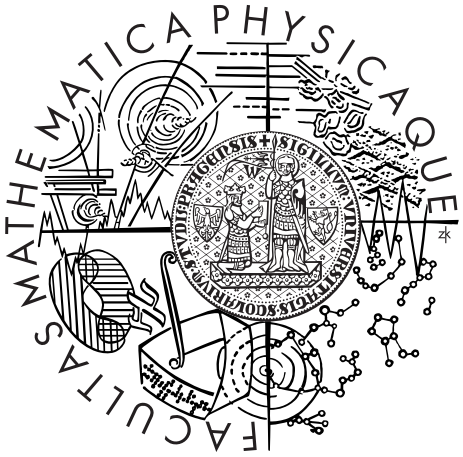
Charles University in Prague
Faculty of Mathematics and Physics

MASTER THESIS



Radovan Duga

Querying NoSQL databases in MPS

Department of Distributed and Dependable Systems

Supervisor of the master thesis: RNDr. Pavel Parízek, Ph.D.
Study programme: Informatics
Specialization: ISS

Prague 2013

Dedication.

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In ..... date ..... signature of the author

Název práce: Dotazování NoSQL databáz v prostředí MPS

Autor: Radovan Duga

Katedra: Katedra distribuovaných a spolehlivých systémů

Vedoucí diplomové práce: RNDr. Pavel Parízek, Ph.D., Katedra distribuovaných a spolehlivých systémů

Abstrakt: S příchodem NoSQL databází se objevila i potřeba pro vznik doménově specifických dotazovacích jazyků. Jednou ze zajímavých domén jsou grafové databáze jako například Neo4j s dotazovací jazykem Cypher. Doménově specifické jazyky (DSLs) může být navržena a snadno použita pomocí speciálních vývojových prostředích zvaných Language Workbenche. Velmi populární Language Workbench je MPS, který implementuje koncept projekčních DSLs.

Tato práce zodpovídá otázku, zda Language Workbenche a projekční DSLs mohou být přínosem v doméně NoSQL databází, vystihnout výhody projekčních DSLs použitím různých typů přístupu. Dalším specifickým cílem je navrhnout a implementovat dotazovací DSL jazyk pro vybranou NoSQL databázi (např. Neo4J nebo Redis) jako případová studie.

Klíčová slova: NoSQL, MPS, dotaz, Cypher

Title: Querying NoSQL databases in MPS

Author: Radovan Duga

Department: Department of Distributed and Dependable Systems

Supervisor: RNDr. Pavel Parízek, Ph.D., Department of Distributed and Dependable Systems

Abstract: With the advent of NoSQL databases, a need for targeted domain-specific query languages has become evident. One of the interesting domains are graph databases, such as Neo4j with the query language Cypher. Domain specific languages (DSLs) can be designed and easily used with the help of special development environments called Language Workbenches. A very popular Language Workbench is MPS, which implements the concept of projectional DSLs.

This work will answer the question whether Language Workbenches and projectional DSLs can make a contribution in the domain of NoSQL databases, and identify the benefits of projectional DSLs over different approaches. An additional specific goal is to design and implement a practical MPS-based query DSL for a chosen NoSQL database (e.g., Neo4J or Redis) as a case study.

Keywords: NoSQL, MPS, query, Cypher

Contents

Introduction

1. NoSQL databases

1.1 Neo4j graph database

1.2 Neo4j Cypher query language

2. Domain specific languages

3. Language workbenches

3.1 How to define DSL in MPS

4. MPS Language workbench

4.1 How to define DSL in MPS

4.1.1 Structure DSL

4.1.2 Editor DSL

4.1.3 Constraints DSL

4.1.4 Typesystem DSL

4.1.5 Intensions and other parts

4.2 MPS Pros and Cons

5. Design of MPSypher language

5.1 Text-like editor

5.2 Graphical extensions

6. Implementation details of MPSypher language

6.1 Implementation of patterns

6.2 Implementation of references

6.3 Integration into BaseLanguage

7. Case Study of MPSypher

Conclusion

Bibliography

List of Tables

List of Abbreviations

Attachments