INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Project-Team ATLAS

Complex Data Management in Distributed Systems

Rennes

THEME SYM

Activity Report

2004

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]

Faculty Members, University of Nantes

Marc Gelgon [Associate Professor]
José Martinez [Professor]
Noureddine Mouaddib [Professor]
Esther Pacitti [Associate Professor]
Guillaume Raschia [Associate Professor]

Technical Staff

Freddy Allilaire [Engineer, CDD Modelware, since nov. 2004]
Grégoire Dupé [Engineer, CDD Carol, until march 2004]
Gaëtan Gaumer [Engineer, CDD MDP2P]
Peter Rosenthal [Engineer, CDD Modelware, since sept. 2004]

Administrative Assistant

Marie Pivaut [Secretary, CDD]

Ph.D. Students

Reza Akbarinia [Fellowship Iranian gov.,since oct. 2003]
Alexandre de Assis Bento Lima [Fellowship CAPES, Brazil, until june 2004]
Cédric Coulon [Fellowship INRIA]
Marcos Del Fabro Didonet [Fellowship Microsoft, since sept. 2004]
Frédéric Jouault [Fellowship CIFRE, TNI]
Erwan Loisant [Fellowship Monbusho, Japan]
Jorge Manjarrez Sanchez [Fellowship Conacyt, Mexico, since sept. 2004]
Vidal Martins [Univ. PUCPR, Brazil]
Antoine Pigeau [Fellowship MENRT]
Jorge Quiane Ruiz [Fellowship Conacyt, Mexico, since sept. 2004]
Régis Saint-Paul [Fellowship MENRT]
Stéphane Tayeb [Fellowship MENRT]
Anthony Ventresque [Fellowship MENRT, since oct. 2004]
Andre Victor [Fellowship CNPQ, Brazil]
Amenel Voglozin [Fellowship MENRT]

2. 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 (such as the Internet or cluster 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 group 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. 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.

3. 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 firsts-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), possibly very large, and described by metadata. Multimedia data management aims at providing high-level capabilities for searching and manipulating multimedia collections efficiently and accurately. To address this objective, we rely on two kinds of techniques: multimedia data analysis and database techniques [5].

Multimedia data analysis is useful to automatically translate multimedia data into sets of features which can be used for indexing and searching. The features can be low-level or higher-level. Low-level features, e.g. color, texture and shape for images, can be obtained from signal processing techniques, e.g., colour histograms for images. Mid- to high-level features are also useful to understand the media content using more complex analysis techniques such as statistical classification or pattern recognition. In addition, media content creators can add metadata information that conveys more semantics.

Database techniques are useful to bridge the semantic gap between the media descriptions that can be automatically produced by analysis and the user requirements to search multimedia data in a natural way. Applying the fundamental principle of data abstraction to multimedia yields new theories and techniques for modeling, indexing and querying multimedia data.

Multimedia modeling is concerned with the definition of a suitable representation for media data, their metadata and the operations to be applied to them. A multimedia data model introduces an abstraction between the physical level (data files and indexes) and the conceptual representation, together with the operations to manipulate the data. An important capability is to capture relationships between content and metadata which is best done using objects and objet references. Thus, most multimedia models use or extend a relational or object-oriented data model.

Indexing is concerned with the physical access to multimedia data. The aim of indexes is to rapidly access the data requested by the query. Multimedia features such as average colors, color histograms and textures are usually modelled as points in a multidimensional space. Thus, multimedia can be efficiently indexed using multidimensional index structures, also called spatial access methods, which organize the multidimensional points as buckets, each corresponding to a disk block. There are two categories of such index structures: tree-based and hashing-based. The major issue faced by these methods is that multimedia features usually have a high number of dimensions. Thus, they suffer from the "dimensionality curse" which says that the performance of indexing (and thus querying) degrades as the data dimensionality increases.

Querying is concerned with the conceptual access to multimedia by the user with a high-level (SQL-like) query language. The common query in multimedia is similarity search where the objects retrieved are ordered according to some scores based on a distance function defined on a feature vector. In the presence of indexes, a similarity query involving two or more features must be decomposed into subqueries whose results need be finally integrated. In the case of image databases, a useful alternative to querying is browsing. By organizing an image collection in a suitable way, e.g. as a Galois' lattice, browsing allows for more interactive and iterative searching.

3.1.3. Fuzzy Logic

The ever growing size of databases makes data summarization needed in order to present the user a concise, yet complete view of the data. Our proposed summarization process [9] can roughly been described as a two step process. The first step is to rewrite the original database data into an homogeneous user-oriented vocabulary. The second step is then to use a concept formation algorithm against the rewritten data. The fuzzy-set theory provides a mathematical foundation to handle these two steps in a more efficient 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 expresses 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 term 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 should the hierarchy 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 algorithm.

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 this operation. 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 permit to handle user queries expressed in vague or imprecise terms.

3.1.4. Model Engineering

A model is a formal description of a design artifact 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 [12]. 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 group 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) [11]. 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

Keywords: *Application Service Provider (ASP)*, *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 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 exploiting data replication and load balancing techniques.

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. Such domestic applications, extended to the video medium mainly, appear as a natural perspective for future TV sets. Currently, the integration of multimedia is effective only with images. >From the usability point of view, open issues are the effective combination of various medias and the adaptability of the indexing process to a specific task or application domain.

5. Software

5.1. ATL (Atlas Transformation Language)

Participants: Jean Bézivin, Frédéric Jouault, Patrick Valduriez.

ATL is a transformation-based model management framework, with metadata management and data mappings as the main application. The ATL language is designed to be general and abstract. We plan to 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 being implemented in Java, and we plan to port major transformation components to the .Net platform. End of 2004, we registered (together with TNI-Software and University of Nantes) the model management system and its components for metadata management to the APP (Agence pour la Protection des Programmes) and released it as Open Source Software under the Eclipse Public Licence (see http://www.eclipse.org/gmt/).

5.2. $Find_{AGE}^{Im}$

Participants: José Martinez, Erwan Loisant.

$Find_{AGE}^{Im}$ is an image search-by-content system. Currently, it is fairly complete from the architectural point of view. It provides three different ways to query images by content: formal querying, interactive querying and browsing. Formal querying is based on the traditional querying approach developed for structured DBMSs. The interactive querying process of $Find_{AGE}^{Im}$ is based on the information retrieval querying process, i.e., when manipulating noisy data it is hardly possible to write down immediately the correct query, if at all. Browsing is a more efficient and effective way to retrieve rapidly visual information such as images. Currently, a first complete implementation variant of this architecture is on the way to be achieved.

5.3. RepDB*

Participants: Cédric Coulon, Gaëtan Gaumer, Esther Pacitti, Patrick Valduriez.

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 furthere 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.

5.4. SaintEtiQ

Participants: Noureddine Mouaddib, Gaëtan Gaumer, Guillaume Raschia, Régis Saint-Paul, Amenel Voglozin.

SAINTETIQ is a data summarisation system providing 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.

The last version of SAINTETIQ has been developed as a Web service. It is accessible at the following address: http://www.simulation.fr/seq.

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. Our recent work focused on the querying of database summaries and the integration of multiple summary sources.

6.1.1. Querying SaintEtiQ summaries

Participants: Noureddine Mouaddib, Guillaume Raschia, W. Amenel Voglozin.

For some years, data summarization techniques have been developed to handle the growth of databases. However these techniques are usually not provided with tools for end-users to efficiently use the produced summaries. As a consequence, users have to directly interpret the summaries, which is reasonable with a few summaries only. In the general case, summarizing a very large amount of data (which is the aim of SAINTETIQ), leads to a large number of summaries and thus requires specific tools.

To address this issue, we proposed a querying tool to exploit the hierarchical tree of summaries produced by our summarization model [46][47]. On one hand, our querying mechanism can be seen as a boolean one, using the vocabulary of the linguistic variables (fuzzy subsets of the attribute domains) defined in the summarization process, and giving anwers with the same vocabulary. It takes advantage of the hierarchical structure of the summaries, which acts as a general multidimensionnal index, to improve response times. On the other hand, as the tuples from the original database can be retrieved from the summaries,the tool is flexible: the query is expressed using linguistic variables, while the answer is tuples associated with satisfaction degrees.

Another trend of research focuses on cooperative querying in case of null answers. Indeed, in case of null answers, the hierarchical structure of the summaries not only allows for very good response times, but also gives immediate information on why the answer is null, and how to find close answers.

The querying process is based on a depth-first search and relies on a property of the hierarchy: the generalization step in the SAINTETIQ model guarantees that any descriptor that exists in a node of the tree also exists in each parent node. Inversely, a descriptor is absent from a summary's intension if and only if it is absent from all subnodes of this summary. This property of the hierarchy permits branch cutting as soon as it is known that no result will be found. Depending on the query, only a part of the hierarchy is explored. In any case, all relevant results, and only relevant results, are captured.

The flexible querying process we propose builts a ranking of results. The summaries hierarchy can be considered as a general index over the data, which can be used to quickly reduce the search space. Thus, a first step (the proposed querying tool) gives answer summaries (leaves of the hierarchy) to a query expressed with linguistic labels. Then a second step retrieves tuples from the summaries extension. Finally, the degrees attached to each tuple can be taken into account to compute a total satisfaction degree of the query.

Besides satisfaction degrees, the SaintEtiq model provides frequency and proportion data attached to descriptors and candidate tuples. Thus, taking advantage of the SaintEtiq summarization process, our tool can perform flexible querying by applying classical operators on the summaries, i.e., on fuzzy data. This is the main distinctive feature of our approach with respect to other flexible querying mechanisms such as $Fquery, SQLF, SummarySQL$.

6.1.2. Integration of multiple summary sources

Participants: Noureddine Mouaddib, Guillaume Raschia, Régis Saint-Paul, Stéphane Tayeb.

Summarization algorithms in SAINTETIQ only operate on structured data organized in large flat tables, regardless of the database structure. A new research direction has then emerged, with the ultimate goal of being able to merge different summaries from distinct data sources in a mediation environment. The first step towards this goal is to propose new algorithms to build global summaries (short materialized views) from local summary sources.

In this context, we have extended our flat file-oriented approach to take into account the schema of the database. Based on the works on database schema matching and data merging in mediation architectures, we defined an original summary hierarchy join operator, which provides an efficient way to exploit referential integrity constraints over the database schema. This usefull operator provides a single global summary hierarchy from two distinct relational tables and will serve as a starting point towards the integration of multiple summary sources.

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. 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 or C. As information systems become more complex and need to support cooperation of heterogeneous applications and components, such manual development of models and transformations is no longer viable.

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 Latex to 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 a more general architecture named 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, Peter Rosenthal, Patrick Valduriez.

Model transformation, e.g. mapping a relational database schema into an XML schema, is probably the most useful and important operation in model management. We have proposed ATL, a combined declarative/imperative language that allows to transform a source model into a target model. The transformation program is itself a model and is based on a given metamodel, like the source and target models, corresponding to the general conceptual unifying scheme. In addition to the language definition, a proof of concept has been proposed with a first implementation on the Sun MDR/NetBeans environment. A second version of the ATL engine has been realized on top of the IBM Eclipse/EMF as GMT open source. In addition to the ATL engine, a complete integrated development environment (IDE) has been built and released as GMT open source. The IDE for ATL allows transformation editing and debugging (syntax coloring, step by step execution, breakpoints, environment observation, etc.) [26] Until now, we have concentrated on the initial definition of the

ATL language and the implementation of its execution engine in collaboration with TNI-Software. Several significant examples of model transformations have been realized, like a XSLT to XQuery transformer.

This initial work is being completed by a conceptual and practical validation of the language. One issue is transformation reusability. This engineering problem has not received yet any satisfactory answer. In the context of the ModelWare European project, we collaborate with SINTEF, Norway, in applying ATL to several case studies.

6.2.2. Model weaving

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

In order to get a regular, minimal and powerful model management platform able to answer our research objectives in data engineering, it is necessary to be able to consider models as first class entities and to be able to store retrieve, access, edit, and manage these models as uniformly as possible. Then we need to consider model transformations as a special kind of models. All this is necessary but remains however insufficient. >From our first experimentations with automatic transformations, we came to the conclusion that we need also to get a complementary model weaving operation, with different properties of arity, automaticity and variability. Weaving is related to transformation but cannot be directly expressed as a composition of transformations. In particular, we must be able to realize the merge operator proposed by Bernstein (Microsoft Research), which is useful in many data management problems such as schema integration. To reach this goal, we need to be able to weave two models or two metamodels. A weaving session produces a weaving model, based on a specific weaving metamodel. This weaving model may be used later to generate an ATL transformation model, but this is only one among many possible use cases.

We consider this model weaving operation as an additional and complementary effort to model transformation. But an important experimental work remains to be done in this area. This is why we have undertaken the building of a model weaving workbench named AMW (ATLAS Model Weaver). We have shown [31] that the model weaving operation is not only of wide applicability in data management (schema matching, etc.), but that it may help solving important problems in software engineering as well. For instance, the so called "Y"-shaped software development cycle to build a software system can be solved by a merge from a business model and a platform model.

6.2.3. Global model management

Participants: Freddy Allilaire, Jean Bézivin, Frédéric Jouault, David Touzet, Patrick Valduriez.

Within a model management environment, the main elements produced or consumed are models, metamodels 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. technical spaces have also to be considered. Special attention should be given to the global management of all these resources. We view a tool as a set of functions taking models as input and producing models as output. These models are explicitly typed by their corresponding metamodels allowing to define the signature of each tool. A platform has the precise and updated knowledge about all the connected tools, each one being characterized by its signature.

All the information about the components known to a given platform is stored in a specific model named "megamodel". A megamodel is some kind of model registry storing reference and metadata information on all accessible resources, including relations between these resources. This notion 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 very well with the general conceptual scheme developed for model management. Experimental validation is being done through the AM3 (ATLAS MegaModel Manager) tool allowing to record and control the global relations betwen model components [32].

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. We have been working on the following forms of multimedia data : image, audio, video, and geo-temporal metadata possibly attached to a document.

6.3.1. Browsing Image databases

Participants: José Martinez, Erwan Loisant.

In our previous works, we showed that Galois' lattices are a useful tool to access image database by browsing rather than querying them. They suffered from a performance problem, the time complexity of building a lattice being quadratic. This has been solved thanks to a pre-processing step using our summarisation tool. Nonetheless, Galois' lattices suffer an additional problem: all the descriptions are orthogonal, which means that the number of incoming and out-going edges of a node can be extremely high, hence a visualisation issue for large image databases. (This problem is somewhat alleviated thanks to the summarisation process.)

In order to solve this second problem for large databases, we have proposed techniques aimed to masking parts of the Galois' lattice during a browsing session [41][20]. In other words, we create dynamic views of the Galois' lattice consisting of coherent sub-lattices obtained either by removing nodes, or edges. In the case of node masking (the investigated case, so far), the algorithm exhibits a linear complexity in the number of masked nodes. When using a lattice of clusters of images, the size of the whole lattice is "only" in $O(\sqrt{n})$, where n is the number of images. This means that this masking technique can be used at run-time even for large databases.

6.3.2. Scalability in image database management systems

Participants: José Martinez, Patrick Valduriez.

An image DBMS should preferably use standard DBMS technology. Even if few very large image sets exist today, e.g., Corbis™ with over 40 millions of images, using multimedia data generates huge volumes of data [49]. We started by evaluating the actual limits of relational DBMSs for managing image metadata (from 100 up to 100 000 images) [42]. Therefore, the challenge is to remain efficient under such heavy conditions. We have identified the following key points:

- First, reducing the initial size of metadata is extremely important. It impacts the space requirements, hence the time complexity. Several transforms exist in the literature.
- Secondly, materialising some redundant information helps to improve queries, especially when the derived information is obtained through a complex formulation, which the database optimiser cannot recognise. A particular case is fuzzy linguistic variables.
- Next, parallelism is unavoidable *in fine*, as an additional improvement when all the other techniques have been used. Nevertheless, it has to be carefully implemented. In our first cluster-based experiments, it was not so complex a trade-off between the local processing runs and the global processing of the final result set.
- Also, distributing images on the nodes of a (cluster) network, based on their visual properties has to be taken into account. Images have to be clustered and assigned to various subsets of the distributed architecture. In addition to a more complex trade-off, which remains to be evaluated, indexing will turn out to be differentiated.
- Finally, we have to take advantage of multi-dimensional indexing techniques, even limited ones, whenever possible. This has to be combined with the specialised subsets of image metadata.

6.3.3. Structuring and querying audio databases

Participants: Marc Gelgon, José Martinez.

In a way, audio data is harder to manipulate than images, from the user's point of view, for it cannot be visualised. Therefore, we started modelling it in order to *query* it, through a standard query language, rather than to browse it [21]. From the modelling point of view, we limited our interest to speaker recognition under no adversary conditions. More precisely, we are interested in indexing mainly spoken radio programmes.

Since audio tracks are temporal data, we are facing a querying language issue which has been solved by introducing an abstract data type for audio tracks with temporal operators. These operators are interval extractions, concatenations, and their generalisation. Additional operators are required to allow the writing of complex queries, namely Allen's predicates and list-based operators. Then, the system can be implemented in an object-oriented DBMS, or less efficiently in an object-relational DBMS.

While the audio stream is partitioned into speaker-homogeneous segments, the database is fed with speaker models (new or updated). Speakers are represented by probability distributions (Gaussian mixtures) over a MFCC (mel-frequency cepstral coefficients) feature space. We observe that the approach leads to a quite accurate structuring of radio news programmes, but it does not yet scale up well to a long, continuous audio stream. It is the goal of current work.

6.3.4. Personal image collection management from mobile devices

Participants: Marc Gelgon, Antoine Pigeau.

Extension of image retrieval systems to address personal image collections appears among emerging needs. In particular, mobile devices such as camera-equipped phones are an interesting case for content creation and retrieval. In this context, we have proposed a unsupervised technique for organising an image collection, based on time and geo-location metadata [23][44][43]. The objective is to recover the natural spatial and temporal structure present in such a data set. This metadata is indeed both reliable and appealing for spatio-temporal browsing, as shown by studies on user requirements. Our proposal is formulated as statistical mixture model-based classification. Dedicated optimality criterion and expectation-maximisation search technique have been devised. An incremental version of the scheme has been defined, as content creation and retrieval are not consecutive in this application. Since the data sometimes exhibits spatial or temporal clusters at multiple scales, an extension of the scheme attempting to recover such a structure is being designed. Expected benefits are two-fold : convenience of user interaction for browsing the document collection, and opportunities to improve the system-level efficiency, especially on a moderately powerful mobile platform.

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 2004, we improved our work on data management in database clusters and pursued three research actions on preventive data replication, transaction load balancing and OLAP query processing. We also started the design of a new system for P2P data management called Atlas Peer-to-Peer Architecture (APPA).

6.4.1. Preventive data replication in a database cluster

Participants: Cédric Coulon, Esther Pacitti, Patrick Valduriez.

Database clusters provide a cost-effective alternative to parallel database systems, i.e. database systems implemented on tighly-coupled multiprocessor computers. A database cluster is a cluster of PC servers, each having its own processor(s) and hard disk(s), and running a "black-box" DBMS. Using a "black-box" DBMS at each node has the advantages of preserving the autonomy of the databases and avoiding expensive data migration (for instance to a parallel DBMS). Database clusters can make new businesses such as Application Service Providers (ASP) economically viable. In the ASP model, customers' databases (including data and DBMS) are hosted at the provider site and should not require any expensive migration.

To obtain high-performance and high-availability in a database cluster, we replicate databases (and DBMSs) at several nodes, so they can be accessed in parallel through applications. Then the main problem is to assure the consistency of autonomous replicated databases. A good solution that scales up is multi-master lazy replication (also called update everywhere). It provides high-availability and high-performance since replicas can be updated in parallel at different nodes. However, conflicting updates at different nodes can introduce replica divergence.

To address this problem, we proposed a new preventive replication strategy that can avoid the occurrence of conflicts, by exploiting the cluster's high speed network, thus providing strong consistency. In [34], we address the limitations of the original solution by providing support for partial replication, where databases are partially replicated at different nodes. Unlike full replication, partial replication can increase access locality and reduce the number of messages for propagating updates to replicas. We also proposed an optimisation to allow for the concurrent execution of the transactions. We implemented our algorithm in our RepDB* prototype [33] running on top of the PostgreSQL Open Source DBMS and performed extensive experiments over the 32-node Linux cluster of the Paris project-team at IRISA. Our experimental results using the TPC-C Benchmark show that it yields excellent scale-up and speed up.

6.4.2. *Transaction and query load balancing in a database cluster*

Participants: Esther Pacitti, Patrick Valduriez.

In a database cluster, data replication is often used to increase both data availability and performance. This solution has been successfully used by Web search engines using high-volume server farms. However, Web search engines are typically read-intensive which makes it easier to exploit parallelism. When applications are update-intensive, as in an ASP context, maintaining replicas mutually consistent, for instance, using preventive replication, can hurt performance. However, there are important cases where consistency can be relaxed. For instance, read-only queries do not always require reading perfectly consistent data and may tolerate inconsistencies. Thus, an interesting solution is to trade consistency for performance based on users' requirements. In most approaches (including ours), consistency reduces to freshness: update transactions are globally serialised over the different cluster nodes, so that whenever a query is sent to a given node, it reads a consistent state of the database. Global consistency can be achieved using either a preventive approach that avoids conflicts such as [34] or an optimistic approach with conflict detection and reconciliation. However, the consistent state may not be the latest one, since update transactions may be running at other nodes. Then the data freshness of a node reflects the divergence between its actual state and the state it would have if all transactions had already been applied to it.

In [38], we proposed a freshness model which allows users to specify freshness requirements for their queries. This model allows defining conflict classes between queries and transactions. This information can be exploited to optimise query load balancing. We implemented our solution in the Leganet prototype [36] on a 4-node Linux cluster running Oracle™8i. Through experiments with the TPC-R benchmark, we showed that freshness control can help increase query throughput significantly. The results give significant improvement when freshness requirements are specified at the relation level rather than at the database level.

In [22], using our freshness model, we proposed a solution to transaction routing and a cost model to estimate replica freshness. Using the Leganet prototype and using simulation up to 128 nodes, our validation based on the TPC-C OLTP benchmark demonstrates the performance benefits of our approach in comparison with existing solutions. This work is done in cooperation with LIP6.

6.4.3. *OLAP query processing in a database cluster*

Participants: Alexandre de Assis Bento Lima, Patrick Valduriez.

OLAP applications require high-performance database support. They typically access large data sets using heavy-weight read-intensive queries. A simple, yet efficient, solution to OLAP query processing in database clusters is virtual partitioning which yields intra-query parallelism. The main requirement for virtual partitioning to work is that each node has access to the entire database. This can be achieved either in a shared-disk architecture or in a shared-nothing architecture with full replication. Each query submitted to the database cluster is rewritten in a number of queries, one for each node, by adding range predicates that correspond to different virtual partitions of a relation. However, virtual partitioning suffers from two problems that can dramatically hurt performance: (1) virtual partitions which are allocated to nodes can get very large; (2) virtual partitions assuming uniform distribution of the partitioning attribute values yield to unbalanced processing in case of skewed data.

In [40], we proposed a first improvement called fine-grained virtual partitioning which addresses these problems. To validate our solution, we implemented a Java prototype on a 16-node cluster and ran experiments with the TPC-H benchmark. Our experimental results show that linear, and sometimes super-linear, speedup is obtained for typical OLAP queries. In many cases, it outperforms traditional virtual partitioning by a factor of 6. However, we used a simple method for partition size determination based on database statistics which are hard to obtain. In [39], we proposed a simpler, yet efficient, solution called adaptive virtual partitioning technique which dynamically tunes partition sizes, without requiring any knowledge about the database and the DBMS. To validate our solution, we implemented a Java prototype on a 32 node cluster system and ran experiments with typical queries of the TPC-H benchmark. The results show that our solution yields linear, and sometimes super-linear, speed-up. In many cases, it outperforms traditional virtual partitioning by factors superior to 10.

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

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

Peer-to-peer (P2P) systems adopt a completely decentralised approach to data sharing. By distributing data storage and processing across autonomous peers in the network, they can scale up without the need for powerful servers. Recent 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. As a potential example of advanced application that can benefit from a P2P system, consider the cooperation of scientists who are willing to share their private data (and programs) for the duration of a given experiment. In this context, users tend to work in dynamic groups and may join or leave a group as they wish.

Advanced applications are likely to need more general replication capabilities such as various levels of replication granularity and multi-master mode, i.e. whereby the same replica may be updated by several (master) peers. For instance, a patient record may be replicated at several medical doctors and updated by any of them during a visit of the patient, e.g. to reflect the patient's new weight. The advantage of multi-master replication is high-availability and high-performance since replicas can be updated in parallel at different peers. However, conflicting updates of the same data at different peers can introduce replica divergence. A more practical solution is optimistic replication which allows the independent updating of replicas and divergence until reconciliation. However, existing optimistic replication solutions do not address important properties of P2P systems such as self-organisation.

In [25], we address replication and query processing in the context of APPA (Atlas Peer-to-Peer Architecture), a new P2P data management system which we are building. The main objectives of APPA are scalability, availability and performance for advanced applications. APPA has a network-independent architecture 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. To capitalise on Web service standards, the shared data are in XML format (which may be interfaced with many data management systems). In APPA, we propose an optimistic replication solution which provides eventual consistency and other useful properties found in distributed systems. We exploit application semantics to reduce conflicts during reconciliation.

We also propose a schema-based query processing strategy that deals with replication in a way that reduces redundant work and optimises peer load balancing.

7. Contracts and Grants with Industry

7.1. 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.

7.2. IBM/Eclipse (2004-2005)

Participant: Jean Bézivin.

The objective is to port the ATL platform in the Eclipse Open Source environment. This was the only french project granted by IBM Eclipse Educational Grant in 2004. A first version of the prototype has been presented at the OOPSLA conference in october 2004 in Vancouver.

7.3. RNTL Modathèque (2004-2005)

Participants: Jean Bézivin, Patrick Valduriez.

In this project, we work with Thales RT (project leader), France Telecom R&D, LIP6 and software engineering tool vendors in France. The objective is to define the conceptual and practical basis for model engineering, in particular, using components of the Model Driven Environment (MDE) of the OMG. In this project, we use our ATL platform for MDE components.

7.4. Caroll Motor (2003-2006)

Participants: Jean Bézivin, Grégoire Dupé, 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.

7.5. RNTL Domus Videum (2002-2004)

Participants: Marc Gelgon, José Martinez, Noureddine Mouaddib, Guillaume Raschia.

In this project, led by Thomson Multimedia, we worked with SFRS, INA, LaBRI, IRCCyN, IRISA (Metiss, Vista, and Temics groups). The main goal of this project was to provide services for a new generation of home multimedia platforms and more specifically, to develop an intelligent Personal Video Recorder (PVR). This kind of PVR is able to filter TV broadcasts depending on a user profile, and offer, through a user-friendly navigation interface, different points of view over programs (full version, summary, digest, best-of). Our contribution to the project was the task of learning the user profile based on metadata of TV programs, and the filtering process.

7.6. IP Modelware (2004-2006)

Participants: Freddy Allilaire, Jean Bézivin, Peter Rosenthal, 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.

8. Other Grants and Activities

8.1. Regional Actions

The Atlas group participates in the COM project funded by the “Region 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.

8.2. National Actions

We are involved in three projects

8.2.1. ACI MDP2P (2003-2006)

Participants: Gaëtan Gaumer, Esther Pacitti, Patrick Valduriez.

The project MDP2P (Massive data management in peer-to-peer systems) of the ACI Masses of Data of the French ministry of research is led by the Atlas group and involves three other INRIA groups: 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 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.

8.2.2. ACI WebSeXy (2004-2007)

Participants: José Martinez, Noureddine Mouaddib, Guillaume Raschia.

The project WebSeXy (Querying the Semantic Web with XQuery) of the ACI Masses of Data 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.

8.2.3. ACI APMD (2004-2007)

Participants: José Martinez, Noureddine Mouaddib, Guillaume Raschia.

The project APMD (Personalised Access to Masses of Data)(2004-2007) of the ACI Masses of Data 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, partially funded by CNPQ and CAPES in Brazil and COFECUB and INRIA in France, with the universities PUC-Rio and UFRJ, Brazil, on distributed data management;
- 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 participates to the MDA work.

Furthermore, we have regular scientific relationships with research laboratories in

- North America: Univ. of Waterloo (Tamer Özsu), University of California Berkeley (Michael Franklin), MIT (Stuart Madnick), New Jersey Institute of Technology (Vincent Oria), Wayne State University (Farshad Foutouhi and Wiliam Grosky), Kettering University (Peter Stanchev);
- Europe: CWI (Martin Kersten), University of Twente (Mehmet Aksit), University of Roskilde (Henrick Larsen);
- Others: University 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 group 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). P. Valduriez is a member of the scientific committee of the ACI GRID and a member of the ACM SIGMOD steering committee.

J. Bézivin was co-chair of a research seminar in Dagsthul (Germany) in 2004 on model engineering. He is a member and co-founder of the steering committee of the ECOOP (AITO) and UML/Models conferences.

In 2004, the Atlas group organised the ACM SIGMOD/PODS conference in Paris (for the first time outside of North America) [50] with the help of the external relations services of INRIA-Rennes and INRIA-Rocquencourt. P. Valduriez was general chair and J. Martinez web site manager. The conference was a great success with about 650 scientists, practitioners, and students from all over the world.

The Atlas group also organised the 12th “Rencontres Francophones sur la Logique Floue et ses Applications” held at the “École polytechnique de l’université de Nantes”, France in November 2004.

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 (IJUFKS) : N. Mouaddib

Participation in conference programme committees :

- Int. Conf. on Very Large Databases (VLDB), 2005: P. Valduriez
- Int. Conf. on Parallel and Distributed Computing (Euro-Par), 2005: P. Valduriez (vice chair, Distributed and Parallel Databases track)
- Journées Bases de Données Avancées (BDA), 2004: N. Mouaddib
- Journées Bases de Données Avancées (BDA), 2005: N. Mouaddib, E. Pacitti
- Conférence sur la Recherche d'Information et Applications (CORIA), 2005: J. Martinez
- ACM SIGMOD Int. Conf. on the Management of Data (SIGMOD), 2004: E. Pacitti
- ACM Int. Symp. on Applied Computing: Multimedia and Visualisation Track (SAC), 2004: J. Martinez
- Int. Conf. on Enterprise Information Systems (ICEIS), 2005: J. Bézivin
- Int. Conf. on the Unified Modeling Language (UML), 2004: J. Bézivin
- Enterprise Distributed Object Computing (EDOC), 2005: J. Bézivin
- Generative Programming (GPCE), 2004: J. Bézivin
- Fundamental Approaches to Software Engineering (ETAPS/FASE), 2005: J. Bézivin
- Journées Recherche d'Information Assistée par Ordinateur (RIAO), 2004: J. Martinez
- Conférence sur la Recherche d'Information et Applications (CORIA), 2004: J. Martinez
- Rencontres francophones sur la Logique Floue et ses Applications (LFA), 2004: N. Mouaddib
- Flexible Query Answering Systems (FQAS), 2004 : G. Raschia, N. Mouaddib
- Computer Vision meets Databases workshop of ACM SIGMOD (CVDB), 2004: N. Mouaddib

9.3. Invited Talks

In 2004, P. Valduriez was a keynote speaker at the Int. Conf. on High Performance Computing for Computational Science (VecPar2004) in Valence (Spain). He gave an invited talk at UFRJ, Rio de Janeiro, on data management in P2P systems. He also gave a seminar at the summer school on distributed algorithms organised by U. Paris 7 in Porquerolles.

In 2004, E. Pacitti gave invited talks on consistency management in P2P systems at UFRJ and PUC-Rio, Rio de Janeiro. She also gave a seminar on consistency in distributed databases at the Druide summer school organised by INRIA in Port-aux-Rocs.

In 2004, J. Bézivin gave a guest talk at the Working Conference on Reverse Engineering (WCRE) in Delft, Netherlands.

9.4. Teaching

All the members of the Atlas group 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] L. BOUGANIM, D. FLORESCU, P. VALDURIEZ. *Load Balancing for Parallel Query Execution on NUMA Multiprocesseurs*, in "Distributed and Parallel Databases", vol. 7, n° 1, 1999.

[2] J. BÉZIVIN, N. PLOQUIN. *Tooling the MDA framework: a new software maintenance and evolution scheme proposal*, in "Journal of Object-Oriented Programming, JOOP", 2001.

[3] M. GELGON, P. BOUTHEMY. *A region-level motion-based graph representation and labeling for tracking a spatial image partition*, in "Pattern Recognition", vol. 33, n° 4, 2000.

[4] J. MARTINEZ. *Database Browsing*, in "Journal of Research and Practice in Information Technology (JRPIT), Australian Computer Society Inc.", vol. 34, n° 1, 2002.

[5] J. MARTINEZ, N. MOUADDIB. *Multimedia and Databases: A Survey*, in "Networking and Information Systems Journal (NISJ)", vol. 2, n° 1, 1999.

[6] E. PACITTI, P. MINET, E. SIMON. *Replica Consistency in Lazy Master Replicated Databases*, in "Distributed and Parallel Databases", vol. 9, n° 3, 2001.

[7] E. PACITTI, E. SIMON. *Update Propagation Strategies to Improve Freshness in Lazy Master Replicated Databases*, in "The VLDB Journal", vol. 8, n° 3-4, 2000.

[8] 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 VLDB2000", vol. 10, n° 2-3, 2001.

[9] G. RASCHIA, N. MOUADDIB. *A Fuzzy Set-based Approach to Database Summarization*, in "Int. Journal of Fuzzy Sets and Systems", vol. 129, n° 2, 2002.

[10] A. TOMASIC, L. RASCHID, P. VALDURIEZ. *Scaling Access to Heterogeneous Data Sources with DISCO*, in "IEEE Trans. on Knowledge and Data Engineering", vol. 10, n° 5, 1998.

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

Books and Monographs

[12] J. BÉZIVIN. *Ingénierie des modèles Logiciels et Systèmes – Ouvrage ARAGO 30*, éditions Lavoisier, May 2004.

Doctoral dissertations and Habilitation theses

[13] A. DE ASSIS BENTO LIMA. *Equilibrage de charge de requêtes OLAP dans une grappe de bases de données.*, Ph. D. Thesis, UFRJ, Rio de Janeiro, December 2004.

Articles in referred journals and book chapters

[14] B. BENATALLAH, M. HASSAN, N. MOUADDIB, H. PAIK, F. TOUMANI. *Building and Querying an e-Catalog Network Using P2P and Data Summarisation Techniques*, in "International Journal of Intelligent Information Systems (JIIS)", to appear, 2004.

[15] G. BERNARD, J. BEN-OTHMAN, L. BOUGANIM, G. CANALS, B. DEFUDE, J. FERRIÉ, S. GANÇARSKI, R. GUERRAOUI, P. MOLLI, P. PUCHERAL, C. RONCANCIO, P. SERRANO-ALVARADO, P. VALDURIEZ. *Mobile Databases: a report on open issues and research directions*, in "ACM SIGMOD Record", 2004, p. 78–83.

[16] J. BÉZIVIN, E. BRETON. *Applying the basic principles of model engineering to the field of process engineering*, in "Novatica – Special Issue on Software Process Technologies", n° 171, September 2004.

[17] J. BÉZIVIN. *In search of a Basic Principle for Model-Driven Engineering*, in "Novatica – Special Issue on UML (Unified Modeling Language)", n° 168, March 2004.

[18] J. BÉZIVIN. *On the Unification Power of Models*, in "Software and System Modeling (SoSym)", vol. 3, n° 4, 2004.

[19] J. BÉZIVIN. *Sur les principes de base de l'ingénierie des modèles logiciels*, in "L'Objet", colloque en l'honneur de J.-F. Perrot, vol. 10, n° 4, 2004.

[20] E. LOISANT, H. ISHIKAWA, J. MARTINEZ, O. MANABU, K. KATAYAMA. *User-Adaptive Navigation Structures for Image Retrieval*, in "Database Japanese Society Letters", vol. 3, n° 1, 2004.

[21] R. LUTFI, M. GELGON, J. MARTINEZ. *Structuring and Querying Documents in an Audio Database Management System*, in "Multimedia Tools and Applications", vol. 24, n° 2, November 2004, p. 105-123.

[22] H. NAACKE, F. D. NGOC, P. VALDURIEZ. *Routage de Transactions dans un Cluster de Bases de Données Autonomes*, in "Ingénierie des Systèmes d'Information", vol. 9, n^{o} 1, 2004, p. 85–107.

[23] A. PIGEAU, M. GELGON. *Organizing a personal image collection with statistical model-based ICL clustering on spatio-temporal camera phone meta-data*, in "Journal of Visual Communication and Image Representation", vol. 15, n^{o} 3, September 2004, p. 425-445.

[24] T. ÖZSU, P. VALDURIEZ. *Handbook of Computer Science and Engineering, 2nd Edition*, chap. Distributed and Parallel Database Systems, CRC Press, 2004, p. 58-1–58-24.

Publications in Conferences and Workshops

[25] R. AKBARINIA, V. MARTINS, E. PACITTI, P. VALDURIEZ. *Replication and Query Processing in the APPA Data Management System*, in "Int. Workshop on Distributed Data and Structures (WDAS'2004), Lausanne", 2004.

[26] F. ALLILAIRE, T. IDRISSI. *ADT: Eclipse development tools for ATL*, in "Second European Workshop on Model Driven Architecture (EWMDA-2), Canterbury, England", 2004.

[27] J. BÉZIVIN. *Model Engineering for Software Modernization*, in "Working Conference on Reverse Engineering (WCRE2004), Delft University of Technology, the Netherlands", nov 2004.

[28] J. BÉZIVIN, S. HAMMOUDI, D. LOPES, F. JOUAULT. *Applying MDA Approach for Web Service Platform*, in "IEEE Enterprise Distributed Object Computing Conference - EDOC 2004, Monterey, California, USA", sep 2004.

[29] J. BÉZIVIN, S. HAMMOUDI, D. LOPES, F. JOUAULT. *Applying MDA Approach to B2B Applications: A Road Map*, in "Workshop on Model Driven Development (WMDD2004), at ECOOP 2004, Oslo, Norway", jun 2004.

[30] J. BÉZIVIN, S. HAMMOUDI, D. LOPES, F. JOUAULT. *B2B Applications, BPEL4WS, Web Services and dotNET in the context of MDA*, in "Int. Conf. on Enterprise Integration and Modelling Technology (ICEIMT 2004), Toronto, Canada", oct 2004.

[31] J. BÉZIVIN, F. JOUAULT, P. VALDURIEZ. *First Experiments with a ModelWeaver*, in "OOPSLA & GPCE Workshop on Best Practices for Model Driven Software Development, Vancouver, Canada", 2004.

[32] J. BÉZIVIN, F. JOUAULT, P. VALDURIEZ. *On the Need for Megamodels*, in "OOPSLA & GPCE Workshop on Best Practices for Model Driven Software Development, Vancouver, Canada", 2004.

[33] C. COULON, G. GAUMER, E. PACITTI, P. VALDURIEZ. *The RepDB* prototype: Preventive Replication in a Database Cluster*, in "Journées Bases de Donnéees Avancéees (BDA'2004)", 2004.

[34] C. COULON, E. PACITTI, P. VALDURIEZ. *Scaling up the Preventive Replication of Autonomous Databases in Cluster Systems*, in "Int. Conf. on High Performance Computing for Computational Science (VecPar 2004), Valencia, Spain", 2004.

[35] L. ESSANNOUNI, M. RZIZA, D. ABOUTAJDINE, M. GELGON, N. MOUADDIB. *Détection de Visages en utilisant les Machines à Vecteurs de Support*, in "International Symposium on Image and Video communications over Fixed and Mobile Networks, ENSTB, Brest, France", July 2004.

[36] S. GANÇARSKI, C. L. PAPE, H. NAACKE, I. SEBEI, P. VALDURIEZ. *The Leganet prototype: Managing Autonomous Databases in a Cluster System*, in "Journées Bases de Donnéees Avancéees (BDA'2004)", 2004.

[37] P. LAMARRE, S. CAZALENS, S. LEMP, P. VALDURIEZ. *A Flexible Mediation Process for Large Distributed Information Systems*, in "Int. Conf. on Cooperative Information Systems (CoopIS 2004), Cyprus", 2004.

[38] C. LEPAPE, S. GANÇARSKI, P. VALDURIEZ. *REFRESCO: Improving Query Performance through Freshness Control in a Database Cluster*, in "Int. Conf. on Cooperative Information Systems (CoopIS 2004), Cyprus", 2004.

[39] A. LIMA, M. MATTOSO, P. VALDURIEZ. *Adaptive Virtual Partitioning for OLAP Query Processing in a Database Cluster*, in "Brazilian Symposium on Databases (SBBD), Brazilia, Brazil", 2004.

[40] A. LIMA, M. MATTOSO, P. VALDURIEZ. *Scaling up the Preventive Replication of Autonomous Databases in Cluster Systems*, in "Int. Conf. on Parallel and Distributed Computing (Euro-Par 2004), Pisa, Italy", 2004.

[41] E. LOISANT, H. ISHIKAWA, J. MARTINEZ, O. MANABU, K. KATAYAMA. *User-Adaptive Navigation Structures for Image Retrieval*, in "Data Engineering Workshop (DEWS'2004), Iseshima, Japan", Database Society of Japan, March 2004.

[42] J. MARTINEZ, P. VALDURIEZ. *Vers un passage à l'échelle dans un SGBD d'images*, in "Conférence en Recherche d'Information et Applications (CORIA'04)", March 2004.

[43] A. PIGEAU, M. GELGON. *Incremental statistical geo-temporal structuring of a personal camera phone image collection*, in "Proceedings of the International Conference on Pattern Recognition (ICPR'2004), Cambridge, U.K.", August 2004, p. 128-132.

[44] A. PIGEAU, M. GELGON. *Organisation statistique spatio-temporelle d'une collection d'images acquises d'un terminal mobile géolocalisé*, in "Congrès Reconnaissance des Formes et Intelligence Artificielle (RFIA'2004), LAAS, Toulouse, France", January 2004, p. 76-84.

[45] P. VALDURIEZ, E. PACITTI. *Data Management in Large-scale P2P Systems*, in "Int. Conf. on High Performance Computing for Computational Science (VecPar 2004), Valencia, Spain", 2004.

[46] W. A. VOGLOZIN, G. RASCHIA, L. UGHETTO, N. MOUADDIB. *Querying the* SAINTETIQ *Summaries – A First Attempt*, in "Int. Conf. on Flexible Query Answering Systems (FQAS 2004), Lyon, France", LNCS/LNAI 3055, Springer Verlag, June 2004, p. 404–417.

[47] W. VOGLOZIN, G. RASCHIA, L. UGHETTO, N. MOUADDIB. *Interrogation de résumés de données et réparation de requêtes*, in "12èmes Rencontres Francophones sur la Logique Floue et ses Applications (LFA'04), Nantes, France", Cépaduès, November 2004, p. 239–246.

Internal Reports

[48] J. BÉZIVIN, S. HAMMOUDI, D. LOPES, F. JOUAULT. *An Experiment in Mapping Web Services to Implementation Platforms*, Technical report, n° 04.01, LINA, Université de Nantes, March 2004, http://www.sciences.univ-nantes.fr/info/recherche/Vie/RR/RR-IRIN2004-01.pdf.

Miscellaneous

[49] M. BOUGHANEM, S. CALABRETTO, J.-P. CHEVALLET, J. MARTINEZ, L. LECHANI-TAMINE. *Rapport final de l'AS-91 du RTP-9 : « Passage à l'échelle dans la taille des corpus »*, http://www.irit.fr/ASVolume/, January 2004, Rapport d'expertise auprès du CNRS.

[50] P. WEIKUM, P. VALDURIEZ. *Report on the 2004 SIGMOD Conference*, vol. 33, n° 4, 2004.