THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Institut National Polytechnique de Toulouse (INP Toulouse)*

Présentée et soutenue le *31/01/2023* par :
Igor FONTANA DE NARDIN

Ordonnancement en ligne de tâches IT et engagement de sources dans un centre de calcul alimenté par des énergies renouvelables

JURY

PREMIER MEMBRE	Professeur d'Université	Rapporteur
SECOND MEMBRE	Professeur d'Université	Rapporteur
TROISIÈME MEMBRE	Professeur d'Université	Examinateur
QUATRIÈME MEMBRE	Professeur d'Université	Examinateur
CINQUIÈME MEMBRE	Professeur d'Université	Examinateur

École doctorale et spécialité :
> *MITT : Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse*

Unité de Recherche :
> *Laplace (UMR 5213) et IRIT (UMR 5505)*

Directeur(s) de Thèse :
> *Patricia STOLF* et *Stéphane CAUX*

Rapporteurs :
> *Premier RAPPORTEUR* et *Second RAPPORTEUR*

Abstract

Résumé

Acknowledgments

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Context

1.2 Problem Statement

1.3 Research Goals

1.4 Summary of Contributions

1.5 Publications and Communication

1.6 Dissertation Outline

The remaining dissertation has the following organization:

Chapter 2 - Context and Related Work: This chapter presents the fundamentals to understand this dissertation. Considering the scope of the topic, the context consists of third parts. First, we explain the context of Datazero2 and the link between this dissertation and the project. The second part describes the IT elements of a data center, including the type of tasks/jobs, scheduling algorithms, and energy consumption. The last part defines the electrical elements, such as batteries, hydrogen, wind turbines, and solar panels. After presenting the context, we introduce a list of works that solve part of our problem, highlighting the existing gaps in the state-of-the-art;

Chapter 3 - Modelling, Data, and Simulation: In this chapter, we describe the model to deal with the several elements that compose

a renewable-only data center. Datazero2 creates a division between Offline and Online decisions. We present the model to deal with offline decisions using predicted power demand and production. Then, we demonstrate the output of Offline used by the Online. Finally, we define the Online model, which englobes the job scheduling and modifications in the Offline plan. After describing the model, we explain the source of the different data (e.g., workload, weather, servers) applied in the simulations. Finally, we present the simulation tools used in this work;

Chapter 4 - Introducing Power Compensations: This chapter describes the proposed optimization to react to power uncertainties. We created four heuristics to find the best place to compensate for battery changes, which aim to reduce the number of killed jobs and the distance between the battery level and the target level. The results presented are related to the publications [1] and [2];

Chapter 5 - Learning Power Compensations: This chapter presents the idea and the results of the introduction of Reinforcement Learning (RL) in the power compensation problem. We propose two RL algorithms (Q-Learning and Contextual Multi-Armed Bandit) to learn the best moment to compensate;

Chapter 6 - Adding Battery Awareness in EASY Backfilling: This chapter explains a heuristic to mix scheduling and power compensation decisions. This heuristic is based on the EASY Backfilling scheduling algorithm but considers the battery's State of Charge to make better decisions;

Chapter 7 - Middleware Integration: This chapter presents the tools, frameworks, and approaches applied to integrate the algorithms in the Datazero2 middleware. Also, we describe the work done to create a docker environment, allowing different actors to execute the middleware;

Chapter 8 - Conclusion and Perspectives: Finally, in this chapter, we summarize the contributions of this work, providing a discussion about future works.

Chapter 2

Context and Related Work

2.1 Datazero2

2.1.1 Offline Decision Modules

2.1.2 Online Decision Modules

2.2 IT Elements

2.2.1 Type of Jobs

2.2.2 Scheduling Algorithms

2.2.3 Energy consumption

2.3 Electrical Elements

2.3.1 Storages

2.3.2 Renewable Production

2.4 Literature Review

2.4.1 Discussion and Classification of the Literature

Chapter 3

Modelling, Data, and Simulation

3.1 Model

3.1.1 Offline Decision Modules

Power Decision Module

IT Decision Module

3.1.2 Offline Plan

3.1.3 Online Decision Modules

Job scheduling

Modifying Power Plan

Modifying IT Plan

3.2 Data

3.2.1 Workload Trace

3.2.2 Weather Trace

3.2.3 Platform Configuration

3.3 Simulation

3.3.1 Simulator

3.3.2 Metrics

3.4 Conclusion

Chapter 4

Introducing Power Compensations

Chapter 5

Learning Power Compensations

5.1 Introduction

5.2 Algorithms

5.2.1 Random

5.2.2 Q-Learning approach

5.2.3 Contextual Multi-Armed Bandit approach

5.3 Results Evaluation

5.4 Conclusion

Chapter 6

Adding Battery Awareness in EASY Backfilling

6.1 Introduction

6.2 Heuristic

6.2.1 Predictions

6.2.2 Job Scheduling

6.2.3 Power compensation

6.3 Conclusion

Chapter 7

Middleware Integration

7.1 Introduction

7.2 Frameworks

7.3 Containerization

7.4 Conclusion

Chapter 8

Conclusion and Perspectives

8.1 Conclusion

8.2 Perspectives

Bibliography

[1] I. F. de Nardin, P. Stolf, and S. Caux, "Mixing offline and online electrical decisions in data centers powered by renewable sources," in *IECON 2022–48th Annual Conference of the IEEE Industrial Electronics Society.* IEEE, 2022, pp. 1–6.

[2] ——, "Analyzing power decisions in data center powered by renewable sources," in *2022 IEEE 34th International Symposium on Computer Architecture and High Performance Computing (SBAC-PAD).* IEEE, 2022, pp. 305–314.