

検証報告書

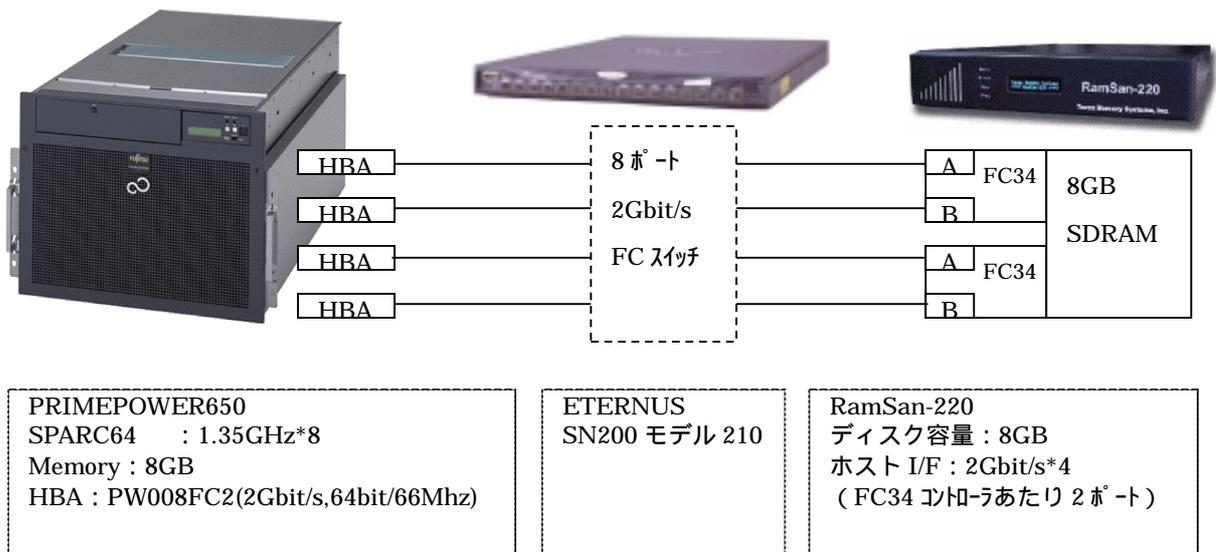
システムクリエイイト株式会社 技術部

1. 検証目的

TexasMemorySystems 社の半導体ディスク装置 RamSan-220 は、200,000IOPS の処理性能と 700MB/s の内部帯域を持ち、データベースや科学技術計算といったアプリケーションでディスク I/O のボトルネックを解消する切り札として導入されています。

今回の検証では、PRIMEPOWER650 を使用して接続確認を行うとともに、サーバーのパワー(8CPU、8GB メモリ)を利用して、このディスクの能力が引き出させるか挑戦します。

2. ハードウェア構成



3. ソフトウェア構成

OS : Solaris 8 OE 2/02

ベンチマーク : STK10 (StorageTek 社開発のフリーウェア)

4. RamSan-220 設定

論理ドライブ数 : 4つ (各 2GB)

ポートマッピング : 各ポートに論理ドライブを一つずつ割り当て

5. 検証項目

検証項目	結果	備考
富士通製ファイバーチャネルカード(PW008FC2)とダイレクトに接続してリンクアップすること	OK	4ポートとも 2Gbit/s FC-AL でリンクアップした
format コマンドで4つの論理ドライブが認識すること	OK	ディスク名は、<TMS-FC34-conXX-C.19 cyl 1021 alt 2 hd 128 sec 32> (XX = 1A, 1B, 2A, 2B)
newfs コマンドでファイルシステムが作成出来ること	OK	論理ドライブ単体、SDS とも OK
ディスク I/O が正常に行えること	OK	dmesg にてエラーが発生していないことを確認
富士通製ファイバーチャネルスイッチ ETERNUS SN200 と接続できること	OK	4ポートとも 2Gbit/s Fabric でログインし、WWN も正しく表示された
ファイバーチャネルスイッチを介して、format、newfs コマンドが正常に実行され、ディスク I/O が正常に行えること	OK	ダイレクト接続と同じく、正常に行えた

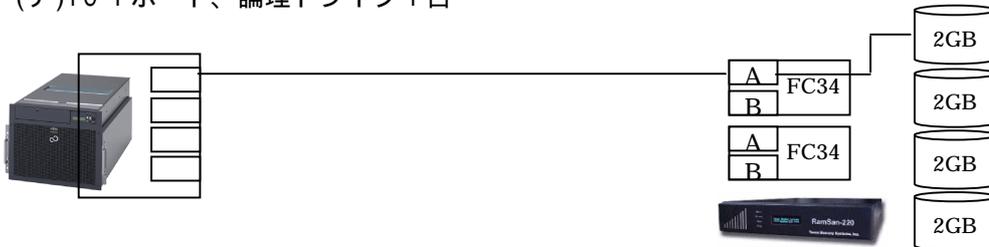
- RamSan-220 と PRIMEPOWER650 とは PW008FC ファイバーチャネルカードとのダイレクトおよび、ETERNUS SN200 を介したいずれの接続形態でも正常に行われました。
- OS からは通常のディスクドライブとして認識され、ファイルシステムの作成、アクセスが出来ることを確認しました。

RamSan-220 は、PRIMEPOWER および富士通製のファイバーチャネルカード、スイッチとの接続は問題ないと判断いたします。

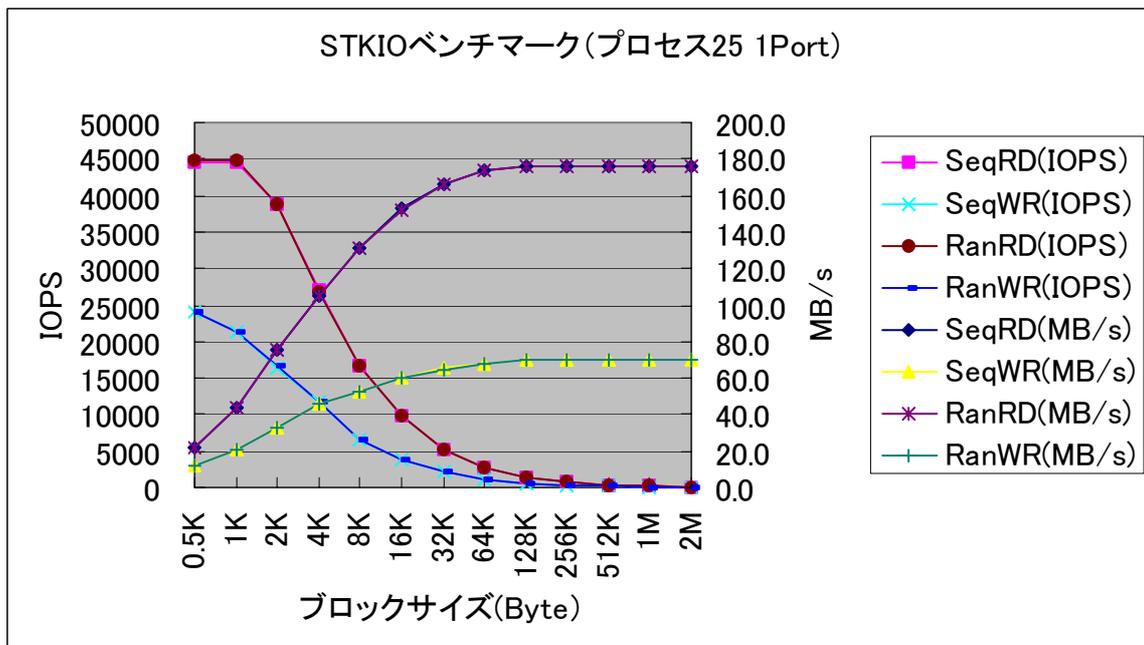
6. ベンチマーク結果

今回、メモリ上のディスクバッファの影響を避けるために、ベンチマークプログラムから RamSan-220 の論理ドライブを Raw デバイスでアクセスを行いました。

(ア) FC 1 ポート、論理ドライブ 1 台



データサイズ(Byte)	SeqRD(MB/s)	SeqRD(IOPS)	SeqWR(MB/s)	SeqWR(IOPS)	RanRD(MB/s)	RanRD(IOPS)	RanWR(MB/s)	RanWR(IOPS)
0.5K	21.7	44530	11.8	24174	21.8	44734	11.8	24152
1K	43.4	44458	20.9	21415	43.7	44756	20.9	21401
2K	75.7	38776	32.3	16512	75.9	38855	32.4	16590
4K	105.2	26941	45.5	11655	104.8	26829	45.5	11655
8K	130.7	16723	52.0	6655	130.6	16720	51.9	6647
16K	152.5	9759	60.0	3839	151.9	9722	60.1	3843
32K	166.3	5322	65.1	2082	166.3	5322	65.0	2078
64K	173.6	2777	68.0	1088	173.5	2776	68.0	1088
128K	176.3	1410	69.5	556	176.4	1410	69.5	556
256K	176.4	705	69.5	278	176.4	705	69.5	278
512K	176.4	352	69.5	139	176.4	352	69.5	138
1M	176.4	176	69.5	69	176.3	176	69.5	69
2M	176.3	88	69.5	34	176.3	88	69.5	34

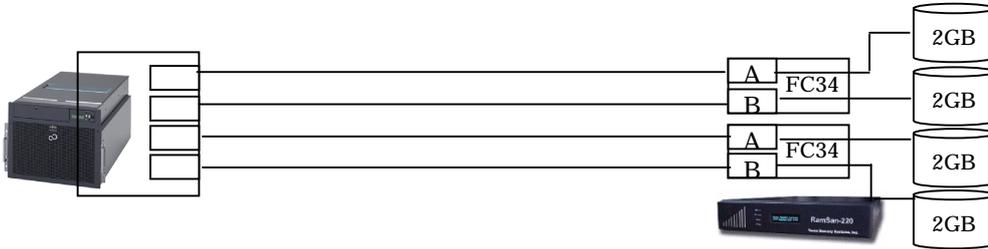


テスト実行例

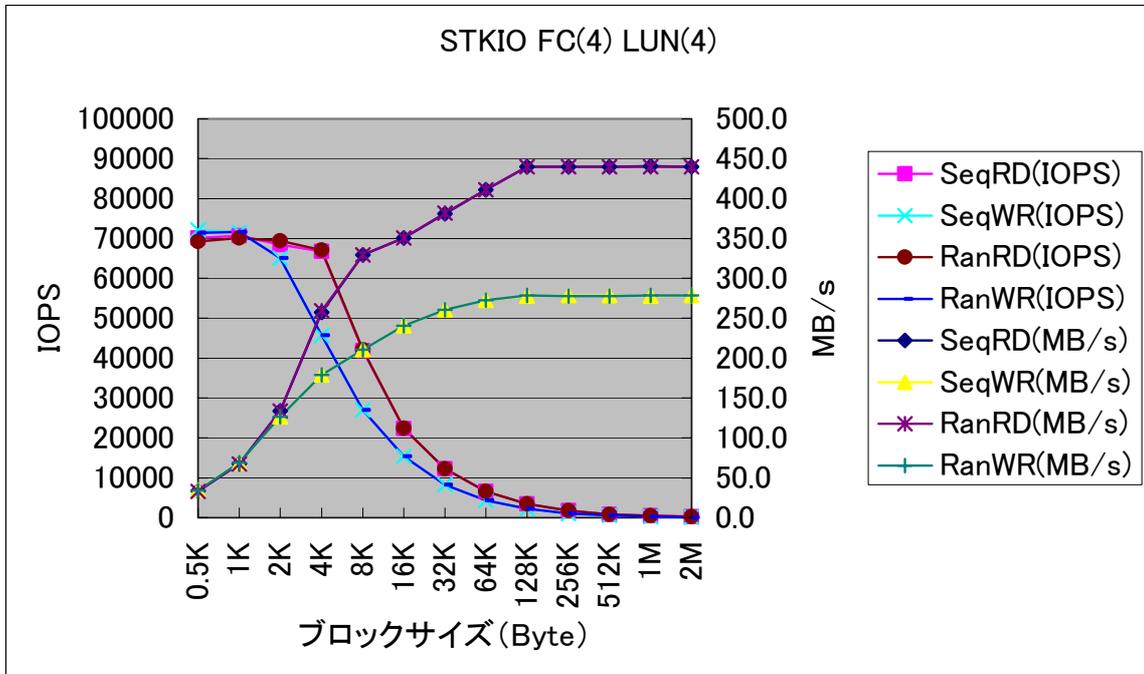
```
./stkio -f /dev/rds/c0t1d0s6 -w -t25 -m1024 -b1024 -l512 -e2097152
```

(イ) FC 4ポート、論理ドライブ4台

論理ドライブ毎に Stkio を実行し、それらの結果を合計



データサイズ(Byte)	SeqRD(MB/s)	SeqRD(IOPS)	SeqWR(MB/s)	SeqWR(IOPS)	RanRD(MB/s)	RanRD(IOPS)	RanWR(MB/s)	RanWR(IOPS)
0.5K	33.4	70075	34.6	72195	33.2	69235	34.3	71432
1K	67.8	70750	68.5	71537	67.3	70128	68.9	71666
2K	133.6	68390	126.4	65040	133.7	69381	126.4	65057
4K	257.5	66713	178.7	45751	259.1	67143	178.6	45736
8K	329.5	41947	210.9	26990	329.3	42157	210.9	26996
16K	350.4	22287	240.6	15394	350.5	22427	240.5	15392
32K	381.4	12203	260.7	8340	381.9	12219	260.7	8341
64K	411.3	6578	272.4	4358	410.9	6574	272.4	4357
128K	440.1	3518	278.2	2223	439.9	3518	278.2	2223
256K	440.1	1758	278.3	1110	439.9	1758	278.2	1110
512K	440.0	878	278.2	554	439.9	879	278.2	554
1M	440.5	438	278.2	276	439.8	439	278.2	276
2M	440.0	218	278.2	136	439.7	219	278.2	136



考察：

- (ア) (イ)とも、リード/ライトの転送速度および IOPS は、シーケンシャル/ランダムアクセスパターンの違いに関係なく同等の数値でした。磁気ディスクは、ディスクの回転待ちやヘッドのシークなどの機械的動作が必要で、ランダムアクセスにおけるパフォーマンスが著しく低下してしまいますが、シリコンディスクの RamSan-220 では、この問題が生じないことが判ります。
- ライトとリード時のパフォーマンスの差が見られますが、TMS 社からの回答では、ライトはリードより遅く、ライトの「ホストが SCSI のライトコマンドを送る Ramsan-220 が ack を返す ホストがデータの転送を開始する RamSan-220 がデータを受け取ったと応答する」というプロセスの中で、 ~ で時間がかかることがあるということでした。但し、これは、FC アナライザを使用して追求する必要があります。
- 1 ポートのリードでは、IOPS が最大 45,000 でした。しかし、FC 4 ポートから RamSan-220 の 1 ポートにアクセスさせると、70,000 IOPS (コントローラ当たりのスペック値は 100,000 IOPS) まで測定できたことから、この値はファイバーチャネルカードの限界値と考えられます。
- FC 4 ポートで 4 つの論理ドライブにアクセスさせた時、IOPS は 70,000 までしか達しませんでした。原因を追求したところ 8 つの CPU を全て使い切っており (アイドル 0%)、CPU が I/O プロセスがこれ以上実行出来ないことが判りました。つまり、200,000 IOPS の性能を引き出すには、さらに CPU 数を増やす (もしくは、複数のマシンに接続する) 必要があると考えられます。
- SDS によって FC 4 ポートを束ねた場合の測定も行っていますが、インターレス 32K とした時シーケンシャルで 31,000 IOPS、ランダムで 55,000 IOPS が最高でした。これは、I/O アクセスがポート間をローテーションして行われるためストライピングの効果 (ランダムだとポートが分散されるので効果が上がる) が出難かったと考えられます。

お問い合わせ先

システムクリエイイト株式会社

営業担当：二見 由紀男

技術担当：白井 紀行

電話：03-5296-3775

FAX：03-5296-3777

Mail：sales@sc-i.co.jp

URL：<http://www.sc-i.co.jp>

※Solaris™ Operating Environment および Solaris™オペレーティング環境を、『Solaris OE』と記述しています。