

2007.1.26/ No.159

新年を迎えて

藤井 孝藏 (PLAIN センター長) /
海老沢 研 (PLAIN センター)

明けましておめでとうございます。2007年最初のPLAINセンターニュースをお送りします。今年もPLAINセンターをよろしくお願ひします。

PLAINセンターの業務の大きな柱は、1. JAXA/ISAS のスパコン運営、2. ネットワーク整備、3. 衛星理工学データベースの運用ですが、それぞれについて、簡単に昨年の動きを振り返り、今年の見通しを述べたいと思います。

1. 一昨年にJAXAの情報・計算工学(JEDI)センターが発足したことにより、スーパーコンピュータ(スパコン)の運用担当はPLAINセンターからJAXA/JEDIセンターに変更されました。JAXAのスパコン管理は一元化され(<http://www.isas.jaxa.jp>)、ユーザーにとって使いやすくなったと思います。

また、2008年度に予定されている機器換装を念頭に、次期システムについての議論も始まっています。

2. JAXA全体のネットワーク整備の一環として、相模原のネットワーク整備を進めてきましたが、4月からより強固でユーザーのニーズに合った新ネットワークシステムを運用予定です。また、PLAINセンターはスーパーSINETのノードとして活動してきましたが、スーパーSINETがSINET3に徐々に移行していくに伴い、さらに高速になったネットワークを生かして、国内の宇宙科学研究拠点に計算機資源の提供をしていく予定です。

3. 昨年は、JAXA/ISASの「あかり」、「ひので」という大型天文衛星が打ち上げられ、記者発

表等で報じられているとおり、続々とすばらしい科学的成果が得られています。PLAINセンターでもこれらの衛星のデータ処理に協力しており、また、いずれは無償で世界中に公開されるデータを、PLAINセンターのデータアーカイブス「DARTS」(<http://darts.isas.jaxa.jp>)に保管する準備を着々と行っています。一昨年に上げられた「すざく」衛星のデータ処理は順調に進んでおり、国内外の研究者がDARTSからデータを取得し、すでに多くの論文発表を行っています。今年の夏には「すざく」アーカイブスが公開される予定です。いろいろな衛星のDARTSと工学データベース「EDISON」を統一的に開発するミドルウェアもほぼ完成し、PLAINセンターが中心となって、複数の衛星のデータ処理、アーカイブスを担当する枠組みも整いつつあります。今後も、各衛星プロジェクトと協力し、JAXA/ISASの科学衛星から最大限の成果が引き出せるようなシステムを構築、運営していく予定です。

PLAINの人事関連では、昨年10月から松崎恵一さんが助手から助教授に昇進したことに伴い、4月からは助手(助教に呼称が変わる予定)として新たなスタッフを迎える予定です。また、2007年度は、5年毎のJAXA中期計画の最終年度になり、2008年度からの次期中期計画に向けての議論がすでに始まっています。日本と世界の宇宙科学の発展のために、PLAINセンターがJAXAの中で何が出来るか?十分に議論を尽くして、次期中期計画をまとめていくつもりです。今後とも皆様のご支援をよろしくお願ひいたします。

「あかり」のデータプロセッシングの現状

山村 一誠 (赤外・サブミリ波天文学研究系)

1. 「あかり」の現状

赤外線天文衛星「あかり」(ASTRO-F)は、2006年2月22日にM-V-8号機で打ち上げられ、4月13日のファーストライト以来、現在まで順調に観測を続けている。「あかり」は、「太陽同期極軌道」を周回しており、約半年間かけて全天を見渡して観測ようになっている。「あかり」の主要任務には、全天のサーベイ観測を行い、赤外線天体の「住所録」を作成することがあるが、本観測を開始した5月から最初の半年間に、月の影響で観測出来ない部分などを除き、約7割の空についてデータを取得している。現在行われている二周目の観測で、より完全なデータを得ることになる。また、「あかり」は特定の天体・領域に望遠鏡を向けて詳細な観測を行う指向観測を行っている。これは二周目の観測が始まった11月頃より本格的に運用に入り、現在最大で一日に18回程度行っている。

全天サーベイを含む、「あかり」のすべての観測装置を用いた運用は、装置を冷却している液体ヘリウムがなくなるまで続けられる。今のところ短くても今年5月末までは観測可能だと考えられている。その後、「あかり」は機械式冷凍機と近赤外線観測装置による観測を続けていくことになっている。

2. 「あかり」の観測データ

「あかり」には、「遠赤外線サーベイヤ」(FIS)と「近・中間赤外線カメラ」(IRC)の二台の観測装置が搭載され、2ミクロンから180ミクロンまでの広い波長域で観測を行える。両者は同時に動作するが、見ている視野や観測の方法、データのフォーマットが異なっているため、処理は独立に行われる。

観測プログラムは提案方法によって、(A)大規模サーベイ(LS:プロジェクトが遂行主体となる。全天サーベイも含む)、(B)ミッションプログラム(MP:プロジェクト関係者の有志が構成する「勉強会」が主体となって提案する)、(C)日本・韓国・ESA関係国からの一般公募観測(OT)、の三種類がある。個々のプログラムで、数十から数百の指向観測を行っている。

全天サーベイ観測のデータは、プロジェクトの責任で処理が行われ、その成果物である赤外線天体カタログなどが提供される。カタログの一般への公開は、サーベイ観測終了の二年後を目標としている。一方、指向観測データは、プロジェクトとの密接な

協力の下に、各観測プログラムの遂行主体の責任で処理・解析が行われる。指向観測データの観測者への配布は、2007年1月に開始される。

すべての配布データには、一年間の優先利用期間が設定され、その間は観測提案者のみが利用出来る。その後、データは世界中に公開される。

3. 「あかり」観測データファイルの作成

観測データの処理は、テレメトリデータがSIRIUSデータベースに登録されてから始まる。観測装置や、姿勢系など、複数の機器からのテレメトリデータを編集し、天文データのフォーマットとして標準的なFITS形式に変換する。

FISに関しては、このデータ作成のシステム構築と定常運用を、PLAINセンターと協力してPLAINセンターの提供するデータ処理基盤DANS上で行っている。作成されたデータは、ISAS-SAN(SNFS)/NFS経由でプロジェクト側のネットワークにある計算機がアクセスし、キャリブレーションなどの処理が行われる。この部分は、大量のメモリとCPU能力を必要とし、また処理プログラムが日々更新されるため、プロジェクト側で管理している。FISのデータは、全天サーベイ観測と指向観測のデータが、ひとつながりに連続して含まれている。我々は、データベースによって時系列データファイルを管理することで、見かけ上区切りのないファイルアクセスを可能にしており、指向観測のデータは、別途もうけている観測計画のデータベースからの情報をもとに、切り出して処理を行っている。一方、IRCについてはデータの変換も含めてプロジェクト側の計算機で行っており、DANS上ではテレメトリデータの切り出しと管理のみを行っている。これらのプロセスの技術的詳細については、別の機会に解説したい。

現在行われている処理プロセスやキャリブレーションは、今後さらに詳細なデータの解析が進むにつれて改訂されていく。指向観測のユーザーには、処理ソフトウェアとキャリブレーションパラメータの更新によって対応するとともに、適当な機会に全データの再処理を行うことを予定している。

4. 「あかり」データの配布

「あかり」観測データ、あるいは処理によって得られた成果(天体カタログなど)のユーザーへの配布は、PLAINセンターが運営するDARTSを唯一

の窓口として行う。ヨーロッパの公募観測のデータのみ、ユーザーへの便宜のため ESAC (European Space Astronomy Centre: ESA の天文ミッションのサポートセンター) にもデータを置くが、すべてのデータを保持・管理し、将来の一般公開に備えるのは、DARTS のみである。

プロジェクト側の計算機で処理が行われた配布データは、DANS 上に再転送され、ファイルの内容などを解説した README ファイルを作成・添付した後に、tar でパッケージされる。さらに、「すざく」の方法を踏襲して、gpg (GnuPGP) で暗号化と圧縮が行われる。暗号鍵はファイル毎ではなく、観測プログラム (複数の指向観測を含む) 毎に設定され、PI に送付される。配布データの作成と暗号鍵も

含めた管理は、PLAIN センターの DARTS チームにゆだねられている。

この原稿を執筆している時点で、FIS, IRC ともにプロジェクト側でのデータ作成と処理の最終段階にあり、1 月中の配布開始を目指している。また、これと平行して、データ処理用のソフトウェアツール、マニュアルの整備も ESAC と協力して進めている。「あかり」の真の意味での成果、すなわち科学的研究の成果が出るのは、データがユーザの手に渡り、解析が始まるこれからである。歴史に残るような素晴らしい成果が上がるように、今後も PLAIN センターの協力を得ながら、努力を続けていく所存である。

情報通信技術を宇宙科学にどう活用するか? (第 5 回)

村田 健史 (愛媛大学総合情報メディアセンター助教授
・宇宙科学情報解析センター客員助教授)

3.3 STARS1 & STARS2 : オブジェクト指向とオブジェクト指向開発技法 (その 2 ~データフォーマットについての考察) のつづき

前回、太陽系プラズマ観測データには、自分自身のデータフォーマットを表現するメタ情報を内包する自己記述型のデータ形式が有効であることを述べた。筆者は、愛媛大学総合情報メディアセンターの木村映善助手 (現在は同大学医学部助教授) や私の研究室の修士課程の学生諸氏 (山本和憲君、石倉聡君、久保卓也君など) と一緒に NASA/GSFC が提唱する自己記述形式 CDF (Common Data Format) について勉強し、CDF が今後の STP 分野の科学衛星観測データには最も有効な自己記述型データフォーマットであると確信している。

前述の SDB 形式で記述されていた GEOTAIL 衛星および AKEBONO 衛星観測データの CDF 形式への変換作業である。変換作業は、篠原助教授と久保君が中心となり各 SI の協力の元で進めており、必ずしも順調とはいえないが、現在も続いている。

2005 年に打ち上げられた REIMEI 衛星については、打ち上げ前から SI と相談し、全ての Level-2 データを CDF 形式で記述することを申し合わせた。各 SI データの CDF 設計を打ち上げ前からはじめたおかげで、表に示すように打ち上げ後 1 年ですべての REIMEI データを CDF 化することができた。(この作業は、宇宙研の笠羽助教授には特にご尽力をいただいた。) GEOTAIL や AKEBONO 衛星の作業が順調には進んでいないことと比較すると、打ち上げ前からデータフォーマットを定義することが L2 データベース構築には効率的であることがこの表から明らかである。この教訓を生かし、我々は SELENE、Planet-C、BepiColombo といった打ち上げが予定されているプラズマ観測衛星のデータ形式についての議論をすでに始めている。

なお、久保君はすでに 10 種類以上の太陽系プラズマ観測データの CDF 化に関わっており、今や CDF のスペシャリストである。彼は優秀であるが、しかし情報工学専攻の学生であり、科学衛星観測データの物理解析の経験はまったくない。データの物理的意味を理解していない彼にデータ構造の設計を任せるのは、やはり重荷であろう。物理知識がある専任の若手研究者の中からデータ構造設計のスペ

[裏へ続く]

表 PLAIN センターが管理する太陽系プラズマ観測データ CDF 化の現状: ○は実装済み、△は実装中、□は CDF の設計中

(2007 年 1 月現在)

AKEBONO 衛星

軌道 (△)、VLF (△)、PWS (□)

GEOTAIL 衛星

軌道 (○)、PW/SFA (□)、PW/WFC (△)、MGF (○)、LEP (□)

REIMEI 衛星

軌道 (○)、EISA (○)、MAC (○)、GAS (○)、CRM (○)

数年前、縁あって PLAIN センターの新 DARTS システムの太陽系プラズマ観測データサイトの設計に関わるようになり、同センターの篠原育助教授や松崎恵一助教授との協力のもとで、DARTS が管理する太陽系プラズマ観測分野の観測データの CDF 化を進めている。表に、これまでに CDF 化を行ったデータの一覧を示す。この中でまず手がけたのが、

シャリストを育成することは、宇宙科学研究本部の責務のひとつである。

このように、太陽系プラズマ観測データの自己記述形式での L2 データベース作成は軌道に乗りつつある。しかしなぜ太陽系プラズマ観測データの CDF 化は、netCDF の気象学、FITS の天文学と比較してこれほど遅れたのであろうか？ひとつの理由は、気象学や天文学と比較すると太陽系プラズマ観測データはデータ構造が複雑であり、また観測モードも多岐にわたるからである。天文学は原則的には 2 次元の天空撮像データが主体となるため、画像形式を主対象とする FITS フォーマットで対応できる。一方、地上観測である気象学はテレメトリと言う衛星独自の制約がないため観測データ量を節約する必要がなく、衛星観測のように（最適な観測データを取得するための）複雑なモード設定が不要である。太陽系プラズマ観測は、観測機器が多く、また機器ごとの観測モードも多い。一つのデータ形式であらゆる観測データに対応する設計には年月を要したのである。NASA/GSFC の CDF に関わった技術者には頭が下がる思いがする。

もうひとつ、忘れてはならないことがある。それは、「コスト」である。残念ながら太陽系プラズマ研究グループには、これまでコストと言う概念が不足していたように思われる。SI が独自にデータ構造を記述する SDB と“あてがいくち”でデータを記述しなくてはならない CDF では、どちらが SI にとって楽なデータフォーマットであろう？もちろん SDB である。自由にデータ構造を記述できるのであるから、SI にとってこれほど都合がよいデータ構造はない。私の経験から言っても、CDF という

フレームワークの中で自分の観測データを記述するのは、正直なところ、各 SI にとっては楽な作業ではない。しかし、その努力を惜しんではならない。なぜなら、一度標準的なデータフォーマットで記述してしまえば、その後の作業は情報処理の専門家にバトンタッチできるからである。近年の IT または ICT（情報通信技術）の普及により、システムエンジニアやプログラマの数は増え、技術力のあるプログラマに比較的安価でシステム構築を依頼することもできるようになってきた。これは広い目で見れば、システム構築の大きなコストダウンである。

わたしは、これまで多くの太陽系プラズマ研究者が「情報処理技術者（たとえばプログラマ）にデータを渡しても、物理的側面の知識がないプログラマがデータ解析ツールを作れるはずがない」という言葉を聞いてきた。しかし、それは間違いである。重要なのは、客観的にデータ構造を表現することである。物理的側面を消してデータの構造を客観的に記述できれば、優秀なプログラマであればデータを処理したりプロットしたりするツールの実装は困難ではない。では、誰がそのデータ表現をするのか。それは、データの物理的側面を理解している研究者の仕事である。（だからこそ、物理知識を持つデータ構造設計のスペシャリストが必要なのである。）現在、篠原助教授と松崎助教授が中心となり実装が進んでいる新 DARTS-STP では、その客観的設計があらゆるところで行われている。その結果、すばらしいシステムが、プログラマやシステムエンジニアとのコラボレーションによって構築されつつある。期待していただきたい。（次号へ続く）

宇宙研計算機、ネットワークに関するお知らせ

三浦 昭 (PLAIN センター)

●解析サーバ、相模原ネット関連

利用案内、申請方法：

解析サーバ

http://plain.isas.jaxa.jp/ana_servers/

ネットワーク利用

<http://www.pub.isas.jaxa.jp/>（相模原ネット内限定）

申請受付：計算機室 山本 (RN.2103, 内線 8388)

下記の各申請を受け付けています。

- ・ ISAS ドメインメールサービス
- ・ 解析サーバ (ISAS 内)

・相模原ネット接続等

計算機等利用上の質問・トラブルなどはシステム・プログラム相談室 (RN 2113・内線 8391) 迄、ネットワーク関係の質問・トラブルなどは PLAIN センター本田秀之 (RN 1261・内線 8073)、長木明成 (RN 2101・内線 8386) 迄お願いします。

●スーパーコンピュータ

利用案内、申請方法：

<https://www.jss.jaxa.jp/>

編集発行：宇宙航空研究開発機構 宇宙科学研究本部 宇宙科学情報解析センター (PLAIN センター)

〒229-8510 相模原市由野台 3-1-1 Tel.042-759-8351 住所変更等 e-mail : news@plain.isas.jaxa.jp

本ニュースはインターネットでもご覧になれます。 <http://www.isas.jaxa.jp/docs/PLAINnews>

●編集後記：冬は鍋です。Event PLAIN あんこう鍋ツアー計画は大洗→水戸→築地と変更になりましたが、食べ放題、飲み放題で盛況でした。(K.E.)