

Diss. ETH No. 19276

Design and Performance Analysis of Multiprocessor Streaming Applications

A dissertation submitted to
ETH Zurich

for the degree of
Doctor of Sciences

presented by
WOLFGANG CHRISTIAN FERDINAND HAID
Dipl. El.-Ing. ETH Zurich
born November 24, 1980
citizen of Austria

accepted on the recommendation of
Prof. Dr. Lothar Thiele, examiner
Prof. Dr. Jan Madsen, co-examiner

2010

Abstract

The processing of regularly structured data streams frequently occurs in the context of real-time audio, video, and digital signal processing in consumer electronics devices, communication systems, and medical systems, to name a few. Software programs developed for this purpose are commonly referred to as streaming applications. The steadily growing computational demand of streaming applications increasingly leads to implementations on heterogeneous single-chip multiprocessors. This trend calls for suitable design automation techniques to deal with the complexity of the resulting systems and raises the questions of how (real-time) streaming applications can be efficiently implemented on multiprocessors and how the timing behavior of these applications can be analyzed.

In this thesis, these questions are tackled by proposing the distributed operation layer (DoL) design flow that automates the implementation and performance analysis of multiprocessor streaming applications. To this end, a streaming application is specified as a dataflow process network, that is, as a network of actors that explicitly communicate over first-in first-out channels. Then, fundamental properties of dataflow process networks are leveraged to automatically transform this high-level specification into a platform-specific implementation and an analytic performance analysis model.

Specifically, the following contributions are presented in this thesis. Regarding the platform-specific implementation, a novel multiprocessor run-time environment for dataflow process networks based on so-called protothreads and windowed first-in first-out channels is introduced which considerably reduces the run-time overhead compared to existing solutions. Regarding the analytic performance analysis, a new method to analyze the timing behavior of actors with complex activation schemes is proposed which extends the modeling scope of compositional performance analysis, a class of state-of-the-art performance analysis methods for distributed real-time systems. Furthermore, approaches for the automated implementation and performance analysis of multiprocessor streaming applications are described. A prototype of the DoL design flow has been developed that is used as a proof-of-concept

to gain insights into the potential and limitations of the proposed design flow.

It is concluded that an approach based on dataflow process networks is a viable strategy for the design and performance analysis of multiprocessor streaming applications. Specifically, the automated implementation and performance analysis relieves software developers and performance analysts of the associated technical details, giving them more time to focus on essential design activities.

Zusammenfassung

Die Verarbeitung regelmässig strukturierter Datenströme kommt häufig bei der (Echtzeit-) Verarbeitung von Audio-, Video- und anderen digitalen Daten vor, beispielsweise in der Unterhaltungselektronik, in Kommunikationssystemen oder in medizinischen Systemen. Software Programme für diesen Zweck werden als "Streaming-Applikationen" bezeichnet. Der stetig steigende Bedarf an Rechenleistung für Streaming-Applikationen führt zunehmend zu Implementierungen auf heterogenen Multiprozessoren. Dieser Trend erfordert geeignete Methoden zur Automatisierung des Entwurfs, um der Komplexität dieser Systeme zu begegnen. Dabei stellen sich insbesondere die Fragen, wie Streaming-Applikationen effizient auf Multiprozessoren implementiert werden können und wie das Zeitverhalten (Performanz) solcher Applikationen analysiert werden kann.

In der vorliegenden Dissertation werden diese Fragen im Rahmen eines Entwurfsablaufs (distributed operation layer — DOL) behandelt, der die Implementierung von Streaming-Applikationen auf Multiprozessoren und deren Performanz-Analyse automatisiert. Dazu wird eine Streaming-Applikation als Prozessnetzwerk spezifiziert, das heisst als Netzwerk von Aktoren, die ausschliesslich über Kanäle kommunizieren. Die fundamentalen Eigenschaften von Prozessnetzwerken werden anschliessend verwendet, um diese Spezifikation automatisch in eine plattformspezifische Implementierung bzw. in ein Modell zur Performanz-Analyse zu transformieren.

Konkret werden die folgende Beiträge in dieser Dissertation präsentiert. Bezüglich der plattformspezifischen Implementierung wird eine Laufzeitumgebung für Prozessnetzwerke vorgestellt, die im Vergleich mit existierenden Lösungen einen deutlich reduzierten Rechenaufwand aufweist. Bezüglich der Performanz-Analyse wird eine neue Methode zur Analyse von Aktoren mit komplexem Aktivierungsschema vorgeschlagen, was den Anwendungsbereich von kompositionellen Analysemethoden, einer aktuellen Klasse von Methoden zur Performanz-Analyse verteilter Echtzeitsysteme, erweitert. Zudem wird die Vorgehensweise für die automatisierte Implementierung und Performanz-Analyse von Streaming-Applikationen beschrieben.

Ein Prototyp des vorgeschlagenen Entwurfsablaufs wurde entwickelt, um einerseits seine Tauglichkeit zu zeigen und andererseits seine Möglichkeiten und Beschränkungen einzuschätzen.

Zusammenfassend wird festgehalten, dass ein auf Prozessnetzwerken basierter Entwurfsablauf eine praktikable Strategie darstellt, um Streaming-Applikationen auf Multiprozessoren zu implementieren und deren Performanz zu analysieren. Insbesondere entlastet die automatische Implementierung und Performanz-Analyse Software-Entwickler und Performanz-Analysten von der Aufgabe, sich mit den diesbezüglichen technischen Details befassen zu müssen, sodass sie sich auf die wesentlichen Entwurfsaufgaben konzentrieren können.