Verification of Large-Scale Distributed Database Systems in the NEOPPOD Project*

Olivier Bertrand[1], Aurélien Calonne[2], Christine Choppy[1], Silien Hong[3], Kais Klai[1], Fabrice Kordon[3], Emmanuel Paviot-Adet[3], Laure Petrucci[1], and Jean-Paul Smets[2]

[1] LIPN, CNRS UMR 7030, Université Paris XIII, 99 avenue Jean-Baptiste Clément, 93430 Villetaneuse, France
[2] NEXEDI, 270 bd Clémenceau, 59700 Marcq-en-Barœul, France
[3] LIP6 - CNRS UMR 7606, Université Pierre & Marie Curie, 4 PLace Jussieu, 75252 Paris Cedex 05, France

1 Introduction

Nowadays, large applications are developed that must access and maintain huge data bases. Examples of such software are e-government, internet based information systems, commerce registries, etc.). They are characterised not only by the huge amount of data they manipulate, but also by a mandatory high level of security and reliability.

Hardware has become rather cheap: 150€ for a 64-bit dual core server, and 100€ for a 1TB disk. A large application and data base server could be composed of hundreds of such disks and servers. It would thus be possible to handle as much as 1PB of data and thousands simultaneous transactions, for a moderate investment of 175,000€. However, this requires to elaborate reliable and safe distributed data base management software.

ZODB, the *Zope Object Database*, has become within a few years the most used object data base. This libre software, associated to the Zope application server is used for a Central Bank, to manage the monetary system of 80 million people in 8 countries [2]. It is also used for accounting, ERP, CRM, ECM and knowledge management. It is now a major libre software as PHP or MySQL is.

However, the current Zope architecture does not apply yet for data as huge as those mentioned earlier. In order to attain such performances, the architecture had to be revisited. It led to the design of an original peer-to-peer transaction protocol: NEO. This protocol must also ensure both safety and reliability, which is not easy to achieve for distributed systems using traditional testing techniques.

The aim of the NEOPPOD project is to formally verify safety and reliability properties for the new NEO protocol. The process thus involves the model design, expected properties verification and eventual revision of the protocol according to the results obtained.

* This work is supported by FEDER Île-de-France/System@tic—libre software.

2 Challenges

The new NEO protocol is expected to handle clusters of 100 to 10,000 server nodes. Therefore, safety and reliability are critical issues. The project aims at considering both qualitative (e.g. data consistency) and quantitative (e.g. performance aspects) characteristics of the system. This requires the use of different formal methods that should be operated from a common specification for consistency purposes.

Modelling issues: Hence, designing an appropriate specification is a first challenge. Starting from the protocol description, a reverse-engineering process allows for extracting step-by-step a corresponding symmetric Petri net mode [1]. Since the original program description is very large and well structured, it is mapped to a modular specification. However, in order to mimic different configurations of the cluster architecture, as well as the different roles of the servers involved, w.r.t. the protocol operation, the model must also be highly parameterised.

Verification issues: Since numerous instances of several actors are involved in the system, the combinatorial explosion of the state space is a major difficulty. Dedicated techniques exploiting characteristics of distributed systems must be elaborated. These techniques rely on both exploiting symmetries and use of compact data structures such as decision diagrams. Finally, the hierarchical architecture must be exploited to separate local actions in the system from those affecting several components.

Since the model is highly modular, compositional and/or modular verification approaches must be investigated. This should be particularly interesting to check the dimensioning of the system elements by means of place bounds.

3 Expected outcome

The expected outcome of this project are :

- a modular specification designed with several levels of abstraction ;
- verification of critical properties of the protocol, such as data consistency, correct fault recovery, detection of bottlenecks ;
- pushing further the limits of verification tools and techniques, enhancing the CPN-AMI platform [3].

References

1. G. Chiola, C. Dutheillet, G. Franceschinis, and S. Haddad. A symbolic reachability graph for coloured Petri nets. *Theoretical Computer Science*, 176(1–2):39–65, 1997.
2. ERP5. *Central Bank Implements Open Source ERP5 in Eight Countries after Proprietary System Failed.* `http://www.erp5.com/news-central.bank`.
3. MoVe-Team. The CPN-AMI Home page, url: `http://www.lip6.fr/cpn-ami`.