

# **DIGITAL IC DESIGN**

## **Project Report**

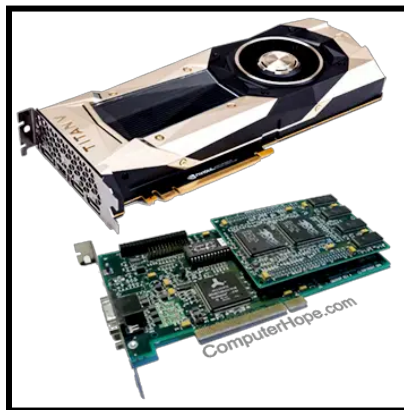
**Title:** Design a custom hardware accelerator for a machine learning algorithm, using SystemVerilog. (Fully Connected Neural Network)

### **What is Hardware Accelerator?**

A hardware accelerator is a specialized computing device that is designed to perform a specific function or set of functions with high efficiency and speed.

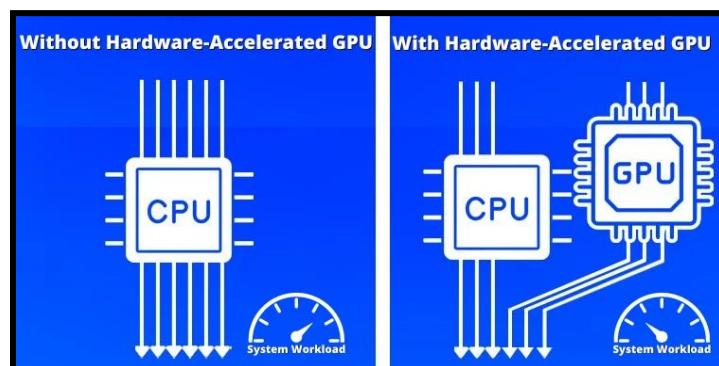
### **Applications of Hardware Accelerator:**

- AI & ML application
- Cryptography and Security
- Image and Video Processing
- Data Analytics



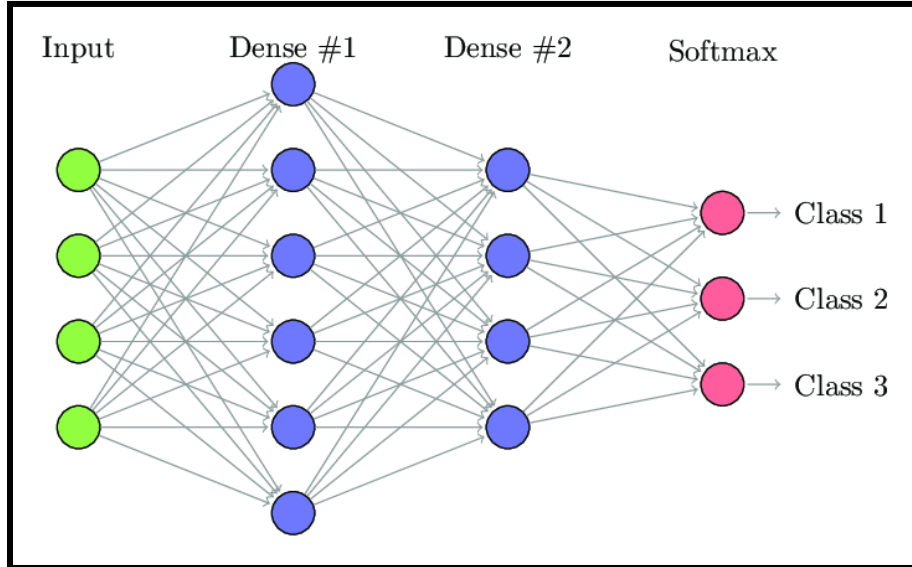
### **How does it work?**

It works in parallel to the CPU to reduce its workload.



## Architecture Followed:

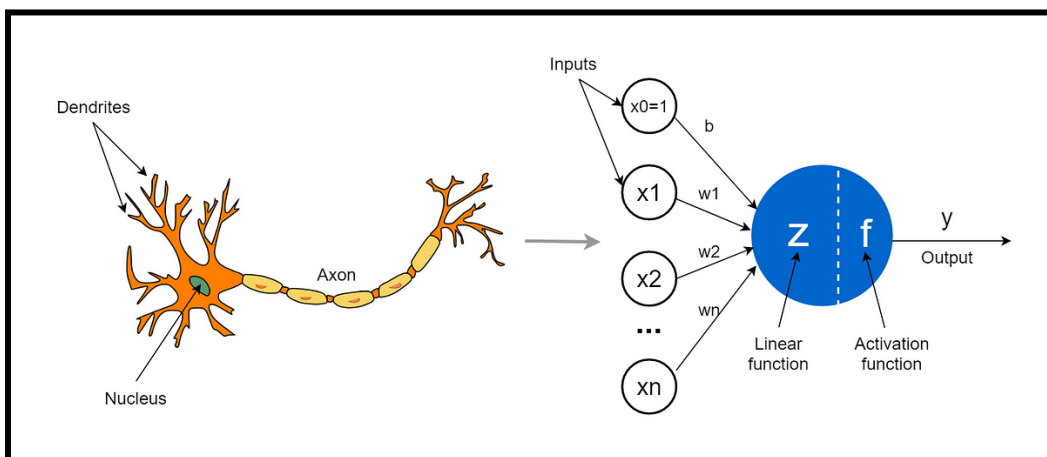
Fully Connected Neural Network (FNN):



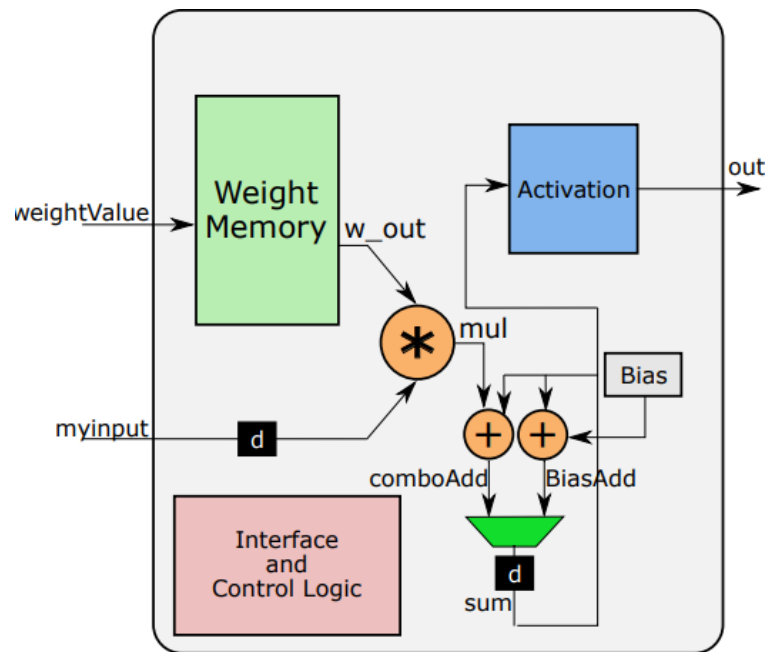
## Implementation of NN on hardware:

- In most cases, we use pre-trained neural networks:
  - Training requires a lot of hardware
  - Algorithms are not very hardware friendly
  - Most networks are trained only once → waste of resources used for training.
  - In most applications, training is not the time-critical path
- So generally we train in computers with all sophistication and find the weight and bias values and directly use them during hardware design

## Basic Building Block: Neuron:



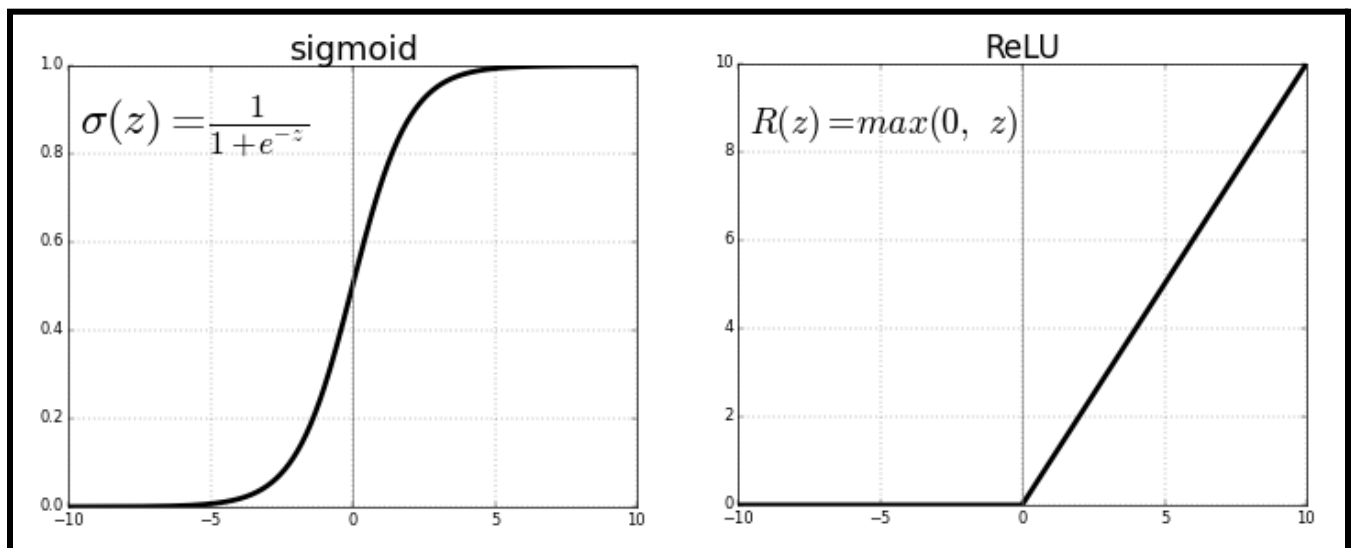
## Architecture of a Neuron:



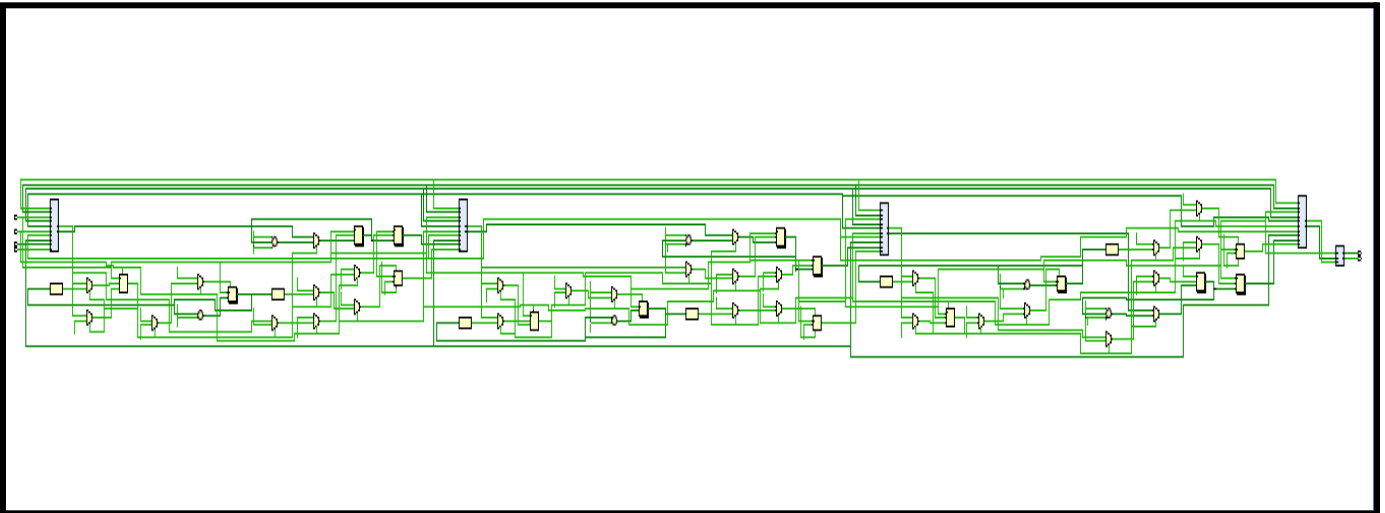
## Components of the FNN:

1. Neuron
2. Weight Memory
3. Bias Memory
4. Activation Unit
5. Hardmax (Maximum Output Finder)

## Activation Functions Used:



## Elaborated Design:



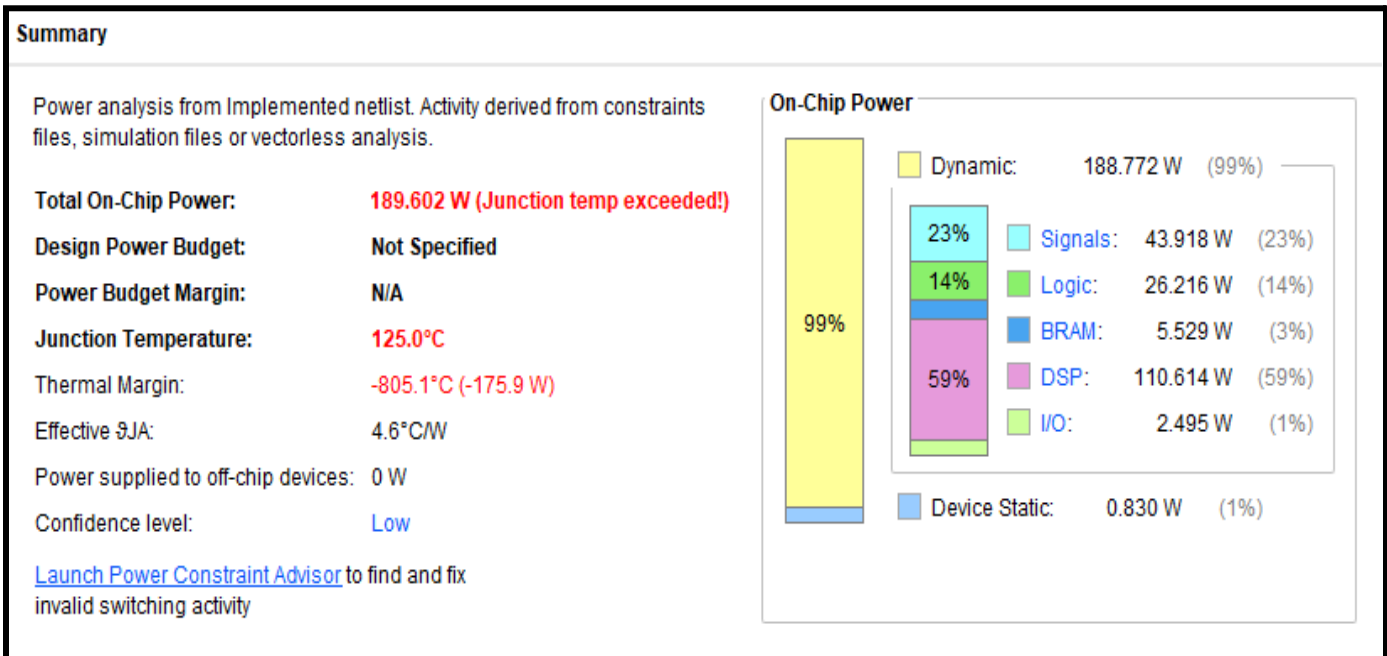
## Input To the Network:

- Presently inputs are given as text files through Testbench
- The weights and Bias of each Neuron are stored in the dedicated memory of each Neuron (Register file)

## Results Obtained:

### Vivado Simulation results:

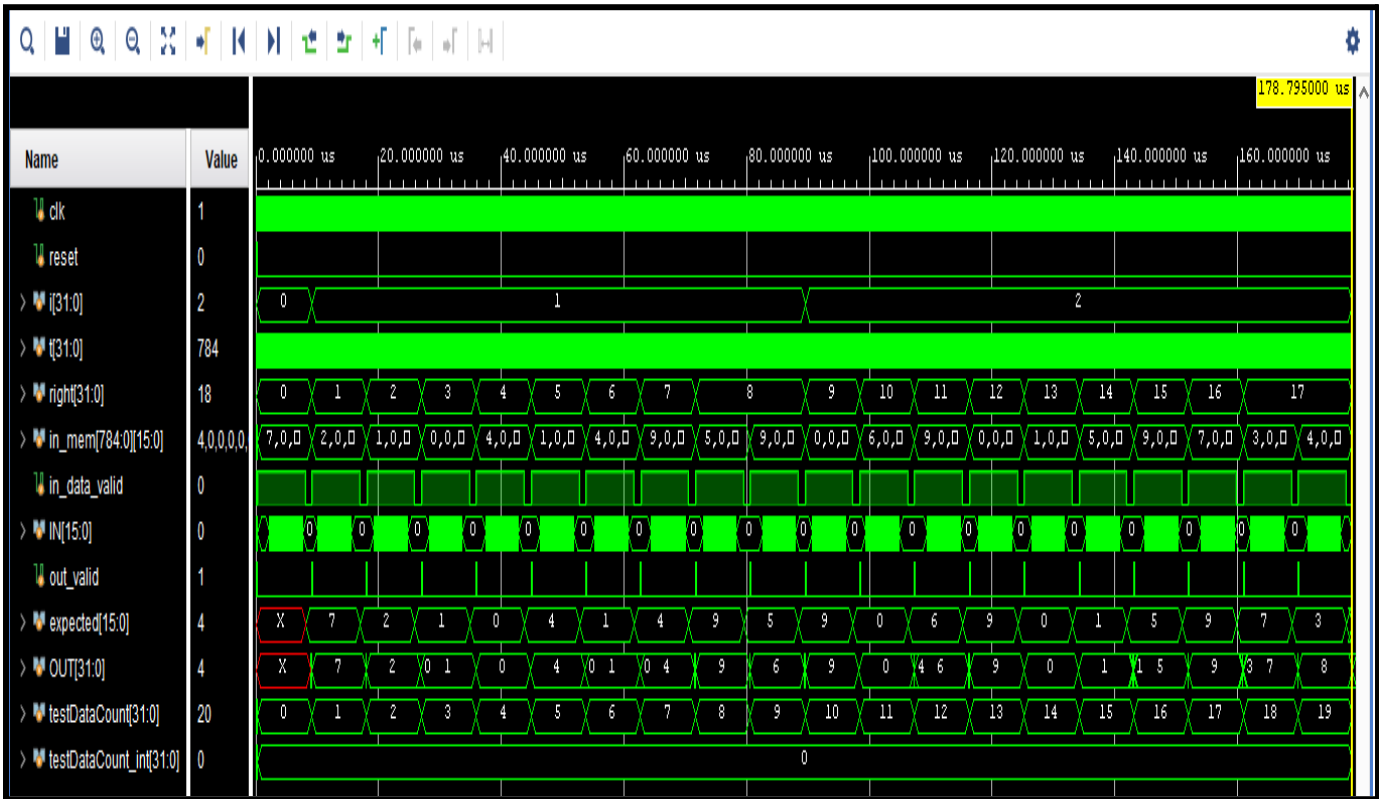
#### Power Analysis:



Utilization Report:

Name	Slice LUTs (63400)	Slice Registers (126800)	F7 Muxes (31700)	Slice (15850)	LUT as Logic (63400)	Block RAM Tile (135)	DSPs (240)	Bonded IOB (210)	BUFGCTRL (32)
top_FNN	5716	4655	15	2005	5716	43	160	52	1
I1 (Layer1)	1926	1382	0	689	1926	30	60	0	0
I2 (Layer2)	2227	1246	0	741	2227	8	60	0	0
I3 (Layer3)	719	412	0	239	719	2.5	20	0	0
I4 (Layer4)	729	394	0	261	729	2.5	20	0	0
mFind (maxFinder)	82	95	15	50	82	0	0	0	0

Simulation Results:



## Accuracy Obtained for MNIST Dataset (for 20 samples):

```
Tcl Console x Messages Log
Time resolution is 1 ps
relaunch_sim: Time (s): cpu = 00:00:14 ; elapsed = 00:02:19 . Memory (MB): peak = 2776.551 ; gain = 2.184
run all
1. Accuracy: 100.000000, Detected number: 7, Expected: 0007
2. Accuracy: 100.000000, Detected number: 2, Expected: 0002
3. Accuracy: 100.000000, Detected number: 1, Expected: 0001
4. Accuracy: 100.000000, Detected number: 0, Expected: 0000
5. Accuracy: 100.000000, Detected number: 4, Expected: 0004
6. Accuracy: 100.000000, Detected number: 1, Expected: 0001
7. Accuracy: 100.000000, Detected number: 4, Expected: 0004
8. Accuracy: 100.000000, Detected number: 9, Expected: 0009
9. Accuracy: 88.888889, Detected number: 6, Expected: 0005
10. Accuracy: 90.000000, Detected number: 9, Expected: 0009
11. Accuracy: 90.909091, Detected number: 0, Expected: 0000
12. Accuracy: 91.666667, Detected number: 6, Expected: 0006
13. Accuracy: 92.307692, Detected number: 9, Expected: 0009
14. Accuracy: 92.857143, Detected number: 0, Expected: 0000
15. Accuracy: 93.333333, Detected number: 1, Expected: 0001
16. Accuracy: 93.750000, Detected number: 5, Expected: 0005
17. Accuracy: 94.117647, Detected number: 9, Expected: 0009
18. Accuracy: 94.444444, Detected number: 7, Expected: 0007
19. Accuracy: 89.473684, Detected number: 8, Expected: 0003
20. Accuracy: 90.000000, Detected number: 4, Expected: 0004
Accuracy: 90.000000
$stop called at time : 178795 ns : File "C:/Users/RAJESH KUMAR/Desktop/Digital_IC_Assign/Final_Project_HW_Acc/Final_Project_HW_Acc.srcs/sim_1/new/
run: Time (s): cpu = 00:00:11 ; elapsed = 00:00:17 . Memory (MB): peak = 2776.551 ; gain = 0.000
```

## Genus Synthesis results:

### Genus Synthesis Area Report:

Generated by: Genus(TM) Synthesis Solution 21.10-p002_1						
Generated on: Apr 13 2023 02:06:45 am						
Module: top_FNN						
Technology library: uk65lscllmvbbr_120c25_tc						
Operating conditions: uk65lscllmvbbr_120c25_tc (balanced_tree)						
Wireload mode: top						
Area mode: timing library						
Instance	Module	Cell Count	Cell Area	Net Area	Total Area	Wireload
top_FNN		244066	648116.280	0.000	648116.280	wl0 (D)
l1	Layer1_NN30_numWeight784_dataWidth16_layerNum1_sig	118000	296593.560	0.000	296593.560	wl0 (T)
n_0_siginst.s1	Sig_ROM_inWidth10_dataWidth16	961	2143.080	0.000	2143.080	wl0 (T)
n_690_siginst.s1	Sig_ROM_inWidth10_dataWidth16_4	961	2143.080	0.000	2143.080	wl0 (T)
n_691_siginst.s1	Sig_ROM_inWidth10_dataWidth16_6	961	2143.080	0.000	2143.080	wl0 (T)
n_695_siginst.s1	Sig_ROM_inWidth10_dataWidth16_11	961	2143.080	0.000	2143.080	wl0 (T)
n_696_siginst.s1	Sig_ROM_inWidth10_dataWidth16_12	961	2143.080	0.000	2143.080	wl0 (T)
n_697_siginst.s1	Sig_ROM_inWidth10_dataWidth16_13	961	2143.080	0.000	2143.080	wl0 (T)
n_698_siginst.s1	Sig_ROM_inWidth10_dataWidth16_17	961	2143.080	0.000	2143.080	wl0 (T)
n_700_siginst.s1	Sig_ROM_inWidth10_dataWidth16_26	961	2143.080	0.000	2143.080	wl0 (T)
n_702_siginst.s1	Sig_ROM_inWidth10_dataWidth16_3	961	2143.080	0.000	2143.080	wl0 (T)
n_704_siginst.s1	Sig_ROM_inWidth10_dataWidth16_8	961	2143.080	0.000	2143.080	wl0 (T)
n_707_siginst.s1	Sig_ROM_inWidth10_dataWidth16_16	961	2143.080	0.000	2143.080	wl0 (T)
n_708_siginst.s1	Sig_ROM_inWidth10_dataWidth16_18	961	2143.080	0.000	2143.080	wl0 (T)
n_712_siginst.s1	Sig_ROM_inWidth10_dataWidth16_25	961	2143.080	0.000	2143.080	wl0 (T)
n_714_siginst.s1	Sig_ROM_inWidth10_dataWidth16_29	961	2142.720	0.000	2142.720	wl0 (T)
n_2_siginst.s1	Sig_ROM_inWidth10_dataWidth16_2	961	2143.080	0.000	2143.080	wl0 (T)
n_180_siginst.s1	Sig_ROM_inWidth10_dataWidth16_22	961	2143.080	0.000	2143.080	wl0 (T)
n_196_siginst.s1	Sig_ROM_inWidth10_dataWidth16_23	961	2143.080	0.000	2143.080	wl0 (T)
n_689_siginst.s1	Sig_ROM_inWidth10_dataWidth16_1	961	2143.080	0.000	2143.080	wl0 (T)
n_692_siginst.s1	Sig_ROM_inWidth10_dataWidth16_7	961	2143.080	0.000	2143.080	wl0 (T)
n_693_siginst.s1	Sig_ROM_inWidth10_dataWidth16_9	961	2143.080	0.000	2143.080	wl0 (T)
n_694_siginst.s1	Sig_ROM_inWidth10_dataWidth16_10	961	2143.080	0.000	2143.080	wl0 (T)
n_699_siginst.s1	Sig_ROM_inWidth10_dataWidth16_21	961	2143.080	0.000	2143.080	wl0 (T)
n_701_siginst.s1	Sig_ROM_inWidth10_dataWidth16_27	961	2143.080	0.000	2143.080	wl0 (T)

### Genus Synthesis Power Report:

Instance: /top\_FNN

Power Unit: W

PDB Frames: /stim#0/frame#0

Category	Leakage	Internal	Switching	Total	Row%
memory	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.00%
register	6.19448e-06	1.05997e-01	2.72155e-02	1.33219e-01	21.60%
latch	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.00%
logic	4.69343e-05	2.41888e-01	2.41716e-01	4.83651e-01	78.40%
bbox	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.00%
clock	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.00%
pad	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.00%
pm	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.00%
Subtotal	5.31288e-05	3.47885e-01	2.68931e-01	6.16870e-01	100.00%
Percentage	0.01%	56.40%	43.60%	100.00%	100.00%

### Genus Synthesis Timing Report:

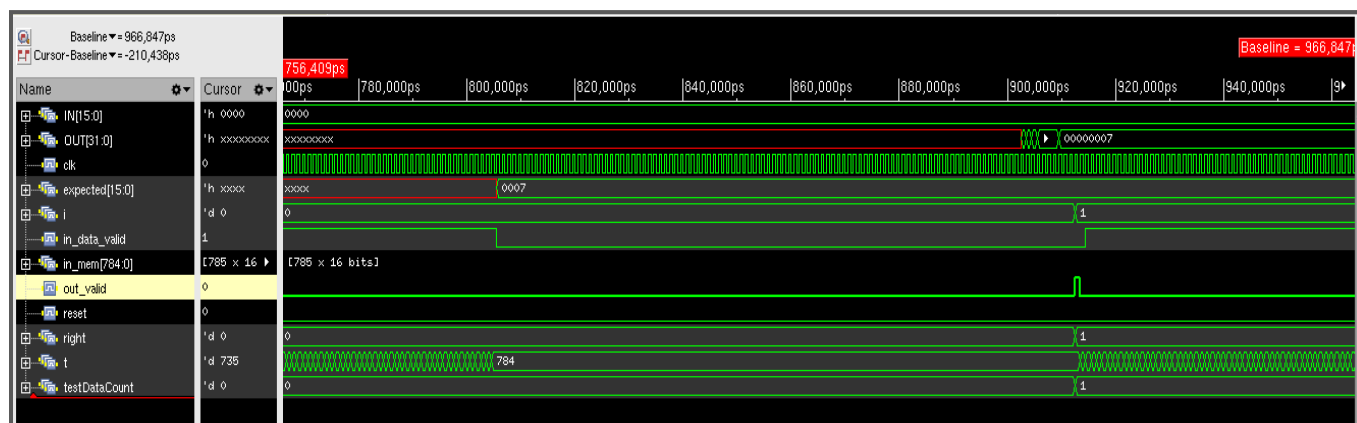
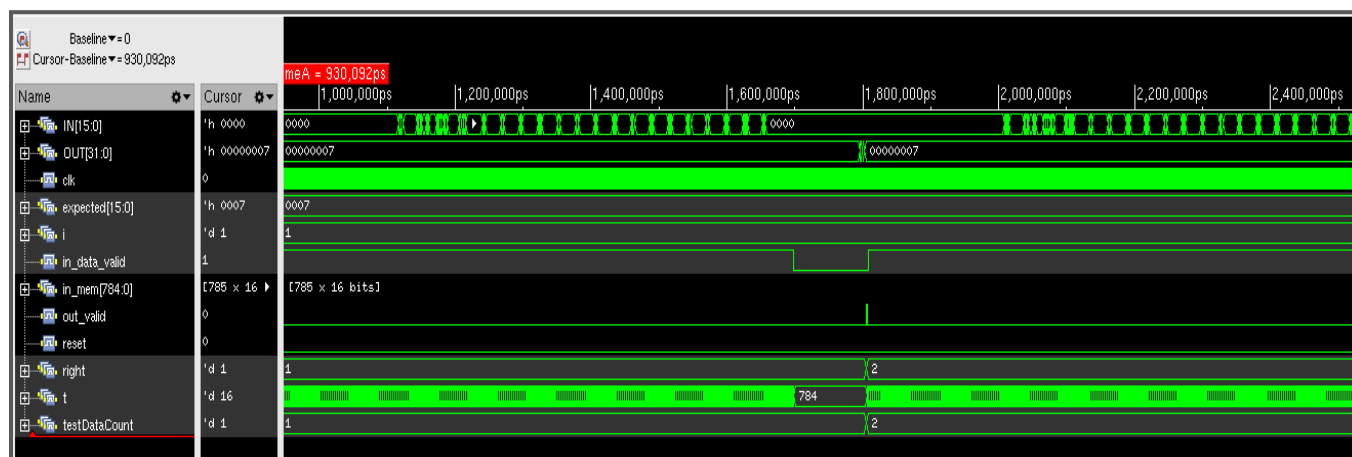
g66489/Z	CKND2M2R	2	3.1	35	+30	700	R
g64943/B					+0	700	
g64943/Z	OAI211M4R	20	27.4	71	+141	841	F
g64805/A					+0	841	
g64805/Z	INVM2R	13	11.3	90	+76	917	R
g62515/NA1					+0	917	
g62515/Z	OAI21B20M2R	1	1.4	32	+66	983	R
n_696_sum_reg[25]/D	<<< DFM2RA				+0	983	
n_696_sum_reg[25]/CK	setup			50	+7	990	R
(clock clk)	capture					1000	R
	uncertainty				-10	990	R
Cost Group : 'clk' (path_group 'clk')							
Timing slack : 0ps							
Start-point : l2/n_696_sum_reg[0]/CK							
End-point : l2/n_696_sum_reg[25]/D							

**Max Clock Frequency:** 1 GHz (Clock Period: 1ns)

## Genus Synthesis Gates Report:

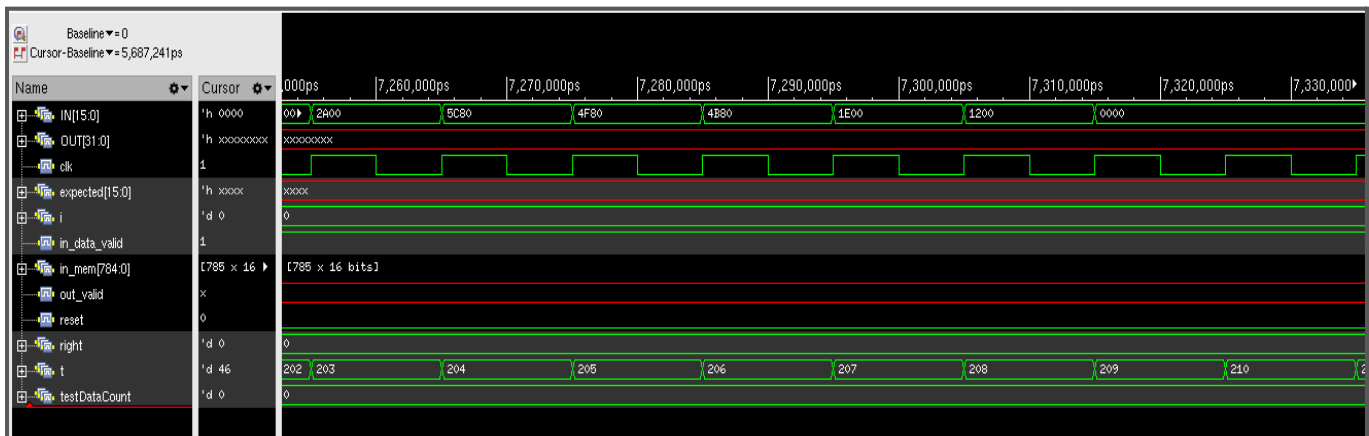
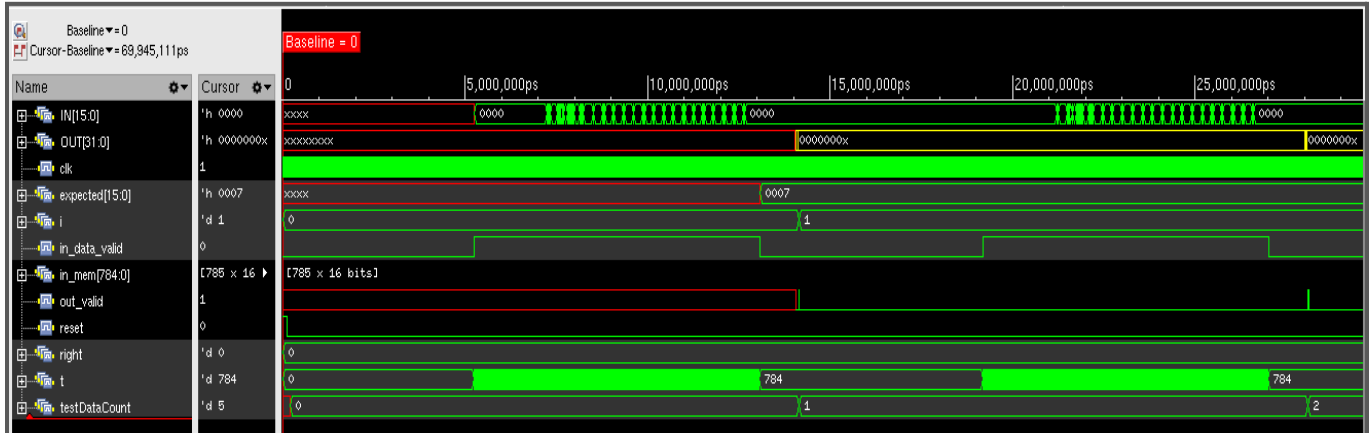
Type	Instances	Area	Area %
sequential	8577	64995.480	10.0
inverter	29879	42611.400	6.6
buffer	1251	3329.280	0.5
logic	204359	537180.120	82.9
physical_cells	0	0.000	0.0
total	244066	648116.280	100.0

## NCSim Simulation: Pre-Synthesis:





## NCSim Simulation: Post-Synthesis:



## Future Works:

→ CPU Interfacing:

- Using IP's present in Xilinx Vivado, the Accelerator can be interfaced with CPU.

→ CTS, Placement, Routing and Layout using tools like Innovus.

Additional Tasks to do:

- Automating the Number of Neurons and Layers using Scripting and making it suitable for various datasets (Presently designed for MNIST dataset).

## **GitHub Repository Link:**

[https://github.com/SharmaPrateek18/Fully\\_Connected\\_Neural\\_Network](https://github.com/SharmaPrateek18/Fully_Connected_Neural_Network)

## **References:**

- <https://www.computerhope.com>
- <https://www.youtube.com>
- <https://www.Stackoverflow.com>
- <http://neuralnetworksanddeeplearning.com>