National College of Ireland
Cloud Competency Centre

H. González–Vélez
Cloud Competency Centre
National College of Ireland
Mayor Street, IFSC, Dublin 1, IRELAND
www.ncirl.ie/cloud
horacio@ncirl.ie

Assessment Guide

Cloud Architectures (CLAR) : MSc Cloud : H. González–Vélez
Version: 3rd February 2023

National College of Ireland© 2022-23

Contents

Administrative

This module assessment guide has been designed to facilitate the understanding of the scope for the project within the *Cloud Architectures (CLAR)* module as part of the MSc Cloud at the School of Computing of the National College of Ireland.
The learning outcomes of this module are:

LO1 Critically compare and contrast distinct parallel and distributed architectures in terms of their functional and non-functional characteristics and associated enabling technologies.

LO2 Demonstrate in-depth knowledge of different types of computing systems for data storing, staging, and processing.

LO3 Evaluate and assess virtualisation and software environments for cloud computing.

LO4 Construct and present a business case for a complex, dynamic high performance computing solution for clouds.

LO5 Apply data governance and ethical frameworks to complex computational problems and recommend cloud-based solutions.

Representing 50% of the final grade awarded, the deadline for submission of this project is:

- **DEADLINE:** 31^{st} March 2023 9am.

> N.B.
>
> This project assesses LO3, LO4, and LO5.

Should any student miss the assessment with a valid reason, (s)he can now apply for an application for coursework Extension/Re-run Form online, via NCI360. PCF forms are no longer in use and will not be accepted by the School of Computing office.
Both assignments have to be electronically submitted using the dropbox provided in the module page in Moodle.
Please:

- ensure that your name in full (as per NCI official documents) and student number are clearly visible on the front page of the written reports; and
- name your files starting with the first letter of your given name followed by the first three letters of your surname, your student id, a dash, and the word "report". No spaces or any other alphanumeric characters should be included in the filename. That is to say, when "Ciara Byrne" with student id 22123456 submits her project report, she ought to name the file:
  `cbyr22123456-report.pdf`.

> N.B.
>
> All reports will be electronically screened for evidence of academic misconduct (plagiarism and collusion).
> The mark for this project represents 50% of the overall mark for this module.

National College of Ireland© 2022-23

Project (50% of Final Mark)

(a) Description

A major European research facility is exploring a move to a petascale architecture [1, 2]–as part of their road to exascale [3, 4]–for their key High Performance Computing (HPC) systems. They have asked you to help them analyse the technical and cost implications of adopting such HPC architecture. The facility managers want to know what the economic benefits of utility computing for HPC are, that is to say, the company may rent/lease their computational capacity (cycles) to different departments within their premises and, potentially, selected partners or other research consortia.

You must prepare a case study to quantify the cost benefits by comparing the economic implications of the proposed petascale architecture. The case study ought to be based on a given HPC challenge i.e. "application domain". For the purposes of this project, the following seven HPC application domains have been selected:

Modulo Result	Applications
0	Global climate modelling
1	Human proteome mapping
2	Drug design
3	New materials modelling
4	Earthquake ground motion
5	Groundwater transport and remediation
6	Land cover dynamics

ege of Ireland© 2022-23

The actual application domain you will focus on will be defined by a simple (hashing) function based on your student id. To calculate it, add all digits of your student number and then apply the modulo operator over the seven domain (%7) e.g. Ciara Byrne with student id 22123456 should work on the "Earthquake ground motion" domain [Explanation: $(2 + 2 + 1 + 2 + 3 + 4 + 5 + 6)\%7 = 4$].

N.B.

- You will **NOT** need to contact any research facility directly. Your assumptions should be based on publicly available information only.
- For the purposes of the case study, it is assumed that the research facility will utilise on-premises HPC systems only.
- The actual attributes of the HPC system must be set by you as a consultant

(b) Assumptions

You are asked to design a multi-variable utility HPC architecture that realistically fulfils the different system requirements such as:

Functional: Computation i.e. Peak and Sustained floating-point operations per second (Flops) to justify a petascale architecture, memory, and data storage needs, number of users, requests,...

Non-Functional Expected SLA, QoS, scalability, and service/maintenance costs, and any other relevant details.

You will compare a traditional HPC vs HPC petascale computing and the analysis must include a basic Total Cost of Ownership (TCO) analysis. You may use a second model or methodology to validate your result.

- Use available HPC architectures including GPU and other offloading devices. Make sure that hidden costs are accounted for in your case study and analyse a 5-year horizon with all costs and any possible expansions.

- Describe any and all possible acceptance criteria other than direct costs: specific benchmark results (e.g. Linpack, [5],HPCG [6],SPEChpc21 [7],...), response times, etc. You must conduct some independent research and include any relevant bibliography.
- State in your conclusions what the best option is. Illustrate an understanding of your clients business needs and how the adoption of a HPC petascale solution would benefit the research facility.

(c) Deliverable

1. A 8-page report formatted using the IEEE double-column template. The report has to include all figures and any references to existing work. Please refer to the following link for the formatting requirements and LaTeX/Word templates: `https://www.ieee.org/conferences/publishing/templates.html`.

(d) Structure of the Report

All relevant findings should be compiled into an accompanying report. It should be structured as follows:

Headline: Title of the report, your name, student number, module, programme, and date

Abstract: a roughly 200-word executive summary of the project and the key results

Introduction: set the scene of the project including the objectives of the project and a generic overview of the application domain and its needs. Conceivable here, you would also discuss how the domain has used HPC historically.

Background: introduce the HPC application domain, core research, and any other elements relevant to your report with special emphasis on the computation, memory and storage specific requirements. That is to say, you are expected to map the application domain intrinsic characteristics to your HPC proposal.

Analysis of the Research Facilitys on-premises HPC proposal: where details of the current system are limited, it is acceptable to make reasonable assumptions based on typical use cases in that domain, but you must provide your rationale for such assumptions. Even if you have details of the system, clearly specify your general assumptions on systems, users, and applications. Try to be as specific as possible by concentrating on assumed HPC system(s) and application(s).

HPC Petascale functional/non-functional details: essentially, how have you addressed any user requirements in your project?. Ethical and data governance consideration on your architecture and their intended use should also be integrated here. Describe how you have built your application workflow, what components and/or forms of HPC you have used and why. Focus on **Scalability & Redundancy Considerations** and provide evidence on how the new number of nodes/processors/cores will provide enable speedup and redundancy. Your considerations should be linked to Amdahl's Law, Gustafson's Law, MTBF/MTTF, and related topics.

Multi-Variable Cost: By using the Dell Enterprise Infrastructure Planning Tool `https://www.dell.com/calc` and any other relevant tools, compute the Total Cost of Ownership for the given variables (configurations, users, etc.). Finally discuss any interesting aspects of your results.

Conclusions and future work: what (in general) did you learn and find out? If you were to do the project again, what would you do differently? If you had more time (e.g. in your final project) what would you do next to extend your work?

Bibliography: a complete list of academic works and/or online materials used in the project. References should be included as in-text citations according to the IEEE citation style (similar to this guide). To find academic works and citation style guidelines, please refer to the NCI Library's Cloud Computing Subject Guide: `http://libguides.ncirl.ie/cloudcomputing`

(e) Marking Grid

National College of Ireland© 2022-23

MARKING GRID– Cloud Architectures (CLAR) , Assessment Guide, H. González–Vélez Due date: 31st March 2023 .

ASSESSMENT CRITERIA	*EXCELLENT / VERY GOOD*	*GOOD*	*SATISFACTORY*	*THRESHOLD*	*FAIL*
Project Objectives, Architectural Considerations, and Ethics & Data Governance: (40% weight)	Challenging project objectives are well presented, met, and thoroughly discussed. Architectural considerations have been well prepared and explored. Assumptions have a high degree of complexity. Comprehensive consideration of ethical and data governance issues.	Clear project objectives are well presented, met, and discussed. Architectural considerations have been prepared and explored. Assumptions have some degree of complexity. Clear consideration of ethical and data governance issues.	Reasonable project objectives are somehow presented, partially met and discussed. Architectural considerations have been prepared and mostly explored. Assumptions have limited degree of complexity. Reasonable consideration of ethical and data governance issues.	There are some objectives, which are at least partially met. Architectural considerations have been prepared but somewhat trivial. Assumptions have very limited degree of complexity. Scant consideration of ethical and data governance issues.	Cannot discern project objectives and if project objectives were met. Architectural considerations are not evident. Assumptions are trivial. No consideration of ethical and data governance issues.
Scalability and Cost model and analysis. (40% weight)	Excellent/very good application of HPC architecture design principles in terms of appropriate methodology and methods for generating and analysing cost data. Excellent/very good scalability considerations.	Good application of HPC architecture design principles in terms of appropriate methodology and methods for generating and analysing cost data. Good scalability considerations.	Adequate application of HPC architecture design principles in terms of appropriate methodology and methods for generating and analysing cost data. Borderline scalability considerations.	Weak application of HPC architecture design principles and limited evidence of understanding of appropriate methodology; and methods for generating and analysing cost data. Weak scalability considerations.	Poor application of HPC architecture design principles and very limited evidence of understanding of appropriate methodology and methods for generating and analysing cost data. Poor scalability considerations.
Identified impact/outcomes, structure, proposal abstract, and referencing. (20% weight)	Excellent/very good consideration of potential research impact/outcomes. Excellent/very good abstract and structure. All referencing consistent and appropriate	Good consideration of potential research impact/outcomes. Good abstract and structure. Most referencing consistent and appropriate.	Adequate consideration of potential research impact/outcomes. Adequate abstract and structure. Adequate consistent and appropriate referencing.	Limited/weak consideration of potential research impact/outcomes. Weak abstract and structure. Frequent inconsistent and/or inappropriate referencing.	Very limited and poor consideration of potential research impact/outcomes. Poor abstract and structure. Very frequent inconsistent and/or inappropriate referencing.
	70-100	*60-69*	*50-59*	*40-49*	*<40*

THE FINAL MARK MUST BE 40% OR ABOVE TO ACHIEVE A PASS

National College of Ireland© 2022-23

References

[1] J. J. Dongarra and D. W. Walker, "The quest for petascale computing," *Computing in Science Engineering*, vol. 3, no. 3, pp. 32–39, 2001. [Online]. Available: https://doi.org/10.1109/5992.919263

[2] G. Bell and J. Gray, "What's next in high-performance computing?" *Communications of the ACM*, vol. 45, no. 2, p. 9195, Feb. 2002. [Online]. Available: https://doi.org/10.1145/503124.503129

[3] D. A. Reed and J. Dongarra, "Exascale computing and big data," *Communications of the ACM*, vol. 58, no. 7, p. 5668, Jun. 2015. [Online]. Available: https://doi.org/10.1145/2699414

[4] D. Schneider, "The exascale era is upon us: The frontier supercomputer may be the first to reach 1,000,000,000,000,000,000 operations per second," *IEEE Spectrum*, vol. 59, no. 1, pp. 34–35, 2022. [Online]. Available: https://doi.org/10.1109/MSPEC.2022.9676353

[5] J. J. Dongarra, P. Luszczek, and A. Petitet, "The LINPACK benchmark: past, present and future," *Concurrency and Computation: Practice and Experience*, vol. 15, no. 9, pp. 803–820, Aug. 2003. [Online]. Available: https://doi.org/10.1002/cpe.728

[6] V. Marjanović, J. Gracia, and C. W. Glass, "Performance modeling of the HPCG benchmark," in *High Performance Computing Systems. Performance Modeling, Benchmarking, and Simulation*, ser. LNCS, vol. 8966. Cham: Springer International Publishing, 2015, pp. 172–192. [Online]. Available: https://doi.org/10.1007/978-3-319-17248-4_9

[7] J. Li, A. Bobyr, S. Boehm *et al.*, "SPEChpc 2021 benchmark suites for modern HPC systems," in *Companion of the 2022 ACM/SPEC International Conference on Performance Engineering*, ser. ICPE '22. Bejing, China: Association for Computing Machinery, 2022, pp. 15–16. [Online]. Available: https://doi.org/10.1145/3491204.3527498

National College of Ireland© 2022-23