

Windows Server 2012 R2 Hyper-V ストレージ QoS 機能による仮想ディスクの I/O 制御

1. はじめに

Windows Server 2012 R2 Hyper-Vでは、仮想ディスクに対するサービスの品質管理機能(ストレージQoS機能)が提供され、仮想マシンの仮想ディスク(VHD/VHDX)に対するI/Oを最小および最大IOPS¹で設定できるようになりました。これにより、従来の仮想ネットワークのQoS機能に加え、ストレージQoS機能による仮想ディスクのI/O制御が可能になります。

今回、当社が実施しているWindows Server 2012 R2 先行評価で得たストレージQoS機能に対する気づきをご紹介します。

2. ストレージ QoS 機能の概要とメリット

ストレージ QoS 機能は、特定の仮想ディスクの I/O 帯域を制御する機能です。Windows Server 2012 以前の Hyper-V で各仮想マシン(以下、VM と略) のディスク I/O 性能を確保するには、仮想ディスクを異なる物理ディスクまたは LUN に配置するなどの考慮が必要でしたが、Windows Server 2012 R2 では、ストレージ QoS 機能により、同一の物理ディスクまたは LUN に複数の仮想ディスクを配置したストレージ構成が可能になります。

ストレージ QoS の設定は GUI 画面(図 1)または CLI(PowerShell)で設定できます。

GUI 画面では、「最小」と「最大」の IOPS を、ブロックサイズ 8KB 単位で設定します。「最小」を設定すると、設定した IOPS が可能な範囲で確保されます。「最大」を設定すると、設定した値以上に IOPS が使われないように I/O 帯域が制限されます。例えば「最大」に 2000 IOPS を設定した場合、 $2000 \times 8KB = 16000KB$ つまり 16MB/sec の I/O 帯域までしか使えないように制限されます。

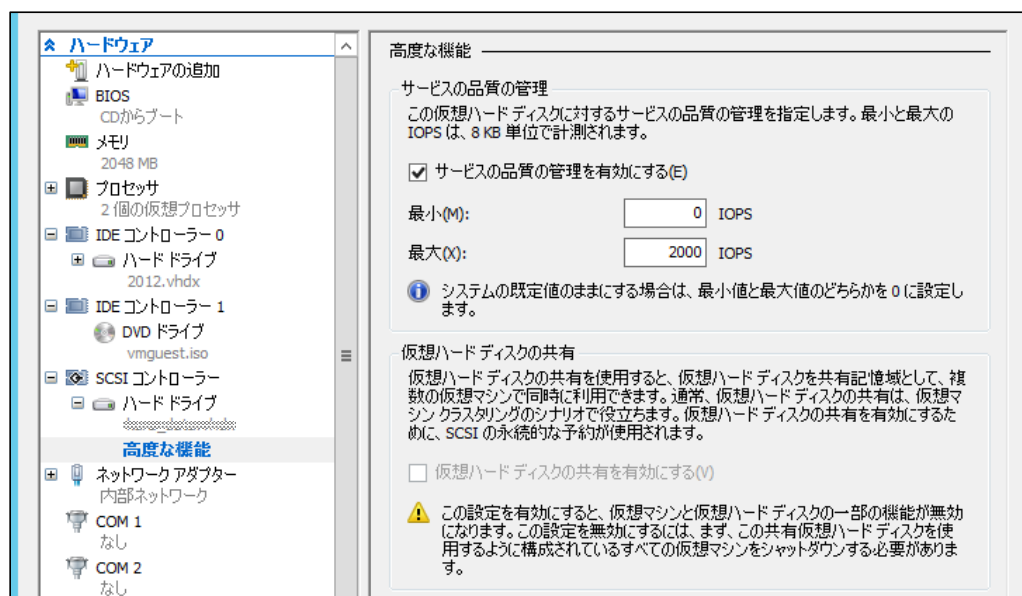


図 1. Windows Server 2012 R2 Hyper-V ストレージ QoS の設定画面

¹ Input/Output Per Second の略。ストレージが 1 秒間に処理できる I/O 数で、この値が大きいかほど I/O アクセス性能が高くなる。

3. VM ごとに仮想ディスクの I/O 帯域を確保する方法

VM ごとに仮想ディスクの I/O 帯域を確保する方法を確認するために、表 1 の環境で動作確認しました。

表 1. Windows Server 2012 R2 Preview のストレージ QoS 動作確認に使用したハードウェア環境

(1)Hyper-V ホスト(単体サーバおよびクラスター構成時のノード 1、ノード 2 共通)

機種	FUJITSU Server PRIMERGY RX300 S7
OS	Windows Server 2012 R2 Datacenter(Build 9431)日本語 Preview 版
CPU	Intel Xeon CPU E5-2690 @ 2.90GHz × 2(合計 32 論理プロセッサ)
メモリ	384GB(16GB × 24socket フル装備) (DDR3-1066RDIMM)
HDD	600GB
RAID 構成	RAID 0(HDD2 本)
CNA カード	Emulex OCe 10102-F 10Gb 2-port PCIe Converged Network Adapter

(2)ストレージ

機種	FUJITSU Storage ETERNUS DX90 S2T
HDD	2729GB
RAID 構成	RAID 1+0(HDD10 本)の 1LUN

(3)VM(10 台)

	VM
OS	Windows Server 2012 R2 Datacenter(Build 9431)日本語 Preview 版
CPU	1 仮想 CPU
メモリ	2048MB
HDD	システム用 VHDX:可変 IDE 128GB データ用 VHDX:固定 SCSI 30GB

上記構成で、以下のとおり測定しました。

- IOPS 設定対象の仮想ディスクはデータ用 VHDX とする。
- 負荷ツール Iometer を使用して全 VM のデータ用 VHDX に I/O 負荷をかける。Iometer のパラメーターは以下のとおり。
 - Transfer Request Size = 512KB
 - Percent Random/Sequential Distribution = Random 100%
 - Percent Read/Write Distribution = Write 100%
 - # of Outstanding I/Os = 1

3.1 仮想ディスクの I/O 帯域を確保するための動作確認(単体サーバ構成)

動作確認した単体サーバ構成は図 2 のとおりです。

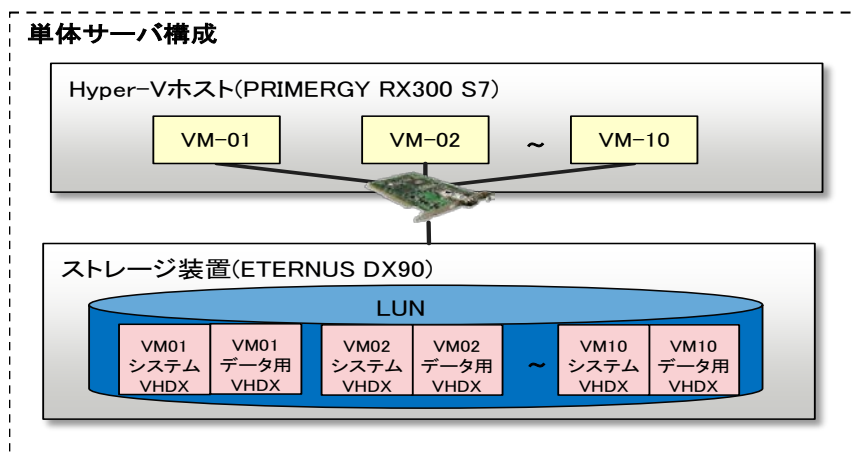


図 2. ストレージ QoS 動作確認環境の単体サーバ構成

「最小」の動作を確認するために、「最大」と組み合わせて IOPS 設定を以下のとおりに変化させて仮想ディスクのスループットを測定しました。

- ① VM01~VM05: IOPS 未設定 VM06~VM10: IOPS 未設定
- ② VM01~VM05: 最大 1000 IOPS を設定(8MB/sec に制限) VM06~VM10: IOPS 未設定
- ③ VM01~VM05: 最大 1000 IOPS を設定(8MB/sec に制限) VM06~VM10: 最小 3000 IOPS を設定(24MB/sec を確保)
- ④ VM01~VM05: IOPS 未設定 VM06~VM10: IOPS 未設定

以下に測定結果を示します。

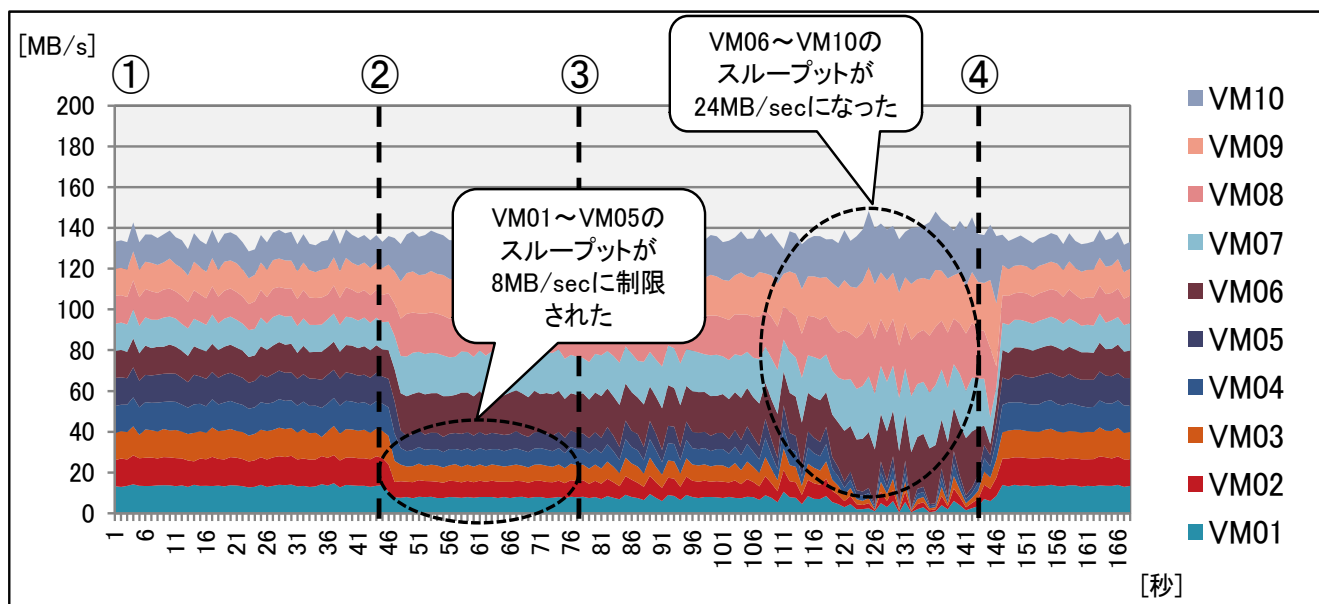


図 3. 単体サーバにおける IOPS の設定変更と仮想ディスクスループットの変化

図 3 の①~④すべての時点において全 VM に対して均等に Iometer による負荷がかかっています。

- ①の時点は、全 VM の仮想ディスクに対して IOPS 未設定の状態です。
- ②の時点で VM01~VM05 の仮想ディスクそれぞれに対して最大 1000IOPS (8MB/sec の制限)を設定したところ、

スループットは 8MB/sec に制限され、未設定の VM06～VM10 のスループットが高くなりました。

次に、③の時点で VM06～VM10 の仮想ディスクに最小 3000IOPS (24MB/sec の確保)を設定したところ、VM06～VM10 は徐々に 24MB/sec のスループットになりました。このように、最小が設定された VM06～VM10 のスループットは、最大が設定された VM01～VM05 の I/O 帯域を奪うようにして確保獲得されました。

3.2 仮想ディスクの I/O 帯域を確保するための動作確認(クラスター構成)

次に、クラスター構成においても、「最小」を設定することで I/O 帯域が確保されるのかを確認するために、図 4 の構成で動作確認しました。この時、VM01～VM05 はノード1側に配置し、VM06～VM10 はノード 2 側に配置しました。

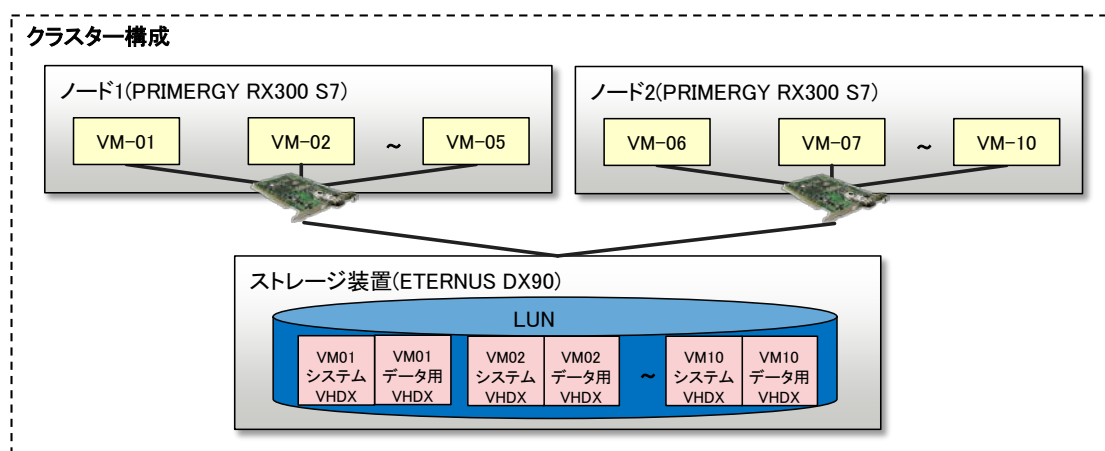


図 4. ストレージ QoS 動作確認環境のクラスター構成

IOPS 設定を以下のとおり変化させて、仮想ディスクのスループットを測定しました。

- | | |
|---|---|
| ① VM01～VM05:IOPS 未設定 | VM06～VM10:IOPS 未設定 |
| ② VM01～VM05:最小 2000 IOPS を設定 (16MB/sec を確保) | VM06～VM10:IOPS 未設定 |
| ③ VM01～VM05:最小 3000 IOPS を設定 (24MB/sec を確保) | VM06～VM10:IOPS 未設定 |
| ④ VM01～VM05:最小 3000 IOPS を設定 (24MB/sec を確保) | VM06～VM10:最小 3000 IOPS を設定 (24MB/sec を確保) |
| ⑤ VM01～VM05:IOPS 未設定 | VM06～VM10:IOPS 未設定 |

以下に測定結果を示します。

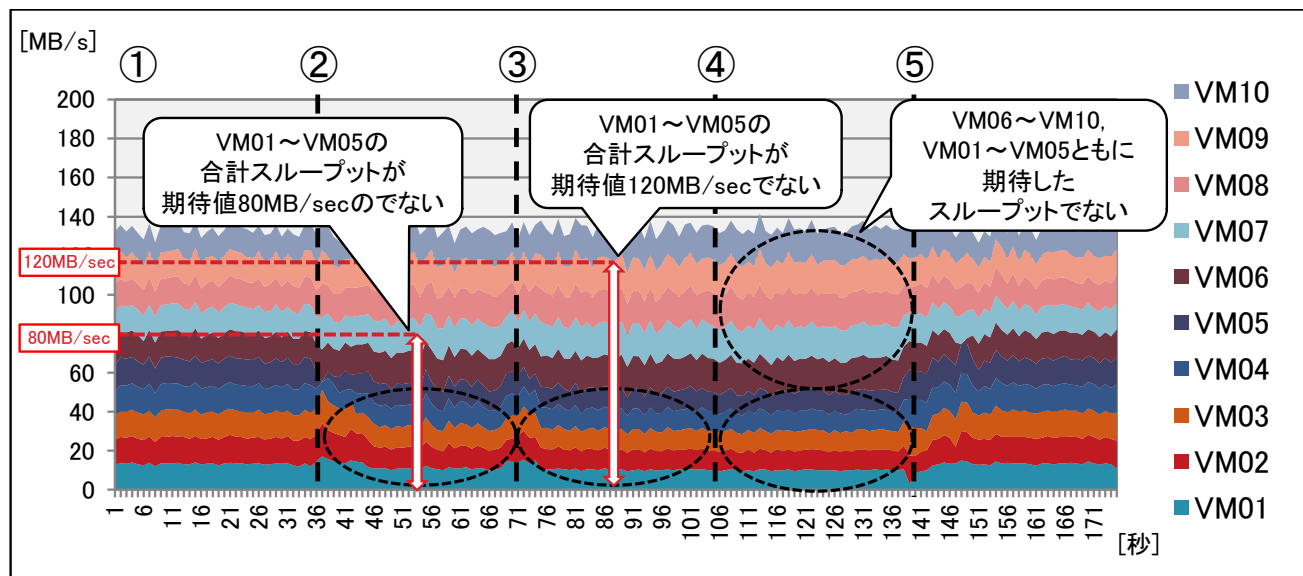


図 5. クラスター構成で全て「最小」を設定した際の仮想ディスクスループットの変化

図 5 の①～⑤すべての時点において全 VM に対して均等に Iometer による負荷がかかっています。

①の時点は、全 VM に対して IOPS 未設定の状態です。

②の時点でノード 1 側の VM01～VM05 の仮想ディスクそれぞれに対して最小 2000IOPS (16MB/sec の確保)を設定し、合計 80MB/sec のスループットを期待しましたが、期待どおりにはなりません。その後、③の時点で最小 3000IOPS (24MB/sec の確保)に設定変更しましたが、ここでも期待どおりの合計値 120MB/sec にはなりません。また、②の時点から約 60 秒経過したところ(95 秒目付近)で、設定したスループットにならないことを示す通知がノード 1 側のイベントログに記録されました。

さらに、④の時点でノード 2 側の VM06～VM10 の仮想ディスクに最小 3000IOPS (24MB/sec の確保)を設定したところ、VM06～VM10 も期待どおりのスループットになりませんでした。

このように、クラスター構成においては I/O 帯域を確保するために「最小」を設定しても、設定したとおりにならない場合があることがわかりました。これは、相手ノード上に存在する VM に対してはスループットを制御できないため、相手ノード上の VM の負荷状況に影響を受けてしまうためと考えられます。

4. クラスター構成において VM ごとに I/O 帯域を確保したい場合の対策

前述の検証では、クラスター構成において単体サーバと同様に「最小」を設定しても、設定したとおりのスループットにはならないことがわかりましたが、重要度が高い VM に対して確実に I/O 帯域を確保したい場合があります。そのような場合には、ストレージ装置の最大 I/O 性能を超えないように、全ての VM に対して、1VM あたり使用可能な「最大」を設定することで、ディスクの I/O 帯域を確保できるのではと推測しました。以下にイメージ図を示します。

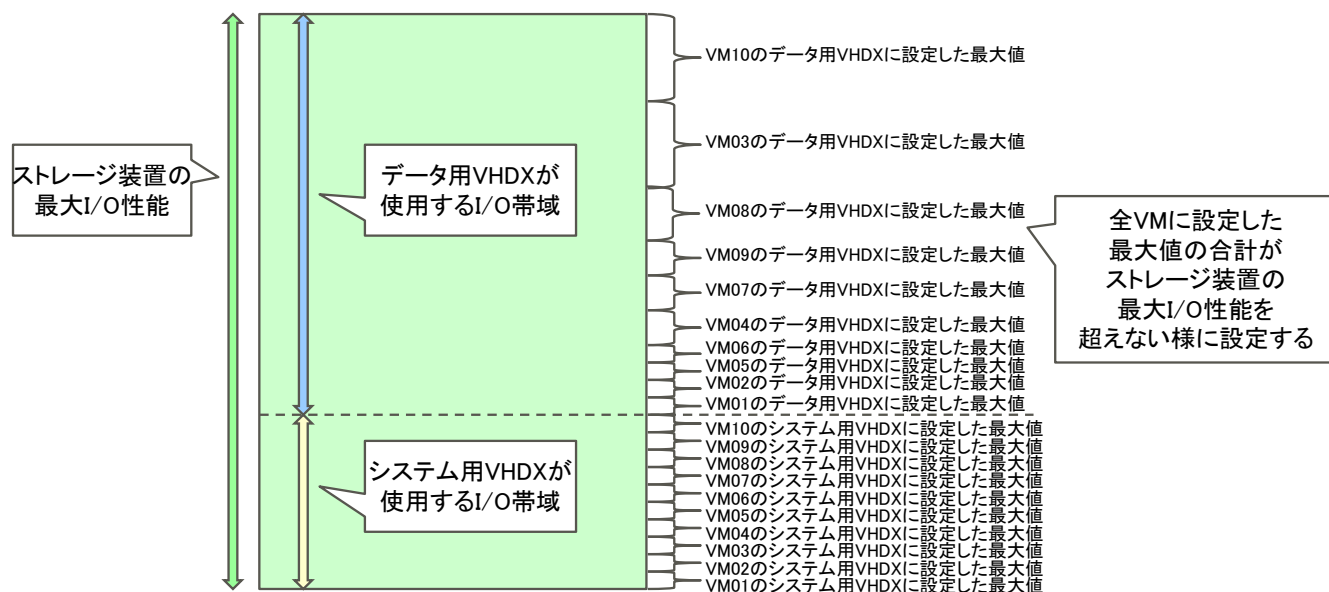


図 6. クラスタ構成で、ストレージ装置の最大 I/O 性能を超えないように全 VM に「最大」を設定したイメージ

そこで、全ての VM に「最大」設定した際に、相手ノード上 VM の負荷状況に影響を受けないことを確認するために、図 4 のクラスタ構成で IOPS 設定を以下のとおり変化させて、仮想ディスクのスループットを測定しました。

- | | |
|--|--|
| ① VM01～VM05:IOPS 未設定 | VM06～VM10:IOPS 未設定 |
| ② VM01～VM05:最大 500 IOPS を設定 (4MB/sec に制限) | VM06～VM10:IOPS 未設定 |
| ③ VM01～VM05:最大 1000 IOPS を設定 (8MB/sec に制限) | VM06～VM10:IOPS 未設定 |
| ④ VM01～VM05:最大 1000 IOPS を設定 (8MB/sec に制限) | VM06～VM10:最大 1000 IOPS を設定 (8MB/sec に制限) |
| ⑤ VM01～VM05:IOPS 未設定 | VM06～VM10:IOPS 未設定 |

以下に測定結果を示します。

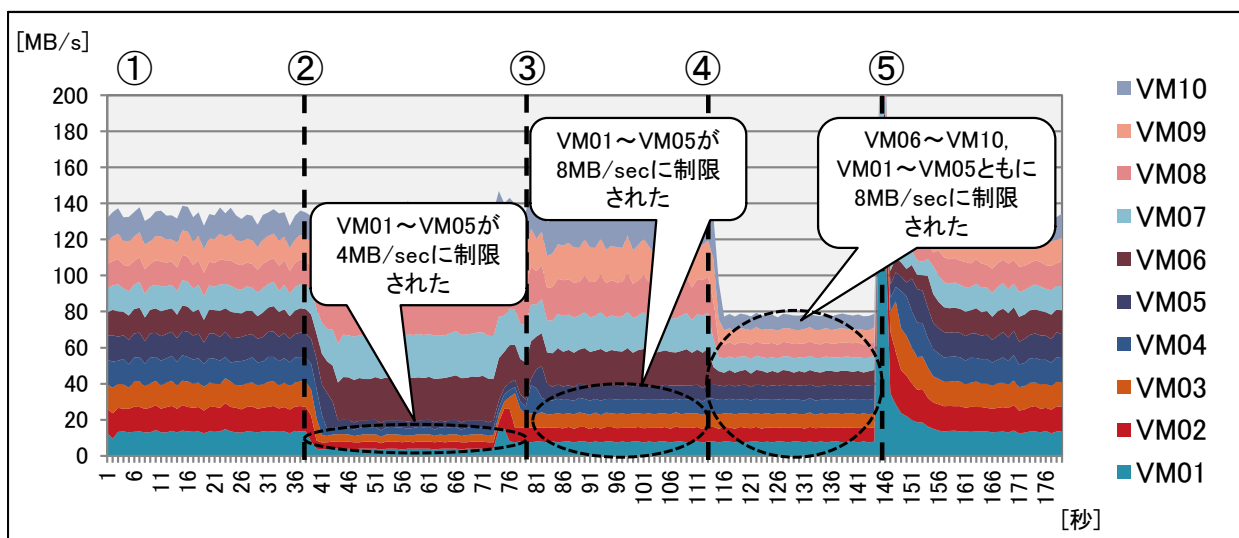


図 7. クラスタ構成で全て「最大」を設定した際の仮想ディスクスループットの変化

図7の①～⑤すべての時点において全 VM に対して均等に Iometer による負荷がかかっています。

①の時点は、全 VM に対して、IOPS 未設定の状態です。

②の時点でノード1側の VM01～VM05 の仮想ディスクそれぞれに対して最大 500IOPS (4MB/sec の制限)を設定したところ、スループットは 4MB/sec に制限され、未設定のノード2側の VM06～VM10 のスループットが高くなりました。次に、③の時点でノード1側の VM01～VM05 の仮想ディスクそれぞれに対して最大 1000IOPS(8MB/sec の制限)を設定したところ、スループットは 8MB/sec に制限され、未設定のノード2側の VM06～VM10 のスループットが高くなりました。さらに、④の時点で VM06～VM10 の仮想ディスクそれぞれに対して最大 1000IOPS (8MB/sec の制限)を設定したところ、全ての VM のスループットが 8MB/sec に制限されました。

これらの結果から、クラスター構成においてVMごとに仮想ディスクI/O帯域を確保するには、ストレージ装置の最大性能値を超えないように²、全てのVMに対して、1VMあたり使用可能な「最大」を設定する方法が有効と言えます。

ディスクI/O性能の測定方法については、「FUJITSU PRIMERGY サーバ ディスクI/O パフォーマンスの基本」を参照ください。

(<http://jp.fujitsu.com/platform/server/primergy/performance/pdf/wp-basics-of-disk-io-performance-ww-ja.pdf>)

FUJITSU Storage ETERNUS のディスク性能についてはカタログ等でご確認ください。

5. まとめ

Windows Server 2012 R2 Hyper-V のストレージ QoS を利用することで、VM ごとのディスクI/O性能を制御でき、安定したパフォーマンスを維持できるようになります。今回の検証では、VM 集約では一般的な構成になりつつあるクラスター構成において、仮想ディスクの I/O 帯域を確保する方法として、全 VM に対して「最大」設定する方法を紹介しました。

(注意) 先行評価した Windows Server 2012 R2 は開発段階にあるため仕様変更の可能性があります。また、測定値については、開発中の機能であることと評価環境により左右されるため、参考値としてご覧ください。

² より安全に使用するためには、最大性能値の 90%程度に設定してください。

PC サーバ FUJITSU Server PRIMERGY につきましては、以下の技術情報を参照願います。

- ・PC サーバ FUJITSU Server PRIMERGY (プライマジー)

<http://jp.fujitsu.com/platform/server/primergy/>

- ・FUJITSU Server PRIMERGY 機種比較表

<http://jp.fujitsu.com/platform/server/primergy/products/lineup/select-spec/>

- ・FUJITSU Server PRIMERGY サーバ選定ガイド

<http://jp.fujitsu.com/platform/server/primergy/products/lineup/select-model/>

PC サーバ FUJITSU Server PRIMERGY のお問い合わせ先。

- ・PC サーバ FUJITSU Server PRIMERGY お問い合わせ

<http://jp.fujitsu.com/platform/server/primergy/contact/>

基幹 IA サーバ FUJITSU Server PRIMEQUEST につきましては、以下の技術情報を参照願います。

- ・基幹 IA サーバ FUJITSU Server PRIMEQUEST (プライムクエスト)

<http://jp.fujitsu.com/platform/server/primequest/>

- ・FUJITSU Server PRIMEQUEST 1000 シリーズ 製品ラインナップ

<http://jp.fujitsu.com/platform/server/primequest/products/>

基幹 IA サーバ FUJITSU Server PRIMEQUEST のお問い合わせ先。

- ・本製品のお問い合わせ

<http://jp.fujitsu.com/platform/server/primequest/contact/>