Wireless Networks 8, 391–402, 2002

Disconnection Modes for Mobile Databases

JOANNE HOLLIDAY
Department of Computer Engineering, Santa Clara University, Santa Clara, CA 95053, USA

DIVYAKANT AGRAWAL and AMR EL ABBADI
Department of Computer Science, University of California at Santa Barbara, Santa Barbara, CA 93106, USA

Abstract. As mobility permeates into todays computing and communication arena, we envision application infrastructures that will increasingly rely on mobile technologies. Traditional database applications and information service applications will need to integrate mobile entities: people and computers. In this paper, we develop a distributed database framework for mobile environments. A key requirement in such an environment is to support frequent connection and disconnection of database sites. We present algorithms that implement this framework in an asynchronous system.

Keywords: databases, data consistency, mobility, replication, disconnected operation

1. Introduction

As mobility permeates into todays computing and communication arena, we envision application infrastructures that will increasingly rely on mobile technologies. Current mobility applications tend to have a large central server and use mobile platforms only as caching devices. We want to elevate the role of mobile computers to first class entities in the sense that they allow the mobile user work/update capabilities independent of a central server. In such an environment, several mobile computers may collectively form the entire distributed system of interest. These mobile computers may communicate together in an ad hoc manner by communicating through networks that are formed on demand. Such communication may occur through wired or wireless networks. At any given time, a subset of the computer collection may connect and would require reliable and dependable access to data of interest.

In this paper, we consider a distributed database that can be made up entirely of mobile components. These component sites are peers and the database may be replicated for fault-tolerance or availability or both. Thus we have a distributed replicated database where all sites must participate in the synchronization of transactions. The capabilities of the distributed replicated database are extended in this paper by allowing a site or team member to plan to disconnect so that the remaining sites are minimally disrupted. A “check-out” mode is described that allows for independent update of a portion of the database by a disconnected member. These updates are automatically synchronized and integrated into the database upon reconnection. A “relaxed check-out” mode permits increased concurrency at the expense of serializability, the traditional notion of correctness in database management systems. An “optimistic check-out” mode further increases concurrency, but requires reconciliation rules. An implementation in an asynchronous system is discussed. Our approach emphasizes the mobile aspect of all system components, and hence the lack of a centralized server.

Walborn and Chrysanthis [27] describe the use of mobile computers in the trucking industry. Each truck has a computer with a satellite or radio link and interacts with the corporate database. Other applications involving remote or disaster areas and military applications have mobile computers forming ad hoc networks without communication with stationary computers. Faiz and Zaslavsky [8] discuss the impact of wireless technologies and mobile hosts on a variety of replication strategies. Distributed replicated file systems such as Ficus [21] and Coda [17] have extensive experience with disconnected operations. The problem of failed sites and the related issue of network partitioning has been investigated extensively and many of the results apply to disconnected operation.

In this paper, we accommodate voluntary disconnection, by using the notion of planned disconnection, which has been shown to be of particular use in asynchronous systems [16]. In section 2, we present several alternative disconnection modes. In section 3, we present our model of a distributed database system. In section 4, we explain in detail how the various modes will operate. In section 5, we show how to implement the modes in an asynchronous system. In section 6, we discuss related work and compare it with ours. We conclude with section 7.

2. Motivation

In the system we are considering, the components of the distributed database are laptop computers belonging to members of a small team. The team members maintain a database which may or may not be replicated at each of their laptop computers. Normal operations ensure that all team members are completely synchronized with each other’s progress. Team members gather together in various locations and form ad hoc networks to perform their work. These ad hoc networks may be wired or wireless. Several varieties of wireless LANs are already available. Examples include NCR’s Wave-