INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team ATLAS

Complex Data Management in Distributed Systems

Rennes

THEME SYM

Activity Report

2006

Table of contents

1. Team

Head of Project Team
Patrick Valduriez [Director of Research, INRIA]

Vice-head of Project Team
Jean Bézivin [Professor, University of Nantes, HdR]

Faculty Members, University of Nantes
Marc Gelgon [Associate Professor]
José Martinez [Professor, HdR]
Noureddine Mouaddib [Professor, HdR]
Esther Pacitti [Associate Professor]
Guillaume Raschia [Associate Professor]

Technical Staff
Freddy Allilaire [Engineer, CDD Modelware and Modelplex]
Hugo Brunelière [Engineer, CDD Modelplex since November]

Administrative Assistant
Élodie Lizé [Secretary, CDD]

Ph.D. Students
Reza Akbarinia [Fellowship Iranian gov.]
Mikael Barbero [Fellowship Modelplex since October]
Cédric Coulon [ATER U. Nantes until September]
Marcos Del Fabro Didonet [Fellowship Microsoft]
Rabab Hayek [Fellowship MENRT]
Manal El Dick [Fellowship MENRT since September]
Jorge Manjarrez Sanchez [Fellowship Conacyt]
Vidal Martins [Univ. PUCPR, Brazil]
Lamiaa Naoum [ATER U. Nantes until September]
Afshin Nikseresht [Fellowship Iranian gov.]
Antoine Pigeau [ATER U. Nantes until September]
Jorge Quiane Ruiz [Fellowship Conacyt]
Amenel Voglozin [Fellowship MENRT]

Post-doctoral fellow
Frédéric Jouault [Fellowship OpenEmbeDD and U. Alabama at Birmingham (USA) since October]
Hanna Kozankiewicz [Fellowship ERCIM since November]
Ivan Kurtev [Fellowship Modelware until October]
Kwangjin Park [Fellowship INRIA since September]

2. Overall Objectives

2.1. Overall Objectives

Today's hard problems in data management go well beyond the traditional context of Database Management Systems (DBMS). These problems stem from significant evolutions of data, systems and applications. First, data have become much richer and more complex in formats (e.g., multimedia objects), structures (e.g., semi-structured documents), content (e.g., incomplete or imprecise data), size (e.g., very large volumes), and associated semantics (e.g., metadata, code). The management of such data makes it hard to develop data-intensive applications and creates hard performance problems. Secondly, data management systems need to scale up to support large-distributed systems (cluster systems, P2P systems) and deal with both fixed and mobile clients. In a highly distributed context, data sources are typically in high number, autonomous and heterogeneous, thereby making data integration difficult. Third, this combined evolution of data and systems gives rise to new, typically complex, applications with ubiquitous, on-line data access: virtual libraries, virtual stores, global catalogs, services for personal content management, services for mobile data management, etc.

The general problem can be summarized as complex data management in distributed systems. The Atlas project-team addresses this problem with the objective of designing and validating new solutions with significant advantages in functionality and performance. To tackle this objective, we separate the problem along four main dimensions which we address in four themes. The theme "database summaries" addresses the issues of data abstraction from large size databases. The theme "model management" addresses the issues of data abstraction from complexity. The theme "multimedia data management" deals with efficient and personalised access to multimedia data. Finally, the theme "distributed data management" addresses the problems of data replication and distributed query processing with complex data.

These dimensions are not independent and we foster cross-fertilization between themes. Examples of inter-theme research activities are: multimedia database summaries, multimedia data management in cluster systems, database summaries in P2P systems, and model management applied to distributed data integration.

3. Scientific Foundations

3.1. Scientific Foundations

Keywords: *Data management*, *database*, *distributed database*, *distributed systems*, *fuzzy logic*, *model engineering*, *multimedia*, *summaries*.

3.1.1. Data Management

Data management is concerned with the storage, organisation, retrieval and manipulation of data of all kinds, from small and simple to very large and complex. It has become a major domain of computer science, with a large international research community and a strong industry. Continuous technology transfer from research to industry has led to the development of powerful DBMSs, now at the heart of any information system, and of advanced data management capabilities in many kinds of software products (application servers, document systems, directories, etc.).

The fundamental principle behind data management is data abstraction, which enables applications and users to deal with the data at a high conceptual level while ignoring implementation details. The relational model, by resting on a strong theory (set theory and first-order logic) to provide data independence, has revolutionized database management. The major innovation of relational DBMS has been to allow data manipulation through queries expressed in a high-level (declarative) language such as SQL. Queries can then be automatically translated into optimized query plans that take advantage of underlying access methods and indices. Many other advanced capabilities have been made possible by data independence : data and metadata modelling, schema management, consistency through integrity triggers, transaction support, etc.

This data independence principle has also enabled DBMS to continuously integrate new advanced capabilities such as objet and XML support and to adapt to all kinds of hardware/software platforms from very small smart devices (PDA, smart card, etc.) to very large computers (multiprocessor, cluster, etc.) in distributed environments.

Following the invention of the relational model, research in data management continued with the elaboration of strong database theory (query languages, schema normalization, complexity of data management algorithms, transaction theory, etc.) and the design and implementation of DBMS. For a long time, the focus was on providing advanced database capabilities with good performance, for both transaction processing and decision support applications. And the main objective was to support all these capabilities within a single DBMS.

Today's hard problems in data management go well beyond the traditional context of DBMS. These problems stem from the need to deal with data of all kinds, in particular, text and multimedia, in highly distributed environments. Thus, we also capitalize on scientific foundations in multimedia data management, fuzzy logic, model engineering and distributed systems to address these problems.

3.1.2. Multimedia Data Management

Multimedia data such as image, audio or video is quite different from structured data and semi-structured (text) data in that it is media-specific (with specific operations) and described by metadata. Furthermore, useful representations of multimedia data, that are involved in storage and computation phases, are possibly voluminous and generally defined in high-dimensional spaces. Multimedia data management aims at providing high-level capabilities for organizing, searching and manipulating multimedia collections efficiently and accurately. To address this objective, we rely on the following research areas which we list in an order corresponding to the data flow: multimedia data analysis and pattern recognition, information retrieval and databases (mostly distributed). The overall architecture remains organised around the three fundamental parts of database design: modelling, querying and indexing. However, they have to be considerably adapted in order to manipulate multimedia data while maintaining the desired abstraction level.

With respect to modelling, multimedia data analysis performs automatic translation of raw multimedia data into sets of discriminant, concise descriptions that are used for indexing and searching. These descriptions range from low-level transforms on the original data (e.g. image texture features), that translate into feature vectors, to more abstract representations (e.g. parametric models), that often attempt to capture a class rather than an instance of multimedia elements. Furthermore, media content creators may add metadata information that conveys more semantics. Briefly stated, multimedia data analysis deals with the design of suitable observations from multimedia and pattern recognition techniques. Its interdependence with information retrieval and databases has encouraged the development of dedicated research branches, since many interesting applications consider multimedia information retrieval on voluminous data. Our work follows this direction.

Querying has been concerned with the conceptual access to data by the user with a high-level (SQL-like) query language on user-defined schemas. In contrast, techniques for querying multimedia data come from the information retrieval community. Athough extensible, each content-based multimedia system relies on a single, well-defined schema (similar to the document-term matrix from textual documents). Similarly, the common query in multimedia is a similarity search where the objects retrieved are ordered according to some scores based on a distance function defined on a feature vector, rather than a boolean expression. Similarly, relevance feedback has been introduced early in content-based systems since it is impossible to provide a concise description of a user's needs. In this respect, multimedia querying becomes mainly an interactive activity. Finally, it appears that several difficulties can be overcome by clustering multimedia data, something which is not new in databases, e.g., datawarehouses, but has to be done in a totally different way.

These important differences lead to reconsidering indexing too. Indexing is concerned with the physical access to multimedia data. The aim of indices is to rapidly access the data requested by the query. Efficient multimedia descriptors often span high dimensional spaces (say, 10 to 1,000 dimensions) since, to some extent, more features means more discriminant. Application of classical indexing structures (tree-based and hashing-based) supplied by database research is not effective, at least not in the straightforward manner, because these structures suffer from the "dimensionality curse problem", which states that the performance of indexing (and thus querying) degrades *severely* as the data dimensionality increases, in particular in the abovementioned dimension range. This particular issue is currently attracting much interest. The general problem is to achieve both high *effectiveness*, i.e., retrieving multimedia data that correspond to the user's needs and *efficiency* in order to scale up to large multimedia databases.

3.1.3. Fuzzy Logic

The ever growing size of databases makes data summarization needed in order to present the user a concise and complete view of the database. Our proposed summarization process [8] can roughly be described as a two step process. The first step is to rewrite the original database records into an unified user-oriented vocabulary. The second step is then to use a concept formation algorithm against the rewritten data. The fuzzy set theory provides mathematical foundations to manage these two steps in a more user-friendly and robust way than can be achieved with first order logic. Fuzzy sets theory was introduced by L.A. Zadeh in 1965 in order to model sets whose boundaries are not sharp. A fuzzy (sub)set F of an universe Ω is defined thanks to a membership function denoted by μ_F which maps every element x of Ω into a degree $\mu_F(x)$ in the unit interval $[0, 1]$. Thus, a fuzzy set is a generalization of regular set (whose membership function is defined on the pair (0,1).

In the first step, database tuples are rewritten using a user defined vocabulary. This vocabulary is intended to match as well as possible the natural language in which users express their knowledge. A database user usually refers to his or her data using a vocabulary appropriate for his field of expertise and understood by his or her fellows. For example, a salary will be said to be high, reasonable or average. This description in fact is an implicit categorization and there is no crisp border line between an average and a high salary. Fuzzy logic offers the mathematical ground to define such a vocabulary in terms of linguistic variables where each data is more or less satisfactorily described by the concept.

In a concept formation algorithm, new data are incorporated into a concept hierarchy using a local optimization criteria to decide how the hierarchy should be modified. A quality measure is evaluated to compare the effect of operators that modify the hierarchy topology namely, creating a new node, creating a new level, merging two nodes, or splitting one. Using fuzzy logic in the evaluation of this measure, our concept formation algorithm is less prone to suffer the well known threshold effect of similar incremental algorithms.

Database query languages are typically based on first order logic. To allow for more flexible manipulation of large quantities of data, we rest on fuzzy logic to handle flexible querying and approximate answering. Using the database summary, queries with too few results can be relaxed to retrieve partially satisfactory subsets of the database. The fuzzy matching mechanism also allows handling user queries expressed in vague or imprecise terms.

3.1.4. Model Engineering

A model is a formal description of a design artefact such as a relational schema, an XML schema, a UML model or an ontology. Data and meta-data modelling have been studied by the database community for a long time. We also witness the impact of similar principles in software engineering. Metamodels are used today to define domain specific languages that may help capturing the various aspects of complex systems. Models are no more viewed as contemplative artefacts, used only for documentation or for programmer inspiration. In the new vision, models become computer-understandable and may be applied a number of precise operations. Among these operations, model transformation is of high practical importance to map business expression onto executable distributed platforms but also of high theoretical interest because it allows establishing precise correspondences between various representation systems without ambiguity and, as such, is leverage for synchronization. Modelling naturally comes along with correspondences and constraints between models, i.e. the representation of a system by a model, the conformance of a model to a metamodel and the relation of one metamodel with another expressed by a transformation. In this area, research focuses on constraint languages and the traceability of transformations.

Considering models, meta-models, and model transformations as first class elements yields much genericity and flexibility to build complex data-intensive systems. A central problem of these systems is data mapping, i.e. mapping heterogeneous data from one representation to another. Examples can be found in different contexts such as schema integration in distributed databases, data transformation for data warehousing, data integration in mediator systems, data migration from legacy systems, ontology merging, schema mapping in P2P systems, etc. A data mapping typically specifies how data from one source representation (e.g. a relational schema) can be translated to a target representation (e.g. another, different relational schema or an XML schema). Generic model management has recently gained much interest to support arbitrary mappings between different representation languages.

3.1.5. Distributed Data Management

The Atlas project-team considers data management in the context of distributed systems, with the objective of making distribution transparent to the users and applications. Thus we capitalise on the principles of distributed systems, in particular, large-scale distributed systems such as clusters, grid, and peer-to-peer (P2P) systems, to address issues in data replication and high availability, transaction load balancing, and query processing.

Data management in distributed systems has been traditionally achieved by distributed database systems which enable users to transparently access and update several databases in a network using a high-level query language (e.g. SQL) [10]. Transparency is achieved through a global schema which hides the local databases' heterogeneity. In its simplest form, a distributed database system is a centralized server that supports a global schema and implements distributed database techniques (query processing, transaction management, consistency management, etc.). This approach has proved effective for applications that can benefit from centralized control and full-fledge database capabilities, e.g. information systems. However, it cannot scale up to more than tens of databases. Data integration systems extend the distributed database approach to access data sources on the Internet with a simpler query language in read-only mode.

Parallel database systems also extend the distributed database approach to improve performance (transaction throughput or query response time) by exploiting database partitioning using a multiprocessor or cluster system. Although data integration systems and parallel database systems can scale up to hundreds of data sources or database partitions, they still rely on a centralized global schema and strong assumptions about the network.

In contrast, peer-to-peer (P2P) systems adopt a completely decentralized approach to data sharing. By distributing data storage and processing across autonomous peers in the network, they can scale without the need for powerful servers. Popular examples of P2P systems such as Gnutella and Kaaza have millions of users sharing petabytes of data over the Internet. Although very useful, these systems are quite simple (e.g. file sharing), support limited functions (e.g. keyword search) and use simple techniques (e.g. resource location by flooding) which have performance problems. To deal with the dynamic behavior of peers that can join and leave the system at any time, they rely on the fact that popular data get massively duplicated.

Initial research on P2P systems has focused on improving the performance of query routing in the unstructured systems which rely on flooding. This work led to structured solutions based on distributed hash tables (DHT), e.g. CAN and CHORD, or hybrid solutions with super-peers that index subsets of peers. Although these designs can give better performance guarantees, more research is needed to understand their trade-offs between fault-tolerance, scalability, self-organization, etc.

Recently, other work has concentrated on supporting advanced applications which must deal with semantically rich data (e.g., XML documents, relational tables, etc.) using a high-level SQL-like query language. Such data management in P2P systems is quite challenging because of the scale of the network and the autonomy and unreliable nature of peers. Most techniques designed for distributed database systems which statically exploit schema and network information no longer apply. New techniques are needed which should be decentralized, dynamic and self-adaptive.

4. Application Domains

4.1. Application Domains

Keywords: *Application Service Provider (ASP)*, *distributed collaborative application*, *large decision-support application*, *multimedia personal database*.

Complex data management in distributed systems is quite generic and can apply to virtually any kind of data. Thus, we are potentially interested in many applications which help us demonstrate and validate our results in real-world settings. However, data management is a very mature field and there are well-established application scenarios, e.g., the On Line Transaction Processing (OLTP) and On Line Analytical Processing (OLAP) benchmarks from the Transaction Processing Council (TPC). We often use these benchmarks for experimentation as they are easy to deploy in our prototypes and foster comparison with competing projects.

However, there is no complete benchmark that can capture all the requirements of complex data management. Therefore, we also invest time in real-life applications when they exhibit specific requirements that bring new research problems. Examples of such applications are Application Service Provider (ASP), large-scale distributed collaborative applications, large decision-support applications or multimedia personal databases.

In the ASP model, customers' applications and databases (including data and DBMS) are hosted at a provider site and need be available, typically through the Internet, as efficiently as if they were local to the customer site. Thus, the challenge for a provider is to manage applications and databases with a good cost/performance ratio. In Atlas, we address this problem using a cluster system and by exploiting data replication and load balancing techniques.

Large scale distributed collaborative applications are getting common as a result of the progress of distributed technologies (GRID, P2P, and mobile computing). Consider a professional community whose members wish to elaborate, improve and maintain an on-line virtual document, e.g. reading or writing notes on classical literature, or common bibliography, supported by a P2P system. They should be able to read/write on the application data. An important aspect of large scale distributed collaborative applications is that user nodes may join and leave the network whenever they wish, thus hurting data availability. In Atlas, we address the issues of replication, query processing and load balancing for such applications assuming a P2P architecture (APPA) that is fully decentralized.

Large decision-support applications need to manipulate information from very large databases in a synthetic fashion. A widely used technique is to define various data aggregators and use them in a spreadsheet-like application. However, this technique requires the user to make strong assumptions on which aggregators are significant. In Atlas, we propose a new solution whereby the user can build a general summary of the database that allows more flexible data manipulation.

A major application of multimedia data management that we are dealing with in Atlas is multimedia personal databases which can help retrieve and classify personal audio-visual material stored either locally on a PC/Settop-box, or a mobile handset. Content-based retrieval from distributed multimedia documents is a second class of applications, which importance is bound to grow.

5. Software

5.1. ATL (Atlas Transformation Language)

Participants: Jean Bézivin, Frédéric Jouault, Patrick Valduriez.

URL: http://www.eclipse.org/gmt/

ATL is a transformation-based model management framework, with metadata management and data mappings as the main applications. The ATL language is designed to be general and abstract. We use it to compile transformations to many different target languages including XSLT and XQuery. The ATL design strives to be consistent with the MDA standards, in particular MOF/QVT. The ATL system is implemented in Java, and we are porting major transformation components to the .Net platform. ATL has been registered in 2004 (together with TNI-Software and the University of Nantes) to the APP (Agence pour la Protection des Programmes) and is released as Open Source Software under the Eclipse Public Licence.

5.2. AMW (Atlas Model Weaver)

Participants: Jean Bézivin, Marcos Didonet Del Fabro, Patrick Valduriez.

URL: http://www.eclipse.org/gmt/

AMW is a component-based platform for model weaving, i.e. establishing and managing correspondences between models. The platform is based on the Eclipse contribution mechanism: components are defined in separated plugins. The plugins are further interconnected to create the model weaver workbench. Components for user interface, matching algorithms and serialization of models may be plugged as necessary. We extended the Eclipse EMF architecture for model manipulation to coordinate the weaving actions. We use the EMF reflective API to obtain a standard weaving editor which adapts its interface according to metamodels modifications. The ATL transformation engine is plugged as the standard transformation platform. ATL has been registered in 2005 (together with the University of Nantes) to the APP (Agence pour la Protection des Programmes) and is released as Open Source Software under the Eclipse Public Licence.

5.3. APPA (Atlas Peer-to-Peer Architecture)

Participants: Reza Akbarinia, Vidal Martins, Esther Pacitti, Patrick Valduriez.

URL: http://www.sciences.univ-nantes.fr/lina/gdd/appa/

APPA is a new P2P data management system that provides scalability, availability and performance for applications which deal with semantically rich data (XML, relational, etc.). APPA provides advanced services such as queries, replication and load balancing. It is being implemented on top of the Open Source JXTA framework and tested on GRID5000.

5.4. RepDB*

Participants: Cédric Coulon, Esther Pacitti, Patrick Valduriez.

URL: http://www.sciences.univ-nantes.fr/lina/ATLAS/RepDB/

RepDB* is a data management component for replicating autonomous databases or data sources in a cluster system. It has been initially designed in the context of the Leg@net RNTL project and further developped in the context of the ACI MDP2P project. RepDB* supports preventive data replication capabilities (multi-master modes, partial replication, strong consistency) which are independent of the underlying DBMS. It uses general, non intrusive techniques. It is implemented in Java on Linux and supports various DBMS: Oracle™, PostGreSQL and BerkeleyDB. It has been validated on the Atlas 8-node cluster and another 64-node cluster at INRIA-Rennes. In 2004, we registered RepDB* (together with the University of Nantes) to the APP (Agence pour la Protection des Programmes) and released it as Open Source Software under the GPL licence.

5.5. SaintEtiQ

Participants: Noureddine Mouaddib, Guillaume Raschia, Amenel Voglozin.

URL: http://www.simulation.fr/seq

SAINTETIQ is a data summarisation system which provides synthetic user-friendly views over large databases. The fuzzy-set based representation of summaries provides an effective way of dealing with uncertainty in data, and natively supports flexible queries. A user-centric approach of a summary-oriented knowledge discovery process has been integrated into the prototype. We also enhanced the implementation with a set of tools to generate the background knowledge required for the summarisation process. Finally, a complete graphical user interface has been developed to support the user manipulating and browsing data, background knowledges and summaries. SAINTETIQ is now available as a Web service. It has been registered in 2005 to the APP (Agence pour la Protection des Programmes).

6. New Results

6.1. Database summaries

DBMS has become a very mature technology that is ubiquitous in information systems. Over time, the extensive use of DBMS technology has had major consequences in large organizations: the production of very large databases, the production of heterogeneous databases, and the increasing requirement of diverse applications to access those very large, heterogeneous databases. This creates difficult technical problems which get worse as DBMS technology improves and is more able to produce very large, heterogeneous databases. The SAINTETIQ system provides a novel solution for representing, querying and accessing large databases. We recently completed our work on summary querying techniques as well as decision support systems. We also pursued our work on summary management over P2P systems.

6.1.1. Summary query evaluation

Participants: Noureddine Mouaddib, Guillaume Raschia, Amenel Voglozin.

We proposed a querying mechanism for users to efficiently exploit the hierarchical summaries produced by SAINTETIQ. The first idea is to query the summaries with their own vocabulary, taking advantage of the hierarchical organization of the summaries [24]. The query evaluation matches summaries in the tree with fuzzy selection predicates of the query. The algorithm performs boolean set comparisons and uses the tree structure to cut branches and prune the search space. This leads to important gains in response time, in particular, in the case of null answers (i.e., of an empty result set), as only a small part of the summary hierarchy must be parsed, instead of the entire database.

As an extension of this work, we proposed to formulate the query predicates with a free user vocabulary rather than with the summary descriptors. We studied the query evaluation including the mapping between user concepts and summaries, using the symbolic-numerical interface of the fuzzy set theory [59].

6.1.2. Querying summaries: multidimensional indexing

Participants: Noureddine Mouaddib, Guillaume Raschia, Amenel Voglozin.

We investigated the area of multidimensional indexing from the point of view of space-partitioning. Through its architectural aspects, a summary hierarchy shares many features with multidimensional indexes (R-Tree, UB-Tree, X-Tree, ...). Current work on flexible querying uses the hierarchy as an index to select the appropriate database records, since in multidimensional indexing, each selection criterion reduces the search space for the other criteria.

Thus, we proposed to use summary hierarchies from the SAINTETIQ system as an index structure for a *PostgreSQL* access method. The objective of this work is to study the feasibility of using summaries as indexes, and determine the parameters that have an impact on the access method's performance. The study is limited to searching because defining a fully functional access method is a tedious task: updates and inserts are not yet supported. The index file is a binary version of the XML file produced by the SAINTETIQ prototype. The point in not modifying the tree structure is to evaluate the prototype's output as faithfully as possible. Although a summary hierarchy is intended for a different purpose and not optimized for querying, it provides acceptable response time for queries other than one-column queries. However, explaining the response time remains difficult. The immediate perspective is to use larger data sets so as to make the influence factors more distinct. Since it does not exist any benchmark data set for evaluating multidimensional indexing techniques, we are working on generating random data with a variable search space occupation ratio. Tuning that ratio will help simulate real data. Once the performance factors are known, it will be possible to adapt the construction of summaries for the purpose of using them as an index structure. very promising.

6.1.3. On-Line Analytical Processing of summaries

Participants: Lamiaa Naoum, Noureddine Mouaddib, Guillaume Raschia.

We proposed a general framework to explore and analyze database summaries built from massive data sets. Summaries are self-descriptive and higher-level views of groups of raw data. The overall on-line summarization processing is then intended to support a new approach to On-Line Analytical Processing of large data sets [13]. It aims at providing an effective and rich tool for visualizing, querying and accessing summaries considered as compressed semantic views of raw data.

Our contributions are as follows. First, we defined a logical data model called *summary partitions*, by analogy with OLAP datacubes. The aim is to provide the end-user with an effective way of presenting a reduced version of the data set as well as to support analysis. Pre-built and ordered partitions are considered on the basis of a process dedicated to the generation of summaries at different levels of granularity. Second, we defined a collection of algebraic operators over the space of summary partitions: relational, granularity and structuring operators are designed for on-line analytical processing of summarized versions of the data [50]. Third, we addressed the issue of representing the summary partitions, especially to make as simple and informative as possible the summaries to the end-user. To achieve this, we tried to build fuzzy prototypes for the summaries, as a pre-visualization mechanism [49].

6.1.4. Summaries over a P2P architecture

Participants: Rabab Hayek, Noureddine Mouaddib, Guillaume Raschia, Patrick Valduriez.

We started to study the integration of a new service for managing summaries in P2P systems. In such a context, summaries have two main virtues. First, they can be directly queried and used to approximately answer a query without exploring the original data. Second, as semantic indexes, they support locating relevant nodes based on data content.

The first idea was to incrementally construct a global summary which describes all the data shared in the network. Distributed storage of such a global summary is, for instance, managed by a dedicated service and peers call that service with the right global summary key. For a given query, the global summary is first used to determine the set of nodes having relevant data. Then, those nodes are directly contacted. Simulation results have shown that the cost of query routing is significantly reduced compared to flooding approaches. However, converging to, and maintaining such a global summary is hard and costly in a P2P environment. Current work consists in retrieving a sort of natural partitioning of unstructured networks in peer domains, each managing its global summary. Our approach relies only on scale-free network properties such as the power law degree distribution and the associated clustering coefficient distribution. The intra-domain links will be used as summary links (i.e. index links) to maintain the global summary, while the inter-domain links will be used as search links to propagate the query among domains. We aim at finding the optimal number of domains that minimizes the total cost of query routing and summary maintenance.

6.2. Model management

A model is a structure that represents a design artefact such as a database schema, an interface definition, an XML type definition, a UML model or a Web document. More generally a model can naturally be represented as a graph-based structure. We have proposed a formal characterization of models as a set of related graphs [38], [41]. All the operational tools built within the project refer to this common conceptual framework which is an improvement on our previous work. Developers of information systems must typically deal with different models and perform transformations between models. Examples of transformations are: mapping heterogeneous data source descriptions in a global schema to perform data warehousing, converting XML documents into HTML, or generating EJB or .Net component definitions from a UML model. Today, most of these transformations are still programmed using specific languages like SQL, XSLT or even Java, Perl, or C. As information systems become more complex and must support cooperation of heterogeneous applications and components, there is an urgent need to propose more systematic ways to develop transformations.

Model management aims at solving this problem by providing techniques and tools for dealing with models and model transformations in more automated ways. It has been studied independently for years by several research communities such as databases, document management, and software engineering. One of the major problems is the multiplicity of input and output format and transformations systems, e.g., from LaTeXto HTML or from SQL to XQuery. There is much to gain if we could handle these various transformations in a more generic way with a coordinated family of languages. To contribute to this evolution, we have continued to refine ATL (Atlas Transformation Language) which is now the basis for AMMA (ATLAS Model Management Architecture). Our research activities in AMMA concern model transformation and weaving, global management of related resources (mainly models, metamodels and transformations), and the integration of these functionalities into an open model management platform.

6.2.1. Model transformation

Participants: Freddy Allilaire, Jean Bézivin, Frédéric Jouault, Ivan Kurtev, Patrick Valduriez.

Model transformation, e.g. mapping a relational database schema into an XML schema, is a very useful and important operation in model management. We have proposed ATL, a combined declarative/imperative language that allows to transform source models into target models. Like the source and target models, the transformation program is itself a model and thus conforms to a given metamodel [29].

We have continued our implementation of ATL on Eclipse as part of the GMT open source project. In addition to the ATL engine, a complete integrated development environment (IDE) has been built and also released as GMT open source. The IDE for ATL allows transformation editing and debugging (syntax coloring, step by step execution, breakpoints, environment observation, etc.). We developped several examples of model transformations as part of a basic library.

In the context of the ModelWare project, ATL has been stabilized and applied to several case studies. ATL is now being used by a strong international research community on more than 200 sites. Several companies including Airbus, CS, oAW, NASA/JPL, SODIUS, etc. are now developing transformations in ATL. The language is partially aligned on the recent QVT normative recommendation [39] and has recently been proposed as one of the standard Eclipse solutions for Model2Model Transformations. We have also shown how ATL allows to bridge the OMG and various other environments like GME or Microsoft DSL Tools. An extensibility scheme has also been proposed for ATL [42].

An original aspect of the ATL implementation is that it is based on the public definition of a portable transformation virtual machine. The specification of this virtual machine has been publicly released on Eclipse. We have also proposed KM3 (Kernel MetaMetaModel), a domain specific language for specifying metamodels [38], for example those describing tools' internal data formats (MS Excel, MS Project, MatLab, Bugzilla, Mantis, etc.).

Taking stock on the ATL implementation framework, we have obtained several original results that are summarized in [12]. We have shown the possibility to express a model verification by a pure transformation. Not only the verification criteria may be expressed by a separate model, but also the diagnostic result may be expressed as a model conforming to a variable metamodel. We have proposed a proof of concept based on our ATL implementation. We are currently extending it to measure models by using a similar model-based organization. Many of our practical applications make use of higher order transformations, i.e. transformations taking transformations as input or/and producing transformations as output. This has been made easy because a transformation is not only a program, but also a model conforming to a precise metamodel [29].

6.2.2. Model weaving applied to data mapping

Participants: Marcos Didonet Del Fabro, Jean Bézivin, Frédéric Jouault, Patrick Valduriez.

Mapping between heterogeneous data is a central problem in many data-intensive applications. A typical data mapping specifies how data from one source representation (e.g. a relational schema) can be translated to a target representation (e.g. an XML schema). Although data mappings have been studied independently in different contexts, there are two main issues involved. The first one is to discover the correspondences between data elements that are semantically related in the source and target representations. This is called schema matching in schema integration and many techniques have been proposed to (partially) automate this task. After the correspondences have been established, the second issue is to produce operational mappings that can be executed to perform the translation. Operational mappings are typically declarative, e.g. view definitions or SQL-like queries. However, using one mapping language causes serious limitations and makes mapping management difficult.

We have proposed a solution based on model weaving which can better control the trade-off between genericity, expressiveness, and efficiency of mappings. In other words, our objective is to support generic data mapping (as in other model management systems but with a different approach) while exploiting specific mapping languages and engines, such as XQuery, SQL or ATL. Our solution considers mappings as models and exploits specific mapping engines. We defined model weaving as a generic way to establish element correspondences. Weaving models may then be used by a model transformation language to translate source model(s) into target model(s). We validated our approach using the ATLAS Model Weaver (AMW) prototype on several application scenarios including the operational interoperability between different tools [34] Our experiments have shown that many different proposals may be unified by our model-based approach. Coupling a weaving facility (like AMW) with a transformation facility (such as ATL) gave us good efficiency and flexibility. We have illustrated the joint use of AMW and ATL (i.e. generating executable transformations from correspondances) by several practical projects [33].

The number of possible applications of model weaving techniques is rapidly rising. A small but convinced community of external users (mainly academics) has started using the Eclipse AMW prototype (www.eclipse.org/gmt/amw/).

6.2.3. Global model management

Participants: Freddy Allilaire, Jean Bézivin, Frédéric Jouault, Hugo Bruneliere, Patrick Valduriez.

Within a model management environment, the main elements produced or consumed are models, metamodels, correspondences or transformations. However, in order to allow for the manipulation of other resources such as XML documents, database tables or flat files, collections of generic importers and exporters are needed. Special attention should be given to the global management of all these resources. These models are explicitly typed by their corresponding metamodels allowing to define the signature of each tool.

In our approach, all the information about the components known to a given platform is stored in a specific model named "megamodel". A megamodel is a kind of model registry that stores reference and metadata information on all accessible resources, including relations between these resources. It allows us to build a minimal and highly extensible infrastructure. In particular, this allows easy extension of a local platform towards a distributed platform such as a P2P system without significant modification of tool interoperability mechanisms. Furthermore the approach fits well within the general conceptual scheme developed for model management. Experimental validation is being done through the AM3 (ATLAS MegaModel Manager) tool to record and control the global relations betwen model components (www.eclipse.org/gmt/am3/). This validation is being pursued within the ModelPlex project, on the use cases defined by our industrial partners: SAP, Thales, Telefonica and Westerngeco/Schlumberger.

As part of the current ModelPlex project, this global management approach has also been applied to the problem of model driven reverse engineering [17]. The idea is to develop tools to extract homogeneous models from legacy programs and data that have been developed since fifty years, under various technologies [30]. Capturing the various kinds of artefacts present in the information system portfolio of a bank for example is very challenging, difficult and useful for software and system modernization or for handling different situation of company merging for example. Legacy systems are highly complex systems and their understanding is currently an open problem. The application of model driven techniques to this problem is one important goal of the ModelPlex project. In addition to the AM3 Eclipse component mentioned above, the new MoDisco collaborative Eclipse project is targeting the area of Model Driven Reverse Engineering.

The problem of global model management also arises when we are dealing with a high number of different Domain Specific languages (DSLs). We are going to apply similar techniques to the problem of DSL coordination in the FLFS project (Families of languages for Families of Systems). The problem of defining abstract syntaxes, concrete syntaxes and semantics to a DSL has been studied in [61], [38], [37] and [60].

6.3. Multimedia data management

The ability to store multimedia information in digital form has spurred both the demand and offer of new electronic appliances (e.g., DVD players, digital cameras, mobile phones connected to the Web, etc.) and new applications (e.g., interactive video, digital photo album, electronic postcard, distance learning, etc.). The increasing production of digital multimedia data magnifies the traditional problems of multimedia data management and creates new problems such as content personalisation and access from mobile devices. The major issues are in the areas of multimedia data modelling, physical storage and indexing as well as query processing with multimedia data.

6.3.1. Scaling up multimedia indexing

Participants: José Martinez, Patrick Valduriez, Jorge Manjarrez.

An image database management system should rely on a DBMS, but we identified several shortcomings to relational DBMSs. Five key points have been determined and some answers provided : reducing the physical size of metadata, introducing replication, exploiting parallelism, classification and distribution of data, differentiated physical indexing. Combining classification and parallel computing enables, in principle, algorithms of sublinear complexity [63]. We are conducting experiments to validate this theoretical result, with several simulations considering various assumptions [62].

6.3.2. Personal image collection management from mobile devices

Participants: Marc Gelgon, Antoine Pigeau, Afshin Nikseresht.

Extension of image retrieval systems to address personal image collections appears among emerging needs in both industrial and academic worlds. In particular, mobile devices such as camera-equipped phones are an interesting case for content creation and retrieval. Our objective is to recover the natural spatial, temporal or spatio-temporal structure present in such a data set. We had previously developped [54] a technique for building and tracking a hierarchical temporal and geographical structure, modelled as a hierarchy of mixture models. We have proposed an alternative to building and tracking a hierarchy of mixture models, which involves a lower computational cost and is more robust to the non-Gaussianity of clusters, as the upper level of the hierarchy is fitted to the lower-levels of the model, rather than to the data. The iterative nature of this fitting enables a predict/update mechanism as new data flows in [55].

6.3.3. Decentralized, distributed learning of multimedia class models

Participants: Marc Gelgon, Afshin Nikseresht.

A fundamental task in multimedia content-based retrieval is class characterization (defining observations from audiovisual data and capturing its variability or criteria discriminating it from other classes). We assume large amounts of data, partly labelled with their class identifier, available on-line on a large scale. Since model estimation for many classes is expensive and the data is assumed initially distributed, we proposed a distributed learning technique. This context fits the flexible, dynamic distribution approach favoured in the Atlas project.

In the case of Gaussian mixture models (one of the most useful models for modelling multimedia data), we proposed a parameter estimation technique based on gossip propagation and aggregation of models through the network. Model aggregation proceeds by minimizing an approximate KL-divergence (loss), at parameter level, which avoids moving heavy multimedia data over the network. Concurrently, the reliability of models being estimated is assessed [52], [51].

6.3.4. Scaling up retrieval in a collection of mixture models

Participants: Jamal Rougui, Marc Gelgon, José Martinez.

When querying a multimedia database for the class identifier, a central quantity to evaluate is the likelihood of the query, given each candidate models stored in the database (analogous to a distance). We focus on the case where models take the form of Gaussian mixtures and we consider, as a practical application and for implementation, the speaker recognition task. While tree-based indexing structures have been intensively studied for speeding up retrieval when database entries are vectors, open issues remain when retrieval operates among probabilistic models, which are fundamental to multimedia application. We proposed techniques for building a tree of Gaussian mixture models as an index [58], [57]. This tree is built by determining groups of models, assigned to parent node, so that evaluation of likelihood, given a parent node, supplies a value as close as possible to the one that would have been computed from its children. Due to the relation between likelihood loss and KL-divergence, optimal grouping of children into parent nodes amounts to optimizing the latter criterion.

6.4. Distributed data management

In a large scale distributed system, data sources are typically in high numbers, autonomous (under strict local control) and very heterogeneous in size and complexity. Data management in this context offers new research opportunities since traditional distributed database techniques need to scale up while supporting high data autonomy, heterogeneity, and dynamicity.

We are interested in database clusters and peer-to-peer (P2P) systems which are good examples of large-scale distributed systems of high practical interest. However, to yield general results, we strive to develop common algorithmic solutions with the right level of abstraction from the context. In 2006, we continued our work on data management in P2P systems with the design of Atlas Peer-to-Peer Architecture (APPA) with new techniques for distributed reconciliation of replicated data and top-k query processing. We have also proposed a new query load balancing strategy.

6.4.1. Design of APPA

Participants: Reza Akbarinia, Vidal Martins, Esther Pacitti, Patrick Valduriez.

APPA (Atlas Peer-to-Peer Architecture) is a new P2P data management system which we are building [25], [14]. Its main objectives are scalability, availability and performance for advanced applications. These applications must deal with semantically rich data (e.g., XML documents, relational tables, etc.) using a high-level SQL-like query language. As a potential example of advanced application that can benefit from APPA, consider the cooperation of scientists who are willing to share their private data (and programs) for the duration of a given experiment. The main originality of APPA's architecture is to be network-independent and organized in terms of basic and advanced services that can be implemented over different P2P networks (unstructured, DHT, super-peer, etc.).. This allows us to exploit continuing progress in such systems. To deal with semantically rich data, APPA supports decentralised schema management, data replication and updates, query processing and load balancing.

We have started the implementation of APPA using the JXTA framework. APPA's advanced services are provided as JXTA community services. Only the P2P network layer of the APPA implementation depends on the JXTA platform. Thus, APPA is portable and can be used over other platforms by replacing the services of the P2P network layer. We validated some of APPA's services on the cluster of Paris team at INRIA, which has 64 nodes. Additionally, in order to study the scalability of these services with larger numbers of nodes, we implemented simulators using Java and SimJava. The current version of the APPA prototype and its service simulators manage data using a Chord DHT. Experimental results showed that simulators are well calibrated and the implemented services have good performance and scale up. Implementing on top of JXTA was relatively easy, but we faced some problems to adapt JXTA core services and to deploy the APPA prototype for tests using the JXTA framework.

6.4.2. Distributed Reconciliation in APPA

Participants: Reza Akbarinia, Vidal Martins, Esther Pacitti, Patrick Valduriez.

An important aspect of collaborative applications is that users may perform updates and may join and leave the network whenever they wish, thus hurting data availability. Data replication can then be used to increase availability. Lazy master replication is not applicable in P2P because a single master peer hurts availability. Thus, we focus on a multi-master approach in which all peers may update replicated data. In particular, we are interested in optimistic solutions based on semantic reconciliation, because it provides more flexibility and supports connections and disconnections. Existing semantic reconciliation solutions are typically performed at a single peer which may become a bottleneck in a large-scale system. In [46], we proposed a Distributed Semantic Reconciliation Algorithm (DSR) that enables optimistic multi-master replication and assures eventual consistency among replicas. DSR works in the context of APPA system. The DSR/APPA integration is presented in [43].

The P2P network costs incurred in data accesses may vary significantly from node to node and have a strong impact on the reconciliation performance. Thus, network costs should be considered to perform reconciliation efficiently. In [45], we proposed the P2P-reconciler protocol that extends DSR by providing a cost model for selecting reconciler nodes based on network latencies. For computing communication costs, P2P-reconciler uses local information and deals with the dynamic behavior of nodes. It also limits the scope of event propagation (e.g. joins or leaves) in order to avoid network overload. The cost model was improved in [44] by taking into account different bandwidths and data transfer costs. We proved the algorithms' correctness and validated them through implementation and simulation. The experimental results showed that cost-based reconciliation outperforms a random approach for selecting reconciler nodes by a factor of 26. In addition, the number of connected nodes does no affect the performance of cost-based reconciliation since the reconciler nodes are as close as possible to data accessed in the reconciliation. Compared with a centralized solution, our algorithm yields high data availability and excellent scalability, with acceptable performance and limited overhead.

6.4.3. Top-k query processing in APPA

Participants: Reza Akbarinia, Vidal Martins, Esther Pacitti, Patrick Valduriez.

High-level queries over a large-scale P2P system may produce very large numbers of results that may overwhelm the users and generate heavy network traffic. This is caused mainly by high numbers of query answers, many of which are irrelevant for the users. One solution to this problem is to use Top-k queries whereby the user can specify a limited number (k) of the most relevant answers. In [15], we proposed FD, a (Fully Distributed) framework for executing Top-k queries in unstructured P2P systems, with the objective of reducing network traffic. FD consists of a family of algorithms that are simple but effective. FD is completely distributed, does not depend on the existence of certain peers, and addresses the volatility of peers during query execution. We validated FD through implementation over a 64-node cluster and simulation using the BRITE topology generator and SimJava. Our performance evaluation shows that FD can achieve major performance gains in terms of communication and response time.

In [27], we considered the problem of top-k query processing in Distributed Hash Tables (DHTs). The most efficient approaches for top-k query processing in centralized and distributed systems are based on the Threshold Algorithm (TA) which is applicable for queries where the scoring function is monotone. However, the specific interface of DHTs, i.e. data storage and retrieval based on keys, makes it hard to develop TA-style top-k query processing algorithms. We proposed an efficient mechanism for top-k query processing in DHTs. Although our algorithm is TA-style, it is much more general since it supports a large set of non monotone scoring functions including linear functions. In fact, it is the first TA-style algorithm that supports linear scoring functions. We proved analytically the correctness of our algorithm and validated it through a combination of implementation and simulation. The results show very good performance, in terms of communication cost and response time.

6.4.4. Query load balancing in large-scale distributed systems

Participants: Jorge Quiane Ruiz, Patrick Valduriez.

We consider dynamic distributed systems, providing access to large numbers of heterogeneous, autonomous information sources [21]. We assume that information sources play basically two roles: consumers that generate queries and providers which perform requests and generate informational answers. In this context, we considered the query allocation problem. Provider sources are heterogeneous, autonomous, and have finite capacity to perform queries. A main objective in query allocation is to obtain good response time. Most of the work towards this objective has dealt with the problem of balancing the query load among providers. But little attention has been paid to satisfy the providers'interests in performing certain queries. In [56], we addressed both sides of the problem. We proposed a query allocation process which allows providers to express their intention to perform queries based on their preference and satisfaction. We compared our solution to both query load balancing and economic approaches. The experimentation results show that our approach yields high efficiency while supporting the providers' preferences in adequacy with the query load. Also, we showed that our approach guarantees interesting queries to providers even under low arrival query rates. In the context of open distributed systems, our solution outperforms traditional query load balancing approaches as it encourages providers to stay in the system, thus preserving full system capacity.

7. Contracts and Grants with Industry

7.1. IP Modelware (2004-2006)

Participants: Freddy Allilaire, Jean Bézivin, Patrick Valduriez.

In this very large european project, we work with Thales (project leader), IBM UK, IBM Israel, France Telecom R&D, LIP6 and the major industrial actors in model engineering in Europe. The objective is to demonstrate within 4 years the industrial application of model engineering.

7.2. IP Modelplex (2006-2009)

Participants: Freddy Allilaire, Jean Bézivin, Mikael Barbero, Hugo Brunelière, Patrick Valduriez.

The ModelPlex projects (with Thales, IBM, Sodifrance, SAP, etc.) aims at defining a coherent infrastructure for the development of complex systems, where the complexity factor corresponds to several factors like size, heterogeneity, dynamic evolution, distribution and subsystem autonomy. One example of highly heterogeneous systems are legacy systems that have been built and adapted on long period of time, using different technologies. Model driven reverse engineering is one important work subject in ModelPlex.

7.3. STREP Grid4All (2006-2008)

Participants: Reza Akbarinia, Vidal Martins, Esther Pacitti, Jorge Quiane, Patrick Valduriez.

The project is with France Telecom R&D (leader), INRIA (Atlas, Grand-Large, Régal, and Sardes), Kungliga Tekniska Hoegskolan, Swedish Institute of Computer Science, ICCS (Greece), University of Piraeus Research Center, Universitat Politècnica de Catalunya and Rededia S.L. (Spain). Atlas and INRIA-Rennes are the INRIA representatives. The goal of Grid4All is to develop a grid infrastructure and middleware for the collaboration of dynamic, small virtual organizations such as communities, schools and families. The main technical innovation is to foster the combination of grid and P2P techniques to provide a light-weight, flexible solution. Atlas contributes to the definition of the P2P infrastructure (which is based on APPA) and to the development of two key services: resource discovery (using our mediation techniques) and optimistic replication (using our semantic reconciliation techniques).

7.4. Microsoft Research (2003-2006)

Participants: Jean Bézivin, Patrick Valduriez.

The objective is to contribute to the development of the AMMA model management framework and foster the dissemination of our results as Open Source Software under a non restrictive license. In particular, we are adapting the AMMA framework to the principles and tools of the Microsoft Software Factory approach (Visual Studio 2005 Team System). Artefacts built by tools as ATL, AM3, AMW should be made available to the Microsoft environment with the help of technical space projectors.

7.5. Caroll Motor (2003-2006)

Participants: Jean Bézivin, Patrick Valduriez.

In the context of the Caroll joint venture between INRIA, CEA and Thales, the objective of the Motor project was to study the interoperability of model transformation languages. In this project, we showed interoperability results based on ATL. More generally, the principles of the AMMA platform are also being studied in this project.

7.6. Programme blanc ANR FLFS (2006-2009)

Participants: Jean Bézivin, Frédéric Jouault.

FLFS means "Families of Languages for Families of Systems" The objective of the FLFS project (with OBASCO, Phoenix, INRIA, ENST) is to study the continuum domain modeling/implementation with the complementary technologies of model engineering, domain specific languages and aspect oriented programming. In this project, the ATLAS-GDD team develops concrete solutions based on the AMMA platform and more specifically on the ATL and TCS languages.

7.7. RNTL OpenEmbeDD (2006-2009)

Participants: Jean Bézivin, Frédéric Jouault, Patrick Valduriez.

The project involves Airbus, Anyware, CEA, CS, FT R&D, LAAS, Thales DAE, Thales TRT, Verimag and Atlas. OpenEmbeDD is an Eclipse open-source platform based on the model engineering principles for the software engineering of embedded systems. The ATLAS project-team provides a model transformation virtual machine, and components for global model management. Furthermore the project-team is working in this project on interoperability between the AMMA and GME platforms.

7.8. OpenDevFactory, Paris Competitivity Cluster "Usine Logicielle" (2006-2008)

Participants: Freddy Allilaire, Jean Bézivin.

In this project, ATLAS-GDD works, in relation with different industrial partners (Thales, EADS, Dassault, Esterel, Softeam, etc.) to develop mature model driven solutions to several software production problems. Based on our generic open source AMMA platform, complex chains of data transformations are considered as an alternative to more conventional development solutions.

8. Other Grants and Activities

8.1. Regional Actions

We are involved in four projects:

8.1.1. COM, Région Pays-de-la-Loire (2000-2006)

The Atlas project-team participates in the COM project funded by the “Région des Pays de la Loire” (2000-2006). The objective of the COM project is to promote research in computer science in the region, in particular the creation of LINA (Laboratoire d’Informatique de Nantes Atlantique), a UMR between CNRS, University of Nantes and École des Mines de Nantes.

8.1.2. MILES Project (2007–2010)

N. Mouaddib coordinates the MILES project, funded by Region Pays-de-la-Loire. MILES is the main Region-funded project on information and communication technologies. Within the MILES project, M. Gelgon is in charge of a sub-project dealing with distributed multimedia systems, involving the Atlas project-team and IRCCyN (IVC group). This sub-project addresses, on one side, multimedia data learning and classification in a distributed computing and storage context and, on the other side, secure, distributed storage with involving techniques specific to multimedia data.

8.1.3. Pôle de compétitivité (2006-2008)

The Atlas project-team is involved in the PC Images & Réseaux through the SafeImage project, jointly with Alcatel, IRCCyN, and QoSmetrics. The Atlas project-team is involved in the ANR Safeimage project (3/2007-2010), dealing with inspection of data in high-speed routers for security purposes. The task devoted to Atlas is classification of multimedia data (examining how to scale up learning and recognition tasks with state-of-the-art classifiers in future routers). This project is further supported by Pôle de Competitivité Images & Réseaux.

8.1.4. GeoPict, Mégalis, Région Pays-de-la-Loire (2005-2006)

Participants: José Martinez, Noureddine Mouaddib.

GeoPict is a joint project between Magic Instinct Software, a startup in Nantes, and three research teams at LINA and IRCCyN. The motivation is to take advantage of high-speed networks in order to create new services to transmit huge amounts of information such as multimedia data. The goal of the project is to provide an on-line service to access and visualise geo-referenced videos connected to a geographic information system (namely, the register – *cadastre* –of each city). More precisely, videos are recorded at 360 degrees while driving along the streets of a city. This information has to be stored, connected to the geographic database thanks to the spatial positioning recorded during the travelling. Next, the video information must be mixed with 3D geographical models in order to reconstruct panoramic views at any point in the city, as well as virtual reality trips.

8.2. National Actions

We are involved in four projects:

8.2.1. ARA Massive Data Respire (2006-2008)

Participants: Reza Akbarinia, Vidal Martins, Esther Pacitti, Jorge Quiane, Patrick Valduriez.

The project Respire involves LIP6 (leader), Paris (IRISA), Regal (INRIA et LIP6) and INT. The objective is to propose a P2P infrastructure for resource and data sharing in large scale networks. We study the following problems: resource catalog, dynamic clustering of peers, replication, query processing and equitable mediation. To validate the infrastructure, the Atlas project-team develops services in the context of the APPA prototype.

8.2.2. ARA Massive Data MDP2P (2003-2006)

Participants: Reza Akbarinia, Vidal Martins, Esther Pacitti, Jorge Quiane, Patrick Valduriez.

The project MDP2P (Massive data management in peer-to-peer systems) is led by the Atlas project-team and involves three other INRIA project-teams: Paris and Texmex in Rennes, and Gemo in Orsay. The main objective of the project is to provide high-level services for managing text and multimedia data in large-scale P2P systems. Similar to database management systems, these services are not limited to file sharing (like current P2P systems) and need be high-level with query capabilities and transactional support (for data consistency). Furthermore, they must provide good access performance which can be obtained through data replication, distributed query optimization, and parallel query processing. To validate our P2P approach and show its wide range of application, we concentrate on two different P2P contexts that we know well: the Web and clusters of PC.

8.2.3. ARA Massive Data SemWeb (2004-2007)

Participants: José Martinez, Noureddine Mouaddib, Guillaume Raschia.

The project SemWeb (Querying the Semantic Web with XQuery) involves PRiSM, Versailles, CNAM, Paris, LIP6, Paris, SIS, Toulon and LINA, Nantes. The project aims at studying problems and providing solutions to XML-based mediators in the context of the Semantic Web using XQuery as the common querying language. Foreseen main problems are scalability of the proposed architecture, integration of heterogeneous sources of information, and dealing with metadata. The results of the project should be an homogeneous mediator architecture, exemplified on typical applications, and delivered as a open-source software.

8.2.4. ARA Massive Data APMD (2004-2007)

Participants: José Martinez, Noureddine Mouaddib, Guillaume Raschia.

The project APMD (Personalised Access to Masses of Data)(2004-2007) involves PRiSM, Versailles, CLIPS-IMAG, Grenoble, IRISA, Lannion, IRIT, Toulouse, LINA, Nantes and LIRIS, Lyon. The goal of the project is to improve the quality of retrieved information thanks to personalisation techniques or, in other words, to personalise the retrieved information in order to improve its quality with respect to the end user. This is of major importance for applications targeted to a large audience, like e-commerce, which have to take into account a large number of parameters: heterogeneous sources of information, various data formats, used languages, large amount of available data, etc. More precisely, the project has to define precisely which are the components of a user's profile, how it can evolve, and then take advantage of these profiles in order to filter and present adaptively the retrieved information, especially when dealing with huge amounts of information.

8.3. International actions

We are involved in the following international actions:

- the Interop European network of excellence (2003-2006) with all the research groups working on model engineering in Europe;

- the Daad (Distributed computing with Autonomous Applications and Databases) project (2003-2007), funded by CAPES in Brazil and COFECUB in France, with UFRJ, Brazil, on distributed data management;
- the GridData project (2005-2008) Databases) project, funded by CNPQ in Brazil and INRIA, with the Gemo project-team and the universities PUC-Rio and UFRJ, Brazil, on data management in Grid environments;
- the STIC multimedia network between France and Morocco, with University Mohammed V of Rabat, EMI, ENSIAS and University of Fès;
- the STIC Software Engineering project between France and Morocco with University Mohammed V of Rabat, EMI, ENSIAS and University of Fes;
- the OMG consortium, in which J. Bézivin contributes to the MDA work.

Furthermore, we have regular scientific relationships with research laboratories in

- North America: Univ. of Waterloo (Tamer Özsu), NYU (Dennis Shasha), New Jersey Institute of Technology (Vincent Oria), Wayne State University (Farshad Foutouhi and Wiliam Grosky), Kettering University (Peter Stanchev), Riad Hammoud (Delphi);
- Europe: Univ. of Madrid (Ricardo Jimenez-Periz), Univ. of Twente (Mehmet Aksit), Univ. of Roskilde (Henrik Larsen), Nokia (Andreas Myka), ;
- Others: Univ. Federal of Rio de Janeiro (Marta Mattoso), Tokyo Metropolitan University (Hiroshi Ishikawa)

9. Dissemination

9.1. Animation of the scientific community

The members of the Atlas project-team have always been strongly involved in organising the French database research community, in the context of the I3 GDR and of the conference Bases de Données Avancées (BDA).

J. Bézivin is a member and co-founder of the steering committee of the ECOOP (AITO) and UML/Models conferences. In 2006, he is a co-organizer of a track on model transformation of the ACM Symposium of Applied Computing to be held in Dijon.

The 20th ECOOP conference took place in Nantes in July 2006. The conference co-chairs are J. Bézivin and Pierre Cointe (Obasco project team in Nantes) who founded ECOOP 20 years ago. In addition to their international recognition, this also demonstrates the strong cooperation between Atlas and Obasco. ECOOP is currently ranked 39th on a total of more than 1200 conferences and computer science journals by CiteSeer.

In 2008, the Atlas project-team will organize the EDBT conference in Nantes. N. Mouaddib is organization chair and P. Valduriez is general chair. EDBT is currently ranked 170th (top 13 percent) on Citeseer.

9.2. Editorial Program committees

Participation in the editorial board of scientific journals:

- Distributed and Parallel Database Systems, Kluwer Academic Publishers: P. Valduriez.
- Internet and Databases: Web Information Systems, Kluwer Academic Publishers: P. Valduriez.
- Ingenierie des Systèmes d'Information, Hermés : N. Mouaddib, P. Valduriez.
- Journal of Object Technology: J. Bézivin.
- SoSyM, Software and System Modeling, Springer Verlag: J. Bézivin.
- IEEE Transactions Journal on Fuzzy Systems : N. Mouaddib.
- International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems: N. Mouaddib.

Participation in conference programme committees :

- Int. Conf. on Very Large Databases (VLDB) 2006, 2007: P. Valduriez.
- ACM-SIGMOD Int. Conf. 2007: P. Valduriez.
- IEEE Int. Conf. on Data Engineering 2007: P. Valduriez (Ind. PC chair), E. Pacitti.
- IEEE Int. Conf. on Distributed Computing Systems (ICDCS) 2007, Data Management track: P. Valduriez (PC chair), E. Pacitti.
- Int. Conf. on Parallel and Distributed Computing (Euro-Par) 2006: P. Valduriez (chair, Distributed and Parallel Databases track).
- International Conference on Cooperative Information Systems (CoopIS) 2006: P. Valduriez.
- Int. Conf. on High Performance Computing for Computational Science (VecPar) 2006: P. Valduriez.
- Int. Workshop on High-Performance Data Management in Grid Environments (HPDGrid 2006), co-located with VecPar 2006: P. Valduriez (co-chair with M. Mattoso, UFRJ), E. Pacitti (PC chair), G. Raschia.
- Journées Bases de Données Avancées (BDA), 2006: N. Mouaddib, E. Pacitti, G. Raschia.
- Conférence sur la Recherche d'Information et Applications (CORIA), 2006: J. Martinez, N. Mouaddib.
- Int. Conf. on Enterprise Information Systems (ICEIS), 2006: J. Bézivin.
- Enterprise Distributed Object Computing (EDOC), 2006: J. Bézivin.
- Fundamental Approaches to Software Engineering (ETAPS/FASE), 2006: J. Bézivin.
- Int. Conf. on Flexible Query-Answering systems (FQAS), 2006: N. Mouaddib, G. Raschia.

9.3. Invited Talks

J. Bézivin was keynote speaker at the LDTA conference in Vienna. M. Gelgon gave a talk in october at Nokia Multimedia Products, Helsinki, in October. In July, E. Pacitti and P. Valduriez gave invited talks on the APPA P2P system at UFRJ, Rio de Janeiro.

9.4. Teaching

All the members of the Atlas project-team teach database management, multimedia, and software engineering at the Bs, Ms and Ph.D. degree level at the University of Nantes.

The book Principles of Distributed Database Systems, co-authored with professor Tamer Özsu, U. Waterloo, published by Prentice Hall in 1991 et 1999 (2nd edition) has become the standard book for teaching distributed databases all over the world. Our Web site features course material, exercises, and direct communication with professors.

10. Bibliography

Major publications by the team in recent years

[1] R. Akbarinia, V. Martins, E. Pacitti, P. Valduriez. *Global Data Management*, R. Baldoni, G. Cortese, F. Davide (editors). , IOS Press, 2006.

[2] S. Gançarski, H. Naacke, E. Pacitti, P. Valduriez. *The Leganet System. Freshness-Aware Transaction Routing in a Database Cluster*, in "Information Systems", to appear, vol. 32, n^{o} 2, 2007, p. 320-343.

[3] M. GELGON, P. BOUTHEMY, J.-P. LE CADRE. *Recovering and Associating the Trajectories of Multiple Moving Objects in an Image Sequence with a PMHT Approach*, in "Image and Vision Computing", vol. 23, n^{o} 1, 2005, p. 19-31.

[4] E. LOISANT, J. MARTINEZ, H. ISHIKAWA, K. KATAYAMA. *Galois Lattices as a Classification Technique for Image Retrieval*, in "IPSJ Transactions on Data", n^{o} 28, 2005.

[5] E. PACITTI, C. COULON, P. VALDURIEZ, T. ÖZSU. *Preventive Replication in a Database Cluster*, in "Distributed and Parallel Databases", vol. 18, n^{o} 3, 2005, p. 223-251.

[6] A. PIGEAU, M. GELGON. *Building and Tracking Hierarchical Partitions of Image Collections on Mobile Devices*, in "ACM Multimedia conference, Singapore", 2005.

[7] P. PUCHERAL, L. BOUGANIM, P. VALDURIEZ, C. BOBINEAU. *PicoDBMS : Scaling down database techniques for the Smartcard*, in "The VLDB Journal, special issue on Best Papers from VLDB 2000", vol. 10, n^{o} 2-3, 2001.

[8] R. SAINT-PAUL, G. RASCHIA, N. MOUADDIB. *General Purpose Database Summarization*, in "Int. Conf. on Very Large Databases (VLDB 2005), Trondheim, Norway", 2005, p. 733–744.

[9] W. A. VOGLOZIN, G. RASCHIA, L. UGHETTO, N. MOUADDIB. *Querying a Summary of Database*, in "Journal of Intelligent Information Systems (JIIS)", vol. 26, n^{o} 1, 2006, p. 59–73.

[10] T. ÖZSU, P. VALDURIEZ. *Principles of Distributed Database Systems, 2nd edition*, Prentice Hall, 1999.

Year Publications

Doctoral dissertations and Habilitation theses

[11] C. COULON. *Réplication préventive dans une grappe de bases de données*, Ph. D. Thesis, Université de Nantes, September 2006.

[12] F. JOUAULT. *Contribution à l'étude des langages de transformation de modèles*, Ph. D. Thesis, Université de Nantes, September 2006.

[13] L. NAOUM. *Un modèle multidimensionnel pour un processus d'analyse en ligne de résumés flous*, Ph. D. Thesis, Université de Nantes, October 2006.

Articles in refereed journals and book chapters

[14] R. AKBARINIA, V. MARTINS, E. PACITTI, P. VALDURIEZ. *Global Data Management*, R. BALDONI, G. CORTESE, F. DAVIDE (editors). , IOS Press, 2006.

[15] R. AKBARINIA, E. PACITTI, P. VALDURIEZ. *Reducing Network Traffic in Unstructured P2P Systems Using Top-K Queries*, in "Distributed and Parallel Databases", vol. 19, n^{o} 2, 2006, p. 67-86.

[16] M. BOUGHANEM, S. CALABRETTO, J.-P. CHEVALLET, J. MARTINEZ, L. LECHANI-TAMINE. *Un nouveau passage à l'échelle en recherche d'information ?*, in "Ingéniérie des Systèmes d'Information (ISI'06), RTSI série ISI-NIS", 2006.

[17] J.-M. FAVRE, J. BÉZIVIN, I. BULL. *Evolution, Rétro-ingénierie, et IDM*, Hermès Sciences, 2006.

[18] D. FAYE, G. NACHOUKI, P. VALDURIEZ. *Un système Pair-à-Pair de médiation de données*, in "Revue Africaine de la Recherche en Informatique et Mathématique Appliquées (ARIMA)", vol. 4, 2006.

[19] C. FURTADO, A. LIMA, E. PACITTI, P. VALDURIEZ, M. MATTOSO. *Adaptive Hybrid Partitioning for OLAP Query Processing in a Database Cluster*, in "Int. Journal of High Performance Computing and Networking", Special Issue on Best Papers from SBAC2005, to appear, 2007.

[20] M. GELGON, R. HAMMOUD. *Building object-based hyperlinks in videos*, R. HAMMOUD (editor). , Springer, 2006.

[21] P. LAMARRE, S. LEMP, S. CAZALENS, P. VALDURIEZ. *A Flexible Mediation Process for Large Distributed Information Systems*, in "Int. Journal of Cooperative Information Systems", to appear, 2007.

[22] E. LOISANT, J. MARTINEZ, H. ISHIKAWA, M. OHTA, K. KATAYAMA. *Galois Lattices as a Classification Technique for Image Retrieval*, in "IPSJ Digital Courier", vol. 2, 2006, p. 1–13.

[23] J. ROUGUI, M. GELGON, D. ABOUTAJDINE, N. MOUADDIB, M. RZIZA. *Organizing Gaussian mixture models into a tree for scaling up speaker retrieval*, in "Pattern Recognition Letters", to appear, 2007.

[24] W. A. VOGLOZIN, G. RASCHIA, L. UGHETTO, N. MOUADDIB. *Querying a Summary of Database*, in "Journal of Intelligent Information Systems (JIIS)", vol. 26, n^{o} 1, 2006, p. 59–73.

Publications in Conferences and Workshops

[25] R. AKBARINIA, V. MARTINS. *Data Management in the APPA P2P System*, in "Int. Workshop on High-Performance Data Management in Grid Environments (HPDGRID)", 2006.

[26] R. AKBARINIA, V. MARTINS, E. PACITTI, P. VALDURIEZ. *Top-k Query Processing in the APPA P2P System*, in "Int. Conf. on High Performance Computing for Computational Science (VecPar)", Springer, 2006.

[27] R. AKBARINIA, E. PACITTI, P. VALDURIEZ. *Top-k Query Processing in DHTs*, in "Journées Bases de Données Avancées (BDA)", 2006.

[28] J. BÉZIVIN, I. KURTEV. *Model-based Technology Integration with the Technical Space Concept*, in "Metainformatics Symposium", 2006.

[29] J. BÉZIVIN, F. BUTTNER, M. GOGOLLA, F. JOUAULT, I. KURTEV, A. LINDOW. *Model Transformations? Transformation Models!*, in "Int. Conf. on Model Driven Engineering Languages and Systems (MoDELS)", LNCS 4199, Springer, 2006.

[30] J. BÉZIVIN. *Model Driven Engineering: An Emerging Technical Space*, in "Generative and Transformational Techniques in Software Engineering, International Summer School, GTTSE 2005, Braga, Portugal, July 4-8, 2005. Revised Papers", R. LÄMMEL, J. SARAIVA, J. VISSER (editors). , LNCS, vol. 4143, Springer, 2006.

[31] J. BÉZIVIN. *sNets: A First Generation Model Engineering Platform*, in "Satellite Events at the MoDELS 2005 Conference", LNCS 3844, Springer, 2006, p. 169-181.

[32] N. DESSAIGNE, J. MARTINEZ. *A Model for Describing and Annotating Documents*, in "Information Modelling and Knowledge Bases", Y. KIYOKI, J. HENNO, H. JAAKKOLA, KANGASSALO (editors). , Frontiers in Artificial Intelligence and Applications, vol. 136, IOS Press, 2006.

[33] M. DIDONET DEL FABRO, P. VALDURIEZ. *Semi-automatic Model Integration using Matching Transformations and Weaving Models*, in "ACM Symposium of Applied Computing (SAC)", to appear, 2007.

[34] M. DIDONET DEL FABRO, J. BÉZIVIN, P. VALDURIEZ. *Model-driven Tool Interoperability: an Application in Bug Tracking*, in "Int. Conf. on Ontologies, DataBases, and Applications of Semantics (ODBASE)", 2006.

[35] D. FAYE, G. NACHOUKI, P. VALDURIEZ. *Intégration de données hétérogènes dans SenPeer*, in "Colloque Africain sur la Recherche en Informatique (CARI)", 2006.

[36] M. GELGON, A. PIGEAU, A. NIKSERESHT. *Suivi de partitions géo-temporelles à partir d'une divergence de Kullback-Leibler modifiée en vue de la navigation dans une collection d'images personnelles*, in "Journées CORESA'2006 (COmpression et REpresentation des Signaux Audiovisuels", 2006.

[37] F. JOUAULT, J. BÉZIVIN, C. CONSEL, I. KURTEV, F. LATRY. *Building DSLs with AMMA/ATL, a Case Study on SPL and CPL Telephony Languages*, in "ECOOP Workshop on Domain-Specific Program Development (DSPD)", 2006.

[38] F. JOUAULT, J. BÉZIVIN. *KM3: a DSL for Metamodel Specification*, in "IFIP Int. Conf. on Formal Methods for Open Object-Based Distributed Systems", LNCS 4037, Springer, 2006.

[39] F. JOUAULT, I. KURTEV. *On the Architectural Alignment of ATL and QVT*, in "ACM Symposium on Applied Computing (SAC)", 2006.

[40] F. JOUAULT, I. KURTEV. *Transforming Models with ATL*, in "Satellite Events at the MoDELS 2005 Conference", LNCS 3844, Springer, 2006.

[41] Y. KURTEV, J. BÉZIVIN, F. JOUAULT, P. VALDURIEZ. *Model-based DSL Frameworks*, in "Onward! track, Companion of the ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications (OOPSLA)", ACM, 2006.

[42] I. KURTEV, K. VAN DEN BERG, F. JOUAULT. *Rule-based Modularization in Model Transformation Languages illustrated with ATL*, in "ACM Symposium on Applied Computing (SAC)", 2006.

[43] V. MARTINS, R. AKBARINIA, E. PACITTI, P. VALDURIEZ. *Reconciliation in the APPA P2P System*, in "IEEE Int. Conf. on Parallel and Distributed Systems (ICPADS)", 2006.

[44] V. MARTINS, E. PACITTI, R. JIMENEZ-PERIZ, P. VALDURIEZ. *Scalable and Available Reconciliation on P2P Networks*, in "Journées Bases de Données Avancées (BDA)", 2006.

[45] V. MARTINS, E. PACITTI. *Dynamic and Distributed Reconciliation in P2P-DHT Networks*, in "European Conf. on Parallel Computing (Euro-Par)", LNCS 4128, Springer, 2006.

[46] V. MARTINS, E. PACITTI, P. VALDURIEZ. *A Dynamic Distributed Algorithm for Semantic Reconciliation*, in "Int. Workshop on Distributed Data and Structures (WDAS), Records of the 7th Int. Meeting", Carleton Scientific, 2006.

[47] B. MIRANDA, A. LIMA, P. VALDURIEZ, M. MATTOSO. *Apuama: Combining Intra-query and Inter-query Parallelism in a Database Cluster*, in "Currents Trends in Database Technology – EDBT 2006", LNCS 4254, Springer, 2006, p. 649-661.

[48] N. MOUADDIB, J. MARTINEZ. *Résumé de bases de données et application aux données multimédias*, in "Rencontres Inter-Associations (AFIA, ARIA, EGC, INFORSID, SFC, SFDS, LMO, ASTI) : La classification et ses applications (RIAs)", 2006.

[49] L. NAOUM. *Représentation de résumés de base de données par prototypes flous*, in "14es Rencontres Francophones sur la Logique Floue et ses Applications (LFA)", 2006.

[50] L. NAOUM, G. RASCHIA, N. MOUADDIB. *Towards On-Line Analytical Processing for Database Summaries: The Core Algebra*, in "IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE)", 2006.

[51] A. NIKSERESHT, M. GELGON. *Agrégation légère de mélange de lois gaussiennes en vue de l'indexation multimédia répartie*, in "Congrès Reconnaissance des Formes et Intelligence Artificielle (RFIA 2006)", 2006.

[52] A. NIKSERESHT, M. GELGON. *Decentralized Distributed Learning of a Multimedia Class for Content-based Indexing*, in "Euromicro Conference on Parallel, Distributed and Network-based Processing (PDP)", IEEE Computer Society, 2006.

[53] Q.-K. PHAM, N. MOUADDIB, G. RASCHIA. *Data Stream Synopsis Using SaintEtiQ*, in "Int. Conf. on Flexible Query Systems (FQAS)", LNCS 4027, Springer, 2006.

[54] A. PIGEAU, M. GELGON. *Construction et suivi d'une hiérarchie de partitions géo-temporelles pour les collections d'images sur appareils mobiles*, in "Congrès Reconnaissance des Formes et Intelligence Artificielle (RFIA)", 2006.

[55] A. PIGEAU, A. NIKSERESHT, M. GELGON. *Fast tracking of hierarchical partitions with approximate KL-divergence for geo-temporal organization of personal images*, in "ACM Symposium of Applied Computing (SAC)", to appear, 2007.

[56] J. QUIANÉ-RUIZ, P. LAMARRE, P. VALDURIEZ. *Satisfaction Based Query Load Balancing*, in "Int. Conf. on Cooperative Information Systems (CoopIS)", 2006.

[57] J. ROUGUI, M. GELGON, M. RZIZA, J. MARTINEZ, D. ABOUTAJDINE. *Fast incremental clustering of Gaussian mixture speaker models for scaling up retrieval in on-line broadcast*, in "IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP 2006)", 2006.

[58] J. ROUGUI, M. GELGON, M. RZIZA, J. MARTINEZ, D. ABOUTAJDINE. *Hierarchical clustering of mixture models for scaling up speaker recognition*, in "ACM Symposium on Applied Computing (SAC)", 2006.

[59] L. UGHETTO, W. A. VOGLOZIN, N. MOUADDIB. *Personalized Database Querying using data summaries*, in "IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)", 2006.

Internal Reports

[60] D. D. RUSCIO, F. JOUAULT, I. KURTEV, J. BÉZIVIN, A. PIERANTONIO. *A Practical Experiment to Give Dynamic Semantics to a DSL for Telephony Services Development*, Technical report, n^{o} 06.03, LINA, 2006.

[61] D. D. RUSCIO, F. JOUAULT, I. KURTEV, J. BÉZIVIN, A. PIERANTONIO. *Extending AMMA for Supporting Dynamic Semantics Specifications of DSLs*, Technical report, n^{o} 06.02, LINA, 2006.

Miscellaneous

[62] J. MANJARREZ, J. MARTINEZ, P. VALDURIEZ. *High-dimensional Data Allocation in a Shared-nothing Cluster (on-going research)*, April 2006, Réunion de travail de l'ACI MDP2P, LINA, Université de Nantes.

[63] J. MARTINEZ. *Classification et parallélisme pour une recherche efficiente par le contenu*, April 2006, Réunion de travail de l'ACI MDP2P, LINA, Université de Nantes.