

キャッシュサーバを用いた分散 Web システムにおける キャッシュ更新最適化の検討

12T252 畠山 侑也（最所研究室）

分散 Web システムにおいて、キャッシュ情報を管理しキャッシュ更新を最適化するための機構の実現方法について行った検討と、キャッシュの対象についての議論を述べる。

1 はじめに

サーバの過負荷によってクライアントへのレスポンスが悪化することや、サーバダウンなどの障害を防ぐため、我々の研究室ではキャッシュサーバを用いて、Web サーバの負荷量に応じて稼働するキャッシュサーバの台数を制御する分散 Web システムの開発が行われている [1]。先行研究では、キャッシュサーバのためのキャッシュ更新機構の設計方法について検討された [2]。

本研究では、キャッシュ更新を最適化する機構を実際のキャッシュサーバで実装するために、Web サーバソフトウェアである Apache のキャッシュ制御に関連するソースコードの解析、またコンテンツの種類に応じたキャッシュ有効期限、動的コンテンツのキャッシュを行う際の問題についての検討、およびそれらの検討に基づいたキャッシュ機構の設計について述べる。

2 Apache におけるキャッシュ

2.1 キャッシュの流れ

Apache には、機能拡張のためのモジュールが標準で用意されている。本研究では、キャッシュ制御に関するいくつかのモジュールをベースにして、分散 Web システムのためのキャッシュ更新機構を開発する予定であり、それらのモジュールを用いた時のキャッシュ管理の振る舞いを調べるための実験を行った。図 1 に Apache のキャッシュの流れを簡易化した図を示す。キャッシュサーバは、ユーザからのアクセスをトリガーにして、ユーザが要求する Web ページのキャッシュを保持しているかをチェックする。保持しているかつキャッシュの有効期限内であれば、そのままキャッシュをユーザへ返却するが、保持していない、またキャッシュの有効期限が切れている場合はキャッシュの取得のためにオリジンサーバへアクセスし、ヘッダ情報と HTTP ステータスコードから、そのコンテンツがキャッシュ可能かどうかを判断する。キャッシュ可能な場合、キャッシュの URI からキーを生成し、キャッシュを保存するディレクトリを決定し、ファイル名のハッシュ化を行った後保存される。キャッシュ不可の場合は取得したコンテンツの返却のみ行う。キャッシュできない例として、そもそもヘッダ情報が破損し

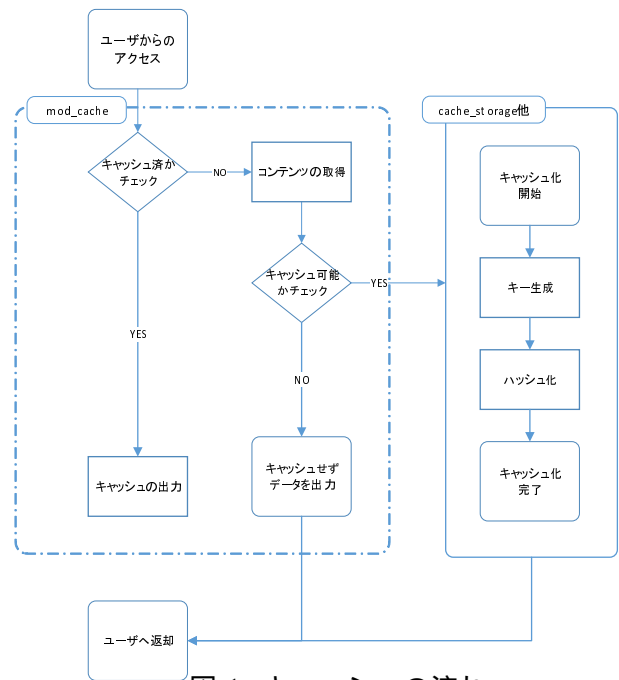


図 1: キャッシュの流れ

ている場合や、Cache-Control によってキャッシュ不可とされている場合がある。

3 コンテンツの種類によるキャッシュの問題

3.1 コンテンツの種類

コンテンツは、静的コンテンツと動的コンテンツの 2 つに分類できる。静的コンテンツは、ファイルを更新しない限り、全く同じ内容のページが表示され続けるものであり、HTML や CSS、画像ファイルなどが挙げられる。動的コンテンツは、Java や Javascript、PHP などの言語で記述された Web ページで、ユーザの入力に応じて生成されるコンテンツが変化するものである。例として検索サイトや、文書や作品を投稿などできるサイトが挙げられる。

動的コンテンツは、ユーザ情報やリクエスト、時刻など、様々な条件に応じて変化するため、Web ページ

のコンテンツをキャッシュすることが負荷軽減に効果的かどうかを考えなければならない。例えばユーザを識別するための情報を扱うものは、キャッシュするとセキュリティの問題が発生するため、キャッシュ不可能である。また動的コンテンツの内容がアクセスするユーザによって変化する場合、そのキャッシュはユーザや、検索結果の数だけ蓄積してしまうことになり、キャッシュストレージが一瞬で埋まってしまう状況も考えられる。こういったコンテンツをキャッシュすることは非効率であり、推奨されない。一方でブログや天気予報などの、一定時間毎に更新するような動的コンテンツは、キャッシュの有効期限とコンテンツの更新時刻を一致させればキャッシュが可能である。また、短期間のキャッシュであれば、更新部分が小さければキャッシュしても問題無い場合が多いと考えられる。

3.2 動的コンテンツの問題

動的コンテンツは、その動作方法から大きく2種類に分けられる。Webサーバからクライアントへプログラムが送られ、Webブラウザ上でプログラムが実行されるものと、クライアントからのリクエストによって、Webサーバでプログラムが実行され、その結果がクライアントへ返却されるものである。前者の場合、javascriptが例として挙げられる。コンテンツであるプログラムのコード自体は変化しないため、キャッシュ可能である。後者はPHP、Perlなどが挙げられる。こちらはSQLを利用しデータベースと連携した運用を行っている場合、データベースから取得した内容によってコンテンツが変化するため、安易にキャッシュできない。動的コンテンツのキャッシュを行う場合、コンテンツに応じた有効期限の設定が重要になる。キャッシュ非推奨なコンテンツの一つにオークションが挙げられる。新しい入札によってデータの更新が行われた場合、それまでのデータの価値はなくなる。しかし、ユーザからのアクセスの度に、オリジンサーバへデータを取得しに行くのは負荷となる。特にオークションの場合は、入札締切直前のアクセスは大きく増加し、それに伴い負荷も増加する。そこで、キャッシュの有効期限をオークションの運営に影響しない程度のごく短い時間に設定する。たとえ有効期限が1秒や2秒の場合でも、大量のアクセスをキャッシュサーバのみで捌けることは有用性があると考えられる。

4 キャッシュ更新機構

キャッシュ更新機構はキャッシュサーバ内に設置される。図2にキャッシュ更新機構の動作を示す。

ユーザからのアクセスをトリガーとしてキャッシュが蓄積される。キャッシュサーバはオリジンサーバへコンテンツを取得するためアクセスし、キャッシュストレージに取得したコンテンツをキャッシュとして保存する。キャッシュ管理機構は、キャッシュから登

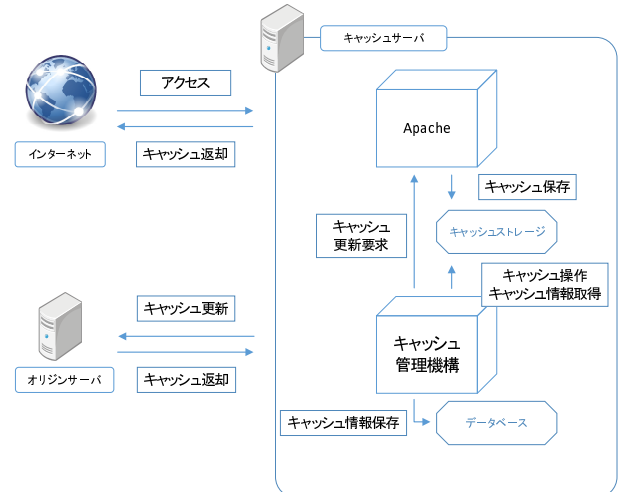


図2: キャッシュ更新機構

録日時、有効期限などの情報を取得し、データベース化する。更にキャッシュファイル毎にアクセス数を記録し、これもデータベースへ追加する。取得した情報からキャッシュの有効期限が切れないような更新スケジュールを作成し、キャッシュ更新の動作がユーザからのアクセスに起因しないようにする。管理機構からのApacheへ更新要求を送ることで、オリジンサーバへキャッシュを取得しに行く。ただし、保存したキャッシュにユーザからのアクセスが無ければ、該当ファイルを削除し、更新は行わない。

5 おわりに

本論文では、Apacheにおけるキャッシュ更新において、キャッシュ更新機構を実際のキャッシュサーバで実装するために、Apacheのキャッシュ操作に関連するソースコードの解析と、機能実装に向けた実験を行った。またコンテンツの種類に応じたキャッシュ有効期限、動的コンテンツのキャッシュを行う際の問題について検討した。

今後の課題として、キャッシュ情報管理データベースの作成、キャッシュ有効期限前に更新するための機能、mod_cacheの解析、また動的コンテンツのキャッシュ実験を行い、必要がある。

参考文献

- [1] 堀内農彦, "クラウドに適したWebシステムの負荷監視機能の改善と評価", 香川大学, 学士論文, 2014
- [2] 岡田 祥太, "キャッシュサーバを用いた分散Webシステムにおけるキャッシュ更新機構の検討", 香川大学, 学士論文, 2015
- [3] Apache Software Foundation, "Apache HTTP Server", <http://httpd.apache.org/>