

NICTサイエンスクラウドを用いたかぐや衛星 WFC-Lからのバイポーラ型波形抽出処理の高速化 High performance data processing for detection of bipolar waveforms from KAGUYA/WFC-L using the NICT Science Cloud

矢木 大介^{1*}; 村田 健史²; 笠原 禎也¹; 後藤 由貴¹
DAISUKE, Yagi^{1*}; MURATA, Ken T.²; KASAHARA, Yoshiya¹; GOTO, Yoshitaka¹

¹ 金沢大学, ² 情報通信研究機構

¹Kanazawa Univ, ²National Institute of Information and Communications Technology

月探査衛星かぐやは、2007年9月に打ち上げられ、2009年6月に月面に制御落下した。かぐや衛星には月周辺のプラズマ波動を観測する波形捕捉器 WFC が搭載されており、特に 100Hz~100kHz の電界波形を観測する WFC-L では、これまでの研究からいくつかのパターンに分類できる特徴的なバイポーラ型の波形が多数確認されている。しかし、WFC-L は 250kHz のサンプリング周波数で波形データを取得することから、そのデータ総量は約 190GB に達する。我々は現在、この波形データからバイポーラ型波形を自動抽出するアルゴリズムを開発中であるが、汎用の PC ワークステーションでは、全観測データから波形抽出を行うのに 1 週間近い処理時間を要する。これでは、より精度よく波形を抽出し、その分類を行うための処理アルゴリズムを試行錯誤するには、大変非効率的である。そこで、情報通信研究機構 (NICT) のサイエンスクラウドを用いて、処理の高速化を図った。NICT サイエンスクラウドは、科学研究目的のために構築されたクラウドシステムで、特にビッグデータサイエンスを主対象の一つにしている。今回は、NICT サイエンスクラウド上でワークフローシステムを用いた並列分散処理による高速化を行い、その効率について評価した結果を報告する。

実際の並列処理は、Pwrake (Parallel Workflow extension for Rake) と呼ばれるツールを用いて、サーバ群にワークフローを与えることで実現した。Pwrake は、Ruby 言語で記述されるビルドツールである Rake をファイル共有システム向けに拡張したものである。Pwrake 上に処理内容と使用するノード及びコア数を記述することで、コマンドをタスクとして各ノードに割り振り並列処理を行うことが可能である。

結果として、10 ノード 24 コアの計算リソースを用いた場合、1 ノード 1 コアでの処理に比べて約 140 分の 1 の時間で処理を終えることが確認できた。処理速度が使用リソース数に比例していないのはハイパースレッドの影響によると思われる。同システムを活用することで、より精度の高い波形抽出アルゴリズムの開発が効率よく行えることが期待できる。今後は、波形抽出アルゴリズムの改善に加えて、更に計算リソースを増やした場合の測定実験についても実施する予定である。

キーワード: 月探査衛星かぐや, 波形捕捉器, NICT サイエンスクラウド, 並列処理

Keywords: Lunar Orbiter KAGUYA, Waveform Capture, NICT Science Cloud, parallel processing