Annotation

Title	Garbage Collection in Software Performance Engineering
Author	Peter Libič
	peter.libic@d3s.mff.cuni.cz
Advisor	doc. Ing. Petr Tůma, Dr.
	petr.tuma@d3s.mff.cuni.cz
Department	Department of Distributed and Dependable Systems
	Faculty of Mathematics and Physics
	Charles University
	Malostranské nám. 25, 118 00 Prague, Czech Republic

Abstract

The increasing popularity of languages with automatic memory management makes the garbage collector (GC) performance key to effective application execution. Unfortunately, performance behavior of contemporary GC is not well understood by the application developers and often ignored by the performance model designers.

In this thesis, we (1) evaluate nature of GC overhead with respect to its effect on accuracy of performance models. We assess the possibility to model GC overhead as a black-box and identify workload characteristics that contribute to GC performance. Then we (2) design an analytical model of one-generation collector and a simulation model of both one-generation and two-generation collectors. These models rely on application characteristics. We evaluate the accuracy of such models and perform an analysis of their sensitivity to the inputs. Using the model we expose the gap between understanding the GC overhead based on knowing the algorithm and the actual implementation. In the course of evaluation we discover important GC issues concerning application developer and suggest how to tackle those issues. Last, we (3) design a model to help the developer predict effects of certain code additions on GC overhead of the whole application. This model is easy to use and only uses readily obtainable inputs and workload characteristics.

Keywords

Garbage collection, performance, modeling, Java