UUM NETWORK TRAFFIC ANALYSIS AND USER' WEBSITE PREFERENCES

MUSTAFA MOHAMMED IBRAHIM AL-KAWAZ

UNIVERSITI UTARA MALAYSIA
2012

UUM NETWORK TRAFFIC ANALYSIS AND USER' WEBSITE PREFERENCES

BY

Mustafa Mohammed Haki Ibrahim

(808988)

KOLEJ SASTERA DAN SAINS
(College of Arts and Sciences)
Universiti Utara Malaysia

PERAKUAN KERJA KERTAS PROJEK
(Certificate of Project Paper)

Saya, yang bertandatangan, memperakukan bahawa
(I, the undersigned, certifies that)

MUSTAFA MOHAMMED IBRAHIM AL-KAWAZ
(808988)

calon untuk Ijazah
(candidate for the degree of) **MSc. (Information Technology)**

telah mengemukakan kertas projek yang bertajuk
(has presented his/her project of the following title)

UUM NETWORK TRAFFIC ANALYSIS AND USERS' WEBSITE PREFERENCES

seperti yang tercatat di muka surat tajuk dan kulit kertas projek
(as it appears on the title page and front cover of project)

bahawa kertas projek tersebut boleh diterima dari segi bentuk serta kandungan dan meliputi bidang ilmu dengan memuaskan.
(that this project is in acceptable form and content, and that a satisfactory knowledge of the field is covered by the project).

Nama Penyelia
(Name of Supervisor) : **DR. MOHD HASBULLAH OMAR**

Tandatangan
(Signature) : Tarikh (Date) : 27/6/2012

Nama Penyelia
(Name of Supervisor) : **MR. ADIB M. MONZER HABBAL**

Tandatangan
(Signature) : Tarikh (Date) : 27/6/2012

DEAN OF AWANG HAD SALLEH GRADUATE SCHOOL
UNIVERSITI UTARA MALAYSIA

PERMISSION TO USE

In presenting this study in partial fulfilment of the requirements for a postgraduate degree from the Universiti Utara Malaysia, I agree that the University Library may make it freely available for inspection. I further agree that permission for copying of this study in any manner in whole or in part, for scholarly purposes may be granted by my supervisor(s) or in their absence by the Dean of Awang Had Salleh Graduate School. It is understood that any copying or publication or use of this study or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my study.

Requests for permission to copy or to make other use of materials in this study, in whole or in part, should be addressed to

Dean of Awang Had Salleh Graduate School
College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok
Kedah Darul Aman
Malaysia

Abstract

The current world is experiencing a revolution in Internet services and networking; a revolution that provided and continues to provide varying features and invaluable tools to computer networks. On the other hand, several problems are being faced by users and global organizations in networking including lack of bandwidth and packet loss during transmission which impacts Internet efficiency and the performance of network. These issues can be rectified through the measurement and analysis of the network's performance. Moreover, for network performance enhancement, it is imperative to study users' behaviour. Therefore, the main objectives of the present study are to identify UUM network performance through Internet traffic and to highlight users' behaviour. A total of three methodological steps are carried out to meet the objectives of the study; the first one is the data collection phase whereby the source for data collection is taken from the presently used main distributed switch in an hour for each working day in a duration of one week; the second one is the data analysis phase where Wireshark is used to provide the statistics of traffic and finally; the third phase is the data presentation where Microsoft Excel is utilized to present data. Study findings presents valuable information of network bandwidth, data loss rates and Ethernet traffic distribution, in addition to the fact that is social websites are the most websites used in UUM. These findings leads to facilitate the enhancement of UUM network performance and Internet bandwidth strategies.

Acknowledgement

First and foremost, Alhamdulillah, All praise is to my Lord, the Compassionate, and the Merciful Subhanahhuwata 'alah; for giving me the will and strength in the completion of this study.

I would like to express my deepest gratitude and appreciation to my respective Supervisor: Dr. Mohd. Hasbullah bin Omar and to my second supervisor: Mr. Adib M. Monzer Habbal and Mr. Khuzairi bin Mohd Zaini for their expertise, kindness, and patience in guiding throughout the production of this Study.

My excessive gratefulness goes to Dr. Mohammed Haki, my spiritual mentor.
The first, last and always, a lasting heartfelt gratitude to the source of my light and pleasure, to the one who enlightens my life, to my dear Mother. Equal gratitude goes out to my precious Sister.

Finally I am also thankful to the people I met in my life who touched my heart and gave me strength to move forward to something better, Mr Adli for his assistance on data collection at the computer centre. My dear brother Houzifa Mohammed Hentaya and my friend Abed-Alsalam Tayara.

Mustafa Mohammed Haki AL-Kawaz
June 15, 2012

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

ARP	Address Resolution Protocol
DEC/RPC	Distributed Computing Environment / Remote Procedure Calls
DNS	Domain Name Service
FTP	File Transfer Protocol
GIF	Graphics Interchange Format
GREP	Generic Routing Encapsulation Protocol
HTTP	Hypertext Transfer Protocol
ICMP	Internet Control Message Protocol
IP	Internet Protocol
IPv4	Internet Protocol Version 4
IPv6	Internet Protocol Version 6
JPEG	Joint Photographic Experts Group
MIME	Multipurpose Internet Mail Extensions

OSPF	Open Shortest Path First
P2P	Peer To Peer
RTMP	Real Time Message Protocol
RTSP	Real Time Streaming Protocol
SMTP	Simple Mail Transfer Protocol
SNMP	Simple Network Management Protocol
SSH	Secure Shell Protocol
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UUM	University Utara Malaysia
VLAN	Virtual Local Area Networks
YMG	Yahoo Messenger Protocol

CHAPTER ONE
INTRODUCTION

1.1 Introduction

The Phenomenal success of the Internet has led to the rapid adoption of the Internet protocol technology to build all types of communication networks, including private corporate networks (intranet), military communication networks, home networks, and the emerging Third-generation cellular networks. Billions of devises worldwide are IP-capable, which allows remote access and control through the Internet. Such rapid and unprecedented convergence of communication through IP presents a host of challenging problems in guaranteeing the required performance in such networks (Jain & Hassan, 2004).

For the monitoring and security of network, it is imperative to acknowledge and expound on how the applications function. Many researchers have concentrated on the characteristics of traffic and network behavior under particular applications including P2P applications (Cao, Liu, & Xue, 2010).

In other words, a network may be defined as a "set of devices (often referred to as nodes) connected by communication links that are built using different physical media" (Marsic, 2010). A node can be represented by a computer, telephone or any device that facilitates the sending and receiving of messages while the medium of communication is referred to as the physical path through which the message flows from sender to receiver. Examples of media are fiber-optic cable, copper wire or air carrying radio waves (Marsic, 2010).

REFERENCES

Acharya, T. (2005). *JPEG2000 Standard for Image Compression Concepts, Algorithms and VLSI Architectures*. New Jersey, USA: Wiley Sons, INC.

Argyraki, K., Maniatis, P., & Singla, A. (2010). Verifiable Network-Performance Measurements. *IEEE/ACM Transactions on Networking ACM , 19*, 1224-1226.

Augustin, B., & Mellouk, A. (2011). On Traffic Patterns of HTTP Applications. *IEEE Communications Society journal, 11*, 2-4.

Blum, R. (2003). *Network Performance Open Source Toolkit*: Wiley Publishing, Inc., Indianapolis, Indiana.

Canali, C., Casolari, S., & Lancellotti, R. (2010). A quantitative methodology to identify relevant users in social networks. *IEEE/ACM Transactions on Networking IEEE, 4*, 3-10.

Cao, Y., Liu, B., & Xue, Y. (2010). Locality Analysis of BitTorrent-Like Peer-to-Peer Systems. *Communications Society journal IEEE,10*, 1-5.

Chang, C.-W., Huang, G., Lin, B., & Chuah, C.-N. (2011). LEISURE: A Framework for Load-Balanced Network-Wide Traffic Measurement. *IEEE Transactions on Wireless communications,10*, 326-628.

Chuan, X., & Hong, T. (2008). Design and complementation of a real time Traffic Measurement System in High-Speed Networks. *IEEE*.

COMER, D. E. (2009). *Computer Networks and Internets*. New Jersey, USA: Pearson Education, Inc.

Crandall, S., & Jasani, H. (2011). ProMix: Linux Promiscuous Wireless Packet Analysis. *IEEE Transactions on Industrial Informatics*, 1-4.

Curtis, N., & Taylor, P. J. (2005). *Network+ ACompTIA Certification* (Fourth Edition ed.): K LLC- CompTIA.

Donahue, G. A. (2011). *Network Warrior* (Second Edition ed.): O'Reilly.

Dong, X., Clark, J. A., & Jacob, J. L. (2009). User Behaviour Based Phishing Websites Detection. *International Multiconference on Computer Science and Information Technology IEEE, 6*, 40-42.

Dulaney, E., & Harwood, M. (2012). *CompTIA Network+*. USA: Paul Boger, Pearson-Inc.

Edwards, J., & Bramante, R. (2009). *Networking Self-Teaching Guide*. Canada: Wiley Publishing, Inc.

Fan, Z., Zhang, L., & Shen, J. (2010). A User's Preference based Method for Web Service Selection. *International Conference on Advances in Computing, Control, and Telecommunication Technologies IEEE, 63*, 784-787.

Flanagan, D. (2011). *JavaScript: The Definitive Guide*. United States of America.: O'Reilly Media, Inc.

Fuentetaja, I. G., & Economou, M. (2009). Analysis of users' access to museums websites. *15th International Conference on Virtual Systems and MultimediaI, EEE, 24*, 123-126.

Governor, J., Hinchcliffe, D., & Nickull, D. (2009). *Web 2.0 Architectures*. United States of America.: O'Reilly.

Gu, T., Hong, S.-J., & Yoo, J.-B. (2009). Personal Preference for Reliable Transaction Identification on Web Service. *Seventh ACM/IEEE Symposium on Architectures for Networking and Communications Systems IEEE, 9*, 250-255.

Gu, T., Yoo, J.-B., & Park, C.-Y. (2008). Consideration of User Preference on Internet-based Overlay Network. *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, 27*(2), 202-203.

Hartpence, B. (2011). *Packet Guide to Core Network Protocols* (First Edition ed.): O’Reilly Media, Inc.

Hassan, H., Garcia, J. M., & Bockstal, C. (2009). Modeling Internet Traffic: Performance Limits. *43rd Annual IEEE/ACM International Symposium on Microarchitecture IEEE, 3*, 4-6.

Iliofotou, M. (2009). Exploring Graph-based Network Traffic Monitoring. *Workshops of International Conference on Advanced Information Networking and Applications IEEE, 58*, 757-758.

Jain, R., & Hassan, M. (2004). *High Performance TCP/IP Networking*: Pearson Education,Inc.

Joseph, V., & Veciana, G. d. (2011). Stochastic Networks with Multipath Flow Control: Impact of Resource Pools on Flow-level Performance and Network Congestion. *International Conference on Control, Automation and Systems ACM 76*, 1613-1615.

Kim, H., Claffy, k., & Fomenkov, M. (2009). Internet Traffic Classification Demystified: Myths, Caveats, and the Best Practices. *International Conference and Workshop on Emerging Trends in Technology ACM, 32*(891-893).

Knoth, A., & Neuhäuser, D. (2010). IPv6-only Nodes in Corporate and Academic Networks. *IEEE Globecom 2010 Workshop on Heterogeneous, Multi-hop Wireless and Mobile Networks, 11*, 142-145.

Kouvatsos, D. D. (2011). *Network Performance Engineering*. Berlin Heidelberg: Springer-Verlag.

Kurose, J. F., & Ross, K. W. (2010). *Computer networking: a top-down approach* (5th ed ed.): Addison Wesley.

Lammle, T. (2007). *CCNA - Cisco Certified Network Associate* (Sixth Edition ed.): Wiley Publishing, Inc., Indianapolis, Indiana.

Lee, K., Mirchandani, D., & Zhang, X. (2010). An Investigation on Institutionalization of Websites of Firms. *15th IEEE/ACM International Symposium on Distributed Simulation and Real Time Applications IEEE, 9*, 81-87.

Liao, Y.-W., Wang, Y.-S., & Tang, T.-I. (2011). Investigating the Influence of the Landscape Preference of Blogs, User Satisfactory and Behavioral Intention. *Eighth International Conference on Information Technology, IEEE*.

Liu, P., & Hu, R. (2009). Research on Evaluation of E-Commerce WebSites Based on linguistic ordered weighted averaging Operator. *Workshop on Knowledge Discovery and Data Mining, IEEE*.

Lucas, M. W. (2010). *Network Flow Analysis*. San Francisco: William Pollock, No Starch Press, Inc.

Mahimkar, A., Song, H. H., Ge, Z., & Shaikh, A. (2010,). Detecting the Performance Impact of Upgrades in Large Operational Networks. *International Conference on Future Networks ACM, 8*, 74-67.

Marsic, I. (2010). *computer networks, performance and quality of service*: Rutgers University.

McFarland, S., Sambi, M., Sharma, N., & Hooda, S. (2011). *IPv6 for Enterprise Networks*. Indianapolis, IN 46240 United States of America: Cisco Press.

Meiss, M., Menczer, F., & Vespignani, A. (March 2011). Properties and Evolution of Internet Traffic Networks from Anonymized Flow Data. *ACM Transactions on Internet Technology, 71*, 825-828.

Mieghem, P. (2009). *Performance Analysis of Communications Networks and Systems*. New York: Cambridge University Press.

Narayan, S., Lutui, P. R., & Vijayakumar, K. (2010). Performance Analysis of Networks with IPv4 and IPv6. *IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, 6*, 232-236.

Orebaugh, A., Ramirez, G., & Burke, J. (2007). *Wireshark & Ethereal Network Protocol Analyzer Toolkit*: Syngress Publishing, Inc.

Oskouei, R. J. (2010). Analyzing Different Aspects of Social Network Usages on Students Behaviors and Academic Performance. *3rd International Conference on Cloud Computing IEEE*, 59-60.

Parziale, L., & Britt, D. T. (2006). *TCP/IP Tutorial and Technical Overview* (Eighth Edition ed.): IBM Corp - International Business Machines Corporation.

Peterson, L. L., & Davie, B. S. (2012). *Computer Networks: A Systems Approach* (Fifth Edition ed.). United States of America: Elsevier, Inc.

Qadeer, M. A., & Khan, A. H. (2010). Bottleneck Analysis and Traffic Congestion Avoidance. *ACM International Conference and Workshop on Emerging Trends in Technology, 15*, 218-219.

Ray, E. T. (2003). *Learning XML, Second Edition*. United States of America: O'Reilly & Associates, Inc.

Sanders, C. (2007). *practical packet analysis using wireshark to solve real-world network problems*. United States of America: William Pollock, No Starch Press, Inc.

Shan, X., & Sun, H. (2011). The Research of Web Users' Behavior Mining Based on Association Rules. *State Natural Sciences Foundation project subsidization IEEE, 43*, 835-837.

Sloan, J. D. (2001). *Network Troubleshooting Tools*: O'Reilly.

Straubhaar, J., & LaRose, R. (2012). *Media Now 2012 Update* (seventh ed.). United State Of Amarica: Michael Rosenberg, Inc.

Team, C. (2006). Router IP Traffic Export Packet Capture Enhancements. Retrieved 14th April 2012 http://www.cisco.com/en/US/docs/ios/12_4t/12_4t11/ht_rawip.pdf

Wamser, F., Pries, R., & Staehle, D. (2010). Traffic characterization of a residential wireless Internet access. *Springer Science+Business Media, LLC, 25*, 682-683.

Wang, J. H., An, C., & Yang, J. (2010). A study of traffic, user behavior and pricing policies in a large campus network. *Symposium on Architectures for Networking and Communications Systems IEEE, 13*, 268-270.

Wang, N. (2010). The Fuzzy Comprehensive Evaluation of User-oriented Government Websites. *IEEE Transactions on vechicular techology 59*, 1821-1823.

Xiaojian, W. (2009). Comprehansive Evaluation on E-commerce Website Applying Improved TOPSIS Method. *International Conference on Electronic Commerce and Business Intelligence, IEEE, 11*, 119-120.

Xu, D., Wang, S., & Yan, S. (2010). Analysis and Application of Wireshark in TCP/IP Protocol Teaching. *Fifth International Conference on Systems and Networks Communications IEEE, 26*, 381-386.

Yang, X., Chen, X., & Jin, Y. (2011). A High-speed Real-time HTTP Performance Measurement Architecture Based on Network Processor. *IFIP International Conference on Network and Parallel Computing IEEE, 21*, 341-342.

Yildirim, E., Suslu, I. H., & Kosar, T. (2009). Which Network Measurement Tool is Right for You? A Multidimensional Comparison Study. *9th Grid Computing Conference, IEEE, 15*, 152-155.