BladeSymphony と Hitachi Storage Solutions を利用した Microsoft System Center Virtual Machine Manager 2012(RC 版)による 動的最適化機能の検証とその評価

第 1.0 版 2012 年 03 月 30 日

株式会社日立製作所 プラットフォームソリューション事業部

変更履歴

項番	版数	内容	更新日
1	1.0 版	新規作成	2012年03月30日

目次

1.	はじ	めに	1
2.	機能	概要	3
3.	通常	運用時におけるユーザーへの影響調査	9
;	3.1.	本章の概要	9
;	3.2.	検証環境(システム構成)	9
;	3.3.	検証環境(動的最適化設定)	12
,	3.4.	検証項目および方法	13
;	3.5.	検証結果	14
,	3.6.	本章のまとめ	18
4.	障害	'時におけるユーザーへの影響調査	19
	4.1.	本章の概要	19
	4.2.	検証環境(システム構成)	19
	4.3.	検証環境(動的最適化設定)	22
	4.4.	移動元サーバー障害が与える影響調査	23
	4.4.	1. 検証項目および方法	23
	4.4.2	2. 検証結果	25
	4.5.	移動先サーバー障害が与える影響調査	27
	4.5.	1. 検証項目および方法	27
	4.5.2	2. 検証結果	29
	4.6.	本章のまとめ	30
5.	まと	めと考察	31
6.	文献	:目録	32
付	録 1.	SCVMM2012 を用いた環境における監視項目	33

用語および略号

Hyper-V	マイクロソフトが提供する仮想化技術の名称。
IOPS	Input Output Per Second:
	1 秒間の入出力処理の回数。
MSFC	Microsoft Failover Cluster:
	Windows Server 2008 のクラスタ技術。
SCVMM	System Center Virtual Machine Manager:
	マイクロソフトが提供する仮想化サーバーの管理ソフトウェア。
ライブマイグレーション	移動対象の仮想マシンのメモリを移動先の物理サーバーに事前に複製する
	ことで、ダウンタイムを発生させずに実行中の仮想マシンを別の物理サー
	バへ移動できる機能。

登録商標および商標について

Microsoft、Windows、Windows Server、System Center、Virtual Machine Manager、Hyper-Vは米国 Microsoft Corporation の米国およびその他の国における登録商標または商標です。

その他、このドキュメントで記載する製品名および会社名は、各社の商標または登録商標です。本文中では、® および ™ は明記しておりません。

1. はじめに

昨今、サーバーの仮想化技術の向上に伴い、クライアントの仮想化も普及の段階に入ってきており、徐々に本番システムにおいても仮想化の波が押し寄せています。そのような状況の中、本番システムのクライアントを仮想化するにあたって解決すべき課題が次第に明らかになってきました。主な課題を以下に挙げます。

● 構成管理

構成管理とは、仮想クライアントのライフサイクル管理のことを指します。仮想化された環境では、従来の物理的なクライアントが仮想クライアントとしてソフトウェアの形式で扱われるため、仮想クライアントの新規作成から構成変更、削除まで、より柔軟かつ迅速な運用が可能になります。また、仮想化の特徴のひとつでもある可搬性の高さも構成管理の運用に影響します。このように構成管理の運用の方式が物理環境とは異なるため、どのような運用が最適か検討する必要があります。

● 監視

仮想化機構(Hyper-V)により物理的なクライアントはエミュレートされるため、従来の監視項目以外も監視する必要があります。また複数の仮想クライアントが1台の物理サーバーに搭載されるため、物理サーバーの障害時の仮想クライアントへの影響を考慮する監視運用の重要性が高まります。

● 可用性

従来の物理的なクライアントでは可用性を考慮していませんでした。仮想化された環境では、物理サーバー が障害で停止すると、その物理サーバーに搭載されている複数の仮想クライアントも停止することになるため、 仮想クライアントの可用性の確保が課題となります。

● バックアップ

従来の物理的なクライアントでは、個別のバックアップを定期的に取得する運用を行うことは多くありません。 しかし、仮想化されることによって、仮想環境特有のバックアップ方式を取ることができ、バックアップ運用方式 のバリエーションが広がります。一方、仮想環境では複数の仮想クライアントで1台の物理サーバーのリソース を共有するため、1台の仮想クライアントのバックアップがその他の仮想クライアントの性能に直接影響を及ぼ します。そのような違いを考慮し、実施の有無を含め最適なバックアップ方式を検討する必要があります。

近日リリースが予定されているマイクロソフトの System Center Virtual Machine Manager 2012(以下、SCVMM2012)においてもこれらの課題を解決することができます。SCVMM2012 の新機能の一つとして可用性向上を目的とした動的最適化機能があります。今回の検証では、動的最適化機能を利用することで、仮想化されたクライアント環境を安定してユーザーに提供するための検証を実施しました。

本書は、Windows Server 2008 R2の Hyper-V サーバー仮想化テクノロジーを活用した効果的なクライアントの仮想化を検討している企業やエンジニアを対象に以下の情報を提供することを目的としています。

- 動的最適化機能概要 動的最適化の機能概要と、SCVMM2012の管理コンソール上で設定する設定値について説明します。
- 通常運用時におけるユーザーへの影響通常運用時に、動的最適化機能が仮想クライアントを利用するユーザーに影響を与えないことを確認します。
- 障害時におけるユーザーへの影響障害時に、動的最適化機能が仮想クライアントを利用するユーザーに与える影響を確認します。
- 障害時における復旧の方法 物理サーバーが停止した際の障害復旧の方法を確認します。

本ホワイトペーパーは、マイクロソフト大手町テクノロジーセンター内に設置した「日立-マイクロソフト総合検証センター」にて、株式会社日立製作所と日本マイクロソフト株式会社の共同で実施した検証に基づき執筆しております。

本検証では、プラットフォームとしてBladeSymphony BS320(以下、BS320)およびHitachi Adaptable Modular Storage 2300(以下、AMS2300)、SCVMM2012は製品候補版(RC)を利用しております。 SCVMM2012はRC版を利用しているため、製品版では本書と仕様が異なる可能性があります。 あらかじめご 了承ください。

本ホワイトペーパーに記載する内容は、弊社環境にて実施した検証結果に基づいており、実運用環境下での動作および性能を保証するものではありません。あらかじめご了承ください。

2. 機能概要

SCVMM2012 の動的最適化機能とは、クラスタを構成するノード間で、自動的に負荷が分散されるようにリソースのバランスを取りながら仮想クライアントを動的にライブマイグレーションする機能です。動的最適化が実行されたときの仮想クライアントの動きの概要を図 2.1~2.4 に記載します。

① Node1、Node2 に仮想クライアントが集中して配置されており、負荷が偏っている状態です。

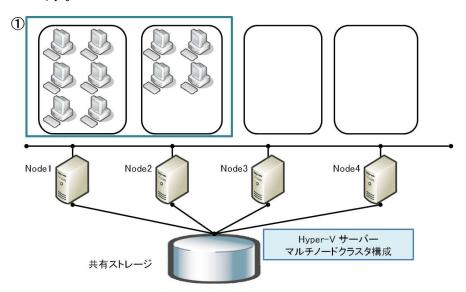


図 2.1 動的最適化イメージ図①

② 負荷が偏っている物理サーバーから、動的最適化対象の仮想クライアントが選択されます(選択される仮想クライアントは負荷状況などによって変更されます)。

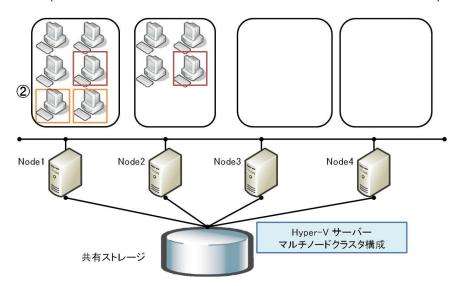


図 2.2 動的最適化イメージ図②

③ 動的最適化対象の仮想クライアントが順次移動します。

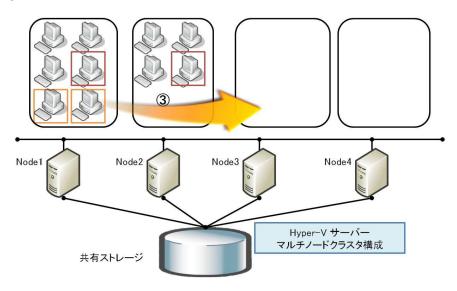


図 2.3 動的最適化イメージ図③

④ 動的最適化対象の仮想クライアントの移動が完了し、Node1~Node4 で負荷分散された状態になります。

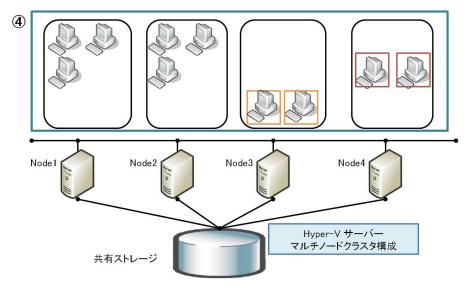


図 2.4 動的最適化イメージ図④

SCVMM2012 において実行する処理の単位を「ジョブ」と呼びます。動的最適化はジョブの一つとして扱われます。通常運用時のジョブの状態推移を図 2.5 に示します。動的最適化のジョブ実行中はステータスが「移動中」になり、動的最適化ジョブが正常に終了すると、ステータスが「完了」になります。なお、動的最適化のジョブ実行中は、動的最適化対象クライアントに対する他のジョブを受け付けません。

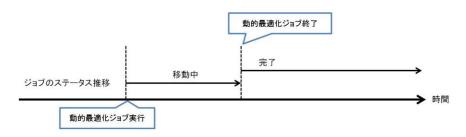


図 2.5 通常運用時における動的最適化ジョブの状態推移

動的最適化の設定範囲について、説明します。動的最適化の設定は、親ホストグループもしくは、子ホストグループのグループ単位で適用することができます。図 2.6 に親ホストグループと子ホストグループの考え方を示します。子ホストグループとは、物理サーバーをまとめたものを示します。親ホストグループとは、子ホストグループをまとめたものを示します。なお、設定値は、親ホストグループ、子ホストグループの設定でき、どちらかの設定のみが適用されます。

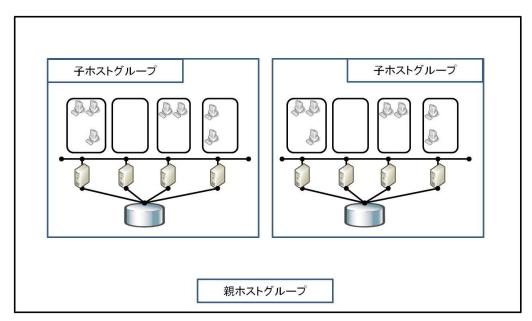


図 2.6 親ホストグループと子ホストグループの考え方

動的最適化の具体的な設定について、説明します。

・設定範囲(子ホストグループでのみ設定可能)

動的最適化の設定は、親ホストグループか子ホストグループのどちらかの設定のみが適用されます。初期の設定では、子ホストグループの設定が反映され、親ホストグループの設定は適用されません。子ホストグループにおいて、親ホストグループの設定を引き継ぐ場合、『親ホストグループの動的最適化を設定する』を選択する必要があります。

•強度

強度は、「高・中・低」を指定できます。「高」を設定した場合、高い効果が得られない場合でも動的最適化を実行します。そのため、ライブマイグレーションの頻度が高くなります。「低」を指定した場合、高い効果が得られる場合に動的最適化を実行します。そのため、ライブマイグレーションは少なくなります。「中」はその中間となります。初期設定は「中」が設定されます。

動的最適化の自動実行間隔

動的最適化を自動的に実行する間隔を設定します。初期値では 10 分です。設定範囲は 10 分~1440 分で す(文献目録 [1])。

・しきい値

しきい値は移動先サーバーの残りリソースが対象となります。動的最適化に用いる各しきい値の初期値と設 定範囲を表 2.1 に示します。

設定項目	初期値(設定範囲)
CPU *1	30% (10%~100%)
メモリ ※2	512MB(256MB~MAX)
ディスク I/O ※3	0 IOPS(0∼MAX)
ネットワーク I/O ※4	0% (0~100%)

表 2.1 動的最適化設定(しきい値)の初期値と設定範囲

※1 しきい値を初期値の 30%とすると、移動先サーバーにおいて CPU の使用率が 70%を下回っている間、動的最適化の移動先対象として選ばれます。

※2 しきい値を初期値の 512MB とすると、移動先サーバーの搭載するメモリの総量(32GB と仮定)から、物理サーバーで使用するメモリ容量(1GBと仮定)としきい値の 512MB を引いた値の 30.5GB まで、仮想クライアントに割り当てることができます。物理サーバーの残りメモリ容量が 30.5GB を下回っている間、動的最適化の移動先対象として選ばれます。

※3,4 ディスク I/O 及びネットワーク I/O を初期値とすると、常に動的最適化の移動先対象として選ばれます。 しきい値を設定するときの考え方は、※2 で述べたメモリと同じになります。

CPU、メモリ、ディスク I/O、ネットワーク I/O のうち 1 つでもしきい値を超えたノードは、移動先の対象から外れます。しきい値を超えないノードが、移動先の対象となります。また、クラスタを構成するすべてのノードのリソースが均等になるように配置されるため、動的最適化の対象となったノードでも仮想クライアントが移動しない場合もあります。

移動先対象の選別について、図 2.7 を用いて説明します。しきい値を初期値とした場合、移動元ノード (Node1)から見て、移動先ノード (Node2)の残りリソースがしきい値を超えていないため、移動先対象となります。SCVMM2012 の動的最適化は、リソースが均等になるように仮想クライアントが移動するため、Node1 の 仮想クライアントが Node2 へ移動します。

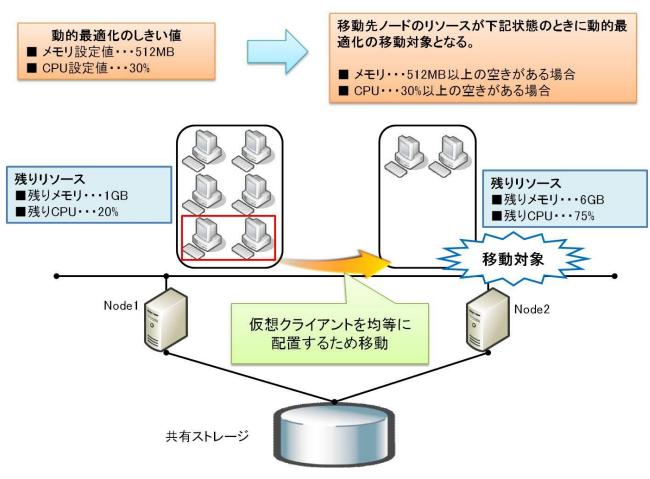


図 2.7 動的最適化移動対象の選別

参考として、表 2.2 に親ホストグループおよび、子ホストグループを設定する場合の設定値を記載します。動的最適化の設定は、親ホストグループか子ホストグループのどちらかの設定のみが適用されます。初期の設定では、子ホストグループの設定が反映され、親ホストグループの設定は適用されません。子ホストグループにおいて、親ホストグループの設定を引き継ぐ場合、『親ホストグループの動的最適化を設定する』を選択する必要があります。

表 2.2 動的最適化の設定値

	衣 2.2 到的最過100設定但	
設定項目	親ホストグル一プの設定を 使用する場合	子ホストグループの個別の設定を 使用する場合
親ホストグループの設定		
強度	中	中
負荷分散するバーチャルマシン	10	10
を自動的に移行する間隔(分)		
CPU[%]	30	30
メモリ[MB]	512	512
ディスク I/O[IOPS]	0	0
ネットワーク I/O[%]	0	0
子ホストグループの設定		
親ホストグループの動的最適化 設定を使用する	使用する	使用しない(初期設定)
強度	中	中
負荷分散するバーチャルマシン	10	10
を自動的に移行する間隔(分)		
CPU[%]	30	30
メモリ[MB]	512	512
ディスク I/O[IOPS]	0	0
ネットワーク I/O[%]	0	0

3. 通常運用時におけるユーザーへの影響調査

3.1. 本章の概要

本章では、SCVMM2012 の新機能である動的最適化機能を検証し、仮想クライアントを使用するユーザーへの影響を調査します。

なお、本検証に用いる環境の構築手順は省略します。

3.2. 検証環境(システム構成)

本検証では図 3.1 に示す検証環境を構築しました。図 3.1 の論理構成は、ハードウェアリソースを有効に活用するように考え、物理サーバー数を少なくする構成にしています。従って、検証に必要なサーバーを全て管理サーバー上に集約し、仮想サーバーとして構築しています。本検証では仮想化ソフトとして、Hyper-V を用います。また、クラスタノード 1 とクラスタノード 2 をクラスタ構成にするために MSFC を用います。仮想クライアントは、SCVMM2012 を用いてクラスタ上に構築します。

本検証で用いる仮想サーバーと各仮想サーバーに追加が必要となる役割を以下に述べます。

- SCVMM サーバー SCVMM2012(VMM サーバー、管理コンソール)をインストールし、Hyper-V と仮想クライアントを管理します。
- AD サーバー SCVMM2012 の前提としてドメインコントローラーと DNS サーバーを構築する必要があるため、[Active Directory ドメイン サービス]と[DNS サーバー]の役割を追加します(文献目録 [2])。

図 3.1 に記載したサーバーとクライアントのハードウェア構成を表 3.1 に、ソフトウェア構成を表 3.2 に示します。 仮想クライアントは、Windows 7 をインストールした後に Hyper-V の統合サービスをインストールした状態であり、仮想クライアントの環境を同一の状態にします。

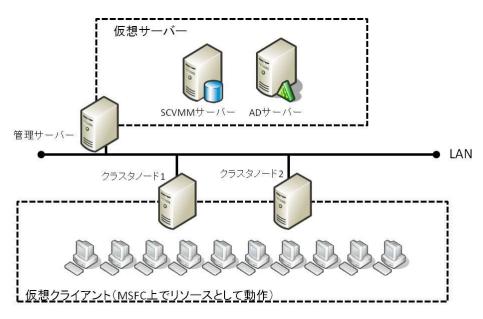


図 3.1 論理構成図

表 3.1 ハードウェア構成

	対象	ハードウェア構成		
管理サーバー		【モデル】BS320 A4		
		【CPU】QuadCore Xeon E5520 2.26GHz (4 コア×2)		
		【メモリ】 32GB RAM		
		【内蔵ディスク】SAS 73GB×2		
	SCVMM サーバー	【CPU】2 コア		
		【メモリ】8GB RAM		
		【ディスク容量】100GB		
	AD サーバー	【CPU】2 コア		
		【メモリ】 2GB RAM		
		【ディスク容量】60GB		
クラス	スタノード 1,2	【モデル】BS320 A3		
		【CPU】QuadCore Xeon E5520 2.26GHz (4 コア×2)		
		【メモリ】8GB RAM		
		【内蔵ディスク】SAS 73GB×2		
	仮想クライアント	【CPU】1 コア		
		【メモリ】 1GB RAM		
		【ディスク容量】20GB		

表 3.2 ソフトウェア構成

	対象	ソフトウェア構成		
管理	サーバー	•Windows Server 2008 R2 Enterprise Edition Service Pack1		
	SCVMM サーバー	Windows Server 2008 R2 Enterprise Edition Service Pack1 NET Framework 3.5 SP1		
		•Windows Power Shell 2.0		
		•Windows Remote Management (WinRM) 2.0		
		•SQL Server 2008 R2 Enterprise Edition		
		・System Center Virtual Machine Manager2012(RC 版)		
	AD サーバー	•Windows Server 2008 R2 Enterprise Edition Service Pack1		
クラスタノード 1,2		·Windows Server 2008 R2 Enterprise Edition Service Pack1		
		Microsoft System Center Virtual Machine Manager Agent (x64)		
	仮想クライアント	•Windows7 Professional Edition Service Pack1		

3.3. 検証環境(動的最適化設定)

本検証における動的最適化の設定値を表 3.3 に示します。なお、設定値が「←」の欄は、初期値を設定値として利用していることを表しています。また、網掛けの部分は、子ホストグループの設定を利用するため、設定値が無視される項目です。

表 3.3 設定値

	公 5.5 					
	設定項目	初期値				
親	ホストグループの設定					
	強度	中	←			
	負荷分散するバーチャルマシンを自動的に移	10	←			
	行する間隔(分)					
	CPU[%]	30	←			
	メモリ[MB]	512	←			
	ディスク I/O[IOPS]	0	←			
	ネットワーク I/O[%]	0	←			
	子ホストグループの設定					
	親ホストグループの動的最適化設定を使	使用しない	←			
	用する	使用しない				
	強度	中	強			
	負荷分散するバーチャルマシンを自動的	10	1440(最大値)			
	に移行する間隔(分)					
	CPU[%]	30	←			
	メモリ[MB]	512	←			
	ディスク I/O[IOPS]	0	←			
	ネットワーク I/O[%]	0	←			

本検証では、基本的に初期値から変更しない方針としますが、一部のみ初期値を変更します。なお、変更する理由は以下のとおりです。

● 強度(中⇒高)

動的最適化の再現性を高めるために、強度を「強」に設定します。

● 負荷分散するバーチャルマシンを自動的に移行する間隔(10分⇒1440分) 動的最適化は手動で実施するため、最大値を設定します。なお、動的最適化を手動で実行しても、自動で実行 しても同様の結果となります。

3.4. 検証項目および方法

本検証では、動的最適化が仮想クライアントを使用しているユーザーへ影響が無いことを確認するため、検証項目 I ~IVを検証します。ユーザーがページの表示を待つのに我慢できる時間が 8 秒であり、それを越えると大部分のユーザーが我慢できずに異なるページへ移動するという調査結果があります(文献目録 [3])。近年の回線速度高速化を考慮し、ユーザーが我慢できる時間は、8 秒より短くなっていると予測できます。よって、本書では、3 秒をユーザーへ影響が無い時間と仮定しました。

- 検証項目I. ユーザーがセッションを繋いでいる場合のセッション断絶時間(リモートデスクトップが接続された状態での、ping 断絶時間)
- 検証項目II. ユーザーがセッションを繋いでいない場合のセッション断絶時間(リモートデスクトップが接続されていない状態での、ping 断絶時間)

【個別検証設定】

仮想クライアントには、シンクライアントで使用される可能性が高い 2GB、3GB、4GB のメモリを搭載します。 【検証方法】

セッション断絶時間は、SCVMM サーバーから ping コマンドを 1 秒間隔で実行し、その戻り値から測定しました。1 秒間隔で ping コマンドの戻り値を測定していますので、戻り値が返ってこない期間をセッション断絶時間とします。以下の ping コマンドを実行し、メモリ搭載量ごとに 5 回ずつ検証します。

ping△-t△-w△1△仮想クライアントの IP アドレス

上記コマンドの△はスペースを示している。

検証項目Ⅲ.動的最適化の処理時間と仮想クライアントに割り当てられたメモリとの関連

【個別検証設定】

仮想クライアントには、シンクライアントで使用される可能性が高い 2GB、3GB、4GB のメモリを搭載します。 【検証方法】

メモリ搭載量ごとに、動的最適化の実行時間を測定し、動的最適化の処理時間と仮想クライアントに割り当てられたメモリとの関連を明確にします。動的最適化の処理時間は、SCVMM2012の管理コンソール上で、動的最適化対象クライアントごとに、動的最適化の開始時間と終了時間から算出します。メモリ搭載量ごとに 5 回ずつ検証します。

検証項目IV. ユーザー作業中における仮想クライアントへの影響

【個別検証設定】

検証項目IIIと比較するため、仮想クライアントには、2GB のメモリを搭載します。

【検証方法】

ユーザー作業中に動的最適化が実行された場合を想定し、Office2010 インストール作業中に動的最適化を 実施します。動的最適化実行後に、正常にインストールが完了することを確認します。また、Office2010 をイン ストールしていない場合(検証項目皿)と比較するため、動的最適化の処理時間を算出します。

3.5. 検証結果

検証項目I. ユーザーがセッションを繋いでいる場合のセッション断絶時間(リモートデスクトップが接続された状態での、ping 断絶時間)

リモートデスクトップが接続された状態での ping 断絶時間の結果を、表 3.4 に示します。

表 3.4 仮想クライアントに対する搭載メモリごとの ping 断絶時間

搭載メモリ	1 回目	2 回目	3 回目	4 回目	5 回目
2GB	1 秒	0 秒	0 秒	0 秒	1 秒
3GB	1 秒	0 秒	1 秒	0 秒	1 秒
4GB	1 秒	0 秒	0 秒	1 秒	1 秒

(リモートデスクトップで接続している場合)

表 3.4 より、ping 断絶時間が 3 秒以下であるため、ユーザーがセッションを繋いでいる場合、動的最適化が 仮想クライアントを使用するユーザーに影響を与えないことを確認できました。

検証項目II. ユーザーがセッションを繋いでない場合のセッション断絶時間(リモートデスクトップが接続されていない状態での、ping 断絶時間)

リモートデスクトップが接続されていない状態での ping 断絶時間の結果を表 3.5 に示します。

表 3.5 仮想クライアントに対する搭載メモリごとの ping 断絶時間

搭載メモリ	1 回目	2 回目	3 回目	4 回目	5 回目
2GB	9 秒	2 秒	8 秒	29 秒	3 秒
3GB	25 秒	3 秒	14 秒	2 秒	3 秒
4GB	5 秒	3 秒	1 秒	13 秒	1 秒

(リモートデスクトップで接続していない場合)

表 3.5 より、セッションを繋いでいない状態だと、ping 断絶時間が 1 秒で完了する場合もありましたが、 30 秒近く伸びる場合もあり、規則性がありませんでした。ping が断絶したタイミングを図 3.2 に示します。 図 3.2 より、SCVMM2012 上の動的最適化ジョブが完了する直前から ping が断絶しています。注意点としては、ping 断絶期間中にリモートデスクトップを接続しても、ping が通るまでリモートデスクトップが繋がらないことです。セッションを繋いでいる状態での ping 断絶時間(検証項目 I)を考慮すると、セッションを繋いでるかを判断して、動的最適化を実行していることが分かります。これは、ユーザーが仮想クライアントにセッションを繋いでいる場合、ユーザーへの影響が少なくなるよう断絶時間を短くしているが、セッションを繋いでいない場合、ユーザーへの影響を考える必要が無いため、断絶時間を短くする必要が無いとシステム上で判断していると考えられます。ただし、セッションを繋いでいない場合、リモートデスクトップに30 秒程度繋がらなくても、元々セッションを繋いていないため、ほぼ問題無いと判断できます。

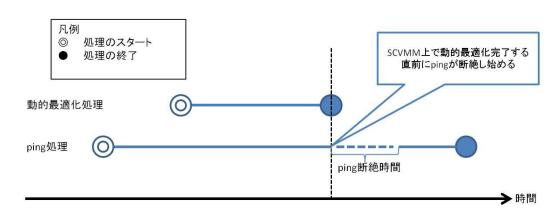


図 3.2 ping 断絶のタイミング図

検証項目Ⅲ. 動的最適化の処理時間と仮想クライアントに割り当てられたメモリとの関連

メモリ搭載量ごとの動的最適化処理時間を表 3.6 に示します。表 3.6 より、メモリ搭載量を増やすと、動的最適化処理時間も増加することが分かります。動的最適化の処理時間とメモリ搭載量の関係を図 3.3 に示します。図 3.3 より、動的最適化の処理時間とメモリ搭載量は比例関係であることが分かります。動的最適化はライブマイグレーションを用いており、メモリ領域を移動先サーバーに複製した後にセッションを切り替えます。メモリ搭載量により移動先へ複製する処理時間が増えるため、処理時間とメモリ搭載量は比例関係になります。

搭載メモリ	1 回目	2 回目	3 回目	4 回目	5 回目	平均
2GB	26 秒	26 秒	25 秒	24 秒	26 秒	25 秒
3GB	35 秒	35 秒	34 秒	35 秒	35 秒	35 秒
4GB	45 秒	46 秒	43 秒	43 秒	43 秒	44 秒

表 3.6 メモリ搭載量ごとの動的最適化処理時間

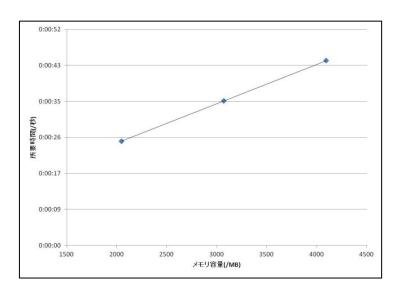


図 3.3 動的最適化の処理時間とメモリ搭載量との関係図

仮想クライアントの搭載メモリ容量が、2GB、3GB、4GB の場合、検証項目 I より、メモリ搭載量が増加しても、セッション断絶時間は 1 秒程度であることが分かっています。また、検証項目皿より、メモリ搭載量が増加した場合、処理時間は増加することが分かっています。以上より、メモリ搭載量が増加した場合、処理時間は増加するが、セッション断絶時間は 1 秒程度であるため、メモリ搭載量に関わらずユーザーに影響が無いことを確認できました。

検証項目IV. ユーザー作業中における仮想クライアントへの影響

インストール作業中に動的最適化が実施してもインストール作業に影響無く、正常終了しました。また、Word や Excel を開いて使用しましたが、問題無く使用することができました。

動的最適化の処理時間は、検証項目皿と比べると、10 秒ほど延びました。インストール作業では、動的最適化中にメモリの内容が更新されるため、処理時間が延びることが確認できました。ping 断絶時間に関しては、1 秒程度であったため、仮想クライアントを操作するユーザーに影響が無いことが確認できました。

3.6. 本章のまとめ

検証を通じて、メモリ搭載量や作業の有無に関わらず、動的最適化中のセッション断絶時間が 3 秒以下であるため、仮想クライアントを使用するユーザーへ影響を与えないことを確認しました。

検証を通じて、動的最適化が図 3.4 の流れで処理することが分かりました。

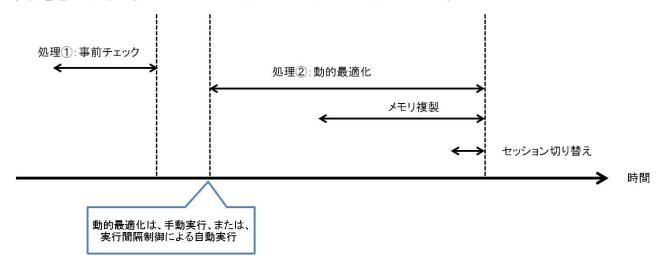


図 3.4 動的最適化の処理の流れ

処理①. 事前チェック

移動元サーバーから見て、残りリソースがしきい値を超えた物理サーバーは、移動先対象となります。 また、リソースが均等になるように移動対象の仮想クライアントを決定します。移動対象の仮想クライアントが存在する場合、動的最適化の実行が決定します。

処理②. 動的最適化

処理①の事前チェックにより、動的最適化の実行が決定した場合、処理②に移動します。動的最適化は、手動実行、または、実行間隔制御により自動実行します。動的最適化では、主にメモリ複製とセッション切り替えを実行します。メモリ複製時間は、メモリ搭載量とユーザーの作業量に連動して増減します。セッションを切り替える時に、セッションが 1 秒程度断絶します。

今回の検証では、ディスク I/O とネットワーク I/O を動的最適化の実行条件から対象外としました。ディスク I/O または、ネットワーク I/O が多く発生する運用(仮想クライアントのウイルス対策ソフトが同時に実行される場合など)が想定される場合は、最適にサイジングする必要があります。

動的最適化の実行条件に設定する4つのしきい値(CPU、メモリ、ディスクI/O、ネットワークI/O)は、移動先の残りリソースを設定します。SCVMM2012(RC版)に記載されている文章では、「リソースがしきい値まで達した場合に、動的最適化が実行される」と読むこともできますので、意味の取り違いに注意する必要があります。

4. 障害時におけるユーザーへの影響調査

4.1. 本章の概要

動的最適化実行中に移動元または移動先のサーバーが障害によって停止した場合を想定し、仮想クライアントを 利用しているユーザーへの影響と障害復旧の方法を確認します。

検証は、メモリ複製中に移動元または移動先のサーバーをシャットダウンまたは電源オフさせて、仮想クライアントを利用しているユーザーへの影響および障害復旧方法を確認します。シャットダウンは、オペレーションミスを想定しています。また、電源オフは、ストレージやブレード等のハードウェア障害を想定しています。さらに、検証から分かった障害時の復旧方法の一例を記載します。

4.2. 検証環境(システム構成)

本検証では図 4.1 に示す検証環境を構築しました。図 4.1 の論理構成は、ハードウェアリソースを有効に活用するように考え、物理サーバー数を少なくする構成にしています。従って、検証に必要なサーバーを全て管理サーバー上に集約し、仮想サーバーとして構築しています。本検証では仮想化ソフトとして、Hyper-V を用います。また、クラスタノード 1~4 をクラスタ構成にするために MSFC を用います。仮想クライアントは、SCVMM2012 を用いてクラスタ上に構築します。

本検証で用いる仮想サーバーと各仮想サーバーに追加が必要となる役割を以下に述べます。

- SCVMM サーバー SCVMM2012(VMM サーバー、管理コンソール)をインストールし、Hyper-V と仮想クライアントを管理します。
- AD サーバー SCVMM2012 の前提としてドメインコントローラーと DNS サーバーを構築する必要があるため、[Active Directory ドメイン サービス]と[DNS サーバー]の役割を追加します(文献目録 [2])。

図 4.1 に記載したサーバーとクライアントのハードウェア構成を表 4.1 に、ソフトウェア構成を表 4.2 に示します。 仮想クライアントは、Windows7 をインストールした後に Hyper-V の統合サービスをインストールした状態であり、仮想クライアントごとの環境を同一の状態にします。

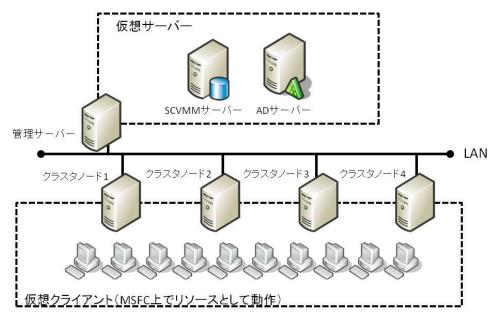


図 4.1 論理構成図

表 4.1 ハードウェア構成

	対象	ハードウェア構成		
管理サーバー		【モデル】BS320 A4		
		【CPU】QuadCore Xeon E5520 2.26GHz (4 コア×2)		
		【メモリ】 8GB RAM		
		【内蔵ディスク】SAS 73GB×2		
	SCVMM サーバー	【CPU】2 コア		
		【メモリ】8GB RAM(静的)		
		【ディスク容量】100GB		
	AD サーバー	【CPU】2 コア		
		【メモリ】2GB RAM(静的)		
		【ディスク容量】60GB		
クラス	スタノード 1,2,3,4	【モデル】BS320 A3		
		【CPU】QuadCore Xeon E5520 2.26GHz (4 コア×2)		
		【メモリ】 8GB RAM		
		【内蔵ディスク】SAS 73GB×2		
	仮想クライアント	【CPU】1 コア		
		【メモリ】1GB RAM(静的)		
		【ディスク容量】20GB		

表 4.2 ソフトウェア構成

対象		ソフトウェア構成	
管理サーバー		-Windows Server 2008 R2 Enterprise Edition Service Pack1	
	SCVMM サーバー	·Windows Server 2008 R2 Enterprise Edition Service Pack1 ·.NET Framework 3.5 SP1 ·Windows Power Shell 2.0 ·Windows Remote Management (WinRM) 2.0 ·SQL Server 2008 R2 Enterprise Edition ·System Center Virtual Machine Manager2012(RC 版)	
	AD サーバー	•Windows Server 2008 R2 Enterprise Edition Service Pack1	
クラスタノード 1,2,3,4		Windows Server 2008 R2 Enterprise Edition Service Pack1 Microsoft System Center Virtual Machine Manager Agent (x64)	
	仮想クライアント	·Windows7 Professional Edition Service Pack1	

4.3. 検証環境(動的最適化設定)

本検証における動的最適化の設定値を表 に示します。なお、設定値が「←」の欄は、初期値を設定値として利用 していることを表しています。なお、網掛けの部分は、子ホストグループの設定を利用するため、設定値が無視され る項目です。

表 4.3 動的最適化の設定値

	設定項目	初期値	設定値
親	ホストの設定		
	強度	中	←
	負荷分散するパーチャルマシンを自動的に移	10	←
	行する間隔(分)		
	CPU[%]	30	←
	メモリ[MB]	512	←
	ディスク I/O[IOPS]	0	←
	ネットワーク I/O[%]	0	←
	子ホストの設定		
	親ホストグループの動的最適化設定を使	使用しない	←
	用する		
	強度	中	強
	負荷分散するパーチャルマシンを自動的	10	1440(最大値)
	に移行する間隔(分)		
	CPU[%]	30	←
	メモリ[MB]	512	←
	ディスク I/O[IOPS]	0	←
	ネットワーク I/O[%]	0	←

本検証では、基本的に初期値から変更しない方針としますが、一部のみ初期値を変更します。なお、変更する理由は以下のとおりです。

● 強度(中⇒高)

動的最適化の再現性を高めるために、強度を「強」に設定します。

● 負荷分散するバーチャルマシンを自動的に移行する間隔(10分⇒1440分) 動的最適化は手動で実施するため、最大値を設定します。なお、動的最適化を手動で実行しても、自動で実行 しても同様の結果となります。

4.4. 移動元サーバー障害が与える影響調査

本節では、移動元サーバーにおいて障害発生を想定した検証を実施します。検証では、メモリ複製中に移動元の サーバーをシャットダウンまたは電源オフさせて、仮想クライアントを使用しているユーザーへの影響および障害復 旧方法を確認します。

4.4.1. 検証項目および方法

以下に、検証手順を図 4.2 および図 4.3 を用いながら示します。

手順Ⅰ~Ⅳに、障害発生の手順を示します。

手順!. 動的最適化実行環境の作成(図 4.2 の①)

クライアントノード1で1台の仮想クライアントを立ち上げ、動的最適化が発生する環境を準備します。

手順II. 仮想クライアントへのセッション確立(図 4.2 の②)

動的最適化対象の仮想クライアントにリモートデスクトップ(RDP)にてアクセスします。

手順III. 動的最適化実行(図 4.2 の③)

SCVMM2012 の管理コンソールから手動で動的最適化を実行します。

手順IV. 障害発生(サーバー停止)(図 4.2 の④)

<シャットダウンの場合>

動的最適化のジョブが完了する前に、移動元サーバーであるクラスタノード 1 を、リモートデスクトップから手動でシャットダウンします。

<電源オフの場合>

動的最適化のジョブが完了する前に、移動元サーバーであるクラスタノード 1 を、ブレードの電源ボタン長押しで強制的に電源オフします。

なお、クラスタノード1に配置されていた仮想クライアントは、MSFCの機能により、他のノードへフェールオーバーされます。

手順V~VIに、障害復旧後の手順を示します。

手順V. 復旧(サーバー起動)(図 4.3 の⑤)

クラスタノード 1 をブレードの電源ボタンを押して起動させます。

手順VI. 動的最適化実行(図 4.3 の⑥)

SCVMM2012 の管理コンソールから手動で動的最適化を実行します。

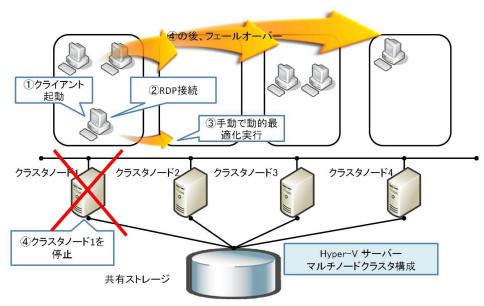


図 4.2 障害(移動元サーバー停止)時の検証手順(停止前)

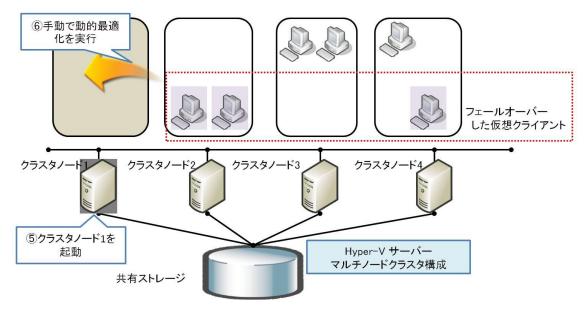


図 4.3 障害(移動元サーバー停止)時の検証手順(停止後)

上記手順を用いて、まず、動的最適化ジョブの状態推移を調査します。動的最適化ジョブが実行状態の間、動的最適化対象クライアントへの操作ができなくなるため、動的最適化対象クライアントに対する復旧作業ができません。よって、動的最適化ジョブが終了するタイミングを明確にする必要があります。次に移動元サーバー障害時におけるユーザーへの影響と障害時の復旧方法を明確にします。最後に、障害対応に必要となるHyper-VとSCVMM2012のイベントログを調査します。

4.4.2. 検証結果

検証結果を以下に示します。

■ 動的最適化ジョブの状態推移

動的最適化ジョブの状態推移を図 4.4 に示します。なお、動的最適化ジョブの状態推移は、移動元サーバー 停止方法がシャットダウンの場合と電源オフの場合で同じ結果になります。

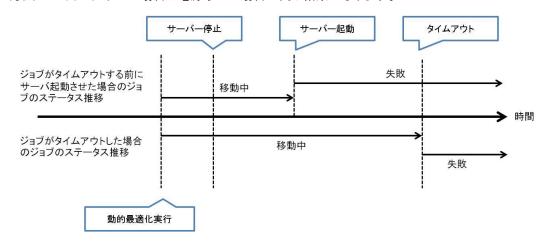


図 4.4 移動元サーバー停止における動的最適化ジョブの推移

図 4.4 より、移動元サーバーが停止しても、ジョブの状態は「移動中」のまま硬直します。タイムアウトすると、 状態が「失敗」に変更され、ジョブが終了します。また、停止したサーバーがタイムアウト前に起動すると、その 時点で状態が「失敗」に変更されます。

■ ユーザーへの影響

<シャットダウンの場合>

10 秒ほどリモートデスクトップ接続が切断されますが、切断は一時的なもので、移動前の状態が移動先サーバーに引き継がれ、作業中の内容に影響はありませんでした。

<電源オフの場合>

仮想クライアントが強制終了し、リモートデスクトップは強制的に終了します。仮想クライアントは、フェールオーバーし、他サーバーで再起動されます。作業中だった内容は失われますが、再起動後は、正常に動作します。

■ 障害時の復旧方法

障害復旧の方法は、移動先サーバー停止方法がシャットダウンの場合と電源オフの場合で同じ結果になります。

移動元サーバーに配置されていた仮想クライアントは、フェールオーバーし、他サーバーで再起動します。再起動した仮想クライアントは MSFC の機能によって移動されたため、SCVMM2012 上では移動したことになっていません。ジョブの状態が「失敗」になった後、ジョブの「再試行」と仮想クライアントの「修復」をすることにより、SCVMM2012 の管理コンソールとフェールオーバークラスタ管理画面とで同じ配置に復元でき、正常に仮想ク

ライアントを管理できるようになります。

起動後は、起動したサーバーに対するジョブが正常に完了するようになります。ただし、起動させたサーバーは、動的最適化の移動先として指定されない状態です。動的最適化の移動先として指定させるためには、起動させたサーバーへ、SCVMM2012 上から仮想クライアントを手動でライブマイグレーションする必要があります。

■ SCVMM2012 と Hyper-V のイベントログ

SCVMM2012とHyper-Vのイベントログは、移動元サーバー停止方法がシャットダウンの場合と電源オフの場合で同じ結果になります。

サーバー停止中には、MSFC、Hyper-V、SCVMM2012 のイベントログにおいて、クラスタを構成するサーバーから停止していることが分かるログが出力されます。また、SCVMM2012 の状態更新ジョブが失敗し続けるため、SCVMM2012 のイベントログには、ジョブの失敗が出力され続けます。

4.5. 移動先サーバー障害が与える影響調査

本節では、移動先サーバーにおいて障害発生を想定した検証を実施します。検証では、メモリ複製中に移動先サーバーをシャットダウンまたは電源オフさせて、仮想クライアントを使用しているユーザーへの影響および障害復旧方法を確認します。

4.5.1. 検証項目および方法

以下に、検証手順を図 4.5 および図 4.6 を用いながら示します。

手順 I ~IVに、障害発生までの手順を示します。

手順I. 動的最適化実行環境の作成(図 4.5 の①)

クライアントノード 1 で 1 台の仮想クライアントを立ち上げ、動的最適化が発生する環境を準備します。

手順II. 仮想クライアントへのセッション確立(図 4.5 の②)

動的最適化対象の仮想クライアントにリモートデスクトップ(RDP)にてアクセスします。

手順Ⅲ. 動的最適化実行(図 4.5 の③)

SCVMM2012 の管理コンソールから手動で動的最適化を実行します。

手順IV. 障害発生(サーバー停止)(図 4.5 の④)

<シャットダウンの場合>

動的最適化のジョブが完了する前に、移動先サーバーであるクラスタノード2を、リモートデスクトップから手動でシャットダウンします。

<電源オフの場合>

動的最適化のジョブが完了する前に、移動先サーバーであるクラスタノード 2 を、ブレードの電源ボタン長押しで強制的に電源オフします。

なお、クラスタノード2に仮想クライアントが配置されていた場合は、フェールオーバークラスタの機能により、 他のノードへフェールオーバーされます。

手順V~VIに、障害復旧後の手順を示します。

手順V. 復旧(サーバー起動)(図 4.6 の⑤)

動的最適化のジョブが終了した後、クラスタノード2をブレードの電源ボタンを押して起動させます。

手順VI. 動的最適化実行(図 4.6 の⑥)

SCVMM2012 の管理コンソールから手動で動的最適化を実行します。

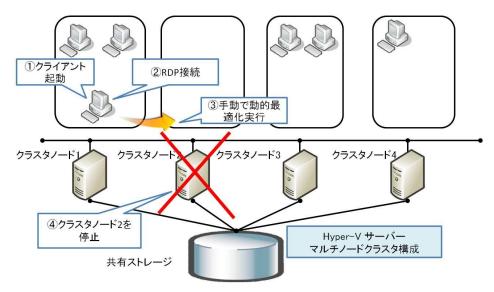


図 4.5 障害(移動先サーバー停止)時の検証手順(停止前)

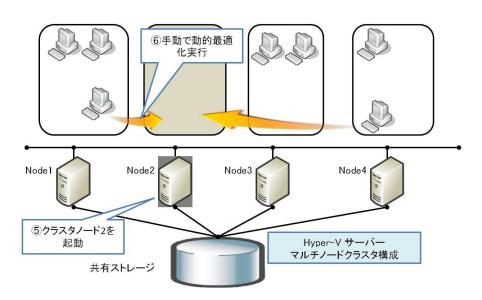


図 4.6 障害(移動先サーバー停止)時の検証手順(停止後)

上記手順を用いて、まず、動的最適化ジョブの状態推移を調査します。動的最適化ジョブが実行状態の間、動的最適化対象クライアントへの操作ができなくなるため、動的最適化対象クライアントに対する復旧作業ができません。よって、動的最適化ジョブが終了するタイミングを明確にする必要があります。次に移動元サーバー障害時におけるユーザーへの影響と障害時の復旧方法を明確にします。最後に、障害対応に必要となるHyper-VとSCVMM2012のイベントログを調査します。

4.5.2. 検証結果

検証結果を以下に示します。

■ 動的最適化ジョブの状態

動的最適化ジョブの状態推移を図 4.7 に示します。なお、動的最適化ジョブの状態推移は、移動先サーバー 停止方法がシャットダウンの場合と電源オフの場合で同じ結果になります。

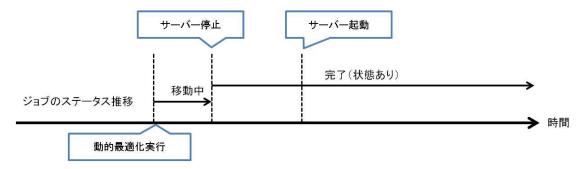


図 4.7 移動先サーバ停止における動的最適化ジョブの状態推移

図 4.7 より、移動先サーバーが停止した時点でジョブは終了し、状態が「完了(状態あり)」に変更されます。 移動先サーバーが停止するため、ジョブが途中で中止し、仮想クライアントは移動しません。

■ ユーザーへの影響

ユーザーへの影響は、移動先サーバー停止方法がシャットダウンの場合と電源オフの場合で同じ結果になります。セッションが断絶しないため、ユーザーへの影響はありません。

■ 障害時の復旧方法

障害復旧の方法は、移動先サーバー停止方法がシャットダウンの場合と電源オフの場合で同じ結果になります。

起動後は、起動したサーバーに対するジョブが正常に完了するようになります。ただし、起動させたサーバーは、動的最適化の移動先として指定されない状態です。動的最適化の移動先として指定させるためには、起動させたサーバーへ、SCVMM2012 上から仮想クライアントを手動でライブマイグレーションする必要があります。

■ イベントログ(SCVMM2012、Hyper-V)の警告とエラー

SCVMM2012 と Hyper-V のイベントログは、移動元サーバー停止方法がシャットダウンの場合と電源オフの場合で同じ結果になります。

サーバー停止中には、MSFC、Hyper-V、SCVMM2012 のイベントログにおいて、クラスタを構成するサーバーが落ちていることが分かるログが多く出力されます。また、SCVMM2012 の状態更新ジョブが失敗し続けるため、SCVMM2012 のイベントログには、ジョブが失敗したことが表示され続けます。

4.6. 本章のまとめ

動的最適化実行中に移動元サーバーまたは移動先サーバーを停止することにより、障害時の仮想クライアント、SCVMM2012 に与える影響を調査しました。検証は、移動元または移動先においてシャットダウンまたは電源オフを実施して、調査しました。検証結果と復旧方法を以下にまとめます。

■ 移動元サーバー障害(ユーザーへの影響)

メモリ複製中の移動元サーバーで障害(電源オフ)が発生した場合、サーバー停止前のメモリ情報は保持されないことがわかりました。ただし、仮想クライアントは、フェールオーバーされた後、他のサーバーで起動されるため、サーバー停止という障害に対して、仮想クライアントを継続して利用することができます。

■ 移動元サーバー障害(障害復旧の方法)

メモリ複製中の移動元サーバーで障害(シャットダウンまたは電源オフ)が発生した場合、フェールオーバーした仮想クライアントはSCVMM2012の管理から外れますが、SCVMM2012上からジョブの再試行および仮想クライアントの修復を実行することにより、再度管理対象にできます。

障害により移動元サーバーが停止すると、起動した後も動的最適化対象から外れる場合があります。 一度、動的最適化対象から外れたサーバーに対して、手動でライブマイグレーションすることにより、動的 最適化対象に加えることができます。

■ 移動先サーバー障害(ユーザーへの影響)

動的最適化の実行が中止になるため、仮想クライアントは移動せず、ユーザーへ影響を与えないことがわかりました。

■ 移動先サーバー障害(障害復旧の方法)

障害により移動先サーバーが停止すると、起動した後も動的最適化対象から外れる場合があります。 一度、動的最適化対象から外れたサーバーに対して、手動でライブマイグレーションすることにより、動的 最適化対象に加えることができます。

5. まとめと考察

SCVMM2012 の動的最適化機能において、以下の項目に関して確認できました。

■ 通常運用時

● 動的最適化機能によるセッション断絶時間は、1 秒程度であるため、仮想クライアントを使用するユーザー へ影響がありませんでした。

■ 障害時

- 障害時動的最適化移動元サーバーにおいて障害(シャットダウン相当)が生じた場合、仮想クライアントを 使用するユーザーへ影響がありませんでした。
- 動的最適化移動元サーバーにおいて障害(電源オフ相当)が生じた場合、仮想クライアントは MSFC の機能により、フェールオーバーし、再起動します。作業中だった内容は失われますが、再起動後は、正常に動作します。
- 動的最適化移動先サーバーにおいて障害(シャットダウン相当、電源オフ相当)が生じた場合、仮想クライアントへ影響がありませんでした。
- サーバー停止障害時には以下の2点を実行すれば SCVMM2012を復旧できます。
 - 移動元サーバーにおいて障害(電源オフ相当)が生じた場合、仮想クライアントはフェールオーバーし、他サーバー上から起動されます。この際、フェールオーバーした仮想クライアントはSCVMM2012 の管理対象から外れますが、SCVMM2012 上からジョブの再試行および仮想クライアントの修復を実行することにより、再度管理対象に加えることできます。
 - ▶ サーバーが停止すると、起動した後も動的最適化対象から外れる場合があります。一度、動的 最適化対象から外れたサーバーに対して、手動でライブマイグレーションすることにより、動的最 適化対象に加えることができます。

動的最適化の実行条件に設定する4つのしきい値(CPU、メモリ、ディスクI/O、ネットワークI/O)は、移動先サーバーの残りリソースを設定します。SCVMM2012(RC版)に記載されている文章だと、「しきい値まで達した場合に、動的最適化が実行される」と読むこともできますので、意味の取り違いに注意する必要があります。

今回の検証では、ディスク I/O とネットワーク I/O を動的最適化の実行条件から対象外としました。ただし、ディスク I/O または、ネットワーク I/O が多く発生する運用(ウイルス対策ソフトが実行される場合など)が想定される場合は、最適にサイジングする必要があると考えられます。また、①サーバーの稼動監視(JP1 など)設定、②お客様の稼動監視ポリシー、③SCVMM2012 のしきい値を同一のポリシーにする必要があります。よって、動的最適化のしきい値は、①、②を踏まえて適切に検討すべきだと考えられます。

6. 文献目録

- 1. **Microsoft Corporation.** How to Configure Dynamic Optimization and Power Optimization. (オンライン) 2011 年 9 月 8 日. http://technet.microsoft.com/ja-jp/library/gg675118(en-us).aspx.
- 2. —. System Requirements for System Center 2012 Virtual Machine Manager. (オンライン) 2011 年 9 月 8 日. http://technet.microsoft.com/ja-jp/library/gg610592(en-us).aspx.
- 3. Zona Research. The Need for Speed. 1999.

付録1. SCVMM2012 を用いた環境における監視項目

検証から判明した SCVMM2012 を用いた環境において監視すべき項目について記載します。監視は、障害に対して素早く発見することが求められます。よって、SCVMM2012 の障害発見のために、SCVMM2012 のログの監視が必要不可欠です。また、SCVMM2012 の障害原因を調べるために、SCVMM2012 が管理する Hyper-V に関してもログを監視するべきです。以下に、SCVMM2012 を用いた環境において監視すべきイベントログを表付.1 に示します。

表 付.1 SCVMM2012 を用いた環境において監視すべきイベントログ

イベント出力製品	ログ格納場所	重要度
SCVMM2012	VM Manager	Critical, Error, Warning
Hyper-V	Microsoft-Windows-Hyper-V-Config-Admin	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-Config-Operational	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-High-Availability-Admin	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-Hypervisor-Admin	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-Hypervisor-Operational	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-Image-Management-Ser	Critical, Error, Warning
	vice-Admin	
	Microsoft-Windows-Hyper-V-Image-Management-Ser	Critical, Error, Warning
	vice-Operational	
	Microsoft-Windows-Hyper-V-Integration-Admin	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-Network-Admin	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-Network-Operational	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-SynthNic-Admin	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-SynthStor-Admin	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-SynthStor-Operational	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-VMMS-Admin	Critical, Error, Warning
	Microsoft-Windows-Hyper-V-Worker-Admin	Critical, Error, Warning
System	system	Critical, Error, Warning
Application	Application	Critical, Error, Warning

表 付.1 のイベントログは出力されるサーバーがそれぞれ異なる場合があります。管理対象ログが増えると、ログの発見が遅れることや、障害原因追求が遅れる可能性があります。よって、Hyper-V、SCVMM2012、システム、アプリケーションのログを集約、監視するために JP1/IM を用いることが望まれます。JP1/IM は、各サーバーに出力されるログを収集し、一元管理できる製品です。JP1/IM を用いることにより、すべてのサーバーの上記イベントログを収集できるため、迅速に障害を見つけることが可能です。

JP1 に関しては以下の URL をご参照ください。

統合システム運用管理 JP1

http://www.hitachi.co.jp/Prod/comp/soft1/jp1/index.html