



富士通サーバPRIMERGYとMellanox ConnectX-3 ProにおけるvMotion性能とスタン時間の検証

2016年2月

株式会社アルティマ

▶ 検証目的

- ◆ 本検証はMellanox社製ConnectX-3 Proを用い、以下2点の検証を実施しました。
 1. VMware ESXi6.0のin-boxドライバ認識検証
 2. 10GE/40GEにおける、vMotion時間及びスタン時間の検証

▶ ご協力いただいたベンダー様

- ◆ 検証サポート、検証結果のレビューについて、メラノックステクノロジーズ様にご協力いただきました。

▶ 検証場所・期間

- ◆ 場所：富士通検証センター(東京・浜松町)
- ◆ 期間：2016年2月8日～2月19日

▶ Mellanox ConnectX-3 Pro 10/40GEアダプターカード



- 型番：MCX314A-BCCT
- デュアルポート 10/40/56GE対応
- QSFPコネクタ
- PCIe3.0 x 8レーン

▶ Mellanox SX1710 10/40/56GEイーサネットスイッチ



- QSFP 36ポート
- 1/10/40/56GE対応
- Single ASIC
- X86 dual-CPU (Management)
- 低レイテンシ 220ns
- 低消費電力：92W (Typical)

▶ ハードウェア

	Type	CPU	MEM	PCIe
Server	PRIMERGY RX2540 M1	Xeon E5-2630v3 2.40GHz/8コア/20MB	256GB	Gen3
	PRIMERGY TX300 S8	Xeon E5-2697v2 2.70GHz/12コア/30MB	32GB	Gen3
Ethernet Switch	Mellanox SX1710, OS version : 3.4.3002			
NIC	Mellanox ConnectX-3 Pro , FW Rev :2.35.5100			

▶ ソフトウェア

- ◆ OS : ESXi6.0 build3073146
- ◆ vCenter Server 6.0
- ◆ NIC Driver : OS inbox

ESXi6.0 in-box ドライバ認証検証

▶ inboxドライババージョン確認

- ◆ OSインストール後、特にベンダードライバをインストールすることなくアダプターカードを使用することができます。
- ◆ 一部Mellanoxの独自機能や最新ドライバを使用する場合は別途、ベンダードライバ、もしくはVMware Certifiedドライバが必要となります。

詳細は下記参照。

Mellanox OFEDドライバ

- http://www.mellanox.com/page/products_dyn?product_family=36&menu_section=34

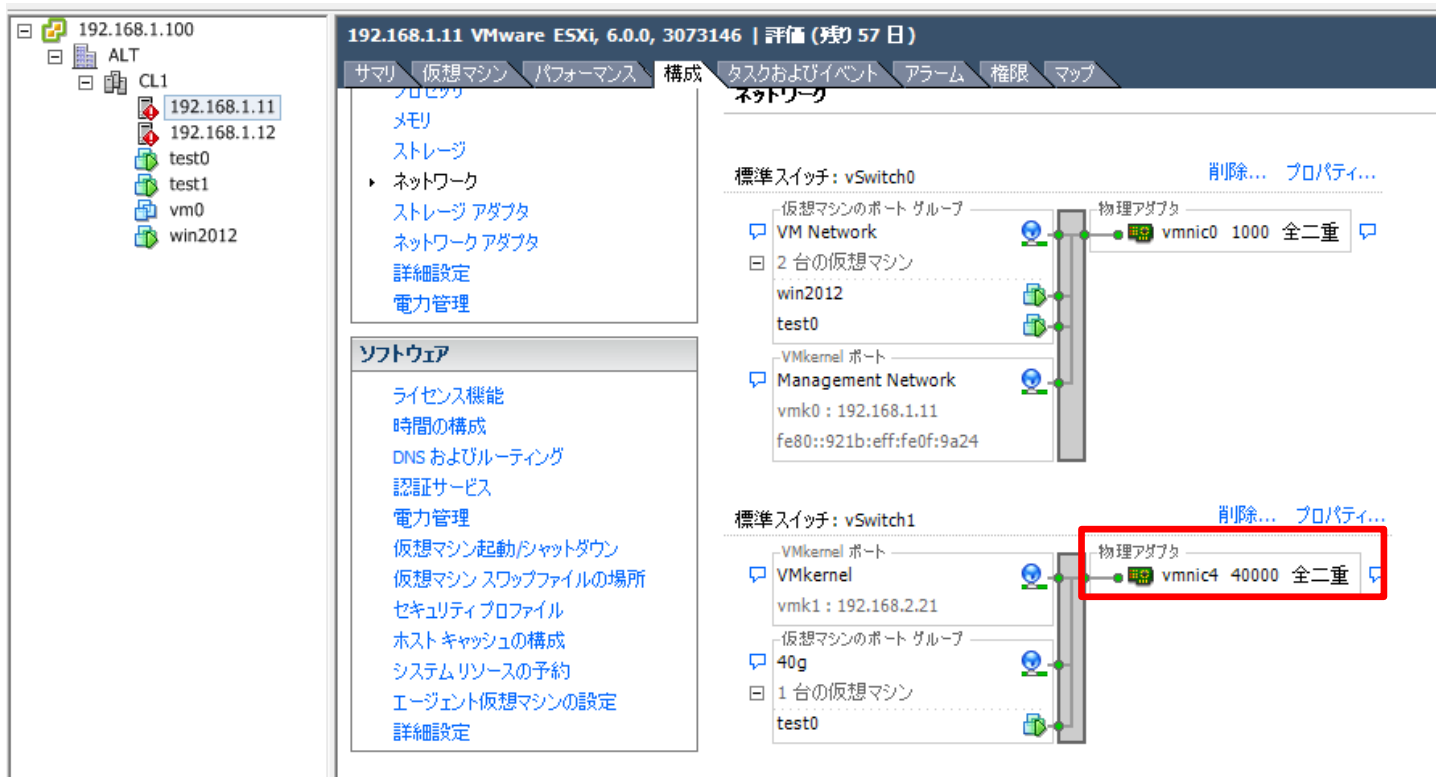
VMware Certified ドライバ

- <https://www.vmware.com/resources/compatibility/search.php>

```
[root@localhost:~] esxcli software vib list | grep -i mlx
```

net-mlx4-core	1.9.7.0-1vmw.600.0.0.2494585	VMware VMwareCertified	2016-01-08
net-mlx4-en	1.9.7.0-1vmw.600.0.0.2494585	VMware VMwareCertified	2016-01-08
nmlx4-core	3.0.0.0-1vmw.600.0.0.2494585	VMware VMwareCertified	2016-01-08
nmlx4-en	3.0.0.0-1vmw.600.0.0.2494585	VMware VMwareCertified	2016-01-08
nmlx4-rdma	3.0.0.0-1vmw.600.0.0.2494585	VMware VMwareCertified	2016-01-08

- ▶ 40GE vNICとして認識し、vSwitchに接続できることを確認



192.168.1.100

ALT

CL1

- 192.168.1.11
- 192.168.1.12
- test0
- test1
- vm0
- win2012

192.168.1.11 VMware ESXi, 6.0.0, 3073146 | 評価 (残り 57 日)

サマリ 仮想マシン パフォーマンス 構成 タスクおよびイベント アラーム 権限 マップ

ネットワーク

標準スイッチ: vSwitch0

仮想マシンのポート グループ

- VM Network
- 2 台の仮想マシン
 - win2012
 - test0

VMkernel ポート

- Management Network
 - vmk0 : 192.168.1.11
 - fe80::921b:eff:fe0f:9a24

物理アダプタ

- vmnic0 1000 全二重

標準スイッチ: vSwitch1

VMkernel ポート

- VMkernel
 - vmk1 : 192.168.2.21

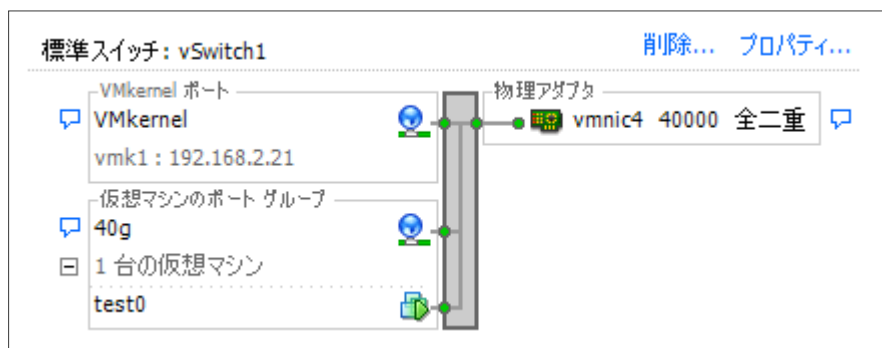
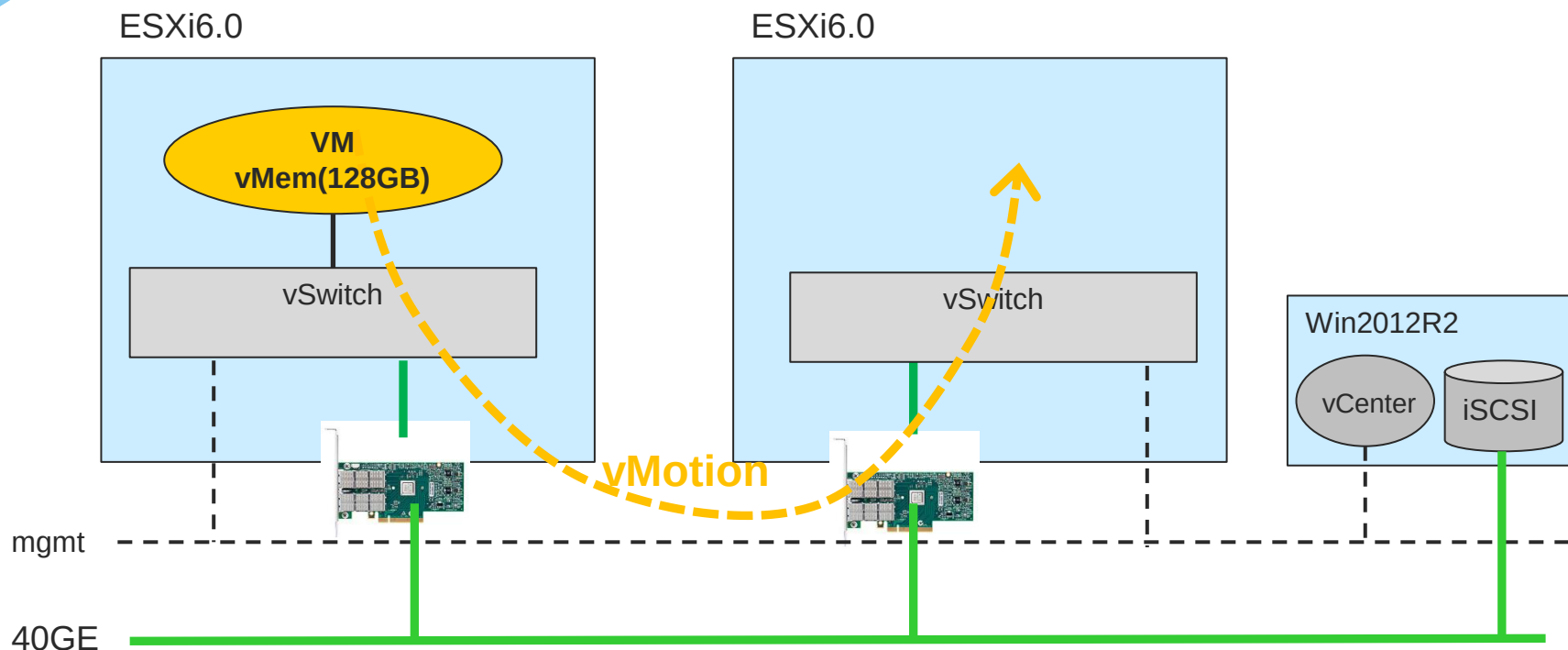
仮想マシンのポート グループ

- 40g
- 1 台の仮想マシン
 - test0

物理アダプタ

- vmnic4 40000 全二重

vMotion時間及びスタン時間の検証



▶ ネットワーク

- ◆ VMkernel(VMotion用ネットワーク)を40GEインターフェースに接続
- ◆ vSwitch, 物理スイッチのMTU=9000

▶ ストレージ

- ◆ iSCSIストレージをマウント

※Win2012R2 = Windows server 2012 R2

vMotionで移行する仮想マシン構成

- ▶ ゲストOS：Windows server 2012 R2
- ▶ vCPU：16
- ▶ メモリ：128GB
- ▶ ストレージ：外部iSCSI接続

全般	
ゲスト OS:	Microsoft Windows Server 2012 (64 ビット)
仮想マシンのバージョン:	11
CPU:	16 vCPU
メモリ:	131072 MB
メモリ オーバーヘッド:	
VMware Tools:	❖ 実行していません (未インストール)
IP アドレス:	
DNS 名:	
EVC モード:	該当なし
状態:	パワーオン
ホスト:	192.168.1.11
アクティブなタスク:	仮想マシンの移行
vSphere HA 保護:	② 該当なし

リソース		
消費されたホスト CPU:	10844 MHz	
消費されたホスト メモリ:	6755.00 MB	
アクティブなゲスト メモリ:	115343.00 MB	
	ストレージ使用率の更新	
プロビジョニングしたストレージ:	168.18 GB	
共有されないストレージ:	40.00 GB	
使用済みストレージ:	168.18 GB	
ストレージ	ステータス	ドライブのタイプ
iscsi	✓ 標準	非 SSD
◀ ▶		
ネットワーク	タイプ	ス
VM Network	標準ポート グループ	✓
◀ ▶		

メモリ負荷のかけかた

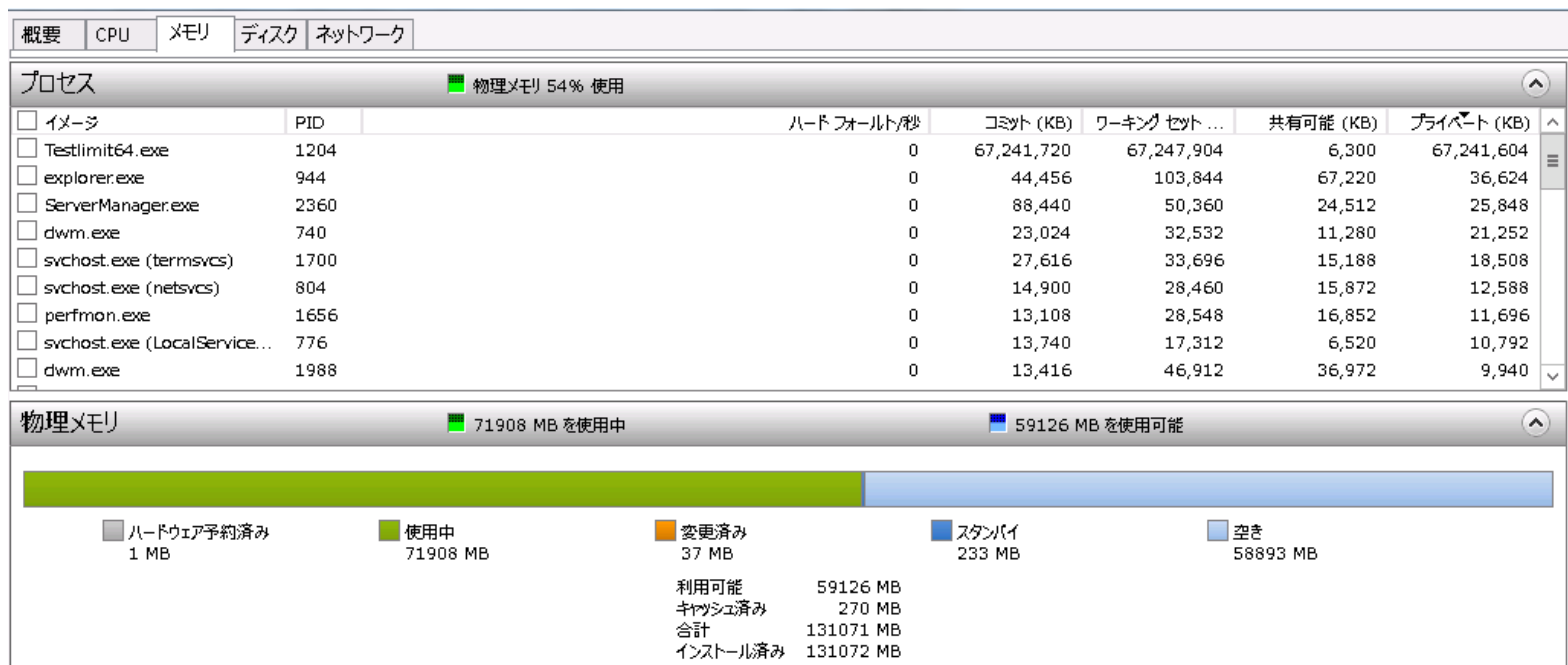
▶ 仮想マシンのメモリ負荷ツール

- ◆ Testlimitツールを使用し64GBの負荷をかける

```
TestLimit¥ Testlimit64.exe -d -c 65536
```

ソース

<https://blogs.msdn.microsoft.com/vijaysk/2012/10/26/tools-to-simulate-cpu-memory-disk-load/>



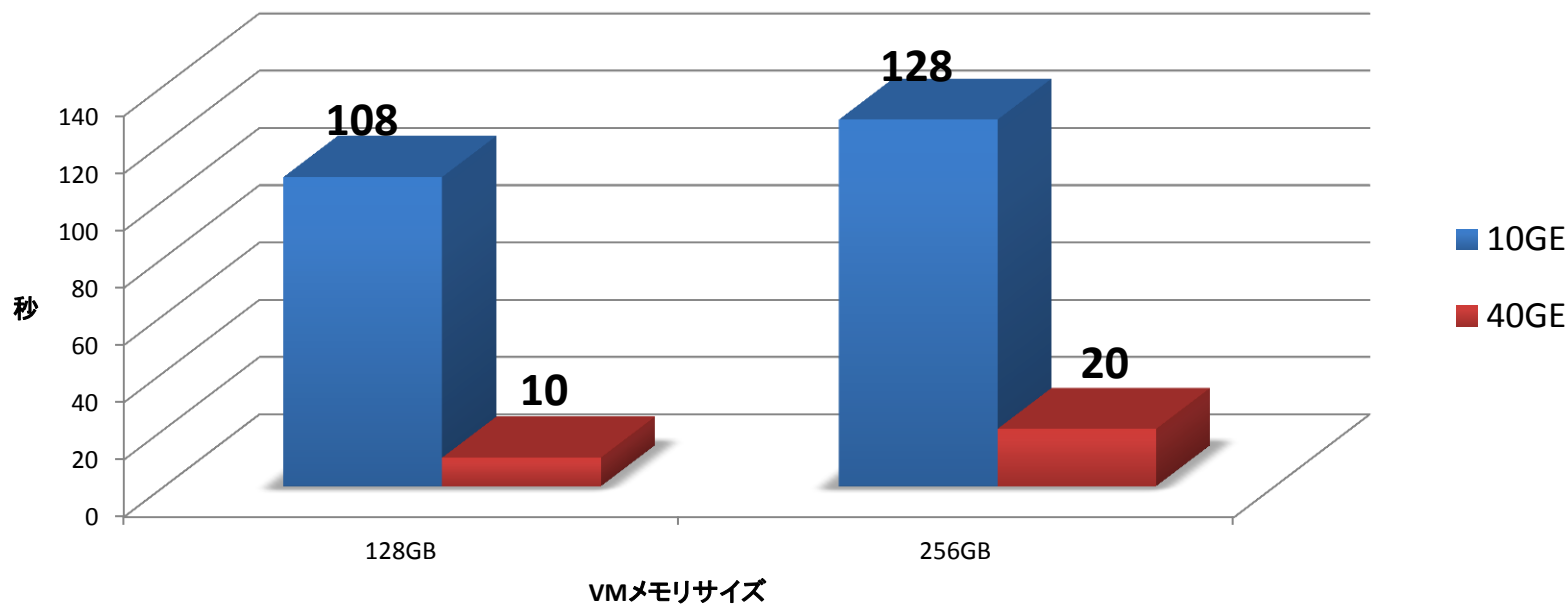
▶ vMotion時間の測定

- ◆ vSphereのvMotion機能を使用し、仮想マシンの移行時間を測定します。
- ◆ ネットワーク帯域が10GEと40GEのときの実行時間の測定結果をまとめます。

▶ 仮想マシンのスタン時間測定について

- ◆ VMにメモリ負荷を掛けた状態でvMotionを実行します。
- ◆ vMotion移行中に仮想マシン上で操作ができなくなってしまう時間を測定します。(スタン時間)
 - 補足：仮想マシンのスタンはアプリケーションから見るとサービス停止状態となります

vMotion実行時間



	128GB/10G	128GB/40G	256GB/10G	256GB/40G
vMotion (Sec)	108	10	128	20
Stun (Sec)	7	<1	19	<1

※検証環境、実行条件により数値は変動する場合があります

▶ 結果まとめ

◆ ドライバ認証検証

- inboxドライバでConnectX-3 Proが問題なく認識し動作することが確認できた。

◆ vMotion時間及びスタン時間の検証

- 40GEネットワークの場合、vMotion実行時間が10GEに比べ約1/10となる結果が得られた。
- 40GEネットワークの場合、スタン時間が1秒未満となる結果が得られた。

▶ 考察

- ◆ Mellanoxの40GEネットワークは10GEに比べ、vMotion実行時間及びVMのスタン時間を短縮することができるソリューションであることが分かった。
- ◆ ConnectX-3 Proでサポートしている、VXLAN offloadやRDMAを使用することがさらなる性能向上を期待することができる。

ご不明点、ご質問がありましたら下記までお問い合わせください。

株式会社アルティマ

Mail : mellanox-support@altima.co.jp

TEL : 045-476-2197

FAX : 045-476-2198