UNIVERSITY
OF TRENTO

DEPARTMENT OF INFORMATION AND COMMUNICATION TECHNOLOGY

38050 Povo – Trento (Italy), Via Sommarive 14
http://www.dit.unitn.it

QUERY ANSWERING IN PEER-TO-PEER DATABASE NETWORKS

Ilya Zaihrayeu

March 2003

Technical Report # DIT-03-012

Query Answering in Peer-to-Peer Database Networks

Ilya Zaihrayeu
PhD thesis proposal

Department of Information and Communication Technology
University of Trento

March 2003

Abstract

Peer-to-Peer (P2P) received significant attention from both industry and academia as a version of distributed computing lying between traditional distributed systems and the web. P2P found lots of its application in file sharing systems, distributed computations, instant messaging systems and so on. Database community investigates P2P as a new decentralized paradigm for distributed management of data, that gives sound advantages over existing database systems, and also raises some challenges for thorough research in this area. The main focuses of the current PhD thesis, supervised by Prof. Fausto Giunchiglia, are the study of the optimal balance of centralization vs. decentralization levels in a P2P database network necessary for efficient query answering; development of logical and physical architecture to support P2P databases; development of a query answering algorithm and a formal theory that would describe P2P databases as well as query answering and its quality; development of a prototype and conducting of the performance study in different initial settings.

1 Introduction

Peer-to-Peer computing is an open ended network of distributed computational nodes or peers which can exchange data and/or services with other nodes, called *acquaintances*[1]. In a P2P network nodes have equivalent functionality for providing and consuming data/services; moreover, in pure P2P settings, nodes are largely independent from each other, and there is no global control in the form of global registry, or global resources management, nor a global schema or data repository.

Actually the peer-to-peer ideas are not new, the P2P has always existed, but it has not been recognized as such up to the latest time. Servers with fixed or resolvable addresses always had a possibility to exchange services with other servers; email and the domain name system were built on the P2P principles. It also worth mentioning the Usenet application, which was created in 1979 and provided a way for two computers to exchange information. One computer could dial another, check for new files and download those files. Finally the system evolved into a massive newsgroup system as it is today.

A lot of attention and its name, P2P received after ICQ [2] (1996), SETI@Home [6] (1996), Groove [1] (1997) and then Napster [4] (1999) popularized this paradigm. ICQ is an application that allows users to be notified when their friends come online, and to send instant messages to them. SETI@Home is a distributed computing project that allows splitting the task of large volumes of radio-telescope data processing into subtasks and sending them to peers for remote computation, thus allowing them to contribute to the search of extraterrestrial life. Groove allows its users to create a secure shared space

[1] Adopted from [7]

for collaborative activity (shared file editing, conferences, etc). Napster is a file sharing application that allows its users to locate music files on other nodes and download files directly from those locations.

The examples given above demonstrate different domains for P2P applications. Up to now, the most popular P2P applications can be classified according to the four major categories: instant messaging, file or data sharing, remote collaboration and distributed computing.

Some of the main advantages of the P2P computing are said to be the robustness of a P2P network, when failure of one or several nodes does not lead to the overall network malfunction; scalability, i.e. new nodes can join and leave the network in runtime without a need of the whole system restructuring. Another advantage of a P2P network, which mainly is exploited in searches, is the ability of peers to use their acquaintance links to transitively propagate data (queries or results, for instance) consequently through a number of nodes.

The ad hoc nature of a P2P network, high level of nodes' autonomy and the absence of global centralization, that constitute the grounds for the advantages of P2P, also lead to some drawbacks. The main disadvantage relates the problem of *resources discovery* (both peers and their resources) in a P2P network. As the bottom line, a P2P network has no any centralization or coordination mechanisms in the resources discovery request propagation. In such a setting the request is propagated on the 'send-to-all-acquaintances' basis, thus the propagation takes the avalanche-like form, involving more and more peers with each consequent propagation. This certainly in most cases leads to a big portion of irrelevant results and may lead to network overload. There is no guarantee of the *Quality of the Service* (QoS) provided by the P2P network.

Another problem is the *meta-data* problem, namely its scope, maintenance, its management and obsoleteness. In order to make the use of the network's resources efficient, a lot of meta-data needs to be produced and maintained. Due to the strong dynamics of P2P network this is a crucial and hard task to perform. The higher the level of meta-data management in the network the higher QoS, and the harder to maintain it. In contrast, the more ill-defined the meta-data, the easier to maintain it, but the QoS can suffer greatly. There certainly should be an optimal tradeoff between the efforts put in meta-data management and the justified efficiency of the network.

Some basic techniques exist to address these problems. One is adding centralization points into the P2P network topology infrastructure. Such a P2P network is then called *hybrid*. For example, meta-data on peers and their resources can be kept and managed at dedicated servers. Peers send requests to these servers to locate other peers. After desirable resource is located, peers establish connections in the P2P fashion by-passing the server.

Another approach which is actually a dimension in meta-data management is the support for peers of being able to (self-) organize into *peer groups* according to a specific subject of their common interest (domain of the data they store, for instance). Peers are free to choose what groups they want to participate in. This increases the mutual relevance of the data stored at peers, and therefore decreases the irrelevant results ratio.

Nowadays hundreds of millions of people around the world are participating in different P2P networks, using dozens of P2P application of different functionality sets and purposes. Nevertheless, there is still space for innovative solutions in the P2P domain. One stands at the interfaces between the P2P paradigms and the database technologies – which is the subject for the current thesis work.

The rest of the proposal is structured as follows: section 2 expands more on the perspectives of P2P DBs and peculiarities of query propagation; section 3 describes the latest scientific achievements in the correlated areas. In section 4 the main objectives of the thesis work are defined; section 5 concludes with the overview of work that has been done so far.

2 P2P and databases

Consider the evolution line of the database systems: original *central databases* were defined and administrated centrally (as it follows from their name). With the appearance of *distributed databases* the level of centralization decreased: administration of the databases can be handled in a physically distributed manner over a computer network, although it was assumed that databases were defined locally. Still, the data distribution over the network was largely transparent for the end users. Traditional distributed database system has a global schema and the same DBMS for all the involved databases. If an element of local autonomy is introduced then such a system is usually called a *federated database system* or a federation. But still, a federation has the notion of globally defined schema. With the presence of heterogeneity between the schemas of involved databases, the system becomes a *multi-database system*.

It is easy to see that the tendency is to go away from the centrally defined and administrated databases to a distributed databases system, where each database is governed locally, but which still can be viewed as one monolithic database represented by means of a global schema. The latter makes the work with the databases system transparent for the end user in respect to the underlying complexity of databases distribution and inter-schema dependencies.

P2P paradigms appear to be very coherent with this databases evolution line. Peer-to-peer database (P2P DB) networks can be a well-founded choice for the next generation of database systems. Databases can reside and be managed locally at peers and can be given large level of autonomy in respect to the databases sitting at other peers. Autonomy in this context means that peers decide on their own how to develop their databases, what DBMS to use, how to store data, etc.

But on the other hand, in order to be consistent with data integration ideas, substantially exploited in the modern database systems, database peers should have well-developed mechanisms to coordinate their databases when answering queries or processing updates. At most, it concerns query evaluation and propagation, since query answering is the fundamental functionality required from a database system.

Ideally, work of the end user on a P2P DB network should not be much different from the one on a multi-database system, for instance. That means that the number of peers, how data is allocated on those peers and semantic interdependencies between them should be irrelevant to the user. Although three fundamental differences can be distinguished in query answering: since P2P network evolves over time, the same query can bring different results if posed at different moments of time; the query response time can be longer and results can stream in continuously as further nodes are reached; and the results can be not as relevant due to the fact that the quality of links between the peers may vary greatly.

2.1 Query propagation

Query propagation in a P2P DB network is the most crucial and challenging task to solve. In centralized database systems as well as in the data integration systems there is a notion of *correctness* and *completeness* of query results. Intuitively it means that results are correct in respect to the inter-database schema mappings and complete in respect to the databases storing relevant to the query data. That notions comes from the fact of the existence of a global schema that allows to view all involved databases as one single database. In the peer-to-peer settings there is no notion of global schema, and mainly all peers know about only of a (small) subset of other peers present on the network. Therefore the notions of correctness and completeness cannot be used in their pure meaning, but can only be approximated for the case of P2P databases.

In most cases user will not get a correct and complete answer, but at least will get some answer which can be recognized as *complete enough* in respect to the set of peers, storing relevant to the query data, which can be reached via acquaintances links; and as *correct enough* in respect to the semantic mappings between peers involved in query answering. In aggregate, a notion of *good enough answer* can be exploited in P2P DB networks, as proposed in [10].

In order to maintain high quality answers from the network, significant efforts towards partial centralization and meta-data management should be done, as it is highlighted in the introduction. The same solutions are also applicable for P2P DB networks. In addition, yet another way to put a certain control on query propagation in database P2P networks is intensively discussed. The main idea is to specify local to a peer set of coordination rules which state what kind of query, when and where should be propagated given the set of available acquaintances and meta-data that captures semantic interdependencies between schemas of databases, for instance. The goodness of this approach in respect to P2P is that it preserves the decentralized nature of peer-to-peer, shifting the accent from data integration to local data coordination. Hereinafter this approach is referenced as *local peer coordination*.

Thus query propagation strategies and consequently the quality of query answering depend on some main characteristics, which are listed bellow:

- The level of centralization (in the case of hybrid topology architecture) and how the centralization points are put together. The latter can mean that centralization points are loosely coupled and inter-operate only occasionally; they tightly coupled or have hierarchical schema of interrelationships.
- The level of meta-data development and management (including peer groups management, resources description, etc);
- The level of local peer coordination;

3 State of the art

3.1 Local peer coordination

There are different approaches in defining local to a peer rules for query propagation. *Local Relational Model* [19] defines a formal way, based on the first-order logic, to describe the set of databases in a P2P network, and the set of *coordination formulas* which are soft inter-databases constraints which in most general case state that if a certain data is present in database i then corresponding data should also be present in database j. The problem of heterogeneity is solved with the help of *domain relations* – a function that maps elements of the domain of database i into corresponding elements of the domain of database j and vice versa.

Data integration techniques were also proposed as formalisms for expressing semantic mappings between peers. There are two major approaches in this area. The first operates with notions of *Local-as-View* (LAV) and *Global-as-View* (GAV) [16]. Both assume the existence of a global schema, of a set of local schemas, and of the mapping between the two. If the mapping defines the sources in terms of the global schema then this approach is LAV. If the global schema is defined in terms of the sources then this is GAV. GAV suffers from the necessity of changing the global schema (in design time) if there is a change in a local schema. In its turn LAV leads to the problem of complex query rewriting techniques implementation.

In the second approach the contents of the sources are described as views over the mediated (global) schema. As a result, the problem of reformulating a user query, posed over the mediated schema, into a query that refers directly to the source schemas becomes the problem of answering queries using views. This approach was thoroughly studied in the literature on data integration, for instance in [12, 17, 20].

The crucial drawback of these two approaches is that they rely on the notion of a global schema, whereas in pure P2P settings the assumption of a global schema existence can not be accepted.

Nevertheless data integration formalisms and related query rewriting techniques can be reused to express interdependencies between two peers or between a peer and a (small) subset of peers as well as for query translation between them. This is the underlying idea proposed in [13]. In this work, a schema mediation approach is exploited as a way to facilitate transitive query propagation and answering within a P2P network. In particular the authors describe a formalism to mediate not only between peer schema and data, but also between peer schemas. GAV and LAV are the basic formalisms used to specify local mappings. The paper also describes a query reformulation algorithm that takes as input a peer's query and the formulas describing semantic relationships between peers, and returns a query that relates only the data tuples stored at relevant peers.

An orthogonal approach to coordinate query and update propagation is studied with respect to the active database technology [9]. Active database technology allows to specify *Event-Condition-Action* (ECA) rules that can be used as local to peer mechanism to handle query and update processing, including propagation to a P2P network. The main idea is that a triggering event can be a query or an update to the underlying database; a condition part checks a certain property a given query or an update may hold; and, if the condition holds then in the action part that query or update is propagated to the condition-dependent subset of peers on the network. In particular, this approach is studied in [10, 14].

When making the decision on what approach to take as the underlying mechanism for local peer coordination, it is important to evaluate their relative expressive power. In other words it is important to know whether all desirable semantic interdependencies can be captured by a particular approach. So far, there is no literature on formal comparison of the expressive power of the above mentioned methods, although some of them are separately studied, for example in [8].

3.2 P2P DB projects

A number of research groups are working in the field of P2P databases. Peer-to-peer database management system, called *Piazza*, is being developed at the University of Washington [5]. Piazza allows peers to define semantic mappings between pairs of peers (or among small subsets of peers) by means of mediated schemas that can relate peer and database schemas as described in [13]. The group is also working on the data placement in a P2P network problem [11] and has a preliminary version of the prototype.

University of Toronto develops *The Hyperion Project* which aims mostly on the precise definition of a peer-to-peer data management architecture and the study of viable data integration, exchange and mapping mechanisms for the P2P environment, as reported in [15].

Project called *PeerDB* is being developed at the University of Singapour. PeerDB is a P2P-based system for distributed data sharing. PeerDB is mostly focused on the architecture development that would support P2P database management, and query processing, mainly assisted by agents. PeerDB has already running prototype and report on performance study as described in [18].

The University of Trento evolves the *Context-to-Context InfoNet* project. The main aim is the development of methodologies, theories, mechanisms, and technologies which enable nodes viewed as information sources, in particular as databases, to efficiently interact in a P2P network. Some of the achieved results describe proposed architecture, query propagation algorithm and a vision on the use of ECA rules as the basic underlying mechanism for local peers coordination as reported in [10].

4 Objectives of the thesis work

The *primary objective* of the thesis work is the development of an efficient query answering algorithm for a P2P DB network. The notion of efficiency in particular includes the good enough answers ratio, reasonable response time, optimal propagation scope and moderate network load.

As it follows from the subsection 2.1, this is not a trivial task to perform, and it involves a number of different issues to be solved. In particular, the future thesis work includes:

- Study of the optimal level of centralization mechanisms incorporated into the network and the modality of their inter-operation;
- Development of the meta-data of a level sufficient to effectively support the P2P network infrastructure, as well as the development of the meta-data management mechanisms. A special attention will be given to the interest groups formation and maintenance, as well as on the resources representation and discovery mechanisms;
- Development of the underlying logical and physical architecture that would support developed meta-data and centralization mechanisms;
- Development of a query propagation algorithm;

- Development of the formal underlying theoretical model that would be able to describe involved databases, local peer coordination, the level of correctness and completeness of query answering in a P2P database network, etc.

- Development of a complete prototype of a P2P DB application based on the developed architecture. The application would implement the query propagation algorithm, enable peers to evaluate queries against their local databases, receive and reconcile query results, locate other peers, learn about peer's resources, etc.

- Performance study for the query answering algorithm. In particular, the following characteristics will be examined: query answering time, number of reached peers as a function of the elapsed time, volumes of data returned, ratio of good enough answers, the overall network load, and some others. The experiments will be conducted in several different initial settings including number of participating nodes, volumes of data stored at nodes, quality of the internet connections etc.

The performance study will serve as the feedback which will be used for the propagation algorithm's parameters tuning and probably for meta-data management mechanisms modification.

The local peer coordination will not be closely studied due to it is a rather big area of research itself. Presumably, in the implementation some light form of local peer coordination will be used to boost the query answering efficiency.

5 What has been done so far

As the joint work with Prof. Giunchiglia, the first level architecture that supports peer databases was developed, which was first described in [7]. This architecture implements four proposed basic notions of a P2P DB network: *Interest Groups*, *Acquaintances*, *Coordination Rules*, and *Correspondence Rules*. Interest Groups are supposed to gather peers which are able to answer about a certain topic. The peer groups are proposed to be centrally managed by means of a *Group Manager*, a special node that keeps meta-data about all other nodes in the group. Group Manager is used to compute the *Query Scope* for any given query – the set of nodes the query should be propagated to. Acquaintances are nodes a node knows about and which are able to answer a specific query. In other words, a node is acquainted with another node only with respect to a certain type of queries. Coordination Rules are used to specify under what condition, when, how and where to propagate a given query. These rules exploit Correspondence Rules which tell how to translate queries and query results when propagating from one node to another, or said differently they solve the semantic heterogeneity problem. A query propagation algorithm was also developed that uses the above mentioned notions as primitives. Finally all these intuitions elaborated into a published paper [10].

The first version of the P2P DB application has been developed. Currently, peers are able to locate other peers, build communication links, send queries to and receive results from other nodes. Accordingly peers can receive queries from the network, evaluate queries over the local database, send the results back, and propagate the query further to their acquaintances.

The architectural choice for the underlying platform was made to the favor of the *JXTA project* [3]. JXTA defines a set of protocols, designated mainly for transport and addressing space support in P2P networks. It also provides mechanisms for peers and other basic resources discovery, gives well-developed tools for meta-data representation, communication links establishment, and so on. In other words, JXTA gives an instrument suite for P2P applications development of arbitrary nature. On top of JXTA one can implement his own procedures defining query propagation strategies, meta-data management mechanisms, etc.

Currently the centralization schema is adopted from JXTA and it is rather lightweight. Peers called rendezvous cache meta-data from other peers and serve as reference points for resources discovery. Although they do not take care about specific queries propagation strategies, instead the application does.

Currently queries can be propagated to only acquaintances chosen by the user. JXTA also supports interest groups formation, and, in the present version of the prototype, all peers belong to one fundamental peer group. Meta-data representation primitives are also adopted from JXTA and currently they can capture basic peer description and provide information required for communication links establishment.

The prototype was written using the JAVA language, which is platform-independent. One of the JXTA's advantages is that it is also platform-independent. Thus the P2P DB application can run on different system platforms (such as the Microsoft Windows or Unix operating systems) and network platforms (such as TCP/IP or Bluetooth).

References

[1] Groove networks, see http://www.groove.net.

[2] Icq, see www.icq.com.

[3] Jxta project, see http://www.jxta.org.

[4] Napster, see www.napster.com.

[5] Piazza peer data management system (pdms), see http://data.cs.washington.edu/p2p/piazza.

[6] Seti@home, see http://setiathome.ssl.berkeley.edu.

[7] P. Bernstein, F. Giunchiglia, A. Kementsietsidis, J. Mylopoulos, L. Serafini, and I. Zaihrayeu. Data management for peer-to-peer computing: A vision. *WebDB*, 2002.

[8] A. Cal, D. Calvanese, G. Giacomo, and M. Lenzerini. On the expressive power of data integration systems. *Proceedings of the International Conference on Conceptual Modeling (ER)*, pages 338–350, 2002.

[9] U. Dayal, E. Hanson, and J. Widom. Active database systems. *Modern Database Systems*, pages 338–350, 1994.

[10] F. Giunchiglia and I. Zaihrayeu. Making peer databases interact - a vision for an architecture supporting data coordination. *6th International Workshop on Cooperative Information Agents (CIA-2002), Universidad Rey Juan Carlos, Madrid, Spain, September 18 -20*, 2002.

[11] S. Gribble, A. Halevy, Z. Ives, M. Rodrig, and D. Suciu. What can databases do for peer-to-peer? *WebDB, Workshop on Databases and the Web*, June 2001.

[12] A. Halevy. Answering queries using views: a survey. *VLDB Journal*, 2001.

[13] A. Halevy, Z. Ives, D. Suciu, and I. Tatarinov. Schema mediation in a peer data management system. *To appear, International Conference on Data Engineering*, 2003.

[14] V. Kantere. A rule mechanism for peer-to-peer data management. *M.Sc. thesis, University of Toronto, Canada*, 2002.

[15] A. Kementsietsidis, M. Arenas, and R. Miller. Data mapping in peer-to-peer systems. *Technical Report CSRG-456, University of Toronto*, July 2002.

[16] M. Lenzerini. Tutorial on "data integration: A theoretical perspective". *21st ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS, Madison, WI, USA*, June 2002.

[17] A. Levy, A. Mendelzon, Y. Sagiv, and D. Srivastava. Answering queries using views. *Proceedings of the 14th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, 1995.

[18] W. Ng, B. Ooi, K. Tan, and A. Zhou. Peerdb: A p2p-based system for distributed data sharing. *International Conference on Data Engineering (ICDE)*, 2003.

[19] L. Serafini, F. Giunchiglia, J. Mylopoulos, and P. Bernstein. The local relational model: Model and proof theory. *IRST Technical Report 0112-23, Istituto Trentino di Cultura*, December 2001.

[20] J. Ullman. Information integration using logical views. *Theoretical Computer Science*, 1997.