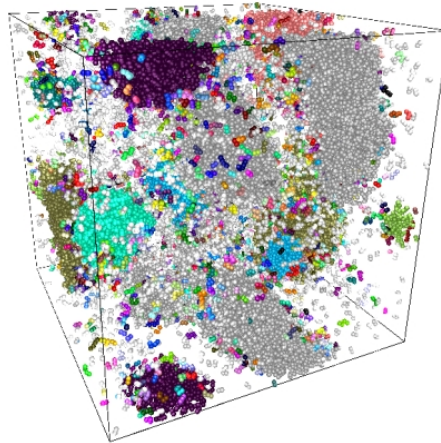


Projektbeschreibung zum IDP “Numerische Aspekte bei Molekulardynamik-Simulationen”

Zur Analyse bzw. Vorhersage thermodynamischer Stoffdaten, wie Druck oder potentielle Energie eines Stoffes, gewinnen Molekulardynamik-Simulationen in der Verfahrenstechnik immer größere Bedeutung. Dabei wird beispielsweise ein Gasgemisch mit Millionen von einzelnen Teilchen auf molekularer Ebene simuliert.



Zu Beginn eines Zeitschritts der Simulation sind die Positionen und Geschwindigkeiten der Moleküle gegeben. Aus den Positionen können Kräfte zwischen den Molekülen berechnet werden, gemäß diesen werden die Moleküle dann bewegt. Das Herzstück des Algorithmus ist dabei das numerische Lösen der Bewegungsgleichung

$$F = m \cdot \ddot{x}.$$

Ein solches Simulationsprogramm wurde mit dem Programm `ls1` realisiert. Um den Berechnungsaufwand bei kurzreichweitigen Paar-Wechselwirkungen gering zu halten, wurde der Linked-Cell-Algorithmus implementiert. Dabei wird das Simulationsgebiet in Zellen unterteilt, sodass zum Finden von wechselwirkenden Teilchen nur noch jeweils benachbarte Zellen durchsucht werden müssen. Die Zeitintegration erfolgt bislang durch den Leap-Frog-Algorithmus [2].

In bisherigen Arbeiten wurden der Programmcode hybrid [4] parallelisiert sowie diverse Lastbalancierungsstrategien erarbeitet [1]. Im Rahmen des IDPs sollen die Voraussetzungen geschaffen werden, dass `ls1` problemlos auf heterogene Architekturen mit Beschleunigerhardware (z.B. Graphikkarten) portiert bzw. auf `ccNuma`-Architekturen optimiert werden kann. Vorbereitend sollen insbesondere die numerischen Aspekte sowie die zugrundeliegende Algorithmik untersucht werden. Dabei ist das IDP in weitere Arbeiten bezüglich der Nutzung moderner Multicore-Architekturen im HPC sowie der Weiterentwicklung der numerischen Algorithmen eingebettet, sodass hier über die einzelnen Arbeitsgebiete hinaus ein konkreter Bezug zur Mathematik als Anwendungsfach hergestellt wird.

Insgesamt gliedert sich das Projekt in drei Teile:

- Vergleich verschiedener numerischer Zeitintegrationsverfahren und Untersuchung ihrer Eignung für die konkrete Anwendung.
- Untersuchung von Graphikkarten als Rechenbeschleuniger hinsichtlich der numerischen Genauigkeit, bezogen auf die Molekulardynamik-Simulation.
- Analyse der algorithmischen Komplexität von Datenstrukturen, die auf effizientes Speicherlayouts optimiert sind.

1 Implementierung und Vergleich weiterer Zeit-Integrationsverfahren

Ein Teil des Projekts besteht aus der Implementierung weiterer Zeit-Integrationsverfahren sowie deren Analyse und Bewertung.

Zur Lösung der o.g. Differentialgleichung existieren verschiedene Algorithmen, die sich insbesondere bezüglich numerischer Stabilität und Rechenaufwand unterscheiden. Im Rahmen der Arbeit sollen die Eigenschaften ausgewählter Verfahren verglichen und diese auf ihre Anwendbarkeit für die gegebene Anwendung hin untersucht werden.

Für die Evaluation und Implementierung der unterschiedlichen Algorithmen ist ein tiefgehendes Verständnis der simulierten physikalischen Phänomene notwendig. Darüber hinaus ist eine detaillierte Kenntnis der Diskretisierung und Lösung von Differentialgleichungen nötig. Zur Bewertung wird grundlegendes Wissen aus der Numerik benötigt.

2 Untersuchung der numerischen Eignung von Graphikkarten

Weiterhin soll die Eignung von Graphikkarten als Beschleuniger für die Molekulardynamik-Simulation evaluiert werden. Graphikkarten werden von diversen Gruppen verstärkt auf diesem Gebiet eingesetzt, ihr Potential soll auch für `ls1` nutzbar gemacht werden.

Zunächst soll die erforderliche Rechengenauigkeit für Molekulardynamik-Simulationen in der Thermodynamik analysiert werden. Danach werden der Algorithmus prototypisch

für Graphikkarten implementiert sowie die numerischen Eigenschaften der CPU / GPU - Implementierungen verglichen. Falls nötig, sollen Vorschläge erarbeitet werden, wie der bestehende Algorithmus zu ändern ist (z.B. durch Verwendung kürzerer Zeitschrittweiten oder durch Verwendung eines Algorithmus aus dem vorstehenden Arbeitspaket).

Hier soll insbesondere dem Umstand Rechnung getragen werden, dass es für eine hohe Rechenleistung auf Graphikkarten nötig ist, nur Fließkomma-Arithmetik in einfacher Genauigkeit zu verwenden. Darüber hinaus soll der Einfluss des auf Graphikkarten verwendeten nicht-fehlerkorrigierenden Speichers auf die Numerik [3] berücksichtigt werden.

3 Komplexitätsanalyse einer speicher-optimierten Datenstruktur

Im dritten Teil soll eine Komplexitätsanalyse einer dynamischen Datenstruktur ähnlich einer `Unrolled Linked List` durchgeführt werden. Derzeit werden die Moleküldaten in einer Liste im Speicher gehalten, sie werden über Referenzen in Zellen einsortiert. Zur effizienten Behandlung auf aktuellen Systemen sollten die Moleküldaten jedoch physisch in Zellen gehalten werden. Da sich Moleküle im Laufe der Simulation bewegen, wechseln sie ihre Zellen. In einer auf Lokalität optimierten Datenstruktur entsteht daher Aufwand zur Reorganisation der Daten (Kopien der Moleküle nötig, Anpassung der Datenstruktur). Dieser zusätzliche Aufwand soll in Abhängigkeit von zwei typischen Anwendungsszenarien geklärt werden. Dafür soll jeweils eine Komplexitätsanalyse durchgeführt werden, die Anhand einer Implementierung validiert wird.

4 Dokumentation

Zur Dokumentation ist zum einen dokumentierter Sourcecode abzugeben. Weiter ist jeweils eine kurze Ausarbeitung, die die wichtigsten theoretischen Erkenntnisse enthält, gefordert. Darüber hinaus muss die Ausarbeitung die Ergebnisse geeigneter Leistungsmessungen beinhalten, welche die theoretischen Erkenntnisse fundieren

5 Vorlesung

Als begleitende Vorlesungen zum IDP werden die Vorlesungen "Introduction to Scientific Computing I" (3 ETCS) und "Introduction to Scientific Computing II" (4 ETCS) ausgewählt. Um eine stärkere Gewichtung der praktischen Arbeit zu ermöglichen, schlagen wir vor, aus "Scientific Computing I" nur einen Teil des Stoffes (Gewöhnliche Differentialgleichungen) für prüfungsrelevant zu erklären und diesen mit lediglich 1 ETCS-Punkt zu gewichten. Damit zählt der gesamte Vorlesungsteil mit genau 5 ETCS.

References

- [1] M. Buchholz. *Framework zur Parallelisierung von Molekulardynamiksimulationen in verfahrenstechnischen Anwendungen*. PhD thesis, Technische Universität München, 2010. to be published.
- [2] D. Fincham. Leapfrog Rotational Algorithms. *Molecular Simulation*, 8:165–178, 1992.
- [3] Imran S. Haque and Vijay S. Pande. Hard data on soft errors: A large-scale assessment of real-world error rates in gpgpu. *CoRR*, abs/0910.0505, 2009.
- [4] Johannes Weißl. Erweiterung einer MPI-basierten Molekulardynamik-Parallelisierung zu einer hybriden Parallelisierung. Bachelorarbeit, Fakultät für Informatik der Technischen Universität München, 2009.