LB/DON/56/02

DCS 03/02 2

Design and Implementation of a Network Traffic Analyzing Utility

This thesis was submitted to the

Department of Computer Science & Engineering

of the

University of Moratuwa

in partial fulfillment of the requirements for the

Degree of Masters of Science

681.3 "02"
681.324

074544
University of Moratuwa

පුස්තකාලය
මොරටුව විශ්ව විද්‍යාලය, ශ්‍රී ලංකාව
මොරටුව

74544

by

Priyantha Pushpa Kumara

Department of Computer Science & Engineering
University of Moratuwa
Sri Lanka
February 2002

TH

74544

Declaration

I, Priyantha Pushpa Kumara hereby certify that the work included in this thesis has not been submitted in part or whole for any other academic qualification at any institution.

P.pushpak

Priyantha Pushpa Kumara

Research Student

UOM Verified Signature

Dr. P G V Dias

Supervisor

Abstract

Developments in Information and Communication Technology (ICT) have created many new applications, which require a large amount of Internet bandwidth for proper application usage. However, Internet bandwidth is an expensive resource especially in this part of the world. Therefore, enterprises that use the Internet for business always need to efficiently use the bandwidth. Internet Service Providers may also need to manage their Internet bandwidth efficiently in order to increase the profit margin of their services.

Knowing the current (and past) usage is an important requirement in estimating the bandwidth requirement of a corporate network. The information about the composition of the use of the bandwidth by applications, user groups, etc. will help an administrator to efficiently manage it. Furthermore, knowing what is happening in the network always helps an administrator to keep the network secure, efficient and reliable all the time.

We studied the requirements of a network administrator by considering LEARN as a test bed for this study. We then assessed, whether the existing network traffic monitoring tools can provide the required information to the administrator. We examined some of the popular network monitoring tools that are widely used in the Internet community for their features and drawbacks. We found that the existing tools for network traffic monitoring are not capable of providing most of the required information by the administrator.

With these results, we identified the features available in those tools, and developed a new tool, LEARNStat, re-using some of the freeware utilities available in order to meet the requirements of a network administrator. We tested LEARNStat in LEARN which is our test-bed, and during the short period of running LEARNStat, we were able to obtain several important results.

In this thesis, we discuss the requirements of a network administrator and how we met those with LEARNStat. We present a brief description of related principles and also discuss the results we obtained with LEARNStat. The thesis also includes possible future enhancements for the LEARNStat.

Acknowledgement

I take pleasure in thanking Dr. Gihan Dias, for supervising this project and using his valuable time to contribute, guide and instruct me in not only revising and writing this thesis but also proposing, commenting and evaluating my work throughout the project. I would also like to express my sincere thanks to Dr. Nalin Wickramarachchi, head of the department of Computer Science & Engineering for his guidance and permission given to me, to use departmental resources for this work. My sincere thanks to Dr. Nalin Ranasinghe, the external supervisor, for his guidance and positive comments for the success of this project.

I am also grateful to the LEARN management committee for allowing me to carry out this project while working in LEARN as well as allowing me to use its resources for this project. Without the sponsorship, this work would not have been possible.

Thanks are also due to Prof. (Mrs.) Ratnayake, Director, Postgraduate Studies, for her guidance throughout the project.

I received valuable comments and guidance from Vishaka Nanayakkara, throughout the project that always encouraged me to successfully complete this project. Without her help, I would not have met Prof. Sven Tafvelin, who gave me valuable guidance as well as a contact for NorduStat. Thank you Vishaka and thank you Prof. Tafvelin.

Thanks to Rasika Amarasiri for his guidance and help given to me whenever I had any problems with Perl or HTML.

Thanks to all staff members of the Department of Computer Science & Engineering and network administrators of LEARN member sites who directly or indirectly helped me to successfully develop and test the LEARNStat.

Thanks to Dave Plonka and Håvard Eidnes for their kindness in allowing me to publish some of the graphs that represent their own network statistics in this thesis.

My special thanks to the final year project group, Karunarathna, Sendanayake, Ms. Wickramasinghe, & Ms. Jayawardana, those who first installed and configured NeTraMet and showed me the capabilities of NeTraMet.

Thanks to Mr. G. H. A. Perera for his help in proof reading and correcting the language mistakes.

Finally, the most important people, my wife Preethi and even little Dilsha endured everything with infinite patience and good grace, during this period. Thank you.

UNIVERSITY OF MORATUWA
LIBRARY
SRI LANKA

Contents

ABSTRACT I

ACKNOWLEDGEMENT II

CONTENTS III

LIST OF FIGURES V

LIST OF TABLES VII

ABBREVIATIONS VIII

1. INTRODUCTION 1

1.1 OBJECTIVES 1
1.2 LEARNSTAT 2
1.3 OUTLINE OF THE THESIS 3

2. LEARN 4

2.1 HISTORY 4
2.2 EXPERIENCE IN LEARN 6
2.2.1 *Network Monitoring tools being used* 6
2.2.2 *Problems encountered with tools used* 7
2.2.3 *Summary* 8

3. NETWORK MANAGEMENT 9

3.1 ISO/OSI NETWORK MANAGEMENT STANDARDS 9
3.2 INTERNET NETWORK MANAGEMENT 10
3.2.1 *Simple Network Management Protocol-SNMP* 10
3.2.2 *TCP/IP* 14
3.3 NETWORK MONITORING 18
3.3.1 *Which layer to monitor?* 18
3.3.2 *Types of monitors* 18
3.4 CASE STUDIES 19
3.4.1 *MRTG v2.0 (Multi Router Traffic Grapher)* 19
3.4.2 *FlowScan* 21
3.4.3 *NeTraMet* 24
3.4.4 *NORDstat* 26

4. LEARNSTAT SYSTEM DESIGN AND ARCHITECTURE 30

4.1 SYSTEM OVERVIEW 30
4.2 LEARNSTAT SYSTEM DESIGN 30
4.3 SYSTEM ARCHITECTURE 32
4.4 USER INTERFACE 33
4.4.1 *User Interface Design* 39

5. LEARNSTAT IMPLEMENTATION 40

5.1 SYSTEM COMPONENTS 40
5.1.1 *LEARNStat Regular Module* 40
5.1.2 *LEARNStat On-line Module* 45
5.2 SYSTEM INTEGRATION DETAILS 48
5.2.1 *Integrating Locally developed LEARNStat components* 49
5.2.2 *Integrating Re-used components to LEARNStat* 50

6. **OUTCOME OF LEARNSTAT** 52

6.1 INTERESTING RESULTS 52
6.1.1 *Discovery of unknown servers* 52
6.1.2 *Discovery of unauthorized cache users* 52
6.1.3 *Discovery of a Loop-Hole in Router Access-lists* 53
6.1.4 *Monitoring Bandwidth Utilization by Applications* 54
6.2 LEARNSTAT SYSTEM VALIDATION & VERIFICATION 56
6.2.1 *LEARNStat Unit Testing* 56
6.2.2 *LEARNStat Integration Testing* 57
6.2.3 *LEARNStat System Testing* 57
6.3 LEARNSTAT RESULT VALIDATION 57
6.3.1 *Validation of LEARNStat Graphs* 58
6.3.2 *Validation of non-graphical outputs* 59
6.3.3 *Validation using known operations* 61
6.4 PROBLEMS ENCOUNTERED 62
6.4.1 *Problems in General* 62
6.4.2 *Problems Encountered in Regular module* 63
6.4.3 *Problems Encountered in On-line Module* 66

7. **CONCLUSION AND RECOMMENDATIONS** 68

7.1 MERITS AND DRAWBACKS OF LEARNSTAT 68
7.1.1 *Merits of LEARNStat* 68
7.1.2 *Drawbacks of LEARNStat* 68
7.2 FUTURE ENHANCEMENTS POSSIBLE TO LEARNSTAT 69
7.2.1 *A new Rules Syntax* 69
7.2.2 *A new Algorithm to match Alarm Rules* 70
7.2.3 *Sending urgent alarms* 71
7.2.4 *Alarms on traffic volumes* 71
7.2.5 *A GUI for Administrators* 71
7.2.6 *Data Validation* 72
7.2.7 *Code re-writing* 72
7.3 LEARNSTAT IN SUMMARY 72

APPENDIX A 73

A.1 INSTALLING AND CONFIGURING RE-USED COMPONENTS 73
A.1.1 *Installing and Configuring NeTraMet* 73
A.1.2 *Configuring NetFlow on Cisco routers* 74
A.1.3 *Installing and Configuring RRDTools* 74
A.2 INSTALLING AND CONFIGURING LOCALLY DEVELOPED COMPONENTS 74
A.2.1 *Installation procedure* 74
A.2.2 *Configuration procedure* 74

APPENDIX B 79

B.1 LEARNSTAT DATA FORMATS 79
B.1 IDENTIFYING DATA FLOWS 80
B.2 FILE FORMATS 80
B.2.1 *Configuration File Format* 80
B.2.2 *Rule File Format* 82

REFERENCES 85

List of Figures

Figure 2.1 LEARN phase 1 network topology 5

Figure 2.2 LEARN topology expected at the eARLY 2002 6

Figure 2.3 An MRTG graph showing Internet bandwidth usage 7

Figure 2.4 An MRTG graph showing a local link usage 7

Figure 2.5 An MRTG graph showing the usage of one of the LEARN backbone links 8

Figure 3.1 OSI Network Management Architecture 9

Figure 3.2 SNMP Management Model 10

Figure 3.3 A part of ASN.1 object naming tree 12

Figure 3.4 TCP/IP protocol suite 15

Figure 3.5 Adding headers and trailers to Application data as it goes down the TCP/IP protocol stack 15

Figure 3.6 IP Datagram format 16

Figure 3.7 IP Address and Netmask 16

Figure 3.8 MRTG graphs with different data resolutions 19

Figure 3.9 MRTG uses SNMP queries to gather data from a remote router 20

Figure 3.10 A graph generated by RRDTools 21

Figure 3.11 FlowScan graph showing Well-known services 22

Figure 3.12 NetFlow version 5 PDU format 23

Figure 3.13 Interaction between Meters and Collector & Manager 24

Figure 3.14 A sample flow data file 25

Figure 3.15 NORDstat graph showing traffic rate in Kbits/s 26

Figure 3.16 A network map produced by NORDstat 27

Figure 3.17 Hourly average graph by NORDstat 27

Figure 3.18 Snapshot of tabular information presented by NORDstat 28

Figure 3.19 Graph obtained by clicking the first link in the table (figure 3.18) 28

Figure 4.1 LEARNStat Software components 31

Figure 4.2 LEARNStat Layered Architecture 32

Figure 4.3 Screen shot of LEARNStat front page 33

Figure 4.4 LEARNStat Usage Graphs 34

Figure 4.5 LEARNStat Port Summary Screens 35

Figure 4.6 LEARNStat Network Users Screens 36

Figure 4.7 Servers listed by services they run 37

Figure 4.8 RRGrapher screens 38

Figure 4.9 LEARNStat Front Page with an Alarm 38

Figure 4.10 LEARNStat User Interface summary 39

Figure 5.1 The LearnStat.pl algorithm in brief 41

Figure 5.2 Summary of makedb.pl program logic 42

Figure 5.3 Logic of makegraphs.pl 43

Figure 5.4 Logic of servers.pl 43

Figure 5.5 Simple logic of alarms.pl 44

Figure 5.6 LEARNStat On-line Module components and their interaction with HTML pages 47

Figure 5.7 A part of a sample HTML template file 48

Figure 5.8 LEARNStat components after integration 49

Figure 5.9 A sample batch job listing 50

Figure 5.10 The updated batch job listing 50

Figure 6.1 The alarm rules file used for beta test 53

Figure 6.2 Alarm details screen 53

Figure 6.3 LEARNStat graph showing video conferencing bandwidth usage 54

Figure 6.4 New lines added to the rules file for identifying video conferencing data flows 54

Figure 6.5. Mirror synchronization bandwidth utilization 55

Figure 6.6 A Comparison of LEARNStat graphs with MRTG graphs 59

Figure 6.7 LEARNStat Data Validation Query and a Result 60

Figure 6.8 LEARNStat result verification using a known operation 62
Figure 6.9. UDP stream of cache-flow data from a router to the meter 62
Figure 6.10 LEARNStat graph that shows a delay in getting updated 63
Figure 6.11 LEARNStat batch job script with time stamps 63
Figure 6.12 LearnStat.pl performance variations 65
Figure 6.13 A script to run LearnStat.pl with two sets of rules 66
Figure 7.1 New alarm rule syntax 69
Figure 7.2 A new algorithm for evaluating alarm rules 70
Figure A.1 Alarm rules as seen from LEARNStat on-line module 78
Figure B.1 Logic of attribute based simple test 81
Figure B.2 Configuration file format 82
Figure B.3 Format of the LEARNStat Rule file 83
Figure B.4 Format of the alarm rule file 84

List of Tables

TABLE 3.1 *SNMP PRIMITIVE DATA TYPES* 11
TABLE 3.2 *OBJECTS GROUPS OF MIB-2* 12
TABLE 3.3 *IP ADDRESS EXAMPLES* 17
TABLE 3.4 *SOME OF THE COMMON TCP/UDP PORTS ASSIGNMENTS* 17
TABLE 5.1 *LEARNSTAT REGULAR MODULE COMPONENTS* 40
TABLE 5.2 *LEARNSTAT ON-LINE MODULE COMPONENTS* 45
TABLE 5.3 *LEARNSTAT PERL MODULE COMPONENTS* 50
TABLE 6.1 *LEARNSTAT UNIT TEST RESULTS* 56
TABLE A.1 *KEYWORDS USED IN LEARNSTAT CONFIGURATION FILE* 75
TABLE B.1 *LEARNSTAT COMPONENTS WITH THEIR INPUTS AND OUTPUTS* 79

UNIVERSITY OF MORATUWA
LIBRARY
SRI LANKA

Abbreviations

CIDR	Classless Inter Domain Routing
CINTEC	Council for INformation TECnology
GUI	Graphical User Interface
IANA	Internet Assigned Number Authority
ICP	Internet Cache Protocol
IETF	Internet Engineering Task Force
IOS	Internet Operating System (Cisco uses this name for their router operating systems)
ISO	International Standard Organization
IX	Internet eXchange
LEARN	Lanka Education And Research Network
MIB	Management Information Base
NARA	National Aquatic Resources Agency
NARESA	Natural Resources Energy and Science Authority (Now, this is named National Science Foundation - NSF)
NORDUnet	This is the Nordic Internet highway to research and education networks in Denmark, Finland, Iceland, Norway and Sweden and provides the Nordic backbone to the global Internet.
PDU	Protocol Data Unit
RFC	Request For Comments
RPN	Reverse Polish Notation
RTFM	Real-time Traffic Flow Measurement
SMS	Short Message Service
SNMP	Simple Network Management Protocol
Sida	Swedish International Development Cooperation Agency
SAREC	The research arm of the Sida