地球観測衛星データ処理におけるJAXAスパコン活用の効果検証

齋藤 紀男*¹, 上田 陽子*¹, 中西 功*¹, 仁尾 友美*¹, 小西 利幸*¹, 斎藤 進*¹, 田中 誠*², 井口 茂*², 井上 淳一*³, 鳥居 雅也*³

Verification of Extreme Time Reduction for Earth Observation Satellite Data Re-processing with JAXA Supercomputer

Norio Saito^{*1}, Yoko Ueda^{*1}, Isao Nakanishi^{*1}, Tomomi Nio^{*1}, Toshiyuki Konishi^{*1}, Susumu Saito^{*1}, Makoto Tanaka^{*2}, Shigeru Iguchi^{*2}, Junichi Inoue^{*3}, Masaya Torii^{*3}

Key Words: Earth Observation, Data Processing, Supercomputer

Abstract

Observation by satellites orbiting around the Earth is an important tool for scientific study of global changes. Observation data are reprocessed almost every year according to progress in computation model to obtain physical information from observation. The amount of observation data grows day by day. This yields longer and longer time to reprocess the data. However, data latency is an important subject because quick access to earth observation data considerably encourages research into global environment changes and their effects to our society. On this point, JAXA Supercomputer System Generation 2 (JSS2) helps us to solve the problem since it can enhance throughput of data processing by using a number of computing nodes. First, we have utilized JSS2 for reprocessing of IBUKI (GOSAT: Greenhouse gases Observing Satellite) observation data. IBUKI was launched on Jan. 23rd, 2009, and the data during 6.5 year period has been reprocessed in Nov. 2015. We obtained incredibly high throughput; 30 nodes (360 cores) of Pre-Post system of JSS2 were assigned and resulted in 33 times faster latency. Next, we have used JSS2 for GPM (Global Precipitation Measurement) Core Observatory that was launched on Feb. 28th, 2014. The data during 2.2 year period has been processed and 4 times faster latency was achieved in June 2016 by using 3 nodes at a maximum. As a conclusion, it is shown that JSS2 is a powerful tool to improve data latency of Earth observation data.

概要

人工衛星による地球観測は、地球環境の変化を科学的に研究するための重要なツールである. 観測データは、1年に1回程度、物理量を算出するための計算モデルが改訂されることに伴って再 処理される.観測データの量は日に日に増加するため,再処理に要する時間は増加し続ける.し たがって、観測データをユーザへ提供するまでのレイテンシが重要な課題となっている.なぜな らば、地球観測データをできるだけ早く入手できることは、地球環境変動の研究の進展に大きく 寄与できるからである. この点で、JAXAが有するスーパーコンピュータシステム (JSS2: JAXA Supercomputer System Generation 2) を活用することが有効である. なぜなら, JSS2の計算ノードを 大量に使用することで,データ処理のスループットを大きく改善できるからである.初めに,温 室効果ガスの全球観測を行っている衛星「いぶき」(GOSAT: Greenhouse gases Observation Satellite) におけるJSS2利用の成果を報告する.2009年1月23日に打ち上げられた「いぶき」の6年半に渡る 観測データを2015年11月に再処理した. JSS2プレポストシステムの30ノード(360コア)を使用し, 従来の手法で再処理した場合に比して33倍のスループットが達成できた. 次に, GPM主衛星 (GPM: Global Precipitation Measurement, 2014年2月28日打上げ) でのJSS2利用について述べる. 2016年6 月に2年3ヶ月の観測データのレベル1処理を最大3ノードを使用して実施し、従来の4倍のスループ ットを達成した. JSS2は地球観測データの提供時間を大幅に短縮するために極めて効果的なツー ルとなることが示された.

doi: 10.20637/JAXA-RR-16-007/0009

^{*} 平成 28 年 11 月 24 日受付(Received November 24, 2016)

^{*&}lt;sup>1</sup> 宇宙航空研究開発機構 第一宇宙技術部門 衛星利用運用センター (Satellite Applications and Operations Center, Space Technology Directorate I, Japan Aerospace Exploration Agency)

^{*2} 日本電気株式会社 (NEC Corporation)

^{*3} 富士通株式会社(FUJITSU Limited)

1 はじめに

人工衛星による地球観測では、日常の定常的な観測を遂行する他に、取得データの解析アルゴリズムを改良し、精度・確度の高い情報をユーザーへ提供する取り組みを日々行っている。アルゴリズム改良に伴い、観測データを解析し直す作業(再処理という)が行われるが、観測データは日々蓄積され増加し続け、その結果、再処理に要する時間は観測期間に比例して増大する。

温室効果ガス観測技術衛星「いぶき」 (GOSAT: Greenhouse gases Observing Satellite)を例にとり、人工衛星から地球 を観測する様子を図1に示す. 衛星の進行 方向を指してアロングトラック方向とい

う.

「いぶき」に搭載された2種類のセンサーは、それぞれ異なる動作をしている. TANSO-CAIという可視画像センサーは進行方向に垂直な方向(クロストラック方向)に帯状に連続観測を行う.一方、衛星フーリエ分光器であるTANSO-FTSは、図1に楕円印で示すように、飛び飛びに複数点を観測する. CAIによる一日分の観測データの例を図2に示す.

次に、いくつかの地球観測データの例を示す。 図3-aは「いぶき」が観測した大気中の二酸化 炭素濃度の月平均を示し、年々、その濃度が上 昇していることを示す。次に図3-bは、GPM主 衛星に搭載された二周波降水レーダ(DPR; Dual frequency Precipitation Radar)の観測例で ある。最後に、図3-cに全球降水観測計画(GPM: Global Precipitation Measurement)から提供して

いるGSMaP(世界の雨分布を示すデータ)の例を示す.地球全体の降水量のデータを,ほぼリアルタイムに提供している.

これらの例のように,地球全 体の観測を継続的に実施することにより地域ごとに時間経過り とにより地域ごとに時間経過り をできるようにな気 をできるようにの気観測で また,局所的な短時間の観測と は得られない知見を得ること は得られないらのデータは ははきる。これらの で気候変動等の理解のために用いられている。

衛星観測データは、各々のプロジェクトを推進している機関のホームページから無償で公開されている.「いぶき」については国立環境研究所から、GPM

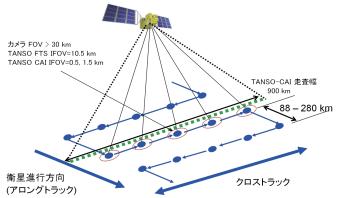


図1 いぶき(GOSAT)による地球観測の様子



図2 TANSO-CAIの観測例(日照域) (JAXA/EORCのホームページより)

	2010年1月 (V02.21)	2011年1月 (V02.21)	2012年1月 (V02.21)	2013年1月 (V02.21)	2014年1月 (V02.21)	2015年1月	2016年1月 (V02.60)
	3			3.4		No Image	-4
	2010年2月 (V02.21)	2011年2月 (V02.21)	2012年2月 (V02.21)	2013年2月 (V02.21)	2014年2月 (V02.21)	2015年2月 (V02.40)	2016年2月 (V02.60)
	2010年3月 (V02.21)	2011年3月 (V02.21)	2012年3月 (V02.21)	2013年3月 (V02.21)	2014年3月 (V02.21)	2015年3月 (V02.40)	Spinol Lab
						The Court	
2009年4月	2010年4月 (V02.21)	2011年4月 (V02.21)	2012年4月 (V02.21)	2013年4月 (V02.21)	2014年4月 (V02.21)	2015年4月 (V02.40)	
No Image							
2009年5月	2010年5月 (V02.21)	2011年5月 (V02.21)	2012年5月 (V02.21)	2013年5月 (V02.21)	2014年5月 (V02.21)	2015年5月 (V02.40)	
No Image					32		
2009年6月 (V02.21)	2010年6月 (V02.21)	2011年6月 (V02.21)	2012年6月 (V02.21)	2013年6月 (V02.21)	2014年6月	2015年6月 (V02.40)	
	TANK				No Image	100	
2009年7月 (V02.21)	2010年7月 (V02.21)	2011年7月 (V02.21)	2012年7月 (V02.21)	2013年7月 (V02.21)	2014年7月 (V02.31)	2015年7月 (V02.40)	
			7	-34	-		
2009年8月 (V02.21)	2010年8月 (V02.21)	2011年8月 (V02.21)	2012年8月 (V02.21)	2013年8月 (V02.21)	2014年8月 (V02.31)	2015年8月 (V02.50)	
				700			

図3-a 「いぶき」観測による二酸化炭素分布の変化(国立環境研究所のホームページより)

関連についてはJAXA第一宇宙技術部門の地球観測研究センター(EORC)から容易に入手できる. また、地球観測データ提供システム(G-Portalという)では、観測時刻や場所をキーにプロダクトを検索し、一般ユーザへ提供するサービスを行っている.

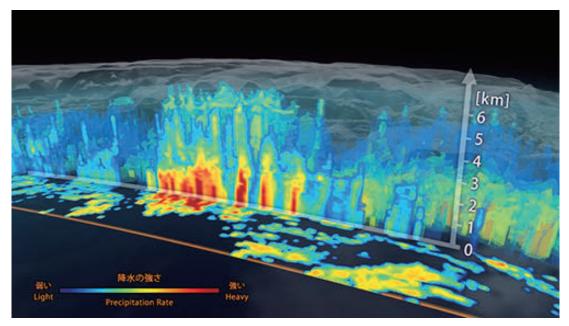


図3-b 二周波降水レーダ (DPR) の観測例 (JAXA/EORCのホームページより)

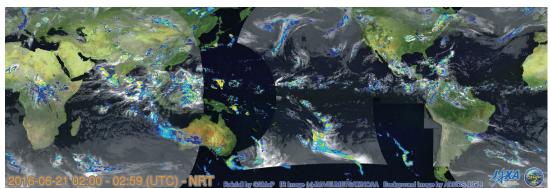


図3-c GSMaPの例 (JAXA/EORCのホームページより)

2 ゴールと背景

地球観測データは、衛星の軌道周回の単位にデータファイルを作り、各種のプロダクトとして提供される.これは定常運用という.図4に、その概念図を示す.

定常運用においては、衛星の各周回で得られたデータを、それぞれの周回単位にいかに早くユーザへ届けるか、ということが重要である.

観測データの蓄積に従い、データの解析処理アルゴリズム(物理的なモデルに基づく)を改善したり、また、処理パラメータの改訂(センサの校正係数の改訂など)が行われる。センサーの経年変化に対応するためのアルゴリズム、パラメータの変更も行われる。その際には、蓄積された観測データの全てを、新しいアルゴリズム、パラメータで再度の処理を行う。これを再処理運用と言っている。

定常処理や再処理に先立って, 処 理アルゴリズムが正しいかどうか, センサの校正係数が妥当かなど,処 理の結果に関する妥当性を検証する 必要が有り, そのために行うデータ 処理を試験処理ということがある. 特に長期間に渡って蓄積されたデー タの全数を処理して検証する場合に は長期試験処理という. その結果, データが妥当であると判断されて 初めて再処理の実施とユーザへの プロダクト提供が行われる.表1に 処理の種類と概要の例を示す. た だし,正確な定義は衛星プロジェ クトによって異なることが有るこ とに注意する必要が有る.

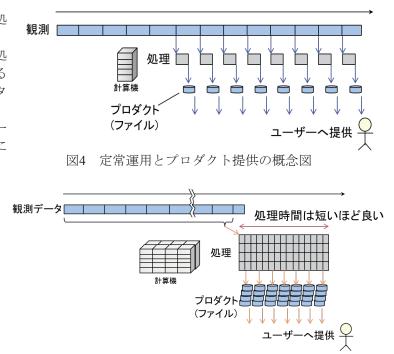


図5 再処理の概念図. 処理期間は短い方が良い.

XI /CE / IT C /AIX					
名称	概要	頻度			
準リアルタイム処理	観測データを周回単位に地上へダウンリンク	1回/軌道周回			
(NRT)	し、即時処理してユーザへ提供する.				
標準処理	軌道データ,姿勢データ等の確定後に再度,デ	軌道データ等の受領後			
(STD)	ータ処理し、ユーザへ提供する.このプロダク				
	トを標準プロダクトとし、保存・管理する.				
再処理	アルゴリズムの改訂,パラメータ(校正係数等)	1回/年			
	の改訂時に累積した観測データを全数、処理				
	し、ユーザへ提供する.				
長期試験処理	再処理に先立ち、アルゴリズム評価のために実	1回/年			
	施する処理. 再処理と同等の長期間の処理を実				
	施する.				

表1 処理の名称と概要

再処理は年に1回あるいはそれ以上、行われる。観測が長期化するとデータ量も膨大になり、全データの再処理を完了するまでに長い期間を要するようになる。例えば、降水量の観測を担ってきた衛星TRMMは2015年4月に観測を終了し、その観測期間は17年間以上に及ぶ。他にも衛星Aquaに搭載されたAMSR-Eは10年間のデータを蓄積した。

定常運用は筑波宇宙センター(以下、『TKSC』という)に整備されたデータ処理設備を用いて実施されるが、その計算資源には限りが有る。その理由はコスト、及び、設置場所の制約である。その制限により、再処理のために大きな設備を抱えるわけにはいかず、必要最小限の設備を整備してきた。そのため、従来は再処理の開始からプロダクトの提供までのレイテンシが1年以上に達したことが有った。その結果、データを必要とする研究者等のユーザには長期間の待ちを余儀なくさせる状況になっていた。

我々の目的は、JSS2を活用することにより、再処理のスループットを大幅に向上させることにある.目標として、従来は年単位の時間を要した再処理を数日から十数日へ短縮させる.これにより、

データ提供を抜本的に早め、防災・減災などへの社会貢献や、地球環境変動の研究利用に対するさらなる貢献を果たすことがゴールである.

現在, 再処理にJSS2を活用することを計画している衛星, センサーについて表2に示す. もちろん, 今後打上げを計画する衛星についてもJSS2活用が計画される.

衛星/センサー	観測期間(2016年6月現在)
TRMM/PR	17年
Aqua/AMSR-E	10年
GCOM-W(しずく)/AMSR2	4年以上(運用中)
GOSAT (いぶき) /FTS	7年以上(運用中)
GPM/DPR	2年以上(運用中)
GSMaP	18年以上
(TRMM, GPM等の複数マイクロ波放射	
計データの複合プロダクト)	

表2 再処理でのJSS2利用対象

なお,以下に記述する「いぶき」に関する成果については,JAXAが主体となって,JAXAのセキュリティ・情報化推進部スーパーコンピュータ活用課の協力を得て,日本電気株式会社に作業を委託¹⁾して得られたものである。GPM/GSMaPに関しては,同様に富士通株式会社に作業を委託²⁾したものである.

3 JSS2利用の方法

TKSCでの運用設備とJSS2では計算機の構成が異なる.したがって、処理エラーが起きないことの確認と計算結果であるデータの妥当性を検証する必要が有った.また、再処理の効率を最大にできるよう、必要な課題を洗い出し、必要な対策を施す準備を進めた.以下、次の項目に沿って今回のJSS2利用の成果を示す.

- 処理プログラムの移植とプロダクト検証
- 運用方式の設計と環境整備
- 3. 1 処理プログラムの移植とプロダクト検証
- (1) 「いぶき」処理プログラムの場合

TKSCで利用しているシステムとJSS2の主な相違点を表3に示す.

Steel and a steel					
項目	JSS2 (SORA-PP)	TKSC運用環境			
CPU	Intel Xeon E5-2643 v2	Intel Xeon E5640			
アーキテクチャ					
OS	RHEL 6.4	RHEL 5.5			
コンパイラ	gcc4.4.7	gcc4.1.2			

表3 プログラミング・実行環境の比較(「いぶき」のケース)

この他,ディレクトリ構成の差異などにより各種PATHの設定やMakefileの修正が必要であった. また,「いぶき」の運用開始時期(2009年)以降に言語規格に発生した変更に対応するためプログラム修正があった. ただし,アルゴリズムに影響するようなプログラム改変は必要なかった. プロダクトの妥当性については,gcc(GNUコンパイラ群)のバージョンの違いによると考えられる計算結果の微小な差異が出たものの,数値計算の誤差の程度と判断され,プロダクト品質への影響は無いと判断した.GOSAT(いぶき)/FTSのポインティング機構が2014年末から2015年初めにかけてA系からB系に切り替えられ,レベル1処理の校正係数が変更されたので,精度検証に当たっては,A系の運用期間のうち1カ月分とB系の1カ月分を対象とした.この検証は,JAXA/EORCの協力を得て実施した.

(2) GPM/GSMaP処理プログラムの場合

「いぶき」の場合と同様に、TKSCで利用しているシステムとJSS2の主な相違点を表4に示す.

項目	JSS2 (SORA-PP)	JSS2 (SORA-MA)	TKSC運用環境	
CPU	Intel Xeon E5-2643 v2	SPARC64XI fx	Intel Xeon E5640	
アーキテクチャ				
OS	RHEL 6.4	Linux	RHEL 6.4	
コンパイラ	Intel Cluster Studio XE 2013	Technical Computing Suit(富士通製)	Intel Cluster Studio XE 2013	

表4 プログラミング・実行環境の比較(GPM/DPR, GSMaPのケース)

JSS2のプレポストシステムという位置づけのSORA-PPを利用する場合は、OS, コンパイラ共に製品とバージョンがTKSCの運用環境と同一である。ディレクトリ構成の差異などにより各種PATHの設定やMakefileの修正が必要となったのは「いぶき」の場合と同様であるが、コンパイラが同一だったので、作成されるプロダクトは完全に一致した。このため、検証にあたってJAXA/EORCの協力を得るには及ばなかった。なお、降水に関する物理量は季節による変動が大きいことから、2014年の4月、7月、10月、12月の連続する各3日間のデータを対象とした。比較ツールとしては、地球観測で用いられることの多いHDFファイルの比較には"hdfcmp"を、それ以外のファイルの比較には"diff"を用いた。

一方、メインの計算システムであるSORA-MAは京コンピュータの後継とされるSPARCプロセッサを用いたシステムであり、主たる言語環境は富士通製である。そのためプログラム移植が必要となり、一部のアルゴリズムについてSORA-MAへの移植を試みた。このケースでは言語環境の変更への対応とともに、バイナリデータのエンディアンの違いに伴うプログラム修正が必要となり、異環境への移植に当たっての教訓、知見を得ることができた。今回、移植とデータ検証の作業規模や期間を考慮して、再処理にはSORA-PPを使用することを選択した。

以上により、「いぶき」、GPMともにJSS2上で再処理するためのプログラム移植・動作確認とプロダクト検証に成功したので、次に、大量のデータの再処理とユーザへのプロダクト提供をスムーズに行うための環境整備に取り組んだ.

3. 2 運用方式の設計と環境整備

(1) 「いぶき」の場合

レベル1処理においては、異なる衛星周回のデータは互いに独立に処理できる。したがって、まず処理アルゴリズムの実行時間をプロダクトーつ毎に単体で測定しておけば、目標とする処理完了までの時間に応じて必要な計算リソースを見積もることができる。プロダクトーつの平均的な処理時間は数分であったため、今回、10日間前後の処理期間を目標としてSORA-PPの30ノード(360コア)を占有利用することとした。

次に課題になるのは、処理の対象となるプロダクトの数が衛星の周回数に比例して増大し、6年半の間の蓄積数は数百万のオーダーとなり、ジョブ投入などの操作をマニュアルで行うのが実質的に不可能なことである。その操作を自動化するとともに、処理プログラムの実行状況に対応してジョブ投入を制御し、スループットの最大化を図れるような手段が必要である。その対策として、ジョ

ブの実行状況の把握、ジョブキューの蓄積状況の把握を行いながら、処理対象のプロダクトを選択 しジョブ実行スクリプトの生成と投入を実現できるシェルスクリプトを開発した

これらの準備作業は2015年10月までに完了し、また、TKSCから調布航空宇宙センター(以下、『CAC』という)のJSS2ヘデータ移行等の作業を行って、2015年11月に6年半分の観測データに対する再処理を実施した.

その結果、TKSCのシステムを利用する従来通りの再処理運用では約1年間を要すると見積もられた再処理を11日間、約9.5万コア時間(コア数 \times H)で完了した。図6にこの成果を示す。

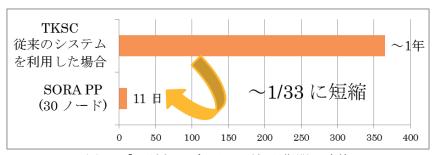


図6 「いぶき」データの再処理期間の改善

計算ノードを占有したことによりジョブの実行待ちもなく、大幅なレイテンシの改善が達成できた. 2016年1月には国立環境研究所等の機関ユーザに対してJSS2上から直接プロダクトを提供している.

その後、センサーの運用状況に応じてパラメータの見直しが行われ、約7年分の再処理が2回実施された.今回は、再処理の実行スクリプトに次のような改善を施している.

- 処理の進捗状況が容易に確認できること.
- 処理エラーが生じたケースについて、その内容を容易に確認・識別できること.
- エラーが生じた処理を容易に識別し、再投入できること.
- 処理ログ等の中間ファイルの自動削除.

2回の再処理の合計で約35.7万コア時間を要する再処理を、シェルスクリプトの改善効果もあって 最大の効率で実施できた。すなわち、占有した計算ノードの利用効率はほぼ100%に達した。今後の の再処理運用においても、JSS2の利用によって従来の運用を凌駕する効率的な再処理が実施できる ことが期待される。

(2) GPM/GSMaP処理の場合

「いぶき」のケースと同様に、処理プログラムの単体性能測定を行い、SORA-PPのノード数30程度を計算リソースの必要量の目途とした。

一部の処理アルゴリズムには並行実行ができないという制約が有る。まず、GPM主衛星に搭載されたセンサーDPRは処理レベル1からレベル3へ向かい順に処理する必要がある。ただし、同一レベルのデータ同士は処理の単位毎に独立している。すなわち、レベル1と2のデータは衛星周回毎に独立であり、レベル3データは日、または、月毎に独立である。そのため、同一レベルの処理においては並行実行が可能であり、したがって、割り当てるノード数に反比例して処理期間を短縮することが可能である。

さらに、JSS2には、MPIライブラリを活用して複数の周回の観測データを同時に処理起動することにより、全体のスループットの向上を図る機能が整備されている。図7はこの機能によるスループット向上の効果を示す。左右どちらのグラフも、横軸は時刻、縦軸は実行中のジョブ数を示す。

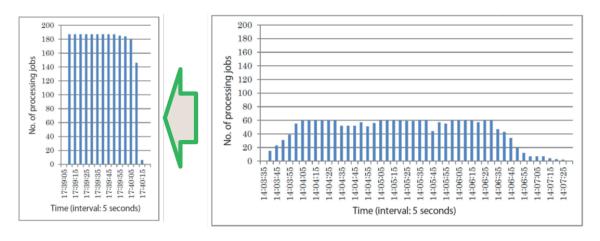


図7 MPI同時起動機能によるスループット向上

通常通りジョブ起動した場合,処理の時間が分程度と短いために,多数のジョブが立ち上がる前に処理完了してしまう.そのため,図7の右図に示したように,同時に稼働するコア数が伸びないという現象が生じた.このままでは多数のノードを同時利用できるというJSS2の効果が発揮できないが,同時起動機能を活用することにより,図7の左図に示すように同時に多数のコアを活用した処理が可能となる.これにより,計算リソースの利用効率が大幅に向上する.

ここでも同様に膨大な量のプロダクトの処理を自動化する手段が必要である. 処理レベル間のデータの依存性を考慮しながら実行ジョブを自動生成し、実行状態の管理や、エラーが生じた場合の再実行をサポートする機能として、JSS2に備わる衛星データ処理ワークフロー制御機能を使用した. この機能を使用するためには、予め処理の依存性を定義する情報をデータベースに登録するとともに、TKSCでの定常処理の記録(ワークオーダと呼ばれる情報とその関連情報)をJSS2に複写する必要がある.

2016年6月,これらの環境を活用し、GPM主衛星打上げ(2014年2月末)後の2年3ヶ月間の観測データの長期試験処理を実施し、再処理用アルゴリズムを評価するためのデータセットを作成した。今回はレベル1のみの試験処理であることから計算ノードの占有は行わず、2~3ノードを使用し、処理期間は二日間かかった。

今後,11月には全ての処理レベル(レベル1,2,及び3)の長期試験処理が計画されている.使用する計算リソースとしては,今回の処理性能から30ノード程度を利用する計画である.これにより,従来よりデータレイテンシを1/20程度に短縮できる見通しである.

4. 考察

再処理運用において同等のスループットを達成するためにTKSCで単独に計算機調達を行うとすると、そのコストは観測期間に比例して大きくなる.詳細は省くが、JSS2への移植作業やデータ移行のコストなどを考慮しても、計算機コストは従来の10倍以上に達すると考えられ、JSS2を活用する効果は大きい.

今後数年間の再処理予定によれば、TKSCとCAC間で輸送を要するデータ量は表5に示す通りである.なお、入力データは多少時間がかかっても事前にTKSCからCACへ輸送ないし伝送しておくことが可能なので、ここではユーザへ早く提供すべき再処理結果である出力データの量に注目している

昨年度までは、TKSCとJSS2間のデータの経路として利用可能なものはJAXAの社内ネットワークであるJAXAnetのみであり、低速(帯域幅100 Mbps)で送受信しなければならなかった。この場合、処理時間は短縮できてもデータ転送に時間を要し、ユーザへの提供を早くするという目的は果たせない。表5のデータ量では少なくとも約2 Gbpsの帯域幅が必要である。これに対する対

衛星,センサ	FY27	FY28	FY29	FY30
GOSAT	55TB	>60TB	未定	未定
TRMM/PR	_	>40TB	>150TB	未定
GPM/DPR	_	>50TB	未定	>90TB
AMSR-E,	12TB	15TB	9TB	未定
AMSR2				

表5 今後数年の再処理予定と出力データ量

策として本年5月初めに、TKSCの地球観測システムとJSS2とをSINET5を介して高速に接続する環境を整備した.現在、TKSCとJSS2は10 Gbpsの帯域を持つネットワークで接続されている.これにより、GPM/DPRの試験処理では、TKSCからユーザへのデータ提供に要するオーバーヘッドを従来の約1/10に短縮できた.このデータ伝送路の高速化により、JSS2からTKSCへの伝送時間を含めても、JSS2で処理した方が、TKSCの既存設備で処理するより、大幅な時間短縮が可能となった.

さらに、一般ユーザは対象ではないが、「いぶき」の機関ユーザ(国立環境研究所など)に対して、JSS2からプロダクトを直接提供する方式を採用した. 従来は必ずTKSCからユーザへ送付しているため、新たなデータの経路を設けたことになる. 外為法に基づく手続きを経る必要が有ったが、今後、定常的なデータ提供手段として検討する価値は非常に大きい.

また、昨年度より、衛星データ処理はJSS2の重点利用課題となった. その効果は大きく、以下の通り地球観測業務にとって大いに助けとなっている.

- 優先度が最高レベルに設定されているため、ジョブ投入後の待ちが少ない.
- 申請によりノードを占有することが可能である.この場合,完全に滞りなくジョブ実行できる.

こうした環境が無ければ、ジョブ実行の待ちにより、スループットの大きな低下が避けられない. TKSCにおけるJSS2利用は、AMSR-E、並びに、GCOM-W衛星に搭載のAMSR2へも拡大される. 6月にはAMSR-Eの10年分のレベル1処理を完了した. さらに、2016年中にAMSR-E及びAMSR2の全処理レベルの再処理を計画している.

他の機関の事例も参考となる.「いぶき」と同様に温室効果ガスを観測する衛星にOCO-2(Orbiting Carbon Observatory-2, 2014年7月2日打上げ)がある.これはNASAが打ち上げ,運用を行っている. 昨年、NASAのAmes Research CenterにあるスーパーコンピュータPleiadesを利用して10か月分の観測データの再処理を行った. 観測データの60%の再処理にHaswell Node (2088ノード、4176CPU)のうち500ノードを利用し、140万CPU時間を費やしたとのことである. NASAのように潤沢な予算を誇る機関であっても、スーパーコンピュータを利用できることの利点は大きく、この規模の計算を短期間に実施できたのはその効果であると、NASA担当者も認めている.

5. 結論

JSS2利用の効果は大きく, GOSAT/TANSO-FTSの再処理を従来の33倍のスループットで実現できた.

GPM/DPRの試験処理では、データ提供を従来の約4倍のスループットを達成できた.11月の長期試験処理や、2017年度に計画されている再処理では、計算ノードを増やすことにより処理期間の大幅な短縮が達成できる見通しである.

JSS2の利用効果は、AMSR-E、及び、AMSR2においても期待される.

2018年に打上げ予定のGOSAT-2は、JSS2の利用を前提として運用することを考えており、そのためのシステム設計、構築を進行中である.

以上の通り、TKSCでの衛星データ処理の業務では、JSS2は業務効率化とユーザの利便性を両立させて実現するための必須のツールであると考えられる.

参考文献

- 1) 日本電気株式会社,「温室効果ガス観測センサ(FTS)処理固有部 ソフトウェアの移植」 (JX-PSPC-421815) 成果報告書(GOSAT-H27移植-提005), 2015, 日本電気株式会社
- 2) 富士通株式会社,「全球降水観測計画 (GPM) プロダクト再処理時のJAXA Supercomputer System 2 (JSS2) 利用検討 (その2)」 (JX-PSPC-410301) 成果報告書 (FJ-GPM-JSS2利用検討-14-020), 2015, 富士通株式会社
- 3) Tsengdar Lee, Climate Data Downscaling Projects at NASA, Presented at SC15 Conference, Nov 16 20, 2015

謝辞

調布のスーパーコンピュータJSS2の遠隔利用において、移植時のプログラム相談やジョブ実行の効率化対策などにおいてJSS2運用窓口の協力を得た、紙面を借り謝辞を申し上げる.