

Invasive Computing - Paralleles Betriebssystem einer SFB/TRR-Projektinitiative

Wolfgang Schröder-Preikschat, FAU Erlangen-Nürnberg

Die Universitäten von Erlangen (FAU), Karlsruhe (KIT) und München (TU) planen einen überregional aufgestellten Sonderforschungsbereich zum Thema *Invasive Computing*, gefördert durch die DFG (SFB/TRR 89). Die von Erlangen ausgehende Initiative befindet sich derzeit in der fortgeschrittenen Antragsphase, die Einrichtungsbegutachtung wird im Frühjahr 2010 an der FAU stattfinden. Zweck dieses Vorhabens ist die gemeinsame Erforschung innovativer Rechnerarchitekturen, mit den Schwerpunkten der Hard- und Softwarearbeiten auf Analyse, Entwurf, Implementierung und Validierung komplexer, heterogener massiv-paralleler Systeme.

Invasive Computing versteht sich als ein Paradigma zur betriebsmittelgewahren (*resource aware*) Programmierung heterogener, massiv-paralleler Systeme. Weit über Mehr- oder Vielkernprozessoren hinausgehend, schließt dieser Ansatz insbesondere auch Spezialhardware zur dedizierten Ausführungsunterstützung paralleler Programme mit ein, wie etwa zur Beschleunigung schleifenparalleler Rechenprozesse vorgesehene schwach programmierbare Rechenfelder. Betriebsmittelgewahre parallele Programmierung soll dabei unabhängig von der jeweils gewählten Abstraktionsebene (d.h., semantischen Lücke) mit syntaktisch ähnlichen, jedoch je nach gewählter Abstraktionsebene eben semantisch unterschiedlichen Elementaroperationen ermöglicht werden. Basisbefehle dabei sind *invade*, *infect* und *retreat*, die in Hardware, Software oder in hybrider Art implementiert sind. Bereitgestellt werden abstrakte und reale Maschinen sowie Programmiersysteme, die die Entwicklung und Ausführung invasiv-paralleler Programme unterstützen. Eingebettet wird dieses Programmiermodell u.a. in X10, einer von IBM entwickelten experimentellen, typsicheren und objektorientierten Programmiersprache (massiv) verteilt-paralleler Systeme.

Die Betriebssystemarbeiten fokussieren auf die Entwicklung eines Laufzeitunterstützungssystems (*runtime support system*) zur invasiv-parallelen Programmierung, einerseits ausgelegt als Stammbetriebssystem und andererseits als Erweiterung/Spezialisierung eines Wirtsbetriebssystems (FreeBSD, Linux). Untersucht werden sollen Methoden, Prinzipien und Mechanismen zur anwendungsorientierten Erweiterung, Konfigurierung und Anpassung invasiv-paralleler Rechensysteme. Betriebssystemabstraktionen und -strukturen für heterogene massiv-parallele Rechner werden mit der Zielrichtung erforscht, durch strikt problemorientierte Softwareinfrastrukturen die funktional dedizierte und kontrolliert-invasive Nutzung der Verarbeitungseinheiten der Hardware zu ermöglichen. Zur lokalen wie auch globalen, gemeinschaftlichen Betriebsmittelverwaltung soll ein agentenbasierter Ansatz zum Einsatz kommen, der anwendungsbezogen ausgerichtet ist. Dabei werden Verhandlungs- und Anreizmodelle erforscht und zur Betriebsmittelvergabe genutzt, um die Systemleistung der invasiv-parallelen Ausführungsplattform für die Benutzerschaft

zu maximieren. Unterstützt durch statische Programm- und dynamische Laufzeitanalyse soll Vorabwissen zum Betriebsmittelbedarf in Raum und Zeit aufgebaut und fortgeschrieben werden, so dass dem Betriebssystem durch geschickte Einplanung und Einlastung die Vorbeugung bzw. Vermeidung von Wettstreitigkeiten unter gleichzeitigen Prozessen weitestgehend möglich ist.

Querschneidender Aspekt in den Betriebssystemarbeiten der ersten Förderphase (Jahre 1-4) ist Vorbeugung/Vermeidung von Wettstreitigkeit (*contention*), durch Schaffung von Lokalitäten, und Latenzverbergung (*latency hiding*), durch Ausnutzung von Parallelität im Betriebssystem selbst. In der zweiten Förderphase (Jahre 5-8) wird allgemein Virtualisierung den Schwerpunkt bilden. Dabei geht es weniger um die Bereitstellung verschiedener Ausführungsumgebungen auf demselben Rechensystem, als viel mehr um die Abstraktion von den physikalischen Verarbeitungseinheiten hinsichtlich Art, Anzahl und Betriebsart. In der letzten Förderphase (Jahre 9-12) steht Heterogenität im Vordergrund der Arbeiten. Hier ist insbesondere die Integration der im Rahmen des Gesamtvorhabens entwickelten Spezialhardware, z.B. schwach programmierbare Rechenfelder, und deren Verwaltung im Zusammenschluss mit über gemeinsamen oder verteilten Speicher gekoppelten (Gruppen von) Mehr- und Vielkernprozessoren unterschiedlicher Hersteller vorgesehen.

Der Vortrag stellt das Gesamtvorhaben kurz vor und skizziert das allen Arbeiten darin zu Grunde liegende Programmiermodell. Darauf aufbauend werden zwei zentrale Betriebssystemabstraktionen diskutiert, nämlich das Anrecht (*claim*) und die Mannschaft (*team*), die die Schnittstelle zum Programmiersystem bilden sollen, um invasiv-parallele Programme technisch zu repräsentieren und auf die jeweils gegebene Hardware abzubilden. Darüberhinaus werden erste Vorstellungen zum Umgang mit Wettstreitigkeit und Verbergung von Latenzen präsentiert.