Submission of Final Version of Thesis after Ph.D. Viva-Voce Examination & Executing Thesis Non-Exclusive License

Name of the Research Scholar: Mohammed Abderehman Adem

Roll Number:	1	6	6	1	0	1	1	0	3

Program of Study: Ph.D.

Thesis Title: Reverse Engineering of High-Level Synthesis and its Applications

Name of Thesis Supervisor(s) : Dr.Chandan Karfa

Name of the Department/ Center to which thesis is submitted : CSE

Date of completion of PhD Viva-Voce Exam : 2nd November 2022

(i) I hereby certify that I am the sole author of the above mentioned thesis.

(ii) I hereby confirm that the thesis contains my original research work done under the supervision of my thesis supervisor(s) and it does not infringe any rights of others. I also confirm that if any third party owned material was included in my thesis which required a written permission from the copyright owners, I have obtained all such permissions from respective copyright owners.

(iii) I understand that I retain copyright ownership and moral rights in my thesis, and that I may deal with the copyright in my thesis consistent with these rights.

(iv) I hereby grant to Indian Institute of Technology Guwahati and Lakshminath Bezbaroa Central Library of IIT Guwahati a non-exclusive, worldwide, irrevocable, royalty free license, in respect of my thesis, to use, reproduce, translate without changing the content, publish, archive, preserve, communicate and distribute, in paper form, in microform, electronically by telecommunication or on the internet, and/or any other formats as may be adopted for such use from time to time. I also authorize IIT Guwahati and L. B. Central Library of IIT Guwahati to sub-license, sub-contract for any of the acts mentioned.

(v) I verified the FINAL VERSION of the Thesis for completeness and for incorporation of all suggestions of Viva-Voce Board. I hereby submit the FINAL VERSION of the printed copy of my thesis and the exact same content in electronic format as a Single PDF file with a copy of short abstract in a separate PDF file to the Academic Affairs Section.

(vi) I have not submitted a requisition to temporarily withhold publication of my thesis for public access (Strike it out, if you have submitted).

Date: **Signature of Research Scholar**

Short Abstract of the Thesis is checked for correctness. Recommended for submission.

High-level synthesis (HLS) is the process of translating an abstract behavioral specification (usually written in C, C++) into a register transfer level (RTL) that realizes the given behavior. The HLS is widely used in the semiconductor industries due to advantages like shorter design cycles, efficient design space exploration, and easy writing specifications at a higher abstraction level. In the context of the quick development of hardware accelerators, the use of HLS is also crucial. In this work, we explore if we can reverse engineer the HLS, i.e., extracting a C code from the HLS generated RTL. The answer is yes as identified in this thesis. Specifically, we take the advantage of the special structure of the HLS generated RTL which consists of the separate datapath and controller finite state machine and automatically generate a concise, cycle accurate, and debug friendly C code called RTL-C from the RTL. We then show several applications of the RTL-C in the context of verification and security of HLS.

At present, the RTL co-simulation is the primary platform used for HLS design verification. Although most of the state-of-art RTL simulators provide an abstracted user friendly platform for verification, they are undesirably slow and sometimes incomprehensible to non-FPGA experts to debug. In the first application, we show that the RTL-C can be used for faster simulation based verification of the HLS. In this thesis, we introduce an automatic cycle accurate simulation tool FastSim for the same. Our simulation tool ensures RTL correctness, provides cycle accuracy, accurate performance estimation and renders on an average around 300 times faster simulation compared to RTL simulators and comparable performance to

that of software C simulators. Experiments on various HLS benchmarks demonstrate the efficiency and scalability of our simulation tool.

The formal verification of the HLS is still an open problem and the HLS tools are not bug free. The primary challenge of the formal verification is the abstraction gap between the input C and its corresponding RTL. As a second application, we show that RTL-C is helpful in reducing this abstraction gap. Specifically, we develop a formal verification tool DEEQ for checking equivalence between the C code against the RTL-C. We have taken a data-driven approach to find the correspondence of traces between two behaviors. We also merge compatible traces within a behavior to reduce the verification complexity. Finally, the equivalence of traces is shown with help of an SMT solver. Experimental results show that our proposed method can prove the end-to-end equivalence for small to medium benchmark designs for a commercial HLS tool.

The variables of a high-level behavior are mapped to hardware registers during the register allocation (RA) step of HLS. Due to possible many-to-many relations between the variables in C and the registers in the RTL, it is not straightforward to identify this mapping automatically. As a third application, we have shown that the RTL-C can be utilized to identify this mapping automatically. Specifically, we have taken the input C/scheduled C code and RTL-C and we come up with two methods through which we can automatically extract this mapping information. In the first approach, the scheduled C code and the RTL-C are combined state-wise and an invariant generator tool Daikon is used to identify the mapping information. In the second approach, we formulate the mapping problem as a Satisfiability (SAT) problem and use Satisfiability Modulo Theory (SMT) solver to obtain the register to variable mapping information. The frameworks are implemented and tested on a commercial HLS tool for several benchmark designs.

A hardware Trojan (HT) is a malicious modification of the design done by a rogue employee or a malicious foundry to leak secret information, create a backdoor for attackers, alter functionality, degrade performance and even halt the system. Recently, a possibility of HTs - specifically, battery exhaustion attack, degradation attack, and downgrade attack insertion, are shown in a compromised HLS tool. As a fourth application, we utilize our RTL-C to detect HTs inserted by HLS tool. Specifically, we have identified a battery exhaustion attack during generation of RTL-C. The degradation attack and the downgrade attack are detected during the C to RTL-C equivalence checking. The experimental results confirm the detection of HTs of the black-hat HLS tool. Overall, this thesis proposes an automatic way to extract a C code from the RTL generated by the HLS tool and show various important applications of this reverse engineering process.

Signature of Thesis Supervisor(s) with Date

Forwarded

Signature of Chairperson, Doctoral Committee with Date

Signature of Chairperson DPPC/ CPPC with Date

- Collected One Printed Copies and Two Softcopies of the thesis.
- Checked the Printed Copy of the Thesis and the Short Abstract.
- Opened & Checked the PDF file of the Thesis and PDF file of the Abstract.
- *Requisition to temporarily withhold publication of thesis for public access is* ***submitted / not submitted.***

Sent the received Printed Copy and Softcopy of the thesis to Lakshminath Bezbaroa Central Library

Signature of Receiving Staff with Date

Signature of Deputy Registrar with Date