Chapter 1
A Semantics-Based Information Distribution Framework for Large Web-Based Course Forum System

Hung Chim
City University of Hong Kong, Hong Kong

Xiaotie Deng
City University of Hong Kong, Hong Kong

ABSTRACT
We propose a novel data distribution framework for developing a large Web-based course forum system. In the distributed architectural design, each forum server is fully equipped with the ability to support some course forums independently. The forum servers collaborating with each other constitute the whole forum system. Therefore, the workload of the course forums can be shared by a group of the servers. With the secure group communication protocol and fault tolerance design, the new distribution framework provides a robust and scalable distributed architecture for the large course forum system. The forum servers can be settled in anywhere as long as a broadband network connection to Internet is provided. Our experimental performance testing results show that the large forum system is a high performance distributed system with very low communication overhead cost. In addition, all course forums are classified by their teaching content relevance. Relevant course forums can be arranged on the same forum server together. Hence our distribution framework also provides a knowledge-based taxonomic storage solution to build a large digital course teaching material library.
INTRODUCTION

Rapid advance of Web technology has changed not only the initially proposed role of the Web as the medium of information communication but also human life in various ways. Web learning has become one of the hot research topics in recent years. Many Web-based education application systems have been introduced and affected the traditional teaching-learning concepts, models and methods. Without the limitations of time zones and geographic locations, these systems provide synchronous or asynchronous interactive learning environment for the teachers and students as well as among the students themselves. We started working on online Web-based Bulletin Board System (BBS) forums in 2003, and have developed a Web-based BBS forum system named Teaching Assistant System (TAS) (Hung Chim, 2004; Hung Chim, 2005). Currently, we are planning to extend the BBS forum system to a large course forum system with the capacity to support the tutorial of all teaching courses in our university. Having reviewed our original TAS system design, we devise an innovative information distribution framework to build a large Web-based course forum system as presented in this article.

Nowadays, almost all Web-based BBS forum systems use quite similar conventional client-server database design shown in Figure 1(a). This kind of design produces a tight system architecture. The biggest benefit from this architecture is the lower maintenance cost. However, this tight architecture apparently has its limitation as all forum servers must be allocated in a protected local network. Consequently, the performance of a forum will be unavoidably affected by the other forums that are sharing the same hardware or network bandwidth.

Our approach provides a solution to overcome the limitation and build up a high-performance, large-course forum system which can work over the Internet. The large forum system consists of several forum servers with the same system architecture. Each forum server (also called a node) is a fully equipped Web-based forum system (similar to Figure 1 (a)) which works independently to support the forums on it. Additionally, a new module called Node Communication Module is developed to provide the communications for data exchange and synchronization among the nodes. Therefore, all nodes collaborating with each other construct a large forum system to hold up all course discussion forums. Certainly, a particular node has to be assigned as a coordinator (called main node) to manage the collaborative communication among the nodes.

We believe that fault tolerance capability is a crucial issue for the distributed forum system. As a mature system design technique which we are using in the Forum Processing Module development, the conventional client-server database design is widely used in Web service applications. Thus we assume that each node in our forum system has sufficient stability in handling all local forum operations and works against the security attacks. On the other hand, nobody can guarantee that the network between two nodes will never be broken or jammed if and when the two nodes are located in two different cities. How to guarantee that each node can provide adequate forum services even if it temporarily loses network connections to other nodes is the major concern in our approach. We solve the problem with two methods. First, we apply a partial data replication in the database model design, the essential data for maintaining the local forum services are replicated in each node. Second, secure group communication protocols are developed to keep the consistency of the replicated data on all nodes. Therefore, our approach provides a robust and high scalable distribution framework to meet the demand and nature characteristic of Web distance education.

Communication is a key issue in distributed system, since efficiency can only be achieved when the communication overhead is small. Based on the results of investigating the ordinary operations of registered users and corresponding