

1. Übung Komplexe Softwaresysteme 2

Abgabe bis 4. März (vor der Übung) durch Hochladen von plain text Dateien in git (eine Datei pro Aufgabe, max. eine A4 Seite pro Aufgabe. Sourcecode Dateien sind von diesem Limit ausgenommen). Gelöste Beispiele müssen präsentiert und erklärt werden können.

1. Wie kann man die Ausführungszeit eines Programms oder einer Funktion messen? Welche Alternativen gibt es? Was sind die Vor- und Nachteile von `time()`, `clock_gettime()`, `clock()`, `rdtsc`? Studieren Sie dafür die entsprechenden man pages (<http://linux.die.net/man/>) bzw. Onlineressourcen. Was ist processor time? Was ist wall clock time? Welche Möglichkeiten gibt es um auch den Speicherverbrauch eines Prozesses zu messen? Experimentieren Sie mit unterschiedlichen Methoden (z.B.: Messen Sie für verschiedene n wie lange `sleep(n)` tatsächlich dauert. Siehe `man 3 sleep`) und schildern Sie ihre Erkenntnisse in Textform (keine Sourcecode Abgabe erforderlich).
2. (Abgabe bis 11. März vor der 2. Übung) Implementieren Sie einen Min-Heap für Werte des Typs `double` mittels eines Arrays (siehe Vorlesung) und vergleichen sie die Ausführungszeit eines einfachen Experiments mit der Implementierung der Standard Template Library (Alternative: Boost an Stelle von STL. Siehe www.boost.org). Das experimentelle Setup lautet wie folgt: Wir interessieren uns für die Performance der push Operation. Messen Sie die Zeit für das Einfügen von 100, 1.000, 10.000, 100.000, 1.000.000 und 10.000.000 Elementen. Plotten Sie einen Graphen der Ihre Implementierung mit der STL Implementierung (bzw. Boost) vergleicht. Die x-Achse sei die Anzahl der Elemente und die y-Achse die Ausführungszeit. Verwenden Sie dafür ein Plotprogramm ihrer Wahl. Im weiteren Verlauf der Übungen werden wir Gnuplot verwenden. Eine Einführung zu Gnuplot folgt in der ersten Übungseinheit. Seien Sie jedoch ermutigt sich bereits in der Zwischenzeit mit Gnuplot zu befassen.
3. Stellen Sie sich folgendes Szenario vor: Sie arbeiten für einen grossen Suchmaschinenbetreiber und bekommen den Auftrag eine Textdatei mit e-Mail Adressen zu sortieren (ein Eintrag pro Zeile, das Ergebnis ist eine neue Datei). Die Datei ist 32 Terabyte gross und über ein Netzwerkdateisystem angebunden. Ihr Auftraggeber möchte von Ihnen wissen wie lange dieser Vorgang in etwa dauern wird und welche Ressourcen Sie dafür benötigen. Wählen Sie realistische Grössen für Durchsatz und Latenz von Netzwerk, Festplatten, Arbeitsspeicher, CPU(s)... Argumentieren Sie Laufzeitverhalten und Speicherverbrauch Ihres Ansatzes.