ECE 520 ASIC Design Project Report

StudentID:200207571 Name : Amaresh Subburaj		
Delay (ns to run provided provided example). Clock period: 7ns # cycles": 1171	Logic Area: 25180.3581 (um^2)	1/(delay.area) ($ns^{-1}.um^{-2}$) 4.844e-9
Delay (TA provided example. TA to complete)		1/(delay.area) (TA)

Abstract

This is an implementation of simplified version of a Convolutional Neural Network, a common algorithm used in machine learning. The Convolutional layers apply a convolution operation to the input, passing the result to the next layer. The input is a 12x12 matrix of 16 bit twos complement number from SRAM convoluted with b_0, b_1, b_2, b_3 which are first stage convolutional vectors. The resultant vector matrix is convoluted with second stage convolutional vectors. These convolution operations involve a series of dot products.

Design and Implementation of simple Convoluted Neural Network using Verilog(RTL).

Amaresh Subburaj

Abstract

This is an implementation of simplified version of a Convolutional Neural Network, a common algorithm used in machine learning. The Convolutional layers apply a convolution operation to the input, passing the result to the next layer. The input is a 12x12 matrix of 16 bit twos complement number from SRAM convoluted with b_0, b_1, b_2, b_3 which are first stage convolutional vectors. The resultant vector matrix is convoluted with eight second stage convolutional vectors. These convolution operations involve a series of dot products.

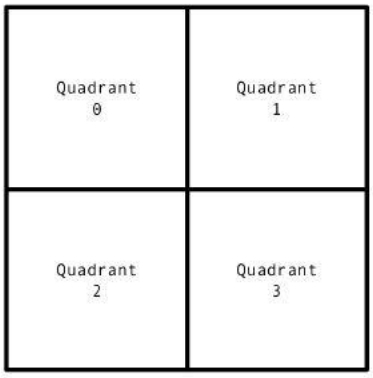
1. Introduction

A two-layered convolutional Neural network algorithm is being implemented. The sample input and first and second convolution layer data are read from dim and bvm SRAM respectively and the output will be written back to dom SRAM.A balanced approach having moderate area while achieving decent clock period and decent throughput is used. and In Micro-Architecture section, high level architecture drawing with interfaces of the design and data path is presented. In Interface specification section, detailed description of top level of the design interface is presented. In Technical Implementation section detailed structure of the design and controller functionality will be explained. Thereafter verification approach, results achieved are discussed.

2. Micro-Architecture

Step one:

12x12 input matrix is divided into four 6x6 matrix as shown in the below diagram,



Each quadrant of 6x6 is subdivided into four 3x3 matrices each of arranged as 1x9 vector. Effectively the entire sample input is subsampled into 16 ,1x9 vectors.

Each of these 16 vectors are convoluted with given convolution filter vectors b_0, b_1, b_2, b_3 . A sample dot product operation between one of 16 subsampled vector and filter vector is given below,

$$b = [b_0, b_1 ... b_8]$$
$$a = [a_0, a_1 ... a_8]$$
$$c = \sum_{n=0}^{8} a_i \cdot b_i$$

For a sub quadrant operated by a filter vector will create four c values which forms 2x2 array. From C values Z values are calculated with below function,

$$z_{i,j}^b = f(c_{i,j}^b) = \begin{cases} c_i & , c_i \geq 0 \\ 0 & , otherwise \end{cases} \text{ where } b = 0 \ldots 3, i = 0, 1, j = 0, 1$$

This would produce 64 values of Z, considering that each sub quadrant will be operated by four b-filter vectors.
The 64 values of Z are ordered as explained below, before passing as an input to second convolution operation.

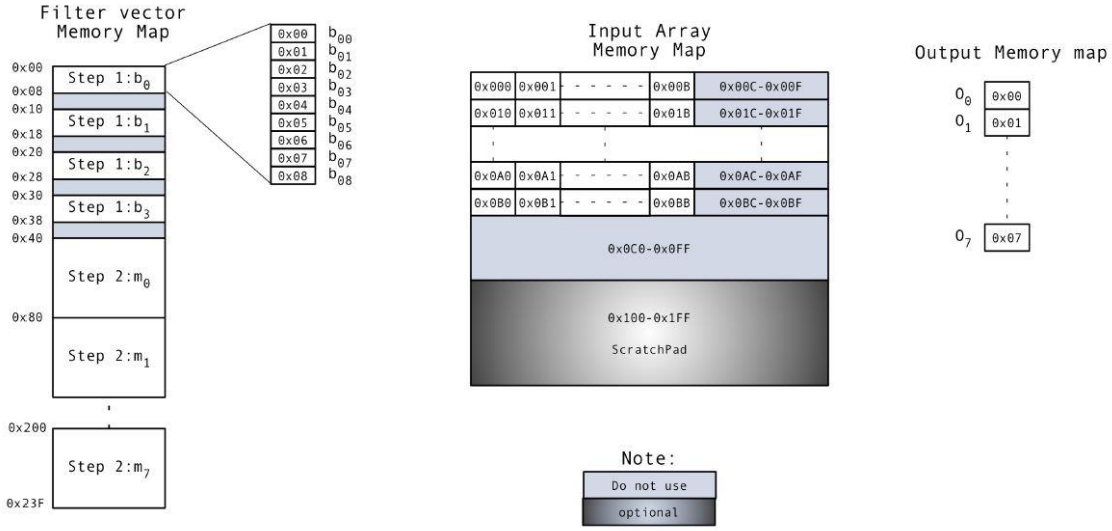
The first layer consists of 16 Z values form the operation with 16 sub quadrants with b0 similarly second layer is formed from the operation with b1, third from b2 ,fourth from b3.

Step two:

In step two, dot products are performed on step one output with
eight supplied [m] vector. Each m vector has 64 16 bit twos complement integer.
The output is 32bit int value which is again converted to 16 bit number using the same function as described in step one.

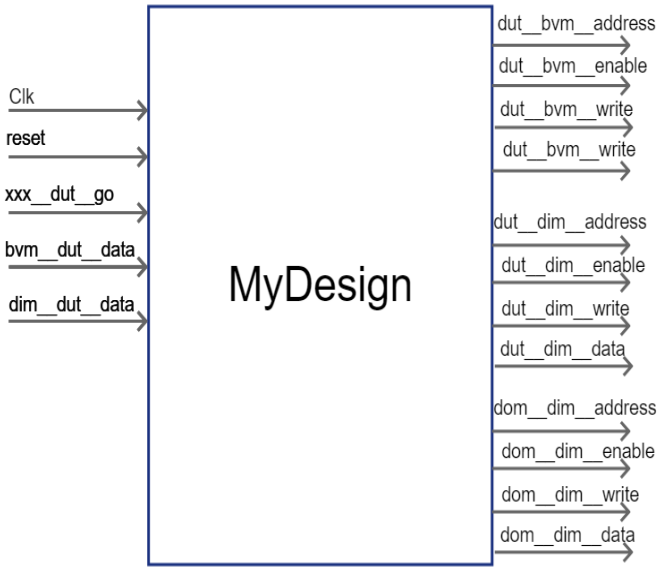
This will result in eight outputs, O and this will be the output of our system.

The storage memory maps for SRAMs are given below,

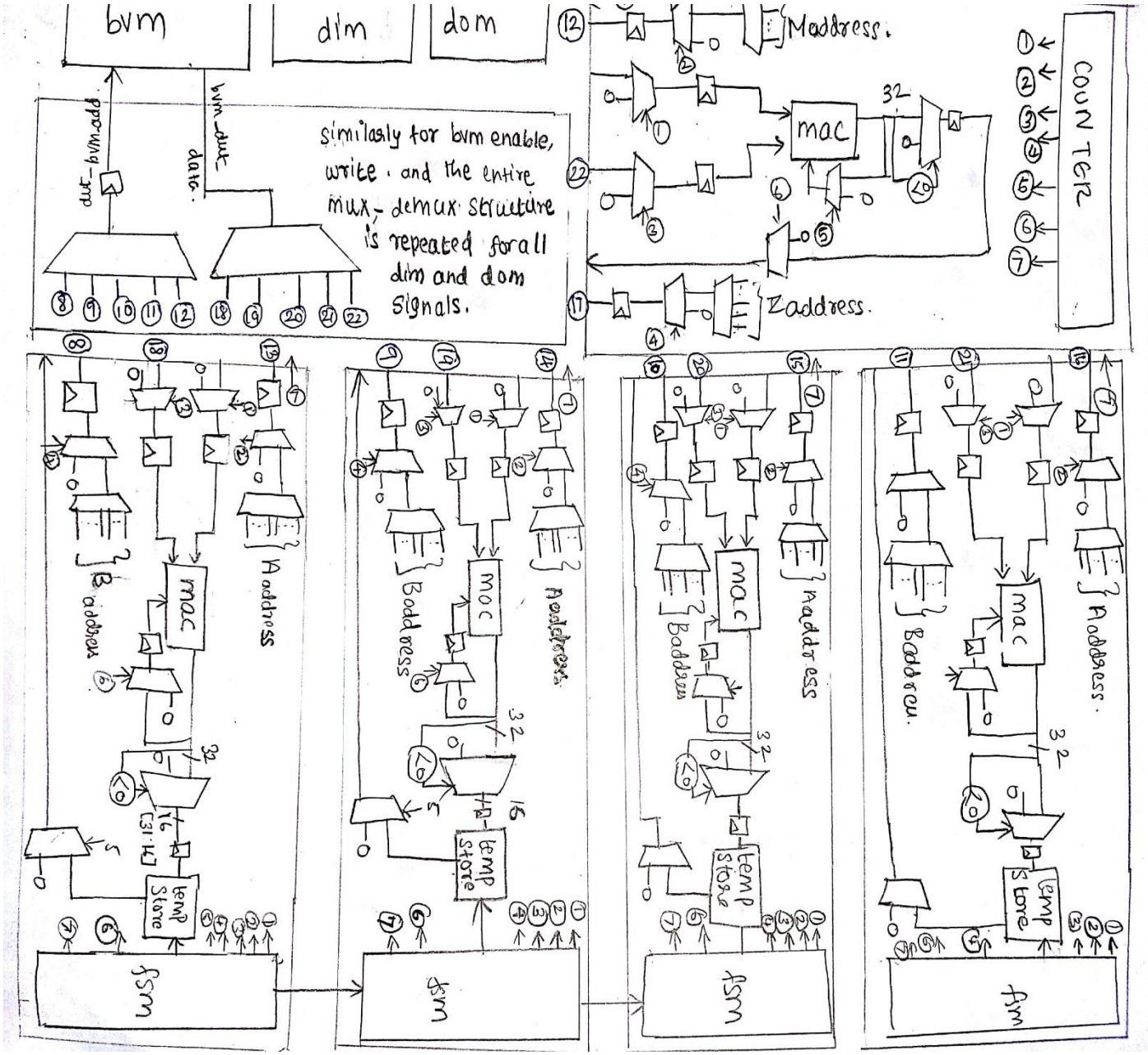


(The images are reproduced from given project technical specification document for better understanding).

High Level Design:



Datapath:



The input to mux in controller is a concatenated signal from all 4 stepone and one steptwo instance. It selects the only active instance at that particular time.

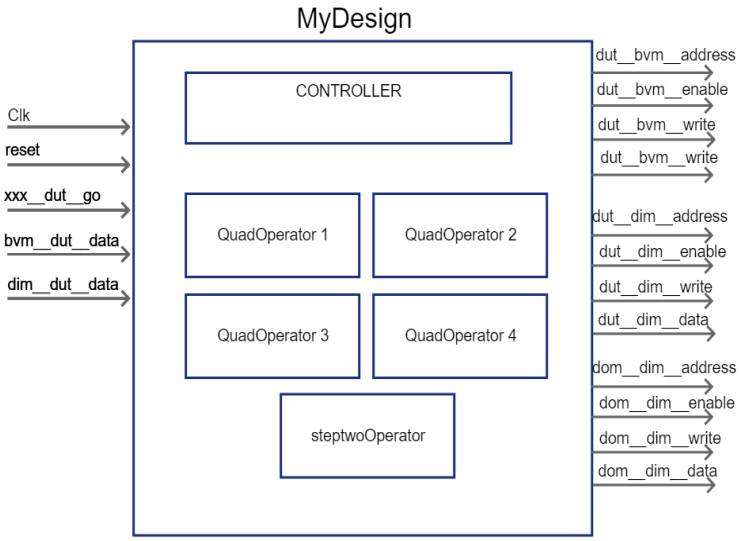
3. Interface Specification

Specification of inputs and outputs

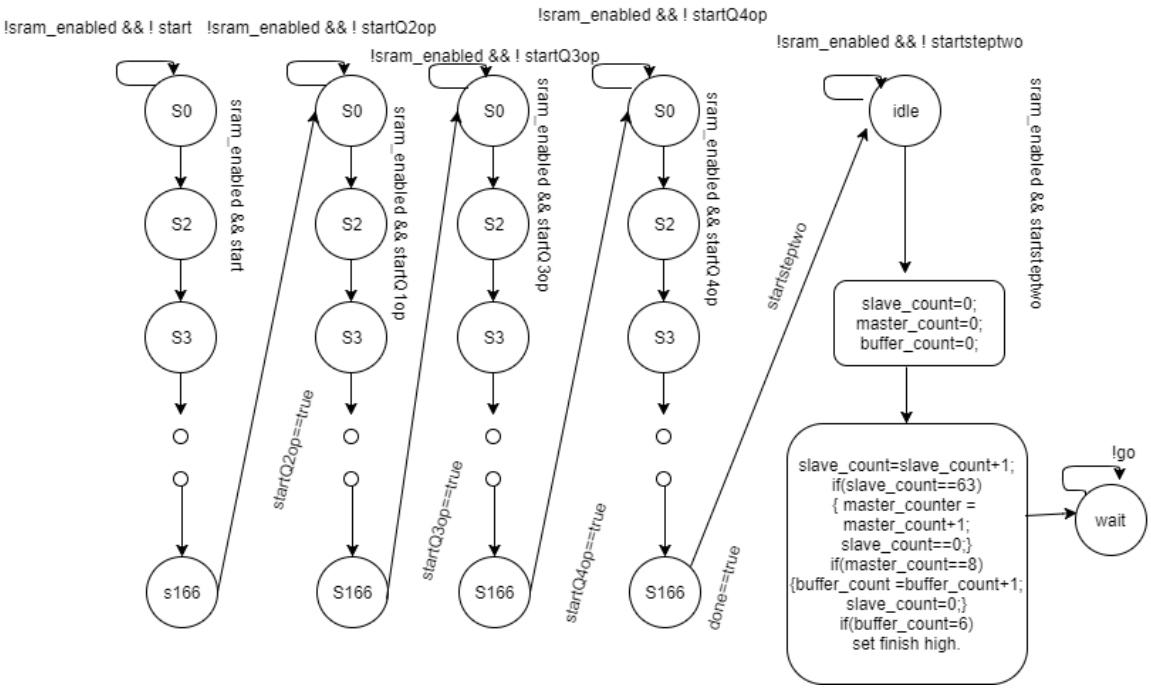
Signals	Direction	Width	Function
dut__xxx__finish	Output	1	Previous Operation has finished
xxx__dut__go	Input	1	Start Operation
dut__bvm__address	Output	10	Read/write to bvm SRAM at this address
dut__bvm__enable	Output	1	enable bvm SRAM
dut__bvm__write	Output	1	Write enable to bvm SRAM
dut__bvm__data	Output	16	Read bus for bvm SRAM
bvm__dut__data	Input	16	Write bus for bvm SRAM
dut__dim__address	Output	9	Read/write to dim SRAM at this address
dut__dim__enable	Output	1	enable dim SRAM
dut__dim__write	Output	1	Write enable to dim SRAM
dut__dim__data	Output	16	Read bus for dim SRAM
dim__dut__data	Input	16	Write bus for dim SRAM
dut__dom__address	Output	3	Read/write to dom SRAM at this address
dut__dom__data	Output	16	Write bus for dom SRAM
dut__dom__enable	Output	1	enable dom SRAM
dut__dom__write	Input	1	Write enable to dom SRAM
clk	Input	1	clock
reset	Input	1	reset

4. Technical Implementation

Design hierarchy



Controller Description



- FSM is used to control states for step one operations. Each QuadOperator has 166 states and as the state moves S0 to S166 all operations for a quadrant is completed.
- While a QuadOperator does not receive start signal it remains in the state S0. Once sram_enabled and start signals are given it moves to next state. At state 166 all operations are performed for the specific quadrant and it sets startQXOP (X=2,3,4) signal as high.StartQXOP shuts down the executed instance and starts the next.
- When startsteptwo signal is made high steptwo instance executes. The controller for steptwo are set of counters and operation of it are explained as pseudo code in above picture. Slave_counter counts from 0 to 63 and main_counter counts whenever Slave is 63 till it reaches 8. when main_counter is 8 buffer_count starts to provide buffer cycles. Since there is a clock delay from instance an address is sent, and the data received. Once all operations are finished, it goes to wait state setting the finish as high awaiting the next go signal.

5. Verification

The results of the simulation were compared against the results of the test bench which simulates three SRAMS loaded with sample input. The test bench was provided. Intermediate values where generated using given test bench and compared against timing diagram of the simulation. Various random sets where generated using seed and output of simulation was validated.

After data validation, the behavior of design when go signal goes high was verified. The test bench provides go signal when finish goes high and hence we have two iterations.

6. Results Achieved

timing_max_slow:

```
Information: Updating design information... (UID-85)
Warning: Design 'MyDesign' contains 1 high-fanout nets. A fanout number of 1000 will be used for delay calculations
involving these nets. (TIM-134)

****************************************
Report : timing
        -path full
        -delay max
        -max_paths 1
Design : MyDesign
Version: K-2015.06-SP1
Date   : Sun Nov 12 19:45:36 2017
****************************************

 # A fanout number of 1000 was used for high fanout net computations.

Operating Conditions: slow   Library: NangateOpenCellLibrary_PDKv1_2_v2008_10_slow_nldm
Wire Load Model Mode: top

  Startpoint: clk_r_REG1275_S5
              (rising edge-triggered flip-flop clocked by clk)
  Endpoint: clk_r_REG1185_S6
            (rising edge-triggered flip-flop clocked by clk)
  Path Group: clk
  Path Type: max

  Point                                    Incr       Path
  -----------------------------------------------------------------------------
  clock clk (rise edge)                  0.0000     0.0000
  clock network delay (ideal)            0.0000     0.0000
  clk_r_REG1275_S5/CK (DFFR_X1)                     0.0000 #   0.0000 r
  clk_r_REG1275_S5/Q (DFFR_X1)           0.4990     0.4990 r
  C1/Q3/U1/A[13] (MyDesign_DW02_mac_17)             0.0000     0.4990 r
  C1/Q3/U1/U1234/ZN (INV_X4)             0.0507     0.5497 f
  C1/Q3/U1/U1167/ZN (INV_X4)             0.3257     0.8754 r
  C1/Q3/U1/U960/Z (XOR2_X2)              0.4466     1.3220 r
  C1/Q3/U1/U957/ZN (INV_X4)              0.1599     1.4819 f
  C1/Q3/U1/U1051/ZN (NAND2_X4)           0.3504     1.8323 r
  C1/Q3/U1/U1371/ZN (OAI22_X2)           0.1810     2.0133 f
  C1/Q3/U1/U417/S (FA_X1)                0.8894     2.9027 f
  C1/Q3/U1/U414/CO (FA_X1)               0.5986     3.5013 f
  C1/Q3/U1/U407/S (FA_X1)                0.7993     4.3006 r
  C1/Q3/U1/U406/S (FA_X1)                0.7010     5.0016 f
  C1/Q3/U1/U1348/ZN (INV_X2)             0.1046     5.1062 r
  C1/Q3/U1/U1346/ZN (NAND2_X2)           0.0589     5.1651 f
  C1/Q3/U1/U1296/ZN (NAND2_X1)           0.0967     5.2618 r
  C1/Q3/U1/U1344/ZN (AND2_X1)            0.1706     5.4324 r
  C1/Q3/U1/U1049/ZN (OAI21_X2)           0.1167     5.5492 f
  C1/Q3/U1/U947/ZN (AOI21_X2)            0.2714     5.8206 r
  C1/Q3/U1/U1270/ZN (OAI21_X2)           0.1244     5.9449 f
  C1/Q3/U1/U1610/ZN (AOI21_X2)           0.3453     6.2902 r
  C1/Q3/U1/U63/Z (XOR2_X2)               0.3652     6.6555 r
  C1/Q3/U1/MAC[30] (MyDesign_DW02_mac_17)           0.0000     6.6555 r
  clk_r_REG1185_S6/D (DFFR_X1)           0.0000     6.6555 r
  data arrival time                                 6.6555

  clock clk (rise edge)                  7.0000     7.0000
  clock network delay (ideal)            0.0000     7.0000
  clock uncertainty                     -0.0500     6.9500
```

```
clk_r_REG1185_S6/CK (DFFR_X1)                    0.0000     6.9500 r
library setup time                              -0.2936     6.6564
data required time                                          6.6564
--------------------------------------------------------------------------
data required time                                          6.6564
data arrival time                                          -6.6555
--------------------------------------------------------------------------
slack (MET)                                                 0.0009
```

timing_max_slow_holdfixed

```
****************************************
Report : timing
        -path full
        -delay max
        -max_paths 1
Design : MyDesign
Version: K-2015.06-SP1
Date   : Sun Nov 12 19:48:48 2017
****************************************

 # A fanout number of 1000 was used for high fanout net computations.

Operating Conditions: slow   Library: NangateOpenCellLibrary_PDKv1_2_v2008_10_slow_nldm
Wire Load Model Mode: top

  Startpoint: clk_r_REG1275_S5
              (rising edge-triggered flip-flop clocked by clk)
  Endpoint: clk_r_REG1185_S6
            (rising edge-triggered flip-flop clocked by clk)
  Path Group: clk
  Path Type: max

  Point                                             Incr       Path
  --------------------------------------------------------------------------
  clock clk (rise edge)                          0.0000     0.0000
  clock network delay (ideal)                    0.0000     0.0000
  clk_r_REG1275_S5/CK (DFFR_X1)                  0.0000 #   0.0000 r
  clk_r_REG1275_S5/Q (DFFR_X1)                   0.4990     0.4990 r
  C1/Q3/U1/A[13] (MyDesign_DW02_mac_17)          0.0000     0.4990 r
  C1/Q3/U1/U1234/ZN (INV_X4)                     0.0507     0.5497 f
  C1/Q3/U1/U1167/ZN (INV_X4)                     0.3257     0.8754 r
  C1/Q3/U1/U960/Z (XOR2_X2)                      0.4466     1.3220 r
  C1/Q3/U1/U957/ZN (INV_X4)                      0.1599     1.4819 f
  C1/Q3/U1/U1051/ZN (NAND2_X4)                   0.3504     1.8323 r
  C1/Q3/U1/U1371/ZN (OAI22_X2)                   0.1810     2.0133 f
  C1/Q3/U1/U417/S (FA_X1)                        0.8894     2.9027 f
  C1/Q3/U1/U414/CO (FA_X1)                       0.5986     3.5013 f
  C1/Q3/U1/U407/S (FA_X1)                        0.7993     4.3006 r
  C1/Q3/U1/U406/S (FA_X1)                        0.7010     5.0016 f
  C1/Q3/U1/U1348/ZN (INV_X2)                     0.1046     5.1062 r
  C1/Q3/U1/U1346/ZN (NAND2_X2)                   0.0589     5.1651 f
  C1/Q3/U1/U1296/ZN (NAND2_X1)                   0.0967     5.2618 r
  C1/Q3/U1/U1344/ZN (AND2_X1)                    0.1706     5.4324 r
  C1/Q3/U1/U1049/ZN (OAI21_X2)                   0.1167     5.5492 f
  C1/Q3/U1/U947/ZN (AOI21_X2)                    0.2714     5.8206 r
  C1/Q3/U1/U1270/ZN (OAI21_X2)                   0.1244     5.9449 f
  C1/Q3/U1/U1610/ZN (AOI21_X2)                   0.3453     6.2902 r
  C1/Q3/U1/U63/Z (XOR2_X2)                       0.3652     6.6555 r
  C1/Q3/U1/MAC[30] (MyDesign_DW02_mac_17)        0.0000     6.6555 r
  clk_r_REG1185_S6/D (DFFR_X1)                   0.0000     6.6555 r
  data arrival time                                         6.6555
```

```
  clock clk (rise edge)                          7.0000     7.0000
  clock network delay (ideal)                    0.0000     7.0000
  clock uncertainty                             -0.0500     6.9500
  clk_r_REG1185_S6/CK (DFFR_X1)                  0.0000     6.9500 r
  library setup time                            -0.2936     6.6564
  data required time                                        6.6564
  --------------------------------------------------------------------------
  data required time                                        6.6564
  data arrival time                                        -6.6555
  --------------------------------------------------------------------------
  slack (MET)                                               0.0009
```

timing_min_fast_holdcheck:

```
Information: Updating design information... (UID-85)
Warning: Design 'MyDesign' contains 1 high-fanout nets. A fanout number of 1000 will be used for delay calculations
involving these nets. (TIM-134)

****************************************
Report : timing
        -path full
        -delay min
        -max_paths 1
Design : MyDesign
Version: K-2015.06-SP1
Date   : Sun Nov 12 19:47:14 2017
****************************************

 # A fanout number of 1000 was used for high fanout net computations.

Operating Conditions: fast   Library: NangateOpenCellLibrary_PDKv1_2_v2008_10_fast_nldm
Wire Load Model Mode: top

  Startpoint: clk_r_REG788_S4
              (rising edge-triggered flip-flop clocked by clk)
  Endpoint: clk_r_REG789_S5
            (rising edge-triggered flip-flop clocked by clk)
  Path Group: clk
  Path Type: min

  Point                                   Incr       Path
  -----------------------------------------------------------
  clock clk (rise edge)                   0.0000     0.0000
  clock network delay (ideal)             0.0000     0.0000
  clk_r_REG788_S4/CK (DFFS_X2)            0.0000 #   0.0000 r
  clk_r_REG788_S4/Q (DFFS_X2)             0.0674     0.0674 r
  U9679/ZN (OAI22_X2)                     0.0154     0.0829 f
  clk_r_REG789_S5/D (DFFR_X1)             0.0000     0.0829 f
  data arrival time                                  0.0829

  clock clk (rise edge)                   0.0000     0.0000
  clock network delay (ideal)             0.0000     0.0000
  clock uncertainty                       0.0500     0.0500
  clk_r_REG789_S5/CK (DFFR_X1)            0.0000     0.0500 r
  library hold time                       0.0022     0.0522
  data required time                                 0.0522
  -----------------------------------------------------------
  data required time                                 0.0522
  data arrival time                                 -0.0829
  -----------------------------------------------------------
  slack (MET)                                        0.0306
```

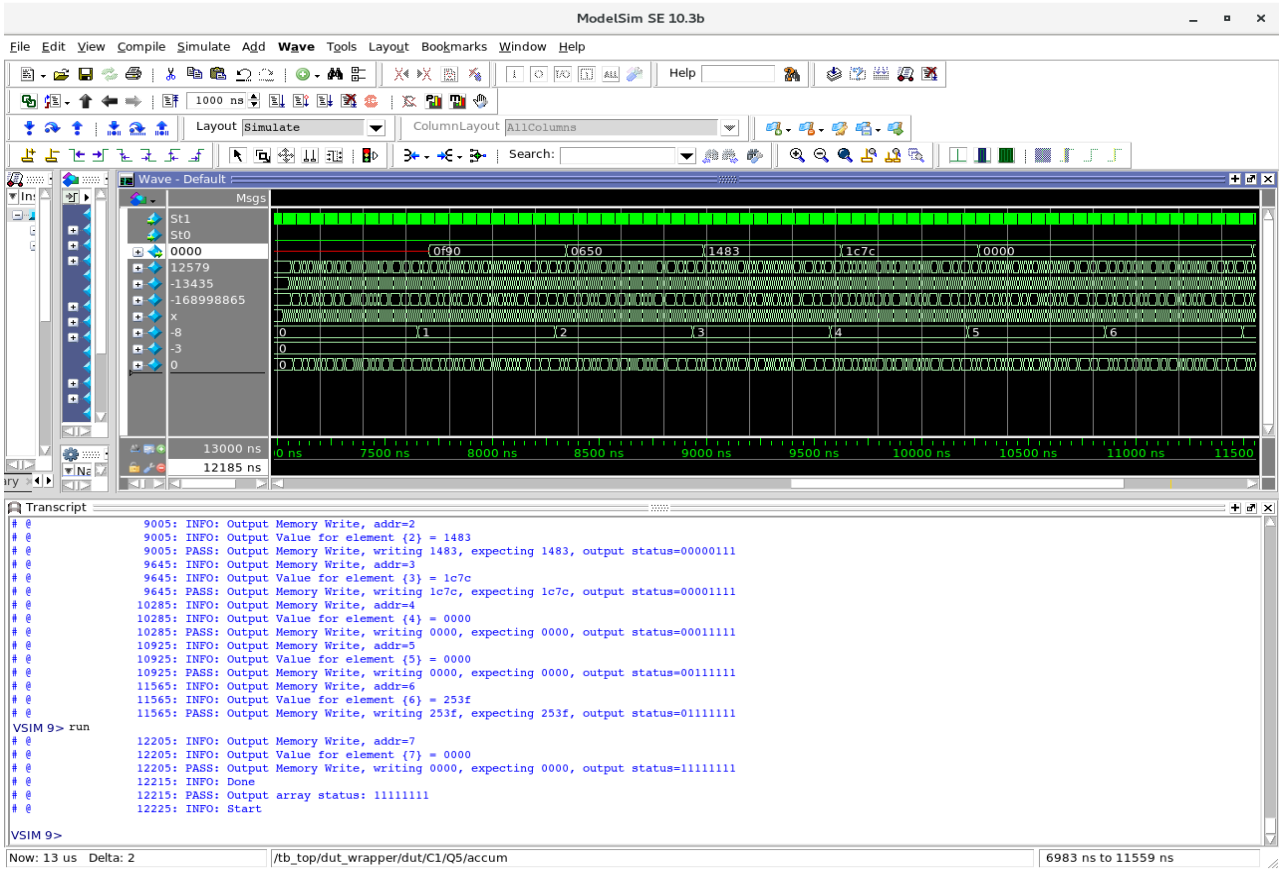
Area Report:

```
Number of ports:                          621
Number of nets:                         16409
Number of cells:                        13967
Number of combinational cells:          12137
Number of sequential cells:              1824
Number of macros/black boxes:               0
Number of buf/inv:                       1694
Number of references:                      70

Combinational area:                15263.346126
Buf/Inv area:                        947.758007
Noncombinational area:              9917.011959
Macro/Black Box area:                  0.000000
Net Interconnect area:      undefined  (No wire load specified)

Total cell area:                   25180.358085
Total area:                 undefined
```

Simulation Result:



7. Conclusions

Below is the summary of the results obtained

- Throughput = 1171*7 = 8197ns.
- Area = 25180.3581(um^2)
- Clock period = 7 ns