

# An Open Architecture for Scalable Database Clustering

Rui Carlos Oliveira  
Universidade do Minho, Portugal  
rco@di.uminho.pt

## Abstract

*The Database Management System (DBMS) used to be a commodity software component, with well known standard interfaces and semantics. However, the performance, scalability and reliability expectations being placed on DBMS have increased the demand for a variety of add-ons, that augment the functionality of the database in a wide range of deployment scenarios, offering support for features such as clustering, replication, and self-management, among others.*

*Recently, several such add-ons have been designed and implemented both in the academia and by leading commercial database providers. Each proposal tends to target certain goals and applications, therefore establishing specific tradeoffs that impair their flexibility. Moreover, it has been a common fundamental assumption that any add-ons should not be intrusive and that the DBMS should be kept unchanged and monolithically handled. While this is a very sensible and pragmatic view due to the complexity of DBMS and the critical role they play in existing information systems, emerging demands on scalability require greater flexibility of the whole data management system so that major functionalities can be realized as autonomous services with specific tradeoffs and quality of service.*

*The GORDA project (EU IST FP6) proposed a general purpose DBMS reflection architecture and interface – GAPI, which supports a number of useful extensions while at the same time admitting efficient implementations. By exposing at the interface an abstract representation of the systems' inner functionality, the later can be inspected and manipulated, thus changing its behavior without loss of encapsulation. DBMS have long taken advantage of this – on the database schema, on triggers, and when exposing the log.*

*In this talk we describe the various aspects and goals that led to GAPI and we illustrate the usefulness of the architecture and interface with concrete examples. GORDA fundamentally emphasizes the modularity of the add-ons, e.g. clustering, replication and management, the DBMS itself and fundamental building blocks such as reliable group communication.*

*This effort clearly seems to be of major relevance for the emerging Cloud storage systems. By easing the development of different add-ons for database systems, it can be used to enrich the current products offered by key providers such as Amazon and Google and enable small providers to jump into this new trend. Cloud storage offers are touted as being able to deal with both very large data volumes as well as large numbers of clients with different storage needs. Per se, these two requirements call for highly scalable and flexible infrastructures. Current general tradeoffs however, favor minimal client interfaces with pretty relaxed consistency guarantees which are not adequate to general applications. Bringing transactional semantics and ACID guarantees to the Cloud appears as a major commercial trend and research challenge.*

## Biography

Rui Oliveira graduated (1991) in Electrotechnic and Computers Engineering, by the University of Porto, has a Master (1994) in Computer Science by the University of Minho and a PhD (2000) in Computer Science by The École Polytechnique Fédérale de Lausanne. He is Associate Professor in Computer Science at Universidade do Minho, teaching dependable distributed systems to Masters and Doctoral Programmes. He currently leads the Computer Science and Technology Center of Univesidade do Minho.

His current research interests are in reliable and large-scale distributed systems, in particular with application to distributed database systems and cloud computing. His work has been focused on distributed fault-tolerant agreement protocols, replication protocols and scalable reliable multicast protocols. Rui Oliveira coordinated several research projects on database replication namely GORDA - Open Replication of Databases (EU IST FP6). He publishes regularly in conferences such as IEEE Dependable Systems and Networks (DSN) and IEEE Symposium on Reliable Distributed Systems (SRDS), and has recently given several invited talks and tutorials. He is a member of ACM and IEEE. His Web page is: <http://gsd.di.uminho.pt/rco>.