



INSTITUT UNIVERSITAIRE DE TECHNOLOGIE LYON 1
Département Informatique
43 Boulevard du 11 Novembre 1918 - 69622 Villeurbanne

RAPPORT DE STAGE - 2^{EME} ANNÉE DUT INFORMATIQUE

Simuler l'empreinte environnementale des centres de données

9 AVRIL - 15 JUIN 2018



LABORATOIRE DE L'INFORMATIQUE DU PARALLÉLISME
École Normale Supérieure
46 Allée d'Italie - 69364 Lyon

Maître de stage
Laurent LEFEVRE

Étudiant
Bastien MARSAUD

Responsable pédagogique
Hamamache KHEDDOUCI

Fiche technique

Le Laboratoire de l'Informatique du Parallélisme

Le Laboratoire de l'Informatique du Parallélisme est un laboratoire de recherche situé sur le site Monod de l'École Normale Supérieure de Lyon. Il regroupe 57 membres permanents, 20 membres temporaires et entre 40 et 50 doctorants autour de sujets très larges liés à l'informatique.

Le sujet du stage

Le sujet du stage est de concevoir un simulateur d'empreinte environnementale des centres de données. Ce stage à lieu dans le cadre d'un projet avec l'Institut d'aménagement et d'urbanisme de la région Île-de-France et l'École d'architecture de la ville et des territoires. Ainsi il permettrait à terme d'aider les architectes dans la construction des centres de données et d'aider les urbanistes dans leur intégration sur le territoire.

L'une des perspective de ce projet serait d'aider à la construction de nouveaux centre de données en Île-de-France afin de répondre aux besoins massifs en traitement de données que nécessiterons les Jeux Olympiques 2024.

L'environnement du stage

Dans le cadre de ce stage je suis intégré au Laboratoire de l'Informatique du Parallélisme dans l'équipe AVALON. Tous les membres de l'équipe sont soit des chercheurs, soit des ingénieurs de recherche, soit des doctorant en informatique. Un camarade de ma promotion **Lucas Besnard** est lui aussi en stage dans l'équipe AVALON mais nous ne travaillons pas sur le même sujet.

Je travaille seul sur le projet, **Laurent Lefèvre**, mon maître de stage est bien entendu présent pour me donner les consignes, m'aiguiller et m'épauler dans ma réflexion mais ne participe pas au développement.

L'environnement de travail

La laboratoire possède des ordinateurs portables, mais comme ils ne sont pas très performants on m'a conseiller d'utiliser mon ordinateur personnel. La laboratoire m'a cependant fourni un deuxième écran.

Comme je n'avais aucune contraintes aux niveaux des technologies j'ai décidé d'utiliser celles avec lesquelles j'étais le plus à l'aise. Le projet en lui-même est développé en **JAVA** en utilisant la technologie **JavaFX** pour l'interface graphique, j'utilise **Maven** pour la gestion des librairies ainsi qu'**Eclipse** en tant qu'**IDE**. Pour versionner le code source j'utilise le protocole **GIT** couplé à un **repository** privé sur **GitHub**. Il était en effet compliqué de me créer un **repository** sur la plateforme interne à cause de formalités administratives.

Méthode de travail

Pour le bon déroulement du projet il était indispensable de faire une recherche bibliographique conséquente avant de commencer la phase de développement afin d'assimiler un certains nombre de notions spécifiques.

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué de près ou de loin à la réussite de mon stage.

Je remercie tout d'abord **Mme. Jocelyne Debouté**, enseignante en expression communication et responsable des stages à l'IUT Lyon 1, qui fait un travail remarquable dans la recherche et le partage d'offres de stage. Grâce à qui j'ai pu postuler à ce stage et qui m'a beaucoup aidé lors de la validation de ma convention.

Je remercie chaleureusement mon maître de stage **M. Laurent Lefèvre**, chercheur à l'INRIA, membre de l'équipe AVALON au Laboratoire de l'Informatique du Parallélisme pour sa confiance et l'indépendance qu'il m'a accordée, mais surtout pour ses conseils avisés et sa bonne humeur à toute épreuve.

Je remercie **M. Sylvain Maurin**, mon voisin de palier, chercheur à l'Institut des Sciences Cognitives, **M. Christophe Jaloux**, enseignant en mathématiques à l'IUT Lyon 1, **Mme Aude Joubert**, enseignante en mathématiques à l'IUT Lyon 1 et **M. Pierre-Antoine Champin**, enseignant-chercheur à l'IUT Lyon 1 pour m'avoir permis, à leurs manières, d'assister à la conférence EcoInfo *Que deviennent nos déchets électroniques ?* du 6 Avril dernier à Grenoble dans laquelle mon maître de stage, Laurent Lefèvre, est intervenu.

Je remercie **M. Issam Raïs**, avec qui je partageais mon bureau, doctorant sous la supervision de Laurent Lefèvre, pour sa bonne humeur entre deux lignes de thèse.

Je remercie **Mlle. Dorra Boughzala**, doctorante sous la supervision de Laurent Lefèvre pour avoir pris le temps de m'expliquer en détail le fonctionnement et les subtilités de la plateforme *Grid' 5000*.

Enfin, je remercie mes camarades de promotion **M. Simon Lecutiez**, **M. Valentin Gilles**, **M. Tom Befieux** et **M. Alex Pertuiset** pour ces deux années de DUT Informatique ainsi que pour leurs conseils avisés durant la rédaction et leur relecture de ce rapport de stage.

Sommaire

Fiche technique	1
Remerciements	2
Introduction	4
1 L'environnement du stage	5
1.1 Le laboratoire et ses relations institutionnelles	5
1.1.1 Présentation générale	5
1.1.2 Organisation générale	5
1.1.3 Des métiers variés	6
1.1.4 Une production conséquente	8
1.2 L'équipe Avalon en détails	9
1.3 Ma place au sein du laboratoire	9
2 La mission	10
2.1 Le contexte	10
2.2 Le enjeux	10
2.3 L'outils attendu	10
2.4 Comment parvenir au résultat ?	10
2.4.1 La recherche bibliographique	10
2.4.2 Le développement	10
3 Mon expérience de stage	11
Conclusion	12
Glossaire	13
Table des figures	14
Liste des tableaux	14
Sources	15
Annexes	16

Introduction

Du 9 Avril au 15 Juin 2018 j'ai effectué mon stage de fin de deuxième année de DUT Informatique à l'IUT Lyon 1 au Laboratoire de l'Informatique du Parallélisme (**LIP**). Le Laboratoire de l'Informatique du Parallélisme est un laboratoire de recherche en informatique situé à l'**ENS** Lyon et regroupant des chercheurs, des ingénieurs et des doctorants autour de problématiques comme l'arithmétique en informatique, les architectures distribuées, l'optimisation des réseaux et des ressources ou encore l'analyse de la compilation. Il a permis, depuis l'année 2000, de mettre au jour environ 2500 publications.

Durant ma recherche de stage, même si j'utilisais des plateformes de recherche d'emplois en ligne, cette offre m'est parvenu via notre enseignante responsable des stages qui en envoyait régulièrement par mail. Cette offre a tout particulièrement attiré mon attention pour plusieurs raisons. Tout d'abord, il s'agissait de la seule opportunité qui permettait de travailler dans un laboratoire de recherche, les autres offres étant majoritairement des entreprises privées. Le milieu de la recherche m'a toujours attiré et comme j'essaye de plus en plus de l'intégrer à mon projet professionnel, cette offre me semblait un bon moyen d'y parvenir. Ensuite, les deux sujets proposés par l'offre portaient sur des problématiques environnementales, étant très attiré par ces questions depuis mon plus jeune âge et ayant suivi mon cursus de lycéen dans un lycée agricole j'étais très enthousiaste à travailler dans ce domaine. Ainsi ce stage m'apparaît comme une bonne opportunité de découvrir le monde de la recherche et de travailler sur un projet complexe, d'utilité publique sur des problématiques qui m'intéressent.

Ma mission durant ce stage est de concevoir un simulateur informatique d'empreinte environnementale des centres de données afin d'aider des architectes et des urbanistes dans leur démarche d'intégration des centres de données sur notre territoire. Pour y parvenir je devrai tout d'abord m'inscrire dans une démarche de recherche bibliographique poussée avant de pouvoir espérer commencer le développement.

TODO : Modifier le plan si il a changé

Dans ce rapport de stage, je présenterai tout d'abord l'environnement du stage : le laboratoire et ses relations institutionnelles, l'équipe AVALON dont je fais partie ainsi que ma place au sein de cette organisation. Ensuite j'expliquerais en détail ma mission, son contexte, les enjeux qu'elle porte, son but ainsi que les étapes que j'ai suivi pour y parvenir. Enfin je terminerai par détailler mon expérience durant ce stage, ce que j'ai appris et découvert techniquement et humainement.

1 L'environnement du stage

L'environnement de travail d'un stage dans un laboratoire de recherche est sans doute très différent d'un stage plus classique dans une entreprise privée. En effet, un laboratoire de recherche s'inscrit dans une organisation complexe et possède de nombreuses relations avec d'autres organisations ou d'autres personnes.

Dans cette partie nous présenterons tout d'abord ce qu'est le Laboratoire de l'Informatique du Parallélisme ainsi que ces relations avec d'autres institutions, ensuite nous présenterons l'équipe AVALON dont je fais partie, et enfin nous terminerons par présenter la place que j'occupe au sein du laboratoire.

1.1 Le laboratoire et ses relations institutionnelles

Le Laboratoire de l'Informatique du Parallélisme (abrégé **LIP**) est un laboratoire de recherche en informatique situé principalement sur le site Monod de l'École Normale Supérieure de Lyon.

1.1.1 Présentation générale

Créé dans les années 80, le Laboratoire de l'Informatique du Parallélisme devient **Unité de Recherche Associé** (URA) de l'ENS Lyon en 1989, puis **Unité Mixte de Recherche** (UMR) en 1999 qui sera complété par l'Université Claude Bernard Lyon 1 en 2003. Le laboratoire est très tôt épaulé par l'Institut national de recherche en informatique et en automatique (**Inria**) qui est le principal acteur de la recherche en informatique et mathématique en France depuis 1967. Ainsi le laboratoire héberge plusieurs équipes-projets communes avec l'**Inria**. [4]

Le laboratoire compte 57 enseignants titulaires et chercheurs, entre 40 et 50 doctorants ainsi qu'une vingtaine de personnes sur des postes non permanents. L'équipe d'administration et l'équipe technique quant-à-elles sont épaulés par 12 ingénieurs.

Le **LIP** possède une bonne visibilité de ses équipes au niveau national et de plusieurs de ses membres au niveau international. Il occupe une place centrale dans le paysage de la recherche en informatique française pour quasiment l'ensemble de ses thématiques. La forte croissance du laboratoire lui a même obligé à s'étendre sur 2 autres sites : au sein de l'Institut Rhône-Alpin des Systèmes Complexes, et au sein de locaux appartenant à l'UCB à Gerland.

1.1.2 Organisation générale

Le laboratoire compose l'**Unité Mixte de Recherche** 5668 avec l'ENS Lyon (EnsL), l'Université Claude Bernard Lyon 1, l'**Inria** et le CNRS. Il est dirigé par **M. Patrick Baillot** qui est secondé par **M. Frédéric Vivien**. Le responsable en charge de l'appel à projets, de la valorisation de recherche et des relations internationales est **M. Eddy Caron** et le responsable en charge des thèses, de l'enseignement et des postes non permanents est **M. Damien Stehlé**.

L'équipe administrative et l'équipe en charge des moyens informatiques quant-à-elles sont composées de différentes personnes issues des institutions qui constituent l'**Unité Mixte de Recherche** 5668.

Le laboratoire héberge sept équipes de recherche dont cinq sont commune avec l'**Inria** : AriC, Avalon, CASH, DANTE, MC2, PLUME et ROMA chacune administrés par un chef d'équipe.

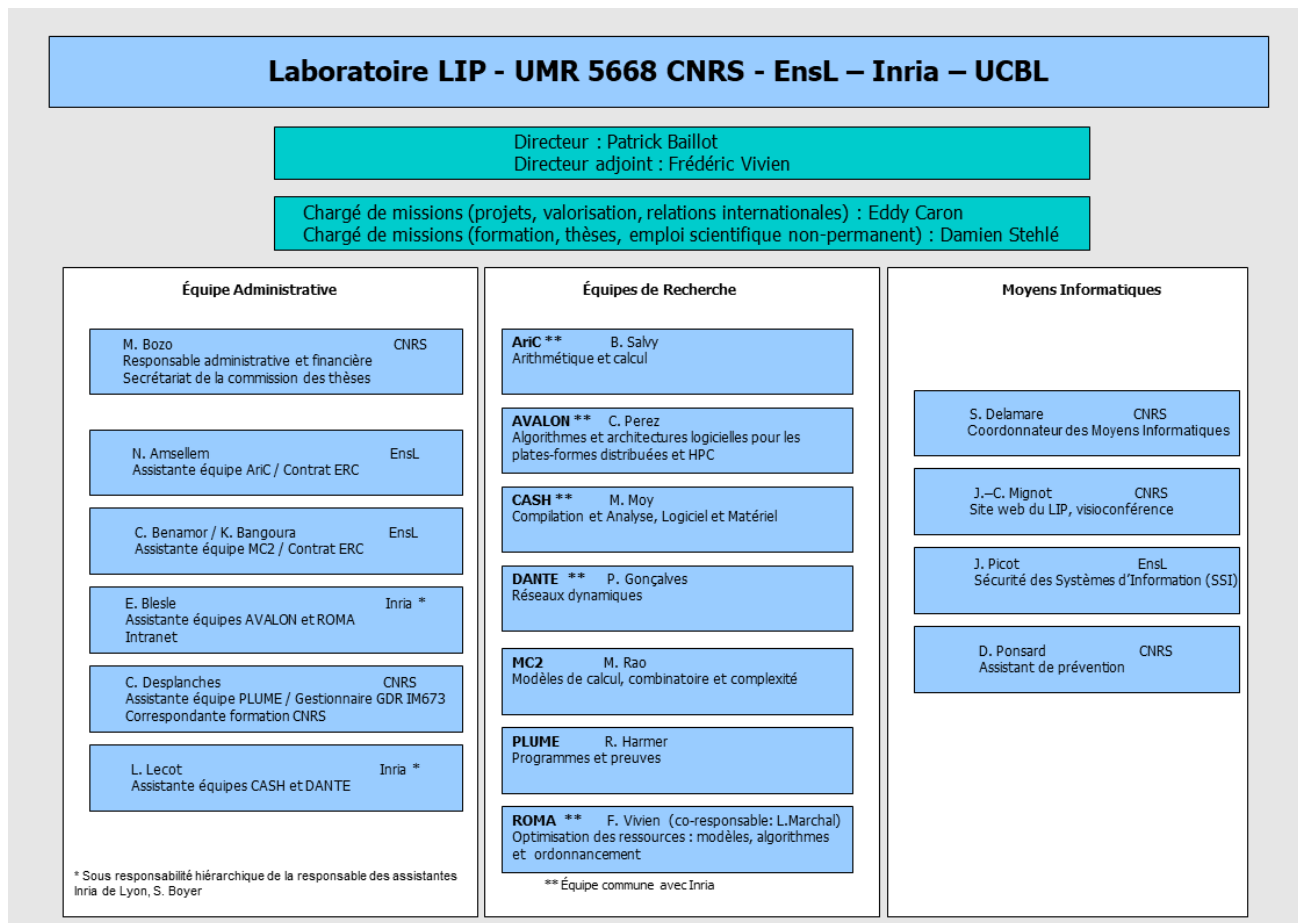


FIGURE 1 – Organigramme du LIP au 10 Avril 2018 [3]

1.1.3 Des métiers variés

Le projet du Laboratoire de l'Informatique du Parallélisme est de mettre en relation l'informatique fondamental et sa mise en œuvre pratique dans les institutions. Ainsi le laboratoire crée un lien fort entre informatique et d'autres sciences comme les mathématiques, les sciences du vivant ou, plus globalement, les sciences fondamentales.[1]

Les chercheurs du **LIP** possèdent un socle commun : l'algorithmie et l'utilisation efficace des ressources. Ils organisent leurs recherches autour de deux grands axes :

- La conception, l'utilisation et l'adéquation aux besoins des applications des futurs architectures de calcul (traitement de données, calcul fondamental) et de communication (réseaux)
- L'étude des modèles et des méthodes en informatique : complexité, algorithmie, développement logiciel et matériel, avancée technologiques etc.

Ainsi, le laboratoire ne compte pas dans ses rang uniquement des expert de l'informatique. D'autres profils bien différents sont mis en valeurs dans les 7 équipes-projets que nous allons présenter.

AriC - Arithmetic and Computing

AriC est une grande équipe projet commune avec l'**Inria** composée d'une vingtaine de membres. Elle a pour but d'améliorer le calcul, en terme de performance, d'efficacité et de fiabilité. Ses 3 principaux projets de recherche portent sur les sujets suivants :

- Les réseaux euclidiens : algorithme et cryptologie,
- Les méthodes d'approximation efficaces en calcul formel,
- Le calcul fiable à haute performance, avec virgule flottante et précision d'au plus une centaine de bits.

Au delà de la recherche cette équipe a une véritable vocation à la diffusion et à la vulgarisation de leurs travaux, ce qui passe par la multiplication des interventions dans les lycées ou autre institutions d'enseignement et la publication d'articles et d'ouvrages [5].

Avalon - Algorithms and Software Architectures for Distributed and HPC Platforms

L'objectif de l'équipe commune **Inria** Avalon est de concevoir des modèles de programmations, des systèmes et des algorithmes pour exécuter des applications sur des ressources tout en satisfaisant les contraintes des utilisateurs (i.e. coût, performances) et des administrateurs (i.e. maximisé l'usage des ressources, minimiser la consommation d'énergie). L'équipe se concentre en particulier sur le profilage et la modélisation d'applications gourmandes en énergie et en données, la gestion des données et l'ordonnancement des applications sur des architectures de supercalculateurs [6].

CASH - Compilation and Analyses for Software and Hardware

La vision de l'équipe commune **Inria** CASH est d'utiliser l'architecture **dataflow** pour le traitement des données par les supercalculateurs. Son but est d'utiliser les caractéristiques particuliers du matériel informatique afin de fournir des couples matériel-logiciel efficaces énergétiquement au développeur final. Pour ce faire elle travaille sur les axes d'études suivants :

- Développer l'architecture **dataflow**,
- Améliorer les algorithmes de compilation,
- Développer la compilation matériel, qui consiste à transformer un programme informatique en un circuit électronique physique,
- Émuler les **systèmes sur puce** pour faciliter leur optimisation [7].

DANTE - Dynamic Network

L'objectif principal de l'équipe commune **Inria** DANTE est de poser des bases solides à la caractérisation des réseaux dynamiques et des processus dynamiques se produisant sur des réseaux à grande échelle. Afin de développer des outils d'une pertinence pratique en situation réelle, elle fonde ses études méthodologiques sur des jeux de données réelles. Ses 3 grands thèmes de recherche sont :

- Le traitement du signal basé sur les graphes,
- La théorie des graphes dynamiques,
- Les **algorithmes distribués** pour les réseaux dynamiques [8].

MC2 - Models of computation, Complexity, Combinatorics

L'équipe MC2 a pour but de comprendre les possibilités et les limitations des algorithmes efficace. Pour ce faire elle crée et analyse des algorithmes jusqu'à leurs limites. Parmi les différents domaines des mathématiques au cœur de ses problématiques, l'équipe MC2 se concentre sur l'algèbre et l'**analyse combinatoire**. Ces deux domaines sont des sources de problèmes algorithmiques qui jouent un rôle clé dans la théorie de la **complexité** [9].

PLUME - Programs and Proof

Les recherches menées par l'équipe PLUME s'articulent autour de deux thèmes fortement imbriqués : les fondements logiques des langages de programmation et la théorie des systèmes informatiques. Elle met au centre de ses recherches la logique mathématique afin de trouver comment écrire des programmes sûrs ou comment vérifier formellement des systèmes informatiques complexes [10].

ROMA - Resource Optimization : Models, Algorithms and Scheduling L'équipe commune **Inria** ROMA vise à concevoir des modèles, des algorithmes et des stratégies d'ordonnement pour optimiser l'exécution d'applications scientifiques sur des supercalculateurs. Plus spécifiquement, ROMA vise à obtenir la "meilleure" performance possible du point de vue de l'utilisateur (i.e. le temps d'exécution de l'application) tout en utilisant les ressources aussi efficacement que possible [11].

Ainsi, le Laboratoire de l'Informatique du Parallélisme possède un impressionnant savoir faire dans de nombreux domaines de l'informatique et produit de nombreuses ressources pour des institutions publiques et privées, comme nous allons le voir.

1.1.4 Une production conséquente

Le Laboratoire de l'Informatique du Parallélisme est plutôt prolifique dans la quantité de production. Depuis 2000, 2502 publications ont été publiées et sont disponibles sur des plateformes en ligne comme les archives ouvertes HAL. En effet le laboratoire est dans une véritable démarche de production de connaissances, mais cela ne l'empêche pas de s'autofinancer grâce à des contrats industriels et des projets institutionnels régionaux, nationaux et internationaux [4]. 2 brevets ont même été déposés.

Le laboratoire encourage également les ambitions d'entrepreneuriat de ses membres avec la création de 5 start-ups dans le domaine du numérique depuis 2010 [1].

Enfin, les différentes équipes du laboratoire produisent également de nombreuses ressources logicielles open-source à destination des institutions, les industriels et même des particuliers que l'on peut retrouver sur les plateformes de partage de code source en ligne.

1.2 L'équipe Avalon en détails

1.3 Ma place au sein du laboratoire

2 La mission

TODO : Rédiger l'introduction de la partie

2.1 Le contexte

2.2 Le enjeux

2.3 L'outils attendu

2.4 Comment parvenir au résultat ?

TODO : Rédiger l'introduction de cette partie

2.4.1 La recherche bibliographique

2.4.2 Le développement

3 Mon expérience de stage

TODO : Rédiger l'introduction de la partie

Conclusion

Glossaire

algorithme distribué Algorithme s'exécutant, généralement en parallèle, sur plusieurs sites. 7

analyse combinatoire Domaine des mathématiques étudiant les configurations de collections finies d'objets ou d'ensembles et le dénombrement. 7

complexité En informatique, désigne la quantité de ressources nécessaire à l'exécution d'un algorithme. 7

dataflow Flux de données, indique que les données sont actives et traversent un programme de manière asynchrone contrairement à une architecture classique où elles attendent leur tour chargées en mémoire [12]. 7

Eclipse IDE multiplateforme et multilangage. 1

ENS École Normale Supérieure. 4, 5

GIT Protocole de gestion de version centralisé, permet de stocker du code source en conservant la chronologie de toutes les modifications. 1

GitHub Plateforme en ligne de gestion de version utilisant le protocole GIT. S'est imposé en tant que réseau social pour développeur. 1

IDE Environnement de développement intégré, ensemble d'outils dédiés au développement regroupés dans un même logiciel. 1

Inria Institut National de Recherche en Informatique et en Automatique. 5–7

JAVA Langage de programmation orienté objet et multiplateforme. 1

JavaFX Bibliothèque interne à JAVA gérant l'interface graphique utilisateur. 1

LIP Laboratoire de l'Informatique du Parallélisme. 4–6

Maven Outils de gestion de production. Facilite la gestion de bibliothèques. 1

repository Un dépôt centralisé et organisé de code source. 1

système sur puce Systeme embarqué sur une seule puce électronique. 7

Unité de Recherche Associé Structure de recherche qui relève d'un autre organisme que le CNRS dans laquelle le CNRS lui-même est impliqué [2]. 5

Unité Mixte de Recherche Structure de recherche placée sous la responsabilité conjointe du ministère de la recherche et du CNRS [2]. 5

Table des figures

1 Organigramme du LIP au 10 Avril 2018 [3] 6

Liste des tableaux

Sources

- [1] Université Claude Bernard Lyon 1. Unité mixte de recherche 5668 - laboratoire de l'informatique du parallélisme (lip). <https://www.univ-lyon1.fr/recherche/entites-de-recherche/laboratoire-de-l-informatique-du-parallelisme-lip--618125.kjsp>, 2018.
- [2] Direction de la recherche de Grenoble INP. Les labels des unités de recherche. pages 1–2, 19 Juillet 2012.
- [3] Laboratoire de l'Informatique du Para. Organigramme du lip. <http://www.ens-lyon.fr/LIP/index.php/organization-chart>, 10 Avril 2018.
- [4] Pierre FRAIGNIAUD. Évaluation du HCERES sur l'unité : Laboratoire de l'Informatique du Parallélisme. *Haut Conseil de la Recherche et de l'Enseignement Supérieur*, pages 5–6, 2015.
- [5] Équipe AriC. Aric : Arithmetic and computing. http://www.ens-lyon.fr/LIP/AriC/wp-content/uploads/2014/09/Slides_eval_Aeres_Aric.pdf, 2018.
- [6] Équipe Avalon. The avalon research team. <https://avalon.ens-lyon.fr/>, 2018.
- [7] Équipe CASH. Cash team. <http://www.ens-lyon.fr/LIP/CASH/>, 2018.
- [8] Équipe DANTE. Team presentation. <https://team.inria.fr/dante/>, 2018.
- [9] Équipe MC2. Mc2 : Modèles de calcul, complexité, combinatoire. <http://www.ens-lyon.fr/LIP/MC2/>, 2018.
- [10] Équipe PLUME. Preuves & langages : Un manège enchanté. <http://www.ens-lyon.fr/LIP/PLUME/>, 2018.
- [11] Équipe ROMA. Roma. <http://www.ens-lyon.fr/LIP/ROMA/>, 2018.
- [12] Communauté Wikipédia. Architecture dataflow. https://fr.wikipedia.org/wiki/Architecture_Dataflow, 29 Novembre 2016.

Annexes

A	Première annexe	17
A.1	Sous section de la première annexe	17
A.2	Deuxième Sous section de la première annexe	17
B	Deuxième annexe	17
B.1	Sous section de la deuxième annexe	17
B.2	Deuxième Sous section de la deuxième annexe	17

A Première annexe

A.1 Sous section de la première annexe

A.2 Deuxième Sous section de la première annexe

B Deuxième annexe

B.1 Sous section de la deuxième annexe

B.2 Deuxième Sous section de la deuxième annexe

Simuler l'empreinte environnementale des centres de données

BASTIEN MARSAUD
9 AVRIL - 15 JUIN 2018

Résumé

Étudiant en deuxième année de DUT informatique à l'IUT Lyon 1, j'ai effectué mon stage de fin de cursus au Laboratoire de l'Informatique du Parallélisme à l'ENS Lyon. Très intéressé par le milieu de la recherche et par les problématiques environnementale, j'ai initié la conception d'un simulateur d'empreinte environnementale des centres de données. Ce simulateur permettra, à terme, d'aider les urbanistes et les architectes dans la conception et l'intégration des centres de données sur le territoire. Il sera capable, à partir d'une multitude de données entrées par l'utilisateur, de générer automatiquement un rapport complet présentant les caractéristiques du centre de données ainsi que ses scores pour les différents indicateurs permettant d'évaluer son empreinte environnementale.

Dans ce rapport de stage vous découvrirez le Laboratoire de l'Informatique du Parallélisme en tant qu'organisation, les différentes étapes du projet : de la phase de découverte du sujet et de conception d'une bibliographie à la phase de développement, mais également mon ressenti sur cette expérience dans le monde de la recherche et ce que j'ai découvert techniquement et humainement.

Mots clefs : informatique, green IT, datacenter, recherche, environnement, énergie, simulation, économie, urbanisme

