Optimized XML Data Management for Mobile Transactions

Stefan Böttcher

University of Paderborn, Computer Science, Fürstenallee 11, 33102 Paderborn, Germany
stb@uni-paderborn.de
http://wwwcs.upb.de/cs/boettcher/

Abstract

This paper summarizes the research issues in the area of XML data management and mobile transactions performed at the University of Paderborn. We shortly describe the problems worked at and the basic solution ideas, and we refer to the further literature that provides detailed descriptions of the elaborated solutions.

1 Introduction

Whenever (XML) data is shared between multiple clients in mobile ad-hoc networks, database technologies face new challenges. For example, transaction management has to consider unpredictable disconnections of clients and even that there may be no stable server – nevertheless applications may require transactional guarantees. Furthermore, limited bandwidth and limited energy supply require new approaches to query processing which includes data exchange, replication, and caching of previous query results. Furthermore, data interchange between (mobile) participants that use different XML data formats requires specific query optimization techniques that limit the amount of data, which has to be transformed and exchanged between different participants. Within this extended abstract, we focus on the research that has been contributed to this field at the University of Paderborn and summarize some of the main results.

2 Bandwidth optimization and caching

We have optimized data exchange needed for query processing and for transaction synchronization, while still guaranteeing correct answers and serializability [1]. When a mobile client submits a query to the server within our approach, the server reconstructs or re-computes the data stored in the client's cache at the server side, such that only XML difference fragments have to be submitted to the client. Furthermore, we have investigated how to integrate XML difference fragments and cached query results with mobile clients [2].

Dagstuhl Seminar Proceedings 04441
Mobile Information Management
http://drops.dagstuhl.de/opus/volltexte/2005/268

3 Synchronization and cache consistency of mobile transactions

As mobile clients may lose their connection to a database server, we do not grant locks to resources accessed by mobile clients. Instead, we use an optimistic approach inspired by ideas of [3] that even integrates with centralized or distributed servers in wired networks that use a lock-based protocol [4]. Our approach allows unprepared disconnections of clients at any time without blocking the server, it allows repairing lost connections, i.e. to continue interrupted transactions, and it allows even for cross-transaction optimization, i.e. to use cached and outdated data of previous transactions [5]. Cache consistency is checked and guaranteed within this approach as a result of the validation step during the transactions' commit-request.

4 Cooperative caching and mobile ad-hoc networks

When we investigated how our approach of reducing data exchange to XML difference fragments can be extended to cooperative caching, we discovered that XPath is too complex a language to describe cached results of mobile clients, because the execution time needed for containment tests and for computing the missing fragments exhausts the mobile devices' power and time resources. Therefore, we have designed XML patterns that represent XPath fragments and support an efficient computation of all operations on XML fragments that are needed to support cooperative caching between neighbor clients [6].

Furthermore, we have integrated data management technologies into the network layer of the PaMANET, i.e. the Paderborn Mobile Ad-hoc NETwork, which offers XML data access to a network of mobile clients that communicate via IPv6 [7].

5 Distributed XPath query optimization based on containment tests and intersection tests of XPath expressions

Further work includes proof techniques for subclasses of XPath expressions, the results of which can applied to access control and synchronization techniques for XML database transactions as well as for XPath query optimization. For the purpose of XPath query optimization, we have implemented a subset test, which checks $Q1 \subseteq Q2$ for two XPath expressions Q1 and Q2 [8], and we have implemented an intersection test checking $Q1 \cap Q2$ [9]. If the data, that we need to answer a query, can be described by Q1, and we have a previous query result Q2 stored in a cache, and we can prove $Q1 \subseteq Q2$, we have the option to use the previous query result Q2 instead of reloading the result for Q1 from a remote location. Similarly, a query to a remote location that contains at most data that is described by an XPath expression Q2 need not be searched for queries that need data which are described by Q1, if we can prove that $Q1 \cap Q2 = \emptyset$. Different from other approaches to containment tests and satisfiability test for XPath expressions (e.g. [10]), we use fast but incomplete testers which include DTD information and use a graph-based search algorithm.

6 Query optimization on XML views

Finally, we have optimized the XML data exchange between platforms that need the same XML data, but in a different XML format. Whenever a given XML document (e.g. a product catalog) in the XML format F1 is imported by a different application (or a different company) that needs the data in a different XML format F2, it is common to use XML transformers (e.g. XSLT or XQuery). Instead of transforming the whole XML document D, it may be advantageous to restrict the data that is transformed by a transformer T from F1 into F2 to the needs of a given query Q. In other words, given Q and T, we compute a query Q2 such that Q(T(D)) = Q(T(Q2(D))). As a result, we have to transform only a much smaller fragment Q2(D), and we have to submit only this smaller fragment to the other company (or to the other application) [11]. In comparison to [12], our work includes the use of filters, the transformation of which is based on work of [13], and different from both, our approach optimizes queries through XPath views. Recent research achieved the result that the optimization of XPath queries on XSLT views can be extended to queries in one of the languages XQuery, XSLT or XPath applied to data that is transformed by a view given either in XSLT or in XQuery [14]. The extension of our query optimization to XQuery views is uses an XQuery to XSLT translation that has been inspired by ideas of [15]. We consider this to be especially helpful for applications where bandwidth is limited and a possible bottleneck.

References

1. Stefan Böttcher, Adelhard Türling. XML Fragment Caching for Small Mobile Internet Devices. 2nd International Workshop on Web-Databases. Erfurt, Germany, October, 2002. Springer, LNCS 2593, Heidelberg, 2003.
2. Stefan Böttcher, Adelhard Türling. Caching XML Data for Mobile Web Clients. International Conference on Internet Computing IC'04, Las Vegas, June 2004.
3. Manuel Reimer: Solving the Phantom Problem by Predicative Optimistic Concurrency Control, 9th VLDB, Florenz, 1983.
4. Stefan Böttcher, Adelhard Türling. Transaction Synchronisation for XML Data in Client Server Web Applications. In: Informatik 2001 , Tagungsband der GI-Jahrestagung 25.-28. September 2001, Vienna, Austria , Volume 1, pages 388-395.
5. Stefan Böttcher. Repairing lost connections of mobile transactions with minimal XML data exchange. IFIP TC8 Working Conference on Mobile Information Systems (MOBIS). Oslo, Norway, September 2004.
6. Adelhard Türling, Stefan Böttcher. Search tree patterns for mobile and distributed XML processing. 2nd Int. VLDB Workshop on Databases, Information Systems and Peer-to-Peer Computing (DBISP2P). Toronto, Canada, August 2004.

7. Christian Schindelhauer, Stefan Böttcher, Franz Ramming, Peter Bleckmann, Eduardas Cesnavicius, Andre Francisco, Tim Daniel Hollerung, Birger Kuhnel, Jing Liu, Sebastian Obermeier, Simon Oberthür, Felix Peter, Gunnar Schomaker, Thorsten Steenweg, Qamar Abas Tarar, Marcel Tiemeyer, Adelhard Türling, Arne Vater. The Design of PAMANET the Paderborn Mobile Ad-hoc Network (Extended Abstract/Poster). ACM International Workshop on Mobility Management and Wireless Access. Philadelphia, USA, September/October 2004.
8. Stefan Böttcher, Rita Steinmetz. Testing Containment of XPath Expressions in order to Reduce the Data Transfer to Mobile Clients. 7th East-European Conference on Advances in Databases and Information Systems, Dresden, Germany, September 2003.
9. Stefan Böttcher. Testing Intersection of XPath Expressions under DTDs. International Database Engineering & Applications Symposium, Coimbra, Portugal, July 2004.
10. Gerome Miklau, Dan Suciu: Containment and Equivalence for an XPath Fragment. PODS 2002: 65-76.
11. Sven Groppe, Stefan Böttcher, Georg Birkenheuer. Efficient Querying of Transformed XML Documents. 6th International Conference on Enterprise Information Systems, Porto, Portugal, April 2004.
12. Amelie Marian, Jerome Simeon: Projecting XML Documents, VLDB 2003.
13. Olteanu, D., Meuss, H., Furche, T., Bry, F., XPath: Looking Forward, *XML-Based Data Management (XMLDM)*, EDBT Workshops, Prague, Czech Republic, 2002.
14. Sven Groppe, Stefan Böttcher. Query Reformulation for the XML Standards XPath, XQuery and XSLT. XSW 2004 - The Workshop on XML Technologies for the Semantic Web, Berlin, Germany, October 2004.
15. Lechner, S., Preuner, G., and Schrefl, M., Translating XQuery into XSLT, In *ER 2001 Workshops*, Yokohama, Japan, 2001.