SARAVÁ: data sharing for online communities in P2P

Marta Mattoso[1], Esther Pacitti[2], Patrick Valduriez[2], Reza Akbarinia[2], Vanessa Braganholo[3], Alexandre A. B. Lima[4]

[1]COPPE, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
[2]INRIA and LINA, Nantes, France
[3]Computer Science Department, Federal University of Rio de Janeiro, Brazil
[4]School of Science and Technology, Unigranrio University, Rio de Janeiro, Brazil

marta@cos.ufrj.br, Esther.Pacitti@univ-nantes.fr, Patrick.Valduriez@inria.fr, Reza.Akbarinia@inria.fr, braganholo@dcc.ufrj.br, abento@unigranrio.com.br

Abstract. *This paper describes SARAVÁ, a research project that aims at investigating new challenges in P2P data sharing for online communities. The major advantage of P2P is a completely decentralized approach to data sharing which does not require centralized administration. Users may be in high numbers and interested in different kinds of collaboration and sharing their knowledge, ideas, experiences, etc. Data sources can be in high numbers, fairly autonomous, i.e. locally owned and controlled, and highly heterogeneous with different semantics and structures. Our project deals with new, decentralized data management techniques that scale up while addressing the autonomy, dynamic behavior and heterogeneity of both users and data sources. In this context, we focus on two major problems: query processing with uncertain data and management of scientific workflows.*

1. Introduction

Online communities such as social networks (e.g. sites like MySpace and Facebook) and professional communities (e.g. scientific communities, online technical support groups) are becoming a major killer application of the web. Community members typically have common interests or purposes and are willing to share their data. Through data sharing, online communities foster mass collaboration whereby large numbers of members can contribute in parallel to a large project, e.g. developing a world-wide encyclopedia such as Wikipedia.

To scale up to very large amounts of data while providing performance and availability for ever growing numbers of users, they can rely on a large data center (e.g. using a cluster computer). However, this solution is expensive (both in hardware/software and electricity), requires skilled staff to administer, tune and repair the system and is hard as it requires centralized integration of heterogeneous data. Furthermore, community members may not trust the administrators of the centralized servers and prefer to maintain their data in their computer under their own control. Thus, radically new data sharing techniques where the user plays a central role are needed. A promising solution is to organize community members in a peer-to-peer (P2P) architecture where each member can share data with the others through a P2P overlay network. The major advantage of P2P is a completely decentralized approach to data sharing which does not require centralized administration.

This paper describes SARAVÁ, a research project that aims at investigating new challenges in P2P data sharing for online communities. In Portuguese, SARAVA means "good luck", that is certainly an interesting name for a project that involves collaboration among Brazilian and French teams that have been working on successful projects over the past years (Ecobase, DAAD and GriData, to cite just a few).

2. Data Sharing in P2P: research challenges

The general problem we address in our project is P2P data sharing for online communities, by offering a high-level network ring [Abiteboul and Polyzotis 2007] across distributed data source owners. Users may be in high numbers and interested in different kinds of collaboration and sharing their knowledge, ideas, experiences, etc. Data sources can be in high numbers, fairly autonomous, i.e. locally owned and controlled, and highly heterogeneous with different semantics and structures. What we need then is new, decentralized data management techniques that scale up while addressing the autonomy, dynamic behavior and heterogeneity of both users and data sources.

To illustrate the data management requirements in this context, consider two representative, rather complex examples of online community applications: collaborative scientific research (e.g. bioinformatics) and social networking systems. As many other online community applications, these two applications have common requirements (e.g. high level data access, update support, data privacy,) and differences. For instance, collaborative scientific research may be quite demanding in terms of quantity of data exchanged while social networks may involve very high numbers of participants with dynamic behavior.

The large quantity of scientific data should be processed as a data warehouse. Traditionally, decision support queries (OLAP) are based on fully integrated data sources. Recent researches investigate techniques for OLAP query processing in computing grids [Furtado 2006, Paes et al. 2008], where data integration is not so strong. The dynamicity of P2P environments makes the data sources very loosely coupled and impose new challenges for OLAP. We intend to adapt the techniques proposed for clusters [Lima et al. 2009] and grids [Kotowski et al. 2008] for such an environment.

Also, collaborative scientific research typically requires input data to be processed by workflows of heterogeneous programs while social networks require high-level query capabilities as well as key-word search. A P2P architecture for data management, e.g. APPA [Akbarinia et al. 2006], provides important advantages like decentralized control and administration, scale up to high numbers of peers and support of the dynamic behavior of peers (who may join or leave the system at will) [Valduriez and Pacitti 2005]. These advantages are important for online communities.

In addition to the traditional requirements of distributed database or data integration systems [Özsu and Valduriez 1999] such as query expressiveness, semantic data integration, data replication and caching, we must deal with some new, hard problems:

- *Query processing with uncertain data*. Some data should not be assumed to be 100% certain, precise or correct, in particular, when coming from peers with different levels of confidence. Query processing techniques designed for P2P systems, e.g. top-k queries [Akbarinia, Pacitti and Valduriez 2007] should be revisited to deal with data uncertainty at all levels. Similarly, the recent extensions of DBMS with support for data uncertainty [Agrawal et al. 2006] should be revisited for P2P.
- *Scientific workflow management*. In a P2P setting, the traditional parallel techniques to manage scientific workflows no longer apply because many workflow components are legacy programs. Therefore, it is not possible to modify their code as they have to be treated like black box components [Meyer et al. 2005, Serra da Cruz et al. 2008]. Furthermore, exploiting the massive scale

of computing power available in desktop peers for workflow management is an open problem [Pacitti, Valduriez and Mattoso 2007].

3. Main Project Tasks

We plan to pursue parallel work on the following three main tasks:

1. **P2P data management architecture**. Architectural work will be ongoing throughout the project as it is very hard to come up with the right architecture upfront, without having solved the research issues and come up with efficient techniques. Thus, in 2009, we will start by adopting the Data Ring vision [[AP07] which is well suited for community data sharing and general enough so we can adapt it to our requirements. In a Data Ring, we can model an online community as a set of *participants* and *relationships* between them. The participants are autonomous data sources (under their own control) in the network and can be heterogeneous (e.g. relational, XML, files) with very different levels of processing capabilities (e.g. ranging from a DBMS to a file system). The relationships are between any two or more participants and indicate how their data sources are related, e.g. one source is a copy of the other, or some sources share the same semantic domain or schema. Based on a selected collaborative scientific application from bio-informatics in Brazil and a social network application, we will analyze the requirements for the data model and query language for uncertain data, and for query processing and workflow management. We will reflect these requirements in application scenarios which should be useful for the two next tasks. For selecting the data model and query language with uncertainty, we will capitalize on probabilistic data models and study the support of imprecision on attribute values using fuzzy logic. In a P2P system with heterogeneous data sources, we need a model that can describe various annotated and structured data.

2. **P2P query processing with uncertain data**. To address P2P query processing with data uncertainty, we must deal with the problems of uncertain query routing (to the nodes holding relevant data), and ranking results with uncertain data. For uncertain query routing, P2P query routing algorithms cannot be used without significant revision. The reason is that a query must be routed only to the nodes which involve relevant data with certainties higher than a given certainty, i.e. that expressed in the query. Another major problem is the ranking of results since for each query, there may be huge quantities of answers, most of them uninteresting for the user. This requires support of top-k queries with new techniques to deal with uncertainty. We will start designing query optimization and query routing techniques and reflect them to task 1, in terms of a query processing service. We will also validate these techniques by implementation and experimentation over the Grid5000 and Planetlab platforms, using synthetic and real data.

3. **P2P workflow management**. One typical scenario in genomic applications is having a set of input data to be processed by a workflow. Bioinformatics programs often generate and process large datasets. Thus, one question that can be asked is how to process this genome workflow in parallel on a high number of heterogeneous peers. There are many alternatives to execute a genome workflow in parallel because programs and data can be distributed among the peers in many different

ways. Choosing the best strategy for parallel execution in a P2P system is difficult because this choice must consider: (i) the dependencies among the components; (ii) the unbalanced execution time of the peers; (iii) the different size of datasets; and (iv) the computational resources available at peers. The choice of the best execution alternative is much more difficult in a P2P system than in a parallel machine or a cluster due to the heterogeneity and the dynamic nature of the peers. Decisions to replicate code and datasets have to be taken either on demand or by pre-staging in order to provide better performance. We have started designing workflow management techniques to exploit parallel execution to reflect them to task 1, in terms of a workflow management service. Preliminary results using VisTrails workflow management system and a cluster machine show significant performance improvements. We will also validate these techniques by implementation and experimentation over the Grid5000 and Planetlab platforms, using synthetic and real data.

4. Concluding Remarks

The general impact of this collaboration on the two partners should be high quality joint papers, co-advising and exchange of Ph.D. students, training of Ms. students, dissemination of results through workshops jointly organized as well as special issues in top journals such as JOGC. Furthermore, the experience of both teams on using regularly the Grid'5000 platform [Furtado et al. 2008, Kotowski et al. 2008, Paes et al. 2008, Lima et al. 2009] fosters our joint work.

We also designed the Saravá project in order to exploit our complementarities and develop much synergy. For the Atlas project-team in particular, Saravá provides the opportunity to make progress on workflow management which is the area of expertise of the Brazilian group and to gain access to real users in bioinformatics and oil industry with which they have good collaboration. For the French side in particular, Saravá provides the opportunity to learn more on P2P in the context of the APPA project and thus exploit new research opportunities for P2P workflow management.

References

Abiteboul, S. and Polyzotis, N. The Data Ring: Community Content Sharing. Conference on Innovative Data Systems Research (CIDR), 154-163, 2007.

Agrawal, P., Benjelloun, O., Das Sarma, A., Hayworth, C., Nabar, S., Sugihara, T. and Widom, J. Trio: A System for Data, Uncertainty, and Lineage. Int. Conf. on Very Large Databases (VLDB), 1151-1154, 2006.

Akbarinia, R., Martins, V., Pacitti, E. and Valduriez, P. Design and Implementation of APPA. Global Data Management (Eds. R. Baldoni, G. Cortese, F. Davide), IOS Press, 2006.

Akbarinia, R., Pacitti, E. and Valduriez, P. Best Position Algorithms for Top-k Queries. Int. Conf. on Very Large Databases (VLDB), 495-506, 2007.

Furtado, C., Lima, A. ,Pacitti, E. ,Valduriez, P. ,Mattoso, M. Adaptive Hybrid Partitioning for OLAP Query Processing in a Database Cluster. Int. Journal of High Performance Computing and Networking, Special issue on Best Papers from SBAC2005, 5(4), p. 251-262, 2008.

Furtado, P. Node Partitioned Data Warehouses: Experimental Evidence and Improvements. Journal of Database Management 17(2), 42-60, 2006.

Lima, A. A. B., Furtado, C., Valduriez, P., Mattoso, M. Parallel OLAP Query Processing in Database Clusters with Data Replication. Distributed and Parallel Databases 25(1-2), 97-123, 2009.

Kotowski, N., Lima, A. A. B., Pacitti, E., Valduriez, P., Mattoso, M. Parallel query processing for OLAP in Grids. Concurrency and Computation. Practice & Experience 20(17), 2039 - 2048, 2008.

Meyer, L., Scheftner, D., Vöckler, J-S., Mattoso, M., Wilde, M. and Foster, I. An Opportunistic Algorithm for Scheduling Workflows on Grids. Int. Conf. on High Performance Computing for Computational Science (VecPar -2004), LNCS 3402, Springer, 1-12, 2005.

Özsu, T. and Valduriez, P. Principles of Distributed Database Systems. 2nd Edition, Prentice Hall, 1999 (3rd Edition forthcoming).

Pacitti, E., Valduriez, P. and Mattoso, M. Grid Data Management: open problems and new issues. Journal of Grid Computing, 5(3), 273-281, 2007.

Paes,M., Lima, A., Valduriez, P. and Mattoso, M. High-performance Query Processing of a Real-world OLAP Database with ParGRES. Int. Conf. on High Performance Computing for Computational Science (VecPar), Best student paper award for M. Paes, LNCS 5336, Springer, 188-200, 2008.

Serra da Cruz, S., Nogueira da Silva, F., Gadelha Jr., L., Reis Cavalcanti, M.C., Machado Campos, M.L., Mattoso, M. A Lightweight Middleware Monitor for Distributed Scientific Workflows. IEEE Int. Symp. on Cluster Computing and the Grid (CCGrid,), 693-698, 2008.

Valduriez, P., Pacitti, E. Data Management in Large-scale P2P Systems. Int. Conf. on High Performance Computing for Computational Science (VecPar -2004), Invited paper, LNCS 3402, Springer, 109-122, 2005.