

Function Proxy: Template-Based Proxy Caching for Table-Valued Functions

Qiong Luo Wenwei Xue
Department of Computer Science
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong, China
{luo, wxue}@cs.ust.hk

1. Introduction

Traditionally, user-defined functions are employed in database applications that handle complex data, such as Geographical Information Systems and Computer Aided Design systems. With the proliferating use of the Web, certain types of database-backed web sites also heavily utilize user-defined functions in answering user queries. While recent research [1] has shown the usefulness of advanced web proxy caching for general database-backed web sites, their techniques are not directly applicable to web sites with *function-embedded queries*, which are queries with embedded calls to user-defined functions.

In this work, we demonstrate advanced proxy-caching techniques for these function-embedded queries. We term our proxy a *function proxy*. It does not only service a function-embedded query that is an exact match to a previous one, but also services a query that can be answered by processing results of previous queries. We focus on *table-valued functions* (which return a set of tuples) as opposed to *scalar functions* (which return a scalar value), because the former brings additional challenges as well as opportunities for our function proxy.

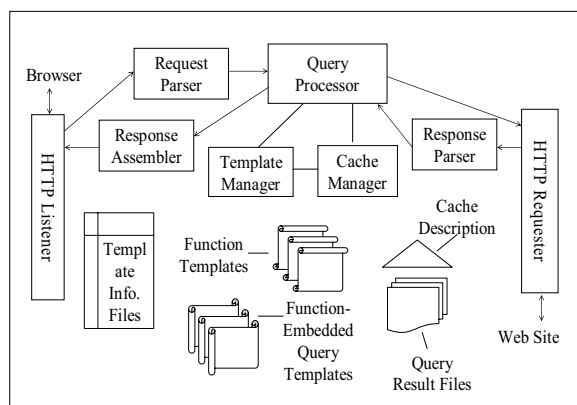


Figure 1. System architecture

Figure 1 illustrates the system architecture of our function proxy. The *function templates* capture the semantics of user-defined functions and the *function-embedded query templates* define the types of queries that the proxy answers. These templates are a natural abstraction of the HTML forms for accepting user queries at the web site, the embedded user-defined functions with parameters, and the parameterized queries programmed in the web site applications. They enable the proxy to decide on whether and how to cache the query results and to answer new queries by processing the previous results. They also allow the proxy to optimize its cache organization and its limited query processing logic accordingly.

2. Demonstration

We implemented our function proxy as a Java servlet on top of the Apache Tomcat servlet engine version 4.1.24. We demonstrate our function proxy using real query traces extracted from the SkyServer [3] web logs. These query traces are from two search forms in the SkyServer which correspond to two table-valued function embedded queries. We run these traces through our proxy to the SkyServer web site. We also run synthetic workloads through our proxy to a personal SkyServer [2] site that is downloaded and set up in our department. We show statistics about the query answering status along with the returned query result, such as the query text, the cache status, the number of result tuples, and the time breakdown in the proxy.

References

- [1] Qiong Luo, Jeffery F. Naughton. Form-Based Proxy Caching for Database-Backed Web Sites. VLDB 2001: 191-200.
- [2] Personal SkyServer. http://research.microsoft.com/~gray/SDSS/personal_sky_server.htm
- [3] The SkyServer web site. <http://skyserver.sdss.org/>