

4U-3

トークンリングを介した高速データ
伝送制御に関する一検討

富田清次 木村行男 星子隆幸

NTT 電気通信研究所

1. はじめに.

プロセッサ間を、ネットワークを介して接続するプロセッサ間通信システムにおいて、長データを転送する方式には、ネットワークアダプタ(以下、アダプタ)内に十分なバッファを用意し、バッファ長までのデータを一括して送信する方式(一括転送方式)が、一般に採用されている。この方式では、プログラムが指定するデータ長は、バッファサイズで制限される。また、アダプタ内のデータバッファに一旦全データを収容した後、送信先アダプタにデータ転送されるため、プロセッサ-プロセッサ間のデータ転送が総て直列に行われることになる。

これに対して、アダプタ内に一定の大きさで複数(一般的には2個用意し、交互に切り替えて使用すれば良い)のデータバッファを設け、ひとつのバッファに対してプロセッサからデータ転送を行っている間に、他のデータバッファの内容をリングを介して送信先アダプタへ転送を行う方式(分割転送方式)が考えられる[2]。この方式では、プログラムが指定するデータ長に制限を加えない。また、送受信双方のアダプタ間のデータ転送、アダプタ-アダプタ間のデータ転送が総て並行して行われるため、電文通過時間を短縮することが可能になる。特に、アダプタ-アダプタ間のデータ転送速度が、プロセッサ-アダプタ間のデータ転送速度より速く、送信権獲得のリアルタイム性が高い高速トークンリングネットワーク(図1)においては有効な方式と考えられる。

しかしながら、分割転送方式の場合、フレームの送信、受信処理の回数は、一括転送方式の場合と比較して、分割転送回数分だけ増加する。このため、分割転送データ長を短くして分割転送回数を増加させると、アダプタの送受信制御に要する内部処理時間とアダプタ-アダプタ間の転送時間の和が、プロセッサ-アダプタ間の1分割データ当りの転送時間に比べて、相対的に長くなり、逆に電文転送時間が長くなる。

本検討では、アダプタの内部処理時間と最適な分割転送データ長との関係を求めることにより、性能的に分割転送方式の適用が有効となる領域を明らかにする。

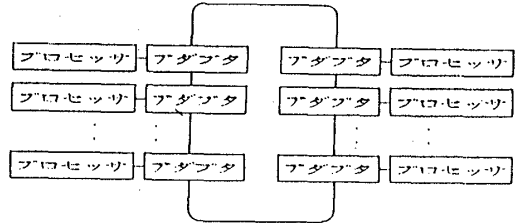
2. 電文転送モデルと電文転送時間算出条件.

1回のデータ転送シーケンスは、分割転送の場合、図2のようになる。データ転送の際にアダプタは、データ入出力処理(T1、T5、T6)、フレーム生成処理(T2)、フレーム送信処理(T3)、フレーム受信処理(T4)等の内部処理を行う。本稿では、アダプタ内での内部処理(T1~T6)に関して、個々の処理単位に同じ時間とした。以下において、内部処理時間とは、この単位時間を指す。

また、データ転送では、プロセッサにおいてアダプタの起動処理待ち(W4)、アダプタにおいてトークン待ちW1等が生ずる。

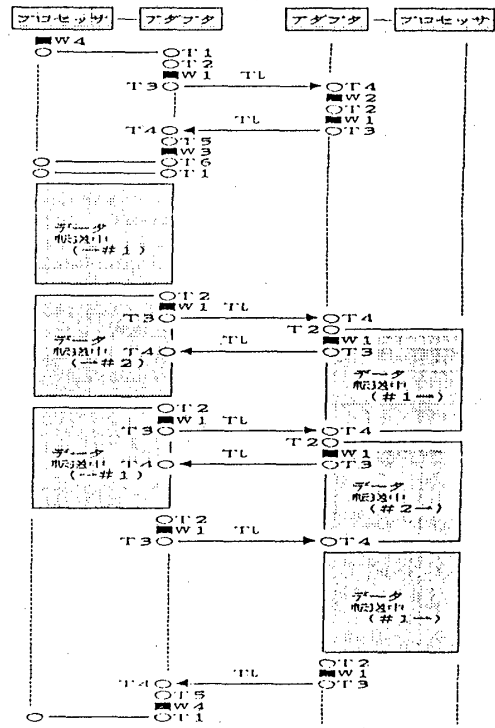
データ転送時間は、各待ち行列の直列接続型を用いて、各待ち時間(W1~W4)、及び、アダプタでのフレーム送受信処理時間(内部処理時間T1~T6)、プロセッサ-アダプタ間のデータ転送時間(TL)の総和で求められる。(算出式等詳細は、[2]を参照のこと。)

本稿の評価に用いた電文転送モデルでは、1回の電文転送は、入力電文(長データ)と出力電文(長データ)、及びフロー制御のための2回の制御情報(100B)の、計4回のデータ転送からなる。電文転送時間は、以上計4回のデータ転送時間の和である。また、全転送データ長とは、入力電文(=出力電文)のデータ長を指すものとする。



光リング
総距離 : L [km]
ノード数 : n
データ転送速度 : V1 [Mb/S]
I/Oインタフェース転送速度 : Vi [MB/S]

【図1 プロセッサ間通信システム】



W1: トークン待ち時間
W2: 受信バッファ獲得待ち時間
W3: 転送待ち時間
W4: 起動待ち時間
T1: プロセッサ-アダプタ間転送準備時間
T2: フレーム生成処理時間
T3: フレーム送信処理時間
T4: フレーム受信処理時間
T5: プロセッサ-アダプタ間転送待ち処理時間
T6: プロセッサ-アダプタ間転送処理時間
TL: フレーム伝搬時間

【図2 データ転送シーケンス(分割転送の場合)】

3. アダプタの処理能力（内部処理時間）と分割転送適用領域の関係

分割転送データ長を変えた場合のトラヒック—電文転送時間の変化の例を図3に示す。

分割転送単位が大きい場合（2kB）には、一括転送の場合（4kB）より電文転送時間は短縮されるが、プロセッサ—アダプタ間のデータ転送動作とアダプタ—アダプタ間のデータ転送動作が並行して行われる部分が小さいため、1kBの場合と比較すると電文転送時間は長くなる。分割転送単位が短くなると、アダプタでのフレームの送受信動作に時間がかかり、トラヒックが高くなると待ち時間が長くなるので、1kBの場合より短くなるに従って、電文転送時間が長くなる。以上から、この場合の最適分割データ長は1kBと見積られる。

以上のようにして求めた、全転送データ長、アダプタ内部処理時間に対する最適分割データ長を表1に示す。内部処理時間が速い100μSの場合、電文転送時間を短縮するためには、分割データ長を250～1kB程度に選ばば良い。内部処理時間が大きい400μS以上では、一括転送の方が電文転送時間が短い。

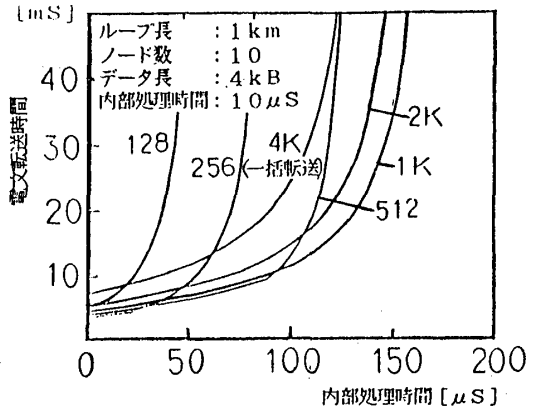
最後に、アダプタ処理能力と分割／一括転送の適用領域の関係を求めた。例として、全送信データ長が4kBの場合を示す。図4は、分割転送データ長をパラメータとした場合である。縦軸には、電文転送時間を採った。分割転送データ長が、128B、256Bと小さい場合には、内部処理時間を小さくしても、一括転送の場合と比較して、電文転送時間が多くかかる。分割転送データ長を1kBに採った場合、内部処理時間が150μS程度まで、分割転送は有利となる。図5は、分割転送単位1kBで、電文転送トラヒックをパラメータとした場合である。縦軸には、一括転送を基準とした電文転送時間の相対値を採った。内部処理時間100μS程度までは、電文転送トラヒックが変化しても分割転送は有効である。以上から、全送信データ長が4kBでは、分割転送データ長を1kBとした場合、内部処理時間を100μS程度以下に選ぶことで分割転送は有効となる。

4. おわりに

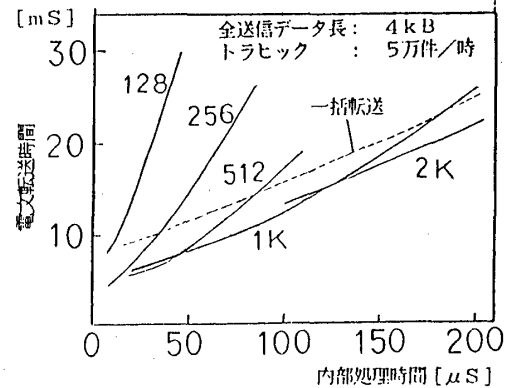
高速のネットワークを介したプロセッサ間通信における長データの転送方式について、実際の作りに着目して、分割転送と一括転送の適用領域を求めた。これにより、高速リングネットワークのアダプタ設計のための指針を与えることが出来ると考えている。

参考文献

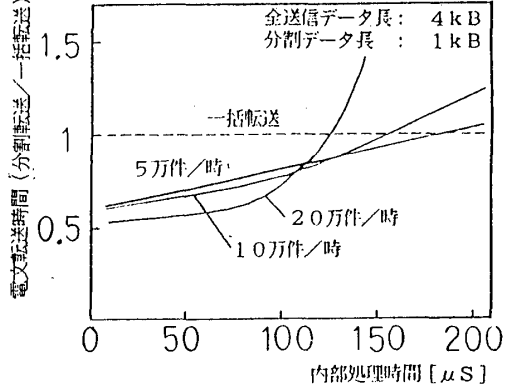
1. 小柳津、魚住、森、星子：高速光トークンリングを用いたプロセッサ間結合システムの性能評価、情処学会論文誌、pp. 128-132 (1986)
2. 星子、魚住、小柳津：100Mb/s光トークンリングを用いたプロセッサ間結合システム、情処学会論文誌、pp. 435-444 (1986)
3. 橋田：ポーリング制御における待ち合わせモデルの解析、昭和43年度信学会全国大会、p. 129 (1968)



【図3】 分割転送データ長によるメッセージ通過時間の変化



【図4】 分割転送の有効領域（分割長をパラメータとした場合）



【図5】 分割転送の有効範囲（トラヒックをパラメータとした場合）

【表1】 最適データ分割長

		全送信データ長					
		1kB	2kB	4KB	8kB	16kB	32kB
アダプタ内部処理時間	10μS	128, 256 (0~30) (30~)	256, 512 (0~24) (24~)	256, 512 (0~7) (7~23) 1k (23~)	256, 512 (0~2) (0~9) 1k 4k (9~19) (19~)	256, 512 (0~6) (6~3)	512, 1k (0~1) (1~2) 2k 4k (2~5) (5~)
	100μS	512, 1k (0~15) (15~)	1k	1k, 2k (0~13) (13~)	1k, 4k (0~1.2) (1.2~)	2k, 4k (0~8.5) (8.5~)	2k, 4k (0~3) (3~)
	200μS	1k	1k, 2k (0~4) (4~)	2k	2k, 4k	2k, 4k	4k
	400μS	1k	2k	2k, 4k	4k	4k	4k

() : トラヒック (万件/時)
記入のないものは、ほぼ全トラヒック領域で最適な場合。