

# Kurzfassung

## FGPU: eine Flexible Soft GPU Architektur für Universelles Rechnen auf FPGAs

von  
Muhammed Al Kadi

Die Realisierung eingebetteter Grafikprozessoren (GPUs) auf Field Programmable Gate Arrays (FPGAs) führt zu einer neuen Prozessorfamilie, die die Effizienz und eine einfache Programmierung einer GPU ermöglicht, aber die Flexibilität und die Rekonfigurierbarkeit einer FPGA-Plattform nutzt. Diese Dissertation beschreibt den Entwurf der Hardware sowie das Tool Flows der FGPA-GPU (FGPU): eine konfigurierbare, skalierbare und portierbare GPU, die für FPGAs speziell entwickelt wurde. Sie ist ein open-source 32bit Prozessor, der mit OpenCL programmierbar ist. FGPU ist für universelles Rechnen auf GPUs (GPGPU) gedacht und sie repliziert keine andere Architektur.

Auf einem mittelgroßen Zynq System-On-Chip (SOC) lassen sich bis zu 64 Verarbeitungselemente (PEs) in einer einzigen FGPU mit 250MHz realisieren. In einem Benchmark von 20 Applikationen übertrifft die größte FGPU, die wir implementieren konnten, einen hart verdrahteten ARM Cortex-A9 Prozessor unterstützt von einem 128bit NEON Vektor-Coprozessor im Durchschnitt um einen Faktor von 4x. FGPU unterstützt Gleitkomma-Arithmetik in einfacher Genauigkeit in Hardware oder als emulierte Instruktionen in Software. Im Falle einer harten Gleitkomma-Unterstützung bietet sie im Durchschnitt 2,9x besseren Durchsatz pro Fläche und 11x weniger Energieverbrauch als jedes MicroBlaze-basierte homogene Multi-Processor SoC (MPSoC) an. Im Falle einer emulierten Gleitkomma-Arithmetik überschreitet die durchschnittliche Leistungsverschlechterung 4,5x nicht. Darüber hinaus wurde ein dedizierter Compiler auf Basis der LLVM Entwicklungsumgebung entworfen. FGPU kann über eine PYNQ-basierte Schnittstelle mit Python-Skripten programmiert und gesteuert werden.

Wir beleuchten die Unterschiede zwischen den Ansätzen von FGPU- und High-Level Synthesis (HLS). Unsere Experimente zeigten, dass beide Lösungen ähnliche Performanz liefern. HLS hat eine bessere Flächen- und Energieeffizienz sofern die Aufgabenparameter zur Synthesezeit festgelegt sind. Ansonsten würde die FGPU HLS mit kompakteren und einfacheren Software-Implementierungen übertreffen. Am Ende dieser Dissertation sind unsere ersten Versuche beschrieben, um eine teilweise rekonfigurierbare GPU zu realisieren, deren Hardware sich zur Laufzeit anpassen lässt.