Wilson Sa, Kevin Rohan

1. System Overview

- Necessity of Improving efficiency in sparse matrix multiplication:
 - Widely used in modern day applications of feature extraction, graph analysis etc.
 - Sparse Matrix multiplication is costly due to the compression format used
- Factors Considered for implementation:
 - Using a compression technique which makes the code more parallel.
 - Testing it on a FPGA which is more programmable to suit the functionality.

2. Approach

 Sparse matrices are represented in following compression formats.

(a)
$$A = \begin{bmatrix} 1 & 4 & 0 & 0 & 0 \\ 0 & 2 & 3 & 0 & 0 \\ 5 & 0 & 0 & 7 & 8 \\ 0 & 0 & 9 & 0 & 6 \end{bmatrix}$$
(b)
$$COO \begin{cases} data = \begin{bmatrix} 1 & 4 & 2 & 3 & 5 & 7 & 8 & 9 & 6 \end{bmatrix} \\ row = \begin{bmatrix} 0 & 0 & 1 & 1 & 2 & 2 & 2 & 3 & 3 \end{bmatrix} \\ col = \begin{bmatrix} 0 & 1 & 1 & 2 & 0 & 3 & 4 & 2 & 4 \end{bmatrix} \end{cases}$$
(c)
$$CSR \begin{cases} data = \begin{bmatrix} 1 & 4 & 2 & 3 & 5 & 7 & 8 & 9 & 6 \end{bmatrix} \\ ptr = \begin{bmatrix} 0 & 2 & 4 & 7 & 9 \end{bmatrix} \\ col = \begin{bmatrix} 0 & 1 & 1 & 2 & 0 & 3 & 4 & 2 & 4 \end{bmatrix} \end{cases}$$
(d)
$$CSC \begin{cases} data = \begin{bmatrix} 1 & 5 & 4 & 2 & 3 & 9 & 7 & 8 & 6 \end{bmatrix} \\ row = \begin{bmatrix} 0 & 2 & 0 & 1 & 1 & 3 & 2 & 2 & 3 \end{bmatrix} \\ ptr = \begin{bmatrix} 0 & 2 & 4 & 6 & 7 & 9 \end{bmatrix}$$

CSC representation is more parallelizable due to columnwise access while multiplication. It is also a compressed format. This representation is used to perform multiplication on FPGA.

Software: Vivado HLS

FPGA: Artix 7

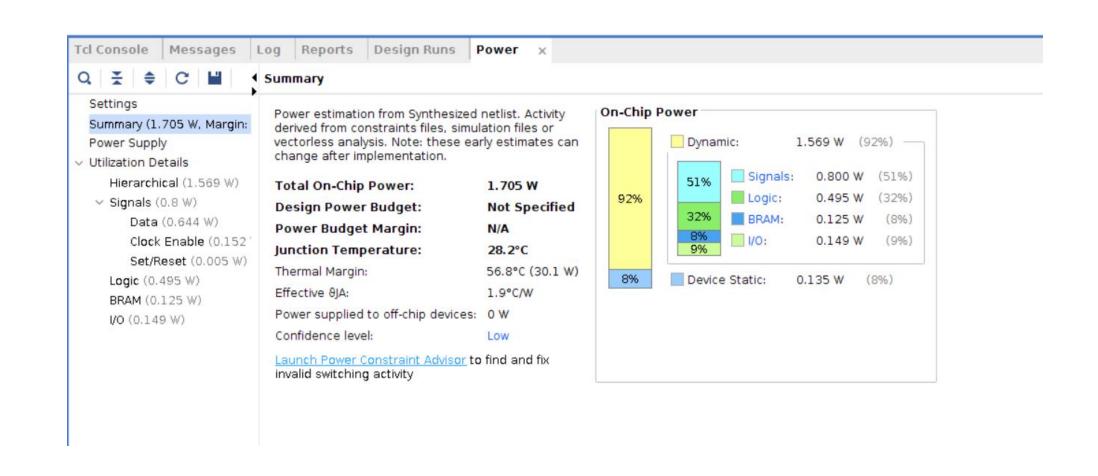
4. Conclusion

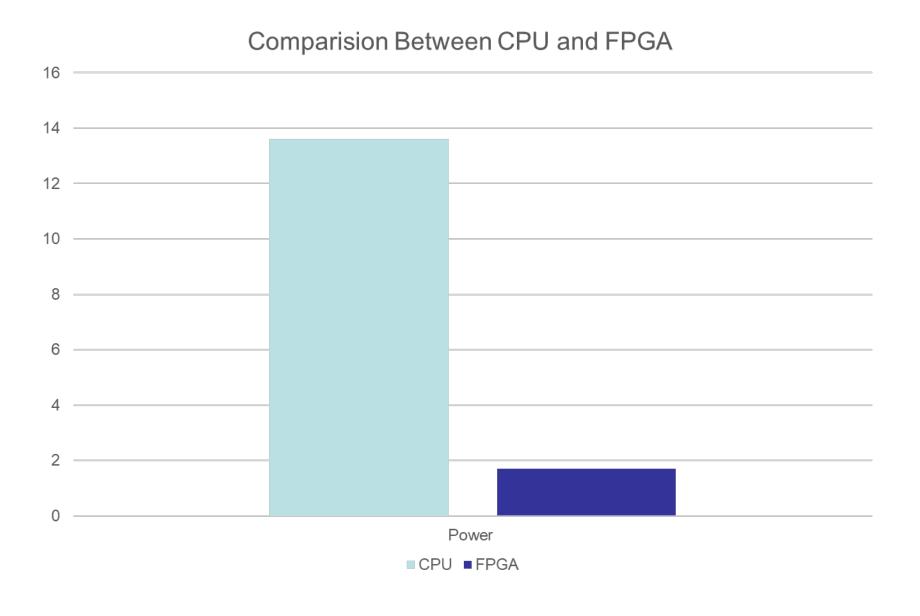
3. Results

Utilization of FPGA resources:

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	0.0	-
Expression	-	-	0	414
FIFO	-	-	-	-
Instance	-	-	-	-
Memory	8	-	0	0
Multiplexer	-	-	0.750	242
Register	-	-	366	-
Total	8	0	366	656
Available	40	40	16000	8000
Utilization (%)	20	0	2	8

- Power utilized on Laptop: 13.632W
- Power Analysis obtained on FPGA:





Using CSC representation and utilizing FPGA as the hardware the power was reduced by 87.5%. This is due to the fact random accesses in memory were made more organized using the CSC representation which provides a more serialized approach and the Look up tables which was used in the FPGA.