

スーパーコンピュータ HITAC S-810 の
オペレーティングシステム

4V-8

大房 義隆・片田 久 (株)日立製作所)

上 政之 (日立ソフトウェアエンジニアリング(株))

1. はじめに

HITAC S-810システムは、近年の科学技術計算の大規模化に応じて開発した超高速のアレイプロセッサシステムである。

S-810システムのハードウェア構成を図1に示す。高速演算の中心となるベクトル処理ユニット(VPU)と、HITAC Mシリーズの汎用プロセッサの持つ命令を実行するスカラ処理ユニット(SPU), さらに、技術計算の中間データを一時的データセットとして高速にアクセスできる大容量の半導体拡張記憶装置(ES)が主な特徴である。

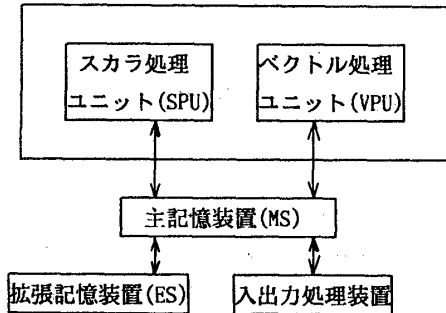


図1 S-810システムのハードウェア構成

S-810システムの制御OSは、スーパーコンピュータを制御するOSとして、次に示すような方針で開発した。

- (1) ベクトル処理ユニット, 拡張記憶装置などがユーザープログラムで容易に利用できること。
- (2) S-810システムの高速度演算能力を最大限に発揮できる資源スケジューリング機能とすること。
- (3) HITAC Mシリーズの汎用システムが持つ機能はそのまま利用可能とし, センタ運用が容易かつ柔軟にできること。

2. S-810システムの制御OS

図2に示すように, S-810システムを制御するOSは, HITAC Mシリーズの汎用OSであるVOS3/ES1をベースに, S-810システム専用OS VOS3/HAP/ESを付加することによりVOS3/ES1の持つ豊富なセンタ運用機能, 疎結合マルチプロセッサシステムなどの柔軟なシステム構成が構築ができる。

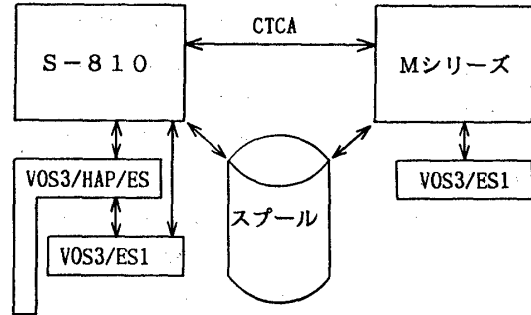


図2 S-810システムの構成例

更にVOS3/HAP/ESの持つS-810システム専用の資源スケジューリング機能によって, S-810システムの高速度演算能力を最大限に発揮させることができる。

3. VOS3/HAP/ESの制御方式

S-810システムの高速度演算能力を最大限に発揮させるための機能として, ベクトル処理ユニットのスケジューリング機能, 31ビットアドレッシング領域でのプログラム実行機能および, 拡張記憶装置による入出力処理の高速化機能がある。

以下に, 各々の制御方式について説明する。

(1) ベクトルジョブのスケジューリング機能

S-810システムで実行するジョブは, すべて, 各々の実行属性に応じてジョブクラスと呼ぶ概念でグループ分けできる。そして, システムの管理者は, このジョブクラス毎に, 当該ジョブクラスで実行するジョブの多重度や実行優先順位, 使用可能な記憶装置や拡張記憶装置の容量さらに, 実行時のディスパッチング方式(自動優先順位方式からラウンドロビン方式)などを制御することができる。これによりS-810システムの負荷状況に合わせたスケジューリングが可能となる。

(2) 31ビットアドレッシング領域でのプログラム実行機能

大規模な科学技術計算を行うためには, 使用可能なプログラム領域も大規模なものが必要になってくる。S-810システムの主記憶装置は, アーキテクチャ上は最大で2ギガバイトまで接続可能である。VOS3/HAP/ESではハードウェア資源の許す限り, ベクトルジョブで使用される大規模な配列のための領域や, ベクトルプログラムの命令語領域などの大半の領域を31ビットアドレッシング領域にローディングできるようにした。

(3) 拡張記憶装置による入出力処理の高速化

大規模な科学技術計算に必要なデータ量は、31ビットアドレッシングで表現できる範囲を超えるものも当たり前になりつつある。このため、これらのデータは、これまでは磁気ディスクのような外部記憶装置に格納して入出力処理を繰り返しながら演算を行う方式が一般的であった。S-810システムでは、このようなベクトルジョブに対して、高速な入出力処理を実現するために大容量の半導体による拡張記憶装置を装備している。

拡張記憶装置の利用は、ベクトルジョブの起動時にファイル単位にJCLで割り当てを要求することができ、FORT77/HAPコンパイラで作成されたベクトルプログラムからは、磁気ディスク装置に対して行っていた入出力ステートメントはそのままプログラム修正せずアクセスできるようにしている。拡張記憶装置の最大転送速度は1000メガバイト/秒と高速であり、これまでの入出力チャンネルによる最大転送速度3メガバイト/秒に比べて大幅な入出力処理の高速化が図れる。また、主記憶装置と拡張記憶装置のデータ転送処理は、専用の入出力アクセスプログラムで行うようにしている。図4で示すようにJCLで割り当てられたファイル対応に付与したファイル番号とファイル内相対アドレスを指定することにより、OSが実際の拡張記憶装置アドレスにアドレス変換してアクセスを行う。これにより拡張記憶装置内のフラグメンテーションの発生防止と、入出力処理におけるCPU処理時間の短縮を図っている。

拡張記憶装置

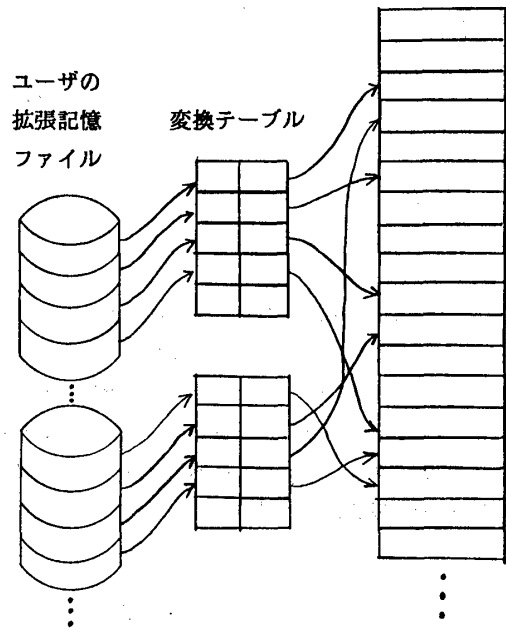


図4 拡張記憶ファイルの管理

4. おわりに

VOS3/HAP/ESはスーパーコンピュータ HITAC S-810システムの持つハードウェア能力を容易に効率よく運用できることを目標に開発した。今後は、さらに多数のTSSユーザから直接ベクトルプログラムを起動できるような方式を強化してゆく必要がある。

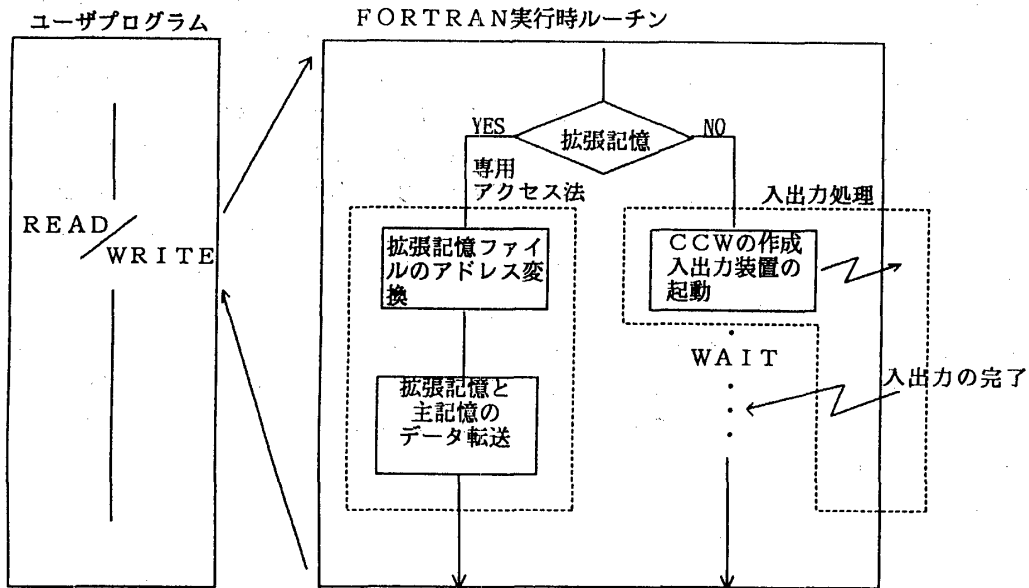


図3 拡張記憶装置のアクセス方式

⋯ : OS