Australian Journal of Basic and Applied Sciences, 10(12) July 2016, Pages: 175-181

AUSTRALIAN JOURNAL OF BASIC AND APPLIED SCIENCES

ISSN:1991-8178 EISSN: 2309-8414
Journal home page: www.ajbasweb.com

A Methodology to Detect Most Effective Compression Technique Based on Time Complexity Cloud Migration for High Image Data Load

[1]Dasaradha Ramaiah K and [2]T. Venugopal

[1]Professor and Head - IT Department,BVRIT,Hyderabad -500082, India.
[2]Assoc. Professor & Head of CSE Department,JNTUCE- Sultanpur,Hyderabad -505501, India.

Address For Correspondence:
Dasaradha Ramaiah K, Professor and Head - IT Department,BVRIT,Hyderabad -500082, India
E-mail: joydasaradh@gmail.com

ARTICLE INFO

Article history:
Received 26 April 2016
Accepted 21 July 2016
Published 30 July 2016

Keywords:
DWT; K-Mean; 3D Spiral JPEG 2000; Cloud Storage Comparator; Microsoft Azure; Google Apps engine; Amazon Web Service; Cloud Storage.

ABSTRACT

Digital image compression technology is of special interest for the fast transmission and real-time processing of digital image. However image compression is a trend of research grown for a long time and there are multiple sophisticated approaches which significantly improve the compression rate and down grade computation time, a basic comparison with the aspect of storage on cloud environment is required. This work analyzes about the background of image compression, including when image compression is needed, categories of techniques and their properties. However compression uses many algorithms that store an exact representation or an approximation of the original image in a smaller number of bytes that can be expanded back to its uncompressed form with a corresponding decompression algorithm. This work majorly focuses on a novel framework for analyzing the impact of generic image compression techniques based on compression ratio and time for high data loads on cloud storage services and find the optimal set of algorithms most suitable for cloud storage.

INTRODUCTION

The digital image is most popular way of representing the information over internet because of its effectiveness of presenting information and the continuous efforts to improve the compression algorithms (Ding, J.J. and J.D. Huang, 2008) for low cost storage over cloud infrastructure. The requirement for high resolution information for large amount of data storage cannot be ignored. The digital image contains significant amount of duplicate, redundant and complex information in high density, hence the compression of the image data cannot be neglected (Gonzalez, R.C. and R.E. Woods, 2002). A set of great work has been conducted in the area of image compression; however a comparative study needs to be conducted to evaluate the performance of most popular image compression algorithms over different cloud storage platforms. The different two categories for image compression is majorly divided based on the information can be recovered with or without loose. Each category is consisting of multiple algorithms. However the detailed comparison is performed in the previous work and this work is focused focuses on a novel framework for analyzing the impact of generic image compression techniques based on compression ratio and time for high data loads on cloud storage services and find the optimal set of algorithms most suitable for cloud storage.

Image Compression Algorithms:

We consider few most popular lossy compression techniques such as DWT, K-Mean and 3D Spiral JPEG.

To Cite This Article: Dasaradha Ramaiah K and T. Venugopal, A Methodology to Detect Most Effective Compression Technique Based on Time Complexity Cloud Migration for High Image Data Load. ***Aust. J. Basic & Appl. Sci.,*** *10(12): 175-181, 2016*

GDI / BMP:

The bitmapped graphics or the BMP format is used by the Microsoft Windows graphics subsystem or GDI as a simple graphic information representation format internally. As the popularity of the platform increases, thus increases the use of this file format. The BMP format does not include any compression technique, thus it remains lossless (Gonzalez, R.C. and R.E. Woods, 2002).

Deflate / png:

The Portable Network Graphics or the PNG is a bitmap format, which uses a lossless data compress. The compression algorithm is called DEFLATE compression algorithm, which is a combination of LZ77 algorithm and Huffman Coding algorithm (Gonzalez, R.C. and R.E. Woods, 2002).

Lzw / tiff:

The Tagged Image File format or TIFF is majorly intended to store image with dense information like photograph and line art. The major reason for popularity is to be used for high color depth in image manipulation tools or image processing tools for optical character recognition. The compression technique used for TIFF is LZW for majority of the cases (Ding, J.J. and J.D. Huang, 2008; Gonzalez, R.C. and R.E. Woods, 2002).

Dwt:

Discrete Wavelet Transform (DWT) has gained wide spread acceptance in signal processing and image compression because it provides an extremely flexible multi-resolution image and can decompose an original image into different sub band images including low- and high- frequencies. The high frequency components provides with the information about the fine edges of the image which is very constructive for recovering the original image at the receiver end. The original image is converted to YCbCr color space so that the Discrete Wavelet transform can be applied to the luminance (Y) component. Then the image containing the low pass sub-band, diagonal sub-band information and Cb& Cr component information is processed further (Pennebaker, W.B., J.L. Mitchell, 1993).

K-Means:

K-means algorithm is a form of vector quantization which makes the use of K-means clustering in which a set of vectors are taken as input. It is a grouping method of bunching analysis which aims at partitioning ‘n’ groups into ‘k’ clusters in which each group of pixels belongs to the cluster with the nearest mean. A similar method is that of LBG Technique which also comes under Vector Quantization. Also the algorithm of LBG technique is quite similar to the below explained K-means algorithm (The JPEG web page: http://www.jpeg.org/.).

3D Spiral JPEG 2000 Encoding:

In 3D Spiral JPEG 2000 encoding, 2-D image is initially divided into a set of 8*8 pixel blocks. Then an 8*8*8 dimensional 3-D cube is formed simply using spiral scanning procedure on each 8*8 pixel block starting from the center of the graphic and going outwards (Wang, H., *et al*., 2012).

Comparisons of Compression Ration and Time Scale:

During the image compression operations is important to identify the data set, which can be allowed to be redundant and which should be eliminated. Here is the compression ratio comparison [Table – 1]:

Table 1: Compression Ratio Comparison

Algorithm Type	Compression Ratio
GDI / BMP	0.000
DEFLATE / PNG	0.996
LZW / TIFF	0.600
DWT	0.050
K-Means	0.600
JPEG	0.300
JPEG 2000	0.200

The results are been analyzed graphically as well [Figure – 1]:

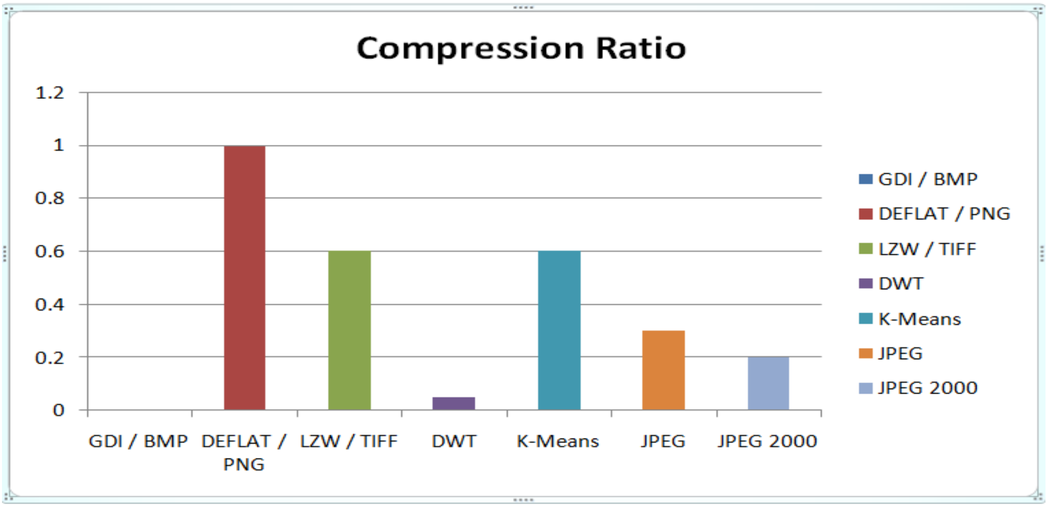

Fig. 1: Compression Ratio Analysis

Based on multiple parameters, considering the effectiveness of the compression the time consumption is calculated below [Table – 2] [Table – 3]:

Table 2: Comparison of Compression Time with Load up to 10000KB (Maximum Comparable Load)

Algorithm Type	RAW Data Size	Compression Time	Avg. Compression Time for 100KB
GDI / BMP	10000 KB	0.000	0.000
DEFLATE / PNG	10000 KB	120 Sec	0.12 Sec
LZW / TIFF	10000 KB	145Sec	0.15 Sec
DWT	10000 KB	260 Sec	0.26 Sec
K-Means	10000 KB	424 Sec	0.43 Sec
JPEG	10000 KB	628 Sec	0.63 Sec
JPEG 2000	10000 KB	628 Sec	0.63 Sec

The results are been analyzed graphically as well [Figure – 2]:

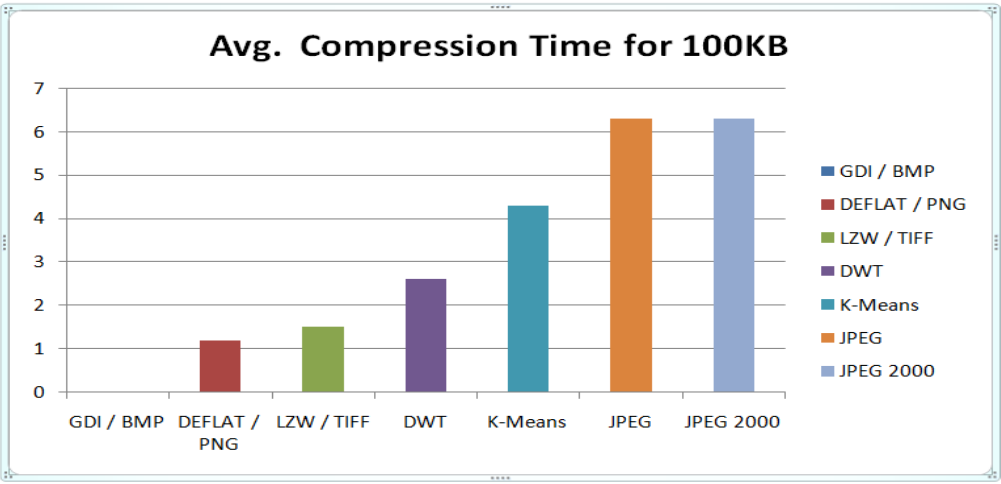

Fig. 2: Compression Time Analysis

Table 3: Comparison of De-Compression Time with Load up to 10000KB (Maximum Comparable Load)

Algorithm Type	RAW Data Size	Decompression Time	Avg. Decompression Time for 100KB
GDI / BMP	10000 KB	0.000	0.000
DEFLATE / PNG	10000 KB	72 Sec	0.072 Sec
LZW / TIFF	10000 KB	88 Sec	0.089 Sec
DWT	10000 KB	106 Sec	0.11 Sec
K-Means	10000 KB	300 Sec	0.30 Sec
JPEG	10000 KB	312 Sec	0.31 Sec
JPEG 2000	10000 KB	312 Sec	0.31 Sec

The results are been analyzed graphically as well [Figure – 3]:

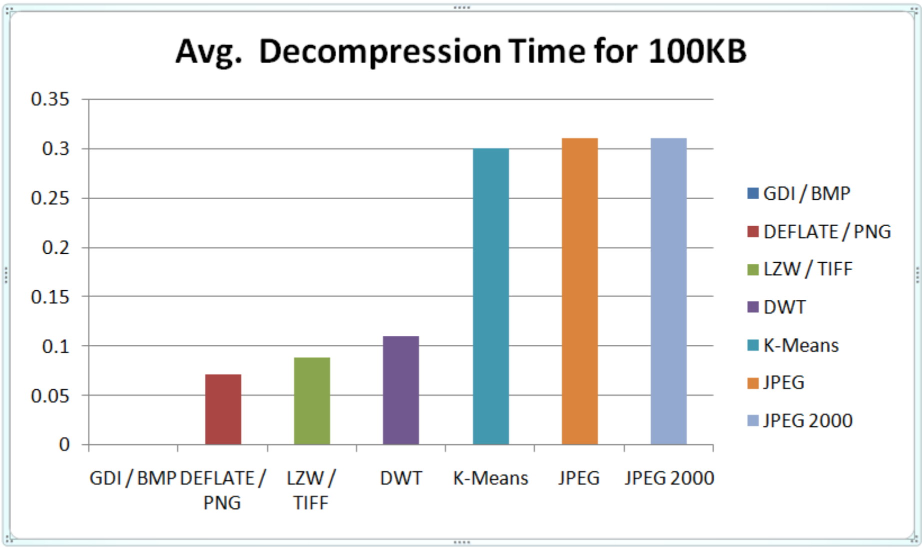

Fig. 3: Decompression Time Analysis

Cloud Storage Environments and Performance Comparison for Image Data:

We understand the maximized world of options for cloud storage (Wang, H., *et al.*, 2012) and selected few most popular for testing the data storage and retrieval for multiple image files load. The most popular cloud based storage solutions are DropBox / AWS S3 based on Amazon Cloud Service, OneDrive (Formerly known as SkyDrive) based on Microsoft Azure and Google Drive based on Google App Engine cloud service (Li, A., *et al.*, 2010; Bermudez, I.N., *et al.*, 2013).

Table 4: Test -1 Comparison of Upload / Download Speed in Seconds for Cloud Storage (Amazon. Cloud Drive v. 2.0.2013; Drago, I., *et al.*, 2012; DropBox / AWS S3. v. 2.0.8. https://www.DropBox; Google. Drive v. 1.9.4536.8202; Google. Network Introduction.; Google. Trends.; Microsoft. SkyDrive v. 17.0.2006.0314)

Image Type and Size	Test 1		
	DropBox / AWS S3	Google Drive	One Drive
10 MB BMP File	15	13	29
10 MB PNG File	15	13	29
10 MB TIFF File	15	13	29
10 MB DWT Compression File	15	13	29
10 MB K-Means Compression File	15	13	29
10 MB JPEG File	15	13	29
10 MB JPEG 2000 File	15	13	29

Table 5: Test - 2Comparison of Upload / Download Speed in Seconds for Cloud Storage (Amazon. Cloud Drive v. 2.0.2013.841; Drago, I., *et al.*, 2012; DropBox / AWS S3. v. 2.0.8; Google. Drive v. 1.9.4536.8202; Google. Network Introduction; Google. Trends; Microsoft. SkyDrive v. 17.0.2006.0314)

Image Type and Size	Test 2		
	DropBox / AWS S3	Google Drive	One Drive
10 MB BMP File	14	14	30
10 MB PNG File	14	14	30
10 MB TIFF File	14	14	30
10 MB DWT Compression File	14	14	30
10 MB K-Means Compression File	14	14	30
10 MB JPEG File	14	14	30
10 MB JPEG 2000 File	14	14	30

Table 6: Test - 3 Comparison of Upload / Download Speed in Seconds for Cloud Storage (Amazon. Cloud Drive v. 2.0.2013.841; Drago, I., *et al.*, 2012; DropBox / AWS S3. v. 2.0.8; Google. Drive v. 1.9.4536.8202; Google. Network Introduction; Google. Trends; Microsoft. SkyDrive v. 17.0.2006.0314)

Image Type and Size	Test 3		
	DropBox / AWS S3	Google Drive	One Drive
10 MB BMP File	16	12	27
10 MB PNG File	16	12	27
10 MB TIFF File	16	12	27
10 MB DWT Compression File	16	12	27
10 MB K-Means Compression File	16	12	27

10 MB JPEG File	16	12	27
10 MB JPEG 2000 File	16	12	27

Table 7: Average of Comparison of Upload / Download Speed in Seconds for Cloud Storage (Amazon. Cloud Drive v. 2.0.2013.841; Drago, I., *et al*., 2012; DropBox / AWS S3. v. 2.0.8; Google. Drive v. 1.9.4536.8202; Google. Network Introduction; Google. Trends; Microsoft. SkyDrive v. 17.0.2006.0314)

Image Type and Size	Average		
	DropBox / AWS S3	Google Drive	One Drive
10 MB BMP File	15	13	28.66
10 MB PNG File	15	13	28.66
10 MB TIFF File	15	13	28.66
10 MB DWT Compression File	15	13	28.66
10 MB K-Means Compression File	15	13	28.66
10 MB JPEG File	15	13	28.66
10 MB JPEG 2000 File	15	13	28.66

Hence the test results show the type of image information file does not imply any difference on the upload and download speed on various storage solutions [Figure – 4].

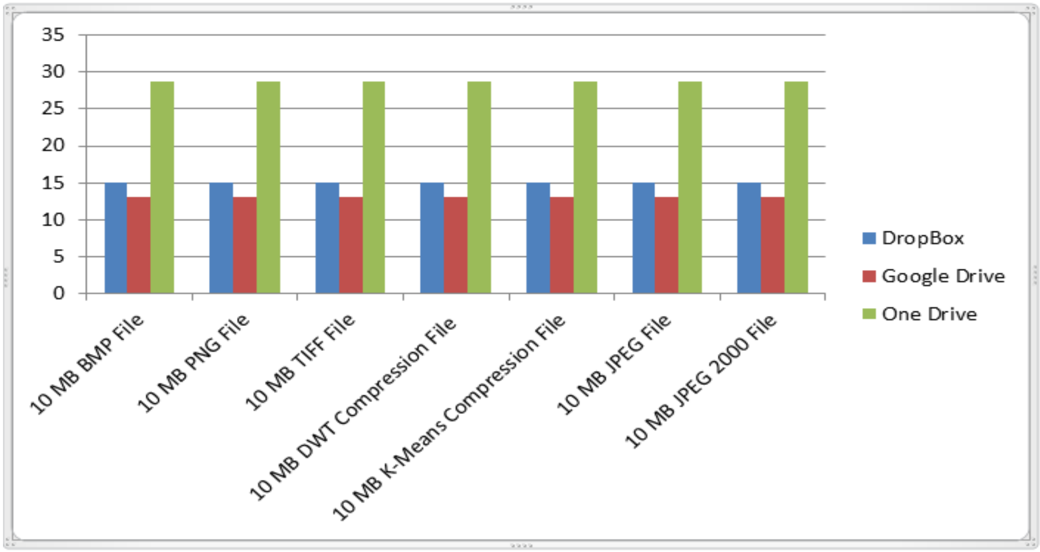

Fig. 4: Average Time for Upload and Download for Cloud Storage Services

Proposed Algorithm for Compression Technique Selection:

The major purpose of this work is to generate an automatic Compression technique to select the best compression technique depending on the average compression time and ratio. Here we propose architecture of the algorithm to decide the most applicable compression technique [Figure - 5].

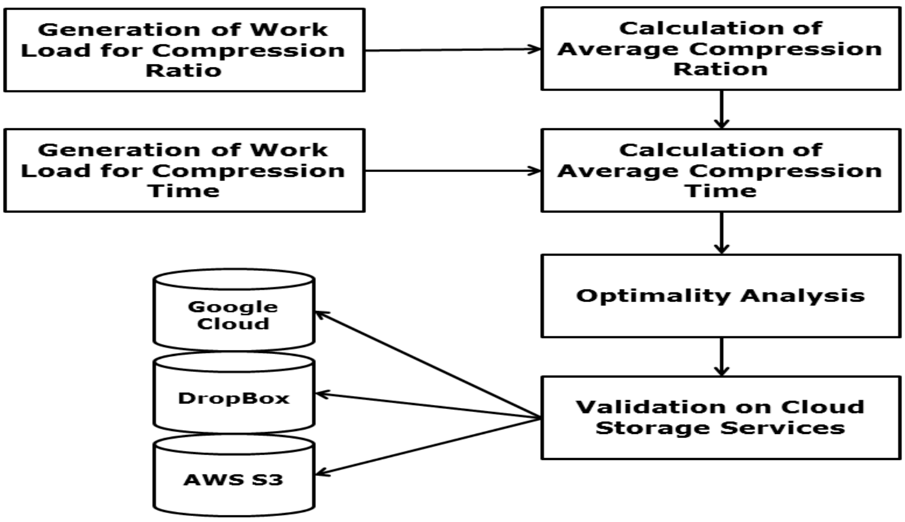

Fig. 5: Auto Selector Application

Australian Journal of Basic and Applied Sciences, 10(12) July 2016, Pages: 175-181

The algorithm for identifying the most suitable compression technique is as following:

Finding the Average Compression Times:

In this part of the work, the generic formula calculates the average compression time.

Firstly the total compression time is been calculated:

$$Comp.Time = \sum_{i=1}^{n} Comp.Time(i) \qquad (1)$$

Then, the average compression time for all considered algorithm is been calculated:

$$Comp.Time_{Avg} = \frac{Comp.Time}{c\left\{\sum_{i=1}^{n} Comp.Time(i)\right\}} \qquad (1)$$

Hence, the result of this phase is the average compression time as furnished below [Table – 8]:

Table 8: Average Compression time for Listed Algorithms (Maximum Comparable Workload)

Algorithm Type	RAW Data Size	Compression Time	Avg. Compression Time for 100KB
BMP	10000 KB	0.000	3.171429 Sec
PNG	10000 KB	120 Sec	
TIFF	10000 KB	145Sec	
DWT	10000 KB	260 Sec	
K-Means	10000 KB	424 Sec	
JPEG	10000 KB	628 Sec	
JPEG 2000	10000 KB	628 Sec	

Finding the Average Compression Ratio:

In this part of the work, the generic formula calculates the average compression ratio.

Firstly the total compression ratio is been calculated:

$$Comp.Ratio = \sum_{i=1}^{n} Comp.Ratio(i) \qquad (2)$$

Then, the average compression time for all considered algorithm is been calculated:

$$Comp.Ratio_{Avg} = \frac{Comp.Ratio}{c\left\{\sum_{i=1}^{n} Comp.Ratio(i)\right\}} \qquad (3)$$

Hence, the result of this phase is the average compression ratio as furnished below [Table – 9]:

Table 9: Average Compression Ratio for Listed Algorithms

Algorithm Type	Compression Ratio	Average Compression Ratio
BMP	0.000	0.392286
PNG	0.996	
TIFF	0.600	
DWT	0.050	
K-Means	0.600	
JPEG	0.300	
JPEG 2000	0.200	

Results and Conclusion:

The work has analyzed the most popular compression algorithms in the space of image compression for the algorithm facts based on the compression ratio and time.

This work finally results into an algorithm for recommending the most suitable image compression technique and also proved that the type of the compression technique will not have any impact on the upload and download speed, however the size of the image data will have impact.

During the result analysis, the work proposes an optimality analysis based on the average compression time and average compression ratio [Figure – 6]:

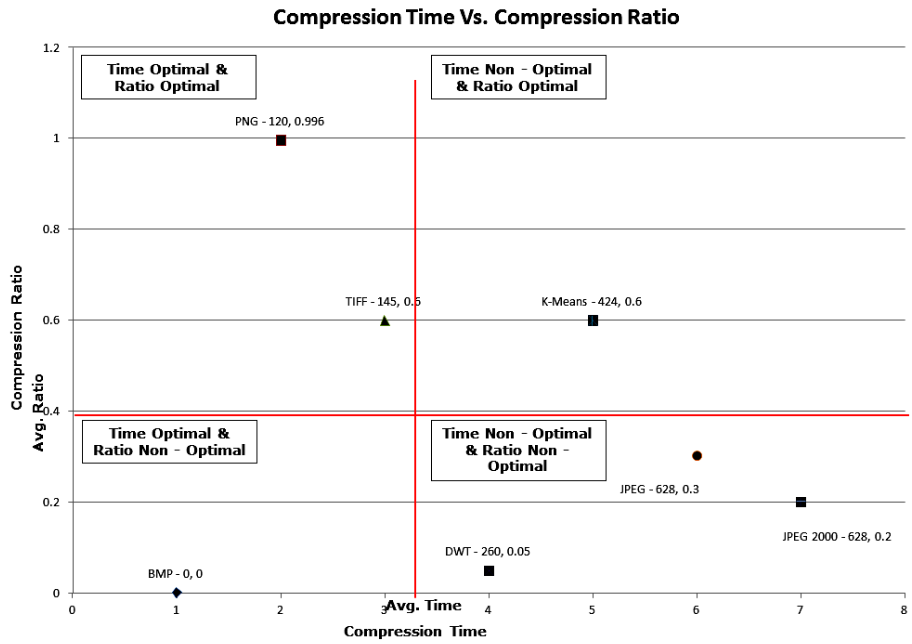

Fig. 6: Optimality Analysis

REFERENCES

Ding, J.J. and J.D. Huang, 2008. “Image Compression by Segmentation and Boundary Description,” June.

Gonzalez, R.C. and R.E. Woods, 2002. Digital Image Processing 2/E. Upper Saddle River, NJ: Prentice-Hall.

Pennebaker, W.B., J.L. Mitchell, 1993. “JPEG Still Image Data Compression Standard,” Van Nostrand Reinhold.

The JPEG web page: http://www.jpeg.org/.

Wang, H., R. Shea, F. Wang and J. Liu, 2012. On the Impact of Virtualization on DropBox / AWS S3-like Cloud File Storage/Synchronization Services. In Proceedings of the IEEE 20th International Workshop on Quality of Service, IWQoS ’12(11): 1–11.

Li, A., X. Yang, S. Kandula and M. Zhang, 2010. CloudCmp: Comparing Public Cloud Providers. In Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement, IMC’10, pp: 1-14.

Amazon. Cloud Drive v. 2.0.2013.841. http://www.amazon.com/gp/feature.html?docId=1000828861.

Bermudez, I.N., S. Traverso, M. Mellia and M.M. Munaf`o, 2013. Exploring the Cloud from Passive Measurements: the Amazon AWS case. In The 32nd Annual IEEE International Conference on Computer Communications, INFOCOM’13.

Drago, I., M. Mellia, M.M. Munaf`o, A. Sperotto, R. Sadre and A. Pras, 2012. Inside DropBox / AWS S3:Understanding Personal Cloud Storage Services. In Proceedings of the 12th ACM Internet Measurement Conference, IMC’12, pp: 481-494.

DropBox / AWS S3. v. 2.0.8. https://www.DropBox / AWS S3.com/release_notes.

Google. Drive v. 1.9.4536.8202. https://tools.google.com/dlpage/drive.

Google. Network Introduction. https://peering.google.com/about/delivery_ecosystem.html.

Google. Trends. http://www.google.com/trends/.

Microsoft. SkyDrive v. 17.0.2006.0314. https://skydrive.live.com/.

http://en.wikipedia.org/wiki?curid=2679981

http://en.wikipedia.org/wiki?curid=3748933

Jin Li “ Image Compression - the Mathematics of JPEG, 2000”, Microsoft Research, Communication Collaboration and Signal Processing.