

## 超並列 Teraflops マシン TS/1

4B-1

～基本構想～\*

小柳 滋 田邊 昇 菅野 伸一 鈴木 真樹†

RWCP‡ 超並列東芝研究室§

## 1 はじめに

我々は通産省 RWC (Real World Computing) プロジェクトの一環として、超並列マシンの研究を行っている。RWCでは、柔らかな情報処理を実現するためのシステム基盤として超並列マシンを位置付けている。我々は光などの新しいデバイス技術や実装技術を積極的に取り込むことにより、柔らかな情報処理のための新たなパラダイムにも適用可能な超並列マシンの開発を目指している。

本研究テーマでは、前期(93年～96年)に電気による超並列マシンの開発を、後期(97年～2001年)に光インタコネクションを用いた超並列マシンの開発を計画している。本稿では、前期に開発する超並列マシン(TS/1)の設計思想と概要について述べる。

## 2 設計思想

超並列マシンの本質的課題は通信ボトルネックをいかに解消するかである。このためにはハードウェアの本質的な制約を考慮した上で応用のモデルを規定し、そのモデルを実現するための最適なアーキテクチャを構築するアプローチが有効である。

我々は、柔らかな情報処理のモデルとして、規則的な通信を内在するモジュールが複数個協調しながら動作するモデルを想定する。すなわち、低レベルでのデータ並列と高レベルでの制御(オブジェクト)並列とを融合したモデルである。将来超並列マシンが対象とする多数の応用分野において、このような階層化は自然な考え方であり、さらに、規則的通信をベースとしてモデル化することにより、超並列マシンの通信ボトルネックを解消することが可能となる。以上のような考え方にに基づき、TS/1のアーキテクチャの設計思想について述べる。

\*Massively Parallel Teraflops Machine "TS/1", - Overview-

†Shigeru OYANAGI, Noboru TANABE, Shin-ichi KANNO, Masaki SUZUKI

‡Real World Computing Partnership(新情報処理開発機構)

§(株)東芝 研究開発センター 内

1. 規則的細粒度処理の究極的性能向上  
従来型の超並列マシンでは演算性能を最高に発揮できる応用は通信が無視できるものに限られている。TS/1では上記のモデルにおける下位部分、すなわち充分な並列性と規則的な通信パターンをもつ問題においてハードウェアの性能がフルに発揮できることを設計目標とする。
2. マルチパラダイム支援  
多様な応用分野において適用可能な超並列マシンを目指すためには、アーキテクチャにおいても可能な限りマルチパラダイムを支援することが望まれる。上記のモデルにおける上位部分では、agent間のメッセージ通信や共有データのアクセスなどの不規則的粗粒度通信が必要となる。TS/1においてはマルチパラダイム支援として以下の機構を想定している。
  - (a) シストリック演算機構 (wave front array)  
ニューラルネットや画像処理、信号処理などの基本演算を最高速に実行する。
  - (b) 汎用プロセッサ+メッセージ通信機構  
オブジェクト指向や agent model などの高次のモデルの通信を高速化する。
  - (c) 分散共有メモリ機構  
共有データのアクセスを支援し、多様な並列プログラミングを許容する。
3. 大規模構成への拡張性  
超並列マシンのもつ大きな利点の一つは拡張性 (scalability) である。拡張性を実現するためには、結合網のトポロジーや実装方法に充分な検討が必要である。TS/1では3次元トラスを基本とするネットワークを採用し、3次元実装による通信バンド幅の確保により64K台までの拡張性をもたせている。
4. 高信頼性  
大規模な超並列マシンにおいてはPEの故障を前提としたフォールトトレランスは不可欠である

る。TS/1では故障したPEを迂回するルーティング機構や代替プロセッサの装備により、信頼性を高めている。

### 3 アーキテクチャの概要

TS/1のPEは、汎用プロセッサ(R4400PC)、高速メモリ(同期DRAM)、および付加プロセッサ(ベクトルプロセッサ、ルータ、制御機構等)から構成される。プロセッサ間結合網は3次元トラスを基本とし、これに加えて1軸方向にクロスバスイッチを経由して代替プロセッサと接続する。TS/1の目標性能諸元を表1に、PEのブロック図を図1に示す。

表 1: 目標性能諸元

最大PE数	65,536 (64 x 64 x 16)
汎用プロセッサ性能	125MIPS, 62.5MFLOPS
付加プロセッサ性能	125MFLOPS(倍精度) 250MFLOPS(単精度)
最大システム性能	20.5TFLOPS(単精度) 12.3TFLOPS(倍精度)
メモリ容量	64MB/PE
メモリ速度	800MB/s/PE
通信性能	250MB/s/port

TS/1アーキテクチャの特徴的な機構について簡単

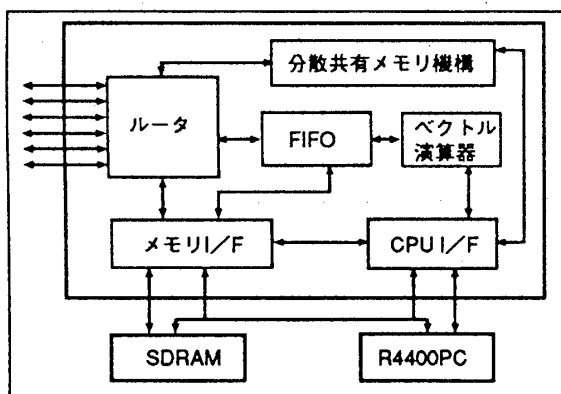


図 1: PE の構成

に説明する。

#### 1. プロセッサ間チェイニング

ルータとベクトルプロセッサとをFIFOにより接続することにより、プロセッサ間にまたがるベクトル演算(シストリック演算)を実質的に通信遅延なく実行することができる。

#### 2. フォールトトレラントルータ

ワームホールルーティングにおいて、デッドロックを回避しつつスループットの向上と故障回避ルーティングを実現している。

#### 3. 分散共有メモリ機構

各PEからTS/1の全メモリ空間をアクセスすることができ、また細粒度通信におけるオーバーヘッドを低減している。

これらについては、引き続き発表[1][2][3]で説明する。

### 4 おわりに

TS/1の設計思想とアーキテクチャの概要について述べた。現在付加プロセッサの機能設計を行っており、96年度に本格的な超並列マシンを構築する予定である。なお、ソフトウェアに関しては現在HPFコンパイラとデバッガ、性能モニタなどの並列プログラミング環境の開発を行っている。これらをベースとして、今後並列オブジェクト指向等の高位のシステムに取り組む予定である。

### 参考文献

- [1] 田邊, 菅野, 鈴木, 小柳: 超並列 Teraflops マシン TS/1 ~Wavefront Array のための結合網アーキテクチャ~, 第48回情報処理学会全国大会, 4B-02, MAR 1994
- [2] 菅野, 田邊, 鈴木, 小柳: 超並列 Teraflops マシン TS/1 ~フォールトトレラントルーティング~, 第48回情報処理学会全国大会, 4B-03, MAR 1994
- [3] 鈴木, 田邊, 菅野, 小柳: 超並列 Teraflops マシン TS/1 ~分散共有メモリアーキテクチャ~, 第48回情報処理学会全国大会, 4B-04, MAR 1994