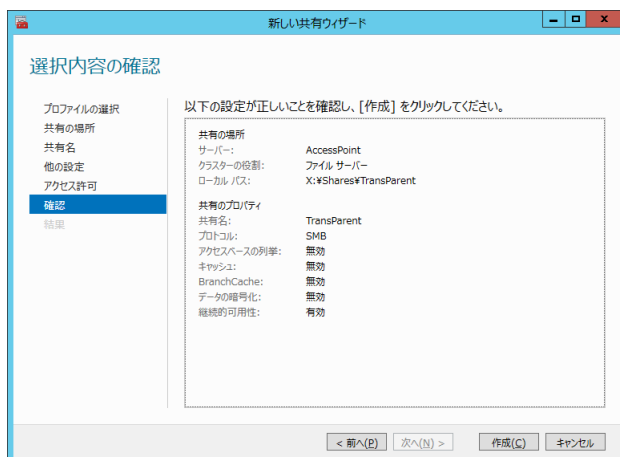
- 4.5. “継続的可用性を有効にする” にチェックが入っていることを確認し、“次へ”を左クリックします。



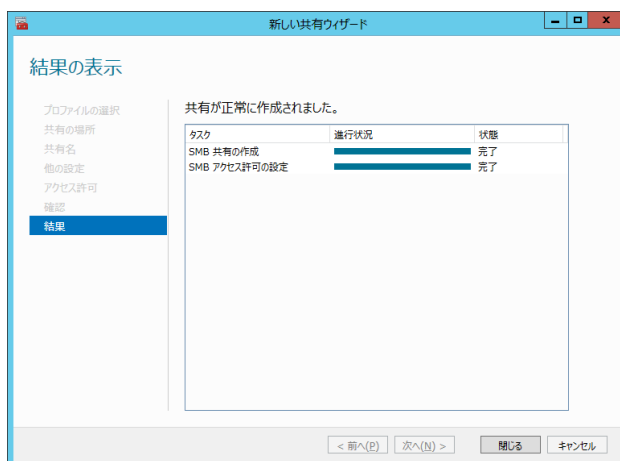
- 4.6. アクセス許可の状態を確認し、必要に応じてアクセス権の変更します。



4.7. ローカルパスなどの設定が意図したものであるかを確認し、“次へ”をクリックします。



4.8. 正常に共有フォルダが作成されるのを確認し、“閉じる”を左クリックします。



以上で透過フェールオーバーの設定は完了です。

クライアントから透過フェールオーバーに対応したフォルダへの接続には、上記の設定ではエクスプローラーなどのファイルパスに、

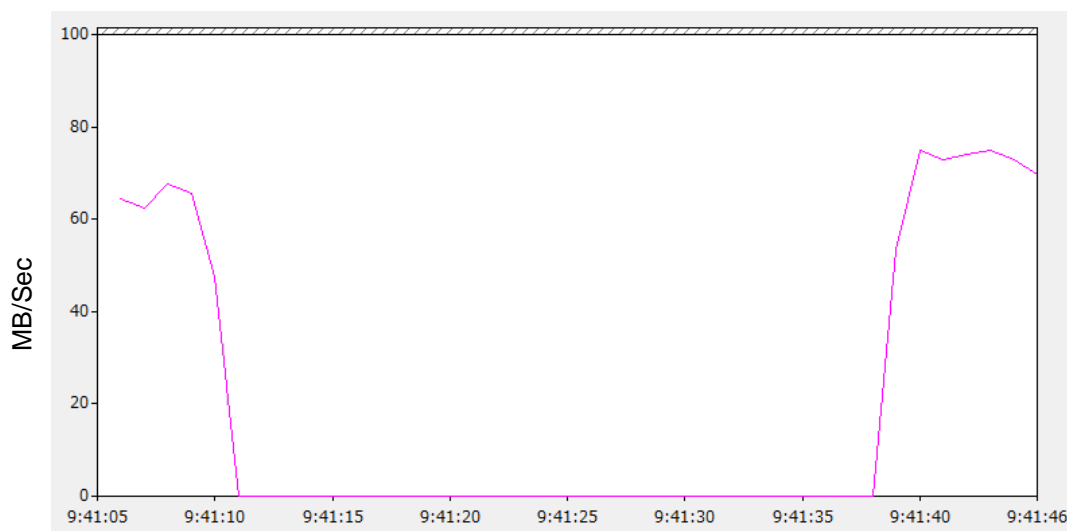
“¥¥AccessPoint¥TransParent” または、

“¥¥10.10.10.210¥TransParent”

で接続することができます。

[結果]

RDMA による通信は、リソースモニタやタスクマネージャーからは確認できないため、SMB 3.0 クライアント上からパフォーマンスカウンターにより RDMA Activity の通信状況を確認しました。



計測値: RDMA Activity: RDMA Outbound Bytes/Sec

最大: 69,691,701 Bytes/Sec

最低: 0 Bytes/Sec

無通信時間: 26 秒

シャットダウンから 30 秒弱ほどクライアントとサーバ間でのネットワークが断絶し、ファイル転送が停止しましたが、その後、フェールオーバーが完了し、フェールオーバー先のノードを経由してファイル転送が再開。ファイルの破損なくコピーが完了しました。

[参考]

Microsoft より修正プログラムの情報がでています。必要に応じて修正プログラムを適用ください。

「Windows Server 2012 では、ドライブのケーブルを切断した後、SMB 3.0 の透過的なフェールオーバー機能が機能しません。」

<http://support.microsoft.com/kb/2884665/ja>

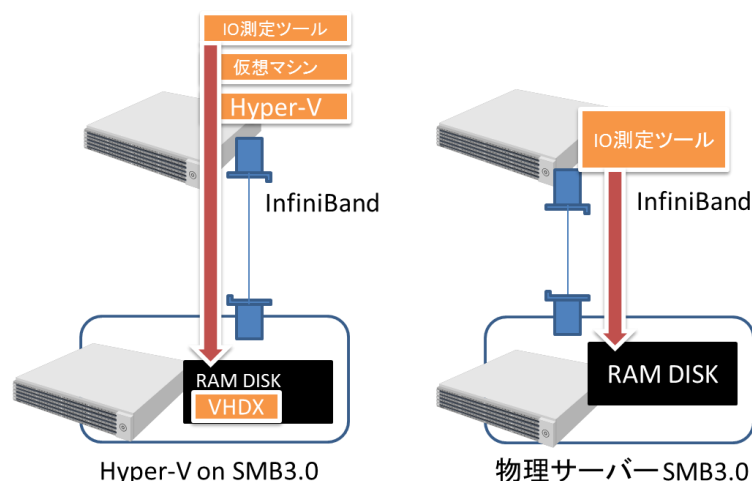
4.4 Hyper-V over SMB3.0

Windows Server 2012 以降の Windows サーバでは、仮想ディスクの格納先に SMB3.0 の共有ディスクを選択することができます。管理も簡単で安価な NAS サーバをストレージとして、利用することができます。

InfiniBand 接続ボードは、RDMA に対応しているので、InfiniBand の高速なネットワークによるディスクアクセス速度、スケールアウト性の向上の他、SMB3.0 マルチチャネルや透過フェールオーバーと組み合わせることで、可用性の面でのメリットも享受することができます。

[検証内容]

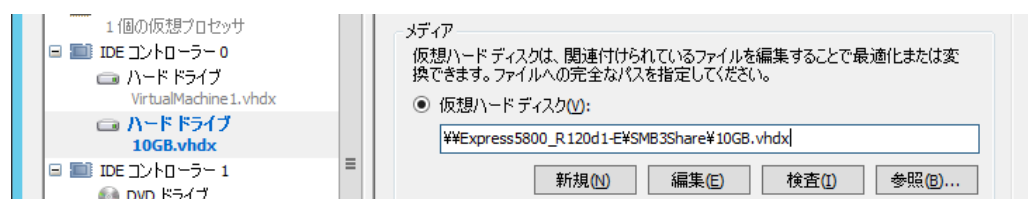
SMB3.0 ファイル共有上に 10GB の仮想ハードディスクを作成し、仮想マシン上の IO 測定ツール(iometer)から負荷をかけて、ディスクの Read 性能を測定しました。比較のため、物理マシンから直接 RAM ディスクへアクセスした場合の性能と比較を行いました。



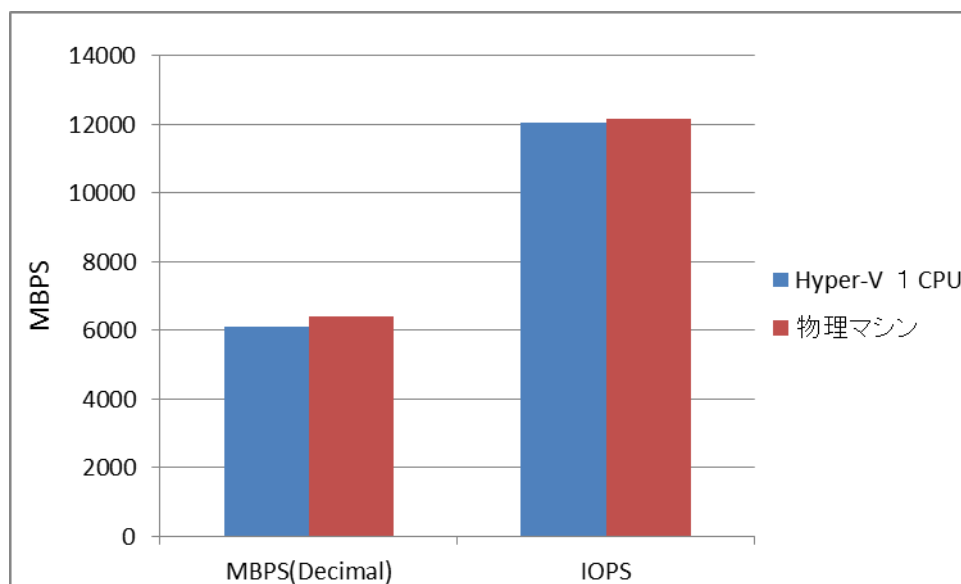
[設定]

共有フォルダのあるサーバと Hyper-V サーバはともにアクティブディレクトリドメインに参加し、仮想ハードディスクを置くための共有フォルダは、Hyper-V からアクセスできるように権限を与える必要があります。

仮想ハードディスクは、仮想マシンの設定から、ファイル共有で用いる"¥¥"で始まる共有フォルダのパスを指定する必要があります。



[結果]



RDMA 有効にし、iometer (1worker 8threads 512KB Seaquential Read)での性能比較

	MBPS(Decimal)	IOPS
Hyper-V 1CPU	6,084.36	12,032.31
物理マシン	6,379.91	12,168.72

Hyper-V 環境においても、物理マシンに近いスループットが出ており、Hyper-V Over SMB3.0 の有効性がわかります。

5 注意・制限事項

本章では、Express5800/100 シリーズ上で仮想ファイバーチャネル環境を構築する際の注意点、および制限事項をご説明します。

5.1 サーバ

5.1.1 最新 BIOS の適用について

使用する OS によっては、BIOS をアップデートする必要があります。公開されている最新バージョンにアップデートすることを強く推奨します。

アップデート手順に関しては、各装置の BIOS ダウンロードページの紹介文、およびダウンロードデータに含まれる「Readme.txt」を参照してください。

BIOS アップデートモジュールの有無については、以下のウェブサイトで確認してください。

Express5800 シリーズ PC (IA)サーバ サポート情報

<http://support.express.nec.co.jp/pcserver/>

- 100 シリーズ / ECO CENTER(i モデル)を選択
- 使用するモデルを選択
- ダウンロードを選択し、モジュールの有無を確認

5.2 RDMA の設定について

5.2.1 パフォーマンスカウンター、リソースカウンターについて

RDMA を用いた通信では、OS を介さず HCA に処理をオフロードするため、通常の RDMA をサポートしないネットワークカードと同様の手順では、パフォーマンスカウンターやリソースカウンターからネットワーク負荷を計測することができません。ネットワーク負荷を測定したい場合、パフォーマンスカウンターから別途、RDMA の通信を計測できる以下のようなカウンターを選択する必要があります。

- RDMA Activity
- SMB Direct Connection
- SMB Client Shares
- SMB Server Shares
- SMB Server Session
- Mellanox Adapter Diagnostic Counters
- Mellanox Adapter QoS Counters
- Mellanox Adapter Traffic Counters

5.2.2 設定の確認

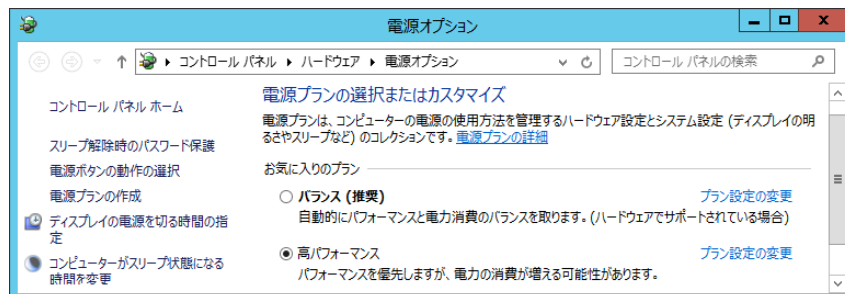
RDMA は標準で有効に設定されていますが、無効になっている場合は、RDMA の機能が使えません。

PowerShell から RDMA が有効になっているか確認し、無効の場合は、設定を有効に変更してください。また、InfiniBand 接続ボードのシングルポートしか使わない場合でも、RDMA の機能を利用する場合には、SMB マルチチャネルの設定を有効にしてください。

Get-NetAdapterRDMA	enabled が True となっていることをご確認ください。HCA、NIC が RDMA 対応でない場合は、結果は表示されません。 False になっている場合は、以下のコマンドにより、RDMA を有効にしてください。 Set-NetAdapterRDMA -enabled \$True
Get-NetOffloadGlobalSetting	NetworkDirect: Enabled となっていることをご確認ください。
Get-NetAdapterHardwareInfo	PCI Express のバス数やバージョン、リンクスピードを確認することができます。(Linkspeed は unknown で取得できないこともあります。) HCA が適切なスロットに接続されているかご確認ください。
Get-SmbClientConfiguration	EnableMultiChannel : True となっていることを確認してください。 False になっている場合は、以下のコマンドにより、RDMA を有効にしてください。 Set-SmbClientConfiguration -enabled \$True
Get-SmbClientNetworkInterface	RDMA Capable: True Speed: 54Gbps となっていることをご確認ください。
Get-SmbServerConfiguration	EnableMultiChannel : True となっていることを確認してください。 False になっている場合は、以下のコマンドにより、RDMA を有効にしてください。 Set-ServerClientConfiguration -enabled \$True
Get-SmbServerNetworkInterface	目的の Scope Name (共有フォルダ) の RDMA Capable が True になっているかと Speed が 54Gbps になっているかご確認ください。
netstat.exe -xan ? {\$_ -match "445"}	RDMA で接続予定のネットワークが Listener になっていることをご確認ください。
Get-SmbConnection	SMB のバージョンが Dialect で確認できます。RDMA の機能を利用するにはサーバ、クライアントともに SMB 3.0 以上である必要があります。 Windows Server 2012 ではバージョン 3.00 Windows Server 2012 R2 では、バージョン 3.02 です。
Get-SmbMultichannelConnection	クライアント、サーバごとの RDMA の利用可能状況を確認することができます。ClientRdmaCapable と ServerRdmaCapable が True になっていることをご確認ください。

5.2.3 電源オプション

サーバの CPU 性能によっては、電源オプションの設定がバランスになっている場合、CPU の周波数が低周波数に固定され、広帯域な RDMA 通信の性能をカバーしきれない場合があります。性能が十分に発揮できない場合には、電源オプションのお気に入りのプランを「高パフォーマンス」に設定することで性能が改善する場合もあります。



6 検証環境について

検証環境で利用したハードウェアとソフトウェアについて紹介します。

6.1 ソフトウェア

検証で用いたソフトウェアを紹介します。

Iometer win64 Version 1.1.0-rc1

Intel が開発した IO 測定用のツールです。現在はオープンソースになっており SourceForge などのページからダウンロードできます。今回は、Version 1.1.0-rc1 Windows 64bit 版を利用しました。

http://sourceforge.jp/projects/sfnet_iometer/

StarWind RAM Disk

StarWind 社が開発した RAM ディスクソフトウェアです。メモリから指定領域を確保し、仮想的に SCSI HDD として利用できます。HDD や SSD に比べ高速なディスクアクセスが実現できます。

<http://www.starwindsoftware.com/high-performance-ram-disk-emulator>

6.2 ハードウェア

検証で用いたサーバーハードウェアを紹介します。

マシン	Express5800/ R120d-1E (2 台)
CPU	1CPU (E5-2420[6Core/1.9GHz])
メモリ	48G バイト
OS	Windows Server 2012 Standard

マシン	Express5800/ R120d-1M
CPU	2CPU(E5-2420[6Core/1.9GHz])
メモリ	48G バイト
OS	Windows Server 2012 Standard

マシン	Express5800/R120d-1M
CPU	2CPU(E5-2690[8Core/2.9GHz])
メモリ	128G バイト
OS	Windows Server 2012 Standard

マシン	Express5800/E120d-1
CPU	2CPU(E5-2407[4Core/2.2GHz])
メモリ	64G バイト
OS	Windows Server 2012 Standard

Express 5800 シリーズのサーバについては、「2.1.1 Express5800 関連」をご覧ください。

Express5800/100 シリーズ
InfiniBand と Windows Server 2012 SMB3.0 を利用した
仮想基盤

2014 年 1 月 初版

日本電気株式会社
東京都港区芝五丁目7番1号

TEL (03) 3454-1111 (大代表)

© NEC Corporation 2014