

並列分散処理による衛星データ解析システムの構築

A Ground-Based System for Processing Information Sent
from the Interferometric Monitor for Greenhouse Gasses

新 留 照 英⁽¹⁾ 川 浦 淳 一⁽¹⁾ 鎌 田 憲 幸⁽²⁾ 佐 治 秀 道⁽³⁾
Teruhide Junichi KAWAURA Noriyuki KAMATA Hidemichi SAJI
NIIDOME

抄 録

宇宙開発事業団(NASDA)が開発した地球観測プラットフォーム技術衛星“ADEOS: Advanced Earth Observation Satellite”(愛称“みどり”)に搭載された通商産業省開発の温室効果ガス観測センサーであるIMG(Interferometric Monitor for Greenhouse Gasses)の地上処理システムを構築したのでその概要を述べた。IMGデータ解析システム(IMGDIS: IMG Data and Information System)は、1993年12月の基本設計開始以降、約3年間の設計・製作・調整期間を経て、1996年8月のNASDAによる衛星打ち上げ後、10月にシステムチェックアウトを完了し、運用を開始した。IMGDISは、従来の衛星データ処理設備の主流であったミニコンピュータやスーパーコンピュータ、メインフレームを主力にした機器構成ではなく、約30台の汎用ワークステーションをFiber Distributed Data Interface(FDDI)ネットワークで接続した分散並列処理システムである。また、独自にミドルウェアを開発しアプリケーションソフトウェアのモジュール化を図り、コストパフォーマンスの高い柔軟性、拡張性に富むシステムとなった。

Abstract

A ground-based system for processing information sent from the Interferometric Monitor for Greenhouse Gasses (IMG) has been constructed, which is a sensor for observing greenhouse effect gasses that was developed by the Ministry of International Trade & Industry and is mounted on the Advanced Earth Observation Satellite (ADEOS) with a hypocoristic name "Midori", the earth observatory platform technical satellite developed by the National Space Development Agency of Japan (NASDA). The IMG Data and Information System (IMGDIS) was drawn up basically first in December, 1993, and it took three years to design, manufacture and condition the basic IMGDIS. In August, 1996, the satellite carrying thus developed IMGDIS was launched by NASDA. Then the IMGDIS was continually checked out till October, 1996, and was put to practice use. The system, IMGDIS, is a distributed parallel processing system consisting of about thirty universal workstations connected by Fiber Distributed Data Interface (FDDI) network, which differs from the conventional satellite data processing system consisting mainly of minicomputers, super-computers and main frame computers. From the viewpoint of software, the application software installed in the IMGDIS is modularized by using the originally developed middle-ware, to bring about a flexible and highly open-ended system with a high cost-performance.

1. 緒 言

IMGは、マイケルソン干渉計を用いて大気中に存在する二酸化炭素やメタン等の温室効果気体を低コストで高精度にかつグローバルに観測する世界初の衛星センサーである。

IMGのミッションの達成には、センサーの開発、センサーデータから大気物理量を導き出すデータ解析アルゴリズムの開発、及びデータ解析システムの開発が不可欠であり、これらの開発は通商産業省の委託を受け、各々、(財)資源探査観測システム研究開発機構

(JAROS)、(財)電力中央研究所(CRIEPI)及び、(財)資源・環境観測解析センター(ERSDAC)で遂行された。新日本製鐵は、ERSDACよりデータ解析システムの開発を受注した新日鉄情報通信システム(株)(以下ENICOMと記す)と協同で開発にあたった。

2. IMGの概要

2.1 センサー概要

IMG開発は、近年の化石燃料消費拡大に伴い発生するガスの温室

⁽¹⁾ エレクトロニクス・情報通信事業部
産業システムソリューション第三部 マネジャー
東京都渋谷区代々木3-25-3 大東京火災新宿ビル
▽ 151-8527 ☎ (03)5352-9347

⁽²⁾ 新日鉄情報通信システム(株) 鉄鋼システム事業部
科学技術システム部
⁽³⁾ (財)資源・環境観測解析センター 技術第一部

効果に対する全地球的なモニタリングと大気の様子によるそれらガスの振る舞いの解析へのニーズに応えることを目的とした。IMGは地表及び大気からの熱赤外放射光を検知するセンサーであり、熱放射スペクトルを連続的かつ全地球的に得ることができる。その中には水蒸気、二酸化炭素、メタン、一酸化二窒素、一酸化炭素、オゾン等の各種温室効果気体スペクトルが含まれており、これらを抽出して気体濃度及び温度の高度分布を求める。

図1にIMGの光学系の構成を示す。地球から放射される赤外光は、ビームスプリッターによって走査鏡側に反射する光と透過して固定鏡側に進む光に二分される。走査鏡側に反射した光は走査鏡を前後に移動させることによって透過した光との間に光路差が発生する。この光路差によって検出器上の光強度が変動する。この変動をインターフェログラムと呼び、入射光のスペクトルをフーリエ変換したものになっている。従って、得られたインターフェログラムをフーリエ逆変換することによって入射光のスペクトルを得る。

IMGを含めて日・米・仏3か国の8センサーを搭載する衛星ADEOSは、約800km上空を約110分で地球を縦に一周する太陽同期準回帰軌道を周回している。これによって、4日間で全地球の粗観測、41日間で詳細観測ができる。

2.2 データ解析アルゴリズム概要

IMGDISで行う主たるデータ解析アルゴリズムには、誤差の少ないスペクトルを求めるための補正処理と、スペクトルから大気に含まれる温室効果気体濃度を求めるリトリバル処理がある。

2.2.1 スペクトル補正処理

1) 位相補正処理

IMGでは、電気的あるいは光学的な原因により信号の位相にエラーが発生する。これらの位相誤差はインターフェログラムを逆複素フーリエ変換して取り除く。

2) 非直線性補正処理

IMGで使用している検出器の内、HgCdTe¹⁾には入出力関係に非直線性が含まれる。この特性を多項式近似を用いて補正する。

3) 干渉効率補正処理

IMGの移動鏡の高周波振動によって干渉効率が変動する。これをアライメントを検知するセンサー出力を用いて補正する。

2.2.2 リトリバル処理

校正されたスペクトルを入力し、気温、物理量を出力するリト

リバル処理は、物理量を推定するインバージョン部と、推定された物理量に対応して観測されるべき放射スペクトルを計算するフォワード部により構成される。

1) インバージョン部

逆問題を解くための行列演算とフォワード部との繰り返し計算の収束判定を行う。

2) フォワード部

与えられた物理量の値に対応して、観測されるべき放射輝度を得るため放射伝達の計算を行う。

2.3 データ解析システム概要

図2に示すように、衛星に搭載されたセンサーよりダウンリンクされたデータは、NASDAの地上処理設備によってセンサー毎にD1テープに記録される。IMGDISは、このD1テープで受領したセンサーデータを処理・解析し、温室効果気体濃度及び温度の高度分布を導出し、研究者を初めとするユーザーへ提供する。従って、IMGDISは、高度な解析アルゴリズムを高速に実行するばかりでなく、継続的なアルゴリズムの改訂に伴うデータの再処理にも対処できなくてはならない。また、広範な研究者からのデータ配布要求にも対応する必要がある。

IMGDISで生産・配布するデータはプロダクトと呼ばれ、ユーザーの幅広い要望にこたえて加工度の異なるプロダクトを作成する(表1参照)。データは加工度によってレベルが異なり、かつ付帯情報によって種別が異なる。表1は基本的なものを示しており、実際は加工方法の違いによる種別を含め、トータル20種類のプロダクトを作成・配布する。

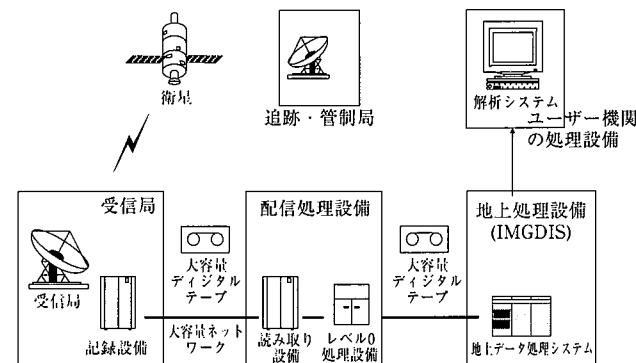


図2 データ解析システム概要

表1 プロダクト一覧

プロダクト名称	データの定義
レベル0	NASDAデータより受領した原データ
レベル0A	観測単位にまとめた工学値変換済みのインターフェログラムデータ
レベル0B	位相・非直線性補正済みのインターフェログラムデータ
レベル1A	感度未校正複素スペクトルデータ
レベル1B	校正済みパワースペクトルデータ
レベル1C	雲情報を含むレベル1Bデータ
レベル1D	気象データを含むレベル1Cデータ
レベル2	温室効果気体の高度分布データ

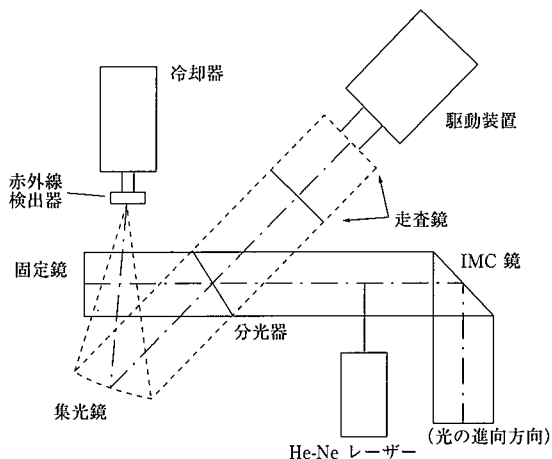


図1 IMG光学系の構成

¹⁾ HgCdTe: 赤外光の検出に用いられる。MCT検出器と呼ばれることも多い。入出力に非直線性があることが知られている。

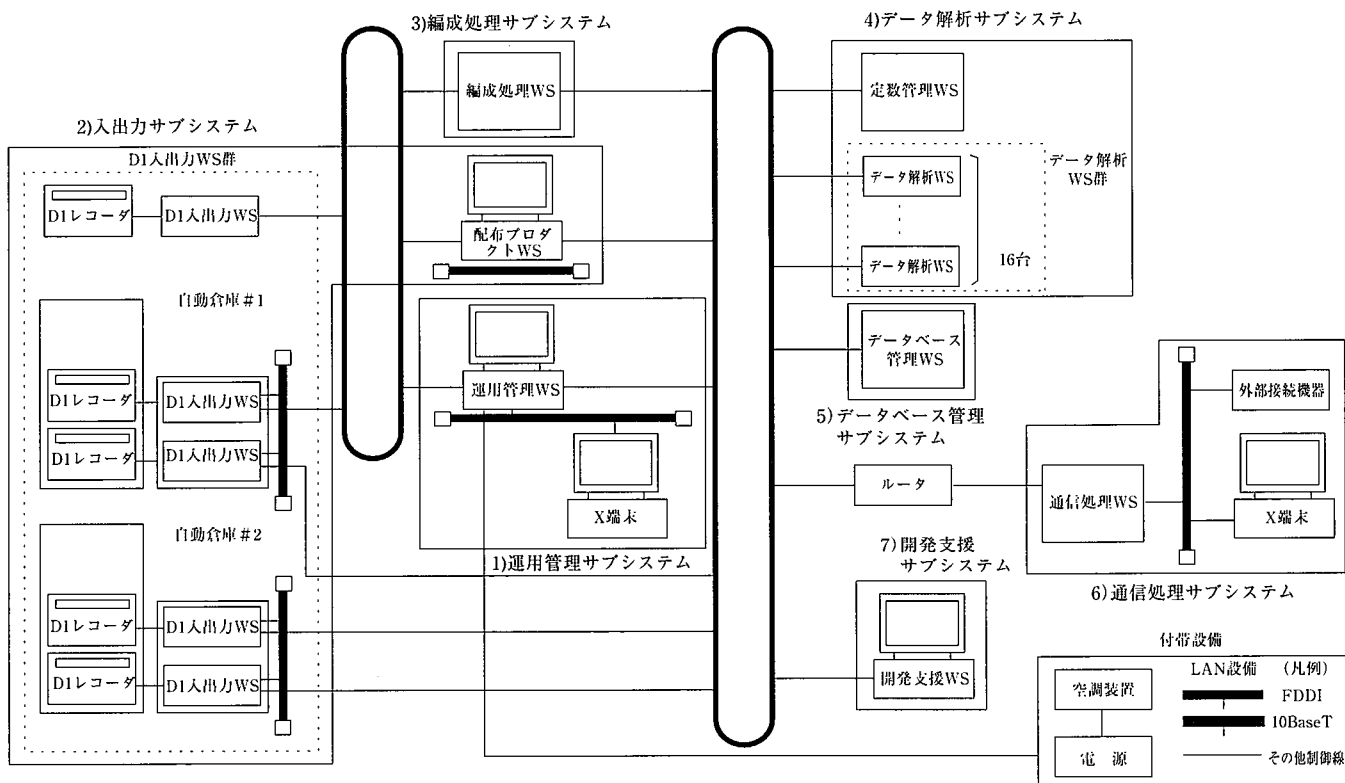


図3 IMGDISのハードウェア構成図

2.3.1 システムハードウェア構成

IMGDISのシステムの構成は、前述のシステム機能を考慮して図3に示すサブシステム構成を取った。

1) 運用管理サブシステム

- ・IMGDISの管理・運用を行う。
- ・IMG観測計画の立案、観測計画の進捗管理を行う。立案した観測計画はNASDAに送付する。
- ・ユーザーからの配布要求を受け付け、プロダクトの作成進捗管理を行う。

2) 入出力サブシステム

- ・NASDAより受領したD1テープのレベル0データの読み込みとデータチェックを行う。
- ・IMGDISで作成したプロダクトデータをD1自動倉庫へ格納する。

3) 編成処理サブシステム

- ・衛星データを工学値変換し、観測ユニット単位に編成する。
- ・観測点の位置を計算し、編成後のデータはデータ解析サブシステムに送られる。

4) データ解析サブシステム

- ・レベル0からレベル2への処理が行われる。実際の処理は以下の4種類に分けられる。

(1) スペクトルデータ作成処理(IMGデータ処理)

インターフェログラム(レベル0 A)からスペクトル(レベル1)への変換処理の中で、センサー特性に基づく補正処理、校正処理を実施する。

(2) リトリバル処理(IMGデータ解析)

ここでレベル1からレベル2への処理が行われる。処理内容はスペクトルから反転法を用いて、気温分布及び気体濃度分布を算出する。

(3) OCTS処理

ADEOS搭載センサーの一つであるOCTS(Ocean Colour and Temperature Scanner)のデータを用いてIMG観測位置での雲の状況を判別し、同時にブラウザ画像²⁾を作成する。

(4) 気象データ処理

(財)気象業務支援センターから受領する客観解析データからIMGデータ処理に必要な気象データを抽出・編集する。

5) データベース管理サブシステム

各種プロダクトの在庫管理及びカタログデータの管理、衛星運用計画・実績ファイルの管理を行う。

6) 通信処理サブシステム

NASDAに対してセンサー運用計画及び実績に関する情報の送受信と生産されたプロダクトのカタログ情報の送信を行う。ユーザーに対してカタログ検索、配布プロダクト要求受け付け、センサー観測情報等を提示する。

7) 開発支援サブシステム

アルゴリズムの新規開発のためのツール、データを一元的に管理する。また、IMGDISのソフトウェアのソース管理の機能を持つ。

2.3.2 ソフトウェア構成

IMGDISのソフトウェア機能は大別すると表2に示すように五つの機能により構成される。

IMGDISでは、新日本製鐵開発のミドルウェアの上に各種サーバーを構築し、データ処理等のアプリケーションは、それらサーバーのApplication Programming Interface(API)を使用する仕組みを採用した。これによって、アプリケーションの外部インタフェースは、全てサーバー経由とすることでアプリケーション固有

²⁾ ブラウズ画像：衛星搭載のセンサーが観測した地域の大気の状態(特に雲の状態)を、ユーザーが確認できるように可視光や赤外光をRGBに割り付けたビットマップ画像。

表2 ソフトウェア機能大項目

データ処理ソフトウェア	リトリーバルアルゴリズムを始めとする処理アルゴリズムにより、レベル0データから各種解析に適した形式のデータに変換し、保存するソフトウェア群。
システム支援ソフトウェア	処理結果として作成される各種プロダクトをユーザーからの要求に応じて配布する業務、センサーのミッション運用要求の作成支援等を行うソフトウェア群。
業務管理ソフトウェア	データ加工の工程を作成、実行、管理するソフトウェア群。
システム運用管理ソフトウェア	本システムを正常に運用させるべく本システム全体の資源を管理・制御するソフトウェア群。
システム内部インフラストラクチャソフトウェア	本システム内の各アプリケーションが共通に利用するインフラストラクチャ的なソフトウェア群。

の設定を排除し、システム上のいかなるホスト³においても等価的に動作可能な構成を確立した。

3. 技術的特徴

以上述べたように、IMGDISの具備すべき要件は広範にわたるが、その中でも以下の項目は本システムを特徴づけるものである。

1) 高速大容量処理

毎日衛星から送られてくる約30GBのデータを一日の作業時間内で、大容量FFT処理⁴及びリトリーバルアルゴリズムの処理を含めて滞留が生じないように実行する必要がある。

2) ワンマンオペレーションの実現

WSの台数が30台規模のシステムでワンマンオペレーションを実現するためには、人手でしかできない作業以外は基本的に自動的に実行可能とする必要がある。

3) 柔軟性・拡張性の確保

データ解析アルゴリズムは開発性の高いものであり、運用期間中を通してより精度の高いアルゴリズムへと開発が繰り返される。従って、より良いアルゴリズムが開発されればそのアルゴリズムに即座に変更可能である必要がある。その際、新しいアルゴリズムの処理時間の増大に伴って、ハードウェアの増設を含めたマシンパワーの増強に柔軟に対応できるシステムが要求される。即ち、アプリケーションの変更柔軟であり、かつデータ処理性能の強化がデータ処理WSの台数を増やすことで容易に実現できるようなスケラビリティに富んだ仕組みを構築する必要があった。

これらの要件に対して、それぞれ以下の対処方法によって、一般的なUNIX⁵のWSを用いて極めて大量なデータを定められた時間内に処理できるシステムを実現した。

1) 高速・大容量処理

高速な処理を実行するには高速な科学技術計算可能なマシンの採用が不可欠であるが、スーパーコンピュータや大容量ミニコンピュータによる実現方法ではシステムの柔軟性・拡張性の面で問題がある。そこで、システムで実施すべき処理を細かく分割し、各WSで分散実行することによって高速かつ大容量の処理が可能なシステムとした。加えて、非常に重い計算を行う処理(例えばスペクトル処理では、26万点のFFTと13万点のFFTを10秒程度で72回実施する必要がある、逆問題求解処理であるリトリーバルアルゴリズムは約2000スペクトルピンの処理を10秒程度で実施する必要がある)は、一台のWSで所定の性能を満足することは不可能であり、並列分散処理を採用することによって対処した。

2) ワンマンオペレーションの実現

ワンマンオペレーション実現のため、処理の自動実行、実行計画作成、実行管理等業務管理機構の開発を行い、オペレータは基本的に作業実行計画の作成・指示、処理の実行監視やメディアの交換・ラベリングといった人手作業に専念できる仕組みを提供した。

3) 柔軟性・拡張性の確保

標準的WSの採用により、ホスト間での特異性を排除し、アプリケーション実行環境として可能な限り等価なシステムとすることによって、アプリケーション実行ホストを自在に変更できるようにした。また、アプリケーションの外部アクセス(ファイル、DB、画面等の入出力)をクライアント・サーバが標準化すると共に、アプリケーションからはサーバの実行ホストを隠蔽化した。これによりアプリケーション変更に対する柔軟性を向上させた。

以下では、IMGDISにおいて構築した特徴的な機構について説明する。

3.1 分散処理の実現

独自開発ミドルウェアにより実現した分散制御のメカニズムは、大きく三つの階層から成り立っている。

1) 業務管理機構

データ処理を分散実行し、進捗管理するためのサーバ群であり、データ加工を一般化した“業務”モデルを自動実行する仕組みを持つ。ミドルウェアの最上位階層にあたる。

2) アクセスサービス

IMGDIS内の各アプリケーションが、汎用的に利用するサービスを“サーバ・クライアントモデル”で実現。ミドルウェアの中間層にあたる。

3) IPC (Inter-process Communication) 機構

IMGDISの分散処理の基盤となる複数プロセス間のプロセス間通信手段で、異なるホスト上のプロセス起動、終了に関する制御も行う。ミドルウェアの最下層にあたるIPC機構は、上記アクセスサーバとクライアントとのインターフェースを全てつかさどる。IPC機構を用いたサーバ・クライアントモデルの概念図を図4に示す。

3.2 自動実行の実現

自動実行機構の実現に当たってまず、“業務”を定義した。“業務”とは、ある目的を達成するためのプログラムの組み合わせ、処理の流れを定義したものである。

全てのデータ加工処理は、業務として定義され、その構成(プログラムの組み合わせ、処理の流れ)は業務コンフィグレーションファイルに記述される。この定義に基づき作業実行監視機能が各作業を開始・終了することになる。

³ ホスト：本システムを構成するワークステーション及びその実行環境のこと。

⁴ FFT処理：Fast Fourier Transform(高速フーリエ変換)の略。離散フーリエ変換を高速に行うための計算処理のこと。

⁵ UNIXはX/Openカンパニーリミテッドがライセンスしている米国ならびに他の国における登録商標。

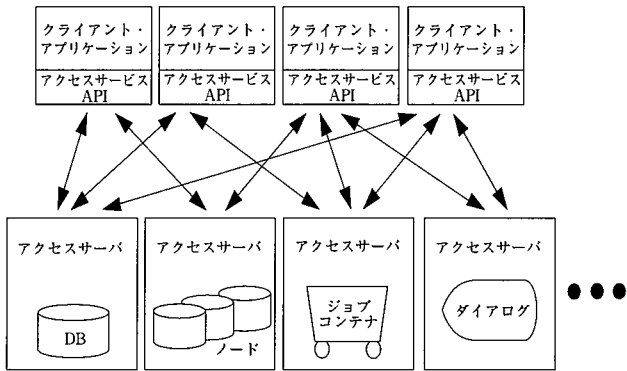


図4 アクセスサービス

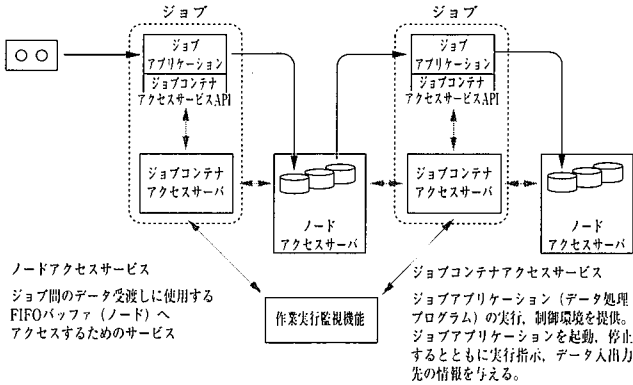


図5 データ処理の実行モデル

作業実行監視機能は作業計画に基づき所定の時刻に対応する業務を立ち上げる。業務の立ち上げはその業務を構成するジョブとサービスを立ち上げることを意味する。また、各ジョブは入力ノードからデータを読み込み、出力ノードに加工後のデータを書き込む作業を終了指示が来るまで繰り返す。各ジョブは、いわゆる処理プログラムであるジョブアプリケーションとその制御を行うジョブコンテナから構成され、作業実行管理機能や各種アクセスサーバーとの複雑なインターフェースはジョブコンテナ側で行いジョブアプリケーション側からは隠蔽されている。これにより変更の多いジョブアプリケーションの開発負荷を軽減した。データ処理の実行モデルを図5に示す。

3.3 並列分散処理の実現

データ解析サブシステムにおいて実現した並列分散処理は、計算負荷が非常に高いFFTやリトリールアルゴリズムを複数の計算機で同時に処理を分散して行うことで処理時間を短縮すると同時に、将来的に処理能力の強化が必要となった場合に計算機の増加のみで対応可能とする。

並列分散処理は、図6に示すように分散処理制御機構がジョブコンテナのクライアントとなり、複数ホスト上で動作する分散処理プログラムのコントロールを行うことで前述の業務の枠組みの中で並列分散処理を実現している。図中の分散処理プログラムがFFT及びリトリールアルゴリズムに対応する。

FFT分散では、図7に示すようにFFT処理を行うレベル0 B処理とレベル1 B処理をFFT処理の開始、終了を管理するフロント、エンド処理とパイプで結合したFFT処理ユニットを分散の一単位としている。即ちホスト1からホストnには全く同じ処理を行う処理ユニットが同時に実行されることになる。この方式ではFFTユニット一つで完結した処理が可能であり、処理の複数実行により処理性能

向上を図る方式、いわばn倍処理方式である。

一方、リトリール分散では、逆の発想により処理性能向上を図っている。即ち、図8に示すように各ホスト上の処理プログラムは、全体の処理の限られた部分部分を実行し、できあがった部分毎の結果を連結することによって全体の結果を導く方式である。即ち、処理の複数分割により性能向上を図る方式、1/n処理方式といえる。

リトリール分散において、このような並列分散方式を採用した理由は、一台のホストで一まとまりの処理を実行するためには大容量のメモリ、ディスク等のリソースが必要となるが、複数台の分割処理であれば、標準的な構成のホストで十分であり、所定の性能を達成するに当たってのコストパフォーマンスに優れるためである。

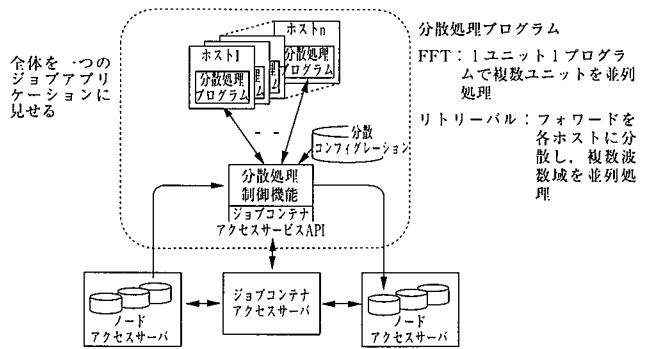


図6 並列分散処理の制御

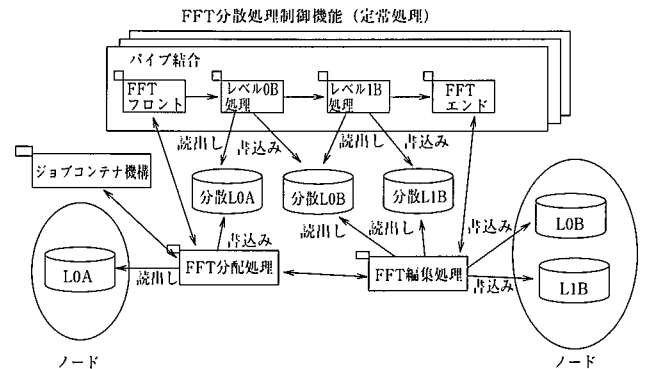


図7 分散処理制御機能

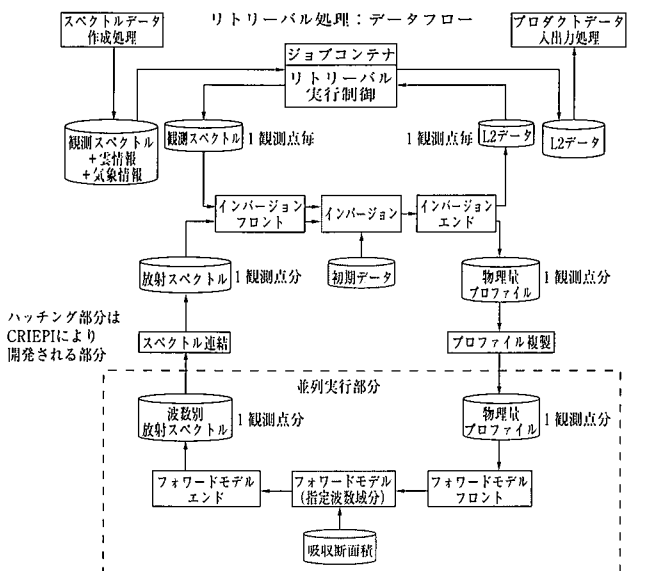


図8 分散処理制御機能

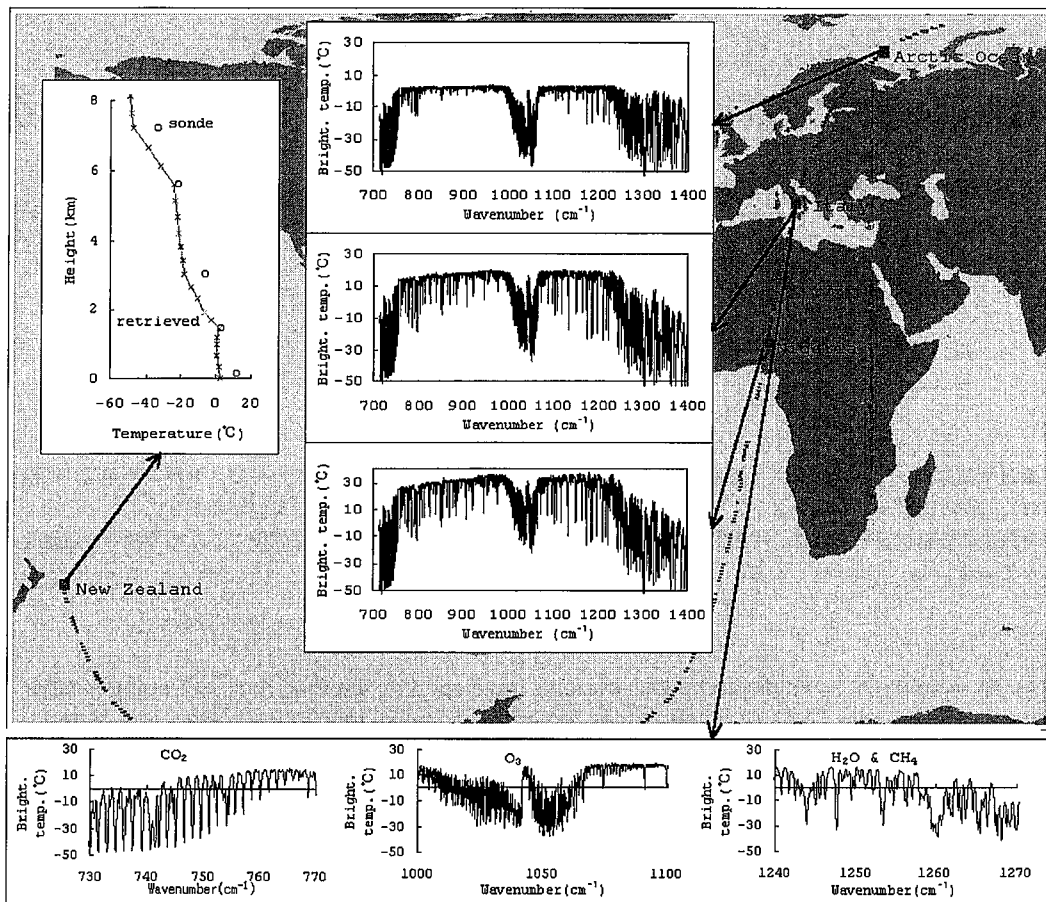


図9 IMGの初観測スペクトルデータ

MITI/NASDA

3.4 ワンマンオペレーションの実現

前述した自動実行機構により、オペレータが行うべき作業は、その日に実行すべき“業務”の集まりに対し、それらの組み合わせと時間枠を設定すること、人手作業であるメディアの挿入・取り替え・ラベリング、及びシステムの状態監視である。これら作業についても支援機能を有しており、運用開始時点からワンマンオペレーションが実現できた。

4. システム運用状況

IMGDISは、1996年8月のADEOS打ち上げ以降順調に稼働している。現在までに約40万点の観測データを作成し、NASDAへのデータ配布をはじめとし、11か国27機関のADEOS特定研究者42名に対してプロダクトの試験配布を行っている。

データ処理アルゴリズムは、打ち上げ以降地上試験に基づいた補正・校正係数等と宇宙での動作環境との差異を各種修正し、精度向上が図られている。現在評価・検証段階の半ばであるが、既にスペクトル作成、リトリーバルアルゴリズム共運用開始以降3～4回のバージョンアップを行い、地上観測データとの比較検証が着々と進んでいる。

図9は、1996年10月10日、日本時間の午後7時1分から午後8時

3分に観測されたIMGの初観測データである。ヨーロッパとアフリカ上空を通り南太平洋に至る軌道上で地球を観測し、それを解析した結果を示す。図中央にスペクトルの輝度温度を示すが緯度の差に応じた温度差が検出されている。下のグラフは、代表的な気体である二酸化炭素、オゾン、水蒸気及びメタンの吸収線の見える波数領域を拡大表示したもので、温室効果気体の濃度変化の大きさはこの領域の変化の大きさに対応している。また、図中左側のグラフは、リトリーバル処理により得られた温度の鉛直分布とゾンデによる観測結果とを比較したものである。

今後、評価・検証を完了したアルゴリズムに基づき観測データを再処理し、高精度のデータセットを作成・保存し、世界の研究者を含むユーザーに配布する予定である。

5. 結 言

高速・大容量のデータ処理システムであるIMGDISを開発した。IMGDISは、C/Sモデルをベースに高速処理を具現化する並列分散処理、自動実行機構等のシステム的新技術を導入し、ワンマンオペレーションを実現すると共に柔軟性・拡張性に富んだシステムの構築を可能とした。