

Functional Instruction Set Simulator (ISS) of a Neural Network (NN) IP with Native BF-16 Generator

Authors : Debajyoti Mukherjee, Arathy B. S., Arpita Sahu, Saranga P. Pogula

Abstract : A Functional Model to mimic the functional correctness of a Neural Network Compute Accelerator IP is very crucial for design validation. Neural network workloads are based on a Brain Floating Point (BF-16) data type. The major challenge we were facing was the incompatibility of gcc compilers to BF-16 datatype, which we addressed with a native BF-16 generator integrated to our functional model. Moreover, working with big GEMM (General Matrix Multiplication) or SpMM (Sparse Matrix Multiplication) Work Loads (Dense or Sparse) and debugging the failures related to data integrity is highly painstaking. In this paper, we are addressing the quality challenge of such a complex Neural Network Accelerator design by proposing a Functional Model-based scoreboard or Software model using SystemC. The proposed Functional Model executes the assembly code based on the ISA of the processor IP, decodes all instructions, and executes as expected to be done by the DUT. The said model would give a lot of visibility and debug capability in the DUT bringing up micro-steps of execution.

Keywords : ISA (instruction set architecture), NN (neural network), TLM (transaction-level modeling), GEMM (general matrix multiplication)

Conference Title : ICVLSIDT 2023 : International Conference on VLSI Design and Technology

Conference Location : New York, United States

Conference Dates : December 11-12, 2023