

IBM FlashSystem 840 と  
富士通 SPARC M10-4/Oracle Database 12c  
における接続検証および性能検証  
検証報告書

2015 年 3 月

日本アイ・ビー・エム株式会社  
株式会社アルゴグラフィックス  
株式会社富士通北陸システムズ  
株式会社富士通エフサス  
ノアキス株式会社

## はじめに

最近、高性能ストレージとして、従来の磁気ディスクに代わってフラッシュストレージが使われ始めています。

そこで、今回は、アクセス時間 100 マイクロ秒という低レイテンシを特長とする「IBM FlashSystem 840」を SPARC M10 に接続し、その導入作業、Oracle Database に対する性能効果などを検証しました。

検証目的は、SPARC M10/Solaris11/Oracle12c という環境に対して、IBM FlashSystem 840 を導入する作業および、その上で Oracle DB の各種操作が問題なく実行できることを確認するものです。

併せて、今回環境準備ができる最少構成、小規模データにおける性能を確認しましたが、今後、IBM FlashSystem の性能をフルに発揮できる環境にて性能検証をしたいと思います。

## 目次

1. 検証の実施概要
2. IBM FlashSystem 840 導入作業
3. 性能評価
4. まとめ
- 付録 1. IBM FlashSystem 840 概要
- 付録 2. お問い合わせ窓口
- 付録 3. 参考情報リンク

## 1. 検証の実施概要

(1) 実施日 2014 年 11 月 4 日～17 日

(2) 実施場所 富士通検証センター（東京・浜松町）

(3) 参加メンバー

日本アイ・ビー・エム株式会社  
株式会社アルゴグラフィックス  
株式会社富士通北陸システムズ  
株式会社富士通エフサス  
ノアキス株式会社

(4) 検証構成

図 1-1 に、検証に使用した機器構成を示す。

構成図では検証結果に関わる機器のみを表しており、管理用端末やネットワークなどは図に含めていない。

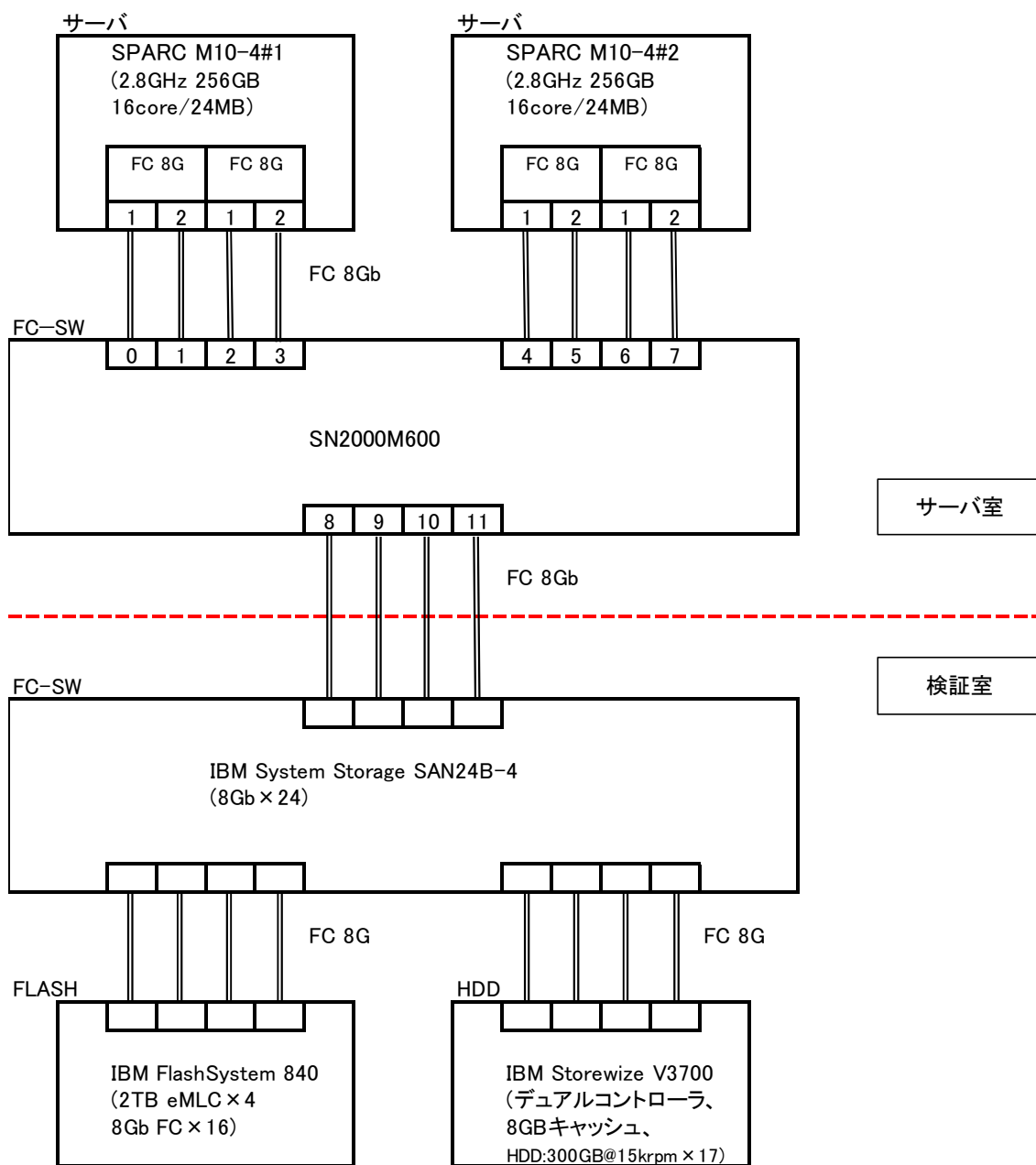


図 1-1 検証用機器構成図

## 2. IBM FlashSystem 840 導入作業

### (1) 設定用 GUI へのアクセス

FlashSystem に IP アドレスなどのネットワーク設定を行い、設定用 GUI へアクセスする。

### (2) ボリューム (LUN) 作成

ボリューム作成画面 (図 2-1) 左側の「ボリューム」アイコンをクリックし、「ボリュームの作成」を押してから名前、容量、ボリュームの数を入力してボリュームを作成する。



図 2-1 ボリューム作成画面

### (3) ホスト定義の作成

ホスト作成画面 (図 2-2) 左側の「ホスト」アイコンをクリックし、「ホストの作成」≫「ポートの追加」でホストを作成する。

このとき、以下の 2 点を留意する必要がある。

- ホストと FlashSystem の接続に FC スイッチを使用する場合、FC スイッチのゾーニング設定やその他の設定を行う。(FC スイッチの設定方法は製品のマニュアルを参照すること。)
- ホスト OS の FC カードドライバやマルチパスの設定を実施する。(設定内容・方法については、ホスト OS によって異なるため、ホスト OS のマニュアルを参照すること。)

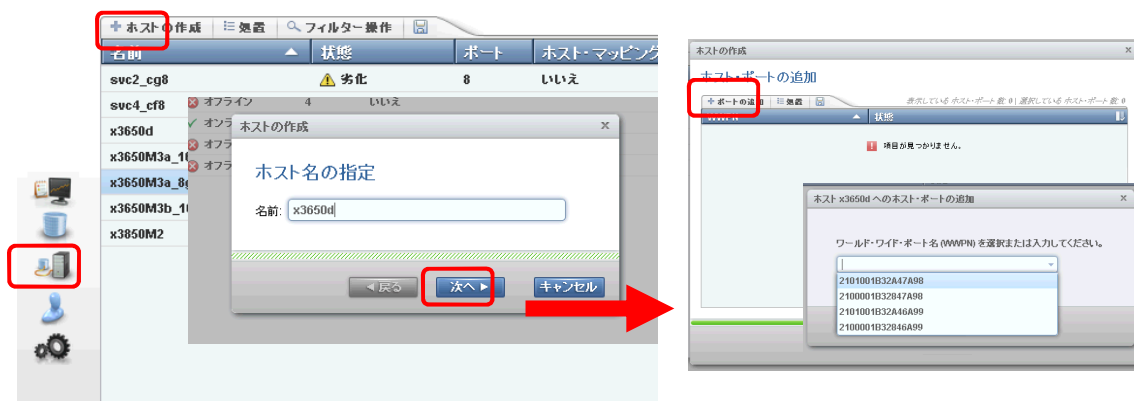


図 2-2 ホスト作成画面

#### (4) ボリューム (LUN) のマッピング設定

作成したボリュームを、アクセス許可するホストにマッピングする。

図 2-3 のように、「ホスト」アイコンをクリックし、「非マップ・ボリューム」から、ボリュームを選択してホストへマップする。

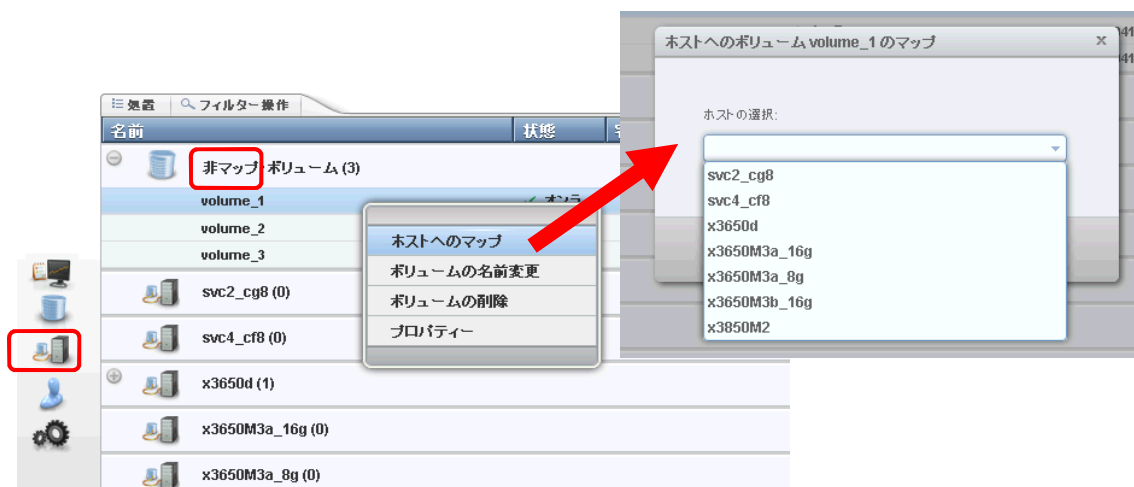


図 2-3 ボリュームのホストへのマッピング

### 3. 性能評価

#### 3.1. 検証環境

図 3-1 に、今回の検証の論理環境を示す。

DB サーバとして、2 ノードの RAC、それに HDD および FLASH ストレージを接続し、各々に検証用 DB を作成し、性能比較を行った。

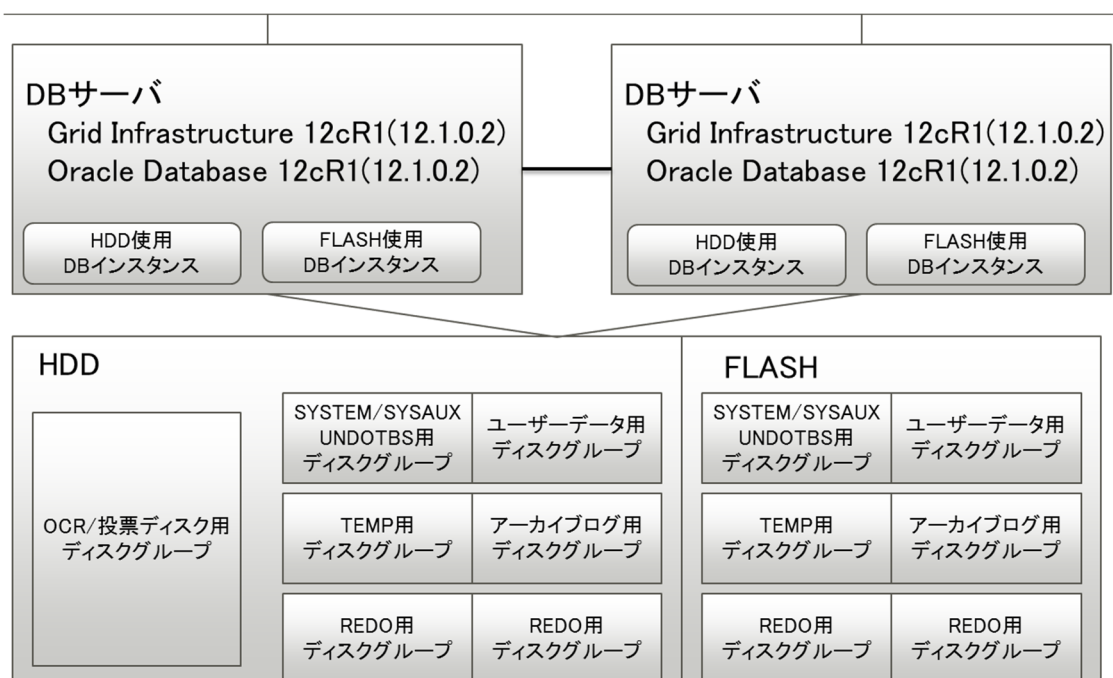


図 3-1 検証環境

#### 3.2. 検証項目

表 3-1 検証項目

| 検証項目          | 検証内容                                                        |
|---------------|-------------------------------------------------------------|
| 基本性能          | JPetStore スキーマ内の特定のテーブルを使用<br>SELECT/UPDATE/DELETE（1000 万件） |
| OLTP 性能       | JPetStore モデルの OLTP（800 スレッド）<br>INSERT+索引検索+UPDATE 処理      |
| バッチ性能         | 複数テーブルを結合した集計処理（全件検索）<br>多重度：1, 2, 4                        |
| RMAN バックアップ   | 多重度：1, 2, 4, 8                                              |
| Export/Import | JPetStore スキーマ内の特定のテーブルを使用（データ量：35GB）                       |

### 3.3. 検証結果

#### (1) 基本性能

DB 処理では基本の SELECT 性能(実行時間)を単純に比較した。

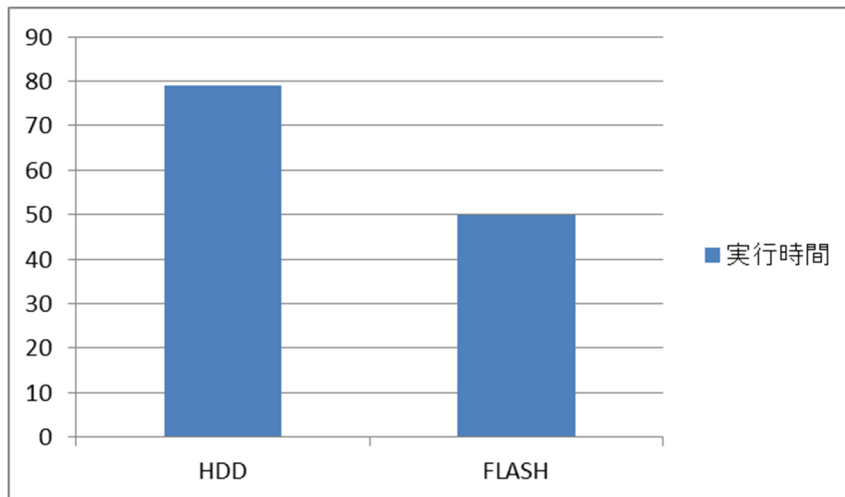


図 3-2 SELECT 性能比較

- HDD に対して、FLASH を使った場合には実行時間が 35%以上短縮できた。
- SELECT 処理における Read I/O のレスポンス時間の差が現れている。

#### (2) OLTP 性能

JPetStore で OLTP を想定した負荷を与え、各トランザクションのレスポンス時間、処理スループットなどを比較した。

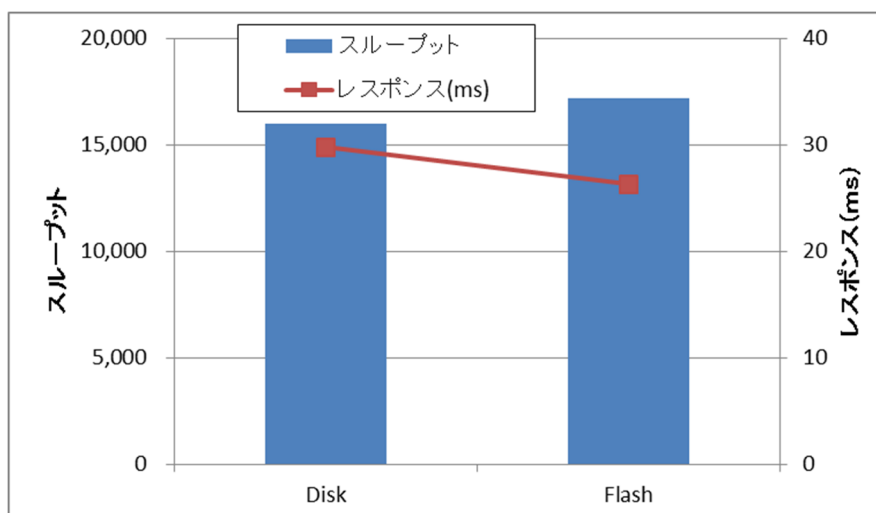


図 3-3 OLTP 性能比較

- FLASH と HDD では、レスポンスで 10%、スループットで 7%程度 FLASH の方が良いが、ストレージ装置自体の性能差ほどの違いは現れなかった。
- JPetStore ベンチマークの OLTP では、処理のほとんどがキャッシュ上で動作したため、FLASH と HDD の差が顕在化しなかったものと思われる。
- 図 3-4 のベンチマーク中のディスクビジー率推移を見ても、HDD では一瞬 100%になるものの、概ね 10%以下で推移している。FLASH ではピークでも 40%程度で、あとは数%以下となっている。

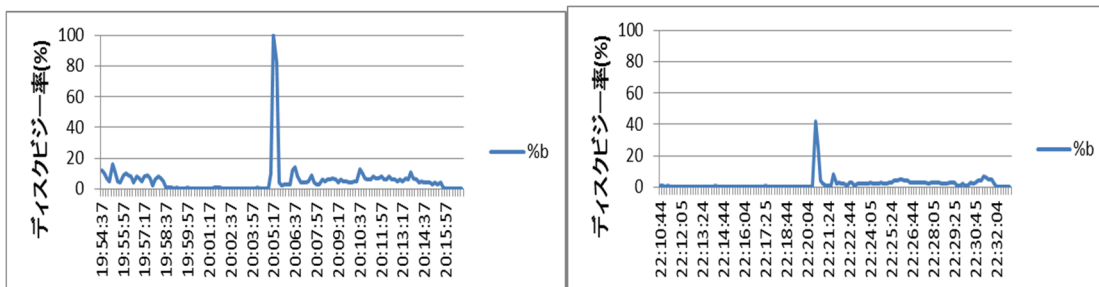


図 3-4(a) HDD ディスクビジー率

図 3-4(b) FLASH ディスクビジー率

### (3) バッチ性能

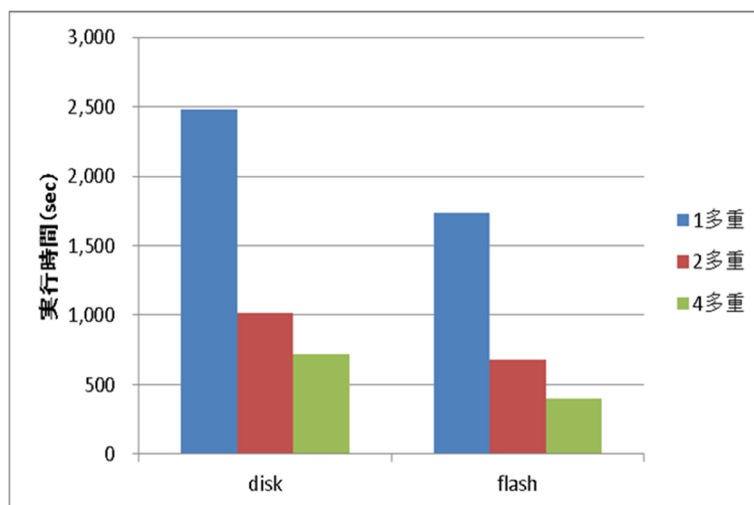


図 3-5 バッチ性能比較

- バッチ性能は実行時間で評価するが、多重度が大きくなるほど改善が大きい。(1 多重△32%、2 多重△40%、4 多重△45%)
- このバッチ処理では Read I/O の負荷が高いため、FLASH と HDD の性能差より大きく結果に現れた。
- 図 3-6 のように、ベンチマーク中の Read 性能は、FLASH が 1000MB/s に対して、HDD が 500MB/s と 2 倍の性能差があった。

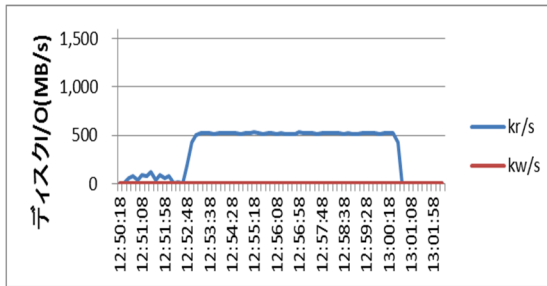


図 3-6(a) HDD データ転送速度

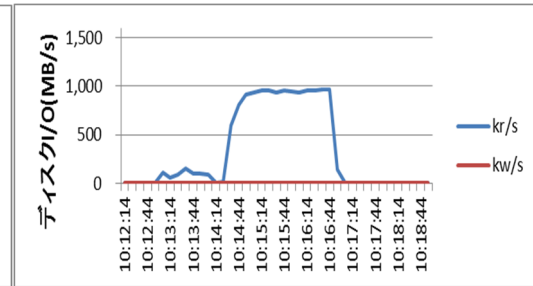


図 3-6(b) FLASH データ転送速度

#### (4) RMAN バックアップ

RMAN によるバックアップでは、HDD の場合は多重で流すとヘッドの動きが干渉して、シングルの場合よりデータ転送速度が低下するが、FLASH では逆に並列転送の効果により全体のデータ転送速度が向上する。8 多重の比較では、FLASH の方が 57%データ転送速度が大きい。(図 3-7)

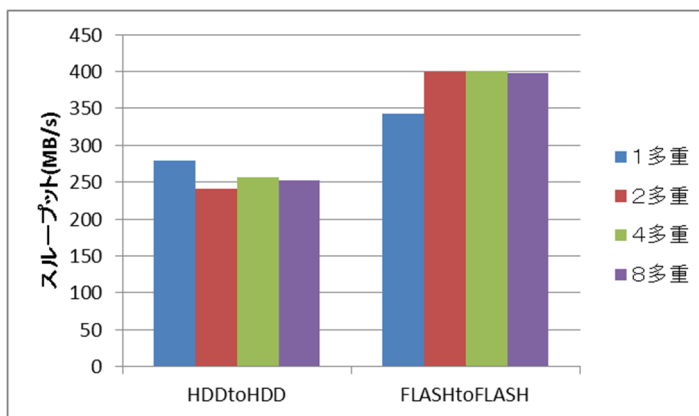


図 3-7 RMAN バックアップ性能比較

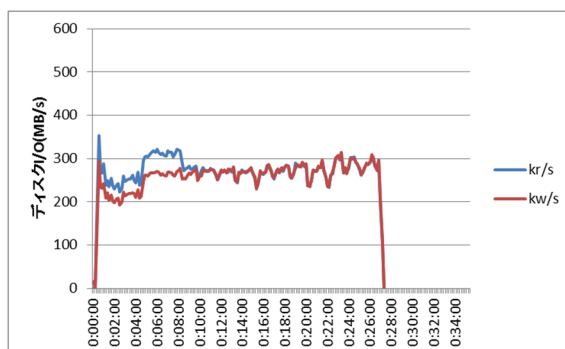


図 3-8(a) RMAN バックアップ時のデータ転送速度推移 (HDD)



図 3-8(b) RMAN バックアップ時のデータ転送速度推移(FLASH)

### (5) EXPORT/IMPORT 性能

Oracle の EXPORT/IMPORT を、HDD と FLASH の各々で行い、性能として実行時間を比較した。なお、データは HDD 同士、FLASH 同士で転送している。

(単位：秒)

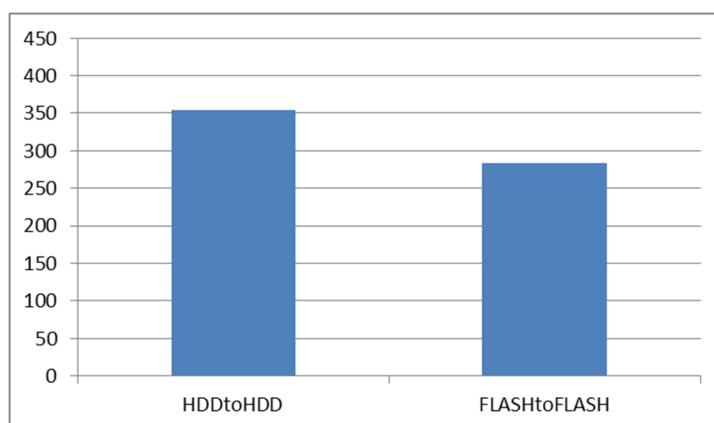


図 3-9(a) EXPORT 性能比較

(単位：秒)



図 3-9(b) IMPORT 性能比較

- EXPORT では FLASH による I/O 性能の効果があり、20%の実行時間短縮となった。
- IMPORT では、書込み側のキャッシュ効果もあり、HDD と FLASH では差がなかった。
- 今回の測定は 1 多重であり、HDD と FLASH の差はあまりなかったが、多重で実行した場合には、HDD の特性として性能が劣化するので、FLASH との差が大きくなると考える。

## 4. まとめ

### (1) セットアップ作業

今回の検証作業では、先ず SPARC M10-4/Oracle 12c に IBM FlashSystem 840 を接続し、セットアップ作業した。FlashSystem 840 では設定用 GUI が用意されており、簡単な設定作業でボリューム作成まで行うことができた。

### (2) 富士通 SPARC M10-4/Oracle Database 12c との接続

IBM FlashSystem 840 を接続した SPARC M10-4/Solaris11.1 に、Oracle 12c をインストールし、動作確認した。特別な追加作業をすることなく、正常に動作することが確認できた。

### (3) Oracle に対する性能

Oracle12c に対して、FlashSystem がどのような振る舞いとなり、限定された条件ではあるが、どのような効果があるのかを性能・移行性などの観点で検証した。

- バッチ処理では、多量のデータをリード/ライトするために大量の I/O 処理が発生するため、FLASH の性能効果が大きく、処理時間が大幅に短縮された。
- OLTP では、小さな I/O が多く発生するが、ほとんどの処理がキャッシュ上で行われるため、FLASH の効果は小さい。
- RMAN バックアップは、FLASH の性能効果が大きかった。
- バッチ処理、RMAN バックアップ、IMPORT/EXPORT などにおいては、多重処理になると、HDD ではヘッドの動きに相互干渉が起こり性能劣化するが、FLASH はその影響が全く無く、性能効果が大きい。

### (4) 移行時の効果

以下のように、FLASH の採用により、移行時のダウンタイムを短縮させることができる。

- RMAN バックアップによる移行では、バックアップセットの格納場所を FLASH にすることで、移行処理時間を短縮させることができる。
- DB 領域を FLASH にすることで、アップグレード・スクリプトの実行時間を短縮させることができる。
- ダンプファイル格納領域を FLASH にすることで、IMPORT の処理時間を短縮させることができる。

(5) 性能に関する考察

IBM FlashSystem 840 は、フラッシュモジュール単位での並列転送が可能  
なため、構成によって性能が変化する。(表 4-1)

今回は、FlashSystem 840 の最小構成 (4 modules) であったため、もう  
少し大きな構成では、特にバックアップなどの性能は、より良い性能値にな  
ったものとする。

表 4-1 FlashSystem 840 の構成と性能

|                          |            | 最大構成<br>(12 modules) | 中間構成<br>(8 modules) | 最小構成<br>(4 modules) |
|--------------------------|------------|----------------------|---------------------|---------------------|
| IOPS                     | 100% read  | 1,100,000            | 1,100,000           | 1,000,000           |
|                          | 100% write | 600,000              | 400,000             | 225,000             |
|                          | R:W=7:3    | 750,000              | 500,000             | 225,000             |
| Large<br>block の<br>連続転送 | read       | 8GB/s                | 8GB/s               | 4GB/s               |
|                          | write      | 4GB/s                | 2.5GB/s             | 1GB/s               |
| レイテン<br>シー               | read       | 135 $\mu$ s          | 135 $\mu$ s         | 135 $\mu$ s         |
|                          | write      | 90 $\mu$ s           | 90 $\mu$ s          | 90 $\mu$ s          |

注1) Oakgate storage test appliance と FC protocol analyzer による測定値

注2) 4TB モジュールを使用。2D RAID 構成。

## 付録1. IBM FlashSystem 840 概要

### (1) 特長

- a 従来のディスクの形をした SSD ではなく、独自のフラッシュモジュールによる高性能・高密度実装
- b コントローラは、ファーム/ソフト制御ではなく、独自の FPGA (ハード回路) を開発することで超高速アクセス (レイテンシー: 0.1ms) を実現
- c データ可用性では、フラッシュメモリの特性に最適化した VSR (Variable Stripe RAID)、および 2D RAID (縦横 2 方向の RAID) により事実上無停止運用
- d ビルディングブロック構造により、2U サイズに最大 12 個のフラッシュモジュール (物理 65.97TB) とコントローラを搭載

### (2) IBM FlashSystem 840 諸元

|                    |                                                                                                                                         |
|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| フラッシュタイプ           | eMLC                                                                                                                                    |
| 記憶容量               | 物理実装容量 : 5.49~65.97TB<br>使用可能容量 : 2~48TB                                                                                                |
| IOPS               | Read : 110 万 IOPS<br>Write : 60 万 IOPS                                                                                                  |
| レイテンシー             | Read : 135 $\mu$ s<br>Write : 90 $\mu$ s                                                                                                |
| RAS 機能             | <ul style="list-style-type: none"><li>• 二次元フラッシュ RAID (2D RAID)</li><li>• 各モジュールの冗長性・ホットスワップ・工具不要交換</li><li>• 並行コード・ロード</li></ul>       |
| 暗号化                | AES-XTS 256                                                                                                                             |
| 接続形態<br>(何れか一つを選択) | 8Gbps FC $\times$ 16<br>16Gbps FC $\times$ 8<br>10Gbps FCoE $\times$ 16<br>10Gbps iSCSI $\times$ 16<br>40Gbps QDR InfiniBand $\times$ 8 |
| 電力                 | 100~240V 50/60Hz 625W                                                                                                                   |
| サイズ・重量             | 高さ 2U $\times$ 幅 445mm $\times$ 奥行 761mm<br>34kg (最大搭載時)                                                                                |

### (3) 外観

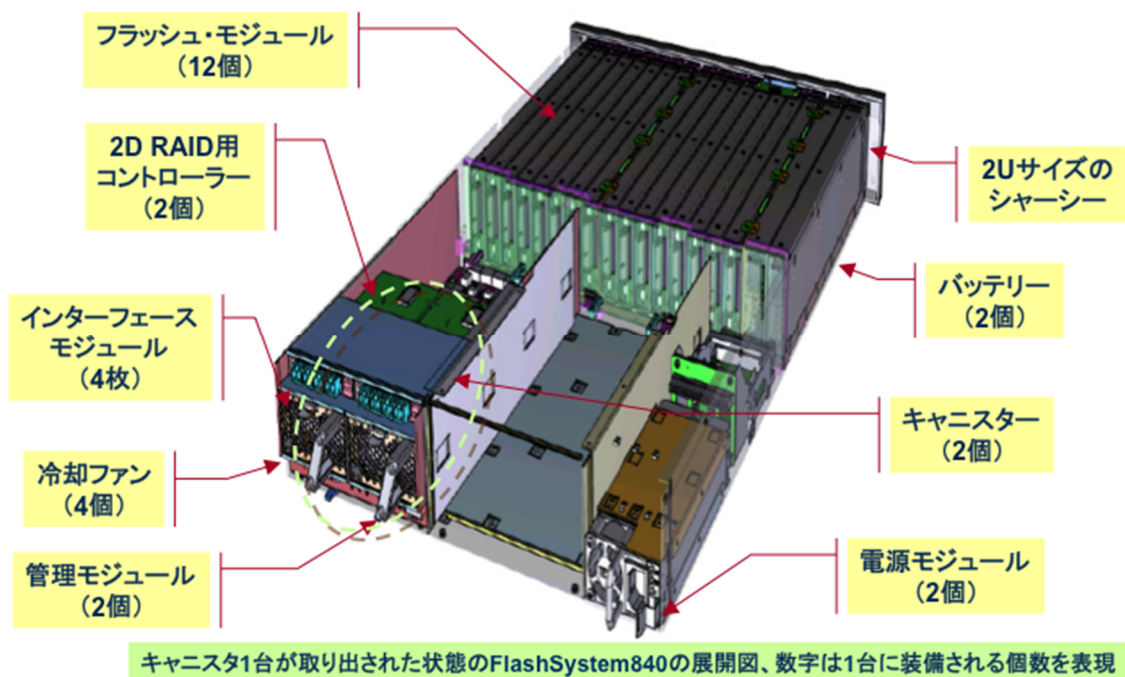


IBM FlashSystem 840 前面



IBM FlashSystem 840 後面

### (4) 内部構造



## (5) 高機能化

IBM FlashSystem840 は、超低レイテンシー、高 IOPS などの性能が特長のフラッシュストレージだが、IBM FlashSystem V840 は、それをベースに高機能化した製品である。



IBM FlashSystem V840

V840 では、以下の機能により、システム運用コストを削減する。

- 自動階層化
- FlashCopy
- (ローカル/リモート) レプリケーション
- リアルタイム圧縮
- シンプロビジョニング
- 暗号化 (注. FlashSystem 840 自体の機能)

## 付録2. お問い合わせ窓口

4. 1.       製品・サービスに関して  
          日本アイ・ビー・エム株式会社  
          アライアンス事業部  
          加藤 正則  
          e-mail : e18677@jp.ibm.com  
          URL : <http://www.ibm.com/jp/ja/>
4. 2.       見積り・オーダーに関して  
          株式会社富士通エフサス  
          サービスビジネス本部  
          お問い合わせ窓口  
          電話 : 044-874-6246  
          e-mail : proj-multi-all@ml.fsas.fujitsu.com  
          URL :  
          <http://www.fujitsu.com/jp/group/fsas/contact/index.html>
4. 3.       Oracle に関して  
          株式会社富士通北陸システムズ  
          データベースソリューション事業本部  
          お問い合わせ窓口  
          電話 : 03-5782-7511  
          e-mail : fjh-dbsupport@cs.jp.fujitsu.com  
          URL : <http://www.fujitsu.com/jp/group/fjh/>
4. 4.       この報告書に関して  
          ノアキス株式会社 (今回、検証作業のとりまとめ、報告書作成を担当)  
          ※ ストレージに関する専門知識をベースに、技術、マーケティングなどのコンサルを業務とする会社  
          金子 悟  
          e-mail : s.kaneko@noachis.com

## 付録3. 参考情報リンク

- (1) IBM FlashSystem 各種資料・ビデオ  
      “IBM FlashSystem” [検索]で、情報ページにアクセスできます。
- (2) 富士通エフサス マルチベンダーカタログ  
      <http://img.jp.fujitsu.com/downloads/jp/group/fsas/services/multi2015/top.html>