

## 知的分散データシステム構築に向けて - 動的データ配置

## Dynamic Data Allocation toward the Intelligent Distributed Data System

# 堀内 一敏[1], 前田 健[2], 大谷津 裕[3], Yingshuo Shen[4], 早稻田 卓爾[5]

# Kazutoshi Horiuchi[1], Takeshi Maeda[2], Yutaka Oyatsu[3], Yingshuo Shen[4], Takuji Waseda[4]

[1] 地球機構, [2] 地球フロンティア, [3] 三菱総研, [4] ハワイ大、IPRC, [5] 地球フロンティア、IPRC  
[1] AESTO, [2] FRSGC, [3] MRI, [4] IPRC

地球環境問題解決に向けて貴重な研究材料を提供する観測研究では、国際協力の下、データ管理・提供が体制的に充実するとともに、WWW、CDROM、DVD等のデータ流通媒体の開発や技術的進展などにより、データ流通基盤が整備されてきた。しかしながら、データ管理・提供が個々のプロジェクト毎に行われているため、地球環境研究者は、研究に必要なデータという視点から、相変わらず、多々ある観測研究プロジェクトのデータ成果物の中からデータを自分で検索し取得したのちに、自分が利用しやすいように編集し直すといった作業が積みまとう状況にある。

また、計算機性能の著しい向上により、より高解像度のモデルの実行が可能になるにつれ、新たな計算スキームの研究開発や高解像度モデルを利用した4次元データ同化も可能になってきたが、高解像度モデルからの膨大な出力結果を解析する際に支障がでるといった新たな問題が発生するようになった。具体的には、現在のハードウェア技術の下では大容量データを容易に蓄積・転送できないことなどである。

このような背景のもと、我々は、以下の要件を満足する「知的分散データシステム」を構築中である。

- (1) 多種多様な現場観測データをプロジェクト横断的に検索でき、かつ、統一的な形式で効率よく取得できること。
- (2) 衛星データ、海洋・大気シミュレーション結果等の膨大な格子点データを、ネットワーク技術を駆使し、効率よく管理し、かつ、効率よく利用できること。
- (3) 可視化・解析プログラムと現場観測データ及び格子点データの管理・提供を直接に連携させて、データの取得・編集の手間を軽減できること。
- (4) 可視化・解析プログラムにおいて、観測データ、格子点データを同時に扱えるようにすること。

ここでは、知的分散データシステムの設計と並行して進めている要素技術検討として、欧米で普及しつつあるDODS(Distributed Ocean Data System)、AS(DODS Aggregation Server)、LAS(Live Access Server)、EPICの機能・性能検証について述べる。また、より効率よく、分散しているデータをアクセスするために検討しているデータキャッシュのネットワーク上における配置方法についても述べる。DODSは、ネットワーク環境下で、どこからでもどのような形式のデータでも取得できるようにするソフトウェアである。ASは、DODSと連携して物理的に分散しているデータを論理的に結合し、ひとつの大きなデータとして扱えるようにするソフトウェアである。LASおよびEPICは、それぞれ、現場観測データ、格子点データの可視化・解析を行うソフトウェアである。