Institut für Systemsoftware

O.Univ.Prof. Dr. Hanspeter Mössenböck

JOHANNES KEPLER
UNIVERSITÄT LINZ
Netzwerk für Forschung, Lehre und Praxis

Optimized Strings for the Java HotSpot™ VM

Diplomaufgabe für Christian Häubl
Matr.-Nr.: 0455105

In Java, a string consists of two parts: the String object itself with metadata such as the string length, and the character array that holds the actual data. Because the character array never changes (strings are immutable), it would be possible to combine the two objects to one object with special support from the virtual machine. Implement this optimization for the Java HotSpot™ VM of Sun Microsystems and the class library of the JDK 7.

Currently, two string objects can share the same character array. This optimizes some string operations, but hinders the grouping of the string and the array. Therefore, it is first necessary to change the string class so that each string has its own character array.

A possible solution for the grouping is a bytecode preprocessing step that modifies all methods that allocate String objects, and all methods that access the contents of a String. This requires changes in the class loader, the interpreter and the client just-in-time compiler.

Dynamic features of Java, like reflection, JNI, and the debugging architecture, are also affected by the optimization. It is not necessary to implement all these parts completely. However, it should be possible to execute benchmarks like SPECjvm98 and SPECjbb2005 with the optimization enabled.

Der Fortgang der Arbeit ist in 14-tägigem Abstand mit dem Betreuer zu besprechen. Für die Ausarbeitung der schriftlichen Arbeit sind die Richtlinien des Instituts für Systemsoftware zu beachten.

Programmiersprache: Java

Nähere Auskünfte: Dipl.-Ing. Christian Wimmer