CachePortal: Technology for Accelerating Database-driven E-commerce Web Sites

Wen-Syan Li   K. Selçuk Candan   Wang-Pin Hsiung   Oliver Po   Divyakant Agrawal
Qiong Luo   Wei-Kuang Wayne Huang   Yusuf Akça   Cemal Yılmaz

Computer and Communication Research Laboratories, NEC USA, Inc.
110 Rio Robles, M/S SJ100, San Jose, CA 95134, USA
Email:wen@ccrl.sj.nec.com   Tel:408-943-3008 Fax:408-943-3099

1   Introduction

Response time is one key point of differentiation among electronic commerce (e-commerce) Web sites. Snafu and slow-downs at major Web sites during special events or peak times demonstrate the difficulty of scaling up e-commerce sites. Such slow response times and down times can be devastating for e-commerce sites as indicated in a recent study by Zona Research[1] on the relationship between Web page download time and user abandonment rate. The study shows that only 2% of users will leave a Web site (i.e. abandonment rate) if the download time is less than 7 seconds. However, The abandonment rate goes up to 70% when the download time is around 12 seconds. This study clearly establishes the importance of fast response times to an e-commerce Web site to retain its customers.

One possible solution to this problem is to deploy network-wide caches so that a large fraction of requests can be served remotely rather than all of them being served from the origin Web site. However, for many e-commerce applications, HTML pages are created dynamically based on the current state of a business, such as product prices and inventory, typically stored in a database. The time to live (TTL) for these dynamic pages can not be estimated in advance. A typical database-driven Web site (illustrated on the right of Figure 1) consists of the following components:

1. A database management system (DBMS) to store, maintain, and retrieve all necessary data and information to model a business.

2. An application server (AS) that incorporates all the necessary rules and business logic to interpret the
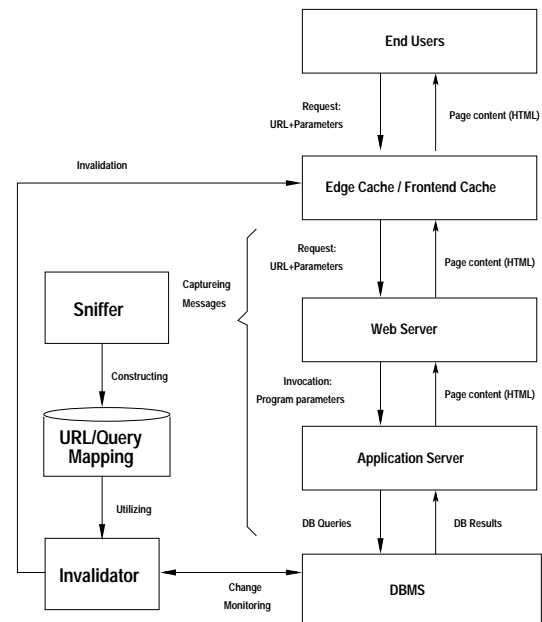
Figure 1: Architecture of a Database-Driven E-Commerce Site with *CachePortal* Technology

data and information stored in the database. AS receives user requests for HTML pages and depending upon the nature of a request may need to access the DBMS to generate the dynamic components of the HTML page.

3. A Web server (WS) which receives user requests and delivers the dynamically generated Web pages.

When a user accesses the Web site, the request and its associated parameters, such as the product name and model number, and cookie information for customization, are passed to an application server. The application server performs necessary computation to identify what kind of data it needs from the database or file system, or external data sources. Then the application server sends appropriate queries to the database or other sources. After the database returns the query results to the application server, the application uses these to prepare a Web page and passes it to the Web server, which then sends it

to the user.

Since cache servers, Web servers, application servers, and databases are independent components, currently there is no coordination mechanism to ensure that database content changes are reflected to the caches. As a result, most e-commerce sites have to specify dynamic contents as non-cacheable. Consequently, each request to an e-commerce site results in both network delay and server delays (i.e. WS delay, AS delay, and DBMS delay) since the request must be processed each time at the Web site.

2 Solutions

CachePortal [2] is a technology developed at NEC for accelerating database-driven e-commerce Web sites. The technology enables dynamic content caching by

- *automated* construction of the relationships between cached pages and database contents (i.e. *URL/Query Mapping*) via a *sniffer*; and

- intelligent monitor of database changes to "eject" (i.e. delete) impacted pages from caches via an *invalidator*.

The architecture of e-commerce Web sites deploying CachePortal is illustrated on the right of Figure 1. Compared with most existing work, where the relationships between Web pages and underlying data are specified manually, CachePortal features a *sniffer* to automatically generate the query/URL mapping. Compared with most existing work, where the invalidation checking is implemented using database trigger functions, the *invalidator* in CachePortal is implemented as an external software component based on incremental view maintenance techniques, which does not add substantial load to the underlying DBMS.

3 System Demonstration

At the conference, we demonstrate an e-commerce Web site in the area of housing. The Web site is implemented using Apache, BEA WebLogic Application Server, and Oracle DBMS and CachePortal is deployed for Web acceleration. The database used in this experiment contains 7 tables and more than 1M records. The update rate is 1 per second and each dynamic content page request results in a query with one join operation to the database. We demonstrate the performance gain of the Web site by adjusting *access rate*, *update rate*, and *hit ratio*. We also demonstrate the performance gain of the Web site based on various architectures. The performance difference between two Web sites, one deploys CachePortal and the other does not, is displayed using a progress bar as shown in Figure 2. We also demonstrate invalidation of affected pages (if any) after an update query is executed. Finally, we are also able to show some benchmark results as shown in Figure 3. In the figure, we see that the system architectures with various cache hit rates (lower three plots) with
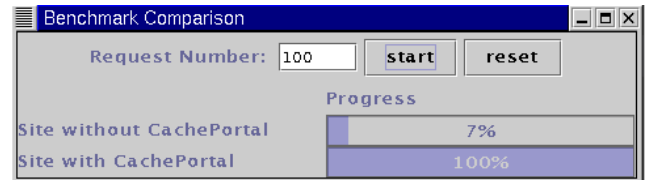

Figure 2: Demonstration on Performance Gain of Serving 100 Requests
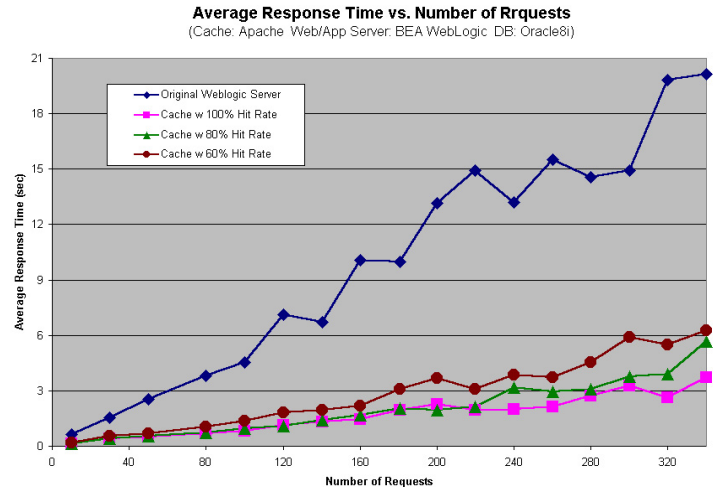

Figure 3: Demonstration of User Response Time Improvement Benchmark

CachePortal technology provide good response time consistently as the access rate increases whereas the system architecture without *CachePortal* technology (the upper plot) have poor performance at higher access rates.

4 Concluding Remarks

It is essential but difficult for e-commerce Web sites to handle the traffic arise at special events or peak times. Being evaluated and compared with many alternative solutions, CachePortal has shown that it can provide many e-commerce applications scalability and user response time improvement. This demonstration highlights many features of CachePortal technology using an e-commerce site built using some of most popular components on the market. It shows a real and useful application of integrating database and Web technologies.

References

[1] Zona Research, http://www.zonaresearch.com/

[2] K. Seluk Candan, Wen-Syan Li, Qiong Luo, Wang-Pin Hsiung, and Divyakant Agrawal. Enabling Dynamic Content Caching for Database-Driven Web Sites. In *Proceedings of the 2001 ACM SIGMOD International Conference on Management of Data, Santa Barbara, CA, USA*, ACM Press, 2001.