

3T-10

並列計算機 P A X における
 ホスト・P U 間交信方式の改良

藤原秀久 朴泰佑 川合敏雄
 (慶應義塾大学)

【1】概要

P A X は科学技術計算を高速に実行することを目的としたマルチプロセッサであり、二次元格子状(トラス状)に結合されたプロセッシング・ユニット(P U)群と、これを制御するホスト計算機から構成される。P U 群が計算を並列に行うのに対しホストとの間のデータ転送は直列に行われるため、ホストとP Uをつなぐアダプタが律速となっていた。この点を改善するためホストのユーザ空間からP Uのメモリを直接アクセスできるようにした。これまでのシステムはホストにPDP-11、バスにQbusを使用していたが、本研究ではホストにNEWS-830、バスにVMEbusを使用した。システム本体はMiPAX-32J、ホストのOSはUNIXである。以下、VMEbusを使用したアダプタをVMADP、Qbusを使用したアダプタをQADPと呼ぶ。

【2】QADPからVMADPへの移行

ホスト・P U間のプログラム・データ転送において、アダプタ・P U間の転送はDMA方式を採用しており、VMADPでも同様である。しかしQADPはホストのI/OとDMAコントローラを介してアクセスされ、ホスト・アダプタ間の直接参照ができないため、アダプタ・P U間のDMA転送の効果が生かされなかった。今回の改善によりホスト・アダプタ間でも直接参照が可能となり、ホスト・P U間全バスを通してのコントローラを用いないDMAが実現し転送速度が向上した。

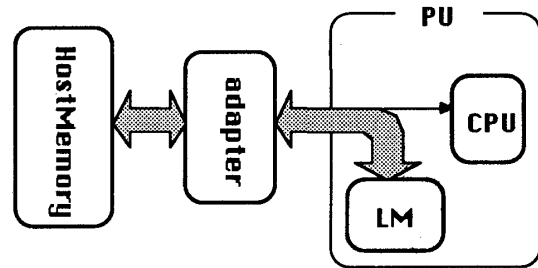


図1 データの流れ

【3】VMADPを用いたメモリ・マッピング

P Uに実装されているローカルメモリ(LM)、コミュニケーションメモリ(CM)は以下に述べるように、P Uメモリ→VME空間(UNIXカーネル空間)→UNIXユーザ空間という流れでマッピングされる。

a) P Uメモリ→VME空間

各P UのメモリはLM128KB、CM8KBの計136KBであるが、1P Uごとに256KBのメモリ空間を与えることにする。現在P Uの数は32であるから、256KB×32=8MBがVME空間にマッピングされる。また、ブロードキャストのために設けた仮想の33番目のP U256KB、レジスタ14BがVMEユーザ空間にマッピングされる。この33番目の仮想空間への書き込みによって全P Uへのブロードキャストが実現できる。

b) VME空間→UNIXユーザ空間

VME空間はUNIXカーネル空間のうちの1GBにマッピングされ、更にユーザプログラム中のmmap()コールによってユーザ空間へマッピングされる。

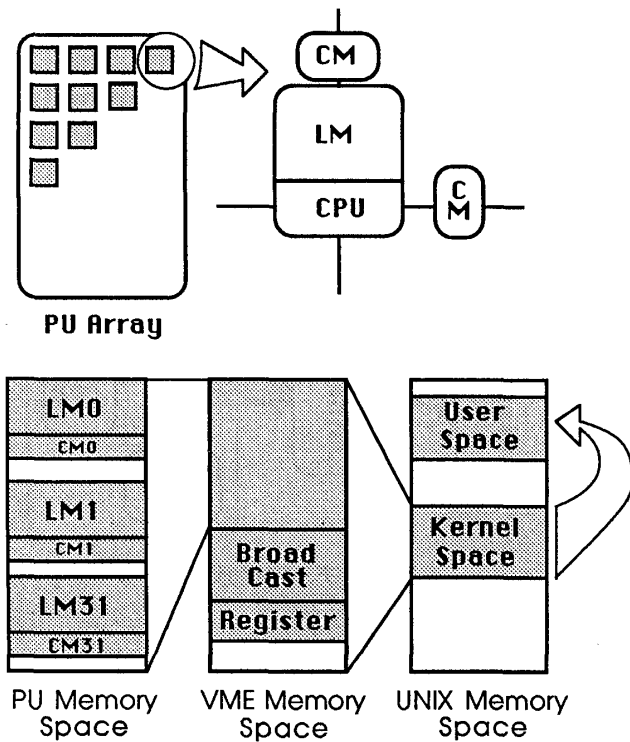


図2 メモリマッピングの流れ

これらの処理は全てVME空間のPUメモリやレジスタにユーザプロセスから直接アクセスすることで容易に実現される。

また、ユーザのプログラミングをより容易にするためPUを制御するライブラリが用意される。データ・プログラムの転送はメモリコピーで済み、コマンドの実行もユーザ空間でレジスタに値を書き込むことで実現される。また処理終了の伝達はホストに対する割り込みによって知らされ、カーネルに依存するのはこの部分だけである。PUのメモリをユーザ空間にマッピングすることによりライブラリの作成・修正が簡単になり、更にその理解も容易になった。

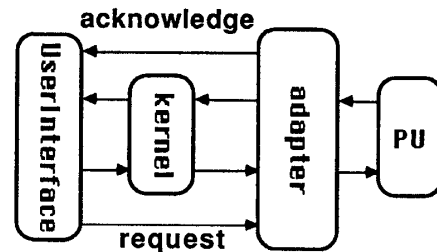


図3 制御の流れ

【4】VMADPを用いた制御

a) 制御の流れ

QADP方式ではユーザからの要求（レジスタ値の読み書き、PUプログラムの起動・停止など）はホストのI/Oからアダプタに送られるため、まずカーネルにシステムコールで要求が出され、処理終了もカーネルから伝えられる。一方VMADPではレジスタがユーザ空間にマッピングされているためカーネルを介さずに直接アダプタにアクセスできる。その結果、制御プログラムが簡単になり、またユーザから見たシステムが単純化され理解しやすくなった。

b) ホストからのPUの制御

VMADPを通してホストがPU群に対して行う操作には以下のようなものがある。

- ① ホスト・PU間のプログラム・データの転送、PU同士の間でのデータの転送
- ② PUにロードされたプログラムの起動・停止
- ③ PU間同期などの処理終了の伝達

【5】おわりに

PUメモリをユーザ空間にマッピングした結果、転送速度の向上・システムの単純化を実現した。今後このシステムをデバッグに応用し、システム操作・管理を容易にすることが課題である。

謝辞

本研究においてご助言を頂いた株式会社リクルートの斉藤隆之氏およびVMADP作成にあたりご協力頂いた、三井造船株式会社に対し深く感謝いたします。

【参考文献】

VME System Architecture Manual, 日本マイクロラ 半導体事業部, 1984年2月。