

大規模データセンター運用最適化フレームワーク構築に向けて

滝澤真一朗¹ 高野了成¹ 松岡聡¹

1. 研究背景

クラウドサービスやスーパーコンピューティングサービスを提供するデータセンターは年々システムの並列度増加に伴い、大規模化、複雑化している。その結果、電力消費量の増加や、従来からの人手による監視・運用におけるコストが増加している。この問題を解決するために、図 1 に示すような、データセンターを構成する各種コンポーネント（計算機群、電力設備、冷却設備など）が生成するログや負荷状況、各種資源の利用率などの時系列のメトリクス（以降、環境情報）と、利用者によるシステムの利用情報（以降、ジョブ情報）を基に、機械学習/深層学習による性能モデル構築を行い、モデル性能が改善するシステムパラメータを導出し、各種機器を制御し運用改善を行う手法が研究開発されている。

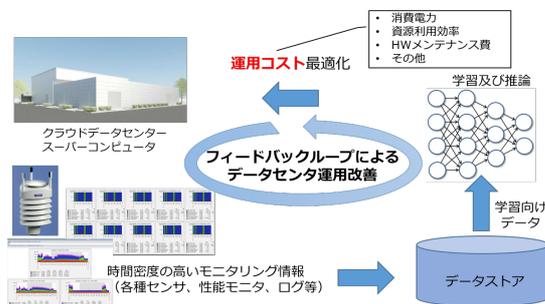


図 1 データセンターのモニタリング情報の活用

例えば Google では機械学習を用いた PUE 予測と、それによる冷却設備の最適化を行なっている [1,2]. Li らは、システムメトリクスを入力とした深層強化学習を用いて、ファイルシステムのパラメータチューニング手法を提案している [3]. 京コンピュータでは過去のジョブの電力消費量の時系列変化を蓄積し、今後実行されるジョブの電力消費量を予測し、最大電力を超えないようジョブスケジューリングを行う機構の研究を進めている [4]. これら

技術は対象としているデータセンター施設に対して個別に開発・実装されているため、他環境への適用性に乏しく、技術移転が困難である。本研究ではこの問題を解決するため、モジュール化された汎用的なデータセンターの運用最適化フレームワークの構築を目指す。

2. 対象とするデータセンター構成

データセンターは、サーバ、ネットワーク機器、各種ストレージ（ディスクベースの共有ストレージ、テープベースのアーカイブストレージ）などの計算機資源や電源設備、冷却設備、照明、住居設備などから構成される。本研究では問題の単純化のために、次に示すミニマムセットの構成を対象とする。

- 計算機資源はサーバ、ネットワーク機器、ディスクベースの共有ストレージから構成される
- 管理サーバ群を除く全サーバが単一のリソースマネージャ、ジョブスケジューラにて管理される
- 電源・冷却設備は自律制御されているものとし、配下の計算機資源の消費電力・熱量に応じて、送電、冷却性能が動的に変動する
- 計算機資源のパラメータ設定を変えることにより、運用改善を図る

3. システム設計

図 1 に示すシステムを構築するためには、次のサブシステムが必要である。

1. 各種環境情報とジョブ情報の収集機構
2. 環境情報とジョブ情報を蓄積するデータストア
3. 環境情報とジョブ情報を元に、運用改善のための性能モデルを構築する機械学習基盤
4. 運用改善のためのシステム設定変更を行うデータセンター制御機構

¹ 産業技術総合研究所

National Institute of Advanced Industrial Science and Technology

1 の環境情報は各種計算機資源が生成する負荷、資源利用率、温度などのメトリクスであり、ジョブ情報は利用者による計算機資源の利用情報（使用した資源種別、資源量、時間など）である。前者は多項目かつ高頻度な時系列データとなるが、後者は多属性を持つ独立したデータの集合となる。このため 2 のデータストアでは、データ種別に応じて適切なデータベースを用いること、及び環境情報とジョブ情報の関連付けが必要である。図 2 に示すとおり、環境情報には時系列データベース、ジョブ情報には関係データベースを用いることを検討している。

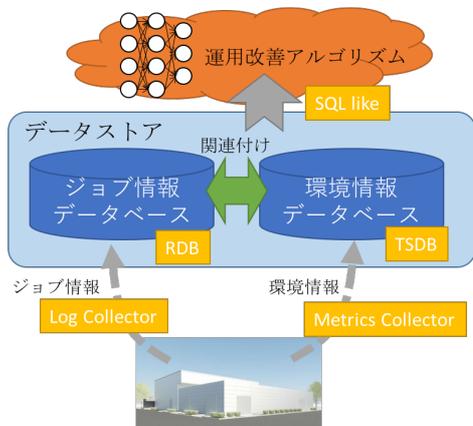


図 2 データストア設計

3 の性能モデルは、データセンターの運用事情に合わせた項目の最適化のために構築する。入力には 1 で収集する環境情報とジョブ情報、出力はシステム設定パラメータとなる。モデル構築のために、scikit-learn などの機械学習フレームワーク、TensorFlow や Chainer などの深層学習フレームワークの各種バージョンが容易に使えるよう、仮想化や Container を用いた機械学習基盤の配備機構を構築する。これは特に、異なるデータセンター・異なるデータセットを元に構築された性能モデルを、別のデータセンターに適用・評価する転移学習のためにも重要である。

3 で出力したシステム設定パラメータを適用することにより運用改善を実施する。運用と並行してリアルタイムに適用可能な項目と、システムの一部を停止した後でない適用できない項目がある。後者については定期メンテナンスなどのタイミングで適用実施する。前者については、システム設定パラメータを入力とし、ジョブスケジューラ及びリソ

ースマネージャを直接制御するためのモジュールを開発する。

4. 現状と今後

現状は上記で示した基本方針・設計を行なった段階であり、フレームワークの開発はこれからである。開発にあたり、データセンター運用を行なっている事業者、大学、研究機関との意見交換を行い、設計にフィードバックさせていきたいと考えている。

産業技術総合研究所では 2018 年度上半期に人工知能の研究開発を目的としたコンピューティングシステム ABCI の運用を開始する。ABCI を対象に、その電力効率向上、資源利用率向上を目的とした運用改善アルゴリズムの開発と実証評価を行う。

謝辞

本研究の成果は、国立研究開発法人新エネルギー・産業技術総合開発機構（NEDO）の委託事業の結果得られたものです。

参考文献

- [1] Jim Gao. Machine Learning Applications for Data Center Optimization. In Google White Paper, 2014.
- [2] Rich Evans and Jim Gao. DeepMind AI reduces energy used for cooling Google data centers by 40%. <https://www.blog.google/topics/environment/deepmind-ai-reduces-energy-used-for/>, 2016.
- [3] Yan Li, Kenneth Chang, Oceane Bel, Ethan L. Miller and Darrell D. E. Long. CAPES: Unsupervised Storage Performance Tuning Using Neural Network-Based Deep Reinforcement Learning. SC17, 2017.
- [4] 宇野篤也, 末安史親, 山本啓二, 肥田元, 池田直樹, 辻田祐一. 消費電力の変動を考慮したジョブスケジューリングの検討. 第 161 回ハイパフォーマンスコンピューティング研究会, 2017.