Contrôle d'accès aux entrepôts de données dans les Cloud Computing

Amina El Ouazzani, Nouria Harbi, Hassan Badir

Contrôle d'accès aux entrepôts de données dans les Cloud Computing

Amina El ouazzani*, Nouria Harbi**
Hassan Badir*

*Université Abdelmalek Essaadi ENSA LabTIC 1818 Tanger MAROC
{ a.elouazzani2000 , h.badir } @gmail.com,
http://www.ensat.ac.ma/
**Université Lumière Lyon 2 Laboratoire ERIC 69635 Lyon, Cedex FRANCE
nouria.harbi@univ-lyon2.fr

Résumé. Un entrepôt de données ou Data Warehouse (DW) peut être utilisé comme un mécanisme puissant pour découvrir les informations cruciales de l'entreprise. les entrepôts de données profitent de l'avènement de l'informatique dans les nuages (Cloud Computing) Afin de bénéficier d'un accès à distance à l'entrepôt de données et des services de haute qualité. Dans ce contexte, il est nécessaire de bien protéger ces entrepôts de données des différents risques et dangers qui sont nés avec l'informatique dans les nuages, ce qui garantirait la confidentialité des données qu'ils contiennent. Dans ce sens, plusieurs propositions ont été présentées, néanmoins, aucune n'est considérée comme un standard dans la gestion des accès aux entrepôts de données. Dans cet article, nous allons d'abord analyser le problème de sécurité au sein d'un entrepôt de données et ses enjeux, ensuite, nous présenterons une architecture d'un entrepôt sécurisée qui consiste à gérer les accès en fonction des profils des utilisateurs dans les Cloud Computing.

1 Introduction

Le système décisionnel représente une nécessité pour les grandes entreprises qui accumulent au cours du temps un grand volume de données qui ne peut pas être satisfait par les systèmes traditionnels de bases de données. Ces entrepôts de données constituent un support efficace pour avoir une vue clair aux décideurs sur les différentes activités de l'entreprise et les aider à prendre des bonnes décisions. Cependant, Les environnements Cloud ont récemment acquis une grande popularité car, ils présentent des nombreux avantages qui les rendent intéressantes pour les entreprises qui veulent offrir des services plus réactifs et améliorer l'efficience de leur informatique à moindres coûts. En effet, la mise en œuvre d'un entrepôt de données dans les nuages représente pour chaque entreprise une bonne solution vue son efficacité et sa rentabilité. Toutefois, comme chaque avancée technologique, l'informatique dans les nuages apporte aussi son lot de risques notamment en termes de sécurité qu'il faut prendre en compte pour pouvoir bénéficier de tous les avantages apportés par cette solution Favre et al. (2013).
Dans cet article nous commençons par la présentation des menaces de sécurité dans les Cloud

Computing et puis un état de l'art sur la sécurité des entrepôts de données dans les nuages. Dans la section 4 nous allons voir une présentation synthétique des principaux travaux. Ensuite, nous décrirons la solution proposée qui permet de résoudre quelques problèmes liés à la sécurité d'accès aux données, puis nous présenterons la concrétisation de cette solution à travers l'implémentation sur un système distribué, et nous terminerons par une conclusion et les perspectives.

2 La sécurité dans les cloud computing

Nos activités quotidiennes de traitement de données créent des quantités massives de données. Les Cloud Computing ont été émergé comme paradigme pour l'hébergement de ces données et avoir une prestation de service sur Internet ce qui permet de conserver des données. Au lieu de les entreposer sur la mémoire interne de son appareil, l'utilisateur les sauvegarde en ligne dans des serveurs de stockage. L'avantage pour l'internaute, c'est qu'il peut accéder à ses contenus numériques depuis n'importe quel ordinateur et qu'en cas de panne il peut les retrouver facilement. Parmi les avantages de l'entreposage de données sur le Cloud Computing :

- Réduction des coûts : Cloud Computing peut réduire la paperasserie, les coûts de transaction, les coûts de matériels et des ressources humaines.
- Adaptabilité : les services de cloud computing sont facturés selon le montant de l'utilisation. Par conséquent, vous ne payez que ce que vous utilisez vraiment et vous pouvez facilement mettre à jour votre service sans avoir à faire des ajouts coûteux en matériels ou en logiciels.
- services selon la taille de l'entreprise : des services de cloud computing sont disponibles dans les petites et moyennes tailles. Cela permettra de réduire le coût des licences de logiciels comme les logiciels de contrôle à distance ainsi que le coût du serveur.
- Plus facile à collaborer : avec le cloud computing, les utilisateurs peuvent accéder aux données de n'importe où, n'importe quand, ce qui permet de collaborer avec les employés distants.

Les menaces de sécurité dans les nuages sont à la fois externes et internes. Beaucoup de menaces extérieures sont similaires aux menaces rencontrées dans les grands centres de données. Les fournisseurs de cloud doivent éviter le vol ou les attaques, et protéger les sources des données des utilisateurs Zunnurhain et V. (2011).

La virtualisation est le principal mécanisme le plus adapté dans les nuages en raison de sa puissante défense et protection contre la plupart des attaques. Cependant, les ressources ne sont pas toutes virtualisées et les environnements de virtualisation ne sont pas tous sans risques, car les logiciels de virtualisation contiennent des bugs. Une mauvaise virtualisation réseau peut permettre l'accès à des parties sensibles de l'infrastructure du fournisseur et aux ressources des utilisateurs.

La responsabilité de problème de sécurité est partagée entre les utilisateurs, et les fournisseurs de cloud computing. Si la sécurité au niveau de l'application est de la responsabilité de l'utilisateur de nuage, le fournisseur est responsable de la sécurité physique et aussi la gestion des politiques de pare-feu externes dont l'objectif est de gérer le risque de perte des données. En outre, si une panne se produit, il n'est pas clair qui est la partie responsable. Une panne peut se produire pour diverses raisons :

- en raison de matériel.

– en raison des logiciels malveillants,
– en raison du mauvais fonctionnement des applications du client.

Quelle que soit la raison, la défaillance peut entraîner des conflits entre le fournisseur et les clients.

3 Etat de l'art

Construire des entrepôts de données trouve une nouvelle expression lorsque l'on se situe dans le nuage parce qu'il devrait tenir compte de nombreuses améliorations pour éviter les risques. Le point crucial à considérer porte sur le contrôle d'accès distribué qui corresponde au contenu distribué et aux applications distribuées. C'est ainsi que de nombreux travaux ont été consacrés à la recherche de solutions pour remédier à ce problème. Dans cette partie nous essayerons de présenter les principaux travaux de recherche concernant 3 axes qui sont la sécurité au niveau conceptuel de l'entrepôt, au niveau exploitation et la sécurité des données dans les nuages.

3.1 La sécurité dans la modélisation des entrepôts (niveau conceptuel)

Parmi les travaux qui ont été développés sur l'intégration de la sécurité dans la modélisation des entrepôts, on trouve :

Fernandez-Medina et al. (2006) ont développé un modèle de contrôle d'accès et d'audit(ACA) spécifique aux entrepôts de données, qui repose sur deux politiques de gestion des accès : MAC et RBAC. Il précise des règles de sécurité lors de la modélisation d'un modèle conceptuel, en intégrant la notion de «profil utilisateur », qui est constitué d'une table isolée contenant toutes les informations des utilisateurs (identité, niveau de classification : top secret, secret, confidentiel ou inconnu). Ce modèle reste un modèle purement théorique car aucune solution concernant son implémentation n'a encore été proposée.

Villarroel et al. (2006) ont défini une extension OCL «Object Constraint Language» en utilisant les mécanismes d'extension UML2.0 pour résoudre les problèmes de la confidentialité, cette extension spécifie les contraintes de sécurité des éléments lors de la modélisation conceptuelle des entrepôts de données.

Soler et al. (2008) ont utilisé des mécanismes d'extension fournis par le CWM (Common Warehouse Metamodel) pour étendre le package relationnel et construire un schéma en étoile, qui représente les règles de sécurité et de vérification capturées pendant la phase conceptuelle de l'entrepôt de données.

Soler et al. (2009) ont développé une méthodologie comprenant quatre phases : analyse, modélisation, implémentation et validation, qui couvrent les cinq niveaux d'abstraction qui sont : analyse des besoins, niveau conceptuel, niveau logique, niveau physique et l'examen post- développement, ce dernier étant une nouvelle discipline introduite par Lujan et Trujillo

(2004) . Cette méthodologie présente toutes les exigences de la sécurité tout au long du cycle de vie de l'entrepôt de données.

Rodriguez et al. (2011) présentent une extension d'UML 2.0 notamment du diagramme d'activité. Cette proposition, libellée comme BPSec (Business Security Process), permet de définir un ensemble d'exigences de sécurité (contrôle d'accès, détection des risques d'attaques, non-répudiation, intégrité, confidentialité et vérification de la sécurité), ce qui améliore l'expressivité des modèles des processus métiers, et permet de sécuriser un entrepôt de données lors de son développement en prenant en considération cette exigence.

3.2 L'intégration de la sécurité dans les entrepôts existants (niveau exploitation)

Online Analytical Processing (OLAP) est devenue de plus en plus une composante importante et répandue des systèmes d'aide à la décision. Le serveur OLAP est censé assurer des accès en fonction des habilitations de chaque utilisateur. Il peut refuser les accès aux données d'une mesure, d'une dimension, et/ou au delà d'un niveau dans une hiérarchie. Les droits d'accès peuvent être explicitement spécifiés sur les tables/colonnes des tables de l'entrepôt de données. Cependant, le serveur OLAP tout seul ne peut pas protéger l'accès aux données interdites. Des travaux ont été réalisés pour renforcer les droits d'accès/habilitations des utilisateurs, et pour interdire tout utilisateur malicieux d'inférer des données qui lui sont interdites à partir des données auxquelles il a accès.

Kirkgoze et al. (1997) ont défini un modèle sécurisé pour les entrepôts de données qui consiste à élaborer un cube personnalisé possédant ses propres dimensions et hiérarchies. Ce modèle repose sur la politique de gestion AMAC. Il s'agit d'une extension du modèle MAC qui permet de spécifier les tâches que l'utilisateur peut exécuter selon son rôle au sein de l'organisation. L'intérêt d'un modèle comme celui-ci, est la flexibilité de l'assignation des rôles aux différents cubes virtuels.

Priebe et Pernul (2000) ont exploré les problèmes de sécurité rencontrés dans le cadre du projet Goal, projet visant à étudier l'intégration d'un système informatique géographique. Au cours de ces recherches, ils ont développé une méthode propre au monde OLAP. Celle-ci reprend la méthodologie classique utilisée pour l'élaboration des bases de données (analyse des besoins, modèle conceptuel, logique et physique), tout en incorporant l'aspect multidimensionnel lors de la phase conceptuelle.

Priebe et Pernul (2001) poursuivent leurs recherches concernant la création des mécanismes de contrôle des accès afin d'assurer la confidentialité des données, et ils ont créé un mécanisme de contrôle d'accès sous forme d'un langage exprimant, au cours de la phase conceptuelle, les contraintes liées à la sécurité. Il s'agit d'un langage basé sur MDX «MulDimensionnale Xpression », celui-ci étant un langage de requête spécialisé dans l'interrogation et la manipulation des données multidimensionnelles. Il est comparable au langage SQL.

Triki et al. (2011) ont proposé une approche qui ne nécessite pas un traitement supplémentaire, après chaque phase d'alimentation de l'entrepôt de données. Elle est basée sur les réseaux Bayésiens. Pour protéger un entrepôt de données contre les inférences, ils utilisent un module de contrôle, qui vise à interdire à un utilisateur d'inférer des données protégées à partir des données qui lui sont accessibles en utilisant les fonctions d'agrégations Min et Max.

Eavis et Althamimi (2012) ont présenté un cadre d'authentification qui s'appuie sur une algèbre spécialement conçue pour OLAP. Il est orienté objet et utilise des règles de réécriture de requêtes afin d'assurer l'accès aux données cohérentes à travers tous les niveaux du modèle conceptuel. Le processus est essentiellement transparent pour l'utilisateur, une notification est fournie dans le cas où un sous-ensemble de la demande initiale est renvoyé. Le résultat final est une approche intuitive et puissante pour l'authentification de base de données qui est uniquement adaptée au domaine OLAP.

3.3 La sécurité des données dans les nuages

Zunnurhain et V. (2011) traitent les attaques sur les nuages que ce soit par redirection des messages échangés entre le navigateur Web et le serveur ou bien par l'utilisation des logiciels malveillants au cours de l'échange de métadonnées dans ce cas le service de cloud computing va souffrir de l'écoute, du blocage et de la saturation. En proposant des solutions théoriques à des problèmes spécifiques tels que la fiabilité des messages SOAP par la génération de la clé RSA, l'utilisation d'un système de fichiers FAT dans les machines virtuelles pour vérifier l'application que le client est en cours d'exécution et détecter les services malveillants d'une manière plus simple.

Karkouda et al. (2012) Ont proposé le partage de chaque donnée stockée dans l'entrepôt sur plusieurs fournisseurs des nuages à travers l'algorithme de secret sharing (Shamir 1979) pour sécuriser le stockage et l'exploitation d'un entrepôt de données dans les nuages, cette dernière est inspirée de l'idée proposée par (Danwei et Yanjun 2010). Cette façon de répartir les données permet de ne pas dépendre d'un seul fournisseur, ce qui minimise le risque de non disponibilité des données.

Jensen et al. (2012) ont énuméré les différentes techniques utilisées dans cloud computing pour sécuriser les accès et ils ont dégagé les lacunes de ces techniques pour mettre en œuvre leur solution qui est basée sur l'utilisation du protocole TLS et la cryptographie XML pour adapter le navigateur web qui présente des problèmes au niveau sécurité

Wang et al. (2012) Ont proposé une solution basée sur le jeton homomorphique pour l'exactitude du stockage et pour localiser les erreurs, il est capable de détecter la corruption des données lors du stockage, afin de garantir la localisation des données erronées, et ils ont utilisés le code « erasure correcting » afin de permettre la redondance et garantir la fiabilité des données.

Chen et He (2010) Ont proposé un algorithme du partage de la clé secrète de Shamir qui est un algorithme de cryptographie, l'algorithme de stockage de données en ligne de Abhishek (Parakh et Kak 2009) et la théorie du nombre, ce qui permet la restauration des données en divisant une donnée en plusieurs parties pour les stocker ultérieurement sur des serveurs choisis aléatoirement. Avec cet algorithme, les données sont prêtes à être transférées, stockées, traitées, en toute sécurité puisqu'elles sont cryptées, parmi les avantage de cette solution est la capacité de restaurer les données même si un ou plusieurs nœuds de stockage ne sont pas disponibles

4 Synthèse des travaux existants

Les travaux présentés sont cohérents et complémentaires. D'abord E Soler et al (2006) ont étendu le méta modèle CWM pour représenter correctement toutes les règles de sécurité et d'audit définies dans la modélisation conceptuelle des entrepôts de données. Dans la suite de leurs travaux, Soler et al. (2008) et Soler et al. (2009) se sont basés sur le modèle MDA pour définir des règles de passage formelles entre le modèle conceptuel de l'entrepôt de données et le modèle logique, en exploitant le query/view/transformations (QVT) proposé par le modèle MDA. Soler et al. (2009) ont fait usage des mécanismes d'extension fournis par le CWM pour étendre le paquet relationnel afin de construire un schéma en étoile, qui représente les règles de sécurité et de vérification capturées pendant la phase de modélisation conceptuelle de l'entrepôt. En général, il n'y a pas une norme pour l'échange et l'interopérabilité des métadonnées pour les entrepôts de données. Bien que la proposition du méta modèle CWM (Common Warehouse Metamodel) basé sur trois standards, à savoir UML, MOF et XML est largement acceptée comme la norme pour l'échange et l'interopérabilité des métadonnées. Pour la sécurité dans les cloud computing nous avons présenté des approches intéressantes qui assurent un niveau de sécurité acceptable et qui restent insuffisants, car ces solutions sont basées sur la cryptographie des données qui n'est pas toujours une solution complète pour protéger les données, en plus le mécanisme de cryptage et de décryptage des données peut engendrer un gaspillage des ressources. Chen et He ont proposé la redondance des données ce qui assure la disponibilité des données, Jensen et al. ont proposé d'utiliser le protocole TLS et adapter le navigateur en intégrant la cryptographie XML. Mais ça ne garantit pas la sécurité de transfert des données à travers le réseau. Pour la proposition de Wang et al. Basée sur le chiffrement homomorphique qui assure de confidentialité des données lors de transfert dans le cloud. Néanmoins, le laboratoire de recherche en cryptographie dans le cloud de Microsoft a annoncé que cette nouvelle façon de chiffrer les données n'en est encore qu'à ces débuts et qu'ils sont loin de pouvoir exécuter sur les machines virtuelles des données cryptées avec le chiffrement homomorphique Zunnurhain et V. (2011).

5 Proposition d'une approche de sécurité des entrepôts

5.1 Profil utilisateur

Le profil de l'utilisateur est une représentation des préférences et les droits d'accès d'un utilisateur individuel. Il contient les informations nécessaires pour l'authentification, et le niveau de sécurité pour accéder aux données de l'entrepôt, alors que l'entrepôt peut utiliser les

détails de catégorie des données pour déterminer le contrôle d'accès. Le modèle de contrôle d'accès à base de rôles adopté est le modèle RBAC étendu. Il s'agit d'une politique d'accès répondant mieux aux besoins des grandes entreprises, gérant beaucoup de permissions pour un plus grand nombre d'utilisateurs. En effet, il est possible, pour certaines personnes, de cumuler plusieurs fonctions et par conséquent, plusieurs rôles. L'objectif est alors d'associer des profils aux profils, voire des rôles au rôle concerné. Il se crée une hiérarchie entre les rôles et donc une relation d'héritage. En effet, le rôle descendant hérite des permissions et restrictions du rôle ascendant.

Exemple. Prenons l'exemple d'un hôpital qui est composé, d'un médecin, d'une infirmière, d'un comptable et de patients. Dans un premier temps, nous pouvons différencier deux catégories de personnes bien distinctes : personnel de l'entreprise et les patients. Le patient pourra avoir accès aux données le concernant uniquement, contrairement au personnel qui se verra autoriser l'accès aux données de différents patients. Au sein du rôle «personnels de l'hôpital », nous recensons trois personnes ne pouvant accéder aux mêmes types de données. Tandis que le médecin et l'infirmière ont la possibilité d'accéder à l'ensemble du dossier médical du patient, le comptable, quant à lui ne peut accéder qu'aux données relatives à son séjour (prix, télé, repas...) et non à ses données relatives à son état de santé. Ainsi, bien que ces derniers figurent dans le même rôle, ils exercent deux tâches différentes. Dans ce cas, nous allons associer à un rôle un autre rôle.

5.2 Big Data

l'augmentation du volume et de la vitesse des données signifie que les organisations devront développer des techniques au niveau de la collecte, l'analyse et l'interprétation des données. Selon IBM Jacobs (2013) , 2,5 trillions d'octets de données sont générés chaque heure. Il est prévu également que pendant l'année 2020 seront générés 35 Zettaoctets, alors que seulement 1 Zettaoctets de données numériques ont été générés par l'humanité entre le début de l'informatique (1940) et 2010.

Formellement, une définition du BIG Data a été proposée pour mieux comprendre ce phénomène de données massives Davenport et Barth (2012). Elle est basée sur les trois dimensions : Volume, Vélocité et Variété (ou 3V).

- Volume : Le volume des données stockées aujourd'hui est en pleine expansion. Ainsi, Twitter génère quotidiennement à l'heure actuelle 7 téraoctets de données et Facebook 10 téraoctets.
- Vélocité : Elle représente à la fois la fréquence à laquelle les données sont générées, capturées et partagées. Par exemple, chaque seconde sont émis plus de 500 tweets, ce qui fait 50 millions chaque jour.
- Variété : En plus des données du types bases de données relationnelles, le Big Data s'intéresse également à un nombre important de formats et variétés de données : image, texte, réseaux sociaux, capteurs.

Les techniques Big Data doivent répondre à un certain nombre de problématiques, comme l'extraction de sens à partir de ces données, et permettre la possibilité de passer d'une analyse reporting (du passé), vers une analyse prédictive (futur).

5.3 L'infrastructure sécurisée à base de profil utilisateur

Les environnements de développement pour le projet de l'informatique de nuage sont difficiles à monter faute de manque de documentation, de stabilité de leurs composantes logicielles et de leurs complexités Klos (2012)

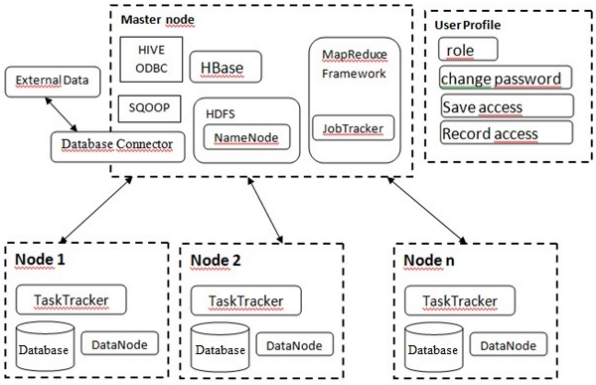

FIG. 1 – *L'infrastructure sécurisée à base de profil utilisateur*

Dans le but de diminuer l'impact de ces contraintes et de faciliter l'accès à un tel environnement, il a été décidé de monter un environnement de développement avec Hadoop et HBase HBase (2009). Hadoop est un environnement d'exécution distribué, performant et scalable, dont la vocation est de traiter des volumes de données considérables. Il repose sur un principe simple : distribuer les données sur un cluster de serveurs (nœuds de données), et migrer les traitements vers ces nœuds. Cette approche assure une mise à l'échelle des algorithmes et des traitements à développer. D'un point de vue logiciel, le big data est souvent associé à la pile technologique Hadoop mise en open source par Google apportant sur un système de stockage distribué, une base de données répartie, ainsi qu'un cadre de programmation et d'exécution des taches de calcul réparties. Dans l'architecture proposée, quelque soit la façon dont les utilisateurs se connectent à la table protégée (via une application, une interface Web ou SQL * Plus), le résultat est le même. Il n'y a pas de problème de sécurité des applications, puisque la politique d'accès est fixée à la table, et ne peut pas être contournée. Pour atteindre les objectifs explicités dans l'architecture, on doit fixer les droits d'accès des utilisateurs de l'entrepôt. Comme dans les systèmes d'information, cette tâche de sécurité peut être traitée au niveau conceptuel ou au niveau logique, et les droits d'accès fixés sont appliqués par le serveur OLAP.

6 Conclusion et perspectives

Dans cet article, nous avons essayés de présenter l'important rôle que jouent les entrepôts de données dans une organisation pour prendre de bonnes décisions stratégiques. Vue la sensibilité de ces ressources, leur sécurité est une nécessité pour les entreprises. C'est pour ce la que

nous avons essayés de cerner les problèmes liés à la sécurité des entrepôts de données dans les cloud computing, en tenant compte de leurs spécificités par rapport à une base de données ordinaire. Nous avons montrés les aspects permettant d'assurer la sécurité d'une information, et les différents modèles d'accès à l'entrepôt. Nous avons fait un tour d'horizon des travaux effectués pour la construction d'un entrepôt de données sécurisé. Et ce, au niveau conceptuel de l'entrepôt et au niveau de l'exploitation(OLAP). En se basant sur des travaux de recherche, nous avons proposés une approche simplifiée et fonctionnelle pour le contrôle d'accès aux entrepôts de données en se basant sur le profil de l'utilisateur,un avantage clé de notre approche est que nous accomplissons la sécurité du DW lors de la modélisation conceptuelle d'une manière indépendante de la plate-forme cible, ce qui permet de mettre en place le DW correspondant sur n'importe quel système de gestion de base de données sécurisé. Notre travail est orienté vers l'avenir, car nous avons l'intention d'appliquer notre approche qui permet de renforcer la sécurité des accès à l'entrepôt dans le cloud computing en fonction des profits utilisateurs de l'entreprise. Etudier la possibilité d'intégrer la cryptographie traditionnelle pour sécuriser le transfert des requêtes sur les réseaux.

Références

Chen, D. et Y. He (2010). A study on secure data storage strategy in cloud computing. *Journal of Convergence Information Technology*.

Davenport et Barth, A. (2012). How big data is different, in mit sloan management review.

Eavis, T. et A. Althamimi (2012). Olap authentication and authorization via queryre-writing. *The Fourth International Conference on Advances in Databases, Knowledge, and Data Applications*, 130–139.

Favre, C., F. Bentayeb, O. Boussaid, J. Darmont, G. Gavin, N. Harbi, N. Kabachi, et S. Loudcher (2013). Les entrepôts de données pour les nuls. . . ou pas.

Fernandez-Medina, E., J. Trujillo, R. Villarroel, et M. Piattini (2006). Access control and audit model for the multidimensional modeling of dws. *Decision Support Systems*, 1270–1289.

Hadoop, A. (2009). http ://hadoop.apache.org/.

HBase, A. (2009). http ://hadoop.apache.org/hbase/.

Jacobs, B. Dinsmore, T. (2013). Delivering value from big data with revolution r enterprise and hadoop, http ://www.revolutionanalytics.com/sites/default/files/driving-value-from-big-data-rre-hadoop.pdf.

Jensen, M., J. Schwenk, N. Gruschka, et L. L. Iacono (2012). On technical security issues in cloud computing. *IEEE International Conference*, 109–116.

Karkouda, K., N. Harbi, J. Darmont, et G. Gavin (2012). Confidentialité et disponibilité des données entreposées dans les nuages. *In 9ème atelier Fouille de données complexes*.

Kirkgoze, R., N. Katic, M. Stolba, et A. Tjoa (1997). A security concept for olap. *Proceedings of the 8th International Workshop on Database and Expert System Applications (DEXA'97)*, 619–626.

Klos, A. (2012). A. optimisation de recherche grace à hbase sous hadoop, rapport technique.

Priebe, T. et G. Pernul (2000). Towards olap security design - survey and research issues. *Proceedings of the 3rd ACM International Workshop on Data Warehousing and OLAP (DOLAP'00)*, 33–40.

Priebe, T. et G. Pernul (2001). A pragmatic approach to conceptual modeling of olap security. *Proceedings of the 20th International Conference on Conceptual Modeling (ER'01) 2224*, 311–324.

Rodriguez, A., E. Fernandez-Medina, J. Trujillo, et M. Piattini (2011). Secure business process model specification through a uml 2.0 activity diagram profile.

Soler, E., J. Trujillo, C. Blanco, et E. Fernandez-Medina (2009). Designing secure data warehouse by using mda and qvt. *Journal of Universal Computer Science 8 15*, 1607–1641.

Soler, E., J. Trujillo, E. Fernandez-Medina, et M. Piattini (2008). Building a secure star schema in data warehouses by an extension of the relational package from cwm. *Computer Standards and Interfaces 30*, 341–350.

Triki, S., H. Ben-Abdallah, N. Harbi, et O. Boussaid (2011). Securing data warehouses: A semi-automatic approach for inference prevention at the design level. *1st International Conference on Model and Data Engineering, Lecture Notes in Computer Science (LNCS) by Springer-Verlag*, 311–324.

Villarroel, R., E. Fernandez-Medina, et M. Piattini (2006). A uml 2.0/ocl extension for designing secure data warehouses. *Journal of Research and Practice in Information Technology 38*, 31–43.

Wang, C., Q. Wang, K. Ren, et W. Lou (2012). Privacy-preserving public auditing for data storage security in cloud computing. *IN FOCOM, 2010 Proceedings IEEE*, 1–9.

Zunnurhain, K. et V. S. V. (2011). Security in cloud computing. *In Proceedings of the International Conference on Security and Management*.

Summary

A data warehouse (DW) can be used as a powerful mechanism for discovering crucial business information. data warehouses benefit of the advent Cloud Computing to get remote access to data warehouse and high quality services. In this context, it becomes necessary to properly protect the data warehouse against different risks and dangers that are born in cloud computing, for ensure the confidentiality of data. In this sense, several proposals have been presented, however, none is considered a standard in the management of access to data warehouses. In this article, we will first analyze the problem of security in a data warehouse and its issues, and we will present an architecture of a secure warehouse is to manage access based on user profiles in the cloud.