

データベース参照のスケジューリングによる  
電子商取引サイトの最適化  
Performance Optimization of an Electronic Commerce Site  
by Database Query Scheduling

赤坂 謙二郎<sup>†</sup>  
Kenjirou Akasaka

布目 淳<sup>‡</sup>  
Atsushi Nunome

平田 博章<sup>‡</sup>  
Hiroaki Hirata

柴山 潔<sup>‡</sup>  
Kiyoshi Shibayama

## 1. まえがき

ウェブショップなどの電子商取引サイト (Electronic Commerce Site: 以下 EC サイト) は、ウェブサーバとそのウェブサーバ上で動作するウェブショップ構築システム、およびデータベースサーバを主なコンポーネントとして構成される。クライアント (顧客) からの HTTP アクセスに対して、EC サイト内では、必要な情報をデータベースサーバに問い合わせ取得し、これをもとに動的にウェブページを生成する。EC サイトを効率良く機能させるためには、ネットワーク回線速度とウェブサーバ、データベースサーバの性能をハードウェアとソフトウェアの両面でバランスさせなければならないが、一般に、これは容易ではない。特に、人気商品の発売開始時やタイムセールなど、平常時とは比較にならないほど多くのアクセスが集中すると、データベースサーバの応答待ちで EC サイトが一時的に機能しなくなる状況も生じる。このような事態を回避するためには、単にメモリ使用量などのシステムパラメータを調整するだけでなく、データベース処理にかかる負荷の軽減も図らなければならない。

本稿では、EC サイトとしての特徴を考慮して、データベースサーバに対するウェブサーバからのクエリをスケジューリングすることにより、データベース処理のスループットを改善する方式を提案する。ウェブショップ構築システムとデータベース管理システムにそれぞれ Zen Cart[1]と MySQL[2][3]を用いた EC サイトを対象として、特にアクセスが集中する状況下において、クライアントに対する EC サイトの応答時間 (データベース処理時間を含むウェブページの生成時間) を短縮する。

## 2. EC サイトにおけるデータベースアクセス

Zen Cart を用いた EC サイトでは、商品に関する情報や在庫数、顧客情報などはすべて MySQL で管理するデータベースに統合して記憶される。

MySQL では、データベースからの読み出しのみを行うクエリ (以下読み出しクエリ) については、複数のクエリを並列 (並行) に処理する。しかし、データの更新や追加、削除など、データベースへの書き込みを伴うクエリ (以下書き込みクエリ) に対しては、原則としてテーブル全体に対するロックをかけて処理を行う。書き込みクエリを受け付けると、データベースサーバはその時点で処理中のクエリの処理がすべて完了するのを待って、書き込み処理を開始する。また、書き込みクエリを受け付けた後は、その書き込み処理が完了するまで、同じテ

ーブルを参照する他のクエリの処理を待たせる。したがって、データベースにおける書き込み処理は、データベースサーバ内部での並列処理の機会を制限する要因となる。

EC サイトのクライアントが商品を買物カゴに入れたり、実際に購入手続きを行うと、買物カゴの内容や商品の在庫数の値を更新するための書き込みクエリが発生する。これはデータベース処理のスループットを低下させる要因となり得るが、この書き込みは商取引上の本質的な処理であり、また、多くの注文が短時間に集中しなければ、性能上の問題とはならない。クライアントからのアクセスの多くは、商品や商品リストのページを要求するもの (商品を購入する場合でも、まずはその商品ページを表示しなければならない) であり、本来、ウェブサーバはデータベースから必要なデータを読み出すだけでよい。しかし、実際には、販促等の目的で商品ページごとにその閲覧回数をカウントすることが多く、この場合、ウェブサーバは、クライアントからのアクセスのたびにデータベース内の閲覧回数の値を更新するクエリを送信する。データベースに対するこのような書き込みクエリが、他の読み出しクエリと混在して送信されるので、それらの書き込みクエリがデータベース処理のスループット低下の要因となる。

以上のように、実際には、データベースサーバ内部での並列処理機能を十分に活用できていない現状にある。そこで、商取引上で不都合を生じない範囲で、ウェブサーバが発する書き込みクエリと読み出しクエリの順序を入れ替え、複数の書き込みクエリを連続してデータベースサーバに送信することにより、スループットの低下を抑止する。

## 3. クエリのスケジューリング

### 3.1 スケジューリング方式の概要

本方式では、図 1 に示すように、ウェブサーバとデータベースサーバとの間にスケジューリングサーバを設置する。スケジューリングサーバは、ウェブサーバに対して自らがデータベースサーバであるかのように振舞い、また、データベースサーバに対しては自らがウェブサーバであるかのように振舞う。ウェブサーバから受信したクエリの順序を入れ替えてデータベースサーバへ転送し、また、データベースサーバからの応答メッセージをそのままウェブサーバへ転送する。

スケジューリングサーバ内には、ウェブサーバから受信した書き込みクエリを一時的に保存するためのキュー (以下書き込みキュー) を設ける。商品ページ等の閲覧回数を更新するための書き込みクエリを書き込みキュー

<sup>†</sup> 京都工芸繊維大学大学院 工芸科学研究科 情報工学専攻  
<sup>‡</sup> 京都工芸繊維大学大学院 工芸科学研究科 情報工学部門  
Dept. of Information Science, Kyoto Institute of Technology

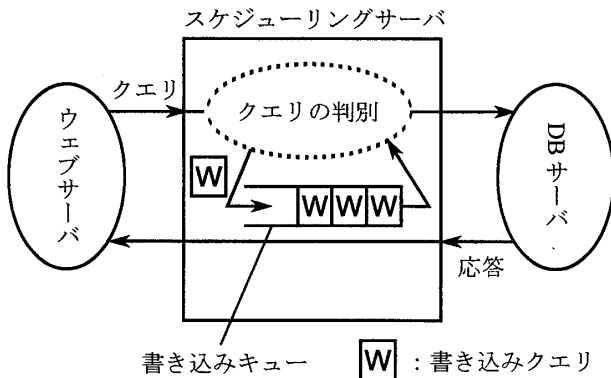


図1 スケジューリングサーバの設置

に一時的に格納し、これに後続する読み出しクエリ等を先にデータベースサーバへ転送する。書き込みキューにクエリが保存されている間は、商品ページの実際の閲覧回数とデータベースに記憶されている閲覧回数との間に不整合が生じるが、このような情報は、商品の在庫数等とは異なり、リアルタイムに更新される必要性はない。したがって、このような書き込みクエリに対する実際のデータベース処理を遅らせても実質的な問題は生じず、逆に、このような書き込みクエリが読み出しクエリの中に散発的に混在することによって引き起こされるスループットの低下を抑えることができる。

### 3.2 スケジューリングサーバの処理内容

スケジューリングサーバは、ウェブサーバからクエリを受信すると、まずそれが読み出しクエリであるか書き込みクエリであるかを判別する。受信したクエリが読み出しクエリの場合はそれをすぐにデータベースサーバへ転送し、書き込みクエリの場合はその書き込みクエリがスケジューリングの対象であるかどうかを判別する。そして、スケジューリング対象の書き込みクエリでなければそのままデータベースサーバへ転送し、スケジューリング対象の書き込みクエリであればスケジューリングサーバ内の書き込みキューに格納する。ただし、ウェブサーバは、書き込みクエリに対する応答を受信するまでページ生成処理を中断するので、スケジューリングサーバは、キューに保存する書き込みクエリに対してデータベースサーバが返すであろう応答メッセージを、あらかじめこの時点でウェブサーバに送信する。

受信した書き込みクエリがスケジューリングの対象かどうかを判別するには、本来、SQL文を解析する必要がある。しかし、スケジューリングサーバはウェブサーバとデータベースサーバとの間に位置し、スケジューリングサーバで扱うクエリはウェブサーバ上で動作するウェブショップ構築システムから送信されるものに限定されるため、書き込みクエリに対して複雑な解析を行う必要はない。よって、スケジューリングサーバは、簡単なパターンマッチングによって、受信した書き込みクエリがスケジューリングの対象かどうかを判別することができる。

書き込みキューに保存したクエリは、以下に挙げるいずれかの時点でデータベースサーバへ送信する。このとき、書き込みキュー内のすべてのクエリを連続して送信する。

- 書き込みキューに最初のクエリを格納してから、あらかじめ設定した時間（最大保存時間）が経過したとき。
- 書き込みキュー内のクエリの数、あらかじめ設定した最大数（最大保存件数）に達したとき。
- ウェブサーバからクエリを最後に受信してから、あらかじめ設定した時間（最大空白時間）が経過したとき。
- 転送したすべてのクエリに対する応答メッセージをデータベースサーバから受信したとき。
- スケジューリングの対象でない書き込みクエリをデータベースサーバに転送した直後。

### 4. 性能評価

Zen Cart と MySQL を用いて実験用の EC サイトを構築し、この EC サイトに多数のクライアントからのアクセスが集中する状況を再現した。この状況下で、クエリのスケジューリングを行わない場合と行う場合のそれぞれについて、ウェブサーバによるウェブページの生成時間（データベース処理に要する時間を含む）を測定した。スケジューリングを行う場合のスケジューリングサーバのパラメータについては、書き込みクエリの最大保存件数を 4,000 件、書き込みクエリの最大保存時間を 5 秒、ウェブサーバからのクエリ受信の最大空白時間を 1 秒とした。クエリのスケジューリングを行わない場合にウェブページの生成時間が平均で約 35.4 秒となる状況に対し、クエリのスケジューリングを行うとページの生成時間が平均で約 24.3 秒となり、約 30%の短縮が確認できた。このことから、クエリのスケジューリングが EC サイトの応答時間の短縮に有効であることがわかる。

### 5. むすび

EC サイトにおいては、データベースに対する書き込みと読み出しの順序を入れ替えても、商取引上で不都合を生じない場合がある点に着目し、クライアントに対する EC サイトの応答時間を短縮するために、ウェブサーバからのクエリをスケジューリングしてデータベースサーバへ送信する方式を提案した。アクセスが集中する状況下において、実際に本方式を用いることで EC サイトの応答時間を短縮できることを確認した。

今後は、書き込みクエリの最大保存時間や最大保存件数、ウェブサーバからのクエリ受信の最大空白時間などのパラメータを含めて、さらに詳細に検討してゆく予定である。

#### 謝辞

本研究の一部は日本学術振興会科学研究費補助金（基盤研究 (C) 21500053, 同 22500046 および若手研究 (B) 21700058) による。

#### 参考文献

- [1] Zen Ventures, LLC, "ecommerce shopping cart software by Zen Cart ecommerce solution", <http://www.zen-cart.com/>
- [2] Sasha Pachev, 伊藤直也, 田中慎司, 吉川英興, 菅野良二, "詳解 MySQL", オライリー・ジャパン (2007)
- [3] Sun Microsystems, "MySQL: The world's most popular open source database", <http://www.mysql.com/>.