Fault-Tolerant Quorum Consensus Scheme for Replication Control in Mobile Distributed Database Systems: FTQC

Siwoo Byun
Songchun Moon

Korea Advanced Institute of Science and Technology

We propose a new replication control scheme for multiple-copy consistency in mobile distributed database systems. Replicating data at multiple sites could contribute to enhance the availability regarding the distributed data. Replicating data, however, inevitably induces the burden of maintaining replica consistency which requires more complex synchronization mechanism, especially in the presence of communication failures. If a communication link fails, all sites in the overall network are divided into two disjoint groups: the major group called quorum partition and the minor group called non-quorum partition. Note that this phenomenon called network partitioning could lead to access starvation in the non-quorum partition. In order for mobile users to access global databases without access starvation even in the presence of network partitioning, we propose a new scheme called fault-tolerant quorum consensus (FTQC). FTQC is based on the idea that quorum formation is allowed even in the non-quorum partition by circulating a single global token. FTQC guarantees a new form of one-copy serializability, inter-partition one-copy serializability, without sacrificing data availability. Based on the results of the performance evaluation, we conclude that the protocols which exploit FTQC scheme outperform the protocols which never exploit FTQC.

Recent advances in computing and networking technologies have made an extensive use of inexpensive portable computers and enabled sharing of on-line information via wireless communication channels. This new computing paradigm, called mobile computing, allows users to perform online transaction processing independent of their physical locations [Alonso and Korth, 1993; Imielinski and Badrinath, 1992; Pitoura and Bhargava, 1994]. Generally, such a mobile computing architecture includes two distinct sets of entities: mobile hosts (MHs) in the wireless network and fixed hosts (FHs) in the wired network (Figure 1).

The MHs can dynamically move within a radio coverage area called a cell or between two cells while retaining their network connection. The average cell size is about two miles in diameter, and an MH is crossing through these cells tens of times a day. The FHs are steadily connected to the wired network and some of them, called mobile support stations (MSSs), are augmented with a wireless interface to communicate with the MHs. Normally, a single MSS is able to support a number of MHs, and is engaged to provide services such as data passing and message interpretation to the MHs positioned only within its cell. A location server (LS) is also a FH that is responsible for keeping track of addresses of MHs to detect the geographical location where the MH is located. Generally, there is a hierarchy of LSs which are connected among themselves and to MSSs by the wired network. Each MH includes several applications such as groupwork tool and one small DBMS which performs basic tasks to manage database consistency regarding transactions issued by the local applications.

The users could request information and receive responses at any places owing to the MH. The MH, however, could suffer from unreliable and ill-timed services due to the characteristics of wireless media. Since the current wireless
network generally has unreliable channel and narrow bandwidth, communications between MHs and an MSS are prone to be delayed due to the low transfer rate, and often disconnected due to unpredictable electronic-noise interventions. One way that reduces possibility of the undesirable services is replicating data at multiple MHs.

Replicating data at multiple independent sites could contribute to enhance the availability regarding the distributed data. Replicating data, however, inevitably induces the burden of maintaining replica consistency which requires more complex and more expensive fault-tolerant synchronization mechanism, especially in the presence of communication failures. If a communication link fails, all sites in the overall network are divided into two disjoint groups: the major group called quorum partition and the minor group called non-quorum partition. In this case, the sites in the same partition may continue to communicate with each other, but unfortunately no communication can occur between sites in different partitions. This phenomenon is called network partitioning. In the presence of network partitioning, the sites in each partition may continue to execute transactions which access only replicas in the same partitions, but executing transactions which access replicas in the different partitions becomes impossible.

Replication control schemes (RCSs) for dealing with multiple-copy consistency in the literature can be divided into two categories: optimistic and pessimistic. Optimistic schemes (Bernstein and Goodman, 1984; Davidson, 1984) achieve high data availability by allowing more than one partition group to access replica, and any resulting inconsistencies are resolved later at the time of partition reconnection. Pessimistic schemes (Agrawal and Abadi, 1990; Cheung et al., 1990; Herlihy, 1986; Jajodia and Mutchler, 1990; Kumar, 1991), on the other hand, achieve multiple-copy consistency.