STMicroelectronics

AI Accelerators Hands-on HW Design

Jul.2025

LAB 0

Systolic Array for Applying Matrix Multiplication

Instructor:

Ahmed Abdelsalam

Due to:

Thu 27 Jul.2025,11:59PM

Objective

The objective of this lab is to assess our skills in writing efficient RTL code and our knowledge about architecture using SV/Verilog/VHDL. Additionally, we will verify their functionality through a testbench.

Prerequisites

- Good understanding of digital logic design.
- Familiarity with VHDL or Verilog/SV HDL.
- Understanding of matrix multiplication and systolic array concepts.

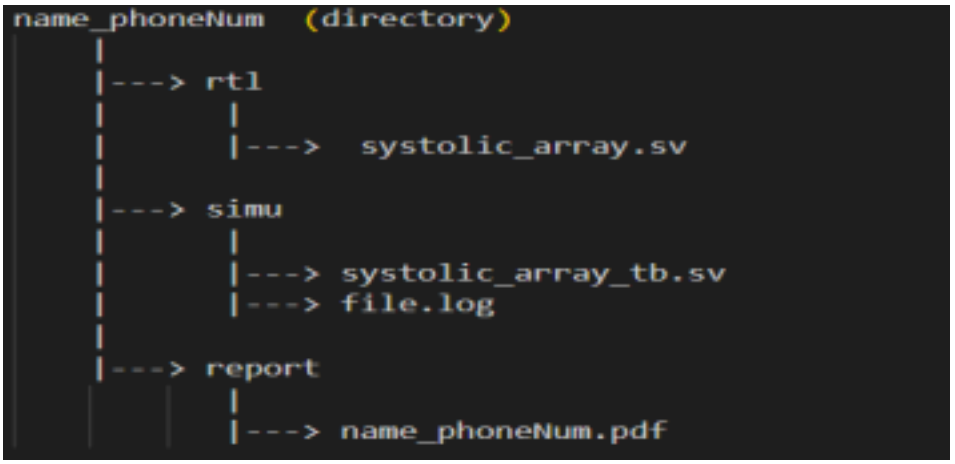
Notes for submission

1. Compress them as "***file.zip***"

2. Rename as “name_phoneNum.zip”
3. upload to your OneDrive / google drive
4. make it sharable (anyone with link can access the file without permission)
5. post the link to this form **FORM LINK**

Requirements

- **Code**
  - **RTL** the Verilog/VHDL code implements the assigned question and comment your code well all the design shall be in one file named “***systolic_array.sv***” you have to respect the naming and port list of the top module that we provide in the document.
  - **Simulation** you need to provide testbench for the module with at least one direct test case and contains display shows input and output matrices used and include the file.log output from the simulation in the zip file
- **Report**
  - **Architecture** detailed diagram shows the architecture of your implementation followed by an explanation for the operation of your design
  - Add snapshots from your code with abstract description
  - **Simulation Snapshots** from the simulation waveform and the log to show your design is working as expected with comments to illustrate the figures
  - If your assignment is not done you need to specify what is done exactly and what is the missing part and reporting your trials
- **Directory structure**



I. Introduction

A ××××××× systolic array consists of a grid of ××^{2} processing elements (PEs), arranged in ×× rows and ×× columns. Each PE performs multiply-accumulate operations for corresponding elements of the input matrices.

- The first matrix (A) elements are fed into the array row-wise from the left.
- The second matrix (B) elements are fed column-wise from the top.
- Each PE multiplies the incoming elements and adds the result to a running sum (partial product).
- Intermediate results (partial sums) flow depends on the Flow you chose for your inputs in our example each output will be kept in same location.
- After sufficient clock cycles, the output matrix elements appear at the output of each PE.

This pipeline enables continuous data flow and parallel computation, significantly speeding up the matrix multiplication process compared to sequential methods

II. Task Description

You are required to implement a (*N_SIZE x N_SIZE*) systolic array perform parallel multiply accumulate operations across processing elements to compute the product of two matrices and result *N_SIZE x N_SIZE* matrix. The implementation must include clearly defined input and output interfaces, synchronized by a clock and reset signals. A detailed report documenting the Task flow, challenges, and simulation results is also required.

1. Parameter List

Port Name	Type	Default Value	Description
DATAWIDTH	Integer	16	Datawidth of elements in in
N_SIZE	Integer	5	The size of the resulting matrix or the Number of PEs in each row or columns we assumed square matrix for simplicity

2. Port List

Port Name	Direction	Width	Description
clk	Input	1-bit	Positive edge clock signal
rst_n	Input	1-bit	Negative edge reset

valid_in	Input	1-bit	Valid signal set to 1 when a valid data are settled on ***'matric_a_in'*** and ***'matric_a_in'*** so the DUT is allowed to sample them
matrix_a_in	Input	N_SIZE* DATAWIDTH	Array of N inputs corresponding to one column of matrix A elements entering the systolic array rows.
matrix_b_in **valid_out**	Input Output	N_SIZE* DATAWID TH 1-bit	Array of N inputs corresponding to one row of matrix B elements entering the systolic array columns. Valid signal set to 1 when a valid row of the result matrix are settled on ***'matric_c_out'***
matrix_c_out	Output	N_SIZE*2* DATAWIDTH	Array of 5 outputs corresponding to one row of matrix C elements resulting from the array multiplication

Note

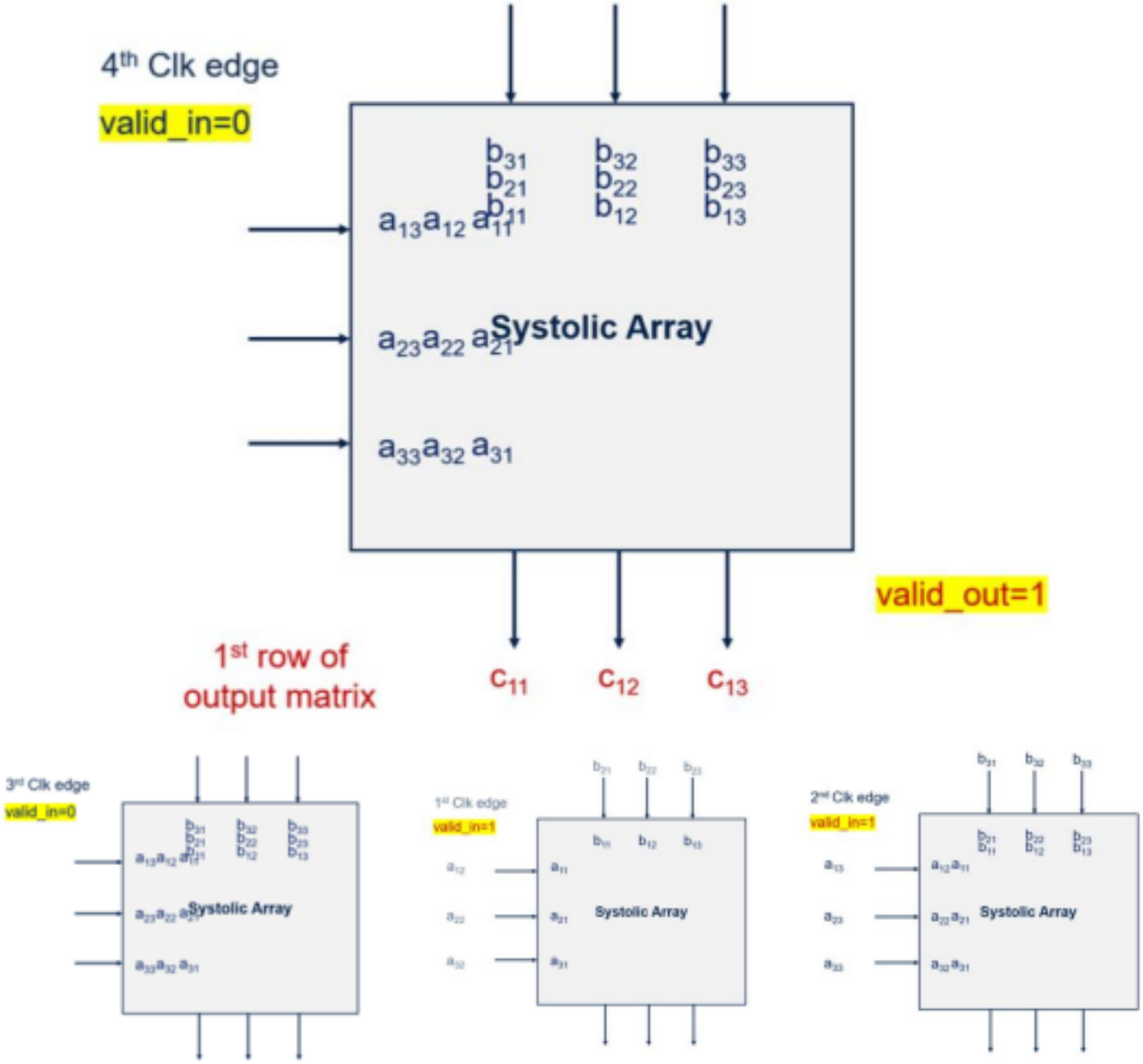
- Define these 2 parameters even if you implement a fixed non-parametrized systolic array and will not use these parameters in the design
- You must respect the ports, Parameters and module naming including upper/lowercase to avoid errors in the testing Phase
- Top Module name: **systolic_array**

Hint

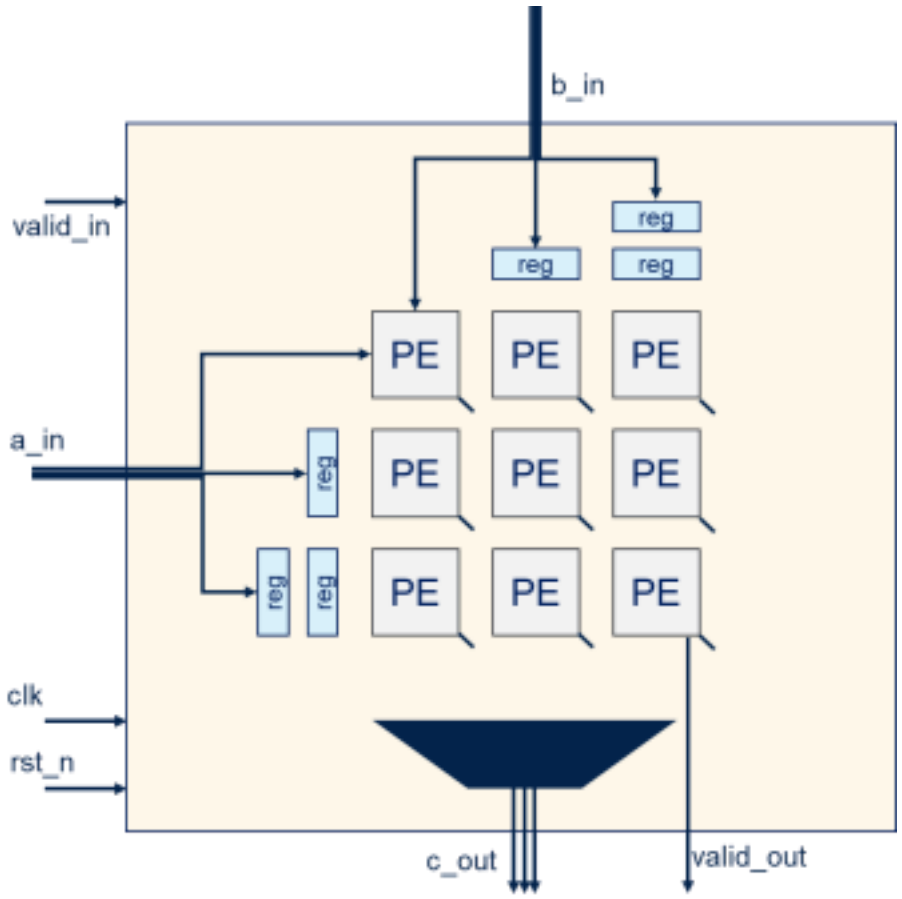
- starts with default values and solve issues till the code is functionally correct them try to parametrize your design.

3. Input/Output Dataflow

The figure below descripes in details how input are fed to the dut and format of the expected output showing the detailed status of the control signal



4. Architecture shows how we can implement this function with some hidden parts you need to fill the missing based on your understanding of the functionality



Rules

- This Task is individual the collaborative work is **not** allowed
- It's fine to use AI tools
- Sharing coding details between colleagues is prohibited
- Any assistance from professor/engineer/colleague is prohibited

Submit Task