Master’s / Diploma Thesis

Extending an Open Source Enterprise Service Bus for PostgreSQL Statement Transformation to Enable Cloud Data Access

Erweiterung eines Open Source Enterprise Service Bus um die Transformation von PostgreSQL Anweisungen und Abfragen für die Unterstützung von transparentem Datenzugriff in die Cloud (German title)

Beginning: Immediately

Background
Cloud computing has become increasingly popular with the industry due to the clear advantage of reducing capital expenditure and transforming it into operational costs [1], [2]. To take advantage of Cloud computing, an existing application may be moved to the Cloud or designed from the beginning to use Cloud technologies [3]. Applications are typically built using a three layer architecture model consisting of a presentation layer, a business logic layer, and a data layer. The presentation layer describes the application-users interactions, the business layer realizes the business logic and the data layer is responsible for application data storage. The data layer is in turn subdivided into the Data Access Layer (DAL) and the Database Layer (DBL). The DAL encapsulates the data access functionality, while the DBL is responsible for data persistence and data manipulation. Each application layer can be hosted using different Cloud deployment models. Possible Cloud deployment models are: Private Cloud, Public Cloud, Community Cloud, and Hybrid Cloud. In this diploma thesis we focus on the database layer. Application data is typically moved to the Cloud because of e. g., Cloud bursting, data analysis, or backup and archiving [6].

As the DAL might be distributed using non-Cloud technologies and Cloud data stores or data services, we identified the requirement of transparent access of the Business Layer to the data in [3]. The distribution of the Database Layer essentially changes the borders of the application compared to traditional applications. On the one hand, transparent access enables loose coupling between the Business Layer and the DAL so that the used Cloud data stores or data services can be changed without affecting the Business Layer. On the other hand, this requires additional functionality in the DAL, because it should be able to determine the data store or data service the request should be forwarded to based on the request sent by the Business Layer.

As this does not only required dynamic routing, but also transformation of the statement send by the Business Layer to the format understood by the corresponding Cloud data store or Cloud data service, in this diploma thesis the student investigates the extension of an existing Open Source Enterprise Service Bus used as DAL [4] for SQL statement transformation focusing on PostgreSQL as source data store.
Tasks

- State of the Art analysis of
  o Existing approaches in research and industry on how to transform SQL statements

- Design and implementation of an extension of a Database Systems Knowledge Base (DBS-KB)
  o Identification of equivalence classes for different traditional relational databases supporting SQL statements focusing on data migration from on-premise relational databases to off-premise relational databases or data services [3].
  o Definition of the transformation functions between the equivalence classes considering the potential transformation directions and especially address the following questions:
    - What is possible related to transformation functions and what are the conditions under which a transformation is possible?
    - What are the limitations what are the implications and effects they have, e.g. on communication with the database or upper architecture layers, i.e. business logic layer?
  o Conceptualization, requirement analysis, specification, and design for an extension of Apache ServiceMix to enable PostgreSQL statement transformation based on the outcomes of the previous tasks and the results of [4].
  o Reference implementation of the Apache ServiceMix extension based on selected established and common mechanisms supported by Java and by the Cloud Data Hosting Solutions based on the results of [4].

- Validation
  o Demonstration of PostgreSQL statement transformation based on several traditional, selected relational databases (MySQL) and Cloud data stores (focus on established providers like Amazon, Google, and Microsoft) following the validation approach in [4].

- Evaluation
  o Measurement of implementation performance of PostgreSQL statement transformation and comparison to the results of the complexity analysis of the algorithm. Performance evaluation has to be geared to and compared to performance evaluation in [4]

- Bonus: Addressing selected open issues identified in future work of [4].

Required previous knowledge and experiences

- Java programming skills and expertise
- Cloud computing [1], [2]
- Cloud Data Migration [3], [5], [6]
- SQL
- ...or the declared intention to deeply dive into these topics in advance

The lectures on Fundamentals of Architecture of Application Systems and Message-based Application Integration as well as [1], [2], [4], [5], [6] provide further reading for the preparation.
A helpful guide for planning and writing a thesis is provided by [8] and [9].
Literature


Contact

Supervisor
Steve Strauch
Room: 1.356
Tel.: +49 711 685-88212
E-Mail: steve.strauch@iaas.uni-stuttgart.de

Examiner
Prof. Dr. Frank Leymann