

SX-8 の高速ディスク装置

High-performance Disk Array for SX-8 Series

滝柳 真澄*
Masumi Takiyanagi

江田 由則*
Yoshinori Eda

斉藤 義洋*
Yoshihiro Saito

萩原 博之*
Hiroyuki Hagiwara

相合 孝雄*
Takao Aigo

要 旨

SX-8 シリーズに接続可能な SAN 対応のディスクアレイ装置 iStorage S2800/S2400 を製品化しました。これらの製品は、先進のテクノロジーにより科学技術計算領域に要求される高性能、高拡張性に加えて、大容量システムを安定稼働させる高信頼、高可用性を実現しています。

本稿では、iStorage S2800/S2400 の特長と製品に搭載されるテクノロジーについて紹介します。

NEC's iStorage S Series are the high-performance and high-available disk array dedicated to Storage Area Network (SAN). The iStorage S2800 and S2400 were designed specifically to provide the scientific and engineering calculation system for high performance and high scalability. Furthermore, the iStorage S2800 and S2400 come to possess many high-reliability and high-availability features, including fully redundancy, and reducing disk drive replacement rate with proven proprietary technology.

1. まえがき

近年、科学技術計算領域においては、CPU の処理性能の高速化や、シミュレーション技術の進歩により、大規模計算が実行されるようになり、データを格納するストレージ装置に対しても、より大量のデータをより高速に処理することが求められています。

SAN (Storage Area Network) は、ファイバチャネルインタフェースの技術を基盤とした、ストレージをネットワーク接続する技術ですが、インタフェース速度が 200M バイト/秒と高速な上、多数の計算機ノードから多くのディスク装置を共有できることから、最近特に科学技術計算分野で多く利用されてきました。

iStorage では、上記のような要求に応えるべく SX-8 向けのストレージ装置を強化しました。

2. SX-8 シリーズ向け製品のラインナップ

従来より、SX シリーズ向けのディスク装置には、RAID (Redundant Array of Independent Disks) 装置が使用されていました。RAID は HDD を冗長化させる技術で、高可用性のメリットがありますが、さらにディスクを並列にアクセスすることで、高い性能が得られるメリットもあります。たとえば RAID-3 は、大規模なアクセスにおいて高いスループットが得られ、RAID-5 では、複数のディスクを個別にアクセスすることで、比較的小規模なデータアクセスにおいても高い応答性能が得られます。iStorage 製品としてはこれまで SX シリーズ向けに、RAID-3 タイプの iStorage S1230 ディスクアレイ装置や、RAID-5 タイプの iStorage S2300 ディスクアレイ装置を提供しています。S2800/S2400 ディスクアレイ装置は、S2300 の後継として、最新テクノロジーを搭載した製品です。

3. iStorage S2800/S2400 ディスクアレイ装置

iStorage S2800 は SAN 対応 iStorage S シリーズのミッドハイレンジモデル、S2400 はミッドレンジモデルです (表)。

(1) iStorage S2800 ディスクアレイ装置

iStorage S2800 は、高さ 3EIA (Electronic Industries Association, EIA 規格 1 ユニット=44.45mm) の筐体に 2 つのコントローラを実装したコントローラ筐体と、同じく高さ 3EIA で 15 台のディスクを搭載可能なディスクエンクロージャ筐体で構成されます (写真)。

36G バイト/73G バイト/147G バイト/300G バイトのディスクを、最大 16 台のディスクエンクロージャに 240 台まで搭載可能で最大容量 55T バイトのスケラビリティを持ったディスクアレイ装置で大規模/大容量のデータを扱う SX システムの用途に適しています。S2800 は RAID-1/5/6/10/50

* システムファイル事業部
System File Products Division

表 S2800/S2400 ディスクアレイ装置の諸元
Table Specifications of S2800/S2400.

項 目		S2800	S2400
ホストインタフェース		200M バイト / 秒 Fibre Channel × 8	200 Mバイト / 秒 Fibre Channel × 4
ドライブインタフェース		200 Mバイト / 秒 Fibre Channel × 8	200 Mバイト / 秒 Fibre Channel × 4
キャッシュメモリ		4 ～ 16GB	4 ～ 8GB
最大装置容量		55T バイト	27.6T バイト
最大HDD搭載台数		240 台	120 台
RAIDタイプ		RAID-1, RAID-5, RAID-10, RAID-50, RAID-6	
冗長コンポーネント		コントローラ, ファン, 電源, BBU (バッテリーバックアップユニット)	
接続ホスト		SX-8 シリーズ	
寸法 (W×D×H)	コントローラ	482 × 596 × 131mm	482 × 596 × 131mm
	ディスクエンクロージャ	482 × 596 × 131mm	482 × 596 × 131mm
質量	コントローラ	36kg 以下	36kg 以下
	ディスクエンクロージャ	45kg 以下	45kg 以下
消費電力		8,564W 以下	4,444W 以下
エネルギー消費効率		0.09	0.1
その他機能		フェニックス技術	



写真 S2800/S2400の外観

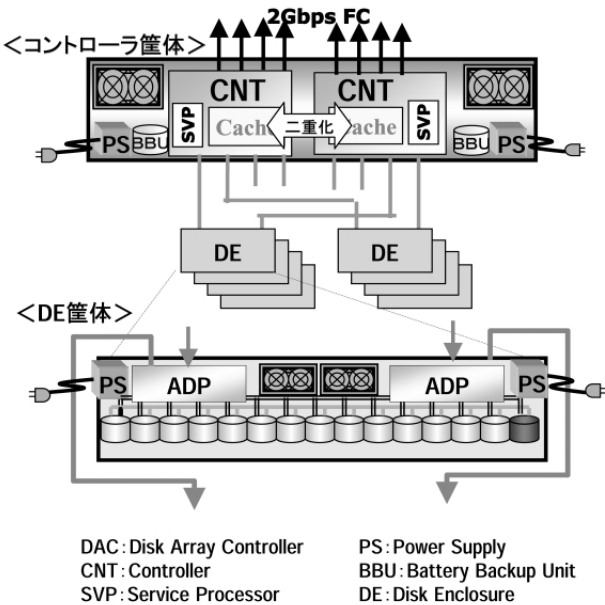
Photo External view of S2800/S2400 disk array unit.

に対応していますが、SXシリーズでは高速性が重視されるため、RAID-5をサポートしています。ホストポートを標準で8ポート装備しており、SXシステム独特のSANによるファイル共有ソリューションであるGFS（Global File System）構成を構築するのに有効です。

また、信頼性、可用性の面においては、RAID構成のディスクを始め、コントローラ、キャッシュメモリ、電源、キャッシュメモリをバックアップするバッテリーなど、すべてのコンポーネントが冗長構成になっており、No SPOF（Single Point of Failure）を実現しています（図1）。

(2) iStorage S2400 ディスクアレイ装置

iStorage S2400 も、RAID-1/5/6/10/50に対応しておりS2800と同様にSXシリーズではRAID-5のディスクアレイ



DAC: Disk Array Controller
CNT: Controller
SVP: Service Processor
PS: Power Supply
BBU: Battery Backup Unit
DE: Disk Enclosure

図1 S2800/S2400 ディスクアレイ装置の構成

Fig.1 Configuration of S2800/S2400 disk array unit.

として使用できます。ホストポートは4本であり、ディスクエンクロージャは最大8台、合計120台のディスクを搭載可能で最大容量27.6Tバイトと、大容量ミッドレンジディスク装置を実現しています。

4 . S2800/S2400 におけるテクノロジー

iStorage S2800/S2400では、高性能、高信頼性、高可用性を実現するため、次のようなテクノロジーを採用しています。

(1) 高速シリアルバス

iStorage S2800/S2400では、コントローラ間のインタフ

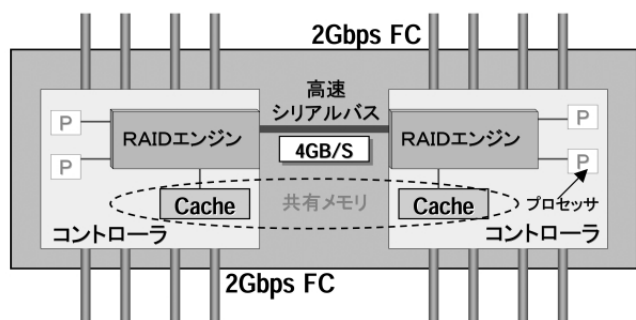


図2 高速シリアルバスとRAID エンジン

Fig.2 High-speed serial-bus and RAID engine.

エース（I/F）に2.5Gbpsの高速シリアルバスを採用し、データ用として8本、制御用として4本を割り当て、最大で2Gバイト/秒のデータ転送が可能です。Writeデータをコントローラ間で冗長化された共有メモリに高速に同時書き込みすることが可能となり、Write性能を大幅に向上させました（図2）。

(2) 新開発RAIDエンジン

高性能、高可用性ディスク装置の心臓部ともいえるRAIDエンジンを新規に開発しました。このRAIDエンジンには4種類の高速度チャンネルがあり、データを高速転送するとともに、データの信頼性を高めるためのチェックコードをデータの転送と同時に付加することができます。各チャンネルの主な機能は以下です。

- ・FWチャンネル：キャッシュメモリをFWで使用するためのチャンネルで2チャンネルあります。
- ・HWチャンネル：256チャンネルのDMAエンジンです。ディスクーキャッシュ間、またはホストーキャッシュ間の転送において、チェックコードの付加/削除/チェック機能を有しており、1チャンネルで同時に2つのキャッシュへの転送が可能です。高負荷/多重転送にも耐えられるよう設計されています。
- ・EXORチャンネル：4チャンネルのEXOR演算エンジンです。RAID-5/-6の演算を行い、パリティの生成とデータ復旧を行います。
- ・MMチャンネル：キャッシュメモリ間専用のDMAエンジンで2チャンネルあります。

以上の4種のチャンネルを使い分けることで高速にデータ演算/転送を実現しています。

(3) スイッチ方式のディスクエンクロージャ

従来のPBC（Port Bypass Circuit）方式に代わって今回スイッチ方式を採用したディスクエンクロージャを新規に開発しました。特長を以下に示します。

- ・従来のPBC方式と比較してArbitrationの時間が短縮され、ディスクアクセス時の応答性能が向上する。
- ・コントローラとHDDの接続が1：1となり、信頼性面においても故障HDDによるファイバチャネルのエラーが他のHDDに伝搬しない。

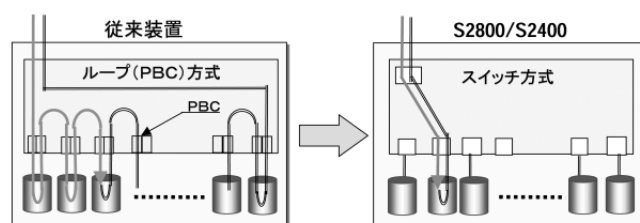


図3 スイッチ方式のディスクエンクロージャ

Fig.3 Switch-type disk enclosure.

- ・HDD挿入時、ポート診断を実施してからファイバチャネルへの組み込みを実施可能なため、故障したHDDをファイバチャネルに組み込むことがない。

この新開発ディスクエンクロージャを採用したことで性能が向上するとともに、HDD起因障害による影響の極小化が図れ、サブシステムの安定稼働を向上できます（図3）。

(4) フェニックス技術

HDDの応答遅延や、一過性の不調を検出した際に、該当HDDを一時的に切り離し、必要なリカバリ処理や、代替セクタへの割付、診断等を行った上で“再組み込み”または“故障”を判断し処置するフェニックステクノロジーを採用し、HDDの縮退率を大幅に低減しています。

この技術により、データのフルコピー処理が行われるため、性能劣化が起こるRAID再構築処理を減らせます。また、RAIDの非冗長時間の短縮を図り、HDDが本来持っている自律回復能力を発揮させ、サーバへのレスポンス性能を維持しながら、RAIDのデータ保護能力を高めます。

5. 今後の強化

iStorage S2800/S2400では、ディスクの大容量化に伴うさらなる信頼性向上のニーズに応じてRAID-6をサポートしています。データに対するパリティを二重化することにより、HDDが1台故障しても冗長状態を保つことが可能で、HDDが2台故障しても継続して業務を行うことが可能となり、HDD故障における信頼性を飛躍的に向上することができました。また、複数のRAID-5にデータをストライピングして記録し、大規模アクセスにおけるスループットをRAID-3相当に向上するRAID-50の機能も搭載しています。今後は、SXシリーズでもRAID-6/RAID-50をサポートし、信頼性/性能の向上を行っていきます。

さらに、RAID-6では仮想的な記憶容量管理を実現したダイナミックプールをサポートしており、これによりRAIDの構成にとらわれない自由な記憶容量管理の実現が可能となります。ダイナミックプールの仮想記憶容量管理には、次のような特長があります。

- ・物理ディスクをまとめて仮想的にプール管理、HDD 1台単位の増設が可能。
- ・物理的なRAID構成の制約を受けずにプールから任意の大きさの論理ディスクを生成。

- ・既存の論理領域をデータ退避せずに容量を拡張可能。
- ・RAID 構成に縛られずに、装置全領域を1論理領域で利用可能。

6. むすび

本稿では、SX-8シリーズ向けのディスクアレイ装置とそのテクノロジーおよび今後のサポート強化予定について述べました。今後の科学技術計算分野においても、オープンシステムと同様にストレージシステムの統合や、バックアップなどを含めた運用の効率化などが期待されています。

ディスクアレイ装置においてもiStorageシリーズを軸に、高機能化、高性能化によるソリューションの強化と、さらなる高信頼性、高可用性の実現を図っていく予定です。

筆者紹介



Masumi Takiyanagi
たきやなぎ ますみ

滝柳 真澄 1986年、NEC入社。現在、第一コンピュータ事業本部システムファイル事業部iStorage製品技術部技術エキスパート。情報処理学会会員。



Yoshinori Eda
えだ よしのり

江田 由則 1983年、NEC茨城入社。現在、NEC 第一コンピュータ事業本部システムファイル事業部第一技術部技術エキスパート。



Yoshihiro Saito
さいとう よしひろ

斉藤 義洋 1991年、NEC入社。現在、第一コンピュータ事業本部システムファイル事業部第一技術部技術エキスパート。



Hiroyuki Hagiwara
はぎわら ひろゆき

萩原 博之 1990年、NEC茨城入社。現在、NEC 第一コンピュータ事業本部システムファイル事業部第一技術部主任。



Takao Aigo
あいごう たかお

相合 孝雄 1993年、NEC入社。現在、第一コンピュータ事業本部システムファイル事業部第一技術部主任。