

NAP-Web を用いた優先アクセス機構の実装と評価

11T275 溝渕 久哲 (最所研究室)

分散型の Web システムにおいて、キャッシュサーバからオリジナルのサーバへのアクセスを、優先的に処理することができる機構を持った Web システムの開発と評価について述べる。

1 はじめに

キャッシュサーバを用いてサービス能力を向上させる Web システムにおいては、オリジナルサーバの過負荷によって、キャッシュサーバのキャッシュ更新が滞ると、サービスの質が低下してしまう。この問題に対処するために先行研究 [2] では、キャッシュサーバからのアクセス (以下優先アクセス) を優先的に処理するために、NAP-Web[1] に組み込む優先アクセス機構の設計及び簡易実装、評価が行われた。本研究では、優先アクセスが通常アクセスに与える影響を詳細に評価し、その問題を抽出し改良を行った。さらに未実装だったキューを実装し、その評価を行った。

2 優先アクセス機構

2.1 概要

優先アクセス機構を組み込んだアクセススケジューリング機構を図 1 に示す。スケジューリング機構は、通常アクセスのためのキューと独立した優先アクセスのためのキューを用意し、それぞれ指定した数のリクエストを同時に処理する。先行研究においては、キューを用いずに優先アクセスを直接処理に回す実装を行っていた。本論文ではこれを旧機構と呼ぶ。

2.2 旧機構を用いた問題点の調査

実験環境 旧機構を組み込んだ NAP-Web を搭載したサーバを用いて実験を行った。優先アクセスを行うクライアントと通常アクセスを行うクライアントから、指定した数の同時アクセスをサーバに対して行う。優

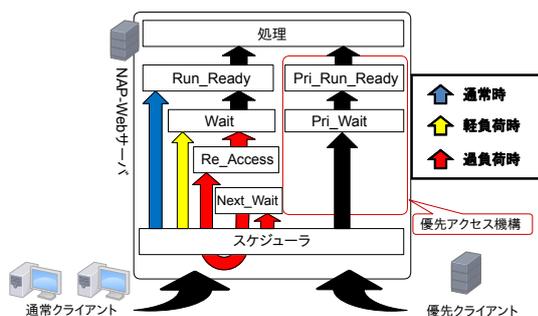


図 1: NAP-Web を用いた優先アクセス機構

先アクセスは通常アクセスが安定する時間 (本実験では 200 秒) 待ってから行う。

実験内容 通常クライアントの同時アクセス数 (通常アクセス数) は 3,000 に固定し、優先クライアントの同時アクセス数 (優先アクセス数) を 0,1,5,10 の 4 通りで実験を行い、各アクセスのスループット及び応答時間を調査する。

実験結果 各応答時間の 20 秒の移動平均をとったグラフを図 2 に示す。優先アクセスを開始直後の通常アクセス応答時間が極端に長く、優先アクセス数が多いほど影響が大きい。優先アクセスに対しては高速に応答している。この結果は、優先処理は実現できているが、通常アクセスに悪影響していることを示す。

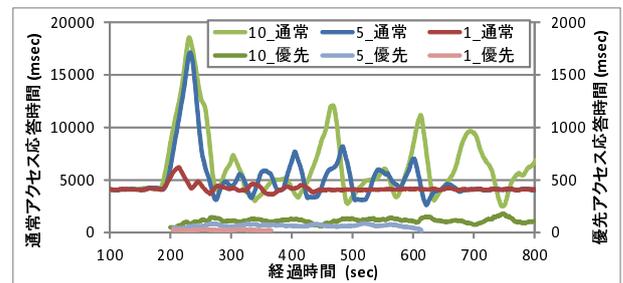


図 2: 旧機構での応答時間

2.3 改善策

2.2 節で述べた問題は、優先アクセスが入ったことにより通常アクセスに割りてられるリソースが減少したためであると考え、同時に処理する優先アクセス数を徐々に増やしていく方法を試した。

実験内容 改善策の有効性を調べるため、2.2 節と同じ実験環境で、優先アクセス数を徐々に増加させる実験を行った。通常アクセス数を 3,000 に固定し、優先アクセス数を一定間隔 (t 秒) 毎に 1 ずつ追加し 5 まで増加させた。t は 0,20,40,60,80 の 5 通りを試した。

結果 各応答時間の 20 秒の移動平均をとったグラフを図 3 に示す。t が 20 と 40 の時は、t が 0 の時と比

べ、応答時間の遅れが半分程度に収まった。t が 60 と 80 の時は、影響はさらに小さくなり、図 2 の同時アクセス数 1 の場合とほぼ同じ値となった。以上のことから、優先アクセスを徐々に増やす方法が有効であることが確認できた。

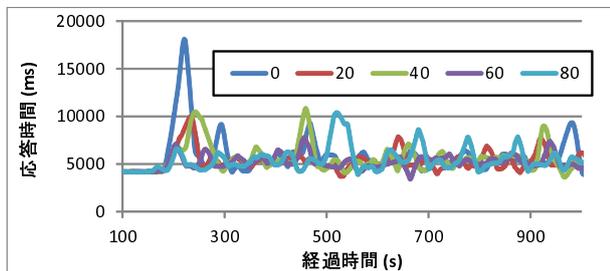


図 3: 改善策を用いた時の通常アクセス応答時間

3 優先アクセス機構の実装及び評価実験

先行研究において設計されたキューを用いた優先アクセス機構 (新機構) を実装した。Run_Ready と Pri_Run_Ready に含まれるプロセス数が、それぞれのアクセスが同時に処理される数となる。この数を調整することで、優先アクセスと通常アクセスの処理割合を調整できると考えた。その確認とキューの効果を評価する実験を行った。

3.1 同時実行数を変更

Run_Ready と Pri_Run_Ready の各キュー数を変更することによる、優先アクセスと通常アクセスの処理割合の調整機能について調べる。

実験内容 実験はここで実装した優先アクセス機構を搭載したサーバを除き、2.2 節と同じ環境で行った。優先アクセスのキューの個数と通常アクセスのキューの個数の組み合わせ (Pri_Run_Ready 数, Run_Ready 数) として、(1,10),(1,20),(2,10),(2,20) の 4 パターンについて実験を行った。

実験結果 各パターンの優先アクセススループットの 20 秒の移動平均をとったグラフを図 4 に示す。なお、通常アクセスのスループットに関しては、各パターンで違いがなく 130 程度となっていた。優先アクセスを処理する Pri_Run_Ready の数が同じである (1,10) と (1,20) を比較すると、全体的に同じ結果となった。(2,10) と (2,20) についても近い値となっている。また、(1,10) と (2,10) を比較すると、明らかにキューの数が多き方がスループットが大きくなっており、これは (1,20) と (2,20) の場合も同様であった。以上から、優先アクセスに関してはキュー数の変更によるスループットの調整が可能であるが、通常アクセスに関

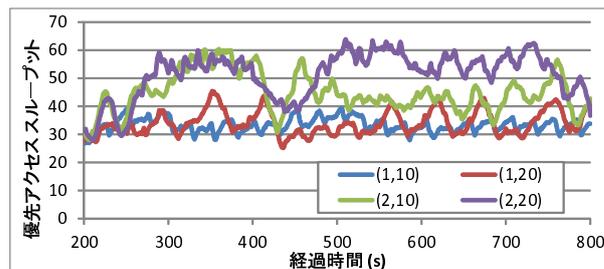


図 4: 各キュー数時の優先アクセススループット

しては行えなかった。通常アクセスが期待した結果と異なった原因については今後の課題とする。

3.2 優先アクセス数を変更

新機構では優先アクセスもキューで処理するため、同時アクセス数が増加しても通常アクセスの処理には影響を及ぼさないはずである。これを実験で確認する。

実験内容 3.1 節と同一の環境で Pri_Run_Ready 数を 2, Run_Ready 数を 20 にした状態で、優先アクセス数を 5,10,20 の 3 パターンについて実験を行った。

実験結果 紙面の都合で結果のグラフは省略するが、旧機構と異なり、本実験では優先アクセス数にかかわらず同じ結果となった。また優先アクセス応答時間は、ほぼ優先アクセス数に比例して長くなっていた。スループットについても比較し、優先、通常の両アクセスとも影響は見られなかった。以上のことから、優先アクセス数が変化しても通常アクセスの処理には影響を及ぼさないことが確認できた。

4 おわりに

旧機構で評価を行い、優先アクセス機構の問題点を指摘した。改善策を提案しその有効性の確認した。またキューを用いた優先アクセス機構を実装し評価実験を行ったところ、優先アクセスについてはリソースの割当量が調整可能であった。また、優先アクセス数が増加しても通常アクセスの処理には影響を及ぼさないことも確認した。

今後の課題として、通常アクセスのスループットが Run_Ready 数に応じて変化しない原因を調査、同時処理数を動的に変更、優先アクセスの同時処理数を段階的に増加させる機構の追加などがある。

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, 博士論文, 2010.
- [2] 曾川直也, "NAP-Web における優先アクセスの設計と評価", 香川大学, 学士論文, 2014