

ECOLE POLYTECHNIQUE DE L'UNIVERSITÉ FRANÇOIS RABELAIS DE TOURS

Département Informatique

64 avenue Jean Portalis

37200 Tours, France

Tél. +33 (0)2 47 36 14 14

polytech.univ-tours.fr

Projet Recherche & Développement

2018-2019

Mise en oeuvre des outils de parallélisme sur un programme donné



POLYTECH[®]
TOURS

Confidentiel

Entreprise
Polytech'Tours



Tuteur entreprise
Patrick MARTINEAU

Étudiant
Ayoub IDEL (DI5)

Tuteur académique
Patrick MARTINEAU

Liste des intervenants

Entreprise

Polytech'Tours
64 Avenue Jean Portalis,
37200 Tours
polytech.univ-tours.fr



Nom	Email	Qualité
Ayoub IDEL	ayoub.idelmahjoub@etu.univ-tours.fr	Étudiant DI5
Patrick MARTINEAU	patrick.martineau@univ-tours.com	Tuteur académique, Département Informatique
Patrick MARTINEAU	patrick.martineau@univ-tours.com	Tuteur entreprise



Avertissement

Ce document a été rédigé par Ayoub IDEL susnommé l'auteur.

L'entreprise Polytech'Tours est représentée par Patrick MARTINEAU susnommé le tuteur entreprise.

L'Ecole Polytechnique de l'Université François Rabelais de Tours est représentée par Patrick MARTINEAU susnommé le tuteur académique.

Par l'utilisation de ce modèle de document, l'ensemble des intervenants du projet acceptent les conditions définies ci-après.

L'auteur reconnaît assumer l'entière responsabilité du contenu du document ainsi que toutes suites judiciaires qui pourraient en découler du fait du non respect des lois ou des droits d'auteur.

L'auteur atteste que les propos du document sont sincères et assument l'entière responsabilité de la véracité des propos.

L'auteur atteste ne pas s'approprier le travail d'autrui et que le document ne contient aucun plagiat.

L'auteur atteste que le document ne contient aucun propos diffamatoire ou condamnable devant la loi.

L'auteur reconnaît qu'il ne peut diffuser ce document en partie ou en intégralité sous quelque forme que ce soit sans l'accord préalable du tuteur académique et de l'entreprise.

L'auteur autorise l'école polytechnique de l'université François Rabelais de Tours à diffuser tout ou partie de ce document, sous quelque forme que ce soit, y compris après transformation en citant la source. Cette diffusion devra se faire gracieusement et être accompagnée du présent avertissement.



Pour citer ce document

Ayoub IDEL, *Mise en oeuvre des outils de parallélisme sur un programme donné*, Projet Recherche & Développement, Ecole Polytechnique de l'Université François Rabelais de Tours, Tours, France, 2018-2019.

```
@mastersthesis{
  author={IDEL, Ayoub},
  title={Mise en oeuvre des outils de parallélisme sur un programme donné},
  type={Projet Recherche \& Développement},
  school={Ecole Polytechnique de l'Université François Rabelais de Tours},
  address={Tours, France},
  year={2018-2019}
}
```

Table des matières

Liste des intervenants	a
Avertissement	b
Pour citer ce document	c
Table des matières	i
I Les premiers trucs dont je vais parler	1
1 Introduction	2
1 Contexte	2
2 Objectifs	3
3 Hypothèses	3
4 Bases méthodologiques	4
2 Description générale	5
1 Environnement du projet	5
2 Caractéristiques des utilisateurs	5
3 Fonctionnalités du système	5
4 Structure générale du système	5
3 État de l’art	6
1	6
2	6
3	6

4	Analyse et conception	7
5	Mise en oeuvre	8
1	Introduction de l'algorithme proposé	8
2	Version initiale de l'algorithme proposé	8
3	Amélioration de l'algorithme proposé	8
4	Version améliorée de l'algorithme.....	8
6	Conclusion	9
1	Bilan du semestre 9	9
2	Planning du semestre 10	9
3	Bilan du semestre 10	9
4	Bilan sur la qualité	9
5	Bilan auto-critique.....	9
	Annexes	10
A	Planification	11
1	Aperçu de gestion de projet	11
2	Découpage des tâches.....	11
B	Spécifications fonctionnelles	12
C	Spécifications non fonctionnelles	13
D	Analyse avec les détails et compléments	14
E	Gestion de projet	15
1	Découpage en tâches	15
1.1	Tache 1 : blabla	15
1.2	Tache 2 : blablabla	15
2	Planning.....	15
2.1	Le planning de recherche	15
2.2	Le planning de développement	15
F	Documents d'installation	16
G	Document d'utilisation	17
H	Document de test	18
1	Introduction.....	19
1.1	Objectifs de test	19

2	Tests unitaires	19
2.1	Tests du fichier X.c.....	19
2.2	Tests du fichier Y.c	19
3	Tests fonctionnels	19
3.1	Test des fichiers de log/erreurs	19
3.2	Vérification de la solution.....	19
4	Tests de performances	19
4.1	Test du temps d'exécution	19
4.2	Test d'occupation de mémoire.....	19
5	Test.....	19
Comptes rendus hebdomadaires		21
Glossaire		23

Première partie

Les premiers trucs dont je vais parler

1

Introduction

Ce document constitue les spécifications de Projet de Recherche et de Développement « Parallélisme » .

Ce sujet de projet est proposé par Monsieur Patrick MARTINEAU qui représente le MOA (Maitrise d'ouvrage) au sein de l'École Polytechnique de l'Université de Tours. Le projet sera réalisé par Ayoub IDEL, élève ingénieur en 5ème année à Polytech Tours.

1 Contexte

Presque tous les calculs effectués au cours des quarante premières années de l'histoire de l'informatique pourraient être appelés séquentiel.

Une des caractéristiques du calcul séquentiel est qu'il utilise un seul processeur pour résoudre un problème (ici, le terme « problème » est utilisé au sens large, c'est-à-dire que l'on exécute une seule tâche). Ces processeurs étaient devenus de plus en plus rapides et moins cher au cours des trois premières décennies, en doublant leur vitesse tous les deux ou trois ans. Cependant, en raison de la limite imposée par la vitesse de la lumière, il semble extrêmement improbable que nous puissions construire des Ordinateurs ne contenant qu'un seul processeur et pouvant atteindre des performances nettement supérieures à 1 000 000 000 d'opérations par seconde (Gflops).

Toutefois, si l'on prend en compte l'appétit sans cesse croissant de la puissance informatique pour le calcul (c'est-à-dire pour résoudre plus rapidement des problèmes plus importants), il devient nécessaire de choisir une autre voie pour le calcul séquentiel. Le calcul parallèle semble être l'alternative la plus prometteuse si ce n'est la seule.

Le calcul parallèle se définit comme la pratique consistant à employer un nombre de processeurs associés, communiquant entre eux pour résoudre rapidement de gros problèmes. Il est rapidement devenu un domaine important de la science informatique. Au cours des cinq dernières années, le calcul parallèle a pris de l'ampleur et confirme que la plupart des recherches menées dans les domaines de la conception et de l'analyse d'algorithmes, de langages informatiques, d'applications informatiques et d'architectures informatiques se situent dans son contexte. De nouvelles machines parallèles à la nouvelle architecture sont construites chaque

an.

Nous allons exploiter le parallélisme dans ce projet afin d'avoir les meilleurs résultats et temps d'exécution possibles d'un algorithme.

Dans ce rapport, le premier chapitre décrit le contexte de la réalisation, y compris l'objectif du logiciel, l'hypothèse et la méthodologie utilisée. Le deuxième chapitre présente les descriptions générales du logiciel sur l'environnement, les utilisateurs ainsi que la structure générale du système, ensuite nous allons exposer en détail les fonctionnalités. Le troisième chapitre présente l'état de l'art relatif à ce projet. Le quatrième chapitre explique les analyses faites sur le problème choisi. Le chapitre 5 explique en détails l'algorithme proposé et des améliorations possibles. Quant au sixième chapitre, nous mettrons en place un bilan générale du projet.

La partie annexe traite les points suivants : planification, spécifications fonctionnelles et non-fonctionnelles détaillées, gestion de projet, documents d'installation et d'utilisation, enfin les différents tests à savoir : tests unitaires, fonctionnels et de performance.

2 Objectifs

sur un programme donné, je dois mettre en oeuvre les différentes technologies du parallélisme : multithread (OpenMP), multicore(CUDA), multiprocessus (MPI).

À l'aide de ces outils, je dois résoudre le problème de la séquence la plus courte en évaluant toutes les combinaisons possibles. La complexité est de $O(n!)$, le problème est NP-difficile.

Dans un premier temps, je dois mettre en place un algorithme séquentiel qui résout le problème, ensuite y affecter quelques modifications afin de le paralléliser.

Dans un deuxième temps, je dois adapter l'algorithme pour qu'il soit fonctionnel sur réseau. En effet, un processus "serveur" gère la file d'attente des séquences à évaluer et les distribues par lot. En ce qui concerne la distribution ou l'évaluation des lots, plusieurs stratégies peuvent être mise en place, comme :

- évaluation par demande
- mécanisme de reporting
- taille des lots
- LIFO ...

Lorsque les différents calculs sont attribués à différents clients, des fichiers de log et d'erreurs doivent exister pour la reprise de ces calculs en cas d'échec d'un client ou d'un serveur.

Dans un troisième temps, je dois programmer une API pour aider au déploiement parallèle des recherches exhaustives au quasi-exhaustives.

3 Hypothèses

1. Quant au projet, nous avons supposé que toutes les machines (Clients et serveur) sont disponibles et fonctionnelles.
2. Les différents programmes sont composés sur un ensemble de fichiers d'instance fixe

3. Le langage de programmation choisi est C, vu que les différentes bibliothèques et frameworks (**OpenMP**, **CUDA**, **MPI**, ...) fonctionnent correctement avec ce langage de programmation.

4 Bases méthodologiques

1. Les outils utilisés

- **Trello** : outil visuel et facile à utiliser pour gérer les différentes tâches du projet. Il est entièrement gratuit
- **Astah UML** : outil gratuit pour produire les différents diagrammes (use case diagram, sequentiel diagram, ...)
- **Gantt** : outil utilisé pour ordonner et visualiser les diverses tâches dans le temps. Il permet de représenter graphiquement l'avancement du projet.

2. Les étapes suivies pour la gestion de projet

- **Compréhension du projet** : Dans première phase d'étude, j'analyse le besoin du client et comprendre les différents points
- **Conception et planification** : Définir avec détails ce qui doit être fait, comment et avec quels moyens, et planifier dans le temps les étapes ainsi que les tâches.
- **Réalisation du projet** : Mettre en oeuvre concrète des éléments planifiés, programmation ...
- **Bilan final** : Rédiger un bilan de synthèse sur le déroulement du projet et les résultats attendus

Le modèle suivi pour la gestion de projet ainsi que la planification sera décrite dans la partie "**Planification**" dans l'Annexe.

2

Description générale

Dans ce chapitre, nous allons présenter l'environnement utilisé pour réaliser le projet, puis les caractéristiques des utilisateurs, ensuite les fonctionnalités du système et enfin la structure générale du système.

1 Environnement du projet

L'environnement utilis

2 Caractéristiques des utilisateurs

3 Fonctionnalités du système

4 Structure générale du système

3

État de l'art

1

2

3

4

Analyse et conception

5

Mise en oeuvre

- 1 Introduction de l'algorithme proposé
- 2 Version initiale de l'algorithme proposé
- 3 Amélioration de l'algorithme proposé
- 4 Version améliorée de l'algorithme

6

Conclusion

- 1 Bilan du semestre 9
- 2 Planning du semestre 10
- 3 Bilan du semestre 10
- 4 Bilan sur la qualité
- 5 Bilan auto-critique

Annexes

A

Planification

1 Aperçu de gestion de projet

2 Découpage des tâches

B

Spécifications fonctionnelles

C

Spécifications non fonctionnelles

D

Analyse avec les détails et compléments

E

Gestion de projet

1 Découpage en tâches

1.1 Tache 1 : blabla

1.2 Tache 2 : blablabla

2 Planning

2.1 Le planning de recherche

2.2 Le planning de développement

F

Documents d'installation



G

Document d'utilisation

H

Document de test

1 Introduction

1.1 Objectifs de test

2 Tests unitaires

2.1 Tests du fichier X.c

2.2 Tests du fichier Y.c

3 Tests fonctionnels

3.1 Test des fichiers de log/erreurs

3.2 Vérification de la solution

4 Tests de performances

4.1 Test du temps d'exécution

4.2 Test d'occupation de mémoire

5 Test

un essai **question**, un autre essai **module**.

voici et une deuxième fois .

Open Source



Comptes rendus hebdomadaires

Compte rendu n°1 du 17/09/2018

Je vous envoie ce mail pour vous tenir au courant de l'avancement de mon PRD. Je cherche toujours une solution pour le problème de la séquence la plus courte. Actuellement, je suis entrain d'analyser deux algorithmes du plus court chemin, à savoir : Johnson's algorithm et Floyd-Warshall algorithm.

Pour Mercredi prochain, je souhaite arriver à la séance en ayant une première version d'algorithme qui sera adaptable au problème.

Compte rendu n°2 du 24/09/2018

J'ai essayé de trouver une bonne approximation au problème. En effet, j'ai implémenté l'algorithme de Floyd-warshall avec une complexité de $O(n^3)$. L'algorithme va me permettre d'évaluer toutes les combinaisons possibles, puis, en déduire une matrice des distances des plus courts chemins entre toutes les paires de sommets.

Celle-ci sera exploitée afin de définir la séquence la plus courte en se basant sur des algorithmes d'approximation du TSP (voyageur de commerce).

Compte rendu n°3 du 01/10/2018

J'ai créé un répertoire GitHub pour que je puisse bien gérer mon projet et que vous ayez une meilleure visualisation de mon état d'avancement. Voici le lien Git : <https://github.com/DevPsksh/PR-D.git>. Quant au projet, j'ai quelques questions à vous poser. J'aimerais bien convenir d'un Rendez-vous. Je vous remercie d'avance.

Compte rendu n°4 du 08/10/2018

Comme nous avons pu échanger la semaine dernière, j'ai fait des recherches sur les différentes technologies en détails afin d'obtenir les meilleurs résultats en performance. J'ai essayé d'analyser les notions de parallélisme afin de trouver un compromis entre équilibrage des calculs intensifs et minimisation des communications entre différents processus (granularité).

Quant au problème de la séquence la plus courte, je suis parti sur la résolution par brute force, c'est-à-dire mettre en place un algorithme qui génère toutes les permutations possibles, déduit le coût total de chaque permutation, puis renvoie la séquence la plus courte. La complexité est $O(n!)$.

Je mettrais en place les différentes technologies de parallélisme pour avoir le meilleur temps d'exécution possible.



Glossaire

module Un module est une entité regroupant un groupe. Par exemple, la création/édition/exécution d'un QCM est un module différent de celui accordeur. 19

Open Source Libre de droit, dont les codes sources sont disponibles et modifiable librementhttps://en.wikipedia.org/wiki/Open_Source

question Une question contient une sortie (texte, image, son, ...) qui demande à l'utilisateur de fournir une entrée (texte, image, son, ...) pour traiter cette entrée comme étant la réponse à la demande faite en sortie. 19

Projet Recherche & Développement

Mise en oeuvre des outils de parallélisme sur un programme donné

Ayoub IDEL

Encadrement : Patrick MARTINEAU



Objectifs

Lorem ipsum dolor sit amet:

- Vestibulum tincidunt ipsum
- In dictum elit condimentum.
- Suspendisse ornare orci.

Aliquam porta mi ut justo pellentesque porta.



Mise en œuvre

Mauris consectetur, et auctor mi fermentum. Etiam venenatis augue neque, ac ullamcorper sodales sit amet.

1	2	3	4
---	---	---	---

Fusce aliquet sed sapien tempor facilisis. Nunc mi urna, cursus a ullamcorper quis, laoreet id felis.



Résultats attendus

- Quisque dictum diam in feugiat
- Proin quam nec, gravida mi
- Suspendisse tristique vehicula magna
- Nulla sagittis diam congue sagittis
- Donec auctor a non nisi
- Vestibulum quis rhoncus risus



Mise en oeuvre des outils de parallélisme sur un programme donné

Ayoub IDEL

Encadrement : Patrick MARTINEAU



En collaboration avec Polytech'Tours

Objectifs

- Lorem ipsum dolor sit amet:
- Vestibulum tincidunt ipsum
 - In dictum elit condimentum.
 - Suspendisse ornare orci.

Aliquam porta mi ut justo pellentesque porta.

Mise en œuvre

Mauris consectetur, et auctor mi fermentum. Etiam venenatis augue neque, ac ullamcorper sodales sit amet.

1	2	3	4
---	---	---	---

Fusce aliquet sed sapien tempor facilisis. Nunc mi urna, cursus a ullamcorper quis, laoreet id felis.

Résultats attendus

- Quisque **dictum** diam in feugiat
- Proin quam nec, gravida mi
- **Suspendisse truci** **bidul** vehicula magna
- Nulla sagittis diam congue sagittis
- Donec auctor a non nisi
- Vestibulum quis rhoncus risus



POLYTECH®
TOURS

dictum elit condimentum



POLYTECH®
TOURS



POLYTECH®
TOURS

*auctor a non nisi
auctor a non nisi*



POLYTECH®
TOURS

Informatique

Mise en oeuvre des outils de parallélisme sur un programme donné

Confidentiel

Résumé

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Mots-clés

mot, clé, deux mots

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Keywords

word, key, two words, fourth word

Entreprise

Polytech'Tours



Tuteur entreprise

Patrick MARTINEAU

Étudiant

Ayoub IDEL (DI5)

Tuteur académique

Patrick MARTINEAU