

仮想計算機システムの運用とその問題点

小川 靖彦, 坂口 瑛, 中山 和彦 (筑波大学)

1. はじめに

筑波大学統合情報処理共有システムは、キャンパス内に多数の単目的の小型コンピュータを分散配置する無駄を省き、大容量ファイルをもつ大型コンピュータを中心として、「何時でも、誰でも、何処からでも」学内全体が有効に利用できることを目的とするシステムである。

この理想を実現するには種々の困難を伴うが筑波大学学術情報処理センター（以下センターと呼ぶ）では、一つの方法としてオンラインシステムを主体とし、日曜、祭日などの休日を除きシステムの連続運転を行っている。

従って、本システムは学術研究、学内事務・学務、図書館、新システム研究開発、運用システム改善作業等々多種多様の処理を行うことになるため、運用上種々の問題が発生する。

例えば、

- (1) 厳重な機密保護を要する処理
- (2) システム全体を占有する必要がある処理
- (3) 複数OSの稼働
- (4) システム資源の有効利用
- (5) 省エネルギーを目的とする縮退運転
- (6) 少人数による的確なオペレーション

等に関するもので、従来からの実計算機を使用して(1), (2), (3)項を満たすためには複数台の計算機を設置するか、或いは一台の計算機を時間帯を分けて使用するなどの方法をとることになる。その結果として、逆に(4), (5), (6)項については不都合となる。このような相反する要請を同時に満たし円滑な運用を行う目的で、センターでは1982年9月にFACOM M-200による仮想

計算機システム (Advanced Virtual Machine ; AVM) を導入したので、約半年間に亘る運用経験を基に現況と幾つかの問題点について述べる。

2. AVM導入の背景

これまで用いて来たACOSシステムでは処理能力、磁気ディスク容量の不足等により、爆発的な需要の増加に対処し切れなくなったため、種々検討の結果、ACOSシステムからFACOM M-200およびM-160Fから成る複合システムへ機種を変更し、1981年4月から運用を開始した。

このシステムは図1に示すようにM-200システム (2CPU: 主記憶12MB) とM-160Fシステム (主記憶6MB) が磁気ディスク (15.2GB) および大容量記憶システム (MSS: 102GB) を共有する疎結合システムで、OSは両者とも、OSIV/F4 (E-40) である。

M-200システムは、データベースおよび学術研究・教育のための処理を、M-160Fシステムは、ACOSシステムに蓄積したソフトウェア資産の、FACOMシステムへの変換をはじめ、学内事務・学務処理、貸出し返却を含む図書館業務処理をそれぞれ主な目的とする。

M-200システムは月曜日午前9時から土曜日午後4時まで (祭日は、当日午前9時まで) の連続運転を実施した。然しながら、この複合システムでは、

- (1) ACOSからFACOMへのコンバージョンを主目的とするM-160Fシステムは、業務の拡大に伴い処理能力の限界に近づいた。
- (2) M-200システムのソフトウェ

ア改良作業（主としてSE作業）を夜間運用停止しているM-160Fシステムを用いて行うことはほとんど不可能である。

- (3) 異なるOSを必要とする大学側のソフトウェア開発作業は、システムの運用時間帯外でしか実施できない。
- (4) 両システムのもつ能力を総合的に活用できず、また、互いに他システムのバックアップを行うことは不可能である。
- (5) CPU能力比で10倍以上も異なるシステムが磁気ディスクを共有している。

など運用上支障を来すことが多く、M-160Fシステムは導入後2年近くを経過し、ACOSからのコンバージョンと云う初期の目的を達成したこと、およびFACOM AVMシステムが運用可能なレベルに達したと考えられたこと、などの理由によりM-160Fシステムを撤去し、M-200システムの主記憶装置をこ32MBに、チャンネル処理装置を2台に、また磁気ディスク装置を40.3GBにそれぞれ増強して、1982年9月からAVMによる運用を開始した。

3. VMによる運用システムの構成

AVMは一つのMシリーズ計算機を複数台のMシリーズ計算機に見せかける機能をもつもので、この各々の仮想計算機をVirtual Machine (VM) と呼ぶ。

実システムの上にVMを構成するとき、システムレジデントボリュームをはじめ、複数のVM間で共有できず、それぞれのVMごとに専有しなければならない周辺装置があり、また、オーバーヘッドの面からも数多くのVMを設けるのは得策ではない。

センターでは現在、主システムを用いる業務を処理形態によって3種類に

大別し、それぞれ異なるVMの配下で処理を行っている。

全体の構成の概要を図2に示す。

VM1は科学技術計算、データベース検索など教職員、学生などが必要とする処理を行うもので、負荷が最も重く、筑波大学統合情報処理共有システムの目標に沿って先に述べた連続運転を行っている。

機密性の高い本学事務・学務処理はVM2の配下で独立に行ない、磁気ディスク装置も他種業務からの干渉を排除する目的で専用の制御装置を用い機密保護に務めている。

また、VM3はシステムレベルアップのためのSE作業や、新システムの開発など、従来運用時間帯外に実施していたものを日常業務と同一時間帯で可能とする目的で設けたものである。この種の作業はシステム負荷は殆ど零に等しいが、通電時間は極めて長いため、VM3を有効に利用できれば、省エネルギーの効果も極めて大きいものと期待される。なお、VM3は必要に応じて起動することにしており、定期的に運用されるものではない。

4. VMによる運用

AVMの適用分野は①新システムへの段階的移行、②システムのレベルアップ③複数OSの同時運用、等であると言われる。

これは、各VMが論理的に互いに独立であることから、定常業務を一つのVMで実行しながら開発的要素を含む作業を他のVMで、定常業務と同一時間帯内で実施できると言う、大きな特長と考えられる。それぞれのVMは他のVMとは無関係に起動・停止を行えるため、実計算機を複数台並べるのに比べて極めてコンパクトであり、かつ、リソースの構成変更もオペレータが走

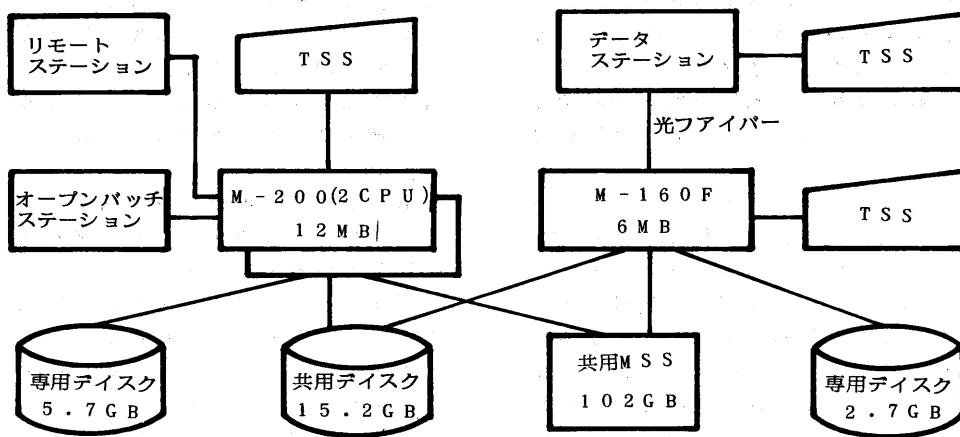


図 1. 複合システム構成

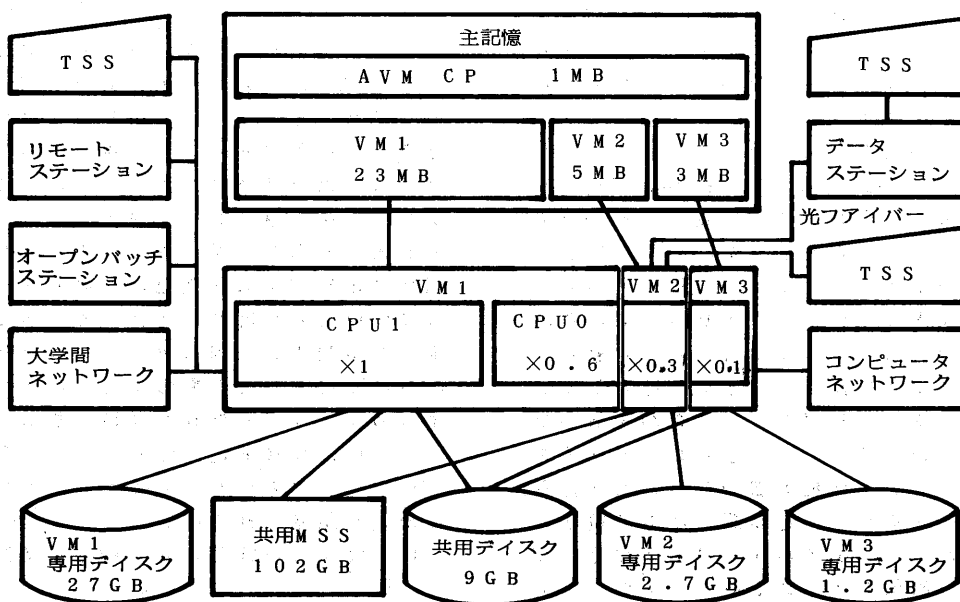


図 2. 仮想計算機システム構成

り回る必要はなく、簡単なオペレーションだけでシステム運用形態の変更が容易であるため非常に柔軟性のあるシステム構成が可能である。事実、当センターではすでに述べたように、一つの業務がVM1の配下で昼夜連続運転を行っているため、この効果は顕著である。

現在、当センターでは、FACOM M-200システムの他にMELCOM COSMO 800Ⅲ/MPおよび、同800Ⅲシステムが別々に稼働しており、これら三種類のシステムを8時30分～17時まで2人、17時～午前1時までを2人、午前1時～8時30分までを1人のオペレータによって運用しているため、オペレーションの簡易化、運用の自動化・省力化は極めておおきな課題である。VMとは直接関係はないが、例えばファイル保護のためのバックアップ作業も大きな問題で、M-200システムVM1の利用者が作成するファイルに関係する部分だけを磁気テープに吸上げるとしても、6250 RPI、2400ftの磁気テープで60巻程度となり、毎日、オペレータが磁気テープを用いてファイルバックアップを行うのは事実上不可能であるため、ADF (Automatic Data-migration Facility) を導入し、自動定時刻起動により、その日に新しく作成された、つたはアップデートされたファイルだけをオペレータ介入なしにMSSへ複写している。このADFによるファイルの自動バックアップ機能をVM2にも適用している。また、磁気ディスクの容量も不足しているので、やはりADFのマイグレーション機能を自動的に、毎日定時刻に起動し、一定期間一度も参照されなかったファイルをMSSへ複写した後、磁気ディスク上のオリジナルファイルを自動的に消去することによりディスクの空エリアの確保に努めており、

オペレーションの自動化・省力化と云う面で大きな役割を果たしている。また、特に機密を要する業務を単独に、一つのVM (VM2) の配下に置くのは、標準的なOSの保護だけでなくVMの独立性を強く期待したためであり、OS自身はVM1と同じOS IV/F4 (E-40) である。

他業務からの独立性を強調する一方で、VM2に共用ディスクのアクセスを許しているのは、例えばソフトウェア開発的要素を含み、特別に機密保護を強化する必要のない利用もふくまれるためである。VM3は、前述のように不定期的に起動されている。我々はまだVMを利用した新システムへの段階的移行の経験をもたないが、近々本年2月から3月にかけて磁気ディスク装置F496 (635MB×2/台) 14台を、F6421 (446MB×4/台) 13台に置き換える計画で作業を進めており、取敢えずVMを利用したファイルの段階的移行作業に注目している所である。

このようにして現在2～3のVMでM-200システムを運用しているが、或るVMを停止したとき、そのVMが専有していたCPU、主記憶などのリソースを他のVMに与えることによりシステムを再構成し、VMとしての能力拡大・縮小を容易に行なえる。また、例えば磁気ディスク装置などのリソースについても利用形態によって分類し、或るVM配下の一連のボリュームが使われなくなったとき電源を切ることにより省エネルギー効果も期待できる。

このような縮退運転として、現在のレベルのものを行っており、大きな効果が期待されている。

- (1) VM2の終了に伴う停止と電源断
 - (2) VM1における縮退運転
- ① オープンバッチステーション

- の入出力装置の切離しと電源断
- ② CPU1の切離しと電源断
- ③ DISKの切離しと電源断
- ・データベース関係、・共有ディスクの一部など。

一方、AVMシステムがまだ開発途上にあるためか、通信制御処理装置、磁気テープ装置などがVM間で共有できないため専用化する必要がある。高価な装置であるだけに、尚一層の開発努力が望まれる。

5. VMによる運用の効果と問題点

すでに述べたように、我々のセンターでは、1つのシステムで異質の業務を処理する必要があること、しかもこれらの業務の間では、通常のOSの機密保護レベルより高い機密性が要求されること、および24時間連続サービスを行うと同時に、新しいシステムの開発やOSのレベルアップなどのシステム作業を並行処理しなければならないことなどの理由により、オーバーヘッドという代償を払って、VMによる運用を、1982.9より、行ってきた。以下ではその効果と問題点について述べる。

(1) 異種業務の独立性

機密性の高い重要なファイルの内容等が、不特定多数の利用者に漏れることは、たとえ、システムの障害に起因するものであっても避けなければならない。もちろん、それぞれの業務に独立なハードウェアシステムを用意すれば、各業務間の干渉は避けられる。しかし、1つのシステムで異種業務を独立に処理しなければならないという条件が与えられた場合には、従来のOSの機能の範囲では、機密性を確保するという点では、まだ十分ではない場合があり、また障害等により、出てはならないところに情報が転送されるところとも生じ得ると考えなければなら

ない。このような場合に、VMの性質、すなわち、各VM間の独立性は、機密性の確保のために、有効であると思われる。我々のセンターにおける約5か月間の運用状況では、この点において、問題は生じていないが、引続き機密性が保持されるよう期待している。VM間の独立性は、機密性の確保ばかりでなく、1システム内で、異なった運用環境の設定が必要な場合にも有効である。しかしながら、次節に述べるように、現状ではオーバーヘッドが大きいのが難点である。

(2) オーバーヘッドについて

VMの導入により、かなりのオーバーヘッドの増加が予想される。オーバーヘッドの主因は、CP(Control Program)およびファームウェアにある。図2は、我々のセンターのシステムにおける例である。このシステムでは、2CPUのうち、1CPU(CPU1)は、VM1の専用CPUとして設定されており、図3は、CPが管理するもう1つのCPU(CPU0)の場合で、このときの各VMに対するCPU資源の割当比率は、VM1:VM2=70:30であった。これからもわかるように、オーバーヘッドの占める割合は、きわめて大きく、とくにCPU割当率の小さいVM2からみると、完全にオーバーヘッドの方が大きくなっている。このことは、利用者からみるとOSの下で得られる計算機能力よりはるかに小さい能力の計算機を使っていることに相当する。さらに、CPU割当率の小さいVM2では、著しくレスポンスタイムが長くなる。たとえば、LOGONコマンドについてある時の例をとってみると、VM1 3.8秒、VM2 9.3秒(TERMコマンドの処理を含む)であった。

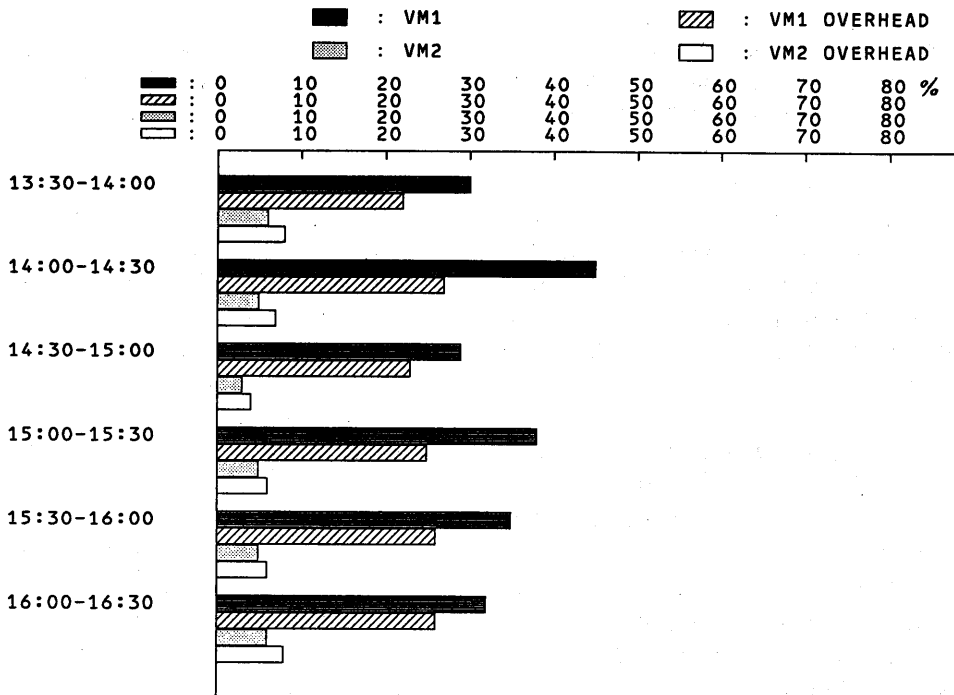


図 3. CPU USAGE OF VM (1982.10.21)

VMにおけるオーバーヘッドの問題は、その成否を制する重要な問題であり、抜本的な改善が望まれる。

(3) システム障害と各VM運用への影響

VMの導入に伴い、新たに制御プログラム(CP)がシステムで動作することになるが、これは各VMを統括して制御するものであるから、特に高い安定度が要求される。すなわち、一般に各VM内の障害は、各VM内に限定され、他のVMには影響を与えないが、CPの障害は、全VMに直接影響を与えるからである。我々のセンターの場合、1982.9~12の4か月間に、システムダウンに至った障害は、のべ19件であったが、これらのうち、CPに起因したものは、5件(26%)であった。この障害は運用中に発生し、全VMのシステムダウンとなった。障害は、一般に、過渡的なものが多い

が、VMを導入したことにより、新たな障害がシステムに持ち込まれたことになり、信頼性あるいはサービス性を一時的に低めている。

(4) 新システムの開発

VMの1つのメリットは、運用と並行して、同一時間帯に新しいアプリケーションシステムの開発やOSのテスト等ができることである。我々のセンターでも、この点を活用して、1つのVM(VM3)では、計算機ネットワークシステムの開発およびテストを行っている。計算機ネットワークシステムは、運用で用いているOSとは、一部分(VTAM)が異なったシステムを用いている。一般に、このような場合には、定常運用終了後に、OSの入換えを行って、作業を進めるのが普通であるが、我々のセンターのように、24時間連続運用を原則とする場合には開発等の作業は、不可能となる。しかし、VMを用いれば、この問題が解決

する。このことは、我々がVMを導入した動機の1つでもある。

(5) VMの導入と応答速度の改善

前述のように、VMの導入により、オーバーヘッドが著しく増加することが予想された。実際、1982.9、VMによる運用開始直後には、TSSコマンド処理のレスポンスがかなり遅く、年末のピーク時が懸念された。このため、VMにおけるCPU資源の割当比率の変更、ページング処理の適正化、コマンドにおけるI/O処理の削減、VIO適用、モジュールのメモリ常駐化等を徹底的に行い、VMオーバーヘッドの影響を極力少なくするようにした。(これらの改善項目のほとんどは、VMとは本質的には関係なく、通常のOSにおいて実現するはずの項目である。)

以下、ややくわしく上記改善項目を列記する。

- ① 当初、VM2におけるレスポンスが極端に遅い現象が見られた。1つの原因は、はじめ設定したVM2に対するCPU割当率(20%)が小さすぎたものと考えられたので、これを30%に増加した。VMに直接関係するものは、これだけである。
- ② ①に対しては、さらにページングが多発していることが判明したので、これに対する処理をした。
- ③ RACF (Resource Access Control Facility) 処理等におけるI/O処理が各コマンドで非常に多かったため、無駄なI/Oを削減した。
- ④ EDITのWORKファイルをVIO化した。
- ⑤ コマンドモジュールやコマンド内のモジュールをメモリ常駐化した。(このため、仮想空間におけ

るユーザ領域が小さくなった。)

⑥ FAIRSのシステムファイルの分散化、FAIRSのモジュールのBLDL化。

以上のような改善を行い、利用者へのサービス性の低下をできるだけ小さくするようにしているが、VMオーバーヘッドの改善がやはり不可欠である。

(6) CPの機能について

各VMは、CP(制御プログラム)のもとで稼働する。したがって、CPの機能は、システムの運用に大きな影響を与える。一例としてセンターの運用上必要なシステムの動作状況やパフォーマンスの監視機能がある。

現在のCPには、各VMの動作パラメータ等を知る機能はあるが、さらに、センターの運用上からは、CPの機能として、各VM内の動作状況を、1回の操作で把握できる機能が必要であろう。また、パフォーマンスの監視についても、現在は、各VMのコンソールから、各々独立に監視プログラムを走らせなければならないが、動作状況と同様に、各VM間の同期をとりながら、各VMのパフォーマンスおよび全体としてのシステムパフォーマンスの監視ができるCP機能が必要である。

6. おわりに

以上、数か月にわたる運用の経験から、少なくともVMの機能の面からは、その存在意義は、十分に認められる。しかし、さらに積極的に使用されるためには、オーバーヘッドを相当量減少させることが不可欠である。

最後に、本稿作成にあたり、資料提供等の協力をいただいた富士通㈱およびFHL㈱の関係者に、謝意を表す次第である。