計算機アーキテクチャ 41－3
（1981. 4. 15）

VOS3疎結合マルチプロセサシステムにおけるファイルシステム

柴宮実　松井洋　東常修也　萩原亘喬
（（株）日立製作所　ソフトウェア工場）

1. 緒言

電子計算機システムが社会活動の中で占める役割は拡大の一途をたどっており，経済性の高い大規模システムの実現が要求されている。この要求を実現するため，HITAC Mシリーズの最上位オペレーティングシステムであるVOS3（Virtual Operating System 3）の下で疎結合マルチプロセサ（Loosely Coupled Multi Processor：以下LCMPと略す）システムを開発した。これは各々個別のOSを持つ複数の計算機をCTCA（Channel to Channel Adaptor）および共用DASD（Direct Access Storage Device）を介して結合し，全体として1システムイメージの計算機システムを実現するものである。

1システムイメージを実現する上でのポイントの1つは，LCMPシステム下のファイルシステムである。複数のOSの下で一元的なファイル管理が行われ，利用者がプロセサを意識せずにファイル操作のできることが1システムイメージの実現のための必要条件である。

本報告では，VOS3のファイルシステムの主要な構成要素について，概要と，LCMPシステム下での一元化方式を述べる。

2. LCMPにおけるファイルシステムの概要

2.1 構成

LCMPシステムでは各プロセサは各々のOSを持ち，単独で動作できるようにしている。このためOSの動作に必要な各種のシステムデータセット（システムのライブラリや補助記憶用のデータセットなど）はプロセサ個有の装置上におく。一方ジョブ入出力のためのスプールデータセット，ユーザデータセット（センタ提供のデータベースなども含む）は各プロセサからアクセス可能な共用DASD上に配置する。ユーザから見てファイルシステムが1システムに見えるように，カタログやSAFE登録簿（システムの利用者情報や機密保護情報の管理簿）も共用DASD上に置き，全プロセサから共用する仕掛としている。MSS（Mass Storage System）は通常ユーザデータセ

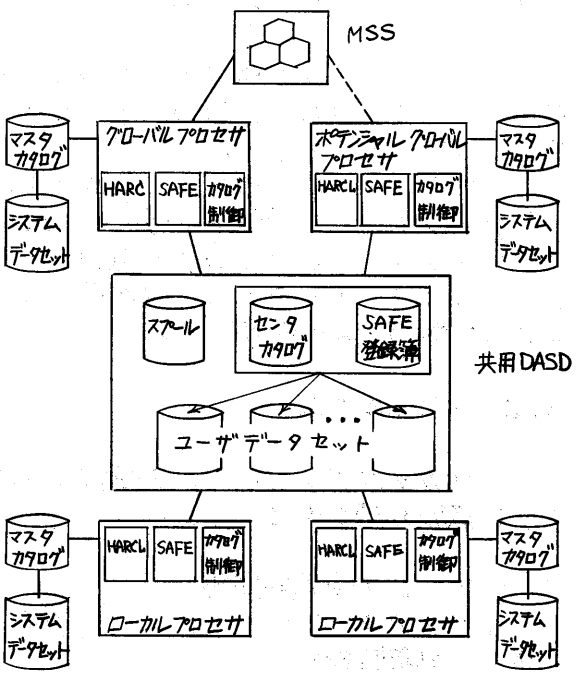


図2.1　LCMP下のファイルシステムの概要

ットのアーカイブ，バックアップ用に使う。アーカイブ，バックアップはグローバルプロセサが全系を管理する仕掛としている。

2.2 ユーザデータセットの排他制御

単一OSのシステムの場合，1つのデータセットに対する複数ジョブの競合についてはOSが排他制御し，データセットの破壊などが起らぬようにしている。LCMPの場合は異なるOSの下で実行されるジョブが，1つのユーザデータセットを競合して使用することがあるため，複数OS間での排他制御が必要である。VOS3 LCMPシステムではシステム間ENQ機能をサポートし，LCMPシステム全体の中でデータセット排他制御を実現している。
これによってジョブがどのプロセサで実行されてもデータセットの排他制御が行われ，単一OSのシステムと同様にジョブを実行できることになる。

このようにLCMPシステム下のファイルシステム一元化のために種々のサポートを行っているが，ここでは主たる役割を演じているセンタカタログ，SAFE（Security Administration FEature），HARC（Hierarchical storage Archive and Retrieve Controller）をとり上げ，特徴と一元化の方式を以下に述べる。

3. センタカタログ

3.1 センタカタログの特徴

LCMPシステムを構成する各OSは，OS個有のシステムデータセットの登録用にマスタカタログを持つ。センタカタログはLCMPシステム全体で1つ存在し，システム内の全てのユーザデータセットを管理する。
センタカタログは単一OSシステムにおけるマスタカタログと同様に扱われ，ユーザデータセットを登録する際に，明示的にセンタカタログを指定する必要はない。データセットをカタログへ登録する時，ユーザデータセットはセンタカタログへ，システムデータセットは登録するジョブを実行するプロセサのマスタカタログへ登録するよう制御する。

3.2 一元化の方式

センタカタログはLCMPを構成する各プロセサから共用する。このため次のような機能を持たせている。

（1）プロセサ間同期制御
センタカタログ内のエントリの参照，更新，追加，削除について各プロセサ間での同期制御を行う。
（2）VSAM構造の採用
複数プロセサからの要求に対しハードウェア上のネックをつくらぬよう，入

出力時のコントローラ，装置の専有時間を短かくできるVSAM構造を採用している。

(3) マルチセンタカタログ

センタカタログを分割して複数のDASDに負荷分散できるようマルチセンタカタログ機能をもつ。

(4) カタログエントリ常駐化機能

DASDに対する負荷低減のため，カタログエントリのメモリ常駐化機能をもつ。

カタログ機能はTSS処理において主要な役割を持っており，その性能はTSS応答時間に大きな影響を与える。LCMPシステムではCPU能力が従来に較べて大幅に増大しており，処理量の増加にともないカタログへの負荷も増大する。これを解決するには負荷の低減と分散が重要な課題である。
VOS3 LCMPシステムではこれをマルチセンタカタログとカタログエントリ常駐化によって対処するよう設計した。以下に詳細を記す。

3.3 マルチセンタカタログ

マルチセンタカタログでは複数個のカタログデータセットによって1つのセンタカタログを形成する。各々のカタログデータセットはインデクスを持つなど構造上それ自身で閉じている。図3.1に示すように各カタログデータセットにはキーレンジが与えられ，キーの範囲に含まれるエントリを登録する。キーレンジは各プロセサ毎にメモリ上で記憶してあり，エントリの参照，更新などの際メモリ上をスキャンし対応するカタログデータセットを求める。複数のカタログデータセットはカタログ制御プログラムが一括管理し，カタログへの要求を行うプログラム側は標準インタフェースでアクセスでき，カタログデータセットが複数個あることを意識しなくてよい。

キーレンジはデータセット名の先頭修飾子（TSS，オープンバッチの場合はユーザ登録名がそれに当る）の中の先頭からの位置と文字数をキーとし設定する。ユーザ登録名の体系が図3.2のような場合，ユーザ登録名の最終1文字をキーに設定するとバランスよく負荷の分散を図れることになる。

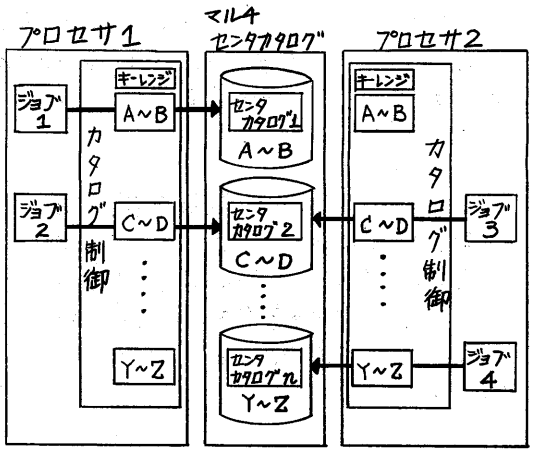


図3.1 マルチセンタカタログ

ユーザ登録名 (X：英字, n：数字) Xnnnnn（最後のnがキー）

キーレンジ	データセット名先頭修飾子(ユーザ登録名)
0～1	A0000, A0001, A0010, ···, Z9991
2～3	A0002, A0003, A0012, ···, Z9993
4～5	A0004, A0005, A0014, ···, Z9995
6～7	A0006, A0007, A0016, ···, Z9997
8～9	A0008, A0009, A0018, ···, Z9999

図3.2 キーレンジ設定の例

3.4 カタログエントリ常駐化

カタログエントリ常駐化機能はカタログアクセスの入出力回数の絶対値を削減するもので，次の2つから成る。

(1) センタ固定データセットのカタログ情報の常駐化

センタ運営上使用頻度の高いシステムデータセット（言語プロセサのライブラリ，HELP，プロシジャライブラリ等）やセンタ独自のライブラリ(センタコマンドライブラリ等）のカタログ情報をメモリ内に常駐化する。対象となるエントリはマスタカタログ，センタカタログのいずれに登録されていてもよい。常駐化を指定されたデータセットのカタログ情報はVOS3のIPL（Initial Program Load）時メモリ上にとり込まれ，以降メモリ内の情報だけを使って処理される。

(2) LOGON時のポインタ情報常駐化

TSSジョブの開始時，該当ユーザのユーザ登録名を先頭修飾子として持つデータセットのカタログ情報へのポインタをメモリ内に常駐化する。TSSセッションの中でカタログを使用する処理は，このメモリ内のポインタを用いてアクセスすることで入出力回数を削減する。常駐化するポインタの数はセンタで設定し，LOGON時にその数のポインタをメモリ内へとり込む。データセットを新たにカタログへ登録した場合もポインタをメモリ上で記憶する。メモリ上のポインタの数が指定した値を越える場合は参照頻度の低いポインタとおきかえを行う。この機能はTSSコマンドバッチ実行機能に対しても有効である。

(3) 効果

カタログエントリ常駐化機能を適用した場合のカタログに対する入出力回数の減少効果を図3.3に示す。図では本機能を用いない場合の入出力回数を1とした時の相対値であらわしている。

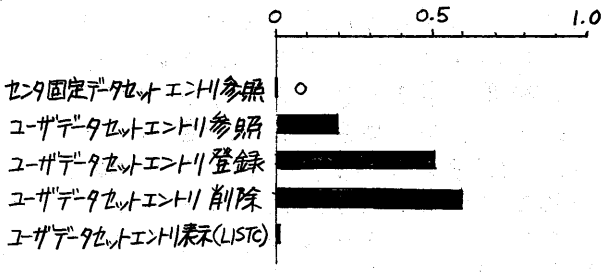


図3.3 入出力回数の減少効果

4. SAFE（Security Administration FEature）

4.1 SAFEの特徴

SAFEはユーザ登録情報，予実算管理情報，データセットやボリュームの機密保護情報等の計算機システムの保全性，機密保護に関する管理情報を1つの登録簿（これをSAFE登録簿と呼ぶ）の中で集中管理し，運用の一元化を図ることを目的としたプログラムである。SAFE登録簿はLCMPシステムの下で1つ存在し，各プロセサから共用され，LCMPシステムの運用の一元化を行っている。SAFEの機能は表4.1に示すとおり大別して4つの機能から成る。

表4.1　SAFEの機能

項番	分類	内容
1	ユーザ管理	• TSS処理，オープンバッチ処理におけるユーザ登録名とユーザパスワードによる識別と検証。 • グループ管理
2	予実算管理	• ユーザ，グループ対応の予実算管理 • データセット上限値管理
3	データセットアクセス制御	• データセットアクセス権限による機密保護 • コマンドによる共用制御
4	ボリュームアクセス制御	• MTボリューム上のデータセット機密保護 • DASDボリューム単位の機密保護

4.2　データセットのアクセス権限

ユーザごとに該当データセットに対してどのようなレベルでアクセスできるかを限定したものをアクセス権限と呼ぶ。SAFEが用意しているアクセス権限を表4.2に示す。

表4.2　アクセス権限

機能＼アクセス権限	MASTER	UPDATE	APPEND	READ	EXECUTE
データセット領域の解放	○	×	×	×	×
データセット名称の変更	○	×	×	×	×
SAFE登録簿中のデータセットコンポーネントの内容変更（DSALTER）	○	×	×	×	×
他グループ，ユーザへの共用許可（PERMIT，ENHIBIT）	○	×	×	×	×
データセットへの出力	○	○	×	×	×
区分データセットへの新メンバの追加	○	○	○	×	×
データセットからの入力	○	○	○	○	×
ロードモジュールライブラリ中のプログラムの実行	○	○	○	○	○

APPEND属性とは区分データセットの既存のメンバの削除，更新は禁止するが，新たなメンバの追加を許すアクセス権限である。EXECUTE属性はロードモジュールライブラリに適用するもので，入力は禁止するがプログラムの実行は許すアクセス権限でセンタのライブラリ等に適用する。

データセットを作成するとそのユーザは自動的にMASTER権限を持つ。他のユーザは通常そのデータセットに対するアクセス権限を与えられていないため，共用を許可されるまでアクセスできない。

データセットの所有者は自分自身のアクセス権限を抑止（ロックをかける）することができる。これは誤ったデータ破壊を防止することを目的としており，DSALTERコマンドで行う。

4.3 データセットの共用

データセットの共用を行う場合は，通常PERMITコマンドで所有者が他のユーザに共用を許可する。許可されたユーザは完全データセット名（先頭修飾子を含む名称）を指定してアクセスする。

共用者が自分の名称をつけてアクセスしたい場合は，SHAREコマンドによって名称を付けることができる。

PERMITしたデータセットはSAFE登録簿の中のアクセスリストに，共用を許可するユーザ登録名とアクセス権限が登録される。（図4.1 参照）

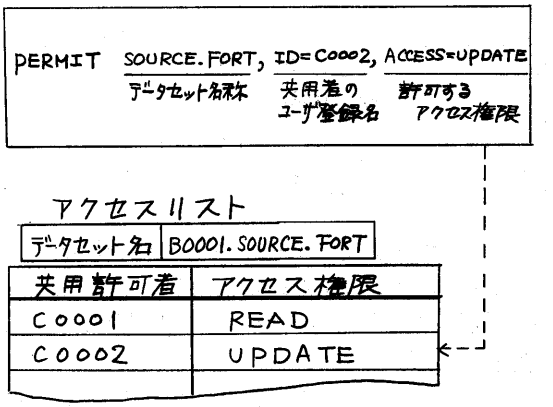


図4.1 データセットの共用

4.4 LCMPシステムにおける一元化

SAFE登録簿はセンタカタログと同様，共用DASD上におき全プロセサから共用する。SAFE登録簿の中にはユーザ管理情報，予実算管理情報，機密保護情報があり，特にジョブの開始，終了時のアクセス頻度が高い。これらの条件に対処できるようSAFEは次の方式をとっている。

(1) 一元管理とプロセサ間同期制御

ユーザ検証はグローバルプロセサで一元管理するようプロセサ間連絡を行う。また機密保護情報については複数プロセサ間の同期制御を行う。

(2) マルチSAFE登録簿

装置に対する負荷分散を行えるよう，SAFE登録簿を複数装置に分散する機能を持つ。

(3) インデクス常駐化機能

DASDに対する負荷低減のため，SAFE登録簿中のインデクス部分を，各OSのメモリ上に常駐化する機能を持つ。インデクス部分の変更の際には，他プロセサに連絡し，矛盾が発生しないようにしてある。

5. HARC (Hierarchical storage Archive and Retrieve Controller)

HARCはDASDとMSSを組み合わせて記憶の階層構造を形成し，外部記憶装置全体の使用効率を向上させ，データセット管理に行う事を目的としたプログラムである。HARCの機能はDASD上の長期間使用しないデータセットをMSSへ保存するアーカイブ機能，MSS中に保存してあるデータセットをDASD上に戻すリコール機能，DASD上のデータセットのバックアップ機能およびリカバリ機能から成る。

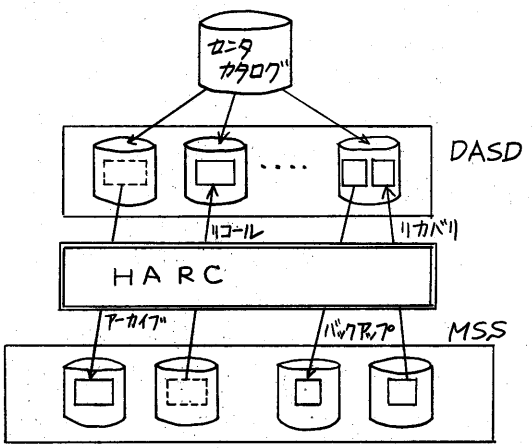


図5.1 HARCの概念図

5.1 アーカイブ／リコール機能

HARCを組込んだシステムでは，データセットはオープンされる都度アクセス日付（最終参照日付）が記録され，この日からの経過日数（未参照日数）によってアーカイブが行われる。
アーカイブにはDASD上の空きスペースの確保と長期間未参照のデータセットの保存という2つの目的があるが，HARCでは〝スペース量アーカイブ〟と〝未参照日数アーカイブ〟と呼ぶ機能でこれらを実現する。
前者はユーザデータセット格納用のDASDにスペース利用率制御ポイント(%)を設定して，スペース使用率がこの値に達するまで未参照日数の大きいデータセットからアーカイブする。後者は日数を指定し，未参照日数が指定日数より大きいデータセットをアーカイブするものである。

リコール動作はアーカイブしてあるデータセットを使用する時行われる。ユーザジョブではデータセットがアーカイブされていることを意識しなくてよい仕掛にしてある。つまりジョブ実行時に使用するデータセットがMSS上にアーカイブされていれば，自動的にリコールされる。

これらの操作はエンドユーザが意識的に起動することもできるようになっており，HARCの各種機能に対するTSSコマンドを用意してある。

5.2 バックアップ／リカバリ機能

DASD上のユーザデータセットをMSS上にコピー管理する機能がHARCのバックアップ機能である。このバックアップ機能では，データセットを単位とする差分バックアップ法を用いている。データセットのオープンの際，内容変更の有無が記録され，DASD上のデータセットのうち変更記録のあるものだけがMSS上へコピーされる。MSS上のコピーは世代管理が可能であり，指定した世代数分のコピーを保存することができる。

5.3 HARCL (HARC for Local processor)

LCMPシステムにおける記憶階層の一元管理のため，ポテンシャルグローバルプロセサ，ローカルプロセサではHARCLが動作する。グローバルプロセサにはHARCがあり，これが全体を管理する。
ローカルプロセサで実行するジョブからHARCに対する要求が発生すると，HARCLがCTCA経由でグローバルプロセサのHARCに連絡を行う。グローバルプロセサでデータ移動を行った後，結果をローカルプロセサに連絡する。
またグローバルプロセサを切りかえる際に備えて，HARCLはHARCの機能を代行できるように設計してある。

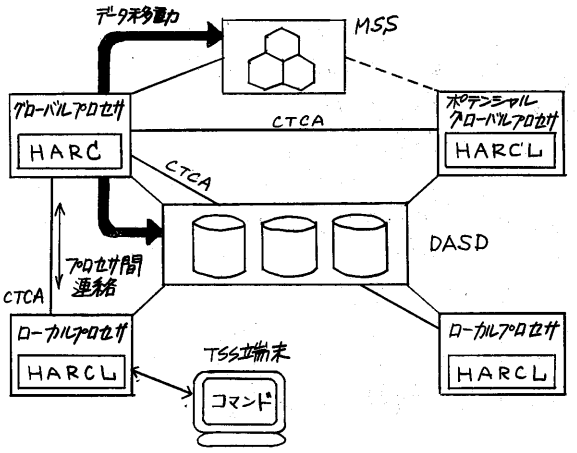


図5.2 LCMPにおけるHARC

6. 結言

以上VOS3 LCMPシステム下のファイルシステムの主要な構成要素と一元化の方式について説明した。VOS3 LCMPシステムは昭和54年10月北海道大学大型計算機センターで稼動を開始し，昭和56年1月からは東京大学大型計算機センターでM200H8台，64MBのシステムが動いており，これらのプログラムはその下で稼動中である。

開発にあたり終始適切な御指導，御協力をいただいた多くの関係各位に対し，深く謝意を表わす次第である。

7. 参考文献

(1) 風間ほか：疎結合マルチプロセサシステム（VOS3．JSS4）の方式，電子通信学会 EC80－80 1981年3月20日

(2) 宮入ほか：疎結合マルチプロセサシステムのジョブ制御方式，情報処理学会 計算機アーキテクチャ研究会 1981年4月15日