Japan Geoscience Union Meeting 2014
(28 April - 02 May 2014 at Pacifico YOKOHAMA, Kanagawa, Japan)

MGI37-01 会場:413 時間:4 月 29 日 09:00-09:15

NICT サイエンスクラウドのビッグデータ処理技術開発と運用
A Report of Big-Data Processing and Operation of the NICT Science Cloud

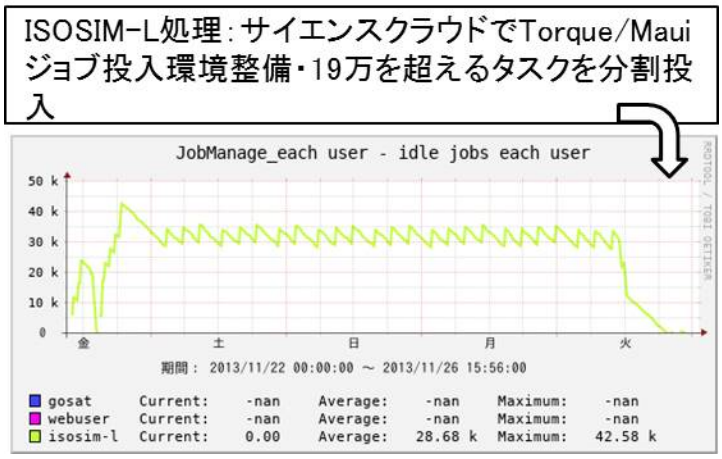
村永 和哉[1*] ; 鵜川 健太郎[1] ; 鈴木 豊[1] ; 村田 健史[2] ; 渡邉 英伸[2] ; 水原 隆道[3] ; 建部 修見[4] ; 田中 昌宏[4] ; 木村 映善[5]
MURANAGA, Kazuya[1*] ; UKAWA, Kentaro[1] ; YUTAKA, Suzuki[1] ; MURATA, Ken T.[2] ; WATANABE, Hidenobu[2] ; MIZUHARA, Takamichi[3] ; TATEBE, Osamu[4] ; TANAKA, Masahiro[4] ; KIMURA, Eizen[5]

[1] 株式会社 セック, [2] 情報通信研究機構, [3] 株式会社　クレアリンクテクノロジー, [4] 筑波大学, [5] 愛媛大学
[1]Systems Engineering Consultants Co., LTD., [2]National Institute of Information and Communications Technology, [3]CLEALINKTECHNO Co.,Ltd., [4]University of Tsukuba, [5]Ehime University

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・データ取集機能、（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はなく、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。

科学研究クラウドシステムは、単に計算機リソースやストレージを接続するだけでは機能しない。様々な科学研究で利活用できる一般的なインフラ整備が必要であるが、インフラは同時に各研究目的にカスタマイズできなくてはならない。NICT サイエンスクラウドは、約 3 年にわたり汎用性と特殊性の両方を満たすクラウドの構築と運用を行ってきた。本発表では、クラウド運用及びそれにかかわる技術を紹介する。また、NICT のリモートセンシング研究を題材として、サイエンスクラウド上で構築した実研究システムについて、ビッグデータ処理の視点から報告する。



Japan Geoscience Union Meeting 2014
(28 April - 02 May 2014 at Pacifico YOKOHAMA, Kanagawa, Japan)

MGI37-02 会場:413 時間:4 月 29 日 09:15-09:30

広域分散ファイルシステム Gfarm と連動した高速ファイル転送ツール
High-speed File Transfer Tool with the Gfarm File System

渡邉 英伸 [1*] ; 黒澤 隆 [2] ; 村田 健史 [1]
WATANABE, Hidenobu[1*] ; KUROSAWA, Takashi[2] ; MURATA, Ken T.[1]

[1] 独立行政法人 情報通信研究機構, [2] 株式会社 日立ソリューションズ東日本
[1]National Institute of Information and Communications Technology, [2]Hitachi Solutions East Japan, Ltd.

巨大なデータ量を扱うハイパフォーマンス・コンピューティング (HPC) は，非常に大きなデータ量を扱う数値シミュレーション等に利用されている．近年，計算クラスタ内に数百から何千ものサーバを含む大規模な処理環境になるまでに至り，大量のストレージリソースが消費されている．さらに，ストレージリソースはエクサバイトオーダ以上のサイズが要求されるまでになり，スケールアウトが可能な広域分散型のストレージシステムが注目を集めている．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 高速バックボーンネットワーク網である 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の Gfarm を用いて約 3PB の広域分散型ストレージシステムを運用している．

HPC 等を想定した広域分散型ストレージシステムは，大容量データに対して高速なデータ I/O とデータ転送が重要となる．Gfarm は，ハイパフォーマンス・コンピューティング・インフラ (HPCI) の共有ストレージに採用されており，高速なデータ I/O を実現することが可能である．一方，データ転送にはインターネットで利用される標準の通信規約である TCP を採用している．TCP は長距離・高遅延のネットワークにおいて伝送遅延の問題が知られており，Gfarm では，TCP マルチストリーミングによってデータ転送の高速化を図っている．しかしながら，ネットワークが長距離・高遅延になればなるほど，高速化の効率が低くなっているのが実情である．

我々は Gfarm と連動する高速ファイル転送ツールを開発した．データ転送の通信プロトコルにオープンソースの通信ライブラリである UDT(UDP-based Data Transfer) プロトコルを採用し，簡易な並列データ転送制御機構を有する．UDT プロトコルは，UDP によるデータのバルク転送と RTT(RoundTrip Time) に依存しない独自のフロー制御や輻輳制御を提供し，長距離・高遅延のネットワークにおいては TCP よりも高速なデータ転送が可能である．本発表では，開発した並列ファイル転送ツールを紹介するとともに，基本的な性能について報告する．

キーワード: 広域分散型ストレージ, Gfarm, 高速ファイル転送, UDT

Japan Geoscience Union Meeting 2014
(28 April - 02 May 2014 at Pacifico YOKOHAMA, Kanagawa, Japan)

MGI37-03 会場:413 時間:4 月 29 日 09:30-09:45

分散ファイルシステムによる並列データ I/O 測定
An Examination of Data I/O Speed on a Parallel Data Storage System

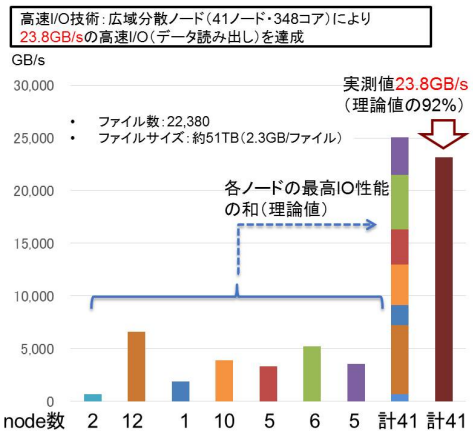
村田 健史 [1*] ; 渡邉 英伸 [1] ; 鵜川 健太郎 [2] ; 村永 和哉 [2] ; 鈴木 豊 [2] ; 建部 修見 [3] ; 田中 昌宏 [3] ; 木村 映善 [4]
MURATA, Ken T.[1*] ; WATANABE, Hidenobu[1] ; UKAWA, Kentaro[2] ; MURANAGA, Kazuya[2] ; YUTAKA, Suzuki[2] ; TATEBE, Osamu[3] ; TANAKA, Masahiro[3] ; KIMURA, Eizen[4]

[1] 情報通信研究機構, [2] 株式会社 セック, [3] 筑波大学, [4] 愛媛大学
[1]National Institute of Information and Communications Technology, [2]Systems Engineering Consultants Co., LTD., [3]University of Tsukuba, [4]Ehime University

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・データ取集機能、（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はなく、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科学研究のビッグデータ処理を行うための基盤技術について議論する。データサイズが大きい場合にクラウドデータ処理で解決すべき問題点の一つはデータ I/O である。例えば、100MB/sec で 100TB のデータを読み出すとすると、1,000,000 秒（約 11.5 日）かかる。すなわち、大規模科学データを処理するためには、高速 I/O 技術が不可欠である。本発表では、並列ファイルシステム（GPFS）と分散ファイルシステム（Gfarm）の 2 つのシステムでのデータ読み出し速度の比較を行い、それらのスケーラビリティーを比較する。



Japan Geoscience Union Meeting 2014
(28 April - 02 May 2014 at Pacifico YOKOHAMA, Kanagawa, Japan)

MGI37-04 会場:413 時間:4 月 29 日 09:45-10:00

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

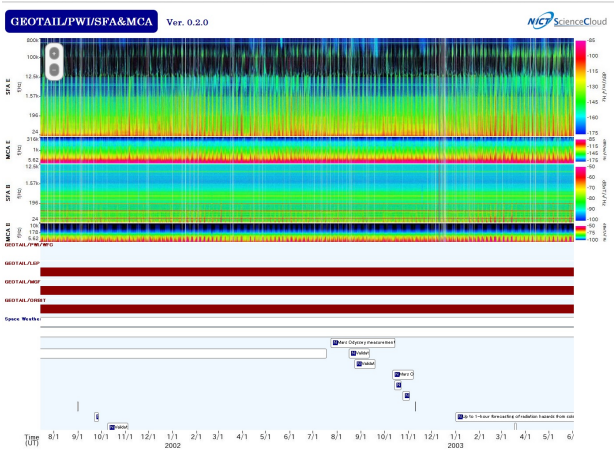
鵜川 健太郎[1*] ; 村永 和哉[1] ; 鈴木 豊[1] ; 村田 健史[2] ; 篠原 育[3] ; 小嶋 浩嗣[4] ; 能勢 正仁[4] ; 渡邉 英伸[2] ; 建部 修見[5] ; 田中 昌宏[5] ; 木村 映善[6]
UKAWA, Kentaro[1*] ; MURANAGA, Kazuya[1] ; YUTAKA, Suzuki[1] ; MURATA, Ken T.[2] ; SHINOHARA, Iku[3] ; KOJIMA, Hirotsugu[4] ; NOSE, Masahito[4] ; WATANABE, Hidenobu[2] ; TATEBE, Osamu[5] ; TANAKA, Masahiro[5] ; KIMURA, Eizen[6]

[1] 株式会社 セック, [2] 情報通信研究機構, [3] 宇宙航空研究開発機構, [4] 京都大学, [5] 筑波大学, [6] 愛媛大学
[1]Systems Engineering Consultants Co., LTD., [2]National Institute of Information and Communications Technology, [3]Japan Aerospace Exploration Agency, [4]Kyoto University, [5]University of Tsukuba, [6]Ehime University

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・データ取集機能、（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はなく、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ーを下げていた。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ーを上げている。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。



Japan Geoscience Union Meeting 2014
(28 April - 02 May 2014 at Pacifico YOKOHAMA, Kanagawa, Japan)

MGI37-05 会場:413 時間:4 月 29 日 10:00-10:15

NICT サイエンスクラウドを用いたかぐや衛星 WFC-L からのバイポーラ型波形抽出処理の高速化
High performance data processing for detection of bipolar waveforms from KAGUYA/WFC-L using the NICT Science Cloud

矢木 大介[1*] ; 村田 健史[2] ; 笠原 禎也[1] ; 後藤 由貴[1]
DAISUKE, Yagi[1*] ; MURATA, Ken T.[2] ; KASAHARA, Yoshiya[1] ; GOTO, Yoshitaka[1]

[1] 金沢大学, [2] 情報通信研究機構
[1]Kanazawa Univ, [2]National Institute of Information and Communications Technology

月探査衛星かぐやは, 2007 年 9 月に打ち上げられ, 2009 年 6 月に月面に制御落下した. かぐや衛星には月周辺のプラズマ波動を観測する波形捕捉器 WFC が搭載されており, 特に 100Hz~100kHz の電界波形を観測する WFC-L では, これまでの研究からいくつかのパターンに分類できる特徴的なバイポーラ型の波形が多数確認されている. しかし, WFC-L は 250kHz のサンプリング周波数で波形データを取得することから, そのデータ総量は約 190GB に達する. 我々は現在, この波形データからバイポーラ型波形を自動抽出するアルゴリズムを開発中であるが, 汎用の PC ワークステーションでは, 全観測データから波形抽出を行うのに 1 週間近い処理時間を要する. これでは, より精度よく波形を抽出し, その分類を行うための処理アルゴリズムを試行錯誤するには, 大変非効率的である. そこで, 情報通信研究機構 (NICT) のサイエンスクラウドを用いて, 処理の高速化を図った. NICT サイエンスクラウドは, 科学研究目的のために構築されたクラウドシステムで, 特にビッグデータサイエンスを主対象の一つにしている. 今回は, NICT サイエンスクラウド上でワークフローシステムを用いた並列分散処理による高速化を行い, その効率について評価した結果を報告する.

実際の並列処理は, Pwrake (Parallel Workflow extension for Rake) と呼ばれるツールを用いて, サーバ群にワークフローを与えることで実現した. Pwrake は, Ruby 言語で記述されるビルドツールである Rake をファイル共有システム向けに拡張したものである. Pwrake 上に処理内容と使用するノード及びコア数を記述することで, コマンドをタスクとして各ノードに割り振り並列処理を行うことが可能である.

結果として, 10 ノード 24 コアの計算リソースを用いた場合, 1 ノード 1 コアでの処理に比べて約 140 分の 1 の時間で処理を終えることが確認できた. 処理速度が使用リソース数に比例していないのはハイパースレッドの影響によると考えられる. 同システムを活用することで, より精度の高い波形抽出アルゴリズムの開発が効率よく行えることが期待できる. 今後は, 波形抽出アルゴリズムの改善に加えて, 更に計算リソースを増やした場合の測定実験についても実施する予定である.

キーワード: 月探査衛星かぐや, 波形捕捉器, NICT サイエンスクラウド, 並列処理
Keywords: Lunar Orbiter KAGUYA, Waveform Capture, NICT Science Cloud, parallel processing

Japan Geoscience Union Meeting 2014
(28 April - 02 May 2014 at Pacifico YOKOHAMA, Kanagawa, Japan)

MGI37-06 会場:413 時間:4 月 29 日 10:15-10:30

気象用格子データ形式の比較：気象庁ではなぜ索引付き順番探査を使うのか
Comparison of grid data formats in meteorology: the reason for indexed sequential access method (ISAM) used in JMA

豊田 英司 [1*]
TOYODA, Eizi[1*]

[1] 気象庁予報部数値予報課
[1]NPD, Japan Meteorological Agency

数値天気予報などのシミュレーションで作成される格子点データにはWMO（世界気象機関）標準の GRIB，OGC 標準の netCDF，気象庁内で使われる NuSDaS など多数の形式がある。Wright and Gao (2008) はデータ形式をファイル構造により順番探査と直接探査に二分し、部分読み込みの高速性とファイルサイズの選択と論じた。しかし気象庁で用いられる NuSDaS（豊田、2001）は索引付き順番探査ファイルであり、高速性とサイズが両立する。これら 3 分類の性能上の長短を比較した（表 1 ）。

現業数値予報でのデータアクセスには、しばしばデータ構造が疎な配列となる、ファイル書き出しは一度に行われて追加が行われない、ファイルの一部を読みだすことが多く性能が問われるという特徴がある。索引付き順番探査ファイルの短所（索引作成）が目立たず、読み書き速度とコンパクトなファイルサイズの両立という利点が享受しやすい。

また、将来データサイズが巨大化してゆくと、単純な順番探査は不利になってゆくことも注意される。

引用文献

Bruce Wright and Feng Gao, 2008: GRIB vs NetCDF: Evaluation of the Technical Aspects. WMO ET-ADRS Doc.2.3(1) http://goo.gl/AFrsls

豊田英司, 2001：気象庁の数値予報ルーチンで用いられているデータセット形式の紹介. 合同大会講演 A2-001 http://goo.gl/JE0a3M

キーワード: GRIB, netCDF, NuSDaS, 格子データ, 索引付き順番探査
Keywords: GRIB, netCDF, NuSDaS, grid data, indexed sequential access method

ファイル構造とその特性

	順番探査	直接探査	索引付き順番探査
一部書き出し処理	単純：ただ先着順	単純：決まったところにシークするだけ	複雑：先着順、その位置を別途索引に保存
ファイルサイズ	コンパクト	疎配列が膨張	比較的コンパクト
一部読み込み処理	遅い：先頭から読まねばならない	速い：決まったところにシークするだけ	比較的早い：索引で決まる場所にシーク
他の利点	ファイル作成時に内容未確定でもよい	単一ファイルへの並列書き出し	
気象格子データの例	GRIB2 GTOOL3	GrADS Binary netCDF	NuSDaS