

大形システム用基本ソフトウェアの技術動向

Trends of the Basic Software Technology for Large Scale Computer Systems

HITAC Mシリーズは、開発以来約15年の歳月を経ようとしている。ここでは、Mシリーズの最上位計算機の基本OS (Operating System) とデータベースの動向について概観する。基本ソフトウェアは、計算機アーキテクチャの進歩と顧客のデータ処理のニーズに合わせた進歩発展を果たしていく責務がある。正にこの15年はOS、データベース、アーキテクチャがそれぞれ劇的な発展を成し遂げた時代にある。そして1990年は16 Tバイトの大容量記憶を提供する新アーキテクチャを実現し、新たな時代に向けたインフラストラクチャが構築された歴史に残る年である。

大形計算機用OSでは、(1) 大容量仮想記憶を利用した入出力削減による高性能化、(2) 高多重プロセッサを効率よく制御する機構と、(3) 数量ともに拡大を続けるファイルの運用機能の開発などを行ってきた。一方、データベース管理システムは、データベース自身の機能開発とともに、大規模・広域化に対応した複合計算機システムの諸機能の開発を行ってきた。

吉澤康文* *Yasufumi Yoshizawa*
米田 茂** *Shigeru Yoneda*
柴宮 実*** *Minoru Shibamiya*
住吉孝史*** *Takashi Sumiyoshi*

1 緒 言

基本ソフトウェアの使命は、計算機システムの持つポテンシャルを可能な限り引き出し、社会活動の多くの分野に広く提供することにある。つまり、計算機が保有する性能をユーザーに十分享受してもらうばかりでなく、システムとしての信頼性の確保を図り、またシステムを容易に構築し運用できる機能を提供しなければならない。

大形計算機システムは、企業や組織の中で情報中枢としての機能を果たしてきた。そこで、上記の目的を担う基本ソフトウェアとしてのOS (Operating System) とデータベースシステムDBMS (Data Base Management System) の動向を以下に述べる。

2 アドレス空間の拡張

2.1 多重仮想記憶

(1) 多重プログラミングの推進

プロセッサの能力を最大限に引き出すために、多重プログラミングの技術が1960年代中ごろのOSから一般的になった。当時は、主記憶が高価で貴重な資源であり、OSは限られた主記憶の中に多くのジョブを詰め込むさまざまな工夫を行ってきた¹⁴⁾。主記憶のページ化やページングを用いた仮想記憶の出現は、その発展した技術の一例である。

大形計算機の分野では、1970年代の前半から仮想記憶方式

が主流となってきた。HITAC Mシリーズ(以下、Mシリーズと略す。)では1975年にVOS 2 (Virtual-storage Operating System 2)を開発し提供してきた。Mシリーズ初期のアーキテクチャは24ビットアドレッシングであり、プログラムからアクセスできるアドレスは16 Mバイト (Mバイト: 10^6) である。VOS 2 では、アーキテクチャの限界である16 Mバイトの論理アドレスを提供する単一仮想記憶方式を実現している。しかし、ジョブ間のデータ機密保護に用いるページ単位のストレーキキーは4ビットであり、このため、ジョブ多重度の限界は16である。実際はOSがキー0を利用するので、15個までのジョブしか実行できないという制約があった。

ジョブ領域の拡大とプロセッサ能力の向上が進む中で、単一仮想記憶方式はジョブ多重度が15という限界から、大形計算機システムのユーザーには不十分なジョブ多重度となってきた。

(2) 単一仮想記憶から多重仮想記憶へ

そこで、各ジョブに16 Mバイトまでの論理的なアドレス空間を提供する多重仮想記憶方式のOSが必要とされるようになる。Mシリーズでは、1977年にVOS 3が多重仮想記憶方式のOSとして開発された。単一仮想記憶方式と多重仮想記憶方式の記憶配置を図1に示す。

この時点の大形計算機であるM180では、利用できる主記憶

* 日立製作所 システム開発研究所 工学博士 ** 日立製作所 システム開発研究所 *** 日立製作所 ソフトウェア開発本部

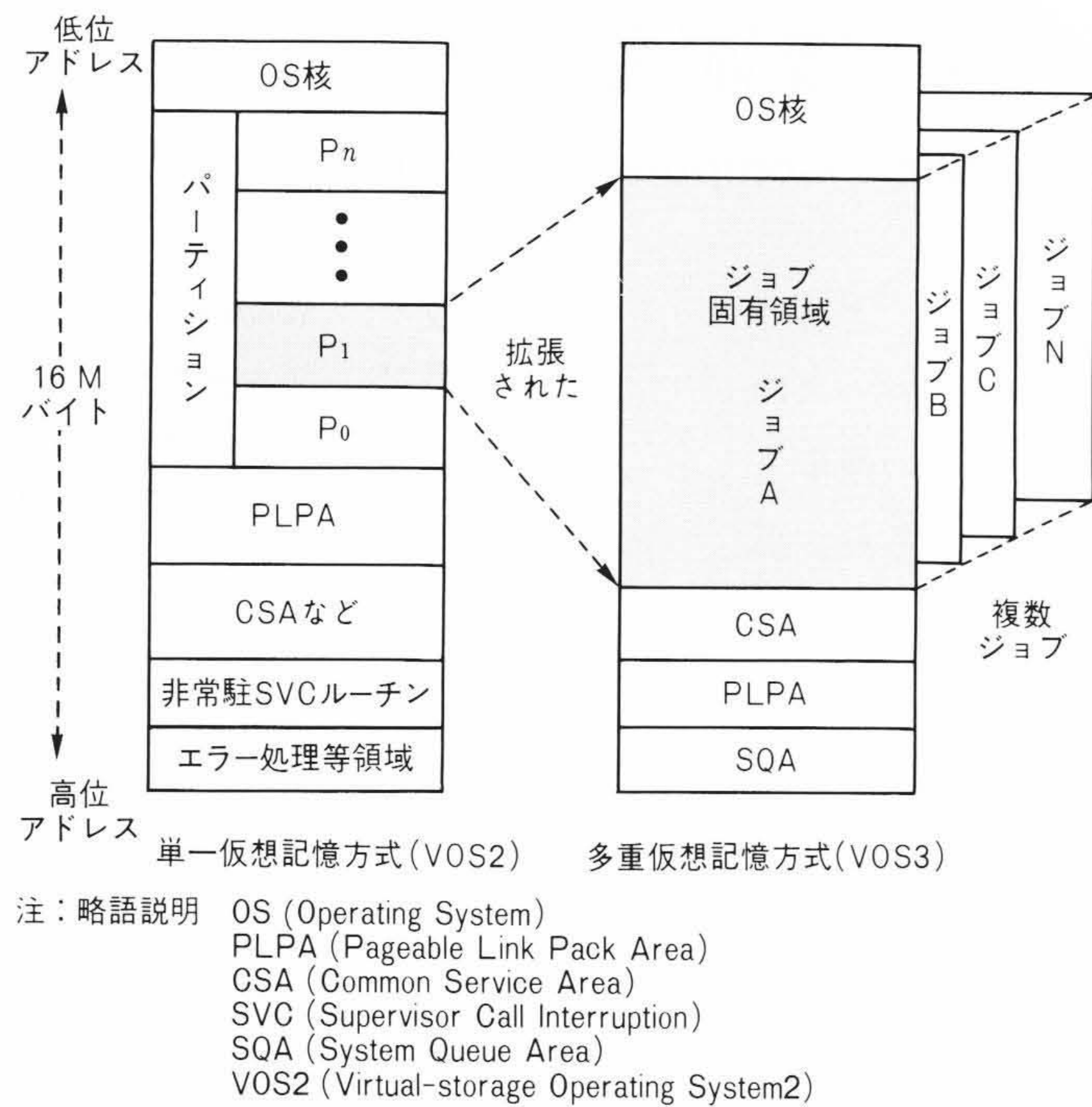


図1 HITAC Mシリーズでの仮想記憶の拡大 プログラム・データの大容量化に伴い飛躍的な仮想記憶の拡大を図り、大量データ処理を可能にしている。

容量は8 Mバイトが限界であり、多重仮想記憶方式を有効に活用するには主記憶が少なく、システムのボトルネックとなることが多かった。特に、TSSのような対話処理での性能劣化問題が生じた^{2),14)}。

しかし、1980年代に入ると半導体技術の進歩によってこの状況は一変し、利用できる主記憶容量が大幅に増大する。Mシリーズでの最上位機の最大実装可能な主記憶容量を図2に示す。各プログラムからアクセスできる論理的なアドレス範囲は16 Mバイトであるが、M200Hでは64 Mバイトの主記憶が利用可能となった。つまり、プログラムからアクセスできる論理的なアドレス以上に主記憶装置の容量が大きくなってきたのである。このような背景から、ジョブ多重度を論理的に無限に拡大できる多重仮想記憶方式が大形計算機システムの主流として定着するようになった。

(3) 実記憶拡張

個々のジョブは24ビットアドレッシングの制約から16 Mバイトまでの論理アドレスしか利用できないが、計算機システム全体では物理的な主記憶を16 Mバイト以上利用できるようになった。このため、アドレス変換機構の拡張によって64 Mバイトまでの主記憶を利用可能とした。この機能をサポートしたのがVOS 3/SP21 (VOS 3/System Product 21)の実記憶拡張機能である。

2.2 拡張アドレッシング

(1) ジョブ固有領域の圧迫

多重仮想記憶方式により、各ジョブが独立した16 Mバイト

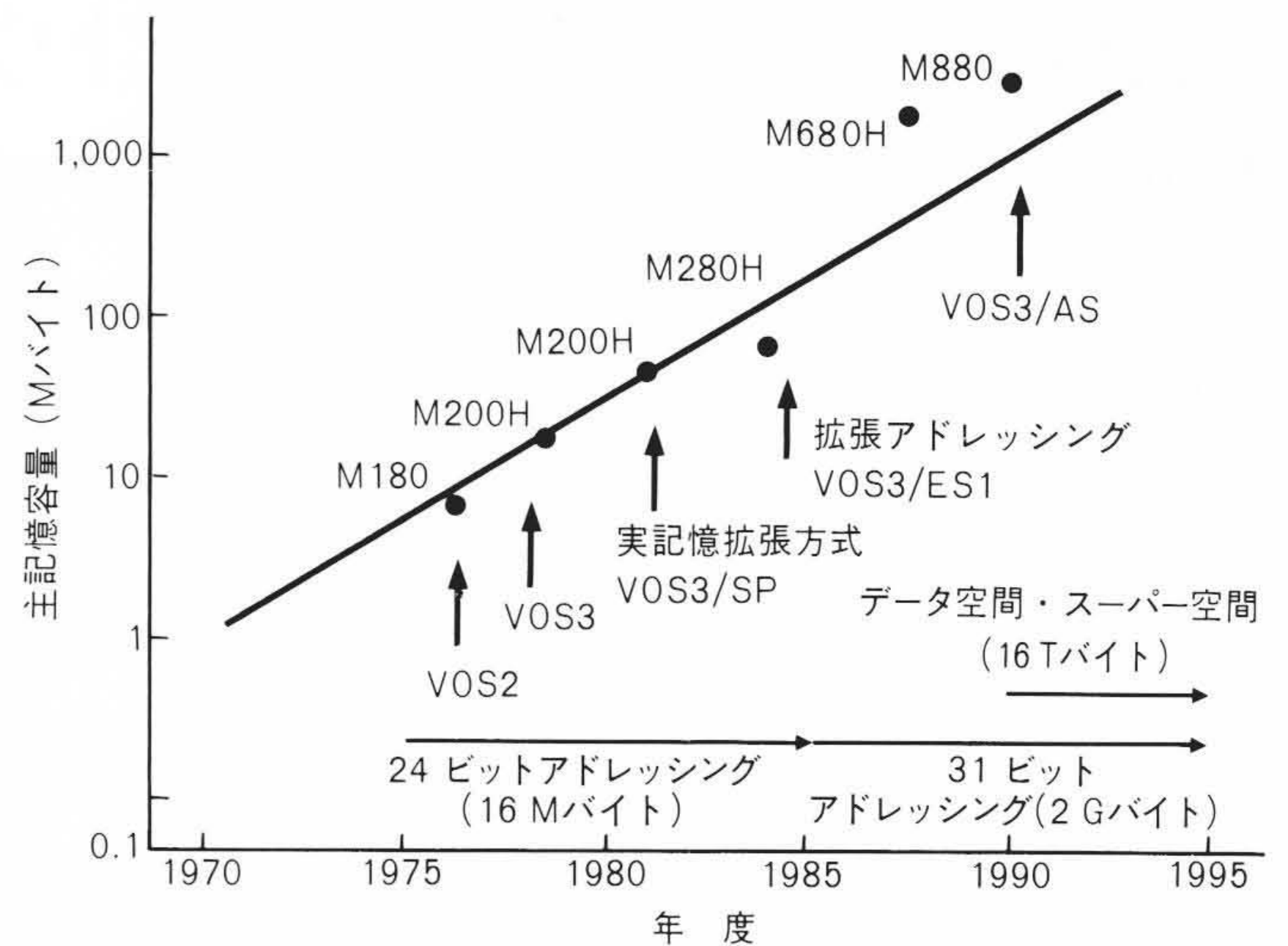
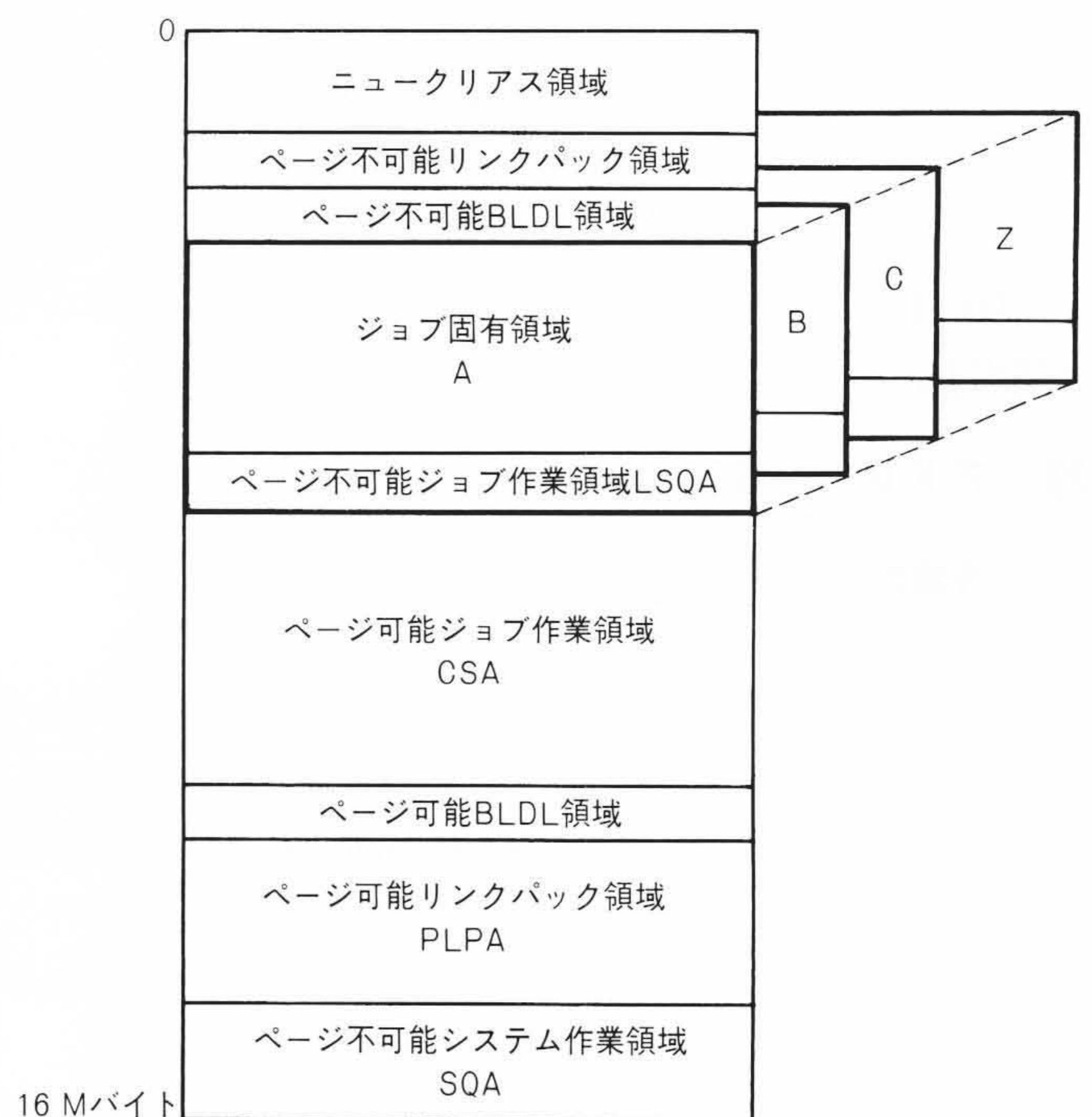


図2 Mシリーズでの最大主記憶容量の推移 利用可能な主記憶容量は、年率約30%以上の勢いで増大している。

の空間を利用することが可能となった。16 Mバイトの仮想記憶領域は、OSをはじめとするシステム共通領域とジョブ固有領域によって占められている。16 Mバイトの仮想記憶の配置を図3に示す。



注：略語説明 BLDL (Build List), LSQA (Local System Queue Area)

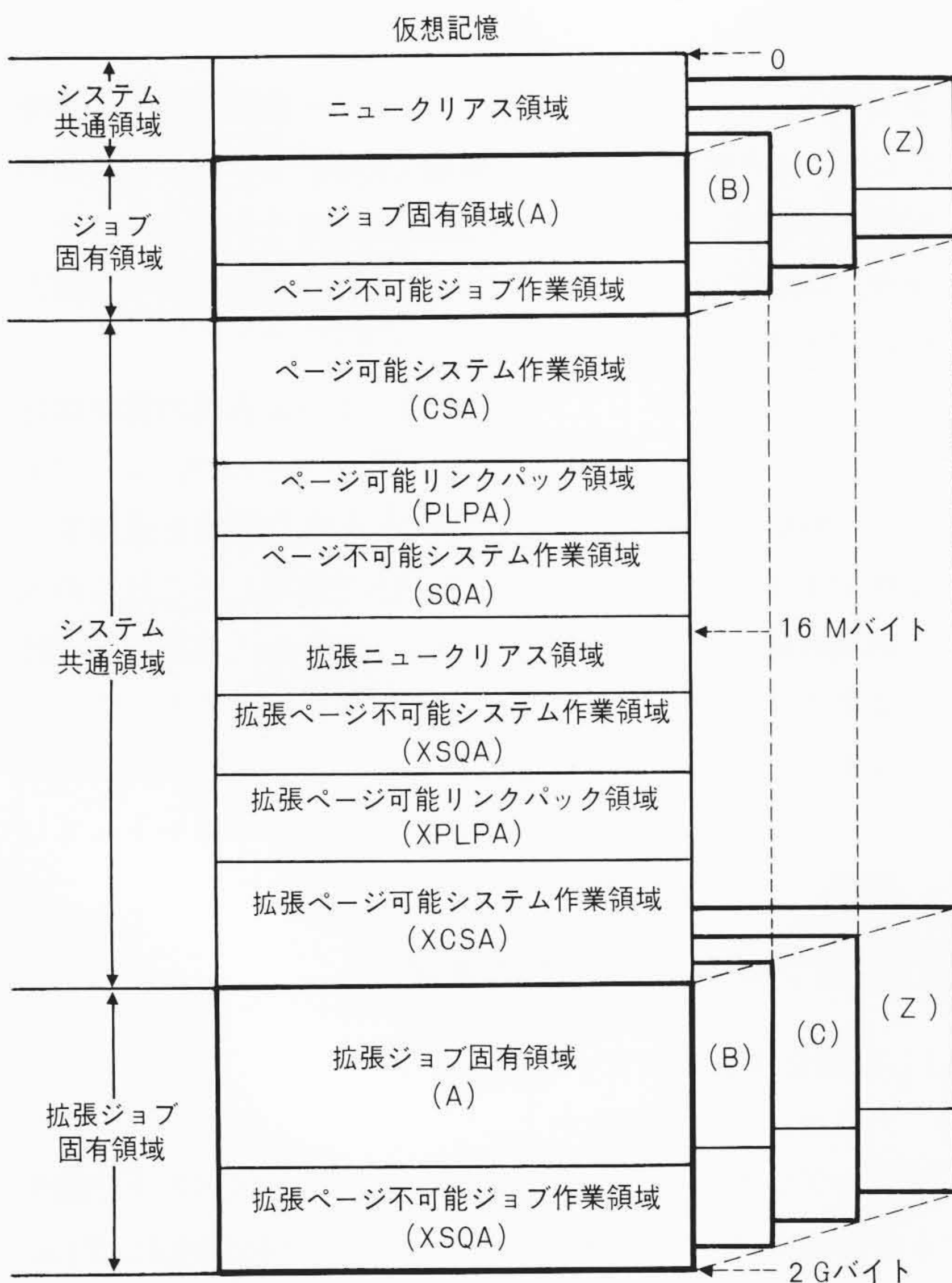
図3 VOS 3での仮想記憶のレイアウト 各ジョブが参照可能な記憶域は、16 Mバイトに拡大された。

大形計算機システムは同時に複数のユーザーが共同利用することが多く、このため共通のプログラムや情報が増大する傾向にある。特に、オンラインシステムOLTP(On Line Transaction Processing)では、規模の拡大に伴ってプログラムと管理情報の増大がシステム共通領域を広げ、結果的にジョブ固有領域が圧迫されるようになった。OSの機能拡張も同様の結果をもたらしてきた。1980年代の中ごろにはこの問題が大きくなり、ユーザー領域拡大のニーズが高まっていた。

(2) 拡張アドレッシング機能

そこで、アドレッシングの大幅な拡張を行うために31ビットのアドレス表現とし、各アドレス空間を2 Gバイト(Gバイト: 10^9)に拡張するアーキテクチャが導入された。1985年に拡張アドレッシング機能を備えたVOS 3/ES 1が開発されている。拡張アドレッシング下の仮想記憶の配置は図4に示すとおりである。

拡張アドレッシング機能により、OSをはじめとするシステムプログラムが16 Mバイト以上の仮想記憶領域に次々と移行され、16 Mバイト未満のジョブ固有領域が拡大することにな



注：略語説明 XSQA (Extended System Queue Area)
XPLPA (Extended Pageable Link Pack Area)
XCSA (Extended Common Service Area)

図4 拡張アドレッシングでの仮想記憶レイアウト VOS 3/ES 1では2 Gバイトに仮想記憶は拡張され、ジョブ固有領域は大幅に拡大された。

った。

(3) M/EA両モード実行

既存のプログラム財産を保護することは当然の使命である。アーキテクチャの変更に伴う問題点は、既存プログラム(既存Mシリーズでの状態で実行されるのでMモードと呼ぶ。)と新プログラム(EAモードと呼ぶ。)の両立性を保つ点にある。そして、新たなEA(Extended Addressing)モードで実行されるプログラムをMモードで実行されるプログラムの中から利用してもらう必要がある。このために用意されたOSの機能がM/EA両モード実行機能である。両モード実行のようすを図5に示す。

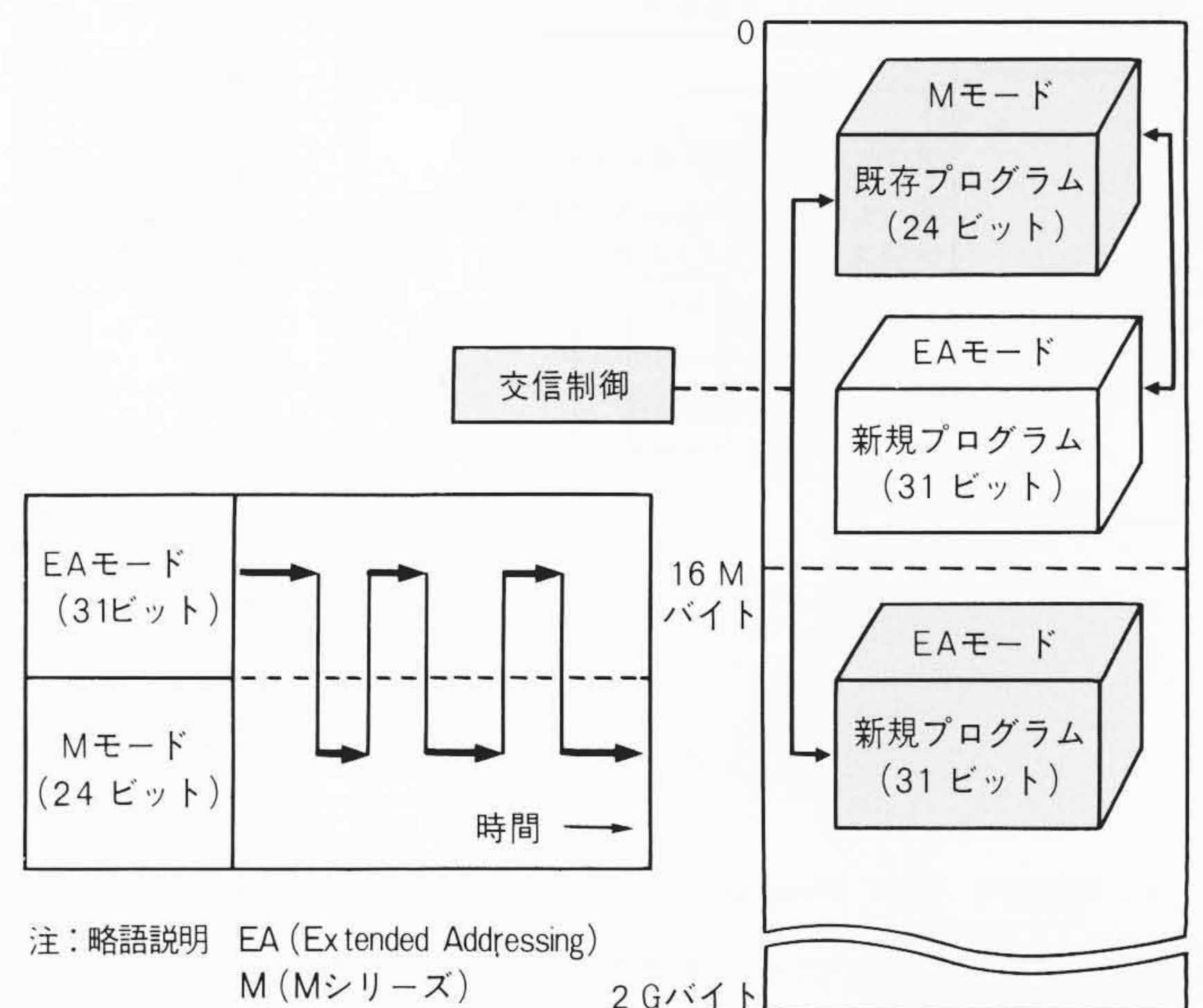
2.3 データ専用空間

(1) 究極の高性能化を実現するには。

プロセッサパワーをエンドユーザーの処理プログラムだけに100%使用できれば最大の高性能化となる。しかし、大形計算機の多くは複数のユーザーが同時に資源を共有するため、一つの処理プログラムだけでプロセッサを独占することはできない。つまり、OSに代表される資源管理とスケジューリングのオーバーヘッドがかかるため、エンドユーザーの処理プログラムに100%のプロセッサパワーを割り付けられない。

(2) 入出力削減

上記の各種オーバーヘッドを削減することはOS研究の課題である。この解決策の一つに、入出力削減がある。つまり、磁気ディスク内のファイルを主記憶上に常駐させれば、ファイルアクセス時のOSオーバーヘッドを少なくすることができる。半導体素子の開発が進み、大形計算機ではこのような方式が可能となってきた。ファイルを主記憶上に格納し、処理プログラムが直接アクセスすれば、入出力に関し究極の高性能化



注：略語説明 EA (Extended Addressing)
M (Mシリーズ)

図5 Mモード・EAモード実行機能 既存ソフトウェアの実行と、拡張アドレッシングのプログラムの同時実行を可能としている。

が図れるようになる。

このためには、データを格納する大容量の仮想記憶領域が必要となる。Mシリーズでは1990年に「データ空間」と「スーパー空間」を実現するM/ASA(M Series Advanced System Architecture)を開発し、VOS 3 /ASによって16 Tバイト(Tバイト：10¹²)のデータ空間およびスーパー空間を実現している(図6)。

データ空間は2 Gバイトの空間を8,192個連結したものであり、新たに設けられたアクセスレジスタを用いてデータの読み込みと書き込みを行う。大容量のデータ処理を必要とするRDB(Relational Data Base)などに適用されると期待される。一方、スーパー空間は、既存の31ビット アドレッシングモードでのプログラミングによって、容易に16 Tバイトをアクセスできる機能であり、拡張記憶を有効利用する新たなデータ格納機能である。

以上、単一仮想記憶、多重仮想記憶、データ空間という一連の記憶領域の拡大と発展について述べてきた。これらを図7に示す。

(3) 入出力の仮想化

究極の高性能化を実現するには、入出力の大幅な削減が第

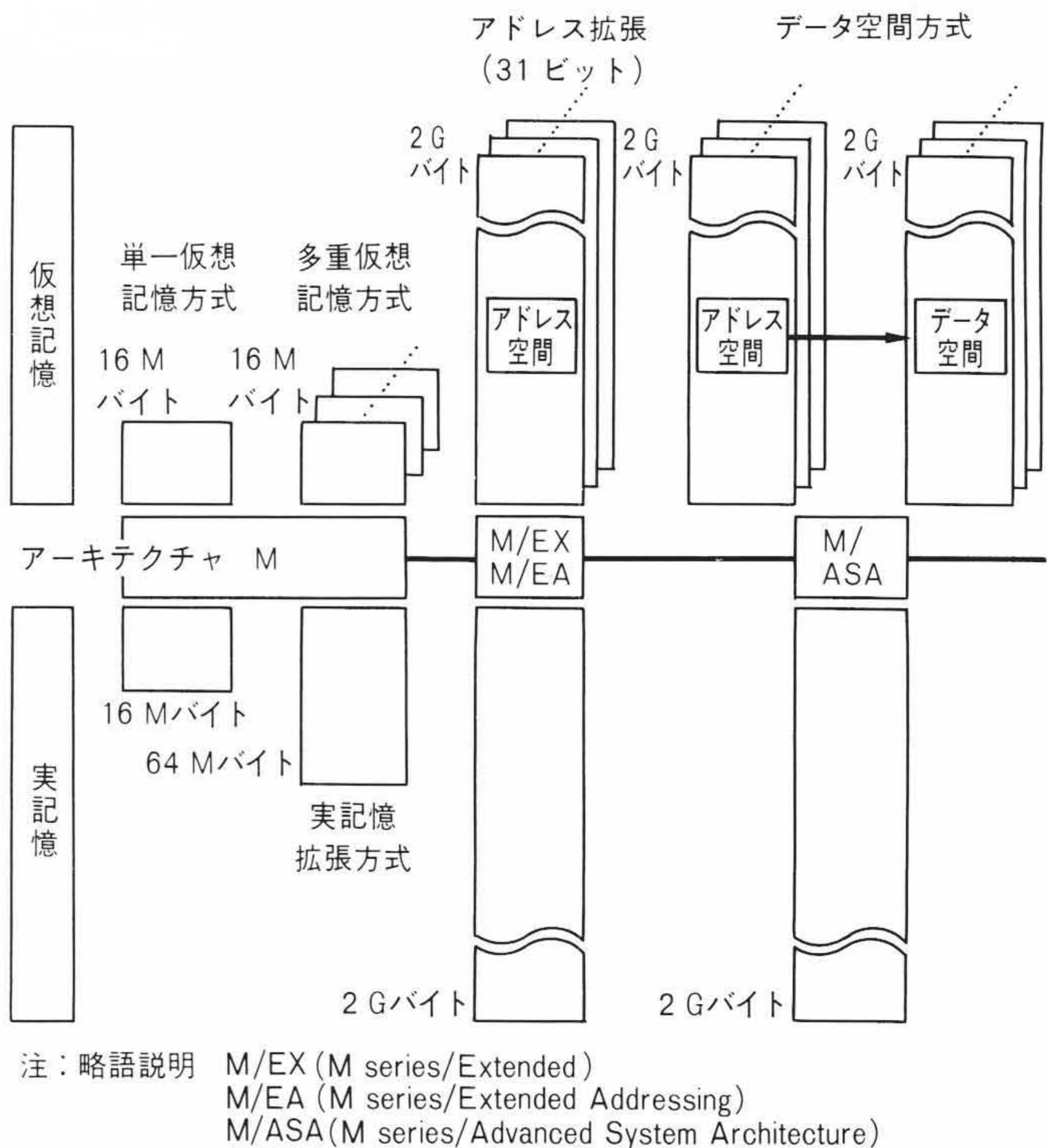


図7 Mシリーズでのアドレッシング機能の拡大 容量は初期の16 Mバイトから2 Gバイトへ、そして16 Tバイトへと飛躍的に増大した。

一步である。上記のデータ空間、スーパー空間は新たに開発されたM/ASA機構を利用した機能である。しかし、大容量の主記憶や半導体記憶である拡張記憶が利用できるようになった現在、既存プログラムからの入出力を大幅に削減する新たな機能が要求される。

そこで、OSやサブシステムが使用している各種の管理情報ファイル、プログラムのローディング用ファイル²⁴⁾、ジョブ実行時の一時的なファイルなどに対する入出力削減を追求する可能性が生まれた。図8はVOS 3 /ASで実現したこれらの入出力削減機能(仮想ファイルシステム)である。これらの機能は、従来のユーザープログラムをまったく変更する必要がなく、今までどおりの入出力を行っていることになるが、OSはプログラムからの入出力要求を記憶上にシミュレートしている。詳細については文献^{3), 7), 8), 15)}に詳しい。

3 多重プロセッシングの進展

3.1 密結合多重プロセッサ

(1) プロセッサ多重度の向上

複数のプロセッサが主記憶を共有し、一つのOSの下で動作する計算機システムがTCMP(Tightly Coupled Multi-Processor：密結合多重プロセッサ)である。Mシリーズは、M180より2台以上のプロセッサを結合したTCMPを提供し、VOS 3によってサポートされてきた。M68XおよびM880では4台までのプロセッサ結合が可能であり、プロセッサパワーの要

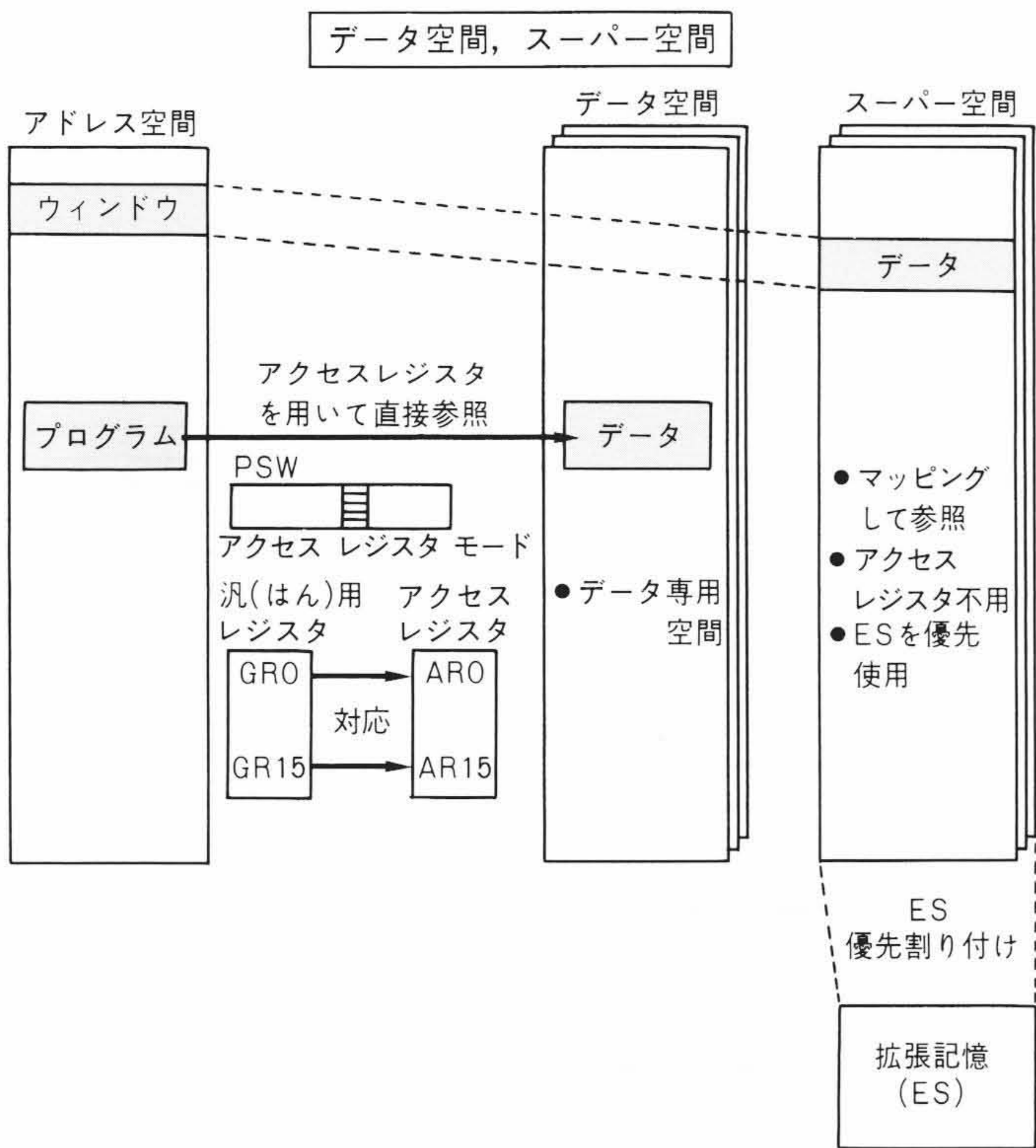
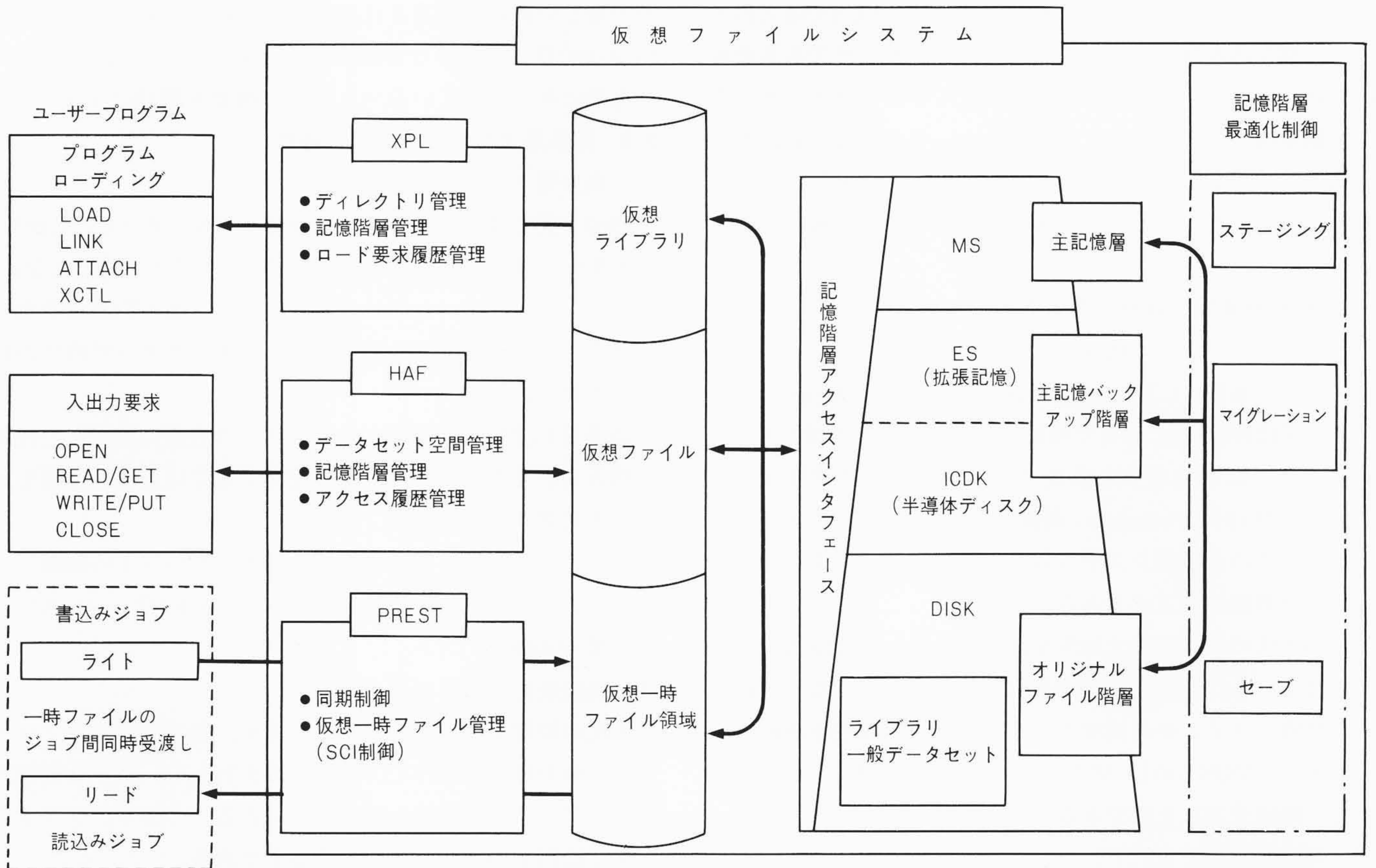


図6 M/ASA機構でのデータ空間・スーパー空間の16 Tバイト

データ空間はアクセスレジスタにより2 Gバイトの仮想記憶を結合してアクセスをし、スーパー空間は仮想記憶のウィンドウから16 Tバイトをアクセスする。



注：略語説明 XPL (Extended Program Loading), HAF (High Performance Access Facility), PREST (Parallel Reference and Synchronous Transfer Facility)
SCI (Synchronous Control for Intermediate Data)

図8 VOS 3/ASに開発された拡張仮想ファイルシステム 入出力削減は究極の高性能化を目指すOSの第一歩である。

求にこたえている。OSでは、このようなプロセッサ多重度の向上に伴う機能の向上が図られてきている。

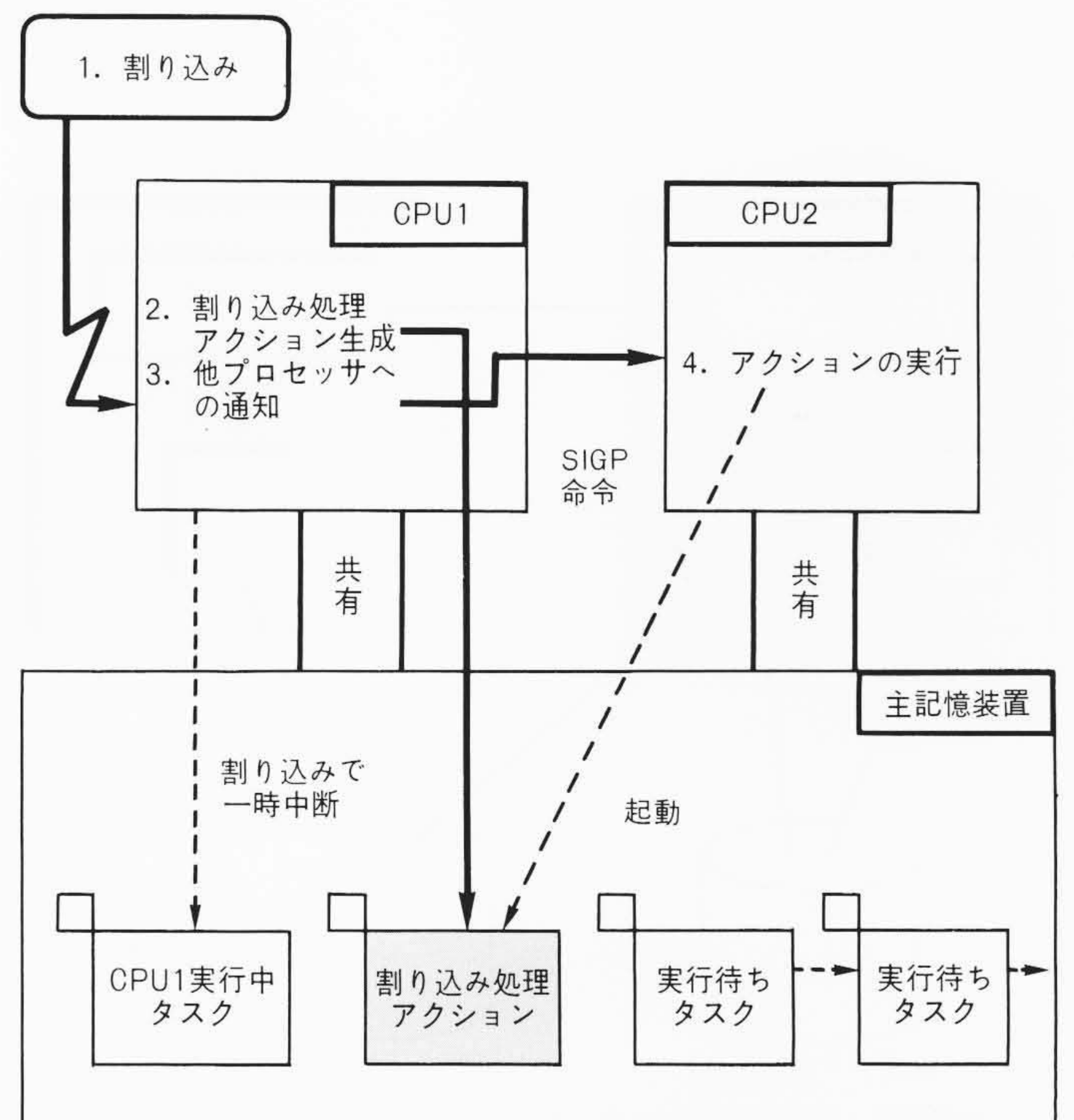
(2) 多重プロセッサ用OS機能

VOS 3は当初から大形計算機を対象として設計されており、多重プロセッサを前提とした制御構造となっている。入出力割り込み、ページング処理、ジョブ間の連絡要求などのイベント発生によってOSが起動されるが、このときOSの処理をタスクに比べ軽量のプロセスであるアクションによって実行できる構造を採用している。

多重プロセッサ用に設計されたOSでの割り込み処理の一例を図9に示す。同図では、割り込み発生によりCPU1が実行中のタスクを中断し、割り込み処理用のアクションを生成する。アクションはタスクと同様にどのプロセッサでも実行できる。そこで、他のプロセッサに対して割り込み処理のアクションを実行させるための通知を行う。これはプロセッサ間連絡用命令(SIGP: Signal Processor)命令によって行われる。この方式によって、OSの核が任意のプロセッサで並列に実行できるようになっている。この結果、OS実行のレスポンスが向上することになる。

(3) 拡張チャネルの処理

拡張アーキテクチャM/EXでは、拡張チャネルシステムECS



注：略語説明 SIGP (Signal Processor)

図9 多重プロセッサを意識したOS処理の例 CPU1への割り込み処理をCPU2にSIGP命令で指示しOS処理の並列化を軽量のアクションによって達成する。

(Extended Channel System)が開発された。ECSでは入出力完了の割り込みは、入出力を発行したプロセッサである必要はない。つまり、チャンネルが接続されているプロセッサの中で、割り込み可能な状態にあるプロセッサならば任意に割り込める。この機能は多重プロセッサの利点を利用したものであり、入出力のレスポンス、スループット向上に貢献している⁶⁾。

3.2 疎結合多重プロセッシング

(1) 最大8システムの構成

複数の計算機をCTCA(Channel To Channel Adapter: チャンネル間結合装置)によって接続し、ジョブ入出力用のスプールボリュームを共有する多重プロセッサをLCMP(Loosely Coupled Multi-Processor: 疎結合多重プロセッサ)と呼んでいる。一つの計算機システム以上の性能を必要とする場合、有効な計算機構成方法である。

1970年代の後半から大規模な計算センターではLCMPが利用されている。LCMPは処理能力が高いばかりでなく、複数の計算機システムから構成されているため信頼性を高める利点もある。VOS 3のLCMPは最大8台の計算機システムを結合し、構成することができる。

(2) 可用性向上

1980年代の中ごろからは、複数の計算機でデータベースを共有することを目的としたLCMPのニーズがOLTPの分野で発生した。このためには、ファイルアクセスの逐次処理とファイルを操作している段階でのシステムダウンを回避するた

めの新しい機能が要求される。ARF(Advanced Reliability Feature)はこのような目的のために開発され、OLTPの分野で有効に利用されている²³⁾。ARFの概要を図10に示す。

3.3 高多重プロセッシングの時代

(1) 高多重プロセッシングの課題

1980年代では2台のTCMPが実用レベルにあった。しかし、1990年代には3台以上のTCMPが真に要求される時代になる。また、メインフレーム各社はそれぞれ高多重のTCMPをセールストークにしており、本格的な高多重プロセッサ時代の到来が予想される。

高多重TCMPの主目標は、プロセッサ台数に比例した性能を得ることである。この目標を達成するには以下の問題を解決する必要がある(図11)。

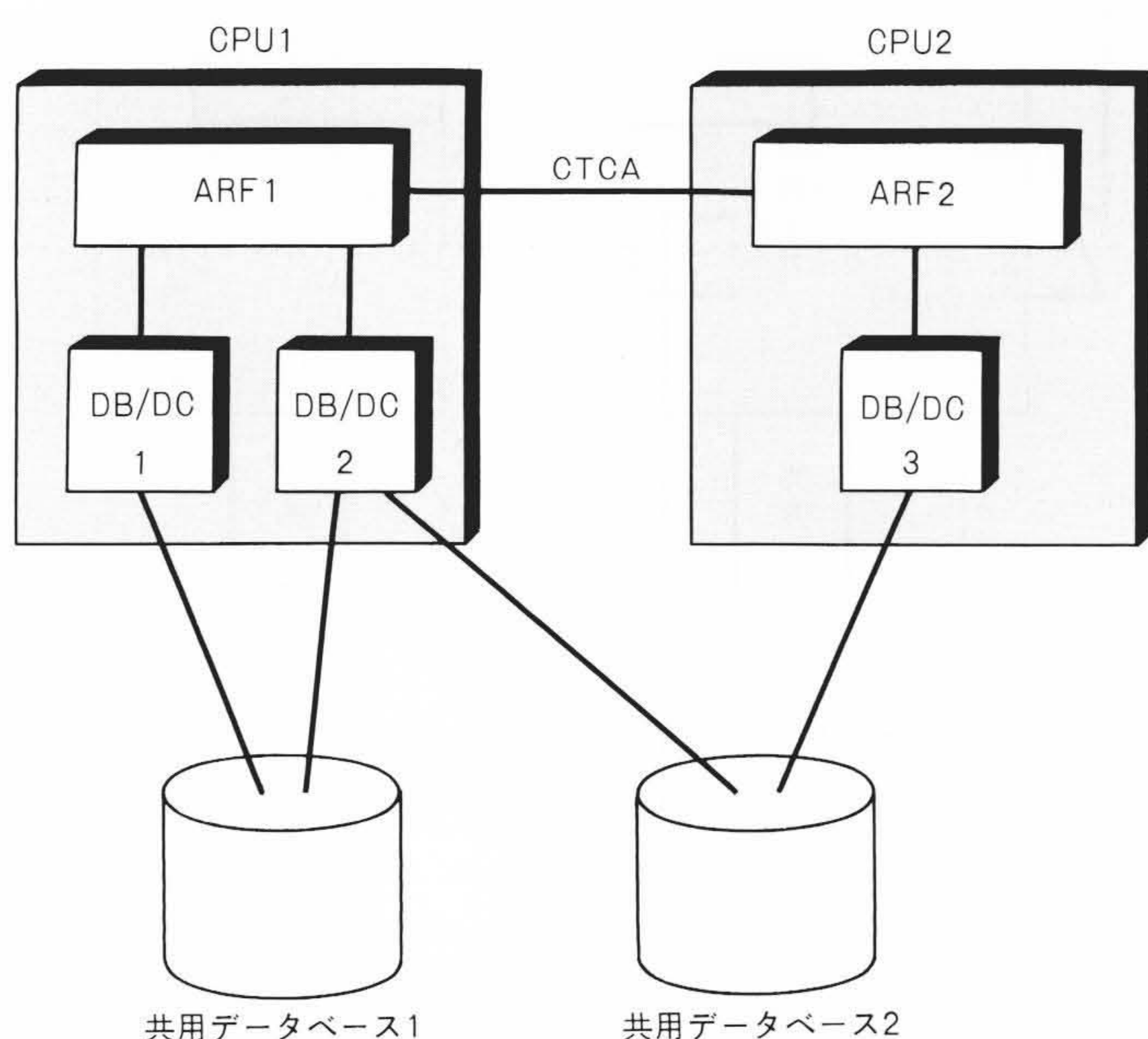
- (a) ソフトウェア資源の排他制御オーバーヘッドの削減
- (b) システム規模増大に伴うオーバーヘッド(ラージシステム効果)の削減

(2) 高性能化へのアプローチ

上記の問題解決は多方面にわたるが、基本的には各プロセッサが極力独立に動作し、相互干渉を行わないような制御方式のソフトウェア構造とすることである。つまり、ソフトウェアでは以下の方策を実現する必要がある。

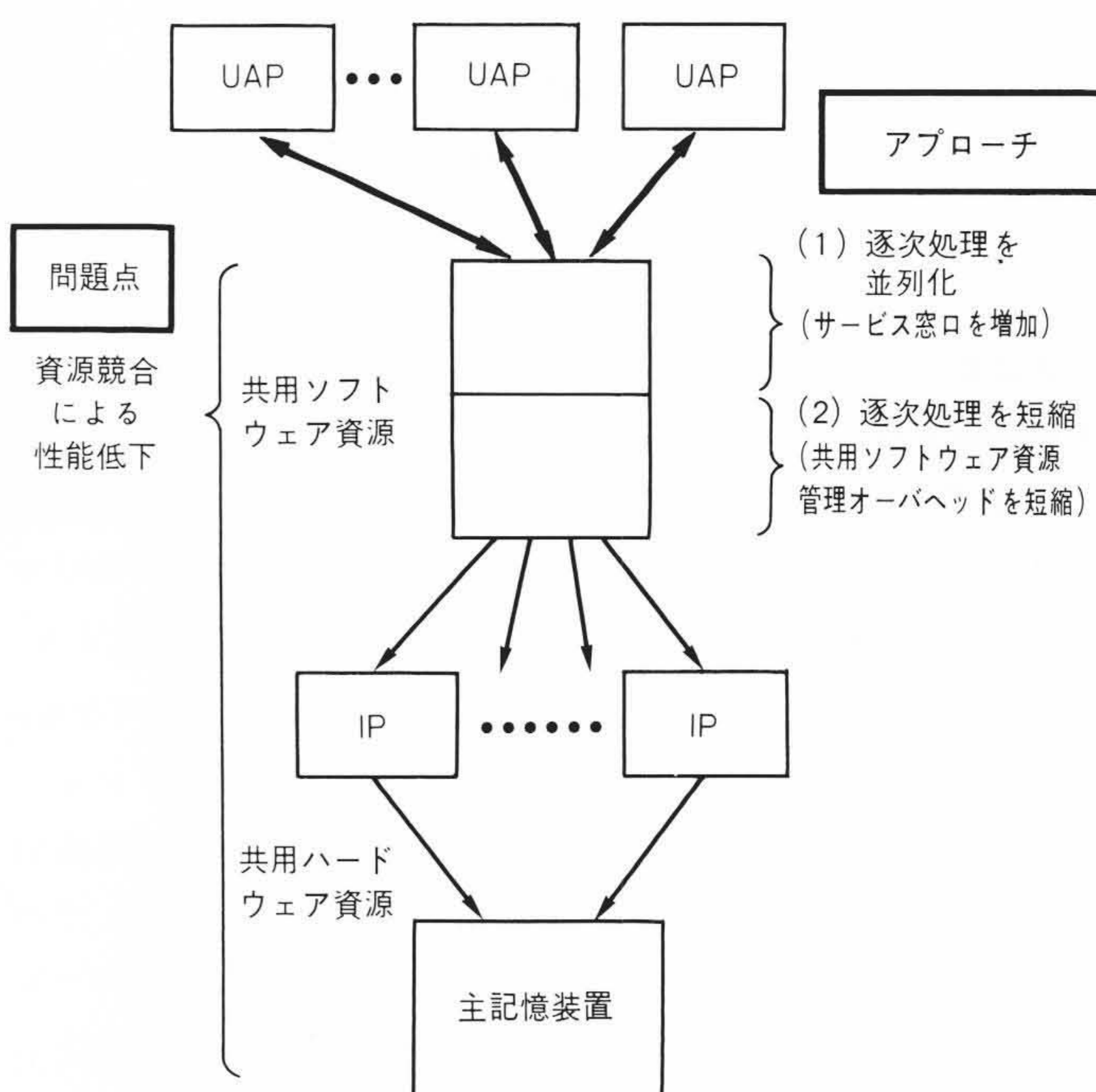
- (a) 排他制御の単位を小さくする。
- (b) プロセッサ単位に管理単位を分割し競合を削減する。

上記以外の改善として、メモリ競合が引き起こすハードウェア性能の劣化の回避を同時に行う必要がある。これらは、



注: 略語説明 ARF (Advanced Reliability Feature)
CTCA (Channel to Channel Adapter)

図10 高信頼なファイル共用形のLCMPを構成するARF ファイルの逐次処理を統一的に取り扱い、プロセッサ停止時に矛盾なくリカバリを達成する。



注: 略語説明 UAP (業務ソフトウェア), IP (命令プロセッサ)

図11 多重プロセッサの性能上の課題 共用情報の逐次使用を必要とするため、制御プログラムの処理が待たされ性能が低下してしまう。

M880およびそのOSであるVOS 3 /ASで幾つかのくふうが行われている。

(3) 高多重LCMPのカップリング

計算機パワーとシステムとしての信頼性をさまざまな要求レベルに応じて実現することが望ましい。機能分散、地域分散などによる高性能化・高信頼化が自由に行えることが今後必要となってくる。

このような分散形態での課題は、運用の煩わしさ、および運用に要する人件費の増大にある。

この問題は集中した運用管理方式の解決法を提起しており、VOS 3 /ASではAOMPLUS(Auto Operation Monitor PLUS)によって解決される¹¹⁾。

今後はデータ処理の大規模化に対応する高多重なLCMPの出現が望まれ、高性能・高信頼化のニーズがより高くなる。高多重なLCMPでは、各計算機どうしの通信が頻繁になり、通信の性能がキーポイントになると予想される。そこで、高速な通信を行える手段として、共有形の拡張記憶装置の利用が期待される。また、使い勝手をよくするためにも仮想記憶上でデータを通信する結合方式VSCM(Virtual Storage Coupled Multi-processor)の実現が期待される。

4 記憶階層の利用技術

4.1 主記憶の大容量化

(1) 大容量主記憶利用の時代

1980年代の中ごろからは128 Mバイト以上の主記憶が利用できるようになり(図2)、主記憶不足による性能劣化の問題は解消される傾向となった。むしろ、大容量の主記憶を利用し、プロセッサ負荷を削減する技術を求める傾向となった。例えば、TSSでは対話処理ごとにアドレス空間を二次記憶にスワップしていたが、余裕のある主記憶でこのオーバーヘッドの最小化を図ろうとするデマンドスワッピング方式などはそのよい例である¹⁶⁾。

(2) 入出力削減に利用

1980年代の後半になると、さらに大容量の主記憶が利用できる状況となる。このような背景から2章で述べたような入出力の削減を図る技術開発がなされ、VOS 3 /ASで図8に示す幾つかの機能が開発された^{3), 7), 9)}。これらから、磁気ディスクはファイルの信頼性を確保するために使用されるアーカイブ機能としての意味しかなくなる傾向にある¹⁾。

4.2 仮想記憶のサポート

(1) 磁気ディスクの限界

仮想記憶方式は主記憶容量以上の記憶域をユーザーに提供できることから、プログラミングを容易にする利点がある。また、ページ化されたメモリによって主記憶を有効利用できる利点もあり、1970年代の初期のころから商用化されてきた。当時、主記憶容量以上となった記憶域を格納する仮想記憶の

バックキングストレージとして、磁気ディスクは有効な装置として利用されてきた。

しかし、1970年の後半から1980年代に入ると、多重仮想記憶が本格的に利用されるようになり、バックキングストレージに対するアクセス頻度が高くなるに従って性能上の問題が生まれてくる¹⁷⁾。特に、対話処理の典型であるTSSではスワッピング時間が応答時間に占める割合が高くなり、磁気ディスク装置とチャンネルの増設を行う必要が生じてきた。

(2) 半導体ディスクの利用

1980年代の中ごろから上記の対話処理での応答時間の問題が生じ、改善が必要とされていた。半導体メモリが比較的安価になり、磁気ディスクのアクセスギャップを埋めるために半導体ディスク装置が開発された。この製品は日立製作所のオリジナル製品であり、ソフトウェアからは従来の磁気ディスクとまったく同様に扱える。OSでは、半導体ディスク装置をページング装置として積極的に利用するために、磁気ディスクよりも優先して使用するアルゴリズムの改善を行っている。この結果、対話処理の応答性向上がなされている。

(3) キャッシュ付き磁気ディスク

半導体ディスクは対話処理の改善やアクセスの頻繁なファイルの格納に利用され、大きな効果があった。しかし、容量が1980年代のはじめでは128 Mバイト程度がコスト的に限界であり、磁気ディスクの容量に比べ小さく、むしろキャッシュ付き磁気ディスクとするほうが有効である。

このアプローチは現在の磁気ディスクの原形となるものである。現在は、ファイルの蓄積と読込みを複雑なプログラム制御機能によって実現する磁気ディスクシステムに発展している¹⁰⁾。

4.3 拡張記憶の利用と高性能化

仮想記憶のバックキングストレージとして半導体ディスク、キャッシュ付きディスクなどが開発されてきた。これらは対話処理の応答性能を向上し、かつバッチ処理のスループット向上に寄与してきた。しかし、より高速の処理を実現するためには、先に述べたファイルの仮想記憶への常駐が必須(す)である。このために、16 Tバイトのデータ空間・スーパー空間が開発されてきたのである。

このような仮想記憶の拡大は、必然的にバックキングストレージの負荷の増大を招き、バックキングストレージの性能がシステム性能を左右するようになる。そこで、大容量かつ高速なアクセスができ、アクセス時にOSオーバーヘッドが小さな装置の必要性が出てくる。ES(Extended Storage: 拡張記憶装置)はこのような背景から開発された^{5), 12)}。

ページングでのOSオーバーヘッドを磁気ディスクの場合とESの場合を比較した結果を図12に示す。仮想記憶管理に代表的なページ読込み(ページイン)と不活性なページの取り上げ処理(ページスチール)を比較したものである。同図から、OSオ

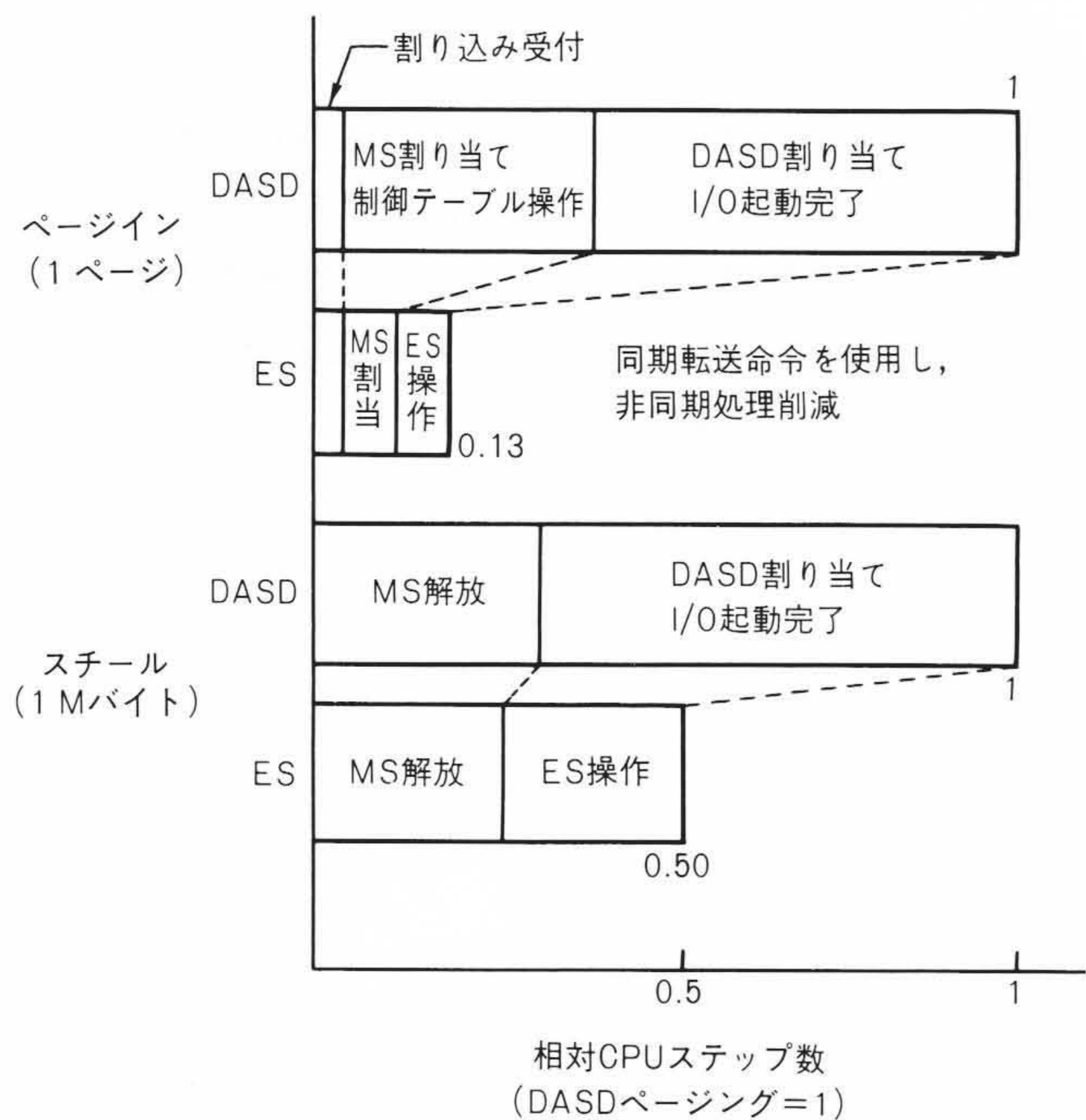


図12 拡張記憶を採用したときのページングオーバーヘッドの効果
アクセスの大幅な短縮だけでなく、OS処理を軽くする拡張記憶は大容量仮想記憶のために利用が期待される。

オーバーヘッドはページ読み込みで約 $\frac{1}{7}$ 、ページスチールで約 $\frac{1}{2}$ となっている。また、アクセス時間は「ミリ秒からマイクロ秒」に短縮している。

4.4 統合記憶管理

(1) ファイル保守・運用

計算機がデータ処理に利用されるようになってから、ユーザーの作り出すデータ量は飛躍的に増大し、また蓄積量も膨大なものになっている。多くの場合、計算機システムのダウン、磁気ディスクのクラッシュなどに備えて、これらのデータのバックアップコピーを採取するというファイル運用を行っているのが通例である。しかし、データ量が膨大になった現在、ファイルのバックアップ作業には多くの人手を要し、また、操作の誤りが介入する、などの問題が生じている。

従来、ファイルの保守・運用はサイトごとの特性を反映した個別の問題としてとらえられていた。そして、おのこの創意くふうによって行われていたのが実情である。しかし、1990年代には従来にもまして大容量の磁気ディスクが出現しており、人手による管理の限界を越えてくる。そこで、ファイルの保守・運用がOSの標準的な機能として必要とされる。

これらを背景として、ファイルの保守をOSの基本機能に組み込むことが不可欠の条件となってきた。VOS 3/ASではDMFHSS(Data Management Facility Hierarchical Control for Storage Services)によってファイル運用・管理を可能にしている¹¹⁾。

(2) スペース割り当て自動化

各種のストレージが開発され利用されるようになってきた。また、データ蓄積量の増大から磁気ディスクボリュームの増大も運営上の大きな問題になっている。したがって、ディスクにどのようなファイルが格納されているか、また、ファイルの容量やアクセス状態はどのようなになっているのかなどの情報を得て監視する機能が必要となっている。

さらに、ファイルの格納では使用目的によって最適な場所への領域割り当てがなされる必要がある。格納する装置が多くなり、かつ新しい装置が導入されていく状況では、おのこのユーザーが最適な格納場所の決定を行うには負担が大きい。そこで、ファイルの特性と各ボリュームの使用状況に合わせた自動的な領域割り当てが必要となってくる。DMFVSS(DMF Volume Maintenance for Storage Services)はこの目的で開発されている。

(3) 磁気テープ操作の省力化

磁気テープは、計算機の初期の段階から大容量のファイル格納に適した安価な記録媒体として利用されてきた。しかし最大の課題は、磁気テープを装置にマウント・デマウントする操作を要するために省力化ができない点にあった。磁気テープのハンドリングを行うロボットは古くからの要求であった。

VOS 3/ASでは磁気テープライブラリ(H-6951)内のテープボリュームを自動的にマウント・デマウントするDMFLSS(DMF Library Support for Storage)が、このロボット機能を実現している。

5 データベースの技術動向

5.1 顧客ニーズの進化

(1) 情報システムの大規模化

DBMSは、情報資源管理の中核として企業情報システムのインフラストラクチャを形成している。銀行・証券の第三次オンラインシステムでは、DBMSは中心的な役割を果たし、処理量、データ量の規模が飛躍的に増大してきている。

このため、複数のプロセッサによる計算機能力、および大容量記憶装置による記憶容量を確保する必要がある。3章で述べた高多重のTCMPとLCMPによるシステム化、およびそれらの統合による大規模システムが、ユーザーの要求に応じて構築される傾向にある。

(2) ネットワークによる広域化・計算機システムの複合化

企業活動のグローバル化、CIM、SIS(Strategic Information System)の構築に対応して、情報システムの統合・複合化が大きな流れとなっている。これにこたえる情報システムの実現には、通信機能の整備、データベースの相互アクセス基盤の確立、情報システムを業務を単位として段階的に構築・拡張運用するための複合サブシステムアーキテクチャの開発が前提となる。これを支える基盤技術が、任意のデータをアクセスするための分散機能と、安心して使用・拡張できるた

めのノーダウン・ノンストップ化機能である。

(3) マルチベンダによるシステム化

最近、業種を越えた提携、M&A (Merge and Acquisition) などにより、マルチベンダによるシステム構築へのニーズが増大している。これに対応するため各種の国際標準化、業界標準の動きが加速している。そしてDBMSは、異機種でのインタオペラビリティを追求し、マルチベンダによるシステムの構築を可能にしなければならない。

(4) エンドユーザー自身によるコンピューティング

ソフトウェアの生産性向上・情報処理部門の人員強化の努力にもかかわらず、バックログが増大し、変化に即応したサービス要求にこたえられていない。この問題を解決するためには、業務担当者自身による容易なプログラミングを可能とする4GL(第4世代言語)や、ユーザーインタフェースツールの開発などが必須である。

(5) 業務の多様化と高度化

基幹業務と情報システムとの融合・共存を行い、付加価値を高めるため、上記SIS、CIMなどのトータルシステムの構築が進められている。そのために、ワークステーション上のOAやDSS (Decision Support System) といった各種サブシステムに大形計算機上のDBMSのサービスを提供しなくてはならない。

強力な計算機パワー、チャネル機能、大容量記憶、ミラーディスクなどを装備する大形計算機は、高信頼・高性能なコンピューティングサーバ、DBMSサーバとしてだけでなく、全体的なシステムの管理を行うコントローラとして位置づけられる。つまり、標準のソフト・ハードとリンクできる機構を持ち、任意のノードのアプリケーション・DBMSに対してアクセスが可能でなければならない。

一方、ユーザーインタフェースの機能は、ワークステーションに負荷分散される傾向にある。しかし、システム全体のアプリケーションは増大するために、大形計算機システムの負荷は増え続けていくことになる。

5.2 データベース実現の課題

5.2.1 構造形データベースからリレーショナルデータベースへ

DBMSの概念は、ユーザーが個別にプログラミングしていたデータ処理機能を、基本ソフトとして汎用化することによって生まれた。商用DBMSは、1964年にC. W. BachmanのIDS (Integrated Data Store) を基本にしたCODASYL (Conference on Data Systems Language) 提案のネットワーク構造DBMSが最初である。その後、階層構造を含め、多くの構造形DBMSが開発され実用に供されてきた。現在の主要なデータベースは、ほとんど構造形DBMSを使用している。

一方1970年、E. F. Coddにより提案されたRDBは、その潜在能力は評価されていたが実用面での実績が乏しかった。この主要因は性能上の問題である。しかし、ハードウェア・

ソフトウェアの進歩により、OLTPへの適用が可能となってきた。

今後のDBMSは、構造形DBMSに比較して、生産性が2倍～2.5倍と言われるRDBが基幹DBMSの役割を果たすものと考えられる²²⁾。RDB用アクセス言語であるSQLの特長は、目的のリレーション (テーブル) にたどりつく過程を使用者が逐一指示するのではなく、RDBマネージャがこのアクセスパスの決定を自動的に行う点にある。この考えが、使用の容易性、簡潔なデータモデルの実現、さらには、4GLなどユーザー言語との親和性、DBMSの拡張性の保証などの効果を実現している。

SQLの性能強化は、今後も継続して行う必要がある。このキーとなる技術が、最適化処理、参照整合性機能の実現、最近の先進ハードウェア機構を活用したシステム資源の常駐化・並列化、DBMS専用ハードウェアなどである。

5.2.2 広域複合サブシステムへ

SIS、CIMの構築では、初めからすべてを開発するのではなく、業務グループ (サブシステム) 単位に修正・拡張を繰り返してトータル化されていく。このサブシステム化の目安となるのが、端末のグループ、データベースのグループを基準とした業務システムのグループ化である。このサブシステムを単位とした情報システムを実現する基本ソフトウェアとして、データベース・データコミュニケーションは以下の機能を備えていなければならない。

(1) システム拡張性とサブシステム単位の運用機能

(a) 複合サブシステム

一つの計算機システムの中でデータベースと端末のグループを作り、これらを複数のデータベース、データコミュニケーションサブシステムに分割する。分割した各サブシステムに対して業務サブシステムを構成し、サブシステム単位の運用 (開始・終了・障害回復など) 機能を実現する。サブシステム間では、データベースの相互アクセス、ジャーナルの一元化運用を可能にする。

(b) ローカル複合サブシステム

上記の各系は同一計算機センター内の通信機能によって結合され、複数の計算機システムに拡張されたローカル複合サブシステムとなる。負荷の増大、高い信頼性、サブシステム間での連携処理などの要求がこの構成で解決される。すなわち、各計算機システム上の複数のデータベース・データコミュニケーション機能により、データベースの共用と分散アクセスがこれらを可能にする。

(c) 広域複合サブシステムの実現

ローカル複合システムが地理的に離れた計算機環境に拡張された系である。単一の計算機システムの運用に比べて複合サブシステム系の運用はきわめて複雑になる。データベース・データコミュニケーションシステム固有の運用機

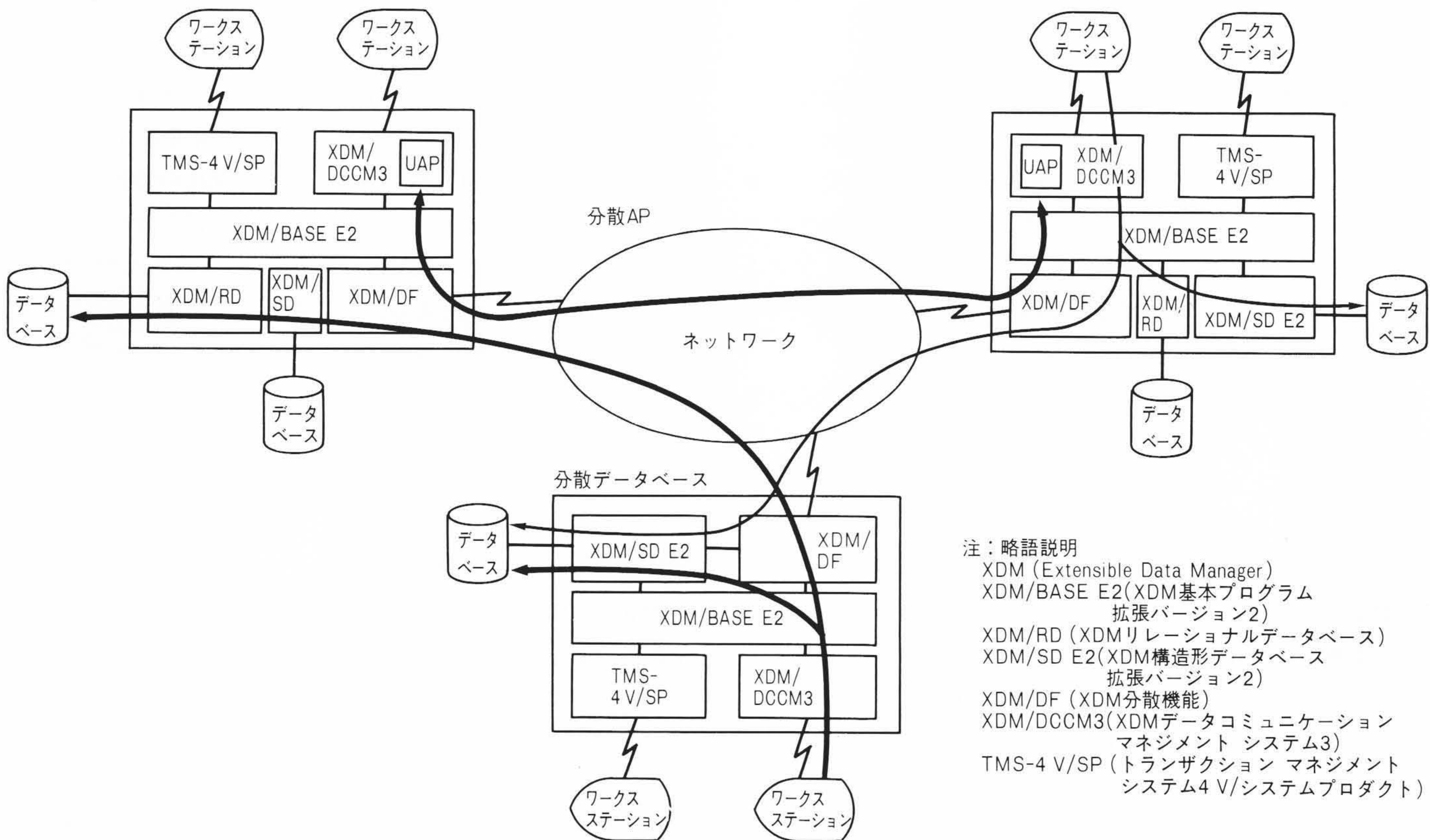


図13 XDMによる広域複合システム 複合サブシステム、ローカル複合サブシステムおよび広域複合サブシステムと、システムを拡大可能とする方式を実現している。

能に加えて、(ローカル・広域)複合サブシステムを統合運用するための機構が必要となってくる。

最新のDBMSであるXDM(Extensible Data Manager)によって実現されている上記の各サブシステムの構成を図13に示す。

(2) 障害の局所化(高信頼化の達成)

上記の複合サブシステムの実現ではきめ細かな高信頼化機能を装備しなければならない。その主なものに以下の機能がある。

- (a) 回復機能
 - (i) ホットスタンバイ回復
 - (ii) QSR(Quick System Restart)回復
 - (iii) システム回復(ウォームスタート)
 - (iv) サブシステム回復
 - (v) トランザクション部分回復
- (b) 閉そく機能
 - (i) 分散ノード閉そく
 - (ii) 共用プロセッサ閉そく
 - (iii) システム閉そく
 - (iv) サブシステム閉そく
 - (v) アプリケーション、トランザクション閉そく
 - (vi) ページ、キーレンジ、ファイル、ボリューム、デー

タベース閉そく

5.2.3 WSと大形計算機システムの複合体へ

OA, DSSといったエンドユーザーコンピューティングの要求は、大形計算機システムが万能の世界から、ワークステーションやパーソナルコンピュータ(以下、パソコンと略す。)を含めた計算機システムへの移行を促してきた。これは、ワークステーション・パソコンと大形計算機システムが独立して機能するのではなく、企業情報が蓄積されている大形計算機システムと、プレゼンテーション、対話処理に向けたワークステーション・パソコンを相互に連携し融合したシステムを意味する。

大形計算機システムは、計算機全体の管理を行うシステム管理サーバ、データベースサーバ、アプリケーションサーバとしての役割を果たす。ワークステーション・パソコンと大形計算機システムとの連携は、統一的な利用インタフェースを定めるHAA(Hitachi Application Architecture)やOSI(Open System Interconnection)などの標準化によって達成される。

5.3 最新データベースXDMの概観

上述した課題に対応するため、Mシリーズ大形計算機の汎用データベース・データコミュニケーションとして開発してきたXDM^{19),20),21)}の諸機能を表1にまとめて示す。

表1 XDMでの重点開発目標 ローカル、広域複合計算機など、どのような大規模構成にも対処可能としている。

項番	項 目	開 発 項 目
1	密結合多重プロセッサでの高速データベース処理	●コンカレント処理機能 ●ロック保持期間の削減
2	疎結合多重プロセッサでのDBMS	●システム間データベース共用機能 ●分散データベース機能 ●ホットスタンバイ機能
3	SQLの高速化(XDM/RD)	●拡張記憶の活用(データベースの常駐化) ●データベース支援ハード(IDP)のサポート
4	NDLの高速化(XDM/SD)	●拡張記憶の活用(データベースの常駐化) ●階層データの高速パス化
5	広域複合システムに対応する分散機能の開発・強化	●分散データベース機能 ●分散AP(リモートAP呼び出し)機能 ●二相同期制御によるデータ保証機能 ●WSからのNDL・SQLデータベースアクセス
6	複合サブシステム	●マルチXDM/RD ●サブシステム単位の構築、運用性、拡張性などの実現 ●デュアルデータベース(SQL, NDLの共存)
7	統合運用機能	●ローカル複合サブシステムの統合運用機能 ●広域複合サブシステムの統合運用機能
8	生産性向上	●システム開発支援EAGLE2との連携 ●OFISシリーズ, ACE3との連携
9	ワークステーション・メインフレームのリンク	●OFISシリーズとHOAPシリーズのリンク ●意思決定支援システムのMML機能
10	国際標準機能の実現	●NDL, SQL, RDAの実現

注：略語説明
NDL(Network Database Language)
EAGLE2(Effective Approach to Achieving High Level Software Productivity 2)
OFIS(Office Automation and Intelligence Support Software)
ACE3(Available Command Language for End user 3)
HOAP(High-Level Object Management and Processing)

6 今後の動向

(1) 万能時代から棲(す)み分けの時代へ

大形計算機は、1980年代に入るとともに使用形態が大きく変化した。マイクロプロセッサの発展によって人間が直接触れるデータ処理分野はパソコン、ワークステーションがとって代わるようになってきた。また、中・小形計算機の処理能力向上によるダウンサイジングが進み、部門でのデータ処理を大形機から切り離す分散処理が進展するようになってきた。このように、必要に応じて計算機を選択することが可能になった。

この結果、大形計算機は大量のデータを管理し、それらを複数のユーザーが同時に高速かつ高信頼にアクセスするデータベースマシンとして、また高速な演算を達成するコンピューティングマシンとして、価値を発揮するようになってきた。つまり、大形計算機が万能計算機として利用された時代から、

ある意味での専用マシンの時代に移ってきた。

(2) 最適化の推進

1990年代に向けて開発された大形計算機の大容量化・高性能化は新たなアプリケーションに道を開くものと期待される。今回開発した各種の機能は、現実の世界で厳しく使い込まれて評価されていく。例えば、データ空間・スーパー空間、高多重プロセッシングなどの新しい利用は、今までにない新たな資源の利用形態であり、今後、フィールドでの使用を通してより有効な機能として定着させていく必要がある。また、データ空間・スーパー空間を支える拡張記憶の制御方式でも、ノウハウの蓄積と最適化を図ることで真の高性能化と高信頼化を達成していく必要がある。

多くの計算機資源が利用可能となっても、フィールドでのニーズは常に高いものがあり、有限の資源を有効利用する最適化の問題は多い。この意味からも、基本ソフトウェアは究極の高性能化を達成すること、高信頼化の確保、使い勝手の向上などの新しい概念を創造していくことがますます期待される。

(3) サーバとしての充実

1990年代では、音声や画像などのメディアが計算機分野で大いに利用されると予想される。これらのデータの多くはワークステーションによって作り出されると予想される。しかし、従来のテキスト情報に比べてこれらのメディアの情報量は比較にならぬほど膨大なものである。ところで、これら大量データを格納し、管理するサーバは大形計算機の得意とする分野であり、マルチメディアのDBMSとしての役割が出てくる。

したがって、大形計算機のデータサーバとしての役割はますます重要になる。しかも、これらの情報のアクセスはマンマシン系の中に組み込まれており、リアルタイム性が強く要求されることになる。このために、大形計算機にはワークステーションとのデータ交換を高速に行う新たな機能が必要であり、高速なLAN、高速なマルチメディアDBMSのアクセス機能が必須の条件となってくる。

上記のニーズを満たすためには、基本ソフトウェアが今までなし遂げてきた高性能化、大容量化および高信頼化を今後とも推進する必要がある。また、ワークステーションなど他の計算機との結合の重要性から、より開かれたシステム間通信機能を充実させていく必要がある。

7 結 言

データ処理の高性能化は、強力なプロセッシングエンジンだけでは達成されない。記憶領域の大容量化と入出力の強化がそろって初めて達成されるものである。ここでは、アーキテクチャの進歩でもあるアドレッシング機能の発展とねらい、また拡大したアドレス空間を効率よく支える主記憶の拡大と

拡張記憶装置の利用方式をみてきた。これらから、Mシリーズの歴史は記憶革命の中にあったことが理解できる。さらに、強大な計算機パワーを発揮する高多重プロセッサの実現など基本ソフトウェアの動向を示し、今後の課題などについて述べた。

現在、各種の計算機が存在しその利用も多様化している。今後もハードウェアおよびソフトウェアの進歩により、ますます多様化の一途をたどるものと予想される。計算機は現在の社会生活に不可欠の道具であり、多様化の中で大形計算機の果たす役割が徐々に明確になってきた。高性能化と大容量化、高信頼化、そして使い勝手の向上は今後とも推進しなければならない。大形計算機が情報システムの中核として果たす役割はますます大きくなっており、ニーズを先取りしたシステム開発が今後とも期待されている。

参考文献

- 1) Ousterhout, J.: Beating the I/O Bottleneck: A Case for Log-Structured File Systems, ACM/Operating System Review, Vol.23, No.1, pp.1~17 (Oct. 30, 1989)
- 2) 新井, 外: 性能評価シミュレータPRIME 3の開発, 情報処理学会第21回全国大会(1980-5)
- 3) 新井, 外: 拡張入出力制御方式XVIO(フェーズ1)の開発, 情報処理学会OS研究会, 84-OS-25-3(1984-12)
- 4) 新井, 外: 大形計算機における大容量記憶を利用した高性能化方式, 情報処理学会第41回全国大会講演論文集(4), 3 D-7(1990-9)
- 5) 片田, 外: 拡張記憶管理機能と仮想記憶管理における拡張記憶利用方式-VOS 3/AS基本機能-, 情報処理学会第41回全国大会講演論文集(4) 4 D-2(1990-9)
- 6) 木下, 外: VOS 3/ES 1における拡張チャンネルシステムのサポート(1)-概説, 割り込み制御方式-, 情報処理学会第34回全国大会(1987-3)
- 7) 長須賀, 外: ジョブ間並列同期転送機能の開発と評価-VOS 3/AS: PREST-, 情報処理学会第41回全国大会講演論文集(4), 4 D-3(1990-9)
- 8) 櫻庭, 外: 汎用入出力仮想化機能HAFの開発, 情報処理学会OS研究会, 90-OS-47-3(1990-6)
- 9) 櫻庭, 外: 入出力仮想化機能と制御方式の開発-VOS 3/AS: HAF-, 情報処理学会第41回全国大会講演論文集(4), 3 D-8(1990-9)
- 10) 久野, 外: 新キャッシュ付きディスク制御装置“H-6581-C3”, 日立評論, 73, 2, 211~216(平3-2)
- 11) 柴宮, 外: オペレーティングシステムVOS 3/ASの大規模システム運用機能, 日立評論, 73, 2, 185~194(平3-2)
- 12) 細内, 外: 拡張記憶を利用したページング方式の性能評価, 情報処理学会第41回全国大会講演論文集(4), 4 D-2(1990-9)
- 13) Madnick, S. E., et al.: Operating Systems, McGraw-Hill Book Company
- 14) 三上, 外: 計算機システム性能解析の実際, 情報処理, 学会創立20周年記念叢書, オーム社(1982)
- 15) 山本, 外: 入出力仮想化機能HAFのアクセス法への適用とその評価, 情報処理学会第41回全国大会講演論文集(4), 3 D-9(1990-9)
- 16) Yoshizawa, Y. et al.: An Analysis of User Behavior and Demand Swapping Policies in Time-Sharing Systems, Journal of Information Processing, Vol.12, No.2, pp. 125~138(1989)
- 17) Yoshizawa, Y. et al.: Adaptive Storage Control for Page Frame Supply in Large Scale Computer Systems, ACM/SIGMETRICS at Santa Fe(May, 1988)
- 18) 吉澤, 外: OS VOS 3/ASの高性能・大容量化方式, 日立評論, 73, 2, 175~184(平3-2)
- 19) 山本, 外: システムの分散, 複合化に対応する統合形DB/DCシステム“XDM”の拡張, 日立評論, 70, 7, 761~768(昭63-7)
- 20) 山下, 外: 大規模化, 分散化に対応した大規模DB/DCシステムXDM, 日経コンピュータ(1990-10)
- 21) 村川, 外: 大規模DB/DCシステムを実現するVOS 3 DB/DC高信頼化機能“ARF”, 日立評論, 70, 7, 769~774(昭63-7)
- 22) 高橋, 外: 発展するリレーショナルDBと今後のDBMSの動向, 日経コンピュータ別冊ソフトウェア(1988-7)
- 23) 新田, 外: 大規模DB/DCの高性能・高信頼化方式の考察, 情報処理学会アドバンスト データベース システム シンポジウム(1988-12)
- 24) Arai, T. et al.: Extended Virtual Input/Output Feature (XVIO) for Program Libraries, Winter Transaction of CMG'90, Orlando, Florida (Dec., 1990)