Webアプリケーションサーバの動作履歴を用いた性能解析

奈良田 徹† 山口 実靖† 淺谷 耕一‡

†工学院大学 ‡工学院大学大学院

1. はじめに

近年，CMS（Content Management System）を用いたブログやWikiなどの普及によりユーザのリクエストに対して動的な応答を返すWebアプリケーションサーバが注目されている．しかし，静的コンテンツと比べ動的コンテンツの生成は負荷が高くサーバへの負荷が問題となっている．Webアプリケーションサーバによる動的コンテンツの生成処理には，CGI(Common Gateway Interface)によるセッション毎のサーバプロセス生成と終了のオーバーヘッド[1]，RDBMSへの接続やデータベース処理に起因するオーバーヘッドなど複数の性能劣化要因があり，これを定量的に評価することは難しい．

本研究では，Webアプリケーションサーバの動作履歴を定量的に解析することで，ボトルネック箇所の発見を行う．

2. Webアプリケーションサーバ

Web アプリケーションサーバとは，CGI 等のプログラムを用いることにより，リクエストに対して決められたデータしか表示することのできない静的ファイルとは異なり，掲示板のような利用者の書き込みやアクセスによって表示内容が変化する動的なコンテンツを扱うことが出来る．

動的ファイルを扱う場合の処理手順は以下のようになる．また，処理手順を図 1 に示す．

①対象となる URL を入力する
②クライアントからサーバへ要求を送る
③サーバ上でプログラムを実行し，データベースへのアクセスを行う
④サーバが Web ページを生成する
⑤サーバが生成したページをクライアントに送信する
⑥結果をクライアント上に表示する

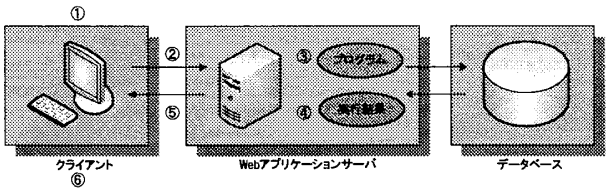

図 1. Webアプリケーションサーバの処理手順

Log-Based Performance Analysis of Web Application Servers
† Toru NARATA,Saneyasu YAMAGUCHI
Kogakuin University
‡ Koichi ASATANI
Graduate School of Kogakuin University

3. Webアプリケーションサーバ性能解析

3.1 レスポンスタイム測定

Web アプリケーションサーバの性能解析として，レスポンスタイムの測定を行った．Web アプリケーションサーバとクライアントの接続図を図 2 に示す．また，性能解析における PC 性能を表 1 に示す．

レスポンスタイム測定は ApacheBench を用いて行った．CGI アプリケーションは RDBMS 内のテーブル全行を SELECT により読み込みを行うものである．本 CGI アプリケーションは Ruby で記述されている．

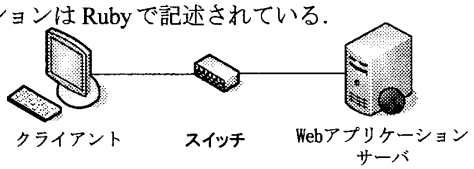

図 2. 実験環境接続図

表 1. 実験環境およびサーバスペック

OS	FedoraCore7
CPU	Intel Celeron D 3.20GHz
Memory	DDR2 1GB
Network	Broadcom 5721 Gigabit NIC
WebServer	Apache2.2.6
DataBase	PostgreSQL8.2.5
Benchmark	ApacheBench2.0.4

本実験において得た Web アプリケーションサーバのレスポンスタイムを基に動作解析を行う．具体的には Web アプリケーションの処理に含まれる個々の処理に要する時間を個別に求め，ボトルネック箇所の特定を行う．レスポンスタイム測定で使用した CGI アプリケーションに時刻を収得する命令文を挿入し，SELECT 文処理，DB 接続，DB 切断時間を求めた．また，CGI オーバーヘッドは 1 バイトの CGI を Ruby で作製し，その処理時間を求めた．

Webアプリケーションのレスポンスタイムと個別の処理時間を図 3 に示す．テーブルの行数を上げる事により全体のレスポンスタイムは上昇しているが，RDBMS のテーブルを読み込む SELECT 文の処理時間が行数の増加に対して増えており，その他の DB 接続時間や CGI 処理時間はほぼ一定となっていることが分かる．また，テーブルの行数が 1000 以下では CGI 起動処理が全処理中で最も時間を要しており，次に RDBMS への接続が時間を要していることが分かる．10000 行の例では，RDBMS の SELECT 文処理時間が全処理時間の 3 分の 2 以上を要している．本手法により，各処理時間が要する時間を定量的に確認することができた．

3.2 アクセス履歴測定

次に動作解析システムを構築し，オープンソースのWebアプリケーションの解析について述べる．

Webアプリケーションとしてブログを構築し，読み込みの際のアクセス履歴を収得した．Webアプリケーションとしてオープンソースの CMS である Drupal を使用した．ブログに対して 1000 回の読み込みを行い，Tcpdump を用いてデータの流れを収得した．本実験環境においてApache と RDBMS は同計算機内に存在するため，eth0 および lo の 2 つの NIC をモニタした．

Web アプリケーションへの読み込みの際のアクセス履歴を図 4 に，図 4 の拡大したものを図 5，6，7，8 に示す．図 4 より 0.02[sec]および 0.045[sec]付近からデータの流れが止まっている．図 5，6 より Apache 側の処理によりデータの流れが止まっていることが分かり，よってレスポンスタイムの多くは CGI 側の処理にあることが分かる．．本手法により，クライアント，Apache，RDBMS 間の詳細なデータの流れを確認することができた．

4. おわりに

本稿では，Web アプリケーションサーバのボトルネック箇所を解析するためレスポンスタイムの詳細な解析および Web アプリケーションへの読み込みの際のアクセス履歴の調査を行った．

今後は詳細なレスポンスタイムおよびアクセス履歴の解析を行い，ボトルネック箇所の発見，改善を行う予定である．

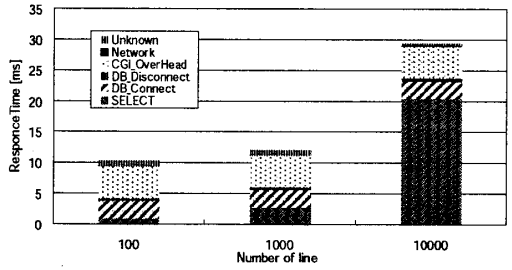

図 3. RDBMS アクセスにおける
レスポンスタイム

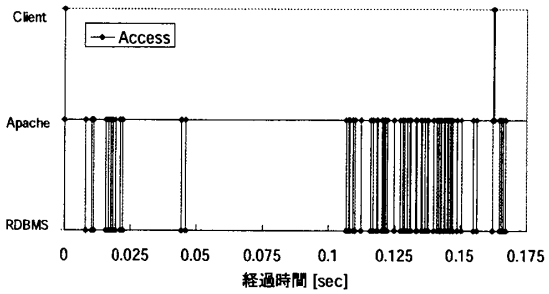

図 4．Web アプリケーションへの書き込み時の
アクセス履歴

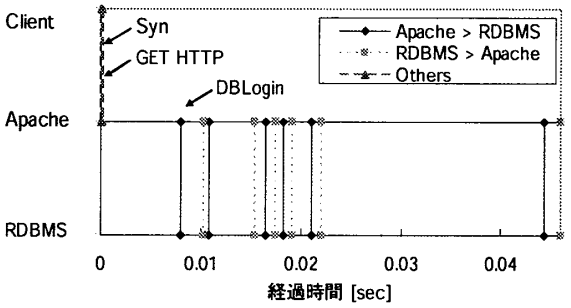

図 5．アクセス履歴拡大図（ｉ）

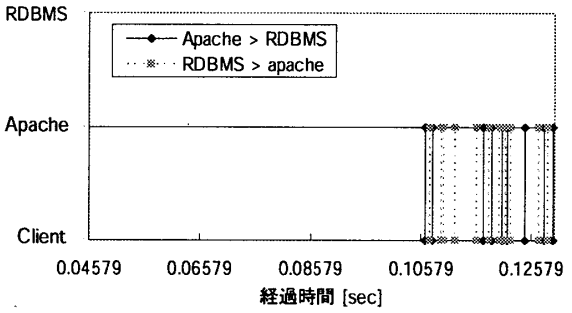

図 6．アクセス履歴拡大図（ii）

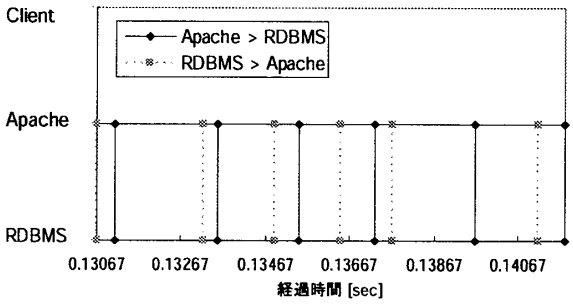

図 7．アクセス履歴拡大図（iii）

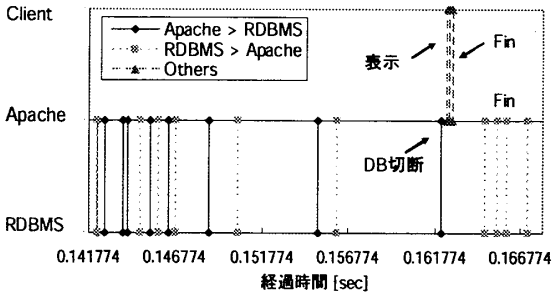

図 8．アクセス履歴拡大図（iv）

参考文献

[1]原大輔,中山泰一 ,"セキュアかつ高性能なウェブサーバの設計と実装," 情報処理学会第 47 回プログラムシンポジウム報告集，pp.71-78，,Jan．2006.