

温室効果ガス観測技術衛星「いぶき」(GOSAT) 研究用計算設備の 取り組み

(Effort of GOSAT Research Computation Facility for Greenhouse Gases Observing Satellite, IBUKI)

HIRAKI, Kaduo MATSUNAGA, Tsuneo YOKOTA, Yasuhiro

National Institute for Environmental Studies

NINOMIYA, Keiichiro

Information and Mathematical Science Laboratory Inc.

OTSUKA, Toshio

MAKINO, Daichi

KUBO, Hiroji

Wang, Xin

NS Solutions Corporation Nippon Dacom Service, INC.

SGI Japan, Ltd.

1. GOSAT プロジェクトについて

温室効果ガス観測技術衛星(いぶき/GOSAT)は、2009年1月に打ち上げられた世界初の二酸化炭素/メタン観測専用衛星である。GOSAT(いぶき)プロジェクトは、環境省(MOE)、宇宙航空研究開発機構(JAXA)、国立環境研究所(NIES)の共同プロジェクトとして推進されている。

NIESは、データの高速処理、検証、配布を担当しており、定常的な高速処理(定常処理)に必要な研究処理も行っている。現在行われている主要な高速処理は、衛星からの観測データを解析し、太陽～地表面～衛星間の大気中における二酸化炭素及びメタンのカラム平均濃度を算出する処理である。処理データ数としては、約3万点/月(ただし、低太陽高度、雲のかかったデータを除く)である。また、研究処理には各種パラメータの調整等も含め、年単位の同一データを複数回処理する必要がある。一般提供用データを算出している定常処理は、研究処理の結果をフィードバックし、研究処理とは独立に行われている。

処理の対象(処理開始済みのものを含む)には、短波長赤外フーリエ変換分光計(SWIR)、熱赤外フーリエ変換分光計(TIR)、雲エアロゾルイメージャ(CAI)、大気輸モデルが予定されている。例えば、SWIRの研究処理では、1観測データあたり3~5分の計算時間が必要となるため、観測済み(2010年11月現在)の約49万点のデータを処理する為には、延べ145万~245万分ほどの計算時間が必要となる。

このような研究処理に対する需要に応える為にNIESではGOSAT研究処理専用スーパーコンピュータ(GOSAT RCF)を2010年3月末に導入した。

2. GOSAT RCF の構成

GOSAT RCFは、160台の計算ノードを中心としたPCクラスタ型のスーパーコンピュータである。各計算ノードには、CPUとして2つのIntel Xeon E5530とCPUと同数のGPGPUが搭載されている。GPGPUは、NVIDIA Tesla C2050を採用した。2組のCPU-GPGPUが同一のシステムボードに搭載されているが、CPU-GPGPU間のそれぞれの通信を干渉させない為に、図1の様に2つのチップセット(I/O HUB)を搭載している。

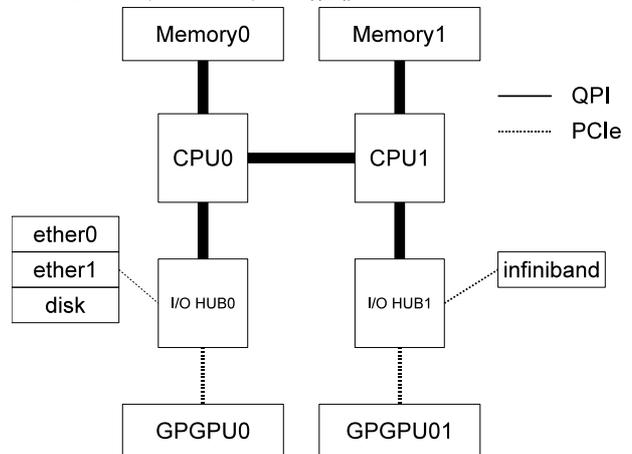


図1: 計算ノードの内部構成(概略)

インターコネクには、infiniband(QDR)によるFull-Bisection接続(図2)を用い、共有ファイルシステムもこのinfiniband上でLustre File Systemにより行われている。共有ディスク容量は、実効約100TBである。

3. Green スパコンに向けて

GOSAT RCFの設計当初から消費電力および床耐荷重が問題となった。

3.1 ハードウェア

省エネ化の方針として、計算ノード数を減らしてでも、効率の良い部品を積極的に採用した。

そこで、回転運動などを必要とする稼働部を出来るだけ少なくする、高効率の電源ユニットを利用する、効率的な空調を採用するという3項目を中心に設計を行った。

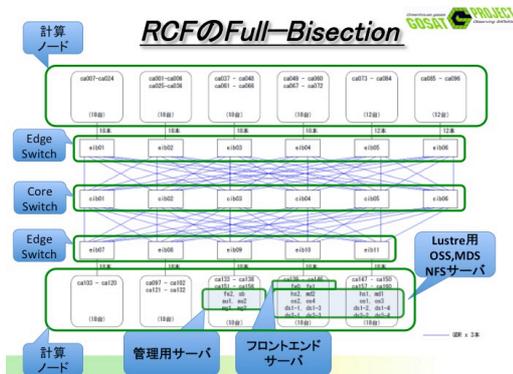


図2:GOSAT RCFのFull-Bisection

その結果、ほぼ全台の起動用ローカルディスクへのSSDの採用、GOLD PLUSの電源ユニットの採用、HACS型の空調システムの採用を決定した。

また、効率的に計算を行い省エネ化することも検討された。観測された光の波長ごとの独立した計算が頻繁に行われる可能性が高いことから、GPGPUを搭載し、計算速度向上と省エネ効果の両方を同時に満たすことにした。

3.2 ソフトウェア

気象条件、衛星のメンテナンスなどによりデータ数が減少する、処理した結果を利用者が検討している間は、ジョブ投入が減少するなど運用上の理由を考慮し、「エコマネージャ」というジョブ投入連動型の計算ノード運用システムを構築した。

このシステムは、キューが空になると一定時間ごとに、設定された台数ずつ計算ノードをシャットダウンし、あらかじめ設定された台数を残して計算ノードを停止させる。ジョブが投入されると処理に必要な台数を確保するように計算ノードを起動し、処理を行う。ここで問題になるのが、死活管理ソフトウェアとの整合性である。GOSAT RCFでは、死活管理にNagiosを利用しており、各計算ノードの死活確認に利用するプラグインがエコマネージャの情報を参照することにより、誤検知を防いでいる。すなわち、エコマネージャが停止させたノードの状態を真とし、それ以外を偽とする。この状態とNagiosのオリジナルのプラグインの真偽の和を評価し、死活判定を行っている。

GOSAT RCFでは、ジョブキューを大きく分けて大量処理用、開発用、実験用の3種類用意している。この省エネ機能は、現時点では、大量処理用キューにのみ設定され、レスポンスが重視

される開発用および実験用キューには、一部の例外を除き設定されていない。この理由は、大量処理を行う場合、計算時間が1日以上かかる場合もあり、初回のジョブ投入時に必要となる数分のレスポンス低下は、問題視されないためである。

4. 計算成果

GOSAT RCFの現状での計算成果は、SWIRの改良である。約49万地点分の観測データは、1024コアを利用し26時間~41時間かけて処理され、その結果から得られた計算方法は、定常処理に活かされている。定常処理の結果は、<http://data.gosat.nies.go.jp/> から一般のデータユーザに向けて配信されている。

本システムのLinpack性能は、74.84TFLOPSであり、その際の総使用電力の最大値は、約140kWであった。また、Green500で指定された計算ノードのみを計測の対象とする方式（

<http://www.green500.org/docs/pubs/tutorial.pdf>参照）では、117.15kWであった。これにより

ワットあたりの性能が636.36MFLOPS/WとなりGreen500では、11位、Top500では、102位を獲得することが出来た。

5. 今後の課題

各研究グループがストレージをGOSAT RCF外に保有しているため、当初のディスク容量の要求は、最小限でよく、GOSAT RCFでの計算に必要なデータを転送して利用するということがあった。しかし、実際には、入力パラメータ違いの計算結果が複数生成され、結果を各研究グループのストレージへ転送する時間が予想以上に必要となった。これを解決する為には、ディスク容量を増強し、入力データの収集・保存、出力データの分析・保存をGOSAT RCFのみで行う必要がある。GOSAT RCFのFull-Bisectionは、図2のようにエッジスイッチ1台分の余裕があるため、そこへinfinibandスイッチとLustre File SystemのOSSとOSTを増設することで、この問題を解決できる。

また、GPGPUの開発環境であるCUDAのバージョンアップが目覚しく、必ずしも最新の機能を使いきれていない。CUDAで利用できる機能の種類やその性能が安定したところで、GPGPUの利用をより積極的に行う必要がある。

最後に、導入に係った全ての皆様に感謝いたします。