MUSYOP: Towards a Query Optimization for Heterogeneous Distributed Database System in Energy Data Management

Zhan Liu, Fabian Cretton, Anne Le Calvé, Nicole Glassey, Alexandre Cotting, Fabrice Chapuis
Institute of Business Information Systems
University of Applied Sciences and Arts Western Switzerland
HES-SO Valais, Sierre, Switzerland
OVERVIEW

• Who Are We?
• Introduction
• Objectives
• Research methodologies
• Architecture of Heterogeneous Distributed Database System
• Implementation and Use Cases
• Conclusion and Future Research
WHO ARE WE?
WHO ARE WE?

- 72 staff members
  - 20 Professors
  - 14 Research scientists
  - 29 Research assistants
  - 9 Administrative collaborators and interns

- 92 Scientific publications

- 291 Projects (2013)

- CHF 8,85 millions (Turnover in 2013)
## CORE COMPETENCIES

<table>
<thead>
<tr>
<th>eEnergy</th>
<th>eGov</th>
<th>eHealth</th>
<th>ERP</th>
<th>eServices</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPM</td>
<td>Cloud Computing</td>
<td>Data Intelligence</td>
<td>Ergonomy, Usability</td>
<td>Intelligent Agents</td>
</tr>
<tr>
<td></td>
<td>Medical Information Processing</td>
<td>Semantic Web</td>
<td>Software Engineering</td>
<td></td>
</tr>
</tbody>
</table>
INTRODUCTION

• Databases in the context of an energy database management system (EDMS) are non-integrated, distributed and heterogeneous

• As a result, information integration is becoming increasingly important

• BUT, it consumes “a great deal of time and money”
OBJECTIVES

• To propose a uniform approach by using Semantic Web technologies for SPARQL querying a heterogeneous distributed database system named MUSYOP for energy data management

• To provide transparent query access to multiple heterogeneous data sources

• Query optimization to speed-up search executions
**RESEARCH METHODOLOGIES**

- **(1) Warehouse approach (centralized system)**
  - ✨ High efficiency and the capability of extracting deeper information for decision making
  - 😞 Must set up “data cleansing” and “data standardization” before actual use
  - 😞 Could take months of planning

- **(2) Federated approach (decentralized system)**
  - ✨ High adaptability to frequent changes of data sources
  - ✨ support of large numbers of data sources and data sources with high heterogeneity

- **MUSYOP focused on the federated approach**
ARCHITECTURE OF HETEROGENEOUS DISTRIBUTED DATABASE SYSTEM
COMPONENTS OF ARCHITECTURE (1)

- **Distributed Query Decomposer**
  - Generates a number of transactions to match remote data sources
  - Assembles transaction results and returns to the user

- **Query Optimizer**
  - Data source optimization: data source selection and building sub-queries
  - Join order optimization: determine the numbers of intermediate return results

- **Distributed Transaction Coordinator**
  - Detects and handles persistent records and manages the communications with databases
COMPONENTS OF ARCHITECTURE (2)

- **D2RQ**
  - Query relational database using SPARQL
  - Custom D2RQ mapping file for describing the relation between an ontology and an relational data model

- **SPARQL Endpoint**
  - Query the knowledge bases, such as TripleStore via SPARQL protocol services

- **SPARQL2XQUERY**
  - Framework provides a generic method for SPARQL to XQuery translation

- **ALLEGROGRAPH**
  - Query MongoDB databases using SPARQL
  - Manual transform datasets from MongoDB to TripleStore
  - **SPARQLverse**
ARCHITECTURE IMPLEMENTATION

- **User's notification settings**
- **Frequent data of consumption**
- **Daily, monthly and yearly consumption data and user alert management data**
- **Building information, cities location information and weather information**
- **User's notification settings**
- **Frequent data of consumption**
ENERGY USE CASE

MySQL
- daily data
- monthly data
- user

XML
- notifications

Notify me and display all the buildings below the altitude of 600 meters when the gas consumption in winter is 15% higher than the previous year for the same month.

MongoDB
- config_port
- frequency data

TripleStore
- buildings
- geography
- weather
PREFIX db: <http://localhost:2020/resource/>
prefix intbatXtra: <http://websemantique.ch/onto/intBatXtra#>
prefix musyopOnto: <http://www.websematique.ch/voc/musyop#>
prefix dbp-ont: <http://dbpedia.org/ontology/>
prefix xsd: <http://www.w3.org/2001/XMLSchema#>

?city ?elevation

  vocab:config_ports_affectation "Gaz";
  vocab:config_ports_batiment_id ?buildingID ;
  vocab:config_ports_config_port_id ?cpID.
  ?donneeMens2011 vocab:donnees_mensuelles_config_port ?config_port;
  vocab:donnees_mensuelles_timestamp "2011-01-01T00:00:00"^^xsd:dateTime;
  ?donneeMens2012 vocab:donnees_mensuelles_config_port ?config_port;
  vocab:donnees_mensuelles_timestamp "2012-01-01T00:00:00"^^xsd:dateTime;
  ?buildingUser rdf:type vocab:building_user;
  vocab:building_user_building_id ?buildingID;
  vocab:building_user_utilisateur_id ?userID;
}}

?build a intbatXtra:Object;
musyopOnto:hasID ?buildingID ;
rdfs:label ?buildingLabel;
musyopOnto:locatedIn ?city.
Filter(?elevation < ?paramElevationMax)
BIND((xsd:double(?cons2)/xsd:double(?cons1)) as ?consDiff).
Filter(?consDiff > ?paramConsDiffMax)
}
ORDER BY ?buildingID ?userID
limit 100
VALUES (?userID) { USERID_VALUES }
GRAPHICAL USER INTERFACES

MUSYOP

Username:

Password:

Login  Clear

Alerts Configuration
Configure the type of alert...

Active Alarms
List of all currently active...

Acquitted Alarms
List of all already acquitted...

New Alarm
Add a new alarm assigned...

Check New Alarms
Check and send new alarm...
GRAPHICAL USER INTERFACES

NOTIFICATION
- Notify me when the outside temperature is low than 12 degrees in my related buildings.
- Notify me and display all the buildings below the altitude of 600 meters, when the gas consumption in winter is 15% higher than the previous year for the same month.
- Notify me when the daily electric consumption of home building has been reached the predefined maximum value.

Submit

ALARMS LIST
Select all the currently active alarms you want to acquit and press the "Acquit" button.

- Alarm Echallens - Building A: The outside temperature is lower than 12 degrees
- Alarm Lausanne - Building B: The outside temperature is lower than 12 degrees
- Alarm Bussigny - Building C: The outside temperature is lower than 12 degrees
- Alarm Lausanne - Building D: The outside temperature is lower than 12 degrees

ALARMS
- Alarm Echallens - Building ...
- Alarm Lausanne - Building ...
- Alarm Lausanne - Building ...

Created: 02.04.2014 at 10:46:28
Acquitted: 02.04.2014 at 10:46:28
Acquitted by: Zhan Liu
GRAPHICAL USER INTERFACES


CONCLUSION AND FUTURE RESEARCH

• Presented an architecture of heterogeneous distributed database system for energy data management

• Integration of a mediator server to access and coordinate with four different databases: relational database, NOSQL, Triple Store and XML

• Proposed an approach for query optimization based on our architecture

• Future work required on the evaluation our system with a large amount of energy data in Switzerland