Deliverable report

"Diego Oniarti": Mat: 257835, diego.oniarti@studenti.unitn.it, GitRepo: https://github.com/diego-oniarti/GPU_homework

***Abstract*—[max 200 words]**
The sparse matrix-dense vector multiplication (SpMV) is a common linear algebra operation involving a sparse matrix and a dense vector. SpMV is widely used in many real-world applications such as ...
This deliverable discusses ...
***Index Terms*—Sparse Matrix, SpMV, CUDA, Parallelization, Storage Format**

I. Introduction

[max 300 words]
Introduce SpMV application and its parallelization challenges ...

II. Problem Statement

Define the problem statement, including a description of the storage format used and a brief discussion of the parallelization approach (e.g., using CUDA).

A. Storage Format

Details about the format (e.g., CSR, COO, etc.) ...

B. Parallelization

Describe the CUDA implementation approach ...

III. State of the Art

Describe the available library solutions for the SpMV problem, including references and citations [1]–[3].

IV. Methodology and Contributions

Describe the methodology used during the analysis, the compared algorithms and the expected outcomes. Use pseudo-codes like Algorithm 1 to describe your own implemented kernels.
Include at least the following implementations:

- Naive CPU implementation
- Optimised CPU implementation based on cache behaviour
- GPU naive implementation

For the analysis include

- Valgrind and runtime comparison between the CPU implementations
- Runtime CPU vs GPU comparison looping over different matrix dimensions

V. System Description and Experimental Set-up

Use this section to describe system, dataset and experimental set-up.

Algorithm 1 Algorithm for the vector addition

Require: The input vectors a and b of size nnz.

```
procedure FUNCTION(a, b, nnz)
    c ← ∅
    for i in {1 ... nnz} parallel do
        c[i] = a[i] + b[i]
    end for
    return c                                  ▷ the result vector
end procedure
```

A. System Description

Describe the used system and the software environment (CUDA version, GCC version, ...). Decide which information are valuable to group into a table like Table I and which are more valuable to be described in the text.

System	Processor	Cores per Socket	RAM	Accelerator
Leonardo	Intel Xeon 8358 CPU	32 at 2.6 GHz	494 GB	NVIDIA A100
Local	2× AMD Rome 7742	64 at 2.25 GHz	1,007 GB	NVIDIA A100

TABLE I
System details

B. Dataset description

Discribe the used dataset and the reasons of your choice. List the used input matrices and all the information that you think are valuable to show (number of non-zero elements, sparsity ratio, ...); Table II gives you a possible example.

C. Experimental Set-up

Explain the benchmark metrics and all the experimental set-up (warm-up cycles, number of performed runs...).

VI. Experimental Results

Present and discuss results. Include plots and tables when required (like Figure 1). Do not simply describe the figures;

Dataset	|V|	|E|	Avg. Degree	Diameter	Benchmark
com-youtube [?]	1,134,890	2,987,624	5.27	20	Naive vs optimized CPU
amazon-ratings [?]	2,146,057	5,743,146	5.35	28	
cit-patents [?]	3,764,117	16,511,741	8.77	26	
com-lj [?]	3,997,962	34,681,189	17.35	17	
dblp [?]	4,000,148	8,649,011	4.32	50	
flickr-link [?]	1,624,992	15,472,576	19.04	24	CPU vs GPU
Wiki-Talk [?]	2,388,953	4,656,682	3.90	9	
com-Orkut [?]	3,072,441	117,185,083	76.28	9	
dbpedia-link [?]	18,265,512	126,888,089	13.98	12	

TABLE II
Caption

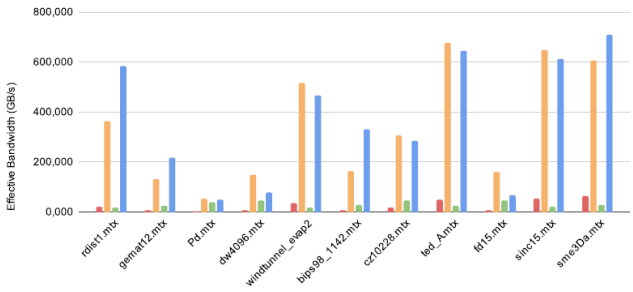

Fig. 1. Caption

criticise the achieved results by underlining how they confirm/differ from the expected outcomes described in Section IV.

VII. Conclusions

[max 200 words]
Summarize findings and future work ...

References

[1] I. Asimov, *I, Robot*. New York: Gnome Press, 1950.
[2] I. Svevo, *La coscienza di Zeno*. Bologna: L. Cappelli Editore, 1923.
[3] L. Pirandello, *Uno, nessuno e centomila*. Firenze: R. Bemporad e Figli, 1926.