

令和 4 年 6 月 15 日現在

機関番号：12601

研究種目：基盤研究(B) (一般)

研究期間：2019～2021

課題番号：19H04121

研究課題名(和文) 機械学習を用いた自律型スマートHPCデータセンター

研究課題名(英文) Autonomous HPC data center using machine learning

研究代表者

松葉 浩也 (Matsuba, Hiroya)

東京大学・情報基盤センター・客員研究員

研究者番号：30444095

交付決定額(研究期間全体)：(直接経費) 13,300,000円

研究成果の概要(和文)：データセンターの運用の自動化のため、データセンターからのデータを取得して仮想空間上に再現する方法と、その仮想空間上で運用ポリシーを最適化する方法を研究した。

前者のデータ取得については、機器によるデータフォーマットの違いやデータの用途に依存することなく、様々な運用に役立つ汎用的なデータを多くのデータセンターから収集できる一般的なデータフォーマットと保存形式を定義することに成功した。後者の運用最適化については、制御方法を自動的に学習する強化学習を用いてジョブスケジューリングを最適化することに成功すると共に、その学習の場としてスケジューリングや冷却設備のシミュレータを実装した。

研究成果の学術的意義や社会的意義

データセンターの運用はより少ない資源で多くの計算需要を満たすための高度な技術であり、スケジューリングポリシーの設定、計算機のメンテナンス、さらには電力空調設備の運用やメンテナンスを含む専門的に高度な業務である。近年の労働人口減少に伴い、このような運用業務に携わる人材の確保は容易でなく、今後、益々増大する計算需要を安定的に満たすためには運用の省力化、少人化は必須の課題である。本研究の成果はこの要請に応じるものであり、設備運用やジョブスケジューリングをシミュレーションで再現する方法、さらにそれを用いて人工知能に最適な運用を試行錯誤させる方法を開発したものである。

研究成果の概要(英文)：To automate data center operations, we studied methods of acquiring data from data centers and reproducing them on a virtual space, as well as methods of optimizing operational policies on that virtual space.

For the former, data acquisition, we succeeded in defining a general data format and storage format that can collect general-purpose data from many data centers useful for various operations without depending on differences in data format by device or data usage. For the latter operational optimization, we succeeded in optimizing job scheduling using reinforcement learning, which automatically learns control methods, and implemented a scheduling and cooling equipment simulator as a place for such learning.

研究分野：計算機科学

キーワード：データセンター 運用 人工知能 スケジューラ ログ収集 強化学習

1. 研究開始当初の背景

本研究を開始した2019年にはすでに深層学習を中心とした人工知能に注目が集まっており、計算資源に対する要求が高まっていた。深層学習を用いた大規模なモデルの学習には大量のデータと計算機を使用して数時間から数日に及ぶ計算を実行することは珍しくない。このような長時間に及ぶ計算は古くから大規模シミュレーションを主な用途と想定してきた大学等のスーパーコンピュータでは珍しくないものの、Webアプリケーションなどの短時間の処理を大量に行うことを主目的にしてきたクラウドサービスには新しいタイプの負荷である。

人工知能に代表される長時間タイプの負荷を効率的に処理するためには、ジョブスケジューリング等の運用においてもそれに見合った技術が必要である。大学等のスーパーコンピュータはその技術を有しているが、運用に精通した技術者の高度な判断に基づいて運用されている側面があり、新たに人工知能の計算負荷に対応する必要に迫られたデータセンター運用者が即座に活用できるものではない。属人的な能力に拠らずに運用が可能な技術に対するニーズが高まることが予想された。

2. 研究の目的

本課題は人工知能に代表される長時間の計算（以下、HPC）を実行するデータセンターにて、データセンターが提供する計算資源量を最大化し消費電力量を最小化する運用を自律的に行うことができる機械学習手法を開発する。そのために以下の3点の研究を実施する。

- (1) データ収集：機械学習での運用最適化のためには学習の元となるデータの収集が必須である。各データセンターが所持するこれらの履歴データは収集に用いているツールや計測地点・計測方法の違いにより、それぞれ独自のフォーマットを持っているためこれらを統一化した共通のデータフォーマットを作成する。
- (2) モデル化：機械学習における運用最適化の場として、ある運用判断を行った場合の結果を算出できるモデル（シミュレータ）を作成する。
- (3) 運用自動化：ジョブ実行効率最大化と最大消費電力の最小化を目指すジョブスケジューリングについて、強化学習にてスケジューリングのパラメータ設定を自動化する方法を研究する。

本課題では理化学研究所、産業技術総合研究所、東京工業大学でそれぞれ運用しているスーパーコンピュータに適用することを想定して研究を推進することにより、特定の計算機構成に依存しない汎用的な手法を開発する。

3. 研究の方法

(1) データ収集

研究分担者の野村・三浦・滝澤を中心に、研究分担者らがアクセスできる東工大・理研・産総研のそれぞれのデータセンターを対象に測定している項目を整理した。また、運用自動化部分が要求する入力を参考に、共通の履歴データフォーマットを作成し、各データセンターのログを共通形式で表した。

(2) モデル化

東工大ではシステム全体の利用効率を向上させるために、特にCPUコアの異ジョブ間での共有を許容するオーバーコミットスケジューリングに注目しオーバーコミット時の挙動を理解するための、異なるジョブがコアを共有したときの性能への影響(多くは速度低下)の研究、本研究課題で行う強化学習によるスケジューリングアルゴリズムの実証プラットフォームとして、オーバーコミットスケジューリングのシミュレータの研究開発、について推進した。

産総研では強化学習活用のためにジョブの傾向を把握すべく、産総研が運用する ABCI のジョブ実行履歴と資源利用量等のシステムが生成するセンサーデータを収集した。このデータを用いて、ジョブの傾向分析を行うとともに、ジョブスケジューリングアルゴリズムを評価する際の入力データを生成した。

(3) 運用自動化

産総研において機械学習・強化学習によるジョブスケジューリングポリシーを作成した。以前より開発していたジョブスケジューラシミュレータを拡張し、機械学習・強化学習システムで使えるようにし、ABCI を用いて多数のシナリオ・パラメータで並列に学習を行った。

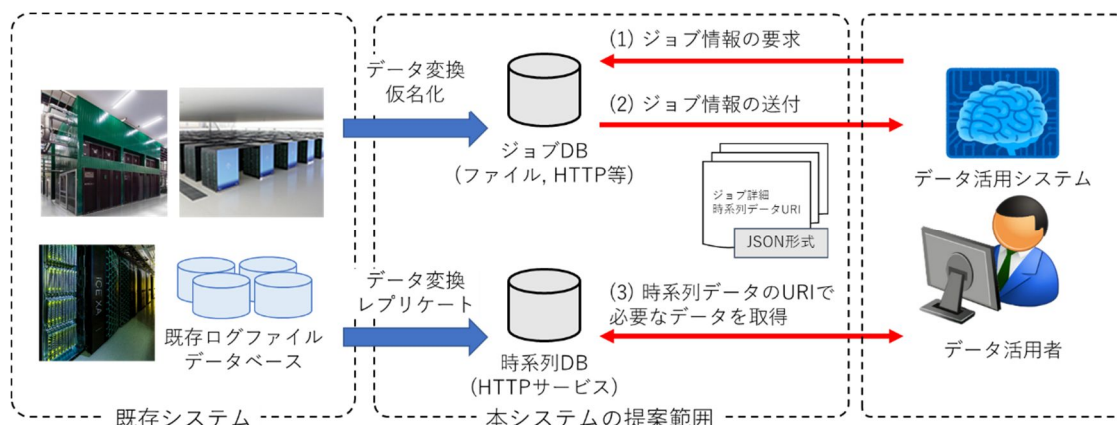
4. 研究成果

(1) データ収集

研究分担者の野村・三浦・滝澤を中心に、ジョブおよびセンサー履歴データの整備に関する検討を行った。それぞれのデータセンターが収集している履歴データは、その収集プラットフォームや計測項目の違いが大きく、それぞれのデータセンターの履歴データをそのまま用いて学習プログラムを作成するとそれぞれの機械学習コンポーネントをデータセンターの数だけ作成することとなり、非現実的である。各データセンターの履歴データフォーマットに依存せずに学習データとして利用するため、各データセンターにおける既存の履歴データをもとに共通の中間形式を経由して扱うこととした。共通形式に求められる要件としては、(1)計測項目を柔軟に追加できること (2)ジョブの履歴データとセンサー情報の履歴データの関連付けを取りつつ、時系列データとなるセンサー情報の履歴データを時系列データのまま、もしくは集約を行った値として利用できることなどが挙げられる。このうちジョブ履歴データフォーマットとしては Parallel Workload Archive で用いられている SWF 形式がデファクトスタンダードとして存在するが、本形式は(1)ノード内並列とノード間並列の区別ができておらず現代のクラスタ型並列計算機における並列性を示すには情報量が不足している (2)CSV をベースとしているために、スキーマ拡張時に任意のデータを含めることが難しく、加えて時系列データを含めることはほぼ不可能 という点で、本プロジェクトの目的には合致せず、これらの要件を満たす形で新たに標準ジョブ履歴スキーマを設計し、各データセンターの履歴データを標準スキーマの形式で提供する変換ゲートウェイを作成することとした。

標準ジョブ履歴スキーマでは、データセンター間の計測項目の違いを吸収しつつも、同じ学習プログラムが履歴データを利用できるよう、構造化された JSON 形式を基本とし、データセンターによらず確実に収集できる項目・データセンターの自律運用において常に必要となるであろう項目を必須項目、それ以外のデータセンターによっては計測できない・記録していない可能性がある項目をオプション項目として整理した。また、計測項目の値の解釈や記法について揺れが生じうる項目について、セマンティクスを定義し今後行われるであろうデータ拡張を含めて、曖昧性を最小化すべく設計した。

履歴データ収集基盤の実装として、東工大の TSUBAME3.0 を対象にプロトタイピングを実施した。TSUBAME3.0 ではジョブ履歴として Univa(現 Altair) Grid Engine(UGE)のログデータ、センサー履歴として主に Ganglia によるノードモニタリングデータおよび Zabbix によるシステムモニタリングデータがデータ源となる。このうち UGE のデータをもとにした標準ジョブ履歴スキーマへの変換を行う Web サービスの実装を行った。産総研では ABCI を構成する各種機器への負荷少なく、機器の各種センサーを収集する基盤を構築した。ABCI では以前より死活監視の一環としてリソースモニタリングを行っていたが、その仕組みを一部流用しつつ、オープンソースソフトウェアの Prometheus・Thanos を用いた蓄積基盤を構築した。センサーデータ量は膨大であるため(圧縮状態で 1.5TB/年程度)、最新の 1 年間分のデータのみ高性能な ABCI 並列ストレージに格納し、過去のデータはオブジェクトストレージにアーカイブとして格納することとした。



作成したログデータ収集基盤の構成

(2) モデル化

オーバーコミットスケジューラ

研究分担者の遠藤・野村が中心となり運用を行う東京工業大学 TSUBAME3 スパコンではインテラクティブジョブ用キューの本格運用を 2020 年度に開始し、これにより各ユーザが待たされることなく計算資源を利用可能となった。一方で、複数ユーザのアプリケーションが同一の CPU コ

ア集合を用いるため、性能低下が起こりうる。このような影響を詳細に調査し、次世代のスケジューリング方式の確立およびその改善を自動化することを目指した研究を行い、下記の成果を得た。

- オーバーコミット時のジョブ性能の調査およびモデル化：まず NAS Parallel Benchmark の複数種プログラムなどを対象として、オーバーコミット時の性能の詳細調査を行った。その結果、N 個のアプリケーションがコアを共有する際に速度が 1/N になるような単純な影響ではなく、メモリ負荷やキャッシュ利用効率、スレッド間同期手法などによって大きく変動すると分かった。その影響を、単独アプリケーションのパフォーマンスカウンタ値などから定量的に推測するモデルを構築し、その精度を評価した。この研究成果について、ジョブスケジューリング分野を専門とする国際会議 Job Scheduling Strategies for Parallel Processing (JSSPP 2021) に論文が採択され、発表を行った。
- オーバーコミットスケジューリングシミュレータの研究開発：本研究課題において機械学習による改善対象の一つとして、オーバーコミットを採用したジョブスケジューリングを設定した。その研究の基盤としてシミュレータが必要であるが、Alea などの既存シミュレータでは機能が大きく不足していることが分かったため、新規に開発することとした。このシミュレータは各ジョブ間に設定された、オーバーコミット時の性能低下率を考慮する。評価には、Parallel workload archive で公開されているジョブ群の履歴を用い、一部は 10 万以上のジョブを含む。シミュレーションの結果により、古典的なスケジューリングに比較して各ジョブの待機時間や slow down 指標を大幅に改良することができた。一方このスケジューリングにおいては、各 CPU コアの共有を許容する最大ジョブ数などの管理者が設定するパラメータが存在し、その調整を機械学習により自動化するという課題を具体化した。

正確な実行時間の指定を優遇するスケジューリングアルゴリズム

大規模システムで実行されるジョブは、利用者が指定する実行時間よりもはるかに短い時間で実行完了することが多くの文献で報告されており、実際(1)の ABCI での分析においては、利用者は平均して 5 倍長い実行時間を指定していることが観測できた。このような実際の実行時間と利用者指定の実行時間の差は、利用者にとってはジョブ実行機会の喪失、システム運用管理者にとっては資源利用率低下の問題につながる。

この問題を解決するために、産総研では、正しくジョブ実行時間指定をするインセンティブをスケジューリングアルゴリズムに組み込む手法を提案した。利用者ごと、ジョブの特徴(並列度、要求実行時間、投入時刻、など)ごとに過去のジョブの実行時間指定の正さを学習し、実行時間指定が正しい利用者のジョブを優先的に実行する手法である。提案手法と、広く使われているスケジューリング手法とで比較を行った結果、利用率が最大 30% 向上することを確認できた。また、提案手法では、ジョブの実行時間指定の正さと平均待ち時間に強い負の相関が確認できた。これはより正しい実行時間を指定すると平均待ち時間が短縮されることを意味し、高いインセンティブを実現できたことを意味する。このような高度なスケジューリングもデータセンターモデルに組み込むことで、最先端の知見を踏まえた自律運用につなげることができる。

冷却系シミュレーション

理化学研究所では東工大の Tsubame3 の冷却系のシミュレータを作成した。理化学研究所「富岳」の冷却系は複雑であるため、初期検討のターゲットとして比較的原理がシンプルな Tsubame3 を選択した。シミュレーション作成のためにはモデリング言語である Modelica を使用した。作成したシミュレータは、計算アプリケーションの実行と共に発生する熱を入力とし、動作中の各部分の冷却水の水温などの冷却系全体の状態を示す様々な物理量を出力する。作成したシミュレーションと実際とを比較した結果、水温などの細部に誤差が残るものの、計算負荷の増大と共に稼働する冷却タワーの数が増加するなどの挙動や全体的に冷却系が正常に機能しているか否かなどの運用判断において重要な挙動は正しく再現できた。

(3) 運用自動化

産総研を中心に深層強化学習を用いたジョブスケジューリングの検討を行った。時事刻々と変動するシステムの状態と利用者の計算要求に応じて、システム状態が常に最適となるよう、スケジューリングアルゴリズム・パラメータを自律的に更新し続けるために、近年では強化学習を用いてスケジューリングアルゴリズムを学習する手法が研究されている。強化学習システムに観測したデータを入力するだけで、スケジューリングアルゴリズムの機械学習モデルが構築され、動的な最適化を自動的に行える可能性を有するためである。しかしながら、既存手法では、ジョブの属性については要求値のみ学習に用いられている。ジョブの実行時間や消費電力量は要求と実測で大きく値が異なりうるため、要求値だけを入力に学習したスケジューリングモデルは、適切な学習が行われていない可能性がある。

我々は、ジョブ実行時間を対象とし、スケジューリングモデル構築時に、要求値だけでなく実際値を入力として誤差を含めた学習を行い、モデル使用時には要求値と予測値を用いてスケジューリングを行う手法の評価を行った。システム利用率最大化を目的とした学習を行った結果、強化学習にて手動設定と同等の性能を得ることに成功した。今回の設定では人間が考案したアルゴリズムを上回る利用率は実現できなかったが、手動設定と同等のポリシーを人工知能によって開発できたことには意義があり、データセンターの自動運用の重要な要素技術となり得ることが示された。

(4) まとめ

4 節(1)項目で述べたデータ収集の仕組みにて、運用自動化に必要なデータの取得が可能となった。これは本研究で用いた3台のスーパーコンピュータに特化した仕組みではなく、汎用性を持つように設計されているため HPC データセンター一般に適用可能である。またこのログ収集の共通仕様は、HPC データセンターにて日常的に取得すべきログの種類を規定しており、新たな並列計算機的设计においてそのログ収集部分の設計指針として活用できる。次に4 節(2)項目で述べたモデル化にて HPC データセンターの挙動が計算機上でシミュレーションの形で再現できるようになった。これらの研究に加え、実行予定ジョブ列から消費電力を予測する既存技術などを加えることにより、HPC データセンターの挙動をサイバー空間に再現した Digital Twin の作成が可能となる。この Digital Twin を学習の場として4 節(3)項目で述べた手法で制御ポリシーを自動学習する技術を適用することで、HPC データセンターの自律運用が可能となる。以上のように本研究では HPC データセンターを自律運用するにあたり必要なデータ収集から運用ポリシーの自動学習までの一通りの要素技術を開発し、HPC データセンターの自律運用が現実的であることを示すことができた。

本研究は HPC データセンターの Digital Twin を作成し、その上で自律的な運用が可能となるような人工知能を学習させるアプローチを採った。この方法はひとつの現実的な方法であるものの、Digital Twin の作成自体は依然人間の能力に依存している。高度な技術を持った人間が多くの時間を割いて開発することが必要な工程が含まれているという意味で、すべてが自動化されているとは言い難く、Digital Twin 作成の自動化や汎用的で応用が容易にできる共通 Digital Twin の作成が今後の課題である。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計18件（うち招待講演 1件 / うち国際学会 8件）

1. 発表者名 Shinichiro Takizawa, Yusuke Tanimura, Hidemoto Nakada, Ryousei Takano, Hirotaka Ogawa
2. 発表標題 ABCI 2.0: Advances in Open AI Computing Infrastructure at AIST
3. 学会等名 IPSJ SIG Technical Reports HPC-180
4. 発表年 2021年

1. 発表者名 Ryousei Takano
2. 発表標題 ABCI 2.0: opportunities and challenges of an open research platform for AI/ML
3. 学会等名 The 2nd International Workshop on Machine Learning Hardware in conjunction with ISC 2021
4. 発表年 2021年

1. 発表者名 滝澤真一朗, 野村哲弘, 松葉浩也
2. 発表標題 深層強化学習を用いたジョブスケジューリングへの不確実なジョブ属性の影響の検証
3. 学会等名 第182回ハイパフォーマンスコンピューティング研究会
4. 発表年 2021年

1. 発表者名 Shohei Minami, Toshio Endo and Akihiro Nomura
2. 発表標題 Measurement and Modeling of Performance of HPC Applications towards Overcommitting Scheduling Systems
3. 学会等名 24th Workshop on Job Scheduling Strategies for Parallel Processing (国際学会)
4. 発表年 2021年

1. 発表者名 滝澤真一朗, 坂部昌久, 谷村勇輔, 小川宏高
2. 発表標題 ABCI上でのジョブ実行履歴の分析による深層学習計算の傾向把握
3. 学会等名 第176回ハイパフォーマンスコンピューティング研究会
4. 発表年 2020年

1. 発表者名 野村哲弘, 滝澤真一朗, 三浦信一, 遠藤敏夫, 松葉浩也
2. 発表標題 センサー情報を意識したジョブスケジューリング実現のための標準ジョブ履歴スキーマの提案
3. 学会等名 第178回ハイパフォーマンスコンピューティング研究会
4. 発表年 2021年

1. 発表者名 安良岡由規, 野村 哲弘, 遠藤 敏夫
2. 発表標題 学内インフラとしてのスパコンの対話的利用による利便性向上
3. 学会等名 大学ICT推進協議会 (AXIES) 年次大会, FA2-3
4. 発表年 2020年

1. 発表者名 野村 哲弘, 遠藤 敏夫, 三浦 信一, 朝倉 博紀, 越野 俊充, 草間 俊博
2. 発表標題 TSUBAME3のインタラクティブ利用の利便性向上にむけた取り組み
3. 学会等名 第175回ハイパフォーマンスコンピューティング研究会
4. 発表年 2020年

1. 発表者名 南 将平, 遠藤 敏夫, 野村 哲弘
2. 発表標題 オーバーコミットスケジュール時のアプリ性能の予備評価
3. 学会等名 第175回ハイパフォーマンスコンピューティング研究会
4. 発表年 2020年

1. 発表者名 Shohei Minami, Toshio Endo, Akihiro Nomura
2. 発表標題 Performance Modeling of HPC Applications on Overcommitted Systems
3. 学会等名 HPC Asia 2021, poster session (国際学会)
4. 発表年 2021年

1. 発表者名 Akihiro Nomura
2. 発表標題 Improving User-friendliness of Interactive Use in TSUBAME3 Supercomputer
3. 学会等名 人工知能CREST合同オンラインワークショップ(ポスター発表) (国際学会)
4. 発表年 2020年

1. 発表者名 Toshio Endo
2. 発表標題 Activity Report from Tokyo Tech:Energy Efficiency of TSUBAME3.0
3. 学会等名 Energy Efficient HPC State of the Practice Kobe Meeting (国際学会)
4. 発表年 2019年

1. 発表者名 Shinichiro Takizawa, Ryousei Takano
2. 発表標題 Effect of an Incentive Implementation for Specifying Accurate Walltime in Job Scheduling
3. 学会等名 International Conference on High Performance Computing in Asia-Pacific Region (国際学会)
4. 発表年 2020年

1. 発表者名 Akihiro Nomura
2. 発表標題 Introducing Container Technology to TSUBAME3.0 Supercomputer
3. 学会等名 ISC High Performance 2019 (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 野村 哲弘, 三浦 信一, 實本 英之, 額田 彰, 遠藤 敏夫
2. 発表標題 TSUBAME3.0におけるストレージ利用効率化のためのファイルシステムベンチマーク
3. 学会等名 情報処理学会研究報告, 2019-HPC-170 No.24
4. 発表年 2019年

1. 発表者名 高野了成, 滝澤真一朗, 三浦信一, 谷村勇輔, 小川宏高
2. 発表標題 ABCIデータセンターにおけるハードウェア障害の傾向
3. 学会等名 コンピュータシステムシンポジウム2019
4. 発表年 2019年

1. 発表者名 Ryousei Takano
2. 発表標題 Activity Report from AIST ~1st anniversary of ABCI~
3. 学会等名 Energy Efficient HPC State of the Practice Kobe Meeting (国際学会)
4. 発表年 2019年

1. 発表者名 Ryousei Takano
2. 発表標題 ABCI and the Energy Efficient Datacenter Operation
3. 学会等名 Energy Efficiency Considerations for HPC Procurements BoF (SC19) (国際学会)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	高野 了成 (Takano Ryousei) (10509516)	国立研究開発法人産業技術総合研究所・情報・人間工学領域・主任研究員 (82626)	
研究分担者	野村 哲弘 (Nomura Akihiro) (30616602)	東京工業大学・学術国際情報センター・助教 (12608)	
研究分担者	三浦 信一 (Miura Shinichi) (50509974)	国立研究開発法人理化学研究所・計算科学研究センター・技師 (82401)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	遠藤 敏夫 (Endo Toshio) (80396788)	東京工業大学・学術国際情報センター・教授 (12608)	
研究分担者	滝澤 真一郎 (Takizawa Shinichiro) (80550483)	国立研究開発法人産業技術総合研究所・情報・人間工学領域・主任研究員 (82626)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関