ISSN:0975-9646

A Dual Framework for Targated Online Data Delivery Algorithms

Ch.Srikanth[1], D.Naveen Kumar[2] , S.P. Anandaraj[3]

[1]M.Tech Student, [2,3]Assistant. Professor
Dept of CSE, SR Engineering College, Warangal, A.P., India

Abstract—A variety of emerging online data delivery applications challenge existing techniques for data delivery to human users, applications, or middleware that are accessing data from multiple autonomous servers. In this paper, we develop a framework for formalizing and comparing pull-based solutions and present dual optimization approaches. The first approach, most commonly used Nowadays, maximizes user utility under the strict setting of meeting a priori constraints on the usage of system resources. We present alternative and more flexible approach that maximizes user utility by satisfying all users. It does this while minimizing the usage of system resources. We discuss the benefits of this latter approach and develop an adaptive monitoring solution Satisfy User Profiles (SUPs). Through formal analysis, we identify sufficient optimality conditions for SUP. Using real(RSS feeds) and synthetic traces, we empirically analyze the behavior of SUP under varying conditions. Our experiments show that we can achieve a high degree of satisfaction of user utility when the estimations of SUP closely estimate the real event stream, and has the potential to save a significant amount of system resources. We further show that SUP can exploit feedback to improve user utility with only a moderate increase in resource utilization.

Index Terms—Distributed databases, online information services, client/server multitier systems, online data delivery.

1. INTRODUCTION

The diversity of data sources and Web services currently available on the Internet and the computational Grid, as well as the diversity of clients and application requirements, poses significant infrastructure challenges. In this paper, we address the task of targeted data delivery. Users may have specific requirements for data delivery, e.g., how frequently or under what conditions they wish to be alerted about update events or update values, or their tolerance to delays or stale information. The challenge is to deliver relevant data to a client at the desired time, while conserving system resources. We consider a number of scenarios including RSS news feeds, stock prices and auctions on the commercial Internet, and scientific data sets and Grid computational resources. We consider an architecture of a proxy server that is managing a set of user profiles that are specified with respect to a set of remote autonomous servers.

Push, pull, and hybrid protocols have been used to solve a variety of data delivery problems. Push-based technologies include Black Berry and JMS messaging, push-based policies for static Web content and push-based consistency in the context of caching dynamic Web content. Push is typically not scalable, and reaching a large number of potentially transient clients is expensive. In some cases, pushing information may over whelm the client with unsolicited information. Pull-based freshness policies have, therefore, been proposed in many contexts such as Web caching and synchronizing collections of objects, e.g., Web crawlers. Several hybrid push-pull solutions have also been presented .We focus on pull-based resource monitoring and satisfying user profiles.

2. MODEL FOR TARGETED DATA DELIVERY

Diverse data types are nowadays available as publications in RSS, including news and weather updates, blog postings, media postings, digital catalog notifications, promotions, white papers, and software updates. The use of RSS feeds is continuously growing and is supported by a pull-based protocol. RSS customization today is provided using specialized RSS readers (also known as RSS aggregators). A user of such a reader can customize her profile by specifying the rate of monitoring each RSS feed. Some readers even allow defining filtering rules over the RSS feed content which support further personalization. Recently, the RSS protocol was extended with special metatags such as server side TTL that hint when new updates are expected. We note that while this improves customization, server side hints such as TTL for static content delivery are not used often in other contexts, and was shown to be inefficient [17, 95]. Despite these features, a client who is only interested in being alerted of updates for a particular item in some news category, whenever the rate of updates increases to be at least twice as much as the usual rate, cannot specify such a profile using standard available RSS readers. This scenario requires further refined personalization that is currently unavailable. Our case study is that of RSS monitoring of CNN News, providing publications of news updates by CNN on various topics such as world news, sports, finance, etc. Typically, only news article titles are provided in an RSS item and a link directs the reader to the original article. Each item also has a time stamp of its publication date, and sometimes, CNN also provides TTL tags.

3. SYSTEM ANALYSIS

The diversity of data sources and Web services currently available on the Internet and the computational Grid, as well as the diversity of clients and application requirements, poses significant infrastructure challenges. In this paper, we address the task of targeted data delivery. Users may have specific requirements for data delivery, e.g., how frequently or under what conditions they wish to be alerted about update events or update values, or their tolerance to delays or stale information. The challenge is to deliver relevant data to a client at the desired time, while

conserving system resources. We consider a number of scenarios including RSS news feeds, stock prices and auctions on the commercial Internet, and scientific data sets and Grid computational resources. We consider architecture of a proxy server that is managing a set of user profiles that are specified with respect to a set of remote autonomous servers.

With this class of problems, user needs are set as the constraining factor of the problem, while resource consumption is dynamic and changes with needs. We present an optimal algorithm in the OptMon2 class, namely, Satisfy User Profiles (SUPs).

OptMon2 can be stated as the following general formulation

- Minimize system resource usage
- Satisfying user profiles(SUP).

SUP is simple yet powerful in its ability to generate optimal scheduling of pull requests. at each time point, it can get additional requests for resource monitoring. Through formal analysis, we identify sufficient conditions for SUP to be optimal given a set of updates to resources.

4. MODULES

- Registration
- Sign-Up
- Notifications
- Weblog

Registration

- It is a registration of people who are interested in finding the information.
- It provides the complete information who is interested to know the information.
- They also provide the profile of various information of the related field, so that to drag the information which is much related to the user.

Sign-Up (Authenticating)

- Only to the registered user sign up and find the wanted information.
- As per the required profile of the user and the admin sends the information directly in the required field with the help of the weblog.
- Thus the required information is gathered based on the required the data.

Weblog (Collecting User Profiles)

Profiles are declarative user specifications for data delivery. A profile should be easy to specify and sufficiently rich to capture client requirements. A profile should have clear semantics and be simple to implement. We assume that every resource r 2 R has a unique identifier (e.g., URI) and can be described using some schema (e.g., Relational Schema, DTD, XML-Schema, RDF-Schema, etc.) A resource can be queried using a suitable query language.

Notifications (Notify user needs)

Clients use notification rules to describe their data needs and express the utility they assign with data delivery. A notification rule extends the Event-Condition- Action (ECA) structure in active databases η and can be modified dynamically by the user. A notification rule η is a quadruple hQ; Tr; T ; Ui.Qis a query written in some suitable query language. T is the epoch in which rules are evaluated. Finally, U is a utility expression specifying the utility client gains by notifications of Q. A notification query Q is specified over a set of resources from the profile domain denoted by Domain; η Þ. Queries are equivalent to actions in the ECA structure.

5. SYSTEM DESIGN

Software design sits at the technical kernel of the software engineering process and is applied regardless of the development paradigm and area of application. Design is the first step in the development phase for any engineered product or system. The designer's goal is to produce a model or representation of an entity that will later be built. Beginning, once system requirement have been specified and analyzed, system design is the first of the three technical activities -design, code and test that is required to build and verify software.

The importance can be stated with a single word "Quality". Design is the place where quality is fostered in software development. Design provides us with representations of software that can assess for quality. Design is the only way that we can accurately translate a customer's view into a finished software product or system. Software design serves as a foundation for all the software engineering steps that follow. Without a strong design we risk building an unstable system – one that will be difficult to test, one whose quality cannot be assessed until the last stage.

During design, progressive refinement of data structure, program structure, and procedural details are developed reviewed and documented. System design can be viewed from either technical or project management perspective. From the technical point of view, design is comprised of four activities – architectural design, data structure design, interface design and procedural design.

6. SCREEN SHOTS

Fig 6.0 login page

Login page is used to login the different users by entering the User Id and Password and click login option to enter into user page. This page also used to sign up the new users.

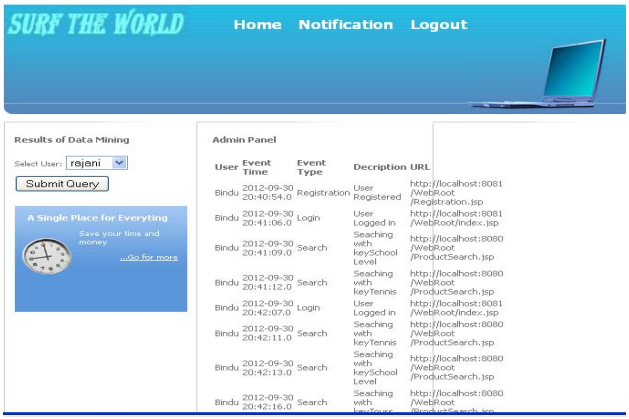

Fig 6.1 Profile tracking

Profile tracking page is used by the admin to sea the details of different user profiles and send the information regarding the user needs.

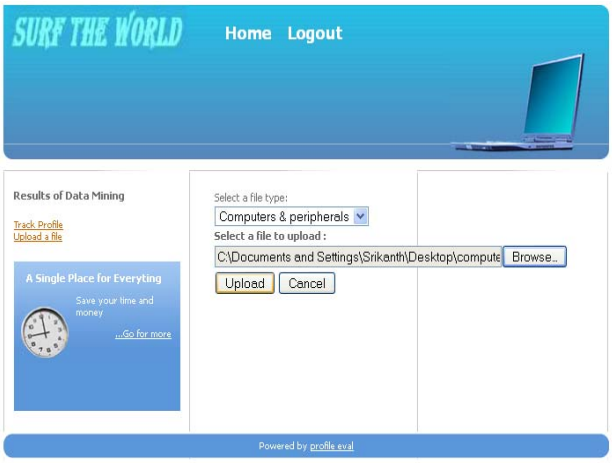

Fig 6.2 Upload Page

Upload page is used to send the information to user profiles. In this page select the file type and browse the upload file and click the upload button.

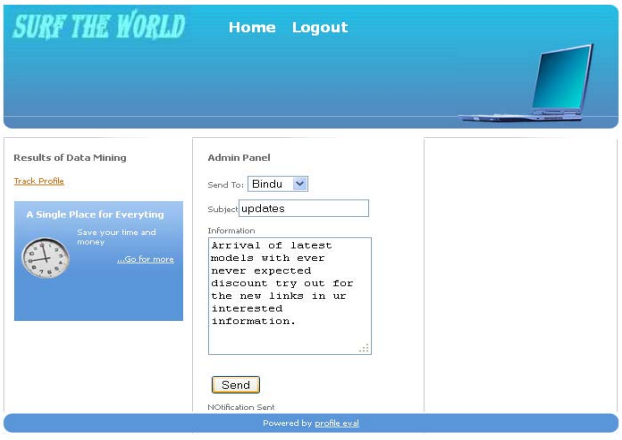

Fig 6.3 Notification Page

The admin can send updated notifications to the users by using the Notification Page. In this page Admin, select the User and type the updated information and click the send button.

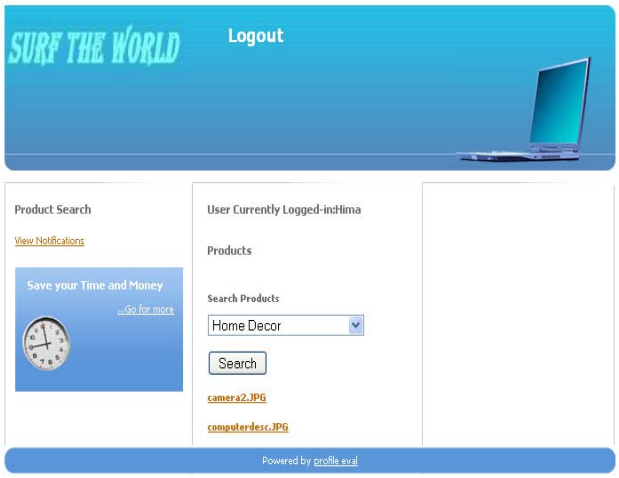

Fig 6.4 Search Page

This page is used by users to search the different product details and also view the notifications.

Fig 6.5 Retrieved page

This page is appeared when the user search for any product information.

Fig 6.6 View Notifications page

This page is used by user for see the notifications sended by admin.

7. SYSTEM TESTING

Information Processing has undergone major improvements in the past two decades in both hardware and software. Hardware has decreased in size and price, while providing more and faster processing power. Software has become easier to use, while providing increased capabilities. There is an abundance of products available to assist both end-users and software developers in their work. Software testing, however, has not progressed significantly. It is still largely a manual process conducted as an art rather than a methodology. It is almost an accepted practice to release software that contains defects.

Software that is not thoroughly tested is released for production. This is true for both off-the-shelf software products and custom applications. Software vendor and in-house systems developers release an initial system and then deliver fixes to the code. They continue delivering fixes until they create a new system and stop supporting the old one. The user is then forced to convert to the new system, which again will require fixes.

In-house systems developers generally do not provide any better level of support. They require the users to submit Incident Reports specifying the system defects. The Incident Reports are then assigned a priority and the defects are fixed as time and budgets permit.

8. CONCLUSION

We focused on pull-based data delivery that supports user profile diversity. Minimizing the number of probes to sources is important for pull-based applications to conserve resources and improve scalability. Solutions that can adapt to changes in source behavior are also important due to the difficulty of predicting when updates occur. We have addressed these challenges through the use of a new formalism of a dual optimization problem (OptMon2), reversing the roles of user utility and system resources. This revised specification leads naturally to a surprisingly simple, yet powerful algorithm (SUP) which satisfies user specifications while minimizing system resource consumption. We have formally shown that SUP is optimal for OptMon2 and under certain restrictions can be optimal for OptMon1 as well. Using RSS data traces as well as synthetic data, that SUP can satisfy user profiles and capture more updates compared to existing policies. SUP is adaptive and can dynamically change monitoring schedules. Our experiments show that using feedback in SUP improves the performance with a moderate increase in the number of needed probes.

We believe that the main impact of this work will be in what is now known as the Internet of things, where sensor data are collected, analyzed, and utilized in many different ways, based on user's needs. With the Internet of things, user profiles, and their satisfaction dictate the way data are utilized, and monitoring sensor data efficiently is a mandatory prerequisite to the creation of any information system that is based on such data. OptMon2 is defined in such a way that satisfaction of a user profile is a hard constraint.

REFERENCES

1. Adi and O. Etzion, "Amit—The Situation Manager," Int'l J. Very Large Data Bases, vol. 13, no. 2, pp. 177-203, May 2004.
2. L. Bright, A. Gal, and L. Raschid, "Adaptive Pull-Based Policies for Wide Area Data Delivery," ACM Trans. Database Systems, vol. 31, no. 2, pp. 631-671, 2006.
3. L. Bright and L. Raschid, "Using Latency-Recency Profiles for Data Delivery on the Web," Proc. Int'l Conf. Very Large Data Bases (VLDB), pp. 550-561, Aug. 2002.
4. D. Carney, S. Lee, and S. Zdonik, "Scalable Application-Aware Data Freshening," Proc. IEEE CS Int'l Conf. Data Eng., pp. 481-492, Mar. 2003.
5. L.S. Chandran, L. Ibarra, F. Ruskey, and J. Sawada, "Generating and Characterizing the Perfect Elimination Orderings of a Chordal Graph," Theoretical Computer Science, vol. 307, no. 2, pp. 303-317, 2003.
6. M. Cherniack, E. Galvez, M. Franklin, and S. Zdonik, "Profile-Driven Cache Management," Proc. IEEE CS Int'l Conf. Data Eng., pp. 645-656, Mar. 2003.
7. J. Cho and H. Garcia-Molina, "Synchronizing a Database to Improve Freshness," Proc. ACM SIGMOD, pp. 117-128, May 2000.
8. J. Cho and A. Ntoulas, "Effective Change Detection Using Sampling," Proc. Int'l Conf. Very Large Data Bases (VLDB), 2002.
9. E. Cohen and H. Kaplan, "Refreshment Policies for Web Content Caches," Proc. IEEE INFOCOM, pp. 1398-1406, Apr. 2001.

AUTHORS

Srikanth Chennuri currently pursuing M.Tech in Computer Science and Engineering at SR Engineering College, Warangal, AndhraPradesh, India. His research interest includes Data mining, Distributed Databases, Client Server multi tier systems.

D.Naveen Kumar has 3+ years of experience in teaching field. His areas of interest include Data Mining, Information Security. He has published research papers in various National and International Journals, National and International Conferences. He also attended many National Seminars/FDP/Workshops Etc

S.P. Anandaraj has 9+ years of experience in teaching field. His areas of interest include Data Mining, Mobile Adoch Networks and Embedded Systems.He has published research papers in various National and International Journals, National and International Conferences. He also attended many National Seminars/FDP/Workshops Etc