



Rapport de Stage 2A

INTRODUCTION AUX TECHNIQUES DE TRAITEMENT DE DONNÉES MASSIVES

RAPHAËL MOULET

ANNÉE 2015-2016

STAGE DE 2E ANNÉE RÉALISÉ DANS L'ENTREPRISE EXCELERATE SYSTEMS

Maître de Stage : Pierre GEORGES
Encadrant Universitaire : Adrien COULET

Déclaration sur l'honneur de non-plagiat

Je soussigné(e),

Nom, prénom : Moulet, Raphaël

Élève-ingénieur(e) régulièrement inscrit(e) en 2^e année à TELECOM Nancy

Numéro de carte de l'étudiant(e) : 31415907

Année universitaire : 2015-2016

Auteur(e) du document, mémoire, rapport ou code informatique intitulé :

Introduction aux techniques de traitement de données massives

Par la présente, je déclare m'être informé(e) sur les différentes formes de plagiat existantes et sur les techniques et normes de citation et référence.

Je déclare en outre que le travail rendu est un travail original, issu de ma réflexion personnelle, et qu'il a été rédigé entièrement par mes soins. J'affirme n'avoir ni contrefait, ni falsifié, ni copié tout ou partie de l'œuvre d'autrui, en particulier texte ou code informatique, dans le but de me l'accaparer.

Je certifie donc que toutes formulations, idées, recherches, raisonnements, analyses, programmes, schémas ou autre créations, figurant dans le document et empruntés à un tiers, sont clairement signalés comme tels, selon les usages en vigueur.

Je suis conscient(e) que le fait de ne pas citer une source ou de ne pas la citer clairement et complètement est constitutif de plagiat, que le plagiat est considéré comme une faute grave au sein de l'Université, et qu'en cas de manquement aux règles en la matière, j'encourrais des poursuites non seulement devant la commission de discipline de l'établissement mais également devant les tribunaux de la République Française.

Fait à _____, le _____

Signature :

Rapport de Stage 2A

INTRODUCTION AUX TECHNIQUES DE TRAITEMENT DE DONNÉES MASSIVES

RAPHAËL MOULET

ANNÉE 2015-2016

STAGE DE 2E ANNÉE RÉALISÉ DANS L'ENTREPRISE EXCELERATE SYSTEMS

Raphaël Moulet
167 avenue du Général Leclerc
54500 Vandœuvre-lès-Nancy
raphael.moulet@telecomnancy.net

TELECOM Nancy
193 avenue Paul Muller
CS 90172 Villers-lès-Nancy
+33 (0)3 83 68 26 00
contact@telecomnancy.eu

Excelerate Systems
ENSEIRB-MATMECA – Bureau P278
Avenue des Facultés, 33400 Talence
05 56 07 23 33

Maître de Stage : Pierre GEORGES
Encadrant Universitaire : Adrien COULET

Excelerate
SYSTEMS

The logo for Excelerate Systems features the word "Excelerate" in a large, grey, sans-serif font. Below it, the word "SYSTEMS" is written in a smaller, all-caps, grey, sans-serif font. To the right of the text, there are three horizontal, blue, brush-stroke-like lines of varying lengths, creating a dynamic, energetic feel.

Résumé

J'ai effectué mon stage de deuxième année chez Excelerate Systems, PME internationale de services et solutions Big Data. Il a tout d'abord démarré par une phase de formation sur les technologies et outils du Big Data - Elastic en particulier. J'ai ensuite travaillé en équipe sur un projet de visualisation des données d'un site - en particulier la création d'une représentation cartographique. Pour cette tâche, j'ai collaboré avec deux autres stagiaires sur une base de code commandée à un tiers. Il a fallu définir les fonctionnalités recherchées, comprendre le code fourni, trouver des outils convenables et les intégrer. Au final, nous avons avancé dans le projet, mais son ampleur ne permettait pas sa complétion avant la fin de mon stage. Je l'ai donc conclu en préparant un transfert des compétences acquises durant la première partie pour mon remplaçant.

Mots-clés : Elastic, Big Data, représentation cartographique de données

Abstract

For my second year internship, I chose Excelerate Systems, an international SME provider of Big Data solutions. It consisted of a formation phase on the technologies and tools of Big Data - with a special emphasis on Elastic. I later worked on a data visualisation project - especially the cartographic representation of a website's data. For this task, I worked with two other interns on a code base commissioned to a third party. It required us to define the features we needed, understand the workings of the code, find the right tools and integrate them into the code. We made good progress, but the scope was too big for it to be completed before the end of my internship. I therefore concluded it by preparing a transfer of skills for my replacement.

Keywords: Elastic, Big Data, cartographic representation

Remerciements

Je tiens à remercier tous ceux qui ont contribué au bon déroulement de ce stage de deuxième année.

Je tiens tout d'abord à remercier l'équipe Bordelaise d'Excelerate Systems pour leur convivialité, et en particulier :

- le vice-président Adel Benyoussef pour son accueil favorable à ma demande,
- mon maître de stage Pierre Georges pour son expertise, son encadrement, son écoute et la formation apportée.

Je tiens également à remercier l'équipe de Télécom Nancy qui a rendu ce stage possible, en particulier mesdames Moufida Memour et Brigitte Thomassin pour la gestion administrative et M. Adrien Coulet pour son suivi.

Je tiens finalement à remercier la famille Moniot pour m'avoir accueilli chez eux et permis à mon stage de se dérouler dans les meilleures conditions.

Table des matières

Introduction	3
1 Présentation de l'entreprise	3
1.1 Excelerate Systems	3
1.2 Organisation	3
1.3 Offre	4
1.3.1 Produits	4
1.3.2 Collaborations et Clients	6
2 Tâches et réalisation	7
2.1 Acquisition de compétences	7
2.1.1 Outils généraux	7
2.1.2 Introduction au Big Data	7
2.2 Stack Elastic	8
2.2.1 Introduction	8
2.2.2 Réalisation	9
2.2.3 Présentation	12
2.3 Projet Reno'Aqt	12
2.3.1 Géocodage	13
2.3.2 Carte et Partitionnement des données	14
2.3.3 Perspectives et fin du stage	15
3 Bilan et Apports personnels	16
3.1 Immersion dans le Big Data	16
3.2 Travail dans une PME	16
Conclusion	18
Bibliographie/Webographie	19
Figures	20
Glossaire	21
Annexes	23

