XML and NoSQL DBMS:Migration and Benchmarking

Master Thesis in fulfillment of the requirements for the degree of
Master of Science (M.Sc.)

Author:

Prakash Thapa

Supervisor:

Prof. Dr. Marc H. Scholl

Konstanz University

October 6, 2014

Abstract

Acknowledgments

The completion of this master thesis would not have been possible without the support of many people. I must express my sincerest thanks towards Dr. Christian Grün for his continuous support and patience during my research. I should confess that I can not imagine a better person to guide and advise me. Besides, I would like to extend my gratitude to my supervisor Prof. Dr. Marc H. Scholl to give me the chance to be involved in this group.

My special appreciation also goes to Prof. ... for answering completely and accurately questions that I would ask frequently. I would like to offer my heartfull gratitude my parents and family for encouraging me throughout my whole study at University. At the end, I am deeply appreciative of my beloved friend M. Ali Rostami which without who it was impossible to it do.

Contents

1 Introduction

The geospatial data is increasingly becoming more important nowadays. example ...???? Different ways of processing this data is studied over years. However, the huge existing data requires more efficient ways nowadays. The time complexity is specially of great interest. We talk about the motivations to .... chera XML khube.......

1.1 Motivation

Few years of time XML [**?**] was de facto data exchange format which enabled people to do previously not that easy thing that time like exchange of content of Microsoft's office documents exchange through HTTP connections. Native XML database

But in recent years a bold transformation has been a foot in the world of Data exchange. The more light weight, less bandwidth consumer JSON(JavaScript Object Notation)[1] has been emerge not just as an alternative to the XML but as rather as potential full Blown successor[2]. Even though these two format has their own pros and cons, the rise of JSON as key in data exchange format, new database technologies so called NoSQL are also emerges and getting success in their own way. The rate of new research papers in these system are increasing in recent years.

1.2 Overview

This essay will first define the concepts needed later in the discussion, called Methodology, in Section **??**. In this section, the basic geospatial data and it's application areas will be introduced. Also, various ways that other works are processing this data will be investigated in Section **??**. BaseX, as a native XML database, which is our focus in this essay, is studied and extended by a module to handle geospatial data in Section **??**. Section **??** improves the efficiency of the module by proposing another novel way considering MongoDB. Finally, new ideas for further works are explained in Section 2.

2 Future Work

storing in the memory

List of Figures

List of Tables

Listings