

科学技術計算専用ロジック組み込み型プラットフォーム・ アーキテクチャの開発 —プロジェクト全体像—

村上和彰 (九州大学), 稲垣祐一郎 (富士総合研究所), 上原正光 (セイコーエプソン),
大谷康昭 (富士総合研究所), 小原 繁 (北海道教育大学), 小関史朗 (三重大学),
佐々木 徹 (アプリア・マイクロシステムズ), 棚橋隆彦 (慶應義塾大学),
中馬 寛 (徳島大学), 塚田 捷 (東京大学), 長嶋雲兵 (産業技術融合領域研究所),
中野達也 (国立医薬品食品衛生研究所)

E-mail: ehpc@star.fuji-ric.co.jp

A Platform Architecture for Embedded High-Performance Computing - Project Overview -

Kazuaki Murakami (Kyushu Univ.), Yuichiro Inagaki (Fuji Research Institute Corp.),
Msamitsu Uehara (Seiko Epson Corp.), Yasuaki Ohtani (Fuji Research Institute Corp.),
Shigeru Obara (Hokkaido Univ. of Education), Shiro Koseki (Mie Univ.),
Toru Sasaki (A Priori Microsystems Inc.), Takahiko Tanahashi (Keio Univ.),
Hiroshi Chuman (Tokushima Univ.), Masaru Tsukada (Univ. of Tokyo),
Umpei Nagashima (NAIR), Tatsuya Nakano (NIHS)

E-mail: ehpc@star.fuji-ric.co.jp

1. はじめに

筆者らは、科学技術庁 平成 12 年度科学技術振興調整費 (総合研究) により「科学技術計算専用ロジック組み込み型プラットフォーム・アーキテクチャの開発」プロジェクトを実施している (平成 16 年度までの予定)。本稿は当プロジェクトの全体像を紹介する。

2. プロジェクト実施の背景

計算化学, 計算物理学, 等に代表される「科学技術計算」, 「数値シミュレーション」は今日, 研究室や実験室のレベルを超えて, 産業界の多くの現場での需要が大きい情報科学技術の基盤技術である。科学技術会議 情報科学技術委員会の「情報科学技術先導プログラムの重点領域の設定について」(平成 11 年 7 月) の中でも次のように指摘している。

すなわち, 「国民生活に役立ち, また産業界にとって有益な実用的な問題について, 信頼性の高いシミュレーションを行うためには, 単に大規模であるだけでなく, ミクロとメゾとマクロ, 気体と液体, 流体と構造, 流体と熱・化学反応など, 異なる物理法則を連成させ, 複雑な系の統合的なシミュレーションを行うことが必要になる。また多くの問題では, 異なる時間スケール・空間スケールの現象が共存し, シミ

ュレーションが困難になるが, こうした困難を解決し, 統合的な解析を行うためには, 膨大な計算速度, 記憶容量を実現することが重要であると同時に, モデル化手法・アルゴリズム・計算手法・並列化手法などの開発が重要である。このような技術は, 物質科学, 生命科学, 医療, 環境, 安全, 製造業など多くの分野で必要とされ, 社会的な重要性は極めて高いものである。」

このように, 統合的な数値シミュレーションの技術開発に対する重要性の認識, ならびに, 当技術に対する産業界からの潜在的な要望が現として存在するにも関わらず, それを実現するソフトウェア, ハードウェアに関わるシステムの現状は必ずしも産業界の要望に応えきれていない。

たとえば, これまでの数値シミュレーションのプラットフォームはもっぱらベクトル計算機に代表されるスーパーコンピュータであったが, これではせいぜい 1 機関に 1 システム程度の導入に留まり, 設計開発現場で技術者がパーソナル・ユースに利用するのは実質的に不可能であった。結局のところ技術者が日々の設計開発業務で数値シミュレーションを実行できるプラットフォームはワークステーションやパーソナルコンピュータというのが実状であり, 如何にパーソナルコンピュータの性能が上がったとは言え, まだその計算能力の非力さ

故に数値シミュレーションの応用が一部にのみ限定されるという問題を抱えている。

3. プロジェクトの目的

本プロジェクトでは上記の状況に鑑み、超高速の数値シミュレーションが産業界の設計開発現場でツールとして簡便に利用できるという状況の実現を目指して、以下の特徴を備えた「専用ロジック組込み型数値シミュレータ構成のためのアーキテクチャ」を開発する。

- ・ 計算物理、計算化学、等の各計算科学分野が有する計算上の独自性を反映した専用/準専用システム
- ・ システム LSI 技術と並列分散処理技術を駆使
- ・ 従来のスーパーコンピュータに比べて飛躍的に優れたコスト・パフォーマンス
- ・ 研究現場、設計開発現場におけるパーソナル・ユースに供することが可能
- ・ 組込み機器システムのようにユーザ・フレンドリーなハードウェア/ソフトウェア一体型システム

具体的に本プロジェクトで対象とする科学技術計算分野は以下の通りである。

- (a) 非経験的分子軌道法等による巨大分子電子状態計算：非経験的分子軌道法，フラグメント分子軌道法，分子力場法，等の分子構造解析に用いるプログラム
- (b) 密度汎関数法等による物質・材料設計計算：実空間有限要素法，Car-Parrinello (CP) 法，Discrete Variational (DV) 法，等の実空間および波数空間で行う密度汎関数法プログラム

本プロジェクトで開発する「専用ロジック組込み型数値シミュレータ」は、具体的には以下の方法論により開発を遂行する。

- ① 計算科学分野における数値シミュレーションに対する共通項目および独自項目を明らかにし、分野毎の共通性に立脚すると同時に独自性をも反映可能とした「専用 LSI 組込み型並列分散処理プラットフォーム (土台、共通基盤)」のアーキテクチャを開発する。
- ② そして、各計算科学分野が有する計算上の特殊性に特化した「専用のロジック」、特に、メモリとプロセッサを混載した「専用 LSI」を上記プラットフォームにプラグインすることで、各分野に専用の数値シミュ

レータを構成する。

上記の「専用ロジック組込み型数値シミュレータ=専用 LSI 組込み型並列分散処理プラットフォーム+専用 LSI」というアーキテクチャ技術により、従来は汎用のスーパーコンピュータで実装していたため非常に高価についた高速かつ大規模な数値シミュレーションを設計開発の現場まで普及させることを可能にする。

4. プロジェクトの目標

本プロジェクトの成果として、以下の事項を達成目標とする。

- ① 業界の研究開発現場においてパーソナル・ユースに供することが可能な程度に優れたコスト・パフォーマンスを提供する。具体的には、既存の汎用コンピュータ・システムの1桁から2桁上のコスト・パフォーマンスを狙う。これにより、
 - ・ 科学技術計算活用機会増大、
 - ・ 科学技術計算ユーザの裾野拡大、
 - ・ 科学技術計算適用分野の拡大、そして、
 - ・ 産業界のあらゆる研究開発 (R&D) 現場における R&D の短 TAT (Turn Around Time) 化、低コスト化、国際競争力強化

に貢献することを目指す。

- ② 実際に専用数値シミュレータを開発するという実践を通して、以下の図式から成る「コスト・パフォーマンスに優れた特定用途向けマシンのアーキテクチャならびにシステム構成論 (メソドロジー)」の確立を目指す。

コスト・パフォーマンスに優れた特定用途向けマシン	
= クラスタ・コンピューティング技術	} 専用ロジック組込み型並列分散処理プラットフォーム技術
+ 並列分散処理技術	
+ コモディティ技術	
+ 組込みシステム技術	
+ システム LSI 技術	

これにより、「高性能科学技術計算分野における組込みシステム」の実現可能性を実証する。また、上記アーキテクチャおよびメソドロジーにより、

- ・ 組込み型科学技術計算プログラムの設計手法ならびに設計容易化技術

- ・ 数値シミュレーション・ユーザに対するユーザ・フレンドリーなシミュレーション環境の構築技術

等のソフトウェア・サイエンス技術を確認することを旨とする。さらには、専用マシンの適用可能分野が科学技術計算分野以外にも拡大することを狙う。

5. 計画の概要

本プロジェクトは以下の2つのサブプロジェクトから成る。

- ① プラットフォームおよび専用ロジックのアーキテクチャ開発
- ② 科学技術計算プログラムのプラットフォーム向き並列分散化および組込みソフトウェア化に関する研究

5.1 プラットフォームおよび専用ロジックのアーキテクチャ開発

「専用ロジック組込み型並列分散処理プラットフォーム」とは、「各種アプリケーションに応じた専用ハードウェア・ロジックを組込み（プラグイン）可能とし、そして当該専用ロジックが提供する特定用途向け計算機能以外の下記機能を共通基盤機能として提供する『専用システムの土台』」である。

- ・ クラスタ・コンピューティング機能
- ・ 並列分散処理機能
- ・ コモディティ機能（汎用プロセッサ、汎用高速ネットワーク、等）
- ・ 組込みシステム機能

その実装法に関しては種々可能であるが、本プロジェクトでは以下の実装法を用いる（図1参照）。

- 専用ロジック⇒専用 LSI
- プラットフォーム⇒汎用 MPU（マイクロプロセッサ）@プリント基板+業界標準バス/ネットワーク規格

将来的にはプリント基板1枚をシステム LSI 1個で実装することも視野に入れて、プラットフォーム・アーキテクチャの研究を行う。

本プロジェクトでは以下の通り、専用 LSI 組込み型並列分散処理プラットフォームのシステム開発、シームレスなプログラム開発環境の構築、ならびに、科学技術計算用途向け専用システム LSI の開発を行う。そして、これら

の方法論によって得られる開発コスト削減効果に関して定量的に評価する。

(a) プラットフォームの・システムの研究開発

本研究では、上記「プラットフォーム」のハードウェアおよびソフトウェアが備えるべき機能、性能を明らかにし、それらをアーキテクチャとして定義する。そして、第1版～第3版に分けて、以下の諸元を有する専用 LSI 組込み型並列分散処理プラットフォームを開発し、アーキテクチャとしての有効性を検証する。

① 第1版：

- ・ 組込みシステム業界標準の Compact PCI 規格に準拠した筐体およびプリント基板（CPCI カード）を採用。
- ・ 「計算ボード」として汎用高性能マイクロプロセッサを4個搭載した CPCI カードを開発。
- ・ 上記計算ボードのソフトウェア環境としては、組込みシステム業界標準の ITRON と C 言語を採用。
- ・ さらに、上記計算ボードをベースに第2版、第3版で FPGA および専用 LSI を搭載できるように、計算ボードは改造可能であること。
- ・ 上記筐体（7計算ボード/筐体）を複数、高速ネットワークで相互結合して全体としてクラスタを構成。

② 第2版：下記以外は、第1版と同じ諸元。

- ・ 上記計算ボードに対して FPGA（フィールド・プログラマブル・ゲートアレイ）を搭載できるように改造して、汎用高性能マイクロプロセッサ1個と複数個の FPGA を搭載したもの。

③ 第3版：下記以外は、第1版と同じ諸元。

- ・ 上記計算ボードに対して専用 LSI を搭載できるように改造して、汎用高性能マイクロプロセッサ1個と複数個の専用 LSI を搭載したもの。

(b) プログラム開発環境の構築

- 上述のように「計算ボード」として、
- ・ 汎用マイクロプロセッサだけのもの
 - ・ （汎用マイクロプロセッサに加えて）FPGA を搭載したもの

・ (汎用マイクロプロセッサに加えて) 専用 LSI を搭載したものの 3 種類を用意するのは、プログラム開発の容易さを考慮に入れた結果である。本研究においては、プログラム開発は以下の 4 ステップを踏んで行われる。

- ① WS/PC クラスタ上でのオリジナルプログラムの並列分散化。
- ② 第 1 版プラットフォームにおいて「計算ボード (汎用プロセッサ版)」上への上記プログラムの実装 (組込みソフトウェア化)。
- ③ 第 2 版プラットフォームの「計算ボード (FPGA 搭載版)」を用いて、上記プログラムの一部のロジックをハードウェアとして FPGA 上に実装。残りのソフトウェア・ロジックは汎用プロセッサ上で実行。
- ④ 第 3 版プラットフォームの「計算ボード (専用 LSI 搭載版)」では、上記③の FPGA を専用 LSI に置き換え。ソフトウェア・ロジックはそのまま汎用プロセッサ上で実行。

上記 4 ステップを通して共通の並列プログラミング・ライブラリを提供するようプログラム開発環境を構築する。これにより、ステップを跨いでシームレスなプログラム開発を可能とする。

(c) 科学技術計算用途向け専用システム LSI の研究開発

本研究では、PPRAM (Parallel Processing RAM) 等のシステム LSI 関連技術に基づき、

- ・ メモリ/ロジック混載システム LSI 技術
- ・ (LSI 内および LSI 間) 並列分散処理技術
- ・ (LSI 内および LSI 間) 標準高速通信インタフェース技術

を駆使して、以下の 3 種類の科学技術計算用途向け専用 LSI アーキテクチャを開発する。

- (a) 分子軌道法: 二電子積分エンジン
- (b) 密度汎関数法: 3次元 FFT エンジン, 有限要素法エンジン

そして、上記のうち二電子積分エンジンを実際に LSI として開発製造する。使用するプロセス技術は $0.18\mu\text{m}$ または $0.14\sim 0.13\mu\text{m}$ を予定。

5.2 科学技術計算プログラムのプラットフォーム向き並列分散化および組込みソ

フトウェア化に関する研究

以下の科学技術計算プログラムのコア計算部分を専用 LSI 組込み型並列分散処理プラットフォーム上にオフロードすると同時に並列分散化する。

- ・ 分子軌道法プログラム関連: 非経験的分子軌道法, フラグメント分子軌道法, 分子力場法, 等の分子構造解析に用いるプログラム
- ・ 密度汎関数法プログラム関連: 実空間有限要素法, Car Parrinello (CP) 法, DVX α 法, 等のプログラム

具体的には、以下の研究を実施する。

- ① 各計算プログラムに関して、そのコア計算部分の認識、および、括り出し。既存のプラットフォーム (スーパーコンピュータ, ワークステーション, パーソナルコンピュータ, WS/PC クラスタ, 等) を用いて定量的にコア計算部分を求める。
- ② 専用 LSI 化される「コア計算部分」を括り出し、第 1 版プラットフォーム上にオフロードすることで、各計算プログラムを一種の「組み込みシステム」化する。これにより、第 2 版~第 3 版プラットフォームにおける FPGA ならびに専用 LSI への組込みを容易にする。
- ③ オフロードされたコア計算部分を並列化および分散化する。この時、プロセッサ性能, 通信性能, メモリ容量の間のトレードオフを考慮に入れた並列化, 分散化を図る。
- ④ 専用 LSI 組込み型並列分散処理プラットフォーム (第 1 版) を用いて各計算プログラムを実行することで、プラットフォームの仕様, プログラムの並列分散化, 組込みシステム化, 等に関する評価およびフィードバックを行う。
- ⑤ また、専用 LSI 組込み型並列分散処理プラットフォーム (第 2 版~第 3 版) を用いて実際に大規模科学技術計算シミュレーションを実行し、既存のプラットフォームでは現実的な時間内で解けなかったような問題を現実的な時間内で解くことで、科学技術計算シミュレーション・ユーザの利用に供する。

(a) 分子軌道法プログラムに関する研究

製薬会社における製薬設計や化学会社にお

ける材料設計などで幅広く利用されている以下の3つの計算プログラムを対象とする。

- ・ **非経験的分子軌道法**: 非経験的分子軌道法は、分子の電子密度分布を求めるとともに、その時どのようなエネルギーを持っているかを求め、これにより分子個々の物理的・化学的性質と分子間相互作用という物質科学の基礎的知見を原子レベル、電子レベルで提供する。とくに、本来は電子分布が重要であるにも関わらず取り扱えるサイズが小さいがために、計算科学が重要視されなかった分野に対して、大きな影響力がある。例えば、薬理学の分野では、膜を通過する際の分子状態を電子分布を考慮して解析することができ、薬の分解・吸収・移動・活性プロセスをより効率よく、安全にする薬の設計に役立てることができる。さらに、遺伝分野についても、100 基質 + 水 10 万個程度を取り扱うことができるため、DNA から t-RNA を通じた情報読み込みとタンパク質合成反応の計算ができ、これにより、ゲノム情報をもちいたシミュレーションで遺伝病などの解明と治療法の設計を大きく進めることが可能となる。また、分子軌道法は、高いエネルギー精度が必要な場合や薬物-受容体相互作用において電子密度分布が本質的な役割を果たす場合の研究を進める上でも大きな役割を果たすものである。
- ・ **フラグメント分子軌道法**: フラグメント分子軌道法は、非経験的分子軌道法に比較してタンパク質のような巨大分子の電子状態を高速に計算することを可能にする手法である。ポストゲノムの大きなテーマの一つとして、ゲノム情報から、タンパク質機能発現を如何に引き出すかがある。このためには、その立体構造を精度良く求めることがこのための一つの焦点であり、これを高速し、研究者がパーソナルに使用できるようになれば、ポストゲノム研究を進展させる大きな影響がある。
- ・ **分子力場法**: 薬物設計の現場で今最も求められているのは、数千~十数万化合物の化合物群(ライブラリー化合物)を対象とした3次元立体構造と物理化学的性質(疎水性、水素結合供与性・受容性、正荷電・負荷電)を考慮した分子設計の方法論である。このような方法が可能となれば、例えば受容体タンパク質に強く結合する化合物群をコンビナトリアル法で合成でき、ハイスループット・スクリーニング (HTS) 法で優れた化合物を検索できるからである。これまでコンビナトリアルライブラリーのデザインは、化合物の2次元

構造に基づいた方法が一般的であったが、2次元構造から3次元構造への変換を高速に、精度よく、かつ準安定な配座まで求めることができれば、薬品開発のスピードが飛躍的に向上するものと考えられる。そしてこの手法が整備されれば、医薬品開発分野(酵素の阻害剤、受容体拮抗剤)のみにとどまらず、例えば、農薬、内分泌攪乱物質、毒性予測、触媒設計、ゲノム配列解析など、化学物質を扱う広範囲な応用分野にも新しい手法となりうることも既に認識されている。

上記の計算プログラムのうち非経験的分子軌道法、フラグメント分子軌道法はともに、そのコア計算部分は計算時間の90%以上を占めている「フォック行列生成」にある。そのフォック行列生成部を小原 繁(北海道教育大学)が考案した高速アルゴリズム、並列分散化、ならびに、専用 LSI を用いることで高速化する方針である。さらに、100 残基を超えるタンパク質について分子構造を固定した状態での電子状態が計算可能なプログラムの開発を行う。また、これに解析的微分ルーチンを組み込み、分子構造の最適化が行えるようにする。現在、フラグメント分子軌道法プログラムでも、(STO-3G 基底関数を用いた場合) 50 残基程度のタンパク質を PC クラスタを用いて約 10 時間で計算できるところまで進んでいるが、100 残基を超えるタンパク質のような巨大分子の電子状態を計算できる方法は存在しない。本研究により、この限界を打破することを目指す。

(b) 密度汎関数法プログラムに関する研究

物質の構造、生成法や、電子的、光学的、力学的、磁氣的、熱的、化学的その他あらゆる性質を統一的に記述するのために、密度汎関数計算、とりわけ第一原理分子動力学法がもっとも系統的、包括的な理論計算法として認知されつつある。これは実用的かつ信頼度が最も高い理論計算であるが、現在は利用し得る計算機資源の限界に阻まれて、原子数が 100 前後までが実用的な限界となっている。本研究ではこれを乗り越え、その 10 倍あるいは 100 倍近い規模の計算を可能とする計算プログラムの開発を実施する。具体的には、以下の3つの解法に対して、並列化のための分析、および、並列化に適した計算アルゴリズムの開発を行う。

- ・ **実空間有限要素法**: 標準的な密度汎関数法計

算では、擬ポテンシャルと平面波展開をもちいるが、このような基底展開は大規模系では常にメリットがあるとは限らない。一方、実空間有限要素法を用い曲線座標系や並列化などの方法と組み合わせると、極めて高精度を保ちしかも大規模な第一原理計算に最適な計算方法となる。水・アルコールなどの分子性液体や、シリコン表面の分子の吸着系などで、高精度で高い並列化効率が得られることが実証されている。さらに電子系の時間発展を記述するための、強力な計算アルゴリズムが開発されている。波数空間と実空間との変換が必要でないため、並列処理化が極めて自然で有効であると考えられる。この計算手法で原子数が数 1000 の系を計算することが可能となれば、物質科学および材料科学へのインパクトは非常に大きいものがある。様々な実験的な知見を必要としない、無限の可能性のある新物質の予言ができるからである。有限要素法は、構造解析、流体解析などで一般的に用いられている手法であり、これを高速化し、専用計算化することの影響は計り知れない。本プロジェクトにおいては、GSMAC 法など流体解析分野における計算プログラムも検討対象とし、組み込み型並列分散処理プラットフォームの汎用化を検証するとともに、専用 LSI 向きのアルゴリズムを抽出し高速化を図る。有限要素法では、離散化ナブラ演算を行う部分も、ボアソンソルバーの部分が比較的高コストであり、この部分の専用 LSI 化による高速化が期待される。

- **Car-Parrinello (CP) 法**：これは平面波展開を用いる標準的な方法であるが、計算理論としては十分に成熟しており、この枠組みの中で大規模計算に最適な計算アルゴリズムとその処理システムを実現することは極めて意義が大きい。世界的にもこの標準的な計算理論を用いている研究者、技術者の数は非常に多く、世界的な波及効果が見込まれるからである。コア計算部分としては、高速フーリエ変換、波動関数の規格・直交化計算部であり、これを並列分散化、専用 LSI 化することにより計算時間の短縮を図ることができる。本研究によって、分子、クラスター、固体結晶、表面など、様々な物質科学研究者および材料開発に携わる技術者が、身近な実験装置のように計算を実行できるようになれば、インパクトは非常に大きい。原子数が 1000 程度の系が手軽にできるようになれば、上記のことは実現するであろう。

- **Discrete Variational (DV) 法**：DV 法は密

度汎関数法の簡便化された計算法である。構造や反応など全エネルギーの計算を必要としない場合、電子状態と関連する性質のみの記述に有効である。この計算方法は並列処理化に特に適した計算法であり、極めて高速で通常の CP 法に比べて遥かに原子サイズの大きい系に適用可能である。これによって、原子数のかなり大きいナノ構造系の電子的、量子的諸性質の予言が可能となれば、任意の固体材料の電氣的磁氣的性質の予測、走査トンネル顕微鏡像のその場シミュレーションなど、実用的に重要な計算が手軽にできるようになり、材料開発へのインパクトは大きい。この計算法では原子数 1 万程度の系の計算を、実行できることが望ましい。

6. おわりに

以上述べた本プロジェクトにより達成される情報科学技術分野の技術的ブレークスルーを整理すると以下のようになる。

- 科学技術計算分野として産業界の需要が大きく緊急度の高いものを厳選し、それらに対する実用に耐え得る並列分散処理プログラムを開発し、一般に提供する。特に、「情報科学技術先導プログラムの重点領域の設定について」の中の「統合シミュレーション技術」の項で挙げられているマルチスケール・モデリングおよびマルチフィジックスという技術課題に対して 1 つの答を提示する。
- 今日のコモディティ技術であるクラスタ・コンピューティング、並列分散処理技術、超高速ネットワーク技術と、新規技術であるシステム LSI 技術とを融合させることで、コスト・パフォーマンスに優れた特定用途向け超高速並列分散システムの構成法、すなわち、アーキテクチャ技術を開発し提供する。これは、「情報科学技術先導プログラムの重点領域の設定について」の中で重点技術として取り上げられている「並列分散ソフトウェア技術」および「アーキテクチャ技術」に対する本プロジェクトからの回答である。