

分散型共有メモリを用いた高速メッセージ通信システムSURE-SXの研究試作

—性能評価—

1 L-10

新家正総 加藤光幾 松平直樹 陣崎 明  
(株)富士通研究所

1 はじめに

富士通のフォールトトレラントコンピュータ SURE SYSTEM2000をFDDI(Fiber Distributed Data Interface)で結合し、装置間で2700メッセージ/秒以上、最大スループット11MByte/秒、最大通信遅延1.8msの性能を実現するメッセージ通信システムSXを研究試作した。本論文ではSXの性能評価結果について述べ、現アーキテクチャでの通信性能を検証し、更にFDDIよりも高速な伝送路を用いた場合のSXの性能予測を行う。

2 SXMのプロセッサ速度と目標性能

SXMは制御プロセッサとしてGMICRO100(12.5MHz、Dhrystone値3283)を用いている(表1)。目標性能はメッセージの送達確認を行った上でメッセージ処理数2500メッセージ/秒以上、最大遅延2.5ms以内、最大スループット10MByte/秒以上を実現することである。

3 性能測定

性能測定は2台のSURE SYSTEM2000をSXで接続し、PM間で相互に通信を行うプログラムを走行させて行った。PM-SXM間を接続するバスの速度は10MByte/秒×4チャンネル、SXM-SXM間を接続するネットワークは12.5MB/秒である。実測の結果SXMの通信制御ソフトウェアがメッセージ処理に要する時間は送信処理に313μs、受信処理に369μsであった。この結果から1秒間当りのメッセージ処理数、最大スループット、最大通信遅延を算出した。以下に算出方法を示す。

スループットはメッセージ長に依存し、ソフトウェア処理がボトルネックになる場合と、ネットワークへのDMAに要する時間がボトルネックになる場合があるため、以下のように計算し最大スループットを求めた。

(1)通信制御ソフトウェアのメッセージ処理時間がネットワークDMA時間よりも短い場合(スループットはDMAネックとなる)

スループット=メッセージ長/ネットワークDMA時間

(2)通信制御ソフトウェアのメッセージ処理時間が

SURE-SX:The High-speed Message Communication System using Distributed Shared Memory -Performance Evaluation-  
Tadafusa Niinomi, Koki Kato, Naoki Matsuhira, Akira Jinzaki  
Fujitsu Laboratories Ltd.

表1 プロセッサ性能

プロセッサ	GMICRO 100/12.5MHz
Dhrystone値(ver.2.1)	3283
MIPS値(VAX MIPS)	1.9

表2 性能測定結果

	目標性能	測定結果	
		送信性能	受信性能
メッセージ当りのソフト処理時間	400	313	369
メッセージ処理能力(メッセージ/秒)	2500	3190	2707
最大スループット(MByte/秒)(メッセージ長4KByte)	10	11.1	
最大通信遅延(ms)	2.5	1.8	

ネットワークDMA転送時間よりも長い場合(スループットは通信制御ソフト処理時間ネックとなる)

スループット=メッセージ長/通信制御ソフト処理時間

また最大通信遅延は最大長の4KByteのメッセージを転送した場合のPM-SXM間のバスのDMA時間、SXMの通信制御ソフトウェアのメッセージ処理時間、ネットワークへのDMA時間、宛先SXMの通信制御ソフトウェアのメッセージ処理時間、宛先PMへのDMA時間を加えることによって得た。

表2に性能測定結果を示す。この測定結果から以下のことが言える。

(1)メッセージ処理数、最大スループット、最大通信遅延の全てについて目標を上回る性能を実現

(2)メッセージ長4KByteでFDDIの物理的な伝送速度限界12.5MByte/秒に近い通信性能(11.1MByte/秒)を実現

これらの性能はFDDIで実測公表された通信性能としては世界最高速の性能である。

以上によりSXは、通信制御のハードウェア化による通信制御ソフトウェアの簡略化によって高速な通信処理を行えることが明らかになった。しかしながら最大スループットからわかるように通信制御ソフトウェアの処理時間はまだ性能のネックになっており、4KByteよりかなり小さなメッセージではFDDIの伝送速度の限界近くで転送を行うことができない。SUREのバスにおいては大量データ

を4KByte単位の packets で転送するモードがあり、4KByte転送時のスループットは重要な要素となっているが、一般的なメッセージ通信では小さなメッセージの高速な転送が要求される。現SXの性能はアーキテクチャの変更なしに容易に性能を向上させることが可能であり、以下では現アーキテクチャでのSXの通信性能改善方法と改善後の性能を検証する。

4 現アーキテクチャでの通信性能の検証

性能改善の方法には以下の3つの方法がある。

- (1) プロセッサクロックとプログラムメモリの高速化(5.7MIPSまでのプロセッサの高速化が可能)
  - (2) 他のより高速なプロセッサの使用
  - (3) C言語で書いたプログラムのチューニング
- (3)による性能向上の効果は現時点でははっきりわからないため、(1)、(2)によるプロセッサ性能の改善をした場合の通信性能を検証した。図1はFDDIの速度限界を使い切る為に必要なメッセージ長である。図1より以下のことが言える。

- (a) プロセッサでクロックやプログラムメモリの高速化を行った場合の最高性能5.7MIPSにおいて、1.5KByteで伝送路速度限界までの性能が実現可能
- (b) 汎用のプロセッサで利用可能な20MIPS程度を用いれば、500Byte程度の小さなメッセージで伝送路速度限界までのスループットを出すことが可能

以上により、SXはアーキテクチャの変更無しに容易に通信性能を向上させることができ、一般にプロセッサ間メッセージ通信で頻度の高い小さなサイズのメッセージについても高速な転送が可能ながことが明らかになった。

5 伝送路速度高速化時の性能予測

現在伝送路の速度は急速に延びており、物理レイヤではギガビットクラスの伝送が可能である。

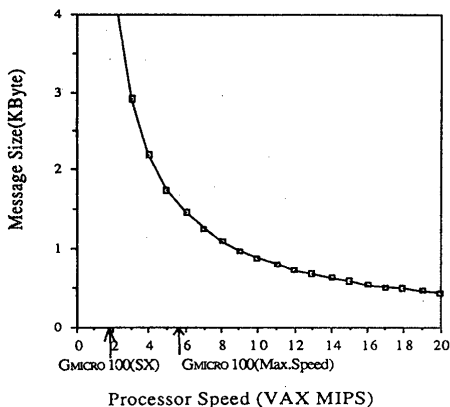


図1 FDDI伝送路速度限界のスループットを得る為のメッセージ長

以下ではSXをギガビットネットワークに適用した場合の性能を予想する。図2に通信制御ソフトウェアの処理パケット数から計算したスループットと、プロセッサ速度との関係を示す。この図から以下のことがわかる。

(1) 4KByteのメッセージ転送では20MIPS程度のプロセッサで1Gbit/秒 (125MByte/秒) の転送速度を得ることができる

(2) 256Byte程度の小さなメッセージについては100MIPS以上のプロセッサが要求される

(1)、(2)よりSXのアーキテクチャはギガビットネットワークにも適用可能であるが、小さなメッセージの転送を行うには非常に高速なプロセッサを要することがわかる。このような高速なプロセッサはコスト面で使用が難しく、ギガビットネットワークでの小さなメッセージの転送では更にソフトウェア処理を削減するようなアーキテクチャを考えることが重要である。

6 おわりに

本論文ではSXの性能評価、現アーキテクチャでの通信性能の検証、更に高速な伝送路を用いた場合の性能予測を行った。今後は更にアーキテクチャの改良を進め、より高速なネットワークを用いた試作を行う予定である。

[参考文献]

- [1] 陣崎他：オブジェクト共有型分散オペレーティングシステムの構想、情処技報DB73-9(1989-9)
- [2] 加藤他：分散型共有メモリを用いた高速メッセージ通信システムSURE-SXの研究試作—システムアーキテクチャー、本大会予稿
- [3] 松平他：分散型共有メモリを用いた高速メッセージ通信システムSURE-SXの研究試作—通信制御方式—、本大会予稿

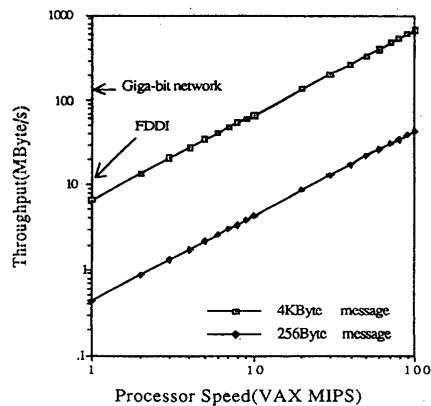


図2 プロセッサ速度とスループットとの関係