

Title	ファジィエントロピーを用いたスライディングウィンドウ回帰に基づくコンテナオートスケーリングシステム
Author(s)	横山, 尚弥
Citation	
Issue Date	2023-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18762
Rights	
Description	Supervisor: 田中 清史, 先端科学技術研究科, 修士(情報科学)

概要

Eコマースサービスなどをクラウド上で提供する際に、サービスレベルを一定に保つ一方でコストをなるべく抑えたい場合、不規則に増減するトラフィックに合わせて提供するサーバ資源量を変化させる必要がある。

時系列予測や負荷予測といった分野の研究は数多くの先行研究はあるが、多くは学習させるデータが定常過程や単位根過程である必要がある古典的な時系列予測や、非常に多くのデータやパラメータを用いる深層学習を用いるアプローチであり、学習と推論に多くのデータや時間を要するため数分程度で急激に増減するバーストトラフィックに対して必要なサーバ資源を提供することは困難である。一方、スライディングウィンドウ学習など細かくデータを区切り、逐次的に学習と推論を繰り返すようなアプローチの先行研究もある。このスライディングウィンドウ学習を応用した既存の自動スケールリング手法では、直近の負荷に加えバーストトラフィックを考慮し、近未来に必要なサーバ資源を提供している。

本研究ではスライディングウィンドウ学習による回帰推定をベースに、ファジィエントロピーを用いたバーストトラフィック検知を取り入れることで、わずかなトレンドの変化にも追従できるようにした動的なウィンドウサイズ変更を伴うトラフィック予測法と、複数の負荷指標に対する決定回帰木により推定された負荷から必要なコンテナ数を適用する新たなオートスケールリングシステムを提案する。

このシステムを動作させる基盤としてはコンテナ仮想化技術を採用した。コンテナ仮想化技術は一つのプロセスに対して他のプロセスとは独立したサーバ資源を割り当てる技術であり、必要に応じたサーバ上のコンピュータ資源を柔軟にプロセスに割り当てることができる。また従来のように1つの仮想マシンに対して1つのアプリケーションサーバを構築するのに比べ、コンテナ仮想化技術を用いることでアプリケーションサーバの起動時間を数秒程度に短縮させることができる。

評価ではシミュレーションではなく実トラフィックを用いた4つの比較実験を行う。比較実験ではウェブ上に公開されているトラフィックデータを用いて、負荷生成器にてトラフィックパターンを再現し、実験対象であるシステムに対して出力する。ベースライン手法としては先行研究2つと、コンテナオーケストレーションツールとしてはデファクトスタンダードである Kubnerets のオートスケールリング機能である Horizontal Pod Autoscaling との比較を行う。

結果、提案手法がベースライン手法と比較してリクエスト失敗数を減少させ、理想的なコンテナ数の推移との平均二乗誤差 (Mean Squared Error, MSE) が平均 150.17 ポイント改善した。