

複数のミラーサーバへの分散アクセスによる 効率的なアクセス機構に関する研究

02G475 藤澤 弘 (最所研究室)

パーソナルコンピュータの急激な普及およびインターネットの急激な広がりには多くの Web ページが登場させ、特にアクセスが多いサイトでは同じサービスを行うミラーサーバを用意している。本研究においては、これらのミラーサーバを有効活用し、特に多くの画像などが表示される Web ページを表示するための効率の向上を図るための信頼性の高い接続および柔軟な接続を実現するプロキシ機構を提案し、その評価を行った。その結果、提案した機構の有効性を確認できた。

1 はじめに

アクセスが多いサーバでは、アクセスの集中を避けるためにサーバをミラーリングしている。これら複数のミラーサーバを用いた高効率高信頼の Web アクセス機構を提案した [1]。提案機構を実装した結果、Web アクセスの信頼性の向上と通信時間の短縮を確認できたが、通信時間は期待したほど短くならなかった。

本研究では、この問題に対処するためにプロキシ機構に HTTP の新しい機能を追加する。特に Web ページを開くための時間を短縮するため HTTP の新しい機能である持続的接続、HTTP パイプラインなどの機能を加えた機構を実装し、インターネット上のミラーサーバおよび大学研究室内のミラーサーバで評価を行う。

2 概要

プロキシ機構に以下の機能を持たせることにより信頼性の向上およびアクセスの効率化を行う。

1. 高信頼アクセス機構
2. 柔軟なアクセス機構

- 選択接続
- 振り分け接続 (図 1)
- 分割接続
- 多重接続

本研究では特に、図 1 に示すように連続したリクエストを、複数のミラーサーバに同時に振り分けることにより Web ページを開くためにかかる時間を短縮する振り分け接続機構を扱っている。

3 HTTP の追加機構

HTTP 1.1 で追加された機能の主なものは以下の 2 である。

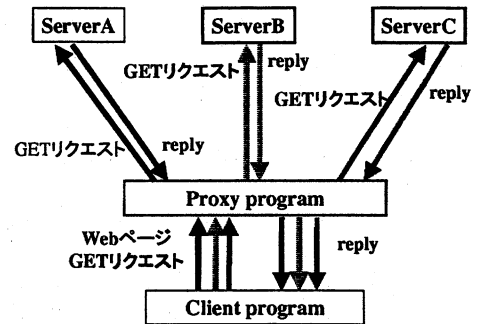


図 1: 振り分け接続

1. 持続的接続 (Persistent Connection)

HTTP 1.0 では、Web サーバに対する 1 回の応答が終了した時点で即座にそのサーバに対する TCP コネクションは切断される。HTTP 1.1 では、持続的接続の機能が追加された。この機能により、従来では HTML のソースコードや画像ファイル、音声ファイルなど、各オブジェクトごとに個別に TCP コネクションをオープンしていたが、これを 1 つの TCP コネクションで連続して取り扱うことができる。同じサーバに対して連続した要求が送られた場合、従来に比較して高速なデータ転送が可能である。

2. HTTP パイプライン

複数の HTTP リクエストを、対応する応答を待つことなくソケットに同時に書き出すことを許可する機能である。この機能を使うことにより、1 つの HTTP リクエストの応答を待つことなく、同時に複数のリクエストを送信することができる。

以上の機能を追加することにより、Web アクセスの効率を向上できる。

4 通信プロトコルの設計

本節では、以上の機能を追加したプロキシプログラムでの通信手順を設計する。ページの都合により、HTTP パイプラインの要求が来たときのみの動作の

概念を図2に示す。

最初の要求(①)でページを読みだし、クライアントに送る。この後、プロキシ機構は次のHTTPパイプラインを用いるアクセスに備え、別のミラーサーバとのTCPコネクションをオープンしておく。クライアントはそのページを表示するために、複数のデータが必要なことを判断し、プロキシにHTTPパイプラインで要求する(②)。プロキシは、HTTPパイプラインの要求を分解し、同時に複数のミラーサーバに対して送る(③、③'、③'')。その後、サーバから帰って来た応答をHTTPパイプラインリクエストの順番通りにクライアントに渡す(④、④'、④'')。

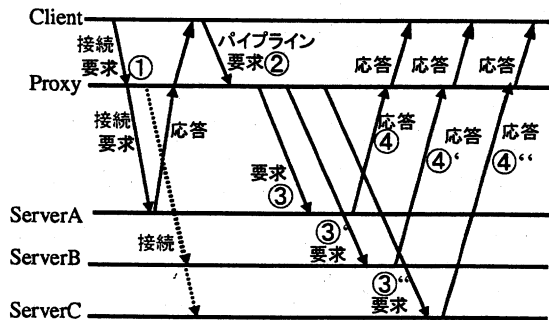


図2: HTTPパイプラインを使った振り分け接続の通信プロトコル

5 評価

開発したプロキシプログラムに時刻を測定する機能を加え、それぞれの通信動作の時刻を測定した。評価にあたり、実際のインターネット環境で運営されているミラーサーバであるグーグル(www.google.co.jp)と、香川大学最所研究室内でミラーサーバを用意した。表示するためのページとしては、HTMLファイルと80個の画像から成る(合計88.9kBytes)。研究室内のサーバにもそのコピーを置いた。

HTTPパイプラインと振り分け接続の機能を使い、グーグルにアクセスした場合の結果の一部を図3に示す。黒い部分がサーバとの接続に要する時間であり、グレーの部分がデータ転送にかかる時間である。

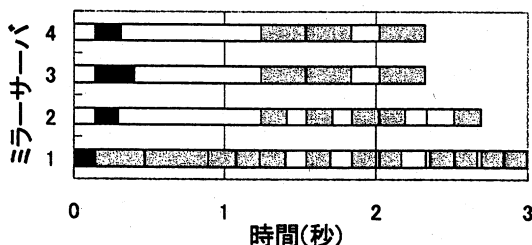


図3: HTTPパイプラインを使った振り分け接続の接続結果

HTTPパイプラインリクエストを複数のミラーサーバに振り分けて要求を行い、データを受け取っていることが分かり、HTTPパイプラインと振り分け接続が有効に働いていることが確認できる。

次に、ページを開くまでにかかった時間を表1に示す。実験1が全ての機能を用いない接続方法、実験2が持続的接続のみを用いる方法、実験3が持続的接続と振り分け接続を利用した方法、実験4が持続的接続とHTTPパイプラインを利用した方法、実験5が持続的接続、HTTPパイプラインおよび振り分け接続の機能を利用した方法である。また、研究室内LANでの実験では、100Base-T(実効速度40Mbps)、10Base-T(実効速度8Mbps)、それとクライアント側で帯域を100kbpsに制限した場合の実験を行った。

表1: 実験の測定結果

環境	かかった時間 [sec]			
	研究室内			グーグル
通信速度	40Mbps	8Mbps	100kbps	5Mbps
実験1	1.87	2.32	72.36	29.5
実験2	1.70	1.98	43.04	15.9
実験3	1.53	1.81	46.59	11.4
実験4	1.50	1.88	42.67	—
実験5	1.31	1.52	44.86	13.6

この中でグーグルにおける実験4の結果がないのは、グーグルのWebサーバであるGWSがHTTPパイプラインの機能に対応していなかったからである。持続的接続、HTTPパイプライン、振り分け接続のすべての機能を用いる実験5が最も効率的であったのは、研究室内の実験の100Base-Tと10Base-Tの場合であった。クライアント側で帯域制御を行った場合、複数のコネクションを持つ場合に効率が落ちるという帯域制御の問題があったため性能が出なかった。インターネット上のグーグルで測定した場合には、HTTPパイプラインが完全に働いていないという問題があり振り分け接続と持続的接続を用いた実験3がもっともよいという結果となった。

6 まとめ

複数のミラーサーバへの効率的なアクセスのために、HTTP 1.1の機能である持続的接続、HTTPパイプラインを用いて分散アクセスを行うよう改良を加えWebページを開くための時間短縮を行った。

今後の課題としては、サーバ選択に関して改良を加え、さらなる通信効率の向上、新たな通信方法の提案をしていきたい。

参考文献

[1] 藤澤弘, “複数のミラーサーバを用いた高信頼高効率 Web アクセス機構の実装および評価”, 香川大学工学部 卒業論文, 2002.