

Zwischenbericht 1

Zusammenfassender Bericht	01IH08003 A-E / SKALB
Lattice-Boltzmann-Methoden für skalierbare Multi-Physik-Anwendungen	
Laufzeit des Vorhabens:	01.01.2009 – 31.12.2011
Berichtszeitraum:	01.01.2009 – 30.06.2009

Das Verbundprojekt *Lattice-Boltzmann-Methoden für skalierbare Multi-Physik-Anwendungen (SKALB)* wird vom Bundesministerium für Bildung und Forschung (BMBF) im Rahmen der Initiative *HPC Software für skalierbare Parallelrechner* unter dem Förderkennzeichen 01IH08003 gefördert. Projektstart von SKALB war der 1. Januar 2009. Während der dreijährigen Förderung haben sich die Projektpartner Regionales Rechenzentrum Erlangen (RRZE, Projektkoordinator) und Lehrstuhl für Systemsimulation (LSS) von der Universität Erlangen-Nürnberg, Hochleistungsrechenzentrum Stuttgart (HLRS) der Universität Stuttgart, Institut für rechnergestützte Modellierung im Bauingenieurwesen (iRMB) der Technischen Universität Braunschweig, Lehrstuhl für Angewandte Mathematik und Numerik zusammen mit dem IT und Medien Centrum der Technische Universität Dortmund (TU-Do) sowie IANUS Simulation GmbH (Dortmund) als KMU vorgenommen, gemeinsam die effiziente Implementierung und Weiterentwicklung von Lattice-Boltzmann basierten Strömungslösern zur Simulation komplexer Multi-Physik-Anwendungen auf Rechnern der Petascale-Klasse voranzutreiben. Dieser Bericht faßt wesentliche Aktivitäten und Ergebnisse des ersten Projekthalbjahres zusammen.

Lattice-Boltzmann-Methoden haben sich in den letzten Jahren als Standardverfahren für numerische Strömungssimulation etabliert. Besonders bemerkenswert ist, dass dieser Ansatz schnell breite Verwendung auch jenseits der klassischen CFD Anwendungen gefunden hat und so zum Beispiel in der Verfahrenstechnik oder der Biomedizin eingesetzt wird. Deutschland konnte sich neben den USA als weltweit führend bei der numerischen Umsetzung der Methode insbesondere im Bereich der Hochleistungsrechner etablieren. Um diese führende Stellung auch bei den nun absehbaren technologischen Veränderungen erhalten zu können, haben fünf Gruppen aus ganz Deutschland mit ausgewiesener Expertise in den Bereichen Lattice-Boltzmann-Methode und Höchstleistungsrechnen für das Verbundprojekt SKALB zusammengeschlossen. Bestehende Lattice-Boltzmann-Verfahren sollen so zur Lösung praxisrelevanter Multi-Physik-Anwendungsfälle auf hochskalierenden Parallelrechnern erweitert und optimiert werden. Dabei werden alle Schritte vom Preprocessing über die Simulation bis hin zur Visualisierung adressiert. Mit Blick auf neue und sich abzeichnende (heterogene) Multi-/Many-Core Architekturen sollen schließlich Potenziale und Programmieransätze exemplarisch evaluiert werden. Die umfangreiche Beteiligung assoziierter Unternehmen sichert dem Projekt den Zugriff auf neueste Hardwaretechnologien sowie den Transfer der methodischen Erkenntnisse und Simulationstechniken in die Wirtschaft.

1 Organisatorisches

Geeignete Projektmitarbeiter zu finden, war für einige Projektpartner schwierig, so dass die Projektstellen z.T. erst verspätet besetzt werden konnten/können. Durch interne Umschichtungen konnten die für das erste Projekthalbjahr gesetzten Ziele aber dennoch erreicht werden, bzw. Anhängigkeiten zwischen einzelnen Projektpartnern durch vorziehen und zurückstellen von Teilaufgaben ausgeglichen werden.

Im Rahmen eines eintägigen, nicht-öffentlichen Kick-Off-Meetings am 16.2.2009 wurden neben organisatorischen Dingen bereits diverse technische Punkte besprochen und abgestimmt, die alle Projektpartner betreffen. Im Rahmen einer großen Videokonferenz am 3.4.2009 erfolgte eine erste Abstimmung hinsichtlich den Anforderungen und Möglichkeiten einer adaptiven Verfeinerung, die im zweiten und dritten Projektjahr eine zentrale Rolle einnehmen wird. Da die Datenstrukturen darauf aber bereits jetzt abgestimmt werden müssen, war der erste Informations- und Erfahrungsaustausch bereits jetzt sehr wichtig. Im Rahmen eines zweitägigen vom LSS organisierten Workshops in Erlangen, an dem LSS, RRZE, HLRS und iRMB mit insgesamt rund 15 Teilnehmern wurde eine einheitliche Begriffsdefinition (vgl. auch Abschnitt 2.2) vereinbart sowie nach Synergien in waLBerla (LSS) und VirtualFluids (iRMB) gesucht. Weitere persönliche oder virtuelle Treffen im kleineren Kreis zu spezifischen Themen ergänzten diese großen Treffen regelmäßig.

Erste im Rahmen von SKALB erzielte Ergebnisse wurden von den einzelnen Gruppen bereits auf internationalen Konferenzen vorgestellt. Besonders hervorzuheben ist der Beitrag von Stefan Donath (LSS), der Ende Mai 2009 auf der Parallel Computational Fluid Dynamics Konferenz in Moffett Field (Kalifornien, USA) den Best Paper Award erhalten hat.

Zur Projektbearbeitung wurden von den einzelnen Gruppen Rechenzeitanträge auf nationaler, europäischer und internationaler Ebene gestellt. Die Anträge von RRZE, LSS und iRMB im Rahmen der DEISA DECI-5 Initiative wurden alle genehmigt, jedoch wurde die beantragte Rechenzeit jeweils um den Faktor vier gekürzt. Rechenzeitanträge für das JUROPA-System am FZ-Jülich wurden z.T. noch stärker gekürzt. In wieweit zumindest die geplanten hochparallelen Skalierungstests trotz dieser massiven Kürzungen durchgeführt werden können, müssen die kommenden Projektphasen zeigen.

2 Wissenschaftliche Ergebnisse

2.1 WP1: Untersuchungen an prototypischen Benchmarkkernen und Optimierung der LB-Applikationen

Codeoptimierung für Standardarchitekturen

Am RRZE und LSS wurde in Q1/Q2 intensiv an der Erstellung benchmarkfähiger Versionen der beiden Produktionscodes ILBDC-Solver und waLBerla gearbeitet. Hintergrund ist

die Anfrage des LRZ München für einen LB basierten Strömungscode als Benchmarkprogramm für die Ausschreibung des nächsten Rechners der im Jahr 2011 installiert werden soll und auf mehrere PFLOPs Spitzenleistung ausgelegt ist. Bei den Arbeiten gelang es u.a. im ILBDC-Kernel erstmalig “non-temporal stores” zu implementieren. Diese Instruktionen reduzieren den Datentransfer zum Hauptspeicher auf Cache-basierten Rechnern um ein Drittel und können analog die Rechenleistung um bis zum 50 % steigern. Die Möglichkeit einer Integration dieser Technik in den Produktionscode wird derzeit evaluiert. Weiters wurde der ILBDC-Produktionscode an den neuen NEC SX9 Vektorrechner angepaßt, wobei hier insbesondere der Einsatz einer undokumentierten Compilerdirektive entscheidend war. Für waLberla wurden umfangreiche Skalierungsuntersuchungen auf bis zu 8.000 Prozessoren der SGI Altix am LRZ durchgeführt.

Evaluierung paralleler Programmiermodelle

Nahezu alle parallelen Applikationen im wissenschaftlich-technischen Umfeld verwenden eine reine MPI Parallelisierung. In SKALB wurden nun verschiedene Arbeiten gestartet die untersuchen sollen inwieweit neue oder alternative Parallelisierungsansätze geeignet sind die zu Grunde liegende Hardwarestruktur effizienter im parallelen Code nutzen zu können. Dabei werden sich ergänzende Aspekte untersucht. Das HLRS hat eine prototypische Version eines LBM-Kernels in CoArray-Fortran (CAF) implementiert und deckt damit den Bereich der PGAS Sprachen ab. Erste verlässliche Performanceergebnisse werden für die nächste Berichtsperiode erwartet, wenn derzeit laufende Arbeiten mit der Firma CRAY abgeschlossen sind. Die hybride MPI/OpenMP Parallelisierung Gesamtlösers wurde an der TUD gestartet und haben zum Ziel u.a. die numerischen Effizienz durch die Verwendung größerer Domänen zu verbessern. Die hybride MPI/OpenMP Parallelisierung der Produktionscodes ILBDC und waLberla wurde in Erlangen im Wesentlichen abgeschlossen. Hier liegt der Fokus darauf bei extrem hohen Prozessorzahlen den hohen Aufwand für die Gitterpartitionierung reduzieren zu können.

Multi-Core spezifische Entwicklungen

Eine charakteristische Eigenschaft moderner Multi-Core-Prozessoren ist ein von mehreren Rechenkernen gemeinsam genutzter Cache. Klassische Parallelisierungsmethoden ignorieren im Wesentlichen die sich daraus ergebenden Möglichkeiten. Ein Ansatz den gemeinsamen Cache für iterative Löser mit regulärem Stencil explizit zu nutzen, ist die Implementierung einer sog. „Wellenfront-Parallelisierung“. Dabei können im Daten im Cache von mehreren Kernen in der gleichen Cachegruppe wiederverwendet werden. Dieses Prinzip wurde zu Beginn der Berichtsperiode am RRZE erstmalig erfolgreich auf einen einfachen Stencil-basierten Code angewendet (Jacobi Iteration in 3D) und anschließend für einen Lattice-Boltzmann-Kernel umgesetzt. Beispielhaft sei der Ansatz kurz anhand der LBM erklärt: Ein Zeitschritt einer Lattice-Boltzmann-Zelle modifiziert insgesamt drei benachbarte Ebenen. Bei der Wellenfront-Parallelisierung lädt Thread 1 diese drei von Thread 0 modifizierten Ebenen direkt aus dem Cache, wie in Abb. 1 zu sehen ist. Im Gegensatz

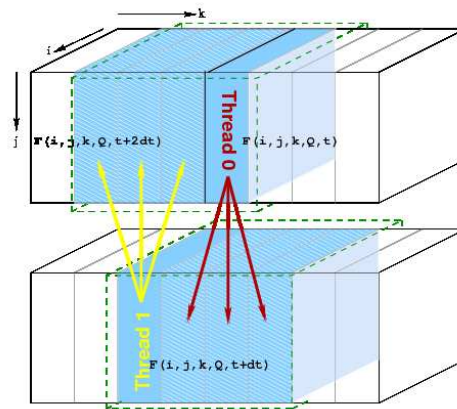


Abbildung 1: Zwei Wellenfronten, die Daten über einen temporären Puffer austauschen.

zu den klassischen zeitlichen Blockingansätzen hat der vorliegende Ansatz den Vorteil, dass keine kubischen Blöcke im Cache gehalten werden müssen, sondern nur vier zusammenhängende Gitterebenen der beiden Zeitschritte. Eine weitere Optimierung lenkt den Zugriff auf den Zeitschritt $t + dt$ auf ein temporäres Feld um (unterer Block in Abb. 1), was die pro Zeitschritt notwendigen Hauptspeicherzugriffe um insgesamt zwei Drittel und den Hauptspeicherbedarf auf nahezu die Hälfte reduziert. Die Ergebnisse zur Wellenfront-Parallelisierung wurden in Vorträgen in Berkeley sowie auf der Conference on Parallel CFD (ParCFD) vorgestellt und wurden als Vortrag sowie zur Publikation auf der IEEE Konferenz COMPSAC 2009 akzeptiert.

Das Thema Performancemodellierung wurde konsequent bei allen Arbeiten im Rahmen des vorliegenden WPs durchgeführt und konnte u.a. bei der „Wellenfront-Parallelisierung“ wichtige Hinweise zur effizienten Implementierungsstrategie geben.

2.2 WP2: Weiterentwicklung von LB-Methoden für praktische Anwendungen auf hochskalierenden Systemen

Datenstrukturen

Zentraler Punkt für die Entwicklung hochparalleler LBM Applikationen mit statischer und/oder dynamischer Lastbalancierung ist die sorgfältige Definition und Implementierung geeigneter Datenstrukturen. In der abgelaufenen Förderperiode wurden von den beteiligten Gruppen entsprechende Analysen der Anforderungen ihrer Applikationen und physikalischen Problemstellungen angefertigt. In einem zweitägigen Projekttreffen aller Partner Anfang Juni in Erlangen wurden diese dann intensiv diskutiert. Ziel ist eine enge

Koordination der verschiedenen Arbeiten. Hinsichtlich ein der Kompatibilität der Datenstrukturen die einen Austausch von Daten und Weiterverwendung von Forschungsergebnissen zwischen den Projekten sicherstellen soll wurde auf dem Treffen ein gemeinsames Glossar entwickelt:

Knoten: Geometrisches Punktelement, das mindestens einen Satz von Verteilungsfunktionen f_i trägt.

Block: Menge von *Knoten*, die jeweils indizierbar durch Indizes ($B(x, y)$ bzw. $B(x, y, z)$) mit homogenem Knotenabstand sind.

Patch: ein Verbund von *Blöcken* (mit identischer Physik), die zumindest über einen *Link* ($e_i \Delta t$) zusammenhängen.

Link: Verbindung von zwei *Knoten*, zwischen denen eine Propagation stattfindet.

Subdomain: Ein Verbund von (nicht notwendigerweise zusammenhängenden) *Patches*, die von genau einem *Prozess* bearbeitet werden.

Domain: Ein Verbund von *Subdomains*, die von einem oder mehreren *Prozessen* bearbeitet werden.

Basierend auf diesem Glossar arbeiten derzeit die Gruppen in Dortmund, Braunschweig und Erlangen an sich ergänzenden Verfeinerungstechniken sowie an adäquaten Abstraktionsmechanismen für den Datenaustausch.

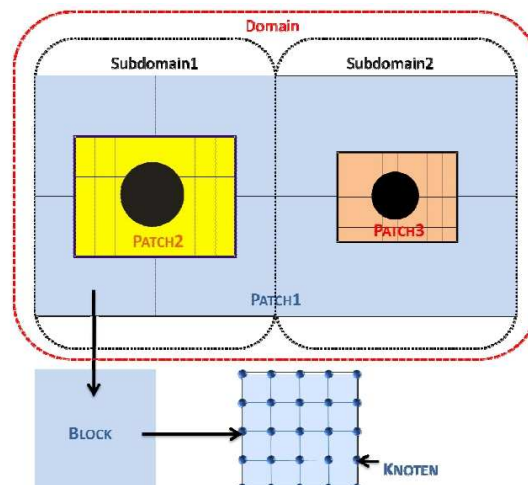


Abbildung 2: Schematische Darstellung der Datenstrukturhierarchie

Unabhängig von der Definition der Datenstrukturen konnte für den ILBDC-Solver bereits mit einem einfachen, alternativen Verfahren zur Gebietszerlegung begonnen werden. Es basiert auf der Idee raumfüllende Kurven als Basis für die Partitionierung von LBM-Strömungslösern zu verwenden. Dieser Ansatz ist herkömmlichen Graph-basierten Partitionierungsmethoden zumindest ebenbürtig ist; er benötigt jedoch wesentlich weniger Rechenleistung und Speicher. Derzeit liegt der Algorithmus in einer seriellen Version vor. Die vollständige Parallelisierung und Integration in den ILBDC-Solver ist im Gange.

Visualisierung

Zur Anbindung des auf C++-basierenden Computational Steering Frameworks *Steereo*¹ an in Fortran implementierte Lattice-Boltzmann-Strömungslöser wurde ein Interface für Fortran90/95 an das Framework geschaffen. Dieses Interface besteht aus einem Wrapper für die notwendigen C++-Funktionen, sowie einer Containerklasse, in denen die Instanzen der grundlegenden Steuerungsklassen verwaltet werden und auf welche die Fortran Simulation, jeweils eine Referenz erhält. Darüberhinaus wurde *Steereo* zur grundlegenden Unterstützung von parallelen Simulationen ausgebaut, wobei jedoch nur der root-Prozess mit der Steuerungsanwendung kommuniziert.

2.3 WP3: Verbesserte numerische Ansätze

Die Realisierung und Analyse “einfacher” (klassischer) LBM Codes auf strukturierten Gittern unter Einsatz von GPUs und CELL Prozessoren wird derzeit im Rahmen von zwei Diplomarbeiten in Dortmund bearbeitet. Für die Flachwasser-Gleichungen wurden detaillierte LBM Simulationsrechnungen durchgeführt, die mit modernen FEM-Zugängen (weitere Diplomarbeit) verglichen werden: Diese Vergleiche haben prototypischen Charakter für die geplante Weiterentwicklung von Alternativmethoden für die LBM-Gleichungen und ihre Evaluierung bzgl. numerischer Effizienz. Flankierend wird dazu ebenfalls aktuell eine 2D-Version von FEATLBM vervollständigt, die für allgemeine, lokal adaptierte Gitter konzipiert wurde und die eine vollimplizite bzw. direkt stationäre Lösung der LB-Gleichung erlaubt.

2.4 WP4: Hardwarenahe Implementierung auf Nicht-Standardprozessoren

In Braunschweig und Dortmund wurde in Q1/Q2 an der Weiterentwicklung von GPU basierten Lösern gearbeitet. Das iRMB hat hierbei eine intranode parallele Implementierung erarbeitet die ein uniformes Simulationsgitter aufgeteilt, gleichzeitig von bis zu 4 GPUs berechnen lässt. Gerade im Gange ist die Erweiterung hinsichtlich der Kommunikation mehrerer solcher GPU Server mit Hilfe von MPI um massiv parallel auf GPUs in kommenden GPU Clustern rechnen zu können. In Dortmund wurden neben Vergleichsrechnungen von LBM Standardansätzen zwischen GPU und CELL bereits voll unstrukturierte Gitter auf GPUs implementiert. Hierdurch wird die Berechnung von allgemeinen Gittern sowie Finite-Elemente-Räume ermöglicht, wodurch sich Alternativmethoden nun auch auf nicht Standard Hardware untersuchen lassen. Weiterhin wurde die Entwicklung der FEAST Kernroutinen bzgl. Linearer Algebra auf GPUs vorangetrieben. Die Erkenntnisse werden in die SBBLAS Bibliothek einfließen.

¹<https://gforge.hlrsl.de/projects/steereo/>

2.5 WP5: Industriekontakte und Definition von Showcases

IANUS und TU-Do sind dabei, zusammen mit Sulzer und BASF die geplanten Showcases zu definieren. Der Testcase „poröse Medien“ wird durch das RRZE in enger Abstimmung mit dem dem Erlanger Exzellenzcluster „Engineering of Advanced Materials“ sowie dem Max-Planck-Institut für Dynamik Komplexer Systeme in Magdeburg erarbeitet.

Dank der Bereitstellung von Nehalem (Quad-Core) und Dunnington (Hexa-Core) „Early-Access“ Systemen durch die Firma Intel bereits im Vorfeld (Q3/2008) des Projektstarts konnte das Projekt zum Teil schon deutlich vor offizieller Verfügbarkeit der Systeme mit Multi-Core spezifischen Arbeiten beginnen.

2.6 Zusammenfassung des Projektfortschritts

Die Akquisition geeigneter Mitarbeiter sowie nicht vorhersehbare personelle Veränderungen haben zu kleineren Verschiebungen gegenüber dem Projektplan notwendig gemacht. Diese haben aber keinen Einfluß auf den Projektfortschritt im Ganzen.

Die Projektentwicklung im abgelaufenen Projektzeitraum rechtfertigt die Entscheidung, vorhandene Codes mit sich ergänzenden Schwerpunkten weiterzuentwickeln und durch enge Kooperation Synergien zu heben. Profitiert hat davon vor allem das WP1, wo durch den personellen Ausbau der Gruppen über die Projektmittel umfangreiche Arbeiten in die Wege geleitet und abgeschlossen werden konnten. Erwartungsgemäß befinden sich die WP 2-4 in der Anlaufphase: Hier sind Forschungsarbeiten im Gange die zum Teil sehr fundamentale Ansätze verfolgen und deren technische Umsetzung erst nach entsprechenden theoretischen Vorarbeiten sinnvoll möglich ist. WP 5 liegt ebenfalls im Projektplan. Hier wurden vor allem Kommunikationsplattformen etabliert. Konkrete Ergebnisse aus den Showcases sind - analog dem Projektplan - nicht vor Mitte 2010 zu erwarten.

Von SKALB ganz oder teilweise geförderte Arbeiten wurden im Berichtszeitraum in zahlreichen Vorträgen vorgestellt. Es sind bereits auch erste wissenschaftliche Publikationen akzeptiert bzw. veröffentlicht worden. Eine Liste ist dem Anhang beigefügt und wird auch auf den Webseiten verfügbar sein.

3 Stand des Vorhabens, Erreichbarkeit der Zeile, Zielsetzung und Verwertungsplan

Gegenüber dem Antrag haben sich keine nennenswerten Veränderungen ergeben.