

CMP メインフレーム・ストレージソリューション

——新世代ストレージ活用事例

Summary of Storage Functionalities on CMP Mainframe System
——Utilization Example of New Generation of Storage Systems

望 月 健 彦

要 約 一世代前までディスク・ストレージに要求されていた機能は大容量化による低価格化とアクセス速度の高速化であった。しかしこのところの飛躍的なディスクデータ量の増加により、メインフレームとオープン系サーバとのストレージコンソリデーションはもとより、ストレージ単体でのソリューション機能の提供が注目されている。従来はテープ装置や LAN/WAN を使用しないと実施できなかった開発機業務の取り込みやメインフレームとオープン系システムとのデータ交換等がストレージを介して可能となった。本稿ではこのストレージの新機能の説明と CMP メインフレームシステムにおけるストレージソリューションについて特に OS 2200 環境での運用形態を中心に記述する。

Abstract Before the previous generation of computer systems, the major functional requirements for a disk storage system were the lower prices due to higher capacity and the higher performance. However, the rapid growth of storage capacity achieved on recent storage system has turned much attention to provision of the opportunity of storage consolidation, as well as solution capabilities based on unconsolidated storage system. For instance, the data transfer from a production host to developmental (secondary) hosts or the data exchange between mainframe systems and open system server, which had been performed conventionally via magnetic tape media or the network such as local area network or wide area network, can be performed through a storage system.

This paper describes the outline of new storage functionalities and solution software with CMP mainframe system, focusing on its operating mode under OS 2200 environment.

1. はじめに

コンピュータシステムの処理性能向上で、取扱うデータ量が増加してきたことに伴い、ディスク・ストレージは大容量化、高速化へと進歩してきたが、近年では単なるデータ格納機能だけではなく、ソリューションソフトウェアによる付加価値をストレージ単独で提供する方向に進んでいる。本稿ではこのストレージの付加価値機能の概要と CMP メインフレームシステムでの運用例を中心に現状のソリューションについて記述する。

2. SANARENA 2000/3000 シリーズ

SANARENA 2000/3000 シリーズは CMP メインフレームシリーズをはじめ、HMP IX/NX シリーズ汎用機や ES 7000、U 10000 等のエンタープライズサーバ向けに提供する高性能、高信頼性を有するストレージ装置で、単一筐体、省スペースタイプのモデルとディスク増設筐体による拡張性の高いモデルがある。

2.1 ハードウェア

2.1.1 アーキテクチャ

すべての HDU (ハードディスクユニット) は RAID (RAID 1 または RAID 5) で構成され、内部プロセッサやキャッシュメモリ、電源装置に至るまでリダンダンシがある。1 世代前の N 8800 制御装置はプロセッサ メモリ間がバス構造になっていたが、SANARENA 2000/3000 シリーズでは階層型スターネットアーキテクチャと呼ばれるクロスバススイッチ構造になり、バス・アービトレーションによる内部処理待ち時間が大幅に軽減され、高負荷状況でも高速処理が可能となった。また、HDU との I/F も SCSI (20 MB/Sec) から Fibre Channel (100 MB/Sec) へと性能向上したことで、装置内部処理についても格段にその処理量は増加しており、1 サブシステムに搭載できる HDU 数も倍増している。

図 1 に N 8800 制御装置のブロック図を、図 2 に SANARENA 2000/3000 シリーズのブロック図を示す。

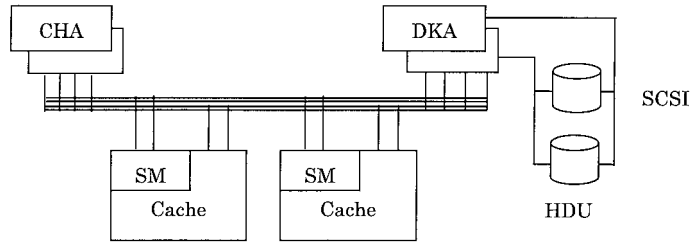
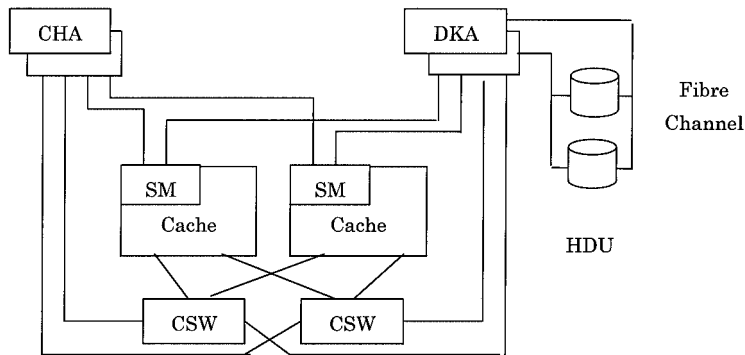


図 1 N 8800 従来形式共通バス機構



CHA: CHannel Adapter, DKA: Disk Adapter, SM: Shared Memory, CSW: Cache Switch

図 2 SANARENA 200/3000 新形式クロスバススイッチ機構

2.1.2 特 色

SANARENA 2000/3000 シリーズでは高性能なサーバ向けにチャンネルインタフェースは SBCON (Single Bytes CONNECTION) と Fibre Channel の二種類が用意されている。一つの筐体内にメインフレームシステムから SBCON Channel で接続したディスクと、ES 7000 等のサーバから Fibre Channel で接続したディスクとを共存することが可能となり、マルチプラットフォームでのストレージコンソリデーションを実現

している。勿論 SAN (Storage Area Network) にも対応している。SAN とは従来のサーバ・ストレージ間の固定経路 (Channel I/F) によるデータ転送方式から脱却し、HUB や SWITCH を用いてデータ転送経路に柔軟性を持たせるようにしたネットワーク通信方式のアーキテクチャである。この技術の根幹を支えるのが Fibre Channel である。その名前から誤解を招き易いが、SAN は OSI (Open System Interconnection) 参照モデルと同種の階層構造を持つネットワークモデルである。サーバとストレージ間は FC 4 (低階層とコマンドセットとの関連付け定義) にて SCSI Protocol で直接コネクションされているので Ethernet, TCP/IP によるデータ転送を余儀なくされるネットワーク・ドライブと比較すると格段に効率的である (この Fibre Channel HUB, SWITCH 群を Fibre Channel Topology と呼ぶ)。

SANARENA 2000/3000 シリーズは Fibre Channel I/F を実装した SAN 対応の製品であり、UNIX, Windows 2000 サーバに対しては SAN 環境での接続が可能である。現在、IX/NX シリーズの Fibre Channel は SAN 環境に対応していないが、近い将来に CS 7000 シリーズで対応予定である (図 3)。

なお、制御装置には SVP (Service Processor) と呼ばれる Note PC が搭載され、GUI によるディスク構成変更や HRC (リモートディスクミラー機能), HMRCF (レプリカ作成分割運用機能), UHMDE (高速データ交換機能) 等のソフトウェアを利用するための各種設定が可能である。

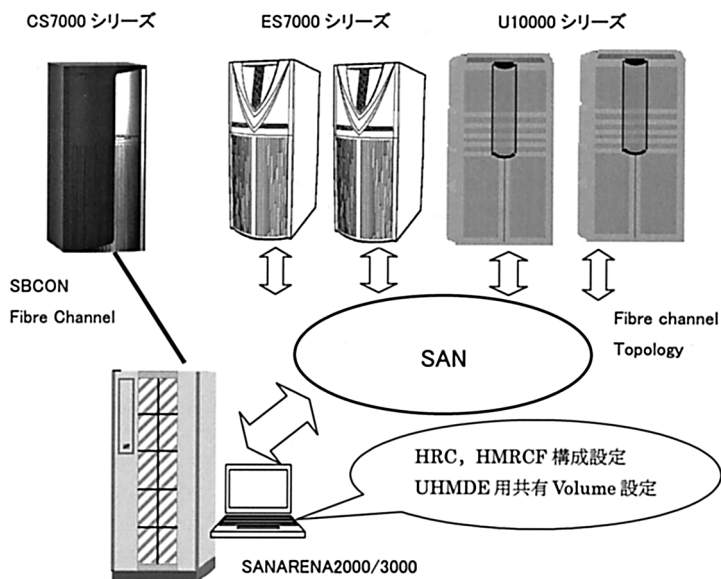


図 3 SAN 構成を使った多品種サーバ環境のストレージコンソリデーション

2.2 ソフトウェア

SANARENA 2000/3000 シリーズは種々のソリューションソフトウェアを有している。

2.2.1 HMRCF (レプリカ作成分割運用機能)

ある論理ボリューム (以降正ボリュームと呼ぶ) のレプリカを同一装置内の別の論

理ディスクに定義（以降副ボリュームと呼ぶ）して、常時更新データを副ボリュームへも反映する機能を HMRCF という。HMRCF は IO 命令とは非同期に HDU 群を制御するプロセッサ（DKA 内に搭載）単独で行われるので IO 命令の処理性能は低下しない。

通常副ボリュームはホストからアクセス不可状態に設定されているが、ある時点で正副ボリューム関係を一時停止し、ペア分割すれば副ボリュームがアクセス可能となり、他のホストでそのデータが利用可能となる。この状態で正ボリュームへデータ更新があった場合も、その更新部分のアドレスを装置内で保存しているので正副関係をペア再生成（リシンク）する場合に更新部だけのコピーで正副関係が復活する。図 4 に HMRCF の動作概略図を示す。

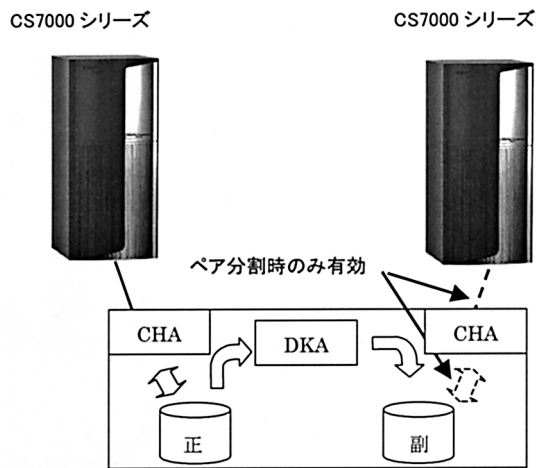


図 4 HMRCF の動作原理

2.2.2 UHMDE（高速データ交換機能）

SANARENA 2000/3000 シリーズの論理ボリュームは通常 SBCON または Fibre Channel のどちらか一方からしかアクセスできないが、双方からアクセス可能な論理ボリューム（以降共有ボリュームと呼ぶ）も定義することが可能である。この共有ボリュームはメインフレーム側（SBCON）からもオープン系サーバ（Fibre Channel）からもアクセス可能となる。

UHMDE ソフトウェアは UNIX（Solaris, HP UX）または Windows 上で動作するユーティリティであり、メインフレーム側で共有ボリューム上に生成されたシステム・データ・ファイル形式（以下 SDF 形式）を UNIX または Windows のファイル形式へ変換して、オープン系サーバ側でそのデータを即時に利用できる機能を提供するものである。これにより従来のネットワークを使ったファイル伝送方式（FTP）によるオープン系サーバ側での待ち時間がなくなり、高速処理が可能となる。

2.2.3 その他の機能

1) HRC（リモートディスクミラー機能）

高速通信回線を利用して制御装置単独で遠隔地へデータを送る機能を HRC と

呼び、オペレーティングシステムへの負荷もなく、ミドルソフトウェア開発も必要としない。

データセンタが被災したことを想定した場合、直前までのデータが別センタに保存されているため、復旧時間が大幅に短縮できるが、同期モード（リモート側のデータが送信完了を待ってI/Oに対して完了報告をする）の場合は適正な処理効率確保のため、構成の設計段階で、回線容量や回線遅延などの影響の見積もりを充分にしておくことが不可欠である。

また、非同期モード（メイン側のI/Oとリモート側のデータ更新が非同期）の場合は、同期モードより低速な回線容量で実現できるが、リモート側へは被災直近データは送信されないことが想定されるため、データ復旧手順を細部まで明確にしておく必要がある。

2) CVS（ボリューム容量可変機能）

SANARENA 2000/3000 シリーズは初期フォーマット時に Emulation Volume Type を必ず選択する事になっており、各々論理容量（最小 2.4 GB）が決まっている。CVSはこの指定された容量よりも小さい論理ボリュームを構成する機能で、後述の HMRCF ユーティリティで使用する制御デバイスなど、大容量を必要としないボリュームの構成時に利用すれば容量効率が上がる。

3) DCR（キャッシュ常駐機能）

超高速処理を期待されるアプリケーション用に Read Hit 率 100% を保証する機能である。DCR は CVS と併用すれば小容量の半導体ディスクを内部構成できるので、最低限必要なファイルだけをキャッシュ上に常駐させることができる。

4) SNMP（Simple Network Message Protocol）Trap 機能

SVP に SNMP Trap 機能を追加することで、SNMP マネージャに詳細情報を含めた障害報告を送信できる。ES 7000 を含め、すべての CMP メインフレームでは、ESR（Electronic Service for Enterprise Servers）に同期した SNMP マネージャ（ヴィーナス）で障害通報を実現している。図 5 に概略図を示す。

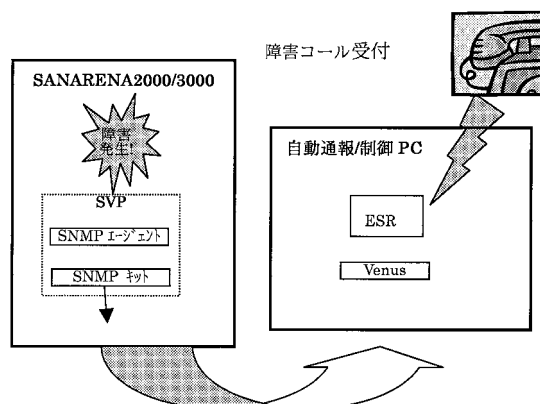


図 5 SANARENA 2000/3000 シリーズの障害通報のしくみ

3. HMRCF (レプリカ作成分割運用機能) の活用

3.1 OS 2200 環境での HMRCF 制御の自動化

HMRCF は前述の通り、SANARENA 2000/3000 装置内に副ボリュームを作成し、ペア分割やペア再生成等の操作を繰返し実行して、副ボリュームのレプリカデータを適宜使用する機能である。正副ペアが確立されている状態で、万一ホスト側から副ボリュームがアクセスされても、それを拒否する [Offline 状態] ことで正副データの一貫性が保たれている。よって、この副ボリュームをホストから使用するには正副ペア関係を一時停止させなければならない。また副ボリュームを使用した場合には、その期間に正ボリューム側に対して更新された差分データを副ボリューム側へコピーして、正副ペア関係を復帰させる指示も必要となる。

これらの操作を実施するため、GUI 形式の Windows AP が用意されているが、定常業務の運用では一連の操作を自動的に実行できることが望ましい。このため、OS 2200 環境にて通常の I/O と同様に、HMRCF 制御コマンドを SBCON I/F から発行し、正副ボリュームのペア操作を行うユーティリティを提供し、自動化を支援することとなった。(注：このユーティリティが必要なのは SBCON 接続時の HMRCF 機能であり、Fibre Channel 接続の場合は NT Node 上に SANARENA 2000/3000 シリーズで提供する HMRCF 制御ソフトウェアを使用してバッチランを作れば同様な機能が提供できる)

次にこのユーティリティの動作概要と仕様設計時に検討した OS 2200 環境への適用に関する留意点について述べる。

1) ユーティリティの概要

このユーティリティは制御するボリュームと HMRCF のコマンドをパラメータ化して、あらかじめ JCL 中に組み込んで設定できる仕様とした。

コマンドの種類と、それを実行した時の正副ボリュームの状態遷移を表 1 に示す。

表 1 ユーティリティで使用するコマンド

	コマンド	コマンドの意味
①	ペア形成 (BUILDPAIR)	論理ボリューム間のペア形成(初期コピー)を実施して 正副ボリュームを同期状態へ遷移させる
②	ペア分割 (SPLITPAIR)	ペア形成された論理ボリューム間の状態を分割状態に遷移させる。 分割状態の副ボリュームは使用可能となる。
③	ペア再生成 (REBUILDPAIR)	正副ボリュームの分割状態中に 正ボリュームへの更新があった場合、その差分だけを副ボリュームへコピーさせる。 完了後同期状態となる。
④	ペア状態表示 (INQUIRYPAIR)	正副ボリューム間の状態を取得して コンソール表示させる。
⑤	ペア解除 (DELETEPAIR)	正副ボリューム間のペア状態を解放して 通常状態に遷移させる。

2) 制御ユーティリティを使用時の自動化運用

2 ホストで制御ユーティリティを使用した環境での接続形態を図 6 に示す。図

6 中の制御デバイスは、制御ユーティリティからの制御コマンドを受領するためにあらかじめ装置内の特定ボリュームを設定したものである。

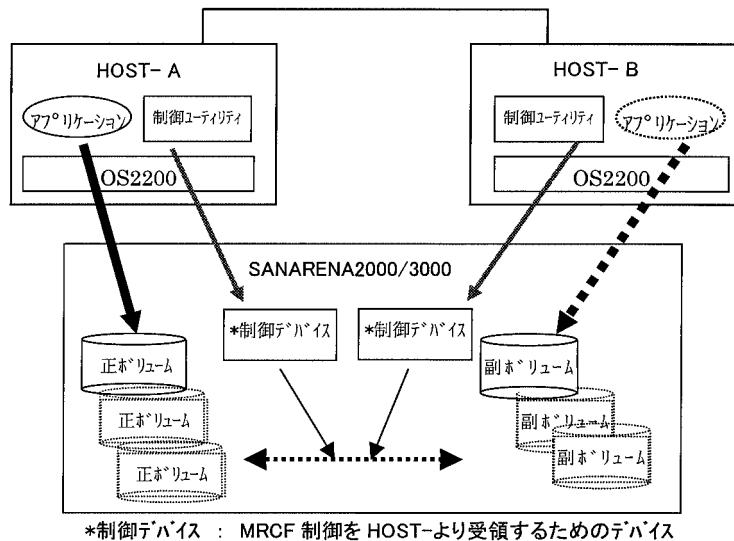


図 6 制御ユーティリティ環境での接続

このデバイスには実際にデータを書き込む訳ではないので、前述の CVS 機能を使用して小容量デバイスを定義し、領域を効率的に使用することができる。通常の Read/Write と同様のパスを経由して実行するので、パスの障害時においても交替パスにて運用の継続が可能であり、SVP 経由に比して可用性も向上している。

次に図 6 の構成で、制御ユーティリティの業務への取り込みの一例を示す。制御ユーティリティそのものには、それを実行しなかった他ホストへの情報伝達機能は有していないが、ホスト間通信ソフトウェアで、ペア操作が完了した時点でその完了を他ホストへ伝達する。

- ① HMRCF 環境設定時に該機能を使用するボリューム(群)のペアを HOST A (あるいは HOST B) にて形成する。
- ② HOST A でのアプリケーション終了後、またはアプリケーションからの正ボリュームの Update 静止時に、HOST A よりボリューム(群)のペア分割を行う。
(通信ソフトウェアを利用してペア分割完了を HOST B へ通知する)
- ③ HOST B に副ボリューム(群)を取り込み、HOST B にて別業務を実行する。
- ④ HOST B での業務終了後、副ボリューム(群)を正ボリューム(群)と同期すべくペア再生成を行う。

3) OS 2200 環境への適用のための考慮

- ① 正、副ボリュームの制御の自動化が目的であるため、制御するボリューム

と MRCF のコマンド等，パラメータ化して，あらかじめ JCL 中に組み込んで設定できる仕様とした。

- ② OS 2200 の品質を損なわないよう，Reserve 状態（OS からはアクセスできない）のデバイスに対して ADH（Arbitrated Disk Handler）を介して制御コマンドを発行する設計とした。
- ③ コマンドが発行された際に SANARENA 2000/3000 からエラーが返ってきた場合は，エラー詳細情報をコンソールに表示させる。また OS 2200 が実行後のタイムアウト等，入出力エラーを検出した場合も，それを知らせるメッセージをコンソール表示させる仕様とした。
- ④ コマンドデバイスは間違っても使用されないよう，PREP（OS 2200 システムフォーマット）を禁止する運用とした。

3.2 HMRCF（レプリカ作成分割運用機能）の実業務への応用

3.2.1 データベースファイル等の静的バックアップ時間の効率化

業後バッチ終了後に副ボリュームを切り離し，別業務機（ホスト）と接続してバックアップを取ることが可能となれば，本番業務機でバックアップを取る必要がなくなり月末/月初のバッチ時間の短縮に寄与できる。

3.2.2 本番業務機（ホスト）の障害保守支援

万一，本番業務機に障害が発生し，緊急に障害保守対応が必要な場合でも，障害発生時ただちに副ボリュームを切り離せば休日機で業務続行が可能となる。また，障害復旧作業終了後も，副ボリュームから正ボリュームに向けて逆方向の更新分コピー動作を実施すれば本番機の回復作業も短縮可能となり，システム全体の可用性向上に寄与できる。

3.2.3 24 時間/365 日運用への対応

コンビニエンスストアの ATM やデビットカードの普及により，金融系の顧客では 24 時間/365 日運用が期待されている。休日稼働機に本番稼働機と同様の環境を準備し，HMRCF 機能を使えば，休日稼働機でのレジストレーションを移行前にあらかじめ実行しておく事と同じ意味をもつ。レジストレーションに要する時間はシステム規模によって異なるが，事前に実施しておくことにより，休日移行時のオンライン停止時間は短縮され，24 時間/365 日運用へ一歩近づいたシステム構築が可能となる。次に HMRCF を使用したシステムの操作を順番に示す。

- 1) 本番業務機のオンライン切替時間の直前に副ボリュームの切り離しを行う。
- 2) 休日稼働機のセットアップを開始(副ボリュームのレジストレーション)する。
- 3) 副ボリューム切り離し後の更新トランザクションについて，本番機から休日機への反映処理を実施する。
- 4) 本番業務機のオンライン業務が停止した後，日替処理を行う。
- 5) 休日稼働機へオンライン業務を切り替える。
- 6) 休日稼働機のオンライン業務終了後，更新抑止処理を実施しセンターカットデータを保存して本番稼働機へ受け渡す。
- 7) 本番稼働機でのセンターカットデータ反映処理終了後，ただちにオンライン業務再開と同時に副ボリュームのペア再生成を指示する。

- 8) これらの動作にて、正ボリュームから副ボリュームへの更新分コピーが開始され、最終的に直前の正ボリュームにセンターカットデータを反映した状態となる。

4. UHMDE (高速データ交換機能) の概要と運用

4.1 UHMDE (高速データ交換機能) ソフトウェア概要

UHMDE ソフトウェアは、前述のとおりオープンサーバ上で動作するソフトウェアであり、SANARENA 2000/3000 装置内に設定された共有ボリュームを介して OS 2200 ホストとのデータ交換を提供する機能である。これは共有ボリュームをファイルシステムで定義されていない raw デバイス (OS のファイルシステムを持たないデバイス) としてハンドリングするので、OS 2200 ファイルシステムを維持したままデータ交換を行うことができる。階層的にテキストデータ処理関数、関数処理を経由してファイルの読み出し/書き込みを行うテキストファイル交換ユーティリティ、および共有ボリューム内のファイル情報をもとにファイル名、サイズ、ロケーションを制御する管理用ユーティリティの三つの機能から構成されている。全体の階層構成を図 7 に示す。

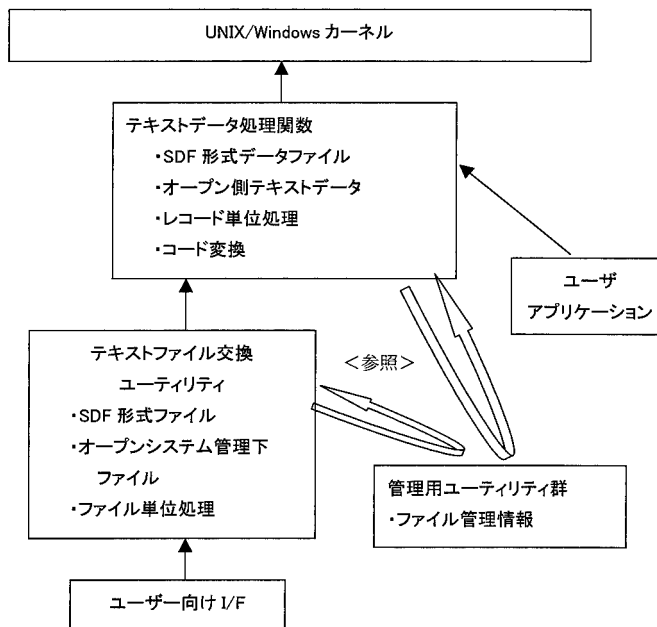


図 7 UHMDE ソフトウェアの階層構造図

1) テキストデータ処理関数

テキストデータ処理関数は主に共有ボリュームへの READ/WRITE 及び日本語 CODE 変換を行うモジュールであるが、OS 2200 とのデータ交換を実現するにあたって以下の考慮が施されている。

① OS 2200 固有のレコードフォーマット変換処理

SANARENA 2000/3000 の 1 論理レコードは 512 Bytes 長であるが、OS 2200

側からは 112 Word (504 Bytes) フォーマットされているため、残り 8 Bytes にはデータ保証用の FLRC(データ保証付加コード 2 Bytes)と ZERO Padding (6 Bytes) が埋め込まれている。テキストデータ処理関数は、読み出し時には、512 Bytes の Raw Data からこの 8 Bytes を除去する。また、書き込み時には FLRC と ZERO Padding データを生成し付加することにより、オープンサーバにて書き込んだデータを OS 2200 側でエラーなしに READ できる。

② ファイルオープン処理

OS 2200 ホストでの Pack ID とファイル名 (QUALIFIER_FILENAME に変換) を引数に、OS 2200 のファイル構造に基づいた検索を行うことにより、オープンサーバ上で OS 2200 のファイルの検索を可能にしている。

③ データ変換処理

オープンサーバ上で扱うデータは 8 bit ASCII 文字であるが OS 2200 では 9 bit ASCII 文字を使用しているため、読み出し時には先頭 1 bit を除去してオープンサーバで扱える ASCII 文字に変換し、書き込み時には 1 bit (=0) を付加して OS 2200 で扱える ASCII 文字に変換している。

2) テキストファイル交換ユーティリティ

テキストデータ処理関数を経由して、オープン側に対してファイル単位での読み込み/書き込みをユーザに提供するユーティリティである。

① オープン向けテキストファイル変換

SDF 形式ファイルをオープン側テキストファイルに変換して指定されたファイル(共有ボリューム以外のオープン側ファイルシステム管理下のファイル)に出力する。OS 2200 ホストでの Pack ID とファイル名、変換する日本語コードの選択及び出力するオープン側のファイル名を引数にして呼び出す。

② OS 2200 向け SDF 形式ファイル変換

オープン側テキストファイルを SDF 未定義ファイルに変換して指定されたファイル(共有ボリューム)に書き込む。

3) 管理用ユーティリティ

共有ボリューム上の OS 2200 ファイル管理情報をオープン側ファイルシステム管理下のファイルに作成し、ファイル管理情報を表示するもので、ファイル情報を初期化するコマンド、ボリューム情報を表示するコマンド、ファイル情報を表示するコマンドが用意されている。

4.2 UHMDE (高速データ交換機能) の運用形態

OS 2200 ホストと Solaris サーバでの UHMDE 運用方法を紹介する。この運用ではホストとサーバのジョブ連携に 2200 側 DDP 1100 (DDP FJT) のジョブ転送機能及び IS プロダクト (IS USF) のジョブ・サービス (2200 に対してジョブの実行・取り消しを要求)、ファイル・サービス (2200 に対してファイルの作成・複写・削除を要求) を利用している。

- 1) OS 2200 ホスト上で JCL を実行して、Solaris サーバに実行させたいシェルを作成し、そのシェルを転送して実行させる。
- 2) 最初にコード変換させたいファイルを OS 2200 ホスト専用ボリュームから共

有ボリュームにコピーし、コード変換後のファイルを UNIX 専用ボリューム上に作る。

- ① 共有ボリューム上にファイルをカタログする。
 - ② OS 2200 ホスト専用ボリュームから共有ボリュームへファイルをコピーする。
 - ③ 共有ボリューム上のファイルを UNIX テキストファイルにコード変換して UNIX 専用ボリュームに出力する。
 - ④ 共有ボリューム上のカタログしたファイルを削除する。
- 3) Solaris 形式テキストファイルでバッチジョブ等を実施する。
- 4) 加工したデータをコード変換し共有ボリュームを経由して、OS 2200 ホスト専用ボリュームへ戻す。
- ① 共有ボリューム上にファイルをカタログする。
 - ② 加工したデータの格納してある UNIX 専用ボリューム上のテキストファイルをコード変換して共有ボリューム上のファイルにコピーする。
 - ③ 共有ボリューム上のファイルを OS 2200 ホスト専用ボリュームへコピーする。
 - ④ 共有ボリューム上のカタログしたファイルを削除する。

5. おわりに

今後 SANARENA 2000/3000 ソリューションソフトウェアを CS 7000 シリーズに適応していく上で、運用面での更なる仕組み作りが課題である。

例えば、現行の UHMDE (高速データ交換機能) には共有ボリュームを OS 2200 側とオープン側とで排他制御する仕組みがない。次期リリース版では UHMDE 専用ファイル上の特定レコードに OS 2200 側/オープン側 JOB 連携情報を共有し、排他制御ができる仕組みが加わり、DataExtractor/IX との連携もよりスムーズになる。

また、Fibre Channel 接続の場合、HMRCF (レプリカ作成・分割) 制御は OS 2200 側のユーティリティから NT node 上で動作するユーティリティに変わることになるので、@Ntshot プロセッサ、Ntshot サービスで OS 2200 側本番業務切替時の NT node との JOB 連携を提供していかなばならない。これらの仕組みは運用形態に左右されることが多いので汎用的なものにすることは難しいが、ある程度方式を決めた上で、更にシームレスな業務切替手順を提案できれば、真の 24 時間/365 日運用ソリューションとして利用して頂けるのではないかと考える。

執筆者紹介 望月 健彦 (Takehiko Mochizuki)

1982年神戸大学工学部卒業。同年日本ユニシス(株)入社。1100/2200シリーズ向けストレージシステムのファームウェア開発設計および保守業務を経て、各種ストレージ・ソリューションのインプリメンテーション業務に従事。現在ブロードバンドビジネス事業部基盤統合コンサルティング室に所属。ISO/SC 25 WG 4 小委員会委員