

**BladeSymphony と Hitachi Storage Solutions を利用した
JP1/Performance Management および JP1/Integrated
Management-System Center Virtual Machine Manager による
Hyper-V 2.0 の可用性構成検証ホワイトペーパー**

第 1.0 版

2010 年 07 月

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
RAID システム事業部

著作権について

この文書は著作権によって保護されています。この文書の内容の一部または全部を、無断で転載することは禁じられています。

Copyright © 2010 Hitachi, Ltd., All rights reserved.

登録商標・商標について

- Microsoft、Windows、Windows Server、Hyper-V、Windows Vista 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。
- Intel、Intel Core、Xeon は米国およ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。

その他、このホワイトペーパー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 ™ は明記していません。

変更履歴

項番	版数	内容	更新日
1	1.0 版	新規作成	2010 年 07 月

目次

1.	はじめに	1
2.	ホストクラスタ構成と JP1/PFM および JP1/IM-SCVMM との連携	3
2.1.	Hyper-V におけるサーバー冗長化構成(ゲストクラスタとホストクラスタ)	3
2.2.	ホストクラスタ構成と JP1/PFM および JP1/IM-SCVMM との連携による 監視付きホストクラスタ構成	7
3.	検証内容	8
3.1.	検証シナリオ	8
3.1.1.	ゲストクラスタ構成の検証	8
3.1.2.	監視付きホストクラスタ構成の検証	10
3.2.	監視付きホストクラスタ構成の障害時の動作	12
3.3.	仮想マシン(SQL サーバー)の設定情報	13
3.4.	SQL の負荷内容・設定	13
3.5.	JP1 の監視項目・設定	15
3.5.1.	検証で使用する JP1 製品群	15
3.5.2.	JP1/PFM の設定項目	16
3.6.	検証に使用したハードウェア	17
4.	検証結果	18
4.1.	ゲストクラスタ構成の検証結果	18
4.2.	ホストクラスタ検証結果	19
4.3.	ゲストクラスタ構成と監視付きホストクラスタ構成の比較	20
5.	監視付きホストクラスタ構成の運用上の考慮点	21
6.	まとめ	24
7.	注意事項	25
8.	【付録】システム構成詳細	26
8.1.	サーバー装置・OS 構成	26
8.2.	ストレージ装置設定	27

用語および略号

Hyper-V	マイクロソフト社が提供するハイパーバイザ型のコンピュータ仮想化技術。Windows Server 2008の主要な機能の一つ。
仮想マシン	コンピュータ仮想化技術により仮想化されたコンピュータ。
ハイパーバイザ	仮想マシン環境とハードウェアとの間に位置する層。ハードウェアをエミュレートし仮想マシン環境を提供するソフトウェア。
iSCSI	Internet Small Computer System Interface : コンピュータと周辺機器を接続する規格であるSCSI のプロトコルをTCP/IP ネットワーク上で使用する規格。
FC	Fiber Channel : コンピュータと周辺機器を接続するためのデータ転送方式の 1 つ。主に、高い性能が必要なサーバーにおいて、外部記憶装置を接続するために利用される。
SAS	Serial Attached Small Computer System Interface : コンピュータと周辺機器を接続する規格である SCSI 規格の一種。シリアル転送を採用した高速な SCSI 規格。
I/O	Input/OutPut の略。
MSFC	Microsoft Failover Cluster : Windows Server 2008 のクラスタ技術。
JP1/PFM	JP1/Performance Management : 日立が提供するサーバー稼働監視製品シリーズ。
JP1/PFM-Manager	JP1/Performance Management - Manager : JP1/PFM の管理マネージャー製品。
JP1/PFM-Agent	JP1/Performance Management - Agent Option for Platform : JP1/PFM-Manager と対応する Agent 製品。
JP1/PFM-Base	JP1/Performance Management - Base : JP1/PFM-Agent の前提製品。
JP1/PFM-RM for SQL	JP1/Performance Management - Remote Monitor for Microsoft SQL Server : マイクロソフト社の SQL Server をリモートで監視するための製品。
JP1/IM-SCVMM	JP1/Integrated Management - System Center Virtual Machine Manager : Hyper-V 環境を管理する製品。

1. はじめに

昨今、サーバーの仮想化技術は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、本番システムのサーバー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、仮想マシン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、従来の物理的なサーバーが仮想マシンとしてソフトウェアの形式で扱われるため、サーバーの新規作成から構成変更、削除まで、より柔軟な運用が可能になります。また、仮想化の特長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

- ホストマシンおよび仮想マシンにおいては、従来の監視項目を監視するだけでは十分な監視は行えません。仮想化機構(Hyper-V)により物理的なサーバーはエミュレートされ、従来の監視項目以外の項目で監視する必要があります。また複数の仮想マシンが1台のホストマシンに搭載されるため、ホストマシン障害時の影響を考慮すると監視運用の重要性が高まります。

- 可用性

- 仮想化された環境では、ホストマシンが障害で停止すると、そのホストマシンに搭載されている複数の仮想マシンも停止することになるため、仮想マシンの可用性の確保が課題となります。また仮想化された環境では、従来の物理的なサーバーの可用性確保の施策方法をそのままでは使えない場合もあるため、個別に検討が必要です。

- バックアップ

- 仮想化されることによって、仮想マシン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マシンで1台の物理マシンのリソースを共有するため、1台の仮想マシンのバックアップがその他の仮想マシン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て最適なバックアップ方式を検討する必要があります。

本ホワイトペーパーでは、仮想環境の可用性に着目し、Hyper-Vにおけるサーバー冗長化構成について検証を行いました。本ホワイトペーパーでは、Hyper-Vにおける従来の可用性施策を補完する構成として、Hyper-Vのサーバー冗長化技術のひとつであるホストクラスタ構成と、JP1/Performance Management(以下、JP1/PFM)による監視機能、およびJP1/Integrated Management - System Center Virtual Machine Manager(以下、JP1/IM-SCVMM)によるHyper-V環境管理機能を組み合わせた構成について検証を実施し、報告しており

ます。

本検証で使用したJP1/IM-SCVMMは、日立の運用管理製品シリーズであるJP1の最新バージョン(JP1 Ver 9.1)で新たに追加された製品です。マイクロソフト社のHyper-V環境管理製品である、System Center Virtual Machine Managerが、新たにJP1に仲間入りしました。JP1 Ver 9.1、およびJP1/IM-SCVMMI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>

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マイクロソフト株式会社が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およびHitachi Adaptable Modular Storage 2300(以下、AMS2300)を利用しております。

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. ホストクラスタ構成と JP1/PFM および JP1/IM-SCVMM との連携

2.1. Hyper-V におけるサーバー冗長化構成(ゲストクラスタとホストクラスタ)

Hyper-Vにおいて、サーバーの冗長化を実現する方式は2つあります。Hyper-V環境では、仮想サーバーのクラスタ技術としてゲストクラスタ方式とホストクラスタ方式が提供されています。図 2-1に2つの方式の模式図を示します。

- ゲストクラスタ方式
 - ゲスト OS でクラスタを構成する方式。アプリケーションのレベルでフェールオーバーを実行する。ゲスト OS 同士でディスク領域を共有するため、仮想マシンから直接共有ディスクに接続する必要があります。そのため、iSCSI 接続のストレージを使用する必要があります。ホスト OS や Hyper-V ハイパーバイザからの干渉なく接続できるストレージは iSCSI デバイスのみとなるため、FC 接続や SAS 接続のストレージは使用できない。
- ホストクラスタ方式
 - ホスト OS でクラスタを構成する方式。仮想マシンのレベルでフェールオーバーを実行する。ホスト OS 同士でディスク領域を共有する。

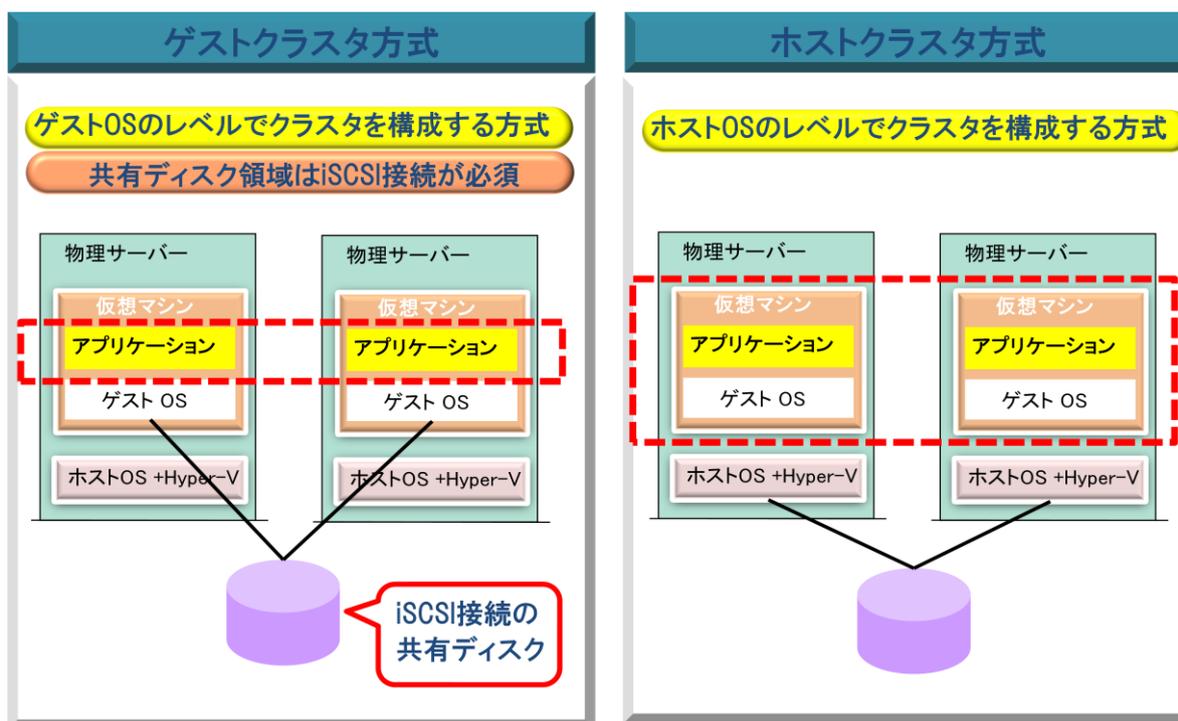


図 2-1 Hyper-V におけるサーバー冗長化方式

次に、それぞれの方式について説明します。

(1) ゲストクラスタ方式

ゲストクラスタ方式は、仮想マシンのゲスト OS 間で、Windows Server のフェールオーバークラスタリング機能 (Microsoft Failover Clustering、以下、MSFC) を用いて、アプリケーションやサービスの冗長化を実現する方式です。ゲスト OS 間でフェールオーバーを構成し、ゲスト OS 上のアプリケーションやサービスの可用性を高めることができます。

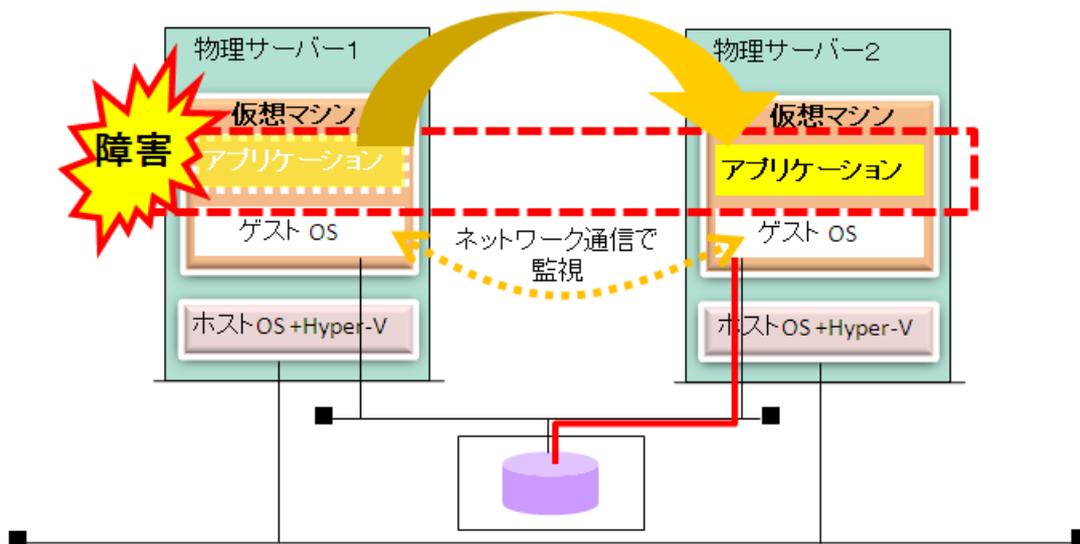


図 2-2 ゲストクラスタ方式の概念図

図 2-2 にゲストクラスタ方式の概念図を示します。仮想マシン1と仮想マシン2でゲストクラスタを構成しており、仮想マシン1のゲスト OS と仮想マシン2のゲスト OS でフェールオーバークラスタリングを構成しています。一方がアクティブノードとなり、他方がパッシブノードになります。仮想マシン同士でディスク領域を共有するため、直接共有ディスクに接続することのできる iSCSI 接続のストレージを接続します。ホスト OS や Hyper-V ハイパーバイザからの干渉なく接続できるストレージは iSCSI デバイスのみとなるため、FC 接続や SAS 接続のストレージは利用できません。

アクティブノードとパッシブノードは、定期的にネットワーク通信を行って状態を確認しています。このネットワーク通信が一定時間以上失われた場合や、監視しているアプリケーションやサービスの状態が障害の状態になった場合には、パッシブ側のゲスト OS がストレージの共有ディスク領域のリソース所有権を獲得しにいく仕組みとなります。

(2)ホストクラスタ方式

ホストクラスタ方式は、ホストマシンのホストOS間で、MSFCによる仮想マシンの冗長化を実現する方式です。ホストOS間でフェールオーバーを構成し、仮想マシンをホストOS上で動くサービスとして構成することで、仮想マシンの可用性を高めることができます。

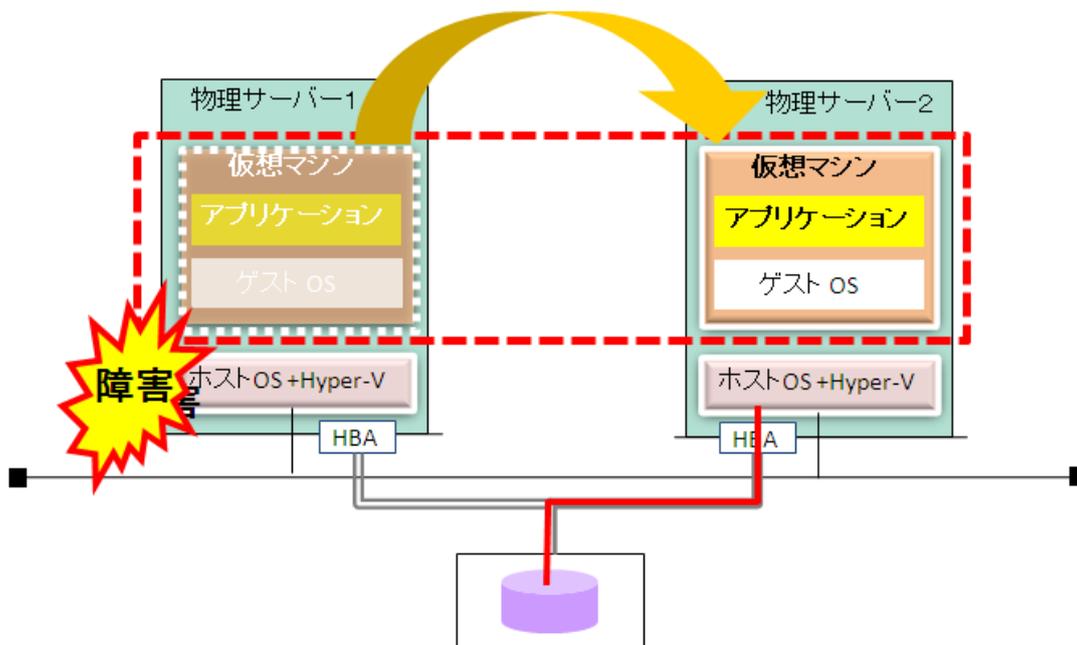


図 2-3 ホストクラスタ方式の概念図

図 2-3 にホストクラスタ方式の概念図を示します。物理サーバー1 と物理サーバー2 でホストクラスタを構成しており、物理サーバー1 のホスト OS と物理サーバー2 のホスト OS でフェールオーバークラスタリングを構成しています。一方がアクティブノードとなり、他方がパッシブノードとなります。ホスト OS 同士でディスク領域を共有するため、ホスト OS はストレージに接続します。

アクティブノードとパッシブノードは、定期的にネットワーク通信を行って状態を確認しています。このネットワーク通信が一定時間以上失われた場合や、監視している仮想マシンが応答しなくなった場合には、パッシブノードがストレージの共有ディスク領域のリソース所有権を獲得しに行く仕組みとなります。

表 2-1 でゲストクラスタ方式とホストクラスタ方式を比較します。

ゲストクラスタ方式は、アプリケーションレベルでの監視が可能であるため、ゲスト OS 上のアプリケーション障害を契機に自動でフェールオーバーすることが可能です。しかし、ゲストクラスタ方式では、ゲスト OS から直接共有ディスク領域に接続する必要があるため、iSCSI 接続ストレージの利用が前提となります。そのため、FC 接続が必要となるような高いディスク I/O 性能が求められるシステムにはあまり向きません。またゲストクラスタリング構成では、共有されているディスク領域は仮想化されていない領域になります。そのため、ゲストクラスタ構成では、仮想化の特長のひとつである可搬性の高さが失われます。仮想マシンをホストマシン間で移行させて動的にリソースを活用する、といった柔軟な運用ができなくなります。

一方、ホストクラスタ方式は、サーバーとストレージ間を FC ケーブルで接続することが可能であるため、高速なディスク I/O が要求されるシステムに向いています。また、仮想マシンをホストマシン間でダウンタイムなしで移行するライブマイグレーション機能が利用できるため、仮想化の特長のひとつである可搬性の高さを生かした運用ができます。しかし、フェールオーバークラスタ機能としては、仮想マシンの死活監視機能しか提供されていません。そのため、ゲスト OS 上のアプリケーションやサービスに障害が発生しても、自動でフェールオーバーされません。

表 2-1 ゲストクラスタ方式とホストクラスタ方式の比較

項目	ゲストクラスタ方式	ホストクラスタ方式
フェールオーバーの方法	<ul style="list-style-type: none"> ゲスト OS 上のアプリケーションやサービスの障害を検知し、フェールオーバーする 	<ul style="list-style-type: none"> 仮想マシンの障害を検知し、フェールオーバーする
メリット	<ul style="list-style-type: none"> ゲスト OS 上のアプリケーションやサービスの障害に対してフェールオーバー可能 	<ul style="list-style-type: none"> MSFC に対応していないアプリケーションも冗長化できる ディスク I/O の性能が高いため、高いディスク性能が求められるシステムにも適用可能 ライブマイグレーション機能が利用可能
デメリット	<ul style="list-style-type: none"> MSFC 対応のアプリケーションのみの利用に限られる 高いディスク I/O 性能が求められるシステムには向かない 仮想マシンのホスト間移行ができない 	<ul style="list-style-type: none"> ゲスト OS 上のアプリケーションやサービスの障害に対してフェールオーバーできない

2.2. ホストクラスタ構成と JP1/PFM および JP1/IM-SCVMM との連携による 監視付きホストクラスタ構成

前節で述べたように、Hyper-V のサーバー冗長化技術である、ゲストクラスタ方式とホストクラスタ方式にはそれぞれにメリットとデメリットがあります。

特に、ゲストクラスタ方式においては、アプリケーションやサービスの障害への対応という比較的高いレベルでの可用性を実現できるのにも関わらず、高いディスクI/O性能が求められるシステムには向かないという、相反するデメリットがあります。また、仮想化の特長のひとつである可搬性を生かした運用をすることが難しくなるため、仮想化することによるメリットが薄くなります。

そこで、ホストクラスタ方式と、アプリケーションまたはサービスを監視する製品とを連携させることで、これらのデメリットを解消するサーバー冗長化の方式を実現できないかを検討しました。以降、このような方式を用いた構成を、監視付きホストクラスタ構成と記載します。図 2-4 に監視付きホストクラスタ構成の仕組みについて示します。図 2-4 は、監視ソフトウェアとして JP1/PFM を使用し、アプリケーションとして SQL Server 2008 を監視した場合の構成になります。監視ソフトウェアの詳細は 3.5.1 節を参照ください。

JP1/PFM-RM for SQL は仮想マシンのゲスト OS 上の SQL Server 2008 の状態を監視します。SQL Server 2008 で障害が発生した場合、JP1/PFM-RM for SQL が障害を検知し、障害発生 の情報を JP1/PFM-Manager に送付します。JP1/PFM-Manager は JP1/IM-SCVMM に対し、仮想マシンのフェールオーバーを実行するように命令します。JP1/IM-SCVMM は仮想マシンのフェールオーバーを実行します。

またこのとき、フェールオーバークラスターマネージャの設定において、仮想マシンでのハートビートモニタリングを無効にしておきます。アプリケーションおよびサービスを監視することにより、仮想マシンでのモニタリングの必要がなくなるためです。

本ホワイトペーパーでは、この方式の実現性について検証しました。次章以降で検証内容と結果について説明します。

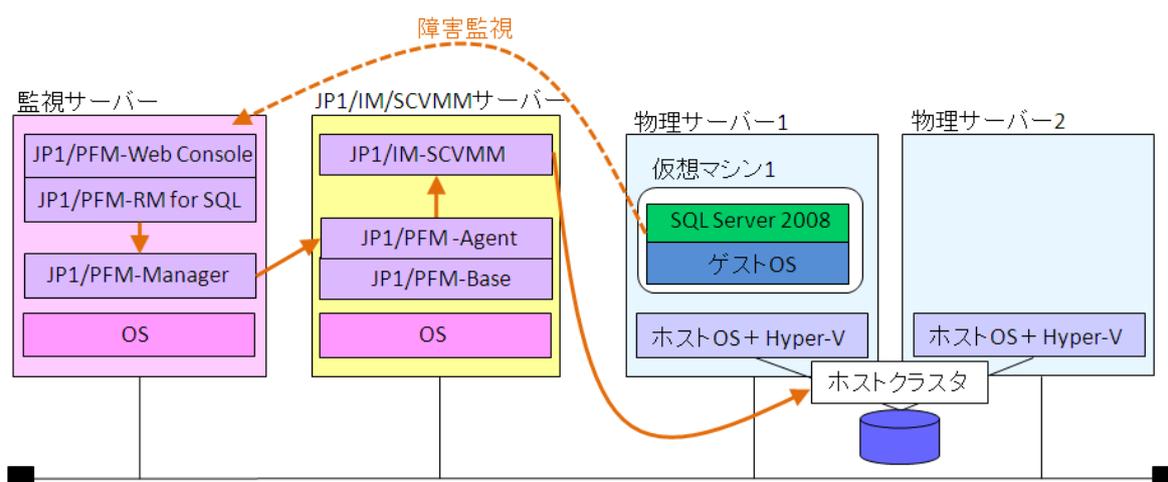


図 2-4 監視付きホストクラスタの仕組み

3. 検証内容

3.1. 検証シナリオ

2.2 節で説明した監視付きホストクラスタ構成について検証を実施しました。本検証では、ゲストクラスタ構成と監視付きホストクラスタ構成の2つの構成について検証を実施し、結果を比較します。

3.1.1. ゲストクラスタ構成の検証

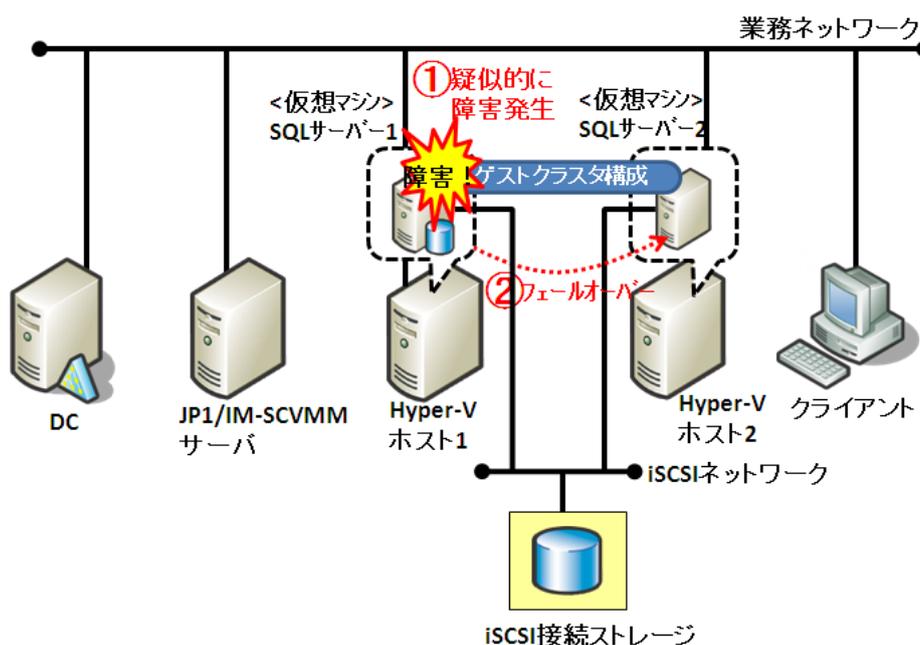


図 3-1 ゲストクラスタ構成の検証 構成図

表 3-1 ゲストクラスタ方式の検証におけるサーバー構成

サーバー	物理/仮想	ソフトウェア	説明
DC	物理サーバー	・Windows Server 2008	ActiveDirectory ドメインコントローラー。
JP1/IM-SCVMM サーバー	物理サーバー	・Windows Server 2008	Hyper-V 環境の管理サーバー。
Hyper-V ホスト 1	物理サーバー	・Windows Server 2008 R2	仮想マシンをホストする Hyper-V ホストマシン。仮想マシンである SQL サーバー 1 をホストする。
Hyper-V ホスト 2	物理サーバー	・Windows Server 2008 R2	仮想マシンをホストする Hyper-V ホストマシン。仮想マシンである SQL サーバー 2 をホストする。

サーバー	物理/仮想	ソフトウェア	説明
SQL サーバー1	仮想サーバー	<ul style="list-style-type: none"> Windows Server 2008 SQL Server 2008 	SQL Server 2008 がインストールされた仮想マシン。SQL サーバー2 とゲストクラスタ構成をとる。
SQL サーバー2	仮想サーバー	<ul style="list-style-type: none"> Windows Server 2008 SQL Server 2008 	SQL Server 2008 がインストールされた仮想マシン。SQL サーバー1 とゲストクラスタ構成をとる。
クライアント	物理サーバー	<ul style="list-style-type: none"> Windows Server 2008 	SQL サーバーに対して負荷を発生させるクライアントマシン。

ゲストクラスタ構成の検証における構成図を図 3-1 に、構成するサーバーの一覧を表 3-1 に示します。詳細な構成は付録をご参照ください。

Hyper-V ホスト 1 には仮想マシン 1、Hyper-V ホスト 2 には仮想マシン 2 が配置されています。仮想マシン 1 と仮想マシン 2 では SQL Server 2008 がインストールされており、ゲストクラスタを構成しています。クライアントは、クラスタ化された SQL Server に対してクエリを発行し、アクセス負荷を与えます。

このような状態で疑似的に障害を発生させて(図 3-1 ①)、自動フェールオーバーを実行させます(図 3-1 ②)。今回は、SQL Server のサービスを停止することで、SQL Server 上の障害を疑似的に発生させます。フェールオーバー時の SQL Server のサービス停止時間等を測定します。測定項目は表 3-2 の通りです。

表 3-2 ゲストクラスタ測定項目

#	測定項目	測定場所	説明および測定方法
1	リソース移動時間	SQL サーバー1	フェールオーバーの処理時間をフェールオーバー管理コンソールから測定する。
2	PING 断時間	クライアント	クライアントマシンから、クラスタ化された SQL Server の IP アドレスに対して ping を打ち続ける。応答がなかった時間を計測する。
3	SQL のサービスの起動・停止時刻	SQL サーバー1 SQL サーバー2	イベントログから SQL のサービスの停止/開始時刻を確認する。

3.1.2. 監視付きホストクラスタ構成の検証

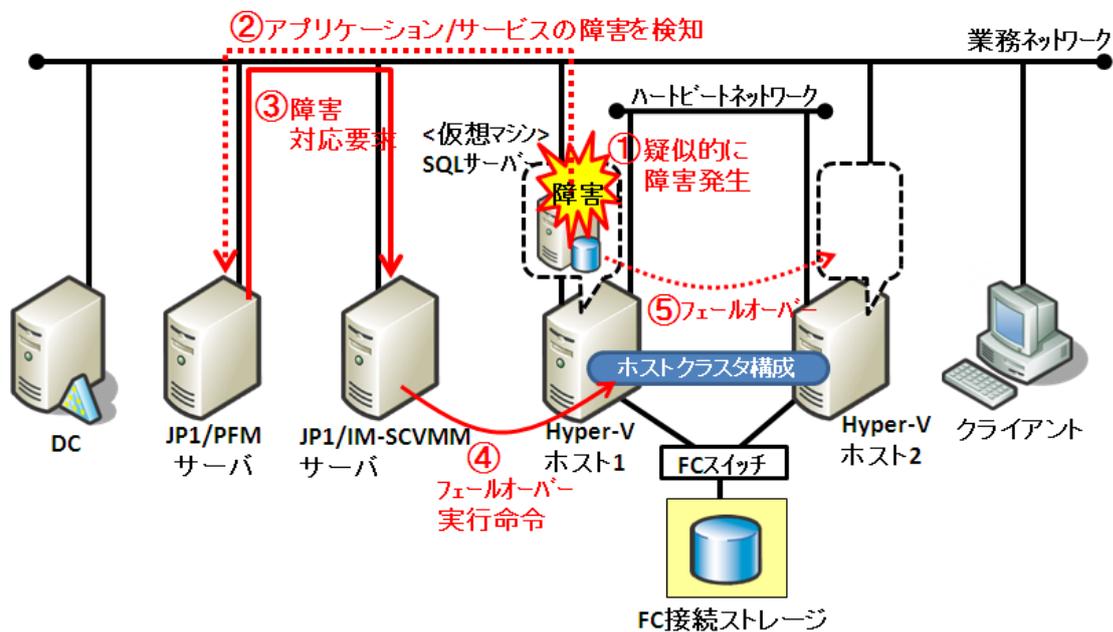


図 3-2 監視付きホストクラスタ構成の検証

表 3-3 監視付きホストクラスタ方式の検証におけるサーバー構成

サーバー	物理/仮想	ソフトウェア	説明
DC	物理サーバー	・Windows Server 2008	ActiveDirectory ドメインコントローラー。
JP1/PFMサーバー	物理サーバー	・Windows Server 2008 ・JP1/PFM-RM for SQL ・JP1/PFM-Manager ・JP1/PFM-Web Console	SQL Server のサービスの稼働を監視するサーバー。
JP1/IM-SCVMMサーバー	物理サーバー	・Windows Server 2008 ・JP1/IM-SCVMM ・JP1/PFM-Base ・JP1/PFM-Agent for Platform	Hyper-V 環境の管理サーバー。ホストクラスタのサービスに対して、フェールオーバーや仮想マシンの再起動などの操作を実行する。
Hyper-V ホスト 1	物理サーバー	・Windows Server 2008 R2	仮想マシンをホストする Hyper-V ホストマシン。Hyper-V ホスト 2 とホストクラスタを構成する。
Hyper-V ホスト 2	物理サーバー	・Windows Server 2008 R2	仮想マシンをホストする Hyper-V ホストマシン。Hyper-V ホスト 1 とホストクラスタを構成する。

サーバー	物理/仮想	ソフトウェア	説明
SQL サーバー	仮想サーバー	<ul style="list-style-type: none"> Windows Server 2008 SQL Server 2008 	ホスト 1 とホスト 2 で構成するホストクラスタ上の仮想マシン。SQL Server 2008 がインストールされている。
クライアント	物理サーバー	<ul style="list-style-type: none"> Windows Server 2008 	SQL サーバーに対して負荷を発生させるクライアントマシン。

ゲストクラスタ構成の検証における構成図を図 3-2 に、構成するサーバーの一覧を表 3-3 に示します。詳細な構成は付録をご参照ください。

Hyper-V ホスト 1 と Hyper-V ホスト 2 でホストクラスタを構成し、ホストクラスタ上のクラスタリソースとして、仮想マシンである SQL サーバーを配置します。JP1/PFM サーバー上の JP1/PFM RM for SQL は、一定の間隔(本検証では 60 秒間隔)で SQL Server へ接続を試みることによって、SQL Server のサービスの稼働監視を行います。クライアントは SQL サーバーに対してクエリを発行し、アクセス負荷を与えます。

このような状態で SQL サーバーに疑似的に障害を発生させます(図 3-2 ①)。今回は、SQL Server のサービスを停止することで、SQL Server 上の障害を疑似的に発生させます。

障害を発生させると、JP1/PFM サーバー上の JP1/PFM RM for SQL が障害を検知します(図 3-2 ②)。そして JP1/PFM Manager を経由して、JP1/IM-SCVMM サーバー上の JP1/IM-SCVMM に対し、フェールオーバーを実行するよう命令します(図 3-2 ③)。JP1/IM-SCVMM はその命令を受け取り、ホストクラスタのサービスに対してフェールオーバーを実行させます(図 3-2 ④⑤)。フェールオーバー時の SQL Server のサービス停止時間等を測定します。測定項目は表 3-4 の通りです。

表 3-4 監視付きホストクラスタ構成の測定項目

#	測定項目	測定場所	説明および測定方法
1	障害発生～JP1 障害検知時間	JP1/PFM サーバー	障害を発生させて JP1 検知するまでの時間を目視で測定する。
2	JP1 障害検知～復旧検知時間	JP1/PFM サーバー	JP1 のイベント履歴で、SQL の障害検知時刻と、復旧確認時刻を確認する。
3	仮想マシン移動時間	JP1/IM-SCVMM サーバー	仮想マシンの移動処理(フェールオーバー)時間を SCVMM 管理コンソールから測定する。
4	PING 断時間	クライアント	クライアントマシンから SQL サーバーに対して ping を打ち続ける。応答がなかった時間を計測する。
5	SQL のサービスの起動・停止時刻	SQL サーバー	イベントログから SQL のサービスの停止/開始時刻を確認する。

3.2. 監視付きホストクラスタ構成の障害時の動作

監視付きホストクラスタ構成の検証において、障害発生からフェールオーバー実行までの処理の流れを図 3-3 に示します。

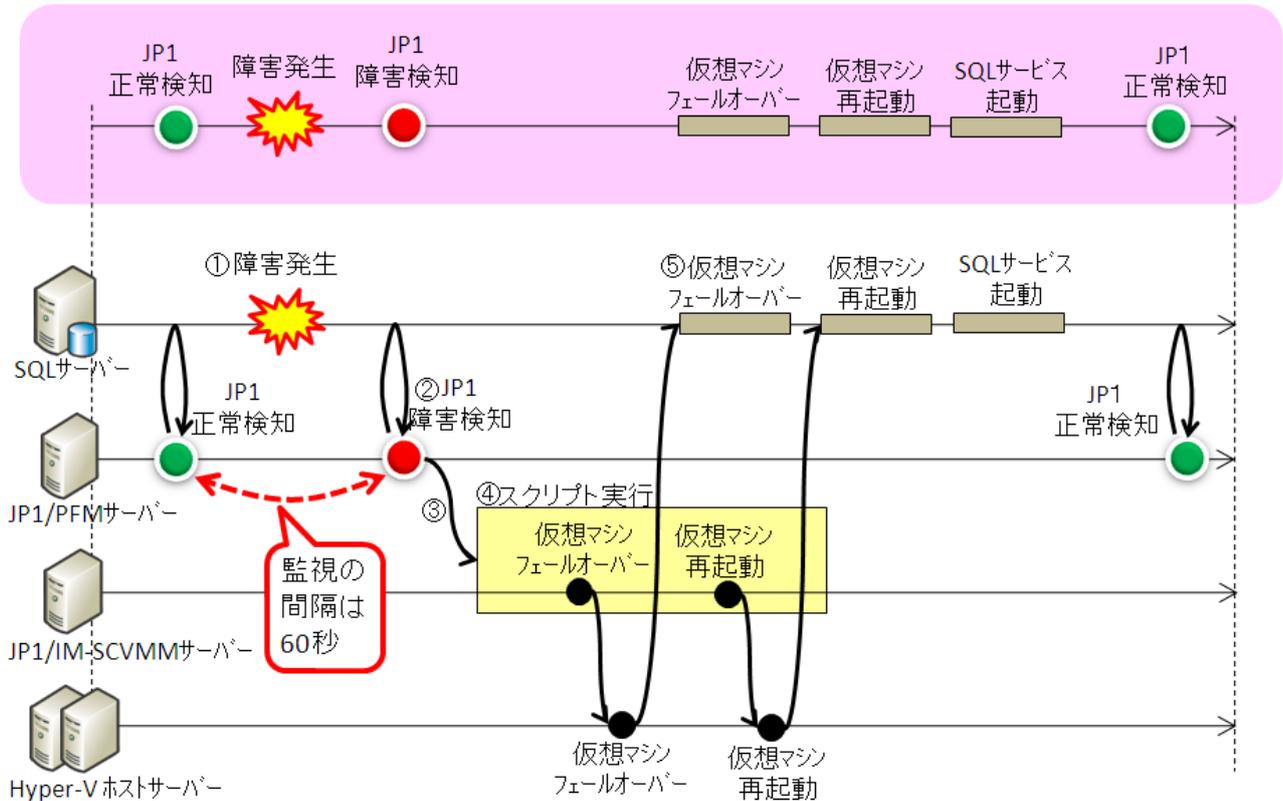


図 3-3 監視付きホストクラスタ方式における障害時の処理のながれ

- ①SQL のサービスに障害が発生します。
- ②JP1/PFM サーバー上の JP1/PFM-RM for SQL が障害を検知します。
- ③JP1/PFM サーバーから JP1/IM-SCVMM サーバーに対して、JP1/IM-SCVMM サーバー上にある PowerShell スクリプトを実行させるためのコマンドを発行します。
- ④JP1/IM-SCVMM サーバーで PowerShell スクリプト(仮想マシンのフェールオーバー後再起動)を実行します。
- ⑤スクリプト従って仮想マシンのフェールオーバーと再起動が実行されます。

3.3. 仮想マシン(SQL サーバー)の設定情報

移行対象とする仮想マシン(SQL サーバー)の設定情報は表 3-5 の通りです。

表 3-5 仮想マシン(SQL サーバー)の設定情報

項目	設定値
CPU	仮想プロセッサ(4 コア)
メモリ	仮想メモリ(8GB)
ドライブ構成	Cドライブ:127GB
VHD サイズ	8.4GB
スナップショットファイルサイズ	9.5GB
仮想マシンの構成ファイル	41.9GB
OS	Windows Server 2008 EE (x64) SP2
OS 以外のソフトウェア	SQL Server 2008

3.4. SQL の負荷内容・設定

データベースの利用状況を、以下の通り定義しました。検証環境の SQL サーバーには、以下の条件に沿った負荷が与えられます。

- データベースには、読取り(select 文)処理と書込み(update 文)処理を含めた負荷をかける

負荷発生用のクライアントツールとして、SQL Management Studio 2008 を使用します。1 台のクライアントからサーバーに対して、結果応答に 5 分を要するトランザクション(ロングトランザクション)となるクエリを発行し、クライアントからデータベースにアクセスします。

この負荷発生においては、1 つのクエリを発行し、トランザクションが終了するまでの処理が完了するまでを実行の 1 単位としています。1 実行単位が終了すると、データベースのコミットが実行されます。1 実行単位で行われるループ処理の概要は以下の通りです。

- ① 読取り処理を実行し、一時保管テーブルにデータを取出す。
- ② 書込み処理を実行する。

データベースに配置したデータを表に示します。データは同一パーティション上に作成し、今回の検証でアクセスされないダミーデータも配置しました。データは任意の ASCII 文字列となっています。検証に使用されるデータの総容量は 1GB 以上とします。

表 3-6 主要データ項目

#	テーブル	項目	データ定義
1	A_売上明細	商品コード	NCHAR(10)
2		備考	NVARCHAR(50)
3	B_商品	商品コード	NCHAR(10)
4		商品名	NVARCHAR(50)

負荷をかけるクエリは以下の通りです。

```

BEGIN TRANSACTION
UPDATE A_売上明細
SET 備考 =(
SELECT 商品名
FROM [データベースクラス名].[dbo]. B_商品
WHERE B_商品.商品コード = A_売上明細.商品コード);
COMMIT TRANSACTION

```

3.5. JP1 の監視項目・設定

3.5.1. 検証で使用する JP1 製品群

本検証で使用する JP1 製品を下記に示します。JP1/PFM-RM for SQL は、監視対象のアプリケーションが SQL Server である場合に限り利用可能です。Microsoft Exchange 等、アプリケーションが異なる場合には、アプリケーションに応じた監視の仕組みを再構成する必要があります。

参考：<http://outside2.soft.hitachi.co.jp/soft/jp1/v9/index.html>

表 3-7 検証で使用する JP1 製品一覧

#	JP1 アプリケーション名	特徴
1	JP1/Performance Management - Remote Monitor for Microsoft SQL Server (JP1/PFM -RM for SQL)	<ul style="list-style-type: none"> 1 つの JP1/PFM - RM for SQL から、複数のホストにある監視対象の Microsoft SQL Server のパフォーマンスをリモートで監視できます。 データベーススペース使用率やキャッシュヒット率の統計情報などのパフォーマンスデータを収集および集計し、その傾向や推移を図示することで、稼働状況の分析が容易にできます。 データベーススペース使用率が増大したり、キャッシュヒット率が低下したりするなどのパフォーマンスの低下が発生した場合、E メールなどを使ってユーザーに通知することで、問題点を早期に発見できます。また、その問題点に関連する情報を図示することで、原因を調査する資料を提供できます。
2	JP1/Performance Management - Manager (JP1/PFM - Manager)	JP1/PFM-RM for SQL を使用するための前提製品です。
3	JP1/Performance Management - Web Console (JP1/PFM - Console)	JP1/PFM-Manager を使用するための前提製品で、Web ブラウザベースで JP1 の監視設定を行います。
4	JP1/Performance Management - Agent Option for Platform (JP1/PFM -Agent)	JP1/PFM-Manager に対する Agent 製品です。
5	JP1/Performance Management - Base (JP1/PFM - Base)	JP1/PFM-Agent for Platform を使用するための前提製品です。

#	JP1 アプリケーション名	特徴
6	JP1/Integrated Management – System Center Virtual Machine Manager (JP1/IM – SCVMM)	<ul style="list-style-type: none"> 仮想マシンの準備、展開、最適な配置、ライブ マイグレーション、P2V (仮想マシン–物理マシン) 変換、V2V (仮想マシン–仮想マシン) 変換、セルフ サービス ポータルによる管理の委任など、仮想化環境のライフサイクル全体をサポートします。 サーバー統合を促進し、ハードウェアの利用率とパフォーマンスの最大化が実現できます。

3.5.2. JP1/PFM の設定項目

JP1/PFM の設定項目を、表 3-8 に示します。

表 3-8 JP1/PFM 監視項目

#	項目	データ定義
1	データベース	<SQL Server のインスタンス名称>@ゲストサーバー1
2	アラート機能	PD_DD (SQL Server の各データベースの応答障害監視)
3	アラームテーブル	<任意のテーブル名>
4	コマンド名	<実行ファイル名> ※フェールオーバー実行のコマンドラインを含む PowerShell ファイル
5	監視対象	SQL Server
6	監視間隔	60 秒

3.6. 検証に使用したハードウェア

検証で共通して使用したハードウェアを以下に示します。

表 3-9 検証で使用したハードウェア

製品名	メーカー	種類
BladeSymphony BS320	日立	ブレードサーバー
Hitachi Adapter Modular Storage 2300(AMS2300)	日立	ストレージ

4. 検証結果

4.1. ゲストクラスタ構成の検証結果

ゲストクラスタ構成の検証結果を示します。

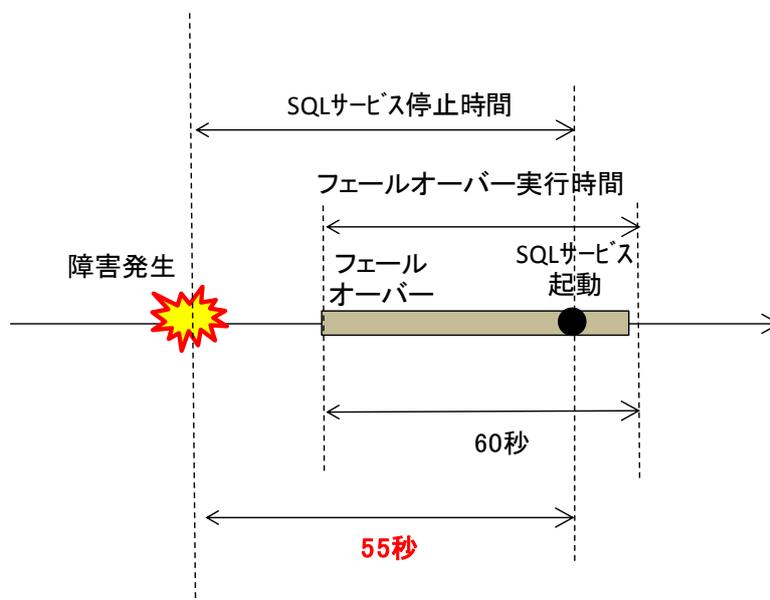


図 4-1 ゲストクラスタデータ取得結果

表 4-1 ゲストクラスタ取得データ表

フェールオーバー時間	PING 不通時間	SQL のサービスの停止-再起動時間
60 秒	22 秒	55 秒

検証の結果、ゲストクラスタ構成では SQL のサービスの停止時間が 55 秒程度でした。

4.2. ホストクラスタ検証結果

監視付きホストクラスタ構成の検証結果を示します。

測定結果を以下に示します(数値は3回の試行の平均値)

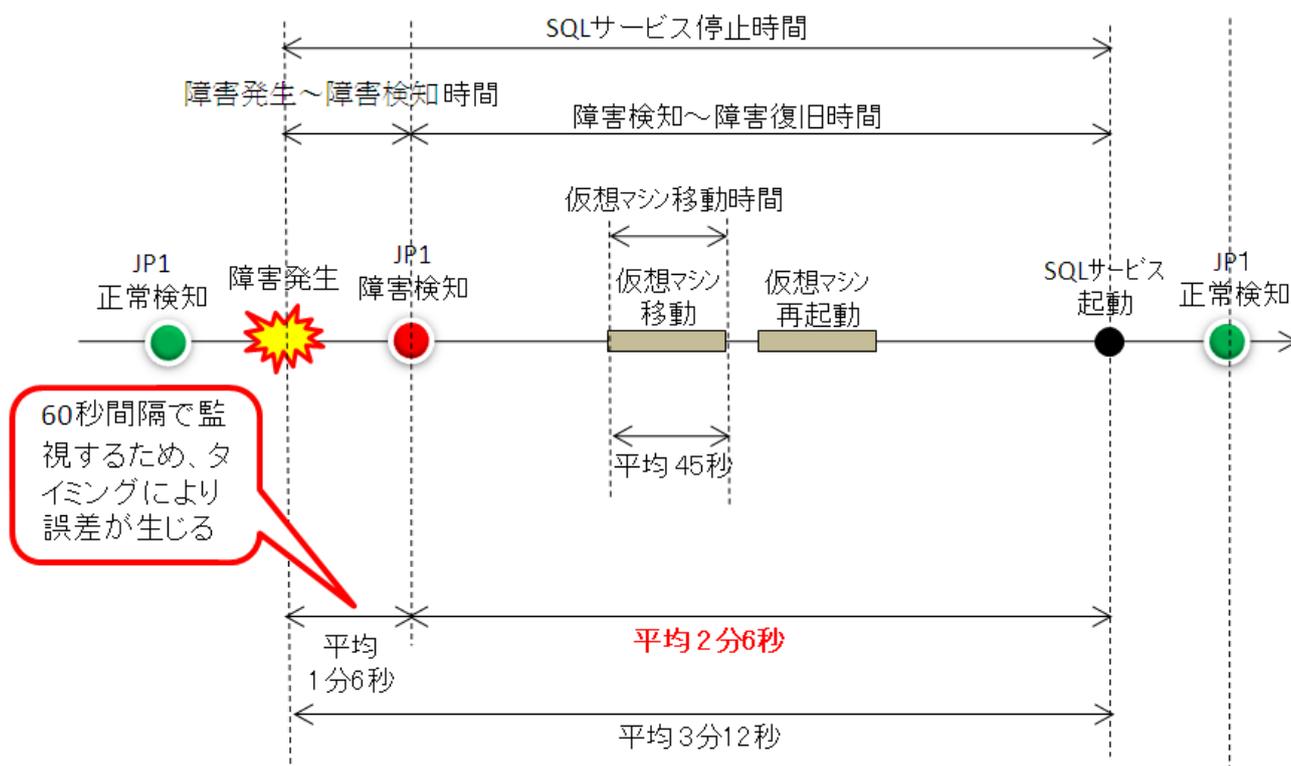


図 4-2 監視付きホストクラスタデータ検証結果

表 4-2 ホストクラスタ取得データ(3回の試行の平均値)

障害発生～JP1 障害検知時間	JP1 障害検知～ 復旧検知時間	クイックマイグレーション 実行時間	Ping 不通時間	SQL のサービスの 停止-再起動時間
1 分 06 秒	2 分 06 秒	45 秒	15 秒	3 分 12 秒

測定の結果は表 4-2 の通りです。3 回の試行の平均値を示しています。障害発生から JP1 障害検知までの時間に関しては、JP1/PFM の監視間隔のタイミングがあるため数十秒～2 分の誤差があります。SQL のサービス停止時間には JP1/PFM の障害監視の間隔による、障害検知までの時間の誤差が含まれるため、幅のある結果となっています。3 回の試行の結果では、SQL のサービスの停止時間は平均して 3 分 12 秒でした。しかし、JP1/PFM の障害検知の間隔を考慮すると、今回の構成では(監視間隔 60 秒)SQL のサービス停止時間は 2 分数十秒～4 分程度になると考えられます。

なお、この測定において、イベントログへ不正なエラーが出ることはありませんでした。

4.3. ゲストクラスタ構成と監視付きホストクラスタ構成の比較

ゲストクラスタ構成と監視付きホストクラスタ構成の結果を比較します。SQL Server のサービス停止時間は、ゲストクラスタの 55 秒(約 1 分)に対し、監視付きホストクラスタでは 192 秒(約 3 分)と測定されました。ただし、監視付きホストクラスタ構成では、JP1/PFM が障害を検知する間隔による誤差があります。そのため SQL サービスの停止時間は、タイミングによって最小 2 分数十秒～最大 4 分程度まで誤差が出ると考えられます。

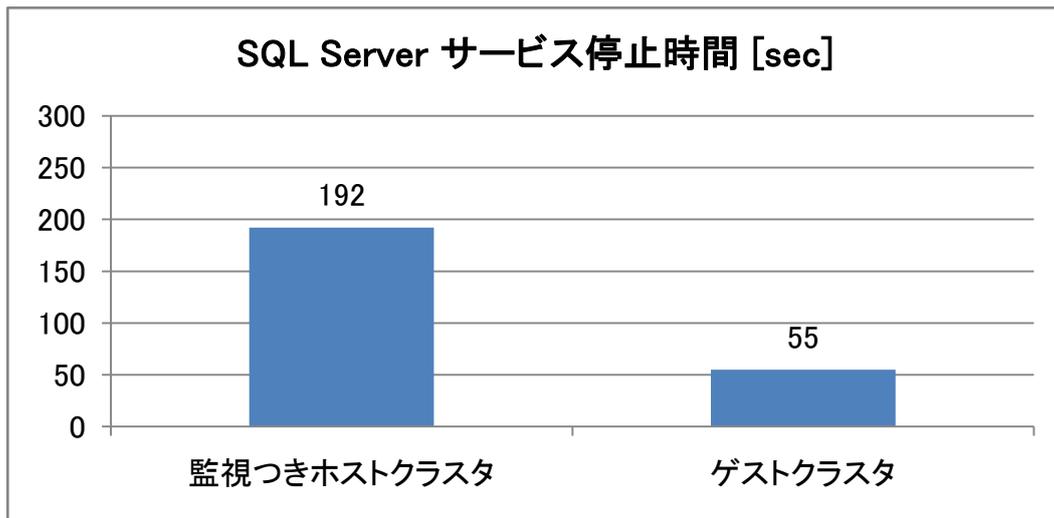


図 4-3 ゲストクラスタ構成と監視付きホストクラスタ構成の比較

5. 監視付きホストクラスタ構成の運用上の考慮点

本検証における、監視付きホストクラスタ構成の障害時の動作について3.2節で説明しました。監視付きホストクラスタでは、JP1/IM-SCVMM に対し障害発生時の処理を定義しておくことで、障害時の自動対応が可能になります。本検証では、JP1/IM-SCVMM に対して、障害発生時の処理を Powershell によるスクリプトで定義し、仮想マシンのフェールオーバーと再起動を実行させる処理を定義しました。

しかし、実際に監視付きホストクラスタ構成を適用する場合には、その目的と障害時の運用によって、障害時の処理を検討する必要があります。

例えば、監視付きホストクラスタは、ホストクラスタ構成を基としているため、ホストマシンでハードウェアの障害が発生した際にはホストクラスタ構成としての自動フェールオーバーが実行されます。そのため、本検証で定義した処理内容では、ハードウェア障害時には、ホストクラスタによる自動フェールオーバーと、JP1/PFM および JP1/IM-SCVMM の連携によるフェールオーバー処理の両方が実行されることとなります。

このケースの例として、本検証の監視付きホストクラスタ構成においてホストマシンの電源ボタンを押下して OFF にした場合の挙動を観測しました。図 5-1 にその結果を示します。フェールオーバーが 2 度実行されてしまい、2 回目のフェールオーバーは失敗しています。但し、SQL のサービスは正常に起動しました。

このようなケースを回避する方法の一つの例を図 5-2 に示します。JP1/IM-SCVMM の障害時の処理において、フェールオーバーを実行する前に、ホストクラスタによる自動フェールオーバーが実行されていないかを確認する処理を追加します。このようにして、現象を回避することができます。

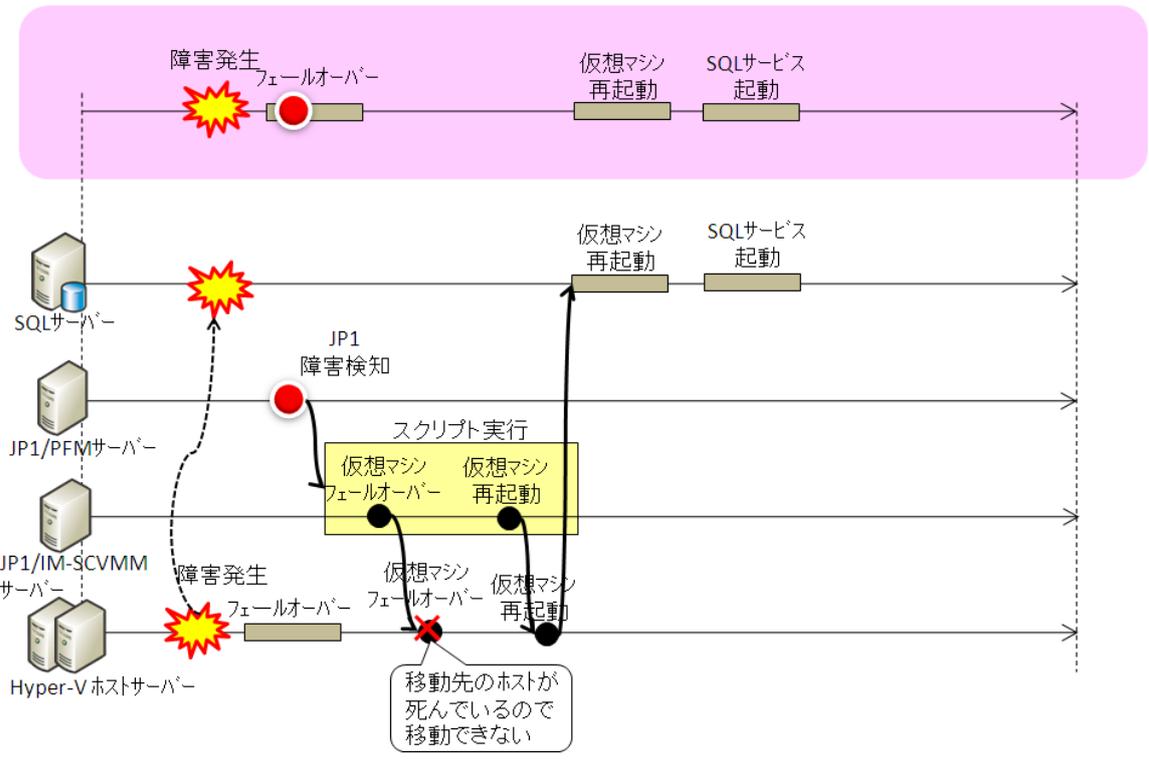


図 5-1 監視付きホストクラスタ構成でホストマシンの電源を OFF にした際の挙動

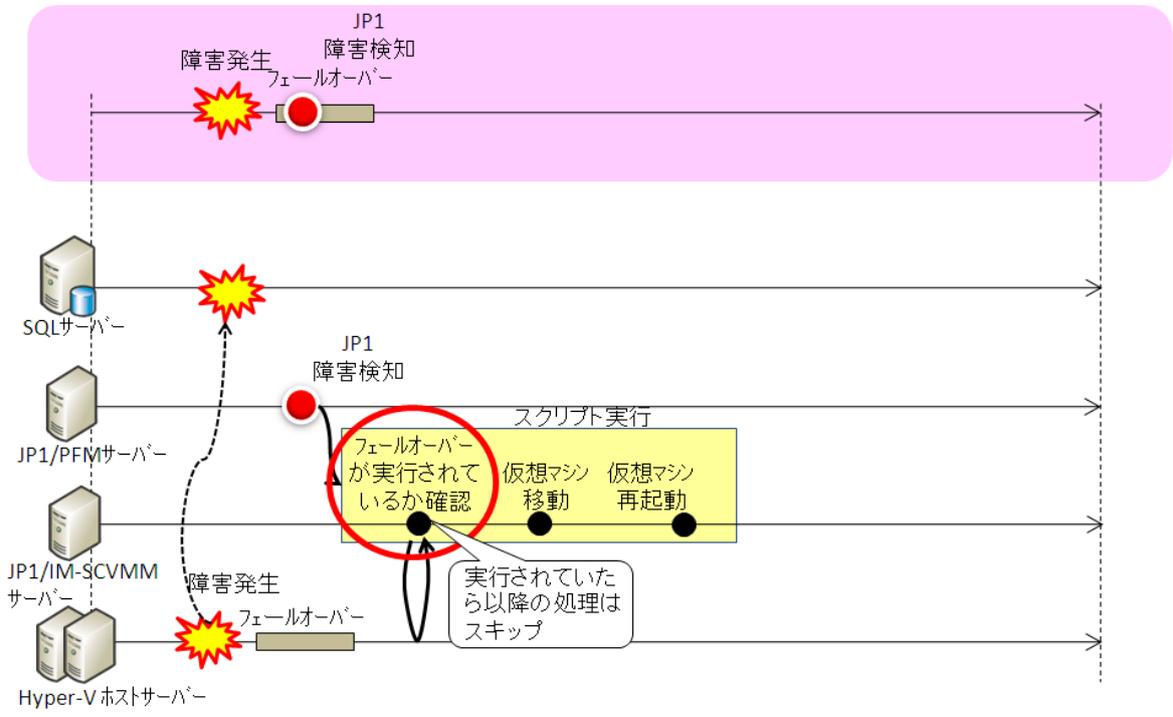


図 5-2 回避方法の例

表 5-1 に、ゲストクラスタ構成、ホストクラスタ構成、監視付きホストクラスタ構成のそれぞれにおいて、自動でフェールオーバー可能な障害の範囲についてまとめます。ホストクラスタ構成と比較して、監視付きホストクラスタ構成では自動フェールオーバーで対応できる障害の範囲が広がっていることが分かります。監視付きホストクラスタ構成では、ホストクラスタによる自動フェールオーバーと、監視機能連携による(JP1/PFM および JP1/IM-SCVMM の連携による)自動フェールオーバーの 2 種類の自動フェールオーバーがあります。また監視付きホストクラスタリングでは、仮想マシンのハートビートモニタリングを無効に設定します。

実際に監視付きホストクラスタ構成を適用する場合には、表 5-1 をもとに、どの障害でどのような運用をしていくか、を十分に検討した上で適用する必要があります。

表 5-1 ゲストクラスタおよび監視付きホストクラスタの対応障害範囲

分類	障害部位	フェールオーバーが自動で発生するか			
		ゲストクラスタ	ホストクラスタ ※1	監視付きホストクラスタ	
				ホストクラスタによる自動フェールオーバー	監視機能連携による自動フェールオーバー
ハードウェア障害	CPU	○	○	○	○
	内蔵 HDD(ホスト OS 領域) ※2	○	○	○	○
	メモリ	○	○	○	○
	電源ユニット ※2	○	○	○	○
	RAID ボード	○	○	○	○
	LAN ボード ※2	×	×	×	×
	マザーボード	○	○	○	○
ソフトウェア障害(ホスト)	ホスト OS	○	○	○	○
	Hyper-V ハイパーバイザ	○	○	○	○
ソフトウェア障害(ゲスト)	リソース不足	○	○	×	○
	ゲスト OS	○	△	×	○
	アプリケーション/サービス	○	×	×	○

(○:自動でフェールオーバーする △:一部自動フェールオーバーしない ×:自動でフェールオーバーしない)

※1:仮想マシンのハートビートモニタリングを有効にした場合

※2:ハードウェアレベルで冗長化することが可能

6. まとめ

本検証では Hyper-V2.0 のホストクラスタ構成と JP1 製品との連携による監視付きホストクラスタ構成の検証を行いました。検証では SQL Server 2008 のサービスを JP1/PFM で監視し、SQL の障害時に、JP1/IM-SCVMM と連携してフェールオーバーをさせる構成について測定を実施しました。結果として、ゲスト OS 上のサービスやアプリケーションに障害が発生した場合、最大で 4 分程度のサービス停止が発生することが分かりました。またこの際に不正なエラー等は検出されませんでした。したがって、数分程度のサービス停止が許容できるシステムにおいて、監視付きホストクラスタ構成を適用できると考えられます。

一方で、本検証では、ハードウェアの障害時に、ホストクラスタによる自動フェールオーバーと、JP1/PFM および JP1/IM-SCVMM の連携によるフェールオーバー処理の両方が実行される問題が起きました。しかし、この問題は、JP1/IM-SCVMM の連携によるフェールオーバー処理の前に、ホストクラスタによる自動フェールオーバー処理が実行されていないかを確認する処理を入れることで回避することが可能です。

このように、実際に監視付きホストクラスタ構成を導入する場合には、目的や障害時の運用を十分考慮した上で、障害時の処理を設計する必要があります。また、監視項目を増やすことで、対応できる障害を変更することもできるため、本ホワイトペーパーで検証した方式以外の応用も考えられます。

7. 注意事項

本検証では、想定したシナリオに基づき計測を実施いたしました。この結果は、システムの構成や利用状況などによっては、これらの傾向が変わる可能性もあるため、注意が必要です。本検証においては、ディスクの構成に余裕がある状態で検証しました。CPU やメモリに大きく負荷がかかっている場合には本結果とは異なる傾向が現れる可能性があります。この注意点以外にも、さまざまな要因によって傾向が変わる可能性があります。

8. 【付録】システム構成詳細

8.1. サーバー装置・OS 構成

サーバー役割	ハードウェア	OS	設定/導入した機能
DC × 1	日立 BladeSymphony BS320(A4) CPU:XeonX5570 (2.93GHz) QuadCore × 2 Memory:32GB NIC:1000Base-T × 4 内蔵 HDD: SAS147GB × 2 (SAS RAID1) 2.5inch 10,000 回転	Windows Server 2008 Enterprise Edition (x64) SP2	<ul style="list-style-type: none"> ▪ OS 設定:導入時既定値 ▪ Windows Firewall:無効 ▪ IPv6 無効
JP1/IM-SCVMM サーバー × 1	日立 BladeSymphony BS320(A4) CPU:XeonX5570 (2.93GHz) QuadCore × 2 Memory:32GB NIC:1000Base-T × 4 内蔵 HDD: SAS147GB × 2 (SAS RAID1) 2.5inch 10,000 回転	Windows Server 2008 Enterprise Edition (x64) SP2	<ul style="list-style-type: none"> ▪ OS 設定:導入時既定値 ▪ Windows Firewall:無効 ▪ IPv6 無効
JP1/PFM サーバー × 1	日立 BladeSymphony BS320(A4) CPU:XeonX5570 (2.93GHz) QuadCore × 2 Memory:32GB NIC:1000Base-T × 4 内蔵 HDD: SAS147GB × 2 (SAS RAID1) 2.5inch 10,000 回転	Windows Server 2003 Enterprise Edition (x64) SP2	<ul style="list-style-type: none"> ▪ OS 設定:導入時既定値 ▪ Windows Firewall:無効 ▪ IPv6 無効
Hyper-V ホスト × 2	日立 BladeSymphony BS320(A3) CPU:XeonE5405 (2.00GHz) QuadCore × 2 Memory:8GB NIC:1000Base-T × 4 内蔵 HDD: SAS73GB × 2 (SAS RAID1) 2.5inch 10,000 回転	Windows Server 2008 R2 Enterprise Edition (x64)	<ul style="list-style-type: none"> ▪ OS 設定:導入時既定値 ▪ Windows Firewall:無効 ▪ IPv6 無効
クライアント × 1	日立 BladeSymphony BS320(A4) CPU:XeonX5570 (2.93GHz) QuadCore × 2 Memory:32GB NIC:1000Base-T × 4 内蔵 HDD: SAS147GB × 2 (SAS RAID1) 2.5inch 10,000 回転	Windows Server 2008 Enterprise Edition (x64) SP2	<ul style="list-style-type: none"> ▪ OS 設定:導入時既定値 ▪ Windows Firewall:無効 ▪ IPv6 無効

8.2. ストレージ装置設定

項目	設定
機種	日立 Adaptable Modular Storage 2300(AMS2300)
コントローラー数	2
ディスクドライブポート数	8ポート/2コントローラー
キャッシュ容量	16Gバイト/装置
ホストインタフェース	FC(最大 8Gbps)×8