Chapter VI

Predicate Based Caching for Large Scale Mobile Distributed On-line Applications

Abhinav Vora, Zahir Tari, and Peter Bertok
RMIT University, Australia

ABSTRACT

Robust mobile middlewares are crucial for online applications as they provide solutions for the core issues of mobility, data interoperability, and security. This chapter describes our experience in designing such middlewares for one of the largest Australian transport companies (CMS Transport Systems). We focus on the design of a predicate-based caching technique for mobile object-based middlewares that optimises the performance of the mobile medium by better utilising the available bandwidth.

Several caching techniques have been proposed to improve system and application performance. Such techniques, along with consistency control mechanisms, are used to reduce the communication load between clients and servers, which is particularly important in wireless networks. Caching techniques are generally classified as either ID-based or predicate-based. In this chapter we propose a predicate-based caching scheme, in which the predicates are used in combination with updates and are broadcast by servers in a set of
appropriate messages called cache invalidation reports. Each report/message contains information about the data items that have been updated in the server during a given period. A function mapping the predicate into binary representation is defined for each attribute. Because not all updates are relevant to a cache, there is a matching algorithm for detecting relevancy between the cache predicate and the predicates in the cache invalidation reports. The predicate-based cache invalidation reports inform the client cache manager concisely about items that need to be refreshed and about those that need to be discarded, and ensure efficient bandwidth usage.

**MOTIVATION**

With mobile computing devices becoming smaller and faster coupled with networks supporting higher bandwidth and reliability, mobile computing is playing an increasingly important part in how applications are designed today. Handheld devices enable users to actively participate in distributed computing while on the move. Communication in such distributed, in many cases wireless, environments is hampered by intermittent and weak connections; in particular, mobile devices encounter wide variations from high bandwidth, low latency through to low bandwidth, high latency and to possibly no connectivity at all (Forma and Zahorjan, 1994).

To avoid high cost of connection or overcome availability problems mobile clients often deliberately disconnect for some time and then reconnect to the network. It is as essential for mobile clients to keep operating while disconnected as it is to support operations when the connection is weak, since there may be clients that never get good connection and still need to operate. Caching is one technique that can be used to support disconnected operations as well as improve performance for mobile clients. Caching objects not only improves response time but also reduces the number of messages between clients and server during operation, which is particularly important due to the high costs and unreliability of connections in a wireless network.

Mobile and portable computing devices have relatively limited memory, processing, and power resources. This limits the ability of the mobile device to cache a large quantity of data. Caching and cache consistency algorithms need to take these factors into concern. Reducing the number and size of messages used for maintaining cache consistency can lead to performance improvement in the form of reduced battery and processing requirements in addition to the saved wireless bandwidth.

Existing caching techniques can be broadly classified into ID-based and predicate-based. Both use cache invalidation reports (server broadcast messages), which are used by the clients for maintaining the consistency of their cache. In addition to other information, a cache invalidation report contains a list of identifiers of data items that have been modified since broadcasting of the last message. ID-
Related Content

E-Government and Its Impact on E-Commerce in LDCs
www.igi-global.com/chapter/government-its-impact-commerce-ldcs/10113?camid=4v1a

Process Simulation for E-Commerce Systems
www.igi-global.com/chapter/process-simulation-commerce-systems/12654?camid=4v1a

Cultural Differences, Information Technology Infrastructure, and E-Commerce Behavior: Implications for Developing Countries
www.igi-global.com/chapter/cultural-differences-information-technology-infrastructure/10115?camid=4v1a