仮想計算機システム"VMS"

Virtual Machine System "VMS"

情報処理技術の急速な進展に伴って、計算機システムの拡張や切替えが頻繁に行なわれるようになってきたが、旧システムから新システムへの切替えは、従来の業務処理と併行しての新業務プログラムの開発や、新旧ハードウェアの併設などコストや期間などの点で問題が多い。日立製作所は、これを解決することを一つのねらいとして、複数システムを1台の実計算機のもとで同時に実行することが可能なVMSを開発した。

この論文では、VMSの実現方式を概説し、特に実時間処理システムにも適用するための性能向上のための諸施策と、高信頼性を得るための諸施策について論述する。

井村淳一* Imura Juńichi

小平光彦* Kodaira Mitsuhiko

若井勝郎** Wakai Katsuo

梅野英典*** Umeno Hidenori

□ 緒 言

情報処理技術の進歩、発展が激しく続けられ、新しい計算機システムが出現する状況下で、旧システムから新システムへの切替えは、ユーザーにとって新旧ハードウェアの併設やプログラム開発などの問題が新たに発生する。

VMS(Virtual Machine System:仮想計算機システム)は、 1台のHITAC Mシリーズ計算機のもとで、機能的に実計算 機と等価なVM(Virtual Machine:仮想計算機)を複数台生 成し、これらのVMを同時に稼動させることを可能にしたシステムである。

VMSの利用により、1台の計算機のもとで新旧システムの同時走行が可能となり、システム切替えに伴うハードウェア設備の併設が不要となるとともに、プログラム移行作業や開発作業を平準化することができる。この論文では、実時間処理システムのOS(Operating System)の実行を可能とするために、高性能、高信頼性を実現したVMSについて述べる。

2 VMSのねらい

2.1 システム移行の支援

VMSは1台の計算機上で複数のOSを同時に稼動できる特長をもつことを利用し、システム切替えに伴う問題の解決に以下のような支援を行なうことができる。

- (1) 新旧ハードウェアの併設を不要とし、ハードウェアコスト、設備コストを低減することができる。
- (2) 旧システムを稼動させながら、プログラムの移行を行なうことができるため、移行作業の平準化が可能となる。
- (3) 新旧両システムを同時に稼動できるため、近い将来不要となる業務プログラムの変換作業を行なう必要がなくなる。
- (4) 独立した仮想計算機を使用して、新システムに対するオペレータ、プログラマなどの教育を行なうことができる。

VMSは以上のようにシステム移行に伴う問題解決を支援する手段となり得るが、これを実用可能なものとするためには当然のことではあるが、旧システムに対して、より高性能のハードウェアを必要とすると同時に、VMCP*1 (Virtual Machine Control Program:仮想計算機制御プログラム)がVMをシミュレーションするための処理オーバヘッドを可能な限り少なくし、新ハードウェアの性能を十分に生かす必要があ

る。また、複数のVMが同時に実行するという、複合計算機 システムでの信頼性の低下要因を極力排除する必要がある。

VMSはこのような観点から、特に高性能化のための諸施策を行ない、また高信頼性を得ることに配慮して開発を行なった。

2.2 システム開発の支援

VMSはVMの主記憶装置の容量や装置アドレスを実計算機の構成に依存せずに設定することができること、また互いのVMは各々独立して実行することができるために、システムの開発テストを通常の業務とは独立に行なうことが可能である。更に、VMの動作を常に監視することができることからシステムのデバッグに便利なオペレータコマンドを用意し、システム開発やテストを支援することをねらいとした。

3 VMSの概念

VMSは実計算機の資源(処理装置,記憶装置,入出力装置など)を配分制御することにより,機能的に実計算機と等価なVMを複数台生成し,これを同時に実行することができる。このための制御を行なうプログラムをVMCPと呼ぶ。

VMが実計算機と機能的に等価であるためには、計算機資源も実計算機と同様に構成され、動作しなければならない。 VMCPは各々の計算機資源の仮想化を行ない、これらのVM 資源を統合することによって、論理的にVMを構成する。

図1にVMSの構成概念を示す。

計算機システムの仮想化は次のように行なわれる。

(1) 処理装置の仮想化

VMCPは複数のVMを同時に並行して実行させるために、OSの多重プログラミングと同様の手法で各VMに実際の処理装置を割り付ける。すなわち、VMCPはスーパバイザ状態で実行し、VMは問題プログラム状態で実行する。したがって、VM上での命令の実行は、非特権命令であれば実計算機上で直接実行し、特権命令はVMCPがシミュレーションする。

^{※1)} VMCP:複数のVMを実行させるために、実計算機資源を配分 制御する制御プログラムのこと。

^{*} 日立製作所ソフトウェア工場 ** 日立製作所神奈川工場 *** 日立製作所システム開発研究所

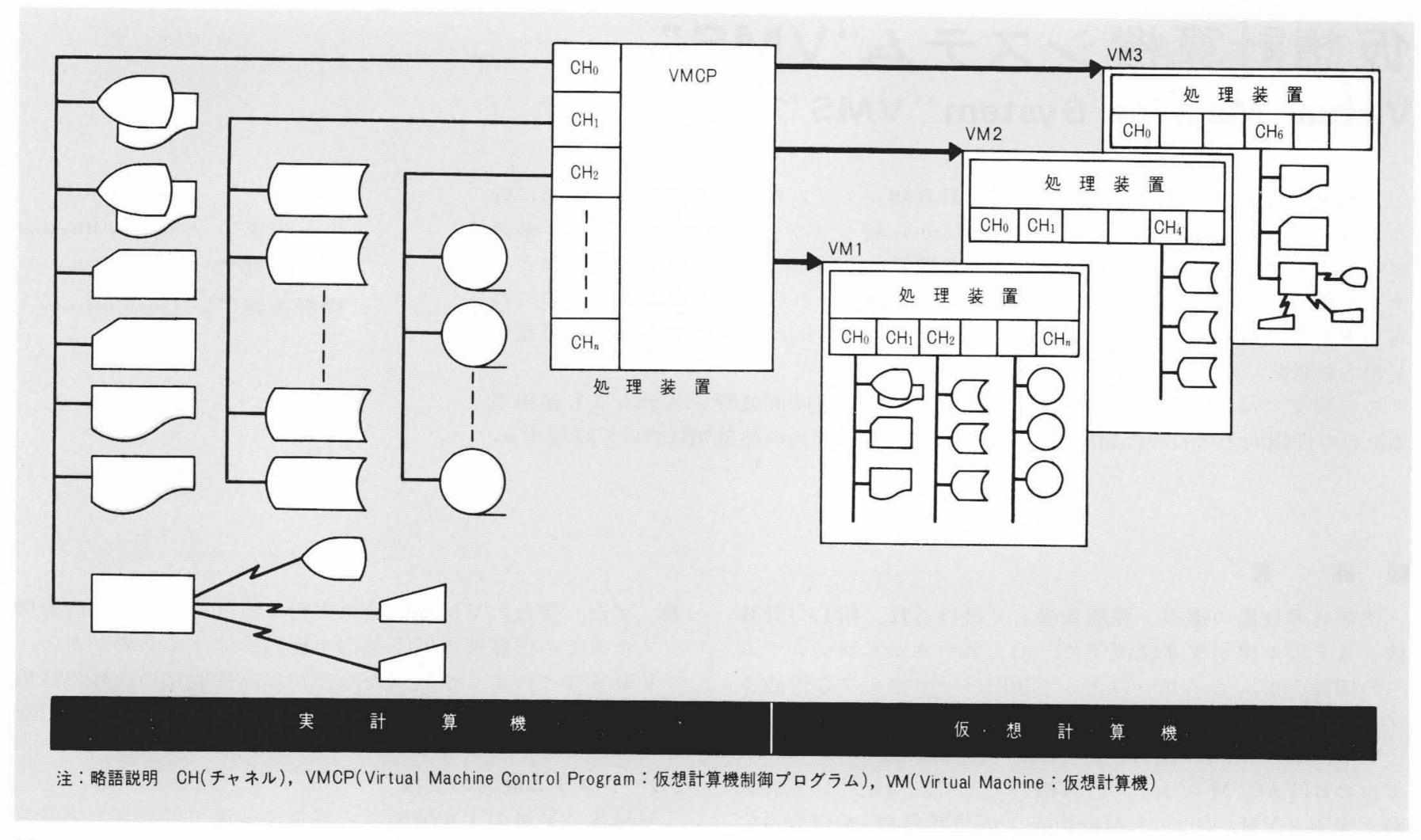


図 I VMSの構成概念図 各々の仮想計算機は、実計算機の計算機資源が配分され動作する。各々の仮想計算機の構成は、システム作成時あらかじめ定義しておく。

(2) 主記憶装置の仮想化

VMの主記憶装置は、その記憶容量分がそのまま一つの仮想記憶装置に割り付けられる。複数のVMが実行されるときVMCPは各々のVMに対してこの仮想記憶装置を割り付け、多重仮想記憶制御により、実記憶装置を動的に配分する。各各の仮想記憶装置は1ページ4,096バイト単位に分割され、必要に応じて実記憶装置上にページインされる。

またVMに対しては実計算機と同様にアドレス変換機構をもたせることができる。この場合、VM上で更に仮想記憶装置が作り出され、VMCPは三つの階層の記憶装置を制御することになる。この様子を図2に示す。実記憶装置とVMの仮想記憶装置のアドレス変換を行なうために、VMCPはシャドーテーブルと呼ぶアドレス変換テーブルを作成し、ハードウェアが直接実行可能な形に変換する。

VMの主記憶装置はこのように仮想化されるため、論理的には実計算機の主記憶装置の容量に依存せずに、任意の容量を与えることができる。

(3) 入出力装置の仮想化

VMを構成する入出力装置には、実計算機の入出力装置を配分するが、各々のVMを構成する入出力装置アドレス(仮想装置アドレス)は、VMCPが実際の入出力装置アドレスに変換して制御するため、実計算機の入出力構成とは論理的に独立して構成することができる。

(4) リモートオペレータパネルの仮想化

実計算機のリモートオペレータパネルの各種スイッチ(LOAD スイッチ、STARTスイッチ、STOPスイッチなど)は各VM に配分することができないため、各VMごとに実際のコンソール装置を割り当て、このコンソール装置を介して、VMCPのオペレータコマンド(VMCPコマンドと呼ぶ。)によりシミュ

レーションする。このためのコンソール装置をVMコンソール装置と呼び、VMはこのVMコンソール装置からLOGONコマンドが投入されたときに生成され、LOGOFFコマンドが投入されたときに消滅する。これは、ちようど実計算機の電源が投入、切断されたことに相当する。

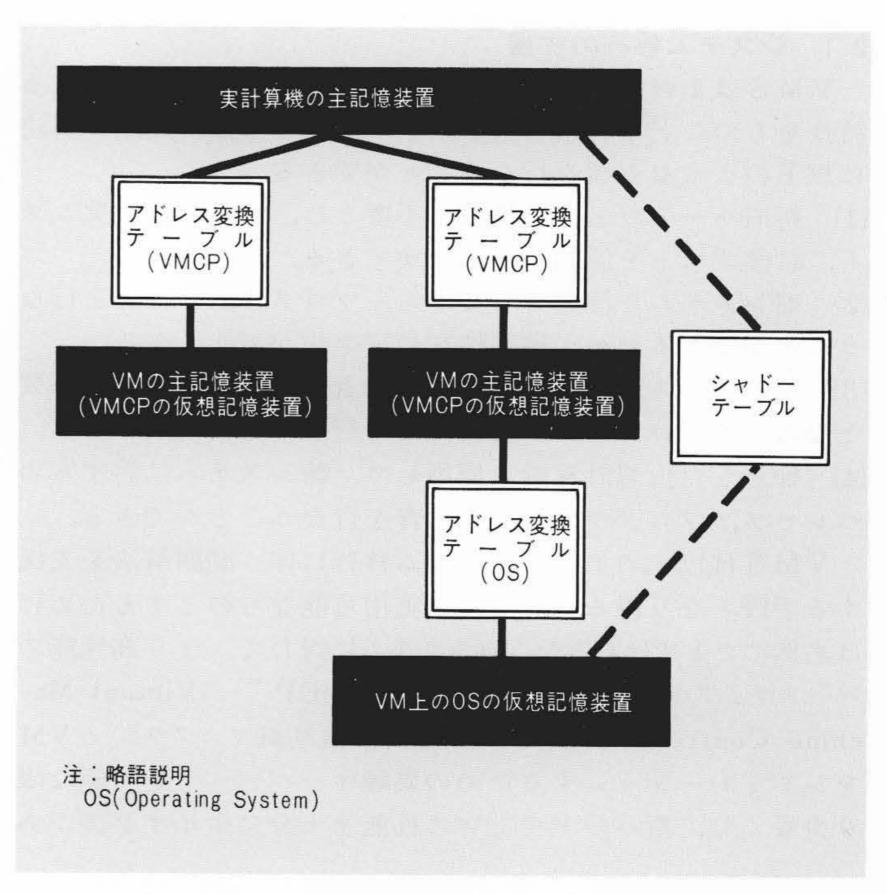


図2 VM実行時の記憶装置の階層 VMSでは実計算機の主記憶装置, VMの主記憶装置及びVMの仮想記憶装置の三つの記憶装置の階層ができる。

4 VMSの特長

VMSは1台のHITAC Mシリーズ計算機のもとで、複数のVMを生成実行できる特長を利用し、旧OSと新OSを同時に実行させシステム移行を円滑に進めようとするものであるが、特に実時間処理業務を行なうOSの実行にも耐えられるよう、性能、信頼性及び操作性に対し特別な配慮を行なった。

4.1 高性能化の実現

VMCPがVMを実現する方法は既に述べたが、VM上でOSを実行させるためにVMCPが介入する処理は、そのままVMの性能に大きな影響を及ぼす。VMSの性能を向上するポイントは、この介入処理を可能な限り小さくすることにある。図3にその要因と対策を示す。

(1) VMA機構の採用

VMCPがVMをシミュレーションするためのオーバヘッドの大部分は、VM上のOSが発行する特権命令のシミュレーション処理である。

VMSではVM上で実行するOSが発行する特権命令のうち、 発行頻度の高い特権命令15個を直接ファームウェアで実行させることにより大幅な性能向上を実現した。このファームウェアはVMA(Virtual Machine Assist)機構と呼び、シミュレーション可能な特権命令は直接VMAで実行してしまい、 VMCPの介入を受けずに処理を行なうことができる。

VMCPの介入はちようど通常のOSのスーパバイザコール 命令の実行と同様に、VMに関する情報の退避、回復やVM のスケジューリング処理が入るが、これをバイパスすること により大幅な性能改善が実現される。この様子を図4に示す。 (2) VMの記憶装置の常駐化

VMCPはVMの主記憶装置に各々仮想記憶装置を割り当て、ページング動作を行ないながらVMを実行させてゆくが、このページング動作による性能の低下を防止する手段として、VMの主記憶装置を実記憶装置上に常駐化させる機能を開発した。図5は実記憶装置上にVMの主記憶装置を常駐化した図である。

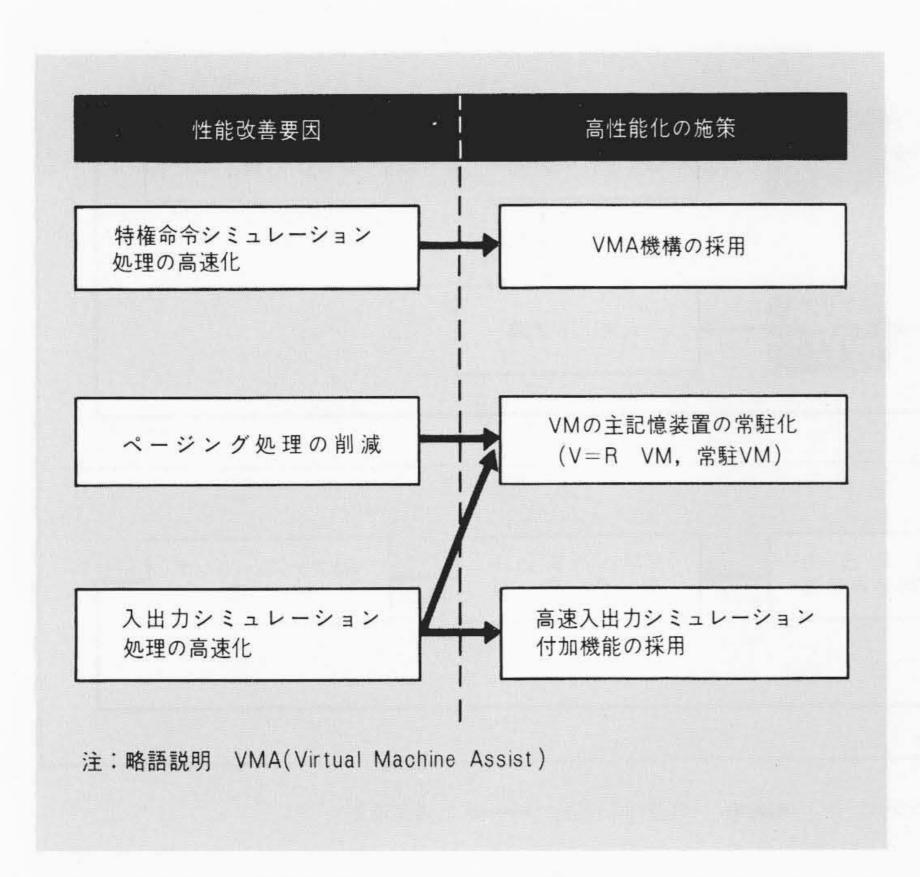


図3 VMS高性能化施策 VMSの高性能化とは、可能な限りVMCPの介入する処理を少なくすることである。

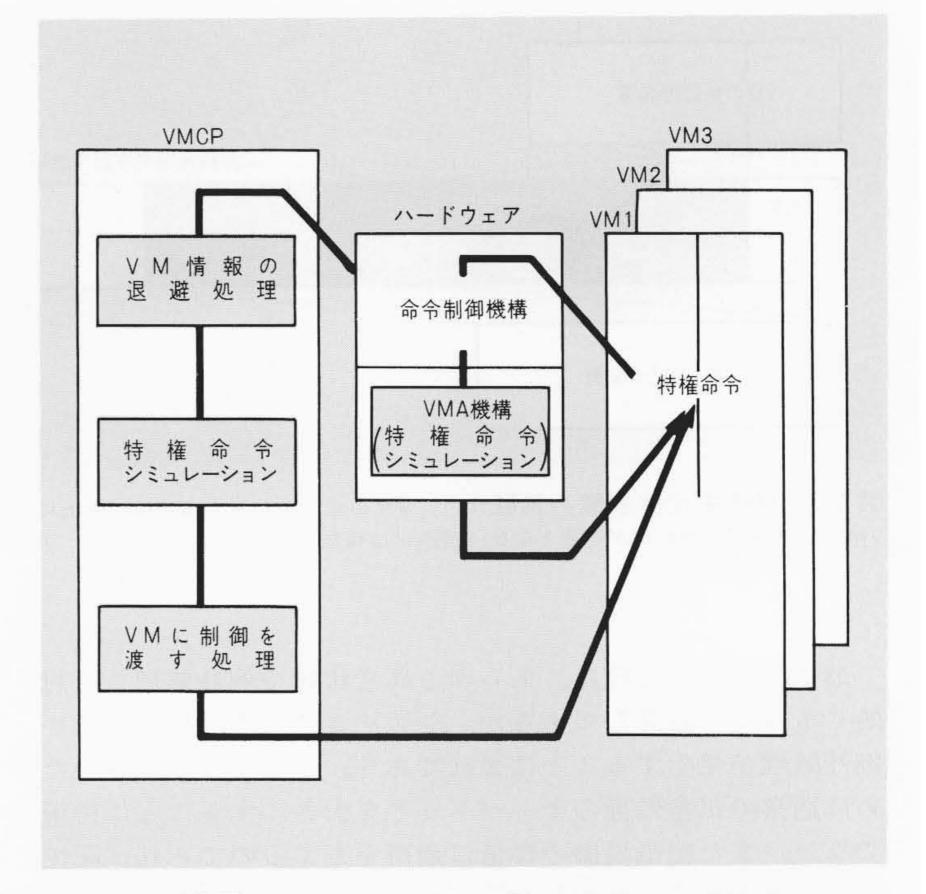


図4 処理装置のシミュレーション例 特権命令は計算機の物理的な制御を行なうための命令であり、VMから発行された特権命令をVMCPがシミュレーションすることにより VMの物理的制御が論理的に実現される。このシミュレーション処理の一部がファームウェアで実現される(VMA機構)。

(a) V = R VM

VMの主記憶装置のアドレスと実記憶装置アドレスが等しくなるように、実記憶装置の低位から連続して常駐化させる機能である。このための実記憶装置の領域をV=R領域と呼び、システムゼネレーション時に定義することによってVMCPのページング処理を不要にすると同時に、VMの入出力シミュレーション時の $CCW(Channel\ Command\ Word: チャネル指令語)変換を行なわずに処理を実行することができる。$

(b) 常駐VM

V=R VMは主記憶装置アドレスの対応関係を全く等しくするという制約から、VMS中でただ一つしか利用できないが、常駐VMとはVMの主記憶装置を実記憶装置上でV=R領域の次から連続的に常駐化する機能である。常駐VMはVMS中で複数個設定可能である。

常駐VMではVMCPのページング処理が不要となるのと同時に、実記憶装置アドレスがVMの主記憶装置アドレスに固定値を加えたものに等しいという性質を利用して、VMの入出力要求に対するCCW変換の方式を簡略化し、VMCPのオーバヘッドを減少できる。

(c) 部分固定VM

VMの主記憶装置の特定のページを実記憶装置上にページ固定する機能であり、特定VMに対するVMCPのページング処理を減少させ、当該VMの性能を改善することができる。

(3) 高速入出力シミュレーション

VMSがVM上でOSを実行するときのVMCPのオーバへッドのうち、VMA機構の処理対象となる特権命令のシミュレーション処理を除けば、その大部分はOSから発行された入出力命令のシミュレーション及び入出力割込みのシミュレーション処理である。

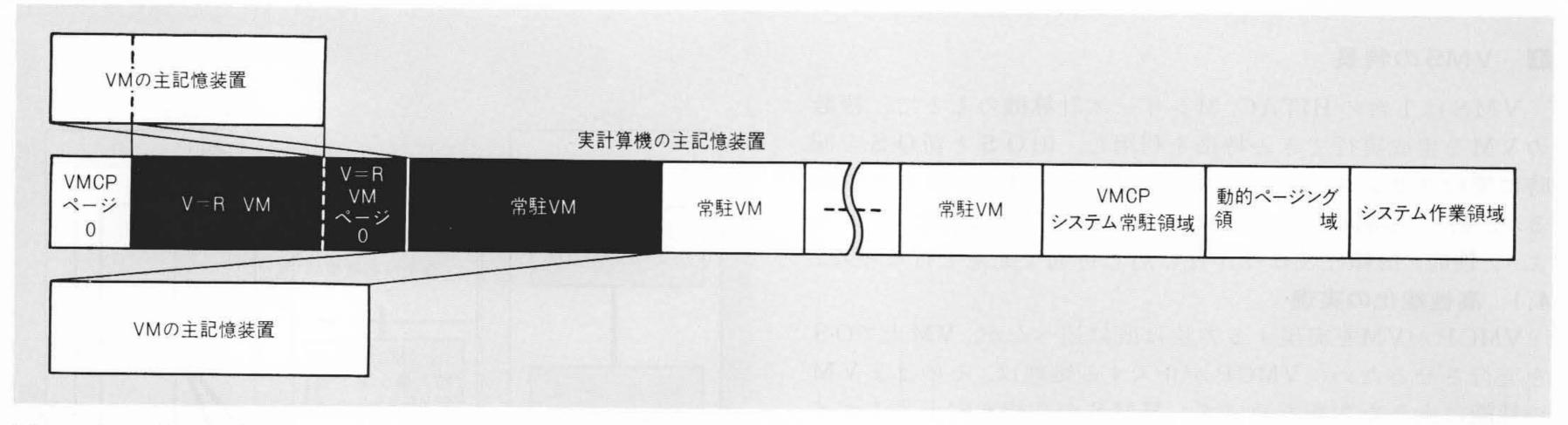


図 5 VMの主記憶装置の常駐化 VMの主記憶装置を実記憶装置上に常駐化することにより、VM実行時に VMCPのページング処理が不要となる。常駐VMは複数個指定できるが、V=R VMはシステムに1個しか指定できない。

通常の入出力処理は、あらゆる異常状態や例外処理にも対処できるよう設計してあるが、システム運転中、異常状態や例外処理が発生することはまれであり、このような処理のために通常の正常処理のオーバヘッドを大きくすることは得策でない。また記憶装置の容量に余裕をもてるのであれば記憶装置を多少余計に使用し、性能を向上させることも十分意味のあることである。VMSは、このような方針のもとで高速入出力シミュレーション機能を付加機能として開発した。

VMSは更に高速入出力シミュレーションの一環としてCCW変換の高速化を図っている。V=R VMに対してはCCW変換は不要であるが、常駐VMについてはCCWの変換処理が必要である。一般のVMの場合と異なり、常駐VMの主記憶装置は連続領域として実記憶装置上にとられている性質を利用し、先頭アドレスの実記憶装置の変位分を単に変換するだけでCCW変換処理の高速化と簡略化が実現された。図6は高速入出力シミュレーション機能を使用したときの処理の流れをブロック図で示したものである。

(4) 性能向上機能の効果

VMSは以上述べた性能向上のための諸機能を開発したが、性能向上のポイントは、VMCPがVMをシミュレーションするために使用するCPU(中央処理装置)使用時間をいかに減少させるかということであった。すなわち、VMCPのCPU使用時間を少なくすることは、VMの性能を実計算機の性能に近づけることと同時に、その分CPUの負荷を軽減できることになり、VMS全体のスループットを高めることになる。

図7は性能向上策対応のVMCPオーバヘッドの減少割合の 測定例を図示したものである。

この結果から、VM上での実時間処理業務の実行にも耐えられる性能が得られたと評価している。

4.2 高信頼性の実現

VMS が実時間処理業務に適用可能であるためには、VMS 下で実行される個々のOSによる運用と同等以上のRAS(Reliability, Availability, Serviceability: 信頼性, 可用性, 保守性)が必要である。

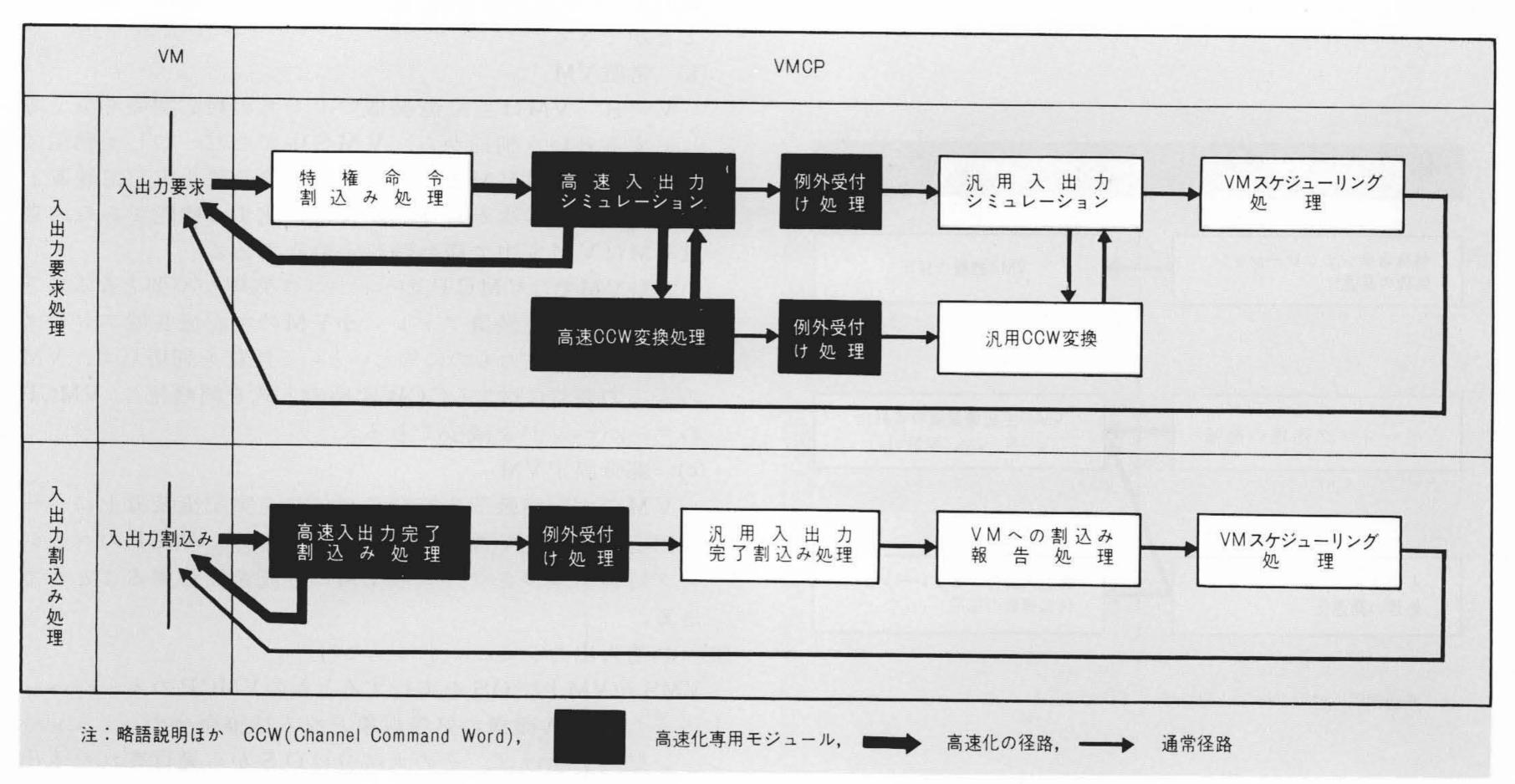


図 6 高速入出力シミュレーション実現図 入出力処理での異常状態処理を汎用入出力シミュレーションに任せ、正常に処理できる入出力処理を専用のシミュレーションモジュールとして分離し、処理の簡略化を図って高速処理を実現した。

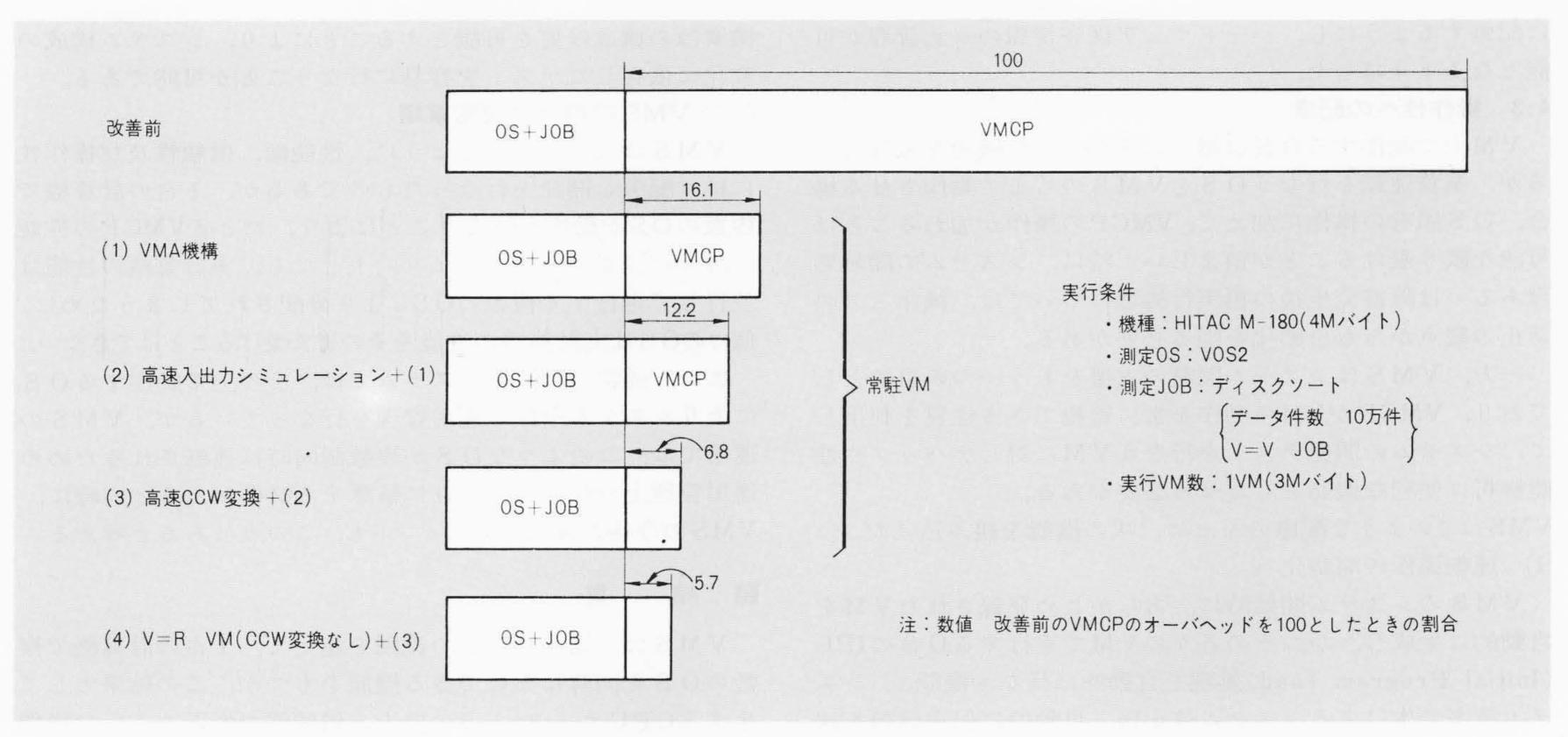


図7 性能向上策による VMCP オーバヘッドの改善割合の測定例 ドに対して、性能改善策を次々に追加適用したときの性能改善の割合を示す。

改善前のVMCPのオーバヘッ

VMSでは、OSが直接処理可能なVMSの構成要素の障害については、VMCPの介入なしに直接OSのRAS機能が適用できるようにした。

また、OS が直接処理可能でないVMCPのシミュレーション処理でのハードウェア障害に対しては、VMCP自身がOS と同等以上のRAS機能を備えることによって対処するようにした。特に、 $I/O(Input/Output: \lambda 出力)$ のハングアップの検出のためのI/Oパトロール機能については、時間監視によるタイムアウト検出後、OS がリトライを行なうことによってトランザクション処理時間が不当に伸びることを防ぐため、入出力装置の閉塞を行なって以後のI/O アクセスをリジェクトする機能を付加機能として開発した。更に、VMコンソー

ル装置の障害に対しては交代コンソール機能を用意したが、 VMS下のOSでの交代コンソールの切替えとVMCPによる 切替えとが同期して行なえるようOSとのハンドシェーク(OS とVMCPとが連絡を取り合って機能を実現する。)機能をサポートした。

以上のような制御方式から、ハードウェアの障害処理はVMS下のOSとVMCPとがそれぞれ行なうことになるが、実計算機システムのハードウェア保守はそれぞれのOSでとられる障害記録情報を寄せ集めてみなければならないのでは大変不便である。そこで、VMS下ではOSの採取するハードウェア障害記録情報、統計情報の記録及びVMCPが採取する記録も、原則としてVMCPがまとめてVMCPのシステム常駐ボリューム

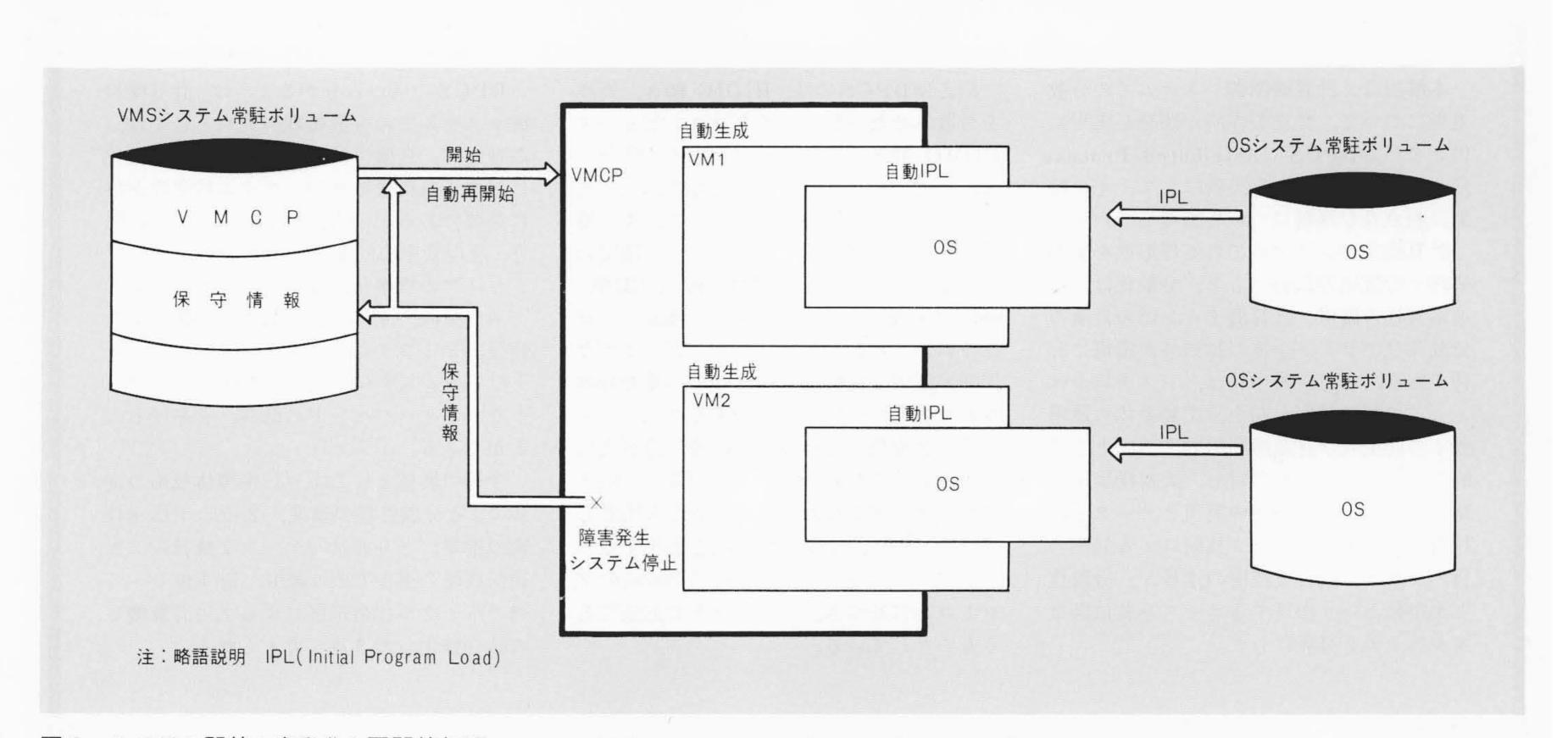


図8 システム開始の自動化と再開始処理 システム開始時にVMを自動生成し、その後指定された装置からOSを自動IPLする。システムを停止しなければならない障害発生時、保守情報採集後システムを自動再開始する。

に記録するようにし、ハードウェア保守情報の一元管理が可 能となるようにした。

4.3 操作性への配慮

VM上で動作するOSは個々に固有の操作機能をもってい るが、業務運転を行なうOSをVMSのもとで動作させる場 合, OS固有の操作に加えて、VMCPの操作が加わることは 可能な限り避けることが望ましい。特に、システムの開始処 理あるいは障害発生後の再実行処理については、操作ミスの 防止の観点からも自動化を図る必要がある。

一方, VMSはシステム開発の支援をもう一つの目的とし ており、VMCPがVMの動作を常に監視できる性質を利用し て、システムの開発テストを行なうVMに対しデバッグや性 能解析に便利な機能をもたせる必要がある。

VMSはこのような配慮のもとに、次の機能を組み込んだ。

(1) 運転操作の簡略化

VMSのシステム開始時に、あらかじめ登録されたVMを 自動的に生成し、かつその各々のVMで実行するOSのIPL (Initial Program Load)処理を自動的に行なう機能と、シス テム障害発生によるシステム停止後, 自動的に保守情報を採 集し、システムを再開始させる機能を組み込み、システム開 始時の操作の簡略化と、障害によるシステム停止から再開始 までの時間の短縮を図ることが可能となった。図8にシステ ムの自動開始と再開始処理の方式を示した。

(2) システム開発テストの支援

VMSでは、システム開発テストを日常の業務運転と独立 して行なうことが可能であり、 VMで実行するシステムのト レース情報の採集機能あるいはテスト対象システムの性能解 析を行なうための機能を開発し付加した。また、VMの計算 機資源の構成変更を可能とすることにより、システム構成の 変化に依存したテストを容易に行なうことが可能である。

4.4 VMS適用上の留意事項

VMSは以上記述したように、性能面、信頼性及び操作性 に種々配慮し開発を行なったものであるが、1台の計算機で 複数のOSが動作を行なうことにより、たとえVMCPの性能 が種々の性能向上施策により向上しても, 実計算機の性能は 並行して走行する複数のOSにより分配されてしまうために, 個々のOSは実計算機の性能をそのまま受けることはできない。

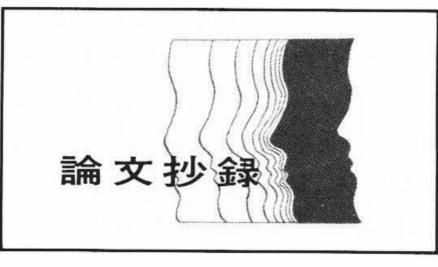
また、通常の計算機システムでは、その上で動作する05 によりシステム全体の運用管理を行なっているが、VMSの 運用では、このような O S が複数個同時に運転されるための 運用管理上の問題がある点に留意する必要があると同時に, VMSの今後の強化上のポイントも、この点にあると考える。

5 結 言

VMSは、従来のOSの範囲を超えて、1台の計算機で複 数のOSを同時に実行できる機能をもつが、この結果として 生ずるCPUオーバヘッドの増大, 信頼性の低下あるいは操作 性の劣化に対し,以上述べてきたような手段で解決を図り, システム移行やシステム開発作業に対する実用に供すること が可能となった。

VMSは昭和54年9月末にリリーズされ、既に数ユーザー で稼動を開始している。

今後、更に利用の拡大が予想されるが、VMSの特長を生 かした適用範囲の拡大, 更には機能の拡張など課題は多い。 広く読者各位からの御意見, 御批判をいただきたいと考えて いる。



制御用分散処理システム

日立製作所 平子叔男・平井浩二・寺田松昭 情報処理学会誌 20-4,346 (昭54-4)

本解説は、計算機制御システムでの分散 処理について、日立製作所が開発し実用に 供しているDPCS (Distributed Process Control System)を実例にして、その特 徴,利点及び課題について整理している。

計算機制御システムでの処理形態を集中 処理と分散処理に分けると,分散化は,高 速応答性の確保, 計算機ダウン時の危険の 分散及びプラント全体の拡張性の実現に有 利である。一方、集中化は、システム全体 としての制御効率の向上,工場全体の運用 効率の向上及び計算機利用効率の向上を目 的としている。これに対し、大規模集積回 路の発達による経済性の実現とデータ伝送 技術, 及びネットワーク技術による効率の 良いシステム統合化技術の進歩が、分散化 と集中化という相反するニーズを共に満た ロトコルに基づき、コンパクトで拡張性あ すシステムを可能にした。

例えばDPCSでは、HIDIC 80と、処理 を分散させた多数のマイクロコンピュータ HIDIC 08とを高速データウエイで結合し て, 分散の利点を生かすとともに, 次に述 べるネットワーク機能を用意して, 集中処 理の利点をも満たしている。保守・運用効 率向上のため、中央のHIDIC 80に、HIDIC 08に対するプログラム開発支援機能、プロ グラムのイニシャルローディングのような 保守・運用支援機能を持たせる。また分散 された処理を有機的に結合するために、メ ッセージ交換のための通信管理, 遠隔入出 力装置共用のためのデータ管理及び遠隔タ スクを制御するためのタスク管理を用意し ている。更に、これらの機能を実現するネ ットワークソフトウェアは、階層構造のプ るものとしている。

DPCSの例から分かるように、計算機制 御システムでの分散処理の利点としては, 高速応答, 危険分散, システムの大規模化 に対する負荷分散及びシステム拡張性が特 に重要である。一方、課題としては、(1)保 守・運用効率の向上, (2) ネットワークソフ トウェアの標準化,(3)ファイルやデータベ ースの分散,(4)システム故障の早期発見と 修復、(5)システム総合性能の把握、(6)ネッ トワークシステムのリソース管理,(7)ソフ トウェアオーバヘッドの低減、を挙げるこ とができる。

今後の展望としては、(1)半導体技術の進 歩による分散処理の進展、(2)データ伝送技 術の進歩, とりわけパケット交換技術, 光 通信技術の進歩とその適用、(3)上位レベル へのトータル化の進展による汎用計算機と の結合強化,の3点が考えられる。