

3T-9

QCDPAXの制御機構と性能評価

内藤 潤 白川 友紀 吉江 友照 星野 力
(筑波大学)

【1】はじめに

QCDPAXは、2次元隣接結合網によるMIMD型高並列計算機である。本発表では、QCDPAXにおける制御機構と、並列処理制御及び表示システムの評価について述べる。

【2】制御機構

①ジョブ処理手順

1つのジョブは、ホストプログラムに記述されている一連の処理手順(PUプログラムのダウンロード、PUへの初期データのロード、PUプログラムの起動、並列処理終了待ち、結果の出力(アップロード)等)に従って実行される。ジョブが実行されるとHPI(Host-PU Interface)¹⁾では、モニタプログラムが走り、PU側でもユーザプログラム実行に先立ちモニタプログラムが走る。

②PUからHPI/ホストへの制御要求

並列演算実行中のPUからの処理要求はHPIによって受け付けられ処理される。PUからHPIへの処理要求には、同期処理、放送、演算終了報告等、がある。同期処理はPUの同期レジスタ、HPIの同期一致判定回路によって行われる。各PUは同期レジスタに同期コードを書きこむことによって一時停止し、HPIでは、全PUの同期コードの一致を検出すると全PUの同期レジスタをリセットし一時停止を解除する。放送や演算終了報告も一種の同期処理であり、特定の同期コードを書き込むことによって行われる。特定のコードで同期がとれたときHPIでは、要求された処理を実行する。PUからHPIには、2本の割り込み信号線(1つはHPIのみ、1つはHPI及び全てのPUへの割り込み用)がある。単独PUによる制御要求は、この信号線を使って行われる。例えばHPIへの割り込みは、同期を使わない放送やPUでのエラー発生報告に使用され、全PUへの割り込みは全PUの処理を強制的に中断させることができるのでデバッグ時に効果を発揮する。

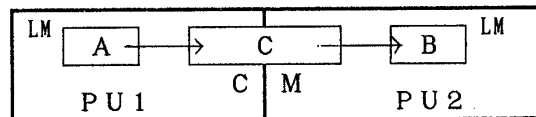
③エラー処理

PU、HPI内でのエラーは、それぞれPU、HPI

内のモニタで受けつける。PU、HPIで処理できるエラーは、PU内(CPUエラー、バリエータ、FPUCエラー)、HPI内(CPUエラー、PUのバリエータ、同期タイムアウトエラー)である。エラー状況報告は、次の順で行われる。(1)PUモニタ(内部処理を行った後HPIへ報告)-->HPI(2)HPIモニタ(内部処理を行なった後ホストへ報告)-->ホスト(3)ホスト(エラー状況をディスプレイに表示)

④PU間通信

PU間通信は32bitバスで、隣接PUとの共有メモリを介して行われる。ユーザはC=A, B=C, という代入文を書くだけでよい。ソフトウェアのオーバーヘッドもなく非常に高速である



【3】並列処理制御の性能

並列処理では、データ転送、それに伴う同期等の逐次処理にないオーバーヘッドが生じる。意味ある結果の得られる量の計算を行うのに、1プロセッサで要する計算時間とそれに伴うデータ通信に要する時間の比がgranularityと言う言葉で定義されている²⁾。我々は、並列計算機のハードウェア、コンパイラ、及びシステムソフトの性能を評価するために Unit-granularity というのを以下の式で定義した³⁾。

Unit-granularity:

$$g = \frac{\text{1回の浮動小数点演算時間}}{\text{1個の浮動小数データ転送時間}}$$

表1にQCDPAXの演算性能とUnit-granularityを示す。ベクタ長は10で計算してある。単一データの転送が、1.44μs、ベクタ長10のデータ転送が、0.92μs/dataである。QCDPAXのgranularityは、スカラ演算で0.8、ベクタ演算では、0.3~1と非常に効率の良いことがわかる。

QCDPAXとPAX64J⁴⁾の同期と放送に要する時間を以下に示す。

	QCDPAX	PAX64J
同期	2.5 μ s	18 μ s
放送(1024byte)	0.56 ms	9.0 ms

同期、放送を使用した例として、4PUのシステムで緩和法による2次元ラプラス方程式(64×64、直接写像)、ガウスジョルダン(GJ)法による連立1次方程式(100元)を解いた結果を表2に示す。放送は、GJ法で使用した。

【4】表示システム

各PUは、ホスト計算機/制御ユニットとの通信用にDualport Memoryを持つ。これは、並列演算実行中の外部とのリアルタイム・インターフェイスとして用いられる。この通信メモリを画像データ受渡しのバッファとして使用し、PUでの画像データ生成とHPIからグラフィックディスプレイへのデータ転送を並行して行うプログラムを作成し評価した。マンデルブrou集合の計算を1000×1000点についておこなった。各PUは、通信メモリ上に1000ワードのバッファを2つ用意し、1000点毎に計算を行う。結果を表3に示す。PUの負荷が均等で軽い(4PU②、area -2-2i~2+2i、loop limit 100) 場合には、表示ネックになるが、他の場合のようにPUの負荷がある程度重い場合は、表示

のオーバーヘッドなく随時計算結果を表示できる。

【5】むすび

PAXのアーキテクチャでは、隣接PU間での通信が非常に高速であり、QCDPAXでは、同期、放送の機能も強化されている。またグラフィックディスプレイへのデータ転送も高速であり計算実行経過の可視化も期待できるであろう。

本研究は、科学研究費(62060001)による。

参考文献

- 1)内藤、白川、星野：QCDPAXの制御ユニット 情報処理学会第36回全国大会7c-3
- 2)Harold S.Stone,High-Performance Computer Architecture, Addison-Wesly,1987
- 3)T.Hoshino,Parallel Computer and Parallel Computing in Scientific Simulations,ICNMF, July 1988
- 4)T.Hoshino,T.Shirakawa and K.Tsuboi, Mesh-connected parallel computer PAX for scientific applications,Parallel Computing 5(1987)

表1. QCDPAXの演算性能とUnit-granularity

式	ピーク性能 (MFLOPS)	半性能長 $n(1/2)$	実行時間 t_c (μ s/datum)	転送時間 t_b (μ s/datum)	Unit-granularity g
A = B + C	0.86	---	1.16	1.44	0.81
A(I)=B(I)+C(I)	5.56	40	0.90	0.92	0.98
S=S+A(I)*B(I)	16.7	38	0.29	0.92	0.32
S=S+A(I)*A(I)	33.3	76	0.26	0.92	0.28

(ベクタ長：10； ピーク性能、tc：予想； tb：実測)

表2. ラプラス方程式(緩和法)と連立一次方程式(GJ法)

4PU	スカラ演算			ベクタ演算		
	実行時間	通信時間	効率	実行時間	通信時間	効率
ラプラス方程式(緩和法)	2.42s	0.0434s	0.94	0.220s	0.0434s	0.80
連立一次方程式(GJ法)	1.44s	0.0348s	0.98	0.214s	0.0348s	0.84

表3. マンデルブrou集合の計算と表示

area	loop limit	SUN3	1PU	4PU①	4PU②	1PU (表示)	4PU① (表示)	4PU② (表示)	表示のみ
-2-2i ~ 2+2i	100	900	22.24	12.41	5.56	26.27	13.82	11.34	7.25
	1000	×	165.5	105.6	41.42	169.5	106.4	46.55	
-1-1i ~ 1+1i	100	×	68.59	31.28	17.17	71.71	32.08	20.10	
	1000	×	559.6	291.9	149.3	599.9	292.7	151.8	

①各PUは、areaを実軸方向に4分割した領域を担当

(時間単位は、Sec)

②各PUは、areaを実軸方向に4点おきに担当