

はじめに

緊急事態宣言下の 2021 年 1 月 21 日、科学技術計算分科会 2020 年度会合がオンラインで開催された。今回のテーマは「富岳スペシャル 2.0～より深く詳しく～」であり、今春から始まる「富岳」の本格稼働に向けて、開発の経緯から実践例まで 4 つの講演が行なわれた。オンラインウェビナーの方式ではあったが、101 名の参加者で活発な議論が繰り広げられた。

まず、企画委員の代表として工学院大学の田中輝雄氏より趣旨説明があった。世界一になった富岳がコロナ対応にも利用がされており、「スーパーコンピュータの結果により」というキーワードが政府やメディアでも使われるようになってきた。新型ウイルスに負けないように HPC の未来を切り開くような議論を望む、という挨拶で本会は幕を開けた。

「スーパーコンピュータ「富岳」が拓く創薬・医療の未来」

奥野 恭史（京都大学）



最初の講演は、富岳のユーザーの視点からコロナの治療薬に関する研究事例の紹介であった。新型コロナワクチンの接種に関するニュースが連日報道されてはいるが、ワクチン

の接種が開始されても全ての感染を防げるわけではないため、治療薬も必要になる。奥野氏らのグループは、2020年4月より新型コロナウイルスの治療薬候補同定のために富岳を利用した。治療薬の探索は、分子シミュレーション(分子動力学計算)を用いて既存の2128種の医薬品(=安全性が確認されている)からコロナの増殖に関連するタンパク質に結合するものを探す。現在進行形で被害が広がっている新型コロナウイルスに対して、ゼロから医薬品を開発していたのでは間に合わない。だからこそスパコンが必要であるという。

シミュレーションは、原子間で働く力を考慮しつつ、体の中の環境に近い状態で、タンパク質に薬剤がドッキングするかを計算している。時間発展幅は2フェムト(500兆分の1秒)であり、1ミリ秒間の動きを再現する場合には、約5000億回の計算が必要になる。京では1マイクロ秒のシミュレーションで数10種類までは計算ができたが、この数十種類の化合物を選ぶことが課題であった。富岳では、数千種類のシミュレーションを実現し、これは世界初の成果となった。

ただ、実際の生体反応はミリ秒オーダーであり、富岳でも1ミリ秒×数千種類は現実的ではない。今回の研究では、薬剤の濃度を高めコンピュータ上での結合の“頻度”をあげることで、アメリカで有力視され臨床実験が行なわれている薬剤と同じものを見つけ出すことができた。当初は2021年からの利用を想定していたため、このシミュレーションに10日間を要したが、現在はアプリケーションのチューニングをすることによって2日程度で可能になった(9.2倍の高速化を実現している)という。

一方で、スパコン利用の別側面としてAIを用いた創薬がある。細胞や動物全体をモデル化することは難しいため、スパコンでシミュレーションをできないが、創薬における医薬品設計のフェーズにAIを適応することで時間を短縮することができる。実際の創薬の現場では、化学構造の最適化(化学合成実験と活性化試験)を何度もしなければならない。このモグラたたきを「化学構造を生成するAI」と「活性を予測するAI」の、2つのAIに繰り返し計算させることによって化学構造をデザインするアプローチが研究されている。

最後に、新型コロナウイルスは、なぜ人によって症状が違うのかという疑問に対する研究について述べた。遺伝子同士の発現制御をネットワークで表現し、健康者と患者のネットワークの差異が感染の重症性を表しているのではないかという仮説をもとに分析を進めている。現時点での富岳の利用は、新型コロナウイルスに関する利用に限っているが、共用開始後は、シミュレーションだけでなく富岳の強みであるAIもシームレスに動かせることも活かして「シミュレーション×AI×生命科学の実験」を実現したいと述べ講演を終えた。

「スーパーコンピュータ「富岳」の開発経緯」

石川 裕 (理化学研究所)

続いて、富岳システムの開発の経緯や成果について、プロジェクトリーダーの視点からの講演が行なわれた。文科省より理化学研究所が委託を受けて富岳(当時はポスト京)の開

発を進めたのは2014年であったが、今回の講演では、実際に富岳に関する話を始めた2010年ごろからの経緯を紹介した。

最初は「ロードマップ策定」のフェーズである。「科学研究ロードマップ白書」や「HPCI技術ロードマップ白書」のようなスパコンの将来像をまとめた白書が策定される。2012年度より具体的な「調査研究」が始まり、スパコンの基本方式とその性能見積もりや、将来の社会的・科学的課題とその解決のためにスパコンで動かすアプリケーションがまとめられた。次の「基本設計」のフェーズでは、富岳の仕様や開発すべき項目を定めた。当初は汎用部+加速部を有するマシンということで提案を進めていたが、予算的な面などから最終的には汎用CPU単体の構成になったことなど、実際の開発のエピソードについても触れていた。

今回の講演の主題は、基本設計で確定した仕様をもとに具体的なハードウェア・ソフトウェアを設計し、試作機を制作した「詳細設計」のフェーズであった。基本設計を終えた時点では、富岳のCPUは10nmの半導体加工技術を利用する予定であったが、7nmに変更になり運用開始予定時期が遅延したことについても言及した。この期間を用いて、AIの性能を高められるCPU機能(半精度浮動小数点演算)の追加をしたことで利用の幅を広げることができたという。その他にも、富岳の設計方針である「課題解決型」「協調設計(コデザイン)」「使い勝手の向上」「Total Cost of Ownership」「拡張性&社会が欲するニーズに即応」の説明があり、特に「コデザイン」については、アプリ側からの改良やシステムへの要請など、開発者側とユーザーのインタラクションなどに関する込み入った話題も提供された。

最後に今冬に終わりを迎える「製造・設置・調整」のフェーズがある。当初計画では2020年度からの正式運用が1~2年遅れる可能性があったものの、2020年の春には(限定的ではあるが)新型コロナ禍対策等への利用提供を実現できた。

講演の結びで石川氏は「(富岳のプロジェクトを通じて)10年後の要素技術の進歩予測に基づいた目標設定の難しさを経験した」と語った。

「スーパーコンピュータ「富岳」を支えるコンパイラの技術」

千葉 修一(富士通株式会社)

科学技術計算分科会 2020年度会合
富岳スペシャル2.0 ～より深く詳しく～

REC



スーパーコンピュータ「富岳」を支える コンパイラの技術

2021年 1月 21日
富士通株式会社
プラットフォームソフトウェア事業本部
千葉 修一

Copyright 2021 FUJITSU LIMITED

休憩をはさみ、次の講演は、富岳のコンパイラ開発を担った千葉氏による「スーパーコンピュータ「富岳」を支えるコンパイラの技術」であった。石川氏の講演にもあった富岳の開発方針であるコデザインの中で、コンパイラがどのように改良されていったかを解説された。本講演は、情報処理技術者試験の問題や全体像を捉えやすいイラストなど、誰にでも分かりやすいスライドが印象的であった。

単純な命令を組み合わせて複雑な要件を実行する RISC アーキテクチャでは、“コンパイラ”の力量が試される。高い処理能力を持つ富岳も、コンパイラによる最適化技術が無いとプログラムの実行性能を引き出すことができない。

コデザインにおけるコンパイラ開発の難しさは、システムやハードウェアが利用できない状態でソフトウェアの設計と開発をしなければならない点であったという。特に京と富岳の差異として、命令セットが SPARC から ARM に変わったことや、浮動小数点のレジスタ数が 256 から 32 に減ったことなどがある。大容量レジスタを活用して最適化を行っていた京の開発時のノウハウを活かすことができるのかを想像しながら開発を進めなければならないところが、コデザインの難しさであったと語っていた。

コンパイラの開発は、一般販売されているコンパイラの改造や ARM のマシンの評価を行なうことで進めていた。基礎性能コードによる測定や、より実際の利用に近いベンチマークおよび数値シミュレーションによる評価が対象である。評価の方法は、最適化の適用の有無で性能がどのように変化するのかと、プロファイラを用いた実行時のコスト分布より性能の傾向を分析することなどであった。比較は、FX 100 のマシンで 128 レジスタの標準モードと仮想的に 32 レジスタに抑えたモード、加えて ARM チップである X-Gene 2 が搭載されたマシンの 3 マシンで行なった。

実アプリケーションを用いた評価に関しては、NICAM(非静力学正 20 面体格子大気モデル)の数値シミュレーションの事例が紹介された。コードの特徴として、メモリへの要求が高い力学過程と計算量や分岐が多い物理過程があり、コンパイラの最適化の評価に使い

やすいアプリケーションである。この中でも特徴的な部分のコードを抜き出して、レジスタの利用や実行時間を比較した。

開発時点ではコデザインの最中であるため、これらの知見を手修正でコードに加えることによって最適化を行ない、実際に性能が向上したという。A64FX のチップができる前に得られたこの成果は、富岳に関する技術として初めての論文になった。

続けて、企業としての事業化について紹介があった。コデザインは実用化や商品化の一部であるが、その先の商品化と事業化においては、より厳しい「品質基準」が課せられる。この際に重要になってくるのは、商品化までの「品質」であるという。品質を担保するためには、無限大に近いテストの組み合わせの中から、テスト範囲を設定する必要があるそうだ。また、社会で利用される製品を作る難しさは、最適化機能の利用頻度や適用範囲、利用者の価値観を踏まえた上で開発をしなければならないことであると述べた。

障害修正のスピード感、最適化オプションの提供、アグレッシブな最適化が欲しいと言ったようなサービス性を重視する利用者側の要望と、品質を重視している富士通側のギャップをいかに埋めていくかは、富岳の本稼働に向けても重要になってくる。富士通側の開発の思想をオープンにして、利用者とディスカッションをするキッカケを作ることが SS 研の意義でもあると語っていた。

「第 3 世代 JAXA スパコンが目指すもの"TOKI"導入の目的と初期性能評価」

藤田 直行(宇宙航空研究開発機構)



The image shows a presentation slide for the 3rd generation JAXA supercomputer project. The slide features the JSS3/TOKI USPACE logo and the title '第3世代 JAXAスパコンが目指すもの -"TOKI"導入の目的と初期性能評価'. Below the title are two photographs: one of a large supercomputer cabinet and another of a person working at a computer. The presenter's name, '藤田 直行', and his affiliation, '国立研究開発法人 宇宙航空研究開発機構', are listed at the bottom. A video feed of the presenter is visible in the top right corner. The slide is part of a presentation from the SS研 2020年度 各イベント 科学技術計算分科会, and the presenter is identified as 講演者4 藤田直行の画面.

第3世代
JAXAスパコンが目指すもの
-"TOKI"導入の目的と初期性能評価

国立研究開発法人 宇宙航空研究開発機構
スーパーコンピュータ活用課長/航空技術部門研究領域主幹
藤田 直行

講演者4 藤田直行の画面

SS研 2020年度 各イベント 科学技術計算分科会

P.1

最後の講演は、2020 年 12 月より稼働した JAXA のスパコン「TOKI」の導入や初期性能評価など、FX 1000 の利用者として特にハードウェアを中心に実践例の報告であった。

JAXA ではスパコンの導入に際し 3 本の柱「航空宇宙分野の国際競争力を強化する数値シミュレーション実施基盤」「大規模データ解析基盤としてのデータセンター機能」「新たなニーズを受け入れる研究開発基盤」を立てた。この背景として、数値流体力学(CFD)が実用の時代になってきたこと、加えてスパコンの利用分野が CFD のみならず、地球観測データ処理や AI 基盤というデータ処理へと範囲が拡大したことがあったという。

TOKI をコンピュータ基盤とする JAXA のスーパーコンピュータシステム JSS3 は、調布航空宇宙センターとつくば宇宙センターの 2 か所に分けて設置されている。調布は非常時に停電の可能性もあるためサブセットを筑波に設置し「止まらないスパコン」を実現したと語った。講演では JSS3 の構成を丁寧に説明されたが、詳細は割愛する。ファイルシステムの同期は今後の検証が必要であるなど、現状と課題についての分析もあわせて述べていた。

続けて、導入時のベンチマークの話に移った。JAXA で利用している主な 5 つのアプリケーションに対するベンチマークの結果は、時間と電力ともにすべて要求水準を満たしていた。TOP500 の結果は、16.592PFLOPS で 19 位(国内 3 位)、実行効率は 85.24%であった。HPCG の結果は、0.61422PFLOPS で 8 位(国内 2 位)であった。Graph500 の結果は、10,813GTEPS で 3 位(国内 2 位)であった。IO500 については今後測定予定である。

最後に、これからの JSS3 として 4 つの新たな利用を述べた。1 つ目は TOKI-FS のランダム IO 性能、特に IOPS を必要とする処理を本格実行する。これは小さなファイルを処理や機械学習時に効果が見込まれる。2 つ目は、TOKI-SORA/RURI の仮想化技術の利用であり、どちらも Singularity を基本として準備を進めている。これによって企業や大学など外部コンピュータ環境との連携を推進しようと考えているという。3 つ目は、TOKI-GP の GPGPU 利用である。これまでは JAXA 内で GPGPU の利用は進んでいなかったが、この利用を促進する。4 つ目は、これまでの大規模計算のさらなる拡大を「大規模チャレンジ」と銘を打ち取り組んでいる。この成果は近いうちに発表するとのことである。質疑応答では、システムの利用者らによって現時点での課題や要望の共有が行なわれた。

終わりに

閉会挨拶の前に、理化学研究所の南氏より「富岳のバーチャルツアー」が行なわれた。360 度カメラで撮影された富岳を、ブラウザ上で操作しながらの説明であったが、ラックの配置や内側まで鮮明にみることができた。

続けて閉会挨拶では、各講演に対する内容の整理とコメントをされた。アプリケーションに対するユーザーからの視点、富岳のプロジェクトを進めた側からの視点、コンパイラの開発からの視点、ハードウェアに対するユーザーからの視点など様々な角度から、活発な議論がなされたことに謝意が述べられ、会合は締めくくられた。