



UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Escola Superior d'Enginyeries Industrial,
Aeroespacial i Audiovisual de Terrassa

Improving the performance of a deep learning framework on high-performance computing (HPC) systems

Document:

Budget

Author:

Martí Llopart Font

Director /Co-director:

Bernardo Morcego / Ramon Pérez

Degree:

MÀSTER UNIVERSITARI EN GESTIÓ
D'EMPRESSES DE TECNOLOGIA I D'ENGINYERIA

Examination session:

Autumn/Winter, 2022-2023

MASTER FINAL THESIS



Table of contents

TABLE OF CONTENTS.....	I
LIST OF TABLES.....	II
LIST OF FIGURES.....	III
1. PLANIFICATION.....	IV - V
2. BUDGETING.....	1

List of tables

Table 11 – Yearly personals costs working at the BSC.....1



List of figures

Figure 1 – This figure displays a Gantt chart of my project, spanning 11 months from March 2022 to January 2023. The blue lines represent time spent with each taskV

1. Planification

This section fully details the steps I took to develop this project, from brainstorming to writing the full thesis. For better visualization of all the steps, I've developed a Gantt diagram that displays the different phases of progression of the project. Including everything, I spent over 12 months with this project, even though the progression wasn't linear. First, I will detail all the phases of this project in a pointed list.

1. Brainstorming: This first step consisted in putting over the table what were the options available within my research team. In this step, I had to do a lot of fundamental learning during this stage, because many concepts were new to me. After that, many ideas appeared.
2. Researching the literature: After elaborating these ideas, the next part was to research the literature if other academics had found a similar or alternative solution to the problem at hand. In addition, I had to research the previous literature our idea was based on.
3. Researching reproducibility: When a good enough theoretical foundation was found, I had to code a program capable of reproducing what was stated in the different studies. This task was very time-consuming because most papers are not well documented for reproducibility, and software packages and platforms are continuously changing and being updated.
4. Coding a use case: Once the literature was reproduced, I had to develop a specific use case including all the necessary features our idea could include.
5. Coding a program implementation: the fifth step consists in implementing the program my team developed to the use case I had come up with using Python.
6. Obtaining the results: Then, after finishing all the coding, it was time to put it to work in the different computing clusters we had available. If the results matched our expectations, we could go on to write the report.
7. Writing the report: The final step was to put all of my work in a report format that could be easily comprehended, given the depth of our study. It was a challenging task, but I believe the results meet the requirements of the university.

GANTT CHART

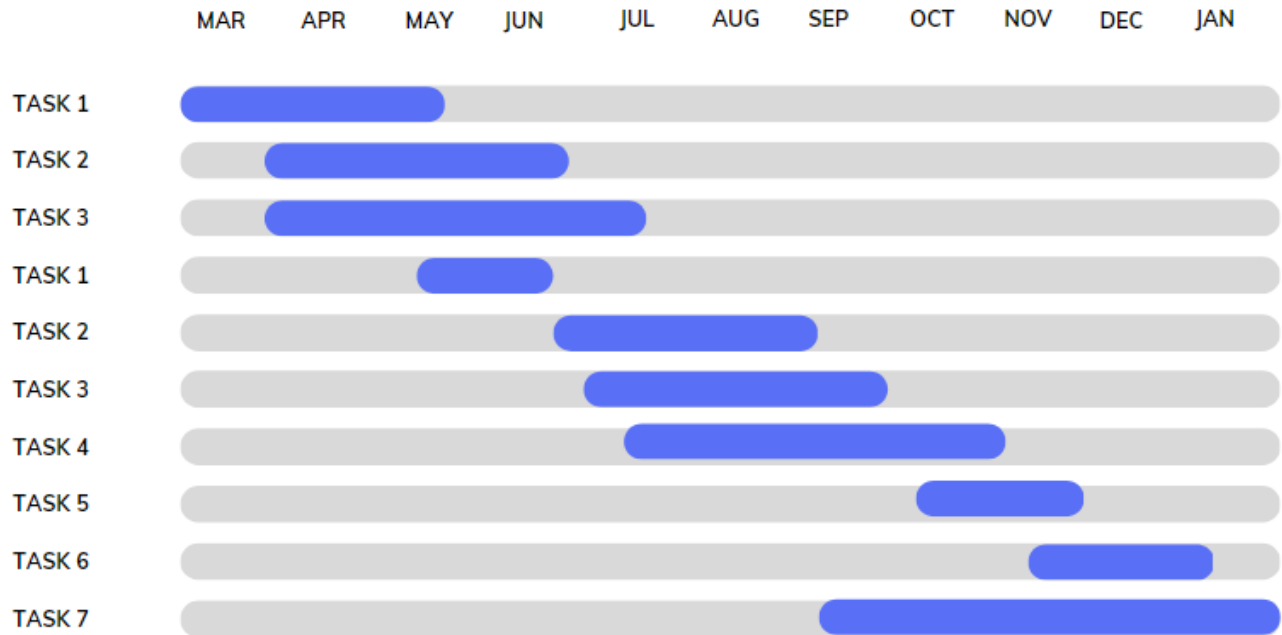


Figure 1 – This figure displays a Gantt chart of my project, spanning 11 months from March 2022 to January 2023. The blue lines represent time spent with each task.

Figure 1 displays a Gantt Chart of my project. It can be seen from the figure how we had to go back to task 1 once. This is because one of our first ideas failed because we couldn't reproduce the results of the literature. Because of that, we had to go back to task 1 and brainstorm again.

2. Budgeting

This study has been done in collaboration with the BSC, where I worked for a year as a Master Student researcher. My company provided all the computing infrastructure and the technical expertise to carry out the experiments. Therefore, there have been no other costs to the project, other than what the BSC has used in terms of resources, which is unavailable to me as an employee. However, I have been able to make some approximate estimations of what would've been the total costs of my investigation, which are present in table 15. These estimations have been made after a literature review, in combination with the AWS prices [21][22][23][24].

	Compute Time	Compute cost (128 CPU)	Worker time	Worker cost	Infrastructure cost
1 Year / Personal	400	1,5€	960 hr.	7,5 €/hr.	4000€ / yr.

Table 12 – Yearly personals costs working at the BSC.

According to the table, in one year I've consumed approximately 11800€ worth of resources, the total cost of this investigation. It must also be mentioned that the BSC is a research center funded by Univesitat Politècnica de Catalunya, the Spanish government, and the Catalan Government.

In terms of the economic feasibility of the study, it would be very impractical to reproduce it without the infrastructure of a supercomputer and the technical expertise to perform experiments on such a platform. However, provided all of that, this study could be reproduced using the code present in the GitHub¹ repository.