

富士通サーバ PRIMERGY/ストレージ ETERNUS と  
VIRIDENT FlashMAX での  
DataCore Software SANsymphony-V による  
ハイブリッドストレージ環境の効果検証

－検証報告書－

Ver. 1.0



## 目次

<b>1</b>	<b>はじめに</b> .....	<b>5</b>
<b>1.1</b>	<b>検証の背景</b> .....	<b>5</b>
1.1.1	Flash SSD について.....	5
1.1.2	FlashSSD を導入するにあたっての課題.....	5
1.1.3	ストレージ仮想化ソフト SAN Symphony-V で解決する .....	6
1.1.4	SAN Symphony-V の導入の効果 .....	6
<b>1.2</b>	<b>検証の狙い目</b> .....	<b>7</b>
<b>1.3</b>	<b>本文での表記</b> .....	<b>7</b>
<b>2</b>	<b>検証システム</b> .....	<b>8</b>
<b>2.1</b>	<b>使用ハードウェアおよび OS</b> .....	<b>8</b>
2.1.1	負荷サーバ .....	8
2.1.2	SSV サーバ .....	8
2.1.3	ストレージ .....	8
2.1.4	FlashSSD.....	9
2.1.5	その他 .....	9
<b>2.2</b>	<b>ソフトウェア</b> .....	<b>9</b>
2.2.1	SAN Symphony-V R9.....	9
2.2.2	ベンチマークツール .....	9
<b>2.3</b>	<b>システムおよび論理ドライブの構成</b> .....	<b>10</b>
2.3.1	素性能を測定する時の論理ドライブの割り当て .....	10
2.3.2	SSV サーバへの論理ドライブの割り当て.....	11
2.3.3	SSV で仮想化された論理ドライブの割り当て.....	12
<b>2.4</b>	<b>論理ドライブまとめ</b> .....	<b>14</b>
2.4.1	構成した論理ドライブ一覧.....	14
2.4.2	負荷サーバからの論理ドライブの認識 .....	15
<b>3</b>	<b>ベンチマークプラン</b> .....	<b>16</b>
<b>3.1</b>	<b>ベンチマークテスト</b> .....	<b>16</b>
3.1.1	IOmeter のテストパターン .....	16
3.1.2	Oracle ORION のテストパターン.....	16
<b>3.2</b>	<b>ベンチマークテストの対象とした論理ドライブ</b> .....	<b>17</b>
<b>4</b>	<b>ベンチマークテスト結果</b> .....	<b>18</b>
<b>4.1</b>	<b>IOmeter for Windows</b> .....	<b>18</b>
4.1.1	DX80S2 の素性能と SSV で仮想化したときのスループット比較 .....	18
4.1.2	Virident SSD の素性能と SSV で仮想化したときのスループット比較 .....	19
4.1.3	DX80S2 の素性能と SSV で 4 種類のパターンで仮想化したときのスループット比較.....	20

4.1.4	IOPS での比較 (その①)	21
4.1.5	IOPS での比較 (その②)	22
<b>4.2</b>	<b>IOmeter for Linux</b>	<b>23</b>
4.2.1	スループットの比較	23
4.2.2	IOPS 比較	24
<b>4.3</b>	<b>Oracle ORION for Linux</b>	<b>25</b>
4.3.1	OLTP リード IOPS	25
4.3.2	OLTP ライト IOPS	25
4.3.3	OLTP リード遅延	26
4.3.4	OLTP ライト遅延	26
4.3.5	DWH リードスループット	27
4.3.6	DWH ライトスループット	28
<b>5</b>	<b>まとめ</b>	<b>29</b>
<b>5.1</b>	<b>IOmeter for Windows</b>	<b>29</b>
5.1.1	スループットでの比較	29
5.1.2	IOPS での比較	29
<b>5.2</b>	<b>IOmeter for Linux</b>	<b>30</b>
<b>5.3</b>	<b>Oracle ORION for Linux</b>	<b>30</b>
5.3.1	OLTP の IOPS での比較	30
5.3.2	OLTP の IO 遅延での比較	30
5.3.3	DWH のスループットでの比較	30
<b>5.4</b>	<b>検証結果を振り返って</b>	<b>31</b>
<b>5.5</b>	<b>本検証に関するお問い合わせ先</b>	<b>31</b>
<b>付録</b>		<b>32</b>
<b>付録 1</b>	<b>Virident Flash MAX の仕様</b>	<b>32</b>
<b>付録 2</b>	<b>Virident Flash MAX を搭載したときのようす</b>	<b>33</b>
<b>付録 3</b>	<b>SSV におけるデータ配置のようす</b>	<b>34</b>
付録 3-1	ストレージの階層構造	34
付録 3-2	ハイブリッドストレージの構成例	35

# 1 はじめに

---

## 1.1 検証の背景

### 1.1.1 Flash SSD について

フラッシュメモリを利用した半導体ディスク(以下、FlashSSD)は、ハードディスクに比べてランダムアクセスの読み出し性能に優れており、発熱が少ない、対衝撃性に優れ故障し難いという利点がある。

また、容量単価は依然としてハードディスクよりは高価ではあるが、コストパフォーマンスの面を加味すると普及価格帯になってきたので、特にデータセンタ向けの業務サーバに導入されるケースが増えてきた。

FlashSSD のインターフェイスは、SATA、USB、PCI-Express に対応したものがあるが、サーバ用途では PCI-Express を使ったものが主流である。

### 1.1.2 FlashSSD を導入するにあたっての課題

スループットの改善や消費電力の低減では、優れた利点を持つ FlashSSD だが PCI-Express のスロットに取り付けて使用する装置なので次のような課題がある。

- ① いわゆる単一故障点 (Single Point of Failure) となる
- ② 複数のサーバから共有ディスクとして利用出来ない (=クラスタリング構成が出来ない)
- ③ 専用のドライバをインストールする必要がある。
- ④ 1U のサーバでは PCI-Express の空きスロットがないことがある。
- ⑤ 容量が限られているので、アクセス頻度の高いホットデータを洗い出して再配置する必要がある。

一般に、スループットの改善が必要となるのはミッションクリティカルな業務サーバなので、①、②、⑤は Flash SSD の導入を検討するときには大きな障壁になってしまう。

### 1.1.3 ストレージ仮想化ソフト SAN Symphony-V で解決する

この問題を解決する手段として、FC や Infiniband、iSCSI インターフェイスを持つ独立した筐体を持つ半導体ディスクや FlashSSD を搭載できる RAID 装置もあるが、これらは PCI-Express インターフェイスの FlashSSD と比べると高価であり、また、⑤のデータの再配置という問題は依然残ってしまう。

これを解決するのが、データコア・ソフトウェア社の SAN Symphony-V (以下、SSV と称す)だ。

SSV は、Windows2008R2 ベースの IA サーバ (以下、SSV サーバと称す) にインストールすれば、そのサーバを仮想ストレージアプライアンスとして機能させることができるソリューションである。

SSV は、配下のストレージを完全に仮想化し、業務サーバに iSCSI または FC 接続の論理ドライブを提供するだけでなく、その論理ドライブには次の機能を持たせることが可能となる。

-  階層型 仮想ディスクプール
-  オンライン スナップショット
-  SNIA準拠のCDP
-  同期ミラーによるHA
-  簡単なデータ移行
-  高速Read/Writeキャッシュ
-  バックエンド ロードバランス
-  リモートレプリケーション
-  シンプロビジョニング

### 1.1.4 SAN Symphony-V の導入の効果

FlashSSD を SSV サーバに搭載して仮想化することで①～⑤の問題は一挙に解決する。

PCI-ExpressタイプのFlashSSDの問題点	SSVの導入による解決
単一故障点 (Single Point of Failure) となる	SSVサーバ間でのHAと同期ミラーリングによる完全な二重化
複数の業務サーバから共有ディスクとして利用出来ない (=クラスタリング構成が出来ない)	業務サーバからは、iSCSI/FC接続のディスク装置として認識させる。 ・専用のドライバは不要である。 ・複数の業務サーバから同じLUNを共有できるのでクラスタリングも可能。
専用のドライバーをインストールする必要がある。	・業務サーバにはFlash SSDを搭載しなくて良い
1UのサーバではPCI-Expressの空きスロットがないことがある。	
容量が限られているので、アクセス頻度の高いホットデータを洗い出して再配置する必要がある。	階層型ストレージプールの機能により、データの再配置は不要である。

さらに、FlashSSD には無かった、スナップショット、最大 48 時間の CDP（継続的データ保護）、リモートレプリケーション、シンプロビジョニングといった機能が付加されるので、より投資効果の高いストレージシステムが構築できることになる。

## 1.2 検証の狙い目

FlashSSD を SSV に搭載して仮想化することで、機能的な問題は解決するが、そのためにスループットが犠牲になってしまうと導入の効果が薄れてしまう。

今回の検証では、sourceforge.jp が開発管理している IOmeter と Oracle 社の ORION というベンチマークツールを使って、FlashSSD および RAID 装置の素性能および SSV で仮想化した場合の性能の比較、ストレージを階層化したときの効果から、SSV の導入の利点と課題を明確化することを目的とする。

なお、タイトルにあるハイブリッドストレージとは、SSV のストレージ階層化の機能を使って、FlashSSD と RAID を組み合わせることで、両者の利点を併せ持ったストレージが構築できることを意味して名付けたものである。（詳細については、2.3.3 も参照のこと）

## 1.3 本文での表記

本資料内では下記ハードウェア／ソフトウェアを以下のように記載する。

	ハードウェア	本資料での表記
1	PRIMERGY RX200 S6	RX200S6
2	PRIMERGY RX300 S6	RX300S6
3	ETERNUS DX80 S2	DX80S2
4	ETERNUS SN200	SN200
5	Virident FlashMAX PCIe Storage Class Memory	Virident SSD

	ソフトウェア	本資料での表記
1	Windows Server 2008 R2	Windows2008R2
2	Red Hat Enterprise Linux 6.1	Red Hat Linux 6.1
3	SAN Symphony-V	SSV

## 2 検証システム

### 2.1 使用ハードウェアおよび OS

#### 2.1.1 負荷サーバ

ベンチマーク (IOmeter および ORION) を実行して、ストレージのパフォーマンスを測定するためのサーバである。OS は Windows2008R2 と RHEL6.1 を使用した。

富士通RX200S6(負荷サーバ#1)  
・Xeon X5680(3.33GHz/6core)×2  
・8GBメモリ  
・8GbFC-HBA 2ポート  
・Windows2008R2



富士通RX200S6(負荷サーバ#2)  
・Xeon X5680(3.33GHz/6core)×2  
・8GBメモリ  
・8GbFC-HBA 2ポート  
・Red Hat Linux 6.1 x86\_64



#### 2.1.2 SSV サーバ

SAN Symphony-V R9 をインストールして、仮想ストレージアプライアンスとして動作させる Windows2008R2 ベースのサーバで、負荷サーバからは FC 接続のストレージ装置として認識される。

富士通RX300S6(SSVサーバ#1)  
・Xeon X5680(3.33GHz/6core)×2  
・32GBメモリ  
・8GbFC-HBA 6ポート  
・Windows2008R2Standard  
・DataCore SSV-R9



富士通RX300S6(SSVサーバ#2)  
・Xeon X5680(3.33GHz/6core)×2  
・32GBメモリ  
・8GbFC-HBA 6ポート  
・Windows2008R2Standard  
・DataCore SSV-R9



#### 2.1.3 ストレージ

SSV サーバにストレージリソース (論理ドライブ) を提供する RAID 装置で、SSV サーバとは FC で接続される。なお、SSV の有無での効果を比較するために負荷サーバにもストレージリソースを提供している。

#### 富士通ETERNUS DX80S2

・8GbFC×4ポート  
・450GB(2.5"/10Krpm)×48本



## 2.1.4 FlashSSD

今回は、SLC（single-level cell）チップを用いた Virident FlashMAX PCIeStorageClassMemory（以下、Virident SSD）を使用した。詳細なスペックについては、付録を参照のこと。

ViridentFlashMAX  
PCIeStorageClassMemory

SLC 300GBモデル



## 2.1.5 その他

これ以外に、8GbFC スイッチ SN200 および 1GbE スイッチを配線に使用している。

## 2.2 ソフトウェア

### 2.2.1 SAN Symphony-V R9

下記の URL より評価版（SAN Symphony-V Release9.0）をダウンロードしてインストールおよび設定を行った。なお、この報告書では具体的な構築手順の説明は省略する。

<http://www.datacore.com/30-Day-Trial.aspx>

なお、評価版は 30 日間の使用制限を除けば、機能的には製品版と同じである。

### 2.2.2 ベンチマークツール

#### IOmeter

テストパターンを作成することで、テストファイルに対して IO を行ってパフォーマンスを測定するベンチマークテストである。

下記の URL より IOmeter をダウンロードしてベンチマークを行った。

[http://sourceforge.jp/projects/sfnet\\_iometer/](http://sourceforge.jp/projects/sfnet_iometer/)

ただし、バージョンは最新でなく IOmeter Version 2008.06.18-RC2 を使用した。

## Oracle ORION

Oracle データベースをインストールすること無しに、Oracle の IO をシミュレートしてパフォーマンスを測定できるベンチマークツールである。

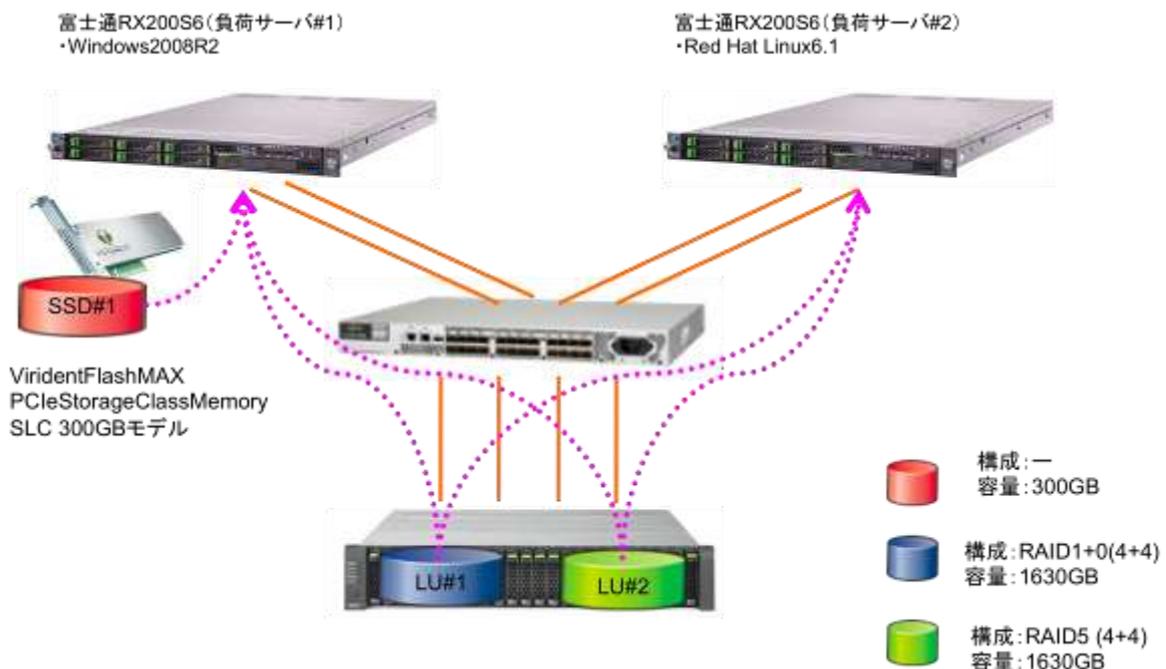
下記の URL より Oracle ORION をダウンロードしてベンチマークを行った。

<http://www.oracle.com/technetwork/jp/topics/index-096484-ja.html>

## 2.3 システムおよび論理ドライブの構成

### 2.3.1 素性能を測定する時の論理ドライブの割り当て

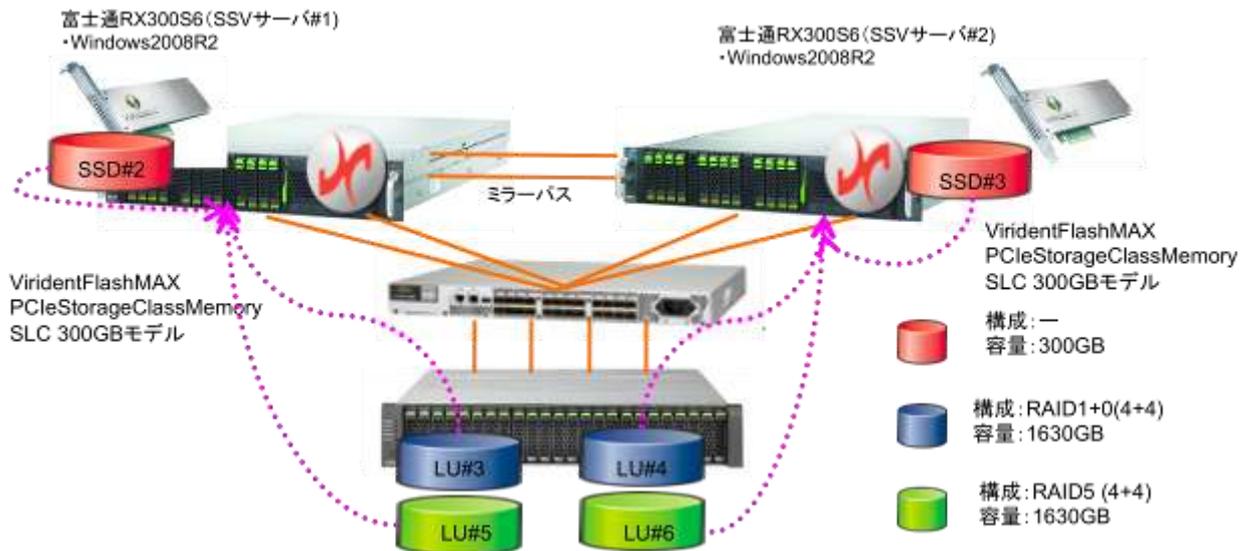
SSV の有無での効果を比較するために、DX80S2 および Virident SSD を負荷サーバに直接接続したときの構成および論理ドライブの割り当てが下図である。



※同じ論理ドライブ（ファイルシステム）は共有できないので、今回は負荷サーバ#1、#2 は論理ドライブ内で100GB 単位のパーティションに分割して、それぞれ異なるパーティションにファイルシステムを作成してマウントさせている。

## 2.3.2 SSV サーバへの論理ドライブの割り当て

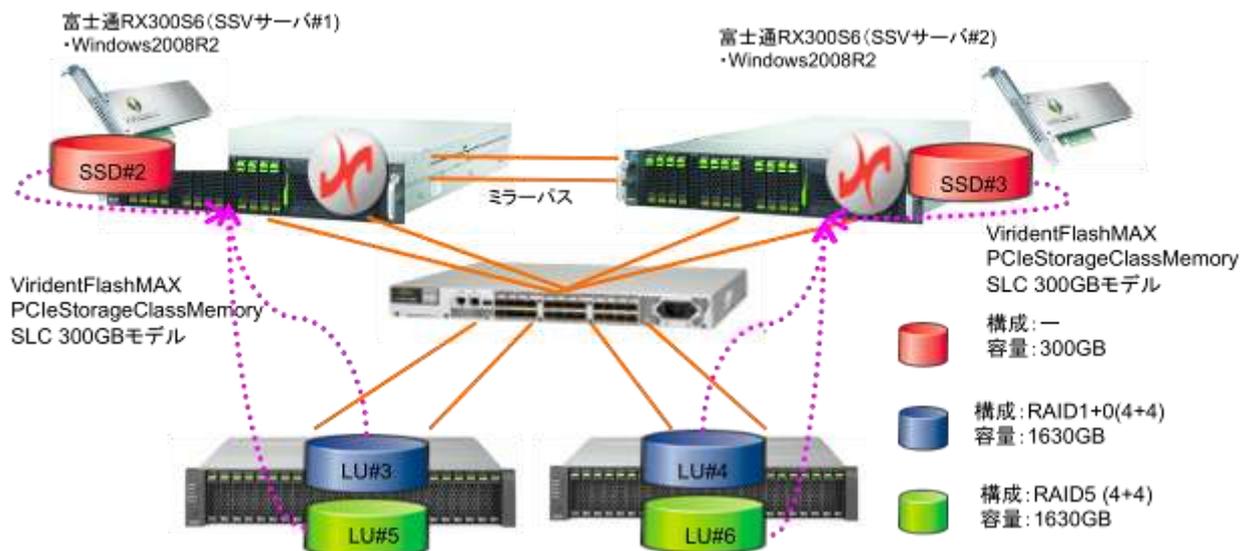
SSV サーバへの DX80S2 と Virident SSD の接続および割り当ては下図の通りで、SSV サーバは、自分に割り当てられた論理ドライブをストレージプールに組み込んで、ストレージの仮想化を行う。



### 今回の構成での留意点

SSV は、独立した筐体のストレージを同期ミラーして二重化したうえで、論理ドライブに様々な機能を付加するソリューションなので推奨される構成は下図である。

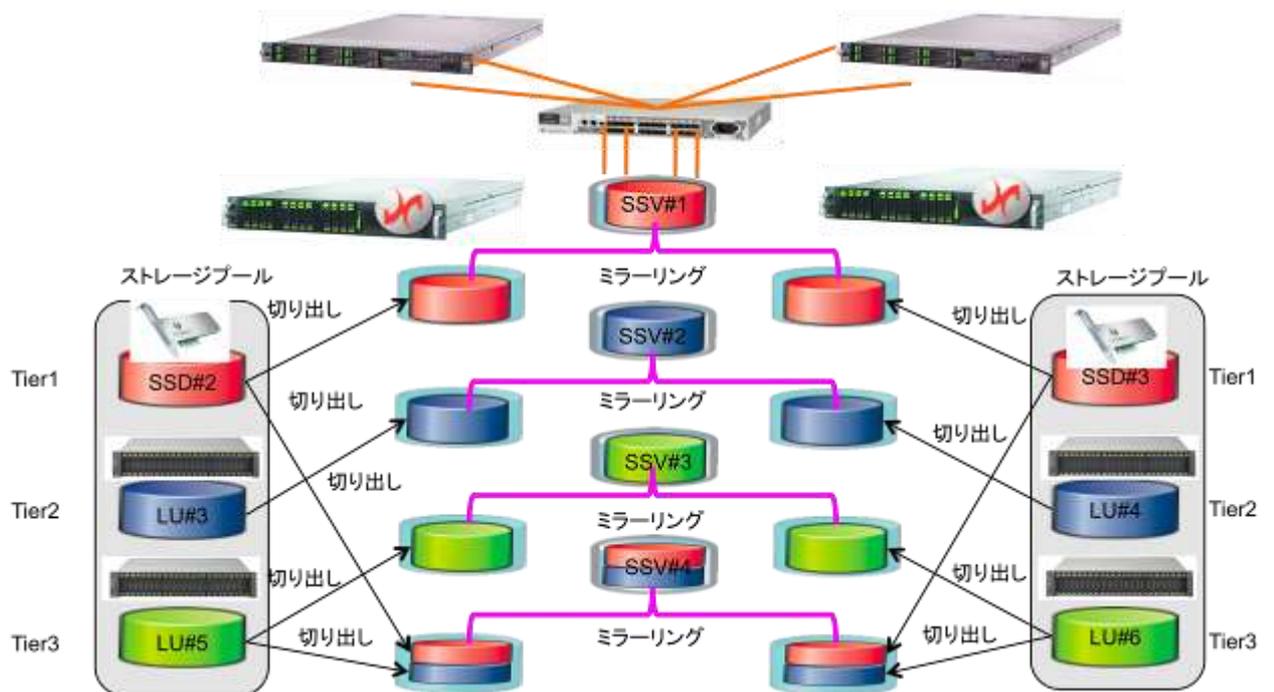
今回は機材の都合で一台の DX80S2 を SSV サーバで共有しているため、この検証結果が SSV 本来の性能の全てでないことは留意しておく必要がある。



### 2.3.3 SSV で仮想化された論理ドライブの割り当て

SSV は、Virident SSD および DX80S2 をストレージプールに組み込んだ際に Tier1~3 で階層化し、このストレージプールから同期ミラーリングおよび適切なプロファイルが設定された仮想ディスクを切り出して、負荷サーバに論理ドライブとして割り当てる。

そのため、負荷サーバからは割り当てられた論理ドライブが二重化されていることを意識する必要がなくなり、もし、いずれかの SSV サーバや DX80S2 の論理ドライブがアクセス不可となったとしても、データへのアクセスは中断されることがない。



#### Tier とプロファイルについて

SSV ではストレージプールを最高 15 段階に分けることができ、Tier 番号が若い程高速なストレージであると見なす。

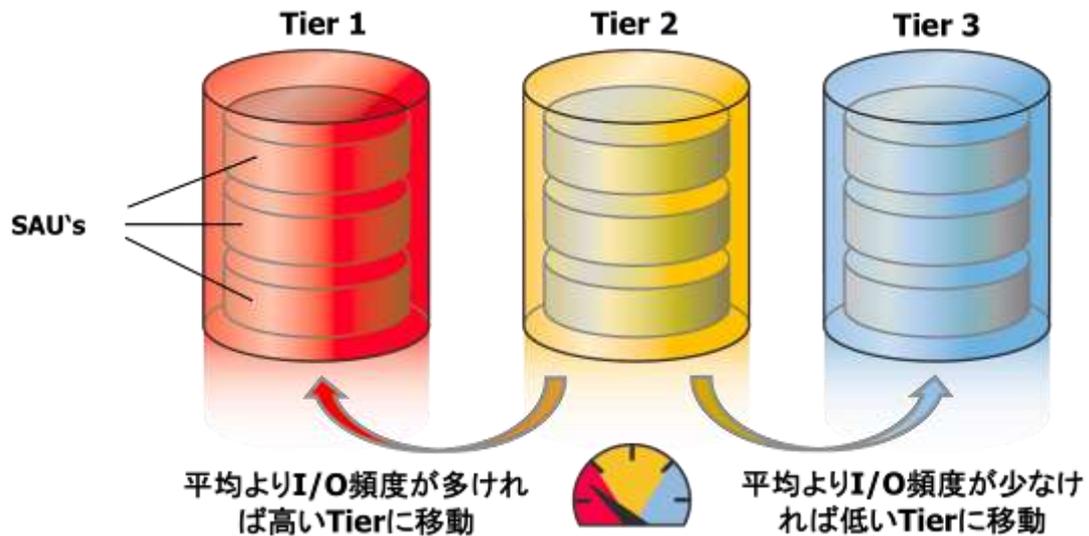
ここでは、Virident SSD を Tier1、DX80S2 の RAID0+1 で構成された論理ドライブを Tier2、RAID5 で構成された論理ドライブを Tier3 としている。

SSV で仮想ディスクを切り出すときは、Critical、High、Normal、Low、Archive の 5 段階のプロファイルを指定することができる。

Normal プロファイルでは、仮想ディスクは Tier1~3 のリソースから論理ドライブを切り出され、運用を開始すると、SSV は頻繁にアクセスされるものは Tier1 に、頻度が少ないものは 3 にデータブロックを自動的に移動する。

これにより、Virident SSD の実際の容量に関わらず、常に、Virident SSD のパフォーマンスを持った大容量の論理ドライブとして扱えるし、もし、全てのデータに Virident SSD のパフォーマンスが必要とするのならプロファイルを **Critical** にしてするだけで良い。

Tier とプロファイルはいつでも変えられるので、非常に柔軟なストレージ管理が実現できるのが SSV を導入する利点である。

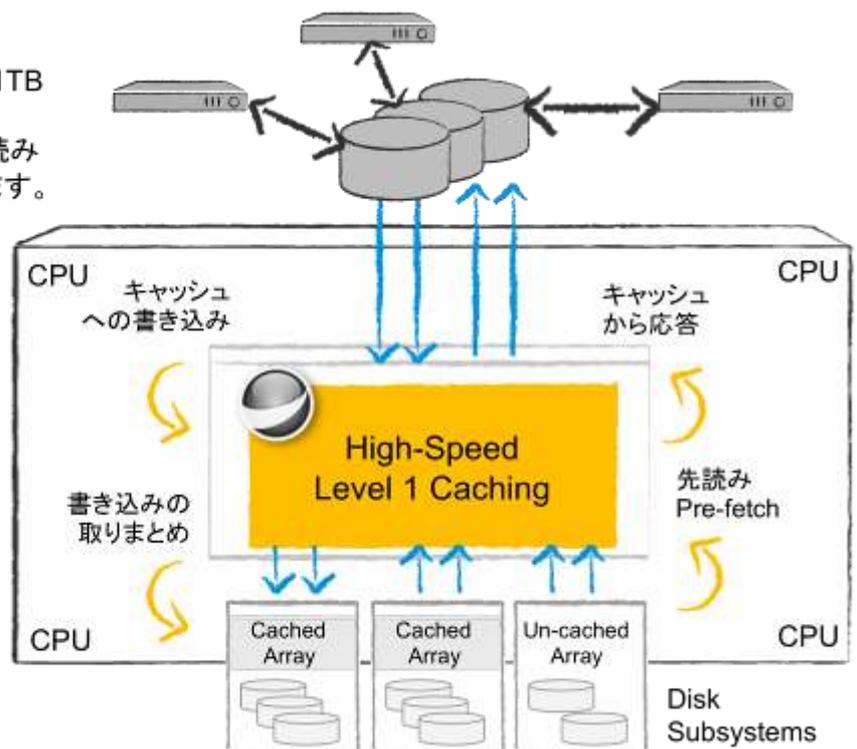


### SSV のキャッシュ機能について

ベンチマーク結果を見る上で欠かせない SSV のキャッシュとは、SSV サーバに搭載されたメモリをキャッシュとして使用することで IO の応答性能を向上させる機能である。

- データコアサーバ毎に、最大1TBのキャッシュを提供可能です。
- 書き込みの取りまとめと、先読みを行い、I/O応答性能を上げます。

Avg. Response ( $\mu$ S)	
DataCore L1 Cache	<20
Array L2 Cache	250 – 300
Disk Drive	4000 – 6000



## 2.4 論理ドライブまとめ

### 2.4.1 構成した論理ドライブ一覧

2.3 章で構成した論理ドライブを下表にまとめた。

論理ドライブ名	容量	ハードウェア	備考
SSD#1	300GB	Virident SSD	負荷サーバ#1のPCIeスロットに搭載、VIRIDENTのドライバーで認識させた論理ドライブ
LU#1	1,680GB	DX80S2 (RAID1+0(4+4))	負荷サーバ#1,2のFC-HBAから直接認識させたDX80S2の論理ドライブ
LU#2	1,680GB	DX80S2 (RAID5(4+1))	負荷サーバ#1,2のFC-HBAから直接認識させたDX80S2の論理ドライブ
SSD#2	300GB	Virident SSD	SSVサーバ#1のPCIeスロットに搭載、VIRIDENTのドライバーで認識させた論理ドライブ(負荷サーバからは見えない)
SSD#3	300GB	Virident SSD	SSVサーバ#2のPCIeスロットに搭載、VIRIDENTのドライバーで認識させた論理ドライブ(負荷サーバからは見えない)
LU#3	1,680GB	DX80S2 (RAID1+0(4+4))	SSVサーバ#1のFC-HBAから直接認識させたDX80S2の論理ドライブ(負荷サーバからは見えない)
LU#4	1,680GB	DX80S2 (RAID5(4+1))	SSVサーバ#2のFC-HBAから直接認識させたDX80S2の論理ドライブ(負荷サーバからは見えない)
LU#5	1,680GB	DX80S2 (RAID1+0(4+4))	SSVサーバ#1のFC-HBAから直接認識させたDX80S2の論理ドライブ(負荷サーバからは見えない)
LU#6	1,680GB	DX80S2 (RAID5(4+1))	SSVサーバ#2のFC-HBAから直接認識させたDX80S2の論理ドライブ(負荷サーバからは見えない)
SSV#1 (Tier1のみ)	1,680GB (シンプロ)	SSD#2+SSD#3	SSVサーバ#1、SSVサーバ#2が認識しているVIRIDENT SSDの論理ドライブから切り出して同期ミラーリング設定を行った論理ドライブ→負荷サーバ#1、#2から認識される
SSV#2 (Tier2のみ)	1,680GB (シンプロ)	LU#3+LU#4	SSVサーバ#1、SSVサーバ#2が認識しているDX80S2の論理ドライブから切り出して同期ミラーリング設定を行った論理ドライブ→負荷サーバ#1、#2から認識される
SSV#3 (Tier3のみ)	1,680GB (シンプロ)	LU#5+LU#6	SSVサーバ#1、SSVサーバ#2が認識しているDX80S2の論理ドライブから切り出して同期ミラーリング設定を行った論理ドライブ→負荷サーバ#1、#2から認識される
SSV#4 (Tier1+2)	1,680GB (シンプロ)	SSD#2+SSD#3 +LU#3+LU#4	SSVサーバ#1、SSVサーバ#2が認識しているVIRIDENT SSDおよびいるDX80S2の論理ドライブから切り出して同期ミラーリング設定を行った論理ドライブ→負荷サーバ#1、#2から認識される

## 2.4.2 負荷サーバからの論理ドライブの認識

下図は負荷サーバ#1 から認識された各論理ドライブである。

ここから 100GB のパーティションを作成してベンチマーク用のドライブとしている。

The screenshot shows the Windows Server Disk Management console for server 'サーバー マネージャー (R0020056-0)'. The main pane displays a list of disks and their partitions. On the right, there is a '操作' (Action) pane with 'ディスクの管理' (Disk Management) selected, showing a vertical stack of logical drive icons: SSV#2 (blue), LU#3 (blue), LU#5 (green), SSV#3 (green), and SSV#1 (red).

ディスク	パーティション	容量	状態
ディスク 0	system (C:)	493.75 GB	正常 (プライマリ パーティション)
ディスク 1	IOmeter用SSV-RAID1+0 (M:)	1530.00 GB	正常 (プライマリ パーティション)
ディスク 2	IOmeter用DX00-RAID1+0 (E:)	1530.00 GB	正常 (プライマリ パーティション)
ディスク 3	IOmeter用DX00-RAID5 (F:)	1530.00 GB	正常 (プライマリ パーティション)
ディスク 4	(G:)	3.90 GB	正常 (プライマリ パーティション)
ディスク 5	IOmeter用SSV-RAID5 (N:)	1530.00 GB	正常 (プライマリ パーティション)
ディスク 6	IOmeter用VSS-SSD (T:)	1530.00 GB	正常 (プライマリ パーティション)

## 3 ベンチマークプラン

---

### 3.1 ベンチマークテスト

#### 3.1.1 IOmeter のテストパターン

- 100GB の全領域を IOmeter のテストデータで書き込み後テスト実施
- Worker (CPU 数) : 1
- テストパターン
  - ◇ データサイズ : 1KB、2KB、4KB、8KB、16KB、32KB、64KB、128KB、256KB、512KB、1MB、2 MB、4MB、8MB
  - ◇ IO パターン : シーケンシャルリード/ライト、ランダムリード/ライト
- IO 間隔 : 15 秒/データサイズ
- Outstanding IO : 1 および 16

#### 3.1.2 Oracle ORION のテストパターン

dd if=/dev/zero of=/mnt/xxxx bs=1024M で 100GB の全領域を書き込み後テスト実施

- OLTP シミュレーション 8KB/リード  
#orion -run advanced -testname \$TESTFILE\$ -num\_disks 16 -size\_small 8 -type rand -simulate raid0 -write 0 -duration 60 -matrix row -num\_large 0
- OLTP シミュレーション 8KB/リード  
#orion -run advanced -testname \$TESTFILE\$ -num\_disks 16 -size\_small 8 -type rand -simulate raid0 -write 100 -duration 60 -matrix row -num\_large 0
- DWH シミュレーション 1024KB/リード  
#orion -run advanced -testname \$TESTFILE\$ -num\_disks 16 -size\_large 1024 -type seq -num\_streamIO 1 -simulate raid0 -write 0 -cache\_size 0 -duration 60 -matrix col -num\_small 0
- DWH シミュレーション 1024KB/リード  
#orion -run advanced -testname DX80-RAID5-DWH-WR -num\_disks 16 -size\_large 1024 -type seq -num\_streamIO 1 -simulate raid0 -write 100 -cache\_size 0 -duration 60 -matrix col -num\_small 0

### 3.2 ベンチマークテストの対象とした論理ドライブ

実際にベンチマークを実施した論理ドライブは下表の通りである。

論理ドライブ名	容量	ハードウェア	IOmeter for Windows	IOmeter for Linux	Orion for Linux
SSD#1	300GB	Virident SSD	○	—	—
LU#1	1,680GB	DX80S2 (RAID1+0(4+4))	○	○	○
LU#2	1,680GB	DX80S2 (RAID5(4+1))	○	○	○
SSV#1 (Tier1)	1,680GB (シンプロ)	SSD#2+SSD#3	○	○	△
SSV#2 (Tier2のみ)	1,680GB (シンプロ)	LU#3+LU#4	○	○	△
SSV#3 (Tier3のみ)	1,680GB (シンプロ)	LU#5+LU#6	○	—	△
SSV#4 (Tier1+2)	1,680GB (シンプロ)	SSD#2+SSD#3 +LU#3+LU#4	○	—	○

※ Oracle ORION は、Windows2008R2 で未サポートだったためベンチマークテストの項目から外した。また、△となっているのは DWH のライトでデータ取得がされていなかったもの。

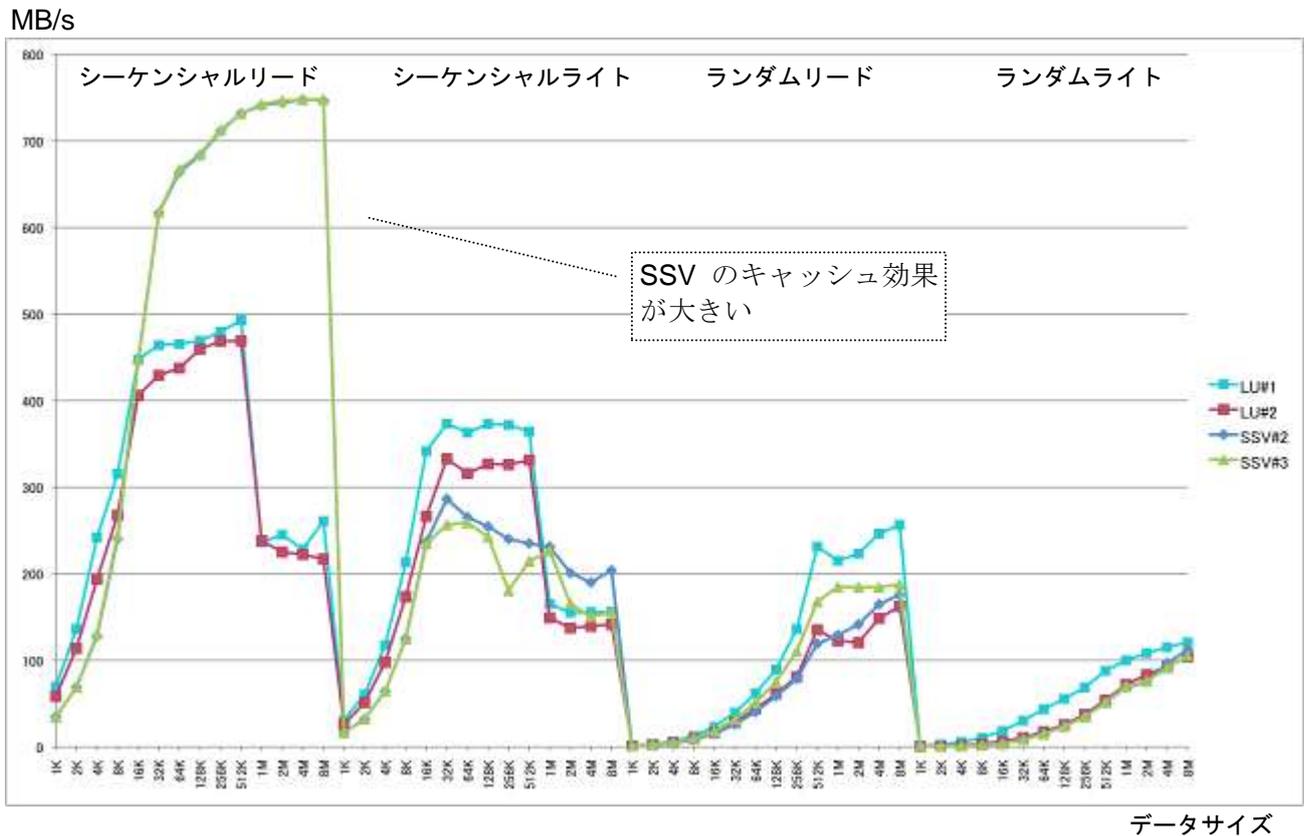
## 4 ベンチマークテスト結果

ここでは、ベンチマークテストの結果とその解説をする。なお、IOmeterについては、3.1.1項に示すテストパターンは全て行っているが、そこから傾向が把握し易いものを取り上げるものとする。

### 4.1 IOmeter for Windows

#### 4.1.1 DX80S2 の素性能と SSV で仮想化したときのスループット比較

DX80S2 の論理ドライブ LU#1、#2 を直接負荷サーバに割り当てたときと SSV サーバに DX80S2 の論理ドライブを割り当てて同期ミラーリングした SSV#2、#3 との比較である。



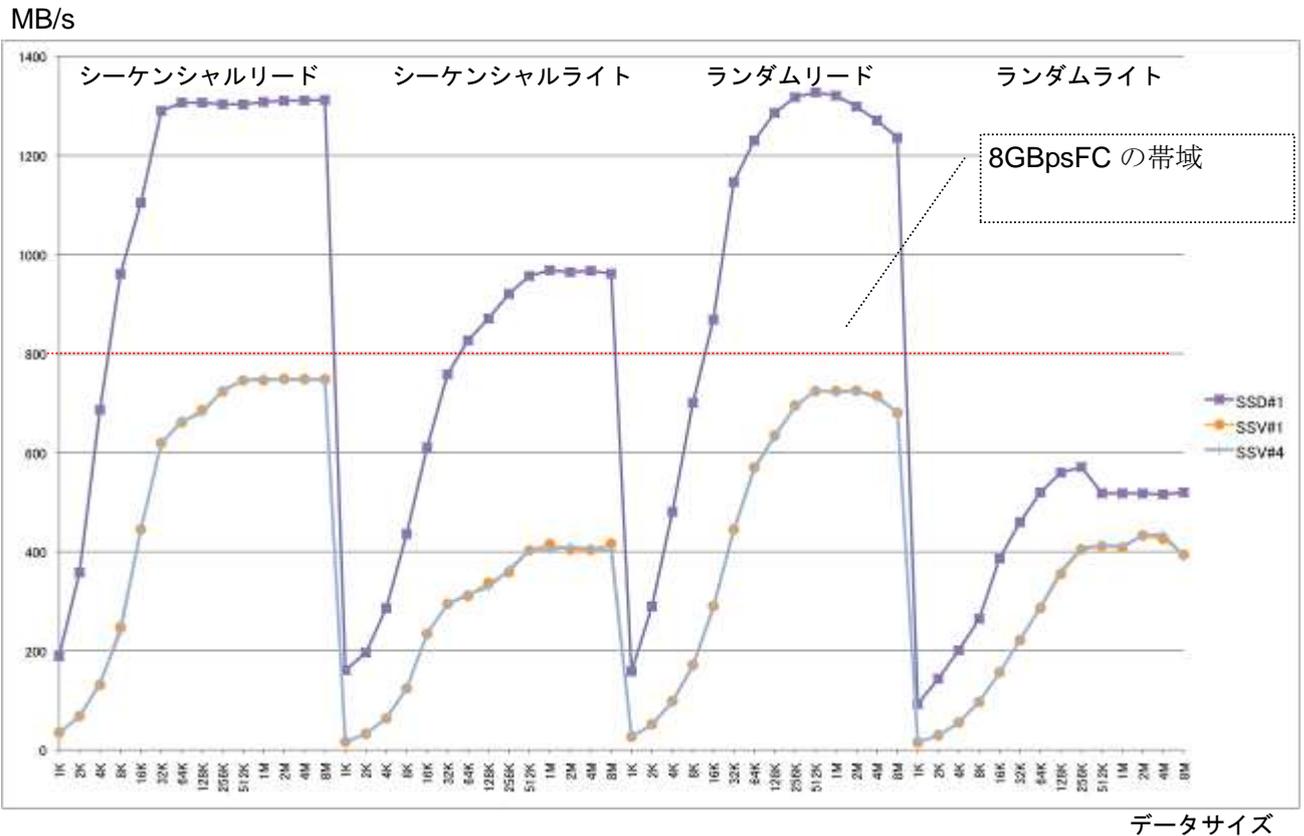
この結果から次のことが分かる。

- LU#1 (RAID1+0) と LU#2 (RAID5) でのスループットの差が見られる
- シーケンシャルリードでは、SSV のキャッシュの効果が大きい。
- シーケンシャルライトに関しては、直接接続よりも低くなっている※

※2.3.2 の留意点にある通り、DX80S2 の構成 (本来は 2 台別筐体をミラー) が影響している可能性が考えられる。

## 4.1.2 Virident SSD の素性能と SSV で仮想化したときのスループット比較

Virident SSD で構成した論理ドライブ SSD#1 を負荷サーバ#1 に直接接続および SSV で同期ミラーリングして構成した SSV#1、さらに DX80S2 の論理ドライブも加えて階層化した SSV#4 のベンチマークの比較である。

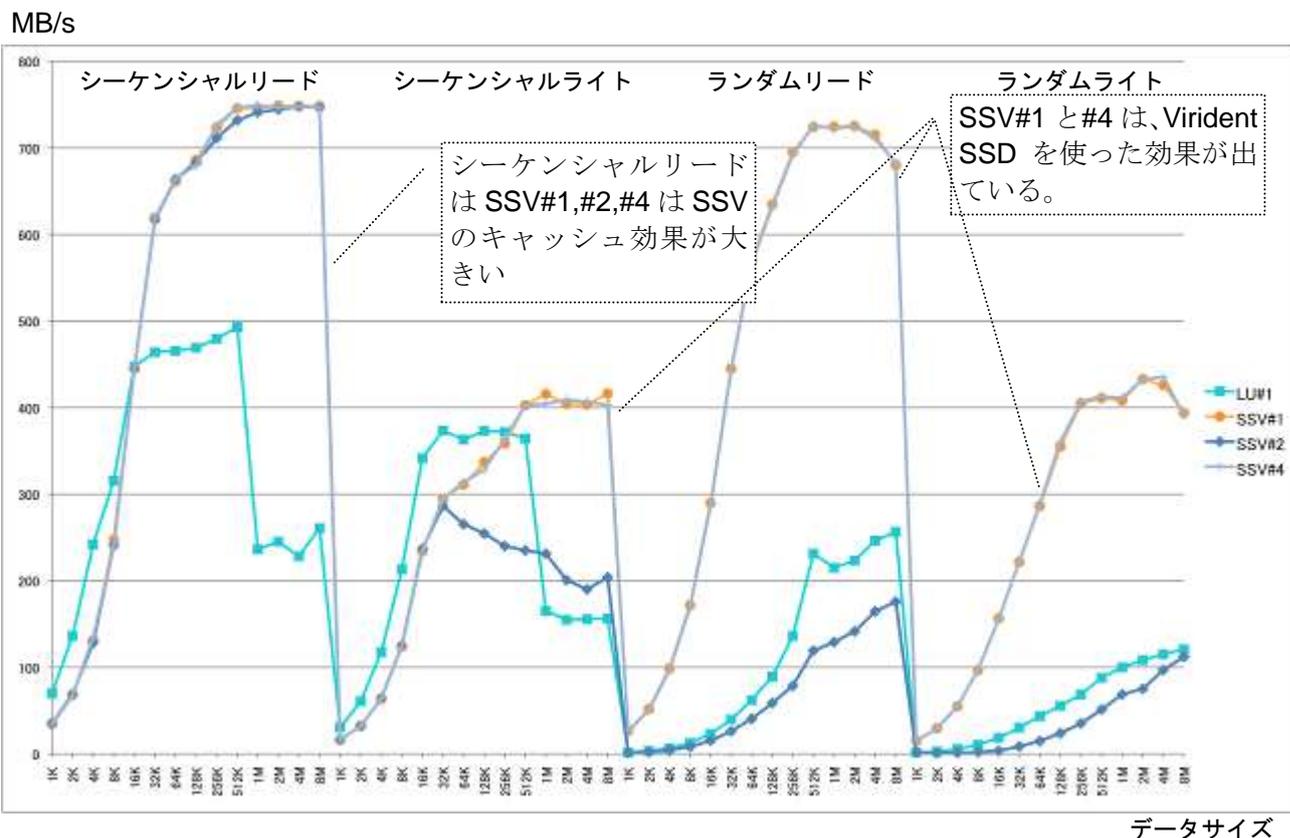


このグラフから以下のことが分かる。

- SSD#1 は、シーケンシャルリードで 1,300MB/S、ライトで 950MB/S なのでほぼカタログ値である。(カタログ値は付録も参照のこと)
- SSV で Virident SSD を同期ミラーリングした SSV#1 と#4 で、リードのスループットが 800MB/s 以下なのは、FC (8Gbps) の帯域によるものである。ただし、これを加味してもシーケンシャルライトのスループットは低い。
- SSV#1 と#4 はほぼ差が見られない。#4 は DX80S2 との階層構造となっているが、実際には全てのデータが Virident SSD 上にあったからと考えられる。

### 4.1.3 DX80S2の素性能とSSVで4種類のパターンで仮想化したときのスループット比較

DX80S2の論理ドライブを直接負荷サーバ#1に接続したLU#1とSSVで同期ミラーリングして構成したVirident SSD (SSV#1)、DX80S2 (SSV#2)、Virident SSDとDX80S2を階層化 (SSV#4)の4種類の仮想ディスクのベンチマーク結果を比較したものである。



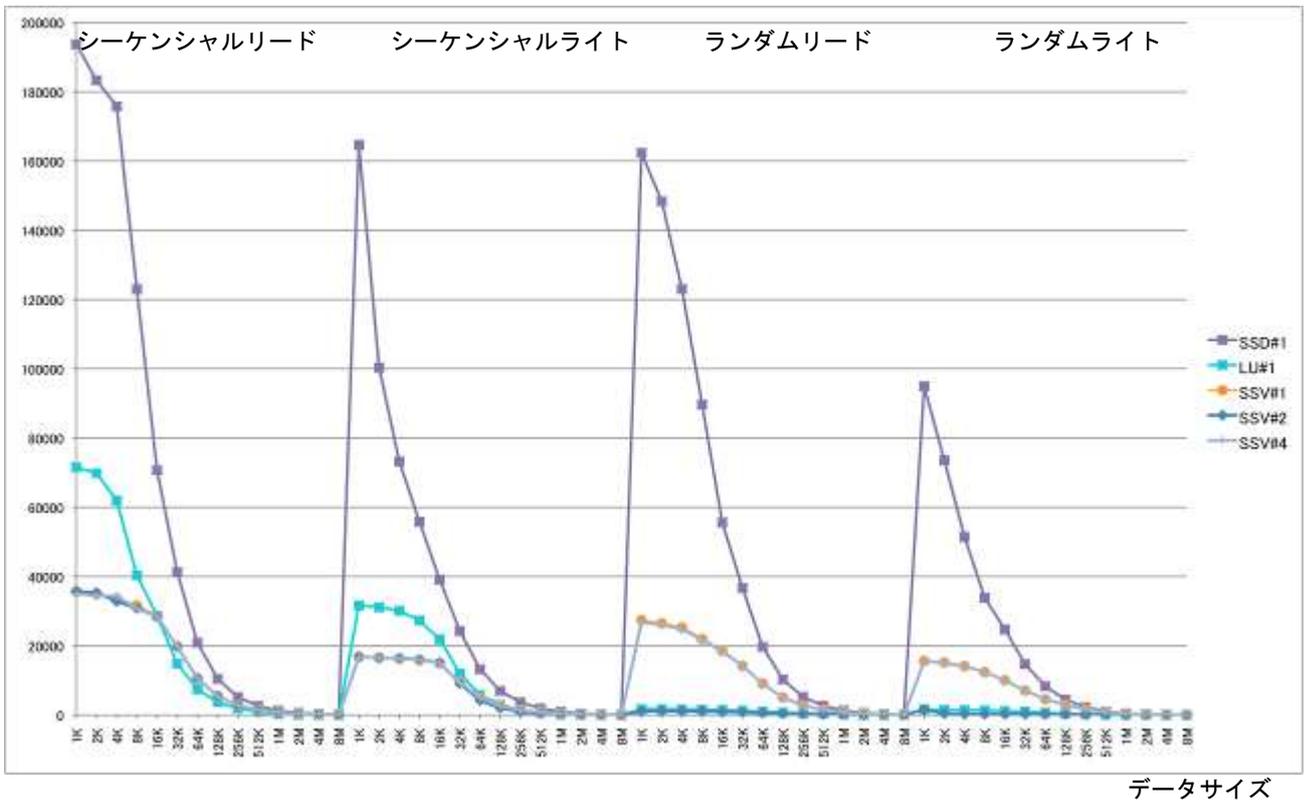
考察結果は、一部、4.1.1、2と重複している。

- シーケンシャルリードは、SSVのキャッシュ効果が得られていることがわかる。(Virident SSDの特性もキャッシュの効果に吸収されている)
- Virident SSDの効果は、シーケンシャルライト、ランダムリード/ライトをDX80S2のみで構成したLU#1とSSV#2を比較すると良く判る。  
シーケンシャルライトでは、LU#1で発生している512KBからのスループットの低下をカバーできているし、ランダムでのスループットの向上は明確である。
- Virident SSDとDX80S2を階層化したSSV#4とVirident SSDのみで構成されたSSV#1においてはスループットの差がほとんどみられないことから、階層化することで、高スループット/大容量というFlashSSDとディスクドライブの特性を活かしたハイブリッドストレージが構築できることがわかる。

#### 4.1.4 IOPS での比較 (その①)

IOPS で Virident SSD のみの SSD#1、DX80S2 のみの LU#1、そして、SSV で仮想化した SSV#1～#3 を比較すると、構成の違いによる IO 特性がより把握し易くなる。

IOPS

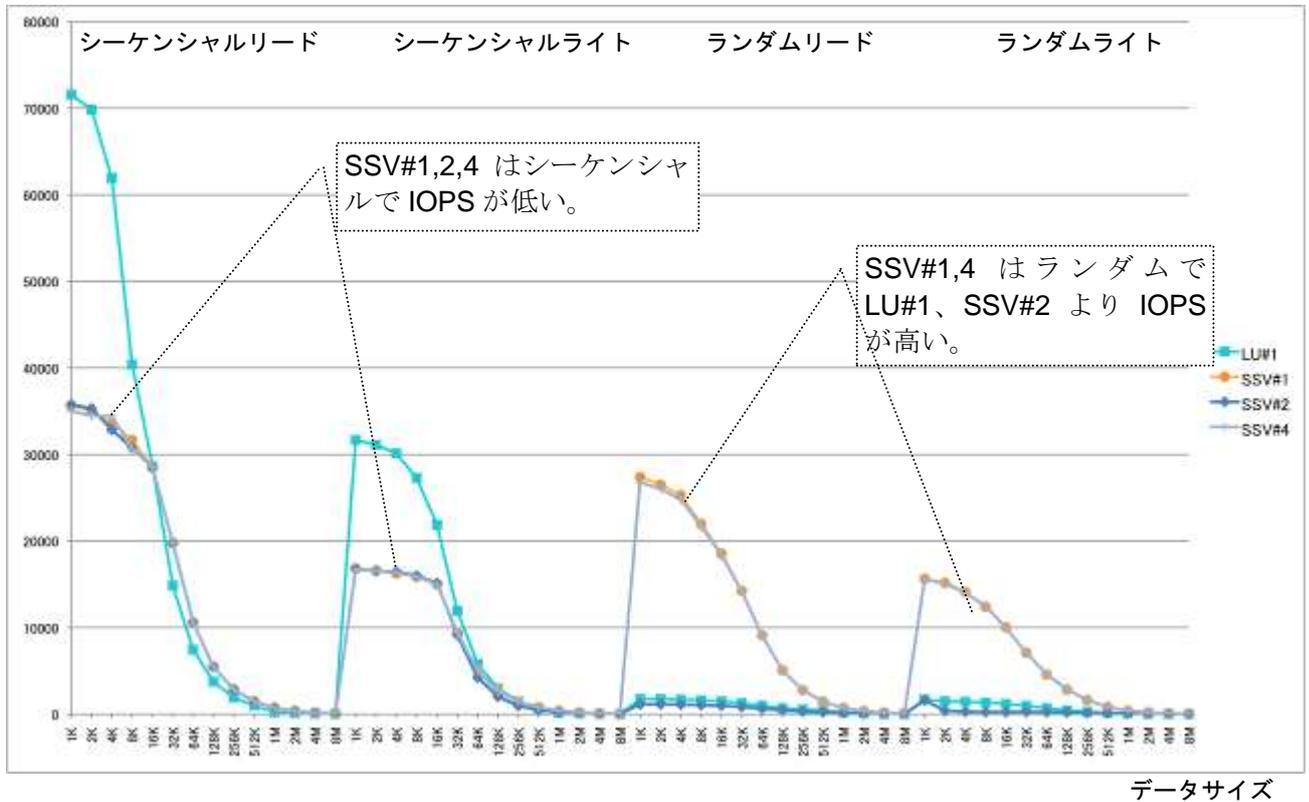


- Virident SSD を負荷サーバに直接接続した SSD#1 は、シーケンシャルリード 1KB で 19.4 万 IOPS が得られている。
- シーケンシャルリード/ライトでは、SSV#1、#2、#4 はいずれもほぼ同じ IOPS 値となった。
- ランダムリード/ライトでは、SSV#1 と #4 はほぼ同様の値である。
- シーケンシャルリード/ライトでは、SSV#1、#2、#4 は LU#1 よりも IOP 値が低く約 50% となった。
- ランダムリード/ライトでは Virident SSD とともに構成された SSV#1、#4 で、IOPS 値は DX80S2 のみで構成された LU#1、SSV#2 よりも最大 15 倍向上した。

#### 4.1.5 IOPS での比較 (その②)

よりパフォーマンスの特性を分かり易くするために、4.1.4 から負荷サーバ#1 に直結した SSD#1 を省いたものである。

#### IOPS



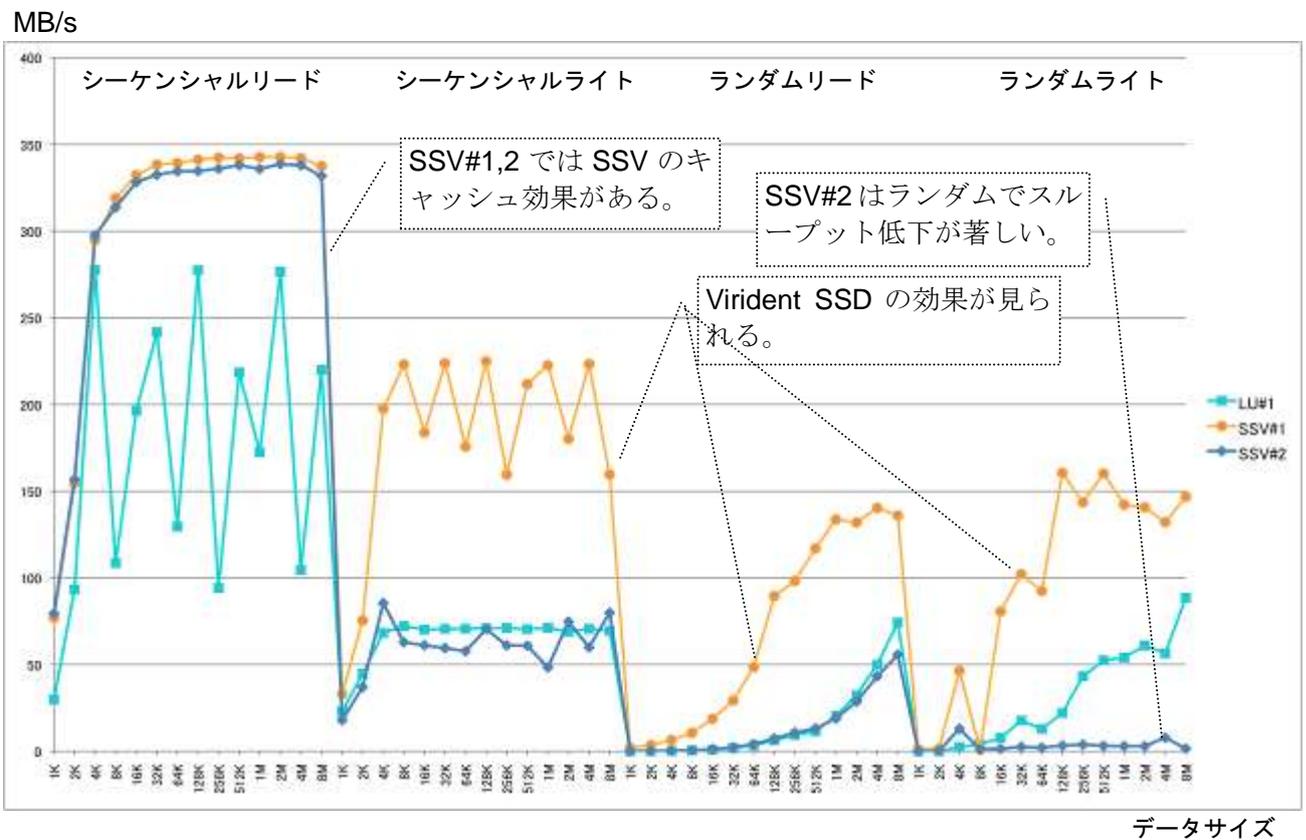
4.1.3 のスループット比較ではシーケンシャル性能は LU#1 よりも SSV#1、#2、#4 よりも良かったのだが、IOPS ではシーケンシャルリードでは 16K 未満、ライトでは 32KB 未満では LU#1 の半分程度となってしまった。

## 4.2 IOmeter for Linux

Linux での IOmeter は時間の都合で全ての構成でテストできなかった。

よって、取得できた DX80S2 を直接負荷サーバ#2 に接続した LU#1 および Virident SSD のみを SSV で同期ミラーリングして仮想化した SSV#1 と DX80S2 のみを仮想化した SSV#2 についてのデータで比較を行うものとする。

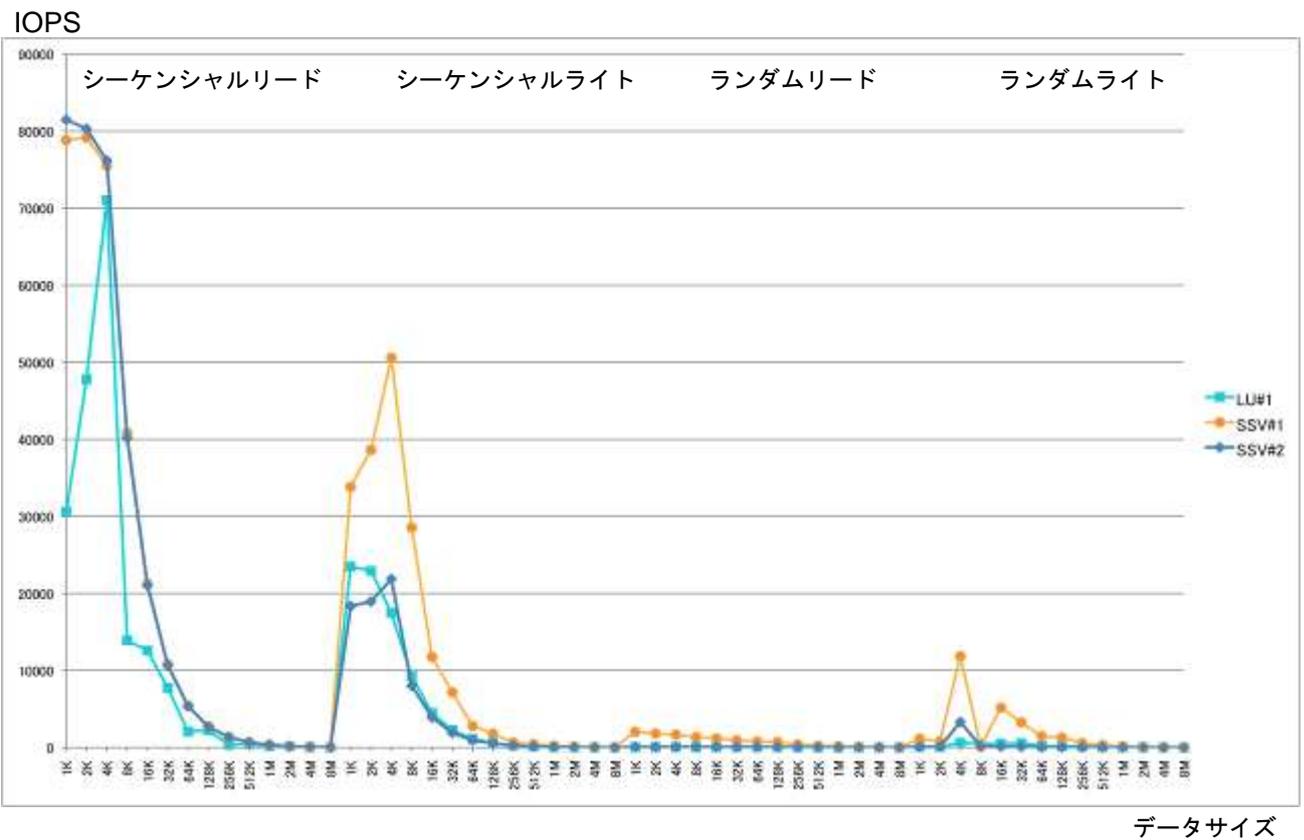
### 4.2.1 スループットの比較



- シーケンシャルリードでは、SSV#1 と#2 の結果から Windows マシンと同様に SSV のキャッシュ効果が得られていることがわかる。
- シーケンシャルライト、ランダムリード/ライトにおいても Virident SSD の効果が得られている。
- SSV で DX80S2 を同期ミラーリングして構成した SSV#2 については、ランダムライトでのスループット低下が著しく見られる※。

※2.3.2 の留意点にある通り、DX80S2 の構成（本来は 2 台別筐体をミラー）が影響している可能性が考えられる。

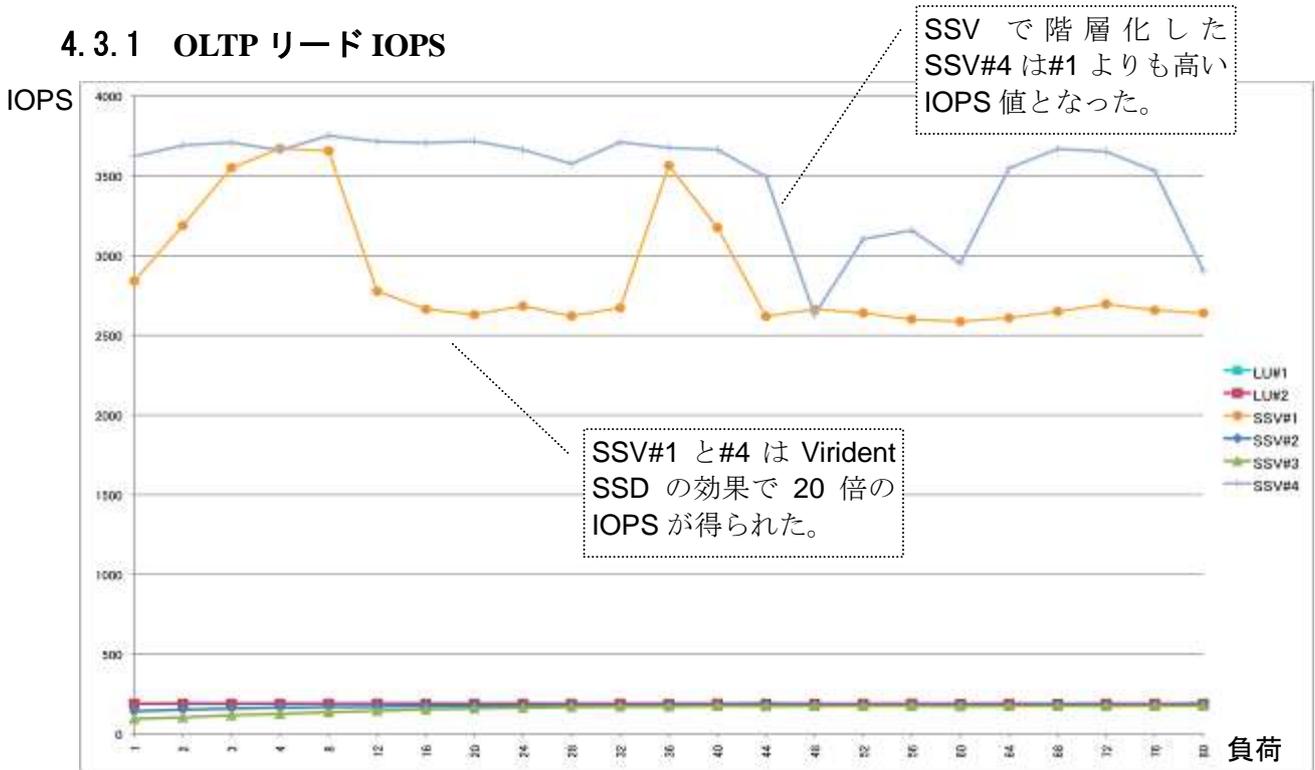
## 4.2.2 IOPS 比較



IOPS の比較でもスループットと同様の傾向が見られる。

### 4.3 Oracle ORION for Linux

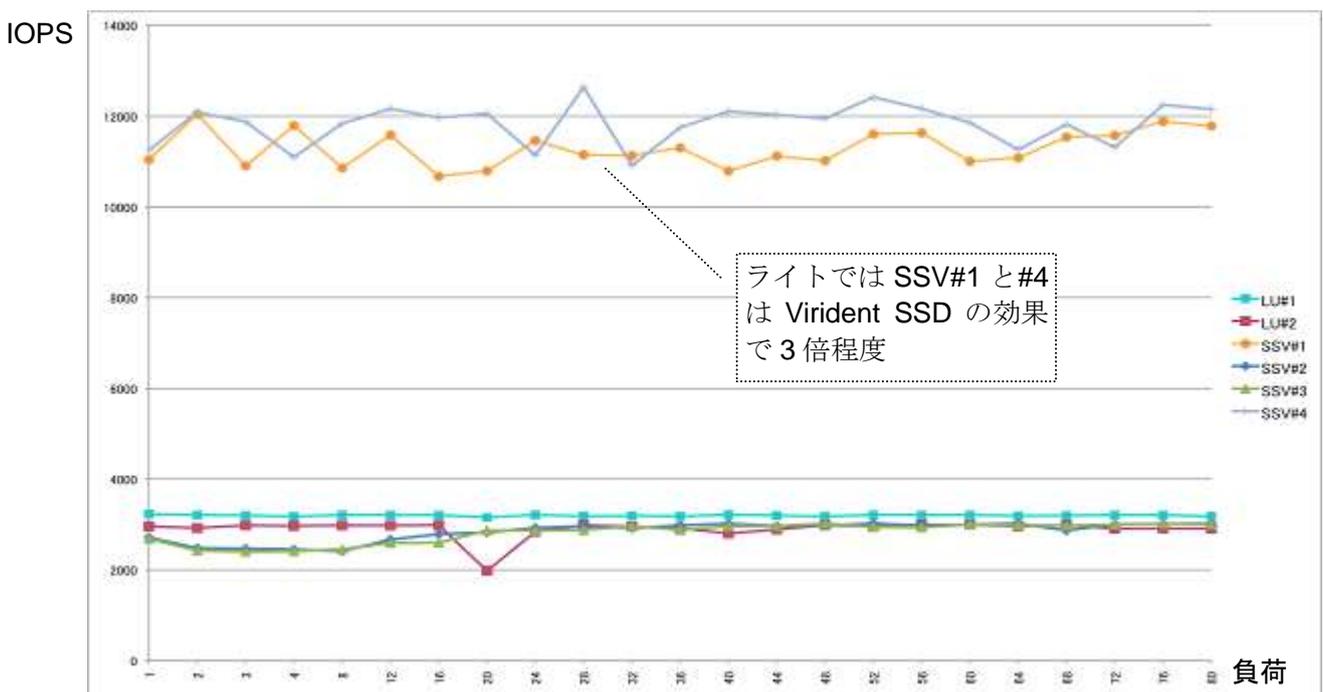
#### 4.3.1 OLTP リード IOPS



ORION で Oracle データベースにおけるデータサイズ 8KB の OLTP をシミュレーションすると Virident SSD は DX80S2 に比べて 20 倍近く IOPS が向上した。

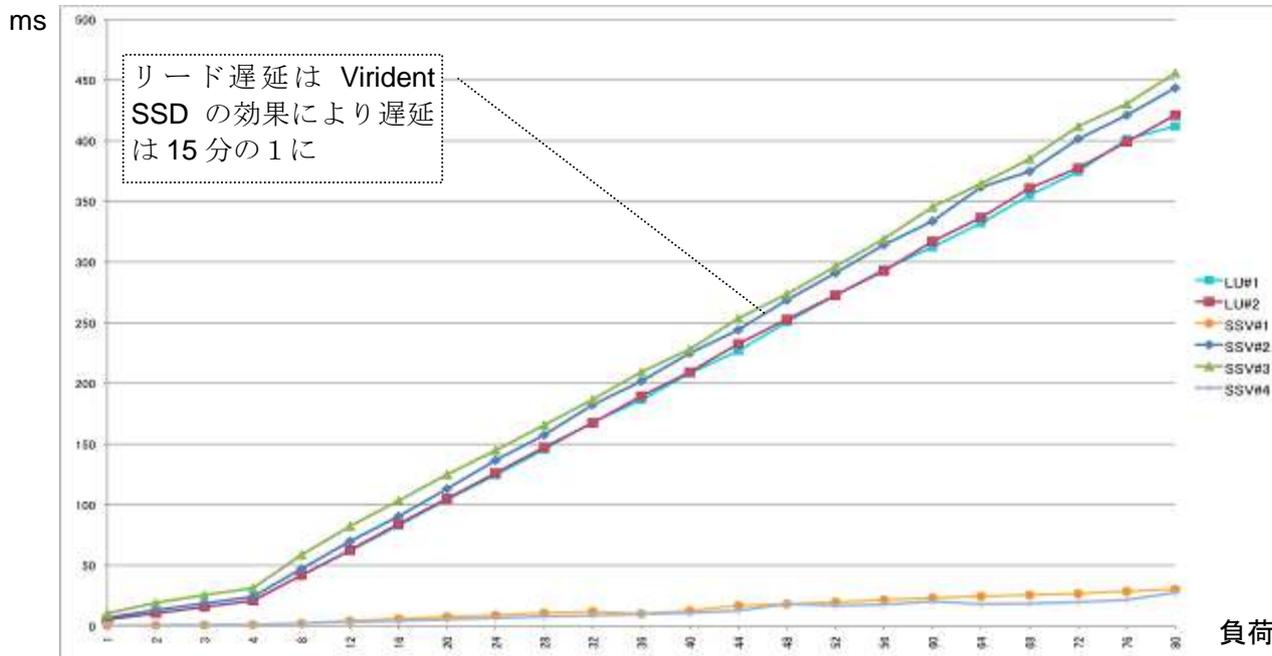
また、Virident SSD と DX80S2 とを階層化した SSV#4 はより高い IOPS 値が得られた。

#### 4.3.2 OLTP ライト IOPS



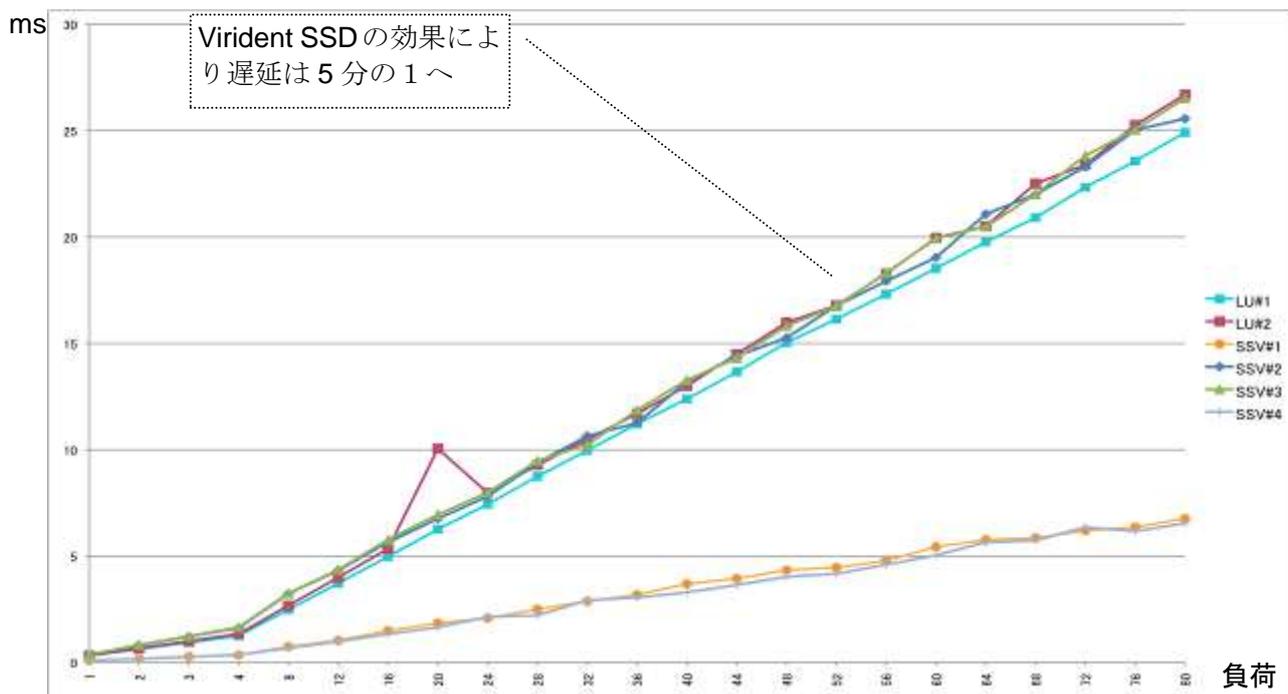
OLTP のライトは 4.3.2 のリードと比較して全体的に IOPS 値が高く、DX80S2 で構成されている LU#1、#2、SSV#2、#3 と Virident SSD で構成されている SSV#1、#4 と比較すると IOPS 値は約 3 倍程度の向上に留まっている。階層化の効果はリード時よりは見られない。

### 4.3.3 OLTP リード遅延



次に IO 遅延をみると、Virident SSD で構成される SSV#1 と #4 負荷の増加に伴っても IO 遅延がほとんど発生しないことが分かる。

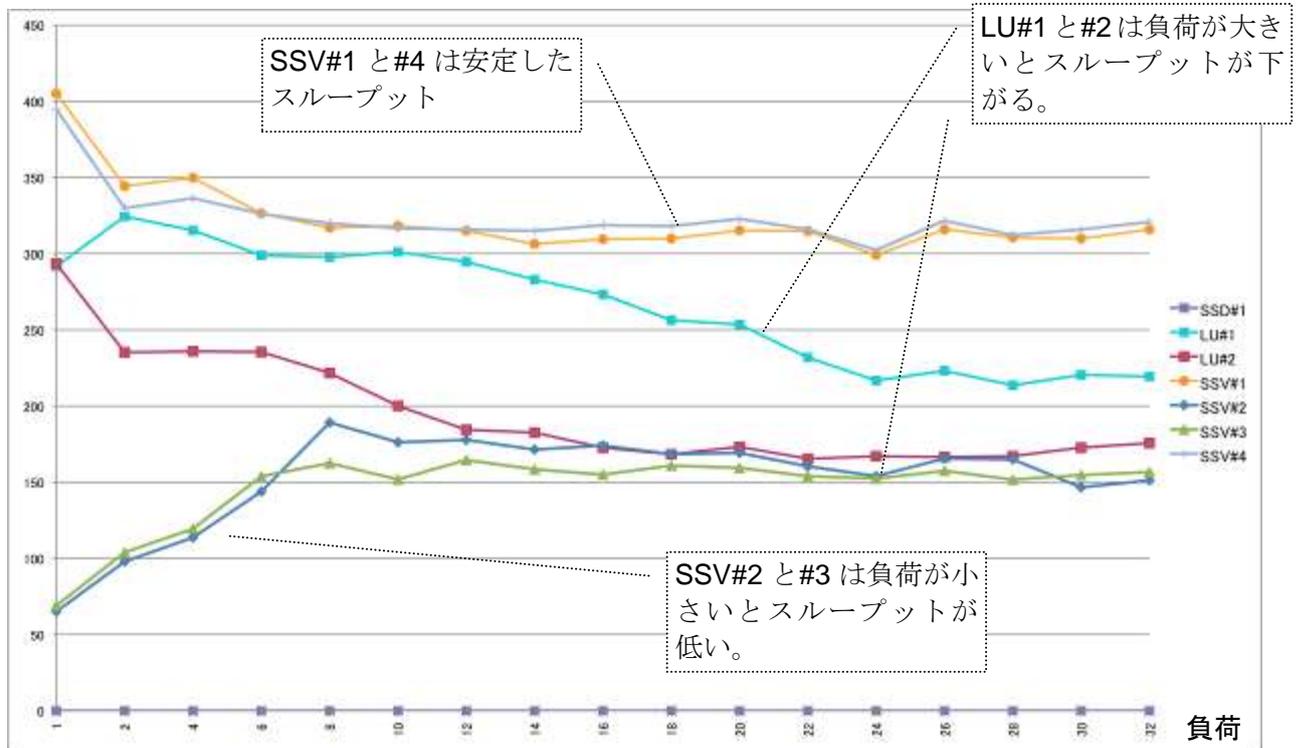
### 4.3.4 OLTP ライト遅延



IOPSで差が出なかったのと同様に5分の1程度では有るが、Virident SSDで構成されるSSV#1、#4の方が負荷に伴ってIO遅延が発生し難いことが分かる。

#### 4.3.5 DWH リードスループット

MB/s



- OrionでOracleデータベースにおけるデータサイズ1024KBのDWHをシミュレーションすると、Virident SSDで構成されるSSV#1,#4が安定した速度をみせる
- DX80S2で構成されるSSV#2と#3については、負荷が小さいうちはスループットが低いことがわかる。
- 負荷サーバ#2に直結したDX80S2の性能に注目するとRAID0+1(LU#1)の方がRAID5(LU#2)よりスループットが高い。また、いずれも、負荷が増えるにつれスループットは低下する傾向にある。

### 4.3.6 DWH ライトスループット

一部データの取得に失敗してしまったため、参考データとなるが、Virident SSD を使った SSV#4 がライトにおいてもスループットが負荷サーバ#2 に直結した LU#1、LU#2 に比べて 2.5 倍のスループットが安定して得られていることが分かる。

MB/s



## 5 まとめ

---

### 5.1 IOmeter for Windows

#### 5.1.1 スループットでの比較

##### SSV のキャッシュ効果

SSV ではサーバに搭載しているメモリをキャッシュとして扱うことができる。これは、2.3.3 項で説明されている通り RAID 装置に搭載されているキャッシュと同じ役割を果たすものであるが、IA サーバ用の低価格なメモリが使えるというのが SSV の大きな利点である。

IOmeter for Windows での測定結果では、特にシーケンシャルリードのスループット改善に大きく寄与することがわかった。

##### Virident SSD の効果

4.1.2 項を見れば分かるように Virident SSD の特性がそのまま SSV にも反映されており、特にランダム IO の性能はディスクドライブと比べるとその効果は歴然である。

スループットの差は PCI-Express と 8GbFC の帯域の違いによるものなので、例えば 8GbFC を 2 本束ねることで、帯域の差を吸収することができる可能性がある。

##### ハイブリッドストレージの効果

今回は全てのデータが Virident SSD 上にあったので、DX80S2 と組み合わせた場合の明確な違いがなかった。

SSV で仮想化すればキャッシュ効果でシーケンシャルリードのスループットは大幅に改善されるが、さらに、SSV の階層化機能でアクセス頻度が高いブロックが Virident SSD に配置されることにより、ランダム IO においても高スループットが得られるハイブリッドストレージが構築できることが期待される。

この SSV の階層化機能によるデータの再配置の動きは、今後、追及して行きたい。

#### 5.1.2 IOPS での比較

IOPS については、Virident SSD を使うことでランダム IO では、DX80S2 だけのときよりも相対的に IOPS の改善が見られたものの、Virident SSD の IOPS を絶対値からみると最大では 20%以下に留まっている。

ただし、SSVには物理ディスクの分割単位（SAU：ストレージアロケーションユニット）が4MB～2048MBの範囲での設定や、ライトバックキャッシュをライトスルーモードにすることが可能なので、IOPS向上を目的としたチューニングの余地は残されている。

## 5.2 IOmeter for Linux

IOmeter for Windows のときと同様に、シーケンシャルリードでのキャッシュ効果、Virident SSD の効果がそれぞれ見られた。

## 5.3 Oracle ORION for Linux

### 5.3.1 OLTP の IOPS での比較

OLTP リードについては、IOmeter でランダムリード 8KB のときをみると、SSV#1 が LU#1 の 21 倍だったので、その比率が ORION でも同様であった。

なお、OLTP リードではハイブリッドストレージのときの差が他のベンチマークテストより明確に出ていることを特記事項としてあげておく。

OLTP ライトについては、IOmeter でランダムライト 8KB のときをみると、SSV#1 の方が IOPS 値は低く ORION と同様の比率にならなかった。ただし、IOmeter の値は 8KB のところのみ SSV#1、#2 とともに一時的に下がっておりベンチマーク固有の問題の可能性がある。

### 5.3.2 OLTP の IO 遅延での比較

エンドユーザがシステムにアクセスをするときにこの IO 遅延は非常に重要である。

見かけ上の転送速度や IOPS は同時 IO 数（IOmeter では Outstanding IO）を大きくすれば増やすことができるからである。

しかし、実際のサービスの観点から考えれば、同時アクセスのユーザが増加したとしても、それにつれて伴って応答時間が長くないことが望ましい。

OLTP の IO 遅延の結果を見ると、Virident SSD を使うことで常に安定した応答が保証されるストレージが構築できることが分かる。

### 5.3.3 DWH のスループットでの比較

DWH では大量のデータをできるだけ短い時間で処理することが求められる。

こちらも Virident SSD を SSV で使うことで、DWH で求められるストレージが構築できることが分かる。

## 5.4 検証結果を振り返って

今回、タイトルに上げたハイブリッドストレージだったが、実際にはデータが Tier1 の Virident SSD にある状態にて測定のみとなってしまった。データが複数階層に跨がった場合にどのような挙動になるかが把握できなかつたので、これは、次回の検証課題としたい。

しかしながら、SSV を導入する効果については、全体を通して、シーケンシャルリード時のキャッシュ効果が得られていることが分かった。

それ以外の IO については、デフォルトのパラメータを使用していることや、2 台の外部ストレージを使用しなかつたことにより、SSV の機能をフルに引き出してはいないが、それでも、今回の結果からみても Virident SSD を組み合わせれば、ほとんどのケースでこの点がカバーされていることは明確であり、更に、チューニングをすることでの向上が期待できる。

そして、何より SSV を用いることで単体のストレージでは解決できなかった問題が、PRIMERGY/ETERNUS/Virident SSD を組合せることで一挙に解決し、柔軟なストレージシステムが構築できることが、本検証より確認できたと結論づけられた。

## 5.5 本検証に関するお問い合わせ先

本検証に関して、ご興味を持たれた方、あるいは、技術的な質問に関しては下記にお問い合わせ下さい。

ネクスト・イット株式会社

Web : <http://nextit.jp> 電話 : 03-5783-0702 e-mail : [info@nextit.jp](mailto:info@nextit.jp)

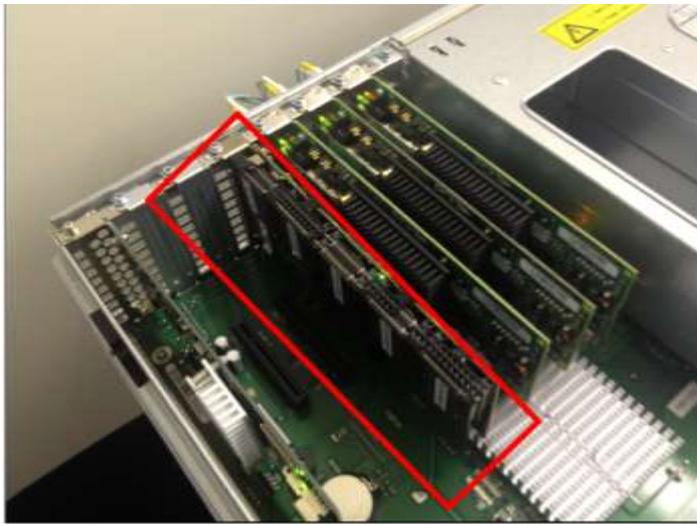
## 付録

### 付録 1 Virident Flash MAX の仕様

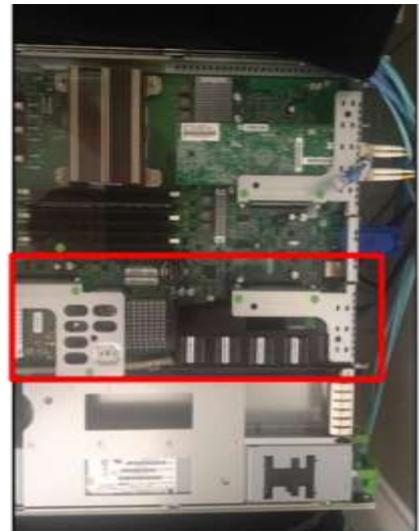
FlashMAXSCMSpecifications		
	MLC	SLC
Capacity (GB)	1000, 1400	300, 400, 800
Form Factor	Low Profile	Low Profile
Read Performance	1.3 GB/s (4 KB) 325 Thousand IOPS (4 KB) 1 Million IOPS (512 byte)	1.4 GB/s (4 KB) 340 Thousand IOPS (4 KB) 1.4 Million IOPS (512 byte)
Write Performance	600 MB/s	1100 MB/s
Sustained Mixed Performance (75/25 r/w) at full capacity	850 MB/s (4 KB) 220 Thousand IOPS (4 KB)	920 MB/s 235 Thousand IOPS (4 KB)
Write Access Latency	19 $\mu$ s	16 $\mu$ s
Read Access Latency	62 $\mu$ s	47 $\mu$ s

Standards and Connectivity	
PCI Express 1.0/2.0	Single slot, low profile, half height, half length
Connectivity	PCI Express electromechanical spec 1.0/2.0x8
Application Interface	Standard block device
Platform Support	Linux: RHEL 6/6.1, SLES 10/11, CentOS 5/6, Oracle EL 6, Debian 4/5/6, Ubuntu 8/9/10/11, Fedora Core Windows: 64-Bit Microsoft Server 2008 R2 SP1 (SLC)

## 付録2 Virident Flash MAX を搭載したときのようす



富士通RX300S6

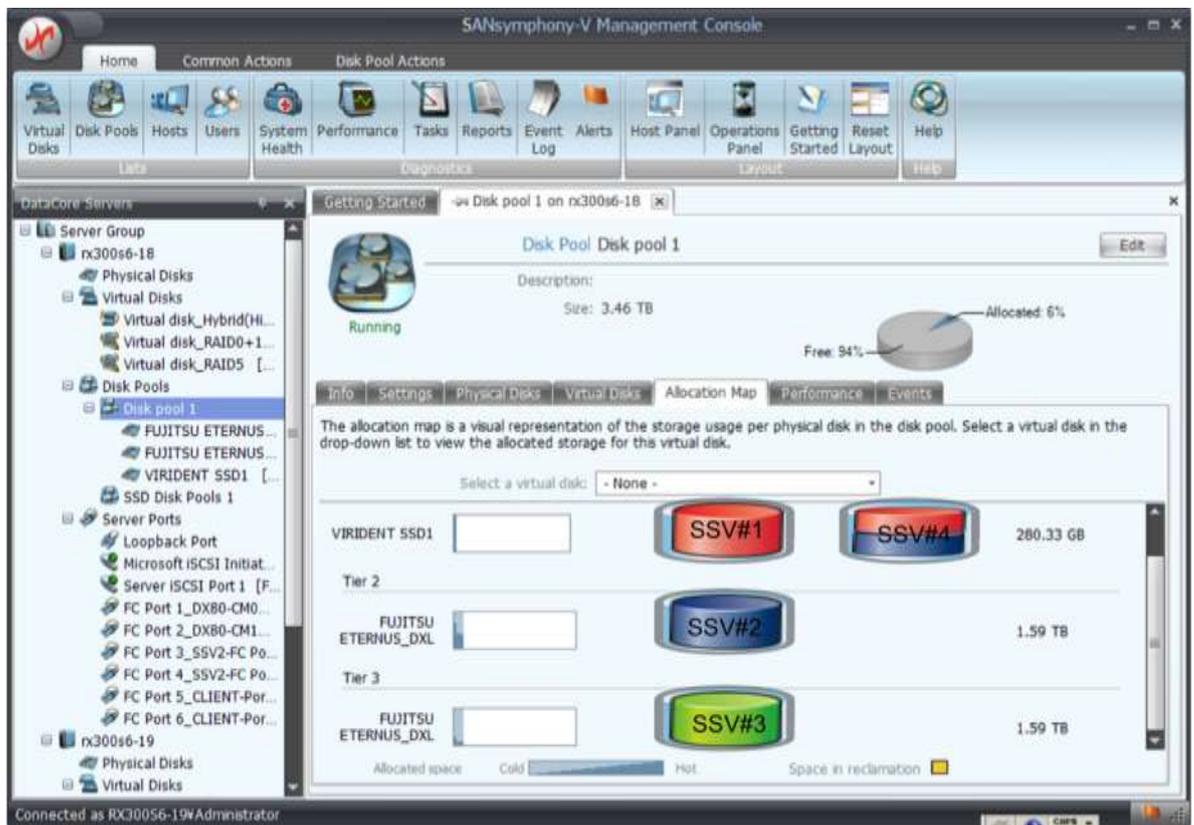


富士通RX200S6

富士通の RX300S6 は 2U サーバでありながら PCI-Express のスロットが 6 個あり、8GbFC×6 ポート+VIRIDENT SSD という構成が組める。これは、FC-HBA、NIC、SSD といった複数のカードを搭載することがある SSV のサーバとして非常に使い易い製品である。

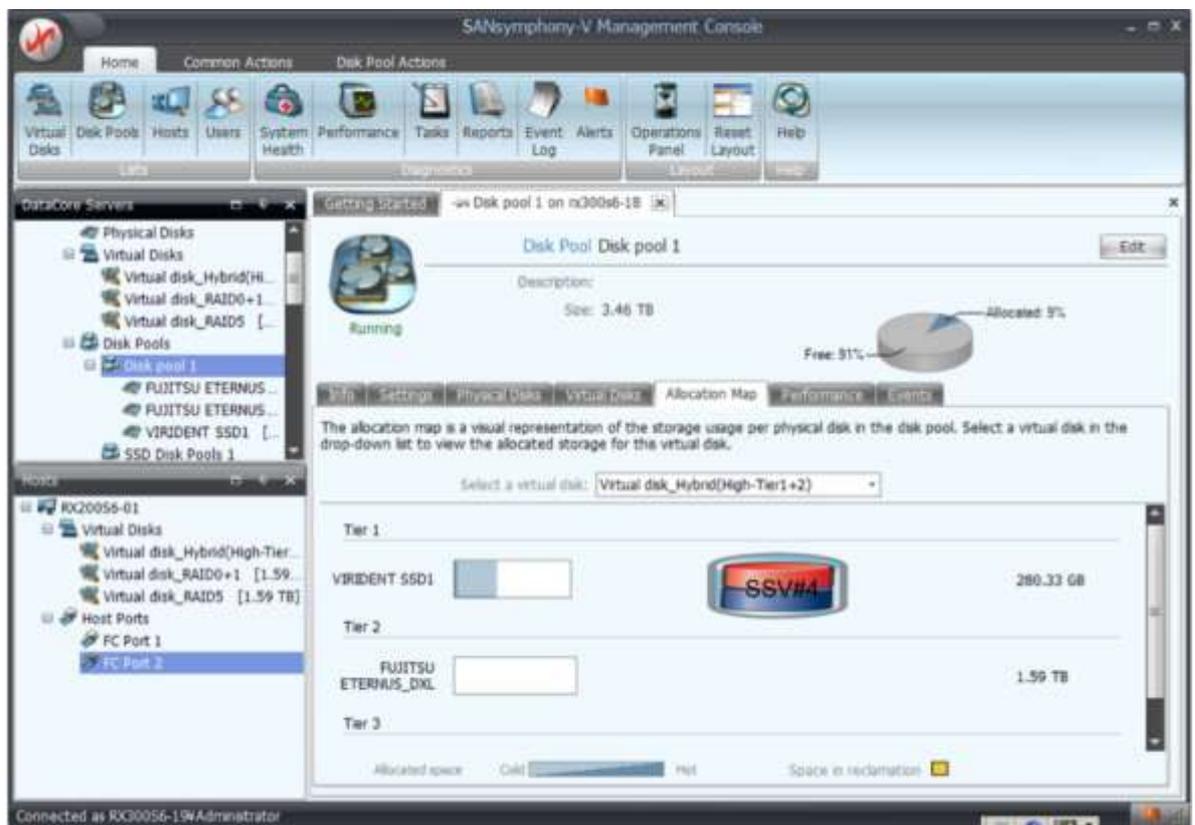
## 付録3 SSV におけるデータ配置のようす

### 付録3-1 ストレージの階層構造



Tier1 が Virident SSD、Tier2 が DX80S2 の RAID1+0、Tier3 が RAID5 の構成である。  
SSV#4 は Tier1 と Tier2 からストレージリソースを配置するようにしている。

### 付録 3-2 ハイブリッドストレージの構成例



SSV#4 に 100GB のデータを書き込んだ結果である。Virident SSD の容量 300GB を全て使い切っていないので、Tier2 には全くデータブロックがないことが分かる。