

# 次回アクセスを保証する Web システム NAP-Web の性能評価と改良

07T276 山田 茂和（最所研究室）

サーバの過負荷時におけるユーザの不満を、整理券を配布し次回アクセス時間の保証を行うことで解消する Web システム NAP-Web の詳細な性能評価を行い、その結果に基づいた改良について述べる。

## 1 はじめに

Web サーバへのアクセス集中によるレスポンス低下は、ユーザにとって大きな不満となる。この問題を解決するために我々の研究室では、整理券を配布することによって次回アクセス時間を通知し、その時間に行われるアクセスを受け付けることを保証する Web システム NAP-Web (Next Access Promised - Web) の開発を行っている。

現在、システムの基本的部分は完成している [1]。しかし、実運用に向けての課題は多く残っており、多くの改良を行う必要がある。本研究では改良の前段階として、NAP-Web の性能を詳細に評価し、実用化するまでの問題点を洗い出す。そして、抽出された問題点を解決するために必要な NAP-Web の改良についての考察を行う。

## 2 NAP-Web の概要

NAP-Web の概要を図 1 に示す。サーバが対応できずにレスポンスを返すことができないアクセスに対して、拒否すると同時に整理券を配布し、再アクセスを保証する。

NAP-Web では動的負荷制御を実現するために、アクセスをサーバの状態に合わせてグループ分けし、スケジューリングを行っている。基本的なアクセスグループは「Run\_Ready」「Wait」「Next\_Wait」「Re\_Access」の 4 つである。

Run\_Ready アクセスグループでは通常の処理が行

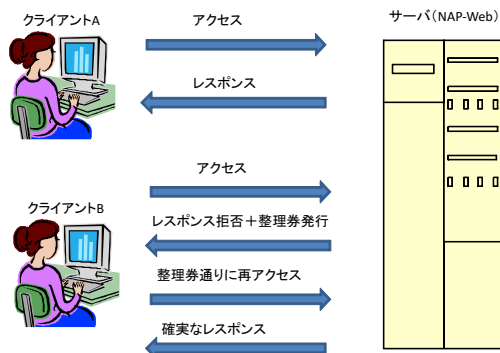


図 1: NAP-Web の基本動作

われる。Wait アクセスグループには、Run\_Ready に入りきらなかったアクセスがキューイングされており、Run\_Ready アクセスグループに空きができれば先頭から移動する。Wait アクセスグループにも入りきらなかったアクセスは Next\_Wait アクセスグループに入り、整理券を渡された上で拒否される。

整理券にはアクセスを識別する ID と、次回アクセス可能時間が記述されており、この時間に再アクセスを行うと、Re\_Access アクセスグループにキューイングされ、Wait アクセスグループに空きができれば先頭から移動する

## 3 NAP-Web の問題点について

NAP-Web の課題として挙げられるのは以下の二点である。

- スケジューリング機構に対する検証
- サービスの公平性

まず、スケジューリング機構に対する検証であるが、現在 Wait アクセスグループの上限数はアクセス数や予測応答時間に基づき動的に決定されているのだが、これに対する評価および考察が十分にされていない。そこで本研究では、Wait アクセスグループの上限数が NAP-Web の振る舞いにどのような影響を与えるのかを調べる。

次回アクセスが可能となる時刻を正確に予測するのは難しいため、整理券には余裕をもたせた時刻が記載されている。しかしながら予測が当たった場合には、この余裕はサーバにとって無駄な時間となってしまふ。この無駄を無くすため、次回アクセス待ちのクライアントが多く存在する場合でも初回アクセスを受け付けるようになっているが、これは公平性に問題がある。本研究では次回アクセス待ちのクライアントが多数存在する場合には初回アクセスの登録を制限するようにして実験を行い、その効果を確認する。

## 4 NAP-Web の性能評価および改良案

Wait アクセスグループの上限数が NAP-Web の振る舞いに与える影響を調べるために、動的コンテンツに対する実験として掲示板、静的コンテンツに対する

実験としてサイズの異なるコンテンツへのアクセス実験を行った。

#### 4.1 動的コンテンツに対する評価実験

Web サーバ上に掲示板プログラムを用意し、1KBの文字列を投稿するリクエストを50個繰り返すスレッドを400個生成するスクリプトを用いて実験を行った。従来設定および Wait の上限数を200に固定した場合での実験結果を図2、3に示す。

図2は、実際にアクセスしてから完了するまでの時間で、Next\_Wait での待ち時間は含まない。アクセスが殺到した時点でのレスポンスタイムは長くなっているが、しばらくすると定常状態となった。Wait アクセスグループの上限数を静的にした場合は、Wait で待つ時間が伸びたため従来手法よりも長くなった。

図3の横軸は初回アクセスの開始時刻、縦軸はアクセス終了時刻である。動的なものよりも、上限数を静的にしたものの方が全体の処理時間は短くなったが、これは Wait の上限数を大きくしたことで、待ち時間の予測精度が上がったためだと思われる。Wait アクセスグループの上限数を静的にしても、アクセスが行われていない空白の時間が発生しており、Wait 上限数の調整のみで解決するのは難しいと言える。

紙面の都合で結果は示さないが、次回アクセス待ちのクライアントが多数存在する場合に、初回アクセスの受け付けを制限するようにして実験を行ったが、レスポンスタイムとアクセス時間の傾向は上限を静的にしたものとほとんど変わらなかった。しかし、整理券の発行枚数は増加しており一定の効果があったことを確認している。

#### 4.2 静的コンテンツに対する評価実験

1K, 10K, 100K, 200K, 500K, 1000KB の5種類のファイルに対して、それぞれアクセスを50回繰り返すスレッドを400個生成するスクリプトを用いて実験を行った。

従来設定の場合と Wait の上限数を静的にした場合の実験結果を、図4に示す。図では、アクセス一つあたりにかかる時間が直感的に分かるようにスループットの逆数をプロットしているが、Wait アクセスグループの上限数とは関係なく、静的コンテンツの大きさと NAP-Web のスループットとは強い相関関係があると言える。これは、次回アクセス可能な時刻の予測にコンテンツの大きさを取り入れることにより、より正確に予測をすることが期待できる。

### 5 まとめと今後の課題

動的コンテンツを用いた実験において、Wait アクセスグループの上限数が NAP-Web に与える影響を確認できた。また、サービスの公平性に対する改良と

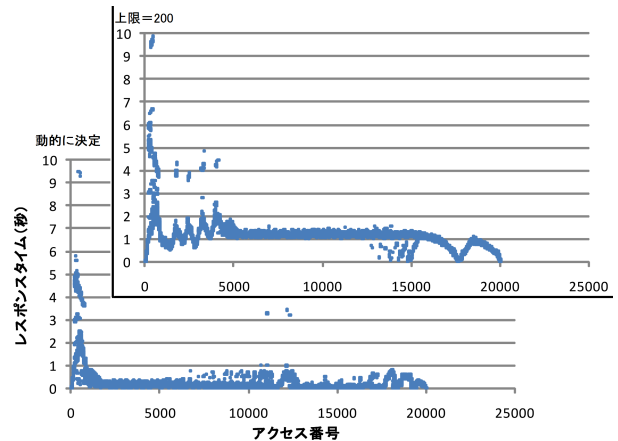


図 2: レスポンスタイム

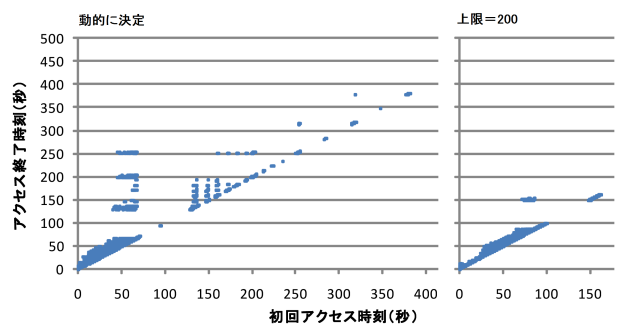


図 3: アクセス時間

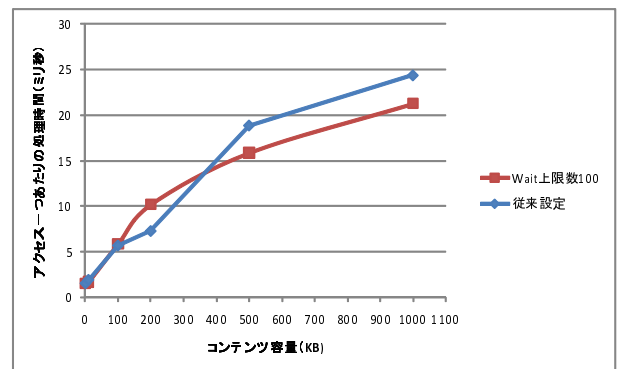


図 4: スループットの逆数

して行った実験も、一定の効果を確認することができた。また、静的コンテンツを用いた負荷実験の結果からは、予測の高精度化の可能性を得ることができた。

今後の課題としては、得られた結果に基づいて整理券に記載する時間の見直しや、Wait アクセスグループに入れる機構のチューニングなどが挙げられる。

#### 参考文献

- [1] 加地智彦, “急激な需要増加による負荷を整理券により時間的に平滑化する Web システムの開発,” 香川大学大学院, 博士論文, 2010 年.