

情報処理学会 第34回全国大会 論文集-4

並列推論マシン PIM

- TM-0262 (1) クラスタ間に渡るゴール管理方式
TM-0263 (2) KLI の抽象命令仕様とコンパイラ
TM-0264 (3) クラスタ内処理のメモリアクセス
TM-0265 (4) クラスタ内実験処理系
TM-0266 (5) 並列キャッシュとロック機構

- (0262) 六沢一昭, 宮崎敏彦, 澁 和男,
(0263) 久門耕一(富士通), 中島 浩(三菱電気), 木村康則, 近山 隆(ICOT)
(0264) 中川貴之, 松本 明, 佐藤正俊, 清水 肇, 後藤厚宏,
(0265) 佐藤正俊, 清水 肇
(0266) 松本 明, 中川貴之, 後藤厚宏

February, 1987

©1987, ICOT

ICOT

Mita Kokusai Bldg. 21F
4-28 Mita 1-Chome
Minato-ku Tokyo 108 Japan

(03) 456-3191~5
Telex ICOT J32964

Institute for New Generation Computer Technology

並列推論マシン PIM

— クラスタ間を渡るゴール管理方式 —

六沢 一昭 宮崎 敏彦 瀧 和男
(財) 新世代コンピュータ技術開発機構

1. はじめに

並列推論マシン PIM は、共有メモリを用いた密結合マルチプロセッサからなるクラスタを連結した構成をとり、クラスタ間に係る処理はすべてメッセージ通信により行なう[1]。従って、K&L1(核言語第一版)を PIM で実行する場合、クラスタ間に渡ったゴールをメッセージ通信によって管理することが必要である。

本稿では、このクラスタ間に渡ったゴールの管理方式について述べる。

2. クラスタ間ゴール木の構成

すべてのゴールは図1に示すゴール木で管理する。システムにはいくつかのメタゴールがあり、ユニークな識別名(これをメタゴールIDと呼ぶ)をもっている。すべてのゴールは1つのメタゴールに属し、メタゴールノードを単方向がインタで指す。

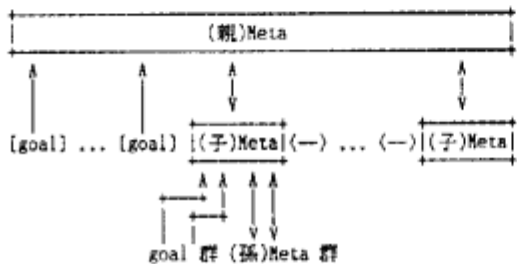


図1. ゴール木

メタゴールには exec / fail の status があり、ゴールがスケジュールされ実行する際チェックする (fail ならば実行しない)。従って、ゴールがクラスタ間に渡るとき、図1のゴール木をそのままクラスタ間にマッピングすると、メタゴールノードとは異なったクラスタにあるゴールがスケジュールされるたびにクラスタ間通信が起こってしまう。

このため、ゴールがクラスタ間に渡る場合のゴール木は、図2のように里親ノードを設け、ゴールは里親ノードを介してメタゴールノードを指す構成とした。里親ノードはメタゴールの status をキャッシュする働きをする。また里親は、1つのメタゴールIDについて1クラスタあたり1つだけ設ける。このクラスタ間に渡るゴール木を「クラスタ間ゴール木」と呼ぶ。以下、ゴール木といった場合、このクラスタ間ゴール木を指す。

メタゴール及び里親は、自分がクラスタ内に抱えているすべてのゴールが成功し、下位の里親がすべて消滅すると消滅する。またゴールが fail した場合は、下位の里親を強制的に消滅させる。このため各ノードは、自分の下位に接続している里親がいくつあるか(下位の里親の消滅検出に使用)、またどのクラスタにあるか(下位の里親消滅要求メッセージ送信先を示す)を覚えておく必要がある。

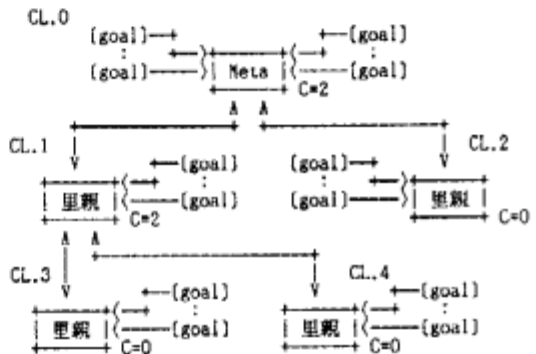
以上のことからノードには、status やメタゴールID、上位ノードの存在するクラスタ#の他に、各クラスタ内の里親との接続 status、接続している里親数を示すフィールドを設けた。ノードの構成を図3に示す。

3. クラスタ間ゴール木の管理

ゴールを他のクラスタへ送出すると、クラスタ間に渡るゴール木ができる。すべてのゴールが終了するとゴール木は消滅する。fail した場合は強制的にゴール木を消滅させる。このようにプログラムの実行はゴール木の生成/消滅を繰返しながら進む。従って、ゴール木の動的な管理が必要である。

またクラスタ間メッセージ(以下、「メッセージ」と略す)はいつ目的クラスタへ到着するか保証がない。このため、他クラスタの状態を知りそれに基づきメッセージを送信しても、メッセージが到着した時は既に状態が変化している可能性がある。

以上の点を考慮したゴール木の生成/消滅方式を以下に示す。ゴール木の生成/消滅は、表1に示すメッセージの送受信及び図3に示したノードの各フィールドを書き替えることにより行なう("x"はメッセージを示す)。



C : 下位に接続している里親数(図3参照)
CL.n : クラスタ n

図2. クラスタ間ゴール木(1メタゴールのみ)

ノードの status
メタゴールID
上位ノードの存在するクラスタ#
クラスタ 0 内の里親との接続 status
⋮
クラスタ N-1 内の里親との接続 status
下位に接続している里親数(C と略す)

ノードの status = exec / fail
exec : 通常状態
fail : fail 状態

接続 status = active / inactive
active : 対応するクラスタ内の里親と接続している
inactive : 対応するクラスタ内の里親と接続していない

図3. メタゴールノード、里親ノードの構成
(クラスタ間管理に関係する部分のみ)

Parallel Inference Machine : PIM
-- Inter-Cluster Handling of Goals --
Kazuaki ROKUSAWA, Toshihiko MIYAZAKI, Kazuo TAKI
Institute for New Generation Computer Technology (ICOT)

3-1. ゴール木の生成(図4参照)

ゴールの送出は、以下の手順で行なう。

① 接続している里親数(以下、Cと略す)+1

② %goal 送信

このゴール送出によって新たな里親が生れる可能性がある。Cを+1してから%goalを送信する。

%goalを受信したクラスタでは以下の処理を行なう。

① メタゴールあるいは里親の存在チェック

② 存在するならば、ゴールをつなぎ、%cancel 返信
存在しないならば、里親を作ってゴールをつなぎ、%ready 返信

再び%goal送信元クラスタでは、以下の処理を行なう。

・%cancelを受信した場合(これは、「ゴールは受取ったが里親は生成しなかった」ことを示す)

① $C := C - 1$ (ゴール送信時の+1を元に戻す)

・%readyを受信した場合(これは、「ゴールを受取り新たに里親を生成した」ことを示す)

① (%ready送信元クラスタ内の里親との)

接続 status := active (接続が生成)

ゴール木は上記の処理を繰り返しながらクラスタ間を広がって行く。

3-2. ゴール木の消滅(図5参照)

ゴール木の消滅は、末端のノードから「消滅 & 消滅を上位へ伝えるメッセージ送信」の手順で行なう。前述したように、メッセージはいつ目的クラスタへ到着するか保証

がない。このため、上位のノードが先に消滅する方式では、図6のように処理が正しく行われない可能性がある。

ゴール木の末端にある里親は、処理するゴールがなくなると消滅し、%termを送信する。

%term受信は、送信元クラスタの里親及び接続消滅を意味するので、%termを受信したクラスタでは、以下の処理を行なう。

① $C := C - 1$ (下位の里親が減った)

② 接続 status := inactive (接続消滅)

末端でない里親あるいはメタゴールは、ゴールがなくなり、かつ下位の里親の消滅を確認する(Cが0になる)のを待って消滅する。

ゴールがfailした場合の強制的なゴール木の消滅は、下位に接続している里親へのfailゴールの送出(%killの送出)により行なう。

4. おわりに

クラスタ間に渡ったゴールを管理するため、クラスタ間ゴール木を示し、それをクラスタ間メッセージにより動的に管理する方式を述べた。

(参考文献)

- 1) 後藤、松本、佐藤 他, "並列推論マシン P I M," 情報処理学会第33回全国大会予備集 3B-5~7.
- 2) 宮崎 他, "multi-PSI における FlatGHC の実現方式," Logic Programming Conference '86, pp.83-92.

表1. クラスタ間メッセージ(ゴール木の生成/消滅に関するもののみ)

名前	意味	送信方向	下位に接続している里親数
%goal	ゴール送出 & 里親生成要求	上位 → 下位	送信時 +1
%ready	里親生成(%goalの返答)	上位 ← 下位	
%cancel	里親既存在(%goalの返答)	上位 ← 下位	受信時 -1
%kill	里親消滅要求(fail ゴール送出)	上位 → 下位	送信時 +1
%fail	fail の伝達	上位 ← 下位	
%term	里親消滅	上位 ← 下位	受信時 -1

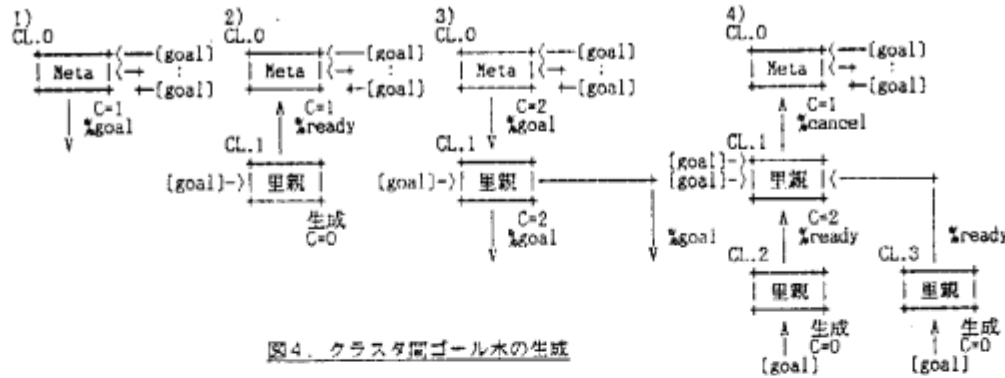


図4. クラスタ間ゴール木の生成

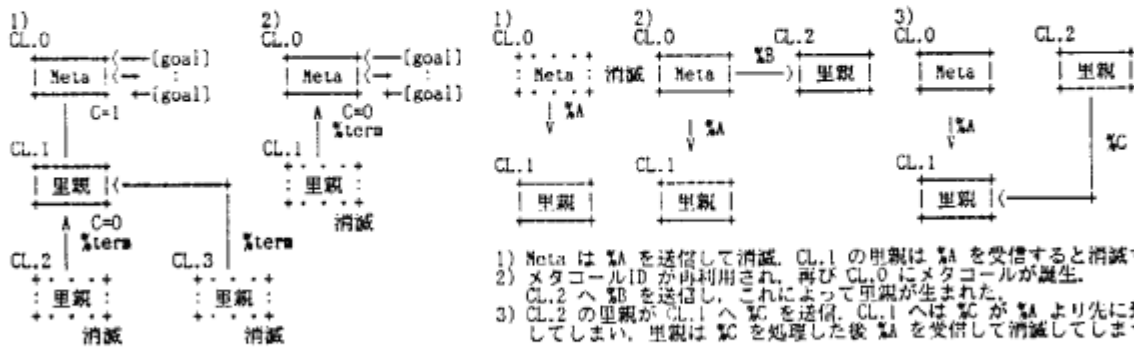


図5. クラスタ間ゴール木の消滅

- 1) Meta は %A を送信して消滅。CL.1 の里親は %A を受信すると消滅する。
- 2) メタゴールID が再利用され、再び CL.0 にメタゴールが誕生。CL.2 へ %B を送信し、これによって里親が生まれた。
- 3) CL.2 の里親が CL.1 へ %C を送信。CL.1 へは %C が %A より先に到着してしまい、里親は %C を処理した後 %A を受信して消滅してしまう。

図6. ゴール木の正しくない消滅手順