Armada: a Model for an Evolving Database

Fabian Groffen* Martin Kersten Stefan Manegold

Centrum voor Wiskunde en Informatica

Soon we face a common repository size scaling into petabytes, filled with data that needs to be stored and processed. However, the rapidly improving technology cannot keep up with the data growth rate, hence data processing becomes more and more an expensive and time-consuming task. This problem is of major concern, since data processing is a core process for many businesses and applications. Yet a real solution to the data growth problem has to be found.

Scaling into multiple machines to process the data is currently successfully applied in grids and distributed databases. However, the centralised scaling technique using a large number of machines, is *fragile* from an availability point of view. All systems depend on the availability of one. Moreover, this single point-of-failure can easily get overloaded, thereby forming the bottleneck in serving a high workload.

A novel reference architecture for a distributed database is urgently needed. It should take site autonomy as both driving force and core feature of a system architecture. A sole central server to guide all interactions is a *dead end* for the scalable solutions required. Instead, several sites may take such a role for a limited period and only for part of the data space.

Within the *Armada* project we aim to create a reference model for a flexible, self-maintaining, efficient distributed database architecture. To achieve this goal, we try to avoid the classical bottlenecks that limit the efficiency of most existing and proposed architectures. These bottlenecks can be seen as the two extreme alternatives of storing and maintaining the metadata that is necessary to ensure correct and efficient handling of the actual data. Classical designs on the one end of the spectrum require a centralised server that holds all metadata, and hence forms a hotspot. Designs on the opposite end of the spectrum avoid this hotspot by fully replicating all metadata — at the expense of requiring that all metadata updates are instantaneously propagated to all sites. The Armada model is a balance between these two extremes. Metadata is only partially replicated over the system. Additionally, sites are able to cope with incomplete or stale metadata.

The model uses data fragmentation, data replication and data fusion as the minimal basis for the lineage of data blocks, that allows maximal autonomy of the nodes co-operating in a distributed architecture. The chosen approach lays the basis for studying further building blocks for an organic database, which is designed to facilitate evolutionary growth in a distributed environment.

In addition to the Armada model, we present a preliminary study on its implementation in an SQL-based system. With SQL being the common language spoken by relational database systems, we aim being as much as possible compatible with existing database systems. We envisage that the Armada model can be implemented on top of SQL in the MonetDB/SQL database.

*Contact Author and PhD Student at CWI, Fabian.Groffen@cwi.nl