الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي و البحث العلمي
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE

Centre de Recherche
Sur l'Information Scientifique et Technique

مركز البحث
في الإعلام العلمي والتقني

RAPPORT DU PROJET DE FIN D'ÉTUDES

Pour l'obtention du diplôme de

Post Graduation Spécialisée en Big Data et Calcul Intensif

Plateforme Big Data pour l'analyse de sécurité du trafic réseau

Réalisé par : FATHI Boualem & ELATSAMENIA Hamid

Encadreur : Dr YAHIAOUI Said & Dr ZEGHACHE Linda

Soutenu le : 13/12/2021. **Devant le jury** composé de :

Président : Mr.BELAZZOUGUI Djamal

Examinateurs : Mr.BOUCENNA Fateh

Année Universitaire : 2019 /2020

.

Remerciements

Avant tout nous remercions Dieu le tout puissant de nous avoir donné la santé, le courage, la force et la volonté d'entamer et d'accomplir ce modeste travail. et de nous avoir éclairé le chemin de la réussite.

En deuxième lieu nos remerciements vont en particulier à nos encadreurs : Dr YAHIAOUI Saïd & Dr ZEGHACHE Linda qu'ils trouvent ici le témoignage de notre profonde gratitude pour avoir proposé ce thème, et pour la confiance qu'ils nous ont accordée et les conseils fructueux qu'ils n'ont cessé de nous prodiguer tout au long de la réalisation de ce travail.

Nous tenons à exprimer nos profondes gratitudes et nos remerciements les plus vifs au président et aux membres de jury pour l'honneur qu'ils nous font en acceptant de juger notre travail.

Enfin, que tous ceux qui nous ont aidé, de près ou de loin, tout au long la réalisation de ce travail, les cadres et les enseignants du CERIST.

Dédicaces

Mes très chers parents : pour leur amour, soutien et énormes sacrifices ;

À ma chère femme Yamna ;

À mes chers enfants : Djohayna, Amani, Khawla et Djawad ;

À mes chers frères et sœurs ;

À tous mes collègues de la formation PGS de l'année 2020/2021 ;

À tous mes Amis.

FATHI Boualem

Mes très chers parents : pour leur amour, soutien et énormes sacrifices ;

À ma chère femme karima ;

À mes chers enfants : hadil, , ritedj et ahmed ;

À mes chers frères et sœurs ;

À tous mes collègues de la formation PGS de l'année 2020/2021 ;

À tous mes Amis.

ELASMENIA Hamid

ملخص

أصبح تحليل بيانات الشبكة من أهم مجالات معالجة البيانات الضخمة، لأن حجم وتعقيد البيانات المتعلقة بتحليل بيانات الشبكة أصبح كبيرًا جدًا بحيث لا يمكن معالجتها بواسطة الأدوات والوسائل التقليدية. بالإضافة إلى ذلك، هناك مجموعة كبيرة من أدوات معالجة البيانات الضخمة المستخدمة في تحليل بيانات الشبكة والتي يتم انتقاؤها حسب الواقع.

في هذا العمل، نقترح معالجة إشكالية إنشاء أرضية خاصة بتحليل بيانات الشبكة تستند على أدوات معالجة البيانات الضخمة، بطريقة مستنيرة تعتمد على دراسة الهدف المسطر والمعلومات المتوفرة حاليا. حيث قمنا بإجراء دراسة مقارنة لمختلف الأرضيات المتوفرة حاليا والمتعلقة بتحليل بيانات الشبكة وبشكل أكثر تحديداً، أدوات المعالجة الخاصة بالبيانات الضخمة، سمحت لنا هذه الدراسة المقارنة بتسليط الضوء على معايير الاختيار لتصميم حل يتم تكييفه حسب متطلبات التطبيق المقصود.

في النهاية، ولشرح الوضع في الخدمة لنظام خاص بتحليل بيانات الشبكة قائم على أدوات معالجة البيانات الضخمة من الناحية العملية، قمنا بتصميم والوضع في الخدمة لتطبيق يتمثل في مراقبة تدفق بيانات الشبكة حيث يتم أولا التقاط وحفظ البيانات بصفة آنية، بعدها يتم تحليلها وعرضها مع إظهار حالة استغلال الشبكة.

Abstract

Network traffic analysis is becoming a Big Data problem along with the growing size and complexity of security data that cannot be manipulated by traditional security tools. In addition, there is panoply of Big Data tools proposed for use in Network traffic analysis and choosing the right ones depends on reality.

In this work, we propose to treat the problem of setting up a Big data-based Network traffic analysis platform, in as knowledgeable as possible way, after an objective and well-documented study of the solutions proposed in the current state of the art. Thus, we made a comparative study of existing Big Data platforms dedicated to Network traffic analysis and more specifically, tools dedicated to the Big Data processing. This comparative study allowed us to highlight selection criteria for the design of a solution that is adapted to the requirements of the intended application.

At the end, and to explain in practical terms the implementation of a Network traffic analysis platform based on Big Data, we implemented a network traffic monitoring application consisting in capturing and saving, in real time, the network flow, and then making some analysis and visualization on the collected data to give an image on the network status.

Résumé

L'analyse du trafic réseau est en train de devenir un problème de Big Data en raison de la taille et de la complexité croissante des données de sécurité qui ne peuvent être manipulées par les outils de sécurité traditionnels. De plus, il existe une panoplie d'outils Big Data proposés pour la cybersécurité et dont le choix dépend de la réalité.

Dans ce travail, nous proposons de traiter la problématique de l'utilisation des outils Big data dans l'analyse du trafic réseau basée sur les données, de la manière la plus informée possible. Après une étude objective et bien documentée des solutions proposées dans l'état actuel des connaissances, nous avons réalisé une étude comparative des plates-formes Big Data existantes dédiées à l'analyse du trafic réseau et plus particulièrement des outils dédiés au traitement du Big Data. Cette étude comparative nous a permis de mettre en évidence des critères de sélection pour la conception d'une solution adaptée aux besoins de l'application envisagée.

Enfin, pour expliquer concrètement la mise en place d'une plateforme d'analyse du trafic réseau basée sur le Big Data, nous avons implémenté une application de surveillance du trafic réseau consistant à capturer et à sauvegarder, en temps réel, le flux du réseau, puis à effectuer des analyses et une visualisation sur les données collectées mettant en évidence l'état d'utilisation du réseau.

Table des matières

Liste des Tableaux

Liste des Figures