

階層化モデルに基づく  
汎用離散系シミュレーション用  
並列処理アーキテクチャ

AN ARCHITECTURE FOR PARALLEL DISCRETE EVENT SIMULATION

工藤知宏 天野英晴

Tomohiro KUDOH Hideharu AMANO

慶応義塾大学理工学部電気工学科

Department of Electrical Engineering, Faculty of Science and Technology, Keio University

あらまし 離散事象モデルのシミュレーションを並列処理するための専用システムEOSのアルゴリズム及びアーキテクチャについて述べる。

EOSでは問題の記述モデルとして待ち行列網をフローグラフを用いて表現する手法を採る。さらに問題の性質を利用してモジュール化したモデルを構築し、時刻管理、事象管理も分散する。

このアルゴリズムでは分散時刻管理において問題となる、事象の処理を行なってよいかどうかの検証をおこなうオーバーヘッドを避けるために、履歴を残しつつ処理を進め、後で不当な処理については取り消すという方法を用いている。

またEOSにおける問題の記述についても簡単に触れる。ここでは、モジュール化とオーバーヘッドの軽減のためにユーザが積極的に情報を与える方法を用いている。

Abstract: Parallel processing algorithm and architecture to process discrete event simulation are proposed in this report.

We adopt the flow graph model to describe target problems. By dividing the flow graph into several modules, virtual time controls and event controls are distributed. To enhance the performance and to guarantee consistency, an anticipation mechanism with history recording is used. If anticipations are illegal, anticipation processes are canceled using the records.

### 1. はじめに

計算機に対する需要の中で離散系シミュレーションの占める割合は大きい。現在離散系シミュレーションはGPSS、Simscrip、Simulaなどの言語で記述され大型汎用計算機上で解かれるのが一般的であるが、離散系シミュレーションは非常に計算機コストを消費するためこれまで大きな問題を解くことは難しかった。離散系シミュレーションを高速に行なうことができれば、これまで扱えなかったような大きな問題をも扱うことができるようになり、その意義は大きい。

一方離散系シミュレーションは並列性を有す

る典型的な問題である。その代表的な例としては、論理回路、通信網、工場の生産ライン、交通、計算機等のシステムなどのシミュレーションが挙げられる。これらの系では系の構成要素は並列に動作している。そこでこの問題固有の並列性を利用して並列処理を行なうことにより処理速度を向上させることが考えられる。

これまで、科学技術計算や推論などを対象とした並列処理は盛んに研究され、すでに実用機も作られているが、離散系シミュレーションの並列処理についてはいくつかの基礎的な研究は報告されているものの、論理回路専用のシミュ

レータを除けば実用的なアーキテクチャは提案されていない。これは、主として離散系シミュレーションには、時刻管理等の並列処理に際し問題となるような点があるためである。

### 1. 1. 事象駆動と時刻駆動

単一プロセッサ上で離散系シミュレーションを行なうもっとも一般的なモデルは、シミュレーション仮想時刻を持つ時計と生起時刻順に並べられた事象の待ち行列（イベントキュー）があるもので、時計の示す時刻の事象の処理がすべて終了すると待ち行列の先頭にある事象（最も近い時刻に起こる事象）の生起時刻まで時計が進められその時刻に生起する事象の処理が行なわれるというものである。このような手法を事象駆動と呼ぶ。これに対し、シミュレーションの対象となる系が一定周期で同期して変化し同一周期に状態が変化する処理体の数が多いときにはシミュレーション仮想時刻を一定間隔づつ進めその都度全処理体について処理を行なう手法が用いられる。これを時刻駆動（タイムスライス法）と呼ぶ。単一プロセッサ上で処理を行なう場合には、事象駆動によるか時刻駆動によるかは問題の性質により効率のよい方をとるのが一般的である。

複数プロセッサ上で処理を行なう場合には、事象駆動と時刻駆動では取り扱い易さが違ってくる。事象駆動ではイベントキュー及び時計の管理が複雑であるため、効率よく高い並列性を得ることが難しい。これに対し時刻駆動ではイベントキューを持つ必要がなく並列処理は比較的容易であるが、時刻管理の融通性が低くまた、各時刻に状態が変化する処理体の数が少なければ無駄が大きくなり効率が上がらない。即ち単一プロセッサ上でも時刻駆動を用いるような問題に対しては有効であるが、一般の問題には適していないといえる。

EOSでは事象駆動を用いて並列処理をおこなう。

### 1. 2. 時刻管理

事象駆動では特に時刻管理の手法が問題であ

る。時刻管理の手法としては大きく分けて、集中時刻管理と分散時刻管理がある。集中時刻管理は処理系全体で一個の時計のみを持つもので、離散系シミュレーションにおける時刻管理の方法としては最も基本的なものである。この方法は、実現は容易であるが時刻管理が律速となる。時刻管理が一カ所で行なわれる場合、イベントキューを分散すると時刻更新の管理が困難になるので通常イベントキューも一カ所に置く。そうするとプロセッサ数をいくら増やしても時計とイベントキューの管理のオーバーヘッドが大きく効率が上がらない。また、イベントキューが一つだと同一時刻のイベントしか並列処理ができず、一般の問題では並列処理をしても効果が小さい。一方、分散時刻管理では、独立な事象は同一時刻のものでなくとも並列に処理することが出来、時刻管理が律速になることもない。しかし、独立な事象であるかどうかを判別することが難しい。

既存の並列離散系シミュレーションのためのアーキテクチャは論理回路等の専用目的のものがほとんどで、集中時刻管理を用いている。これらのシステムはキューを管理する部分と数個の高速な処理装置からなり処理装置の個数はあまり多くない。これは処理装置の数を多くしても集中時刻管理では並列に処理できるのは同一時刻に生起する事象のみであること、時刻管理と各処理装置の処理能力のバランスから、処理装置の数だけ増やしても時刻管理が律速になるという理由からである。

同一時刻に生起する事象以外のものも並列に処理しようとするれば分散時刻管理を用いることになる。分散時刻管理ではAという事象を処理するとき他の部分ではまだAよりも前の時刻の事象Bを処理していないこともあり得、AがBに影響を受けることがないことが保証されている場合のみそのAの処理を行なうことができる。この保証のためのアルゴリズムは既に提案され、その正当性が証明されているが[1]高速に処理することは困難である。

ところで離散系シミュレーションの対象となるような問題には多くの場合モジュール性、階

属性が存在する。例えば交通のシミュレーションでは町ごとにモジュールとすることができるしその町も駅、道路といった単位に分割できる。計算機システムのシミュレーションではプロセッサ、入出力装置といった構成要素ごとにモジュール化できるしこれらの構成要素もさらに分割して扱えることが多い。

そこでEOSではこの問題の性質を利用して階層的に時刻を管理する集中時刻管理と分散時刻管理を組み合わせた手法を用いる。

### 1. 3. フローグラフ

シミュレーションの系を示すためにEOSではフローグラフに基づくモデルを用いる。このモデルでは系全体はフローグラフと呼ばれる有向グラフを用いて表される。(図1)このグラフは現実の待ち行列網システムにおける能動的要素である窓口、交換機等をノードにより表現し、客の通り道や通信回線をノード間を結合するリンクとして表現する。また、受動的要素である客や通信の呼はグラフ内を流れるトークンで表現する。トークンにはシミュレーション仮想時刻により時刻印が押される。時間消費はノード内のみで行われる。

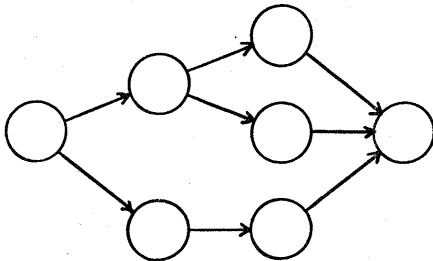


図1 フローグラフ

### 2. 処理モデルとアルゴリズム

EOSではフローグラフは問題の性質に従いいくつかのモジュールに分割され、さらに各々のモジュール内も分割してモジュールを構成することにより多重に階層化できる。この分割階層化された系はモジュール間の包含関係にしたがって木構造を構成する。この木の根は系全体を表す。木の葉はフローグラフの最小構成要素であるノードを1つ以上含んでいる。(図2)

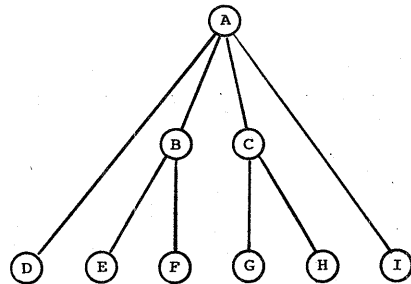
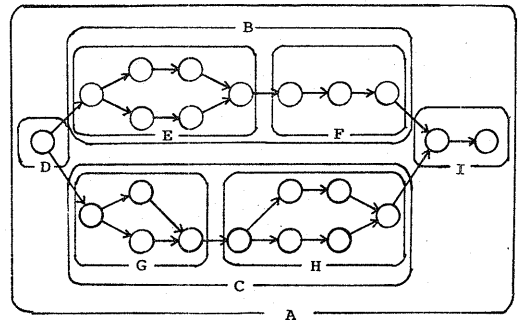


図2 フローグラフの階層、モジュール化

### 2. 1. 処理系とアルゴリズムの概要

EOSの処理系も処理モデルに対応して木構造をとる。ここで各葉を担当する処理プロセッサと各節を担当する同期プロセッサ、根を担当するルートプロセッサを想定する。(図3)何れのプロセッサもシミュレーション仮想時刻を示す時計VTを持つ。ルートプロセッサのVTは系全体の基準となり、GV Tと呼ばれる。葉を担当する処理プロセッサは、VTとともにイベントキューEQを持っている。詳しいアーキテクチャについては後述する。

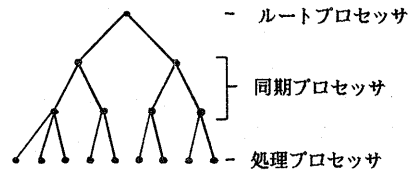


図3 EOSの処理系

事象の処理は処理プロセッサによって行なわれる。処理プロセッサはまずGV T時刻の事象を処理する。これは一般の集中時刻管理においても行なわれることで、この処理は安全である。GV T時刻の処理をすべて終わると、処理プロ

セッサは、各々のEQの内容を各々のVTを更新しながら処理する。これは外部から更新後のVT時刻よりも古い時刻印を持ったトークンが到着する可能性があるので、安全ではない。この処理を先取り処理と呼ぶ。より古い時刻印を持ったトークンが到着した場合、先取り処理は非合法であったという。先取り処理時には履歴が残され、処理が非合法であった場合には履歴により非合法的な処理以前の状態に戻る。これを回復という。

処理プロセッサ外とのトークンのやりとりは親の同期プロセッサを通して行なわれる。GVTはすべての処理プロセッサのVTの最小値に更新できるので、同期プロセッサは子プロセッサのVTの最小値を親のプロセッサに伝える。

GVTが更新されればそれ以前の履歴は不用となる。そこでGVTが更新されると新しいGVTの値が全プロセッサにブロードキャストされる。

プロセッサ間の情報のやり取りはすべてメッセージにより行われる。トークンもメッセージとして受け渡され、これをトークンメッセージという。

## 2. 2. 処理プロセッサのアルゴリズム

処理プロセッサは事象の処理を行う。もし、自身の担当しないノードと接続しているリンクへのトークンの出力がある場合にはトークンメッセージが親の同期プロセッサに送られる。

シミュレーションの開始時にはVTはGVTと同じ時刻を示している。処理プロセッサはこの後次のような処理を行う。

(1) まずVT時刻の事象をEQから取り出し処理を行う。これは安全な処理である。親からトークンメッセージを受け取ったらEQに登録する。

(2) VT時刻のイベントを全て処理し終わると、EQに事象があれば先取り処理を行う。この際VTを更新するごとに、新しいVTの値を親である同期プロセッサに送る。EQが空になればその旨のメッセージを親である同期プロセッサに送る。ここで、

・処理プロセッサは全ての事象の履歴を保存している。

・親のプロセッサからトークンメッセージを受け取り、そのメッセージの時刻印が現在のVTよりも小さければ、VTをその時刻印の時刻まで回復し、履歴により自身の担当するノード、トークンの状態をその時刻以前の状態まで回復し、VTを回復する旨のメッセージを親プロセッサに送る。

・親から受け取ったトークンメッセージの時刻印がVTよりも大きければEQに登録する。

・既に受け取ったトークンメッセージを取り消すというメッセージを親から受け取った場合には、取り消すメッセージがEQ内にまだ登録されていればそれを取り消し、既に処理されていれば処理以前の状態まで回復する。

(3) ルートプロセッサから新しいGVTの値を受け取ると、GVT時刻以前の履歴を消去する。また、VTがGVTよりも小さければVTの値をGVTと等しくする。

(4) VTがGVTと等しければ(1)に、そうでなければ(2)にもどる。

## 2. 3. 同期プロセッサのアルゴリズム

同期プロセッサは処理プロセッサ間のトークンメッセージのやり取りと、時刻の管理を行う。具体的には以下の処理を随時行う。

(1) 子プロセッサからのトークンメッセージは受け手が自分の子である場合には、その子に送る。そうでなければ、さらに親のプロセッサに送る。この時メッセージの履歴は保存する。

(2) 親プロセッサからメッセージを受け取った場合には、宛先のノードを担当する子プロセッサにそのメッセージを送る。

(3) トークンメッセージを送った子プロセッサからVTを回復する旨のメッセージを受け取った場合には、履歴を検索し、その子プロセッサが非合法的な処理中に出した

トークンメッセージがあれば、宛先のプロセッサにそのトークンメッセージを取り消すという内容を持ったメッセージを伝達する。また、回復した後のVTがそのプロセッサ自身のVTよりも小さければ、自分自身のVTを回復し、その旨を親のプロセッサに伝える。

(4) 同期プロセッサは子プロセッサのVTの値を常に知っており、EQが空でない全ての子のVTの中で最小のものを自身のVTとする。全ての子プロセッサのイベントキューが空の場合には、その旨を親プロセッサに知らせる。子プロセッサのVTの変化により自身のVTが変化した場合、VTの値を自身の親であるプロセッサに知らせる。

(5) ルートプロセッサがGVTをブロードキャストしたときには、GVT以前の履歴を消去し、もしその値がVTよりも大きければ、VTをGVTと等しくする。

## 2. 4. ルートプロセッサのアルゴリズム

ルートプロセッサはGVTの管理と子プロセッサ間のトークンメッセージの伝達を行う。

(1) ルートプロセッサは子プロセッサのVTの値を常に知っており、全てのEQが空でない子プロセッサのVTの中で最小のものを自身のGVTとする。全ての子プロセッサのEQが空ならば、そのシミュレーションは終了である。子プロセッサのVTの変化により自身のGVTが変化した場合にはその値を全てのプロセッサにブロードキャストする。

(2) 子プロセッサからのトークンメッセージは受け手を担当する子プロセッサに送る。

(3) 同期プロセッサと同様にトークンメッセージの履歴を持ち、子プロセッサからVTの回復を知らされた場合には、履歴を検索し、子プロセッサが非合法的な処理中に出したトークンメッセージがあればそれを取り消すメッセージを宛先の子プロセッサ

に送る。但し回復したVTがGVTよりも小さいことは有り得ない。

## 3. アーキテクチャ

### 3. 1. 設計方針

2章に述べたアルゴリズムは問題によってはかなり大きな並列性を持っている。しかし、ある種の科学技術計算とは異なり、その並列性は平均的には数十から数百のオーダであることが予測される。このため、EOSのアーキテクチャは以下のような方針で設計される。

(1) 比較的少数(数十台)のプロセッサを用いる。

(2) 各プロセッサは処理機能毎に専用化し、機能の一部をハードウェア化する。また、プロセッサ内の処理における細かいレベルの並列性もできる限り利用する。

### 3. 2. 全体構成

EOSアーキテクチャの全体構成は図3に従う。EOSアーキテクチャは多分岐ツリー構造を持つ。ツリーの根をルートプロセッサといい、中間節を同期プロセッサ、葉は処理プロセッサという。原則としてプロセッサ間のメッセージ交換はツリーのリンクによって行われるが、根から全プロセッサに対してのグローバルバスが存在し、これが唯一の大域交信路である。

ツリーの最小システムは2階層であり、同期プロセッサを持たない。プロセッサ数が多い場合は、同期プロセッサを増やすことで、階層のレベルを大きくする事ができる。各プロセッサは2章で既に述べた処理を高速に実行するため以下のような構造を持つ。

#### (1) 処理プロセッサ

処理プロセッサはフローグラフに沿って、シミュレーションを実行するのが主な役割であるが、その全体の処理量のかかなりの部分はEQ、VTの管理が占める。このことに対応するため、処理プロセッサは2つの独立した処理単位(図4)から構成され、これらの分散処理により高速化を図る。これらは以下の機能を持つ。

・システム管理エンジン

E Q, V T の管理、同期プロセッサに対するメッセージの授受の制御を行う。

「回復」の際の履歴の検索等は非常な高速処理が要求されるため、このエンジンはマイクロプログラム制御による高速専用処理を行う。

・ノード実行エンジン

ユーザにより記述された実際のプログラムを実行する。コンパイラの作成を容易にするため汎用マイクロプロセッサを用いる。

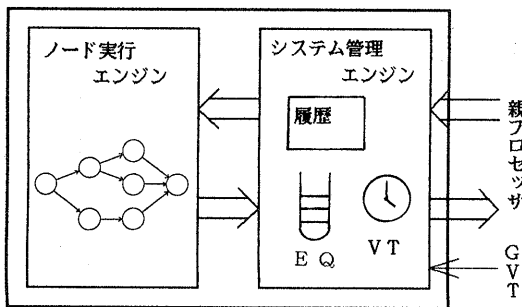


図4 処理プロセッサ

(2) 同期プロセッサ

同期プロセッサの構造を図5に示す。同期プロセッサの主な役目はメッセージのルーティング、メッセージの履歴の管理と検索、子プロセッサのV Tの収集、管理である。このため同期プロセッサはメッセージルータ、メッセージ履歴管理、V T管理の3つの部分から構成される。これらの操作は専用化され高速性が要求されるため、同期プロセッサは、処理プロセッサのシステム管理エンジン同様マイクロプログラム制御の専用プロセッサである。

子または親プロセッサからメッセージが到着すると、まずメッセージルータによって受け付けられ、シミュレーションのプログラム上で用いられるトークンメッセージか、V T更新のメッセージかチェックされる。V T更新のメッセージであった場合、子のV Tの値は保存され、他の子のV Tの

値と比較されて必要に応じて自分自身のV Tの値の更新が行われる。そうでなければ、メッセージは宛先にしたがってルーティングされるとともに履歴が保存される。「回復」の際はこの履歴が高速に検索される。

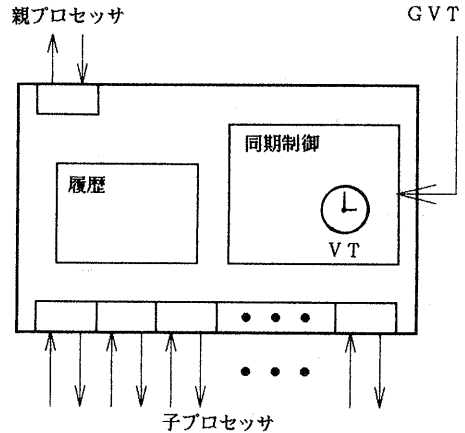


図5 同期プロセッサ

(3) ルートプロセッサ

構造はほぼ同期プロセッサと同様であるが、グローバルバスにGVTをブロードキャストする機能を持つ。

これらのプロセッサの数の設定と、具体的な構造を定めるためにはこのアーキテクチャ自身のシミュレーションを行う必要があり、現在評価検討中である。

4. 問題の記述

4. 1. 従来言語の問題点

EOSの時刻管理アルゴリズムはGPSSを始めとするイベント駆動型の離散系シミュレーション用言語全てに適用することができる。しかし、現在用いられている言語はEOSにおいて問題の記述に用いるには次のような問題点がある。

(1) GPSSはモジュール性、階層性が弱い。ユーザの書いたプログラムを自動的にモジュール化し、プロセッサにマッピングする必要がある。この際、うまくモジュール化できないと、回復のための作業が頻

繁になり効率が低下する。

(2) Simula, Mayはモジュール性を持つが、言語自体柔軟性が大きいため実装上の問題が多く複数のプロセッサでの高速実行は困難である。またフローモデルとの適合性にも欠けている。

EOSにおけるシミュレーションは対象とする実世界に既にモジュール性、階層性がある場合に効率高い。そこで、我々は、やや記述における柔軟性を犠牲にしても実装が容易であり、かつEOSに適した問題に対しては十分な記述力を持つ言語HNCC (Hierarchical Node Oriented Concurrent C) を提案する。プログラマがHNCCを用いて記述すると、自然にモジュール化、階層化が行われる。このため、このモジュール、階層の単位でプロセッサに対してマッピングを行えば、EOSにおいて効率よく実行されるプログラムを書くことができる。

#### 4. 2. HNCC

HNCCは、並行プロセス言語NCCを階層化し、離散系シミュレーション用の機能と関数を加えたものである。NCCは静的なノードとネットワークによって系を表現する。ノードはポートを介して他のノードとの交信を行うが、各ノードは自分のポート名のみを知っていればよく、ポート間の結合は独立に記述できる点がこの言語の大きな特徴である。また、交信は1対複数の転送即ちマルチキャストで行われる。完全な同期通信(ランデヴ)はマルチキャストの効率を悪化させるためNCCにおいては受信ポートにサイズ1のバッファを設ける。NCCの詳細は[2]を参照されたい。ここでは並列計算機のシミュレーションプログラムを例に簡単にHNCC特有の機能に関して解説する。

図6にシミュレーションを行う対象の並列計算機システムを示す。このシステムはバスによりプロセッサと共有メモリが結合された非常に簡単な構造を持つ。各プロセッサは共有メモリに対し、PREADの確率で読出し要求,PWRITEの確率で書き込み要求を出し、それ以外の場合は共有メモリをアクセスしない。

このシステムに関してHNCCで記述したシ

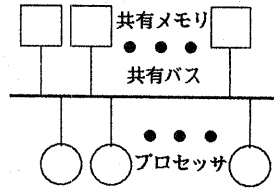


図6 並列計算機システムの例

```

nodeatom PU {
  inport datain; outport request;
  program pu(); }

nodeatom MEM {
  inport access; outport dataout;
  program p_memory(); }

nodeatom BUS {
  inport puin[SIZE],memoryin[MSIZE];
  outport puout[SIZE],memoryout[MSIZE];
  program p_bus(); }

nodeset{
  node{
    PU pu[i:SIZE](i);
    MEM memory[i:MSIZE](i);
    BUS bus; }
  connect{ /* connect nodeatoms */
    int i;
    for(i=0;i<SIZE;i++){
      con(pu[i].request, bus.puin[i]);
      con(bus.puout[i], pu[i].datain); }
    for(i=0;i<MSIZE;i++){
      con(memory[i].dataout, bus.memoryin[i]);
      con(bus.memoryout[i], memory[i].access); }
  }
} main();

program{
  pu(){
    double op; int m_num;
    while(currenttime()<ENDTIME) {
      op=rand()/MAXNUM;
      if(op<PREAD) {
        m_num=rand()%MSIZE;
        send(request,m_num); send(request,READ);
        receive(datain); }
      else if(op<PREAD+PWRITE){
        m_num=rand()%MSIZE;
        send(request,m_num);
        send(request,WRITE);
        receive(datain); }
      else wait(ACCESS_TIME); }
  }
}

memory(){
  int request;
  while(currenttime()<ENDTIME) {
    receive(access,&request);
    wait(ACCESS_TIME);
    if(request==READ) send(dataout); }
}

bus(){
  int i,m_num,request;
  while(currenttime()<ENDTIME) {
    entry {
      for(i=0;i<SIZE;i++) {
        (puin[i],&m_num):
          receive(puin[i],&request);
          wait(BUSCYCLE);
          send(memoryout[m_num],request);
          if(request==READ) {
            receive(memoryin[m_num]);
            wait(BUSCYCLE); send(puout[i]); }
          else send(puout[i]);
          break; } } }
}

```

図7 HNCCによるプログラム例

シミュレーションプログラムを図7に示す。なお、このプログラムはHNCCの考え方を示すためのものなので、必ずしも正確ではなく、統計情報収集なども含まれていない。

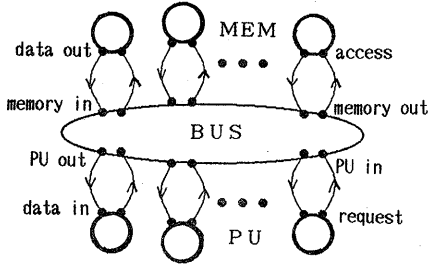


図8 例題のフローグラフ

HNCCはNCC同様、ノードの型定義、ノードの動作定義、ノード間の接続を記述する。この例では3つの基本ノード(nodeatom), PU, MEM, BUSを図8に示すように接続され、全体として複合ノード(nodeset) mainを構成している。離散系シミュレーションに対応するため、次のような関数利用されている。

- wait (時間);  
シミュレーション仮想時間を消費する。
- currenttime ();  
現在のシミュレーション仮想時刻を返す。

仮想時間の消費はwaitのみで行われ、その他の操作では時間は消費されない。また、先に述べたように受信ポートにはバッファが存在するためメッセージ間の時刻管理は次のように行われる。全てのメッセージには時刻印がおされているものとする。

• send側

相手方の受信ポートにメッセージが存在しなければメッセージをその場で送り、そのまま処理を実行する。

一方、受信ポートにデータが存在する場合は待ち状態になる。ここで、バッファが空いてメッセージが送られた時、もしも受信側の仮想時刻が送信側より進んでいた場合は、その時刻を自分自身の時刻とする。

• receive側

既に受信ポートにメッセージが存在する

場合もしそのメッセージの時刻が自分自身の時刻より進んでいる場合、その時刻を自分自身の時刻とする。受信ポートにメッセージが存在しない場合は送られてくるまで待つて同様の操作を行う。

これらの操作は2章で述べたアルゴリズムで自動的に行われるので、プログラマはこのような管理を気にする必要はない。また、多数のポートで同時にメッセージを待っている場合(BUSにおけるenrey文)、到着するメッセージは仮想時刻順になっていることが保証される。

HNCCにおいては複合ノード同士の組合せから更に上位階層の複合ノードを構成することが可能である。一例としては図7のmainに多少手を加え接続する事により、クラスタ構造を持つ並列計算機のシミュレーションを容易に記述することができる。

5. まとめ

EOSの性能上の問題点としては不当な処理の回復に伴うオーバーヘッドが大きくなりすぎる可能性が挙げられる。これに対しては処理プロセッサがトークンメッセージを送らなければならないようになったら先取り処理をやめ、GVTがVVTと等しくなるのを待つという方法、ユーザの指定や実行時の統計的に、不当である可能性が高い処理は行わないという方法などが考えられる。

また、履歴が大きくなりすぎる場合には先取り中の一部の時刻の履歴のみを残し、回復するときにはそれらの履歴のある時刻まで戻るようにすることも考えられる。

現在EOSのアーキテクチャの詳細設計を進めている。各プロセッサの期待できる速度を現在のデバイス技術に照らして求めた後シミュレーションによって効率を確認する予定である。

【参考文献】

[1] 吉田隆一、所真理雄：離散系シミュレーションの並列処理、コンピュータソフトウェア Vol. 4、No. 1 (1987)、pp23-33  
 [2] 高橋範朗、朴泰佑、天野英晴、相磯秀夫：並列計算機(SM)<sup>2</sup>-11のための並行プロセス言語NCC、32回情処全大、pp513-514