

Tailor-made Data Management – A Challenge for Database Technology?

Gunter Saake
School of Computer Science
University of Magdeburg, Germany
e-mail: saake@ovgu.de

The need for tailor-made data management is apparent in database research. There is an impressive body of research on extensible or customizable DBMS like kernel systems or component DBMS. Nevertheless, all these approaches have drawbacks, e.g., regarding customizability or performance, and there is no general solution for efficient development of tailorable DBMS. Novel software engineering techniques (e.g. software product lines and *feature-oriented programming*), can help in developing customizable DBMS and reduce the complexity of building variable solutions as we could show for embedded systems [2]. Applying such an approach to other domains like stream processing, column stores, XML databases, etc. is promising to create DBMS that provide only needed functionality, can be better tuned and maintained, and provide reuse within a domain and even between different domains.

Customizability is also required on the level of the query language, i.e., SQL, which grows with every new standard. Independent of the actual task, which may be as simple as querying a single value, developers are confronted with the full arsenal of SQL. Furthermore, applications only use a small subset of the standard and require special extensions, e.g., for sensor networks or stream processing. This results in *SQL dialects* that are used in different domains and are decoupled from the SQL standard. We could show that SQL (i.e., the grammar and the parser) can be decomposed with a feature-oriented approach. This results in a family of SQL dialects from which tailor-made SQL parsers, e.g., for sensor networks, can be generated [1].

By combining tailor-made SQL dialects and customizable DBMS we think that one can build DBMS that can be fully tailored to an application domain or a special use case within a domain while providing high reuse. Creating such fully customizable systems is a highly challenging task. It involves customizability on all levels of a DBMS such as the query processor and optimizer, the transaction subsystem, or the storage system. Because of interdependencies between those subsystems an architecture has to be developed that is able to handle the resulting complexity. For example, changing the query language affects the whole DBMS including the query optimizer which is highly connected with all other parts of a DBMS. New extensibility mechanisms as provided by feature-oriented programming can help in building such variable systems. However, this is probably not sufficient and the community has to investigate in an architecture and new mechanisms for handling the variability that keep maintainability and performance of DBMS.

Literatur

- [1] M. Rosenmüller, C. Kästner, N. Siegmund, S. Sunkle, S. Apel, T. Leich, and G. Saake. *Sql à la carte – toward tailor-made data management*. In *13. GI-Fachtagung Datenbanksysteme für Business, Technologie und Web (BTW)*, pages 117–136, March 2009.
- [2] M. Rosenmüller, N. Siegmund, H. Schirmeier, J. Sincero, S. Apel, T. Leich, O. Spinczyk, and G. Saake. *FAME-DBMS: Tailor-made Data Management Solutions for Embedded Systems*. In *EDBT'08 Workshop on Software Engineering for Tailor-made Data Management (SETMDM)*, pages 1–6, March 2008.