

可変ルーティングを行う 分散制御スイッチング・ユニット

坂井修一 計宇生 田中英彦 元岡達
(東京大学 工学部)

1. はじめに

LSI技術の進歩により高並列計算機への期待が高まってきたが、こうした計算機における相互結合網の役割は重要である。我々は既に固定ルーティング方式の分散制御スイッチング・ユニットを試作・評価した[1]が、今回はさらにこれを改良して、可変ルーティング方式のスイッチング・ユニットを設計したので報告する。

2. 可変ルーティングを行う相互結合網

格子型結合網やある種の多段結合網等では、可変ルーティングを行うことで転送効率を著しく向上させることが可能である。例えば、 $\log_4 N$ 段多段結合網(1ルート)に1段分のスイッチング・ユニット(SU)を付加する多ルート化で、場合によっては約54%スループットが向上することが示されている。[2]

可変ルーティングは1箇所ですべて集中的に制御することも可能であるが、その場合中央の制御装置が過負荷になりやすい。そこで、今回は各SUにこの制御を行わせる方式にした。

3. 設計システムの特徴

SUの基本構成を図1に示す。SUはm個の入力ポート(IP)とn個の出力ポート(OP)およびこれらを接続する内部バスより成る。IP内にはFIFOバッファがあり(蓄積交換)、またIPコントローラではルーティングを行っている。

各IPコントローラ内にはルーティング・テーブル(RT)があり、行先アドレス対応にOPが選択される。可変ルーティングは、RTに複数のコラム(エントリ)を設け、転送可能なOPをエントリの順番に探していく方式で実現する。(図2)今回は適応型ルーティングは考へなかった。

また、[1]ではFIFOの読出しと書込みが同時にはできなかったが、RAMを2面用いることで、1クロックで1回読出しおよび書込みが同時にできるように改良した。

転送するパケットは可変長とし、1語9bit、うちデータ8bit、EOP(End of Packet)信号1bitとした。転送はパイプライン転送方式[3]で行う。デッドロック対策としては、既にデッドロック・フリーなルーティングを精査した[3]が、動的な検出・回復についても検討中である。

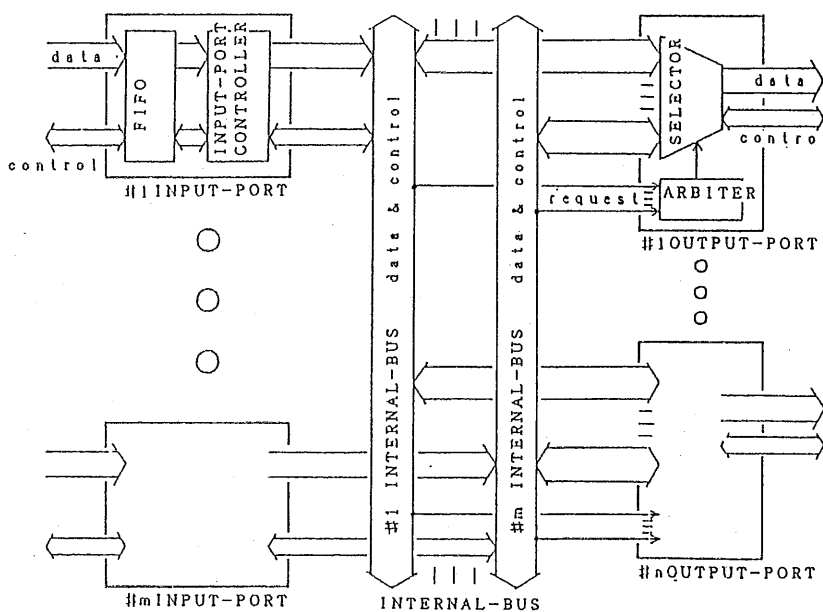


Fig.1 Overview of 1 SU

なお、設計システムは結合網を同一のクロックで動作させる同期方式である。

4. 評価・検討

4.1 転送性能

次の3つの時間をクロック数で求める。本システムでは、転送速度(クロック周波数)はFIFOの性能で決まる。

<データ転送遅延> パケット(長さP)の先頭がIPに到着してから、Q回のルート変更の後そのSUのOPから出てくるまでの最短時間: D_d

<パケット通過時間> 同じ時刻から、パケットが出終わるまでの最短時間: P_p

<パケット転送時間> s段のSUを経由し、総計A回のルート変更を行ってパケットが転送される最短時間: P_t

$$D_d = 2a + 3$$

$$P_p = p + 2a + 4$$

$$P_t = 3s + p + 2A + 1$$

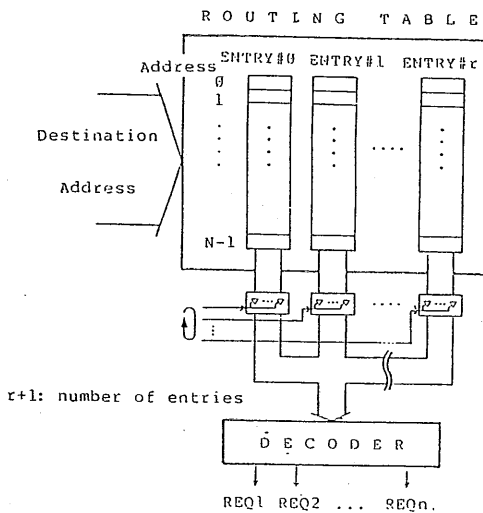


Fig.2 Routing Strategy

4.2 ハードウェア量

表1に、SSIとMSIを用いてSUを設計した場合のハードウェア量を示す。その他にも、内部バスの総配線数(5入力5出力で80本)、内部バスOP間の総接続線数(同300本)を考慮に入れなければならない。

今、5x5SUのLSI化を考えると、必要なピン数は114本である。その場合、内部配線数とRT用のRAM(表1)がやや大きいですが、十分に1チップ化が可能と思われる。

4.3 ルーティング・テーブルの構成

送り先でトラフィックに偏りが生じないように、またSU内での衝突ができる限り少くなるようにRTの各要素は決定されねばならない。現在、数種類のトポロジーの網に対して、RTの最適な構成法を検討中である。

5. おわりに

可変ルーティングを行うSUの設計に関して述べた。今後、本SUを試作するとともに、これを用いた大規模な結合網のシミュレーション評価を行う予定である。

<参考文献>

- [1] 南, 服部 他, 「プロセッサ間結合網に於けるスイッチング・ユニットの試作と評価」第27回情処全大 1983
- [2] 坂井 他, 「高並列計算機に於ける相互結合網の構成」第27回情処全大 1983
- [3] 服部 他, 「プロセッサ間ネットワークの性能評価」第26回情処全大 1983

	1 Port		5 Ports	
Number of Gates in IP	770	(670)	3850	(3350)
FIFO - RAM in IP	2.30kbits	(2.30kbits)	11.5kbits	(11.5kbits)
Routing Table in IP	768(r+1)bits	(768bits)	3.84(r+1)kbits	(3.84kbits)
Number of Gates in OP	110	(120)	550	(600)

(x) : x is the value of [1]

r+1 : number of entries

Table 1 Hardware Complexity of SU