

ミニコン複合体システム TMCS-40

三木雄作 東芝府中工場
産業用電算機設計部

1. はじめに

計算機制御の普及、オンラインリアルタイムシステムの普及に伴い、計算機システムの信頼性、処理性の各種要求に、従来の単一プロセッサシステムをもとにしては、対処できなくなりつつある。

例えば、従来でも信頼性の向上を計るためにデュアル、デュプレックスなどのシステムが使われてきたが、現用機と同規模の余分の機器を必要とし、ほぼ2倍の設備投資をしなければならぬ。これに対し、複合体は、この余分の部分をも積極的に活用しようというものである。

計算機複合体の歴史は古く、60年代初頭にさかのぼるが、ミニコンが出現し、各種ハードウェア技術の発達と相俟て、大量生産による性能/価格比の優れたハードウェアが供給されるようになり、近年再び注目をあびるようになった分野である。

東芝では、ミニコンTOSBAC-40Cを複数台用いたミニコン複合体システムTMCS-40 (Toshiba Minicomputer Complex System by T-40)を開発した。

TMCS-40は、プロセス・コントロールをはじめ各種オンラインリアルタイム処理を主目的とした汎用性のあるミニコン複合体システムである。

以下に、その概要を紹介する。

2. TMCS-40の処理方式

ミニコン複合体の具体例は、既に発表されたものも数多く、その方式は極めて多様であるが、その殆んどがハードウェアの結合様式を設計構想の中心としたもので、ソフトウェア体系まで当初より考慮されたシステムは少ない様である。

我々は、むしろ複合体システムの設計構想をソフトウェアの側から展開し、具体化することを目標に掲げ、TMCS-40の基本概念の設定を行なった。以下の項に詳述する様に、我々の複合体システムの骨格となっているものは、同一種のミニコン複数台を有機的に結合する「ホモジニアス (Homogeneous)」の概念、及び各ミニコンが自主的に仕事を取り出して処理してゆく「オートノマス (Autonomous)」の概念である。

この両概念により、

- あるミニコンの障害は他の正常なミニコンがフォールバックする
- ある部分の障害が他に波及しない
- ミニコンの台数を増やすことにより容易に機能拡張が行なえる

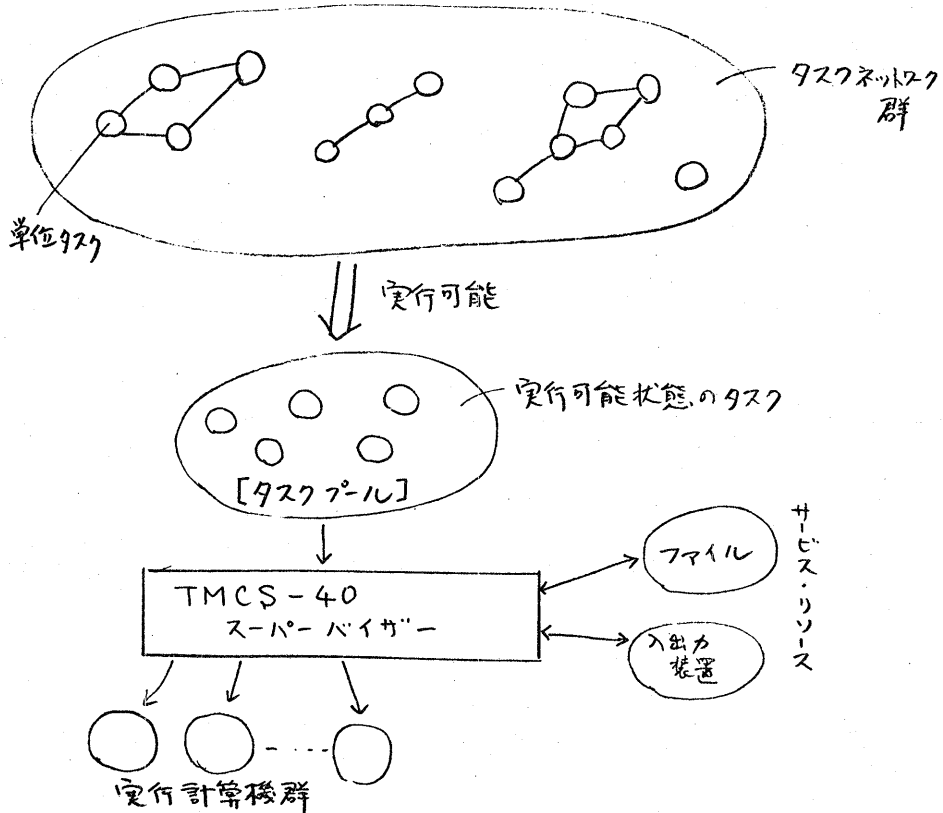
などを可能にしている。

(1) アルゴリズム

ユーザの目的プログラムや、システム機能の計算機処理に於ける実行基本単位をタスクと呼び、一連の業務の遂行に必要なタスクの集合が処理可能な実行順序の前後関係を、つながりとしてまとめたものをタスクネットワークと呼ぶ。この様なタスクネットワーク群は、通常プログラムリソースとして（フィジカルには、ディスク等にオブジェクトとして）用意されている。これらのうち、実行可能状態になったタスクの集合をタスクプールと称する。即ち、タスクプール中にあるタスクは計算機が割当てられればいつでも実行することのできるタスクである。

一方、フィジカルな処理機能を持つリソースとしては、同一機種の数台のミニコン（実行計算機と称する）がある。

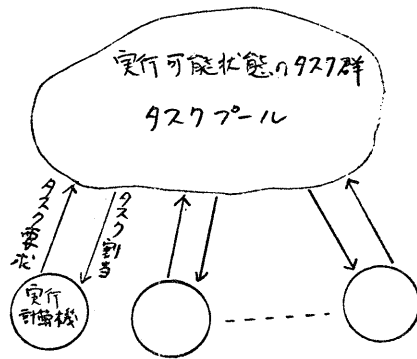
この両者を、タスクネットワークのトポロジーを守りながら、システムの基本管理機能（TMCS-40のスーパーバイザー）が結合してゆくことにより、システムの複合体としての処理が遂行されてゆく。



この結合は、次のようにして行なわれてゆく。即ち、実行計算機群の各計算機は自己の負荷状況を判定し、更に他のタスクを実行する余裕があれば、タスクプールからタスクを取り出し実行する。（具体的には、スーパーバイザーにタスク割当て要求を出す。スーパーバイザーはタスクプール中のタスクのうち最も優先度の高いタスクを要求を出した実行計算機に割当てる。）このように、各計算機は自分の負荷状態に応じて自主的にタスクを取り出して処理してゆくので、スーパー

CA
17

バイパーが特別なスケジューリングを行わなくても、システム全体の負荷は、各計算機にダイナミックにロードシェアされることになる。このようなアルゴリズムを「タスクプール方式」と呼んでいる。



タスクプール方式は、複合体の一部の計算機が故障した場合や、システムの機能拡張にも有利である。

即ち、故障した計算機はシステムから切離され、タスクプール中にあるタスクは残りの正常な計算機によって引続き処理される。故障した計算機が修理完了後システムに復帰すると再びタスクプールからタスクの取出しを開始する。

同様に、システムの処理能力を増すための拡張は単に実行計算機を追加するだけで行なえる。

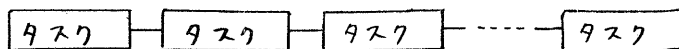
本アルゴリズムに於いて、各計算機がタスクプールからタスクを取出るとき各計算機は、そのタスクが一連のタスクの流れ（ジョブ）の中でどのように位置づけられているかについては全く関知していない。従って、ひとつのジョブの中で並列処理可能な部分をタスクとして構成すれば、それらは同時に複数の計算機上で並列処理されることになり、処理速度を向上させることができる。

また、内容的には同一のジョブを別タスクとして、別々の計算機に割当て、同期を取りながら実行すれば、信頼性を得るためのデュアル処理になる。

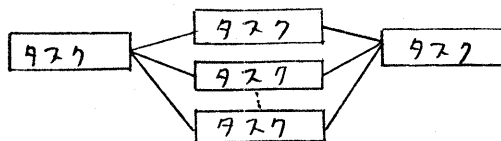
このように、並列処理、同期処理など、仕事の各種形態に柔軟に対応できることが、複合体のひとつの特長である。

<仕事の各種処理形態>

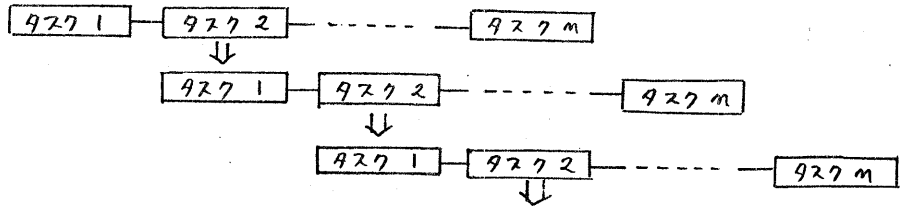
- ・ シリアル・タスク・プロセッシング



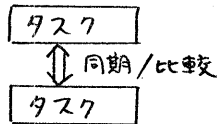
- ・ パラレル・タスク・プロセッシング



・ バイブライン・タスク・プロセッシング



・ デュアル・タスク・プロセッシング



(2) システム構成

前述のアルゴリズムに於て、各計算機は全く対等であり特別な機能を持ったものは存在しない。しかし、処理対象となるタスク、周辺装置、ファイル等のリソースはシステム全体で共通である。問題は、これを総合的に管理するシステムスーパーバイザを、ハードウェア上及びソフトウェア上どのような形態で実現するかであり、この辺が具体的なシステムの構成が大きく分れてくる。

TMCS-40に於いては、

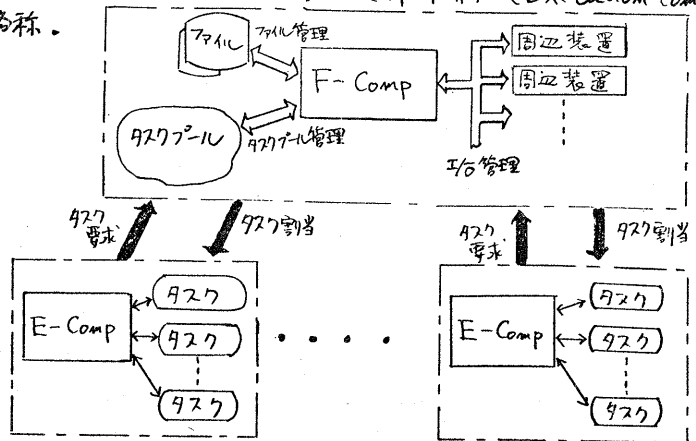
- ・ 各プロセッサのオーバードの節減
 - ・ 各種システムサービスのためのソフトウェアの簡潔化
- を図るため、システムスーパーバイザを特定のプロセッサ上に置く構成とした。即ち、
- ・ 計算機システムに課せられたタスク群がファイルされているディスクファイルを背後に持ち、タスクの割当てと、システムで共通な入出力装置の管理を主業務とする1台の「ファイル計算機 (File Computer)」

F-Comp と略称。

- ・ タスクの実行を主業務とする複数台の「実行計算機 (Execution Computer)」

E-Comp と略称。

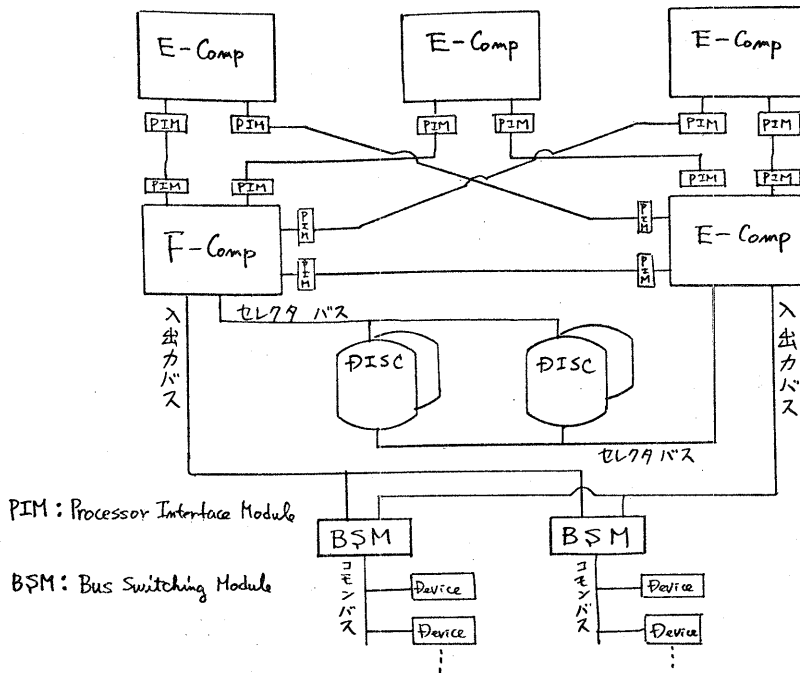
とから構成される。



C
A
17

3. ハードウェア構成

下図は、TMCS-40のハードウェア構成図である。



(1) 要素計算機

要素計算機として用いたミニコンTOSBAC-40Cの概略仕様は次の通りである。

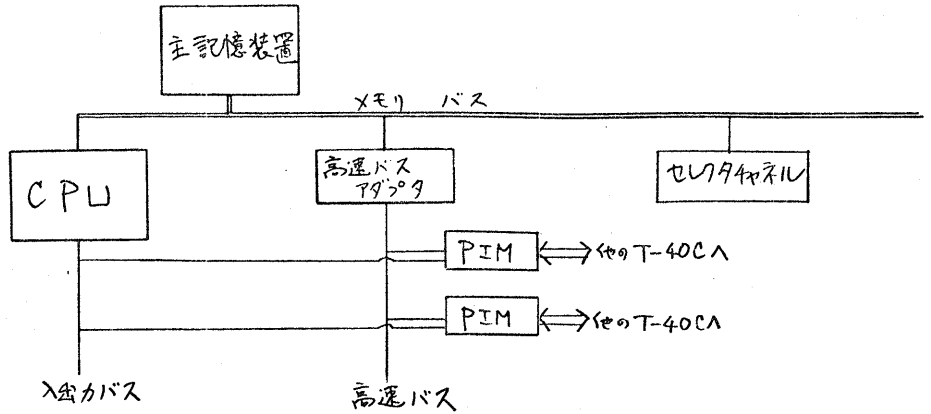
主メモリ	16ビット/〜64KB	割込レベル	内部 6
サイクルタイム	0.8 μ sec		外部 2位
命令数	113		
16個の汎用レジスタ			

(2) 計算機間の結合

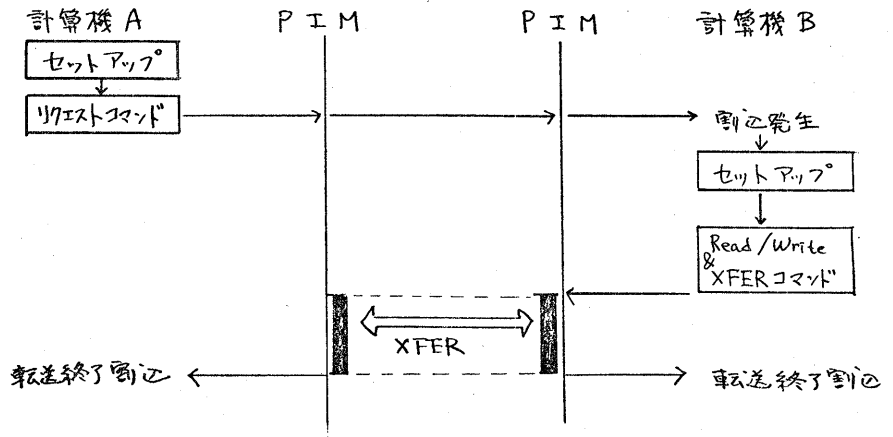
計算機間を結合するハードウェアとしてPIMを用いている。PIMの概略仕様は次の通りである。

・データ転送方式	16ビット並列転送
・転送速度	DMA (Direct Memory Access) モード 最大 500KB/秒
・転送誤り検出	ハードウェアによる垂直パリティの 付加と検出
・転送サイズ	2バイト単位で任意 (最大 64KB)
・転送距離	標準 20m

2台の計算機の接続には、1対のPIMを必要とする。接続図を次に示す。



PIMによる情報交換の流れを次図に示す。



CA
17

(3) 周辺装置の接続

周辺装置が接続されている計算機が故障した場合でも装置の利用が可能のように、バス切換装置(BSM)により、2台の計算機に接続している。BSMはプログラムによって切換えられる。BSMの概略仕様は次の通りである。

切換方式	ICによる半導体スイッチング方式 Auto / Manual の 2種類の切換が可能。
切換表示出力	プリント板内の発光ダイオードによる接続プロセッサ表示 及びリレー出力

(4) 共通ファイル

ファイルは、二重化が可能であり、複数台の計算機からアクセスできる。

4. ソフトウェア構成

TMCS-40のOS (オペレーティング・システム) 作成に当っては、次の点を設計目標として考慮した。

- ・ 耐障害機能を持つこと
- ・ 複合体であることを利用者に意識させないこと
- ・ 計算機の台数に左右されない構造であること

(1) OSの構造

TMCS-40は、F-Comp、E-Comp の2種類の計算機から構成され、各計算機には、その目的に応じたOSが備えられている。このOSをTMCS/RTS (Real Time System) と呼ぶ。TMCS/RTSの構成を次図に示す。



F-Comp OS

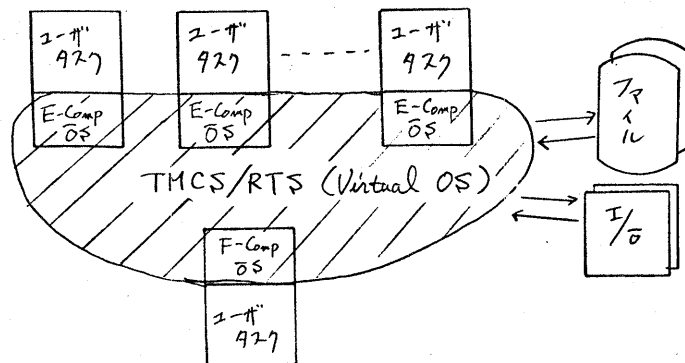


E-Comp OS

OSは「OS核」といくつかのシステムタスク、スーパーバイザール(SDC)および割込ドライバーから構成される。OS核はF-Comp E-Compとも同一でタスクのスケジューリング機能を持つ。F-Comp、E-CompのOSの違いは、OS核に結合されているシステムタスクの種類が違っただけである。システムタスクを変えたことによりF-CompにもE-Compにも成り得る。このような構成にしたのはOS作製の簡略化、OS核的の拡張性などのためである。

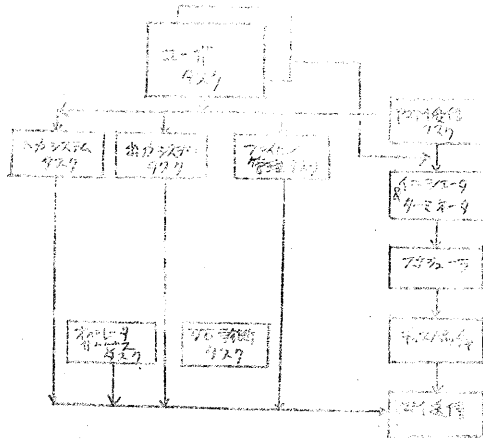
ユーザの作製したアプリケーションプログラムはユーザタスクと呼ばれ、OS核は、システムタスク、ユーザタスクの区別なくマルチタスク制御を行なう。

TMCS/RTSは、内部処理上F-Comp OSとE-Comp OSとに分割されているが、ユーザは各OS間のコミュニケーションなどの内部処理は一切関係しないので良い。ユーザが見た場合、TMCS/RTSは、1つのVirtualなOSである。

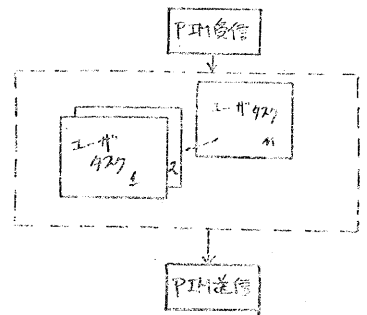


(2) システムタスク

システムタスクはOSの主要な機能を実現するOS構成要素である。OSの機能はシステムタスクの一連の流れとして実現される。F-Comp、E-Compに於ける主要なシステムタスクの流れを次図に示す。



F-Comp システムタスクの流れ



E-Comp システムタスクの流れ

(3) 計算機間通信 (F-E間通信)

計算機間通信はPPIを經由して行なわれる。通信モードには、16バイトの固定長レコードの転送と、任意長レコードの転送の2種類がある。

PPI通信は、OSの内部処理として必要の場合に行なわれるもので、ユーザが直接利用することはない。本システムに於いては計算機間通信はユーザにとって透明な状態になっている。

(4) 障害処理

計算機、ファイル、入出力装置等、施設群を構成する各要素の故障に対しフェールバック処理を行なっている。各要素が故障した場合、その要素をソフトウェア的に隔離し、オンライン保守ができるようにする。修理完了後のシステムへの復帰はオペレータとの合意処理により行なう。このように、オンライン状態での隔離/保守/復帰が行なえるため、システムMTBFは単一プロセッサシステムの数10倍と、飛躍的に向上している。

TMC5-40は、構成図が示すようにF-Compに成り得るのは2台で、このうち1台が通常F-Compとして働いており、他の1台はF-Compが故障した場合にのみF-Compとなり、番替はE-Compとして働いている。従ってこの2台が同時にダウンするとシステムダウンとなるが、信頼性解析より、二重以上のダウンの確率は極めて低く、これを許すために、ハードウェア及びソフトウェアを複雑にすることは実用的でないとの判断で、このような構成にした。

① 計算機の障害処理

計算機の故障には、F-CompとE-Compの二通りの場合がある。E-Compの故障は、その計算機が隔離されるだけで、他は継続して処理を行なう。以後の作業は残りの正常計算機が分担して行なう。F-Comp故障の場合には、E-Compの2台もまた新たなF-Compとなるため、OSの自動再編成が行なわれる。その後ユーザのイニシャライズプログラムが実行される。この場合でもイニシャライズプログラムの作り手による判断実行が可能である。

② ファイルの障害処理

ファイルは二重化されており、障害に対処している。

③ 入出力装置の障害処理

CA 17

故障した装置を分離し、オペレータにメッセージで知らせる。出力装置の場合代替装置が指定してあれば、自動的にとぎれに対して出力を行なう。

④ 時向監視処理

すべての周辺装置、ファイル、PEMに対して時向監視を行っており、一定時向内にハードウェアの応答がない場合、故障と見做して障害処理を行なう。

⑤ 自動再起動処理

電源断後、回復した場合、自動再起動モジュールにより、OSがイニシャルロードされ稼働を始める。

(5) TMCS-40の利用形態

ユーザプログラムの記述は、通常の単体計算機システムの場合と本質的な差はない。即ち、FORTRAN言語、アセンブラ言語、スーパーバイザーコール(サブマクロ)などを使って記述する。

但し、複合体システムが有する複数台の計算機を有効に利用した並列処理、パイプライン処理、同期処理などを行なう場合は、タスク向の同期制御、データの受渡しなどが必要となる。特に同期制御は複合体の特長の1つである。

TMCS-40では、タスク向の同期制御の手段として、セマフォアのP-オペレーション、V-オペレーションを用意している。この2つの基本命令の組合せにより、一斉同期、順序同期、排他的同期などの処理を簡単に記述することができよう。

TMCS-40の下でのプログラムの言語処理、ライブラリの登録、テストランはフリータイムシステム(TMCS/FTS)により行なう。

FTSはリアルタイム処理の空き時向(フリータイム)を利用して、言語処理などを行なうため、リアルタイムジョブに対する影響は殆んどない。

TMCS-40はユーザに複合体であることを意識させないことを前提に開発した。そのためデバッグガイドについても、TMCS-40の内蔵構造とは独立したものを用意した。

これらのソフトウェアに依り、ユーザは1台の計算機システムを利用する感覚で、複合体システムを利用することができよう。

5. おわりに

以上、ミニコン複合体システムTMCS-40の概要を述べた。ミニコン複合体システムについては、未だ定着した概念はないとされており、各所が種々の形態のものがある。我々は、実用ということを一に開発を進め、断片として考えた部分もある。これらは今後、実用上の経験を基にし、更に検討を重ねてゆきたいと考えている。

複合体システムの特長的な利用法の1つに並列処理があるが、計算機がこの世に現れて以来、利用者はシリアル的な利用方法に慣らされてしまった。今後、複合体システムをより有効に使うためには、並列的利用技術及び並列プロセスを容易に記述できる高級言語の開発などが望まれている。既にいくつかの試みが行われているが、これらは複合体システム一般の今後の課題である。

参考文献

- 1) 村上他：小特集コンピュータ・コンプレックス，情報処理，vol 15, No 7
pp 524-564 (1974)
- 2) E. W. Dijkstra: The Structure of the "THE"-Multiprogramming
System, CACM, vol 11, No 5, pp 341-346 (1968)
- 3) C. V. Ramamoorthy, M. J. Gonzalez: A Survey of techniques for
recognizing Parallel processable streams in computer programs, FJCC (1969)
- 4) 田丸他：マルチ・ミニコンピュータ・システム，情報処理学会第15回プ
ログラミング・シンポジウム報告集，pp 240-251
- 5) 板本：制御用マルチコンピュータの一方式における応答性改善について
，電気学会論文誌 95-C 8 (1975)
- 6) 山中他：プロセス制御用複合体システム，昭和49年電気学会全国大会
シンポジウム報告集 pp 271-274 (1974)
- 7) E. Youdon: Reliability of Real-time Systems, MODERN DATA Jan.
Feb, March (1972)