

COMICS: A Global Constraint Manager for Interactive Component Database Systems

Ludmila Himmelspach, Mehmet Kolac, Krasimir Kutsarov, Alexander Chernin,
Christopher Popfinger, and Stefan Conrad

Institute of Computer Science
University of Düsseldorf
D-40225 Düsseldorf, Germany

Abstract. When several heterogeneous database systems are combined to a federation it is necessary to deal with the problem of maintaining the consistency of semantically interrelated data. The integrity constraints that are defined on the schemas of the federation's members must be monitored and ensured. In this paper we introduce the COMICS system, a global **C**onstraint **M**anager for **I**nteractive **C**omponent **D**atabase **S**ystems, that is able to ensure strict consistency in a collection of autonomous and heterogeneous databases. The system uses the functionality of enhanced active databases to trigger a global integrity check from a local update operation.

1 Introduction

In the modern industrial and scientific world large amounts of data are being manipulated daily. Although the sources of this data are established independently, it is often necessary to combine them into federations. Thereby one must ensure that a high degree of autonomy is preserved and that federated data remains consistent. During the last two decades, various systems that handle this problem have been developed, but to our knowledge none of them ensures strict consistency of data. In this paper we present COMICS, a system which interconnects a collection of autonomous and heterogeneous relational databases to ensure strict consistency among interdependent databases. COMICS is specifically designed for Enhanced Active Databases, which are able to detect immediately all changes on their data stock and to actively notify a constraint manager about the modifications. The constraint manager checks the global integrity constraints by querying affected remote component systems and returns a permission to commit or reject the corresponding local transaction. In this way we can be sure that strict consistency is preserved without the need of recovery mechanisms.

The paper is structured as follows. First we introduce Enhanced Active Databases as prerequisite for our system. Section 3 describes the architecture of COMICS. Section 4 gives a detailed explanation of the process of checking global integrity constraints and an application scenario, followed by evaluation results and optimizations in Section 5. Section 6 discusses related work, while Section 7 concludes and catches up future work.

2 Enhanced Active Databases

The COMICS system is specifically designed for *Enhanced Active Databases (EADBS)* [1], which are a subclass of active databases. An EADBS is an active database system that is able to execute programs or methods from within its DBMS to interact with software or hardware components beyond its system border. The execution of a program or method in this context is called an *External Program Call (EPC)*. External programs can be called during the execution of a trigger or as a stored procedure to add new functionalities to the database system. This allows a program call to be executed as part of a database transaction triggered by a certain local event. An EADBS that is integrated into a federated information system as an *Active Component Database System (ACDBS)* [1] can use this enhanced activity to interact with other components of the federation to coordinate their actions. In particular, the extended functionality can be used to ensure consistency of interdependent data and to enforce business rules in the form of global integrity constraints. Within recent commercial database systems a commonly supported programming language that meets this requirement is Java (e.g. Java Stored Procedures or Java UDFs [2]). Besides JDBC, as a database connectivity framework, it contains libraries for various communication channels, such as sockets and the Remote Method Invocation Framework (RMI). Although we cast our work in the context of relational databases using Java and RMI, the concept adapts to other types of Enhanced Active Databases supporting different languages that fulfill the requirement just mentioned.

3 The COMICS System

The paper [1] has already presented an architecture for global integrity checking in heterogeneous information systems using active component systems. Global integrity constraints are enforced using constraint checking mechanisms which are implemented directly on the ACDBSs. In this way, the integrity constraint checks are processed on each component system independently from the constraint checks on the remaining CDBSs. This approach works well as long as the number of interrelated component systems is small, however with an increasing number of components in the system the risk of deadlocks also increases if relations in the CDBSs are updated concurrently. To overcome this drawback, we introduce an external constraint manager component, which enables and controls the communication among the ACDBSs during the global integrity constraint checking. Furthermore, the management of schema mappings and interdependencies between the data sources is shifted from the ACDBSs to an external mapping repository. One advantage of this approach is that adding or removing of mapping information of CDBs can be done quickly at a single point, namely in the central mapping repository. In this section we outline the architecture of COMICS shown in Figure 1 and describe its functionality.

In the COMICS system, global constraints are checked by the constraint manager (CM). Each global constraint is decomposed into a set of *partial integrity*