ERCIM "Alain Bensoussan" Fellowship Scientific Report

Fellow: Balakrishna Prabhu
Visited Location : VTT, Espoo, Finland
Duration of Visit: From November 01, 2006 until July 31, 2006.

I - Scientific activity

(1 page at maximum)

Traffic measurements in the Internet suggest that file sharing and distribution over Peer-to-Peer (P2P) networks is becoming increasing popular among Internet users. The most attractive feature of a P2P network is its scalability which results from the active participation of each node in distributing the file. In practice, however, there is usually some "free-riding", i.e., some users leave the system as soon as they download the file, and this can degrade the system performance. Our research activity involved the design and performance evaluation of distributed file sharing applications based on random encounters and free-riding users. File sharing applications like BitTorrent divide a large file into smaller pieces so as to make the file distribution more efficient. Hence, the goal of our research is to broadcast a set of C messages to a set of N nodes in minimal time. The nodes contact each other at random and can exchange one message in one time unit. A simple message exchange policy is to select a message at random. This policy requires a small overhead. In the first instance we studied the dynamics of the number of copies of each message in the system. For a file to be efficiently distributed, it is important for the copies of each message to grow at similar rates. Using an analogy with Polya urn model, and by simulations, we showed why random message selection policy may not be efficient in distributing a file. Hence, other message selection policies such as rarest-first (already used in BitTorrent) are more suitable for distribution of files in the presence of free-riding user . However, the rarest-first policy requires nodes to maintain the current state information on the number of copies of each message in the system, and this incurs a high overhead in terms of signaling messages. We proposed a new message selection policy, called the "deterministic last K messages" policy which has the same overhead as the random message selection policy but has a better performance in the presence of free-riding users.

II- Publication(s) during your fellowship

Please insert the title(s), author(s) and abstract(s) of the published paper(s). You may also mention the paper(s) which were prepared during your fellowship period and are under reviewing.

Title : Flash crowd in a file sharing system based on random encounters.
Authors : Ilkka Norros,Balakrishna Prabhu, and Hannu Reittu
Abstract : BitTorrent revolutionized the technique of distributing a very large file to a very large number of recipients. The file is chopped into small chunks that the recipients can immediately upload further. In the original design, a ``tracker'' keeps certain centralized control over the chunk transfer process. This paper studies a BitTorrent-like ``information diffusion'' system that has a fully distributed and symmetric architecture. The peers join a Distributed Hash Table -based overlay network and contact each other randomly. This kind of designs have been implemented and analysed recently. A trackerless BitTorrent system has been introduced which can be regarded as one based on random encounters --- the participating nodes contact each other at random and download missing chunks. On the analytical front, Massoulie and Vojnovic showed that a random encounter based system has surprisingly good performance without any chunk preference strategies, with the condition that each peer gets its first chunk from a sufficiently uniform distribution. In this paper, we focus on a scenario where this condition cannot be guaranteed, and show that a ``rare chunk phenomenon'' easily occurs, if both the encounters and the chunk selection are random. Classic urn models give some mathematical understanding of this phenomenon. We then discuss various techniques for alleviating the rare chunk problem and propose a simple distributed chunk selection policy that reduces the imbalance in the distribution of chunks within the network.

III -Attended Seminars, Workshops, and Conferences

Please identify the name(s), date(s) and place(s) of the events in which you participated during your fellowship period.