



République algérienne démocratique et populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université des Sciences et de la Technologie Houari Boumediene U.S.T.H.B

Faculté d'Electronique et d'Informatique

Département d'Informatique

Mémoire

En vue de l'obtention du diplôme de Master

En Informatique

Option

Système Informatique Intelligent SII

Thème

***Intégration d'une application
d'indexation dans un environnement
Cloud Computing***

Réalisé par :

M^{lle} Radia ABDI

Encadré par :

M^{lle} Nour El Houda Boulkrinat

Devant le jury composé de :

Mme Nawel Gharbi	Maitre de conférences	Présidente
Mme Dalila Boughaci	Maitre de conférences	Examinatrice
Mr Fayçal Bouyakoub	Maitre de conférences	Examineur

N° Projet 09/2012

Résumé

Le concept *Cloud Computing* a connu un développement remarquable ces dernières années. Son principe est de déporter le traitement et le stockage des données des utilisateurs sur des infrastructures accessibles depuis internet de n'importe où et à n'importe quel moment. Le *Cloud Computing* compte plusieurs domaines d'utilisation y compris le domaine de la recherche d'informations. Ce dernier a pour objectif de fournir à un utilisateur un accès facile à l'information qu'il l'intéresse. Cependant l'un des problèmes majeurs qui y sont liés étant la masse d'information à indexer, qui a engendré en conséquence, l'accroissement de l'index (espace de stockage) et le temps d'indexation. A cet effet, nous avons essayé dans ce travail de régler le problème de stockage de l'index et cela par l'exploitation d'une nouvelle technologie qui est le *Cloud Computing*.

Dans ce travail, nous avons intégré dans un environnement *Cloud Computing* une application dédiée à l'indexation des informations et des données dans le cadre de la recherche d'information. Après comparaison entre les différentes plates-formes et open sources du *Cloud Computing*, nous avons détaillé et présenté les différentes étapes pour la réalisation de cette intégration, et nous avons étendu l'application de l'indexation existante par le développement des nouvelles classes à savoir : le Worker.

Mots clés: Cloud Computing, PaaS, IaaS, indexation, Amazon Web Services AWS.

Abstract

The Cloud Computing system has experienced a remarkable development in the last few years. Its main principle consists of uploading both the processing and storage of user's data to some infrastructures accessible via internet from anywhere and at any time. Cloud Computing has many fields of use; information retrieval as well. The latter aims at providing the user an easy access to the information they are interested in. However one of the major problems related to that is the huge amount of information to index, this has subsequently increased the index (storage space) and the time indexation. For that purpose, we have tried in this study to solve the index storage problem using a new technology; the Cloud Computing.

In this study, we have included in a cloud computing environment an application program dedicated to the information and data indexing in the context of information retrieval. After comparing the different platforms and open sources of cloud computing, we have given details and steps to follow to this integration. We have also extended the existing indexing program by developing new classes namely the Worker.

Keywords : Cloud Computing, PaaS, IaaS, indexing, Amazon Web Services AWS.

Table des matières

Introduction Générale	1
Chapitre I: Le Cloud Computing	
I.1/ Introduction.....	3
I.2/ Concepts de base autour du Cloud Computing.....	4
I.3/ Définitions	5
I.4/ Cloud vs Grilles	8
I.5/ Les différents modèles de service du cloud computing	9
I.5.1/ IaaS (Infrastructure as a Service).....	9
I.5.2/ PaaS (Platform as a Service).....	10
I.5.3/ SaaS (Software as a Service)	11
I.5.4/ XaaS (anything As a Service)	12
I.6/ Les différents modèles de déploiement d'un Cloud Computing.....	13
I.6.1/ Le cloud public	13
I.6.2/ Le cloud privé	13
I.6.3/ Le cloud Hybride [Abdoul, 11].....	13
I.6.4/ Le cloud Communautaire.....	14
I.7/ Les avantages et inconvénients du cloud computing	14
I.8/ Grid'5000.....	15
I.9/ Le Cloud Computing et l'open source	15
I.9.1/ Eucalyptus.....	16
I.9.1.1/ Définition	16
I.9.1.2/ Composants.....	17
I.9.2/ OpenNebula	18
I.9.2.1/ Présentation.....	18
I.9.2.2/ Composants.....	19
I.9.3/ OpenStack	20
I.9.3.1/ Définition	20
I.9.3.2/ Composants.....	20

I.9.4/ Comparaison entre les solutions open source	21
I.10/ Les acteurs de marché de cloud computing	23
I.10.1/ Amazon Web Services	23
I.10.2/ Microsoft Azure	24
I.10.3/ Google App Engine.....	24
I.10.4/ Google Apps	24
I.10.5/ Force.com	24
I.11/ Conclusion	25

Chapitre II: Concepts de base de l'indexation dans la recherche d'information

II.1/ Introduction	27
II.2/ Notions de base de la recherche d'information	27
II.2.1/ La requête	28
II.2.2/ Le document	28
II.2.3/ Le système de recherche d'information (SRI)	28
II.2.4/ Modèle de représentation.....	29
II.2.5/ La pertinence	29
II.3/ Processus de recherche d'information	29
II.4/ L'indexation	32
II.4.1/ Définition.....	32
II.4.2/ Les étapes de l'indexation	32
II.4.2.1/ Analyse lexicale.....	32
II.4.2.2/ Élimination des mots vides	32
II.4.2.3/ Lemmatisation	33
II.4.2.4/ La pondération	33
II.5/ Modes d'indexation	34
II.5.1/ Indexation manuelle	34
II.5.2/ Indexation automatique	35
II.5.3/ Indexation semi automatique.....	35
II.6/ Conclusion.....	35

Chapitre III: Conception

III.1/Introduction	36
--------------------------	----

III.2/ Définition des besoins	37
III.2.1/ L'application d'indexation ESIndex	37
III.2.2/ Choix d'une plate-forme de « <i>Cloud Computing</i> »	38
III.2.2.1/ Amazon Web services AWS	38
III.2.2.2/ Microsoft Windows Azure	39
III.2.2.3/ Google App Engine	40
III.2.2.4/ Synthèse	41
III.3/ Schéma global du système	41
III.4/ Les étapes d'intégration	43
III.4.1/ Préparation de l'environnement	43
III.4.1.1/ Cloud Computing AWS	43
III.4.1.2/ Fournisseur	46
III.4.2/ Intégration de l'application dans le cloud	48
III.5/ Les classes Java du projet <i>IndexCloud</i>	48
III.6/ Architecture du système	49
III.7/ Langage de modélisation objet UML	51
III.7.1/ Définition	51
III.7.2/ Diagramme de séquence.....	52
III.8/ Conclusion.....	566
Chapitre IV: Réalisation	
IV.1/ Introduction.....	577
IV.2/ Description des outils de travail.....	577
IV.2.1/ Eclipse.....	577
IV.2.2/ PHP	588
IV.2.3/ Importation des classes de l'application ESIndex.....	599
IV.3/ L'organigramme général de l'application.....	59
IV.4/ Description de l'application	60
IV.5/ Conclusion	64
Conclusion Générale	65

Bibliographie

Annexe