Chapter 7.32
Ensuring Serializability for Mobile-Client Data Caching

Shin Parker
University of Nebraska at Omaha, USA

Zhengxin Chen
University of Nebraska at Omaha, USA

IMPORTANCE OF ENSURING SERIALIZABILITY IN MOBILE ENVIRONMENTS

Data management in mobile computing has emerged as a major research area, and it has found many applications. This research has produced interesting results in areas such as data dissemination over limited bandwidth channels, location-dependent querying of data, and advanced interfaces for mobile computers (Barbara, 1999). However, handling multimedia objects in mobile environments faces numerous challenges. Traditional methods developed for transaction processing (Silberschatz, Korth & Sudarshan, 2001) such as concurrency control and recovery mechanisms may no longer work correctly in mobile environments. To illustrate the important aspects that need to be considered and provide a solution for these important yet "tricky" issues in this article, we focus on an important topic of data management in mobile computing, which is concerned with how to ensure serializability for mobile-client data caching. New solutions are needed in dealing with caching multimedia data for mobile clients, for example, a cooperative cache architecture was proposed in Lau, Kumar, and Vankatesh (2002). The particular aspect considered in this article is that when managing a large number of multimedia objects within mobile client-server computing environments, there may be multiple physical copies of the same data object in client caches with the server as the primary owner of all data objects. Invalid-access prevention policy protocols developed in traditional DBMS environment will not work correctly in the new environment, thus, have to be extended to ensure that the serializability involving data updates is achieved in mobile environments. The research by Parker and Chen (2004) performed the analysis, proposed three extended protocols, and conducted experimental studies under the invalid-access prevention policy

in mobile environments to meet the serializability requirement in a mobile client/server environment that deals with multimedia objects. These three protocols, referred to as extended server-based two-phase locking (ES2PL), extended call back locking (ECBL), and extended optimistic two-phase locking (EO2PL) protocols, have included additional attributes to ensure multimedia object serializability in mobile client/server computing environments. In this article, we examine this issue, present key ideas behind the solution, and discuss related issues in a broader context.

BACKGROUND

In a typical client-server computing architecture, there may exist multiple physical copies of the same data object at the same time in the network with the server as the primary owner of all data objects. The existence of multiple copies of the same multimedia object in client caches is possible when there is no data conflict in the network. In managing multiple clients' concurrent read/write operations on a multimedia object, no transactions that accessed the old version should be allowed to commit. This is the basis of the invalid-access prevention policy, from which several protocols have been proposed. The purpose of these protocols is to create an illusion of a single, logical, multimedia data object in the face of multiple physical copies in the client/server network when a data conflict situation arises. When the server becomes aware of a network-wide data conflict, it initiates a cache consistency request to remote clients on behalf of the transaction that caused the data conflict. Three well-known invalid-access prevention protocols are Server-based Two-Phase

Figure 1. CBL failure analysis tree in a mobile environment