

クラウドを用いた分散 Web システムにおけるオートスケールアルゴリズムの改良と評価

17G480 松田 正也（最所研究室）

クラウド環境において負荷量に応じてキャッシュサーバ数を動的に増減させることで、コストを低減しつつ応答性を高める分散 Web システムにおけるオートスケールアルゴリズムの改良および、時系列解析による改良、異なる性能のサーバ利用のための動的重み付け振分アルゴリズムについて述べる。

1 はじめに

クラウドで提供されている仮想マシンをキャッシュサーバとして用いることで容易に負荷分散が行えるようになり、応答性の改善が可能になった。しかし、負荷量に応じてキャッシュサーバ数を増減できなければ応答性の不足あるいは過剰なコストが発生する。そこで、当研究室ではクラウド環境を対象にした負荷量に応じて動的に仮想キャッシュサーバ (VC サーバ) 数を増減させることで、応答性を確保しつつ運用コストを低減する分散 Web システムを開発している [1][2]。本稿では、先行研究で開発した稼働率を用いたオートスケールアルゴリズム (以下、稼働率アルゴリズム) を再評価し、判明した問題点を解決するスループットを用いたオートスケールアルゴリズム (以下、スループットアルゴリズム) の開発とその評価および、時系列解析による予測精度の向上、異なる性能のサーバ利用のための動的重み付け振分アルゴリズムについて述べる。

2 分散 Web システムの概要

図 1 に分散 Web システムの概要を示す。本システムは拡張ロードバランサ、大元のコンテンツを提供するオリジンサーバ、オリジンサーバから取得したキャッシュを提供する VC サーバ群から構成される。拡張ロードバランサはソフトウェアロードバランサに、サーバの負荷量を監視する負荷監視機能、負荷量に応じて VC サーバを起動・停止するキャッシュサーバ管理機能、VC サーバ数の増減に合わせてアクセスの振分先を更新する振分先設定機能を持つ拡張プログラムを追加したものである。負荷量の監視および VC サーバの増減は拡張プログラムの機能で行い、リクエストの制御はソフトウェアロードバランサの機能を用いて行う。

3 稼働率アルゴリズムの再評価

卒業研究 [2] で開発した稼働率アルゴリズムの評価実験を行った結果を図 2 に示す。稼働率の上昇を検知しスケールアウトした時点 (50 秒と 200 秒付近) で、応答時間が既に大きく伸びている。このとき、平均スループットは秒間リクエスト数の上昇に比例して上昇していることから、スループットを用いることでより早期に負荷上昇を検知できると考えた。

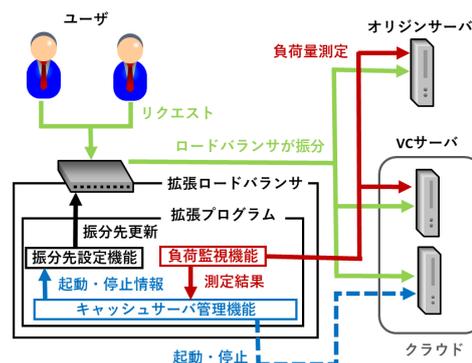


図 1: 分散 Web システムの概要

4 スループットアルゴリズムの開発

前節の結果をふまえて、スループットアルゴリズムを開発した。稼働率アルゴリズムと同じ実験環境で実験した結果を図 3 に示す。応答時間が 5 秒を超えることは無く、大きく応答性が改善されたことを確認できた。しかし、スループットアルゴリズムでは応答性の改善のためにスケールアウト性能の向上に着目していたため、秒間リクエスト数が低下しても過剰にサーバが稼働しており、改善が必要である。

5 時系列解析の利用検討

クラウドサービスにおける仮想マシンなどは従量課金制であるため、スケールイン性能を向上させコストを抑えることは重要な課題であるが、前節に示したように実装しているアルゴリズムでは困難である。このため、時系列解析を用いてリクエストの周期性を予測し、予め適切な数の仮想キャッシュサーバを起動するオートスケールアルゴリズムについて検討した。

香川大学総合情報センターのホームページへのアクセスログに対して、時系列解析モデルである ARMA モデル [3] を用いて予測を行った結果を図 4 に示す。いずれの予測値も 1 日目のみ優位であり、1 週間のデータを用いた予測値が最も誤差が少なかったが、ある程度誤差が出ることが分かった。これより、時系列解析とスループットアルゴリズムを組み合わせ、事前に予測値に合わせた台数の VC サーバを起動し、予測値と実測値の誤差がある程度大きくなった場合、スルー

プットアルゴリズムによりオートスケールすることを検討した。

6 動的重み付け振分アルゴリズムの開発

これまでの研究では単一性能の VC サーバを利用することを想定してきたが、様々な性能の仮想マシンを容易に利用できるクラウド環境に対応しているとはいえない。そのため、性能の異なる VC サーバに対して負荷状況に応じて動的に重み付けを行い、振分量を調整するアルゴリズムを開発した。表 1 に示す性能の VC サーバを用いて実験を行った結果を図 5 に示す。均等な振分状態で開始したが、実験開始後すぐに CPU 性能毎に重みが変わっており、アルゴリズムの有効性を確認できた。

7 まとめ

スループットを用いたオートスケールアルゴリズムを開発し、評価実験により稼働率を用いたオートスケールアルゴリズムよりも応答性の改善ができることを確認した。スケールイン性能を向上させ、コストを抑えるために時系列分析の利用を検討した。異なる性能の VC サーバを用いるために動的重み付け振分アルゴリズムを開発し、有効性を確認した。

今後の課題として、動的重み付け振分アルゴリズムと時系列解析による予測を、スループットアルゴリズムに組み合わせクラウド環境への対応を進めることや、ロードバランサへのリクエスト流入量の利用、負荷測定プログラムの改善がある。

参考文献

- [1] 堀内辰彦, “分散 Web システムにおけるオートスケールアルゴリズムの改良と評価”, 香川大学 修士論文, 2016
- [2] 松田正也, 最所圭三 : “クラウドに適した Web システムにおけるオートスケールアルゴリズムの改良と評価”, 香川大学 卒業論文, 2017
- [3] George Box, Gwilym M. Jenkins, ”Time Series Analysis: Forecasting and Control, second edition,” in San Francisco : Holden-Day, 1976

表 1: 仮想マシンの性能

サイズ名	CPU コア数 [コア]	メモリ量 [GB]
DS1_v2	1	3.5
DS2_v2	2	7
DS3_v2	4	14
DS11_v2	2	14
DS12_v2	4	28

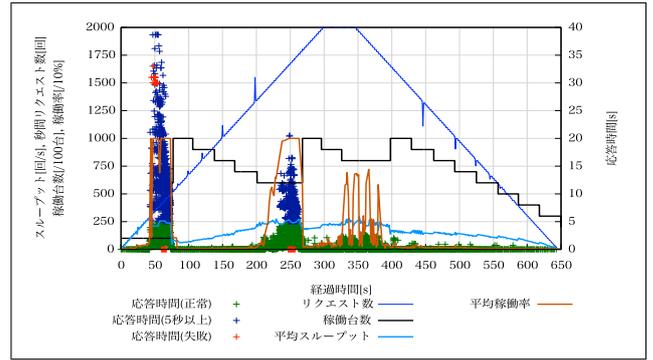


図 2: 稼働率アルゴリズムの実験結果

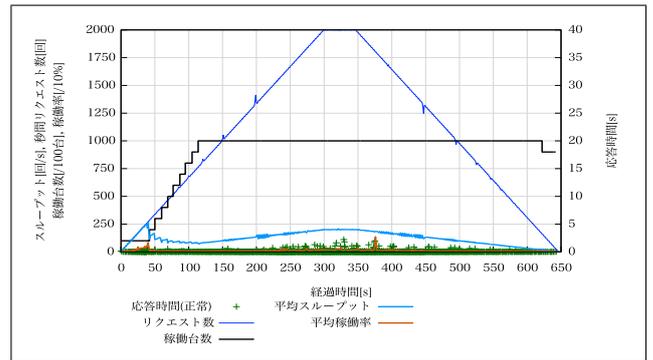


図 3: スループットアルゴリズムの実験結果

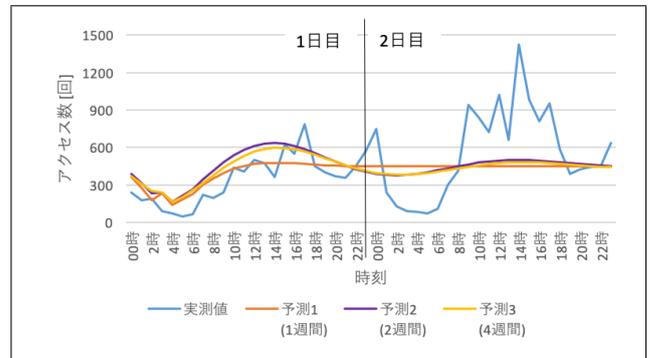


図 4: 実測値と時系列解析による予測値

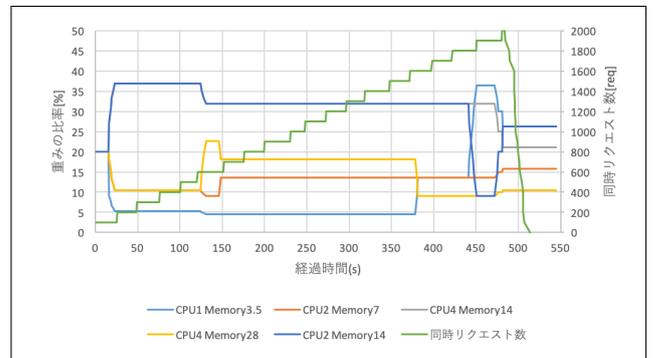


図 5: 動的パターンの重みの比率の変遷