

クラウド上のコンテナで動作する Web アプリケーションの 性能推定手法の評価

1265089 荒武 佑磨 【高度プログラミング研究室】

Evaluation of Performance Estimation Methods for Web Applications in Containers on the Cloud

1265089 Yuma Aratake 【High-Level Programming Lab.】

1 はじめに

近年、クラウド上でのコンテナベースのアプリケーション運用が増加している。アプリケーションの本番運用においては、アプリケーションの性能を正確に把握し、適切なコンテナインスタンスを選択することが重要である。しかし、クラウドは性質上、その性能がランダムに変動してしまう [1] ため、クラウド上で正確な性能指標を得るのは非常に困難である。Sen らが発表した論文 [1] では、性能変動のある仮想マシン環境において性能指標を得る方法として、ブートストラップ法やブロックブートストラップ法を用いた手法が提案された。特に、ブロックブートストラップ法は、性能変動の内部依存性を考慮することができ、より高い精度で性能予測を行えることが示された。そこで本研究では、クラウド上で運用される Web サイトの性能推定を目的とし、具体的な負荷テストとその結果に対して、ブロックブートストラップ法やブートストラップ法を適用する。これにより負荷テストデータからクラウド上の Web アプリケーションの性能推定を行う手法を評価する。

2 性能推定手法

先行研究で用いられた性能推定手法では、テストデータをブロックブートストラップ法 (BBS) やブートストラップ法 (BS) を用いてサンプリングを行う。サンプリングするデータの単位は 1 日分として、一日毎の精度の誤差を調べるため、初めは 1 日分と 2 日分のサンプリング結果の比較、その後は 1 日分増やして n 日分と $n+1$ 日分を比較というように 1 日分ずつの増やしてサンプリング結果を比較する。比較した結果の誤差が事前に定義した最大許容誤差より小さい場合、性能結果は十分に正確であるとみなされ、性能試験を停止する。最後に、最大許容誤差を下回った時点でのテストデータとグランドトゥールズ用のデータを用いて相対誤差の計算を行い、精度を計算する。

3 負荷テストの計画と実施

本研究の目標は、性能推定精度と、クラウド環境におけるテストコストのバランスをとることである。そのた

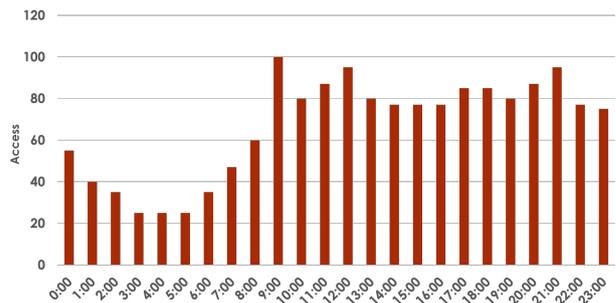


図 1 負荷量の推移

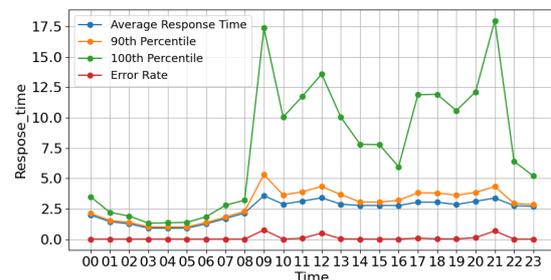


図 2 1 日の各性能の推移 (15 日目)

めに、性能推定結果が正確となる十分な負荷テスト量を、確実かつコスト効率良く特定する手法を提供する。

本研究では、負荷対象の Web アプリケーションとして、ログイン機能や、特定のデータ群の一覧をデータベースから取得して表示するような一般的な機能を持った Web アプリケーションを実装した。負荷テスト時にかかる負荷量の推移は図 1 の通りである。実際の負荷を想定して負荷の量を 1 時間ごとに変動させ、ピーク時で Web アプリケーションからのレスポンスタイムが 3 秒程度になるように負荷を設定した。負荷テストを実施した期間は 4 週間で、頻度としては毎日一時間毎、1 回の負荷テストを行う時間は 3 分間とした。得られた 4 週間分の負荷テストデータは、性能推定を行うため、2 週間分のテストデータと 2 週間分のグランドトゥールズ用のデータに分割して精度推定に利用する。

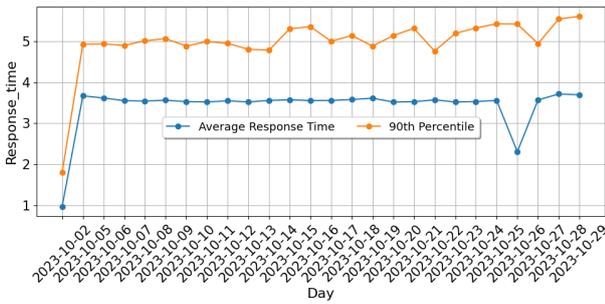


図3 各日の9時時点の性能

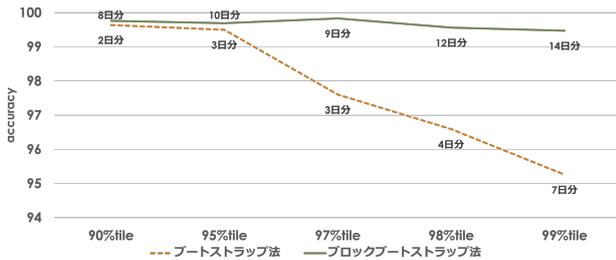


図4 1日の性能の推定結果

負荷テスト4週間の内の1日分のデータを図2に示す。前述した負荷量の推移と比較して、実際に負荷量に合わせて性能が変動していることが分かる。特定の時間の日別の推移を図3に示す。初日は少し負荷テストが安定しない時間帯が存在したが、その後は安定して負荷テストが行えていることがわかる。また、負荷が多い時間帯の方が、少ない時間帯に比べ90%tileと平均値の差が開いており、性能変動が大きいことが示されていた。

4 性能推定手法の評価

本研究で推定する性能は、一日を通してと、各時間におけるレスポンスタイムの各パーセンタイル値とした。これらは、一般的にWebアプリケーションのパフォーマンスの一貫性の把握や、システムのボトルネックの特定を行う際に測られる性能である。

図4は、一日を通しての負荷テストデータに対して、ブートストラップ法とブロックブートストラップ法を用いて最大許容誤差5パーセントで性能推定を行った結果を示している。図4が示すように各パーセンタイル値の性能推定においてブロックブートストラップ方がブートストラップ法よりも高い精度で性能の推定ができていたことが示された。しかし、推定に必要な日数はブートストラップ法に比べ多くなっていることが分かった。

表1は各時間の性能データに対して、ブートストラップ法とブロックブートストラップ法を用いて最大許容誤差5パーセントで性能推定を行った結果を示している。表1が示すように、負荷量が少ない時間帯に比べ、負荷量が多い時間帯の方が性能の推定に必要な日数が多いことが分かった。また、ブロックブートストラップ法

表1 特定の時間帯の性能推定結果

%tile	1:00 BBS	1:00 BS
90%tile	99.97%/2日	99.96%/2日
95%tile	99.55%/2日	99.55%/2日
97%tile	98.70%/2日	98.70%/2日
98%tile	98.96%/3日	98.96%/3日
99%tile	99.25%/6日	99.30%/5日
%tile	9:00 BBS	9:00 BS
90%tile	97.52%/3日	95.99%/2日
95%tile	96.13%/3日	95.65%/2日
97%tile	98.14%/5日	99.20%/4日
98%tile	97.49%/5日	97.49%/5日
99%tile	99.04%/7日	99.02%/6日

とブートストラップ法で推定精度に大きな違いは見られなかった。このことから、各時間の精度を推定する場合は、性能の推定に必要な日数が少ないブートストラップ法の方が性能推定に向いているという事が分かった。

5 まとめ

本研究では、クラウド上で運用されるWebサイトの性能推定を目的とし、具体的な負荷テストとその結果に対して、ブロックブートストラップ法やブートストラップ法を適用し、負荷テストを早期終了するための適切なタイミングを推定する手法を提案した。

その結果、本研究で得られたような、時間的な内部依存性の高い一日を通しての負荷テストデータに対してはブロックブートストラップの方が日数はかかるが、ブートストラップ法に比べ高い精度で性能推定が行えることが分かった。これにより、99%tileのように値のブレ幅が大きい性能を調べたい場合でも、負荷テストの誤った早期終了を防ぐことが可能であることが示された。

時間毎の性能推定では、ブロックブートストラップ法とブートストラップ法で推定精度に大きな差は見られなかったが、ブロックブートストラップ法の方が推定に日数を多く使っていることが分かった。これにより、時間的な内部依存性の少ない時間毎の性能推定においてはブートストラップ法の方が適していることが示された。

参考文献

- [1] Sen, H. Tianyi, L. Palden, L. Jaewoo, L. In, K. and Wei, W. :Performance Testing for Cloud Computing with Dependent Data Bootstrapping, *2021 IEEE/ACM International Conference on Automated Software Engineering* (2021).
- [2] 荒武 佑磨, 松崎 公紀 :クラウド上のコンテナで動作するWebアプリケーションの性能推定手法の評価 第65回 プログラミング・シンポジウム (2024).