

複合計算機 FSK-II の結合方式

林 岑男、藤野卓男、中原義大、(三菱フレイションKK)

堀川勇壮、渡辺 顕、原田公一、(航空宇宙技術研究所)

§ 1. はしごき

ミニコンピュータ複合化の主なめらいは、ミニコンピュータが単体では処理しきれないような高度な応用分野に、複合化することによって対処しようとするものである。言うまでもなく、そのような分野に大型計算機を導入するよりも、はるかに経済的、即ち、価格性能比がすぐれていることが前提となるが、特殊な応用分野によっては、単一の大型計算機でも不十分な場合がある。

ミニコンピュータと複合化することによって、経済性ばかりでなく、計算機ストラクチャとしての柔軟性、拡張性、信頼性、可用性が、比較的容易に得られることも重要なめらいである。

筆者等が製作した複合計算機、FSK-IIの用途は、システムに要求されるその実時間性の観点から、まさにミニコンピュータ複合体の良さと、充分に活用できる分野であり、複合体システムの並列処理機能による速度と、分散同時処理による性能とに主要なポイントを置いて、システム作りを試みた。

基本的な構成は、ミニコンピュータ、MELCOM-70を要素機とする同種結合、超高速データ転送を意図して集中化したセルラ、システム、そして要素機に大巾な手を加えないという制約の下に、主記憶を共有しないバス接続方式を採用した。結果として、計算機間のデータ転送が 826 KW/s の高速で、安定に動作するシステムが得られたので、以下に報告する。

§ 2. 結合方式

複合体の性格を、最も端的に特徴づけるのは言うまでもなく、その結合方式である。複合体の定型はないので、システム構成に際しては、利用目的に応じた接続方式を先ず決定し、それに最適制御方式を決めることになる。

接続方式を定めるに当たって、考慮した事項は、以下のような諸点である。

- (1) 複合体が比較的広域に分散してネットワークを構成する場合と、相互の伝送路長が、数〜数十米以内に配置される場合とでは、問題点が大きく異なる。FSK-II は後者に属するので、データ転送速度を優先し、密な相互結合と目的とした集中型とすること。
- (2) 特定の要素機を制御中枢と定め、それに制御プログラム (SYSTEM CONTROL PROGRAM) を格納してシステムを運転する集中制御方式では、一般に階層構成をとることになり、この場合、ハードウェアの構成は簡略化できる。

一方、制御プログラムを特定の要素機に固定せず、複数に分散させる分散制御方式がある。ハードウェアの配慮があれば、集中制御方式は、分散制御方式の1つの運用形態として包含することができ、システムの柔軟性、可用性に優るので、FSK-II では分散制御と考慮する。

- (3) 分散制御とすることにより制御権の競合、その確立が問題となって来る。FSK-II の運用では、大規模なオンライン・シミュレーションに対しては、階層制御とし、その他の用途では、2つ以上のフォアランド・ジョブと、1つのバックランド・ジョブを同時多重に遂行できる機能を持たせる要求があり、そのために、後述するグループリングの機能を付加して、制御権の制御と円滑に行なうこととした。
- (4) もっと小さなレベルで、ハードウェアの機能を考えると、インターコンピュータリングとして、制御を集中的に行なう方式と、コンピュータ・アダプタの形式で、制御を分散させる方式とがある。
集中制御の場合は、ソフトウェアの手続き(プロトコル)を含めて、制御が比較的簡単になるが、信頼性や、拡張性に限界がある。しかし、複合体として密な結合を意図すれば、必然的に集中制御となり、むしろ積極的にソフトウェアのオーバーヘッドを、ハードウェアで吸収するように考慮した方が良く考えられる。
- (5) 結合路の形成は、多重構成とする。オンライン・シミュレーションでは、特にI/Oデータの処理が問題となるが、ジョブの分散を行なった場合には、同時多重転送(フロートキャスト)の機能が有効となる。結合路は、低、高速バスとも、多重化をほかる。
- (6) 転送路の確立、要素機間に分散されたジョブの同期は、割込によって行なう。実時間性の稀薄なジョブの同期は、勿論メッセージ同期方式等によって可能である。データ転送は、低速バスは一語単位、高速バスは、ブロック転送とする。ブロック転送のデータ・サイズは、バスの占有、待時間の問題から当然可変長とし、問題に応じて分割転送を考慮する。
- (7) 複合体としてのI/O周辺機器の配分も重要な課題となる。資源の共有を意図すれば、I/O制御機と特定の要素機に定めるか、或いは要素機へ自由にスイッチング配分と許す方式が必要となる。自由度の大きさから考えると後者の方が希ましく、且つ利用の多重度を考慮すれば、I/Oバスの多重化をほかった方が希ましい。

以上、主要な事項を検討して定めたFSK-IIの接続制御方式の要点は、以下のようになるものとなった。

- ・ バス接続方式(センターバス・カプラ方式)
- ・ 遷移可能な分散制御方式
- ・ 多重転送路の形成(選択可能なフロートキャスト方式)
- ・ 割込による転送路の確立(独立要求方式)
- ・ データ転送は非同期、バースト転送
- ・ 複数セレクトによるシステム資源の多重配分

§ 3. システムの構成

全体の構成ブロック図は、関連報告に示した通りである。要素計算機間の接続部をICCL (INTER COMPUTER COMMUNICATION LOGIC) と呼び、最大8台の要素機が接続できる機能を持たせ、実装は5台である。

図-1には、結合路の略図を示してある。即ちICCLには、LSBC (LOW SPEED BUS COUPLER, ---- PIO バス) と、HSBC (HIGH SPEED BUS COUPLER ---- DMA

バス)を独立に持ち、計算機間に2つの完全な並列転送路を設けた。更にこの2つのバスカプラの他に、独立した3重のデバイス・バス・セレクトを設けて、要素機とデバイス・バスの自由な接続を可能にした。この3つのセレクトの1つには、FSK-II個有の運用要求からCRTを含むオペレータ・コンソールだけを接続し、後述するような制御機能を持たせ、他の2つには、それぞれ独立して充分機能するような機器配分で周辺機器を接続している。オンラインI/Oリンクージが2ヶあり、それぞれ要素機に直結してある。I/Oリンクージは当初デバイス・セレクトに接続していたが、実時間シミュレーションに際してデバイス・バスを占有してしまうことが明らかになったので、シミュレーション時の制御機に直結、固定することにした。

I/Oリンクージには、ANALOG IN/OUT それぞれ約250CH、DIGITAL IN/OUT それぞれ約350CHが実装され、DMAベースで計算機とインターフェイスがとられている。システムの運用を円滑に行ない、且つインタラクティブなマン・マシン・コミュニケーションを実現するために、FSK-II個有のオペレーション・コンソールが設けられているが、この機能には、本システム運用の基本的な考え方が盛り込まれているので、簡単にその概要を述べておく。

(1) デバイス・バスの手動、自動切換機能。3つのデバイス・バスをそれぞれ任意の要素機に手動接続する機能がある。オペレータ・コンソール自身も例外ではなく、この制御下におかれる。自動モードでは、制御機の切換指令により動作する。

(2) システム全体の運転指令、各要素機の電源スイッチ、NITロック、IPLスイッチランステータス、エラーステータスをリモート化して設けてあり、一括して各要素機の起動、停止、運転の制御が可能である。

(3) システムの生成、オペレータ・コンソールが接続されている要素機(制御機となる)のIPLを起動する。その結果、システム・ファイルを持つドラムより、制御プログラム(SCP)が、

その要素機へロードされ、ロード終了と同時に要素機はシステム制御機となってモニタ・モードで運転を開始する。その状態で他の要素機のIPLを起動すると、その要素機はタミー・ランを開始し、DMA経由で制御機からの転送が受け付け可能な状態となる。以下SCPのプログラム・ジェネレーション機能と、特殊ファンクションスイッチの操作により、自由なシステム生成が行なわれる。システム生成には、ロード・シェア、要素機の組合せリンク、運転指令などが含まれる。

(4) グループリンク
既に述べた如く、99重同時

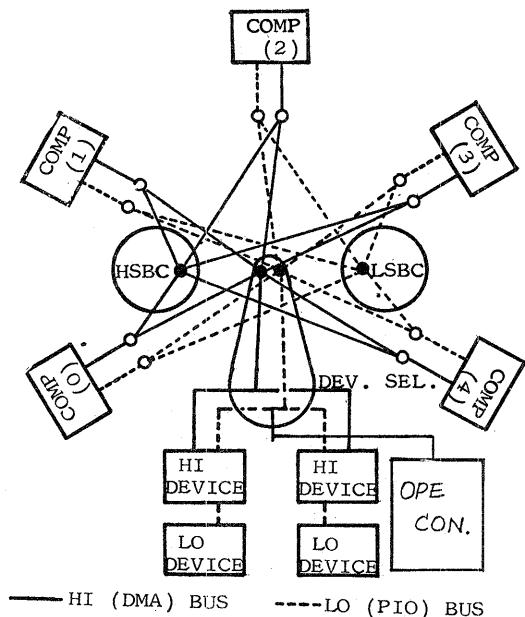


図-1

処理を行なう場合には、要素機をシェアし、資源を共有することになるが、このような運用でプログラムの暴走その他により、他のジョブを妨げるようなことがあってはならない。妨害が発生する可能性は、ICCLの制御権を行使した場合に限られるので、その使用を制限し、異常を検知する手段としてグループ設定機能を設けた。既に、任意の要素機を指定してグループに分ける（最大3グループまで）機能と、各グループに1ヶだけ制御機たり得ることを許し、指定する機能と、グループに属さない要素機への転送アクセスを検知する機能とを設けた。グループを解除することは任意である。

§ 4. 複合化のサブ・システム

4.1 LSBC

すべての要素機のPIOバスは、LSBCを通じて並列に接続されている。転送制御の手段として、各要素機側にステータス・フラグ・コントローラと呼ぶモジュールが設けられており、それぞれ機器アドレスが割当てられている。このモジュールには、1語の出力レジスタ、転送ルート設定回路、割込発生回路が含まれている。(図-2)にしたがって、各要素機は表-1に示すPIO命令と用いて任意の要素機へ1語の情報を転送することが出来る。各モジュールが要素機毎に独立しているため多重転送が可能になり、転送の様子は図-3に示す如くである。即ち、

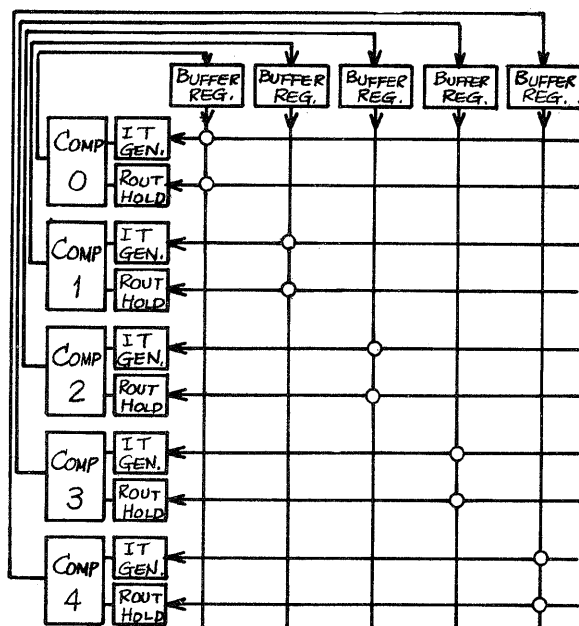


図-2

表-1

命令	内容
SNS0	ステータス・フラグに向けてデータを送出可能か、確認
SNS1	ステータス・フラグからデータ読取可能か、確認
RDA	ステータス・フラグから1語のデータを読取る。
RDB	ステータス・フラグへ転送されたデータの転送先取込状態と読取る。
WRA	ステータス・フラグへ1語のデータを書込む。
WRB	ステータス・フラグへデータ転送先の指定と書込む。
CNT2	ステータス・フラグの状態と初期状態にセットする。
AKIT	割込受付
MKIT	割込マスク セット
RDMK	割込マスク 読取。

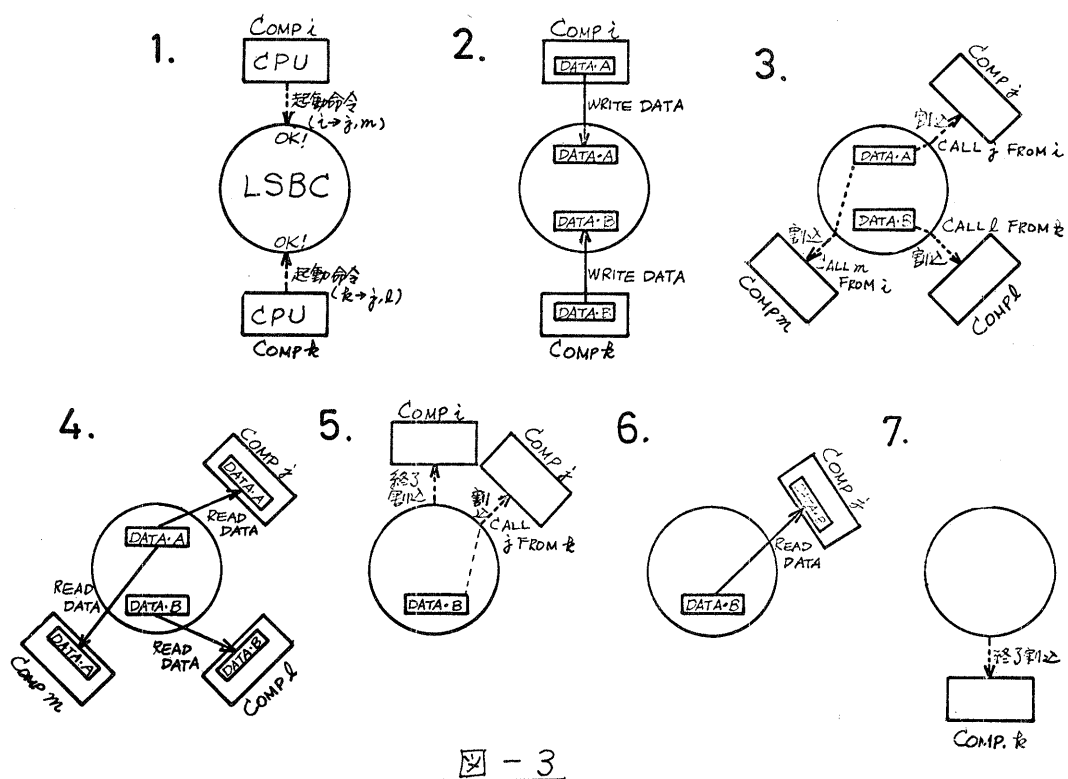


図 - 3

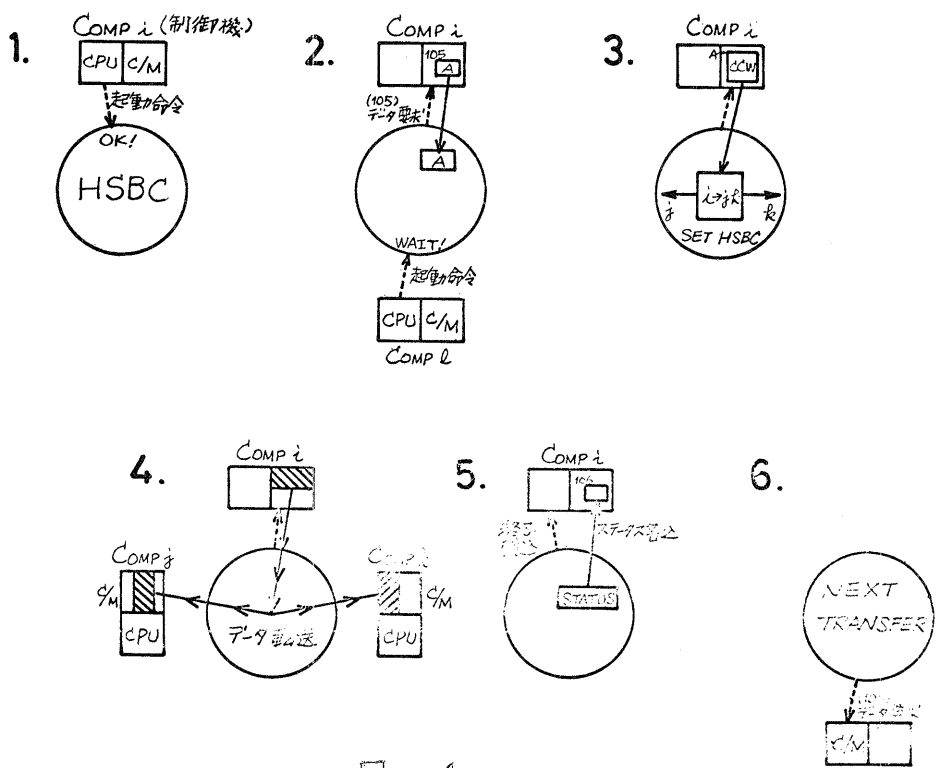


図 - 4

- (1) 接続要求元が、接続先情報を送出する。
- (2) カプラはその情報に従って接続先を決定し、一方要求元は1語のデータを送出する。
- (3) データが送出されるとカプラは転送先へ割込をかけて知らせる。
- (4) 転送先では、これを受けてデータを読取り、カプラは要求元へ終了割込をかける。
- (5) 接続要求の競合が発生した場合には、先着順、同時であれば要素栈ナンバーの若い順に優先度がつけられており、低いものは待ちに入る。
- (6) 各要素栈は割込マスク、ビットを制御することによりデータ転送を禁止することが出来る。また割込モードではなくセンスモードでの転送が可能である。

以上で明らかなように、各要素栈は自己の関与しない接続を制御することは出来ない。即ち直接制御である。また何等かの理由(転送路のフェイル、転送の禁止など)により、転送が完結しない場合には、一定時間応答かない時間弁別によって検知する。

4.2 HSBC

LSBCの場合と同じように、すべての要素栈のDMAバスは、HSBCを通じて並列に接続されている。LSBCの場合と異なり、HSBCには栈番アドレスが唯一だけ割当てられていて、制御のために表-2に示す9種の命令が用意されている。データ転送の制御は、制御栈内のメモリに形成されたCCW(CHANNEL CONTROL WORD)に従って行なわれ、(表-3参照)要点は、図-4に示す如く以下の通りである。

- (1) DMAバス転送路は、同一時間内では、1つの要素栈によってのみ制御される。それを制御栈と名付けるが、すべての要素栈は、制御栈の宣言を行うことによって、制御権をもつことができる。(ただし、制御栈たり得ることを認可する機能は、前記オペレータ・コンソールに持たせている。)
- (2) 制御権の宣言は カプラに対して、起動命令を出すことにより行なわれる。起動命令に競合が発生した場合には、2レベルに分けて優先度が定められており、同じレベルであれば要素栈ナンバーの若い順に優先度が定められる。
- (3) 制御栈が決定すると制御栈のDMA転送サカプラに向って開始される。先ず制御栈の固定番地(105)からCAW(CHANNEL ADDRESS WORD)を読出す。カプラは更に制御栈のCAW番地より12語のCCWを読出す。

表-2

命令	内 容
SNSO	カプラのBUSYフラグの確認
SNSI	カプラのDONEフラグの確認
RDB	制御栈許可状態の読取、DMAインビット・フラグの確認
RDC	カプラ・ステータスの読取
WRC	カプラの起動、DMAインビット・フラグのセット・リセット
CNTZ	待ちに入っている起動命令のキャンセル、その他のステータス・リセット
AKIT	割込受付
MKIT	割込マスク・セット
RDMK	割込マスク読取

- (4) CCWによって規定されるものは、転送先(1)および転送先(多)計算機の指定、転送の種類および語長、定数転送の場合の定数、データ転送の各要素機毎のトップアドレス、転送チェイニング情報である。
- (5) カプラはこのCCWを解説して、所定のバス結合を行ない、制御機とは無関係にシステムのデータ転送を開始する。制御機の役割はCCWを形成し、起動命令を出すのみであり、転送に関与するか否かに係わらない。即ち間接制御である。転送が完了すると、制御機へ終了割込みが返される。
- (6) データ転送は、DMAベースで行なわれるため、CCWに指定された各要素機は、転送の開始と終了を感知し得ない。このためにも、先にのべたグルーピングの効果があるが、各要素機は独自に転送禁止フラグをHSBCにたてること出来る。
- (7) 高速データ転送を実現するために、

図-5に示すようなバッファレジスタが設けられている。転送元計算機から送出されたデータは、入力ポインタの指示により8ケのバッファレジスタに順次送られ、カプラは転送先計算機へ出力ポインタを示して読取要求を出す。出力ポインタは入力ポインタを追う形で循環する。計算機間の非同期性に起因するロスはこのよって吸収し、ほぼ要素機のDMA転送能力に近い速度を維持している。

- (8) 転送中のエラー検出は表-4に示した範囲で行っており、エラー・ステータスは、DMA経由で制御機の固定番地に書込んだ後、カプラが制御機へ割込みをかけて通知する。

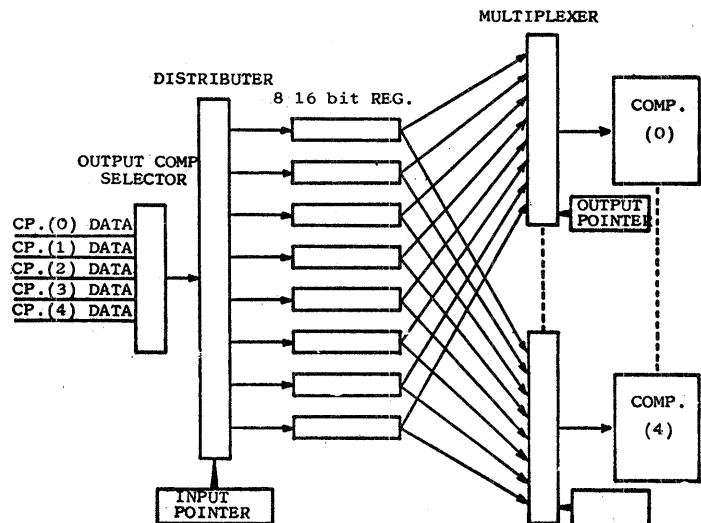


図-5

WORD 1	$s_7, s_6, s_5, s_4, s_3, s_2, s_1, s_0, d_7, d_6, d_5, d_4, d_3, d_2, d_1, d_0$
WORD 2	m_1, m_0 TRANSFER LENGTH
WORD 3	COMP(0) C/M TOP ADDRESS
WORD 4	COMP(1) C/M TOP ADDRESS
WORD 5	COMP(2) C/M TOP ADDRESS
WORD 6	COMP(3) C/M TOP ADDRESS
WORD 7	COMP(4) C/M TOP ADDRESS
WORD 8	COMP(5) C/M TOP ADDRESS
WORD 9	COMP(6) C/M TOP ADDRESS
WORD 10	COMP(7) C/M TOP ADDRESS
WORD 11	CONSTANT NUMBER
WORD 12	c NEXT CCW ADDRESS if chain

NOTE: $s_0 - s_7$ OUTPUT COMPUTER
 $d_0 - d_7$ INPUT COMPUTER
 m_0, m_1 MODE
c CHAINING FLAG
COMP(5,6,7) FOR EXPAND

表-3

表-4

エラーの種類	内 容
Rate Error	転送速度の異常。1語毎の転送間隔を測定し、或る値をこえた時。
Memory Parity	要素栈か、データ送出時にパリティ・エラーを生じた時、カプラにて判定。
ICCL Parity 1	要素栈(センタ)とカプラ間に発生したエラー
" 2	カプラと要素栈(レシーバ)間に発生したエラー
DMA Inhibit	制御栈が指定したレシーバのDMAインヒビット・フラグが立っている時
Illegal END	転送終了時に、送受の語数が一致しない時
Illegal Command	コントロールワードが不正な指定を行なった時
COMP Combination	オペコン上で指定されたフル・フーズ群をこえて転送指令があった時
Illegal Control	制御栈としてオペコン上で許可されていない要素栈が制御権を要求

4.3 デバイス・バス・セレクト

既に述べたように、多重同時利用に、周辺機器を効果的に利用する目的で、3ヶの独立したデバイス・バス・セレクトを設けた。このようなセレクトにより共同利用を計るのが良いのか、或いは、要素栈と周辺機器、特にドラムやディスク等高速度機器を密結合にした方が良いのかは議論の余地がある。FSK-IIに於けるこのセレクトは、バスカプラと異なって、1つの要素栈と1つのデバイス・バスの転送路を、物理的に接続するスイッチング機能をもたせたものである。

このセレクトには高低両バスを含んでおり、1つのセレクトに最大10ヶまでのデバイスを接続することができ、セレクト・モードには、手動と自動接続のモードがあり、自動モードでは、自己の関与しない接続を制御することもできる。

5.5 制御プログラムの構成

複合計算栈システムに要求される機能を満足させるためには、構成要素をソフトウェア的に結合し、定められた計算栈間情報処理方式にしたがって、システムを制御するプログラムが必要となる。このような管理プログラムは、計算栈間の情報転送を行ない、要素栈の資源を管理するために、それぞれの要素栈に於て動作することが要求される。しかしこれらは必ずしも同じ機能を持つ必要はなく、システムのジョブ処理に最も適したものであることが望まれる。

我々が採用した方式は、このような管理プログラムのサブ・セットをあらかじめ用意し、この中から必要に応じた管理プログラムをとり出して組合せ、システムを生成する方式である。管理プログラムの中でシステムを生成し、システムの資源を管理し、各管理プログラムの初期設定を行ない、且つインターフェイスを取る機能を持たせたものをシステム制御プログラム(SCP)と呼んでいこう。

SCPは、幾つかの計算栈が、あるジョブを実行中であっても影響を与えずに、必要に応じた管理プログラムと、それに制御されるタスク・プログラムとを、他の計算栈に生成することができる。他の主要な管理プログラムには、実時間の飛行シミュレーションに際して、最適なタスク制御を行なうためのRTSMなどがある。複合計算栈に於けるソフトウェア的結合とは、これらの管理プログラムに於て計算栈間情報処理の方式を規定し、インターフェイスをとることに他ならない。計算栈間情報処理としては、(1)計算栈ナンバーの識別、(2)動作開始、終了の中枢、(3)コマンドの転送、(4)データの転送、が主要となる。

即ち 計算機ナンバーを識別することによって 管理プログラムは任意の時点
でそのプログラムが動作する計算機を検知し、コマンド或いはデータ転送に際し
て転送路の設定を行なうことができる。(前項参照)

動作の中継機能とは、計算機に生成されたプログラムの動作が開始するとき、
或いは動作が終了し消滅するとき、計算機のラン状態を維持し(WAIT命令を
実行)動作の中継を行なうことを指す。FSK-IIでは、オンライン実時間処理
に重点を置いているので、応答処理効率をあげるため、コマンドは1語(16ビット)
を原則とし、管理プログラム動作指令、タスク制御指令、入出力処理指令、デー
タ転送指令、管理プログラム間の同期、異常通報など、32種のコマンドを規定し
ている。これらのコマンドは、計算機間通信のいわばホットライン的要素をまっ
ので、最高順位の優先度を与えている。

データ転送の機能は、入出力データや1つのジョブが幾つかの計算機に分散さ
れた時の関連データを転送することである。転送領域および転送数などのデータ属
性は、ランゲージ・プロセッサ、リングローダによってあらかじめ与えられ、管理
プログラムはこの属性にしたがってデータ転送を制御する。

なほ、コマンドおよびデータ転送は、ブロードキャスト機能を有効に利用して
行なっている。基本的な計算機間結合のソフトウェア上の配慮は以上の通りである。
複合計算機としての汎用性、処理能力などは、ここに述べた管理プログラムのサブ
セットをどのように作成し、どのように各計算機に配分するかにかかわれると思わ
れる。

6. おまけ

複合計算機として構成したFSK-IIは、最終的に以上述べたように、多重並列
の転送路を設け、かなりの制御機能を含ませたハードウェア・カプラによって、所期の
密結合を実現した。主な用途である実時間のシミュレーション用計算機としては、処
理の高速性、ジョブの分散による処理能力の大きさなど、充分満足すべきものが得
られた。現在なほ各種の試験を実施している段階であるが、接続バスの競合、シ
ステム資源としてのI/O機器の競合など、ハードウェア的には、期待通りの性能が得
られた。個々の要素計算機がミニコンピュータであることの制約は、例えば処理速度
の面から眺めると浮動小数点演算が極めて遅く、ジョブによっては大型機に遠く及
ばないというように、汎用高性能ではあり得ない。これらはミニコンピュータ複合体
としては当然の限界であり、問題何性能を評価すべきであろうと考えている。

複合体としてのリングローダや、SCP、或いは問題何言語プロセッサなど
FSK-IIを機能させるための基本ソフトウェアの開発を余儀なくされ、これらは当
初システムを計画した時点に考えたものよりはるかに充大なものとなった。

今後の運用によって、ソフトウェアの体系など多くの手直しを要すると思われるが
本システムの機能を最大にするため、複合体に於けるジョブの分散、並列処理に関
するアルゴリズムの確立、異常状態の生起に対するシステムの再構成などが当面
の課題である。

参考文献

- (1) 勝又、元屈： 共通母線による計算機群の結合方式。

電通学会計算機研究会資料 EC-71-22 (昭46.9)

- (2) T.S. Miller, et al : Multiprocessor Computer System Study
NASA Report CR-108654, 1970
- (3) W.A. Wulf, et al : Comp-A Multi-Miniprocessor,
Proc of FJCC, 1972
- (4) 堀川・林 : フライト・シミュレータを含む多目的用複合体システム
昭49 電気学会全国大会予種集 (昭49.3)
- (5) 藤野・原田他 : 複合計算機システムの結合方式・その1 ハードウェア
昭49 情報処理学会第15回大会 (昭49.12)
- (6) 中原・堀川他 : 複合計算機システムの結合方式・その2 ソフトウェア
昭49 情報処理学会第15回大会 (昭49.12)
- (7) 原田・富山 : 複合計算機用リニター・ジ・エディタ
昭49 情報処理学会第15回大会 (昭49.12)
- (8) 原田・堀川 : シミュレーション用複合計算機のデータ転送
昭49 情報処理学会第15回大会 (昭49.12)