

アクセス情報を用いたWebサービスのサービス時間の予測

09G468 新田 淳二 (最所研究室)

本研究では、NAP-Webの課題である連続アクセスへの対応とサービスの公平性を解決するために、アクセス待ち時間予測手法についての提案を行う。具体的には、情報を取得するページ情報取得システム的设计を行い、効率よく情報を収集するアクセスランキング手法とサンプリング手法を提案し、有効性を確かめた。そして、実際に運用されているサーバのログに基づいて生成したログを用いてシステムの実装と評価を行った。

1. はじめに

我々の研究室では、アクセス集中によって人気 Web ページを見ることができない問題を解決するための Web システム (NAP-Web) [1]の開発を行っている。

NAP-Web は、サーバに対するアクセスを、直ちにサービスするアクセスとしないアクセスに分類する。そして、サービスしないアクセスを發したユーザに対しては、待ち時間を書いた整理券を渡し、その時間のアクセスを保証する。現在 NAP-Web は、確実にサービスするために、待ち時間に余裕を持たせている。また、HTML コンテンツなどの複数のリソースから構成される連続アクセスに十分に対応していない[2]。このため、システムの利用効率が悪く、連続アクセスできないユーザから不満が出る問題がある。この問題を解決するためには、サービス可能な時間の予測精度を上げる必要がある。

本研究では、上記の問題を解決するために、Web ページの構成やサービスにかかる時間などの Web コンテンツ情報を収集し、その情報をもとに待ち時間の予測を向上する「ページ情報取得システムの開発」を行っている。

2. 提案手法

NAP-Web のアクセス待ち時間の予測精度が低い理由と連続アクセスに対応できない理由を述べ、解決策を述べる。本稿では、Web ページそのものをコンテンツと呼び、構成されるコンテンツのメインファイルをメインリソース、その他のファイルをサブリソースと呼ぶ。

2.1. アクセス待ち時間予測手法

現在は、アクセス単位で待ち時間を予測しており、図 1 に示すようにアクセスを複数の間隔 (スロット) に分け、スロットごとに待ち時間を予測している。スロットに入るアクセス数は、アクセスされるリソースの種類を考慮することなく、アクセス時間の平均値を用いて決めている。

このため、平均と異なるアクセス時間のファイル群へのアクセスを複数受けた場合、サービス可能な時間が予測と異なってくる。現時点では、予防策として予測の 2 倍近くを通知することにより、確実にサービスできるようにしている。しかし、これにより、各スロットの空き時間が多くなり、無駄が発生し、サーバの利用効率が低下するという問題が生じる。

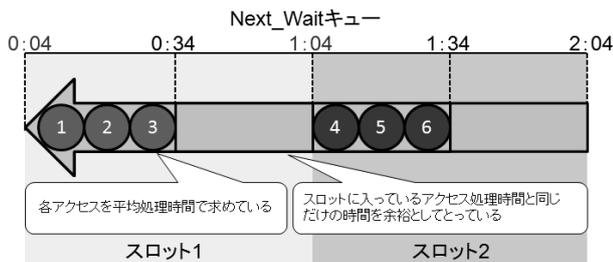


図1 現在のアクセス待ち時間予測手法

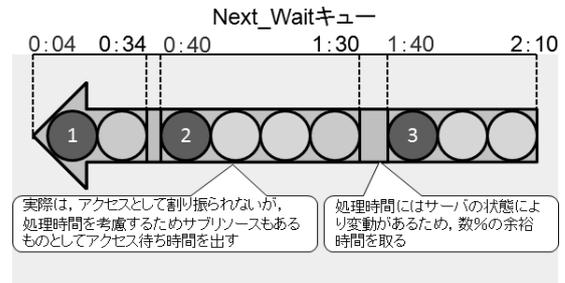


図2 提案するアクセス待ち時間予測手法

提案手法では、アクセスされるリソースの情報を把握することにより、各リクエスト (アクセス) 単位の処理時間を得ることができ、連続アクセスを要するコンテンツへのアクセスがきた場合でも、サービス可能な時間を正確に予測することが可能となる。

ただし、サーバの状態により処理時間が変化するので、それを吸収するための空き時間は必要になるが、サーバの利用効率は向上する (図 2)。

2.2. Web アクセススケジューラ

サービスの公平性を向上させ、連続アクセスを必要とするコンテンツへ対応するために、これまでの Web スケジューラに Prim_Access を導入した (図 3)。Prim_Access では、連続アクセスを必要とするコンテンツのサブリソースを保持し、優先的に処理をする。

これにより、メインリソースに付随するサブリソースを高速に提供する。

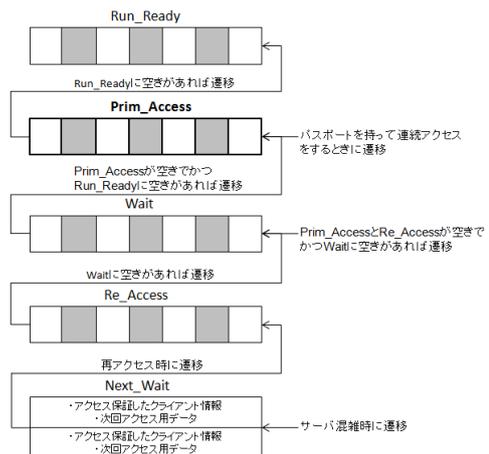


図3 Web アクセススケジューラ

3. ページ情報取得システム

コンテンツは複数のリソースにより構成されているので、構成 (以下、ページ構成) やリソースごとの処理時間などの情報が必要となる。ページ情報取得システムは、情報を収集するシステムである。時間予測に必要な Web ページの

ページ構成、処理時間およびサイズをサーバのアクセスログ情報とリソースディレクトリから集める。

3.1. システムの改良

卒業研究で作成したプロトタイプを改良することにより処理時間を改善できたが、実用化には程遠かった。原因は、以下の2つであった。

- ① 膨大なログファイルのソート時間
- ② DB書き込みによるオーバーヘッド

これらの問題を解決するための改善策として以下の2つを導入することにした。

- ① サンプルング手法を用いた探索対象の削減
- ② アクセスランキング手法を用いた扱うデータの削減

3.2. サンプルング・アクセスランキング手法の有効性

図4は、香川大学の代表サーバにおけるログをサンプルングした情報でコンテンツのアクセスランキングを生成し、ランキング順位とそれらのリソースの全アクセスに対するアクセス割合を示したものである。図4より、このサーバの場合には、サンプルング率が5%でも母集団とほとんど変わらないことが分かる。これにより、サンプルングをすることにより人気コンテンツを把握し、全体のアクセス傾向を十分に把握できることが分かった。

また、ランキング上位のコンテンツにアクセスが集中していることから、これらを把握することにより全体を把握できる。

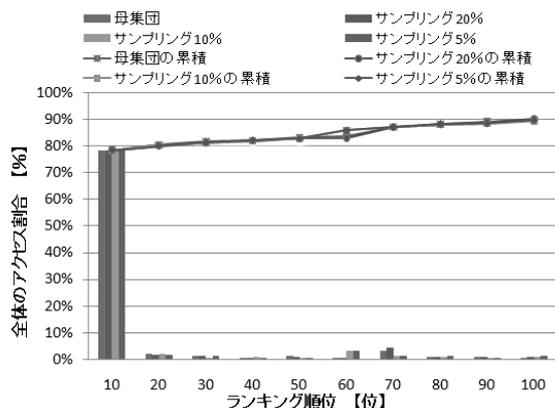


図4 サンプルング率とアクセス把握割合の関係

3.3. ページ情報取得システムの概要

図5は、ページ情報取得システムの概要である。本システムの動作としては、初めに、サンプルングモジュールがNAP-Webのログファイルをサンプルングして、サンプルングデータを生成する。次に、アクセスランキング生成モジュールが

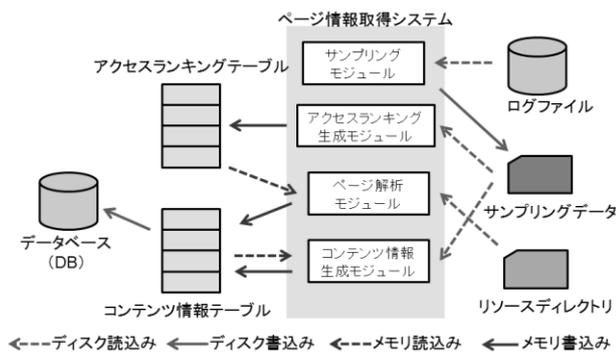


図5 ページ情報取得システムの概要

ジュールが人気コンテンツを判別してアクセスランキングを生成する。そして、ページ解析モジュールが人気コンテンツの構成を解析してリソースの構成情報を取得し、コンテンツ情報生成モジュールで、人気コンテンツを構成するリソースの処理時間とサイズをサンプルングしたログから取得し、コンテンツ情報テーブルに格納する。また、コンテンツ情報は、DBにも保存するが、メモリ中に常駐させる。

4. 評価

本システムは、NAP-Webと同じサーバ上で動作するため、サーバ資源を多く使ってはならない。そのため、システムの処理時間としては数分以内に完了する必要があると考える。これは、仮に1分間CPUを占有したとしても、24時間でのサーバの負荷としては0.1%未満に抑えることができる。このため、評価としては処理時間(CPUタイム)を指標とした。

図6は解析対象のログ数を500,000行、サンプルング率を1%として、アクセスランキング適用順位を変化させた場合の処理時間の推移である。

アクセスランキングの適用順位により、全体の処理時間が増大するものの1.6秒程度であり、十二分に要件を満足して、サンプルング率、ログ数を変えて処理時間を測定したが、全て1分以内であり、いずれの場合も要件を満足していた。

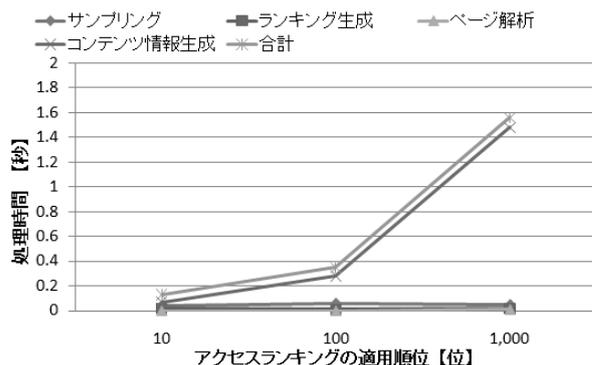


図6 システムの処理時間の推移

5. まとめ

本稿ではNAP-Webの問題を挙げ、その解決策としての手法を提案し、提案手法に必要なコンテンツ情報を集めるページ情報取得システムについて述べた。以下では、今後の課題を述べる。

- 提案した手法のNAP-Webへの組み込み、及び有効性を調査
- 処理時間だけでなく、スループットなどの別の基準を用いたアクセス待ち時間の予測
- NAP-Webによるリアルタイムでのサンプルングログの生成による、サンプルング処理の除去

参考文献

- [1] 加地智彦, 最所圭三, "NAP-Webにおける次回アクセスの効率化", 電子情報通信学会論文誌, Vol.J92-B, No.10, pp.1600-1610, 2009
- [2] 加地智彦, 最所圭三, "NAP-Webへのページスケジューラの導入", 第7回情報科学技術フォーラム講演論文集 (FIT 2008), L-031, Vol.4, PP.159-160, 2008