Title: Implementation of parallel query processing in PostgreSQL
Author: Bc. Daniel Vojtek
Department: Department of Software Engineering
Supervisor: Mgr. Július Štroffek
Supervisor's e-mail address: julo@stroffek.cz

Abstract: Parallel query processing can help with processing of huge amounts of data stored in database systems. The aim of this diploma thesis was to explore the possibilities, analyze, design and finally implement parallel query processing in open source database system PostgreSQL. I used a Master/Worker design pattern, in which standard PostgreSQL backend process is a master. As workers I used processes created from postmaster. In the thesis I focused on preparing an infrastructure necessary for parallel processing. I defined a new top level memory context over shared memory, which allows efficient and convenient memory allocations. Then I implemented creation of new worker processes, based on master process requirements. To be able to control these workers I defined controlling structures using state machines. Then I implemented parallel sort operation and SQL operator UNION ALL using this infrastructure. The result of this diploma thesis is not only implementation of infrastructure and some parallel operations, but also description of the problems encountered during the implementation with their possible solutions. I also outlined further extensions of our implementation.

Keywords: Parallel query processing, PostgreSQL, Master Worker paradigm