Analysis and Optimization of Fast and Accurate SoC Platform Models

Gunar Schirner (hschirne@uci.edu), Advisor: Prof. Rainer Doemer
Center for Embedded Computer Systems, University of California Irvine
Expected Graduation: December 2007, Topic not presented at DATE or ASP-DAC

I. Introduction

Designing embedded systems and System-On-Chip (SoC) becomes increasingly challenging. The design space to be explored grows with increasing complexity, while at the same time shorter product life cycles require a shorter time-to-market. Addressing this gap has been the aim of system-level research, aiming to raise the level of abstraction.

Abstract models have been introduced to tackle the complexity in the design process. For one, abstract models exhibit tremendous gains in simulation speed, allowing fast validation and extensive design space exploration. For communication in particular, Transaction Level Modeling (TLM) has been proposed. TLM abstracts the communication in a system to whole transactions, abstracting away low level details about pins, wires and waveforms. This results in models that execute dramatically faster than synthesizable, bit-accurate models. This benefit, however, usually comes at the price of low accuracy. In general, TLMs pose a trade-off between an improvement in simulation speed and a loss in accuracy.

Although TLM has been generally accepted as one solution to tackle SoC design complexity, TLMs and their abstraction levels are not clearly defined in the current literature. Furthermore, the TLM trade-off has not been examined in detail.

This thesis contributes with three major aspects. First, it addresses TLM organization and analysis. It proposes a set of abstraction levels, defines test setups and metrics for systematically analyzing TLM and provides detailed analysis of three bus systems. Second, it proposes a novel modeling technique, called Result Oriented Modeling (ROM), which removes the inaccuracy drawback of TLM. Third, it addresses the abstraction of computation and introduces an approach to abstract processor modeling in the context of multi-processor architectures[1].

II. TLM Analysis

Transaction Level Models can be classified according to the granularity with which they handle arbitration and data transfers. We have identified three levels, shown in Figure 1, which correlate with layers of the OSI reference model [2].

Our most abstract model, which we call the TLM, is based on user transactions (data blocks of an arbitrary size, like messages). Our next more fine grained model, the ATLM, is based on bus transactions (bus primitives, such as words or a burst). While these two models are cycle approximate, our Bus Functional Model (BFM) is a cycle accurate and pin accurate bus model that simulates based on bus cycles.

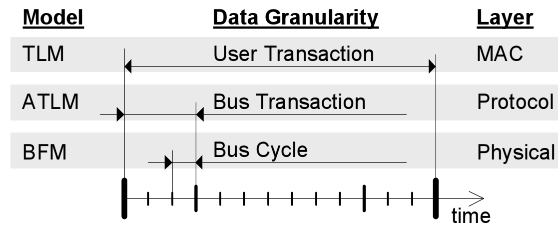

Fig. 1. Model classes and their granularity.

We define appropriate metrics and setups for a systematic quantitative analysis of TLM with respect to simulation performance and timing accuracy. We propose a two node setup to measure model speed in dependency of transaction size. For the accuracy analysis, we define a generic 4 node setup. With a standardized traffic set, me measure accuracy of the simulated timing in dependency of bus contention.

Addressing several classes of embedded communication protocols, we apply our analysis to three common bus architectures, the industry-standard AMBA Advanced High-performance Bus (AHB) as an on-chip parallel bus, the Controller Area Network (CAN) as an off-chip serial bus, and the Motorola ColdFire Master Bus as an example for a custom embedded processor bus. We modeled, validated, and systematically analyzed each bus using our metrics.

Our analysis shows that an abstraction based on a decreasing (coarsening) granularity yields at least an order of magnitude improvement per granularity level. However, TLM abstraction results in a serious loss in accuracy, which defines the TLM trade-off. In general, a model is either fast or accurate. Our fast TLM models with up to 100 MBytes/s simulation bandwidth show an error of up to 47%. Accurate models, on the other hand, are slow. Our Bus Functional Models (BFMs) simulate with less than 0.2 MByte/s bandwidth. More detailed information can be found in [5], [8], [4], [3].

III. Result Oriented Modeling

As analyzed before, traditional abstract modeling poses a trade-off between speed and accuracy. As one major contribution, this thesis introduces a novel modeling technique, Result Oriented Modeling (ROM), that escapes the TLM trade-off. ROM[2] is a modeling approach similar to TLM that hides internal states of bus communication. Unlike traditional TLMs that incrementally model the bus state, ROM avoids modeling of intermediate bus states.

[1] The work presented in the thesis has been integrated into SCE [1], an interactive environment for system-level design. Based on an abstract C-style application specification, SCE automatically generates simulation models according to the architecture and mapping choices made by the designer. It guides the designer through the design space exploration, and generates the final implementation over the selected target platform.

[2] The journal article [9], which describes ROM, has been submitted as the supporting document. It has been accepted for publication in IEEE TCAD.

As shown in Figure 2, ROM uses an *optimistic prediction* approach. Right at the beginning of a transfer, it predicts the end result. ROM constructs a bus schedule taking pending transactions into account, and determines the earliest possible finish time for the requested transaction. While waiting for the predicted time, ROM records any *disturbing influence* (i.e. preemptions by higher priority transfers). If a preemption has occurred, ROM re-calculates the bus schedule at the end of the predicted time, adjusts the prediction as a *corrective measure*, and waits again to achieve 100% accuracy.

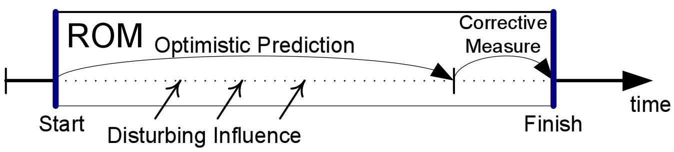

Fig. 2. Generic ROM concept.

To demonstrate the ROM concept, we apply ROM two complementary busses: the on-chip AMBA AHB and off-chip CAN, and compare them against traditional layer-based models. Figure 3 summarizes the results by showing the inaccuracy in simulated timing in dependency of the achievable simulation speed (see [9] for details).

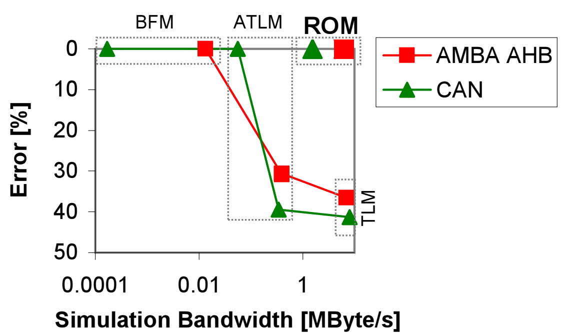

Fig. 3. ROM beats the TLM Trade-Off.

Our experimental results show the clear benefits of ROM. Traditional models are either accurate but slow with less than 0.4 MBytes/s, or they are fast (i.e. the TLMs) but exhibit an error of more than 35%. The proposed ROM, on the other hand, escapes the TLM trade-off for both bus systems. ROM delivers highest speed (reaching near TLM performance) and 100% accuracy at the same time.

Abstract models based on the ROM technique allow rapid design space exploration with high fidelity. ROM also opens the opportunity to utilize the TLM benefits for real-time sensitive applications, where 100% timing accuracy is required. The relevant publications are [9], [7], [6].

IV. Abstract Processor Modeling

With the growing system complexity and the ever increasing software content, the abstraction of software execution gains importance. Traditional ISS-based validation becomes infeasible, due to the complexity. Faster, abstract models are necessary for performance evaluation and functional validation.

To address this need, we introduce our approach of abstract processor modeling in the context of multi-processor architectures. We use a layer based approach to construct versatile, multi-faceted processor models. Our models provide target timed execution of software, abstract RTOS modeling, processor external communication, and synchronization with interrupts (which are correctly scheduled to preempt task execution).

We validate our approach using an industrial strength cellphone example. It executes multiple subsystems (GSM 06.60 speech transcoder, MP3 decoder, and JPEG encoder) on a heterogeneous multi-processor SoC supported by custom hardware accelerators. Our experimental results show the feasibility and benefits of abstract processor modeling. Our TLM based system model reaches a speedup of 549 times over a dual-ISS based simulation, with an accuracy of 2%. Please see [10], [11] for more details.

V. Conclusion

Transaction Level Modeling has become the main approach to tackle the complexity challenges in todays system level design. This thesis provides guidelines for the model designer when developing abstract communication models. It guides in categorization and systematically analyzing communication models. It introduces ROM, a novel technique for communication modeling, which escapes the traditional TLM trade-off. It furthermore, presents an approach for abstracting processors.

The presented analysis of both the communication and the computation models aid the system designer to navigate the TLM trade-off effectively and choose the most suitable model for a given application.

References

[1] S. Abdi, J. Peng, H. Yu, D. Shin, A. Gerstlauer, R. Dömer, and D. Gajski. System-on-chip environment (SCE version 2.2.0 beta): Tutorial. Technical Report CECS-TR-03-41, Center for Embedded Computer Systems, University of California, Irvine, July 2003.

[2] ISO. *Reference Model of Open System Interconnection (OSI)*. Internation Organization for Standardization (ISO), second edition, 1994. ISO/IEC 7498 Standard.

Relevant Publications

[3] G. Schirner. System Level Modeling of an AMBA Bus. Master's thesis, Electrical Engineering and Compuer Science, University of California, Irvine, Mar. 2005.

[4] G. Schirner and R. Dömer. Abstract Communication Modeling: A Case Study Using the CAN Automotive Bus. In A. Rettberg, M. Zanella, and F. Rammig, editors, *From Specification to Embedded Systems Application*, Manaus, Brazil, August 2005. Springer.

[5] G. Schirner and R. Dömer. Abstract Communication Modeling and Analysis. *ACM Transactions on Embedded Computer Systems (under review)*, 2006.

[6] G. Schirner and R. Dömer. Accurate yet Fast Modeling of Real-Time Communication. In *CODES+ISSS*, Seoul, Korea, Oct. 2006.

[7] G. Schirner and R. Dömer. Fast and Accurate Transaction Level Models using Result Oriented Modeling. In *ICCAD*, San Jose, CA, Nov. 2006.

[8] G. Schirner and R. Dömer. Quantitative Analysis of Transaction Level Models for the AMBA Bus. In *DATE*, Munich, Germany, March 2006.

[9] G. Schirner and R. Dömer. Result Oriented Modeling a Novel Technique for Fast and Accurate TLM. *IEEE Transactions on CAD of Integrated Circuits and Systems (accepted for publication)*, 2007. [Online:] www.cecs.uci.edu/~hschirne/cv/ROM_TCAD_DACFORUM.pdf.

[10] G. Schirner, A. Gerstlauer, and R. Doemer. Abstract, Multifaceted Modeling of Embedded Processors for System Level Design. In *ASPDAC*, Yokohama, Japan, Jan. 2007.

[11] G. Schirner, G. Sachdeva, A. Gerstlauer, and R. Dömer. Embedded Software Development in an System-Level Design Flow: Case study for an ARM Processor. In *Proceedings of the International Embedded Systems Symposium*, Irvine, CA, June 2007.