الجمهورية العربية السورية
جامعة البعث
كلية الهندسة المعلوماتية
قسم هندسة النظم والشبكات الحاسوبية

دراسة سرية البيانات وسلامتها في الحوسبة السحابية وتحسينها

دراسة أُعدت لنيل درجة الماجستير في هندسة النظم والشبكات الحاسوبية

إعداد

المهندس: عمار سهيل مقصود

إشراف

الدكتور المهندس: سهيل الحمود

مدرس في قسم هندسة النظم والشبكات الحاسوبية /جامعة البعث/

١٤٣٧ هـ – ٢٠١٦ م

الملخص

تُعتبر الحوسبة السحابية Cloud Computing نموذجاً للتكنولوجيا المتطورة، تُسهل توفير الخدمات والموارد الحاسوبية بشكل متنوع وديناميكي للزبائن، ويُعد التخزين السحابي Cloud Storage أحدَ أهمّ الخدمات المقدمة في هذا المجال، حيث يقدم للزبائن من مؤسسات وأفراد طريقة لتخزين المعلومات الخاصة بهم عن بعد. على الرغم من الفوائد التي يقدّمها التخزين السحابي، إلا أن الأخطار الأمنية المرافقة لاستخدامه ظلت تشكّل عائقاً أمام اعتماده كحل بالنسبة للزبائن ذوي المعلومات الهامة والحساسة. وقد نالت قضية تحقيق الأمن في التخزين السحابي نصيباً وافراً من الدراسات والأبحاث الحديثة التي تقدم مجموعةً واسعةً من الحلول المقترحة لأبرز المتطلبات الأمنية، لكن من الصعوبة تحديد كفاءة هذه الحلول من الناحية العملية ومن وجهة نظر مستخدمي التخزين السحابي، لذلك فإن أحد الموضوعات التي يقدّمها هذا البحث هو تسليط الضوء على أهم الدراسات التي تعالج مسألتي سرية البيانات المخزنة سحابياً وسلامتها، وتقييم هذه الحلول استناداً على معايير خاصة بالأداء والأمن وغيرها، مما يساعدنا على تحديد محدوديتها ونقاط القوة والضعف فيها.

من ناحية أخرى، قدّم البحث تقييماً و مقارنةً للحلول التي تعالج مسألة سرية البيانات، حيث تبيّن أنه في غالبية الحلول التقليدية لحماية سرية البيانات المخزنة عن بعد، يتم تشفير البيانات على حواسب المستخدمين قبل إرسالها إلى السحابة وفك تشفيرها عند استقبالها من أجل القيام بأي عمليات عليها، وهذا ما يحمّل مالك البيانات أعباءً إضافية، لذا اتجهت بعضُ الأبحاث نحو استخدام أنماطاً متقدمةً من التشفير والتي تناسب أنواعاً خاصةً من البيانات مثل البيانات العددية. ناقشَ البحث بعض هذه الدراسات، وقدم اقتراحاً لبروتوكول يؤمن سرية البيانات السحابية باستخدام التشفير التشابهي، يتضمن هذا البروتوكول مخططات الحماية التقليدية التي تؤمن وظائف التعمية والتوقيع الالكتروني والبصمة الالكترونية، إلى جانب تقنية التشفير التشابهي. تم توظيف البروتوكول الأمني المقترح في حماية بيانات مؤسسة، قمنا بتحقيق الحل المقترح ضمن بنية Map-Reduce في مشروع Hadoop، وهو نظام يقدم خدمات معالجة البيانات وتخزينها بشكل موزع. تؤكد نتائج هذا البحث إمكانية ترحيل العمليات الحسابية المطبقة على البيانات المخزنة سحابياُ من طرف مالك البيانات إلى طرف السحابة من دون أن يكون لتشفيرها أي تأثير سلبي لناحية زيادة حجم المعالجة المطلوبة من قبل السحابة لهذه البيانات، ودون ارتباط حجم المعالجة المطلوب في السحابة مع كمية البيانات المشفرة تشابهياً. وبالتالي يمكن لأي مستخدم تشفير بياناته قبل تخزينها ثم إنجاز عمليات معينة عليها دون الحاجة لتحميلها وفك تشفيرها. في سياقٍ آخر، قام هذا البحث بدراسة ومناقشة أهمّ الأبحاث التي تعالج مسألة سلامة البيانات في الحوسبة السحابية، حيث تم مقارنتها باعتماد مجموعة من المعايير الخاصة بالأمن و الأداء، ثم تم مناقشة أحد حلول سلامة البيانات المقترحة في الحوسبة السحابية والذي يعتمد على فكرة إلحاق بعض المعلومات الوصفية بالملفات والتأكد من سلامتها لاحقاً بدلاً من تحميل كامل الملفات إلى ناحية الزبون. لم نلحظ لهذا الحل تحقيقاً عملياً فقمنا بتنفيذه وتحقيقه بعد إضافة بعض

التحسينات، و أيضاً قمنا بنمذجته نظرياً و اختباره و تقييمه عملياً. تم تحديد إمكانيات الحل المقترح، ومقارنة النتائج النظرية الممثلة بالعلاقات المستنتجة مع النتائج العملية و إثبات تطابق أو صحة النموذج الرياضي الذي وضعناه لوصف هذا الحل.

Abstract

Cloud Computing is considered to be an instance of an advanced technology that facilitates providing customers with divers and with dynamic computer services and resources. Cloud Storage is considered to be one of the most prominent structure for providing services in this domain, as it provides institutions and individuals with a way to store/retrieve their data remotely. Despite the benefits of cloud storage, the security risks associated with its use have been an obstacle to its adoption as a solution for customers with significant and with sensitive information. The issue of security in cloud storage has had its fair share of modern studies and researches that provide a wide range of proposed solutions to the most prominent security requirements. It is -however- difficult to determine the efficiency of these solutions regarding the practice and regarding users' perspectives. Thus, one of the topics offered by this research is to highlight the most important studies that address the issues of confidentiality and integrity of cloud stored data, and then, to evaluate these solutions based on performance, security and other criteria. Such work will hopefully help identifying the limitations, strengths and weaknesses in solutions offered by reviewed studies.

Also, this research evaluates and compares solutions in the data confidentiality, where it was found that in the majority of traditional solutions used to protect the confidentiality of data stored remotely, the data is encrypted on the user computer before being sent to the cloud, and when needed for processing it must be downloaded first and decrypted. This burdens the data owner with higher client side processing and bandwidth costs. Therefore, some researches lean towards the use of advanced encryption patterns that fit specialized data types such as numerical data. This research discusses some of these studies, and proposes using homomorphic encryption to resolve the issue of data confidentiality. This solution includes traditional protection schemas, which provide the functions of cryptography, digital signature, along with homomorphic encryption. The proposed security protocol is employed and is used to protect data of an institution while keeps it feasible for processing by - possibly untrusted - third party. The proposed solution was implemented within the Map-Reduce infrastructure using Apache Hadoop. The results of this research confirm the possibility of deporting the calculations applied to data from the data owner side to the cloud side without

negative impact to the amount of processing required regardless of the size of the homomorphically-encrypted data. Therefore any user can encrypt his/her data before storing it, and then once uploaded he can perform certain operations on the data without the need to download/decrypt it first.

Also, this work study most recent researches that address the issue of data integrity in cloud computing. The related literature of several research papers was reviewed and compared using set of security and performance criteria. This research focuses on one of the proposed -and not implemented yet- solution for data integrity which inserts meta data section the uploaded data and uses it for integrity verification with small bandwidth overhead. This work implements the proposed solution with improvements. In addition to the improved implementation, a new theoretical statistical model is proposed to evaluate the solution performance with different configurations. Several experiments were carried out to proof the validity of the theoretical model.