MACHINE LEARNING ACCELERATOR

Negin Safari 810197525

Dr. Navvabi

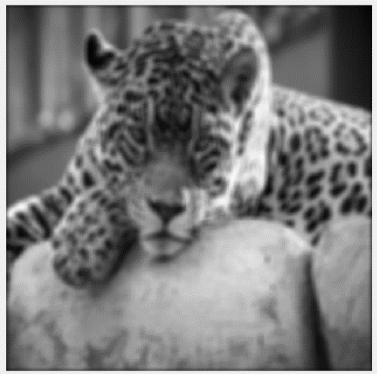
Accelerators

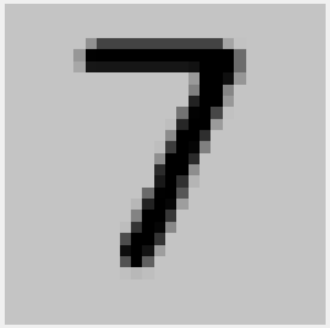
- Higher performance
- Higher speed
- Higher throughput
- Lower power

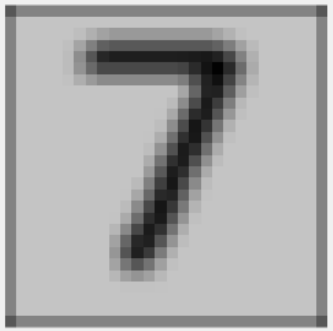
Convolution

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Matlab blurring using convolution







```
rgb = imread('yelsev.png');

gray = rgb2gray(rgb);
subplot(1,2,1), imshow(gray);
result = zeros(size(gray));
m = 3;
wgray = zeros(size(gray) + 2*floor(m/2));
wgray((floor(m/2) + 1):(28+floor(m/2)),(floor(m/2) + 1):(28+floor(m/2))) = gray;
k = ones(m);
k = k/(m*m);
y = 0;
ik = 0; jk=0;
for row = (floor(m/2) + 1):(28+floor(m/2)) %2
    for col = (floor(m/2) + 1):(28+floor(m/2)) %3
        y = 0;
        ik = 0;
        jk=0;
        for i = (row - floor(m/2)) : (row + floor(m/2)) %1 - 3
            ik= ik + 1;
         for j = (col - floor(m/2)) : (col + floor(m/2)) %2 - 4
            jk= jk + 1;
            y = y + k(ik,jk) * wgray(i,j);
         end
          jk = 0;
        end
        result(row-floor(m/2),col-floor(m/2)) = y;

    end
end
tt = uint8(result);
subplot(1,2,2),imshow(uint8(result));
```

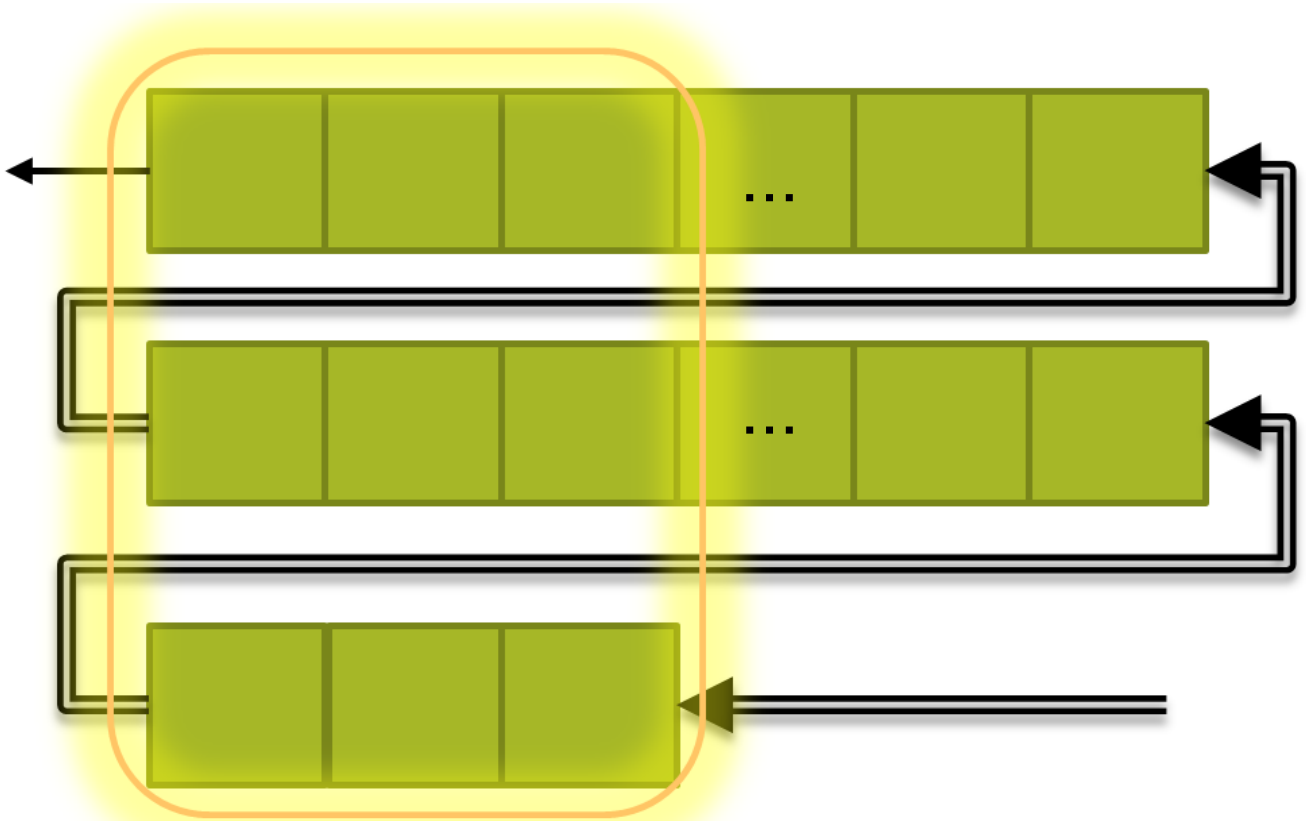
Input, kernel and output matrices

	1	2	3	4	5	6	7	8	9	10	11
1	196	196	196	196	196	196	196	196	196	196	196
2	196	196	196	196	196	196	196	196	196	196	196
3	196	196	196	196	196	196	196	196	196	196	196
4	196	196	196	196	196	196	196	145	67	67	67
5	196	196	196	196	196	196	151	1	0	0	0
6	196	196	196	196	196	196	167	23	22	22	22
7	196	196	196	196	196	196	196	196	196	196	196
8	196	196	196	196	196	196	196	196	196	196	196
9	196	196	196	196	196	196					
10	196	196	196	196	196	196					
11	196	196	196	196	196	196					
12	196	196	196	196	196	196					
13	196	196	196	196	196	196					
14	196	196	196	196	196	196					
15	196	196	196	196	196	196					
16	196	196	196	196	196	196					

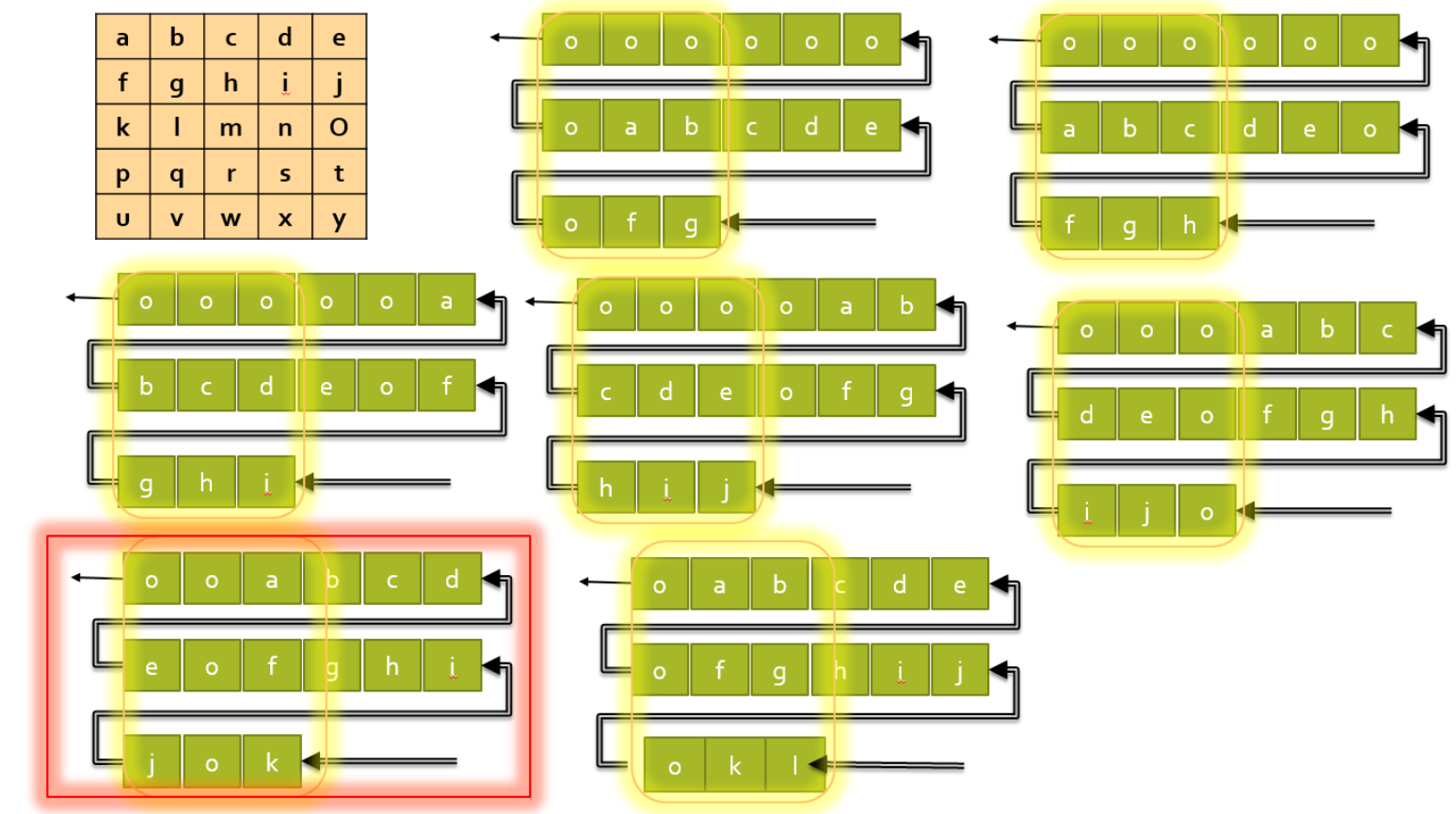
0.1111	0.1111	0.1111
0.1111	0.1111	0.1111
0.1111	0.1111	0.1111

	1	2	3	4	5	6	7	8	9
1	87.1111	130.6667	130.6667	130.6667	130.6667	130.6667	130.6667	130.6667	130.666
2	130.6667	196.0000	196.0000	196.0000	196.0000	196.0000	196.0000	196.0000	196.000
3	130.6667	196.0000	196.0000	196.0000	196.0000	196.0000	190.3333	176	161.666
4	130.6667	196.0000	196.0000	196.0000	196.0000	191.0000	163.6667	127.5556	96.444
5	130.6667	196.0000	196.0000	196.0000	196.0000	187.7778	141.2222	85.7778	38.555
6	130.6667	196.0000	196.0000	196.0000	196.0000	187.7778	146.8889	105.7778	72.888
7	130.6667	196.0000	196.0000	196.0000	196.0000	192.7778	173.5556	154.2222	138.111
8	130.6667	196.0000	196.0000	196.0000	196.0000	196.0000	196.0000	196.0000	196.000
9	130.6667	196.0000	196.0000	196.0000	196.0000	196.0000	196.0000	196.0000	196.000
10	130.6667	196.0000	196.0000	196.0000	196.0000	196.0000	196.0000	196.0000	196.000

Registers

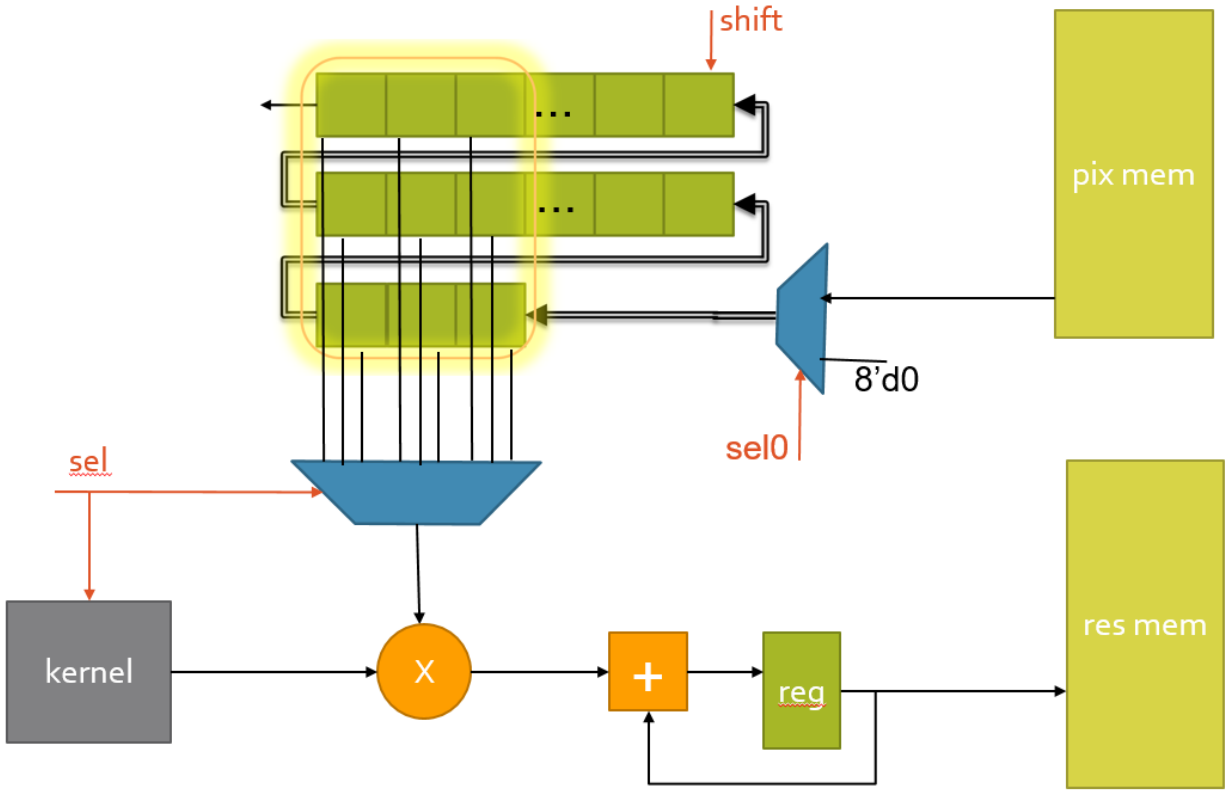


Calculation

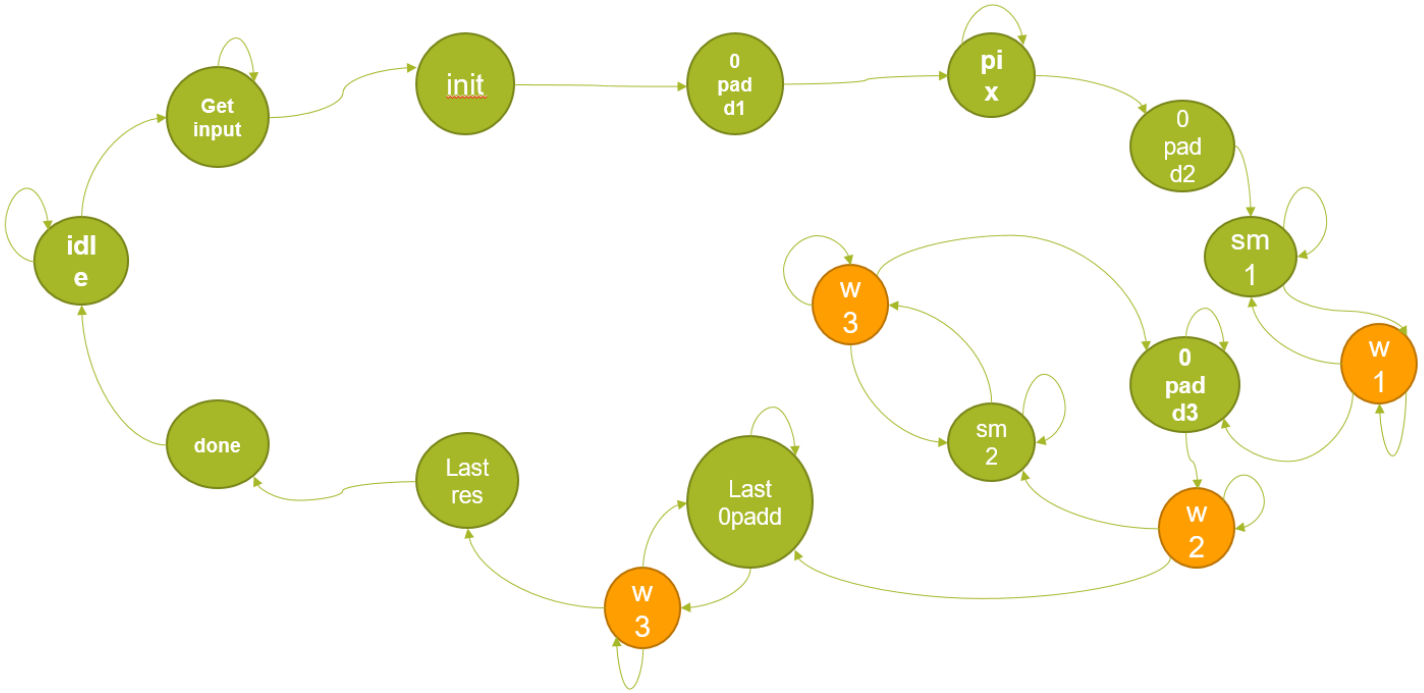


Convolution with one multiplier

Data path

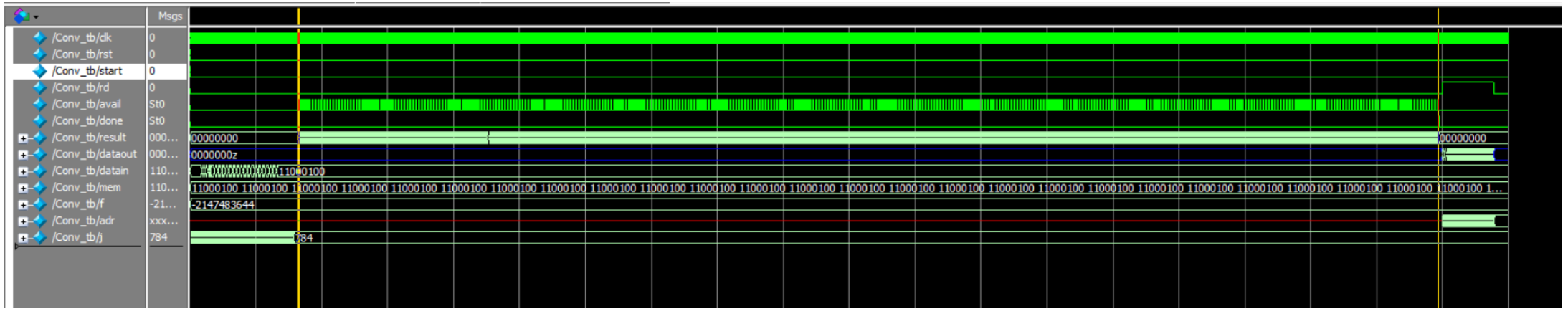


Control unit



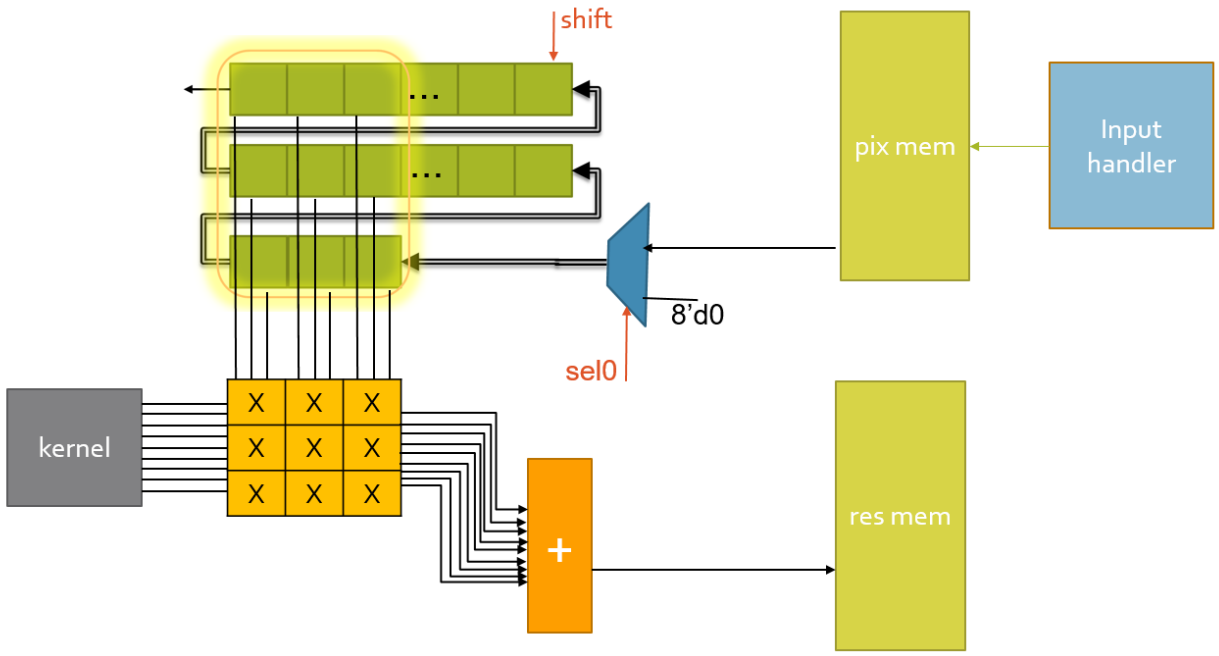
Wave form

- Number of cycles: **9440**

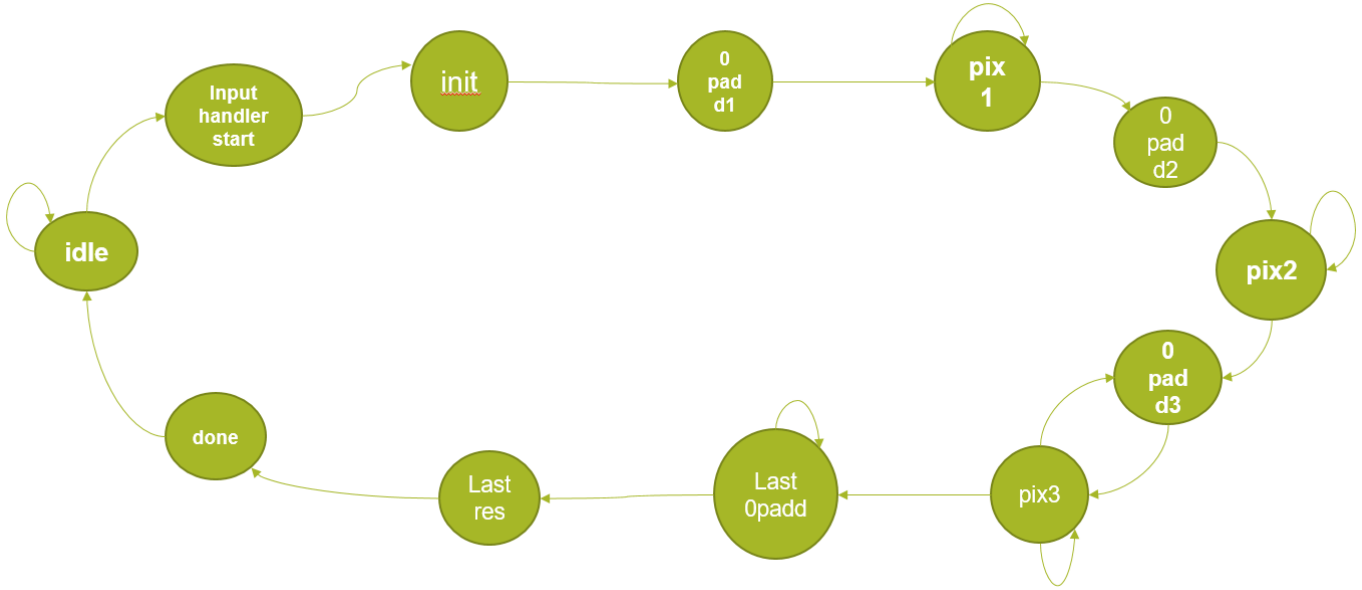


Convolution with nine multiplier

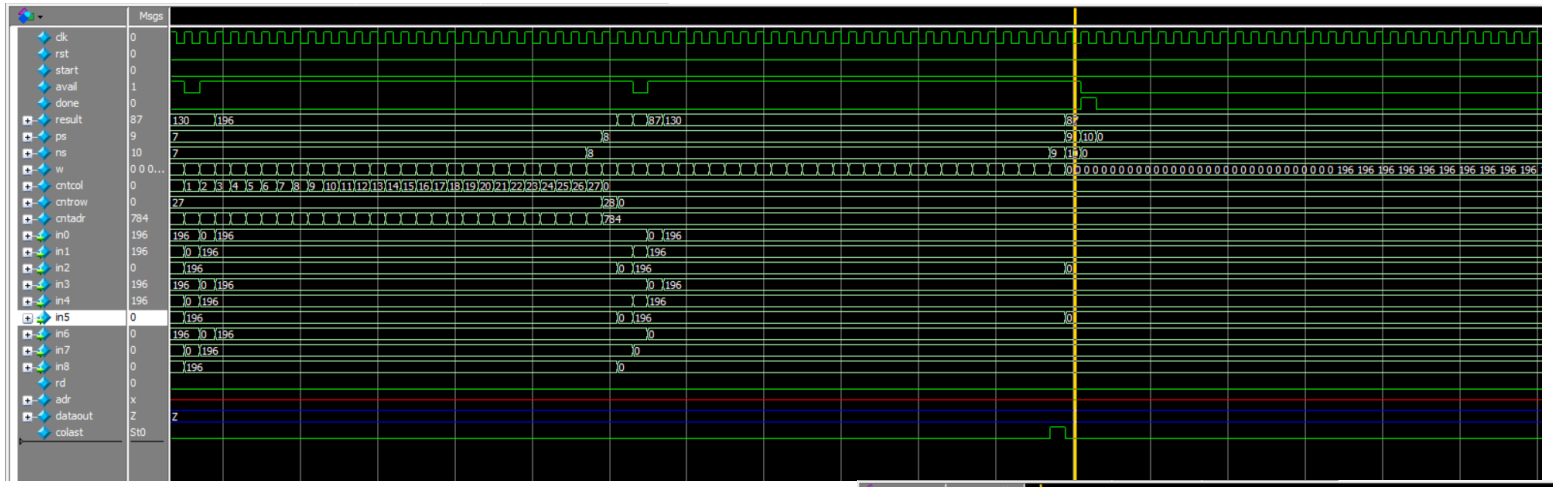
Data path

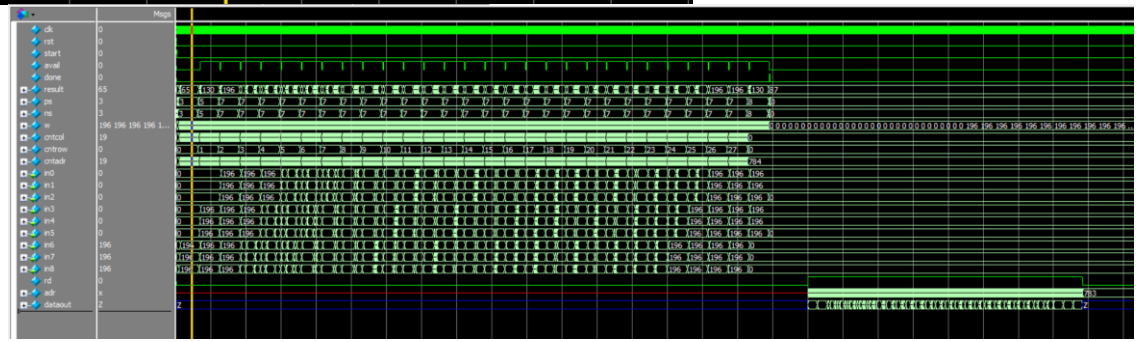

shift
pix mem
Input handler
8'd0
sel0
kernel
X X X
X X X
X X X
+
res mem

Control unit



Wave forms

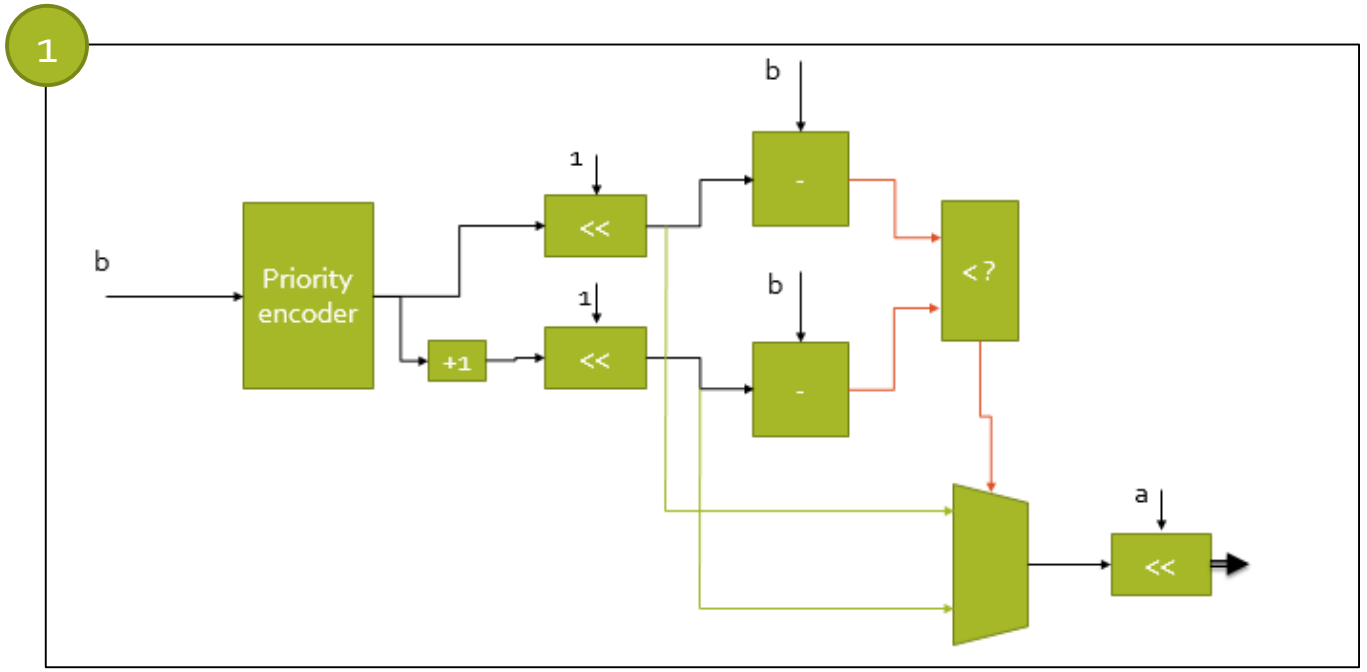




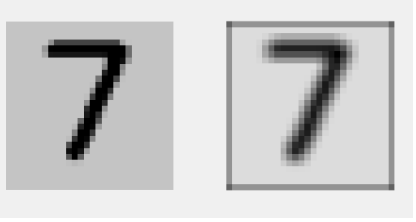
Number of cycles: 840

Lower accuracy multipliers

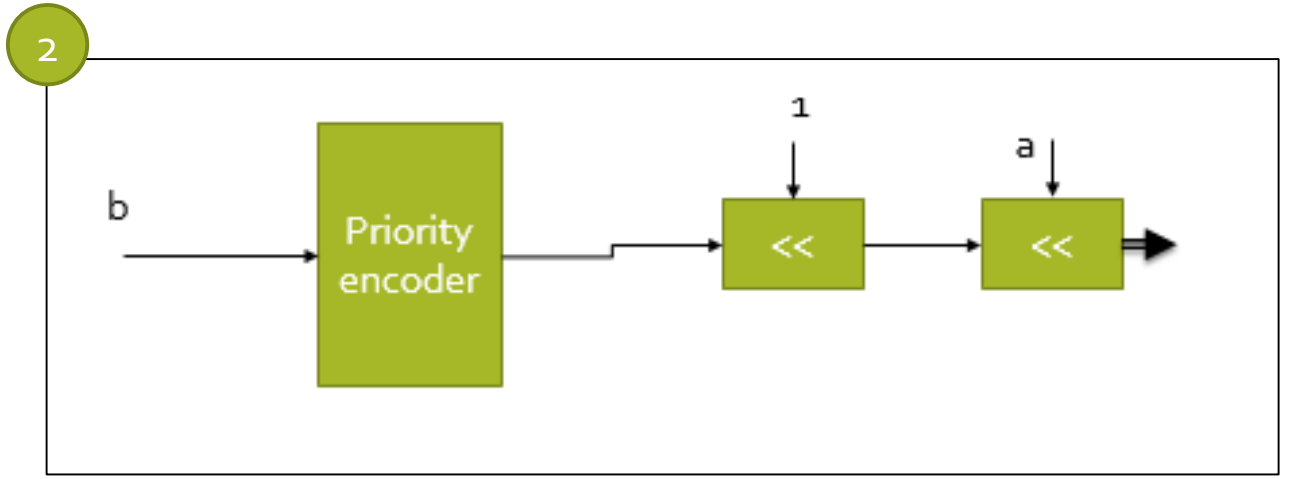
1: Round to nearest power of two



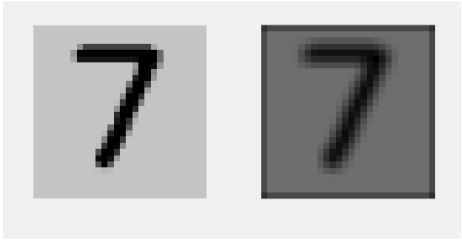
Result:



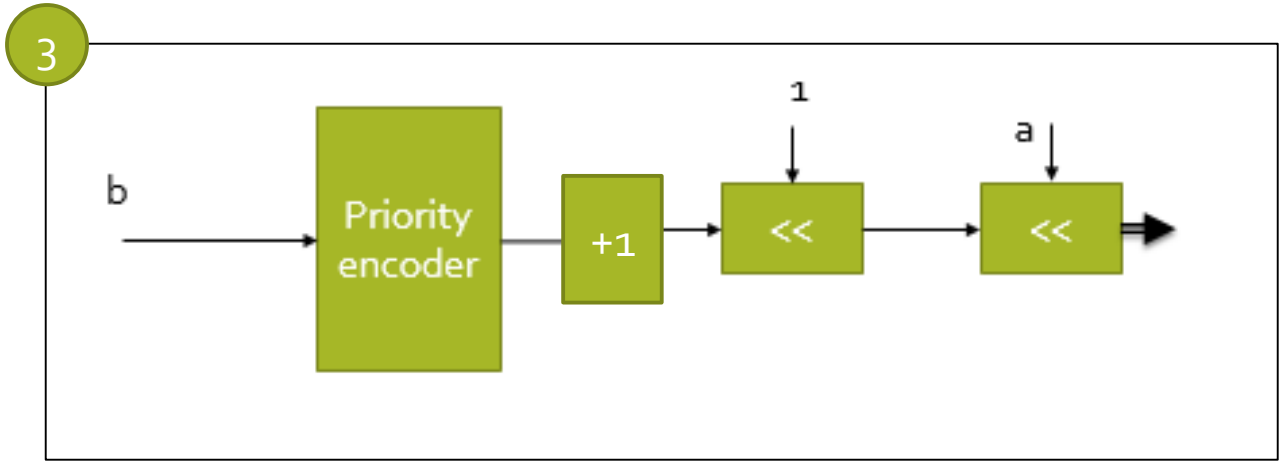
2: Round down to nearest power of two



Result:

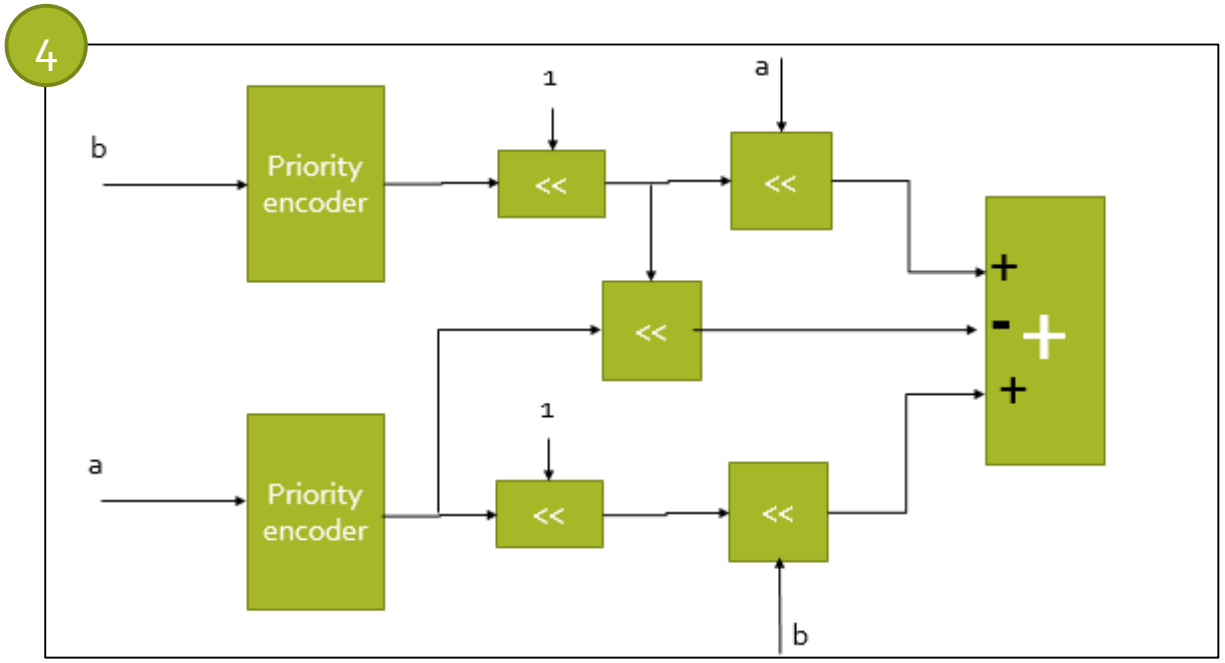


3: Round up to nearest power of two



Result:

4: Rounding- Based Approximate Multiplier(ROBA)



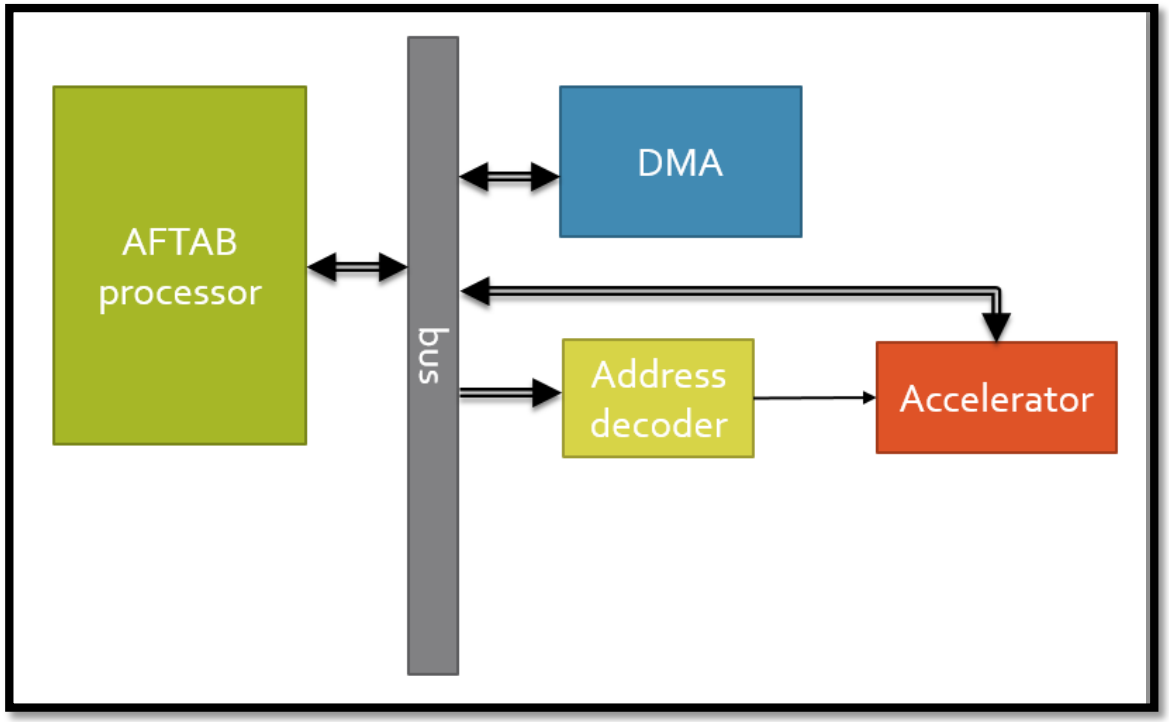
Result:

Lower accuracy multipliers

multiplier	RMSE	area
1	21.9665	100
2	74.8395	82
3	21.9665	91
4	27.0707	170

Area of a combinational 8 bit multiplier: 103

Processor and off-chip accelerator communication



Thank you for your attention