

時系列データダイナミックレビュー用 Web アプリケーション紹介 STARS touch: A web-application for time-dependent observation data

村田 健史^{1*}; 渡邊 英伸¹
MURATA, Ken T.^{1*}; WATANABE, Hidenobu¹

¹ 情報通信研究機構

¹National Institute of Information and Communications Technology

現在、多くの科学研究分野ではデータのほとんどがデジタル化され、その量および種類は大規模化の一途をたどっている。これからますます大規模化・複雑化するデータ指向型科学時代を踏まえて、ビッグデータ処理がより容易に、また一元的に行うことができるクラウドシステムが求められている。

NICTサイエンスクラウドは、地球惑星科学を含む様々な科学研究データおよびソーシャルデータのためのクラウドシステムである。NICTサイエンスクラウドでは(1)データ伝送・データ収集機能、(2)データ保存・データ管理機能、(3)データ処理・データ可視化機能の3つの柱(機能)から構成されている。それぞれの機能についての基盤技術を開発するだけでなく、複数の基盤技術を組み合わせることでシステム化を行うことができる。システムを実際に科学研究に応用・適用することで、様々な分野でのビッグデータ科学・データインテンシブ科学が可能となる。

本研究では、NICTサイエンスクラウド上で開発した時系列データ表示ツール(開発名:STARS touch)について紹介する。これまでの多くの時系列データ表示用科学データ Web アプリケーションは、Web アプリ用のミドルウェアなどによりデータの読み込みと画像表示を行ってきた。その多くは、日時やデータ選択を行う手間やデータ処理を行う処理時間がユーザビリティを下げていた。STARS touchはクラウド上のデータ収集システム(NICTY/DLA および WONM システム)により収集した科学データを Gfarm/Pwrake 等により並列処理することで画像化した時系列画像データを用いる。また、Ajax やキャッシュプログラムにより閲覧しているデータに近いデータを優先的に読み込む非同期処理を導入することでユーザビリティを上げている。発表では、STARS touch のデモを行うと同時に STARS touch のバックエンドの技術を紹介する。

キーワード: Web アプリケーション, WDS, 時系列データ, 集合知, 衛星観測, 気象観測

Keywords: Web application, WDS, time-dependent data, Collective intelligence, spacecraft observation, Meteorological observation

NICTサイエンスクラウドの広域分散型ストレージにおけるセキュリティの取り組み

Security approaches of a distributed storage system in the NICT Science Cloud

渡邊 英伸^{1*}; 鈴木 豊²; 村永 和哉²; 鶴川 健太郎²; 村田 健史¹

WATANABE, Hidenobu^{1*}; SUZUKI, Yutaka²; MURANAGA, Kazuya²; UKAWA, Kentaro²; MURATA, Ken T.¹

¹ 独立行政法人 情報通信研究機構, ² 株式会社セック

¹National Institute of Information and Communications Technology, ²Systems Engineering Consultants Co., LTD.

近年、科学分野においてオープン化やグローバル引用・参照化に関する国際的な取り組みが進められており、科学データがインターネット上で誰でも参照可能なサービスまで提供されてきている。膨大な科学データが扱える中、クラウドコンピューティング技術によって新しい知見獲得や分野横断的研究を発掘するデータ指向科学が期待されており、地球惑星情報学でも議論が活発である。

一方、データ指向科学の成功事例は非常に少なく、その理由のひとつに、多くの科学研究者が巨大なデータ量を安心かつ安全に保管し、そのデータにいつでもアクセスできる無償のストレージや大量の高性能計算機をいつでも無償で占有して利用できる処理環境が少ないことが挙げられる。特に、科学研究者がクラウドストレージサービスを利用しない一番の理由は、情報漏えいなどセキュリティに不安を感じている点であり、文部科学省平成25年度国家課題対応型研究開発推進事業「コミュニティで紡ぐ次世代大学 ICT 環境としてのアカデミッククラウド」で実施されたアンケート調査結果でも示されている。

情報通信研究機構 (NICT) は、観測データやシミュレーションデータなど、あらゆる科学データを収集・蓄積すると同時に解析環境も提供するデータ指向科学研究向けのクラウドシステム (NICT サイエンスクラウド) を構築している。NICT サイエンスクラウドは、国内5地区 (東京, 名古屋, 京都, 大阪, 沖縄) にあるデータセンターに分散配置した計算機を10Gbps の L2 ネットワーク (JGN-X) で接続し、オープンソースの広域分散ファイルシステムである Gfarm を用いて約500 の CPU コア、約3PB のストレージを有した広域分散型のコンピューティング環境を提供しており、利用者は無償で利用できる。

本発表では、Gfarm を用いた NICT サイエンスクラウドの広域分散型ストレージにおけるセキュリティの取り組みを紹介する。特に、ある時刻においてそのデータは本物であることをいつでも検証できるアプリケーションやデータの操作イベント (作成、更新、削除など) をいつでも追跡できるシステムについて報告する。加えて、個々の分野を越えたデータ指向科学に貢献する事項について議論する。

キーワード: データ指向科学, クラウドサービス, 広域分散型ストレージシステム, 情報セキュリティ, オープンデータ

Keywords: data-intensive science, cloud service, distributed storage system, security, open data

気象観測報の表参照通報式 BUFR・CREX への移行 Migration to BUFR/CREX of meteorological observation reports

豊田 英司^{1*}
TOYODA, Eizi^{1*}

¹ 気象庁数値予報課

¹Numerical Prediction Division, Japan Meteorological Agency

世界気象機関（WMO）では世界天気監視計画の枠組みで交換される気象観測報を伝統的の文字形式通報式（TAC、電報に由来する英数字列でデータを表現する方式）から表参照通報式（TDCF）に移行しようとしている。主要なデータ種別について2014年11月が移行期限とされ、実際に文字形式の配信を停止する国が現れ始めた。

表参照通報式にはバイナリのBUFRとテキスト環境での代替表現CREXの2種類がある（他に格子点データ用のGRIBがある）。どちらもメッセージ先頭でデータの構造を要素記述子（番号）の列で表現する。それぞれは記述子は、単位付数値、外部符号表を参照する列挙型またはフラグ、あるいはASCII文字列を表現する。

記述子列は一部を繰り返すことができるので、XMLのような半構造データ、つまり同種のデータの集まりに全体についての補足情報を付加したものを一括して表現できる。

新種のデータ構造もメッセージを見るだけで読み解いていける点では自己記述的だが、記述子定義はメッセージ外の表を参照する必要がある。XMLと比べてバイナリ、番号ベース、外部表の参照といった特徴は、開発の始まった1990年代に通信帯域が高価だったためメッセージ長を短くするため可読性を犠牲にした、としばしば説明されるが、現業的システムの安定運用のため自由度を減らすという効用もある。

表参照通報式への移行は衛星観測データについてはほぼ完了しているが、地上観測点ベースのデータでは世界各国が対応しなければならないことから時間を要している。特に高層気象観測では、移行の機に通報内容を充実させようとした一方で、TACから変換して作られたBUFRが流通しており、利用には注意が必要である。

キーワード: 世界気象機関, 世界天気監視計画, 高層気象観測, 表参照通報式

Keywords: WMO, World Weather Watch Programme, Upper-air observation, Table-driven code form, BUFR, TEMP

日本シームレス地質図(地理院地図版, 仮称)の開発と公開 Development and publishing of the Seamless Geological Map of Japan (Chiriinchizu edition, tentative name)

西岡 芳晴^{1*}; 長津 樹理¹; 北尾 馨²
NISHIOKA, Yoshiharu^{1*}; NAGATSU, Juri¹; KITAO, Kaoru²

¹ 産業技術総合研究所地質情報研究部門, ² キューブワークス
¹Institute of Geology and Geoinformation, AIST, ²CubeWorks Inc.

コードを全面的に書き換えて、20万分の日本シームレス地質図の次期ビューアを開発した。このビューアは、ベースマップに地理院地図を採用している。このビューアは、さらに、このビューアは独自に開発したオフラインで地図を利用できる仕組みであるポータブルマップ等の最先端技術を随所に適用している。また、地図描画部分を分離しており今後の3D化等への柔軟性も有している。我々はこのビューアを今年の5月に公開する予定である。

キーワード: シームレス地質図, Leaflet.js, 地理院タイル, ポータブルマップ, Web サイト
Keywords: Seamless Geological Map, Leaflet.js, Chiriin tile, portable map, Web site

サイエンスクラウドを活用したプラズマ波形データの並列分散処理 Parallel Distributed Processing of Plasma Waveform Using Science Cloud

笠原 禎也^{1*}; 矢木 大介¹; 村田 健史²; 後藤 由貴¹
KASAHARA, Yoshiya^{1*}; YAGI, Daisuke¹; MURATA, Ken T.²; GOTO, Yoshitaka¹

¹ 金沢大学, ² 情報通信研究機構

¹Kanazawa University, ²National Institute of Information and Communications Technology

科学衛星上でのプラズマ波動の波形観測は、宇宙プラズマ中の物理過程を理解するうえで非常に重要とされるが、波形データはデータ容量が大きく、その中から特徴的な波動を人手で抽出・解析することが大変困難である。この問題に対し、我々は計算機の手でシステムティックに様々な特徴的な波動を自動抽出するアルゴリズムの開発に長年取り組んできたが、同処理は雑音除去や特定周波数帯域の信号抽出など、検出精度の向上にむけたアルゴリズムの改良と試行に、非常に多くのターンアラウンドを要する。今回、我々は NICT サイエンスクラウドに代表される大規模な並列計算機を活用し、ヘテロタイプの時系列データ処理に適したタスクスケジューリング法を利用することで、並列化のために既存解析ソフトに大幅な修正を加えることなく、解析効率の向上を実現した。本発表ではその成果を報告するとともに、今後、同クラウドの活用によって期待される新たなサイエンスアウトプットについて議論する。

キーワード: 並列分散処理, サイエンスクラウド, プラズマ波動, 信号処理, 波形解析

Keywords: Parallel Distributed Processing, Science Cloud, Plasma Wave, Signal processing, Waveform analysis