

ベクトル演算に基づく高速データ解析ライブラリの開発と ALMA データ解析への応用 Development of High-Performance Data Analysis Library Based on Vector Operation and Its Application to ALMA

中里 剛^{1*}; 杉本 香菜子¹; 川崎 渉¹; 川上 申之介¹; 國吉 雅也¹; 中村 光志¹; 小杉 城治¹; 前川 淳¹
NAKAZATO, Takeshi^{1*}; SUGIMOTO, Kanako¹; KAWASAKI, Wataru¹; KAWAKAMI, Shinnosuke¹;
KUNIYOSHI, Masaya¹; NAKAMURA, Kohji¹; KOSUGI, George¹; MAEKAWA, Jun¹

¹ 国立天文台

¹National Astronomical Observatory of Japan

Atacama Large Millimeter/submillimeter Array (ALMA) は、東アジア、北米、ヨーロッパの国際協力のもとで運用されている大型電波干渉計である。世界最高レベルの感度と分解能を活かし、ALMA では太陽系天体から初期宇宙にいたるまで幅広い分野で画期的な成果が得られている。ALMA の観測データは定常運用時には年間 200 テラバイト程度となることが想定されており、このような大量のデータを効率的に解析するためには、解析用サーバーの性能向上やクラスターシステムの構築等ハードウェア面の処理能力強化とともに、計算機資源の性能を最大限活用するような解析アプリケーション側の工夫が必須である。中でも、並列計算は処理を大幅に高速化する可能性があるため、大規模データの処理においては並列計算を有効活用することが重要になる。特に、並列計算の一形態であるベクトル演算は、CPU の組み込み機能なのでノート PC から大規模なクラスターシステムにわたる様々な環境で普遍的に利用できること、コーディングに気をつければコンパイラの機能を使って手軽に利用できること、の 2 点からきわめて有望な高速化の手法である。このような背景から、我々はベクトル演算を軸として ALMA の標準データ解析アプリケーション Common Astronomy Software Applications (CASA) の高速化に取り組んでいる。高速化に際しては、CASA とは独立した比較的汎用性の高いデータ解析ライブラリを開発し、CASA からそれらの機能を利用することにより処理の高速化を図ることとした。そして、そのためのデータ解析ライブラリを Sakura と名付けて開発を進めている。Sakura の特長は、ベクトル演算の徹底活用による処理の高速化である。Sakura を利用するアプリケーションは、最適化されたベクトル演算処理を手軽に利用することができる。ベクトル演算の活用の際に問題になるのは、最適化と汎用性のバランスである。ベクトル演算命令の拡張命令セットはいくつかの世代を持ち、一般に新しい命令セットほど高速な処理になる。一方で、新しい命令を使うことにより、古い世代の CPU ではライブラリが動作しなくなってしまうという弊害がある。Sakura では、最適化と汎用性の両立のため、様々な世代の拡張命令セットに最適化されたコンパイル方法をサポートしている。これにより、実行環境に応じて最適化された機能を提供したり、汎用性と最適化の両立をアプリケーションレベルで実現したりすることを可能にしている。CASA への組み込みに先立ち、我々は Sakura を基盤として CASA 上で動作する単一電波望遠鏡の観測データを解析するプログラムのプロトタイプを作成し、その性能評価を行った。Sakura の活用に加えて処理フローの見直しや不要なデータ入出力の低減等の高速化を行った結果、プロトタイプのパフォーマンスは既存の機能に対して最大で 20 倍という結果を得た。この結果をもとに CASA 開発グループで検討を行い、2015 年 4 月リリース予定の CASA 4.4 で単一電波望遠鏡データの解析機能に対する Sakura の導入が承認された。本稿ではプロトタイプの実装手法や性能測定結果の概要、および CASA 4.4 に向けた Sakura の導入作業の進捗と初期の性能評価結果について報告する。今後 CASA の開発において Sakura をさらに活用していくためには、干渉計のデータ解析をサポートすることが不可欠である。そこで、干渉計データの解析に対する Sakura の導入計画についても述べる。

キーワード: ALMA, データ解析, ベクトル演算

Keywords: ALMA, Data Analysis, Vector Operation