

**A COMPARISON STUDY ON DATA RETRIEVAL SPEED BETWEEN
HASHING METHOD AND ENCRYPTION METHOD FOR SPEED
PERFORMANCE IN CLOUD COMPUTING**

**By
NG YANG YANG**



**- LIBRARY -
INFRASTRUCTURE UNIVERSITY
KUALA LUMPUR**

**Project Paper Submitted in Partial Fulfilment as the Requirement for the
Master In Information Technology In Faculty of Creative Media And
Innovative Technology**

IUKL

2016

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Infrastructure University Kuala Lumpur or any other institution.



NG YANG YANG

10 OCTOBER 2016

**A COMPARISON STUDY BETWEEN HASHING METHOD AND
ENCRYPTION METHOD FOR SPEED PERFORMANCE IN CLOUD
COMPUTING**

By

NG YANG YANG

December 2016

Chair: Dr Abudhahir Buhari

Faculty: Faculty of Creative Media and Innovative Technology

There is an increasingly amount of data in today IT world. This data comes from all kinds of sources that are use for data analysis or just video types of files which is extremely large per file. This data when it is store into cloud storage it will be needed to be encrypt in order to keep the data integrity to a certain level. The amount of time and computation level needed to encrypt a data is varied depending on the amount and size of the data. The verifier used to compute the cryptographic hash key as well as the secret key will be store together into cloud storage during encryption and will be used to decrypt the data during decryption. While this scheme is easy to be implemented, it is not usable for big data since one hash key is used to decrypt one section of data during decryption time. For this problem, a comparison study was done between encryption technique (Blow Fish) and hashing method (SHA) on the amount of time it takes to decrypt large data. Blow Fish needs one primary key in order to encrypt and decrypt the whole data while SHA used several keys in order to decrypt the whole data. This testing was done using two cloud platform provider, Google Cloud Platform and Amazon Web Service. Several groups of data was uploaded into both platform after encrypting the data. After that, both data was retrieve back along with decrypting the data. The total amount that was taken during the retrieving and decrypting the data. It is found that using a single key (Blow Fish) was considering much faster in decrypting the data compare to SHA who use several keys to decrypt the data.

ACKNOWLEDGEMENT

I would like to say a very much appreciation to my family who provide support to me all the time towards settling this master thesis. Without their support, its impossible for me to finish this thesis.

Secondly, I would like to say thank you to my supervisor, Dr Abudhahir Buhari who has guided me all the way towards completing this thesis with all his patient and time.

APPROVAL PAGE

We have examined this manuscript and verify that it meets the programme and University requirements for Master In Information Technology.

Name of Supervisor: Dr Abudhahir Buhari

Name of Faculty: Faculty of Creative Media and Innovative Technology

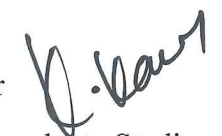
IUKL

Name of Internal Examiner: Mr Haslin Hasan

Name of Faculty: Faculty of Creative Media and Innovative Technology

IUKL

Kamaljeet Kaur



Centre for Postgraduate Studies

IUKL

Date: 16/12/16

TABLE OF CONTENT	PAGE
DECLARATION	ii
ABSTRACT	iii
ACKNOWLEDGEMENT	iv
APPROVAL PAGE	v
LIST OF TABLE	ix
LIST OF FIGURES	x

CHAPTER

1 INTRODUCTION	
1.1 Problem Statement	1
1.2 Objectives	2
1.3 Scope of Study	3
2 LITERATURE REVIEW	
2.1 Origin of Cloud Computing	5
2.1.1 1970s	6
2.1.2 1990s	6
2.1.3 2000s	6
2.2 Service Models	7
2.2.1 Infrastructure as a Service (IaaS)	7
2.2.2 Platform as a service (PaaS)	9
2.2.3 Software as a service (SaaS)	10
2.2.3.1 Characteristics of SaaS models:	11
2.2.3.2 Layers Of SaaS:	12
2.3 Big Data	12
2.3.1 Classification Of Big Data	15
2.3.2 Relationship between Cloud Computing and Big Data	16
2.3.3 Challenges of Big Data	18
2.4 Deployment Models	21

2.5 Blowfish and SHA Hash Function.....	22
2.5.1 Blowfish.....	22
2.5.2 SHA Hash Function	23
2.6 Security Concerns	25
2.6.1 Data Integrity	26
2.6.2 Data Intrusion.....	27
2.6.3 Service Availability	27
2.6.4 Current Resolving Study.....	28
2.7 Review of existing approaches	29
2.8 Current Works By Researchers	32
2.9 How Encryption Works.....	37
2.10 Current Research On Cloud Computing Security.....	39
2.10.1 Security and Privacy	39
2.10.1.1 Identity Management	39
2.10.1.2 Physical Security.....	40
2.10.1.3 Personnel Security	40
2.10.1.4 Privacy	40

3 **METHODOLOGY**

3.1 Introduction.....	41
3.2 Research Methodology	41
3.2.1 First Phase.....	41
3.2.2 Second Phase	42
3.2.3 Third Phase	42
3.2.4 Analysis.....	42
3.3 Conclusion	42

4 **RESULTS AND DISCUSSION**

4.1 Setting up Cloud Environment.....	44
4.2 Evaluation of the Performance Result	45

5	SUMMARY, CONCLUSION AND FURTHER RESEARCH	
	5.1 Conclusion	47
	5.2 Further Research	48
	References	49

LIST OF TABLE

Table 2.1: Total Execution Time for Cloud Network for SHA algorithm.....34

Table 2.2: Comparison Of Speed Up Ratio of the algorithm for different output36

LIST OF FIGURES

Figure 2.1: Information As A Service (IaaS)	9
Figure 2.2: Big Data Classification.....	16
Figure 2.3: Cloud Computer Usage In Big Data.....	17
Figure 2.4: Blowfish Procedure	23
Figure 2.5: SHA-1.....	24
Figure 2.6: SHA-2.....	24
Figure 2.7: SHA: Input Size VS Total Execution Time	35
Figure 2.8: Performance Comparison between Encryption.....	37
Figure 3.1: Testing Flow	43
Figure 4.1: Speed Performance of BlowFish and SHA1 (Low Traffic)	45
Figure 4.2: Speed Performance of BlowFish and SHA1 (High Traffic).....	46

CHAPTER 1

INTRODUCTION

This thesis is written for the purpose of further studying on the security of cloud computing. Nowadays, Cloud computing is one of the major feature in the computer world where it is involving with many useful feature to help user in service, library and etc. With the increasing of more and more user who is using this service, more data will be input and output in the cloud computing. By the end of this thesis, it is hope that this study can help to add more security to the large volume of data in the cloud computing.

1.1 Problem Statement

It is become increasingly difficult to encrypt data files when the data is extremely large. Computational power will be very low. Storage itself will be very big since all the sentinels are needed to be stored together. So it is not a practical solution to just download the file to check for its integrity since it required a high cost of transmission across the network.

The verifier used to computes the cryptographic hash of F using $hk(F)$ before storing the data file ion the cloud storage. The hash as well as the secret key to decrypt the data was stored along as well. When the user wanted to access the data, the verifier releases the secret key in order to check the integrity of the file and asked it to compute and return the value of $hk(F)$. This can be done for many times.

References

Hassan, Q. (2011). Demystifying cloud computing. *The Journal of Defense Software Engineering*, pp16-21.

Mell, P., & Grance, T. (2011). *The NIST definition of cloud computing*.

Haghighat, M., Zonouz, S., & Abdel-Mottaleb, M. (2015). CloudID: trustworthy cloud-based and cross-enterprise biometric identification. *Expert Systems with Applications*, 42(21), 7905-7916.

What is Cloud Computing? (2013, March 19). Retrieved from Amazon Web Services, <https://aws.amazon.com/what-is-cloud-computing/>

Cloud Computing: Clash of the clouds. (2009, October 15). Retrieved from *The Economist*, <http://www.economist.com/node/14637206>

Conn, S. (2008, June 26). Cloud computing: Clash of the clouds. Retrieved from Gartner, <http://www.gartner.com/newsroom/id/707508>

Eric Knorr. (2008, April 07). What cloud computing really means. Retrieved from InfoWorld, <http://www.infoworld.com/article/2683784/cloud-computing/what-cloud-computing-really-means.html>

Dealey, C. "Cloud Computing Working Group". Retrieved from Network Centric Operations Industry Consortium, <https://www.ncoic.org/technology/technical-team/cloud-computing-wg>

The economy is flat so why are financials Cloud vendors growing at more than 90 percent per annum? (2013, March 05). Retrieved from FSN, http://www.fsn.co.uk/channel_outsourcing/the_economy_is_flat_so_why_are_financials_cloud_vendors_growing_at_more_than_90_percent_per_annum#.UbmTsPIJPGA/

Liu, X. (2012). *Software reuse in the emerging cloud computing era*. Boca Raton, FL, United States: Information Science Reference

Schmidt, E., & Rosenberg, J. (2014). *How Google works*. Grand Central Publ

National Science Foundation. (1981). Diagram of CSNET. Retrieved from http://gu.friends-partners.org/Bookwriting/PART_I/Chapter_I/Total/Insertions/NFS/CSNET/CSNET.html

Antonio Regalado. (2011, October 31). Who Coined "Cloud Computing"? Retrieved from MIT Technology Review, <https://www.technologyreview.com/s/425970/who-coined-cloud-computing/>

James E. White. (1971, March). Network Specifications for Remote Job Entry and Remote Job Output Retrieval at UCSB. Retrieved from Network Working Group, <https://tools.ietf.org/html/rfc105>

"July, 1993 meeting report from the IP over ATM working group of the IETF". CH: Switch. Retrieved 2010-08-22. Retrived from <http://mirror.switch.ch/ftp/doc/ietf/ipatm/atm-minutes-93jul.txt>

Fernando J. Corbató, Marjorie Merwin Daggett, & Robert C. Daley. (1962, May 03). AN EXPERIMENTAL TIME-SHARING SYSTEM. Retrieved from <http://larch-www.lcs.mit.edu:8001/~corbato/sjcc62/>

Rochwerger, B., Breitgand, D., Levy, E., Galis, A., Nagin, K., Llorente, I. M., ... Galan, F. (2009). The reservoir model and architecture for open federated cloud computing. *IBM Journal of Research and Development*, 53(4), 4:1–4:11. doi:10.1147/jrd.2009.5429058

Kyriazis, D; A Menychtas; G Kousiouris; K Oberle; T Voith; M Boniface; E Oliveros; T Cucinotta; S Berger (November 2010). "A Real-time Service Oriented Infrastructure". *International Conference on Real-Time and Embedded Systems (RTES 2010)*. Singapore.

Gogouvitis, S., Konstanteli, K., Waldschmidt, S., Kousiouris, G., Katsaros, G., Menychtas, A., ... Varvarigou, T. (2012). Workflow management for soft real-time interactive applications in virtualized environments. *Future Generation Computer Systems*, 28(1), 193–209. doi:10.1016/j.future.2011.05.017

Amy Schurr. (2008, August 07). Keep an eye on cloud computing. Retrieved from Network World, <http://www.networkworld.com/article/2281563/data-center/keep-an-eye-on-cloud-computing.html>

STAMFORD, Conn. (2008, August 18). Gartner Says Worldwide IT Spending On Pace to Surpass \$3.4 Trillion in 2008. Retrieved from Gartner, <http://www.gartner.com/newsroom/id/742913>

Oracle Cloud, Enterprise-Grade Cloud Solutions: SaaS, PaaS, and IaaS. (2014, October 12). Retrieved from Oracle Cloud, <https://cloud.oracle.com/home>

Bob Evans. (2012, October 09). Larry Ellison Doesn't Get the Cloud: The Dumbest Idea of 2013. Retrieved from Forbes, <http://www.forbes.com/sites/oracle/2012/10/09/larry-ellison-doesnt-get-the-cloud-the-dumbest-idea-of-2013/#127e5bde6f69>

Bob Evans. (2012, October 03). Oracle Disrupts Cloud Industry with End-to-End Approach. Retrieved from Forbes, <http://www.forbes.com/sites/oracle/2012/10/03/oracle-disrupts-cloud-industry-with-end-to-end-approach/#1f0104996ef9>

Kavian, Y., & Rashvand, H. (Eds.). (2014). *Using cross-layer techniques for communication systems*. Boca Raton, FL, United States: Information Science Reference.

Cavanillas, J. M., Curry, E., & Wahlster, W. (Eds.). (2016). *New horizons for a data-driven economy*. doi:10.1007/978-3-319-21569-3

Data, data everywhere. (2010, February 25). Retrieved from The Economist, <http://www.economist.com/node/15557443>

Joe Hellerstein. (2009, November 09). Parallel Programming in the Age of Big Data. Retrieved from Gigaom, <https://gigaom.com/2008/11/09/mapreduce-leads-the-way-for-parallel-programming/>

Hammerbacher, J., & Segaran, T. (2009). *Beautiful data: The stories behind elegant data solutions*. United States: O'Reilly Media, Inc, USA.

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information Systems*, 47, 98-115.

Labrinidis A, Jagadish HV (2012) Challenges and opportunities with big data. *Proc VLDB Endowment* 5(12):2032-2033

Chaudhuri S, Dayal U, Narasayya V (2011) An overview of business intelligence technology. *Commun ACM* 54(8): 88-98

Agrawal D, Bernstein P, Bertino E, Davidson S, Dayal U, Franklin M, Gehrke J, Haas L, Halevy A, Han J et al (2012) Challenges and opportunities with big data. A community white paper developed by leading researchers across the United States.

John Foley. (2010, August 22). Private Clouds Take Shape. Retrieved from Information Week, <http://www.informationweek.com/news/services/business/showArticle.jhtml?articleID=209904474>

Margaret Rouse. What is public cloud. Retrieved from TechTarget, <http://searchcloudcomputing.techtarget.com/definition/public-cloud>

Tom Bittman. (2012, September 24). Mind the Gap: Here Comes Hybrid Cloud. Retrieved from Gartner, http://blogs.gartner.com/thomas_bittman/2012/09/24/mind-the-gap-here-comes-hybrid-cloud/

Kaur, N., & Singh, H (2013). Efficient and Secure Data Storage in Cloud Computing Through Blowfish, RSA and Hash Function.

Garfinkel, S. (2007). An evaluation of amazon's grid computing services: EC2, S3, and SQS.

Zisis, D., & Lekkas, D. (2012). Addressing cloud computing security issues. *Future Generation computer systems*, 28(3), 583-592.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., ... & Zaharia, M. (2010). A view of cloud computing. *Communications of the ACM*, 53(4), 50-58.

Cox, M., & Ellsworth, D. (1997, August). Managing big data for scientific visualization. In *ACM Siggraph* (Vol. 97, pp. 146-162).

James Manyika, M. C., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A.

H. Big data: The next frontier for innovation, competition, and productivity.

Berman, J. J. (2013). Principles of big data: preparing, sharing, and analyzing complex information. Newnes.

Mohta, A., & Awasti, L. K. (2012). Cloud data security while using third party auditor. *International Journal of Scientific & Engineering Research*, 3(6), 1.

Ateniese, G., Burns, R., Curtmola, R., Herring, J., Kissner, L., Peterson, Z., & Song, D. (2007, October). Provable data possession at untrusted stores. In *Proceedings of the 14th ACM conference on Computer and communications security* (pp. 598-609). Acm.

Wang, C., Wang, Q., Ren, K., & Lou, W. (2010, March). Privacy-preserving public auditing for data storage security in cloud computing. In *INFOCOM, 2010 Proceedings IEEE* (pp. 1-9). Ieee.

Juels, A., Kaliski Jr, B. S., Bowers, K. D., & Oprea, A. M. (2013). U.S. Patent No. 8,381,062. Washington, DC: U.S. Patent and Trademark Office.

As'habi, K., Vafabakhsh, A., & Borji, S. (2016). Data Transmission Security In Cloud Computing (pp. pp. 37-45): *Indian Journal of Fundamental and Applied Life Sciences* ISSN: 2231- 6345.

VSingh, V. K., & Dutta, D. M. (2014). Analyzing Cryptographic Algorithms For Secure Cloud Network (Vol. IJASCSE Volume 3, Issue 6): *International Journal of advanced studies in Computer Science and Engineering*.

Yu, S., Wang, C., Ren, K., & Lou, W. (2010). Achieving Secure, Scalable, and Fine-grained Data Access Control in Cloud Computing

Bhardwaj, S., Jain, L., & Jain, S. (2010). Cloud Computing: A Study Of Infrastructure As A Service (IaaS). *International Journal of Engineering and Information Technology*, IJEIT 2010, 2(1), 60-63

Satyanarayana, S. (2012). Cloud Computing: SaaS. *GESJ: Computer Science and Telecommunications*, 4(36), 2012.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data. The management revolution. *Harvard Bus Rev*, 90(10), 61-67.

Agrawal, D., Das, S., & El Abbadi, A. (2011, March). Big data and cloud computing: current state and future opportunities. In *Proceedings of the 14th International Conference on Extending Database Technology* (pp. 530-533). ACM.

O'Driscoll, A., Daugelaite, J., & Sleator, R. D. (2013). 'Big data', Hadoop and cloud computing in genomics. *Journal of biomedical informatics*, 46(5), 774-781.

Chen, M., Mao, S., & Liu, Y. (2014). Big data: a survey. *Mobile Networks and Applications*, 19(2), 171-209.

C. Cachin, I. Keidar and A. Shraer (2009). Trusting the cloud, *ACM SIGACT News*, 40, pp. 81-86.

Grembowski, T., Lien, R., Gaj, K., Nguyen, N., Bellows, P., Flidr, J., ... & Schott, B. (2002, September). Comparative analysis of the hardware implementations of hash functions SHA-1 and SHA-512. In *International Conference on Information Security* (pp. 75-89). Springer Berlin Heidelberg.

Thakur, J., & Kumar, N. (2011). DES, AES and Blowfish: Symmetric key cryptography algorithms simulation based performance analysis. *International journal of emerging technology and advanced engineering*, 1(2), 6-12.

