

分散 Web システムのための TCP バックログと 応答時間の分布を用いた負荷監視機能の開発

20T306 小林拓実 (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費用対効果の高い分散 Web システムを開発している。本稿では、TCP バックログと応答時間の分布を基にした負荷監視機能とそれを組み込んだ分散 Web システムの開発及び評価について述べる。

1. はじめに

インターネット上で様々なサービスが Web で展開されており、それらにアクセスが集中することで、サーバからの応答時間が長くなることもある。当研究室では、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、それらの数を負荷量に応じて増減できる費用対効果の高い分散 Web システムを開発している。

先行研究では、負荷の指標として応答時間や CPU、メモリ使用率などを用いていた。この方法では、負荷が上がっていると判断した時点で、サーバが過負荷寸前の状態になっていた。そのため、その時点で仮想マシンに対して起動指示を出しても起動完了までは応答時間の劣化をあまり改善できないという問題があった。本研究では、過負荷になり応答時間が増加する前に TCP バックログに接続が溜まることに着目し、TCP バックログを負荷指標として用いる負荷監視機能を開発した。

本稿では、TCP バックログと応答時間の分布を用いる負荷監視機能とそれを組み込んだ分散 Web システムの開発及び評価を述べる。

2. 分散 Web システムの概要

当研究室で開発している分散 Web システムは、図 1 に示すように、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キャッシュサーバ群から構成される。なお、図中の TCP バックログに関

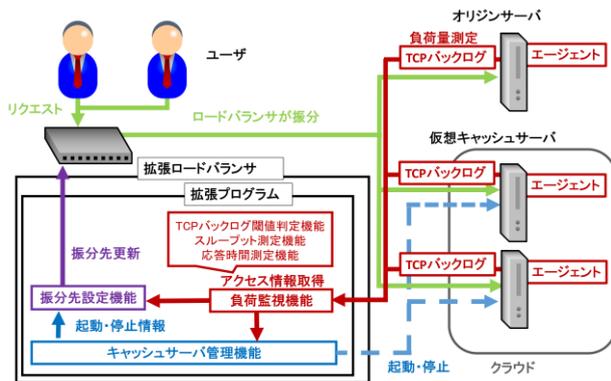


図 1 分散 Web システムの概要

する部分は本研究で開発する機能である。

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、キャッシュサーバの増減に合わせてアクセスの振り分先を更新する振り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。

3. TCP バックログを用いた分散 Web システム

3.1 TCP バックログ

TCP バックログはサーバに accept されていない接続要求数である。サーバに負荷がかかると接続要求はすぐに accept されず待ち行列に繋がる。その状態はサーバの処理が間に合わず、負荷がかかっていると判断できる。この待ち行列に繋がる accept 待ち数を負荷指標として用いる。

3.2 負荷監視機能

各サーバに TCP バックログを取得し、拡張ロードバランサに送信するエージェントを設置する。受信した TCP バックログやロードバランサによって記録されるアクセスログを基に負荷を監視する。

TCP バックログ閾値判定機能：TCP バックログの値やその移動平均がそれぞれの閾値を超えているかどうかを判定する

スループット測定機能：アクセスログを用いて各サーバのスループットを測定する

応答時間測定機能：アクセスログを用いて各サーバの応答時間とその分布を測定する

3.3 振り分先設定機能

起動したサーバのうち、最後に起動したサーバ以外が過負荷にならないようにアクセスを配分する。本研究では、サーバが過負荷にならないスループットを運用スループットとしている。

3.4 キャッシュサーバ管理機能

最後に起動したサーバの TCP バックログの値またはその移動平均や応答時間の分布がそれぞれの閾値を超えていれば、新たにキャッシュサーバを起動し、アクセスを振り分ける。逆に、最後に起動し

たキャッシュサーバへの振り分けが一定時間なければ停止する。

4 評価実験

4.1 評価基準

運用スループットが変動する実験を行う。TCP バックログを用いることで、変動する運用スループットに追従でき、かつ応答時間を一定以上確保できることを確認する。ここでは応答時間が 1000ms 以下であれば応答時間を確保できているものとする。TCP バックログと応答時間の関係性を調べる予備実験の結果より、負荷がかかった際の TCP バックログが 2 以上になっていたため、これを新規キャッシュサーバの起動条件として設定した。さらに負荷がかかり始めると TCP バックログの出現頻度が増えたため、4 秒間の移動平均が 0.75 以上も条件に加えた。またアクセス数と応答時間の関係性を調べる予備実験の結果より、負荷がかかると応答時間の長いアクセスの割合が増えたため、TCP バックログの 10 秒間の移動平均が 0.2 以上かつ応答時間の最頻値の分布が全体の 80% 以下も条件に加えた。以上の条件を満たした時のスループットを最後に起動したキャッシュサーバの運用スループットにする。

TCP バックログが 2 以上または TCP バックログの 4 秒間の移動平均が 0.75 以上になった場合に、この条件を満足しなくなるまで徐々に振分量を減らす。この時のスループットを新たな運用スループットとする。

4.2 評価手法

本実験で、各サーバは CPU に負荷のかかるサービス S1, S2 を提供している。オリジンサーバ単体では、S1 は毎秒 230 アクセス、S2 は毎秒 140 アクセスを処理できる。

初期状態ではオリジンサーバのみが起動しており、ロードバランサの重みの比はオリジンサーバが 100、キャッシュサーバが 0 である。

負荷テストツール Gatling を使用して以下のシナ

リオを実行する。S1 へ毎秒 240 アクセスまでかけ、一定期間アクセスを保ち、期間経過後アクセスを減らす。次に毎秒 200 アクセスまでかけ、200 アクセス到達後にアクセスを減らす。S2 へ実験開始 190 秒経過後から、毎秒 120 アクセスまで増加させる。120 アクセス到達後にアクセスを減らす。

4.3 評価結果

オリジンサーバとして Azure の Standard D2s v3, キャッシュサーバとして Standard DS1 v2 を用いた実験結果を図 2 に示す。①アクセスの遷移、②各サーバへのアクセス数、③各サーバの TCP バックログ、④各サーバの最大応答時間を表す。①の水色が S1 へのアクセス数、黄色が S2 へのアクセス数、紫色が総アクセス数を表す。キャッシュサーバの起動指示が発令された時点を③に矢印で示す。

オリジンサーバで毎秒 180 アクセスが発生した時点で TCP バックログが閾値を超え、キャッシュサーバの起動指示が発令された。キャッシュサーバの起動完了時、オリジンサーバには毎秒約 220 アクセスがあり、上限を超える前に起動が行えた。

S2 へのアクセスが始まり、オリジンサーバへの負荷が増え、TCP バックログが閾値を超えたため、オリジンサーバへのアクセス数を減少させた。この時点のオリジンサーバの最大応答時間は 54ms であった。さらに負荷が増え、2 台目のキャッシュサーバが起動された。

アクセス数の減少に伴い、キャッシュサーバへの振り分けが減少し、キャッシュサーバが停止された。その後アクセスが増え、700 秒付近で再びキャッシュサーバが起動された。

これらの結果より、TCP バックログを用いることで、変動する運用スループットへの追従及び応答時間を低下させずに振り分けが行えることが分かった。

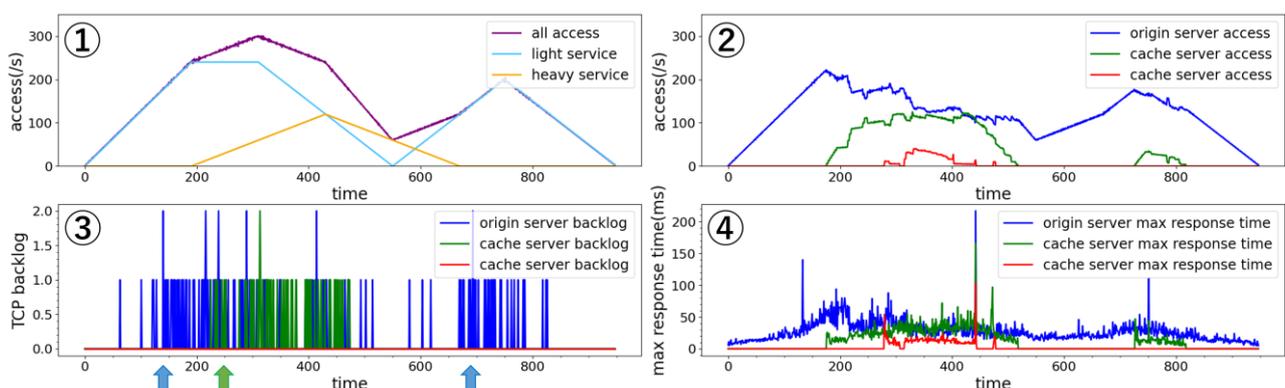


図 2 実験結果