

Euro-Par 2011 参加報告

東京大学情報基盤センター
片桐 孝洋、實本 英之

1. はじめに

本報告は、フランスのボルドーで 2011 年 8 月 29 日から 9 月 2 日に開催された国際会議 Euro-Par2011 の参加報告である。本稿の 2 節と 4 節は片桐により執筆され、3 節は實本により執筆された。全体の取りまとめは、片桐が執り行った。

はじめに、国際会議 Euro-Par とは何かを説明したい。Euro-Par は、並列と分散コンピューティングの科学技術を取り扱う国際会議である。1995 年から毎年、欧州で開催されている。扱う分野は、ソフトウェアとハードウェアに対するアルゴリズムや理論の問題であり、科学と技術の双方の対象を包含している。近年重要性が増してきた、モバイルやクラウドを用いた科学技術への適用についても考慮している。

国際会議 Euro-Par2011 の開催前には、Euro-Par に関連する分野のワークショップが開催されている。著者らは Euro-Par2011 に加え、この関連ワークショップのいくつかにも参加している。そこで本稿では、このワークショップの内容についても紹介することとした。

2. 生産性と性能ワークショップ (Productivity and Performance (PROPER2011))

8 月 30 日に開催された Euro-Par 併設のワークショップ PROPER2011 に参加した。今回は第 4 回目の開催であった。3 つのセッション、合計で 8 件の発表があった。そのうち、Programming Interfaces セッションで発表されていた「Scout: A Source-to-Source Transformator for SIMD-Optimizations」Olaf Krzikalla and Kim Feldhoff (ドイツ ドレスデン大学) の発表が興味深かった。

この発表は、著者らが携わっている Hi-CFD Project において開発している、SIMD 命令に対応する自動並列化の研究発表であった。ソースコードに注釈 (プラグマ) を挿入することで行う自動ベクトル化に加えて、ループアンローリングなどの最適化を実現することが出来る。これらの「プラグマ」は、従来からコンパイラが提供している機能であり新規性がない。しかしここでの特徴は、<小さなループ>に対する最適化機能の強化にある。この背景として、近年、マルチメディア命令を搭載している計算機が普及していることがある。かつそれらの計算機アーキテクチャは、従来のベクトル計算機のような単純構成ではない。階層化されたキャッシュメモリ、非均質なメモリアクセスなど、複雑化された計算機アーキテクチャの普及がある。その結果、アプリケーションの特性に応じ、きめ細かなチューニングをしないと高性能化が達成できなくなった。このような背景から、小さなループでの最適化強化が特に重要となってい

る。

そのほかの発表では、「Auto-tuning for Energy Usage in Scientific Applications」Ananta Tiwari, Michael A. Laurenzano, Laura Carrington and Allan Snavely (米国サンディエゴ・スパコンセンター) のように、電力最適化のための自動チューニングの話題が面白かった。今後 2020 年頃にかけて開発されると見込まれているエクサスケールのスーパーコンピューターでは、「電力の壁」が問題視されている。これらは計算機アーキテクチャだけの問題ではなく、科学技術計算アプリケーションにおいても電力最適化を行うことは重要と考えられ、次世代のスーパーコンピューティング環境で決定的な要因となると予想される。

3. Resilience ワークショップ

3.1 概要

Euro-Par には同時開催される幾つかのワークショップがあり、Grid/Cloud コンピューティングや GPU/MIC 等のヘテロアーキテクチャ、さらに仮想化、生産性、耐故障性などのエクサスケールに向けたホットトピックについての萌芽研究が発表された。

この中で、近年エクサスケールコンピューティングの耐故障を実現する手段として注目されている Resiliency についてのワークショップ "Resiliency in High Performance Computing (Resilience) in Clusters, Clouds, and Grids" に参加した。

ワークショップは 12 件の発表があり、40~50 人程度が収容可能な講義室で行われたが、常時 30 人ほどの参加者がいた。また各発表につき、2, 3 件の質問があり、活発な論議が交わされていた。近年注目される GPU/MIC や Cloud コンピューティングに比して劣らぬ関心を集めており、エクサスケールコンピューティングにおいて、どのように耐故障を実現すれば良いのかは今後も重要な課題と成っていくと考えられる。

3.2 研究発表の様子

発表は故障解析、難検知故障 (ソフトウェア故障)、耐故障評価環境、チェックポイント/リスタートについてグループ化されていた。いずれも、エクサスケールコンピューティングにおいて特に重要となってくる技術であり、エクサスケールへのロードマップを定めた、IESP においても言及されている分野である。

いずれの発表も、従来手法を適切に組み合わせ発展させる事によるもので、非常に実現性の高い研究であった。近年はアプリケーションとシステムが協調することにより、コストの低い耐故障技術を実現し、エクサスケール環境に対応するという路線が一般的で、Resiliency の研究というと、アプリケーションによる故障対処を指すことが多い。しかし、本ワークショップは、ほとんどがシステムレベルでの耐故障性研究であった。アプリケーションレベルの故障対

処は、アプリケーション開発者に耐故障技術の実装を行わせるため非常に実装コストが高くなると考えられており、本ワークショップのように、システムレベルの耐故障を限界まで突き詰めることも重要であるとする。

このような内容から、質問の内容は、システム開発者の立場で有ることが多く、利用法などのアプリケーション開発者の目線にたったものが少なかったように感じられた。しかし今後、アプリケーション開発者に耐故障実装の一部を課すことは十分考えられる。システム開発者の立場から、より一層、アプリケーション開発者が耐故障技術に関心をもつよう、アピールしていく必要性も感じた。各グループの概要を以下に紹介する。

3.2.1 故障解析

エクサスケールのシステムは、ファイルシステム等の複数のサブシステムが複雑に結びついた形態となると考えられている。このため、故障に複雑な過程があり、その原因を特定することが難しい。しかし、故障要因を特定することにより、復旧範囲の限定や故障予防を行うことが可能であるため、システムやアプリケーションの挙動からどのような故障が起こったのかを解析することや、故障・復旧のモデル化を行うことが重要である。本トピックでは、現存するシステムを元に、エクサスケールシステムで起きると考えられる故障の性質についての考察や、膨大なシステムログからの故障要因特定を容易にするツール等について発表・公演が行われた。

3.2.2 難検知故障

システムの大規模化に伴い、メモリやネットワーク上でのデータ化けが起きる確率は上昇していくと考えられている。このような故障はアプリケーションが一見正常に終了し、実際は間違った答えを出力するというような振る舞いをするため、システムレベルでは検知が難しい。本トピックでは、メモリ更新を記録しておくことで予期せぬデータ化けを検出するライブラリやレプリケーションによるネットワーク上でのデータ化け検出法等が発表された。

3.2.3 耐故障評価環境

耐故障技術の有効性を検証するためには、故障発生器を始めとした評価環境や、耐故障性能を評価する尺度の検討が必要である。本トピックでは、階層構造メモリを利用したときのアプリケーションの頑健性の評価手法や、VMを用いたデータ化け故障のシミュレータ等について発表が行われた。

3.2.4 チェックポイント／リスタート

チェックポイント／リスタートはスタンダードな耐故障手法として広く使われているが、エ

クサスケールコンピューティングでは規模の増大によりチェックポイントコストが増加し、破綻すると予想されている。本トピックでは、その予想が本当に起こりうるのかの検証や、規模増大に対してコストが増加しづらいような効率の良いチェックポイント取得手法について発表が行われた。

4. Euro-Par2011 セッションでの発表

Euro-Par2011 の発表は、8月31日から9月2日まで行われた。基調講演3件と、26件のセッションからなり、最大で4セッションが並列で実施される。したがって、すべてを聴講することはできない。片桐が参加してみた限りにおける全体の印象としては、突飛で革新的な技術発表ではなく、既存の手堅い研究をきっちりおこなっている発表が多い印象をうけた。

たとえば、Session E3 - Multicore and Manycore Programming で発表された「Fully Empirical Autotuned Dense QR Factorization For Multicore Architectures」Emmanuel Agullo (INRIA / LaBRI, France); Jack Dongarra (University of Tennessee, Knoxville, USA); Rajib Nath (University of California, San Diego, USA); Stanimire Tomov (University of Tennessee, USA)では、数値計算ライブラリ LAPACK における QR 分解ルーチンにおいて、近年問題となっているスレッド並列化時の性能劣化問題を改善する方法の提案をしていた。LAPACK では、BLAS と呼ばれる基本線形代数ルーチンをスレッド並列化している。しかしライブラリ設計の制約上、十分なスレッド並列性が抽出できないことが問題となっている。そこで、LAPACK レベルの数値計算アルゴリズム上の並列性を、専用のジョブスケジューラで抽出し、従来では 20GFLOPS 程度しか性能が出ない問題においても 100GFLOPS 程度にまで高性能化できる方法が提案されている。これらは PLASMA プロジェクトで開発されている既存技術であるが、性能向上の要となる性能パラメータは「タイルサイズ」と呼ばれる、キャッシュ容量を勘案したライブラリ上の性能パラメータである。このタイルサイズを、現在のマルチコア CPU 上でうまく調整する経験的な手法を提案していた。

一方、Session Distinguished Papers - Parallel and Distributed Data Management - Parallel and Distributed Programming - Parallel Numerical Algorithms で発表された、「Communication-optimal parallel 2.5D matrix multiplication and LU factorization algorithms」Edgar Solomonik (University of California at Berkeley, USA); James Demmel (University of California at Berkeley, USA)では、連立一次方程式の求解アルゴリズムである LU 分解において、行列のコピー数を増やして通信回数を減らす通信削減方式を提案していた。一般に、2次元分散に比べ3次元分散にすると、データが重複しすぎて計算量が増える欠点がある。しかし、2次元と3次元の間である「2.5次元分散」にすると、データ重複が増えるが、通信量がルートオーダーで減る。結果的に超並列の実行環境で有利となる。今後來たる 100

万～1000 万並列実行環境を見通した並列アルゴリズムといえる。

以上はごく限られた研究発表の紹介であった。しかし、Euro-Par のメインセッションにおいても、エクサフロップス・コンピューティングに向けた超並列化、厳しい電力制約、高生産性を実現する並列プログラミング手法、GPU コンピューティングなどのヘテロジニアス環境への適応が、現在のホットトピックといえ、それらの発表が多かった。

5. おわりに

来年開催される Euro-Par2012 は、ギリシャのロードス島で 2012 年 8 月 27 日～31 日に開催される予定である。なお Euro-Par2011 の詳細は、以下のホームページで公開されている。興味ある読者は一読されたい。

<http://europar2011.bordeaux.inria.fr/>

以上