Chapter 12
Verification of Super-Peer Model for Query Processing in Peer-to-Peer Networks

J. Pourqasem
University of Guilan, Iran

S.A. Edalatpanah
Ayandegan Institute of Higher Education, Iran

ABSTRACT

Equal peers in peer-to-peer (P2P) networks are the drawbacks of system in term of bandwidth, scalability and efficiency. The super-peer model is based on heterogeneity and different characteristics of peers in P2P networks. The P2P networks and large- scale distributed systems based on P2P networks use the super-peer model to design the query processing mechanism. This chapter first reviews the query processing methods in P2P networks, in which the authors classify theses query processing approaches in Unstructured and Structured mechanisms. Furthermore, the query processing techniques in distributed systems based on P2P networks are discussed. Afterward, authors concentrate on super-peer model to process the query of peers in P2P networks. Authors present the query processing methods in P2P-based distributed systems using the super node. Finally, the chapter provides some examples of each of the presented query processing techniques, and then illustrates the properties of each of them in terms of scalability and performance issues.

INTRODUCTION

P2P networks are distributed systems in which resources are shared by direct exchange between autonomous nodes. The shared resources contain documents, storage capacity, bandwidth, and CPU cycles. Each peer links to a small subset of other peers, so logical overlay networks are formed on top of the physical one (usually the Internet). Existing P2P systems have already had advantages of scalability, load balancing, self-organization, adaptation and fault tolerance (Xiao, Zhuang, & Liu, 2005; Hawa, As-Sayid-Ahmad, & Khalaf, 2013; Torkestani, 2012). In pure P2P systems, all peers play equal roles and

DOI: 10.4018/978-1-5225-0287-6.ch012

take same responsibilities regardless of their capabilities. Moreover, in query flooding process, any peer could be a query sender and the query receiver (Navimpour & Milani, 2015). However, as the size of P2P network grows, weak peers will seriously limit the scalability of P2P systems and become bottlenecks.

The peers in P2P networks usually have different characteristics with respect to their capabilities, e.g. available bandwidth, storage space or processing power (Nejdl et al., 2003; Cholvi, Felber, & Biersack, 2004). A discussion about the different capabilities of peers lead to an efficient network topology in P2P networks (Yang & Garcia-Molina, 2002). There is a small subset of peers with high capacity, called super-peers that takes over specific responsibilities for peer aggregation, query routing, query processing and possibly mediation (Cao, Li, & Liu, 2008). In this topology, the symmetry of pure P2P systems is broken by assigning additional responsibilities to super nodes (Garbacki, Epema, & van Steen, 2007). Consequently, super-peers make the decentralized networks more efficient by exploiting heterogeneity of peers and distributing load to machines that can handle the burden.

The super-peer network comprises the two-layer architecture (Awan, Ferreira, Jagannathan, & Grama, 2005). In the low layer, the super-peer is a node that acts as a centralized server to a subset of connected clients. When the clients need resource/file, submit queries to their super-peer and receive results from it (Yang & Garcia-m, 2003; Tan, Lü, & Lin, 2012). In the higher-layer, the super-peers are connected to each other and form the overlay network. The super-peers route messages over this overlay network, and submit and answer queries on behalf of their clients (Mastroianni, Talia, & Verta, 2005b). The super-peers are equal in terms of search, and all the client peers are equal in terms of download. Consequently, the client peers with low capacity are shielded from massive query traffic, which improves the scalability of the system. The peers use resources of network more efficiently because the super-peers provide the efficient and reliable query processing function.

The authors of (Garbacki et al., 2010; Keller & Martin-Flatin, 2006; Loser, Naumann, Siberski, Nejdl, & Thaden, 2004; Yu & Li, 2008; Yang & Garcia-m, 2003; Jelasity, Montresor, Babaoglu, 2009; Xiao et al., 2005) have proposed the super-peer networks. The authors of (Kirk, 2015; Yang & Garcia-m, 2003) addressed the principles and guidance of designing a super-peer overlay network. The authors of (Kaller & Martin-Flatin, 2006; Yu & Li, 2008; Jelasity et al., 2009) used network proximity for building a super-peer overlay, in which the client peers are connected to super-peers based on their distances. These studies considered decreasing the communication latency between nodes by exploring the network proximity. However, these studies have not discussed the efficiency a super-peer overlay by using the high capacity nodes. Super-peer selection has to make a trade-off between latency and efficiency features. The authors of (Garbacki et al., 2010; Loser et al., 2004; Nejdl et al., 2003) investigated the semantic similarity of peers in building process of a super-peer overlay. In selection process based on semantic similarity, the client peers that share the same interest are connected to the same super-peers. Furthermore, there are several design issues with the super-peer networks such as what should be the number of clients that a super peer serves? What should be the topology among super peers (Yang & Garcia-m, 2003)?

Every client peer in a super-peer network is connected to a fixed, very small number (usually one) of super-peers randomly and statically. The static assignment cannot adapt to the changes in the network structure or to the peer characteristics and capabilities. Moreover, single super-peer makes load balancing among the super-peers difficult and becomes a performance and scalability bottlenecks. The super-peer network suffers from super node crashing, in which restoring the system structure back to a consistent situation requires a considerable effort. Moreover, the super-peer systems are more vulnerable to attacks, as there is usually a single super-peer in the network (Yang & Garcia-m, 2003; Mahdy, Deogun & Wang,