

DB/DCシステムの密結合マルチプロセッサにおけるリソース競合に関する一考察

2H-7

根岸 和義

塚野 博実

緒方 博通

(日立システム開発研究所) (日立マイクロコン

ピュータエンジニアリング)

1. はじめに

近年コンピュータシステムに対するオンライン・ニーズが増大している。これとともにシステムの処理能力に対する要求も高水準化している。これに対処する方策の一つとして、複数のCPUで主記憶を共用する密結合マルチプロセッサ上でのDB/DCオンライン処理がある。このシステムでは、複数CPUで同時に実行されるトランザクション (TR) 間のリソース要求の競合が性能に大きく影響を与えることが予想される。我々はリソースの競合をモデル化し、CPUの処理能力を予測する方策を検討してきた。本報告では、リソースのうちで競合発生時に別のTRをディスパッチすることのできるディスパッチ可リソース (例DB/DC内のリソース) について考察を行った。

2. モデル化の前提条件

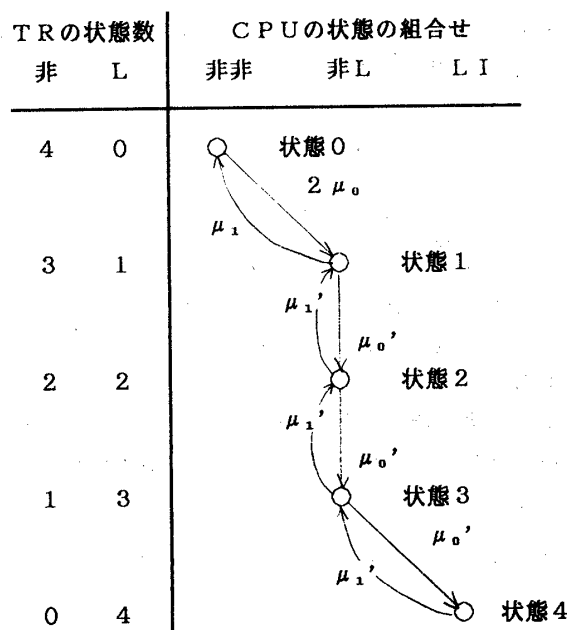
処理能力評価モデルを設定する際の条件を以下に示す。

- (1) 各CPUは切れ目なくTRの処理を行う。(一つ終わったらすぐ次のTRの実行を開始する)。
- (2) TRはリソースの占有/非占有状態の処理を同じパターンで繰り返す同種のTRとする。
- (3) TRの開始終了およびI/Oの発行によるTRの入れ替わりの回数はTRのリソース要求の回数と比較して十分少でこれらによる状態の変化は無視できるものとする。
- (4) リソースは単一の排他リソースとする。
- (5) TRによるリソースの占有/非占有処理時間はそれぞれ平均 $1/\mu_1$, $1/\mu_0$ の指数分布とする。
- (6) 二台以上のCPUがリソースを要求すると、最初の一台以外はこのリソースを使用しない別のTRの処理に切り替えられる。このようなTRのない場合、そのCPUはアイドル状態となる。

3. 性能評価モデル

(1) ディスパッチオーバーヘッドを考えない場合
各CPUはそれぞれ占有、非占有、アイドルのいずれかの状態をとると考える。個々のCPUを区別する必要のないこと、占有状態のCPUは一台のみであることから順序を無視したCPU状態の組合せをシステムの状態として考える。

二台のCPUから成るシステムにおけるCPUの状態およびこの間の状態遷移 (TR数4の例) を図1に示す。ただし、 $1/\mu_0'$, $1/\mu_1'$ はここでは $1/\mu_0$, $1/\mu_0$ と同じとする。



[注] 非: ロック非占有状態
L: ロック占有状態, I: アイドル状態

図1 システム状態遷移図

(2) ディスパッチオーバーヘッドを考えた場合
リソース競合発生時に別TRへのディスパッチが発生する。このディスパッチのオーバーヘッド時間を $1/\mu_{ov}$ とする。ディスパッチを伴う状態遷移に関するロックの占有/非占有サービス時間を、下記を平均とする指数分布とする。

$$\frac{1}{\mu_0'} = \frac{1}{\mu_0} + \frac{1}{\mu_{ov}}$$

$$\frac{1}{\mu_1'} = \frac{1}{\mu_1} + \frac{1}{\mu_{ov}}$$

TRおよびCPUの状態の組合せはTR4個の場合、図1のようになる。
各状態0, 1, 2, 3, 4の確率を P_0, P_1, P_2, P_3, P_4 とすれば、これらの関係は次のように表される。

$$P_1 \cdot \mu_1 = P_0 \cdot 2\mu_0$$

$$P_2 \cdot \mu_1' = P_1 \cdot \mu_0'$$

$$P_3 \cdot \mu_1' = P_2 \cdot \mu_0'$$

$$P_4 \cdot \mu_1' = P_3 \cdot \mu_0'$$

また、

$$\sum_{i=1}^n P_i = 1 \text{ である。}$$

これらより、 P_0, P_1, P_2, P_3, P_4 の値を求めることができる。

$$P_0 = \frac{a a'^3}{G_4}$$

$$P_i = \frac{2 a'^{4-i}}{G} \quad (i \geq 1)$$

ここで、 $a = \mu_1 / \mu_0, a' = \mu_1' / \mu_0'$

$$G_4 = (a+2) a'^3 + 2 a'^2 + 2 a + 2$$

である。

また、処理能力 Power は、オーバーヘッドのないCPU一台の処理能力を1とした時下記のようになる。

$$\text{Power} = 2 P_0 + \left(\frac{\mu_0'}{\mu_0} + \frac{\mu_1'}{\mu_1} \right) \times (P_1 + P_2 + P_3) + \frac{\mu_1'}{\mu_1} P_4$$

さらに、 $1/\mu_0 + 1/\mu_1$ に対する $1/\mu_1, 1/\mu_{ov}$ の割合をそれぞれ b, b_0 とすると、

$$\mu_{ov} = O_{ov} \cdot \mu_0$$

$$O_{ov} = (1-b) / b_0$$

$$a' = \frac{a(1+O_{ov})}{a+O_{ov}}, a = \frac{1-b}{b}$$

と表される。(bはロック処理の割合を表す)

同様に、TR数m個の場合の処理能力は以下の通りである。

$$G_m = a a'^{m-1} + \frac{2(a'^m - 1)}{a' - 1}$$

$$\text{Power} = \left[2 - \left(\frac{\mu_0'}{\mu_0} + \frac{\mu_1'}{\mu_1} \right) + \left(1 - \frac{\mu_1'}{\mu_1} \right) \frac{2}{a} \right] P_0 - \frac{\mu_0'}{\mu_0} P_m + \left(\frac{\mu_0'}{\mu_0} + \frac{\mu_1'}{\mu_1} \right)$$

ここで、 $P_0 = \frac{a a'^{m-1}}{G_m}, P_m = \frac{2}{G_m}$

である。

図2に $b_0 = 1/4$ および $b_0 = 0$ の場合の処理能力を示す。

処理能力 (Power)

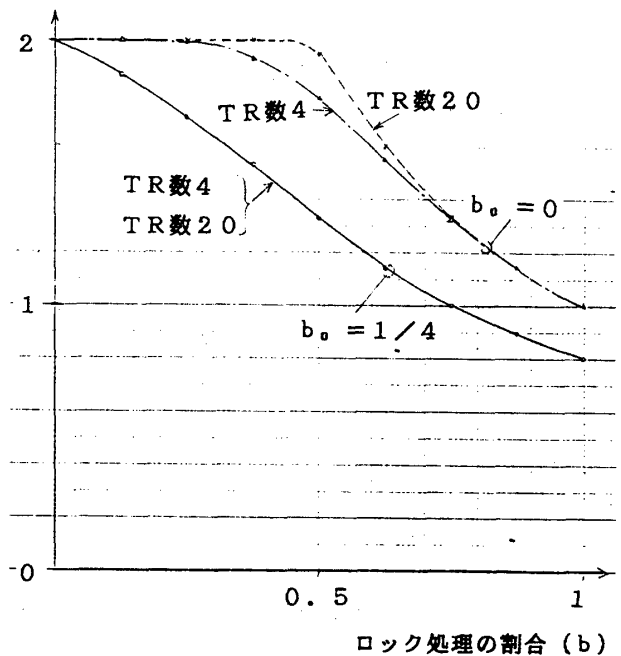


図2 ロック率に対するシステムの処理能力

4. まとめ

密結合マルチプロセッサにおけるディスパッチ可リソースの排他競合時の処理能力(スループット)を状態遷移モデルにより求めた。また、ディスパッチ発生時のオーバーヘッドを考えた場合の評価をおこなった。

〔参考文献〕

- 1) L. Kleinrock: 待ち行列システム理論(上), マグロウヒル好学社, 1979
- 2) 根岸他「DB/DCシステムの密結合マルチプロセッサにおけるリソース競合に関する性能評価」情報処理学会第32回全国大会2C-8