



Storage Bridge Bayからの出発

～SBB規格がもたらすストレージの未来像～

エムアイシー・アソシエーツ株式会社

ここに記載された内容は更新される可能性があります。この文書に記載されている内容はこの文書の発行時点におけるエムアイシー・アソシエーツ株式会社の見解を述べたものです。エムアイシー・アソシエーツ株式会社が、この文書に記載された内容の実現に関して確約するものではありません。また発行日以降については、この文書に記載された内容の正確さは保証しません。

この文書は情報の提供のみを目的としており、明示的または黙示的に関わらず、この文書の内容について エムアイシー・アソシエーツ株式会社はいかなる保証をするものでもありません。

エムアイシー・アソシエーツ株式会社は、本書に記載してあるすべて、または、一部の記載内容に関し、許可なく転載、または、引用することを禁じます。

CLARiX AX4はEMC Corporation 社の登録商標です。

OneStorはXyratex社の登録商標です。

SBBはStorage Bridge Bay Working Group Inc社の商標登録です。

XScale、Nehalem、Jasper ForestはIntel社の登録商標です。

その他、記載されている会社名、製品名は各社が所有する商標の場合があります。

バージョン	作成日付	旧バージョンからの	総ページ数
3.0	2010/07/29	・ SBB2.0からみえるストレージ製品 追加	11
2.0	2010/06/24	・ SBB2.0とJasper Forest 追加 ・ Storage Processor : Jasper Forest 追加	8
1.0	2010/05/27	初版	7

本書作成、編集、管理



エムアイシー・アソシエーツ株式会社
〒103-0004 東京都中央区東日本橋3-12-12
櫻正宗東日本橋ビル9F
Tel. 03-5614-3757 Fax. 03-5614-3752

目次

SBBとは	1
SBB規格による内部バスの標準化	2
SBB規格によるストレージ製品	2
SBB標準化の示す未来のストレージ	3
RAIDコントローラ用CPUの動向	3
Storage Processer : Jasper Forest	4
SBB2.0とJasper Forest	5
SBB2.0からみえるストレージ製品	5

SBBとは

SBBという言葉をご存知でしょうか？実は、ストレージ業界では広く知られた言葉です。SBBとは「Storage Bridge Bay」の頭文字で、ストレージハードウェアを低位なレベルで標準化した国際的な共通仕様のことです。

2006年3月、Dell、EMC、LSI、Intel、Xyratex 他、8社が集まり、エントリーレベルからミッドレンジのストレージにおいてJBOD [Just a bunch of disks (ストレージ筐体)] や、RAID装置製品等のストレージ製品の筐体各部品サイズの標準化、コネクタピンアサインの標準化など、物理的、電氣的、低位レベルの仕様を提案するStorage Bridge Bay Working Group Inc. (略称、SBB) を立ち上げました。そこで決定された標準化仕様をSBB規格とといいます。

SBBによるハードの標準化

仕様標準化により、ストレージコントローラ等を設計するベンダは、迅速に、より高度なソリューションを提供できるようになりました。また、従来エンタープライズレベルのストレージに限られていた技術やソリューションを全てのストレージユーザが使用でき、より安価な利用ができることを目指しています。

右図はコントローラ、IOモジュールを安全に挿入するためのハンドルラッチ機能のラッチ部分を定めた図面です。SBB仕様では各社のIOモジュールやRAIDコントローラを安全に筐体に挿入し、筐体とIOモジュール、コントローラ間のピンコネクタが安全に勘合することができるよう、挿入されたIO

モジュールを筐体にラッチし、スライドさせるレバーを使用します。SBB1.0の規格ではこのレバーの先端にあるラッチ機構のコーナー半径や、フックの長さ等、細かい寸法を詳細に定め、筐体ベンダと異なるベンダが開発したIOモジュールや、コントローラでも物理的には接続できるように規定しています。

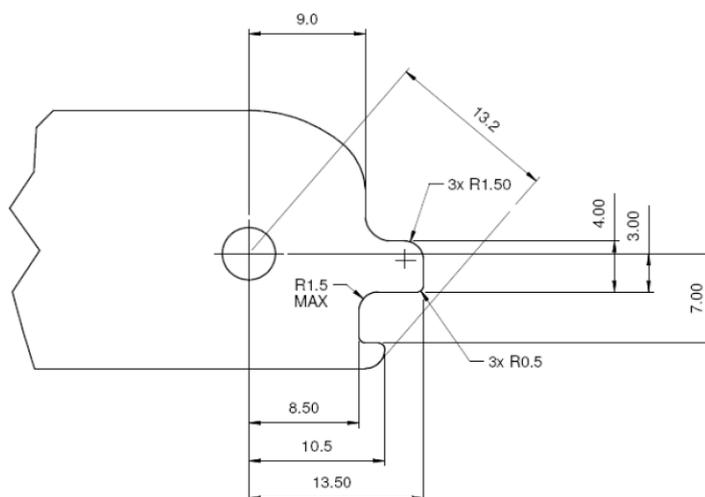


Figure 10: SBB Lever Mechanism Dimensions

SBB規格による内部バスの標準化

ミッドレンジからエンタープライズレベルのRAID製品において2002年頃にはプリント基板を簡素化できるメリットから、内部バスがパラレルSCSIからFibre Channelに移行し、1GbのFC-AL (Fibre Channel Arbitration Loop)が内部インタフェースに実装される様になりました。

2004年、業界ではSAS(シリアルアタッチドSCSI)というメタルケーブルを使用したインタフェースが標準化され、主にストレージ製品の内部インタフェースとして使用されるようになりました。SASインタフェースは3Gbのバンド帯域を持つバスが4レーン分が1つの物理インタフェースを構成し、合計で12Gb/秒のデータを転送することができます。また、FC-ALと同様にデュアルチャンネルをサポートしていますので、ストレージ内部バスのHA化を容易に図ることができます。

一方、SASインタフェースはFC-ALとは異なり、Expanderと呼ばれるスイッチ機能を持つデバイスに総てのノードが放射上に接続されるため、FC-ALの様に一つのノードの障害がループ全体の性能低下を引き起こす問題から開放されます。この為、現在殆どのミッドレンジからエンタープライズレベルのストレージの内部バスはSASインタフェースを採用するようになってきました。

SBB-1.0、2.0ではストレージ製品の内部バスをSAS、FC、PCIeの3つの種類のインタフェースの仕様を大まかに定めていますが、実情を反映し大半はSASに関する記述にあてられています。SASインタフェースは今日3Gbから6Gbへの世代が交替する時期に当たり、2008年1月に発表されたSBB2.0ではSAS 6Gbの仕様が定義されています。

SBB規格によるストレージ製品

SBBワーキンググループの立ち上げとほぼ同時期に、EMCとNECはSBB準拠のエントリーレベルのストレージ製品を共同開発することに合意し、その2年後の2008年2月にEMCはCLARiX AX4というモデルを発表しました。ホストインタフェースはFC 4Gbpsの他に、1GbpsのiSCSIインタフェースを持ち、ネットワークメディアに対して柔軟に対応できるストレージ製品になりました。

一方、Xyratex社では2008年10月にOneStor Extensible Storage PlatformファミリーとしてこのSBB2.0仕様に基づくストレージ製品を発表しました。OneStorはモジュラー化されたSBBベースのアーキテクチャーを採用した上で、高いドライブ実装密度、SASホストインタフェース、インテリジェントなプラットフォームマネジメント、高効率電源仕様等、業界屈指の多用性と適合性を持つストレージ装置です。OneStorファミリーとしては3.5”HDDを12台搭載する2UのSP1212、24台搭載する4U SP1424。更に2.5”HDDを2U筐体に24台搭載するSP1224があります。OneStor Storage Platformファミリーは、ホストとの接続インタフェースや、筐体間の接続インタフェースを装備するIO

モジュール、電源、冷却ファンモジュールがファミリーのどの筐体でも共通化されています。2008年のSBB2.0の標準化仕様の発表までに、SBB Working Group への参加企業が43社にのぼりました。

SBB標準化の示す未来のストレージ

OSレベルでサーバの仮想化環境が整いつつある現在、ユーザデータの多様化や、非構造化によってストレージに対する要求はますます高まっています。

SBBの1.0仕様が発表された頃には、IDCや、ガートナー等の大手調査会社は従来の企業基幹システムで使われる構造化データに比較して、非構造型のユーザデータが今後数年間で大きく増大し、これらデータに対応して行く必要を訴えていました。特に映像や画像などのデータを高信頼、高性能、かつ、安価なストレージに保存し、どう再利用するかというテーマに対し、ストレージ業界はSBBという低レベルの標準化から急激に増大する非構造化データに対するストレージの姿を示して行く必要がありました。

SBB仕様に準拠し、ストレージ製品メーカーはこのようなユーザシステム環境の急激な変化に対して新しいストレージ提案の準備を進めています。

RAIDコントローラ用CPUの動向

さて、いままで、SBBはStorage Bridge Bay Working Groupという組織で標準化作業が行なわれて来たことをお話しました。そして、その組織のキープレーヤとしてEMCや、Xyratex等のストレージベンダ以外にどういう訳かIntel社がその発起人に名を連ね、Working Group設立時から参加しています。

Intel社はノートブック、デスクトップやサーバのCPUを開発製造する世界最大のメーカーであることはご周知の通りですが、同時に、ストレージ業界においてもコントローラ用の組み込み用プロセッサの最大供給ベンダです。1990年代後半にDECのStrongARMという32ビットRISC プロセッサの事業を買収したことにより、2000年から低消費型の組み込みプロセッサをXScaleという製品で提供しはじめました。

XScaleのラインアップには携帯電話等の通信機器で使用するファミリーの他、ネットワーク機器やストレージコントローラ等に仕様されるファミリーがありました。ストレージコントローラ用CPUにおいてはRAID-5に使用するXOR(排他論理和) 機能を内蔵しており、それまで、ASIC等でRAIDコントローラの制御部分をカスタムで起こしていたRAIDコントローラベンダにとっては、自分たちが開発するファームウェアからRAIDのXOR 機能をオフロードし、よりフレキシブルなRAIDのセットアップや、よりカスタマイズしたエラーハンドリング等、高度なストレージ制御を行うことができるようになりました。

2005年にリリースされたXScale IOP333からインテル独自のP+Q data protectionというRAID-6の機能が追加されました。従来RAID-6のダブルパリティの機能はカスタムのASICや、ファームウェアに追加コードを起こして行なっていたオーバーヘッドが大幅に軽減されることになり、多くのRAIDコントローラに標準で実装されるようになりました。

しかし、XScaleにはIOプロセッサファミリー、ネットワーク機器用のプロセッサ以外に携帯ワイヤレス端末用等の低消費電力のプロセッサファミリーがあり、大きなシェアを持っていました。ところがIntel社は2006年にIOプロセッサとネットワーク機器用プロセッサ以外のXScale製品事業をマーベルセミコンダクターに売却してしまいました。

それから3年、昨年4月にIntel社はNehalemアーキテクチャと呼ばれるCore 2マイクロアーキテクチャの後継になるプロセッサラインアップを発表しました。また、同年9月にはJasper Forestと呼ばれる、Nehalem EPアーキテクチャでシングル、デュアル、クワッドコア仕様で、24Wから最大65Wの低消費電力型のCPUを発表しました。このJasper ForestはCPUの他、PCI Express Gen2 コントローラを内蔵し、外部に16 Lane(5GT x 16)の高バンド帯域のIOバスを提供します。また、別にIOHチップを通してPCI Express 2.0のバスを 36 Laneを使用することができますので、大変パワフルな外部へのインタフェースが揃うことになります。

Storage Processor : Jasper Forest

Intel社は更に、Jasper ForestにStorage Processor として位置づける機能を付加しています。それは、RAID-5や、RAID-6のRAIDコントローラ機能をCPUの一部に割当て高速に処理する他、電源障害時に最大48GBまでのメモリ中のデータをSSD等の高速デバイスにフラッシュする機能です。これらの機能は XScaleをリプレースするに十分なものです。

このように、Jasper Forestは新世代のストレージコントローラ用プロセッサと位置付けられています。現在、Xyratex社をはじめ、HP社やEMC社等がエンタープライズ仕様の次世代のストレージをこのJasper Forest を使用し開発を行なっているようです。また、更にIntel社のJasper Forestは、SBB 2.0とピンアサインに互換性があります。直接SBBのコントローラベイのスロットにJasper Forest を実装したコントローラを挿入することができるということです。

少し回り道を致しましたが、SBBはIntel社のJasper Forestを搭載したストレージを映像データや、画像データが急激に増大する時期に合わせて開発されたと考えるのは筆者だけでしょうか？

SBB2.0とJasper Forest

以降は筆者の予測になります。業界関係者としてこのようなデザインフォーキャストを出して良いものか悩みましたが、現在の雲の先に少しでも晴れ間が見えるといいのではないかと思います勝手に予想をしてみたいと思います。

まず、Jasper Forestによって現在のIOP333や、348に比べてどの位の性能向上が期待できるかを予想する必要があります。

一般的にCore 2マイクロアーキテクチャとNehalemマイクロアーキテクチャの同一クロック性能のサーバのパフォーマンステストを行なうと、Nehalemの方が9倍程度高速になる、又は、9分の1のサーバ台数で同一レベルの処理ができると言われています。この比をそのままXScaleとJasper ForestのSingle CPU性能に持ってくる事が妥当とは思われませんが、少なくとも、この比の半分は保証されるのではないかと推測されます。則ち、現在のXScaleベースのストレージコントローラの4倍程度の処理能力とPCIe 1.0(2.5GHz) 4 lane からPCIe 2.0 (5.0GHz) 16 lane程度の高速なIO性能を持つ事になります。以上の推測から、Jasper Forestを搭載したSBB2.0ベースのコントローラは現在我々が日常使用しているRAIDコントローラの数倍の仕事をする事ができます。

SBB2.0からみえるストレージ製品

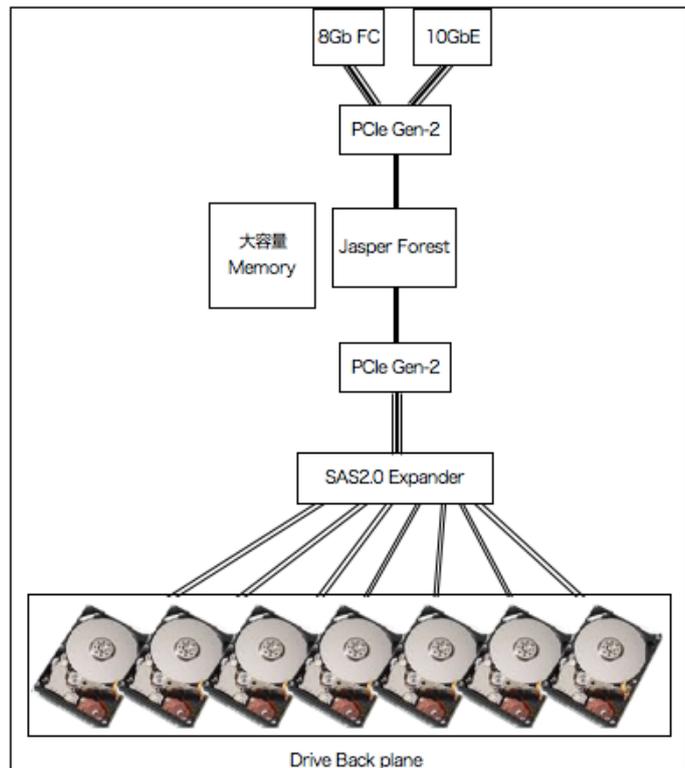
今までSBB2.0におけるSAS2.0規格の 6Gb インタフェース、FC-8Gb インタフェース、Intel 社のNehalemアーキテクチャをベースにしながらXOR機能を内蔵したストレージコントローラとしても使用できるCPUのことをご紹介してまいりました。

また、大規模ストレージシステムを構成するセル的な部品としてのハードディスクにおいてはSAS(Serial Attached SCSI)インタフェースのドライブとSATA(Serial ATA)インタフェースのドライブで着実な進化が起きています。エンタープライズクラスのハードディスクインタフェースはSAS2.0となり、全二重で書き込みも、読み込みもそれぞれ600MB/秒(6Gb/秒)の性能で同時に実行されます。また、SAS 2.0仕様のドライブはIO性能を向上させるための大容量リードライトキャッシュや、コマンドを事前に受付処理し、128コマンドまで受け付けることができるコマンドキューイング(Command

Queuing)、回転待ち時間、ヘッドシークを削減する等、より効率のよいリードライトを実現する上で大切な機能が充実しています。一方、SATAもSATA2.0の仕様となり、SASドライブと比較すると諸機能で劣りますが、ミッドレンジクラスのストレージ製品として充分使用可能な機能が備わってきています。

以上のように、SBB2.0筐体内部には高性能、高速なドライブ、SAS2.0仕様のバックプレーン、Jasper Forestレベルの高性能CPU、大容量メモリ、更に、PCI Express Gen-2の5Gb 4レーンや、8レーンの広帯域ローカルバス上のスロットに挿入されたFC、SASや、10Gb NICカード等、高速データ転送を保証するハードウェアを装備することが可能になります。

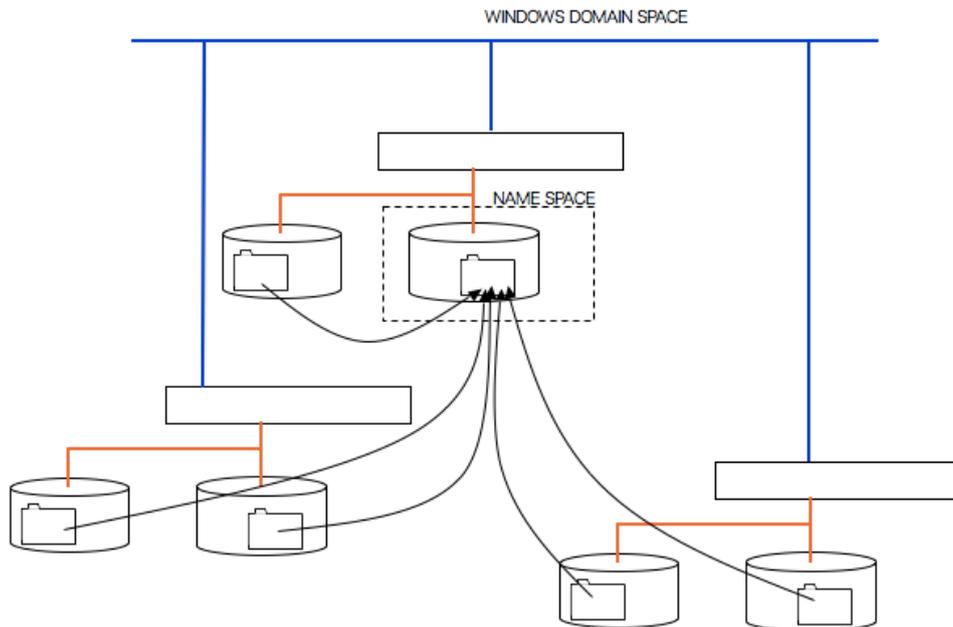
左図はSBB2.0ストレージ筐体の内部で実現可能と想定される高速なRAID装置のブロック図です。高速HDDとSAS2.0 Expander を結ぶバスは全二重の6Gb SASです。また、大容量メモリは最大48GB程度は実装可能です。このように高性能ストレージシステムにおいては、従来のように特殊なOSをRAIDのファームウェアを搭載するだけでなく、x86アーキテクチャで稼動可能な汎用のWindows OSや、Linux OSとそれらのOS上で稼動する仮想OSやアプリケーションも搭載することがきることでしょう。



現在、CloudやGridの名前でよばれるストレージの仮想化技術は今後ますますユーザに利用されると考えられます。例えば、Windows 2008においてはDFS/Name Spaceというストレージの仮想化の概念が既に内包されていることは読者の多くの方がご承知の通りです。

これは、SANストレージや、内蔵ディスク、ネットワークストレージのディレクトリー上の位置情報を同一ネームスペースとして管理し、ネームスペースを管理するサーバホストにアクセスすることで、そこに収容されたパス情報に基づき、ディレクトリコントローラ配

下にある他のネットワークストレージ上のデータに対して更新したり、読み出したりすることができる分散ファイルシステムです。(下図参照)



また、Linuxを代表するRedHat 5.0でもGFSをベースにしたストレージのグリッド化が進化しています。GNBDとよばれるストレージノードをGNBDサーバで管理し、このGNBDサーバにアクセスすることで、ネットワーク上の他のGNBDストレージノード内のファイルにアクセスすることができます。

このように、近年のOSには既にストレージの仮想化機能が実装されつつありますので、SBB2.0筐体に搭載されたストレージシステムではこれらの分散型のファイル管理システムをRAIDシステム中で実現することができるようになると予想されます。

一方、基本OS上にゲストOSを複数搭載するサーバ仮想化に関しても、RAIDコントローラ内で大容量ストレージに対する直接的なアクセスを保証することで、複数OSからの単一ストレージに対する同時にアクセスする際に発生するアクセス性能の大幅な劣化(Congestion)の問題を解消することが可能になります。

また、既に多くのエンタープライズストレージ製品で実装されはじめた

- ハードウェアSnapshotに基づくディスク内でのデータのリアルタイムバックアップ
- SSDや、SAS、SATAドライブ等で構成されるストレージの階層化(Multi-Tier

Storage)でのデータのタイプ、アクセス頻度等のポリシーベースでのデータマイグレーション

- 重複データの削除(De-duplication)
- 必要に応じストレージ容量を追加したり、削減したりすることが可能なストレージプール機能

これらの新規機能を容易にRAIDコントローラ内に実装することができます。

SBB2.0ベースのストレージ筐体は今年末頃からストレージユーザーは利用することが可能になるでしょう。また、Jasper ForestベースのRAIDコントローラも2011年中には市場に投入され、各ストレージベンダーの技術と知恵を凝らした新しいストレージ製品として紹介されてくると予想されます。

IDCやGartnerといった有名調査会社でのフォーキャストでも今後ますますユーザの非構造化データが増大し、これに対応してネットワークストレージは進化を遂げて行くことでしょう。