ObjectGlobe: Open Distributed Query Processing Services on the Internet *

Reinhard Braumandl* Markus Keidl* Alfons Kemper* Donald Kossmann†
Stefan Seltzsam* Konrad Stocker*

*University of Passau
⟨last name⟩@db.fmi.uni-passau.de

†TU Munich
kossmann@in.tum.de

1 Vision

Today, there is a vast discrepancy between data publication and data processing capabilities on the Internet. Virtually anybody can publish data, e.g., in HTML- or XML-format. Consequently, a wide range of (different quality) data sources exist on the Internet, ranging from personal documents to real estate offers to product catalogs, to name just a few. In a way, the Internet could be viewed as a large distributed database. However, today's web sites are merely "dumb" page servers which are only capable to send the data sitting behind a particular URL/URI. At best, web sites can process local queries if they are backed by a database system and a query interface is published via, e.g., a forms interface. But true distributed query processing plans as enabled by homogeneous distributed databases with interacting distributed subplans are not supported. Our goal is to create a query processing server that can be deployed throughout the Internet. These query servers can then be used in a federation to execute truly distributed query processing plans composed of completely unrelated *query processing services* which are offered on the Internet in an open market. These services could be specialized on providing data, resources for the query execution itself (CPU power, storage area) or functions which can be embedded in the execution. Such an open system could vastly ease the interaction in business-to-business and business-to-customer applications like shopping portals, electronic marketplaces or virtual enterprises. For example, somebody could search for real estate offers which fulfill some constraints with regard to the building, its location and the corresponding ambient data. The respective query should then access data from several commercial realtor databases, a geographical information system and a server with global ambient data. Additionally, the query should use a ranking function specialized for real estate data and provided in the form of *mobile code* by a third party specialized in that particular business area.

1.1 The Requirements

The differing demands of *data providers* and users with respect to such a global query processing system show why current architectures for distributed databases ([CDF+94]) and mediator systems ([HKWY97, PGGMU95, JKR99]) are not sufficient. Data providers are interested in

Bulletin of the IEEE Computer Society Technical Committee on Data Engineering

*This research is supported by the German National Research Foundation under contract DFG Ke 401/7-1.