

# 分散 Web システムにおけるサーバの応答時間を用いた性能測定機能の開発

18T310 岡本 涼太郎（最所研究室）

当研究室では，サーバの負荷量に応じて動的にキャッシュサーバを増減させることで応答性の確保とコスト軽減を満たす分散 Web システムを開発している．本項では，サーバの応答時間を用いた新しいスループット測定機能の提案及び，動的なサーバ性能測定方法についても述べる．

## 1 はじめに

当研究室では，ソフトウェアロードバランサに，サーバを監視し，負荷量に応じた動的なキャッシュサーバ数の増減及び処理能力の柔軟な制御を行う機能を追加することでこれらの問題点を解決する分散 Web システムを提案している [1]．本研究では，負荷の指標として新たに各サーバのリクエストに対する応答時間を用いてスループットの測定を行う機能を開発した．更に，動的に最大スループットが変化するサービスに対して有効であるかの検証を行った．

## 2 分散 Web システム概要

図 1 に当研究室で開発している分散 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，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仮想キャッシュサーバ群から構成されている．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，負荷量に応じて仮想キャッシュサーバを起動・停止するキャッシュサーバ管理機能，仮想キャッシュ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更新機能を持つ拡張プログラムを追加したものである．負荷量の監視及び仮想キャッシュサーバの増減は拡張プログラムの機能で行い，リクエストの振り分けはソフトウェアロードバランサの機能で行う．

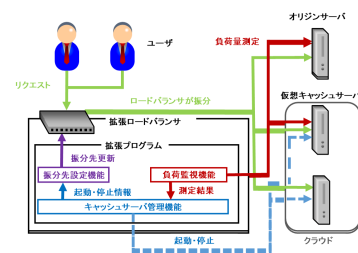


図 1: 分散 Web システムの概要

## 3 スループットと応答時間の関連性

負荷量を増やしながらスループットと応答時間を測定する実験を行った．この実験でロードバランサのログから得られた結果を図 2 に示す．なお，結果は省略するがロードバランサ側で測定されたスループットとサーバ側で測定されたスループットはほぼ等しいという結果が得られた．しかし，サーバ側の応答時間は，負荷に関係なく低い値を示していた．このことからロードバランサ側の応答時間がサーバの上昇見積りに使用することができると考えられる．図 2 を見ると，スループットが 200 を超えた時点で応答時間が急激に増加している．それ以降はスループットは増えていなかった．そのため，応答時間がここを超えないようにする必要があり，この時点でのスループットを上限スループットとして用いることにした．ロードバランサ側でスループットと応答時間を測定することができるようになったことで，これまで問題になっていたサーバの過負荷時にサーバからスループットを取得出来ないという問題を解決

することも出来るようになった。

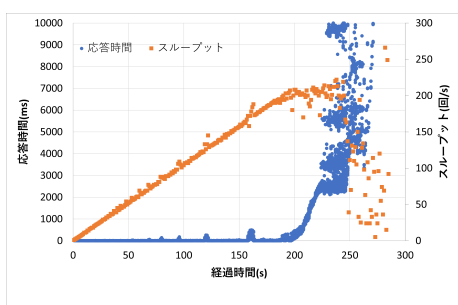


図 2: サーバのロードバランサでの応答時間とスループット

#### 4 サーバの応答時間を用いた上限スループット測定機能

ロードバランサの一定区間のログから、その区間の平均応答時間とスループットを算出する。上限スループットの測定時には、各サーバを起動し、徐々に負荷を増やしながらか以下の処理を行う。

1. 単位時間内のロードバランサのログの取得
2. 取得したログから対象サーバのログのみに絞り平均応答時間を算出する
3. 平均応答時間とその基準値を比較する。(基準値を超えていなければ2に戻る)
4. 取得しているログを基に、スループットを算出し、上限スループットとする

#### 5 上限スループット測定機能を用いたサーバの動的な性能測定及び振分量調整機能

図 3 に示す環境を用いて、オリジンサーバに対する上限スループットの測定及び振分量調整機能について説明する。

応答時間が基準値を超えた場合、この時のスループットを上限スループットとし、オリジンサーバへの振分量を少なくする。

応答時間が一定の値以下である状態が続いた場合、軽負荷とし、オリジンサーバへの振分量を多くする。

この機能の有効性を確認する実験を行った。

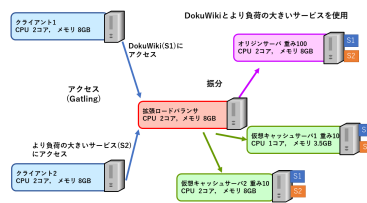


図 3: 上限スループットと応答時間

各サーバに負荷の軽いページ (スループット 180) にアクセスするクライアント 1 と負荷の重いページ (スループット 80) にアクセスするクライアント 2 を用意し、二つのクライアントからアクセス量を変えながらアクセスを行った結果を図 4 に示す。

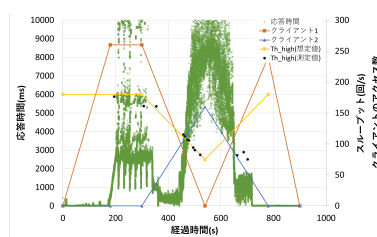


図 4: 上限スループットと応答時間

上限スループット (黒丸) が想定値 (黄色線) とほぼ一致しているため、上限スループットの変化にも追従できていることがわかる。しかし、負荷が軽い場合には追従できていなかったため、改善が必要である。

#### 6 まとめ

新たな負荷の指標として応答時間を用いた、サーバの上限スループット測定機能を開発し、性能測定が可能であることを確認した。負荷の上昇に対する上限スループットの変化に追従できていることを確認した。今後は軽負荷時にも追従するようにする必要がある。

#### 参考文献

[1] 畑 智裕, “分散 Web システムにおける複数パラメタを用いたオートスケールアルゴリズム”, 香川大学修士論文, 2021