

並列推論マシン PIM/c

— 負荷分散支援機構 —

2L-4

井門徳安^{*1}、前田浩光^{*2}、垂井俊明^{*1}、中川貴之^{*1}、杉江 衛^{*1}
 (*1 日立製作所 中央研究所、*2 日立製作所 神奈川工場)

1. はじめに

並列推論マシン PIM (Parallel Inference Machine)[1]の動的負荷分散方式として、スマート・ランダム方式[2]が提案されている。

PIM/c[3]では、スマート・ランダム方式を支援する機構をクラスタ・コントローラ[3]内のネットワークインタフェース・ユニットとネットワークであるルータ・ユニットに設けている。本稿では、PIM/cの負荷分散支援機構について述べる。

2. スマート・ランダム方式実現状況の課題

スマート・ランダム方式は、負荷の分配先を無作為に決定し、分配先の負荷が分配元より小さい時のみ分配を実行する。こうした制御の実現には、負荷分配時に他のクラスタ[3]の負荷値を知る必要がある。

クラスタ間を通常のネットワークを用いて接続した場合、メッセージ通信により、負荷値の問い合わせ/返答を行わなければならない。この場合、ネットワークの閉塞等により、負荷値が得られるまでの時間が大きくなる恐れがある。負荷値問い合わせ時間の増大は、負荷分散のオーバーヘッドを大きくしかねず、また、問い合わせの間に負荷量に変化してしまっており、正しい値が転送されないという問題も発生する。PIM/cでは、ネットワークインタフェース・ユニット (NU) とルータ・ユニット (RU) 内のハードウェアにより上記課題の解決を図る。

3. 負荷分散支援機構

図1にPIM/cのネットワーク関連部分の構成を示す。この部分の機能は、以下の二つである。

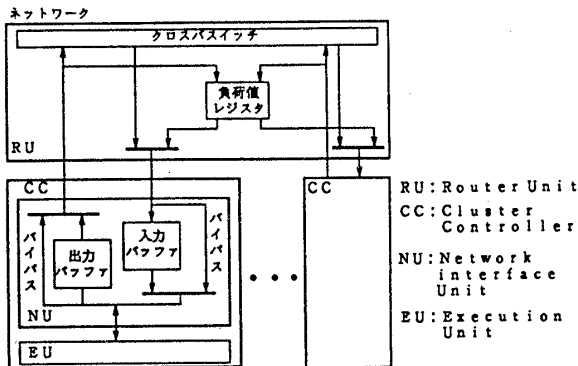


図1. ネットワーク関連部分の構成

(1) メッセージの転送

(2) 負荷値の問い合わせ/返答

(1)の機能は、NU内の64ワードの入出力バッファとRU内のクロスバスイッチによって実現する。

(2)が負荷分散支援機構で、これはNU内の入出力バッファのパイパスとRU内の負荷値レジスタによって実現している。実行ユニット (EU) [4]から出力された負荷値は、NU内のパイパスを通してRU内の負荷値レジスタに入力される。負荷値の問い合わせ/返答も、このパイパスを通して速やかに行われる。図2に負荷値の問い合わせ/返答動作のタイミングを示す。問い合わせは、EUからの出力後、3サイクルでNUからRUに出力される。この4サイクル後、RUからNUに負荷値が出力され、更に、4サイクル後EUに返答される。

4. おわりに

PIM/cでは、通常のメッセージ転送パスにパイパスを設けることで、高速の負荷値問い合わせ機能を実現した。

最後に、日頃ご指導いただいたICOT第4研究室内田俊一室長に深謝する。

なお、本研究は、ICOTからの委託研究の一環として実施された。

参考文献

- 1) 後藤他、"Toward a high performance parallel inference machine -The intermediate stage plan of PIM-", Future Parallel Computers, LNCS272, Springer-Verlag, 1986。
- 2) 杉江他、"Load-Dispatching Strategy on Parallel Inference Machines", Proc. of FGCS'88 Vol.3, PP.987-993
- 3) 後藤他、"並列推論マシン PIM/c 一概要一"、本大会発表予定。
- 4) 田中他、"並列推論マシン PIM/c -CPUについて-"、本大会発表予定。

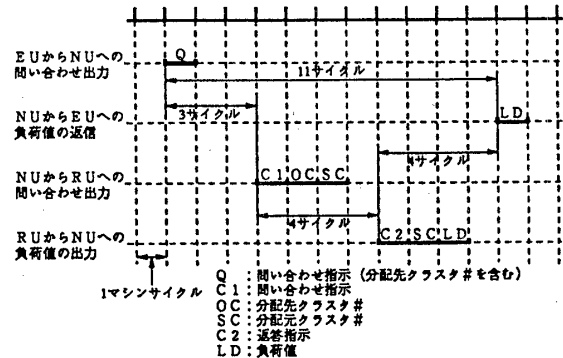


図2. 負荷値問い合わせ/返答のタイミング

Parallel Inference Machine PIM/c - Mechanism to support dynamic load balancing -

Noriyasu Ido, Hiromitsu Maeda, Toshiaki Tarui, Takayuki Nakagawa and Mamoru Sugie
 Hitachi, Ltd.