

複数リソースからなるページを用いた分散 Web システムの評価

17T289 吉田 悠(最所研究室)

当研究室で開発している分散 Web システムに対して、複数リソースからなるテスト用 Web ページを用いて実用性の評価を行う。

1 はじめに

当研究室では、サーバの負荷量が大きくなった時にキャッシュサーバを動的に起動・停止することで、サーバにかかる負荷を分散する分散 Web システムを開発している[1][2]。

これまで、分散 Web システムの評価実験は単一のリソースからなる Web ページを用いて行われていた。しかし、実際にインターネット上にある Web ページのほとんどは、画像ファイルなどの複数のリソースを組み合わせで構成されている。そのため、当研究室で開発した分散 Web システムが実用的なものであるかどうか知るために、複数のリソースからなる Web ページに対する評価を行った。

本稿では、評価実験およびその結果について述べる。

2 分散 Web システム

図 1 に分散 Web システムの構成を示す。

本システムはコンテンツの提供元であるオリジンサーバと、負荷状況に応じて起動・停止されるキャッシュサーバ、リクエストを起動中のサーバに分散させるロードバランサで構成されている。ロードバランサに、負荷状況に応じてキャッシュサーバの起動停止および、それらへのリクエストの振り分けを指示するプログラムを追加している。

3 テスト用 Web ページ

テスト用 Web ページは、インターネット上にある Web ページの構造を調査した結果に基づいて構築したページの構造を表 1 に示す。ペ

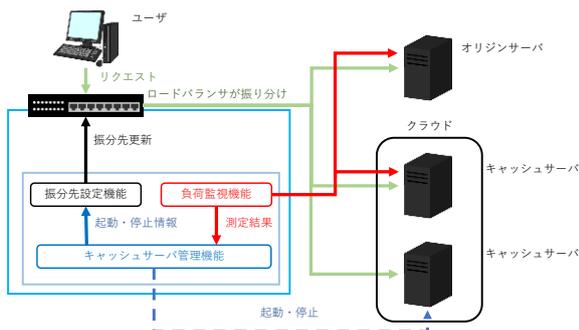


図 1: システムの構成

ージ 1 は本学 moodle を、ページ 2 は YouTube トップページを、ページ 3 については Amazon を、ページ 4 は Yahoo!ショッピングをページ 5 は Twitch を参考にして構築した。テストページは画像ファイル群と、その画像ファイルをすべて参照する html ページによって構成されている。

表 1: テスト用 Web ページ

ページ名	全体ファイルサイズ (MB)	画像ファイル数	ファイル 1 個当たりのサイズ (MB)
ページ 1	2	10	0.2
ページ 2	20	1	20
ページ 3	4	100	0.04
ページ 4	8	100	0.08
ページ 5	8	50	0.16

4 単一 Web サーバでの応答性の調査

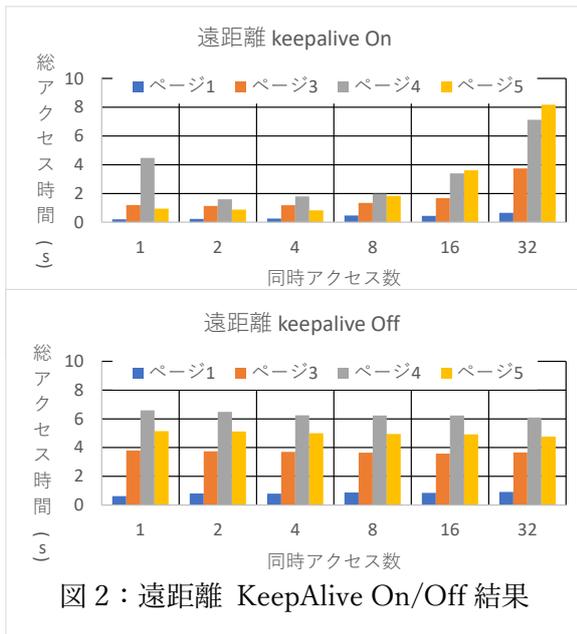
分散 Web システムに対する評価実験を行う前に、単一の Web サーバに対しての評価実験を行う。本実験では、クライアントと Web サーバ間のインターネット上の距離が応答時間に与える影響と、KeepAlive On/Off が応答時間に与える影響を確認する。表 2 に実験に使用する機器の性能を示す。

表 2: 実験使用機器

	台数 (台)	CPU コア数	メモリ量 (GiB)
サーバ	2	1	3.5
クライアント	1	1	3.5

実験では、Microsoft の Azure が提供しているクラウドプラットフォーム上にサーバとクライアントを構築した。サーバを Azure の西日本地域に、クライアントを西日本と東日本にそれぞれ設置した。これにより、西日本のサーバへアクセスしたときは近距離、東日本のサーバへアクセスしたときは遠距離となる。本実験では、遠距離と近距離のサーバに対して KeepAlive On/Off の設定でアクセス実験を行う。実験の流れは以下のようになる。

1. 指定したプロセス数でページ全体を取得する wget を実行するプログラムを用意
2. そのプログラムを 1,2,4,8,16,32 のプロセス数で同時に実行



3. 応答時間の平均をグラフにプロット
 遠距離の実験結果を図 2 に示す。KeepAlive On の時は同時アクセス数と応答時間が比例しているが、Off の時は同時アクセス数に影響されずほぼ一定になっていることがわかる。近距離については、KeepAlive On/Off に関係なく遠距離 KeepAlive On の時とほとんど同じ傾向であった。これは、KeepAlive Off 時は、TCP のコネクション確立のオーバーヘッドやスロースタートが影響していると考えられる。

5 分散 Web システムの評価実験

分散 Web システムに対して評価実験を行った。

本実験での評価項目を以下に示す。

- 稼働率：サーバが処理に割いているリソースを示す。本実験ではサーバの最大クライアント数に対する現在のクライアントの割合。
- 応答時間：ユーザが送ったリクエストが返ってくるまでの時間。
- 通信量 (bps)：サーバにおける単位時間当たりの通信量

本実験で使用した機器の情報を表 3 に示す。

表 3：実験使用機器

	台数(台)	CPU コア数	メモリ量 (GiB)
サーバ	5	1	3.5
クライアント	5	1	3.5
ロードバランサ	1	8	28

本実験では、先行研究と実験環境を同じにするために同一のサブネット内で行った。そのため、通信遅延は非常に小さいと考えられる。実験の流れは以下ようになる。

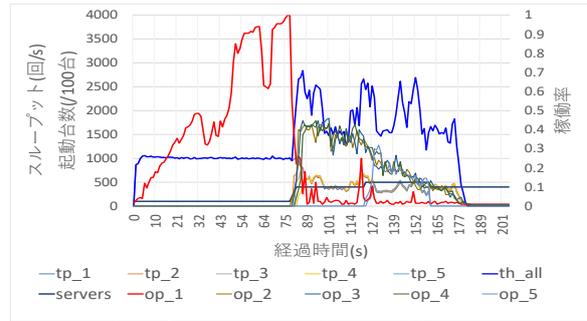


図 3：稼働率とスループットに関するグラフ

- 1 クライアントを 5 台起動する
- 2 起動したサーバは連続して wget を実行するプロセスを決められた実行間隔で起動。
- 3 起動してから実行時間分の時間が経過したプロセスから停止していく。

図 3 に実行間隔を 1.5 秒、実行するプロセス数を 50、実行時間を 105 秒にした時の結果を示す。図 3 からわかるようにオリジンサーバの稼働率(赤線)は同時アクセス数に比例して伸びていた。一方、全体のスループット(青線)はすぐに一定になっており同時アクセス数の影響を受けなかった。紙面の都合で図を載せることができなかったが、応答時間については、稼働率と同様の傾向を示していた。さらに、キャッシュサーバの台数が増えても全体のスループットが伸びなかった。この時、どのキャッシュサーバも通信量が 700[Mpbs] 前後であった。このことから、サーバ側の通信路がボトルネックであると考えられる。

4 おわりに

本稿では、以下のことについて述べた。

- 単一サーバ上の、複数のリソースからなる Web ページに同時アクセスした時の応答時間について調査した。
- 同じページを用いて分散 Web システムを評価した。

複数の静的リソースからなるページへ低遅延のネットワークでアクセスすると、サーバの通信路がボトルネックになることがわかった。

今後の課題として通信路がボトルネックになることへの対応や、CGI によって生成されるページを用いた評価実験を行うことなどが考えられる。この実験を行うことで、条件に応じて変化する画像ファイルを持つ Web ページを想定した評価実験を行うことができる。

5 参考文献

- [1] 畑 智裕 “分散 Web システムにおけるロードバランサを用いたオートスケールアルゴリズムの改良と評価”，香川大学 卒論 2020
- [2] 松田 正也 “クラウドに適した分散 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”，香川大学 修論 2019