Communication Protocol Analysis of Transaction-Level Models using Satisfiability Modulo Theories

Che-Wei Chang
Center for Embedded Computer Systems
University of California, Irvine
Irvine, CA 92697-2625, USA
e-mail: cheweic@uci.edu

Rainer Dömer
Center for Embedded Computer Systems
University of California, Irvine
Irvine, CA 92697-2625, USA
e-mail: doemer@uci.edu

Abstract—A critical aspect in SoC design is the correctness of communication between system blocks. In this work, we present a novel approach to formally verify various aspects of communication models, including timing constraints and liveness. Our approach automatically extracts timing relations and constraints from the design and builds a Satisfiability Modulo Theories (SMT) model whose assertions are then formally verified along with properties of interest input by the designer. Our method also addresses the complexity growth with a hierarchical approach. We demonstrate our approach on models communicating over industry standard bus protocol AMBA AHB and CAN bus. Our results show that the generated assertions can be solved within reasonable time.

I. INTRODUCTION

In system-level design, a transaction level model (TLM) describes the system components, their abstract computation behavior, and in particular the system communication over busses at an abstract functional level. Typically, the functionality and timing of a TLM is validated through simulation. In this paper, we formally verify the model and propose a method using Satisfiability Modulo Theories (SMT) [4] to statically analyze the TLM and verify features of interest. In particular, our main focus is here on the timing constraints in the communication protocols. As illustrated in Fig 1, we perform multiple rounds of verification using SMT, following a designer-in-the-loop methodology.

A. Designer Augmented Assertions

Our proposed methodology in this paper is to automatically extract timing relations and constraints from a given design and build a corresponding SMT model as verification framework. Then the designer can verify the properties of interest on the framework by augmenting the SMT model with assertions reflecting his points of interest. For example, to verify that the execution time of the application is always less than 100 time units, the designer can augment the SMT model with an assertion asking "Can the execution time be more than 100 time units?". If this is found unsatisfiable, the application will execute in 100 time units or less, taking all conditions into account. In other words, the execution time is proven to meet the timing constraint. On the other hand, if it is found satisfiable, the tool will also lists the conditions satisfying the assertions so that the designer can examine the situation.

B. TLM with Communication Timing

In a top-down system design flow, the system architecture model is further refined into a TLM. The main objective of TLM refinement is to choose and parameterize a bus protocol to implement the communication between the processing elements in the system. The communication protocol is specified by the inserted transaction level bus model which specifies the detailed communication timing, including synchronization and delays compliant to the chosen protocol. Compared with the previous model, TLM with communication timing better represents the real-world design in which communication does take time. The communication timing also has great influence on the execution time and even liveness of the design. Fig 2 shows two TLM examples and their corresponding timing diagrams. In both designs, the corresponding timing information for AMBA Advanced High-performance Bus (AHB) and Controller Area Network (CAN) bus protocol is specified in the TLM channel[10]. By formally extracting the timing rela-
C. Related Work

There is significant work in the realm of formal verification of system-level design, and one research method in this area is to convert the semantics of a behavioral model into another well-defined representation and make use of existing tools to validate the properties of interest. In [8] designs in SystemC are transformed into UPPAAL timed automata and verified by UPPAAL model checker; in [7] and [6] a method to convert SystemC into state machines for verification is proposed; [9] proposed to translate SystemC models into a Petri-net based representation for embedded systems (PRES+) for model checking; [2] proposes a multi-layer modeling to represent SystemC design in a predictive synchronization dependency graph (PSDG) and extended Petri net is proposed for formal deadlock checking. [1] translates SystemC to Kripke structure and applies symbolic model checking for verification. In contrast, Our approach acts as an interactive property checking tool which brings uncertainties about critical properties and corner cases to the attention of the designer. In our method, the designer verifies points of interest by adding corresponding assertions to the extracted SMT model. The satisfiability report obtained through our method highlights often special situations, such as missed acknowledge signals or unsatisfied condition, and assists the designer in verification of the model for all cases. In this paper, we use SpecC language [11] to create the system level model.

II. TIMING RELATIONS IN MODELS

A. Time Interval Model

In system design, functionality is not the only concern. Timing constraints are critical as well, especially for real-time systems and communication protocols. Therefore, the notion of time is an important aspect of the model. In this paper, we use a time interval \( (T_{\text{start}}, T_{\text{end}}) \) \[5\] to represent the start and end time of the execution of a statement \( s \) in the model. To properly reflect the discrete event semantics with delta cycles, we make every time stamp a 3-tuple \( (\text{Time}(t), \text{Delta}(d), \text{Order}(o)) \). Note that we use the third member, called \( \text{order} \), to distinguish statements that otherwise happen at the same time and delta cycle. The ordering is determined based on the timing relation between statements and assigned automatically by the solver. For such time stamps, we define a set of operations as listed in Table 1, describing the relations equality and greater-than, as well as time advance by wait-for-time.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>OPERATIONS OF TIME STAMP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operation</td>
<td>Definition</td>
</tr>
<tr>
<td>( T_A = T_B )</td>
<td>( T_A.t = T_B.t, T_A.d = T_B.d, T_A.o = T_B.o )</td>
</tr>
<tr>
<td>( T_A &gt; T_B )</td>
<td>( T_A.t &gt; T_B.t ) or ( T_A.d &gt; T_B.d ) or ( T_A.o &gt; T_B.o )</td>
</tr>
<tr>
<td>( T_A ) wait for ( N )</td>
<td>( T_A.t = T_A.t + N, T_A.d = 0 )</td>
</tr>
</tbody>
</table>

Exact timing, such as delay or execution time of computation and communication, can be specified by using wait-for-time statements that carry a time argument of integral constant type. When a wait-for-time statement is executed, the current behavior is suspended from execution for the specified time. Fig 3(A) shows an example with \( \text{waitfor}, \text{wait}, \) and \( \text{notify} \) statements. Here, statement A is executed \( \text{val} \) time units before statement B. Formally \( T_{\text{start}}(\text{stmt}_B) = T_{\text{end}}(\text{stmt}_A) + \text{val}, T_{\text{start}}(\text{stmt}_B) = 0 \) will be generated in the SMT model.

B. Timing Constraints

Minimum or maximum bounds on the time between two statements in the model are called timing constraints. To meet real-time constraints imposed on the application by the environment, e.g. for communication, such constraints need to be specified with the design model so that it can be implemented accordingly.

In the SpecC language [11], timing constraints can be specified in the model with a special \textbf{do-timing} construct, with which the timing constraints can be checked during simulation or, in our case, be extracted to assertions for formal verification. The syntax of timing constraints contains two parts: the do block specifies a set of labeled statements, whereas the timing block contains the actual constraints. In the do block, the statements whose timing the designer wants to check are given a unique label and in the following timing block the labels are used to set the constraints. Constraints are specified with the \textbf{range} construct, which takes four arguments. The first two arguments specify the labels and the last two the lower and upper bounds of the timing constraint, respectively. A \textbf{do-timing} example is shown in Fig 3(B). There are three labels in the do block, and two constraints are specified with range constructs in the timing block. Note that label \L2 is attached to a compound statement which contains two child behavior calls. The following condition must hold for the constraints in this example:

\[
0 \leq T_{\text{start}}.(L2) - T_{\text{start}}.(L1) \leq 100 \\
0 \leq T_{\text{start}}.(L3) - T_{\text{start}}.(L2) \leq 300
\]
Since $T_{start}(L1) = T_{start}(i_A)$ and $T_{start}(L2) = T_{end}(i_A)$, the first condition limits the execution time of $i_A$, and the second sets the constraint for the total execution time of $i_B$ and $i_C$.

III. TIMING RELATION EXTRACTION

A system model is composed of multiple computation blocks (modules, behaviors) with communication (channel) between them. We distinguish two types of behaviors: 1) Leaf and 2) Hierarchical behavior, which implements the computation and specifies the composition of instances respectively.

![Behavior Diagram](image)

Fig. 3. Two types of timing specification in SpecC language[11]

In our proposed method, we utilize the logic of uninterpreted functions with linear arithmetic (QF_LIA) which incorporates the Core and Ints theories to generate the assertions. The Core theory contains the basic types and operations for Boolean logic, and the Ints theory defines the integer type and basic functions for integer arithmetic and comparison. We use a function symbol (in SMT-LIB2 language) to represent each time stamp in the model and convert the timing relations between those stamps into assertions. For a newly introduced function symbol, the user can define the number of arguments, and the data type of the argument and the return value. In our method, the return value of an uninterpreted function is seen as the value of a time stamp, and the arguments of the function are used to specify the number of iterations a block is executed (if in a pipelined or loop structure). Take the waitfor statement in Fig 3(A) as an example. In the example, stmtn_A and stmtn_B will be executed once only and no argument is needed in the function symbol declaration for these two statements, and there is a delay of val time units between the execution of stmtn_A and stmtn_B inserted by the waitfor val statement. Thus, the assertions below are generated for the timing relation above. Our tool will name the symbol with the full hierarchy path to ensure the uniqueness of each function symbol.

\[(\text{declare - fun Tend_tStmtnA } () \text{ Int})\]
\[(\text{declare - fun Tstart_tStmtnB } () \text{ Int})\]
\[(\text{assert } (= \text{Tstart_tStmtnB (+ Tend_tStmtnA val)}))\]

A. Timing Relation for Hierarchical Behaviors

In SpecC, the child behavior instantiation implies a function call to the child behavior. For a behavior $S$ consisting of a set of child behavior instances $\langle s_1, s_2, s_3, ... s_n \rangle$, the following condition holds:

$$\forall i \in \{1, 2, 3, ... n\}, T_{start}(S) \leq T_{start}(s_i), T_{end}(S) \geq T_{end}(s_i)$$

The timing relation between the child behaviors is dependent on the execution type specified in the parent behavior. In this paper, we support sequential, parallel, pipelined, and loop behaviors.

1) Sequential Execution of statements is defined by ordered time intervals that do not overlap. Formally, for a statement $S$ consisting of a sequence of sub-statements $\langle s_1, s_2, ... s_n \rangle$, the following conditions hold:

$$\forall i \in \{1, 2, ..., n\}, T_{start}(s_i), T_{end}(s_i), T_{start}(S), T_{end}(S)$$

$$T_{start}(S) < T_{start}(s_i) \leq T_{end}(s_i) \leq T_{end}(S)$$

2) Parallel Execution can be specified by par or pipe statements. Formally, for a par statement $S$ consisting of concurrent child statements $\langle s_1, s_2, ... s_n \rangle$, the following conditions hold:

$$\forall i \in \{1, 2, ..., n\}, T_{start}(S) \leq T_{start}(s_i), T_{end}(S) \geq T_{end}(s_i)$$

$$T_{start}(S) < T_{end}(s_i)$$

3) Pipelined Execution of statements is a special form of concurrent execution. Formally, for a pipe statement $S$ executed for $m$ iterations, let $s_{i,j}$ represents the $j$-th iteration of the execution of statement $s_i$. Then, the following conditions hold:

$$\forall i, x \in \{1, 2, ..., n\}, j, y \in \{1, 2, ..., m\} :$$

$$T_{start}(s_{i,j}) < T_{end}(s_{i,j}),$$

$$T_{start}(s_{i,j}) = T_{start}(s_{x,y}), \quad \text{if } i + j = x + y$$

$$T_{end}(s_{i,j}) = T_{end}(s_{x,y}), \quad \text{if } i + j = x + y$$

$$T_{end}(s_{i,j}) \leq T_{start}(s_{x,y}), \quad \text{if } i + j < x + y$$

A limitation of our approach is that the number of iterations $m$ has to be a known integer. If it is statically unknown (i.e. a variable), our tool will prompt the designer to input an upper bound for the loop.

4) Loop Execution can be regarded as a special case of pipelined execution with only one stage. As above, we assume that the number of iterations is a finite constant.

B. Timing Relation Extraction for Leaf Behaviors

We pay significant attention in this paper to analyze the timing information specified in leaf behaviors and channels, which is critical in order to capture communication timing in TLMs. Fig. 4(A) highlights the statements which are analyzed in the source code as well as the rules to extract the corresponding timing relations for the static analysis. The rule for the waitfor statement has been introduced already. We now describe the others.

1) Conditional Execution: When conditional execution, such as a if statement or if-else statement, is used in the model, we create a time interval $\langle T_{if_start}, T_{if_end} \rangle$ and a logic stamp $C_{if}$ which represents the logic condition (for if-else, we also create a tuple $\langle T_{else_start}, T_{else_end} \rangle$). Fig. 5 illustrates the timing relations for the conditional execution. Here, $T_{prev}$ and
$T_{next}$ represent the time stamps before and after the conditional execution. As shown with the selection structure in Fig. 5, the value of $T_{next}$ is dependent on the binary value of $C_{if}$. Note that $T_{never}$ is a time stamp with a very large value representing infinity. We represent the situation that a statement will never be executed by giving the corresponding time stamp this large value (there is no way to represent infinity in the SMT-LIB language). Any time stamp greater or equal to $T_{never}$ means that the corresponding situation will never happen. Note that our tool will not analyze the specified condition in an if-statement, but only create the conditional assertions as listed in the illustration. It is the SMT solver’s job to find an assignment for the condition and time stamps that satisfy the assertions.

2) Loop Unrolling: To limit the verification space and the execution time of the solver, for each loop with undefined iteration count (i.e., the condition is variable), our tool will prompt the designer to provide an upper bound for the loop, and then unroll the loop to multiple if statements. Fig. 4(B) illustrates the loop unrolling performed by our tool. It also shows that the tool creates implication assertions for the conditions generated by loop unrolling.

3) wait-notify synchronization: In order to analyze a TLM with synchronization among multiple concurrent behaviors, we support events and the corresponding wait-notify synchronization. When a wait statement is executed, it suspends the current thread from execution until the event is triggered by a notify. A time interval $(T_{start}, T_{end})$ is generated as for other statements. For a wait statement $W$ triggered by a notify statement $N$, the following conditions hold:

\[
T_{start}(W) \leq T_{start}(N),
\]
\[
T_{end,d}(W) = T_{end,d}(N) + 1
\]

Note that $T_{start}$ equals $T_{end}$ for a notify statement. Also, to analyze the satisfiability of the specified timing constraints, we have to determine the mapping between wait and notify statements, i.e., which notify wakes up which wait. Our proposed method to generate the assertions for the wait-notify mapping is illustrated in Fig. 6.

![Timing relation extraction for the wait-notify statement](image)

In this example, all behaviors are executed in parallel except for the two behavior pairs (B1, B2) and (B6, B7) executed sequentially. Our method consists of two steps:

1) for every event in the model, our approach generates the assertions to sort the time stamps of the notify statements which trigger the event. This step is illustrated in the upper part of Fig. 6. Note that if the notify statement is inside a conditional statement, the value of its time stamp is dependent on the condition. For example, $T_{start,t}$ for the notify statement in behavior B4 in Fig. 6 will be greater than $T_{never}$ if the logic condition is false.

2) for every wait in the model, we generate the assertions to "search" the sorted time stamps of the notify statements and find one that is greater than and the closest to $T_{start}$ of the wait, and set the time and delta cycle of the wait using the condition we listed above. This step is illustrated in the lower part of Fig. 6.

4) Channel Interface Function Call: In a TLM, the timing information of the target bus protocol and the synchronization mechanism between communicating parties are specified in the interface methods defined in the channel. The communication between the behaviors takes place by calling those interface functions.

To generate assertions for the SMT solver, our approach traverses down to the interface method in the channel when it is called. Consequently, the timing information specified in the channel model is taken into consideration during the timing analysis of the behavior.

C. Liveness and Deadlock

For a multi-PEs system model, improper execution order or communication may lead to problems, including deadlock. In our method, a deadlock caused by circular waiting in the model...
will be reported to the designer in the form of unsatisfiable assertions since there are conflicts in the timing relations. Another potential deadlock would be a wait statement missing the wake-up signal. Fig. 6 also shows examples for two cases. Behavior B7 shows one case in which wait X misses all notification for X therefore it will never be waken up. Behavior B7 illustrates another case. wait Y can not wake up if the condition for notify Y in behavior B4 is not true. Both situations are covered by our tool and reported to the designer.

D. Hierarchical Timing Analysis

The number of assertions generated by our method increases with the complexity of the model. To keep the number of assertions manageable and limit the run time of the SMT solver, our method addresses the complexity growth by analyzing the timing constraints in a hierarchical manner. Timing constraints verified at a lower hierarchy level are regarded as the prerequisite conditions for the verification of the higher level. Verified timing constraints can be specified by use of the do-timing construct in the model. When our method finds a do-timing construct during the design traversal, it will take the constraints as they listed and not traverse further down the hierarchy. Thus, the assertions needed for model verification at the higher hierarchical level are greatly reduced.

Fig. 7. Hierarchical timing analysis of CAN bus protocol

Take the CAN bus protocol as an example. The bit time generated by the bit time logic for each engine control unit (ECU) can vary due to different local operating frequencies. Thus, the time needed for transmission can differ from one frame to another. To verify the timing constraint of the frame transmission, we use the pre-verified lower and upper bound of the bit time as prerequisite conditions. Fig. 7 illustrates the hierarchical timing analysis of CAN protocol from the bit time via frame time up to the application.

IV. EXPERIMENTS

We use two standard bus protocols widely used in industry to demonstrate our approach. As shown in Table II, both models are of reasonable size with practical analysis times. The first example is a three-ECU communication over a CAN bus protocol[10]. In this automotive example, the RPMcompute ECU issues a request to an RPMsensor using a remote frame. Upon the reception of the request, the sensor initiates an operation to read revolutions per minute (RPM) from the engine and sends it back to RPMcompute using a data frame. After receiving the raw RPM from the sensor, the RPMcompute ECU calculates the average RPM and sends that to Dashboard ECU for displaying. The procedure is illustrated in Fig. 2(A), and the detailed bus TLM is shown in Fig 8. Note that the bit time units required for each communication step in CAN bus protocol are specified with do-timing construct as prerequisite. In this example, the timing is analyzed on assumptions that reading RPM value from engine takes 40 bit time units and computing average RPM takes 10 to 20 bit time units.

Our second example is a producer-consumer model communicating over an AMBA AHB protocol. Here, the producer and consumer call interface functions send and receive, respectively, to transfer data through an AMBA AHB channel specified at TLM abstraction [10]. Fig. 9 illustrates the TLM with the detailed bus model. Note that a parallel behavior PollFlag is created to respond for the slave (Consumer) to all polling requests from the master (Producer). The procedure contains three steps as the timing diagram shown in Fig. 2(B): master reads the flag in PollFlag, master resets the flag, and then sends the data to slave. In this model, the delays compliant to the AHB reference are specified by waitfor statements.

Fig. 8. TLM of automotive example using CAN bus

Fig. 9. TLM of Producer-Consumer example using AMBA AHB

The statistics of TLM timing analysis for both bus protocols
are listed in Table II. In the experiments, we verify the satisfiability of liveness and timing constraints. Exp.1 and Exp.7 check if there is any conflict caused by circular waiting in the model, and the others verify the liveness and timing constraints in various scenarios specified with user augmented assertions. Exp.6 in the table shows a scenario where we allow the model to utilize the bus up to 60% maximum, that is, on average over 5 slots only 3 may be used. The number of assertions, lines of code (LOC) for those assertions and the run time of the solver are also listed in the table. According to the measured time, the satisfiability searching for these two models and the constraints we added is reasonably fast, and as is often expected, unsatisfiable solutions are faster obtained than satisfiable ones. Note that for Exp.1 the solver gave no answer after two hours of calculation. To reduce the search time, we added an assumption that the entire transaction finishes in finite time (test case 2). All experiments have been performed on a host PC with a 4-core CPU (Intel(R) Core(TM)2 Quad) at 3.0 GHz with Microsoft Z3 solver (version 4.1). Both experiments contain initial designer augmented assertions (3 for the CAN TLM and 6 for the AHB TLM) which are inserted to specify the real use case. Taking the CAN bus as an example, the three user augmented assertions reflect the timing relations shown in Fig. 2(A): the end time of remote frame transmission in RPMcompute equals the end time of the remote frame reception in RPMsensor; the end of frame transmission in RPMsensor equals the end of the frame reception in RPMcompute; and the end of the frame transmission in RPMcompute equals the end of the frame reception in Dashboard. The initial 6 user assertions in the AHB example reflect a similar situation.

V. CONCLUSION

In this paper, we have proposed an approach to verify liveness and timing constraints by extracting timing relations from a TLM design model and using a SMT solver to verify the satisfiability of the corresponding assertions. We verify the timing information specified in computation as well as in communication. Also, we introduce a hierarchical method to cope with the complexity growth of the model. We demonstrated our approach with two standard bus protocols AMBA AHB and CAN bus. Our approach utilizes the designer’s augmented assertion reflecting the properties of interest. In future work, we plan to improve the interaction between the designer and the SMT assertion generator.

REFERENCES


<table>
<thead>
<tr>
<th>exp</th>
<th>Constraint</th>
<th>Condition</th>
<th>#of assertions</th>
<th>LOC</th>
<th>Time</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T_end(DashDisp)&lt; T_never</td>
<td>No Circular Waiting</td>
<td>382 (3 augmented)</td>
<td>24963</td>
<td>&gt; 2hr</td>
<td>UNKNOWN</td>
</tr>
<tr>
<td>2</td>
<td>T_end(RPMcmp)&lt; 200 units</td>
<td>Minimum execution time</td>
<td>383 (4 augmented)</td>
<td>24972</td>
<td>189s</td>
<td>SAT</td>
</tr>
<tr>
<td>3</td>
<td>T_end(RPMcmp)&lt; 500 units</td>
<td>Write always fails in the 1st attempt</td>
<td>384 (5 augmented)</td>
<td>24975</td>
<td>50s</td>
<td>UNSAT</td>
</tr>
<tr>
<td>4</td>
<td>T_end(RPMcmp)&lt; 300 units</td>
<td>Write succeeds in the 1st attempt</td>
<td>385 (8 augmented)</td>
<td>24984</td>
<td>284s</td>
<td>SAT</td>
</tr>
<tr>
<td>5</td>
<td>T_end(RPMcmp)&lt; 500 units</td>
<td>Bus utilization &lt;= 60%</td>
<td>386 (5 augmented)</td>
<td>25054</td>
<td>135s</td>
<td>SAT</td>
</tr>
<tr>
<td>6</td>
<td>None</td>
<td>No Circular Waiting</td>
<td>240 (6 augmented)</td>
<td>19389</td>
<td>32s</td>
<td>SAT</td>
</tr>
<tr>
<td>7</td>
<td>T_end(Prod)&lt; T_never</td>
<td>No Circular Waiting</td>
<td>241 (7 augmented)</td>
<td>19392</td>
<td>31s</td>
<td>SAT</td>
</tr>
<tr>
<td>8</td>
<td>T_end(Prod)&lt; 6 cycles</td>
<td>Minimum execution time</td>
<td>242 (8 augmented)</td>
<td>19395</td>
<td>4s</td>
<td>UNSAT</td>
</tr>
<tr>
<td>9</td>
<td>T_end(Prod)&lt; 10 cycles</td>
<td>Polling succeeds in the first 2 attempts</td>
<td>243 (9 augmented)</td>
<td>19401</td>
<td>33s</td>
<td>SAT</td>
</tr>
<tr>
<td>10</td>
<td>T_end(Prod)&lt; 10 cycles</td>
<td>Polling succeeds in the first 2 attempts</td>
<td>243 (9 augmented)</td>
<td>19401</td>
<td>114s</td>
<td>UNSAT</td>
</tr>
</tbody>
</table>