

# Bachelor Thesis Review

Faculty of Mathematics and Physics, Charles University

**Thesis author** Martin Koreček  
**Thesis title** A Comparison of Strategies for Database Caching  
**Year submitted** 2021  
**Study program** Computer Science  
**Study branch** General Computer Science

**Review (co)author** William Brown Advisor  
**Department** SUSE Labs, Brisbane, Australia  
**Review (co)author** Petr Tůma Consultant  
**Department** Department of Distributed and Dependable Systems

## Overall

good OK poor insufficient

	good	OK	poor	insufficient
Assignment difficulty	X			
Assignment fulfilled		X		
Total size <i>... text and code, overall workload</i>		X		
<p>Researching and measuring an already heavily researched topic (caches) while adding in the complexity of concurrent behavior and analysis is a challenging area—there is a huge amount of existing work for single-threaded structures, but there is very little to draw on for how to approach the concurrent aspects. Many tools, systems and measurements assume single-threaded behavior and analysis. This meant that Martin had to overcome some very difficult challenges, create new ideas, and new approaches to achieve this. While there are limitations, he presented a solid foundation and investigation of the space. In the process he also was able to learn and realise more about the problem, and his own strengths and limitations in his approach.</p> <p>The exploratory nature of the assignment required more background study and speculative work than usual. Obviously, this is only partially reflected in the text of the thesis, but should also be considered.</p>				

## Thesis Text

good OK poor insufficient

	good	OK	poor	insufficient
Form <i>... language, typography, references</i>	X			
Structure <i>... context, goals, analysis, design, evaluation, level of detail</i>		X		
Problem analysis			X	
Developer documentation			X	
User documentation			X	

The text is reasonably well written and generally above average compared to other theses written by non-native speakers. The structure reflects the goals of the thesis, but the analytical part struggles to convey the complex interactions inherent to practical aspects of cache behavior (this is at least in part due to the challenging topic and not necessarily an indication of sub par skills).

The code accompanying the thesis was created mostly as a tool to conduct experiments, rather than an end-result of the thesis, and is therefore not explicitly covered by standard documentation. That said, the documentation could have been provided to help future experimenters.

**Thesis Code**

good    OK    poor    insufficient

Design	<i>... architecture, algorithms, data structures, used technologies</i>	X			
Implementation	<i>... naming conventions, formatting, comments, testing</i>		X		
Stability			X		

Martin was able to implement all of the presented algorithms in Rust (which may not have had some of these cache strategies before), as well as implementing concurrent data-structures without any prior experience in the space. There are some rough edges in the implementation but none of these are severe enough to bring doubt into the implementations. Further testing of the implementations would have helped gain a stronger confidence in their correct behavior, but on reading they appear correct implementations of the algorithms.

It is worth noting that some of the behaviors such as the concurrent caching strategies have very little prior art in the space. To implement this is a major achievement of its own. A hidden gem of this thesis was his work into concurrent tries for hashmaps, which may prove more effective than external approaches using b+trees. This is an area that could be further pursued and investigated.

**Overall grade**    Very Good  
**Award level thesis**    No

Date

Signature