P2P-DIET: One-Time and Continuous Queries in Super-Peer Networks

Stratos Idreos, Manolis Koubarakis, and Christos Tryfonopoulos

Dept. of Electronic and Computer Engineering
Technical University of Crete, GR73100 Chania, Greece
{sidraios,manolis,trifon}@intelligence.tuc.gr

1 Introduction

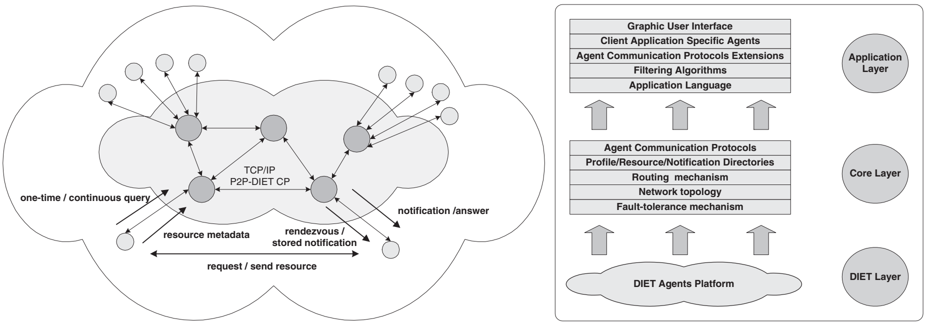
In peer-to-peer (P2P) systems a very large number of autonomous computing nodes (the *peers*) pool together their resources and rely on each other for *data* and *services*. P2P systems are application level *virtual* or *overlay networks* that have emerged as a natural way to share data and resources. The main application scenario considered in recent P2P data sharing systems is that of *one-time querying*: a user poses a query (e.g., "I want music by Moby") and the system returns a list of pointers to matching files owned by various peers in the network. Then, the user can go ahead and download files of interest. The complementary scenario of *selective dissemination of information (SDI)* or *selective information push* is also very interesting. In an SDI scenario, a user posts a *continuous query* to the system to receive notifications whenever certain *resources* of interest appear in the system (e.g., when a song of Moby becomes available). SDI can be as useful as one-time querying in many target applications of P2P networks ranging from file sharing, to more advanced applications such as alert systems for digital libraries, e-commerce networks etc.

At the Intelligent Systems Laboratory of the Technical University of Crete, we have recently concentrated on the problem of SDI in P2P networks in the context of project DIET (`http://www.dfki.de/diet`). Our work, summarized in [3], has culminated in the implementation of P2P-DIET, a service that unifies one-time and continuous query processing in P2P networks with super-peers. P2P-DIET is a direct descendant of DIAS, a Distributed Information Alert System for digital libraries, that was presented in [4]. P2P-DIET combines one-time querying as found in other super-peer networks and SDI as proposed in DIAS. P2P-DIET has been implemented on top of the open source DIET Agents Platform (`http://diet-agents.sourceforge.net/`) and it is available at `http://www.intelligence.tuc.gr/p2pdiet`.

2 The System P2P-DIET

A high-level view of the P2P-DIET architecture is shown in Figure 1(a) and a layered view in Figure 1(b). There are two kinds of nodes: *super-peers* and *clients*. All super-peers are equal and have the same responsibilities, thus the

E. Bertino et al. (Eds.): EDBT 2004, LNCS 2992, pp. 851–853, 2004.



(a) (b)

Fig. 1. The architecture and the layered view of P2P-DIET

super-peer subnetwork is a *pure* P2P network. Each super-peer serves a fraction of the clients and keeps *indices* on the resources of those clients.

Clients can run on user computers. Resources (e.g., files in a file-sharing application) are kept at client nodes, although it is possible in special cases to store resources at super-peer nodes. Clients are equal to each other only in terms of download. Clients download resources directly from the resource owner client. A client is connected to the network through a single super-peer node, which is the *access point* of the client. It is not necessary for a client to be connected to the same access point continuously since *client migration* is supported in P2P-DIET. Clients can connect, disconnect or even leave from the system silently at any time. To enable a higher degree of decentralization and dynamicity, we also allow clients to use *dynamic IP addresses*. *Routing* of queries (one-time or continuous) is implemented using *minimum weight spanning trees* for the super-peer subnetwork. After connecting to the network, a client may *publish* resources by sending resource metadata to its access point, *post an one-time query* to discover matching resources or *subscribe* with a continuous query to be notified when resources of interest are published in the future. A user may *download* a file at the time that he receives a notification, or save it in his *saved notifications folder* for future use. Additionally a client can download a resource even when he has *migrated* to another access point. The feature of *stored notifications* guarantees that notifications matching disconnected users will be delivered to them upon connection. If a resource owner is disconnected, the interested client may arrange a *rendezvous* with the resource. P2P-DIET also offers the ability to add or remove super-peers. Additionally, it supports a simple fault-tolerance protocol based on *are-you-alive* messages. Finally, P2P-DIET provides message authentication and message encryption. For the detailed protocols see [5].

The current implementation of P2P-DIET to be demonstrated supports the model $\mathcal{AWP}$ [4] and it is currently been extended to support $\mathcal{AWPS}$ [4]. Each super-peer utilises efficient query processing algorithms based on indexing of resource metadata and queries and a hierarchical organisation of queries (poset) that captures query subsumption as in [1]. A sophisticated index that exploits commonalities between continuous queries is maintained at each super-peer, enabling the quick identification of the continuous queries that match incoming resource metadata. In this area, our work extends and improves the indexing algorithms of SIFT [6] and it is reported in [2].

References

1. A. Carzaniga and D.S. Rosenblum and A.L. Wolf. Design and evaluation of a wide-area event notification service. *ACM Transactions on Computer Systems*, 19(3):332–383, August 2001.
2. C. Tryfonopoulos and M. Koubarakis. Selective Dissemination of Information in P2P Networks: Data Models, Query Languages, Algorithms and Computational Complexity. Technical Report TUC-ISL-02-2003, Intelligent Systems Laboratory, Dept. of Electronic and Computer Engineering, Technical University of Crete, July 2003.
3. M. Koubarakis and C. Tryfonopoulos and S. Idreos and Y. Drougas. Selective Information Dissemination in P2P Networks: Problems and Solutions. *ACM SIGMOD Record, Special issue on Peer-to-Peer Data Management, K. Aberer (editor)*, 32(3), September 2003.
4. M. Koubarakis and T. Koutris and C. Tryfonopoulos and P. Raftopoulou . Information Alert in Distributed Digital Libraries: The Models, Languages and Architecture of DIAS. In *Proceedings of the 6th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2002)*, volume 2458 of *Lecture Notes in Computer Science*, pages 527–542, September 2002.
5. S. Idreos and M. Koubarakis. P2P-DIET: A Query and Notification Service Based on Mobile Agents for Rapid Implementation of P2P Applications. Technical Report TUC-ISL-01-2003, Intelligent Systems Laboratory, Dept. of Electronic and Computer Engineering, Technical University of Crete, June 2003.
6. T.W. Yan and H. Garcia-Molina. Index structures for selective dissemination of information under the boolean model. *ACM Transactions on Database Systems*, 19(2):332–364, 1994.