

GRACEプロトタイプシステムにおける  
プロセッシングモジュールの制御方式

3C-9

鈴木 孝<sup>+</sup> 伏見 信也<sup>+</sup> 喜連川 優<sup>+</sup> 田中 英彦<sup>+</sup> 元岡 達<sup>+</sup>

(<sup>+</sup>三菱電機 <sup>+</sup>東京大学工学部 <sup>+</sup>東京大学生産技術研究所)

1. はじめに

関係データベースマシンGRACE プロトタイプシステムにおけるプロセッシングモジュール(以下PMと呼ぶ)は、汎用計算機を用いたホストマシン上のソフトウェアプロセスによって制御される。ホストマシン上では、OS中のI/Oハンドラが直接PMを制御しているが、今回ソフトウェアインターフェースとして入出力命令レベルでの仕様を定めた。さらにそれを基に、システムバスアダプタ(以下SBAと呼ぶ)を設計し、ホスト上の主記憶とPM間のデータ転送方式について検討を行った。

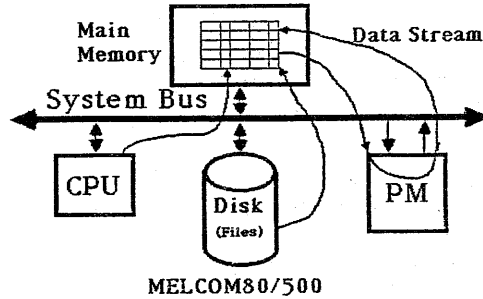


Fig. 1 Hardware Configuration of Prototype System

2. GRACE プロトタイプとプロセッシングモジュール

Fig.1 にプロトタイプのハードウェア構成図を示し、Fig.2 にプロセッシングモジュールの内部構成を示す。

GRACE はhashとsortを基本アルゴリズムとして、Diskからのデータストリームに沿って各種演算が実行される。プロトタイプでは、データストリームはホストマシン上の各種ソフトウェアプロセスによって制御される[1][2]。

PMはハードウェアソータを中心とする各種ハードウェアとその制御プログラムによって実装され、sort、join、aggregation等の演算を実行する[3]。PMはホスト上の主記憶からデータを受け取り、各演算実行後、結果を再び主記憶に返す。この制御を実行する部分がFig.2 中のBus Interface Unit (BIU) 中の入出力ポートSBAI、SBAOであり、ホストによるSIO(Start I/O)命令を解釈し入出力動作を行う。PMはこの他、ハードウェアソータとその駆動系からなるSort Unit(SU) と68000CPUを中心とするStream Processing Unit(SPU) の各ユニットから構成される。

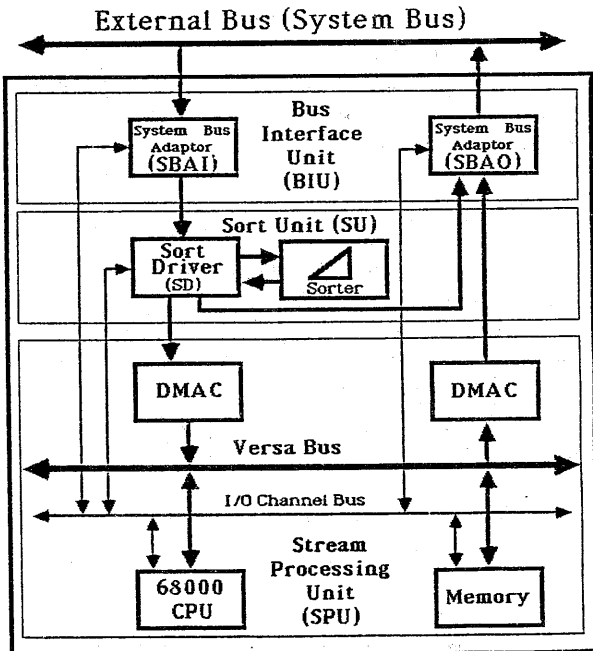


Fig.2 Processing Module

3. システムバスインターフェース

3-1. ソフトウェアインターフェイス

PMへのデータ入出力は、ホストマシン上のOS内のI/Oハンドラが発行するSIO 命令中のCCW(Channel Command Word)によって定義される。

ホストマシンインターフェース決定において、ホストから見て、できるだけ高レベルのインターフェースを与えることが目標となる。一方、既存の計算機上にインプリメントすることを考えると、そのシステムが持つ入出力アーキテクチャとの親和性も考慮する必要がある。今回、それらを念頭におき、さらにできるだけ単純化した形で汎用性を持たせるべく、入出力命令の仕様を定義した。

Fig.3 がその一般化したCCW 形式である。CCW はCCWA(CCW Address)で示される主記憶上にあり、SIO 発行時に各SBA はCCWAを受け取る。次にSBA はそのアドレスからCCWを読み込み、UA、ORDを解釈し、各種動作を実行し、TCBAで示されるアドレスにステータスを書き込み終了割り込みを返す。データ転送は、DTAで示されるアドレスからDTCで示される数だけ行なわれる。

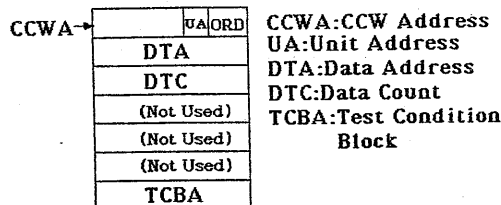


Fig.3 CCW Format

Device-SBAI			Device-SBA0		
UA	ORD	Operation	UA	ORD	Operation
0	Read	(SPU Read)	0	Read	SPU Read(Join)
0	Write	Initialize	0	Write	(SPU Write)
1	Read	Undefined-NOP	1	Read	SU Read(Sort)
1	Write	Data Out to SU	1	Write	Undefined-NOP

UA 0: SPU Select  
1: SU Select

Fig.4 I/O Commands for PM

Fig.4 にPMの入出力動作を示す。入出力ポートSBAI,SBA0 は独立しており、別デバイスとして同時に入出力動作が可能となっている。ユニットアドレス(UA)を、0がSPU、1をSUと定義している。

入力ポート(SBAI)に対しては、初期化命令と、SUへのデータ転送命令が用いられる。Initialize動作は、SBAIがDTAで示されるアドレスを間接的にアクセスし、その中にある各種情報(レコード情報、key情報、join情報等)を基に68000CPUがPMを初期化する。

出力ポート(SBA0)に対しては、SPUからのjoin結果等の読み込みと命令と、sort結果の読み込み命令が用いられる。

### 3-2. システムバスアダプタ

Fig.5 にSBAI,SBA0の内部構成を示す。この両者は、SU/SPU Interfaceを除いて、ほとんど同一構造を有している。

CCWの解読、CCWA,DTA,DTC等の各レジスタや各種ハードウェアリソースの制御、68000CPUとの通信等は、SBA中のMC68120(Intelligent Peripheral Controller)のプログラムによって実行することにより、汎用性、柔軟性を持たせている。

内部には、32KBの高速バッファを持ち、主記憶との間で高速転送(12MB/S)ができ、初期化データやステータス等、ホストとPMの情報の受け渡しに十分な余裕を持っている。

主記憶のアクセスは、DMA制御で実行され、SU/SPUとは、3MB/Sでデータ転送を行う。

PMへのデータ転送単位は、CCW中のDTA,DTCによって示されるメモリブロックで行なわれ、複数回のSIOに分けて転送することができ、主記憶上の特別な構造を意識する必要はない。また物理的に離れたブロックをData Chaining機構を用いて、一度の入出力動作で行ったり、Channel Address変換機構を用い論理アドレスアクセスすることも可能である。

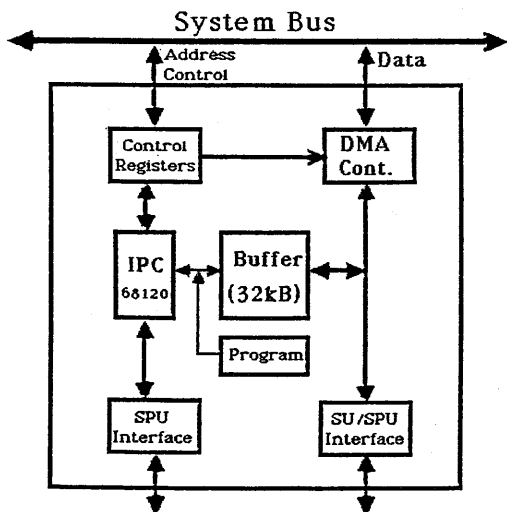


Fig.5 System Bus Adaptor

### 4. ホスト、PM間データ転送方式

ホストマシンのSIO命令によって、主記憶上のレコードデータがどのように転送され、PM内で処理され、結果が主記憶に返されるかをjoin演算を例にして以下に説明する。(Hはホスト、PはPMの処理を示す。)

- H (1) SIO Initialize
- |
- P (2) 68000CPUによるPM初期化
- |
- H (3) SIO Data Write (主記憶->SBAI->SU)
- |
- P (4) Sort Driver (4)によるデータ変換
- |
- P (5) ソート結果を68000メモリーに転送
- |
- P (6) 68000によるjoin処理
- |
- H (7) SIO Data Read (join結果:SPU->SBA0->主記憶)
- |
- (8) 終了

(3),(7)のSIO命令は、レコードデータ書き込み、結果の読み込みを独立して非同期に発行可能であり、SBA側で出力の同期化機構を有する。

このようにホストマシンからは、必要最小限のSIO命令で、join等の負荷の重い演算を高スループットで受け取ることが可能となっている。

### 5. おわりに

ホストマシンから見た入出力装置としてのプロセッシングモジュールのインターフェース、システムバスアダプタの構成、データ転送手順について述べた。

入出力システム的设计がシステムスループットに与える影響は大きいと考えられる。今後さらに効率の良い入出力方式について検討を加えていく予定である。

### <参考文献>

- [1] 中山、伏見、喜連川、田中、元岡：  
「GRACEプロトタイプシステムにおける記憶管理プロセスの構成」  
情報処理学会第32回全国大会 3C-8、1986
- [2] 中山、伏見、喜連川、田中、元岡：  
「GRACEプロトタイプシステムにおけるソフトウェア構成」  
情報処理学会第31回全国大会 1B-7、1985
- [3] 鈴木、伏見、喜連川、田中、元岡：  
「GRACEプロトタイプシステムにおけるプロセッシングモジュールの設計」  
情報処理学会第31回全国大会 1B-8、1985
- [4] 鈴木、伏見、喜連川、田中、元岡：  
「ハードウェアマージソータの駆動系の設計」  
情報処理学会第30回全国大会 1D-9、1985