A New Data Compression Technique for Event Based Program Traces

A. Knüpfer

Abstract

The paper presents an innovative solution to the problem of the very huge data sets that are regularly produced by performance tracing techniques - especially on HPC programs. It designs an adapted data compression scheme that takes advantage of regularities frequently found in program traces. Algorithms to reveal repetition patterns in a programs call structure and run time behavior are discussed in detail, solutions to some problems arising on practical application are addressed as well. Two examples demonstrate the capabilities of the approach and document its behavior. Finally, some thoughts are given regarding how the patterns revealed in the process of data compression may assist the automatic analysis of traces.