

## 導入事例

インテル® Xeon® プロセッサ・ファミリー

インテル® Cluster Studio XE

インテル® Xeon® プロセッサ・ファミリーを搭載したスーパーコンピューターの導入

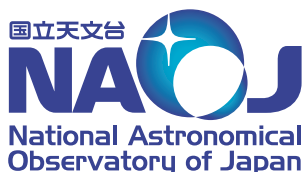


# インテル® ノード・マネージャーによって 消費電力を可視化し、 スーパーコンピューターの消費電力を最適化

天文学のシミュレーション計算に 1Pflps のスーパーコンピューターを導入



写真提供: 国立天文台



自然科学研究機構 国立天文台

所在地: 東京都三鷹市大沢 2-21-1

設立: 1888 年 (明治 21 年) 6 月

(東京天文台の設置)

事業内容:

天文学の観測・研究・開発および、

全国の研究者との共同利用

<http://www.nao.ac.jp/>

## 課題

- ・ 国内外の利用者による数値シミュレーション研究を支援
- ・ 消費電力のモニタリングとパワー・キャッピング
- ・ コードの最適化とアプリケーションの高速化、クラスターのパフォーマンスの最適化

## ソリューション

- ・ インテル® Xeon® プロセッサ・ファミリーを搭載したクラスター型スーパーコンピューター
- ・ インテル® ノード・マネージャー
- ・ C++/Fortran プログラマーおよび MPI アプリケーション開発者向け総合ツール「インテル® Cluster Studio XE」

## スーパーコンピューターの共同利用 により、シミュレーション天文学の 研究者を支援

コンピューターシミュレーションによって、宇宙のさまざまな謎を解明する「シミュレーション天文学」を推進する自然科学研究機構 国立天文台 (以下、国立天文台) 天文シミュレーションプロジェクト (CfCA: Center for Computational Astrophysics)。その目的は、スーパーコンピューターや重力多体問題専用計算機、中小規模の PC クラスタなどから構成される計算機システムを共同利用形式で運用し、国内外にいる研究者のシミュレーション研究を支援することにあります。現在はそのミッションを達成するため、ハードウェアとソフトウェアの両面から新しいシミュレーション方法の開発に取り組んでいます。

一口に天文学と言っても、現在は大きく分けて 3 種類の研究分野が確立されています。第 1 が望遠鏡を使った最も歴史の古い「観測天文学」で、第 2 が宇宙の物理を考える「理論天文学」です。20 世紀後半までは「観測」と「理論」の 2 つが天文学の主流でしたが、コンピューターの性能向上と共に、コンピューターを使って実験的に研究することが可能となり、第 3 の天文学として「シミュレーション天文学」が誕生しました。

天文学で行われるシミュレーションは主に、星団や銀河などの振る舞いを調べる「多体計算」、ガスでできた恒星や星間雲などの振る舞いを調べる「流体計算」、光の伝わり方を調べる「輻射輸送計算」の 3 種類あります。

天文シミュレーションプロジェクトプロジェクト長の小久保英一郎氏は、「シミュレーションで扱う空間スケールは、小は惑星から大は宇宙全体まで多岐にわたり、時間スケールも 1 秒から宇宙年齢の 138 億年までにおよびます。つまり、シミュレーションは宇宙におけるあらゆる現象を対象としています」と語ります。

国立天文台では現在、シミュレーション天文学の研究者が、惑星系の形成、星の誕生、ブラックホール降着円盤、超新星爆発など、多彩なジャンルの研究を、スーパーコンピューターなどの計算機システムを利用して行っています。

## 天文学専用としては世界最速 性能は 500TFLOPS から 1PFLOPS へ

CfCA では天文シミュレーションを行うスーパーコンピューターとして、2008 年 4 月に Cray 製の並列スカラ型計算機「XT4 システム」と、NEC 製のベクトル型計算機「SX-9 シ

# インテル® Xeon® プロセッサ・ファミリーが 搭載する「インテル® ノード・マネージャー」が 電力管理を効率化



「スーパーコンピューター本体と電気料金の合計で予算が決まる運用体制において、最大限のパフォーマンスを発揮するためには、インテル® ノード・マネージャーのリアルタイム電力監視機能と、消費電力の上限を設定できるパワー・キャッピングを活用した電力の最適化が欠かせません」

自然科学研究機構 国立天文台  
天文シミュレーションプロジェクト  
プロジェクト長  
博士(学術)  
小久保 英一郎 氏

システム」を導入して運用してきました。2013年4月にシステム更新を迎え、新たにCray製の大規模並列スカラ型計算機「XC30システム」を導入し、2013年4月1日より本格的な運用を開始しています。XC30を含む、CfCAが保有する共同利用計算機システムは、日本の研究機関に所属する研究者および日本で学位を取得した海外の研究機関に所属する研究者であれば誰でも申請ができ、審査によって価値が認められれば無料で利用が可能です。現在は約100名がXC30を利用中で、その稼働率は90%を超えています。

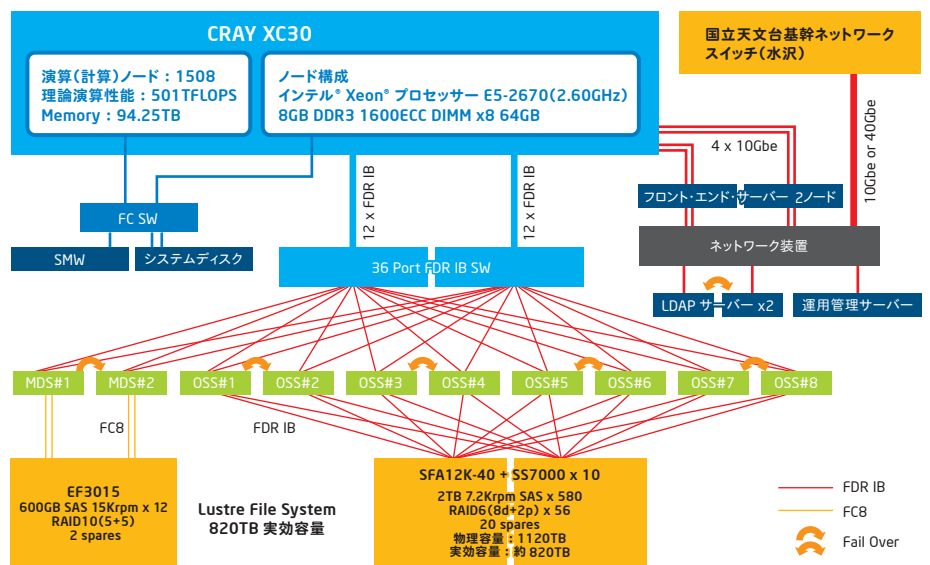
XC30は、Sandy Bridge（開発コード名）版のインテル® Xeon® プロセッサ E5-2670 (2.60GHz) を搭載し、16個のCPUコアで構成されるノードが、システム全体で1,512ノード、2万4,192CPUコアとなるシステムです。メモリー容量は1ノード64GB、システム全体では94.25TBで、さらに数値シミュレーションのデータを格納するために820TBのストレージ(lustre)と接続されています。理論演算の性能は現時点で501.8TFLOPSですが、2014年9月までにノードをインテルの次世代マイクロアーキテクチャーであるHaswell（開発コード名）版インテル® Xeon® プロセッサ・ファミリーに交換することで、1PFLOPS以上を実現する計画です。「XC30は天文学専用のスーパーコンピューターとしては現時点で世界最速であり、2013年6月に発表された世界全体の性能ランキングにおいても、63位の性能を誇っています」と小久保氏は語ります。

XC30は、CfCAによって「アテルイ」という愛称が付けられています。この名前は、平安時代の初期、征夷大將軍 坂上田村麻呂が行った東北の蝦夷(エミシ)討伐の際、蝦夷を率いて朝廷軍と戦った英雄「阿弭流為(アテルイ)」に由来するものです。XC30が岩手県奥州市の国立天文台 水沢 VLBI 観測所に設置されていることから、水沢にゆかりのある英雄にちなんで「アテルイ」と名付けられました。

XC30を水沢に設置した理由を小久保氏は、「設置場所を検討した時期の電気料金が、東京電力より東北電力の方が安価であったことが1つです。そして、東北地方は外気温が低いことから冷却効率に優れ、空調維持費を抑えることもできます」と説明します。アテルイは水沢 VLBI 観測所に設置されていますが、それ以外の計算機はすべて東京都三鷹市の国立天文台に置かれ、水沢と三鷹間は10Gbpsの光ファイバーで接続されています。

## インテル® ノード・マネージャーを用いて消費電力と積算電力の推移を監視し、急激な消費電力の増加に対応

XC30の特筆すべきポイントは、インテル® Xeon® プロセッサ・ファミリーが搭載する電力制御機能「インテル® ノード・マネージャー」を活用して消費電力をモニタリングしていることです。さらに、一定の電力量を超えないように制御するパワー・キャッピングを設定することもできます。インテル® ノード・



XC30システム構成概要図

マネージャーでXC30の電力をモニタリングする目的について、小久保氏は次のように語ります。

「近年、スーパーコンピューターの調達は、コンピューター本体の価格ばかりでなく、電気料金も含めた、合計金額の枠組みで予算化されるようになりました。予算総額からコンピューター本体の価格を引き、残った金額を5年で割った金額がXC30の稼働で許される1年間の電気料金となります。それゆえ、電気料金が高額になればなるほど、本来行うべきシミュレーションが制限されることになるため、電気料金の最適化はスーパーコンピューターの調達において重要なカギを握るのです」。

現在CfCAでは、水沢VLBI観測所の運用管理者が、Cray社から提供されたモニタリングツールを用いて、定期的に消費電力の日次変化や、月の累積消費電力をチェックしています。そして、消費電力が何らかの原因で急激に増加したり、電力会社と取り決めた月の契約消費電力量をオーバーしたりするような、急を要するケースにはアラートが三鷹のCfCAに伝達される体制となっています。1カ月の最大需要電力が契約電力を超えると、電力会社に違約金を支払うことになるため、使用電力量の監視は欠かせません。

消費電力のピークは、1年間で空調をフルで使う夏と冬の2回あり、1カ月の中でも定期メンテナンス後に割り当てられている全システムを用いての計算の際に、契約電力を超す恐れがあると言います。

「通常は、月に1回、Cray社から提供される消費電力の推移と、積算電力のグラフを見ながら、定例会議において電力会社との契約電力の見直しなど、次の月の電力対策を話し合っています。2013年4月の運用開始から半年経った現時点では、インテル® ノード・マネージャーでパワー・キャッピングをかけることなくフルパワーでの運用が続いていますが、今後電力料金の値上げが続くと、パワー・キャッピングでCPUのクロック数を下げたり、ノードの一部を停止したり、ユーザーの使用制限を設けたりといった対策が必要になる可能性があります」(小久保氏)。

また、2014年9月のシステム・アップグレードによって、理論演算の性能が1PFLOPSを超えてくると、より消費電力の最適化が重要となると想定しています。



自然科学研究機構 国立天文台  
天文シミュレーションプロジェクト  
プロジェクト長  
博士(学術)  
小久保 英一郎 氏

### 多体計算などの並列処理の高速化にコンパイラーや数値計算ライブラリーを活用

今回のXC30の導入では、スーパーコンピューターの調達と並行して、シミュレーション・ソフトのコンパイルなどの開発用途に、Linux\* 版の開発ツール「インテル® Cluster Studio XE」を導入しました。インテル® Cluster Studio XEには、インテルの高度なクラスターツール、高速化・並列化のためのコンパイラー、ライブラリー、パフォーマンス分析ツール、エラー検出ツールなどがバンドルされています。

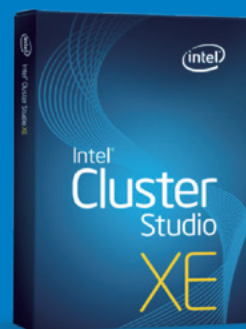
ユーザーの利用用途について小久保氏は、「100人のユーザーがいたら、百人百様の使い方があり、それぞれがインテルのコンパイラーや他社のコンパイラー、フリーのコンパイラーなどと併用しながら、並列計算処理の高速化を行っています」と語ります。XC30を利用する研究者は、ほとんどが自作のシミュレーション・ソフトを使うため、アプリケーションの構造もすべてばらばらで、最適化の手法も異なります。どのコンパイラーを使うかは、ユーザーの好みに依存することが多く、コンパイラーによって多体計算が得意、流体計算が得意、輻射輸送計算が得意といったクセがあるほか、ユーザーの書くコードのクセによっても性能を発揮するコンパイラーは変わってくると言います。「高度なスキルを持つユーザーは、コンパイラーの吐き出したアセンブラコードを読み解き、個々のコン

### インテル® ノード・マネージャー

インテル® ノード・マネージャーは、CPUのチップセットに実装された電力監視機能です。システム側のBMC (Baseboard Management Controller)と協調して動作し、プロセッサとメモリー・サブシステムの消費電力だけでなく、システムレベルの消費電力を監視したり、システム、プロセッサ、メモリーの各レベルで消費電力を制限したりする機能を備えています。インテリジェントな電力管理機能をプラットフォーム・レベルで使用することにより、電力冷却運用コストを適切に管理してTCOを削減しながら、高いパフォーマンスを維持することができます。主な機能は次のとおりです。

- リアルタイムの電力監視
- プラットフォームのパワー・キャッピング (消費電力の上限設定)
- 電力のしきい値アラート

### インテル® Cluster Studio XE



インテル® Cluster Studio XEは、MPIアプリケーションのパフォーマンスとスケラビリティ向上を支援する総合開発ツールです。インテル® C++/Fortranコンパイラー、パフォーマンス分析ツール、スレッド化アドバイザー、パフォーマンス・ライブラリー、エラー検出ツールなどをバンドルしています。バージョン2013では、マルチスレッド化すべき箇所を検出するインテル® Advisor XEが追加され、新たにインテル® MICアーキテクチャー (インテル® Xeon Phi™ コプロセッサ) をサポートしています。

### インテル® Cluster Studio XEが同梱する製品一覧

- インテル® C++ Composer XE
- インテル® Fortran Composer XE
- インテル® MKL
- インテル® IPP
- インテル® TBB
- インテル® MPI ライブラリー
- インテル® Trace Analyzer/Trace Collector
- インテル® VTune Amplifier XE
- インテル® Inspector XE
- インテル® Advisor XE

パイラーの性能やクセをチェックしながら、最適なコンパイラーを利用しているようです」と小久保氏は説明します。

コンパイラーと同様、ユーザーは自作のライブラリー群に加えて、複数の接続に対応する Intel® MPI ライブラリー、スレッド化を簡素化する Intel® TBB、パフォーマンスの向上を図る数値演算ライブラリーの Intel® MKL などを活用。さらに、アプリケーションを最適化するチューニングツールである Intel® VTune Amplifier XE や、複数のノードをまたいでパフォーマンスのボトルネックを発見する Intel® Trace Analyzer/Trace Collector などを適宜利用しながら開発を行っています。

「ヘビーユーザーになると、SIMD 命令をアセンブラで書いてクロック数を減らす工夫をしたり、オリジナル数値計算ライブラリーを作成して公開したりしています。私自身も、多体問題の重力計算を行う際には、こうしたオリジナル・ライブラリーや、Intel の数値計算ライブラリーを活用しながら、CPU の

ロック数を下げる工夫を重ねています」(小久保氏)。



写真提供:  
国立天文台天文シミュレーションプロジェクト

Intel® Xeon® プロセッサー・ファミリー搭載  
Cray 製の大規模並列スカラ型計算機「XC30 システム」

### 高い稼働率を維持しながら シミュレーション天文学の発展に貢献

今後の計画については、2014 年 9 月のシステム・アップグレードを計画通りに進めるほか、4 年後の次期スーパーコンピューターの更新に向けて、検討を重ねていくとしてい

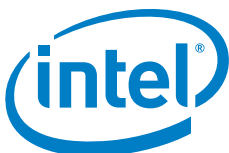
ます。スーパーコンピューターの運用について小久保氏は、「国策としてフラッグシップ・モデルの高性能化を追究していくことも重要ですが、それと同時に小回りの利く、高性能のスーパーコンピューターを揃えて活用頻度を高めていくことが、天文学およびあらゆる学問の発展には必要と考えています。科学論文をまとめるためには、アテルイ級のスーパーコンピューターで何度も何度も計算を繰り返すことが必要です。そのためには CfCA のような研究機関が、アテルイのようなスーパーコンピューターの運用を継続することが、真の科学を進める原動力になるはずだ」と提言しています。

さらに Intel に対しては、「チップベースで電力を制御する Intel® ノード・マネージャーに続き、ネットワーク機能などのチップ化に期待しています」と述べました。Intel は今後も Intel® Xeon® プロセッサー・ファミリーの機能強化と開発ツール群の技術革新を通して、CfCA が目指すスーパーコンピューターの運用効率化と、アプリケーション開発の高度化に貢献していきます。

Intel® ノード・マネージャーに関する詳しい情報は、下記のサイトをご覧ください。

<http://www.intel.co.jp/content/www/jp/ja/data-center/data-center-management/node-manager-general.html>

ソリューション・プロバイダー



性能に関するテストや評価は、特定のコンピューター・システム、コンポーネント、またはそれらを組み合わせて行ったものであり、このテストによる Intel 製品の性能の概算の値を表しているものです。システム・ハードウェア、ソフトウェアの設計、構成などの違いにより、実際の性能は掲載された性能テストや評価とは異なる場合があります。システムやコンポーネントの購入を検討される場合は、ほかの情報も参考にして、パフォーマンスを総合的に評価することをお勧めします。Intel 製品の性能評価についてさらに詳しい情報をお知りになりたい場合は、「Intel パフォーマンス・ベンチマークの限界」を参照してください。Intel は、本資料で参照している第三者のベンチマーク・データまたは Web サイトの設計や実装について管理や監査を行っていません。本資料で参照している Web サイトまたは類似の性能ベンチマーク・データが報告されているほかの Web サイトも参照して、本資料で参照しているベンチマーク・データが購入可能なシステムの性能を正確に表しているかを確認されるようお勧めします。

Intel® コンパイラーは、互換マイクロプロセッサー向けには、Intel 製マイクロプロセッサー向けと同等レベルの最適化が行われない可能性があります。これには、Intel® ストリーミング SIMD 拡張命令 2 (Intel® SSE2)、Intel® ストリーミング SIMD 拡張命令 3 (Intel® SSE3)、ストリーミング SIMD 拡張命令 3 補足命令 (SSSE3) 命令セットに関連する最適化およびその他の最適化が含まれます。Intel では、Intel 製ではないマイクロプロセッサーに対して、最適化の提供、機能、効果を保証していません。本製品のマイクロプロセッサー固有の最適化は、Intel 製マイクロプロセッサーでの使用を目的としています。Intel® マイクロアーキテクチャーに非固有の特定の最適化は、Intel 製マイクロプロセッサー向けに予約されています。この注意事項の適用対象である特定の命令セットの詳細は、該当する製品のユーザー・リファレンス・ガイドを参照してください。

この文書は情報提供のみを目的としています。この文書は現状のまま提供され、いかなる保証もいたしません。ここにいう保証には、商品適格性、他者の権利の非侵害性、特定目的への適合性、また、あらゆる提案書、仕様書、見本から生じる保証を含みますが、これらに限定されるものではありません。Intel はこの仕様の情報の使用に関する財産権の侵害を含む、いかなる責任も負いません。また、明示されているか否かにかかわらず、また禁反言によるとよらずにかかわらず、いかなる知的財産権のライセンスも許諾するものではありません。

Intel、Intel、Intel ロゴ、Xeon、Xeon Inside、Intel Xeon Phi は、アメリカ合衆国および / またはその他の国における Intel Corporation の商標です。

Microsoft、Windows は、米国 Microsoft Corporation および / またはその関連会社の商標です。

\* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。

Intel 株式会社

〒100-0005 東京都千代田区丸の内 3-1-1

<http://www.intel.co.jp/>

©2014 Intel Corporation. 無断での引用、転載を禁じます。  
2014 年 1 月

329848-001JA  
JPN/1401/PDF/SE/ESS/FN