

マルチプロセッサ方式におけるプロセッサ間通信方式と通信特性について†

池 原 悟††

平等マルチプロセッサ方式は、単一 CPU 方式に比べて高性能化、低価格化が期待される反面、各種分割損や通信オーバーヘッドなどの不利な面も存在する。

ここでは平等マルチプロセッサシステムの各プロセッサ間で行われる通信方式の違いによる通信オーバーヘッドの差異を評価する。そのため、まず通信方式を分類して4つの代表的な方式モデルを抽出し、その各々にて選ばれる通信量、通信オーバーヘッドによるスループットの低下、通信待ち時間などを求めて比較した。

従来、マルチプロセッサにおけるプロセッサ間の結合方式の性能評価では、結合媒体（メモリやチャネル）へのアクセス競合を評価するモデルが主として用いられてきた。しかし、通信方式によっては受信側のプロセッサの状態を無視することはできない。本論文では通信要求元、通信媒体、通信相手の3者を含めたモデルによって評価した。

その結果によれば、たとえば各プロセッサからの通信要求の小さい場合は通信媒体の共用、非共用の差は小さく、むしろ通信時に相手プロセッサが独自の仕事を続けられるかもしくは通信処理に専念しなければならないかと言った方式上の差異の方が大きく性能に影響することなどが分かった。

1. ま え が き

複数のプロセッサで構成された複合計算機システムをすべてのプロセッサが平等に処理を分担する平等マルチプロセッサ方式と、各プロセッサの分担する処理内容の異なる分散処理方式に分けると、前者は単一プロセッサでは得られない高性能が期待され、後者はそのほかにプロセッサの特殊化によるシステムの価格性能比の向上が期待される^{1),2)}。

一方、これらの方式ではシステムの性能低下を招く原因としてプロセッサ分割損³⁾、制御表などのソフトウェア競合^{3),4)}、プロセッサ間通信オーバーヘッドなどが考えられる。本論文は通信オーバーヘッドを評価する。

複合計算機システム内のプロセッサ間通信については、プロセッサ間が完全群結合されているか、共通バスやチャネルと交換スイッチなどを有するか、もしくは共通メモリを介しているか、さらにはプロセッサ間の結合ではなくてメモリ間の結合かと言った各種の結合形態が考えられる。これらの結合形態によっては受信側のプロセッサの状態に通信の可否が依存するため、通信要求の発生元と通信線などの通信媒体だけでなく受信側の状態をも考えに入れる必要が生じる。

また、受信側の状態を考慮し、交信状態、通信待ち状態を分類して状態方程式を求める際、通信要求の発生順序が問題となり、この点、FCFSのヘテロジニアスな有限線群と似た解析上の困難さが生じる。そこで本論文では比較的小数のプロセッサから構成されるシステムに限定して各プロセッサはホモジニアスであることを仮定し、通信相手は1つのプロセッサとする2者間通信に限定してシステムの通信オーバーヘッドを評価する。

まず、第2章ではプロセッサ間の通信方式を分類し、その中から解析上同等な扱いの可能な方式は除外し、4つの異なる通信方式のモデルを作成する。現実のシステムでのプロセッサ間通信方式としては種々の方法が考えられており^{5),6)}、それらの分類法として、通信媒体（バス、チャネル、メモリなど）とその制御方式に着目した Anderson などの方法⁷⁾があるが、ここでは解析モデルの同等性に着目した分類を行う。したがって、バスを用いるかチャネルを用いるかの差を考えるのではなく、通信媒体が共用されるか否かなどに着目する。

次に、第3章では抽出された4つの通信モデルの通信特性を解析する。通信が片側通信で、通信相手がビジーか否かに無関係にメッセージを送信できる場合は従来の有限線群や CPU サイクルモデルが適用できて、たとえば文献 8), 9) にみられるように比較的容易に解析結果を得ることができる。これに対して、双

† Performance Degradation of MP System by Communications among Processors by SATORU IKEHARA (Yokosuka Electrical Communication Laboratories, NTT.).

†† 日本電信電話公社横須賀電気通信研究所データ通信研究部

方向通信などで通信時に通信相手が通信のため同時に保留されるモデルでは解析が困難であり、3台のCPUをメモリ結合した場合¹⁰⁾のほか解析例がない。一方、電話網における通信モデルで、高トラヒック状態の解析として通信相手の電話使用状況を考慮に入れた例¹¹⁾があるが、そこでは通信相手のビジー状態に発せられた呼は呼損として扱われており、計算機システムへそのまま適用することはできない。そこで、第3章では状態方程式から(一部シミュレーションも含まれる)システム全体で運ばれる呼量、通信オーバーヘッドで低下する処理能力、通信待ち時間などを評価し、4つのモデルで代表される通信方式の得失を論じる。

2. 通信方式と通信モデル

2.1 通信方式の分類

複合計算機システムにおけるプロセッサ間の結合方式として、バス結合、チャンネル結合、メモリ結合などがあげられるが、ここではこれらの通信媒体の種類よりもむしろ通信媒体へのアクセス競合の有無に着目し、待ちモデルとしてみたときの構造上の違いを考慮して通信方式を分類する。

通信の形態としては2者間の通信を考え、同時に3者以上が交信する会議形式の通信や1つのプロセッサがほかの2つ以上のプロセッサに同一のメッセージを送る放送形式の通信は考えない。また、2者間の通信でも片方向にのみメッセージの送られる片側通信と双方向にメッセージのやり取りを行う双方向通信があるが、通信時に2つのプロセッサが同時保留される場合はそのいずれでも良いが、メモリ結合などで送信時に受信側のプロセッサが受信処理に専有されないときは片側通信を考える。

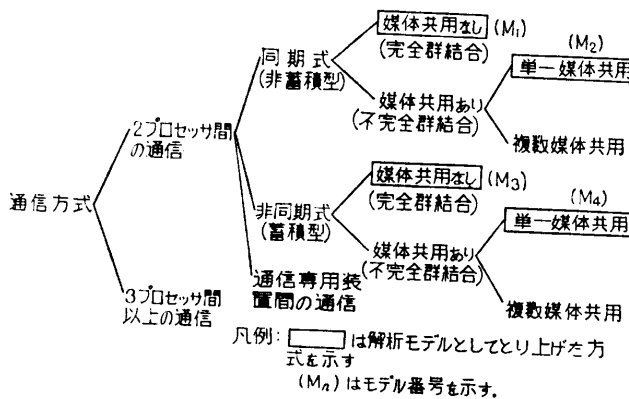


図1 プロセッサ間通信方式の分類

Fig. 1 Classification of the communication method.

(1) 同期式通信と非同同期式通信

以上の通信を通信時の通信相手側プロセッサの保留の有無に着目して、同期式通信と非同同期式通信に分ける。同期式通信方式は通信時に送信側、受信側のプロセッサが共に通信処理のため保留される方式であり、双方向通信の行われるとき、プロセッサ間が直接接続されるときはこれに相当する。これに対して、非同同期式通信方式は通信を行うプロセッサ間に蓄積型の通信媒体(メモリ)が存在し、通信は送信側のプロセッサから通信用のメモリに一方的にメッセージを書くことで行われ、受信側のプロセッサはその間もほかの独自の処理を実行できる方式である。通信用メモリとしてはプロセッサ間の共用メモリ、プロセッサごとのデイクレーティド・メモリ、通信バッファなどが考えられる。

このほか、通信が送信側、受信側のプロセッサの固有のメモリ間で行われたり、各プロセッサがCCPなどの通信専用の別装置をもち、それらの装置間で通信を行う方法も考えられるが、それらの場合は通信時に送信側、受信側のプロセッサ共に通信以外の処理が実行でき、両プロセッサは通信のためのオーバーヘッドを特に負担しないので検討の対象から除く。

(2) 通信媒体の共用と非共用

次に、通信方式を通信媒体(手段)へのアクセス競合の有無に着目して、媒体を共用する方式と共用しない方式に分ける。媒体を共用しない方式は通信の行われるすべてのプロセッサ間が独自の回線もしくはメモリを介して接続されている完全群結合方式で、通信回線へのアクセス競合はなく、プロセッサへの競合(プロセッサが個別のメモリを介して結合されているときはメモリへの競合となる)が主たるオーバーヘッドとなる。

これに対して通信媒体を共用する方式ではプロセッサ間の異なる通信で共通の通信媒体を介して通信が行われるため、プロセッサへの競合だけでなく共用する通信媒体への競合も問題となる。

媒体を共用する場合は媒体の共用の仕方によって種々のモデル構成が考えられるが、ここでは簡単のため単一の媒体がすべての通信で共用される場合を考える。

以上の分類の結果を図1に示す。図中本論文でとりあげる方式はM1~M4の4つの方式である。

なおM1~M4の方式名は、

M1……完全群結合同期式通信方式

M2……不完全群結合同期式通信方式

M3……完全群結合非同期式通信方式

M4……不完全群結合非同期式通信方式

とする。また、これらの通信方式のモデルを図2に示す。

2.2 通信モデルの構成条件

前節で取り上げた4つの通信モデルに対して以下の条件を仮定する。

(1) 全モデル共通の条件

① 通信回線競合, 受信装置競合のないときのプロセッサ i から j への通信要求の発生は平均呼空間隔 $1/\lambda_{ij}$ のランダム生起とし, 通信のための回線保留時間は平均 $h_{ij}(=1/\mu_{ij})$ の指数分析とする。

② すべてのプロセッサは通信に関して均質 (ホモジニアス) であり, $i \neq j$ なるすべての i, j に対して $\lambda_{ij} \equiv \lambda, h_{ij} \equiv h(=1/\mu)$ とする。

③ 各プロセッサに対する通信待ち要求は先着順に処理される。

④ 各通信回線における信号の遅延および回線切り換え時間は無視する。

(2) 通信モデル固有の条件

⑤ プロセッサ i から j への通信要求が発生するとプロセッサ i は j への通信媒体を確保する。通信媒体がすでに保留中のときは開放されるまでプロセッサ i は待ち状態となる (M2, M4)。

⑥ M1, M2 では通信媒体確保後, プロセッサ i は j を確保する。プロセッサ j がほかのプロセッサと通信中のときはその通信の終了を待つが, 通信以外の処理中のときは割込む。

⑦ M3, M4 では受信プロセッサ j の確保はせず, 通信媒体が確保されるとそこへ通信文を書き込む。したがって, 通信中もプロセッサ j は独自の処理の続行が可能であり, プロセッサ j は任意の時点で通信文を読み取ることができる。

3. 各通信モデルの評価

3.1 評価項目について

すべての通信モデル (M1~M4) の間で通信能力に差が生じるのはプロセッサ台数 N が4以上の場合である。 $N=1$ では通信はなく, $N=2, 3$ では M1 と M2 および M3 と M4 はそれぞれ縮退しているが $N=4$ 以上でこれらの縮退はとれる。そこで各モデルについて $N=2, 3, 4$ の各場合について状態方程式を解き, システム全体の通信量 C (アーラン), 通信オーバーヘッドによるスループットの低下率, 各プロセ

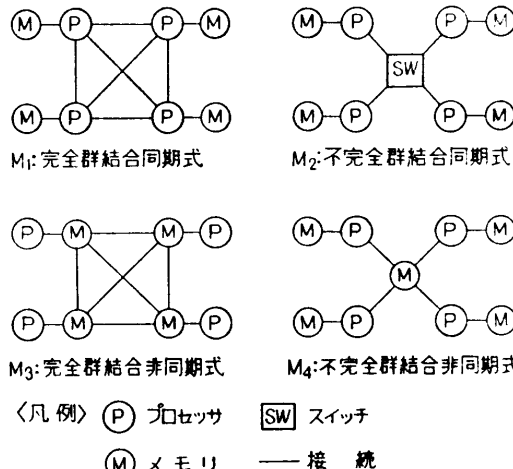


図2 通信方式を代表する4つのモデル (プロセッサ数=4の場合)

Fig. 2 Representative communication models.

ッサの受信負荷, 通信待ち時間を評価する。

状態方程式をたてるに当っては, 通信中の通信要求の数を n_1 , 通信待ち状態の通信要求の数を n_2 とし, システムの状態を (n_1, n_2) で表わし, 状態 (n_1, n_2) に含まれる異なる状態の数を $G(n_1, n_2)$ とする。また状態 (n_1, n_2) に含まれる各異なる状態の平衡状態における状態確率はすべて等しい*とし, これを $P(n_1, n_2)$ で表わす。したがって状態確率の規格化条件は,

$$\sum_{(n_1, n_2)} G(n_1, n_2) P(n_1, n_2) = 1 \quad (1)$$

となる。

次に各通信モデルの通信特性を表現する量を定義する。まず, システム全体の通信量 C を次式で定義する。

$$C = \sum_{(n_1, n_2)} n_1 G(n_1, n_2) P(n_1, n_2). \quad (2)$$

各通信方式において通信要求の競合が無く, 通信待ち時間が零の場合を考えると, この場合はプロセッサ1台当たり $(N-1)\lambda h / \{1+(N-1)\lambda h\}$ アーランの通信が運ばれるから, プロセッサ N 台のシステムでは $N(N-1)\lambda \cdot h / \{1+(N-1)\lambda h\}$ アーランの通信が処理される。次に各プロセッサが完全にスタンド・アローンで動いている場合を考えると, この場合は通信の必要はなく, 各プロセッサはその能力を送受信にふり向けることはないため, システム全体の処理能力は上述の場合の $1+(N-1)\lambda h$ 倍に増大し, 仮想的には

* 状態確率が等しい状態の中からほかの状態への遷移確率が等しくなる状態を選び出し, それを1つの状態の集合 (n_1, n_2) としてまとめる。

$N(N-1)\lambda h$ アーランの通信要求が処理されている場合に相当する。そこで、各方式のスループットとスタンド・アロン方式のスループットの比をスループット達成率と呼び R で表わすと、 R は次式で与えられる。

$$R = \frac{C}{N(N-1)\lambda h} \quad (3)$$

また、各プロセッサの受信負荷を ρ_c 、通信要求の平均待ち時間を W/h で表わすと、

$$\rho_c = C/N \quad (4)$$

$$W/h = L/C \quad (5)$$

となる。ただし、 L は平均通信待ち数で、

$$L = \sum_{(n_1, n_2)} n_2 G(n_1, n_2) P(n_1, n_2) \quad (6)$$

で与えられる。

3.2 完全群結合同期式通信モデル (M1)

(1) $N=2$ の場合

平衡状態での状態方程式と規格化条件、

$$p(1, 0) = k p(0, 0) \quad (7)$$

$$p(0, 0) + 2p(1, 0) = 1 \quad (8)$$

$$(k = \lambda h)$$

を解いて、

$$\left. \begin{aligned} p(0, 0) &= 1/(1+2k) \\ p(1, 0) &= k/(1+2k) \end{aligned} \right\} \quad (9)$$

よって表1の結果を得る。

(2) $N=3$ の場合

状態方程式と規格化条件は、

$$\left. \begin{aligned} p(1, 0) &= k p(0, 0) \\ k p(0, 0) + 2p(1, 1) &= (1+2k)p(1, 0) \\ k p(1, 0) &= p(1, 1) \end{aligned} \right\} \quad (10)$$

$$p(0, 0) + 6p(1, 0) + 12p(1, 1) = 1. \quad (11)$$

表1 M1の通信特性

Table 1 Communication characteristics (M1).

プロセッサ数	$N=2$	$N=3$	$N=4$
特性			
システム全体で運ばれる通信量 C	$2k$	$6k+12k^2$	$[12k + \frac{24k^2(6+39k+69k^2)}{2+9k+7k^2}] p(0,0)^*$
呼源呼量	$2k$	$6k$	$12k$
スループット達成率 R	$\frac{1}{1+2k}$	$\frac{1+2k}{1+6k+12k^2}$	$\frac{C}{12k}$
プロセッサ当りの受信負荷 ρ_c	$\frac{1}{1+2k}$	$\frac{2k+4k^2}{1+6k+12k^2}$	$\frac{C}{4}$
通信待ち時間 W/h	0	$\frac{12k^2}{6k+12k^2}$	$\frac{4k(2+16k+35k^2)}{2+21k+79k^2+138k^3}$

$$k = \lambda h$$

* $1/p(0,0) = 1 + 12k + 12k^2 \frac{10+61k+101k^2}{2+9k+7k^2}$

表2 M1における (1,1), (1,2) の状態の分離

Table 2 State separation for (1,1) and (1,2).

状態ベクトル	状態数	意味
(1,1A)	48	送受信中のプロセッサに対してほかのプロセッサから通信要求が発生
(1,1B)	48	送受信中のプロセッサからのほかのプロセッサへの通信要求が保留中
(1,2A)	48	送受信中のプロセッサのいずれかにほかのプロセッサからの2つの通信要求が発生
(1,2B)	48	送受信中のプロセッサの双方へほかのプロセッサからの通信要求が発生
(1,2C)	48	受信プロセッサから第3のプロセッサへの通信要求と第4のプロセッサから受信プロセッサへの通信要求が保留中
(1,2D)	96	受信プロセッサから第3のプロセッサへの通信要求と、第3のプロセッサから送受信いずれかのプロセッサへの通信要求が保留中
(1,2E)	48	受信プロセッサから第3のプロセッサへの通信要求と、第4のプロセッサから送信プロセッサへの呼が保留中

これを解いて、

$$\left. \begin{aligned} p(0, 0) &= 1/(1+6k+12k^2) \\ p(1, 0) &= k p(0, 0), \quad p(1, 1) = k^2 p(0, 0) \end{aligned} \right\} \quad (12)$$

を得るから、表1の結果を得る。

(3) $N=4$ の場合

この場合は状態は(0,0), (1,0), (1,1), (2,0), (2,1), (1,2) の6通りであるが、このうち(1,1), (1,2)は表2の通りさらに分離され、状態遷移図は図3の通りとなるから、状態方程式は以下の通り。

$$\left. \begin{aligned} p(1, 0) &= k p(0, 0) \\ k p(0, 0) + 4p(1, 1A) + 4p(2, 0) + 4p(1, 1B) &= (1+6k)p(1, 0) \\ k p(1, 0) + p(1, 2A) + p(2, 1) + p(1, 2C) &= (1+3k)p(1, 1A) \\ k p(1, 0) + 2p(1, 2B) + 2p(1, 2E) &= p(2, 0) \\ k p(1, 1A) &= p(1, 2A) = p(1, 2B) \\ k p(1, 1A) + 2k p(1, 1B) &= 2p(2, 1) \\ k p(1, 1B) &= p(1, 2C) = p(1, 2D) \\ p(2, 1) + 2p(1, 2D) &= (1+6k)p(1, 1B) \end{aligned} \right\} \quad (13)$$

これを(1)式と組み合わせて解くと、

$$\left. \begin{aligned} 1/p(0, 0) &= 1 + 12k + 12k^2(10 + 61k + 101k^2)/(2 + 9k + 7k^2) + 9k + 7k^2 \\ p(1, 0) &= k p(0, 0) \\ p(1, 1A) &= \{k^2(2+6k)/(2+9k+7k^2)\} p(0, 0) \\ p(1, 1B) &= \{k^3/(2+9k+7k^2)\} p(0, 0) \\ p(2, 0) &= [k^2(2+13k+21k^2)/\{2(2+9k+7k^2)\}] p(0, 0) \\ p(2, 1) &= \{k^3(1+4k)/(2+9k+7k^2)\} p(0, 0) \\ p(1, 2A) &= p(1, 2B) = \{k^3(2+6k)/(2+9k+7k^2)\} p(0, 0) \\ p(1, 2C) &= p(1, 2D) = p(1, 2E) \\ &= \{k^4/(2+9k+7k^2)\} p(0, 0) \end{aligned} \right\} \quad (14)$$

よって、各特性は表1の通りとなる。

3.3 不完全群結合同期式通信モデル (M2)

本モデルは $N=2, 3$ の場合は M1 と等価となるから $N=4$ の場合を解析する。

状態間の遷移関係を求めるため、この場合も状態 $(1, 1), (1, 2)$ を表 3 の通り分離する。状態間の遷移確率が $\lambda \Delta t, \mu \Delta t$ の整数倍となるようにするにはさらに状態 $(1, 2D), (1, 2E), (1, 3)$ を分離することが必要であるが、ここではこれ以上は分離しない。

図 4 にこれらの状態間の状態遷移図を示す。図からも分かる通り、本モデルでは単一の通信媒体が共用されるため同時に 2 つ以上の通信が実行されることはない。

図 4 から状態方程式を書き出すと、

$$\begin{aligned}
 p(1, 0) &= k p(0, 0) \\
 k p(0, 0) + 3p(1, 1A) + 4p(1, 1B) \\
 &\quad + 2p(1, 1C) = (1 + 6k)p(1, 0) \\
 2p(1, 2D) + 2p(1, 2E) + 2p(1, 2C) \\
 &= (1 + 6k)p(1, 1A) \\
 2p(1, 2D) + 2p(1, 2B) + 2p(1, 2E) \\
 &\quad + k p(1, 0) = (1 + 3k)p(1, 1B) \\
 k p(1, 0) + 2p(1, 2D) + 2p(1, 2E) + p(1, 2A) \\
 &= (1 + 3k)p(1, 1C) \\
 k p(1, 1A) + p(1, 3) &= (1 + 3k)p(1, 2D) \\
 k p(1, 1C) + p(1, 3) &= p(1, 2C) \\
 k p(1, 1B) + p(1, 3) &= p(1, 2B) \\
 2k p(1, 1B) + 2p(1, 3) &= p(1, 2A) \\
 p(1, 3) &= (1 + 3k)p(1, 2E) \\
 k p(1, 2D) + k p(1, 2E) &= p(1, 3)
 \end{aligned}
 \tag{15}$$

となる。正規化特性は、

$$\begin{aligned}
 p(0, 0) + 12p(0, 0) + 36p(1, 1A) + 48p(1, 1B) \\
 + 24p(1, 1C) + 216p(1, 2D) + 72p(1, 2C) \\
 + 16p(1, 2B) + 24p(1, 2A) + 216p(1, 2E) \\
 + 648p(1, 3) = 1
 \end{aligned}
 \tag{16}$$

よって (15), (16) 式を数式処理言語 AL¹²⁾ を用いて解くと、

$$\begin{aligned}
 1/p(0, 0) &= 1 + 12k + \frac{k^2}{1 + 5k} (872k^2 + 578k + 72) \\
 &\quad + \frac{k^4(36, 372k^3 + 26, 420k^2 + 6, 056k + 432)}{(k+1)(3k+1)(5k+1)(7k+1)} \\
 p(1, 0) &= k p(0, 0) \\
 p(1, 1A) &= \{2k^3/(5k+1)\} p(0, 0) \\
 p(1, 1B) &= p(1, 1C) = \{k^2(4k+1)/(5k)
 \end{aligned}$$

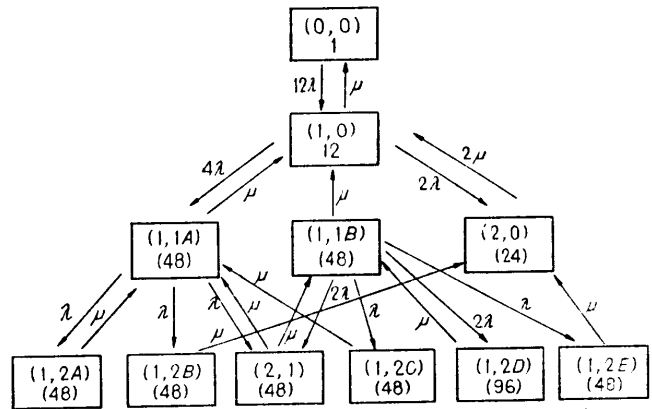


図 3 M1 における $N=4$ の場合の状態遷移図

Fig. 3 State transition diagram of M1 for $N=4$.

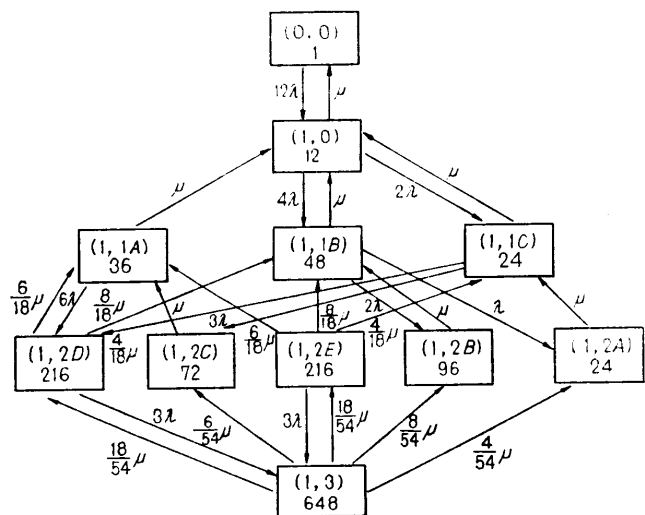


図 4 M2 における $N=4$ の状態遷移図

Fig. 4 State transition diagram of M2 for $N=4$.

表 3 M2 における $(1, 1), (1, 2)$ の状態の分離

Table 3 State Separation of $(1, 1)$ and $(1, 2)$.

状態ベクトル数	意味
(1, 1A) 36	受信中のプロセッサからのほかのプロセッサへの通信要求が保留中
(1, 1B) 48	送受信中のいずれかのプロセッサへほかのプロセッサからの通信要求が保留中
(1, 1C) 24	ほかのプロセッサ間での通信要求が保留中
(1, 2A) 24	送受信中のプロセッサ双方へのほかのプロセッサからの通信要求が保留中
(1, 2B) 96	送受信中のプロセッサのいずれか一方へのほかのプロセッサからの 2 つの通信要求が保留中。もしくは送受信中のプロセッサへのほかのプロセッサからの通信要求(先着)とほかのプロセッサ間での通信要求(後着)が保留中
(1, 2C) 72	第 3 プロセッサからの通信要求(後着)と第 4 プロセッサから第 3 プロセッサへの通信要求(先着)が保留中
(1, 2D) 216	受信中のプロセッサからの通信要求(後着)が保留中
(1, 2E) 216	受信中のプロセッサからの通信要求(後着)とそのほかの任意の 1 つの通信要求(先着)が保留中

ほかのプロセッサ: 送受信以外のプロセッサ 2 つを言う
 第 3 のプロセッサ: ほかのプロセッサのうちの 1 つ
 第 4 のプロセッサ: ほかのプロセッサのうちの 1 つ

$$\begin{aligned}
 &+1\} p(0, 0) \\
 p(1, 2A) &= 2p(1, 2B) = 2p(1, 2C) \\
 &= [k^3(4k+1)/(5k+1) + 4k^5/\{(k+1)(5k+1)\}] p(0, 0) \\
 p(1, 2D) &= [k^4(2k+1)/\{(k+1)(3k+1)(5k+1)\}] p(0, 0) \\
 p(1, 2E) &= [2k^5/\{(k+1)(3k+1)(5k+1)\}] p(0, 0) \\
 p(1, 3) &= [2k^5/\{(k+1)(5k+1)\}] p(0, 0)
 \end{aligned} \tag{17}$$

そこで平均通信待ち呼数 L は,

$$L = \left\{ 72k^2 + \frac{432k^2(6k+1)}{5k+1} + \frac{3,888k^5}{(k+1)(5k+1)} \right\} p(0, 0). \tag{18}$$

この場合は同時に2つ以上の通信要求は処理されないからシステム全体の通信量 C は通信媒体の使用率に等しく $C=1-p(0, 0)$ で与えられる。また各プロセッサの受信負荷 ρ_c , スループット達成率は $\rho_c=C/4$, $R=C/(12k)$ であり, 平均通信待ち時間は(5)式から求まる。

3.4 完全群結合非同期式通信モデル (M3)

$N=2, 3$ の場合は本モデルは次節で述べる $M4$ に一致するから, ここでは $N=4$ の場合を評価する。

本方式は受信側のプロセッサが通信中も独自の処理を実行することができ, 受信中のプロセッサからも通信要求が発生することがあるため, $M1, M2$ に比べて通信および通信待ち状態は複雑となる*。そこで, ここでは GPSS を用いたシミュレーションによって評価する。本モデルのシミュレーション・プログラムは各プロセッサの固有メモリに通信時に同時保留があるため, デッド・ロックに入らぬよう注意すれば容易に書ける。

3.5 不完全群結合非同期式通信モデル (M4)

通信文を書き込む蓄積型の通信媒体(メモリ)に容量制限は無いものとする, 2.2 の条件 ⑦ から, 本モデルは通信媒体をサーバとし, N 台のプロセッサを呼源とする入線有限待時式線群 $M(N)/M/1$ で表わされる。

したがって, システム全体の通信量 C , 平均通信待ち時間 W/h は,

$$C = 1 - 1 / \left\{ \sum_{i=0}^N (N)_i (\lambda h)^i \right\} \tag{19}^{**}$$

* 状態数は 1,681, これを 64 の状態に分類すると 64 元の連立方程式を解くプログラムは数式処理言語 AL で容易に書けるが, 実行時のメモリーオーバーフローが生じ, 数式解を得るに至らない。

** $(N)_i = N \cdot (N-1) \cdots (N-i+1)$, $(N)_0 = 1$

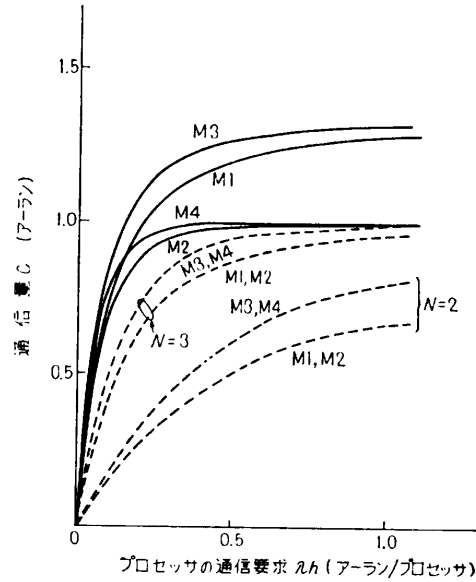


図5 プロセッサ間で運ばれる全通信量
Fig. 5 Total call (erl).

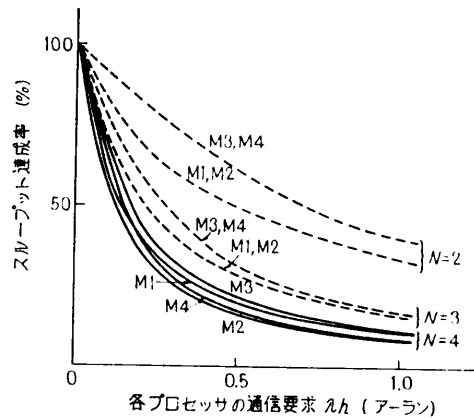


図6 スループット達成率
Fig. 6 Throughput degrade.

$$W/h = \frac{\sum_{i=1}^{N-1} i(N-1)_i (\lambda h)^i}{\left\{ \sum_{i=0}^{N-1} (N-1)_i (\lambda h)^i \right\}} \tag{20}$$

で表わされる。プロセッサ受信負荷 ρ_c , スループット達成率 R は(4), (3)式から得られる。

4. 各通信モデルの特性比較

各通信モデルの通信量 C , スループット達成率 R , 平均通信待ち時間 W/h をそれぞれ図5~7に示す。

4.1 全通信量 C の評価

(1) $N \leq 3$ の場合

通信媒体の共有の有無による通信特性の差はプロセッサ数 $N \leq 3$ では生じない。すなわち, $M1$ と $M2$

および M3 と M4 はそれぞれ縮退している。この場合は通信の実行中に受信側のプロセッサが通信処理のため保留されるか否かで全体の通信要求発生量に差が生じ、実際に運ばれる通信量にも差が出る。受信処理のため受信側プロセッサが保留されるモデル (M1, M2) は、通信中も受信プロセッサが独自の処理を実行できるモデル (M3, M4) に比べて運ばれる通信量は 10~20% 程度小さい。すなわち、M1, M2 の方式は M3, M4 の方式に比べて通信オーバーヘッドが大きく、システム性能は 10~20% 低い。

(2) $N \geq 4$ の場合

プロセッサ台数 N が大きくなると M1, M2, M3 のモデルは解析が急速に困難となる。ここでは 3 章で得られた $N=4$ の結果をもとに $N > 4$ についても若干の考察を行う。

$N \geq 4$ では図 5 から分かるように、各プロセッサの通信要求 (λh) の大小によって各方式の通信量の大小関係が変わる。 λh の小さいときは通信媒体を共用するモデルでも通信媒体へのアクセス競合は小さく、各方式の通信量 C の間には (1) と同様の関係が成り立つ。これに対して、 λh が大きくなると、通信媒体を共用する方式 (M2, M4) はアクセス競合が大きくなり、通信媒体を共用しない方式 (M1, M3) に比べて運ばれる通信量 C は小さくなる。

すなわち、通信量 C は常に M3 の方式が最大、M2 の方式が最小であり、M1, M4 はそれらの中間にあって交差し、 λh が小のとき M4 の通信量の方が大きい。

4.2 スループット達成率

図 6 のスループット達成率 R はすべてのプロセッサが通信を行わずスタンド・アロン方式で稼働するときのスループットに比べて、通信を行うことで送受信負荷、通信待ちが生じ、スループットの低下する割合を示したものである。

スループット達成率についても通信量 C と同様のことが言えるが、プロセッサ間の通信がシステム性能に大きな影響を与えることが分かる。

4.3 プロセッサの受信負荷

プロセッサの受信負荷 ρ_c は通信量 C を N で割った値である。 ρ_c が最大となるのは $N=2$ の場合で最大値は 50% である。通信のオーバーヘッド*は送信

* 通信のない状態で 100% 稼働しているプロセッサの能力が通信による処理または待ち時間などによって何 % 能力を低下させるかを言う。ただし、ここでは送信前の送信文作成処理、受信後の受信文解析処理は見積りにくい面もありオーバーヘッドから除外している。

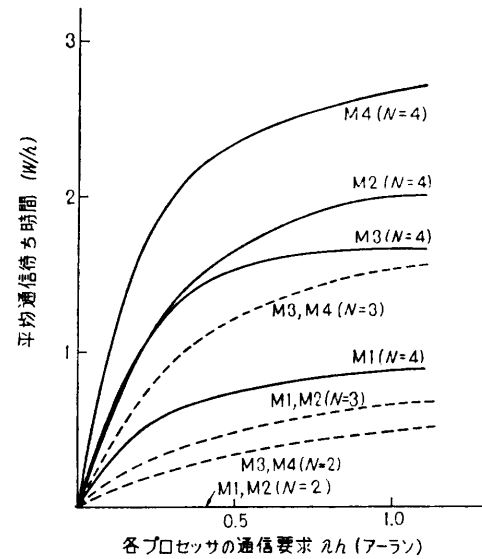


図 7 平均通信待ち時間

Fig. 7 Average waiting time of communication call.

処理、受信処理、送信待ちから構成されるものと考え、受信負荷 ρ_c の 2 倍以上が通信オーバーヘッドと言うことになる。プロセッサ間通信のあるシステムでは通信要求 λh を小さくすることがプロセッサ間の結合形態を工夫することにも増して重要と思われる。

なお、M4 では受信処理は受信プロセッサの空き時間に行うことができる点が有利である。

4.4 平均通信待ち時間

図 7 から分かるように、 $N \leq 3$ では平均通信待ち時間 W/h は、非同期式通信モデル (M3, M4) は同期式通信モデル (M1, M2) に比べてかなり大きい。また、 $N \geq 4$ ではメモリの介在しない、完全群結合方式 (M1) に比べてほかの方式の W/h は大きい。しかし、M1 は通信時に受信側のプロセッサが保留され、その間発信能力を失うため、システム全体の発呼量が小さく、 W/h が小さいのであり、図 5 から分かるように通信量 C は M3 の方が大きい。

本論文では通信待ちとなった通信要求を発生したプロセッサは停止状態 (スピンロック) で待つことを仮定しているが、通信要求の多いときはサスペンド・ロックなどの方法によって回線使用率を上げ、スループットの低下を防ぐことが大切である。

5. あとがき

複合計算機におけるプロセッサ間の通信方式を分類して通信方式を代表する 4 つの通信モデルをとりあ

げ、各モデルの通信特性を求めた。

スタンド・アローンのシステムに比べ、複数のプロセッサ間で通信のあるシステムでは通信オーバーヘッドとして、送信処理、受信処理のほかに通信媒体や受信プロセッサへの競合による通信待ちを負擔しなければならぬ。これらの通信オーバーヘッドはプロセッサ間の通信量が増せば急速に増大する。通信媒体への競合を無くすためプロセッサ間を完全群結合にしたとしても受信プロセッサへの競合などのため、通信媒体を増やした割にはスループットは改善されない。したがってこのようなシステムではプロセッサ間の結合方式を工夫するよりもまずプロセッサ間の通信量を減らすことが大切である。

本論文では各プロセッサは通信に関して均質（ホモジニアス）であることを仮定し、各通信モデルの性能上の差異の現れるプロセッサ4台までの通信特性を求めたが、今後、プロセッサ数5以上の場合やヘテロジニアスな通信モデルについての評価も必要である。

おわりに、日頃ご指導頂いている本研究所新井データ通信方式研究室長ならびに村岡調査役に深謝する。

参 考 文 献

- 1) 池原：待ち行列モデルによる機能分散型分散処理方式の性能評価，情報処理，Vol. 18, No. 11, pp. 1102-1107 (1977).
- 2) Ikehara, S.: Effects of the Dynamic Processor Scheduling in a Function Distributed System, J. of the Information Processing, Vol. 1, No. 2, pp. 85-91.
- 3) 池原：分散処理方式の性能評価，研究実用化報告，Vol. 27, No. 10, pp. 2225-2247.
- 4) 池原：パッシブ・サーバをもつネットワーク型待ち行列を用いた計算機の性能評価法，情報処理，Vol. 20, No. 2, pp. 105-112 (1979).
- 5) Maekawa, M.: Interprocess Communications in a High Diversified Distributed System, IF-IP '77, pp. 149-154.
- 6) Anderson, G. A.: Interconnecting a Distributed Processor System for AVIONICS, Symposium on Computer Architecture.
- 7) Anderson, G. A. and Jensen, E. D.: Computer Interconnection Structure: Taxonomy, Characteristics and Examples, Computing Surveys, Vol. 7, No. 4, pp. 197-213 (Dec. 1975).
- 8) Wong, J. W. and Unsoy, M. S.: Analysis of Flow Control in Switched Data Networks, Information Processing 77, pp. 315-320.
- 9) 池原：マルチプロセッサ方式における共用メモリアクセス競合の解析，信学論 D Vol. 63-D, No. 4, pp. 334-341.
- 10) Skinner, C. E. and Asher, J. R.: Effects of Storage Contention on System Performance, IBM SYST. J., No. 4, pp. 319-333 (1969).
- 11) 村田，秋山：高密度トラヒックに関する検討，電子通信学会交換研究会，SE 77-78.
- 12) 池原，岡田：数式処理言語 AL とその評価，情報処理学会論文誌，Vol. 20, No. 2, pp. 158-165 (昭 54-3).

(昭和 55 年 1 月 21 日受付)

(昭和 55 年 9 月 18 日採録)