

# 最新情報技術を活用した超大規模天文データ解析機構の研究開発

大石 雅寿<sup>†</sup> 水本 好彦<sup>†</sup> 白崎 裕治<sup>†</sup> 大江 将史<sup>†</sup> 高田 唯史<sup>†</sup> 田中 昌宏<sup>†</sup> 安田 直樹<sup>‡</sup>

<sup>†</sup>国立天文台, {masatoshi.ohishi, mizumoto.y, yuji.shirasaki, masafumi.oe, tadafumi.takata, masahiro.tanaka}@nao.ac.jp, <sup>‡</sup>東京大学 宇宙線研究所, yasuda@icrr.u-tokyo.ac.jp

## 1. 概要

天文学では、ペタバイトクラスの規模であらゆる波長の観測データがデータベースとして格納されて研究に活用されている。

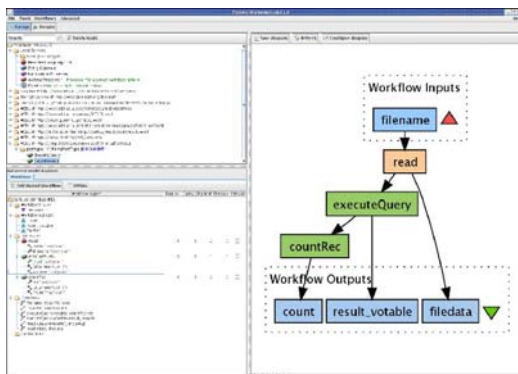
しかしこれらのデータベースは、波長や天文台ごとに独立している。宇宙の諸現象を理解するためには、多波長のデータを有機的に連携することが重要である。高速ネットワークを介したヴァーチャル天文台（天文データグリッド）を構築することで、世界各地の天文データベースやデータ解析サーバ等を相互利用し、これまでは困難であった大規模データを用いた天文学研究が可能になる。

平成 19 年度は、これまでの研究成果を踏まえ、ヴァーチャル天文台におけるデータ検索の効率を画期的に高速化する手法の開発に成功すると共に、検索データを解析するためのワークフローを誰でも利用できるようにするためのワークフロービルダ プロトタイプ の構築を行った。

## 2. 手法・アプローチ

少なくとも 10 億天体を対象とする高速検索については、レコード数が多いデータベースを効率的に検索するための手法として、Table Partitioning を用いた。天文検索では天球座標による検索が基本であることから、天球座標による Table Partitioning を行った。天球座標のインデクス化の手法として、HTM (Hierarchical Triangular Mesh) を用いた。本手法を Java で実装した。

また、天文学者が容易にワークフロー(WF)を利用できるようにするため、bioinformatics 用に開発されたワークベンチ Taverna を利用した WF ビルダ プロトタイプ を構築した。その例を図に示す。



## 3. 議論・考察

我々の手法による高速検索性能を測定した結果を表に示す。半径 3 度という広い検索範囲でも 1 秒以下で検索できることがわかった。さらに、PostgreSQL のパーティショニング機能を用いた場合と比較した結果、最大 150 倍高速であるという結果を得た。

検索半径 分角	天体数 個	経過時間(秒)			HTM 条件数	
		Postgre SQL	独自方式	比	Postgre SQL	独自方式
1	2	6.46	0.04	154	32	32
10	165	3.81	0.03	127	16	16
60	6697	6.47	0.11	60	32	32
100	26720	2.02	0.31	7	4	16
180	57246	9.04	0.71	13	48	72

このように、我々の手法は大規模な天文データベースにおいても十分な性能を持つことがわかった。

WF ビルダは、WFDL を知らなくても WF を実行でき、また、実行状況を視認できるため、遠隔地の解析サービス実行状況を用いるために解析プログラムをウェブサービス化しておけば利用価値が高い。さらに、同様の仕組みを他分野が利用することも可能であると考えられる。

## 4. 結論

国立天文台では、2008 年 3 月にデータサービスの運用を開始した(<http://jvo.nao.ac.jp/portal/>)。今後は、ワークフローの天文 VO における標準策定に貢献するため、実際の天文学研究のユースケースを設定し、そのユースケースを実行するという前提で遠隔地の計算資源をも利用できるワークフローとして実用性を高めてゆきたい。

### 成果（一部）

- [1] Tanaka, M. et al., “Construction of Multiple-Catalog Database for JVO”, Astronomical Data Analysis Software & Systems XVII (2007).
- [2] Shirasaki Y. et al., “Constructing the Subaru Advanced Data and Analysis Service on the VO”, Astronomical Data Analysis Software & Systems XVI (2007).
- [3] Ohishi, M. et al., “Construction of Virtual Observatories through Global Collaboration and Standardization”, UN/ESA/NASA Workshop on Basic Space Science and the International Heliophysical Year 2007” (2007).